Mutation-specific reporter for optimization and enrichment of prime editing

Imre Schene (✉ i.f.schene-2@umcutrecht.nl)

Wilhelmina Children's Hospital, University Medical Center Utrecht

Indi Joore

University Medical Center Utrecht

Jan Baijens

Utrecht University Graduate School of Life Sciences <https://orcid.org/0000-0003-0397-8533>

Remi Stevelink

Utrecht University

Gautam Kok

Department of Metabolic Diseases, Wilhelmina Children's Hospital, University Medical Center Utrecht

Sawsan Shehata

Division of Pediatric Gastroenterology, Wilhelmina Children's Hospital, University Medical Center Utrecht

Eveline Ilcken

Department of Metabolic Diseases, Wilhelmina Children's Hospital, University Medical Center Utrecht

Dian Bolhuis

Department of Metabolic Diseases, Wilhelmina Children's Hospital, University Medical Center Utrecht

Rosalie van Rees

Department of Metabolic Diseases, Wilhelmina Children's Hospital, University Medical Center Utrecht

Sacha Spelier

Department of Pediatric Respiratory Medicine, Wilhelmina Children's Hospital, University Medical Center

Hubert van der Doef

University Medical Center Groningen

Jeffrey Beekman

Regenerative Medicine Center Utrecht

Roderick Houwen

Division of Pediatric Gastroenterology, Wilhelmina Children's Hospital, University Medical Center Utrecht

Edward Nieuwenhuis

University Medical Center Utrecht

Sabine A. Fuchs

University Medical Center Utrecht <https://orcid.org/0000-0001-9147-2406>

Keywords: fluoPEER, prime edit guide RNAs, protospacer adjacent motif recognition

Posted Date: September 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-563512/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on March 1st, 2022. See the published version at <https://doi.org/10.1038/s41467-022-28656-3>.

Mutation-specific reporter for optimization and enrichment of prime editing

I.F. Schene^{1,2,3*}, I.P. Joore^{1,2,3*}, J.H.L. Baijens⁴, R. Stevelink⁵, G. Kok^{1,2,3}, S. Shehata^{1,2,3}, E.F. Ilcken^{2,3}, D.P. Bolhuis^{2,3}, R.C.M. van Rees^{2,3}, S.A. Spelier^{3,6}, P.J. van der Doef⁷, J.M. Beekman^{3,6}, R.H.J. Houwen¹, E.E.S. Nieuwenhuis^{1,8}, S.A. Fuchs^{2,3,a}

*: these authors contributed equally

a: Lead contact, on behalf of United for Metabolic Diseases (UMD)

1: Division of Pediatric Gastroenterology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Lundlaan 6, 3584 EA Utrecht, the Netherlands

2: Department of Metabolic Diseases, Wilhelmina Children's Hospital, University Medical Center Utrecht, Lundlaan 6, 3584 EA Utrecht, the Netherlands

3: Regenerative Medicine Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands

4: Utrecht University Graduate School of Life Sciences, Heidelberglaan 8, 3584 CS Utrecht, the Netherlands

5: Department of Genetics, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands

6: Department of Pediatric Respiratory Medicine, Wilhelmina Children's Hospital, University Medical Center, Utrecht University, 3584 EA Utrecht, the Netherlands

7: Department of Pediatric Gastroenterology, University Medical Center Groningen, Hanzeplein 1, 9713 GZ Groningen, the Netherlands

8: Department of Sciences, University College Roosevelt, Lange Noordstraat 1, 4331 CB Middelburg, the Netherlands

Abstract

Prime editing is a versatile genome-editing technique. However, some genomic sites remain difficult to edit and optimal design of prime-editing tools is elusive. We present a fluorescent prime editing and enrichment reporter (fluoPEER), which can be tailored to any genomic target site. This system rapidly and faithfully ranked the efficiency of prime edit guide RNAs (pegRNAs) and any prime editor protein, including novel variants with flexible protospacer adjacent motif (PAM) recognition. We applied fluoPEER to instruct correction of pathogenic variants in patient cells and found that plasmid-editing enriched for editing at the genomic target site. Transcriptomic analysis of reporter-edited cells revealed that successful prime editing was associated with expression of DNA repair genes. Together, our results show that fluoPEER can be employed for rapid and efficient correction of patient cells, selection of gene-edited cells, and elucidation of cellular mechanisms needed for successful prime editing.

Introduction

Prime editing is a precise genome-editing technique to generate any moderately sized genetic variation without the need for double-strand DNA breaks (Anzalone et al., 2019). This technique enables the repair of patient mutations, generation of disease models, and in vivo editing (Schene et al., 2020; Liu et al., 2021). The requisite protospacer adjacent 5'-NGG-3' motif (NGG-PAM) sequence near edit sites limits the capacity of the original prime editor (PE2) to edit mutations. Availability of only one nearby NGG-PAM on average to target the pathogenic mutations described in the ClinVar database (Supplementary Fig. 1) limits pegRNA design to often suboptimal spacer and primer binding sites and thereby reduces gene-editing efficiency. PE2 variants with flexible PAM recognition (Nishimasu, 2018; Walton, 2018; Kweon, 2021) provide at least nine PAM-sites to target the pathogenic mutations described within the ClinVar database and thereby greatly improve gene-editing potential.

The mechanisms required for efficient prime editing are poorly understood. More specifically, optimal design of the prime edit guide RNA (pegRNA) remains largely elusive. However, it is known that primer binding site (PBS) and reverse transcriptase template (RTT) variations have marked effects on editing efficiency (Anzalone et al., 2019; Lin et al., 2021). A deep learning strategy has further specified important pegRNA characteristics, and can now be used to instruct 'NGG' pegRNA design (Kim et al., 2021). However, this deep learning algorithm failed to predict a pegRNA efficiency score for >80% of the mutations in the

ClinVar database and does not support non-‘NGG’ pegRNA prediction (Supplementary Fig. 2a and 2b). As expected, for those pathogenic mutations within the scope of the prediction algorithm, a higher number of available PAMs per mutation resulted in a higher predicted maximum pegRNA efficiency (Supplementary Fig. 2c). To address the shortcomings of current prediction methods and harness the strength of new prime editor variants, we present fluoPEER: a fluorescent prime editing and enrichment reporter. This reporter can be tailored to any genomic target site and allows high-throughput analysis of pegRNA designs and prime editors using fluorescence-activated cell sorting (FACS) within seven days from the start of cloning.

Results

FluoPEER guides pegRNA design for efficient genomic editing

We cloned genomic target regions (45-100 nucleotides) spanning all possible pegRNAs between an (e)GFP and an (m)Cherry cassette under constitutive expression (Fig. 1a). When the target mutation did not contain a native stop codon or frameshift, a stop codon or frameshift was inserted (Fig. 1a). We hypothesized that this adaptation would not limit the predictive power of fluoPEER since the substitution or insertion of single nucleotides in the RTT-region has minor effects on prime-editing efficiency (Supplementary Fig. 3). Importantly, this system allows targeting of the reporter and the corresponding genomic locus with the same pegRNA design.

We transfected fluoPEER together with prime editing machinery into HEK293T cells and analyzed reporter editing using FACS after 3 days (Fig. 1b). Based on the ratio of Cherry to GFP signal, an efficiency score was established for each editing condition, enabling ranking of PE2 variants and pegRNAs (Fig. 1c). To verify the reliability of this reporter system, we used the same combinations of PE2 variants and pegRNAs to convert a genomically integrated GFP gene to BFP. When ranking the editing conditions based on efficiency for both strategies, correlation between genomic editing and reporter prediction was very strong ($r=0.94$, Fig. 1d, Supplementary Fig. 4). This correlation further confirms that insertion of a 1-nucleotide frameshift mutation in the reporter does not hamper the predictive capacity. Importantly, Sanger sequencing could only detect genomic editing in conditions with >10% BFP+ cells (Supplementary Fig. 4), supporting the need for a sensitive reporter system when targeting difficult-to-edit loci.

Figure 1: fluoPEER instructs pegRNA design for genomic editing

a, The fluoPEER plasmid uses a 45-100 nucleotide genomic region containing a stop codon or frameshift. In case the genomic region does not contain a naturally occurring stop codon or frameshift, one is added. The prime edit machinery targets and edits the genomic insert, removing the insertion or stop codon. The same prime edit machinery, including the same pegRNA design, is capable of editing the genomic DNA. A finished fluoPEER plasmid contains a GFP-P2A-Genomic insert-P2A-Cherry. **b**, Editing of the genomic insert is visualized by Cherry signal and quantified using flow cytometry. **c**, fluoPEER distinguishes between the efficiency of different PE2 variants. This is quantified as the total Cherry signal over the total GFP signal, which gives a measure of editing per transfected plasmid. **d**, A comparison between the generation of a GFP to BFP conversion using prime editing on the genomic DNA and on the fluoPEER plasmid. HEK293T cells containing a lentivirally integrated genomic GFP cassette are transfected with the prime editing machinery to convert GFP to BFP in 24 conditions, while the same conditions are applied to HEK293T cells cotransfected with the fluoPEER plasmid for the GFP to BFP conversion. Comparing the efficiency ranking of editing conditions, a strong spearman correlation is found (0.94) between fluoPEER and genomic DNA editing. **e**, Prime editing on fluoPEER yields efficiency read-outs for 7 edits for 22 pegRNAs in total. **f**, Using the optimal pegRNAs based on fluoPEER data, several pathogenic mutations in patient-derived organoids are genetically corrected and organoids with biallelic *IARS1* mutations are generated in wildtype liver organoids. Error bars represent standard deviations from the mean.

Next, we compared the DeepPE algorithm to fluoPEER. In contrast to the algorithm, fluoPEER ranked pegRNA efficiencies correctly at each genomic locus (Supplementary Fig. 5). Furthermore, we confirmed the robustness of the Cherry over GFP ratio as a read-out for efficiency by varying the input reporter plasmid concentration (Supplementary Fig. 6a-b). Finally, by varying the time point of read-out, we further corroborated that the ratio-based ranking was stable over a period of 6 days (Supplementary Fig. 6c).

Using the reporter system, we tested pegRNAs and PE2 variant combinations for many different genomic loci within 7 days from the start of cloning (Fig. 1e). Using available knowledge of characteristics influencing editing efficiency (Kim et al., 2021, Lin et al., 2021), we designed pegRNAs containing spacers with high Cas9 sgRNA binding scores, PBSs of 9-15 nucleotides length with GC-counts of at least 5, and RTTs of 10-20 nucleotides length (Supplementary Table 1b). Selecting the pegRNA+PE2 condition with the highest reporter prediction score and combining it with a nicking sgRNA (PE3) enabled quick and efficient repair of disease-causing mutations in patient-derived organoids (Fig. 1f). These included common mutations causing cystic fibrosis (*CFTR*^{F508del} and *CFTR*^{G542X}) that are difficult to edit without the use of flexible prime editors. In addition, we corrected mutations causing progressive familial intrahepatic cholestasis (PFIC) type 3 (*ABCB4*^{E1012X}), and PFIC type 1 (*ATP8B1*^{R600Q}). Furthermore, we used fluoPEER ranking to efficiently mutate cytosolic isoleucyl-tRNA synthetase (*IARS1*) in liver-derived organoids. This confirmed that biallelic, but not monoallelic, *IARS1*^{I1174N} mutations impaired organoid expandability, reflecting the failure to thrive observed in patients with biallelic *IARS1* mutations (Kopajtich et al., 2016; Fuchs et al., 2019; Supplementary Fig. 7a-b). Importantly, pegRNAs targeting a difficult-to-edit locus (low GC-content) to repair a common methylmalonic acidemia mutation (*MUT*^{R369H}) did not edit on the reporter (Fig. 2e) and also failed in organoids, demonstrating accurate negative prediction (Supplementary Fig. 7c). Lastly, fluoPEER correctly predicted pegRNA efficiencies to edit *CTNNB1* as observed in earlier organoid experiments (Schene et al., 2020) (Supplementary Fig. 7d).

FluoPEER enables characterization of various genome editing techniques

To better characterize new PE2 variants, we adapted the genomic target region of fluoPEER to test PAM-specificities of each flexible prime editor. We found that the SpG- and NG-PE2 complexes have comparable flexible PAM preferences, but that the SpG has higher efficiency scores overall (Fig. 2a,b). The SpRY-PE2 displayed a PAM-specificity pattern highly similar to the SpRY-Cas9 protein (Walton et al., 2020), essentially functioning as a PAM-less prime editor. Interestingly, a guanine nucleotide on the 4th PAM position led to higher editing efficiency compared to cytosine (Fig 2b). We also tested PE2* variants with

Figure 2: FluoPEER enables characterization of various genome editing techniques

a, Schematic of the design of 16 different fluoPEER plasmids, each containing the same pegRNA spacer binding site (HEK3) containing a stop codon at position 18-20 followed by a variable 4-nucleotide PAM sequence. Prime editing using a pegRNA that converts this stop codon to an arginine-encoding codon results in Cherry signal. **b**, Heatmaps of Cherry to GFP ratio for the 16 plasmids in (a), using four PE2 variants and replicates ($n=2$) per condition. We find that PAM-specificities of PE2 variants are highly similar to those described for corresponding cutting Cas9 variants (Walton et al., 2020). Note that a guanine (G) nucleotide on position 4 of the PAM sequence is associated with higher overall editing efficiency. **c**, The left panel shows a summary of all fluoPEER scores, expressed as the fold change (FC) of the PE2* variants relative to the corresponding PE2 variants. The right panel shows a summary of the conversion efficiency of genomic GFP to BFP using various PE2 and PE2* variants, expressed as the FC between the PE2* and the PE2 variant. Note that the PE2* NLS adaptations did not result in an overall increase of fluoPEER scores nor an overall increase in genomic editing efficiency. See also Supplementary Fig. 8. **d**, Prime editing on fluoPEER allows testing of PE3b-nicking sgRNAs. **e**, FluoPEER can report base editing, when editing of the target nucleotide resolves a stop codon in any possible reading frame. **f**, Patient-derived fibroblasts with biallelic *POLG*^{A467T} mutations are shown as an example. Error bars represent standard deviations from the mean.

improved nuclear localization (Liu et al., 2021), but these variants did not display higher efficiency of reporter editing and genomic editing than standard PE2 variants (Fig. 2c and Supplementary Fig. 8).

We continued to characterize the ability of the fluoPEER system to report genome editing techniques beyond PE2. Nicking the unedited strand with a sgRNA that only binds after successful editing, known as PE3b, can increase prime editing efficiency without creating unwanted insertions and deletions (Anzalone et al., 2019). We reasoned that fluoPEER should support this technique and indeed found that these PE3b designs, including non-NGG-PAM sgRNAs, increased reporter editing while others did not (Fig. 2d). Finally, we applied fluoPEER to base editing, which is possible when base editing resolves a stop codon in any of the six possible reading frames. We integrated the genomic region of the common *POLG*^{A467T} mutation into the fluoPEER plasmid with a reporter-shifted (+1) reading frame that resulted in a stop codon (Fig. 2e). As expected, adenine base editing (ABE8e-TadA) corrected this mutation on both the plasmid and the genomic DNA of patient-derived fibroblasts (Fig. 2e).

FluoPEER enriches for genomic editing

Selection of cells with a detectable edit has been shown to enrich for a simultaneously introduced edit at another locus (Katti et al., 2020; Li et al., 2021). To test whether fluoPEER allows such enrichment, we sorted transfected cells that were plasmid-edited (GFP⁺Cherry⁺) and found that these cells showed large increases (>2-fold) in genomic editing compared to plasmid-unedited cells (GFP⁺Cherry⁻) (Fig. 3a). The same was observed for editing in patient-derived organoids, corroborating that this system can function as an enrichment tool for repair of patient mutations (Fig. 3a). Finally, we targeted the *POLG*^{A467T} mutation in patient-derived skin fibroblasts, using a PE3 approach and cotransfection with a corresponding fluoPEER plasmid. After sorting and 19 days of culturing, plasmid-edited fibroblasts showed editing efficiencies of >90%, which was >2-fold higher than plasmid-unedited fibroblasts (Supplementary Fig. 9).

We hypothesized that enrichment of genomic editing was caused by higher transfection rates of the prime-editing plasmids in GFP⁺Cherry⁺ cells, but we determined that >90% of transfected cells received all plasmids (Fig. 3b, Supplementary Fig. 10). This suggested that a subpopulation of cells is more susceptible to prime editing. Hence, we compared the transcriptome of reporter-edited versus non-edited cells and found enrichment of DNA repair genes and cell cycle-associated genes (Fig. 3c,d and Supplementary Fig. 11). We therefore evaluated the reporter editing of cells during different cell cycle stages using FACS. We

Figure 3: FluoPEER enriches for genomic editing

a, Sorting based on presence (GFP⁺Cherry⁺) and absence (GFP⁺Cherry⁻) of reporter editing shows a strong enrichment of genomic editing in the reporter-edited population in HEK293T cells and organoids. Cells were sorted 24 (PE2) or 72 (PE3) hours after transfection. **b**, Cotransfection of 2, 3, or 4 different fluorescent plasmids shows that >90% of the cells successfully transfected with a single plasmid also received all other plasmids. **c**, PCA plot and **d**, gene set enrichment analysis of RNA sequencing of the GFP⁺Cherry⁺ and GFP⁺Cherry⁻ populations show expression-based differences in DNA repair- and cell cycle- associated genes. **e**, Hoechst staining in HEK293T cells transfected with the fluoPEER plasmid and prime editing machinery shows enrichment of reporter editing signal in G2. **f**, Sanger sequencing of HEK293T cells sorted based on cell cycle shows differences in genomic editing efficiency. Error bars represent standard deviations from the mean.

found a large increase (>40%) of Cherry signal in cells that were in G2 (Fig. 3e). This indicates an effect of cell cycle-associated mechanisms, but the precise cell cycle stages of importance for prime editing are difficult to infer due to the delay in Cherry expression. We confirmed cell cycle associated differences in prime-editing efficiency when looking at genomic DNA editing (Fig. 3f). Due to the time required for gene editing and the concurrent progression in cell cycle, the specific effects of cell cycle stage on prime editing need further elucidation. Nevertheless, these findings indicate a cell cycle dependency for prime-editing efficiency. Together, these results establish fluoPEER as a highly dynamic prime-editing read-out, enabling enrichment of genomic editing based on plasmid co-editing.

Discussion

With our studies, we show that fluoPEER is a customizable tool to guide pegRNA design, and to select and optimize prime editor proteins for almost any mutation in patient-derived cells. Furthermore, fluoPEER provides a transient selection method to enrich genomic editing more than two-fold. In contrast to previously developed base-editing reporter systems (Katti et al., 2020; Li et al., 2021), this versatile tool does not require transfection of a second pegRNA, does not rely on genomic integration of a selection cassette, nor requires generation of an additional genomic edit. Based on these characteristics, fluoPEER will contribute to facilitate effective prime editing of disease-causing mutations in any relevant cell type.

Methods

ClinVar database computational analysis

Information for all pathogenic mutations shorter than 51 bp was obtained from the ClinVar database (Landrum et al., 2018), accessed October 2020. Genomic sequences flanking these mutations were obtained from RefSeq (Pruitt et al., 2005) accessed October 2020, using the SPDI data model (Holmes et al., 2020) and a custom python script. The -10 to +4 bp region around the target sites were searched for NGG, NAN or NGN PAMs. The efficiency of prime editing using NGG PAMs was predicted using PE_Position and PE_type random forest models, provided by (Kim et al., 2021). Figures were made in python using Matplotlib (Hunter et al., 2007). The code used for this analysis is available on GitHub (https://github.com/JBaijens/PE_prediction).

Study approval and human subjects

The study was approved by the responsible local ethics committees (Institutional Review Board of the University Medical Center Utrecht and University Medical Center Groningen (STEM: 10-402/K; TcBio 14-008; Metabolic Biobank: 19-489)). For cystic fibrosis organoids, collection of patient tissue and data was performed following the guidelines of the European Network of Research Ethics Committees (EUREC). Tissue biopsies from the liver of a patient with ABCB4 deficiency (PFIC3) was obtained during a liver transplant procedure in the UMCG, Groningen. Rectal biopsies used for intestinal organoid culture from a patient with ATP8B1 deficiency (PFIC1) and skin biopsies used for fibroblasts culture from a patient with homozygous *POLG*^{A467T} mutations were obtained at the outpatient clinic in the UMCU, Utrecht. Biobanked intestinal organoids are stored and cataloged (<https://huborganoids.nl/>) at the foundation Hubrecht Organoid Technology (<http://hub4organoids.eu>). All biopsies were used after written informed consent.

Plasmid cloning

FluoPEER plasmids were cloned using the backbone of the pmGFP-P2A-K0-P2A-RFP (Addgene #105686) plasmid, which was a gift from Ramanujan Hegde. This plasmid was cut with Sall and Acc65I for 16 hours at 37 °C, creating 'TCGA' and 'GTAC' overhangs, after which the 6 kb fragment was isolated from gel. Genomic insert oligos containing 5' 'TCGACC' and 3' 'G' overhangs on the top oligo, and 5' 'GTACC' and 3' 'GG' overhangs on the bottom oligo were annealed and inserted using a conventional ligation protocol (see Supplementary Note 1). Cloning of pegRNA plasmids was performed according to previously described protocols (Anzalone et al., 2019). In brief, the pU6-pegRNA-GG-Vector (Addgene #132777) was digested for 16 hours with BsaI-HFv2 (NEB), after which the 2.2 kb fragment was isolated from gel. Oligonucleotide duplexes of the pegRNA spacer, pegRNA extension, and pegRNA scaffold sequences were ordered containing the appropriate overhangs and subsequently annealed. The annealed pegRNA duplexes were ligated into the pU6-pegRNA-GG-Vector using Golden Gate assembly with BsaI-HFv2 (NEB) and T4 DNA ligase (NEB) in a protocol of 12 cycles of 5 min at 16 °C and 5 min at 37 °C. For cloning of sgRNAs used for PE3, we replaced the BsmBI restriction sites of the BPK1520 plasmid by BbsI restriction sites using PCR, which allowed direct ligation of sgRNA-spacer duplexes (Ran et al., 2013). All fluoPEER insert, pegRNA, and sgRNA sequences used in this work are listed in Supplementary Table 1 and were synthesized by Integrated DNA Technologies (IDT). pCMV-PE2 (Addgene #132775) and pU6-pegRNA-GG-acceptor (Addgene #132777) were gifts from David Liu; BPK1520 (Addgene #65777) was a gift from Keith Joung.

Cloning of flexible PE2s and PE2*

Using PCR and In-Fusion cloning (Takara Bio), the 'NGG' PAM-recognition domain of the prime editor protein (PE2) was replaced with the corresponding domains in NG-ABEmax (Huang et al., 2019), SpG-ABEmax, or SpRY-ABEmax (Walton et al., 2020). NG-ABEmax was a gift from David Liu (Addgene #124163). SpG- and SpRY-ABEmax were gifts from

Benjamin Kleinstiver (Addgene plasmids #140002 and #140003). PE2* variants with improved nuclear localization sequences (NLSs) were adapted from the NGG-PE2* developed by Liu et al. (2021) and were cloned by PCR and In-Fusion cloning (Takara Bio). Successful cloning of all plasmids was confirmed by Sanger sequencing.

Organoid culture

Liver and intestinal organoids were cultured according to previously described protocols. In short, liver organoids were plated in matrigel (Corning) and maintained in liver expansion medium (EM), consisting of AddMEM/F12 (GIBCO) supplemented with 2% B27 without vitamin A (GIBCO), 1.25 mM N-Acetylcysteine (Sigma), 10 mM Nicotinamide (Sigma), 10 nM gastrin (Sigma), 10% RSPO1 conditioned media (homemade), 50 ng/ml EGF (Peprotech), 100 ng/ml FGF10 (Peprotech), 25 ng/ml HGF (Peprotech), 5 mM A83-01 (Tocris), and 10 mM FSK (Tocris). Small intestine and colon organoids were plated in matrigel and maintained in SI-EM, consisting of AddMEM/F12 supplemented with 50% WNT3A-, 20% RSPO1-, and 10% NOG(gin)-conditioned medium (all homemade), 2% B27 with vitamin A (GIBCO), 1.25 mM N-Acetylcysteine, 10 mM Nicotinamide, 50 ng/ml murine-EGF (Peprotech), 500 nM A83-01, and 10 mM SB202190 (Sigma). The medium was changed every 2–4 days and organoids were passaged 1:4 –1:8 each week. After thawing, organoids were passaged at least once before electroporation.

Cell culture, lentiviral production, and lentiviral transduction

HEK293T cells were maintained and split every 4-5 days. For production of lentivirus, HEK293T cells were plated in a 145mm CELLSTAR dish (Corning) and transfected 24 hours later (at 50-60% confluence) with a mix of 10 µg of the pLenti-CMV-GFP-Hygro plasmid, 5 µg of psPAX2, 5 µg of pMD2.G, and 60 µl of polyethylenimine (1 mg/ml). pLenti-CMV-GFP-Hygro was a gift from Eric Campeau & Paul Kaufman (Addgene #17446), psPAX2 and pMD2.G were gifts from Didier Trono (Addgene plasmid # 12260 and # 12259). 24 hours after transfection, the medium was replaced with DMEM containing pen/strep supplemented with 10% (vol/vol) FBS, and virus containing medium was harvested at 48 and 96 hours after transfection. Medium was centrifuged at 1,000 r.p.m. for 5 min and supernatants were filtered through a 0.22-µm filter, after which virus particles were concentrated by ultracentrifugation at 50,000 x g for 2 hours and resuspension in 1 ml DMEM. HEK293T cells were transduced at an MOI of 2 for 24 hours in the presence of polybrene (8 µg/µl) and analyzed by FACS 14 days after transduction to confirm stable GFP expression.

Transfection of HEK293T cells

HEK293T cells were plated at a density of 50,000 cells per well in a 96 wells plate one day prior to transfection. HEK293T cells were transfected with 0.1 µg fluoPEER plasmid, 0.25 µg PE2 plasmid, 0.1 µg pegRNA plasmid and 0.05 µg nicking-gRNA plasmid in a mix of 25 µl OptiMEM and 0.5 µl lipofectamine 2000 for each well.

Electroporation of organoid cells

Before electroporation, organoids were grown under standard culture conditions. Four wells containing organoids were then dissociated for each condition using TrypLE for 4–5 min at 37 °C, after which mechanical disruption was applied through pipetting. Cells were washed once using Advanced DMEM/F12, resuspended in 80 µl OptiMEM containing Y-27632 (10 µM), and 20 µl DNA mixture was added. For prime editing, the DNA mixture contained 4 µg fluoPEER, 12 µg PE2 plasmid, 4 µg pegRNA plasmid and 2 µg nicking sgRNA plasmid. The cell-DNA mixture was transferred to an electroporation cuvette and electroporated using a NEPA21 electroporator (NEPA GENE) with 2× poring pulse (voltage: 175 V, length: 5 ms, interval: 50 ms, polarity: +) and 5× transfer pulse (voltage: 20 V, length: 50 ms, interval: 50 ms, polarity ±), as previously described (Fujii, 2015). Cells were removed from the cuvette and transferred into 500 µl OptiMEM containing Y-27632 (10 µM). After 20 minutes, cells were plated in 120 µl matrigel divided over four wells. Upon polymerization of the Matrigel, medium was added containing Y-27632 (10 µM).

Transfection of fibroblast cells

Skin-derived fibroblasts were grown under standard culture conditions and plated on 12-well plates 3 days prior to transfection so that confluency was 60-70% at transfection. For prime editing experiments, fibroblasts were transfected with 0.12 µg fluoPEER plasmid, 0.19 µg PE2 plasmid, 0.05 µg pegRNA plasmid and 0.05 µg nicking-gRNA plasmid in a mix of 8.4 µl OptiMEM, 0.6 µl lipofectamine 3000, and 0.4 µl P3000 reagent. For base editing experiments, the DNA mix consisted of 0.14 µg fluoPEER plasmid, 0.2 µg ABE8e-TadA (V106W) plasmid, and 0.06 µg sgRNA. ABE8e (TadA-8e V106W) was a gift from David Liu (Addgene #138495).

FACS

Organoids and cell lines were harvested and digested to single cells, after which cells were resuspended in FACS buffer (phosphate-buffered saline with 2 mM ethylenediaminetetraacetic acid and 0.5% bovine serum albumin). Prior to FACS, cells were filtered through a 5 ml Falcon polystyrene test tube (Corning). Flow cytometry was performed on the FACS Fortessa (BD) and sorting was performed on the FACS FUSION (BD) using FACS Diva software (BD). Sorted cells were collected in culture medium and spun down. Gating strategy for cells included for fluoPEER analysis is shown in Supplementary Fig. 12.

Genotyping

Sorted cells were harvested using the Quick-DNA microprep kit (Zymogen) according to manufacturer's protocols. PCR was performed on the genomic region of interest using the Phusion polymerase (ThermoFisher) and purified using the QIAquick PCR Purification Kit (Qiagen) according to manufacturer instructions. The PCR product was sent for Sanger sequencing to EZSeq Macrogen Europe. All PCR and sequencing primers used are listed in Supplementary Table 1.

RNA sequencing

HEK293T cells were transfected with fluoPEER, PE2, pegRNA, and ngRNA plasmids and FACS sorted after 48 hours. Total RNA was isolated using Trizol LS reagent (Invitrogen) and stored at -80°C until further processing. mRNA was isolated using Poly(A) Beads (NEXTflex). Sequencing libraries were prepared using the Rapid Directional RNA-Seq Kit (NEXTflex) and sequenced on a NextSeq500 (Illumina) to produce 75 base long reads (Utrecht DNA Sequencing Facility). Sequencing reads were mapped against the reference genome (hg19 assembly, NCBI37) using BWA (Li and Durbin, 2009) package (mem -t 7 -c 100 -M -R). Raw reads were further analyzed as described under 'Data analysis'.

Cell cycle analysis

200 ng/ml nocodazole was added to HEK293T cells 20 hours before FACS analysis as a control for cells in G2 phase. 24 hours before fluoPEER read-out and genomic sanger sequencing, HEK293T cells were transfected. 60 minutes before harvesting for FACS analysis, 10 µg/ml Hoechst 33342 (ThermoFisher) was added to the culture medium. Gating strategy for G1, S, and G2 phase of the cell cycle is shown in Supplementary Fig. 12.

Data analysis

Flow cytometry data was analyzed using FlowJo™ Software. RNA sequencing was analyzed using DESeq2 in RStudio (Love et al., 2014), gene set enrichment analysis (Subramanian et al., 2005), and enrichR (Chen et al., 2013). All figures were made in Prism (GraphPad Software) or GGPlot2 (Wickham et al., 2012) in RStudio. Sanger sequencing was quantified using EditR (Kleusner et al., 2018) or TideR (Brinkman et al., 2018). Sanger sequencing chromatograms were made in Benchling.

Statistics and reproducibility

No pre-specified effect size was calculated, and no statistical method was used to predetermine sample size. For comparisons of multiple groups, an ordinary one-way ANOVA with Holm–Sidak correction for multiple comparisons was used and performed in Prism (GraphPad Software). Statistical tests were appropriate for comparisons being made; assessment of variation was carried out but not included. Experiments were not randomized. Reproducibility: Fig. 1c representative of three replicates from one experiment. In Fig. 1d, each point in the dot-plot represents the mean of three (reporter rank) or two (genomic rank) replicates in two independent experiments. Each pegRNA-PE2-fluoPEER combination from Fig. 1e was tested in at least two independent transfection experiments. Fig. 1e representative of two or three technical replicates from one characteristic experiment. Fig. 1f representative of 1) 14/20 clonally picked intestinal organoids from two different patients with biallelic *CFTR*^{F508del} mutations, that showed swelling after addition of 1 μ M forskolin to the medium, 2) 2/10 clonally picked liver-derived organoids from a patient with biallelic *ABCB4*^{E1012X} mutations, 3) 3/10 clonally picked intestinal organoids from a patient with biallelic *ATP8B1*^{R600Q} mutations, and 4) 2/10 clonally picked liver-derived organoids from a healthy control, in which biallelic *IARS1*^{I1174N} were created; 4/10 clonally picked organoids from the same experiment showed monoallelic *IARS1*^{I1174N} mutations. Fig. 2b representative of two replicates in two independent experiments. Fig. 2c and Supplementary Fig. 8a-b representative of two to three replicates per condition. Data from all conditions in Supplementary Fig. 8a-b was pooled for calculation of fold changes between PE2* and PE2 variants in Fig. 2c. Fig. 2d representative of two or three technical replicates from one characteristic experiment. Fig. 2f and Supplementary Fig. 9 show one representative experiment. Fig. 3a representative of two to four replicates from one experiment. Fig. 3b representative of six replicates from two independent experiments. Fig. 3c-d and Supplementary Fig. 11 representative of two replicates for two conditions (Cherry⁺ vs. Cherry⁻) in two prime editing strategies. Fig. 3e representative of six replicates in two independent experiments. Fig. 2f representative of three replicates in one experiment. Supplementary Fig. 4c representative of three replicates per condition in two independent experiments. Supplementary Fig. 5a representative of two replicates in a single experiment that was characteristic of 2 repeated experiments. Supplementary Fig. 6 representative of two replicates per condition in one experiment; for representation of percentage of GFP⁺ cells, data from different conditions was pooled. Supplementary Fig. 7a-b representative of clonally picked liver organoids with monoallelic (two clones) or biallelic (two clones) *IARS1*^{I1174N} mutations. Supplementary Fig. 7c representative of 20 clonally picked liver-derived organoids from a patient with biallelic *MUT*^{R329H} mutations. Supplementary Fig. 10 shows FACS plots of single conditions that are representative for two independent transfection experiments.

Data and material availability

All plasmids created for this study are available upon reasonable request. Raw counts of the RNA-sequencing experiment will be made publicly available upon publication.

Acknowledgements

The authors thank the Kim lab of the Department of Pharmacology, Yonsei University College of Medicine for supplying the PE_prediction code. This work was supported by Metakids funding (to S.A.F.), a Clinical Fellows grant from The Netherlands Organisation for Health Research and Development Health Institute (40-00703-97-13537 to S.A.F.), and a Boost Grant Child Health from the WKZ, UMC Utrecht (to I.F.S.).

Competing interest

Authors declare that they have no competing interests.

Author contributions

I.F.S., I.P.J. and S.A.F. designed the project; J.H.L.B. analysed the pathogenic mutations in the ClinVar database using PE_prediction; H.P.J.D. and R.H.J.H. helped establishing the

biobank of patient-derived stem cell organoids used in this study; I.F.S., I.P.J., R.S., G.K., S.S., E.F.I., D.P.B. and R.C.M.R performed experiments and analyses; S.A.S. and J.M.B. provided the cystic fibrosis samples used in this study; I.F.S., I.P.J., E.E.S.N., and S.A.F. wrote the manuscript.

References:

- Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., ... & Liu, D. R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, 576(7785), 149-157.
- Katti, A., Foronda, M., Zimmerman, J., Diaz, B., Zafra, M. P., Goswami, S., & Dow, L. E. (2020). GO: a functional reporter system to identify and enrich base editing activity. *Nucleic acids research*, 48(6), 2841-2852..
- Kim, H. K., Yu, G., Park, J., Min, S., Lee, S., Yoon, S., & Kim, H. H. (2021). Predicting the efficiency of prime editing guide RNAs in human cells. *Nature Biotechnology*, 39(2), 198-206.
- Kweon, J., Yoon, J. K., Jang, A. H., Shin, H. R., See, J. E., Jang, G., ... & Kim, Y. (2021). Engineered prime editors with PAM flexibility. *Molecular Therapy*.
- Li, S., Akrap, N., Cerboni, S., Porritt, M. J., Wimberger, S., Lundin, A., ... & Maresca, M. (2021). Universal toxin-based selection for precise genome engineering in human cells. *Nature Communications*, 12(1), 1-14.
- Lin, Q., Jin, S., Zong, Y., Yu, H., Zhu, Z., Liu, G., ... & Gao, C. (2021). High-efficiency prime editing with optimized, paired pegRNAs in plants. *Nature Biotechnology*, 1-5.
- Liu, P., Liang, S. Q., Zheng, C., Mintzer, E., Zhao, Y. G., Ponnienselvan, K., ... & Xue, W. (2021). Improved prime editors enable pathogenic allele correction and cancer modelling in adult mice. *Nature Communications*, 12(1), 1-13.
- Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O.O., Gootenberg, J.S., Mori, H. (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science*, 361, 1259-1262
- Schene, I. F., Joore, I. P., Oka, R., Mokry, M., van Vugt, A. H., van Boxtel, R., ... & Fuchs, S. A. (2020). Prime editing for functional repair in patient-derived disease models. *Nature communications*, 11(1), 1-8.
- Walton, R. T., Christie, K. A., Whittaker, M. N. & Kleinstiver, B. P. (2020) Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science*, 368, 290-296.

Method section references

- Brinkman, E. K., Kousholt, A. N., Harmsen, T., Leemans, C., Chen, T., Jonkers, J., & van Steensel, B. (2018). Easy quantification of template-directed CRISPR/Cas9 editing. *Nucleic acids research*, 46(10), e58-e58.
- Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., Ma'ayan, A. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*, 128(14).
- Fujii, M., Matano, M., Nanki, K., & Sato, T. (2015). Efficient genetic engineering of human intestinal organoids using electroporation. *Nature protocols*, 10(10), 1474.
- Holmes, J. B., Moyer, E., Phan, L., Maglott, D., & Kattman, B. (2020). SPDI: data model for variants and applications at NCBI. *Bioinformatics*, 36(6), 1902-1907.
- Huang, T.P., Zhao, K.T., Miller, S.M., Gaudelli, N.M., Oakes, B.L., Fellmann, C., Savage, D.F., Liu, D.R. (2019) Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nature Biotechnology*, 37(6), 626-631.
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *IEEE Annals of the History of Computing*, 9(03), 90-95.
- Kim, H. K., Yu, G., Park, J., Min, S., Lee, S., Yoon, S., & Kim, H. H. (2021). Predicting the efficiency of prime editing guide RNAs in human cells. *Nature Biotechnology*, 39(2), 198-206.
- Kluesner, M. G., Nedveck, D. A., Lahr, W. S., Garbe, J. R., Abrahante, J. E., Webber, B. R., & Moriarity, B. S. (2018). EditR: a method to quantify base editing from Sanger sequencing. *The CRISPR journal*, 1(3), 239-250.
- Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., ... & Maglott, D. R. (2018). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic acids research*, 46(D1), D1062-D1067.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows – Wheeler transform. *Bioinformatics* 25, 1754–1760.
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15(12), 1-21.
- Pruitt, K. D., Tatusova, T., & Maglott, D. R. (2005). NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic acids research*, 33(suppl_1), D501-D504.
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature protocols*, 8(11), 2281-2308.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., ... & Mesirov, J.P. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102(43), 15545-50.
- Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.

Figure legends

Figure 1: FluoPEER instructs pegRNA design for genomic editing

- a**, The fluoPEER plasmid uses a 45-100 nucleotide genomic region containing a stop codon or frameshift. In case the genomic region does not contain a naturally occurring stop codon or frameshift, one is added. The prime edit machinery targets and edits the genomic insert, removing the insertion or stop codon. The same prime edit machinery, including the same pegRNA design, is capable of editing the genomic DNA. A finished fluoPEER plasmid contains a GFP-P2A-Genomic insert-P2A-Cherry.
- b**, Editing of the genomic insert is visualized by Cherry signal and quantified using flow cytometry.
- c**, fluoPEER distinguishes between the efficiency of different PE2 variants. This is quantified as the total Cherry signal over the total GFP signal, which gives a measure of editing per transfected plasmid.
- d**, A comparison between the generation of a GFP to BFP conversion using prime editing on the genomic DNA and on the fluoPEER plasmid. HEK293T cells containing a lentivirally integrated genomic GFP cassette are transfected with the prime editing machinery to convert GFP to BFP in 24 conditions, while the same conditions are applied to HEK293T cells cotransfected with the fluoPEER plasmid for the GFP to BFP conversion. Comparing the efficiency ranking of editing conditions, a strong spearman correlation is found (0.94) between fluoPEER and genomic DNA editing.
- e**, Prime editing on fluoPEER yields efficiency read-outs for 7 edits for 22 pegRNAs in total.
- f**, Using the optimal pegRNAs based on fluoPEER data, several pathogenic mutations in patient-derived organoids are genetically corrected and organoids with biallelic *IARS1* mutations are generated in wildtype liver organoids. Error bars represent standard deviations from the mean.

Figure 2: FluoPEER enables characterization of various genome editing techniques

- a**, Schematic of the design of 16 different fluoPEER plasmids, each containing the same pegRNA spacer binding site (HEK3) containing a stop codon at position 18-20 followed by a variable 4-nucleotide PAM sequence. Prime editing using a pegRNA that converts this stop codon to an arginine-encoding codon results in Cherry signal.
- b**, Heatmaps of Cherry to GFP ratio for the 16 plasmids in (a), using four PE2 variants and replicates (n=2) per condition. We find that PAM-specificities of PE2 variants are highly similar to those described for corresponding cutting Cas9 variants (Walton et al., 2020). Note that a guanine (G) nucleotide on position 4 of the PAM sequence is associated with higher overall editing efficiency.
- c**, The left panel shows a summary of all fluoPEER scores, expressed as the fold change (FC) of the PE2* variants relative to the corresponding PE2 variants. The right panel shows a summary of the conversion efficiency of genomic GFP to BFP using various PE2 and PE2* variants, expressed as the FC between the PE2* and the PE2 variant. Note that the PE2* NLS adaptations did not result in an overall increase of fluoPEER scores nor an overall increase in genomic editing efficiency. See also Supplementary Fig. 8.
- d**, Prime editing on fluoPEER allows testing of PE3b-nicking sgRNAs.
- e**, FluoPEER can report base editing, when editing of the target nucleotide resolves a stop codon in any possible reading frame.
- f**, Patient-derived fibroblasts with biallelic *POLG*^{A467T} mutations are shown as an example. Error bars represent standard deviations from the mean.

Figure 3: FluoPEER enriches for genomic editing

- a**, Sorting based on presence (GFP⁺Cherry⁺) and absence (GFP⁺Cherry⁻) of reporter editing shows a strong enrichment of genomic editing in the reporter-edited population in HEK293T cells and organoids. Cells were sorted 24 (PE2) or 72 (PE3) hours after transfection.
- b**, Cotransfection of 2, 3, or 4 different fluorescent plasmids shows that >90% of the cells successfully transfected with a single plasmid also received all other plasmids.

c, PCA plot and **d**, gene set enrichment analysis of RNA sequencing of the GFP⁺Cherry⁺ and GFP⁺Cherry⁻ populations show expression-based differences in DNA repair- and cell cycle-associated genes.

e, Hoechst staining in HEK293T cells transfected with the fluoPEER plasmid and prime editing machinery shows enrichment of reporter editing signal in G2.

f, Sanger sequencing of HEK293T cells sorted based on cell cycle shows differences in genomic editing efficiency. Error bars represent standard deviations from the mean.

Supplementary Information

Supplementary Figures 1-12

Supplementary Note 1: fluoPEER cloning

Supplementary Table 1: oligos used in this study

Supplementary Figures

Supplementary Figure 1: Analysis of the targeting scope of (NGG-)PE2 and flexible PE2 variants within the ClinVar database

a, The number of NGG-, NAN-, or NGN-PAMs within the genomic region surrounding pathogenic variants in the ClinVar database. Only mutations shorter than 51 bp were considered as correctable by prime editing and PAM sites in a window of 10bp upstream to 4bp downstream of the target site were considered usable, based on Kim et al., (2021).

b, Cumulative histogram of the number of available PAMs shows that flexible PE2 variants (NGN+NAN) have at least 9 PAMs available for the repair of ClinVar mutations.

Supplementary Figure 2: Evaluation of reparable mutations in the ClinVar database that are within the scope of the PE_prediction machine learning algorithm

a, Types of mutations of pathogenic variants in the ClinVar database.

b, Number (left panel) and type (right panel) of pathogenic variants in the ClinVar database for which the PE_prediction machine learning algorithm could predict the efficiency of a mutation-correcting pegRNA.

c, Maximum predicted efficiency for repair of ClinVar pathogenic variants for mutations that fall within the scope of the DeepPE deep learning algorithm. Increase in available PAMs for pathogenic variants results in a higher median maximum predicted efficiency.

Supplementary Figure 3: Correlation of prime-editing efficiency of 1bp deletions versus 1bp insertions

a, Data from Kim et al. (2021) was re-analyzed to estimate the correlation of prime editing efficiencies of pegRNAs that create a 1bp deletion (x-axis) to the efficiencies of corresponding pegRNAs with the same spacer and PBS sequence that create a 1bp insertion. No data was available for the comparison of 1bp deletion- to 1bp substitution-edits since Kim et al. performed 1bp substitutions with much longer RTTs than 1bp deletions.

Supplementary Figure 4: Comparison of genomic editing of GFP→BFP to editing on the corresponding fluoPEER plasmid

a, Design of three different pegRNAs converting GFP to BFP (TAC to CAT). The same three pegRNAs remove a single nucleotide insert on the corresponding fluoPEER plasmid, resolving a frameshift mutation upstream of the Cherry cassette. Note that the PAM sequences (purple) of the pegRNA 1, 2, and 3 are NGG, NGN, and NAN, respectively.

b, HEK293T cells containing a lentivirally integrated genomic GFP cassette were transfected with the prime editing machinery to convert GFP to BFP. FACS plot shows occurrence of BFP⁺ cells 14 days after transfection with pegRNA1 (NGG-PAM) and NGG-PE2. Sanger sequencing confirms TAC to CAT conversion in BFP⁺ cells (blue outlines), but not in GFP⁺ cells (green outlines).

c, Editing of integrated genomic GFP (left panel) and editing of the corresponding fluoPEER plasmid (middle panel) in HEK293T cells using the three pegRNA designs from (a) and 7 different PE2 variants ('no PE2' is included in Fig 1d as an additional PE2-condition). Sanger sequencing of the genomic target region (right panel) in the same cells as shown in the left panel illustrates that this technique is not sensitive enough to quantify GFP to BFP conversion in less than 5% of cells. Related to figure 1d.

Supplementary Figure 5: Comparison of fluoPEER ranking to predictions by the DeepPE algorithm

a, Comparison of the ranking of editing efficiencies by fluoPEER (left panel) to the prediction score of the DeepPE algorithm (middle panel) and to actual prime editing efficiencies as reported by Kim et al., 2021. (right panel).

Supplementary Figure 6: Dynamics of fluoPEER under various transfection conditions and read-out times

a, Percentage of GFP+ cells and average of Cherry to GFP ratio in GFP+ HEK293T cells, with decreasing amounts of total transfected plasmid DNA (PE2+pegRNA+fluoPEER). **b**, Percentage of GFP+ cells and average of Cherry to GFP ratio in GFP+ HEK293T cells, with decreasing amounts of the fluoPEER (f-PR) plasmid DNA but equal amounts of PE2 and pegRNA plasmid DNA. As expected, lowering fluoPEER concentration led to decreased percentages of GFP+ cells, but increased reporter editing (Cherry to GFP ratio). **c**, Percentage of GFP+ cells and average of Cherry to GFP ratio in transfected cells, as measured 3 and 6 days after transfection of PE2+pegRNA+fluoPEER.

Supplementary Figure 7: FluoPEER-predicted prime editing in organoids

a, Sanger sequencing of clonal liver-derived organoids with monoallelic and biallelic I1174N mutations in exon 32 of *IARS1*.

b, Clonal organoid lines with monoallelic (2 clones) and biallelic (2 clones) *IARS1*^{I1174N} mutations were continually passaged for 32 days. Clones with biallelic, but not monoallelic, *IARS1*^{I1174N} mutations had lower organoid reconstitution capacity and could be passaged less often due to slower growth.

c, Correction of pathogenic *MUT*^{R369H} mutations in liver organoids derived from an MMA patient could not be corrected using two different pegRNA designs. This corresponds to very low fluoPEER prediction scores for these targeting strategies as shown in Fig. 2e.

d, Organoid reconstitution in medium without Wnt-pathway agonists Rspo1 or Wnt, after prime-editing mediated introduction of stabilizing *CTNNB1* mutations that cause Wnt-independent growth. Note that pegRNA19 significantly outperformed other pegRNA designs in both the fluoPEER analysis (Fig. 2d) and genomic editing of *CTNNB1*. Data in (d) was adapted from earlier experiments within our lab (Schene et al. 2020).

Supplementary Figure 8: Characterization of PAM-specificity and NLS adaptations of novel PE2 variants using fluoPEER

a, Comparison of fluoPEER scores for flexible PE2 variants and the corresponding PE2* variants with nuclear localization sequence (NLS) adaptations as described by Liu et al. (2021).

b, Conversion of genomic GFP to BFP using various PE2 and PE2* variants. Related to Fig. 2c. Note that the PE2* NLS adaptations did not result in an overall increase of fluoPEER scores nor an overall increase in genomic editing efficiency. The data shown in Extended Data Fig. 4c was reused for (b).

Supplementary Figure 9: FluoPEER allows enrichment of gene-edited patient-derived fibroblasts

a, Skin-derived fibroblasts from a patient with biallelic *POLGA*^{467T} mutations were cotransfected with prime editing (PE3) or

b, base editing machinery and corresponding fluoPEER plasmids. Reporter-edited (GFP⁺Cherry⁺) and unedited (GFP⁺Cherry⁻) cells were sorted based on fluoPEER reporter-only conditions (left panels), plated, and cultured for 19 days. Sanger sequencing confirmed increased genome editing in GFP⁺Cherry⁺ versus GFP⁺Cherry⁻ cells for both the prime editing (>90% versus ~45%) and the base editing (~30% versus 0%) conditions.

Supplementary Figure 10: Dynamics of cotransfection of four fluorescent plasmids

a, Mixtures of plasmids (5kb) containing CMV-mTurquoise2, CMV-eGFP, CMV-mKO2, and/or CMV-mCherry were transfected into HEK293T cells and inspected by FACS after 2 days. FACS plots show proper compensation in the 'fluorescence minus one' (FOI) controls

(top row) and strong correlation between fluorescent intensities from different plasmids in the condition containing all 4 plasmids (bottom row), indicating very high cotransfection efficiency. This strongly suggests that 2-4 fold enrichment of genomic editing in GFP⁺Cherry⁺ versus GFP⁺Cherry⁻ cells (Fig. 2a) cannot be the result of inefficient cotransfection of plasmids encoding prime editing machinery (PE2+pegRNA) in GFP⁺Cherry⁻ cells.

Supplementary Figure 11: RNA sequencing analysis of fluoPEER-edited versus unedited cells

a, HEK293T cells were transfected with PE2 and pegRNA plasmids creating either *IARS1*^{I1174N} mutations or *CTNNB1* 6-bp deletions, together with corresponding fluoPEER plasmid. 48 hours after transfection, GFP⁺Cherry⁺ and GFP⁺Cherry⁻ cells were sorted and genomic editing was estimated by Sanger sequencing.

b, RNA-sequencing was performed in replicates of each population from (a). MA-plot shows significant (adjusted *p*-value <0.05) up- and downregulation of 279 and 164 genes, respectively in GFP⁺Cherry⁺ versus GFP⁺Cherry⁻ cells.

c, The left panel shows the enrichment plot of the 'DNA repair' pathway, which was the most enriched 'Reactome Database' pathway in GFP⁺Cherry⁺ versus GFP⁺Cherry⁻ cells. The right panel shows the expression heatmap of enriched genes within this pathway.

d, EnrichR analysis of GFP⁺Cherry⁺ versus GFP⁺Cherry⁻ cells.

Supplementary Figure 12: Flow cytometry gating strategies

a, Example of gating strategy for a fluoPEER experiment in HEK293T cells. The 'P1+P2+P3' gate selects for single cells; the 'live' gate selects for DAPI negative, live cells. The GFP⁺ gate selects for cells that were transfected with fluorescent reporter plasmid. Lower left panel shows the Cherry/GFP plot and population hierarchy for an ineffective (*CTNNB1*, pegRNA1) and lower right panel for an effective (*CTNNB1*, pegRNA3) prime editing strategy.

b, Example of gating strategy for cell cycle phase based on Hoechst signal intensity.

Supplementary Note 1: Plasmid cloning

pegRNA

PegRNAs should be designed according to the following guiding principles:

- 1) pegRNA spacers with high Cas9 sgRNA binding scores. Currently available Cas9 prediction algorithms have been developed using SpCas9 (NGG-recognizing), but do not discriminate between NGG and non-NGG PAMs. For this study, we selected spacers with an on-target score of at least 20, based on an algorithm developed by Doench, Fusi et al. (2016) which was employed in the Benchling web interface.
- 2) pegRNA primer binding sites (PBSs) with length 9-15 and a GC-count of at least 5 and a GC-content of at least 40%. If this is not possible, for instance when targeting a mutation in an AT-rich region, pick a PBS with a GC-count of at least 4 and GC-content >30%.
- 3) pegRNA reverse transcriptase templates (RTTs) with length 9-20, spanning the genomic region at least 5 bp downstream of the intended edit. For example, when the intended edit is located at position +6 from the pegRNA nicking site, the RTT should be at least 11 nucleotides in length. When the RTT-length is less than 12, if possible, pick a G for the last edit.
- 4) If possible, the RTT should create a (silent) PAM mutation to increase editing efficiency. When using the flexible SpRY-PE2, PAM mutations are not expected to increase editing efficiency.

FluoPEER

STEP 1

Digest the pmGFP-P2A-K0-P2A-RFP plasmid by making the following mix:

Component	Volume
1 µg pmGFP-P2A-K0-P2A-RFP plasmid	X
Sall (NEB)	1 µl
Acc65I (NEB)	1 µl
Buffer R3.1 (NEB)	3 µl
Distilled water	To 50 µl

Put the mix at 37 °C for 16 hours and isolate the 6 kb band from gel.

fluoPEER needs a STOP codon or frameshift in the genomic insert to function. In case the genomic target already contains either, 45-100 bp around the mutation of interest can be selected. The genomic insert should at least contain the full spacer and extension (PBS+RTT) sequences of all designed pegRNAs, including a margin of 5bp upstream and 5bp downstream. In case none such frameshift nor STOP codon is present, one will have to be inserted. Important notes:

- 1) The fluoPEER plasmid and not the pegRNA design should be modified, such that the same pegRNAs can be used to edit both the fluoPEER plasmid and the genomic target.
- 2) The PAM site and spacer sequence (including PBS) of all pegRNAs should not be altered by the addition of frameshift mutation or creation of a stop-mutation. In most cases, there is a significant difference in GC-content between the DNA sequence upstream and downstream of the target mutation, offering an obvious optimal region for PBS and spacer binding. This means that most of the tested pegRNAs will bind

the same strand and can therefore be tested on the same fluoPEER plasmid. In case this is unavoidable, design two fluoPEER plasmids for the mutation instead.

STEP 2

Once the genomic insert has been selected, two oligos need to be ordered:

TOP: 5'TCGACC'-[GENOMIC INSERT]-'G'3

BOTTOM: 5'GTACC'[GENOMIC INSERT REVERSE COMPLEMENT]-'GG'3

Anneal the oligos according to the following protocol:

Component	Volume
TOP oligo (100 µM)	1 µl
BOTTOM oligo (100 µM)	1 µl
PNK (NEB)	1 µl
PNK buffer 10x (NEB)	1 µl
Distilled water	6 µl
Total	10 µl

Heat the mix up to 95 °C for 3 minutes and ramp down to 20 °C at 5 °C / min.

Dilute the oligos 1:10 by adding 90 µl water to the mix.

STEP 3

Ligate the plasmid and annealed oligos in the following mix:

Component	Volume
Digested pmGFP-P2A-K0-P2A-RFP plasmid	30 ng
Annealed genomic insert	1 µl @ 1 µM
T4 DNA ligase (NEB)	1 µl
T4 DNA ligase buffer 10X (NEB)	1 µl
Distilled water	To 10 µl

Set at 21 °C for 60 minutes and transform.

References:

Doench, J., Fusi, N., Sullender, M. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 34, 184–191 (2016).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [20210526FluoPEERSupplementaryTables1.pdf](#)
- [20210526FluoPEERSupplementaryFigures.pdf](#)