

# Engineering Oncogenic Heterozygous Gain-of-Function Mutations in Human Hematopoietic Stem and Progenitor Cells

Tommaso Sconocchia<sup>1</sup>, Johannes Foßelteder<sup>1</sup>, Thomas Köhnke<sup>2</sup>, Ravindra Majeti<sup>2</sup>, Andreas Reinisch<sup>1,3</sup>

<sup>1</sup>Division of Hematology, Department of Internal Medicine, Medical University of Graz <sup>2</sup>Division of Hematology, Stanford Cancer Institute, Stanford Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine <sup>3</sup>Department of Blood Group Serology and Transfusion Medicine, Medical University of Graz

## Corresponding Author

Andreas Reinisch

a.reinisch@medunigraz.at

## Citation

Sconocchia, T., Foßelteder, J., Köhnke, T., Majeti, R., Reinisch, A. Engineering Oncogenic Heterozygous Gain-of-Function Mutations in Human Hematopoietic Stem and Progenitor Cells. *J. Vis. Exp.* (193), e64558, doi:10.3791/64558 (2023).

## Date Published

March 10, 2023

## DOI

10.3791/64558

## URL

jove.com/video/64558

## Abstract

Throughout their lifetime, hematopoietic stem and progenitor cells (HSPCs) acquire somatic mutations. Some of these mutations alter HSPC functional properties such as proliferation and differentiation, thereby promoting the development of hematologic malignancies. Efficient and precise genetic manipulation of HSPCs is required to model, characterize, and better understand the functional consequences of recurrent somatic mutations. Mutations can have a deleterious effect on a gene and result in loss-of-function (LOF) or, in stark contrast, may enhance function or even lead to novel characteristics of a particular gene, termed gain-of-function (GOF). In contrast to LOF mutations, GOF mutations almost exclusively occur in a heterozygous fashion. Current genome-editing protocols do not allow for the selective targeting of individual alleles, hampering the ability to model heterozygous GOF mutations. Here, we provide a detailed protocol on how to engineer heterozygous GOF hotspot mutations in human HSPCs by combining CRISPR/Cas9-mediated homology-directed repair and recombinant AAV6 technology for efficient DNA donor template transfer. Importantly, this strategy makes use of a dual fluorescent reporter system to allow for the tracking and purification of successfully heterozygously edited HSPCs. This strategy can be employed to precisely investigate how GOF mutations affect HSPC function and their progression toward hematological malignancies.

## Introduction

With the development of the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 technology, a new and extremely powerful instrument has been added to the

toolkit of scientists. This technology allows for the precise engineering of the genome and has proven itself to be extremely useful not only for research purposes (reviewed

in Hsu et al.<sup>1</sup>) but more recently has also been successfully translated into the clinical setting<sup>2,3,4</sup>. CRISPR/Cas9 editing strategies rely on the activity of a Cas9 protein and a single-guide RNA (sgRNA)<sup>5,6,7</sup>. In the host cell, the Cas9 protein is guided to a specific site in the DNA that is complementary to the sgRNA sequence and will introduce a DNA double-strand break (DSB). Once a DSB is generated, there are two main and competing repair mechanisms that can occur: non-homologous end joining (NHEJ) and homology-directed repair (HDR). NHEJ is an error-prone, predominantly used repair mechanism leading to insertions and deletions (indels), whereas HDR, by using the sister chromatid as a repair template, is very precise but limited to the S or G2 phase of the cell cycle<sup>8</sup>. In genome engineering, HDR can be utilized for the targeted modification of DNA by providing a donor template that is flanked by homology arms identical to both DNA ends of the Cas9-induced DSB (**Figure 1**). The type of donor template that is used for HDR can greatly impact editing efficiency. For genetic engineering in human HSPCs, adeno-associated virus serotype 6 (AAV6) has recently been described to be an excellent vehicle for the delivery of single-stranded DNA templates<sup>9,10</sup>.

CRISPR/Cas9 genome engineering can be employed therapeutically to correct deleterious mutations<sup>11</sup> but can also be utilized to introduce pathogenic mutations into the DNA to model cancer development<sup>12</sup>. Blood cancer, such as leukemia, develops *via* the sequential acquisition of somatic mutations in healthy HSPCs<sup>13,14</sup>. Early genetic events lead to a clonal proliferative advantage, resulting in clonal hematopoiesis of indeterminate potential (CHIP)<sup>15,16</sup>. Further acquisition of mutations will eventually lead to leukemic transformation and the development of the disease.

Somatic mutations can be found in genes controlling self-renewal, survival, proliferation, and differentiation<sup>17</sup>.

Introducing individual mutations *via* genome engineering into healthy HSPCs allows for precisely modeling this stepwise leukemogenic process. The limited number of recurrent mutations found in myeloid neoplasms such as acute myeloid leukemia (AML)<sup>18,19</sup> makes this disease particularly amenable to being recapitulated using genome engineering tools.

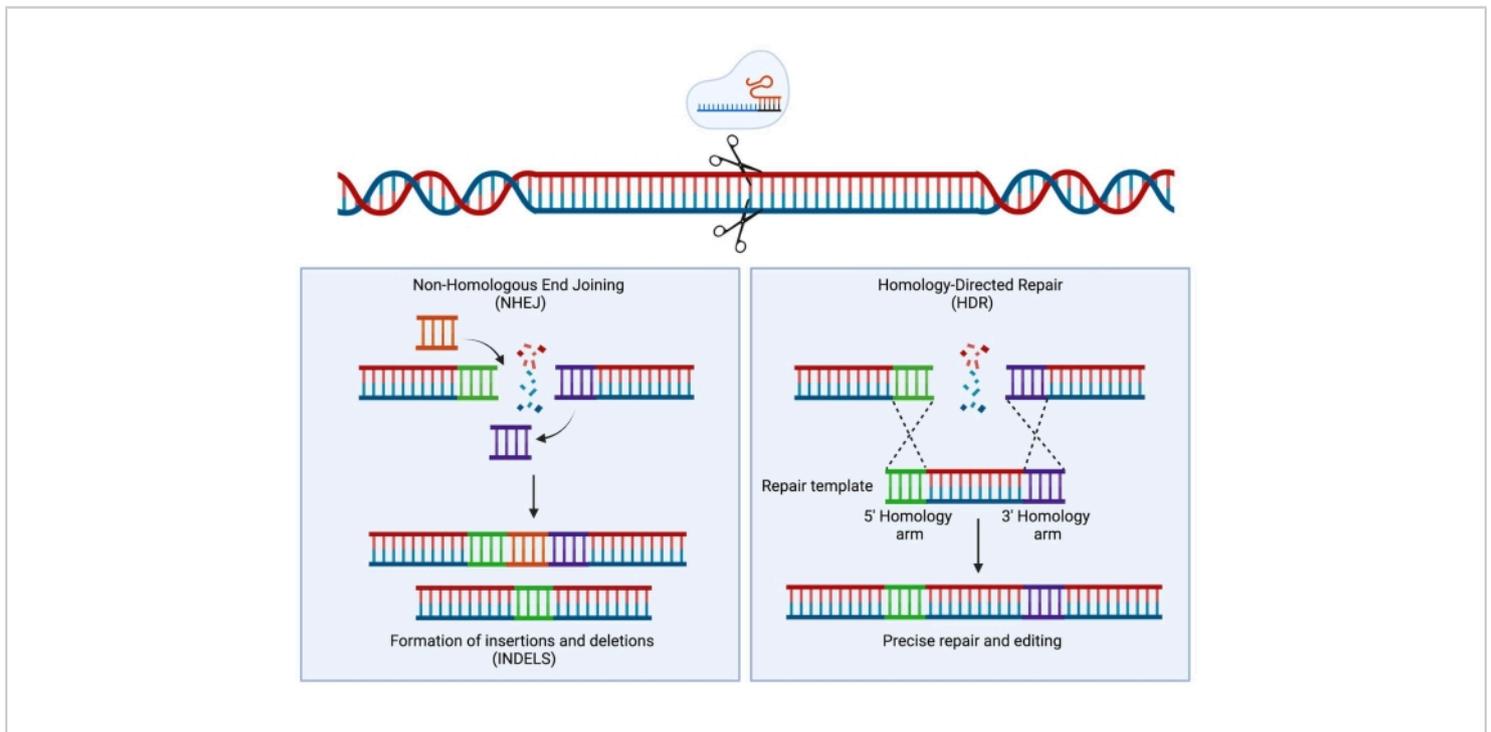
Somatic mutations can emerge on only one allele (monoallelic/heterozygous mutations) or on both alleles (biallelic/homozygous mutations) and can have profound effects on the gene's function, which could cause a loss-of-function (LOF) or a gain-of-function (GOF). LOF mutations lead to a reduced (if one allele is affected) or complete (if both alleles are affected) LOF of the gene, whereas GOF mutations lead to an increased activation or novel function of the gene. GOF mutations are typically heterozygous<sup>20</sup>.

Importantly, the zygosity (hetero- vs. homozygous) has major implications for the attempt to faithfully model a mutation; therefore, the targeted manipulation of only one allele of a gene is necessary to engineer heterozygous hotspot GOF mutations. Error-prone NHEJ leads to indels of different lengths<sup>21</sup> that can lead to varying, unpredictable biological consequences. However, since NHEJ is the predominant repair program employed by cells after the introduction of a DSB, most CRISPR/Cas9 platforms currently used to manipulate HSPCs do not allow for precisely predicting the genetic outcome<sup>22,23</sup>. In contrast, the introduction of a CRISPR/Cas9-mediated double-strand break (DSBs), combined with the use of recombinant adeno-associated virus (rAAV) vector-based DNA donor templates for genome engineering *via* HDR, allows for the allele-specific insertion

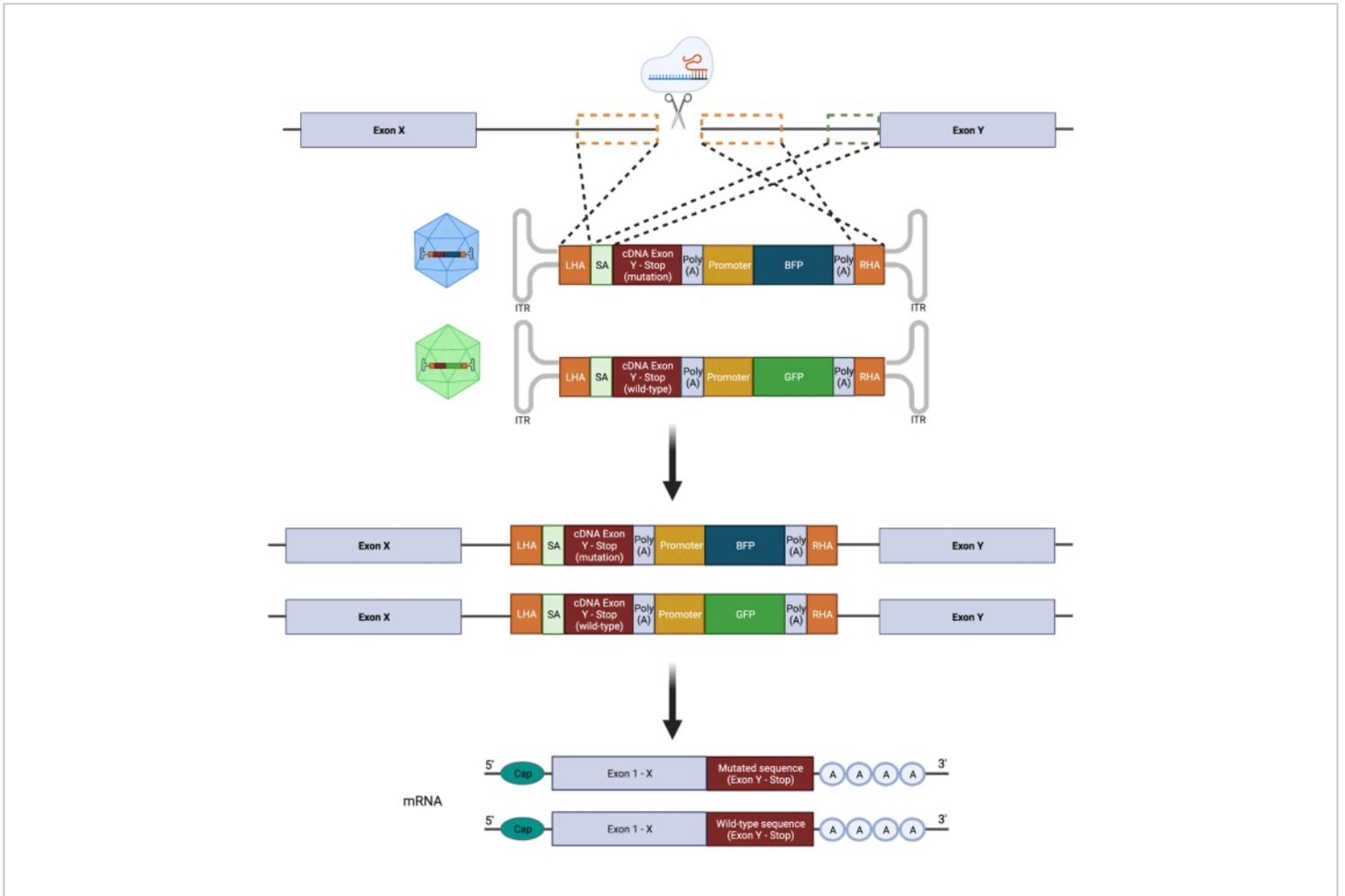
of mutations in human HSPCs<sup>11,24</sup>. The simultaneous integration of a mutant and a wild-type (WT) sequence coupled with distinct fluorescent reporters on the individual alleles can be performed to select for a heterozygous genotype (**Figure 2**). This strategy can be leveraged as a powerful tool to precisely characterize the effects of recurrent, leukemic, heterozygous GOF hotspot mutations on HSPC function, disease initiation, and progression.

In this article, a detailed protocol for the efficient engineering of recurrently mutated heterozygous GOF mutations in primary human HSPCs is provided. This strategy combines the use of CRISPR/Cas9 and a dual AAV6 transduction to provide WT and mutant DNA donor templates for the prospective generation of the heterozygous GOF mutation. As an example, engineering of the recurrent type 1 mutations (52 bp deletion) in the calreticulin (*CALR*) gene will be

shown<sup>25</sup>. The heterozygous GOF mutation in exon 9 of *CALR* is recurrently found in myeloproliferative disorders such as essential thrombocythemia (ET) and primary myelofibrosis (PMF)<sup>26</sup>. *CALR* is an endoplasmic reticulum resident protein that has primarily a quality control function in the folding process of newly synthesized proteins. Its structure can be divided into three main domains: an amino (N)-terminal domain and a proline-rich P domain, which are involved in the chaperone function of the protein, and a C-domain, which is involved in calcium storage and regulation<sup>27,28</sup>. *CALR* mutations cause a +1 frameshift, leading to the transcription of a novel extended C-terminal end and the loss of the endoplasmic reticulum (ER)-retention signal (KDEL). Mutant *CALR* has been shown to bind the thrombopoietin (TPO) receptor, thereby leading to TPO-independent signaling with increased proliferation<sup>29</sup>.



**Figure 1: NHEJ and HDR repair.** Simplified schematic representation of NHEJ and HDR repair mechanisms following the introduction of a double-stranded break in the DNA. [Please click here to view a larger version of this figure.](#)



**Figure 2: Schematic overview of biallelic HDR editing strategy.** Schematic representation showing the integration of the donor templates into the targeted alleles followed by their translation into functioning mRNAs. The orange dotted boxes indicate the regions corresponding to the left homology arm (LHA) and right homology arm (RHA). The ideal size of the HAs is 400 bp each. The green dotted box represents the region corresponding to the SA sequence. The size of the SA is 150 bp. [Please click here to view a larger version of this figure.](#)

## Protocol

This protocol requires the use of healthy donor-derived CD34<sup>+</sup> HSPCs and requires ethical approval from the local institutional review boards (IRB) and signed informed consent. The CD34<sup>+</sup> HSPCs used in this protocol were isolated from the umbilical cord blood (UCB) of term deliveries (>34 weeks of gestation). Informed consent was obtained

from the mothers prior to delivery, and ethical approval for the collection of UCB was obtained (IRB approval: 31-322 ex 18/19) from the Medical University of Graz. A full list of materials used in this protocol can be found in the **Table of Materials**.

## 1. sgRNA design and evaluation of cutting efficiencies

1. Search for the location of the desired mutation and the correct transcript on an online database tool (e.g., COSMIC, <https://cancer.sanger.ac.uk/cosmic>).
2. Select an sgRNA adjacent to the mutation (exonic targeting) or in the previous intron (intrinsic targeting) by using an sgRNA design tool. In this protocol, the online tool from Benchling is used. See **The Table of Materials** for a list of possible online tools for sgRNA design.
3. Select the **+** (**create**) option **> DNA Sequence > Import DNA Sequences > Import from Databases**. Enter the desired gene e.g., CALR and select human as the species **> Search >** Select the correct transcript (e.g., CALR-201 -ENST00000316448) **> Import**. Select the region of interest in which to introduce the DS break (e.g., intron 7).
4. Select the **CRISPR** option on the right-hand side of the screen and select **Design and analyze guides**. Select **single guide**, keep the guide length to 20 bp, and PAM sequence for editing with SpCas9 (NGG).
5. Pick a guide with a high on-target score (high chance for editing at the desired locus) and a high off-target score (low chance of editing at unwanted loci). Pick at least three guides to test in order to find the best performing sgRNA. Order the sgRNA as a chemically modified synthetic sgRNA from a commercial vendor.  
**NOTE:** It is advisable at this stage to order a small amount of the synthetic sgRNAs for the initial screen. Once a good performing sgRNA has been identified, proceed with a larger order of the selected sgRNA.

6. Thaw  $2 \times 10^5$ - $5 \times 10^5$  CD34<sup>+</sup> HSPCs. Transfer the cells into 10 mL of pre-warmed RPMI1640 supplemented with 1% antibiotics (e.g., penicillin/streptomycin).  
**NOTE:** CD34<sup>+</sup> HSPCs with a purity >90% should be used to achieve the best and most reproducible results.
7. Centrifuge at 350 x g at room temperature (RT) for 10 min. Count the cells with a hemocytometer and suspend the cells in SFEM II medium supplemented with 0.2% penicillin/streptomycin (P/S), 100 ng/mL thrombopoietin (TPO), 100 ng/mL stem cell factor (SCF), 100 ng/mL FMS-like tyrosine kinase 3 ligand (FLT3L), 100 ng/mL interleukin-6 (IL-6), 35 nM UM171, and 0.75 μM StemRegenin1 (SR1) to a concentration of  $2.5 \times 10^5$  cells/mL. Incubate at 37 °C/5% CO<sub>2</sub> for 48 h - 72 h.  
**NOTE:** The complete medium will be referred to from now on as the HSPC retention medium.
8. Collect the cells in a 15 mL tube and count the cells. Check the cell viability by trypan blue exclusion.
  1. Before starting, turn on the transfection system and select the option for the cuvettes. Select the appropriate program and nucleofection solution for HSPCs (see **Table of Materials**). Between  $2 \times 10^5$  and  $5 \times 10^6$  cells can be nucleofected in one 100 μL cuvette. Set aside a small number of cells (e.g.,  $1 \times 10^5$ - $2 \times 10^5$ ) to keep in culture for 48 h to be used as a WT control.
9. Prepare the RNP complex. In a 1.5 mL tube, add 15 μg of Cas9 and 8 μg of sgRNA (molar ratio 1:2.5) and incubate at 25 °C for 10 min in a heating block.
10. While the RNP complex is incubating, centrifuge the cells at 350 x g at RT for 5 min and discard the supernatant

- by using a pipette. Suspend the cells in 100  $\mu\text{L}$  of nucleofection solution.
11. Mix the cells with the RNP complex and transfer into the cuvette. Gently tap the cuvette to remove any eventual air bubbles that might have formed during the transfer.
  12. Insert the cuvette into the holder of the transfection system and electroporate the cells with the DZ-100 program.
  13. Immediately after electroporation, add 400  $\mu\text{L}$  of pre-warmed HSPC retention medium without P/S.
  14. Transfer the cells with a fine transfer pipette into a culture plate containing pre-warmed HSPC retention medium without P/S. Depending on the cell number, use an appropriate culture plate (24-, 12-, or 6-well plate) in order to reach a density between  $0.25 \times 10^6$  and  $1 \times 10^6$  cells/mL.
  15. Transfer the plate to the incubator at  $37^\circ\text{C}/5\% \text{CO}_2$ . Incubate the nucleofected cells for 6-8 h.
  16. Following 6-8 h, remove the old medium and replace with fresh pre-warmed HSPC retention medium supplemented with P/S. Suspend the cells at a concentration between  $2.5 \times 10^5$  and  $5 \times 10^5$  and transfer to a cell culture plate (24-, 12-, or 6-well plate). Incubate the cells for 48 h at  $37^\circ\text{C}/5\% \text{CO}_2$ .
  17. Harvest  $2 \times 10^5$  cells and centrifuge at  $350 \times g$  at RT for 5 min. Before starting, set the heating blocks at  $65^\circ\text{C}$  and  $98^\circ\text{C}$ .
  18. Discard the supernatant and suspend the cells in 1 mL of 1x DPBS in a 1.5 mL tube. Centrifuge at  $350 \times g$  at RT for 5 min.
  19. Discard the supernatant and suspend the cells in 50  $\mu\text{L}$  of DNA extraction solution. Vortex for 15 s.
  20. Incubate for 6 min at  $65^\circ\text{C}$ . Vortex for 15 s and incubate for 2 min at  $98^\circ\text{C}$ .
  21. Amplify the region of the DSB by PCR using primers that generate an amplicon of about 400-600 bp with the DSB in the center. Use 0.5-1  $\mu\text{L}$  of extracted DNA for the PCR reaction.
 

**NOTE:** Due to the composition of the DNA extraction solution, the DNA concentrations cannot be accurately quantified by spectrophotometry.
  22. Run the PCR product with a DNA ladder on a 1.5 % agarose gel at 100 V for 45-60 min. Place the gel on a blue light or UV transilluminator.
  23. Extract the DNA band of the correct size from the gel by using a commercially available kit according to the manufacturer's instructions (see **Table of Materials**).
  24. Sequence the samples by Sanger sequencing using the forward primer or reverse primer from the PCR. For PCR products of 400-600 bp in length, 75 ng of DNA is required for sequencing with a concentration of 5 ng/mL in a total volume of 15  $\mu\text{L}$ .
  25. Analyze the editing efficiency of the sgRNAs by uploading the sequencing files to a program designed to calculate the editing efficiency by identifying the insertion and deletions produced by the sgRNA. Select the best-performing sgRNA to proceed with.
 

**NOTE:** Dedicated online tools are listed in the **Table of Materials**. This analysis requires the .ab1 files of the transfected and WT HSPCs, the sgRNA sequence, and the PAM sequence.

## 2. Homology-directed repair (HDR) vector construction

1. HDR template design

**NOTE:** Two HDR templates should be designed: a template for the WT sequence and a template for the mutated sequence.

1. Import the genomic sequence and the coding sequence (CDS) of the desired gene into an appropriate software for molecular cloning (dedicated tools can be found listed in the **Table of Materials**). From the genomic sequence file, design the left homology arm (HA) by selecting ideally 400 bp on the 5' end of the DSB. Paste this sequence into a new file.
2. If an intron is targeted with the sgRNA, a splice acceptor (SA) sequence needs to be included. Select the last 150 bp of the targeted intron and paste it after the left HA. If the exon is directly targeted, this sequence is not necessary (**Figure 3**).
3. From the CDS file, select the cDNA of interest. In case an exonic sequence is targeted, ensure the cDNA starts immediately downstream of the DSB site and includes all the following exons of the gene of interest as well as the stop codon. In case an intron is targeted, ensure the cDNA starts with the first codon of the following exon. Codon-optimize the cDNA and insert it after the SA sequence. Use dedicated online tools for this purpose (**Table of Materials**).
4. Insert a 3' polyadenylation (PolyA) signal after the cDNA of interest (e.g., SV40 or bGH) (**Table 1**). Following the PolyA, insert a promoter sequence (e.g., spleen focus forming virus [SFFV] or polyubiquitin C [UBC], **Table 1**) for the fluorescent protein.

5. Insert the sequence for the fluorescent protein (i.e., either GFP, BFP, or mCherry). Insert a second but different PolyA signal.

**NOTE:** The insertion of a different PolyA sequence will avoid problems such as bacterial recombination of the plasmid or problems with sequence aligning because of two identical sequences in the template.

6. From the genomic sequence file, design the right HA by selecting ideally 400 bp on the 3' end of the DSB. Insert this sequence after the second PolyA.
7. Create a copy of the whole template and modify the cDNA of interest so that it contains the sequence of the desired mutation.
8. Exchange the fluorescent protein with a different fluorescent protein. For example, if GFP was used for the WT template, then exchange it with BFP or mCherry for the mutant template. Construct and clone the HDR templates into the pAAV-MCS plasmid (or other suitable backbones).  
**NOTE:** The fragments required for the assembly can be produced either by PCR or ordered commercially and need to contain overlapping sequences with their neighboring fragments. If the mutation of interest is a point mutation, a small insertion, or a small deletion, the mutated cDNA can be produced by PCR by performing a site-directed mutagenesis on the HDR template containing the WT cDNA.
9. Transform competent *Escherichia coli* with the assembled product by using the heat shock method. Before starting, place the bacteria to thaw on ice for 10 min.
10. Add 2  $\mu$ L of the assembled products to 50  $\mu$ L of bacteria. Mix the contents by gentle flicking.

11. Place the samples on ice for 30 min. Transfer the samples to a thermoblock set to 42 °C for 30 s.
  12. Transfer the samples on ice for 5 min. Add 450 µL of room temperature SOC medium and incubate the samples at 37 °C for 1 h.
  13. Spread the samples onto LB agar plates containing ampicillin. For each sample, spread 100 µL of three different dilutions of the bacterial solution (undiluted, 1:5, and 1:10) in order to obtain LB agar plates with single colonies that can be picked. Incubate the plates overnight at 37 °C.
  14. The next day, pick three colonies per sample. Transfer the colonies to 15 mL tubes with caps containing 4 mL of LB medium supplemented with ampicillin and incubate overnight in a shaker at 37 °C.
  15. Aliquot 500 µL of the bacterial solution from each colony and store it in the fridge for later use. Perform a mini-prep to extract plasmid DNA as per the manufacturer's instructions.
  16. Send the samples for Sanger sequencing to confirm the correct assembly of the plasmids. Use enough primers distributed throughout the plasmid to ensure uninterrupted sequence confirmation, ideally covering each region twice with forward and reverse reads.
  17. Add 200 µL of the bacterial solution containing the correctly assembled plasmid to 200 mL of LB media supplemented with ampicillin and incubate in a shaker overnight at 37 °C.
  18. Perform a midi- or maxi-prep to extract the plasmid DNA as per the manufacturer's instructions. Store the plasmid DNA at -20 °C.
2. Recombinant AAV6 preparation
 

**NOTE:** Two separate AAV6 preparations will need to be performed: one for the rAAV6 with the WT HDR template and one for the rAAV6 with the mutated HDR template. This section describes the steps needed for the preparation of only one virus.

    1. Thaw HEK293T cells, transfer the cells into 10 mL of pre-warmed DMEM supplemented with 10% FBS, 1% P/S, and 25 mM HEPES, and centrifuge at 350 x g at RT for 5 min.
    2. Suspend the cells at a concentration of 1 x 10<sup>5</sup> cells/mL and transfer to a suitable flask (e.g., a 175 cm<sup>2</sup> flask). Transfer the flask to an incubator at 37 °C/5% CO<sub>2</sub>.
 

**NOTE:** HEK293T cells should be thawed in advance in order to allow the cells to fully recover and to obtain a sufficient number of cells (11 x 10<sup>7</sup> cells are necessary for the production of one virus). It is recommended to use the cells after a minimum of three passages and to keep the cells below 20 passages. HEK293T cells should be split three times a week. While maintaining the cells, these should not exceed 70%-80% confluency. The DMEM used in the AAV preparation should also already be supplemented with L-glutamine and sodium pyruvate.
    3. Collect the cells in a 50 mL tube and centrifuge at 350 x g at RT for 5 min.

4. Suspend the cells in 20 mL of DMEM supplemented with 10% FBS, 1% P/S, and 25 mM HEPES and count with a hemocytometer. Use trypan blue for dead cell exclusion.
5. Seed  $3 \times 10^6$  cells in 20 mL of DMEM supplemented with 10% FBS, 1% P/S, and 25 mM HEPES in a 175 cm<sup>2</sup> flask. For the production of one virus, prepare at least four 175 cm<sup>2</sup> flasks. Incubate the cells at 37 °C/5% CO<sub>2</sub> for 3 days.

**NOTE:** This step is best performed on a Friday as it allows the cells to expand during the weekend.

6. Harvest and count the HEK293T cells.
7. For the preparation of one of the viruses, prepare ten 150 mm dishes each containing  $11 \times 10^6$  HEK293T in 20 mL of DMEM supplemented with 10% FBS, 1% P/S, and 25 mM HEPES.
8. Place the dishes in the incubator at 37 °C/5% CO<sub>2</sub> for 24 h. Carefully discard the old medium and replace with 20 mL of antibiotic-free DMEM supplemented with 10% FBS, 25 mM HEPES, and 1 mM sodium butyrate. Before proceeding, check that the confluency of the cells is not above 80%.
9. Prepare two 15 mL tubes that will contain the transfection mixes. To Tube 1, add 5 mL of reduced serum medium, 60 µg of rAAV6 mutated HDR plasmid, and 220 µg of pDGM6 (helper plasmid). To Tube 2, add 5 mL of reduced serum medium and 1120 µL of 1 mg/mL polyethylenimine solution (PEI, transfection reagent).
10. Add the contents of Tube 2 to Tube 1 and vortex for 30 s. Incubate for 15 min at RT to ensure proper encapsulation of the DNA into the PEI micelles.

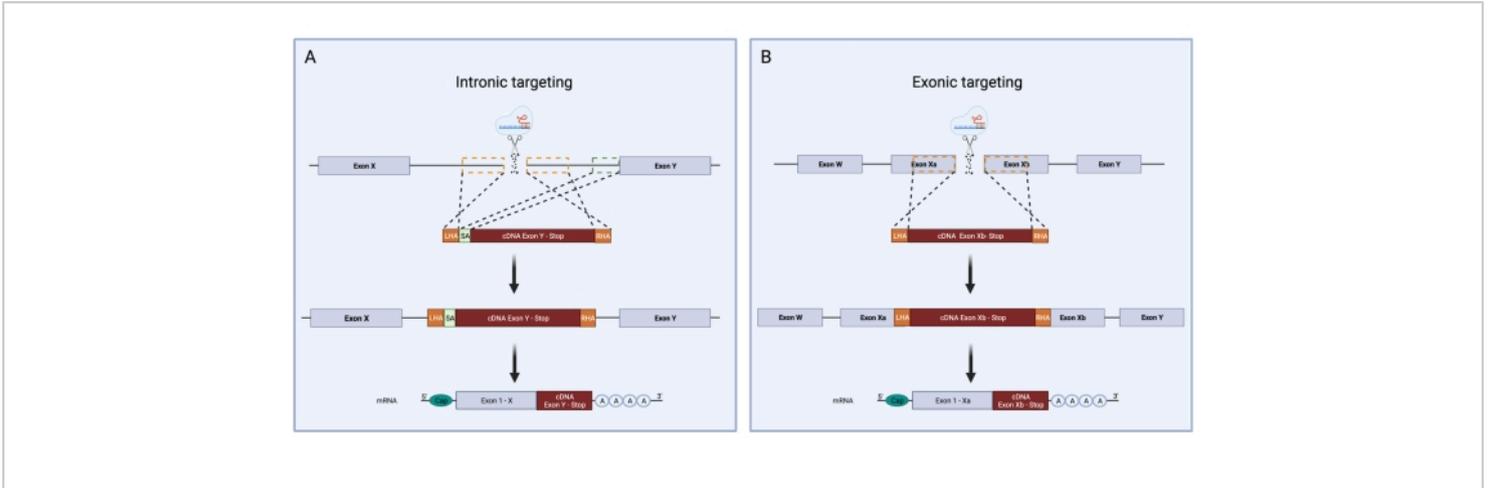
**NOTE:** Do not keep the solution for more than 20 min.

11. Carefully add dropwise 1.1 mL of the solution to each dish and distribute by swirling gently. Place the dishes in the incubator at 37 °C/5% CO<sub>2</sub> for 48 h.
  12. After 48 h, add 250 µL of 0.5 M EDTA to each dish and place the dishes in the incubator for 10 min. Harvest the cells by washing them off the dish and transfer to a 500 mL centrifuge tube or multiple 50 mL tubes.
  13. Centrifuge at 2,000 x g for 10 min at 4 °C. Discard the supernatant. To ensure the complete removal of the supernatant, centrifuge again at 2,000 x g for 1 min at 4 °C.
  14. Discard any remaining supernatant. Any residual supernatant could impair the virus purification.
  15. Loosen the pellet by vortexing and extract the virus by using an AAV purification kit (see **Table of Materials**) according to the manufacturer's instructions. Alternatively, perform extraction by iodixanol gradient ultracentrifugation<sup>30,31</sup>.
  16. Aliquot the purified virus and store at -80 °C. Titrate the AAV functionally or quantify the AAV titer by digital droplet PCR (ddPCR)<sup>32</sup> in order to determine the optimal transduction conditions.
  17. Repeat this process for the preparation of the rAAV6 WT HDR plasmid.
3. AAV titration by ddPCR
    1. Before starting, set the heating blocks at 65 °C and 98 °C.
    2. Extract viral DNA by adding 15 µL of DNA extraction solution to 5 µL of the virus. Vortex for 15 s.

3. Incubate for 6 min at 65 °C. Vortex for 15 s. Incubate for 2 min at 98 °C.
4. Use the extracted DNA (which is a 1:4 dilution of the viral DNA) to prepare serial dilutions (1:400, 1:40,000, 1:160,000, 1:640,000) with nuclease free (nf) H<sub>2</sub>O for the ddPCR.
 

**NOTE:** The dilutions can be stored at -20 °C if the ddPCR will not be immediately performed.
5. Select three of the dilutions for the ddPCR. The choice of the dilution to use depends on the concentration of the virus. Preferably, use three different dilutions (e.g., 1:40,000, 1:160,000, and 1:640,000) to find the best dilution that is within the detection limit of the machine.
6. Keep the reagents on ice while working. Prepare a master mix (all the samples will be measured in duplicates). For each reaction, prepare the following: 12.5 µL of ddPCR Supermix for Probes, no dUTP, 1.25 µL of PrimeTime Std qPCR Assay, AAV-ITR (see **Table of Materials**), and 6.25 µL nf-H<sub>2</sub>O.
7. Mix 5 µL of DNA with 20 µL of the master mix. Perform this for the 1:40,000, 1:160,000, and 1:640,000 dilutions.
8. Place a cartridge in the cartridge holder. Add 70 µL of droplet generation oil to the probes in the row denominated as Oil. Be careful not to generate any bubbles.
9. Add 20 µL of the sample volume in the row denominated as Sample. Be careful not to generate any bubbles.
10. Seal the cartridge with the gasket and place it in the droplet generator. Generate the droplets by pressing the start button on the machine, remove the gasket, and carefully transfer, with a multi-channel pipette, 40 µL of the generated droplets into a 96-well plate. Work slowly to avoid destruction of the droplets and air bubbles.
11. Seal the 96-well plate with pierce foil (the red stripe facing up) using the PCR plate sealer. Place the plate in a thermocycler.
12. Set the volume to 40 µL, the lid temperature to 105 °C, and the ramp rates to 2 °C/s. Run the PCR program as described in **Table 2**.
13. Turn on the droplet reader 30 min before use. Open the software on the desktop.
14. Enter the plate layout and select **ABS** in the experiment tab and **ddPCR Supermix** in the Supermix tab. Then, first press **PRIME**, followed by clicking **FLUSH SYSTEM** on the software to initiate the system.
15. Enter the plate into the reader. Start the run and, when finished, export the data as a CSV file for subsequent analysis as a spreadsheet.
 

**NOTE:** The numbers generated by the software are genome copies (GC) per µL. These need to be multiplied by the dilution factors: GC/µL x 5 (dilution of the DNA in the master mix) x initial dilution (i.e., 40,000 or 160,000).



**Figure 3: Schematic overview of intronic and exonic targeting during CRISPR/Cas9 HDR knock-in.** Schematic comparison between intronic and exonic targeting strategies for CRISPR/Cas9 HDR knock-in. **(A)** During intronic targeting, a double-stranded break is introduced in an intron of the DNA. The HDR template is comprised of an LHA, cDNA sequence, and RHA. The intronic targeting requires, additionally, the presence of a splice acceptor containing the 3' splice site, branch point, and the polypyrimidine tract. This allows for correct splicing. The green dotted box represents the region corresponding to the SA sequence. The size of the SA is 150 bp. **(B)** Exonic targeting relies on the production of a double-stranded break directly in the exon. The HDR template is comprised of an LHA, cDNA sequence, and RHA. The orange dotted boxes indicate the regions corresponding to the LHA and RHA. The ideal size of the HAs is 400 bp each. [Please click here to view a larger version of this figure.](#)

| <b>Promoters</b>                       |               |  |
|--|---------------|--|
| <b>Name</b>                            | <b>Length</b> | <b>Description</b>   |
| SFFV                                   | 492 bp        | Spleen Focus Forming Virus promoter. Strong mammalian promoter. Constitutively expressed   |
| UBC                                    | 400 bp        | Promoter derived from the human ubiquitin C gene. Constitutively expressed.  |
| CMV                                    | 508 bp        | Promoter derived from the Cytomegalovirus. It may contain an enhancer region. Constitutively expressed. Strong mammalian promoter. Can be silenced.  |
| EF-1a                                  | 1182 bp       | Human eukaryotic translation elongation factor 1 alpha promoter. Constitutively expressed. Strong mammalian promoter.  |
| EFS                                    | 200-300 bp    | EF-1 alpha intron-less short form  |
| CAG                                    | 584 bp        | Hybrid mammalian promoter containing the CMV early enhancer (C), the chicken beta actin promoter (A), and the splice acceptor for the rabbit beta-globin gene (G). Constitutively expressed. |
| <b>Polyadenylation (PolyA) signals</b> |               |  |
| <b>Abbreviation</b>                    | <b>Length</b> | <b>Description</b>   |
| SV40 PolyA                             | 82-122 bp     | Simian virus 40 polyadenylation signal   |
| bGH PolyA                              | 224 bp        | Bovine growth hormone polyadenylation signal   |
| rbGlob PolyA                           | 56 bp         | Rabbit beta globin polyadenylation signal  |

**Table 1: Promoters and polyadenylation signals.**

| Step                | Temperature | Time       | Cycles |
|---------------------|-------------|------------|--------|
| Enzyme activation   | 95°C        | 10 minutes | 1x     |
| Denaturation        | 94°C        | 30 seconds | 42x    |
| Annealing           | 60°C        | 1 minute   |        |
| Extension           | 72°C        | 30 seconds |        |
| Enzyme deactivation | 98°C        | 10 minutes | 1x     |
| Hold                | 4°C         | ∞          |        |

**Table 2: Digital droplet PCR program.**

### 3. Editing of HSPCs

#### 1. Transfection and transduction of HSPCs

1. Thaw CD34<sup>+</sup> HSPCs and transfer the cells into 10 mL of pre-warmed RPMI supplemented with 1% P/S. Centrifuge at 350 x *g* at RT for 10 min.
2. Suspend the cells in HSPC retention medium to a concentration of 2.5 x 10<sup>5</sup> cells/mL and incubate at 37 °C/5% CO<sub>2</sub> for 48 h - 72 h.
3. Harvest the cells in a 15 mL tube. Count the cells and check the cell viability by trypan blue exclusion. Between 2 x 10<sup>5</sup> and 5 x 10<sup>6</sup> cells can be nucleofected in a single cuvette. Prepare the appropriate number of cuvettes depending on your calculated cell number.
4. Prepare the RNP complex. In a 1.5 mL tube, add 15 µg of Cas9 and 8 µg of sgRNA (molar ratio 1:2.5) and incubate at 25 °C for 10 min in a heating block.
5. While the RNP complex is incubating, centrifuge the cells at 350 x *g* for 5 min and discard the supernatant.

6. Suspend the cells in 100 µL of nucleofection solution. Mix the cells with the RNP complex and transfer into the cuvette.
7. Gently tap the cuvette to remove any residual air bubbles that might have formed during the transfer.
8. Insert the cuvette into the holder of the transfection system and electroporate the cells as previously described in the sgRNA design section.
9. Immediately after electroporation, add 400 µL of pre-warmed HSPC retention medium without P/S and transfer the cells with a fine transfer pipette into a tissue culture plate containing 500 µL of pre-warmed HSPC retention medium without P/S. Depending on the cell number, use an appropriate culture plate (24-, 12-, or 6-well plate) in order to reach a density between 0.25 x 10<sup>6</sup> and 1 x 10<sup>6</sup> cells/mL. Transfer the plate to the incubator at 37 °C/5% CO<sub>2</sub>.
10. Thaw the vials containing the frozen rAAVs on ice. Transduce the cells by pipetting the optimal amount of each rAAV6 to the cell suspension. Perform the transduction within 20-30 min after the electroporation to get high transduction efficiencies.

**NOTE:** The optimal concentration of each virus (one virus for the WT HDR template and another separate virus for the mutated HDR template) should be determined experimentally. Usually, 5,000-10,000 GC/cell (e.g.,  $5 \times 10^9$  GC for  $1 \times 10^6$  cells) result in high transduction.

11. Gently mix the cell suspension with the pipette. Incubate the transduced cells for 6-8 h at 37 °C/5% CO<sub>2</sub>. Following 6-8 h, collect the cells in a tube and centrifuge at 350 x g at RT for 5 min.
12. Discard the old medium and replace with fresh pre-warmed HSPC retention medium supplemented with P/S.
13. Suspend the cells at a concentration between  $2.5 \times 10^5$ - $5 \times 10^5$  cells/mL and transfer to a tissue-treated cell culture plate (24-, 12-, or 6-well plate). Incubate the cells for 48 h at 37 °C/5% CO<sub>2</sub> before proceeding with sorting.

2. Flow sorting of the engineered HSPCs bearing the heterozygous GOF mutation

1. Harvest the cells in a 15 mL tube and centrifuge at 350 x g at RT for 5 min. Remove the supernatant, suspend the cells in 1 mL of DPBS containing 0.1% BSA, and centrifuge again at 350 x g at RT for 5 min.
2. Discard the supernatant and suspend the cells in an appropriate volume of DPBS + 0.1% BSA depending on the cell number.

**NOTE:** It is recommended to suspend the cells at a minimum volume of 200 µL and not to exceed  $1 \times 10^7$  cells/mL in order to reduce the risk of clogging the sorter.

3. Transfer the cells to a sterile FACS tube equipped with a cap. Add 7-AAD or other viability dye (depending on their compatibility with the fluorescent proteins) to the cell suspension for live/dead cell exclusion.
4. Sort the live cells that are double positive for the fluorescent reporter proteins into a collection tube containing 200 µL of HSPC retention medium.
 

**NOTE:** Appropriate controls consisting of cells only transduced with AAVs should be added to ensure proper gating during the sorting procedure.
5. Centrifuge the sorted cells at 350 x g at RT for 5 min and suspend the cells in HSPC retention medium at a concentration of  $2.5 \times 10^5$  cells/mL for further expansion in culture or use the cells directly for functional assays.

#### 4. Confirmation of successful gene editing

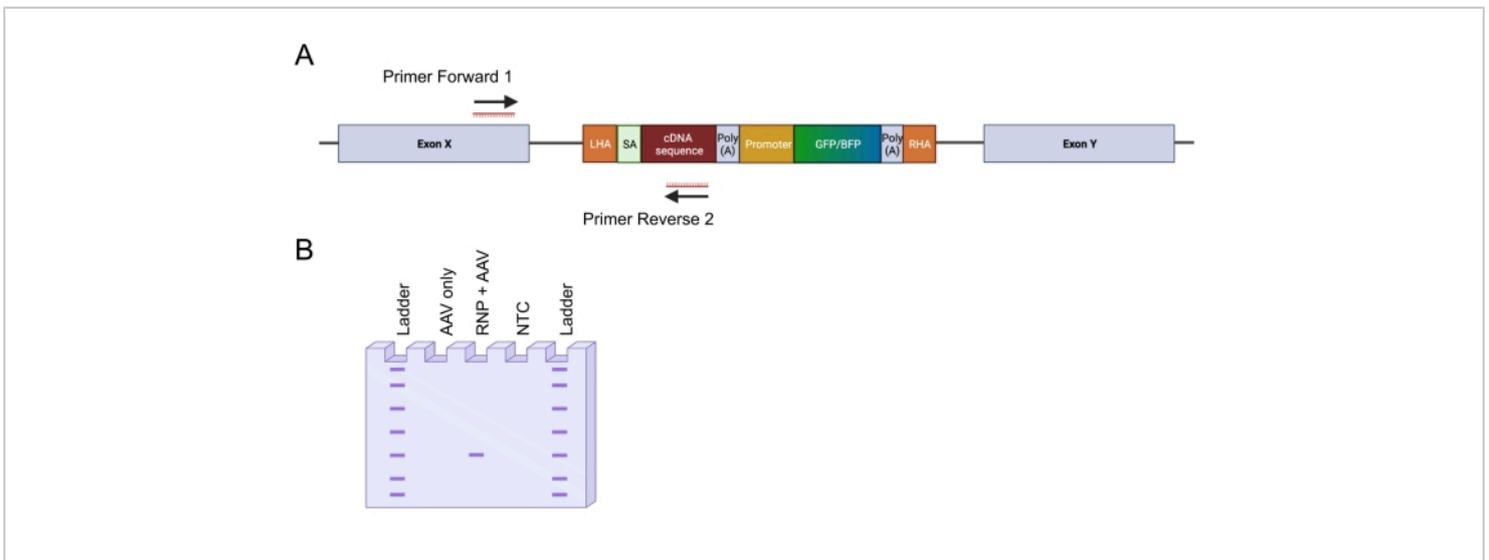
1. Genomic DNA extraction
  1. Before starting, set the heating blocks at 65 °C and 98 °C. Harvest  $2 \times 10^5$  genome-edited and sort-purified cells and centrifuge at 350 x g for 5 min.
  2. Extract the gDNA as previously described. Either use the DNA solution directly for PCR or store at -20 °C.
2. In-Out PCR
  1. Design two primers to amplify the 5' insertion site (**Figure 4A**): primer forward 1 targets the genomic locus outside of the left homology arm; primer reverse 2 targets the integrated sequence.

- For the in-out PCR, prepare the following PCR mix per reaction:
  - 10  $\mu\text{L}$  of PCR Master Mix (2x; see **Table of Materials**)
  - 1  $\mu\text{L}$  of primer forward 1 (10  $\mu\text{M}$  stock)
  - 1  $\mu\text{L}$  of primer reverse 2 (10  $\mu\text{M}$  stock)
  - 0.5-1.0  $\mu\text{L}$  of extracted DNA
  - Nuclease-free  $\text{H}_2\text{O}$  to a final volume of 20  $\mu\text{L}$

**NOTE:** For more than one reaction, it is advisable to prepare a master mix that contains one extra reaction to account for pipetting errors. It is important to include a non-template control and template DNA from a mock-treated sample.
- Run the PCR reactions with the appropriate thermal cycling program (see the manufacturer's instructions).

**NOTE:** It is advisable to first test out the primer pair for the optimal annealing temperature. This can be done by calculating the  $T_m$  and then by running the PCR with a thermocycler that can perform a gradient temperature.

- Run the PCR products with a DNA ladder on a 1.5% agarose gel at 100 V for 45-60 min (**Figure 4B**). Place the gel on a blue light or UV transilluminator. Excise the bands from the gel.
- Extract the DNA from the bands by using a DNA gel extraction kit. Send the extracted PCR samples with appropriate primers for Sanger sequencing to confirm the correct and seamless integration of the desired cDNA at the endogenous gene locus.



**Figure 4: Validation of genomic integration by in-out PCR.** (A) Schematic representation of the in-out PCR strategy. In the depicted strategy, two primers were designed. The primer forward 1 targets the genomic locus outside the LHA, and the primer reverse 2 targets the codon-optimized sequence. (B) Schematic representation of an agarose gel electrophoresis. Only successfully edited cells (RNP + AAV) will generate a PCR product during the in-out PCR, whereas the unedited

samples (AAV only) will not generate a PCR product. Abbreviation: NTC = non-template control. [Please click here to view a larger version of this figure.](#)

## Representative Results

By applying the above-described protocol, heterozygous type 1 *CALR* mutations were reproducibly introduced in cord-blood derived HSPCs. This mutation consists of a 52 bp deletion in exon 9 (the last exon of *CALR*), which results in a +1 frameshift, leading to the translation of a novel positively charged C-terminal domain<sup>26,33</sup>. To introduce the *CALR* mutation at the endogenous gene locus, an intronic targeting strategy upstream of exon 9 was adopted, since this would circumvent any unwanted changes in the coding sequence in cases where the Cas9-induced DSB was not repaired *via* the HDR mechanism. In this specific case, an sgRNA for intron 7 was designed due to the availability of high on-target and low off-target sequences in combination with favorable homology arms (lack of sequence repeats; **Figure 5A**).

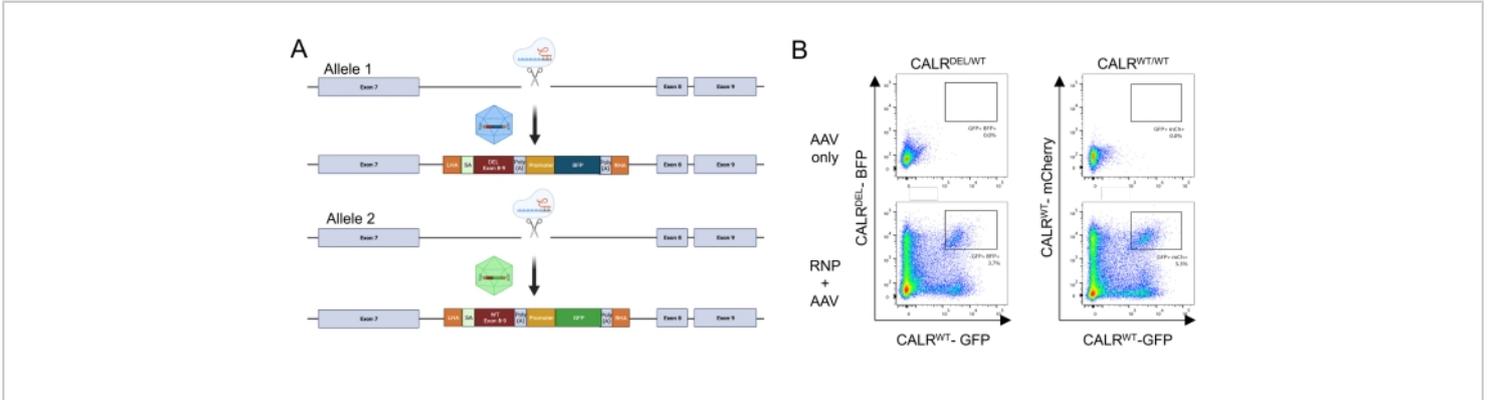
Two donor templates were then designed and packaged in AAV6 vectors. In order to enable correct splicing from the endogenous exons to the integrated cDNA, the donor templates contain (i) a SA sequence including the 3' splice-site, branch point, and the polypyrimidine tract, (ii) the codon-optimized cDNA sequence of exons 8-9, either containing the WT (*CALR*<sup>WT</sup>) or the mutated sequence (*CALR*<sup>DEL</sup>) including a stop codon, (iii) a simian virus 40 (SV40) polyA signal, (iv) a sequence encoding for a fluorescent protein under the control of the separate internal promoter, the spleen focus-forming virus (SFFV) promoter, followed by (v) a bovine growth hormone (bGH) polyA signal. The donor template containing the *CALR*<sup>WT</sup> cDNA was designed to contain a GFP cassette, whereas the donor template containing the *CALR*<sup>DEL</sup> cDNA sequence was designed to contain a BFP

cassette. The whole construct was flanked by a left and right HA (**Figure 5A**).

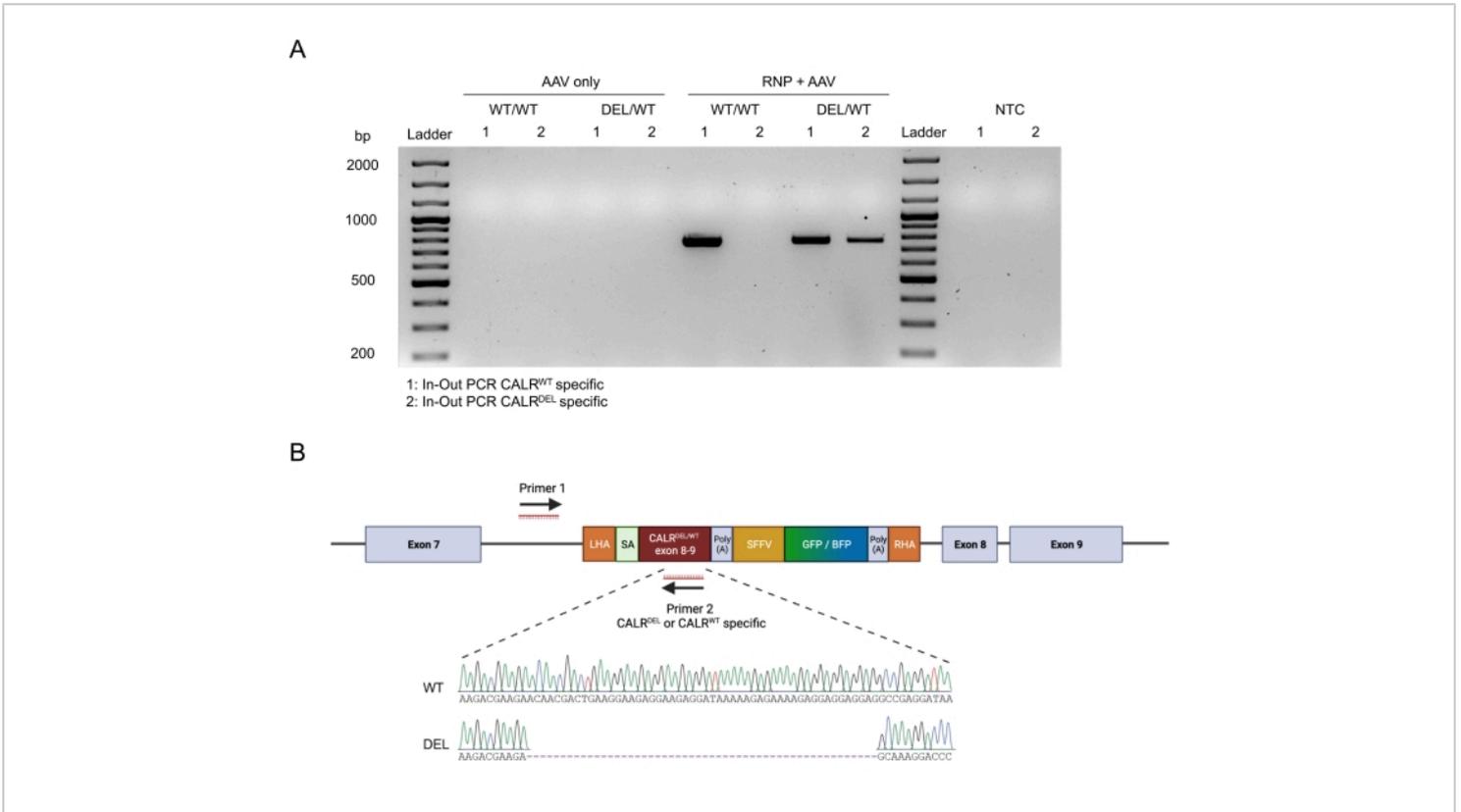
Two days following transfection with the RNP complex and transduction with the rAAV6 viruses, the cells were analyzed by flow cytometry. Four main populations could be detected: (i) cells expressing neither the GFP nor the BFP, representing cells with no HDR-based genome editing, (ii) cells positive only for GFP, representing those that had integrated only the WT construct, (iii) cells positive only for BFP, representing those that had integrated only the mutated construct, and (iv) GFP and BFP double-positive cells, representing the cells that had integrated both the WT and mutated sequences (**Figure 5B**). In order to obtain pure populations of HSPCs bearing the heterozygous type 1 *CALR* mutation, the double-positive cells were sorted by flow cytometry. HSPCs in which two WT sequences were knocked-in were used as control cells (GFP+ mCherry+; **Figure 5B**). A valid alternative to use as control cells would be HSPCs with a biallelic integration of the fluorescent proteins in a safe harbor locus (i.e., AAVS1; not shown). Cell counting by trypan blue exclusion performed on the sorted HSPCs indicated that more than 90% of the cells were viable.

Seamless on-target integration of the constructs was confirmed by applying the in-out PCR strategy (**Figure 6A**). In this specific case, we performed two separate in-out PCRs, one for the knocked-in *CALR*<sup>WT</sup> sequence (lane 1 of the gel electrophoresis in **Figure 6A**) and one for the knocked-in *CALR*<sup>DEL</sup> sequence (lane 2 of the gel electrophoresis of **Figure 6A**). Sanger sequencing performed on the DNA extracted from the gel bands confirmed the correct insertion of

the WT and mutated sequences in the  $CALR^{DEL/WT}$  HSPCs (Figure 6B).



**Figure 5: Generation of HSPCs bearing the heterozygous CALR mutation.** (A) Representative scheme depicting the editing strategy for the insertion of the heterozygous *CALR* mutation. The RNP complex targets the intron between exon 7 and exon 8 of the *CALR* gene. Two AAVs, one containing the mutated exons 8-9 and a BFP and the other containing the WT exons 8-9 and a GFP, will serve as donor repair templates and will promote the integration of the mutated sequence in one allele and the integration of the WT sequence in the remaining allele. (B) Representative flow cytometry plots depicting the expression of GFP and BFP or GFP and mCherry 48 h following the transfection and transduction of HSPCs. [Please click here to view a larger version of this figure.](#)



**Figure 6: Validation of successful heterozygous CALR mutation in HSPCs.** (A) Gel electrophoresis from the products of the in-out PCR performed on genomic DNA extracted from AAV controls, CALR<sup>WT/WT</sup>, and CALR<sup>DEL/WT</sup>. A 100 bp DNA ladder was used. Abbreviation: NTC = non-template control. (B) Sanger sequencing results obtained from the in-out PCRs performed on CALR<sup>DEL/WT</sup> confirming successful integration of the WT and mutated sequences. [Please click here to view a larger version of this figure.](#)

## Discussion

The efficient and precise genetic manipulation of human primary HSPCs represents a great opportunity to explore and understand the processes influencing normal hematopoiesis, and most importantly, the leukemic transformation of hematopoietic cells.

In this protocol, an efficient strategy to engineer human HSPCs to express recurrent heterozygous GOF mutations was described. This procedure took advantage of CRISPR/Cas9 technology and rAAV6 vectors as donors for DNA

templates to precisely insert WT and mutant DNA sequences into their endogenous gene loci. Coupling the engineered cDNAs (WT and mutant) with separated fluorescent reporter proteins allows for the enrichment and tracking of cells with a definitive heterozygous state.

This strategy presents several advantages in comparison to the frequently used lentiviral (LV)-based methods. One main advantage is that the CRISPR/Cas9-based system allows for precise editing in the endogenous loci, resulting in the preservation of the endogenous promoters and regulatory elements. This leads to homogeneity in the expression of

the edited gene in the cells, a goal hardly achievable when an LV-based method is used. Gene transfer with LV vectors leads to semi-random integration of the gene with preference for transcriptionally active sites<sup>34</sup>. This can translate into overexpression of the transferred gene and heterogeneity between the edited cells, eventually resulting in difficulties to investigate and analyze the role of mutations and gene interactions. A second advantage is that the described system, being a site-specific editing system, eliminates the risks of insertional mutagenesis<sup>35</sup>.

The dual fluorescent reporter strategy allows for the precise enrichment and tracking of cells that were successfully edited on both alleles, with one allele integrating the WT cDNA and the other allele integrating the mutated cDNA sequences. Cells only expressing a single reporter represent either only monoallelic integration or biallelic integration of HDR templates with the same fluorescent reporter. Both scenarios can only be precisely distinguished if single cell-derived clones are produced and individually analyzed. However, HSPCs have only limited proliferative capacity *in vitro*, and when kept in culture for extended periods of time, HSPCs start differentiating into more mature progeny and lose their self-renewal and engraftment capacity. This makes selecting and expanding single-cell clones harboring the desired heterozygous mutation unfeasible. The application of the dual fluorescent protein strategy and enrichment by flow cytometry for cells bearing the heterozygous mutation allows for bypassing the problems induced by extended *in vitro* culture.

In this specific example, it was successfully demonstrated that HSPCs could be efficiently engineered and sorted in order to obtain pure populations of HSPCs carrying the heterozygous *CALR*<sup>DEL/WT</sup> mutation.

However, this system is not limited to engineering heterozygous frameshift mutations but can also easily be adopted to create other mutation types, including missense and nonsense mutations. By applying different combinations of AAVs containing WT or mutated sequences with different fluorescent reporter proteins, this system can also be utilized for the introduction of homozygous mutations (simultaneous transduction with two rAAVs both carrying mutant cDNA but different fluorescent reporters) or even the correction of mutations (simultaneous transduction with two AAVs both carrying WT cDNA but different fluorescent reporters). Additionally, it is important to mention that this strategy is not limited to the introduction of oncogenic GOF mutations. In fact, the described protocol can be utilized for multiple alternative strategies including gene knock-out, gene replacement<sup>36,37</sup>, targeted knock-in of transgenes (i.e., chimeric antigen receptors)<sup>38</sup>, and even for the correction of disease-causing mutations<sup>11,39</sup>.

The strategy of combining CRISPR/Cas9 and AAV6 with multiple fluorescent reporters has also been shown to be applicable in many other cell types including T-cells, plasmacytoid dendritic cells, induced pluripotent stem cells, neuronal stem cells, and airway stem cells<sup>24,38,40,41,42,43,44</sup>. This strategy can be implemented for the production of superior chimeric antigen receptor (CAR) T cells. For example, it was recently published that CRISPR/Cas9-mediated knock-out of the *TGFBR2* gene in CAR T cells greatly increases their function in the suppressive TGF- $\beta$  rich tumor microenvironment<sup>45</sup>. Such an approach could provide a one-step protocol to both engineer the T cells to express the CAR and to knock out the *TGFBR2* gene by site specifically inserting the CAR into both alleles of the *TGFBR2* gene. Moreover, this approach could also be useful to generate

universal CAR T cells by integrating the CAR in the T cell receptor alpha constant (TRAC) gene<sup>46,47</sup>.

To increase the reproducibility and to guarantee efficient editing of the cells, some important considerations need to be taken care of. The main critical points for ensuring successful editing of the cells reside in (i) the selection of the sgRNA, (ii) the design of the HDR template, and (iii) the rAAV6 production.

The selection of a good-performing sgRNA is crucial as it will determine the maximum number of alleles in which the HDR template can be integrated. Due to numerous software that are now available, the search for candidate sgRNAs has been simplified. By selecting the region of interest, the software can propose a series of sgRNAs with an on-target score and an off-target score that indicate the chances for editing at the desired locus and unwanted loci, respectively. These scores are calculated based on previously published scoring models<sup>48,49</sup>. Although this is a good starting point for selecting a good-performing sgRNA, the performance of the sgRNA needs to be confirmed as its predicted performance *in silico* does not always correspond to an efficient sgRNA *in vitro*. Therefore, it is highly recommended to design and test out at least three sgRNAs to increase the chances of finding the best sgRNA. Once a true good-performing sgRNA has been identified, then it is suggested to proceed with the design of the HDR template.

Precautions should be taken into consideration when designing the HDR template. The left and right homology arms (LHA and RHA, respectively) should each span 400 bp upstream and downstream of the sgRNA cut site, respectively, as shorter HAs could result in reduced HDR frequencies. The size of the cDNA that can be introduced *via* HDR is dependent on the packaging capabilities of AAVs,

which is roughly 4.7 kb. Due to the numerous elements mandatory within the HDR template (LHA, RHA, SA, PolyA, promoter, and fluorescent reporter sequence), the remaining space for the mutated or WT cDNA is limited. This is unproblematic if the desired mutation is located near the 3' end of a gene or in genes with an overall short CDS. However, in cases where the mutation is located near the transcriptional start site (TSS) of the genes with a long CDS (exceeding the remaining packing space of the AAV), this described approach may not be feasible. To circumvent this problem, a strategy that relies on splitting the HDR template into two AAVs has been recently developed by Bak and colleagues. This strategy relies on two separate HDR-mediated integrations to obtain the final seamless integration of a large gene<sup>50</sup>.

The quality of the virus and its titer are additional factors that can make or break the successful genome engineering of the cells. For an optimal yield, it is important to not let the HEK293T reach full confluency while being maintained in culture. Ideally, the HEK293T cells should be split when 70%-80% confluency is reached. Additionally, the HEK293T should not be cultured for long periods of time as this can decrease their ability to produce virus. New HEK293T cells need to be thawed after 20 passages. Obtaining high virus titers is important for increasing the efficiency and reproducibility of the experiments. Low viral titers will translate to large volumes of rAAV solution required for the transduction of the HSPCs. As a general rule, the rAAV solution added to the nucleofected cells should not exceed 20% of the total volume of the HSPC retention medium. Higher volumes of AAV solution may lead to increased cell death, lower proliferation, and impaired transduction efficiencies. In the

case of low virus titers, it is, therefore, recommended to further concentrate the virus.

In summary, this protocol offers a reproducible approach to manipulate human HSPCs precisely and efficiently through the simultaneous use of CRISPR/Cas9 and rAAV6 donor templates with additional dual fluorescent reporters. This approach has proven to be a great tool in studying normal hematopoietic stem cell biology and the contributions that mutations make to leukemogenesis.

## Disclosures

The authors have nothing to disclose.

## Acknowledgments

This work is supported by grants from the Austrian Science Fund (FWF; number P32783 and I5021) to A.R. Additional funding to A.R. is also provided by the Austrian Society of Internal Medicine (Joseph Skoda Fellowship), the Austrian Society of Hematology and Oncology (OeGHO; Clinical Research Grant), and MEFOGraz. T.K. is a Special Fellow of the Leukemia & Lymphoma Society.

## References

- Hsu, P. D., Lander, E. S., Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. **157** (6), 1262-1278 (2014).
- Gillmore, J. D. et al. CRISPR-Cas9 in vivo gene editing for transthyretin amyloidosis. *New England Journal of Medicine*. **385** (6), 493-502 (2021).
- Frangoul, H. et al. CRISPR-Cas9 gene editing for sickle cell disease and  $\beta$ -thalassemia. *New England Journal of Medicine*. **384** (3), 252-260 (2021).
- Stadtmauer, E. A. et al. CRISPR-engineered T cells in patients with refractory cancer. *Science*. **367** (6481), eaba7365 (2020).
- Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. **337** (6096), 816-821 (2012).
- Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. **339** (6121), 819-823 (2013).
- Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*. **8** (11), 2281-2308 (2013).
- Heyer, W. D., Ehmsen, K. T., Liu, J. Regulation of homologous recombination in eukaryotes. *Annual Review of Genetics*. **44**, 113-139 (2010).
- Veldwijk, M. R. et al. Pseudotyped recombinant adeno-associated viral vectors mediate efficient gene transfer into primary human CD34+ peripheral blood progenitor cells. *Cytherapy*. **12** (1), 107-112 (2010).
- Song, L. et al. High-efficiency transduction of primary human hematopoietic stem cells and erythroid lineage-restricted expression by optimized AAV6 serotype vectors in vitro and in a murine xenograft model in vivo. *PLoS One*. **8** (3), e58757 (2013).
- Dever, D. P. et al. CRISPR/Cas9  $\beta$ -globin gene targeting in human haematopoietic stem cells. *Nature*. **539** (7629), 384-389 (2016).
- Drost, J. et al. Sequential cancer mutations in cultured human intestinal stem cells. *Nature*. **521** (7550), 43-47 (2015).
- Jan, M. et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid

- leukemia. *Science Translational Medicine*. **4** (149), 149ra118 (2012).
14. Corces-Zimmerman, M. R., Hong, W. J., Weissman, I. L., Medeiros, B. C., Majeti, R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proceedings of the National Academy of Sciences of the United States of America*. **111** (7), 2548-2553 (2014).
  15. Jaiswal, S. et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *New England Journal of Medicine*. **377** (2), 111-121 (2017).
  16. Genovese, G. et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *New England Journal of Medicine*. **371** (26), 2477-2487 (2014).
  17. Papaemmanuil, E. et al. Genomic classification and prognosis in acute myeloid leukemia. *New England Journal of Medicine*. **374** (23), 2209-2221 (2016).
  18. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *New England Journal of Medicine*. **368** (22), 2059-2074 (2013).
  19. Ball, M., List, A. F., Padron, E. When clinical heterogeneity exceeds genetic heterogeneity: thinking outside the genomic box in chronic myelomonocytic leukemia. *Blood*. **128** (20), 2381-2387 (2016).
  20. Varmus, H. E. The molecular genetics of cellular oncogenes. *Annual Review of Genetics*. **18**, 553-612 (2003).
  21. Cox, D. B. T., Platt, R. J., Zhang, F. Therapeutic genome editing: Prospects and challenges. *Nature Medicine*. **21** (2), 121-131 (2015).
  22. Tothova, Z. et al. Multiplex CRISPR/Cas9-based genome editing in human hematopoietic stem cells models clonal hematopoiesis and myeloid neoplasia. *Cell Stem Cell*. **21** (4), 547-555.e8 (2017).
  23. Mandal, P. K. et al. Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell*. **15** (5), 643-652 (2014).
  24. Bak, R. O., Dever, D. P., Porteus, M. H. CRISPR/Cas9 genome editing in human hematopoietic stem cells. *Nature Protocols*. **13** (2), 358-376 (2018).
  25. Foßelteder, J. et al. Human gene-engineered calreticulin mutant stem cells recapitulate MPN hallmarks and identify targetable vulnerabilities. *Leukemia*. (2023).
  26. Nangalia, J. et al. Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2. *New England Journal of Medicine*. **369** (25), 2391-2405 (2013).
  27. Merlinsky, T. R., Levine, R. L., Pronier, E. Unfolding the role of calreticulin in myeloproliferative neoplasm pathogenesis. *Clinical Cancer Research*. **25** (10), 2956-2962 (2019).
  28. Belčič Mikič, T., Pajič, T., Zver, S., Sever, M. The contemporary approach to CALR-positive myeloproliferative neoplasms. *International Journal of Molecular Sciences*. **22** (7), 3371 (2021).
  29. How, J., Hobbs, G. S., Mullally, A. Mutant calreticulin in myeloproliferative neoplasms. *Blood*. **134** (25), 2242-2248 (2019).
  30. Grieger, J. C., Choi, V. W., Samulski, R. J. Production and characterization of adeno-associated viral vectors. *Nature Protocols*. **1** (3), 1412-1428 (2006).

31. Zolotukhin, S. et al. Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Therapy*. **6** (6), 973-985 (1999).
32. Aurnhammer, C. et al. Universal real-time PCR for the detection and quantification of adeno-associated virus serotype 2-derived inverted terminal repeat sequences. *Human Gene Therapy Methods*. **23** (1), 18-28 (2011).
33. Klampfl, T. et al. Somatic mutations of calreticulin in myeloproliferative neoplasms. *New England Journal of Medicine*. **369** (25), 2379-2390 (2013).
34. Bulcha, J. T., Wang, Y., Ma, H., Tai, P. W. L., Gao, G. Viral vector platforms within the gene therapy landscape. *Signal Transduction and Targeted Therapy*. **6** (1), 1-24 (2021).
35. Montini, E. et al. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *The Journal of Clinical Investigation*. **119** (4), 964-975 (2009).
36. Vaidyanathan, S. et al. Targeted replacement of full-length CFTR in human airway stem cells by CRISPR-Cas9 for pan-mutation correction in the endogenous locus. *Molecular Therapy*. **30** (1), 223-237 (2022).
37. Cromer, M. K. et al. Gene replacement of  $\alpha$ -globin with  $\beta$ -globin restores hemoglobin balance in  $\beta$ -thalassemia-derived hematopoietic stem and progenitor cells. *Nature Medicine*. **27** (4), 677-687 (2021).
38. Wiebking, V. et al. Genome editing of donor-derived T-cells to generate allogenic chimeric antigen receptor-modified T cells: Optimizing  $\alpha\beta$  T cell-depleted haploidentical hematopoietic stem cell transplantation. *Haematologica*. **106** (3), 847-858 (2021).
39. Wilkinson, A. C. et al. Cas9-AAV6 gene correction of beta-globin in autologous HSCs improves sickle cell disease erythropoiesis in mice. *Nature Communications*. **12** (1), 1-9 (2021).
40. Dever, D. P. et al. CRISPR/Cas9 genome engineering in engraftable human brain-derived neural stem cells. *iScience*. **15**, 524-535 (2019).
41. Laustsen, A. et al. Interferon priming is essential for human CD34+ cell-derived plasmacytoid dendritic cell maturation and function. *Nature Communications*. **9** (1), 1-14 (2018).
42. Bak, R. O. et al. Multiplexed genetic engineering of human hematopoietic stem and progenitor cells using CRISPR/Cas9 and AAV6. *eLife*. **6**, e27873 (2017).
43. Nakauchi, Y. et al. The cell type-specific 5hmC landscape and dynamics of healthy human hematopoiesis and TET2-mutant preleukemia. *Blood Cancer Discovery*. **3** (4), 346-367 (2022).
44. Vaidyanathan, S. et al. High-efficiency, selection-free gene repair in airway stem cells from cystic fibrosis patients rescues CFTR function in differentiated epithelia. *Cell Stem Cell*. **26** (2), 161-171.e4 (2020).
45. Tang, N. et al. TGF- $\beta$  inhibition via CRISPR promotes the long-term efficacy of CAR T cells against solid tumors. *JCI Insight*. **5** (4), e133977 (2020).
46. Georgiadis, C. et al. Long terminal repeat CRISPR-CAR-coupled "universal" T cells mediate potent anti-leukemic effects. *Molecular Therapy*. **26** (5), 1215-1227 (2018).
47. Ren, J. et al. Multiplex genome editing to generate universal CAR T cells resistant to PD1 inhibition. *Clinical Cancer Research*. **23** (9), 2255-2266 (2017).

48. Hsu, P. D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*. **31** (9), 827-832 (2013).
49. Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology*. **34** (2), 184-191 (2016).
50. Bak, R. O., Porteus, M. H. CRISPR-mediated integration of large gene cassettes using AAV donor vectors. *Cell Reports*. **20** (3), 750-756 (2017).