Preliminary Exploration of a DNA Editing Strategy to Treat Cystic Fibrosis using Homology-Independent Targeted Integration and Receptor-Targeted Nanoparticles

Carina Graham
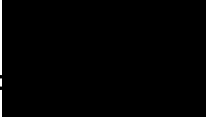
Great Ormond Street Institute of Child Health

University College London

Thesis submitted for the degree of Doctor of Philosophy

September 2023

Declaration

I, Carina Graham, declare that the work presented in this thesis is my own. Any contributions from collaborators, or where information has been derived from other sources, is explicitly referenced in the text.

Signature:

Date: 12 September 2023

Abstract

Cystic fibrosis (CF) is a common life-limiting disorder. Dysfunctional cystic fibrosis transmembrane conductance regulator (CFTR) protein hinders mucociliary clearance in the airways, leading to a cycle of infection, inflammation, and tissue damage. Pharmaceutical intervention is available for patients of the most common CF-causing genotype (F508del), but burden of care remains high, and patients carrying non-F508del variants are ineligible for drug use. CRISPR-Cas9 is a promising technology to eliminate underlying causes of genetic disorders. We hypothesised firstly that a CRISPR approach based in homology independent targeted integration could be used to generate a "one size fits all" treatment strategy for CF patients harbouring "undruggable" genotypes ineligible for pharmaceutical intervention, and secondly that receptor-targeting nanoparticles (RTNs) could be used to deliver CRISPR machinery to differentiated airway epithelial cells.

We characterised a CF mutant cell line homozygous for an undruggable variant in order to provide a phenotypic baseline for future treatment studies. Characterisation via Ussing chamber and scanning ion conductance microscopy confirmed that variant cells exhibited lack of ion transport, dehydrated airway surface liquid, and decreased mucosal elasticity. To provide proof of concept for a CRISPR-based strategy using homology-independent targeted integration (HITI), CRISPR-Cas9 machinery and donor were transfected via plasmids and used to edit an eGFP reporter construct into the AAVS1 safe harbour locus. The reporter construct inserted with an efficiency of ~40%. RTNs were then investigated for their ability to transfect airway epithelial cells; to enhance editing efficiency, RTNs were packaged with CRISPR ribonucleoprotein (RNP). RTNs were assayed for their biophysical characteristics (size and charge) and their ability to transfect airway epithelial

cells *in vitro* using luciferase-based reporter assays. RTNs' ability to penetrate human airway mucus was investigated *in vitro*, establishing that CF mucus provides a significant barrier to RTN mobility in fluid. When transfection of differentiated air-liquid interface (ALI) cultures was attempted, we found negligible transfection using RTNs.

These results lay the groundwork for future study. The editing efficiency achieved using HITI exceeds the predicted rate required for rescue of CF pathology; therefore, replacing the reporter construct donor with a CFTR cDNA donor could potentially represent a variant-agnostic treatment for CF. RTNs' promising performance in submerged cells and isolated non-CF mucus, but not ALI-differentiated non-CF cells, provides insight into the barriers to *in vivo* RTN-mediated transfection. Future work with these RTNs will require further optimisation of their formulation (addition of polyethylene glycol to lipid formulations and repeat of phage display in ALI to identify new targeting peptides) and ALI transfection protocols (applications of mucolytics, chemicals to loosen tight junctions and increase accessible membrane surface area, paralysis of cilia).

Impact Statement

While great gains have been made in cystic fibrosis (CF) treatment in the past decade, it remains a life-limiting disease. In particular, while recent pharmaceutical advances can alleviate disease for an estimated 90% of patients, there remains a population of patients unable to benefit from current available drugs. Additionally, even in patients eligible for the drug, burden of care remains high.

Due to the monogenic nature of the disease, CF is good candidate for gene-based medicine. However, many prior attempts to treat CF with genetic medicine have not been approved following clinical trial. The current pharmaceutical treatments available to patients leave a high burden of care and are only applicable to patients harbouring a specific genotype; therefore, a variant-agnostic approach is required. Therefore, CRISPR-mediated integration of a full CFTR cDNA is a promising option. When it comes to CRISPR-Cas9 based approaches to treating CF, there are three major obstacles. Firstly, many precision editing strategies are inherently low-efficiency, and the high number of CF-causing variants makes mutation-specific approaches an ineffective strategy. Secondly, the non-dividing nature of the target tissue makes transfection difficult and eliminates the possibility of relying on homology-directed repair. The research presented in this thesis attempts to address both issues. Firstly, the CRISPR-based editing strategy used herein accomplishes high-precision insertion of a reporter construct without relying on homology-directed repair, meaning that it is usable in both dividing and non-dividing cells (such as those making up the airway epithelium). This strategy was found to yield high editing efficiency and provides proof of concept for a potential "one size fits all" solution to rescue CF regardless of patient genotype. This is an exciting prospect because the current major treatment for CF (the

triple-combination therapy drug Trikafta) is incapable of rescuing CF pathology in patients who do not carry a specific class of genotype. Secondly, receptor-targeted nanoparticles (RTNs) are investigated for their ability to more effectively transfect airway epithelial cells than non-targeting nanoparticles, assayed for motility in airway mucus, and tested for their ability to transfect fully differentiated models of the airway epithelium. Both the RTNs and the DNA editing strategy can be more broadly applied to other monogenic diseases, as the target locus for editing is a genomic safe harbour rather than the endogenous disease-causing locus, and the targeting elements of RTNs are modular and modifiable.

The work herein has been presented at conferences internationally, and manuscripts are in preparation or under review for publication.

UCL Research Paper Declaration Form
referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

   a) **What is the title of the manuscript?**

   CRISPR-Cas9 Gene Editing Therapies for Cystic Fibrosis

   b) **Please include a link to or doi for the work**

   10.1080/14712598.2021.1869208

   c) **Where was the work published?**

   Expert Opinion on Biological Therapy

   d) **Who published the work?** (e.g. OUP)

   Taylor & Francis

   e) **When was the work published?**

   2021

   f) **List the manuscript's authors in the order they appear on the publication**

   Graham, C.

   Hart, S.

   g) **Was the work peer reviewed?**

   Yes

   h) **Have you retained the copyright?**

   No, but I retain rights to reproduce it in a thesis

   i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

   No

   If 'No', please seek permission from the relevant publisher and check the box next to the below

   statement:

   ☒

*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

   a) **What is the current title of the manuscript?**
   b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

Click or tap here to enter text.

   c) **Where is the work intended to be published?** (e.g. journal names)

Click or tap here to enter text.

   d) **List the manuscript's authors in the intended authorship order**

Click or tap here to enter text.

   e) **Stage of publication** (e.g. in submission)

Click or tap here to enter text.

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

   Carina Graham wrote the body and generated figures; Stephen Hart provided guidance and advice for structuring, sources to include, and responses to reviewer comments.

4. **In which chapter(s) of your thesis can this material be found?**

   1 Introduction

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Carina Graham

Date:

5 September 2023

Supervisor/ Senior Author (where appropriate)

Click or tap here to enter text.

Date

Click or tap here to enter text.

Acknowledgements

Thanks firstly to the four cornerstones of my mental health: my family, Haley, Jas, and sertraline.

To the members of the Hart lab. To Steve for all the guidance, and for tirelessly insisting that I was going to be okay. To Ruhina for keeping the whole lab afloat, Amy J for the pep talks, and Miriam and Yiman for the laughter and commiseration. To Amy W and Emily for welcoming me to the lab and to London.

To the other wonderful scientists who have supported me, through collaboration or emotional support (or both): Issie Rose, Dr. Debbie Baines, and Dr. Max Woodall for Ussing; Arthur Mitchell, Dr. David Benton, and Dr. Guy Moss for SICM; Dr. Dale Moulding for microscopy; Machaela Palor, Panagiotta Constandinou, and Dr. Ayad Eddouadi for flow cytometry; to Dr. Patrick Harrison for all the HITI advice. To the G20 gang for daily support.

To my Eckerd mentors. To the NIH crew, especially Andy – may your memory be a blessing. You all told me this was worth a shot. Let's see if you were right.

Thank you to those who have made London feel like home, especially Jas (love you to pieces, Lightning BitQueen). I can't believe I went from knowing no one on this entire continent to this richest universe of friends.

Thanks of course to the folks back home. To Mom, Dad, Marisa, and McKenna: nothing on this earth I love more than you. To Amy and Hop for all the help writing. To Lexi for a lifetime of unabashed silliness. To Haley: thank goodness we found each other. I can't wait to see who we turn out to be.

Thanks finally to Mrs. Virginia Brown. I hated science before I took your class, and sometimes I think I still might, but at least now I have a decent shot at getting to make people call me "Doctor."

Publications

Graham, C., & Hart, S. (2021). CRISPR/Cas9 gene editing therapies for cystic fibrosis. *Expert Opinion on Biological Therapy*, *21*(6), 767-780.

Walker, A., **Graham, C.,** *et al.* Molecular and Functional Correction of a Deep Intronic Splicing Mutation in *CFTR* by CRISPR/Cas9 Gene Editing. *Molecular Therapy.* Manuscript under review.

Graham, C. *et al.* Receptor-targeted nanoparticles for the treatment of cystic fibrosis. Manuscript in preparation.

Table of Contents

Abbreviations

AAV	Adeno-associated virus
AAVS1	Adeno-associated virus integration site 1
ABC	ATP-binding cassette
ABE	Adenine base editor
ALI	Air-liquid interface
ANOVA	Analysis of variance
AsCas12a	Cas12a nuclease derived from *Acidaminococcus sp.*
ASL	Airway surface liquid
ATP	Adenosine triphosphate
BFP	Blue fluorescent protein
BMI-1	B-cell-specific Moloney murine leukaemia virus integration site 1
bp	Base pair
C14	DTDTMA/DOPE
C16	DHDTMA/DOPE
C18	DOTMA/DOPE
$CaCl_2$	Calcium chloride
CAG	CMV enhancer fused to chicken beta-actin promoter
Cas	CRISPR-associated system
CBE	Cytidine base editor
CBF	Ciliary beat frequency
CBh	Chicken beta actin short promoter
cDNA	Complementary DNA

CF	Cystic fibrosis
CFBE	Cystic fibrosis bronchial epithelial (cell)
CFNE	Cystic fibrosis nasal epithelial (cell)
CFTR	Cystic fibrosis transmembrane conductance regulator
Cl^-	Chloride ion
CO_2	Carbon dioxide
Cpf1	CRISPR from *Prevotella* and *Francisella* 1
CRISPR	Clustered regularly interspaced palindromic repeats
crRNA	CRISPR RNA
DHDTMA	1,2-Di-((Z)-hexadec-11-enyloxy)-N,N,N trimethylammonium propane iodide
DMEM	Dulbecco's modified Eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOPE	dioleoyl L-α phosphatidyl ethanolamine
DOTMA	N-[1- (2,3-dioleyloxy)propyl]-N,N,N-trimethlylammonium chloride
DSB	Double-strand break
dsDNA	Double-stranded DNA
DTDTMA	1,2-Di-((Z)-tetradec-11-enyloxy)-N,N,N-trimethylammonium propane chloride
E. coli	*Escherichia coli*
ENaC	Epithelial sodium ion channel

F508del	Deletion of phenylalanine at amino acid position 508 of CFTR
GFP	Green fluorescent protein
gRNA	Guide RNA
h	Hours
HDR	Homology-directed repair
HITI	Homology-independent targeted integration
HIV	Human immunodeficiency virus
iPSCs	Induced pluripotent stem cells
Isc	Short-circuit current
K_2HPO_4	Dipotassium phosphate
kb	Kilobase pair
KH_2PO_4	Monopotassium phosphate
L2K	Lipofectamine 2000
LTR	Long terminal repeat
M	Molar
MEM	Minimum essential media
$MgCl_2$	Magnesium chloride
MMEJ	Microhomology-mediated end joining
MOI	Multiplicity of infection
mRNA	Messenger RNA
MSCs	Mesenchymal stromal cells
MSD	Membrane-spanning domain
mV	Millivolt

Na$^+$	Sodium ion
NaCl	Sodium chloride
NaCO$_3$	Sodium carbonate
NBD	Nucleotide binding domain
ng	Nanogram
NGS	Next-generation sequencing
NHBE	Normal human bronchial epithelial (cells)
NHEJ	Non-homologous end joining
NHNE	Normal human nasal epithelial (cell)
NLS	Nuclear localisation signal
nm	Nanometer
ns	Not significant
O$_2$	Oxygen
p	Probability value
PAM	Protospacer-adjacent motif
PCR	Polymerase chain reaction
PDI	Polydispersity index
PE	Prime editor
PEG	Polyethylene glycol
Peptide E	K$_{16}$GACSERSMNFCG
Peptide Y	K$_{16}$GACYGLPHKFCG
PKA	Protein kinase A
PT	Post-transfection

RD	Regulatory domain
RFP	Red fluorescent protein
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
ROCK	rho-associated protein kinase
RT-PCR	Reverse transcription PCR
RTN	Receptor-targeted nanoparticle
SD	Standard deviation
siRNA	Small interfering RNA
*Sp*Cas9	*Streptococcus pyogenes* Cas9
ssDNA	Single-stranded DNA
ssODN	Single-stranded oligonucleotide
TAE	Tris-acetate-EDTA buffer
TALEN	Transcription activator-like effector nuclease
TEER	Transepithelial electrical resistance
TIDE	Tracking of indels by decomposition
TMD	Transmembrane domain
tracrRNA	Trans-activating crRNA
TransCRISTI	Transposase-CRISPR mediated targeted integration
U	Units
UTR	Untranslated region
V	Volt

v:v	Volume:volume	
WT	Wild-type	
ZFN	Zinc finger nuclease	
µL	Microliter	
µm	Micrometer	
µM	Micromolar	

Table of Figures

Table of Tables

Introduction

1.1 Cystic Fibrosis

1.1.1 The *CFTR* gene and protein biosynthesis

Cystic fibrosis (CF) is a severely life-limiting genetic disease caused by mutations in the *CFTR* gene that normally encodes cystic fibrosis transmembrane conductance regulator protein. The *CFTR* gene is located on chromosome 7 and consists of 27 exons across ~190 kb (not including promoter) encoding a ~4.5 kb cDNA [1-3]. Several alternatively spliced forms of *CFTR* mRNA have been detected, but these isoforms are typically tissue-specific and in many cases their precise function is unknown [4]. For example, an isoform excluding exon 5 is co-expressed in the heart alongside the full-length transcript [5]. In rabbits, relative expression of each isoform was found to fluctuate from balanced to imbalanced in favour of the exon 5$^-$ isoform over the course of development [6]. Similarly, an isoform expressed in the kidney during nephrogenesis includes a 165-bp deletion inducing a frameshift and premature stop codon; in human fetal kidney cells, this isoform was found to increase in expression relative to full-length CFTR over the course of development [7].

During translation of the polypeptide chain by the ribosome, the ribosome nascent chain complex (RNC) is translocated to the endoplasmic reticulum (ER) [8]. Folding begins cotranslationally, but final domain assembly is achieved in the ER with assistance from chaperone molecules including calnexin [9]. Also in the ER, the protein undergoes an initial N-glycosylation stage [10]. From the ER, properly folded protein is translocated to the Golgi apparatus for final glycosylation modification. Via these two glycosylation stages, the protein grows in size from 130 kDa to 150 kDa to 170 kDa [10]. Relative intensity of bands at these three distinct sizes (termed Bands A, B, and C respectively) is often used in Western blotting to assess successful protein processing and maturation [11-13]. Once the protein

has reached its mature form, it is trafficked via the trans-Golgi network (TGN). Golgi-secreted vesicles are trafficked through cytoskeletal actin fibres via interaction with Myo5b, deposited at the apical cell membrane, and embed their transmembrane (TM) domains to form active ion channels [14].

The most common CF-causing mutation is known as F508del. A three base pair deletion results in absence of a critical phenylalanine residue, which prevents proper inter-domain association and correct protein folding [15-17]. It has been shown that F508del protein does not successfully exit the ER for transport to the Golgi apparatus. Instead, the misfolded protein is bound by Hsp70 and targeted for degradation in a pre-Golgi compartment [18, 19].

1.1.2 CFTR protein structure and molecular function

CFTR encodes CFTR, an adenosine triphosphate (ATP)-gated chloride-bicarbonate ion channel protein belonging to the ATP-binding cassette (ABC) transporter protein superfamily [20]. Human CFTR consists of 1480 amino acids comprising five domains: two transmembrane domains (TMD1 and 2, sometimes also called membrane-spanning domains or MSDs), each associated with one of two nucleotide-binding domains (NBD1 and 2), and one regulatory domain (RD). The two TMDs together form the ion channel pore, which allows or disallows passage of chloride anions via its opening and closing. Opening and closing of the channel is mediated by hydrolysis of ATP by the NBDs [20, 21]. Activation of the NBDs requires dual phosphorylation of the RD by protein kinase A (PKA) [20]. At the apical membrane, CFTR turnover is mediated via cycles of endocytosis and re-trafficking. Normally, a single CFTR molecule has a half-life of approximately 30 hours at the apical membrane [22].
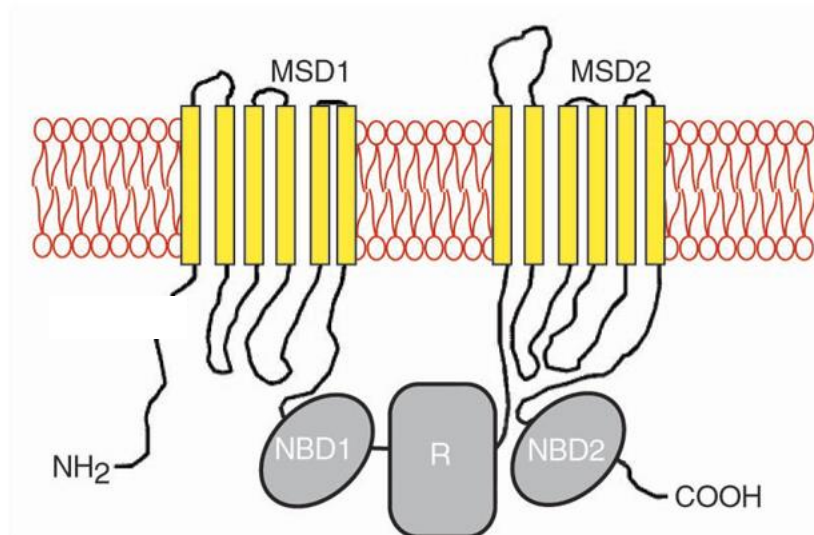
Fig. 1.1: Illustration of the protein structure of CFTR. From Cystic Fibrosis Mutation Database (http://www.genet.sickkids.on.ca/cftr/). NH_2 is N-terminal amine group; MSD is membrane-spanning domain, NBD is nucleotide-binding domain, R is regulatory domain, COOH is C-terminal carboxyl group. Domains shown as grey or yellow shapes; intervening peptide sequence shown as black line. Cell membrane shown as red structure.

1.1.3 CFTR and CF

CFTR is expressed in various mucus-producing exocrine tissues including the airway, the pancreas, and the *in-utero* vas deferens [23]. The ion channel protein's primary role in these tissues is in regulating the viscosity of mucus. Trafficking of anions across the cell membrane creates osmotic pressure at the epithelial surface, which draws water into mucus to maintain appropriately thin consistency [24].

CFTR activity is essential for a variety of organ-specific functions; hence, CF is a highly pleiotropic disease. For example, in the pancreas, inappropriately thick mucus can clog bile ducts and prevent secretion of digestive enzymes, leading to malnutrition [25, 26].

Thickened mucus in the intestines also results in inhibited digestion and nutrient absorption [25]. During gestation, formation of mucus plugs in the vas deferens lead to degradation of the tissue, resulting in infertility [26]. However, the most severe impacts of CFTR dysfunction are found in the airway.

1.1.4 CF in the airway

In the healthy airway, the epithelium is coated in thin, watery airway surface liquid (ASL). ASL consists of two layers: the periciliary layer (PCL) and the mucus layer [27]. Mucus is responsible for trapping dirt, pathogens, and any other foreign substances. Cilia move within the PCL, which is made of 96% water [28]. Their movement is coordinated in a directional sweeping motion that transports mucus up and out of the airway; this phenomenon is commonly referred to as mucociliary clearance (MCC) or the mucociliary escalator [29]. Additionally, the ASL contains bacteria-killing cells and antibacterial compounds [30].

In people with CF, thickened mucus and reduced ASL layer height create a wide variety of health problems. Mucus forms sticky plugs that impede free airflow [31]. Bronchiectasis impedes ciliary movement, meaning that dust and pathogens are not removed from the lung [32]. Dysregulation of ion trafficking alters ASL pH, which in turn inhibits the bacteria-killing activity of immune cells [33]. This non-removal and uninhibited growth of bacteria leads to chronic treatment-resistant infection [34-36], which contributes to eventual mortality. Infection in turn leads to inflammation. Swelling of the airway tissues causes further inhibition of airflow. Furthermore, chronic recruitment of neutrophils creates downstream harm. Neutrophils release elastases which damage the collagen matrices of the lungs, further contributing to inflammation [37]. Dead neutrophils also release their DNA in

structures known as neutrophil extracellular traps (NETs), which contribute further to thickening of mucus and function as anchoring points for bacterial biofilms [34]. Thusly, thickened mucus caused by CFTR dysfunction causes a cycle of damage and inflammation in airway tissues. Progressive loss of lung elasticity, chronic bacterial infection, and mucosal blockages in the airway lead to impeded lung function; this is typically measured by forced expiratory volume in one second (FEV_1), which in later stages of lung disease can decrease to 30% of non-CF patients [38]. At the last stage of lung disease, patients require lung transplant as a lifesaving intervention. In the UK, median age of survival is 48.4 years [39]; however, in countries with less developed healthcare systems, median survival age can be as low as 12 [40].
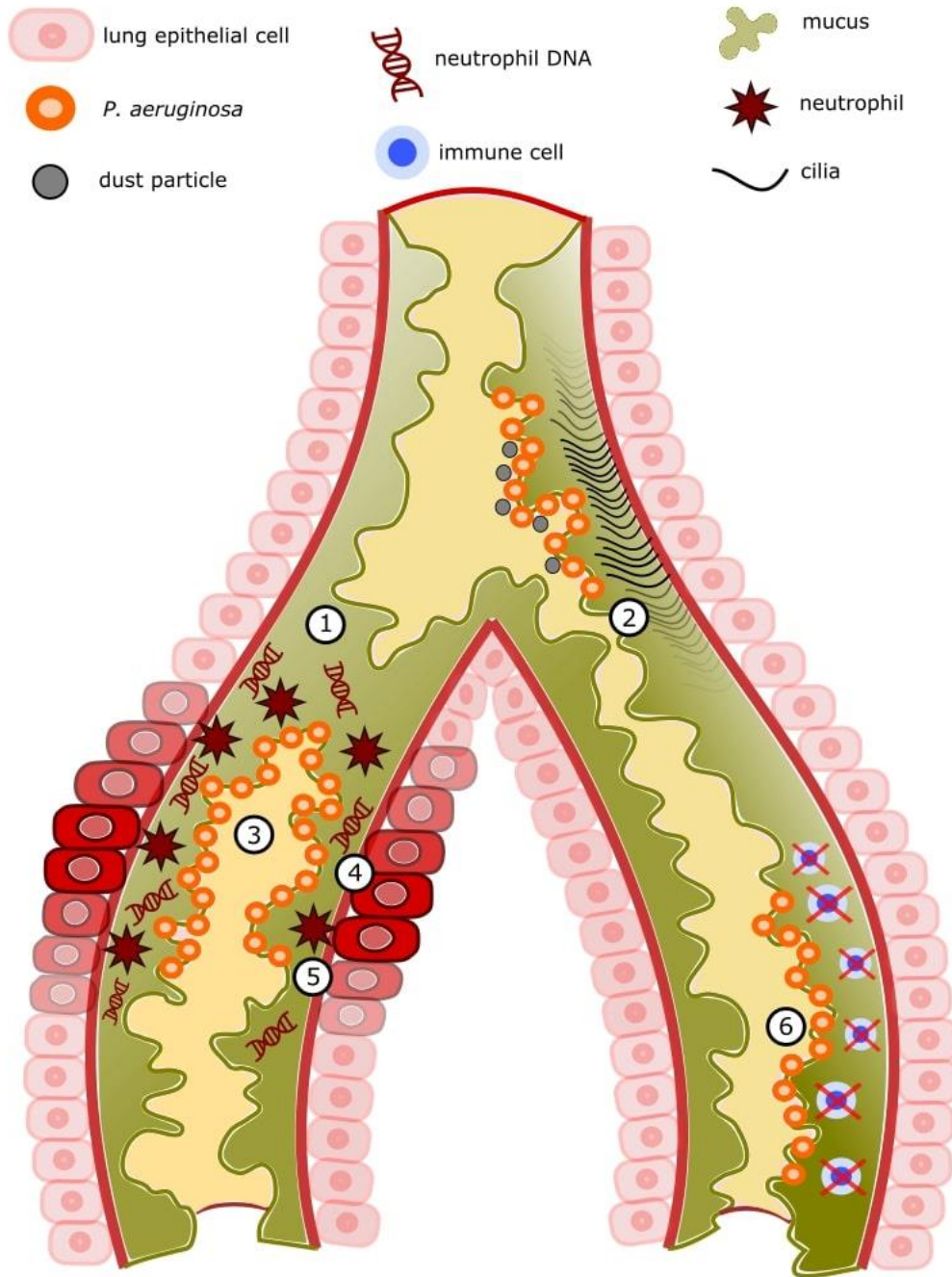
Fig. 1.2: Schematic of the CF airway. Large branched structure represents subsection of human airway with epithelial cells shown as pink rounded rectangles.

1) Mucus plugs (green globular shape) block airways.

2) Cilia movement (black curved lines) is impeded.

3) Lack of mucociliary clearance leads to chronic bacterial infection (orange circles).

4) Tissue damage leads to swelling, which further restricts airflow.

5) Neutrophils (red jagged shapes) release elastases, damaging collagen
   matrices, and DNA (red double helices), further thickening mucus.

6) Altered ASL pH reduces bacterial killing ability of immune cells (blue circles).

1.1.5 CF mutation classes

CF mutations have historically been categorised into six mutation classes [41].
Mutations resulting in no protein production whatsoever, often caused by premature stop
codons and resulting in the most severe clinical phenotypes, belong to Class I. The most
common CF-causing variant, F508del, results in misfolded protein and subsequent
degradation by the ER-associated degradation pathway; mutations of this type are known as
Class II. Class III mutations result in dysregulated CFTR channel activity, leading to reduced
open probability. Class IV mutations impact conductance of the ion channel pore, reducing
ion transport. Alterations in splicing pattern, which result in a decreased quantity of
functional protein product, are caused by Class V mutations. Finally, Class VI mutations
result in conformational changes that reduce retention at the apical membrane, leading to
accelerated turnover.
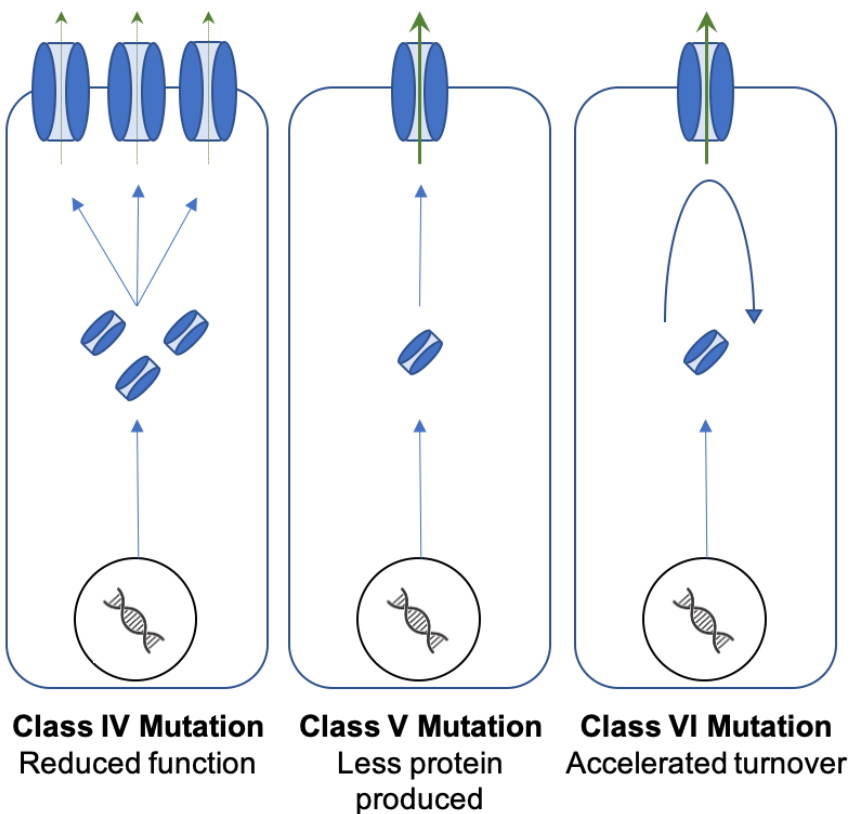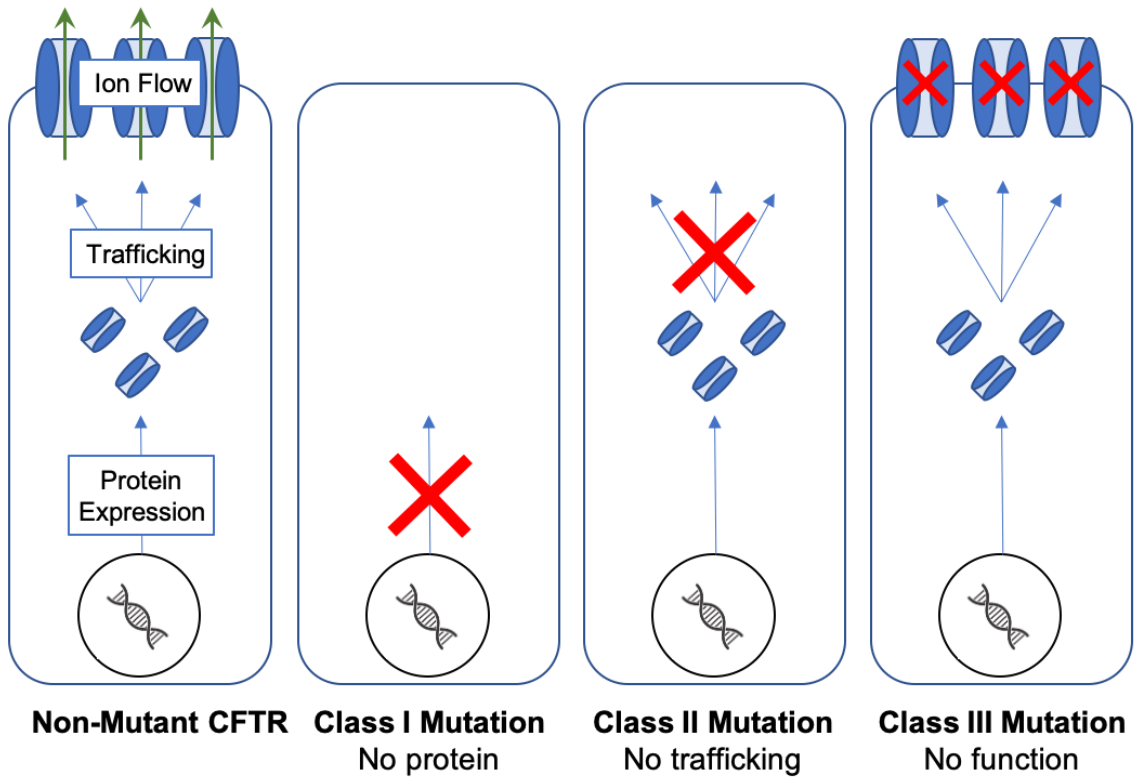
Fig. 1.3: Illustration of CFTR dysfunction due to different classes of mutation.

CFTR protein (blue shapes) is expressed from DNA and trafficked to cell

membrane (rounded rectangle). Ion flow indicated by green arrows; development of CFTR from DNA to mature protein to properly trafficked protein indicated by blue arrows; loss of indicated phenomenon indicated by red X.

In a study of CF variants in Europe and Israel, it was found that 51.3% of patients harboured a Class II/Class II genotype, with the second most frequent combination (Class I/Class II) trailing at 11.2% [42]. This is explained by the fact that F508del, the most common *CFTR* variant, is a Class II mutation. However, relative proportion of mutation classes varied with geographic population. For example, Class I variants were markedly more prevalent in Israel, while Class II was most prevalent in Ireland. Israel was the only country surveyed in which F508del was not the most common allele; among the Ashkenazi population, the most common CF-causing allele is the variant W1282X, a Class I mutation [43].

As previously stated, the most common *CFTR* mutation is F508del, a three-nucleotide deletion falling under Class II; this variant also exhibits class III (gating) and VI (peripheral stability) defects [17]. Nearly 90% of people with CF carry at least one copy of this allele [15]; therefore, pharmaceutical interventions have primarily focused on addressing this mutation.

1.1.6 Genetic modifiers and environmental factors impacting CF

Many attempts have been made to examine the association of non-*CFTR* genes' impacts on CF pathology. However, several studies have reported either conflicting or non-replicatable results. In a 2019 systematic review of reports of CF gene modifiers, Shanthikumar *et al* screened 140 reports for genes that have been positively identified by

multiple studies and not contradicted by any other studies [44]. No genes were identified

that directly interact with *CFTR* based on these criteria; however, several genes involved in

other CF-related processes were identified and reviewed for functional relevance.

Increased levels of endothelin receptor type A (*ENDRA*), a proinflammatory peptide,

was found to increase airway constriction and therefore lung disease severity. *TGFB1* and

MIF were implicated in inflammatory responses. Overexpression of the cytokine *TGFB1* has

been associated with increased inflammation and fibrosis; similarly, reduced expression of

the proinflammatory mediator *MIF* has been associated with increased lung function and

decreased rates of infection. *HFE* and *HMOX1* were both associated with infection

susceptibility. Loss of function of *HFE*, which encodes a protein crucial for maintaining iron

homeostasis, was found to be associated with more severe and faster-progressing loss of

lung function.

Several genes not directly related to inflammation were also identified. Methylation

of *GCLC*, a rate-limiting enzyme in synthesis of the antioxidant glutathione, and of the

glutathione transferase *GSTM3*, were implicated in more severe lung disease. Methylation

of *HMOX1*, which encodes the antioxidant agent heme oxygenase 1, was found to correlate

with more severe lung disease. The genes encoding the solute carrier proteins SLC9A3 and

SLC6A14 were identified via multiple genome-wide association studies and were found to

be associated with pleiotropic effects across the lung and digestive system. SLC9A3 is

implicated in sodium homeostasis, and SLC6A14 is a sodium- and chloride-dependent

transporter of amino acids; dysregulation of both sodium and chloride ions are a hallmark of

CF.

Furthermore, a systematic review of environmental impacts on CF disease identified

proximity to roadways, exposure to secondhand smoke, and general air pollution as

impactors of FEV_1, while warmer climates and indoor allergens were associated with higher incidence of lung infection [45]. Additionally, factors such as distance to treatment centres and income were also implicated in progression of lung disease and mortality rate.

1.1.7 Modulator therapies for CF

The latest pharmaceutical advancement in CF treatment combines three drugs of two different classes (two correctors and a potentiator). Corrector molecules are used to enhance protein folding into proper conformation for trafficking to and retention at the apical membrane [26]; the corrector Tezacaftor inserts among four alpha helices that make up the TMD1 domain, therefore stabilising their assembly (thought to occur during biogenesis in the ER) [46, 47]. Potentiator molecules function by increasing open probability of the ion channel pore at the cell surface [26]; the potentiator Ivacaftor binds to CFTR at the site of a critical "hinge" region in TMD1, stabilising the open conformation [48].. The drug Elexacaftor is thought to act as both corrector and potentiator. In its correcting capacity, Elexacaftor binds at a later stage than Tezacaftor does, associating with TMD helices to stabilise their interface with NBD1; this confers resistance to degradation by proteases in the ER [46]. It has also been shown to enhance ion transport in non-CF cells, suggesting that plays a role in potentiation as well as correction [49].

Ivacaftor was first of the three above compounds to receive FDA approval in 2012. It was identified by Vertex Pharmaceuticals via a screen to identify compounds increasing anion transport in cells harbouring the Class III *CFTR* gating mutant G551D; in theory, it could rescue other gating mutations as well [50]. It was later combined with Tezacaftor, also developed by Vertex; the drugs' use in combination was then approved for treatment of

patients harbouring the F508del variant in 2018 [51]. In 2019, a drug combining all three molecules (Tezacaftor, Elexacaftor, and Ivacaftor) was approved for use in patients in the USA [52]. This was followed shortly by its introduction to UK patients in 2020. This "triple-combination" drug is specifically engineered to correct defects caused by the F508del mutation. In a phase 3, randomised, double-blind, placebo-controlled trial, patients receiving the drug were found to have significant amelioration of pulmonary health as measured by various metrics: FEV_1, sweat chloride concentration, BMI, and frequency of pulmonary exacerbations requiring hospital visits and/or antibiotics [53].

However, for the approximate 10% of patients carrying Class I mutations that result in no rescuable protein product (i.e. nonsense, frameshift, and many splicing mutations), triple-combination therapy is not a viable treatment option [54]. For patients carrying "undruggable" mutations, molecular genetic approaches to treatment represent potential lifesaving care.

1.1.8 *In vitro* methods of modelling and studying CF

1.1.8.1 Cell lines

One of the most popular non-primary cell models in which to study CFTR and CF is the 16HBE14o- (hereafter 16HBE) cell line. Originally isolated from the bronchus of a one-year-old heart-lung transplanted patient, these cells were immortalised via transformation with a plasmid encoding a replication-defective simian virus 40 genome [55, 56]. Surviving cells were found to be functionally immortal and express CFTR. Furthermore, multiple cell lines have been derived from 16HBEs via gene editing, commonly referred to as 16HBEge [57]. These lines are often generated to carry specific CF-causing mutations for modelling and study. For example, lines harbouring F508del, G542X, and W1282X mutations have

been generated via nucleofection of CRISPR RNP and ssODN HDR template [57]. The F508del 16HBEge line was shown to exhibit negligible CFTR-mediated ion transport, an effect rescuable by treatment with a CFTR corrector. 16HBEge lines harbouring nonsense variants (W1282X and G542X) exhibited nonsense-mediated decay which was shown to be rescuable by treatment with molecules inducing readthrough. A more recent study utilised 16HBEge cells carrying early (Y122X), middle (G542X), or late (W1282X and R1162X) nonsense variants to provide evidence that NMD pathway inhibition could have an additive effect on rescue by readthrough therapy [58],

1.1.8.2 Primary cell types and collection methods

Primary cells for CF study are commonly harvested from the upper (nasal) and lower (tracheal and bronchial) airways via brushings. Nasal brushings are simple to perform: a brush is inserted into the nose, epithelial surfaces briefly swabbed, and brush withdrawn [59]. There is typically no endoscope or anesthesia required, and the procedure can be performed in a few minutes. In lower airway sampling, target cells are much less accessible, and thus a more complex procedure is required. An endoscope is typically inserted prior to the brush in order to ensure proper location of sampling. So that no upper airway cells are picked up by the brush during insertion or withdrawal, brushes can be housed in a retractable sheath, or patients can be intubated to provide a physical barrier around the advancing brush. During this time, the volunteer is normally placed under general or local anesthetic [60, 61]. Because of the more complex and demanding nature of the procedure, bronchial samples are commonly collected while a patient is already undergoing an invasive procedure, whereas nasal samples can be obtained during a routine appointment.

1.1.8.3 Extending proliferative potential

Samples of primary airway epithelial cells are invaluable tools for modelling and studying the human airway. However, such cell populations have very limited proliferative potential, and tend to be an effective model for only 2-4 passages, representing 3-4 population doublings [62].

The proliferative potential of keratinocyte, mammary, and prostate epithelial cells have all been successfully extended via co-culture with growth-arrested feeder cells and treatment with ROCK inhibitors [63, 64]. Feeder cells that have been growth-arrested via irradiation or chemical inhibition form a layer of non-dividing cells. These cells secrete compounds that stimulate growth; while the exact identities of these compounds are not entirely elucidated, they are known to include insulin-like growth factor 2 [65]. Alongside feeder cells, supplementation of cell media with chemical ROCK inhibitors has been shown to promote expression of telomerase, as well as remodelling of the actin cytoskeleton, in a manner that promotes proliferation [63, 64]. However, long-term sustainment of this specialised cell culture is complex, time-consuming, and requires mixing of cells of interest with feeder cells. Therefore, a single-dose genetic modification of primary cells enabling long-term culture is favourable.

Lentiviral transduction with the mouse polycomb complex gene *BMI-1* has been shown to significantly extend primary cells' proliferative and differentiation potential *in vitro*. Munye *et al* reported that *BMI-1* transduced normal human bronchial epithelial (NHBE BMI-1) cells maintained healthy morphology, karyotype, and doubling time for up to 25 passages [66]. Air-liquid interface (ALI) cultures were differentiated from NHBE BMI-1 cells at passage 2.>10 (transduced at passage 2, maintained until passage 2 + >10) were found to form pseudostratified epithelia. These epithelia exhibited cilia beat frequency,

transepithelial electric resistance (TEER), and ion channel activity comparable to those of untransduced passage 1 and 2 NHBEs [66]. Additionally, transcriptomic comparison of cells with or without *BMI-1* expression showed that 97.6% of genes showed no significant differences in expression; those that did show differences were primarily related to cell cycle regulation or epithelial-mesenchymal transition [67].

1.1.8.4 Air-liquid interface (ALI) culture

Underneath the outermost epithelial layer lies a population of airway basal cells, which are capable of proliferation and differentiation [68]. ALI culture is a method of stimulating undifferentiated airway basal cells to form a structure more closely resembling the airway epithelium [69]. The technique relies on "transwell" inserts, which slot into tissue culture plate wells. Transwells contain micropores (typically 0.4 µm in diameter) that allow movement of liquid but not cells across the membrane.

Basal cells are initially seeded on transwells and submerged in growth media. Filling the basal chamber also allows for liquid to contact cells' basolateral sides through the membrane pores. Once cells have formed a confluent monolayer, the submerging apical media is removed and the growth media in the basal chamber is replaced with a specialised differentiation-stimulating media; this step is commonly referred to as "airlifting" of cultures. Over approximately 28 days, basal cells develop into an electrically resistant pseudostratified epithelium [69]. This structure consists of multiple cell types: basal cells, which are capable of proliferation and multipotent differentiation; club cells, which can differentiate into ciliated and goblet cells, as well as de-differentiate into basal cells; goblet cells, which generate and secrete mucus; ciliated cells, whose cilia sweep dirt and pathogens out of the lungs; tuft cells, which are thought to have both chemosensory and

immunological function; and ionocytes, which exhibit high CFTR expression [70, 71].

Transcriptomic analysis of these cultures has revealed them to be highly similar to those found in airway biopsies [72, 73]. However, it has also been shown that use of different differentiation-stimulating media can lead to variation in phenotypic profiles. When commonly used media were compared for differentiation of NHBE cells, PneumaCult medium was found to most closely resemble the *in vivo* airway in terms of mucosal production and composition, proportional representation of cell types, and cilia length and beat frequency [74].
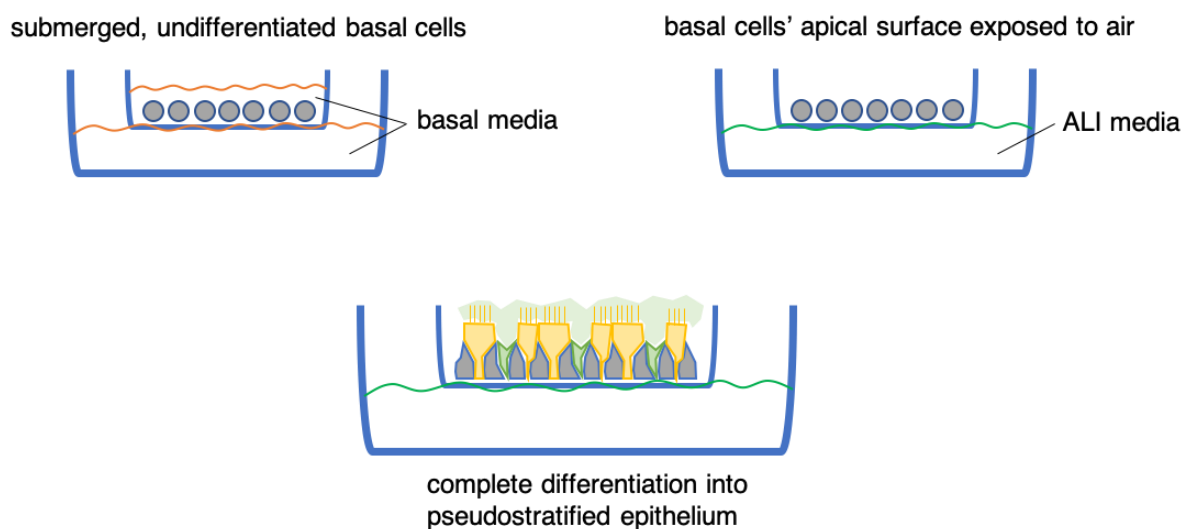
Fig. 1.4: Schematic of ALI culture. Basal cells are seeded on porous transwells such that their basolateral surfaces contact liquid and are apically submerged. Once a confluent monolayer is formed, apical media is aspirated, and basolateral media replaced with differentiation-stimulating media. Over 4 weeks, a pseudostratified epithelium forms.

ALI cultures derived from primary nasal and bronchial epithelial cells have been compared using various functional readouts. One study used paired samples (*i.e.* nasal and bronchial samples obtained from the same individuals)[75]. Cells were differentiated on ALI and various parameters were examined. Investigated parameters included trans-epithelial electrical resistance (TEER), representation of different cell types, and CFTR ion transport ability (measured via Ussing chamber analysis, a technique described in Section 1.1.7.5). A significant amount of variation was found in the differences between individuals' paired bronchial and nasal cell types. For example, some individuals' nasal vs. bronchial TEER was significantly different, while others showed no significant difference. This trend was also observed in measurements of CFTR ion transport. However, pooled comparison of all individuals' nasal vs. bronchial epithelial samples revealed no significant difference between cell types in ion transport or TEER [75].

1.1.8.5 Intestinal organoids: a non-airway model

One commonly used model is the intestinal organoid, or mini-gut. Organoids are formed from self-renewing stem cells residing in intestinal crypt structures. In an initial study, single mouse crypts were isolated and cultured, and within 14 days, had developed into closed multi-crypt structures [76]. Soon after, a study was performed investigating CFTR activity in such organoids [77]. CFTR protein abundance was assayed via immunoblotting in freshly biopsied crypts and week-old organoids and found to be highly similar. CFTR activity was assayed by detecting changes in membrane potential after drug-induced stimulation or inhibition. Treatment with forskolin (a drug that stimulates *CFTR*-mediated ion transport by increasing intracellular cAMP) resulted in depolarisation of the organoid membrane; subsequent treatment with CFTRihh-172 (a drug that allosterically inhibits CFTR protein in a

manner that favours the closed-channel state) resulted in repolarisation. Conversely, organoids from CFTR-knockout mice displayed no response to forskolin or CFTRinh-172 [77]. This led to the development of the forskolin-induced swelling (FIS) assay. Treatment with forskolin leads to CFTR-mediated influx of ions into the lumen of organoids, generating osmotic pressure, which causes an influx of water and swelling of the entire organoid structure [78]. This can be used to assess response to different treatment strategies including drugs [79] and genetic medicine [80-83].

1.1.9 *In vivo* methods for studying CF

Beyond cell and organoid models, there are limited *in vivo* systems for preclinical trialling of CF treatment strategies. One model of CF lung pathology relies on overexpression of another ion channel protein, the epithelium sodium channel ENaC. Like CFTR, ENaC resides at the apical membranes of airway epithelial cells and is responsible for trafficking sodium ions through the epithelial surface in order to generate transepithelial osmotic pressure [84]. However, while CFTR primarily expels Cl^- ions into ASL, ENaC absorbs Na^+ ions from ASL. This reciprocal activity ties CFTR and ENaC together in an ionic balancing act that plays a major role in regulation of ASL and mucus consistency. An ENaC-overexpressing mouse model has been found to recapitulate pathogenic hallmarks of CF, including thickened plug-forming mucus, inflammation, and death due to lung disease [85]. This is an invaluable tool for testing mucolytics, modulator drugs, and receptor-targeted nanoparticle (RTN) delivery. However, because this model does not rely on mutant *CFTR*, it is less useful for DNA editing strategies. Some conclusions could certainly be drawn from the effect of adding an extra *CFTR*-expressing cDNA construct to supplement endogenous WT expression;

however, this is not necessarily a faithful readout of such a strategy's ability to compensate for dysfunctional endogenous CFTR.

Various animal models have been generated via knockout of endogenous *CFTR*, but none thus far recapitulate all phenotypes of human disease. For example, *CFTR*-knockout (KO) rats do not display evidence of enhanced airway inflammation or lung infection, despite exhibiting evidence of mucosal plugging in lower airways, reduced ASL height, and reduced mucociliary clearance [86]. Mouse models with disrupted CFTR generally do not develop spontaneous lung disease. A congenic KO mouse line did display lung dysfunction, but its use is complicated by high mortality rate due to intestinal issues, necessitating tissue-specific rescue via expression of an hCFTR construct in the intestine [87]. Larger animal models show more promise: pigs and ferrets show reduced mucociliary clearance, progressive inflammation, mucus plugging, and chronic infection [86]. However, study in these models is more demanding in terms of money, space, time, and ethical approval when compared with smaller animals. Additionally, in CF mutant pigs, meconium ileus typically prove fatal to neonates; therefore, intestine-specific *CFTR* expression cassettes are required to maintain animals until pulmonary symptoms arise [88]. More recently, CF sheep and rabbits have been investigated; despite recapitulating intestinal and select upper airway phenotypes, rabbits did not exhibit lower airway disease [89]; mutant sheep exhibited meconium ileus at similar rates to pigs, and therefore would likely require similar conditional rescue [90]. Therefore, ALI culture remains a crucial model to assay DNA editing's effects on CFTR function in the airway.

1.1.10 CFTR's interactions with an epithelial sodium channel (ENaC)

It has long been proposed that ENaC is inhibited by CFTR and that loss of CFTR function therefore leads to ENaC overactivity [91]. Proteolytic cleavage by two additional CF-relevant agents have also been found to stimulate ENaC: elastases produced by neutrophils [92] and proteases released from *Pseudomonas aeruginosa*, one of the most common bacterial invaders of the CF lung [93]. Unbalanced sequestering of Na^+ contributes further to dehydration of ASL and mucus, as evidenced by the fact that ENaC overexpression in the mouse airway leads to a phenotypic profile nearly identical to that of CFTR loss in humans [94]. Furthermore, hypomorphic alleles of ENaC have been found to correlate with slower lung disease progression in CF patients carrying the F508del mutation [95].

Given these findings, it has been proposed that inhibition of ENaC might ameliorate the CF lung phenotype. ENaC inhibition by siRNA has also been shown to rescue CF phenotypes in ALI-differentiated cells, such as ASL height, cilia beat frequency, and fluid absorption rate [96]. It has further been proposed that ENaC inhibition could complement the effects of CFTR modulators by contributing to polarisation of apical cell membranes, thereby stimulating activity of any rescued CFTR at the surface [84]. A 2020 study investigated a novel ENaC inhibitor (BI 1265162) on CF and healthy ALI cultures [97]. Treatment with BI 1265162 or the CFTR modulators ivacaftor/lumacaftor partially rescued water retention at the apical surface. Furthermore, combining BI 1265162 with the modulators rescued water retention to a level similar to that seen in non-CF cells treated with modulators. Taken together, this evidence suggests that ENaC inhibition could provide relief to CF patients struggling with lung disease regardless of their eligibility for modulator therapies and could enhance eligible patients' response to the drugs.

1.2 DNA Editing

1.2.1 Approaches relying on non-Cas9 nucleases

Broadly, genome editing can be defined as any process that makes permanent alterations to the sequence of genomic DNA. In recent years, CRISPR-Cas9 (described in detail in 1.2.2) has become the dominant force in the genome editing field. However, other methods of DNA editing have existed for decades and are still in use.

1.2.1.1 Restriction enzymes and meganucleases

One of the earliest forms of DNA editing was achieved using restriction enzymes (REs). REs are nuclease proteins, derived from bacteria, that bind and cleave DNA. The first RE capable of cleaving at specific sequence, leaving behind predictable ends, was described in 1970 [98]. Since then, thousands of such REs have been described [99]. The process of RE digestion and subsequent ligation of different DNA molecules has been used for decades to generate recombinant DNA. For example, RE digestion and subsequent ligation remains an extremely common technique for plasmid subcloning [100, 101].

Meganucleases and homing endonucleases were also invaluable tools for early genome editing and remain in use to some extent today. These proteins recognise and bind longer sequences of DNA than REs (up to 40 bp), meaning that they are capable of recognising unique sequences in the genome [102]. The LAGLIDADG family (named for a conserved sequence of amino acids in the active site) includes two commonly used nucleases, I-CreI and I-SceI. I-CreI, first characterised in the algae *Chlamydomonas reinhardtil*, acts as a homodimer to recognise a 22 bp sequence; later studies identified the amino acids crucial for DNA sequence recognition and engineered versions of the protein

specific [103-105]. An engineered version of this endonuclease forms the basis of the ARCUS system, which has been used in chimeric antigen receptor T cell production [106] and elimination of latent hepatitis B virus [107]. The I-SceI nuclease, originally characterised in *Saccharomyces cerevisiae*, underwent a directed evolution process to yield versions of the protein with altered DNA recognition specificity [108]. Recently, I-SceI itself as a directed evolution tool to assess the cutting efficiency of other nucleases [109]. A plasmid encoding I-SceI was expressed in *E. coli* such that its activity would disrupt the bacterial genome and kill the cell; a secondary plasmid encoding another nuclease was then introduced such that the activity of the second nuclease would destroy the I-SceI expression construct, thus rescuing the bacteria.

1.2.1.2 Zinc finger nucleases (ZFNs)

ZFNs combine the DNA recognition abilities of zinc fingers (originally described in *X. laevis*) and the DNA cleavage abilities of the *FokI* endonuclease (originally derived from flavobacterium) [110]. Zinc fingers are small protein structural motifs with "finger" domains that extend to bind with DNA. The amino acid sequence of these finger domains can be used to direct binding to specific DNA sequences. Typically, one ZF targets 3-4 nucleotides worth of DNA; therefore, a tandem array of six ZFs can uniquely bind an 18-bp sequence of DNA, which theoretically represents uniqueness in the genome [111]. These engineered DNA recognition proteins are fused to *FokI* nucleases to form ZFNs. *FokI* activity requires dimerisation; therefore, two ZFNs are designed to target upstream and downstream of the desired cut site. Upon binding to the DNA in the proper orientation, each ZFN contributes its *FokI* cargo to a dimer, which induces a DSB in the DNA at the desired locus [111, 112].

1.2.1.3 Transcription activator-like effector (TALE) proteins

The DNA-binding domain of TALE proteins contains a repeat domain. Within the repeat domain, two hypervariable amino acids dictate specificity to a nucleotide [113, 114]. Much like in the case of ZFNs, this represents an opportunity to construct arrays of TALE domains that specify binding to a unique DNA sequence of the genome. Also similarly to ZFNs, pairs of customised TALE proteins can be paired with *FokI* nuclease to form TALE-nucleases (TALENs) [115, 116].

1.2.2 Fundamentals of CRISPR-Cas9

As early as the mid-2000s, researchers had begun studying a naturally occurring bacterial adaptive immunity system that essentially allowed bacterial cells to create customisable endonucleases [117]. This system, dubbed "CRISPR-Cas9," was adapted over the next decade for use in mammalian cells [118-120]. This technology revolutionised the field of DNA editing with its precision and efficiency.

1.2.2.1 Origins of CRISPR-Cas9

The CRISPR-Cas9 DNA editing system is derived from a system of adaptive immunity found in prokaryotes [121]. Upon infection by agents introducing foreign DNA (i.e. bacteriophages), cells incorporate short lengths of invasive viral genetic material (spacers) into sections of the genome called clustered regularly interspaced palindromic repeats (CRISPR). The RNAs produced by these incorporated sequences are used as "guides" to search for homology to previously encountered foreign DNA. Complexed with these CRISPR RNAs (crRNAs) are CRISPR-associated nuclease proteins (in *S. pyogenes*, Cas9). Their association is stabilised by a scaffolding molecule (trans-activating CRISPR RNA or tracrRNA);

the combination of tracrRNA and crRNA is commonly called a guide RNA (gRNA, sometimes called single guide RNA or sgRNA) [118]. The Cas9 ribonucleoprotein (RNP) searches the genome until it finds a site with homology to the spacer sequence encoded by the gRNA. Upon binding to a protospacer-adjacent motif (PAM) site (NGG in wild type *Sp*Cas9), Cas9 cuts both strands of the DNA three nucleotides upstream of the PAM, creating a double-strand break (DSB). This is typically followed by some degree of resection of the cut strands [122].
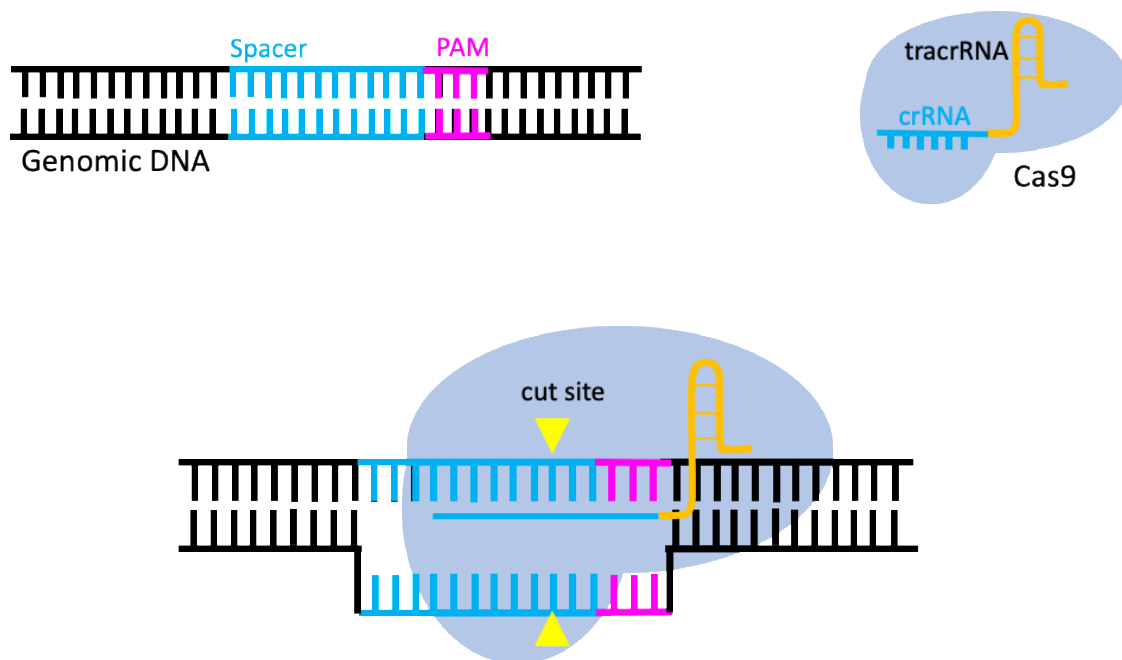
Fig. 1.5: Schematic of CRISPR-Cas9 binding and cutting genomic DNA based on homology to gRNA sequence. Cas9 protein (blue shape) complexes with guide RNA, which itself consists of a tracrRNA (yellow hairpin) and crRNA (blue). DNA is cut three bp upstream of the PAM (pink). Yellow triangles indicate site of DSB formation.

By 2013, the necessary components of a CRISPR reaction (tracrRNA, crRNA, and Cas9), had been elucidated in both bacterial and *in vitro* free DNA [118, 123, 124]. That year, two groups simultaneously reported successful cleavage of genomic DNA in mammalian cells using the CRISPR system [119, 120]. By designing a custom crRNA that targeted the desired locus in the genome, Cas9 was shown to function as a customisable restriction enzyme to create DSBs at any genomic position.

Substantial elucidation of the CRISPR-mediated DSB formation has been performed. Cas9 gains nuclease activity only when complexed with a gRNA. Upon association of the gRNA with the DNA strand complementary to the crRNA sequence, Cas9 cuts both strands three base pairs upstream of the PAM motif. The RuvC domain cuts the noncomplementary strand, and the HNH domain cuts the complementary strand [125]. Cas proteins have also been characterised and adapted from other bacterial hosts, including AsCas12a (*Acidaminococcus sp.*, PAM sequence TTTV, generates sticky ends) and SaCas9 (*Staphlococcus aureus*, PAM sequence NNGRRT) [126]. Furthermore, engineering of Cas proteins has led to nucleases with relaxed PAM specificity [127] and smaller protein size for ease of delivery [128]. A major concern with use of any Cas protein is that of immunogenicity, although studies on the prevalence of antibodies against Cas9 in the general population have yielded drastically different results, ranging from <5% to >90% [129]. Proposed methods to avoid this issue include engineering Cas9 to remove immunogenic epitopes and performing editing *ex vivo* [130].

1.2.2.2 DNA repair after DSBs

Mammalian cells use two main pathways to repair DSBs: homology-directed repair (HDR) [131] and non-homologous end joining (NHEJ) [132].

NHEJ consists of ligation of the ends of the DSB, with repair not based on intact template DNA, but rather based on micro-homologies to single-stranded overhangs at the DSB site [133]. Upon DSB formation, the Ku heterodimer recognises and binds the blunt ends of the DSB. DNA-dependent protein kinases (DNA-PKcs) stabilises the DSB by keeping the ends in close proximity. The Artemis protein exerts exonuclease activity on any hairpin structures and/or single-stranded overhangs, while DNA polymerases fill in missing nucleotides. Finally, ligation occurs, mediated by Ligase IV [134]. If overhangs together preserve the original sequence found at the DSB site, then base-perfect repair is possible. However, this is not always the case, and NHEJ often results in significantly less precise repair, with unpredictable insertions and/or deletions (indels) frequently occurring.

A non-canonical form of NHEJ uses microhomology to repair DNA. This pathway, known as microhomology-mediated end joining (MMEJ) tends to result in larger deletions [133]. MMEJ is initiated when MRE11 resects one DNA strand from the 5' end; this nick allows infiltration by nucleases including EXO1 and DNA2 that generate sections of ssDNA. ssDNA flaps anneal based on sequence complementation, and proofreading-deficient polymerases fill in gaps before a final ligation step joins the two ends. This frequently results in larger deletions than those observed in canonical NHEJ [135].

A few studies have attempted to elucidate CRISPR-mediated NHEJ and MMEJ. One such study found that target sequence can determine preference for NHEJ vs. MMEJ in a manner conserved across cell types, and that MMEJ deletion frequency served as a predictor of HDR efficiency [136]. Another study recapitulated that NHEJ is sequence-dependent, and further characterised indels to provide evidence that the process is more error-prone at the PAM-distal end than the PAM-proximal [137]. It has also been shown that NHEJ-mediated 1 bp insertions are typically templated by the nucleotide located

immediately upstream of the cut. The identity of this nucleotide also correlated with whether an insertion or deletion was formed, with T/A associated with insertions and G/C associated with deletions. These trends were leveraged to predict repair outcomes by programming microhomology into target sequences [138].

HDR uses the homologous chromosome as a template to rebuild damaged DNA. Upon DSB formation, a complex consisting of MRE11, RAD50, and NBS1 (the MRN complex) binds exposed ends, lending stability to the DSB and preventing breakage of the chromosome. This complex also recruits ataxia telangiectasia mutated (ATM), which phosphorylates H2A histone family member X (H2AX); this in turn recruits DNA damage checkpoint protein 1 (MDC1). After this initial stabilisation step, 5' exonucleases (CtIP or EXO/BLM) resect ends to generate single-stranded 3' overhangs. The resultant ssDNA is then coated by human replication protein A (RPA), which is subsequently replaced by DNA repair protein RAD51, a process aided by BRCA1. RAD51 catalyses strand invasion and annealing of either the sister chromatid or an experimenter-provided template molecule, forming a D-loop structure. D-loops are then extended by DNA synthesis and ligated via the double Holliday junction pathway, yielding precise homology-directed repair [131, 134, 139, 140]. This generates precise and base-perfect repair of the damaged DNA. Experimenters can provide a customised template of DNA with homology arms that mimic the homologous chromosome and altered intervening sequence, leading to incorporation of any desired edits [141]. HDR has been used successfully to induce a variety of edits, from single-nucleotide substitutions [142], to large-scale insertion [143] or deletion/replacement edits [144]. However, HDR only occurs in actively dividing cells, as it relies on presence of a sister chromatid alongside the DSB, which only occurs during S/G2 phases of the cell cycle [134]. Furthermore, even in actively dividing cells, NHEJ tends to dominate [134].
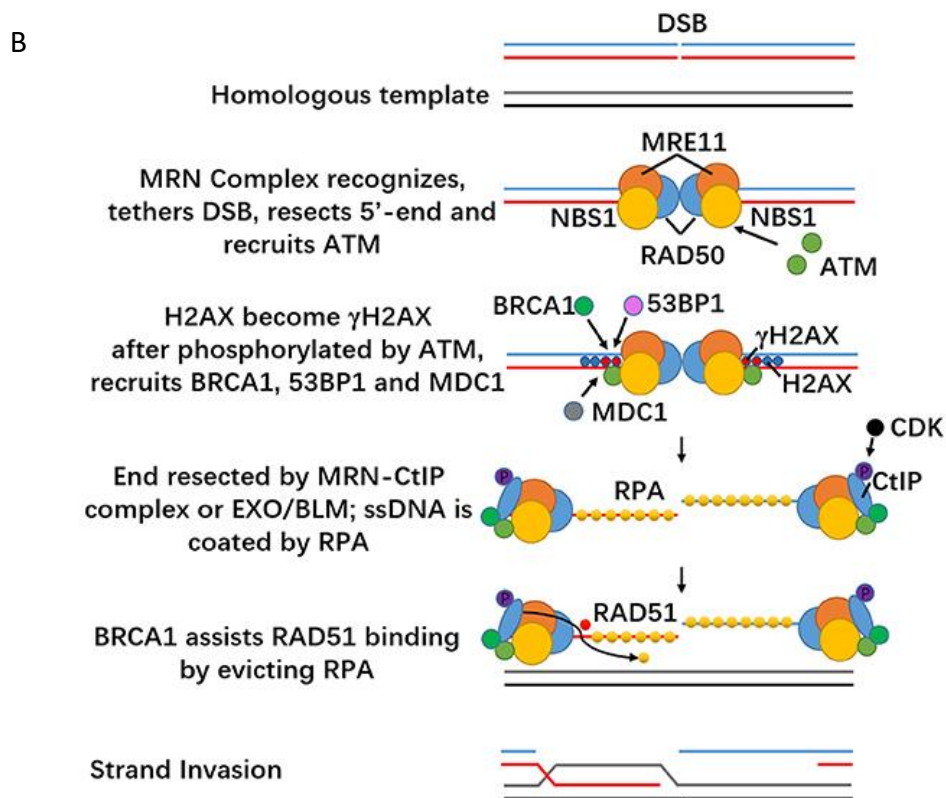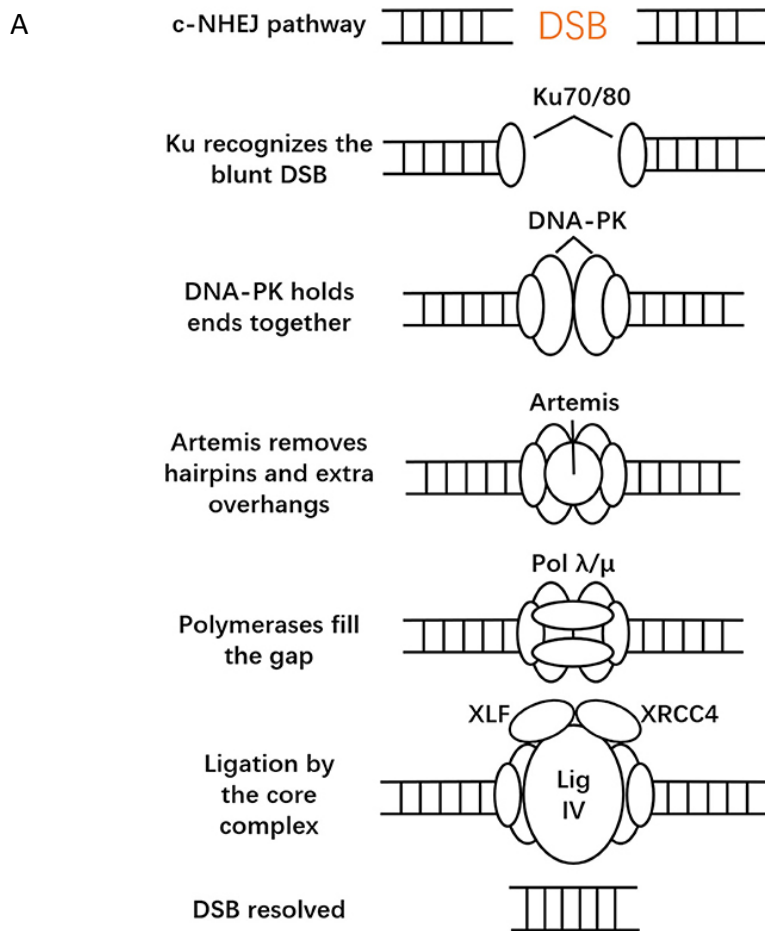
Fig. 1.6: Schematics of main DNA repair pathways. A) Schematic of NHEJ. DNA (ladder shape) is bound by proteins to stabilise the DSB before other enzymes trim, fill in, and ligate strands. B) Schematic of HDR. Proteins (coloured shapes) stabilise and resect ends of DNA (solid lines) before recruiting further machinery to facilitate strand invasion with the homologous template. DNA is rebuilt according to homology with the template. From Liu, M., *et al.*, *Methodologies for Improving HDR Efficiency.* Front Genet, 2018. **9**: p. 691[134].

NHEJ is useful for knocking out genes. Introducing two ribonucleoproteins (RNPs) carrying two gRNAs distant from one another can essentially "cut out" large sections of DNA, with the two resulting DSBs being resolved into one truncated repair product [144, 145]. Alternatively, indel formation can be relied upon to some extent to introduce frameshift [146] and/or splicing [147] mutations using a single gRNA targeting a locus early in a gene. In addition to knocking out genes, the dual RNP strategy can be used to excise deleterious genetic features. For example, myotonic dystrophy type 1 is caused by repeat expansion of a CTG motif in the gene *DMPK*. Researchers have developed a dual Cas9 approach to excise these repeats via injection of two recombinant AAV vectors carrying CRISPR machinery. This strategy was shown to be effective in restoring normal *DMPK* transcripts in patient-derived muscle cell lines and a myotonic dystrophy mouse model [148], and is now currently under clinical testing [149].

It is also possible to use NHEJ to insert large DNA constructs into the genome by introducing constructs as end ligation is taking place [143, 150]. However, due to the lack of

homology-directed repair in canonical NHEJ, it is generally not possible to ensure that constructs insert in the proper orientation [151].

1.2.2.3 Homology-independent targeted integration: a more precise NHEJ-based approach

Homology-independent targeted integration (HITI) is an NHEJ-based technique for inserting large constructs of DNA in a specific orientation. The technique relies on inclusion of a reverse-complemented gRNA sequence (RC-gRNA) flanking the desired DNA insert. When donor template is introduced alongside RNP, the RNP cuts both the target genomic locus and the donor construct. If the donor inserts in the reverse orientation, the residual ends of the RC-gRNA restore the spacer sequence recognisable by the RNP, and Cas9 can re-bind and re-excise the construct. However, if the donor inserts in the proper orientation, the spacer sequence at either end of the insert is destroyed, and Cas9 can no longer bind and cut. Therefore, in theory, the donor construct will be continually excised until it successfully inserts in the proper orientation [152].
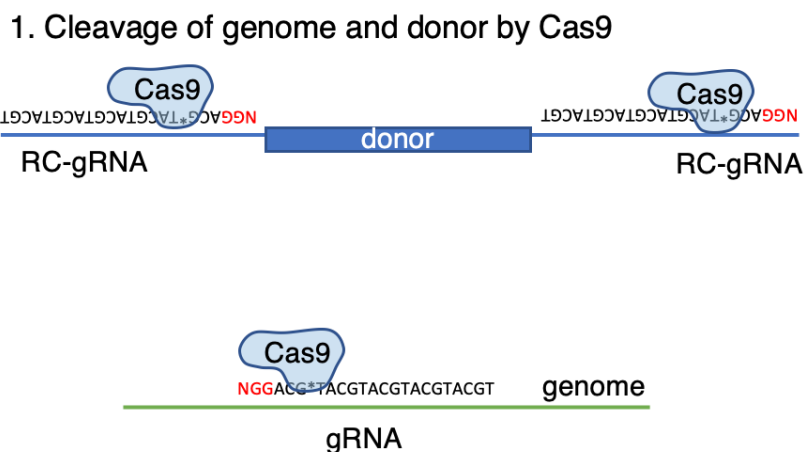
2. Genomic integration of cleaved donor
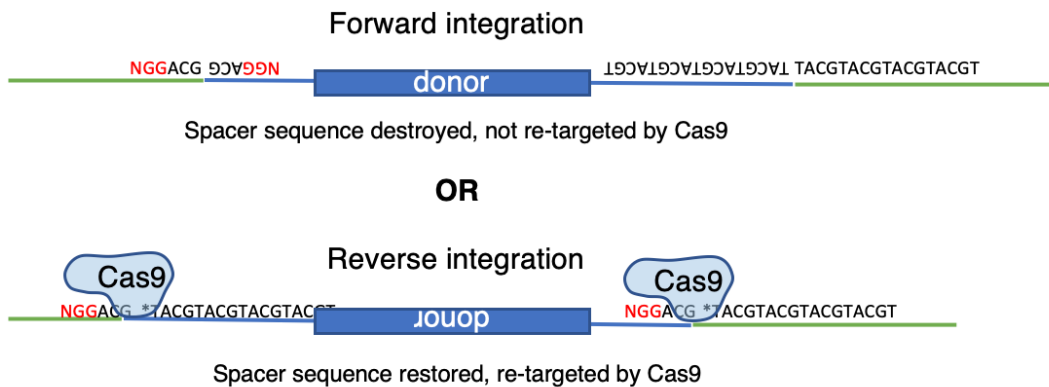
Fig. 1.7: Schematic of HITI. Donor molecule (blue rectangle) includes two copies of the reverse-complemented target sequence (RC-gRNA) flanking the desired insert. This acts as a substrate for Cas9, along with the genome. Upon reverse integration, Cas9 re-cuts the genome; upon forward integration, the spacer sequence is no longer recognisable by the RNP.

In its initial report, HITI was used to successfully insert a GFP cassette into non-dividing mouse primary neurons. After editing, 56% of successfully transfected cells expressed GFP [152]. Since then, HITI has been used to insert partial or full-length cDNAs to cell models of various genetic disorders, which has been shown to be a valid strategy for rescue of disease phenotypes; several such studies have been reported in 2023 alone. For example, a superexon of the *CTNS* gene was integrated via HITI into proximal tubule epithelial cells mutant for *CTNS* (a disease model for nephropathic cystinosis); stable expression of mRNA was detected for >40 days, and a decrease in lysosomal cysteine accumulation was found, indicating functional rescue [153]. In another study, exons 1-19 of the *DMD* gene were integrated into a disease model for Duchenne's muscular dystrophy; 1.4% editing efficiency in the heart led to 11% rescue of dystrophin levels, well above the

predicted threshold for clinical benefit [154]. Another study used HITI to knock in the *RHO*

gene, which encodes the light-sensing receptor rhodopsin, with up to 43% efficiency;

successfully edited cells exhibited higher light sensitivity and rescue of retina structure

[155].

1.2.3 Fusion of Cas9 with preexisting customisable nucleases

Catalytically dead Cas9 (dCas9) has been successfully fused with Fok1 dimers (the

nuclease proteins forming the basis of TALEN and ZFNs, described in Section 1.2.1) [156].

Because Fok1 operates as a dimer, two fusion proteins would need to localise to sites just

upstream and downstream of the desired cut site. This in theory provides fewer

opportunities for off-target binding and cleaving in the genome. Indeed, editing with paired

dCas9-Fok1 constructs yielded higher ratios of on-target:off-target editing; however, overall

on-target editing efficiency was significantly decreased. Additionally, it was revealed that

spacing between the editing constructs was very stringent; the number of PAM-containing

loci with the appropriate space between them represents a limitation on targetable sites.

1.3 CRISPR-Cas9 in practice

Two of the major considerations in the methodology of CRISPR-based DNA editing

are the design of gRNAs and prevention/detection of off-target edits (OTEs).

1.3.1 gRNA design to enhance editing

Multiple web-based tools have been developed for *in silico* design of gRNAs to target

specific genomic sites. These algorithms typically take into account a set of

recommendations that have been empirically established over the years: optimal GC content, predicted formation of secondary structures, and high specificity to the target locus.

1.3.1.1 GC content

In a study examining the editing efficiency of over 1,800 gRNAs, gRNAs with ~50% GC content were found to yield the highest editing efficiency [157]. Specifically, later studies established the importance of GC residues specifically at the PAM-distal end of the guide sequence, with "GG" the optimal dinucleotide immediately adjacent to the PAM [158, 159]. This was hypothesized to be due to preferential Cas9 loading [160, 161], as association of Cas9 and gRNA generates an activating conformational change in Cas9 [118]. Despite the fact that enhanced on-target editing does not necessarily correlate with reduced OTEs, it was established in another study that GC content greater than 70% significantly increased incidence of OTE events; therefore, GC content of around 50% are optimal for enhancing on-target specificity and reducing OTEs [162].

1.3.1.2 Secondary structure

Multiple studies have examined the self-folding Gibbs free energy of high- and low-efficiency gRNA seed sequences. One such study calculated the potential for secondary structure formation by examining self-folding Gibbs free energy for known high-efficiency and low-efficiency gRNAs [163]. Another study examined low-efficiency gRNAs that formed stable secondary structures, generated point mutations to decrease their folding potential, and subsequently re-assessed their editing efficiency *in vitro* [164]. Both methods found that decreased secondary structure formation correlated with higher DNA cleavage

efficiency. Furthermore, it was predicted that nucleotides at positions 21-50 (i.e. just beyond the seed sequence) formed a stable stem-loop structure. However, extension of this stem-loop structure via complementarity between nucleotides 51-53 and 18-20 meant that the final nucleotides of the seed sequence were rendered inaccessible [163].

Another study hypothesized that, if gRNAs were allowed to fold as they were transcribed, unfavourable secondary structures could form. Heating and cooling of transcribed gRNAs in order to allow for refolding was found to significantly enhance editing efficiency of previously inactive gRNAs [165].

Taken together, these results suggest that while some secondary structure is essential (i.e. the tracrRNA stem loop) for Cas9 loading, any potential for folding that impedes access to the seed sequence is unfavourable for DNA cleavage. Structure of DNA at the target site is also important to take into consideration, as chromatin structure has been shown to impact accessibility of DNA to CRISPR machinery, as well as playing a role in repair pathway choice [164, 166].

1.3.2 Quantification of editing efficiency in pooled samples

Analysis of editing events in individual single-cell clones (SSCs) can be used to estimate editing efficiency. However, generation of SSCs is a time-consuming and expensive process, and many cell types cannot tolerate the biophysical stresses involved with SSC generation (serial dilution or machine-mediated cell sorting). Therefore, for early proof-of-concept experiments (such as those establishing gRNA cleavage efficiency), analysis of genomic DNA derived from pooled cells is a valuable approach.

1.3.2.1 T7E1 assay

The T7E1 assay relies on a nuclease that specifically cleaves DNA at mismatched nucleotides. After a DNA editing experiment, genomic DNA is isolated from pools of treated cells and the target locus is PCR amplified. Amplicons are heated and cooled in order to form heteroduplexes of edited and unedited DNA. Treatment with the mismatch sensitive T7 endonuclease cleaves these heteroduplexes. Upon running the resultant fragments on a gel, densitometry can be used to estimate the proportion of edited vs. unedited amplicons [167].

1.3.2.2 Sanger sequencing-based approaches

Following PCR amplication of target loci, Sanger sequencing can be performed. Sanger-derived chromatogram traces can be analysed *in silico* for heterogeneity compared to an unedited control trace. This heterogeneity is taken as a measurement of the proportion of sequencing reads exhibiting alterations to the control sequence. Further analysis can identify the proportion of reads carrying specific sizes and orientations of indels, thereby identifying the highest-contributing edits in the pooled population. Two major methods utilize this approach: Incidence of CRISPR Analysis (ICE) and Tracking of Indels by Decomposition (TIDE). These methods have been verified as detecting editing efficiency at a similar level to next-generation sequencing, which allows for enhanced depth and breadth of sequencing [168-170].

1.3.2.3 Digital droplet PCR (ddPCR)

ddPCR relies on water-oil emulsion droplet formation to separate individual DNA template molecules into sequestered droplets for amplification [171]. Similarly to other

fluorescence-based PCR assays such as qPCR, fluorescent probes are designed to have complementarity to specific sequence features (edited or un-edited DNA). These probes bind to DNA during PCR extension steps and are subsequently released upon cleavage by polymerase. Therefore, upon amplification, templates carrying certain sequence features will yield fluorescent signal. Individual droplets are analysed for fluorescence, with the proportion of fluorescent droplets taken as a quantitative measure of the original template mix's composition. This can be used to establish what percentage of a pooled group of cells were successfully edited, either via precise directed repair (probe contains small edit or spans junction of desired indel) or NHEJ (presence or absence of WT sequence surrounding cut site) [172].

1.3.3 Minimising off-target effects (OTEs)

Off-target editing is a major concern with DNA editing. In theory, a 20-bp gRNA sequence should be unique in the genome. However, Cas9 is capable of tolerating mismatches provided a PAM site is present. Any formation of a DSB is a major threat to cell health, and formation of multiple DSBs in the same cell during CRISPR-Cas9 editing would increase risk of apoptosis [173]. Indels formed at off-target sites can also be disastrous if they interrupt function of essential genes or important regulatory sequences [174]. Even small-scale changes formed without DSBs (like the edits performed by technologies discussed in Section 1.3.4) can present a threat to non-target loci.

Some of the first studies investigating this phenomenon found several instances of off-target editing reaching comparable levels to on-target editing, with CRISPR machinery tolerating up to five mismatches in spacer sequence [175-179]. Since then, analysis of OTEs

has been an essential component of CRISPR research. Many tools have been developed for the prediction and detection of OTEs.

1.3.3.1 Prediction of vulnerable off-target loci

Many web-based tools exist for prediction of off-target edits. Most of these tools offer rankings of potential sites for OTEs that are based on a variety of criteria:

1) Mismatches between on- and off-target spacer sequences. The lower the number of mismatched nucleotides, the higher the likelihood of gRNA binding and Cas9 cleaving.

2) Location of mismatches in the seed sequence. Cas9 has been shown to be more tolerant of mismatches at the PAM-distal end of gRNA sequences, with the nucleotide at position 1 shown to be highly tolerant of mismatch and replaceable by a G nucleotide for efficient expression by a U6 promoter [180].

3) Predicted effect of OTEs. OTEs occurring at intergenic loci are unlikely to be pathogenic, even if they occur at high rates. On the other hand, even low-efficiency OTEs can prove threatening to cell health if they occur in coding regions or promoters of essential genes.

A recent study investigated the relative strength of three candidate *in silico* prediction methods in contrast with bench-based surveys (CHANGE-Seq, CIRCLE-Seq, DISCOVER-Seq, GUIDE-Seq, and SITE-Seq – discussed in detail below) [181]. The study found that, despite significant overlap in predicted and detected OTEs, no *in silico* or bench-based method found all 11 of the total OTE detected OTE events. However, not all available

methods were used for either prediction or detection of OTEs.

1.3.3.2 Detection of OTEs

The same tools used for detection of editing at on-target loci can also be used to detect OTEs by using them to investigate *in-silico*-predicted sites. This is sufficient for proof-of-concept or early *in vitro* work. However, as DNA editing strategies get closer to clinical application, a more thorough and unbiased approach is called for.

GUIDE-Seq (Genome-wide, Unbiased Identification of DSBs Enabled by Sequencing) relies on NHEJ-mediated integration of a double-stranded oligonucleotide donor (dsODN) after DSB formation [162]. dsODNs are introduced to cells alongside RNP and integrate into DSB sites. Genomic DNA is sheared and adapters ligated. Amplification is performed using primers specific to adapters and the dsODN, thereby selecting for genomic fragments that have undergone DSBs and subsequent dsODN integration. This strategy is unbiased, but relies on NHEJ-mediated integration of the dsODN into every DSB formed, which cannot be guaranteed. A later technique developed by the same group, CIRCLE-Seq (Circularisation for In vitro Reporting of Cleavage Effects by Sequencing), addresses this issue. CIRCLE-Seq relies on RNP treatment of pre-sheared, circularised genomic DNA [182]. Cas9 cleavage re-linearises DNA, converting it to a substrate for adapter ligation, amplification, and sequencing. This does not depend on integration events, which theoretically cannot be relied upon as a 1:1 readout of DSB formation (*ie*, not every DSB is guaranteed to be repaired via integration of the construct). A third technique from this group, CHANGE-Seq, also relies on Cas9 cleavage of circularized gDNA; however, the initial generation of circularized DNA relies on tagmentation, significantly reducing workload [183].

SITE-Seq (Selective enrichment and Identification of Tagged genomic DNA Ends by Sequencing) relies on biotin-mediated isolation of fragments carrying DSBs [184]. gDNA is treated with RNP, and exposed ends are biotinylated. gDNA is fragmented, and fragments conjugated with biotin are isolated. Isolated fragments are then amplified, indexed, and sequenced. Again, however, this technique requires alteration of cut DNA ends that is not guaranteed to occur at every DSB.

DISCOVER-Seq relies on chromatin immunoprecipitation followed by sequencing (ChIP-Seq) to detect loci where MRE11 has been recruited to DSBs. This technique has been further improved by inhibiting DNA-PKAcs. DNA-PKAcs are crucial for NHEJ; their inhibition leads to domination of HDR and MMEJ pathways, which rely on MRE11 and therefore enhance its retention at DSB sites [185]. BLISS-Seq allows for *in situ* detection of DSBs via fixation and permeabilization of edited cells, DSB blunting, and subsequent adapter ligation for *in vitro* transcsription and library prep [181]. High-throughput, genome-wide translocation sequencing (HTGTS) detects translocations occurring after DSB formation; however, translocation does not always occur as a result of DSBs [186].

1.3.4 Beyond "Conventional" CRISPR-Cas9

1.3.4.1 The need for non-DSB based approaches

Following the advent of the first iterations of CRISPR technology, major concerns arose about the danger that DSBs present to cellular health [173]. This is typically mediated by the p53 pathway, which is an essential agent in preventing tumour formation [187]. In various studies, suppression of the p53 pathway has been shown to enhance CRISPR-Cas9 editing efficiency [188, 189]. Therefore, researchers have theorised that DSB-based editing

approaches may inadvertently create selection pressure that favours cells with insufficient p53 pathways, leaving edited tissues at elevated risk for cancer development. In the last 5 years, researchers have devoted significant effort to the development of technologies derived from CRISPR but not relying on DSB formation.

1.3.4.2 Tandem paired nicking

By inducing a mutation at a critical residue of either of the nuclease domains of Cas9, a "nickase" enzyme (nCas9) can be generated that cuts only one DNA strand at a gRNA-specified site [190]. The tandem paired nicking strategy uses paired nCas9 constructs to cut the same strand of genomic DNA at adjacent sites. A donor plasmid also contains identical nicking sites surrounding the desired edit. Researchers have shown that this approach yields fewer indels, lower levels of off-target effects, and on-target editing efficiency comparable to Cas9 nuclease. Additionally, suppression of p53 did not affect editing efficiency (as it does in Cas9 nuclease), suggesting that nCas9 activity is less activating of p53.

nCas9 pairs have also been used *in trans* to induce DSBs with sticky ends [191]. This provides the opportunity to introduce donors with complementary sticky ends, thus eliminating the need for HDR and allowing for repair by simple base pairing (similarly to DNA resolution after treatment with restriction enzymes that generate complementary sticky ends). This strategy has been used to edit constructs into the AAVS1 site with higher editing efficiency than a conventional Cas9 nuclease/HDR approach.

1.3.4.3 Cytidine and adenine base editing

Cytidine base editing (CBE) also relies on nCas9 [192]. In this approach, nCas9 is fused with a deaminase capable of converting cytidine nucleotides into thymine. This

alteration is followed by treatment with a nickase to induce damage to the complementary DNA strand, prompting intracellular machinery to repair the damaged strand according to homology with the edited; this is enhanced by inhibition of uracil DNA glycosylase. Thusly, the desired change is permanently incorporated into the genome. Retention of the edit is also enhanced by inclusion of a uracil glycosylase inhibitor (UGI), which inhibits uracil excision.

Soon after the advent of CBE, adenine base editing (ABE) was reported [193]. This relied on the same basic principles of CBE; however, no known deaminase was capable of converting adenine residues to guanine. Via directed evolution, a deaminase protein was developed capable of creating the required change. This novel enzyme was paired with nCas9 in the same fashion as CBE.
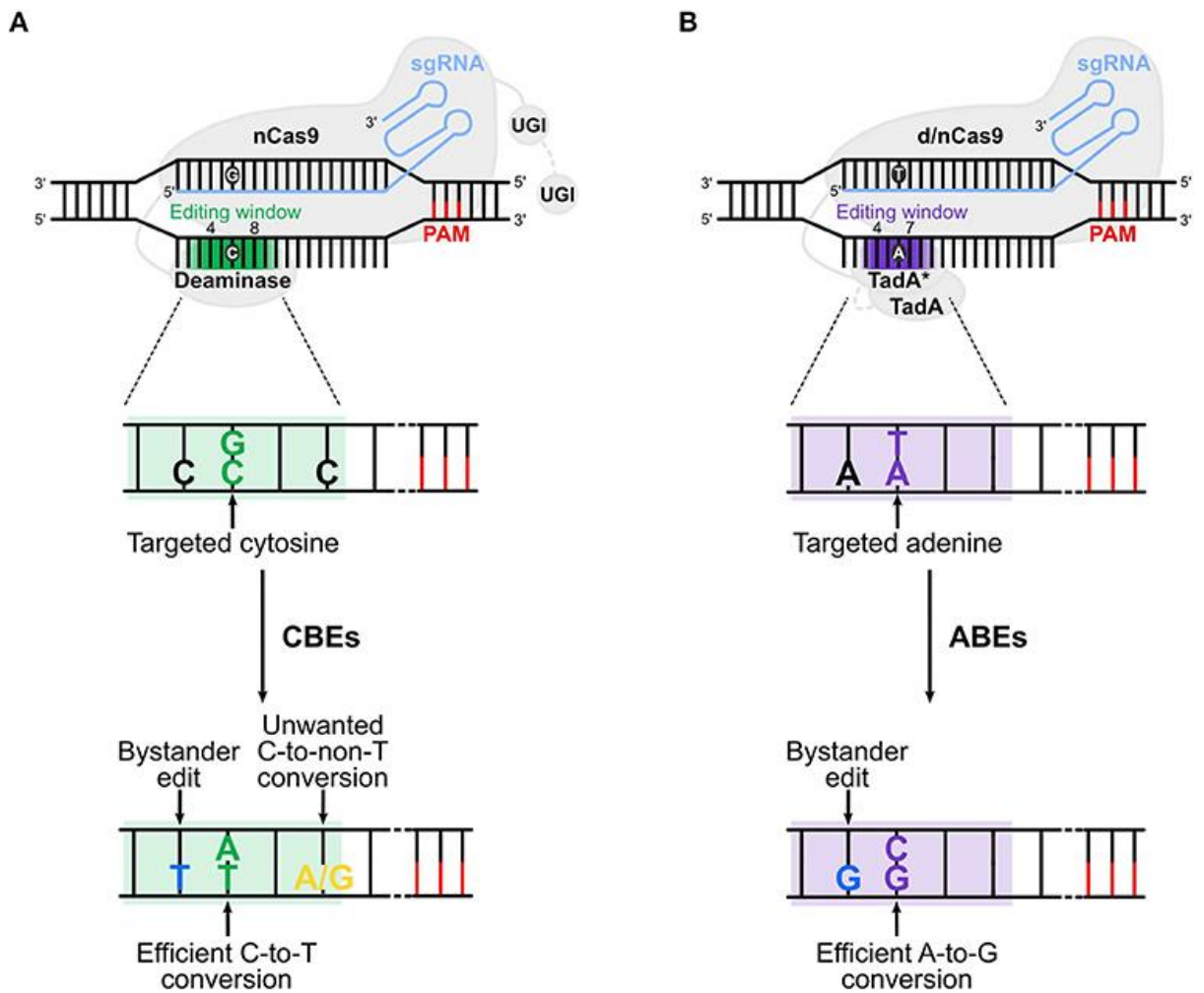
Fig. 1.8: Schematic of A) base cytidine and B) adenine base editors. nCas9

nicks the target strand according to homology with the sgRNA. Residues in

the editing window are deaminated. In CBEs, uracil glycosylase inhibition

(UGI) enhances permanent integration of the desired edit. From Antoniou, P.

et al. "Base and Prime Editing Technologies for Blood Disorders." *Frontiers in*

Genome Editing, vol. 3, Jan. 2021, p. 618406. *DOI.org (Crossref),*

https://doi.org/10.3389/fgeed.2021.618406 [194].

Both base editors have been shown to achieve high levels of editing efficiency (albeit with some variation in outcomes for different target loci) and minimal off-target effects (OTEs) [192, 193]. They have also undergone optimisation since their initial introduction in order to further enhance editing efficiency [174, 195-198]. However, they are not without disadvantages. Firstly, they can generate only single nucleotide modifications. Secondly, there have been reports of BE machinery causing unwanted edits on mRNA [199, 200], though this has been alleviated by further optimisation and engineering of BE machinery [201, 202]. Thirdly, there is the potential for creation of "bystander" mutations, in which a non-target cytosine or adenine residue in the editing window is deaminated, generating an unwanted point mutation [203]; techniques to eliminate this often rely on restriction of the editing window by introducing amino acid substitutions at critical residues, which also restricts flexibility based on PAM availability [204].

1.3.4.4 Prime editing

Prime editing (PE) offers significantly more flexibility in editing outcomes than ABEs or CBEs. Rather than relying on enzymes to reconstruct individual nitrogenous bases at the submolecular level, this technology uses a reverse transcriptase (RT) to reconstruct nicked and resected DNA using a customised RNA-based primer and repair template [205].

As in base editing, nCas9 is used to target the desired locus, but instead of a normal gRNA, it carries a prime editing gRNA (pegRNA). pegRNAs contain a protospacer sequence, a scaffolding sequence, a primer binding sequence (PBS), and a repair template. As with conventional CRISPR and base editing, the protospacer sequence has homology to the target locus, and directs the nickase to the correct site for strand cutting. Upon cutting of the target DNA and exposure of a strand end, the primer binding sequence of the pegRNA

anneals to the genomic DNA and acts as a primer for the RT enzyme fused to the nCas9 (fused nCas9 and RT are commonly referred to as "PE2"). The RT extends the exposed end of the DNA according to the template incorporated in the pegRNA; this template includes the desired edit. This extension generates a 3' "flap" containing the edit. The 3' flap equilibrates with the residual 5' flap, which does not contain the edit. Upon exposure of the 5' flap, endogenous intracellular machinery cleaves the flap, a process that naturally occurs during lagging-strand DNA synthesis. Once the competing un-edited flap is removed, the edited strand re-ligates [205].

Just as in base editing, the product is double-stranded DNA carrying an edit on only one strand. Incorporation of the edit into the unedited strand is enhanced as in base editing, by use of a secondary nickase targeting the non-edited strand (an approach referred to as "PE3"), and subsequent endogenous DNA repair to match the previously edited strand, thereby permanently incorporating the edit [205].

Prime Editor (PE)

Cas9 nickase

Reverse Transcriptase

Prime Editing Guide RNA (pegRNA)

+

5'

gRNA Target Sequence

3'

Primer Binding Sequence (PBS)

Template Containing Edited RNA Sequence (red)
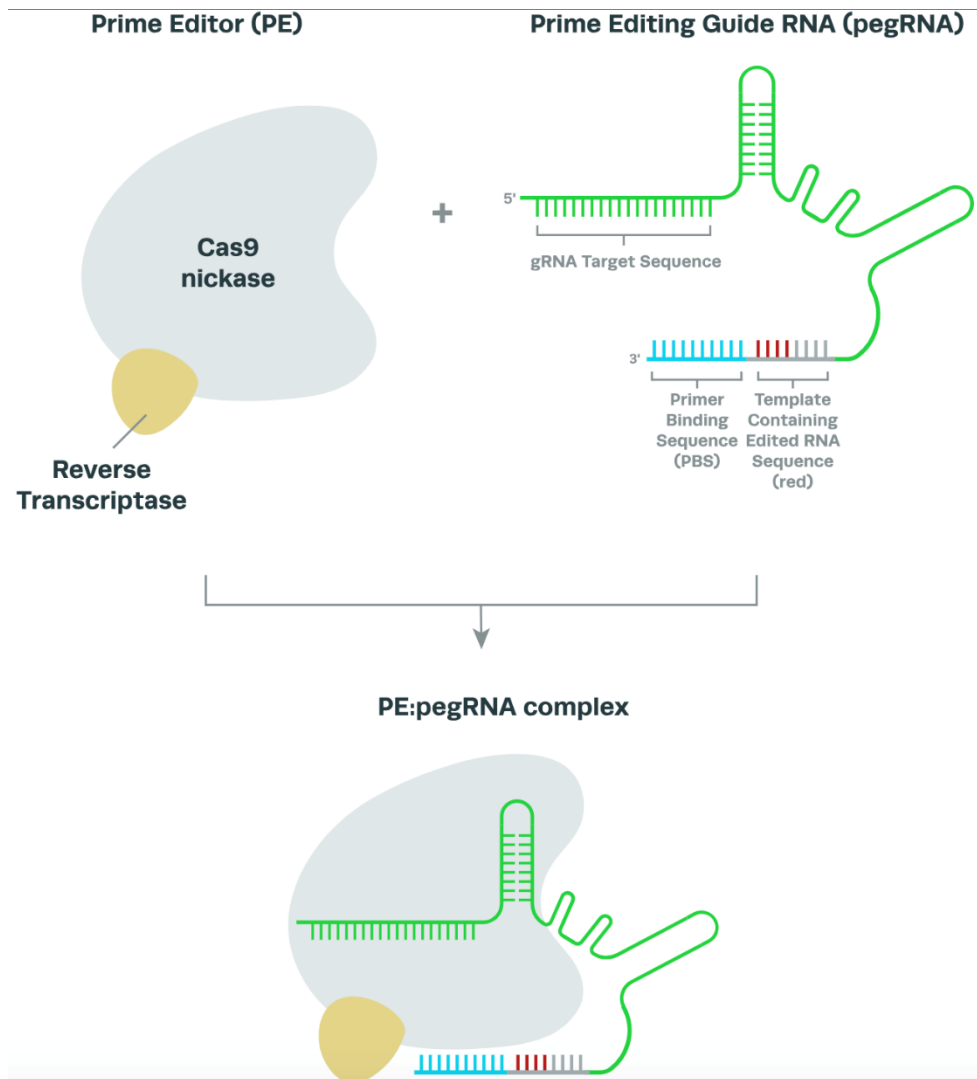
PE:pegRNA complex

1. PE:pegRNA Complex Binds to Target DNA
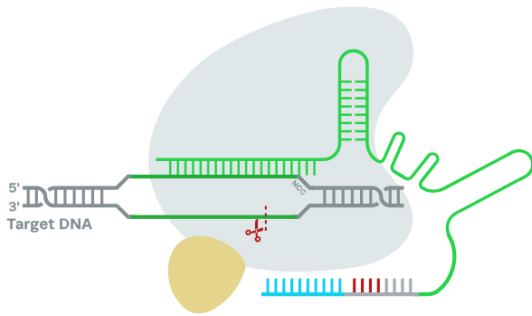
2. Cas9 Nickase Cuts One Strand of DNA

5'
3'
Target DNA

Cut DNA
on one strand

5' 3'
3' 5'
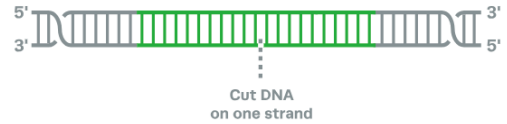
3. Reverse Transcription of RNA Incorporates Desired Sequence into Target DNA
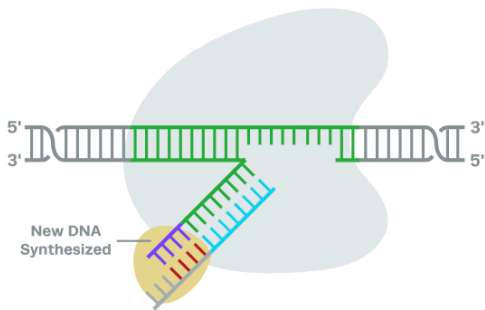
4. Edited Strand is Incorporated. Original DNA is Cleaved by Cellular Endonuclease

5' 3'
3' 5'

New DNA
Synthesized

Unedited
Strand

5' 3'
3' 5'
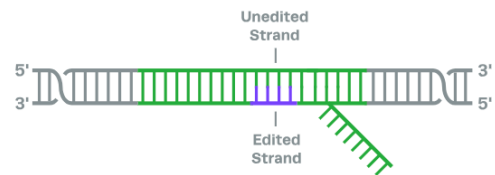
Edited
Strand

5. Unedited Strand is Repaired to Match Newly Edited Sequence*

5' 3'
3' 5'

Both Strands
Contain the
Desired Edit

Fig. 1.9: Schematic of the components and process of prime editing. Cas9 nickase (grey shape) nicks target DNA according to homology with spacer sequence in pegRNA. After nicking, the primer binding sequence of the pegRNA acts as a primer for reverse transcriptase, which extends DNA according to the template section of the pegRNA. This generates a "flap," which is excised by cellular machinery, and re-ligates. From Synthego (https://www.synthego.com/guide/crispr-methods/prime-editing).

Prime editing has been shown to achieve high levels of editing efficiency and low levels of OTEs [205]. Additionally, alterations to pegRNA design (called engineered pegRNAs or epegRNAs) can yield even higher editing efficiencies by conferring protection from intracellular degradation [206].

1.3.4.5 Use of CRISPR to induce changes in gene expression

CRISPR's ability to precisely locate and bind to unique genomic loci has been co-opted to activate and inhibit expression. Both CRISPR-mediated inhibition (CRISPRi) and activation (CRISPRa) utilise a catalytically dead Cas9 (dCas9) protein guided by gRNA to transcriptional start sites of target genes.

In CRISPRi, dCas9 can either be applied on its own to sterically hinder binding of RNA polymerases at gRNA-specified sites [207], or be fused to a ZFN-derived Kruppel-associated box (KRAB) domain, which promotes formation of heterochromatin [208]. Further enhancement of the CRISPRi system has been achieved by fusion of methyl-CpG binding protein 2 (MeCP2) to the KRAB domain [207], and by inclusion of the KRAB domain ZIM3, which is a more potent inhibitor of expression [209]. CRISPRi has been used in high-throughput screening for gene essentiality in bacteria and eukaryotes [210], elucidating functions of human pseudogenes [211], and identifying long non-coding RNAs (lncRNAs) in glioblastoma cells for therapeutic targeting [212]. Potential applications for direct clinical benefit include multiple eye diseases, enhanced bone healing, lung cell squamous carcinoma, and high cholesterol (reviewed by Peddle *et al.* 2020) [213].

CRISPR-mediated activation (CRISPRa) can be achieved by multiple approaches. The first iterations relied on fusing dCas9 with the ZFN-derived VP64 acidic transactivation domain, which led to substantial upregulation of target genes [214]. Since then, CRISPRa has been enhanced by combining dCas9 with multiple activators fused together [215]. Researchers have also enhanced CRISPRa by using modified sgRNAs to generate scaffolding gRNAs that are capable of bearing aptamers that recruit specific transcriptional effectors [216]. Translational applications of CRISPRa have included muscular dystrophy, retinitis pigmentosa, and diabetes mellitus (reviewed by Becirovic 2022) [217].

Another method of modifying gene expression is by targeting mRNA directly [218]. A

Cas13 protein (*Psp*Cas13b) was found to effectively target and silence target transcripts with

comparable efficiency to a shRNA, but with fewer off-target effects. Furthermore, fusion of

the Cas13 with a deaminase enzyme allowed for specific sequence editing of mRNA,

allowing for rescue of disease alleles without editing native DNA.

1.3.5 Pros and cons of commonly used DNA editors

A brief comparison of DNA editors discussed is shown in Table 1.1.

Editor	Description	Pros	Cons	Examples of Clinical Uses
Meganucleases	Sequence-specific nuclease enzymes recognising ~20 bp	Highly specific	DSB formation, require custom engineering for unique targeting	CAR-T cell generation
TALEN	Can be built into an array to target specific sequences, and paired with nuclease	Highly specific	DSB formation, require custom engineering for unique targeting	CAR-T cell generation
ZFN	Can be built into an array to target specific sequences, and paired with nuclease	Highly specific	DSB formation, require custom engineering for unique targeting	Hemophilia B Mucopolysaccharidosis
CRISPR-Cas system	Cas nuclease targeted by guide RNA	Customization of targeting requires alterations to	DSB formation, potential for indels at on-	Multiple blood disorders Multiple eye disorders

		small RNA sequence, not protein itself	and off-target loci	CAR-T cell generation Antibacterial treatments
Base editors	Cas9 nickase paired to deaminase for alterations of individual nucleotides	No DSB formation, reduced off-target editing compared to original CRISPR-Cas9 system	Can make only small-scale edits, bystander mutations common within editing window	Sickle cell disease Beta-thalassemia
Prime editors	Cas9 nickase paired to reverse transcriptase; gRNA extended to include short repair template	No DSB formation, reduced off-target editing compared to original CRISPR-Cas9 system	Can make only small-scale edits	None at present

Table 1.1: Comparison of DNA editing technologies. All clinical trials listed in detail at

https://crisprmedicinenews.com/clinical-trials under the respective technologies

used [219].

1.4 Use of CRISPR technologies to treat CF

1.4.1 Generation and Correction of Individual Variants

1.4.1.1 Cas9/HDR: Editing Efficiencies and Delivery Strategies

Some of the first experiments using CRISPR-Cas9 to treat CF relied on HDR to correct

individual variants, most commonly F508del.

In one such study, fibroblast cells from a CF F508del-homozygous patient were

expanded and reprogrammed into induced pluripotential stem cells (iPSCs) [220]. iPSCs

were nucleofected with codon-optimised Cas9 encoded in a plasmid under the Eukaryotic Translation Elongation Factor 1 Alpha 1 (EEF1A1) promoter, while gRNAs targeting the F508del variant region were encoded in a plasmid under the U6 promoter [220]. Donor vectors included template for correction of the F508del variant and puromycin selection marker [220]. After puromycin selection, 6 of 36 clones were found to carry the F508del correction in at least one allele, giving an editing efficiency of 16.7%. Of note, the puromycin resistance cassette in the donor oligo was flanked by piggyBac transposase sites, allowing researchers to scarlessly excise the selection marker after successful editing.

HDR of Cas9-induced DNA damage can also be achieved using a single-stranded oligodeoxynucleotide (ssODN) containing homology arms. In a recent study, pairing Cas9 RNP with ssODN repair enabled researchers to both generate and repair common CF variants in human iPSCs [221]. In order to induce *CFTR* variants in wild-type (WT) cells, Cas9 RNP incorporating gRNAs targeting the native amino acids F508, G542, or G551 were electroporated into WT iPSCs. Simultaneously to this RNP construct, ssODN repair templates were electroporated. To induce the F508del variant in WT iPSCs, sense or antisense ssODN templates were compared, both of which carried near-symmetrical homology arms of 50 and 52 bases. The sense template yielded an editing efficiency of 5%, while the antisense template yielded editing efficiency of <1% [221]. To induce the G551D variant in WT iPSCs, ssODN sense templates with asymmetrical homology arms of 118 and 82 bases edited 19% of alleles [221]; to induce the G542X variant in WT iPSCs, ssODN sense template with symmetrical homology arms of 80 bp edited 27% of alleles [221]. Overall, in this study, repair efficiency was inconsistent, and did not strongly correlate with symmetry, length, or sense vs. antisense homology. However, controlled direct comparison of such design factors were not explored for the same alleles. Notably, studies in optimisation of ssODN design

have suggested that both length and degree of symmetry in ssODN homology arms have significant impacts on editing efficiency; however, different alleles yield different results. For example, the study summarised above reported that optimal ssODN orientation differed depending on allele targeted, while another reported that for single-nucleotide edits, antisense orientation yielded superior editing rates [142].

Recombinant adeno-associated virus (rAAV) vectors have been used to deliver repair templates for HDR in human cells [222, 223]. Using Cas9 RNP, researchers targeted the F508del variant in upper airway basal stem cells (UABCs) and human bronchial epithelial cells (HBEs) from CF patients either homozygous or compound heterozygous for the F508del variant [223]. Cells were collected, expanded, electroporated with the RNP, and then immediately transduced with an AAV6 vector carrying a repair template. Template included the F508 correction alongside six silent mutations spanning the PAM and spacer sequence; inducing silent mutations would preserve coding sequence, but prevent Cas9 from re-targeting successfully edited loci [223]. Using this strategy, variant correction was observed in 28 ± 5% of F508del-homozygous UABCs, 42 ± 15% of compound heterozygous UABCs, and 41 ± 4% of homozygous HBEs [223]. Corrected cells were then expanded and seeded in ALI culture. ALI cultures were analysed for electrophysiological properties using an Ussing chamber. Corrected UABCs displayed 31 ± 5% current of non-CF controls, while corrected HBEs displayed 51 ± 3% current of non-CF controls [223]. Previously, it has been estimated that a 10% increase in CFTR ion channel activity may be clinically beneficial to CF patients [224].

HDR was also used to correct the G551D variant in primary airway basal cells obtained from a ferret model [225]. Basal cells were edited to stably express *Sp*Cas9 via lentiviral transduction. A single rAAV6 vector was then used to deliver gRNA and donor

template. Analysis of cloned DNA from edited cells revealed 13.95% HDR-mediated correction and 23.26% indels. In order to enrich for successfully edited cells, a reporter cell line was generated via lentiviral integration of a GFP construct. The GFP contained a point mutation that eliminated fluorescence; editing of the mutation would restore fluorescence, allowing for FACS-based enrichment. After enrichment, a pool of cells containing 18.4% G551D correction was differentiated on ALI. Ussing analysis revealed 70.4% rescue of CFTR activity. However, lentiviral introduction of either Cas9 or GFP are unfeasible in a translational setting.

In another study, intestinal organoids generated from two CF patients, each homozygous for the F508del variant [226], were transfected with plasmids encoding both Cas9 and gRNAs targeting the F508 locus, as well as a puromycin resistance gene to enable selection of corrected cells [226]. The study showed that 18% of puromycin-resistant transfected organoids showed successful editing of the F508del variant, although transfection efficiency was only 1.5%.

1.4.1.2 Adenine base editing

The ability to edit adenine to guanine represents an opportunity to transform stop codons (UGA, UAG, UAA) into codons encoding tryptophan (UGG). This is potentially invaluable for CF, as the current most promising CF pharmaceutical intervention is unable to rescue defects caused by premature stop codons, but is effective in rescuing most missense mutations (which a STOP>Trp would most likely represent) [52]. Additionally, by targeting the noncoding strand, a broader range of changes can be made.

A biobank in the Netherlands contains intestinal organoids of genotypes representing nearly 50% of the Dutch CF population. 21% of these variants were found by

bioinformatic analysis to be potentially rescuable using ABEs [227]. Several such alleles were targeted for treatment by ABEs including the R785X variant; this yielded an editing efficiency five-fold higher than that of Cas9-mediated DSB and HDR [227]. Subsequently, two more variants (W1282X and R553X) were corrected in clonal cell lines using ABEs with an efficiency of 1.43% for both variants [227]. In this study, editing efficiency was quantified by phenotypic screening (namely the forskolin-induced (FIS) assay); it is possible that more cells received the desired edit, but also acquired deleterious off-target mutations that prevented FIS response or organoid formation in general. Analysis of off-target effects revealed no significant alterations at predicted off-target sites, nor were increases observed in the occurrence of A>G changes across the entire genome [227].

The editing efficiency of ABEs has recently been increased by incorporating chemical modifications into the ABE mRNA and guide RNAs. Uridine-depleted (UD) mRNA encoding the codon-optimised ABE RA6.3 was created by silently altering codons in order to minimize the number of uridines in the transcript, while remaining uridines were replaced with 5-methoxyuridine to create 5-methoxyuridine-modified (5moU) mRNA, which is a more stable template for translation [228]. 5moU ABE mRNA yielded unexpectedly low A>G editing rates when targeted to a genomic site known to have high editing rates upon delivery of RA6.3 encoded by plasmid DNA (pDNA) [228]. However, combining the 5moU ABE mRNA electroporated with 2'O-methyl 3'-phosphorothioate modified sgRNA yielded editing rates that exceeded plasmid encoded ABE RA6.3 and sgRNA at three separate genomic sites [228]. This approach was then used to target the *CFTR* W1282X variant, which is caused by a G>A mutation. Two A residues lay within the editing window, the editing target (A9G) and a non-target bystander (A5G). Modified sgRNA was electroporated alongside either unmodified or 5moU-modified ABE mRNA. Bystander editing rates at A5G was significantly

higher with the 5moU ABE mRNA [228], although editing efficiency was higher at A9G than at A5G with both 5moU and unmodified mRNA. Base editing at A5G modifies the glutamine residue encoded by the codon immediately preceding the target codon (Q1281) to an arginine, and so the phenotypic effect of this Q1281R substitution was assessed in clones. Clones containing the corrected nonsense mutation with or without the Q1281R mutation exhibited normal CFTR channel activity, suggesting that ABE editing restored CFTR function and that the presence of the Q1281R mutation in the same allele did not hinder CFTR function [228].

Another study used ABEs to correct two nonsense (R553X and W1282X) and one splicing (3849+10kbC>T) variant [229]. These variants produce truncated or absent protein product and are therefore not rescuable by existing modulator therapies. In this study, ABE enzyme machinery and gRNAs were delivered via preformulated RNP. RNP was electroporated into ALI-differentiated primary CF airway epithelial cells compound heterozygous for R553X/G85E. One week post treatment, cells were genotyped via NGS of target site PCR amplicons. Initial tests yielded editing efficiency of 54.5% with negligible bystander edits or indels. Next, editing was attempted in primary cells compound heterozygous for W1282X/ CFTR-del2,3 (a 21 kb deletion resulting in the loss of exons 2 and 3). In this case, a bystander adenine was present in the editing window. Treatment yielded an editing efficiency of 38% at the target adenine. However, this was accompanied by 16.6% editing of the bystander residue. This bystander change creates the mutation R1283G, which is a significant alteration to amino acid composition in terms of both size and charge. Finally, editing was attempted to remedy the 3849+10kbC>T mutation. While this mutation cannot directly be targeted for deamination, the cryptic splice site that it generates can be removed by deamination of a neighbouring adenine. Using this approach, ABE editing

efficiency was 29.1%, with <2% bystander edits. For all editing experiments, no genomic

OTEs were found upon sequencing of predicted off-target loci. Upon Ussing analysis of

mixed edited populations, it was revealed that higher editing efficiency did not always relate

linearly to functional rescue and was instead mutation dependent [229].

A more recent study targeted the 2789+5G>A mutation [83]. In cells with this

mutation, aberrant splicing results in an out-of-frame truncated protein product. Firstly,

minigene models were generated for the mutation site and the surrounding exons, with the

mutation either present or absent in the construct. Aberrant splicing was observed in

HEK293T cells transfected with the mutant minigenes. In order to repair the mutation,

multiple ABE strategies were tested, with varying optimised ABE proteins and gRNA

sequences delivered in plasmid form. Additionally, a variant of *Sp*Cas9 was used with

relaxed PAM specificity, allowing recognition of NG rather than NGG. The most efficient

gRNA-ABE combination yielded ~40% editing of the minigene construct, as detected by

sequencing and Sanger trace analysis via EditR software. Efficiency was further increased to

up to ~70% using all-in-one delivery of ABE machinery via lentiviral transduction; however,

bystander editing was also detected. In intestinal organoids, the same lentiviral approach

yielded ~10% editing of the target adenine; however, bystander effects were once again

observed. Splicing defects were modestly rescued, with correct splicing increasing from

~65% pre-treatment to ~75% post-treatment. In rectal organoids, ABE enzymes and gRNAs

were delivered via electroporation of mRNA and sgRNA respectively. ~5% editing efficiency

was observed, with <1% editing of a bystander adenine. Restoration of correct splicing was

comparable to that achieved by the lentiviral delivery approach. Two weeks post treatment,

edited organoids were responsive in FIS assays. Finally, in primary bronchial epithelial cells,

electroporation with the same RNA constructs yielded comparable editing efficiency of the

target site; however, nearly four-fold editing was observed at a bystander adenine. Successfully edited cells were expanded into ALI cultures, which were found to achieve 13-18% of WT CFTR activity via Ussing chamber analysis [83].

Overall, ABEs have shown promising editing efficiency depending on the model in which editing is attempted (organoids, cell lines, primary cells) and the method of delivery (virus, plasmid, RNA, RNP). However, bystander edits can significantly outweigh on-target editing, meaning that ABEs may yet only be appropriate for mutations with no critical neighbouring adenines.

1.4.1.3 Prime editing

Geurts *et al.* attempted prime editing in intestinal organoids homozygous for the F508del variant [81]. Organoids were electroporated with PE2 and pegRNA/PE3 guide-expressing plasmids. Two weeks post-treatment, an FIS assay was used to assess functional response. One organoid showed the desired response and was isolated for further analysis. Sanger sequencing of the F508del locus revealed heterozygous repair of the mutation. In order to improve PE efficiency, the experiment was repeated with various RT and PBS lengths, as well as varying distances from cut site to desired edit. Eight distinct editing strategies were compared with conventional CRISPR with two HDR tested donors. CRISPR/HDR resulted in over 100 FIS-responsive clones with either donor, whereas prime editing never resulted in more than four clones. However, in all edited clones, FIS response was restored to WT levels. Sanger sequencing of one HDR clone and two PE clones revealed heterozygous variant correction; however, the HDR clone and one PE clone also contained indels at the uncorrected allele [81].

PE was then compared with BE of the R785X variant. ABE outperformed both CRISPR/HDR and PE at this locus, yielding 9.1% corrected clones versus 1.2% with CRISPR/HDR and 0-5.7% with PE (variation observed between pegRNA/PE3 designs) as measured by FIS response. Three repaired clones from each editing strategy were genotyped via Sanger sequencing to reveal heterozygous variant correction. In two of three clones from PE, and three of three from ABE, the uncorrected allele was undamaged. The third PE clone, as well as the HDR-repaired clones, contained indels at the uncorrected allele. However, no significant genome wide OTEs were detected via WGS for any prime edited clonal populations [81].

1.4.1.4 NHEJ

Despite its imprecision, some CF mutations may be corrected by NHEJ strategies. One such candidate is the deep intronic variant c.3718+2477C>T, which creates a cryptic splice acceptor that results in the inclusion of a pseudoexon which includes a premature stop codon [230]. A minigene plasmid was generated containing a contiguous segment of the *CFTR* gene, including the relevant intron; this sequence was either c.3718+2477C>T variant or WT. A second plasmid was generated encoding Cas9 and a pair of gRNAs, which targeted sites immediately upstream and downstream of the c.3718+2477C>T variant in order to excise it. Plasmids were co-transfected into HEK293T cells, and minigene splicing was subsequently analysed. As measured by PCR and agarose gel densitometry, the variant was excised in 61% of transfected cells. RT-PCR was performed on RNA isolated from transfected cells; electropherogram analysis revealed that correct splicing had been restored at 83% efficiency. While these results provide promising proof of concept, cleavage of a minigene in HEK293T cells may not effectively predict chromosomal editing in airway

epithelial cells. A more robust study design would involve characterising editing efficiency in the native genome of the cell type of interest. Another splicing variant, c.3849+10kbC>T, has been corrected in airway epithelial cells' endogenous genomic DNA using a similar dual cut approach [231]. RNPs were transfected with two guides that targeted just up and downstream of the variant, and after four sequential transfections of RNP (48 hours between transfections), 62% of Sanger sequencing reads were found to harbour the desired deletion by ICE analysis. This translated to 41% rescue of WT splicing (from 25% in unedited variant cells to 66% in edited).

1.4.2 cDNA integration strategies

There are over 200 known *CFTR* variants associated with CF, as reported by the CF variant database CFTR2 [232] . Due to the stringent trialling and approval process required of novel treatments, it is not feasible to generate and test DNA editing strategies for each individual CF-causing variant. Therefore, a universal treatment approach would be much more valuable.

In a 2019 study, a full-length CFTR expression cassette was delivered to *CFTR* knockout pig airway basal cells using a helper-dependent adenovirus (HD-Ad) system [233]. An HD-Ad vector was designed including all required components of the editing reaction: a gRNA targeting the porcine *GGTA1* locus (a known safe harbour for genomic integration), Cas9, and an 8.7 kB CFTR expression cassette with poly-A tail. Successful integration, as measured by ddPCR, was found to occur most at the highest MOI tested (MOI 50, 10% efficiency). RT-PCR was performed on transduced cells at passage 0 (immediately after transduction), 10, and 20. Interestingly, the trend of CFTR mRNA expression was inversed from the editing efficiency, with the highest mRNA expression level achieved with the lowest MOI tested

(MOI 10, 7.94% of P0 expression at P10, 2.53% of P0 expression at P20). Western blotting was not sufficient enough to detect CFTR protein product at any MOI at or beyond P10. Despite this, functional analysis by membrane potential-sensitive dye-based FLIPR assay revealed significant differences between transduced and untransduced cells' responses to CFTRinh-172 and forskolin [233]. Additionally, HD-Ad vectors have previously been shown to successfully deliver transgenes to pig airway basal cells *in vivo*, making them a promising method for delivery [234].

Another study utilised two AAV vectors to deliver each half of a full-length *CFTR* cDNA to UABCs and HBEs carrying *CFTR* variants [235]. The first AAV vector contained a homology arm, the first half of the *CFTR* cDNA, and a randomised 400 bp "stuffer" sequence with no homology to any target in the human genome. The second AAV vector contained a sequence homologous to the stuffer, the second half of the *CFTR* cDNA, a polyA tail, and tCD19 for enrichment of edited cells. Primary UABCs were electroporated with RNP, and incubation with both donor AAVs immediately followed. In CF cells homozygous or compound heterozygous for F508del, treatment led to 4% editing efficiency pre-enrichment as measured by tCD19-mediated FACS. In seven populations of HBEs carrying a variety of genotypes (including undruggable nonsense variants), 10% editing efficiency was achieved pre-enrichment. Enriched populations of edited UABCs and HBEs were differentiated on ALI. Western blot analysis of ALI lysates revealed an increase in mature CFTR production. In an Ussing chamber, edited cells exhibited rescue of CFTRinh-172 response to 72% (UABCs) or 88% (HBEs) of non-CF levels. OTEs were examined at 47 *in silico* predicted loci. Wild type *Sp*Cas9 (WT-Cas9) was found to exhibit significant OTEs at one locus; however, use of a HiFi Cas9 (carrying a single point mutation thought to inhibit Cas9 transitioning from inactive to

active state when mismatches are present) significantly redu ced these OTEs from >50% to

<2% of alleles tested (both Cas9 variants obtained from Integrated DNA Technologies) [235].

Study	Cell Type Targeted	Cas9/ gRNA Delivery	Repair Template Delivery	Selection Marker	Indel Formation	Editing Efficiency
Firth, A.L., et al. 2015	iPSCs	Plasmid	dsDNA	Puromycin resistance	NR	16.7% repair of F508del-homozygous iPSCs (post-selection)
Sanz, D.J., et al. 2017	HEK293T	Plasmid	N/A	N/A	34–47% at c.3718-2477C locus 30% at c.3140-26A locus 28–33% at c.1679+1634A locus	N/A
Zhou, Z.,et al. 2019	Intestinal porcine enterocyte cells (IPECs)	HD-Ad	HD-Ad	N/A	45-66% (dependent on MOI)	3-10% (dependent on MOI)
Ruan, J., et al. 2019	iPSCs	RNP	ssODN	N/A	30% at F508 locus 51% at G542 locus 48% at G551 locus	12% repair of F508del-homozygous iPSCs
Vaidyana-than, S., et al. 2019	Upper airway basal stem cells (UABCs) and human bronchial epithelial cells (HBECs)	RNP	AAV6	N/A	80% at F508 locus	28 ± 5% repair of F508del-homozygous UABCs 42 ± 15% repair of compound heterozygous UABCs 41 ± 4% repair of F508del-homozygous HBECs
Vaidyana-than, S., et al. 2022	UABCs and HBECs	RNP	AAV6	tCD19 enrichment tag	NR	57-89% (post-enrichment, dependent on variant and cell type)
Yan, Z. et al. 2022	Ferret airway basal cells (FABCs)	Cells engineered to express Cas9	AAV2.6	GFP integration for enrichment	NR	18.4% (post-enrichment) correction of G551D
Walker, A., et al. 2023	HBECs	RNP	N/A	N/A	86% after 4 rounds of transfection	62% desired deletion of 3849+10kb C>T 14% undesired indels (ie only one sgRNA cut, leaving indels upstream or downstream of variant)

Table 1.2: Editing efficiencies obtained by relevant studies on using CRISPR-Cas9 to treat CF. RNP: ribonucleoprotein. N/A: not applicable. GFP: green fluorescent protein. NR: not reported.

1.5 Non-Cas9 Approaches to Gene Editing of CFTR

1.5.1 Transcription-Activator-Like Effector Nucleases (TALENs)

Additionally, non-CRISPR-Cas9 gene editing therapies have recently been explored for CF. TALEN-mediated homologous recombination was recently used to correct the F508del variant in iPSCs, which were then expanded into intestinal organoids [82]. A TALEN plasmid was nucleofected alongside a donor vector carrying the desired variant correction as well as a puromycin resistance selection cassette flanked by piggyBac transposon sites; after correction and selection, the selection cassette was scarlessly removed via the piggyBac system [82]. Corrected iPSCs showed CFTR protein abundance and localisation comparable to non-CF cells; corrected and uncorrected iPSCs were then expanded into intestinal organoids [82]. CFTR function was assayed using FIS assay; upon exposure to forskolin, corrected organoids increased their surface area to 177%, while uncorrected organoids increased to 103%, with 100% representing the pretreatment surface area [82]. Additionally, corrected organoids were found to be responsive to the clinically approved double-combination drug VX-770/VX-809 (marketed as Orkambi), suggesting that DNA editing-based therapies can be combined with pharmaceutical intervention to enhance clinical outcome [82].

1.5.2 AsCas12a/Cpf1

The Cas protein AsCas12a (also referred to as Cpf1), derived from *Acidaminococcus*, recognises the PAM sequence TTTV. This nuclease has recently been used *in lieu* of Cas9 in order to excise intronic splice altering *CFTR* mutations. Due to AsCas12a's tendency to create deletions >4bp, a single gRNA rather than a pair was used to target the desired mutation; this is potentially advantageous because fewer DSBs leads to fewer opportunities for apoptosis induction and off-target indel formation. The variants 3272-26A>G and 3849+10kbC>T *CFTR* were targeted using Cas12a in first a minigene model and then in primary airway cells and intestinal organoids. In primary airway cells derived from a 3272-26A>G/F508del compound heterozygote, lentiviral delivery of Cas12a with a single guide yielded 30% cutting efficiency and 13-fold rescue of aberrant splicing without puromycin selection [236]. In human intestinal organoids derived from a 3272-26A>G/4218insT compound heterozygote, lentiviral delivery of Cas12a with a single gRNA resulted in restoration of proper splicing from 69% to 95% [236]. FIS response in edited organoids increased 2.8-fold compared to unedited organoids, similar to the result obtained by rescue using lentiviral introduction of non-variant *CFTR* cDNA [236]. These protocols were repeated with a guide targeting the 3849+10kbC>T allele; in compound heterozygous 3849+10kbC>T/F508del primary airway cells, a 20% indel rate was achieved; in intestinal organoids with the same genotype, a 30% indel rate was achieved, and FIS swelling was restored to the same level as that achieved by lentiviral introduction of non-mutant *CFTR* cDNA [236]. A pair of guide RNAs accompanied by Cas9, also introduced via lentivirus into Caco-2 epithelial cells, yielded significantly lower rates of FIS responsiveness than both Cas12a-corrected organoids and WT *CFTR* cDNA-expressing organoids [236]. However, long-term expression of Cas9 from lentiviral integration raises safety concerns as it has been

shown to exacerbate off-target DSBs throughout the genome [237], and so comparison with more clinically viable delivery strategies is necessary in the future.

Cas12a was directly compared with Cas9 in a study aiming to correct the W1282X variant using ssODN-mediated HDR [238]. Different gRNAs were used for each enzyme due to their differing PAM requirements; however, in both cases, the closest possible cut site to the mutation was targeted. Using guidance from previous studies, asymmetrical ssODNs were designed: 5'-91nt/36nt-3' complementary to non-target strand for Cas9; 5'-77nt/37nt-3' complementary to target strand for Cas12a. The PAM site in the Cas12a donor was modified with a silent mutation to prevent retargeting by the nuclease; in the Cas9 donor, the W1282X modification was immediately 5' of the cut site, and therefore represents an alteration to the spacer sequence, preventing retargeting. Cas9/Cas12a RNPs were electroporated, with or without ssODNs, into a 16HBEge cell line engineered to carry the W1282X variant [57]. 72 hours post treatment, genomic DNA was isolated from cells, PCR performed to amplify the cut site, and Sanger sequencing executed. Inference of CRISPR Edit (ICE) software was used to quantify HDR and indels based on heterogeneity of Sanger sequence chromatogram traces [238]. Cas9 RNP was found to generate significantly higher editing efficiency and a higher ratio of repair:indels than Cas12a. In the case of Cas12a HDR, three times the amount of ssODN was found to be required to achieve half the editing efficiency; this was also accompanied by 1.75-fold higher indel frequency than Cas9/HDR treatment. At the mRNA level, up to 60-70% of correct transcripts were observed after Cas9 treatment, whereas only 50% restoration was observed after Cas12a treatment. Two clonal homozygous Cas9/HDR-repaired populations were expanded and analysed; in these populations, *CFTR* mRNA was restored to WT levels. An HS-YFP quenching assay was

performed on these clonal populations, and quenching was observed at non-CF levels; in the mixed edited pool, 18% of non-CF quenching levels was observed.

1.5.3 Zinc Finger Nucleases (ZFNs)

Integration of a *CFTR* superexon encoding exons 11-27 has also been studied using zinc finger nucleases (ZFNs) [239]. The plasmids (encoding ZFNs or containing the superexon donor construct) were transfected into an F508del-homozygous cell line (CFBE41o-) [239]. The superexon was integrated at the target locus (just upstream of the F508 locus) with a targeting efficiency of approximately 10% [239]. Monoallelic integration was found in five out of 48 sequenced clones; three of these five clones produced chimeric CFTR mRNA as measured by allele-specific RT-PCR after drug-induced demethylation of the *CFTR* locus [239]. One successfully edited clonal cell line was then expanded into a pseudostratified epithelium in ALI culture, and subsequently analysed using an Ussing chamber [239]. While unedited CFBE41o- cells yielded no significant transepithelial current, edited cells displayed current comparable to wild-type airway epithelial cells [239].

ZFNs have also been used to correct CFTR defects in CF patient-derived airway basal cells of various genotypes. Firstly, ZFN mRNA targeting the F508del variant was electroporated alongside a 200-bp ssODN donor molecule consisting of the restoring edit flanked by approximately 100 bp of homology on either side [240]. Electroporation of the targeting ZFN mRNA alone yielded 44.6 ± 2.4% indel formation while co-delivery with the repair oligo yielded 10.6 ± 2.6% correction of the F508del variant (measured by percent of sequenced alleles harbouring each result), yielding an HDR:indel ratio of 1:5 [240]. In order to increase rates of HDR, a longer DNA donor containing two 1 kb homology arms was introduced via AAV6 transduction. This approach increased the editing efficiency to 31.0 ±

4.0% and an improved HDR:indel ratio of 3:2 [240]. Cells corrected using the AAV6 donor

strategy were expanded on ALI culture and assayed using an Ussing chamber, along with

non-CF cells and F508del-homozygous cells treated with the CF double-combination drug

VX-809/VX-770. When exposed to the CFTR activator forskolin, corrected cells exhibited

40.2% of non-CF increase in current, compared with 26.7% of the normal CFTR current in CF

F508del cells which had been treated with VX-809/VX-770- [240]. When exposed to the

CFTR inhibitor CFTRinh-172, corrected cells exhibited 49.4% of non-CF reduction in current,

compared with 31.3% non-CF current reduction exhibited by VX-809/VX-770-treated cells

[240].

Secondly, in the same study, ZFN mRNA targeting intron 8 was electroporated

followed by AAV6 transduction of the *CFTR* superexon cassette *SA-CFTR$_{9-27}$-pA*. This

construct encodes wild-type *CFTR* exons 9-27; provided that patient variants occur in or

after exon 9, targeted addition of this superexon in intron 8 will create a functional *CFTR*

gene. Electroporation of the targeting ZFN mRNA alone yielded 86.5% indel formation; co-

delivery with AAV6 donor yielded 56.5 ± 7.4% of NGS-surveyed alleles carrying the

successfully integrated superexon, and an integration:indel ratio of 5:3 [240]. In two

independent experiments, Ussing analysis was performed using the same comparison

populations (non-CF and F508del-homozygous drug-treated) as in the prior correction

strategy. Across the two experiments, there was some variation in the non-CF and

superexon-integrated cells' responses to forskolin and CFTRinh-172. However, in both

experiments, superexon-integrated cells either matched or outperformed VX-809/VX-770-

treated CF cells, as measured by the percent of non-CF current change achieved in response

to both inhibition and activation.

Finally, the AAV6 superexon integration approach was used in G542X/R785X compound heterozygous primary airway basal cells for an integration rate of 61.8%±6.0% and an integration:indel ratio of 1:2. Corrected cells were expanded on ALI and analysed via Ussing chamber using the same comparison populations as above. When exposed to forskolin, superexon-integrated cells exhibited 33.8% of non-CF increase in current; when exposed to CFTRinh-172, superexon-integrated cells exhibited 30.6% of non-CF reduction in current; CF cells treated with VX-809/VX-770 did not respond to the CFTR activator or inhibitor [240].

1.6 Challenges of editing in the CF airway environment

1.6.1 Challenges to editing epithelial cells *in vivo*

One of the major obstacles to precise genome editing is the fact that HDR does not occur in non-dividing cells [241, 242]. The post-mitotic cells of the airway epithelium, therefore, cannot be precisely edited using HDR. Additionally, these cells turn over approximately every 30 to 50 days [243]; therefore, even if editing is achieved, it must be repeated routinely, which is not possible with immunogenic vectors like AAV [244].

1.6.2 Towards editing basal cells *in vivo*

Underneath the epithelial layer lies a population of airway basal cells, which are capable of proliferation and differentiation [68]. These multipotent stem cells represent an attractive option for targeted CRISPR, as they can be repaired by HDR and will continuously give rise to edited progeny cells. However, in order to access these basal cells *in vivo*, the outer epithelial layer of cells must first be stripped away. Multiple strategies have been proposed for achieving this.

Firstly, mechanical brushing of airway and/or nasal surfaces is commonly used to harvest airway epithelial cells from patients with airway diseases [245]. Scraping can clear swaths of epithelial cells while leaving basal cells undamaged [246]. Chemical stripping of the epithelium is difficult *in vivo* without risking the overall health of the airway. Previous studies of airway healing after chemical stripping of the epithelium have used chlorine gas, proteolytic enzymes, naphthalene, and irradiation [68, 247, 248]. Another potential avenue is the use of common, low-risk rhinoviruses to irritate the epithelium. Rhinovirus C and H1N1 influenza viruses, which are known to target ciliated cells in the airway, result in infected cells being shed from the epithelium [249]. Therefore, deliberate infection could potentially create access to progenitor cells, although this approach carries risk, particularly in the CF airway, where chronic activation of the immune system is already a major issue.

A less harmful approach is loosening of the tight junctions between epithelial cells to allow for leakage of CRISPR machinery into the basal layer. For example, treatment of the airways with lysophosphatidylcholine (LPC), a mild detergent, exposed the basal cells and led to significantly higher rates of lentiviral transduction than mice pretreated with phosphate-buffered saline (PBS) [250]. Mice pretreated with PBS yielded an average of $10 \pm 8.59\%$ transduced cells per counted region, whereas mice pretreated with 0.1% and 0.3% LPC yielded an average of $39.7 \pm 9.2\%$ and $24.35 \pm 8.23\%$ respectively [250]. While the majority of transduced cells were ciliated epithelial cells, treatment with LPC did allow some instances of successful transduction in basal cells [250]. However, accurate quantification of LPC-mediated basal cell penetration would require either longer timescale experiments to evaluate presence of transduction reporter (in this case X-gal) in the next generation of surface epithelial cells post-turnover, or co-labelling with basal cell markers. Additionally, differently sized particles may permeate with different efficiency through LPC-loosened

tight junctions; therefore, investigation into the effect of particle size on permeation is also

necessary to assess versatility in introducing different biochemical machinery, such as

differently sized viral vectors, nanoparticle-encapsulated plasmids, CRISPR RNPs, etc.

1.6.3 Transplantation of *ex vivo*-edited cells

Ex vivo gene editing is an attractive option due to the fact that actively dividing

progenitor cells can be edited via HDR before being differentiated into epithelial cells, and

altered genomes assessed for off-target edits before being introduced into the body. The

success of this approach will depend on the development of safe and efficient methods for

engraftment of edited cells back into the airways and subsequent successful differentiation

into the pseudostratified epithelium. Successful implantation of basal epithelial cells has

been shown to be enhanced by airway pre-injury, exposing the basement membrane and

creating a niche for engraftment of edited cells. NHBE cells were transduced via lentivirus to

express GFP and intratracheally introduced into the airways of mice [251]. Mice were

pretreated with 2% polidocanol administered into the airway, a procedure previously shown

to cause extensive shedding of epithelial airway cells, resulting either in exposure of basal

cells or complete denuding of the basement membrane [252]. Twenty-four hours after

polidocanol treatment, GFP+ NHBE cells were administered and 48 hours later, cell

retention efficiency was assessed in lung homogenates and reported as 10.48 ± 2.86% [251].

In a similar study, 2% polidocanol was administered to mice prior to introduction of

human airway basal cells (hABCs) transduced via lentivirus with luciferase. After one week,

luciferase fluorescence intensity was quantified via *in vivo* bioluminescence imaging [253].

Mice pretreated with PBS showed no bioluminescence at any timepoint. Three of seven

mice pretreated with polidocanol two hours prior to hABC cell delivery showed significant

bioluminescence compared with the PBS control at both three and five weeks before dropping to insignificant levels at five and eight weeks, indicating that transduced and implanted cells had died off or their transgene expression silenced. Additionally, three of the ten initial hABC-treated animals died during cell delivery [253].

Another potential chemical for pretreatment is naphthalene (NA), as shown by successful implantation of a mixture of murine lung tissue stem cells (TSCs), a rare basal cell subtype, and progenitor cells [254]. TSCs cultured *in vitro* were found to create clones consisting of a mixture of TSCs and basal progenitor cells. Attempts to purify TSCs from these mixed pools resulted in inhibition of TSC self-renewal [254] and so transplantation was attempted with the mixed population. Similarly to the study described above, airway pre-injury was found to significantly enhance cell implantation. No post-transplant repopulation occurred in NA-untreated mouse airway, nor in areas of the airway that were undamaged by NA treatment, such as the parenchyma [254]. LacZ-labelled TSC/basal progenitor cell mixtures were transplanted into the NA-injured airways of NOD/SCID/TCR-γ mice. Two weeks post-transplant, implanted cells were found to have repopulated the injured trachea with basal and secretory cells at frequencies similar to the host-derived populations, although differentiation into ciliated cells was significantly decreased compared to host populations [254].

Subsequently, transplantation was tested with GFP-marked human-derived TSC/basal progenitor cells. Two weeks post-transplantation, GFP+ cells were found in the trachea and intrapulmonary regions; as with LacZ-labelled mouse-derived cells, no successful implantation occurred in non-preinjured parenchyma [254]. In the majority of transplanted airways, transplant-derived cells were less frequent compared to host-derived cells. Host differentiation ratios were incongruent with host-derived populations, with an

overrepresentation of mucus-producing cells and underrepresentation of basal and ciliated cells [254]. Overall, despite representing a negligible portion of basal cells in bronchial airways, transplant-derived cells represented approximately half of epithelial secretory and ciliated cell types [254].

In another study, embryonic cell engraftment was explored as an option for repopulation of injured airway. In both human and mouse samples, tissues harvested from the canalicular stage of embryonic lung development were found to contain high concentrations of epithelial, endothelial, and mesenchymal progenitor cells [248]. Similarly to the study above, C57BL/6 mouse airways were injured by NA prior to transplantation. However, 48 hours after NA exposure, mice were also exposed to 6 Gy total body irradiation to limit endogenous host cell repopulation, which allowed transplanted cells to more successfully compete for niche engraftment [248]. Several hours after completion of this preconditioning regimen, GFP+ canalicular cells were transplanted into the airway. After 8 weeks, transplant-derived cells were found to occupy an average of 11 ± 3% of recipient lungs. After 16 weeks, transplant-derived cells occupied 28 ± 6% of lungs [248]. Of note, upon staining for different marker proteins of lung cell types, several cells were found to express CFTR, though protein functionality and ion conductance were not assayed [248]. Researchers then attempted to use the preconditioning regimen to prime NOD/SCID mouse lungs for transplant with cells derived from human embryos. However, due to the increased sensitivity to radiation in the mouse line, a lower dose was used [248]. After transplant, antibody staining revealed significantly reduced integration of human donor cells compared to mouse embryonic cell transplantation [248]. This was hypothesized to be due to the lower dose of radiation, or due to differences between human and mouse cytokines, which has previously been implicated as a hindrance to chimerism [248].

It has also been proposed that basal cells from the sinuses of CF patients could be edited *ex vivo* to repair the *CFTR* mutation; edited cells would then be engrafted back into the sinuses using a biomaterial to promote engraftment [255]. Vaidyanathan *et al.* reported that edited airway basal cells were successfully implanted on a porcine-derived membrane, after which they retained their capability to proliferate and differentiate. The authors proposed that the membrane, which had already received approval from the United States Food and Drug Administration, could then be surgically engrafted into patient airways [255].

1.7 Delivery formats for DNA editing machinery

As outlined in the studies discussed above, various molecular formats can be used to deliver Cas9 and gRNA. The most commonly used options are DNA (plasmid or viral), mRNA, or RNP.

Plasmid DNA is a reasonably effective method of delivering Cas9 and gRNAs; however, it comes with several disadvantages. Firstly, there is the delay in expression due to the time it takes for protein expression of Cas9 and RNA expression of gRNA cassettes. Secondly, plasmids yield longer-term expression of Cas9, which can increase the risk of off-target DSBs, creating unwanted indels and risk of apoptosis [141, 237, 256]. Exacerbating this risk is the potential for plasmid DNA to integrate into the host genome [256]. Thirdly, in both relevant studies discussed above, plasmids included cassettes selecting for antibiotic resistance to aid in selection of edited cells; use of antibiotic resistance selection markers is commonly discouraged in clinical settings and *in vivo* research [257].

In order to avoid the drawbacks of plasmid delivery, Cas9 may be delivered as *in vitro* transcribed messenger RNA (mRNA) mixed with synthetic gRNA, or as purified Cas9 protein

pre-loaded with gRNA in a ribonucleoprotein complex (RNP) [256, 258]. In addition to the transient activity of purified mRNA and RNPS, these approaches enable the use of synthetic gRNAs which may be chemically modified to greatly enhance targeting efficiency and reduce cytotoxicity [259-262]. RNP transfection efficiency outperforms plasmid encoded Cas9/gRNAs transfection in multiple cell types, including human iPSCs [221, 256]. Transfection of RNPs via electroporation has been shown to efficiently cut the *CFTR* gene at the F508, G542, and G551 loci in wild-type iPSCs, with indel formations of 30%, 51%, and 48% respectively [221]. In human upper and lower airway stem cells, electroporation of optimised RNP yielded over 80% cutting (resulting in 38% indel formation and 43% homologous recombination) at the F508 locus [223].

However, electroporation is not a viable option for *in vivo* airway delivery. Lipid-based nanoparticles are a promising option due to their low immunogenicity, high efficiency, and biodegradability.

1.8 Delivery to the lungs via lipid nanoparticles (LNPs)

Delivery of gene editing reagents to the lungs can be accomplished using LNPs. Careful selection of lipid composition and conjugation with targeting peptides can significantly impact LNPs' transfection efficiency of difficult-to-transfect tissues. LNPs have been investigated for delivery of CRISPR machinery as mRNA and RNP, discussed in depth below.

1.8.1 General considerations for enhancing LNP delivery

LNP delivery involves four major considerations: cell targeting, cell penetration, endosomal escape, and cargo release. Much work has been done to optimise LNP formulation at each of these critical stages.

1.8.1.1 Receptor targeting and cell-penetrating peptides

Different cells express different profiles of receptor proteins, and each receptor protein comes with a set of molecules to which it preferentially binds. Researchers have taken advantage of this by conjugating LNPs with peptide moieties that bind specific cell types, enhancing LNPs' targeting and penetration in therapeutically relevant tissues.

One effective tool for discovering such peptides is biopanning accomplished by phage display [263]. In this approach, a library of phages is generated that display randomised short sequences of amino acids. The library is introduced to a cell type of interest, and non-binding phages are eliminated in a series of stringent washing steps. Only phages displaying strongly bound peptides remain and are subsequently eluted for purification and sequencing. The result is a population of peptides that bind to an unknown receptor expressed by the cell type of interest. Additionally, different purification protocols can distinguish between intracellular (i.e. membrane-penetrating) and extracellular (i.e. membrane-affiliated) peptides.

A phage display study performed on airway cells yielded a peptide ("Peptide Y") with sequence similarity to a protein of unknown function found in the bacterium *Legionella pneumophila*, which is known to infect airway cells. Peptide Y was found to significantly enhance transfection efficiency of pDNA to human airway epithelial cells when conjugated to LNPs compared to a sequence-scrambled peptide control [264]. This peptide was more

recently used in combination with LNPs to successfully deliver therapeutic siRNA to CF

mutant bronchial epithelial cells [265].

1.8.1.2 Endosomal escape

Upon uptake by the cell, LNPs are typically encapsulated via endocytosis [266, 267].

Endosomes containing foreign material are normally directed to the lysosome; therefore,

LNP escape from endosomes is crucial to avoid degradation of cargo [267, 268]. The use of

"helper lipids" to aid in endosomal escape has been documented for over 25 years [269].

For example, the lipid dioleoylphosphatidylethanolamine (DOPE) is frequently used for this

purpose. DOPE has a molecular cone-shaped geometry, which promotes assembly into non-

lamellar hexagonal 3D structures [270, 271]. These hexagonal structures are fusogenic with

endosomal membranes, aiding in endosomal escape [268, 271-273].

1.8.1.3 Optimising lipid discovery pipelines

A major consideration is that efficient LNP delivery *in vitro* does not always translate

to the *in vivo* environment, especially when particles are subjected to nebulisation for

delivery [274, 275]. To address this issue, a 2021 study attempted to generate an *in vivo*

protocol for LNP discovery [276]. Three characteristics (lipids used, molar ratios of lipids,

and charge) were selected and treated as axes in an N-dimensional chemical space. Eight to

twelve LNPs were formulated at either extreme of each axis. LNPs in these "extreme

groups" assessed for their size and polydispersity (i.e. consistency of size), stability, and

transfection efficiency in the murine lung. Traits found to be favourable from each extreme

group were combined into LNPs situated at the intersection of relevant axes and

subsequently re-assessed. Using this pipeline, the following "rules" were established: 1)

PEGylated lipids tend to form smaller and more stable LNPs than non-PEGylated, 2) helper lipids and PEG tend to enhance transfection efficiency, 3) inclusion of neutral helper phospholipids reduces the need for PEG. The pool of LNP formulations generated by these rules exhibited variable stability through the nebulisation process. Despite this, the pipeline was successfully used to design and optimise a novel LNP formulation that efficiently delivered mRNA to the murine lung. This formulation, consisting of high-molarity PEG and cationic lipid, was used to deliver mRNA encoding an H1N1 antibody to mice; mice treated with the LNP exhibited significantly higher survival rate when subsequently exposed to H1N1 [276].

1.8.2 CRISPR mRNA delivery to the lung *in vivo*

mRNA replacement therapy delivered via LNPs has been investigated for several diseases, including CF. However, delivery of Cas9 mRNA/sgRNA can be used to permanently correct mutant cells, representing a potential one-dose fix.

A 2017 study examined the comparative kinetics of sgRNA and Cas9 mRNA delivery, and found that *in vivo* co-delivery in the same zwitterionic LNP yielded widespread (lung, liver, and kidney) but low-efficiency (<4% of cells by FACS) editing [277]. Another study further characterised the fusogenic properties of lipids capable of assuming the hexagonal HII phase. In this phase, lipids pack such that polar head groups face inward and hydrocarbon tails face outward, stacking to create tube-like formations. These tubes are themselves capable of stacking together. The exposed hydrophobic hydrocarbon tails are then capable of fusing with membranes, aiding in endosomal escape [278]. Additionally, the cationic helper lipids used in this study led LNPs to accumulate in the lungs after intravenous administration, and delivery of Cas9 mRNA/sgRNA generated 28.3% editing in the lungs as

measured by T7E1 assay [279]. However, each of these studies used intravenous (systemic) delivery. In the case of direct pulmonary delivery, one study utilising cationic LNPs observed 7.2% editing efficiency upon intratracheal injection [280].

Delivery of CRISPR machinery via mRNA has many advantages over pDNA, as previously discussed. However, it also has disadvantages when compared to RNP. Firstly, Miller *et al.* demonstrated that sgRNA activity and Cas9 expression in the cell peaked at different times [277]. This could be alleviated by staggered delivery of each component; however, co-delivery in the same NP is more advantageous to ensure that sgRNA and Cas9 are present in the same cells. Additionally, the delay in translation of mRNA into Cas9 allows for more opportunity for the cell to recognise and destroy foreign genetic material, and RNA is more susceptible to intracellular enzymatic degradation than a pre-translated protein.

1.8.3 CRISPR RNP delivery

Delivery of CRISPR RNP to airway cells via LNP has been investigated in a small number of studies and has achieved only modest editing efficiency upon systemic *in vivo* delivery. In most LNP studies, intravenous delivery leads to accumulation in the liver [281]. One study sought to alter lipid composition in order to enhance tissue specificity. Selective organ targeting (SORT) nanoparticles were initially developed for mRNA delivery to specific tissues by titrating up the quantity of cationic lipid in LNP assembly [282]. Higher concentration (~50%) of the cationic lipid 1,2-dioleoyl-3-trimethylammonium-propane (DOTAP) was found to lead to higher levels of accumulation in the lungs compared to liver and spleen upon systemic *in vivo* delivery. Upon packaging and delivery of CRISPR RNP, 5.3% editing was observed in the lungs.

To avoid the issue of liver accumulation, aerosolised LNPs can be delivered directly to the airway via intratracheal injection or nebulisation. However, the CF lung presents unique difficulties in delivery of aerosolised particles due to the accumulation of thick, sticky mucus.

1.8.4 A potential avenue for mucus penetration

Cationic lipids are favourable for their association with negatively charged cargo and negatively charged components of cell membranes. However, the negatively charged mucins, mucopolysaccharides, and neutrophil DNA in CF mucus can lead to retardation of positively charged particles [283]. In multiple studies, researchers have proposed conjugation of lipids with the neutral molecule poly(ethylene glycol) (PEG) as a solution for this issue.

Chitosan, a cationic linear polysaccharide, is a derivative of the naturally occurring biodegradable molecule chitin. Due to its positive charge, chitosan naturally complexes with negatively charged nucleic acids. However, as with cationic lipids, this also confers mucoadhesive properties. A 2018 study investigated conjugation of chitosan with PEG to enhance mobility in mucus [284]. PEG-COOH was conjugated to chitosan to generate PEGylated chitosan. In a mucus permeation assay, it was demonstrated that DNA mobility through mucus-coated transwells was undetectable with chitosan-DNA complexes, but mobility was improved by PEGylation. PEG-chitosan-DNA nanoparticles were also found to have favourable size (diameter <200 nm at optimal chitosan:DNA ratio) before and after nebulisation. However, transfection efficiency never reached >20% of that achieved by the commercially available transfection reagent Lipofectamine in nebulised or unnebulised conditions [284].

Another study investigated PEGylation of cell penetrating peptides (CPPs) for *in vivo* use in the CF airway [285]. Glycosaminoglycan (GAG)-binding enhanced transduction (GET) peptides consist of cell-targeting and cell-penetrating domains. The peptide used in the study (termed FGF2B) carries a cationic charge at pH ≤ 7. Furthermore, for transfection purposes, it was fused to an amphiphilic sequence and a cell-penetrating octa-arginine, generating a GET with a significant net positive charge. PEGylation was investigated as a method to make GET-DNA nanocomplexes more motile in mucus. Firstly, researchers established a ratio of PEG:GET-DNA nanocomplexes that did not significantly affect particle size, aggregation, or protection of DNA from enzymatic degradation. However, transfection efficiency of *in vitro* cells was found to inversely correlate with % PEGylation. 40% PEGylation was selected as the optimal ratio for balancing transfection efficiency with stability in biological fluids. In CF sputum, PEG-GET-DNA complexes were found to exhibit six-fold motility (measured by square of distance travelled by an individual particle in a given time interval) compared to non-PEGylated complexes. In an *in vivo* healthy mouse lung, aerosolised PEG-GET-DNA complexes were found to achieve significantly higher epithelial coverage and parenchymal distribution than non-PEGylated complexes. Additionally, there was no significant difference between inflammatory response between PEGylated and non-PEGylated complexes [285].

1.9 Aims of this project

The aims of this thesis are as follows:

1. To characterise a *BMI-1* transduced primary airway epithelial cell line homozygous for a *CFTR* variant. This will provide a baseline for development of a therapeutic treatment for an "undruggable" variant. Furthermore, scanning ion conductance microscopy has not yet been performed for cells harbouring undruggable *CFTR* variants [286].

   Hypothesis: A pseudostratified epithelium made up of CFNE BMI-1 R709X variant cells will display negligible CFTR-mediated ion trafficking, reduced ASL height, and decreased mucosal elasticity compared to non-variant cells.

2. To assess the efficiency of HITI in delivery of a reporter gene to a safe harbour locus of the genome, and to evaluate the effects of repeat dosage and subsequent off-target editing effects of CRISPR RNP delivery to airway epithelial cells. Variant-specific DNA editing strategies are not favourable for a disease with hundreds of causative/associated variants [232], so laying the groundwork for a full gene replacement represents an important step forward for CF treatment. Furthermore, because HITI uses NHEJ pathways [152], it can be performed in the target tissue of interest for an *in vivo* therapy (fully differentiated airway epithelial cells).

   Hypotheses: 1) HITI at the AAVS1 safe harbour locus will result in rates of editing efficiency that satisfy those predicted to be required for restoration of CFTR function; 2) The native PPP1R12C promoter will provide constitutive expression of

integrated constructs.

3. To examine the biophysical properties of cationic lipid-based receptor-targeted nanoparticles (RTNs) and their ability to transfect populations of patient-derived airway epithelial cells. RTNs have been thoroughly characterised for delivery of plasmid and RNA to airway epithelial cells [96, 265, 272, 287-290], but substantially less work has been done in delivery of preformulated CRISPR ribonucleoprotein. Additionally, RTN transfection of airway epithelial cells is more frequently performed in submerged, undifferentiated basal cells [57, 273, 287-289]. Transfection of ALI culture is typically hindered by the biophysical properties of differentiated cells, such as cilia, tight junctions, and (most relevant in CF) mucus [291]. Therefore, RTNs competent in transfection of submerged cells must also be verified in ALI culture.

    Hypotheses: 1) Nanoparticles made up of cationic lipids will efficiently self-assemble with anionic RNP cargo; 2) Conjugation with a receptor-targeting peptide will enhance transfection efficiency; 3) RTNs will be more motile in normal human mucus than CF mucus; 4) If motile in mucus, RTNs will effectively transfect a pseudostratified epithelium of airway epithelial cells.

2. Materials and Methods

2.1 In vitro receptor-targeted nanoparticle (RTN) formulation and characterisation

2.1.1 CRISPR RNP assembly

sgRNAs (tracrRNA pre-complexed with crRNA) were ordered from Integrated DNA Technologies (IDT) and diluted to 100 ng/uL in nuclease-free H_2O. *Sp*Cas9 was ordered from various sources (IDT, GenScript, Thermo Fisher) depending on desired modification (RFP-tagged, GFP-tagged, or untagged, respectively). RNPs were complexed at a weight ratio of 4:1 Cas9:sgRNA. In 24 well plates 70-80% confluent with cells, 625 ng of RNP was used (500 ng Cas9 + 125 ng gRNA). Components were mixed in OptiMEM for transfections or in CRISPR RNP buffer (20 mM Tris, 20% glycerol) for cell-free RTN characterisation experiments. RNPs were allowed to self-assemble at room temperature for 5 minutes before downstream application.

2.1.2 RTN Assembly

Lipid, peptide, and cargo (RNP or plasmid) were diluted in equal volumes (typically to 50 µL each) of sterile water or OptiMEM as appropriate for the desired application. Unless otherwise stated, components were mixed at a 1:4:1 lipid:peptide:cargo weight ratio, and mixed in the order cargo-lipid-peptide. For example, in a 24 well plate receiving 625 ng RNP per well, 625 ng of lipid and 2500 ng peptide were used. Components were mixed by thorough pipetting up and down at each addition. RTNs were then left to self-assemble for 15-30 minutes before dilution to the required final volume (for example, 500 µL for a 24-

well plate) unless otherwise stated. The final concentrations were as follows: 1.25 ng/μL

lipid, 1.25 ng/μL RNP, and 5 ng/uL peptide.

Structures of lipids used are as follows:

Cationic lipids

DTDTMA (C14)

DHDTMA (C16)

DOTMA (C18)
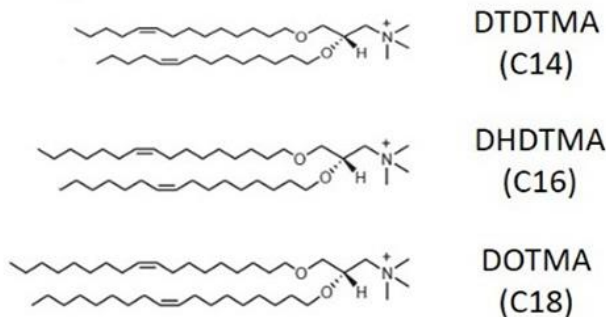
Phospholipids
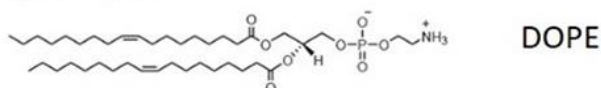
DOPE

Fig. 2.1: Chemical structure of lipids used.

DTDTMA: Ditetradecyl trimethyl ammonium propane. DHDTMA:

Dihexadecenyl trimethyl ammonium propane. DOTMA: Dioctadecenyl

trimethyl ammonium propane. DOPE: 1,2-dioleoyl-sn-glycero-3-

phosphoethanolamine.

From Grant-Serroukh, D., M. R. Hunter, R. Maeshima, A. D. Tagalakis, A. M.

Aldossary, N. Allahham, G. R. Williams, M. Edbrooke, A. Desai and S. L. Hart

(2022). "Lipid-peptide nanocomplexes for mRNA delivery in vitro and in vivo."

J Control Release 348: 786-797.

Cationic lipids were mixed with DOPE at a 1:1 ratio to form C14DOPE, C16DOPE, and

C18DOPE.

Peptide sequence: K_{16}GACYGLPHKFCG (Peptide Y).

2.1.3 Size and charge measurement

Minimum 2 μg of RTN was assembled as above, diluted to 900 μL with water, and loaded into a ZetaSizer cartridge (Malvern). Size and charge (zeta potential) were measured on a Malvern Zetasizer Nano ZS via dynamic light scattering. Size was measured before charge.

2.1.4 Quenching assays

Assembled RTNs were diluted to a final volume of 100 μL (for non-heparin conditions) or 90 μL (to be mixed with 10 μL heparin) using Tris buffer (50 mM pH 7.4). Samples were mixed with 100 μL of RiboGreen (diluted v:v 1:2,000) (Thermo Fisher #R11490, all conditions) and 10 μL of 10 U/mL heparin (for only heparin-containing conditions). For positive control, naked RNP was diluted instead of the RTNs in both absence and presence of heparin. For negative control, 100 μL of pure Tris was used in place of RTNs or RNPs. Heparin used was heparin sodium salt from porcine intestinal mucosa (Sigma-Aldrich #H3149).

2.1.5 Mucus penetration assay

A plastic container was lined with clean paper towels soaked in sterile water. This was left to equilibrate at 37°C for 30 minutes in order to create a humidity chamber. 0.4 μm ThinCerts (Greiner Bio-One #662641) were set up in a 24-well plate. 600 μL of Tris buffer (50 mM pH 7.4) was added to the basal chamber, and 200 μL to the ThinCert. To this apical layer

of Tris, 2 μL of mucus (Epithelix) was added. The plate was then placed in the humidity chamber and allowed to incubate at 37°C for a further 30 minutes. Nanoparticles were created as described and allowed to incubate at room temperature for 30 minutes. After incubation, size and PDI were measured as previously described.

Once apical Tris had drained to the basal chamber, leaving behind a layer of mucus, 10 μL of RNP-containing RTN was added to the ThinCert. Before and between timepoints, plate remained in the humidity chamber in a 37°C incubator. At each timepoint, 200 μL of Tris from the basal chamber was extracted. In a 96 well plate with opaque walls, 100 μL of sample was mixed with 10 μL of heparin and 100 μL of RiboGreen; 100 μL was mixed only with RiboGreen. Fluorescence was measured after 2 minutes of incubation in the dark at room temperature using a FluoStar Optima plate reader with excitation filter wavelength 492 nm, emission filter wavelength 520 nm, and gain 2000. For positive control, 10 μL of RTN in 90 μL of Tris was mixed directly with either RiboGreen and heparin or just RiboGreen, and fluorescence measured as described. For negative control, solely Tris was mixed with either RiboGreen and heparin or just RiboGreen, and fluorescence measured.

2.2 Cell culture

All cells were cultured at 37°C and 5% CO_2.

2.2.1 Collagen coating of plasticware

All coating was performed in a laminar flow tissue culture hood. Plasticware was coated with an appropriate volume of 10% PureCol collagen solution (Advanced BioMatrix #5005) in sterile phosphate-buffered saline (PBS) (Thermo Fisher #10010023). Plasticware

was incubated with collagen solution at room temperature for 10 minutes, rinsed with pure

PBS, and allowed to air dry. Plasticware was used within six weeks of coating.

2.2.2 BMI-1 lentiviral transduction

On day one, 1×10^5 primary nasal epithelial cells carrying the R709X variant were

seeded in collagen-coated 6-well plates in 1 mL of PneumaCult EX-Basal medium (hereafter

referred to as "basal media") (Stemcell Technologies #05008). On day two, lentivirus

containing a BMI-1 expression cassette was diluted in OptiMEM (Thermo Fisher #15392402)

before further dilution to 1 mL with basal media for a multiplicity of infection (MOI) of one

based on number of cells seeded 24 hours prior and doubling time. Basal media was

replaced with virus-containing media. On day three, another 1 mL of basal media was

added, but virus-containing media was not removed. On day five, virus-containing media

was removed and replaced with 1 mL of basal media. This basal media was replaced every

48 hours until cells reached 90% confluency, at which point they were trypsinised (described

below) and transferred to a collagen-coated tissue culture flask for further expansion.

2.2.3 General cell maintenance (BMI-1 transduced airway basal cells)

Cells were cultured in collagen-coated flasks with basal medium prepared according

to manufacturer's instruction. Upon reaching 70-80% confluency, cells were passaged. Basal

media was aspirated and cells were rinsed using a volume of 0.05% trypsin-EDTA (hereafter

referred to as "trypsin") (Thermo Fisher #25300062) 10% of the flask volume and then

aspirated. After the initial trypsin rinse, another 10% volume of trypsin was added and

incubated on cells at 37°C for ~5 minutes or until cells had lifted. Trypsinised cells were

collected in basal medium, pelleted at 1000 x g for 5 minutes, and resuspended in fresh culture medium before diluting with fresh culture medium into a new collagen-coated flask, seeded into collagen-coated plates, or frozen.

2.2.4 General cell maintenance (cell lines not derived from primary cells)

Cells were cultured in uncoated flasks with appropriate medium (listed below). Upon reaching 70% confluency, cells were passaged. Growth media was aspirated, and cells were rinsed using a volume of PBS 20% of the flask volume. After the PBS rinse, 20% volume of trypsin was added and incubated on cells at 37°C for ~5 minutes or until cells had lifted. Trypsinised cells were collected in basal medium, pelleted at 1000 x g for 5 minutes, and resuspended in fresh culture medium before diluting with fresh culture medium into a new collagen-coated flask, seeded into collagen-coated plates, or frozen.

Media for 16HBE14o- cells: Minimum Essential Medium (Sigma-Aldrich #M7278)

Media for HEK293T cells: Dulbecco's Modified Eagle Medium, high glucose (Thermo Fisher #11965092)

2.2.5 Air-liquid interface (ALI) culture

Basal cells were seeded on the apical side of collagen-coated polyester transwells (for microscopy) or snapwells (for SICM and Ussing analysis) (Corning #3460 and #3801 respectively). Basal media was added to the basal chamber and cells were left submerged until they formed a confluent monolayer (3-5 days), changing both apical and basal chamber media every 48 hours. Once confluence was achieved, cells were airlifted. The apical layer of

media was aspirated, and the basal media was aspirated and replaced with PneumaCult ALI

Basal medium (herefore referred to as ALI media) prepared according to manufacturer's

instruction (Stemcell Technologies #05001). For 28 days, ALI media in basal chamber was

changed every 48 hours. Once cells began producing mucus, excess mucus was washed off

at each feeding: 200 µL PBS was added to the apical surface, incubated at 37°C for 30

minutes, and aspirated.

2.2.6 Freezing and thawing of cells

Cells for freezing were trypsinised and collected as described above. Cells were

counted using a haemocytometer and aliquots of 1×10^6 cells were collected in 900 uL. 100

µL of dimethylsulfoxide (DMSO, Thermo Fisher #20688) was added and mixed thoroughly

before transferring into cryosafe screw-cap tubes. Cryoprotectant was permitted to

permeate cells for 10 minutes before transfer to -80°C. Cells were transferred from -80°C to

liquid nitrogen within 5 days.

Cells for thawing were removed from liquid nitrogen and thawed in a 37°C water

bath with gentle agitation. Once cells had thawed, the entire 1 mL volume was diluted to 10

mL with warm growth media. Cells were pelleted at 700 x g for 5 minutes. Media and

cryoprotectant were aspirated and cells were resuspended in appropriate growth media.

Cells were then transferred to an appropriate flask (collagen-coated or uncoated as per

normal culturing conditions) and maintained as normal.

2.2.7 Serial dilution of cells for single-cell cloning

Single-cell cloning was executed in 96 well plates via serial dilution according to the "Cell Cloning by Serial Dilution in 96 Well Plates" protocol available from Corning Life Sciences.

https://www.corning.com/catalog/cls/documents/protocols/Single_cell_cloning_protocol.pdf

2.3 Transfections

All transfections were performed in a laminar flow tissue culture hood.

2.3.1 Transfection of submerged HEK293T cells with pDNA in Lipofectamine 2000

Cells were seeded on an uncoated 24-well plate 24 hours in advance to be 70-80% confluent at time of transfection. An amount of Lipofectamine 2000 (L2K) appropriate for a 3:1 weight ratio of L2K:plasmid was diluted in 50 μL OptiMEM per well, and incubated for 5 minutes per manufacturers' instructions (Thermo Fisher #11668019). In a 24 well plate, 500 ng Cas9/gRNA plasmid and 250 ng donor plasmid were used. Each plasmid component was diluted as appropriate in a 50 μL working volume per well (whether donor and Cas9 plasmid were combined for packaging in one nanoparticle, or kept divided for packaging as separate nanoparticles, depended on experimental design). L2K and plasmid were combined and incubated at room temperature for 20 minutes. After incubation, media was aspirated from cells and replaced with 400 μL of fresh media. 100 μL of L2K/plasmid mix was added to each well.

2.3.2 Transfection of submerged 16HBE cells with pDNA in RTNs (for luciferase assay)

Cells were seeded on uncoated 96-well plates 24 hours in advance to be 70-80% confluent at time of transfection. RTN components were diluted to working solutions in 50 µL of OptiMEM before being mixed in the order lipid, cargo, peptide for a total of 150 µL incubation volume. RTNs were incubated for 30 minutes at room temperature to allow self-assembly. After incubation, RTNs were diluted to a final working volume of 200 µL per well with OptiMEM. Culture medium was aspirated from seeded cells and replaced with RTN-containing OptiMEM. The plate was spun at 1000 x g for 5 minutes and subsequently left to incubate at 37°C and 5% CO_2 for four hours. After incubation, OptiMEM was aspirated and replaced with fresh culture medium.

2.3.3 Transfection of submerged BMI-1 airway basal cells with RTNs

Cells were seeded on 24-well collagen-coated plates 24 hours in advance to be 70-80% confluent at time of transfection. RTN components were diluted to working solutions in 50 µL of OptiMEM before being mixed in the order lipid, cargo, peptide for a total of 150 µL incubation volume. RTNs were incubated for 30 minutes at room temperature. After incubation, RTNs were diluted to a final working volume of 500 µL per well with OptiMEM (added in 500 µL increments, mixing thoroughly by pipetting up and down with each addition). Culture medium was aspirated from seeded cells and replaced with RTN-containing OptiMEM. The plate was spun at 1000 x g for 5 minutes and subsequently left to

incubate at 37°C and 5% CO_2 for four hours. After incubation, OptiMEM was aspirated and replaced with fresh culture medium.

2.3.4 Transfection of ALI cultures

200 µL PBS was added to the apical surface of ALI cultures and incubated at 37°C for 30 minutes before aspiration in order to remove excess mucus. RTNs were assembled and incubated in 150 µL as above. After incubation, RTNs were diluted to a final working volume of 150 µL per well with OptiMEM. 150 µL was added to the apical surface to submerge cells in RTN-containing solution. Cells were left to incubate at 37°C and 5% CO_2 for four hours. After incubation, OptiMEM was aspirated.

2.4 Post-transfection *in vitro* assays (Bradford and luciferase)

Bradford assays were performed using the Pierce Coomassie (Bradford) Protein Assay Kit (Thermo Fisher # 23200). Luciferase assays were performed using the Promega Luciferase Assay System (Promega #E1500).

Cell media was removed from cells, and cells rinsed twice with PBS. After rinsing, cells were treated with 5x Reporter Lysis Buffer (Bradford) diluted to 1X in nuclease-free H_2O according to manufacturer's instruction. Cells were flash-frozen at -80°C to fully lyse.

Half of cell lysate was reserved for luciferase assay, and half for Bradford assay. Both assays were performed according to manufacturers' instructions. Fluorescence (Bradford) and luminescence (luciferase) were measured via a FluoStar Optima. Relative luminescence units per mg of protein (RLU/mg) was used as a measure of transfection efficiency.

2.5 Molecular biology and Sanger sequencing

2.5.1 DNA extraction from cultured cells

Cells were trypsinised, pelleted, and resuspended in 40 μL PBS. DNA was extracted using DNEasy Blood and Tissue Kit (Qiagen #69504). Manufacturer's instruction was followed until the final elution step. DNA was eluted with 70 μL sterile H_2O; the elution step was repeated for a total of 140 μL eluted into two separate tubes.

2.5.2 Polymerase chain reaction (PCR)

Primers were designed using either A Plasmid Editor (ApE) or the National Center for Biotechnology Information (NCBI) design tools. 100 ng of template DNA and 0.75 μL of each 10-μM primer was used per 25-uL reaction volume. Q5 Hi-Fi 2X PCR MasterMix (New England Biolabs (NEB) #M0492) was used according to manufacturers' instruction.

PCR was performed in a C1000 Touch Thermal Cycler (Bio-Rad). Primer annealing temperature (T_a) was determined using NEB's annealing temperature calculator (https://tmcalculator.neb.com/). For the extension step, 1 minute per kb was used. The following protocol was used for every PCR. The set of steps in bold was repeated for 35 cycles.

98°C – 30 seconds

98°C – 10 seconds

T_a – 30 seconds

72°C – 1 min per kb

72°C – 2 minutes

2.5.3 Gel electrophoresis

1.5% agarose (Thermo Fisher #16500500) gels were made in tris-acetate-EDTA (TAE) (Thermo Fisher #B49, diluted to 1X) buffer. Agarose was melted and allowed to cool to ~55°C. SYBR-Safe DNA gel stain (Thermo Fisher #S33102) was added per manufacturer's instruction. DNA samples were mixed with loading dye (Thermo Fisher #R0611) for loading. Samples were run at 100V alongside GeneRuler 1 kB Plus DNA Ladder (Thermo Fisher #SM1331) until dye front had reached the bottom of the gel. Gels were imaged with the ChemiDoc system (Bio-Rad).

2.5.4 Sanger sequencing (cultured human cells or plasmid DNA)

Sanger sequencing was performed by GeneWiz (Azenta Life Sciences). Samples were prepared according to company guidelines.

PCR was performed on samples as described above, and a small aliquot (~100 ng DNA) was run on an agarose electrophoresis gel to confirm presence of expected band and absence of contamination/primer dimers. Desired amplicons were isolated from the remaining PCR product using column cleanup (New England BioLabs #T1030) and eluted in nuclease-free water. Concentration of cleaned product was verified via NanoDrop (260 nm wavelength), and samples were diluted as appropriate (dependent on amplicon or plasmid size, per https://www.genewiz.com/en-GB/Public/Resources/Sample-Submission-Guidelines/Sanger-Sequencing-Sample-Submission-Guidelines/Sample-Preparation#sanger-sequence) with nuclease-free water.

2.5.6 Inference of CRISPR Edit (ICE) analysis

Synthego's ICE software was used to quantify CRISPR-mediated editing efficiency. ICE uses heterogeneity of Sanger-derived chromatogram traces to infer indel formation. Sanger traces were uploaded and analysed according to instruction (https://ice.synthego.com). gDNA from untransfected cells was used to generate control sequence.

2.5.7 Plasmid cloning, transformation, and sequencing

Parental plasmids were generated by previous members of the Hart lab. Restriction enzymes (REs) were obtained from New England Biolabs (NEB). Plasmids were digested with appropriate REs according to manufacturer's instruction, run on agarose gels, and appropriate fragments were excised and cleaned up using the Monarch DNA Gel Extraction Kit (NEB #T1020). Ligations were performed using T4 DNA ligase and buffer (NEB #M0202) according to manufacturer's instruction.

OneShot TOP10 Chemically Competent *E. coli* cells (Thermo Fisher #C404010) were used for transformation per manufacturer's instructions. Bacterial cells were expanded from single colonies in Vegitone LB broth (Sigma-Aldrich #28713) or on LB-agar plates (Sigma-Aldrich #19344) as needed, with appropriate antibiotics included. To make frozen stocks, 500 µL of overnight liquid culture and 500 µL of 50% glycerol were combined and kept at -80°C. To recover bacteria, frozen stock was scraped with a sterile loop, swiped on an agar plate with appropriate antibiotics, allowed to grow overnight, and a single colony inoculated in overnight culture.

Plasmid DNA was extracted using GeneJet Plasmid Miniprep or Maxiprep Kits (Thermo Fisher K0502 and K0491, respectively) depending on the quantity of plasmid

required for downstream application. Plasmids were Sanger sequenced by GeneWiz Azenta

Plasmid-EZ. Plasmid maps and details of cloning are shown in Chapter 4.

2.6 Biophysical assays for ALI cultures

2.6.1 Ussing chamber analysis

Ussing chamber experiments were performed in the lab of Dr. Deborah Baines at St.

George's Hospital.

Ussing chambers were filled with Krebs Bicarbonate Ringers solution. Buffer was

prepared in sterile distilled H_2O and consisted of 115 mM NaCl, 25 mM $NaCO_3$, 2.4 mM

K_2HPO_4, and 0.4 mM KH_2PO_4, pH 7.4. Buffer was bubbled at 37°C with 21% O_2, 5% CO_2, and

74% N_2 gas before addition of $CaCl_2$ (1.2 mM) and $MgCl_2$ hexahydrate (1.2 mM). All

chemicals from Sigma-Aldrich.

Upon placement of fully differentiated ALI cultures within the Ussing chamber,

cultures were kept in open-circuit conditions until trans-epithelial voltage potential (V_t) and

resistance (R_t) had stabilised. Subsequently, current was applied via DVC-4000

voltage/current clamp until V_t had been cancelled out to zero. This current (short circuit

current, or Isc) was briefly dropped every 30 seconds in order to establish spontaneous V_t

and R_t. Current was recorded via PowerLab software.

Drug application was as follows:

Drug	Purpose	Final concentration in chamber	Apical or basolateral addition?

Amiloride	Inhibit ENaC	10 μM	Apical
Forskolin	Stimulate CFTR	25 μM	Both
CFTRinh-172	Inhibit CFTR	10 μM	Apical

All drugs acquired from Sigma-Aldrich and stock solutions prepared in DMSO.

2.6.2 Scanning ion conductance microscopy (SICM)

SICM experiments were performed under the guidance of Drs. Guy Moss and David Benton, and Mr. Arthur Mitchell (UCL Division of Biosciences). Cells were differentiated on ALI culture as described for four weeks, until cilia movement and mucus production were visible. ALI transwells of differentiated cells were placed in a bath reservoir, which was filled with saline (NaCl 116 mM NaCl, 5.1 mM KCl, 10 mM glucose, 25 mM $NaHCO_3$, 1.2 mM $CaCl_2$, 1.2 mM $MgCl_2$, pH 7.4) until the bottom of the transwell was covered but no liquid entered the apical layer. This bath was placed in a chamber at 37°C, 5% CO_2, and 99% humidity for the duration of measurements. A glass nanopipette probe was generated from pulling a borosilicate capillary tube (1.0 mm outside diameter, 0.5 mm inside diameter) with a P-2000 laser puller (Sutter Instruments) and filled with saline; a wire electrode was then threaded into the back of the probe. The probe was then mounted above the cells in the humidified chamber.

Feedback controls were operated via a program developed by the Moss lab [286]. In brief, the probe was lowered toward cells until the tip became immersed in airway surface liquid (ASL), thus completing a circuit between the bath electrode and probe electrode, which automatically triggered the probe to halt its approach. Subsequently the approach

could be resumed until the probe tip reached the ciliated cell surface, at which point it halted once more, and the distance between the two stopping points was taken as a readout of ASL depth. 21 measurements were recorded for each ALI well assayed, measured at points every 50 nm in a cross formation.

2.7 Microscopy

Confocal microscopy was performed using the Zeiss LSM710 Confocal Microscope. Fluorescence and brightfield microscopy were performed using the Nikon Ti-2 Inverted Microscope. Because cells either endogenously expressed fluorescence, or were transfected with fluorophore-tagged proteins, no antibodies or stains were used.

2.8 Cell sorting and flow cytometry

All flow cytometry was performed on the CytoFlex or CyAN ADP (Beckman Coulter). Cell sorting was performed on the BDFACSAria III (BD Biosciences) or MoFlo XDP (Beckman Coulter). Which machine was used for each application was determined solely by availability, and all parameters (gain, voltage, gating) were kept as consistent as possible across machines. Size-based gates were set to exclude cell fragments and aggregates. Fluorescence-based gates were set according to untreated non-fluorescing cells. Where indicated, cells transfected with promoter-driven fluorophores were used as positive controls.

Samples were prepared as follows. Media was aspirated from cells, and cells rinsed gently with 1 volume (i.e. 500 uL for a 24 well plate) of PBS. 20% volume of 0.5% trypsin-

EDTA was added and incubated on cells for 3-5 minutes at 37°C until >90% of cells had lifted. Trypsin was diluted and inactivated by adding 80% volume of serum-containing growth media. Cell suspensions were collected and centrifuged in Eppendorf tubes at 1000 x g for 5 minutes. Media and trypsin were aspirated, and cell pellets were resuspended in 200 uL PBS and kept on ice until analysis.

2.9 Data analysis

Image analysis was performed in Fiji (specific macros indicated where used in Results). Statistical analysis and graphing was performed in GraphPad Prism (specific statistical tests indicated where used in Results).

2.10 Sequences of Nucleic Acids

2.10.1 gRNAs (Sense)

A78 GTCACCAATCCTGTCCCTAG (-)

Targets the AAVS1 safe harbour locus (intron 1 of *PPP1R12C*, chromosome 19q13.4). This gRNA was the most efficient of six AAVS1-targeting gRNAs tested by Ma *et al.* [292].

GFP GGTGGTGCAGATGAACTTCA (-)

Targets and knocks out function of GFP.

3849 U3 CTTGATTTCTGGGACCACA (+)

3849 D1 TTGATCCAACATTCTCAGGG (+)

Target *CFTR* intron 22 upstream (U3) and downstream (D1) of a mutation causing a cryptic splice site.

gRNAs sourced from IDT.

2.10.2 Primers

Primers were designed using A Plasmid Editor (ApE) software or NCBI PrimerBlast.

AAVS1 F AGTCTTCTTCCTCCAACCCG

AAVS1 R CAAACTGCTTCTCCTCTTGG

Surround the cut site of the AAVS1 locus-targeting gRNA A78 used in Chapter 4 (gRNA sequence above).

GFP Insert Junction R CTCACAATTCCACACAACA

Targets junction of AAVS1 locus and GFP-coding construct used in Chapter 4. When paired with AAVS1 F, amplifies band spanning genomic DNA into inserted construct after successful genomic integration.

A78 OTS 1 F CACAGTGACTGGTGCACTGA

A78 OTS 1 R CAGGCAGGCTGGAAGAAAGA

A78 OTS 2 F GCTACCCAGGCTGATCTCTAATCC

A78 OTS 2 R GACTCCAGCACTTCATCAGCTT

A78 OTS 3 F CAGAGGTTGCAGGAGATCCCATAT

A78 OTS 3 R GACTCCAGCACTTCATCAGCTT

A78 OTS 4 F GGAGGTTGTAGTGAGCGAAGATCA

A78 OTS 4 R CATCAACGGTGCTGAGCATG

A78 OTS 5 F GTTTCCATTTCTCTGCTGAGGCT

A78 OTS 5 R CAAACCCAGGTCTCTCTGCCAT

Amplify predicted off-target genomic sites potentially targeted by A78 gRNA.

CFTR int22 F TCCTCTCAAAATGCCTACTGG

CFTR int22 R TCATCCAGTCTTTCCCAACC

Amplify section of *CFTR* intron 22 including the sites targeted by U3 and D1 gRNAs.

3849 U3 OTS 1 F GCAGTGTGATGAACGTGGTGA

3849 U3 OTS 1 R GGAGGATGTGACAGATTGATTGCA

3849 U3 OTS 2 F CCACCAAGCCCAGCAGATTT

3849 U3 OTS 2 R GAAAGGAAGGAAGATGAGCAGTGG

3849 U3 OTS 3 F CCAAGTGTGCTCAAGGTTCACC

3849 U3 OTS 3 R GCAACACAGCAAGACTCCATC

3849 U3 OTS 4 F GAAGGGAGGATCATCTGAGTCCAG

3849 U3 OTS 4 R GACAGAGGACTTATGGAAGGCG

3849 U3 OTS 5 F CTGGAGAGGAGATTTCGGGCAA

3849 U3 OTS 5 R CCTCATCCATGCTTCCCAGAA

Amplify predicted off-target genomic sites potentially targeted by U3 gRNA.

3849 D1 OTS 1 F GACCCAGGAGTAAGCACTCACAA

3849 D1 OTS 1 R CTTTCTCTGCACCCTCTATAAGAGC

3849 D1 OTS 2 F GGATACGTCAAGCCTAATGAGAGT

3849 D1 OTS 2 R AAGACCTTGAGGGAGGGAGAAATTG

3849 D1 OTS 3 F CCTGAGCAAGCCTTAGTGGTTC

3849 D1 OTS 3 R TCATGTGAGAGAGAGCCTGAGTTAG

3849 D1 OTS 4 F CAACTGGCATAAAGAGGTCTGGG

3849 D1 OTS 4 R TCATGGCTTTGCTATCTCCCAG

3849 D1 OTS 5 F CAGAGCAGACTACGTGCTTACA

3849 D1 OTS 5 R CTGAAGCGTGGAGAAGTGAAGG

Amplify predicted off-target genomic sites potentially targeted by D1 gRNA.

R709X genotyping F CTCATGGGATGTGATTCTTTCGACC

R709X genotyping R GAAAGACAACAGCATCCACACGAA

Amplify region surrounding the *CFTR* R709X variant.

Primers sourced from IDT.

2.11 Materials

2.11.1 Equipment

LSM710 Confocal Microscope (Zeiss)

Ti-2 Inverted Microscope (Nikon)

CytoFlex (Beckman Coulter)

CyAN ADP (Beckman Coulter)

BDFACSAria III (BD Biosciences)

MoFlo XDP (Beckman Coulter)

NanoDrop Spectrophotometer (Thermo Fisher)

C1000 Touch Thermal Cycler (Bio-Rad)

ChemiDoc XRS+ (BioRad)

Zetasizer (Malvern)

Horizontal electrophoresis tanks and power supplies (Bio-Rad)

2.11.2 Kits and Reagents

DNEasy Blood and Tissue Kit (Qiagen)

Monarch PCR and DNA Cleanup Kit (New England BioLabs)

Monarch DNA Gel Extraction Kit (New England BioLabs)

Q5 HiFi 2X MasterMix (New England BioLabs)

GeneRuler 1kb Plus (Thermo Fisher)

6X TriTrack DNA Loading Dye (Thermo Fisher)

Lipofectamine 2000 (Thermo Fisher)

Trypsin EDTA 0.5% with Phenol Red (Fisher Scientific)

PureCol Solution (Advanced BioMatrix)

SOC Medium (Thermo Fisher)

TOP10 Competent Cells (Thermo Fisher)

T4 DNA Ligase (New England BioLabs)

TrueCut Cas9 v2 (Thermo Fisher)

Alt-R CRISPR Custom sgRNAs (Integrated DNA Technologies)

Alt-R *S.p.* Cas9-GFP nuclease (Integrated DNA Technologies)

Alt-R *S.p.* Cas9-RFP nuclease (Integrated DNA Technologies)

GenCrispr NLS-Cas9-EGFP Nuclease (GenScript)

SYBRSafe DNA Gel Stain (Thermo Fisher)

2.11.3 Chemicals

Tris HCl (Sigma Aldrich)

Agarose UltraPure™ Low Melting Point Agarose (Thermo Fisher)

Heparin ammonium salt from porcine intestinal mucosa (Sigma Aldrich)

Phosphate-buffered saline (Thermo Fisher)

Glycerol (Sigma Aldrich)

2.11.4 Cell Media

Gibco™ DMEM, high glucose (Thermo Fisher)

Opti-MEM Reduced Serum Medium (Thermo Fisher)

Minimum Essential Medium (Sigma Aldrich)

L-Glutamine 200 mM (Thermo Fisher)

PneumaCult™-Ex Medium, plus 50x Ex Basal Supplement (StemCell Technologies)

Hydrocortisone 200x Stock Solution (StemCell Technologies)

Heparin Solution (StemCell Technologies)

PneumaCult™-ALI Medium, plus 10x ALI Supplement and 100x Maintenance Supplement

(StemCell Technologies)

Fetal Bovine Serum (various sources depending on availability)

Penicillin-Streptomycin (Thermo Fisher)

2.11.5 Media Recipes

Media Base	Additives	Cells

Dulbecco's Modified Eagle Medium	10% fetal bovine serum, 1% penicillin-streptomycin	HEK293T
Minimum Essential Media	1% L-glutamine, 1% penicillin-streptomycin, 10% fetal bovine serum	16HBE
PneumaCult™-Ex Medium	10% 50x Ex Basal Supplement, 0.1% 200x hydrocortisone stock solution	BMI-1 transduced basal epithelium (submerged culture)
PneumaCult™-ALI Medium	• 500 mL of medium combined with 10% 10x ALI supplement and frozen at -20°C in 40 mL aliquots • One aliquot at a time thawed and supplemented with 1% 100x maintenance supplement, 0.05% 200x hydrocortisone solution, and 0.01% heparin solution	BMI-1 transduced basal epithelium (ALI culture)

Results: Characterisation of an undruggable cystic fibrosis variant cell line

Introduction

Recent years have provided promising advances in pharmaceutical interventions for CF. However, these drugs' mechanism of action relies on correction of misfolded protein product and gating properties; therefore, variants that result in truncated, misspliced, or no protein product are unrescuable by these means [53].

From the Living Airway Biobank (IRAS ID 125793), we acquired primary nasal epithelial cells from a patient homozygous for the R709X (2257C>T) variant, which introduces a stop codon in exon 14. This variant falls approximately midway through the CFTR amino acid sequence, within the regulatory (R) domain [293]. This variant, first reported in 1995, is one of many CFTR nonsense variants. A study of four CFTR nonsense variant alleles (including W1282X, which theoretically encodes a nearly full-length mRNA) found that all evaluated alleles underwent nonsense-mediated decay (NMD) that reduced protein product by 50-80%, and residual protein product was truncated [58].

This variant accounts for a relatively small proportion of CF alleles [294]; according to the UK CF Trust's 2021 Registry Annual Data Report, it accounts for only 0.2% of the UK CF population [295]. However, the sum of patients with undruggable variants account for approximately 10% of the CF population. Therefore, these cells are representative of a significant portion of patients for whom modulator drugs are not a viable treatment option. They are, therefore, a good candidate for development of a "one size fits all" CF treatment strategy. Before exploration of a potential treatment strategy, however, we sought to thoroughly characterise the cells' CF phenotypes.

We sought first to make the primary cells more amenable to long-term study. Primary basal airway cells lose capability to proliferate and differentiate after ~4 passages

[62]. Multiple methods exist for extending their lifespan, such as treatment with ROCK inhibitors and co-culture with feeder cells [63, 64]. However, these strategies come with the disadvantage of long-term modified culture conditions and co-culture with cells that are not subject of study. The human *BMI-1* oncogene has been implicated in self-renewal of mammalian stem cells. Overexpression of BMI-1 has been shown to lead to extended self-renewal in various stem cell types, including mesenchymal [138] and neuronal [296]. Lentiviral transduction of primary airway epithelial cells with *BMI-1* has been shown to increase longevity of primary airway epithelial cells for culture from ~4 to ~20 passages [66, 67]. Therefore, *BMI-1* transduction was performed as described in Methods section 2.2.2. Following transduction, cells were expanded and differentiated via ALI culture for functional analysis. The clinical hallmarks of CF in patients are decreased ion transport, viscous and dehydrated mucus, and decreased ASL height. We therefore sought to examine these characteristics in ALI-differentiated R709X cells.

Ussing analysis has been used for decades to characterise CF cells and has been called the "gold standard" methodology for *in vitro* ion transport studies [297]. In this technique, an ALI-differentiated pseudostratified epithelium is sealed between buffer-filled chambers such that activity of ion channel proteins is the only means for passage of ions between the respective chambers; differences in ion concentration from one side to the other creates voltage potential. Current is applied across the membrane until the voltage potential is cancelled out; the required "short circuit" current functions as a readout of ion channel activity. Drugs stimulating or inhibiting specific ion channels can be added to chamber buffers in order to quantify specific ion channels' activity. In the context of CF, it has been used extensively to assess phenotypic severity of various CF genotypes for prognostic purposes [298, 299] and trial efficacy of modulator therapies [300-302].

Comparison of Ussing-measured, drug-induced responses in cells of different CF genotypes is a standard method for assessing variant severity and treatment response. Drugs typically used include direct inhibitors and activators of CFTR (CFTRinh-172 and forskolin, respectively) as well as inhibitors of other ion channel proteins (amiloride for ENaC and ouabain for Na^+/K^- channels). We therefore sought to use Ussing analysis to characterise the functional impact of the R709X variant.
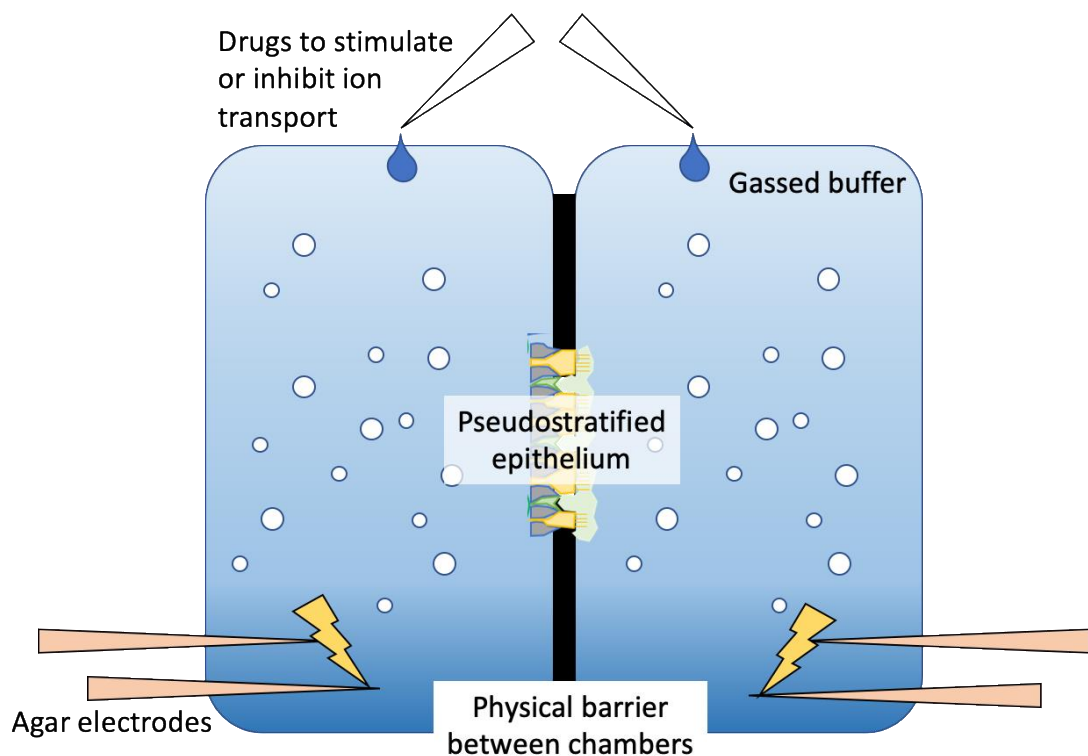
Fig. 3.1: Schematic of Ussing chamber setup. Two chambers (blue rounded rectangles) are filled with gassed buffer. The pseudostratified epithelium is sealed between the two chambers and acts as the only way that ions can pass from one chamber to the other. Agar electrodes (pink triangles) apply current (yellow lightning bolts) to cancel out transepithelial potential difference. Drugs (dark blue droplets) can be added to chambers to stimulate or inhibit specific ion channels.

Another functional readout of CFTR activity is altered ASL and mucus. Scanning ion conductance microscopy (SICM) is a useful tool for measuring microscopic topographical features of fragile biological materials. It relies on fluctuations in electric current based on proximity of an electrode to a surface coated in electrolyte-containing fluid. The methodology relies on a pair of electrodes. One electrode is housed in a saline-filled nanopipette (the "probe" electrode), which approaches the electrolyte-coated surface until it becomes immersed. Immersion creates a circuit between the probe electrode and a "bath" electrode, which is itself permanently immersed in the buffer bath in which the sample sits during analysis. The strength of current between bath and probe electrode then decreases as the pipette tip approaches a solid surface. Because ASL by nature contains electrolytes, SICM can be used to assess its depth over the underlying cell surface. In this technique, initial immersion of the probe triggers the feedback controls to cease its approach, and its position is recorded as the top of the ASL. The probe can then be prompted to continue its approach until it reaches the cell surface, at which point it automatically stops again; its position is again recorded, and the physical distance between these two positions is a readout of ASL depth. SICM has been previously performed on ALI cultures to assess differences in ASL depth of CF (homozygous F508del) and WT samples after *BMI-1* transduction and ALI differentiation, and it was found that CF ALI samples had significantly reduced ASL depth [286]. Additionally, as the probe retreats from the cell surface, mucus typically sticks to the pipette tip, generating a thread that stretches until it snaps. The distance this mucus thread can stretch before snapping is a functional readout of mucosal spinnability.
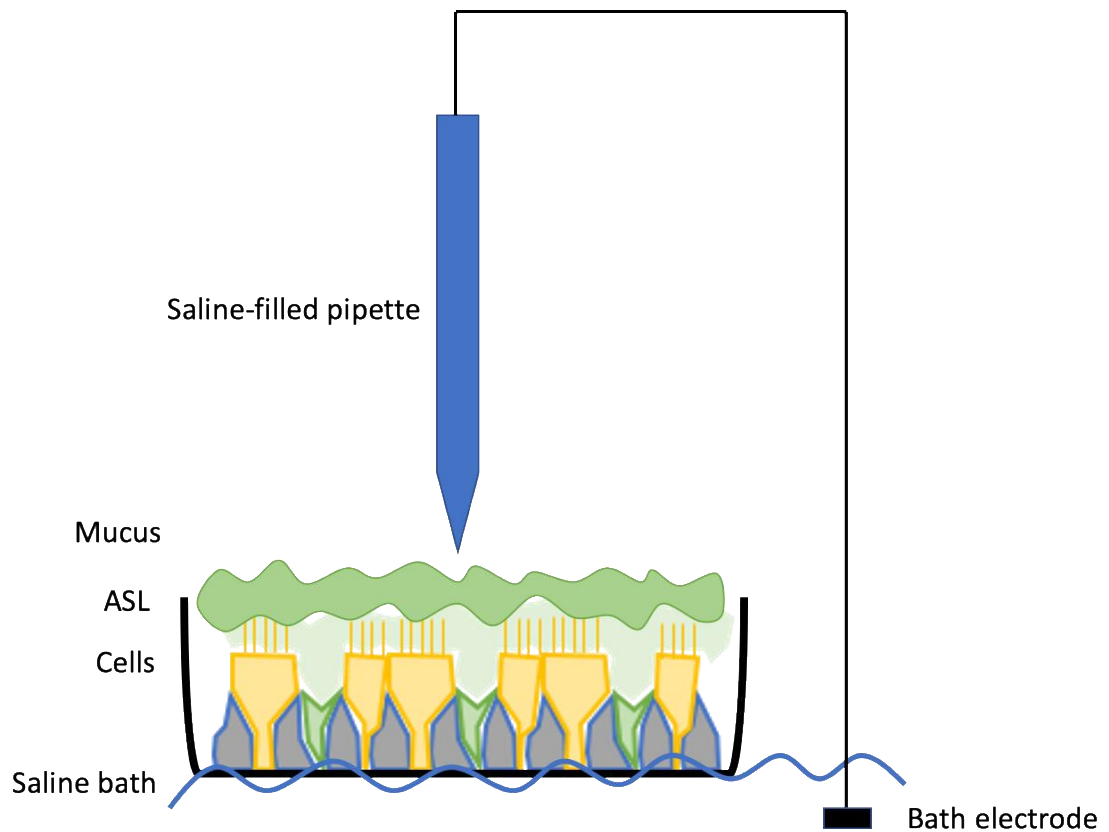
Fig. 3.2: Schematic of SICM. A saline-filled pipette containing a probe electrode (blue vertical shape) is suspended over a pseudostratified epithelial sample housed in a transwell (bottom structure). The transwell rests in a reservoir of saline containing a bath electrode (black rectangle).

By using assays that provide direct readouts of CF phenotypes, we sought to establish a baseline for treatment development in this undruggable cell model.

Aims

In this chapter we aim to establish a baseline for unedited CF variant cells via the following functional analyses:

1) Analyse CFTR-mediated ion trafficking in variant and WT cells by Ussing chamber, as defective ion trafficking is the main cause behind CF pathogenesis.

2) Analyse the airway surface liquid (ASL) depth of ALI cultures derived from CF variant and WT cells by SICM, as dehydrated ASL is one of the major hallmarks of CF lung disease.

Results

3.1.1 Genotypic validation of *CFTR* R709X variant

We first sought to validate the presence of the R709X variant in *BMI-1* transduced CF human nasal epithelial cells (hereafter referred to as CFNE BMI-1 R709X). Cells were expanded and DNA was isolated for Sanger sequencing of the relevant genetic locus following PCR amplification (Fig. 3.3). Sequence data was compared to canonical *CFTR* cDNA sequence according to NCBI (CCDS5773.1). Primer sequences used for PCR are detailed in Methods under "R709X Genotyping F" (also used for Sanger sequencing) and "R709X Genotyping R."
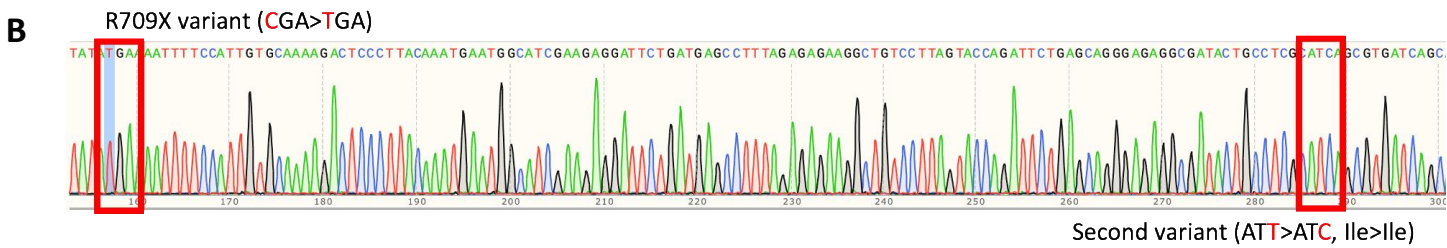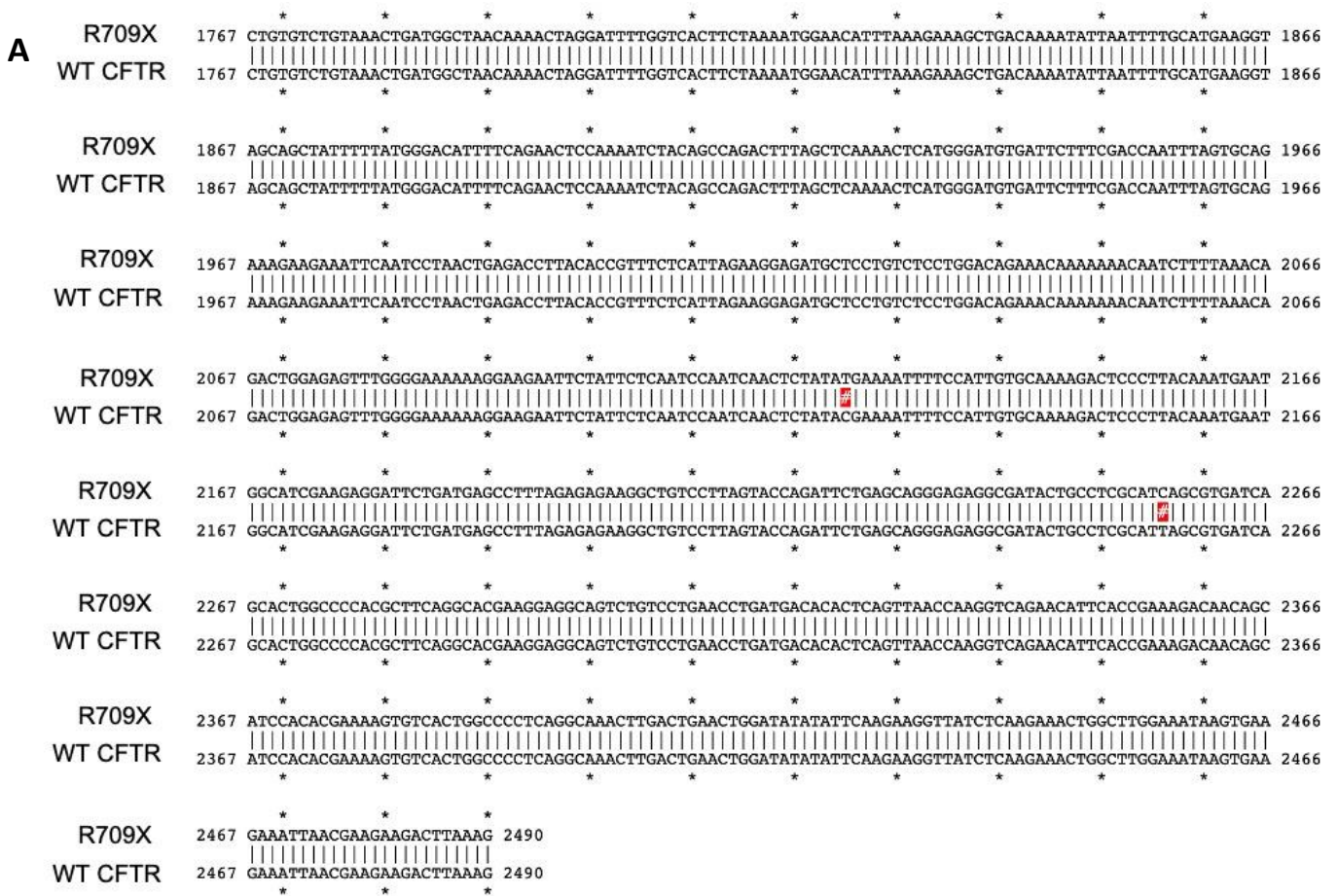
Fig. 3.3: Sanger sequencing confirmation of predicted variant R709X.

A) Sequence alignment of *CFTR* exon 14 with two mismatches highlighted. The region around the predicted variant was PCR amplified and Sanger sequenced.

B) Chromatogram of Sanger sequencing with variant codons highlighted.

The expected R709X-causing variant was present (CGA>TGA), but it was accompanied by a second variant downstream in exon 14. This second variant has not been reported in CFTR2 database; therefore, its potential phenotypic consequences are undocumented. However, the variant sequence (ATT>ATC) represents a silent mutation, as both canonical and variant codon encode isoleucine. Additionally, because this variant is downstream of the stop codon introduced by R709X, it would theoretically not be encoded into an amino acid. Therefore, it is reasonable to predict that this variant will carry no consequences at the protein level, and that any phenotypic differences in these cells is due principally to the R709X variant.

3.1.2 Ussing analysis of CFNE BMI-1 R709X ALI cultures

Ussing analysis was performed on ALI cultures derived from two cell populations: *BMI-1* transduced NHNE cells (NHNE BMI-1) and CFNE BMI-1 R709X (Fig. 3.4). Passage numbers of BMI-1 transduced cells take the format A.B, where A is the passage number at which transduction was performed, and B is the number of passages since transduction. ALI cultures were seeded from cells at passage 2.7 (CFNE BMI-1 R709X) or 2.9 (NHNE BMI-1) and permitted to grow for four weeks, at which point mucus was visible atop cultures and cilia movement visible under light microscope.

A Representative Ussing Traces

B Drug Responses
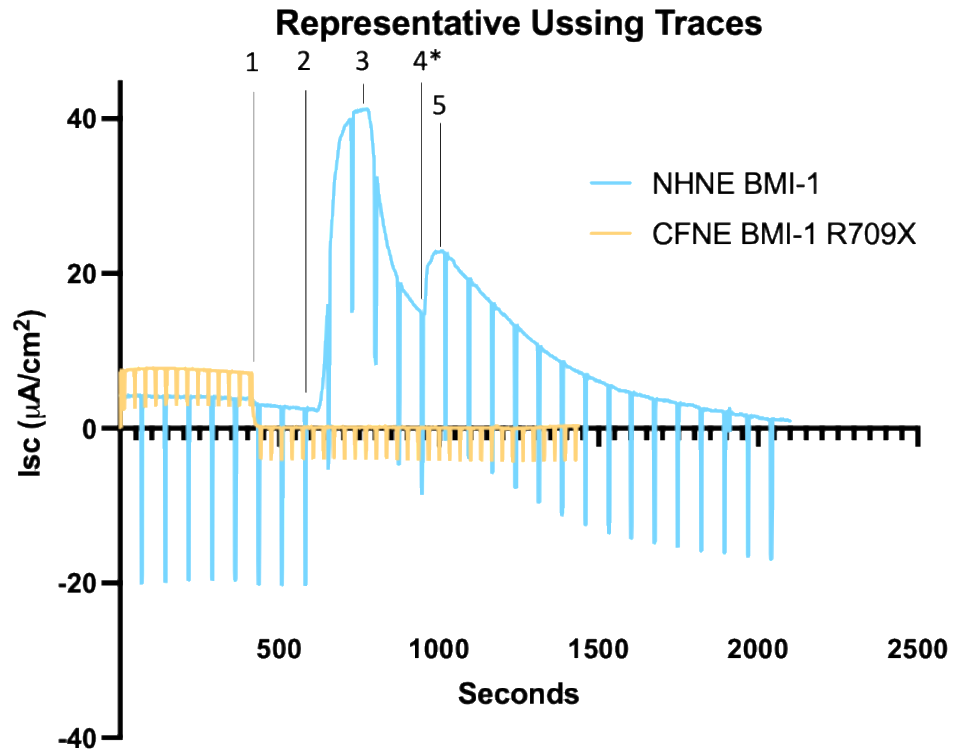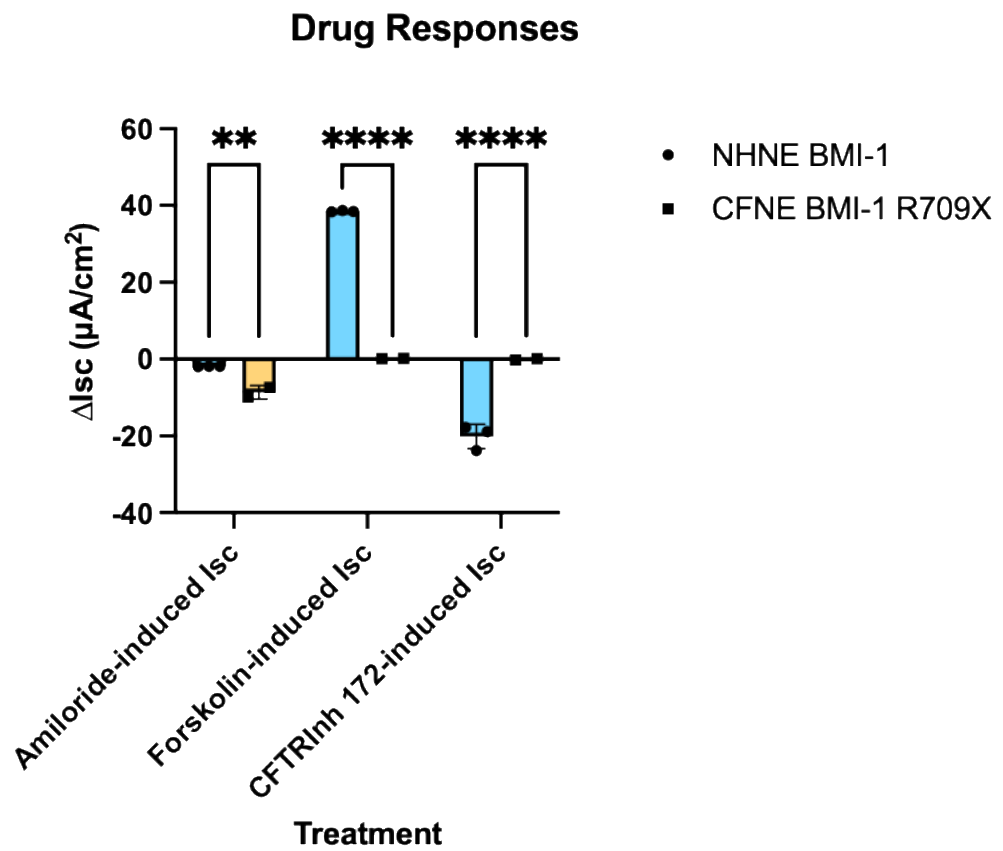
Fig. 3.4: Ussing chamber analysis of WT and CFNE BMI-1 R709X ALI cultures. n = 3 (NHNE BMI-1) or n = 2 (CFNE BMI-1 R709X) replicate ALI wells per genotype. Cells were differentiated on ALI for four weeks, at which point mucus and cilia movement were detectable. Ussing chamber analysis was run as described in Methods section 2.6.1.

A) Representative Ussing traces of NHNE *BMI-1* and CFNE *BMI-1* R709X ALI cultures. 1) amiloride, 2) forskolin, 3) CFTRinh-172, 4*) ATP (applied only to NHNE cells, not to CFNE cells), 5) ouabain. Traces were normalised to ouabain response by subtracting the lowest post-ouabain treatment short-circuit current (Isc) from all values for each transwell sample.

B) Graphical analysis of normalised changes in current after application of relevant drugs. Data displayed as mean with SD. Each dot represents one ALI well of indicated genotype. p-value calculated via two-way ANOVA with Tukey test. **** $p < 0.0001$; ** $p < 0.005$

The rationale for drug application was as follows:

1) Amiloride (10 µM on apical side) was used to inhibit ENaC for two reasons: firstly to examine how CFTR activity affected ENaC-mediated ion flow, and secondly to remove ENaC-controlled current as a confounding factor.

2) Forskolin (25 µM on apical and basolateral sides) was used to stimulate CFTR. The drug functions by increasing intracellular cAMP via stimulation of adenylyl cyclase, activating protein kinase A (PKA). PKA then phosphorylates the regulatory domain of CFTR, leading to its activation.

3) CFTRinh-172 (10 μM on apical side) inhibits CFTR by directly interacting with the protein and favouring the closed channel state.

4) ATP (100 μM on apical side) binds to purinergic receptors P2RY2 leading to intracellular Ca_2^+ mobilisation and hence activation of TMEM16A (ANO1) chloride channels.

5) Ouabain (100 μM on basolateral side) inhibits sodium and potassium ion transport. When drugs are applied in this order, the post-ouabain trace should represent net zero ion transport (hence its use in normalising the data above) [303-305].

NHNE BMI-1 cells showed negligible response to amiloride, indicating low ENaC activity (consistent with high CFTR activity, as CFTR activity is known to downregulate ENaC activity [91]), and dramatic responses to CFTR inhibition and activation. CFNE BMI-1 R709X cells, on the other hand, showed negligible response to CFTR inhibition or activation, and high response to ENaC stimulation (indicating a high baseline level of activity, consistent with low CFTR activity). This provides evidence that CFTR ion transport activity is high in ALI cultures derived from NHNE BMI-1 cells, but nonexistent in CFNE BMI-1 R709X cells.

3.1.3 Dehydration of ASL in CFNE BMI-1 R709X ALI cultures

After 28 days of differentiation on ALI culture, airway epithelial cells begin producing ASL. If left unwashed, the mucus layer of ASL rapidly accumulates on cultures until it becomes visible to the naked eye as a shiny textured layer. ALI cultures were left to differentiate for five weeks with no washing, and transwells were photographed to capture mucosal accumulation (Fig. 3.5).
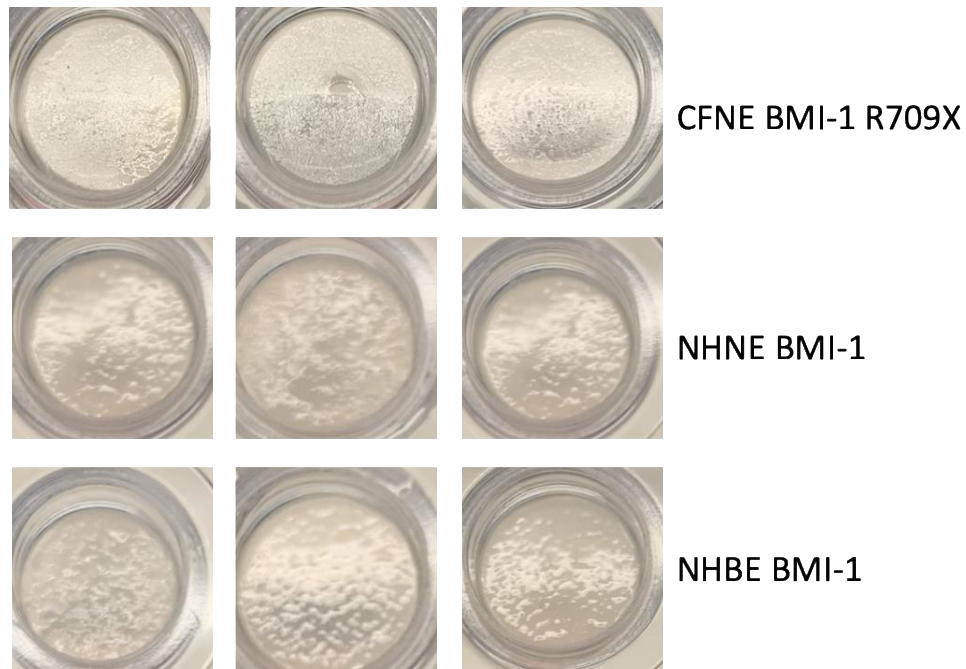
Fig. 3.5: Photographs of unwashed mucus-producing ALI cultures. Basal cells of indicated genotypes were allowed to differentiate on ALI for four weeks.

CF cells produced dry, grainy, low-luster mucus. WT cells produced mucus of higher moisture content with visible globules of shiny mucus. This recapitulates a hallmark pathogenicity of CF.

3.1.4 SICM analysis of CFNE BMI-1 R709X ALI cells

In order to quantify dehydration of ASL, SICM analysis was performed on ALI samples derived from three cell types: *BMI-1* transduced normal human bronchial epithelial (NHBE BMI-1) cells (passage 2.10), NHNE BMI-1 (passage 2.9), and CFNE BMI-1 R709X (passage 2.9). Cells were differentiated on ALI for 5 weeks (a process not primarily involving doubling of cells, but rather differentiation and self-renewal of basal cells); their mucosal production

is pictured above (Fig. 3.5), and cilia beating was detectable under light microscope. Cells were assayed for ASL depth and mucosal spinnability.

ASL Depth

Mucus Spinnability
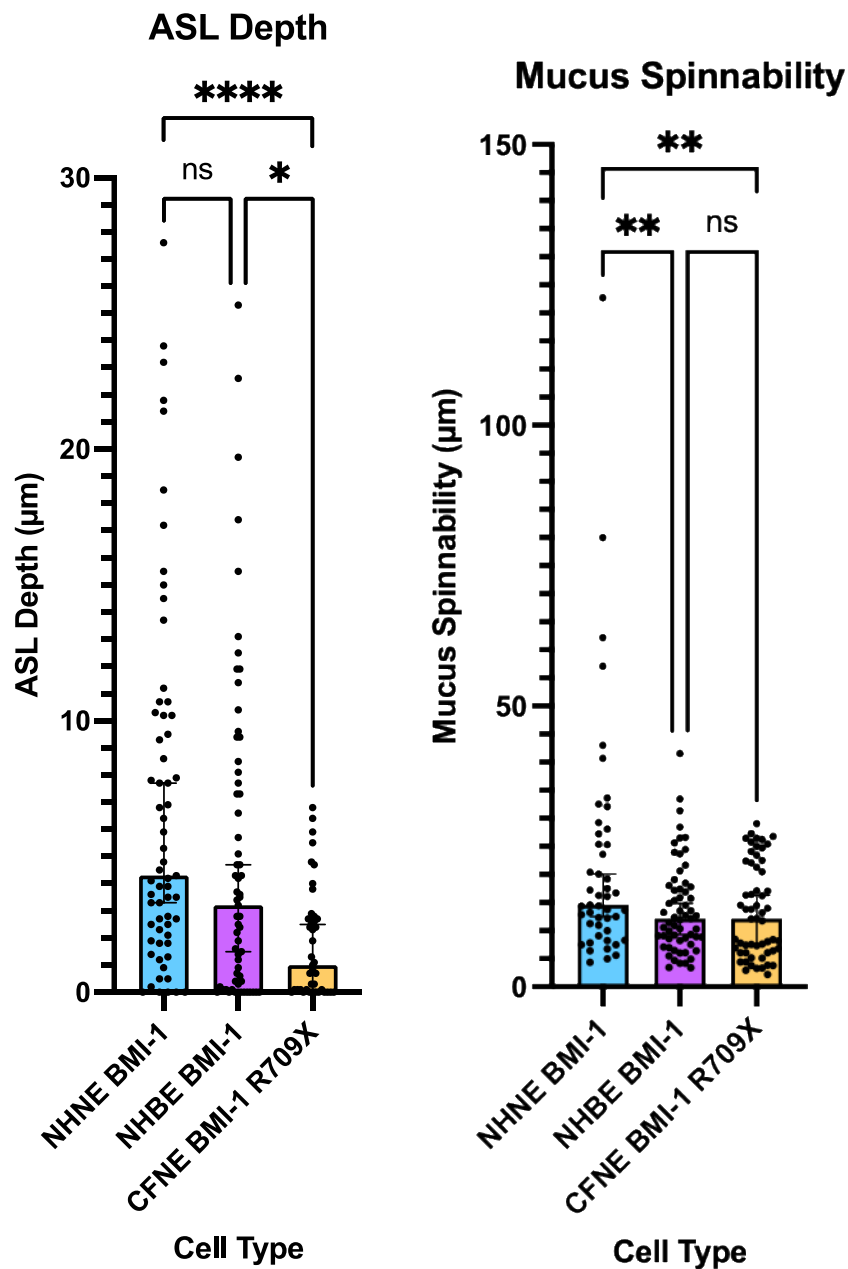
Fig. 3.6: SICM analysis of CF nasal, WT nasal, and WT bronchial epithelial cells. Plotted as median ± 95% CI. Each data point represents one measurement. For ASL depth, 21 measurements were performed for each of 3 wells of each genotype for a total of n = 63 per cell type; in the case of R709X cells, one well measured showed

abnormally high ASL height and was not plotted. For mucosal spinnability, 72 measurements were initially recorded, and any measurement that did not display detectable breakage of a mucosal thread was discarded. Ordinary one-way ANOVA with Tukey test. **** $p < 0.0005$; ** $p < 0.005$; * $p < 0.05$; ns $p > 0.05$

As predicted, CFNE BMI-1 R709X cells showed significantly reduced median ASL depth compared with both bronchial and nasal epithelial cells (1.0 μm for CF variant nasal, 3.2 μm for WT bronchial, 4.3 μm for WT nasal). One well of R709X cells showed abnormally high ASL height, falling outside the range of any wild-type cells (data in Appendix). This phenomenon has been observed before in ALI wells that have been infected with bacterial or fungal contaminants, typically a day before infection becomes visually apparent (cloudy media, visible contaminant, etc.) (unpublished). Because the wells assayed herein were not returned to ALI culture and maintained, it is unclear whether or not they were infected. However, given that these aberrant cells more closely matched data from infected wells than either of the wild-type genotype or the other R709X replicate wells assayed, the data was excluded from analysis. CF variant nasal cells also displayed significantly decreased mucosal spinnability compared with WT nasal cells (14.59 μm for WT, 12.13 μm for CF variant, $p < 0.005$), but not compared with WT bronchial epithelial cells (12.15 μm). This is consistent with CF pathogenicity, in which CF patient sputum exhibits high viscosity and low elasticity.

Discussion

In order to develop a treatment strategy, baseline phenotypic properties of *in vitro* disease models must first be established. Via ALI culture, *CFTR* variant basal cells can be differentiated into pseudostratified epithelia that exhibit properties analogous to human CF pathogenicity [56]. Two of the major hallmarks of CF are defective ion trafficking and subsequent defects in ASL regulation, which manifests as dehydrated mucus and decreased periciliary liquid layer height [306].

Measurement of ALI cultures' performance in Ussing chambers has been used for years as a predictor of *CFTR* variants' phenotypic severity, as well as CF treatments' efficacy, including recent breakthrough modulator therapies [300-302]. Loss of CFTR activity has been shown to lead to increased ENaC activity in CF patients, and ENaC inhibition has been discussed as a therapy for CF by both pharmaceuticals and nucleic acid therapy [96]. Therefore, we expected to see a noticeable difference between amiloride-sensitive current in variant vs. WT cells. Indeed, while NHNE BMI-1 cells showed negligible response to amiloride (ΔIsc = 1.915 μA/cm^2), CFNE BMI-1 R709X variant cells showed significantly higher response (ΔIsc = -8.64 μA/cm^2, *p* < 0.005). Additionally, as expected, variant cells showed no detectable response to drugs either stimulating or inhibiting CFTR activity. This is consistent with lack of activity or complete absence of protein at the cell surface due to premature stop codons in *CFTR* [58, 307, 308]. In future studies, following advice from Dr. Luis Galietta, ATP will be included for both genotypes. Additionally, ouabain will be added after the effect of ATP has depleted. In the presented timing of drug applications, it is unclear how much of Isc decay is due to ATP effect rundown, and how much is due to ouabain inhibition; however, this does not undermine the data suggesting that CFTR is inactive in these cells.

Unlike Ussing analysis, which has a robust history in the field of CF research, SICM is a relatively novel technique. However, technologies with similar aims (such as confocal microscopy) have long been used as a readout of phenotypes resulting from *CFTR* variants [306]. Unlike other microscopy techniques, however, SICM also allows for concurrent measurement of other CF-relevant phenotypes including mucosal spinnability. Because increased mucosal viscosity is one of the main underlying causes of lung disease in CF, SICM is a valuable tool for simultaneous multifaceted characterisation of differentiated cells' ASL. The mean ASL depth observed in SICM analysis of NHBE BMI-1 (5.1 μm) was comparable with those previously reported in the literature (~5.8 μm measured via SICM) [286]. The mean ASL depth observed in CFNE BMI-1 R709X cells (1.8 μm) was lower than that previously reported for CF variant cells (~3 μm for CFBE BMI-1 F508del cells measured via SICM) [286]. This could have been due to differences in either cell type (nasal vs. bronchial) or genotype (missense vs. nonsense variant). This could of course be elucidated by comparing different cell types carrying the same variant, or different variants in the same cell types; however, this is restricted by availability of patient cells carrying specific genotypes.

Baseline characterisation of disease model cells is an essential first step in developing a treatment strategy. In our hands, basal cells carrying the undruggable R709X variant (as well as a previously unreported silent variant) exhibit multiple phenotypes consistent with CF pathology, and therefore are a strong model of CF disease. The following chapter therefore investigates a proof of concept DNA editing strategy that can be used to insert a full-length *CFTR* cDNA into a safe harbour of the genome.

Results: Proof of concept of homology independent targeted integration at the AAVS1 *safe harbour locus*

4.1　Introduction

Recessive monogenic diseases can potentially be rescued by introduction of a healthy copy of the causative gene through direct modification of the genome. Genomic insertion of large DNA constructs via CRISPR can be accomplished via homology-directed repair (HDR) or non-homologous end joining (NHEJ), as discussed in Introduction section 1.2.2.2. In non-dividing cells, such as those making up the airway epithelium, HDR does not occur [241]. Therefore, an NHEJ-based approach is necessary to integrate DNA in this target tissue.

Because NHEJ-mediated repair is not dictated by homology, there is the potential for donor constructs to integrate in the reverse of the desired direction. Homology-independent targeted integration (HITI) prevents this. In HITI, a donor construct is used that includes a reverse complemented gRNA sequence (RC-gRNA) on either side of the desired insertion (see Fig. 1.9) [152]. If the donor integrates in the reverse direction, these RC-gRNA sequences restore the spacer sequence, allowing Cas9 to re-bind and re-cut DNA. If the donor integrates in the forward direction, however, the spacer sequence is destroyed, and Cas9 cannot re-target. This has been shown to enhance rates of desired integration in various cell types, including non-dividing cells [152].

Another major consideration for introduction of compensatory gene copies is integration site of the genome. Adeno-associated virus (AAV), as its name suggests, requires coinfection with adenovirus to successfully propagate in cells. In the absence of adenovirus, AAV integrates into the host genome at a section of euchromatin on chromosome 19, where it remains latent until adenovirus becomes available [309]. Despite the fact that this site lies within a protein-coding gene (*PPP1R12C*), no cellular pathogenic effects have been reported

after integration of the viral genome at this locus. Therefore, this site is regarded as a "safe harbour" for integration of large DNA constructs into the genome, and has been used for integration and stable expression of transgenes using multiple DNA editing approaches including TALENs [310], ZFNs [311-313], and CRISPR [314, 315]. Kelly *et al.* utilised HITI to insert fluorescent reporter constructs at the AAVS1 locus with higher efficiency than HDR approaches. Edited cells isolated via FACS were single-cell cloned and clones assayed for editing; 11.8% of clones were positive using HDR, compared to 36.1% using HITI [315].

We aimed to target the *AAVS1* site over the endogenous *CFTR* locus for two primary reasons. Firstly, the native gene at the *AAVS1* site, Protein Phosphatase 1 Regulatory Subunit 12C (*PPP1R12C*), has been characterised as a robustly and constitutively expressed gene (Fig. 4.1). Specifically, in respiratory epithelial (both bronchial and nasopharyngeal) cells, expression of Protein Phosphatase 1 Regulatory Subunit 12C (PPP1R12C) is predicted to be even higher than that of CFTR [316, 317]. Single-cell RNA sequencing of the airway epithelium revealed homogenous expression of PPP1R12C in basal cells, goblet cells, ciliated cells, and ionocytes (in which CFTR expression is highest) [70].
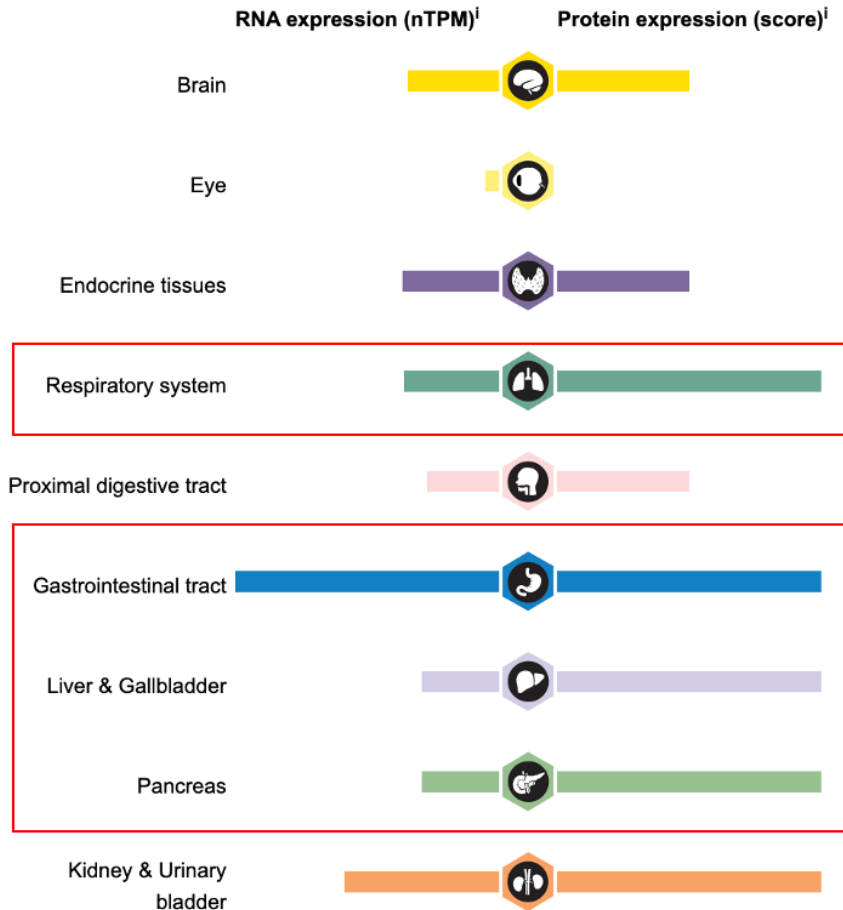
Fig. 4.1: Protein expression of *PPP1R12C* in CF relevant tissues (red boxes). nTPM: normalised transcripts per million. Protein expression score is based on "immunohistochemical data manually scored with regard to staining intensity (negative, weak, moderate or strong) and fraction of stained cells (<25%, 25-75% or >75%)." Tissues in red boxes are relevant to CF pathology, and have received an expression score of "high," indicating "strong [signal] combined with either 25-75% [fraction of positive cells], or 75% [fraction of positive cells]." [316]

It is unreasonable to expect that every cell will be successfully transfected and edited in an *in vivo* strategy. Therefore, elevated expression of a transgene by a subpopulation of

cells could compensate for <100% transfection and editing efficiency. There are of course some genes for which this strategy would not work (i.e. genes whose protein product functions intracellularly). However, because CFTR's role is to contribute ions to the general extracellular environment, an overexpressing cell could theoretically compensate for its neighbours' lack of activity. Multiple studies have examined the proportion of CFTR-expressing cells required to see amelioration of CF pathology, based on varying approaches and criteria. An early study used Ussing analysis to measure ion transport in mixtures of corrected (via retroviral cDNA delivery) and uncorrected F508del homozygous cells; it was found that 10% corrected cells was sufficient to restore ion trafficking to WT levels [318]. Another study, using the RNA virus parainfluenza to deliver CFTR, found that successful expression in 25% of an ALI-differentiated epithelium was sufficient to restore normal ASL height and mucociliary clearance in F508del homozygous cells [319]. Other studies have focused not on the proportion of CFTR-expressing cells, but rather on the proportion of correctly spliced *CFTR* transcripts. An early study on a splicing variant (the 5T allele in intron 8) suggested that >25% of transcripts must be correctly spliced in order for patients to achieve FEV1 comparable to non-CF levels [320]. Another study on the same splicing variant estimated that ~30% normal transcripts required for normal lung function [321]. In clinical trials for the triple-combination therapy Trikafta, *in vitro* study revealed that Trikafta treatment rescued chloride ion transport by approximately sixfold, which translated to 10-15% increase in FEV_1 relative to non-CF individuals (variable depending on genotype) [322]. Overall, while it is difficult to know exactly how much of WT CFTR is required (in terms of the percent of cells expressing protein, the percentage of transcripts rescued by editing, and the clinical criteria being investigated), most estimates have fallen between 10 and 30%.

Another reason to target the AAVS1 locus was because we did not wish to target the endogenous *CFTR* locus. Despite HITI's co-opting of the NHEJ pathway for high integration efficiency, there remains the risk of integration in the reverse direction, or indel formation. Constructs can still integrate in the reverse orientation and not be excised, or resected DNA can be re-annealed before integration can occur, albeit at reduced rate compared to non-HITI NHEJ. Therefore, unsuccessful editing events can present a threat to the targeted locus.

There are many *CFTR* genotypes which retain some phenotypic function (which can in some cases be enhanced by modulator therapies) [323]. An unsuccessful edit and resulting indel at the endogenous locus could potentially lead to loss of any residual function. This is especially risky if the strategy involves integration of a promoterless construct immediately downstream of the native promoter, as indel formation in a promoter or early in a gene could potentially generate a full knockout. Despite this, a full gene replacement strategy is still an attractive option, as it represents a single-administration treatment for theoretical full rescue.
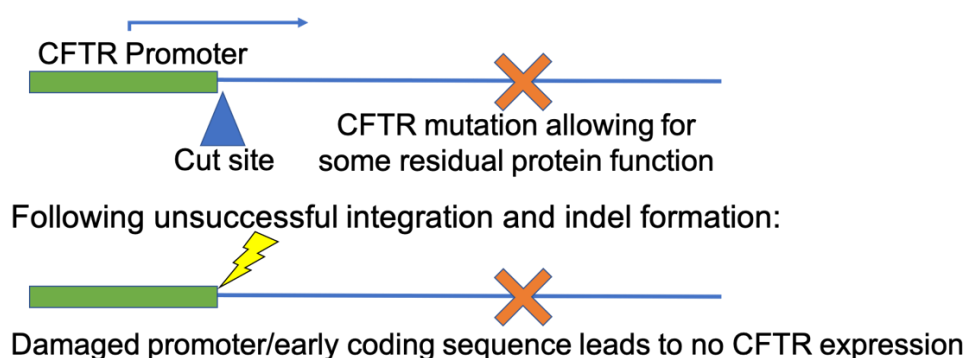
Fig. 4.2: Rationale for not targeting the *CFTR* locus. *CFTR* promoter shown as green box; gene expression shown as blue arrow; cut site shown as blue triangle; *CFTR* mutation shown as red X; Cas9-induced DNA damage shown as lightning bolt.

CFTR superexons (either complete or incomplete) have been utilised by several groups as a "one size fits most" approach [235, 239, 240]. Unlike the current small molecule therapies available, or any DNA editing strategy that corrects individual mutations, this approach could provide rescue regardless of patient genotype. Prior to this strategy, however, validation of the HITI strategy and activity of the native *PPP1R12C* promoter is vital. Integration of fluorescent reporters and subsequent analysis of expression is a valuable proof of concept for further uses of the HITI/AAVS1 strategy, such as integration of cDNA for clinically relevant genes. Therefore, in this chapter, we aim to characterise the on- and off-target editing efficiency of HITI at the AAVS1 locus using the fluorescent reporter GFP.

4.2 Aims

In this chapter we aim to:

1) Develop a HITI-based strategy for inserting large coding sequences at the AAVS1 safe harbour locus.

2) Insert a reporter construct in HEK293T cells using this strategy.

3) Assess predicted off-target editing events in a clonally isolated population of edited cells, and a mixed population of repeatedly edited cells.

4.3 Results

4.3.2 Generation of edited cells

4.3.2.1 Generation of HITI donors

The HITI approach requires two main components: the editing machinery (*Sp*Cas9 + gRNA) and donor construct for integration. For our purposes, we used gRNA A78 (sequence in Methods). This gRNA was initially designed by previous Hart lab member Dr. Ileana Guerrini, and was later identified in an independent study as the most efficient of six gRNAs tested at the AAVS1 locus (measured by Universal CRISPR Activity assay) [292]. Two donor constructs were tested. The first, called the "pLinear" donor, contained the following features (Fig. 4.3):

1) One copy of the reverse-complemented A78 gRNA target site and adjacent PAM, located just upstream of the donor construct. Upon treatment with Cas9, plasmid becomes linearised.

2) A donor construct consisting of eGFP, preceded by a splice acceptor site and T2A site. The splice acceptor allows GFP to be placed under the control of the endogenous *PPP1R12C* promoter, eliminating the need for a promoter in the plasmid itself. The T2A site (derived from *Thosea asigna* virus) prompts ribosomal machinery to split the GFP peptide chain from the PPP1R12C chain, preventing generation of a fusion protein at the N-terminal end.

3) A stop codon, SV40-derived termination sequence, and poly-A tail, prompting termination of translation after the GFP coding sequence and preventing generation of a fusion protein at the C-terminal end.

4) An unrelated gRNA sequence (T7) – this is an artefact left over from a previous round of cloning, and not targeted by any gRNA used in the HITI experiments described herein.

5) Kanamycin resistance construct used in cloning and transformation.

A
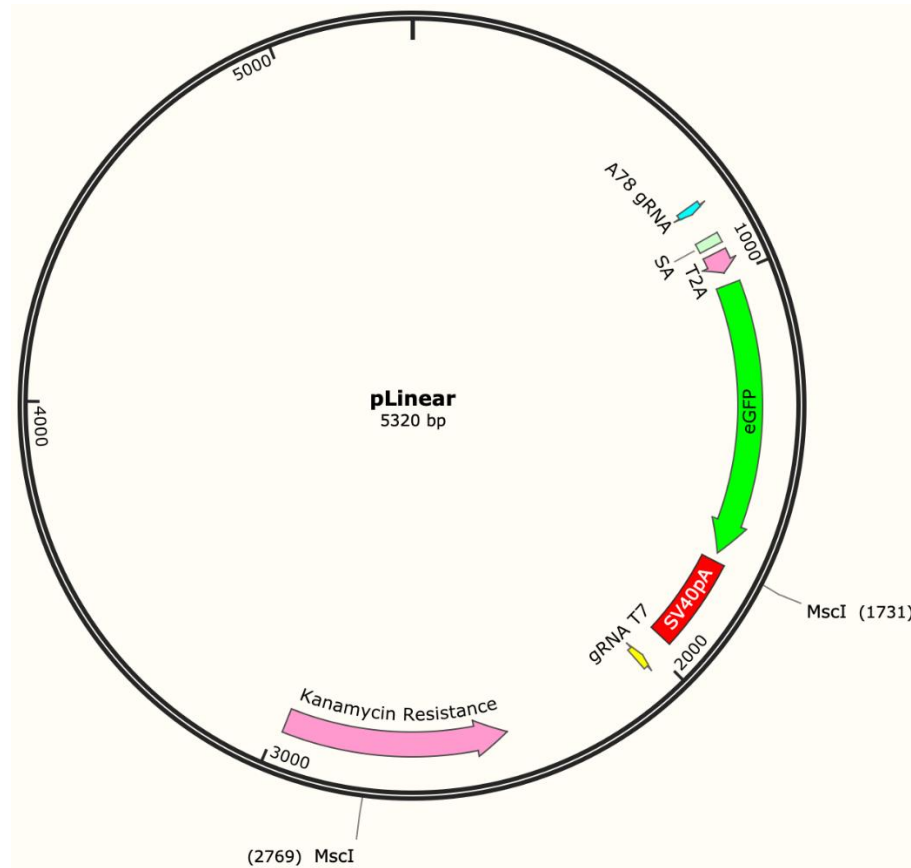
B

1. Cleavage of genome and A78-T7 donor by Cas9

Chr19 — Intron 1 — PPP1R12C — A78 gRNA

A78-T7 — RC-A78 gRNA — RC-T7 gRNA

2. Genomic integration of cleaved donor

SA-T2A-GFP-pA — backbone (~5.5 kb)

Chr19
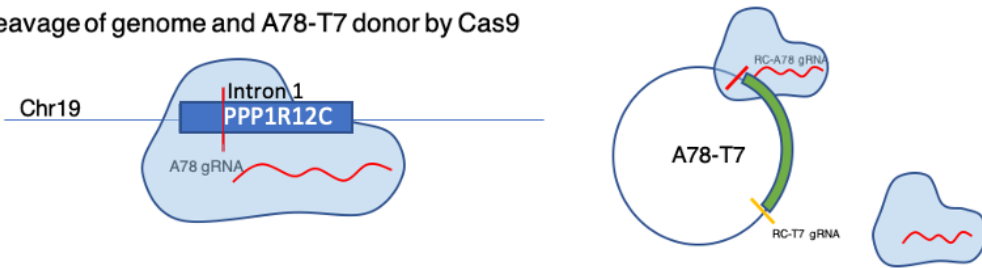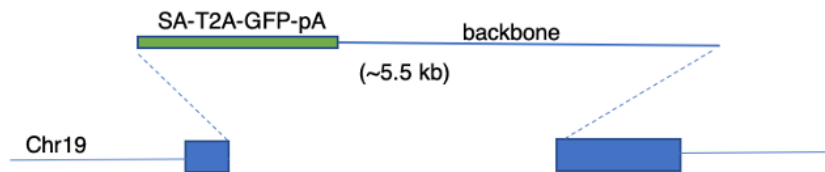
Fig. 4.3: Schematics of HITI approach with pLinear donor.

A) Plasmid map of pLinear donor. A78 gRNA: reverse-complemented spacer sequence and PAM for A78 guide. SA: splice acceptor. T2A: signals ribosome to split peptide chain. eGFP: eGFP cDNA. SV40pA: poly-A tail from simian virus. gRNA T7: reverse-complemented spacer sequence and PAM for an unrelated gRNA.

B) Schematic of donor excision and integration upon treatment of donor and genome with Cas9. AAVS1 target locus shown as blue box; Cas9 shown as translucent blue shape; gRNA shown as red line; section of plasmid encoding SA-T2A-GFP-polyA construct shown as green box. This plasmid was generated by Dr. Ileana Guerrini, from a plasmid originally sourced from System Biosciences (catalogue no. MN511-A1).

However, genomic integration of plasmid backbone DNA, including antibiotic resistance genes, is not favourable for translation into the clinic. We therefore sought to

create a "pExcised" donor plasmid, which had the desired insert flanked by two A78 gRNA

cut sites. Treatment with Cas9 would therefore excise the GFP cDNA from the backbone.

Another plasmid had previously been generated (also by Dr. Guerrini) that had two

copies of the RC-A78 sequence appropriately flanking the eGFP construct.
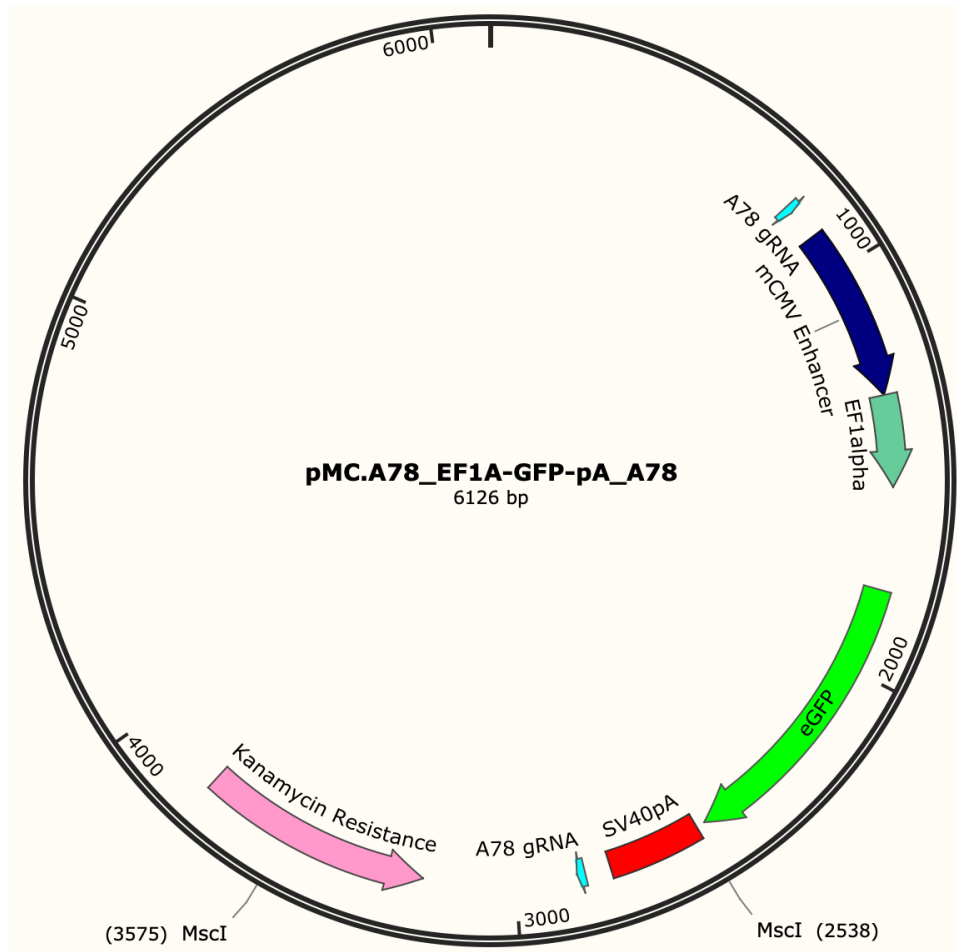
Fig. 4.4: Plasmid map of pMC.A78_EF1A-GFP-pA_A78, one of the parental

plasmids used alongside pLinear in order to generate the second HITI donor

called "pExcised." "A78" refers to the reverse-complement of gRNA sequence

and includes adjacent PAM.

However, this plasmid also contained a CMV enhancer and EF1A promoter sequence, and lacked the SA and T2A features (Fig. 4.4). We wished to remove the enhancer and promoter from this parental plasmid. The reasons for this were twofold. Firstly, we wanted the donor construct to be small enough to potentially be packaged in an AAV vector in downstream applications. Full-length CFTR cDNA is ~4.5 kb, which itself nearly meets the packaging capacity of an AAV vector. The HITI donor would also have to include a splice acceptor, T2A sequence, and poly-A tail in order to generate functional mRNA; inclusion of these features would not be possible in an AAV vector alongside a promoter. Secondly, we wished to confirm that promoterless constructs could be reliably expressed under the native *PPP1R12C* promoter. This would also necessitate the inclusion of the SA and T2A sequences.

By digesting the two plasmids with the restriction enzyme MscI (highlighted in Fig. 4.3 and 4.4), it was possible to isolate the sections of each plasmid that contained the desired features and ligate together as described in Methods section 2.5.7. The resulting plasmid, deemed "pExcised," is detailed in Fig. 4.5B, and contains the following features:

1. Reverse-complemented A78 gRNA sequence

2. SA and T2A   From pLinear plasmid

3. GFP cDNA

4. SV40 poly-A tail   From pMC.A78_EF1A-GFP-pA_A78

5. Reverse-complemented A78 gRNA sequence

Each plasmid was digested with MscI and run on an agarose gel. The desired bands were excised from the gel, DNA purified, and fragments ligated together (as described in Methods section 2.5.7). In order to ensure that the insert integrated in the proper

orientation, sequencing was performed on single colonies (full plasmid sequence in

Appendix); a colony with the expected sequence was then grown up overnight and saved as

a glycerol stock as described in Methods. The features listed above were confirmed in the

resultant plasmid via sequencing (full plasmid sequence in Appendix with features

highlighted) (Fig. 4.5).
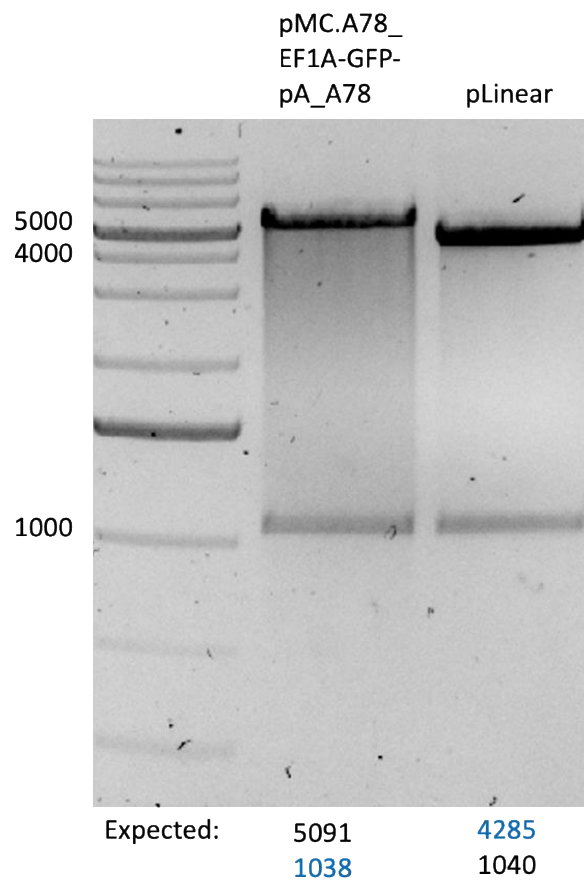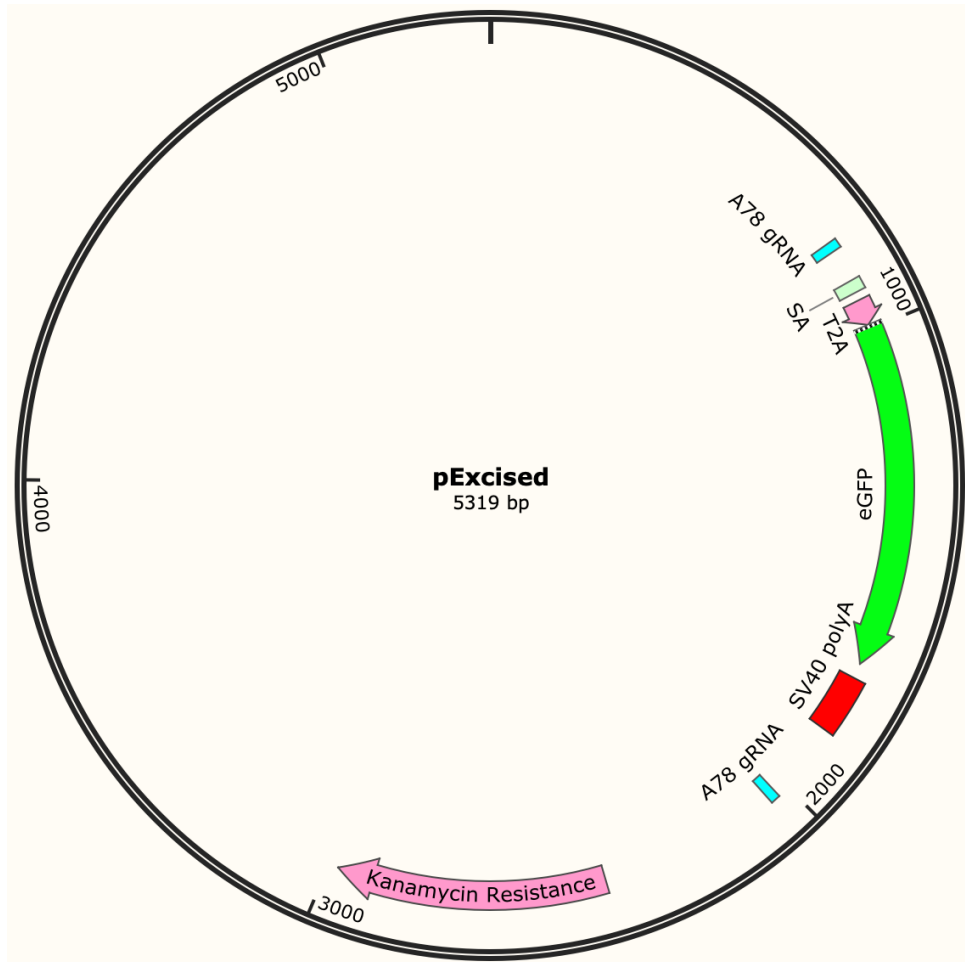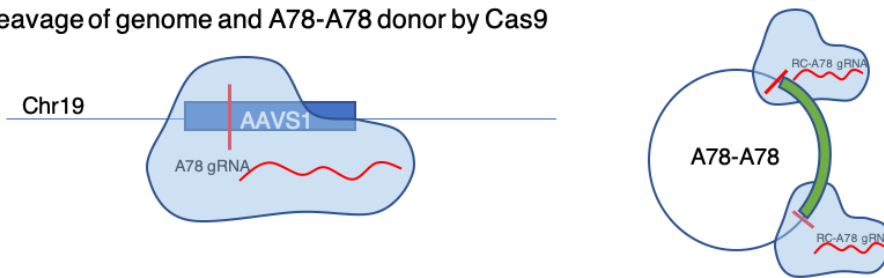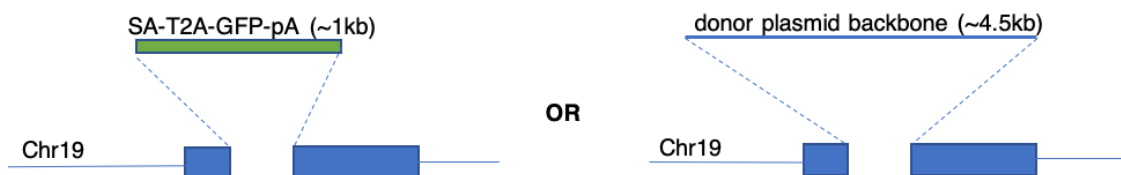
Fig. 4.5: Schematics of HITI approach with pExcised donor.

A) Gel showing two parental plasmids after digestion with MscI. Expected band sizes indicated. The desired bands (highlighted in blue) were excised from the gel and ligated together as described in Methods.

B) Plasmid map of pExcised donor with dual A78 gRNA sequences (full sequence available in Appendix). "A78" refers to the reverse-complement of gRNA sequence and includes adjacent PAM.

C) Schematic of donor excision and integration upon treatment of donor and genome with Cas9. Upon treatment with Cas9, both genome and donor will be cut; unlike pLinear, the pExcised donor will be cut twice, thus completely excising donor construct from backbone.

The final plasmid used in initial HITI experiments contained an expression cassette for the A78 gRNA and tracrRNA as well as Cas9, henceforth referred to as pCas9-BFP. This plasmid's features are illustrated in Fig. 4.6.
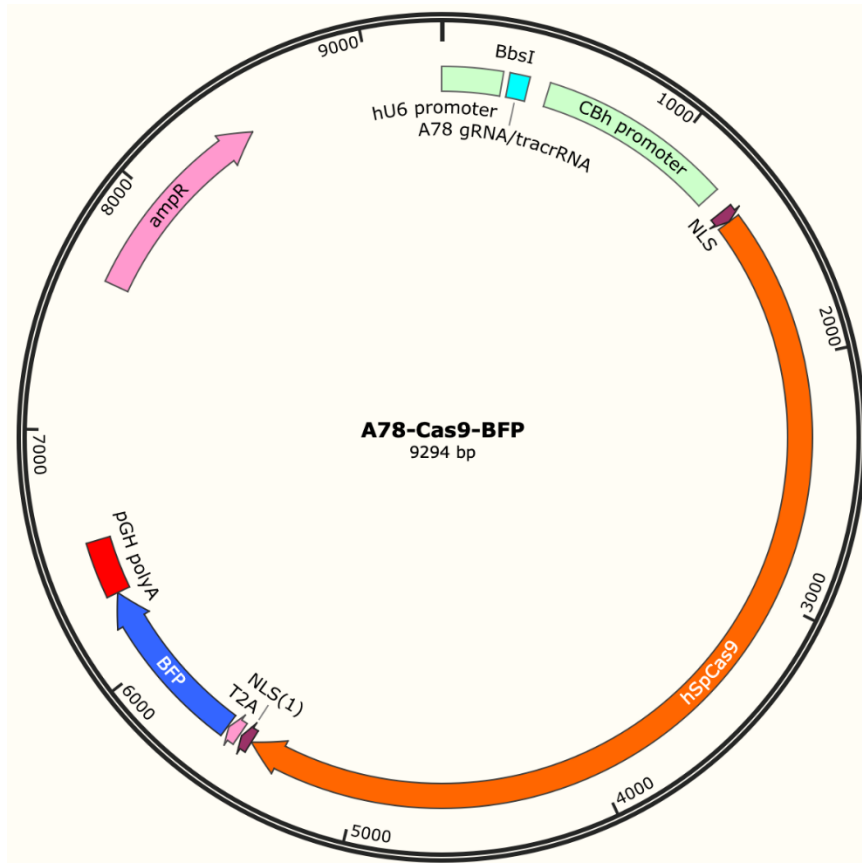
Fig. 4.6: Plasmid map of pCas9-BFP. The plasmid contains the following features:

gRNA (A78 crRNA and tracrRNA) expression construct under the hU6 promoter (a

human-derived promoter commonly used to drive expression of small RNAs);

*Sp*Cas9 under the CBh promoter (derived from chicken B-actin promoter,

modified with a murine virus of mice hybrid intron); a nuclear localisation

sequence (NLS) to direct expressed Cas9 to intranuclear space; a T2A sequence

to prevent generation of a Cas9-BFP fusion protein; blue fluorescent protein

(BFP), functioning as a reporter of Cas9 expression; BGH termination sequence

and poly-A tail; ampicillin resistance construct used in cloning and

transformation.

When delivered together, pCas9-BFP and either of the donor plasmids would have the potential to result in the AAVS1 target site being cut, and donor construct integrating via NHEJ.

4.3.2.2   Delivery of pCas9-BFP and Donor plasmids

In order for HITI to take place, Cas9 must be expressed and donor must be present in the same cell. We therefore initially wished to test whether transfection of pCas9-BFP and pLinear plasmids could be achieved by co-delivery in the same nanoparticle (NP). Initial experiments used the commercially available reagent Lipofectamine 2000 (L2K) for delivery. Work by previous Hart lab member Dr. Amy Walker revealed that the optimal transfection amount of donor and Cas9 plasmid in 24-well plates was 250 ng donor:500 ng Cas9 [324]. Therefore, these amounts were delivered to 70-80% confluent HEK293T cells in 24 well plates, in either the same or separate L2K NPs. L2K NPs were formulated at a 3:1 weight ratio of L2K:DNA. Forty-eight hours post-transfection, flow cytometry was performed in order to quantify BFP and GFP expression as the percentage of GFP+ and/or BFP+ live singlet cells in each condition (Fig. 4.7).

A

pCas9-BFP + pLinear, same nanoparticle

BFP(19.51%)

GFP-BFP(15.03%)

GFP(0.15%)

pCas9-BFP + pLinear, separate nanoparticles

BFP(25.44%)

GFP-BFP(17.87%)

GFP(0.45%)

Cas9 Expression

B

** (significance)

% of BFP+ Live Singlets

pCas9-BFP + pLinear Same NP
pCas9-BFP + pLinear Separate NP
pCas9-BFP Only
pLinear Only
eGFP
UT
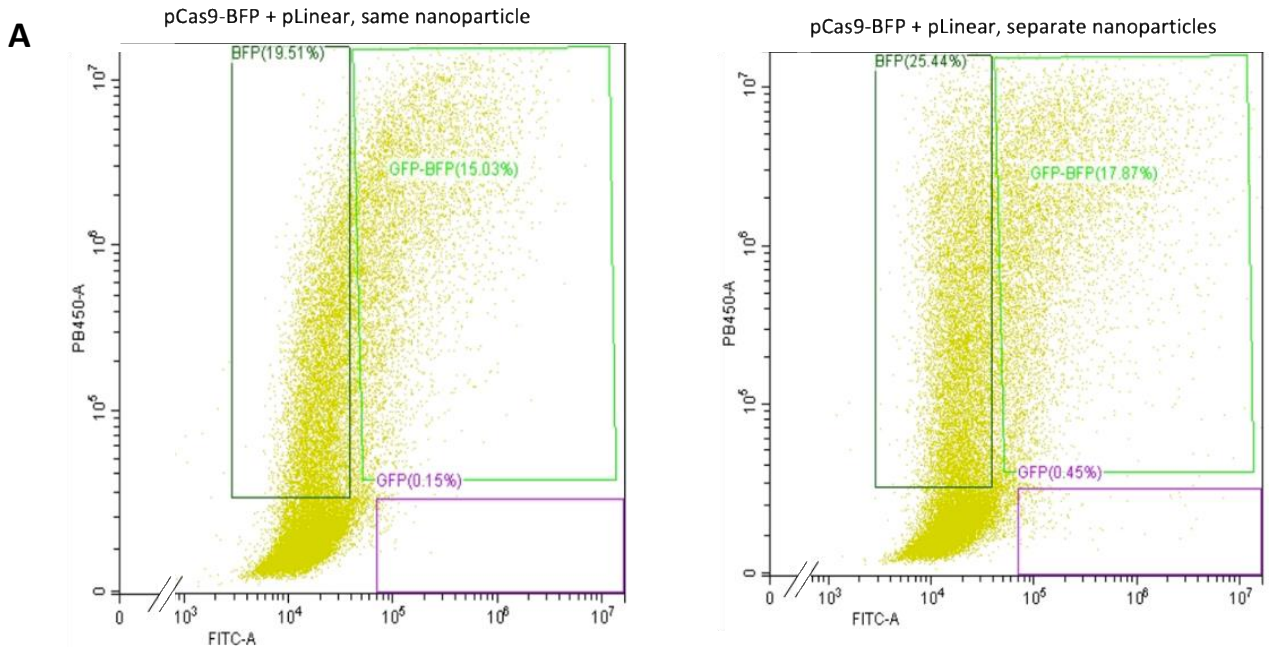
Donor Expression
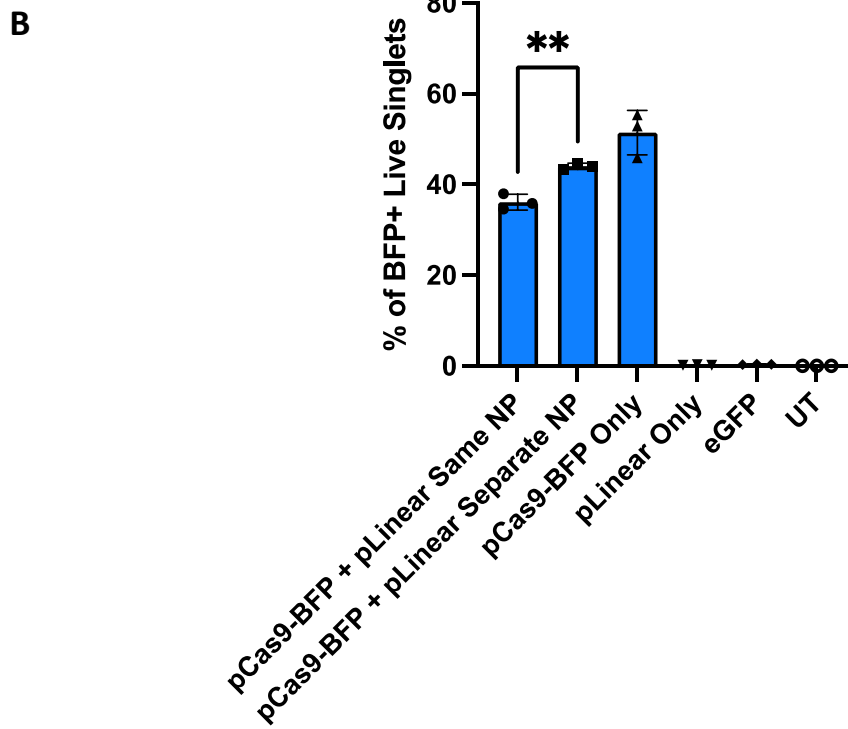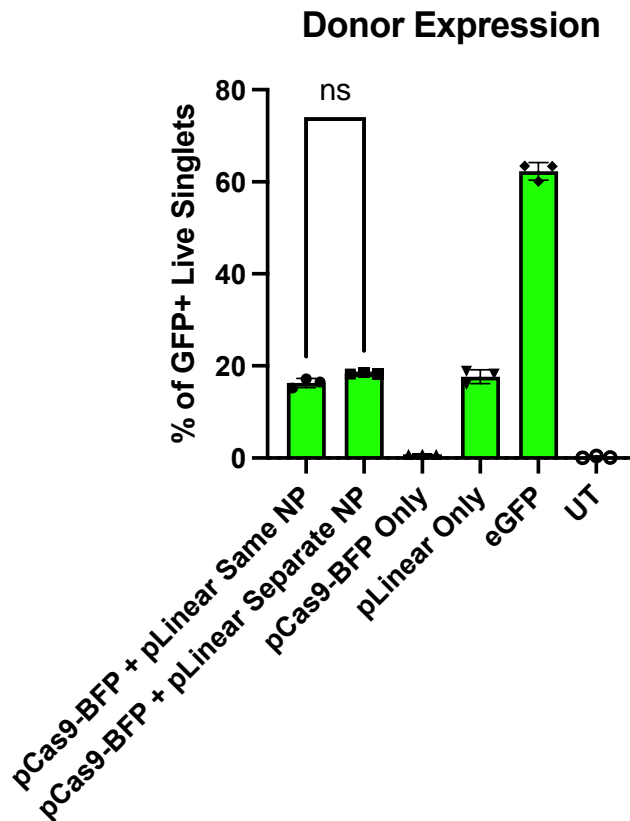
Fig. 4.7: Flow cytometric quantification of transfection efficiency. HEK293T cells were transfected with pCas9-BFP and pLinear donor using Lipofectamine 2000. 48 hours later, flow cytometry was performed. Gates were set for live singlets, and then for GFP/BFP expression based on negative (untransfected HEK293T) control population (details of gating in Appendix). ≥30,000 live singlets were sorted for each condition shown.

A) Representative flow plots for pCas9-BFP and pLinear donor plasmids delivered in the same or separate nanoparticle (one replicate transfected well shown per condition).

B) Quantified transfection efficiency of pCas9-BFP and pLinear donor plasmids. BFP expression was taken as a readout of Cas9 expression, due to both being under the same promoter within the plasmid. n = 3

replicate transfected wells. Plotted as mean $\pm$ SD. One-way ANOVA with Tukey multiple comparisons test. ** $p < 0.01$

A statistically significant difference was observed in total BFP (and therefore Cas9) expression between cells receiving pCas9-BFP in the same vs. separate NPs, with co-delivery in the same NP achieving approximately 8% higher Cas9 expression. No significant difference was observed in total GFP expression between cells receiving pLinear in the same vs. separate NPs (Fig. 4.7B).

Formation of double strand breaks (DSBs) without an available donor for repair will most likely lead to indels. Therefore, it is generally unfavourable to have Cas9 expression in cells that have not also received a donor construct. In order to ascertain the percentage of cells expressing pCas9-BFP or pLinear, but not both, 48 hours post-transfection, we examined the rates of singly fluorescent cells with each treatment by examining the percentage of cells falling within BFP+/GFP- and BFP-/GFP+ gates (shown in Fig. 4.7).

Singly Fluorescent Cells
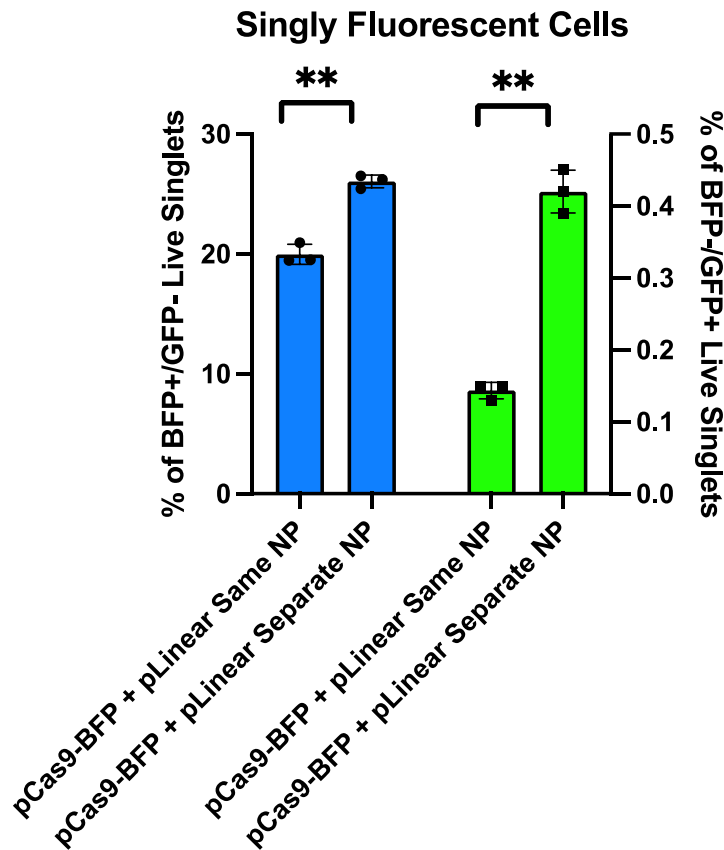
Fig. 4.8: Percentage of singly fluorescent (i.e. receiving either only pLinear or only pCas9-BFP, according to populations within GFP+/BFP- or GFP-/BFP+ gates in Fig. 4.7) cells. n = 3 replicate transfected wells. Plotted as mean ± SD. Paired t-test performed for each fluorescence readout. ** $p < 0.005$

>99% of cells that showed evidence of Cas9 expression also showed evidence of donor integration, as evidenced by coexpression of GFP and BFP (Fig. 4.7). However, there were a notable amount of cells (≥20%) that showed evidence of Cas9 expression only. Again, expression of Cas9 without an available donor construct could lead to unnecessary DSB formation and indels. Despite no pathogenic effects being reported as a result of PPP1R12C loss of function, DSB formation in general represents a threat to cellular health, and extraneous formation should be avoided. A small but statistically significant reduction in

single fluorescence of the pCas9-BFP plasmid was observed upon co-delivery. For this reason, it was decided that co-delivery in the same NP was favourable.

Furthermore, two unexpected results were observed in this data. Firstly, there was a significant difference in GFP expression ($p < 0.0005$) between pLinear and a positive control eGFP plasmid used for gating. This is explained by the presence of a strong enhancer (CAG) and promoter (CMV) present in the eGFP positive control, resulting in a high number of copies per cell of GFP. It is likely that this combination of enhancer and promoter were stronger than the native *PPP1R12C* promoter, under which any integrated constructs would be expressed. Additionally, using HITI, integration of a construct should eliminate the ability of Cas9 to re-cut, which theoretically restricts integration to a single copy per cell. Therefore, by setting flow cytometry parameters according to the plasmid with the stronger promoter, it is possible that successfully edited but dimly fluorescent cells were missed in the Cas9 + donor population.

Secondly, the donor-only control was not expected to show any fluorescence due to the lack of a promoter before the GFP within the donor. In order to examine whether this donor-only GFP expression was transient, it was decided that in subsequent experiments, GFP expression would be measured 48 hours post-transfection and also 12 days post-transfection.

4.3.2.3   Comparison of pLinear vs. pExcised donors

We next examined the difference in editing efficiency between the pLinear and pExcised donor constructs. As described above, upon Cas9 digestion, the pExcised plasmid generates two potential molecules for integration (the desired SA-T2A-GFP-polyA construct and the plasmid backbone) while the pLinear donor generates one (the entire plasmid,

linearised just upstream of the SA-T2A-GFP-polyA construct). When using the pExcised

donor, potential for integration of plasmid backbone rather than desired donor construct is

obviously unfavourable; however, there was the possibility that the desired donor would

integrate more readily than the backbone due to its smaller size (1.2 kb versus 4.2 kb).

In order to assess this, pCas9-BFP and each donor were transfected as described

above into HEK293T cells with L2K (same NP). This time, cells were sorted for co-expression

of BFP and GFP in order to isolate cells that had been successfully transfected. Then, 10 days

after the initial sort (12 days post-transfection), cells were sorted for GFP expression.

A

Initial Expression of GFP

B

Stable Integration of GFP
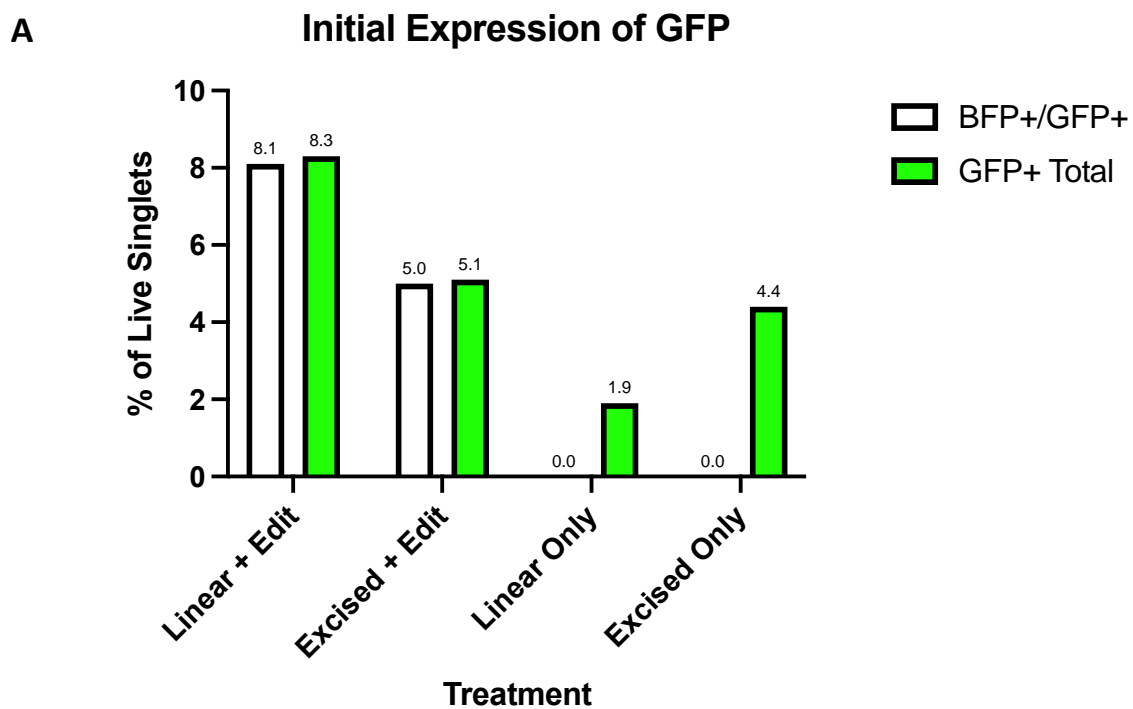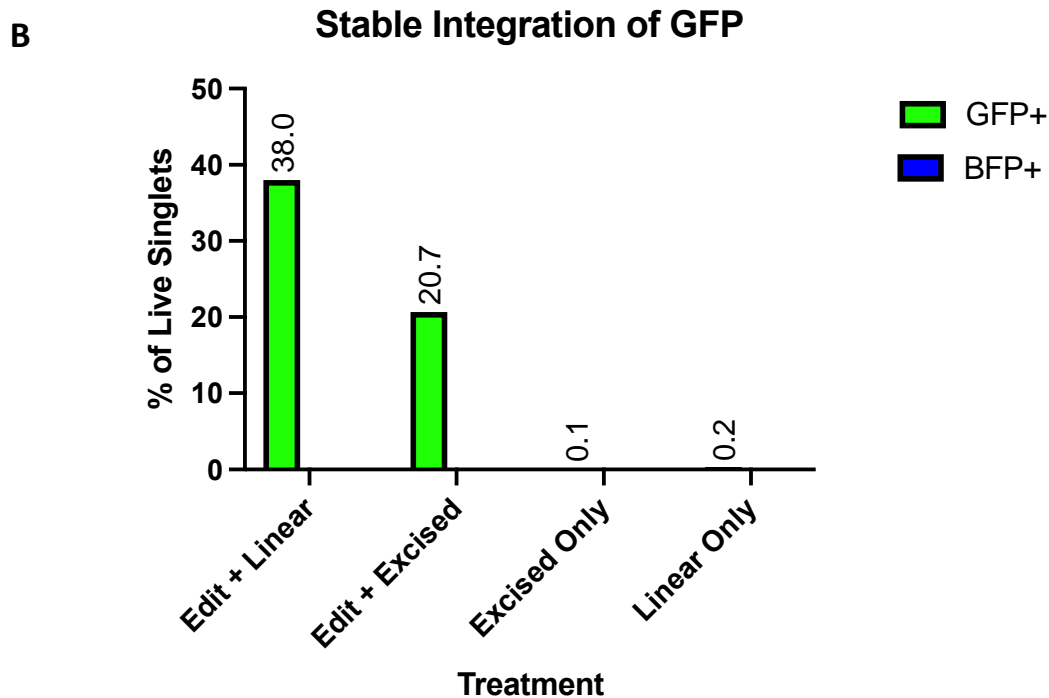
Fig. 4.9: Cell sorting of GFP+/BFP+ cells 48 hours post-transfection. pCas9-BFP, pLinear, and pExcised plasmids were transfected into HEK293T cells and sorted at indicated timepoints. n = 1 replicate transfected well per condition.

A) Results (percentage of fluorescent live singlets) of initial sort 48 hours post-transfection.

B) Results of final sort 12 days post-transfection.

Because pCas9-BFP was a non-integrating construct, its expression was expected to diminish over time. As expected, BFP expression was transient and had diminished after 12 days. Similarly, in the absence of Cas9, donor constructs are also non-integrating, and therefore also expected to diminish. Indeed, after 12 days, the GFP expression observed in the donor-only controls had diminished.

Interestingly, the pLinear donor yielded approximately double the integration of GFP compared to the pExcised donor (38% vs. 20.7% GFP+ cells respectively). This is likely

explained by the fact that cleavage of the pExcised donor yields two potential constructs for integration (the desired GFP expression construct and the plasmid backbone) while cleavage of the linear donor only yields one. This suggests that integration of the backbone fragment was not necessarily hindered by its size (1.2 kb for the GFP-encoding fragment, and 4.2 kb for the backbone fragment).

Given that the GFP expression of pExcised was approximately half that of pLinear (Fig. 4.9B), it is reasonable to assume that 50% of successfully edited pExcised cells integrated the desired GFP-encoding construct (with or without plasmid backbone still attached) and the remaining 50% either resulted in indel formation, integrated solely the bacterial plasmid backbone, or integrated the GFP construct in the reverse orientation. On the other hand, 100% of edited cells receiving the pLinear donor received bacterial DNA, including a kanamycin resistance construct.

4.3.1 Characterisation of edited cells

4.3.2.1 Stability of GFP expression in edited cells

We next wished to examine the consistency of GFP expression after successful integration under the control of the *PPP1R12C* promoter. In order to achieve this, transfection with pExcised donor and pCas9-BFP was repeated in HEK293T cells with L2K as before, and the following workflow was used for analysis:

On Day 0, cells were transfected with pCas9-BFP and pExcised plasmids, or donor-only control. On days two and 12 post-transfection, cells were sorted and returned to culture for expansion. Flow cytometry was then performed every 5 days until one month had passed since the initial sort.

A

Day 0   Day 2   Day 12

Day 17   Day 22   Repeat Until Day 32

B

- pExcised + pCas9-BFP
- pExcised Only
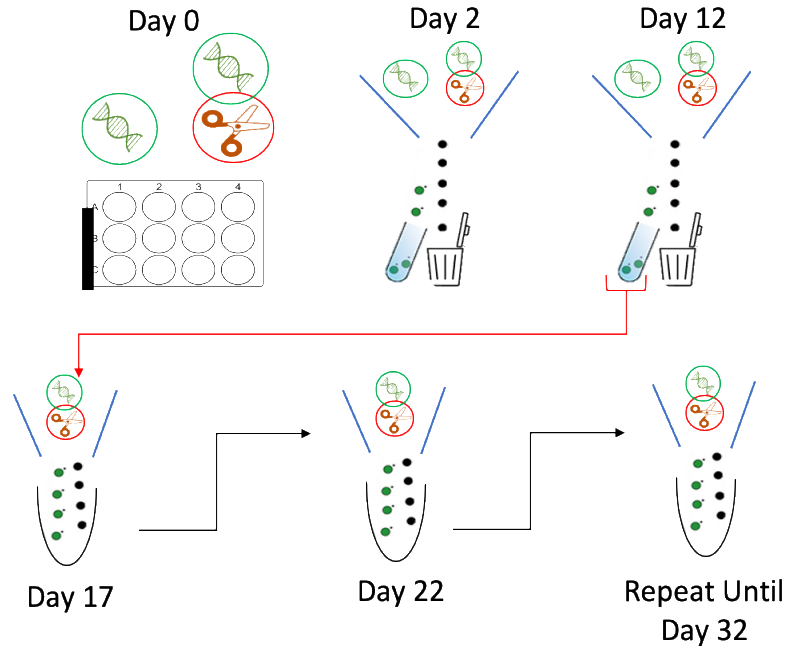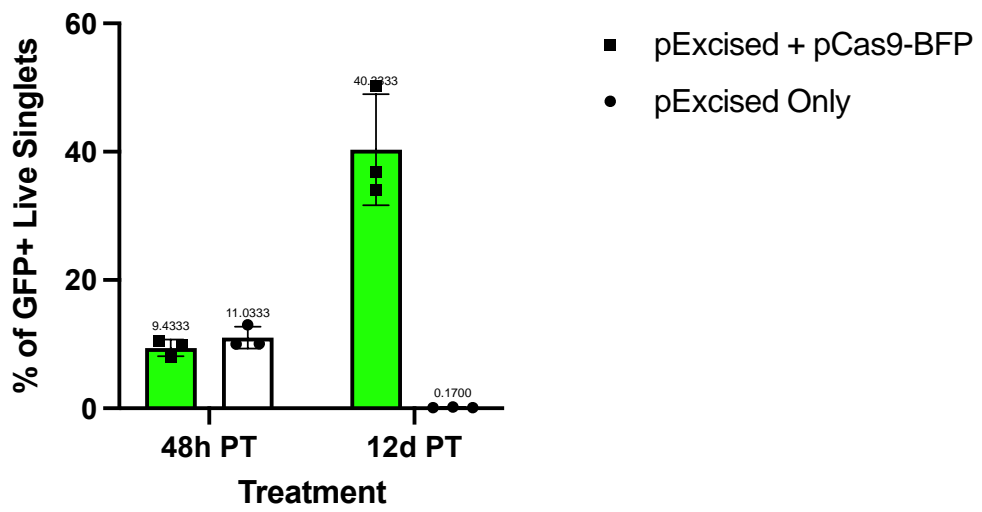
C
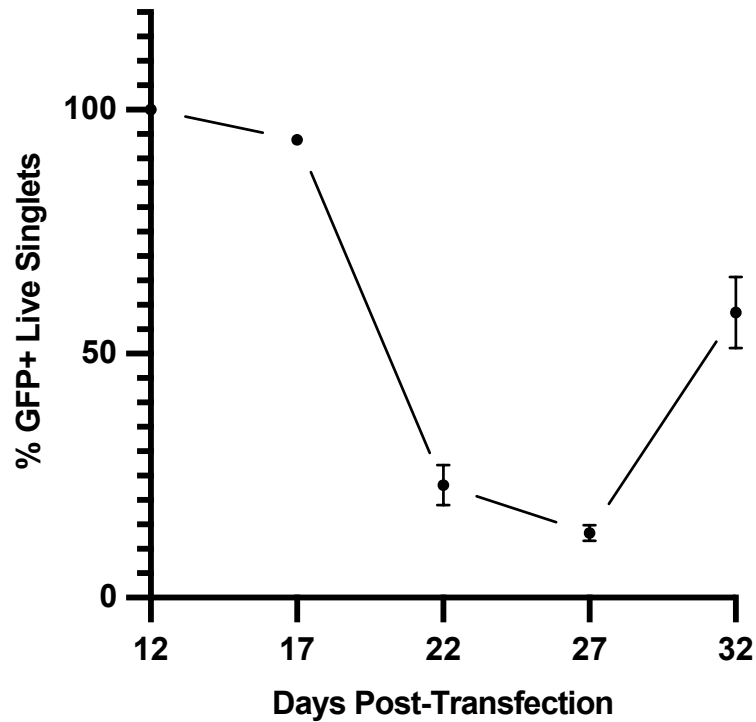
Fig. 4.10: Quantification of GFP expression over time in edited cells.

A) Illustrated workflow for experiment. pExcised represented by green DNA; pCas9-BFP represented by red scissors. On day zero, HEK293T cells were transfected with pExcised only or pExcised + pCas9-BFP using L2K. Cells were sorted 48 hours and 12 days post-transfection. Cells were then assayed by flow cytometry every five days until 30 days had passed since the initial sort (32 days post-transfection).

B) Initial sorts of successfully transfected (48h) and stably edited (12d) cell populations. n = 3 replicate wells per condition. Plotted as mean ± SD.

C) Tracking of %GFP+ live singlet cells over time. n = 3 replicate wells per condition. Plotted as mean ± SD.

Again, initial expression of GFP was low (~10% for each condition), and GFP expression was detected in donor-only controls despite lack of promoter. However, as

observed previously, transient donor-only GFP expression had ceased entirely by 12 days post-transfection, and permanently integrated GFP had increased substantially (~40%). This increase is likely due to the accumulation of GFP protein in cells, which increased overall brightness to above the threshold used for detection.

Upon tracking of GFP expression over time, there was a steep decline in expression in all three replicates 22 days post-transfection. This was unexpected, as the endogenous promoter at the AAVS1 site was presumed to be constitutively active as previously reported [317, 325]. It was initially assumed that the decline in GFP expression was due to selection against edited cells, and that the (presumably minute) subpopulation of GFP- cells that had been errantly sorted as GFP+ were outcompeting edited cells and growing to dominate the general population. However, at 32 days post-transfection, GFP expression recovered. This suggested that the transient loss of expression was not (or not entirely) due to cell die-off, but rather due to fluctuations in expression. This in turn suggested that GFP expression may have been an underrepresentation of actual editing efficiency, and that >40% editing efficiency may have initially been achieved. This also potentially serves to explain the inconsistency between editing efficiency in the first experiment using the excised donor (Fig. 4.9) and this independent experiment. Additionally, the fact that all replicates clustered relatively close together through the fluctuations suggests that this could be a programmed feature of the *PPP1R12C* promoter in HEK293T cells rather than random variation between replicates. In order to confirm that GFP+ cells actually contained the integrated donor construct, genotypic validation was performed.

4.3.2.2 Clonal isolation and genotyping of GFP-expressing cells

In order to ascertain the rate of targeted integration at the AAVS1 site, a PCR

strategy was designed to genotype pools of single-cell clones. HITI relies on NHEJ, and NHEJ

often leads to indels due to factors including resection of cut ends after DSB formation [122,

326, 327]. NHEJ-mediated blunt-end insertion of donors could therefore presumably lead to

genomic indels on either or both sides of successfully inserted donor constructs.

Clones were derived from each of the three replicate GFP+ edited populations

isolated in the above set of experiments (from a transfection of HEK293T cells with pExcised

and pCas9-BFP, using L2K). ≥10 clones per replicate were isolated via serial dilution as

described in "Cell Cloning by Serial Dilution in 96 Well Plates"

(https://www.corning.com/catalog/cls/documents/protocols/Single_cell_cloning_protocol.

pdf). Three PCR primers were designed: one just 5' of the cut site, one just 3' of the cut site,

and one spanning the junction of genomic DNA and the GFP expression construct. Primer

sequences available in Methods section under "AAVS1 F" and "AAVS1 R" (to amplify WT

genomic band) and "GFP Insertion Junction R" (to amplify junction band when paired with

"AAVS1 F"). DNA isolated from clones was amplified with both sets of primers, yielding a

543 bp amplicon (unedited genomic DNA) and/or a 250 bp amplicon (base-perfect

integration) (Fig. 4.11).

B

pExcised + pBFP-Cas9 Replicate 1 (17 clones)

2000 bp
1500 bp

500 bp

200 bp

Replicate 2 (11 clones)

Replicate 3 (13 clones)
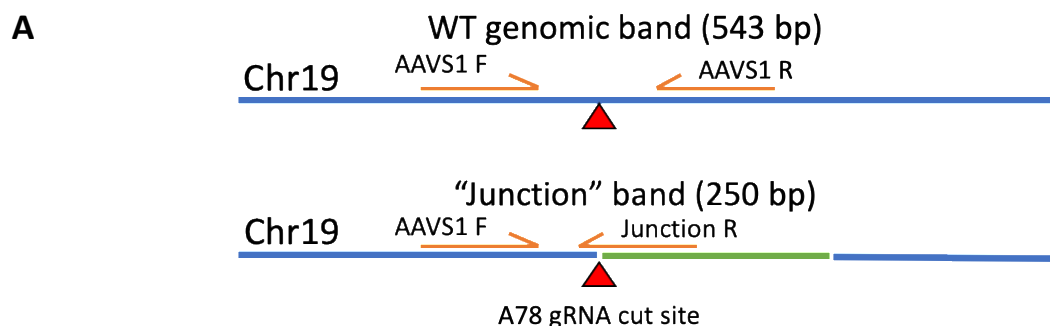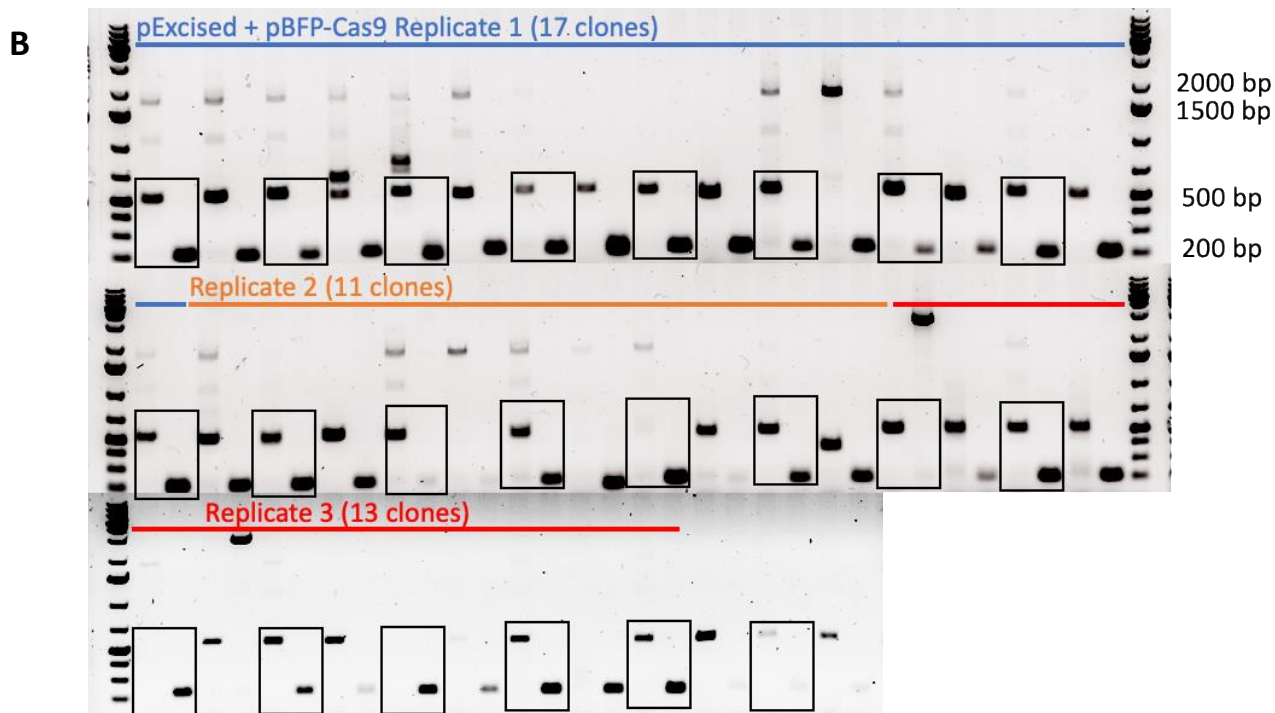
C

Single Cell Cloning of GFP+ Cells

- ● Excised Donor Plus Cas9
- ▲ Untreated

% of Clones

100

50

0

% of Clones with Any Integration

% of Clones with Biallelic Integration
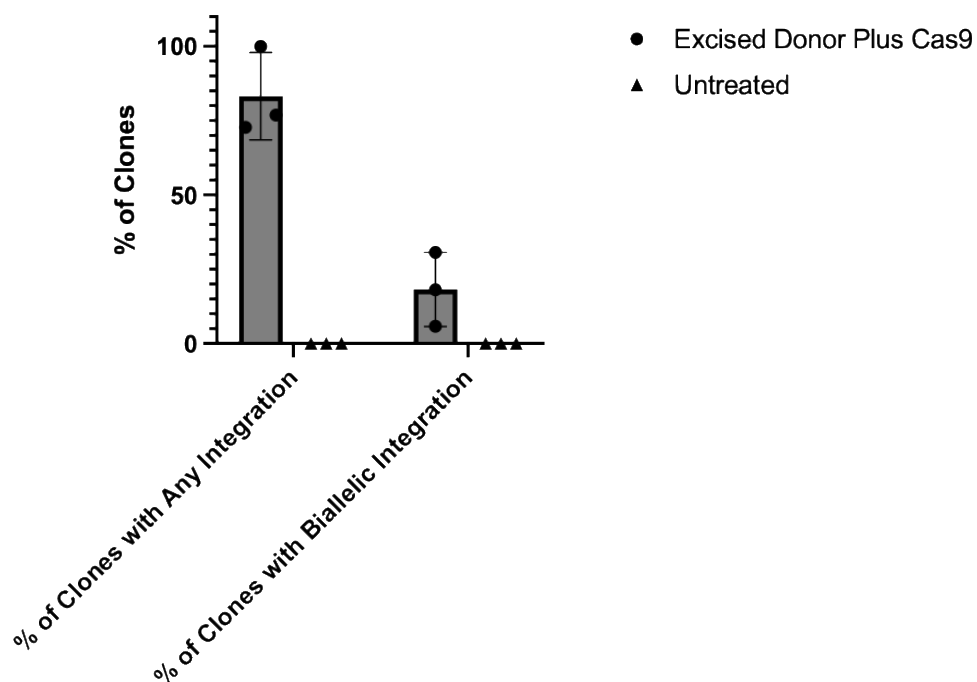
Fig. 4.11: PCR of single-cell clones isolated after HITI editing.

A) "Junction" PCR strategy yielding edited and unedited amplicons. Cut site shown as red triangle, primers shown as orange arrows, GFP insert shown as green bar.

B) Agarose gel showing results of junction PCR amplification. Single-cell clones were isolated from HEK293T populations found to be GFP+ 12 days after L2K-mediated transfection with the pCas9-BFP and pExcised donor constructs. Once clones had expanded to 80-90% confluency in a 96-well plate (8-10 days), genomic DNA was extracted as described in Methods. Clones derived from single transfection replicates are indicated by coloured bars. Boxes around pairs of bands indicate genomic DNA amplified from a single-cell clonal population. Upper band indicates presence of a WT allele, lower band indicates GFP insertion; both bands indicate heterozygotes.

C) Quantification of clones showing any (monoallelic or biallelic) or biallelic integration. n = 3 replicate transfected wells. Plotted as mean ± SD.

35/41 (85.4%) clones on the gel revealed the predicted 200 bp band, indicating targeted integration in at least one allele. 6/41 (14.6%) of clones revealed a band at 200 bp but no band at 500 bp, indicating targeted integration at both alleles. For multiple clones of the non-junction PCR, a faint band at 1.7 kb was detected. This is the expected size of the AAVS1 F/AAVS1 R amplicon upon integration of the desired donor construct; it is also detected in several of the clones where biallelic editing is indicated. In clone 5 of replicate 2, this 1.7 kb band is visible despite no 500 or 200 bp band. This indicates biallelic integration of the construct, but that the integration was not base-perfect at either allele, meaning that the junction R primer could not successfully bind. This brings the total number of suggested biallelic editing events to 7/41, or 17%. In two clones, there was a strong band at ~5000 bp for the junction PCR. This could potentially be explained by integration of the entire donor plasmid. In order for this to happen, the plasmid would need to be linearised via a Cas9-

mediated cut occurring at only the post-GFP RC-gRNA sites in the donor. Importantly, if only the pre-GFP RC-gRNA was cut, the resultant junction band would still be 200 bp. Therefore, in future experiments, it would be wise to examine both junctions of the expected integration site, and/or Sanger sequence resulting amplicons. There were also very faint bands at ~1000 bp in several lanes; this does not correspond to any predicted amplicon size, and is therefore likely to be a PCR artifact.

On average, 92% of clones exhibited targeted integration on at least one allele, and 16% showed targeted homozygous editing. For several clones, there were multiple bands present in the "genomic" amplicon. This suggests formation of random insertions at the non-HITI allele. However, because only one functional copy of *CFTR* is required for functional rescue, monoallelic repair is sufficient. Additionally, because disruption of the *PPP1R12C* gene has no reported pathogenic effect [328], indels at this locus are predicted to have negligible effect on overall cell health. Indel formation at the native *CFTR* locus could be much more harmful, as it could eliminate any residual function of the individual's CF mutation.

While high rates of on-target editing efficiency are obviously favourable, they are often accompanied by formation of off-target edits. Therefore, it is essential to examine sites of the genome with homology to the on-target spacer sequence to assay for off-target editing events.

4.3.2.3 Off-target analysis of clones with biallelic editing

The web-based tool CRISPOR was used to predict potential off-target sites (OTSs) for the A78 gRNA. Primers were designed to target the top five predicted OTSs of the genome (all OTS primer sequences in Methods). In order to ensure that GFP had not integrated at

these sites, PCR was run with an extension time sufficient to include GFP insertion, and

expected band sizes validated on agarose gels. PCR was performed using gDNA from all

seven homozygote edited clones (as determined by electrophoresis, see Fig. 4.11) to amplify

these sites, and Synthego's Incidence of CRISPR Edit (ICE) analysis was used to detect any

indel formation at those loci. ICE analysis compares the heterogeneity of Sanger-derived

chromatogram traces from edited and unedited samples in order to infer rates of indel

formation.

Fig. 4.12: Representative traces from the same OTS site (OTS 5) in two

homozygote clones where zero indel formation was detected.

A) Traces with indel % = 0 and R^2 = 1. Edited and control chromatogram traces are functionally identical, with no heterogeneity surrounding the cut site.

B) Traces with indel % = 0 and R^2 = 0.95. Some heterogeneity is present in the edited chromatogram; however, heterogeneity does not align with the cut site, suggesting that this was background due to impurity of sample rather than an editing event.

Off-Target Site (OTS)	Genetic Feature	gRNA (# Mismatches)	% Indels Detected	Confidence (R^2)
Target Site Chr19: 55115745-55115764	Intronic (*PPP1R12C*)	**GTCACCAATCCTGTCCCTAG (0)**		
1 Chr22: 48160908-48160927	Intergenic	G**G**CAC**T**AATCCT**A**TCCCTA**A** (4)	0	1
2 Chr4: 98714021-9874040	Intergenic	GTCAC**A**AATCCT**A**TC**ACA**AG (4)	0	1
3 Chr2: 178603025-178603044	Intronic (*TTN-AS1*)	G**A**CAC**TG**ATCCTGTCCCTA**A** (4)	0	1
4 Chr12: 105200588-105200607	Intronic (*APPL2*)	Predicted: GT**A**ACCAA**GA**CTGTC**A**CTAG (4) Experimental: GT**A**ACCAA**GA**CTGTC**A**CTA**C** (5)	0	0.99
5 Chr8: 141870522-141870541	Intergenic	GTCAC**AAAA**CCTGTC**A**CT**T**G (4)	0	0.95

Table 4.1: Indel formation detected at OTSs by ICE analysis for one representative homozygous edited clone. DNA was extracted from clones and PCR performed in order to amplify off-target sites. Sanger sequencing traces were used as input into ICE software, which reported indel frequency at off-target sites. In gRNA column, mismatches from on-target spacer sequence are highlighted in red, and number of mismatches from on-target spacer sequence are shown in parentheses. For OTS 4, "Predicted" refers to the sequence predicted by CRISPOR software, and "Experimental" refers to the sequence detected in untransfected cells' genomic DNA. Chr: chromosome; numbers following indicate range of bases on that chromosome covering the OTS (per UCSC Genome Browser, reference genome GRCh38/hg38) [329].

Zero indel formation was detected at any OTS in any clone tested (Table 4.1). At OTS 4, the canonical predicted gRNA sequence was not present in any of the tested samples, including the untreated control. Upon further investigation, there was a single base substitution immediately 3' of the PAM in every clonal amplicon, as well as the untreated control. Because this substitution was also present in the untreated control, it was presumed to not be caused by CRISPR editing. Therefore, the canonical gRNA ICE input sequence was altered to match the untreated control so that ICE analysis could be performed; upon ICE analysis, zero indels were detected at this OTS in any clone assayed.

Given these results, it was concluded that a single round of editing successfully integrated a cDNA at the AAVS1 locus with no detectable off-target editing events.

4.4   Discussion

In this chapter, I have demonstrated a proof of concept for integration of large coding construct into the native *AAVS1* locus under control of the native promoter. I have also used ICE analysis to demonstrate a lack of off-target editing effects in both singly and repeatedly edited cell populations.

Firstly, for proof of concept, GFP cDNA was integrated at the *AAVS1* locus of HEK293T cells. Unexpectedly, despite lacking a promoter, the donor-only control showed some evidence of GFP expression. It was theorised that the GFP expression in the donor-only control was due to recruitment of RNA polymerases after DSB formation. It has been shown that transcription at DSB sites can be initiated by recruitment of RNA polymerases II and III [326, 330, 331]. Typically, this recruitment results in DSB-induced long noncoding RNAs (dilncRNAs), which do not translate into functional protein products. However, DSB formation can also recruit other members of a pre-initiation complex, forming functional promoters at cut sites [331]. Due to the absence of any Cas9 in the donor-only control, it was theorised that the DSBs may have been formed by an endogenous nuclease. Despite normally acting on single-stranded DNA and RNA, the endogenous S1 nuclease has previously been shown to form double-strand breaks in supercoiled pDNA in the cytosol four hours after pDNA transfection [332].

When cells were sorted for dual expression of GFP (donor) and BFP (Cas9), the resulting population of successfully transfected cells could be followed over time to assess permanent GFP expression as a readout of successful integration of the GFP cDNA donor construct. By this parameter, 20% editing efficiency was achieved with the pExcised donor, and 38% with the pLinear donor. However, pLinear donor includes a kanamycin resistance

construct. Permanent integration of bacterial DNA, particularly encoding antibiotic resistance, is generally not considered translatable to the clinic [257]. Additionally, Johnson *et al.* found that, when wild-type (WT) and CF-variant cells were mixed and assayed for ion transport activity, as low as 6-10% of WT cells were sufficient to rescue ion trafficking to WT levels. Therefore, if a CFTR cDNA construct were to integrate into cells with the editing efficiency achieved with the pExcised donor (20%), this would be more than sufficient to have a therapeutic effect in CFTR-defective cells [318].In a subsequent experiment performed identically, the pExcised editing efficiency increased to 40%. This was initially alarming, as it represented lack of repeatability in the results. However, further analysis of the integrated GFP's expression over time revealed that GFP expression alone is not a reliable readout of actual editing efficiency at the molecular level. After successful DNA editing, it was found that eGFP under the native *PPP1R12C* promoter was not expressed constitutively (as measured by %GFP+ cells in population over time) in HEK293T cells. Significant dips in GFP expression were detected, which later recovered. Another report did not observe this phenomenon; however, this used a different editing approach (ZFNs) and a different cell type (K562 cells) [325]. A recent report provided evidence that transcriptional fluctuations occur at the AAVS1 site even when a foreign promoter is integrated alongside the transgene [333]. Inderbitzin *et al.* used CRISPR to integrate barcoded constructs into Jurkat cells at the AAVS1 locus and compared barcode prevalence in cDNA versus gDNA. Significant transcriptional variation was detected even with a non-native promoter. It was proposed that this could be due to the HIV-1 LTR promoter's sensitivity to local chromatin; this suggests that the *PPP1R12C* site undergoes fluctuations in chromatin structure, contrary to previous reports [313]. This could have resulted from DSB formation, as DSBs have been shown to promote transient methylation after DSB formation in *PPP1R12C* in HEK293T cells

[334]. The fact that fluctuations were consistent across three replicate edited populations support the idea that this is a programmed response to editing rather than random fluctuations in individual cells or populations.

More research is needed into expression of foreign cDNAs under this native promoter in order to ascertain whether its activity would be sufficient to rescue CF pathology if a *CFTR* cDNA were placed under its control. Timescale analysis of PPP1R12C expression in CF ALI cultures would yield valuable insight. Alternatively, *CFTR* cDNA would need to be paired with a strong and constitutively active foreign promoter to provide clinical benefit. Oceguera-Yanez *et al.* reported that, when a reporter construct was integrated at the AAVS1 locus under the CAG promoter, expression remained stable for up to six months in iPSCs [314]. However, there is also precedence for caution in overexpressing CFTR. It has been shown that integration of codon optimised CFTR with the strong spleen focus-forming virus (SFFV) promoter actually yielded less rescue of CFTR function as measured by Ussing chamber [335] However, in this experiment, delivery was achieved using lentiviral integration, and therefore multiple copies could have integrated per cell. It was theorised that, upon generation of excess CFTR protein, some protein could errantly embed in cells' basolateral membranes, thus counteracting the effects of apically localised protein. In theory, because successful HITI integration destroys spacer sequence, only one copy per cell should integrate, thus alleviating this potential issue when paired with an appropriate promoter. The CFTR promoter could potentially be included in the donor construct, allowing for appropriate control without threatening the cell's native CFTR promoter; however, as discussed in Section 4.1, a stronger promoter could compensate for lower editing efficiency.

Another major hurdle to editing efficiency is transfection efficiency and activity of Cas9 within cells. The editing experiments described in this chapter were performed in HEK293T

cells, which are known to be amenable to transfection [336]. Primary cells are more difficult to transfect, especially when differentiated on ALI due to cilia movement and formation of tight junctions [291]. A simpler approach would be to target the AAVS1 locus in submerged, undifferentiated BMI-1 transduced cells, and then differentiate after successful editing; however, because the ultimate goal of this strategy is to develop an *in vivo* treatment, ALI transfection is a more faithful model. Therefore, for future experiments using this strategy to insert *CFTR* cDNA, several steps can be taken to enhance transfection and editing efficiency. Firstly, RNP formulated with synthetic gRNA has been shown to outperform plasmid delivered Cas9 and gRNA [221, 256]. As discussed in Introduction section 1.8.3, delivery of RNP is favourable over plasmid due to its more transient half-life in cells and the biochemical modifiability of synthetically generated gRNA.

In addition, the structure of donor molecules can also enhance editing efficiency. As outlined in Fig. 4.5, the pExcised donor plasmid yields two potential constructs for integration. A more favourable approach would be delivery of a blunt-ended DNA oligo. Single-stranded DNA constructs (ssODNs) have been shown to outperform double-stranded constructs (dsODNs) in HDR experiments [337]. However, this discrepancy is theorised to be due to differing repair pathways active in HDR in the presence of a dsODN vs. ssODN, upon which HITI does not rely [338]. Because HITI relies on NHEJ, a dsODN would likely be optimal for direct conjugation to exposed DSB ends. Commercial synthetic generation of a full-length *CFTR* cDNA dsODN is extremely costly; therefore, cheaper approaches could be used to generate the desired molecule. Firstly, restriction enzyme digestion could be used on the pExcised plasmid, and gel electrophoresis could be performed followed by band excision and purification. However, yield would be limited by both plasmid input (as the desired construct represents only ~20% of the plasmid by mass) and cleanup capacity. Additionally,

it is highly unlikely that a given donor construct would have a unique restriction enzyme site at the exact Cas9 cut point in the reverse complemented-gRNA. Generation of overhangs would mean further generation of multiple unwanted free DNA constructs upon Cas9-mediated "trimming" of excess sequence. Another option is direct PCR amplification off the plasmid. Because gRNA sequences tend to follow similar rules to plasmid design (20 bp, ~50% GC content), primers could be designed that yield an amplicon identical to that generated by Cas9 cleavage. However, due to the presence of multiple RC-gRNA sequences in the pExcised plasmid template, nonspecific primer binding may occur. This could be alleviated by using a generic plasmid including only the desired expression construct, and appending RC-gRNA overhangs via primer extensions. Notably, even high-fidelity Taq polymerases may introduce point mutations in multi-kB amplicons, potentially yielding a mixed population of constructs. Integration of a *CFTR* cDNA carrying mutated sequence would be unlikely to provide clinical benefit. Kelly *et al.* used minicircle plasmids for delivery of their constructs, which, while lacking plasmid backbone and antibiotic resistance constructs, still relies on simultaneous cutting of Cas9 and donor for integration to occur [315].

Another highly relevant study targeting the AAVS1 locus relied not on CRISPR-Cas9 but on TALEN-mediated DNA editing [339]. In a 2019 report, TALEN homodimer constructs and donor DNA (either LacZ reporter, or *CFTR* cDNA) were delivered via HD-Ad vector to a CFBE (F508del/W1282X) cell line. The donor sequence was flanked by symmetrical 4 kb homology arms. At MOI 100, the TALEN construct was found to cut target DNA with an efficiency of 66% (measured by TIDE assay). The LacZ reporter construct was found to integrate at the target locus with an efficiency of 5%. When *CFTR* cDNA was integrated, successfully edited cells displayed rescue of protein production (via Western blot) and apical

ion transport (via FLIPR assay). This methodology did not achieve the level of editing observed via CRISPR/HITI, and carries two further limitations to its application. Firstly, codelivery of donor and nuclease in the same vector risks genomic integration of a nuclease-expressing construct, which is not clinically translatable. Secondly, reliance on HDR limits the targetable cell type to basal cells, as this editing strategy cannot occur in the differentiated apical epithelial layer.

Another important consideration in CRISPR-based treatments is off-target editing. Zero off-target editing was detectable at five predicted OTS sites after HITI-mediated cDNA insertion construct into HEK293T cells via plasmid-expressed Cas9. However, if a DNA editing strategy is to be taken to the clinic, more robust investigation of off-target editing effects is required. *In silico* prediction methods, while having been shown to be reliable experimentally, are not an adequate substitute for in-depth full-genome sequencing. In next steps, NGS or an unbiased method (such as CIRCLE-Seq, described in Introduction Section 1.3.3.2) would be a vital aspect of experimental design to further validate the lack of off-target editing events in successfully edited cells. Another independent study using this gRNA performed no off-target analysis, and overall integration efficiency was not reported [292].

Overall, these preliminary results reveal integration of a reporter construct in at least 40% of transfected cells, and zero detectable off-target editing at five surveyed loci. Future study should take place in primary airway epithelial cells, and carefully consider the structure of delivered Cas9 RNP and donor molecules. Transfection of human airway epithelial cells with RNP using receptor-targeted nanoparticles is further explored in the following results chapter.

Results: Characterisation of cationic lipid based receptor-targeted nanoparticles to the airway epithelium

5.1 Introduction

One of the major concerns with DNA editing applications is delivery [340]. Any *in vivo* approach must rely on either systemic or targeted delivery. Systemic delivery, usually accomplished via administration to the bloodstream, is favourable for accessing a wider range of organs and tissues. However, despite the fact that some systemic editing approaches have reached clinical trials, no current DNA editing technology is completely without risk of off-target edits or DNA damage-induced apoptosis (and downstream risks therein, as discussed in Introduction Section 1.3.1.1). Despite some DNA editing treatments already approved or in trials, follow-up studies have recently been proposed to comprehensively profile their safety [341]. Therefore, if only a specific population of cells requires the therapy, it is safest to target only those cells and avoid any unnecessary editing of other tissues. This can be accomplished via a variety of means. Firstly, researchers can choose to target only specific tissues with their delivery approach. These strategies include direct injection into the target organ (such as the eye) [342], topical application [343], or inhalation [344, 345]. However, the composition of RTNs themselves can be used to enhance cell-specific targeting via conjugation with cell-targeting peptides.

Lipid-based RTNs have been used by our group in the past to deliver nucleic acids to airway epithelial cells *in vitro* and *in vivo*. The receptor-targeting peptide used was identified via phage display and confirmed to significantly enhance transfection of human bronchial epithelial cells [264]. The lipids used in these RTNs are one of three cationic lipids with differing alkyl tail lengths (C14, C16, and C18) and the neutral helper lipid DOPE. C14 (DTDTMA) and C18 (DOTMA) were directly compared for delivery of reporter (GFP or luciferase) plasmids to various cultured mammalian smooth muscle cell types and explants

of rabbit aorta [346]. It was initially assumed that the shorter alkyl chain length in C14 would lead to more structural flexibility, allowing for low-temperature phase changes that aid in endosomal escape by C14-formulated RTNs. Each lipid was assembled into an RTN with the same targeting peptide (Peptide Y, K_{16}GACYGLPHKFCG) at the same weight ratio to DNA cargo (Peptide Y consists in part of a string of 16 lysine residues, which carry a cationic charge, allowing them to complex with cargo; three amino acids serve as a spacer sequence; the remainder of the peptide is exposed on the surface of the nanocomplex and allows for targeting and association with cell membrane receptors [264]). RTNs were also assembled with or without inclusion of the neutral helper lipid DOPE at a 1:1 weight ratio; DOPE has a cone-shaped structure that allows for assembly into an inverted hexagonal phase [347]; this structure has been associated with higher rates of endosomal escape after transfection [273, 348]. In cultured cells (receiving luciferase plasmid), RTNs containing C14 outperformed those containing C18 in every cell type [346]. Additionally, formulations including DOPE outperformed those lacking DOPE by 25-33% in all cases. However, in rabbit aorta explants, C18DOPE outperformed C14DOPE, highlighting differences between *in vitro* cultured cells and organ models [346].

In vitro experiments using a different targeting peptide (K_{16}GACSERSMNFCG, or Peptide E) determined that C16DOPE was optimal for transfection of human airway epithelial cells (CF variant and normal human) with luciferase-encoding plasmid. RTNs formulated with C16DOPE and Peptide E were intratracheally instilled (direct injection of RTNs in 50 μL of vehicle) in mice and found to efficiently transfect bronchial epithelial cells *in vivo* [287]. This optimised gene transfer formulation was later nebulised for murine inhalation. Luciferase plasmid was once again packaged in RTNs consisting of C16DOPE and Peptide E before nebulisation. 71% of mice receiving an aerosolised dose of RTNs showed

luciferase expression, primarily in ciliated tracheal epithelial cells [345]. C18DOPE was also compared with C14DOPE and C16DOPE for delivery of siRNA to various mammalian cell types, including human airway epithelial cells, in RTNs formulated with Peptide Y [288]. In these cell types, contrary to the ones used in the previous study, C14DOPE RTNs were optimal for transfection with luciferase-encoding plasmid. However, when cargo was changed to siRNA, C18DOPE RTNs were optimal, indicating that maximal transfection efficiency is somewhat cargo-dependent.

Taken together, these studies highlight the efficacy of cationic lipids and receptor-targeting peptides in transfecting various cell types both *in vitro* and *in vivo*. However, differences in cell types targeted and RTN cargo have an impact on which lipid formulation is optimal in specific cases.

Extensive characterisation has been performed of these RTNs in delivering DNA and RNA [96, 265, 273, 290, 345, 349]. However, less substantial characterisation has been done for RNP delivery. The biophysical properties typically investigated are size, polydispersity index (PDI), and charge. Nanoparticles with uniform size distribution (PDI ~0.2-0.3) are considered the standard for clinical application [350]. Cationic charge is also associated with favourable interaction with anionic cell membrane components [351]. Therefore, biophysical characterisation of RNP-containing RTNs is critical for optimising ability to transfect clinically relevant cells.

5.2 Aims

In this chapter we aim to:

1) Investigate the role of different components in RTN assembly and targeting.

2) Investigate biophysical properties of RTNs under different assembly conditions.

3) Investigate ability of RTNs to transfect ALI-differentiated cells.

4) Analyse RTNs' mucus penetration capabilities.

5.3 Results

5.3.1 RTN components and basic assembly

5.3.1.1 Nanoparticle components

RTNs used to target airway epithelial cells consisted of the following components (Fig. 5.1):

- o One of three different cationic lipids

    - Ditetradecyl trimethyl ammonium propane (DTDTMA, C14)

    - Dihexadecenyl trimethyl ammonium propane (DHDTMA, C16)

    - Dioctadecenyl trimethyl ammonium propane (DOTMA, C18)

- o The neutral phospholipid 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), at a 1:1 weight ratio with one of the cationic lipids above

- o Targeting peptide initially discovered via phage display (Peptide Y, K_{16}GACYGLPHKFCG) [264]

Cationic lipids
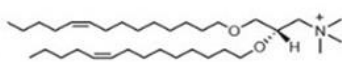
DTDTMA (C14)

DHDTMA (C16)

DOTMA (C18)
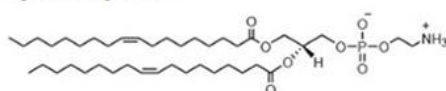
Phospholipids

DOPE

Fig. 5.1: Lipid components of RTNs. DTDTMA (C14), DHDTMA (C16), and DOTMA (C18) are cationic ionisable lipids. DOPE is a neutral "helper" phospholipid. From Grant-Serroukh, D., M. R. Hunter, R. Maeshima, A. D. Tagalakis, A. M. Aldossary, N. Allahham, G. R. Williams, M. Edbrooke, A. Desai and S. L. Hart (2022). "Lipid-peptide nanocomplexes for mRNA delivery in vitro and in vivo." J Control Release 348: 786-797.

The rationale for each component was based on previous studies, as described in the introduction to this chapter. In brief, the lipids C14, C16, and C18 (each named for their respective alkyl chain lengths) carry cationic charge that allow them to interact electrostatically favourably with negatively charged cargo such as nucleic acids and anionic RNP, as well as with anionic components of cell membranes.

5.3.1.2 Development of an assay to quantify RTN encapsulation efficiency of RNPs

We first aimed to analyse RTNs' ability to successfully encapsulate RNP cargo. In order to develop a suitable assay, we first attempted use of fluorescently tagged Cas9 in our RNP formulations. Two forms of GFP-tagged Cas9 (expressed as fusion proteins) were purchased and assessed: one from GenScript, and one from Integrated DNA Technologies (Fig. 5.2). Because GFP and Cas9 protein both carry a net negative charge, association with cationic lipid and peptide should not be hindered by inclusion of GFP. RNP was assembled by combining gRNA and Cas9 at a 1:4 weight ratio and allowing to self-assemble for five minutes at room temperature; this was the ratio found to be optimal for editing of NHBE BMI-1 cells by previous Hart lab member Dr. Amy Walker [324]. Subsequently, RNP was either encapsulated in RTN in the mixing order RNP-lipid-peptide or left naked. RTNs and

RNP were diluted in Tris buffer (50 mM pH 7.4) in a 96-well plate in six replicate wells.

Heparin was then added to three of the six wells (10 uL of 10 U/mL in 200 uL assay volume

for a final working concentration of 0.5 U/mL) before measurement. Heparin, a densely

negatively charged molecule, outcompetes RNP cargo for association with cationic lipid

components, thus dissociating the RTN [352, 353].
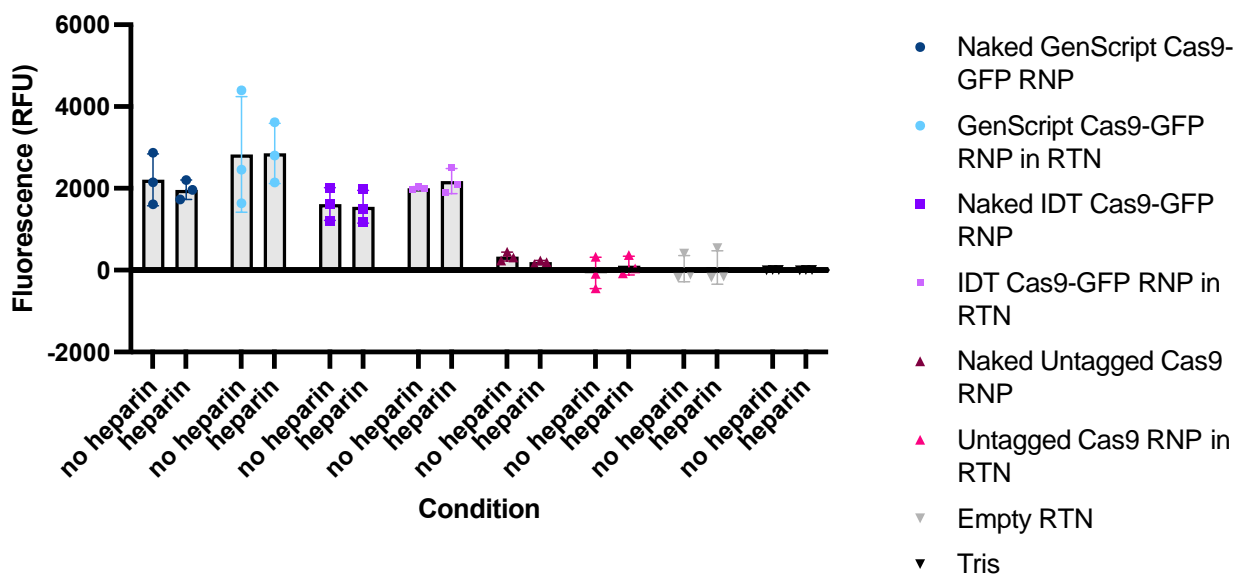
Fig. 5.2: Fluorescence of GFP in RNP upon encapsulation and subsequent

release by RTNs. RTNs were formulated at a 1:4:1 lipid:peptide:RNP mass

ratio. Fluorescence was measured via FluoStar Optima plate reader with

excitation filter wavelength 492 nm, emission filter wavelength 520 nm, and

gain 2000. n = 3 replicate wells per condition. Plotted as mean $\pm$ SD.

No significant difference was detected between the two brands of GFP-tagged Cas9.

However, GFP fluorescence was detected with and without addition of heparin, meaning that it could not be used as a readout of encapsulation efficiency of RTNs. We therefore sought to use a fluorescent readout that would be quenched by RTN encapsulation. Staining gRNA rather than using fluorescently tagged Cas9 protein would theoretically allow for differentiation between successfully encapsulated and unencapsulated RNP. RiboGreen, a fluorescent RNA stain, has previously been used by our lab to stain mRNA and siRNA cargo [288], and we sought to adapt this approach into staining gRNA-complexed RNP (Fig. 5.3).

A

B

RiboGreen Quenching
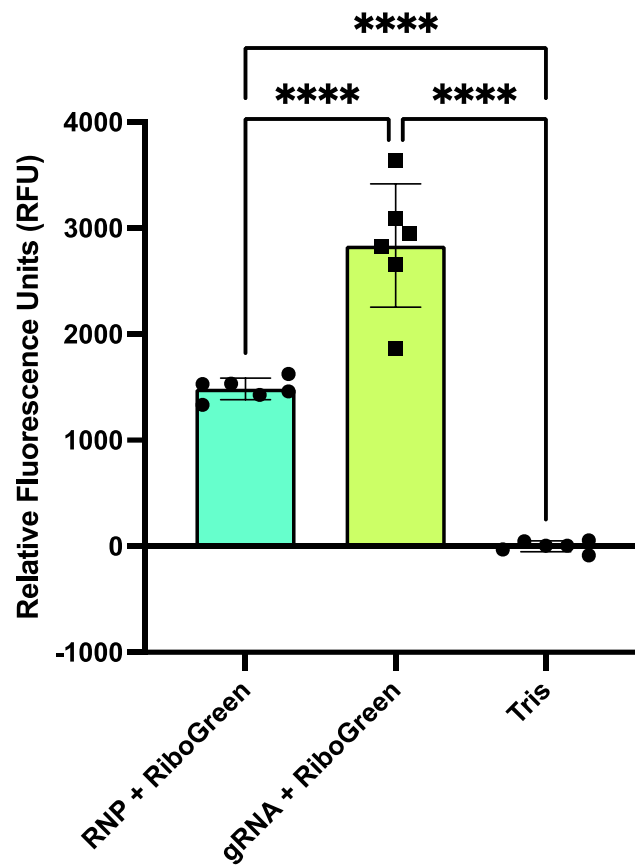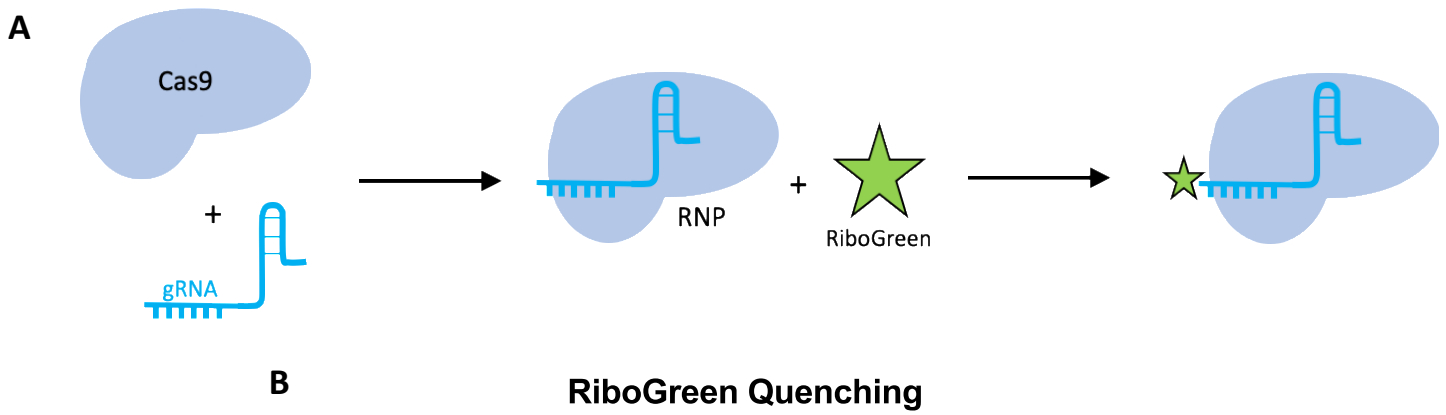
Fig. 5.3: RiboGreen assay for gRNA-complexed Cas9.

A) Schematic of RiboGreen staining approach. Cas9 and gRNA were

combined at a 4:1 weight ratio, incubated at room temperature for five

minutes, diluted to 100 uL in Tris buffer (50 mM, pH 7.4) and incubated

with 100 uL RiboGreen working solution prepared per manufacturer's instructions.

B) Fluorescence readout of RiboGreen staining approach. Fluorescence was measured via FluoStar Optima plate reader with excitation filter wavelength 492 nm, emission filter wavelength 520 nm, and gain 2000. Plotted as mean $\pm$ SD. One-way ANOVA with multiple comparisons. n = 6 replicate wells per condition. **** $p < 0.0001$

gRNA was shielded somewhat from RiboGreen staining by complexing with Cas9. Even when complexed with Cas9, the segment of a gRNA molecule with homology to the spacer sequence must remain exposed so that it is free to complex with target DNA [354]. Presumably, RiboGreen is able to stain only this exposed section of the gRNA; when the gRNA is freed from Cas9, its entire length can be stained.

It is possible that some fluorescence could have come from staining free gRNA (*ie* unbound by Cas9). However, this is unlikely due to the fact that Cas9 is in excess compared to gRNA. The molar mass of recombinant Cas9 protein is ~162,000 g/mol, and the molar mass of the average synthetic sgRNA is ~32,000 g/mol [355]. Therefore, for a 1:1 molar ratio of Cas9:gRNA, a 1:5 mass ratio should be used. A molar ratio of 1:4 means that Cas9 is in excess, making it mathematically unlikely that any free gRNA remains unbound.

Having ascertained that RiboGreen successfully stained the gRNA in RNP, we sought to investigate whether this staining would be quenched by encapsulation in RTN. All conditions were formulated at a 1:4:1 lipid:peptide:RNP ratio, treated with the indicated quantity of heparin, and then incubated with RiboGreen before fluorescence was detected via plate reader (Fig. 5.4).
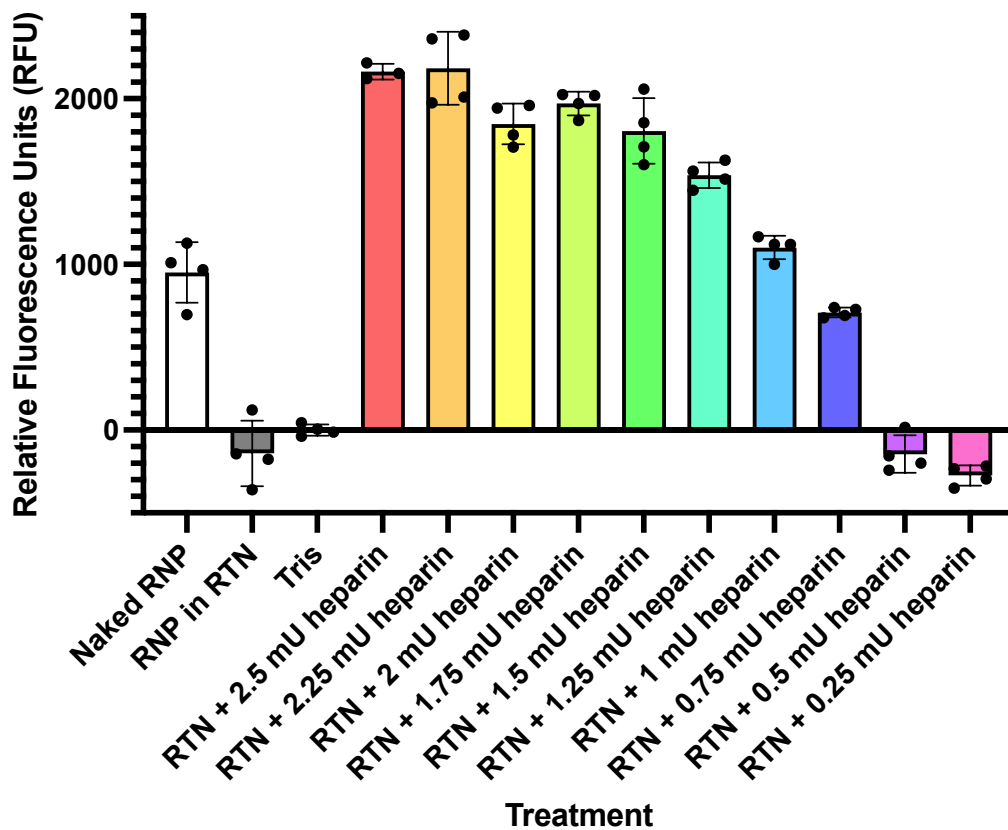
Fig. 5.4: Quenching of RiboGreen fluorescence by RTN assembly. RTNs were assembled at a lipid:peptide:RNP ratio of 1:4:1 and subsequently treated with the indicated quantity of heparin in 200 µL volume. RiboGreen fluorescence detected using FLUOStar Optima plate reader with excitation filter wavelength 492 nm, emission filter wavelength 520 nm, and gain 2000. n = 4 replicate wells per condition. Plotted as mean ± SD.

Fluorescence of RiboGreen labelled gRNA in the RNP was completely quenched by packaging in RTN. For heparin quantities >0.00175 U, the ratio of pre-dissociation:post-dissociation fluorescence was similar to the ratio of naked gRNA:RNP-complexed gRNA

observed in the previous experiment. This suggested that heparin not only dissociated the RTN but also the RNP cargo itself, freeing gRNA from Cas9. In all future experiments, 0.1 U heparin was added per 200 μL for a saturating quantity in order to assess encapsulation efficiency. The RiboGreen assay described above was then used to analyse the role of lipid and peptide components in the formation and stability of the RTN.

5.3.1.3 Role of targeting peptide in assembly

RTNs were generated via formulation of lipid nanoparticles with targeting Peptide Y (K_{16}GACYGLPHKFCG). The 16 positively charged K residues allow for electrostatic interaction with negatively charged cargo, resulting in a complex of lipid, peptide, and cargo (in this case, RNP). In order to investigate Peptide Y's effect on RTN assembly, the RiboGreen assay was performed on RTNs formulated with varying proportions of targeting peptide (Fig. 5.5): lipid proportion was kept at a 1:1 ratio with RNP. RTNs were assembled in the mixing order lipid-RNP-peptide and either treated with heparin or left intact.

A

Peptide Ratios for Assembly (Lipid:Peptide:RNP)

B
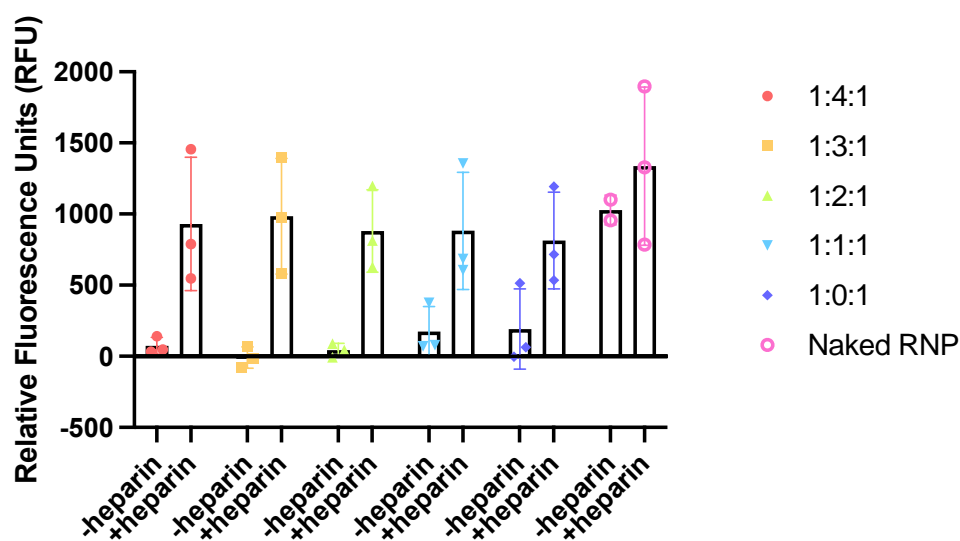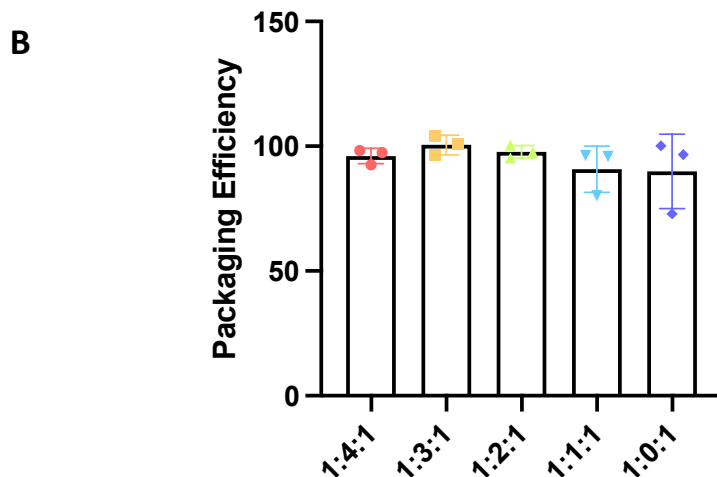
Fig. 5.5: Packaging efficiency of RTNs with differing ratios of Peptide Y. RNP and RTNs were assembled with decreasing amounts of targeting peptide, diluted to 100 uL with Tris buffer (50 mM, pH 7.4), and incubated with RiboGreen at a 1:1 v:v ratio per manufacturer's instructions. 0.5 U/mL heparin was included or excluded as indicated (where heparin was included, the corresponding volume of Tris was removed as a diluent in order to keep all concentrations consistent).

A) RiboGreen fluorescence before and after heparin addition to RTNs with differing lipid:peptide:cargo ratios (such that the overall amount of RTN varied accordingly to altered amounts of indicated components). Plotted as mean ± SD. Each point is the average of six replicate formulations from n = 3 experimental repeats (or n = 2 repeats, for naked RNP condition).

B) Quantification of encapsulation efficiency. Efficiency was calculated using the following formula: efficiency = 100 − (RNP in RTN without heparin treatment / naked RNP without heparin treatment). Plotted as mean ± SD. Ordinary one-way ANOVA; no significant differences found.

No significant difference was detected in packaging efficiency differing ratios of

targeting peptide, suggesting that peptide neither enhances nor decreases packaging of

RTNs when a 1:1 weight ratio of lipid:RNP is used.

However, peptide-only nanoparticles have also been used in the past [356]. In order

to validate whether peptide alone could contribute to packaging in the absence of lipid, the

experiment was repeated, this time keeping peptide ratio consistent (4:1 peptide:RNP) and

titrating down the amount of lipid included (Fig. 5.6).

A

B
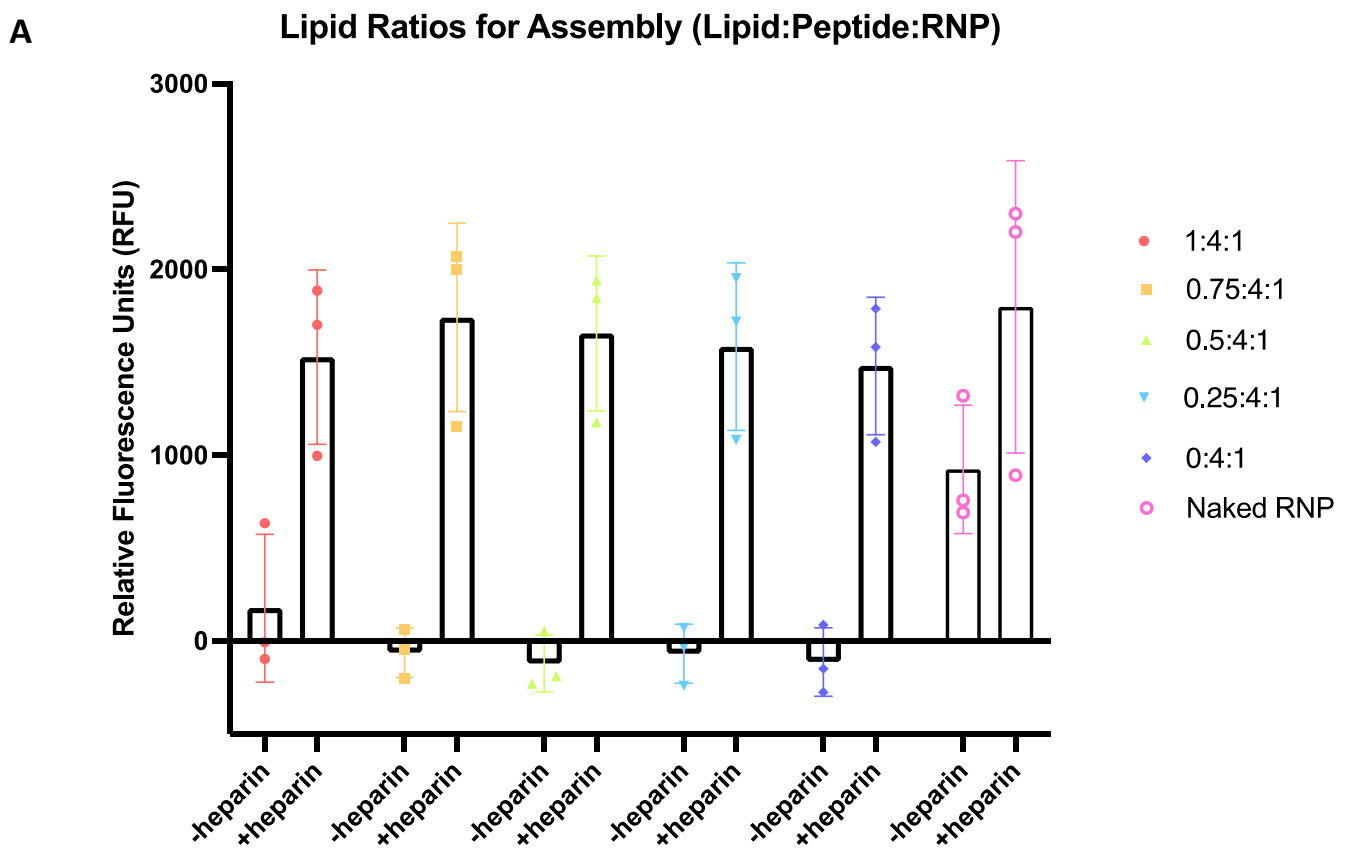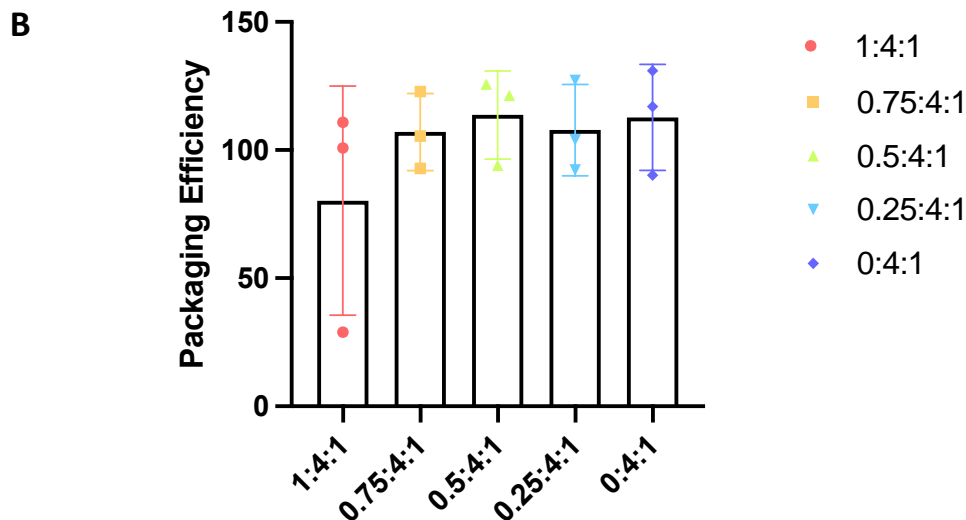
Fig. 5.6: Packaging efficiency of RTNs with differing ratios of C16DOPE. RNP and RTNs were assembled with decreasing amounts of lipid, diluted in 100 uL Tris buffer (50 mM, pH 7.4), and incubated with RiboGreen at a 1:1 v:v ratio per manufacturer's instructions. 0.5 U/mL heparin was included or excluded as indicated.

A) RiboGreen fluorescence before and after heparin addition to RTNs with differing lipid:peptide:cargo ratios. Plotted as mean $\pm$ SD. Each point is the average of six replicate formulations from one of two (for 1:4:1) or three (for all other conditions) biological repeats.

B) Quantification of encapsulation efficiency, calculated as follows: efficiency = 100 – (RNP in RTN without heparin treatment / naked RNP without heparin treatment). Plotted as mean $\pm$ SD. Ordinary one-way ANOVA; no significant differences found.

In the first experimental repeat, the 1:4:1 lipid ratio appeared to have only 29% encapsulation efficiency of RNP. This directly contradicts previous encapsulation

experiments' results, in which a 1:4:1 ratio was sufficient for >90% encapsulation efficiency (Figs. 5.5 and 5.6); it also contradicts the subsequent two experimental repeats of this same experiment, in which encapsulation efficiency was shown to be ~100% (leftmost column on graph in Fig. 5.6B). Therefore, it is most likely that this aberrant result was due to human error.

Much like the 1:1 ratio of lipid:RNP, a 1:4 ratio of peptide:RNP was sufficient to prevent RiboGreen staining, with a calculated packaging efficiency of >90%. The run of positively charged amino acids (K_{16}) present in the peptide are likely sufficient to allow for total encapsulation of anionic cargo. This has been observed before with nucleic acid cargo [356]. However, experiments by a previous Hart lab member, Dr. Amy Walker, showed that a 1:1 lipid:RNP ratio was optimal for transfecting RNP into NHBE BMI-1 cells [324], and therefore this ratio was used for all future experiments.

Taken together, these experiments indicate that both lipid and peptide are individually capable of packaging RNP cargo, as measured by quenching of fluorescence from labelled gRNA. Quenching was achieved in part by complexing with Cas9, and completely by incorporation into lipid-peptide RTNs. Treatment of intact RTNs with the densely negatively charged molecule heparin successfully induced release of RNP cargo, supporting the idea that RTNs are held together primarily by electrostatic interaction between cationic lipid, cationic section of peptide, and anionic cargo. Therefore, the order of mixing of each charged component may have an impact on RTNs' stability in solution.

5.3.1.4 Mixing order and incubation time of RTN components

We next wished to ascertain how mixing order and incubation time of different RTN components could affect their stability by analysing their biophysical properties, as low PDI

and consistent size are indicative of stable (*ie* unlikely to dissociate) RTN formation [350].

RTN size and PDI were measured using a Malvern ZetaSizer at six timepoints after initial

mixing, with intervening incubation taking place at room temperature (Fig. 5.7).
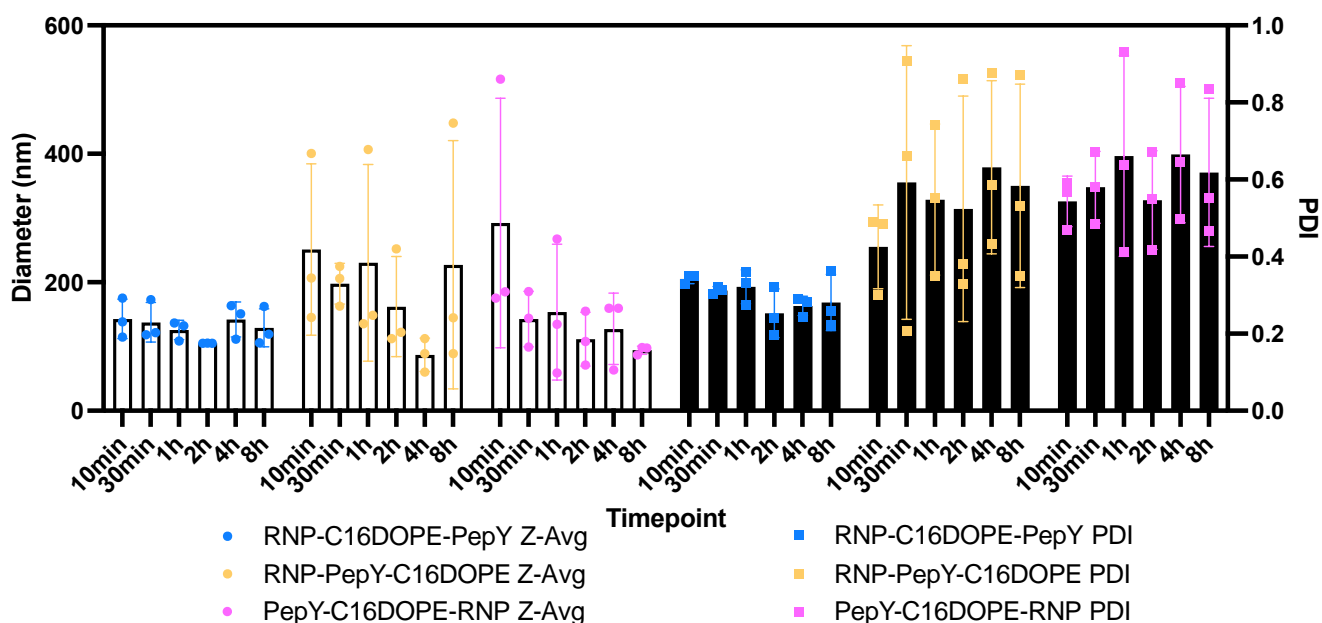
Fig. 5.7: Size (diameter in nm) and PDI of RTNs at each timepoint for each mixing order indicated. RTNs were mixed at a 1:4:1 lipid:peptide:RNP mass ratio in each of the orders indicated and incubated at room temperature for the indicated times before ZetaSizer analysis. Each point represents the average of three replicate measurements from one of three biological repeats. Plotted as mean $\pm$ SD. PDI is a unitless measure calculated via the square of the standard deviation of particle diameters divided by the mean particle diameter.

No significant difference was detected in mixing orders at each timepoint (two-way ANOVA test with multiple comparisons). However, the variation across technical replicates

and repeats was very high for the mixing orders RNP-peptide-lipid and peptide-lipid-RNP as compared to RNP-lipid-peptide. Additionally, the PDI for mixing orders RNP-peptide-lipid and peptide-lipid-RNP was greater than 0.3 in nearly all cases, meaning that multiple subpopulations of different sized particles were present in these populations. This indicated that RTNs mixed in the order RNP-lipid-peptide were more inherently stable than the other mixing orders, and did not dissociate as readily in solution.

	RNP-Lipid-Peptide					
	Peak 1		Peak 2		Peak 3	
	Contribution	Size (nm)	Contribution	Size (nm)	Contribution	Size (nm)
Replicate 1	100%	130.9	N/A	N/A	N/A	N/A
Replicate 2	96.4%	121.9	3.6%	4175	N/A	N/A
Replicate 3	97.5%	117.6	2.5%	4393	N/A	N/A

	RNP-Peptide-Lipid					
	Peak 1		Peak 2		Peak 3	
	Contribution	Size (nm)	Contribution	Size (nm)	Contribution	Size (nm)
Replicate 1	50.6%	405.2	28.7%	49.28	20.6%	4301
Replicate 2	59.3%	320.5	25.8%	36.79	14.9%	4768
Replicate 3	70.6%	1203	29.4%	51.66	N/A	N/A

	Peptide-Lipid-RNP					
	Peak 1		Peak 2		Peak 3	
	Contribution	Size (nm)	Contribution	Size (nm)	Contribution	Size (nm)
Replicate 1	51.6%	346.4	37.2%	49.93	11.2%	4683

Replicate 2	66.2%	416.3	30%	47.64	3.8%	4644
Replicate 3	59.7%	414.1	34.2%	54.28	6.1%	4448

Table 5.1: Relative abundance of different sized RTNs in the total population from one representative timepoint (2 hours). RTNs were mixed at a 1:4:1 lipid:peptide:RNP mass ratio in each of the orders indicated and incubated at room temperature for two hours before ZetaSizer analysis. Three replicate measurements from one independent experiment shown per condition.

A

Size Distribution by Intensity

RNP-Lipid-Peptide

RNP-Peptide-Lipid

Peptide-Lipid-RNP

Intensity (relative %)

Size (diameter, nm)
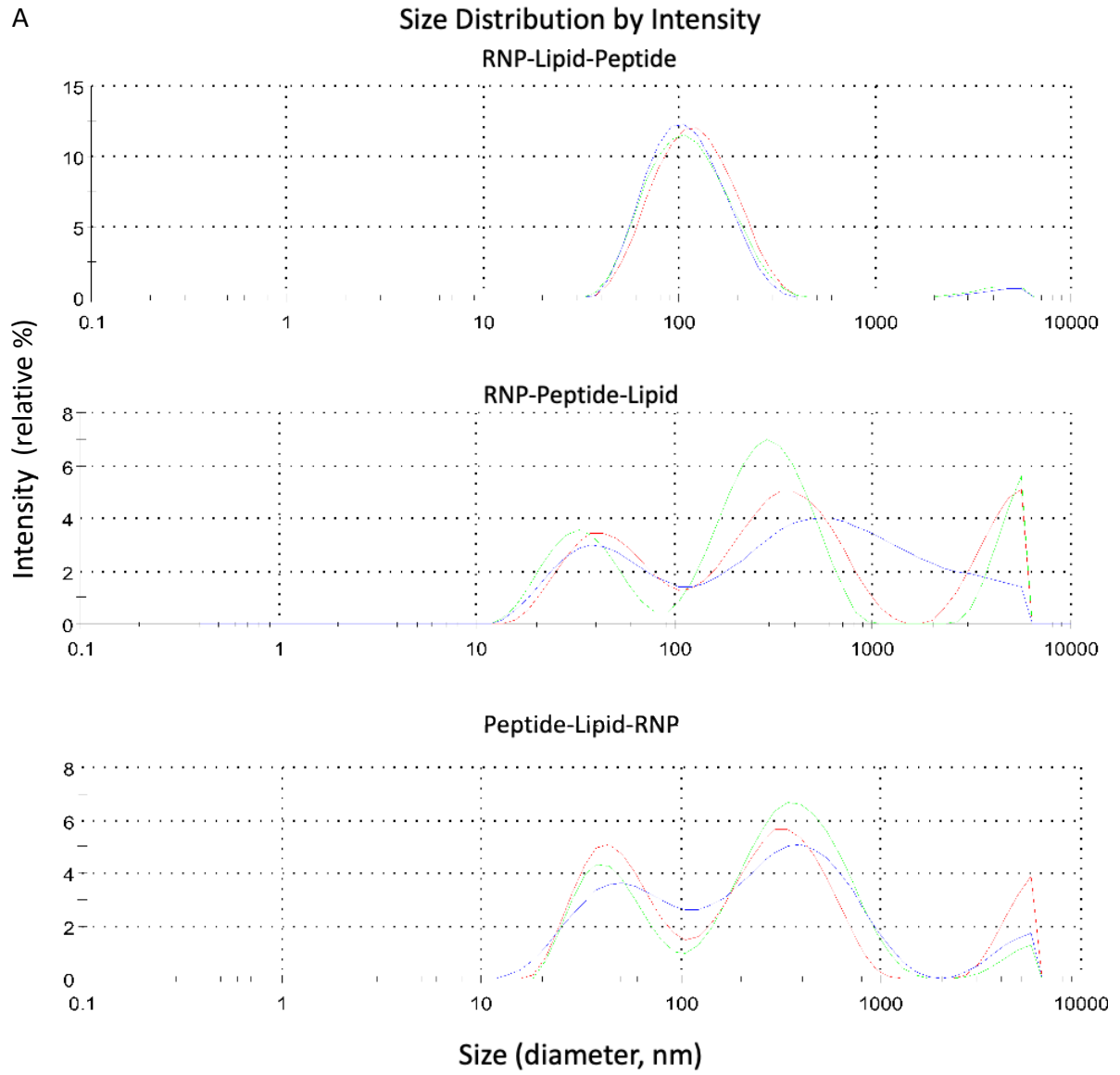
B
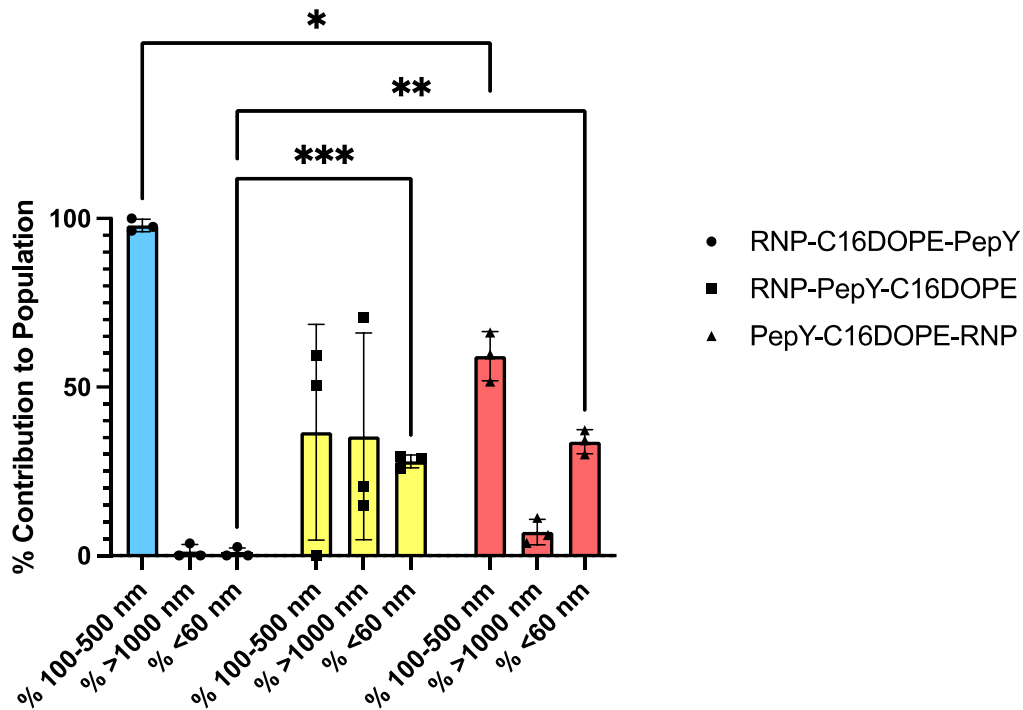
Fig. 5.8: Analysis of size distribution at different mixing orders. RTNs were mixed at a 1:4:1 lipid:peptide:RNP mass ratio in each of the orders indicated and incubated at room temperature for two hours (timepoint chosen at random) before ZetaSizer analysis. A) Sample graphs displaying contribution of different sized particles to overall population. Three replicate measurements, indicated by three different coloured lines, overlaid per mixing order. "Intensity" refers to the percentage of RTNs at a certain size that contribute to the overall population (100% = area under the curve for the entire graph). B) Percent contribution of different subpopulations to the overall RTN population. n = 3 replicate measurements. Plotted as mean $\pm$ SD. Two-way ANOVA with Tukey test. * $p < 0.05$. ** $p < 0.005$. *** $p < 0.0005$.

Upon further investigation, it was found that an average of 35.36% or 7.03% of the particles detected in the RNP-peptide-lipid and peptide-lipid-RNP populations, respectively,

were contributed by populations presumed to be aggregates (>1000 nm). An average of 27.97% (RNP-peptide-lipid) or 33.8% (peptide-lipid-RNP) were <60 nm. On the other hand, the mixing order RNP-lipid-peptide yielded a near-uniform population of RTNs, with >97% of detected particles across three replicates falling between 117.6 and 130.9 nm diameter. Therefore, the mixing order RNP-lipid-peptide was used in all subsequent experiments.

In order to ascertain whether RTNs possessed cationic charge, RTNs were formulated in the mixing order RNP-lipid-peptide at a ratio of 1:4:1 lipid:peptide:RNP and incubated for 30 minutes at room temperature before charge analysis via ZetaSizer. Alongside these RTNs, a commercially available cationic lipid-based transfection reagent (Lipofectamine CRISPRMax) was also assayed for charge (Fig. 5.9).
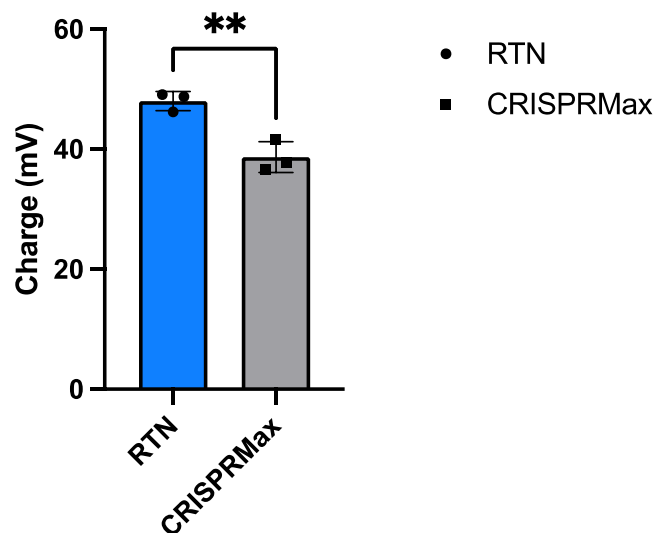
Fig. 5.9: ZetaSizer charge analysis of nanoparticles formulated as RTNs or using commercially available cationic lipid-based transfection reagent. RTNs were formulated using a 1:4:1 ratio of C16DOPE:PepY:RNP and in the mixing order RNP-lipid-peptide, and were allowed to self-assemble at room

temperature for 30 minutes. n = 3 technical replicate measurements. Plotted

as mean $\pm$ SD. Unpaired t-test. ** *p* = 0.0061

RTNs were found to have an overall cationic charge (~48 mV), indicating that the

cationic charge of lipids and peptides outweighed the anionic charge of the RNP. A zeta

potential of +40 to +60 mV has been previously reported to be effective in delivery of RNA

[357]. Additionally, uniform positive charge means that particles will repel one another,

further preventing aggregation.

Having shown evidence of small size, small PDI, and cationic charge of RTNs, it was

anticipated that they would be efficient in transfecting cells.

5.3.2 RTN-mediated transfection

5.3.2.1 Role of targeting peptide in transfection efficiency

Having shown that packaging efficiency is unaffected by inclusion of targeting

peptide, we investigated how targeting peptide affects transfection efficiency (Fig. 5.10).

Luciferase-encoding plasmid was transfected into 16HBE cells using RTNs composed of

different lipids (C14DOPE, C16DOPE, and C18DOPE at a 1:1 ratio of cationic:DOPE)

conjugated with different ratios of Peptide Y. RTNs were mixed in low-serum OptiMEM

medium at a 1:4:1 ratio of lipid:peptide:RNP and allowed to self-assemble for 30 minutes at

room temperature; OptiMEM containing RTNs was applied to cells in replacement of

standard growth medium, centrifuged gently (five minutes at 1000 x g) to pull RTNs to cell

surface, and incubated at 37°C for 4 hours before replacement with growth medium as

described in Methods.
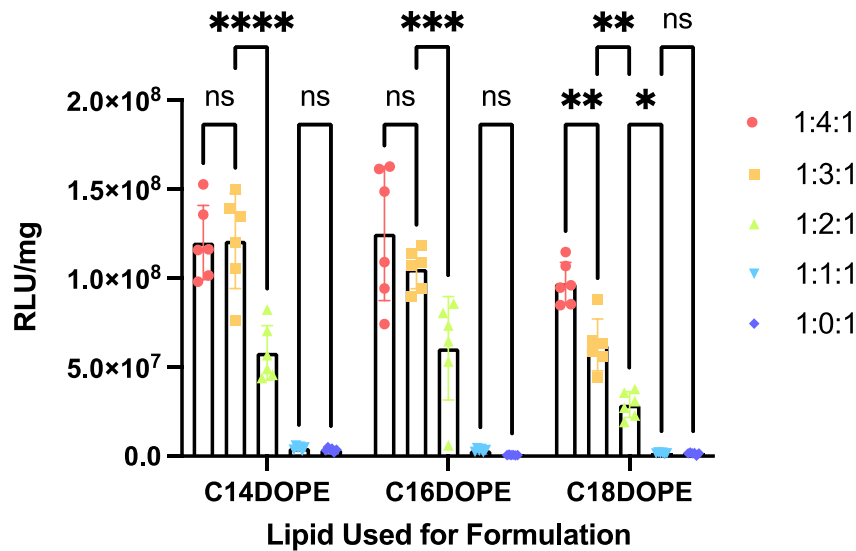
Fig. 5.10: Transfection efficiency of RTNs formulated with differing ratios of lipid (as indicated):PepY:RNP. Luciferase-encoding plasmid was transfected into 16HBE14o- cells using RTNs. Cells were lysed 24 hours post-transfection and lysate incubated with luciferase substrate before luminescence was quantified via plate reader (detailed in Methods section 2.4). Luminescence was subsequently standardised to total protein concentration as determined by BCA assay. Plotted as mean $\pm$ SD. n = 6 replicate wells per condition. Two-way ANOVA with multiple comparisons. ns $p > 0.05$; * $0.01 < p < 0.05$; ** $0.001 < p < 0.01$; *** $p < 0.0001$

Peptide Y was found to significantly enhance transfection efficiency, as its inclusion at higher ratios progressively enhanced transfection efficiency until reaching a plateau (Fig. 5.10). Due to its discovery via phage display, it is presumable that Peptide Y has a positive association with some receptor present on the surface of airway epithelial cells. Because Peptide Y does not have high sequence homology to any known airway cell receptor

agonists, it is unknown which receptor precisely is being targeted. Regardless, the highest transfection efficiency was achieved using RTNs composed of C16DOPE lipid with a lipid:peptide:cargo ratio of 1:≥3:1. The efficiency appeared to increase between 1:3:1 and 1:4:1, although not to a statistically significant degree, suggesting that this was nearing a plateau of efficiency. Therefore, all subsequent experiments were performed using the ratio 1:4:1.

However, performance in submerged cells does not necessarily translate to the *in vivo* environment (or models thereof). Airway epithelial cells differentiated on air-liquid interface (ALI) culture form tight junctions, reducing cell membrane surface area available to RTN; furthermore, physical barriers such as cilia and mucus are present on ALI but not in submerged culture [59, 291]. Specifically, the mucus produced by CF variant (CFBE) cells is significantly more viscous than normal (NHBE) cells [358]. Therefore, it vital to assess RTNs' motility in mucus.

5.3.3   Mucus penetration

5.3.3.1 Mucus penetration capability *in vitro*

Mucus samples were obtained from ALI cultures of the variant genotype F508del or WT CFTR. Firstly, a BCA assay was performed in order to investigate inherent differences in protein quantity in the respective mucus samples (Fig. 5.11). Due to the dehydration of mucus, the same amount of mucosal proteins produced by epithelial cells will exist in a smaller volume of liquid, thus increasing their concentration. These negatively charged mucosal proteins could potentially retard the movement of cationic RTNs through mucus.

Protein Concentration in Mucus
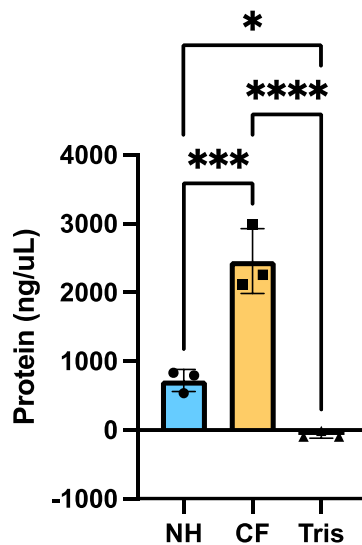
Fig. 5.11: BCA assay on mucus from normal human (NH) or CFTR variant (CF) ALI cultures. n = 3 replicate wells per condition. Plotted as mean $\pm$ SD. Two-way ANOVA with multiple comparisons. * $0.01 < p < 0.05$; *** $0.0001 < p < 0.001$; *** $p < 0.0001$

As expected, CF mucus contained significantly higher concentrations of general protein than NH mucus, presumably due to dehydration from osmotic dysregulation typical of CF (*ie* normal amounts of protein secreted by cells, but into a smaller volume). Also of note, CF mucus, which can be pictured as a gel containing matrices of proteins and sugars, has been shown to be smaller than that of NH mucus. The pores in CF sputum have been estimated at around 100-200 nm in diameter via analysis of penetration of non-mucoadhesive polymer beads [359]; other microscopy-based methods approximated different pore sizes (in one study, between 160 and 1440 nm, with much variation therein), but these results contradicted penetration assays, and were theorised to be inaccurate due to alteration of mucosal matrices by sample preparation for microscopy [358, 360]. The RTNs used in this

assay were ~120-130 nm in diameter according to ZetaSizer measurement (see Table 5.1), and therefore close to the theoretical steric limit for passage through the smallest predicted pore size of CF mucus.

In order to assess how CF mucus affected RTN diffusion, a mucus penetration assay was performed by coating porous transwell inserts with either normal human (NHBE) or cystic fibrosis (CFBE) mucus (isolated from ALI cultures) and applying RTNs (Fig. 5.12). As RTNs passed through the mucus and transwell over time, they accumulated in the Tris buffer (50 mM, pH 7.4) located in the basal chamber. Fractions were taken from the basal chamber at timepoints over several hours, stained with RiboGreen (before and after heparin dissociation), and fluorescence detected on a plate reader. As a positive control, RTNs were pipetted directly into the basal chamber buffer before fraction collection and measurement.

A
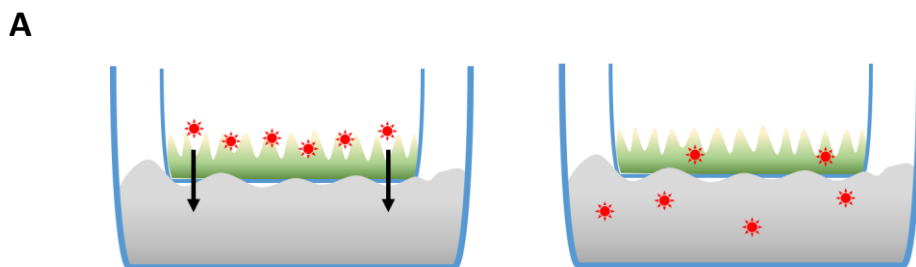
B

Pre-Heparin Treatment

Post-Heparin Treatment

- NH mucus / RNP in RTN
- NH mucus / naked RNP
- CF mucus / naked RNP
- CF mucus / RNP in RTN
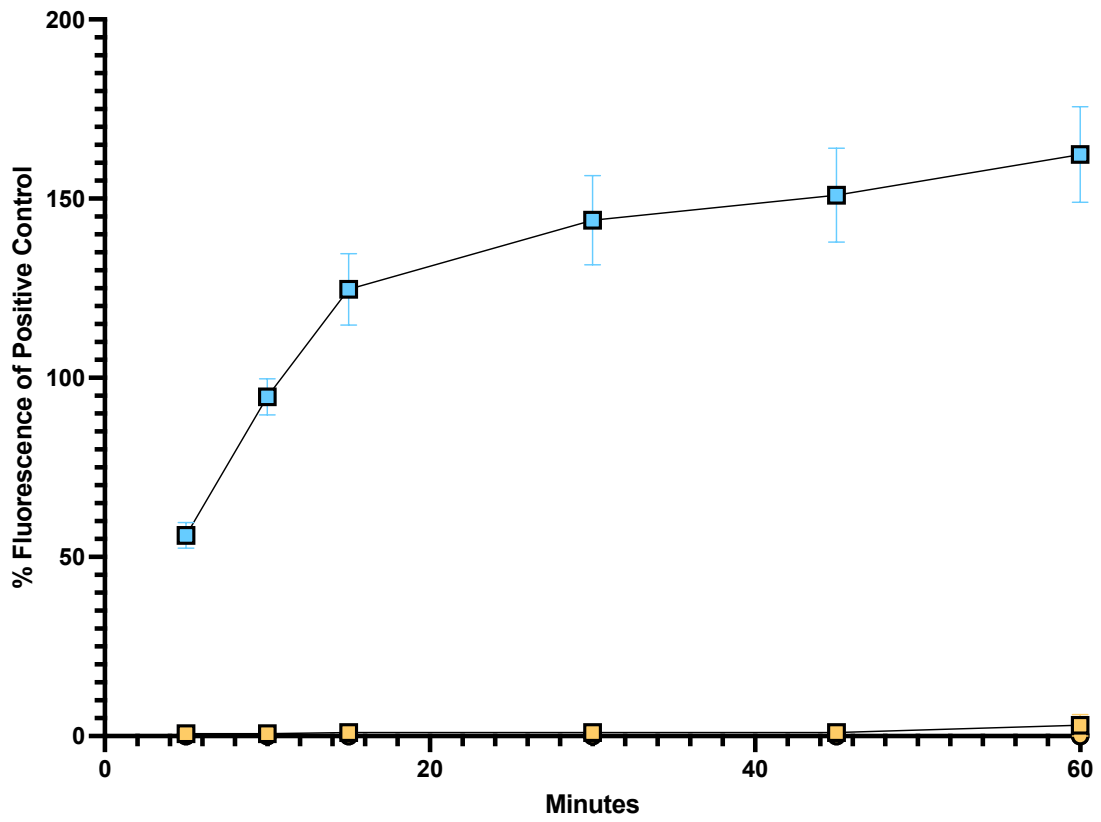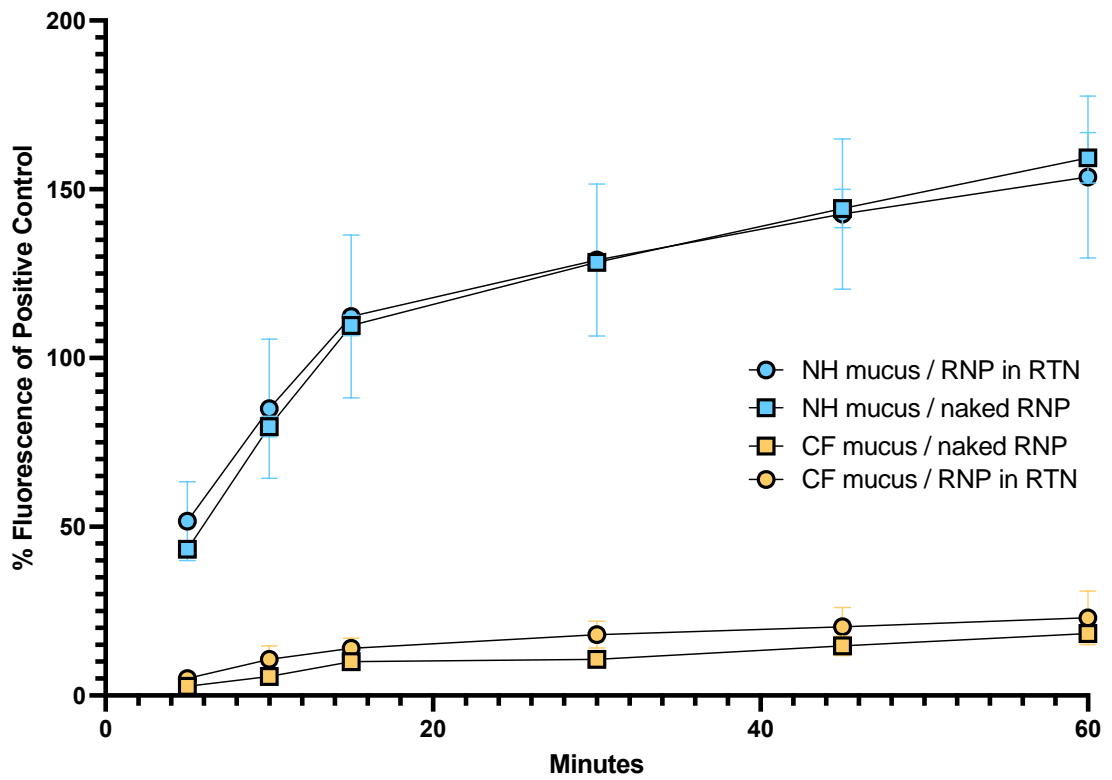
Fig. 5.12: *In vitro* mucus penetration assay.

A) Schematic of assay. RTNs (red) added to transwells coated in 2 μL of either CFBE or NHBE mucus (green) and allowed to pass into Tris buffer (grey) over time.

B) Mucus penetration of RTNs. RTNs were formulated as normal, applied to the mucus-coated transwells (750 ng RTN per well; 1:4:1 ratio of C16DOPE:PepY:RNP), and collected from basal chamber (containing Tris buffer, 50 mM pH 7.4) at timepoints indicated. Collected volumes were split in two, with half dissociated by heparin and half left intact, before incubation with RiboGreen and measurement on plate reader. Fluorescence detected at each timepoint (detected via FLUOStar Optima with excitation filter wavelength 492 nm, emission filter wavelength 520 nm, and gain 2000) was divided by the fluorescence detected in positive controls (750 ng RTN injected directly into basal chamber buffer and incubated with RiboGreen) and multiplied by 100 to calculate % fluorescence of positive control.

Naked RNP was freely mobile in NHBE mucus, with the majority of particles passing through the transwell membrane within the first 20 minutes. Although no fluorescence was detected from media containing RTN-encapsulated RNPs, upon treatment with heparin, fluorescence was detected at levels nearly identical to those of naked RNP. This suggested that RTNs not only successfully penetrated mucus, but also remained fully intact for the entirety of their journey. This is favourable due to the fact that the RTNs' lipid and peptide components are what allow for cell penetration, and therefore transfection efficiency would

likely be significantly hindered if they dissociated upon exposure to mucus. For CFBE mucus, on the other hand, a negligible amount of fluorescence was detected, indicating that particles were unable to traverse mucus and permeate the transwell.

Interestingly, over 100% of positive control fluorescence was seen in RTNs penetrating NH mucus. It was theorised that some mucosal components could be leaking through the transwells, possibly carried by the RTNs themselves, as the cationic lipids could "collect" anionic nucleic acids and/or mucins on their journey through the mucus. However, anionic RNP (unlikely to associate with anionic components) also showed fluorescence of >100% of positive controls. Airway mucus is known to contain DNA, with NHBE mucus containing an average of 10 µg/mL DNA, and CFBE mucus containing >100 µg/mL [361, 362]. Because the mucus used in this experiment was harvested from ALI cultures containing only airway epithelial cells, no neutrophils would likely be present. Therefore, it is likely that the effect seen here is due to mucins rather than nucleic acids; however, in an *in vivo* environment, both anionic components would likely be present. Extraction of proteins from mucus and Western blotting for mucin proteins could potentially elucidate this further.

In order to ascertain whether there was material in the mucus stainable by RiboGreen, mucus alone was dissolved in Tris buffer (50 mM pH 7.4), allowed to equilibrate for 30 minutes at 37°C, and incubated with an equal volume of RiboGreen prepared according to manufacturer's instruction (Fig. 5.13).
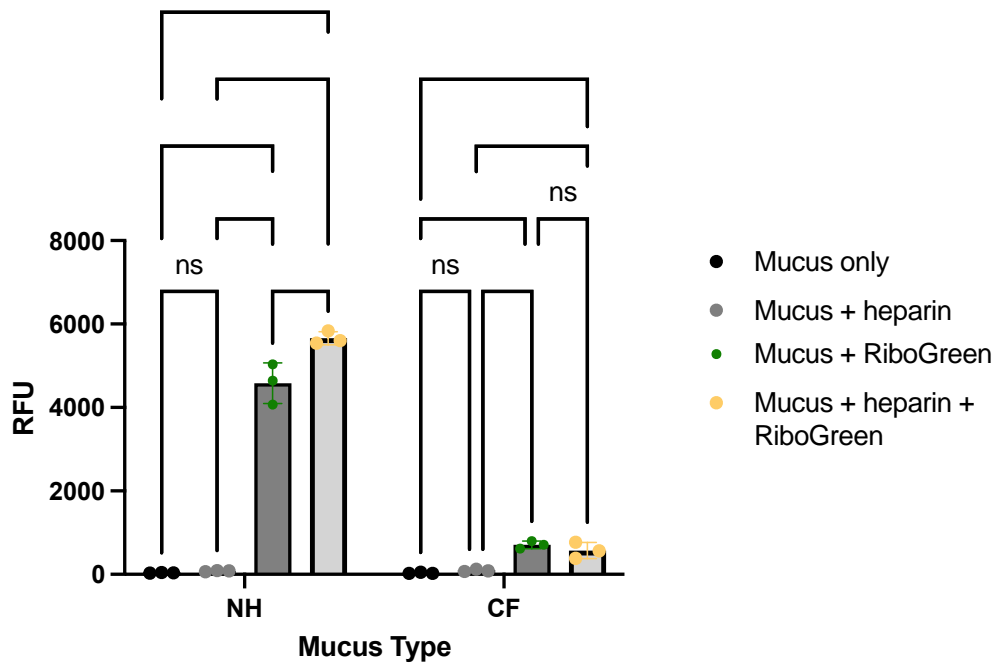
Fig. 5.13: Fluorescence of mucus upon treatment with indicated assay components. Mucus was diluted in 100 uL Tris buffer and allowed to equilibrate for 30 minutes at 37°C before incubation with 100 uL RiboGreen. Plotted as mean ± SD. Two-way ANOVA with multiple comparisons. ns $p > 0.05$; * $0.01 < p < 0.05$; ** $0.001 < p < 0.0$; *** $p < 0.0001$

No autofluorescence was detected in NHBE or CFBE mucus with or without heparin. Upon addition of RiboGreen, significant signal was detected in both types of mucus; however, this increase in signal was markedly higher in the NHBE mucus over the CFBE mucus. Additionally, addition of heparin had a significant impact on fluorescence of NHBE mucus. This phenomenon is difficult to explain given that is unclear what exactly in the mucus is being stained by RiboGreen. Regardless, RTNs' ability to transfect in an *in vivo* environment is reliant on more than solely mucus. Therefore, RTNs were tested for transfection efficiency on ALI-differentiated airway epithelial cells.

5.3.3.2 Mucus penetration and transfection of ALI-differentiated cells

A GFP-expressing airway epithelial cell model had been developed by previous Hart lab member Dr. Amy Walker. Briefly, NHBE BMI-1 cells were lentivirally transduced to express GFP under the hPGK promoter (viral schematic in Appendix). GFP-expressing cells were isolated via FACS to yield NHBE BMI-1 GFP+ cells.

In my hands, these cells were transfected with 1:4:1 C16DOPE:PepY:RNP RTNs at a range of dosages (500-1000 ng Cas9 per well). The RNP consisted of RFP-tagged Cas9 and a gRNA that had previously been shown to effectively silence GFP expression in submerged GFP+ NHBE BMI-1 cells [324].

Twenty-four hours post-transfection, inverted fluorescent microscopy was used to detect RFP (RNP) and GFP (cell) signal. If cells were successfully transfected, then RFP and GFP should colocalize (both transfected and endogenously expressed proteins respectively present in the interior of the cell, with GFP in cytoplasm and RFP in endosomes or nucleus). Fiji's "Plot Z Axis Profile" function was applied to z-stacks in order to identify which slices in each z-stack had the highest RFP and GFP signal. In all but one sample, RFP signal was detected most strongly at z-planes above where GFP signal was strongest, suggesting that the majority of RTNs were not able to penetrate into cells, and remained trapped in airway surface liquid. In the case of the only sample where RFP signal appeared below GFP, the RFP signal colocalised with the collagen coating on the plastic transwell, indicating that the ALI culture had been damaged such that a gap had formed in the cells, allowing RTNs to penetrate below; therefore, this sample was removed from further analysis.
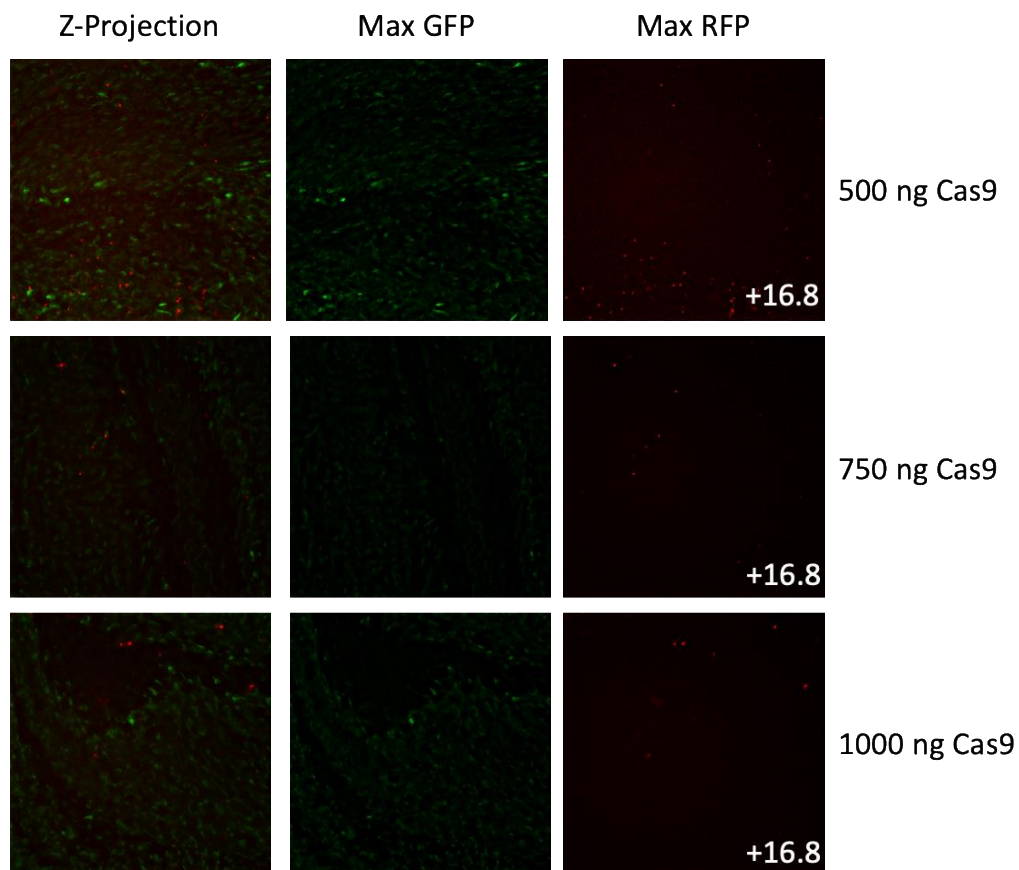
Fig. 5.14: Sample microscopy images. One representative image of 3 replicate ALI wells, containing GFP-expressing NHBE cells, shown per condition. RFP-tagged Cas9 was complexed into RNP, packaged in RTNs at a 1:4:1 ratio of lipid:peptide:RNP in the mixing order lipid-RNP-peptide, and applied to ALI wells. Wells were imaged 24 hours post-transfection.

Z-projection: maximum value from each slice projected into a 2D image. Max GFP: single slice with highest GFP signal. Max RFP: single slice with highest RFP signal. Discrepancy between max GFP and RFP slices shown (i.e. "+16.8" signifies max RFP signal is located 16.8 µm above max GFP signal). Images taken at 20X magnification. GFP excitation wavelength 480 nm and

emission wavelength 509 nm. RFP excitation wavelength 540 nm and emission wavelength 583.

In order to further visualise this, subsections of each image were identified in which RFP signal was visible directly above GFP signal (i.e. where RTNs had been caught in sputum directly above strongly fluorescent cells). Stacks were re-sliced and Z-projected to show cells and RTNs side-on (Fiji macro available in Appendix) (Fig. 5.15).
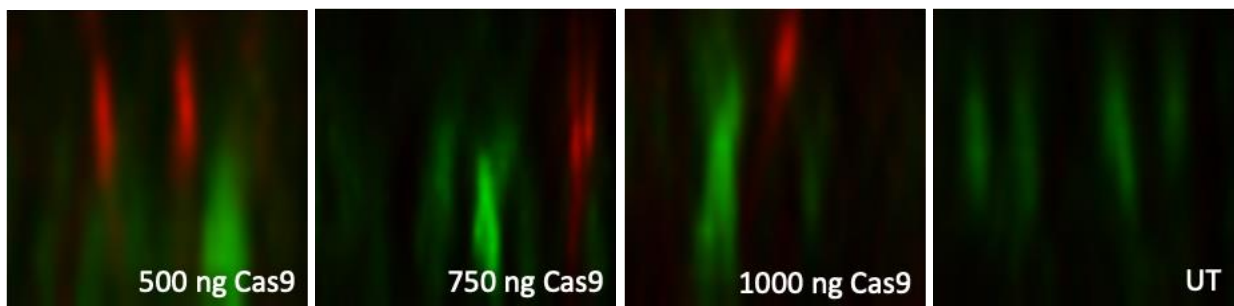
Fig. 5.15: Re-sliced z-projections of RTNs and cells. RFP-tagged Cas9 was complexed into RNP, packaged in RTNs at a 1:4:1 ratio of lipid:peptide:RNP in the mixing order lipid-RNP-peptide, and applied to ALI wells containing differentiated GFP-expressing NHBE cells. Wells were imaged 24 hours post-transfection.

The majority of RFP signal is located above or between cells, suggesting that RTNs did not successfully enter cells to a significant degree. Transfection of submerged cells in parallel using the same RTNs and gRNA could have further confirmed that differentiated cells are inherently more difficult to transfect.

Again, Fiji's "Plot Z Axis Profile" function was applied to these same z-stacks in order to generate a profile of GFP and RFP signal intensity versus height from transwell surface. Additionally, the physical distance between highest-GFP and highest-RFP slices in each z-stack are plotted below, representative of the physical distance between cells and RNP.
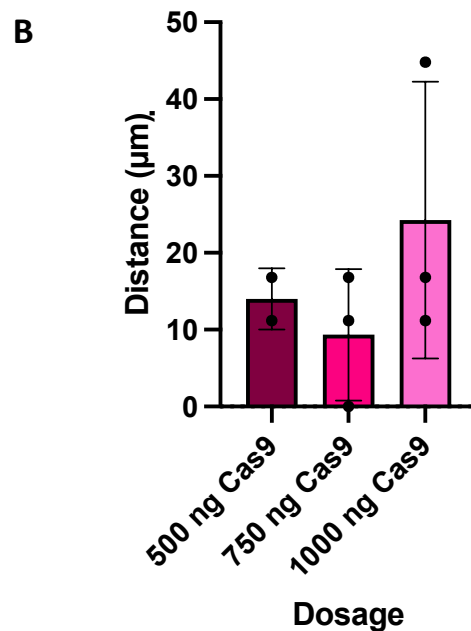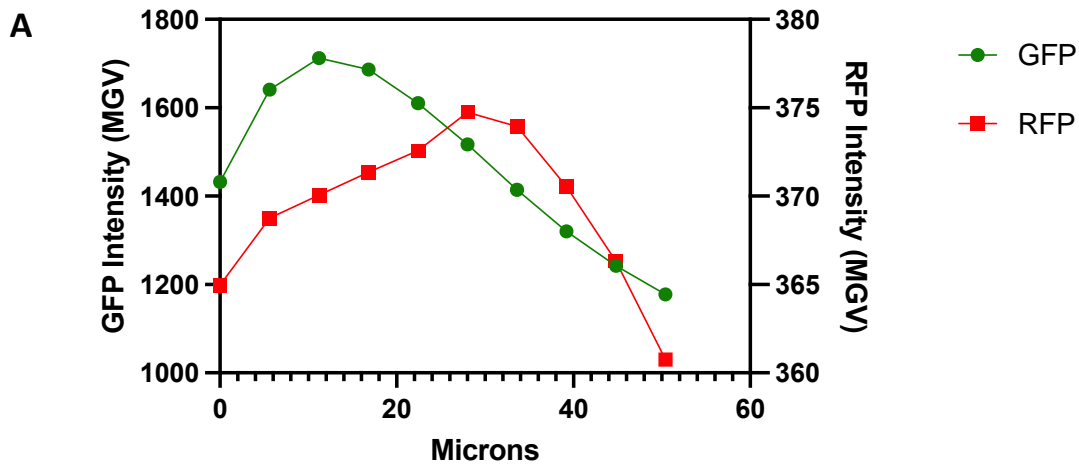
Fig. 5.16: Distance between max RFP and GFP signal in z-stacks of ALI wells containing GFP-expressing NHBEs imaged after transfection. RFP-tagged Cas9 was complexed into RNP and then packaged in RTNs at a 1:4:1 ratio of lipid:peptide:RNP in the mixing order lipid-RNP-peptide. RTNs were applied to ALI cultures, and wells were imaged 24 hours post-transfection.

A) Z-profile plot from one representative well. Average signal intensity of each slice for each channel (RFP or GFP) was calculated via Fiji's "Plot Z Axis Profile" function and plotted. MGV = mean grey value.

B) Distances between max GFP signal and max RFP signal. n = 3 replicate transfected wells (n = 2 for 500 ng Cas9; one well removed from analysis due to bacterial contamination). Plotted as mean $\pm$ SD. One-way ANOVA with Tukey test; no significant differences detected.

In order to ascertain whether a significant number of cells were eventually transfected and GFP silenced, these same transfected ALI wells were subsequently maintained for 8 days (to allow pre-existing GFP protein to be degraded in the event of successful knockout) and re-imaged using confocal microscopy. Fiji was used to calculate GFP signal from average Z-projections.
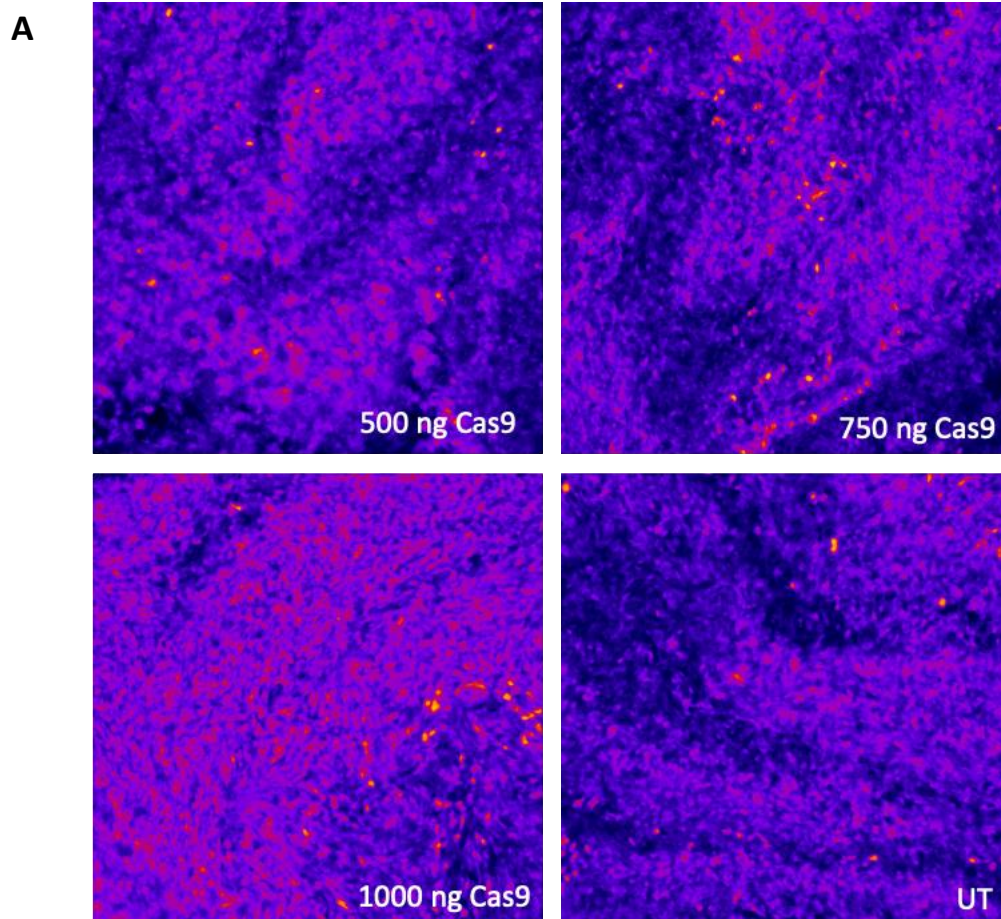
Pixel value (fluorescence intensity): 0 ▮▮▮▮▮▮▮ 65535
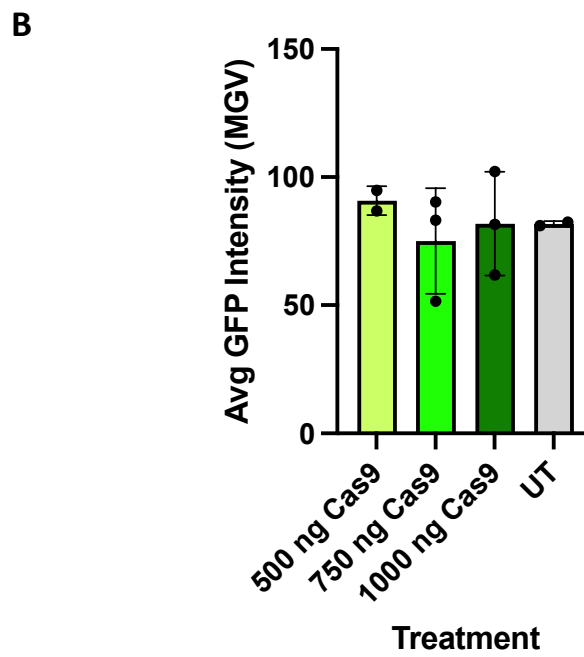
Fig. 5.17: GFP signal 8 days post-transfection.

A) Sample Z-projections of wells with different RFP-Cas9 RNP dosages, shown as heat maps of GFP fluorescence intensity (Fiji "Fire" lookup table). Scale from pixel value 0 (black) to 65535 (white, for 16-bit image) as a readout of GFP intensity. One well shown per condition. RTNs were assembled at a 1:4:1 ratio of lipid:peptide:RNP in the order lipid-RNP-peptide, applied to wells, and imaged 8 days post-transfection. MGV = mean grey value.

B) Average signal intensity from z-projections. n = 3 replicate wells for 750 and 1000 ng Cas9; n = 2 replicate wells for 500 ng Cas9 and untreated (UT). Plotted as mean $\pm$ SD. One-way ANOVA followed by Tukey test; no significant differences detected.

No significant reduction was detected between treated and untreated samples (calculated in Fiji via analysis of mean grey value and integrated density), suggesting that transfection of ALI cultures was not efficient. Because identical RTNs (1:4:1 C16DOPE:PepY:RNP) were previously shown to be freely motile in NHBE-derived mucus (Fig. 5.12), and successfully transfected submerged non-CF airway epithelial cells, it is unlikely that mucus alone is responsible for the RTNs' inability to transfect cells.

5.4 Discussion

Delivery of DNA editing machinery is a vital consideration when the aim of editing is disease treatment. We therefore sought to characterise RTNs' biophysical properties *in vitro* and test their performance in transfecting both submerged and ALI-differentiated airway epithelial cells.

Firstly, RTNs' ability to encapsulate RNP was examined using a fluorescent dye assay. It has previously been shown that RiboGreen dye can be used to stain RNA cargo, but that packaging in RTNs shields this staining, resulting in no detectable fluorescence [288]. Additionally, another group has used the RIboGreen assay to stain ssODN donors complexed with Cas9 RNPs [363]. However, to my knowledge, no group has adapted this assay to stain the gRNA in CRISPR RNPs. Using this approach, we showed that gRNA is stainable to a detectable level, even when complexed with Cas9; association into an RNP reduces fluorescence by only ~50%. Furthermore, this fluorescence is completely eliminated by packaging of RNP into RTNs. When examining the role of different RTN components (lipid vs. targeting peptide), it was shown that either component is sufficient to shield RNP from RiboGreen, provided that other component is present in sufficient proportion. Multiple studies have used either lipid or peptide to formulate RNPs capable of transfecting cells with nucleic acid cargo [356]. However, both components were required for effective transfection efficiency.

Small and uniform size distribution is favourable in RTNs. Smaller particles are more motile in matrix-like mucus gels [359], and more readily taken up by endocytosis [364]. Uniformity of size (measured via polydispersity index, or PDI) is also necessary for consistency of transfection, and is considered a critical element of quality control for clinical

application [350]. Therefore, we examined how different aspects of RTN formulation could affect these properties. Mixing order of components and incubation time were examined for RTNs consisting of lipid, targeting peptide, and cargo, and the optimal mixing order was identified as RNP-lipid-peptide. RTNs mixed in this order have previously been imaged via cryogenic transmission electron microscopy (cryo-TEM), revealing that empty liposomes formed spherical structures; upon addition of RNP, multi-lamellar structures were formed; addition of peptide resulted in darker patches on lamellar surfaces, which were theorised to be the K_{16} portion of the peptide inserting into the lipid bilayer [324]. No other mixing orders were imaged using cryo-TEM.

We next investigated RTNs' ability to penetrate mucus, as *in vivo* application would require. We had initially theorised that movement of cationic RTNs would be retarded by interaction with anionic mucosal elements, particularly with negatively charged mucin proteins present at higher concentrations in CF mucus. If this electrostatic interaction alone was sufficient to retard RTN diffusion, then anionic naked RNP would move more freely through both CFBE and NHBE mucus than cationic RTNs. However, naked RNP was also unable to penetrate CF mucus despite moving freely in NH mucus, confirming that there is more than electrostatic inhibition at play. Another factor likely to impact RTNs' reduced mobility in CF mucus is the size of pores in CF mucus gel matrix, as described in Section 5.3.3.1. CF gel matrix pores have been measured at 100-200 nm in diameter, where our RTNs were 120-130 nm in diameter (as measured by ZetaSizer, reported in Table 5.1) [359].

Also unexpectedly, we saw more than 100% of the fluorescence of the positive controls in NHBE mucus conditions. In the experiment, fluorescence of RTN and naked RNP were injected directly into basal chamber buffer and measured for fluorescence to serve as a positive control for 100% translocation through the mucus-coated transwell. The increase in

fluorescence above 100% of positive controls in the NHBE samples may have been due to anionic mucosal elements clinging to the cationic RTNs on their passage through mucus into the basal chamber; the lack of penetration of RTNs through CF mucus likely eliminated them as a vehicle for transport of mucosal elements stainable by RiboGreen. However, two points challenge this theory. Firstly, RTNs collected after passage through NHBE mucus did not display any RiboGreen staining until they were dissociated with heparin; if RiboGreen was staining elements that had clung to the outside of the RTN, then it would be visible pre-dissociation. Secondly, naked RNP also displayed >100% staining compared to positive controls; anionic RNP would not attract anionic mucosal elements. Therefore it is unclear why >100% staining was observed. However, the general trend of the data (RTNs moving significantly more freely in NHBE than CFBE mucus) supported our theories about CF mucus retarding RTN movement.

Importantly, if association with anionic mucosal elements neutralises the RTNs' cationic charge, this could potentially affect diffusion rates, and the transfection efficiency of cells. We chose to explore this by transfecting mucus-producing ALI-differentiated GFP+ NHBE BMI-1 cells with RTNs containing RFP-tagged Cas9 complexed with a gRNA shown to knock out GFP in submerged cells (data in Appendix) [324]. Upon initial transfection and imaging, several trends were observed. Firstly, endogenous GFP expression in cells was not uniform. This is likely explained by the fact that lentiviral transduction can lead to multiple copy numbers integrating into cells. After initial transduction, the GFP+ population was isolated via FACS by previous Hart lab member Dr. Amy Walker, meaning that any cell fluorescing above a specified threshold was collected, with no differentiation between cells with single and multiple copies of GFP [324]. In the case of ALI culture, in which basal cells multiply before differentiation, this would theoretically result in a mosaic effect, with zones of

brighter and dimmer cells. This could have been mitigated by dilution cloning before ALI culture; however, single-cell cloning of BMI-1 transduced cells is difficult. Ideally, the same field of view would be imaged at each time point (initial post-transfection image and 8 days post-transfection); however, this was beyond the capabilities of the microscopes used. Additionally, general image quality was sacrificed due to A) the long focal distance due to transwell structure, and B) noise introduced by the movement of beating cilia. Another surprising phenomenon observed was a lack of correlation between RFP signal and amount of RFP-tagged Cas9 introduced, as well as noticeable variation in amount of RFP signal observed between replicates receiving the same amount (see Fig. 5.14). This could have been due in part to cilia movement "sweeping" RTNs into areas of higher and lower concentration, making direct comparison difficult. It has been shown before that cilia movement is capable of moving fluorescently labelled beads across the surfaces of ALI cultures; it is reasonable to assume that the same effect would be had on RTNs [365].

Eight days post-transfection, no reduction of GFP signal was detected compared to untreated controls. Transfection of ALI cultures with RTNs has previously been shown to be significantly reduced compared to submerged cells [265]. Possible causes could be tight junctions reducing available cell membrane surface area, and movement of cilia sweeping RTN-containing sputum away from cell surfaces. There is also evidence that transfection of cells by cationic nanoparticles is somewhat dependent on cell cycle [366]; therefore, postmitotic cells (like those making up the topmost layer of ALI cultures) may not take up RTNs as readily, even with the aid of receptor-targeting peptides. However, transfection of ALI cultures with RNA has been reported previously [96]. Therefore, inclusion of eGFP siRNA should be used in future experiments as a positive control to help elucidate whether the lack of transfection was cargo-specific. Additionally, editing efficiency estimated via

microscopy is not the most sensitive assay. Due to time restrictions, no further characterisation of these cells could be performed. However, the data reported would have been strengthened by follow-up analysis of these cells including Western blot of harvested protein to quantify eGFP expression, ICE analysis of harvested genomic DNA, and qRT-PCR of harvested RNA.

Interestingly, as shown in Fig. 3.4, the median [ASL + mucus] height of ALI-differentiated NHBE *BMI-1* cells is ~3 μm above cell surface, and red signal was strongest at a median of 14 μm (see Fig. 5.16). This could potentially be explained by the same ciliary movement that may have "swept" RTNs into patches of higher and lower concentration. It stands to reason that such an effect would also be had on mucus; therefore, areas of high RTN concentration would likely coincide with areas of accumulated globules of mucus that protrude higher off of the cell surface.

We have developed a method for generating small, low-PDI, cationically charged RTNs that are capable of transfecting submerged airway epithelial cells and appear to be motile in NHBE mucus *in vitro*. However, the *in vivo* environment presents more barriers, particularly in the context of CF. The issue of mucus penetration can potentially be addressed by formulating liposomes with polyethylene glycol (PEG), as described in Introduction section 1.8.4. Another potential avenue could be identification of a new targeting peptide identified by biopanning not in submerged cells but in ALI cultures themselves; it is possible and in fact likely that different receptors are expressed in different cell types post-differentiation. Overall, more work is required to optimise RTNs for RNP delivery to the *in vivo* environment.

6. *Conclusions*

6.1 Key findings

The key findings of this thesis are as follows:

- Nasal epithelial cells harbouring the *CFTR* variant causing R709X (CFNE BMI-1 R709X cells) show no CFTR function as measured by ion transport, ASL height, and mucosal spinnability.

- Using homology-independent targeted integration (HITI), a reporter cDNA can be integrated at the AAVS1 genomic safe harbour locus with high efficiency and no off-target effects detected at five *in silico* predicted loci; however, expression from the native promoter fluctuated.

- Cationic lipid-based receptor-targeted nanoparticles (RTNs) successfully encapsulate CRISPR RNP; however, they do not appear to transfect ALI-differentiated cells.

6.2 Routes for further study: characterisation of CF cell models

CF airway pathologies stem from defective ion trafficking, hence the historic dependence on Ussing chamber analysis as a modelling tool [297]. However, there are several other techniques that can assay the downstream effects of dysregulated ion transport. For example, inflexible mucus in the CF airway leads to reductions in mucociliary clearance [32]. This can be assayed via examining movement of fluorescently labelled beads applied to the ciliated surfaces of ALI cultures [306, 365]. Matsui *et al.* took long-exposure images of ALI cultures after application of fluorescent beads [306]. NHBEs images showed vortex-like streaks of fluorescence, suggesting a rotational movement of mucus generated by cilia beating. CFBEs, on the other hand, showed stationary clumps of fluorescence,

suggesting minimal movement of mucus. Wu *et al.* recaptured these results, and additionally found that this effect was partially rescued by application of the CF modulator drug Orkambi [365].

Mucociliary clearance is dependent in part on cilia beating, and CF cultures have been shown to have altered cilia beat frequency (CBF). Allan *et al.* reported that CF ALI cultures showed significantly reduced CBF compared to healthy controls (6.75 ± 0.17 vs. 7.61 ± 0.11 Hz respectively) as measured by high-speed video microscopy [367]. Thomas *et al.* reported similar results by analysing slides containing fresh nasal brushings from CF patients vs. healthy volunteers (9.5 vs. 13.4 Hz respectively) [368]. However, Alikadic *et al.* reported contradictory results with fresh nasal brushings, finding higher CBF in CF patients vs. healthy volunteers (15.4 ± 3.03 vs. 12.92 ± 2.37 Hz respectively) [369]. Rather than directly measuring and comparing CBF, it may be more revelatory to compare the effect of different treatments on CBF in cells of different genotypes. For example, Choccioli *et al.* compared the CBF of washed vs. unwashed ALI cultures of CF variant or normal cells, and reported a higher ΔCBF in CF cultures, suggesting that CF mucus impedes ciliary movement [370]. Schmid *et al.* reported that, despite no significant difference in baseline CBF in normal vs. CF cells, significantly increased ΔCBF was observed in CF cells in response to apical bicarbonate addition ($-23.6 \pm 1.3\%$ in normal cells vs. $-30.3 \pm 1.7\%$ in CF cells) [371].

Additionally, airway models exist beyond ALI culture. Airway epithelial organoids are a more recent development in the field. Similarly to intestinal organoids, these are closed semi-spherical structures whose inside surface represents the apical layer *in vivo*, including ciliated and mucus-producing cells; like intestinal organoids, airway epithelial organoids showed forskolin-induced swelling that was compromised in CFTR variant cells [372, 373]. However, Sachs *et al.* reported variation in FIS from different organoids. This was theorised

to be due to the fact that pools of organoids were not synchronised in differentiation [373]. Therefore, the group created another organoid model derived not directly from nasal brushings, but from ALI-differentiated epithelial layers [374]. Embedding of epithelia in MatriGel induced spherification of cultures that maintained ciliation and mucus production. 48 wells' worth of ALI cultures could be converted thusly and subjected simultaneously to FIS assay, versus the enormous amount of time it would take to analyse this number of wells individually in an Ussing chamber. These organoids were then assayed for their responses to CF modulator drugs, and FIS was found to be enhanced post-treatment.

Exploration of mucociliary clearance and CBF in ALI-differentiated CFNE BMI-1 R709X cells, as well as generation of airway organoids and subsequent FIS assays, could be used to further characterise these CF variant cells' phenotype and provide baselines for treatment responses.

6.3 Routes for further study: Integration of large constructs into the genome

In this thesis, we chose to target the AAVS1 site rather than the endogenous *CFTR* locus. The rationale for this is explained in detail in the introduction to Chapter 4. The editing strategy reported herein could be improved by altering the structure of donor and Cas9 delivered. As previously mentioned, preformulated RNP is advantageous over plasmid delivery due to its transient expression, inability to integrate into the genome, and ability to be paired with synthetic gRNAs modified for enhanced editing efficiency [141, 256]. Additionally, high rates of editing efficiency have been shown using donors packaged in AAV vectors. Vaidyanathan *et al.* have already shown promising results in delivery of CRISPR RNP alongside AAV donor (specifically AAV6) in multiple reports, albeit for HDR rather than HITI. In one study, an AAV6 donor containing homology arms was used to correct bronchial

epithelial cells homozygous for the F508del variant with ~41% efficiency [223]. In another, a

full-length CFTR cDNA was integrated at the endogenous locus using multiple AAV6

constructs with ~72% efficiency after enrichment and expansion of edited bronchial

epithelial cells carrying nonsense mutations [235]. Crucially, in this more recent study, CFTR

cDNA was delivered in pieces via multiple AAV transductions. One of our major motivations

for using a promoterless cDNA (thus placing the insert under the control of the native

PPP1R12C promoter) was to avoid exceeding the packaging capacity of an AAV vector. Using

the multi-virus approach could therefore allow for inclusion of a promoter. Although

Vaidyanathan *et al.*'s work relied on HDR rather than HITI, AAV vectors have also been used

for delivery of HITI donors to non-dividing tissues [375]. Tornabene *et al.* accomplished low

levels of HITI after direct injection of Cas9- and donor-encoding AAVs into pig

photoreceptors; the low editing efficiency was proposed to have been due to low

transduction efficiency.

Recently, three new homology-independent editing strategies were reported. First,

Bazaz *et al.* paired Cas9 with piggyBac transposase (Cas9.PB) in order to integrate donor

molecules in the form of transposons in a technique termed TransCRISTI [376]. In this

approach, the donor molecule is excised from its delivery plasmid via the transposase and

"carried" to the target site of the genome. There, Cas9 cuts as usual, and the transposon

borne by the PB integrates. This is of course an idealised scenario relying on multiple

molecular processes occurring with comparable efficiencies and in a particular order;

however, there are a similar number of factors at play in HITI. In both approaches, donor

DNA must be linearised before or during the presence of a DSB; in cases using a plasmid

donor carrying both desired insert and backbone, donor construct must be cut at both ends

(either cut by Cas9 in HITI or excised by transposase in transCRISTI) for proper excision from

plasmid backbone. Overall, the major advantage of this approach over HITI appears to be the fact that the Cas9.PB fusion allows for proximity of donor molecule to cut genomic DNA. Bazaz *et al.* also attempted insertion of a promoterless construct at the AAVS1 locus. They recovered only ~4% reporter-expressing cells after TransCRISTI; however, TIDE analysis detected nearly 90% indel formation at the target locus, and PCR analysis of single-cell clones returned 72% with the desired insert (compared to 44% using HITI). This represents another dramatic example of a reporter gene's expression under the native *PPP1R12C* serving as an unreliable readout of actual editing efficiency at the molecular level.

The second recent breakthrough is programmable addition via site-specific targeting elements (PASTE) [377]. This technique combines prime editing (discussed in Introduction section 1.3.1.3) with a virus-derived serine integrase to insert up to 36kb at specified sites in the genome. In nature, serine integrases function by depositing DNA containing an *attP* site at sites in the host genome that contain an *attB* attachment point. Because *attB* sites are typically <50 nt, prime editing can be used to edit them into any target site of the genome. Following this insertion of a "beacon" for the serine integrase, a double-stranded circular DNA donor containing an *attP* site can be integrated into the genome. Although editing efficiency was site dependent, Yarnall *et al.* demonstrated insertion of 36 kb into the ACTB locus with ~10% efficiency. In non-dividing primary hepatocytes, a 13.3 kb construct was integrated at ~5%. This exceeded levels of HITI achieved by Yarnall *et al*. In my hands, HITI achieved ~40% efficiency, albeit with a smaller donor construct; additionally, the methods used for detection of integration differ (digital droplet PCR vs. FACS, cloning, and PCR confirmation). Additionally, because I performed PCR on cells enriched for reporter expression, it is unknown the true rates of indel formation in the transfected population; in their hands, indels greatly outweighed successful integration [377].

A third recent breakthrough is click editing (CE), which was not reported until after this thesis's initial submission date [378]. CE bears some similarities to prime editing, but relies on a DNA polymerase rather than a reverse transcriptase to install desired edits. In CE, a gRNA-guided Cas9 nickase (achieved by inducing the H840A mutation in the HNH-like nuclease domain) is fused with two additional agents: a DNA-dependent polymerase (DDP) and an ssDNA-tethering domain (STD). The DDP builds DNA from free dNTPs according to complementarity with an existing ssDNA (occurring naturally during DNA replication). The STD is derived from an HUH endonuclease (named for the histidine-hydrophobic-histidine motif that these endonucleases share in their active sites), which is capable of forming a 5' covalent bond between ssDNA and the protein itself. Fusion of these three agents (nCas9, DDP, and tethering domain) yield the click editor enzyme. The ssDNA (referred to as a clkDNA) tethered to this enzyme is analogous to the PBS and RT template of a pegRNA, as it contains both homology to the nicked strand (to act as a primer) and a template for DDP containing the desired edit. By optimizing the design of ssDNAs to include additional silent substitutions surrounding the desired edit (proposed to help evade DNA mismatch repair machinery), da Silva *et al.* reported 26.6% editing efficiency in HEK293Ts upon transfection of plasmids encoding all CE reaction components. Some variation was observed in editing efficiency at different genomic loci, but rates at different loci could be enhanced via customization of the clkDNA and gRNA. CE was outperformed by PE2, but matched or outperformed PE1 (a version of PE utilising a wild-type reverse transcriptase). The authors therefore suggest that enhancement of the DDP via directed evolution and/or rational design could enhance CE to levels comparable to the highest-efficiency prime editing machinery.

There are many potential permutations of editing strategy and physical structures of Cas9/donor delivered. Many factors must be taken into consideration, including on-target efficiency, off-target indel formation, feasibility and scalability of approach, and risk of apoptosis via DSB formation. PASTE and CE have undergone minimal optimisation since their initial reports, but it is reasonable to assume they will each soon undergo evolutions in the hands of other labs, just as base editors and prime editors did after their initial introduction. Furthermore, if CE is developed to a degree at which it outperforms PE at specific or all loci, it could be used in place of PE to install *attB* sequences for PASTE integration. The appeal of an editing approach capable of inserting enormous amounts of DNA, without generating Cas9-induced DSBs, represents a significant step forward for the field of genetic medicine.

6.4 Routes for further study: Receptor-targeted nanoparticles for *in vivo* delivery

One option is to try a different receptor-targeting peptide. The peptide used herein (Peptide Y) was initially identified via phage display performed on submerged cells [264]. Because differentiated cells have altered expression of receptor proteins [72], it is likely that a use of a different peptide in differentiated cells would be beneficial. This could be accomplished by repeating phage display directly on ALI cultures. There would likely be similar challenges in performing phage display as RTN-mediated transfection. The phage library particles would need to contend firstly with mucus, and secondly with cilia. The mucociliary system evolved specifically to defend airway epithelial cells from invading particles and thus present a major barrier. However, cells used for phage display need

primarily to retain cell receptors, not to faithfully replicate the *in vivo* environment. Therefore, mucus could be removed via thorough washing. Alternatively, mucosal viscosity could be reduced via siRNA-mediated knockdown [379] or CRISPR-mediated knockout of mucin-producing genes. Similarly, phage display could be performed in differentiated cells harbouring mutations that cause ciliary dyskinesia, which exhibit negligible ciliary movement on ALI culture [380].

At the time of this thesis's initial submission, phage display on ALI had not yet been attempted in the CF field. However, the lab of Dr. Debadyuti Ghosh has since reported performance of phage display on ALI-differentiated primary CFBEs pooled from seven different CF patients of the genotype F508del/F508del. A cysteine constrained T7 7-mer peptide-presenting phage library was panned against ALI cultures for five rounds of selection. With each successive round, cells were lysed, internalised peptides recovered, and recovered peptides used in the following round on fresh cells. With successive rounds of selections, peptides tended to grow more positively charged and more hydrophilic, being rich in arginine and lysine. The top-performing peptide (CTSTRKKQC, referred to as Clone C), displayed ~454-fold improvement in uptake by cells compared to naïve library. Clone C was then conjugated into a peptide-lipid with the fatty acid myristic acid. This peptide-lipid was complexed into RTNs with SM-102 (the ionisable cationic lipid used in the COVID-19 vaccine), DSPC (a helper phospholipid comparable to DOPE), and PEG. RTN formulations with Clone C yielded particles with small diameter (~60 nm), low PDI (~0.1), and high encapsulation efficiency (>90%). This size, PDI, and encapsulation efficiency was comparable to RTNs with a peptide-lipid previously shown by the group to penetrate mucus (peptide sequence CPSSSREKC, called CPS, also conjugated to myristic acid); this peptide was identified via phage display on CF mucus alone. However, Clone C significantly

outperformed CPS in transfection of reporter mRNA both in ALI-differentiated CFBEs and in mouse airways *in vivo* (via intratracheal instillation of 40 μL).

It is difficult to directly compare the respective performances of RTNs bearing Clone C and CPS to Peptides E and Y. Firstly, the RTNs used by the Ghosh lab were formulated using different cationic ionisable and helper lipids (SM-102/DSPC vs. DHDTMA/DOPE, respectively). Secondly, the Ghosh lab's RTNs were complexed via direct conjugation of peptides to myristic acid to yield peptide lipids. Conversely, Peptide Y and E contain an N-terminal 16-lysine run of amino acids that allow them to assemble with RTN components via electrostatic interaction, and also possess a linker sequence of three amino acids (GAC).

When examining the properties of the peptides themselves (Table 6.1), Clone C and CPS are smaller (7mers) and more hydrophilic (GRAVY score <-2) than Peptides E and Y (9mers, GRAVY score >-1). Clone C is also more cationic at pH 7, carrying a charge of +3, versus +1 from CPS and Peptide Y. The elevated performance of Clone C in CF mucus counteracts over its less cationic counterparts contradicts my initial hypothesis that more positively charged RTNs might be retarded in mucus, which contains many anionic mucins and mucopolysaccharides. Synthesis of a K_{16}-GAC-Clone C peptide, and/or myristic conjugation of Peptide Y, would allow for direct comparison of each peptide's mucus penetration and transfection capability on ALI culture.

Table 6.1 compares the four peptides; K_{16} and linker, and flanking cysteines are omitted.

Peptide	Phage Protein	Selected on	Sequence	MW (g/mol)	IP (pH)	CpH7	GRAVY
Clone C	T4	ALI CFBEs	TSTRKKQ	874	14	+3	-2.57
CPS	T4	CF mucus	PSSSREK	815.87	11.38	+1	-2.27

Peptide E	M13	Submerged 16HBEs	SERSMNFCG	1056.18	6.14	0	-0.7
Peptide Y	M13	Submerged 16HBEs	YGLPHKFCG	1047.23	9.06	+1	-0.18

Table 6.1: Comparison of peptides from three different phage display experiments. Peptides' properties were calculated based on their amino acid sequences using the Peptide Property Calculator [381] and GRAVY Calculator [382]. MW: molecular weight, IP: isoelectric point, CpH7: charge at pH 7, Sol.: solubility. GRAVY is a measure of hydropathy, with more negative values indicating more hydrophilic properties and vice versa.

In order to avoid the issues of cilia and mucus entirely, editing could be performed *ex vivo* in submerged basal cells, which can then be introduced back into patient airways. This could either be achieved by intratracheal instillation of edited basal cells, or by seeding biocompatible membranes with edited cells for surgical implantation. The former option typically requires preinjury of the airway. For example, Farrow *et al.* pretreated mouse airways with 2% polidocanol in order to strip away the epithelial layer and basal cells [253]. After this acute injury, the airway was seeded with luciferase-expressing basal epithelial cells; luciferase activity was detected in upper airways three weeks after administration. However, eight weeks after administration, luciferase activity was no longer detectable. This suggests that the implanted basal cells did not serve as a permanent substitute for the basal layer of proliferating stem cells. Additionally, given that the CF airway is already prone to inflammation and tissue damage, acute injury is unfavourable to perform in human patients. Chemical preinjury in the ENaC-overexpressing mouse model, which replicates CF-induced lung phenotypes, could be revelatory of the magnitude of the risk involved. The latter

option (surgical implantation) is described by Vaidyanathan *et al.*, who used CRISPR and HDR to correct the F508del mutation in upper airway bronchial cells (UABCs) and embedded them on a porcine small intestinal submucosal membrane (pSIS), which had already been approved for use by the United States Food and Drug Administration (FDA) [255]. Membrane-embedded cells retained their ability to differentiate, meaning that this approach could be used to introduce basal cells to differentiate within the airway *in vivo*, or cells could be differentiated *ex vivo* and a mature epithelial layer implanted. However, this again requires a highly invasive medical procedure in an already vulnerable tissue in CF patients. I believe that repeat administration of RTNs to the airways via nebulisation is a gentler and safer option. Mucus plugs would likely hinder penetration into lower airways on first administration. However, successfully edited cells would begin alleviating the mucosal phenotype, allowing RTNs to penetrate progressively deeper into the lungs with each successive administration.

Delivery of donor via virus is also a promising avenue, as discussed in the future directions for Aim 2 above. Extensive characterisation has been performed of AAV serotypes that successfully target airway epithelial cells. One study reported that AAV6 viruses showed enhanced motility in CF sputum when compared to AAV1 [383]. Using particle tracking, Duncan *et al.* recorded fluorescently labelled viral particles' movement through sputum expectorated from CF patients and found mean squared displacement (MSD1) of AAV6 significantly higher than AAV1 by 3- to 10-fold. However, an obvious drawback of AAV vectors is their immunogenicity [384, 385]. As the airway epithelium is made up of non-dividing cells, repeat administration would be required as edited cells naturally die off and are replaced by unedited cells from the basal layer, a process that typically takes place every 30-50 days [386]. This would not be feasible using AAV vectors. One potential avenue to

address this is by encapsulating anionic viral particles themselves in cationic lipid-based nanoparticles, essentially "hiding" the AAV from agents of the host immune system. The targeting ability of the vector would likely be sacrificed, but this loss could be compensated for via receptor-targeting peptides (such as the one described herein). The intracellular trafficking of AAVs post endosomal escape could potentially then outperform that of released RTN cargo.

In addition to size, PDI, and charge, it would be useful to examine other biophysical properties of RTNs. For example, characterisation of RTNs surfaces could elucidate peptide display. Spectrometry/spectroscopy could be used to elucidate the role of helper lipids in pH-dependent fusogenic phase changes. Three-dimensional size and shape could give further clues as to how RTNs migrate through pores in mucus. Some popular RTN characterisation methods, as reviewed by Mourdikoudis *et al.*, are summarised in Table 6.2 [387].

Category	Examples	Information Gained
X-ray	X-ray diffraction	Crystal structure, composition, crystalline grain size
	X-ray absorption spectroscopy	Chemical species, interatomic distances
	X-ray photoelectron spectroscopy	Elemental composition, ligand binding
Microscopy	Transmission electron microscopy	Size, dispersity, shape
	Electron diffraction microscopy	Crystal structure
	Scanning electron microscopy	Size and shape
	Atomic force microscopy	3D size and shape
	Fourier transform infrared spectroscopy	Functional groups

Non-X-Ray Spectroscopy/ Spectrometry	Nuclear magnetic resonance	Ligand density/arrangement, atomic composition, size
	Mass spectroscopy	Elemental composition, functional groups
	UV-Vis spectroscopy	Size, some information on shape
	Secondary ion mass spectrometry	Functional groups, surface topography
Light scattering	Dynamic light scattering	Hydrodynamic size, polydispersity, aggregation
	Nanoparticle tracking analysis	Size, size distribution
	Elliptically polarised light scattering	Size, shape
Chromatography	Size exclusion chromatography	Size, size distribution
	Hydrodynamic chromatography	Size, size distribution

Table 6.2: Examples of RTN characterization methods [387].

Finally, delivery of RTNs to the target tissue is a major consideration. Systemic delivery of RTNs generally leads to accumulation in the liver. Therefore, targeted delivery to the lung via aerosolization is a promising avenue. A recent study utilised the lipid SM-102 (the clinically approved ionisable cationic lipid used in COVID-19 vaccine developed by Moderna) to formulate LNPs to transfect airway epithelia with reporter mRNA [388]. LNPs were aerosolised and applied to ALI cultures or *in vivo* mouse airways. Aerosolised LNPs were found to selectively transfect airway epithelial cells over immune and endothelial cells *in vivo*. Interestingly, TEM of LNPs before and after aerosolization revealed that the particles yielding highest transfection efficiency displayed higher rates of aerosolization-induced morphological defects, compared to other formulations that were more structurally robust under aerosolization but offered lower transfection efficiency. Upon aerosolization, LNPs were also found to form respiratory droplets that fell in the optimal size range for

penetration of the entire airway upon inhalation. These results support ionisable cationic lipids as promising vehicles for tissue- and cell-specific delivery *in vivo*.

Overall conclusions

Overall, the results presented in this thesis firmly lay the groundwork for future experimental work. The characterisation of an undruggable CF mutation provides a baseline for measuring the effects of a DNA editing-based treatment strategy. The editing efficiency achieved via HITI, if replicated using a CFTR cDNA, is more than sufficient to provide a rescuing effect in CF variant cells. The RTNs used are highly effective in transfecting submerged cells; their composition could be altered to more effectively transfect differentiated cells, or they could form the basis for an *ex vivo* editing and reimplantation strategy.

1. Geddes, D.M. and R. Shiner, *Cystic fibrosis--from lung damage to symptoms.* Acta Paediatr Scand Suppl, 1989. **363**: p. 52-6; discussion 56-7.

2. Kerem, B., et al., *Identification of the Cystic Fibrosis Gene: Genetic Analysis.* Science, 1989. **245**: p. 1073-1080.

3. Riordan, J.R., et al., *Identification of the Cystic Fibrosis Gene: Cloning and Characterization of Complementary DNA.* Science, 1989. **245**: p. 1966-1073.

4. Vankeerberghen, A.C., H.; Cassiman, J.J., *The cystic fibrosis transmembrane conductance regulator: an intriguing protein with pleiotropic functions.* J Cyst Fibros, 2001. **1**: p. 13-29.

5. Warth, J.D., et al., *CFTR chloride channels in human and simian heart.* Cardiovasc Res, 1996. **31**(4): p. 615-624.

6. Davies, W.L., et al., *Cardiac expression of the cystic fibrosis transmembrane conductance regulator involves novel exon 1 usage to produce a unique amino-terminal protein.* J Biol Chem, 2004. **279**(16): p. 15877-87.

7. Devuyst, O., et al., *Developmental regulation of CFTR expression during human nephrogenesis.* Am J Physiol Ren, 1996. **271**(3): p. F723-F735.

8. Walter, P. and G. Blobel, *Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein.* J Cell Bio, 1981. **91**: p. 551-556.

9. Pind, S., J.R. Riordan, and D.B. Williams, *Participation of the endoplasmic reticulum chaperone calnexin (p88, IP90) in the biogenesis of the cystic fibrosis transmembrane conductance regulator.* Journal of Biological Chemistry, 1994. **269**(17): p. 12784-12788.

10. Gregory, R.J., et al., *Expression and Characterization of the Cystic Fibrosis Transmembrane Conductance Regulator.* Nature, 1990. **347**: p. 382-386.

11. Cheng, S.H., et al., *Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis.* Cell, 1990. **63**(4): p. 827-834.

12. Kalin, N., et al., *DeltaF508 CFTR protein expression in tissues from patients with cystic fibrosis.* J Clin Invest, 1999. **103**(10): p. 1379-89.

13. Becq, F., et al., *The rescue of F508del-CFTR by elexacaftor/tezacaftor/ivacaftor (Trikafta) in human airway epithelial cells is underestimated due to the presence of ivacaftor.* Eur Respir J, 2022. **59**(2): p. 2100671.

14. Swiatecka-Urban, A., et al., *Myosin Vb is required for trafficking of the cystic fibrosis transmembrane conductance regulator in Rab11a-specific apical recycling endosomes in polarized human airway epithelial cells.* J Biol Chem, 2007. **282**(32): p. 23725-36.

15. Wang, X.R. and C. Li, *Decoding F508del misfolding in cystic fibrosis.* Biomolecules, 2014. **4**(2): p. 498-509.

16. McDonald, E.F., et al., *Structural Comparative Modeling of Multi-Domain F508del CFTR.* Biomolecules, 2022. **12**(3): p. 471.

17. Thibodeau, P.H., et al., *The cystic fibrosis-causing mutation deltaF508 affects multiple steps in cystic fibrosis transmembrane conductance regulator biogenesis.* J Biol Chem, 2010. **285**(46): p. 35825-35.

18. Yang, Y., et al., *The common variant of cystic fibrosis transmembrane conductance regulator is recognized by hsp70 and degraded in a pre-Golgi nonlysosomal compartment.* Proc Natl Acad Sci U S A, 1993(20): p. 9480-9484.

19.   Lukacs, G.L., et al., *Conformational maturation of CFTR but not its mutant counterpart (delta F508) occurs in the endoplasmic reticulum and requires ATP.* EMBO J, 1994. **13**(24): p. 6076-86.

20.   Lukasiak, A. and M. Zajac, *The Distribution and Role of the CFTR Protein in the Intracellular Compartments.* Membranes (Basel), 2021. **11**(11): p. 804.

21.   Amaral, M.D., et al., *CFTR processing, trafficking and interactions.* J Cyst Fibros, 2020. **19 Suppl 1**: p. S33-S36.

22.   McShane, A.J., et al., *Targeted proteomic quantitation of the absolute expression and turnover of cystic fibrosis transmembrane conductance regulator in the apical plasma membrane.* J Proteome Res, 2014. **13**(11): p. 4676-85.

23.   Scotet, V., C. L'Hostis, and C. Ferec, *The Changing Epidemiology of Cystic Fibrosis: Incidence, Survival and Impact of the CFTR Gene Discovery.* Genes (Basel), 2020. **11**(6): p. 589.

24.   Turcios, N.L., *Cystic Fibrosis Lung Disease: An Overview.* Respir Care, 2020. **65**(2): p. 233-251.

25.   Niv, Y., S.B. Ho, and T. Rokkas, *Mucin Secretion in Cystic Fibrosis: A Systematic Review.* Dig Dis, 2021. **39**(4): p. 375-381.

26.   Bierlaagh, M.C., et al., *A new era for people with cystic fibrosis.* Eur J Pediatr, 2021. **180**(9): p. 2731-2739.

27.   Jiang, C., et al., *Altered fluid transport across airway epithelium in cystic fibrosis.* Science, 1993. **262**: p. 424-427.

28.   Hanssens, L.S., J. Duchateau, and G.J. Casimir, *CFTR Protein: Not Just a Chloride Channel?* Cells, 2021. **10**(11): p. 2844.

29.   Sanchis, J., et al., *Pulmonary mucociliary clearance in cystic fibrosis.* N Engl J Med, 1973. **288**: p. 651-654.

30.   Singh, P.K., et al., *Synergistic and additive killing by antimicrobial factors found in human airway surface liquid.* Am J Physiol Lung Cell Mol Physiol, 2000. **279**(5): p. 799-805.

31.   Amodio, J.B., et al., *Cystic fibrosis in childhood: pulmonary, paranasal sinus, and skeletal manifestations.* Seminars in Reontgenology, 1987. **22**: p. 125-135.

32.   Rutland, J. and P.J. Cole, *Nasal mucociliary clearance and ciliary beat frequency in cystic fibrosis compared with sinusitis and bronchiectasis.* Thorax, 1981. **36**: p. 654-658.

33.   Pezzulo, A.A., et al., *Reduced airway surface pH impairs bacterial killing in the porcine cystic fibrosis lung.* Nature, 2012. **487**(7405): p. 109-13.

34.   Walker, T.S., et al., *Enhanced Pseudomonas aeruginosa biofilm development mediated by human neutrophils.* Infect Immun, 2005. **73**(6): p. 3693-701.

35.   Prayle, A. and A.R. Smyth, *Aminoglycoside use in cystic fibrosis: therapeutic strategies and toxicity.* Curr Opin Pulm Med, 2010. **16**(6): p. 604-10.

36.   Shah, V.S., et al., *Airway acidification initiates host defense abnormalities in cystic fibrosis mice.* Science, 2016. **351**(6272): p. 503-7.

37.   Venaille, T.J., G. Ryan, and B.W.S. Robinson, *Epithelial cell damage is induced by neutrophil-derived, not pseudomonas-derived, proteases in cystic fibrosis sputum.* Respir Med, 1998. **92**(2): p. 233-240.

38.   Ramos, K.J., et al., *Lung transplant referral for individuals with cystic fibrosis: Cystic Fibrosis Foundation consensus guidelines.* J Cyst Fibros, 2019. **18**(3): p. 321-333.

39.   McBennett, K.A., P.B. Davis, and M.W. Konstan, *Increasing life expectancy in cystic fibrosis: Advances and challenges.* Pediatr Pulmonol, 2022. **57 Suppl 1**(Suppl 1): p. S5-S12.

40.   Hinds, D.M., et al., *Cystic fibrosis in El Salvador.* Pediatr Pulmonol, 2019. **54**(4): p. 369-371.

41.   Lopes-Pacheco, M., *CFTR Modulators: The Changing Face of Cystic Fibrosis in the Era of Precision Medicine.* Front Pharmacol, 2019. **10**: p. 1662.

42.   De Boeck, K., et al., *The relative frequency of CFTR mutation classes in European patients with cystic fibrosis.* J Cyst Fibros, 2014. **13**(4): p. 403-9.

43.   Shoshani, T., et al., *Association of a Nonsense Mutation (WI 282X), the Most Common Mutation in the Ashkenazi Jewish Cystic Fibrosis Patients in Israel, with Presentation of Severe Disease.* Am J Hum Genet, 1992. **50**: p. 222-228.

44.   Shanthikumar, S., et al., *Gene modifiers of cystic fibrosis lung disease: A systematic review.* Pediatr Pulmonol, 2019. **54**(9): p. 1356-1366.

45.   Szczesniak, R., et al., *Influences of environmental exposures on individuals living with cystic fibrosis.* Expert Rev Respir Med, 2020. **14**(7): p. 737-748.

46.   Fiedorczuk, K. and J. Chen, *Molecular structures reveal synergistic rescue of D508 CFTR by Trikafta modulators.* Science, 2022. **378**: p. 284-290.

47.   Fiedorczuk, K. and J. Chen, *Mechanism of CFTR correction by type I folding correctors.* Cell, 2022. **185**(1): p. 158-168 e11.

48.   Liu, F., et al., *Structural identification of a hotspot on CFTR for potentiation.* Science, 2019. **364**(6446): p. 1184-1188.

49.   Shaughnessy, C.A., P.L. Zeitlin, and P.E. Bratcher, *Elexacaftor is a CFTR potentiator and acts synergistically with ivacaftor during acute and chronic treatment.* Sci Rep, 2021. **11**(1): p. 19810.

50.   Hadida, S., et al., *Discovery of N-(2,4-di-tert-butyl-5-hydroxyphenyl)-4-oxo-1,4-dihydroquinoline-3-carboxamide (VX-770, ivacaftor), a potent and orally bioavailable CFTR potentiator.* J Med Chem, 2014. **57**(23): p. 9776-95.

51.   Guerra, L., et al., *The preclinical discovery and development of the combination of ivacaftor + tezacaftor used to treat cystic fibrosis.* Expert Opin Drug Discov, 2020. **15**(8): p. 873-891.

52.   Bear, C.E., *A Therapy for Most with Cystic Fibrosis.* Cell, 2020. **180**(2): p. 211.

53.   Middleton, P.G., et al., *Elexacaftor-Tezacaftor-Ivacaftor for Cystic Fibrosis with a Single Phe508del Allele.* N Engl J Med, 2019. **381**(19): p. 1809-1819.

54.   Zaher, A., et al., *A Review of Trikafta: Triple Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Modulator Therapy.* Cureus, 2021. **13**(7): p. e16144.

55.   Gruenert, D.C., et al., *Characterization of human tracheal epithelial cells transformed by an origin-defective simian virus 40.* Prot Natl Acad Sci USA, 1988. **85**(16): p. 5951-5955.

56.   Cozens, A.L., et al., *CFTR expression and chloride secretion in polarized immortal human bronchial epithelial cells.* Am J Respir Cell Mol Biol, 1993. **10**: p. 38-47.

57.   Valley, H.C., et al., *Isogenic cell models of cystic fibrosis-causing variants in natively expressing pulmonary epithelial cells.* J Cyst Fibros, 2019. **18**(4): p. 476-483.

58.   Sanderlin, E.J., et al., *CFTR mRNAs with nonsense codons are degraded by the SMG6-mediated endonucleolytic decay pathway.* Nat Commun, 2022. **13**(1): p. 2344.

59. Schogler, A., et al., *Characterization of pediatric cystic fibrosis airway epithelial cell cultures at the air-liquid interface obtained by non-invasive nasal cytology brush sampling.* Respir Res, 2017. **18**(1): p. 215.

60. McNamara, P.S., et al., *Comparison of techniques for obtaining lower airway epithelial cells from children.* Eur Respir J, 2008. **32**(3): p. 763-8.

61. Looi, K., et al., *Bronchial brushings for investigating airway inflammation and remodelling.* Respirology, 2011. **16**(5): p. 725-37.

62. Kicic, A., et al., *Intrinsic biochemical and functional differences in bronchial epithelial cells of children with asthma.* Am J Respir Crit Care Med, 2006. **174**(10): p. 1110-8.

63. Chapman, S., et al., *Human keratinocytes are efficiently immortalized by a Rho kinase inhibitor.* J Clin Invest, 2010. **120**(7): p. 2619-26.

64. Liu, X., et al., *ROCK inhibitor and feeder cells induce the conditional reprogramming of epithelial cells.* Am J Pathol, 2012. **180**(2): p. 599-607.

65. Suzuki, D. and M. Senoo, *Dact1 Regulates the Ability of 3T3-J2 Cells to Support Proliferation of Human Epidermal Keratinocytes.* J Invest Dermatol, 2015. **135**(11): p. 2894-2897.

66. Munye, M.M., et al., *BMI-1 extends proliferative potential of human bronchial epithelial cells while retaining their mucociliary differentiation capacity.* Am J Physiol Lung Cell Mol Physiol, 2017. **312**(2): p. L258-L267.

67. O'Loughlin, J., et al., *Extended lifespan of bronchial epithelial cells maintains normal cellular phenotype and transcriptome integrity.* ERJ Open Res, 2021. **7**(1): p. 00254.

68. Tuck, S.A., et al., *Time course of airway remodelling after an acute chlorine gas exposure in mice.* Respir Res, 2008. **9**: p. 61.

69. Baldassi, D., B. Gabold, and O. Merkel, *Air-liquid interface cultures of the healthy and diseased human respiratory tract: promises, challenges and future directions.* Adv Nanobiomed Res, 2021. **1**(6): p. 2000111.

70. Plasschaert, L.W., et al., *A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte.* Nature, 2018. **560**(7718): p. 377-381.

71. Davis, J.D. and T.P. Wypych, *Cellular and functional heterogeneity of the airway epithelium.* Mucosal Immunol, 2021. **14**(5): p. 978-990.

72. Ghosh, B., et al., *Strong correlation between air-liquid interface cultures and in vivo transcriptomics of nasal brush biopsy.* Am J Physiol Lung Cell Mol Physiol, 2020. **318**(5): p. 1056-1062.

73. Bukowy-Bieryllo, Z., et al., *In vitro differentiation of ciliated cells in ALI-cultured human airway epithelium - The framework for functional studies on airway differentiation in ciliopathies.* Eur J Cell Biol, 2022. **101**(1): p. 151189.

74. Leung, C., et al., *Structural and functional variations in human bronchial epithelial cells cultured in air-liquid interface using different growth media.* Am J Physiol Lung Cell Mol Physiol, 2020. **318**(5): p. L1063-L1073.

75. Brewington, J.J., et al., *Brushed nasal epithelial cells are a surrogate for bronchial epithelial CFTR studies.* JCI Insight, 2018. **3**(13): p. e99385.

76. Sato, T., et al., *Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche.* Nature, 2009. **459**(7244): p. 262-5.

77. Liu, J., et al., *Functional Cftr in crypt epithelium of organotypic enteroid cultures from murine small intestine.* Am J Physiol Cell Physiol, 2012. **302**(10): p. C1492-503.

78. Dekkers, J.F., et al., *A functional CFTR assay using primary cystic fibrosis intestinal organoids.* Nat Med, 2013. **19**(7): p. 939-45.

79.   Dekkers, J.F., et al., *Characterizing responses to CFTR-modulating drugsusing rectal organoids derived from subjectswith cystic fibrosis.* Sci Transl Med, 2016. **8**(344).

80.   Geurts, M.H., et al., *CRISPR-Based Adenine Editors Correct Nonsense Mutations in a Cystic Fibrosis Organoid Biobank.* Cell Stem Cell, 2020: p. 503-510.

81.   Geurts, M.H., et al., *Evaluating CRISPR-based prime editing for cancer modeling and CFTR repair in organoids.* Life Sci Alliance, 2021. **4**(10): p. e202000940.

82.   Fleischer, A., et al., *iPSC-Derived Intestinal Organoids from Cystic Fibrosis Patients Acquire CFTR Activity upon TALEN-Mediated Repair of the p.F508del Mutation.* Mol Ther Methods Clin Dev, 2020. **17**: p. 858-870.

83.   Amistadi, S., et al., *Functional restoration of a CFTR splicing mutation through RNA delivery of CRISPR adenine base editor.* Mol Ther, 2023: p. 1647-1660.

84.   Mall, M.A., *ENaC inhibition in cystic fibrosis: potential role in the new era of CFTR modulator therapies.* Eur Respir J, 2020. **56**(6): p. 2000946.

85.   Zhou, Z., et al., *The ENaC-overexpressing mouse as a model of cystic fibrosis lung disease.* J Cyst Fibros, 2011. **10 Suppl 2**: p. S172-82.

86.   McCarron, A., M. Donnelley, and D. Parsons, *Airway disease phenotypes in animal models of cystic fibrosis.* Respir Res, 2018. **19**(1): p. 54.

87.   McCarron, A., D. Parsons, and M. Donnelley, *Animal and Cell Culture Models for Cystic Fibrosis: Which Model Is Right for Your Application?* Am J Pathol, 2021. **191**(2): p. 228-242.

88.   Stoltz, D.A., et al., *Intestinal CFTR expression alleviates meconium ileus in cystic fibrosis pigs.* J Clin Invest, 2013. **123**(6): p. 2685-93.

89.   Xu, J., et al., *Phenotypes of CF rabbits generated by CRISPR/Cas9-mediated disruption of the CFTR gene.* JCI Insight, 2021. **6**(1): p. e139813

90.   Fan, Z., et al., *A sheep model of cystic fibrosis generated by CRISPR/Cas9 disruption of the CFTR gene.* JCI Insight, 2018. **3**(19): p. e123529.

91.   Mall, M., et al., *The Amiloride-inhibitable Na1 Conductance Is Reduced by the Cystic Fibrosis Transmembrane Conductance Regulator in Normal But Not in Cystic Fibrosis Airways.* J. Clin. Invest., 1998. **102**: p. 15-21.

92.   Caldwell, R.A., R.C. Boucher, and M.J. Stutts, *Neutrophil elastase activates near-silent epithelial Na+ channels and increases airway epithelial Na+ transport.* Am J Physiol Lung Cell Mol Physiol, 2005. **288**(5): p. L813-9.

93.   Butterworth, M.B., et al., *Activation of the epithelial sodium channel (ENaC) by the alkaline protease from Pseudomonas aeruginosa.* J Biol Chem, 2012. **287**(39): p. 32556-65.

94.   Mall, M., et al., *Increased airway epithelial Na+ absorption produces cystic fibrosis-like lung disease in mice.* Nat Med, 2004. **10**(5): p. 487-93.

95.   Agrawal, P.B., et al., *The Epithelial Sodium Channel Is a Modifier of the Long-Term Nonprogressive Phenotype Associated with F508del CFTR Mutations.* Am J Respir Cell Mol Biol, 2017. **57**(6): p. 711-720.

96.   Tagalakis, A.D., et al., *Effective silencing of ENaC by siRNA delivered with epithelial-targeted nanocomplexes in human cystic fibrosis cells and in mouse lung.* Thorax, 2018. **73**(9): p. 847-856.

97.   Nickolaus, P., et al., *Preclinical evaluation of the epithelial sodium channel inhibitor BI 1265162 for treatment of cystic fibrosis.* ERJ Open Res, 2020. **6**(4): p. 00429

98.   Kelly, T.J. and H.O. Smith, *A restriction enzyme from Hemophilus influenzae.* J. Mol. Biol., 1970. **51**: p. 393-409.

99.   Roberts, R.J., et al., *REBASE--a database for DNA restriction and modification: enzymes, genes and genomes.* Nucleic Acids Res, 2010. **38**(Database issue): p. D234-6.

100.   Struhl, K., *Subcloning of DNA fragments.* Curr Protoc Mol Biol, 2001. **Chapter 3**: p. Unit3 16.

101.   *NEBCloner.* 2016 2023; Available from: https://nebcloner.neb.com/.

102.   Thierry, A. and B. Dujon, *Nested chromosomal fragmentation in yeast using the meganuclease l-Sce 1: a new method for physical mapping of eukaryotic genomes.* Nucleic Acids Res, 1992. **20**(21): p. 5625-5631.

103.   Seligman, L.M., et al., *Nested chromosomal fragmentation in yeast using themeganuclease l-Sce 1: a new method for physical mapping of eukaryotic genomes.* Nucleic Acids Res, 2002. **30**(17): p. 3870—3879.

104.   Sussman, D., et al., *Isolation and characterization of new homing endonuclease specificities at individual target site positions.* J Mol Biol, 2004. **342**(1): p. 31-41.

105.   Smith, J., et al., *A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences.* Nucleic Acids Res, 2006. **34**(22): p. e149.

106.   MacLeod, D.T., et al., *Integration of a CD19 CAR into the TCR Alpha Chain Locus Streamlines Production of Allogeneic Gene-Edited CAR T Cells.* Mol Ther, 2017. **25**(4): p. 949-961.

107.   Gorsuch, C.L., et al., *Targeting the hepatitis B cccDNA with a sequence-specific ARCUS nuclease to eliminate hepatitis B virus in vivo.* Mol Ther, 2022. **30**(9): p. 2909-2922.

108.   Doyon, J.B., et al., *Directed Evolution and Substrate Specificity Profile of Homing Endonuclease I-SceI.* J Am Chem Soc, 2005. **128**(2477-2484).

109.   Lee, K.Z., et al., *Repurposing the Homing Endonuclease I-SceI for Positive Selection and Development of Gene-Editing Technologies.* ACS Synth Biol, 2022. **11**(1): p. 53-60.

110.   Podhajska, A.J. and W. Szybalski, *Conversion of the FokI endonuclease to a universal restriction enzyme: cleavage phage M13mp7 DNA at predetermined sites.* Gene, 1985. **40**(2-3): p. 175-182.

111.   Kim, H.J., et al., *Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly.* Genome Res, 2009. **19**(7): p. 1279-88.

112.   Smith, J., et al., *Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains.* Nucleic Acids Res, 2000. **28**(17): p. 3361-3369.

113.   Moscou, M.J. and A.J. Bogdanove, *A simple cipher governs DNA recognition by TAL effectors.* Science, 2009. **326**(5959): p. 1501.

114.   Boch, J., et al., *Breaking the code of DNA binding specificty of TAL-type III effectors.* Science, 2009. **326**: p. 1509-1512.

115.   Miller, J.C., et al., *A TALE nuclease architecture for efficient genome editing.* Nat Biotechnol, 2011. **29**(2): p. 143-8.

116.   Cermak, T., et al., *Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting.* Nucleic Acids Res, 2011. **39**(12): p. e82.

117.   Mojica, F.J.M. and L. Montoliu, *On the Origin of CRISPR-Cas Technology: From Prokaryotes to Mammals.* Trends Microbiol, 2016. **24**(10): p. 811-820.

118.   Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012: p. 816-821.

119. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**: p. 819-823.

120. Mali, P., et al., *RNA-guided human genome engineering via Cas9.* Science, 2013. **339**(6121): p. 823-6.

121. Bolotin, A., et al., *Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin.* Microbiology (Reading), 2005. **151**(Pt 8): p. 2551-2561.

122. Canela, A., et al., *DNA Breaks and End Resection Measured Genome-wide by End Sequencing.* Mol Cell, 2016. **63**(5): p. 898-911.

123. Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**(7340): p. 602-7.

124. Garneau, J.E., et al., *The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA.* Nature, 2010. **468**(7320): p. 67-71.

125. Chen, H., J. Choi, and S. Bailey, *Cut site selection by the two nuclease domains of the Cas9 RNA-guided endonuclease.* J Biol Chem, 2014. **289**(19): p. 13284-94.

126. Xu, Y. and Z. Li, *CRISPR-Cas systems: Overview, innovations and applications in human disease research and gene therapy.* Comput Struct Biotechnol J, 2020. **18**: p. 2401-2415.

127. Ren, B., et al., *Cas9-NG Greatly Expands the Targeting Scope of the Genome-Editing Toolkit by Recognizing NG and Other Atypical PAMs in Rice.* Mol Plant, 2019. **12**(7): p. 1015-1026.

128. Xu, X., et al., *Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing.* Mol Cell, 2021. **81**(20): p. 4333-4345 e4.

129. Ewaisha, R. and K.S. Anderson, *Immunogenicity of CRISPR therapeutics-Critical considerations for clinical translation.* Front Bioeng Biotechnol, 2023. **11**: p. 1138596.

130. Ferdosi, S.R., et al., *Multifunctional CRISPR-Cas9 with engineered immunosilenced human T cell epitopes.* Nat Commun, 2019. **10**(1): p. 1842.

131. Liang, F., et al., *Homology-directed repair is a major double-strand break repair pathway in mammalian cells.* Proceedings of the National Academy of Science, 1998. **95**: p. 5172-5177.

132. Burma, S., B.P. Chen, and D.J. Chen, *Role of non-homologous end joining (NHEJ) in maintaining genomic integrity.* DNA Repair (Amst), 2006. **5**(9-10): p. 1042-8.

133. Chang, H.H.Y., et al., *Non-homologous DNA end joining and alternative pathways to double-strand break repair.* Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.

134. Liu, M., et al., *Methodologies for Improving HDR Efficiency.* Front Genet, 2018. **9**: p. 691.

135. Sfeir, A. and L.S. Symington, *Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway?* Trends Biochem Sci, 2015. **40**(11): p. 701-714.

136. Tatiossian, K.J., et al., *Rational Selection of CRISPR-Cas9 Guide RNAs for Homology-Directed Genome Editing.* Mol Ther, 2021. **29**(3): p. 1057-1069.

137. Song, B., et al., *Analysis of NHEJ-Based DNA Repair after CRISPR-Mediated DNA Cleavage.* Int J Mol Sci, 2021. **22**(12): p. 6397.

138. Chen, G., et al., *Bmi1 Overexpression in Mesenchymal Stem Cells Exerts Antiaging and Antiosteoporosis Effects by Inactivating p16/p19 Signaling and Inhibiting Oxidative Stress.* Stem Cells, 2019. **37**(9): p. 1200-1211.

139. Riesenberg, S. and T. Maricic, *Targeting repair pathways with small molecules increases precise genome editing in pluripotent stem cells.* Nat Commun, 2018. **9**(1): p. 2164.

140. Tang, X.D., et al., *Methods for Enhancing Clustered Regularly Interspaced Short Palindromic Repeats/Cas9-Mediated Homology-Directed Repair Efficiency.* Front Genet, 2019. **10**: p. 551.

141. Liang, X., et al., *Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection.* J Biotechnol, 2015. **208**: p. 44-53.

142. Okamoto, S., et al., *Highly efficient genome editing for single-base substitutions using optimized ssODNs with Cas9-RNPs.* Sci Rep, 2019. **9**(1): p. 4811.

143. He, X., et al., *Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair.* Nucleic Acids Res, 2016. **44**(9): p. e85.

144. Zheng, Q., et al., *Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells.* Biotechniques, 2014. **57**(3): p. 115-24.

145. Guo, T., et al., *Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing.* Genome Biol, 2018. **19**(1): p. 170.

146. Lalonde, S., et al., *Frameshift indels introduced by genome editing can lead to in-frame exon skipping.* PLoS One, 2017. **12**(6): p. e0178700.

147. Kapahnke, M., A. Banning, and R. Tikkanen, *Random Splicing of Several Exons Caused by a Single Base Change in the Target Exon of CRISPR/Cas9 Mediated Gene Knockout.* Cells, 2016. **5**(4): p. 45.

148. Lo Scrudato, M., et al., *Genome Editing of Expanded CTG Repeats within the Human DMPK Gene Reduces Nuclear RNA Foci in the Muscle of DM1 Mice.* Mol Ther, 2019. **27**(8): p. 1372-1388.

149. Pascual-Gilabert, M., R. Artero, and A. Lopez-Castel, *The myotonic dystrophy type 1 drug development pipeline: 2022 edition.* Drug Discov Today, 2023. **28**(3): p. 103489.

150. Lackner, D.H., et al., *A generic strategy for CRISPR-Cas9-mediated gene tagging.* Nat Commun, 2015. **6**: p. 10237.

151. Suzuki, K. and J.C. Izpisua Belmonte, *In vivo genome editing via the HITI method as a tool for gene therapy.* J Hum Genet, 2018. **63**(2): p. 157-164.

152. Suzuki, K., et al., *In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration.* Nature, 2016. **540**(7631): p. 144-149.

153. Sendino Garví, E., et al., *Gene surgery as a potential treatment option for Nephropathic Cystinosis in vitro.* bioRxiv, 2023.

154. Stephenson, A.A., et al., *CRISPR-Cas9 homology-independent targeted integration of exons 1-19 restores full-length dystrophin in mice.* Mol Ther Methods Clin Dev, 2023. **30**: p. 486-499.

155. Hoang, D.A., et al., *Mutation-independent gene knock-in therapy targeting 5'UTR for autosomal dominant retinitis pigmentosa.* Signal Transduct Target Ther, 2023. **8**(1): p. 100.

156. Guilinger, J.P., D.B. Thompson, and D.R. Liu, *Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification.* Nat Biotechnol, 2014. **32**(6): p. 577-582.

157. Doench, J.G., et al., *Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation.* Nat Biotechnol, 2014. **32**(12): p. 1262-7.

158. 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.

159. Labuhn, M., et al., *Refined sgRNA efficacy prediction improves large- and small-scale CRISPR-Cas9 applications.* Nucleic Acids Res, 2018. **46**(3): p. 1375-1385.

160. Wang, T., et al., *Genetic Screens in Human Cells using the CRISPR-Cas9 System.* Science, 2014. **343**(6166): p. 80-84.

161. Moreno-Mateos, M.A., et al., *CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo.* Nat Methods, 2015. **12**(10): p. 982-8.

162. Tsai, S.Q., et al., *GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases.* Nat Biotechnol, 2015. **33**(2): p. 187-197.

163. Wong, N., W. Liu, and X. Wang, *WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system.* Genome Biol, 2015. **16**: p. 218.

164. Jensen, K.T., et al., *Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency.* FEBS Lett, 2017. **591**(13): p. 1892-1901.

165. Thyme, S.B., et al., *Internal guide RNA interactions interfere with Cas9-mediated cleavage.* Nat Commun, 2016. **7**: p. 11750.

166. Janssen, J.M., et al., *The Chromatin Structure of CRISPR-Cas9 Target DNA Controls the Balance between Mutagenic and Homology-Directed Gene-Editing Events.* Mol Ther Nucleic Acids, 2019. **16**: p. 141-154.

167. Mashal, R., J. Koontz, and J. Sklar, *Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases.* Nat Gen, 1995. **9**: p. 177-183.

168. Hsiau, T., et al., *Inference of CRISPR Edits from Sanger Trace Data.* CRISPR J, 2022. **5**(1): p. 123-130.

169. Brinkman, E.K., et al., *Easy quantitative assessment of genome editing by sequence trace decomposition.* Nucleic Acids Res, 2014. **42**(22): p. e168.

170. Sentmanat, M.F., et al., *A Survey of Validation Strategies for CRISPR-Cas9 Editing.* Sci Rep, 2018. **8**(1): p. 888.

171. Hindson, B.J., et al., *High-throughput droplet digital PCR system for absolute quantitation of DNA copy number.* Anal Chem, 2011. **83**(22): p. 8604-10.

172. Findlay, S.D., et al., *A Digital PCR-Based Method for Efficient and Highly Specific Screening of Genome Edited Cells.* PLoS One, 2016. **11**(4): p. e0153901.

173. Krenning, L., J. van den Berg, and R.H. Medema, *Life or Death after a Break: What Determines the Choice?* Mol Cell, 2019. **76**(2): p. 346-358.

174. Naeem, M., et al., *Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing.* Cells, 2020. **9**(7): p. 1608.

175. Lin, Y., et al., *CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences.* Nucleic Acids Res, 2014. **42**(11): p. 7473-85.

176. Hsu, P.D., et al., *DNA targeting specificity of RNA-guided Cas9 nucleases.* Nat Biotechnol, 2013. **31**(9): p. 827-32.

177. Fu, Y., et al., *High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells.* Nat Biotechnol, 2013. **31**(9): p. 822-6.

178. Cho, S.W., et al., *Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases.* Genome Res, 2014. **24**(1): p. 132-41.

179. Pattanayak, V., et al., *High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity.* Nat Biotechnol, 2013. **31**(9): p. 839-43.

180. Anderson, E.M., et al., *Systematic analysis of CRISPR-Cas9 mismatch tolerance reveals low levels of off-target activity.* J Biotechnol, 2015. **211**: p. 56-65.

181. Cromer, M.K., et al., *Comparative analysis of CRISPR off-target discovery tools following ex vivo editing of CD34(+) hematopoietic stem and progenitor cells.* Mol Ther, 2023. **31**(4): p. 1074-1087.

182. Tsai, S.Q., et al., *CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets.* Nat Methods, 2017. **14**(6): p. 607-614.

183. Lazzarotto, C.R., et al., *CHANGE-seq reveals genetic and epigenetic effects on CRISPR-Cas9 genome-wide activity.* Nat Biotechnol, 2020. **38**(11): p. 1317-1327.

184. Cameron, P., et al., *Mapping the genomic landscape of CRISPR-Cas9 cleavage.* Nat Methods, 2017. **14**(6): p. 600-606.

185. Zou, R.S., et al., *Improving the sensitivity of in vivo CRISPR off-target detection with DISCOVER-Seq.* Nat Methods, 2023. **20**(5): p. 706-713.

186. Chiarle, R., et al., *Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells.* Cell, 2011. **147**(1): p. 107-19.

187. Aubrey, B.J., et al., *How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression?* Cell Death Differ, 2018. **25**(1): p. 104-113.

188. Ihry, R.J., et al., *p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells.* Nat Med, 2018. **24**(7): p. 939-946.

189. Haapaniemi, E., et al., *CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response.* Nat Med, 2018. **24**(7): p. 927-930.

190. Hyodo, T., et al., *Tandem Paired Nicking Promotes Precise Genome Editing with Scarce Interference by p53.* Cell Rep, 2020. **30**(4): p. 1195-1207 e7.

191. Chen, X., et al., *In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting.* Nat Commun, 2017. **8**(1): p. 657.

192. Komor, A.C., et al., *Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage.* Nature, 2016. **533**(7603): p. 420-4.

193. Gaudelli, N.M., et al., *Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage.* Nature, 2017. **551**(7681): p. 464-471.

194. Antoniou, P., A. Miccio, and M. Brusson, *Base and Prime Editing Technologies for Blood Disorders.* Front Genome Ed, 2021. **3**: p. 618406.

195. Komor, A.C., et al., *Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity.* Sci Adv, 2017: p. eaao4774.

196. Zafra, M.P., et al., *Optimized base editors enable efficient editing in cells, organoids and mice.* Nat Biotechnol, 2018. **36**(9): p. 888-893.

197. Yuan, T., et al., *Optimization of C-to-G base editors with sequence context preference predictable by machine learning methods.* Nat Commun, 2021. **12**(1): p. 4902.

198. Talas, A., et al., *BEAR reveals that increased fidelity variants can successfully reduce the mismatch tolerance of adenine but not cytosine base editors.* Nat Commun, 2021. **12**(1): p. 6353.

199. Grunewald, J., et al., *Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors.* Nature, 2019. **569**(7756): p. 433-437.

200. Zuo, E., et al., *Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos.* Science, 2019. **364**(6437): p. 289-292.

201. Rees, H.A., et al., *Analysis and minimization of cellular RNA editing by DNA adenine base editors.* Sci Adv, 2019. **5**.

202. Zhou, C., et al., *Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis.* Nature, 2019. **571**(7764): p. 275-278.

203. Rees, H.A. and D.R. Liu, *Base editing: precision chemistry on the genome and transcriptome of living cells.* Nat Rev Genet, 2018. **19**(12): p. 770-788.

204. Chen, L., et al., *Engineering a precise adenine base editor with minimal bystander editing.* Nat Chem Biol, 2023. **19**(1): p. 101-110.

205. Anzalone, A.V., et al., *Search-and-replace genome editing without double-strand breaks or donor DNA.* Nature, 2019. **576**(7785): p. 149-157.

206. Nelson, J.W., et al., *Engineered pegRNAs improve prime editing efficiency.* Nat Biotechnol, 2022. **40**(3): p. 402-410.

207. Bikard, D., et al., *Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system.* Nucleic Acids Res, 2013. **41**(15): p. 7429-37.

208. Gilbert, L.A., et al., *CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes.* Cell, 2013. **154**(2): p. 442-51.

209. Alerasool, N., et al., *An efficient KRAB domain for CRISPRi applications in human cells.* Nat Methods, 2020. **17**(11): p. 1093-1096.

210. Ghavami, S. and A. Pandi, *CRISPR interference and its applications.* Prog Mol Biol Transl Sci, 2021. **180**: p. 123-140.

211. Sun, M., et al., *Systematic functional interrogation of human pseudogenes using CRISPRi.* Genome Biol, 2021. **22**(1): p. 240.

212. Liu, S.J., et al., *CRISPRi-based radiation modifier screen identifies long non-coding RNA therapeutic targets in glioma.* Genome Biol, 2020. **21**(1): p. 83.

213. Peddle, C.F., et al., *CRISPR Interference-Potential Application in Retinal Disease.* Int J Mol Sci, 2020. **21**(7): p. 2329.

214. Maeder, M.L., et al., *CRISPR RNA-guided activation of endogenous human genes.* Nat Methods, 2013. **10**(10): p. 977-9.

215. Chavez, A., et al., *Highly efficient Cas9-mediated transcriptional programming.* Nat Methods, 2015. **12**(4): p. 326-8.

216. Konermann, S., et al., *Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex.* Nature, 2015. **517**(7536): p. 583-8.

217. Becirovic, E., *Maybe you can turn me on: CRISPRa-based strategies for therapeutic applications.* Cell Mol Life Sci, 2022. **79**(2): p. 130.

218. Cox, D.B.T., et al., *RNA editing with CRISPR-Cas13.* Science, 2017. **358**: p. 1019-1027.

219. *CRISPR Clinical Trials*. 2023; Available from: https://crisprmedicinenews.com/clinical-trials/.

220. Firth, A.L., et al., *Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs.* Cell Rep, 2015. **12**(9): p. 1385-90.

221. Ruan, J., et al., *Efficient Gene Editing at Major CFTR Mutation Loci.* Mol Ther Nucleic Acids, 2019. **16**: p. 73-81.

222. Martin, R.M., et al., *Highly Efficient and Marker-free Genome Editing of Human Pluripotent Stem Cells by CRISPR-Cas9 RNP and AAV6 Donor-Mediated Homologous Recombination.* Cell Stem Cell, 2019. **24**(5): p. 821-828 e5.

223. Vaidyanathan, S., et al., *Highly efficient repair of the ΔF508 mutation in airway stem cells of cystic fibrosis patients with functional rescue of the differentiated epithelia.* BioRxiv, 2019.

224. Cutting, G.R., *Cystic fibrosis genetics: from molecular understanding to clinical application.* Nat Rev Genet, 2015. **16**(1): p. 45-56.

225. Yan, Z., et al., *Recombinant Adeno-Associated Virus-Mediated Editing of the G551D Cystic Fibrosis Transmembrane Conductance Regulator Mutation in Ferret Airway Basal Cells.* Hum Gene Ther, 2022. **33**(19-20): p. 1023-1036.

226. Schwank, G., et al., *Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients.* Cell Stem Cell, 2013. **13**(6): p. 653-8.

227. Geurts, M.H., et al., *CRISPR-Based Adenine Editors Correct Nonsense Mutations in a Cystic Fibrosis Organoid Biobank.* Cell Stem Cell, 2020.

228. Jiang, T., et al., *Chemical modifications of adenine base editor mRNA and guide RNA expand its application scope.* Nat Commun, 2020. **11**(1): p. 1979.

229. Krishnamurthy, S., et al., *Functional correction of CFTR mutations in human airway epithelial cells using adenine base editors.* Nucleic Acids Res, 2021. **49**(18): p. 10558-10572.

230. Sanz, D.J., et al., *Cas9/gRNA targeted excision of cystic fibrosis-causing deep-intronic splicing mutations restores normal splicing of CFTR mRNA.* PLoS One, 2017. **12**(9): p. e0184009.

231. Walker, A.J., et al., *Molecular and Functional Correction of a Deep Intronic Splicing Mutation in CFTR by CRISPR/Cas9 Gene Editing.* Molecular Therapy - Methods & Clinical Development, 2023.

232. *The Clinical and Functional TRanslation of CFTR (CFTR2).* 2011, last update 2023, Johns Hopkins University and the Hospital for Sick Children.

233. Zhou, Z.P., et al., *In Vitro Validation of a CRISPR-Mediated CFTR Correction Strategy for Preclinical Translation in Pigs.* Hum Gene Ther, 2019. **30**(9): p. 1101-1116.

234. Cao, H., et al., *Transducing Airway Basal Cells with a Helper-Dependent Adenoviral Vector for Lung Gene Therapy.* Hum Gene Ther, 2018. **29**(6): p. 643-652.

235. Vaidyanathan, S., et al., *Targeted replacement of full-length CFTR in human airway stem cells by CRISPR-Cas9 for pan-mutation correction in the endogenous locus.* Mol Ther, 2022. **30**(1): p. 223-237.

236. Maule, G., et al., *Allele specific repair of splicing mutations in cystic fibrosis through AsCas12a genome editing.* Nat Commun, 2019. **10**(1): p. 3556.

237. Chen, Y., et al., *A Self-restricted CRISPR System to Reduce Off-target Effects.* Mol Ther, 2016. **24**(9): p. 1508-10.

238. Santos, L., et al., *Comparison of Cas9 and Cas12a CRISPR editing methods to correct the W1282X-CFTR mutation.* J Cyst Fibros, 2022. **21**(1): p. 181-187.

239. Bednarski, C., et al., *Targeted Integration of a Super-Exon into the CFTR Locus Leads to Functional Correction of a Cystic Fibrosis Cell Line Model.* PLoS One, 2016. **11**(8): p. e0161072.

240. Suzuki, S., et al., *Highly Efficient Gene Editing of Cystic Fibrosis Patient-Derived Airway Basal Cells Results in Functional CFTR Correction.* Mol Ther, 2020. **28**(7): p. 1684-1695.

241. Brandsma, I. and D.C. Gent, *Pathway choice in DNA double strand break repair: observations of a balancing act.* Genome Integr, 2012. **3**(1): p. 9.

242. Mao, Z., et al., *Comparison of nonhomologous end joining and homologous recombination in human cells.* DNA Repair (Amst), 2008. **7**(10): p. 1765-71.

243. Bowen, D.H., *Cell turnover in the lung.* Am Rev of Resp Dis, 1983. **128**(2): p. 2.

244. Verdera, H.C., K. Kuranda, and F. Mingozzi, *AAV Vector Immunogenicity in Humans: A Long Journey to Successful Gene Transfer.* Mol Ther, 2020. **28**(3): p. 723-746.

245. Ghosh, B., et al., *Strong correlation between air-liquid interface cultures and in vivo transcriptomics of nasal brush biopsy.* Am J Physiol Lung Cell Mol Physiol, 2020.

246. Erjefalt, J.S., et al., *In vivo restitution of airway epithelium.* Cell Tissue Res., 1995. **281**(2): p. 305-16.

247. Machado, M.N., et al., *Papain-induced experimental pulmonary emphysema in male and female mice.* Respir Physiol Neurobiol, 2014. **200**: p. 90-6.

248. Rosen, C., et al., *Preconditioning allows engraftment of mouse and human embryonic lung cells, enabling lung repair in mice.* Nat Med, 2015. **21**(8): p. 869-79.

249. Griggs, T.F., et al., *Rhinovirus C targets ciliated airway epithelial cells.* Respir Res, 2017. **18**(1): p. 84.

250. Cmielewski, P., et al., *Gene therapy for Cystic Fibrosis: Improved delivery techniques and conditioning with lysophosphatidylcholine enhance lentiviral gene transfer in mouse lung airways.* Exp Lung Res, 2017. **43**(9-10): p. 426-433.

251. Gui, L., et al., *Efficient intratracheal delivery of airway epithelial cells in mice and pigs.* Am J Physiol Lung Cell Mol Physiol, 2015. **308**(2): p. L221-8.

252. Leblond, A., et al., *Developing Cell Therapy Techniques for Respiratory Disease: Intratracheal Delivery of Genetically Engineered Stem Cells in a Murine Model of Airway Injury.* Hum Gene Ther, 2009. **20**(11): p. 1329-43.

253. Farrow, N., et al., *Epithelial disruption: a new paradigm enabling human airway stem cell transplantation.* Stem Cell Res Ther, 2018. **9**(1): p. 153.

254. Ghosh, M., et al., *Transplantation of Airway Epithelial Stem/Progenitor Cells: A Future for Cell-Based Therapy.* Am J Respir Cell Mol Biol, 2017. **56**(1): p. 1-10.

255. Vaidyanathan, S., et al., *High-Efficiency, Selection-free Gene Repair in Airway Stem Cells from Cystic Fibrosis Patients Rescues CFTR Function in Differentiated Epithelia.* Cell Stem Cell, 2020. **26**(2): p. 161-171 e4.

256. Kim, S., et al., *Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins.* Genome Res, 2014. **24**(6): p. 1012-9.

257. Mignon, C., R. Sodoyer, and B. Werle, *Antibiotic-free selection in biotherapeutics: now and forever.* Pathogens, 2015. **4**(2): p. 157-81.

258. Liu, J., et al., *Fast and Efficient CRISPR/Cas9 Genome Editing In Vivo Enabled by Bioreducible Lipid and Messenger RNA Nanoparticles.* Adv Mater, 2019. **31**(33): p. e1902575.

259. Ryan, D.E., et al., *Improving CRISPR-Cas specificity with chemical modifications in single-guide RNAs.* Nucleic Acids Res, 2018. **46**(2): p. 792-803.

260. Cromwell, C.R., et al., *Incorporation of bridged nucleic acids into CRISPR RNAs improves Cas9 endonuclease specificity.* Nat Commun, 2018. **9**(1): p. 1448.

261. Yin, H., et al., *Partial DNA-guided Cas9 enables genome editing with reduced off-target activity.* Nat Chem Biol, 2018. **14**(3): p. 311-316.

262. Basila, M., M.L. Kelley, and A.V.B. Smith, *Minimal 2'-O-methyl phosphorothioate linkage modification pattern of synthetic guide RNAs for increased stability and efficient CRISPR-Cas9 gene editing avoiding cellular toxicity.* PLoS One, 2017. **12**(11): p. e0188593.

263. Jaroszewicz, W., et al., *Phage display and other peptide display technologies.* FEMS Microbiol Rev, 2022. **46**(2): p. fuab052.

264. Writer, M.J., et al., *Targeted gene delivery to human airway epithelial cells with synthetic vectors incorporating novel targeting peptides selected by phage display.* J Drug Target, 2004. **12**(4): p. 185-93.

265. Manunta, M.D.I., et al., *Delivery of ENaC siRNA to epithelial cells mediated by a targeted nanocomplex: a therapeutic strategy for cystic fibrosis.* Sci Rep, 2017. **7**(1): p. 700.

266. Wang, M., et al., *Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles.* Proc Natl Acad Sci U S A, 2016. **113**(11): p. 2868-73.

267. Sousa de Almeida, M., et al., *Understanding nanoparticle endocytosis to improve targeting strategies in nanomedicine.* Chem Soc Rev, 2021. **50**(9): p. 5397-5434.

268. Habrant, D., et al., *Design of Ionizable Lipids To Overcome the Limiting Step of Endosomal Escape: Application in the Intracellular Delivery of mRNA, DNA, and siRNA.* J Med Chem, 2016. **59**(7): p. 3046-62.

269. Farhood, L., N. Serbina, and L. Huang, *The role of dioleoyl phosphatidylethanolamine in cationic liposome mediated gene transfer.* Biochimica et Biophysica, 1995. **1235**: p. 289-295.

270. Johnsson, M. and K. Edwards, *Phase Behavior and Aggregate Structure in Mixtures of Dioleoylphosphatidylethanolamine and Poly(Ethylene Glycol)-Lipids.* Cell Biophys J, 2001. **80**: p. 313-323.

271. Semple, S.C., et al., *Rational design of cationic lipids for siRNA delivery.* Nat Biotechnol, 2010. **28**(2): p. 172-6.

272. Writer, M., et al., *Analysis and optimization of the cationic lipid component of a lipid/peptide vector formulation for enhanced transfection in vitro and in vivo.* J Liposome Res, 2006. **16**(4): p. 373-89.

273. Du, Z., et al., *The role of the helper lipid on the DNA transfection efficiency of lipopolyplex formulations.* Sci Rep, 2014. **4**: p. 7107.

274. Tam, A., et al., *Lipid nanoparticle formulations for optimal RNA-based topical delivery to murine airways.* Eur J Pharm Sci, 2022. **176**: p. 106234.

275. Zhang, H., et al., *Aerosolizable Lipid Nanoparticles for Pulmonary Delivery of mRNA through Design of Experiments.* Pharmaceutics, 2020. **12**(11): p. 1042.

276. Lokugamage, M.P., et al., *Optimization of lipid nanoparticles for the delivery of nebulized therapeutic mRNA to the lungs.* Nat Biomed Eng, 2021. **5**(9): p. 1059-1068.

277. Miller, J.B., et al., *Non-Viral CRISPR/Cas Gene Editing In Vitro and In Vivo Enabled by Synthetic Nanoparticle Co-Delivery of Cas9 mRNA and sgRNA.* Angew Chem Int Ed Engl, 2017. **56**(4): p. 1059-1063.

278. Jouhet, J., *Importance of the hexagonal lipid phase in biological membrane organization.* Front Plant Sci, 2013. **4**: p. 494.

279. Liu, S., et al., *Membrane-destabilizing ionizable phospholipids for organ-selective mRNA delivery and CRISPR-Cas gene editing.* Nat Mater, 2021. **20**(5): p. 701-710.

280. Li, B., et al., *Combinatorial design of nanoparticles for pulmonary mRNA delivery and genome editing.* Nat Biotechnol, 2023: p. 1410-15.

281. Zhang, Y.N., et al., *Nanoparticle-liver interactions: Cellular uptake and hepatobiliary elimination.* J Control Release, 2016. **240**: p. 332-348.

282. Cheng, Q., et al., *Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR-Cas gene editing.* Nat Nanotechnol, 2020. **15**: p. 313-320.

283. Leal, J., H.D.C. Smyth, and D. Ghosh, *Physicochemical properties of mucus and their impact on transmucosal drug delivery.* Int J Pharm, 2017. **532**(1): p. 555-572.

284. Zhang, H., et al., *PEGylated Chitosan for Nonviral Aerosol and Mucosal Delivery of the CRISPR/Cas9 System in Vitro.* Mol Pharm, 2018. **15**(11): p. 4814-4826.

285. Osman, G., et al., *PEGylated enhanced cell penetrating peptide nanoparticles for lung gene therapy.* J Control Release, 2018. **285**: p. 35-45.

286. Ivanova, R., et al., *A Nanosensor Toolbox for Rapid, Label-Free Measurement of Airway Surface Liquid and Epithelial Cell Function.* ACS Appl Mater Interfaces, 2019. **11**(9): p. 8731-8739.

287. Tagalakis, A.D., et al., *A receptor-targeted nanocomplex vector system optimized for respiratory gene transfer.* Mol Ther, 2008. **16**(5): p. 907-15.

288. Tagalakis, A.D., et al., *Receptor-targeted liposome-peptide nanocomplexes for siRNA delivery.* Biomaterials, 2011. **32**(26): p. 6302-6315.

289. Munye, M.M., et al., *Minicircle DNA Provides Enhanced and Prolonged Transgene Expression Following Airway Gene Transfer.* Sci Rep, 2016. **6**: p. 23125.

290. Grant-Serroukh, D., et al., *Lipid-peptide nanocomplexes for mRNA delivery in vitro and in vivo.* J Control Release, 2022. **348**: p. 786-797.

291. Lin, H., et al., *Air-liquid interface (ALI) culture of human bronchial epithelial cell monolayers as an in vitro model for airway drug transport studies.* J Pharm Sci, 2007. **96**(2): p. 341-50.

292. Ma, X., et al., *Establishment of a type II insulin-like growth factor receptor gene site-integrated SKBR3 cell line using CRISPR/Cas9.* Oncol Lett, 2020. **20**(6): p. 354.

293. UniProt, *CFTR_HUMAN*. UniProt.

294. Bonizzato, A., et al., *Analysis of the complete coding region of the CFTR gene in a cohort of CF patients from North-Eastern Italy: identification of 90% of the mutations* Hum Genet, 1995. **95**: p. 397-402.

295. Sokhi, S., et al., *UK Cystic Fibrosis Registry 2021 Annual Data Report*. 2022, Cystic Fibrosis Trust.

296. He, S., et al., *Bmi-1 over-expression in neural stem/progenitor cells increases proliferation and neurogenesis in culture but has little effect on these functions in vivo.* Developmental Biology, 2009. **328**(2): p. 257-272.

297. Ramalho, A.S., et al., *Assays of CFTR Function In Vitro, Ex Vivo and In Vivo.* Int J Mol Sci, 2022. **23**(3): p. 1437.

298. Noel, S., et al., *Correlating genotype with phenotype using CFTR-mediated whole-cell Cl(-) currents in human nasal epithelial cells.* J Physiol, 2022. **600**(6): p. 1515-1531.

299. Silva, I.A.L., et al., *Assessment of Distinct Electrophysiological Parameters in Rectal Biopsies for the Choice of the Best Diagnosis/Prognosis Biomarkers for Cystic Fibrosis.* Front Physiol, 2020. **11**: p. 604580.

300. Taylor-Cousar, J.L., et al., *Clinical development of triple-combination CFTR modulators for cystic fibrosis patients with one or two F508del alleles.* ERJ Open Res, 2019. **5**(2): p. 82.

301. Park, J.K., et al., *Functional Profiling of CFTR-Directed Therapeutics Using Pediatric Patient-Derived Nasal Epithelial Cell Models.* Front Pediatr, 2020. **8**: p. 536.

302. Nick, H.J., et al., *Measurements of spontaneous CFTR-mediated ion transport without acute channel activation in airway epithelial cultures after modulator exposure.* Sci Rep, 2021. **11**(1): p. 22616.

303. Woollhead, A.M., et al., *Pharmacological activators of AMP-activated protein kinase have different effects on Na+ transport processes across human lung epithelial cells.* Br J Pharmacol, 2007. **151**(8): p. 1204-15.

304. Greenwood, I.A., et al., *KCNQ-encoded channels regulate Na+ transport across H441 lung epithelial cells.* Pflugers Arch, 2009. **457**(4): p. 785-94.

305. Tan, C.D., I.A. Selvanathar, and D.L. Baines, *Cleavage of endogenous gammaENaC and elevated abundance of alphaENaC are associated with increased Na(+) transport in response to apical fluid volume expansion in human H441 airway epithelial cells.* Pflugers Arch, 2011. **462**(3): p. 431-41.

306. Matsui, H., et al., *Evidence for Periciliary Liquid Layer Depletion, Not Abnormal Ion Composition, in the Pathogenesis of Cystic Fibrosis Airways Disease.* Cell, 1998. **95**(7): p. 1005-1015.

307. Sharma, J., et al., *A small molecule that induces translational readthrough of CFTR nonsense mutations by eRF1 depletion.* Nat Commun, 2021. **12**(1): p. 4358.

308. Kim, Y.J., et al., *Gene-specific nonsense-mediated mRNA decay targeting for cystic fibrosis therapy.* Nat Commun, 2022. **13**(1): p. 2978.

309. Kotin, R.M., R.M. Linden, and K.I. Berns, *Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination.* EMBO J, 1992. **11**(13): p. 5071-8.

310. Hockemeyer, D., et al., *Genetic engineering of human pluripotent cells using TALE nucleases.* Nat Biotechnol, 2011. **29**(8): p. 731-4.

311. Hockemeyer, D., et al., *Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases.* Nat Biotechnol, 2009. **27**(9): p. 851-7.

312. Tiyaboonchai, A., et al., *Utilization of the AAVS1 safe harbor locus for hematopoietic specific transgene expression and gene knockdown in human ES cells.* Stem Cell Res, 2014. **12**(3): p. 630-7.

313. Lombardo, A., et al., *Site-specific integration and tailoring of cassette design for sustainable gene transfer.* Nat Methods, 2011. **8**(10): p. 861-9.

314. Oceguera-Yanez, F., et al., *Engineering the AAVS1 locus for consistent and scalable transgene expression in human iPSCs and their differentiated derivatives.* Methods, 2016. **101**: p. 43-55.

315. Kelly, J.J., et al., *Safe harbor-targeted CRISPR-Cas9 homology-independent targeted integration for multimodality reporter gene-based cell tracking.* Sci Adv, 2021. **7**(4): p. eabc3791.

316. *The Human Protein Atlas*.

317. Uhlen, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.

318. Johnson, L., et al., *Efficiency of gene transfer for restoration of normal airway epithelial function in cystic fibrosis.* Nat Gen, 1992. **2**: p. 21-25.

319. Zhang, L., et al., *CFTR delivery to 25% of surface epithelial cells restores normal rates of mucus transport to human cystic fibrosis airway epithelium.* PLoS Biol, 2009. **7**(7): p. e1000155.

320. Rave-Harel, N., et al., *The Molecular Basis of Partial Penetrance of Splicing Mutations in Cystic Fibrosis.* Am. J. Hum. Genet., 1997. **60**: p. 87-94.

321. Mak, V., et al., *Higher proportion of intact exon 9 CFTR mRNA in nasal epithelium compared with vas deferens.* Hum Mol Genet, 1997. **6**(12): p. 2099-2107.

322. Keating, D., et al., *VX-445-Tezacaftor-Ivacaftor in Patients with Cystic Fibrosis and One or Two Phe508del Alleles.* N Engl J Med, 2018. **379**(17): p. 1612-1620.

323. Han, S.T., et al., *Residual function of cystic fibrosis mutants predicts response to small molecule CFTR modulators.* JCI Insight, 2018. **3**(14): p. e121159.

324. Walker, A.J., *Delivery of CRISPR/Cas9 by receptor-targeted nanoparticles as a corrective therapy for Cystic Fibrosis*, in *Institute of Child Health*. 2019, University College London.

325. DeKelver, R.C., et al., *Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome.* Genome Res, 2010. **20**(8): p. 1133-42.

326. Liu, S., et al., *RNA polymerase III is required for the repair of DNA double-strand breaks by homologous recombination.* Cell, 2021. **184**(5): p. 1314-1329 e10.

327. Lees-Miller, S.P. and K. Meek, *Repair of DNA double strand breaks by non-homologous end joining.* Biochimie, 2003. **85**(11): p. 1161-73.

328. Hayashi, H., et al., *Efficient viral delivery of Cas9 into human safe harbor.* Sci Rep, 2020. **10**(1): p. 21474.

329. *UCSC Genome Browser*. 2002, University of California Santa Cruz.

330. Long, Q., Z. Liu, and M. Gullerova, *Sweet Melody or Jazz? Transcription Around DNA Double-Strand Breaks.* Front Mol Biosci, 2021. **8**: p. 655786.

331. Pessina, F., et al., *Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors.* Nat Cell Biol, 2019. **21**(10): p. 1286-1299.

332. Rattan, R., A.U. Bielinska, and M.M. Banaszak Holl, *Quantification of cytosolic plasmid DNA degradation using high-throughput sequencing: implications for gene delivery.* J Gene Med, 2014. **16**(3-4): p. 75-83.

333. Inderbitzin, A., et al., *Quantification of transgene expression in GSH AAVS1 with a novel CRISPR/Cas9-based approach reveals high transcriptional variation.* Mol Ther Methods Clin Dev, 2022. **26**: p. 107-118.

334. Ayrapetov, M.K., et al., *DNA double-strand breaks promote methylation of histone H3 on lysine 9 and transient formation of repressive chromatin.* Proc Natl Acad Sci U S A, 2014. **111**(25): p. 9169-74.

335. Woodall, M., *Functional analysis of human cystic fibrosis epithelial models to investigate the efficacy of genetic therapy techniques*, in *Institute of Infection and Immunity*. 2020, St. George's University London.

336. Thomas, P. and T.G. Smart, *HEK293 cell line: a vehicle for the expression of recombinant proteins.* J Pharmacol Toxicol Methods, 2005. **51**(3): p. 187-200.

337. Renaud, J.B., et al., *Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases.* Cell Rep, 2016. **14**(9): p. 2263-2272.

338. Kan, Y., et al., *Mechanisms of precise genome editing using oligonucleotide donors.* Genome Res, 2017. **27**(7): p. 1099-1111.

339. Xia, E., et al., *TALEN-Mediated Gene Targeting for Cystic Fibrosis-Gene Therapy.* Genes (Basel), 2019. **10**(1): p. 39.

340. Ates, I., et al., *Delivery Approaches for Therapeutic Genome Editing and Challenges.* Genes (Basel), 2020. **11**(10): p. 1113.

341. Roy, S. *FDA panel says Vertex/CRISPR to assess safety risks of gene therapy in follow-up study*. 2023.

342. Jain, A., et al., *CRISPR-Cas9-based treatment of myocilin-associated glaucoma.* Proc Natl Acad Sci U S A, 2017. **114**(42): p. 11199-11204.

343. Wan, T., Q. Pan, and Y. Ping, *Microneedle-assisted genome editing: a transdermal strategy of targeting NLRP3 by CRISPR-Cas9 for synergistic therapy of inflammatory skin disorders.* Sci Adv, 2021. **7**(11): p. eabe2888.

344. Alton, E.W.F.W., et al., *Repeated nebulisation of non-viral CFTR gene therapy in patients with cystic fibrosis: a randomised, double-blind, placebo-controlled, phase 2b trial.* The Lancet Respiratory Medicine, 2015. **3**(9): p. 684-691.

345. Manunta, M.D., et al., *Nebulisation of receptor-targeted nanocomplexes for gene delivery to the airway epithelium.* PLoS One, 2011. **6**(10): p. e26768.

346. Irvine, S.A., et al., *Receptor-targeted nanocomplexes optimized for gene transfer to primary vascular cells and explant cultures of rabbit aorta.* Mol Ther, 2008. **16**(3): p. 508-15.

347. Koltover, I., et al., *An inverted hexagonal phase of cationic liposome-DNA complexes related to DNA release and delivery.* Science, 1998. **281**: p. 78-81.

348. Caracciolo, G., et al., *Efficient escape from endosomes determines the superior efficiency of multicomponent lipoplexes.* J Phys Chem B, 2009. **113**(15): p. 4995-7.

349. Munye, M.M., et al., *Role of liposome and peptide in the synergistic enhancement of transfection with a lipopolyplex vector.* Sci Rep, 2015. **5**: p. 9292.

350. Danaei, M., et al., *Impact of Particle Size and Polydispersity Index on the Clinical Applications of Lipidic Nanocarrier Systems.* Pharmaceutics, 2018. **10**(2): p. 57.

351. Rehman, Z., I.S. Zuhorn, and D. Hoekstra, *How cationic lipids transfer nucleic acids into cells and across cellular membranes: recent advances.* J Control Release, 2013. **166**(1): p. 46-56.

352. Kwok, A. and S.L. Hart, *Comparative structural and functional studies of nanoparticle formulations for DNA and siRNA delivery.* Nanomedicine, 2011. **7**(2): p. 210-9.

353. Tagalakis, A.D., et al., *Comparison of nanocomplexes with branched and linear peptides for siRNA delivery.* Biomacromolecules, 2013. **14**(3): p. 761-70.

354. Jinek, M., et al., *Structures of Cas9 endonucleases reveal RNA-mediated conformational activation.* Science, 2014. **343**: p. 1215.

355. *Dharmacon™ Edit-R™ synthetic guide RNA resuspension protocol*. 2019.

356. Hart, S.L., L. Collins, and J.W. Fabre, *Integrin-mediated transfection with peptides containing arginine-glycine-aspartic acid domains.* Nature, 1997. **4**: p. 1225-1230.

357. Liu, X., et al., *Cationic lipid nanoparticle production by microfluidization for siRNA delivery.* Int J Pharm, 2021. **605**: p. 120772.

358. Sanders, N.N., et al., *Cystic fibrosis sputum: a barrier to the transport of nanospheres.* Am J Respir Crit Care Med, 2000. **162**(5): p. 1905-11.

359. Suk, J.S., et al., *The penetration of fresh undiluted sputum expectorated by cystic fibrosis patients by non-adhesive polymer nanoparticles.* Biomaterials, 2009. **30**(13): p. 2591-7.

360. Broughton-Head, V.J., et al., *Actin limits enhancement of nanoparticle diffusion through cystic fibrosis sputum by mucolytics.* Pulm Pharmacol Ther, 2007. **20**(6): p. 708-17.

361. Piva, T.C., et al., *Extracellular DNA in sputum is associated with pulmonary function and hospitalization in patients with cystic fibrosis.* Respir Med, 2020. **172**: p. 106144.

362. Rouillard, K.R., et al., *Effects of Mucin and DNA Concentrations in Airway Mucus on Pseudomonas aeruginosa Biofilm Recalcitrance.* mSphere, 2022. **7**: p. e0029122.

363. Onuma, H., Y. Sato, and H. Harashima, *Lipid nanoparticle-based ribonucleoprotein delivery for in vivo genome editing.* J Control Release, 2023. **355**: p. 406-416.

364. Zhang, S., et al., *Size-Dependent Endocytosis of Nanoparticles.* Adv Mater, 2009. **21**: p. 419-424.

365. Wu, Y.S., et al., *ORKAMBI-Mediated Rescue of Mucociliary Clearance in Cystic Fibrosis Primary Respiratory Cultures Is Enhanced by Arginine Uptake, Arginase Inhibition, and Promotion of Nitric Oxide Signaling to the Cystic Fibrosis Transmembrane Conductance Regulator Channel.* Mol Pharmacol, 2019. **96**(4): p. 515-525.

366. Mortimer, I., et al., *Cationic lipid-mediated transfection of cells in culture requires mitotic activity.* Nature, 1999. **6**: p. 403-411.

367. Allan, K.M., et al., *Collection, Expansion, and Differentiation of Primary Human Nasal Epithelial Cell Models for Quantification of Cilia Beat Frequency.* J Vis Exp, 2021. **177**.

368. Thomas, B., et al. *Ciliary function of the nasal and bronchial epithelium in children with cystic fibrosis*. in *European Respiratory Society Annual Congress 2012*. 2012. Vienna, Austria.

369. Alikadic, S., et al. *Ciliary beat frequency in nasal and bronchial epithelial cells in patients with cystic fibrosis*. in *European Respiratory Society Annual Congress 2011*. 2011. Amsterdam, Netherlands.

370. Chioccioli, M., et al., *Phenotyping ciliary dynamics and coordination in response to CFTR-modulators in Cystic Fibrosis respiratory epithelial cells.* Nat Commun, 2019. **10**(1): p. 1763.

371. Schmid, A., et al., *Decreased soluble adenylyl cyclase activity in cystic fibrosis is related to defective apical bicarbonate exchange and affects ciliary beat frequency regulation.* J Biol Chem, 2010. **285**(39): p. 29998-30007.

372. Liu, Z., et al., *Human Nasal Epithelial Organoids for Therapeutic Development in Cystic Fibrosis.* Genes (Basel), 2020. **11**(6): p. 603.

373. Sachs, N., et al., *Long-term expanding human airway organoids for disease modeling.* EMBO J, 2019. **38**(4): p. e100300.

374. Amatngalim, G.D., et al., *Measuring cystic fibrosis drug responses in organoids derived from 2D differentiated nasal epithelia.* Life Sci Alliance, 2022. **5**(12): p. e202101320.

375. Tornabene, P., et al., *Therapeutic homology-independent targeted integration in retina and liver.* Nat Commun, 2022. **13**(1): p. 1963.

376. Rezazade Bazaz, M., M.M. Ghahramani Seno, and H. Dehghani, *Transposase-CRISPR mediated targeted integration (TransCRISTI) in the human genome.* Sci Rep, 2022. **12**(1): p. 3390.

377. Yarnall, M.T.N., et al., *Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases.* Nat Biotechnol, 2023. **41**(4): p. 500-512.

378. da Silva, J.F., et al., *Click editing enables programmable genome writing using DNA polymerases and HUH endonucleases.* bioRxiv, 2023.

379. Bartman, C.M., et al., *Passive siRNA transfection method for gene knockdown in air-liquid interface airway epithelial cell cultures.* Am J Physiol Lung Cell Mol Physiol, 2021. **321**: p. 280-286.

380. Hirst, R.A., et al., *Culture of primary ciliary dyskinesia epithelial cells at air-liquid interface can alter ciliary phenotype but remains a robust and informative diagnostic aid.* PLoS One, 2014. **9**(2): p. e89675.

381. *Peptide Property Calculator*. 2015; Available from: https://pepcalc.com/.

382. Fuchs, S. *GRAVY Calculator*. 2011; Available from: https://www.gravy-calculator.de/.

383. Duncan, G.A., et al., *An Adeno-Associated Viral Vector Capable of Penetrating the Mucus Barrier to Inhaled Gene Therapy.* Mol Ther Methods Clin Dev, 2018. **9**: p. 296-304.

384. Fitzpatrick, Z., et al., *Influence of Pre-existing Anti-capsid Neutralizing and Binding Antibodies on AAV Vector Transduction.* Mol Ther Methods Clin Dev, 2018. **9**: p. 119-129.

385. Mingozzi, F., et al., *CD8(+) T-cell responses to adeno-associated virus capsid in humans.* Nat Med, 2007. **13**(4): p. 419-22.

386. Ruysseveldt, E., K. Martens, and B. Steelant, *Airway Basal Cells, Protectors of Epithelial Walls in Health and Respiratory Diseases.* Front Allergy, 2021. **2**: p. 787128.

387. Mourdikoudis, S., R.M. Pallares, and N.T.K. Thanh, *Characterization techniques for nanoparticles: comparison and complementarity upon studying nanoparticle properties.* Nanoscale, 2018. **10**(27): p. 12871-12934.

388. Lewis, M.M., et al., *Optimization of ionizable lipids for aerosolizable mRNA lipid nanoparticles.* Bioengineering & Translational Medicine, 2023. **8**(6).

1. Appendix

ASL depths (measured via SICM) of nasal epithelial cells. Three replicate wells shown

per genotype, n = 21 measurements per well. One CFNE BMI-1 R709X well showed

abnormally high ASL depth measurements and was discarded from the overall analysis

depicted in Fig. 3.4.

ASL Depth
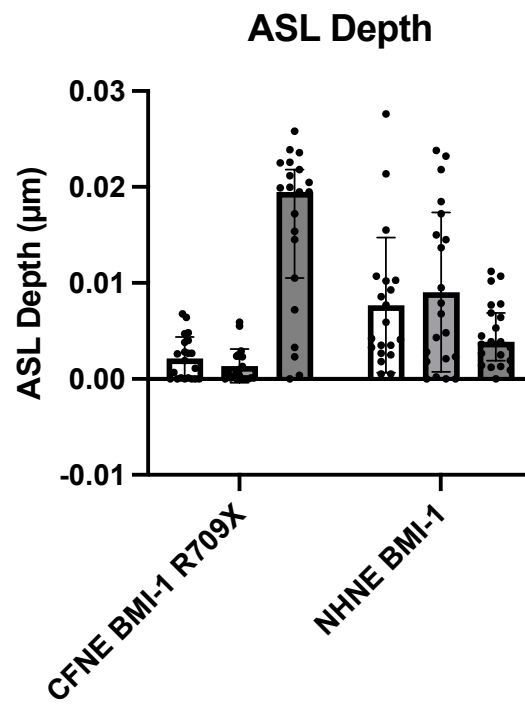
Sequence of pExcised plasmid used in Chapter 5. Sequence was confirmed via GeneWiz

Plasmid-EZ sequencing. Relevant features are highlighted.

ACCCTGTTATCCCTAGATGACATTACCCTGTTATCCCAGATGACATTACCCTGTTATCCCTAGATACAT

TACCCTGTTATCCCAGATGACATACCCTGTTATCCCTAGATGACATTACCCTGTTATCCCAGATGACAT

TACCCTGTTATCCCTAGATACATTACCCTGTTATCCCAGATGACATACCCTGTTATCCCTAGATGACAT

TACCCTGTTATCCCAGATGACATTACCCTGTTATCCCTAGATACATTACCCTGTTATCCCAGATGACAT

ACCCTGTTATCCCTAGATGACATTACCCTGTTATCCCAGATGACATTACCCTGTTATCCCTAGATACAT

TACCCTGTTATCCCAGATGACATACCCTGTTATCCCTAGATGACATTACCCTGTTATCCCAGATGACAT

TACCCTGTTATCCCTAGATACATTACCCTGTTATCCCAGATGACATACCCTGTTATCCCTAGATGACAT

TACCCTGTTATCCCAGATGACATTACCCTGTTATCCCTAGATACATTACCCTGTTATCCCAGATGACAT

ACCCTGTTATCCCTAGATGACATTACCCTGTTATCCCAGATGACATTACCCTGTTATCCCTAGATACAT

TACCCTGTTATCCCAGATGACATACCCTGTTATCCCTAGATGACATTACCCTGTTATCCCAGATAAACT

CAATGATGATGATGATGATGGTCGAGACTCAGCGGCCGCGGTGCCAGGGCGTGCCCTTGGGCTCCC

CGGGCGCGACTAGTGAATTCCTGCAGGACTAGTGACGTCGTCACCAATCCTGTCCCTAGTGGTCGAC

CTTAATAGCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACTGCAGCTGACC

TCTTCTCTTCCTCCCACAGGCCCGATCGGGCAGCGGAGAGGGCAGAGGAAGTCTTCTAACATGCGG

TGACGTGGAAGAGAATCCCGGCCCTATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGC

CCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAG

GGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCT

GGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAA

GCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAG

GACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT

CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACT

ACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGA

TCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCG

GCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCC

CAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCAT

GGACGAGCTGTACAAGTAAAGCGGCCTCGACGCTAGCTGGCCAGACATGATAAGATACATTGATGA

GTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATT

GCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTT

CAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAA

TTCGCGGCCGCCGCCAACTCGGTACCATAACTTCGTATAGTTGATAATTCACTGGCCGTCGAGCTCGT

CACCAATCCTGTCCCTAGTGGGGCCCAGTCGCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGG

GGGTAATCAGCATCATGATGTGGTACCACATCATGATGCTGATTATAAGAATGCGGCCGCCACACTC

TAGTGGATATCGAGTTAATAATTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCG

AATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAG

CAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGAT

GAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCACGAC

GAGATCCTCGCCGTCGGGCATGCTCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTG

ATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGC

GATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCATTGCAT

CAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTT

CGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAGGAACGCC

CGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTTGCAGTTCATTCAGGGCACCGGACAGGTC

GGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGC

CGATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTG

CAATCCATCTTGTTCAATCATGCGAAACGATCCTCATCCTGTCTCTTGATCAGAGCTTGATCCCCTGCG

CCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGCAGGGCTTCCCAACCTTACCAGAG

GGCGCCCCAGCTGGCAATTCCGGTTCGCTTGCTGTCCATAAAACCGCCCAGTCTAGCTATCGCCATG

TAAGCCCACTGCAAGCTACCTGCTTTCTCTTTGCGCTTGCGTTTTCCCTTGTCCAGATAGCCCAGTAGC

TGACATTCATCCGGGGTCAGCACCGTTTCTGCGGACTGGCTTTCTACGCGCTCGAGGGGGGGCCAA

ACGGTCTCCAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGA

ACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCC

CATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGT

AGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCT

GTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGA

AGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAATTAAGCAG

AAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTTGTTTATTTTTCTAAATACATTC

AAATATGTATCCGCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCG

TAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAA

AAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAA

CTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTT

CAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTG

GCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGG

GCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACC

TACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTA

AGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCANGGGGAAACGCCTGGTATCTTTA

TAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG

AGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCA

CATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAC

CGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTG

ATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAAT

CTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTG

CGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTA

CAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGC

GCGAGGCAGCAGATCAATTCGCGCGCGAAGGCGAAGCGGCATGCATAATGTGCCTGTCAAATGGA

CGAAGCAGGGATTCTGCAAACCCTATGCTACTCCGTCAAGCCGTCAATTGTCTGATTCGTTACCAATT

ATGACAACTTGACGGCTACATCATTCACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGGCTGGCC

CCGGTGCATTTTTTAAATACCCGCGAGAAATAGAGTTGATCGTCAAAACCAACATTGCGACCGACGG

TGGCGATAGGCATCCGGGTGGTGCTCAAAAGCAGCTTCGCCTGGCTGATACGTTGGTCCTCGCGCC

AGCTTAAGACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGACAAGCAAA

CATGCTGTGCGACGCTGGCGATTCTTCTCTCATCCGCCAAAACAGCCAAGCTGGAGACCGTTTGACA

TTACCCTGTTATCCCTAGATACATTACCCTGTTATCCCAGATGACAT

Plasmid backbone

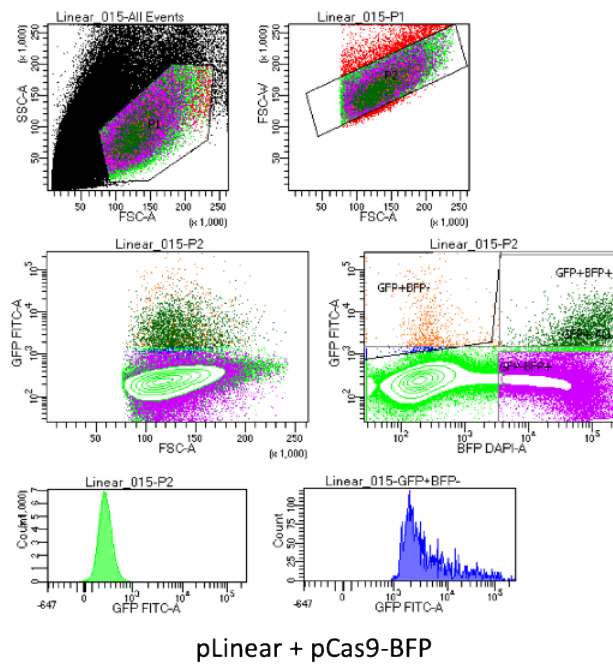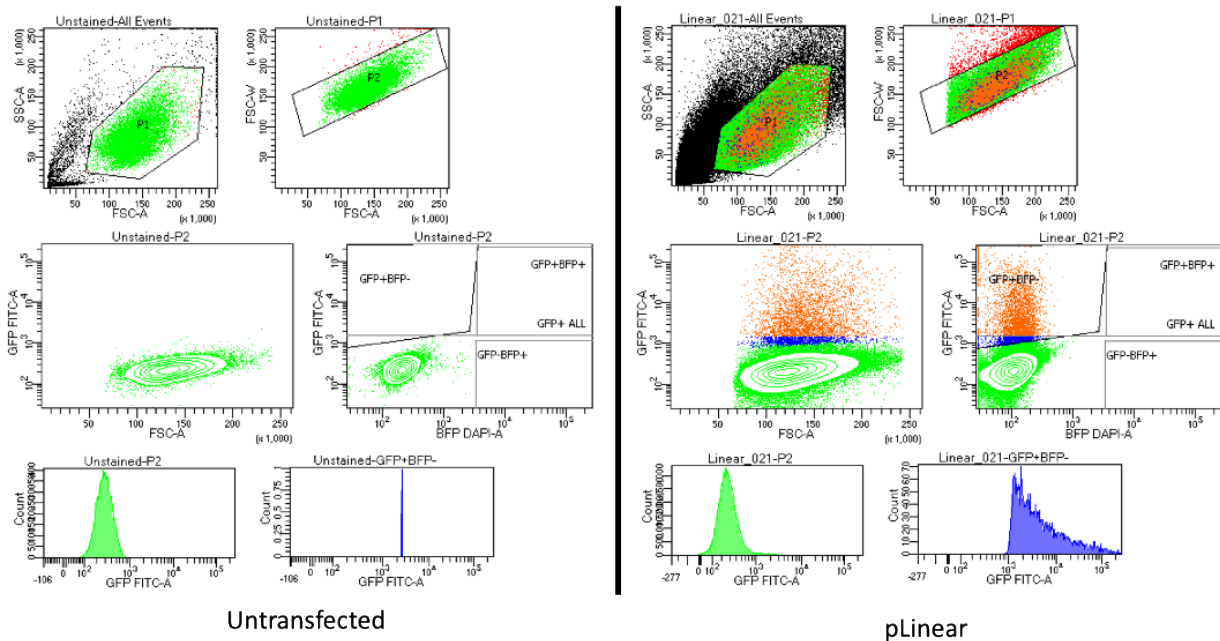A78 gRNAs (one upstream and one downstream of GFP construct)

Splice acceptor

T2A sequence

eGFP cDNA

SV40 poly-A tail

Kanamycin resistance

Untransfected

pLinear

pLinear + pCas9-BFP

Example of gating for flow cytometry and cell sorting. One replicate shown for

untransfected and transfected (pLinear only and pLinear + pCas9-BFP) cells. Gate P1

was set by forward (area) and side (area) scatter to isolate cells and exclude debris.

Gate P2 was set by forward scatter (area and width) to isolate singlets. Fluorescence-

based gates (GFP+BFP-, GFP-BFP+, GFP+BFP+, and GFP+ ALL) set based on

untransfected control. Unstained: untransfected cells. "Linear_021": cells transfected with pLinear donor only. "Linear_015": cells transfected with pLinear and pCas9-BFP. All gates for flow/sorting were set using this method, with gates kept consistent between replicates.

Generation and use of NHBE BMI-1 GFP cells. A) Schematic of construct integrated into cells' genomes using lentivirus. GFP was under the control of the hPGK promoter. 14 days post-transduction, GFP+ cells were isolated via FACS. B) Validation of gRNAs to knock down GFP in submerged NHBE BMI-1 cells transduced lentivirally with GFP. Cells were transfected with Cas9 mRNA using the lipid C18DOPE and Peptide Y at a weight ratio of 1:3:4 RNP:lipid:peptide. gRNA 3 was used for ALI transfection experiments in Chapter 6. From Walker, A.J., *Delivery of CRISPR/Cas9 by receptor-targeted nanoparticles as a corrective therapy for Cystic Fibrosis*, in *Institute of Child Health*. 2019, University College London.
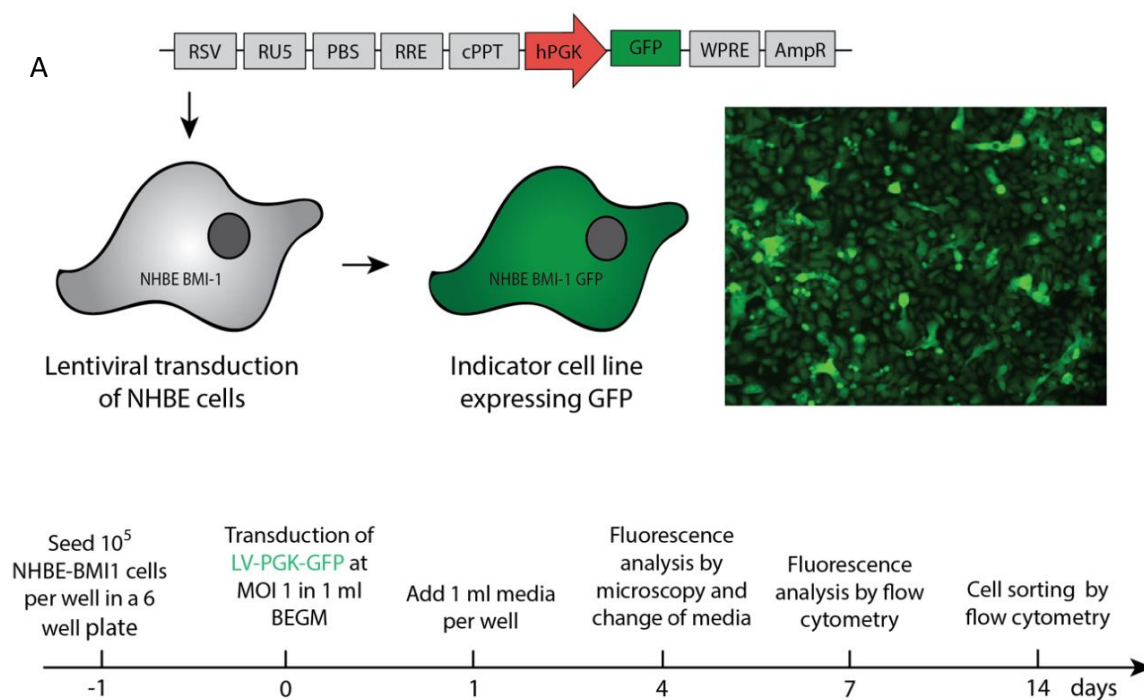
B
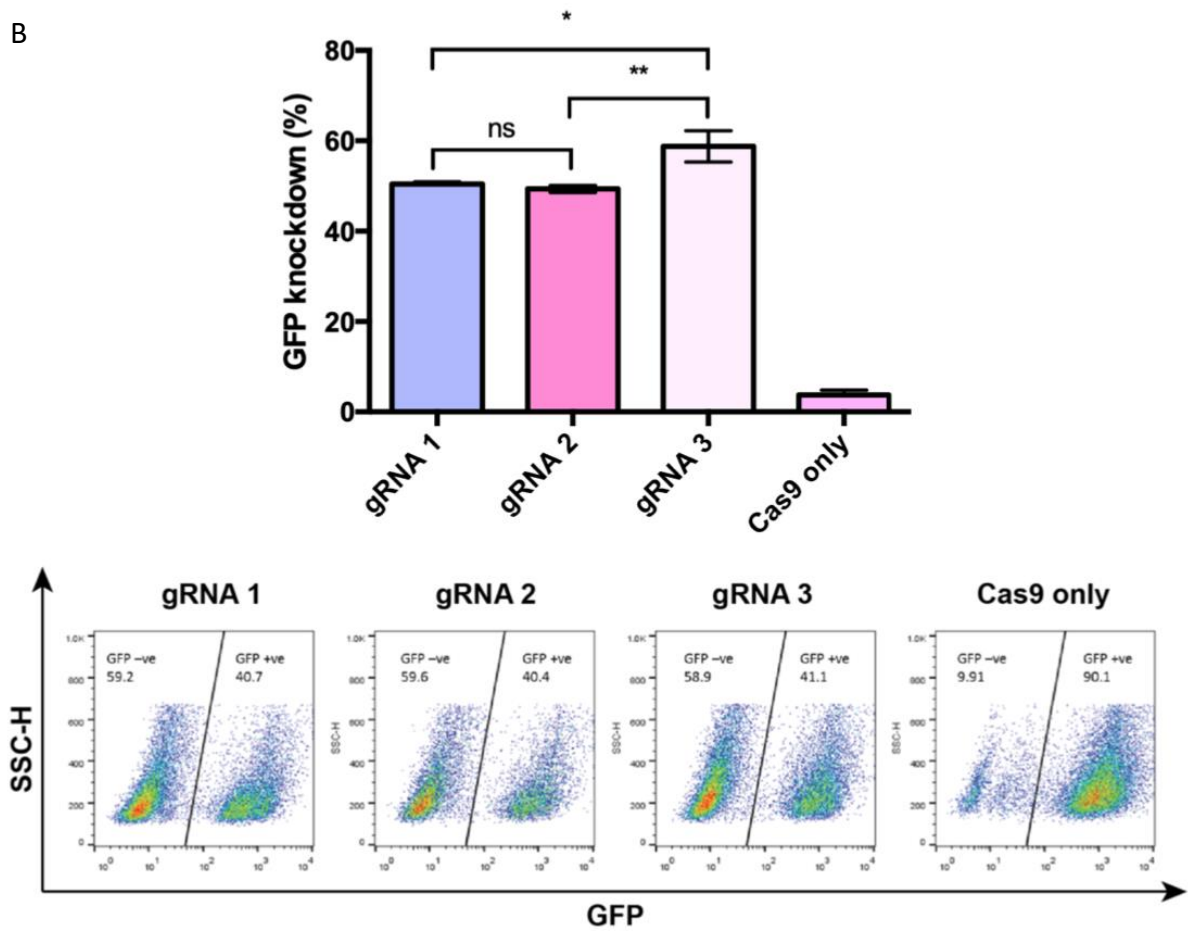
Macro used in Fiji to re-slice and z-project images of ALI cultures of GFP-expressing NHBE BMI-1 cells that had been transfected with RTNs carrying RFP-tagged Cas9 RNPs (Fig. 5.15). A horizontal section of each image 10 pixels high and 1012 wide (the full width of the image) was selected using the rectangular selection tool; sections were selected that had RFP and GFP signal overlapping.

```
run("Duplicate...", "duplicate");

run("Subtract Background...", "rolling=50 sliding disable");

run("Reslice [/]...", "output=5.600 start=Top flip avoid");

run("Scale...", "x=1.0 y=12.6 z=1.0 width=1024 height=189 depth=13 interpolation=Bilinear
average create");

run("Z Project...", "projection=[Max Intensity]");

run("Gaussian Blur...", "sigma=2");
```