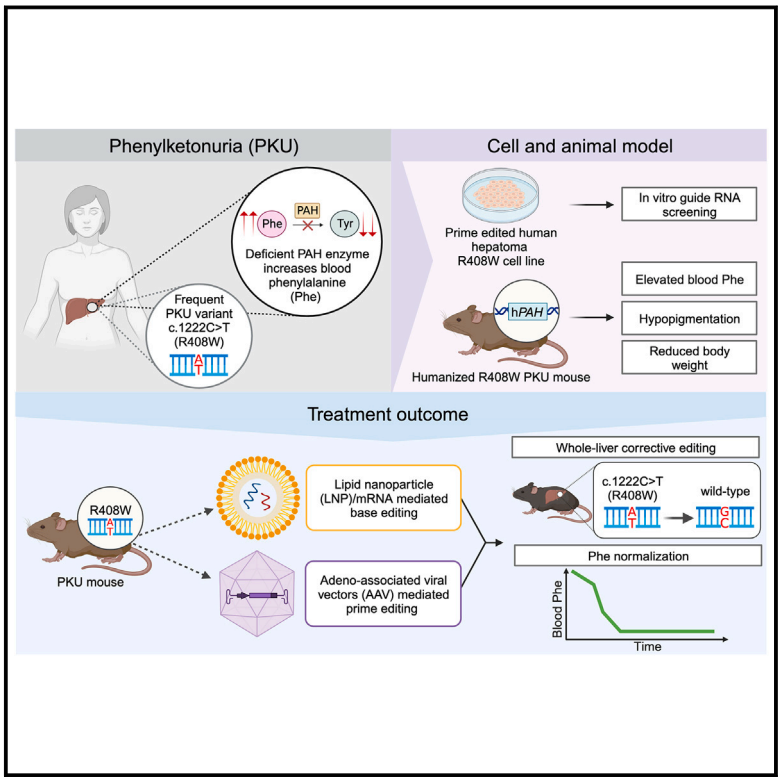
Efficient *in vivo* prime editing corrects the most frequent phenylketonuria variant, associated with high unmet medical need

Graphical abstract



Authors

Dominique L. Brooks, Madelynn N. Whittaker, Ping Qu, Kiran Musunuru, Rebecca C. Ahrens-Nicklas, Xiao Wang

Correspondence

kiranmusunuru@gmail.com

The *PAH* c.1222C>T variant is the most frequent variant reported in phenylketonuria (PKU) individuals. Here, the authors show these individuals have poor metabolic control, and they use prime editing, delivered by adeno-associated viral vectors, to directly correct the variant in the liver in humanized PKU mice and definitively treat the disease.

Brooks et al., 2023, The American Journal of Human Genetics *110*, 2003–2014
December 7, 2023
https://doi.org/10.1016/j.ajhg.2023.10.005

CellPress

Efficient *in vivo* prime editing corrects the most frequent phenylketonuria variant, associated with high unmet medical need

Dominique L. Brooks,[1,2,3] Madelynn N. Whittaker,[1,2,3,4] Ping Qu,[1,2,3] Kiran Musunuru,[1,2,3,7,*] Rebecca C. Ahrens-Nicklas,[3,5,6,7] and Xiao Wang[1,2,3,7]

Summary

The c.1222C>T (p.Arg408Trp) variant in the phenylalanine hydroxylase gene (*PAH*) is the most frequent cause of phenylketonuria (PKU), the most common inborn error of metabolism. This autosomal-recessive disorder is characterized by accumulation of blood phenylalanine (Phe) to neurotoxic levels. Using real-world data, we observed that despite dietary and medical interventions, most PKU individuals harboring at least one c.1222C>T variant experience chronic, severe Phe elevations and do not comply with Phe monitoring guidelines. Motivated by these findings, we generated an edited c.1222C>T hepatocyte cell line and humanized c.1222C>T mouse models, with which we demonstrated efficient *in vitro* and *in vivo* correction of the variant with prime editing. Delivery via adeno-associated viral (AAV) vectors reproducibly achieved complete normalization of blood Phe levels in PKU mice, with up to 52% whole-liver corrective *PAH* editing. These studies validate a strategy involving prime editing as a potential treatment for a large proportion of individuals with PKU.

Introduction

Phenylketonuria (PKU [MIM: 261600]) is an autosomal-recessive disorder caused by mutations in the gene encoding phenylalanine hydroxylase (PAH), resulting in the accumulation of phenylalanine (Phe) to neurotoxic levels because of the inability to convert Phe into tyrosine. Untreated PKU can manifest in high blood Phe levels of >1200 μmol/L, whereas current treatment guidelines state that blood Phe levels should ideally be maintained in the range of 120–360 μmol/L.[1] The existing treatment options range from a strict low-Phe diet to an oral medication (sapropterin, a cofactor of PAH) to an injectable enzyme-substitution therapy (with pegvaliase). All are chronic, everyday interventions to which affected individuals must adhere for their lifetimes if they are to prevent cognitive impairment and a range of neuropsychiatric complications. The burden of intensive Phe control and monitoring faced by individuals with PKU and their families is profound, especially with treatments that can be unpalatable and cost-prohibitive. Lack of adherence to the guidelines on the part of many individuals with PKU is not only a serious risk, but also to be expected. Studies have shown that more than 70% of adults with PKU are noncompliant with therapy.[2] As such, durable and, ideally, curative therapies would be needed to optimally address the medical needs of individuals with PKU.

More than 1,000 *PAH* variants have been cataloged in PKU individuals.[3] These vary in their consequences for PAH activity, from having little or no effect to eliminating PAH activity completely. The most frequently occurring pathogenic *PAH* variant worldwide is the c.1222C>T (p.Arg408Trp) variant (GenBank: NM_000277.3) (MIM: 612349), which is most prevalent in European populations (e.g., present in 98.9% of PKU individuals in Estonia, 89.2% in Poland, 75.7% in Russia, 43.6% in Sweden, and 35.7% in Germany), Australia (34.7%), and the United States (32.9%).[3] Individuals homozygous for this variant have the most severe form of PKU and do not respond to sapropterin,[4] limiting their treatment options. Individuals with the c.1222C>T variant do respond to pegvaliase, a once-daily injected bacterial-derived enzyme that directly catabolizes Phe, albeit with a substantial risk of anaphylaxis. However, in clinical trials, pegvaliase had only a mean Phe reduction of 51% at one year after initiation (from 1,233 μmol/L to 565 μmol/L), meaning that a large proportion of the individuals did not achieve the guideline-directed Phe goal of <360 μmol/L.[5]

A safe one-time therapy that permanently normalizes blood Phe levels would be transformative for the care of individuals with PKU. Previous studies have suggested that reverting only ≈10% of the *PAH* alleles to wild-type in the liver, where *PAH* is largely expressed, would achieve this goal.[6–9] *In vivo* CRISPR editing is an emerging therapeutic approach that makes changes in DNA sequence in

[1]Cardiovascular Institute, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA; [2]Division of Cardiovascular Medicine, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA; [3]Department of Genetics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA; [4]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA; [5]Metabolic Disease Program, Division of Human Genetics, Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; [6]Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA
[7]These authors contributed equally
*Correspondence: kiranmusunuru@gmail.com
https://doi.org/10.1016/j.ajhg.2023.10.005.

organs such as the liver in an individual's body. Nuclease editors have been demonstrated to be effective in non-human primates and in human individuals, but their use is limited to disruption or deletion of genes or regulatory elements in the genome.[10] Base editors have the ability to directly and precisely correct disease-causing variants, but their utility is highly dependent on target-sequence context and is generally restricted to single-nucleotide transitions (C→T, A→G).[11,12] Prime editors have the precision of base editing, with the advantage that they can make all possible single-nucleotide changes, can correct indels, and have greater target-site flexibility.[13]

A prime editor comprises a Cas9 variant that nicks one DNA strand at the target site, a reverse transcriptase enzyme fused to Cas9, and a guide RNA (called prime-editing guide RNA, or pegRNA) that in complex with Cas9 specifies the target site (via a matching protospacer sequence at the DNA target site, just upstream of the Cas9 protein's preferred protospacer-adjacent motif [PAM], typically an NGG sequence) and has a 3′ extension with (1) a primer-binding site (PBS) that hybridizes with the nicked DNA strand upstream of the nick and (2) a reverse transcriptase template (RTT) that serves as a template for the reverse transcriptase to build a DNA flap anchored to the nicked DNA strand at the nick site.[13] The desired edit is encoded within the RTT sequence and is thus incorporated into the synthesized DNA flap. The DNA flap competes with the original sequence on the nicked DNA strand downstream of the nick for incorporation into the final repaired double-strand DNA product. Mismatch repair (MMR) favors the original sequence by acting against the desired prime-editing outcome.[14] Successful prime editing is facilitated by: (1) using an optimized prime-editor protein (e.g., PEmax); (2) using an engineered pegRNA with a stabilizing structured RNA motif at its 3′ end; (3) inhibiting MMR, e.g., by adding extra mismatches into the RTT, namely synonymous edits that will not compromise the correction of a disease-causing variant; (4) using a nicking guide RNA (ngRNA) that directs Cas9 to nick the DNA strand opposite the DNA flap, instigating a nick repair process that favors incorporation of the DNA flap into the final product; and (5) using an ngRNA whose protospacer sequence overlaps with the site of the edit(s).[13–15]

In this work, we sought to define the degree of adherence to Phe treatment and monitoring goals by PKU individuals with the *PAH* c.1222C>T variant. We used real-world data from a PKU cohort managed at a metabolic specialty clinic at an academic medical center. Upon establishing a high degree of unmet medical need among these individuals, we created an *in vitro* cellular model and *in vivo* humanized PKU mouse models in order to develop and validate a corrective prime-editing therapeutic for the *PAH* c.1222C>T variant. Using adeno-associated viral (AAV) vectors for delivery of a lead prime-editing candidate, we demonstrated efficient corrective editing and normalization of blood Phe levels in homozygous and compound-heterozygous mice.

Subjects and methods

Enrollment

PKU-affected individuals born prior to January 1, 2021, and followed in the Children's Hospital of Philadelphia (CHOP) Metabolism Clinic were considered for inclusion in the study. All available genotype information was reviewed, and subjects who were either homozygous or compound-heterozygous for the *PAH* c.1222C>T variant were included in the final study cohort.

This study was reviewed by the CHOP institutional review board (IRB 23–021115) and found to meet exemption criteria for human subjects research per 45 CFR 46.104(d) 4(iii). In addition, a waiver of HIPAA authorization was granted per 45 CFR 164.512(i)(2)(ii) for accessing subjects' *PAH* genotypes and longitudinal Phe levels.

Vectors

The following prime-editor-expressing plasmids were obtained from Addgene as gifts from Dr. David Liu: pCMV-PEmax-P2A-hMLH1dn (Addgene plasmid #174828; http://n2t.net/addgene:174828; RRID: Addgene_174828), pU6-tevopreq1-GG-acceptor (Addgene plasmid #174038; http://n2t.net/addgene:174038; http://n2t.net/addgene:174038; RRID: Addgene_174038), v3em-Nterm-PE2max (Addgene plasmid #198734; http://n2t.net/addgene:198734; RRID: Addgene_198734), and v3em-Cterm-PE2-max-ΔRNaseH-dualU6 (Addgene plasmid #198735; http://n2t.net/addgene:198735; RRID: Addgene_ 198735).[14–16] Each pegRNA was expressed from pU6-tevopreq1-GG-acceptor, with the spacer sequence (plus an additional 5′ G to facilitate U6 expression) and PBS and RTT sequences indicated in Table S1 cloned into the plasmid. Each ngRNA was expressed from pGuide (Addgene plasmid #64711; http://n2t.net/addgene:64711; RRID: Addgene_64711); the spacer sequence (plus an additional 5′ G to facilitate U6 expression) indicated in Table S2 was cloned into the plasmid. For the lead pegRNA/ngRNA combination (P56/N19), the sequences were cloned into the v3em-Cterm-PE2max-ΔRNaseH-dualU6 plasmid (Table S3). This last vector and the v3em-Nterm-PE2max vector were transferred to PackGene Biotech to generate AAV vectors with the AAV8 serotype on a fee-for-service basis. Any of the newly generated plasmids is readily available from the authors via a "transfer of research material" agreement with the University of Pennsylvania.

Culture and transfection of HuH-7 cells

The methods were similar to those previously described.[9] HuH-7 cells were maintained in Dulbecco's modified Eagle's medium (containing 4 mM L-glutamine and 1 g/L glucose) with 10% fetal bovine serum and 1% penicillin-streptomycin at 37°C with 5% CO_2. HuH-7 cells were seeded on 6-well plates (Corning) at 3.5×10^5 cells per well. At 16–24 h after seeding, cells were transfected at approximately 80%–90% confluency with 9 μL TransIT-LT1 Transfection Reagent (MIR2300, Mirus), 1,500 ng prime-editor plasmid (pCMV-PEmax-P2A-hMLH1dn), 750 ng pegRNA plasmid, and 750 ng ngRNA plasmid per well according to the manufacturer's instructions. Cells were cultured for 72 h after transfection, and then media were removed, cells were washed with 1× DPBS (Corning), and genomic DNA was isolated with the DNeasy Blood and Tissue Kit (QIAGEN) according to the manufacturer's instructions.

Generation of *PAH* c.1222C>T homozygous HuH-7 cells

The methods were similar to those previously described.[9] HuH-7 cells in a well of a 6-well plate were transfected with 9 μL

TransIT-LT1 Transfection Reagent, 1,500 ng prime editor plasmid (pCMV-PEmax-P2A-hMLH1dn), 750 ng pegRNA plasmid, and 750 ng ngRNA plasmid. Cells were dissociated with trypsin 48 h after transfection and replated onto 10 cm plates (5,000 cells/plate) with conditioned medium to facilitate recovery, and genomic DNA was isolated from the remainder of the cells as a pool for PCR and Sanger sequencing of the *PAH* c.1222C>T site. Single cells were permitted to expand for 7–14 days so that clonal populations could be established. Colonies were manually picked and replated into individual wells of a 96-well plate. Genomic DNA was isolated from individual clones, and PCR and Sanger sequencing were performed for identification of c.1222C>T homozygous HuH-7 clones. One representative clone was expanded for use in subsequent studies. This c.1222C>T homozygous HuH-7 cell line is readily available from the authors via a "transfer of research material" agreement with the University of Pennsylvania.

Generation of humanized PKU mouse models

Humanized mice with either the *PAH* c.1222C>T (p.Arg408Trp) variant (hereafter referred to as R408W mice) or the *PAH* c.1066−11G>A variant (hereafter referred to as 1066 mice) were generated by Ozgene on a fee-for-service basis. The schemes are outlined in Figures S1, S2, and S3. In brief, the humanized c.1222C>T allele was created by knocking in human exon 12 and approximately 500 bp of genomic DNA both 5′ and 3′ of the knockin exon in place of the corresponding genomic region of murine *Pah* via homologous recombination in C57BL/6 mouse embryonic stem cells. The knockin exon included the c.1222C>T variant, introduced by alteration of the arginine codon CGG to the tryptophan codon TGG. The *neo* antibiotic selection cassette used for facilitating knockin was subsequently removed by Cre recombinase. The humanized c.1066−11G>A allele was created by knockin of genomic DNA of human *PAH* from approximately 500 bp upstream of exon 10 through approximately 600 bp downstream of exon 12 in place of the corresponding genomic region of murine *Pah* via homologous recombination in C57BL/6 mouse embryonic stem cells. The knockin human region included the c.1066−11G>A variant, introduced by alteration of the corresponding position in intron 11 from a G to an A. The *neo* antibiotic selection cassette used for facilitating knockin was subsequently removed by Cre recombinase. After the generation of chimeric mice by blastocyst injection, the humanized alleles were bred to heterozygosity, homozygosity, or compound heterozygosity through crosses of the founder mice and subsequent generations with wild-type mice or with each other. Each of the humanized mouse models is readily available from the authors via a "transfer of research material" agreement with the University of Pennsylvania.

Mouse studies

The methods were similar to those previously described.[9] All procedures used in mouse studies were approved by the Institutional Animal Care and Use Committee at the University of Pennsylvania (protocol #805887), where the studies were performed, and were consistent with local, state, and federal regulations, including the National Institutes of Health Guide for the Care and Use of Laboratory Animals, as applicable. Mice were maintained on a 12 h light/12 h dark cycle in a temperature range of 65°F–75°F and a humidity range of 40%–60% and were fed *ad libitum* with a chow diet (LabDiet, Laboratory Autoclavable Rodent Diet 5010). Homozygous and compound-heterozygous humanized PKU mice, as well as heterozygous humanized non-PKU mice, were generated as littermates or colony mates via timed breeding; in some cases wild-type C57BL/6J mice (stock no. 000664; RRID:IMSR_JAX:000664) from The Jackson Laboratory were used. Genotyping was performed via PCR amplification from genomic DNA samples (prepared from clipped tails and ears) followed by Sanger sequencing or by next-generation sequencing. Age-matched female and male colony mates were used for experiments at 6 weeks of age or 10 weeks of age; animals were randomly assigned to various experimental groups when applicable, and collection and analysis of data were performed in a blinded fashion when possible. AAV vectors were generated from the v3em-Nterm-PE2max and v3em-Cterm-PE2max-ΔRNaseH-dualU6 plamids, as described above. The AAV vectors were administered as 1:1 mixes to the mice at total doses of 4×10^{11} vg, 8×10^{11} vg, or 1×10^{12} vg via retro-orbital injection under anesthesia with 1%–2% inhaled isoflurane. The AAV dose range was chosen on the basis of a prior study with similar vectors;[16] the exact doses were determined by the sizes of aliquots provided by PackGene. Mice were euthanized at about 6 weeks after AAV treatment. Euthanasia in all instances was achieved via terminal inhalation of carbon dioxide followed by secondary euthanasia through cervical dislocation or decapitation, consistent with the 2020 American Veterinary Medical Association Guidelines on Euthanasia. For all treated mice, eight liver samples (two from each lobe) were obtained on necropsy and processed with the DNeasy Blood and Tissue Kit (QIAGEN) as per the manufacturer's instructions for isolation of genomic DNA. Next-generation sequencing results from the liver samples were averaged to provide quantification of whole-liver editing. Blood samples were collected via the tail tip at various timepoints (pre-treatment to establish the baseline level, week 1, week 2, week 3, week 4, and week 5) in the early afternoon to account for diurnal variation in blood phenylalanine levels.

Measurement of blood analytes

The methods were similar to those previously described.[9] The blood phenylalanine levels were measured by an enzymatic method with the Phenylalanine Assay Kit (MAK005, Millipore Sigma) according to the manufacturers' instructions. Plasma samples were deproteinized with a 10 kDa MWCO spin filter (CLS431478-25EA, Millipore Sigma) and pre-treated with 5 μL of tyrosinase for 10 min at room temperature prior to the start of the assay. Infrequently, a sample was lost during preparation, and the value for that sample was omitted from the final dataset for that reason. Reaction mixes were made according to the manufacturers' instructions, and the fluorescence intensity of each sample was measured ($\lambda_{ex} = 535/\lambda_{em} = 587$ nm). Aspartate aminotransferase (AST) (MAK055-1KT, Millipore Sigma) and alanine aminotransferase (ALT) (MAK052-1KT, Millipore Sigma) activities were measured according to the manufacturers' instructions.

RNA production

For ONE-seq, a 120-mer pegRNA was chemically synthesized under solid-phase synthesis by a commercial supplier (Integrated DNA Technologies) with end modifications: 5′-mA*mC*mU*U
UGCUGCCACAAUACCUGUUUUAGAGCUAGAAAUAGCAAGU
UAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCAC
CGAGUCGGUGCGAGAAGGGACGUGGUAUUGUG*mG*mC*

mA-3′, where "m" and * indicate 2′-*O*-methylation and phosphorothioate linkage, respectively.

ONE-seq and *in silico* nomination of candidate off-target sites

The methods were similar to those previously described.[9,17] The human ONE-seq library for the protospacer sequence of the lead pegRNA (5′-ACTTTGCTGCCACAATACCT-3′) was designed with the GRCh38 Ensembl v98 reference genome. In comparison to the on-target site, sites with up to four mismatches; up to three mismatches plus up to one DNA or RNA bulge; and up to two mismatches plus up to two DNA or RNA bulges were identified with Cas-Designer v1.2.[18] The final oligonucleotide sequences were generated with a script,[17] and the oligonucleotide library was synthesized by Twist Biosciences. The library was PCR-amplified and subjected to 0.9× AMPure XP (Beckman Coulter) bead purification. After incubation at 25°C for 10 min in NEBuffer r3.1 (New England Biolabs), ribonucleoproteins from 17 nM recombinant Cas9 protein (New England Biolabs) and 34 nM pegRNA (Integrated DNA Technologies) were mixed with 30 ng of the purified library and incubated at 37°C for 2 h. Proteinase K was added to quench the reaction at 37°C for 45 min, followed by 1.2× AMPure XP bead purification. The reaction was then incubated with Phusion High-Fidelity Polymerase (New England Biolabs) at 72°C for 10 min, followed by 1.2× AMPure XP bead purification. The reaction was ligated with an annealed adaptor oligonucleotide duplex at 25°C for 10 min to facilitate PCR amplification of the cleaved library products, and 2× AMPure XP bead purification followed. Size selection of the ligated reaction was performed on a BluePippin system (Sage Science) so that DNA of 150–200 bp would be isolated on a 3% agarose gel cassette; two subsequent rounds of PCR amplification generated a barcoded library, which underwent paired-end sequencing on an Illumina MiSeq System as described below. The analysis pipeline[17] used for processing the data assigned a score quantifying the cleavage efficiency for each potential off-target site, relative to the on-target site. Sites were ranked on the basis of this ONE-seq score, and the mean ONE-seq score between duplicate experiments was used for site prioritization. The top 24 ONE-seq sites had scores ranging from 1.0 (on-target site) to below a threshold of 0.01 (1/100th of the cleavage efficiency observed with the on-target site). In addition to the top 24 ONE-seq sites, we selected all remaining genomic sites (55 in total) with up to three mismatches, up to two mismatches plus up to one DNA or RNA bulge, and up to one mismatch plus up to two DNA or RNA bulges, in comparison to the on-target site, as identified in silico with Cas-Designer v1.2. The results of next-generation sequencing (see below) for these sites are listed in Table S4.

Next-generation sequencing

The methods were similar to those previously described.[9] PCR reactions were performed with NEBNext Polymerase (NEB) and the primer sets listed in Table S5; these were designed with Primer3 v4.1.0 (https://primer3.ut.ee/). The following program was used for all genomic DNA PCRs: 98°C for 20 s, 35 × (98°C for 20 s, 57°C for 30 s, 72°C for 10 s), and 72°C for 2 min. PCR products were visualized via capillary electrophoresis (QIAxcel, QIAGEN) and then purified and normalized via an NGS Normalization 96-Well Kit (Norgen Biotek Corporation). A secondary barcoding PCR added Illumina barcodes (Nextera XT Index Kit V2 Set A and/or Nextera XT Index Kit V2 Set D) through the use of ~15 ng of first-round PCR product as a template, followed by purification and normalization. Final pooled libraries were quantified on a Qubit 3.0 Fluorometer (Thermo Fisher Scientific); then, after denaturation, dilution to 10 pM, and supplementation with 15% PhiX, libraries underwent single-end or paired-end sequencing on an Illumina MiSeq System. We analyzed the amplicon sequencing data with CRISPResso2 v2[19] and scripts to quantify editing. For on-target editing, corrective T-to-C editing was quantified at the site of the c.1222C>T variant. For candidate off-target sites, we considered all aligned reads in which there was any mismatch to the reference sequence in any position within a window spanning from the position 5′ of the non-target-strand nick site to the 3′ end of the portion of the non-target strand overlapping with the RTT, 14 positions downstream of the nick site—thus accounting for single-nucleotide substitutions and indels within the window—to be potentially edited. In some cases, PCR amplicons were subjected to Sanger sequencing, performed by GENEWIZ, instead of next-generation sequencing.

Data analysis

Sequencing data were analyzed as described above. Other data were collected and analyzed with GraphPad Prism v10.0.2.

Results

PKU individuals with *PAH* c.1222C>T variants demonstrate very poor metabolic control

Among 129 PKU-affected individuals followed by the Children's Hospital of Philadelphia (CHOP) Metabolic Clinic, 32 (24.8%) were found to be compound heterozygous for the *PAH* c.1222C>T allele, and four (3.1%) were homozygous for the variant. This might be an underestimate of true c.1222C>T prevalence because genotype information was not available for some older individuals with PKU.

Consensus management guidelines from the American College of Medical Genetics recommend individuals with PKU maintain Phe levels in the 120–360 μmol/L range.[1] Levels above 600 μmol/L can be neurotoxic and are associated with poor psychiatric and neurocognitive outcomes.[20–22] In our study cohort, 33 of 36 (91.6%) treated individuals had at least a single Phe level above 360 μmol/L, and 25 of 36 (69.4%) had at least a single level above 600 μmol/L (Figure 1). Furthermore, 10 of 36 (27.7%) had lifetime average Phe levels above 360 μmol/L. The four individuals homozygous for the c.1222C>T allele demonstrated especially poor metabolic control.

In addition to high Phe levels, individuals with c.1222C>T variants demonstrated poor adherence to Phe monitoring schedules. Although exact recommended intervals can vary by clinical scenario, general guidelines include weekly Phe monitoring prior to 1 year of age, every-2-week monitoring for ages 1–12 years, and every-2–4 weeks monitoring in adolescents and adults (Figure 2, gray box).[1] We calculated the interval since the last recorded value for every Phe measurement in our cohort (Figure 2, black dots). The interval between monitoring labs was longer than recommended for 1,556 of 2,535 (61.3%) of the Phe values measured. Collectively, these

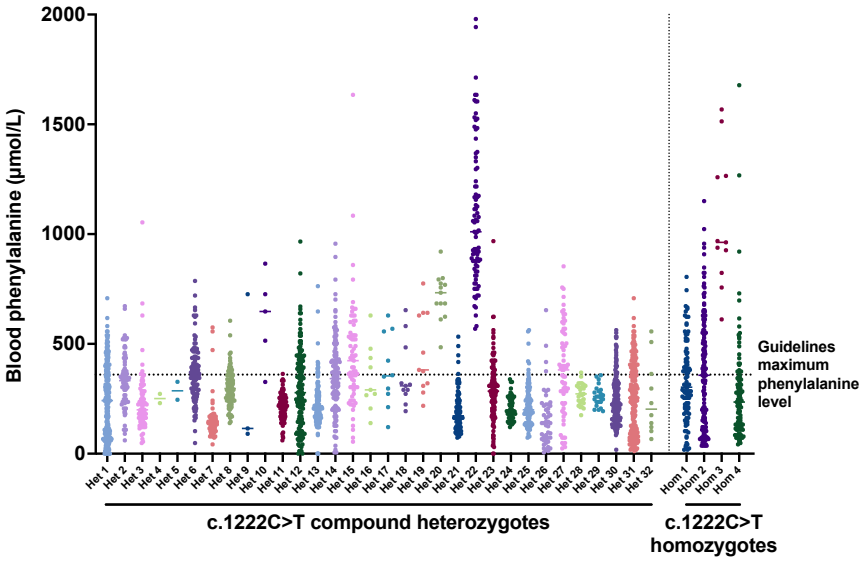


Figure 1. Lifetime recorded measurements of blood phenylalanine concentrations in PKU individuals with *PAH* c.1222C>T variants

All available Phe concentrations from 32 PKU individuals harboring the c.1222C>T variant in *trans* with a second pathogenic allele are shown on the left, where all available levels from four individuals homozygous for c.1222C>T are shown on the right. Most individuals had several recorded concentrations above 360 µmol/L, the maximum recommended concentration per PKU guidelines, as indicated by the dotted line.

data demonstrate that most individuals with c.1222C>T variants have poor metabolic control and are at risk for chronic neurologic damage.

Screening and optimization of prime editing in an edited c.1222C>T hepatocyte cell line

Motivated by the severe unmet medical need in PKU individuals with *PAH* c.1222C>T variants, we set out to determine whether prime editing could productively correct the variant in cellular and animal models and thus potentially serve as the basis of a future therapeutic intervention for these individuals. Lacking access to c.1222C>T-individual-derived hepatocytes, we used prime editing to introduce the c.1222C>T variant into HuH-7 human hepatoma cells, a commonly used hepatocyte-like substitute for primary human hepatocytes. Four pegRNA protospacer sequences have NGG PAMs with positioning near the site of the c.1222C>T variant and amenable to prime editing of the variant (Figures 3 and 4). We found that our initial attempt at the use of PEmax with an engineered pegRNA (designated P1)—harboring one of the four protospacer sequences, with a PBS length of 12 nt and an RTT length of 12 nt (lengths consistent with recommendations by previous publications on prime editing)—and either of two ngRNAs, designated N1 and N2, could achieve 34% and 25% insertion of the c.1222C>T variant, respectively, into bulk HuH-7 cells when delivered via plasmid transfection (Figures 3 and 4, insertion set). We used single-cell cloning to expand edited HuH-7 cell lines and identified lines that were homozygous for the c.1222C>T variant.

The c.1222C>T variant is the result of a C→T change that could in principle be reversed with adenine base editing on the antisense strand. However, the lack of an appropriately positioned NGG PAM relative to the variant, as well as the possibility of unwanted bystander editing, militated against the use of a base editor. We therefore pursued prime editing as a therapeutic strategy.

With a c.1222C>T homozygous HuH-7 cell line, we screened a variety of pegRNA and ngRNA candidates for corrective activity with iterative experiments. We initially focused on the four candidate pegRNA protospacer sequences by using a PBS length of 12 nt and an RTT length adjusted on the basis of the distance of the PAM from the site of the c.1222C>T variant, paired with ngRNAs (Figure 4, correction set #1). For the pegRNAs with the two protospacer sequences with PAMs closest to the site of the c.1222C>T variant, we also assessed the effect of introducing one or more synonymous variants (limited to transition variants) to inhibit mismatch repair, in the hope of enhancing the prime editing efficiency. We found that three of the protospacer sequences supported corrective editing activity with double-digit percentages; the best result of 24% resulted from the use of the same protospacer with which the c.1222C>T homozygous HuH-7 cell lines were generated, along with two synonymous variants closely flanking the site of the c.1222C>T variant. We selected this protospacer, along with an RTT harboring the two synonymous variants (pegRNA P4) and a matching ngRNA (N5), for further optimization.

Our next step was to vary the PBS and RTT lengths, ranging from 10 to 14 nt each (Figure 4, correction set #2A). Among these PBS/RTT combinations, we found that a PBS length of 10 nt and RTT length of 14 nt (pegRNA/ngRNA combination P31/N5) performed best, with corrective editing activity of 47%, although several other combinations also had activity in the 40%–50% range, suggesting that there would be no single optimal PBS/RTT configuration that would substantially outperform all others. We extended this work by testing a PBS length of 10 nt with a wider range of RTT lengths (as high as 20 nt) and an RTT length of 14 nt with a wider range of PBS lengths (as low as 9 nt and as high as 16 nt) (Figure 4, correction set #2B). None of these additional combinations exceeded the corrective editing activity of the P31/N5 combination. Taking advantage of the longer-length RTTs in some pegRNAs, which had RTTs of 17–20 nt, we also tested the addition of a third

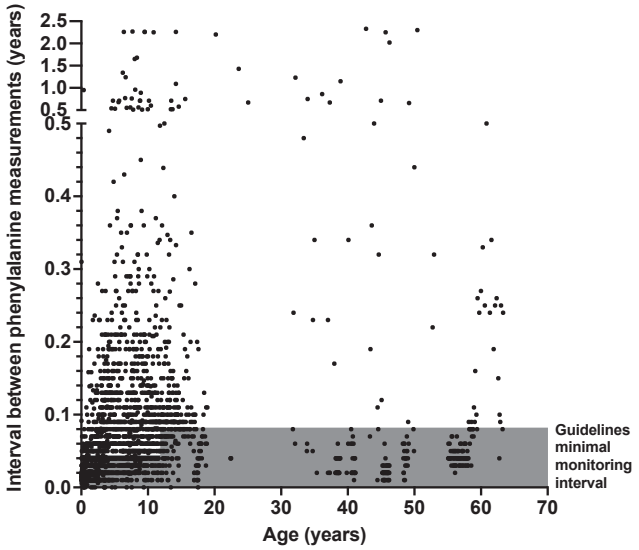


Figure 2. Intervals between measurements of blood phenylalanine concentrations in PKU individuals with *PAH* c.1222C>T variants

The durations of intervals between Phe-monitoring checks are plotted by age. Recommended guidelines for how often Phe should be checked varies with age and is indicated in gray. Most individuals did not check Phe concentrations frequently enough, as evidenced by the fact that most monitoring intervals fell outside the recommended range.

synonymous variant and found that it improved the corrective editing activity beyond that achieved by two synonymous variants (Figure 4, correction set #2B); a PBS length of 10 nt and RTT length of 19 nt (P47/N7 combination) had 48% activity.

We undertook further optimization of the P31/N5 combination by testing all of the nine possible sets of substitutions (whether transition or transversion) in the positions of the two synonymous variants, along with matching ngRNAs (Figure 4, correction set #3). We found that the identity of the substitutions had a meaningful effect on corrective editing activity and that distinct trends were apparent in the two positions. The activity ranged from 37% to 55%. We selected the pegRNA/ngRNA combination (P56/N19) that yielded 55% corrective editing activity in HuH-7 cells for testing *in vivo*.

Generation of humanized PKU mouse models

To generate humanized PKU mice with the *PAH* c.1222C>T variant (R408W mice), we undertook homologous recombination in C57BL/6 embryonic stem cells to knockin the human *PAH* exon 12 sequence harboring the c.1222C>T variant, as well as ~500 bp of flanking genomic sequence both 5′ of the exon (intron 11) and 3′ of the exon (intron 12) and thereby replaced the orthologous portion of the endogenous mouse *Pah* (Figure S1). Homozygous R408W mice were generated from the recombinant embryonic stem cells and confirmed to have elevated blood Phe levels, as well as hypopigmentation and reduced weight, consistent with the mice's having PKU (Figures 5A, 5B, and S2).

We sought to assess therapeutic prime editing not only in homozygous R408W mice—which would model human PKU individuals homozygous for the c.1222C>T allele—but also in mice harboring one c.1222C>T allele and another loss-of-function *Pah* allele—which would model human PKU individuals compound heterozygous for the c.1222C>T allele; such individuals far outnumber homozygous individuals (Figure 1). We reasoned that using an existing loss-of-function mouse *Pah* allele, such as the allele in the commonly used Pah^{enu2} PKU mouse model, would not be optimal; this allele would not have the wild-type human sequence at the site of the c.1222C>T variant in the endogenous *Pah* exon 12, and interactions with the prime editor would not be reflective of the interactions of a wild-type *PAH* exon 12 sequence in human cells. Accordingly, we also generated humanized PKU mice with a different severe, frequent *PAH* allele, the c.1066−11G>A variant (1066 mice).

We undertook homologous recombination in C57BL/6 embryonic stem cells to knock in the human *PAH* exon 11 sequence harboring the c.1066-11G>A variant, as well as the entirety of the flanking introns (10 and 11), the flanking exons (10 and 12), and ≈500–600 base pairs (bp) of flanking genomic sequence both 5′ of exon 10 (intron 9) and 3′ of exon 12 (intron 12), replacing the orthologous portion of the endogenous mouse *Pah* (Figure S3). Mice bearing the humanized c.1066−11G>A allele were generated from the recombinant embryonic stem cells and then crossed with humanized R408W mice, resulting in compound-heterozygous R408W/1066 offspring that were confirmed to have elevated blood Phe levels consistent with the mice's having PKU (Figure 5C).

In vivo corrective editing of the c.1222C>T variant in humanized PKU mice

To deliver the prime editor into the mouse liver, we used a recently reported dual-AAV system: one vector encodes the N-terminal part of the PEmax protein fused to a C-terminal split-intein; the other vector encodes the C-terminal part of the PEmax protein (lacking the RNase H domain) fused to the N-terminal split-intein, an expression cassette for the pegRNA, and an expression cassette for the ngRNA (Figure S4).[16] We produced AAV8 vectors for delivery of the optimized pegRNA/ngRNA combination described above (P56/N19) by mixing the two vectors in equal proportion for intravenous administration to mice.

In our initial *in vivo* study, we treated four age-matched (6 weeks of age) homozygous R408W (PKU) mice each with a combined 8×10^{11} viral genomes (vg) dose and three age-matched homozygous R408W (PKU) mice with a combined 4×10^{11} vg dose; three untreated homozygous R408W (PKU) colony mates and three untreated heterozygous R408W (non-PKU) colony mates served as controls (Figure 5A). In aggregate, the PKU mice had a mean baseline blood Phe level of 1,376 μmol/L; the non-PKU mice, 93 μmol/L. All of the treated PKU mice displayed substantially decreased Phe levels by 2 weeks after treatment and

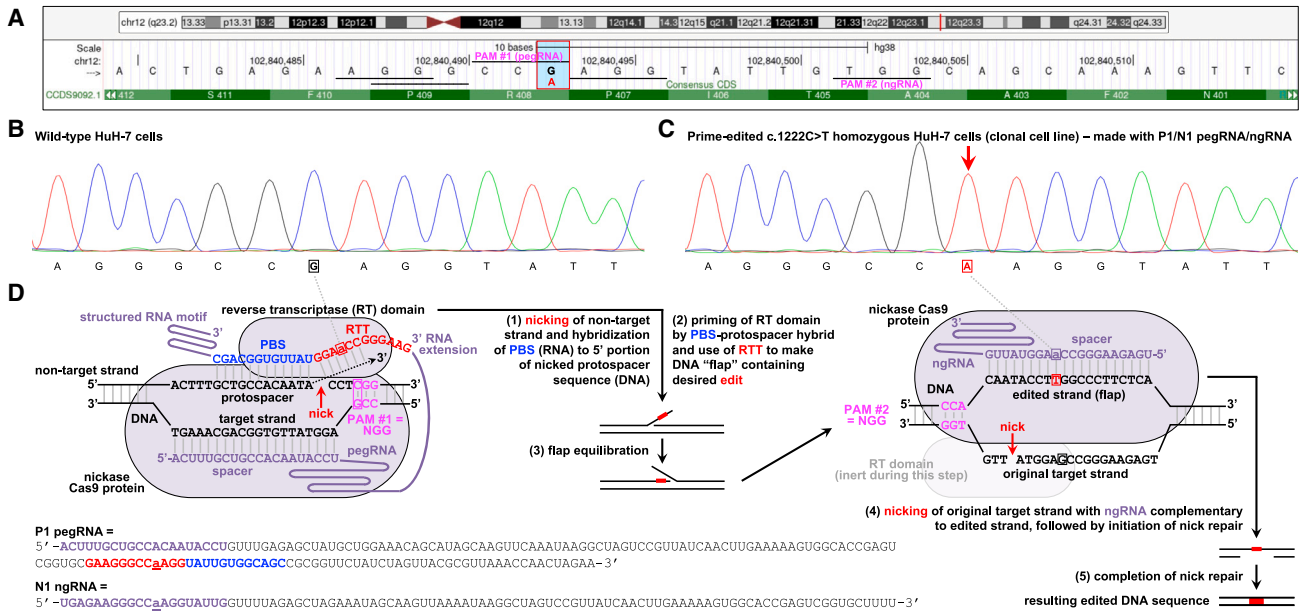


Figure 3. Prime editing to introduce *PAH* c.1222C>T variant in human hepatocytes *in vitro*

(A) Schematic depiction of the genomic site of the *PAH* c.1222C>T variant, adapted from the UCSC Genome Browser (GRCh38/hg38). The vertical blue bar outlined by the red box indicates the G altered to A (in red) by the variant on the antisense strand. The four horizontal lines near the site of the variant indicate four potential NGG PAM sequences for use with pegRNAs or ngRNAs; the horizontal line to the far right of the variant indicates a potential NGG PAM for use with an ngRNA only. The labeled horizontal lines (PAM #1, PAM #2) correspond to the PAMs used in the schematic shown in (D).

(B) Sanger sequencing chromatogram of the genomic site of the *PAH* c.1222C>T variant in wild-type HuH-7 cells. The boxed nucleotide is at the site of the variant.

(C) Chromatogram of the site in *PAH* c.1222C>T homozygous HuH-7 cell line. The boxed nucleotide is at the site of the edited variant, in the position indicated by the red arrow.

(D) Schematic depiction of prime editing with the P1/N1 pegRNA/ngRNA combination to insert the c.1222C>T variant into the *PAH* locus in the human genome. Boxes or underlines indicate nucleotides at the site of the variant.

were well below the threshold of 360 μmol/L by 5 weeks; the higher-dosed mice were at 109 μmol/L, and the lower-dosed mice at 165 μmol/L. Upon necropsy around 6 weeks, next-generation sequencing of genomic DNA from whole-liver samples to determine corrective editing activity showed a range from 31% to 52% (mean 41%) in the higher-dosed mice and from 17% to 35% (mean 26%) in the lower-dosed mice (Figure 5D).

In a second study with older mice, we treated three age-matched (10 weeks of age) homozygous R408W (PKU) mice each with a combined 8×10^{11} vg dose (Figure 5B). Starting with a mean baseline blood Phe level of 1,948 μmol/L, once again all of the treated PKU mice displayed substantially decreased Phe levels by 2 weeks after treatment and, at 151 μmol/L, were well below the threshold of 360 μmol/L by 5 weeks. Whole-liver editing (genomic DNA) ranged from 17% to 52% (mean 33%) in the treated mice (Figures 5D and 5E).

In a third study, we treated three age-matched (6 weeks of age) compound-heterozygous R408W/1066 mice with a combined 1×10^{12} vg dose (Figure 5C). Starting with a mean baseline blood Phe level of 1,867 μmol/L, yet again all of the treated PKU mice displayed substantially decreased Phe levels by 2 weeks after treatment and, at 154 μmol/L, were well below the threshold of 360 μmol/L by 5 weeks. Because the pegRNA/ngRNA combination had the potential to edit wild-type alleles (i.e., introduction of two synonymous variants only), we were only able to estimate the corrective editing (genomic DNA) of the c.1222C>T alleles in the livers of the compound-heterozygous mice; estimates ranged from 22% to 56% (mean 40%) (Figure 5D).

In all three studies, blood aspartate aminotransferase (AST) and alanine aminotransferase (ALT) levels in the treated mice remained normal throughout the study period (Figure S5).

Off-target assessment

To evaluate off-target editing of our optimized pegRNA/ngRNA combination (P56/N19), we performed ONE-seq[17] with a synthetic human genomic library selected by homology to the pegRNA protospacer sequence, treated with recombinant Cas9 protein and the appropriate pegRNA. We reasoned that off-target prime editing activity would be unlikely to occur at a genomic site if Cas9 nuclease activity was not supported at that site. We assessed the top 24 ONE-seq-nominated sites (including the on-target *PAH* site, which was the top-ranked site) with next-generation sequencing of targeted PCR amplicons from pegRNA/ngRNA-transfected versus control-transfected c.1222C>T homozygous HuH-7 genomic DNA samples (Figure 6). To supplement

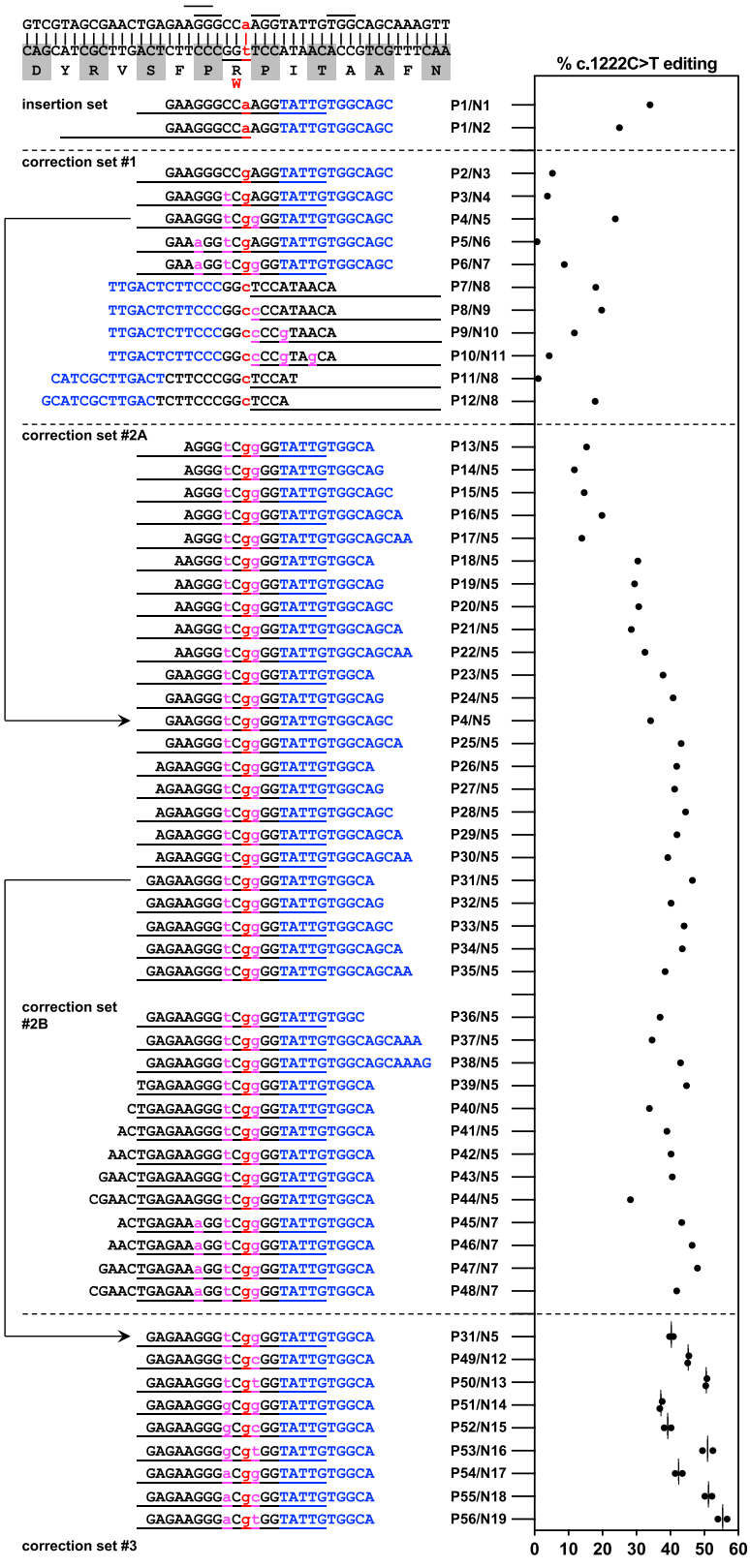


Figure 4. Prime editing to correct the *PAH* c.1222C>T variant in human hepatocytes *in vitro*

Efficiencies of *PAH* c.1222C>T editing by various pegRNA/ngRNA combinations (Px/Nx, refer to Tables S1 and S2 for sequences); the first two combinations in the table/graph were for insertion of the c.1222C>T variant in wild-type HuH-7 cells (used to generate the *PAH* c.1222C>T homozygous HuH-7 cell line in Figure 3), and the remainder for correction of the variant in homozygous HuH-7 cells. For each letter sequence, blue indicates the PBS of the pegRNA, and black indicates the RTT of the pegRNA. Red lowercase letters indicate edits at the site of the variant; magenta lowercase letters indicate additional synonymous edits. The short lines at the top of the table indicate NGG PAM sequences used for pegRNAs and/or ngRNAs. Each long line immediately below a pegRNA PBS/RTT sequence in the table indicates the span of the protospacer sequence of the ngRNA used with that pegRNA. n = 1 biological replicate for each pegRNA/ngRNA combination except for the last set, for which n = 2 biological replicates. Lines in the graph = mean values.

the off-target analysis, we assessed an additional 55 candidate genomic sites nominated by in silico prediction based on sequence similarity to the on-target *PAH* site (Figure 6). We observed no evidence of off-target sequence alteration, either single-nucleotide substitutions or indels, at any of the interrogated sites.

Discussion

We and others have reported the use of base editing to correct PKU pathogenic variants *in vivo*, either the non-human variant present in the homozygous *Pah*^{enu2} PKU mouse model[6,7] or the human c.842C>T (p.Pro281Leu) variant (one of the most frequent *PAH* variants worldwide, albeit much less frequent that the c.1222C>T variant) in a humanized PKU mouse model.[9] In our work on the c.842C>T variant, we found delivery of an adenine base editor mRNA and a guide RNA via lipid nanoparticles (LNPs) achieved efficient corrective editing and the complete normalization of blood Phe levels in humanized PKU mice within a few days of treatment.[9] In a previous effort to use a prime editor delivered either via dual AAV vectors or a single adenoviral vector to correct the non-human variant present in the homozygous *Pah*^{enu2} PKU mouse model, AAV treatment did not significantly modify blood Phe levels (minimal corrective editing), and adenoviral vector treatment moderately reduced blood Phe levels but generally not below the 360 μmol/L threshold (<10% corrective editing).[8] In this current study, using an optimized prime editor configuration for a highly clinically relevant human *PAH* variant, we achieved reduction of blood Phe levels to well below the 360 μmol/L threshold in every single treated PKU mouse with the variant, even

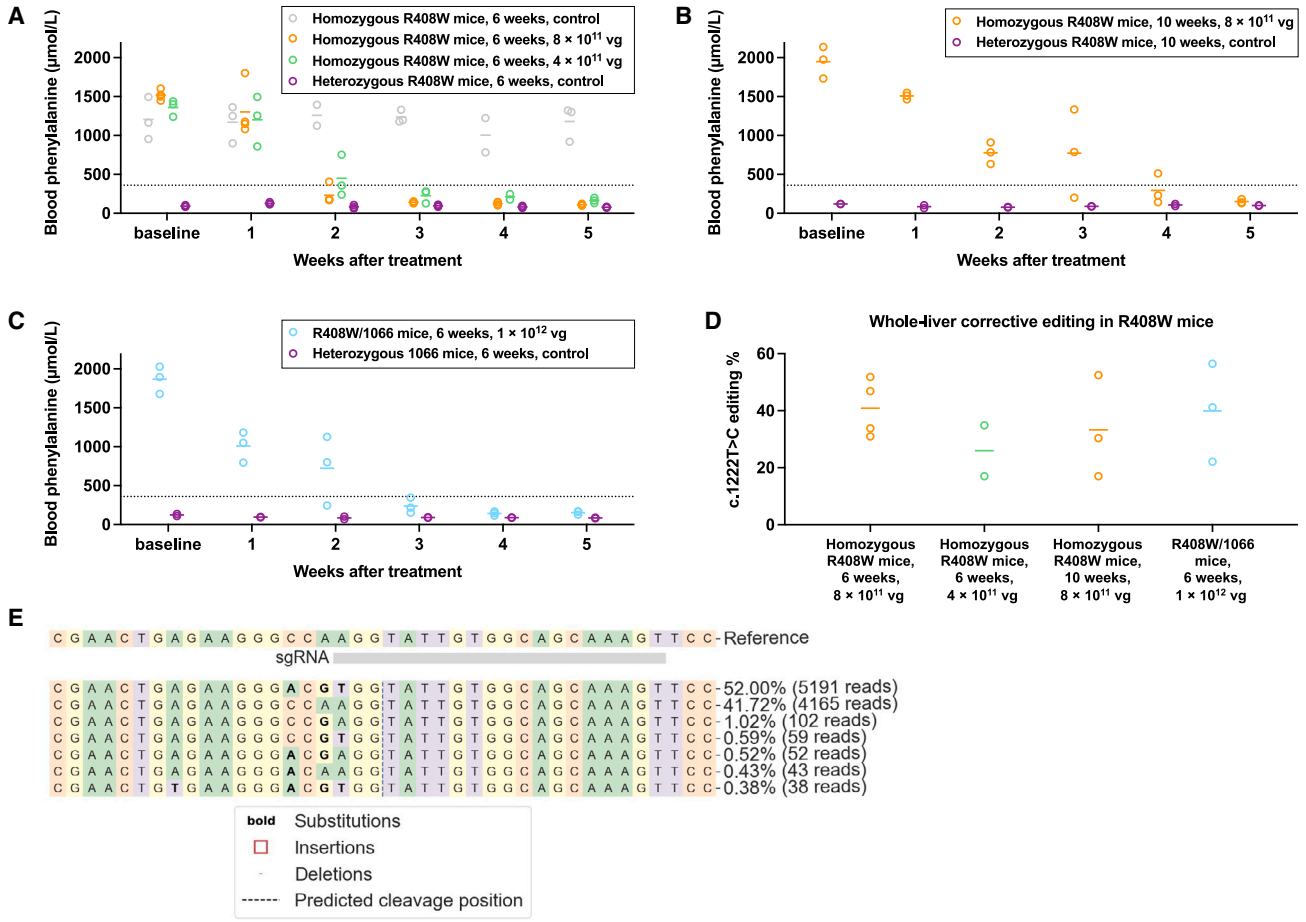


Figure 5. Prime editing to correct *PAH* c.1222C>T variant in humanized mice

(A) Changes in blood phenylalanine levels in homozygous PKU mice after treatment with 8×10^{11} vg AAV dose (n = 4 animals) or with 4×10^{11} vg AAV dose (n = 3 animals); concentrations at various timepoints up to 5 weeks after treatment were compared to concentrations in untreated homozygous PKU control (n = 3 animals) and untreated heterozygous non-PKU control (n = 3 animals) age-matched (6 weeks of age) colony mates (one blood sample per timepoint).

(B) Changes in blood phenylalanine concentrations in homozygous PKU mice after treatment with 8×10^{11} vg AAV dose (n = 3 animals). Concentrations at various timepoints up to 5 weeks after treatment were compared to concentrations in untreated heterozygous non-PKU control (n = 2 animals) age-matched (10 weeks of age) colony mates (one blood sample per timepoint).

(C) Changes in blood phenylalanine concentrations in compound-heterozygous PKU mice after treatment with 1×10^{12} vg AAV dose (n = 3 animals). Concentrations at various timepoints up to 5 weeks after treatment were compared to concentrations in untreated heterozygous non-PKU control (n = 2 animals) age-matched (6 weeks of age) colony mates (one blood sample per timepoint).

(D) Corrective *PAH* c.1222C>T editing (determined from genomic DNA) in the whole liver (mean value from eight liver samples) in each of the treated groups of mice. For compound-heterozygous mice, each displayed number is the estimated percentage of edited c.1222C>T alleles (editable alleles).

(E) Standard CRISPResso output for a liver sample from the treated homozygous PKU mouse with the highest level of editing. Lines in graphs = mean values.

in compound-heterozygous mice with only half the editable alleles of homozygous mice.

We acknowledge the limitations of this study. AAV delivery has disadvantages in comparison to LNP delivery when applied for clinical use. The use of AAV vectors entails (1) the prolonged expression of the editor, which could elicit cytolytic immune responses and exacerbate off-target editing, (2) the risk of vector sequence integration, especially at any double-strand-break sites that should be induced by the editor or should spontaneously occur, and (3) the inability to re-dose the therapy if an initial treatment does not achieve the ~10% corrective editing threshold. LNP delivery shares none of these detrimental aspects of AAV therapy. Current obstacles to LNP delivery of a prime editor include (1) synthesis and encapsulation of the substantially longer mRNA encoding the prime editor protein, in comparison to an mRNA encoding Cas9 nuclease or a base editor, (2) the challenges imposed with regard to scale and purity by the chemical synthesis of a pegRNA that is substantially longer than a standard guide RNA, and (3) the need to co-deliver two guide RNAs rather than one guide RNA. As described in a parallel manuscript,[23] our initial attempt to deploy the P56/N19 prime editing configuration via LNP delivery in humanized R408W PKU mice was unsuccessful, whereas deployment of a base editing configuration to correct the c.1222C>T variant via LNP delivery

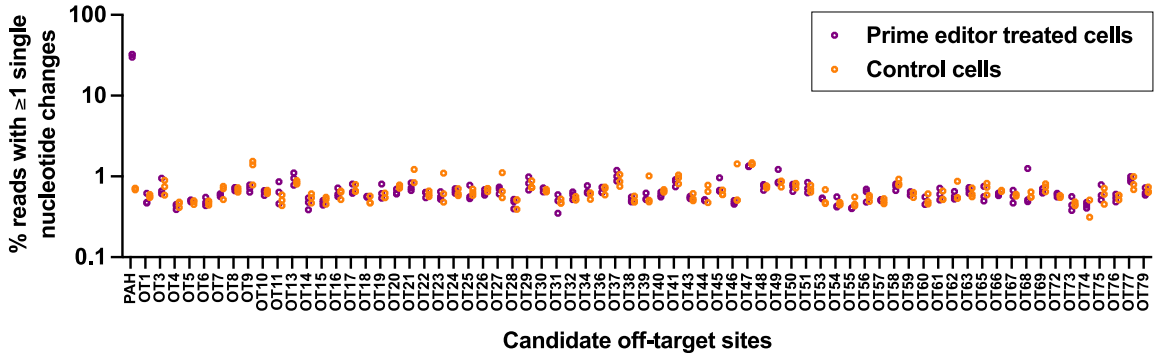


Figure 6. Assessment of off-target editing

On-target or off-target editing at top ONE-seq-nominated or in-silico-nominated candidate sites calculated as the proportion of aligned sequencing reads with alteration of ≥1 base pair within the target sequence (ranging from the position proximal to the pegRNA nick site to the distal position spanned by the RTT sequence, capturing both single-nucleotide substitutions and indels) in *PAH* c.1222C>T homozygous HuH-7 cells that underwent plasmid transfection with the lead pegRNA/ngRNA (P56/N19) combination (n = 3 treated and 3 untreated biological replicates; y axis is $\log_{10}$ scale). Sites with unsuccessful sequencing are omitted. Refer to Table S4 for candidate site sequences and numerical values.

achieved substantial whole-liver editing and normalized Phe levels in humanized R408W PKU mice.

Another limitation is uncertainty as to the optimal approach to off-target assessment for prime editors. Most prior studies have used Cas9 nuclease-oriented techniques as the means to predict sites of off-target prime editing, as we have done in this study, although very new approaches have been designed to assess the reverse transcriptase activity inherent to prime editors.[24,25] We did not perform multiple replicates for each pegRNA/ngRNA combination when screening for corrective editing activity in c.1222C>T homozygous HuH-7 cells because our objective was not accuracy with respect to absolute values of editing activity but rather efficiency in iteratively arriving at a highly active prime-editing configuration that could be tested *in vivo*. We were only able to assess a small subset of all possible pegRNA/ngRNA combinations, and it is possible that combinations with more optimal editing activity than P56/N19 exist. Finally, we did not include vehicle-treated control mice or control mice treated with AAVs encoding an unrelated non-*PAH*-targeting pegRNA/ngRNA in our experiments; we reasoned that saline vehicle or AAV vectors would not intrinsically cause substantial reductions in blood Phe levels or achieve corrective editing.

Our analysis of real-world clinical data from a group of individuals with *PAH* c.1222C>T variants unequivocally demonstrate the shortcomings of the current treatment regime with respect to metabolic control and highlight the high unmet medical need of these individuals. On the basis of the results reported here, we envision a prime-editing therapy that uses the same or a similar pegRNA/ngRNA combination as the one optimized in this study and that is demonstrated to achieve as high as >50% whole-liver corrective editing of the c.1222C>T variant. (That high level of editing, when adjusted for a liver hepatocyte fraction of ~70%, signifies corrective editing of >70% of the c.1222C>T alleles in the hepatocytes; this percentage is well over the threshold needed to address any recessive inborn error of metabolism). Such a therapy, deployed as a treatment to normalize blood Phe levels durably and even permanently, could go a long way toward definitively addressing the medical needs of a large plurality of individuals with PKU. We also anticipate that the workflow demonstrated in this study can be productively applied to the development of prime-editing therapeutics for other PKU variants and for pathogenic variants causing other inborn errors of metabolism.

Data and code availability

The accession number for the next-generation sequencing data reported in this paper is Sequence Read Archive: PRJNA1015257. Other data are available in the Supplemental Data. The GRCh38 Ensembl v98 reference genome annotation was used.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2023.10.005.

Acknowledgments

This work was supported by grants R35-HL145203, R01-HL148769, and U19-NS132301 from the U.S. National Institutes of Health (K.M.); a career development award from the American Heart Association (X.W.); a graduate research fellowship from the U.S. National Science Foundation (M.N.W.); the Children's Hospital of Philadelphia Research Institute Frontier Program (R.C.A.-N.); and the Winkelman Family Fund in Cardiovascular Innovation. The graphical abstract was created with BioRender.com.

Author contributions

K.M., R.C.A.-N., and X.W. supervised the work. D.L.B., M.N.W., P.Q., K.M., and X.W. contributed to wet laboratory experiments.

R.C.A.-N. performed clinical data analyses. K.M. performed bioinformatic analyses. D.L.B., K.M., R.C.A.-N., and X.W. drafted the manuscript, and all authors contributed to the editing of the manuscript.

Declaration of interests

K.M. is an advisor to and holds equity in Verve Therapeutics and Variant Bio and is an advisor to LEXEO Therapeutics. R.C.A.-N. is an advisor to Latus Bio. All other authors have no financial conflicts of interests to disclose.

Received: September 15, 2023
Accepted: October 6, 2023
Published: November 2, 2023

References

1. Vockley, J., Andersson, H.C., Antshel, K.M., Braverman, N.E., Burton, B.K., Frazier, D.M., Mitchell, J., Smith, W.E., Thompson, B.H., Berry, S.A.; and American College of Medical Genetics and Genomics Therapeutics Committee (2014). Phenylalanine hydroxylase deficiency: diagnosis and management guideline. Genet. Med. *16*, 188–200. https://doi.org/10.1038/gim.2013.157.
2. Jurecki, E.R., Cederbaum, S., Kopesky, J., Perry, K., Rohr, F., Sanchez-Valle, A., Viau, K.S., Sheinin, M.Y., and Cohen-Pfeffer, J.L. (2017). Adherence to clinic recommendations among patients with phenylketonuria in the United States. Mol. Genet. Metabol. *120*, 190–197. https://doi.org/10.1016/j.ymgme.2017.01.001.
3. Hillert, A., Anikster, Y., Belanger-Quintana, A., Burlina, A., Burton, B.K., Carducci, C., Chiesa, A.E., Christodoulou, J., Đorđević, M., Desviat, L.R., et al. (2020). The genetic landscape and epidemiology of phenylketonuria. Am. J. Hum. Genet. *107*, 234–250. https://doi.org/10.1016/j.ajhg.2020.06.006.
4. Leuders, S., Wolfgart, E., Ott, T., du Moulin, M., van Teeffelen-Heithoff, A., Vogelpohl, L., Och, U., Marquardt, T., Weglage, J., Feldmann, R., and Rutsch, F. (2014). Influence of PAH genotype on sapropterin response in PKU: results of a single-center cohort study. JIMD Rep. *13*, 101–109. https://doi.org/10.1007/8904_2013_263.
5. Burton, B.K., Longo, N., Vockley, J., Grange, D.K., Harding, C.O., Decker, C., Li, M., Lau, K., Rosen, O., Larimore, K., et al. (2020). Pegvaliase for the treatment of phenylketonuria: results of the phase 2 dose-finding studies with long-term follow-up. Mol. Genet. Metabol. *130*, 239–246. https://doi.org/10.1016/j.ymgme.2020.06.006.
6. Villiger, L., Grisch-Chan, H.M., Lindsay, H., Ringnalda, F., Pogliano, C.B., Allegri, G., Fingerhut, R., Häberle, J., Matos, J., Robinson, M.D., et al. (2018). Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. Nat. Med. *24*, 1519–1525. https://doi.org/10.1038/s41591-018-0209-1.
7. Villiger, L., Rothgangl, T., Witzigmann, D., Oka, R., Lin, P.J.C., Qi, W., Janjuha, S., Berk, C., Ringnalda, F., Beattie, M.B., et al. (2021). In vivo cytidine base editing of hepatocytes without detectable off-target mutations in RNA and DNA. Nat. Biomed. Eng. *5*, 179–189. https://doi.org/10.1038/s41551-020-00671-z.
8. Böck, D., Rothgangl, T., Villiger, L., Schmidheini, L., Matsushita, M., Mathis, N., Ioannidi, E., Rimann, N., Grisch-Chan, H.M., Kreutzer, S., et al. (2022). In vivo prime editing of a metabolic liver disease in mice. Sci. Transl. Med. *14*, eabl9238. https://doi.org/10.1126/scitranslmed.abl9238.
9. Brooks, D.L., Carrasco, M.J., Qu, P., Peranteau, W.H., Ahrens-Nicklas, R.C., Musunuru, K., Alameh, M.-G., and Wang, X. (2023). Rapid and definitive treatment of phenylketonuria in variant-humanized mice with corrective editing. Nat. Commun. *14*, 3451. https://doi.org/10.1038/s41467-023-39246-2.
10. Gillmore, J.D., Gane, E., Taubel, J., Kao, J., Fontana, M., Maitland, M.L., Seitzer, J., O'Connell, D., Walsh, K.R., Wood, K., et al. (2021). CRISPR-Cas9 in vivo gene editing for transthyretin amyloidosis. N. Engl. J. Med. *385*, 493–502. https://doi.org/10.1056/NEJMoa2107454.
11. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature *533*, 420–424. https://doi.org/10.1038/nature17946.
12. Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature *551*, 464–471. https://doi.org/10.1038/nature24644.
13. Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., and Liu, D.R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. Nature *576*, 149–157. https://doi.org/10.1038/s41586-019-1711-4.
14. Chen, P.J., Hussmann, J.A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.F., Chen, C., Nelson, J.W., Newby, G.A., Sahin, M., et al. (2021). Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. Cell *184*, 5635–5652.e29. https://doi.org/10.1016/j.cell.2021.09.018.
15. Nelson, J.W., Randolph, P.B., Shen, S.P., Everette, K.A., Chen, P.J., Anzalone, A.V., An, M., Newby, G.A., Chen, J.C., Hsu, A., and Liu, D.R. (2022). Engineered pegRNAs improve prime editing efficiency. Nat. Biotechnol. *40*, 402–410. https://doi.org/10.1038/s41587-021-01039-7.
16. Davis, J.R., Banskota, S., Levy, J.M., Newby, G.A., Wang, X., Anzalone, A.V., Nelson, A.T., Chen, P.J., Hennes, A.D., An, M., et al. (2023). Efficient prime editing in mouse brain, liver and heart with dual AAVs. Nat. Biotechnol., 1–12. https://doi.org/10.1038/s41587-023-01758-z.
17. Petri, K., Kim, D.Y., Sasaki, K.E., Canver, M.C., Wang, X., Shah, H., Lee, H., Horng, J.E., Clement, K., Iyer, S., et al. (2021). Global-scale CRISPR gene editor specificity profiling by ONE-seq identifies population-specific, variant off-target effects. Preprint at bioRxiv. https://doi.org/10.1101/2021.04.05.438458v1.
18. Bae, S., Park, J., and Kim, J.S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics *30*, 1473–1475. https://doi.org/10.1093/bioinformatics/btu048.
19. Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E., and Pinello, L. (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat. Biotechnol. *37*, 224–226. https://doi.org/10.1038/s41587-019-0032-3.
20. Romani, C., Manti, F., Nardecchia, F., Valentini, F., Fallarino, N., Carducci, C., De Leo, S., MacDonald, A., Palermo, L., and

Leuzzi, V. (2019). Adult cognitive outcomes in phenylketonuria: explaining causes of variability beyond average Phe levels. Orphanet J. Rare Dis. *14*, 273. https://doi.org/10.1186/s13023-019-1225-z.

21. Ashe, K., Kelso, W., Farrand, S., Panetta, J., Fazio, T., De Jong, G., and Walterfang, M. (2019). Psychiatric and cognitive aspects of phenylketonuria: the limitations of diet and promise of new treatments. Front. Psychiatr. *10*, 561. https://doi.org/10.3389/fpsyt.2019.00561.
22. Thomas, L., Olson, A., and Romani, C. (2023). The impact of metabolic control on cognition, neurophysiology, and well-being in PKU: a systematic review and meta-analysis of the within-participant literature. Mol. Genet. Metabol. *138*, 106969. https://doi.org/10.1016/j.ymgme.2022.106969.
23. Brooks, D.L., Whittaker, M.N., Said, H., Dwivedi, G., Qu, P., Musunuru, K., Ahrens-Nicklas, R.C., Alameh, M.-G., and Wang, X. (2023). A base editing strategy using mRNA-LNPs to correct the most frequent phenylketonuria variant in vivo. HGG Adv *2023*, 2003–2014. https://doi.org/10.1016/j.xhgg.2023.100253.
24. Kwon, J., Kim, M., Bae, S., Jo, A., Kim, Y., and Lee, J.K. (2022). TAPE-seq is a cell-based method for predicting genome-wide off-target effects of prime editor. Nat. Commun. *13*, 7975. https://doi.org/10.1038/s41467-022-35743-y.
25. Liang, S.Q., Liu, P., Ponnienselvan, K., Suresh, S., Chen, Z., Kramme, C., Chatterjee, P., Zhu, L.J., Sontheimer, E.J., Xue, W., and Wolfe, S.A. (2023). Genome-wide profiling of prime editor off-target sites in vitro and in vivo using PE-tag. Nat. Methods *20*, 898–907. https://doi.org/10.1038/s41592-023-01859-2.

The American Journal of Human Genetics, Volume *110*

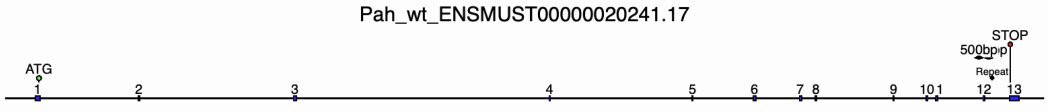
Supplemental information

Efficient *in vivo* prime editing corrects the most frequent phenylketonuria variant, associated with high unmet medical need

Dominique L. Brooks, Madelynn N. Whittaker, Ping Qu, Kiran Musunuru, Rebecca C. Ahrens-Nicklas, and Xiao Wang

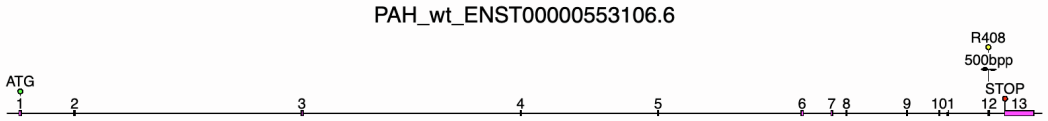
Figure S1. Scheme for Generation of Humanized R408W Mice.

Wildtype Locus of the Murine Pah Gene (standard project)



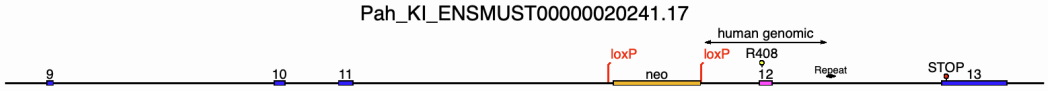
blue boxes: murine exons

Wildtype Locus of the Human Pah Gene (standard project)



pink boxes: human exons

Gene Targeted Locus of the Pah Gene (standard project)



blue boxes: mouse exons; pink boxes: human exons; neo: neomycin cassette for selection in ES cells; loxP: recognition sequence for cre recombinase mediated neo removal

Gene Targeted Locus after cre mediated neo removal (included in project)

Figure S2. Phenylketonuria Phenotypes in Humanized R408W Mice.

(A) Homozygous R408W mouse with PKU as evidenced by hypopigmentation of the fur, on left, compared to heterozygous R408W non-PKU littermate, on right.

(B) Blood phenylalanine levels and weights of homozygous R408W PKU (n = 8 animals) and heterozygous R408W non-PKU (n = 6 animals) age-matched (6 weeks of age) colonymates. Lines = mean values. P values calculated with Mann-Whitney U test.

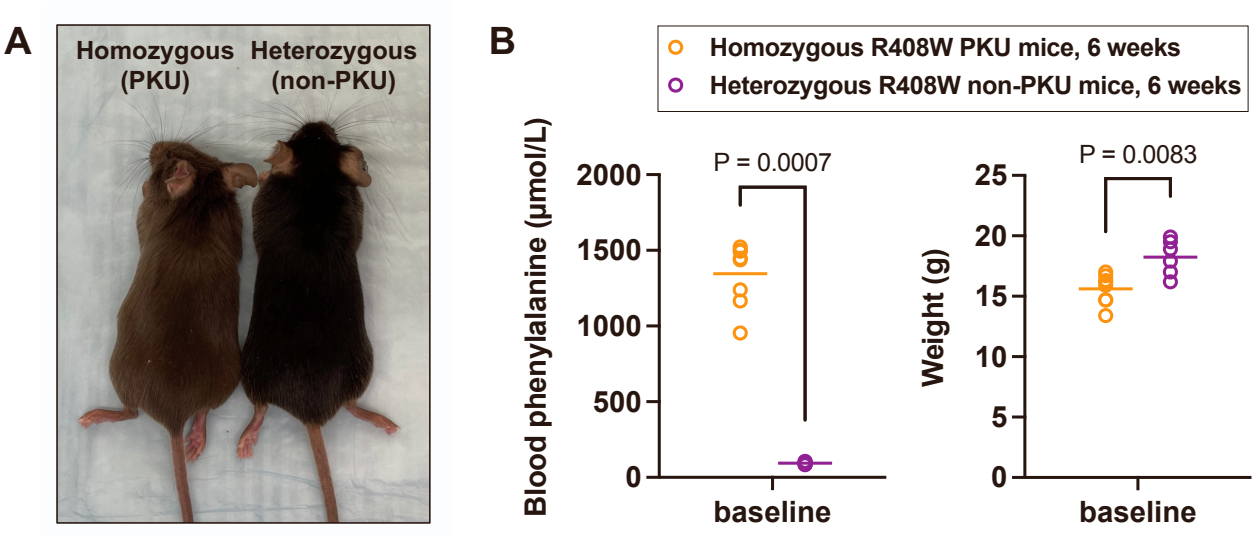
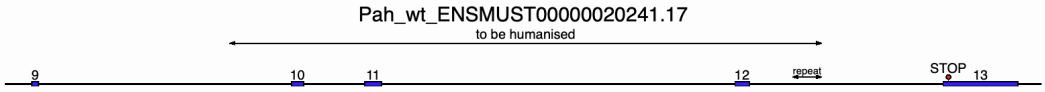


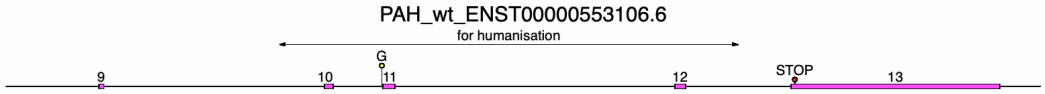
Figure S3. Scheme for Generation of Humanized 1066 Mice.

Wildtype Locus of the Murine Pah Gene (standard project)



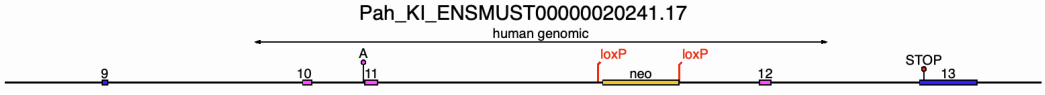
blue boxes: exons

Wildtype Locus of the Human Pah Gene (standard project)



pink boxes: human exons

Gene Targeted Locus of the Pah Gene (standard project)



blue boxes: murine exons; pink boxes: human exons; neo: neomycin cassette for selection in ES cells; loxP: recognition sequence for cre recombinase mediated neo removal

Gene Targeted Locus after cre mediated neo removal (included in project)

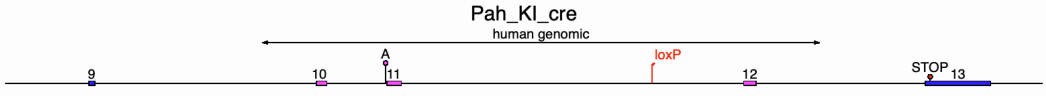


Figure S4. Schematic of Dual-AAV Prime Editing System.

PE3 = prime editor 3; AAV = adeno-associated virus; SpPE2 max = *Streptococcus pyogenes* prime editor 2 max; ITR = inverted terminal repeat; Cbh promoter = chicken β-actin promoter with cytomegalovirus early enhancer element; SpCas9max = *Streptococcus pyogenes* Cas9 max; Npu-N = *Nostoc punctiforme* N-terminal split intein; Npu-C = *Nostoc punctiforme* C-terminal split intein; MMLV = Moloney murine leukemia virus; RT = reverse transcriptase; RNaseH = ribonuclease H; SV40 = simian virus 40; pA = polyadenylation signal; mU6 = mouse U6 promoter; sgRNA = single guide RNA; hU6 = human U6 promoter; epegRNA = engineered prime editing guide RNA. This figure was reproduced without changes from reference 16 [Davis, J.R., Banskota, S., Levy, J.M., Newby, G.A., Wang, X., Anzalone, A.V., Nelson, A.T., Chen, P.J., Hennes, A.D., An, M., et al. (2023). Efficient prime editing in mouse brain, liver and heart with dual AAVs. Nat. Biotechnol. 10.1038/s41587-023-01758-z] in accordance with the terms of the Creative Commons Attribution 4.0 International License: https://creativecommons.org/licenses/by/4.0/.

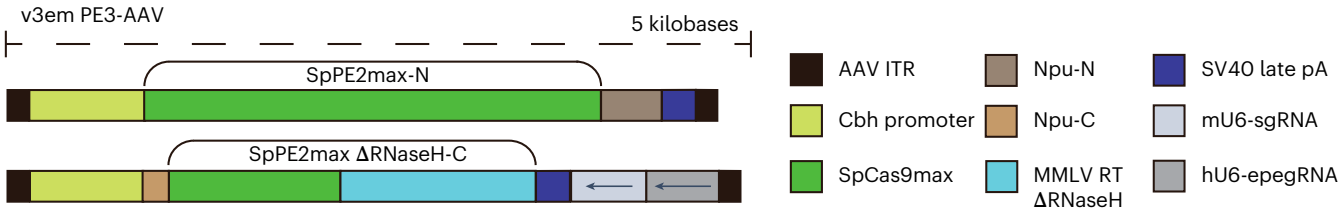


Figure S5. Liver Function Tests in Treated PKU and Control Non-PKU Groups of Mice, Related to Figure 5.

(A) Blood aspartate aminotransferase (AST) levels and alanine aminotransferase (ALT) levels in homozygous PKU mice following treatment with 8×10^{11} vg AAV dose ($n = 4$ animals) or with 4×10^{11} vg AAV dose ($n = 3$ animals), comparing levels at various timepoints up to 5 weeks following treatment to levels in untreated heterozygous non-PKU control ($n = 3$ animals) age-matched (6 weeks of age) colonymates (1 blood sample per timepoint).
(B) AST levels and ALT levels in homozygous PKU mice following treatment with 8×10^{11} vg AAV dose ($n = 3$ animals), comparing levels at various timepoints up to 5 weeks following treatment to levels in untreated heterozygous non-PKU control ($n = 2$ animals) age-matched (10 weeks of age) colonymates (1 blood sample per timepoint).
(C) AST levels and ALT levels in compound heterozygous PKU mice following treatment with 1×10^{12} vg AAV dose ($n = 3$ animals), comparing levels at various timepoints up to 5 weeks following treatment to levels in untreated heterozygous non-PKU control ($n = 2$ animals) age-matched (6 weeks of age) colonymates (1 blood sample per timepoint).
Lines in graphs = mean values.

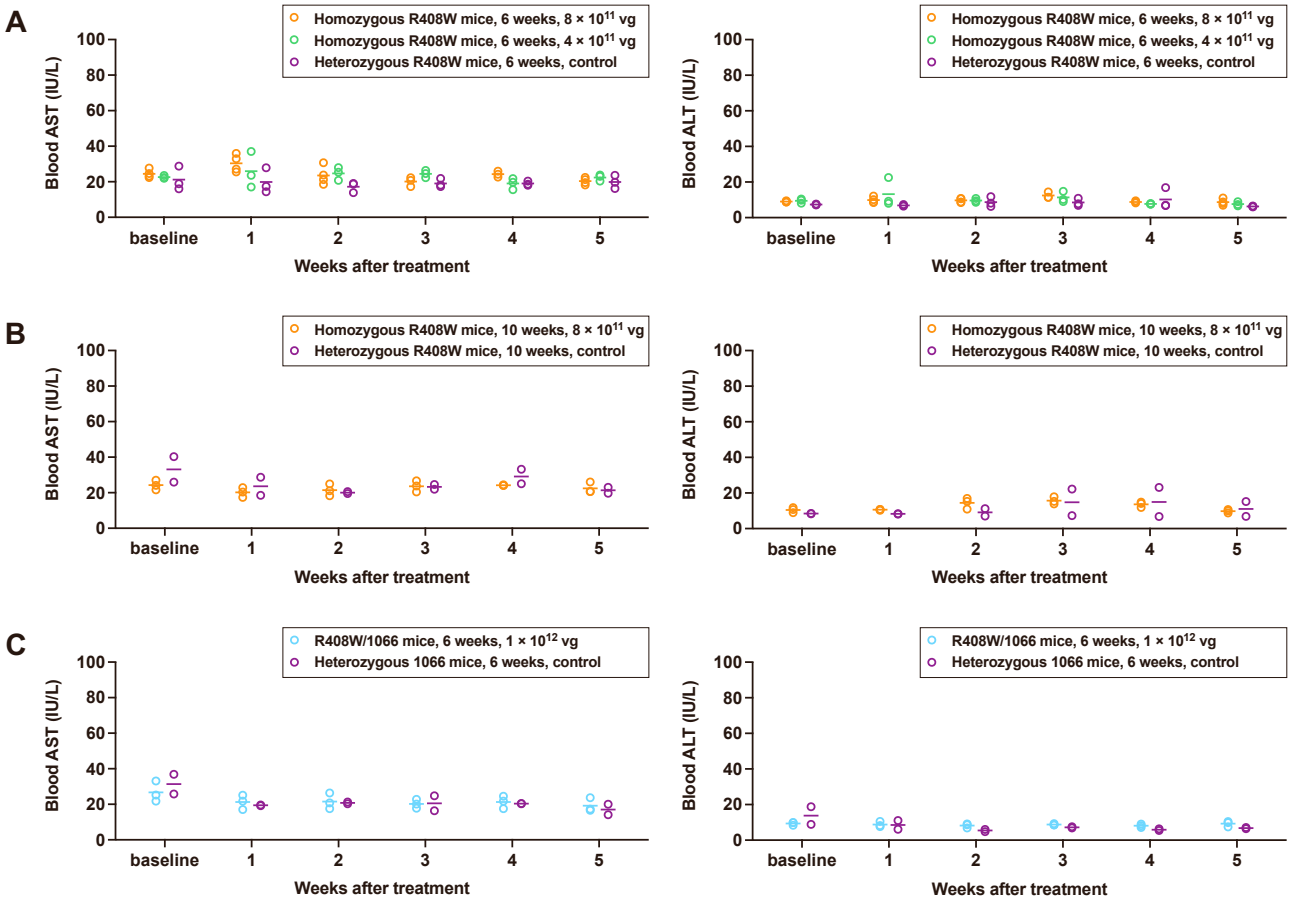


Table S1. pegRNAs.

P1	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAAGGGCCaAGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P2	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAAGGGCCgAGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P3	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAAGGGuCgAGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P4	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAAGGGuCggGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P5	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAAaGGuCgAGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P6	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAAaGGuCggGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P7	**UAGCGAACUGAGAAGGGCCA**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **ACAAUACCUcGGCCCUUCUCAGUU**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P8	**UAGCGAACUGAGAAGGGCCA**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **ACAAUACCccGGCCCUUCUCAGUU**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P9	**UAGCGAACUGAGAAGGGCCA**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **ACAAUgCCccGGCCCUUCUCAGUU**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P10	**UAGCGAACUGAGAAGGGCCA**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **ACgAUgCCccGGCCCUUCUCAGUU**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P11	**GGGUCGUAGCGAACUGAGAA**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **UACCUcGGCCCUUCUCAGUUCGCUAC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P12	**UGGGUCGUAGCGAACUGAGA**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **ACCUcGGCCCUUCUCAGUUCGCUACG**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P13	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P14	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGGGuCggGGUAUUGUGGCAG**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P15	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGGGuCggGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P16	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGGGuCggGGUAUUGUGGCAGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P17	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGGGuCggGGUAUUGUGGCAGCAA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P18	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P19	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AAGGGuCggGGUAUUGUGGCAG**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P20	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AAGGGuCggGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P21	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AAGGGuCggGGUAUUGUGGCAGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P22	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AAGGGuCggGGUAUUGUGGCAGCAA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P23	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P24	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAAGGGuCggGGUAUUGUGGCAG**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P25	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAAGGGuCggGGUAUUGUGGCAGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P26	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGAAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P27	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGAAGGGuCggGGUAUUGUGGCAG**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P28	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGAAGGGuCggGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P29	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGAAGGGuCggGGUAUUGUGGCAGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P30	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AGAAGGGuCggGGUAUUGUGGCAGCAA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P31	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P32	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCggGGUAUUGUGGCAG**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P33	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCggGGUAUUGUGGCAGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P34	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCggGGUAUUGUGGCAGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P35	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCggGGUAUUGUGGCAGCAA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P36	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCggGGUAUUGUGGC**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P37	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCggGGUAUUGUGGCAGCAAA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P38	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCggGGUAUUGUGGCAGCAAAG**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P39	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **UGAGAAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA

P40	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **CUGAGAAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P41	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **ACUGAGAAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P42	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AACUGAGAAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P43	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAACUGAGAAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P44	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **CGAACUGAGAAGGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P45	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **ACUGAGAAaGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P46	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **AACUGAGAAaGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P47	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAACUGAGAAaGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P48	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **CGAACUGAGAAaGGuCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P49	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCgcGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P50	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGuCguGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P51	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGgCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P52	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGgCgcGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P53	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGgCguGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P54	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGaCggGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P55	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGaCgcGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA
P56	**ACUUUGCUGCCACAAUACCU**GUUUGAGAGCUAUGCUGGAAACAGCAUAGCAAGUUCAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC **GAGAAGGGaCguGGUAUUGUGGCA**CGCGGUUCUAUCUAGUUACGCGUUAAACCAACUAGAA

Black bold = spacer sequence; red bold = RTT; blue bold = PBS; lower case indicates nucleotide substitution to be introduced by RTT.

Table S2. ngRNAs.

N1	**UGAGAAGGGCCaAGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N2	**UAGCGAACUGAGAAGGGCCa**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N3	**UGAGAAGGGCCgAGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N4	**UGAGAAGGGuCgAGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N5	**UGAGAAGGGuCggGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N6	**UGAGAAaGGuCgAGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N7	**UGAGAAaGGuCggGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N8	**ACUUUGCUGCCACAAUACCU**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N9	**ACUUUGCUGCCACAAUACCc**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N10	**ACUUUGCUGCCACAAUgCCc**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N11	**ACUUUGCUGCCACgAUgCCc**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N12	**UGAGAAGGGuCgcGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N13	**UGAGAAGGGuCguGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N14	**UGAGAAGGGgCggGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N15	**UGAGAAGGGgCgcGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N16	**UGAGAAGGGgCguGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N17	**UGAGAAGGGaCggGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N18	**UGAGAAGGGaCgcGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
N19	**UGAGAAGGGaCguGGUAUUG**GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU

Black bold = spacer sequence; lower case indicates position complementary to nucleotide substitution introduced by RTT.

Table S3. Sequence of C-terminal AAV Vector With pegRNA/ngRNA P56/N19.

CCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCCTTA
ATTAACCTAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATA
GCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCA
GCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCC
CTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGA
CGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCT
ATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTAT
TTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCG
TGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTA
CATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCG
TATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAG
AGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGA
TCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATT
AACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTAT
TGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGC
AACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACT
TCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGAT
CAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCT
TTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCT
CGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAAC
GGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGA
CAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGA
GCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTT
TCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGC
GGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAAT
TAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTA
TGACCATGATTACGCCAGATTTAATTAAGGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGC
GAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTGCGGCCTCTAGATCAGGGTACCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCG
CCCAACGACCCCCGCCCATTGACGTCAATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTG
TATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTGTGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTAC
GTATTAGTCATCGCTATTACCATGGTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCTCCCCACCCCCAATTTTGTATTTATTTATTTTTTAATTAT
TTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGCGGGGCGGGGCGAGGAGGTGCGGCGGCGGCAGCCAATCAGAG
CGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGCGGGCGGGAGTCGCTGCGCGCTGCCTTCGCCCCGTGCC
CCGCTCCGCCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCTGAGCAA
GAGGTAAGGGTTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTGGAGCACCTGCCTGAAATCACTTTTTTTCAGGTTGGACCGGTGCCACCATGAAACGG
ACAGCCGACGGAAGCGAGTTCGAGTCACCAAAGAAGAAGCGGAAAGTCATCAAGATTGCTACACGGAAATACCTGGGAAAGCAGAACGTGTACGACATCGGCGTGGAGCGG
GATCACAACTTCGCCCTGAAGAATGGCTTTATCGCCAGCAATTGTTTCAACGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTCAAG
ACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGG
AAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCC
AGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAG
AGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATC
ATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATAT
GTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATC
ATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCC
GAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTG
GACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGTGACTCCGGCGGAAGCTCTGGTGGCAGCAAGCGGACCGCC
GACGGCTCTGAATTCGAGAGCCCTAAGAAGAAAAGAAAGGTGAGCGGAGGCTCTAGCGGCGGAAGCACCCTGAACATTGAAGACGAGTATAGACTGCATGAAACAAGCAAG
GAACCCGACGTGTCCCTGGGCTCCACCTGGCTGTCCGACTTTCCCCAGGCCTGGGCCGAGACAGGAGGAATGGGCCTGGCCGTGCGGCAGGCACCCCTGATCATCCCTCTG
AAGGCCACCTCTACACCCGTGAGCATCAAGCAGTACCCTATGTCTCAGGAGGCCAGACTGGGCATCAAGCCTCACATCCAGAGGCTGCTGGACCAGGGCATCCTGGTGCCA
TGCCAGAGCCCCTGGAACACACCACTGCTGCCCGTGAAGAAGCCAGGCACCAATGACTATAGACCCGTGCAGGATCTGAGAGAGGTGAACAAGAGGGTGGAGGATATCCAC
CCCACCGTGCCCAACCCTTACAATCTGCTGTCCGGCCTGCCCCCTTCTCACCAGTGGTATACAGTGCTGGACCTGAAGGATGCCTTCTTTTGTCTGAGACTGCACCCTACC
AGCCAGCCACTGTTCGCCTTTGAGTGGAGGGACCCTGAGATGGGCATCTCTGGCCAGCTGACCTGGACACGCCTGCCTCAGGGCTTCAAGAATAGCCCAACACTGTTTAAC
GAGGCCTGCACCGCGACCTGGCAGATTTCCGGATCCAGCACCCAGATCTGATCCTGCTGCAGTACGTGGACGATCTGCTGCTGGCCGCCACCAGCGAGCTGGATTGCCAG
CAGGGAACACGCGCCCTGCTGCAGACCCTGGGAAACCTGGGATATAGGGCATCCGCCAAGAAGGCCCAGATCTGTCAGAAGCAGGTGAAGTACCTGGGCTATCTGCTGAAG
GAGGGCCAGAGATGGCTGACAGAGGCCAGGAAGGAGACAGTGATGGGCCAGCCAACACCCAAGACCCCAAGACAGCTGAGGGAGTTCCTGGGCAAAGCAGGATTTTGCAGG
CTGTTCATCCCAGGATTCGCAGAGATGGCAGCACCTCTGTACCCACTGACCAAGCCGGGCACCCTGTTTAATTGGGGCCCTGACCAGCAGAAGGCCTATCAGGAGATCAAG
CAGGCCCTGCTGACAGCACCAGCCCTGGGCCTGCCAGACCTGACCAAGCCTTTCGAGCTGTTTGTGGATGAGAAGCAGGGCTACGCCAAGGGCGTGCTGACCCAGAAGCTG
GGACCATGGAGACGGCCCGTGGCCTATCTGTCCAAGAAGCTGGACCCAGTGGCAGCAGGATGGCCACCATGCCTGAGGATGGTGGCAGCAATCGCCGTGCTGACAAAGGAT
GCCGGCAAGCTGACCATGGGACAGCCACTGGTCATCCTGGCACCACACGCAGTGGAGGCCCTGGTGAAGCAGCCTCCAGATCGCTGGCTGTCTAACGCCCGGATGACACAC
TACCAGGCCCTGCTGCTGGACACCGATCGCGTGCAGTTTGGCCCTGTGGTGGCCCTGAATCCAGCCACCCTGCTGCCTCTGCCAGAGGAGGGCCTGCAGCACAACTGTCTG
GACTCTGGCGGCTCAAAAAGAACCGCCGACGGCAGCGAATTCGAGTCTCCCAAGAAGAAGAGGAAAGTCTAAGTCGACCTTTATTTGTGAAATTTGTGATGCTATTGCTTT
ATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGCTTGCAAA
AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC**CAATACCACGTCCCTTCTCA**CAAACAAGGCT
TTTCTCCAAGGGATATTTATAGTCTCAAAACACACAATTACTTTACAGTTAGGGTGAGTTTCCTTTTGTGCTGTTTTTTAAAATAATAATTTAGTATTTGTATCTCTTATA
GAAATCCAAGCCTATCATGTAAAATGTAGCTAGTATTAAAAAGAACAGATTATCTGTCTTTTATCGCACATTAAGCCTCTATAGTTACTAGGAAATATTATATGCAAATTA
ACCGGGGCAGGGGAGTAGCCGAGCTTCTCCCACAAGTCTGTGCGAGGGGGCCGGCGCGGGCCTAGAGATGGCGGCGTCGGATCGCGAATTCAAAAAAATTCTAGTTGGTTT
AACGCGTAACTAGATAGAACCGCG**TGCCACAATA**CCAcGtCCCTTCTCGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTGAACTTGCTATGC
TGTTTCCAGCATAGCTCTCAAAC**AGGTATTGTGGCAGCAAAGT**CGGTGTTTCGTCCTTTCCACAAGATATATAAAGCCAAGAAATCGAAATACTTTCAAGTTACGGTAAGC
ATATGATAGTCCATTTTAAAACATAATTTTAAAACTGCAAACTACCCAAGAAATTATTACTTTCTACGTCACGTATTTTGTACTAATATCTTTGTGTTTACAGTCAAATTA
ATTCTAATTATCTCTCTAACAGCCTTGTATCGTATATGCAAATATGAAGGAATCATGGGAAATAGGCCCTCTTCCTGCCCGACCTTGCGGCCGCGAGGAACCCCTAGTGATG
GAGTTGG

Black bold = spacer sequence; red bold = RTT; blue bold = PBS; lower case indicates nucleotide substitution to be introduced by RTT.

Table S4. Assessment of Off-target Editing in R408W Homozygous HuH-7 Cells.

Genomic site	Protospacer/PAM sequence[a]	Amplicon (chromosome: position range)[b]	Treated sample 1[c]	Treated sample 2	Treated sample 3	Control sample 1	Control sample 2	Control sample 3
PAH	ACTTTGCTGCCACAATACCTTGG	chr12:102840439-102840600	32.32	29.99	32.40			
OT1	ttTaTGCTGCCA-AATACCTAGG	chr8:78462251-78462495	0.62	0.47	0.48	0.55	0.60	0.57
OT2	AagcgtgTGCCACAATACCTGGG	chr2:13589808+13589996	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT3	tCTaTGCTGCCA-AATACCTGGG	chr4:169715128-169715285	0.95	0.61	0.66	0.59	0.75	0.89
OT4	AaTcTaCTGCCA-AATACCTGGG	chr8:76707015-76707213	0.43	0.39	0.45	0.41	0.48	0.47
OT5	AggTTG-TaCCACAATACCTTGG	chr10:113490901-113491091	0.51	0.51	0.48	0.49	0.50	0.45
OT6	AgcTgGCTG-CACAATACCTTGG	chr13:56216964-56217162	0.55	0.43	0.47	0.50	0.44	0.47
OT7	tCTgcGgTGCCACAATACCTTGG	chr12:98762684+98762848	0.59	0.61	0.56	0.75	0.52	0.71
OT8	AgagTGCaGCCACAATACCTAGG	chr9:28448755+28448930	0.71	0.73	0.67	0.64	0.72	0.68
OT9	ACaaTaCTGCCA-AATACCTAGG	chr6:45052043-45052196	0.71	0.64	0.78	1.53	0.78	1.39
OT10	AtTTATGaaGCCACAATACCTGGG	chr12:90309031+90309183	0.67	0.64	0.58	0.64	0.67	0.61
OT11	gtTTTGCTGCCACAAgACCTAGG	chr6:46216177-46216276	0.86	0.64	0.46	0.43	0.59	0.51
OT12	tgTTTGtTGCCA-AATACCTTGG	chr7:136729999-136730159	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT13	ACTTcaCTGCCAtAATACCTGGG	chr2:195824652-195824785	0.95	0.78	1.11	0.89	0.81	0.84
OT14	ACTTTatATGaCACAATACCTGGG	chr9:117378303+117378482	0.48	0.54	0.39	0.54	0.61	0.47
OT15	tCTcTGCT-CCACAATACCTAGG	chr3:25365542+25365698	0.50	0.44	0.47	0.52	0.45	0.55
OT16	tCTgTGCTGCC-CAATACCTTGG	chr13:35884818+35884996	0.72	0.56	0.60	0.64	0.66	0.52
OT17	cCTcTtCT-CCACAATACCTAGG	chr21:27362728-27362863	0.80	0.64	0.62	0.65	0.79	0.66
OT18	gCTTTcCTGCCAC-ATACCTGGG	chr1:84363379+84363574	0.55	0.56	0.57	0.47	0.57	0.47
OT19	Ag-aTGCTGCaACAATACCTAGG	chr18:3456273-3456460	0.80	0.54	0.61	0.62	0.54	0.55
OT20	AggTTcCTcCCACAATACCTGGG	chr5:2961395+2961503	0.70	0.61	0.62	0.78	0.76	0.72
OT21	ACTTTGCaGCCACAtTACCTGGG	chr17:61009795+61009954	0.67	0.72	0.83	0.84	n.d.	1.22
OT22	A-TgTGaTcCCACAATACCTGGG	chr12:5650929-5651072	0.54	0.63	0.64	0.66	0.60	0.56
OT23	AtTcT-CTGCCAtAATACCTGGG	chr8:108304580-108304753	0.52	0.56	0.65	1.10	0.48	0.61
OT24	cCTTTGCTcCCtCAATACCTAGG	chr12:53011957-53012129	0.70	0.62	0.66	0.61	0.71	0.58
OT25	ACTTTGCaGCCACAtgACCTGGG	chr17:75547288+75547430	0.56	0.53	0.78	0.69	0.58	0.63
OT26	ACTTTGCTtCCAaAATAgCTGGG	chr13:76387376-76387540	0.69	0.59	0.62	0.71	0.65	0.70
OT27	ACTTTGttCCACAATAgCTAGG	chr13:77055944-77056107	0.69	0.61	0.74	1.11	0.54	0.65
OT28	ACTTTGCTGCtACAATcCCTAGC	chr13:81974451+81974555	0.40	0.49	0.52	0.39	0.51	0.51
OT29	ACTTTGtTGCCACAcgACCTGGG	chr10:73226422+73226613	0.82	0.68	0.99	0.73	0.73	0.88
OT30	ACaTTGCTaCCACAATACCTAGT	chr11:1493022+1493184	0.72	0.71	0.64	0.65	0.69	0.64
OT31	AaTTTGCTGaCACAAgACCTTGG	chr11:10677426+10677544	0.50	0.59	0.35	0.46	0.47	0.53
OT32	AgTTTGCTtCCACAATACaTAGG	chr11:36513516+36513638	0.59	0.52	0.64	0.56	0.51	0.51
OT33	taTTTGCTGCCAgAATACCTTGG	chr9:73068606-73068772	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT34	ACagTGCTGCCAaAATACCTGGG	chr15:78935858+78935995	0.62	0.64	0.77	0.61	0.52	0.63
OT35	ACTTgcCTGCCACcATACCTGGG	chrY:7231179+7231337	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT36	ACTTTGaTGCCACAATAtCaTGG	chr7:24649482-24649678	0.64	0.63	0.73	0.59	0.74	0.71
OT37	ACTTTtCTGCCtCcATACCTTGG	chr3:70299910+70300073	1.19	0.88	1.00	0.76	1.06	0.86
OT38	ACTTTGCT-CCACAATgCCTGGG	chr1:227324114-227324224	0.50	0.56	0.48	0.50	0.57	0.48
OT39	ACTTTGCTGCCACtgGTACCTGGG	chr4:20389722-20389855	0.62	0.52	0.53	0.51	1.02	0.48
OT40	ACTcTGCT-CCACAATAtCTGGG	chr4:101789193-101789357	0.58	0.64	0.56	0.69	0.64	0.66
OT41	tCTTTGCT-CCACAATAtCTGGG	chr4:129408329+129408552	0.74	0.91	0.87	0.81	0.96	1.04
OT42	ACTTgGCTGCCACcA-ACCTGGG	chr18:23806155+23806259	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT43	ACTTTGCTG-CACcATACCTGGC	chr1:70805762-70805954	0.57	0.55	0.53	0.61	0.53	0.50
OT44	ACTTTGCTGCCAC-ATACCaTGT	chr1:216169709+216169886	0.51	0.50	0.51	0.78	0.47	0.65
OT45	AC-TTGCTGCCAaAATgCCTTGG	chr5:52278218-52278389	0.66	0.96	0.67	0.59	0.68	0.60
OT46	ACTTTGCTTGCCAaAATACCTGGA	chr5:59665136-59665275	0.45	0.50	0.47	1.43	0.51	0.51
OT47	ACTTTGCTGCC-CAAgACtTTGG	chr19:49082325+49082466	1.33	1.35	1.33	1.48	1.41	1.42
OT48	ACaTTGCTG-CAgAATACCTTGG	chr12:50837080-50837226	0.79	0.77	0.67	0.72	0.76	0.75
OT49	AC-TgGCTGCCACcATACCTGGG	chr12:79983294-79983445	1.22	0.85	0.83	0.88	0.86	0.74
OT50	ACTTT-CTGCCACAtgACCTGGG	chr12:126286065-126286233	0.65	0.80	0.76	0.72	0.82	0.79
OT51	ACTTTGCTGatAC-ATACCTGGG	chr16:50829626+50829775	0.63	0.71	0.84	0.78	0.69	0.64
OT52	ACTTTGgTGCCAaACATACCTGGG	chr16:67891109-67891286	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT53	ACcTTGCTG-CAgAATACCTTGG	chr22:41811842+41811995	0.54	0.53	n.d.	0.48	0.69	0.46
OT54	ACTTCTGCTGgCACAAaACCTTGG	chr8:10512237+10512356	0.56	0.42	0.43	0.44	0.45	0.48
OT55	ACTTaGCTGCCAgAA-ACCTTGG	chr8:130760741+130760888	0.40	0.41	n.d.	0.45	0.43	0.56
OT56	AC-TTGCTGCCACAATAaCTTTG	chr14:93610395+93610551	0.65	0.48	0.70	0.55	0.49	0.58
OT57	AaTTTGCTGCCACAATt-CTAGG	chr10:20442761-20442922	0.51	0.50	0.52	0.49	0.46	0.53
OT58	ACTTTGCT-CCACAtTACCcAGG	chr11:58176050-58176199	0.80	0.76	0.67	0.78	0.81	0.92
OT59	tCTTTGCTGCCA-AtTACCTGGG	chr15:29093086+29093232	0.65	0.58	0.59	0.55	0.63	0.62
OT60	ACTTTGCTG-CAaAATACCTTGC	chr21:39088389+39088584	0.45	0.56	n.d.	0.49	0.46	0.61
OT61	ACcTTGCTGCCA-AAcACCTTGG	chr2:160662807-160662949	0.72	0.60	0.51	0.66	0.52	0.52

OT62	ACaTTGCTGgCACAAT-CCTGGG	chr2:213255950+213256111	0.65	0.57	0.52	0.87	0.54	0.55
OT63	ACTTTGgTtCC-CAATACCTTGG	chr2:196036434-196036580	0.68	0.62	0.73	0.56	0.60	0.72
OT64	ACTcT-CTGCCACAATACCTAGA	chrY:13436340-13436535	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT65	ACTTTtCT-CCACAtTACCTTGG	chr6:39041870+39042064	0.56	0.75	0.50	0.59	0.82	0.73
OT66	ACcTTGCTG-CAgAATACCTTGG	chr6:27785457+27785633	0.58	0.62	0.59	n.d.	0.67	0.65
OT67	AgTTTGCTGCCACAA-ACCaAGG	chr6:9683015-9683167	0.47	0.57	0.67	0.60	0.57	n.d.
OT68	ACTTTGg-GCCAgAATACCTAGG	chrX:13570761-13570950	0.51	0.49	1.25	0.64	0.56	0.55
OT69	ACcTTGtTGCCA-AATACCTTGG	chrX:48090226-48090407	0.62	0.64	0.71	0.81	0.65	0.75
OT70	AtTTTGCTGCCACAA-ACCaAGG	chrX:106057761-106057923	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT71	ACTgTGCTGCaACAA-ACCTTGG	chrX:143235479-143235619	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT72	ACTTTGCT-CCACAATgCCTGGT	chr7:155177942+155178085	0.60	0.56	0.62	0.58	0.57	0.55
OT73	AC-TTGaTGCCACAATACtTGGG	chr3:72392649-72392845	0.44	0.56	0.38	0.45	0.48	0.43
OT74	ACTcTGCTGCCACAgT-CCTAGG	chr3:133006042+133006198	0.48	0.40	0.44	0.51	0.31	n.d.
OT75	ACTcTGaTGCCACAGATACCTGGG	chr3:171524879+171525051	0.58	0.79	0.51	0.60	0.45	0.72
OT76	tCTTTGCT-CCACAATAtCTGGG	chr3:169677452-169677627	0.48	0.54	0.60	0.58	0.59	0.51
OT77	tCTTTGCTGCCAC--TACCTTGG	chr1:184534886-184535019	0.91	0.86	0.99	0.99	0.69	0.80
OT78	ACTTTGCTGC--CAATACtTTGG	chr15:40233992+40234119	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
OT79	ACcTTGCTGCCAC--TACCTGGG	chr2:98883002-98883164	0.72	0.59	0.62	0.74	0.65	0.63

[a]The *PAH* on-target sequence and top ONE-seq-nominated and in-silico-nominated candidate off-target sequences with mismatched positions and bulges indicated by lower case letters and dashes.

[b]Coordinates in the GRCh38/hg38 assembly of the human genome; "+" and "-" indicate forward or reverse orientation of protospacer/PAM sequence.

[c]n.d. = PCR was unsuccessful for the genomic site.