# Co-opting regulation bypass repair (CRBR) as a gene correction strategy for monogenic diseases

Jingjie Hu[1], Rebecca A. Bourne[1], Barbara C. McGrath[1], Alice Lin[2], Zifei Pei[1], and Douglas R. Cavener[1]

Affiliations:

[1] Department of Biology, The Pennsylvania State University, University Park, PA, 16802, USA

[2] Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, 16802, USA

Correspondence should be addressed to Douglas R. Cavener (drc9@psu.edu)

ABSTRACT

With the development of CRISPR/Cas9-mediated gene editing technologies, correction of disease-causing mutations has become possible. However, current gene correction strategies preclude mutation repair in post-mitotic cells of human tissues, and a unique repair strategy must be designed and tested for each and every mutation that may occur in a gene. We have developed a novel gene correction strategy, Co-opting Regulation Bypass Repair (CRBR), which can repair a spectrum of mutations in mitotic or post-mitotic cells and tissues. CRBR utilizes the non-homologous end-joining (NHEJ) pathway to insert a coding sequence (CDS) and transcription/translation terminators targeted upstream of any CDS mutation and downstream of the transcriptional promoter. CRBR gene repair results in simultaneous co-option of the endogenous regulatory region and bypass of the genetic defect. We demonstrated the potential of CRBR strategy for human gene therapy by rescuing a mouse model of the Wolcott-Rallison syndrome (WRS) with permanent neonatal diabetes caused by either large deletion or nonsense mutation in the PERK (EIF2AK3) gene. Additionally, we expressed a GFP CDS-terminator cassette that was integrated downstream of the human insulin promoter in cadaver pancreatic islets of Langerhans which paves the way for autologous cell-tissue replacement therapy for gene repair in beta cells.

## Introduction

Conventional treatment of genetic diseases has relied upon long-term drug therapy or organ transplantation which necessitates the use of immunosuppressive drugs that lead to an increased risk of infections and cancer. Because these therapeutic approaches entail severe and debilitating side-effects, strategies to permanently repair the underlying genetic defect have been sought. Gene therapy was pioneered through the use of viral expression vectors to overcome gene deficiency[1-3], either by overexpressing a wild-type cognate to the deficient gene or with a heterologous gene that leads to metabolic compensation. Major drawbacks of viral vector gene expression are a lack of normal temporal, spatial and quantitative gene regulation and continued expression of the mutant gene.

The advent of CRISPR/Cas9 based technologies[4-7] provided an immediate solution to the problems inherent in existing gene therapies, namely targeted correction of genetic disease-causing mutations. Expression of both Cas9 endonuclease and single guide RNA (sgRNA) in eukaryotic cells induces a double-strand break (DSB) at a target site in the genome. The DSB can be repaired by two major pathways: error-prone non-homologous end joining (NHEJ) or homology directed repair (HDR). Although the HDR pathway has been shown to repair genes precisely in mouse models of human disease[8-14], it is dependent upon cellular homologous recombination functions that are only expressed during cell division. Therefore, HDR is not capable of gene repair in post-mitotic cells[15, 16]. Base editing approaches[17-20] provide precise genome editing in post-mitotic tissues, but both HDR and base editing are limited because the components provided in trans must be engineered and tested for each specific mutation. Given that many single-gene genetic diseases[21-23] may be caused by a spectrum of mutations throughout the coding sequence, a gene therapy method that utilizes a single design to repair any one of several possible mutations would be highly advantageous.

Herein, we present a novel gene editing strategy, Co-opting Regulation Bypass Repair (CRBR), which provides the means to repair a spectrum of mutations in mitotic or post-mitotic cells/tissues. CRBR is based on the efficient NHEJ repair pathway that is induced upon CRISPR/Cas9-mediated targeted

DSB. Normally, NHEJ DSB repair results in the rejoining of two genomic DNA fragments cut by Cas9. However, if a heterologous DNA molecule is present, NHEJ repair can ligate the heterologous DNA to the cut ends of the genomic DNA, thereby inserting a large, heterologous DNA fragment at a specific sgRNA/Cas9 target site. Suzuki and coworkers[24] showed that this mechanism can be used to insert a wild-type exon to repair a truncated exon with a specific splicing defect, but their method cannot be generalized to repair mutations within an exon or a spectrum of mutations. The CRBR strategy that we have developed enables targeted insertion of a normal coding sequence (CDS) and transcription/translation terminator cassette upstream of a deleterious disease-causing gene mutation. Expression of the CRBR cassette rescues gene deficiency by restoring normal expression of the wild-type CDS under its native promoter and other regulatory elements, such as enhancers, while bypassing the downstream mutated region. Because a single CRBR CDS-terminator cassette contains all of the coding sequence, it can therefore be used to rescue any coding sequence mutation, as well as splice-site mutations.

To test the efficacy of CRBR, we targeted two genes, eukaryotic translation initiation factor 2 alpha kinase 3 (PERK) and insulin (INS), which are both critically important for pancreatic beta cell functions and maintenance of glucose homeostasis. Using CRBR, we successfully integrated a complete PERK CDS-terminator cassette into the 5'UTR and showed that its expression rescued two independent *Perk* KO alleles in mice, one with a large three-exon deletion and the other with a nonsense mutation. Notably, all of the severe anomalies[25, 26] including neonatal diabetes, growth retardation, necrotic death of the exocrine pancreatic, and skeletal dysplasia were absent in the CRBR allele rescued *Perk* KO mice. We also demonstrated the potential of CRBR for human gene therapy by integrating a GFP CDS-terminator cassette downstream of the human insulin gene by both plasmid transfection and AAV transduction of human cadaver islets. We observed a large number of pancreatic beta cells within these islets that expressed high levels of GFP driven by the insulin promotor. It should now be possible to use CRBR gene repair to correct deficiencies in genes critical for insulin synthesis and secretion by autologous cell-tissue replacement therapy.

## Results

### CRBR-mediated *in vitro* PERK CDS integration in *Perk* KO cell line

The CRBR strategy features a genome editing process that generates a Cas9/sgRNA targeted DSB at a non-coding region in the genome, either within the 5'UTR or an intron. The same Cas9/sgRNA cut sites are engineered in the donor to promote the insertion of a wild-type coding sequence with transcription termination into the genomic DSB (Figure 1A). The CRBR-edited allele expresses the inserted CDS-terminator cassette under control of the endogenous promoter and bypasses expression of the downstream mutation.

We first tested the CRBR strategy in a *Perk* KO mouse embryonic fibroblast (MEF) cell line ($Perk^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$) in which exons 7-9 have been deleted. A partial CDS (~2.2kb) containing the 3' end of intron 6 and exons 7-17 of rat *Perk* followed by a heterologous polyadenylation signal (bGHpA) was designed to integrate into the endogenous intron 6 to restore normal PERK expression. A Cas9/sgRNA target cut site identified within intron 6 was engineered into the donor plasmid with reverse orientation flanking the 3'in6-rPERKex7to17-bGHpA cassette (Figure 1B). The rPERKex7-17-2cut CRBR cassette can be integrated in two possible orientations: the correct 5'- 5'/3' -3' orientation and the incorrect, "flipped" 5'- 3'/5' -3' orientation. We designed the cassette cut sites in reversed orientation so that the correctly oriented integrants would not regenerate the cut sites whereas the incorrectly oriented integrants would restore it. Consequently, incorrectly oriented integrants could be re-excised for possible re-insertion in the correct orientation. *Perk* KO MEF cells co-transfected with the Cas9/sgRNA plasmid and the rPERKex7-17-2cut plasmid were positive for the 5' junction and 3' junction PCRs (Figure 1C), indicating the presence of correctly edited cells within the population. The chimeric mouse-rat *Perk* mRNA was also detected in this mixed cell population (Figure 1D).

This mixed population was then sorted into single cells and expanded to create 96 independent cell lines with two possible *Perk* alleles. Among the 96 single sorted cell lines, thirty-three cell lines were positive for the 5' junction PCR (Figure S1A). In order to test for functional PERK restoration in the

CRBR-edited *Perk* KO MEF cells, we subjected these cell lines to thapsigargin treatment, which induces ER stress by PERK auto-phosphorylation and phosphorylation of its major substrate eIF2α. Cell line #3 had detectable levels of both PERK-P and eIF2α-P, indicating that a functional chimeric PERK protein was expressed in this cell line (Figure 1E). CRBR-editing was confirmed in 7 other single sorted cell lines at the genome level (Figure S1B) but PERK protein expression could not be detected in these lines (data not shown). In these cases, we suspect that the 5' junction within the intron 6 of CRBR-edited *Perk* altered the splicing signal between the mouse exon 6 and rat exon 7-17 CDS of the cassette. Cell line #3, which exhibited PERK expression, had an 11bp deletion at the 5' junction that removed an unintended cryptic splice-acceptor site (AG/G), which fortuitously reversed the splicing defect. The 5' junction of the other 7 non-expressing cell lines occurred as designed (either a clean joint or 1-2bp indels) but retained the splice-acceptor (Figure S1C). The resulting alternative mature transcript in these non-expressing cell lines contained an extra 135bp intronic sequence that encoded a stop codon, which likely resulted in nonsense-mediated mRNA decay. These results show that a CRBR-mediated partial-CDS gene editing can restore *Perk* gene expression and gene function in *Perk* KO cell line, but the introduction of cryptic splice sites must be avoided.

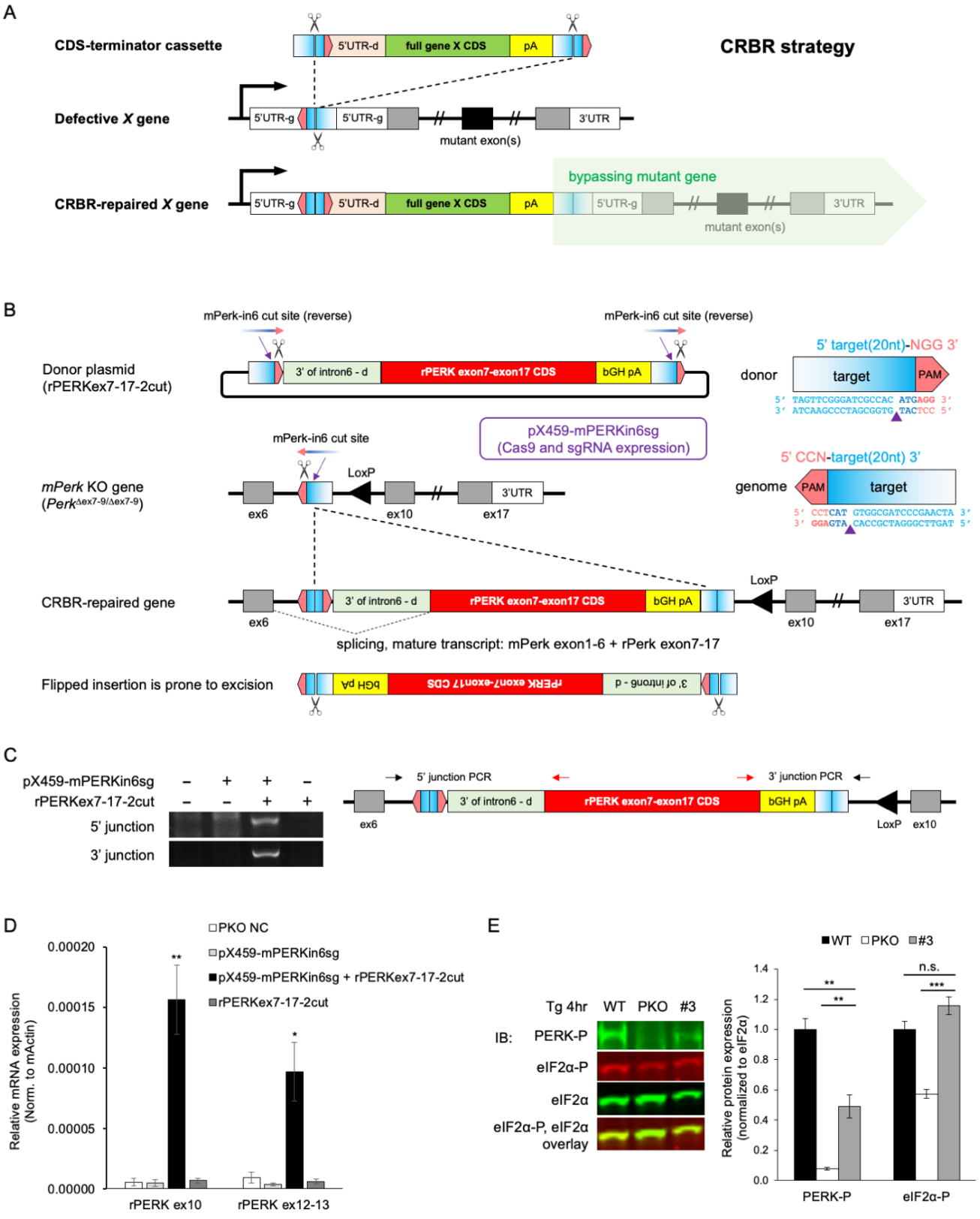



**Figure 1. CRBR-mediated *in vitro* partial PERK CDS integration in *Perk* KO cell line.**
A. Schematic of CRBR strategy. CDS-terminator cassette is flanked by Cas9/gRNA target sites in reverse orientation as in the genome. Correct integration of the CRBR cassette is expressed under the native promoter, with the 5'UTR region having small changes resulted from residue target site

from the donor. Salmon pentagon, PAM site (3nt). Rectangle with blue gradient, Cas9/gRNA targeted protospacer sequence (20nt); Cas9 cleavage locates at 17nt to the white side, 3nt to the blue side. 5'UTR-g, 5'UTR region in the genome. 5'UTR-d, 5'UTR region engineered in the donor.
B. Schematic of CRBR-PartialCDS strategy for *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ genome. Donor plasmid provides a 3'intron6-rPERKex7-17CDS-bGHpA cassette that is flanked by Cas9/gRNA target sites in reverse orientation (5' 20nt-NGG 3') as identified within the *mPerk* intron 6 (5' CCN-20nt 3'). Expression of Cas9 and mPERKin6-sgRNA leads to the cleavage of the mPerk-in6 cut sites that are engineered in the donor to generate the CRBR cassette, and also a targeted DSB at genomic *mPerk* intron 6. Correct integration of the CRBR cassette is retained while the incorrect integrant is prone to excision. Small changes at 5' junction should be spliced out with intron 6 and mature transcript results in a chimeric mouse-rat *Perk* sequence.
C and D. *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ MEF cells ($3\times10^6$ cells) were electroporated with 1.8μg of pX459-mPERKin6sg, 1.6μg of rPERKex7-17-2cut donor or both in 100μl using MEF 2 Nucleofector Kit. Puromycin (1μg/ml) was used to enrich transfected cells (with pX459-mPERKin6sg treatment) for 3 days. (C) Genomic DNA was harvested 6d post transfection for 5' junction PCR and 3' junction PCR analysis. Primers were designed to flank the junction sites. (D) Chimeric mouse-rat *Perk* mRNA expression levels were quantified in sub-cultured *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ (PKO) MEF cells (mixed cell population). Relative gene expression was normalized to mActin. Quantification represents n=3 per treatment. Statistical significance was calculated relative to the no treatment control, pX459-mPERKin6sg only and rPERKex7-17-2cut donor only; *p<0.05, **p<0.01.
E. Protein expression levels were quantified in *Perk*$^{+/+}$ (WT) and *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ (PKO) MEF cells and CRBR edited #3 cell line (*Perk*$^{\text{CRBR-rPERKex7-17/backbone integration}}$) treated with 1μM thapsigargin (Tg) for 4hrs. Quantification represents n=4 per cell line. Relative protein expression was normalized to eIF2α. Statistical significance was calculated relative to *Perk* WT or *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ MEF cells; **p<0.01, **p<0.001, n.s., not significant.

**rPERK-CRBR-edited *Perk* allele completely rescues *Perk* KO mice**

To circumvent the RNA splicing defects that might be generated during NHEJ-DSB repair at the 5' junction, we modified the CRBR strategy so that an entire fully-spliced rat PERK CDS carrying a c-terminal myc tag was targeted to the 5'UTR of the mouse *Perk* gene. The rPERKmyc-2cut CRBR cassette consists of the intact mouse *Perk* 5'UTR, a rPERK CDS (~3.4kb) with a myc tag, bGHpA terminator, and a Cas9/sgRNA target site engineered in reverse orientation (Figure 2A). This modified CRBR strategy preserves the sequence of mouse *Perk* 5'UTR to ensure normal translation initiation. The *Perk* KO nonsense mutant MEF cell line (*Perk*$^{C528X/C528X}$) co-transfected with the Cas9/sgRNA plasmid and the rPERKmyc-2cut plasmid was positive for both 5' junction and 3' junction PCRs (Figure 2B), which confirmed the CRBR-Full-CDS integration at the intended target site in the genome *in vitro*.

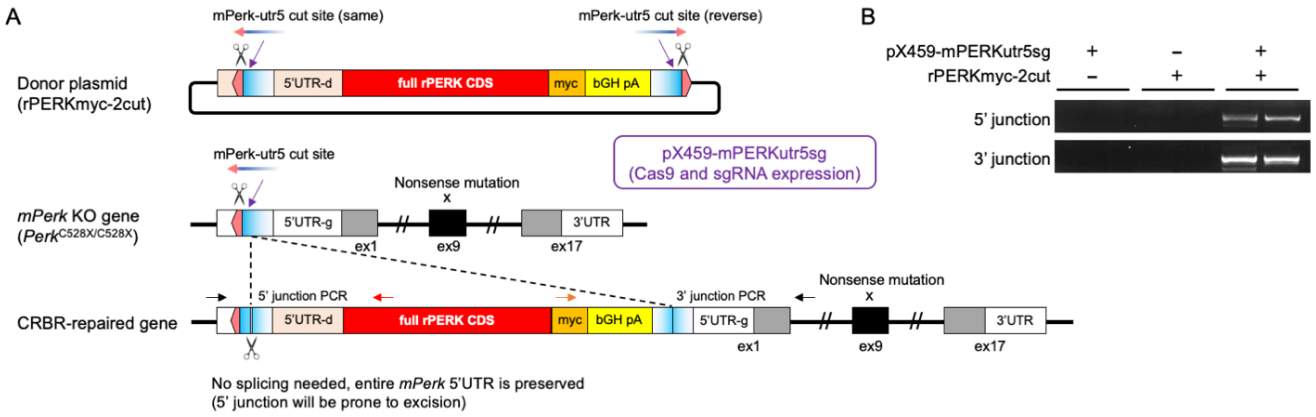



**Figure 2. CRBR-mediated *in vitro* full PERK CDS integration in *Perk* KO cell line.**
A. Schematic of CRBR-FullCDS strategy. Donor plasmid provides a full rPERKmyc CDS-bGHpA cassette that is flanked by a wild-type 5'UTR of *mPerk* and a Cas9/gRNA target site in reverse orientation as identified within the *mPerk* 5'UTR. Expression of Cas9 and mPERKutr5-sgRNA leads to the cleavage of the mPerk-utr5 cut sites that are engineered in the donor to generate the CRBR cassette, and also a targeted DSB at genomic *mPerk* 5'UTR. Correct integration of the CRBR cassette preserved the wild-type sequence of *mPerk* 5'UTR but also resume the mPerk-utr5 cut site that is prone to excision. Small indels could retain the integration of the rPERKmyc CRBR cassette and no splicing is required to achieve a mature transcript of rat *Perk* from the CRBR-edited genome. B. $Perk^{C528X/C528X}$ MEF cells ($1\times10^5$ cells) were electroporated with 1μg of pX459-mPERKutr5sg, 1μg of rPERKmyc-2cut donor or both using the 10μl Neon transfection system (n=2). Genomic DNA was harvested 2d post transfection for 5' junction PCR and 3' junction PCR analysis. Primers were designed to flank the junction sites.

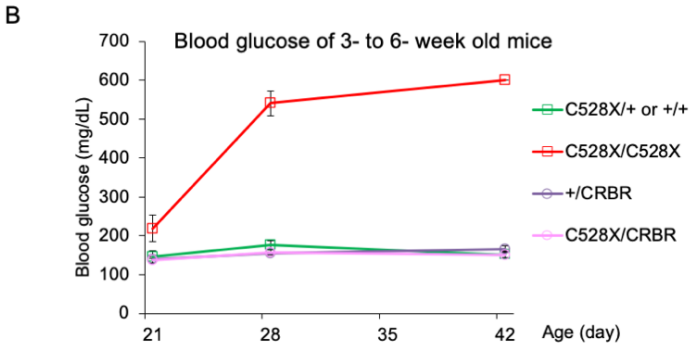
To demonstrate that the CRBR-edited allele can be expressed and regulated normally at the mRNA and protein level during development, we designed an *in vivo* proof-of-concept experiment to test if an engineered rPERK-CRBR-edited allele could rescue a *Perk* KO allele in mice. A key assumption of this strategy is that the integration of the CRBR cassette in a wild type *Perk* allele will generate a complete loss-of-function insertional mutation of the endogenous allele while simultaneously introducing a functional CRBR cassette under the endogenous promoter. The CRBR cassette-insertional mutation can be genetically crossed to a mouse bearing any other type of *Perk* null mutation to generate offspring that carry the CRBR cassette-insertional mutation on one chromosome and a *Perk* null mutation on the other. If these mice express PERK only from the correctly targeted CRBR cassette and are phenotypically normal with respect to the WRS phenotype, the ability of CRBR to rescue PERK expression and function *in vivo* would be confirmed.

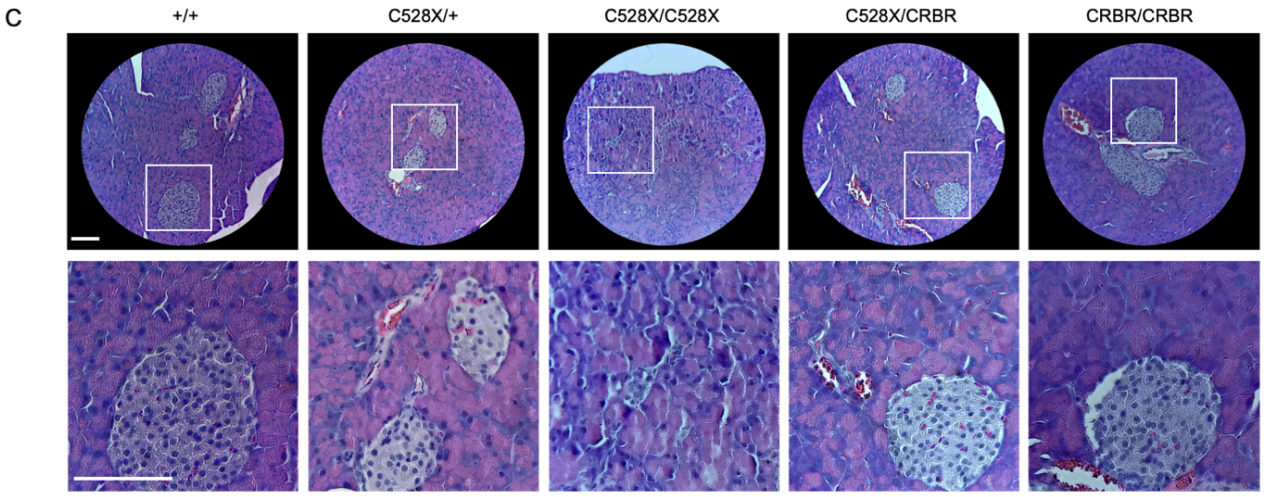
The SpCas9 protein, mPERK-utr5-sgRNA, and the rPERKmyc-2cut plasmid were microinjected into zygotes to create transgenic mice with the rPERKmyc-CDS integrated into the 5'UTR of the wild-type mouse *Perk* allele. Out of the 21 transgenic mice generated, one was positive for both 5' and 3' junction PCRs. After further genotyping, this mouse turned out to be mosaic at the *Perk* locus (WT/4bpDel/rPERK-CRBR/flipped-backbone-CRBR), with the rPERK-CRBR allele having small indels in the 5'UTR region (Figure 3A). To validate the rPERK-CRBR allele is transmissible into the germline, the mouse carrying the mosaic alleles was crossed to wild-type mice to produce F1 offspring that carried the rPERK-CRBR allele. The F1 $Perk^{+/rPERK\text{-}CRBR}$ mice were then crossed to mice heterozygous for a *Perk* null allele ($Perk^{C528X/+}$ or $Perk^{\Delta ex7\text{-}9/+}$). Some of these F2 offspring were KO/rPERK-CRBR heterozygotes ($Perk^{C528X/rPERK\text{-}CRBR}$ or $Perk^{\Delta ex7\text{-}9/rPERK\text{-}CRBR}$). *Perk* KO mice exhibit high neonatal lethality (50-99%), and those mice that survive exhibit severe growth retardation, low pancreatic beta cell mass, exocrine pancreas atrophy, and extreme hyperglycemia by four weeks of age[26-29]. The rPERK-CRBR allele showed complete phenotypic rescue of both the *Perk* nonsense null mutant (n>30, Figure 3B-C) and the *Perk* Δex7-9 deletion mutant (n>8, data not shown) with respect to survivorship, growth, beta cell mass, exocrine pancreas viability, and glucose homeostasis.

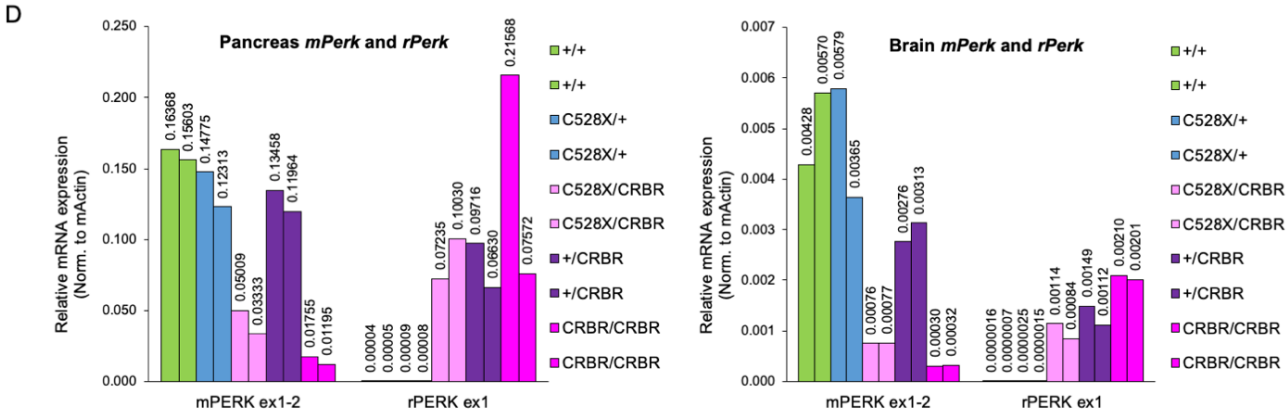
*Perk* mRNA levels from both the rPERK-CRBR cassette and the endogenous mouse *Perk* were analyzed to determine if the CRBR-integrated *rPerk* was expressed and if the CRBR insertion blocked expression of the downstream *mPerk* mRNA as expected from the experimental design. The rPERK-CRBR cassette was robustly expressed in the pancreas and brain in genotypes carrying one or two rPERK-CRBR alleles and was absent in mice lacking the rPERK-CRBR cassette (Figure 3D). Similarly, *mPerk* expression was seen in genotypes carrying one or two copies of the wild-type mouse *Perk* allele, with reduced expression in genotypes carrying the C528X nonsense mutation. The reduction of mouse *Perk* mRNA in the latter is likely caused by nonsense-mediated mRNA decay (NMD). The insertion of the CRBR cassette in the wild-type mouse allele resulted in a ~95% reduction in mouse *Perk* mRNA as can be seen when comparing mice homozygous for the rPERK-CRBR and wild-type alleles. Therefore, we estimate that ~5% of the primary transcripts in the CRBR alleles are

transcriptional read-through of the rPERKmyc-bGHpA terminator within the CRBR cassette resulting in low-level expression of the downstream mouse *Perk* sequence. This conclusion was confirmed by sequencing the cDNA. The small fraction of transcripts that leak through the bGH polyA terminator would be bicistronic, comprised of rPERK-myc and mPERK. It is very unlikely that the mPERK sequences within this hybrid CDS would be translated, however, because normal cap-dependent translation initiates only at the first CDS (i.e. rPERK-myc, would be translated). Any translation of the downstream mPERK CDS would require that the 40S ribosome either remain on the mRNA after translation termination of the rPERK-myc CDS with subsequent translation re-initiation or bind internally upstream of mPERK CDS in a cap-independent mechanism[30]. Both of these possibilities are highly unlikely as they require specialized sequence contexts[31] that are absent in this case.

Consistent with their phenotypic rescue, the C528X/CRBR and Δex7-9/CRBR mice expressed a substantial level of the *rPerk* mRNA derived from the CRBR cassette. Low-level detection of *mPerk* mRNA in these mice was contributed by the KO mutant allele (leftover from NMD) and by the CRBR allele (leaky transcriptional read-through), neither of which are competent for normal translation. We conclude, therefore, that the CRBR rescue of *Perk* null mutations is due solely to the expression of the rPERK protein translated from rPERK-CRBR cassette. Cassette-derived rPERK protein expression was confirmed by immunoblotting with antibodies directed to the myc tag and both rat and mouse PERK (Figure 3E). Critically, the cassette-encoded myc-tagged rPERK showed strong expression in all genotypes bearing a rPERK-CRBR allele but not in other genotypes. Altogether, these results demonstrate that a CRBR-edited allele can rescue a null allele in a living organism. Additionally, they suggest the expression of the CDS-terminator cassette in a CRBR-repaired cell can be regulated normally under the endogenous promoter and provide therapeutic effect *in vivo*.

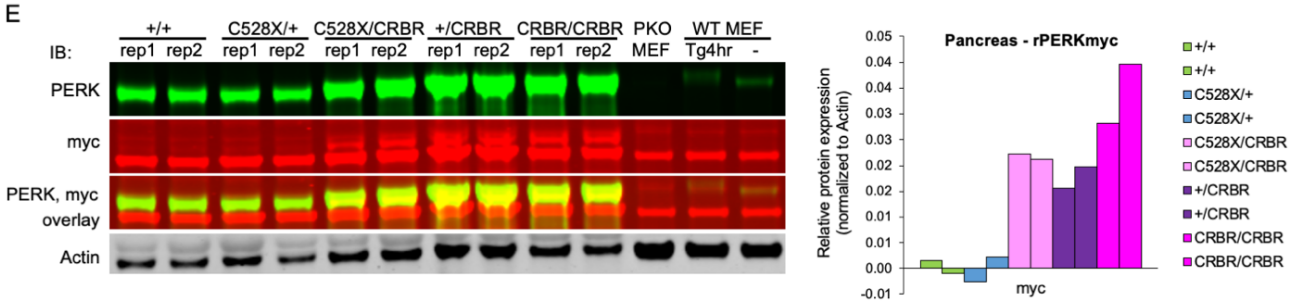



**Figure 3. CRBR-edited *Perk* allele rescues *Perk* KO allele in a proof-of-concept mouse model.**
A. Schematic of rPERK-CRBR allele (in a wild-type mouse *Perk* background) from the transgenic mouse. Sequence of the 5' junction at 5'UTR, see Figure S9.
B. Blood glucose levels were monitored on P21, P28 and P42 mice with genotypes indicated in the chart. Normal blood glucose levels were observed in all *Perk*$^{C528X/rPERR\text{-}CRBR}$ individuals as of manuscript submission (n>30), including ones presented here (n=7). Student's *t*-test showed no significance difference of blood glucose between C528X/CRBR mice (pink, n=7) and littermate +/CRBR mice (purple, n=5) or independent litters with at least one wild-type mouse *Perk* allele (C528X/+ or +/+, green, n=6) at all three age points. Only *Perk* KO mice (C528X/C528X, red, n=8) become diabetic before P28 and exceeding the glucometer upper limit (600mg/dL) by P35. C528X/CRBR or +/CRBR mice were offspring from *Perk*$^{+/rPERR\text{-}CRBR}$ crossing to *Perk*$^{C528X/+}$ mice. *Perk* KO (C528X/C528X) or littermates (C528X/+ or +/+) were offspring from *Perk*$^{C528X/+}$ mice self-crossing. These control glucose data points were from diabetic *Perk* KO mice that were injected with three dosages of CRBR AAV vector mixture, on P21, P28 and P35, but no reversal of the diabetic phenotypes; while C528X/+ or +/+ littermates were injected with saline.
C. Representative Hematoxylin and Eosin staining images from the pancreas of *Perk*$^{+/+}$(P62), *Perk*$^{C528X/+}$(P53), *Perk*$^{C528X/C528X}$(P34), *Perk*$^{C528X/rPERK\text{-}CRBR}$ (P46) and *Perk*$^{rPERK\text{-}CRBR/rPERK\text{-}CRBR}$ (P46) mice. The *Perk*$^{C528X/C528X}$ pancreas has typical *Perk* KO defects such as very small islets with small beta cell mass, disorganized acinus structure, some degranulated cells (white), clear halo around nuclei and gaps between acinar cells, which are not seen in the pancreas of the *Perk*$^{C528X/rPERK\text{-}CRBR}$ and *Perk*$^{rPERK\text{-}CRBR/rPERK\text{-}CRBR}$ mice. Bright field, 20× objective, scale bar, 100μm.
D and E. Two replicate mice with the same genotype were sacrificed at P38 (*Perk*$^{+/+}$, from *Perk*$^{+/rPERK\text{-}CRBR}$ crossing), P58 and P30 (*Perk*$^{C528X/+}$, from *Perk*$^{C528X/+}$ cross *Perk*$^{C528X/rPERK\text{-}CRBR}$), and P46 (*Perk*$^{C528X/rPERK\text{-}CRBR}$, *Perk*$^{+/rPERK\text{-}CRBR}$ and *Perk*$^{rPERK\text{-}CRBR/rPERK\text{-}CRBR}$, from *Perk*$^{C528X/rPERK\text{-}CRBR}$ cross *Perk*$^{+/rPERK\text{-}CRBR}$) for pancreas and brain tissues. The mRNA expression levels (D) of endogenous *mPerk* and *rPerk* from CRBR-edited allele in pancreas and brain were quantified using *mPerk*- and *rPerk*-specific primers and normalized to mActin. Both mPERK and rPERK protein expression levels in pancreas (E) were measured by PERK antibody, while rPERK was also recognized by a myc tag antibody. Negative control, *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ (PKO) MEF cells. Positive control for PERK, *Perk*$^{+/+}$ (WT) MEF cells treated with or without 1μM thapsigargin (Tg) for 4hrs. Relative protein expression was normalized to Actin. rPERKmyc quantitation was further subtracted by the average background signal of the two *Perk*$^{+/+}$ replicates.

## CRBR-mediated *in vitro* and *in vivo* gene editing in mouse pancreatic beta cells

To better assess and visualize the protein expression from a CRBR-edited allele, we applied a similar two-cut CRBR strategy to introduce a GFP CDS into the *Insulin* gene locus, the most highly expressed gene within pancreatic beta cells. Again, we designed the Cas9/sgRNA cut sites in the reverse orientation relative to the native cut site in the 5'UTR target site of the mouse *Ins2* gene (Figure 4A) to increase the likelihood that the EGFP-CDS-pA cassette (~1.1kb) remained stably integrated. This design feature, however, did alter the 5'UTR from the wild-type sequence. To avoid

potential interference with translation, we selected an integration site within a region that is not conserved among mammals, and we avoided introducing new ATG codons within the CRBR-edited 5'UTR that could incorrectly initiate translation of the resulting mRNA. We first tested this strategy in MIN6 mouse beta cells by co-transfecting them with the Cas9/sgRNA plasmid and the EGFP-2cut plasmid. EGFP-positive cells were visible by two days post-transfection and continued to increase in number until 15 days, whereas donor-only treated cells remained EGFP-negative over the same time period (Figure 4B). 5' and 3' junction analysis of the integrants confirmed CRBR-editing at the genome level (Figure 4C). Single cell sorting revealed that the mixed population contained 2.5% GFP-positive cells (Figure S2A); the low percentage of positive cells reflected the relatively poor transfection efficiency of MIN6 cells (~25%).

A subset of GFP-positive cells was clonally isolated for further characterization (Figure S2B). Junction PCRs and sequence analyses showed that cell lines #8, #10, #13 and #14 had one CRBR-edited allele and one allele with small indels at the genomic cleavage site. The cell line #15 had one CRBR-edited allele and one whole donor plasmid integrated allele (Figure S2C). EGFP mRNA expression was confirmed in the sorted GFP-positive cell lines (Figure 4D). We also expected that the native mouse *Ins2* expression would be reduced as a consequence of the insertion of the EGFP CRBR cassette. Indeed, we found that the mouse *Ins2* mRNA levels were reduced compared to wild-type MIN6 cells (Figure 4E). These results suggest the CRBR-integrated EGFP-CDS-pA cassette is expressed and can bypass the endogenous mouse *Ins2* transcription.

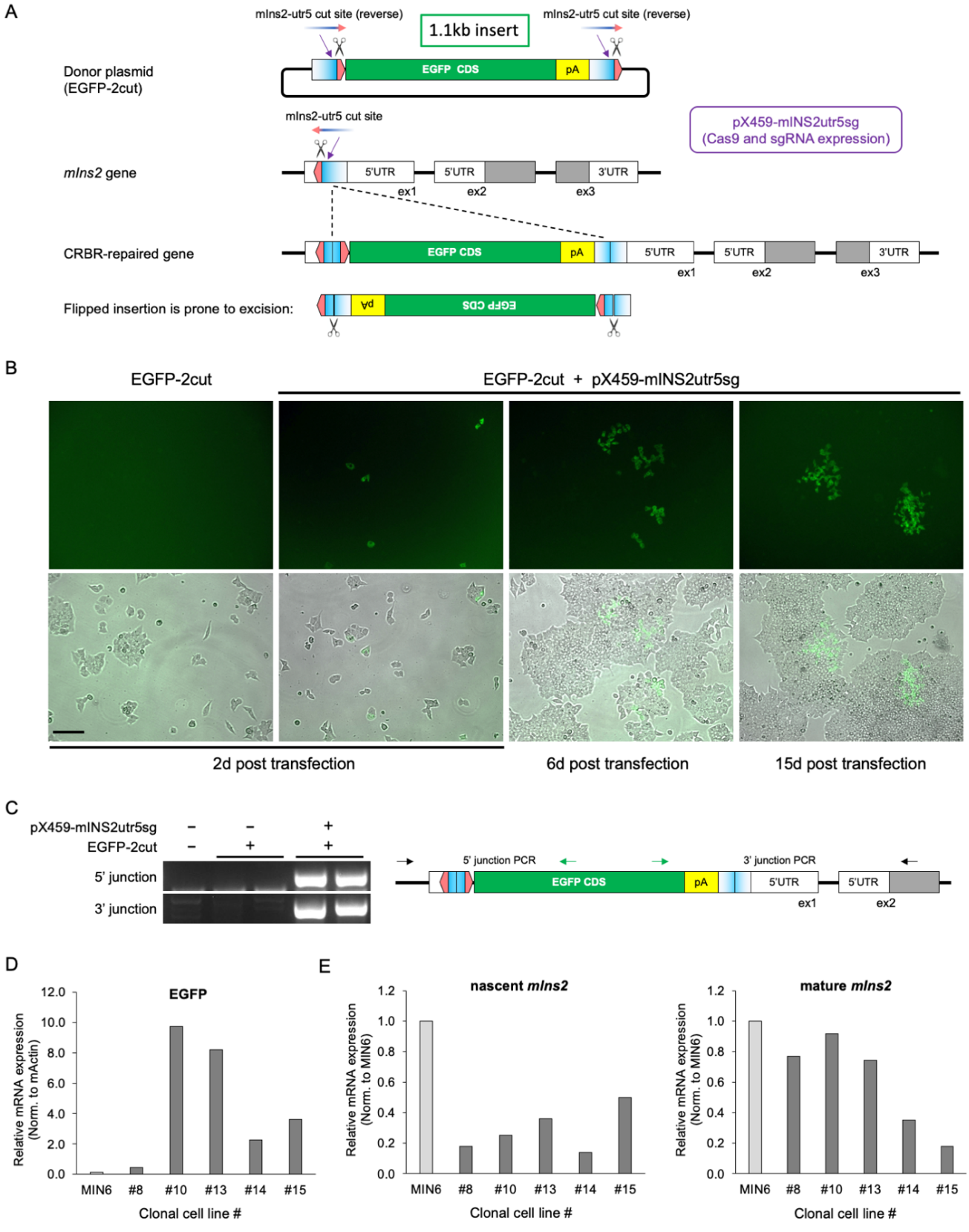



**Figure 4. CRBR-mediated *in vitro* EGFP CDS integration in mouse *Ins2* gene.**

A. Schematic of CRBR-EGFP-2cut strategy for wild-type *mIns2* genome. Donor plasmid provides an EGFP CDS-pA cassette that is flanked by Cas9/gRNA target sites in reverse orientation (5' 20nt-NGG

3') as identified within the *mIns2* 5'UTR in exon 1 (5' CCN-20nt 3'). No *mIns2* 5'UTR sequence is engineered between the 5' cut site and the start codon of EGFP. Expression of Cas9 and mINS2utr5-sgRNA leads to the cleavage of the mIns2-utr5 cut sites that are engineered in the donor to generate the CRBR cassette, and also a targeted DSB at genomic *mIns2* 5'UTR. Correct integration of the CRBR cassette will be retained while incorrect integrant is prone to excision.
B and C. MIN6 cells ($1\times10^6$ cells) were electroporated with 1μg of EGFP-2cut donor with or without 1μg of pX459-mINS2utr5sg in 100μl using Nucleofector V Kit. Cells were imaged (B) as live cultures 2d, 6d, and 15d post transfection at 10× objective, scale bar, 100μm. Genomic DNA (C) was harvested 6d post transfection for 5' junction PCR and 3' junction PCR analysis. Primers were designed to flank the junction sites.
D and E. EGFP mRNA expression levels from the CRBR-edited allele (D) were quantified in five sorted GFP-positive cells (#8, 10, 13, 14, and 15) by normalizing to mActin. Nascent *mIns2* and mature *mIns2* mRNA expression levels (E) were quantified by normalizing to mActin first, and then the relative fold change in expression was calculated relative to wild-type MIN6 cells.

To evaluate the capability of CRBR-mediated gene editing in the mouse pancreas *in vivo*, an AAV carrying the EGFP-CDS-pA cassette and U6-driven mINS2-utr5 sgRNA cassette (AAV-sgRNA-CDS) was systemically delivered to the Rosa26-CAG-Cas9-EGFP mouse strain, which constitutively expresses Cas9 nuclease throughout the body. Using a Cas9 expressing mouse strain reduces the variability of Cas9 delivery in trans via an additional viral vector. For comparison, we also delivered the same AAV-sgRNA-CDS into wild-type mice in combination with another AAV that supplies Cas9 (AAV-Cas9) (Figure 5A). Liver and pancreas tissues from P14 Cas9-EGFP mice were harvested 30-day post retro-orbital (*r.o.*) injection of the AAV8-sgRNA-CDS. Junction PCRs and DNA sequence analyses revealed substantial CRBR-mediated gene editing at the genome level in the liver and a detectable level in the pancreas (Figure 5B). Some individuals had detectable EGFP transcription from the mouse *Ins2* gene locus in the pancreas RNA (Figure 5C). The mouse *Ins2* promoter is not active in the liver, and as expected, EGFP transcription from *Ins2* gene locus in the liver was not observed.

Previous experiments[32-34] suggested that AAV serotypes DJ and 8 would be the most appropriate for delivery into the pancreas. Eight-week old Cas9-EGFP mice were subjected to tail vein injection of AAV-sgRNA-CDS of either serotype DJ or 8. Both serotypes had substantial CRBR-mediated gene editing at the genome level in the liver, with AAV-DJ being more efficient (Figure 5D). Junction PCRs

analysis indicated that the tail vein injected AAV8-sgRNA-CDS was capable of targeting the pancreas, with some individuals having detectable CRBR-editing at the genome level. However, similarly administered AAV-DJ-sgRNA-CDS showed no indication of pancreatic CRBR editing. These results show that systematic delivery of the AAV8-CRBR-construct via intravenous injection can result in CRBR editing at the genome level in the liver and pancreas, as well as CRBR-mediated EGFP mRNA expression in pancreatic beta cells under the control of the *Ins2* promoter.

We next tested whether providing both the sgRNA-CDS and Cas9 via AAV-DJ vectors could also elicit gene editing in wild-type mice (lacking endogenous Cas9). CRBR-mediated gene editing was achieved in the liver (Figure 5E) but not as efficiently as seen in the previous experiment where Cas9 was endogenously expressed (Figure 5D). This result suggests that CRBR-mediated gene editing is feasible *in vivo* via dual AAV delivery once both viral vectors are successfully transduced in the host cell. For human therapy purposes, a more direct intra-organ injection route may improve AAV delivery to the pancreas or other tissues that are challenging to target by intravenous injection.

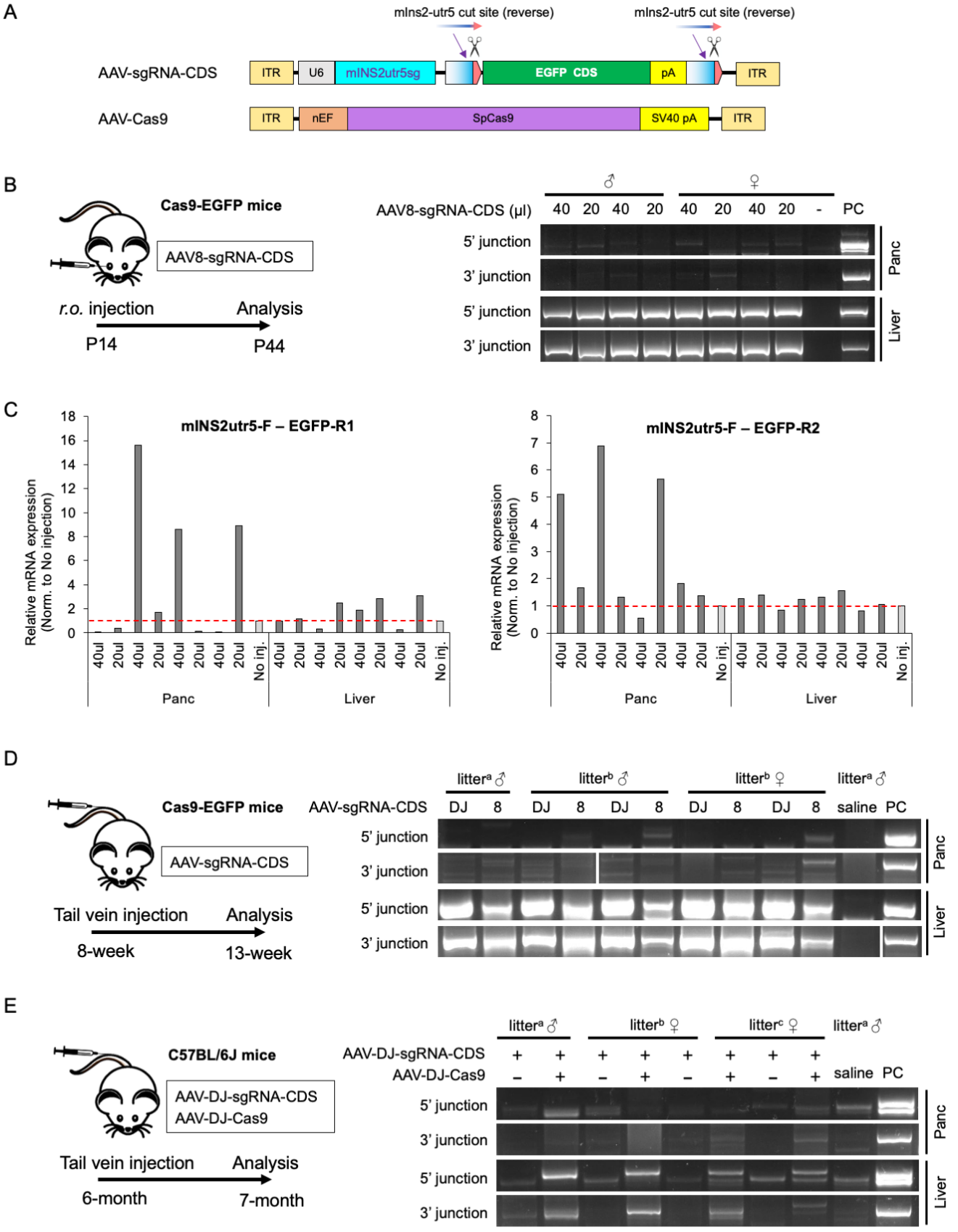



**Figure 5. CRBR-mediated *in vivo* EGFP CDS integration in mouse *Ins2* gene.**

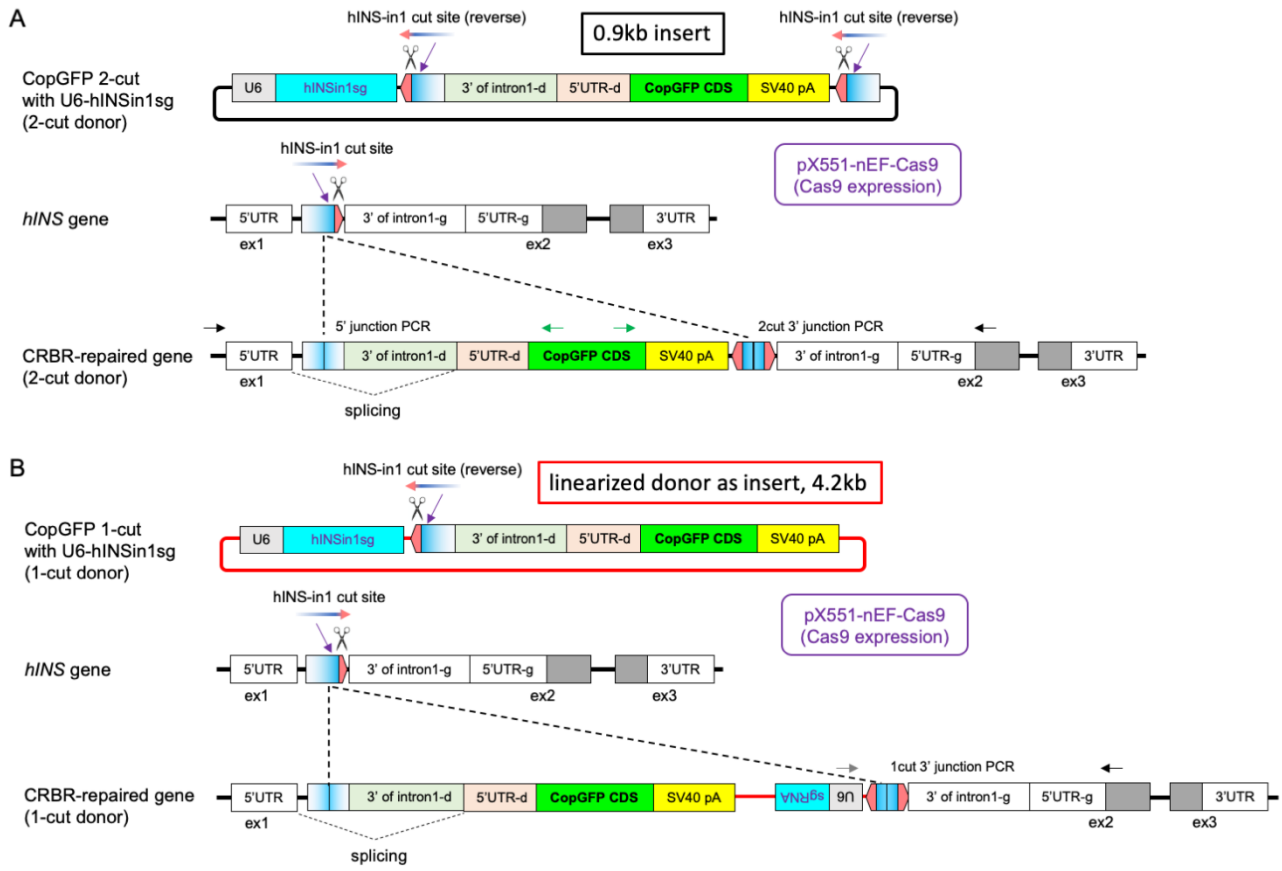
A. Schematic of CRBR AAV vectors used in AAV delivery to Cas9-EGFP mice or wild-type mice. AAV vector provides the same EGFP CRBR cassette as in the EGFP-2cut donor plasmid but also includes a U6-driven mIns2utr5-sgRNA. Cas9 expression was considered to be introduced in all tissues under the universal promoter CAG in the Cas9-EGFP mice.
B and C. P14 Cas9-EGFP mice from one litter (four males and five females) were administered with 40µl, 20µl or no injection of AAV8-U6-mINS2utr5sg-EGFP-2cut via *r.o.* injection. DNA and RNA from pancreas and liver were harvested 30d post injection. Genomic DNA (B) was tested for 5' junction PCR and 3' junction PCR analysis. EGFP mRNA expression levels (C) from the CRBR-edited *mIns2* gene were measured by using a forward primer targeting *mIns2* 5'UTR and a reverse primer (R1 or R2) targeting EGFP to avoid picking up signals from the endogenous EGFP of the Cas9-EGFP mouse strain. The relative fold changes were quantified by normalizing to mActin first and calculated relative to the no injection control.
D. Eight-week old Cas9-EGFP mice from two litters (litter[a or b], gender is indicated in the figure) were administered with 50µl of AAV-U6-mINS2utr5sg-EGFP-2cut in serotype DJ or 8, or a saline control via tail vein injection. Genomic DNA from pancreas and liver was harvested 35d post injection. CRBR editing at genome level was tested by 5' junction PCR and 3' junction PCR analysis.
E. Six-month old C57BL/6J mice from three litters (litter[a, b or c], gender is indicated in the figure) were administered with 50µl of AAV-U6-mINS2utr5sg-EGFP-2cut with or without 50µl of AAV-nEF-Cas9 in serotype DJ, or saline via tail vein injection. Genomic DNA from pancreas and liver was harvested 35d post injection. CRBR editing at genome level was tested by 5' junction PCR and 3' junction PCR analysis.
For B, D and E, all primers were designed to flank the junction sites. PC, positive control, was genomic DNA from MIN6 cells co-transfected with EGFP-2cut donor and pX459-mINS2utr5sg.

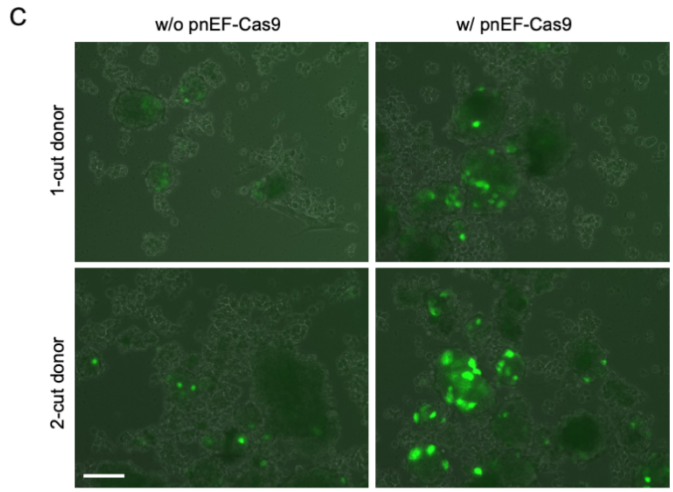
**CRBR-mediated *ex vivo* gene editing in human islets**

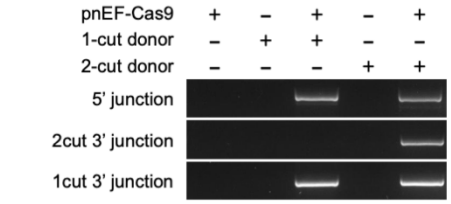
To further validate the CRBR strategy as a potential human gene therapeutic, we similarly targeted GFP to the insulin (*INS*) gene in isolated human islets. Primary human cadaveric islets were transfected or AAV infected with CRBR constructs containing CopGFP (alternative GFP reporter) CDS and targeting the *INS* gene. The CopGFP CRBR cassette was designed to insert into intron 1 between the two exons containing the 5'UTR and upstream of the insulin start codon (Figure 6A). The CRBR cassette contains the 3' half of homologous intronic sequences and the 5'UTR of exon 2 to allow splicing of the endogenous 5'UTR in exon 1 and exogenous 5'UTR in exon 2 from the CRBR cassette. Any unforeseen indels generated during integration/repair will be removed as intronic sequence from the resultant mRNA. We designed and tested both one-cut and two-cut donors to determine which strategy would be more efficacious (Figure 6A-B). The one-cut strategy generates only one insert linearized from the 1-cut donor, with one correct integrant out of two possible

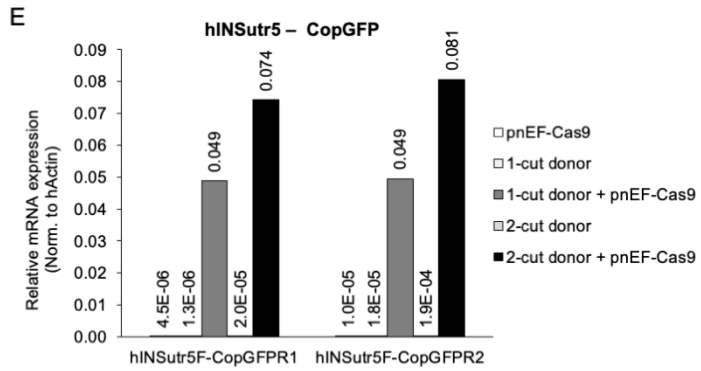
outcomes; whereas the two-cut strategy generates four possible inserts that may be integrated in two orientations, with two correct integrants out of eight possible outcomes (Figure S3). The one-cut strategy requires the integration of a much larger fragment (4.2kb) in contrast to the two-cut strategy (0.9kb CRBR cassette), as it includes the extraneous vector sequences. However, these sequences should not interfere with gene expression because they are downstream of the transcription/translation terminators in the CRBR cassette.

This CRBR-CopGFP strategy was first tested in an easily transfected human cell line, AD293, to determine the optimal sgRNA within intron 1 and to optimize the donor design before testing on human islets (Figure S4A). The CopGFP 1-cut (or 2-cut) donor plasmids were engineered with the U6-hINSin1sg cassette that express the optimal sgRNA. The plasmid expressing SpCas9 and the 1-cut (or 2-cut) donor were co-transfected into human islets. Six days post transfection, many CopGFP-positive cells were observed, which are pancreatic beta cells that possess an active insulin promoter and comprise 45-70% of the total cadaver islet cell population (Figure 6C and Table S1). Typically, about 60% of the cells in a human islet are insulin-secreting beta cells, whereas the remaining 40% secrete other metabolically important peptide hormones[35]. While these cells would be edited with equal frequency compared to beta cells, their insulin promoter is inactive, and we would therefore not expect those cells to express the CopGFP CRBR cassette. The 5' and 3' junction analyses confirmed CRBR editing of the human *INS* locus at the genome level (Figure 6D), and transcription of CopGFP from the human *INS* promoter was also detected (Figure 6E). Furthermore, we observed a modest reduction of human *INS* mRNA expression (Figure 6F), as expected. No biological replicates from the same batch of human islets were analyzed since the samples produce only enough genomic DNA or total RNA for one replicate per treatment. However, CopGFP integration at the genome level, CopGFP transcription, and reduction of human *INS* mRNA expression were seen in all human islet experiments using independent batches of islets (Figure S4B-E). Collectively, these results demonstrate that CRBR-mediate gene correction via plasmid transfection is feasible in human islets if a wild-type coding sequence is targeted downstream of a mutant gene's promoter.



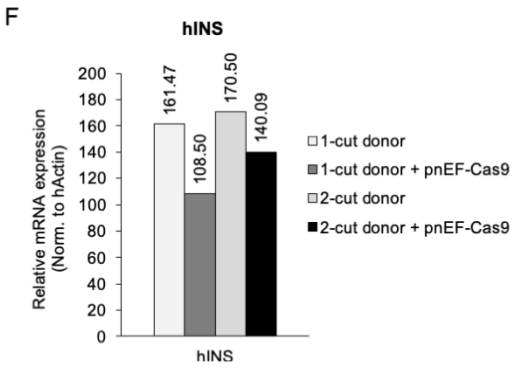



**Figure 6. CRBR-mediated *ex vivo* CopGFP CDS integration in human *INS* gene via plasmid transfection.**

A. Schematic of CRBR-CopGFP-2cut strategy for wild-type *hINS* genome. Donor plasmid provides a 3'intron1-utr5(in exon2)-CopGFP-SV40pA cassette that is flanked by Cas9/gRNA target sites in reverse orientation (5' CCN-20nt 3') as identified within the *hINS* intron 1 (5' 20nt-NGG 3'), and a U6-driven hINSin1-sgRNA. Expression of Cas9 from pnEF-Cas9 and hINSin1-sgRNA from the donor leads to the cleavage of the hINS-in1 cut sites that are engineered in the donor to generate the CRBR cassette, and also a targeted DSB at genomic *hINS* intron 1 between exon 1 and 5'UTR in exon 2.
B. Schematic of CRBR-CopGFP-1cut strategy for wild-type *hINS* genome. The 1-cut donor plasmid is the same as the 2-cut donor except for removing the 3' cut site. Expression of Cas9 and hINSin1-sgRNA leads to the cleavage of the hINS-in1 cut site that is engineered in the donor, linearizing the donor, as well as a targeted DSB at genomic *hINS* intron 1. The 1-cut insert is 4.2kb, much larger than the 2-cut insert which is only 0.9kb. In both 1-cut and 2-cut strategies, correct integration of the CRBR cassette will be retained while incorrect integrant is prone to excision; the 5' junction in the CRBR-edited *hINS* intron 1 should be spliced out and results in a wild-type 5'UTR for normal translation initiation of CopGFP.
C-F. Human cadaveric islets (500 IEQs) were electroporated with 1μg of pnEF-Cas9, 1μg of pU6-hINSin1sg-CopGFP-1cut, 1μg of pU6-hINSin1sg-CopGFP-2cut, or either donor in combination with pnEF-Cas9 using Neon transfection system. Six days post transfection, human islets were imaged (C) as live cultures at 10× objective, scale bar, 100μm. Genomic DNA (D) was harvested for 5' junction PCR, 2cut 3' junction PCR, and 1cut 3' junction PCR analysis. Primers were designed to flank the junction sites. CopGFP mRNA expression levels from the CRBR-edited *hINS* gene (using a forward primer targeting *hINS* 5'UTR and a reverse primer [R1 or R2] targeting CopGFP) (E) and *hINS* mRNA expression levels (F) were quantified by normalizing to hActin. Donor RRID: SAMN14255441.

Previous reports of AAV transduction of human islets has shown limited success[36, 37]. However, our previous success in using AAV to edit the insulin gene in the mouse pancreas (Figure 5B-D) motivated us to evaluate various serotypes of AAV for their ability to deliver CRBR components into human islets and edit the human insulin gene. We tested GFP overexpressing AAV serotypes 2, 5, 6, 8, 9, EB, and DJ for their ability to transduce human islets, and found that AAV-DJ infection led to the most GFP-positive cells (Figure S5A). To test the ability of CRBR-mediated gene editing in human islets *ex vivo* via AAV-DJ transduction, human islets were co-incubated with AAV-DJ-sgRNA-CDS-1cut (or 2cut) (Figure 7A) along with AAV-DJ-Cas9. CopGFP-positive cells were observed at 6-day post infection (Figure 7B). By 10- and 16-day post infection, these cells dramatically increased in both number and fluorescence intensity (Figure 7B and S5B). This indicates that living and functional human beta cells can maintain insulin expression for up to 16-day post infection and that CRBR-editing via AAV transduction did not affect beta cell proliferation *ex vivo*. When CRBR integration was

analyzed at the genome level, the expected 5' junction PCR was observed, however a few larger fragments were also amplified (Figure 7C). DNA sequence analysis revealed that the larger fragments contained the left ITR and U6-driven hINS-in1 sgRNA cassette, which could still be spliced out, resulting in a wild-type 5'UTR for normal translation initiation. The transcription of CopGFP from the human *INS* locus was measured at 18-day post infection (Figure 7D), with that of the one-cut strategy slightly exceeding that of the two-cut strategy. The 1-cut AAV integration size (1.1kb) is not much bigger than the insert from the 2-cut AAV strategy (0.9kb). Therefore, AAV transduction using 1-cut CRBR donor appears to be more efficient. In conclusion, these results indicate CRBR-mediated gene editing via AAV transduction works effectively with human host DNA repair machinery and that AAV serotype DJ is a promising candidate vector for gene therapy in human pancreatic beta cells.

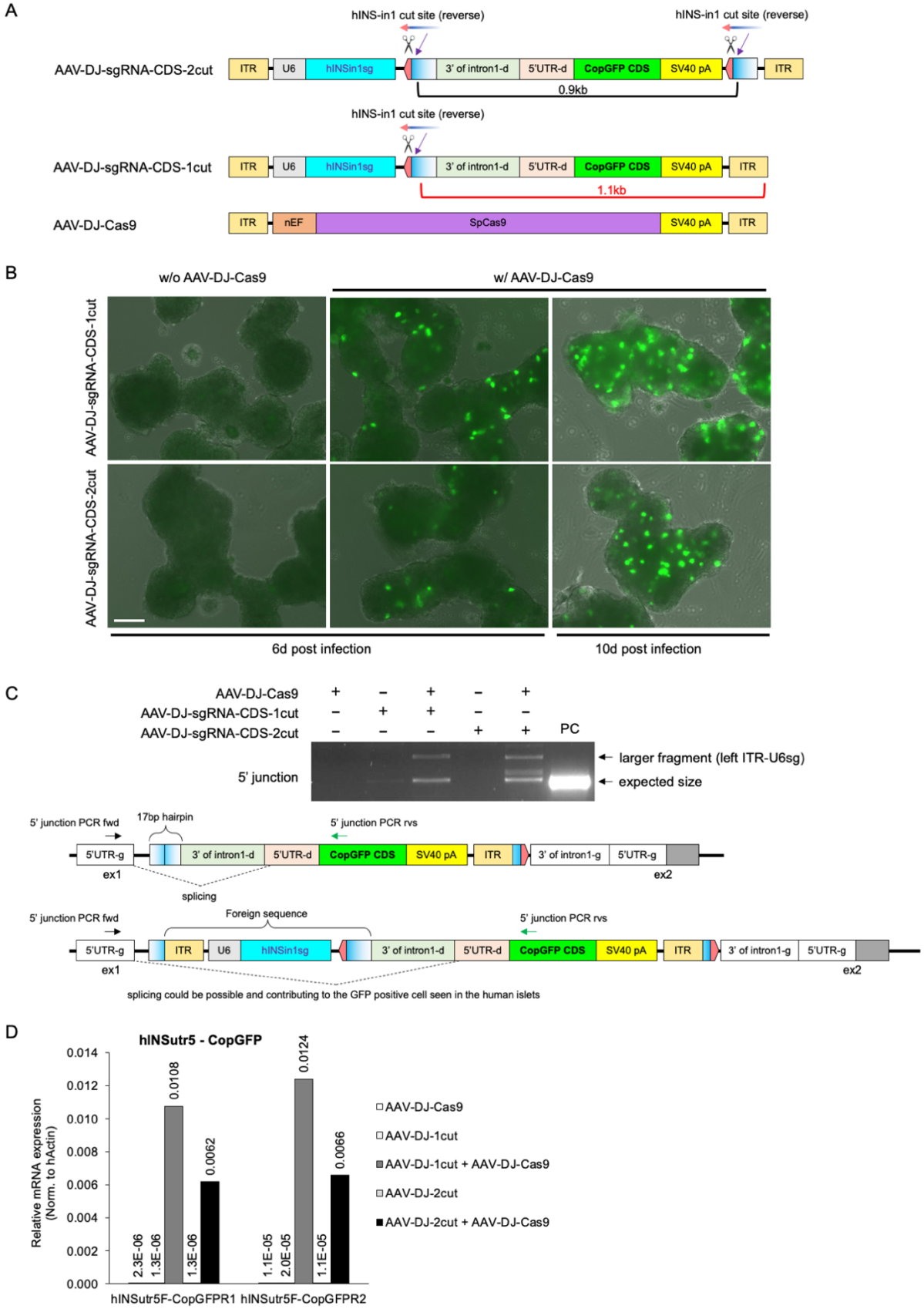



**Figure 7. CRBR-mediated *ex vivo* CopGFP CDS integration in human *INS* gene via AAV-DJ transduction.**

A. Schematic of CRBR AAV vectors used in the CopGFP-2cut and CopGFP-1cut strategies for wild-type *hINS* genome targeting.
B and C. Human cadaveric islets (300 IEQs) were incubated with AAV-DJ-nEF-Cas9, AAV-DJ-U6-hINSin1sg-CopGFP-1cut, AAV-DJ-U6-hINSin1sg-CopGFP-2cut, or either donor AAV vector in combination with AAV-DJ-nEF-Cas9 at 60000 MOI. Human islets were imaged (B) 6d and 10d post infection as live cultures at 10× objective, scale bar, 100μm. Genomic DNA (C) was harvested 16d post infection for 5' junction PCR. Primers were designed to flank the 5' junction site. PC, positive control, was genomic DNA from AD293 cells co-transfected with CopGFP-2cut donor and pX459-hINSin1sg. Donor RRID: SAMN15518672. Resultant genome diagrams show two possible AAV-1cut integrations: expected 5' junction generates a nascent mRNA with a 17bp hairpin which will be spliced out; in the case of Cas9/sgRNA cleavage failure, the whole AAV vector integrant will generate a nascent mRNA with the left ITR-U6sg in the intronic region, which could also be spliced out.
D. A different batch of human cadaveric islets were treated the same as B and C, and RNA was harvested 18d post infection. CopGFP mRNA expression levels from the CRBR-edited *hINS* gene were quantified by normalizing to hActin. Donor RRID: SAMN15314807.

## Discussion

Overexpression of a deficient gene via systemic injection of a recombinant adeno-associated virus (rAAV) derived vector was previously shown to provide a safe and efficient approach for gene therapy[38, 39]. Delivering CRISPR-based therapeutics using AAV vectors[40] has been the favored approach for targeted gene correction *in vivo* in mitotically active tissues. Studies[41, 42] aimed at improving efficiency of HDR in post-mitotic cells offer one solution, however the NHEJ-based repair pathway has provided an alternative strategy that is feasible in both mitotic and post-mitotic cells. Three groups independently[43-45] employed a NHEJ-based strategy to excise an exon of the Duchenne muscular dystrophy gene (*Dmd*) containing a deleterious mutation, which reversed muscular dystrophy in mice. However, the *Dmd* gene is atypical in its tolerance for exon loss, therefore, this strategy cannot be generalized to most other mutations. Consequently, a gene editing strategy that can repair a spectrum of mutations without requiring the design and testing of a specific repair template for each mutation is highly desirable.

Here we describe a Co-opting Regulation Bypass Repair (CRBR) strategy that can be generalized to different kinds of monogenic diseases where traditional treatments or current gene therapy are not feasible or practical. The complete wild-type CDS used in CRBR strategy targets a non-coding region

between the promoter and the downstream mutated region, thereby bypassing any mutation that may exist in the coding sequence. Once validated, the CRBR repair cassette should be able to rescue any deleterious or loss-of-function mutation that might exist in that gene. Currently, the efficiency of CRBR may be too low to directly repair genetic diseases systemically in humans where a large fraction of an organ or tissue may require repair to restore normal function. However, CRBR should be a highly effective method to repair mutations in autologous cells and tissues for transplantation back into the patient from which they were derived. Mutations in *Perk*, which result in severe and permanent neonatal diabetes, present a particularly difficult challenge because very few beta cells exist due to a severe postnatal cell proliferation defect[27] and a block in proinsulin trafficking and processing[46]. Consequently, there may not be enough beta cells present in a WRS patient's islets to repair. A more promising route would be to derive patient specific induced pluripotent stems cells (PS-iPSCs) from a WRS patient, perform CRBR gene repair, screen for CRBR corrected PS-iPSCs, and differentiate them into functional beta cells using the Maxwell protocol[47]. These beta cells could then be transplanted back into the original patient. Repairing a defective gene in a patient's own cells would avoid transplantation rejection and the use of immunosuppressive drugs. Overall, CRBR gene repair combined with autologous cell replacement therapy should be generally applicable to a wide range of human genetic diseases.

While CRBR gene repair offers significant advantages, there are potential pitfalls that must be considered in the design and execution. Because CRBR relies upon the error-prone NHEJ repair pathway, small indels at the integration site of the CRBR cassette are common. It is thus important to restrict the integration site to non-coding and non-regulatory sequences. Ideally, the integration site should be either in the 5'UTR or within an intron upstream of the coding sequence of the subject gene. The introduction of translational start codons or strong secondary mRNA structure (Figure S6) in the 5' UTR and alternative splice sites in an intron must also be avoided. However, because the nature of the indels at the integration site cannot be predetermined, mutations may be generated that result in alternative translational and splicing regulatory sequences that interfere with normal gene expression. We and others have found that a small set of specific indels will be generated for any

given CRISPR-Cas9 experiment (data not shown). Therefore, testing the design in cell culture first can help identify the specific array and frequency of indels that are likely to occur. If necessary, the design may be modified to avoid mutations that interfere with gene expression and regulation. Alternatively, if a gene repair autologous cell replacement (GR-ACR) strategy is used, a specific cell line can be clonally isolated that is devoid of interfering mutations.

Although other delivery methods[48, 49] can be used, rAAV vectors are currently the safest delivery vectors for *in vivo* genome editing. However, AAV vectors have a limited packaging capacity of 4kb. The CRBR strategy, which necessitates delivery of a large multi-element cassette (5'UTR/intronic sequences, CDS with stop codon, and heterologous polyA signal/ transcriptional terminator), will be constrained by this size limitation for viral packaging as well as genomic integration efficiency. Fortunately, about 95% of human proteins are encoded by genes that are less than 4kb. For genes that exceed 4kb, a partial CRBR CDS can be designed for integration into introns upstream of the defective coding exons. Whether or not the integration of a partial CDS cassette will provide a general solution for repairing a spectrum of mutations that exist among patients with a genetic disease depends upon the distribution of the mutations across the coding sequence.

To reduce the size of the CRBR repair cassette, the intronic sequences separating the CDS exons are excluded. However, this approach could be problematic for rare cases where alternative spliced transcripts are essential for normal gene function. In addition, important transcriptional regulatory elements such as enhancers may exist within intronic sequences and would be absent in the CRBR CDS-terminator cassette. In most cases this should not pose a problem since these cis-acting regulatory elements would still exist downstream in the endogenous mutant gene and could still potentially serve to regulate gene transcription. As with all gene therapy strategies, thoroughly testing repair efficacy in cell culture and/or model organisms is essential. A distinct advantage of CRBR gene repair is that, testing and validation need only be performed for a single design which can then be used to repair a spectrum of mutations among a population of human patients thus substantially reducing the cost of treatment.

## Materials and Methods

### Transgenic mice - *Perk* KO (c.1584C>A; p.Cys528X)

SpCas9 mRNA (5meC, Ψ) was purchased from TriLink (San Diego, CA). *In vitro* transcription and purification of mPERKex9-sgRNA (sequence see Construction of plasmids) was as previously described[50]. Repair template (200nt ssODN, 4 nmole Ultramer DNA Oligo) was purchase from Integrated DNA Technologies (IDT, Coralville, IA) (cagcccccactacagcaagaacatccgcaagaaggaccctatcctcctgctgcactggtggaaggagatattcgggacgatcctgctt**tgA** atcgtGgccacAacGttTatcgtgcgcaggcttttccatcctcagccccacagggtaagatgctctgtcaacctaatgtgcttccaagtggttgc tgtgtaggaaacct).

A nonsense mutation was introduced by a C to A mutation on the ssODN template via HDR, 14bp from the Cas9/sgRNA cleavage. Three synonymous mutations were designed 2bp, 5bp and 8bp from PAM site to prevent re-excision of the HDR repaired genome. SpCas9 mRNA, sgRNA and ssODN were sent to the Harvard Genome Modification Facility for microinjection into C57BL/6J zygotes and implantation into pseudo pregnant females. Fifty-seven individuals survived to weaning age from one injection experiment; thirteen individuals carried the *Perk* KO allele (C528X).

### Transgenic mice – rPerk-CRBR (rPERKmyc integration at 5'UTR of *mPerk*)

SpCas9 protein was purchased from IDT. A synthetic mPERKutr5-sgRNA (target sequence, see Construction of plasmids) was purchased from Synthego (Redwood City, CA). A rPERK CDS with Myc tag at the C-terminus was designed to be integrated into mouse *Perk* 5'UTR region using the CRBR strategy as described in Results. The SpCas9 protein, sgRNA and rPERKmyc-2cut donor plasmid were sent to Harvard Genome Modification Facility for microinjection into C57BL/6J zygotes and implantation into pseudo pregnant females. Twenty-one individuals survived to weaning age from two injection experiments; one individual carried the CRBR-edited allele (rPERK-CRBR).

### Genetic strains

B6J.129(Cg)-Gt(ROSA)$^{26Sortm1.1(CAG-cas9*,-EGFP)Fezh}$/J (Cas9-EGFP), C57BL/6J (wild-type) and 129S1/SvImJ (wild-type) mice were purchased from the Jackson Laboratory. *Perk* KO allele (Δex7-9) were generated as previously described[26]. *Perk*$^{\Delta ex7\text{-}9/+}$ strain (used to cross with *Perk*$^{rPERK\text{-}CRBR/+}$) was congenic for C57BL/6J. *Perk*$^{C528X/+}$, *Perk*$^{rPERK\text{-}CRBR/+}$ and offspring (Figure 3) were of mixed C57BL/6J and 129S1/SvImJ background. The Cas9-EGFP strain (Figure 5) was of mixed C57BL/6J and 129S1/SvImJ background. Blood glucose was measured from tail blood using OneTouch UltraMIni glucometer (LifeScan, Malvern, PA). Mice were sacrificed by $CO_2$ euthanization. All animal studies

were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of the Pennsylvania State University.

**Construction of plasmids**

The vectors expressing SpCas9 and sgRNAs targeting *mPERK*, *mIns2* and *hINS* genes were cloned into the pX459 plasmid (pSpCas9(BB)-2A-Puro V2.0, (Addgene, Watertown, MA, plasmid #62988, deposited by Feng Zhang) as previously described[51]. Oligonucleotides for BbsI site cloning see Table S2. The Cas9/sgRNA genomic target sequences (20nt + PAM) on sense (+) or antisense strand (-) used in this study include: mPerk-ex9, CCTGCGCACGATGAAGGTCGTGG (-);

mPerk-in6, TAGTTCGGGATCGCCACATGAGG (-);

mPerk-utr5, AGACATCGCCCATTGAGCGAGGG (-);

mIns2-utr5, TGTAGCGGATCACTTAGGGCTGG (-);

hINS-in1 (or hINS-in1-Reverse in Figure S4), GCCCCAGCTCTGCAGCAGGGAGG (+);

hINS-in1-Same (Figure S4), TGGGCTCGTGAAGCATGTGGGGG (+).

Each of these target sequences were determined from 2-3 possible choices (top scored from crispr.mit.edu or benchling.com/crispr/) by Surveyor Assay (IDT) or T7 Endonuclease I (T7E1) Assay (New England Biolabs, Ipswich, MA). To construct rPERKex7-17-2cut (Figure S7), rPERKex7-17CDS-bGHpA were amplified by mega primer adding 3' cut site to the amplicon from pcDNA-rPERK (in house) and TA-cloned into pCR2.1 (Invitrogen, Carlsbad, CA), followed by subcloning of the 3' part of *mPerk* intron 6 and a 5' cut site by PCR amplification into the pCR2.1-rPERKex7-17CDS-bGHpA-3pCUT. A rPERK-2cut was first generated by cloning ITR-mPERKutr5-rPERK-CDS-bGHpA-3x3pCUT-ITR into pBluescript II KS (+) through PciI and SalI (synthesized by GenScript, Piscataway, NJ). The rPERKmyc-2cut (Figure S8), was then generated by cloning mPERK(450bp)-myc from pcDNA-mPERK-9E10 (in house) into rPERK-2cut through SapI and XhoI to replace rPERK(450bp). The 150aa C-terminus is conserved between mPERK and rPERK. The EGFP-2cut (Figure S10) for *mIns2* targeting was generated by cloning ITR-U6-mINS2utr5sg-5pCUT-EGFP-CDS-pA-3pCUT-ITR into pUC57-Kan through EcoRV (synthesized by GenScript). A short (49bp) polyadenylation signal was used as previous described[24]. AAV-U6-mINS2utr5sg-EGFP-2cut in serotype 8 or DJ was packaged using EGFP-2cut. CopGFP-CDS-SV40pA sequence were obtained from Lonza of its pmaxGFP plasmid. The CopGFP-2cut (Figure S11) for *hINS* targeting was generated by cloning ITR-U6-BbsI-scaffold-hINSin1(flipped cut site for sg-Reverse)-CopGFP-CDS-SV40pA-3×3pCUT-ITR into pUC57-Kan through EcoRV (synthesized by GenScript). The CopGFP-1cut was generated by MfeI double digestion to remove the 3×3pCUT from the CopGFP-2cut (Figure S11). The CopGFP-1cut (or 2cut) with U6-hINSin1sg (Figure S12 and S13) was constructed by cloning the hINSin1sg-Reverse

into BbsI site and was then used either in plasmid experiment or to package AAV-DJ-U6-hINSin1sg-CopGFP-1cut (or 2cut). pAAV-nEF-Cas9 was purchased from Addgene (plasmid #87115, deposited by Juan Belmonte) and was used either in plasmid experiment or AAV-nEF-Cas9 packaging in serotype DJ.

**Cell culture**

Mouse embryonic fibroblasts (MEF) cells were immortalized from *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ [52] and *Perk*$^{C528X/C528X}$ mice using a plasmid carrying the SV40 large T antigen (SV40 1: pBSSVD2005, Addgene, plasmid #21826, deposited by David Ron). Following immortalization, MEF cells were maintained in Dulbecco's Modified Eagle Medium, DMEM (Gibco, Gaithersburg, MD) supplemented with 10% fetal bovine serum, FBS (Gemini, West Sacramento, CA) and 1× Penicillin-Streptomycin (Pen-Strep) at 100U/ml-100µg/ml (Gibco). Mouse MIN6 (Dr. Jun-Ichi Miyazaki, Osaka University, Japan) beta cells and human AD293 cells (Agilent, Santa Clara, CA) were cultured under the same conditions as MEF cells. Primary human cadaveric islets were obtained from Prodo Labs of Integrated Islet Distribution Program (IIDP). Upon receipt, islets were transferred from shipping media to CMRL 1066 (Connaught Medical Research Laboratories, Toronto, Canada; purchased from Gibco) supplemented with 10% FBS, 1× Pen-Strep and 2mM L-Glutamine (Gibco) at a concentration of 800-1000 islet equivalents (IEQ) per milliliter in a non-tissue culture treated 6 cm dish and cultured overnight. All cells were cultured in a humidified, 5% $CO_2$ incubator at 37°C.

**Plasmid transfection via electroporation**

*Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ MEF cells were transfected with CRISPR/Cas9 and CRBR donor constructs by electroporation using the MEF 2 Nucleofector Kit (Lonza, Basel, Switzerland), program T-20 in Nucleofector™ 2b Device (Lonza) according to the manufacturer's protocol. MIN6 cells were similarly electroporated using Nucleofector Kit V (Lonza), program G-16. The pmaxGFP plasmid provided in the Nucleofector Kit was used as transfection positive control in all plasmid electroporation experiments. To achieve higher electroporation efficiency, the Neon Transfection system (Invitrogen) was used for the following cells in a 10µl electroporation system (Invitrogen) with no more than 1µg plasmid DNA per 10µl treatment: *Perk*$^{C528X/C528X}$ MEF cells, $1\times10^7$ cells/ml, 1650V, 20ms, 1 pulse; AD293 cells, $5\times10^6$ cells/ml, 1245V, 10ms, 3 pulses; human islets, 500 IEQs/10ul, 1050V, 40ms, 1 pulse. The Neon procedure for electroporation of human islets was adapted from previously described protocols[53, 54]. Briefly, about 1000 IEQs for two replicates of one treatment was transferred to a 1.5 ml tube and centrifuged for 1 min at 100 g and washed with PBS and re-centrifuged. The islets were then incubated with Accutase (Gibco) for 2 min at 37°C to partially dissociate them, and then washed with PBS and resuspended in 20µl R buffer with 2µg of each plasmid DNA needed for the treatment.

About 500 IEQs in 10µl with 1µg plasmid DNA were electroporated with 1 pulse at 1050V for 40ms and then cultured individually in a non-tissue culture treated 24-well plate.

**AAV production and titration**

AAVs carrying hGFAP::Cre and CAG::FLEx-GFP for serotype testing in human islets were as previously described[55]. AAV8-U6-mINS2utr5sg-EGFP-2cut ($6.15\times10^{13}$GC/ml) was produced and purified by Penn Vector Core. AAV-DJ-U6-mINS2utr5sg-EGFP-2cut ($2.92\times10^{12}$GC/ml), AAV-DJ-U6-hINSin1sg-CopGFP-2cut ($1.83\times10^{13}$GC/ml), AAV-DJ-U6-hINSin1sg-CopGFP-1cut ($6.02\times10^{12}$GC/ml), and AAV-DJ-nEF-Cas9 ($3.83\times10^{12}$GC/ml) were produced and purified as described below. Briefly, recombinant AAVs were produced in 293AAV cells (Cell Biolabs, San Diego, CA). Polyethylenimine (PEI, linear, MW 25,000) was used for transfection of triple plasmids: the pAAV vector constructs, pAAV2/8-RC (Penn Vector Core) or pAAV-DJ (Cell Biolabs) and pHelper (Cell Biolabs). Cells were scrapped in their medium and centrifuged, frozen, and thawed four times by placing it alternately in dry ice-ethanol and 37°C water bath, 72hrs post transfection. AAV crude lysate was purified by centrifugation at 54,000 rpm for 1hr in discontinuous iodixanol gradients with a Beckman SW55Ti rotor. The virus-containing layer was extracted, and viruses were concentrated by Millipore Amicon Ultra Centrifugal Filters (Millipore-Sigma, Bedford MA). Virus titers were determined by qPCR according to Addgene protocol.

**AAV transduction of human islets**

AAV-DJ-U6-hINSin1sg-CopGFP-2cut, AAV-DJ-U6-hINSin1sg-CopGFP-1cut and AAV-DJ-nEF-Cas9 were added to 300 IEQs cultured overnight in 200µl CMRL1066 medium with reduce FBS (2%) at a final titer of $9.0\times10^{10}$ GC/ml. If 1 IEQ is considered to be 1000 cells, the AAV incubation of human islets was at 60,000 MOI. CMRL1066 medium with 10% FBS was added to the sample at 1d post infection.

**AAV administration via intravenous injection**

P14 Cas9-EGFP mice were injected with 20µl or 40µl of AAV8-U6-mINS2utr5sg-EGFP-2cut, via retro-orbital (*r.o.*) injection. Eight-week old Cas9-EGFP mice were injected with 50µl of AAV8-U6-mINS2utr5sg-EGFP-2cut or AAV-DJ-U6-mINS2utr5sg-EGFP-2cut, or 50µl saline solution via tail vein injection. Six-month old C57BL/6J mice were injected with 100µl of AAV-DJ mixture (AAV8-U6-mINS2utr5sg-EGFP-2cut, with or without AAV-DJ-nEF-Cas9), or 100µl saline solution via tail vein injection.

**Single cell sorting**

MEF cells and MIN6 cells were single cell sorted according to size configuration or GFP signal using Beckman Coulter MoFlo Astrios (Beckman-Coulter, Brea, CA) performed by Flow Cytometry Facility at the Huck Institutes of the Life Sciences at Penn State University. Cells were dissociated using 0.25% Trypsin-EDTA solution for 5 min at 37°C and warm DMEM medium supplemented with 10% FBS was added to stop digestion. Cells were then transferred into a 15ml tube and centrifuged at 200g for 1 min at room temperature. The cells were re-suspended thoroughly in DMEM medium with 1× Pen-Strep as single cells and were sorted into 96-well plate with full DMEM medium.

**Genomic DNA extraction and diagnostic PCR analysis**

Genomic DNA was extracted from cell culture or mouse tissue by digesting in lysis buffer (5mM EDTA, 0.2% SDS, 200mM NaCl, and 100mM Tris-HCl, pH8.5) with 100μg/ml proteinase K overnight at 50°C and then precipitated with 1 volume of isopropanol and dissolved in TE buffer (10mM Tris-HCl, 1.0mM EDTA, pH8.0). Blood DNA was extracted using Monarch Genomic DNA Purification Kit (New England Biolabs). Diagnostic PCRs were performed using GoTaq Master Mix (Promega, Madison, WI). Five percent of DMSO was added to improve amplification of GC-rich sequences. PCR product purification was carried out using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). Gel purification to recover PCR fragments after electrophoretic separation was performed using the Zymoclean Gel DNA Recovery Kit (Zymo, Irvine, CA). Sanger sequencing of the PCR products was performed by Genomics Core Facility at the Huck Institutes of the Life Sciences at Penn State University. DNA sequencing results were analyzed using the SnapGene software. See Table S3 for primer sequences.

**RNA isolation and quantitative PCR analysis**

Total RNA from cell lines and mouse tissues other than pancreas were extracted using the Quick-RNA Miniprep Kit (Zymo). Pancreas RNA was extracted as previously described by Robert C. De Lisle (10.3998/panc.2014.9). Human islet RNA was extracted using AllPrep DNA/RNA/Protein Mini Kit (Qiagen). Reverse transcription was performed using qScript cDNA SuperMix (Quanta, Beverly, MA). Quantitative mRNA measurement was carried out using PerfeCTa SYBR Green SuperMix ROX (Quanta) with the StepOnePlus Real-time PCR system (Applied Biosystems, Foster City, CA). Gene expression levels were normalized to endogenous mouse Actin (*Actb*) or human Actin (*ACTA1*) levels of the same sample. The relative fold change in expression was calculated using the ΔΔCt method. See Table S4 for primer sequences.

**GFP imaging and histological analysis**

MIN6 cells and human islets were imaged as live cultures using the FITC and Transillumination channels of the ECHO Revolve microscope (Echo Labs, San Diego, CA). Fluorescence images were

captured with the ECHO Revolve microscope. Whole pancreata were harvested as previously described[56]. Slides were dewaxed and Hematoxylin and Eosin stained by Leica Autostainer ST5010 XL (Wetzlar, Germany) and bright field images were captured with the ECHO Revolve microscope.

**Immunoblot analysis**

Total cell lysates were made from mouse pancreatic tissue using RIPA buffer (1% Nonidet P40, 0.5% sodium doxycholate, 0.1% SDS, 1× PBS, pH 8.0) with 1× Protease Inhibitor cocktails and 1× Phosphatase Inhibitor cocktail 2 and 3 (Millipore-Sigma). Tissue lysates or MEF cells were denatured by boiling the lysates in 2× SDS sample buffer prior to electrophoresis on NuPAGE 8% Bis-Tris Midi gel (Invitrogen). The separated proteins were transferred to nitrocellulose membranes (0.45 µm, Thermo Scientific, Waltham, MA) in carbonate transfer buffer using wet transfer conditions (Criterion Blotter, BioRad, Hercules, CA). Primary antibodies (diluted in 5% BSA-TBST) used include: Phospho-PERK (Thr980) (#3179, Cell signaling, Danvers, MA), PERK (#3192, Cell Signaling), Phospho-eIF2α (Ser51) (#9721, Cell signaling), eIF2α (#AHO1182, Invitrogen), Myc Tag (#R950-25, Invitrogen) and Actin (#A5060, Millipore-Sigma). Appropriate IRDye-conjugated secondary antibodies were used, and IR fluorescence was detected using the LI-COR Odyssey CLx Imaging System and quantified using the LI-COR Image Studio Software (LI-COR, Lincoln, NE).

**Statistical analysis**

Numerical data were represented as mean +/- SE. Statistical significance was determined using Student's *t*-test, where appropriate.

Author Contributions

Conceptualization, J.H., R.A.B., B.C.M., D.R.C.; Methodology, J.H., R.A.B., D.R.C.; Validation, J.H., R.A.B., B.C.M., D.R.C.; Formal Analysis, J.H., A.L., R.A.B.; Investigation, J.H., R.A.B., A.L., ; Resources, Z.P.; Writing – Original Draft, J.H.; Writing – Review & Editing, J.H., R.A.B., B.C.M., D.R.C.; Visualization – J.H.; Funding Acquisition, J.H., R.A.B., B.C.M., and D.R.C.

Acknowledgments

Human pancreatic islets were provided by the NIDDK-funded Integrated Islet Distribution Program (IIDP) (RRID:SCR _014387) at City of Hope, NIH Grant # 2UC4DK098085. We thank the Harvard Genome Modification Facility for providing the mouse zygote microinjection service (Job 1772 for *Perk*$^{C528X}$ allele generation and Job 2182 for *Perk*$^{CRBR\text{-}rPERKmyc}$ allele generation). We thank the Genomics Core Facility, Flow Cytometry Facility and Microscopy Core Facility at the Huck Institutes of

the Life Sciences at Penn State University for providing the Sanger Sequencing service, Cell Sorting service, and Histology lab equipment for tissue processing. This work was supported by National Institutes of Health Grant R01 DK88140.

## References

1. Wilson, JM (2019). Cycling at the Frontiers of Gene Therapy. *Human gene therapy Clinical development* **30**: 47-49.
2. Dunbar, CE, High, KA, Joung, JK, Kohn, DB, Ozawa, K, and Sadelain, M (2018). Gene therapy comes of age. *Science (New York, NY)* **359**.
3. Lundstrom, K (2018). Viral Vectors in Gene Therapy. *Diseases* **6**.
4. Jinek, M, Chylinski, K, Fonfara, I, Hauer, M, Doudna, JA, and Charpentier, E (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**: 816-821.
5. Cong, L, Ran, FA, Cox, D, Lin, S, Barretto, R, Habib, N*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819-823.
6. Mali, P, Yang, L, Esvelt, KM, Aach, J, Guell, M, DiCarlo, JE*, et al.* (2013). RNA-guided human genome engineering via Cas9. *Science* **339**: 823-826.
7. Cho, SW, Kim, S, Kim, JM, and Kim, JS (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**: 230-232.
8. Yin, H, Xue, W, Chen, S, Bogorad, RL, Benedetti, E, Grompe, M*, et al.* (2014). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nature biotechnology* **32**: 551-553.
9. Yin, H, Song, C-Q, Dorkin, JR, Zhu, LJ, Li, Y, Wu, Q*, et al.* (2016). Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nature biotechnology* **34**: 328-333.
10. Tran, NT, Graf, R, Wulf-Goldenberg, A, Stecklum, M, Strauß, G, Kühn, R*, et al.* (2020). CRISPR-Cas9-Mediated ELANE Mutation Correction in Hematopoietic Stem and Progenitor Cells to Treat Severe Congenital Neutropenia. *Mol Ther*.
11. Ohmori, T, Nagao, Y, Mizukami, H, Sakata, A, Muramatsu, SI, Ozawa, K*, et al.* (2017). CRISPR/Cas9-mediated genome editing via postnatal administration of AAV vector cures haemophilia B mice. *Sci Rep* **7**: 4159.
12. Wang, L, Yang, Y, Breton, CA, White, J, Zhang, J, Che, Y*, et al.* (2019). CRISPR/Cas9-mediated in vivo gene targeting corrects hemostasis in newborn and adult factor IX-knockout mice. *Blood* **133**: 2745-2752.
13. Vagni, P, Perlini, LE, Chenais, NAL, Marchetti, T, Parrini, M, Contestabile, A*, et al.* (2019). Gene Editing Preserves Visual Functions in a Mouse Model of Retinal Degeneration. *Front Neurosci* **13**: 945.
14. Cai, Y, Cheng, T, Yao, Y, Li, X, Ma, Y, Li, L*, et al.* (2019). In vivo genome editing rescues photoreceptor degeneration via a Cas9/RecA-mediated homology-directed repair pathway. *Sci Adv* **5**: eaav3335.
15. Cox, DB, Platt, RJ, and Zhang, F (2015). Therapeutic genome editing: prospects and challenges. *Nat Med* **21**: 121-131.

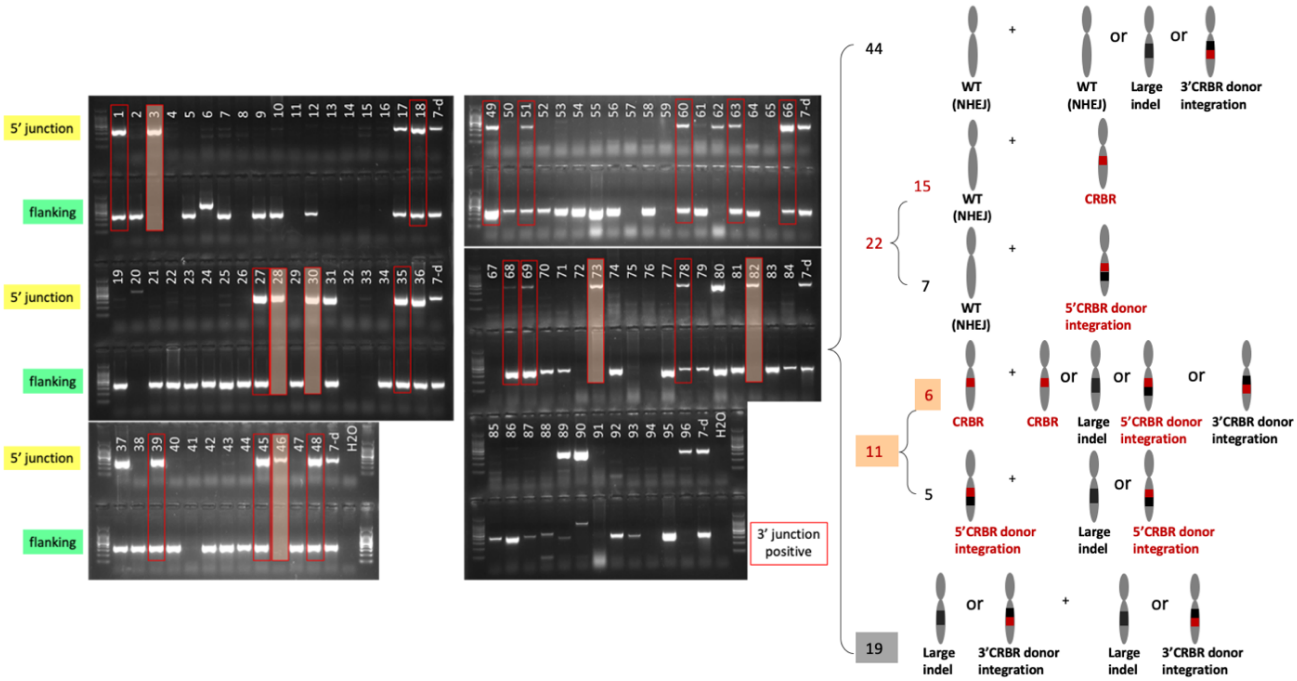
16. Panier, S, and Durocher, D (2013). Push back to respond better: regulatory inhibition of the DNA double-strand break response. *Nat Rev Mol Cell Biol* **14**: 661-672.
17. Komor, AC, Kim, YB, Packer, MS, Zuris, JA, and Liu, DR (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**: 420-424.
18. Gaudelli, NM, Komor, AC, Rees, HA, Packer, MS, Badran, AH, Bryson, DI*, et al.* (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**: 464-471.
19. Yeh, WH, Chiang, H, Rees, HA, Edge, ASB, and Liu, DR (2018). In vivo base editing of post-mitotic sensory cells. *Nat Commun* **9**: 2184.
20. Villiger, L, Grisch-Chan, HM, Lindsay, H, Ringnalda, F, Pogliano, CB, Allegri, G*, et al.* (2018). Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. *Nat Med* **24**: 1519-1525.
21. Bansal, V, Gassenhuber, J, Phillips, T, Oliveira, G, Harbaugh, R, Villarasa, N*, et al.* (2017). Spectrum of mutations in monogenic diabetes genes identified from high-throughput DNA sequencing of 6888 individuals. *BMC Med* **15**: 213.
22. Rebbeck, TR, Friebel, TM, Friedman, E, Hamann, U, Huo, D, Kwong, A*, et al.* (2018). Mutational spectrum in a worldwide study of 29,700 families with BRCA1 or BRCA2 mutations. *Hum Mutat* **39**: 593-620.
23. Julier, C, and Nicolino, M (2010). Wolcott-Rallison syndrome. *Orphanet J Rare Dis* **5**: 29.
24. Suzuki, K, Tsunekawa, Y, Hernandez-Benitez, R, Wu, J, Zhu, J, Kim, EJ*, et al.* (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**: 144-149.
25. Harding, HP, Zeng, H, Zhang, Y, Jungries, R, Chung, P, Plesken, H*, et al.* (2001). Diabetes mellitus and exocrine pancreatic dysfunction in perk-/- mice reveals a role for translational control in secretory cell survival. *Molecular cell* **7**: 1153-1163.
26. Zhang, P, McGrath, B, Li, Sa, Frank, A, Zambito, F, Reinert, J*, et al.* (2002). The PERK eukaryotic initiation factor 2 alpha kinase is required for the development of the skeletal system, postnatal growth, and the function and viability of the pancreas. *Molecular and cellular biology* **22**: 3864-3874.
27. Zhang, W, Feng, D, Li, Y, Iida, K, McGrath, B, and Cavener, DR (2006). PERK EIF2AK3 control of pancreatic beta cell differentiation and proliferation is required for postnatal glucose homeostasis. *Cell metabolism* **4**: 491-497.
28. Li, Y, Iida, K, O'Neil, J, Zhang, P, Li, Sa, Frank, A*, et al.* (2003). PERK eIF2alpha kinase regulates neonatal growth by controlling the expression of circulating insulin-like growth factor-I derived from the liver. *Endocrinology* **144**: 3505-3513.
29. Iida, K, Li, Y, McGrath, BC, Frank, A, and Cavener, DR (2007). PERK eIF2 alpha kinase is required to regulate the viability of the exocrine pancreas in mice. *BMC cell biology* **8**: 38.
30. Hellen, CU, and Sarnow, P (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* **15**: 1593-1612.
31. Gunisova, S, Hronova, V, Mohammad, MP, Hinnebusch, AG, and Valasek, LS (2018). Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. *FEMS Microbiol Rev* **42**: 165-192.
32. Cheng, H, Wolfe, SH, Valencia, V, Qian, K, Shen, L, Phillips, MI*, et al.* (2007). Efficient and persistent transduction of exocrine and endocrine pancreas by adeno-associated virus type 8. *J Biomed Sci* **14**: 585-594.

33. Rehman, KK, Trucco, M, Wang, Z, Xiao, X, and Robbins, PD (2008). AAV8-mediated gene transfer of interleukin-4 to endogenous beta-cells prevents the onset of diabetes in NOD mice. *Mol Ther* **16**: 1409-1416.
34. Mulder, NL, Havinga, R, Kluiver, J, Groen, AK, and Kruit, JK (2019). AAV8-mediated gene transfer of microRNA-132 improves beta cell function in mice fed a high-fat diet. *J Endocrinol* **240**: 123-132.
35. Da Silva Xavier, G (2018). The Cells of the Islets of Langerhans. *J Clin Med* **7**.
36. Rehman, KK, Wang, Z, Bottino, R, Balamurugan, AN, Trucco, M, Li, J*, et al.* (2005). Efficient gene delivery to human and rodent islets with double-stranded (ds) AAV-based vectors. *Gene Ther* **12**: 1313-1323.
37. Craig, AT, Gavrilova, O, Dwyer, NK, Jou, W, Pack, S, Liu, E*, et al.* (2009). Transduction of rat pancreatic islets with pseudotyped adeno-associated virus vectors. *Virol J* **6**: 61.
38. Park, K, Kim, WJ, Cho, YH, Lee, YI, Lee, H, Jeong, S*, et al.* (2008). Cancer gene therapy using adeno-associated virus vectors. *Front Biosci* **13**: 2653-2659.
39. Kattenhorn, LM, Tipper, CH, Stoica, L, Geraghty, DS, Wright, TL, Clark, KR*, et al.* (2016). Adeno-Associated Virus Gene Therapy for Liver Disease. *Human gene therapy* **27**: 947-961.
40. Wang, D, Zhang, F, and Gao, G (2020). CRISPR-Based Therapeutic Genome Editing: Strategies and In Vivo Delivery by AAV Vectors. *Cell* **181**: 136-150.
41. Canny, MD, Moatti, N, Wan, LCK, Fradet-Turcotte, A, Krasner, D, Mateos-Gomez, PA*, et al.* (2018). Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR-Cas9 genome-editing efficiency. *Nat Biotechnol* **36**: 95-102.
42. Nishiyama, J, Mikuni, T, and Yasuda, R (2017). Virus-Mediated Genome Editing via Homology-Directed Repair in Mitotic and Postmitotic Cells in Mammalian Brain. *Neuron* **96**: 755-768.e755.
43. Long, C, Amoasii, L, Mireault, AA, McAnally, JR, Li, H, Sanchez-Ortiz, E*, et al.* (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science (New York, NY)* **351**: 400-403.
44. Nelson, CE, Hakim, CH, Ousterout, DG, Thakore, PI, Moreb, EA, Castellanos Rivera, RM*, et al.* (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science (New York, NY)* **351**: 403-407.
45. Tabebordbar, M, Zhu, K, Cheng, JKW, Chew, WL, Widrick, JJ, Yan, WX*, et al.* (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science (New York, NY)* **351**: 407-411.
46. Sowers, CR, Wang, R, Bourne, RA, McGrath, BC, Hu, J, Bevilacqua, SC*, et al.* (2018). The protein kinase PERK/EIF2AK3 regulates proinsulin processing not via protein synthesis but by controlling endoplasmic reticulum chaperones. *The Journal of biological chemistry* **293**: 5134-5149.
47. Maxwell, KG, Augsornworawat, P, Velazco-Cruz, L, Kim, MH, Asada, R, Hogrebe, NJ*, et al.* (2020). Gene-edited human stem cell-derived β cells from a patient with monogenic diabetes reverse preexisting diabetes in mice. *Sci Transl Med* **12**.
48. Wilbie, D, Walther, J, and Mastrobattista, E (2019). Delivery Aspects of CRISPR/Cas for in Vivo Genome Editing. *Acc Chem Res* **52**: 1555-1564.
49. Yin, H, Song, CQ, Suresh, S, Wu, Q, Walsh, S, Rhym, LH*, et al.* (2017). Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. *Nat Biotechnol* **35**: 1179-1187.
50. Yang, H, Wang, H, and Jaenisch, R (2014). Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat Protoc* **9**: 1956-1968.

51. Ran, FA, Hsu, PD, Wright, J, Agarwala, V, Scott, DA, and Zhang, F (2013). Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**: 2281-2308.
52. Jiang, HY, Wek, SA, McGrath, BC, Scheuner, D, Kaufman, RJ, Cavener, DR*, et al.* (2003). Phosphorylation of the alpha subunit of eukaryotic initiation factor 2 is required for activation of NF-kappaB in response to diverse cellular stresses. *Mol Cell Biol* **23**: 5651-5663.
53. Tamaki, S, Nye, C, Slorach, E, Scharp, D, Blau, HM, Whiteley, PE*, et al.* (2014). Simultaneous silencing of multiple RB and p53 pathway members induces cell cycle reentry in intact human pancreatic islets. *BMC Biotechnol* **14**: 86.
54. Lefebvre, B, Vandewalle, B, Longue, J, Moerman, E, Lukowiak, B, Gmyr, V*, et al.* (2010). Efficient gene delivery and silencing of mouse and human pancreatic islets. *BMC Biotechnol* **10**: 28.
55. Chen, YC, Ma, NX, Pei, ZF, Wu, Z, Do-Monte, FH, Keefe, S*, et al.* (2020). A NeuroD1 AAV-Based Gene Therapy for Functional Brain Repair after Ischemic Injury through In Vivo Astrocyte-to-Neuron Conversion. *Mol Ther* **28**: 217-234.
56. Gupta, S, McGrath, B, and Cavener, DR (2010). PERK (EIF2AK3) regulates proinsulin trafficking and quality control in the secretory pathway. *Diabetes* **59**: 1937-1947.

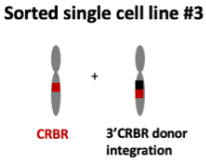
A

| 5' junction and flanking PCR | 96 sorted cells | Flanking PCR | 5' CRBR | 3' CRBR (within 5'CRBR) | Large indels |
|---|---|---|---|---|---|
| WT/NHEJ: only Flanking positive | 44 | 66 | - | - | - |
| Het: 5' CRBR and Flanking positive | 22 | | 33 | 15 (out of 22) | 7 |
| Homo: only 5' CRBR positive | 11 | - | | 6 (out of 11) | 5 |
| Both no band | 19 | - | - | - | 19 |




B

| Cell ID | Mouse *Perk* intron6 at Ch6 | 5' junction PCR | 3' junction PCR | Flanking PCR (WT/small NHEJ) |
|---|---|---|---|---|
| 1 | CRBR/small NHEJ indel | clean joint | 4bp deletion (-4 to -1) | 2bp deletion (GT at +1 to +2) |
| 3 | CRBR/backbone integration | 11bp deletion (-1 to +10) | clean joint (3' seq mixture) | no band |
| 28 | CRBR/large indel | 2bp insertion (TG at +1) | clean joint | no band |
| 35 | CRBR/small NHEJ indel | 1bp mutation (T to C mutation at -1) | clean joint | 2bp deletion (GT at +1 to +2) |
| 46 | CRBR/large integration | 1bp insertion (C at +1) | clean joint | no band |
| 66 | CRBR/small NHEJ indel | clean joint | 4bp deletion (-4 to -1) | 2bp deletion (GT at +1 to +2) |
| 73 | CRBR/large indel | 2bp insertion (TG at +1) | clean joint | no band |
| 82 | CRBR/backbone integration | 2bp insertion (TG at +1) | clean joint (5' seq mixture) | no band |




C


**Figure S1. Allele analysis of single sorted *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ MEF cell lines with CRBR editing.**

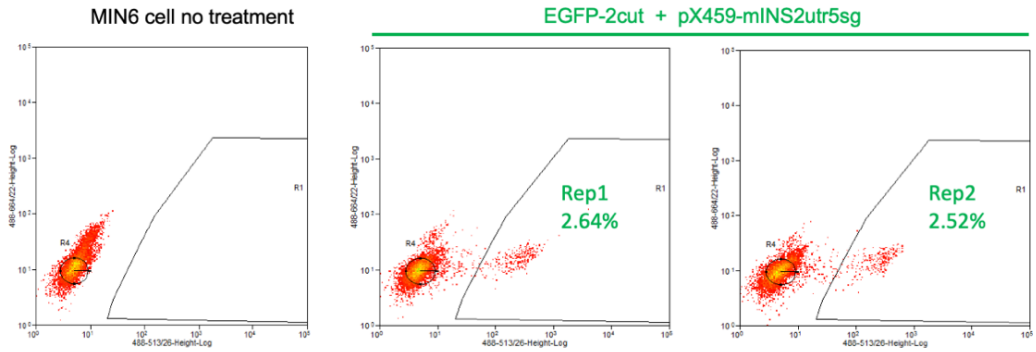
A. Possible composition of the two alleles in the single sorted *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ MEF cell lines according to the PCR analysis. Positive 5' junction PCR in total 96 cell lines showed 33 sorted cells have CRBR editing in at least one allele. Positive flanking PCR showed 22 out of the 33 have the other allele cleanly rejointed or with small NHEJ indels at the DSB in the genome. Positive 3' junction PCR found

in the 33 cell lines (that were 5' junction positive) are highlighted with red rectangle on the DNA gel. 7-d, positive control, was genomic DNA of *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ MEF cells co-transfected with pX459-mPERKin6sg and rPERKex7-17-2cut.

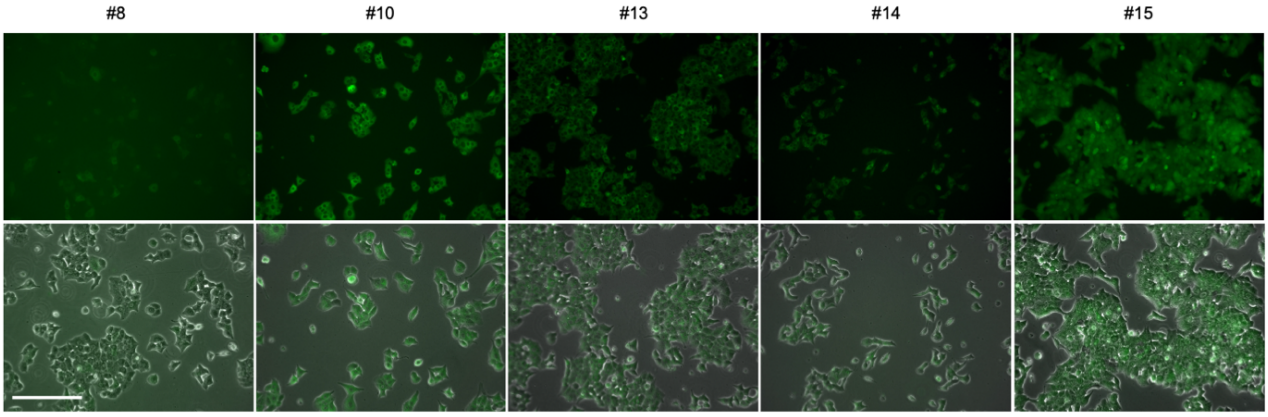
B. Eight cell lines (positive for 5' junction PCR) were subjected to sequence analysis for 5' junction PCR, 3' junction PCR and flanking PCR. Sequencing of PCR product that contained mixed signals on either side of the junction suggests the unknown large indel on the second allele could be a whole donor integration.

C. Alignment of the sequencing of the 5' junction PCR of 8 cell lines to the CRBR edited genome. Except #3 cell line, all other 7 cell lines retained the AG/G from the donor cut site residues, which may be recognized as a splice-acceptor 135bp upstream of the intended slicing site between the 3' of *mPerk* intron 6 and *rPerk* exon 7. The extra 135bp intronic sequence encodes a stop codon in the 7 cell lines that would cause a nonsense-mediated mRNA decay of the chimeric mouse-rat *Perk* transcript.

A




B




C

| Cell ID | Mouse *Ins2* 5'UTR at Ch7 | 5' junction PCR | 3' junction PCR | Flanking PCR (WT/small NHEJ) | Whole plasmid integration | Visible GFP |
|---|---|---|---|---|---|---|
| **8** | CRBR/small NHEJ indel | clean joint | 102bp insertion (btw -1 and +1) | 2bp deletion (+1 to +2) | no band | **mixed; ~5%** |
| **10** | CRBR/small NHEJ indel | clean joint | clean joint | 2bp del + 1bp ins (btw -1 and +1) | no band | **all cells; high** |
| **13** | CRBR/small NHEJ indel | clean joint | 4bp deletion (-1 to +3) | 4bp deletion (+1 to +2) | no band | **all cells; high** |
| **14** | CRBR/small NHEJ indel | 1bp insertion (btw -1 and +1) | clean joint | 2bp deletion (+1 to +2) | no band | **all cells; medium** |
| **15** | CRBR/backbone integration | clean joint | clean joint | no band | 10bp del + 275bp ins at 3' junction | **all cells; high** |

**Figure S2. Single cell sorting of CRBR-edited GFP-positive MIN6 cells.**

A. Single cell sorting revealed about 2.5% of MIN6 cells in the mixed cell population at 21d post co-transfection of EGFP-2cut plasmid and Cas9/sgRNA plasmid had EGFP CRBR cassette integration and GFP protein expression.

B and C. Five single sorted GFP-positive MIN6 cell lines (#8, 10, 13, 14, and 15) were (B) imaged as live cultures at 20× objective, scale bar, 100μm. Same exposure time and light intensity was applied. Sequence analysis (C) for 5' junction PCR, 3' junction PCR, flanking PCR and whole plasmid integration PCR of the same sorted cell lines are described in the table.

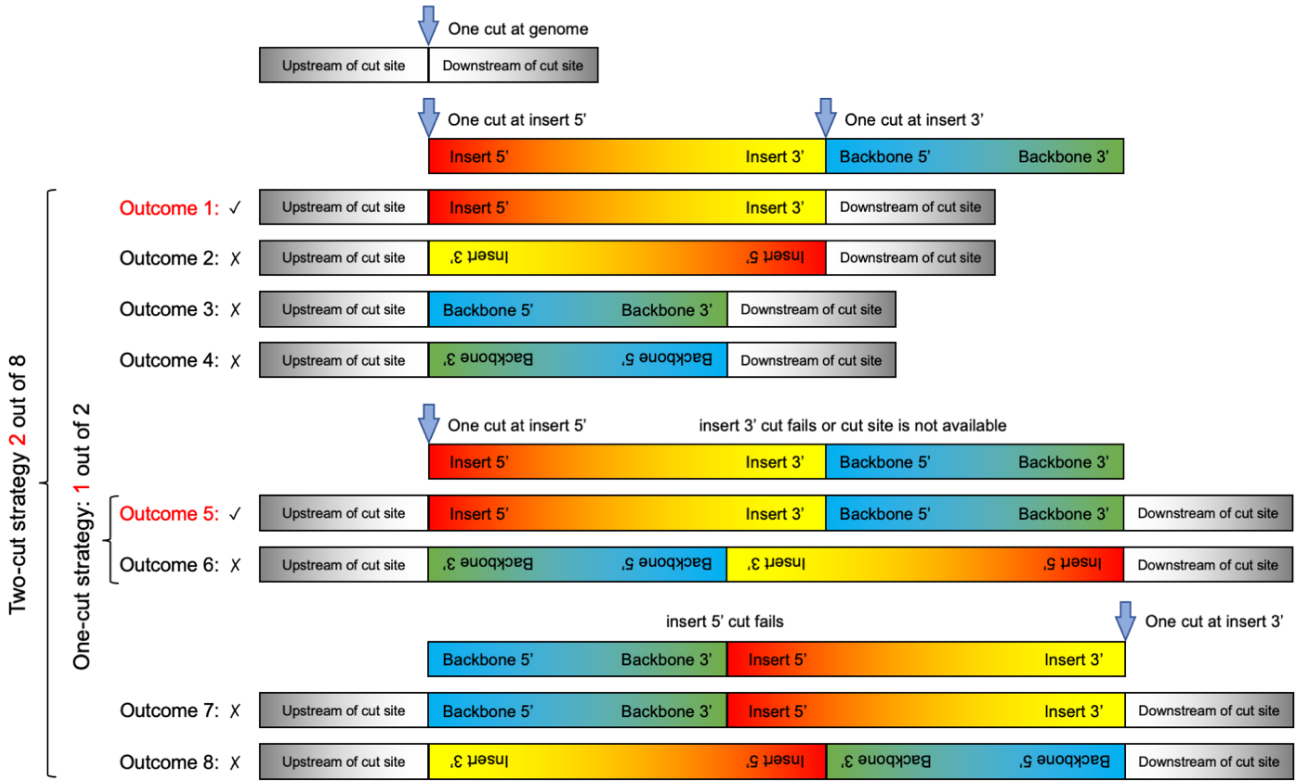



**Figure S3. Schematic of possible integrants from CRBR 2-cut strategy and CRBR 1-cut strategy.**

Plasmid donor with two Cas9/sgRNA target sites (2-cut donor) can generate insert fragment and backbone fragment when both cuts take place, and four integrants (outcome 1~4) could be generated randomly if not considering the engineered reverse-orientated cut site effect. When either cleavage fails, the 2-cut donor could be linearized in two ways and integrated into genome in two orientations (outcome 5~8). Only outcome 1 and 5 would express the CRBR cassette correctly. Plasmid donor with one Cas9/sgRNA target site (1-cut donor) at the 5' of the insert could be integrated as a linearized donor in two orientations, so 1 out of 2 outcomes is the correct integrant for CRBR editing.

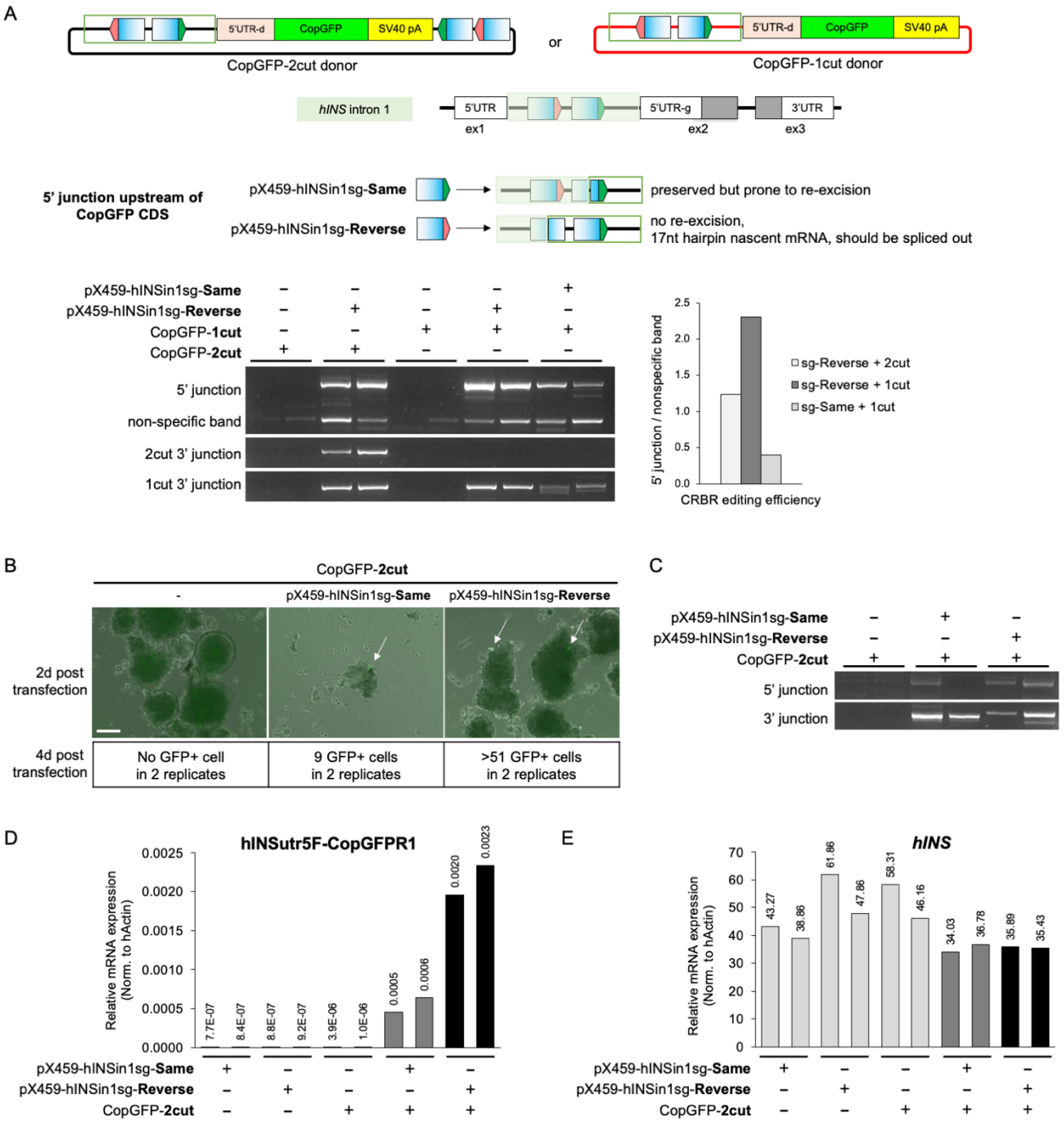



**Figure S4. Reverse-orientated vs same-orientated target sites comparison; 2-cut donor vs 1-cut donor comparison.**

A. Two Cas9/sgRNA targeted sites in *hINS* intron 1 were identified, both of which have NGG on the sense strand (5' 20nt-NGG 3'). The upstream targeted site (red PAM) was engineered in the reverse orientation in the donor (termed as hINSin1sg-Reverse) and the downstream targeted site (green PAM) was engineered in the same orientation in the donor (termed as hINSin1sg-Same). This design allowed the 3' intron 1 downstream of the cut site to be wild-type to the genome, no matter which

sgRNA was finally determined to be used in later study. AD293 cells were electroporated with 1μg of CopGFP-1cut donor in combination with 1μg of pX459-hINSin1sg-Same or pX459-hINSin1sg-Reverse; or pX459-hINSin1sg-Reverse in combination with CopGFP-2cut donor or CopGFP-1cut donor. Genomic DNA was harvested 3d post transfection for 5' junction PCR, 2cut 3' junction PCR and 1cut 3' junction PCR analysis. 5' junction PCR band intensity was calculated relative to the nonspecific band.

B-E. Human cadaveric islets (500 IEQs) were electroporated with 1μg of CopGFP-2cut donor with 1μg of pX459-hINSin1sg-Same or pX459-hINSin1sg-Reverse using 10ul Neon transfection system with two replicates. Two- or four-day post transfection, human islets were imaged (B) as live cultures at 10× objective, scale bar, 100μm. White arrows, GFP positive cells. Five-day post transfection, genomic DNA (C) was harvested for 5' junction PCR and 3' junction PCR analysis. CopGFP mRNA expression levels from the CRBR-edited *hINS* gene (D) and *hINS* mRNA expression levels (E) were quantified by normalizing to hActin. Donor RRID: SAMN13254972.

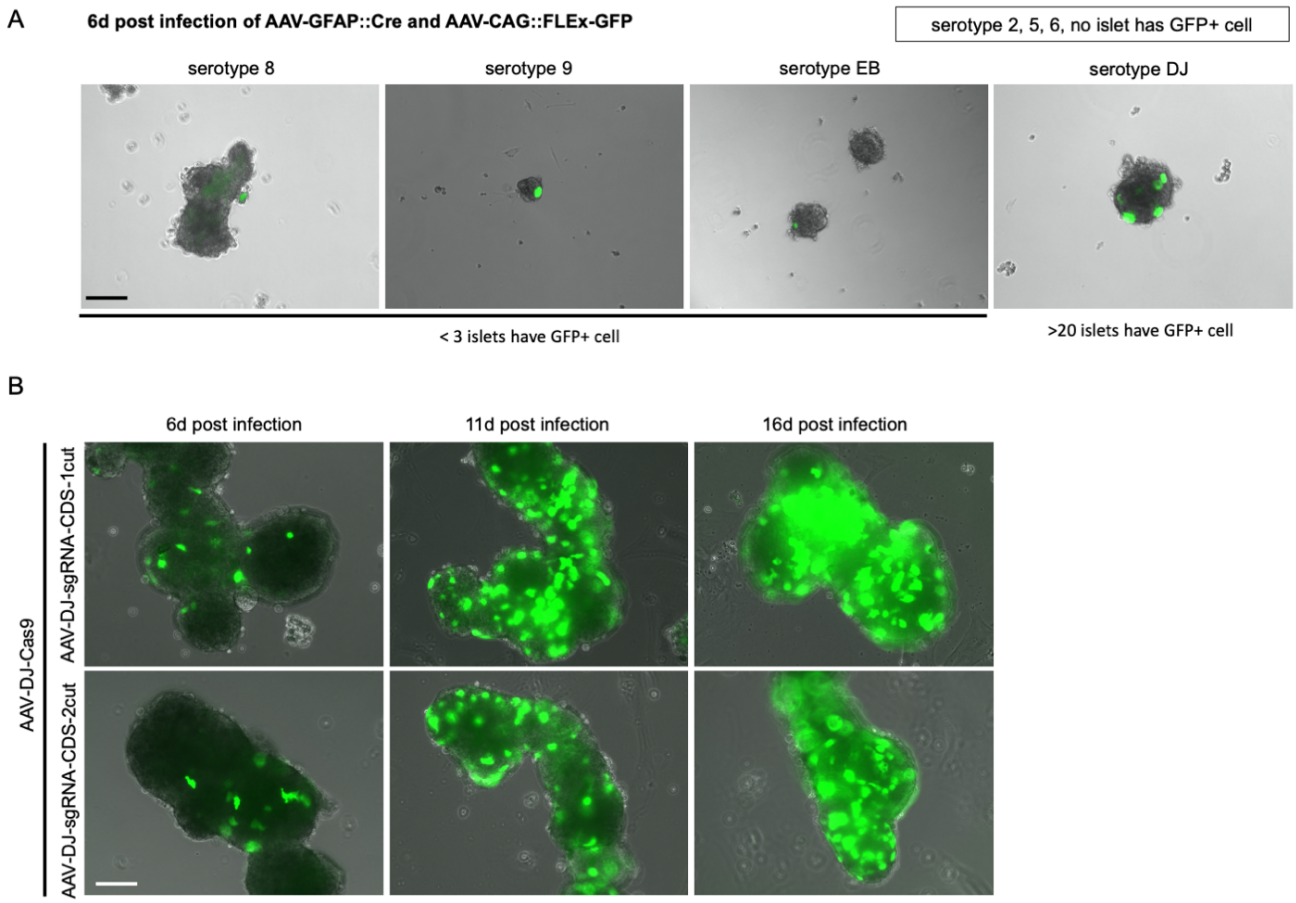



**Figure S5. AAV serotype testing in human islets and long time GFP expression via CRBR-mediated AAV-DJ based gene targeting and expression.**

A. Human cadaveric islets (about 30 IEQs) were co-incubated with AAV-GFAP::Cre and AAV-CAG::FLEx-GFP in serotype 2, 5, 6, 8, 9, EB, and DJ, MOI>60000. Human islets were imaged 6d post infection as live cultures at 10× objective, scale bar, 100μm. Donor RRID: SAMN13515839.

B. Human cadaveric islets (about 30 IEQs) were co-incubated with AAV-DJ-nEF-Cas9 and AAV-DJ-U6-hINSin1sg-CopGFP-1cut or AAV-DJ-U6-hINSin1sg-CopGFP-2cut, MOI>60000. Human islets were imaged 6d, 11d, and 16d post infection as live cultures at 10× objective, scale bar, 100μm. Same exposure time and light intensity was applied. Donor RRID: SAMN14255441.

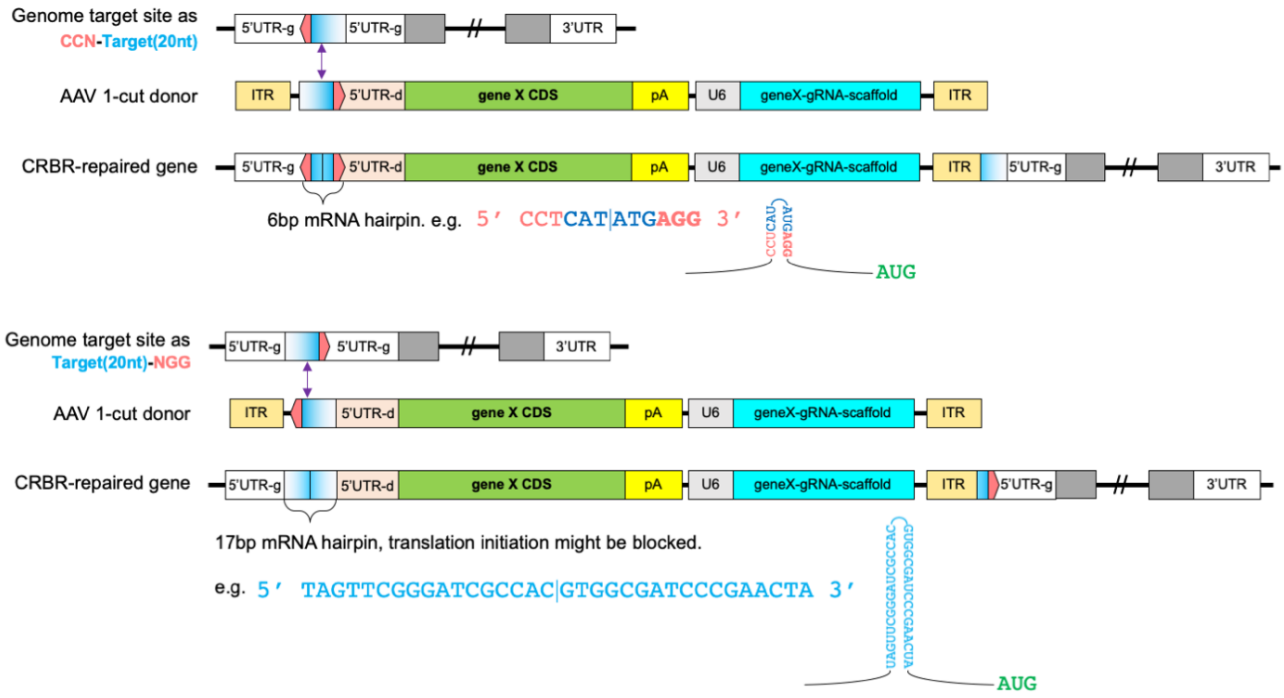
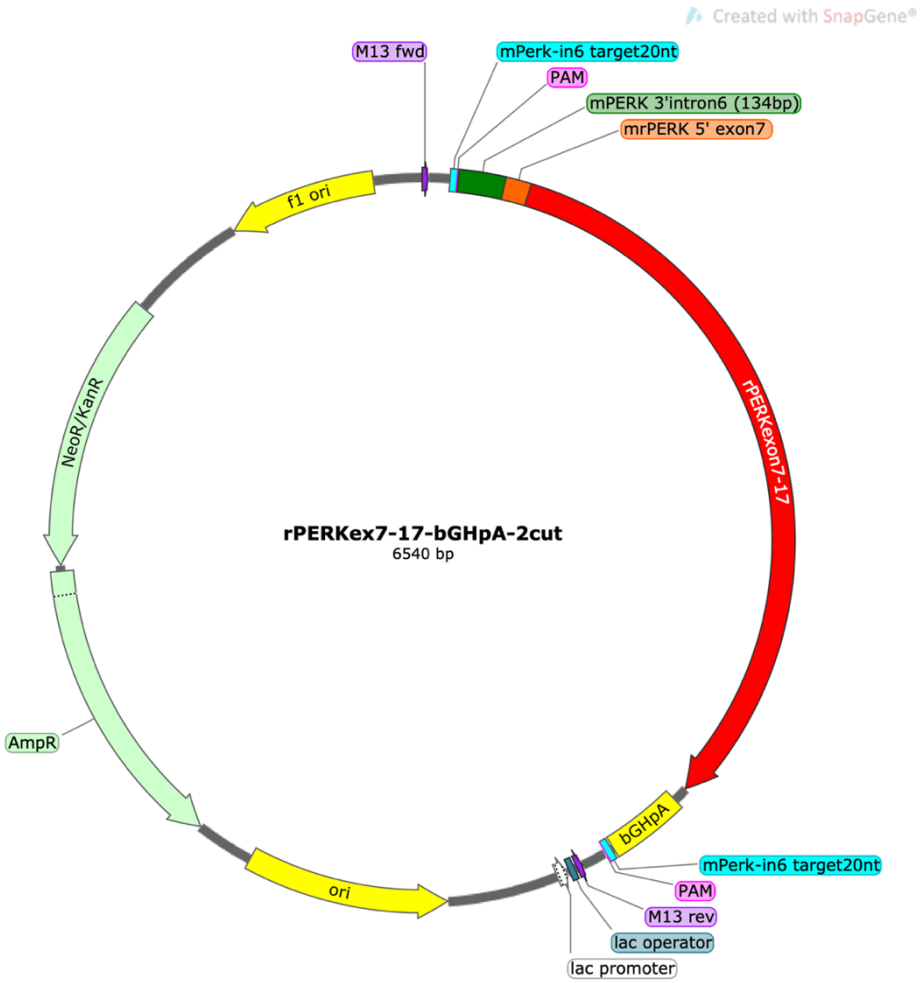



**Figure S6. Two scenarios of genome target site and corresponding result genome using reverse-orientated target site in the donor.**

Top scenario, when the PAM site (NGG) of the genome target site is located on the antisense strand (5' CCN-20nt 3'), the result genome will have a 6bp mRNA hairpin and probably will not block translation initiation; Bottom scenario, when the PAM site (NGG) of the genome target site is located on the sense strand (5' 20nt-NGG 3'), the result genome will have a 17bp mRNA hairpin which may block the translation initiation. In a EGFP targeting human *INS* gene experiment (data not shown), translation of an EGFP cassette was not observe in the human islets, even though CRBR editing at both genome level and transcription level were confirmed, whereas a similar CRBR strategy was successful in EGFP targeting mouse *Ins2* gene. It is believed that the 17bp hairpin might be blocking the translation, because the only difference between the CRBR targeting of *hINS* and *mIns2* is that the *hINS* strategy generated a 17bp hairpin and *mIns2* generated a 6bp hairpin in the mRNA transcript very close to the start codon.




>pCR2.1-3'mPERKintron6-rPERKexon7-17-bGHpA-2cut (6540 bp)

GTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGCCCTCTAGATGCATGCTCGAGCGGCCgc
TAGTTCGGGATCGCCACATGAGGAATGCTTTGAGCATCTTCCCTGGAAAGAGCTATGTAGCAAAGGAATAACATTTGTCT
AAGTGCATTAAATCAGAGCAAACTTTTAAGAACTCCATCATGTGACTGACATGTTTTTGACCATGTGGTTGACCCAGGGA
TGTACAGAGGCCAGCTGTACCTGCAGTCGTCCGTCAGGGTCTCAGAAAAGTTCCCTACAAGCCCAAAGGCCTTGGAGTCT
GTAAATGGAGAAAGTGCCATTATTCCTCTGCCGACGATCAAATGGAAGCCCTTAATCCATTCTCCGTCTAGGACTCCTGT
CTTGGTTGGGTCTGATGAATTTGACAAGTGTCTAAGTAATGATAAATATTCCCATGAAGAATACAGTAATGGTGCGCTTT
CAATCCTACAGTATCCATATGACAACGGTTACTATCTGCCTTATTACAAGAGAGAAAGGAATAAGCGGAGCACGCAGATC
ACGGTCAGGTTCCTGGACAGCCCGCACTACAGCAAGAACATCCGCAAGAAGGACCCCATCCTCCTCCTGCACTGGTGGAA
GGAAATATTTGGAACGATCTTGCTTTGCATTGTGGCCACGACCTTTATCGTGCGCAGGCTTTTCCATCCTCAGCCCCACA
GGCAACGGAAGGAGTCTGAAACTCAGTGCCAAACTGAAAGTAAATACGACTCTGTGAGTGCTGATAACAGCGACAACAGC
TGGAACGACATTAAACACTCAGGATACGTGTCCCGATACCTAACAGATTTTGAGCCAATTCAGTGCATGGGTCGAGGCGG
CTTTGGCGTTGTCTTTGAAGCTAAGAACAAAGTAGATGACTGCAATTACGCCATCAAGAGGATCCGTCTCCCAAACAGGG
AGCTGGCACGGGAGAAGGTAATGCGGGAAGTCAAAGCCTTGGCTAAGCTGGAGCACCCAGGCATCGTGAGGTATTTCAAT
GCGTGGCTGGAAACACCACCAGAGAAGTGGCAAGAGGAGATGGATGAAATCTGGCTGAAGGACGAAAGCACAGACTGGCC
ACTCAGCTCCCCTAGCCCAATGGATGCCCCATCTGTTAAGATCCGGCAGATGGATCCTTTCTCTACAAAAGAACAGATTG
AAGTCATAGCTCCTTCTCCTGAAAGAAGTCGGTCTTTCTCAGTGGGCATTTCCTGTGGCCGGACAAGCTCATCTGAGAGC
CAGTTCTCTCCCCTGGAGTTCTCAGGGACAGACTGTGGAGACAACAGTGACTCAGAGGACGCAGCCCACAACCTCCAGGA
CAGCTGCCTGACAGATTGCGACATGGAAGATGGTACAGTGGACGGCGATGATGAGGGACACTCCTTTGAACTTTGTCCTT
CTGAAGCTTCTCCCTATACCAGGTCAAGGGAGGGAACGTCCTCTTCCATAGTGTTTGAAGACTCTGGCTGTGATAATGCG
TCCAGTAAGGAGGACCCCAGAATGAATCGGCTGCATAATGGCCACCATTACGTTAATAAGCTAACTGAGTTCAAGCACTC
CAGCAGCAGGTCTTCTTCCGAAGCCACCTTGTCTACCTCTCCTACCAGGCCAACCACCCTAAGTTTAGATTTCACCAGGA
ACACGGTGGACCGGCTCCAGCCCAGCTCCCCCAAGGTGTACCTGTACATCCAGATGCAGCTGTGCAGGAAGGAGAACCTT
AAAGACTGGATGAACCGGCGCTGCAGCATGGAGGACAGGGAGCACAGAGTGTGCCTGCACATCTTTCTGCAGATCGCAGA
GGCAGTGCAGTTTCTGCACAGCAAGGGACTCATGCACAGGGACCTCAAGCCTTCCAACATATTCTTCACAATGGATGATG
TGGTCAAAGTTGGGGACTTTGGGCTAGTGACTGCAATGGACCAGGATGAAGAAGAGCAGACAGTACTGACTCCAATGCCA
GCCTATGCTACACACACGGGACAAGTCGGGACCAAGCTGTACATGAGCCCAGAACAGATTCATGGAAACAACTACTCCCA

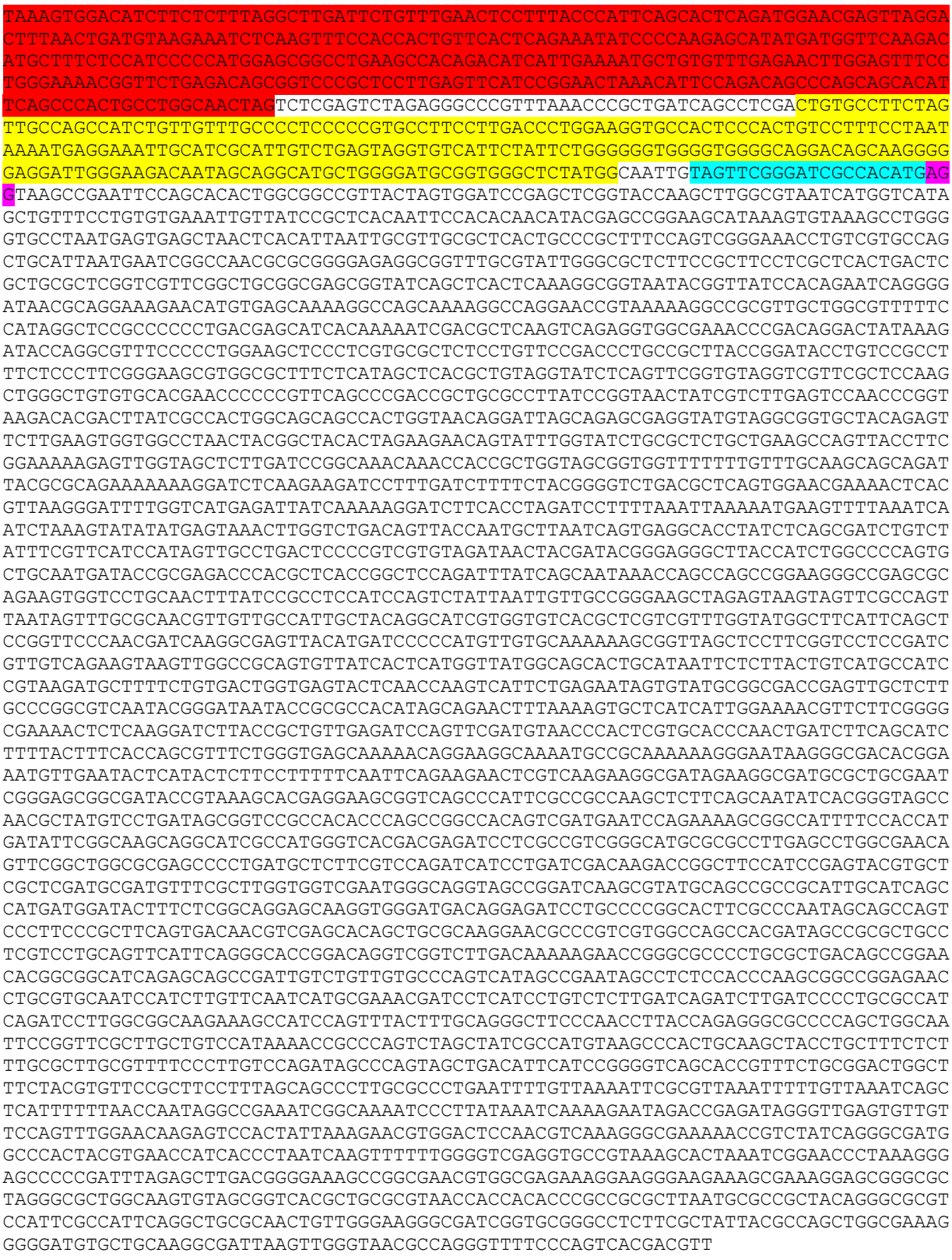
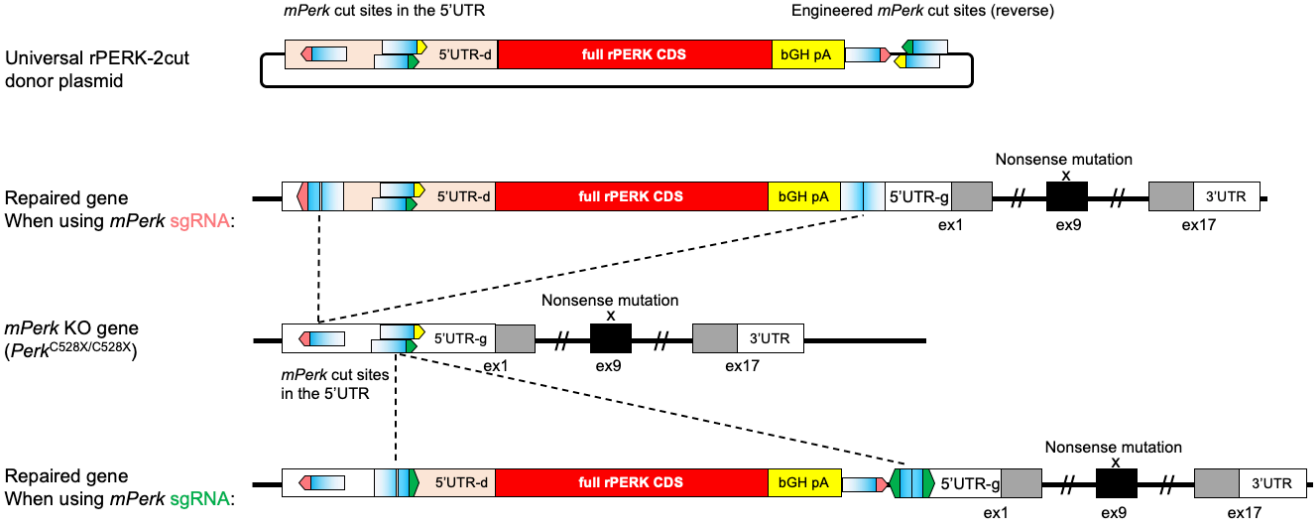
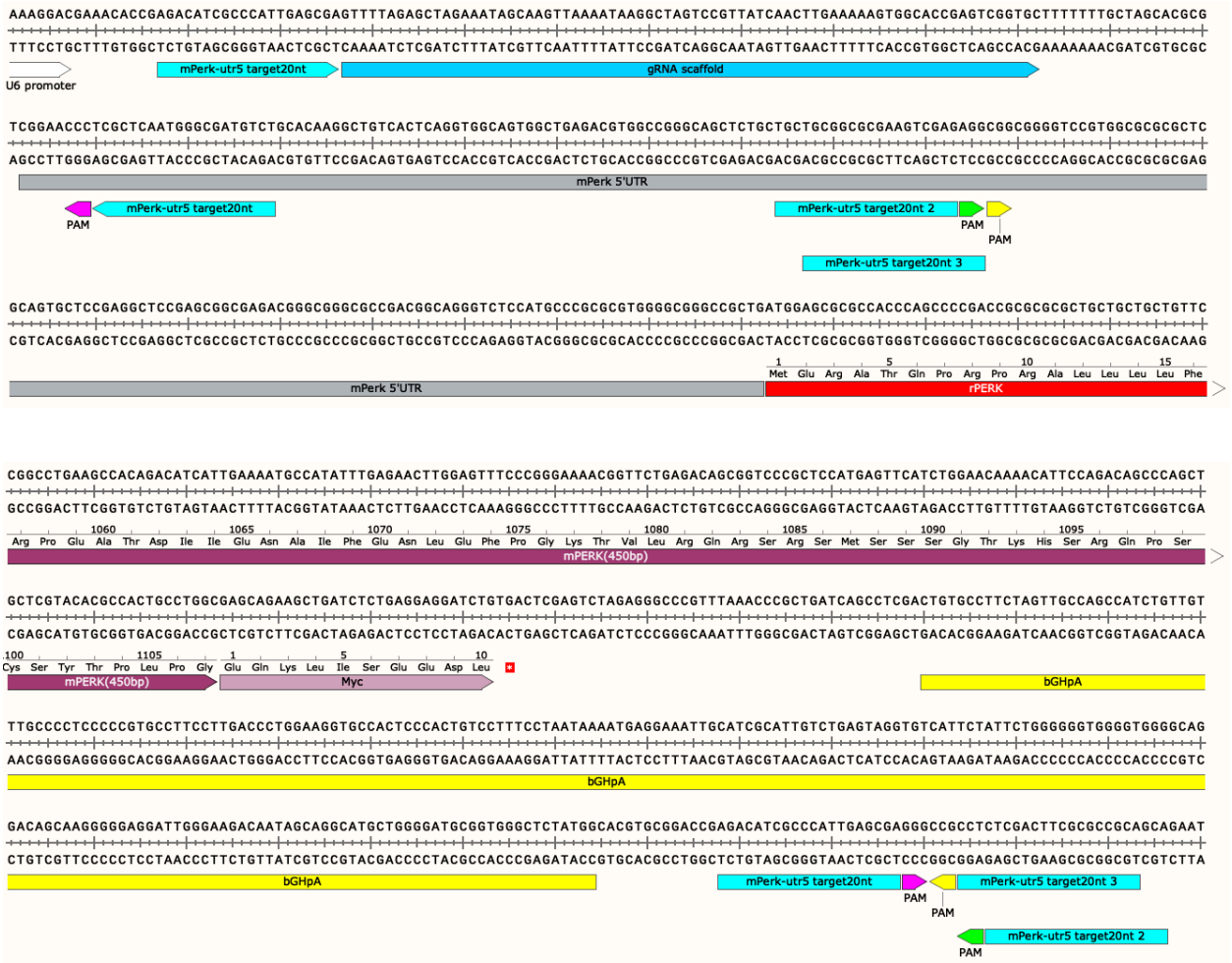



**Figure S7. Map and nucleotide sequence of the rPERKex7-17-2cut plasmid.**

Used for plasmid electroporation in *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ MEF cells (Figure 1).




>pBluescriptII-KS(+)-AAV-U6-mPERKutr5sg-mPERK5'UTR-rPERKmyc-bGHpA-2cut (7120 bp)

GATATCCCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTC
AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTGCGGCCGCACGCGTCAATTG**GAGG**
**GCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTA**
**AACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTT**
**TAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGAC**GAA
ACACCGAGACATCGCCCATTGAGCGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGA
AAAAGTGGCACCGAGTCGGTGCTTTTTTTGCTAGCACGCGTCGGAACCCTCGCTCAATGGGCGATGTCTGCACAAGGCTG
TCACTCAGGTGGCAGTGGCTGAGACGTGGCCGGGCAGCTCTGCTGCTGCGGCGCGAAGTCGAGAGGCGGCGGGGTCCGTG
GCGCGCGCTCGCAGTGCTCCGAGGCTCCGAGCGGCGAGACGGGCGGGCGCCGACGGCAGGGTCTCCATGCCCGCGCGTGG
GGCGGGCCGCTGATGGAGCGCGCCACCCAGCCCCGACCGCGCGCGCTGCTGCTGCTGTTCCTGCTGCTGGGCTGCGCGGC
GGGGATCTCAGCGGTCGCGCGCGCCCGCAGCTTGCTTGCTCCCACATCGGATACGGCATTTGGCTTGGGGGCAGCAGCCG
CCCCTACTTCAGCCGCGCGGGTGCCTGCGGTGGCTACGGCCGAAGTGACCGTGGAAGACGCCGAGGCATTGCCCGCTGCC
TCCGGCGAACAGGAGTCACGCGCGACCGAGTCCGATGACGACGTGGAACTGCGGCCTCGCGGCAGGTCCTTAGTAATCAT
CAGCACTTTAGATGGACGAATTGCCGCACTGGATGCCGAGAATCATGGGAAAAAGCAGTGGGATTTGGACGTGGGATCTG
GTTCCTTGGTTTCATCTAGCCTTAGCAAGCCAGAGGTGTTTGGGAACAAGATGATCATTCCTTCCCTGGACGGAGACCTC
TTCCAGTGGGACCGAGACCGAGAGAGCATGGAGGCCGTCCCTTTCACTGTGGAGTCCCTTCTCGAATCTTCATACAAGTT
TGGAGATGATGTTGTTCTGGTTGGAGGAAAGTCTCTGACCACATACGGGCTCAGTGCATATAGTGGAAAGCTGAGGTATA
TCTGTTCTGCCTTGGGATGTCGCCGATGGGATAGTGATGAAATGGAAGAAGAGGAAGACATCTTGCTTCTGCAACGCACG
CAGAAGACTGTGCGAGCTGTCGGGCCTCGAAGCGGCAGTGAGAAGTGGAATTTCAGTGTTGGCCACTTTGAACTTCGGTA
TATTCCAGACATGGAAACTAGAGCCGGATTCATTGAAAGCACCTTTAAACTGGGTGGAAACAAAGAAGACTCTAAAATCA
TTTCAGATGTGGAAGAACAAGATGTGGACACAGTGATAAAAGTTTCCGTTGCTGATTGGAAGGTCATGGCGTTTAGTAAG
AAGGGAGGCCGCCTAGAATGGGAGTACCAGTTTTGTACTCCCATTGCGTCTGCCTGGCTGGTGAGGGATGGTAAAGTCAT
CCCCATCAGTCTGTTTGATGACACAAGTTACACAGCCAATGAGGAAGTTTTGGAAGATGAAGAAGACATTGTAGAGGCTG
CTCGGGGAGCCACAGAGAACAGCGTGTACTTAGGGATGTACAGAGGCCAGCTGTACCTGCAATCATCCGTCAGGGTCTCG
GAAAAGTTCCCTACAAGGCCGAAGGCCTTGGAGTCTGTAAATGGAGAAAGTGCCATTATTCCTCTGCCGACGATCAAATG
GAAGCCCTTAATCCATTCTCCGTCTAGGACTCCTGTCTTGGTTGGGTCTGATGAATTTGACAAGTGTCTAAGTAATGATA
AATATTCCCATGAAGAATACAGTAATGGTGCGCTTTCAATCCTACAGTATCCATATGACAACGGTTACTATCTGCCTTAT

**Figure S8. Map and nucleotide sequence of the rPERKmyc-2cut plasmid with U6-mPERKutr5sg.**

Used for plasmid electroporation in *Perk*$^{C528X/C528X}$ MEF cell (Figure 2) and microinjection to generate rPERK-CRBR transgenic mice (Figure 3).

Schematic of universal CRBR 2-cut donor design, using *Perk*$^{C528X/C528X}$ genome editing as an example. Multiple Cas9/gRNA target sites can be tested for optimal 5' junction at the 5'UTR region, because context sequence around the 5' junction region could favor certain unexpected indels that may introduce unexpected new "ATG" in the 5'UTR region. The universal donor plasmid provides a rPERK-bGHpA cassette that is flanked by a wild-type 5'UTR of *mPerk* and three Cas9/gRNA target sites engineered together in reverse orientation as identified within the *mPerk* 5'UTR. It is a universal design to test out three different sgRNAs using one donor to compare the efficiency of CRBR editing at genome level and validate the transcription and translation from the repair cassette of the CRBR-edited allele.

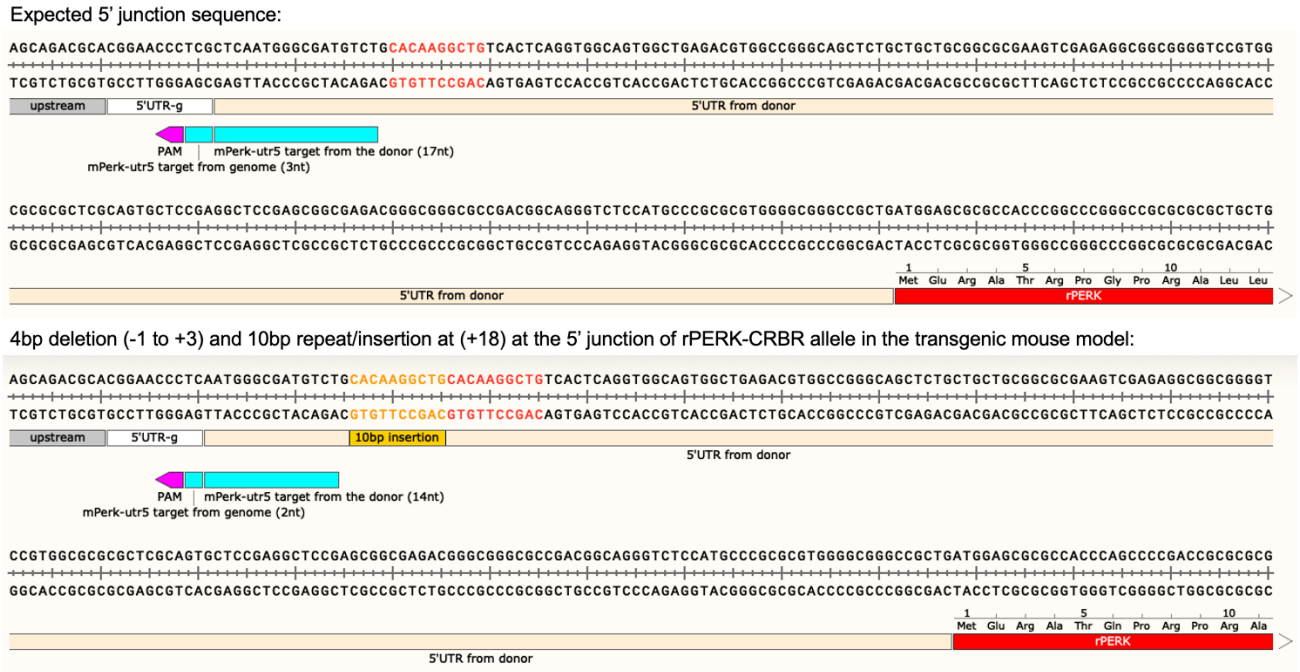
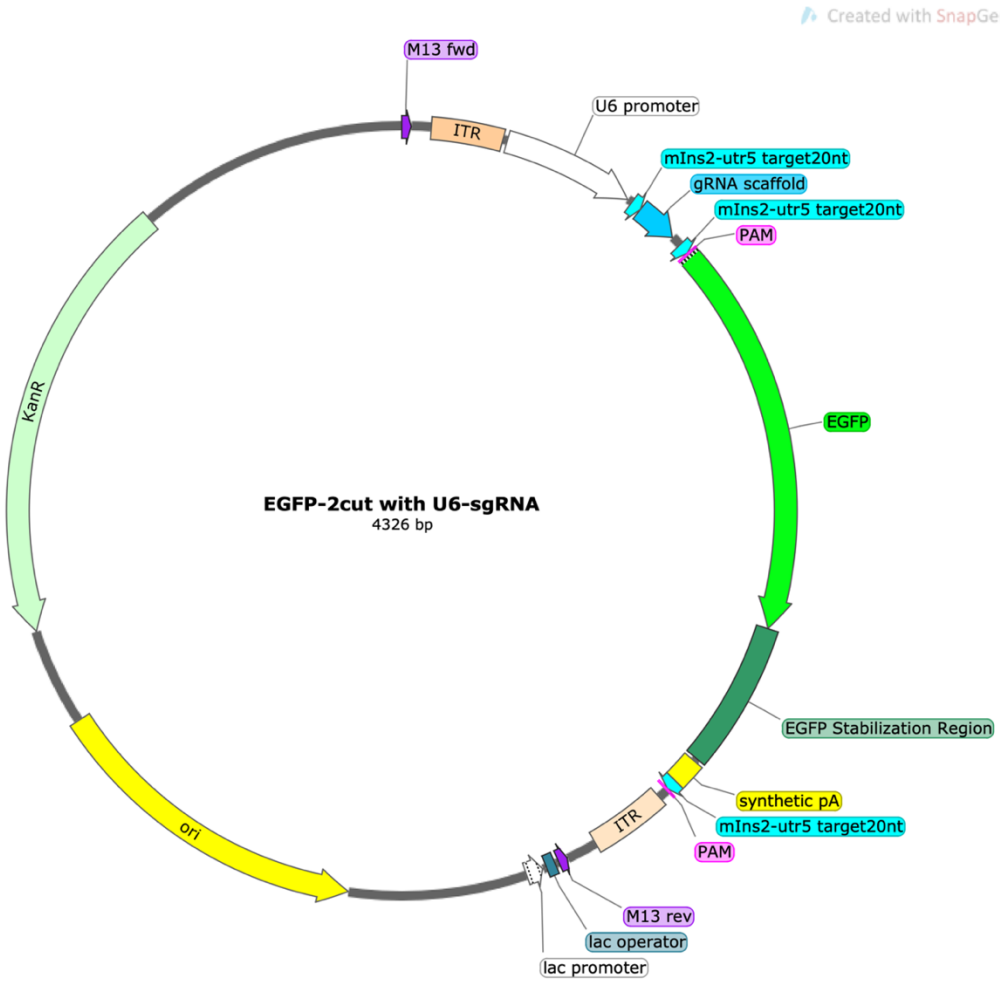



**Figure S9. The 5' junction sequence analysis of rPERK-CRBR allele.**

The 5' junction sequence of CRBR editing, expected vs rPERK-CRBR allele. Sequencing of 5' junction PCR showed 4bp deletion occurred at the joint site and a 10bp repeat sequence was inserted 18bp downstream of the joint site in the rPERK-CRBR allele from the transgenic mice.



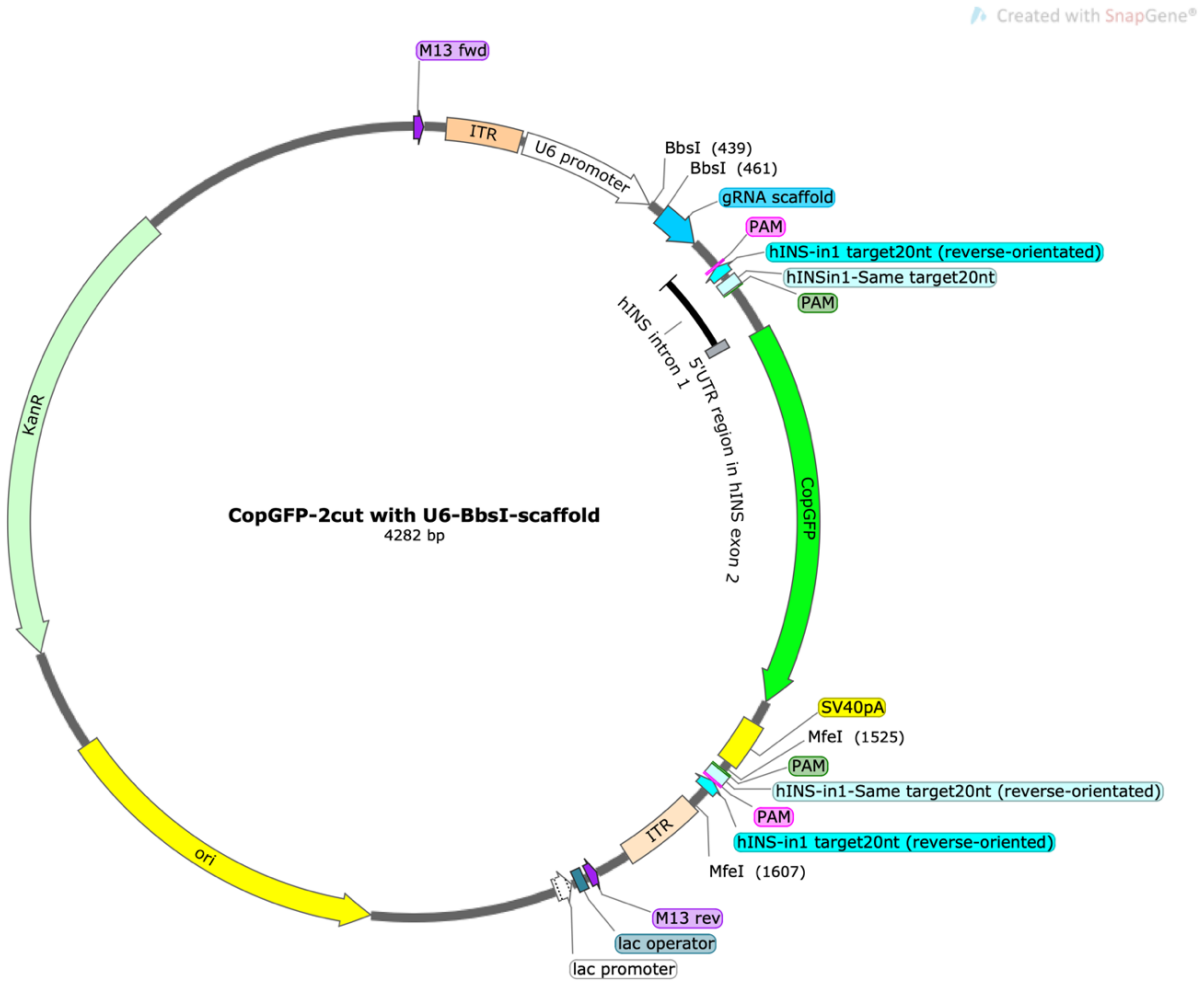

>pUC57-Kan-AAV-U6-mIns2utr5sg-EGFP-pA-2cut (4326 bp)

GTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCTCGCGAATGCATCTAGATCCTGCAGGCAGCTGCGCGCTCGCTCGC
TCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGC
CAACTCCATCACTAGGGGTTCCTGCGGCCGC**GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGG**
**CTGTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAA**
**TTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGA**
**TTTCTTGGCTTTATATATCTTGTGGAAAGGAC**GAAACACCGTGTAGCGGATCACTTAGGGCGTTTTAGAGCTAGAAATAG
CAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCCATGGGCTAGCTGTA
GCGGATCACTTAGGGCTGGATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACG
GCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTC
ATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCG
CTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCT
TCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAG
GGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCAT
GGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCG
ACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCC
CTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCAT
GGACGAGCTGTACAAGTCCGGACTCAGATCTCGAGAGGAGGAGGAGGAGACAGACAGCAGGATGCCCCACCTCGACAGCC
CCGGCAGCTCCCAGCCGAGACGCTCCTTCCTCTCAAGGGTGATCAGGGCAGCGCTACCGTTGCAGCTGCTTCTGCTGCTG
CTGCTGCTCCTGGCCTGCCTGCTACCTGCCTCTGAAGATGACTACAGCTGCACCCAGGCCAACAACTTTGCCCGATCCTT
CTACCCCATGCTGCGGTACACCAACGGGCCACCTCCCACCTACGAATTCAATAAAAGATCTTTATTTTCATTAGATCTGT
GTGTTGGTTTTTTGTGTGTGTAGCGGATCACTTAGGGCTGGATCGATGTTAACCAATTG**AGGAACCCCTAGTGATGGAGT**
**TGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGG**
**GCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG**ATCGGATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCA
AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGG
AAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCC
AGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCT
TCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAT

ACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGG
CGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC
GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTT
CGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAAC
TATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG
GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGT
TTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA
CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAAGCCCAATCTGAATAATGTTACAACCAATTAACCAATTCTGATTAGAAAAACTCATC
GAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAG
GAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATA
CAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAA
TGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAA
CCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGA
ATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTG
GAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAA
GAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTC
AGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCA
TTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATAACAC
CCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAG
AGATTTTGAGACACGGGCCAGAGCTGCATCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGA
GACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTC
GGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGC
GTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTC
TTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCAC
GACGTT

**Figure S10. Map and nucleotide sequence of the EGFP-2cut plasmid with U6-mINS2utr5sg.**

Used for plasmid electroporation in MIN6 cells and AAV intravenous injection (packaged in serotype 8 or DJ, AAV-U6-mINS2utr5sg-EGFP-2cut) in Cas9 or wild-type mice (Figure 4 and 5).




>pUC57-Kan-AAV-U6-BbsI-scaffold-hINSintron1-CopGFP-SV40pA-2cut (4282 bp)

GTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCTCGCGAATGCATCTAGATATCCCTGCAGGCAGCTGCGCGCTCGCT
CGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
GGCCAACTCCATCACTAGGGGTTCCTGCGGCCGCGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACA
AGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAA
TAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTT
CGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGGTCTTCGAGAAGACCTGTTTTAGAGCTAGAAATAG
CAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCCATGGGCTAGCGTGG
GCTCAGGATTCCAGGGTGGCTGGACCCCAGCCTCCCTGCTGCAGAGCTGGGGCACGTGGCTGGGCTCGTGAAGCATGTGG
GGGTGAGCCCAGGGGCCCCAAGGCAGGGCACCTGGCCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCCAGATCACTGTCCT
TCTGCCATGCCCGCCATGAAGATCGAGTGCCGCATCACCGGCACCCTGAACGGCGTGGAGTTCGAGCTGGTGGGCGGCGG
AGAGGGCACCCCCGAGCAGGGCCGCATGACCAACAAGATGAAGAGCACCAAAGGCGCCCTGACCTTCAGCCCCTACCTGC
TGAGCCACGTGATGGGCTACGGCTTCTACCACTTCGGCACCTACCCCAGCGGCTACGAGAACCCCTTCCTGCACGCCATC
AACAACGGCGGCTACACCAACACCCGCATCGAGAAGTACGAGGACGGCGGCGTGCTGCACGTGAGCTTCAGCTACCGCTA
CGAGGCCGGCCGCGTGATCGGCGACTTCAAGGTGGTGGGCACCGGCTTCCCCGAGGACAGCGTGATCTTCACCGACAAGA
TCATCCGCAGCAACGCCACCGTGGAGCACCTGCACCCCATGGGCGATAACGTGCTGGTGGGCAGCTTCGCCCGCACCTTC
AGCCTGCGCGACGGCGGCTACTACAGCTTCGTGGTGGACAGCCACATGCACTTCAAGAGCGCCATCCACCCCAGCATCCT
GCAGAACGGGGGCCCCATGTTCGCCTTCCGCCGCGTGGAGGAGCTGCACAGCAACACCGAGCTGGGCATCGTGGAGTACC
AGCACGCCTTCAAGACCCCCATCGCCTTCGCCAGATCTCGAGCTCGATGAGTTTGGACAAACCACAACTAGAATGCAGTG
AAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACA
ACAACAATTGCCCCCACATGCTTCACGAGCCCACCTCCCTGCTGCAGAGCTGGGGCCCACCCTGGAATCCTGAGCCCAC
ATCGATCAATTGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGGGATATCG
GATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG
TTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC
TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAA
CGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCT

GCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT
GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGAC
GAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGG
AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG
CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCC
CCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT
GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT
ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCT
TGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC
TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA
GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAGCCCAATCTGAATAATGTTA
CAACCAATTAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAA
TACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGG
TATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGA
GAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCC
AGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAAT
ACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAAT
ATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCAT
CATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCT
GTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGCGATAGAT
TGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCG
GCCTCGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCAT
GATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACGGGCCAGAGCTGCATCGCGCGTTTCGGTGA
TGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAG
CCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGA
GTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCT
GCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGC
GATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT

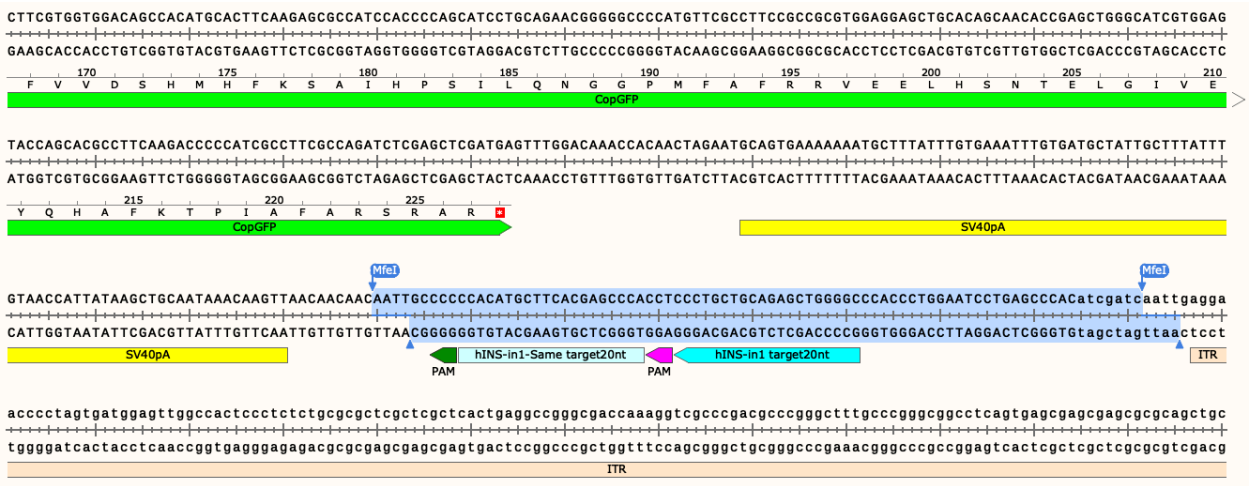
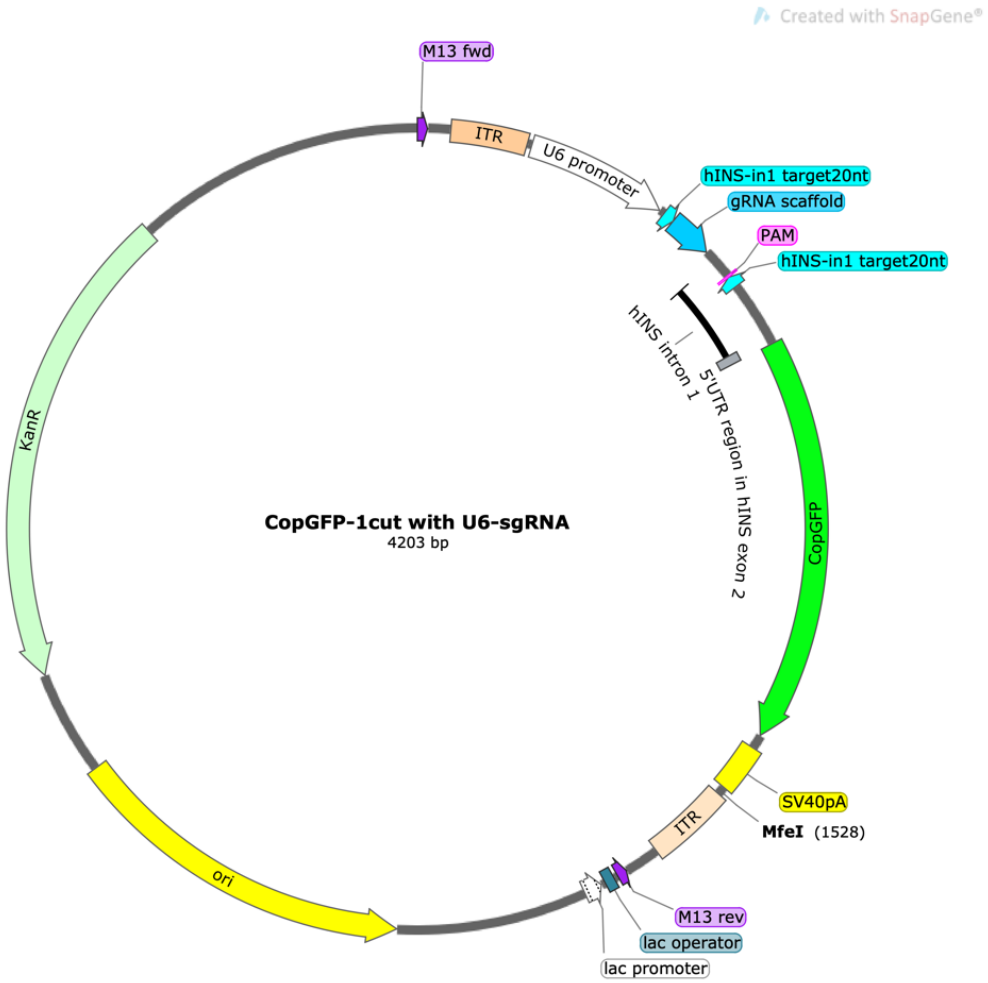



**Figure S11. Map and nucleotide sequence of the CopGFP-2cut plasmid.**

This universal donor was used to test sgRNA hINSin1sg-Reverse and hINSin1sg-Same in Figure S4. hINSin1sg-Reverse showed better CRBR editing in AD293 cells and human islets, so hINSin1sg-Reverse target(20nt) was chosen to clone into BbsI site for U6 driven expression in the same donor. MfeI double digestion on the plasmid removed the two targeted sites from the 3' of the CopGFP CRBR cassette and generated the CopGFP-1cut donor.



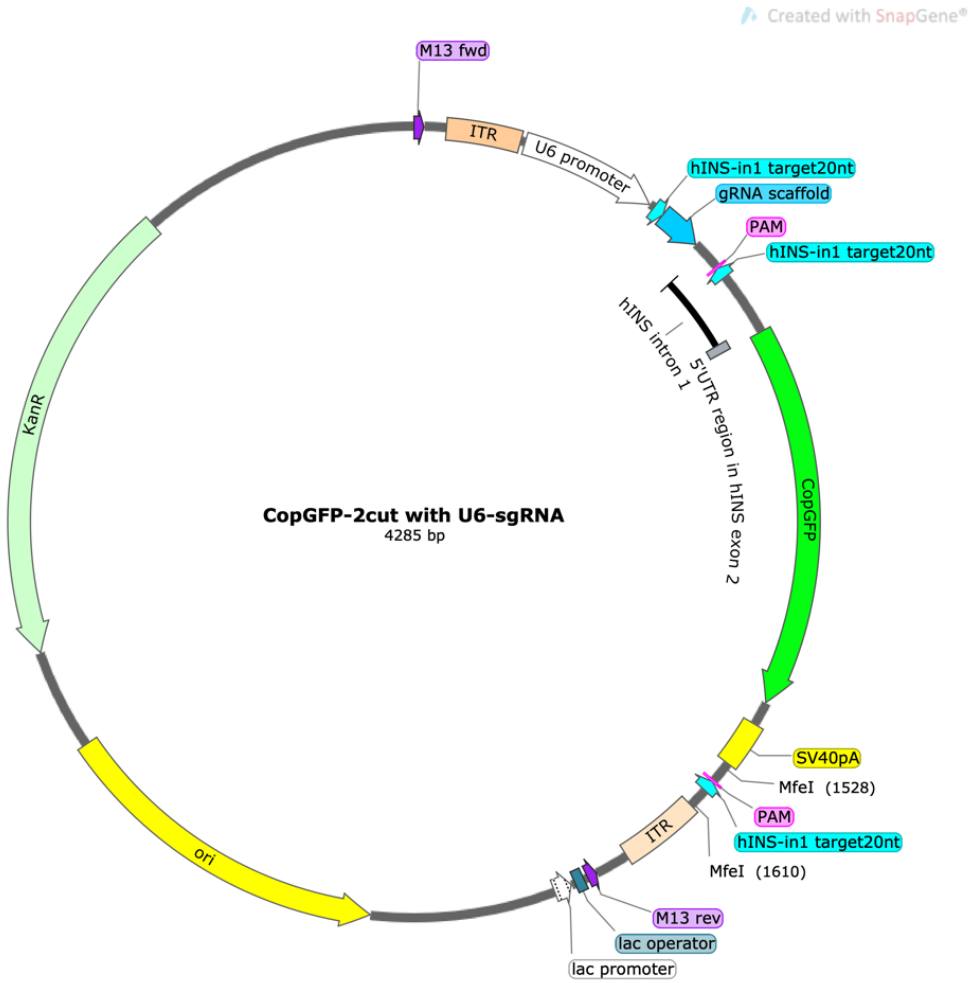

>pUC57-Kan-AAV-U6-hINSin1sg-hINSintron1-CopGFP-SV40pA-1cut (4203 bp)

GTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCTCGCGAATGCATCTAGATATC**CCTGCAGGCAGCTGCGCGCTCGCT**
**CGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT**
**GGCCAACTCCATCACTAGGGGTTCCT**GCGGCCGC**GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACA**
**AGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAA**
**TAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTT**
**CGATTTCTTGGCTTTATATATCTTGTGGAAAGGAC**GAAACACCG**GCCCCAGCTCTGCAGCAGGGGTTTTAGAGCTAGAAA**
**TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC**TTTTTTTCCATGGGCTAGCG
TGGGCTCAGGATTCCAGGGTGGCTGGACCCCAGCCTCCCTGCTGCAGAGCTGGGGCACGTGGCTGGGCTCGTGAAGCATG
TGGGGGTGAGCCCAGGGGCCCCAAGGCAGGGCACCTGGCCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCCAGATCACTGT
CCTTCTGCCATGCCCGCCATGAAGATCGAGTGCCGCATCACCGGCACCCTGAACGGCGTGGAGTTCGAGCTGGTGGGCGG
CGGAGAGGGCACCCCCGAGCAGGGCCGCATGACCAACAAGATGAAGAGCACCAAAGGCGCCCTGACCTTCAGCCCCTACC
TGCTGAGCCACGTGATGGGCTACGGCTTCTACCACTTCGGCACCTACCCCAGCGGCTACGAGAACCCCTTCCTGCACGCC
ATCAACAACGGCGGCTACACCAACACCCGCATCGAGAAGTACGAGGACGGCGGCGTGCTGCACGTGAGCTTCAGCTACCG
CTACGAGGCCGGCCGCGTGATCGGCGACTTCAAGGTGGTGGGCACCGGCTTCCCCGAGGACAGCGTGATCTTCACCGACA
AGATCATCCGCAGCAACGCCACCGTGGAGCACCTGCACCCCATGGGCGATAACGTGCTGGTGGGCAGCTTCGCCCGCACC
TTCAGCCTGCGCGACGGCGGCTACTACAGCTTCGTGGTGGACAGCCACATGCACTTCAAGAGCGCCATCCACCCCAGCAT
CCTGCAGAACGGGGGCCCCATGTTCGCCTTCCGCCGCGTGGAGGAGCTGCACAGCAACACCGAGCTGGGCATCGTGGAGT
ACCAGCACGCCTTCAAGACCCCCATCGCCTTCGCCAGATCTCGAGCTCGATGAGTTTGGACAAACCACAACTAGAATGCA
GTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTA
ACAACAACAATTG**AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCG**
**ACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG**GATATC
GGATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT
GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA
CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC
TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG

TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGA
CGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG
GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG
GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACC
CCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC
TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC
TACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC
TTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT
CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG
AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAGCCCAATCTGAATAATGTT
ACAACCAATTAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCA
ATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTG
GTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTG
AGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGC
CAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAA
TACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAA
TATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCA
TCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATC
TGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGCGATAGA
TTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGC
GGCCTCGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCA
TGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACGGGCCAGAGCTGCATCGCGCGTTTCGGTG
ATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAA
GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAG
AGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGC
TGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGG
CGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT

**Figure S12. Map and nucleotide sequence of the CopGFP-1cut plasmid with U6-hINSin1sg.**

Used for plasmid electroporation and AAV transduction (packaged in serotype DJ, AAV-DJ-U6-hINSin1sg-CopGFP-1cut) in human islet experiments (Figure 6 and 7).




>pUC57-Kan-AAV-U6-hINSin1sg-hINSintron1-CopGFP-SV40pA-2cut (4285 bp)

GTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCTCGCGAATGCATCTAGATATCCCTGCAGGCAGCTGCGCGCTCGCT
CGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
GGCCAACTCCATCACTAGGGGTTCCTGCGGCCGCGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACA
AGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAA
TAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTT
CGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGCCCCAGCTCTGCAGCAGGGGTTTTAGAGCTAGAAA
TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCCATGGGCTAGCG
TGGGCTCAGGATTCCAGGGTGGCTGGACCCCAGCCTCCCTGCTGCAGAGCTGGGGCACGTGGCTGGGCTCGTGAAGCATG
TGGGGGTGAGCCCAGGGGCCCCAAGGCAGGGCACCTGGCCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCCAGATCACTGT
CCTTCTGCCATGCCCGCCATGAAGATCGAGTGCCGCATCACCGGCACCCTGAACGGCGTGGAGTTCGAGCTGGTGGGCGG
CGGAGAGGGCACCCCCGAGCAGGGCCGCATGACCAACAAGATGAAGAGCACCAAAGGCGCCCTGACCTTCAGCCCCTACC
TGCTGAGCCACGTGATGGGCTACGGCTTCTACCACTTCGGCACCTACCCCAGCGGCTACGAGAACCCCTTCCTGCACGCC
ATCAACAACGGCGGCTACACCAACACCCGCATCGAGAAGTACGAGGACGGCGGCGTGCTGCACGTGAGCTTCAGCTACCG
CTACGAGGCCGGCCGCGTGATCGGCGACTTCAAGGTGGTGGGCACCGGCTTCCCCGAGGACAGCGTGATCTTCACCGACA
AGATCATCCGCAGCAACGCCACCGTGGAGCACCTGCACCCCATGGGCGATAACGTGCTGGTGGGCAGCTTCGCCCGCACC
TTCAGCCTGCGCGACGGCGGCTACTACAGCTTCGTGGTGGACAGCCACATGCACTTCAAGAGCGCCATCCACCCCAGCAT
CCTGCAGAACGGGGGCCCCATGTTCGCCTTCCGCCGCGTGGAGGAGCTGCACAGCAACACCGAGCTGGGCATCGTGGAGT
ACCAGCACGCCTTCAAGACCCCCATCGCCTTCGCCAGATCTCGAGCTCGATGAGTTTGGACAAACCACAACTAGAATGCA
GTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTA
ACAACAACAATTGCCCCCCACATGCTTCACGAGCCCACCTCCCTGCTGCAGAGCTGGGGCCCACCCTGGAATCCTGAGCC
CACATCGATCAATTGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGG
CGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGGGATA
TCGGATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAA
TTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCT
AACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGC
CAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCG

GCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACA
TGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCT
GACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCC
TGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCG
TGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAA
CCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC
ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA
ACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGC
TCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGG
ATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCA
TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAGCCCAATCTGAATAATG
TTACAACCAATTAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTAT
CAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCC
TGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAG
TGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAG
GCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGA
AATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAAC
AATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATG
CATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCA
TCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAGCGATA
GATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATC
GCGGCCTCGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTT
CATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACGGGCCAGAGCTGCATCGCGCGTTTCGG
TGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGAC
AAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTG
AGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAG
GCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAA
GGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT

**Figure S13. Map and nucleotide sequence of the CopGFP-2cut plasmid with U6-hINSin1sg.** Used for plasmid electroporation and AAV transduction (packaged in serotype DJ, AAV-DJ-U6-hINSin1sg-CopGFP-2cut) in human islet experiments (Figure 6 and 7).

**Table S1. RRID of human cadaveric islets used for experiments**.

| Figure | RRID | Endocrine Tissue (%) | Beta Cells (%) |
|---|---|---|---|
| Fig. S4B-E | SAMN13254972 | 92 | 45 |
| Fig. S5A | SAMN13515839 | 95 | 43 |
| Fig. 6C-F, S5B | SAMN14255441 | 97 | 68 |
| Fig. 7D | SAMN15314807 | 71 | 59 |
| Fig. 7B-C | SAMN15518672 | 96 | 67 |

**Table S2. Oligonucleotides for BbsI site cloning into sgRNA expression constructs.**

| Target site | Oligonucleotide sequences (5'→3') | |
|---|---|---|
| mPerk-ex9 | Top oligo | 5'-CACCgCCTGCGCACGATGAAGGTCG-3' |
| | Bottom oligo | 5'-AAACCGACCTTCATCGTGCGCAGGc-3' |

| | | |
|---|---|---|
| mPerk-in6 | Top oligo<br>Bottom oligo | 5'-CACCgTAGTTCGGGATCGCCACATG-3'<br>5'-AAACCATGTGGCGATCCCGAACTAc-3' |
| mPerk-utr5 | Top oligo<br>Bottom oligo | 5'-CACCgAGACATCGCCCATTGAGCGA-3'<br>5'-AAACTCGCTCAATGGGCGATGTCTc-3' |
| mIns2-utr5 | Top oligo<br>Bottom oligo | 5'-CACCgTGTAGCGGATCACTTAGGGC-3'<br>5'-AAACGCCCTAAGTGATCCGCTACAc-3' |
| hINS-in1<br>(hINS-in1 Reverse) | Top oligo<br>Bottom oligo | 5'-CACCgGCCCCAGCTCTGCAGCAGGG-3'<br>5'-AAACCCCTGCTGCAGAGCTGGGGCc-3' |
| hINS-in1-Same | Top oligo<br>Bottom oligo | 5'-CACCgTGGGCTCGTGAAGCATGTGG-3'<br>5'-AAACCCACATGCTTCACGAGCCCAc-3' |

**Table S3. PCR primers for junction PCR on genomic DNA.**

| Cassette, target, junction | Primer location | Oligonucleotide sequences (5'→3') |
|---|---|---|
| rPERKex7-17-2cut, mPerk-in6, 5' junction | mP-intron6-F<br>m/rP-exon7-R | CACAAGTTAGAAAGGCTCA<br>TGCAGGTACAGCTGGCCTCT |
| rPERKex7-17-2cut, mPerk-in6, 3' junction | rP-exon15-F<br>mP-intron9-R | CATTCAGCACTCAGATGGAACG<br>GAAGGCAGATAGGAAAGTC |
| rPERKex7-17-2cut, mPerk-in6, flanking | mP-intron6-F<br>mP-intron9-R | CACAAGTTAGAAAGGCTCA<br>GAAGGCAGATAGGAAAGTC |
| rPERKex7-17-2cut, mPerk-in6, flipped 5' junction | mP-intron6-F<br>rP-exon15-F | CACAAGTTAGAAAGGCTCA<br>CATTCAGCACTCAGATGGAACG |
| rPERKmyc-2cut, mPerk-utr5, 5' junction | upstream-mP-utr5-F<br>rP-exon1-R | CTTCCTAGCCACAAATTGGGC<br>CCCAAGCCAAATGCCGTATC |
| rPERKmyc-2cut, mPerk-utr5, 3' junction | Myc-F<br>mP-intron1-R | AGCAGAAGCTGATCTCTGAGG<br>GGTGCCCCGACCCTAGTTC |
| EGFP-2cut, mIns2-utr5, 5' junction | upstream-mIns2-utr5-F<br>EGFP-R | AGCCAAGGACAAAGAAAGCA<br>TTCAGGGTCAGCTTGCCGTAG |
| EGFP-2cut, mIns2-utr5, 3' junction | EGFP-F<br>mIns2-exon2-R | CACTACCTGAGCACCCAGTC<br>GTGGGACTCCCAGAGGAAG |
| EGFP-2cut, mIns2-utr5, flanking | upstream-mIns2-utr5-F<br>mIns2-exon2-R | AGCCAAGGACAAAGAAAGCA<br>GTGGGACTCCCAGAGGAAG |
| EGFP-2cut, mIns2-utr5, whole plasmid integration | U6-F<br>mIns2-in1ex2-R | GAGGGCCTATTTCCCATGATT<br>CTGGAAGATAGGCTGGGTTG |
| CopGFP-1cut(or 2cut), hINS-in1, 5' junction | upstream-hINS-utr5-F2<br>CopGFP-R | GTGCTGACGACCAAGGAGAT<br>GAAGATCACGCTGTCCTCGG |
| CopGFP-1cut(or 2cut), hINS-in1, 2cut 3' junction | CopGFP-F<br>hINS-exon2-R | CCGAGGACAGCGTGATCTTC<br>CCCCGCACACTAGGTAGAGA |
| CopGFP-1cut(or 2cut), hINS-in1, 1cut 3' junction | U6-F<br>hINS-exon2-R | GAGGGCCTATTTCCCATGATT<br>CCCCGCACACTAGGTAGAGA |
| CopGFP-1cut(or 2cut), hINS-in1, 5' junction (for AAV constructs) | upstream-hINS-utr5-F1<br>CopGFP-qR2 | GAGGAAGAGGTGCTGACGAC<br>TGATGCGGCACTCGATC |

**Table S4. Quantitative realtime PCR primers for mRNA expression.**

| Gene | Primer location | Oligonucleotide sequences (5'→3') |
|---|---|---|
| mActin (Actb) | exon4-qF<br>exon5-qR | GCCCTGAGGCTCTTTTCC<br>TGCCACAGGATTCCATACCC |
| rPERKex10 | rP-exon10-qF<br>rP-exon10-qR | GTGAGTGCTGATAACAGCG<br>CGTATCCTGAGTGTTTAATGTCG |
| rPERKex12-13 | rP-exon12-qF<br>rP-exon13-qR | TGGATGAAATCTGGCTGAAGG<br>GGGAGCTGAGTGGCCAGTC |
| EGFP | EGFP-qF<br>EGFP-qR | GCAAAGACCCCAACGAGAAG<br>TCACGAACTCCAGCAGGACC |
| nascent mIns2 | intron2-qF<br>intron2-qR | CCTAGGTGTGGAGGGTCTCG<br>CCAGAAACGTGTCCCCACTC |
| mature mIns2 | exon2-qF<br>exon3-qR | GGGGAGCGTGGCTTCTTCTA<br>AGTGCCAAGGTCTGAAGGTC |
| mINS2utr5-F – EGFP-R1,<br>mINS2utr5-F – EGFP-R2 | exon1(utr5)-qF<br>EGFP-qR1<br>EGFP-qR2 | GGGACCCAGTAACCACCAG<br>CGGTGAACAGCTCCTCG<br>GTTTACGTCGCCGTCCAG |
| hActin (ACTA1) | exon5-qF<br>exon5-qR | CGCGACATCAAGGAGAAGCT<br>CTCGTTCTCGAAGTCCAGGG |
| hINSutr5-F – CopGFP-R1,<br>hINSutr5-F – CopGFP-R2 | exon1(utr5)-qF<br>CopGFP-qR1<br>CopGFP-qR2 | GCATCAGAAGAGGCCATCAA<br>CGGCACTCGATCTTCATG<br>TGATGCGGCACTCGATC |
| hINS | exon2-qF<br>exon2-qR | CAGCCTTTGTGAACCAACAC<br>GGTCTTGGGTGTGTAGAAGAAG |
| mPERKex1-2 | mP-exon1-qF<br>mP-exon2-qR | GTCACGCGCGACGGAGC<br>CTCGGCATCCAGTGCAGCG |
| rPERKex1 | rP-exon1-qF<br>rP-exon1-qR | CCACATCGGATACGGCA<br>CCCGCGCGGCTGAAGTA |