A Universal Approach to Target Various *HBB* Gene Mutations in Hematopoietic Stem/Progenitor Cells for Beta-Thalassemia Gene Therapy by CRISPR/Cas9 and the rAAV6 Vector

Li-Na He
The Third Affiliated Hospital of Guanhzhou Medical Uiversity
Yi Yang
Third Affiliated Hospital of Guangzhou Medical College
Yi Cheng
Guangzhou Medical University
Han Wu
Guangzhou Institutes of Biomedicine and Health
Shou-Heng Lin
Guangzhou Medical University
Bing Song
Guangzhou Medical University
Neng-Qing Liu
Guangzhou Medical University
Di-Yu Chen
Guangzhou Medical University
Dian Lu
Guangzhou Medical University
Ying-Hong Yang
Guangzhou Medical University
Juan Zeng
Guangzhou Medical University
Yong Fan
Guangzhou Medical University
Sun Xiaofang (xiaofangsun@gzhmu.edu.cn)
The Third Affiliated Hospital of Guangzhou Medical University

Research

Keywords: β-thalassemia, CRISPR/Cas9, hematopoietic stem and progenitor cells, rAAV6

Posted Date: September 9th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-68414/v1

Abstract

Background

Engineered nuclease-mediated gene targeting through homology-directed repair (HDR) in autologous hematopoietic stem and progenitor cells (HSPCs) has the potential to cure β-thalassemia (β-thal). Although previous studies have precisely corrected site-specific *HBB* mutations by HDR in vitro and in vivo, targeting the various *HBB* mutations in β-thal is still challenging. Here, we devised a universal strategy to achieve repaired most types of *HBB* mutations through the CRISPR/Cas9 and the rAAV6 donor.

Methods

Using cord blood-derived HSPCs from health donors, we tested the strategy to achieved highly efficient targeted integration by optimizing design and delivery parameters of a ribonucleoprotein (RNP) complex comprising Cas9 protein and modified single guide RNA, together with a rAAV6 donor. We assessed the edited HSPCs function in vitro by methylcellulose colonies assay, CFU assay, differentiation experiment and Wright-Giemsa staining. Edited HSPCs transplanted into NSI mice to assess the long-term reconstitution in vivo. Whole-genome sequencing was used to analysis the off-target mutagenesis of edited HSPCs.

Results

Edited HSPCs exhibited normal multilineage formation and erythroid differentiation abilities without off-target mutagenesis and retained the ability to engraft. Moreover, we used the strategy to efficiently correct the β-CD41/42 mutation of patient-derived HSPCs, erythrocytes differentiation from which expressed more *HBB* mRNA than uncorrected cells.

Conclusion

This strategy demonstrated a universal approach to correct most types of *HBB* gene mutations in β-thal.

Introduction

β-thalassemia (β-thal), one of the most common genetic diseases worldwide, is caused by over 200 different types of mutations in the β-globin (*HBB*) gene[1]. Normally, *HBB* pairs with α-globin (HBA) in a one-to-one ratio to form the tetrameric hemoglobin molecule, and with the insufficient production of *HBB*, unpaired α-globin chains precipitate, thereby causing toxic death to the developing erythrocyte or erythrocyte precursor and leading to the insufficient formation of mature red blood cells (RBCs)[2]. Ineffective erythropoiesis leads to anemia, and severe anemia can cause a high level of mortality or

shortened life expectancy if left untreated. β-thal affects millions of people worldwide, and approximately 3 of 1,000 new births worldwide are affected with a severe form of β-thal[1, 3].

The only curative treatment for β-thal is allogeneic hematopoietic stem cell transplantation (allo-HSCT). However, it is limited because of a lack of immunologically matched donors and graft-versus-host disease[4]. In recent decades, scientists have developed an alternative approach of gene therapy for treating β-thal and this approach relies on genome-inserting lentiviral vectors that carry the functional *HBB* gene, inserting them into permanently the genome of autologous hematopoietic Stem and Progenitors Cells (HSPCs), which will home into the patient's marrow after bone marrow (BM) transplantation, differentiate into erythrocytes and express a high level of the added *HBB* gene[5, 6]. Although many clinical trials have been implemented to determine the balance of efficacy and risks for gene therapy with lentiviral vectors for β-thal, the "semi-random" integration nature of the lentiviral vector is always a potential risk. [7–9]. Currently, studies have moved to achieve precise genome editing through homology-directed repair (HDR) of an *HBB* mutation. Unlike viral-vector-based gene transfer methods, it can preserve endogenous promoters and regulatory gene expression to mediate spatiotemporal gene expression [10, 11].

HDR genome editing is the precise modification of the nucleotide sequence of the genome, it requires engineered nucleases to create DNA double-strand breaks (DSBs) at a specific genomic site and a DNA donor template to repair the damaged site through a "copy and paste" mechanism[12]. The Cas9 nuclease, guided by a single guide RNA (sgRNA), can be programmed to cut a target locus within the genome with rapid iteration and optimization[13]. Recent studies have demonstrated efficient targeted integration of sickle cell disease (SCD) point mutations in exon 1 of the *HBB* gene in HSPCs by combining CRISPR/Cas9 with exogenous HR donors delivered via single-stranded oligonucleotides (ssODNs) [14, 15] or recombinant adeno-associated viral vectors of serotype 6 (rAAV6), which show very positive results in vitro, especially for the rAAV6 HDR donor, which can achieve an average of 29% HDR efficiency[16]. Moreover, the rAAV6 donor has the optimal packing capacity compared to the ssODNs.

However, a previous study focused on a specific mutation, such as the SCD locus in exon 1 of the *HBB* gene and correcting various *HBB* mutations is more beneficial for future clinical applications[17, 18]. It is highly desirable to develop a universal strategy to correct most types of *HBB* mutations by validated CRISPR guide RNA and one DNA donor template for HDR. Therefore, we devised a universal strategy to achieve the normal expression of *HBB* in various mutations. First, a validated sgRNA with a high indel frequency was used as a ribonucleoprotein (RNP) complex to create a DSB in intron 2 of the *HBB* gene. Then, the rAAV6 donor combined with the 1.7 kb left homologous arm and 507 bp right homologous arm is targeted insertion of 3 exons and intron 1 of the *HBB* gene in the DSB locus while simultaneously excluding intron 2. We also linked an EGFP reporter downstream to the *HBB* gene so that the expression of EGFP is indicative of successful insertion of the *HBB* gene into the genome. Using cord blood-derived HSPCs (CB HSPCs) from health donors to test the strategy, we indicate that it can achieve highly efficient targeted insertion and that edited CB HSPCs have normal function compared to noncorrected cells in vitro. Whole-genome sequencing analysis and off-target results indicated that corrected CB HSPCs

exhibited a minimal mutational load and no off-target mutagenesis. Moreover, the edited CB HSPCs retained the ability to engraft when transplanted into immunodeficient nonobese diabetic (NOD)-severe combined immunodeficiency (SCID) IL2rg−/−(gamma) mice (NSI mice); more importantly, the universal approach can correct the β-CD41/42 mutation and improve *HBB* mRNA expression. In addition, it provides an experimental system to screen the small-molecule compounds to improve HDR efficiency in HSPCs based on the coexpression of the EGFP reporter. These results may also have implications in the treatment of other monogenic diseases.

Materials And Methods

Cell culture

CD34+ HSPCs from cord blood and fetal liver were obtained from the Department of Obstetrics and Gynecology at The Third Affiliated Hospital of Guangzhou Medical University, which was approved by the ethics committee of the hospital. HSPCs were purified within 24 h of scheduled apheresis. Briefly, whole cord blood was mixed with PBS in a proportion of 1:1 (v/v), and then the mononuclear fraction was separated by density gradient separation using Ficoll. CD34+ HSPCs were extracted from the mononuclear fraction using a CD34 Microbeads Kit (Miltenyi Biotech, CD34 MicroBead Kit UltraPure, human) according to the manufacturers' protocol. Cells were stained for CD34 using APC anti-human CD34 (clone 563; BD) to test the purity. All CD34+ HSPCs were cultured in StemSpan SFEMII (StemCell Technologies) supplemented with SCF (100 ng ml^{-1}), TPO (100 ng ml^{-1}), Flt3 ligand (100 ng ml^{-1}), IL-6 (100 ng ml^{-1}), Stem Regenin1 (0.75 μM) and UM171 (35 nM). Cells were cultured at 37 °C and 5% CO_2.

AAV vector production

AAV vector plasmids were cloned in the ssAAV-MCS plasmid (PackGene Biotech) containing inverted terminal repeats (ITRs) from AAV serotype 2 (AAV2) using Gibson Assembly Mastermix (New England Biolabs). The *HBB* AAV6 donors contained arms homologous to the beta-globin locus of 1,708 bp on the left side and 507 bp on the right side (Figure 1), the donor also contained BGH polyA, EGFP, and SFFV promoters. AAV6 vectors were produced as follows: briefly, $1X10^7$ 293T cells were seeded per 15-cm dish before transfection; each 15-cm dish was transfected with 6 μg of ssAAV-MCS plasmid containing the donor, 7.5 μg of pAAVcap6 containing the AAV6 cap genes and AAV2 rep genes and 7.5 μg of adenovirus helper genes using polyethylenimine (PEI). After incubating for 72 h, cells were lysed by three freeze-thaw cycles and then incubated with TurboNuclease (Abnova) at 250 U/ml for 45 min. AAV6 particles were purified by iodixanol density gradient centrifugation at 237,000 g for 2 h at 18 °C. AAV6 vectors were extracted at the 60-40% iodixanol interface and then exchanged in PBS with 5% sorbitol using either a molecular weight cut off (MWCO) Slide-A Lyzer G2 dialysis cassette (Thermo Fisher Scientific) following the manufacturer's instructions. AAV6 vectors were titered using quantitative PCR to measure the number of vector genomes. The vectors were stored at -80 °C.

Design and in vitro transcription of sgRNAs

The sgRNAs targeting intron 2 of the *HBB* gene were designed online (https://crispr.cos.uni-heidelberg.de/), and sgRNA oligonucleotide sequences complementary to the *HBB* gene were annealed and cloned into the Bbsl site of PX458 (Addgene 48138). The T7 promoter was then added to the sgRNA template by PCR amplification of sgRNA expression plasmids using the corresponding primers. T7-sgRNA PCR products were purified and used as templates for the synthesis of sgRNAs using the MEGAshortscript T7 Transcription Kit (AM1354, Life Technologies, USA). All sgRNAs were purified by the RNeasy MinElute Cleanup Kit (74204, QIAGEN).

Electroporation and transduction of cells

The *HBB* sgRNAs were produced from in vitro transcription, and the modified sgRNA4, which had 2'-O-methyl-3'-phosphorothioate modifications at the three terminal nucleotides of the 5' and 3' ends, was purchased from Thermo Fisher Scientific. The sequence for sgRNA4 is as follows: 5'-GACGAATGATTGCATCAGTGTGG-3'. Cas9 protein was purchased from Thermo Fisher Scientific, and RNP was made by complexing the Cas9 protein with sgRNA at a molar ratio of 1:1 to 1:3 (Cas9 protein: sgRNA) at room temperature for 10-30 min before electroporation. CD34+ HSPCs were electroporated 2-3 days after thawing or isolation. CD34+ HSPCs were electroporated using a Neon Transfection System, and the electroporation parameters were 1,400 V-1,650 V, 10 ms, and 3 pulses. The following conditions for the 10 μl system were used: 2×10^5 cells, 1.5 μg cas9 protein (1 μg/μl) complexed with sgRNA at a 1:1-1:3 molar ratio. The following conditions for the 100 μl system were used: 1×10^6 cells, 15 μg cas9 protein (5 μg/μl) complexed with sgRNA at a 1:2.5 molar ratio. Following electroporation, AAV6 was transduced into cells immediately upon plating after electroporation at an MOI of 1×10^3-1×10^6. Then, all cells were cultured at 37 °C and 5% CO_2.

T7E1 and TIDE assays

The PCR amplicon spanning the Cas9-sgRNA cleavage site was diluted in Buffer 2 (NEB) and hybridized slowly in a thermal cycler based on the manufacturer's instructions. Hybridized fragments were then digested with 1 μl T7 endonuclease I (NEB) for 10 min at 37 °C. Then, polyacrylamide gel electrophoresis was used to separate digested fragments. Band intensities were analyzed using ImageJ software. The cleavage ratio was calculated by the ratio of the intensities of the cleaved bands to uncleaved bands. For accurate quantification of the editing efficiency, the PCR products spanning the Cas9-sgRNA cleavage site were sent for standard Sanger sequencing with both forward and reverse primers, and TIDE software (tracking of indels by decomposition) was used to quantify the indel frequencies. Primers used for amplifying PCR fragments for TIDE at the beta-globin locus are as follows: *HBB*-in2-DNA4F (forward primer) 5'-GAGTGAGCTGCACTGTGACAA-3' and *HBB*-in2-DNA4R (reverse primer) 5'-AGAATGGTGCAAAGAGGCATGATAC-3'.

Measuring the targeted integration of fluorescent AAV6 donors and methylcellulose CFU assay

Rates of targeted integration of GFP donors were measured by flow cytometry 4 days after electroporation. The GFP^{high} populations were sorted into 96-well plates containing MethoCult Optimum (Stem Cell Technologies) by FACS. After 14 days, colonies were counted under an inverted microscope and scored in a blinded fashion based on morphological features of colony forming units-erythroid (CFU-E), erythroid burst forming units (BFU-E), colony forming unit- granulocytes, monocytes (CFU-GM), and colony forming unit-multipotential cells (CFU-GEMM).

Genotyping of methylcellulose colonies

Colonies formed in methylcellulose were extracted from FACS sorting of single cells into 96-well plates. Briefly, PBS was added to the 96-well plates, and the colonies were mixed with PBS and transferred to a V-bottomed 96-well plate. Then, the cells were pelleted by centrifugation at 300 g for 5 min at room temperature, and the cells were resuspended in 250 μl of PBS after removing the supernatant. Then, the cells were pelleted by centrifugation at 300 g for 5 min again. Finally, cells were resuspended in 10 μl DNA Extraction Solution (a gift from GIBH) and transferred to PCR plates, which were incubated at 65 °C for 60 min followed by 95 °C for 10 min. In-out PCR was used to detect the integrated and nonintegrated alleles, and the integrated (one primer in the insert and one primer outside right homology arm) primers were as follows: *HBB*-3ARM-in-out-F1: 5'-TCCCCCTGAACCTGAAACATAAAAT-3', *HBB*-3ARM-in-out-R: 5'-TTTGGGGTGGGCCTATGACA-3'. The nonintegrated (primer in each homology arm) primers were as follows: *HBB*-3ARM-in-out-F2: 5'-TAAAAAGGGAATGTGGGAGGTCA-3', *HBB*-3ARM-in-out-R: 5'-TTTGGGGTGGGCCTATGACA-3'.

Whole-genome sequencing and Sanger sequencing

The whole-genome sequencing library was established with genomic DNA samples derived from HSPCs by next-generation sequencing (NGS) facility at the Biomarker Technologies Company. All libraries were sequenced with paired-end 150-bp reads in two Illumina Rapid Run flow cells using a HiSeq X instrument (Illumina). In contrast to the human genome Hg19, the data were analyzed through BWA.A genomic analysis toolkit (GATK, version 2.8.1). All the results of Sanger sequencing in the study were analyzed by IGE company.

Differentiation of CD34+ HSPCs into erythrocytes in vitro

HSPCs were cultured in three phases for differentiation 4 days after electroporation and transduction with AAV6. In the first phase, corresponding to days 0-7, cells were cultured at 10^5 cells/ml in SFEMII media supplemented with 100 U/ml penicillin/streptomycin, 2 mM L-glutamine, 100 ng/ml SCF (PeproTech), 10 ng/ml IL-3 (PeproTech), 0.5 U/ml erythropoietin (PeproTech), and 200 μg/ml transferrin (Sigma Aldrich). In the second phase, corresponding to days 7-11, erythropoietin was increased to 3 U/ml, and the cells were maintained at 10^5 cells/ml. In the third phase, corresponding to days 11-21, cells were cultured at 10^6 cells/ml, erythropoietin was maintained at 3 U/ml and transferrin was increased to 1 mg/ml. Erythrocyte differentiation was assessed by flow cytometry using the following antibodies: anti-CD71-PE and anti-CD235a-APC.

Assessment of the mRNA levels in differentiated erythrocytes

RNA was extracted from 100,000 differentiated cells between days 16-21 of erythroid differentiation with TRIzol reagent (Invitrogen). The relative quantity of mRNA was determined by real-time polymerase chain reaction (RT-PCR). GAPDH was chosen as the reference gene. The primer sequences were as follows: *HBB* primer (F): ATGGTGCATCTGACTCCTGA, *HBB* primer (R): TGGACAGATCCCCAAAGGAC; HBG primer (F): CATGATGGCAGAGGCAGAG, HBG primer (R): TGAATGTGGAAGATGCTGGA; GAPDH primer (F): TCAACGACCACTTTGTCAAGCTCA, GAPDH primer (R): GCTGGTGGTCCAGGGGTCTTACT. Quantitative mRNA expression was measured with ABI Prism 7500 Software v2.0.6 and calculated based on the comparative CT method. The expression level of each mRNA was normalized to that of GAPDH, and expression was expressed as an n-fold difference relative to the control.

Transplantation of CD34+ HSPCs into NSI mice

Six- to eight-week-old immunodeficient NSI mice were used for in vivo studies. To clear the mouse BM niche, mice were sublethally irradiated with 100 cGy 12–24 h before transplantation. Four days after electroporation/transduction, 1,000,000 cells were directly administered by tail-vein injection into the mice using an insulin syringe with a 27 gauge×0.5 inch (12.7 mm) needle for the negative control (NC), AAV only, and RNP+AAV groups. Mice were randomly assigned to each experimental group and evaluated in a blinded fashion. There were 3 mice in the NC group, 3 mice in the AAV only group and 4 mice in the RNP+AAV group. Twelve weeks after transplantation, the mice were euthanized, and the BM was harvested to determine the engraftment potential and the rate of targeted HSPCs.

Assessment of human engraftment

Twelve weeks after transplantation, mice were euthanized, and the bones were harvested from the mice (2× femur, 2× tibia, sternum, 2× pelvis, and spine) and crushed using a mortar and pestle. Mononuclear cells were enriched by using Ficoll gradient centrifugation (Ficoll-Paque Plus, GE Healthcare) for 25 min at 2,000 g at room temperature. After centrifugation, mononuclear cells were collected from the Ficoll layer and blocked to prevent nonspecific antibody binding (TruStain FcX, BioLegend) and stained (30 min, 4 °C, dark) with monoclonal antibodies: anti-human CD45 V450 (HI30; BD Biosciences); anti-HLA-ABC APC-Cy7 (W6/32; BioLegend); anti-CD19 APC (HIB19; BD Biosciences); and anti-CD33 PE (WM53, BD Biosciences). Stained cells were then washed and resuspended in FACS buffer and analyzed on a BD FACS Sort II Aria. HLA-ABC+/CD45+ cells represent human engraftment. Normal multilineage engraftment was defined by the presence of myeloid cells (CD33+) and B-cells (CD19+) within engrafted human CD45+HLA-ABC+ cells. GFP expression within engrafted human CD45+HLA-ABC+ cells was also analyzed by flow cytometry and represents the percentage of targeted cell engraftment.

Statistics

All statistical analyses were performed using SPSS Statistics (SPSS). Student's t-tests, one-way ANOVA were used to analyze the data. $P < 0.05$ was considered statistically significant (* means $P < 0.05$; **

means $P < 0.01$). Figures were prepared using GraphPad Prism (GraphPad Software).

Results

Development of a universal approach targeting the *HBB* gene

To develop a universal strategy to correct most *HBB* mutations, we tested the scheme shown in Figure 1. Briefly, we designed a strategy to achieve HDR in intron 2 of the *HBB* locus. Site-specific DSB in intron 2 was created by the ribonucleoprotein (RNP) complex consisting of sgRNA and Cas9 protein, and HDR was achieved using an rAAV6 homologous donor as a repair template. To facilitate testing the universal DNA donor template, we further added an EGFP reporter gene downstream of the *HBB* gene.

Optimizing the delivery of Cas9/sgRNA RNP into hematopoietic stem/progenitor cells

In our study, we selected five sgRNAs with high scores targeting the second intron of *HBB* using the web-based search tool CCTop. sgRNA4 had the highest on-target indel rate in pools of 293T cells electroporated with a plasmid (Figure 2A, Figure S1A), and there were no observable off-targets in the top seven off-targets predicted by CCTop (Figure S1B). Therefore, we chose sgRNA4 as the optimized sgRNA targeting the second intron of *HBB*. Some studies have clarified that the Cas9/sgRNA system delivered as an RNP complex by electroporation is the most effective method for creating DSBs and stimulating HR in HSPCs[14, 19], so we optimized the delivery of the Cas9/sgRNA RNP complex into cord blood-derived HSPCs to acquire the highest indel efficiency while minimizing cell death. First, we compared the different molar ratios of Cas9 protein and sgRNA from 1:1 to 1:3. We found that the indel frequency increased with the ratio until the molar ratio was 1:2.5, and there was no significant difference in the indel frequency between 1:2.5 and 1:3 (Figure 2B). We next compared the different electroporation parameters, and the results illustrated that 1,600 V was the most optimized electroporation parameter for the high indel frequency and cell survival rate (Figure 2C). It has previously been shown that 2’-O-methyl 3’phosphorothioate (MS) modifications to the three terminal nucleotides at the 5’ and 3’ ends of the sgRNA significantly increase the ability of the Cas9/sgRNA system to induce DSBs in HSPCs[16, 20]. Based on previous work, we determined whether modified sgRNA was more effective than in vitro transcription sgRNA (IVT sgRNA). By introducing the Cas9/sgRNA RNP into HSPCs by electroporation and harvesting cells four days later, we found a significant indel ratio increase generated at the *HBB* intro locus when the Cas9/sgRNA system was delivered as an RNP with modified guides versus unmodified guides (Figure 2D). Above all, the indel ratio targeting the *HBB* intro locus can reach an average of 80%.

Optimizing rAAV6 donor transduction into HSPCs to achieve consistently high levels of HDR

We next optimized rAAV6 donor transduction by titering the MOI or vector genomes/cell (vgs/cell) that stimulated the highest rates of HDR in HSPCs with an appropriate cell survival rate. Following the delivery of Cas9/sgRNA RNP to cells by electroporation, CB HSPCs were transduced with the rAAV6 donor at MOIs ranging from 1×10^3-1×10^6 vgs/cell. We found that the rates of HDR reached an average of 12% at an MOI of 1×10^4 (Figure 3A), and the rates of HDR were not increased with increasing MOI, especially for 1×10^5

and 1×10^6, because of the very low cell viability (Figure 3B). Additionally, we found that the fluorescence intensity of HSPCs after delivery of RNP and transduction with rAAV6 reached a plateau at day 4 (Figure 3C); above all, we chose an MOI of 1×10^4 for the following experiments. To further verify the HDR targeting the *HBB* locus, we detected the targeted integration of approximately nine samples through in-out PCR, and the results illustrated that all samples had targeted integration at the *HBB* intro locus (Figure S2A and S2B). Previous studies have clarified that the small-molecule compounds Scr7 and L755507 can improve the HDR ratio[21, 22]. To confirm the effect of the two small-molecule compounds, CB HSPCs transfected with RNP and rAAV6 were cultured in medium containing different concentrations of scr7 or L755507 for 24 h, and the data suggested that the HDR efficiency increased but the cell viability decreased with increasing concentration (Figure S3A and S3B).

Identify the genotype of edited clones at the *HBB* locus

To identify the genotype of *HBB*-targeted HSPCs, we performed single-cell methylcellulose cloning of populations targeted at the *HBB*-a process that took 2-3 weeks (Figure 4A). In experiments using only a GFP-expressing rAAV6 donor, we observed that HSPCs receiving only the rAAV6 donor expressed very low levels of GFP, but HSPCs that received Cas9 RNP and the rAAV6 donor simultaneously at day 4 after electroporation expressed much higher levels of GFP (Figure S4A). We hypothesized that the GFP^{high} population was enriched for *HBB*-targeted cells. Therefore, we sorted and cultured the GFP^{high} population in methylcellulose medium for 2 weeks, and the HSPCs sorted by flow cytometry can differentiate into multiprogenitors with consistently high GFP expression after 14 days of culture (Figure S4B), suggesting that the GFP^{high} population was indeed *HBB*-targeted. To confirm that the GFP^{high} population was enriched for on-target integration and analysis of whether cells possessed momo- or biallelic integrations, we used “in-out PCR” to determine allelic distribution in methylcellulose clones derived from the GFP^{high} population (23 clones). A total of 96% of clones had targeted integration, with 35% containing biallelic integrations and 61% containing monoallelic integrations (Figure 4B and 4C). Collectively, our data suggested that the Cas9 RNP and rAAV6 platform can deliver the *HBB* gene to a specific locus and achieve a high level of HDR followed by cells sorted wtih flow-cytometry.

Functional analysis of CB HSPCs after delivery with RNP and rAAV6 donor

To observe whether the *HBB*-targeted strategy can influence the function of HSPCs, we performed a hematopoietic progenitor CFU assay to show the forming ability of lineage-restricted progenitors (BFU-E, CFU-E, and CFU-GM) and multipotent progenitors (CFU-GEMM). The data demonstrated that there was no significant difference between the NC and RNP+AAV-GFP^{high} groups (Figure S5A). Similarly, the directional erythroid differentiation ability on days 10 and 15 was not significantly different between the two groups (Figure S5B). Additionally, Wright-Giemsa staining was used to show the cell morphological changes on erythroid differentiation days 10 and 15. The results suggested that both groups of cells had normal cell morphology, and they formed late erythroblasts with smaller cells and nuclear pyknosis with light cytoplasm (Figure S5C). Above all, our data suggested that the *HBB*-targeted strategy did not influence the function of HSPCs.

Whole-genome sequencing analysis of gene-targeted HSPCs

CRISPR/Cas9 has highly improved gene targeting efficiency, but it has the potential risk of causing off-target mutagenesis. Although we predicted the potential off-target sites using CCTop and confirmed no observable off-targets in the top seven sites through T7E1 assays, because of the poor sensitivity of the T7E1 assay to discover rare off-target events, high-throughput sequencing was applied to sequence the whole genomes of targeted cells. Due to the importance of exon regions, which affect most cell biological functions, we focused on the analysis of the sequencing results from these regions. Genomic DNA from mixed cells in the RNP+AAV group and sorted cells in the RNP+AAV-GFP group was detected with whole-genome sequencing. Compared with the untargeted HSPCs, cells in the RNP+AAV-GFP group were detected to have 23 single-nucleotide variants (SNVs) (Figure 5A) and 6 indels (Figure 5B). In the RNP+AAV group, 72 SNVs and 22 indels were detected. The variations in the RNP+AAV-GFP group were then compared with the sequences in the NC group to enable the generation of a list of potential variations that may arise during the gene-editing process (Figure 5C; Table S1). We should consider how many of those variations were the results of off-target mutagenesis by the edited program. All sequences of the variations, including SNVs and indel sites, in the RNP+AAV-GFP group were compared with the gRNA target sequence, and all of them were not within a potential off-target region, which strongly suggests that these variations were randomly accumulated during regular cell expansion and are not direct results of off-target activities by gene editing.

In vivo assessment of CB HPSCs treated with Cas9 RNP and the rAAV6 donor template

To determine whether gene-targeted CB HPSCs retain the ability to engraft, we performed an engraftment study using CB HSPCs derived from healthy donors with the delivery of Cas9 RNP and rAAV6 donors by tail-vein injection into NSI mice. Four days after the gene-editing reagents were delivered, 1,000,000 viable cells were transplanted into NSI mice. These mice were euthanized 12 weeks after transplantation, and we harvested the BM to determine the engraftment potential. The data in the NC and AAV only groups were derived from 3 transplanted mice, and the data in the RNP+AAV group were from 4 transplanted NSI mice. All transplanted mice displayed human engraftment in the BM as measured by the presence of hCD45/HLA-ABC double-positive cells (Figure 6A). Compared to the NC group, the treatment groups were observed to have a decrease in the rates of human cell chimerism, and the two groups of AAV only and RNP+AAV displayed similar chimerism rates (Figure 6B), which can be explained by the different viability of transplanted cells. The viability of the NC group reached over 90%; however, the viability of treated cells was only 40-60% (data not shown). We further analyzed the percentage of GFP expression in human cell chimerism, which displayed the engraftment potential of gene-targeted CB HSPCs (Figure 6A). We found that there was a significant decrease in the percentage of RNP+AAV in vitro targeting frequencies (12% in the CB HSPCs) compared to the percentage of GFP+ cells in the BM at week 12 after transplantation (2.1%) (Figure 6C). This decrease was consistent with a previous study[15, 17]. Additionally, we observed myeloid (CD33) and lymphoid (CD19) reconstitution with an average of 2.1% and 1.6% GFP+ cells, respectively (Figure 6A, Figure 6C), implicating targeting of multipotent HSPCs. The data confirm that the gene-targeted HSPCs had engraftment potential and lineage reconstitution ability.

HBB gene correction of β-CD41/42 HSPCs derived from fetal liver

To test whether our strategy can correct the homozygous β-CD41/42 mutation in HSPCs derived from fetal liver and whether the edited β-CD41/42 HSPCs maintained their erythroid differentiation, we subjected the HSPCs in the NC and RNP+AAV groups to a 21-day in vitro erythroid differentiation protocol. We detected the correction of the CD41/42 mutation by Sanger sequencing (Figure S6A and S6B). Flow cytometry analyses after erythroid differentiation showed a higher proportion of CD235a+CD71+ cells in the RNP+AAV group than in the NC group (Figure 7A), indicating a much greater presence of mature differentiated erythroids that express hemoglobin in edited cells. To confirm that *HBB* mRNA was transcribed from edited *HBB* alleles, we performed reverse transcription-quantitative PCR (RT-qPCR) on erythrocytes differentiated from HSPCs in the NC and RNP+AAV groups. The results suggested that the expression of *HBB* in edited cells was nearly 6 times higher than that in the NC cells, along with a decrease in HBF mRNA (Figure 7B), which indicated that our strategy can correct the β-CD41/42 mutation and increase the efficiency of erythroid differentiation in edited cells. Notably, erythrocytes differentiated from edited HSPCs expressed more *HBB* mRNA, validating this strategy.

Discussion

In this study, we established an efficient and universal repair system targeting common *HBB* mutation sites using the Cas9 RNP in combination with rAAV6 homologous donor vectors. This universal strategy enables correction of not only *HBB* mutations but also some upstream and downstream mutations in the *HBB* gene.

The rAAV6 donor in our universal strategy reserves intron 1 of the *HBB* gene, which is different from the cDNA donor that is used in a common treatment strategy for the correction of diseases with loss-of-function mutations scattered throughout a particular gene[23, 24]. Because *HBB* intron 1 has an important regulatory function in the expression of the *HBB* gene and rAAV6 has a 4.7 kb packaging capacity, packaging intron 1 into the donor is feasible; therefore, retaining intron 1 minimized the impact of *HBB* expression.

In the process of optimizing the delivery of our RNP and rAAV6 donor, we discovered that the modified synthetic sgRNA is significantly higher than IVT sgRNA in indel ratio and cell viability. The indel ratio averages 80%, which is beneficial because the modified synthetic sgRNA is not degraded by the cellular immune response, and chemical modification provides greater stability and protection from exonucleases. This is a finding that our group and others have also demonstrated[24, 25]. A higher indel ratio is the basis of high HDR efficiency, and it is nearly 12% HDR efficiency in our study. To further improve HDR efficiency, we tested two different small-molecule compounds, Scr7 and L755507. Scr7 is a specific DNA ligase IV inhibitor that can improve HDR efficiency by inhibiting the nonhomologous end-joining (NHEJ) pathway[26], and L-755507 is a selective β3 adrenergic receptor partial agonist that enhances CRISPR-mediated HDR efficiency[21]. Two of these small-molecule compounds have been proven to improve the HDR ratio in mammalian cells, mouse embryonic stem cells and pluripotent stem

cells in other studies[22, 27]. We found that the effects of these two compounds on HDR efficiency are concentration-dependent and that the HDR efficiency increases with the concentration of the compounds; however, cell viability decreases significantly with increasing concentration, which is not conducive to subsequent functional studies and NSI mouse transplantation experiments. More suitable small-molecule compounds improving HDR efficiency need to be further tested to increase the feasibility of clinical applications in the future.

We observed that there is no off-target effect that is related to the repair system by whole-genome sequencing and off-target analyses. We mainly analyzed the sequencing results from exon regions due to the important effect of exons on cell biological functions. Another reason is that the targeted HSPCs used for sequencing are mixed cells, which are not cell populations from a single clone cell; therefore, heterogeneity between cells may lead to a large number of false-positive results in introns and intergenic regions.

Our results demonstrated that gene-targeted CB HSPCs were able to engraft in NSI mice. However, there was a reduction in the percentage of chimeric cells at 12 weeks post-transplantation in the RNP + AAV group, which was consistent with the results of other studies and may reflect a reduced engraftment capacity of edited mixed cells[16, 17]. The reason that the engraftment capacity of edited mixed HSPCs is decreased compared with that of unedited cells may be that the HDR donors are themselves toxic or that HSCs are intrinsically more difficult to edit than other cells[28]. We further demonstrated that GFP + HSPCs representing the gene-targeted HSPCs can be detected in the BM of NSI mice, which indicated that some of the targeted HSPCs were stable for up to 12 weeks post-transplantation, although the in vivo gene-targeted HSPCs had a nearly 6-fold reduction compared with the in vitro gene-targeted HSPCs. To improve the percentage of gene-targeted long-term HSCs in transplanted NSI mice, the HDR efficiency of long-term HSCs in vitro should be further improved by some small-molecule compounds that can control the cell cycle because HDR is normally most active during the S or G2 phase of the cell cycle.

We also demonstrated that the β-CD41/42 mutation can be repaired by this strategy; moreover, the gene-corrected HSPCs have a higher erythroid differentiation efficiency and a 6-fold increase in *HBB* mRNA expression. Due to the limitations of patient-derived HSPCs, we used only β-CD41/42 HSPCs to confirm the feasibility of our strategy, and more β-thal patient-derived HSPCs should be further applied to illustrate the function of the repair system, which is important for clinical translation.

In conclusion, we provided an efficient and universal repair system targeting most *HBB* mutations, including exons, upstream and downstream of the *HBB* gene, using Cas9 RNP combined with an rAAV6 donor. Although our study used EGFP as a selection marker that is an unsuitable marker for gene therapy, we can use the truncated nerve growth factor receptor (tNGFR) marker or other similar signaling-inert cell surface markers to enrich the gene-corrected HSPCs to increase the feasibility of clinical applications. Overall, our strategy lays a theoretical foundation for the clinical application of gene therapy for β-haemoglobinopathy based on gene editing.

Conclusion

In conclusion, our study provides a universal strategy to correct most types of *HBB* gene mutations in β-thal. The edited HSPCs have normal function in vitro and engraftment potential. patient-derived HSPCs corrected by the approach expressed more *HBB* mRNA than uncorrected cells. In addition, it provides an experimental system to screen the small-molecule compounds to improve HDR efficiency in HSPCs based on the coexpression of the EGFP reporter.

Abbreviations

HSPCs: Hematopoietic Stem and Progenitor Cells; CRISPR/Cas: Clustered reg□larly interspaced short palindromic repeats; RNP: Ribonucleoprotein complex; rAAV6: recombinant adeno-associated virus6; sgRNA: small guide RNA; HDR: homology-directed repair; CFU: Colony-Forming Unit; tNGFR: the truncated nerve growth factor receptor; NHEJ: non-homology end joining; DSB: double-strand break. ssODNs: single-strand oligodeoxynucleotides; FACS: Fluorescent-activated cell sorting.

Declarations

Acknowledgements

We thank the reviewers for their constructive comments.

Authors' contributions

Li-Na He designed the protocol, analyzed results and finished the paper; Yi Yang denoted to the study design; Yi Cheng, Han Wu, Shou-Heng Lin, and Bing Song participated in the process of experiment; Neng-Qing Liu and Ying-Hong Yang collected the sample; Di-Yu Chen and Dian Lu performed the statistical analysis; Juan Zeng reviewed the manuscript; Yong Fan provided parts of experimental materials; Xiao-Fang Sun held the quality control of this study.

Funding

This study was supported by the Guangzhou City Science and Technology Key Topics Project (201904020025), the National Key R&D Program of China (2019YFA0110804), National Natural Science Foundation of China (31872800), Clinical Innovation Research Program of Guangzhou Regenerative Medicine and Health Guangdong Laboratory (2018GZR0201002), Natural Science Foundation of Guangdong Province (2014A030312012), Science Foundation of Guangdong Province (2016B030229008), Foundation of Guangzhou Science and Information Technology of Guangzhou Key Project (201803040009), Guangdong Province Science and Technology Project (2017A030310380), the Guangzhou Science and Technology Project (201803010048).

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

All participants in this study provided written informed consent for donating blood. The Ethics Committee of The Third Affiliated Hospital of Guangzhou Medical University (Guangzhou, China) approved experiments using human cells. The animal ethics were approved by The Third Affiliated Hospital of Guangzhou Medical University. All methods were performed in accordance with the approved guidelines.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Raffaella O. β-Thalassemia. genetics inmedicine. 2017; 19:609–619.
2. Shah FT, Sayani F, Trompeter S, Drasar E, Piga A. Challenges of blood transfusions in β-thalassemia. Blood reviews. 2019;37:1–13.
3. Cavazzana M, Antoniani C, Miccio A. Gene Therapy for β-Hemoglobinopathies. Molecular therapy: the journal of the American Society of Gene Therapy. 2017;25:1142–54.
4. Jagannath VA, Fedorowicz Z, Al Hajeri A, Sharma A. Hematopoietic stem cell transplantation for people with ß-thalassaemia major. The Cochrane Database of Systematic Reviews. 2016;11:CD008708.
5. Brendel C, Williams DA. Current and future gene therapies for hemoglobinopathies. Curr Opin Hematol. 2020;27:149–54.
6. Lamsfus-Calle A, Daniel-Moreno A, Ureña-Bailén G, Raju J, Antony J, Handgretinger R, et al. Hematopoietic stem cell gene therapy: The optimal use of lentivirus and gene editing approaches. Blood reviews. 2019;15:100641.
7. Magrin E, Miccio A, Cavazzana M. Lentiviral and genome-editing strategies for the treatment of β-hemoglobinopathies. Blood. 2019;134:1203–13.
8. Thompson AA, Walters MC, Kwiatkowski J, Rasko JEJ, Ribeil JA, Hongeng S, et al. Gene Therapy in Patients with Transfusion-Dependent beta-Thalassemia. N Engl J Med. 2018;378:1479–93.
9. Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science. 2003;300:1749–51.
10. Biffi A. Gene Therapy as a Curative Option for beta-Thalassemia. N Engl J Med. 2018;378:1551–2.

11. Dever DP, Porteus MH. The changing landscape of gene editing in hematopoietic stem cells: a step towards Cas9 clinical translation. Curr Opin Hematol. 2017;24:481–8.
12. Zhang J, Li XL, Li GH, Chen W, Arakaki C, Botimer GD, et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. Genome biology. 2017;18:35.
13. Jacinto FV, Link W, Ferreira BI. CRISPR/Cas9-mediated genome editing: From basic research to translational medicine. J Cell Mol Med. 2020;00:1–13.
14. Dewitt MA, Corn JE, Carroll D. Genome editing via delivery of Cas9 ribonucleoprotein. Methods. 2017;121:9–15.
15. Hoban MD, Lumaquin D, Kuo CY, Romero Z, Long J, Ho M, et al. CRISPR/Cas9-Mediated Correction of the Sickle Mutation in Human CD34 + cells. Molecular therapy: the journal of the American Society of Gene Therapy. 2016;24:1561–9.
16. Dever DP, Bak RO, Reinisch A, Camarena J, Washington G, Nicolas CE, et al. CRISPR/Cas9 beta-globin gene targeting in human haematopoietic stem cells. Nature. 2016;539:384–9.
17. Mark AD, Wendy M, Nicolas LB. Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. Science translational medicine. 2016;8:360ra134.
18. Megan DH, Gregory JC, Matthew CM. Correction of the sickle cell disease mutation in human hematopoietic. Gene Ther. 2015;29:234–9.
19. Lattanzi A, Meneghini V, Pavani G, Amor F, Ramadier S, Felix T, et al. Optimization of CRISPR/Cas9 Delivery to Human Hematopoietic Stem and Progenitor Cells for Therapeutic Genomic Rearrangements. Molecular therapy: the journal of the American Society of Gene Therapy. 2019;27:137–50.
20. Chakrabarti AM, Henser-Brownhill T, Monserrat J, Poetsch AR, Luscombe NM, Scaffidi P. Target-Specific Precision of CRISPR-Mediated Genome Editing. Molecular cell. 2019;73:699–713.
21. Liu Y, Yang Y, Kang X, Lin B, Yu Q, Song B, et al. One-Step Biallelic and Scarless Correction of a beta-Thalassemia Mutation in Patient-Specific iPSCs without Drug Selection. Molecular therapy Nucleic acids. 2017;6:57–67.
22. Ma Y, Chen W, Zhang X, Yu L, Dong W, Pan S, et al. Increasing the efficiency of CRISPR/Cas9-mediated precise genome editing in rats by inhibiting NHEJ and using Cas9 protein. RNA Biol. 2016;13:605–12.
23. Cai L, Bai H, Mahairaki V, Gao Y, He C, Wen Y, et al. A Universal Approach to Correct Various HBB Gene Mutations in Human Stem Cells for Gene Therapy of Beta-Thalassemia and Sickle Cell Disease. Stem cells translational medicine. 2018;7:87–97.
24. Bak RO, Dever DP, Porteus MH. CRISPR/Cas9 genome editing in human hematopoietic stem cells. Nature protocols. 2018;13:358–76.
25. Scott T, Soemardy C, Morris K. Development of a Facile Approach for Generating Chemically Modified CRISPR/Cas9 RNA. Molecular therapy Nucleic acids. 2020;19:1176–85.

26. Vartak SV, Swarup HA, Gopalakrishnan V, Gopinatha VK, Ropars V, Nambiar M, et al. Autocyclized and oxidized forms of SCR7 induce cancer cell death by inhibiting nonhomologous DNA end joining in a Ligase IV dependent manner. FEBS J. 2018;285:3959–76.
27. Vartak SV, Raghavan SC. Inhibition of nonhomologous end joining to increase the specificity of CRISPR/Cas9 genome editing. FEBS J. 2015;282:4289–94.
28. Wagenblast E, Azkanaz M, Smith SA, Shakib L, Mcleod JL, Krivdova G, et al. Functional profiling of single CRISPR/Cas9-edited human long-term hematopoietic stem cells. Nature communications. 2019;10:10–8.

Figures

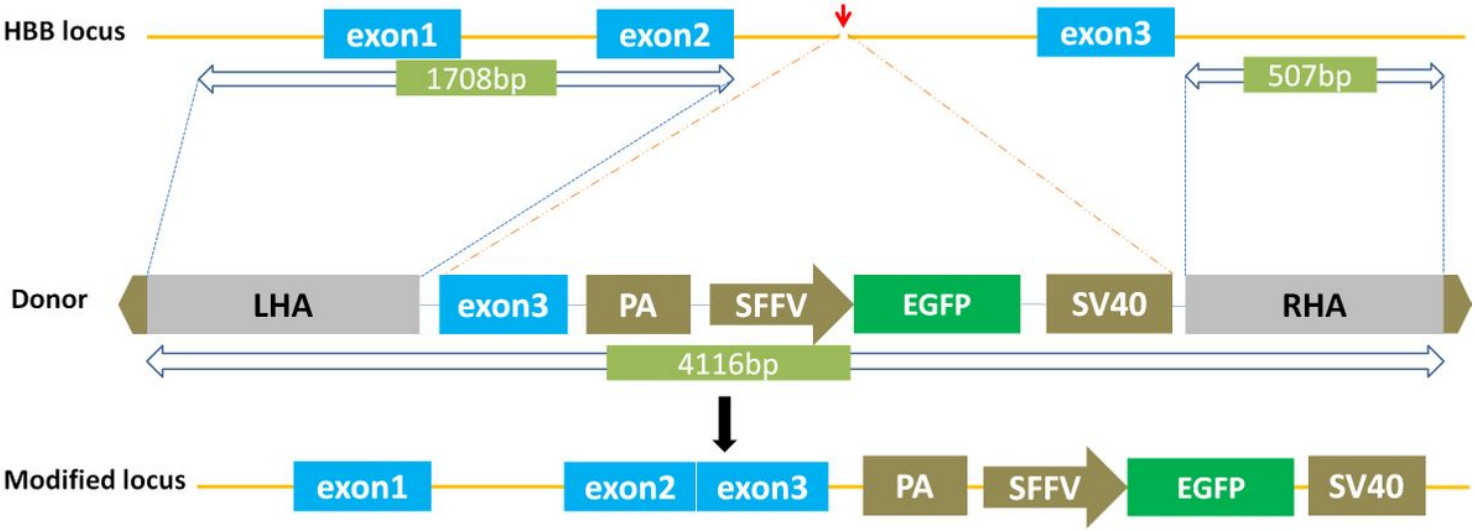


Figure 1

Schematic of targeted genome editing at the HBB locus using CRISPR/Cas9 and AAV6. Site-specific DSBs are created by CRISPR/Cas9 (red arrow). A DSB stimulates homologous recombination (HR) using the rAAV6 homologous donor as a repair template. Blue boxes: HBB exons; light gray boxes: homology arms; green boxes: EGFP selection marker.

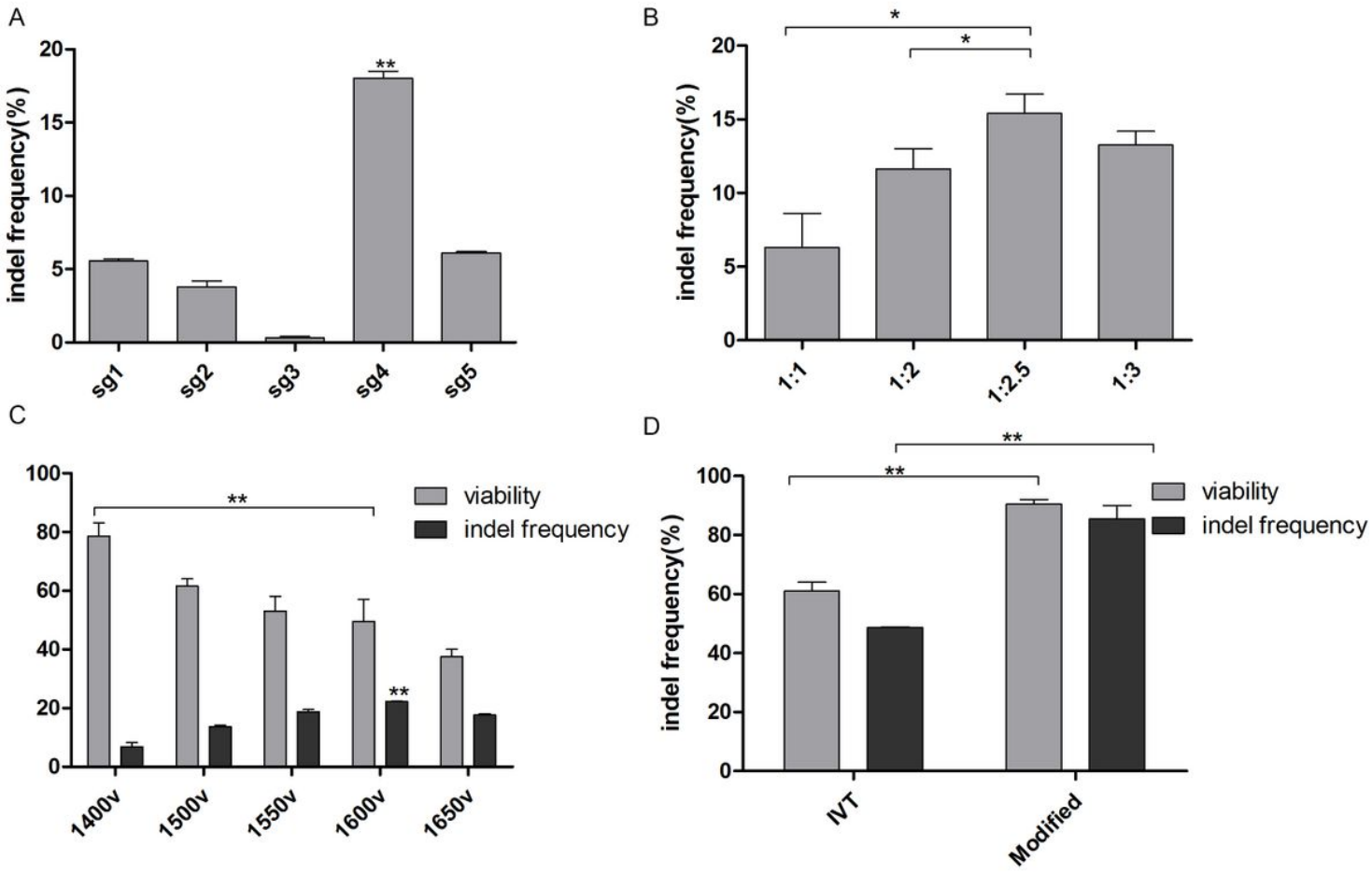


Figure 2

Highly efficient sgRNA screening targeted on HBB intron locus and optimizing the delivery of Cas9/sgRNA RNP into hematopoietic stem/progenitor cells. (A) The indel frequency of five sgRNAs targeted to the HBB intron locus in pools of 293T cells was assessed by TIDE software. (B) The indel frequency of different molar ratios of Cas9/sgRNA RNP was assessed by TIDE software after electroporation in HSPCs. (C) The indel frequency of different electroporation parameters was assessed by TIDE software after electroporation in HSPCs. (D) The indel frequency of different sgRNA treatments, in vitro transcription (IVT) or modified sgRNA. All data represent mean ± SD. *P<0.05, **P<0.01 by Student's t test in data (D) and others by one-way ANOVA test.

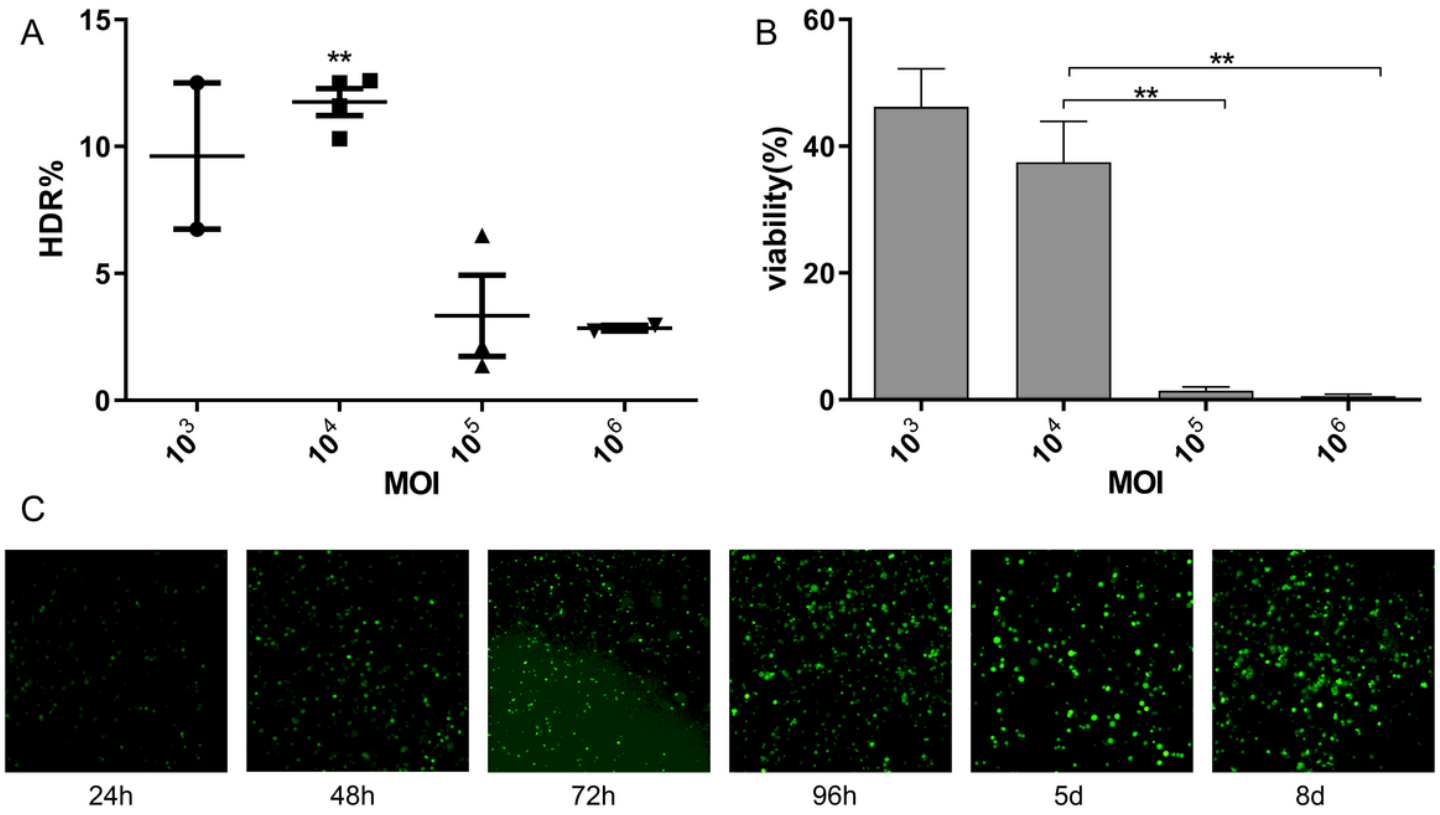


Figure 3

Optimizing HDR efficiency in cord blood-derived CD34+ HSPCs. (A) The HDR efficiency of different rAAV6 MOIs was analyzed by flow cytometry 4 days after the delivery of RNP and rAAV6 into HSPCs. (B) The viability of different rAAV6 MOIs was analyzed by trypan blue staining 4 days after the delivery of RNP and rAAV6 into HSPCs. (C) The fluorescence intensity of HSPCs at different times after the delivery of RNP and rAAV6 into HSPCs; the rAAV6 MOI was 104. All data represent mean ± SD. **P<0.01 by one-way ANOVA test.

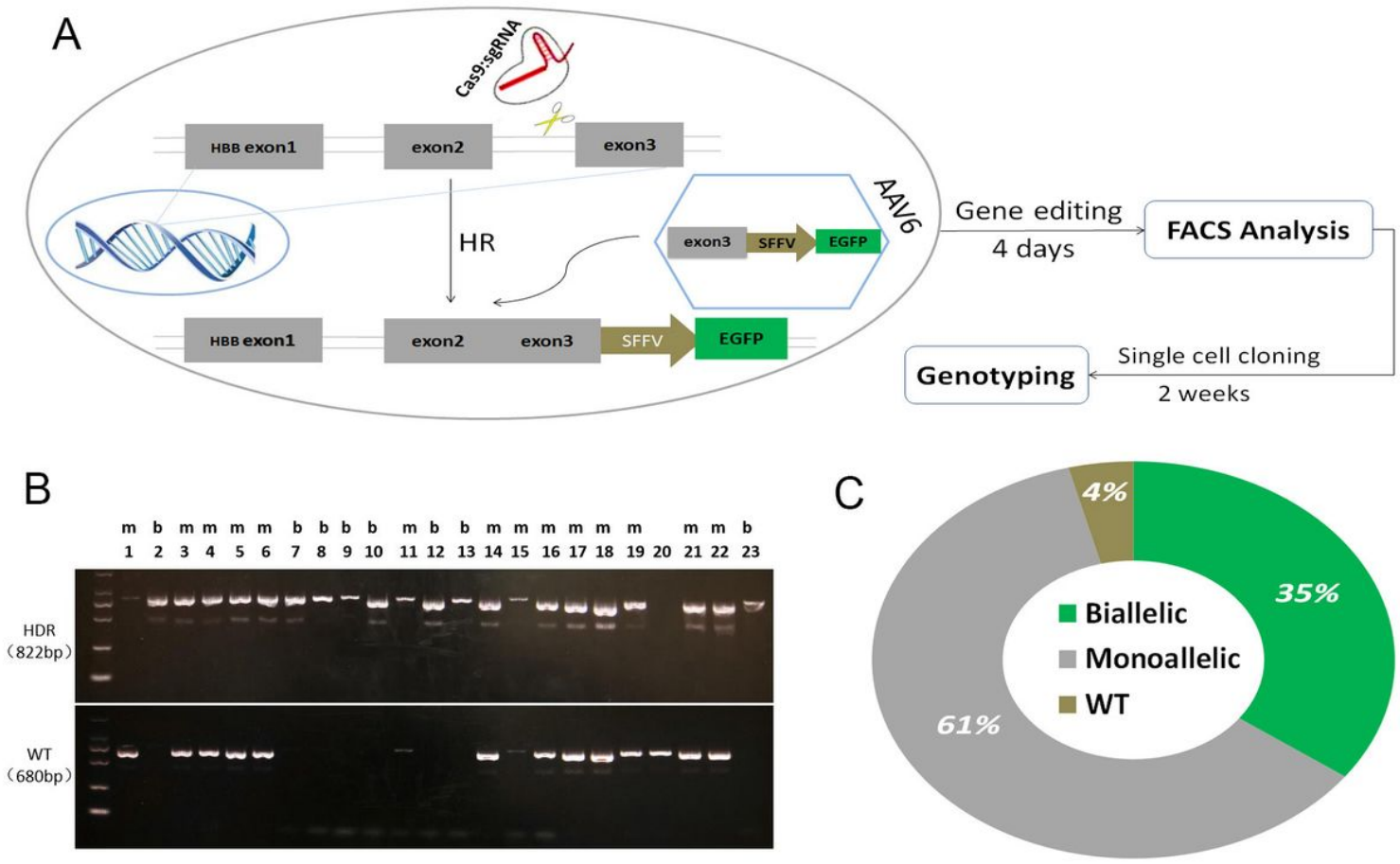
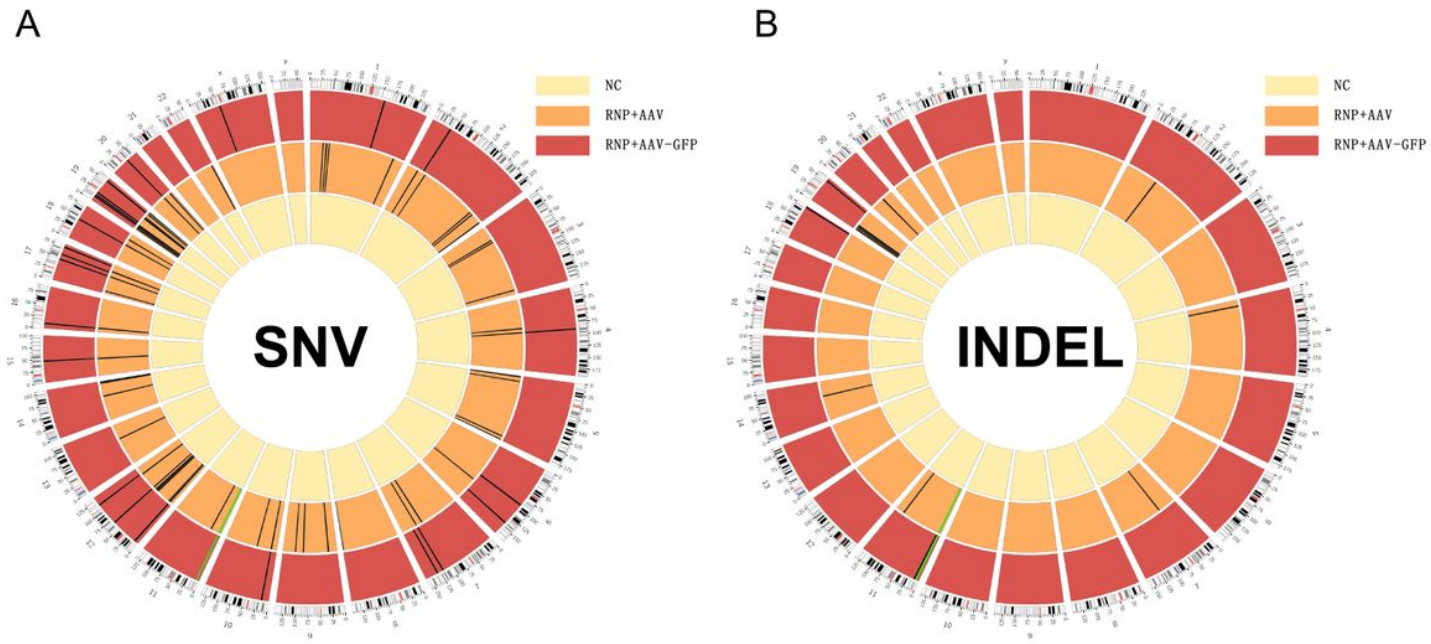


Figure 4

Identify the genotype of methylcellulose colonies from GFPhigh HSPCs. (A) Schematic depicting the genome editing process delivering Cas9 RNP with rAAV6 donor followed by flow cytometry analysis 4 days posttransduction, allowing genotyping of clones within 2 weeks. (B) Agarose gel images show genotypes of 23 clones targeted at the HBB. (C) Pie chart summarizing the results B. b: Biallelic; m: Monoallelic.



C

Indels				gRNA target site PAM
Loci		Gene	Indel	5’ -GACGAATGATTGCATCAGTGTGG-3’
11	5153033	OR52A5	G	TGAGAAAAGGTGGGACTAACAGGTAAAGATGTTGACAAGAGGATATGAATATATGGTGGTA
18	72997849	TSHZ1	TG	AGCTCCGCCCCCACCCCCACACCCCCCACCTGCCCCGTCAGCACCACTGGCCCCACCACGA
18	72997855	TSHZ1	GTCAGCA	GCCCCCACCCCCACACCCCCCACCTGCCCCGTCAGCACCACTGGCCCCACCACGAGCACGC
18	72997864	TSHZ1	ACTGG	CCCACACCCCCCACCTGCCCCGTCAGCACCACTGGCCCCACCACGAGCACGCCCAGCACCA
19	50813024	MYH14	GCCC	TGGGTCTGGGCCATCCCCGGAGCCTGAGGGGCCCGTCCCCACCAGCCCACCCCCAGTGACC
19	50813032	MYH14	AG	GCCATCCCCGGAGCCTGAGGGGTCCCCACCAGCCCACCCCCAGTGACCCTACCCTGTCCCC

Figure 5

Whole-genome sequencing of the gene-untargeted and gene-targeted HSPCs. (A) SNVs in exon regions of HSPCs via genome sequencing. The black bars represent individual SNVs. Compared with the NC group (untargeted HSPCs), HSPCs in the RNP+AAV-GFP group (targeted HSPCs) had 23 SNVs, and HSPCs in the RNP+AAV group (mixed HSPCs) had 72 SNVs. (B) Indels in the exon regions of HSPCs via genome sequencing. Compared with those in the NC group, HSPCs in the RNP+AAV-GFP group had 6 indels, and HSPCs in the RNP+AAV group had 22 indels. (C) Indel sequences detected in the RNP+AAV-GFP group via genome sequencing. Compared with the gRNA target site, no homologous sequences were found in the indel locus. The green lines in pictures A and B indicate the locus of donor insertion.

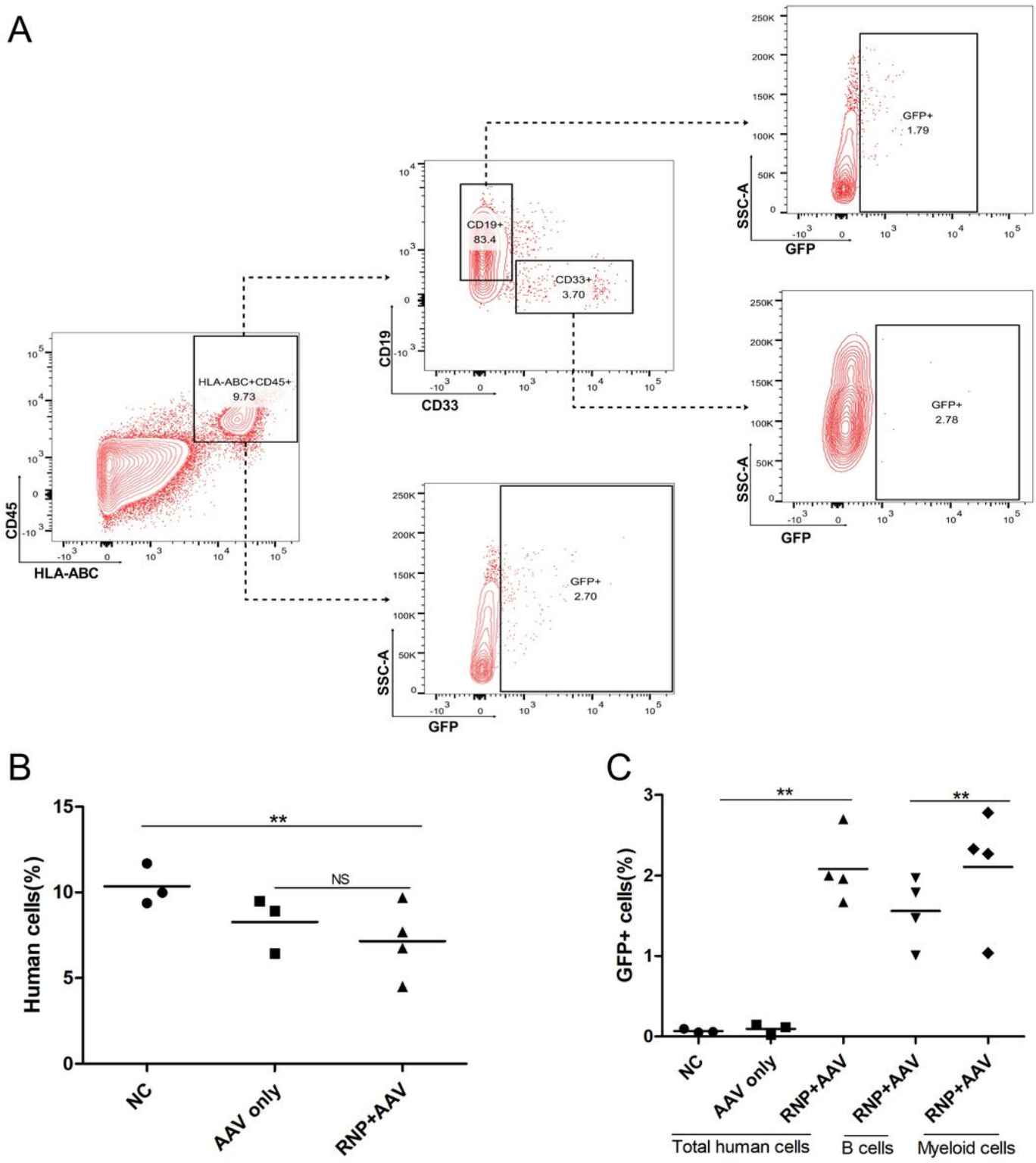


Figure 6

Genome-edited human HSPCs display long-term reconstitution in NSI mice at week 12 after transplantation. (A) Representative plots showing the gating scheme for analyses of NSI mice transplanted with human cells from one mouse in the RNP+AAV group. (B) Human engraftment in NSI mice from all experimental groups. (C) Percentage of GFP+ cells in the total human population, B-cell lineage, and myeloid lineage. All data represent mean ± SD. **P<0.01 by one-way ANOVA test.

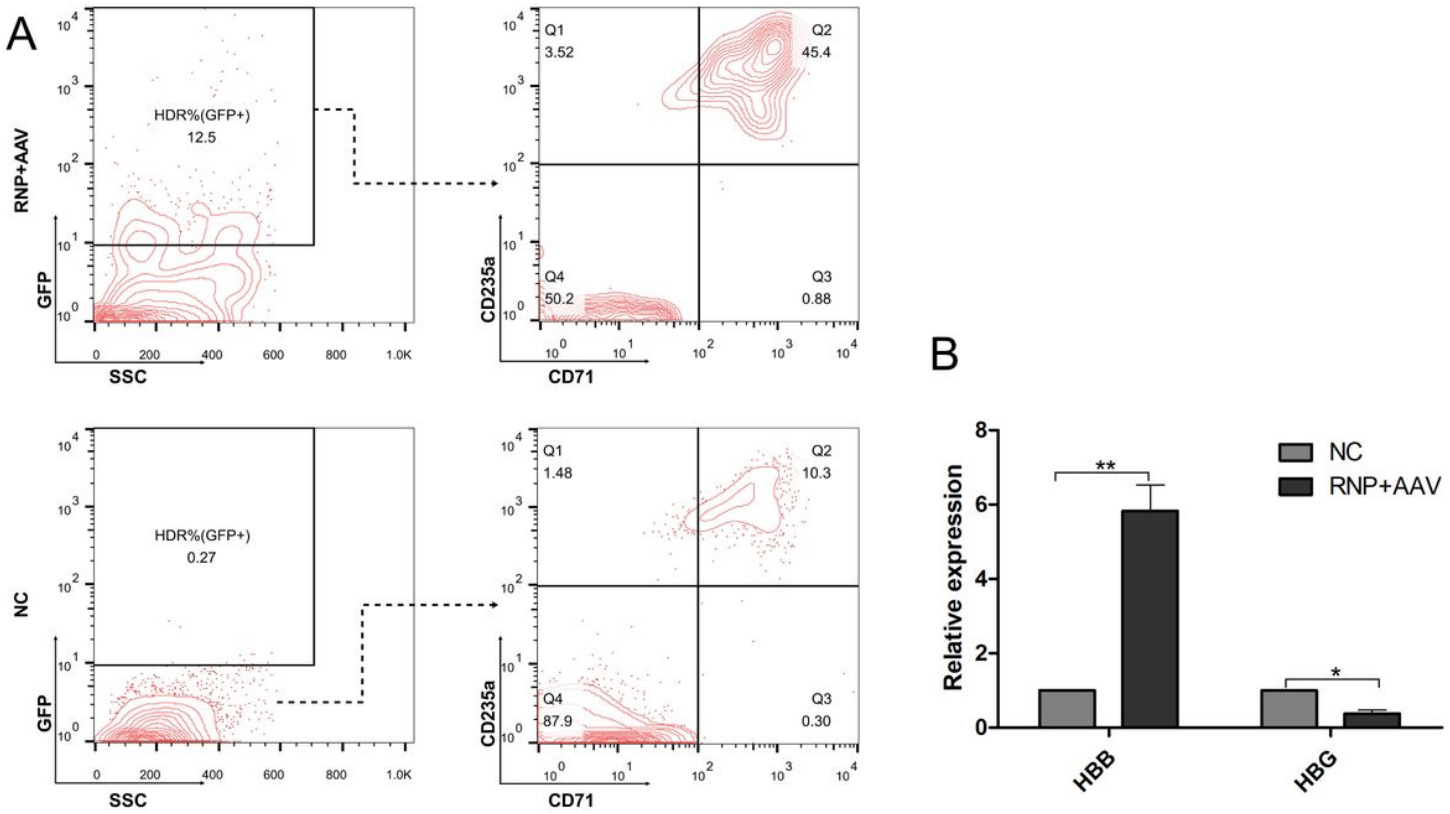


Figure 7

Gene correction of β-CD41/42 mutation in fetal liver-derived HSPCs. (A) HSPCs of the NC (untargeted β-CD41/42 HSPCs) and RNP+AAV (targeted β-CD41/42 HSPCs) groups were differentiated into erythrocytes in vitro. Representative FACS plots from day 21 of differentiation show GFP expression and the erythrocyte cell surface markers CD71 and CD235a. (B) HBB and HBF mRNA expression was quantified by RT-qPCR in erythrocytes differentiated from HSPCs in the NC and RNP+AAV groups, and all mRNA transcript levels were normalized to that of the GAPDH input control. All data represent mean ± SD. *P<0.05, **P<0.01 by Student's t test.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigS1S6.ppt
- TableS1.docx