UNIVERSITY OF TRENTO - Italy
Department CIBIO

International PhD Program in Biomolecular Sciences

Department of Cellular, Computational and Integrative Biology
32nd Cycle

*Harnessing CRISPR technology for the treatment of **cystic fibrosis***

Tutor

Anna Cereseto

Department CIBIO

University of Trento

Zeger Debyser

Department of Pharmaceutical and Pharmacological Sciences

KU Leuven

Co-promoter

Marianne Carlon

Department of Pharmaceutical and Pharmacological Sciences

KU Leuven

Ph.D. Thesis of

Giulia Maule

Department CIBIO

University of Trento

Academic Year 2018-2019

Declaration of Authorship

I, Giulia Maule, confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Trento, 3rd February 2020

INDEX

ABSTRACT……7

INTRODUCTION……8

CYSTIC FIBROSIS……8

THE CFTR CHANNEL ……9

FROM GENOTYPE TO PHENOTYPE: CFTR MUTATION CLASSES……10

3272-26A>G mutation……11

3849+10kbC>T mutation……11

THERAPIES TARGETING CFTR……12

Pharmacological therapies……12

Gene therapy for CF……14

OTHER CF THERAPIES……17

THE CRISPR/CAS SYSTEM……18

Classification of CRISPR/Cas systems……19

CRISPR type II: Cas9 endonuclease……20

CRISPR type V: Cas12a endonuclease……21

APPLICATIONS OF CRISPR NUCLEASES……22

GENE THERAPY WITH CRISPR……25

SPLICING……27

Splicing regulatory sequences……27

Splicing mechanism……29

AIM OF THE THESIS……30

METHODS……31

RESULTS……36

Set up of a genome editing strategy to correct the 3272-26A>G splicing mutation in the CFTR gene……36

Specificity of AsCas12a based correction in 3272-26A>G models……42

3272-26A>G splicing correction in primary airway cells……45

AsCas12a editing restored CFTR wt transcripts in intestinal organoids……49

CFTR channel function is restored after AsCas12a-crRNA+11 editing……52

SpCas9-based correction of 3849+10kbC>T mutation……54
AsCas12a with a single crRNA efficiently correct 3849+10kbC>T mutation……..…58

DISCUSSION……64

BIBLIOGRAPHY……68

APPENDIX……..……..……..77

The data presented in this thesis are published in "Allele specific repair of splicing mutations in cystic fibrosis through AsCas12a genome editing", Nature Communication, 2019[1].

ABSTRACT

Cystic fibrosis is an autosomal recessive disease caused by mutations in the *CFTR* gene. A significant number of mutations (~13%) alter the correct splicing of the *CFTR* gene, causing the transcription of aberrant transcripts resulting in the production of a non-functional CFTR channel. We focus our research on two intronic CF causing mutations, 3272-26A>G and 3849+10kbC>T that create a new acceptor and donor splice site, respectively, generating in the inclusion of intronic portions into the mRNA.

We developed a new genome editing approach to permanently correct the abovementioned mutations by means of CRISPR nucleases. We exploited the use of either *Streptococcus pyogenes* Cas9, SpCas9, or *Acidaminococcus sp. BV3L6*, AsCas12a, to edit the aberrant splicing sites and restore the production of the correct transcript, avoiding modifications of the *CFTR* coding sequence.

A comparative analysis between SpCas9 and AsCas12a revealed that the use of AsCas12a with a single crRNA efficiently edits the target loci, producing correctly spliced mRNAs in both 3272-26A>G and 3849+10kbC>T mutations. Furthermore, this genetic repair strategy proved to be highly specific, exhibiting a strong discrimination between the mutated and the wild-type allele and no detectable off-target activity with genome-wide analysis.

The selected crRNAs were tested in patients derived primary airway cells and intestinal organoids compound heterozygous for the 3272-26A>G or 3849+10kbC>T mutations, that are considered relevant CF models for translational research. The efficient splicing repair and the complete recovery of CFTR channel activity observed confirmed the goodness of the proposed gene editing strategy.

These results demonstrated that allele-specific genome editing with AsCas12a can correct aberrant *CFTR* splicing mutations, paving the way for a permanent splicing correction in genetic diseases.

CYSTIC FIBROSIS

Cystic fibrosis is an autosomal recessive disorder caused by mutations in the Cystic Fibrosis Transmembrane Regulator (*CFTR*) gene[2,3]. It is the most common lethal genetic disease in the Caucasian population, occurring in approximately 1/3500 births but varying according to the population studied[4,5].

The *CFTR* gene encodes a transmembrane chloride (Cl^-) and bicarbonate (HCO_3^-) ion channel expressed on the apical surface of epithelial cells in multiple exocrine organs, where it regulates salt and fluid homeostasis[6,7]. The impact of *CFTR* mutations differs among different affected organs, such as the airways, exocrine glands of the pancreas, the small and large intestine, sweat glands and bile ducts[8]. Despite the variability of CF symptoms, lung disease remains the main cause of morbidity and mortality in patients[9].

In the respiratory epithelia, mutations in CFTR impair chloride secretion, sodium transport mediated by epithelial Na+ channel (ENaC)[10] and water transport that results in cellular alkalinity and luminal acidification. The acidification and dehydration of the mucus layer cause a decreased mucociliary clearance and consequent mucus accumulation, resulting in a viscous gel phase, very difficult to be removed, and obstruction of the airways[11]. This predisposes CF patients to infections, inflammation and biofilm formation that destroys the pulmonary epithelia causing fibrosis and respiratory failure (Figure 1)[12].

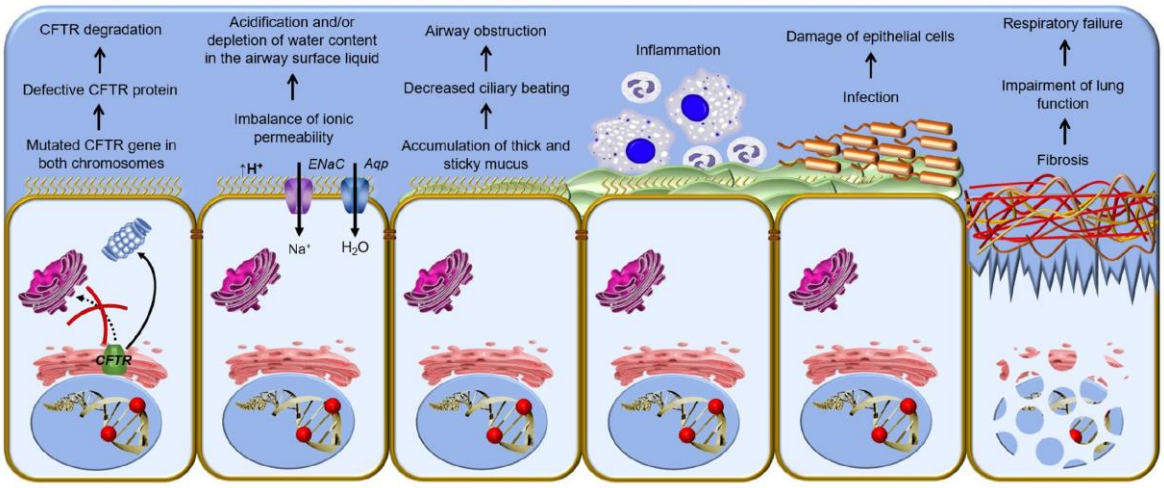


Figure 1. **Physiopathological cascade of CF in airways epithelia**. Dysfunction of CFTR cause an imbalance of anionic permeability leading to the accumulation of thick mucus on the airway epithelia. Infection and inflammation ultimately cause tissue fibrosis and respiratory failure. *Adapted from Lopes-Pacheco, Front. Pharmacol, 2016.*

CFTR Channel

The *CFTR* gene is located on the q31.2 locus of chromosome 7, is 230000 nucleotide in length, with 27 exons and 26 introns, and encodes for a 1480-residue long (170 kDa) membrane protein member of the ATP-binding cassette (ABC) transporter family[13]. Proteins belonging to this family use ATP hydrolysis to pump substrates across cellular membranes. While other members transport substrates against their chemical gradients, CFTR is the only known ABC family member that functions as an ion channel that conducts anions down their electrochemical gradient[14]. ABC transporters have a common architecture that consists of two transmembrane domains (TMD1 and TMD2), that form the channel pore, and two cytoplasmic nucleotide-binding domains (NBD1 and NBD2), involved in ATP hydrolysis[13]. Moreover, CFTR has an additional unique regulatory domain (R), its phosphorylation by protein kinase A (PKA) determines the opening of the channel, and longer extensions at the N- and C- terminus (80 or 30 nucleotides, respectively), probably involved in the regulation of channel function or interaction with cellular components[15].

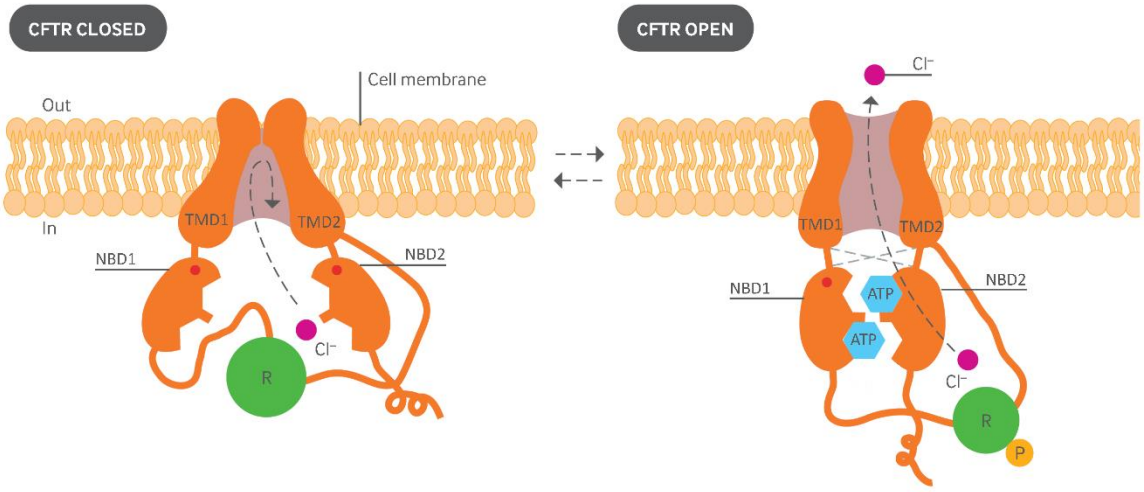


Figure 2: **Configuration of the closed and open state of CFTR channel.** The TMD1 and TMD2 form the transport pathway, whose gaiting is regulated by the NBD1, NBD2 and R domain. Phosphorylation of the R domain and hydrolysis of ATP determine the open conformation of the channel. *Adapted from Quon and Rowe, BMJ. 2016.*

FROM GENOTYPE TO PHENOTYPE: CFTR MUTATION CLASSES

More than 2,000 mutations have been described in the *CFTR* gene[16]. To date, only 412 mutations have been fully characterized and 346 have been identified as disease causing mutations[17]. These variants are grouped into seven classes according to the defect caused in the CFTR protein (Figure 2)[18,19].

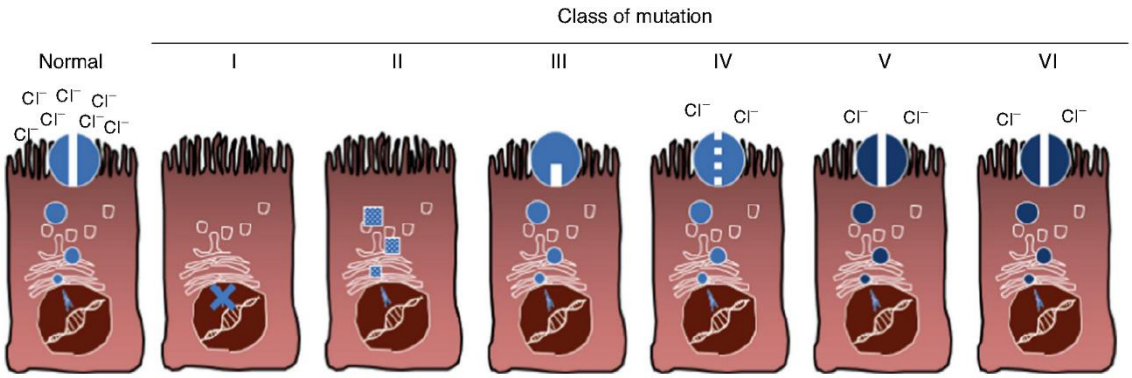
Class I accounts for nonsense and frameshift mutations that cause the formation of a premature termination codon (PTC). This results in the loss of CFTR function at the apical layer of epithelial cells due to both nonsense-mediated mRNA decay (NMD)[20] or interruption of ribosomal translation and the production of a truncated unstable CFTR protein.

Class II mutations result in defective processing and trafficking of the CFTR protein causing its degradation by the proteasome and severely reducing the amount of protein at the apical membrane. This class includes the most common CF causing mutation, the ΔF508, about 80% of patients carry it on one allele[21]. The ΔF508 is characterized by the deletion of a phenylalanine at 508 position. This protein is not correctly released from the endoplasmic reticulum resulting in its premature degradation.

Class III mutations produce a CFTR channel that localizes at the cell membrane but with a defective regulation of the gating process, leading to a severe impairment of the open state of the ion channel.

Class IV grouped defects implicated in the formation of the channel pore. These mutations cause a defective ion conductance, where negatively charged ions poorly pass through the CFTR channel, even in the open conformation.

Class V mutations lead to a decreased amount of functional CFTR at the apical membrane, due to a reduced correctly spliced CFTR transcript. These mutations are often located in introns creating alternative splice sites or modifying splicing silencer or enhancer motifs. The severity of disease in class V patients has been postulated to be inversely related to the level of correctly spliced transcripts. Class VI mutations are characterized by membrane instability of CFTR, as a result of premature recycling at the plasma membrane and degradation by lysosomes. Class VII has been added recently and includes all mutations, like large deletions and frameshift mutations that cannot be treated with pharmacotherapy[19,22–25].



Molecular defect	No synthesis	Block in processing	Block in regulation	Reduced conductance	Reduced synthesis	Reduced half-life
Functional abnormality	Protein is not synthesized	Folding defect	Channel opening defect	Ion transport defect	Decreased protein synthesis	Decreased half-life of the protein
Main mutations	Gly542X Trp128X Arg553X 621+1G→T	Phe508del Asn1303Lys Ile507del Arg560Thr	Gly551Asp Gly178Arg Gly551Ser Ser549Asn	Arg117His Arg347Pro Arg117Cys Arg334Trp	3849+10kbC→T 2789+5G→A 3272-26A→G 5T	4326delTC Gln1412X 4279insA

Figure 3. **Cystic fibrosis mutation classes**. *Adapted from Quintana-Gallego et al., Arch. Bronconeumol., 2014.*

3272-26A>G mutation

3272-26A>G is a class V point mutation in intron 19 that leads to the creation of an alternative acceptor splice site competing with the canonical one during RNA processing[26]. The inclusion of 25 extra nucleotides from intron 19 into the mRNA causes a frame shift and the formation of a premature termination codon. The use of alternative splice site causes a reduction in the mRNA level and consequently in the protein level[27]. Patients carrying this mutation present a phenotypic variability in age at diagnosis, lung function and pancreatic disease. This milder and variable phenotype found in these patients is probably due to the partial production of correctly spliced transcripts (about 5%). This mutation is widespread in Europe with the highest frequency in Portugal (2%)[28].

3849+10kbC>T mutation

The 3849+10kbC>T mutation is a nucleotide substitution that creates a novel donor splice site deep in intron 22 of *CFTR* gene. The activation of a cryptic acceptor splice site 84 nucleotides upstream leads to the insertion of a new cryptic exon into the mRNA, containing an in frame premature termination codon[29]. This mutation

belongs to class V and it is associated with a relatively mild form of CF with high symptoms variability, including age at diagnosis, pulmonary and pancreatic disease. The 3849+10kbC>T mutation is twelfth most common mutation worldwide, with a frequency of 0.2% and a prevalence in subpopulations such as Polish (4%) and Ashkenazi Jewish (6%)[30].

THERAPIES TARGETING CFTR

Since the discovery that CF is caused by mutations in the *CFTR* gene, several advancements have been made in understanding CFTR processing, trafficking and function at the plasma membrane, allowing the development of new therapies[11]. Despite the numerous improvements obtained nowadays, scientists are still searching for a cure for CF patients.

Pharmacological therapies

The available pharmacological therapies mainly aimed at the reduction of clinical symptoms, like chronic bacterial infections, inflammation and intestinal-airway blockages[5]. Nevertheless, a new class of small molecules has been identified by high-throughput screening (HTS), called CFTR modulators, that directly target the molecular defects to increase CFTR activity[31,32]. Based on the pharmacodynamics of these drugs, different groups have been identified: read-through agents, correctors, potentiators and stabilizers[11].

Read-through agents intend to rescue the protein synthesis in CF patients presenting a premature stop codons, PTCs (Class I), promoting the ribosomal translation process till the normal end of the trancripts[33]. First observation of read-through effect was observed with aminoglycoside antibiotics, like gentamicin and tobramycin, which are commonly used to eradicate Gram negative bacteria, such as *P. aeruginosa*[33,34]. These antibiotics impair the proofreading of ribosomes allowing the erroneous codon-anticodon pairing and the addition of an amino acid at the PTC site. Unfortunately, the toxic effect of aminoglycoside antibiotics makes them not suitable for CF patients treatment. Among the new drugs identified by HTS, Ataluren was found capable to suppress the G542X nonsense mutation, however,

it demonstrated limited efficacy in CF rectal organoids[35] and in the first long-term phase III trial[36].

Correctors have been developed to rescue protein folding and trafficking of the CFTR protein to the plasma membrane by acting as pharmacological chaperones, directly binding to the protein, or by modulating the CFTR interactome[37–39]. Principal targets of this group of drugs are class II mutations, where CFTR is misfolded and prematurely degraded. Lumacaftor, tezacaftor and elexacaftor are all correctors that improve maturation and conformational stability of CFTR and are all been approved for patients therapy[40,41].

While rescuing transfer of the protein to the plasma membrane is fundamental to guarantee its function, potentiators are necessary when CFTR is present at the plasma membrane but exhibits reduced or no gating activity, enabling the restoration of channel conductance[8]. Ivacaftor was identified through HTS screening as an effective potentiator, able to increase CFTR activity from 5% to 50% of wild-type levels, stabilizing its open state[33]. Several studies have demonstrated that combinations of drugs with different mechanisms of action are required to obtain a therapeutic effect in CF patients[23,42–44], leading to the approval for clinical use of different combinations of correctors and potentiators for several different mutations, with the last US FDA approval (October 2019) of Trikafta™ (Elexacaftor/Ivacaftor/ Tezacaftor)[40].

Stabilizers increase the stability of the CFTR channel both at the plasma membrane and inside the cells, diminishing its degradation rate. These modulators can be used to correct class VI mutations[11], that account for an increased instability of the protein, or to stabilize the restored activity of CFTR by other modulators, for example in patients with the ΔF508 mutation[44]. N91115 (Cavosonstat) is now in preclinical studies in combination with lumacaftor/ivacaftor for patients homozygous for ΔF508 mutation[45], and with ivacaftor for patients heterozygous for ΔF508 and gating mutations[46].

CFTR modulators represent a huge promise for the treatment of at least 33 different mutations, accounting for approximately 90% of CF patients. However, 90% of disease-causing mutations cannot be yet repaired by any available small molecules[47]. Moreover, life-long drugs uptake is not ideal for patients, due to tolerability, side effects and for the elevated cost of the treatment, posing the need of other strategies aimed to correct the genetic defects.

Finally, for class V CF mutations, that include all those defects that alter the correct splicing pattern of the *CFTR* gene, aberrant transcripts can be corrected by antisense oligonucleotides (ASOs) therapy, allowing the modulation of the splicing reaction impeding the recognition of the alternative splice site created by the mutation. The binding of the ASO on the pre-spliced RNA leads to the restoration of the correct transcript and thus the translation of a functional protein[30,48,49].

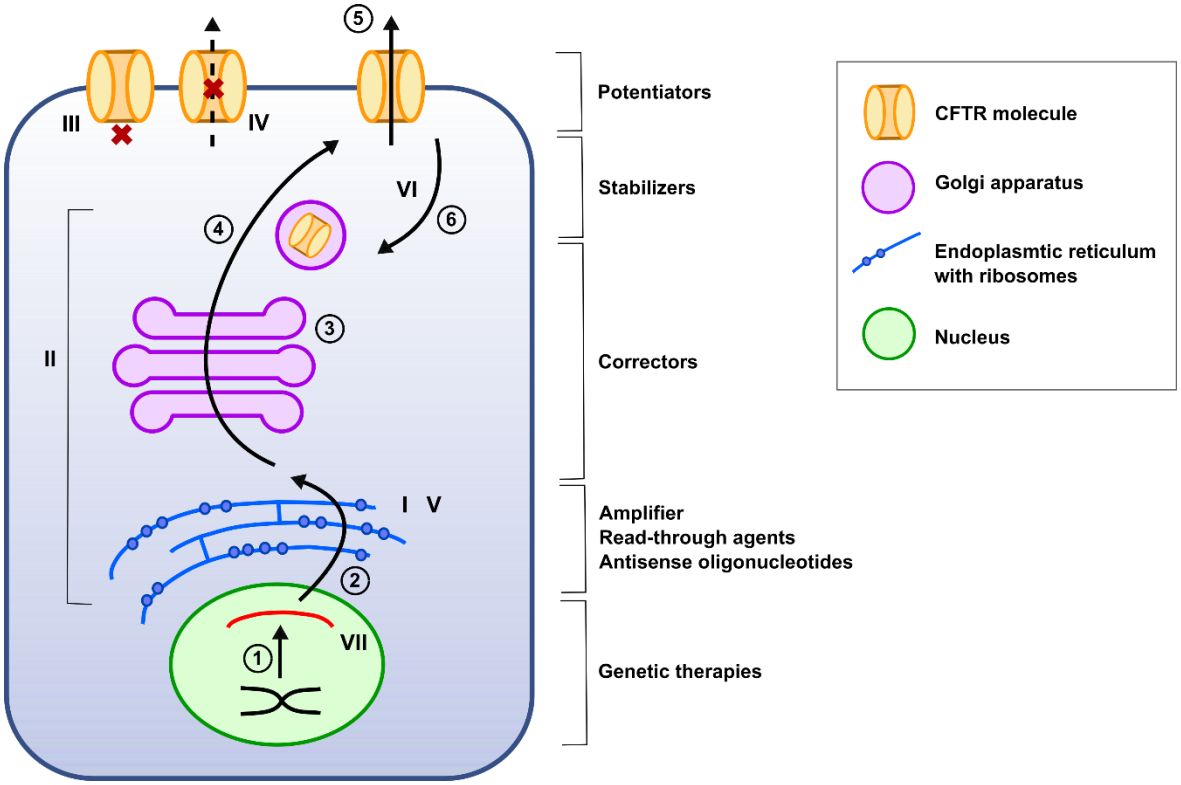


Figure 4. **CFTR mutation classes.** Arabic numbers indicate cellular mechanisms impaired by CFTR mutations. 1, transcription; 2, translation; 3, posttranslational modification; 4, protein trafficking; 5, expression of CFTR; 6, CFTR turnover. Roman numbers indicate the class of CFTR mutations. *Adapted from Cuyx and De Boeck, Semin. Respir. Crit. Care Med., 2019.*

Gene therapy for CF

Since the identification of the *CFTR* sequence, a large number of studies have been performed to develop a gene therapy approach for CF. The recent improvements in the delivery systems, both viral and non-viral, and the development of new technologies to modify our genetic code have highly accelerated the progress of gene correction for CF[50].

The delivery of the wild type *CFTR* cDNA to compensate the genetic defects has been extensively studied providing a mutation independent approach[6]. One of the principal issues of this strategy is the delivery efficiency of the *CFTR* gene into the

target cells. Among the non-viral vectors, cationic liposome demonstrated a safety profile, allowing monthly repeated administration[51]. Moreover, they do not present strict packaging capacity limits, enabling the incorporation of large transgenes. A Phase I/IIa gene therapy trial was performed using pGM169/GL67A gene-liposome formulation[52], demonstrating few side effects and an improvement in lung function compare to the placebo group with repeated doses of the formulation[53]. However, liposomal vectors present some limitations such as inefficient transduction and nuclear entry of the cargo[54].

On the opposite, viral vectors offer the possibility to efficiently enter the target cells and, eventually, the nucleus. The first studies using viral delivery of *CFTR* cDNA were done exploiting recombinant adenoviral vectors (rAdVs)[55,56]. Beside the transient expression of the transgene, these vectors produced an elevated immunogenicity, making them not suitable for repeated administrations[57]. At the same time, recombinant adeno-associated vectors (rAAVs) were tested for gene delivery into the airways, demonstrating less immunological host-response. Initial studies, exploiting the rAAV2 serotype, resulted in no significant clinical effects in patients due to the inefficient transduction of lung epithelia by this vector, thus no restored expression of CFTR[58]. Subsequently, several serotypes have been tested for their ability to transduce the airway epithelia, bringing to the identification of rAAV6 as the most promising for lung delivery[59,60]. However, the main disadvantage of this type of vectors for CF gene therapy is their limited packaging capacity (4.6kb), which set constrains regarding the choice of transgene promoter or the insertion of regulatory regions, having *CFTR* a 4.4kb cDNA sequence[61]. This problem was overcome by the use of a truncated *CFTR* (CFTRΔR) and a synthetic promoter, able to rescue the channel function in the upper airways of CF mice[62] and in the lung of *CFTR*-null pigs[58].

Even though only preclinical studies have been conducted, lentiviral vectors (LVs) represent an attractive delivery system for their ability to transduce both dividing and non-dividing cells, their packaging capacity and long term expression of the transgene achieved by integration in the host cells[63–65]. The integration profile was evaluated for self-inactivating (SIN) lentiviral vectors showing preferences of integration into intronic sequences, proving their safety[66]. Furthermore, LVs can be pseudotyped changing the envelope glycoproteins to achieve different tropisms for specific tissues or cells[67,68]. To date, lentiviral delivery to airways was tested in CF

mouse models, demonstrating stable transgene expression and recovery of disease phenotype for at least 12 months[69], and CF pigs, measuring an increased Cl^- current when stimulated with cAMP 2 weeks after treatment, suggesting the restoration of CFTR activity[70]. Further studies will strengthen the feasibility of LVs for CF gene therapy.

Although many clinical trials have been performed to evaluate the efficacy of *CFTR* cDNA delivery in the airways, the clinical benefits for patients were mainly limited to the stabilization of lung function, underling the need for further enhanced gene delivery efficiency[50,61,71,72].

Beside cDNA addition, gene editing strategies have been developed to correct the endogenous mutated gene, with a renewed interest thanks to the recent developments of precise and efficient CRISPR-nucleases. Programmable nucleases enable the cleavage of the double strand DNA of specific genomic sequences. The break is can be repaired through Non-Homologous End-Joining (NHEJ) pathway, introducing indels into the sequence, or by Homology Directed Repair (HDR) if a donor template is present. The feasibility of using HDR to correct mutations of endogenous genes has been extensively studied for many diseases. What makes this approach so attractive is the possibility to substitute the endogenous mutated sequence with the wild type one without scar formations, representing ideally the perfect cure for any disease. Positive results have been achieved using ZNF nucleases in patients derived induced pluripotent stem cells (iPSCs) for the HDR mediated correction of the ΔF508 deletion, achieving CFTR channel activity restoration in iPSC-derived epithelial monolayers[73]. A similar approach was tested in murine and human intestinal organoids homozygous for the ΔF508 mutation using CRISPR-Cas9 nuclease to achieve the specific cleavage that enhance the HDR. The isolation of corrected clones demonstrated normal CFTR function, measured by organoid swelling[74]. The disadvantage of these approches was the low HDR efficiency achieved in cell models related to the low frequency of such events[75]. Recently, a new strategy, based on HDR correction of the ΔF508 mutation was proposed, exploiting the delivery of Cas9 and the donor DNA through AAV6 vectors[76]. An efficient gene correction (30%–50%) was measured in upper-airway basal stem cells (UABCs) obtained from CF patients, resulting in restoration on CFTR activity. Furthermore, they focused on the development of a transplantation system to set up an ex vivo approach for CF treatment. The

possibility to transplant autologous CFTR-corrected cells will allow to overcome the barriers that render in vivo gene transfer inefficient, like the thick mucus, the inflammatory environment and also the possible immune reaction against Cas9[77,78]. Initial work demonstrated that edited airway basal cells can be embedded in a pSIS (porcine small intestinal submucosal) membrane, an FDA approved bioscaffold that is already in clinical use. UABCs embedded in this membrane retain the expression of phenotypic markers and differentiation capability, producing a differentiated epithelium with active CFTR channel. These results set the basis for in vivo studies for an ex vivo CF therapy[76].

However, due to the large number of defects that cause CF, mutations specific approaches can be challenging. An alternative approach is represented by super-exon integration, offering the possibility to treat the vast majority of patients with a unique approach. Super-exon integration consisting in the precise insertion of a *CFTR* cDNA fragment, encompassing exons 11-27, into exon 11 of the endogenous *CFTR* gene mediated by ZFN cleavage. This produces the expression of correct transcripts that produce active CFTR chloride channels in CFBE41o- cell model[79].

OTHER THERAPIES

An alternative approach for CF targets another ion channel, the epithelial sodium channel (ENaC), reported to be involved in CF lung pathology[80]. The inhibition of this channel could correct hydration of the airway surface liquid, achievable through small molecules or RNA interference (RNAi) and antisense oligonucleotide (ASOs). Whereas a number of small molecules have been studied for this purpose, a clinical therapy has not been proposed yet[81]. Nevertheless, downregulation of the expression of ENaC channel by RNAi approach was achieved by using a siRNA targeting the αENaC chain. The level of silencing was sufficient to restore the mucociliary defect in ALI cultures, moreover the nanoparticle-mediated delivery results in and efficient silencing in vivo[82]. Alternatively, ASOs have been developed to reduce ENaC mRNA through RNase H1-dependent degradation. The positive effects measured in murine CF-like models brought ENaC ASO to phase I clinical trial[71].

THE CRISPR/CAS SYSTEM

The CRISPR/Cas (clustered regularly interspaced short palindromic repeats/ CRISPR-associated proteins) systems has been identified as the adaptive immunity that bacteria and archaea have evolved as a protection against viral infections or exogenous DNA[83,84]. The genomes of bacteria and archaea present a characteristic locus that was identified as the CRISPR locus, usually located in proximity of *cas* genes[85]. These loci are characterized by a pattern of repeated sequences of DNA alternated by spacer sequences[83]. The observation that spacers were similar to sequences of bacteriophage genomes to which the bacterium was resistant[86,87] and that *cas* genes encode for nuclease and helicase proteins[85] allowed the discovery of CRISPR/Cas function.

Three stages characterize the mechanism of action of CRISPR/Cas systems in bacteria: acquisition, expression and interference[88]. A new invading DNA is first recognized in a PAM (proto-spacer-adjacent motif) dependent manner[89], and subsequently fragmented in into spacer precursor of a defined size. Cas1 and Cas2 enzymes catalyse the insertion of the new spacer into the CRISPR locus with a specific orientation determined by PAM location, providing a genetic record of the infection. Spacer integration enable the host to prevent subsequent invasion by the same DNA. During the expression stage, a long precursor transcript (pre-crRNA) is generated from the CRISPR array, that is subsequently processed with different mechanisms depending on the type of CRISPR system[84,90]. The Cas proteins and the mature crRNA assemble forming the crRNP complex[91]. In the interference step, the second time invading nucleic acid is cleaved by the Cas nuclease. The recognition of the target sequence, complementary to the crRNA, rely on the identification of the PAM region, allowing the discrimination between self and non-self DNA. This is followed by base pairing of the seed region (7-8 nucleotides of the spacer) and consequential extension, to eventually obtain complete hybridization between the spacer and the protospacer and strand displacement[92]. The base pairing of the crRNA and the target sequence activate the nuclease activity of Cas nuclease, producing the cleavage of the foreign DNA[84].

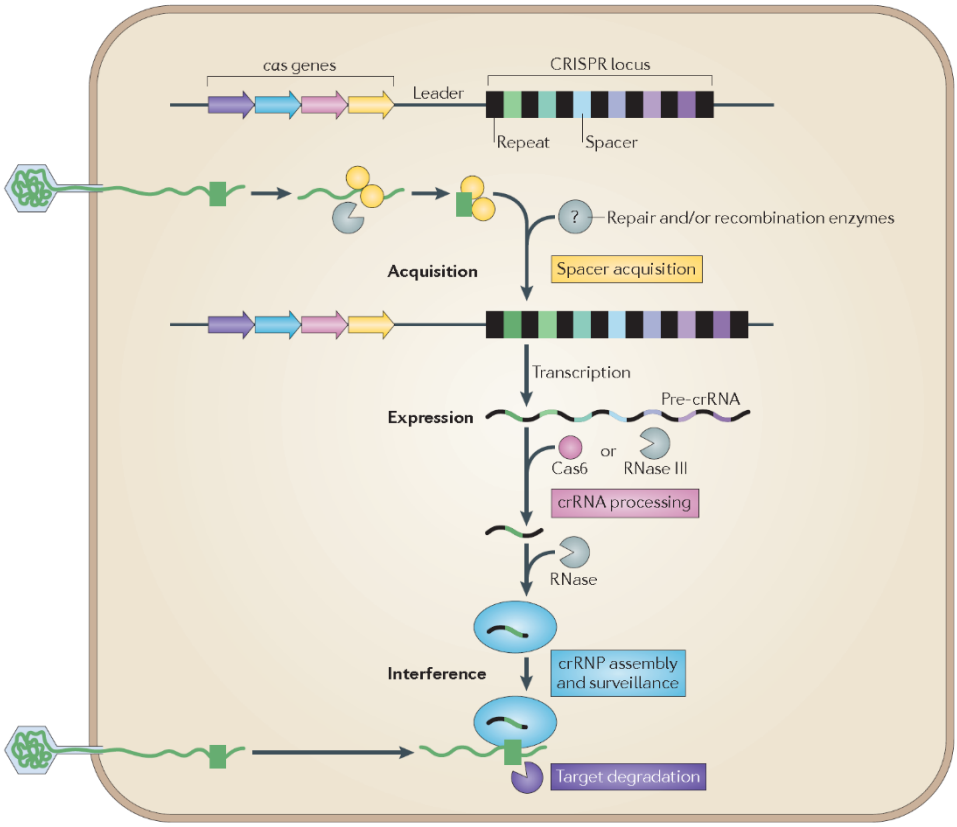


***Figure 5*. Mechanism of action of CRISPR/Cas system as bacteria adaptive immunity**. Integration of the invading DNA into the CRISPR locus is mediated by cas genes. These DNA fragments, or spacers, are transcribed and used by Cas proteins in the interference step. *Adapted from van der Oost et al., Nat. Rev. Micro., 2014.*

Classification of CRISPR/Cas systems

A large number of different CRISPR/Cas loci have been identified and divided into two classes depending on the number of Cas proteins involved in the degradation of the invading nucleic acids[88]. Class 1 have multiple Cas effector modules, that together with the crRNA bind and process the target, and include types I, III and IV. Class 2, instead, have a single large multidomain Cas protein binding the crRNA, that substitute the effector complexes of class 1. This class include types II, V and VI. One of the main features that divide CRISPR system into different types is the presence of characteristic Cas protein. Type I systems encode for Cas3, type II is depicted by Cas9 protein, type III contains Cas10, type V systems are associated with Cas12a and type VI with C2c2[90]. Type IV systems are highly variable and lack the nuclease module require for interference, suggesting a different mechanism of action[93]. These six types of CRISPR systems further divide in many different subtypes. Each system is also characterized by its particular crRNA.

CRISPR type II: Cas9 endonuclease

Type II CRISPR systems were the first to be characterized and are among the most studied[94]. These systems are composed by a single DNA endonuclease, Cas9, whose role is the recognition and cleavage of double strand DNA sequences. Each strand is cleaved by a different Cas9 domain: the HNH and the RuvC domains, responsible of the cut of the complementary and the non-complementary sequence in the target DNA, respectively[95,96]. Moreover, they present an additional small non coding trans-activating RNA (tracrRNA) necessary for crRNA maturation, that by 24 nucleotides-long base pair interaction with the crRNA forms the RNA complex associated with Cas9 nuclease[97]. The dual-RNA/Cas9 system enable the cleavage of any DNA sequence, provided that a complementary protospacer sequence (20 nt) flanked by a PAM is present into the target DNA. The cleavage happens 3 nucleotides upstream the PAM producing blunt ends fragments that are quickly degraded[98].

The type II CRISPR/Cas9 system from *Streptococcus pyogenes* (SpCas9) has been extensively studied and adapted to allow targeted genome editing in eukaryotic cells. SpCas9 recognize either 5'-NGG or 5'-NAG PAM and to increase the number of possible targets, several protein variants recognizing different PAM sequences have been developed, overcoming the limit of specific PAM requirement[99]. Furthermore, to simplify the experimental design for biotechnological applications, a simpler system has been created by fusing the crRNA and the tracrRNA into a single RNA transcript, called single guide RNA (sgRNA). This two-component system retains the entire Cas9 function with the advantage that, by changing the spacer sequence of the sgRNA, every DNA sequence can be possibly targeted[98].

One major concern related to type II systems, and also to CRISPR nucleases in general, is the off-target activity of Cas9. Due to its ability to tolerate mismatches between sgRNA and the protospacer sequence, SpCas9 can produce unwanted modifications in the genome causing unpredictable effects on cellular metabolism and viability. To overcome this problem, several high fidelity variants have been identified[100–103], together with strategies to limit the expression of the nuclease into the cells, associated with a reduce chance of off-target cleavages[104,105].

CRISPR type V: Cas12a endonuclease

CRISPR type V systems are characterized by Cas12a nuclease. While Cas12a belongs to the same class of Cas9, these two enzymes present different structural architecture and thus distinct molecular mechanism[106]. Cas12a is associated with a single crRNA 42 nucleotides long, which presents a 3' complementary sequence of about 23-25 nucleotides with the protospacer of the target DNA[107,108]. The maturation of the crRNA does not require a tracrRNA. Moreover, these types of nucleases naturally recognize a T-rich PAM sequence (TTTN) located at the 5' of the protospacer, engineered version with altered specificities have been also established[109]. In contrast to type II, Cas12a contain a single nuclease domain, RuvC-like, whose cleavage generates staggered DNA ends with 4 or 5 nucleotides 5' overhang in the distal part of the protospacer[110]. Several Cas12a orthologs have been identify as bacteria adaptive immune systems, among them, *Acidaminococcus sp. BV3L6* Cas12a (AsCas12a) exhibit robust genome editing in mammalian

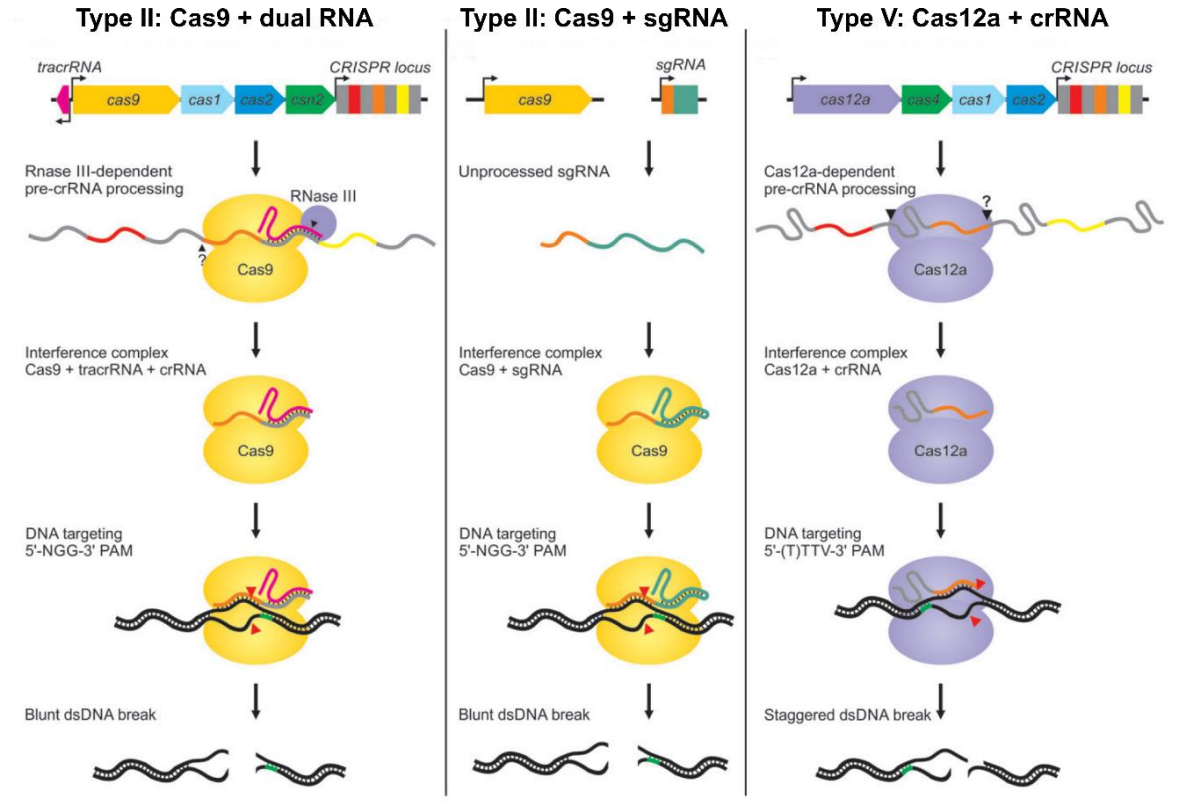


Figure 6. **CRISPR type II and V.** In the dual RNA system, the tracrRNA (magenta) bind the repeats sequences (gray) of the pre-crRNA transcript. This structure is recognized by Cas9 and cleaved by RNase III. The recognition of a PAM sequence (green) in the target DNA produce a blunt end cleavage. The sgRNA allows the expression of a single transcripts that fuse together the tracrRNA and crRNA. In type V, Cas12a recognize the repeated sequence of the pre-crRNA (gray) and bind to it, catalysing also its processing. Upon recognition of the PAM (green), Cas12a cut the target DNA generating a staggered break. *Adapted from Swarts and Jinek, Wiley Interdiscip Rev RNA, 2018.*

cells[107]. In addition, it has been demonstrated that AsCas12a is intrinsically more specific, compare to SpCas9, demonstrating less off target cleavage[111].

APPLICATIONS OF CRISPR NUCLEASES

The extreme low cost and simplicity of experimental design made CRISPR programmable nucleases very attractive for the biotechnological field, with the fast development of several genome engineering approaches.

The value of CRISPR/Cas arises from its ability to induce site-specific DNA cleavage into the genome. The double strand breaks (DSBs) created are then repaired by different cellular mechanisms that produce modifications of the target sequence. DSBs are preferentially repaired through Non Homologous End-Joining (NHEJ) pathway that produce an error-prone repair by inserting or deleting (indels) small portion of the DNA sequence[112,113]. Indels formations in gene coding sequences usually cause frameshifts that can be exploited to achieve gene knockout. This approach is very useful to study the function of genes into tissues or cells.

Alternatively, in the presence of a donor DNA with flanking homologous regions to the target site, the Homology Directed Repair (HDR) pathway can be activated[75,114].

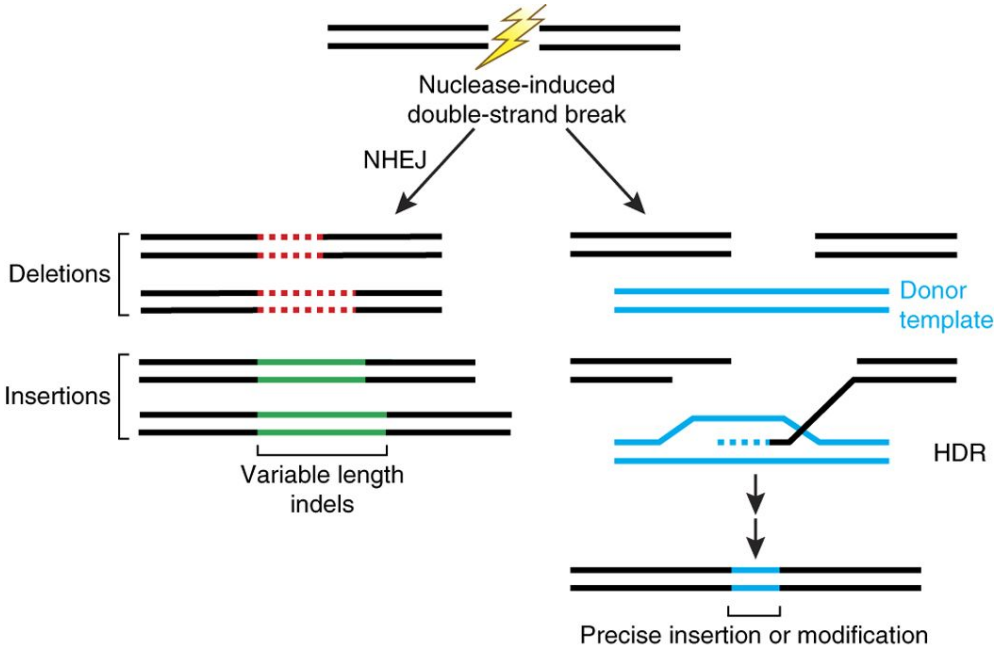


Figure 7. **Repair pathways after nuclease double strand break.** DSBs can be repaired through NHEJ, producing the insertion of deletion of variable length, or by HDR, in the presence of a donor template that is used to modify the sequence. *Adapted from Sander and Joung, Nat Biotech, 2014.*

Even though this process is highly inefficient in mammalian cells and tissues and limited to the late S and G2 phases of cell cycle, it has been shown that DSBs enhance the efficiency of HDR by at least 2-3 order of magnitude[115,116]. This pathway is extensively studied because it could allow the modification of any sequence into the genome, thus understanding the exact mechanism and the players involved may further improve HDR efficiency[117]. Moreover, it can be used to integrate sequences in a site-specific manner relying on the presence of homology arm flanking the sequence, as opposed to viral delivery that cause random integration into the genome[117].

Several biotechnological tools have been developed for the study of cell biology. For example, a dead version of the Cas9 nuclease, dCas9, was created mutating the two nuclease domains (D10A and H840A mutations)[98]. This inactive nuclease can be considered as an RNA-guided DNA-binding protein providing a platform for transcriptional or epigenetic modifications of the genome without the creation of DSBs and consequent indels formation in the DNA sequence[118]. To achieve transcriptional modulation effects, dCas9 was fused to transcriptional effector proteins that acts as repressors (CRISPRi)[119], such as Kruppel-associated box (KRAB) domain, or activators (CRISPRa)[120], like VP16 domains. Moreover, different epigenetic-modifying enzymes were combined to dCas9 to produce epigenetic modification of targeted genomic loci. Some examples are represented by the fusion of dCas9 with the core catalytic domain of the human acetyltransferase p300, to upregulate gene expression following acetylation of histone H3 Lys27[121], or with the lysine demethylase 1 (LSD1), that instead reduce the acetylation of the same histone[122].

Another powerful use of this technology is the study of genomic functions by high throughput screening using library of hundreds of thousands of sgRNAs targeting the desire set of genomic sequences[123]. These sgRNAs libraries can be associated with an active Cas9, to knockout genes[124], or to dCas9 fusion proteins, to regulate gene transcription[125,126].

Given its ability to target specific sequences, Cas9 has been used as imaging tool to visualize genomic organisation and dynamics in cells[127–129], as a tool for cell lineage tracing[130] and for the generation of animal models[118].

One of the last CRISPR technologies developed are DNA base editors. These complexes are constituted by a catalytically impaired Cas nuclease fused to a base

modification enzyme capable to deaminate bases on a single-stranded DNA, like cytosine or adenine deaminases[131]. This results in the modification of specific nucleotides in the target sequence. In the case of type II CRISPR systems, the efficiency of base editor is improved by the use of a Cas9 nickase, mutated only in one of the two nuclease domains (D10A mutation). The creation of a nick in the non-edited strand induces the cell repair machinery to use the edited strand as a template to repair the nick allowing the incorporation of the new mutation. These tools allow base conversion without the creation of DSBs or donor DNA template and in a more efficient manner compared to the HDR[132].

Recently a new genome editing strategy have been reported and defined as prime editing[133]. This technology involves a Cas9 nickase fused to an engineered reverse transcriptase that together with a particular sgRNA, prime editing guide RNA (pegRNA), writes new genomic information into the target sequence. The pegRNA determines both the target sequence and the desired editing event. The prime editing has been demonstrated very efficient and precise, considerably expanding the targets for genome editing[133].

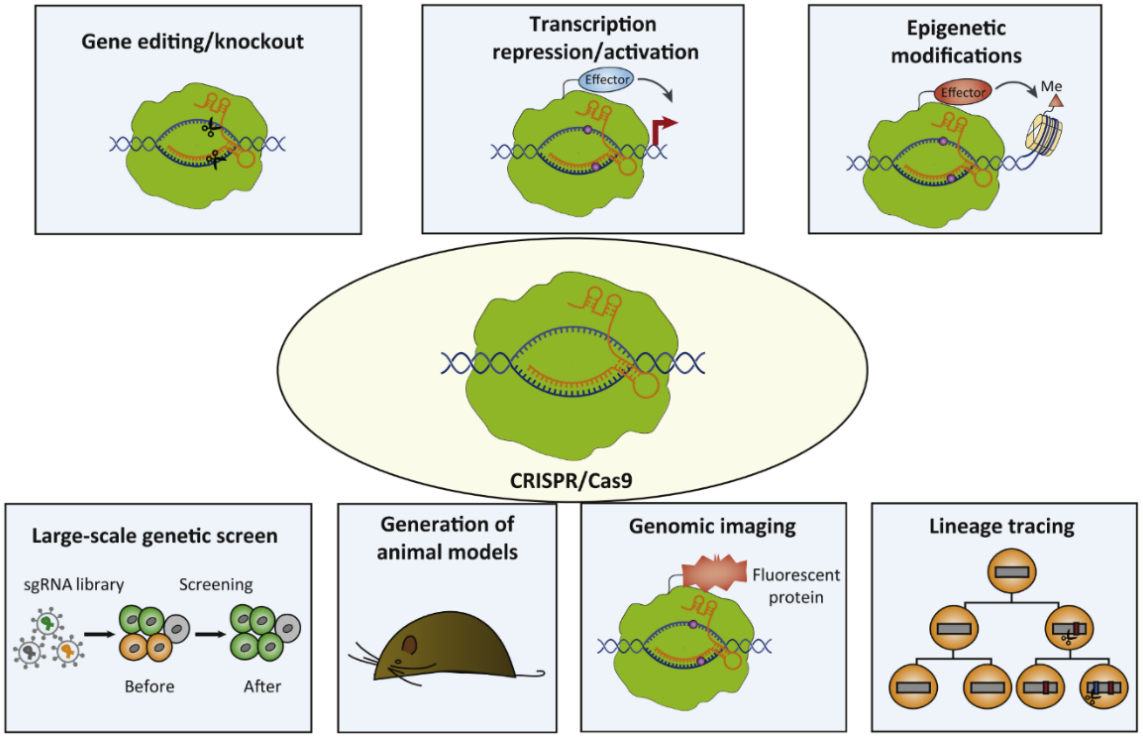


Figure 8. **Applications of CRISPR-Cas nucleases.** Several biotechnological applications have been developed. Among them there are gene editing, transcriptional regulation, epigenetic modifications, genetic screening, generation of animal models and genomic imaging. *Adapted from Wang and Qi, Trends Cell Biol., 2016.*

GENE THERAPY WITH CRISPR

The possibility offered by CRISPR nucleases to generate targeted deletions, insertions and modifications at a single base level has revolutionized the gene therapy approaches for genetic diseases. The results achieved in the improvements of CRISPR-Cas efficiency and its delivery either *in vivo* or *ex vivo* have given rise to a number of therapeutic applications and their clinical translation[134,135].

The first-in-human CRISPR trial was proposed in 2016 for the treatment of melanoma, myeloma and sarcoma by *ex vivo* engineering of autologous T cells in combination with chemotherapy agents. Patient's T cells were edited in order to express NY-ESO-1 TCR and eliminate endogenous PD-1 and TRC expression[136].

Another example of *ex vivo* CRISPR therapy is represented by the editing of hematopoietic stem and progenitor cells (HSPCs). These cells can be modified *ex vivo* by electroporation of CRISPR nucleases and guide RNAs, representing a possible treatment for haematological diseases. Currently a phaseI/II clinical trial for sickle cell disease and β-thalassemia is ongoing exploiting the autologous transplantation of HSPCs after Cas9-sgRNA disruption, mainly through NHEJ, of the BCL11A expression[137]. This allows the production of fetal γ-globin and amelioration of the disease phenotype[138]. The possibility of an *ex vivo* therapy has been proposed also for the treatment of Duchenne Muscular Dystrophy (DMD)[139]. Several mutations-specific approaches have been studied, such as knock in of a deleted exon, or deletion of an exon to restore the frame of the protein[139,140]. It has been demonstrated that the editing of induced pluripotent stem cells (iPSCs) can restore the correct expression of the dystrophin complex in several differentiated cells type, posing the basis for an *ex vivo* DMD therapy[139].

The possibility to modify cells in vitro and transplant them back to the patients to achieve a therapeutic effect represents an advantageous strategy for CRISPR therapy. While it requires more passages (collection of cells, isolations, expansions, modification and selection of the corrected clones, transplantation), it increases the safety of CRISPR approaches, mainly in terms of off-targets activity and, in addition, prevents the challenges given from the *in vivo* delivery[135].

Nevertheless, *in vivo* gene therapy approaches have been developed exploiting the use of either viral or non-viral vectors. A first example is represented by the delivery of Cas9 through AAV9 vector for the treatment of DMD. Two different approaches have been proposed based on HDR, in the presence a donor template[141], or

exploiting the more efficient NHEJ repair, using two sgRNAs to perform a deletion of a portion of the coding sequence in order to restore the correct frame[142,143], to restore dystrophin expression in DMD models. Furthermore, the delivery of Cas9-sgRNA through AAV2/8Cas9 followed by NHEJ repair was efficiently used to disrupt the expression of the *Pcsk9* gene in human hepatocytes and in mouse liver for the treatment of hypercholesterolemia[144,145].

A Phase I/II clinical trial started in 2019 for Leber congenital amaurosis type 10 (LCA10)[146], a retinal degenerative disease caused by mutations in *CEP290* gene. To correct the IVS26 mutation, creating an aberrant splice site into intron 26 of *CEP290* gene, CRISPR-Cas9 with two sgRNA was delivered through AAV5 vector to delete the intronic region containing the mutation. Subretinal injection in non-human primate (NHP) showed an efficient deletion of the aberrant splice site and restoration of therapeutic level of CEP290 protein[147].

Another ongoing *in vivo* phase I clinical trial proposes the elimination of E6 and E7 oncogenes from the human papillomavirus (HPV), responsible for cervical cancer[148,149].

The deep knowledge and the continuous and fast improvements of CRISPR systems allowed the development of all these therapeutic approaches and the beginning of a number of clinical trials[134]. These results suggest that CRISPR technology could be used to treat a broad range of diseases in the future[135].

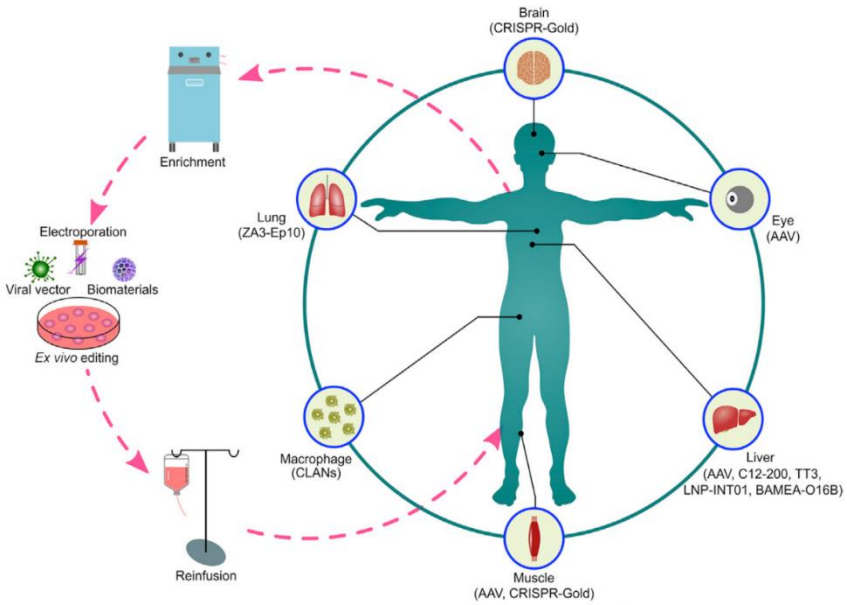


Figure 9. **CRISPR ex vivo and in vivo therapy**. *Adapted from Li et al., Trends in Pharmacological Sciences, 2019.*

SPLICING

In eukaryotes, the coding sequences of genes, the exons, are interrupted by other sequences, the introns, that do not contain protein coding information. Thus, introns must be removed during the transcription of the precursor messenger RNA (pre-mRNA) in order to give rise to the mature messenger RNA (mRNA) that contains the information necessary for the protein production. This is a fundamental process for cells viability and is defined as splicing[150]. Mutations in splicing regulatory elements cause a misregulation of the splicing reaction leading to disease[151].

Splicing regulatory sequences

The boundary between exons and introns relies on important short conserved sequences, the 5' splice site, the 3' splice site, the branch point and the poly-pyrimidine tract[152].

The 5' splice site, or donor splice site, presents a highly conserved dinucleotides GU sequence at the beginning of the intron, whereas the surrounding regions are less conserved (spanning from the 3 last nucleotides of the exon and the 6 first of the intron)[153]. The 3' splice site, or acceptor splice site, is less conserved and is defined by an AG dinucleotide that ends the intronic sequence. Nevertheless, another class of introns has been identified, having AT and AC as consensus sequences at the 5' and 3' splice site respectively[154]. The branch point and the poly-pyrimidine tract are also needed to define the intronic region. The branch point exhibits a conserved adenosine (A) nucleotide surrounded by highly degenerated sequences and is located 20-50 nucleotide before the beginning of the exon. Between the branch point and the 3' splice site there is a stretch of pyrimidine residues (8 bases on average) that defines the poly-pyrimidine tract[155,156].

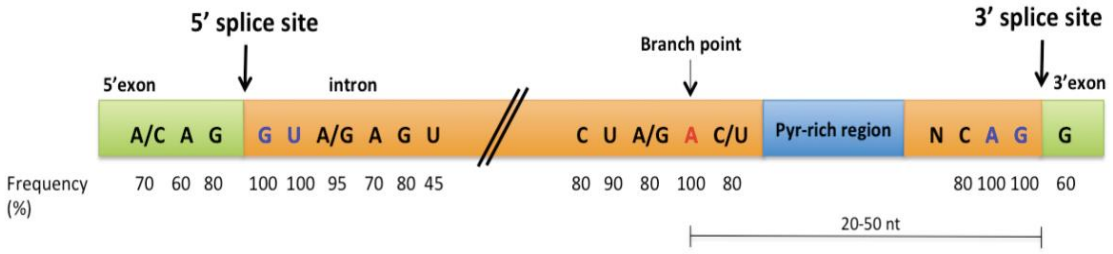


Figure 10. **Conserved splicing regulatory sequences**.

All these sequences are recognized by the splicing machinery, the spliceosome, with high fidelity among the multitude of potential, but incorrect, sites. Thus, the precision of the splicing reaction depends also on auxiliary intronic and exonic sequences that act as *cis*-elements that recruit *trans*-acting splicing factors able to activate or repress the use of splicing regulatory sequences in the proximity[157,158]. These elements present highly degenerated sequences making their identification difficult. Moreover, they are named Exonic Splicing Enhancer (ESE) or Silencer (ESS) and Intronic Splicing Enhancer (ISE) or Silencer (ISS), depending on their position and function[159].

Exonic splicing enhancers are short sequences usually recognized by Serine/Arginine-rich (SR) proteins that help the definition of the exon sequence and promote its inclusion into the mRNA[160]. On the opposite, exonic splicing silencers mainly recruit inhibitory factors belonging to the hnRNP protein family that induce the spliceosome to ignore pseudo-exons or cryptic splice sites promoting exon skipping[161]. Intronic regulatory sequences are generally less characterized. Intronic splicing enhancers are also involved in exon recognition and probably in alternative splicing mechanisms[162]. In contrast, it has been shown that intronic splicing silencers are recognized by poly-pyrimidine tract binding protein (PTB) and hnRNP A1 that act as splicing repressor, impeding the recognition of 3' splice sites[151,163].

These splicing regulatory elements create a fine regulation of the splicing reaction, influencing the spliceosome assembly on the pre-mRNA, the identification of exon/intron boundary and maturation of the transcripts, therefore, playing a fundamental role in the regulation of gene expression[164–166].

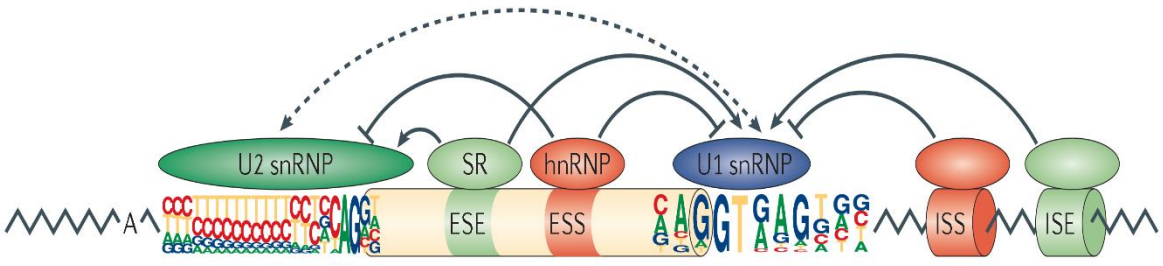


Figure 11. **cis-acting and trans-acting factors that regulate cellular splicing.** *Adapted from Matera and Wang, Nat. Rev. Mol. Cell Biol., 2014.*

Splicing mechanism

The intron removal from the pre-mRNA is catalysed by a ribonucleoprotein complex (RNP), the spliceosome, involving a two-step transesterification reactions[167]. The spliceosome is a complex machinery composed by 5 small nuclear RNAs (snRNAs) and more than 100 proteins. Each snRNA (U1, U2, U4, U5 and U6) interacts with a set of proteins, forming small nuclear ribonucleoprotein complexes (snRNPs)[150]. During the first step, the U1 snRNP recognizes the 5' splice site, thanks to base pair interactions, and the U2 snRNP the BP, positioning them in close proximity. Subsequently, the preassembled U4/U6-U5 snRNPs complex associate to form the complete spliceosome. Several rearrangements are then needed to form the catalytically active spliceosome, including the substitution of U1 snRNP by the U6, that form the active site of the spliceosome, and the release of U4 complex. Thus, the mature spliceosome is catalytically active and ready to perform the two transesterification reactions. In the first one, the 2' hydroxyl group of the adenosine at the BP operates a nucleophilic attack to the phosphodiester group of the first intronic residue at the 5' splice site, generating a lariat intron intermediate. In the second reaction, the 3' hydroxyl group of the 5' splice site reacts with the phosphodiester group at the 3' splice site, joining the two exons and releasing the lariat intron[150,167,168].

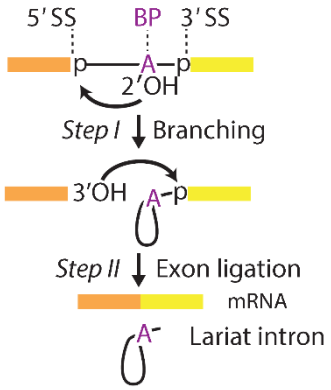


Figure 12. **Splicing reaction**. *Adapted from Wilkinson et al., Annu. Rev. Biochem., 2019.*

AIM OF THE THESIS

Cystic fibrosis (CF) is an autosomal recessive disease caused by mutations in the *CFTR* gene, representing one of the most lethal disorders. More than 300 mutations have been identified in the *CFTR* sequence as CF-causing, giving rise to different phenotypic clinical onset. 3272-26A>G and 3849+10kbC>T are intronic defects that cause the alteration of the splicing pattern and the production of a non-functional CFTR channel.

Despite the development of CFTR modulator drugs represented a great improvement for patients, the need of life-long drug uptake, side effects and elevated cost pose the need of the development of new strategies that correct the genetic defects.

The aim of the thesis was to assess the possibility to restore the correct splicing pattern of the 3272-26A>G and 3849+10kbC>T defects by modifying the intronic sequences nearby the mutations, using the programmable nucleases SpCas9 and AsCas12a. Indeed, the discovery of CRISPR nucleases transformed the field of gene therapy allowing to easily edit specific sequences in a precise and efficient manner. This approach aims to provide a permanent correction of the endogenous *CFTR* gene without the alteration of the coding sequence.

Moreover, this approach can potentially be expanded to other CF-causing splicing mutations or other diseases caused by the alteration of the splicing pattern.

METHODS

Plasmids. Wild type and mutated minigenes for 3272-26A>G and 3849+10kbC>T mutations were cloned into previously published pcDNA3 and pCI plasmid respectively[169,170]. Wild type minigenes (pMG3272-26WT and pMG3849+10kbWT) were obtained by PCR amplification and cloning of target regions from *CFTR* gene of HEK293T cell. Primers used are listed in Appendix 1. The pMG3272-26WT plasmid includes the last 23 bases of exon 18, full length exons 19 and 20 and intron 19. The addition of the last part of exon 18 was performed by semi-nested PCR: the first PCR was done with oligo TM exon 18 exon 19 hCFTR fw and oligo exon 20 hCFTR rev; the second PCR was done with the forward primer oligo KpnI-AgeI exon 18-19 hCFTR fw and the same reverse oligo. The pMG3849+10kbWT plasmid contains full length exons 22 and 23, and portions of intron 22 as previously described[29]. Mutated minigenes (pMG3272-26A>G and pMG3849+10kbC>T) were obtained by site directed mutagenesis of pMG3272-26WT and pMG3849+10kbWT, for both mutations. Guide RNAs were cloned into pY108 lentiAsCas12a (Addgene Plasmid 84739) or lentiCRISPR v1 (Addgene Plasmid 49535) using BsmBI restriction sites as previously described[171]. These plasmids allow simultaneous delivery of the RNA-guided nuclease and the sgRNA to target cells and contain puromycin as selection marker. Guide RNA sequence is listed in Appendix 2.

Cell lines. Human colorectal adenocarcinoma cells (Caco-2), HEK293T cells and HEK293 cells stably expressing pMG3272-26WT (HEK293/pMG3272-26WT) or 3272-26A>G cells (HEK293/pMG3272-26A>G) were cultured in Dulbecco's modified Eagle's medium (DMEM; Life Technologies) supplemented with 10% fetal bovine serum (FBS; Life Technologies), 10 U/ml antibiotics (PenStrep, Life Technologies) and 2 mM L-glutamine at 37°C in a 5% CO_2 humidified atmosphere. Puromycin selection was performed using 10 µg/ml for Caco-2 cells, and 2 µg/ml for HEK293T or HEK293 cells. HEK293T, HEK293 and Caco-2 cells were obtained from American Type Culture Collection (ATCC; www.atcc.org). Stable minigene cell lines (HEK293/3272-26A>G and HEK293/3272-26WT) were produced by transfection of Bgl-II linearized minigene plasmids (pMG3272-26WT or pMG3272-26A>G) in HEK293 cells. Cells were selected with 500 µg/ml of G418, 48h after

transfection. Single cell clones were isolated and characterized for the expression of the minigene construct.

Transfection and lentiviral transduction of cell lines. Transfection experiments were performed in HEK293T cells seeded (150000 cells/well) in a 24 well plate and transfected using PEI (polyethylenimine) with 100 ng of minigene plasmids and 700 ng of plasmid encoding for nuclease and sgRNA/crRNA (pY108 lentiAsCas12a or lentiCRISPR v1). After 16h incubation cell medium was changed, and samples were collected at three days from transfection.
Lentiviral particles were produced in HEK293T cells at 80% confluency in 10 cm plates. 10 µg of transfer vector (pY108 lentiAsCas12a or lentiCRISPR v1) plasmid, 3.5 µg of VSV-G and 6.5 µg of Δ8.91 packaging plasmid were transfected using PEI. After over-night incubation the medium was replaced with complete DMEM. The viral supernatant was collected after 48h and filtered in 0.45 µm PES filter. Lentiviral particles were concentrated and purified with 20% sucrose cushion by ultracentrifugation for 2 hours at 4°C and 150000xg. Pellets were resuspended in appropriate volume of OptiMEM. Aliquots were stored at -80°C. Vector titres were measured as Reverse Transcriptase Units (RTU) by SG-PERT method[172].
For transduction experiments HEK293/pMG3272-26WT or HEK293/pMG3272-26A>G and Caco-2 cells were seeded (300000 cells/well) in a 12 well plate and the day after were transduced with 3 RTU of lentiviral vectors. 48 hours later, cells were selected with puromycin (2 µg/ml for HEK293 or 10 µg/ml for Caco-2 cells) and collected 10 days from transduction.

Transcripts analysis. RNA was extracted using TRIzol™ Reagent (Invitrogen) and resuspended in DEPC-ddH2O. cDNA was obtained starting from 500 ng of RNA using RevertAid Reverse Transcriptase (Thermo Scientific), according to manufacturer's protocol. Target regions were amplified by PCR with Phusion High Fidelity DNA Polymerase (Thermo Fisher). Oligonucleotides are listed in Appendix 1. Uncropped and unprocessed scans are available in the Source Data files.

Detection of nuclease induced genomic mutations. Genomic DNA was extracted using QuickExtract DNA extraction solution (Epicentre) and the target locus amplified by PCR using Phusion High Fidelity DNA Polymerase (Thermo

Fisher). Oligos are listed in Appendix 1. To evaluate indels resulting from cleavage of one gRNA, purified PCR products were sequenced and analyzed using the TIDE or the SYNTHEGO ICE software[173,174]. In some experiments DNA editing was measured also by T7 Endonuclease 1 (T7E1) assay (New England BioLabs) following manufacturer's instructions and as previously described[104].

Primary airway epithelial cell culture and transduction. Primary airway (bronchial) epithelial cells derived from CF patients compound heterozygous for 3272-26A>G splicing mutation (3272-26A>G/ΔF508, n=1) and for 3849+10Kb C>T splicing mutation (3849+10Kb C>T/ΔF508, n=1) (kindly provided by the Primary Cell Culture Service of the Italian Cystic Fibrosis Research Foundation) were cultured in LHC9/RPMI 1640 (1:1) without serum[175]. The day before the transduction, 50000 cells, at passage 3, were seeded into a 24 well plate previously treated with collagen. Transduction was performed with 2 RTU of lentiviral vectors. Puromycin selection, where indicated, was performed with 2 µg/ml, 48h post transduction for 72h. Cells were collected for analysis after 15 days.

Human intestinal organoids culture and transduction. Human intestinal organoids of CF subjects compound heterozygous for 3272-26A>G splicing mutation (3272-26A>G/4218insT, n=1, CF-86) and for 3849+10Kb C>T mutation (3849+10Kb C>T/□F508, n=1, CF-110) were derived from fresh rectum suction biopsies and cultured as previously described[176]. The Ethics Committee of the University Hospital Leuven approved this study and informed consent was obtained from all participating CF subjects. Organoids cultures, at passages 10-15, were trypsinized to single cell using trypsin 0.25% EDTA (Gibco), 30000-40000 single cells were resuspended with 25 µl of lentiviral vector (0.25-1 RTU) and incubated for 10 min at 37°C[177]. The same amount of Matrigel (Corning) was added and the mix plated in a 96-well plate. After polymerisation of the Matrigel drops (37°C for 7 minutes), they were covered with 100 µl of complete organoid medium[176] containing 10 µM of Rock inhibitor (Y-27632 2HCl, Sigma Aldrich, ref. Y0503) for the first three days to ensure optimal outgrowth of single stem cells[178]. Medium was replaced every 2-3 days until the day of organoid analysis.

Analysis of CFTR activity in intestinal organoids. Fourteen days after viral vector transduction, organoids were incubated for 30 minutes with 0.5 µM calcein-green (Invitrogen, ref. C3-100MP) and analyzed by confocal live cell microscopy with a 5X objective (LSM800, Zeiss, with Zen Blue software, version 2.3). Steady-state organoids area was determined by calculating the absolute area (xy plane, μm^2) of each organoid using ImageJ software through the Analyse Particle algorithm. Defective particles with an area less than 1500 or 3000 µm for 3272-26A>G or 3849+10Kb C>T, respectively, were excluded from the analysis. Data were averaged for each different experiment and plotted in a box plot representing means ± SD.

The Forskolin Induced Swelling (FIS) assay was performed by stimulating organoids with 5 µM of forskolin and analyzed by confocal live cell microscopy at 37 °C for 60 min (one image every 10 min). The organoid area (xy plane) at different time points was calculated using ImageJ, as described above.

GUIDE-seq. GUIDE-seq experiments were performed as previously described[179]. Briefly, $2x10^5$ HEK293T cells were transfected using Lipofectamine 3000 transfection reagent (Invitrogen) with 1 µg of lenti Cas12a plasmid (pY108) and 10 pmol of dsODNs[179]. The day after transfection cells were detached and selected with 2µg/ml puromycin. Four days after transfection cells were collected and genomic DNA extracted using DNeasy Blood and Tissue kit (Qiagen) following manufacturer's instructions. Using Bioruptor Pico sonicatin device (Diagenode) genomic DNA was sheared to an average lenght of 500 bp. Library preparations were performed with the original adapters and primers according to previous work. Libraries were quantified with the Qubit dsDNA High Sensitivity Assay kit (Invitrogen) and sequenced with the MiSeq sequencing system (Illumina) using an Illumina Miseq Reagent kit V2 - 300 cycles (2 × 150 bp paired-end). Raw sequencing data (FASTQ files) were analyzed using the GUIDE-seq computational pipeline[179]. After demultiplexing, putative PCR duplicates were consolidated into single reads. Consolidated reads were mapped to the human reference genome GrCh37 using BWA-MEM; reads with mapping quality lower than 50 were filtered out. Upon the identification of the genomic regions integrating double-stranded oligodeoxynucleotide (dsODNs) in aligned data, off-target sites were retained if at most seven mismatches against the target were present and if absent in the

background controls. Visualization of aligned off-target sites is available as a color-coded sequence grid [100,105].

Targeted deep sequencing. The loci of interest were amplified using Phusion high-fidelity polymerase (Thermo Scientific) from genomic DNA extracted from human intestinal organoids (3272-26A>G/4218insT) 14 days after transduction with lentiAsCas12a-crRNA +11, +14 or Ctr, either in organoids (on-target) or airway cells (on and off-target). Amplicons were indexed by PCR using Nextera indexes (Illumina), quantified with the Qubit dsDNA High Sensitivity Assay kit (Invitrogen), pooled in near-equimolar concentrations and sequenced on an Illumina Miseq system using an Illumina Miseq Reagent kit V3–150 cycles (150 bp single read). Primers used to generate the amplicons are reported in Appendix 1. Raw sequencing data (FASTQ files) were analyzed using CRISPResso online tool[180], by setting Windows size =3, Minimum average read quality (phred33 scale) =30 and minimum single bp quality (phred33 scale) =10.

In silico off-target analysis. Off-target for crRNA +11 and +14 were analyzed by Cas-OFFinder online algorithm, by selecting: AsCpf1 from *Acidaminococcus* or LbCpf1 from *Lachnospiraceae* 5'-TTTV-3', Mismatch Number ≤4, DNA bulge size =0 and as a target genome the Homo sapiens (GRCh38/hg38) – Human.

In silico splicing prediction. For wild-type, mutated and edited CFTR gene sequences, a region of 400 bp spanning either the 3272-26A>G or 3849+10Kb C>T locus was analyzed by HSF and MaxEnt prediction algorithms[181,182] available at the Human Splicing Finder website (www.umd.be/HSF3/). Splice sites score were normalized to the score of the 3272-26A>G, or 3849+10Kb C>T splice site.

Statistical analyses. Statistical analyses were performed by GraphPad Prism version 6. For organoids experiments ordinary one-way analysis of variance (ANOVA) was performed. For the *in silico* prediction analyses Shapiro-Wilk test was used to verify the distribution of the data. Significance of the data was calculated by two-tailed Wilcoxon signed-rank test. Differences were considered statistically different at $P<0.05$.

RESULTS

Set up of a genome editing strategy to correct the 3272-26A>G splicing mutation in the CFTR gene

CFTR 3272-26A>G is a splicing mutation that forms an alternative acceptor splice site 26 bases before the canonical one of exon 20, causing the inclusion of 25 extra nucleotides into the mature transcripts (schematized in Fig 13a). We wanted to evaluate whether genetic perturbation through CRISPR technology could correct the aberrant transcription. We considered that the genome editing strategy could be applied either to delete the mutation or to alter sequences near the mutation that could play a role in CFTR splicing regulation of this region.

To verify our hypothesis, we generated minigene models to mimic the splicing pattern of the *CFTR* gene and test the efficacy of CRISPR nucleases for splicing correction. Indeed, minigenes are useful models for studying RNA splicing regulation[29,183,184]. These constructs are commonly used to study specific cis-acting elements and their binding factors involved in constitutive or alternative splicing regulation[185].

Minigene models for 3272-36A>G mutation include the region encompassing part of exon 18, full length exons 19 and 20, and intron 19 (legacy name: exon 16, 17a, 17b and intron 17a), either wild-type (pMG3272-26WT) or carrying the 3272-26A>G mutation (pMG3272-26A>G) (Fig.13a).

The splicing pattern produced by the mutated or wild-type minigenes, respectively, was evaluated by RT-PCR analysis in transfected HEK293T cells showing the different splicing pattern of the wild-type and mutated minigenes (Fig.13b). As reported in Figure 13b the pMG3272-26WT produces 100% of correct splicing, whereas the pMG3272-26A>G transcripts undergo alternative splicing, including 25 bases from intron 19. The splicing products identified by PCR were analyzed by Sanger sequencing to verify the presence of the expected pattern from the wild type or mutated minigenes (Fig. 13c).

We then designed single guide RNA (sgRNA) for SpCas9 (Appendix 2) targeting intron 19 where the 3272-26A>G mutation is located. The sgRNAs were programmed to generate either single or double cleavages in the non coding sequence to obtain isolated indels or deletions within this region. To evaluate the

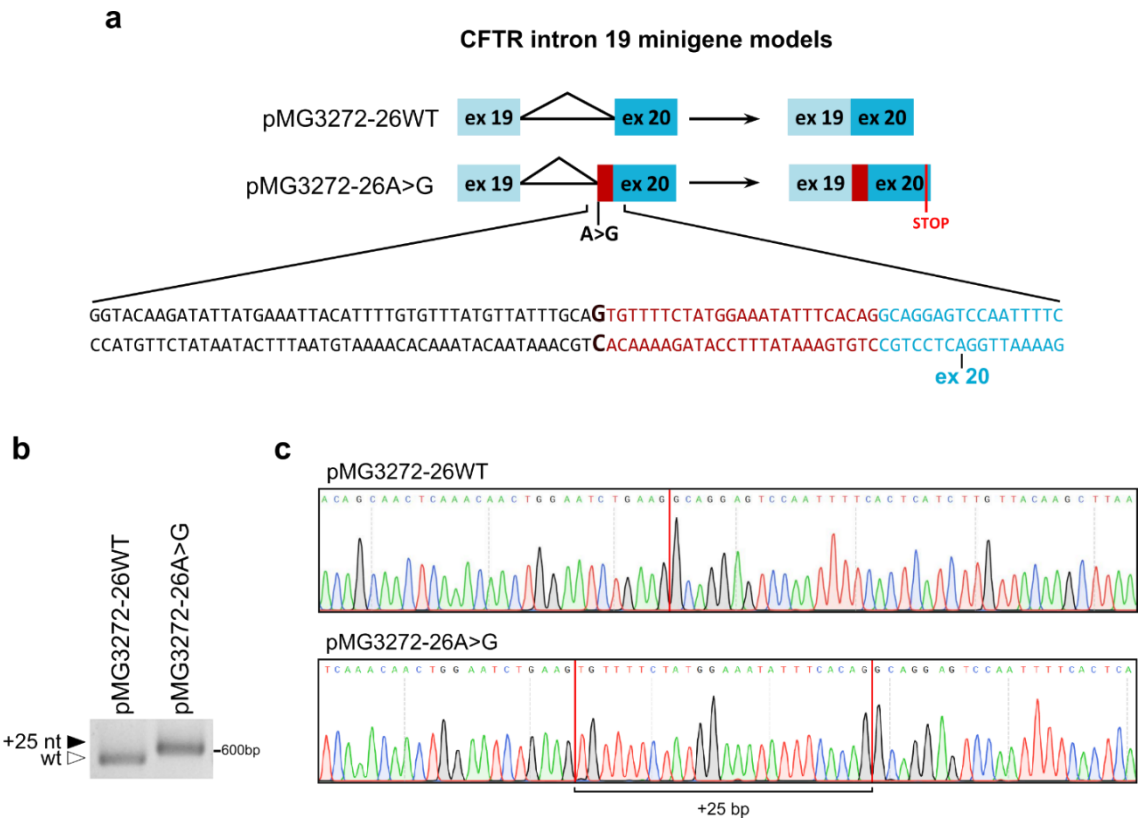


Figure 13. **pMG3272-26WT and pMG3272-26A>G *CFTR* minigene models.**
a, Scheme of the *CFTR* minigenes containing a sequence (~1.3 Kb) corresponding to the *CFTR* region extending from exon 19 to 20 either wild type (pMG3272-26WT) or 3272-26A>G mutated (pMG3272-26A>G). Exons are shown as boxes, introns as lines; the expected spliced transcripts are represented on the right according to the presence or absence of the 3272-26 A>G mutation. The lower panel shows nucleotide sequence and intron-exon boundaries nearby the 3272-26A>G mutation (labelled in bold). **b**, Splicing pattern of *CFTR* wild type (pMG3272-26WT) and mutated (pMG3272-26A>G) minigene models, transfected in HEK293T, by agarose gel electrophoresis analysis of RT-PCR products. Black-solid arrow indicates aberrant splicing, white-empty arrow indicates correct splicing. **c**, Sanger sequencing chromatograms of minigene splicing products from (**b**). Red lines represent the boundary between exons.

restoration of the correct splicing pattern in pMG3272-26A>G model, cells were transiently transfected with the minigene construct and SpCas9 in combination with a single or a pair of sgRNA. We observed an increased level of the correct transcript with four different sgRNA pairs (-52/+9, -47/-0, -47/+9, -47/+10), but not when they were used as a single sgRNA (Fig. 14a), suggesting that indels generated by SpCas9 and only one sgRNA are not sufficient to correct the mutated splicing. The efficiency of the deletion induced by selected sgRNA pairs was evaluated by PCR showing the expected DNA fragments after SpCas9 double cleavage (Fig. 14b).
To further investigate this splicing correction strategy, we generated a cell line stably transfected with the pMG3272-26WT or pMG3272-26A>G (HEK293/3272-26WT or

HEK293/3272-26A>G, respectively), in order to test cleavage efficiency and splicing correction in a more physiological chromatin context. Surprisingly, RT-PCR analysis of the transcripts revealed that none of the sgRNA pairs tested was able to correct the splicing pattern in this context (Fig. 14c). Moreover, the evaluation of DNA deletion efficiency by PCR resulted in no cleaved products (~50bp smaller, Fig. 14d), demonstrating that the inefficient splicing correction was due to the lack of Cas9 editing in the 3272-26A>G locus.

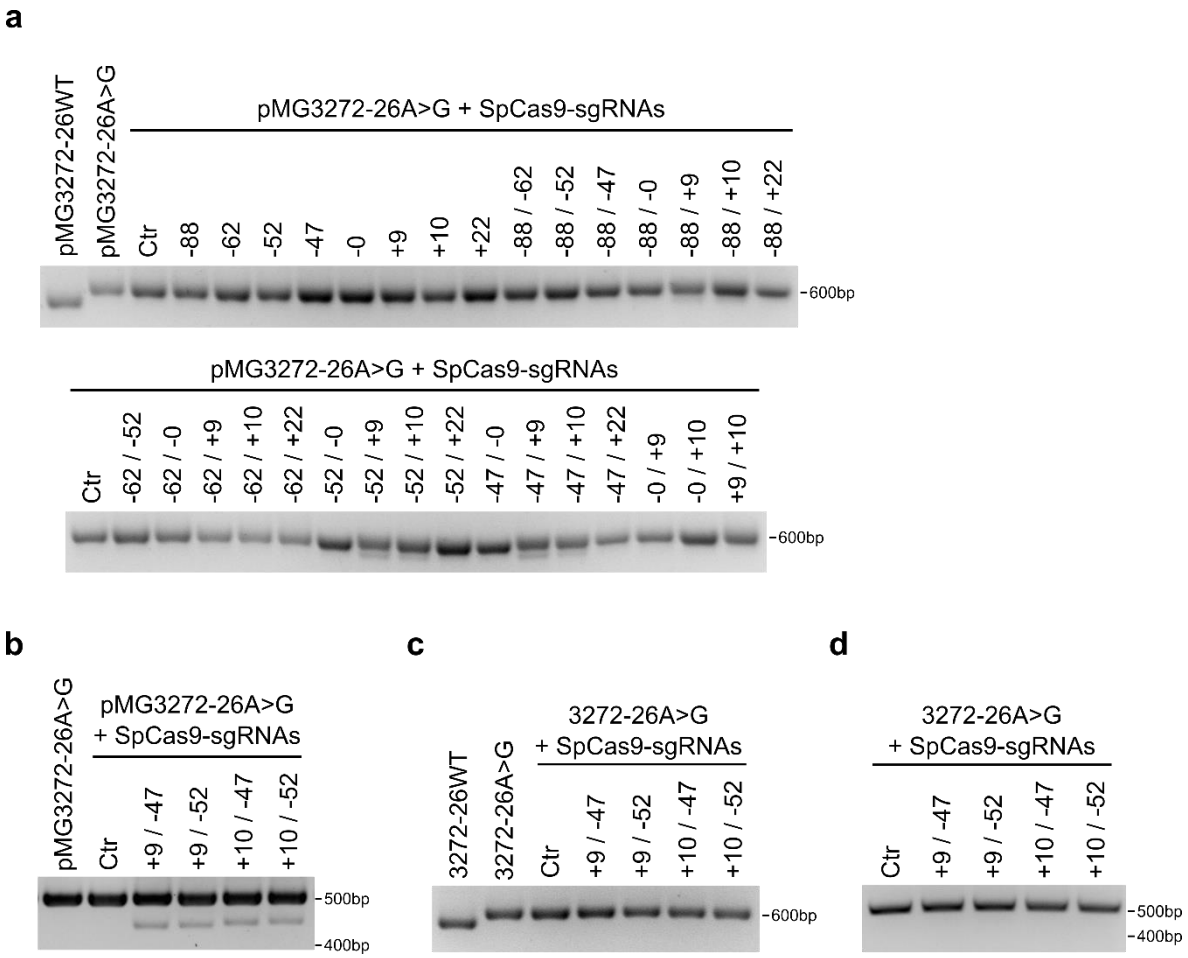


Figure 14. **SpCas9 sgRNA functional screening for splicing correction of 3272-26A>G minigene.**

a, SpCas9-sgRNAs screening based on the ability to restore the correct splicing pattern of *CFTR* 3272-26A>G minigene. The nuclease and sgRNAs single or in pair were transfected in HEK293T with pMG3272-26A>G, RT-PCR products were analyzed by agarose gel electrophoresis. pMG3272-26WT was used as a reference for correct intron 19 splicing. **b**, Agarose gel electrophoretic analysis of targeted deletions in 3272-26A>G minigene after cleavage by SpCas9-sgRNA pairs from (**a**) measured by PCR. The bigger band represents non-edited minigene sequences, the smaller band is the expected deletion product. **c,d**, Agarose gel electrophoresis of RT-PCR products (**c**) and PCR of targeted deletion (**d**) for SpCas9-sgRNA pairs selected from (**a**) in HEK293 having stable genomic integration of 3272-26A>G minigene (HEK293/pMG3272-26A>G cells).

In consideration of the poor genome editing obtained by SpCas9 we moved to another CRISPR nuclease, the AsCas12a. We design a pool of crRNAs that target the region surrounding the mutation (Appendix 2) and, as performed with SpCas9, we tested them both as single and combination with AsCas12a in transient co-transfection with pMG3272-26A>G.

The splicing analyses showed that AsCas12a together with crRNA -2 or +11, used individually or in combination with another crRNA, corrected the splicing pattern in the minigene model (Fig. 15a). Strikingly, DNA analysis of the locus revealed the absence of deletions in samples with crRNA pairs restoring the correct splicing. Weak deletion products were observed only in samples were the splicing was not

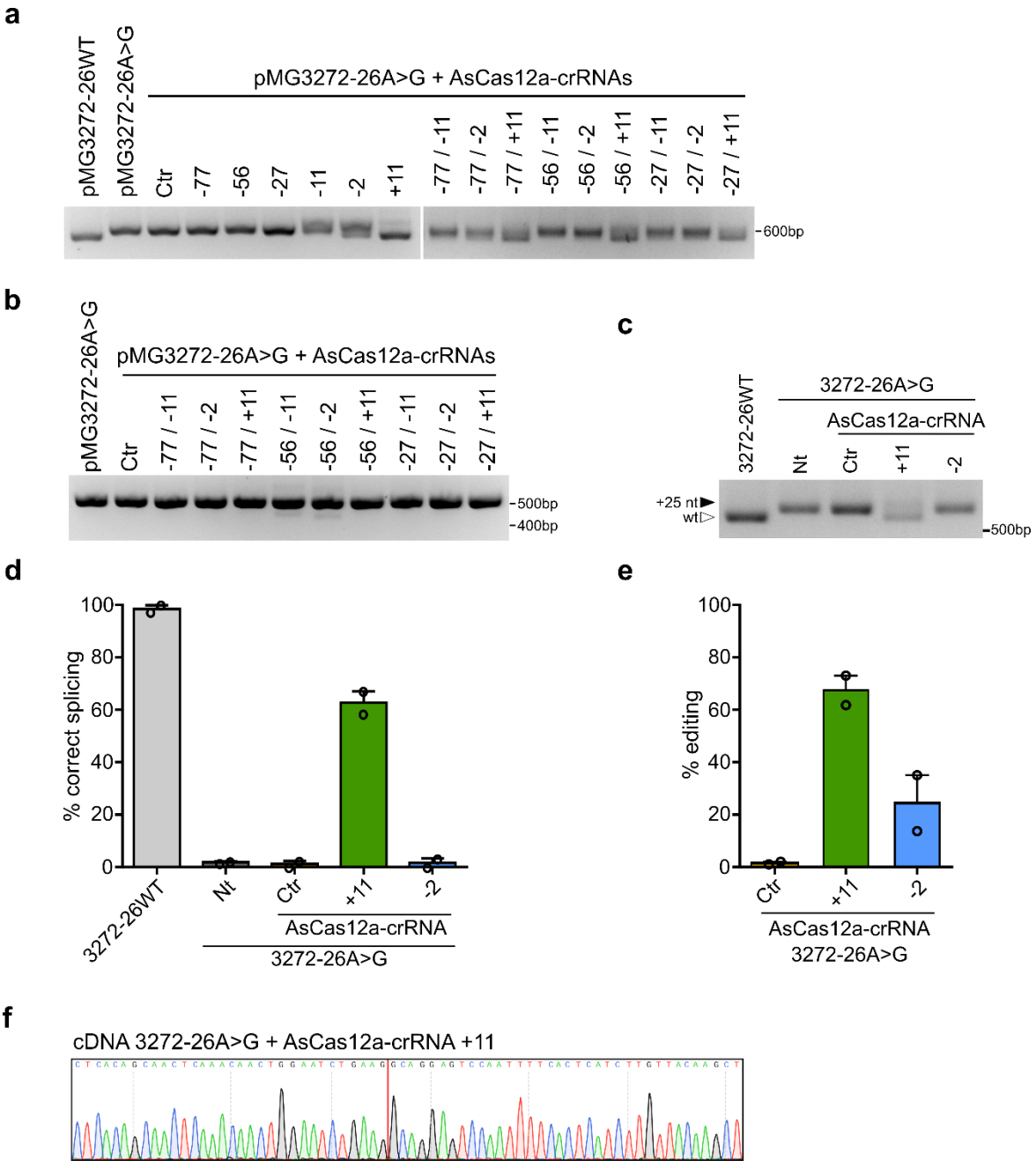


Figure 15. **AsCas12a-crRNA functional screening for splicing correction of 3272-26A>G minigene.**
a, AsCas12a-crRNAs screening based on the ability to restore the correct splicing pattern of *CFTR* 3272-26A>G minigene. Nucleases and crRNAs single or in pair were transfected in HEK293T with pMG3272-26A>G, RT-PCR products were analyzed by agarose gel electrophoresis. pMG3272-26WT was used as a reference for correct intron 19 splicing. **b**, Agarose gel electrophoretic analysis of targeted deletions in 3272-26A>G minigene after cleavage by AsCas12a-crRNA pairs from (**a**) measured by PCR. The bigger band represents non-edited minigene sequences, the smaller band is the expected deletion product. **c**, Splicing pattern analyzed by RT-PCR in HEK293/pMG3272-26A>G cells following treatments with AsCas12a-crRNA control (Ctr) or specific for the 3272-26A>G mutation (+11 and -2). Black-solid arrow indicates aberrant splicing, white-empty arrow indicates correct splicing. Representative data of n=2 independent experiments. **d**, Percentages of correct splicing measured by densitometry and **e**, editing efficiency analyzed by TIDE in cells treated as in **c**. Data are means ± SEM from n=2 independent experiments. **f**, Sanger sequencing chromatogram of correct intron 19 splicing from 3272-26A>G integrated minigene after AsCas12a-crRNA+11 editing from Fig.1b. Red line represents the boundary between exons 19-20.

repaired (compare Fig. 15a and 15b). These results suggest that the editing by a single crRNA (-2 or +11) is sufficient to correct the splicing pattern in the mutated minigene.

We validated the efficacy of AsCas12a correction in HEK293/3272-26A>G stable cell line. Conversely to what we obtained with SpCas9, the AsCas12a-crRNA+11 generated more than 60% of correct transcripts (Fig. 15c,d,f) and efficient DNA editing (67.4%) of the transgene (Fig. 15e), which is consistent with the degree of restored splicing observed.

To analyse the editing efficacy by a single crRNA, we used the tracking of indels by decomposition (TIDE) analysis[174]. This tool allows the estimation of relative abundance of all the possible indels performed by CRSIPR nucleases in the target sequence, decomposing the quantitative sequence trace (chromatogram) data from Sanger sequencing into its individual components.

The editing analyses of the integrated minigenes following editing with AsCas12a-crRNA+11, revealed a heterogeneous pool of deletions (Fig. 16a). To analyse the contribution of each single editing events to the restoration of the correct transcript, edited sequences were cloned into the pMG3272-26A>G minigene. Single editing events and their splicing patterns were characterized (Fig. 16b,c). We can observe that editing of the locus is mainly represented by deletion events of different length, ranging from -1 nt to -26 nt (Fig. 16b), which is consistent with the editing profile of AsCas12a reported in literature (deletions bigger than 4 bases)[186]. In particular, the 18 nucleotides deletion (that occur with a high frequency of 9/34 clones) completely

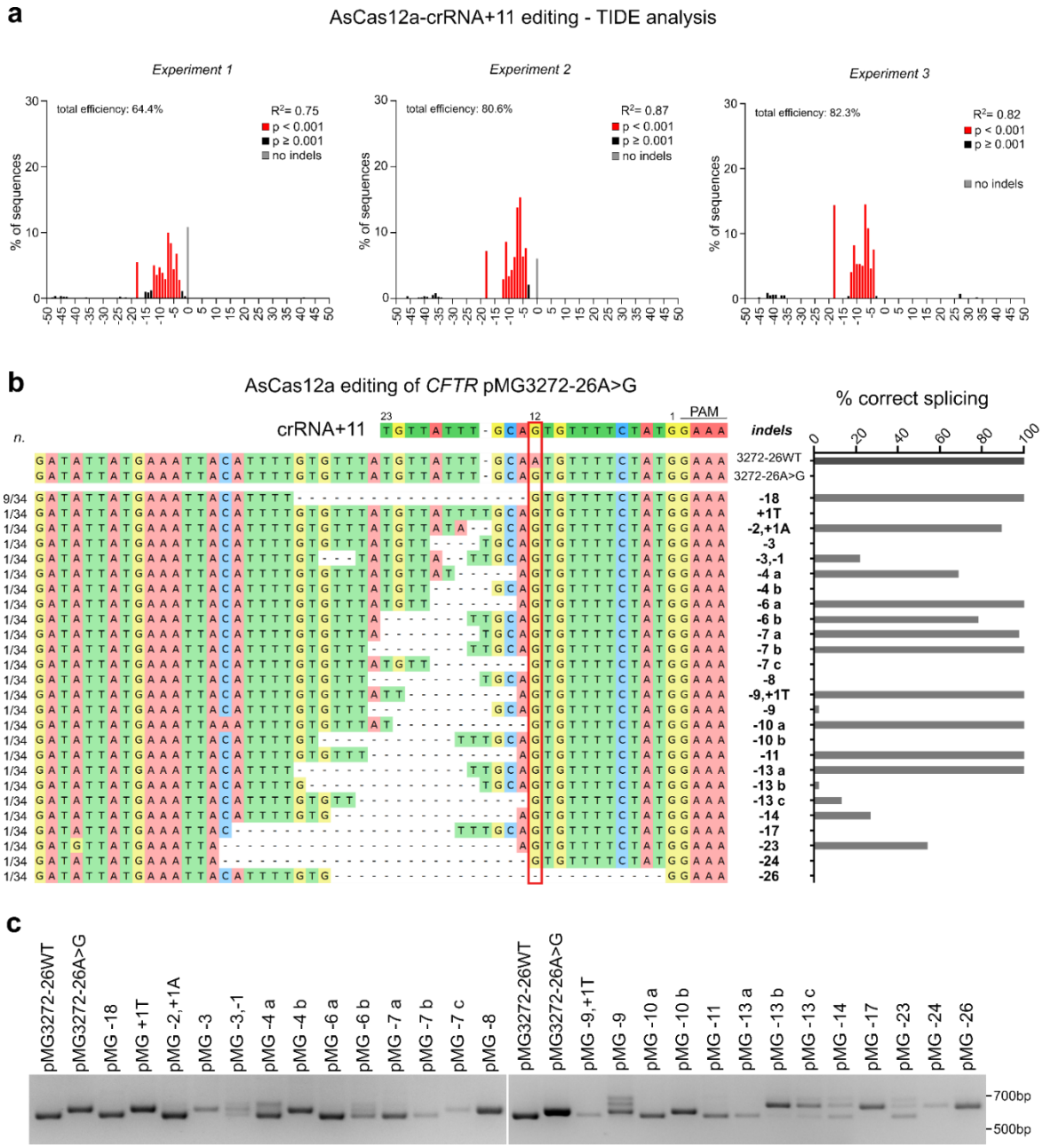


Figure 16**. Repair pattern after AsCas12a-crRNA+11 cleavage**
a, Indels spectrum by TIDE analysis from HEK293/pMG3272-26A>G cells after AsCas12a-crRNA+11 editing from n=3 independent experiments. **b**, Indels triggered by AsCas12a-crRNA+11. The 3272-26A>G locus from cells edited by crRNA +11 were amplified, cloned in the minigene backbone and Sanger sequenced (34 different clones, left panel), or analyzed as in **b** for produced splicing pattern (right panel and Supplementary Fig. 3a). pMG3272-26WT and pMG3272-26A>G were used as references. **c**, Splicing pattern of edited sites cloned into the minigene plasmid. Minigenes were transfected in HEK293T cells and RT-PCR products were analyzed by agarose gel electrophoresis.

restored the correct splicing pattern of the pMG3272-26A>G minigene (Fig. 16b,c). The vast majority of the remaining events, occurring at a low frequency (1/34 clones), contributed to splicing correction. In some cases, we can observe the

appearance of additional splicing products (Fig. 16c), attributable to an artefact of the minigene model. Overall, we measured a correction of the splicing pattern in 68% of the events analysed.

Interestingly, sequence analyses revealed the persistence of the 3272-26A>G mutation after AsCas12a-crRNA+11 cleavage. For this reason, we performed an *in silico* analysis to evaluate the effects of the indels caused by AsCas12a-crRNA+11 on the mutated splice site. All the indels with a frequency higher than 1% were analysed by HSF and MaxEnt prediction algorithms[181,182], showing that the large majority of events correlated with decreased strength of the aberrant splice site induced by the 3272-26A>G mutation (Fig. 17).

In conclusion, AsCas12a in combination with the single crRNA+11 efficiently correct the splicing of the 3272-26A>G CF mutation in a minigene model by producing small deletions upstream the mutated splice site, while preserving the mutation.

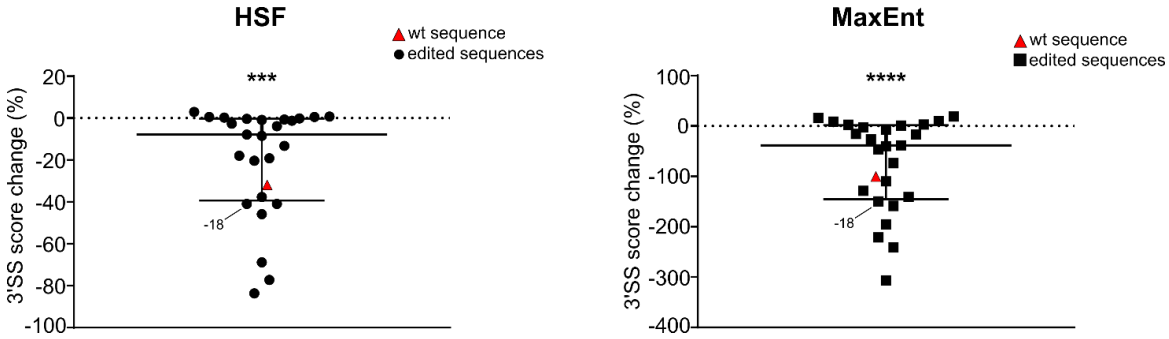


Figure 17**. In silico splicing score prediction.**
Prediction of the score change of AsCas12a-crRNA+14 modification on the mutated 3'splice site (ss) by bioinformatics tools (HSF and MaxEnt prediction algorithms, see Methods section). Data are median with interquartile range. Statistical analysis was performed using two-tailed Wilcoxon signed-rank test; ***P<0.001, ****P<0.0001. Score of wild type sequence was given only as reference in the graph.

Specificity of AsCas12a based correction in 3272-26A>G models

An important issue relative to the use of CRISPR nucleases is their off-target activity against similar sequences. Indeed, it has been reported that CRISPR nucleases may induce cleavages at non-target sites due to the possibility of tolerating mismatches between the guide RNA and the target genomic sequence.

Another important aspect that should be considered when designing a gene therapy approach is the allelic discrimination when the therapeutic goal is limited to a single

allele. In the case of CF with 3272-26A>G mutation, the majority of patients are compound heterozygous, thus making highly relevant the accurate evaluation of the possible AsCas12a-crRNA+11 cleavage and modification of the second mutant allele, carrying the wild-type 3272-26 sequence.

We initially tested the cleavage properties of AsCas12a-crRNA+11 in HEK293/3272-26WT or HEK293/3272-26A>G cell lines, carrying the wt or mutated intronic sequences. In Fig. 18a we can observe that while crRNA+11 efficiently cleaves the mutated locus (as already shown before), it does not cleave with the same efficiency the wild-type sequence, showing a drop of the editing from 77% to only 7.5%. This results in a 10-fold difference of cleavage between the mutated and wild-type allele. To further evaluate the specificity of AsCas12a, we designed a crRNA targeting the CFTR 3272-26WT sequence, crRNA+11/wt, thus identical to the crRNA+11 except for the mutated nucleotide. AsCas12a-crRNA+11/wt showed high cleavage efficiency in HEK293/3272-26WT cells (86.5%) and low editing in HEK293/3272-26A>G cells (12%, Fig. 18a), confirming the allelic discrimination by AsCas12a with the selected crRNA. The splicing pattern after AsCas12a-crRNA+11 and crRNA+11/wt editing was evaluated by RT-PCR to exclude any negative alteration of the produced transcripts, following genome editing modification of the intronic region. As we can observe in Fig. 18b, except for the splicing correction by crRNA+11, no alteration of the transcripts was observed in HEK293/3272-26WT after AsCas12a-crRNA+11/wt cleavage.

The experiments so far reported were performed in HEK293 cells, which are commonly used for genome editing set up. We next move to more relevant cellular model used in CF research. Caco-2 cell are colorectal adenocarcinoma epithelial cells that express the endogenous wild-type *CFTR*. We evaluated the specificity of the proposed strategy by delivering AsCas12a-crRNA+11 with lentiviral vectors in order to achieve long term nuclease expression. This will allow us to evaluate the off target activity on the wild-type allele, since the non-specific cleavages are usually associated with long term expression of CRISPR nucleases[104], and to consider a delivery systems that could be potentially applied for CF in vivo gene therapy. Ten days after transduction no editing above TIDE background levels (about 1%)[174] was measured with AsCas12a-crRNA+11, while AsCas12a-crRNA+11/wt was very efficient in the cleavage of the wt *CFTR* sequence in Caco-2 cells (86.3%, Fig. 18c).

Transcripts analyses in Caco-2 cells did not reveal any major alteration in the splicing pattern of the wt *CFTR* after intron modifications (Fig. 18d).

In order to perform an unbiased analysis of the off targets at genome wide scale, we performed the GUIDE-seq analysis[179,187]. This technique is based on the integration of a 34 bp-long oligonucleotide in the genomic loci cleaved by Cas nuclease, allowing their amplification during library preparation and subsequent

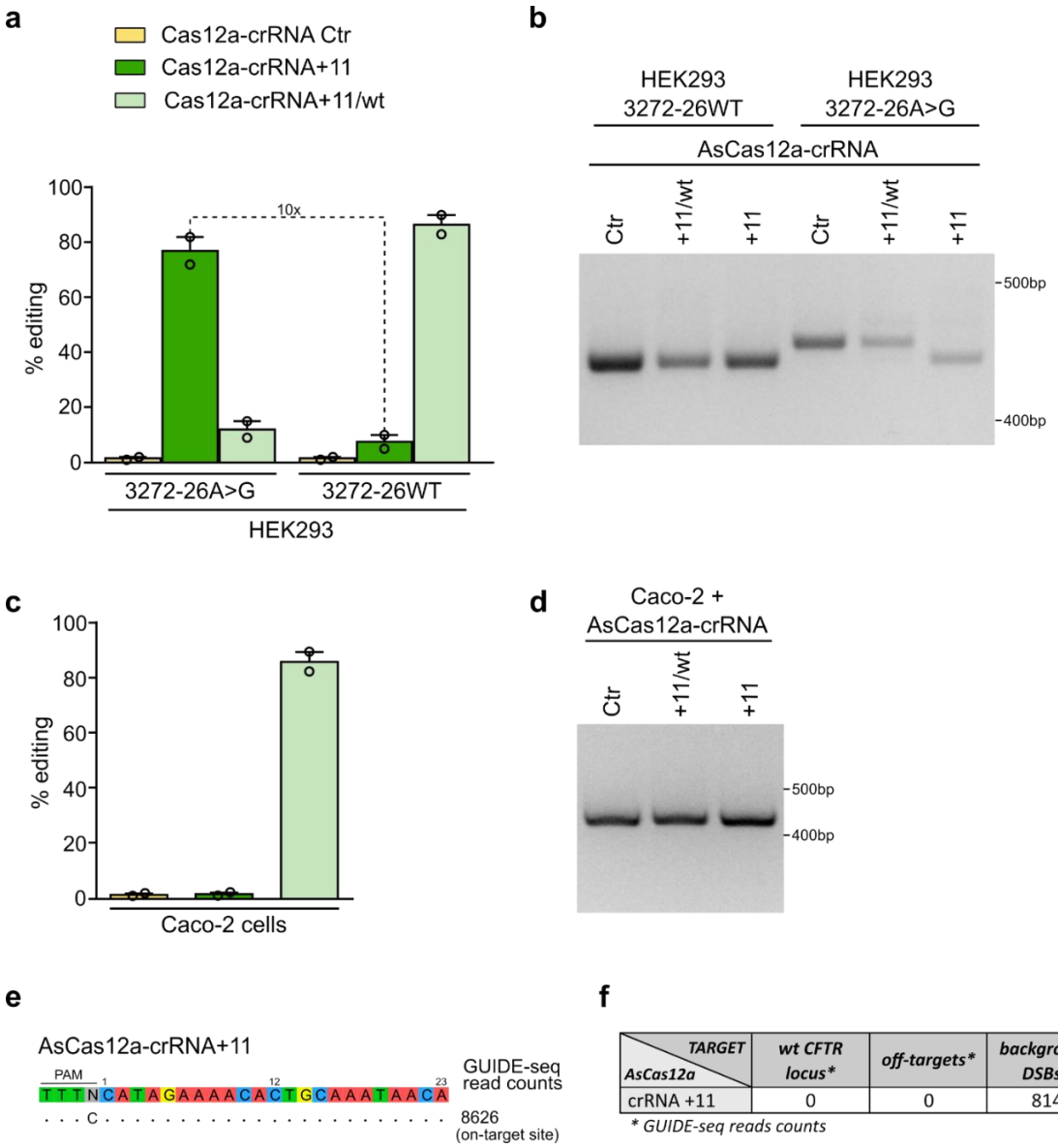


Figure 18. **Target specificity of AsCas12a-crRNA+11.**
a, editing efficiency by TIDE analysis in HEK293/pMG3272-26WT or HEK293/pMG3272-26A>G following lentiviral transduction of AsCas12a-crRNA+11 or +11/wt, as indicated, and selected with puromycin for 10 days. **b**, RT-PCR products analysis of cells from (**a**). **c**,**d**, Editing efficiency by TIDE analysis (**c**) and RT-PCR products analysis (**d**) in Caco-2 cells, having WT *CFTR* sequence. Data are means ± SEM from n=2 independent experiments. Images are representative of two independent experiments. **e**, GUIDE-seq analysis of crRNA+11.

identification through next generation sequencing. Off-target profiling of AsCas12a-crRNA+11 showed very high specificity for the mutated sequence in HEK293/3272-26A>G cells and no cleavage detection of unwanted sequences (Fig. 18e,f), supporting the safety and specificity of the proposed strategy.

3272-26A>G splicing correction in primary airway cells

The AsCas12a-crRNA+11 genome editing strategy that we set up in HEK293 cells was subsequently validated in cellular models more relevant for CF. To this aim we used human primary airway epithelial cells that are considered a relevant model system for CF studies. Indeed, two-dimensional (2-D) cultures allow disease modelling and preclinical testing of new therapies[188,189]. Primary airways epithelial cells derived from a patient compound heterozygous for the 3272-26A>G mutation (3272-26A>G/ΔF508) were kindly provided by the Italian Cystic Fibrosis Research Foundation.

The analysis of the transcripts via RT-PCR of the 3272-26 region, showed the presence of two different splicing patterns (Fig. 19a). The longer transcripts corresponded to the splicing pattern of the 3272-26A>G mutated allele (+25 nt), while the shorter ones showed a wild-type splicing pattern. These wild-type transcripts are derived either from the ΔF508 allele, having a normal splicing in the region under analysis, or from the 3272-26A>G allele, partially producing correct spliced transcripts. The difference in size and abundance of these transcripts is consistent with the cell heterozygosity and in agreement with previous data[190].

Following lentiviral delivery of the AsCas12a-crRNA+11 treatments, we decided to either treat or not the cells with puromycin selection to evaluate the efficiency of lentiviral transduction in airway cells. Splicing analyses in Fig. 19a showed the disappearance of the mutated transcripts after the treatment with AsCas12a-crRNA+11, which did not differ much between samples selected or not with puromycin. The abundance of the mutated transcripts was measured by chromatogram deconvolution analyses, that allowed the quantification of the different traces, corresponding to the two splicing patterns, after Sanger sequencing of the RT-PCR products. This analysis confirmed the correction of the splicing by

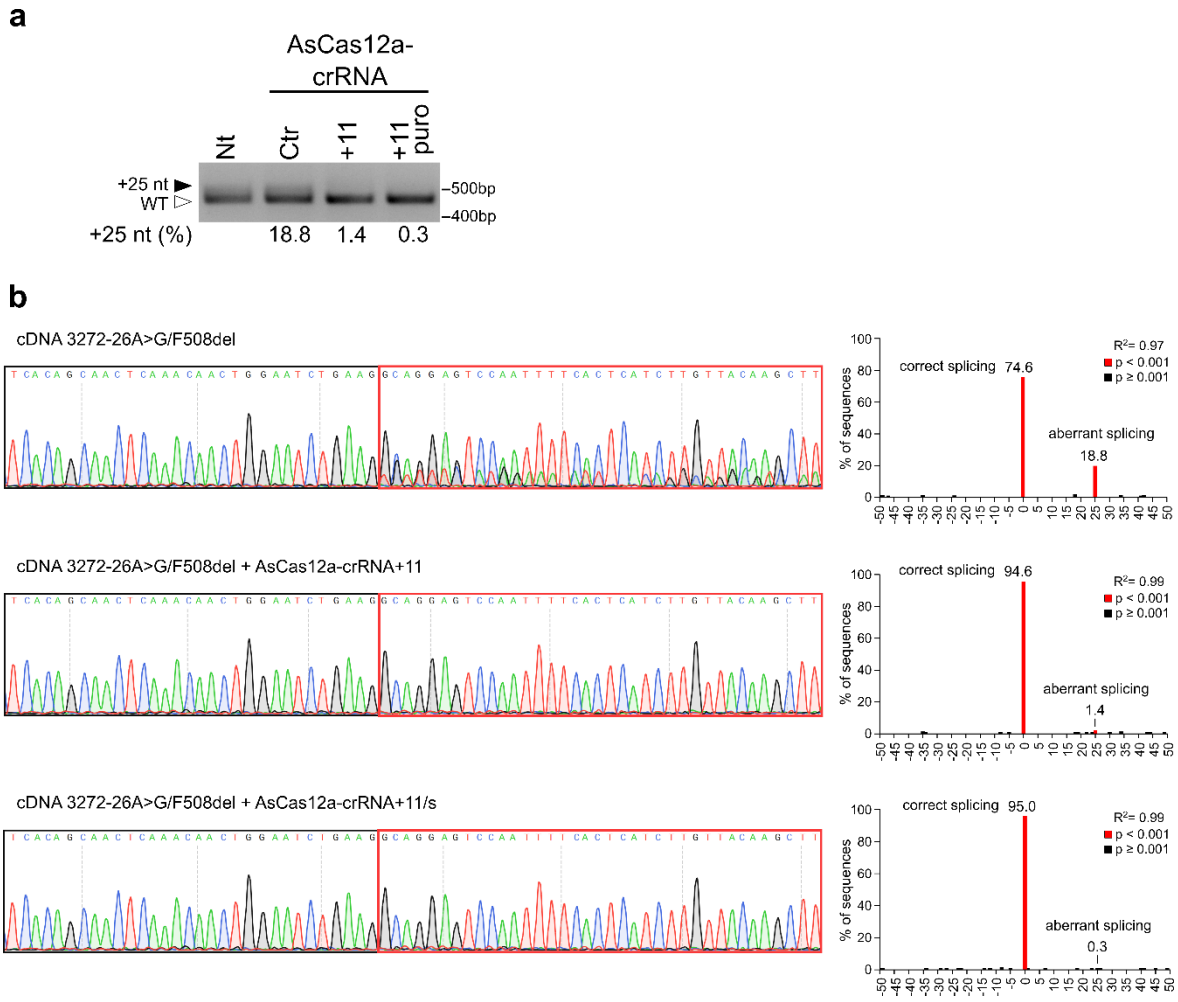


Figure 19. **AsCas12a-crRNA+11 genome editing analysis in 3272-26A>G models.**
a, Splicing pattern analysis by RT-PCR in 3272-26A>G primary airway cells following lentiviral transduction (15 days) of AsCas12a-crRNA control (Ctr) or specific for the 3272-26A>G mutation (+11). Puromycin selection was performed for 72h in +11 puro. Black-solid arrow indicates aberrant splicing, white-empty arrow indicates correct splicing. **b**, Left panels, chromatograms of RT-PCR products from Ctr, AsCas12a-crRNA+11, AsCas12a-crRNA+11-puro samples, as indicated, form (**a**), showing two different transcripts of 3272-26A>G/ΔF508 primary airway cells. Red box indicates chromatogram area after exon19-exon20 junction. Right panels, chromatograms deconvolution analysis was used to evaluate the amount of mutated splicing (inclusion of +25 nt from intron 19) before and after AsCas12a-crRNA+11 cleavage.

AsCas12a-crRNA+11 (+25 nt isoform: 18.8% control sample, 1.4% crRNA+11, 0.3% crRNA+11 puro, Fig. 19b).

To correlate the level of correct splicing with the amount of indels induced, we initially performed TIDE analyses revealing that the editing efficiency is 30% without puromycin, and up to 42% following puromycin selection (Fig. 20a,b). Based on the data obtained in minigene models and Caco-2 cells, we expected that crRNA+11 cleavages occurred mainly at the mutated allele. However, through TIDE analysis is not possible to distinguish whether the editing occurred at the mutated 3272-

26A>G or at the 3272-26WT (ΔF508) *CFTR* alleles. Thus, based on specificity data of crRNA+11, we inferred that the editing efficiency of the mutated allele is double.

To verify the allelic discrimination in primary airway cells, we performed a targeted deep sequencing analysis. This method offers the advantages of analysing different targets simultaneously with a high coverage, allowing the detection of low abundance editing events.

Deep sequencing revealed high editing of the 3272-26A>G allele, 78.7%, and the complete absence of indels in the second *CFTR* allele, 3272-26WT (Fig. 21a,b). We can observe that the indel profile of crRNA+11 in these cells is similar to the one obtained into HEK293/3272-26A>G cells (Fig. R4b), with a high frequency of the 18 nucleotides deletion (Fig. 21a).

The off-target profile of AsCas12a-crRNA+11 was further evaluated in primary airway cells. *In silico* off-target prediction was performed by Cas-OFFinder online algorithm, to identify sequences with high similarity with the crRNA+11. Based on

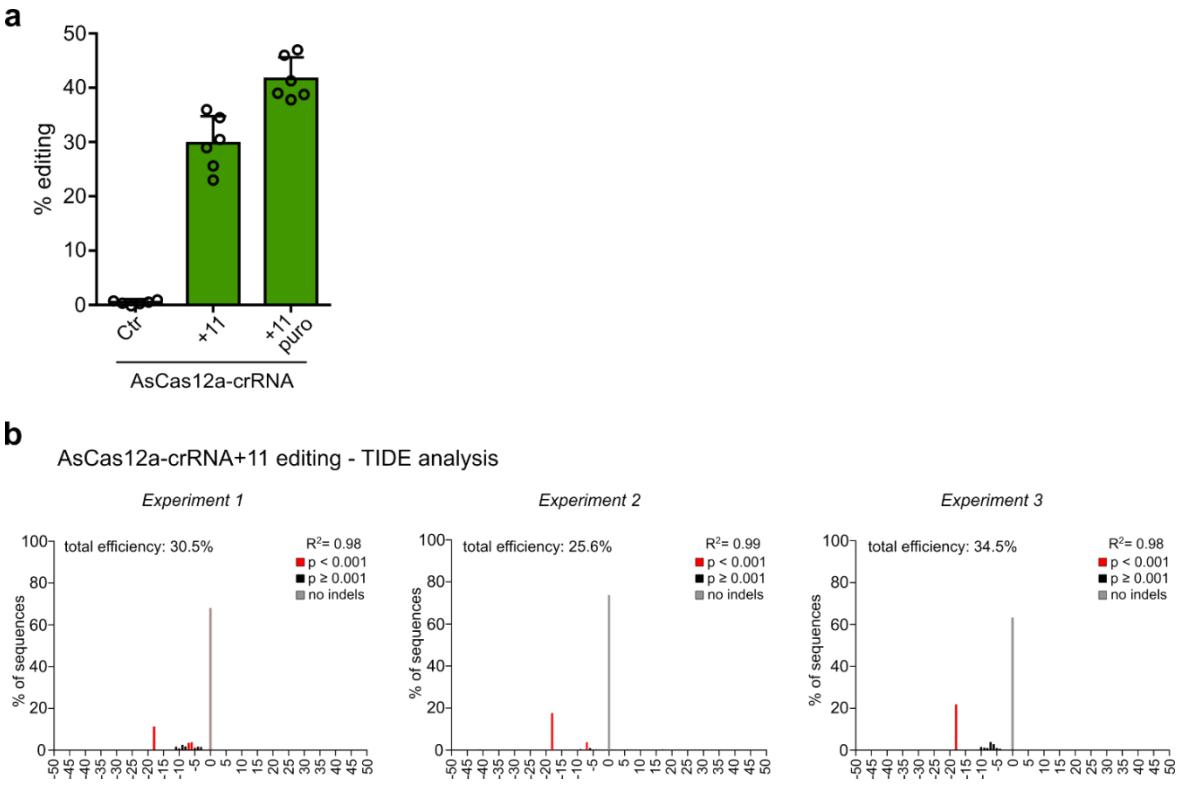


Figure 20. **AsCas12a-crRNA+11 editing pattern in primary airway cells**
a, Percentages of indels generated in intron 20 by crRNA+11 in cells following lentiviral transduction (15 days) of AsCas12a-crRNA+11 measured by TIDE analysis. Data are from n=2 independent experiments. **b**, Indels spectrum by TIDE analysis from 3272-26A>G/ΔF508 primary airway cells after AsCas12a-crRNA+11 editing from n=3 independent experiments.

literature data reporting the high fidelity of AsCas12a[111], we decided to consider all the sequences up to 4 mismatches with the original target.

The 12 identified possible off-target site were analysed by deep sequencing showing no off-target cleavages by AsCas12a-crRNA+11 (Fig. 21c), consistent with GUIDE-seq results in HEK293/3272-26A>G cells (Fig. 21e,f).

c

AsCas12a-crRNA+11		Chromosome	Position	Mismatches	% editing crRNA+11	total reads crRNA+11	% editing crRNA CTR	total reads crRNA CTR
*3272-26A>G**	TTTCCATAGAAAACACTGCAAATAACA	chr7	117611543	0	78.7	129892	0.9	238374
3272-26WT	TTTCCATAGAAAACAtTGCAAATAACA	chr7	117611543	1	0.0	157201	0.9	241291
OFF 1	TTTCCATAGAtAAtACTGCAcATAACt	chr8	50202652	4	0.5	464287	0.4	462812
OFF 2	TTTCCAgAGAAAACACTaCAAATAAtg	chr3	176656937	4	0.4	332209	0.3	331609
OFF 3	TTTGCATAaAAAAgtCTaCAAATAACA	chr7	138131106	4	0.5	579560	0.4	498759
OFF 4	TTTACtTAGAAAAgAaTGaAAATAACA	chr4	132046364	4	0.7	713622	0.7	625136
OFF 5	TTTGCATAGAAAACACTtCcAcaAACA	chr5	25017613	4	0.7	592811	0.6	568162
OFF 6	TTTACATgGAAAAgACTGaAAATAgCA	chr5	160120802	4	0.7	434289	0.6	413415
OFF 7	TTTGCATAGAgAAtACTGCAAAaAACc	chr1	143617739	4	9.2	729449	10.1	487345
OFF 8	TTTAtATAtAAAAaACTtCAAATAACA	chr2	52000516	4	0.5	259045	0.5	581289
OFF 9	TTTACATAGAAAAaACTaaAAATgACA	chr15	70700393	4	0.7	45008	0.6	193068
OFF 10	TTTCCATAGAAAACAtTtCAAATgAaA	chr10	105528793	4	0.9	547443	1.0	451210
OFF 11	TTTACATgaAAAACACTGCAAgaAACA	chr6	110538141	4	0.9	637391	0.7	491770
OFF 12	TTTACAaAGAcAcCACTGaAAATAACA	chr9	37009449	4	0.5	211688	0.5	174784

*on-target site

Figure 21. **Deep sequencing analysis of in silico predicted off-target sites for crRNA+11 in primary airway epithelial cells.**

a, Deep sequencing analysis of the *CFTR* on-target locus after AsCas12a-crRNA+11 transduction (15 days) of the 3272-26A>G primary airway cells. **b**, Percentage of deep sequencing reads of the edited and non-edited 3272-26A>G or WT alleles from (**a**). **c**, off-target analysis of in silico predicted sites. All the predicted sites with 4 or less mismatches (12 sites) were analyzed in 3272-26A>G primary airway cells treated as in (**a**).

Therefore, we concluded that, in a relevant cellular model, AsCas12a-crRNA+11 is specific in terms of allelic discrimination as well as genome-wide precision with no detectable off target sites even after long term nuclease expression through lentiviral delivery.

AsCas12a editing restored CFTR wt transcripts in intestinal organoids

Recent advances in adult stem cell biology have resulted in the development of organoid cultures, three-dimensional (3-D) multi-cellular structures that recapitulate tissue features of the parental organ and are usually grown from donor tissue fragments. Intestinal organoids from CF patients allow the quantification of CFTR function by recapitulating essential features of the in vivo tissue architecture, representing a near-physiological model for translational research[176,178,191].

Encouraged by the splicing correction and allele specificity obtained in minigene models and in primary airway epithelial cells, we next evaluated the rescue potential of the CF 3272-26A>G phenotype by AsCas12a-crRNA+11 in human intestinal organoids compound heterozygous for the 3272-26A>G mutation (3272-26A>G/4218insT). The 4218insT mutation present the insertion of a T in exon 25 of the CFTR gene and it has been found in Belgian patients with mild respiratory symptoms and elevated sweat chloride[16]. The allele carrying this mutation present a correct splicing pattern in the region of the 3272-26A>G mutation, therefore in our analyses we considered it as a reference for the wild-type splicing.

AsCas12a-crRNA+11 was delivered in organoids through lentiviral vectors and no puromycin selection was performed in this model. As observed in primary airway cells, the sequence analyses of the transcripts via RT-PCR analysis showed two different products (Fig. 22a): the wild-type *CFTR* intron 19 splicing pattern and +25 nt transcripts derived from the 3272-26A>G mutation. Two weeks after transduction, we observed an efficient correction of the aberrant splicing (+25 nt) as demonstrated by the disappearance of the mutated transcripts (upper band, Fig. 22a). Chromatogram deconvolution confirmed the restoration of the correct splicing pattern in treated organoids (+25 nt isoform: 22.8% control sample, 0.9% crRNA+11, Fig. 22b). As a positive control for the experiments in intestinal

organoids, we delivered the wild-type CFTR cDNA through lentiviral vectors (gold standard of gene therapy approches), that produce only wt *CFTR* transcripts.

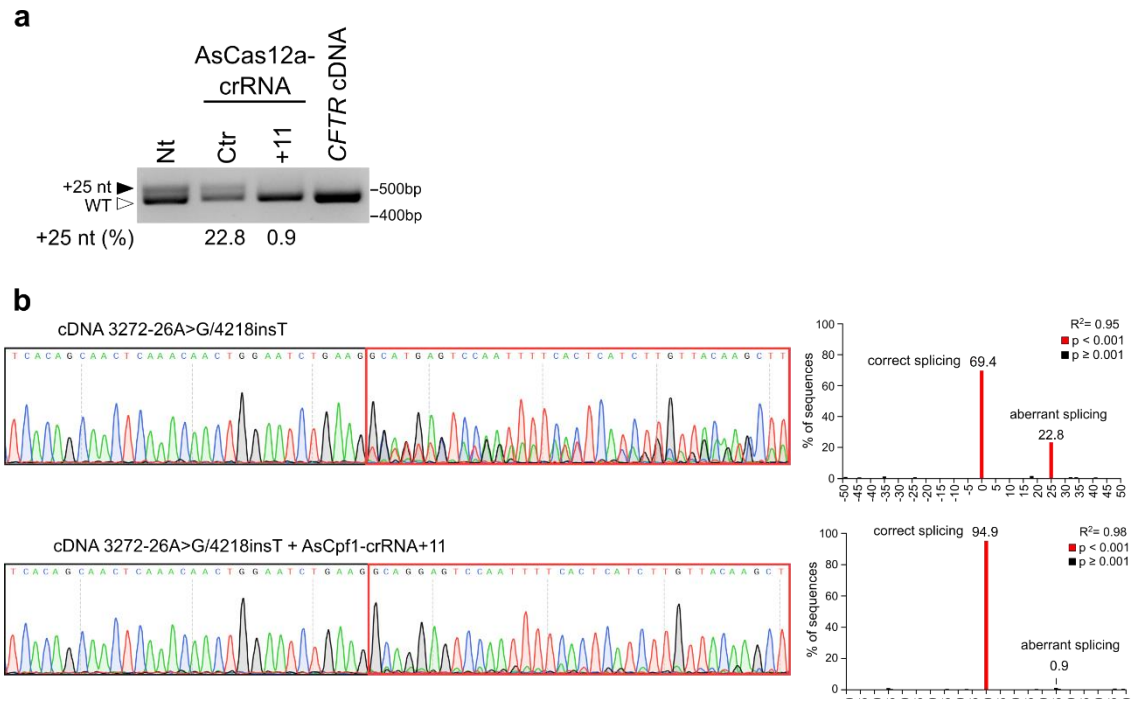


Figure 22. **CFTR splicing of 3272-26A>G mutated CF patient's organoids after genome editing with AsCas12a-crRNA+11.**
a, Splicing pattern analysis by RT-PCR in 3272-26A>G organoids following lentiviral transduction (14 days) of AsCas12a-crRNA control (Ctr) or specific for the 3272-26A>G mutation (+11) or with the *CFTR* cDNA. Black-solid arrow indicates aberrant splicing, white-empty arrow indicates correct splicing. **b**, Left panel, chromatogram of RT-PCR products from (**a**). Upper chromatogram represents the mixed population of mRNA transcripts of 3272-26A>G/4218insT organoids, the lower chromatogram shows transcripts after AsCas12a-crRNA+11 editing in these organoids. Red box indicates area after exon19-exon20 junction. Right panel, chromatogram deconvolution analysis was used to evaluate the amount of mutated splicing (inclusion of +25 nt from intron 19) before and after AsCas12a-crRNA+11 cleavage.

Editing efficiency was measured by T7 Endonuclease I assay (T7E1). This endonuclease cleaves heteroduplex sequences obtained by re-hybridization of amplicons of the target locus. In the presence of edited sequences, containing a variety of different indels, the annealing of the amplicons will result in heteroduplex formation. The digested product is separated in agarose gel, allowing the quantification of editing events. An editing efficiency of more than 30% (Fig. 23a) was measured by T7E1 assay in these organoids.

The high fidelity of AsCas12a-crRNA+11 prompted us to evaluate indels formation at the 3272-26 locus by targeted deep sequencing analysis as performed in primary epithelial cells. Results in Figure 23b,c confirmed the preferential and efficient

cleavage of the mutated allele by AsCas12a-crRNA+11: 39.77% of edited reads in the 3272-26A>G allele with an overall editing efficiency of 40.25% (only 0.48% of editing within the other allele, that is comparable to background). Moreover, if we consider the two allele separately, the indels formation on the 3272-26A>G allele is 84.9%, while only 0.9% of indels were found in the 3272-26WT allele (Fig. 23b), thus indicating a 94-fold allelic discrimination by AsCas12a-crRNA+11 in patient derived intestinal organoids.

In agreement with previous reports[192,193] and what was observed in minigene models and primary airway cells, the cleavage of intron 19 by AsCas12a-crRNA+11 produced variable length deletions, with the 18 nucleotide deletion as the most frequent events. Interestingly, all the editing events, with a frequency above 0.5% of the total DNA repair events, were already isolated in pMG3272-26A>G model (Fig. 16b,c), contributing to splicing correction.

In conclusion these data demonstrated that AsCas12a-crRNA+11 efficiently and precisely edited the mutated allele leading to the production of correct spliced transcripts.

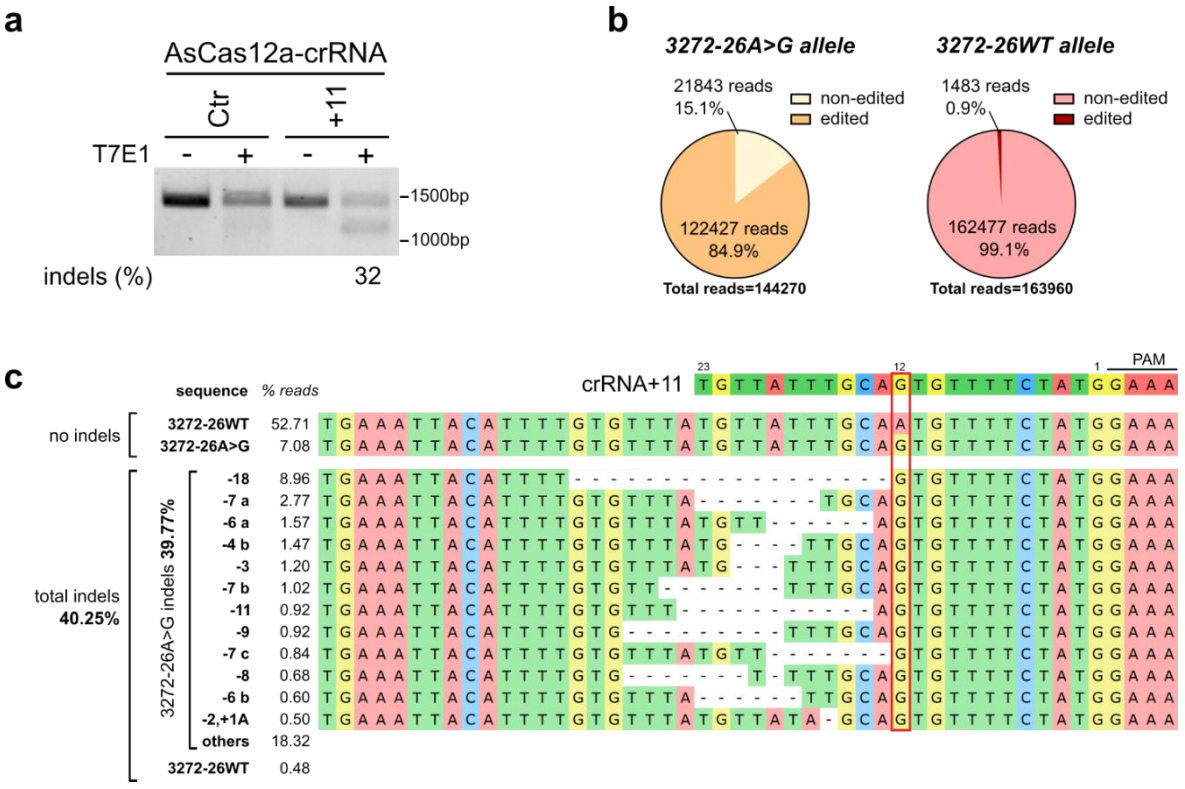


Figure 23. **Efficiency of crRNA+11 in intestinal organoids**
a, Editing efficiency in 3272-26A>G organoids measured by T7E1 assay following lentiviral transduction as in Fig. R10a. **b**, Percentage of deep sequencing reads of the edited and non-edited 3272-26A>G or WT alleles from (**a**). **c**, Deep sequencing analysis of the *CFTR* on-target locus after AsCas12a-crRNA+11 transduction of the 3272-26A>G organoids (average from n=2 independent experiments) (Supplementary Data 2).

CFTR channel function is restored after AsCas12a-crRNA+11 editing

The CFTR channel activity can be measured in intestinal organoids derived from CF patients taking advantage of the dependence of organoids size and lumen formation (swelling) on CFTR function[176](Fig. 24a). This functional analysis in organoids is widely considered a valuable tool to test the efficacy of therapeutic approaches under investigation.

Organoids area was evaluated 14 days post treatment with AsCas12a-crRNA. In Figure 24b we can observe that in control conditions organoids showed a small size and a small lumen, indicating that the mutations 3272-26A>G/4218insT produce a non-functional CFTR channel. Instead, following AsCas12a-crRNA+11 treatment the area of the organoids increased by 2.5-folds at the steady state (T=0 min, Fig. 24b,c), thus indicating restored channel function. Noteworthy, there was no significant difference in organoids area between treatments with AsCas12a-crRNA+11 or transduction of wild-type *CFTR* cDNA (Fig. 24c), indicating the remarkable efficiency of the AsCas12a-crRNA+11 genetic editing in 3272-26A>G phenotypic reversion.

To further assess the recovery of CFTR channel function following AsCas12a-crRNA+11 treatment, we performed the Forskolin Induced Swelling (FIS) assay[176] (Fig. 24b,d,e). Forskolin raises the level of cyclic AMP in the cell, therefore increasing CFTR activity. Consistent with the data obtained at the steady state, we measured an increase of 2.8-fold in organoids area following AsCas12a-crRNA+11 treatment, with no significant difference with the lentiviral delivery of wild-type CFTR cDNA (Fig. 24d,e).

In light of these results, we concluded that AsCas12a-crRNA+11 modifications of the 3272-26A>G defect in patient's organoids allow the repair of the intron 19 splicing defect, leading to full recovery of the endogenous CFTR protein function.

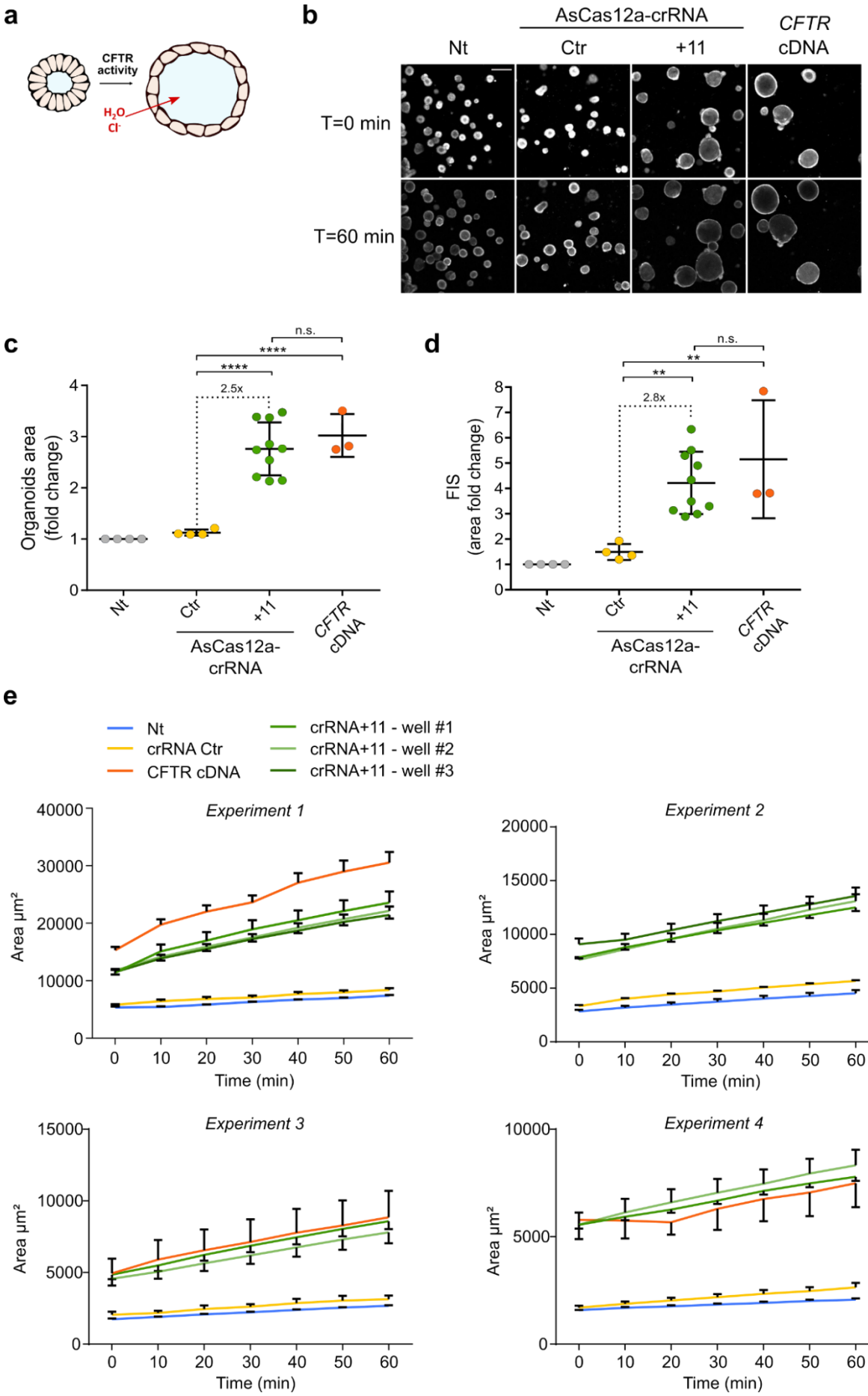

a
CFTR activity
H_2O
Cl^-
b
AsCas12a-crRNA
Nt
Ctr
+11
CFTR cDNA
T=0 min
T=60 min
c
Organoids area (fold change)
n.s.
****
****
2.5x
Nt
Ctr
+11
CFTR cDNA
AsCas12a-crRNA
d
FIS (area fold change)
n.s.
**
**
2.8x
Nt
Ctr
+11
CFTR cDNA
AsCas12a-crRNA
e
Nt
crRNA Ctr
CFTR cDNA
crRNA+11 - well #1
crRNA+11 - well #2
crRNA+11 - well #3
Experiment 1
Experiment 2
Experiment 3
Experiment 4
Area µm²
Time (min)

Figure 24. **CFTR activity is restored by AsCas12a-crRNA+11 editing**
a, Schematic representation of CFTR dependent swelling in primary intestinal organoids. **b**, Representative confocal images of calcein green labelled 3272-26A>G organoids before (T=0 min) and after (T=60 min) Forskolin Induced Swelling (FIS) assay (5 µM forskolin). Scale bar = 200 µm. **c**, Quantification of organoids area following lentiviral transduction of AsCas12a-crRNA Ctr, AsCas12a-crRNA+11 or with *CFTR* cDNA as indicated. Each dot represents the average organoid area analyzed for a single well (number of organoids per well: 25-300) from n=4 independent experiments. **d**, Fold change in organoid area before (T=0 min) and after (T=60 min) the FIS (Forskolin Induced Swelling) assay, each dot represents the average increase in organoid area analyzed per well (number of organoids per well: 25-300) from n=4 independent experiments. Data are means ± SD. Statistical analysis was performed using one-way ANOVA; **P<0.01, ****P<0.0001, n.s. non-significant. **e**, FIS assay of n=4 independent experiments; each line represents one well (n=25-300). Data are means ± SD.

SpCas9-based correction of 3849+10kbC>T mutation

Encouraged by the efficacy of the 3272-26A>G splicing correction of the developed technology, we then decided to explore this genome editing approach to repair another *CFTR* splicing mutation. *CFTR* 3849+10KbC>T mutation causes the inclusion of a new 84 bases cryptic exon coming from intron 22 (Fig. 25a). We generated minigene models containing exon 22, part of intron 22 and exon 23 (legacy name: exon 19, intron 19 and exon 20) that mimic either wild-type (pMG3849+10kbWT) or +84nt aberrant (pMG3849+10kbC>T) *CFTR* splicing (Fig. 25b,c). Sequencing analysis of the RT-PCR products was performed to assess correct splicing of the transcripts (Fig. 25d).
These minigene models produce also smaller splicing products. Sequencing analysis of these transcripts showed that in both constructs a new 5' splice site in exon 22 was activated, which determined the inclusion of only 49 bases of the exon into the mRNA (Fig. 25c).
As for the 3272-26A>G mutation, we initially designed a pool of sgRNA for SpCas9 nuclease targeting the region surrounding the 3849+10kbC>T mutation (Appendix 2). In order to evaluate whether the deletion of the mutation or indels nearby could correct the splicing, SpCas9 with single or pairs of sgRNAs were transiently delivered in HEK293T cells together with the pMG3849+10kbC>T. As observed with pMG3272-26A>G (Fig. R2a), analyses of the transcripts by RT-PCR showed that SpCas9 restored the correct splicing pattern if used in combination with two sgRNAs (Fig. 26a). Agarose gel of the PCR products from the edited samples revealed

efficient DNA deletion of the intronic portions containing the mutation (Fig. 26b). The observation of a very high splicing correction in the minigene model (100% of correct transcripts), prompted us to further validate all the sgRNA pairs in Caco-2 cells, carrying the wt *CFTR* gene. Of note, in this approach the two sgRNAs that were not specific for the mutation, thus cleaving the wt sequence of *CFTR*. Transcript

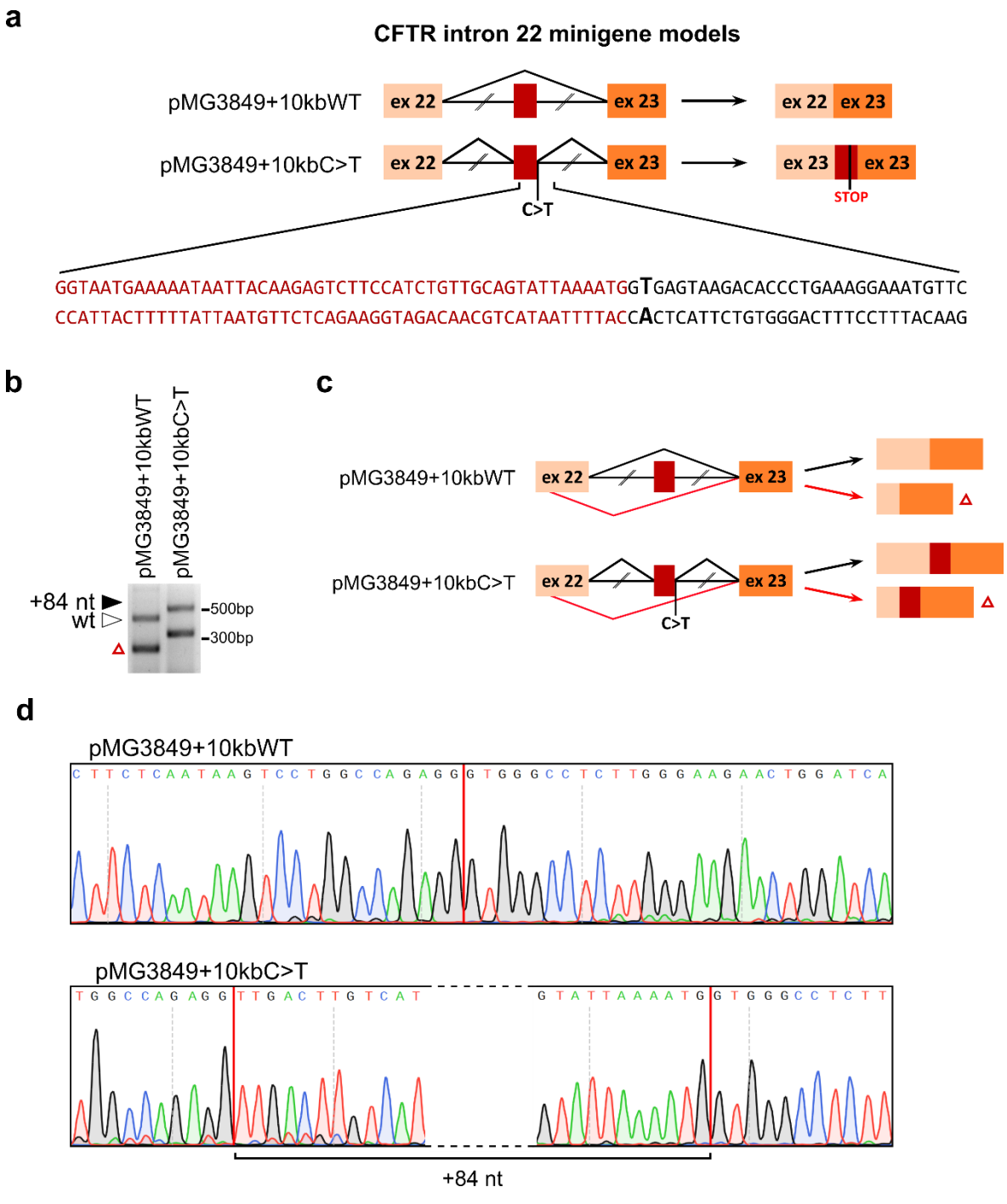


Figure 25. **Validation of intron 22 splicing in pMG3849+10kbWT and pMG3849+10kbC>T *CFTR* minigene models.**
a, Scheme of *CFTR* wild type (pMG3849+10kbWT) and 3849+10KbC>T (pMG3849+10kbC>T) minigenes carrying exon 22, portions of intron 22 encompassing the 3849+10KbC>T, and exon 23 of the *CFTR* gene. Exons are shown as boxes and introns as lines; the expected spliced transcripts are represented on the right according to the presence or absence of the 3849+10kbC>T mutation. The lower panel shows nucleotide sequence nearby 3849+10kbC>T mutation (labelled in bold) and the AsCas12a-crRNA+14 target position (underlined, with the PAM in red). **b**, Splicing pattern of *CFTR* wild type (pMG3849+10kbWT) and mutated (pMG3849+10kbC>T) minigene models, transfected in HEK293T, by agarose gel electrophoresis analysis of RT-PCR products. Black-solid arrow indicates aberrant splicing, white-empty arrow indicates correct splicing and red triangle indicates a minigene splicing artifact. **c**, Scheme of minigenes splicing products. The minigene splicing artifact was caused by the use of an alternative donor splice site in exon 22, causing the inclusion of only 49 bases of exon 22 into the mature transcripts of the minigene models. **d**, Sanger sequencing chromatogram of minigene splicing products from (**b**). Red lines represent the boundary between exons.

analyses in Fig. 26c showed that the correct splicing of intron 22 is not altered, whereas SpCas9 efficiently delete the target intronic sequences (Fig. 26d).

We identified the best performing sgRNA pair (+119/-95) to be tested in intestinal organoids, based on deletion efficiency in Caco-2 cells. As before organoids were

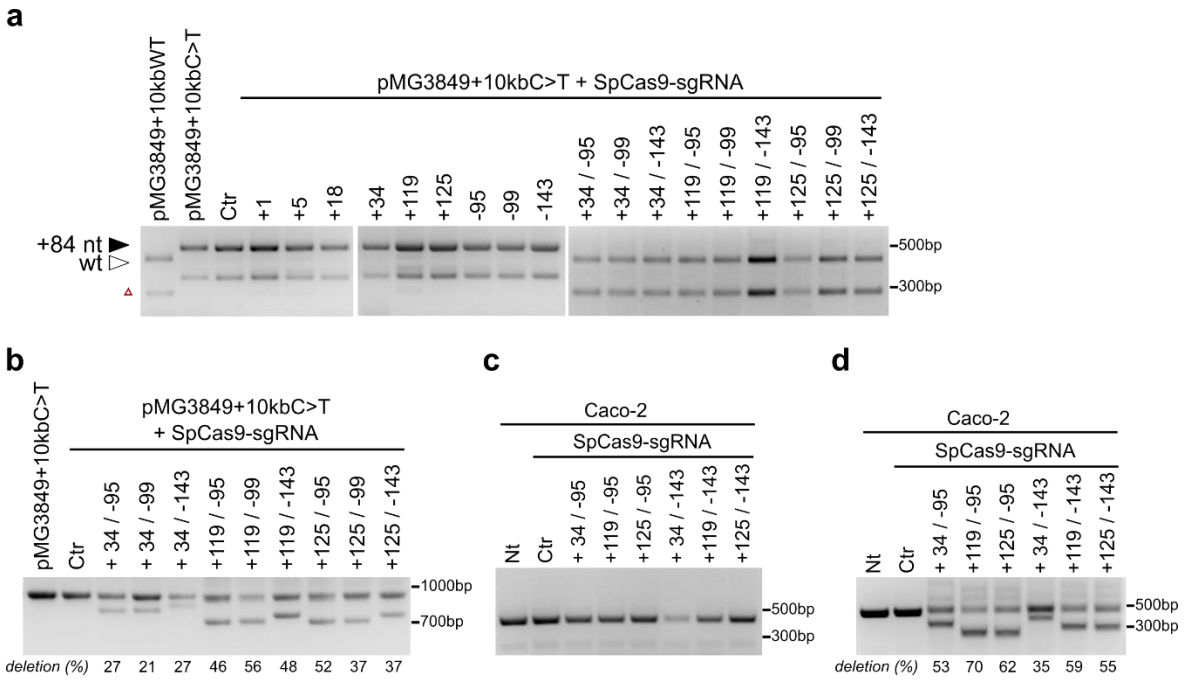
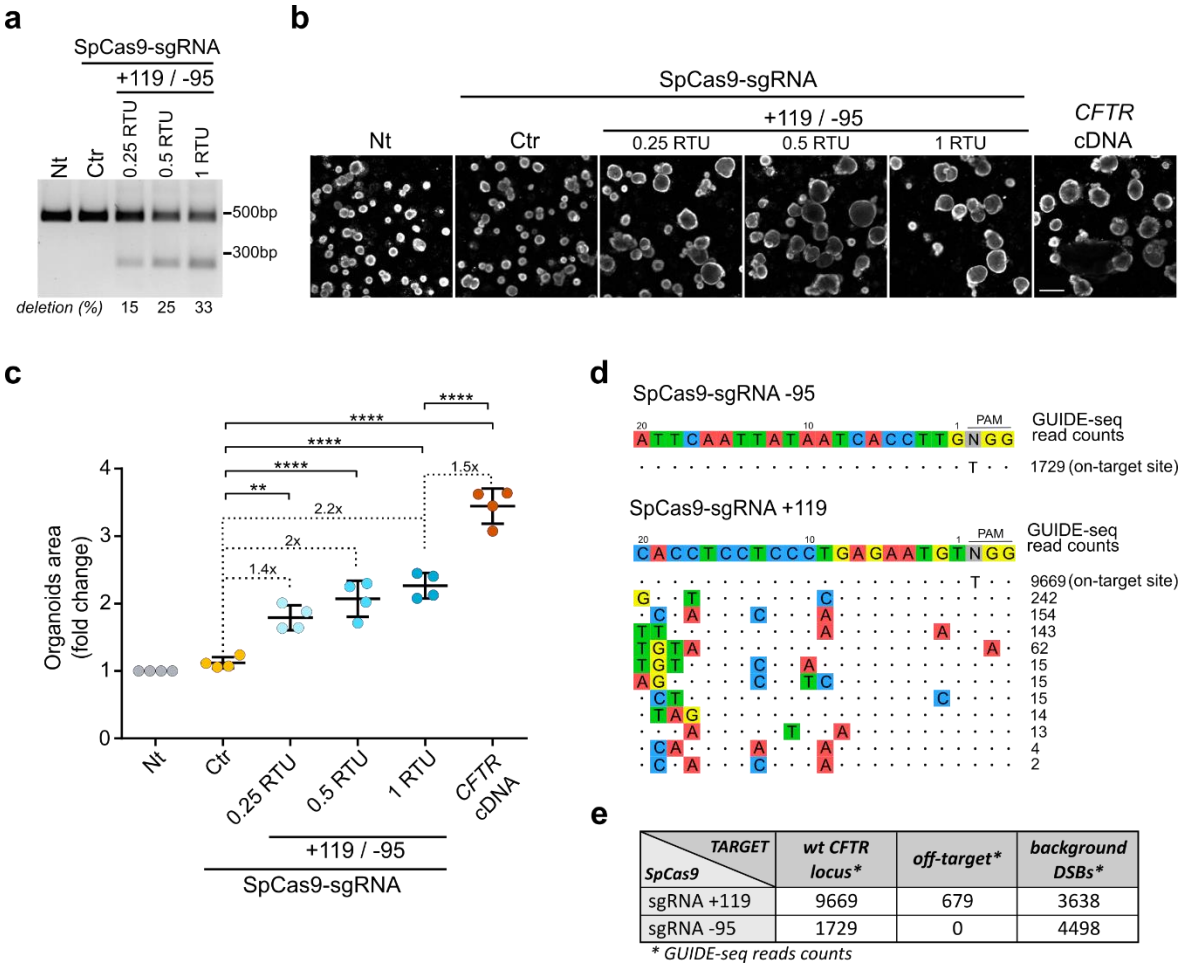


Figure 26. **SpCas9-sgRNA correction of 3849+10kb splicing defect in a minigene model**
a, Screening of SpCas9-sgRNA single or in pairs in pMG3849+10kbC>T transfected in HEK293T cells. RT-PCR products were analyzed by agarose gel electrophoresis. Black-solid arrow indicates aberrant splicing, white-empty arrow indicates correct splicing and red-arrow indicates a minigene splicing artifact. **b**, Agarose gel electrophoretic analysis of targeted DNA deletions in pMG3849+10kbC>T after cleavage of SpCas9-gRNA pairs. The bigger band represents non-edited minigene sequences, the smaller band is the expected deletion product. **c**, RT-PCR products and **d**, targeted deletions in Caco-2 cells transduced with a SpCas9-sgRNA lentiviral vectors and 10 days of puromycin selection.

derived from a compound heterozygous patient carrying the 3849+10kbC>T mutation (3849+10kbC>T/ΔF508).

Two different lentiviral vectors were delivered to the organoid culture, containing SpCas9 and one sgRNA each, using different quantities (0.25, 0.5 or 1 RTU). As expected, the deletion efficiency correlated with the amount of SpCas9-sgRNA+119/-95 (up to 33%, Fig. 27a) used in the experiments. The organoids area, which is indicative of CFTR channel function, strongly correlated with the efficiency of editing (Fig. 27b,c). We can notice that the effect on organoids area of SpCas9-sgRNA+119/-95 deletions is significantly reduced compared to the delivery of the wt CFTR cDNA (Fig. 27c), thus suggesting a lower efficacy than the one obtained with



SpCas9 \ TARGET	wt CFTR locus*	off-target*	background DSBs*
sgRNA +119	9669	679	3638
sgRNA -95	1729	0	4498

** GUIDE-seq reads counts*

Figure 27. **SpCas9-sgRNA editing restore CFTR function in 3849+10kb CF patient derived intestinal organoids.**

a, Editing in patient organoids analyzed by agarose gel electrophoresis. **b**, Confocal images of calcein green labelled CF 3849+10kbC>T organoids at T=0 min transduced with 0.25, 0.5 or 1 RTU of SpCas9-sgRNAs-95/+119. Scale bar = 200 µm. **c**, Quantification of steady-state organoid area; each dot represents the average area of organoids from one well (3-30 organoids per well) from n=1 experiment. Data are means ± SD. Statistical analysis was performed using one-way ANOVA; **P<0.01, ****P<0.0001. **d,e**, GUIDE-seq analysis of sgRNA-95 and sgRNA+119.

AsCas12a for the 3272-26A>G mutation. As mentioned before, sgRNAs +119 and -95 are not specific for the mutated sequence, producing deletions also in the second allele (carrying the ΔF508). Therefore, the final genome editing efficacy should be considered diluted over the two alleles. Only the deletions produced in the 3849+10kbC>T allele will lead to restoration of the correct splicing and production of a functional CFTR protein, thus correlating with the decreased effect of this dual-guides strategy.

Finally, we evaluated the genome wide off-target profile of the two sgRNA through the GUIDE-seq method. Although our sgRNA pool was designed in *silico to* minimize the probability of SpCas9 off-target activity[194], sgRNA+119 revealed 11 off-target sites throughout the genome (Fig. 27d,e).

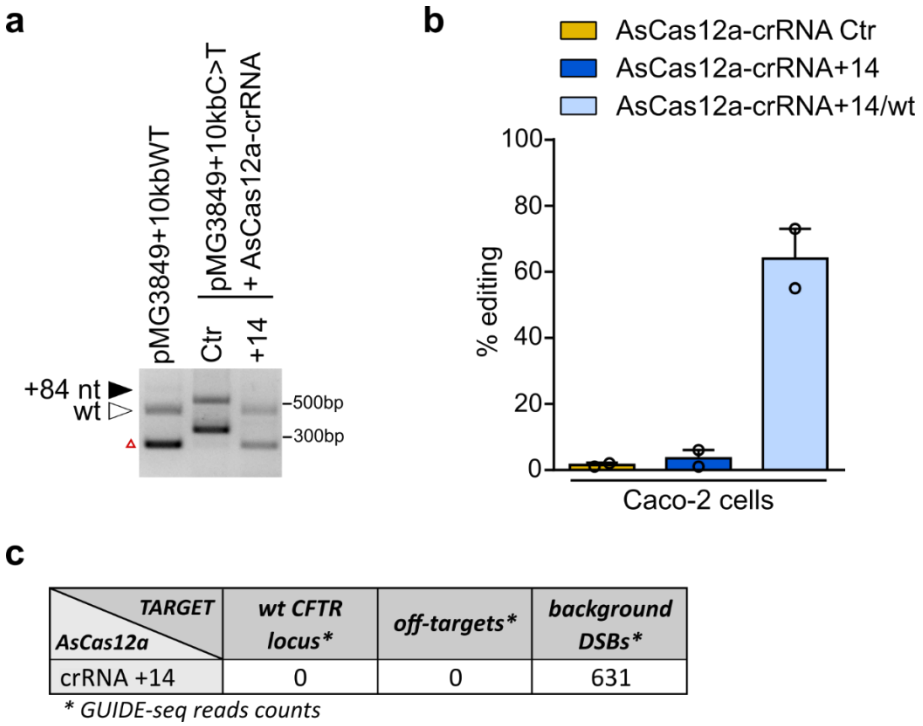
In conclusion, the intronic deletion produced by SpCas9-sgRNA+119/-95 restored the activity of the CFTR channel, while without reaching the level of the delivery of CFTR cDNA.

AsCas12a with a single crRNA efficiently correct 3849+10kbC>T mutation

Learning from the results obtained for both the CF mutations studied, we wanted to verify aberrant splicing correction of 3849+10kbC>T mutation by means of a single crRNA for AsCas12a. We designed the crRNA+14 that specifically target the mutated sequence and test its efficacy in splicing correction in the pMG3849+10kbC>T model. Transient delivery of the minigene and AsCas12a-crRNA+14 followed by analyses of the RNA products by RT-PCR showed a complete recovery of the correct transcripts (Fig. 28a).

We then decided to evaluate the specificity of AsCas12a-crRNA+14 for the 3849+10kbC>T mutation. Caco-2 cells were transduced with a lentiviral vector encoding for either AsCas12a-crRNA+14 or AsCas12a-crRNA+14/wt, targeting the wild-type sequence. Figure 28b showed that crRNA+14 does not modify wild-type CFTR gene in Caco-2 cells, generating indels near background levels (3.5%), while crRNA+14/wt produced 64% *CFTR* editing, thus indicating specificity of the AsCas12a-crRNA+14 towards the mutated allele.

To further verify AsCas12a-crRNA+14 specificity in terms of genome-wide off-target activity, GUIDE-seq analysis was performed in HEK293T cells, showing the



c

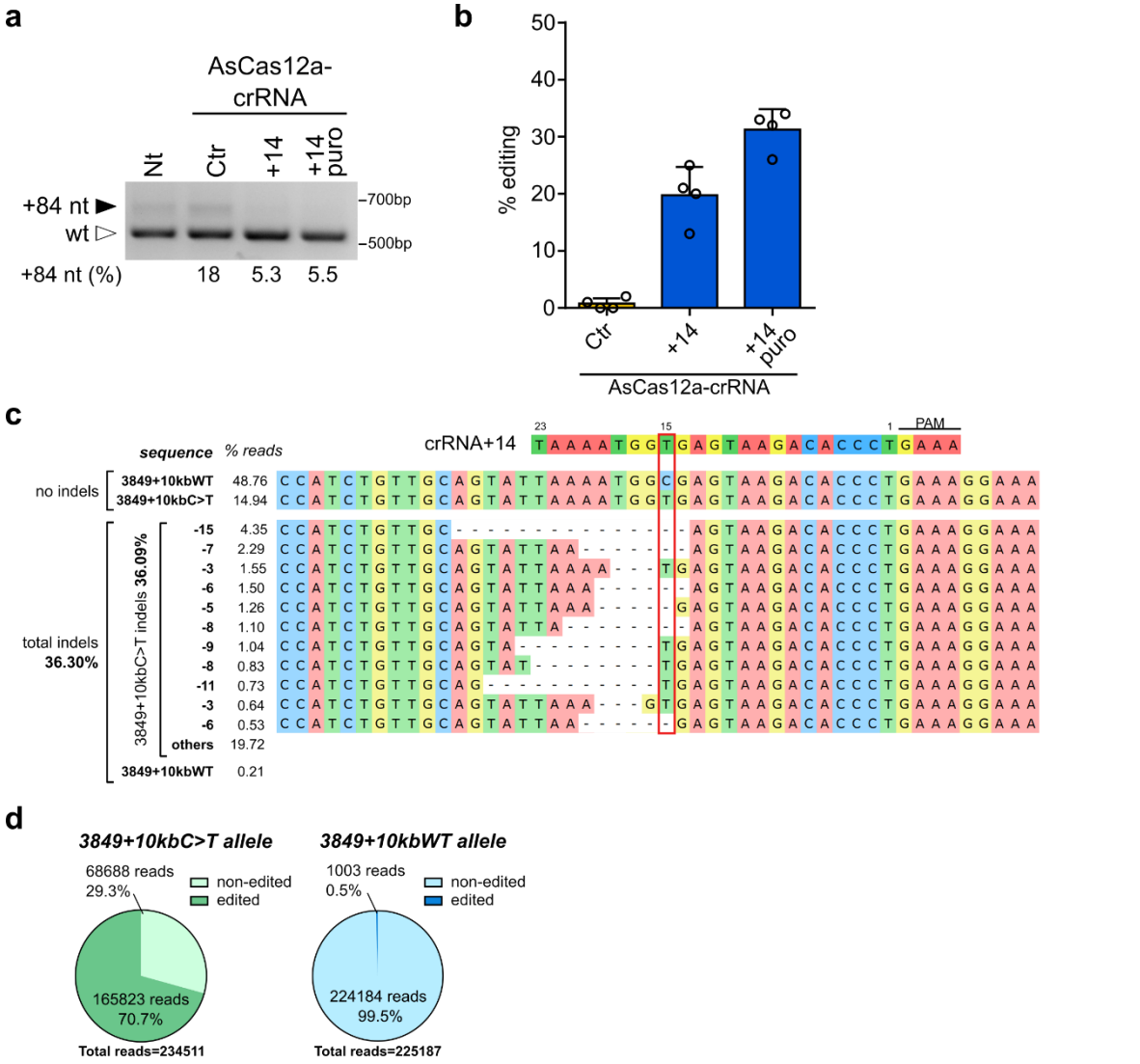
TARGET / AsCas12a	wt CFTR locus*	off-targets*	background DSBs*
crRNA +14	0	0	631

* GUIDE-seq reads counts

Figure 28. **AsCas12a-based correction for 3849+10kbC>T mutation.**
a, Splicing pattern analyzed by RT-PCR in HEK293/pMG3849+10kbC>T cells following treatments with AsCas12a-crRNA control (Ctr) or specific for the 3272-26A>G mutation (+14). Black-solid arrow indicates aberrant splicing, white-empty arrow indicates correct splicing and red triangle indicates a minigene splicing artifact. **b**, Caco-2 cells lentivirally transduced with AsCas12a-crRNA+14 or +14/wt were analyzed for editing in *CFTR* intron 22 by SYNTHEGO ICEanalysis. Data are means ± SEM from n=2 independent experiments.
c, GUIDE-seq analysis for AsCpf1-crRNA+14. The double strand breaks (DSBs) are spontaneous cleavages grouped by the analysis as spurious cleavages (not resulting from AsCas12a or SpCas9 cleavages) that are nevertheless indicative of properly executed GUIDE-seq assay.

complete absence of sequence reads in the *CFTR* locus or in any other off-target site; the 631 sequencing reads corresponding to spontaneous DNA breaks are indicative of the proper execution of the GUIDE-seq assay[179] (Fig. 28c).

In view of these results, the efficacy of this crRNA was analysed in primary airway epithelial cells derived from a compound heterozygous patient carrying the 3849+10kbC>T and the ΔF508 mutations. Transcripts analyses of intron 22 splicing revealed two different patterns (Fig. 28a), the wild-type one and the +84nt transcripts produced from the 3849+10kbC>T allele. As for 3272-26A>G mutation, some correct transcripts are produced by the 3849+10kbC>T allele, reflecting the abundance measured in the RT-PCR analyses and previous data[28].



e

AsCas12a-crRNA+14		Chromosome	Position	Mismatches	% editing crRNA+14	total reads crRNA+14	% editing crRNA CTR	total reads crRNA CTR
3849+10kbC>T*	*TTTC*AGGGTGTCTTACTCACCATTTTA	chr7	117639952	0	70.7	234511	0.3	232560
3849+10kbWT	*TTTC*AGGGTGTCTTACTC**g**CCATTTTA	chr7	117639952	1	0.5	225187	0.3	242654
OFF 1	*TTTG*AGGGTGTCTTA**tg**CAtCA**a**TTTA	chr4	94977071	4	1.0	485018	0.0	713721
OFF 2	*TTTC*AG**a**GTGTCTTACTCAC**t**ATTT**ct**	chr13	35225035	4	1.0	334723	0.6	448459
OFF 3	*TTTC*AG**a**GT**t**T**ta**TACTCACCATTTTA	chr6	152271172	4	0.5	474836	0.3	564122

*on-target site

Figure 29. **AsCas12a-crRNA+14 splicing correction and editing efficiency in primary airway cells.**

***a**, Splicing pattern analyzed by RT-PCR in 3849+10kbC>T primary airway cells following lentiviral transduction (15 days) of AsCas12a-crRNA control (Ctr) or specific for the 3849+10kbC>T mutation (+14). Puromycin selection was performed for 72h in +14 puro. The percentage of aberrant splicing (84 nt insertion into mRNA) was measured by densitometric analysis. **b**, Percentages of indels in 3849+10kbC>T primary airway cells measured by TIDE analysis following lentiviral transduction as in (**a**). Data are from n=2 independent experiments. **c**, Deep sequencing analysis of the CFTR on-target locus after AsCas12a-crRNA+14 transduction (15 days) of the 3849+10kbC>T primary airway cells. **d**, Percentage of deep sequencing reads of the edited and non-edited 3849+10kbC>T or WT alleles in primary airway cells treated as in (**a**). **e**, off-target analysis of in silico predicted sites. All the predicted sites with 4 or less mismatches (3 sites) were analyzed in 3849+10kbC>T primary airway cells treated as in (**a**).*

AsCas12a-crRNA+14 was delivered through lentiviral vectors and cells were either treated or not with puromycin. The aberrant splicing pattern was not detected after treatment, as demonstrated by the disappearance of the upper band in RT-PCR analyses. The amount of mutated splicing in each sample was evaluated through densitometric analysis resulting in 18% of +84 nt isoform in the control sample, 5.3% with crRNA+14 and 5.5% with crRNA+14 puro (Fig. R29a). Indels formation calculated by TIDE analysis reveals 20% of efficiency (up to 30% after puromycin selection, Fig. 29b), correlating with the level of splicing correction.

The allelic discrimination, previously observed in Caco-2 cells, was further verified by targeted deep sequencing analyses in primary airway cells. The overall efficiency of AsCas12a-crRNA+14 was about 36.30%, with all indels located on the target mutated allele and fully preserved second *CFTR* allele, 3849+10kbWT (70.7% editing of the 3849+10kbC>T mutated allele, Fig. 29c,d).

The off-target profile of AsCas12a-crRNA+14 was evaluated after long term expression (14 days) of AsCas12a in airway cells by targeted deep sequencing analysis. In silico off-target prediction by Cas-OFFinder reveals 3 possible sites (considering sequences up to 4 mismatches). Data in Figure 29e. confirm the results of the GUIDE-seq analysis in HEK293T cells (Fig. 29c), showing no cleavage of the predicted off-target loci.

Similarly to crRNA+11, the editing of intron 22 by crRNA+14 produced deletions of variable length and the mutated allele was either deleted or not (Fig. 29c). To evaluate the effect of indels on the 3849+10kbC>T mutation, in silico analysis by

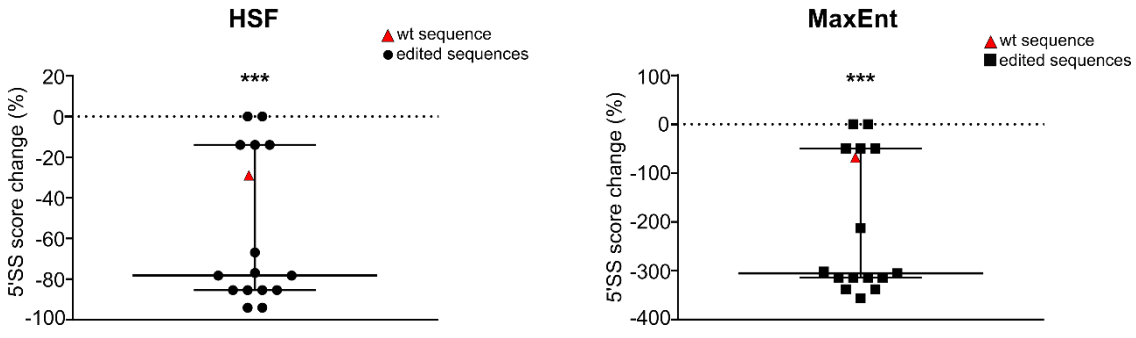


Figure 30. **Figure R18. In silico splicing score prediction for 3849+10kbC>T.**
Prediction of the score change of AsCas12a-crRNA+14 modification on the mutated 5'splice site (ss) by bioinformatics tools (HSF and MaxEnt prediction algorithms, see Methods section). Data are median with interquartile range. Statistical analysis was performed using two-tailed Wilcoxon signed-rank test; ***P<0.001. Score of wild type sequence was given only as reference in the graph.

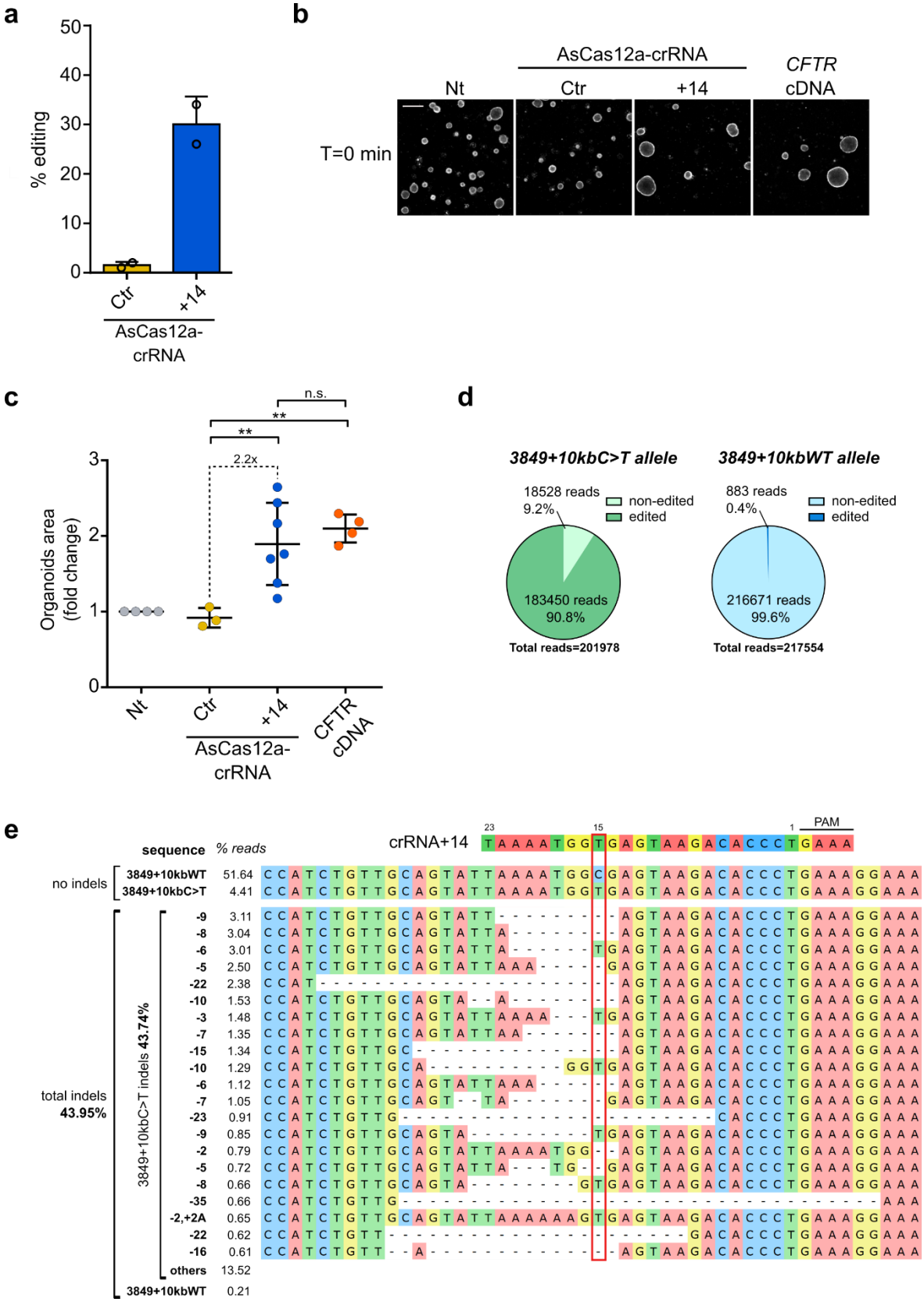
HSF and MaxEnt prediction algorithms was performed[181,182], showing that edited sequences probably decrease the strength of the aberrant splice site, due to deletion of regulatory regions or deletion of the mutation itself (Fig. 30).

Finally, to evaluate the effect of splicing correction measured on CFTR channel activity, we transduced compound heterozygous organoids (3849+10kbC>T/ ΔF508) with AsCas12a-crRNA+14 lentiviral vector. We obtained high percentages of indels, 30%, at the target locus (Fig. R19a) that correlated with the restoration of CFTR function, as measured by the 2.2-fold increase in organoids area (Fig. R19b,c). The functional recovery of CFTR by AsCas12a-crRNA+14 was similar to what observed with wt CFTR cDNA (Fig. R19c). The allelic discrimination of crRNA+14 was confirmed by deep sequencing analysis in intestinal organoids, obtaining an editing efficiency of 43.74% of the 3849+10kbC>T sequence (90.8% if we consider only the mutated allele) and no editing of the other allele (Fig. R19d,e).

In conclusion, similarly to the splicing repair of the 3272-26A>G variant, the correction of the CFTR 3849+10kbC>T splicing defect was efficiently and precisely obtained by using AsCas12a combined with a single allele specific crRNA in different CF models derived from patients, such as primary airway cells and intestinal organoids. As observed with the 3272-26A>G mutation, this strategy was proven superior to the conventional SpCas9 induced genetic deletion obtained using sgRNA pairs.

Figure 31. **CFTR activity in intestinal organoids is restored by AsCas12a-crRNA+14**

a, 3849+10Kb C>T patient derived intestinal organoids were lentivirally transduced with AsCas12a-crRNA control (Ctr) or crRNA+14 and analyzed for intron 22 editing by SYNTHEGO ICE. Data are from n=1 experiment. **b**, Confocal images of calcein green labelled 3849+10KbC>T organoids transduced with AsCas12a-crRNA+14 or *CFTR* cDNA. Scale bar 200 µm. **c**, Quantification of organoids area; each dot represents the average area of organoids analyzed in each well (number of organoids per well: 3-30) from n=1 experiment. Data are means ± SD. Statistical analysis was performed using one-way ANOVA; **P<0.01, n.s. non-significant. **d**, Percentage of deep sequencing reads of the edited and non-edited 3849+10kbC>T or WT alleles from (**a**). **e**, Deep sequencing analysis of the *CFTR* on-target locus after AsCas12a-crRNA+14 transduction (15 days) of the 3849+10kbC>T organoids.


a
% editing
50
40
30
20
10
0
Ctr
+14
AsCas12a-crRNA
b
AsCas12a-crRNA
Nt
Ctr
+14
CFTR cDNA
T=0 min
c
n.s.
**
**
2.2x
Organoids area (fold change)
3
2
1
0
Nt
Ctr
+14
CFTR cDNA
AsCas12a-crRNA
d
3849+10kbC>T allele
18528 reads
9.2%
non-edited
edited
183450 reads
90.8%
Total reads=201978
3849+10kbWT allele
883 reads
0.4%
non-edited
edited
216671 reads
99.6%
Total reads=217554
e
sequence
% reads
crRNA+14
23
15
1
PAM
T A A A A T G G T G A G T A A G A C A C C C T G A A A
no indels
3849+10kbWT 51.64
3849+10kbC>T 4.41
total indels 43.95%
3849+10kbC>T indels 43.74%
-9 3.11
-8 3.04
-6 3.01
-5 2.50
-22 2.38
-10 1.53
-3 1.48
-7 1.35
-15 1.34
-10 1.29
-6 1.12
-7 1.05
-23 0.91
-9 0.85
-2 0.79
-5 0.72
-8 0.66
-35 0.66
-2,+2A 0.65
-22 0.62
-16 0.61
others 13.52
3849+10kbWT 0.21

DISCUSSION

Since the discovery of the *CFTR* gene sequence several gene therapy approaches have been developed to achieve a cure for cystic fibrosis. The most advanced treatment available rely on the delivery of the wild type *CFTR* cDNA to the airway epithelia to compensate the defective endogenous *CFTR* gene[72]. A vast number of clinical trials has been performed exploiting the use of different vectors, aimed at evaluating the delivery efficiency in lung tissue and transgene expression[50]. Unfortunately, the clinical benefit observed in CF patients was reduced to the stabilization of lung function due to inefficient transduction and loss of the therapeutic gene due to the rapid turnover of pulmonary epithelial cells[195].

The main obstacle to CF gene therapy approaches are CF symptoms themselves. Indeed, the thick mucus layer, the infection and inflammation status of the lung represent the major obstacle for delivering therapeutic agents to the airways of CF patients[12,196]. Among the viral vector tested for CF, AAV have been extensively studied, bringing to the identification of different serotypes able to efficiently transduce the lung epithelia[58,60,197]. Nevertheless, they produce a transient expression of the transgene necessitating multiple treatments, while re-administration will result in immune response against the vector[57]. An attractive alternative is represented by lentiviral vectors allowing the transduction of both dividing and non-dividing cells and long-term expression of the carried transgene. Moreover, the possibility to pseudotype the envelope makes them suitable to achieve specific tropism for cells and tissues[63,65]. In order to achieve an efficient transduction of lung epithelia, lentiviral vectors have been pseudotype with Sendai virus (SeV) envelope[66,68]. Pre-clinical studies in animal models revealed that in mouse models gene expression is maintained for the life of the animal[198–200] and that re-administration of lentiviral vectors do not cause an immune response in the airway epithelia[198,201]. These data together with tests in CF animal models[69,70] demonstrated an encouraging efficacy of this delivery system for the treatment of CF, leading to the preparation of a first-in-man lentivirus trial in patients with CF[66]. In addition, lentiviral vectors recently reached the clinic for the treatment of a severe combined immunodeficiency[202].

The discovery of CRISPR nuclease as a genome editing tool produced an incomparable excitement in the gene therapy field, offering the opportunity to

permanently repair the endogenous mutated *CFTR* locus, thus restoring the expression of the CFTR channel under physiological conditions. Genetic correction approaches by CRISPR systems don't need long-term expression of the nuclease, that need to be active only the time necessary to modify the target DNA. This strategy avoids the problem of the continuous expression of the transgene required in cDNA addiction therapies. On the contrary, several methods have been exploited to avoid long term expression of Cas protein into the target cells after lentiviral delivery, considering long term expression either as referred in this thesis work or considering possible clinical applications. These systems account for natural promoter silencing[203,204] and self-limiting circuits for the inactivation of CRISPR-Cas activity[104].

Nevertheless, to achieve a permanent and lasting correction, the targets for gene therapy need to be cells with progenitor capacity. It has been proved that lentiviral vectors are suitable for transduction of some progenitor cells of the airway epithelia, both in vitro and in vivo[66,205]. Moreover, recent studies demonstrated the differential expression of CFTR in different cell types present in the airways[206–208], demonstrating the need of further knowledge to better understand their role for CFTR correction[208].

Another crucial aspect is also the number of transduced cells needed to obtain a therapeutically relevant correction of the CFTR channel. Some studies reported that restoring channel activity in 10 to 50% of cells produce a chloride transport comparable to wild-type levels[209–213]. Therefore, the expression of CFTR in only 10% of airway epithelia may be therapeutically relevant for CF treatment[65].

Recently, an ex-vivo therapy approach based on CRISPR technology has been proposed, setting the basis for transplantation of autologous CFTR-corrected cells into CF patients[76]. This strategy will offer several advantages, such as, the possibility to select corrected cells, to control side effects and to transiently express Cas9 protein. The possibility of an ex vivo therapy for CF is further supported by the demonstration that human iPSCs derived from lung epithelia can engraft in a murine lung injury model[214]. Nevertheless, some challenges need still to be faced, for example the derivation of a pure population, sufficient in number, of airway cells to be engrafted, the possible accumulation of somatic mutations due to the long culture time needed for modification and expansion, and finally yet importantly the engraftment into lung epithelia[196].

CRISPR-SpCas9 nucleases has already been employed for the correction of the ΔF508 mutations by homology directed repair (HDR) in intestinal organoids[74] and upper-airway basal stem cells (UABCs) from CF patients[76], with different degree of correction. Nevertheless, HDR is very inefficient in differentiated cells and the need to additionally deliver to the tissues a donor template to correct the endogenous gene pose several hurdles for a future clinical translation.

However, it has been shown in minigene models that CRISPR-SpCas9 can be used in combination with a dual sgRNAs system to correct splicing mutations by deletion of big intronic portions (around 150-200 bp) containing the mutation[215].

In our study, we proposed an innovative genome editing approach, based on the use of AsCas12a, to permanently correct two of the most frequent CF-causing splicing mutations, the 3272-26A>G and the 3849+10kbC>T. A major advantage of the developed strategy is the use of a single crRNA that, by creating small deletions (about 4-26 nucleotides) in the intronic sequences near the mutations, efficiently correct the aberrant splicing. In particular, for the 3272-26A>G mutation sequencing analyses of the edited locus showed the persistence of the mutated nucleotides (G) into the majority of the isolated repair events. These results suggested that AsCas12a-crRNA+11 possibly modify regulatory sequences necessary to the aberrant 3' splice site, such as the branch point or the polypyrimidine tract, rather than deletion of the splice site itself (compare Fig. 13a and 16b). This hypothesis is supported by the *in silico* analysis of the most frequent editing events, correlating these sequences with a reduced strength of the 3272-26A>G splice site, while preserving the mutation. On the contrary, from sequences analyses of the 3849+10kbC>T edited sites we inferred that, while the mutation was either deleted or not, crRNA+14 destroyed the consensus sequence that define the aberrant 5' splice sites (in this case TG/GT, where the last nucleotide represents the mutation, see Fig. 25a and 29c). Similarly, *in silico* analyses of the most frequent repair events showed a decrease strength of the mutated splice site, reinforcing our hypothesis.

Moreover, we compare AsCas12a nuclease with the most frequently used SpCas9 nuclease. We operated a screening of sgRNAs, used as a single or as pair, that demonstrated that SpCas9 can restore the correct splicing pattern in minigene models only if used in combination with two sgRNAs. Although for the 3272-26A>G the use of SpCas9 was subsequently excluded due to the inefficient cleavage of the endogenous locus (integrated minigene), in the case of the 3849+10kbC>T defect

we identified a pair of sgRNAs that, by performing a deletion of the intronic region surrounding the mutation, corrected the splicing pattern also in relevant CF cell models. Nevertheless, this approach applied to intestinal organoids could not reach the degree of functional correction obtained with the gold standard gene therapy strategy (delivery of the CFTR cDNA, Fig. 27). The disadvantages of this dual-guide system are the need of co-deliver two different vectors containing one sgRNAs each and the need of concomitant cleavage by the sgRNA pair to produce the desired intronic deletion. In addition, the sgRNAs selected are not specific for the mutated sequence, thus able to cleave both the alleles. Only the deletions of the 3849+10kbC>T mutation will produce correct spliced transcripts and a functional CFTR protein, further explaining the decreased efficacy of this system.

Off-targets are one of the major concerns regarding the clinical applications of CRISPR technology. Unwanted cleavages can result in genomic modification (either point mutation or chromosomal translocation) that can alter cell functions and viability. AsCas12a is considered one of the most precise naturally occurring CRISPR nuclease[111]. Indeed, genome wide off-target analyses demonstrated the high fidelity of this nuclease, whereas SpCas9 showed to be more prone to off-target site cleavage, at least for one of the two sgRNA studied (compare Fig. 18e,f; 27d,e; R16c). The specificity of AsCas12a was confirmed by targeted deep sequencing analyses of *in silico* predicted off-target sites in primary airway cells after long term expression, in reference to experimental timing (14 days, Fig. 21c; 29e), usually associated with an increased off-target activity. Of note, all the identified off target sites were located outside coding sequences, reducing the chance of possible side effects caused by unwanted editing. Furthermore, we design two crRNAs (+11 and +14) whose target sequence include the mutated nucleotides. This characteristic allowed us to obtain allelic discrimination in both the mutations (Fig. 21a,b; 23b,c; 29c,d; 31d,e), providing the relevant advantage of minimizing the potential risk of on-target chromosomal translocations. Despite the absence of detectable off-target cleavages in the cell models used, AsCas12a-crRNA+11 and +14 fidelity should be thoroughly evaluated while moving to the clinical application.

The efficacy of the proposed gene correction strategy was confirmed in primary airway epithelial cells and intestinal organoids derived from patients compound heterozygous for the mutations. These models are considered valuable for preclinical testing of new therapies for translational research. The efficient editing

and splicing correction obtained in primary airway cells correlated with the amount of restored CFTR channel activity measured in intestinal organoids (Fig. 24 and 31). These 3-dimensional structures recapitulate the tissue feature of the parental organ and are very used to evaluate the efficacy of specific therapies for different mutations, especially for patients with rare mutations that are too low in number for proper clinical trials, establishing a personalized medicine approach. The possibility to predict ex vivo the therapeutic efficacy in CF patients of a treatment provide an important milestone for future clinical trials for the proposed strategy.

Even though CFTR modulators therapy has been approved also for splicing mutations, including 3272-26A>G and 3849+10kbC>T (Kalydeco from Vertex therapeutic), life-long drug uptake is arduous for patients, due to tolerability, side effects and the elevate cost of the treatment. Moreover, in the case of splicing mutations, the efficacy of modulators relies on the level of residual correct spliced transcripts produced by the mutated *CFTR*, which is extremely reduced and variable among patients. This explains the urgent need of new strategies aimed at permanently correcting the genetic defects at the endogenous level, restoring the correct functionality of the CFTR channel in its physiological environment. The strategy presented in this study could possibly represent a valuable genome editing strategy to correct at least two genetic defects that cause CF.

In conclusion, this work set a robust proof of concept for a CF gene therapy approach. We demonstrated the correction of two different deep intronic mutations by sequence modification through AsCas12a nuclease gene editing, which corrected the splicing pattern and restored CFTR function in relevant disease models. Furthermore, this approach can be broadened to other splicing defects in CF and even to other genetic disease caused by splicing alterations.

BIBLIOGRAPHY

1. Maule, G. *et al.* Allele specific repair of splicing mutations in cystic fibrosis through AsCas12a genome editing. *Nat. Commun.* **10**, (2019).
2. Kerem, B. *et al.* Identification of the Cystic Fibrosis Gene: Genetic Analysis. *Science (80-.).* **121**, (1989).
3. Rommens, J. M. *et al.* Identification of the Cystic Fibrosis Gene: Chromosome Walking and Jumping. *Science (80-.).* **1059**, 1–9 (1989).
4. Mehta, G., Macek, M. & Mehta, A. Cystic fibrosis across Europe: EuroCareCF analysis of demographic data from 35 countries. *J. Cyst. Fibros.* **9**, S5–S21 (2010).
5. Spielberg, D. R. & Clancy, J. P. Cystic Fibrosis and Its Management Through Established and Emerging Therapies. *Annu. Rev. Genomics Hum. Genet.* **17**, 155–175 (2016).
6. Mention, K., Santos, L. & Harrison, P. T. Gene and base editing as a therapeutic option for cystic fibrosis-learning from other diseases. *Genes (Basel).* **10**, 1–17 (2019).
7. Linsdell, P. Architecture and functional properties of the CFTR channel pore. *Cell. Mol. Life Sci.* **74**, 67–83 (2017).
8. Virant-Young, D. *et al.* Cystic Fibrosis: A Novel Pharmacologic Approach to Cystic Fibrosis Transmembrane Regulator Modulation Therapy. *J. Am. Osteopath. Assoc.* **115**, 546 (2015).
9. Stoltz, D. A., Meyerholz, D. K. & Welsh, M. J. Origins of Cystic Fibrosis Lung Disease. *N Engl J Med* **372**, 351–362 (2015).
10. Gentzsch, M. *et al.* The cystic fibrosis transmembrane conductance regulator impedes proteolytic stimulation of the epithelial Na+ channel. *J. Biol. Chem.* **285**, 32227–32232 (2010).
11. Lopes-Pacheco, M. CFTR modulators: Shedding light on precision medicine for cystic fibrosis. *Front. Pharmacol.* **7**, 1–20 (2016).
12. Turcios, N. L. Cystic Fibrosis Lung Disease: An Overview. *Respir. Care* **0**, 1–19 (2019).
13. Gadsby, D. C., Vergani, P. & Csanády, L. The ABC protein turned chloride channel whose failure causes cystic fibrosis. *Nature* **440**, 477–483 (2006).
14. Liu, F., Zhang, Z., Csanády, L., Gadsby, D. C. & Chen, J. Molecular Structure of the Human CFTR Ion Channel. *Cell* **169**, 85–92 (2017).
15. Hwang, T. C. *et al.* Structural mechanisms of CFTR function and dysfunction. *J. Gen. Physiol.* **150**, 539–570 (2018).
16. Cystic Fibrosis Mutation Database. Available at: http://www. genet.sickkids.on.ca/. Accessed December 2019.
17. CFTR2. Available at: https://www.cftr2.org/. Assessed January 2020.
18. Pranke, I., Golec, A., Hinzpeter, A., Edelman, A. & Sermet-Gaudelus, I. Emerging therapeutic approaches for cystic fibrosis. From gene editing to personalized medicine. *Front. Pharmacol.* **10**, 1–21 (2019).
19. Cuyx, S. & De Boeck, K. Treating the underlying CFTR defect in patients with cystic fibrosis. *Semin. Respir. Crit. Care Med.* **40**, in press (2019).
20. Popp, M. W. & Maquat, L. E. Leveraging rules of nonsense-mediated mRNA decay for genome engineering and personalized medicine. *Cell* **165**, 1319–1322 (2016).
21. Patient Registry Annual ECFS. Data Report 2016. 2016. Available at: https://www.ecfs.eu/sites/default/files/general-content-images/ working-groups/ecfs-patient-registry/ECFSPR_Report2016_06062 018.pdf. Accessed December 2019.
22. De Boeck, K. & Amaral, M. D. Progress in therapies for cystic fibrosis. *Lancet Respir. Med.* **4**, 662–674 (2016).
23. Veit, G. *et al.* From CFTR biology toward combinatorial pharmacotherapy: Expanded classification of cystic fibrosis mutations. *Mol. Biol. Cell* **27**, 424–433 (2016).

24. Quintana-Gallego, E., Delgado-Pecellín, I. & Calero Acuña, C. CFTR Protein Repair Therapy in Cystic Fibrosis. *Arch. Bronconeumol.* **50**, 146–150 (2014).
25. Van Willigen, M. *et al.* Folding-function relationship of the most common cystic fibrosis-causing CFTR conductance mutants. *Life Sci. Alliance* **2**, 1–14 (2019).
26. Beck, S. *et al.* Cystic fibrosis patients with the 3272-26A→G mutation have mild disease, leaky alternative mRNA splicing, and CFTR protein at the cell membrane. *Hum. Mutat.* **14**, 133–144 (1999).
27. Amaral, M. D. *et al.* Cystic fibrosis patients with the 3272-26A>G splicing mutation have milder disease than F508del homozygotes: a large European study. *J. Med. Genet.* **38**, 777–83 (2001).
28. Ramalho, A. S. *et al.* Five percent of normal cystic fibrosis transmembrane conductance regulator mRNA ameliorates the severity of pulmonary disease in cystic fibrosis. *Am. J. Respir. Cell Mol. Biol.* **27**, 619–627 (2002).
29. Nissim-Rafinia, M., Chiba-Falek, O., Sharon, G., Boss, A. & Kerem, B. Cellular and viral splicing factors can modify the splicing pattern of CFTR transcripts carrying splicing mutations. *Hum. Mol. Genet.* **9**, 1771–1778 (2000).
30. Friedman, K. J. *et al.* Correction of aberrant splicing of the cystic fibrosis transmembrane conductance regulator (CFTR) gene by antisense oligonucleotides. *J.Biol.Chem.* **274**, 36193–36199 (1999).
31. Elborn, J. S. Personalised medicine for cystic fibrosis: Treating the basic defect. *Eur. Respir. Rev.* **22**, 3–5 (2013).
32. Goor, F. Van *et al.* Correction of the F508del-CFTR protein processing defect in vitro by the investigational drug VX-809. *PNAS* **7**, 159–163 (2011).
33. Quon, B. S. & Rowe, S. M. New and emerging targeted therapies for cystic fibrosis. *BMJ* **352**, 1–14 (2016).
34. Du, M. *et al.* Aminoglycoside suppression of a premature stop mutation in a Cftr-/- mouse carrying a human CFTR-G542X transgene. *J. Mol. Med.* **80**, 595–604 (2002).
35. Zomer-van Ommen, D. D. *et al.* Limited premature termination codon suppression by read-through agents in cystic fibrosis intestinal organoids. *J. Cyst. Fibros.* **15**, 158–162 (2016).
36. Kerem, E. *et al.* Ataluren for the treatment of nonsense-mutation cystic fibrosis: A randomised, double-blind, placebo-controlled phase 3 trial. *Lancet Respir. Med.* **2**, 539–547 (2014).
37. Sinha, C. *et al.* Capturing the Direct Binding of CFTR Correctors to CFTR by Using Click Chemistry. *ChemBioChem* **16**, 2017–2022 (2015).
38. Rapino, D. *et al.* Rescue of NBD2 mutants N1303K and S1235R of CFTR by small-molecule correctors and transcomplementation. *PLoS One* **10**, 1–17 (2015).
39. Mijnders, M., Kleizen, B. & Braakman, I. Correcting CFTR folding defects by small-molecule correctors to cure cystic fibrosis. *Curr. Opin. Pharmacol.* **34**, 83–90 (2017).
40. Hoy, S. M. Elexacaftor/Ivacaftor/Tezacaftor: First Approval. *Drugs* **79**, 2001–2007 (2019).
41. Boyle, M. P. *et al.* A CFTR corrector (lumacaftor) and a CFTR potentiator (ivacaftor) for treatment of patients with cystic fibrosis who have a phe508del CFTR mutation: A phase 2 randomised controlled trial. *Lancet Respir. Med.* **2**, 527–538 (2014).
42. Lopes-Pacheco, M. *et al.* Combination of correctors rescue AF508-CFTR by reducing its association with Hsp40 and Hsp27. *J. Biol. Chem.* **290**, 25636–25645 (2015).
43. Bose, S. J. *et al.* Towards next generation therapies for cystic fibrosis: Folding, function and pharmacology of CFTR. *J. Cyst. Fibros.* (2020). doi:10.1016/j.jcf.2019.12.009
44. Phuan, P. W. *et al.* Potentiators of defective DF508-CFTR gating that do not interfere with corrector action. *Mol. Pharmacol.* **88**, 791–799 (2015).
45. Donaldson, S. H. *et al.* Pharmacokinetics and safety of cavosonstat (N91115) in healthy and cystic fibrosis adults homozygous for F508DEL-CFTR. *J. Cyst. Fibros.* **16**, 371–379 (2017).

46. ClinicalTrials.gov. *Study of Cavosonstat (N91115) in CF Patients Who Are Heterozygous for F508del-CFTR and a Gating Mutation and Being Treated With Ivacaftor (SNO-7)* Available at: https://clinicaltrials.gov/ct2/show/NCT02724527.
47. Clancy, J. P. *et al.* CFTR modulator theratyping: Current status, gaps and future directions. *J. Cyst. Fibros.* **18**, 22–34 (2019).
48. Igreja, S., Clarke, L. A., Botelho, H. M., Marques, L. & Amaral, M. D. Correction of a Cystic Fibrosis Splicing Mutation by Antisense Oligonucleotides. *Hum. Mutat.* **37**, 209–215 (2016).
49. Bonini, J. *et al.* Small-scale high-throughput sequencing-based identification of new therapeutic tools in cystic fibrosis. *Genet. Med.* **17**, 796–806 (2015).
50. Carlon, M. S., Vidović, D. & Birket, S. Roadmap for an early gene therapy for cystic fibrosis airway disease. *Prenat. Diagn.* **37**, 1181–1190 (2017).
51. Alton, E. W. F. W. *et al.* The safety profile of a cationic lipid-mediated cystic fibrosis gene transfer agent following repeated monthly aerosol administration to sheep. *Biomaterials* **34**, 10267–10277 (2013).
52. Eric W. F. W. Alton, A. Christopher Boyd, David J. Porteous, Gwyneth Davies, Jane C. Davies, Uta Griesenbach, Tracy E. Higgins, Deborah R. Gill, Stephen C. Hyde, and J. A. I. A Phase I/IIa Safety and Efficacy Study of Nebulized Liposome-mediated Gene Therapy for Cystic Fibrosis Supports a Multidose Trial. *Am. J. Respir. Crit. Care Med.* **192**, 1389–92 (2015).
53. Alton, E. 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. *Lancet Respir Med* **3**, 684–691 (2015).
54. Yin, H. *et al.* Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.* **15**, 541–555 (2014).
55. Zabner, J. *et al.* Adenovirus-mediated gene transfer transiently corrects the chloride transport defect in nasal epithelia of patients with cystic fibrosis. *Cell* **75**, 207–216 (1993).
56. Crystal, R. G. *et al.* Administration of an adenovirus containing the human CFTR cDNA to the respiratory tract of individuals with cystic fibrosis. *Nat. Genet.* **8**, 42–51 (1994).
57. Zabner, J. *et al.* Repeat administration of an adenovirus vector encoding cystic fibrosis transmembrane conductance regulator to the nasal epithelium of patients with cystic fibrosis. *J. Clin. Invest.* **97**, 1504–1511 (1996).
58. Steines, B. *et al.* CFTR gene transfer with AAV improves early cystic fibrosis pig phenotypes. *JCI Insight* **1**, e88728 (2016).
59. Halbert, C. L., Allen, J. M. & Miller, a D. Adeno-Associated Virus Type 6 (AAV6) Vectors Mediate Efficient Transduction of Airway Epithelial Cells in Mouse Lungs Compared to That of AAV2 Vectors Downloaded from http://jvi.asm.org/ on March 1 , 2015 by SERIALS CONTROL Lane Medical Library. **75**, 6615–6624 (2001).
60. Carlon, M. *et al.* Efficient gene transfer into the mouse lung by fetal intratracheal injection of rAAV2/6.2. *Mol. Ther.* **18**, 2130–2138 (2010).
61. Miah, K. M., Hyde, S. C. & Gill, D. R. Emerging gene therapies for cystic fibrosis. *Expert Rev. Respir. Med.* **13**, 709–725 (2019).
62. Vidovic, D. *et al.* rAAV-CFTRDeltaR Rescues the Cystic Fibrosis Phenotype in Human Intestinal Organoids and Cystic Fibrosis Mice. *Am J Respir Crit Care Med* **193**, 288–298 (2016).
63. Castellani, S. & Conese, M. Lentiviral vectors and cystic fibrosis gene therapy. *Viruses* **2**, 395–412 (2010).
64. Milone, M. C. & O'Doherty, U. Clinical use of lentiviral vectors. *Leukemia* **32**, 1529–1541 (2018).
65. Marquez Loza, L., Yuen, E. & McCray, P. Lentiviral Vectors for the Treatment and Prevention of Cystic Fibrosis Lung Disease. *Genes (Basel).* **10**, 218 (2019).
66. Alton, E. W. F. W. *et al.* Preparation for a first-in-man lentivirus trial in patients with cystic fibrosis. *Thorax* **72**, 137–147 (2017).

67. Cronin, J., Zhang, X.-Y. & Reiser, J. Altering the Tropism of Lentiviral Vectors through Pseudotyping. *Curr. Gene Ther.* **5**, 387–398 (2005).
68. Mitomo, K. *et al.* Toward gene therapy for cystic fibrosis using a lentivirus pseudotyped with sendai virus envelopes. *Mol. Ther.* **18**, 1173–1182 (2010).
69. Limberis, M., Anson, D. S., Fuller, M. & Parsons, D. W. Recovery of airway cystic fibrosis transmembrane conductance regulator function in mice with cystic fibrosis after single-dose lentivirus-mediated gene transfer. *Hum. Gene Ther.* **13**, 1961–1970 (2002).
70. Cooney, A. L. *et al.* Lentiviral-mediated phenotypic correction of cystic fibrosis pigs. *JCI Insight* **1**, e88730 (2016).
71. Crosby, J. R. *et al.* Inhaled ENaC antisense oligonucleotide ameliorates cystic fibrosis-like lung disease in mice. *J. Cyst. Fibros.* 1–10 (2016). doi:10.1016/j.jcf.2017.05.003
72. Alton, E. W. F. W. *et al.* Genetic medicines for CF: Hype versus reality. *Pediatric Pulmonology* **51**, S5–S17 (2016).
73. Crane, A. M. *et al.* Targeted correction and restored function of the CFTR gene in cystic fibrosis induced pluripotent stem cells. *Stem Cell Reports* **4**, 569–577 (2015).
74. Schwank, G. *et al.* Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* **13**, 653–658 (2013).
75. Kim, H. & Kim, J. S. A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.* **15**, 321–334 (2014).
76. 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* 1–11 (2019). doi:10.1016/j.stem.2019.11.002
77. Wagner, D. L. *et al.* High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. *Nat. Med.* **25**, 242–248 (2019).
78. Charlesworth, C. T. *et al.* Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254 (2019).
79. Bednarski, C., Tomczak, K., Hövel, B. Vom, Weber, W. M. & Cathomen, T. Targeted integration of a super-exon into the CFTR locus leads to functional correction of a cystic fibrosis cell line model. *PLoS One* **11**, 1–15 (2016).
80. Mall, M., Grubb, B. R., Harkema, J. R., O'Neal, W. K. & Boucher, R. C. Increased airway epithelial Na+ absorption produces cystic fibrosis-like lung disease in mice. *Nat. Med.* **10**, 487–493 (2004).
81. Boyd, A. C. *et al.* New approaches to genetic therapies for cystic fibrosis. (2020). doi:10.1016/j.jcf.2019.12.012
82. 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* **73**, 847–856 (2018).
83. Barrangou, R. *et al.* CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science (80-.).* **315**, 1709–1712 (2007).
84. Sorek, R., Lawrence, C. M. & Wiedenheft, B. CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea. *Annu. Rev. Biochem.* **82**, 237–266 (2013).
85. Haft, D. H., Selengut, J., Mongodin, E. F. & Nelson, K. E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **1**, 0474–0483 (2005).
86. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653–663 (2005).
87. Bolotin, A., Quinquis, B., Sorokin, A. & Dusko Ehrlich, S. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551–2561 (2005).
88. Van Der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **12**, 479–492 (2014).

89. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733–740 (2009).
90. Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **9**, 467–477 (2011).
91. Stan J. J. Brouns, M. M. J. *et al.* Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science (80-.).* 960–965 (2008).
92. Semenova, E. *et al.* Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10098–10103 (2011).
93. Makarova, K. S. *et al.* Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, (2019).
94. Jiang, F. & Doudna, J. A. CRISPR – Cas9 Structures and Mechanisms. *Annu. Rev. Biophys* **46**, 505–529 (2017).
95. Chen, H., Choi, J. & Bailey, S. Cut site selection by the two nuclease domains of the Cas9 RNA-guided endonuclease. *J. Biol. Chem.* **289**, 13284–13294 (2014).
96. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 2579–2586 (2012).
97. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).
98. Jinek, M. *et al.* A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.).* **337**, 816–821 (2012).
99. Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).
100. Casini, A. *et al.* A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat. Biotechnol.* **36**, 265–271 (2018).
101. Slaymaker, I. M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science (80-.).* **351**, 84–88 (2016).
102. Kleinstiver, B. P. *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
103. Chen, J. S. *et al.* Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407–410 (2017).
104. Petris, G. *et al.* Hit and go CAS9 delivered through a lentiviral based self-limiting circuit. *Nat. Commun.* **8**, 1–9 (2017).
105. Montagna, C. *et al.* VSV-G-Enveloped Vesicles for Traceless Delivery of CRISPR-Cas9. *Mol. Ther. - Nucleic Acids* **12**, 453–462 (2018).
106. Swarts, D. C. & Jinek, M. Cas9 versus Cas12a/Cpf1: Structure – function comparisons and implications for genome editing. *Wiley Interdiscip Rev RNA* **9**, e1481 (2018).
107. Zetsche, B. *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163**, 759–771 (2015).
108. Jeon, Y. *et al.* Direct observation of DNA target searching and cleavage by CRISPR-Cas12a. *Nat. Commun.* **9**, 2777 (2018).
109. Gao, L. *et al.* Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* **35**, 789–792 (2017).
110. Zaidi, S. S. e. A., Mahfouz, M. M. & Mansoor, S. CRISPR-Cpf1: A New Tool for Plant Genome Editing. *Trends Plant Sci.* **22**, 550–553 (2017).
111. Kleinstiver, B. P. *et al.* Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* **34**, 869–874 (2016).
112. Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013).
113. Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–350 (2014).
114. Capecchi, M. R. Altering the genome by homologous recombination. *Science (80-.).*

236, (1989).
115. Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol.* **14**, 8096–8106 (1994).
116. Bibikova, M. *et al.* Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Mol. Cell. Biol.* **21**, 289–297 (2001).
117. Lombardo, A. *et al.* Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* **8**, 861–869 (2011).
118. Wang, F. & Qi, L. S. Applications of CRISPR Genome Engineering in Cell Biology. *Trends Cell Biol.* **26**, 875–888 (2016).
119. Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* **8**, 2180–2196 (2013).
120. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661 (2014).
121. Hilton, I. B. *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517 (2015).
122. Kearns, N. A. *et al.* Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* **12**, 401–403 (2015).
123. Evers, B. *et al.* CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat. Biotechnol.* **34**, 631–633 (2016).
124. Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Sciencecience* **343**, 84–87 (2014).
125. Shi, J. *et al.* Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* **33**, 661–667 (2015).
126. Korkmaz, G. *et al.* Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat. Biotechnol.* **34**, 192–198 (2016).
127. Ratz, M., Testa, I., Hell, S. W. & Jakobs, S. CRISPR/Cas9-mediated endogenous protein tagging for RESOLFT super-resolution microscopy of living human cells. *Sci. Rep.* **5**, 1–6 (2015).
128. Deng, W., Shi, X., Tjian, R., Lionnet, T. & Singer, R. H. CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11870–11875 (2015).
129. Ma, H. *et al.* Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3002–3007 (2015).
130. McKenna, A. *et al.* Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science (80-.).* **353**, (2016).
131. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
132. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* **19**, 770–788 (2018).
133. Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, (2019).
134. Li, B., Niu, Y., Ji, W. & Dong, Y. Strategies for the CRISPR-Based Therapeutics. *Trends Pharmacol. Sci.* **41**, 55–65 (2019).
135. Li, Y., Glass, Z., Huang, M., Chen, Z. Y. & Xu, Q. Ex vivo cell-based CRISPR/Cas9 genome editing for therapeutic applications. *Biomaterials* **234**, 119711 (2020).
136. ClinicalTrials.gov. *Identifier: NCT03399448*
137. ClinicalTrials.gov. *Identifier: NCT03655678 and NCT03745287*
138. Wu, Y. *et al.* Highly efficient therapeutic gene editing of human hematopoietic stem cells. *Nat. Med.* **25**, 776–783 (2019).
139. Young, C. S. *et al.* A Single CRISPR-Cas9 Deletion Strategy that Targets the Majority of DMD Patients Restores Dystrophin Function in hiPSC-Derived Muscle Cells. *Cell Stem Cell* **18**, 533–540 (2016).

140. Li, H. L. *et al.* Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Reports* **4**, 143–154 (2015).
141. Long, C. *et al.* Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science (80-.).* **345**, 1184–1188 (2014).
142. Jin, Y., Shen, Y., Su, X., Weintraub, N. & Tang, Y. CRISPR/Cas9 Technology in Restoring Dystrophin Expression in iPSC-Derived Muscle Progenitors. *J. Vis. Exp.* 1–9 (2019). doi:10.3791/59432
143. Min, Y. L. *et al.* CRISPR-Cas9 corrects Duchenne muscular dystrophy exon 44 deletion mutations in mice and human cells. *Sci. Adv.* **5**, 1–13 (2019).
144. Ding, Q. *et al.* Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. *Circ. Res.* **115**, 488–492 (2014).
145. Wang, X. *et al.* CRISPR-Cas9 targeting of PCSK9 in human hepatocytes in vivo - Brief report. *Arterioscler. Thromb. Vasc. Biol.* **36**, 783–786 (2016).
146. ClinicalTrials.gov. *identifier: NCT03872479*
147. Maeder, M. L. *et al.* Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. *Nat. Med.* **25**, 229–233 (2019).
148. Oppel, F., Schurmann, M., Goon, P., Albers, A. E. & Sudhoff, H. Specific targeting of oncogenes using CRISPR technology. *Cancer Res.* **78**, 5506–5512 (2018).
149. ClinicalTrials.gov. *Identifier: NCT03057912*
150. Wilkinson, M. E. & Charenton, C. RNA Splicing by the Spliceosome. *Annu. Rev. Biochem.* **89**, 1–30 (2019).
151. Black, A. J., Gamarra, J. R. & Giudice, J. More than a messenger: Alternative splicing as a therapeutic target. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1862**, 194395 (2019).
152. Robberson, B. L., Cote, G. J. & Berget, S. M. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* **10**, 84–94 (1990).
153. Roca, X., Krainer, A. R. & Eperon, I. C. Pick one, but be quick: 5′ splice sites and the problems of too many choices. *Genes Dev.* **27**, 129–144 (2013).
154. Padgett, R. A. New connections between splicing and human disease. *Trends Genet.* **28**, 147–154 (2012).
155. Hang, J., Wan, R., Yan, C. & Shi, Y. Structural basis of pre-mRNA splicing. *Sci.* **349**, 1191–1198 (2015).
156. Bartys, N., Kierzek, R. & Lisowiec-Wachnicka, J. The regulation properties of RNA secondary structure in alternative splicing. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1862**, 194401 (2019).
157. Desmet, F. O. *et al.* Human Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37**, 1–14 (2009).
158. Bourgeois, C. F., Lejeune, F. & Ste, J. *Broad Specificity of SR (Serine / Arginine) Proteins in the Regulation of Alternative Splicing of Pre-Messenger RNA I . Introduction More than 25 years after the discovery of introns in higher eukaryotes , the*. (2004).
159. Zhang, X. H. F., Leslie, C. S. & Chasin, L. A. Computational searches for splicing signals. *Methods* **37**, 292–305 (2005).
160. Blencowe, B. J. Exonic splicing enhancers: Mechanism of action, diversity and role in human genetic diseases. *Trends Biochem. Sci.* **25**, 106–110 (2000).
161. Zhu, J. *et al.* Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol. Cell* **8**, 1351–61 (2001).
162. Venables, J. P. Downstream intronic splicing enhancers. *FEBS Lett.* **581**, 4127–4131 (2007).
163. Wang, Z. & Burge, C. B. Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *Rna* **14**, 802–813 (2008).
164. Barash, Y. *et al.* Deciphering the splicing code. *Nature* **465**, 53–59 (2010).
165. Ule, J. & Blencowe, B. J. Alternative Splicing Regulatory Networks: Functions,

Mechanisms, and Evolution. *Mol. Cell* **76**, 329–345 (2019).
166. Matera, A. G. & Wang, Z. A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* **15**, 108–121 (2014).
167. Sperling, R. The nuts and bolts of the endogenous spliceosome. *Wiley Interdiscip. Rev. RNA* **8**, (2017).
168. Plaschka, C., Lin, P. C. & Nagai, K. Structure of a pre-catalytic spliceosome. *Nature* **546**, 617–621 (2017).
169. Cesaratto, F., López-Requena, A., Burrone, O. R. & Petris, G. Engineered tobacco etch virus (TEV) protease active in the secretory pathway of mammalian cells. *J. Biotechnol.* **212**, 159–66 (2015).
170. Alanis, E. F. *et al.* An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. *Hum. Mol. Genet.* **21**, 2389–2398 (2012).
171. Shalem, O. *et al.* Genome - scale CRISPR - Cas9 knockout screening in human cells. *Science (80-.).* **343**, 84–87 (2014).
172. Casini, A., Olivieri, M., Vecchi, L., Burrone, O. R. & Cereseto, A. Reduction of HIV-1 Infectivity through Endoplasmic Reticulum-Associated Degradation-Mediated Env Depletion. *J. Virol.* **89**, 2966–2971 (2015).
173. Hsiau, T. *et al.* Inference of CRISPR Edits from Sanger Trace Data. *bioRxiv* **Jan. 20,** 1–14 (2018).
174. Brinkman, E. K., Chen, T., Amendola, M. & Van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, 1–8 (2014).
175. Sondo, E. *et al.* Pharmacological Inhibition of the Ubiquitin Ligase RNF5 Rescues F508del-CFTR in Cystic Fibrosis Airway Epithelia. *Cell Chem. Biol.* **25**, 891-905.e8 (2018).
176. Dekkers, J. F. *et al.* A functional CFTR assay using primary cystic fibrosis intestinal organoids. *Nat. Med.* **19**, 939–945 (2013).
177. Vidovic, D. *et al.* rAAV-CFTRDR rescues the cystic fibrosis phenotype in human intestinal organoids and cystic fibrosis mice. *Am. J. Respir. Crit. Care Med.* **193**, 288–298 (2016).
178. Sato, T. *et al.* Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762–1772 (2011).
179. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–198 (2015).
180. Pinello, L. *et al.* Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat. Biotechnol.* **34**, 695–697 (2016).
181. Hamroun, D. & Lalande, M. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37**, 1–14 (2009).
182. Yeo, G. & Burge, C. B. Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals. *J. Comput. Biol.* **11**, 377–394 (2004).
183. Buratti, E., Stuani, C., De Prato, G. & Baralle, F. E. SR protein-mediated inhibition of CFTR exon 9 inclusion: Molecular characterization of the intronic splicing silencer. *Nucleic Acids Res.* **35**, 4359–4368 (2007).
184. Kashima, T., Rao, N., David, C. J. & Manley, J. I. hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. *Hum. Mol. Genet.* **16**, 3149–3159 (2007).
185. Cooper, T. A. Use of minigene systems to dissect alternative splicing elements. *Methods* **37**, 331–340 (2005).
186. Kim, D. *et al.* Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* **34**, 863–868 (2016).
187. Kleinstiver, B. P. *et al.* Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* **34**, 869–874 (2016).
188. Awatade, N. T. *et al.* Human Primary Epithelial Cell Models: Promising Tools in the Era of Cystic Fibrosis Personalized Medicine. *Front. Pharmacol.* **9**, 1429 (2018).
189. Gianotti, A., Delpiano, L. & Caci, E. In vitro methods for the development and analysis

of human primary airway epithelia. *Front. Pharmacol.* **9**, 1–12 (2018).
190. Amaral, M. D. *et al.* Cystic fibrosis patients with the 3272-26A>G splicing mutation have milder disease than F508del homozygotes: a large European study. *J. Med. Genet.* **38**, 777–83 (2001).
191. Dekkers, J. F. *et al.* Characterizing responses to CFTR-modulating drugs using rectal organoids derived from subjects with cystic fibrosis. *Sci. Transl. Med.* **8**, 344ra84 (2016).
192. Shen, M. W. *et al.* Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **563**, 646–651 (2018).
193. van Overbeek, M. *et al.* DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. *Mol. Cell* **63**, 633–646 (2016).
194. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
195. Hart, S. L. & Harrison, P. T. Genetic therapies for cystic fibrosis lung disease. *Curr. Opin. Pharmacol.* **34**, 119–124 (2017).
196. Berical, A., Lee, R. E., Randell, S. H. & Hawkins, F. Challenges facing airway epithelial cell-based therapy for cystic fibrosis. *Front. Pharmacol.* **10**, 1–12 (2019).
197. Moss, R. B. *et al.* Repeated Aerosolized AAV-CFTR for Treatment of Cystic Fibrosis: A Randomized Placebo-Controlled Phase 2B Trial. *Hum. Gene Ther.* **18**, 726–732 (2007).
198. Griesenbach, U. *et al.* Assessment of F/HN-pseudotyped lentivirus as a clinically relevant vector for lung gene therapy. *Am. J. Respir. Crit. Care Med.* **186**, 846–856 (2012).
199. Sinn, P. L., Burnight, E. R., Hickey, M. A., Blissard, G. W. & McCray, P. B. Persistent Gene Expression in Mouse Nasal Epithelia following Feline Immunodeficiency Virus-Based Vector Gene Transfer. *J. Virol.* **79**, 12818–12827 (2005).
200. Stocker, A.G.; Kremer, K.L.; Koldej, R.; Miller, D.S.; Anson, D.S.; Parsons, D. W. Single-dose lentiviral gene transfer for lifetime airway gene expression. *J. Gene Med.* **11**, 861–867 (2009).
201. Patel, M.; Giddings, A.M.; Sechelski, J.; Olsen, J. C. High efficiency gene transfer to airways of mice using influenza hemagglutinin pseudotyped lentiviral vectors. *J. Gene Med.* **15**, 51–62 (2013).
202. Mamcarz, E. *et al.* Lentiviral Gene Therapy Combined with Low-Dose Busulfan in Infants with SCID-X1. *N. Engl. J. Med.* **380**, 1525–1534 (2019).
203. Herbst, F. *et al.* Extensive methylation of promoter sequences silences lentiviral transgene expression during stem cell differentiation in vivo. *Mol. Ther.* **20**, 1014–1021 (2012).
204. Gill, D. R. *et al.* Increased persistence of lung gene expression using plasmids containing the ubiquitin C or elongation factor 1α promoter. *Gene Ther.* **8**, 1539–1546 (2001).
205. Farrow, N. *et al.* Role of Basal Cells in Producing Persistent Lentivirus-Mediated Airway Gene Expression. *Hum. Gene Ther.* **29**, 653–662 (2018).
206. Montoro, D. T. *et al.* A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* (2018). doi:10.1038/s41586-018-0393-7
207. Plasschaert, L.W.; Zilionis, R.; Choo-Wing, R.; Savova, V.; Knehr, J.; Roma, G.; Klein, A.M.; Jaffe, A. B. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **560,377**–**38**, (2018).
208. Kreda, S.M.; Mall, M.; Mengos, A.; Rochelle, L.; Yankaskas, J.; Riordan, J.R.; Boucher, R. C. Characterization of Wild-Type and deltaF508 Cystic Fibrosis Transmembrane Regulator in Human Respiratory Epithelia. *Mol Biol Cell* **16**, **2154**–**2**, (2005).
209. Johnson, L. G. *et al.* Efficiency of gene transfer for restoration of normal airway epithelial function in cystic fibrosis. *Nat. Genet.* **2**, 21–25 (1992).
210. Goldman, M. J., Yang, Y. & Wilson, J. M. Gene therapy in a xenograft model of cystic fibrosis lung corrects chloride transport more effectively than the sodium defect. *Nat.*

Genet. **9**, 126–131 (1995).
211. 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.* **7**, (2009).
212. Farmen, S. L. *et al.* Gene transfer of CFTR to airway epithelia: Low levels of expression are sufficient to correct Cl- transport and overexpression can generate basolateral CFTR. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **289**, 1123–1131 (2005).
213. Dannhoffer, L., Blouquit-Laye, S., Regnier, A. & Chinet, T. Functional properties of mixed cystic fibrosis and normal bronchial epithelial cell cultures. *Am. J. Respir. Cell Mol. Biol.* **40**, 717–723 (2009).
214. Miller, A. J. *et al.* In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells. *Stem Cell Reports* **10**, 101–119 (2018).
215. Sanz, D. J., Hollywood, J. A., Scallan, M. F. & Harrison, P. T. Cas9/gRNA targeted excision of cystic fibrosis-causing deep-intronic splicing mutations restores normal splicing of CFTR mRNA. *PLoS One* **12**, 1–13 (2017).

APPENDIX 1. Oligonucleotides for PCR

Oligos for the 3272-26A>G mutation analyses

Oligos for minigene cloning, ***RT-PCR and PCR for TIDE analysis***

oligo KpnI-AgeI exon 18-19 hCFTR fw	ATggtaccggtgaccttctgcctcTTACCATATTTGACTTCATCCAGTTG
oligo TM exon 18 exon 19 hCFTR fw	ttaccatatttgacttcatccagTTGTTATTAATTGTGATTGGAGCTATAG
oligo exon 20 hCFTR rev	TGtAgaattcttaggatccctcgcCTGTTGTTAAAATGGAAATGAAGGTAACAG

Oligos for site directed mutagenesis on wt minigene plasmid

oligo MUT 3272-26 A>G fw	ATGGTCTCAgTGTTTTCTATGGAAATATTTCAC
oligo MUT 3272-26 A>G rev	ATGGTCTCaacAcTGCAAATAACATAAACACAAAATG

Oligo for RT-PCR and PCR for TIDE analysis

oligo BGH rev	TAGAAGGCACAGTCGAGG

Oligos for deep sequencing

DS 3272-26 ON fw	tcgtcggcagcgtcagatgtgtataagagacagGCTTGTAACAAGATGAGTGAAAATTGGA
DS 3272-26 ON rev	gtctcgtgggctcggagatgtgtataagagacagATATCTATTCAAAGAATGGCACCAGTGT
DS 3272-26 OFF 1 fw	tcgtcggcagcgtcagatgtgtataagagacagAAGGGTAAATTTACTTTAACTCTCCTTAC
DS 3272-26 OFF 1 rev	gtctcgtgggctcggagatgtgtataagagacagACTACTGAATACTGTTCTATTGCACAG
DS 3272-26 OFF 2 fw	tcgtcggcagcgtcagatgtgtataagagacagGGAAATAACATGTATTTGCTCTATTTAACTAG
DS 3272-26 OFF 2 rev	gtctcgtgggctcggagatgtgtataagagacagCTACTATTCAGGTCATGTTAGTAAAATGG
DS 3272-26 OFF 3 fw	tcgtcggcagcgtcagatgtgtataagagacagATACATTACTTGTATAAAAGCATGACTAAGC
DS 3272-26 OFF 3 rev	gtctcgtgggctcggagatgtgtataagagacagTCTTATTTGTGGCATTTACTAACTTTCTTAG
DS 3272-26 OFF 4 fw	tcgtcggcagcgtcagatgtgtataagagacagGTTTCCAAGATGTGATAACAATACTTGC
DS 3272-26 OFF 4 rev	gtctcgtgggctcggagatgtgtataagagacagTTTCTATGATTTTGATATAGTAACTTACTCAAAC
DS 3272-26 OFF 5 fw	tcgtcggcagcgtcagatgtgtataagagacagTTGCCGGCTGACTAGCACTTC
DS 3272-26 OFF 5 rev	gtctcgtgggctcggagatgtgtataagagacagAATAGGAACCATTACAATTAACATATTTTTGC

DS 3272-26 OFF 6 fw	tcgtcggcagcgtcagatgtgtataagagacagGTTACTCGGCTGGTAAATTGCATAG
DS 3272-26 OFF 6 rev	gtctcgtgggctcggagatgtgtataagagacagGTATTTTTAAATGATATTCTGAAACTTCCCTG
DS 3272-26 OFF 7 fw	tcgtcggcagcgtcagatgtgtataagagacagTCTCTGTTGTATGCTAACAGTTCTTC
DS 3272-26 OFF 7 rev	gtctcgtgggctcggagatgtgtataagagacagGATGAGTATTGACTGGCAAGGAC
DS 3272-26 OFF 8 fw	tcgtcggcagcgtcagatgtgtataagagacagTATGCAGTTCTTAAAATAAATGGAGCAAATC
DS 3272-26 OFF 8 rev	gtctcgtgggctcggagatgtgtataagagacagGTGATAGTTCCTCATTGTTACTATTATTATTG
DS 3272-26 OFF 9 fw	tcgtcggcagcgtcagatgtgtataagagacagCATTCTATACAGAAAGCAATTCATTATGTTTATA
DS 3272-26 OFF 9 rev	gtctcgtgggctcggagatgtgtataagagacagTTTATAAAAACAATACATGTGGAATGTAAAAAGG
DS 3272-26 OFF 10 fw	tcgtcggcagcgtcagatgtgtataagagacagTGTAGATCATTCTTCCTCTCTGCTG
DS 3272-26 OFF 10 rev	gtctcgtgggctcggagatgtgtataagagacagTAACCATTCTACTCTAGAGATAACCTTG
DS 3272-26 OFF 11 fw	tcgtcggcagcgtcagatgtgtataagagacagCATTCACCATATTTTTTCAGCCCACC
DS 3272-26 OFF 11 rev	gtctcgtgggctcggagatgtgtataagagacagAGGTTACAAATACGGAAAAATCAGACAG
DS 3272-26 OFF 12 fw	tcgtcggcagcgtcagatgtgtataagagacagAAATATGAATATTTTAAGACGGAAGCTAGTC
DS 3272-26 OFF 12 rev	gtctcgtgggctcggagatgtgtataagagacagCTCTTCATCTGCAAATGGGACTAAG

Oligos for the 3849+10kbC>T mutation analyses

Oligos for minigene cloning, ***RT-PCR and PCR***

oligo XhoI ex 22 CFTR fw1	ATATctcgagATGCGATCTGTGAGCCGAGTCTTTAA
oligo BsmBI_int22_new ex fw2	tacgtctcATAtATTCAGTGGGTATAAGCAGCATATTCTC
oligo BamHI-BsmBI_ex 23 fw3	tatggatccagatcgtctcgAAAGGTCAGTGATAAAGGAAGTCTGCAT
oligo int22 BsmBI-BamHI rev1	tatggatccagatcgtctcgATATAGGTTCAGGACTCTGCAAATTAAATTTC
oligo new ex_BsmBI rev2	atcgtctctCTTtAGGCTTCTCAGTGATCTGTTGAATAAG
oligo ex 23_NotI rev3	atagtgcggccgcCTGTGGTATCACTCCAAAGGCTTTC

Oligos for site directed mutagenesis on wt minigene plasmid

oligo MUT 3849 10kb C-T fw	CCATCTGTTGCAGTATTAAAATGGtGAGTAAGACACCCTGAAAGG
oligo MUT 3849 10kb C-T rev	CCTTTCAGGGTGTCTTACTCaCCATTTTAATACTGCAACAGATGG

Oligo for RT-PCR minigene

oligo T7F2 (x pCI)	TACTTAATACGACTCACTATAGGCTAGCCTCG

Oligos for PCR for TIDE analysis

oligo CFTR 10kb int 22 fw	CTGCTTTCTCCATTTGTAGTCTCTTG
oligo CFTR 10kb int 22 rev	TGCTGGTAATGCATGATATCTGACAC

Oligos for deep sequencing

DS 3849+10kb ON fw	tcgtcggcagcgtcagatgtgtataagagacagGTGCTAGCTGTAATTGCATTGTACC
DS 3849+10kb ON rev	gtctcgtgggctcggagatgtgtataagagacagCATAAAAGATTCAATTATAATCACCTTGTGG
DS 3849+10kb OFF 1 fw	tcgtcggcagcgtcagatgtgtataagagacagTCTTTGTTTCCTTTCAGTTCCAGGTG
DS 3849+10kb OFF 1 rev	gtctcgtgggctcggagatgtgtataagagacagTCAACAACCTAACTACCAGAACCATC
DS 3849+10kb OFF 2 fw	tcgtcggcagcgtcagatgtgtataagagacagTTAAATTTCTCTACTTTTACTTTGTGTTTAGTTAG
DS 3849+10kb OFF 2 rev	gtctcgtgggctcggagatgtgtataagagacagAGGACAAAAGAGAACCCTTGAACAC
DS 3849+10kb OFF 3 fw	tcgtcggcagcgtcagatgtgtataagagacagCAACCTAACTGCCCACAGAGAC
DS 3849+10kb OFF 3 rev	gtctcgtgggctcggagatgtgtataagagacagTTAGCTTCAAGCACCGACACCG

APPENDIX 2. Sequences of sgRNAs spacers and relative target sites

Guide RNA name is the distance of the PAM from the mutation, + and - indicate the gRNA position, before or after the mutation respectively.

AsCas12a spacer sequences used to target CFTR 3272-26 locus

Name	Protospacer (1)	Target (*)
-77	TGATATGATTATTCTAATTTAGT	aca**TTTG**TGATATGATTATTCTAATTTAGTctt
-56	GTCTTTTTCAGGTACAAGATATT	taa**TTTA**GTCTTTTTCAGGTACAAGATATTatg
-27	ATAATATCTTGTACCTGAAAAAG	taa**TTTC**ATAATATCTTGTACCTGAAAAAGact
-11	TGTTATTTGCA**G**TGTTTTCTATG	gtg**TTTA**TGTTATTTGCAGTGTTTTCTATGgaa
-2	CA**G**TGTTTTCTATGGAAATATTT	tta**TTTG**CA**G**TGTTTTCTATGGAAATATTTcac
+11	CATAGAAAACA**C**TGCAAATAACA	ata**TTTC**CATAGAAAACA**C**TGCAAATAACAtaa
+11/wt	CATAGAAAACA**T**TGCAAATAACA	ata**TTTC**CATAGAAAACA**T**TGCAAATAACAtaa

SpCas9 spacer sequences used to target CFTR 3272-26 locus

Name	Protospacer (1)	Target (*)
-88	AATCATATCACAAATGTCAT	aatAATCATATCACAAATGTCAT**TGG**tta
-62	GTACCTGAAAAAGACTAAAT	cttGTACCTGAAAAAGACTAAAT**TAG**aat
-52	ATAATATCTTGTACCTGAAA	ttcATAATATCTTGTACCTGAAA**AAG**act
-47	ATTCTAATTTAGTCTTTTTC	attATTCTAATTTAGTCTTTTTC**AGG**tac
-0	TTTTGTGTTTATGTTATTTG	acaTTTTGTGTTTATGTTATTTG**CAG**tgt
+9	CTGCCTGTGAAATATTTCCA	ctcCTGCCTGTGAAATATTTCCA**TAG**aaa
+10	GTTATTTGCA**G**TGTTTTCTA	tatGTTATTTGCA**G**TGTTTTCTA**TGG**aaa
+22	TTTTCTATGGAAATATTTCA	gtgTTTTCTATGGAAATATTTCA**CAG**gca

AsCas12a spacer sequences used to target CFTR 3849+10Kb locus

Name	Protospacer (2)	Target (*)
+14	AGGGTGTCTTACTC**A**CCATTTTA	tcc**TTTC**AGGGTGTCTTACTC**A**CCATTTTAata
+14/wt	AGGGTGTCTTACTC**G**CCATTTTA	tcc**TTTC**AGGGTGTCTTACTC**G**CCATTTTAata

SpCas9 spacer sequences used to target CFTR 3849+10Kb locus

Name	Protospacer (2)	Target (*)
-95	ATTCAATTATAATCACCTTG	aagATTCAATTATAATCACCTTG**TGG**atc
-99	AACTGAAATTTAGATCCACA	gtcAACTGAAATTTAGATCCACA**AGG**tga
-143	CTTGATTTCTGGAGACCACA	catCTTGATTTCTGGAGACCACA**AGG**taa
+1	TGTTGCAGTATTAAAATGG**T**	atcTGTTGCAGTATTAAAATGG**TGAG**taa
+5	GCAGTATTAAAATGG**T**GAGT	gttGCAGTATTAAAATGG**T**GAGT**AAG**aca
+18	GG**T**GAGTAAGACACCCTGAA	aatGG**T**GAGTAAGACACCCTGAA**AGG**aaa
+34	GAAAGGAAATGTTCTATTCA	cttGAAAGGAAATGTTCTATTCA**TGG**tac
+119	CACCTCCTCCCTGAGAATGT	atgCACCTCCTCCCTGAGAATGT**TGG**atc
+125	TTGATCCAACATTCTCAGGG	atcTTGATCCAACATTCTCAGGG**AGG**agg

(1) 3272-26A>G mutation position is highlighted in bold
(2) 3849+10KbC>T mutation position is highlighted in bold
(*) PAM is in bold. Nucleotides around target site are in lowercase.