ISSN 0006-2979, Biochemistry (Moscow), 2025, Vol. 90, No. 6, pp. 773-785 © Pleiades Publishing, Ltd., 2025.
Russian Text © The Author(s), 2025, published in Biokhimiya, 2025, Vol. 90, No. 6, pp. 833-846.

Selection of Optimal pegRNAs to Enhance Efficiency of Prime Editing in AT-Rich Genome Regions

Olga V. Volodina[1,a]*, Anna G. Demchenko[1], Arina A. Anuchina[1], Oxana P. Ryzhkova[1], Valeriia A. Kovalskaya[1], Ekaterina V. Kondrateva[1], Vyacheslav Y. Tabakov[1], Alexander V. Lavrov[1], and Svetlana A. Smirnikhina[1]

[1]*Research Centre for Medical Genetics, 115522 Moscow, Russia*
[a]*e-mail: volodold@gmail.com*

Received December 27, 2024
Revised March 18, 2025
Accepted March 26, 2025

Abstract—Prime editing is a highly promising strategy for treating hereditary disorders due to its superior efficiency and safety profile compared to the conventional CRISPR-Cas9 systems. This study is dedicated to development of a causal therapy for cystic fibrosis by targeting the F508del variant of the *CFTR* gene using prime editing, as this specific deletion accounts for a substantial proportion of cystic fibrosis cases. While prime editing has shown remarkable precision in introducing targeted genetic modifications, its application in AT-rich genomic regions, such as the one containing the F508del variant, remains challenging. To overcome this limitation, we systematically evaluated 24 pegRNAs designed for two distinct prime editing systems, PEmax and PE2-NG. Efficiency of the F508del variant correction reached 2.81% (without normalization for transfection efficiency) in the airway basal cells from the patients with homozygous F508del mutation. However, the average transfection efficiency was only 11.9%, emphasizing critical need for the advancements in delivery methodologies. These findings highlight potential of prime editing as an approach for treating cystic fibrosis, while also underscoring necessity for further optimization of both editing constructs and delivery vectors to achieve clinically relevant correction levels.

DOI: 10.1134/S0006297924604672

Keywords: prime editing, *CFTR*, genome editing, pegRNA, F508del

INTRODUCTION

Genome editing methods have been often used recently for the development of gene therapy techniques, which allow introducing targeted changes to the genome such as correcting genetic variants and gene knockouts. At present the system based on the clustered regularly interspaced short palindromic repeats, CRISPR, associated with the Cas9 protein (CRISPR/Cas9 system) has been used most often [1], because, unlike its progenitors (transcription activator-like effector nucleases (TALEN), zinc finger nucleases (ZFN) [2], and meganucleases [3]) it is easy to use and optimize for the research needs, it is sufficiently effective and simple. In the course of editing with the standard CRISPR/Cas9 method, a single guide RNA (sgRNA) that forms a ribonucleoprotein complex with the Cas9 nuclease binds to the DNA molecule at the editing locus via complementary base pairing, and next Cas9 introduces a double-strand break to the DNA molecule at a precise location of 3 nucleotides (nt) from the protospacer adjacent motif (PAM), which is followed by the repair according to one of the main mechanisms. The first mechanism – non-homologous end joining (NHEJ) – is a more mutagenic one, as it could result in undesirable insertion and deletions of nucleotides in the process of DNA repair, and also could be the cause of chromosome rearrangements. The second mechanism – homology-directed repair (HDR) – allows introducing specific changes to the genome. In order to activate this mechanism introduction of an additional molecular template (donor DNA molecule) carrying desired changes is required. However, even in the case of presence of such molecule,

* To whom correspondence should be addressed.

NHEJ is the predominant repair pathway. Low efficiency of HDR and high percent of undesirable changes do not allow using CRISPR/Cas9 for gene therapy involving directed correction of pathological variants. Moreover, CRISPR/Cas9 could exhibit non-target activity, which results in formation of double strand DNA breaks outside of the editing locus. Such undesirable effects as chromosome rearrangements and additional mutations also limit use of the CRISPR/Cas9 system for the gene therapy development [4].

New variants of editing platforms based on modified CRISPR/Cas9 technology are being developed by the researchers to enhance efficiency of editing and to avoid double strands brakes in the DNA molecule [5]. One of such modification, prime editing (PE), has been chosen for this study, which is composed of several required components (Fig. 1):

(i) nCas9 nickase – a mutant form of the wild type Cas9 nuclease, it introduces a single-strand break to the DNA molecule instead of a double-strand one [6];

(ii) MMLV-RT – Moloney murine leukemia virus reverse transcriptase fused with Cas9 nickase through a linker;

(iii) prime editing guide RNA (pegRNA).

During prime editing guide RNA (pegRNA) interacts with the DNA molecule at the targeted editing site, Cas9 nickase introduces a single-strand break at a distance of 3 nt from PAM, one strand is released and binds complementary to the primer binding site (PBS) in the pegRNA, and next the strand is completed with the help of reverse transcriptase (RT) based on the reverse transcriptase template (RTT) with the desired change [5]. HDR and NHEJ repair mechanisms are not involved in the mechanism of prime editing, and the site of pegRNA serves as a molecular template, hence, it could be assumed that the efficiency of genomic DNA correction should be higher. Moreover, considering that only single-strand DNA break is introduced in the course of PE, safety of this approach is potentially higher than in the case of classic variant of CRISPR/Cas9.

After the PE discovery various modification of the system have been developed to increase efficiency of introduction of changes. The following modifications have been selected for this study: PE2-NG – second generation of prime editing using nickase that recognizes PAM NG, as well as PEmax modification – improved PE2 editor with nickase recognizing PAM NGG.

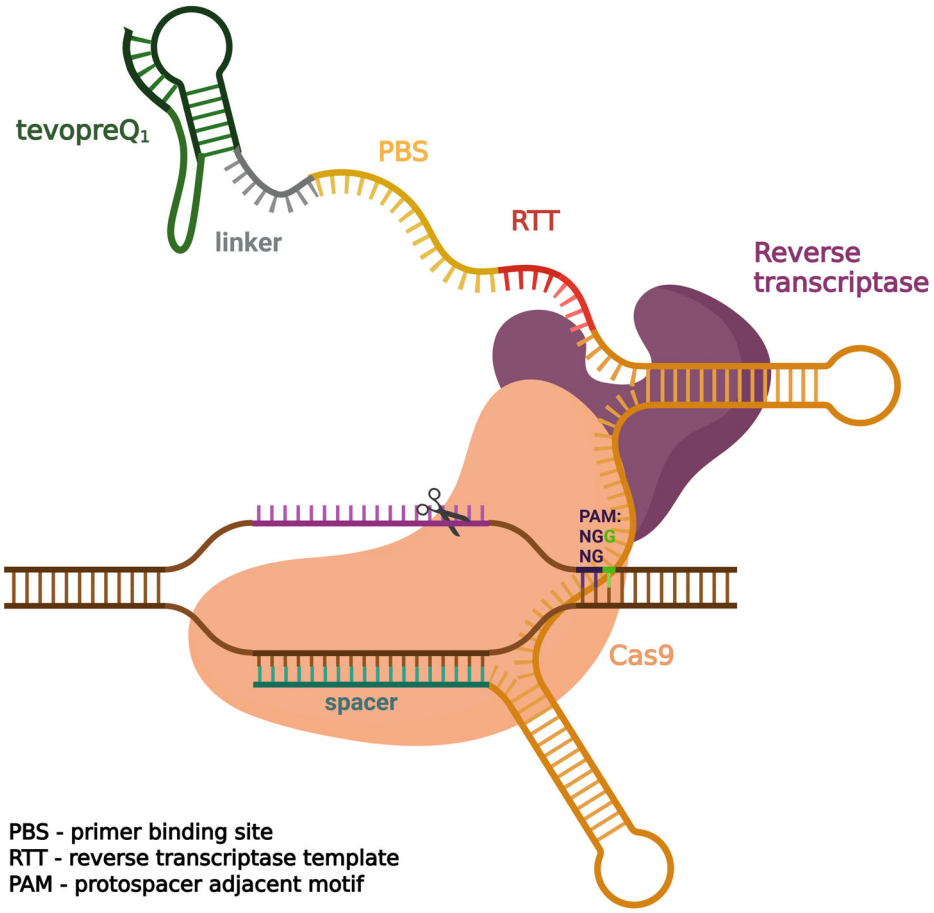

Fig. 1. Ribonucleoprotein complex for prime editing, details in the text.

pegRNA plays an important role in prime editing, because both efficiency and accuracy of the method depend on the selected sequence. pegRNA is composed from a scaffold, spacer site, PBS, RTT, and protective hairpin structure (tevopreQ$_1$) at the 3′-end connected to the construct via the linker (Fig. 1).

The spacer site is bound to the target DNA at the editing locus, its optimal length is 20 nt with added G nucleotide at the 5′-end [7], and optimal GC-composition should be in the range from 20 to 80% [8]. The PBS fragment is involved in complementary binding of the single strand, which is released after introduction of a single-strand break to the DNA; in the case of PBS not binding to the chain, the changes will not be introduced. RTT contains the desired change, the single-strand DNA is synthesized on its basis.

The 3′-end of pegRNA containing PBS and RTT sites is prone to degradation by exonucleases, hence, the pegRNA design includes the tevopreQ$_1$ hairpin at the 3′-end. The hairpin is connected to the main sequence of pegRNA through a linker. When selecting the linker sequence, it is important to avoid its potential interactions with PBS and spacer site of pegRNA [9].

Both length and composition of the pegRNA fragments are important. It could be suggested based on the literature data that editing would be successful in the case of GC content above 20%. Selection of pegRNA for AT-rich regions (GC-content below 40%) is challenging because at present there are no recommendation in the literature on selection of pegRNAs for these types of loci. This study is devoted to prime editing of such region. The pathological genetic variant F508del in the gene of cystic fibrosis transmembrane conductance regulator (*CFTR*) was selected for this study, which is located at the AT-rich region of the genome. This variant is associated with cystic fibrosis (CF) – prevalent monogenic autosomal recessive disease; average incidence rate in Russia is 1 : 10,250 [10], in Europe – from 1 : 3000 to 1 : 6000 of newborns [11].

The most common pathogenic variant of the *CFTR* gene in the world is the 3-nucleotide deletion F508del; in the European population this mutation was detected in 70% of alleles of the patients of European ancestry with CF [12], while in Russia frequency of this allele is 51.6% [13]. F508del causes disruption of the process of folding of the trans-membrane protein transporter of chloride ions CFTR resulting in lack of protein transport to apical membrane and its degradation by active proteases. Ion disbalance leads to disruption of functions of many organs, however, accumulation of thick mucus in the respiratory airways is the most serious manifestation of this disorder, which facilitates development of bacterial and viral infections [14] that eventually result in fibrotic changes in lungs and respiratory failure causing death of the CF patients. Median life expectancy of the patients in Europe is 51.7 years [15], and in Russia – 25 years [16]. At present pathogenetic treatment of this disease has been developed, but the therapeutics have serious side effects, are not suitable for all patients, and they must be taken for life. Hence, development of causal therapy for CF is important [17-19].

As has been mentioned above, selection of pegRNAs for AT-rich genome regions potentially containing pathogenic variants associated with serious diseases for which development of gene therapy is vital, such as F508del variant in the *CFTR*, is especially challenging. Furthermore, it is impossible to predict most optimal pegRNA *in silico* without experimental testing. The goal of this study was selection of most effective conditions for correction of the F508del variant in the *CFTR* gene. The most optimal pegRNAs for correction of the F508del mutation have been identified in the course of *ex vivo* experiments.

MATERIALS AND METHODS

Cell culture. Basal cells (BCs) of airway epithelium derived from the induced pluripotent stem cells (IPSCs) of the patients with CF with homozygote pathogenic variant F508del in the *CFTR* gene (patients P1 and P7) were used in the experiments [20, 21]. All participant or their legal representatives provided voluntary informed consent to participate in the study, which was approved by the ethics committee of the Research Centre for Medical Genetics (protocol no. 1, January 28, 2016). Protocol for BCs generation has been described previously [22]. Medium for cultivation of BCs cells consisted of a PneumaCult™-Ex Plus Medium (StemCell Technologies, Canada) supplemented with 1 μM A83-01 (Tocris, United Kingdom) and 1 μM DMH1 (Tocris).

Selection and synthesis of oligonucleotides, generation of a plasmid for editing. Twenty-four variants of pegRNA were selected for prime editing of F508del mutation (Table S1 in the Online Resourse 1) differing in variable sites (Table 1). Selection was carried out with the help of the internet resource PE Designer (CRISPR RGEN Tools) [8]. Sequences of pegRNAs included modified 3′-end with the tevopreQ$_1$ site (5′-CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA-3′) for their protection from degradation; the protective fragment is connected with each of the pegRNA via a unique linker. Linkers for each of 24 pegRNAs were selected using the internet resource pegLIT [9]. Total length of pegRNA after modification was more than 170 nt, while the standard amidophosphite method of oligonucleotide sequence synthesis allows error-free synthesis of molecules with length no more than 130 nt, hence, it was decided to divide each plasmid insert into two parts for cloning. For each pegRNA

2 pairs of oligonucleotides were synthesized (Evrogen, Russia), which were complementary to each other for further annealing, as well as they contained protruding ends for ligation both between each other and into the plasmid at the restriction sites. To assembly plasmids containing pegRNA, the pU6-pegRNA-GG-acceptor plasmid was used [a gift from David Liu (Addgene plasmid #132777; http://n2t.net/addgene:132777; RRID:Addgene_132777)]. The plasmid contained the gene of red fluorescent protein (*mRFP*), which was deleted via treatment of the plasmid with the restriction endonuclease Bso31 I (Sibenzim, Russia). After purification the plasmid backbone was ligated with pegRNA using T4-ligase (New England Biolabs, USA); prior to that the complementary parts of pegRNA were annealed in an Eppendorf 5332 Mastercycler Personal PCR Thermal Cycle (Eppendorf, Germany) and treated with a T4 polynucleotide kinase (T4 PNK, New England Biolabs) (Fig. 2). Presence of mRFP allows selection of the clones with successful insertion, because the colonies with initial plasmid were red, while the colonies with specific insertion were colorless. The obtained ligase mixture was used for transformation *Escherichia coli* using a standard protocol. Testing of colonies for successful insertion of pegRNA was carried out using PCR (primers: saCas9-sgRNA-seq-F – 5′-TGGACTATCATATGCTTACCG-3′; mRFP.R – 5′-GTACCTCGAGCGGCCCA-3′). PCR results were confirmed using agarose gel electrophoresis. After confirming correct length of the fragment, samples were subjected to Sanger sequencing using the same primers to verify sequence of the insertion.

Selection of modifications for prime editing. At present several platforms have been developed for PE. The PEmax platform (pCCF-PEmax plasmid was a kind gift from David Liu (Addgene plasmid #174820; http://n2t.net/addgene:174820; RRID:Addgene_174820) [23] and the PE2-NG platform (pCCF-PE2-NG plasmid was a kind gift from Yongsub Kim (Addgene plasmid #159977; http://n2t.net/addgene:159977; RRID:Addgene_159977) [18] were selected for editing the F508del variant of the *CFTR* gene. The PE2 system, unlike the original PE1, contains several mutations in the M-MLV RT for increasing efficiency, and the PE2-NG contains Cas9 recognizing PAM NG [24]. The PEmax system has been developed based on PE2, this variant includes NLS sequences, additional mutations in the SpCas9 nickase, and codon-optimized reverse transcriptase. For screening 24 pegRNA were selected: 15 – for PEmax and 9 – for PE2-NG.

Electroporation of airway BCs. Electroporation of two lines of airway BCs homozygous for the F508del variants of the *CFTR* gene derived from the patients P1 and P7 was carried out using a Neon™ Transfection System (Thermo Fisher Scientific, USA). For this purpose, BCs were detached from a polymer surface using a Versene solution (PanEko, Russia), cells were counted using an automatic cell counter Countess II (Thermo Fisher Scientific) and centrifuged at 150*g* for 5 min. Cell deposit was resuspended in an Opti-MEM™ medium (Gibco™, USA) to concentration 0.5×10^6 cells per 100 µl. Plasmids were added to cell suspension at the amount 400 ng per 0.5×10^6 cells: 200 ng of the plasmid with editor and 200 ng of the plasmid with pegRNA. The following controls were used: plasmid encoding PEmax without pegRNA (400 ng per well); mix plasmid encoding PEmax with non-target pegRNA (totally 400 ng per well); plasmid encoding PE2-NG without pegRNA (400 ng per well); non-transfected control. Electroporation of the suspension of BCs and plasmid was carried out in a 10-µl volume with two 20-ms pulses at voltage 1290 CF. Next cells were placed

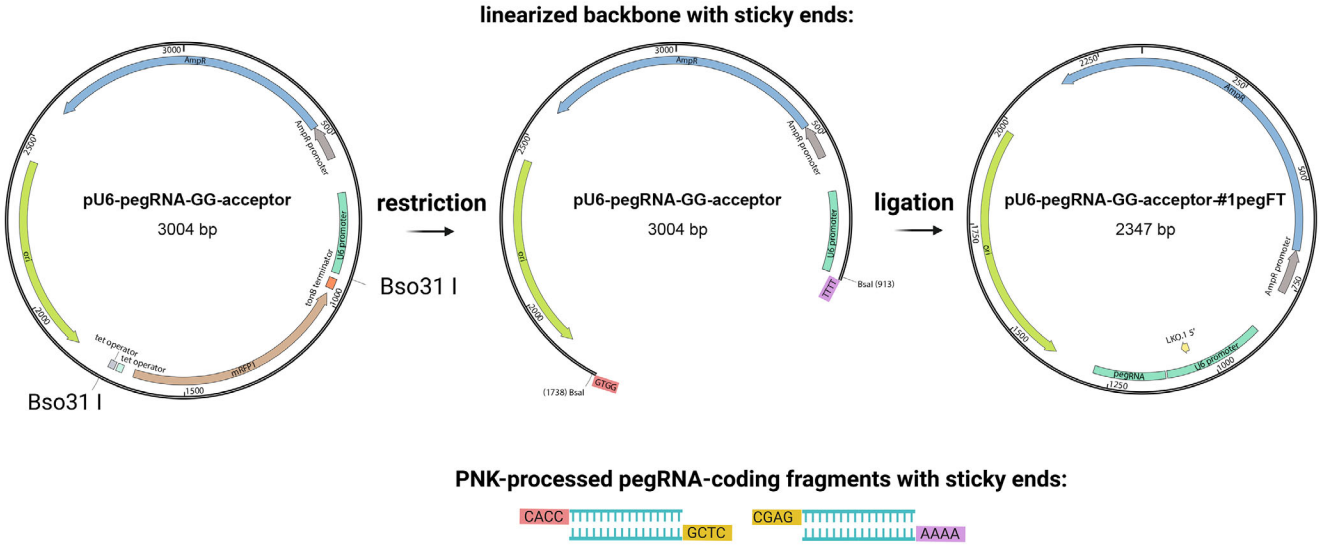

Fig. 2. Scheme of plasmid assembly with pegRNA.

Table 1. Variable regions of pegRNA for prime editing

No.	Spacer site – site hybridizing with DNA molecule (5′ → 3′)	RTT – reverse transcription template (5′ → 3′)	PBS – primer binding site (5′ → 3′)	Linker sequence for tevopreQ$_1$ (5′ → 3′)
#1	ACCATTAAAGAAAATATCAT	ACCAAAGATG	ATATTTTCTTTA	TACCAACT
#2	ACCATTAAAGAAAATATCAT	ACACCAAAGATG	ATATTTTCTTTA	CTAACTTT
#3	ACCATTAAAGAAAATATCAT	AACACCAAAGATG	ATATTTTCTTTA	CCCACCGA
#4	ACCATTAAAGAAAATATCAT	GAAACACCAAAGATG	ATATTTTCTTTA	CCACCCAC
#5	ACCATTAAAGAAAATATCAT	AGGAAACACCAAAGATG	ATATTTTCTTTA	CACATACG
#6	ACCATTAAAGAAAATATCAT	ACCAAAGATG	ATATTTTCTTTAA	CTACAACA
#7	ACCATTAAAGAAAATATCAT	CACCAAAGATG	ATATTTTCTTTAA	CTACCCAG
#8	ACCATTAAAGAAAATATCAT	AACACCAAAGATG	ATATTTTCTTTAA	ACCCTATT
#9	ACCATTAAAGAAAATATCAT	GAAACACCAAAGATG	ATATTTTCTTTAA	CCCAACGC
#10	ACCATTAAAGAAAATATCAT	AGGAAACACCAAAGATG	ATATTTTCTTTAA	CATAAGCA
#11	ACCATTAAAGAAAATATCAT	ACCAAAGATG	ATATTTTCTTTAAT	ACTAATAG
#12	ACCATTAAAGAAAATATCAT	CACCAAAGATG	ATATTTTCTTTAAT	ACTAATAG
#13	ACCATTAAAGAAAATATCAT	AACACCAAAGATG	ATATTTTCTTTAAT	ACTAATAG
#14	ACCATTAAAGAAAATATCAT	GAAACACCAAAGATG	ATATTTTCTTTAAT	TATACCAA
#15	ACCATTAAAGAAAATATCAT	AGGAAACACCAAAGATG	ATATTTTCTTTAAT	ACTAACAG
#16	ATTATGCCTGGCACCATTAA	AAAGATGATATTTTCTTTA	ATGGTGCCAGGCAT	AGAATAGT
#17	CCATTAAAGAAAATATCATT	GGAAACACCAAAGAT	GATATTTTCTTT	CTACCAAG
#18	CCATTAAAGAAAATATCATT	GGAAACACCAAAGAT	GATATTTTCTTTA	CTCAAACA
#19	CCATTAAAGAAAATATCATT	GGAAACACCAAAGAT	GATATTTTCTTTAA	CCAATAAC
#20	TTCATCATAGGAAACACCAA	ATCATCTTTG	GTGTTTCCTATG	ATCGTAAT
#21	TTCATCATAGGAAACACCAA	TATCATCTTTG	GTGTTTCCTATG	TATAATTA
#22	TTCATCATAGGAAACACCAA	AATATCATCTTTG	GTGTTTCCTATG	TCTCAGTC
#23	TTCATCATAGGAAACACCAA	AAAATATCATCTTTG	GTGTTTCCTATG	CTCCTTCC
#24	TTCATCATAGGAAACACCAA	AGAAAATATCATCTTTG	GTGTTTCCTATG	CCTTTAAT

onto a polymer surface in a medium for BCs. Efficiency of transfection was evaluated indirectly based on the number of GFP$^+$-cells transfected with an *eGFP* (AAT-PB-CG2APtk F508del) plasmid with the size of 10,067 bp in a separate well using a CytoFLEX S flow cytometer (Beckman Coulter, USA) 72 h after electroporation. Indirect evaluation is required, because there was no possibility to add reporter gene into the plasmid with editor due to its size, because increase of the plasmid size could significantly reduce efficiency of both transfection and editing.

Lipofection of airway BCs. Lipofection of two lines of BCs, P1 and P7, was carried out with the help of a commercial Lipofectamine LTX kit (Thermo Fisher

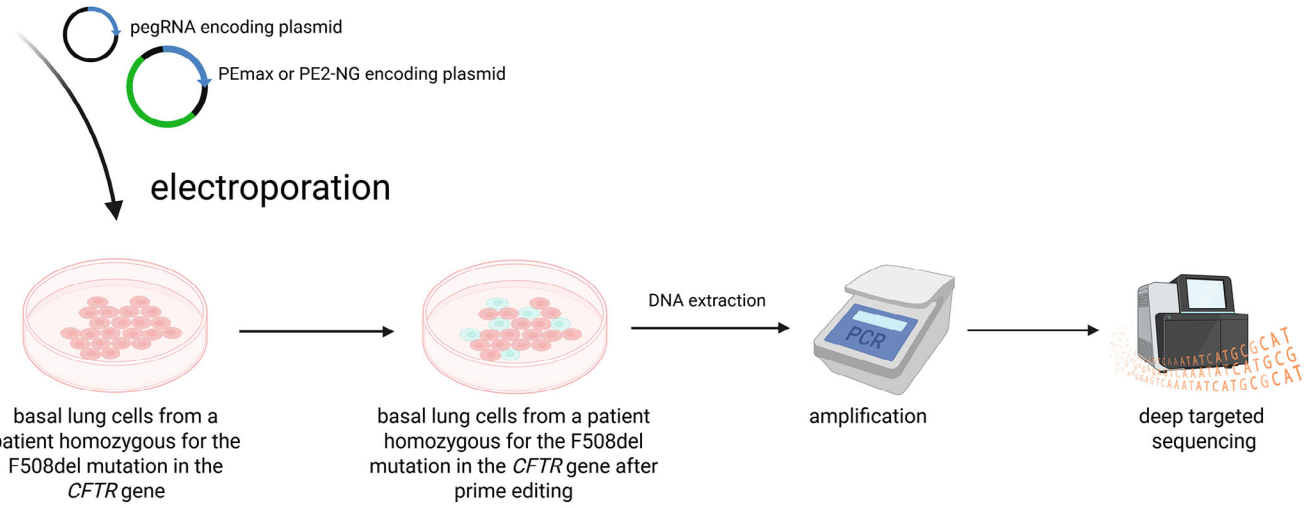

Fig. 3. Scheme of experiments for screening of 24 pegRNAs in airway BCs from two patients with homozygous F508del variant of the *CFTR* gene.

Scientific). One day prior to lipofection BCs were passaged into 24-well plates (SPL Lifesciences, Korea) at the amount of 40×10^3 cells per well in the medium for BC. Transfection conditions: 2.5 µl of a Lipofectamine LTX and 700 ng plasmid DNA (570 ng of the plasmid with editor and 130 ng of the plasmid with pegRNA) were added to a well (40×10^3 cells). Efficiency of transfection was evaluated as described above.

Evaluation of editing efficiency. DNA for amplification and deep targeted sequencing of the *CFTR* fragment with NGS (next-generation sequencing) was isolated using phenol-chloroform extraction technique according to the standard protocol followed by amplification (primers: 5F – 5′-TGGAGCC TTCAGAGGGTAAAAT-3′ and 8R – 5′-TGGCATGCTTTGA TGACGCT-3′). Next the obtained amplicons were subjected to deep targeted sequencing at the Center for Collective Use "Genome" at the Research Centre for Medical Genetics. Deep targeted sequencing was carried out with a new-generation sequencer NextSeq500 (Illumina, USA) using pair-end sequencing reads method (2×150 bp). A library preparation kit for amplicon sample SG GM Ampli (Raissol, Russia) was used for library generation according to the manufacturer's protocol. Number of cycles for final amplification was reduced to 4. Primers for library preparation with Illumina platform were used for dual indexing. Concentration of the produced libraries was measured with the help of a Qubit 2.0 fluorimeter (Invitrogen, USA) using a Qubit dsDNA HS Assay kit (Invitrogen). Processing of sequencing data was carried out with the NGSdata program developed in the department of Bioinformatics of the Research Centre for Medical Genetics (registration no. 2021662113). The obtained results were analyzed with the help of CRISPResso2 program [25].

Statistical data processing. Dunn's test was used for statistical processing of the data using the GraphPad Prism program. Differences with negative control were considered significant at $p < 0.05$.

RESULTS

Screening of 24 variants of prime editing of airway BC. In order to correct pathogenic variant F508del in the *CFTR* gene the PE method was chosen, which is a genome editing method that allows introducing a broad range of changes into the DNA molecule not involving generation of a double-strand DNA break. The desired correction suggests introduction of a 3-nucleotide insertion at the site of deletion. F508del is located in the AT-rich locus, which complicates selection of pegRNA for prime editing. Screening of 24 pegRNAs differing in composition of spacer sequences, PBS, and RTT, as well as in lengths of PBS and RTT was carried out to select most efficient sequences. From those, 15 pegRNAs were designed for PEmax, which is a PE variant containing Cas9 nickase recognizing PAM NGG, and 9 – for PE2-NG, which involves Cas9 nickase recognizing PAM NG. Experiments were performed with airway BCs derived from two patients with homozygous F508del variant of the *CFTR* gene – P1 and P7. Plasmids with the editor and pegRNA were co-transfected using electroporation, 72 h after transfection DNA was isolated, targeted region was amplified, and deep targeted sequencing was performed (Fig. 3). The sequencing data were analyzed with the help of online resource CRISPResso2, and statistical analysis was performed using the GraphPad Prism9 program.

First, the initial screening was carried out – two experiments with airway BCs from two patients using

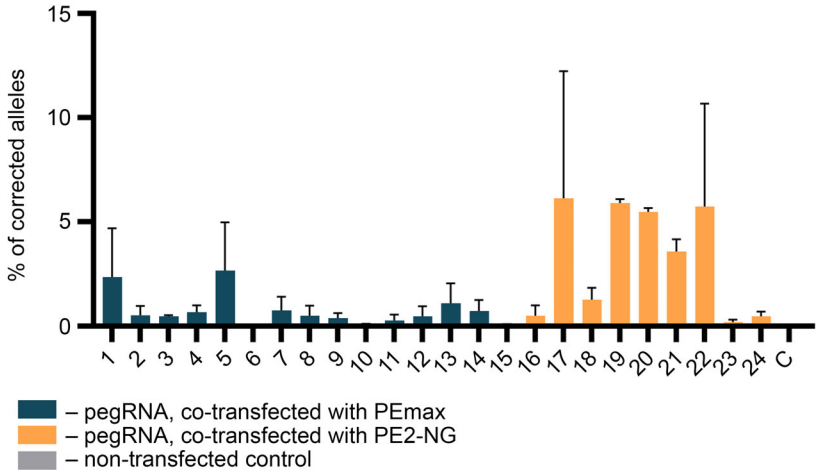

Fig. 4. Efficiency of correction of the pathogenic variant after prime editing using 24 pegRNAs in the airway BCs from two patients with homozygous F508del variant of the *CFTR* gene. Genetic constructs were delivered with electroporation. Average efficiency of correction of the pathogenic variant in two experiment is shown in the plot. Error bars for Y-axis represent standard error of the mean; C, non-transfected control.

all 24 variants of pegRNA. The results of NGS revealed that the pegRNA1 and pegRNA5 are potentially most effective for PEmax, and pegRNA17, pegRNA19, pegRNA20, pegRNA21, and pegRNA22 – for PE2-NG: average efficiency of editing in two experiments was from 2.4 to 6.1% without accounting for transfection efficiency (Fig. 3), which was 25.4% on average. Fractions of the alleles with the desired CTT insertion to the region of the F508del in the *CFTR* gene are presented in Fig. 4; in the process this insertion could be accompanied with introduction of additional non-target changes in the same allele. It could be seen from the data presented in Fig. 4 that there is high variability of editing in the case of electroporation, which, likely, could be explained by high variability of plasmid delivery with the help of this method.

Testing of most effective variants of pegRNAs. Further experiments were carried out using pegRNAs shown to be most effective in the screening experiments. Two biological and three technical replicates were performed in the experiment with the BCs from the patients P1 and P7. Plasmids with the editor and pegRNA were co-transfected using lipofection technique. In the case of PE2-NG the highest efficiency of correcting pathogenic variant was observed for pegRNA19, pegRNA20, pegRNA21, and pegRNA22; it was on average 2.50% (p = 0.03), 2.81% (p = 0.0005), 2.06% (p = 0.01), and 1.79% (p = 0.03), respectively. All pegRNAs selected for the PEmax platform did not demonstrate high efficiency, and differences between the efficiencies of correction of the F508del variant in the *CFTR* gene in the test sample and in the control were not statistically significant ($p > 0.05$). The differences between the correction efficiency of this variant for each pegRNA with other pegRNAs were statistically insignificant ($p > 0.05$) (Fig. 5). The data are shown without adjustment for the transfection efficiency, on average it was 11.9%.

Evaluation of undesirable changes in the editing locus. All undesirable changes in the 20-nt window around the editing locus were assessed in the experiment, sequences of 10-nt before and after the

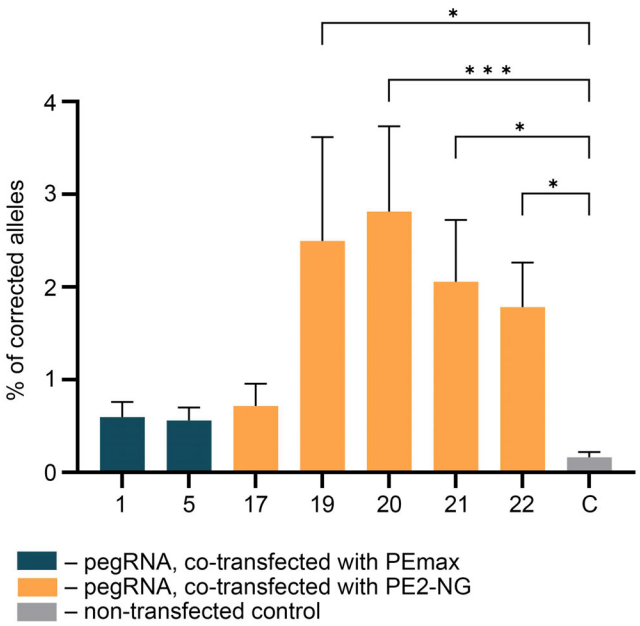

Fig. 5. Efficiency of correction of the pathogenic variant using prime editing in the airway BCs from two patients with homozygous F508del variant of the *CFTR* gene with 7 most effective pegRNAs. Average efficiency of editing for each pegRNA calculated based on the combined data from two biological and three technical replicates of the experiment with airway BCs from two patients is shown on the graph. Dunn index was used for statistical analysis. Y-axis error bars show standard error of the mean; C, non-transfected cells; * $p < 0.05$; *** $p < 0.001$.

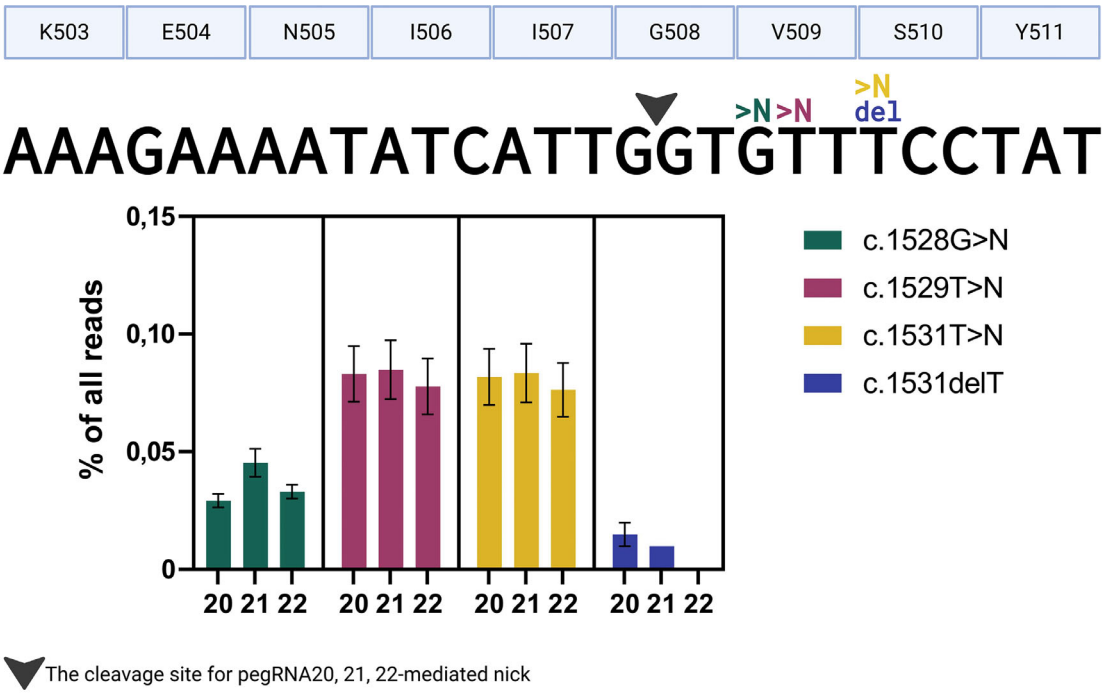

Fig. 6. Evaluation of non-target changes in DNA after prime editing in the samples without targeted CTT insertion. The analyzed DNA sequence (with F508del variant) is shown and corresponding amino acid sequence (amino acid is shown together with its position in the polypeptide chain) as well as site of introduction of single-strand break in the process of editing by pegRNA20, -21, and -22. Distribution of single-nucleotide substitutions c.1528G>N, c.1529T>N, c.1531T>N, and c.1531delT is shown on the graph, which were introduced during editing using pegRNA20, -21, and -22. Y-axis error bars show standard error of the mean.

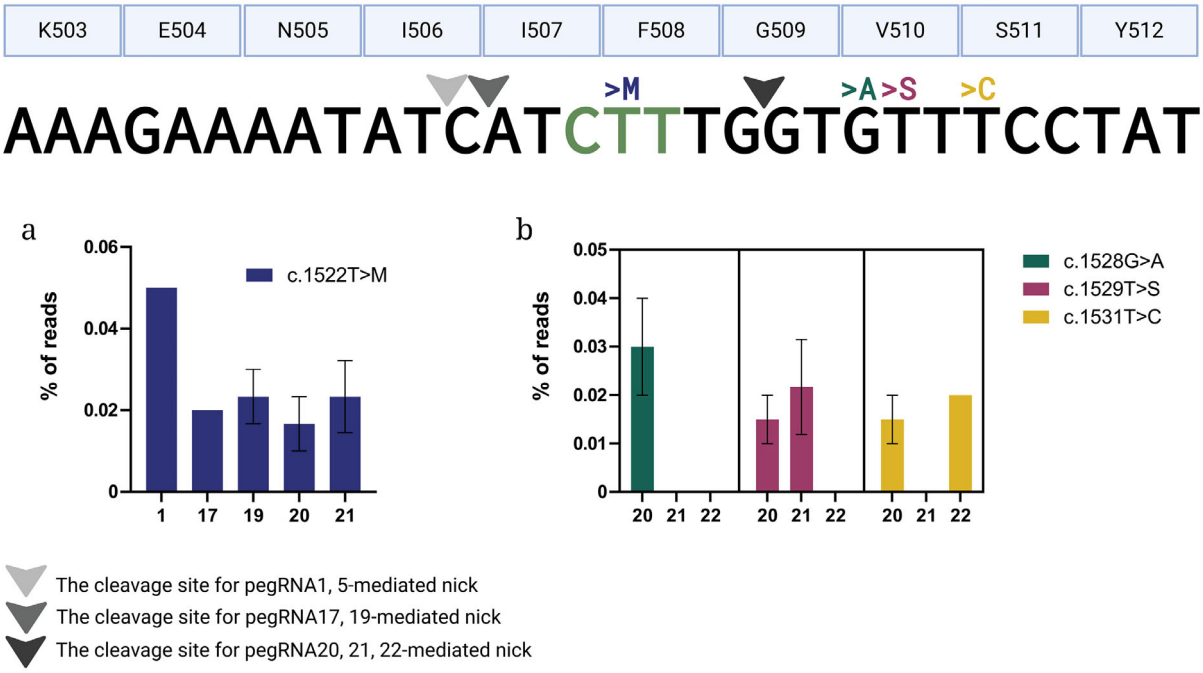

Fig. 7. Evaluation of non-target changes in DNA after prime editing of the samples with CTT targeted insertion. The analyzed DNA sequence is shown (without the F508del variant), as well as corresponding amino acid sequence (amino acid is shown together with its position in the polypeptide chain) and sites of introduction of single-strand breaks in the process of editing by various pegRNAs. Y-axis error bars show standard error of the mean. a) Introduction of changes into the target insertion during edition with the help of pegRNA1, -17, -19, and -20. b) Distribution of single-nucleotide substitutions c.1528G>A, c.1529T>S, and c.1531T>C introduced during editing with the help of pegRNA20, -21, and -22.

single-strand break site were analyzed. In most of the cases, frequency and spectrum of the introduced changes were the same for the transfected and control samples, although some exceptions were observed.

Non-target changes that are absent in the non-transfected control were revealed in the samples subjected to PE without the target CTT insertion that were transfected with pegRNA20, -21, and -22. These pegRNAs contain identical spacer sequence and, correspondingly, the same site of single-strand break. Among the non-target changes not only synonymous substitutions (c.1530T>N) were observed, but also variants with undefined significance: c.1528G>N observed with average frequencies 0.03%, 0.05%, and 0.03%; c.1529T>N observed with average frequencies 0.08%, 0.08%, and 0.08%; c.1531T>N also observed with average frequencies 0.08%, 0.08%, and 0.08% from total number of reads in the samples edited with the help of pegRNA20, -21, and -22, respectively. In addition, a single-nucleotide probably pathogenic deletion c.1531delT was observed with mean frequency of 0.02% in the samples edited with pegRNA20, and 0.01% – in the samples edited with pegRNA21 (Fig. 6).

The variants with undefined significance were also observed in the samples with successful correction edited with the help of pegRNA20, pegRNA21, and pegRNA22: c.1528G>A with average frequency 0.04% for pegRNA20; c.1529T>S with average frequencies 0.02% and 0.02% for pegRNA20 and -21, respectively; c.1531T>C with average frequencies 0.02% and 0,02% for pegRNA20 and -22, respectively. The fraction was calculated with respect to total number of reads.

The substitutions c.1522T>M in the insert itself also were observed in the cases of editing with the help of pegRNA1, -17, -19, -20, and -21, their frequencies were on average 0.05%, 0.02%, 0.02%, 0.02%, and 0.04% of total number of reads, respectively. These variants were also of unidentified significance (Fig. 7).

DISCUSSION

The F508del variant of the *CFTR* gene is located in the AT-rich genome region, which complicates its effective correction with the help of prime editing. That is why it was important to identify effective editing systems allowing selection of the largest number of pegRNAs. The PEmax and PE2-NG systems were selected for the study. Both variants introduce only one break to DNA, unlike in the case of more effective PE3 system involving introduction of two breaks, which is accompanied by the significant increase of the frequency of undesired insertions and deletions at the editing locus [26, 27]. Moreover, both these variants do not use supplementary proteins affecting repair of the introduced breaks, hence, they have smaller size (and higher efficiency of intracellular delivery) than, for example, more effective system PE4 [23]. The PEmax editor contains nCas9 recognizing PAM NGG, which limits the possibilities of pegRNA selection. Hence, the PE2-NG editor, which demonstrates lower efficiency in this system, was also selected [23], however, nCas9 in this case recognizes PAM NG, which allows selection of other pegRNAs with higher GC-content, which proved to be very important in this study, because these variants demonstrated the highest efficiency.

In all pegRNAs selected for the PEmax editor content of GC nucleotides in the spacer composition was 20%, and in the PBS composition – was below 10%. It was possible to select guide RNAs for the PE2-NG system with higher GC-content in both spacer and PBS fragments. The highest efficiency was observed with the pegRNA20 (2.81%), in this case GC content of the spacer was 35%, and of the PBS – 42%. We suggest that the higher efficiency could be associated with the increase of GC content in the variable regions of pegRNAs.

Experiments in this study involved two steps. In the first step initial screening of the pegRNAs was carried out, which allowed selecting the most effective pegRNAs for each editor. In this step transfection was performed with the help of electroporation. However, the obtained results (efficiency of transfection and editing) were not sufficiently reproducible between the technical and biological replicates, hence, the following experiments were carried out using lipofection as more reproducible technique [28]. As a result, the most effective pegRNA for correction of the pathogenic F508del variant of the *CFTR* gene in the AT-rich genome region was identified. Efficiency of this mutation correction with the help of pegRNA20 reached 2.81%. Despite the fact that the exact efficiency of transfection of the plasmid with the genome editor is unknown, based on the indirect data it is predicted to be 11.9%, and, thus, it could be expected in this case that the efficiency of editing under optimal conditions is 23.6% of all alleles. Furthermore, it is known from the literature data that 6-10% of the CFTR-expressing cells is sufficient for normalization of chlorine ion transport [29], hence, we assume that if the efficiency of pegRNA20 delivery could be increased, this approach could a form a basis for the development of CF therapy.

At present, delivery of the prime editing system is one of the main limitations of this method. Its low efficiency is determined by the large size of the prime editor. In particular, the size of plasmid construct of the editor is up to 10 kb, as it includes Cas9 nickase consisting of approximately 4100 nt, reverse transcriptase with size ~2000 bp, as well as other required sequences including regulatory sequences, origin of replication, and antibiotic resistance cassette [26, 30, 31]. In the case of airway BCs, as has been noted by

Bulcaen et al. [32], the delivery is even more complicated due to existence of the cell barriers preventing penetration of genetic constructs. In the cited study prime editing was used for correction of the pathogenic variants L227R and N1303K in the *CFTR* gene. The authors reported restoration of CFTR function in the intestine organoids and in the primary cells of nasal epithelium, but there were complications with the delivery of this system to the airway cells.

In one of the studies devoted to correction of the W1282X mutation in the *CFTR* gene in the airway BCs with the help of PE the authors selected a helper-dependent adenovirus, which demonstrated high efficiency of transduction, however, correction of W1282X was only 2.4 ± 0.6% [33].

In one of the studies the authors demonstrated significant results in correction of the F508del variant in the *CFTR* gene. Sousa et al. reported [34] that the PE6 system in combination with improved pegRNAs, additional guide RNAs, and MLH1dn protein demonstrated efficiency of correction up to 51% in the immortalized airway cells. High efficiency of transfection with electroporation also has been demonstrated in this study. Unfortunately, not all of the used conditions could be used for further development of gene therapy for mucoviscidosis. Firstly, high percentage of introduced undesirable changes in the editing locus was observed – 13%. Likely this was due to the use of additional guide RNA and introduction of one more break to the DNA molecule, as well as due to suppression of the system of repair of noncomplementary nucleotides. Secondly, electroporation is not suitable for introduction of genetic constructs *in vivo* [34].

In some studies efficiency of prime editing was shown to be comparable with the CRISPR/Cas9 system, which contradicts the previously obtained data demonstrating higher efficiency of the PE system [35]. Nevertheless, PE remain a preferable method for introduction of targeted changes to nucleotide sequence, due to the fact that it is the safest. In the studies associated with correction of the F508del mutation in the *CFTR* gene with the help of CRISPR/Cas9, the standard system was shown to be ineffective and resulted in high percentage of undesirable changes in the editing locus because NHEJ was predominant repair pathway [36, 37]. Another limitation for the use of CRISPR/Cas9 system for development of gene therapy is the fact that the HDR mechanism is active only in dividing cells. And majority of airway cells are in the state of quiescence and are not dividing actively, which increases probability of the repair via NHEJ when CRISPR/Cas9 is used *in vivo* [38].

Prime editing, on the other hand, demonstrates very low percentage of undesirable changes in the editing locus, although it does not demonstrate high efficiency due to peculiarities of delivery. We were able to make such conclusion based on the results of analysis of frequencies of single-nucleotide substitutions, insertions, and deletions. In most of the cases frequency of such changes was the same as in the non-transfected controls and samples. The errors could appear and accumulate at the stage of amplification, as well as at the stage of deep targeted sequencing [39, 40]. The single-nucleotide substitutions not observed in the controls were mostly observed in the samples with unsuccessful insertion, and fraction of such changes did not exceed 0.08%. In the process only one single-nucleotide deletion was observed, and its fraction did not exceed 0.02% of total number of reads.

CONCLUSIONS

Despite the fact that prime editing is a very promising tool for development of gene therapy, its use could be limited in the AT-rich genome regions. In this study we conducted screening of 24 pegRNA in the airway BC from the patients with homozygous F508del variant of the *CFTR* gene. pegRNAs were selected for two different editors, PEmax and PE2-NG, and the most effective molecule, pegRNA20, was identified as a result. According to the literature data, the editor PE2-NG, for which this pegRNA was designed, is less effective editor than the PEmax, however, for correction of the AT-rich region flexibility of PAM was found to be more important.

Efficiency of correction of the pathogenic variant with the help of pegRNA20 was only 2.81% without consideration of transfection efficiency. This could be due to the fact that the genetic construct encoding the editor is large, and it is difficult to deliver it with the help of existing delivery systems. In order to develop mucoviscidosis therapy based on prime editing it is necessary to find effective and safe method for delivering large genetic constructs to the cell.

Abbreviations. BCs, basal cells; Cas, CRISPR-associated protein; CF, cystic fibrosis; *CFTR*, gene of cystic fibrosis transmembrane conductance regulator, mucoviscidosis transmembrane conductance regulator; CRISPR, system of clustered regularly interspaced short palindromic repeats; HDR, homology-directed repair; nt, nucleotide; NGS, next-generation sequencing; NHEJ, non-homologous end joining; PAM, protospacer adjacent motif; PBS, primer binding site; PE, prime editing; pegRNA, prime editing guide RNA; PEmax and PE2-NG, two different systems for prime editing; RT, reverse transcriptase; RTT, reverse transcriptase template.

Supplementary information. The online version contains supplementary material available at https://doi.org/10.1134/S0006297924604672.

Acknowledgments. The authors express their gratitude to Elena Lvovna Amelina (PhD, Scientific-Research Institute of Pulmonology, FMBA of Russia) for providing initial cell material from the patients. Images were created with the help of BioRender (https://www.biorender.com/).

Contributions. O. V. Volodina, A. A. Anuchina, A. V. Lavrov, and S. A. Smirnikhina – concept and supervision of the study; O. V. Volodina, A. G. Demchenko, A. A. Anuchina, and E. V. Kondrateva – conducting experiments; O. P. Ryzhkova and V. A. Kovalskaya – conducting deep targeted sequencing; E. V. Kondrateva, V. Yu. Tabakov and A. G. Demchenko – preparation of cell material; O. V. Volodina and S. A. Smirnikhina – discussion of the results of the study, writing text of the paper; S. A. Smirnikhina – procuring funding for the study.

Funding. This study was financially supported by the State Budget Project of the Ministry of Science and Higher Education of the Russian Federation. Work with pegRNA1, pegRNA5, pegRNA17, pegRNA19, pegRNA20, pegRNA21, pegRNA22 was financially supported by the LLC "MK Development" with information support provided by AVVA Pharmaceuticals Ltd.

Ethics approval and consent to participate. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Voluntary informed consent was obtained from all participants.

Conflict of interest. The authors of this work declare that they have no conflicts of interest.

REFERENCES

1. Xiao-Jie, L., Hui-Ying, X., Zun-Ping, K., Jin-Lian, C., and Li-Juan, J. (2015) CRISPR-Cas9: a new and promising player in gene therapy, *J. Med. Genet.*, **52**, 289-296, https://doi.org/10.1136/jmedgenet-2014-102968.
2. Gaj, T., Gersbach, C. A., and Barbas, C. F. (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering, *Trends Biotechnol.*, **31**, 397-405, https://doi.org/10.1016/j.tibtech.2013.04.004.
3. Tröder, S. E., and Zevnik, B. (2022) History of genome editing: from meganucleases to CRISPR, *Lab. Anim.*, **56**, 60-68, https://doi.org/10.1177/0023677221994613.
4. Uddin, F., Rudin, C. M., and Sen, T. (2020) CRISPR gene therapy: applications, limitations, and implications for the future, *Front. Oncol.*, **10**, 1387, https://doi.org/10.3389/fonc.2020.01387.
5. Kantor, A., McClements, M. E., and MacLaren, R. E. (2020) CRISPR-Cas9 DNA base-editing and prime-editing, *Int. J. Mol. Sci.*, **21**, 6240, https://doi.org/10.3390/ijms21176240.
6. Lee, J., Lim, K., Kim, A., Mok, Y. G., Chung, E., Cho, S.-I., Lee, J. M., and Kim, J.-S. (2023) Prime editing with genuine Cas9 nickases minimizes unwanted indels, *Nat. Commun.*, **14**, 1786, https://doi.org/10.1038/s41467-023-37507-8.
7. Chow, R. D., Chen, J. S., Shen, J., and Chen, S. (2021) A web tool for the design of prime-editing guide RNAs, *Nat. Biomed. Eng.*, **5**, 190-194, https://doi.org/10.1038/s41551-020-00622-8.
8. Hwang, G.-H., Jeong, Y. K., Habib, O., Hong, S.-A., Lim, K., Kim, J.-S., and Bae, S. (2021) PE-Designer and PE-Analyzer: web-based design and analysis tools for CRISPR prime editing, *Nucleic Acids Res.*, **49**, W499-W504, https://doi.org/10.1093/nar/gkab319.
9. Nelson, J. W., Randolph, P. B., Shen, S. P., Everette, K. A., Chen, P. J., Anzalone, A. V., An, M., Newby, G. A., Chen, J. C., Hsu, A., and Liu, D. R. (2022) Engineered pegRNAs improve prime editing efficiency, *Nat. Biotechnol.*, **40**, 402-410, https://doi.org/10.1038/s41587-021-01039-7.
10. Kiseleva, A., Klimushina, M., Sotnikova, E., Skirko, O., Divashuk, M., Kurilova, O., Ershova, A., Khlebus, E., Zharikova, A., Efimova, I., Pokrovskaya, M., Slominsky, P. A., Shalnova, S., Meshkov, A., and Drapkina, O. (2020) Cystic fibrosis polymorphic variants in a Russian population, *Pharmgenomics Pers. Med.*, **13**, 679-686, https://doi.org/10.2147/PGPM.S278806.
11. Scotet, V., L'Hostis, C., and Férec, C. (2020) The changing epidemiology of cystic fibrosis: incidence, survival and impact of the CFTR gene discovery, *Genes (Basel)*, **11**, 589, https://doi.org/10.3390/genes11060589.
12. Bobadilla, J. L., Macek, M., Fine, J. P., and Farrell, P. M. (2002) Cystic fibrosis: a worldwide analysis of CFTR mutations – correlation with incidence data and application to screening, *Hum. Mutat.*, **19**, 575-606, https://doi.org/10.1002/humu.10041.
13. Registry of Patients with Cystic Fibrosis in Russia [in Russian], URL: https://mukoviscidoz.org/mukovistsidoz-v-rossii.html.
14. De Boeck, K. (2020) Cystic fibrosis in the year 2020: a disease with a new face, *Acta Paediatr.*, **109**, 893-899, https://doi.org/10.1111/apa.15155.
15. McKone, E. F., Ariti, C., Jackson, A., Zolin, A., Carr, S. B., Orenti, A., van Rens, J. G., Lemonnier, L., Macek, M., Keogh, R. H., Naehrlich, L., and European Cystic Fibrosis Society Patient Registry (2021) Survival estimates in European cystic fibrosis patients and the impact of socioeconomic factors: a retrospective registry cohort study, *Eur. Respir. J.*, **58**, 2002288, https://doi.org/10.1183/13993003.02288-2020.
16. Amelina, E. L., Cherniak, A. V., and Cherniaev, A. L. (2001) Cystic fibrosis: current survival [in Russian], *Pulmonologiya*, **3**, 61-64.

17. Spoletini, G., Gillgrass, L., Pollard, K., Shaw, N., Williams, E., Etherington, C., Clifton, I. J., and Peckham, D. G. (2022) Dose adjustments of Elexacaftor/Tezacaftor/Ivacaftor in response to mental health side effects in adults with cystic fibrosis, *J. Cyst. Fibros.*, **21**, 1061-1065, https://doi.org/10.1016/j.jcf.2022.05.001.
18. Papadakis, L., Stander, T., Mombourquette, J., Richards, C. J., Yonker, L. M., Lawton, B., Hardcastle, M., Zweifach, J., Sicilian, L., Bringhurst, L., and Neuringer, I. P. (2025) Heterogeneity in reported side effects following initiation of elexacaftor-tezacaftor-ivacaftor: experiences at a quaternary CF care center, *Pediatr. Pulmonol.*, **60**, e27382, https://doi.org/10.1002/ppul.27382.
19. Safety and Possible Side Effects|TRIKAFTA® (elexacaftor/tezacaftor/ivacaftor and ivacaftor), URL: https://www.trikafta.com/safety-side-effects.
20. Kondrateva, E., Adilgereeva, E., Amelina, E., Tabakov, V., Demchenko, A., Ustinov, K., Yasinovsky, M., Voronina, E., Lavrov, A., and Smirnikhina, S. (2020) Generation of induced pluripotent stem cell line (RCMGi001-A) from human skin fibroblasts of a cystic fibrosis patient with p.F508del mutation, *Stem Cell Res.*, **48**, 101933, https://doi.org/10.1016/j.scr.2020.101933.
21. Kondrateva, E., Demchenko, A., Slesarenko, Y., Yasinovsky, M., Amelina, E., Tabakov, V., Voronina, E., Lavrov, A., and Smirnikhina, S. (2021) Derivation of iPSC line (RCMGi002-A) from dermal fibroblasts of a cystic fibrosis female patient with homozygous F508del mutation, *Stem Cell Res.*, **53**, 102251, https://doi.org/10.1016/j.scr.2021.102251.
22. Demchenko, A., Belova, L., Balyasin, M., Kochergin-Nikitsky, K., Kondrateva, E., Voronina, E., Pozhitnova, V., Tabakov, V., Salikhova, D., Bukharova, T., Goldshtein, D., Kondratyeva, E., Kyian, T., Amelina, E., Zubkova, O., Popova, O., Ozharovskaia, T., Lavrov, A., and Smirnikhina, S. (2024) Airway basal cells from human-induced pluripotent stem cells: a new frontier in cystic fibrosis research, *Front. Cell. Dev. Biol.*, **12**, 1336392, https://doi.org/10.3389/fcell.2024.1336392.
23. Chen, P. J., Hussmann, J. A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.-F., Chen, C., Nelson, J. W., Newby, G. A., Sahin, M., Osborn, M. J., Weissman, J. S., Adamson, B., and Liu, D. R. (2021) Enhanced prime editing systems by manipulating cellular determinants of editing outcomes, *Cell*, **184**, 5635-5652.e29, https://doi.org/10.1016/j.cell.2021.09.018.
24. Kweon, J., Yoon, J.-K., Jang, A.-H., Shin, H. R., See, J.-E., Jang, G., Kim, J.-I., and Kim, Y. (2021) Engineered prime editors with PAM flexibility, *Mol. Ther.*, **29**, 2001-2007, https://doi.org/10.1016/j.ymthe.2021.02.022.
25. Clement, K., Rees, H., Canver, M. C., Gehrke, J. M., Farouni, R., Hsu, J. Y., Cole, M. A., Liu, D. R., Joung, J. K., Bauer, D. E., and Pinello, L. (2019) CRISPResso2 provides accurate and rapid genome editing sequence analysis, *Nat. Biotechnol.*, **37**, 224-226, https://doi.org/10.1038/s41587-019-0032-3.
26. Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., and Liu, D. R. (2019) Search-and-replace genome editing without double-strand breaks or donor DNA, *Nature*, **576**, 149-157, https://doi.org/10.1038/s41586-019-1711-4.
27. Kim, D. Y., Moon, S. B., Ko, J.-H., Kim, Y.-S., and Kim, D. (2020) Unbiased investigation of specificities of prime editing systems in human cells, *Nucleic Acids Res.*, **48**, 10576-10589, https://doi.org/10.1093/nar/gkaa764.
28. Fajrial, A. K., He, Q. Q., Wirusanti, N. I., Slansky, J. E., and Ding, X. (2020) A review of emerging physical transfection methods for CRISPR/Cas9-mediated gene editing, *Theranostics*, **10**, 5532-5549, https://doi.org/10.7150/thno.43465.
29. Johnson, L. G., Olsen, J. C., Sarkadi, B., Moore, K. L., Swanstrom, R., and Boucher, R. C. (1992) Efficiency of gene transfer for restoration of normal airway epithelial function in cystic fibrosis, *Nat. Genet.*, **2**, 21-25, https://doi.org/10.1038/ng0992-21.
30. Ribeiro, S., Mairhofer, J., Madeira, C., Diogo, M. M., Lobato da Silva, C., Monteiro, G., Grabherr, R., and Cabral, J. M. (2012) Plasmid DNA size does affect nonviral gene delivery efficiency in stem cells, *Cell Reprogram.*, **14**, 130-137, https://doi.org/10.1089/cell.2011.0093.
31. Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M., and Danielsen, M. (1987) Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure, *Proc. Natl. Acad. Sci. USA*, **84**, 7413-7417, https://doi.org/10.1073/pnas.84.21.7413.
32. Bulcaen, M., Kortleven, P., Liu, R. B., Maule, G., Dreano, E., Kelly, M., Ensinck, M. M., Thierie, S., Smits, M., Ciciani, M., Hatton, A., Chevalier, B., Ramalho, A. S., Casadevall i Solvas, X., Debyser, Z., Vermeulen, F., Gijsbers, R., Sermet-Gaudelus, I., Cereseto, A., and Carlon, M. S. (2024) Prime editing functionally corrects cystic fibrosis-causing CFTR mutations in human organoids and airway epithelial cells, *Cell Rep. Med.*, **5**, 101544, https://doi.org/10.1016/j.xcrm.2024.101544.
33. Li, C., Liu, Z., Anderson, J., Liu, Z., Tang, L., Li, Y., Peng, N., Chen, J., Liu, X., Fu, L., Townes, T. M., Rowe, S. M., Bedwell, D. M., Guimbellot, J., and Zhao, R. (2023) Prime editing-mediated correction of the CFTR W1282X mutation in iPSCs and derived airway epithelial cells, *PLoS One*, **18**, e0295009, https://doi.org/10.1371/journal.pone.0295009.
34. Sousa, A. A., Hemez, C., Lei, L., Traore, S., Kulhankova, K., Newby, G. A., Doman, J. L., Oye, K., Pandey, S., Karp, P. H., McCray, P. B., and Liu, D. R. (2025) Systematic optimization of prime editing for the efficient

functional correction of CFTR F508del in human airway epithelial cells, *Nat. Biomed. Eng.*, **9**, 7-21, https://doi.org/10.1038/s41551-024-01233-3.

35. Averina, O. A., Permyakov, O. A., Grigorieva, O. O., Starshin, A. S., Mazur, A. M., Prokhortchouk, E. B., Dontsova, O. A., and Sergiev, P. V. (2021) Comparative analysis of genome editors efficiency on a model of mice zygotes microinjection, *Int. J. Mol. Sci.*, **22**, 10221, https://doi.org/10.3390/ijms221910221.
36. Bassai, L., Shigunov, P., and Soares, L. (2021) Correction of the mutation cause of cystic fibrosis (F508DEL) in pulmonary epithelius line cells BY CRISPR/CAS9, *Cytotherapy*, **23**, 8, https://doi.org/10.1016/j.jcyt.2021.02.028.
37. Smirnikhina, S. A., Kondrateva, E. V., Adilgereeva, E. P., Anuchina, A. A., Zaynitdinova, M. I., Slesarenko, Y. S., Ershova, A. S., Ustinov, K. D., Yasinovsky, M. I., Amelina, E. L., Voronina, E. S., Yakushina, V. D., Tabakov, V. Y., and Lavrov, A. V. (2020) P.F508del editing in cells from cystic fibrosis patients, *PLoS One*, **15**, e0242094, https://doi.org/10.1371/journal.pone.0242094.
38. Graham, C., and Hart, S. (2021) CRISPR/Cas9 gene editing therapies for cystic fibrosis, *Expert Opin. Biol. Ther.*, **21**, 767-780, https://doi.org/10.1080/14712598.2021.1869208.
39. Pfeiffer, F., Gröber, C., Blank, M., Händler, K., Beyer, M., Schultze, J. L., and Mayer, G. (2018) Systematic evaluation of error rates and causes in short samples in next-generation sequencing, *Sci. Rep.*, **8**, 10950, https://doi.org/10.1038/s41598-018-29325-6.
40. Ma, X., Shao, Y., Tian, L., Flasch, D. A., Mulder, H. L., Edmonson, M. N., Liu, Y., Chen, X., Newman, S., Nakitandwe, J., Li, Y., Li, B., Shen, S., Wang, Z., Shurtleff, S., Robison, L. L., Levy, S., Easton, J., and Zhang, J. (2019) Analysis of error profiles in deep next-generation sequencing data, *Genome Biol.*, **20**, 50, https://doi.org/10.1186/s13059-019-1659-6.