Molecular Therapy
Original Article

Co-opting regulation bypass repair 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]

[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

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 results in simultaneous co-option of the endogenous regulatory region and bypass of the genetic defect. We validated the CRBR strategy for human gene therapy by rescuing a mouse model of Wolcott-Rallison syndrome (WRS) with permanent neonatal diabetes caused by either a large deletion or a nonsense mutation in the PERK (EIF2AK3) gene. Additionally, we integrated a CRBR GFP-terminator cassette downstream of the human insulin promoter in cadaver pancreatic islets of Langerhans, which resulted in insulin promoter regulated expression of GFP, demonstrating the potential utility of CRBR in human tissue gene repair.

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 Cas9 endonuclease with a 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), and homology-directed repair (HDR). Although the HDR pathway has been shown to repair genes precisely in mouse models of human disease,[8–14] this pathway 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, Suzuki et al.[24] have shown that NHEJ repair pathway can ligate heterologous DNA to the two cut ends generated by sgRNA/Cas9 double-strand cleavage. This mechanism, denoted as homologous-independent targeted insertion (HITI), can be used to insert large DNA fragments. We have used the HITI method to develop CRBR as a novel gene therapy strategy, whereby an entire gene coding sequence (CDS) and transcription/translation terminator cassette is inserted downstream of a gene's promoter but upstream of a deleterious disease-causing mutation. Expression of the CRBR cassette, which contains the normal coding sequence of the gene being repaired, can

Received 30 October 2020; accepted 15 April 2021;
https://doi.org/10.1016/j.ymthe.2021.04.017.

Correspondence: Douglas R. Cavener, Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA.
E-mail: drc9@psu.edu

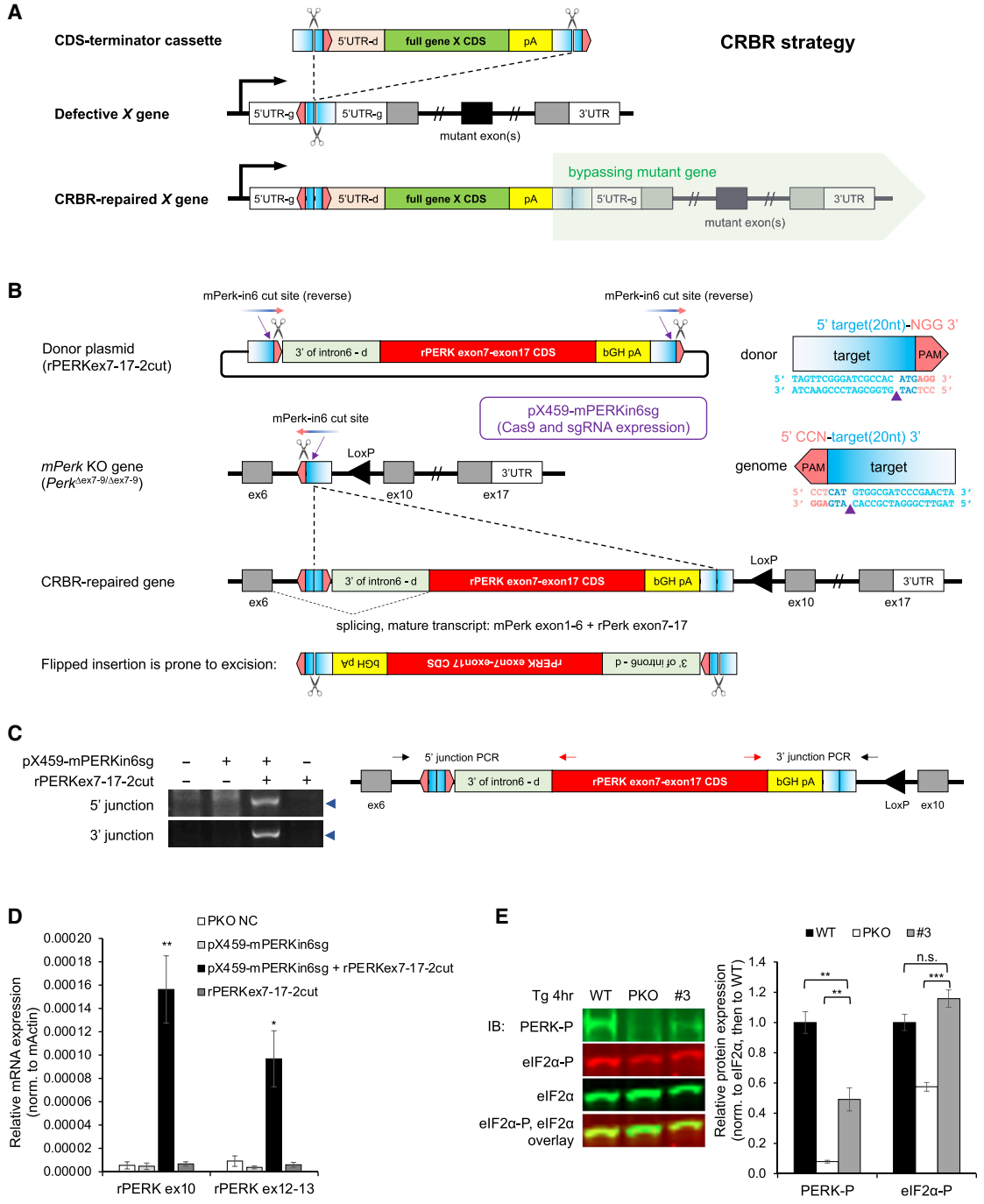


Figure 1. CRBR-mediated *in vitro* partial PERK CDS integration in *Perk* KO cell line

(A) Schematic of CRBR strategy. The CDS-terminator cassette is flanked by Cas9/gRNA target sites in reverse orientation of the genome. Correct integration of the CRBR cassette is expressed under the native promoter, with the 5′ UTR having small changes resultant from residue target site from the donor. Salmon pentagon: PAM site (3 nt). Rectangle with blue gradient: Cas9/gRNA targeted protospacer sequence (20 nt); Cas9 cleavage locates at 17 nt to the white side, 3 nt to the blue side. 5′UTR-g: 5′ UTR in the genome. 5′UTR-d: 5′ UTR engineered in the donor. (B) Schematic of CRBR-Partial-CDS strategy for $Perk^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ genome. The donor plasmid provides a 3′intron6-rPERKex7-17CDS-bGHpA cassette that is flanked by Cas9/gRNA target sites in reverse orientation (5′ 20 nt-NGG 3′) as identified within the *mPerk* intron 6 (5′ CCN-20 nt 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 Cas9 excision. Small changes at the 5′

(legend continued on next page)

rescue its deficiency by restoring normal expression of the wild-type CDS under its native promoter and other regulatory elements while bypassing the downstream mutated region. Because a single CRBR CDS-terminator cassette contains all of the wild-type 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* knockout (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 pancreas, 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. The CRBR gene repair may be used in the future as the basis for a strategy 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. The *Perk* gene is highly conserved in rodents, and the rat *Perk* gene has previously been shown to be fully functional in mice;[26] therefore, using the rat *Perk* CDS was advantageous for distinguishing between endogenous mouse *Perk* and the CRBR integrated rat *Perk*. 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 them. Consequently, incorrectly oriented integrants could be re-excised by Cas9 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′ and 3′ junction diagnostic 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, 33 cell lines were positive for the 5′ junction diagnostic PCR (Figure S1A). In order to test for functional PERK restoration in the CRBR-edited *Perk* KO MEF cells, we chose and subjected eight cell lines to thapsigargin treatment, which induces endoplasmic reticulum (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 seven 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 expressed functional PERK, had an 11 bp 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

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 (3×10^6 cells) were electroporated with 1.8 μg of pX459-mPERKin6sg, 1.6 μg of rPERKex7-17-2cut donor, or both in 100 μL using MEF 2 Nucleofector Kit. Puromycin (1 μg/mL) was used to enrich transfected cells (with pX459-mPERKin6sg treatment) for 3 days. Genomic DNA (C) was harvested 6 days post-transfection for 5′ and 3′ junction diagnostic PCRs. Primers were designed to flank the junction sites (triangle mark: 5′, 254 bp; 3′, 890 bp). Chimeric mouse-rat *Perk* mRNA expression levels (D) 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* first and then to PKO MEF cells. Quantification represents n = 3 per treatment. Data are represented as mean ± SE. 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 the CRBR-edited cell line #3 ($Perk^{CRBR\text{-}rPERKex7\text{-}17/backbone\ integration}$) treated with 1 μM thapsigargin (Tg) for 4 h. Relative protein expression was normalized to eIF2α first and then to WT MEF cells. Quantification represents n = 4 per cell line. Data are represented as mean ± SE. Statistical significance was calculated relative to the *Perk* WT or $Perk^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ MEF cells; **$p < 0.01$, ***$p < 0.001$; n.s., not significant.

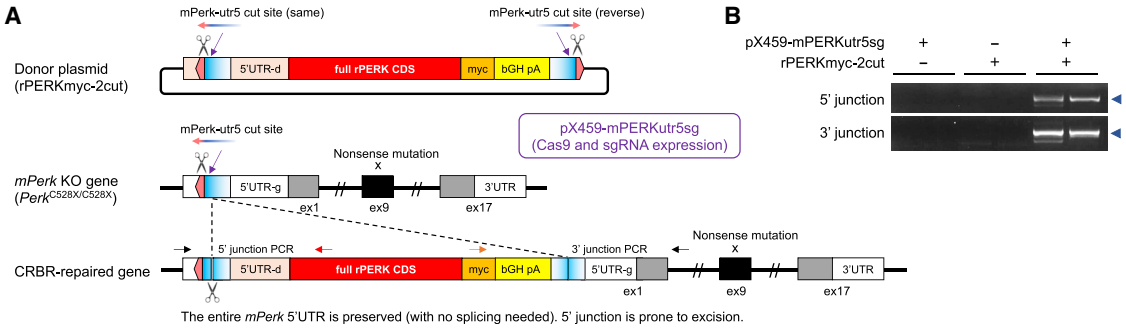


Figure 2. CRBR-mediated *in vitro* full PERK CDS integration in *Perk* KO cell line

(A) Schematic of CRBR-Full-CDS strategy. The 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 preserves the wild-type sequence of *mPerk* 5′ UTR but also resumes the mPerk-utr5 cut site, making it 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 (1×10^5 cells) were electroporated with 1 μg of pX459-mPERKutr5sg, 1 μg of rPERKmyc-2cut donor, or both using the 10 μL Neon transfection system in two replicates. Genomic DNA was harvested 2 days post-transfection for 5′ and 3′ junction diagnostic PCRs. Primers were designed to flank the junction sites (triangle mark: 5′, 921 bp; 3′, 857 bp). The lower-molecular-weight bands seen in one replicate reflect that part of the CRBR-edited alleles had large NHEJ deletions at the junction.

1–2 bp indels) but retained the splice acceptor (Figure S1C). The resulting alternative mature transcript in these non-expressing cell lines contained an extra 135 bp intronic sequence that encoded a stop codon, which likely resulted in nonsense-mediated mRNA decay (NMD). 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 needs to be avoided.

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 (~3.4 kb) with a myc tag, a bGHpA terminator, and a Cas9/sgRNA target site engineered in reverse orientation (Figure 2A). This modified CRBR strategy preserves the sequence of the 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′ and 3′ junction diagnostic PCRs (Figure 2B), which confirmed the CRBR-Full-CDS integration at the intended target site in the genome *in vitro*.

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 utero* 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 into 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 diagnostic PCRs. Further genotyping of F1 offspring from this founder mouse crossed to a wild-type mouse revealed the founder 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 (Figure 3A). 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 genotyped to be KO/rPERK-CRBR heterozygotes ($Perk^{C528X/rPERK\text{-}CRBR}$ or $Perk^{\Delta ex7\text{-}9/rPERK\text{-}CRBR}$), healthy and fertile. *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 4 weeks of age.[26–29] The rPERK-CRBR allele showed complete phenotypic rescue of both the *Perk* nonsense null mutant (Figures 3B, 3C, S2A, and S2D) and the *Perk* Δex7-9 deletion mutant (Figures S2B–S2D) with respect to survivorship, growth, beta cell mass, exocrine pancreas viability, and glucose homeostasis.

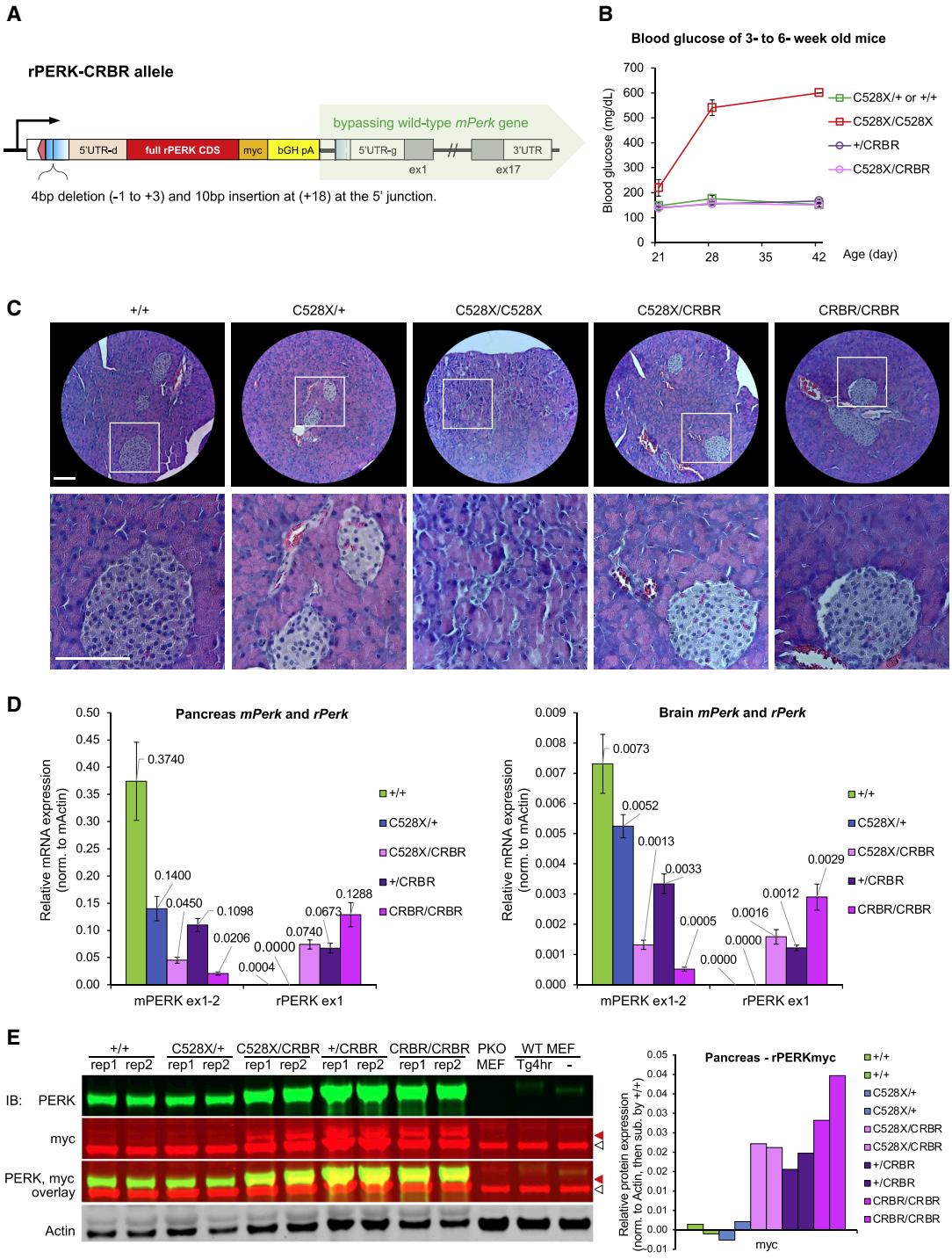


Figure 3. CRBR-edited *Perk* allele rescues *Perk* KO allele in a proof-of-concept mouse model

(A) Schematic of rPERK-CRBR allele (in a wild-type mouse *Perk* background) from the transgenic mouse. See Figure S11 for sequence of the 5′ junction at 5′ UTR. (B) Blood glucose levels were monitored at P21, P28, and P42 of mice with genotypes indicated in the chart. Normal blood glucose levels were observed in $Perk^{C528X/rPERR\text{-}CRBR}$ mice at all ages (see Figure S2A for data of 5-month-old mice). Data are represented as mean ± SE. Student's t test showed no significant difference in 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 exceeded the glucometer upper limit (600 mg/dL) by P35. C528X/CRBR or +/CRBR mice were offspring from $Perk^{+/rPERR\text{-}CRBR}$ crossed to $Perk^{C528X/+}$ mice. *Perk* KO (C528X/C528X) or littermates (C528X/+ or +/+) were offspring from

(legend continued on next page)

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 NMD. The insertion of the CRBR cassette into the wild-type mouse allele resulted in a ~95% reduction in mouse *Perk* mRNA. Therefore, we estimate that ~5% of the primary transcripts in the CRBR alleles are transcriptional read-through of the rPERK-myc-bGHpA terminator within the CRBR cassette resulting in a low level of the downstream mouse *Perk* mRNA transcript. This small fraction of transcripts generated by failure to terminate at the bGH polyA terminator are bicistronic, composed of rPERK-myc followed by mPERK. It is very unlikely that the mPERK sequences within this hybrid CDS would be translated, because normal cap-dependent translation initiates only at the first CDS, which, in this case, is the rPERK-myc CDS. 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. Consequently, a low level of transcriptional read-through in a CRBR engineered gene-correction scheme should not interfere with the CRBR strategy to bypass translation of the downstream endogenous coding sequence.

Consistent with their phenotypic rescue, the C528X/CRBR and Δex7-9/CRBR mice expressed a substantial level of *rPerk* mRNA derived from the CRBR cassette. Low-level detection of *mPerk* mRNA in these mice was contributed by the KO mutant allele and by the CRBR allele (leaky transcriptional read-through), neither of which is 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 the rPERK-CRBR cassette. Cassette-derived rPERK protein expression was confirmed by immunoblotting with a myc antibody as well as an antibody that recognizes 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 effects *in vivo*.

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

To more directly 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. 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 (~1.1 kb) remains stably integrated. This design feature, however, did alter the 5′ UTR from the wild-type sequence with small changes resulting from the residue target site in the donor. 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 donor plasmid. EGFP-positive cells were visible by 2 days post-transfection and continued to increase in number through 15 days, whereas donor-only treated cells remained EGFP-negative over the same time period (Figure 4B). 5′ and 3′ junction analyses 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 S3A); the low percentage of positive cells reflects the relatively poor transfection efficiency of MIN6 cells (~25%).

A subset of GFP-positive cells was clonally isolated for further characterization (Figure S3B). Junction PCRs and DNA sequence analyses showed that cell lines #8, #10, #13, and #14 had one CRBR-edited

$Perk^{C528X/+}$ mice intercross. (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 had typical *Perk* KO defects, such as very small islets with reduced beta cell mass. The disorganized acinus structure contained some degranulated cells (white), clear halos around the nuclei, and gaps between acinar cells, which were 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, 100 μm. (D) The mRNA expression levels of endogenous *mPerk* and *rPerk* from CRBR-edited allele in pancreas and brain of adult mice (1- to 5-month) were quantified using *mPerk*- and *rPerk*-specific primers and were normalized to *mActin*. $Perk^{+/+}$, n = 6; $Perk^{C528X/+}$, n = 6; $Perk^{C528X/rPERK\text{-}CRBR}$, n = 9; $Perk^{+/rPERK\text{-}CRBR}$, n = 7; $Perk^{rPERK\text{-}CRBR/rPERK\text{-}CRBR}$, n = 8. $Perk^{+/+}$ and $Perk^{C528X/+}$ mice had no detectable *rPerk* signal (Ct value > 36, used 40 for calculation if undetermined) in pancreas and brain. Data are represented as mean ± SE. (E) Two replicate mice with the same genotype were sacrificed at P38 ($Perk^{+/+}$, from $Perk^{+/rPERK\text{-}CRBR}$ intercross), 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 $Perk^{+/rPERK\text{-}CRBR}$). Both mPERK and rPERK protein expression in pancreas were detected by immunoblotting using an anti-PERK antibody. The rPERK-myc protein was also recognized by a myc tag antibody. Solid red triangle marks the true myc signal, while the hollow triangle marks a nonspecific band recognized by the myc tag antibody. Negative control was $Perk^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ (PKO) MEF cells. Positive control was $Perk^{+/+}$ (WT) MEF cells treated with or without 1 μM thapsigargin (Tg) for 4 h. Relative rPERK-myc protein expression was normalized to actin first and then obtained by background subtraction of the average signal of the two $Perk^{+/+}$ replicates.

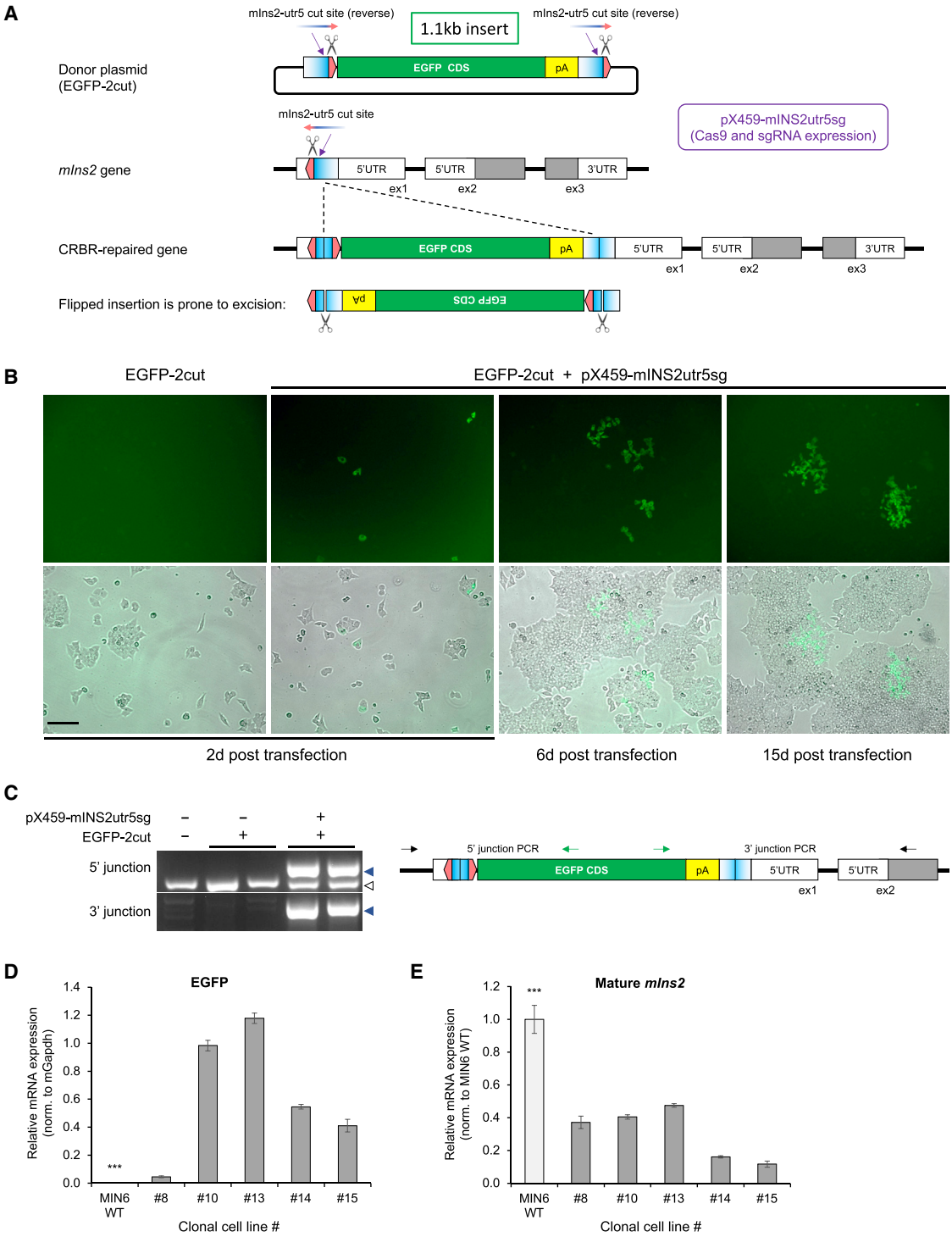


Figure 4. CRBR-mediated *in vitro* EGFP CDS integration in mouse *Ins2* gene

(A) Schematic of CRBR-EGFP-2cut strategy for wild-type *mIns2* genome. The donor plasmid provides an EGFP CDS-pA cassette that is flanked by Cas9/gRNA target sites in reverse orientation (5′ 20 nt-NGG 3′) as identified within the *mIns2* 5′ UTR in exon 1 (5′ CCN-20 nt 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 as well as a targeted DSB at genomic *mIns2* 5′ UTR. Correct integrants will retain the CRBR cassette, while incorrect integrants are prone to excision. (B and C) MIN6

(legend continued on next page)

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 S3C). 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 that the CRBR-integrated EGFP-CDS-pA cassette is expressed and can bypass the endogenous mouse *Ins2* transcription.

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 substantially reduces the variability when compared to 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 does supply Cas9 in *trans* (AAV vectors; Figure 5A). Liver and pancreas tissues from Cas9-EGFP mice were isolated 30 days post retro-orbital (r.o.) injection of the AAV8-sgRNA-CDS vector. Junction PCRs and digital droplet PCR (ddPCR) quantitation revealed substantial CRBR-mediated gene editing at the genome level in the liver (4.16% of chromosome 7 edited with CRBR integration of EGFP CDS) and a detectable level (0.64%) 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, therefore, and, as expected, EGFP transcription from the *Ins2* gene locus in the liver was not observed.

Previous experiments[32–35] 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 (8.39%) being more efficient (Figure 5D). The tail-vein-injected AAV8-sgRNA-CDS was capable of targeting the pancreas (0.84%), with some individuals having detectable CRBR editing at the genome level by junction PCRs. However, similarly administered AAV-DJ-sgRNA-CDS showed less pancreatic CRBR editing (0.29%). These results show that systemic 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 separate 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) by dual AAV administrations (0.56%), however, not with the same efficiency as was seen when Cas9 was endogenously expressed (Figure 5D). Leaky expression of the promoterless EGFP CDS from AAV vector (ITR) was not observed in the liver of mice with AAV-DJ-sgRNA-CDS (Figure S4), although it is known that the ITR of AAV has weak promoter activity.[36,37] Overall, these results suggest that CRBR-mediated gene editing is feasible *in vivo* via dual AAV delivery once both viral vectors are successfully transduced in the host cell. Most importantly, the CRBR cassette expression is restricted to pancreatic beta cells under the insulin promoter.

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 encoding the 5′ UTR and upstream of the insulin start codon (Figure 6A). The CRBR cassette contains sequences homologous to the 3′ half of the endogenous intron 1 as well as a region homologous to the 5′ UTR encoded by exon 2, which contains an acceptor splice site that is needed for proper splice excision of the newly integrated intron 1. By this design, any unforeseen indels generated during CRBR integration are spliced out of the resulting mature mRNA. In addition to the 2-cut donor, we introduced a 1-cut donor to determine which strategy was more efficacious (Figure 6B). A 1-cut strategy generates only one insert linearized from the 1-cut donor, with one correct integrant out of two possible outcomes (50%); whereas the 2-cut strategy generates four possible inserts that may be integrated in two orientations, with two correct integrants out of eight possible outcomes (25%) (Figure S5). For the 1-cut strategy, a much larger fragment (4.2 kb) must be integrated. By contrast, the 2-cut strategy integrates a much smaller fragment (0.9 kb, CRBR cassette only), as it excludes the extraneous vector sequences. However, these extraneous vector sequences should not interfere with gene expression

cells (1×10^6 cells) were electroporated with 1 μg of EGFP-2cut donor with or without 1 μg of pX459-mINS2utr5sg in 100 μL using Nucleofector V Kit in two replicates. Cells were imaged (B) as live cultures 2 days, 6 days, and 15 days post-transfection at 10× objective; scale bar, 100 μm. Genomic DNA (C) was harvested 6 days post-transfection for 5′ and 3′ junction diagnostic PCRs. Primers were designed to flank the junction sites (solid triangle: 5′, 452 bp; 3′, 690 bp). The hollow triangle marks a nonspecific band recognized by 5′ junction PCR primers. (D and E) EGFP mRNA expression levels from the CRBR-edited allele (D) were quantified in five sorted GFP-positive MIN6 cells (#8, 10, 13, 14, and 15) by normalizing to *mGapdh*, while the wild-type (WT) MIN6 control cell line had no detectable EGFP signal (Ct value > 36, used 40 for calculation if undetermined). Mouse *Ins2* mRNA expression levels (E) were quantified by normalizing to *mGapdh* first, and then the relative fold change in expression was calculated relative to MIN6 WT cells. Quantification represents n = 4 per sorted cell line. Data are represented as mean ± SE. All five GFP-positive cell lines were significantly different from the wild-type MIN6 control for both EGFP and *mIns2* expression levels; ***$p < 0.001$.

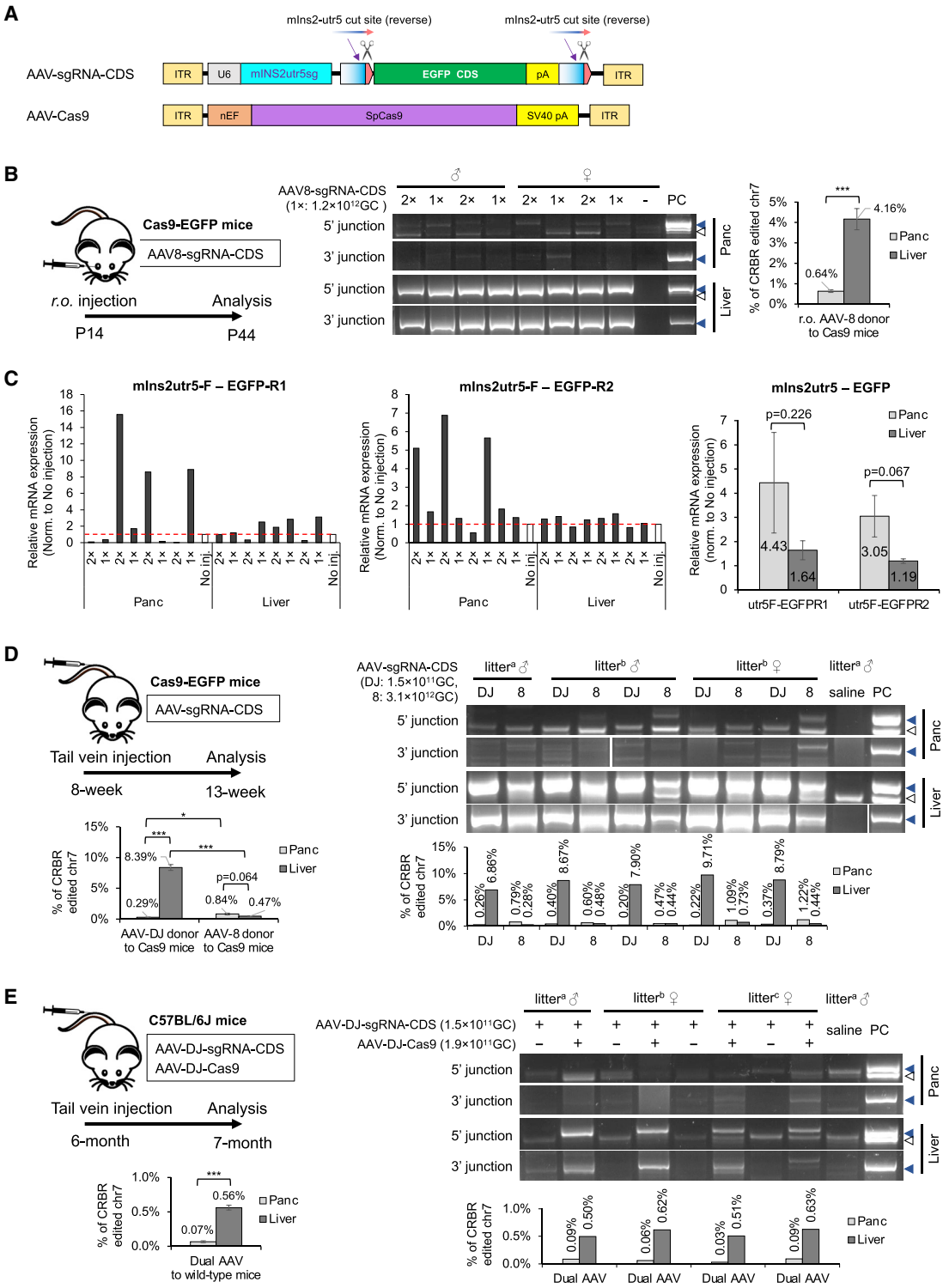


Figure 5. CRBR-mediated *in vivo* EGFP CDS integration in mouse *Ins2* gene

(A) Schematic of CRBR AAV vectors used in AAV delivery to Cas9-EGFP mice or wild-type mice. The AAV vector provides the same EGFP CRBR cassette as in the EGFP-2cut donor plasmid but also includes a U6-driven mIns2utr5-sgRNA. Cas9 is expressed in all tissues under the universal promoter CAG in the Cas9-EGFP mice. (B and C)

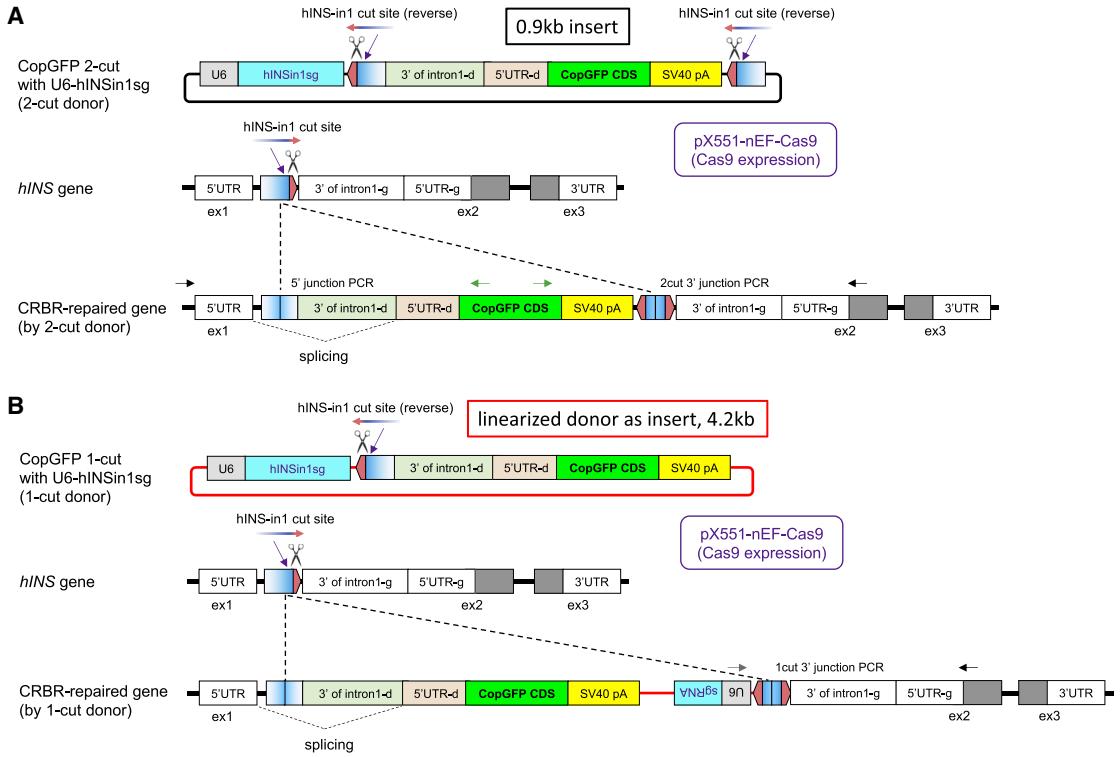
(legend continued on next page)

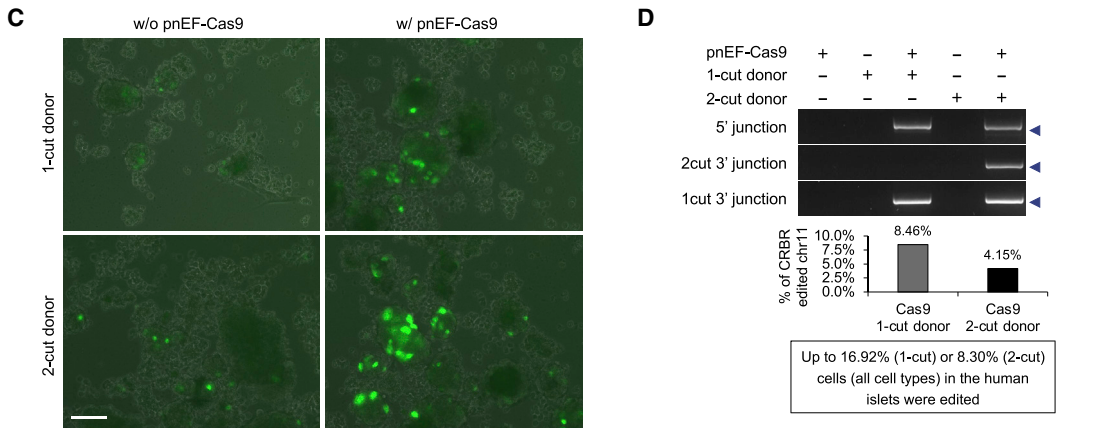
because they are downstream of the transcription/translation terminators in the CRBR cassette.

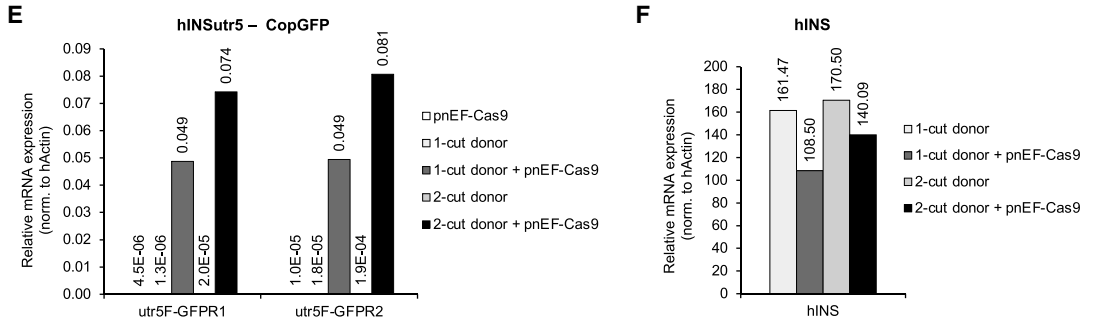
This CRBR-CopGFP strategy was first tested in an easily transfected human cell line, AD293, to identify the optimal sgRNA target site within intron 1 and to optimize the donor design before testing in human islets (Figure S6A). We found that the reverse-oriented sgRNA (12.75%) outperformed the same-oriented sgRNA (4.56%) in CRBR integration (Figure S6B). Six off-targets of the reverse-oriented sgRNA were then tested for possible off-target integration of the CopGFP CDS. Of these, three showed detectable off-target integrations (0.78%–1.60%) (Figures S6C and S6D; Table S1). Both the CopGFP 1-cut and 2-cut donor plasmids were engineered with a U6-hINSin1sg cassette, which expresses the optimized reverse-oriented sgRNA. The SpCas9-expressing plasmid and the 1-cut or 2-cut donor plasmid were co-transfected into human islets. Six days post-transfection, many CopGFP-positive islet cells were observed (Figure 6C). This result indicates successful targeting to the pancreatic beta cells, as they are the only islet cell type with an active insulin promoter and comprise 45%–70% of the total cadaver islet cell population (Table S2). The remaining islet cells secrete other metabolically important peptide hormones.[38] While these non-beta cell types should likely be edited with equal frequency compared to beta cells, their insulin promoter is inactive, and therefore we would not expect those cells to express the CopGFP CRBR cassette. Junction PCRs confirmed CRBR editing of the human *INS* locus at the genome level (Figure 6D), with 8.46% (1-cut) or 4.15% (2-cut) of chromosome 11 edited, 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 (Figures S6E–S6H). Collectively, these results demonstrate that CRBR-mediated 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.

Previous reports of AAV transduction of human islets have shown limited success.[39,40] However, our success in using AAV to edit the insulin gene in the mouse pancreas (Figures 5B–5E) 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 S7A). To test the ability of CRBR-mediated gene editing in human islets *ex vivo* via AAV-DJ transduction, human islets were co-infected with AAV-DJ-sgRNA-CDS-1cut (or 2cut) (Figure 7A) along with AAV-DJ-Cas9. CopGFP-positive cells were observed at 6 days post-infection (Figure 7B). By 10 and 16 days post-infection, these cells dramatically increased in both number and fluorescence intensity (Figures 7B and S7B). This indicates that living and functional human beta cells can at least maintain insulin expression for 16 days. When CRBR integration was analyzed at the genome level, the expected 5′ junction diagnostic PCR was observed with 3.21% (1-cut) or 0.75% (2-cut) of chromosome 11 edited; 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. Single-cell sorting of CRBR-treated human islets showed that 1.97% (1-cut strategy) or 0.96% (2-cut strategy) of the islet cells had undergone CopGFP integration and expression (Figures 7D and S7C). By analyzing beta-cell-specific transcription factors (PDX1 and GLUT2) and alpha-cell-specific (glucagon) and delta-cell-specific (somatostatin) markers, we confirmed that the GFP-positive cells were largely, if not exclusively, beta cells (Figures 7E and 7F). The transcription of CopGFP from the human *INS* locus in an independent batch of human islets was measured 18 days post-infection

Two-week-old Cas9-EGFP mice from one litter (four males and five females) were injected with two doses or one dose (2×: 40 μL, or 1×: 20 μL) of AAV8-U6-mINS2utr5sg-EGFP-2cut via r.o. injection, with un-injected mice serving as a control. DNA and RNA from pancreas and liver were isolated 30 days post-injection. Genomic DNA (B) was tested by 5′ and 3′ junction diagnostic PCRs and by ddPCR quantification of the CRBR integration of EGFP CDS into chromosome 7 (chr7). The percentage of CRBR editing was calculated by normalizing the 5′ junction event to an internal control (*mRpp30* on chr19, two copies per pancreatic cell, four copies per hepatocyte). EGFP mRNA expression (C) from the CRBR-edited *mIns2* gene was 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 then calculated relative to the no-injection control. Quantification represents n = 8 (mice with two different dosages of injection showed no dosage effect and therefore were pooled together for liver and pancreas comparison). Data are represented as mean ± SE. (D) Eight-week-old Cas9-EGFP mice from two litters (litter^{a} or litter^{b}, gender is indicated in the figure) were injected with 50 μL 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 isolated 35 days post-injection. CRBR editing at the genome level was tested by 5′ and 3′ junction diagnostic PCRs and by ddPCR quantification of the CRBR integration as in (B), n = 5. (E) Six-month-old C57BL/6J mice from three litters (litter^{a}, litter^{b}, or litter^{c}, gender is indicated in the figure) were injected with 50 μL of AAV-U6-mINS2utr5sg-EGFP-2cut with or without 50 μL of AAV-nEF-Cas9 in serotype DJ, or saline via tail vein injection. Genomic DNA from pancreas and liver was isolated 35 days post-injection. CRBR editing at the genome level was tested by 5′ and 3′ junction diagnostic PCRs and by ddPCR quantification of the CRBR integration as in (B), n = 4. For (B), (D), and (E), all primers were designed to flank the junction sites, the same as Figure 4 for the MIN6 cell line (solid triangle: 5′, 452 bp; 3′, 690 bp). The hollow triangle marks a nonspecific band recognized by 5′ junction PCR primers. PC, positive control, was genomic DNA from MIN6 cells co-transfected with EGFP-2cut donor and pX459-mINS2utr5sg. Statistically significant differences in CRBR editing efficiency at the genome level were seen between pancreas and liver, or between AAV serotypes; *$p < 0.05$, ***$p < 0.001$. Total genome copies (GC) used is indicated in the figure. Titer of AAV used: AAV8-U6-mINS2utr5sg-EGFP-2cut, 6.15×10^{13} GC/mL; AAV-DJ-U6-mINS2utr5sg-EGFP-2cut, 2.92×10^{12} GC/mL; AAV-DJ-nEF-Cas9, 3.83×10^{12} GC/mL.







(legend on next page)

(Figure 7G), with that of the 1-cut strategy slightly exceeding that of the 2-cut strategy. The consistent better performance of the 1-cut strategy when looking at CRBR integration efficiency at the genome level, the fraction of GFP-positive cells, and GFP mRNA expression suggest that using the 1-cut donor is more efficient than the 2-cut donor via AAV transduction. The second cut downstream of the cassette donor is not necessary, because AAV vector does not have a large backbone as in plasmid vector. 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.

DISCUSSION

Delivering CRISPR-based therapeutics 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. Suzuki et al.[46] had recently developed an intercellular linearized single homology arm donor mediated intron-targeting integration (SATI), which has great applications for targeting of a broad range of mutations and cell types by utilizing both NHEJ and HDR pathways. However, SATI strategy also requires a specific design for each mutation variant. 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 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. A more direct intra-organ injection route may improve the delivery to the pancreas or other tissues that are challenging to target by intravenous injection. Mutations in *Perk*, which result in severe and permanent neonatal diabetes in WRS patients, 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.[47] 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 stem 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.[48] 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 need for immunosuppressive drugs. Overall, CRBR gene repair combined with autologous cell replacement therapy (GR-ACR) 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 therefore 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 S8) 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

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. The 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-20 nt 3′) as identified within the *hINS* intron 1 (5′ 20 nt-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 4.2 kb, much larger than the 2-cut insert, which is only 0.9 kb. 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 1 μg of pnEF-Cas9, 1 μg of pU6-hINSin1sg-CopGFP-1cut, 1 μg 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, 100 μm. Genomic DNA (D) was harvested for diagnostic PCRs of the 5′, 2cut 3′, and 1cut 3′ junctions. Primers were designed to flank the junction sites (triangle: 5′, 820 bp; 2cut 3′, 722 bp; 1cut 3′, 654 bp). The percentage of CRBR editing (ddPCR quantification of the CRBR integration of CopGFP CDS into chr11) was calculated by normalizing the 5′ junction event to an internal control (*hRPP30* on chr10, two copies per cell). CopGFP mRNA expression levels (E) from the CRBR-edited *hINS* gene (using a forward primer targeting *hINS* 5′ UTR and a reverse primer [R1 or R2] targeting CopGFP) and *hINS* mRNA expression levels (F) were quantified by normalizing to *hActin*.

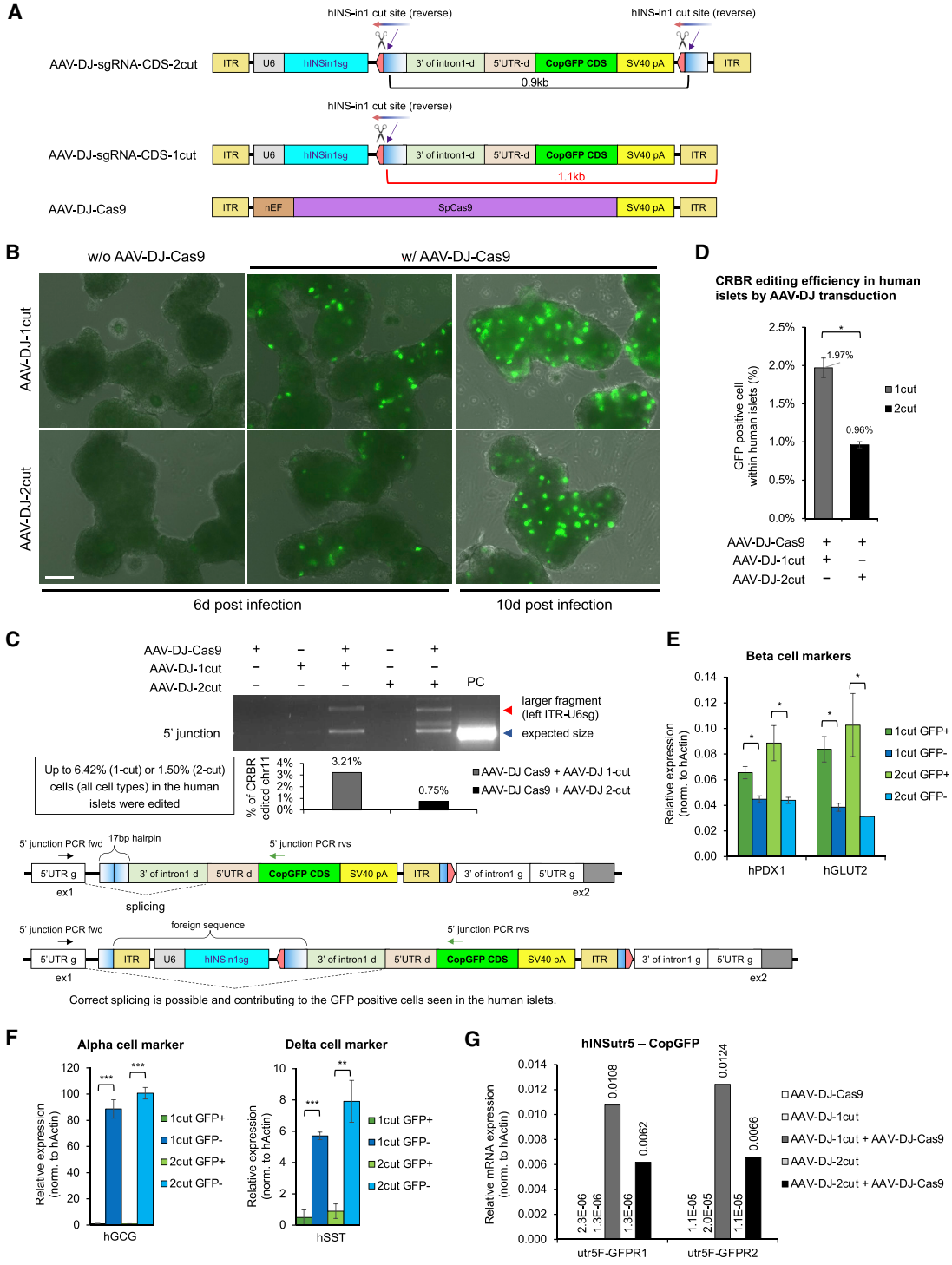


Figure 7. CRBR-mediated *ex vivo* CopGFP CDS integration in human *INS* gene via AAV-DJ transduction

(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 infected 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 60,000 MOI. Human islets were imaged (B) 6 days and 10 days post-infection as live cultures at 10× objective; scale bar, 100 μm. Genomic DNA (C) was harvested 16 days post-infection for 5′ junction PCR. Primers were designed to flank the 5′ junction site and amplify a 476 bp fragment (blue triangle). The solid red

(legend continued on next page)

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 GR-ACR strategy is used, a specific cell line can be clonally isolated that is devoid of interfering mutations.

Although other delivery methods[49,50] 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 4 kb. 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 4 kb. For genes that exceed 4 kb, 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. An additional limitation of using rAAV vectors for CRISPR-based gene editing is the persistent expression of Cas9, which may result in mutagenic and immunological complications.[51] To mitigate this problem, Cas9 mRNA or protein could be delivered by a non-viral vector along with the CRBR cassette and sgRNA delivered by an AAV vector. Alternatively, a self-deleting Cas9 could be employed to limit the expression of Cas9.[52]

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, thorough testing of repair efficacy in cell culture and/or model organisms is essential. A distinct advantage of CRBR gene-correction strategy 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)

A transgenic mouse model with a nonsense mutation in exon 9 of mouse *Perk* gene (c.1584C>A; p.Cys528X) was generated by CRISPR-Cas9-mediated genome editing via HDR in mouse zygote with a 200 nt single-stranded oligodeoxynucleotide (ssODN) template containing one nonsense mutation and four synonymous mutations. SpCas9 mRNA (5meC, Ψ) was purchased from TriLink (San Diego, CA, USA). *In vitro* transcription and purification of mPERKex9-sgRNA (see Construction of plasmids for sgRNA sequence) were as previously described.[53] Repair template (200 nt ssODN, 4 nmol Ultramer DNA Oligo) was purchase from Integrated DNA Technologies (IDT, Coralville, IA, USA) (5′-cagcccccactaca gcaagaacatccgcaagaaggaccctatcctcctgctgcactggtggaaggagatattcgggacg atcctgctt**tgA**atcgtGgccacAacGttTatcgtgcgcaggcttttccatcctcagcccca cagggtaagatgctctgtcaacctaatgtgcttccaagtggttgctgtgtaggaaacct-3′). A nonsense mutation was introduced by a C to A mutation on the ssODN template, 14 bp from the Cas9/sgRNA cleavage. Three synonymous mutations were designed 2 bp, 5 bp, and 8 bp from the 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*)

A transgenic mouse model with the rPERK-CRBR allele (rPERKmyc integration at 5′ UTR of *mPerk*) was generated by CRBR-mediated gene editing in mouse zygote. A rPERK CDS with a myc tag at the C terminus was designed to integrate into the mouse *Perk* 5′ UTR using the CRBR strategy as described in the Results (Figure 2A). SpCas9 protein was purchased from IDT. A synthetic mPERKutr5-sgRNA (see Construction of plasmids for sgRNA sequence) was purchased from Synthego (Redwood City, CA, USA). The rPERKmyc-2cut donor plasmid was constructed as described in Construction of plasmids. The SpCas9 protein, sgRNA, and the rPERKmyc-2cut donor

triangle marks a larger fragment that is only present in Cas9 + sgRNA CDS donor treatments. Sequencing of this additional fragment revealed it to encode the left ITR and U6-sgRNA regions of the AAV vector. PC, positive control, was genomic DNA from AD293 cells co-transfected with CopGFP-2cut donor and pX459-hINSin1sg. The percentage of CRBR editing (ddPCR quantification of the CRBR integration of CopGFP CDS into chr11) was calculated by normalizing the 5′ junction event to an internal control (*hRPP30* on chr10, two copies per cell). Resultant genome diagrams show two possible AAV-1cut integrations: expected 5′ junction generates a nascent mRNA with a 17 bp 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 can also be spliced out. (D–F) A second batch of human cadaveric islets (800 IEQs per replicate) was infected with AAV-DJ-U6-hINSin1sg-CopGFP-1cut or AAV-DJ-U6-hINSin1sg-CopGFP-2cut in combination with AAV-DJ-nEF-Cas9 at 60,000 MOI. Single-cell sorting of 1cut- or 2cut-treated human islets was performed at 11 days post-infection. The percentage of GFP-positive cells (D) among total cells sorted (alpha [~25%], beta [~60%], delta [~8%], and other cell types within islet cell cluster) were calculated. RNA was harvested from GFP-positive and GFP-negative sorted cells. mRNA expression of marker genes for pancreatic endocrine cells (E and F) was quantified by normalizing to *hActin*. Quantification represents n = 3 per treatment. Data are represented as mean ± SE. Statistical significances were shown as marked: *p < 0.05, **p < 0.01, ***p < 0.001. (G) A third batch of human cadaveric islets was treated the same as (B) and (C), and RNA was harvested 18 days post-infection. CopGFP mRNA expression levels from the CRBR-edited *hINS* gene were quantified by normalizing to *hActin*.

plasmid were sent to the 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. The generation of the *Perk* KO allele (Δex7-9) has been 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, USA). Mice were sacrificed by CO_2 asphyxiation. 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, USA; plasmid #62988, deposited by Feng Zhang) as previously described.[54] See Table S3 for oligonucleotides for BbsI site cloning. The Cas9/sgRNA genomic target sequences (20 nt + PAM) on sense (+) or antisense strand (−) used in this study include:

mPerk-ex9, 5′-CCTGCGCACGATGAAGGTCGTGG-3′ (−);

mPerk-in6, 5′-TAGTTCGGGATCGCCACATGAGG-3′ (−);

mPerk-utr5, 5′-AGACATCGCCCATTGAGCGAGGG-3′ (−);

mIns2-utr5, 5′-TGTAGCGGATCACTTAGGGCTGG-3′ (−);

hINS-in1 (or hINS-in1-Reverse in Figure S6), 5′-GCCCCAGC TCTGCAGCAGGGAGG-3′ (+);

hINS-in1-Same (Figure S6), 5′-TGGGCTCGTGAAGCATGTGG GGG-3′ (+).

Each of these target sequences was determined by Surveyor Assay (IDT) or T7 Endonuclease I (T7E1) Assay (New England Biolabs, Ipswich, MA, USA) from 2∼3 candidates with top on-target scores identified from https://zlab.bio/guide-design-resources or https://benchling.com:443/crispr/. To construct rPERKex7-17-2cut (Figure S9), rPERKex7-17CDS-bGHpA was 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, USA), 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, USA). The rPERKmyc-2cut (Figure S10) was then generated by cloning mPERK(450 bp)-myc from pcDNA-mPERK-9E10 (in house) into rPERK-2cut through SapI and XhoI to replace rPERK(450 bp). The 150 amino acid (aa) C terminus is conserved between mPERK and rPERK. The EGFP-2cut (Figure S12) 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 (49 bp) 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 S13) 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 S13). The CopGFP-1cut (or 2cut) with U6-hINSin1sg (Figures S14 and S15) was constructed by cloning the hINSin1sg-Reverse into the BbsI site and was then used either in the 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 the plasmid experiment or AAV-nEF-Cas9 packaging in serotype DJ.

Cell culture

MEF cells were immortalized from *Perk*$^{\Delta ex7\text{-}9/\Delta ex7\text{-}9}$ embryos[55] and *Perk*$^{C528X/C528X}$ embryos 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's medium (DMEM) (Gibco, Gaithersburg, MD, USA) supplemented with 10% fetal bovine serum (FBS) (Gemini, West Sacramento, CA, USA) and 1 × penicillin-streptomycin (Pen-Strep) at 100 U/mL–100 μg/mL (Gibco). Mouse MIN6 (Dr. Jun-Ichi Miyazaki, Osaka University, Japan) beta cells and human AD293 cells (Agilent, Santa Clara, CA, USA) 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, ON, Canada; purchased from Gibco) supplemented with 10% FBS, 1× Pen-Strep, and 2 mM L-glutamine (Gibco) at a concentration of 800–1,000 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 10 μL electroporation system (Invitrogen) with no more than 1 μg plasmid DNA per 10 μL treatment: *Perk*$^{C528X/C528X}$ MEF cells, 1 × 10^7 cells/mL, 1,650 V, 20 ms, 1 pulse; AD293 cells, 5 × 10^6 cells/mL, 1,245 V, 10 ms, 3 pulses; human islets, 500 IEQs/10 μL, 1,050 V, 40 ms, 1 pulse. The Neon procedure for electroporation of human islets was adapted from previously described protocols.[56,57] Briefly, about 1,000 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 20 μL R buffer with 2 μg of each plasmid DNA needed for the treatment. About 500 IEQs in 10 μL with 1 μg plasmid DNA were electroporated with 1 pulse at 1,050 V for 40 ms 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.[58] AAV8-U6-mINS2utr5sg-EGFP-2cut (6.15 × 10^{13} GC/mL) was produced and purified by Penn Vector Core. AAV-DJ-U6-mINS2utr5sg-EGFP-2cut (2.92 × 10^{12} GC/mL), AAV-DJ-U6-hINSin1sg-CopGFP-2cut (1.83 × 10^{13} GC/mL), AAV-DJ-U6-hINSin1sg-CopGFP-1cut (6.02 × 10^{12} GC/mL), and AAV-DJ-nEF-Cas9 (3.83 × 10^{12} GC/mL) were produced and purified as described below. Briefly, recombinant AAVs were produced in 293 AAV cells (Cell Biolabs, San Diego, CA, USA). Polyethylenimine (PEI, linear, MW 25,000) was used for transfection of three plasmids: the pAAV vector constructs, pAAV2/8-RC (Penn Vector Core) or pAAV-DJ (Cell Biolabs), and pHelper (Cell Biolabs). At 72 h post-transfection, cells were scrapped in their medium, centrifuged, and then frozen and thawed four times by placing alternately in dry ice-ethanol and a 37°C water bath to lyse the cells and release the virus. The resulting AAV crude lysate was purified by centrifugation at 54,000 rpm for 1 h 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, USA). Virus titers were determined by quantitative PCR (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 200 μL CMRL1066 medium with reduce FBS (2%) at a final titer of 9.0 × 10^{10} GC/mL. If 1 IEQ is considered to be 1,000 cells, the AAV incubation of human islets was at 60,000 MOI. CMRL1066 medium with 10% FBS was added to the sample at 1 day post-infection.

AAV administration via intravenous injection

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

Single-cell sorting

MEF cells and MIN6 cells were single-cell sorted according to size configuration or GFP fluorescent signal using Beckman Coulter MoFlo Astrios (Beckman-Coulter, Brea, CA, USA) performed by the 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 supplemented with 10% FBS was added to stop trypsinization. Cells were then transferred into a 15 mL tube and centrifuged at 200 × *g* for 1 min at room temperature. The cells were re-suspended thoroughly in DMEM with 1× Pen-Strep as single cells and were sorted into 96-well plate with full DMEM.

Genomic DNA extraction and diagnostic PCR analysis

Genomic DNA was extracted from cultured cells or mouse tissue by digesting in lysis buffer (5 mM EDTA, 0.2% SDS, 200 mM NaCl, and 100 mM Tris-HCl [pH 8.5]) with 100 μg/mL proteinase K overnight at 50°C. DNA was then precipitated with 1 volume of isopropanol and dissolved in TE buffer (10 mM Tris-HCl, 1.0 mM EDTA [pH 8.0]). Blood DNA was extracted using Monarch Genomic DNA Purification Kit (New England Biolabs). Diagnostic PCRs were performed using GoTaq Master Mix (Promega, Madison, WI, USA). 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, USA). Sanger sequencing of the PCR products was performed by the 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 S4 for primer sequences.

RNA isolation and qPCR analysis

Total RNA from cell lines and mouse tissues other than pancreas was extracted using the Quick-RNA Miniprep Kit (Zymo). Pancreas RNA was extracted as previously described https://www.pancreapedia.org/tools/methods/isolation-of-pancreatic-rna). Human islet RNA was extracted using AllPrep DNA/RNA/Protein Mini Kit (QIAGEN). Reverse transcription was performed using qScript cDNA SuperMix (Quanta, Beverly, MA, USA). 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, USA). 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 S5 for primer sequences.

ddPCR

Quantification of CRBR editing efficiency at the genomic DNA level was performed by ddPCR[59,60] using a QX200 ddPCR system (Bio-Rad, Hercules, CA, USA). The ddPCR reaction contained final concentrations of the following components: 1× EvaGreen Supermix (Bio-Rad), 150 nM of each primer, 0.13 U/μL of HindIII-HF (New England Biolabs), and template DNA (human AD293 cell or human islet DNA, 50 ng/reaction; mouse tissue DNA, 200 ng/reaction). Formation of droplet emulsions was performed by mixing 20 μL of PCR reaction and 70 μL of EvaGreen droplet generation oil (Bio-Rad) with the Automatic Droplet Generator (Bio-Rad) and was dispensed into 96-well plate. The emulsions containing approximately 20,000 droplets were cycled to amplicon saturation using a C1000 Thermal Cycler (Bio-Rad) operating at the following conditions: for 5 min at 95°C, 40 cycles of 30 s at 94°C, and for 1 min at 59°C–63.3°C (optimized for each primer set), for 5 min at 4°C, for 5 min at 90°C, and a 4°C hold. Amplitude of fluorescence by amplicons in each cycled droplet was measured using flow cytometry on a QX200 Droplet Reader (Bio-Rad) set on the EVA channel. The QuantaSoft droplet reader software (v1.4.0.99; Bio-Rad) was used to cluster droplets into distinct positive and negative fluorescent groups and fit the fraction of positive droplets to a Poisson algorithm to determine the starting concentration (copies/μL) of the input DNA sample. CRBR editing efficiency was calculated by the ratio of the 5′ junction concentration (including clean CRBR integration and 5′ CRBR whole donor integration) to the reference gene concentration. The reference genes in mouse and human genome, *mRpp30* (chr19) or *hRPP30* (chr10), have the same copy number as the chromosomal alleles to be edited, mouse *Ins2* locus on chr7 or human *INS* locus on chr11. See Table S6 for primer sequences.

GFP imaging and histological analysis

MIN6 cells and human islets were imaged as live cultures, and images were captured using the fluorescein isothiocyanate (FITC) and Trans-illumination channels of the ECHO Revolve microscope and the associated software (Echo Labs, San Diego, CA, USA). Whole pancreata were harvested and paraffin embedded as previously described.[26] Sectioned (6 μm in thickness) slides were dewaxed and hematoxylin and eosin stained by Leica Autostainer ST5010 XL (Wetzlar, Germany). 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). Lysate proteins from tissues or MEF cells were denatured by boiling the lysates in 2× SDS sample buffer for 5 min 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, USA) in carbonate transfer buffer using wet transfer conditions (Criterion Blotter, Bio-Rad). Primary antibodies (diluted in 5% BSA-TBST) used include: phospho-PERK (Thr980) (#3179, Cell Signaling, Danvers, MA, USA), 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 Biosciences, Lincoln, NE, USA).

Statistical analysis

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

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.ymthe.2021.04.017.

ACKNOWLEDGMENTS

Human pancreatic islets were provided by the NIDDK-funded Integrated Islet Distribution Program (IIDP) (RRID:SCR _014387) at City of Hope, NIH grant no. 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*$^{rPERK\text{-}CRBR}$ 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. We thank the Makova Lab at Penn State University for providing the ddPCR system and Dr. Tomaszkiewicz for her help with optimizing the assay. This work was supported by NIH grant no. R01 DK88140.

AUTHOR CONTRIBUTIONS

Conceptualization, J.H., R.A.B., B.C.M., and D.R.C.; Methodology, J.H., R.A.B., and D.R.C.; Validation, J.H., R.A.B., B.C.M., and D.R.C.; Formal analysis, J.H., A.L., and R.A.B.; Investigation, J.H., R.A.B., and A.L.; Resources, Z.P.; Writing – original draft, J.H.; Writing – review & editing, J.H., R.A.B., B.C.M., and D.R.C.; Visualization, J.H.; Funding acquisition, J.H., R.A.B., B.C.M., and D.R.C.

DECLARATION OF INTERESTS

J.Hu, R. A. Bourne, B.C. McGrath, A. Lin, and D.R. Cavener have submitted a provisional patent related to this work.

REFERENCES

1. Wilson, J.M. (2019). Cycling at the Frontiers of Gene Therapy. Hum. Gene Ther. Clin. Dev. *30*, 47–49.
2. Dunbar, C.E., High, K.A., Joung, J.K., Kohn, D.B., Ozawa, K., and Sadelain, M. (2018). Gene therapy comes of age. Science *359*, eaan4672.
3. Lundstrom, K. (2018). Viral Vectors in Gene Therapy. Diseases *6*, 42.
4. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.
5. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

6. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

7. Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (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, R.L., Benedetti, E., Grompe, M., Koteliansky, V., Sharp, P.A., Jacks, T., and Anderson, D.G. (2014). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat. Biotechnol. *32*, 551–553.

9. Yin, H., Song, C.Q., Dorkin, J.R., Zhu, L.J., Li, Y., Wu, Q., Park, A., Yang, J., Suresh, S., Bizhanova, A., et al. (2016). Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. Nat. Biotechnol. *34*, 328–333.

10. Tran, N.T., Graf, R., Wulf-Goldenberg, A., Stecklum, M., Strauß, G., Kühn, R., Kocks, C., Rajewsky, K., and Chu, V.T. (2020). CRISPR-Cas9-Mediated ELANE Mutation Correction in Hematopoietic Stem and Progenitor Cells to Treat Severe Congenital Neutropenia. Mol. Ther. *28*, 2621–2634.

11. Ohmori, T., Nagao, Y., Mizukami, H., Sakata, A., Muramatsu, S.I., Ozawa, K., Tominaga, S.I., Hanazono, Y., Nishimura, S., Nureki, O., and Sakata, Y. (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, C.A., White, J., Zhang, J., Che, Y., Saveliev, A., McMenamin, D., He, Z., Latshaw, C., 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, L.E., Chenais, N.A.L., Marchetti, T., Parrini, M., Contestabile, A., Cancedda, L., and Ghezzi, D. (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., Zhao, H., Bao, J., Zhang, M., Qiu, Z., and Xue, T. (2019). In vivo genome editing rescues photoreceptor degeneration via a Cas9/RecA-mediated homology-directed repair pathway. Sci. Adv. *5*, eaav3335.

15. Cox, D.B., Platt, R.J., 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, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature *533*, 420–424.

18. Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage. Nature *551*, 464–471.

19. Yeh, W.H., Chiang, H., Rees, H.A., Edge, A.S.B., and Liu, D.R. (2018). In vivo base editing of post-mitotic sensory cells. Nat. Commun. *9*, 2184.

20. Villiger, L., Grisch-Chan, H.M., Lindsay, H., Ringnalda, F., Pogliano, C.B., Allegri, G., Fingerhut, R., Häberle, J., Matos, J., Robinson, M.D., et al. (2018). Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. Nat. Med. *24*, 1519–1525.

21. Bansal, V., Gassenhuber, J., Phillips, T., Oliveira, G., Harbaugh, R., Villarasa, N., Topol, E.J., Seufferlein, T., and Boehm, B.O. (2017). Spectrum of mutations in monogenic diabetes genes identified from high-throughput DNA sequencing of 6888 individuals. BMC Med. *15*, 213.

22. Rebbeck, T.R., Friebel, T.M., Friedman, E., Hamann, U., Huo, D., Kwong, A., Olah, E., Olopade, O.I., Solano, A.R., Teo, S.H., et al.; EMBRACE; GEMO Study Collaborators; HEBON (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, E.J., Hatanaka, F., Yamamoto, M., Araoka, T., Li, Z., et al. (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature *540*, 144–149.

25. Harding, H.P., Zeng, H., Zhang, Y., Jungries, R., Chung, P., Plesken, H., Sabatini, D.D., and Ron, D. (2001). Diabetes mellitus and exocrine pancreatic dysfunction in perk-/- mice reveals a role for translational control in secretory cell survival. Mol. Cell *7*, 1153–1163.

26. Zhang, P., McGrath, B., Li, S., Frank, A., Zambito, F., Reinert, J., Gannon, M., Ma, K., McNaughton, K., and Cavener, D.R. (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. Mol. Cell. Biol. *22*, 3864–3874.

27. Zhang, W., Feng, D., Li, Y., Iida, K., McGrath, B., and Cavener, D.R. (2006). PERK EIF2AK3 control of pancreatic beta cell differentiation and proliferation is required for postnatal glucose homeostasis. Cell Metab. *4*, 491–497.

28. Li, Y., Iida, K., O'Neil, J., Zhang, P., Li, S., Frank, A., Gabai, A., Zambito, F., Liang, S.H., Rosen, C.J., and Cavener, D.R. (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, B.C., Frank, A., and Cavener, D.R. (2007). PERK eIF2 alpha kinase is required to regulate the viability of the exocrine pancreas in mice. BMC Cell Biol. *8*, 38.

30. Hellen, C.U., and Sarnow, P. (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev. *15*, 1593–1612.

31. Gunišová, S., Hronová, V., Mohammad, M.P., Hinnebusch, A.G., and Valášek, L.S. (2018). Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. FEMS Microbiol. Rev. *42*, 165–192.

32. Cheng, H., Wolfe, S.H., Valencia, V., Qian, K., Shen, L., Phillips, M.I., Chang, L.J., and Zhang, Y.C. (2007). Efficient and persistent transduction of exocrine and endocrine pancreas by adeno-associated virus type 8. J. Biomed. Sci. *14*, 585–594.

33. Rehman, K.K., Trucco, M., Wang, Z., Xiao, X., and Robbins, P.D. (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, N.L., Havinga, R., Kluiver, J., Groen, A.K., and Kruit, J.K. (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. Grimm, D., Lee, J.S., Wang, L., Desai, T., Akache, B., Storm, T.A., and Kay, M.A. (2008). In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. J. Virol. *82*, 5887–5911.

36. Flotte, T.R., Afione, S.A., Solow, R., Drumm, M.L., Markakis, D., Guggino, W.B., Zeitlin, P.L., and Carter, B.J. (1993). Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter. J. Biol. Chem. *268*, 3781–3790.

37. Haberman, R.P., McCown, T.J., and Samulski, R.J. (2000). Novel transcriptional regulatory signals in the adeno-associated virus terminal repeat A/D junction element. J. Virol. *74*, 8732–8739.

38. Da Silva Xavier, G. (2018). The Cells of the Islets of Langerhans. J. Clin. Med. *7*, E54.

39. Rehman, K.K., Wang, Z., Bottino, R., Balamurugan, A.N., Trucco, M., Li, J., Xiao, X., and Robbins, P.D. (2005). Efficient gene delivery to human and rodent islets with double-stranded (ds) AAV-based vectors. Gene Ther. *12*, 1313–1323.

40. Craig, A.T., Gavrilova, O., Dwyer, N.K., Jou, W., Pack, S., Liu, E., Pechhold, K., Schmidt, M., McAlister, V.J., Chiorini, J.A., et al. (2009). Transduction of rat pancreatic islets with pseudotyped adeno-associated virus vectors. Virol. J. *6*, 61.

41. Canny, M.D., Moatti, N., Wan, L.C.K., Fradet-Turcotte, A., Krasner, D., Mateos-Gomez, P.A., Zimmermann, M., Orthwein, A., Juang, Y.C., Zhang, W., 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.e5.

43. Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science *351*, 400–403.

44. Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Castellanos Rivera, R.M., Madhavan, S., Pan, X., Ran, F.A., Yan, W.X., et al. (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science *351*, 403–407.

45. Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A., et al. (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science *351*, 407–411.

46. Suzuki, K., Yamamoto, M., Hernandez-Benitez, R., Li, Z., Wei, C., Soligalla, R.D., Aizawa, E., Hatanaka, F., Kurita, M., Reddy, P., et al. (2019). Precise in vivo genome editing via single homology arm donor mediated intron-targeting gene integration for genetic disease correction. Cell Res. *29*, 804–819.

47. Sowers, C.R., Wang, R., Bourne, R.A., McGrath, B.C., Hu, J., Bevilacqua, S.C., Paton, J.C., Paton, A.W., Collardeau-Frachon, S., Nicolino, M., and Cavener, D.R. (2018). The protein kinase PERK/EIF2AK3 regulates proinsulin processing not via protein synthesis but by controlling endoplasmic reticulum chaperones. J. Biol. Chem. *293*, 5134–5149.

48. Maxwell, K.G., Augsornworawat, P., Velazco-Cruz, L., Kim, M.H., Asada, R., Hogrebe, N.J., Morikawa, S., Urano, F., and Millman, J.R. (2020). Gene-edited human stem cell-derived β cells from a patient with monogenic diabetes reverse preexisting diabetes in mice. Sci. Transl. Med. *12*, eaax9106.

49. Wilbie, D., Walther, J., and Mastrobattista, E. (2019). Delivery Aspects of CRISPR/Cas for in Vivo Genome Editing. Acc. Chem. Res. *52*, 1555–1564.

50. Yin, H., Song, C.Q., Suresh, S., Wu, Q., Walsh, S., Rhym, L.H., Mintzer, E., Bolukbasi, M.F., Zhu, L.J., Kauffman, K., et al. (2017). Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. Nat. Biotechnol. *35*, 1179–1187.

51. Ates, I., Rathbone, T., Stuart, C., Bridges, P.H., and Cottle, R.N. (2020). Delivery Approaches for Therapeutic Genome Editing and Challenges. Genes (Basel) *11*, E1113.

52. Li, A., Lee, C.M., Hurley, A.E., Jarrett, K.E., De Giorgi, M., Lu, W., Balderrama, K.S., Doerfler, A.M., Deshmukh, H., Ray, A., et al. (2018). A Self-Deleting AAV-CRISPR System for *In Vivo* Genome Editing. Mol. Ther. Methods Clin. Dev. *12*, 111–122.

53. Yang, H., Wang, H., and Jaenisch, R. (2014). Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. Nat. Protoc. *9*, 1956–1968.

54. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

55. Jiang, H.Y., Wek, S.A., McGrath, B.C., Scheuner, D., Kaufman, R.J., Cavener, D.R., and Wek, R.C. (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.

56. Tamaki, S., Nye, C., Slorach, E., Scharp, D., Blau, H.M., Whiteley, P.E., and Pomerantz, J.H. (2014). Simultaneous silencing of multiple RB and p53 pathway members induces cell cycle reentry in intact human pancreatic islets. BMC Biotechnol. *14*, 86.

57. Lefebvre, B., Vandewalle, B., Longue, J., Moerman, E., Lukowiak, B., Gmyr, V., Maedler, K., Kerr-conte, J., and Pattou, F. (2010). Efficient gene delivery and silencing of mouse and human pancreatic islets. BMC Biotechnol. *10*, 28.

58. Chen, Y.C., Ma, N.X., Pei, Z.F., Wu, Z., Do-Monte, F.H., Keefe, S., Yellin, E., Chen, M.S., Yin, J.C., Lee, G., 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.

59. Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., Bright, I.J., Lucero, M.Y., Hiddessen, A.L., Legler, T.C., et al. (2011). High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal. Chem. *83*, 8604–8610.

60. Tomaszkiewicz, M., Rangavittal, S., Cechova, M., Campos Sanchez, R., Fescemyer, H.W., Harris, R., Ye, D., O'Brien, P.C., Chikhi, R., Ryder, O.A., et al. (2016). A time- and cost-effective strategy to sequence mammalian Y Chromosomes: an application to the de novo assembly of gorilla Y. Genome Res. *26*, 530–540.