In vivo genome editing with prime editors allows for the correction of mutations and phenotypes in adult mice with liver and eye diseases

Hyewon Jang

Department of Medical Science

The Graduate School, Yonsei University

In vivo genome editing with prime editors allows for the correction of mutations and phenotypes in adult mice with liver and eye diseases

Hyewon Jang

Department of Medical Science

The Graduate School, Yonsei University

In vivo genome editing with prime editors allows for the correction of mutations and phenotypes in adult mice with liver and eye diseases

Directed by Professor Hyongbum Kim

The Doctoral Dissertation
submitted to the Department of Medical Science,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Doctor of Philosophy

Hyewon Jang

December 2021

This certifies that the Doctoral Dissertation of Hyewon Jang is approved.

Thesis Supervisor: Hyongbum Kim

Thesis Committee Member#1: Sung-Rae Cho

Thesis Committee Member#2: Hong Koh

Thesis Committee Member#3: Dong Jin Joo

Thesis Committee Member#4: Sanghoon Lee

The Graduate School
Yonsei University

December 2021

ACKNOWLEDGEMENTS

I am indebted to so many people and I express great gratitude to all of them. First of all, I sincerely thank professor Hyongbum Kim, my thesis supervisor, for his guidance in research. As my mentor, he has given me sincere advice and encouragement all the time. This work would not have been possible without his intellectual commitment.

I would like to extend my sincere thanks to professors Sung-Rae Cho, Hong Koh, Dong Jin Joo, and Sang Hoon Lee for their advice and comments in the process of reviewing this thesis. Each of my Dissertation Committee members has provided me with professional guidance that led me to develop this study.

In addition, I sincerely appreciate the help of Goosang Yu, Ramu Gopalappa, Sehyuk Kwon, Jung Hwa Seo, professor Jeong Hun Kim, professor Dong Hyun Jo, and professor Daesik Kim in the completion of this thesis. My thanks and appreciations also go to my colleagues in the lab who always help and share their expertise. I express deep and sincere gratitude to all professors and colleagues of the Department of Pharmacology for sharing their perspectives and knowledge. I am also thankful to Younghye Kim, Seonmi Park for assisting with the experiments. Thanks to my dear friend Yurim Jeon who always cheer and encourage me.

Finally, I wish to thank my parents and brother for their love and encouragement. Thank you and I love you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SUPPLEMENTARY TABLES

ABSTRACT

In vivo genome editing with prime editors allows for the correction of mutations and phenotypes in adult mice with liver and eye diseases

Hyewon Jang

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Hyongbum Kim)

The performance of prime editing — a gene-editing technique that induces small genetic changes without the need for donor DNA and without causing double-stranded breaks (DSBs) — in correcting pathogenic mutations and phenotypes needs to be validated in animal models of human genetic diseases. Here, we report the performance of prime editors 2 and 3, delivered by hydrodynamic injection, in mice with the genetic liver disease hereditary tyrosinemia type I, and of prime editor 2, delivered by an adeno-associated viral vector, in mice with the genetic eye disease Leber congenital amaurosis (LCA). For each pathogenic mutation, we identified an optimal prime editing guide RNA by using cells transduced by lentiviral libraries of guide-RNA-encoding sequences paired with the corresponding target sequences. The prime editors corrected the disease-causing mutations in a highly precise manner and led to the amelioration of the disease phenotypes in the mice without detectable off-target effects. This study suggests that prime editing will be a promising approach for treating genetic diseases.

Key words: Prime editing, Genome editing, Hereditary tyrosinemia, Leber congenital amaurosis, Off-target effects

In vivo genome editing with prime editors allows for the correction of mutations and phenotypes in adult mice with liver and eye diseases

Hyewon Jang

Department of *Medical Science*
The Graduate School, Yonsei University

(Directed by Professor Hyongbum Kim)

I. INTRODUCTION

Precise genome editing strategies are required to correct disease-relevant genetic variants including insertions, deletions, and point mutations. Since the development of CRISPR/Cas9 system, Cas9-mediated homology-direct repair (HDR) has been a robust gene editing strategy to correct genetic diseases using an exogenous donor DNA[1]. However, non-homologous end joining (NHEJ) occurs more frequently than HDR in double-stranded break (DSB) induced by Cas9, resulting in the introduction of undesired indels in the targetx. Moreover, HDR is cell cycle-dependent, making it inefficient for many therapeutically relevant cell types, especially in post-mitotic cells[2-4].

Base editing has exceptional promise as a therapeutic means to revert disease-causing point mutations to wild type. Base editing does not require the induction of DSB, thereby preventing indel formation while allowing us to convert a specific nucleotide in the target genomic region. Current base editors consist of either a cytidine deaminase[5-7]or an adenosine deaminase[8] fused to dead Cas9 or Cas9 D10A nickase so that it can convert either a targeted cytosine (C) to thymine (T) or a targeted adenine (A) to guanine (G), respectively. So far, many studies have shown the correction of pathogenic

mutations in animal models using base editors[9-11]. However, there are two major hurdles in the base editing system. First, the bases to be converted must be in an activity window for editing, approximately located from positions 4–8 for cytosine base editors and 4-7 for adenine base editors, counting the PAM as positions 21–23[12]. Second, for target sites with multiple cytosines or adenines within the activity window, the conversion of "bystander" nucleotides can also be induced and lead to the generation of undesired alleles[12].

Prime editor (PE) can induce any small-sized genetic change, including insertions, deletions, and all 12 possible point mutations, without donor DNA or double-strand breaks[13]. PE is a catalytically impaired SpCas9 H840A nickase fused with a Moloney murine leukemia virus (M-MLV) reverse transcriptase (RT) protein. Prime editing guide RNA (pegRNA) is a single guide RNA(sgRNA) with a 3' extension encoding primer binding site and reverse transcriptase template. By forming a complex with the pegRNA, PE binds to a specific target site and directly copy the new sequence from the pegRNA into the target genome[1]. There are four types of prime editors (PEs): PE1, PE2, PE3, and PE3b[13]. PE1 and PE2 are composed of PE and a pegRNA, but PE2 has five mutations within the M-MLV RT, which improves prime editing efficiency. PE3 and PE3b use one additional sgRNA to nick the non-edited DNA strand to increase editing efficiency[13].

Prime editing has been used in cultured mammalian cells[13,14], organoids[15], plants[16], and mouse embryos[17] to introduce genetic changes in a targeted manner. Before prime editing can be applied to treat genetic diseases in human patients, correction of both pathogenic mutations and phenotypes of animal models of human genetic diseases must first be demonstrated. In this study, we show that prime editors delivered into mouse models of genetic liver and eye diseases can precisely correct the disease-causing mutation, leading to functional improvement.

II. MATERIALS AND METHODS

1. Construction of plasmid vectors

To prepare pLenti-NG-PE2-BSD, the Lenti-Split-BE4-N-Blast plasmid[18] was digested with restriction enzymes AgeI and BamHI (New England Biolabs (hereafter, for brevity, NEB)) and a MEGAquick-spin total fragment DNA purification kit (iNtRON Biotechnology) was used to gel purify the linearized plasmid. Fragments of PE2-encoding sequence (Addgene #132775) and NG-Cas9-encoding sequence (Addgene #124163) were amplified by PCR using Phusion Polymerase (NEB). The amplicons and the linearized plasmid were assembled using an NEBuilder HiFi DNA assembly kit (NEB). The assembled plasmid was named pLenti-NG-PE2-BSD.

To prepare px601-PE2, the PE2-encoding sequence (Addgene #132775) and the WPRE sequence (Addgene #52962) were amplified by PCR and cloned into the pX601 plasmid using an NEBuilder HiFi DNA assembly kit (NEB).

To construct px552-pegRNA-CAG-mCherry, gRNA scaffold sequence (Addgene #104174) and CAG-mCherry (Addgene #108685) fragments were PCR-amplified and cloned into the px552 plasmid (Addgene #60958) to make px552-U6-CAG-mCherry. Next, pegRNA sequences were synthesized and cloned into the linearlized px552-U6-CAG-mCherry plasmid, which had been digested with SapI and SpeI (NEB), using an NEBuilder HiFi DNA assembly kit (NEB). These manipulations generated px552-pegRNA-CAG-mCherry.

An sgRNA-expressing vector (Addgene #104174) was digested with BsmBI. sgRNA oligomers were annealed, phosphorylated with T4 PNK, and ligated with the linearized vector to construct gN19-nicking sgRNA.

To generate the trans-splicing PE2 plasmid, the sequence encoding the PE2 N terminus (PE2-N term) (Addgene #132775) and a splicing donor (SD) (Addgene#112734) were PCR-amplified using primers listed in Supplementary Table 7 and cloned into the px601 plasmid using an

NEBuilder HiFi DNA assembly kit (NEB). Similarly, a splicing acceptor (SA) (Addgene#112876), the sequence encoding the PE2 C terminus (PE2-C term) (Addgene #132775), and the WPRE sequence (Addgene #52962) were PCR-amplified and cloned into the pX601 plasmid using an NEBuilder HiFi DNA assembly kit (NEB). These manipulations generated TS-PE2-N plasmid and TS-PE2-C-WPRE plasmid, respectively.

To construct U6-*Rpe65* pegRNA-CAG-mCherry, pegRNA-encoding, sgRNA-encoding, and H1 promoter sequences (Addgene#61089) were either synthesized or PCR-amplified and then cloned into the px552-U6-CAG-mCherry plasmid using an NEBuilder HiFi DNA assembly kit (NEB).

2. Construction of a plasmid library of prime editing guide RNA and target sequence pairs

Oligonucleotides were designed and the library of pegRNA-encoding and target sequence pairs was generated as previously described[14]. Briefly, an oligonucleotide pool containing 435 pairs of *Fah* pegRNA-encoding sequences or 561 pairs of *Rpe65* pegRNA-encoding sequences and target sequences was synthesized by Twist Bioscience (San Francisco, CA). Each oligonucleotide included the following elements: a 19-nt guide sequence, BsmBI restriction site #1, a 15-nt barcode stuffer sequence, BsmBI restriction site #2, the RT template sequence, the PBS, a poly T sequence, an 18-nt barcode sequence (identification barcode), and a corresponding 43~47-nt (*Fah*) or 49~89-nt (*Rpe65*) wide target sequence that included a PAM and an RT template binding region. Oligonucleotides that contained other, unintended BsmBI sites were excluded. The barcode stuffer was later excised by digestion with BsmBI; the identification barcode (located upstream of the target sequence) allowed individual pegRNA and target sequence pairs to be identified after deep sequencing.

The plasmid library containing pairs of pegRNA-encoding and target sequences was prepared using a two-step cloning process[14]. This method effectively prevents uncoupling between paired guide RNA and target sequences during PCR amplification of oligonucleotides[19].

3. Production of lentivirus

8×10^6 HEK293T cells were seeded on 150-mm cell culture dishes containing Dulbecco's Modified Eagle Medium (DMEM). After incubation for 16 hours, the DMEM was exchanged with fresh medium containing 25 μM chloroquine diphosphate; the cells were then incubated for 4 more hours. The plasmid library, psPAX2 (Addgene #12260), and pMD2.G (Addgene #12259) were mixed to yield a total of 40 μg of the plasmid mixture. HEK293T cells were then transfected with this mixture using polyethyleneimine. Fifteen hours later, cultures were refreshed with maintaining medium. At 48 hours after transfection, the supernatant (containing lentivirus) was collected, filtered through a Millex-HV 0.45-μm low protein-binding membrane (Millipore), and aliquoted. Serial dilutions of a viral aliquot were prepared and transduced into HEK293T cells so that the virus titer could be determined. Untransduced cells and cells treated with the serially diluted virus were both cultured in the presence of 2 μg/ml puromycin (Invitrogen). The lentivirus titer was quantified by counting the number of living cells at the time when all the untransduced cells died as previously described[20].

4. Cell library generation and transfection

The cell library was generated as previously described[14]. Briefly, HEK293T cells were seeded on eighteen 150-mm dishes (at a density of 1.2×10^7 cells per dish) and incubated overnight. The lentiviral library was transduced into the cells at a multiplicity of infection of 0.3 to achieve >500× coverage

relative to the initial number of oligonucleotides. After incubation of the cells overnight, untransduced cells were removed by maintaining the cultures in 2 μg/ml puromycin for the next 5 days. The cell library was maintained at a count of at least 7.2×10^7 cells for the entire study to preserve library diversity. Next, a total of 7.2×10^7 cells (from six 150-mm culture dishes, each with 1.2×10^7 cells) were transfected with the pLenti-NG-PE2-BSD plasmid (80 μg per dish) using 80 μl Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. Six hours later, the culture medium was replaced with DMEM supplemented with 10% fetal bovine serum and 20 μg/ml blasticidin S (InvivoGen). Five days later, the cells were harvested for genomic DNA extraction.

5. Genomic DNA preparation and deep sequencing

Genomic DNA was extracted from harvested cells with a Wizard genomic DNA purification kit (Promega) and from mouse tissue using a DNeasy Blood & Tissue kit (Qiagen). Target sequences were PCR-amplified using 2X Taq PCR Smart mix (SolGent). The PCR primers used for the experiment are shown in Supplementary Table 7.

In preparation for evaluating the activities of the pegRNAs in a high-throughput manner, an initial PCR included a total of 350 μg genomic DNA; assuming 10 μg genomic DNA per 10^6 cells, coverage would be more than 1000× over the library. 3.64 μg of genomic DNA per reaction (with a total of 96 reactions) was amplified using primers that included Illumina adaptor sequences, after which the products were pooled and gel-purified with a MEGAquick-spin total fragment DNA purification kit (iNtRON Biotechnology). Next, 100 ng purified DNA was amplified by PCR using primers that included barcode sequences. After gel purification, the amplicons were analyzed using the NovaSeq platform (Illumina, San Diego, CA).

For evaluation of prime editing at endogenous sites, the first PCR for the amplification of the target sequence was performed in a 50-μL reaction volume that contained 3 μg of the initial genomic DNA template for liver samples and 500 ng of DNA template for RPE samples. The second PCR to attach the Illumina barcode sequences was then performed using 50 ng of the purified product from the first PCR in a 30-μl reaction volume. The resulting amplicons were sequenced after gel purification using MiniSeq (Illumina, San Diego, CA) according to the manufacturer's protocols.

In experiments involving *rd12* mice, genomic DNA was extracted from the RPE with an Allprep DNA/RNA mini kit (Qiagen). The on-target *Rpe65* site and potential off-target sites were then PCR-amplified using Phusion® High-Fidelity DNA Polymerase (NEB). For quantifying the prime editing efficiency, an initial PCR was performed to amplify the target sequence in a 30-μL reaction volume that contained 50 ng of genomic DNA from the RPE. To attach Illumina index sequences, a second round of PCR was performed using the purified product from the first PCR (50 ng) in a 30-μl reaction volume. After gel purification, the resulting amplicons were sequenced using the MiniSeq (Illumina) and MiSeq(Illumina) platforms.

6. Analysis of prime editing efficiencies

Previously reported Python scripts[14] were used to analyze the high-throughput results. Each pegRNA and target sequence pair was identified using a 22-nt sequence (the 18-nt barcode and 4-nt sequence located upstream of the barcode). Reads that included the specified edits but not unintended mutations within the wide target sequence were considered to contain NG-PE2-induced mutations. To exclude the background mutations generated during the oligo synthesis and PCR amplification steps, we subtracted the background prime editing frequencies determined in the cell library that had not been treated with NG-PE2 from the observed prime editing frequencies as

previously described[14]. The pegRNA and target sequence pairs with deep sequencing read counts below 100 and those with background prime editing frequencies above 5% were removed from the analyses as previously described[21].

To quantify the prime editing frequencies at endogenous sites, amplicon sequences were aligned to reference sequences using Cas-analyzer[22]. The frequencies of intended and unintended edits, including substitutions and indels, were calculated as the percentage of (number of reads with the edit/number of total reads). The frequencies of such edits in the experimental groups were normalized by subtracting the average frequency of such edits in the control groups to exclude errors originating from PCR amplification and sequencing.

7. Generation of mutant *Fah* target sequence-containing cells

The target region of the *Fah* gene was PCR-amplified from genomic DNA from the $Fah^{mut/mut}$ mouse and cloned into a lentivirus shuttle vector derived from a hygromycin reporter plasmid[23]. Next, lentivirus was produced from the vector and transduced into HEK293Tcells at a multiplicity of infection of 0.5 so that the majority of transduced cells would have a single copy target sequence per cell. After incubation of the cells overnight, untransduced cells were removed by supplementing the culture medium with 2 μg/ml hygromycin for the next 5 days.

8. AAV production

The AAV vectors delivered to *rd12* mice were produced as previously described[24]. Briefly, TS-PE2-N plasmids, TS-PE2-C-WPRE plasmids, or U6-*Rpe65* pegRNA-CAG-mCherry plasmids were co-transfected along with AAV2-capsid plasmids and helper plasmids into HEK293T cells. After three days, cells were lysed; high-grade AAVs were then obtained via iodixanol

(Optiprep, Sigma; D1556) gradient ultra-purification. Finally, viruses were concentrated in phosphate-buffered saline to obtain high-grade AAVs using a centricon (Vivaspin® 20, Sartorius).

9. Animals

All animal study protocols relevant to *Fah*$^{mut/mut}$ mice were approved by the Institutional Animal Care and Use Committee (IACUC) of Yonsei University Health System (Seoul, Korea) and all animal study protocols relevant to mouse eye-related experiments were approved by the IACUC of Seoul National University and Seoul National University Hospital. *Fah*$^{mut/mut}$ mice were provided with water containing 10 mg/L NTBC unless specified. Mating pairs of *rd12* mice (stock no. 005379, The Jackson Laboratory), maintained under conditions of a 12-hour dark-light cycle, produced offspring that were used for subsequent experiments. Subretinal injection was performed on the right eyes of 3-week-old *rd12* mice. C57BL/6 mice, obtained from Central Laboratory Animal, were maintained under a 12-hour dark/light cycle.

10. Gene expression analysis by quantitative RT-PCR

Total RNA was purified using TRIzol (Invitrogen) and reverse-transcribed using AccuPower RT PreMix (Bioneer, Korea). Quantitative RT-PCR was performed using SYBR Green (Applied Biosystems, Carlsbad, CA); the primers are listed in Supplementary Table 7. Gene expression levels were normalized to *Gapdh* levels. All experiments were performed in triplicate.

11. Immunofluorescence

Fah$^{mut/mut}$ mice were euthanized by carbon dioxide asphyxiation. Livers were fixed in 4% paraformaldehyde overnight at 4°C, followed by immersion in 6% sucrose overnight and immersion in 30% sucrose the next day. The livers were embedded in OCT compound (Leica, Wetzlar, Germany), cryosectioned

at 16 μm, and stained with hematoxylin and eosin (H&E) for pathology studies. The liver cryostat sections were washed three times with 1X phosphate-buffered saline and incubated in blocking buffer for 1 hour at room temperature. The sections were then incubated with primary anti-Fah antibody (1:200, Abcam, Cambridge, UK) at 4°C overnight. After washing with 1X phosphate-buffered saline, the sections were incubated with Alexa Fluor® 488 goat anti-rabbit secondary antibody (1:400, Invitrogen, Carlsbad, CA, USA) for 1 hour at room temperature. Cells were mounted on glass slides and nuclei were visualized using fluorescent mounting medium containing 4',6-diamidino-2-phenylindole (DAPI; Vector, Burlingame, CA).

Images were captured using a confocal microscope (LSM700, Zeiss, Gottingen, Germany). Quantification of FAH^+ hepatocytes in the liver was performed for 3-5 mice per group, from >4 liver regions per mouse, with ZEN Imaging software (Blue edition, Zeiss). *rd12* mice were euthanized 6 weeks after subretinal injection and RPE-choroid-scleral tissues were obtained from enucleated eyes. Immunostaining of whole mount tissues with anti-RPE65 antibody (1:100; cat. no. NB100-355AF488, Novus) was done using standard techniques. Nuclei were stained using DAPI (Sigma). Immunostained tissues were observed using a confocal microscope (Leica).

12. Electroretinography

Prior to electroretinography (ERG), mice were kept in the dark overnight. Deep anesthesia was induced, and then a Tropherin ophthalmic solution containing phenylephrine hydrochloride (5 mg/ml) and tropicamide (5 mg/ml) was topically administered to dilate pupils. A universal testing and electrophysiologic system 2000 (UTAS E-2000, LKC) was used for full-field ERG. The light-induced responses to a 0 dB Xenon flash were recorded at a gain of 2 k using a notch filter at 60 Hz; responses were bandpass filtered between 0.1 and 1500 Hz. Graphs were visualized and amplitudes were

estimated using Prism 8 (GraphPad). The a-wave amplitudes were determined by measuring from the baseline to the lowest negative-going voltage, and the b-wave amplitudes were measured from the a-wave trough to the highest peak of the positive b-wave.

13. Optomotor response

A virtual-reality optokinetic system (OptoMotry HD, CerebralMechanics) was used to measure grating acuity visual thresholds, according to the manufacturer's instructions and original publications about the system[25,26]. Briefly, mice were placed on a platform where they were exposed to views of a virtual rotating cylinder on monitors surrounding the enclosure; mice then tracked the grating with head movements. Visual thresholds were determined with a staircase procedure to produce the maximum spatial frequency (cycles/degrees) above which the mice did not respond to the rotating stimuli.

14. Analysis of off-target effects

Potential off-target sites were experimentally identified using Digenome-seq[27] and nDigenome-seq[28], and computationally identified using Cas-OFFinder[29] and CRISPOR[30]. For Cas-OFFinder, genomic sites containing up to 3-bp mismatches compared to the pegRNA were considered and analyzed by targeted deep sequencing. From the CRISPOR results, the four top-ranking predicted off-target sites of the used pegRNA and those of the used sgRNA were selected and analyzed by targeted deep sequencing. Cas-analyzer was used to analyze indel frequencies at off-target sites and the frequencies of the intended edits were analyzed with the same method that was used to analyze frequencies at the on-target sites described above. The sequences of the off-target sites are provided in Supplementary Table 4. The primers used for deep sequencing are shown in Supplementary Table 7.

15. Digenome-seq and nDigenome-seq

Digenome-seq and nDigenome-seq were performed as previously described[27,28]. Briefly, recombinant Cas9 nuclease (100 nM, for Digenome-seq) or Cas9 H840A nickase (100 nM, for Digenome-seq) was incubated with three sgRNAs (each 100 nM, targeting *Fah*, *Rpe65*, and *Atp7b*), which share guide sequences with the pegRNAs of interest, at room temperature for 10 minutes. Next, the resulting Cas9 nuclease or Cas9 H840A nickase/sgRNA complexes were mixed with 20 μg of genomic DNA isolated from NIH3T3 cells in a reaction buffer (100 mM NaCl, 50 mMTris–HCl, 10 mM MgCl2, 100 μg/ml bovine serum albumin, at pH 7.9) and incubated for 8 h at 37°C. Digested genomic DNA was then incubated with RNase A (50 μg/ml) and protease K to remove the sgRNA and the nuclease or nickase, after which a DNeasy tissue kit (Qiagen) was used to purify the DNA again. A Covaris system (Life Technologies) was used to shear the resulting genomic DNA (1 μg) to generate fragments of about 500 bp in size, which were then blunt-ended using End Repair Mix (illumina). To prevent self-ligation, the DNA fragments were adenylated. The fragments were then ligated with adapters using TruSeq DNA Library Prep Kits (illumina). The resulting libraries were then subjected to whole genome sequencing at a sequencing depth of 30x – 40x using an Illumina HiSeq X. Issac aligner was used to align the reads to hg19.

For Digenome-seq, DNA cleavage scores were calculated with previously used source codes (https://github.com/chizksh/digenome-toolkit2)[27]. For nDigenome-seq, aligned sequence data were separated into forward and reverse strands, and *in vitro* single-strand break sites were analyzed using previously used source code (https://github.com/snugel/digenome-toolkit) as previously described[28]. Double- and single-strand break sites with high cleavage scores were subjected to validation by targeted deep sequencing

using genomic DNA extracted from mice treated with prime editing components.

16. Statistics and reproducibility

P-values were determined by Student's t-test using GraphPad Prism8. To determine Pearson and Spearman correlation coefficients, we used Microsoft Excel (version 16.0, Microsoft Corporation). For high-throughput evaluation of pegRNA efficiencies, we combined the data from two replicates independently transfected by two different experimentalists.

III. RESULTS

1. A mouse model of tyrosinemia and high-throughput evaluation of prime editing guide RNAs

As a prototypic disease model for prime editing-based therapeutic genome editing in adult animals, we first chose a mouse model of hereditary tyrosinemia type 1 (*Fah*$^{mut/mut}$), which is caused by a loss-of-function mutation in the fumarylacetoacetate hydrolase (*Fah*) gene[31,32]. The mouse model contains a homozygous G-to-A point mutation at the last nucleotide of exon 8, which causes exon 8 skipping and results in loss-of-function of FAH (**Figure 1A**). Pharmacological inhibition of 4-hydroxyphenylpyruvate dioxygenase, an enzyme that acts upstream of FAH in the tyrosine catabolic pathway, with 2-(2-nitro-4-trifluoromethylbenzoyl)-1,3-cyclohexanedione (NTBC) reduces the accumulation of toxic metabolites and thus prevents hepatic injury[31]. We and others have used this mouse model for testing various *in vivo* genome editing approaches, including a method based on homology directed repair (HDR)[33,34], microhomology-mediated end joining[35], and base editing[10]. In addition, because hydrodynamic injections have commonly been used to deliver genome editing components for Cas9-directed HDR[33], microhomology-mediated end joining[35], and base editing[10] in this mouse model, use of the same delivery method for prime editors should facilitate a comparison of prime editing with the other genome editing approaches.

To find a pegRNA that could induce efficient prime editing to correct the pathogenic point mutation, we first identified PE2 target sequences near the mutation site, which were located at positions ranging from -38 base pairs (bp) to +41 bp from the mutation site. Because only two target sequences with an NGG protospacer adjacent motif (PAM) were available near the mutation site, we generated PE2 based on SpCas9-NG (**Figure 1B**), a SpCas9 variant that has wider PAM compatibility than SpCas9[36-38]. We identified nine target

sequences for SpCas9-NG-based PE2 (for brevity, hereafter, NG-PE2) for the intended editing of the point mutation (**Figure 1C**). Given that Cas9 nuclease activity is one of the most important factors determining PE2 efficiency[14], we first calculated the predicted nuclease activity of SpCas9-NG at the nine sites using DeepSpCas9-NG[38]. One (target id 6) of the nine target sequences showed a very low predicted activity (Supplementary Table 1), so we removed that target sequence from subsequent pegRNA evaluations. For the remaining eight target sequences, we designed pegRNAs containing primer binding sites (PBSs) of five different lengths (7, 9, 11, 13, or 15 nucleotides (nt)) and 10 to 11 different RT template lengths (6 to 40 nt) (**Figure 1D**). To undertake a high-throughput evaluation of the resulting 435 pegRNAs, we generated a lentiviral library containing the 435 pegRNA-encoding sequences paired with the corresponding target sequences[14] (**Figure 1E**). Next, this lentiviral library was transduced into HEK293T cells to make cell libraries. These cell libraries were then transiently transfected with plasmids encoding NG-PE2. Five days after the transfection, the target sequences were analyzed by deep sequencing. The highest prime editing efficiency (9.0%) was observed with a pegRNA (pegRNA id 135) that targets a sequence with an NGT PAM (**Figure 1F**, Supplementary Table 2). The second (7.9%) and third (6.9%) highest prime editing efficiencies were also observed with pegRNAs (pegRNA ids 136 and 137) that target the same sequence targeted by pegRNA id 135. The fourth and fifth highest prime editing efficiencies were 6.6% and 6.5%, which were observed with two pegRNAs (pegRNA ids 88 and 89 in Supplementary Table 2) targeting a common sequence (target sequence id 2) with an NGG PAM. However, according to DeepSpCas9-NG[38], the predicted SpCas9-NG nuclease activity at the NGT PAM-containing target sequence that corresponds to pegRNA id 135 was only 17.0, which was substantially lower than 52.4, the SpCas9 nuclease activity predicted by DeepSpCas9[39] at the NGG PAM-containing target sequence (Supplementary Table 2). Thus, we

expected that the prime editing efficiencies of PE2 with pegRNA id 88 or 89 would be higher than that of NG-PE2 with pegRNA id 135. To test this expectation, we performed the same high-throughput evaluation experiments using PE2 instead of NG-PE2. As expected, PE2-driven prime editing efficiencies were substantially higher than those induced by NG-PE2 at target sequences with NGG PAMs (mean and median fold increases were 3.9- and 3.6-folds, respectively; for accuracy of the fold-increase calculation, pegRNAs that showed NG-PE2-induced efficiencies lower than 0.5% were excluded from the calculation) (**Figure 1G**). The top four highest prime editing efficiencies induced by PE2 were 18.5%, 17.1%, 15.6%, and 15.5% (obtained using pegRNA ids 88, 90, 92, and 89, respectively), which were all higher than the highest prime editing efficiency of 9.0% obtained using NG-PE2 (**Figure 1F**).

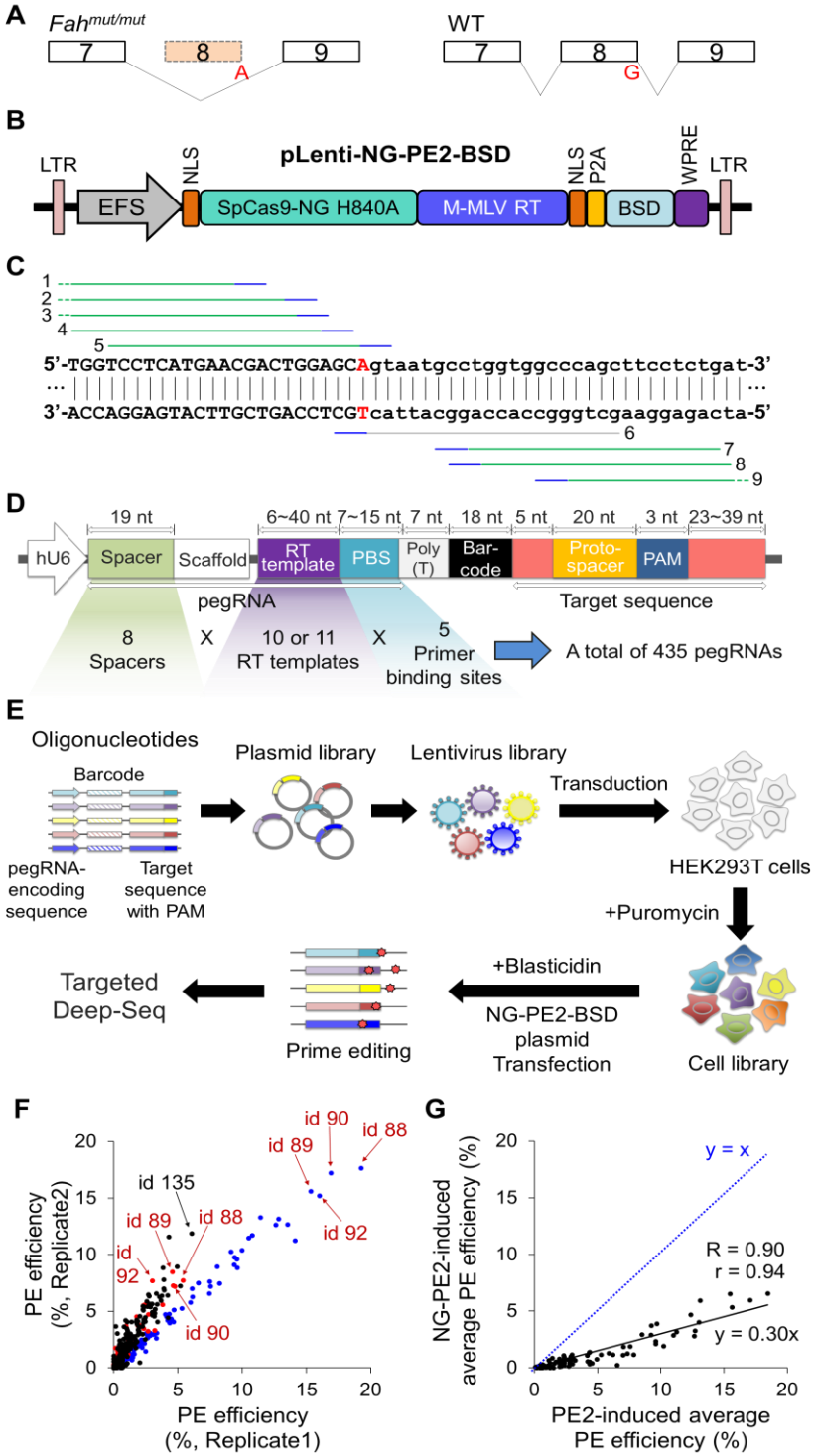

Figure 1. **A mouse model of tyrosinemia and high-throughput evaluation of pegRNAs.** (**A**) Exon 8 skipping in *Fah*$^{mut/mut}$ mice. The G-to-A point mutation (red) at the last nucleotide of exon 8 of the *Fah* gene leads to exon 8 skipping during splicing. (**B**) A schematic representation of the plasmid encoding NG-PE2. NG-PE2, a fusion of SpCas9-NG H840A nickase with Moloney murine leukemia virus reverse transcriptase (M-MLV RT), is expressed from the EFS promoter. The protein encoded by the blasticidin-resistance gene (*BSD*) is co-expressed as a fusion with PE2, from which it is cleaved by the self-cleaving P2A. LTR, long terminal repeat; EFS, elongation factor 1α short promoter; NLS, nuclear localization sequence; P2A, porcine teschovirus-1 2A; WPRE, woodchuck hepatitis virus posttranscriptional regulatory element. (**C**) SpCas9-NG-based PE2 target sequences for correction of the disease-causing point mutation in *Fah*. The protospacer and PAM of each target sequence are represented with green and blue lines, respectively. The numbers on the left and right indicate the target sequence ids. Exon and intron sequences are shown in upper- and lower-case letters, respectively. (**D**) A schematic representation of the lentiviral library of pegRNA-target sequence pairs. Each pegRNA is paired with a wide target sequence that includes a protospacer, a PAM, and neighboring sequences. pegRNA expression is driven by the human U6 promoter (hU6). The library included a total of 435 pegRNA-target pairs, with the pegRNAs containing PBSs of five different lengths (7, 9, 11, 13, or 15 nt) and 10 to 11 different RT template lengths (6 to 40 nt). Spacer, guide sequence of pegRNA; RT, reverse transcriptase; PBS, primer binding site. (**E**) Schematic representation of the high-throughput evaluation of pegRNA activities. A lentiviral plasmid library was prepared from a pool of oligonucleotides that contained pairs of pegRNA-encoding sequences and corresponding target sequences. Next, HEK293T cells were transduced with lentivirus generated from the plasmid library to construct a cell library and untransduced cells were removed by

puromycin selection. This cell library was then transfected with a plasmid encoding NG-PE2, and untransfected cells were removed by blasticidin selection. Five days after the transfection, genomic DNA was isolated from the cells, PCR-amplified, and subjected to deep sequencing to determine prime editing efficiencies. (**F**) Prime editing efficiencies in replicates independently transfected with the NG-PE2- or PE2-encoding plasmid. The red and black dots indicate NG-PE2-induced editing using pegRNAs with corresponding target sequences with NGG and NGH PAMs, respectively, whereas the blue dots represent PE2-induced editing using pegRNAs with target sequences with NGG and NGH PAMs. The ids of pegRNAs that showed high efficiencies using NG-PE2 and PE2 are indicated with red and black arrows. The number of pegRNA and target sequence pairs n = 339 for experiments in which NG-PE2 was used (black and red dots) and n = 339 for experiments in which PE2 was used (blue dots) (out of 435 pairs, those evaluated in both replicates after filtering out pairs with insufficient read counts or high levels of background editing (Methods) are shown). (**G**) Comparison of PE2- and NG-PE2-induced prime editing efficiencies at target sequences with NGG PAMs. The position where PE2-and NG-PE2-induced efficiencies would be the same is shown using a blue dashed line (y = x). The Spearman (R) and Pearson (r) correlation coefficients are shown. The number of pegRNA and target sequence pairs n = 100.

2. Evaluation of prime editor 2 induced editing efficiencies using target sequence-containing cells

We next tested the three pegRNAs that showed the highest editing efficiencies with NG-PE2 (ids 135, 136, and 137) and the three pegRNAs associated with an NGG PAM that showed the highest efficiency with NG-PE2 (ids 88, 89, and 90) by individual evaluations. We generated

HEK293T cells containing the target sequence by transduction of target sequence-containing lentiviral vector (**Figure 2A**). The cells were transiently transfected with plasmids encoding NG-PE2 and pegRNA 135, 136, or 137 or plasmids encoding PE2 and pegRNA 88, 89, or 90. Deep sequencing showed that the average PE2-directed prime editing efficiencies of pegRNAs 88, 89, and 90 were 17.6%, 18.7%, and 12.9%, respectively, which were higher than the NG-PE2-induced prime editing efficiencies of pegRNAs 135, 136, and 137 (3.9%, 3.6%, and 6.3%, respectively) (**Figure 2B**). Thus, we chose pegRNA 89, which showed the highest editing efficiency, for subsequent studies.

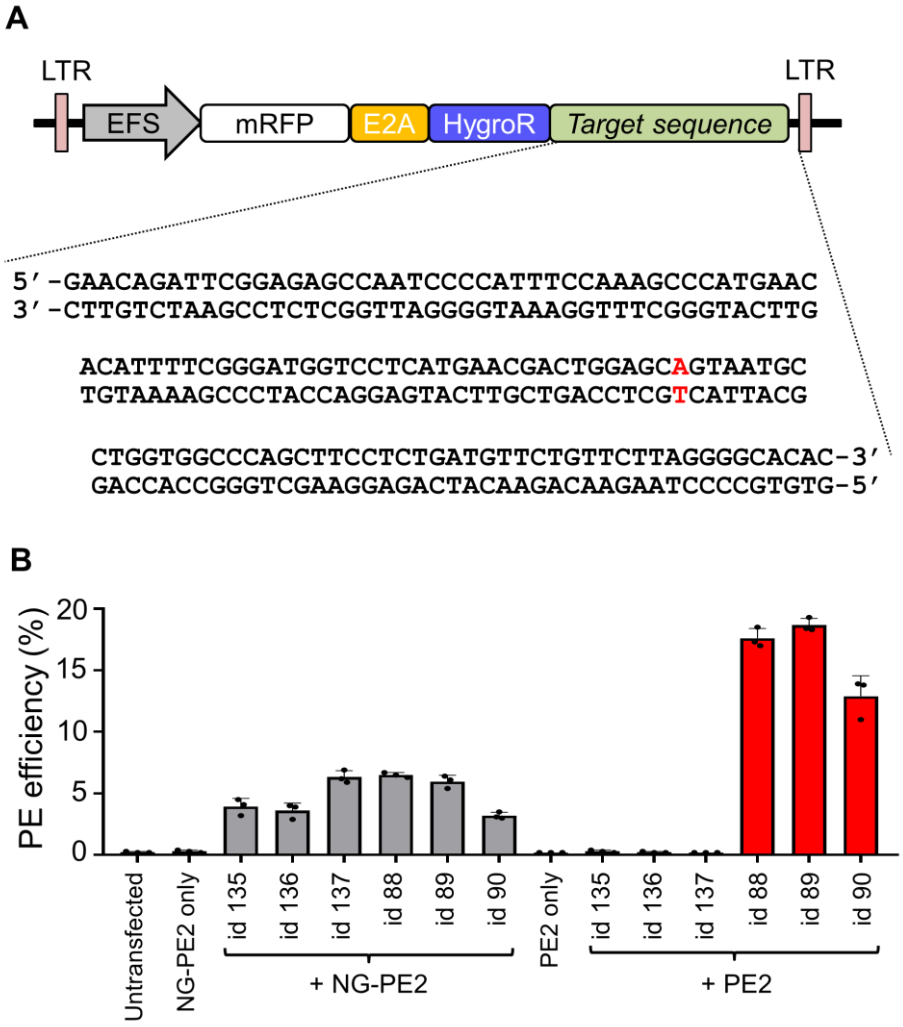

Figure 2. Evaluation of prime editor 2 efficiencies using target sequence-containing cells. (A) Schematic representation of the lentiviral vector containing the mutant *Fah* target sequence found in *Fah*$^{mut/mut}$ mice. The target sequence is shown; red indicates the mutant base pair. HEK293T cells were transduced with this lentiviral vector to generate target sequence-containing cells. (**B**) Prime editing efficiencies of six pegRNAs (ids 88, 89, 90, 135, 136, and 137) in the target sequence-containing cells when

combined with NG-PE2 or PE2. Frequencies were measured five days after transfection of a pegRNA-encoding plasmid and a PE2- or NG-PE2-encoding plasmid. The id of the pegRNA used is shown on the x-axis. Data are mean ± s.d. The number of independent transfections $n = 3$.

3. Prime editor 3 corrects the disease mutation and phenotype in *Fah*$^{mut/mut}$ mice

The initial study of prime editing suggested that the editing efficiency of PE3 and PE3b would generally be, albeit not always, higher than that of PE2 when tested in cultured mammalian cells[13]. Thus, we attempted to use PE3 or PE3b for *in vivo* genome editing by adding an sgRNA. Based on the DeepSpCas9 score[39], we selected a highly active sgRNA (id 1) that enables PE3 (**Figure 3A**, Supplementary Table 3).

We next delivered plasmids encoding the selected pegRNA, the sgRNA, and PE2 (**Figure 3B-D**) into 5- to 7-week-old *Fah*$^{mut/mut}$ mice using hydrodynamic injection, after which the mice were treated with NTBC for 7 days (**Figure 4A**). After discontinuation of NTBC, all mice that had received PE3 (i.e., PE2, pegRNA, and sgRNA) survived until the end of the experiment (40 days), whereas all mice injected with phosphate-buffered saline as a negative control showed substantial weight loss and died before 30 days (**Figure 4B**). This extended survival and the prevention of weight loss suggest PE3-induced amelioration of the disease phenotype, which is consistent with the results of previous studies involving genome editing in this mouse model[10,31-33,35].

To evaluate whether prime editing rescues exon 8 skipping, at the end of the experimental period (40 days) we conducted reverse-transcription PCR (RT-PCR) using liver mRNA as the template and primers binding exons 5 and 9[10]. A 305-bp PCR amplicon, which indicates exon 8 skipping, was observed for *Fah*$^{mut/mut}$ mice, whereas a single 405-bp PCR amplicon, which indicates

that exons 5 to 9 are intact, was seen for wild-type mice (**Figure 4C**). All five mice injected with PE3 showed both 305- and 405-bp amplicons, suggesting that exon 8 skipping was rescued in a fraction of hepatocytes. Sequencing of the 405-bp amplicon confirmed that the mutant sequence was corrected at the mRNA level (**Figure 4D**). Quantitative RT-PCR revealed that the relative average level of exon 8-containing *Fah* mRNA in PE3-treated *Fah*$^{mut/mut}$ mice was 12% of that in wild-type mice, whereas such mRNA was not detectable in control *Fah*$^{mut/mut}$ mice (**Figure 4E**), corroborating that PE3 corrected the exon 8-skipping mutation.

We next quantified the frequency of FAH$^+$ cells in the livers of PE3-treated mice. Immunofluorescence staining showed that FAH$^+$ cells were present at an average frequency of 0.07% (range, 0.01% to 0.12%) at day 0 (the day NTBC was discontinued) and at an average frequency of 61% (range, 45% to 75%) at day 40 (**Figure 4F and Figure 5**). Deep sequencing of liver DNA revealed that the intended edit was not detectable at day 0 (data not shown), which is in line with a previous study of HDR-based genome editing using this mouse model[33]. The intended edit was present at an average frequency of 11.5% (range, 6.7% to 18%) at day 40 (**Figure 6A**). The reason that the frequency of FAH$^+$ cells is higher than the frequency of editing at the DNA level would be because the majority of hepatocytes are polypoid[40] and because nonparenchymal cell DNA is mixed with that of hepatocytes; similar results were observed in the previous genome editing studies using this mouse model[10,33]. The observed editing efficiencies are comparable to those obtained with previous approaches using HDR (9.3% at 33 days after the delivery of genome editing components and 30 days after NTBC withdrawal)[33], microhomology-mediated end joining (5.2% at 37 days after the delivery of genome editing components and 30 days after NTBC withdrawal)[35], and base editing (9.5% at 38 days after the delivery of genome editing components and 32 days after NTBC withdrawal)[10] in this mouse model when the genome

editing components were delivered using hydrodynamic injections, although direct comparisons are difficult due to the differences in the time points at which the editing efficiencies were analyzed and at which NTBC was discontinued.

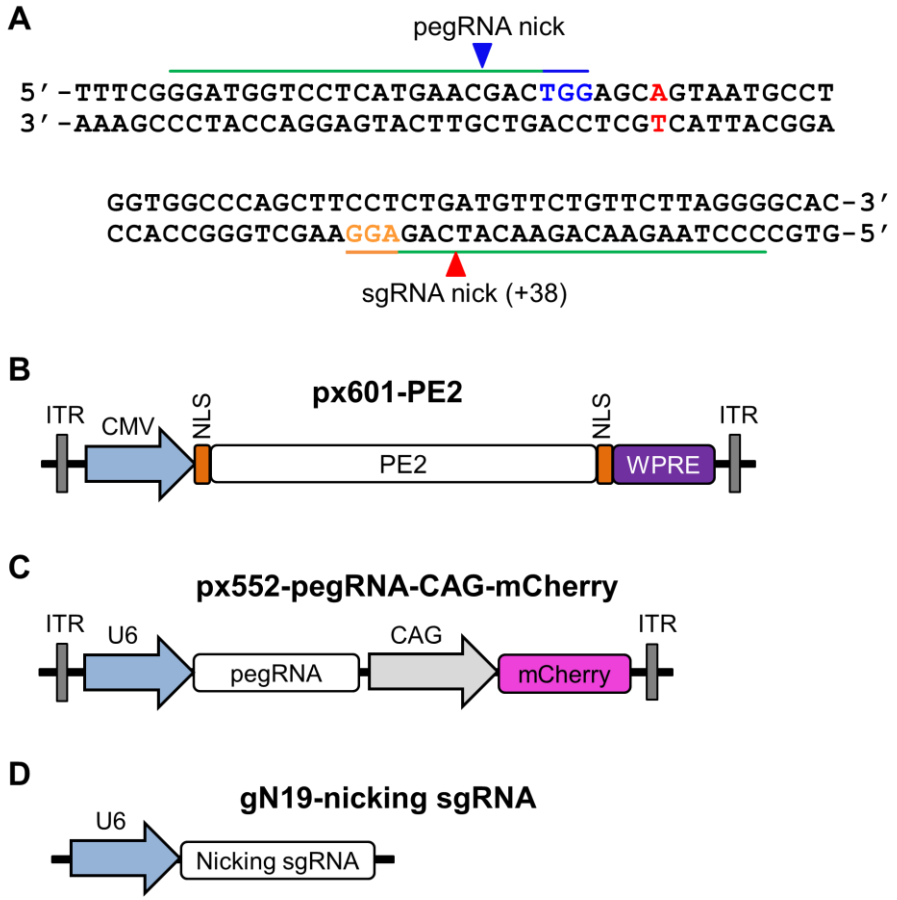

Figure 3. The target sequence in *Fah*$^{mut/mut}$ mice and maps of vectors encoding PE3 components used for treatment. (**A**) The target and neighboring sequences. Green lines represent the pegRNA and sgRNA spacers. The PAMs of the pegRNA and sgRNA target sequences are highlighted in blue and orange, respectively. The disease-causing G>A point mutation is shown in red. (**B-D**) Vector maps of plasmids encoding PE2 (**B**), pegRNA (**C**),

and sgRNA (**D**). ITR, inverted terminal repeat; CMV, cytomegalovirus promoter; NLS, nuclear localization sequence; WPRE, woodchuck hepatitis virus posttranscriptional regulatory element; CAG, CMV early enhancer/chicken β actin promoter; U6, human U6 promoter.

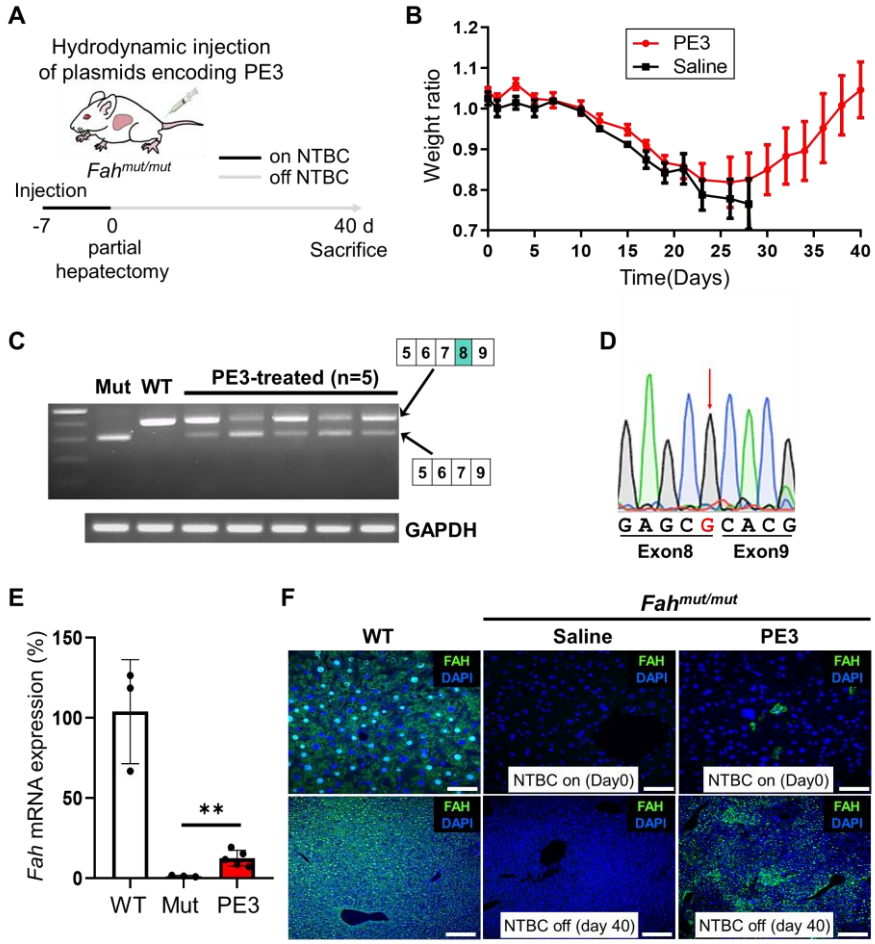

Figure 4. **Prime editor 3 corrects the disease mutation and phenotype in *Fah*$^{mut/mut}$ mice. (A)** A schematic representation of the experiments. *Fah*$^{mut/mut}$ mice underwent hydrodynamic injection of plasmids encoding prime editor 3 components (i.e., prime editor 2, pegRNA, and sgRNA) and were kept on

water containing NTBC for 7 days. The day on which NTBC was withdrawn is defined as day 0. At day 0, a partial hepatectomy was performed to collect liver tissue. At 40 days, the PE3-treated mice were euthanized and analyzed. (**B**) Body weight of *Fah*$^{mut/mut}$ mice injected with PE3 or phosphate-buffered saline (Saline, control). Body weights were normalized to the pre-injection weight. The number of mice n = 5 for the PE3 group and n = 3 for the saline group. Data are mean ± s.e.m. (**C**) Representative RT–PCR using RNA isolated from the liver at 40 days. The primers hybridized to sequences in exons 5 and 9. The wild-type *Fah* (+/+) amplicon is 405 bp in length and the mutant amplicon (which lacks exon 8) is 305 bp. *Gapdh* was used as a control. (**D**) Representative results from Sanger sequencing of the 405-bp RT–PCR band from the PE3-treated mice shown in (**C**). The red arrow indicates the corrected G nucleotide, which is shown in red type. (**E**) The amount of wild-type *Fah* mRNA in the liver was measured by quantitative RT-PCR using primers that hybridze to sequences in exons 8 and 9. WT, wild-type mice; Mut, *Fah*$^{mut/mut}$ mice; PE3, *Fah*$^{mut/mut}$ mice injected with plasmids encoding PE3 components. Data are mean ± s.d. The number of mice n = 3 (WT), 3 (Mut), and 5 (PE3). ***P*-value = 0.0093. (**F**) Immunofluorescence staining of FAH protein. Saline, phosphate-buffered saline. Scale bars, upper panels, 50 μm; lower panels, 200 μm.

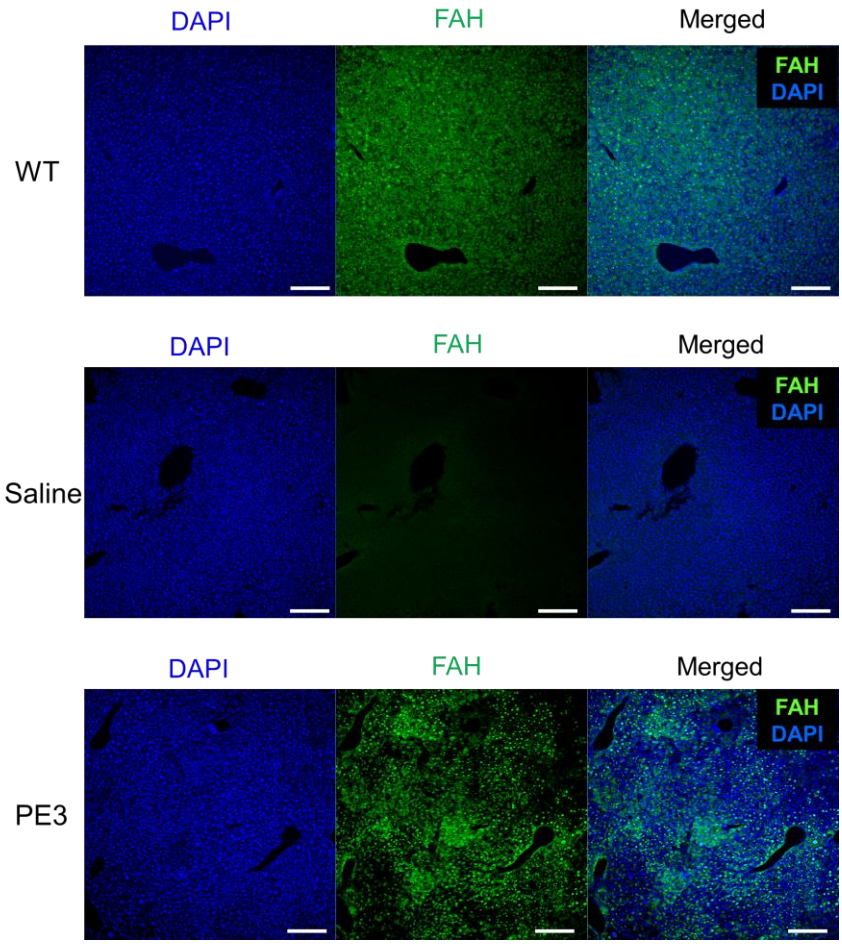

Figure 5. Representative immunofluorescence images. FAH protein expression in wild-type (WT), phosphate-buffered saline-treated *Fah*$^{mut/mut}$ (Saline), and PE3-treated *Fah*$^{mut/mut}$ (PE3) mice. Merged images are shown on the right. Scale bar, 200 μm.

4. Prime editor 3 corrects the disease-causing mutation in a highly precise manner

We next determined whether PE3 induced any unintended editing including indels in or near the target sequence in the mouse liver. Deep sequencing

revealed that unintended substitutions at or near the targeted nucleotide were not detected in any of the mice and that the level of indels ranged from only 0.4% to 1.2% (on average, 0.78%) (**Figure 6A and 6B**). In the case of the Cas9-directed HDR approach, the frequency of indels observed at the target site was 26%, which is 33-fold higher than the level induced by PE3 in this study. When adenosine base editor was used, the frequency of unintended substitutions observed near the target nucleotide was 1.9%[10]. Thus, these data suggest that prime editing can be more precise than the other approaches in mouse somatic cells.

To quantify the off-target effects of PE3, we identified 11 potential off-target sites of the used pegRNA using Digenome-seq and nDigenome-seq, unbiased experimental methods to find potential off-target sites for a pegRNA[27,28], and CRISPOR[30], a computational method, and four potential off-target sites of the used sgRNA using CRISPOR[30] (Supplementary Table 4). Given that prime editor is based on H840A Cas9 nickase, nDigenome-seq should be sufficient as an unbiased experimental method for identifying potential off-target sites. However, given that most studies, including the initial study of prime editing, used Cas9 nuclease to experimentally[13,41] or computationally[17,42] identify potential off-target sites of prime editing, using both Digenome-seq and nDigenome-seq would be a more inclusive approach, which could lead to more thorough analysis of potential off-target sites and enable possible comparisons with results from the other studies based on Cas9 nuclease. Deep sequencing revealed no off-target effects, including intended substitutions or indels, at any of the 15 sites (**Figure 6C-E**). This highly specific editing by prime editor 3 is in line with recent results showing that off-target effects were not detectable when PE3 was used in mouse embryos[17] or human organoids[15] and is also compatible with the low frequency of off-target effects of PE3 in cultured mammalian cells[13,28].

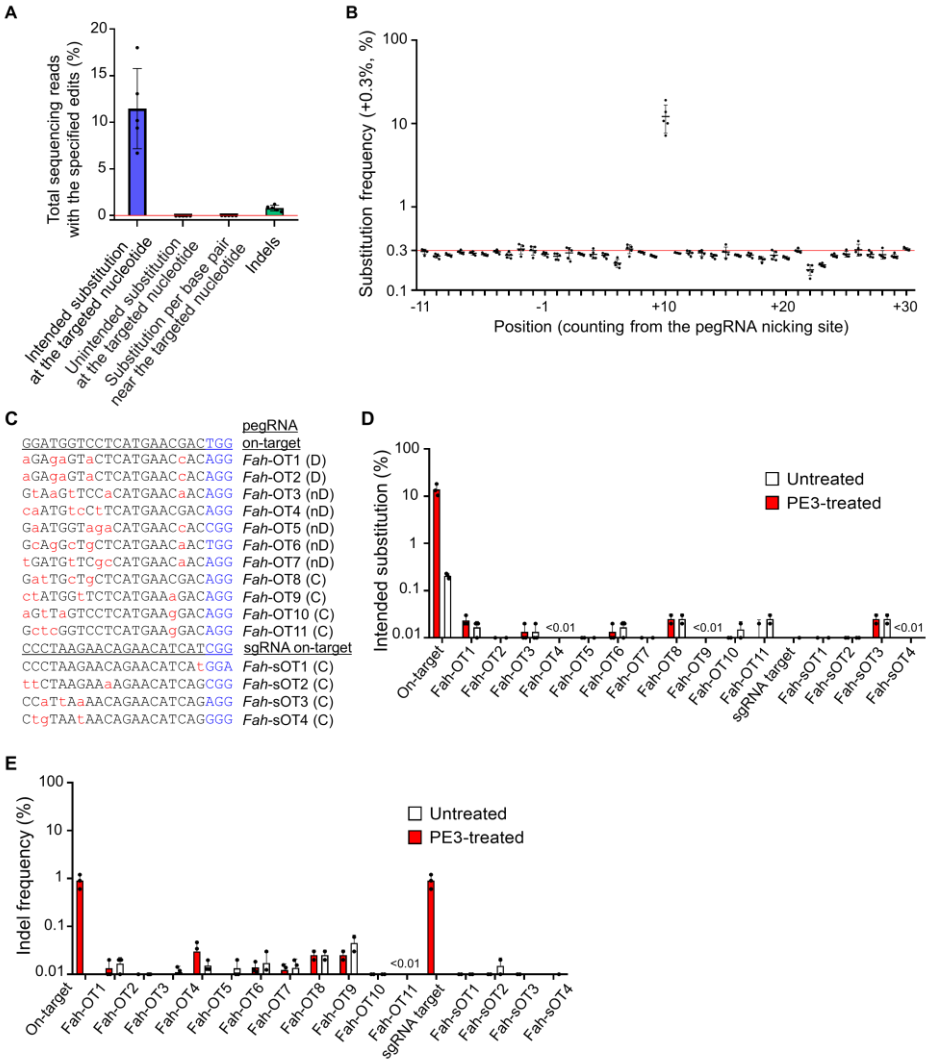

Figure 6. **Prime editor 3 corrects the disease-causing mutation in a highly precise manner.** (**A**) Frequencies of intended and unintended edits in the livers of PE3-treated mice. The frequencies were normalized by subtracting the average frequency of such editing in the control group injected with phosphate-buffered saline to exclude errors originating from PCR amplification and sequencing. Substitutions near the targeted nucleotide were evaluated over a 40-bp range centered on the targeted nucleotide. Indels were

assessed over a 136-bp range centered on the pegRNA nicking site. The red horizontal line represents the point where the normalized frequency = 0. Data are mean ± s.d. The number of mice n = 5. (**B**) Substitution frequencies at positions ranging from -20 bp to +20 bp of the target nucleotide in PE3-treated mice. The frequencies were normalized by subtracting the average edit frequencies in the control group injected with phosphate-buffered saline to exclude errors originating from PCR amplification and sequencing. The red horizontal line represents the location where the normalized frequency = 0. Positions are numbered from the pegRNA nicking site. The targeted position is at +10. Data are mean ± s.d. The number of mice n = 5. (**C**) Potential off-target sites experimentally captured by Digenome-seq (D) and nDigenome-seq (nD) or computationally predicted by CRISPOR (C). Nucleotides in red indicate mismatched sequences and nucleotides in blue represent PAM sequences. (**D**, **E**) Frequencies of the intended substitution (**D**) and indels (**E**) at the 11 potential off-target sites for the used pegRNA (id 89) (OT1, ..., OT11) and at the four potential off-target sites for the used sgRNA (sOT1, …, sOT4) in PE3-treated liver tissues. Genomic DNA isolated from the livers of *Fah*$^{mut/mut}$ mice without PE3 treatment was used as the negative control (Untreated). The number of mice n = 2 or 3.

5. Prime editor 2 corrects the disease mutation and phenotype in *Fah*$^{mut/mut}$ mice

We also performed similar experiments using PE2, which does not require an sgRNA, instead of PE3 (**Figure 7A**). *Fah*$^{mut/mut}$ mice treated with PE2 (not NG-PE2) and the pegRNA used for the PE3 experiments described above survived until the end of the experiment (60 days after the initial NTBC withdrawal), whereas the control *Fah*$^{mut/mut}$ mice all died within 30 days (**Figure 7B**). Quantitative RT-PCR showed that the expression level of *Fah*

mRNA containing intact exon 8 in PE2-treated mutant mice was on average 6.9% that in wild-type mice, but that the expression in untreated mutants was undetectable (**Figure 7C**). Immunofluorescence staining showed that an average of 33% of liver cells from PE2-treated mutant mice were FAH^+ at 60 days (**Figure 7D**).

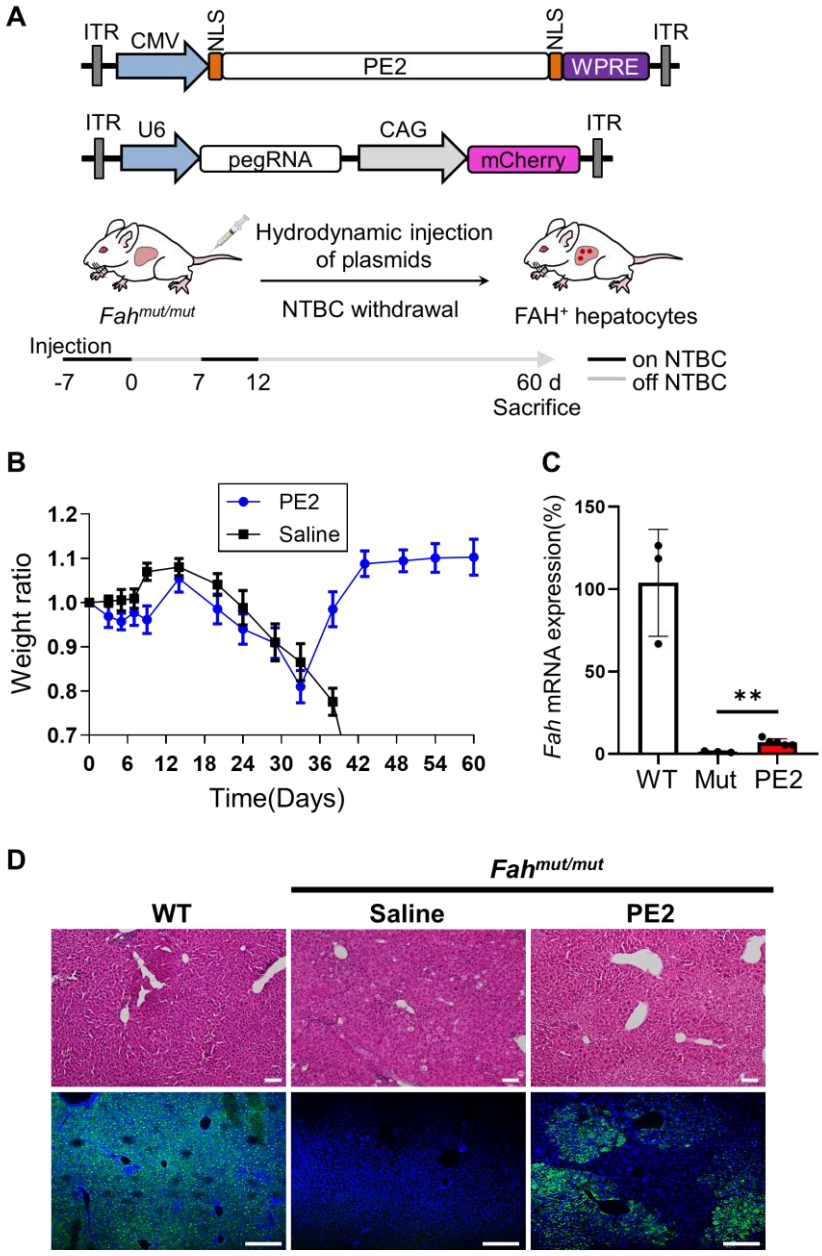

Figure 7. Prime editor 2 corrects the disease mutation and phenotype in *Fah*$^{mut/mut}$ mice. (A) Maps of vectors encoding PE2 components and a schematic representation of the experiments. *Fah*$^{mut/mut}$ mice underwent

injection of plasmids encoding prime editor 2 components (i.e., prime editor 2 and pegRNA) and were kept on water containing NTBC for 7 days. The day on which NTBC was initially withdrawn is defined as day 0. The mice were again provided with NTBC for five days, from day 7 to day 12, after the initial withdrawal of NTBC at day 0. At 60 days, the PE2-treated mice were euthanized and analyzed. ITR, inverted terminal repeat; CMV, cytomegalovirus promoter; NLS, nuclear localization sequence; WPRE, woodchuck hepatitis virus posttranscriptional regulatory element; CAG, CMV early enhancer/chicken β actin promoter; U6, human U6 promoter. (**B**) Body weight of *Fah*$^{mut/mut}$ mice injected with PE2 or phosphate-buffered saline (Saline, control). Body weights were normalized to the pre-injection weight. The number of mice $n = 5$ for the PE2 group and $n = 3$ for the saline group. Data are mean ± s.e.m. (**C**) The level of wild-type *Fah* mRNA in the liver measured by quantitative RT-PCR using primers that hybridize to exons 8 and 9. WT, wild-type mice; Mut, *Fah*$^{mut/mut}$ mice; PE2, *Fah*$^{mut/mut}$ mice injected with plasmids encoding PE2 components. The number of mice $n = 3$ (WT), 3 (Mut), and 5 (PE2). $^{**}P = 0.0032$. (**D**) H&E staining (upper panels) and immunofluorescent staining for FAH protein (lower panels) in liver sections. Scale bars, upper panels, 100 μm; lower panels, 200 μm.

6. Prime editor 2 corrects the disease-causing mutation without any detectable unintended substitutions, indels, bystander effects, or off-target effects

We analyzed genomic DNA isolated from the PE2-injected mice. When the mice were treated with NTBC up to day 0 (7 days after PE2 injection) without partial hepatectomy, the intended editing was not detected (data not shown). When analyzed at the end of the experiments (60 days after the initial NTBC withdrawal) by deep sequencing, we found that the intended edit was present

in an average of 4.0% (range 2.4% to 5.9%) of the total sequencing reads (**Figure 8A**). No unintended edits, including indels, unintended substitutions at the targeted nucleotide, and bystander nucleotide edits, were detectable (**Figure 8A and 8B**). Furthermore, when we evaluated the 11 potential off-target sites of the used pegRNA using Digenome-seq[27], nDigenome-seq[28], and CRISPOR[30] (Supplementary Table 4), no off-target effects, including substitution mutations or indels, were detected, either (**Figure 8C and 8D**). Taken together, these results suggest that PE2-mediated prime editing in mice can correct the disease-causing mutation in a highly precise and specific manner, which could not have been achieved using other previously used genome editing approaches based on either a CRISPR nuclease or a base editor.

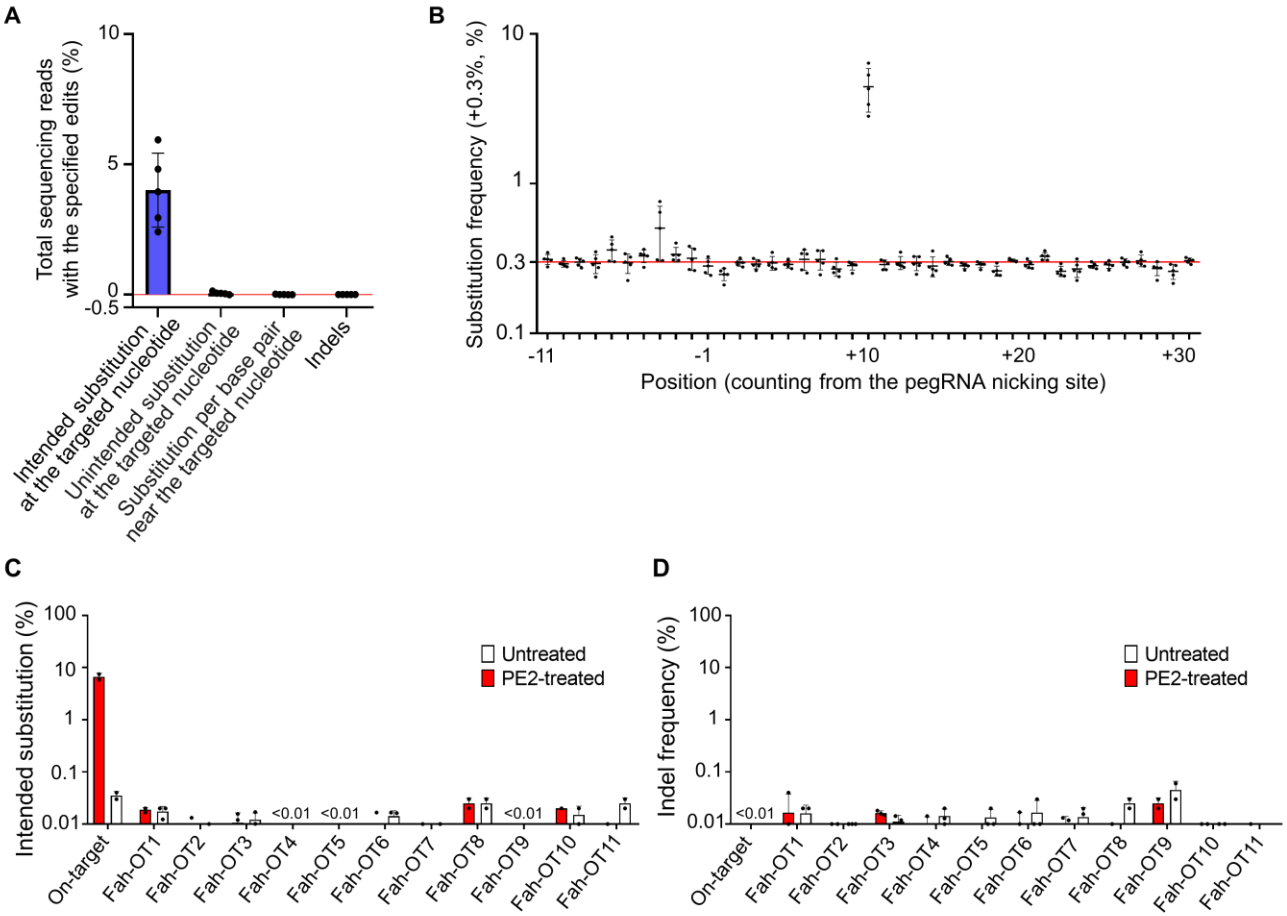

Figure 8. **Prime editor 2 corrects the disease-causing mutation without any detectable unintended substitutions, indels, bystander effects, or off-target effects.** (**A**) Frequencies of intended and unintended edits in the livers of PE2-treated mice. The frequencies were normalized by subtracting the average frequency of such editing in the control group injected with phosphate-buffered saline to exclude errors originating from PCR amplification and sequencing. Substitutions near the targeted nucleotide were evaluated over a 40-bp range centered on the targeted nucleotide. Indels were assessed over a 136-bp range centered on the pegRNA nicking site. The red horizontal line represents the location where the normalized frequency = 0. Data are mean ± s.d. The number of mice n = 5. (**B**) Substitution frequencies at positions ranging from -20 bp to +20 bp of the target nucleotide in PE2-treated mice. The frequencies were normalized by subtracting the average edit frequencies in the control group injected with phosphate-buffered saline to exclude errors originating from PCR amplification and sequencing. The red horizontal line represents the location where the normalized frequency = 0. Positions are numbered from the pegRNA nicking site. The targeted position is at +10. Data are mean ± s.d. The number of mice n = 5. (**C, D**) Frequencies of the intended substitution (**C**) and indels (**D**) at the 11 potential off-target sites for the used pegRNA in PE2-treated liver tissues. The sequences of the 11 potential off-target sites are shown in Fig. 4c. Genomic DNA isolated from the livers of *Fah*$^{mut/mut}$ mice without PE2 treatment was used as the negative control (Untreated). The number of mice n = 2 or 3.

Maps of vectors encoding PE2 components and a schematic representation of the experiments. *Fah*$^{mut/mut}$ mice underwent

7. Identification of prime editing guide RNAs for *in vivo* genome editing in a mouse model of Leber congenital amaurosis (LCA)

We next investigated whether prime editing could be achieved in another disease model. Leber congenital amaurosis (LCA) encompasses a group of monogenic genetic eye diseases involving retinal degeneration that causes severe early-onset visual deterioration[43,44]. So far, mutations in at least 18 genes have been reported to be associated with LCA. One representative gene, *RPE65*, encodes an isomerohydrolase that produces 11-cis retinal, which is essential in the visual cycle[43,45]. Several therapeutic approaches have been tested to rescue the pathology of LCA. Subretinal injection of AAV encoding wild-type RPE65 improved visual function in human patients[46]. However, we cannot rule out the possibility that the exogenous transgene might be silenced after a long period of time[47,48]. CRISPR-Cas9 nuclease and antisense oligonucleotides have been used to bypass a splicing defect-inducing mutation in *CEP290*, another representative LCA-causing gene, in primates and human patients, respectively[49,50]. However, the antisense oligonucleotide approach required repeated injections and the Cas9-based approach that involved bypassing a splicing defect can induce unintended indels and is restricted to mutations causing splicing defects. A more promising approach would be to correct the mutation, generating the wild-type sequence. As potential methods for correcting an LCA-associated mutation, CRISPR-Cas9-directed HDR[51] and adenine base editing[52,53] have been evaluated in *rd12* mice, a mouse model of *RPE65*-related LCA. However, these approaches resulted in limited precision, causing substantial indel frequencies[51] or unintended bystander editing[52,53]. A method for the precise correction of the disease mutation in eyes would be a promising option for treating LCA and other genetic diseases.

To determine whether delivery of a PE using a clinically applicable method could both precisely correct the causal mutation and rescue the phenotype of a genetic eye disease, we used *rd12* mice, which possess a nonsense mutation

caused by a C-to-T transition at position 130 in exon 3 of the *Rpe65* gene (p.R44X)[54] (**Figure 9A**). To find an efficient pegRNA for prime editing of this point mutation, we identified possible target sequences with NGN PAMs at positions ranging from -52 bp to +59 bp away from the mutation site (**Figure 9B**). Among the resulting ten target sequences, six had NGG PAMs and four had NGH PAMs. We designed a total of 561 pegRNAs containing 5 to 6 different lengths (7, 9, 11, 13, 15, or 17 nt) of PBSs and 9 to 17 different lengths (6 to 62 nt) of reverse transcriptase (RT) templates for the ten target sequences (**Figure 9C**). To evaluate the efficiencies of these pegRNAs, we constructed a lentiviral library containing the 561 pairs of pegRNA-encoding and corresponding target sequences as similarly performed above. HEK293T cells were sequentially transduced with the lentiviral library and transfected with plasmids encoding NG-PE2 as conducted for the $Fah^{mut/mut}$ model. Deep sequencing of the treated cells showed that the highest (15%) and second highest (14%) average prime editing efficiencies were achieved with two pegRNAs (pegRNA ids 157 and 198, respectively) that recognized the same target sequence with an NGG PAM (id 5) and that had the same PBS length (9 nt) (**Figure 9D**, Supplementary Table 5). The only difference between the two pegRNAs was the RT template length (13 and 14 nt in ids 157 and 198, respectively). Because the first replicate study showed that pegRNA id 198 had the highest efficiency, we chose it for the subsequent quick application of prime editing in *rd12* mice; however, the average of results from two replicates later showed that pegRNA id 157 had the highest efficiency. In addition, we used PE2 instead of NG-PE2 because SpCas9 showed higher activities than SpCas9-NG at target sequences with an NGG PAM[38] and PE2 showed higher general activities than NG-PE2 at targets with an NGG PAM[55].

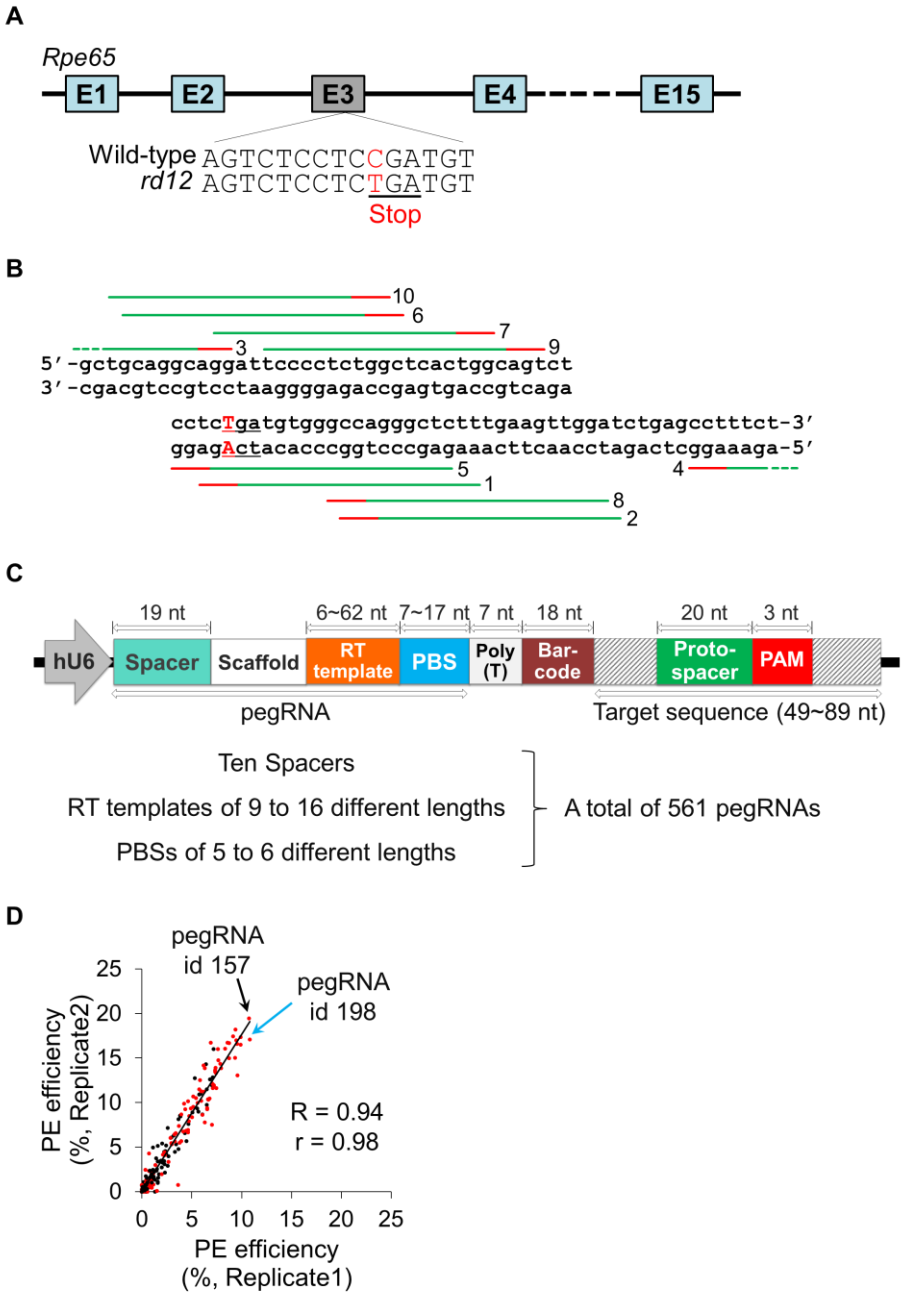

Figure 9. High-throughput evaluation of pegRNA candidates for the correction of the LCA-causing point mutation in *rd12* mice. (**A**) The *rd12* mouse model has a homozygous C-to-T nonsense mutation (red) in exon 3 of

the *Rpe65* gene, leading to a premature stop codon (underlined). (**B**) NG-PE2 target sequences. The protospacer and PAM of each target sequence are represented with green and red lines, respectively. The numbers next to the lines indicate target ids. (**C**) A schematic representation of the lentiviral library of pegRNA and target sequence pairs. Each pegRNA is paired with a wide target sequence that includes a protospacer, a PAM, and neighboring sequences. pegRNA expression is driven by the human U6 promoter (hU6). The library included a total of 561 pegRNA-target pairs, with the pegRNAs containing PBSs of 5 to 6 different lengths (7, 9, 11, 13, 15, or 17 nt) and RT templates of 9 to 16 different lengths (6 to 62 nt). Spacer, guide sequence of the pegRNA; RT, reverse transcriptase; PBSs, primer binding sites. (**D**) The correlation between prime editing efficiencies in replicates independently transfected with NG-PE2-encoding plasmids. The red and black dots indicate pegRNAs with corresponding target sequences with NGG and NGH PAMs, respectively. The blue arrow indicates the pegRNA selected for subsequent experiments. The Spearman (R) and Pearson (r) correlation coefficients are shown. A trend line is shown. The number of pegRNA-target sequence pairs n = 309.

8. AAV-Prime editor 2 precisely corrects the disease mutation in a mouse model of LCA

To investigate the therapeutic potential of prime editing in *rd12* mice, we used the adeno-associated virus (AAV) system with serotype2 to deliver PE2 and the selected pegRNA (pegRNA id 198) because AAV has been used to efficiently deliver other genome editing tools including engineered nucleases and base editors. Given that the coding sequence of PE2 is 6,273 bp, which is longer than the cargo size limit of AAV, we used trans-splicing AAV (tsAAV) vector[9,56,57]. We subretinally injected this system, named AAV-PE2, into

3-week-old *rd12* mice (**Figure 10A**). When the mice were analyzed six weeks after the injection, we observed that on average, 23% (range, 17% to 30%) of the whole retinal pigment epithelium (RPE) area in *rd12* mice was mCherry-positive (**Figure 10B-D**), which reflects the AAV-PE2 delivery efficiency[58].

6 weeks after the injection, the RPE tissue was isolated and genomic DNA was extracted. Deep sequencing of genomic DNA showed an average prime editing efficiency of 6.4% (range, 4.1% to 7.4%) (**Figure 11A**). Given that AAV-PE2 was delivered to only 23% of the RPE, the editing efficiency in regions that were exposed to prime editing components can be estimated to be, on average, 28%. Importantly, no unintended edits, substitutions, or indels near the mutation site were detectable (**Figure 11A and 11B**), suggesting highly precise genome editing in a disease model using a clinically applicable delivery approach. In our previous attempt using a CRISPR/Cas9- mediated HDR approach in the same mouse model, the correction efficiency was 1.2% ± 0.3% with an indel frequency of 17% ± 8%[51]. Lentiviral and AAV-mediated delivery of adenine base editors resulted in editing efficiencies of 16% ± 3% and 11% ± 7%, respectively, with substantial (up to 17% (lentiviral) and 7.7% ± 5% (AAV)) frequencies of bystander edits within the base editor editing windows[52,53]. We cannot rule out the possibility that such unintended edits could cause adverse effects, including unwanted immune responses against the protein variants encoded by genes affected by the unintended editing.

When we evaluated off-target effects at 20 potential off-target sites identified using Digenome-seq[27], nDigenome-seq[28], and Cas-OFFinder[29] (Supplementary Table 4), neither indels nor intended editing were observed (**Figure 11C and 11D**). This lack of off-target effects is in line with a failure to find off-target effects of prime editing in mouse embryos[17] and human organoids[15] even when PE3 was used (Off-target effects of PE3 would be, at a minimum, a combination of the off-target effects of PE2 and those of the

sgRNA). Taken together, our results show that subretinal injection of AAV-PE2 induced almost perfectly precise genome editing without bystander editing, generation of indels, or off-target effects at the genomic DNA level in the mouse RPE.

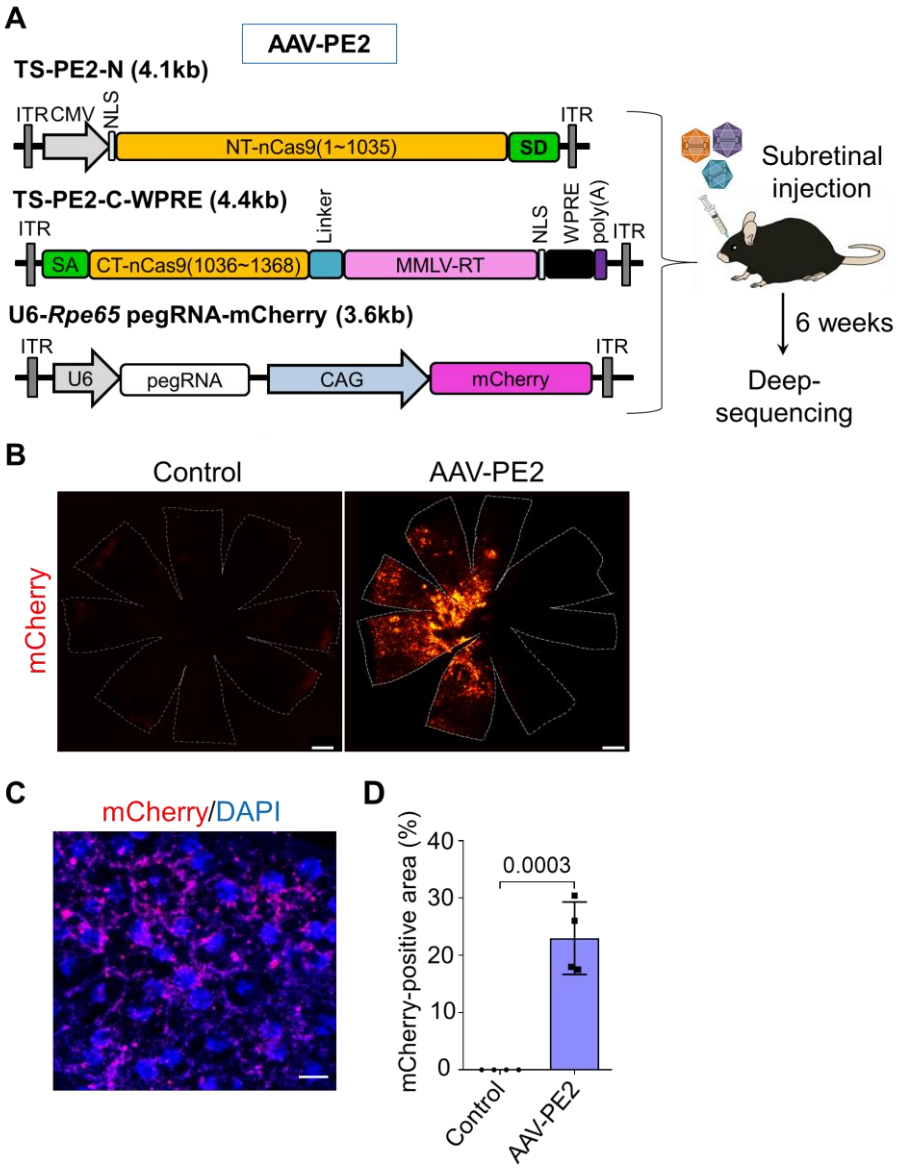

Figure 10. **Transduction efficiency of AAVs encoding PE2 and the pegRNA (AAV-PE2) using mCherry fluorescence as a surrogate marker in the RPE of *rd12* mice six weeks after subretinal injection.** (**A**) A schematic representation of rd12 mice experiments together with maps of the AAV vectors. The lengths of the sequences between the two ITRs in each vector are shown in parentheses. Two trans-splicing PE2-expressing AAVs (TS-PE2-N and TS-PE2-C-WPRE) were delivered together with an AAV expressing pegRNA into the RPE via subretinal injection. The RPE cells were harvested for deep sequencing at 6 weeks after the subretinal injection. SD, splicing donor; SA, splicing acceptor; NLS, nuclear localization sequence; ITR, inverted terminal repeat; WPRE, woodchuck hepatitis virus posttranscriptional regulatory element. (**B**-**C**) Representative low- (**B**) and high- (**C**) magnification photomicrographs of the RPE from *rd12* mice after subretinal injection of AAV-PE2. Scale bars, 500 μm (**B**), 10 μm (**C**). (**D**) Quantified mCherry-positive areas in the RPE from *rd12* mice after subretinal injection of AAV-PE2. Data are mean ± s.d. The number of mice $n = 4$. The p-value shown is based on a Student's t-test.

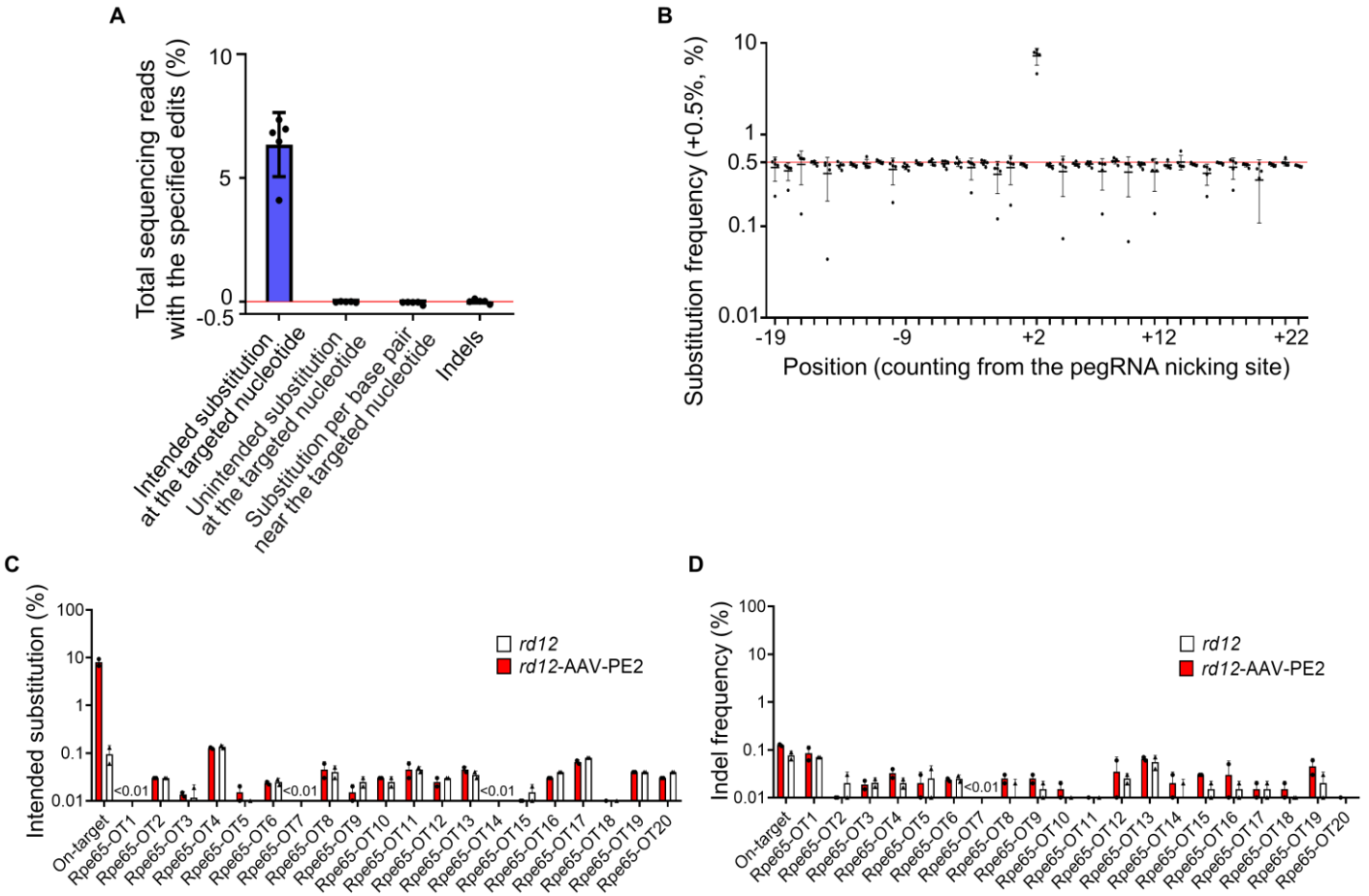

Figure 11. **Subretinal injection of AAV-PE2 efficiently corrects the disease-causing mutation without any detectable off-target effects in *rd12* mice.** (**A**) Frequencies of intended and unintended edits in the RPE of AAV-PE2-treated *rd12* mice. The frequencies were normalized by subtracting the average frequency of such editing in the control group without AAV-PE2 injection to exclude errors originating from PCR amplification and sequencing. Substitutions near the targeted nucleotide were evaluated over a 40-bp range centered on the target nucleotide. Indels were measured over a 60-bp range centered on the pegRNA nicking site. The red horizontal line represents the location where the normalized frequency = 0. Data are mean ± s.d. The number of mice n = 5. (**B**) Substitution frequencies at positions ranging from -20 bp to +20 bp from the target nucleotide, in the RPE of PE2-treated *rd12* mice. The frequencies were normalized by subtracting the average edit frequencies in the RPE of *rd12* mice without PE2 treatment to exclude errors originating from PCR amplification and sequencing. The red horizontal line represents the location where the normalized frequency = 0. Positions are numbered from the pegRNA nicking site. The targeted position is at +2. Data are mean ± s.d. The number of mice n = 5. (**C, D**) Frequencies of intended edits (**C**) and indels (**D**) at the predicted off-target sites for the pegRNA (id 198) in the RPE of PE2-treated *rd12* mice. Genomic DNA isolated from the RPE of *rd12* mice without PE2 treatment was used as the negative control (*rd12*). Data are mean ± s.d. The number of mice n = 2.

9. AAV-Prime editor 2 rescues the visual function in a mouse model of LCA

We next determined if this highly precise and specific genome editing would lead to functional correction of the disease. Confocal microscopic images of immunofluorescent staining showed a distinct membranous and cytoplasmic

expression of the RPE65 protein in RPE tissue from *rd12* mice treated with AAV-PE2, but not in tissue from untreated control *rd12* mice, at 6 weeks after the subretinal injection (**Figure 12A**). As previously reported[51,54], electroretinography (ERG) revealed a lack of dark-adapted light-induced electrical responses in control *rd12* mice. However, AAV-PE2 treatment rescued dark-adapted ERG responses (**Figure 12B**). Amplitudes of scotopic a- and b-waves from AAV-PE2-treated *rd12* mice were respectively on average 59% (range, 48% to 67%) and 27% (range, 24% to 30%) of those of wild-type C57BL/6 mice (**Figure 12C**). Compatible with these results, spatial thresholds to virtual rotating stimuli were significantly increased in AAV-PE2-treated *rd12* mice on optomotor response measurements (**Figure 12D**). Taken together, these results show that AAV-PE2-mediated prime editing in mouse RPE substantially improved visual function in *rd12* mice.

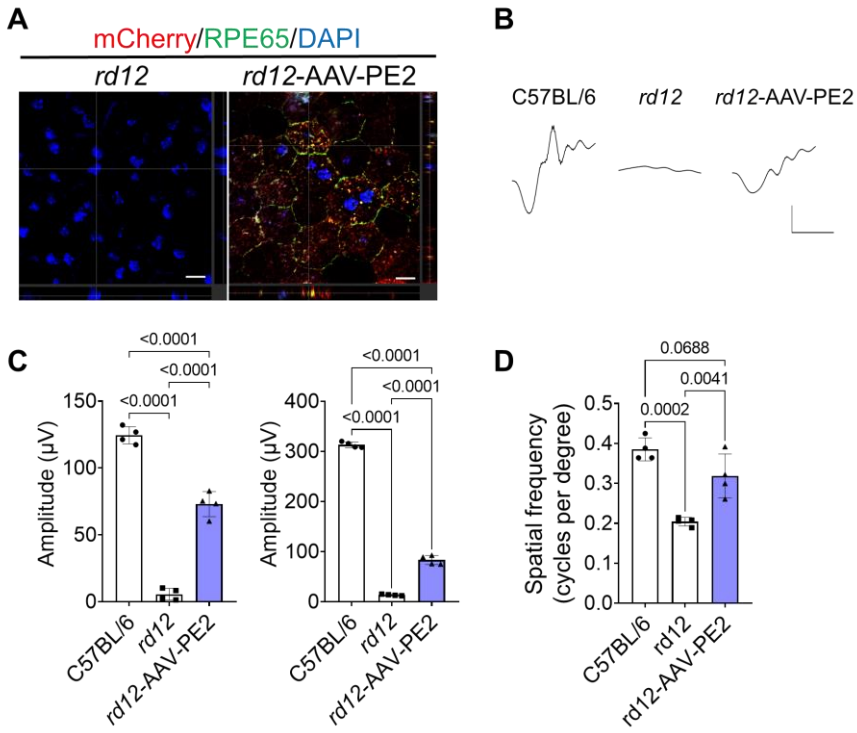

Figure 12. **Restoration of RPE65 expression and improvement of visual function in *rd12* mice after subretinal injection of AAV-PE2.** (**A**) Representative confocal photomicrographs showing RPE65 protein expression in RPE cells in *rd12* mice at 6 weeks after the subretinal injection of AAV-PE2. The uninjected negative control is shown on the left. Scale bars, 10 μm. (**B**) Representative waveforms of dark-adapted ERG responses at 0 dB in wild-type (C57BL/6), uninjected control (*rd12)*, and *rd12* mice injected with PE2-expressing AAV (*rd12*-AAV-PE2). Scale bars, 30 ms (x-axis) and 50 μV (y-axis). (**C**) Amplitudes of a-waves (left) and b-waves (right) of ERG reponses of C57BL/6 and *rd12* mice. Data are mean ± s.d. The number of mice n = 4. *P*-values from one-way ANOVA with post-hoc Tukey's multiple comparison tests are shown. *rd12*-AAV-PE2, *rd12* mice treated with AAV-PE2. (**D**) Optomotor response test results. Data are mean ± s.d. The number of mice n = 4. *P*-values from one-way ANOVA with post-hoc Tukey's multiple comparison tests are shown.

IV. DISCUSSION

Our results demonstrate that prime editing can generate intended edits in a highly precise manner in the liver and eye in mouse models of genetic liver and eye diseases when components are delivered using hydrodynamic injection and AAV-mediated methods, respectively. We also showed that replacing the SpCas9 domain of PE2 with a variant that recognizes a different set of PAMs, such as SpCas9-NG, can expand the list of target sequences for prime editors.

Cells containing disease-causing mutations are often not readily available, which makes it difficult to evaluate the efficiencies of genome editing tools including prime editors at the mutant target sequence. Furthermore, hundreds or thousands of pegRNAs can be designed to induce an intended edit at a target sequence. Although previously reported rules for designing optimal pegRNAs and computational models that predict pegRNA activities can help narrow down the list of optimal pegRNAs[14], the most thorough and definitive method for identifying the optimal pegRNA would be to evaluate all possible pegRNAs or all pegRNAs suggested by the rules and computational models.

For clinical applications of prime editing, the most efficient pegRNAs corresponding to the mutations found in patients would need to be identified. Determination of pegRNA activities using lentiviral libraries of paired pegRNA and target sequences as we describe above would be useful, especially when mutant sequence-containing cell lines are not available or when a large number of pegRNAs are evaluated.

The accuracy of high-throughput evaluations will be increased by using a large number of cells relative to the number of analyzed pegRNAs (high coverage), homogenous delivery of prime editing components, and high sequencing read depth per pegRNA. To identify the pegRNA that will work best for a specific application, we recommend the selection of several top

ranking pegRNAs from the high-throughput evaluations and the subsequent individual evaluation of their activities at endogenous loci.

When we compared PE2- and PE3-based correction of the disease-causing mutation in *Fah*$^{mut/mut}$ mice using hydrodynamic injections with those based on the hydrodynamic injections of Cas9 nuclease[33,35] or adenosine base editor[33], the efficiencies were overall comparable although exact comparisons are difficult, at least in part, due to differences in the time points used for analysis (Supplementary Table 6). However, the most notable difference is the precision of the editing. Cas9 nuclease induced a substantial frequency of indels (26%) at the target sites and a detectable, albeit lower than 0.3%, frequency of indels at off-target sites[33], whereas adenosine base editor induced considerable bystander effects (1.9%). However, PE3, but not PE2, induced only a low level of indels (on average 0.78%) and bystander or off-target effects were not observed for either PE2 or PE3 although we cannot rule out the possibility of off-target effects that could be identified using more advanced methods that might be developed in the future. The level of precision in genome editing we observed for PE2, in particular, has not been achieved using any of the previous methods of genome editing in this mouse model.

The initial PE2- and PE3-induced editing efficiencies, before NTBC withdrawal from *Fah*$^{mut/mut}$ mice, were under the detection limit. Although these low initial editing efficiencies functionally rescued the mice owing to the selective expansion of gene-corrected cells in this model of tyrosinemia, these efficiencies would not be sufficient to obtain satisfactory therapeutic effects in other liver diseases in which gene-corrected cells do not selectively expand. When we delivered PE2 and PE3 using the AAV vector into *Fah*$^{mut/mut}$ mice, the initial editing efficiencies before NTBC withdrawal, measured at the DNA level by deep sequencing, were also below the detection limit (data not shown). Thus, further optimization of the methods for prime editor delivery

could be required for the application of prime editing in human patients with liver diseases, especially those in which the gene-edited cells do not selectively expand.

In this study, we showed that highly precise and specific genome editing induced by the delivery of PE2 using a clinically applicable delivery approach (i.e., AAV) effectively rescued the visual function of *rd12* mice. It has very recently been reported that AAV-mediated delivery of split-intein prime editor 3 induced correction of a pathogenic gene at efficiencies of 0.6% (2 weeks after injection), 2.3% (6 weeks after injection), and 3.1% (10 weeks after injection) in the liver of an alpha-1 antitrypsin deficiency mouse model[59]. However, this approach induced unintended indels (about 22% of all edited alleles) and, more importantly, did not achieve functional correction. We estimate that the editing efficiency in our experiment was, on average, ~28% in the AAV-PE2-transduced region. In human patients, ophthalmologists can inject AAV-PE2 near the macula, to augment the editing efficiency in the cells at and near the macula, which is responsible for the major portion of our vision[46,51]. Given that subretinal injection of AAV has been successfully used in clinical trials[46-48], our study will lay a foundation for clinical translation of prime editing for LCA. Furthermore, given that AAV has been used as a delivery method for genetic diseases in organs other than eyes[60,61], we envision that AAV-mediated delivery of PE2 could be used to treat patients with genetic diseases that affect other organs.

This high precision of prime editing in somatic cells of mice raises the possibility that prime editing could be used for genome editing in human patients. We envision that *in vivo* prime editing, together with previously reported base editing and engineered nuclease-based approaches, will be a promising tactic for genome editing therapy for genetic diseases.

V. CONCLUSION

Prime editors can mediate genome editing without the need for donor DNA and without generating DSBs. However, the correction of both pathogenic mutations and phenotypes by prime editors in adult animal models has not been examined. Here, we identified an optimal prime editing guide RNA for correction of the disease-causing mutation by high-throughput evaluation of hundreds of pegRNAs. Further, we delivered prime editors 2 and 3 with the optimal pegRNA in a mouse model of hereditary tyrosinemia type I via hydrodynamic injection and highly precise and specific genome editing of the pathogenic mutation was achieved. Finally, we established a trans-splicing AAV system, a clinically applicable delivery approach, to deliver prime editors *in vivo*. Following a single injection of the AAV system encoding trans-splicing PE2, the disease-causing mutation was effectively corrected without detectable off-target editing and this led to the amelioration of the disease phenotypes in *rd12* mice. In summary, our results demonstrate the potential scope of prime editing for precise *in vivo* genome editing and should facilitate the clinical application of prime editing to treat a variety of genetic diseases.

REFERENCES

1. Kim H, Kim JS. A guide to genome engineering with programmable nucleases. Nat Rev Genet 2014;15:321-34.
2. Panier S, Boulton SJ. Double-strand break repair: 53BP1 comes into focus. Nat Rev Mol Cell Biol 2014;15:7-18.
3. Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, et al. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature 2016;540:144-9.
4. Pickar-Oliver A, Gersbach CA. The next generation of CRISPR-Cas technologies and applications. Nat Rev Mol Cell Biol 2019;20:490-507.
5. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 2016;533:420-4.
6. Kim YB, Komor AC, Levy JM, Packer MS, Zhao KT, Liu DR. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nat Biotechnol 2017;35:371-6.
7. Ma Y, Zhang J, Yin W, Zhang Z, Song Y, Chang X. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. Nat Methods 2016;13:1029-35.
8. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 2017;551:464-71.
9. Ryu SM, Koo T, Kim K, Lim K, Baek G, Kim ST, et al. Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. Nat Biotechnol 2018;36:536-9.

10. Song CQ, Jiang T, Richter M, Rhym LH, Koblan LW, Zafra MP, et al. Adenine base editing in an adult mouse model of tyrosinaemia. Nat Biomed Eng 2020;4:125-30.
11. Villiger L, Grisch-Chan HM, Lindsay H, Ringnalda F, Pogliano CB, Allegri G, et al. Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. Nat Med 2018;24:1519-25.
12. Rees HA, Liu DR. Base editing: precision chemistry on the genome and transcriptome of living cells. Nat Rev Genet 2018;19:770-88.
13. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 2019;576:149-57.
14. Kim HK, Yu G, Park J, Min S, Lee S, Yoon S, et al. Predicting the efficiency of prime editing guide RNAs in human cells. Nat Biotechnol 2021;39:198-206.
15. Schene IF, Joore IP, Oka R, Mokry M, van Vugt AHM, van Boxtel R, et al. Prime editing for functional repair in patient-derived disease models. Nat Commun 2020;11:5352.
16. Lin Q, Zong Y, Xue C, Wang S, Jin S, Zhu Z, et al. Prime genome editing in rice and wheat. Nat Biotechnol 2020;38:582-5.
17. Liu Y, Li X, He S, Huang S, Li C, Chen Y, et al. Efficient generation of mouse models with the prime editing system. Cell Discov 2020;6:27.
18. Song M, Kim HK, Lee S, Kim Y, Seo SY, Park J, et al. Sequence-specific prediction of the efficiencies of adenine and cytosine base editors. Nat Biotechnol 2020;38:1037-43.
19. Du D, Roguev A, Gordon DE, Chen M, Chen SH, Shales M, et al. Genetic interaction mapping in mammalian cells using CRISPR interference. Nat Methods 2017;14:577-80.

20. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 2014;343:84-7.
21. Kim HK, Song M, Lee J, Menon AV, Jung S, Kang YM, et al. In vivo high-throughput profiling of CRISPR-Cpf1 activity. Nat Methods 2017;14:153-9.
22. Park J, Lim K, Kim JS, Bae S. Cas-analyzer: an online tool for assessing genome editing results using NGS data. Bioinformatics 2017;33:286-8.
23. Ramakrishna S, Cho SW, Kim S, Song M, Gopalappa R, Kim JS, et al. Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations. Nat Commun 2014;5:3378.
24. Deverman BE, Pravdo PL, Simpson BP, Kumar SR, Chan KY, Banerjee A, et al. Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. Nat Biotechnol 2016;34:204-9.
25. Prusky GT, Alam NM, Beekman S, Douglas RM. Rapid quantification of adult and developing mouse spatial vision using a virtual optomotor system. Invest Ophthalmol Vis Sci 2004;45:4611-6.
26. Douglas RM, Alam NM, Silver BD, McGill TJ, Tschetter WW, Prusky GT. Independent visual threshold measurements in the two eyes of freely moving rats and mice using a virtual-reality optokinetic system. Vis Neurosci 2005;22:677-84.
27. Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nat Methods 2015;12:237-43.
28. Kim DY, Moon SB, Ko JH, Kim YS, Kim D. Unbiased investigation of specificities of prime editing systems in human cells. Nucleic Acids Res 2020;48:10576-89.

29. Bae S, Park J, Kim JS. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 2014;30:1473-5.
30. Haeussler M, Schonig K, Eckert H, Eschstruth A, Mianne J, Renaud JB, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol 2016;17:148.
31. Paulk NK, Wursthorn K, Wang Z, Finegold MJ, Kay MA, Grompe M. Adeno-associated virus gene repair corrects a mouse model of hereditary tyrosinemia in vivo. Hepatology 2010;51:1200-8.
32. Aponte JL, Sega GA, Hauser LJ, Dhar MS, Withrow CM, Carpenter DA, et al. Point mutations in the murine fumarylacetoacetate hydrolase gene: Animal models for the human genetic disorder hereditary tyrosinemia type 1. Proc Natl Acad Sci U S A 2001;98:641-5.
33. Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, et al. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat Biotechnol 2014;32:551-3.
34. Yin H, Song CQ, Dorkin JR, Zhu LJ, Li Y, Wu Q, et al. Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. Nat Biotechnol 2016;34:328-33.
35. Shin JH, Jung S, Ramakrishna S, Kim HH, Lee J. In vivo gene correction with targeted sequence substitution through microhomology-mediated end joining. Biochem Biophys Res Commun 2018;502:116-22.
36. Nishimasu H, Shi X, Ishiguro S, Gao L, Hirano S, Okazaki S, et al. Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science 2018;361:1259-62.

37. Kim N, Kim HK, Lee S, Seo JH, Choi JW, Park J, et al. Prediction of the sequence-specific cleavage activity of Cas9 variants. Nat Biotechnol 2020;38:1328-36.
38. Kim HK, Lee S, Kim Y, Park J, Min S, Choi JW, et al. High-throughput analysis of the activities of xCas9, SpCas9-NG and SpCas9 at matched and mismatched target sequences in human cells. Nat Biomed Eng 2020;4:111-24.
39. Kim HK, Kim Y, Lee S, Min S, Bae JY, Choi JW, et al. SpCas9 activity prediction by DeepSpCas9, a deep learning-based model with high generalization performance. Sci Adv 2019;5:eaax9249.
40. Wilkinson PD, Delgado ER, Alencastro F, Leek MP, Roy N, Weirich MP, et al. The Polyploid State Restricts Hepatocyte Proliferation and Liver Regeneration in Mice. Hepatology 2019;69:1242-58.
41. Gao P, Lyu Q, Ghanam AR, Lazzarotto CR, Newby GA, Zhang W, et al. Prime editing in mice reveals the essentiality of a single base in driving tissue-specific gene expression. Genome Biol 2021;22:83.
42. Petri K, Zhang W, Ma J, Schmidts A, Lee H, Horng JE, et al. CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. Nat Biotechnol 2021;1-5.
43. Cideciyan AV. Leber congenital amaurosis due to RPE65 mutations and its treatment with gene therapy. Prog Retin Eye Res 2010;29:398-427.
44. Sahel JA. Spotlight on childhood blindness. J Clin Invest 2011;121:2145-9.
45. Redmond TM, Yu S, Lee E, Bok D, Hamasaki D, Chen N, et al. Rpe65 is necessary for production of 11-cis-vitamin A in the retinal visual cycle. Nat Genet 1998;20:344-51.
46. Russell S, Bennett J, Wellman JA, Chung DC, Yu ZF, Tillman A, et al. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in

patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. Lancet 2017;390:849-60.

47. Bainbridge JWB, Mehat MS, Sundaram V, Robbie SJ, Barker SE, Ripamonti C, et al. Long-Term Effect of Gene Therapy on Leber's Congenital Amaurosis. 2015;372:1887-97.

48. Jacobson SG, Cideciyan AV, Roman AJ, Sumaroka A, Schwartz SB, Heon E, et al. Improvement and decline in vision with gene therapy in childhood blindness. N Engl J Med 2015;372:1920-6.

49. Maeder ML, Stefanidakis M, Wilson CJ, Baral R, Barrera LA, Bounoutas GS, et al. Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. Nat Med 2019;25:229-33.

50. Cideciyan AV, Jacobson SG, Drack AV, Ho AC, Charng J, Garafalo AV, et al. Effect of an intravitreal antisense oligonucleotide on vision in Leber congenital amaurosis due to a photoreceptor cilium defect. Nat Med 2019;25:225-8.

51. Jo DH, Song DW, Cho CS, Kim UG, Lee KJ, Lee K, et al. CRISPR-Cas9-mediated therapeutic editing of Rpe65 ameliorates the disease phenotypes in a mouse model of Leber congenital amaurosis. Sci Adv 2019;5:eaax1210.

52. Suh S, Choi EH, Leinonen H, Foik AT, Newby GA, Yeh WH, et al. Restoration of visual function in adult mice with an inherited retinal disease via adenine base editing. Nat Biomed Eng 2021;5:169-78.

53. Jo DH, Jang H-K, Cho CS, Han JH, Ryu G, Jung Y, et al. Therapeutic adenine base editing corrects nonsense mutation and improves visual function in a mouse model of Leber congenital amaurosis. bioRxiv. in press 2021.

54. Pang JJ, Chang B, Hawes NL, Hurd RE, Davisson MT, Li J, et al. Retinal degeneration 12 (rd12): a new, spontaneously arising mouse model for human Leber congenital amaurosis (LCA). Mol Vis 2005;11:152-62.
55. Kweon J, Yoon JK, Jang AH, Shin HR, See JE, Jang G, et al. Engineered Prime Editors with PAM flexibility. Mol Ther 2021;29.6:2001-7
56. Sun L, Li J, Xiao X. Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization. Nat Med 2000;6:599-602.
57. Lai Y, Yue Y, Liu M, Ghosh A, Engelhardt JF, Chamberlain JS, et al. Efficient in vivo gene expression by trans-splicing adeno-associated viral vectors. Nat Biotechnol 2005;23:1435-9.
58. Levy JM, Yeh WH, Pendse N, Davis JR, Hennessey E, Butcher R, et al. Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. Nat Biomed Eng 2020;4:97-110.
59. Liu P, Liang SQ, Zheng C, Mintzer E, Zhao YG, Ponnienselvan K, et al. Improved prime editors enable pathogenic allele correction and cancer modelling in adult mice. Nat Commun 2021;12:2121.
60. Wang D, Tai PWL, Gao G. Adeno-associated virus vector as a platform for gene therapy delivery. Nat Rev Drug Discov 2019;18:358-78.
61. Li C, Samulski RJ. Engineering adeno-associated virus vectors for gene therapy. Nat Rev Genet 2020;21:255-72.

Table 1. Predicted activities of SpCas9 and SpCas9-NG at nine target sequences for the prime editing of the tyrosinemia-causing mutation

Target Sequence ID	Target Sequence (30bp)	Guide Sequence (20bp)	DeepSpCas9-NG Score	DeepSpCas9-NG Score Rank	PAM	Position and type of editing
1	CATTTTCGGGATGGTCCT CATGAACGACTG	TTCGGGATGGTCCTCATGAA	31.917	6	NGH	+14 A to G
2	TTCGGGATGGTCCTCATG AACGACTGGAGC	GGATGGTCCTCATGAACGAC	38.165	2	NGG	+10 A to G
3	TCGGGATGGTCCTCATG AACGACTGGAGCA	GATGGTCCTCATGAACGACT	42.666	1	NGH	+9 A to G
4	GGGATGGTCCTCATGAA CGACTGGAGCAGT	TGGTCCTCATGAACGACTGG	37.155	3	NGH	+7A to G
5	ATGGTCCTCATGAACGA CTGGAGCAGTAAT	TCCTCATGAACGACTGGAGC	16.969	8	NGH	+4 A to G
6	GGAAGCTGGGCCACCAG GCATTACTGCTCC	GCTGGGCCACCAGGCATTAC	8.695	9	NGH	+4 T to C
7	ACATCAGAGGAAGCTGG GCCACCAGGCATT	CAGAGGAAGCTGGGCCACCA	32.242	5	NGH	+12 T to C
8	AACATCAGAGGAAGCTG GGCCACCAGGCAT	TCAGAGGAAGCTGGGCCACC	21.849	7	NGG	+13 T to C
9	AGAACAGAACATCAGAG GAAGCTGGGCCAC	CAGAACATCAGAGGAAGCTG	36.729	4	NGH	+20 T to C

Table 2. Measured prime editing efficiencies for the tested *Fah* pegRNAs using a paired library approach

Rank	pegRNA ID	Target id	PE efficiency (%, Replicate 1)	PE efficiency (%, Replicate 2)	Average PE efficiency (%)	Spacer sequence	RT template and PBS sequence	PAM	Deep SpCas9-NG score	Deep SpCas9 score
1	135	5	6.05	11.90	8.97	CCTCATGAACGAC TGGAGC	GGCATTACCGCTCCAGT CG	NGT	17.0	No score
2	136	5	4.29	11.60	7.95	CCTCATGAACGAC TGGAGC	GGCATTACCGCTCCAGT CGTT	NGT	17.0	No score
3	137	5	4.94	8.96	6.95	CCTCATGAACGAC TGGAGC	GGCATTACCGCTCCAGT CGTTCA	NGT	17.0	No score
4	88	2	5.38	7.74	6.56	GATGGTCCTCATG AACGAC	GGCATTACCGCTCCAGT CGTTCATGAGGACC	NGG	38.2	52.4
5	89	2	4.56	8.48	6.52	GATGGTCCTCATG AACGAC	GGCATTACCGCTCCAGT CGTTCATGAGGACCAT	NGG	38.2	52.4
6	138	5	3.84	8.85	6.35	CCTCATGAACGAC TGGAGC	GGCATTACCGCTCCAGT CGTTCATG	NGT	17.0	No score
7	38	4	5.17	7.24	6.21	GGTCCTCATGAAC GACTGG	ACCAGGCATTACCGCTC CAGTCGTTCATGAGG	NGC	37.2	No score
8	90	2	4.75	7.21	5.98	GATGGTCCTCATG AACGAC	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGAC C	NGG	38.2	52.4
9	91	2	4.62	7.27	5.95	GATGGTCCTCATG AACGAC	TTACCGCTCCAGTCGTT CATGAGGACCAT	NGG	38.2	52.4
10	139	5	3.88	7.80	5.84	CCTCATGAACGAC TGGAGC	ACCAGGCATTACCGCTC CAGTCG	NGT	17.0	No score
11	39	4	3.88	7.02	5.45	GGTCCTCATGAAC GACTGG	GGCATTACCGCTCCAGT CGTTCATGAGGAC	NGC	37.2	No score
12	140	5	3.93	6.89	5.41	CCTCATGAACGAC TGGAGC	ACCAGGCATTACCGCTC CAGTCGTTCA	NGT	17.0	No score
13	92	2	3.01	7.70	5.35	GATGGTCCTCATG AACGAC	TTACCGCTCCAGTCGTT CATGAGGACC	NGG	38.2	52.4
14	141	5	4.19	6.40	5.29	CCTCATGAACGAC TGGAGC	GGCCACCAGGCATTACC GCTCCAGTCG	NGT	17.0	No score
15	142	5	3.72	6.39	5.06	CCTCATGAACGAC TGGAGC	ACCAGGCATTACCGCTC CAGTCGTTCATG	NGT	17.0	No score

16	143	5	3.31	6.59	4.95	CCTCATGAACGAC TGGAGC	GGCCACCAGGCATTACC GCTCCAGTCGTTCATG	NGT	17.0	No score
17	40	4	3.54	6.27	4.91	GGTCCTCATGAAC GACTGG	GGCATTACCGCTCCAGT CGTTCATGAGG	NGC	37.2	No score
18	230	3	3.68	5.94	4.81	ATGGTCCTCATGA ACGACT	GGCATTACCGCTCCAGT CGTTCATGAGGAC	NGA	42.7	No score
19	93	2	3.81	5.58	4.70	GATGGTCCTCATG AACGAC	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGAC CAT	NGG	38.2	52.4
20	144	5	2.69	6.49	4.59	CCTCATGAACGAC TGGAGC	ACCAGGCATTACCGCTC CAGTCGTT	NGT	17.0	No score
21	145	5	2.76	6.35	4.55	CCTCATGAACGAC TGGAGC	TTACCGCTCCAGTCG	NGT	17.0	No score
22	146	5	3.36	5.63	4.49	CCTCATGAACGAC TGGAGC	CATTACCGCTCCAGTCG	NGT	17.0	No score
23	276	7	4.23	4.45	4.34	AGAGGAAGCTGG GCCACCA	GAGCGGTAATGCCTGGT GGCCCAGC	NGC	32.2	No score
24	147	5	2.96	5.65	4.31	CCTCATGAACGAC TGGAGC	GGCATTACCGCTCCAGT CGTTCATGAG	NGT	17.0	No score
25	41	4	2.58	5.87	4.23	GGTCCTCATGAAC GACTGG	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGAC	NGC	37.2	No score
26	148	5	2.80	5.64	4.22	CCTCATGAACGAC TGGAGC	CAGGCATTACCGCTCCA GTCG	NGT	17.0	No score
27	231	3	3.21	5.18	4.19	ATGGTCCTCATGA ACGACT	GGCATTACCGCTCCAGT CGTTCATGAGGACCA	NGA	42.7	No score
28	149	5	2.84	5.40	4.12	CCTCATGAACGAC TGGAGC	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA G	NGT	17.0	No score
29	42	4	2.63	5.42	4.03	GGTCCTCATGAAC GACTGG	TTACCGCTCCAGTCGTT CATGAGG	NGC	37.2	No score
30	94	2	2.43	5.60	4.02	GATGGTCCTCATG AACGAC	GGCATTACCGCTCCAGT CGTTCATGAGGA	NGG	38.2	52.4
31	150	5	3.17	4.83	4.00	CCTCATGAACGAC TGGAGC	GGCCACCAGGCATTACC GCTCCAGTCGTTCA	NGT	17.0	No score
32	232	3	2.99	4.92	3.96	ATGGTCCTCATGA ACGACT	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGAC CA	NGA	42.7	No score
33	95	2	2.28	5.52	3.90	GATGGTCCTCATG AACGAC	CAGGCATTACCGCTCCA GTCGTTCATGAGGACC	NGG	38.2	52.4

34	151	5	2.37	5.40	3.88	CCTCATGAACGAC TGGAGC	ACCAGGCATTACCGCTC CAGTCGTTCATGAG	NGT	17.0	No score
35	277	7	2.26	5.47	3.86	AGAGGAAGCTGG GCCACCA	GAGCGGTAATGCCTGGT GGCCCAGCTTCC	NGC	32.2	No score
36	152	5	3.17	4.55	3.86	CCTCATGAACGAC TGGAGC	GGCCACCAGGCATTACC GCTCCAGTCGTT	NGT	17.0	No score
37	278	7	3.01	4.66	3.84	AGAGGAAGCTGG GCCACCA	GAGCGGTAATGCCTGGT GGCCCAGCTT	NGC	32.2	No score
38	153	5	3.40	4.07	3.74	CCTCATGAACGAC TGGAGC	TTACCGCTCCAGTCGTT CATG	NGT	17.0	No score
39	96	2	2.71	4.73	3.72	GATGGTCCTCATG AACGAC	CATTACCGCTCCAGTCG TTCATGAGGACC	NGG	38.2	52.4
40	43	4	2.26	5.16	3.71	GGTCCTCATGAAC GACTGG	CATTACCGCTCCAGTCG TTCATGAGG	NGC	37.2	No score
41	44	4	2.29	5.10	3.69	GGTCCTCATGAAC GACTGG	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA GGAC	NGC	37.2	No score
42	45	4	2.44	4.48	3.46	GGTCCTCATGAAC GACTGG	TTACCGCTCCAGTCGTT CATGAGGAC	NGC	37.2	No score
43	233	3	2.39	4.29	3.34	ATGGTCCTCATGA ACGACT	TTACCGCTCCAGTCGTT CATGAGGAC	NGA	42.7	No score
44	234	3	2.33	4.27	3.30	ATGGTCCTCATGA ACGACT	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGAC	NGA	42.7	No score
45	154	5	2.02	4.57	3.29	CCTCATGAACGAC TGGAGC	TTACCGCTCCAGTCGTT	NGT	17.0	No score
46	46	4	2.44	4.12	3.28	GGTCCTCATGAAC GACTGG	CAGGCATTACCGCTCCA GTCGTTCATGAGG	NGC	37.2	No score
47	97	2	3.20	3.30	3.25	GATGGTCCTCATG AACGAC	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA GGACC	NGG	38.2	52.4
48	47	4	2.30	4.01	3.16	GGTCCTCATGAAC GACTGG	TTACCGCTCCAGTCGTT CATGA	NGC	37.2	No score
49	98	2	1.77	4.54	3.16	GATGGTCCTCATG AACGAC	CAGGCATTACCGCTCCA GTCGTTCATGAGGACCA T	NGG	38.2	52.4
50	279	7	1.77	4.23	3.00	AGAGGAAGCTGG GCCACCA	TGGAGCGGTAATGCCTG GTGGCCCAGCTTCC	NGC	32.2	No score
51	155	5	1.87	4.12	3.00	CCTCATGAACGAC TGGAGC	CATTACCGCTCCAGTCG TT	NGT	17.0	No score

52	280	7	2.24	3.68	2.96	AGAGGAAGCTGG GCCACCA	TGGAGCGGTAATGCCTG GTGGCCCAGC	NGC	32.2	No score
53	99	2	2.65	3.24	2.94	GATGGTCCTCATG AACGAC	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA GGACCAT	NGG	38.2	52.4
54	156	5	1.82	3.99	2.91	CCTCATGAACGAC TGGAGC	CAGGCATTACCGCTCCA GTCGTTCATG	NGT	17.0	No score
55	157	5	1.74	4.07	2.91	CCTCATGAACGAC TGGAGC	TTACCGCTCCAGTCGTT CA	NGT	17.0	No score
56	100	2	2.32	3.47	2.89	GATGGTCCTCATG AACGAC	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG ACC	NGG	38.2	52.4
57	101	2	2.00	3.77	2.88	GATGGTCCTCATG AACGAC	TTACCGCTCCAGTCGTT CATGAGGA	NGG	38.2	52.4
58	281	7	1.80	3.88	2.84	AGAGGAAGCTGG GCCACCA	TGGAGCGGTAATGCCTG GTGGCCCAGCTT	NGC	32.2	No score
59	158	5	1.94	3.69	2.81	CCTCATGAACGAC TGGAGC	CAGGCATTACCGCTCCA GTCGTT	NGT	17.0	No score
60	48	4	1.98	3.62	2.80	GGTCCTCATGAAC GACTGG	GGCATTACCGCTCCAGT CGTTCAT	NGC	37.2	No score
61	102	2	1.86	3.74	2.80	GATGGTCCTCATG AACGAC	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGA	NGG	38.2	52.4
62	159	5	1.38	4.08	2.73	CCTCATGAACGAC TGGAGC	CATTACCGCTCCAGTCG TTCATG	NGT	17.0	No score
63	49	4	1.90	3.38	2.64	GGTCCTCATGAAC GACTGG	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG	NGC	37.2	No score
64	282	7	1.91	3.35	2.63	AGAGGAAGCTGG GCCACCA	AACGACTGGAGCGGTAA TGCCTGGTGGCCCAGC	NGC	32.2	No score
65	50	4	2.23	2.99	2.61	GGTCCTCATGAAC GACTGG	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA GG	NGC	37.2	No score
66	160	5	1.95	3.24	2.60	CCTCATGAACGAC TGGAGC	TTACCGCTCCAGTCGTT CATGAG	NGT	17.0	No score
67	51	4	1.98	3.19	2.58	GGTCCTCATGAAC GACTGG	CAGGCATTACCGCTCCA GTCGTTCATGAGGAC	NGC	37.2	No score
68	161	5	0.94	4.21	2.57	CCTCATGAACGAC TGGAGC	GCTGGGCCACCAGGCAT TACCGCTCCAGTCG	NGT	17.0	No score
69	162	5	1.03	4.08	2.56	CCTCATGAACGAC TGGAGC	TGGGCCACCAGGCATTA CCGCTCCAGTCG	NGT	17.0	No score

70	235	3	1.94	3.14	2.54	ATGGTCCTCATGAACGACT	TTACCGCTCCAGTCGTTCATGAGGACCA	NGA	42.7	No score
71	236	3	1.98	3.04	2.51	ATGGTCCTCATGAACGACT	GGCCACCAGGCATTACCGCTCCAGTCGTTCATGAGGAC	NGA	42.7	No score
72	237	3	2.46	2.52	2.49	ATGGTCCTCATGAACGACT	GGCCACCAGGCATTACCGCTCCAGTCGTTCATGAGGACCA	NGA	42.7	No score
73	163	5	1.88	2.92	2.40	CCTCATGAACGACTGGAGC	CCACCAGGCATTACCGCTCCAGTCGTTCATG	NGT	17.0	No score
74	103	2	0.97	3.70	2.33	GATGGTCCTCATGAACGAC	CATTACCGCTCCAGTCGTTCATGAGGACCAT	NGG	38.2	52.4
75	283	7	1.00	3.46	2.23	AGAGGAAGCTGGGCCACCA	GAGCGGTAATGCCTGGTGGCCCAGCTTCCTC	NGC	32.2	No score
76	284	7	0.80	3.65	2.23	AGAGGAAGCTGGGCCACCA	TGAACGACTGGAGCGGTAATGCCTGGTGGCCCAGCTTCC	NGC	32.2	No score
77	52	4	1.63	2.79	2.21	GGTCCTCATGAACGACTGG	TTACCGCTCCAGTCGTTC	NGC	37.2	No score
78	238	3	1.06	3.29	2.17	ATGGTCCTCATGAACGACT	CATTACCGCTCCAGTCGTTCATGAGGAC	NGA	42.7	No score
79	164	5	1.30	3.02	2.16	CCTCATGAACGACTGGAGC	CATTACCGCTCCAGTCGTTCA	NGT	17.0	No score
80	239	3	1.60	2.72	2.16	ATGGTCCTCATGAACGACT	CATTACCGCTCCAGTCGTTCATGAGGACCA	NGA	42.7	No score
81	165	5	1.29	3.00	2.15	CCTCATGAACGACTGGAGC	CAGGCATTACCGCTCCAGTCGTTCA	NGT	17.0	No score
82	177	8	1.54	2.69	2.11	CAGAGGAAGCTGGGCCACC	GAGCGGTAATGCCTGGTGGCCCAGCT	NGG	21.8	51.6
83	53	4	1.38	2.84	2.11	GGTCCTCATGAACGACTGG	CATTACCGCTCCAGTCGTTCATGAGGAC	NGC	37.2	No score
84	240	3	1.53	2.66	2.09	ATGGTCCTCATGAACGACT	CAGGCATTACCGCTCCAGTCGTTCATGAGGAC	NGA	42.7	No score
85	54	4	1.67	2.51	2.09	GGTCCTCATGAACGACTGG	CCACCAGGCATTACCGCTCCAGTCGTTCATGAGGAC	NGC	37.2	No score
86	55	4	1.30	2.86	2.08	GGTCCTCATGAACGACTGG	TTACCGCTCCAGTCGTTCAT	NGC	37.2	No score
87	1	1	1.82	2.33	2.07	TCGGGATGGTCCTCATGAA	GGCATTACCGCTCCAGTCGTTCATGAGGACC	NGA	31.9	No score

88	285	7	1.37	2.68	2.03	AGAGGAAGCTGG GCCACCA	TGAACGACTGGAGCGGT AATGCCTGGTGGCCCAG C	NGC	32.2	No score
89	56	4	1.05	2.92	1.98	GGTCCTCATGAAC GACTGG	GGCATTACCGCTCCAGT CGTTCATGA	NGC	37.2	No score
90	286	7	0.96	2.97	1.97	AGAGGAAGCTGG GCCACCA	AACGACTGGAGCGGTAA TGCCTGGTGGCCCAGCT TCC	NGC	32.2	No score
91	57	4	1.72	2.22	1.97	GGTCCTCATGAAC GACTGG	GGCATTACCGCTCCAGT CGTTC	NGC	37.2	No score
92	2	1	1.68	2.24	1.96	TCGGGATGGTCCT CATGAA	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA GGACCAT	NGA	31.9	No score
93	287	7	0.15	3.67	1.91	AGAGGAAGCTGG GCCACCA	TCCTCATGAACGACTGG AGCGGTAATGCCTGGTG GCCCAGCTTCC	NGC	32.2	No score
94	104	2	1.57	2.22	1.90	GATGGTCCTCATG AACGAC	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG ACCAT	NGG	38.2	52.4
95	288	7	1.32	2.47	1.89	AGAGGAAGCTGG GCCACCA	CTCATGAACGACTGGAG CGGTAATGCCTGGTGGC CCAGC	NGC	32.2	No score
96	105	2	1.57	2.19	1.88	GATGGTCCTCATG AACGAC	CAGGCATTACCGCTCCA GTCGTTCATGAGGA	NGG	38.2	52.4
97	166	5	1.14	2.60	1.87	CCTCATGAACGAC TGGAGC	CCACCAGGCATTACCGC TCCAGTCGTTCA	NGT	17.0	No score
98	241	3	1.15	2.59	1.87	ATGGTCCTCATGA ACGACT	TTACCGCTCCAGTCGTT CATGAGG	NGA	42.7	No score
99	167	5	1.10	2.63	1.86	CCTCATGAACGAC TGGAGC	CCACCAGGCATTACCGC TCCAGTCG	NGT	17.0	No score
100	106	2	1.10	2.57	1.83	GATGGTCCTCATG AACGAC	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA GGA	NGG	38.2	52.4
101	289	7	1.27	2.39	1.83	AGAGGAAGCTGG GCCACCA	AACGACTGGAGCGGTAA TGCCTGGTGGCCCAGCT T	NGC	32.2	No score
102	168	5	1.28	2.37	1.82	CCTCATGAACGAC TGGAGC	CAGGCATTACCGCTCCA GTCGTTCATGAG	NGT	17.0	No score
103	242	3	1.51	2.11	1.81	ATGGTCCTCATGA ACGACT	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG AC	NGA	42.7	No score

104	3	1	1.31	2.28	1.79	TCGGGATGGTCCT CATGAA	GGCATTACCGCTCCAGT CGTTCATGAGGACCAT	NGA	31.9	No score
105	243	3	1.20	2.38	1.79	ATGGTCCTCATGA ACGACT	CAGGCATTACCGCTCCA GTCGTTCATGAGGACCA	NGA	42.7	No score
106	290	7	0.91	2.66	1.78	AGAGGAAGCTGG GCCACCA	ACTGGAGCGGTAATGCC TGGTGGCCCAGC	NGC	32.2	No score
107	4	1	1.58	1.96	1.77	TCGGGATGGTCCT CATGAA	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA GGACC	NGA	31.9	No score
108	58	4	1.72	1.78	1.75	GGTCCTCATGAAC GACTGG	ACCAGGCATTACCGCTC CAGTCGTTCAT	NGC	37.2	No score
109	5	1	1.32	2.12	1.72	TCGGGATGGTCCT CATGAA	TTACCGCTCCAGTCGTT CATGAGGACC	NGA	31.9	No score
110	59	4	1.21	2.17	1.69	GGTCCTCATGAAC GACTGG	ACCAGGCATTACCGCTC CAGTCGTTCATGA	NGC	37.2	No score
111	60	4	0.26	3.05	1.66	GGTCCTCATGAAC GACTGG	GCTGGGCCACCAGGCAT TACCGCTCCAGTCGTTC ATGA	NGC	37.2	No score
112	244	3	1.29	2.01	1.65	ATGGTCCTCATGA ACGACT	GGCATTACCGCTCCAGT CGTTCATGAGG	NGA	42.7	No score
113	61	4	0.76	2.54	1.65	GGTCCTCATGAAC GACTGG	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA	NGC	37.2	No score
114	291	7	0.96	2.18	1.57	AGAGGAAGCTGG GCCACCA	TGGAGCGGTAATGCCTG GTGGCCCAGCTTCCTC	NGC	32.2	No score
115	245	3	1.07	1.94	1.50	ATGGTCCTCATGA ACGACT	ACCAGGCATTACCGCTC CAGTCGTTCATGAGG	NGA	42.7	No score
116	6	1	1.29	1.71	1.50	TCGGGATGGTCCT CATGAA	GGCATTACCGCTCCAGT CGTTCATGAGGACCATC C	NGA	31.9	No score
117	292	7	0.47	2.43	1.45	AGAGGAAGCTGG GCCACCA	ACTGGAGCGGTAATGCC TGGTGGCCCAGCTTCC	NGC	32.2	No score
118	169	5	0.91	1.94	1.43	CCTCATGAACGAC TGGAGC	CCACCAGGCATTACCGC TCCAGTCGTTCATGAG	NGT	17.0	No score
119	62	4	0.67	2.11	1.39	GGTCCTCATGAAC GACTGG	CATTACCGCTCCAGTCG TTC	NGC	37.2	No score
120	107	2	1.11	1.63	1.37	GATGGTCCTCATG AACGAC	ACCGCTCCAGTCGTTCA TGAGGACC	NGG	38.2	52.4
121	7	1	1.03	1.71	1.37	TCGGGATGGTCCT CATGAA	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGAC CAT	NGA	31.9	No score

122	170	5	0.95	1.77	1.36	CCTCATGAACGAC TGGAGC	CCACCAGGCATTACCGC TCCAGTCGTT	NGT	17.0	No score
123	63	4	0.75	1.96	1.36	GGTCCTCATGAAC GACTGG	CATTACCGCTCCAGTCG TTCATGA	NGC	37.2	No score
124	108	2	0.74	1.96	1.35	GATGGTCCTCATG AACGAC	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG A	NGG	38.2	52.4
125	293	7	1.10	1.53	1.32	AGAGGAAGCTGG GCCACCA	CATGAACGACTGGAGCG GTAATGCCTGGTGGCCC AGC	NGC	32.2	No score
126	109	2	0.78	1.81	1.30	GATGGTCCTCATG AACGAC	CATTACCGCTCCAGTCG TTCATGAGGA	NGG	38.2	52.4
127	294	7	0.88	1.68	1.28	AGAGGAAGCTGG GCCACCA	CATGAACGACTGGAGCG GTAATGCCTGGTGGCCC AGCTT	NGC	32.2	No score
128	295	7	0.45	2.08	1.27	AGAGGAAGCTGG GCCACCA	ACTGGAGCGGTAATGCC TGGTGGCCCAGCTT	NGC	32.2	No score
129	8	1	0.75	1.77	1.26	TCGGGATGGTCCT CATGAA	TTACCGCTCCAGTCGTT CATGAGGACCAT	NGA	31.9	No score
130	9	1	2.02	0.47	1.25	TCGGGATGGTCCT CATGAA	CATTACCGCTCCAGTCG TTCATGAGGACCATCC	NGA	31.9	No score
131	178	8	0.95	1.54	1.25	CAGAGGAAGCTG GGCCACC	GAGCGGTAATGCCTGGT GGCCCAGCTTC	NGG	21.8	51.6
132	296	7	0.93	1.49	1.21	AGAGGAAGCTGG GCCACCA	TGAACGACTGGAGCGGT AATGCCTGGTGGCCCAG CTT	NGC	32.2	No score
133	64	4	0.84	1.57	1.21	GGTCCTCATGAAC GACTGG	CAGGCATTACCGCTCCA GTCGTTC	NGC	37.2	No score
134	297	7	0.42	1.97	1.19	AGAGGAAGCTGG GCCACCA	CATGAACGACTGGAGCG GTAATGCCTGGTGGCCC AGCTTCC	NGC	32.2	No score
135	179	8	0.64	1.68	1.16	CAGAGGAAGCTG GGCCACC	AACGACTGGAGCGGTAA TGCCTGGTGGCCCAGCT	NGG	21.8	51.6
136	180	8	1.12	1.20	1.16	CAGAGGAAGCTG GGCCACC	TGGAGCGGTAATGCCTG GTGGCCCAGCT	NGG	21.8	51.6
137	181	8	0.87	1.44	1.15	CAGAGGAAGCTG GGCCACC	GAGCGGTAATGCCTGGT GGCCCAGCTTCCT	NGG	21.8	51.6
138	182	8	0.73	1.55	1.14	CAGAGGAAGCTG GGCCACC	TGGAGCGGTAATGCCTG GTGGCCCAGCTTCCT	NGG	21.8	51.6

139	10	1	0.93	1.34	1.14	TCGGGATGGTCCT CATGAA	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG ACCAT	NGA	31.9	No score
140	183	8	0.84	1.40	1.12	CAGAGGAAGCTG GGCCACC	TGGAGCGGTAATGCCTG GTGGCCCAGCTTC	NGG	21.8	51.6
141	298	7	0.86	1.38	1.12	AGAGGAAGCTGG GCCACCA	CGACTGGAGCGGTAATG CCTGGTGGCCCAGC	NGC	32.2	No score
142	184	8	0.49	1.68	1.08	CAGAGGAAGCTG GGCCACC	AACGACTGGAGCGGTAA TGCCTGGTGGCCCAGCT TCCT	NGG	21.8	51.6
143	185	8	0.69	1.47	1.08	CAGAGGAAGCTG GGCCACC	GAGCGGTAATGCCTGGT GGCCCAGCTTCCTCT	NGG	21.8	51.6
144	299	7	0.95	1.20	1.08	AGAGGAAGCTGG GCCACCA	AACGACTGGAGCGGTAA TGCCTGGTGGCCCAGCT TCCTC	NGC	32.2	No score
145	246	3	0.56	1.58	1.07	ATGGTCCTCATGA ACGACT	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG ACCA	NGA	42.7	No score
146	171	5	1.05	1.01	1.03	CCTCATGAACGAC TGGAGC	CATTACCGCTCCAGTCG TTCATGAG	NGT	17.0	No score
147	300	7	0.30	1.73	1.01	AGAGGAAGCTGG GCCACCA	ACTGGAGCGGTAATGCC TGGTGGCCCAGCTTCCT C	NGC	32.2	No score
148	11	1	0.47	1.55	1.01	TCGGGATGGTCCT CATGAA	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG ACC	NGA	31.9	No score
149	65	4	0.70	1.32	1.01	GGTCCTCATGAAC GACTGG	ACCGCTCCAGTCGTTCA TGAGG	NGC	37.2	No score
150	110	2	0.87	1.12	1.00	GATGGTCCTCATG AACGAC	ACCGCTCCAGTCGTTCA TGAGGACCAT	NGG	38.2	52.4
151	12	1	0.78	1.20	0.99	TCGGGATGGTCCT CATGAA	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGAC CATCC	NGA	31.9	No score
152	13	1	0.76	1.14	0.95	TCGGGATGGTCCT CATGAA	CAGGCATTACCGCTCCA GTCGTTCATGAGGACC	NGA	31.9	No score
153	186	8	0.12	1.76	0.94	CAGAGGAAGCTG GGCCACC	TGAACGACTGGAGCGGT AATGCCTGGTGGCCCAG CTTC	NGG	21.8	51.6
154	172	5	0.53	1.28	0.90	CCTCATGAACGAC TGGAGC	ACCGCTCCAGTCG	NGT	17.0	No score
155	247	3	0.62	1.13	0.87	ATGGTCCTCATGA ACGACT	CAGGCATTACCGCTCCA GTCGTTCATGAGG	NGA	42.7	No score

156	187	8	0.30	1.45	0.87	CAGAGGAAGCTG GGCCACC	CTCATGAACGACTGGAG CGGTAATGCCTGGTGGC CCAGCTTC	NGG	21.8	51.6
157	66	4	0.66	1.08	0.87	GGTCCTCATGAAC GACTGG	CAGGCATTACCGCTCCA GTCGTTCATGA	NGC	37.2	No score
158	67	4	0.50	1.17	0.83	GGTCCTCATGAAC GACTGG	CAGGCATTACCGCTCCA GTCGTTCAT	NGC	37.2	No score
159	14	1	0.53	1.03	0.78	TCGGGATGGTCCT CATGAA	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGAC C	NGA	31.9	No score
160	188	8	0.48	1.07	0.78	CAGAGGAAGCTG GGCCACC	TCCTCATGAACGACTGG AGCGGTAATGCCTGGTG GCCCAGCT	NGG	21.8	51.6
161	189	8	0.38	1.14	0.76	CAGAGGAAGCTG GGCCACC	AACGACTGGAGCGGTAA TGCCTGGTGGCCCAGCT TC	NGG	21.8	51.6
162	248	3	0.44	1.07	0.76	ATGGTCCTCATGA ACGACT	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA GG	NGA	42.7	No score
163	68	4	0.68	0.82	0.75	GGTCCTCATGAAC GACTGG	GGCCACCAGGCATTACC GCTCCAGTCGTTCAT	NGC	37.2	No score
164	111	2	0.52	0.98	0.75	GATGGTCCTCATG AACGAC	GCTGGGCCACCAGGCAT TACCGCTCCAGTCGTTC ATGAGGACC	NGG	38.2	52.4
165	173	5	0.57	0.91	0.74	CCTCATGAACGAC TGGAGC	ACCGCTCCAGTCGTTCA TG	NGT	17.0	No score
166	112	2	0.31	1.17	0.74	GATGGTCCTCATG AACGAC	ACCGCTCCAGTCGTTCA TGAGGA	NGG	38.2	52.4
167	69	4	0.45	1.02	0.73	GGTCCTCATGAAC GACTGG	CATTACCGCTCCAGTCG TTCAT	NGC	37.2	No score
168	70	4	1.05	0.40	0.73	GGTCCTCATGAAC GACTGG	GCTGGGCCACCAGGCAT TACCGCTCCAGTCGTTC ATGAGGAC	NGC	37.2	No score
169	301	7	0.64	0.76	0.70	AGAGGAAGCTGG GCCACCA	GAGCGGTAATGCCTGGT GGCCCA	NGC	32.2	No score
170	249	3	0.42	0.92	0.67	ATGGTCCTCATGA ACGACT	CATTACCGCTCCAGTCG TTCATGAGG	NGA	42.7	No score
171	302	7	0.67	0.66	0.67	AGAGGAAGCTGG GCCACCA	TCCTCATGAACGACTGG AGCGGTAATGCCTGGTG GCCCAGC	NGC	32.2	No score
172	190	8	0.25	1.07	0.66	CAGAGGAAGCTG GGCCACC	CATGAACGACTGGAGCG GTAATGCCTGGTGGCCC	NGG	21.8	51.6

							AGCTTCCT			
173	191	8	0.31	1.01	0.66	CAGAGGAAGCTG GGCCACC	TGAACGACTGGAGCGGT AATGCCTGGTGGCCCAG CT	NGG	21.8	51.6
174	113	2	0.56	0.75	0.66	GATGGTCCTCATG AACGAC	TTACCGCTCCAGTCGTT CATGAG	NGG	38.2	52.4
175	174	5	0.38	0.94	0.66	CCTCATGAACGAC TGGAGC	ACCGCTCCAGTCGTTCA	NGT	17.0	No score
176	250	3	0.56	0.75	0.65	ATGGTCCTCATGA ACGACT	ACCGCTCCAGTCGTTCA TGAGGACCA	NGA	42.7	No score
177	192	8	0.20	1.11	0.65	CAGAGGAAGCTG GGCCACC	TGAACGACTGGAGCGGT AATGCCTGGTGGCCCAG CTTCCT	NGG	21.8	51.6
178	303	7	0.47	0.82	0.65	AGAGGAAGCTGG GCCACCA	GCGGTAATGCCTGGTGG CCCAGCTTCCTC	NGC	32.2	No score
179	15	1	0.97	0.26	0.61	TCGGGATGGTCCT CATGAA	TTACCGCTCCAGTCGTT CATGAGGACCATCC	NGA	31.9	No score
180	304	7	0.36	0.86	0.61	AGAGGAAGCTGG GCCACCA	TGGAGCGGTAATGCCTG GTGGCCCA	NGC	32.2	No score
181	71	4	0.35	0.80	0.57	GGTCCTCATGAAC GACTGG	CCACCAGGCATTACCGC TCCAGTCGTTCATGA	NGC	37.2	No score
182	251	3	0.53	0.61	0.57	ATGGTCCTCATGA ACGACT	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG	NGA	42.7	No score
183	175	5	0.47	0.66	0.57	CCTCATGAACGAC TGGAGC	ACCGCTCCAGTCGTT	NGT	17.0	No score
184	305	7	0.68	0.45	0.56	AGAGGAAGCTGG GCCACCA	CTCATGAACGACTGGAG CGGTAATGCCTGGTGGC CCAGCTTCC	NGC	32.2	No score
185	193	8	0.30	0.79	0.55	CAGAGGAAGCTG GGCCACC	ACTGGAGCGGTAATGCC TGGTGGCCCAGCT	NGG	21.8	51.6
186	72	4	0.39	0.70	0.55	GGTCCTCATGAAC GACTGG	ACCGCTCCAGTCGTTCA TGAGGAC	NGC	37.2	No score
187	306	7	0.48	0.60	0.54	AGAGGAAGCTGG GCCACCA	GCGGTAATGCCTGGTGG CCCAGC	NGC	32.2	No score
188	307	7	0.00	1.08	0.54	AGAGGAAGCTGG GCCACCA	CTCATGAACGACTGGAG CGGTAATGCCTGGTGGC CCAGCTT	NGC	32.2	No score
189	73	4	0.51	0.55	0.53	GGTCCTCATGAAC GACTGG	CCACCAGGCATTACCGC TCCAGTCGTTCAT	NGC	37.2	No score

190	194	8	0.25	0.80	0.52	CAGAGGAAGCTG GGCCACC	ACTGGAGCGGTAATGCC TGGTGGCCCAGCTTC	NGG	21.8	51.6
191	195	8	0.23	0.80	0.52	CAGAGGAAGCTG GGCCACC	CTCATGAACGACTGGAG CGGTAATGCCTGGTGGC CCAGCT	NGG	21.8	51.6
192	308	7	0.47	0.54	0.51	AGAGGAAGCTGG GCCACCA	CGACTGGAGCGGTAATG CCTGGTGGCCCAGCTT	NGC	32.2	No score
193	252	3	0.46	0.55	0.50	ATGGTCCTCATGA ACGACT	ACCGCTCCAGTCGTTCA TGAGGAC	NGA	42.7	No score
194	196	8	0.15	0.86	0.50	CAGAGGAAGCTG GGCCACC	TGGAGCGGTAATGCCTG GTGGCCCAGCTTCCTCT	NGG	21.8	51.6
195	16	1	0.49	0.50	0.49	TCGGGATGGTCCT CATGAA	CATTACCGCTCCAGTCG TTCATGAGGACCAT	NGA	31.9	No score
196	17	1	0.29	0.65	0.47	TCGGGATGGTCCT CATGAA	ACCGCTCCAGTCGTTCA TGAGGACCATCC	NGA	31.9	No score
197	197	8	0.12	0.77	0.44	CAGAGGAAGCTG GGCCACC	TCCTCATGAACGACTGG AGCGGTAATGCCTGGTG GCCCAGCTTCCT	NGG	21.8	51.6
198	18	1	0.58	0.30	0.44	TCGGGATGGTCCT CATGAA	CAGGCATTACCGCTCCA GTCGTTCATGAGGACCA TCC	NGA	31.9	No score
199	198	8	0.21	0.61	0.41	CAGAGGAAGCTG GGCCACC	CGACTGGAGCGGTAATG CCTGGTGGCCCAGCTTC	NGG	21.8	51.6
200	74	4	0.00	0.81	0.41	GGTCCTCATGAAC GACTGG	TGGGCCACCAGGCATTA CCGCTCCAGTCGTTC	NGC	37.2	No score
201	114	2	0.42	0.39	0.41	GATGGTCCTCATG AACGAC	GGCATTACCGCTCCAGT CGTTCATGAG	NGG	38.2	52.4
202	199	8	0.32	0.49	0.40	CAGAGGAAGCTG GGCCACC	TGGAGCGGTAATGCCTG GTGGCCCAG	NGG	21.8	51.6
203	200	8	0.19	0.61	0.40	CAGAGGAAGCTG GGCCACC	CATGAACGACTGGAGCG GTAATGCCTGGTGGCCC AGCT	NGG	21.8	51.6
204	75	4	0.32	0.47	0.39	GGTCCTCATGAAC GACTGG	ACCGCTCCAGTCGTTCA TGA	NGC	37.2	No score
205	76	4	0.71	0.07	0.39	GGTCCTCATGAAC GACTGG	ACCAGGCATTACCGCTC CAGTCGTTC	NGC	37.2	No score
206	253	3	0.26	0.52	0.39	ATGGTCCTCATGA ACGACT	GGCATTACCGCTCCAGT CGTTCATGA	NGA	42.7	No score
207	309	7	0.40	0.39	0.39	AGAGGAAGCTGG GCCACCA	CGACTGGAGCGGTAATG CCTGGTGGCCCAGCTTC	NGC	32.2	No score

							C			
208	115	2	0.24	0.53	0.38	GATGGTCCTCATG AACGAC	CATTACCGCTCCAGTCG TTCATGAG	NGG	38.2	52.4
209	201	8	0.10	0.66	0.38	CAGAGGAAGCTG GGCCACC	ACTGGAGCGGTAATGCC TGGTGGCCCAGCTTCCT	NGG	21.8	51.6
210	19	1	0.56	0.20	0.38	TCGGGATGGTCCT CATGAA	CAGGCATTACCGCTCCA GTCGTTCATGAGGACCA T	NGA	31.9	No score
211	202	8	-0.05	0.79	0.37	CAGAGGAAGCTG GGCCACC	ACTGGAGCGGTAATGCC TGGTGGCCCAGCTTCCT CT	NGG	21.8	51.6
212	77	4	0.16	0.54	0.35	GGTCCTCATGAAC GACTGG	ACCGCTCCAGTCGTTCA T	NGC	37.2	No score
213	203	8	0.21	0.48	0.35	CAGAGGAAGCTG GGCCACC	GAGCGGTAATGCCTGGT GGCCCAG	NGG	21.8	51.6
214	20	1	0.64	0.02	0.33	TCGGGATGGTCCT CATGAA	CATTACCGCTCCAGTCG TTCATGAGGACC	NGA	31.9	No score
215	176	5	0.22	0.44	0.33	CCTCATGAACGAC TGGAGC	ACCGCTCCAGTCGTTCA TGAG	NGT	17.0	No score
216	78	4	0.18	0.47	0.32	GGTCCTCATGAAC GACTGG	CCACCAGGCATTACCGC TCCAGTCGTTC	NGC	37.2	No score
217	204	8	0.13	0.51	0.32	CAGAGGAAGCTG GGCCACC	CGACTGGAGCGGTAATG CCTGGTGGCCCAGCT	NGG	21.8	51.6
218	254	3	0.15	0.49	0.32	ATGGTCCTCATGA ACGACT	ACCGCTCCAGTCGTTCA TGAGG	NGA	42.7	No score
219	116	2	0.55	0.05	0.30	GATGGTCCTCATG AACGAC	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA G	NGG	38.2	52.4
220	255	3	0.49	0.11	0.30	ATGGTCCTCATGA ACGACT	TTACCGCTCCAGTCGTT CATGA	NGA	42.7	No score
221	21	1	0.06	0.52	0.29	TCGGGATGGTCCT CATGAA	ACCGCTCCAGTCGTTCA TGAGGACCAT	NGA	31.9	No score
222	327	9	0.00	0.57	0.29	AGAACATCAGAGG AAGCTG	CGACTGGAGCGGTAATG CCTGGTGGCCCAGCTTC CTC	NGC	36.7	No score
223	22	1	0.12	0.44	0.28	TCGGGATGGTCCT CATGAA	CCACCAGGCATTACCGC TCCAGTCGTTCATGAGG A	NGA	31.9	No score

224	23	1	0.56	0.00	0.28	TCGGGATGGTCCT CATGAA	GGCATTACCGCTCCAGT CGTTCATGAGGACCATC CCG	NGA	31.9	No score
225	310	7	0.13	0.37	0.25	AGAGGAAGCTGG GCCACCA	GCGGTAATGCCTGGTGG CCCAGCTTCC	NGC	32.2	No score
226	24	1	0.28	0.22	0.25	TCGGGATGGTCCT CATGAA	CGCTCCAGTCGTTCATG AGGACC	NGA	31.9	No score
227	205	8	0.33	0.14	0.23	CAGAGGAAGCTG GGCCACC	GCGGTAATGCCTGGTGG CCCAGCTTCCT	NGG	21.8	51.6
228	117	2	-0.26	0.71	0.23	GATGGTCCTCATG AACGAC	TGGGCCACCAGGCATTA CCGCTCCAGTCGTTCAT GAGGA	NGG	38.2	52.4
229	118	2	0.19	0.27	0.23	GATGGTCCTCATG AACGAC	CGCTCCAGTCGTTCATG AGGACC	NGG	38.2	52.4
230	311	7	0.14	0.32	0.23	AGAGGAAGCTGG GCCACCA	ACTGGAGCGGTAATGCC TGGTGGCCCA	NGC	32.2	No score
231	206	8	0.13	0.31	0.22	CAGAGGAAGCTG GGCCACC	CTCATGAACGACTGGAG CGGTAATGCCTGGTGGC CCAG	NGG	21.8	51.6
232	256	3	0.17	0.27	0.22	ATGGTCCTCATGA ACGACT	ACCAGGCATTACCGCTC CAGTCGTTCATGA	NGA	42.7	No score
233	25	1	0.10	0.34	0.22	TCGGGATGGTCCT CATGAA	ACCGCTCCAGTCGTTCA TGAGGACC	NGA	31.9	No score
234	328	9	0.10	0.33	0.22	AGAACATCAGAGG AAGCTG	GAGCGGTAATGCCTGGT GGCCCAGCTTCCTCTG	NGC	36.7	No score
235	119	2	0.22	0.18	0.20	GATGGTCCTCATG AACGAC	CATTACCGCTCCAGTCG TTCATG	NGG	38.2	52.4
236	207	8	0.05	0.34	0.20	CAGAGGAAGCTG GGCCACC	CGACTGGAGCGGTAATG CCTGGTGGCCCAGCTTC CT	NGG	21.8	51.6
237	208	8	0.16	0.22	0.19	CAGAGGAAGCTG GGCCACC	CATGAACGACTGGAGCG GTAATGCCTGGTGGCCC AGCTTC	NGG	21.8	51.6
238	312	7	0.27	0.10	0.18	AGAGGAAGCTGG GCCACCA	TCCTCATGAACGACTGG AGCGGTAATGCCTGGTG GCCCA	NGC	32.2	No score
239	313	7	0.14	0.21	0.17	AGAGGAAGCTGG GCCACCA	GCGGTAATGCCTGGTGG CCCAGCTT	NGC	32.2	No score
240	257	3	0.07	0.25	0.16	ATGGTCCTCATGA ACGACT	CGCTCCAGTCGTTCATG A	NGA	42.7	No score

241	258	3	0.17	0.15	0.16	ATGGTCCTCATGA ACGACT	CGCTCCAGTCGTTCATG AGGAC	NGA	42.7	No score
242	120	2	0.14	0.17	0.16	GATGGTCCTCATG AACGAC	GGCATTACCGCTCCAGT CGTTCATG	NGG	38.2	52.4
243	26	1	0.08	0.24	0.16	TCGGGATGGTCCT CATGAA	GGCATTACCGCTCCAGT CGTTCATGAGGA	NGA	31.9	No score
244	329	9	0.04	0.26	0.15	AGAACATCAGAGG AAGCTG	GGTAATGCCTGGTGGCC CAGCTTCCTC	NGC	36.7	No score
245	259	3	0.15	0.14	0.15	ATGGTCCTCATGA ACGACT	CCACCAGGCATTACCGC TCCAGTCGTTCATGA	NGA	42.7	No score
246	121	2	0.00	0.30	0.15	GATGGTCCTCATG AACGAC	CGCTCCAGTCGTTCATG AGGACCAT	NGG	38.2	52.4
247	209	8	0.04	0.26	0.15	CAGAGGAAGCTG GGCCACC	ACTGGAGCGGTAATGCC TGGTGGCCCAG	NGG	21.8	51.6
248	210	8	0.06	0.23	0.14	CAGAGGAAGCTG GGCCACC	AACGACTGGAGCGGTAA TGCCTGGTGGCCCAGCT TCCTCT	NGG	21.8	51.6
249	211	8	0.13	0.15	0.14	CAGAGGAAGCTG GGCCACC	AACGACTGGAGCGGTAA TGCCTGGTGGCCCAG	NGG	21.8	51.6
250	79	4	0.14	0.13	0.14	GGTCCTCATGAAC GACTGG	ACCGCTCCAGTCGTTC	NGC	37.2	No score
251	212	8	0.15	0.11	0.13	CAGAGGAAGCTG GGCCACC	CATGAACGACTGGAGCG GTAATGCCTGGTGGCCC AG	NGG	21.8	51.6
252	213	8	0.14	0.11	0.12	CAGAGGAAGCTG GGCCACC	GCGGTAATGCCTGGTGG CCCAGCT	NGG	21.8	51.6
253	214	8	0.07	0.18	0.12	CAGAGGAAGCTG GGCCACC	CGACTGGAGCGGTAATG CCTGGTGGCCCAG	NGG	21.8	51.6
254	260	3	0.13	0.11	0.12	ATGGTCCTCATGA ACGACT	CATTACCGCTCCAGTCG TTCATGA	NGA	42.7	No score
255	215	8	0.12	0.12	0.12	CAGAGGAAGCTG GGCCACC	GCGGTAATGCCTGGTGG CCCAGCTTC	NGG	21.8	51.6
256	330	9	0.24	0.00	0.12	AGAACATCAGAGG AAGCTG	TGGAGCGGTAATGCCTG GTGGCCCAGCTTCCTCT G	NGC	36.7	No score
257	216	8	0.23	0.00	0.12	CAGAGGAAGCTG GGCCACC	TCCTCATGAACGACTGG AGCGGTAATGCCTGGTG GCCCAGCTTC	NGG	21.8	51.6
258	27	1	0.12	0.10	0.11	TCGGGATGGTCCT CATGAA	CGCTCCAGTCGTTCATG AGGACCAT	NGA	31.9	No score

259	122	2	0.09	0.13	0.11	GATGGTCCTCATG AACGAC	CAGGCATTACCGCTCCA GTCGTTCATG	NGG	38.2	52.4
260	314	7	0.06	0.17	0.11	AGAGGAAGCTGG GCCACCA	CTCATGAACGACTGGAG CGGTAATGCCTGGTGGC CCA	NGC	32.2	No score
261	261	3	0.13	0.09	0.11	ATGGTCCTCATGA ACGACT	ACCGCTCCAGTCGTTCA T	NGA	42.7	No score
262	315	7	0.07	0.14	0.11	AGAGGAAGCTGG GCCACCA	GGTAATGCCTGGTGGCC CAGCTTCC	NGC	32.2	No score
263	316	7	-0.07	0.29	0.11	AGAGGAAGCTGG GCCACCA	CTCATGAACGACTGGAG CGGTAATGCCTGGTGGC CCAGCTTCCTC	NGC	32.2	No score
264	317	7	0.02	0.19	0.11	AGAGGAAGCTGG GCCACCA	GGTAATGCCTGGTGGCC CAGCTTCCTC	NGC	32.2	No score
265	318	7	0.00	0.21	0.10	AGAGGAAGCTGG GCCACCA	GGTAATGCCTGGTGGCC CAGCTT	NGC	32.2	No score
266	28	1	0.03	0.17	0.10	TCGGGATGGTCCT CATGAA	CATTACCGCTCCAGTCG TTCATGAGGA	NGA	31.9	No score
267	217	8	0.00	0.21	0.10	CAGAGGAAGCTG GGCCACC	GGTAATGCCTGGTGGCC CAGCTTCCTCT	NGG	21.8	51.6
268	262	3	0.04	0.17	0.10	ATGGTCCTCATGA ACGACT	ACCGCTCCAGTCGTTCA TGA	NGA	42.7	No score
269	319	7	0.20	0.00	0.10	AGAGGAAGCTGG GCCACCA	CGACTGGAGCGGTAATG CCTGGTGGCCCAGCTTC CTC	NGC	32.2	No score
270	263	3	0.10	0.10	0.10	ATGGTCCTCATGA ACGACT	CGCTCCAGTCGTTCAT	NGA	42.7	No score
271	123	2	0.03	0.17	0.10	GATGGTCCTCATG AACGAC	ACCAGGCATTACCGCTC CAGTCGTTCATGAG	NGG	38.2	52.4
272	124	2	0.02	0.17	0.10	GATGGTCCTCATG AACGAC	ACCGCTCCAGTCGTTCA TGAG	NGG	38.2	52.4
273	264	3	0.11	0.08	0.09	ATGGTCCTCATGA ACGACT	TTACCGCTCCAGTCGTT CAT	NGA	42.7	No score
274	125	2	0.10	0.08	0.09	GATGGTCCTCATG AACGAC	CAGGCATTACCGCTCCA GTCGTTCATGAG	NGG	38.2	52.4
275	218	8	0.18	0.00	0.09	CAGAGGAAGCTG GGCCACC	CTCATGAACGACTGGAG CGGTAATGCCTGGTGGC CCAGCTTCCT	NGG	21.8	51.6
276	126	2	0.09	0.09	0.09	GATGGTCCTCATG AACGAC	CGCTCCAGTCGTTCATG AG	NGG	38.2	52.4

277	219	8	0.04	0.13	0.09	CAGAGGAAGCTG GGCCACC	GGTAATGCCTGGTGGCC CAGCTTCCT	NGG	21.8	51.6
278	265	3	0.14	0.04	0.09	ATGGTCCTCATGA ACGACT	GGCATTACCGCTCCAGT CGTTCAT	NGA	42.7	No score
279	266	3	0.11	0.05	0.08	ATGGTCCTCATGA ACGACT	CGCTCCAGTCGTTCATG AGGACCA	NGA	42.7	No score
280	220	8	0.07	0.09	0.08	CAGAGGAAGCTG GGCCACC	GGTAATGCCTGGTGGCC CAG	NGG	21.8	51.6
281	80	4	0.06	0.09	0.08	GGTCCTCATGAAC GACTGG	CGCTCCAGTCGTTCATG AGG	NGC	37.2	No score
282	29	1	0.13	0.00	0.07	TCGGGATGGTCCT CATGAA	ACCGCTCCAGTCGTTCA TGAGGACCATCCCG	NGA	31.9	No score
283	127	2	0.09	0.05	0.07	GATGGTCCTCATG AACGAC	TTACCGCTCCAGTCGTT CATG	NGG	38.2	52.4
284	320	7	0.11	0.02	0.07	AGAGGAAGCTGG GCCACCA	GGTAATGCCTGGTGGCC CAGC	NGC	32.2	No score
285	267	3	0.02	0.10	0.06	ATGGTCCTCATGA ACGACT	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA	NGA	42.7	No score
286	30	1	-0.08	0.20	0.06	TCGGGATGGTCCT CATGAA	CGCTCCAGTCGTTCATG AGGACCATCC	NGA	31.9	No score
287	128	2	0.07	0.05	0.06	GATGGTCCTCATG AACGAC	CGCTCCAGTCGTTCATG	NGG	38.2	52.4
288	129	2	0.01	0.11	0.06	GATGGTCCTCATG AACGAC	CGCTCCAGTCGTTCATG AGGA	NGG	38.2	52.4
289	331	9	0.10	0.00	0.05	AGAACATCAGAGG AAGCTG	GCGGTAATGCCTGGTGG CCCAGCTTCCTCTG	NGC	36.7	No score
290	81	4	0.03	0.07	0.05	GGTCCTCATGAAC GACTGG	GGCCACCAGGCATTACC GCTCCAGTCGTTC	NGC	37.2	No score
291	31	1	0.05	0.04	0.04	TCGGGATGGTCCT CATGAA	CGCTCCAGTCGTTCATG AGGA	NGA	31.9	No score
292	321	7	0.05	0.03	0.04	AGAGGAAGCTGG GCCACCA	GGTAATGCCTGGTGGCC CA	NGC	32.2	No score
293	32	1	-0.02	0.10	0.04	TCGGGATGGTCCT CATGAA	CGCTCCAGTCGTTCATG AGGACCATCCCG	NGA	31.9	No score
294	322	7	0.08	0.00	0.04	AGAGGAAGCTGG GCCACCA	TGAACGACTGGAGCGGT AATGCCTGGTGGCCCA	NGC	32.2	No score
295	33	1	0.11	-0.04	0.03	TCGGGATGGTCCT CATGAA	TTACCGCTCCAGTCGTT CATGAGGA	NGA	31.9	No score

296	221	8	0.05	0.01	0.03	CAGAGGAAGCTG GGCCACC	GGTAATGCCTGGTGGCC CAGCT	NGG	21.8	51.6
297	82	4	-0.04	0.09	0.03	GGTCCTCATGAAC GACTGG	CGCTCCAGTCGTTC	NGC	37.2	No score
298	332	9	0.05	0.00	0.02	AGAACATCAGAGG AAGCTG	GGTAATGCCTGGTGGCC CAGCTTCCTCTG	NGC	36.7	No score
299	83	4	0.05	-0.01	0.02	GGTCCTCATGAAC GACTGG	CGCTCCAGTCGTTCATG A	NGC	37.2	No score
300	222	8	0.04	0.00	0.02	CAGAGGAAGCTG GGCCACC	GGTAATGCCTGGTGGCC CAGCTTC	NGG	21.8	51.6
301	323	7	-0.06	0.09	0.01	AGAGGAAGCTGG GCCACCA	AACGACTGGAGCGGTAA TGCCTGGTGGCCCA	NGC	32.2	No score
302	223	8	0.03	0.00	0.01	CAGAGGAAGCTG GGCCACC	TGAACGACTGGAGCGGT AATGCCTGGTGGCCCAG	NGG	21.8	51.6
303	130	2	0.00	0.03	0.01	GATGGTCCTCATG AACGAC	ACCAGGCATTACCGCTC CAGTCGTTCATG	NGG	38.2	52.4
304	224	8	0.05	-0.03	0.01	CAGAGGAAGCTG GGCCACC	GCGGTAATGCCTGGTGG CCCAG	NGG	21.8	51.6
305	268	3	-0.02	0.04	0.01	ATGGTCCTCATGA ACGACT	CCACCAGGCATTACCGC TCCAGTCGTTCAT	NGA	42.7	No score
306	84	4	0.08	-0.07	0.01	GGTCCTCATGAAC GACTGG	CGCTCCAGTCGTTCATG AGGAC	NGC	37.2	No score
307	324	7	0.01	0.00	0.00	AGAGGAAGCTGG GCCACCA	GCGGTAATGCCTGGTGG CCCA	NGC	32.2	No score
308	225	8	0.00	0.00	0.00	CAGAGGAAGCTG GGCCACC	CATGAACGACTGGAGCG GTAATGCCTGGTGGCCC AGCTTCCTCT	NGG	21.8	51.6
309	226	8	0.00	0.00	0.00	CAGAGGAAGCTG GGCCACC	TGAACGACTGGAGCGGT AATGCCTGGTGGCCCAG CTTCCTCT	NGG	21.8	51.6
310	227	8	0.00	0.00	0.00	CAGAGGAAGCTG GGCCACC	CGACTGGAGCGGTAATG CCTGGTGGCCCAGCTTC CTCT	NGG	21.8	51.6
311	228	8	0.00	0.00	0.00	CAGAGGAAGCTG GGCCACC	GCGGTAATGCCTGGTGG CCCAGCTTCCTCT	NGG	21.8	51.6
312	229	8	0.00	0.00	0.00	CAGAGGAAGCTG GGCCACC	TCCTCATGAACGACTGG AGCGGTAATGCCTGGTG GCCCAG	NGG	21.8	51.6
313	333	9	0.00	0.00	0.00	AGAACATCAGAGG AAGCTG	ACTGGAGCGGTAATGCC TGGTGGCCCAGCTTCCT	NGC	36.7	No score

							CTGAT			
314	334	9	0.00	0.00	0.00	AGAACATCAGAGGAAGCTG	GCGGTAATGCCTGGTGGCCCAGCTTCCTCTGAT	NGC	36.7	No score
315	335	9	0.00	0.00	0.00	AGAACATCAGAGGAAGCTG	AACGACTGGAGCGGTAATGCCTGGTGGCCCAGCTTCCTC	NGC	36.7	No score
316	336	9	0.00	0.00	0.00	AGAACATCAGAGGAAGCTG	ACTGGAGCGGTAATGCCTGGTGGCCCAGCTTCCTC	NGC	36.7	No score
317	337	9	0.00	0.00	0.00	AGAACATCAGAGGAAGCTG	TGGAGCGGTAATGCCTGGTGGCCCAGCTTCCTC	NGC	36.7	No score
318	338	9	0.00	0.00	0.00	AGAACATCAGAGGAAGCTG	GCGGTAATGCCTGGTGGCCCAGCTTCCTC	NGC	36.7	No score
319	269	3	0.00	0.00	0.00	ATGGTCCTCATGAACGACT	GCTGGGCCACCAGGCATTACCGCTCCAGTCGTTCATGAGG	NGA	42.7	No score
320	85	4	0.00	0.00	0.00	GGTCCTCATGAACGACTGG	GCTGGGCCACCAGGCATTACCGCTCCAGTCGTTCATGAGG	NGC	37.2	No score
321	86	4	0.00	0.00	0.00	GGTCCTCATGAACGACTGG	TGGGCCACCAGGCATTACCGCTCCAGTCGTTCAT	NGC	37.2	No score
322	270	3	-0.02	0.01	0.00	ATGGTCCTCATGAACGACT	CGCTCCAGTCGTTCATGAGG	NGA	42.7	No score
323	131	2	0.01	-0.03	-0.01	GATGGTCCTCATGAACGAC	CCACCAGGCATTACCGCTCCAGTCGTTCATG	NGG	38.2	52.4
324	271	3	0.03	-0.10	-0.03	ATGGTCCTCATGAACGACT	GGCCACCAGGCATTACCGCTCCAGTCGTTCAT	NGA	42.7	No score
325	132	2	-0.02	-0.05	-0.04	GATGGTCCTCATGAACGAC	ACCGCTCCAGTCGTTCATG	NGG	38.2	52.4
326	272	3	0.02	-0.10	-0.04	ATGGTCCTCATGAACGACT	CATTACCGCTCCAGTCGTTCAT	NGA	42.7	No score
327	87	4	-0.05	-0.06	-0.05	GGTCCTCATGAACGACTGG	CGCTCCAGTCGTTCAT	NGC	37.2	No score
328	273	3	0.00	-0.12	-0.06	ATGGTCCTCATGAACGACT	CAGGCATTACCGCTCCAGTCGTTCAT	NGA	42.7	No score
329	274	3	-0.06	-0.08	-0.07	ATGGTCCTCATGAACGACT	ACCAGGCATTACCGCTCCAGTCGTTCAT	NGA	42.7	No score
330	34	1	-0.09	-0.06	-0.07	TCGGGATGGTCCTCATGAA	CAGGCATTACCGCTCCAGTCGTTCATGAGGA	NGA	31.9	No score

331	275	3	-0.13	-0.07	-0.10	ATGGTCCTCATGA ACGACT	CAGGCATTACCGCTCCA GTCGTTCATGA	NGA	42.7	No score
332	325	7	-0.10	-0.13	-0.11	AGAGGAAGCTGG GCCACCA	CGACTGGAGCGGTAATG CCTGGTGGCCCA	NGC	32.2	No score
333	35	1	-0.11	-0.14	-0.13	TCGGGATGGTCCT CATGAA	ACCGCTCCAGTCGTTCA TGAGGA	NGA	31.9	No score
334	36	1	-0.10	-0.15	-0.13	TCGGGATGGTCCT CATGAA	ACCAGGCATTACCGCTC CAGTCGTTCATGAGGA	NGA	31.9	No score
335	133	2	-0.19	-0.13	-0.16	GATGGTCCTCATG AACGAC	CCACCAGGCATTACCGC TCCAGTCGTTCATGAG	NGG	38.2	52.4
336	339	9	-0.15	-0.19	-0.17	AGAACATCAGAGG AAGCTG	GAGCGGTAATGCCTGGT GGCCCAGCTTCCTC	NGC	36.7	No score
337	134	2	-0.08	-0.31	-0.20	GATGGTCCTCATG AACGAC	GGCCACCAGGCATTACC GCTCCAGTCGTTCATG	NGG	38.2	52.4
338	37	1	-0.39	-0.34	-0.37	TCGGGATGGTCCT CATGAA	GGCCACCAGGCATTACC GCTCCAGTCGTTCATGA GGA	NGA	31.9	No score
339	326	7	-0.31	-0.48	-0.40	AGAGGAAGCTGG GCCACCA	CATGAACGACTGGAGCG GTAATGCCTGGTGGCCC A	NGC	32.2	No score

Table 3. Predicted activities of the sgRNA candidates for the PE3-directed correction of the tyrosinemia-causing mutation

ID	Location from the pegRNA nick site	Target Sequence (30bp)	Guide Sequence (20bp)	DeepSpCas9 Score	PE3 or PE3b
1	+38	GTGCCCCTAAGAACAGAACATCAGAGGAAG	CCCTAAGAACAGAACATCAG	69.531	PE3
2	+30	AAGAACAGAACATCAGAGGAAGCTGGGCCA	ACAGAACATCAGAGGAAGCT	60.865	PE3
3	+22	AACATCAGAGGAAGCTGGGCCACCAGGCAT	TCAGAGGAAGCTGGGCCACC	51.586	PE3
4	+31	TAAGAACAGAACATCAGAGGAAGCTGGGCC	AACAGAACATCAGAGGAAGC	23.167	PE3
5	-4	GCATTACCGCTCCAGTCGTTCATGAGGACC	TACCGCTCCAGTCGTTCATG	43.183	PE3b

Table 4. Potential off-target sites evaluated in the study

Fah							
Target site	Chr	Location	DNA sequence at cleavage site	Number of mismatched nucleotides	Predicted by Digenome-seq	Predicted by nDigenome-seq	Predicted by CRISPOR
on-target	chr7	84595468	GGATGGTCCTCATGAACGACTGG	0			
Fah-OT1	chr1	90482995	aGAgaGTaCTCATGAACcACAGG	5	Yes	No	No
Fah-OT2	chr1	90485057	aGAgaGTaCTCATGAACcACAGG	5	Yes	No	No
Fah-OT3	chr18	25554349	tGATGGTCtcCATGAACGcCAGG	4	No	Yes	No
Fah-OT4	chr9	22667145	GtAaGtTCCaCATGAACaACAGG	5	No	Yes	No
Fah-OT5	chr3	29850736	caATGtcCtTCATGAACGACAGG	5	No	Yes	No
Fah-OT6	chr18	41986725	GaATGGTagaCATGAACcACCGG	5	No	Yes	No
Fah-OT7	chr17	46981418	GcAgGcTgCTCATGAACaACTGG	5	No	Yes	No
Fah-OT8	chr3	132846040	GatTGcTgCTCATGAACGACAGG	4	No	No	Yes
Fah-OT9	chr13	33487838	ctATGGtTCTCATGAAaGACAGG	4	No	No	Yes
Fah-OT10	chr11	39836760	aGtTaGTCCTCATGAAgGACAGG	4	No	No	Yes
Fah-OT11	chr4	73415802	GctcGGTCCTCATGAAgGACAGG	4	No	No	Yes
sgRNA on-target	chr7	84244619	CCCTAAGAACAGAACATCAGAGG	0			
Fah-sOT1	chr11	35554393	CCCTAAGAACAGAACATCAtGGA	1	No	No	Yes

Fah-sOT2	chr5	62342873	ttCTAAGAAaAGAACATCAGCGG	3	No	No	Yes
Fah-sOT3	chr9	99822853	CCaTtAaAACAGAACATCAGAGG	3	No	No	Yes
Fah-sOT4	chr13	46881500	CtgTAAtAACAGAACATCAGGGG	3	No	No	Yes

Rpe65

Target site	Chr	Location	DNA sequence at cleavage site	Number of mismatched nucleotides	Predicted by Digenome-seq	Predicted by nDigenome-seq	Predicted by CRISPOR
on-target	chr3	159307186	AGAGCCCTGGCCCACATCAGAGG	0			
Rpe65-OT1	chr3	159601555	AGAGCCCTGGCCCACATCgGAGG	1	Yes	Yes	Yes
Rpe65-OT2	chr11	21648838	AGAGCCCTGtCCCAgATCAGAGG	2	Yes	Yes	Yes
Rpe65-OT3	chr3	153824076	AGAGCtCTGGCtCACAgCAGCGG	3	Yes	Yes	No
Rpe65-OT4	chr1	164409774	AaAGCCCTGGgCCACAgCAGAGG	3	Yes	No	No
Rpe65-OT5	chr6	91092394	AGgGCCCTGGgCCACAgCAGGGG	3	Yes	Yes	No
Rpe65-OT6	chr2	101041865	AcAGCCCTGGCCaAaATCAGTGG	3	Yes	Yes	No
Rpe65-OT7	chr18	75673981	AGtGCCCTGGCCCcaATCAGTGG	3	Yes	No	No
Rpe65-OT8	chr16	22650452	GGAaCCCTGGCCCAgAcCAGGGG	3	No	No	Yes
Rpe65-OT9	chr15	100412547	GGAGCgCTGGCCCACAggAGTGG	3	No	No	Yes
Rpe65-OT10	chr11	6772263	GaAGgCaTGGCCCACATCAGAGG	3	No	No	Yes
Rpe65-OT11	chr11	108941790	GGgGaCCTGGCgCACATCAGTGG	3	No	No	Yes
Rpe65-OT12	chr11	118086362	GGAGCCCcaGCCCACcTCAGAGG	3	No	No	Yes

Rpe65-OT13	chr12	5245047	GGAGCCCTGGCCCtgcTCAGAGG	3	No	No	Yes
Rpe65-OT14	chr12	9259564	GGAGCCtgGGCCCACATCtGTGG	3	No	No	Yes
Rpe65-OT15	chr12	105187363	GGAGCCCTGtCCCACATaAcTGG	3	No	No	Yes
Rpe65-OT16	chr13	56434803	GGcGCCCaGGaCCACATCAGAGG	3	No	No	Yes
Rpe65-OT17	chr13	89009358	GtAGtCCTtGCCCACATCAGGGG	3	No	No	Yes
Rpe65-OT18	chr6	123856088	GGAGCCtTGGCCaACATCtGTGG	3	No	No	Yes
Rpe65-OT19	chr1	192223740	GGAGCCCTGGCtCtCAgCAGAGG	3	No	No	Yes
Rpe65-OT20	chr10	127519207	GcAGaCaTGGCCCACATCAGTGG	3	No	No	Yes

Table 5. Prime editing efficiencies of the *Rpe65* pegRNAs measured using a paired library approach

Target ID	PE efficiency (%, Replicate 1)	PE efficiency (%, Replicate 2)	Average PE efficiency (%)	Rank	pegRNA ID	Spacer sequence	RTPBS sequence	PAM
5	10.733	19.458	15.096	1	157	GAGCCCTGGCCCACATCAG	GCAGTCTCCTCCGATGTGGGCC	NGG
5	10.830	17.067	13.949	2	198	GAGCCCTGGCCCACATCAG	GGCAGTCTCCTCCGATGTGGGCC	NGG
5	9.366	18.208	13.787	3	159	GAGCCCTGGCCCACATCAG	TCCTCCGATGTGGGCC	NGG
5	9.880	17.336	13.608	4	188	GAGCCCTGGCCCACATCAG	GCAGTCTCCTCCGATGTGGGCCAG	NGG
5	9.138	17.434	13.286	5	153	GAGCCCTGGCCCACATCAG	TCCTCCGATGTGGGCCAG	NGG
5	9.474	16.963	13.218	6	201	GAGCCCTGGCCCACATCAG	CTCCTCCGATGTGGGCC	NGG
5	9.888	16.474	13.181	7	187	GAGCCCTGGCCCACATCAG	GGCAGTCTCCTCCGATGTGGGCCAG	NGG
5	9.559	16.598	13.078	8	156	GAGCCCTGGCCCACATCAG	ACTGGCAGTCTCCTCCGATGTGGGCC	NGG
5	8.755	16.661	12.708	9	190	GAGCCCTGGCCCACATCAG	TCTCCTCCGATGTGGGCCAG	NGG
5	8.567	16.735	12.651	10	200	GAGCCCTGGCCCACATCAG	TCTCCTCCGATGTGGGCC	NGG
5	8.321	16.054	12.188	11	199	GAGCCCTGGCCCACATCAG	AGTCTCCTCCGATGTGGGCC	NGG
5	9.343	15.005	12.174	12	197	GAGCCCTGGCCCACATCAG	TGGCAGTCTCCTCCGATGTGGGCC	NGG
5	8.736	14.963	11.849	13	151	GAGCCCTGGCCCACATCAG	TGGCAGTCTCCTCCGATGTGGGCCAG	NGG
5	6.919	16.706	11.812	14	160	GAGCCCTGGCCCACATCAG	CCTCCGATGTGGGCC	NGG
5	7.615	16.009	11.812	15	184	GAGCCCTGGCCCACATCAG	TCCTCCGATGTGGGCCAGGG	NGG
1	7.170	16.004	11.587	16	129	AAGAGCCCTGGCCCACATC	TCCGATGTGGGCC	NGA

5	9.576	13.025	11.300	17	149	GAGCCCTGGCCCACATCAG	ACTGGCAGTCTCCTCCGATGTGGGCCAG	NGG
5	7.645	14.765	11.205	18	144	GAGCCCTGGCCCACATCAG	TCTCCTCCGATGTGGGCCAGGG	NGG
5	7.846	14.364	11.105	19	177	GAGCCCTGGCCCACATCAG	TCCTCCGATGTGGGCCAGGGCT	NGG
5	7.941	13.862	10.901	20	168	GAGCCCTGGCCCACATCAG	TCCTCCGATGTGGG	NGG
5	7.386	13.918	10.652	21	183	GAGCCCTGGCCCACATCAG	CTCCTCCGATGTGGGCCAGGG	NGG
5	7.332	13.642	10.487	22	167	GAGCCCTGGCCCACATCAG	CTCCTCCGATGTGGG	NGG
5	7.337	13.461	10.399	23	209	GAGCCCTGGCCCACATCAG	CCTCCGATGTGGG	NGG
1	6.432	14.310	10.371	24	109	AAGAGCCCTGGCCCACATC	CTCCTCCGATGTGGGCCAGGG	NGA
5	7.408	13.313	10.360	25	154	GAGCCCTGGCCCACATCAG	GCTCACTGGCAGTCTCCTCCGATGTGGGCC	NGG
5	6.178	14.151	10.164	26	152	GAGCCCTGGCCCACATCAG	AGTCTCCTCCGATGTGGGCCAG	NGG
5	6.302	13.844	10.073	27	158	GAGCCCTGGCCCACATCAG	CAGTCTCCTCCGATGTGGGCC	NGG
1	7.120	12.844	9.982	28	127	AAGAGCCCTGGCCCACATC	CTCCTCCGATGTGGGCC	NGA
1	6.829	12.596	9.713	29	111	AAGAGCCCTGGCCCACATC	TCCGATGTGGGCCAGGG	NGA
5	7.151	12.023	9.587	30	135	GAGCCCTGGCCCACATCAG	TCTCCTCCGATGTGGGCCAGGGCT	NGG
5	7.021	12.148	9.584	31	141	GAGCCCTGGCCCACATCAG	TGGCAGTCTCCTCCGATGTGGGCCAGGG	NGG
5	7.233	11.908	9.571	32	196	GAGCCCTGGCCCACATCAG	CTGGCAGTCTCCTCCGATGTGGGCC	NGG
5	7.461	11.629	9.545	33	143	GAGCCCTGGCCCACATCAG	GCAGTCTCCTCCGATGTGGGCCAGGG	NGG
1	7.037	11.945	9.491	34	125	AAGAGCCCTGGCCCACATC	GGCAGTCTCCTCCGATGTGGGCC	NGA

5	7.087	11.610	9.348	35	147	GAGCCCTGGCCCACATCAG	GCTCACTGGCAGTCTCCTCCGATGTGGGCCAG	NGG
5	5.819	12.497	9.158	36	191	GAGCCCTGGCCCACATCAG	CTCCTCCGATGTGGGCCAG	NGG
1	5.314	12.730	9.022	37	110	AAGAGCCCTGGCCCACATC	CCTCCGATGTGGGCCAGGG	NGA
1	6.506	11.472	8.989	38	118	AAGAGCCCTGGCCCACATC	CTCCTCCGATGTGGGCCAG	NGA
5	6.208	11.226	8.717	39	185	GAGCCCTGGCCCACATCAG	TCACTGGCAGTCTCCTCCGATGTGGGCCAG	NGG
5	6.793	10.447	8.620	40	142	GAGCCCTGGCCCACATCAG	GGCAGTCTCCTCCGATGTGGGCCAGGG	NGG
5	5.729	11.497	8.613	41	192	GAGCCCTGGCCCACATCAG	CCTCCGATGTGGGCCAG	NGG
5	5.942	11.242	8.592	42	208	GAGCCCTGGCCCACATCAG	TCTCCTCCGATGTGGG	NGG
5	6.635	10.363	8.499	43	179	GAGCCCTGGCCCACATCAG	ACTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGG
5	6.676	10.296	8.486	44	133	GAGCCCTGGCCCACATCAG	GCAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
1	5.909	11.004	8.457	45	107	AAGAGCCCTGGCCCACATC	GGCAGTCTCCTCCGATGTGGGCCAGGG	NGA
1	5.738	10.900	8.319	46	120	AAGAGCCCTGGCCCACATC	TCCGATGTGGGCCAG	NGA
1	6.504	9.758	8.131	47	126	AAGAGCCCTGGCCCACATC	CAGTCTCCTCCGATGTGGGCC	NGA
5	5.941	10.198	8.069	48	150	GAGCCCTGGCCCACATCAG	CTGGCAGTCTCCTCCGATGTGGGCCAG	NGG
5	6.541	9.397	7.969	49	132	GAGCCCTGGCCCACATCAG	TGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
5	5.402	10.492	7.947	50	145	GAGCCCTGGCCCACATCAG	CCTCCGATGTGGGCCAGGG	NGG
5	5.142	10.679	7.911	51	155	GAGCCCTGGCCCACATCAG	TCACTGGCAGTCTCCTCCGATGTGGGCC	NGG
5	6.558	9.226	7.892	52	138	GAGCCCTGGCCCACATCAG	GCTCACTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGG

5	5.563	9.974	7.769	53	181	GAGCCCTGGCCCACATCAG	CAGTCTCCTCCGATGTGGGCCAGGG	NGG
1	5.527	9.317	7.422	54	106	AAGAGCCCTGGCCCACATC	CTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGA
5	4.621	10.147	7.384	55	207	GAGCCCTGGCCCACATCAG	AGTCTCCTCCGATGTGGG	NGG
1	6.067	8.684	7.375	56	116	AAGAGCCCTGGCCCACATC	GGCAGTCTCCTCCGATGTGGGCCAG	NGA
5	6.059	8.636	7.347	57	173	GAGCCCTGGCCCACATCAG	ACTGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
5	5.138	9.521	7.330	58	174	GAGCCCTGGCCCACATCAG	GGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
1	5.656	8.992	7.324	59	108	AAGAGCCCTGGCCCACATC	CAGTCTCCTCCGATGTGGGCCAGGG	NGA
5	7.037	7.478	7.257	60	130	GAGCCCTGGCCCACATCAG	GGCTCACTGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
5	4.887	9.312	7.100	61	182	GAGCCCTGGCCCACATCAG	AGTCTCCTCCGATGTGGGCCAGGG	NGG
5	4.234	9.884	7.059	62	140	GAGCCCTGGCCCACATCAG	TCACTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGG
5	6.291	7.780	7.036	63	169	GAGCCCTGGCCCACATCAG	GCTCACTGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
1	5.061	8.873	6.967	64	128	AAGAGCCCTGGCCCACATC	CCTCCGATGTGGGCC	NGA
5	4.326	9.234	6.780	65	189	GAGCCCTGGCCCACATCAG	CAGTCTCCTCCGATGTGGGCCAG	NGG
5	5.011	8.352	6.681	66	180	GAGCCCTGGCCCACATCAG	CTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGG
5	4.754	8.471	6.612	67	176	GAGCCCTGGCCCACATCAG	CTCCTCCGATGTGGGCCAGGGCT	NGG
5	4.710	7.904	6.307	68	137	GAGCCCTGGCCCACATCAG	GGCTCACTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGG
5	3.953	8.660	6.307	69	134	GAGCCCTGGCCCACATCAG	AGTCTCCTCCGATGTGGGCCAGGGCT	NGG
1	3.841	8.524	6.183	70	124	AAGAGCCCTGGCCCACATC	CTGGCAGTCTCCTCCGATGTGGGCC	NGA

1	3.878	8.460	6.169	71	119	AAGAGCCCTGGCCCACATC	CCTCCGATGTGGGCCAG	NGA
1	4.647	7.577	6.112	72	115	AAGAGCCCTGGCCCACATC	CTGGCAGTCTCCTCCGATGTGGGCCAG	NGA
1	3.675	8.136	5.906	73	102	AAGAGCCCTGGCCCACATC	TCCGATGTGGGCCAGGGCT	NGA
5	4.621	6.929	5.775	74	170	GAGCCCTGGCCCACATCAG	CTCACTGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
5	4.438	6.847	5.643	75	139	GAGCCCTGGCCCACATCAG	CTCACTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGG
5	4.617	6.661	5.639	76	131	GAGCCCTGGCCCACATCAG	CTGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
5	3.695	7.380	5.538	77	205	GAGCCCTGGCCCACATCAG	GCAGTCTCCTCCGATGTGGG	NGG
1	4.342	6.722	5.532	78	103	AAGAGCCCTGGCCCACATC	GGCTCACTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGA
5	4.145	6.728	5.436	79	148	GAGCCCTGGCCCACATCAG	CTCACTGGCAGTCTCCTCCGATGTGGGCCAG	NGG
1	4.297	6.544	5.420	80	105	AAGAGCCCTGGCCCACATC	CACTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGA
1	4.456	5.671	5.064	81	98	AAGAGCCCTGGCCCACATC	GGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGA
5	3.377	6.576	4.976	82	171	GAGCCCTGGCCCACATCAG	TCACTGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
1	2.591	7.128	4.860	83	93	AAGAGCCCTGGCCCACATC	TCCGATGTGGGCCAGGGCTCT	NGA
5	3.266	6.446	4.856	84	166	GAGCCCTGGCCCACATCAG	TGGCAGTCTCCTCCGATGTGGG	NGG
5	3.939	5.644	4.792	85	175	GAGCCCTGGCCCACATCAG	CAGTCTCCTCCGATGTGGGCCAGGGCT	NGG
1	3.182	6.302	4.742	86	117	AAGAGCCCTGGCCCACATC	CAGTCTCCTCCGATGTGGGCCAG	NGA
5	3.767	5.494	4.631	87	193	GAGCCCTGGCCCACATCAG	GGCTCACTGGCAGTCTCCTCCGATGTGGGCC	NGG
1	2.939	6.057	4.498	88	100	AAGAGCCCTGGCCCACATC	CTCCTCCGATGTGGGCCAGGGCT	NGA

5	3.307	5.623	4.465	89	146	GAGCCCTGGCCCACATCAG	GGCTCACTGGCAGTCTCCTCCGATGTGGGCCAG	NGG
5	2.876	6.034	4.455	90	136	GAGCCCTGGCCCACATCAG	CCTCCGATGTGGGCCAGGGCT	NGG
1	3.437	5.322	4.380	91	113	AAGAGCCCTGGCCCACATC	CTCACTGGCAGTCTCCTCCGATGTGGGCCAG	NGA
5	2.962	5.447	4.204	92	194	GAGCCCTGGCCCACATCAG	CTCACTGGCAGTCTCCTCCGATGTGGGCC	NGG
1	3.644	4.761	4.202	93	104	AAGAGCCCTGGCCCACATC	CTCACTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGA
5	2.996	5.299	4.148	94	178	GAGCCCTGGCCCACATCAG	CACTGGCAGTCTCCTCCGATGTGGGCCAGGG	NGG
9	3.302	4.434	3.868	95	278	TCCCCTCTGGCTCACTGGC	CCCACATCGGAGGAGACTGCCAGTGAGCCAGAGG	NGT
9	2.407	5.162	3.784	96	279	TCCCCTCTGGCTCACTGGC	CACATCGGAGGAGACTGCCAGTGAGCCAGAGG	NGT
1	2.013	5.375	3.694	97	101	AAGAGCCCTGGCCCACATC	CCTCCGATGTGGGCCAGGGCT	NGA
1	2.362	4.532	3.447	98	97	AAGAGCCCTGGCCCACATC	CTGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGA
9	1.583	5.106	3.345	99	269	TCCCCTCTGGCTCACTGGC	GGCCCACATCGGAGGAGACTGCCAGTGAGCCAGAGGGG	NGT
1	2.229	4.455	3.342	100	121	AAGAGCCCTGGCCCACATC	GGCTCACTGGCAGTCTCCTCCGATGTGGGCC	NGA
1	2.431	4.248	3.339	101	112	AAGAGCCCTGGCCCACATC	GGCTCACTGGCAGTCTCCTCCGATGTGGGCCAG	NGA
5	2.244	4.396	3.320	102	204	GAGCCCTGGCCCACATCAG	GGCAGTCTCCTCCGATGTGGG	NGG
1	2.649	3.946	3.297	103	89	AAGAGCCCTGGCCCACATC	GGCAGTCTCCTCCGATGTGGGCCAGGGCTCT	NGA
5	2.284	4.101	3.193	104	206	GAGCCCTGGCCCACATCAG	CAGTCTCCTCCGATGTGGG	NGG
5	2.092	4.164	3.128	105	195	GAGCCCTGGCCCACATCAG	CACTGGCAGTCTCCTCCGATGTGGGCC	NGG
9	1.100	4.966	3.033	106	280	TCCCCTCTGGCTCACTGGC	CATCGGAGGAGACTGCCAGTGAGCCAGAGG	NGT

5	2.652	3.345	2.998	107	172	GAGCCCTGGCCCACATCAG	CACTGGCAGTCTCCTCCGATGTG GGCCAGGGCT	NGG
9	2.221	3.421	2.821	108	276	TCCCCTCTGGCTCACTGGC	CTGGCCCACATCGGAGGAGACT GCCAGTGAGCCAGAGG	NGT
9	2.147	3.344	2.746	109	286	TCCCCTCTGGCTCACTGGC	GGCCCACATCGGAGGAGACTGC CAGTGAGCCAGA	NGT
1	2.564	2.872	2.718	110	114	AAGAGCCCTGGCCCACATC	CACTGGCAGTCTCCTCCGATGTG GGCCAG	NGA
1	1.918	3.457	2.687	111	91	AAGAGCCCTGGCCCACATC	CTCCTCCGATGTGGGCCAGGGC TCT	NGA
1	2.279	3.054	2.667	112	99	AAGAGCCCTGGCCCACATC	CAGTCTCCTCCGATGTGGGCCA GGGCT	NGA
1	2.055	3.229	2.642	113	122	AAGAGCCCTGGCCCACATC	CTCACTGGCAGTCTCCTCCGATG TGGGCC	NGA
9	1.404	3.760	2.582	114	277	TCCCCTCTGGCTCACTGGC	GGCCCACATCGGAGGAGACTGC CAGTGAGCCAGAGG	NGT
1	1.764	3.364	2.564	115	123	AAGAGCCCTGGCCCACATC	CACTGGCAGTCTCCTCCGATGTG GGCC	NGA
2	0.691	4.276	2.484	116	33	ATCCAACTTCAAAGAGCCC	GGCTCACTGGCAGTCTCCTCCGA TGTGGGCCAGGGCTCTTTGAA	NGG
1	1.583	3.373	2.478	117	88	AAGAGCCCTGGCCCACATC	CTGGCAGTCTCCTCCGATGTGGG CCAGGGCTCT	NGA
9	1.978	2.721	2.349	118	270	TCCCCTCTGGCTCACTGGC	CCCACATCGGAGGAGACTGCCA GTGAGCCAGAGGGG	NGT
9	2.307	2.124	2.216	119	271	TCCCCTCTGGCTCACTGGC	CACATCGGAGGAGACTGCCAGT GAGCCAGAGGGG	NGT
2	3.631	0.758	2.194	120	14	ATCCAACTTCAAAGAGCCC	GGCTCACTGGCAGTCTCCTCCGA TGTGGGCCAGGGCTCTTTG	NGG
5	1.353	2.968	2.160	121	186	GAGCCCTGGCCCACATCAG	CACTGGCAGTCTCCTCCGATGTG GGCCAG	NGG
9	1.520	2.517	2.019	122	272	TCCCCTCTGGCTCACTGGC	CATCGGAGGAGACTGCCAGTGA GCCAGAGGGG	NGT
7	1.967	1.942	1.954	123	236	GGATTCCCCTCTGGCTCAC	CTGGCCCACATCGGAGGAGACT GCCAGTGAGCCAGAGGGGAA	NGG
1	1.477	2.369	1.923	124	95	AAGAGCCCTGGCCCACATC	CTCACTGGCAGTCTCCTCCGATG TGGGCCAGGGCT	NGA

8	1.701	1.948	1.824	125	51	TCCAACTTCAAAGAGCCCT	TCCGATGTGGGCCAGGGCTCTTTGAAGTTGG	NGC
9	1.096	2.452	1.774	126	287	TCCCCTCTGGCTCACTGGC	CCCACATCGGAGGAGACTGCCAGTGAGCCAGA	NGT
9	1.213	2.215	1.714	127	289	TCCCCTCTGGCTCACTGGC	CATCGGAGGAGACTGCCAGTGAGCCAGA	NGT
1	0.890	2.321	1.605	128	90	AAGAGCCCTGGCCCACATC	CAGTCTCCTCCGATGTGGGCCAGGGCTCT	NGA
5	1.133	2.064	1.598	129	165	GAGCCCTGGCCCACATCAG	CTGGCAGTCTCCTCCGATGTGGG	NGG
1	1.616	1.566	1.591	130	86	AAGAGCCCTGGCCCACATC	CTCACTGGCAGTCTCCTCCGATGTGGGCCAGGGCTCT	NGA
9	0.780	2.364	1.572	131	268	TCCCCTCTGGCTCACTGGC	CTGGCCCACATCGGAGGAGACTGCCAGTGAGCCAGAGGGG	NGT
7	0.768	2.308	1.538	132	244	GGATTCCCCTCTGGCTCAC	CTGGCCCACATCGGAGGAGACTGCCAGTGAGCCAGAGGGG	NGG
5	0.713	2.313	1.513	133	164	GAGCCCTGGCCCACATCAG	ACTGGCAGTCTCCTCCGATGTGGG	NGG
1	1.120	1.905	1.512	134	96	AAGAGCCCTGGCCCACATC	CACTGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGA
9	0.931	1.957	1.444	135	295	TCCCCTCTGGCTCACTGGC	GGCCCACATCGGAGGAGACTGCCAGTGAGCCA	NGT
1	1.200	1.638	1.419	136	92	AAGAGCCCTGGCCCACATC	CCTCCGATGTGGGCCAGGGCTCT	NGA
9	1.150	1.603	1.376	137	298	TCCCCTCTGGCTCACTGGC	CATCGGAGGAGACTGCCAGTGAGCCA	NGT
7	0.330	2.420	1.375	138	245	GGATTCCCCTCTGGCTCAC	GGCCCACATCGGAGGAGACTGCCAGTGAGCCAGAGGGG	NGG
7	1.055	1.604	1.330	139	248	GGATTCCCCTCTGGCTCAC	CATCGGAGGAGACTGCCAGTGAGCCAGAGGGG	NGG
9	1.569	1.085	1.327	140	275	TCCCCTCTGGCTCACTGGC	CCCTGGCCCACATCGGAGGAGACTGCCAGTGAGCCAGAGG	NGT
9	1.027	1.625	1.326	141	296	TCCCCTCTGGCTCACTGGC	CCCACATCGGAGGAGACTGCCAGTGAGCCA	NGT
8	1.299	1.344	1.322	142	49	TCCAACTTCAAAGAGCCCT	CTCCTCCGATGTGGGCCAGGGCTCTTTGAAGTTGG	NGC

9	1.180	1.455	1.317	143	288	TCCCCTCTGGCTCACTGGC	CACATCGGAGGAGACTGCCAGT GAGCCAGA	NGT
9	1.575	0.961	1.268	144	285	TCCCCTCTGGCTCACTGGC	CTGGCCCACATCGGAGGAGACT GCCAGTGAGCCAGA	NGT
9	0.814	1.631	1.222	145	292	TCCCCTCTGGCTCACTGGC	AGCCCTGGCCCACATCGGAGGA GACTGCCAGTGAGCCA	NGT
5	0.808	1.461	1.135	146	161	GAGCCCTGGCCCACATCAG	GGCTCACTGGCAGTCTCCTCCGA TGTGGG	NGG
7	0.978	1.276	1.127	147	246	GGATTCCCCTCTGGCTCAC	CCCACATCGGAGGAGACTGCCA GTGAGCCAGAGGGG	NGG
8	0.760	1.434	1.097	148	59	TCCAACTTCAAAGAGCCCT	TCCGATGTGGGCCAGGGCTCTTT GAAGTT	NGC
7	0.931	1.193	1.062	149	239	GGATTCCCCTCTGGCTCAC	CATCGGAGGAGACTGCCAGTGA GCCAGAGGGGAA	NGG
8	0.394	1.711	1.052	150	67	TCCAACTTCAAAGAGCCCT	TCCGATGTGGGCCAGGGCTCTTT GAAG	NGC
9	0.507	1.519	1.013	151	284	TCCCCTCTGGCTCACTGGC	CCCTGGCCCACATCGGAGGAGA CTGCCAGTGAGCCAGA	NGT
9	0.966	1.018	0.992	152	294	TCCCCTCTGGCTCACTGGC	CTGGCCCACATCGGAGGAGACT GCCAGTGAGCCA	NGT
7	1.149	0.775	0.962	153	235	GGATTCCCCTCTGGCTCAC	CCCTGGCCCACATCGGAGGAGA CTGCCAGTGAGCCAGAGGGGAA	NGG
5	0.716	1.108	0.912	154	162	GAGCCCTGGCCCACATCAG	GCTCACTGGCAGTCTCCTCCGAT GTGGG	NGG
7	0.844	0.980	0.912	155	237	GGATTCCCCTCTGGCTCAC	GGCCCACATCGGAGGAGACTGC CAGTGAGCCAGAGGGGAA	NGG
10	0.440	1.333	0.886	156	215	TGCAGGCAGGATTCCCCTC	GGAGGAGACTGCCAGTGAGCCA GAGGGGAATCCT	NGG
2	0.471	1.241	0.856	157	23	ATCCAACTTCAAAGAGCCC	TCCGATGTGGGCCAGGGCTCTTT GAAGTTGGA	NGG
1	0.888	0.823	0.855	158	87	AAGAGCCCTGGCCCACATC	CACTGGCAGTCTCCTCCGATGTG GGCCAGGGCTCT	NGA
7	0.640	1.058	0.849	159	255	GGATTCCCCTCTGGCTCAC	CACATCGGAGGAGACTGCCAGT GAGCCAGAGG	NGG
2	0.878	0.774	0.826	160	6	ATCCAACTTCAAAGAGCCC	TCCGATGTGGGCCAGGGCTCTTT GAAGTTG	NGG

2	1.510	0.071	0.791	161	3	ATCCAACTTCAAAGAGCCC	CTCACTGGCAGTCTCCTCCGATGTGGGCCAGGGCTCTTTGAAGTTG	NGG
9	0.535	1.006	0.770	162	305	TCCCCTCTGGCTCACTGGC	CCCACATCGGAGGAGACTGCCAGTGAGC	NGT
5	0.470	0.994	0.732	163	202	GAGCCCTGGCCCACATCAG	TCACTGGCAGTCTCCTCCGATGTGGG	NGG
8	0.755	0.671	0.713	164	65	TCCAACTTCAAAGAGCCCT	CTCCTCCGATGTGGGCCAGGGCTCTTTGAAG	NGC
7	0.980	0.413	0.697	165	232	GGATTCCCCTCTGGCTCAC	CACATCGGAGGAGACTGCCAGTGAGCCAGAGGGGAATC	NGG
8	0.379	0.945	0.662	166	70	TCCAACTTCAAAGAGCCCT	CTGGCAGTCTCCTCCGATGTGGGCCAGGGCTCTTTGA	NGC
7	0.371	0.949	0.660	167	247	GGATTCCCCTCTGGCTCAC	CACATCGGAGGAGACTGCCAGTGAGCCAGAGGGG	NGG
1	1.293	0.000	0.647	168	94	AAGAGCCCTGGCCCACATC	GGCTCACTGGCAGTCTCCTCCGATGTGGGCCAGGGCT	NGA
2	0.782	0.471	0.626	169	31	ATCCAACTTCAAAGAGCCC	TCCGATGTGGGCCAGGGCTCTTTGAAGT	NGG
5	0.691	0.561	0.626	170	163	GAGCCCTGGCCCACATCAG	CTCACTGGCAGTCTCCTCCGATGTGGG	NGG
2	0.416	0.828	0.622	171	21	ATCCAACTTCAAAGAGCCC	CTCCTCCGATGTGGGCCAGGGCTCTTTGAAGTTGGA	NGG
8	0.274	0.931	0.602	172	66	TCCAACTTCAAAGAGCCCT	CCTCCGATGTGGGCCAGGGCTCTTTGAAG	NGC
9	0.440	0.751	0.596	173	297	TCCCCTCTGGCTCACTGGC	CACATCGGAGGAGACTGCCAGTGAGCCA	NGT
8	0.137	1.046	0.591	174	50	TCCAACTTCAAAGAGCCCT	CCTCCGATGTGGGCCAGGGCTCTTTGAAGTTGG	NGC
8	0.384	0.794	0.589	175	62	TCCAACTTCAAAGAGCCCT	CTGGCAGTCTCCTCCGATGTGGGCCAGGGCTCTTTGAAG	NGC
7	0.461	0.681	0.571	176	252	GGATTCCCCTCTGGCTCAC	CTGGCCCACATCGGAGGAGACTGCCAGTGAGCCAGAGG	NGG
9	0.346	0.754	0.550	177	303	TCCCCTCTGGCTCACTGGC	CTGGCCCACATCGGAGGAGACTGCCAGTGAGC	NGT

2	0.412	0.627	0.519	178	34	ATCCAACTTCAAAGAGCCC	CTCACTGGCAGTCTCCTCCGATG TGGGCCAGGGCTCTTTGAA	NGG
2	0.362	0.588	0.475	179	22	ATCCAACTTCAAAGAGCCC	CCTCCGATGTGGGCCAGGGCTC TTTGAAGTTGGA	NGG
8	0.519	0.379	0.449	180	61	TCCAACTTCAAAGAGCCCT	CACTGGCAGTCTCCTCCGATGTG GGCCAGGGCTCTTTGAAG	NGC
9	0.267	0.574	0.421	181	281	TCCCCTCTGGCTCACTGGC	TCGGAGGAGACTGCCAGTGAGC CAGAGG	NGT
2	0.212	0.609	0.411	182	27	ATCCAACTTCAAAGAGCCC	CCTCCGATGTGGGCCAGGGCTC TTTGAAGTTG	NGG
9	0.522	0.287	0.405	183	293	TCCCCTCTGGCTCACTGGC	CCCTGGCCCACATCGGAGGAGA CTGCCAGTGAGCCA	NGT
2	0.424	0.376	0.400	184	5	ATCCAACTTCAAAGAGCCC	CTCCTCCGATGTGGGCCAGGGC TCTTTGAAGTTG	NGG
7	0.492	0.255	0.373	185	256	GGATTCCCCTCTGGCTCAC	CATCGGAGGAGACTGCCAGTGA GCCAGAGG	NGG
2	0.000	0.730	0.365	186	20	ATCCAACTTCAAAGAGCCC	GGCAGTCTCCTCCGATGTGGGC CAGGGCTCTTTGAAGTTGGA	NGG
5	0.250	0.452	0.351	187	203	GAGCCCTGGCCCACATCAG	CACTGGCAGTCTCCTCCGATGTG GG	NGG
8	0.517	0.185	0.351	188	47	TCCAACTTCAAAGAGCCCT	GGCAGTCTCCTCCGATGTGGGC CAGGGCTCTTTGAAGTTGG	NGC
2	0.130	0.569	0.349	189	24	ATCCAACTTCAAAGAGCCC	CTGGCAGTCTCCTCCGATGTGGG CCAGGGCTCTTTGAAGTTG	NGG
2	0.084	0.609	0.346	190	8	ATCCAACTTCAAAGAGCCC	CACTGGCAGTCTCCTCCGATGTG GGCCAGGGCTCTTTGAAGT	NGG
6	0.000	0.691	0.346	191	231	GCAGGCAGGATTCCCCTCT	GGAGGAGACTGCCAGTGAGCCA GAGGGGAAT	NGC
2	0.317	0.363	0.340	192	11	ATCCAACTTCAAAGAGCCC	CTGGCAGTCTCCTCCGATGTGGG CCAGGGCTCTTTGAA	NGG
10	0.206	0.442	0.324	193	216	TGCAGGCAGGATTCCCCTC	CTGGCCCACATCGGAGGAGACT GCCAGTGAGCCAGAGGGGAATC	NGG
9	0.288	0.348	0.318	194	290	TCCCCTCTGGCTCACTGGC	TCGGAGGAGACTGCCAGTGAGC CAGA	NGT
9	0.052	0.533	0.293	195	304	TCCCCTCTGGCTCACTGGC	GGCCCACATCGGAGGAGACTGC CAGTGAGC	NGT

7	0.059	0.524	0.292	196	249	GGATTCCCCTCTGGCTCAC	TCGGAGGAGACTGCCAGTGAGC CAGAGGGG	NGG
9	0.274	0.293	0.283	197	306	TCCCCTCTGGCTCACTGGC	CACATCGGAGGAGACTGCCAGT GAGC	NGT
2	0.076	0.479	0.277	198	1	ATCCAACTTCAAAGAGCCC	CAGTCTCCTCCGATGTGGGCCA GGGCTCTTTGAAGTTGGA	NGG
8	0.193	0.360	0.277	199	57	TCCAACTTCAAAGAGCCCT	CTCCTCCGATGTGGGCCAGGGC TCTTTGAAGTT	NGC
7	0.545	0.000	0.272	200	243	GGATTCCCCTCTGGCTCAC	CCCTGGCCCACATCGGAGGAGA CTGCCAGTGAGCCAGAGGGG	NGG
9	0.187	0.317	0.252	201	299	TCCCCTCTGGCTCACTGGC	TCGGAGGAGACTGCCAGTGAGC CA	NGT
8	0.077	0.418	0.248	202	55	TCCAACTTCAAAGAGCCCT	GGCAGTCTCCTCCGATGTGGGC CAGGGCTCTTTGAAGTT	NGC
7	0.191	0.294	0.243	203	251	GGATTCCCCTCTGGCTCAC	CCCTGGCCCACATCGGAGGAGA CTGCCAGTGAGCCAGAGG	NGG
9	0.276	0.206	0.241	204	302	TCCCCTCTGGCTCACTGGC	CCCTGGCCCACATCGGAGGAGA CTGCCAGTGAGC	NGT
8	0.260	0.192	0.226	205	60	TCCAACTTCAAAGAGCCCT	CGATGTGGGCCAGGGCTCTTTG AAGTT	NGC
2	0.123	0.325	0.224	206	4	ATCCAACTTCAAAGAGCCC	CACTGGCAGTCTCCTCCGATGTG GGCCAGGGCTCTTTGAAGTTG	NGG
8	0.052	0.383	0.217	207	74	TCCAACTTCAAAGAGCCCT	CCTCCGATGTGGGCCAGGGCTC TTTGA	NGC
8	0.432	0.000	0.216	208	46	TCCAACTTCAAAGAGCCCT	CTGGCAGTCTCCTCCGATGTGGG CCAGGGCTCTTTGAAGTTGG	NGC
2	0.142	0.267	0.205	209	25	ATCCAACTTCAAAGAGCCC	GGCAGTCTCCTCCGATGTGGGC CAGGGCTCTTTGAAGTTG	NGG
8	0.112	0.296	0.204	210	83	TCCAACTTCAAAGAGCCCT	CCTCCGATGTGGGCCAGGGCTC TTT	NGC
2	0.111	0.290	0.200	211	17	ATCCAACTTCAAAGAGCCC	CTCCTCCGATGTGGGCCAGGGC TCTTTG	NGG
7	0.280	0.117	0.198	212	254	GGATTCCCCTCTGGCTCAC	CCCACATCGGAGGAGACTGCCA GTGAGCCAGAGG	NGG
9	0.261	0.131	0.196	213	307	TCCCCTCTGGCTCACTGGC	CATCGGAGGAGACTGCCAGTGA GC	NGT

2	0.205	0.176	0.191	214	45	ATCCAACTTCAAAGAGCCC	CGATGTGGGCCAGGGCTCTTTG	NGG
2	0.245	0.129	0.187	215	38	ATCCAACTTCAAAGAGCCC	CTCCTCCGATGTGGGCCAGGGCTCTTTGAA	NGG
8	0.297	0.056	0.177	216	58	TCCAACTTCAAAGAGCCCT	CCTCCGATGTGGGCCAGGGCTCTTTGAAGTT	NGC
7	0.041	0.293	0.167	217	263	GGATTCCCCTCTGGCTCAC	CCCACATCGGAGGAGACTGCCAGTGAGCCAGA	NGG
7	0.101	0.229	0.165	218	250	GGATTCCCCTCTGGCTCAC	GGAGGAGACTGCCAGTGAGCCAGAGGGG	NGG
9	0.066	0.251	0.159	219	282	TCCCCTCTGGCTCACTGGC	GGAGGAGACTGCCAGTGAGCCAGAGG	NGT
8	0.019	0.294	0.156	220	75	TCCAACTTCAAAGAGCCCT	TCCGATGTGGGCCAGGGCTCTTTGA	NGC
9	0.087	0.225	0.156	221	291	TCCCCTCTGGCTCACTGGC	GGAGGAGACTGCCAGTGAGCCAGA	NGT
8	0.047	0.263	0.155	222	63	TCCAACTTCAAAGAGCCCT	GGCAGTCTCCTCCGATGTGGGCCAGGGCTCTTTGAAG	NGC
8	0.153	0.152	0.153	223	82	TCCAACTTCAAAGAGCCCT	CTCCTCCGATGTGGGCCAGGGCTCTTT	NGC
2	0.206	0.098	0.152	224	44	ATCCAACTTCAAAGAGCCC	TCCGATGTGGGCCAGGGCTCTTTG	NGG
10	0.110	0.179	0.144	225	212	TGCAGGCAGGATTCCCCTC	GGCCCACATCGGAGGAGACTGCCAGTGAGCCAGAGGGGAATCCTGC	NGG
6	0.275	0.000	0.137	226	227	GCAGGCAGGATTCCCCTCT	CCCACATCGGAGGAGACTGCCAGTGAGCCAGAGGGGAAT	NGC
8	0.076	0.192	0.134	227	64	TCCAACTTCAAAGAGCCCT	CAGTCTCCTCCGATGTGGGCCAGGGCTCTTTGAAG	NGC
8	0.191	0.073	0.132	228	71	TCCAACTTCAAAGAGCCCT	GGCAGTCTCCTCCGATGTGGGCCAGGGCTCTTTGA	NGC
7	0.137	0.123	0.130	229	257	GGATTCCCCTCTGGCTCAC	TCGGAGGAGACTGCCAGTGAGCCAGAGG	NGG
8	0.000	0.248	0.124	230	54	TCCAACTTCAAAGAGCCCT	CTGGCAGTCTCCTCCGATGTGGGCCAGGGCTCTTTGAAGTT	NGC

7	0.240	0.000	0.120	231	241	GGATTCCCCTCTGGCTCAC	GGAGGAGACTGCCAGTGAGCCAGAGGGGAA	NGG
9	-0.081	0.305	0.112	232	301	TCCCCTCTGGCTCACTGGC	AGCCCTGGCCCACATCGGAGGAGACTGCCAGTGAGC	NGT
9	0.120	0.103	0.112	233	273	TCCCCTCTGGCTCACTGGC	TCGGAGGAGACTGCCAGTGAGCCAGAGGGG	NGT
8	0.040	0.180	0.110	234	68	TCCAACTTCAAAGAGCCCT	CGATGTGGGCCAGGGCTCTTTGAAG	NGC
7	0.054	0.163	0.109	235	240	GGATTCCCCTCTGGCTCAC	TCGGAGGAGACTGCCAGTGAGCCAGAGGGGAA	NGG
2	0.079	0.132	0.105	236	30	ATCCAACTTCAAAGAGCCC	CCTCCGATGTGGGCCAGGGCTCTTTGAAGT	NGG
2	0.196	0.005	0.101	237	13	ATCCAACTTCAAAGAGCCC	CGATGTGGGCCAGGGCTCTTTGAA	NGG
2	0.116	0.075	0.096	238	32	ATCCAACTTCAAAGAGCCC	CGATGTGGGCCAGGGCTCTTTGAAGT	NGG
10	0.178	0.000	0.089	239	218	TGCAGGCAGGATTCCCCTC	CATCGGAGGAGACTGCCAGTGAGCCAGAGGGGAATC	NGG
9	0.062	0.111	0.087	240	308	TCCCCTCTGGCTCACTGGC	TCGGAGGAGACTGCCAGTGAGC	NGT
7	0.071	0.094	0.083	241	267	GGATTCCCCTCTGGCTCAC	GGAGGAGACTGCCAGTGAGCCAGA	NGG
2	-0.062	0.225	0.081	242	10	ATCCAACTTCAAAGAGCCC	CTCCTCCGATGTGGGCCAGGGCTCTTTGAAGT	NGG
8	0.162	0.000	0.081	243	80	TCCAACTTCAAAGAGCCCT	GGCAGTCTCCTCCGATGTGGGCCAGGGCTCTTT	NGC
2	0.092	0.062	0.077	244	29	ATCCAACTTCAAAGAGCCC	CAGTCTCCTCCGATGTGGGCCAGGGCTCTTTGAAGT	NGG
9	0.067	0.077	0.072	245	274	TCCCCTCTGGCTCACTGGC	GGAGGAGACTGCCAGTGAGCCAGAGGGG	NGT
2	0.102	0.042	0.072	246	12	ATCCAACTTCAAAGAGCCC	CCTCCGATGTGGGCCAGGGCTCTTTGAA	NGG
2	0.152	-0.012	0.070	247	43	ATCCAACTTCAAAGAGCCC	CCTCCGATGTGGGCCAGGGCTCTTTG	NGG
6	0.140	0.000	0.070	248	229	GCAGGCAGGATTCCCCTCT	CATCGGAGGAGACTGCCAGTGAGCCAGAGGGGAAT	NGC

2	0.043	0.096	0.069	249	16	ATCCAACTTCAAAGAGCCC	GGCAGTCTCCTCCGATGTGGGC CAGGGCTCTTTG	NGG
2	0.118	0.000	0.059	250	36	ATCCAACTTCAAAGAGCCC	GGCAGTCTCCTCCGATGTGGGC CAGGGCTCTTTGAA	NGG
8	0.104	0.000	0.052	251	78	TCCAACTTCAAAGAGCCCT	CACTGGCAGTCTCCTCCGATGTG GGCCAGGGCTCTTT	NGC
8	0.008	0.095	0.051	252	52	TCCAACTTCAAAGAGCCCT	CGATGTGGGCCAGGGCTCTTTG AAGTTGG	NGC
7	0.086	0.010	0.048	253	258	GGATTCCCCTCTGGCTCAC	GGAGGAGACTGCCAGTGAGCCA GAGG	NGG
2	0.044	0.048	0.046	254	7	ATCCAACTTCAAAGAGCCC	CGATGTGGGCCAGGGCTCTTTG AAGTTG	NGG
10	0.085	0.000	0.042	255	221	TGCAGGCAGGATTCCCCTC	TCGGAGGAGACTGCCAGTGAGC CAGAGGGGAATC	NGG
6	0.082	0.000	0.041	256	230	GCAGGCAGGATTCCCCTCT	TCGGAGGAGACTGCCAGTGAGC CAGAGGGGAAT	NGC
2	-0.009	0.071	0.031	257	28	ATCCAACTTCAAAGAGCCC	GGCAGTCTCCTCCGATGTGGGC CAGGGCTCTTTGAAGT	NGG
9	0.078	-0.017	0.031	258	309	TCCCCTCTGGCTCACTGGC	GGAGGAGACTGCCAGTGAGC	NGT
7	0.048	0.011	0.030	259	262	GGATTCCCCTCTGGCTCAC	GGCCCACATCGGAGGAGACTGC CAGTGAGCCAGA	NGG
7	0.055	0.000	0.027	260	260	GGATTCCCCTCTGGCTCAC	CCCTGGCCCACATCGGAGGAGA CTGCCAGTGAGCCAGA	NGG
2	0.050	0.005	0.027	261	39	ATCCAACTTCAAAGAGCCC	TCCGATGTGGGCCAGGGCTCTTT GAA	NGG
9	0.058	-0.010	0.024	262	300	TCCCCTCTGGCTCACTGGC	GGAGGAGACTGCCAGTGAGCCA	NGT
7	0.044	0.001	0.023	263	266	GGATTCCCCTCTGGCTCAC	TCGGAGGAGACTGCCAGTGAGC CAGA	NGG
7	0.067	-0.031	0.018	264	253	GGATTCCCCTCTGGCTCAC	GGCCCACATCGGAGGAGACTGC CAGTGAGCCAGAGG	NGG
8	0.082	-0.081	0.001	265	48	TCCAACTTCAAAGAGCCCT	CAGTCTCCTCCGATGTGGGCCA GGGCTCTTTGAAGTTGG	NGC
3	0.000	0.000	0.000	266	210	ATCTCACTTTGCTGCAGGC	CCCACATCGGAGGAGACTGCCA GTGAGCCAGAGGGGAATCCTGC	NGG

							CTGCAGCAAAGT	
4	0.000	0.000	0.000	267	18	CATCAAACAGGTGATAGAA	GGCAGTCTCCTCCGATGTGGGC CAGGGCTCTTTGAAGTTGGATCT GAGCCTTTCTATCACC	NGG
6	0.000	0.000	0.000	268	224	GCAGGCAGGATTCCCCTCT	AGCCCTGGCCCACATCGGAGGA GACTGCCAGTGAGCCAGAGGGG AATCC	NGC
6	0.000	0.000	0.000	269	226	GCAGGCAGGATTCCCCTCT	GGAGGAGACTGCCAGTGAGCCA GAGGGGAATCC	NGC
6	0.000	0.000	0.000	270	225	GCAGGCAGGATTCCCCTCT	GGCCCACATCGGAGGAGACTGC CAGTGAGCCAGAGGGGAATCC	NGC
6	0.000	0.000	0.000	271	223	GCAGGCAGGATTCCCCTCT	TCGGAGGAGACTGCCAGTGAGC CAGAGGGGAATCCTG	NGC
7	0.000	0.000	0.000	272	259	GGATTCCCCTCTGGCTCAC	AAAGAGCCCTGGCCCACATCGG AGGAGACTGCCAGTGAGCCAGA	NGG
7	0.000	0.000	0.000	273	242	GGATTCCCCTCTGGCTCAC	AAAGAGCCCTGGCCCACATCGG AGGAGACTGCCAGTGAGCCAGA GGGG	NGG
7	0.000	0.000	0.000	274	238	GGATTCCCCTCTGGCTCAC	CACATCGGAGGAGACTGCCAGT GAGCCAGAGGGGAA	NGG
7	0.000	0.000	0.000	275	234	GGATTCCCCTCTGGCTCAC	GGAGGAGACTGCCAGTGAGCCA GAGGGGAATC	NGG
9	0.000	0.000	0.000	276	283	TCCCCTCTGGCTCACTGGC	AGCCCTGGCCCACATCGGAGGA GACTGCCAGTGAGCCAGA	NGT
10	0.000	0.000	0.000	277	220	TGCAGGCAGGATTCCCCTC	AAAGAGCCCTGGCCCACATCGG AGGAGACTGCCAGTGAGCCAGA GGGGAATC	NGG
10	0.000	0.000	0.000	278	211	TGCAGGCAGGATTCCCCTC	GGAGGAGACTGCCAGTGAGCCA GAGGGGAATCCTGCCT	NGG
10	0.000	0.000	0.000	279	213	TGCAGGCAGGATTCCCCTC	GGCCCACATCGGAGGAGACTGC CAGTGAGCCAGAGGGGAATCCT	NGG
10	0.000	0.000	0.000	280	214	TGCAGGCAGGATTCCCCTC	TCGGAGGAGACTGCCAGTGAGC CAGAGGGGAATCCT	NGG
7	-0.078	0.072	-0.003	281	264	GGATTCCCCTCTGGCTCAC	CACATCGGAGGAGACTGCCAGT GAGCCAGA	NGG

2	-0.059	0.040	-0.010	282	2	ATCCAACTTCAAAGAGCCC	CGATGTGGGCCAGGGCTCTTTG AAGTTGGA	NGG
7	-0.031	-0.015	-0.023	283	265	GGATTCCCCTCTGGCTCAC	CATCGGAGGAGACTGCCAGTGA GCCAGA	NGG
2	0.020	-0.079	-0.030	284	15	ATCCAACTTCAAAGAGCCC	CTGGCAGTCTCCTCCGATGTGGG CCAGGGCTCTTTG	NGG
7	-0.019	-0.059	-0.039	285	261	GGATTCCCCTCTGGCTCAC	CTGGCCCACATCGGAGGAGACT GCCAGTGAGCCAGA	NGG
8	-0.159	0.078	-0.040	286	76	TCCAACTTCAAAGAGCCCT	CGATGTGGGCCAGGGCTCTTTG A	NGC
8	0.114	-0.196	-0.041	287	79	TCCAACTTCAAAGAGCCCT	CTGGCAGTCTCCTCCGATGTGGG CCAGGGCTCTTT	NGC
8	-0.001	-0.084	-0.042	288	73	TCCAACTTCAAAGAGCCCT	CTCCTCCGATGTGGGCCAGGGC TCTTTGA	NGC
8	-0.216	0.131	-0.043	289	69	TCCAACTTCAAAGAGCCCT	CACTGGCAGTCTCCTCCGATGTG GGCCAGGGCTCTTTGA	NGC
8	-0.037	-0.057	-0.047	290	72	TCCAACTTCAAAGAGCCCT	CAGTCTCCTCCGATGTGGGCCA GGGCTCTTTGA	NGC
8	0.057	-0.171	-0.057	291	85	TCCAACTTCAAAGAGCCCT	CGATGTGGGCCAGGGCTCTTT	NGC
2	0.068	-0.184	-0.058	292	37	ATCCAACTTCAAAGAGCCC	CAGTCTCCTCCGATGTGGGCCA GGGCTCTTTGAA	NGG
2	-0.107	-0.018	-0.063	293	42	ATCCAACTTCAAAGAGCCC	CAGTCTCCTCCGATGTGGGCCA GGGCTCTTTG	NGG
2	-0.031	-0.120	-0.076	294	41	ATCCAACTTCAAAGAGCCC	CACTGGCAGTCTCCTCCGATGTG GGCCAGGGCTCTTTG	NGG
8	0.164	-0.328	-0.082	295	77	TCCAACTTCAAAGAGCCCT	CTCACTGGCAGTCTCCTCCGATG TGGGCCAGGGCTCTTT	NGC
8	-0.272	0.076	-0.098	296	56	TCCAACTTCAAAGAGCCCT	CAGTCTCCTCCGATGTGGGCCA GGGCTCTTTGAAGTT	NGC
10	-0.141	-0.211	-0.176	297	222	TGCAGGCAGGATTCCCCTC	GGAGGAGACTGCCAGTGAGCCA GAGGGGAATC	NGG
2	0.000	-0.404	-0.202	298	26	ATCCAACTTCAAAGAGCCC	CAGTCTCCTCCGATGTGGGCCA GGGCTCTTTGAAGTTG	NGG
8	-0.166	-0.254	-0.210	299	81	TCCAACTTCAAAGAGCCCT	CAGTCTCCTCCGATGTGGGCCA GGGCTCTTT	NGC

2	-0.166	-0.303	-0.235	300	9	ATCCAACTTCAAAGAGCCC	CTGGCAGTCTCCTCCGATGTGGG CCAGGGCTCTTTGAAGT	NGG
2	-0.022	-0.477	-0.250	301	40	ATCCAACTTCAAAGAGCCC	CTCACTGGCAGTCTCCTCCGATG TGGGCCAGGGCTCTTTG	NGG
2	-0.173	-0.327	-0.250	302	35	ATCCAACTTCAAAGAGCCC	CACTGGCAGTCTCCTCCGATGTG GGCCAGGGCTCTTTGAA	NGG
2	-0.258	-0.362	-0.310	303	19	ATCCAACTTCAAAGAGCCC	CTGGCAGTCTCCTCCGATGTGGG CCAGGGCTCTTTGAAGTTGGA	NGG
8	-0.300	-0.370	-0.335	304	84	TCCAACTTCAAAGAGCCCT	TCCGATGTGGGCCAGGGCTCTTT	NGC
6	-0.342	-0.342	-0.342	305	228	GCAGGCAGGATTCCCCTCT	CACATCGGAGGAGACTGCCAGT GAGCCAGAGGGGAAT	NGC
10	-0.307	-0.463	-0.385	306	219	TGCAGGCAGGATTCCCCTC	GGAGGAGACTGCCAGTGAGCCA GAGGGGAATCCTGC	NGG
10	-0.503	-0.503	-0.503	307	217	TGCAGGCAGGATTCCCCTC	CACATCGGAGGAGACTGCCAGT GAGCCAGAGGGGAATC	NGG
7	-0.516	-0.598	-0.557	308	233	GGATTCCCCTCTGGCTCAC	TCGGAGGAGACTGCCAGTGAGC CAGAGGGGAATC	NGG
8	-1.045	-0.517	-0.781	309	53	TCCAACTTCAAAGAGCCCT	CTCACTGGCAGTCTCCTCCGATG TGGGCCAGGGCTCTTTGAAGTT	NGC

Table 6. Comparison of efficiencies and preciseness of genome editing methods when genome editing tools were delivered using hydrodynamic injections in a mouse model of hereditary tyrosinemia

Genome editor	Cas9	Cas9	ABE	ABE	PE2	PE3
Reference	Ref. 8 (Yin et al., 2014)	Ref. 10 (Shin et al., 2018)	Ref. 11 (Song et al., 2020)	Ref. 11 (Song et al., 2020)	The current study	The current study
Editing strategy	Homology-directed repair	Microhomology-mediated end joining	Base editing	Base editing	Prime editing	Prime editing
Vector	Cas9, sgRNA plasmid, ssDNA donor	Cas9, sgRNA plasmid, ssDNA donor	ABE6.3, sgRNA plasmid	RA6.3, sgRNA plasmid	PE2, pegRNA plasmid	PE2, pegRNA, nick sgRNA plasmid
Delivery	Hydrodymic injection	Hydrodymic injection	Hydrodynamic injection	Hydrodynamic injection	Hydrodymic injection	Hydrodymic injection
Initial efficiency	0.40% Fah+cells (day6) / Not detectable in DNA (day6)	No data	1% Fah+cells (day7) / <0.1% A9 to G editing (day7)	3.5% Fah+cells (day7) / ~0.3% A9 to G editing (day7)	No data	partial liver tissue. 0.07% Fah+cells (day7) / partial liver tissue. Not detectable in DNA (day7)
Efficiency after withdrawal of NTBC	9.3% A to G editing (30 days without NTBC)	5.2% A to G editing (30 days without NTBC)	9.5% A9 to G editing (32 days without NTBC)	No data	4.0% A to G editing (60 days after the initial NTBC withdrawal)	11.5% A to G editing (40 days without NTBC)

	33.5% Fah+cells (30 days without NTBC)	64.4% Fah+cells (30 days without NTBC)	Widespread Fah+ cells. Not quantified (32 days without NTBC)		32.8% Fah+cells (60 days after the initial NTBC withdrawal)	61% Fah+cells (40 days without NTBC)
Indels	~26% (30 days without NTBC)	No data	0.05% (32 days without NTBC)	No data	<0.1% (60 days after the initial NTBC withdrawal)	0.78% (40 days without NTBC)
Off-target effects (substitutions)	No data	No data	<0.1% A to G conversion	No data	<0.1%	<0.1%
Off-target effects (indels)	<0.3%	No data	No data	No data	<0.1%	<0.1%
Bystander effects	No data	No data	0.1% A6 to G editing (day7) 1.9% A6 to G editing (32 days without NTBC)	0.3% A6 to G editing (day7)	Not detectable (60 days after the initial NTBC withdrawal)	Not detectable (40 days without NTBC)
Functional recovery	recovered body weight	recovered body weight	recovered body weight	No data	recovered body weight	recovered body weight

Table 7. Oligonucleotides used in the study

Purpose	FP/RP	Primer sequence (5' to 3') FP: forward primer, RP: reverse primer
pLenti-NG-PE2-BSD cloning	FP	GTTTGCCGCCAGAACACAG GACCGAGGTGCAGACAGG GACTCTGGAGGATCTAGCGGA GAAGCCGCCTGTCTGCAC
	RP	TCCGCTAGATCCTCCAGAGTCGCCTCCCAGCTGAGA CAGAGAGAAGTTTGTTGCGC
Oligonucelotide pool amplification	FP	TTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC
	RP	GAGTAAGCTGACCGCTGAAGTACAAGTGGTAGAGTAGAGATCGCGGTACGCCAAGCT
sgRNA scaffold amplification	FP	TATTCTTCAAGTAAGTATCGTCTCTGTTTTAGAGCTAGAAATAG
	RP	TCGTTAAATAGAACACTTCGTCTCTGCACCGACTCGGTGCCACT
Amplification of pegRNA-barcode-target sequence coding region (1st PCR)	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGAAAAAGTGGCACCGAGTCG ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTTGAAAAAGTGGCACCGAGTCG ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGCTTGAAAAAGTGGCACCGAGTCG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAATACTGCCATTTGTCTCAAGA GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTAATACTGCCATTTGTCTCAAGA GTGACTGGAGTTCAGACGTGTGTCTTCCGATCTAATACTGCCATTTGTCTCAAGA
Illuimina indexing (2nd PCR)	FP	AATGATACGGCGACCACCGAGATCTACAC (8bp barcode) ACACTCTTTCCCTACACGAC
	RP	CAAGCAGAAGACGGCATACGAGAT (8bp barcode) GTGACTGGAGTTCAGACGTGT

px601-PE2 cloning	FP	AGCAGAGCTCTCTGGCTAACTACCGGTGAGAGCCGCCACCATGAAAC
		CGGCCTCTGATCGAGACAAA
		AATCAACCTCTGGATTACAAAAT
		TTTGTCTCGATCAGAGGCCG
	RP	ATTTTGTAATCCAGAGGTTGATTTGGTGATGATGACCGGTTAGAC
		CTCCATCACTAGGGGTTCCTGCGGCCGCCCATAGAGCCCACCGCAT
gN19-nick sgRNA cloning	FP	CACCGCCTAAGAACAGAACATCAG
	RP	AAACCTGATGTTCTGTTCTTAGGC
px552-pegRNA-CAG-mCherry cloning	FP	TTGTGGAAAGGACGAAACACCGGATGGTCCTCATGAACGACGTTTTAGAGCTAGAAATA
		TCGTTCATGAGGACCATTTTTTTTGACATTGATTATT
	RP	ATGGTCCTCATGAACGACTGGAGCGGTAATGCCGCACCGACTCGGTGCCA
		ACCCCGTAATTGATTACTAT
TS-PE2-N cloning	FP	GCAGAGCTCTCTGGCTAACTACCGGTGCCACCATGAAACGGACAG
		GTAAGTATCAAGGTTACAAGACAG
	RP	CTGTCTTGTAACCTTGATACTTACGTACTTGGCGGTAGCCTT
		CATCACTAGGGGTTCCTGCG
TS-PE2-C-WPRE cloning	FP	TCACTAGGGGTTCCTGCGGCCTCTAGACTTGTCGAGACAGAGAAGAC
		CTTTGCCTTTCTCTCCACAGTTCTTCTACAGCAACATCATGAAC
		AATCAACCTCTGGATTACAAAAT
		TTTGGGCCGCCTCCCCGCCTGTGCCTTCTAGTTGCCAG
		CTGTGGAGAGAAAGGCAAAG
	RP	ATTTTGTAATCCAGAGGTTGATTTGGTGATGATGACCGGTTAGAC
		GCGGGGAGGCGGCCCAAA
		CTCCATCACTAGGGGTTCCTGCGGCCGCCCATAGAGCCCACCGCAT

Fah targeted deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTTCTTCGTAGGCCCTG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGGCCTCCTGTCCCATAC
Rpe65 targeted deep sequencing (1st PCR)	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTTTCTTCGTAGGCCCTG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGGCCTCCTGTCCCATAC
Illuimina indexing for deep sequencing (2nd PCR)	FP	AATGATACGGCGACCACCGAGATCTACAC (8bp barcode) ACACTCTTTCCCTACACGAC
	RP	CAAGCAGAAGACGGCATACGAGAT (8bp barcode) GTGACTGGAGTTCAGACGTGT
FAH RT-PCR	FP	TTCTACTCTTCTCGGCAGCA
FAH qPCR	FP	AGAGCCAATCCCCATTTCCA
FAH RT-PCR&qPCR	RP	CGGGGAGATTGTGGTTCCAA
Fah-OT1 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTTCAAGGCCATCCTCAGCT
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCCCGCCAGTCACATGTATG
Fah-OT2 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATAGACCCATCTGTCCCGC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCCACACATACCAGTACACCA
Fah-OT3 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCCCGCCTTGGGTCCTAA
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCATTAGAACACCCACATGCTT
Fah-OT4 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTTGCTAGTCCATTGTAGGTG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGGACCTGCAGCCAATTTTTA
Fah-OT5 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGTGAAACTTCGTCTCCAG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGCTTGTGTTAAATGCCATGT
Fah-OT6 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGACCCCTGGCAGTGGAG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCCGTTTCCCTCATTAGCG

Fah-OT7 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATCAAAACTGAGGGAAAAGCCC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCACAATGCTATGACTCTGGGC
Fah-OT8 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGTGGTCTTTTGCCCTGG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCCATCTTGCTTCTCTTCTTT
Fah-OT9 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATTCACTGGGAAGGAAGGTC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCCTGAGAAGGAAAAGATGAA
Fah-OT10 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCTGTACCACTGAAGAAAC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTAGGAAAGTTGGAATGGTGCA
Fah-OT11 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTTAGACACTGGAGGCCAG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATGTCAGAGAGGCCATGTTC
Fah-OT12 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTGAGGAGTGGAATTCTTGG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACCTGCTACTCAGATGTCCTG
Fah-OT13 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTTTGTGTGGTAAAAGTGC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTTGATGAGATGTGAATTCATTCTGG
Fah-OT14 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTAGTGCCTTGGATTGCTTCCAT
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTGCACCATCAATAAGGTCCCA
Fah-OT15 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGGAAGATCGGTCTCCATATG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGTTTTACGATGTCAGTTGAGGC
Rpe65-OT1 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGACAAATAACAAATAGGCAC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGCCCTCCTTGAAGTCAAAC
Rpe65-OT2 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTGTGTATTGTGCATGGTGC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCTCCTAGCATTGGAATGGA

Rpe65-OT3 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTGGCAAAGTTCAGGGTT
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAAAGCCTGCAGGACGGATT
Rpe65-OT4 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCCTTCCTTGTTGCAAAAAGC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACCCCATGCTTGACACTGAG
Rpe65-OT5 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTCCAGGGAGCAGGAGTTG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCAGCATGTCACTCTTGGGT
Rpe65-OT6 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAATGCCCAACAGTGAACCCA
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACACACCCTCTCCTCTCTCC
Rpe65-OT7 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCAGTGTCACTGTCCCCAAA
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTGAAGGGCAACTCAGTCCC
Rpe65-OT8 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCTGTCACTGAACACAGCC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATGAGATTGGAGAAGGAGCTG
Rpe65-OT9 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTAAGGTTGGGGGCTAAC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATTTGGGGTGATATTGAGGC
Rpe65-OT10 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGGGGAGATCGACTTCTAAT
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGCTTCCTTCTCTGCCTG
Rpe65-OT11 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTCTCCATGAGGCCACAGA
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGCAGAAACCAAGCACCAT
Rpe65-OT12 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCAAGTTAAGAAATCCCC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATCCTAGAAGGACTTCTGTG
Rpe65-OT13 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGGTGGAAAAGACAACTGGG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCAACGTCTGCTTTGCCTGGA

Rpe65-OT14 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTTCTCTGACAGTGACAC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCCAACCCATGCCTTTCCA
Rpe65-OT15 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCATCCTGTTTCTCACCTGGG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTTCTGTGGGGTCTTACTTGG
Rpe65-OT16 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCAAGGCCACAGAATCCCCA
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTCCATGGCTCCAATGTC
Rpe65-OT17 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACTCTGCATTTTCATTGATAG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGGATCTTGGAGAAGAATCTAT
Rpe65-OT18 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTGGATTTGTGAATGGA
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAAAACTCTAAGAGACAGAAGG
Rpe65-OT19 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTGAGGGCTTATTCTGCTC
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCAATTTTGGAAAAAGGGAGC
Rpe65-OT20 deep sequencing	FP	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGTGTCTTGTGAGTGGTGG
	RP	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCATCCCGGTCTTCCCCAGG

ABSTRACT(IN KOREAN)

프라임 편집기를 활용한 생체내 유전자 교정에 따른 유전성 간질환 및 유전성 안질환 생쥐모델 치료

<지도교수 김 형 범 >

연세대학교 대학원 의과학과

장 혜 원

이중가닥 절단(DSB)을 일으키지 않고 공여 DNA(Donor DNA) 없이도 특이적으로 유전자를 교정할 수 있는 기법인 프라임 편집(Prime editing)의 성능은 인간의 유전질환을 갖는 동물 모델에서 검증될 필요가 있었다. 본 연구에서는 유전성 타이로신혈증 및 유전성 레버선천성흑암시(LCA)에 대한 병원성 돌연변이에 대해 수백개의 프라임 편집 가이드 RNA(Prime editing guide RNA)를 고안하여 렌티 바이러스 라이브러리 형태로 제작하였고 이를 감염시킨 세포주를 이용한 소규모의 라이브러리 실험을 통해 최적의 프라임 편집 가이드 RNA를 선정하였다. 이후, 유전성 타이로신혈증 생쥐모델에 유체역학 주입을 통해 프라임 편집기(Prime editor) 2 또는 3을 전달하여 병원성 돌연변이 교정 효율을 확인하였고, 레버선천성흑암시를 가진 생쥐에는 아데노 관련 바이러스 벡터(Adeno-associated virus)로 프라임 편집기 2를 전달하여 유전자 교정 효율을 평가하였다. 생체 내에서 프라임 편집기는 질병 유발 돌연변이를 매우 정교한 방법으로 수정하였고, 오프-타겟 효과의 유발 없이 생쥐의 질병 표현형을 개선하였다. 따라서, 본 연구는 프라임 편집이 유전병 치료에 유망한 접근법이 될 것임을 시사하고 있다.

핵심되는 말: 프라임 편집, 유전자 교정, 유전성 타이로신혈증, 레버선천성흑암시, 오프-타겟 효과

PUBLICATION LIST

1. Song M, Kim HK, Lee S, Kim Y, Seo S-Y, Park J, Choi JW, **Jang H**, Shin, JH et al. Sequence-specific prediction of the efficiencies of adenine and cytosine base editors. *Nature Biotechnology* (2020)