Published in final edited form as:
Nature. 2016 November 17; 539(7629): 384–389. doi:10.1038/nature20134.

CRISPR/Cas9 Beta-globin Gene Targeting in Human Hematopoietic Stem Cells

Daniel P. Dever[1,*], **Rasmus O. Bak**[1,*], **Andreas Reinisch**[2], **Joab Camarena**[1], **Gabriel Washington**[1], **Carmencita E. Nicolas**[1], **Mara Pavel-Dinu**[1], **Nivi Saxena**[1], **Alec B. Wilkens**[1], **Sruthi Mantri**[1], **Nobuko Uchida**[4], **Ayal Hendel**[1], **Anupama Narla**[3], **Ravindra Majeti**[2], **Kenneth I. Weinberg**[1], and **Matthew H. Porteus**[1]

[1]Department of Pediatrics, Stanford University, Stanford, CA 94305, USA

[2]Department of Medicine, Division of Hematology, Cancer Institute, and Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, CA 94305, USA

[3]Division of Hematology/Oncology, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA

[4]Stem Cells, Inc. 7707 Gateway Blvd., Suite 140, Newark, CA 94560 USA

Abstract

The β-hemoglobinopathies, including sickle cell disease (SCD) and β-thalassemia, are caused by mutations in the β-globin gene (*HBB*) and affect millions of people worldwide. A curative strategy for the β-hemoglobinopathies would be *ex vivo* gene correction in patient-derived hematopoietic stem cells (HSCs) followed by autologous transplantation. Here we report the first CRISPR/Cas9 gene-editing platform for achieving homologous recombination (HR) at the *HBB* gene in HSCs by combining Cas9 ribonucleoproteins and rAAV6 HR donor delivery. Notably, we devise an enrichment paradigm to purify a population of HSPCs with >90% targeted integration. We also show efficient correction of the SCD-causing E6V mutation in patient-derived HSPCs that after differentiation into erythrocytes, express adult β-globin (HbA) mRNA, confirming intact transcriptional regulation of edited *HBB* alleles. Collectively, these preclinical studies outline a CRISPR-based methodology for targeting HSCs by HR at the *HBB* locus to advance the development of next generation therapies for β-hemoglobinopathies.

Correspondence should be addressed to M.H.P. (mporteus@stanford.edu).
*Co-first author

Competing Financial Interests:
MHP is a consultant and has equity interest in CRISPR Tx, but CRISPR Tx had no input into the design, execution, interpretation or publication of the results herein. Nobuko Uchida is an employee of Stem Cells, Inc, but they had no input into this manuscript.

Author Contributions:
D.P.D. and R.O.B. contributed equally to this work as well as performed and designed most of the experiments. D.P.D., R.O.B., A.R., and C.E.N. designed, executed and analyzed engraftment studies. J.C. performed some *HBB* HR, NHEJ and tNGFR experiments. G.W. performed *in vitro* erythrocyte differentiation experiments. MPD assisted with large-scale electroporation experiments. N.S. ran and analyzed HSPC immunophenotyping experiments. A.B.W. scored and harvested methylcellulose clones. N.U. sorted freshly isolated HSCs prior to targeting. S.M. purified $CD34^+$ HSPCs from peripheral blood of SCD patients. A.N. was the attending physician who oversaw transfusions of SCD patients. A.H., R.M., and K.I.W. contributed to experimental design and data interpretation. M.H.P directed the research and participated in the design and interpretation of the experiments and the writing of the manuscript. D.P.D and R.O.B. wrote the manuscript with help from all authors.

Europe PMC Funders Author Manuscripts

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) highlights the idea that transplantation of HSCs with only a single wild-type b-globin gene (*HBB*) can cure the b-hemoglobinopathies; however, allo-HSCT is limited because of graft versus host disease and lack of immunologically matched donors. An alternative to using allogeneic HSCs to cure the b-hemoglobinopathies is to use homologous recombination (HR) to directly modify the *HBB* gene in the patient's own HSCs1, 2. The first step towards this was described in 1985 when Smithies and colleagues were able to modify the human *HBB* gene by HR in a human embryonic carcinoma cell line, albeit at an extremely low frequency3 (10^{-6}). The subsequent discoveries that a site-specific DNA double-strand break (DSB) could stimulate HR-mediated correction of a reporter gene and that engineered nucleases could be used to induce this DSB, formed the foundation of using HR-mediated genome editing using engineered nucleases to directly modify the *HBB* gene4, 5. The ease of engineering as well as the robust activity of the CRISPR/Cas9 RNA-guided endonuclease system makes it a promising tool to apply to the ongoing challenge of developing effective and safe HR-mediated genome editing to cure b-hemoglobinopathies7, 8.

The CRISPR/Cas9 complex consists of the Cas9 endonuclease and a 100-nucleotide (nt) single guide RNA (sgRNA). Target identification relies first on identification of a 3-base pair protospacer adjacent motif (PAM) and then hybridization between a 20-nucleotide stretch of the sgRNA and the DNA target site, whivh triggers Cas9 to cleave both DNA strands9. DSB formation activates two highly conserved repair mechanisms: canonical non-homologous end-joining (NHEJ) and homologous recombination10 (HR). Through iterative cycles of break and NHEJ repair, insertions and/or deletions (INDELs) can be created at the site of the break. In contrast, genome editing by HR requires the delivery of a DNA donor molecule to serve as a homologous template, which the cellular HR machinery uses to repair the break by a 'copy and paste' method11. For gene editing purposes, the HR pathway can be exploited to make precise nucleotide changes in the genome4. One of the key features of precise genome editing, in contrast to viral vector-based gene transfer methods, is that endogenous promoters, regulatory elements, and enhancers can be preserved to mediate spatiotemporal gene expression1, 12–14. The CRISPR/Cas9 system is highly effective at stimulating DSBs in primary human HSPCs when the sgRNA is synthesized with chemical modifications, precomplexed with Cas9, and then electroporated into cells15.

HSCs have the ability to repopulate an entire hematopoietic system16, and several genetic17–19 and acquired20 diseases of the blood could potentially be cured by genome editing of HSCs. Recent studies have demonstrated efficient targeted integration in HSPCs by combining ZFN expression with exogenous HR donors delivered via single stranded oligonucleotides (ssODN)6, integrase-defective lentiviral vectors (IDLV)21, or recombinant adeno-associated viral vectors of serotype 6 (rAAV6)22, 23. In most of these studies, however, the high editing frequencies *in vitro* did not result in high frequencies of edited cells following transplantation into immunodeficient mice. Moreover, in some of these studies the HSPCs used were derived from fetal liver, which is a non-clinically relevant HSPC source compared to cells derived from bone marrow or mobilized peripheral blood. Collectively, these studies suggest that targeting HSCs by HR at disease-causing loci is difficult in clinically relevant HSPCs.

In this study, we achieve efficient *in vitro* HR-mediated editing frequencies at the *HBB* locus in $CD34^+$ HSPCs derived from mobilized peripheral blood (mPB) using Cas9 ribonucleoproteins (RNPs) combined with rAAV6 homologous donor delivery. In brief, we demonstrate: 1) Cas9 and rAAV6-mediated *HBB* targeting in HSCs characterized by the identification of modified human cells in secondary transplants of immunodeficient mice, 2) efficient correction of the SCD-causing E6V mutation in multiple SCD patient-derived HSPCs, and 3) development of a purification scheme using either FACS or magnetic bead enrichment to create HSPC populations in which >85% of the cells have been modified by HR-mediated targeted integration This purification can be performed early in the manufacturing process when HSCs are still preserved, and may prove valuable in a clinical setting for removing untargeted HSPCs that will be in competition with HR-edited HSPCs for engraftment and re-population following transplantation.

CRISPR/Cas9 targeted *HBB* gene editing in HSPCs

As other groups have reported, we confirmed high transduction of HSPCs using a self-complementary AAV6 (scAAV6) with an SFFV-GFP expression cassette24 (Extended Data Fig 1A). *HBB*-specific single-stranded AAV6 (ssAAV6) vectors were then produced containing SFFV-GFP flanked by arms homologous to *HBB* (Fig 1a). To achieve gene editing at *HBB*, we employed two different CRISPR platforms, which we have previously shown to be highly active in primary cells15. Both platforms use sgRNAs chemically modified at both termini with 2' O-Methyl 3′phosphorothioate (MS sgRNA) and are either delivered in conjunction with Cas9 mRNA or as a ribonucleoprotein (RNP) complex. Both platforms yielded high INDEL frequencies when electroporated into HSPCs, with the RNP showing higher activity (Fig 1b). By supplying ssAAV6 *HBB* donors after electroporation of Cas9 RNP we achieved stable GFP expression in an average of 29% of HSPCs (Fig 1c). We observed lower efficiencies using the mRNA platform (15%) (Fig 1c). Cytotoxicity and off-target cleavage activity (of a reported off-target site)15 was significantly decreased using the RNP system (Extended Data Fig 1B-D).

Because AAV genomes can be captured at the site of an off-target DSB via NHEJ22, 25, 26, we performed experiments mismatching a nuclease with a non-homologous donor to see if this occurs with our methodology. While ~20% cells that received matched *HBB* nuclease and *HBB* donor maintained GFP expression following 18 days in culture, *IL2RG* nuclease combined with *HBB* donor resulted in 0.8% GFP^+ cells (Extended Data Fig 1E-G), which was not significantly higher than the 0.5% GFP^+ cells observed when using the *HBB* AAV donor alone. These results demonstrate that end-capture of the *HBB* donor is an infrequent event using this system in human HSPCs. Furthermore, the observed GFP expression in 0.5% of HSPCs with AAV donor alone suggests that random integration of rAAV6 is limited in human HSPCs. In fact, previous reports have shown AAV-mediated targeted integration without a double-strand break27, 28 and thus, these GFP expressing cells may be the result of on-target events.

SCD is caused by a single nucleotide mutation (adenine to thymine), which changes an amino acid (E6V) at codon 6 of the *HBB* gene29. We created a 4.5kb rAAV6 donor template that would introduce the E6V mutation along with six other silent SNPs to prevent Cas9 re-

cutting of HR alleles(Extended Data Fig 2A-C). Using this rAAV6 donor with Cas9 RNP delivery we measured an average allelic modification frequency of 19% in six different HSPC donors (Fig 1d). These results confirm that the combined use of CRISPR with rAAV6 can precisely change the nucleotide at the position of the mutation that causes SCD..

Early enrichment of *HBB*-targeted HSPCs

Since HSCs differentiate and progressively lose their long-term repopulating capacity upon culturing, the identification of gene-edited HSPCs early in the manufacturing process would be of great use. Accordingly, in experiments using a GFP-expressing rAAV6 donor we observed that while HSPCs receiving only rAAV6 donor expressed low levels of GFP, HSPCs that also received Cas9 RNP generated a population at Day 4 post-electroporation that expressed much higher GFP levels (Fig 2a, left panel). We hypothesized that this GFP^{high} population was enriched for *HBB*-targeted cells. We therefore sorted and cultured the GFP^{high} population as well as the GFP^{low} and GFP^{neg} populations. While sorted GFP^{low} and GFP^{neg} populations were respectively ~25% and ~1% GFP^{+} after 15-20 days in culture, the GFP^{high} population was >95% GFP^{+}, strongly suggesting that this population was indeed *HBB*-targeted (Fig 2a, right panel). In fact, linear regression showed that the percentage of GFP^{high} expressing HSPCs at Day 4 post-electroporation strongly correlates with the total percentage of GFP^{+} cells at Day 18 (Extended Data Fig 3).

To confirm that the GFP^{high} population was enriched for on-target integration, we used ‘In-Out PCR’ (one primer binding the *HBB* locus outside the region of the homology arm and the other binding the integrated insert) to determine integration frequencies and allelic distribution in methylcellulose clones derived from the GFP^{high} population (95 clones). A total of 92% of clones had a targeted integration, with 38% harboring biallelic integrations (Fig 2b, Extended Data Fig 4A-C, and Supplemental Figure 1A). We note, however, that this assay generates colonies from progenitor cells and that the biallelic integration rates could be different in HSCs. Nonetheless, these data highly suggest that the log-fold transgene expression shift following rAAV6 and RNP delivery is due to HR at the intended locus and that the shift allows FACS-based enrichment of *HBB*-targeted HSPCs.

While GFP is not a clinically relevant reporter gene, the truncated nerve growth factor receptor (tNGFR), in which the cytoplasmic intracellular signaling domain is removed, could be used to enrich for targeted HSPCs. tNGFR is expressed on the cell surface, thereby allowing antibody-mediated detection of gene marking and it has already been used in human clinical trials30–33. We examined if we could enrich *HBB*-targeted HSPCs using tNGFR magnetic bead-based separation technology. HSPCs that received RNP and rAAV6 donor (with a tNGFR expression cassette) produced a $tNGFR^{high}$ population that was not present in cells transduced with rAAV6 alone, consistent with our findings with a GFP cassette (Fig 2c, left panel). We then enriched $tNGFR^{high}$ cells using anti-NGFR magnetic microbeads and after 18 days in culture, an average of 84% of HSPCs were $tNGFR^{+}$ (Fig 2c, right panel). We then performed ‘In/Out PCR’ on $tNGFR^{high}$ methylcellulose clones to determine on-target integration frequencies, and found that 86% of clones had a targeted integration with 30% having biallelic integrations (Fig 2d, Extended Data Fig 4D-F, and Supplemental Figure 1B).

We next evaluated progenitor cell capacity of GFP^{high} HSPCs using the colony forming unit (CFU) assay. *HBB* GFP^{high} HSPCs formed all types of colonies (erythroid, granulocyte/macrophage, and multi-lineage) to a comparable extent as Mock electroporated (Extended Data Figure 5A-C). We also evaluated frequencies of GFP^{high} cells in subpopulations of HSPC16, 35, 36 by immunophenotypic analysis, and observed a significant negative correlation between targeting frequencies and the immunophenotypic primitiveness of the analyzed population (Extended Data Figure 5D and Supplemental Figure 2). To confirm these findings, we employed another strategy to evaluate targeting rates in the primitive HSC population. HSPCs or HSCs were sorted from fresh cord blood, then subjected to HR experiments and we observed a 38% reduction in targeting efficiencies in the HSCs compared to the heterogeneous HSPC population (Extended Data Figure 5E). We next tested whether inefficient targeting of primitive cells could be due to reduced rAAV6 donor availability. HSCs and multi potent progenitors (MPPs) were transduced with scAAV6-SFFV-GFP, and results showed a 5-fold reduction in the number of GFP$^+$ HSCs and MPPs compared to the bulk CD34$^+$ population (Extended Data Fig 5F). Collectively, our data suggest that while HSCs are more difficult to target than progenitor cells, we are achieving HR-editing frequencies above 4% and usually above 10%.

HBB targeting in long-term repopulating HSCs

The current gold standard for HSC function, defined by the capacity to self-renew and form differentiated blood cells, is *in vivo* engraftment into immunodeficient non-obese diabetic (NOD)-SCID-gamma (NSG) mice. We used HSPCs derived from mobilized peripheral blood (mPB) for such studies because of their high clinical relevance, although these cells have been shown to have reduced engraftment capacity in NSG mice compared to HSPCs derived from fetal liver, cord blood, and bone marrow22, 38. All transplanted mice displayed human engraftment in the bone marrow as measured by the presence of hCD45/HLA-ABC double positive cells 16 weeks post-transplant (Fig 3a and Extended Data Fig 6A-C). While we observed a decrease in human cell chimerism for all treatment groups compared to Mock, all groups with nuclease-treated cells displayed similar chimerism to the rAAV6 only group. We did measure a small, but not statistically significant, decrease for the RNP+AAV GFP^{high} group compared to RNP+AAV, which can be explained by transplantation of fewer total cells and fewer phenotypically identified LT-HSCs (Extended Data Fig 5D). There was a significant decrease from RNP+AAV input targeting frequencies (16% in the CD34$^+$ mPB HSPCs, Extended Data Fig 5D) compared to the percent GFP$^+$ cells in the bone marrow at week 16 following transplantation (3.5%) (Fig 3c). This decreases is consistent with previous publications, and immunophenotyping of input cells did in fact show an average of 4% targeting in the CD34$^+$/CD38$^-$/CD90$^+$/CD45RA$^-$ population (Extended Data Fig 5D). Despite these observed reductions *in vivo*, our median rates of *HBB* gene targeting in human cells in the BM (3.5%) appears to be higher than what was found by others, particularly Hoban *et al.*, and Genovese *et al.*, in which most mice appeared to have <1% gene-modification following transplant using ZFNs targeting *HBB* and *IL2RG*, respectively6, 21. In contrast, mice transplanted with RNP GFP^{high} cells had a median of 90% GFP$^+$ human cells at Week 16 after transplant, with three mice showing >97% GFP$^+$ human cells (Fig 3a and 3c). We also observed both myeloid (CD33$^+$) and lymphoid (CD19$^+$) reconstitution with

a median of 94% and 83% GFP^+ cells, respectively (Fig 3c), implicating targeting of multipotent HSCs. We detected 5% and 49% GFP^+ human HSPCs ($CD34^+/CD10^-$) in the BM of mice transplanted with RNP+AAV and RNP+AAV GFP^{high} cells, respectively (Extended Data Fig 7A-C). Multi-parameter flow cytometric analysis showed no perturbations in lineage reconstitution and no evidence of abnormal hematopoiesis, a functional assessment of the safety of the editing procedure. To experimentally determine if we targeted *HBB* in HSCs, we performed secondary transplants for the RNP+AAV and RNP +AAV GFP^{high} groups. For both groups, we detected human cells in the bone marrow of secondary recipients at Week 12-14 after transplant with 7% and 90% GFP^+ cells for the RNP+AAV and RNP+AAV GFP^{high} group, respectively (Fig 3d). More importantly, we confirmed *HBB* on-target integration events in human cells sorted from the bone marrow of secondary recipients from GFP^{high} and RNP+AAV experimental groups (Fig 3e and Supplemental Figure 3). Collectively, these data confirm that our strategy targets the *HBB* gene in human HSCs.

We next scaled up the genome editing process to resemble a more clinically relevant HSPC starting cell number. We electroporated 80 million mPB-derived HSPCs with the *HBB* RNP system, transduced them with either SFFV-GFP or SFFV-tNGFR rAAV6, and then transplanted bulk RNP+AAV, sorted GFP^{high}, and $tNGFR^{high}$ (enriched by FACS or magnetic microbeads). At 16 weeks post-transplant, all mice displayed engraftment of edited human cells in the BM (Fig 3f). We note that using this large-scale procedure the human cell engraftment was equivalent to “Mock” in our prior experiment (Fig. 3b). While we observed reductions when comparing editing rates in the input cells to engrafted cells *in vivo*, the *HBB*-tNGFR mice showed a lower reduction (12% *in vitro vs.* 7.5% *in vivo*) than the *HBB*-GFP mice (10% *in vitro* vs. 1.9% *in vivo*), suggesting tNGFR could be a better transgene to evaluate editing of HSCs *in vivo* (Fig 3g). Furthermore, mice transplanted with enriched targeted HSPCs displayed human cell editing frequencies of 10-75% (three mice) with human engraftment levels ranging from 4-30% (Fig 3f-g). These data highly suggest our methodology can be translated to perform large-scale genome editing in HSCs at the *HBB* locus.

HBB gene correction in SCD HSPCs

We next tested our methodology to correct the disease-causing E6V mutation in SCD patient-derived $CD34^+$ HSPCs. We first confirmed high frequencies of INDELs (Fig 4a) and HR using an SFFV-GFP donor (Fig 4b) at the *HBB* locus in SCD HSPCs. We then produced a therapeutic rAAV6 donor (corrective SNP donor) designed to revert the E6V mutation, while also introducing silent mutations to prevent Cas9 re-cutting and premature strand cross over (Extended Data Fig 8A). Targeting SCD HSPCs with the corrective SNP donor reverted an average of 50% of the E6V (HbS) alleles to WT (HbA) alleles (Fig 4c), **and analysis of**methylcellulose clones showed that an average of 45% of clones had at least one HbA allele (Extended Data Fig 8B). We next created an anti-sickling *HBB* cDNA therapeutic donor (HbAS339) using previously reported strategies13, 14 of knocking in divergent cDNAs into the gene start codon to preserve endogenous promoter/enhancer function, followed by a clinically relevant promoter (EF1α) driving tNGFR expression to allow for enrichment of edited cells (Supplemental Fig 4A). Using this donor, we targeted an average

of 11% of SCD patient-derived HSPCs (Fig 4d) and confirmed seamless integration by In-Out PCR (Supplemental Figure 4B). Notably, we observed a tNGFR^{high} population as described previously (Fig. 2c), indicating the ability to enrich SCD-corrected HSPCs early in the culture process. Overall, these studies conclude that our methodology can correct the E6V mutation using two different donor designs.

We next tested if the *HBB*-edited SCD HSPCs maintained their erythroid differentiation potential bysubjecting tNGFR^{high} and Mock HSPCs to a 21-day *in vitro* erythroid differentiation protocol40, 41. Flow cytometric analyses post-erythroid differentiation showed a high proportion of GPA^{+}/CD45^{-}/CD71^{+}/CD34^{-} cells, indicating the presence of mature differentiated erythrocytes that express hemoglobin42 (Fig 4e and Extended Data Fig 9). Finally, to confirm that adult β-globin (HbA) or HbAS3 mRNA was transcribed from edited *HBB* alleles, we performed RT-qPCR on erythrocytes differentiated from edited SCD HSPCs. Erythrocytes edited with the corrective SNP donor expressed 56% HbA mRNA out of total β-globin mRNA, while erythrocytes edited with the cDNA donor (bulk) expressed 20% HbAS3 mRNA (Fig 4f). Thus the percentage of HbAS3 mRNA (20%) matched or exceeded the percentage of cells modified by the tNGFR cassette (11%) suggesting functional expression of the AS3 cDNA from the endogenous HBB promoter. Notably, erythrocytes differentiated from enriched tNGFR^{high} HSPCs expressed 70% HbAS3 mRNA, confirming an enrichment of functionally corrected HSPCs.

Discussion

Our data support the conclusion that HSCs are more resistant to targeted integration by HR than lineage-committed progenitor cells confirming previously published results21. Consequently, enrichment of targeted cells resulted in the removal of the majority of HSCs, leading to an overall 8-fold decrease in the total number of HSCs in the transplanted enriched population (Extended Data Fig 6C). Even though we transplanted 8-fold fewer total HSCs in the enriched population compared to the non-enriched (GFP^{high} vs. RNP+AAV), we found there was still an average of 5-fold more absolute number of edited cells in the bone marrow of the mice from the enriched group 16 weeks post-transplant (Extended Data Fig 6C). Hence, our enrichment strategy not only yields higher frequencies of modified cells in the transplanted mice, but the absolute number of modified human cells in the mice was also higher. Thus, the enrichment strategy can ameliorate the problem of inefficient HSC targeting. However, it is important to note that it remains to be determined by future studies in relevant animal models whether engraftment levels of *HBB*-modified cells would be clinically beneficial for the b-hemoglobinopathies. Nonetheless, recent advances in *ex vivo* HSC expansion protocols and identification of small molecule drugs, like UM17143 that expands HSCs, might be combined with our strategy to generate a large and highly enriched population of genome-edited HSCs. Future transplant studies may help determine if HSC expansion would be required for clinical translation of our enrichment paradigm.

Our methodology sets the framework for CRISPR-mediated *HBB* targeting in HSCs that has the power to be translated to the clinic. While GFP is an unsuitable marker for gene therapy, our enrichment protocol using tNGFR (Fig 2c and Fig 3f-g) (or other similar signaling-inert cell surface markers) represents a strategy for the next generation of β-hemoglobinopathy

therapies that are based on gene editing. Our studies show that this methodology can enrich corrected SCD patient-derived HSPCs that can differentiate into erythrocytes that express *HBB* anti-sickling mRNA from the endogenous *HBB* promoter. This tNGFR selection strategy has the potential advantage over chemoselection strategies that it avoids exposing edited cells and patients to potentially toxic chemotherapy14. The strategy of knocking in a *HBB* cDNA along with a selectable marker to enrich for modified cells would be applicable to both SCD and almost all forms of b-thalassemia. Furthermore, because we can efficiently scale up the genome editing process to clinically relevant starting numbers, future studies will address developing a GMP-compatible process for editing the *HBB* locus in HSPCs.

In conclusion, we believe that the presented methodology lays the foundation for CRISPR/Cas9-mediated genome editing therapies not only for the b-hemoglobinopathies, but also for a range of other hematological diseases and generally advances HSC-based cell and gene therapies.

Materials and Methods

AAV vector production

AAV vector plasmids were cloned in the pAAV-MCS plasmid (Agilent Technologies, Santa Clara, CA) containing ITRs from AAV serotype 2. The *HBB* rAAV6 GFP and tNGFR donor contained promoter, MaxGFP or tNGFR, and BGH polyA. The left and right homology arms for the GFP and tNGFR *HBB* donors were 540bp and 420bp, respectively. The E6V rAAV6 donor contained 2.2kb of sequence homologous to the sequence upstream of E6V. The nucleotide changes are depicted in Extended Data Figure 2. Immediately downstream of the last nucleotide change was 2.2kb of homologous *HBB* sequence. *HBB* cDNA contained same homology arms as GFP and tNGFR donors above except the left homology arm was shortened to end at the sickle mutation. Sequence of full *HBB* cDNA is depicted in Supplemental Figure 4B. The sickle corrective donor used in the SCD-derived HSPCs in Figure 4 had a total of 2.4kb sequence homology to *HBB* with the SNPs shown in Extended Data Figure 8a in the center. Self complementary AAV6 (scAAV6) carrying the SFFV promoter driving GFP was kindly provided by Hans-Peter Kiem (Fred Hutchinson Cancer Research Center). AAV6 vectors were produced as described with a few modifications44. Briefly, 293FT cells (Life Technologies, Carlsbad, CA, USA) were seeded at $13x10^6$ cells per dish in ten 15-cm dishes one day before transfection. One 15-cm dish was transfected using standard PEI transfection with 6 μg ITR-containing plasmid and 22 μg pDGM6 (a kind gift from David Russell, University of Washington, Seattle, WA, USA), which contains the AAV6 cap genes, AAV2 rep genes, and adenovirus helper genes. Cells were incubated for 72 hrs until harvest of AAV6 from cells by three freeze-thaw cycles followed by a 45 min incubation with TurboNuclease at 250 U/mL (Abnova, Heidelberg, Germany). AAV vectors were purified on an iodixanol density gradient by ultracentrifugation at 48,000 rpm for 2 h at 18 °C. AAV vectors were extracted at the 60-40% iodixanol interface and dialyzed three times in PBS with 5% sorbitol in the last dialysis using a 10K MWCO Slide-A-Lyzer G2 Dialysis Cassette (Thermo Fisher Scientific, Santa Clara, CA, USA). Vectors were added pluronic acid to a final concentration of 0.001%, aliquoted, and stored at -80°C until use.

Europe PMC Funders Author Manuscripts

AAV6 vectors were titered using quantitative PCR to measure number of vector genomes as described here45.

CD34+ hematopoietic stem and progenitor cells

Frozen $CD34^+$ HSPCs derived from bone marrow or mobilized peripheral blood were purchased from AllCells (Alameda, CA, USA) and thawed according to manufacturer's instructions. $CD34^+$ HSPCs from cord blood were either purchased frozen from AllCells or acquired from donors under informed consent via the Binns Program for Cord Blood Research at Stanford University and used fresh without freezing. CD34+ HSPCs from SCD patients were purified within 24 hours of the scheduled apheresis. For volume reduction via induced rouleaux formation, whole blood was added 6% Hetastarch in 0.9% Sodium Chloride Injection (Hospira, Inc., Lake Forest, IL, USA) in a proportion of 5:1 (v/v). Following a 60 to 90-minute incubation at room temperature, the top layer, enriched for HSPCs and mature leukocytes, was carefully isolated with minimal disruption of the underlying fraction. Cells were pelleted, combined, and resuspended in a volume of PBS with 2mM EDTA and 0.5% BGS directly proportional to the fraction of residual erythrocytes – typically 200-400 mL. MNCs were obtained by density gradient separation using Ficoll and $CD34^+$ HSPCs were purified using the $CD34^+$ Microbead Kit Ultrapure (Miltenyi Biotec, San Diego, CA, USA) according to manufacturer's protocol. Cells were cultured overnight and then stained for CD34 and CD45 using APC anti-human CD34 (Clone 561; Biolegend, San Jose, CA, USA) and BD Horizon V450 anti-human CD45 (Clone HI30; BD Biosciences, San Jose, CA, USA), and a pure population of HSPCs defined as $CD34^{bright}/CD45^{dim}$ were obtained by cell sorting on a FACS Aria II cell sorter (BD Biosciences, San Jose, CA, USA). All $CD34^+$ HSPCs were cultured in StemSpan SFEM II (Stemcell Technologies, Vancouver, Canada) supplemented with SCF (100 ng/ml), TPO (100 ng/ml), Flt3-Ligand (100 ng/ml), IL-6 (100 ng/ml), and StemRegenin1 (0.75 mM). Cells were cultured at 37°C, 5% CO_2, and 5% O_2.

Electroporation and transduction of cells

The *HBB* and *IL2RG* synthetic sgRNAs used were purchased from TriLink BioTechnologies (San Diego, CA, USA) with chemically modified nucleotides at the three terminal positions at both the 5′and 3′ends. Modified nucleotides contained 2′O-Methyl 3′phosphorothioate and the sgRNA was HPLC-purified. The genomic sgRNA target sequences with PAM in bold are: *HBB*: 5'-CTTGCCCCACAGGGCAGTAA**CGG**-3'46, 47; *IL2RG*: 5'-TGGTAATGATGGCTTCAACA**TGG**-3'. Cas9 mRNA containing 5-methylcytidine and pseudouridine was purchased from TriLink BioTechnologies. Cas9 protein was purchased from Life Technologies. Cas9 RNP was made by incubating protein with sgRNA at a molar ratio of 1:2.5 at 25°C for 10 min immediately prior to electroporation. $CD34^+$ HSPCs were electroporated 1-2 days after thawing or isolation. $CD34^+$ HSPCs were electroporated using the Lonza Nucleofector 2b (program U-014) and the Human T Cell Nucleofection Kit (VPA-1002, Lonza) as we have found this combination to be superior in optimization studies. The following conditions were used: $5x10^6$ cells/ml, 300 ug/ml Cas9 protein complexed with sgRNA at 1:2.5 molar ratio, or 100 μg/ml synthetic chemically modified sgRNA with 150 μg/ml Cas9 mRNA (TriLink BioTechnologies, non-HPLC purified). Following electroporation, cells were incubated for 15 min at 37°C after

which they were added AAV6 donor vectors at an MOI (vector genomes/cell) of 50,000-100,000 and then incubated at 30°C or 37°C overnight (if incubated at 30°C, plates were then transferred to 37°C) or targeting experiments of freshly sorted HSCs (Extended Data Figure 5E), cells were electroporated using the Lonza Nucleofector 4D (program EO-100) and the P3 Primary Cell Nucleofection Kit (V4XP-3024). For the electroporation of 80 million $CD34^+$ HSPCs, the Lonza 4D-Nucleofector LV unit (program DZ-100) and P3 Primary Cell Kit were used. Subsequently, we have found no benefit to the 30°C incubation and now perform all of our manufacturing at 37°C.

Measuring targeted integration of fluorescent and tNGFR donors

Rates of targeted integration of GFP and tNGFR donors was measured by flow cytometry at least 18 days after electroporation. Targeted integration of a tNGFR expression cassette was measured by flow cytometry of cells stained with APC-conjugated anti-human CD271 (NGFR) antibody (BioLegend, clone: ME20.4). For sorting of GFP^{high} or $tNGFR^{high}$ populations, cells were sorted on a FACS Aria II SORP using the LIVE/DEAD Fixable Blue Dead Cell Stain Kit (Life Technologies) to discriminate live and dead cells according to manufacturer's instructions.

Positive selection and enrichment of tNGFR+ HSPCs

Positive selection of targeted HSPCs was performed using the CD271 (tNGFR) Microbead Kit (Miltenyi Biotech, Auburn, CA USA), according to the manufacturers' instructions 72hr post electroporation. Briefly, $tNGFR^+$ cells were magnetically labeled with CD271 Microbeads after which the cell suspension was loaded onto an equilibrated MACS column inserted in the magnetic field of a MACS separator. The columns were washed three times, and enriched cells were eluted by removing the column from the magnetic field and eluting with PBS. Enrichment was determined by flow cytometry during culture for 2-3 weeks by FACS analysis every 3 days.

Immunophenotyping of targeted HSPCs

Harvested wells were stained with LIVE/DEAD Fixable Blue Dead Cell Stain (Life Technologies) and then with anti human CD34 PE-Cy7 (581, BioLegend), CD38 Alexa Fluor 647 (AT1, Santa Cruz Biotechnologies, Santa Cruz, CA, USA), CD45RA BV 421 (HI100, BD Biosciences), and CD90 BV605 (5E10, BioLegend) and analyzed by flow cytometry. For sorting of $CD34^+$ or $CD34^+/CD38^-/CD90^+$ cells, CB-derived $CD34^+$ HSPCs were stained directly after isolation from blood with anti human CD34 FITC (8G12, BD Biosciences), CD90 PE (5E10, BD Biosciences), CD38 APC (HIT2, BD Bioscience), and cells were sorted on a FACS Aria II (BD Bioscience), cultured overnight, and then electroporated with *HBB* RNP and transduced with *HBB* GFP rAAV6 using our optimized parameters.

Measuring targeted integration of the E6V donor

For assessing the allele modification frequencies in samples with targeted integration of the E6V rAAV6 donor, PCR amplicons spanning the targeted region (see Extended Data Figure 2A) were created using one primer outside the donor homology arm and one inside:

*HBB*_outside 5'-GGTGACAATTTCTGCCAATCAGG -3' and *HBB*_inside: 5'-GAATGGTAGCTGGATTGTAGCTGC-3'. The PCR product was gel-purified and re-amplified using a nested primer set (*HBB*_nested_fw: 5'-GAAGATATGCTTAGAACCGAGG-3' and *HBB*_nested_rw: 5'-CCACATGCCCAGTTTCTATTGG-3') to create a 685bp PCR amplicon (see Extended Data Figure 2A) that was gel-purified and cloned into a TOPO plasmid using the Zero Blunt TOPO PCR Cloning Kit (Life Technologies) according to the manufacturer's protocol. TOPO reactions were transformed into XL-1 Blue competent cells, plated on kanamycin-containing agar plates, and single colonies were sequenced by McLab (South San Francisco, CA, USA) by rolling circle amplification followed by sequencing using the following primer: 5′-GAAGATATGCTTAGAACCGAGG-3′. For each of the six unique CD34+ donors used in this experiment, 100 colonies were sequenced. Additionally, 100 colonies derived from an AAV only sample were sequenced and detected no integration events.

Measuring INDEL frequencies

INDEL frequencies were quantified using the TIDE software48 (Tracking of Indels by Decomposition) and sequenced PCR-products obtained by PCR of genomic DNA extracted at least four days following electroporation as previously described15.

Methylcellulose colony-forming unit (CFU) assay

The CFU assay was performed by FACS sorting of single cells into 96-well plates containing MethoCult Optimum (Stemcell Technologies) four days after electroporation and transduction. After 12-16 days, colonies were counted and scored based on their morphological appearance in a blinded fashion.

Genotyping of Methylcellulose colonies

DNA was extracted from colonies formed in methylcellulose from FACS sorting of single cells into 96-well plates. Briefly, PBS was added to wells with colonies, and the contents were mixed and transferred to a U-bottomed 96-well plate. Cells were pelleted by centrifugation at 300xg for 5min followed by a wash with PBS. Finally, cells were resuspended in 25 μl QuickExtract DNA Extraction Solution (Epicentre, Madison, WI, USA) and transferred to PCR plates, which were incubated at 65°C for 10 min followed by 100°C for 2 min. Integrated or non-integrated alleles were detected by PCR. For detecting *HBB* GFP integrations at the 3' end, two different PCRs were set up to detect integrated (one primer in insert and one primer outside right homology arm) and non-integrated alleles (primer in each homology arm), respectively (see Extended Data Figure 4A). *HBB*_int_fw: 5'-GTACCAGCACGCCTTCAAGACC-3', *HBB*_int_rv: 5'-GATCCTGAGACTTCCACACTGATGC-3', *HBB*_no_int_fw: 5'-GAAGATATGCTTAGAACCGAGG-3', *HBB*_no_int_rv: 5'-CCACATGCCCAGTTTCTATTGG-3'. For detecting *HBB* tNGFR integrations at the 5' end, a 3-primer PCR methodology was used to detect the integrated and non-integrated allele simultaneously (see Extended Data Figure 4D). *HBB*_outside_5'Arm_fw: GAAGATATGCTTAGAACCGAGG, SFFV_rev: ACCGCAGATATCCTGTTTGG *HBB*_inside_3'Arm_rev: CCACATGCCCAGTTTCTATTGG. Note that for the primers assessing non-integrated alleles, the Cas9 cut site is at least 90bp away from the primer

binding sites (PBSs) and since CRISPR/Cas9 generally introduces INDELs of small sizes, the PBSs should only very rarely be disrupted by an INDEL.

Transplantation of CD34$^+$ HSPCs into NSG mice

For *in vivo* studies, 6 to 8 week-old NOD *scid* gamma (NSG) mice were purchased from the Jackson laboratory (Bar Harbor, ME USA). The experimental protocol was approved by Stanford University's Administrative Panel on Lab Animal Care. Sample sizes were not chosen to ensure adequate power to detect a pre-specified effect size. Four days after electroporation/transduction or directly after sorting, 500,000 cells (or 100,000-500,000 cells for the GFP^{high} group) were administered by tail-vein injection into the mice after sub-lethal irradiation (200 cGy) using an insulin Syringe with a 27 gauge x 1/2" needle. Mice were randomly assigned to each experimental group and evaluated in a blinded fashion. For secondary transplants, human cells from the RNP+AAV group were pooled and CD34$^+$ cells were selected using a CD34 bead enrichment kit (MACS CD34 MicroBead Kit UltraPure, human, Miltenyi Biotec), and finally cells were injected into the femurs of female secondary recipients (3 mice total). Because GFP^{high} mice had low engraftment, they were not CD34$^+$-selected, but total mononuclear cells were filtered, pooled, and finally injected into the femur of two secondary recipients.

Assessment of human engraftment

At Week 16 post transplantation, mice were sacrificed, total mouse bone marrow (BM) (2x femur, 2x tibia, 2x humerus, sternum, 2x pelvis, spine) was collected and crushed using mortar and pestle. Mononuclear cells (MNCs) were enriched using Ficoll gradient centrifugation (Ficoll-Paque Plus, GE Healthcare, Sunnyvale, CA, USA) for 25min at 2,000xg, room temperature. Cells were blocked for nonspecific antibody binding (10% vol/vol, TruStain FcX, BioLegend) and stained (30min, 4°C, dark) with monoclonal anti human CD45 V450 (HI30, BD Biosciences), CD19 APC (HIB19, BD Biosciences), CD33 PE (WM53, BD Biosciences), HLA-ABC APC-Cy7 (W6/32, BioLegend), anti mouse CD45.1 PE-Cy7 (A20, eBioScience, San Diego, CA, USA), anti mouse PE-Cy5 mTer119 (TER-119, eBioscience) antibodies. Normal multi-lineage engraftment was defined by the presence of myeloid cells (CD33$^+$) and B-cells (CD19$^+$) within engrafted human CD45$^+$/ HLA-ABC$^+$ cells. Parts of the mouse BM were used for CD34-enrichment (MACS CD34 MicroBead Kit UltraPure, human, Miltenyi Biotec) and the presence of human hematopoietic stem and progenitor cells (HSPCs) was assessed by staining with anti human CD34 APC (8G12, BD Biosciences), CD38 PE-Cy7 (HB7, BD Biosciences), CD10 APC-Cy7 (HI10a, BioLegend), and anti mouse CD45.1 PE-Cy5 (A20, eBioScience) and analyzed by flow cytometry. The estimation of the total number of modified human cells in the bone marrow (BM) at week 16 post-transplant was calculated by multiplying percent engraftment with percent GFP$^+$ cells among engrafted cells. This number was multiplied by the total number of MNCs in the BM of a NSG mouse ($1.1x10^8$ per mouse) to give the total number of GFP$^+$ human cells in the total BM of the transplanted mice. The total number of MNCs in the BM of a NSG mouse was calculated by counting the total number of MNCs in one femur in four NSG mice. The total number of MNCs in one mouse was then calculated assuming one femur is 6.1% of the total marrow as found in49.

Differentiation of CD34+ HSPCs into erythrocytes *in vitro*

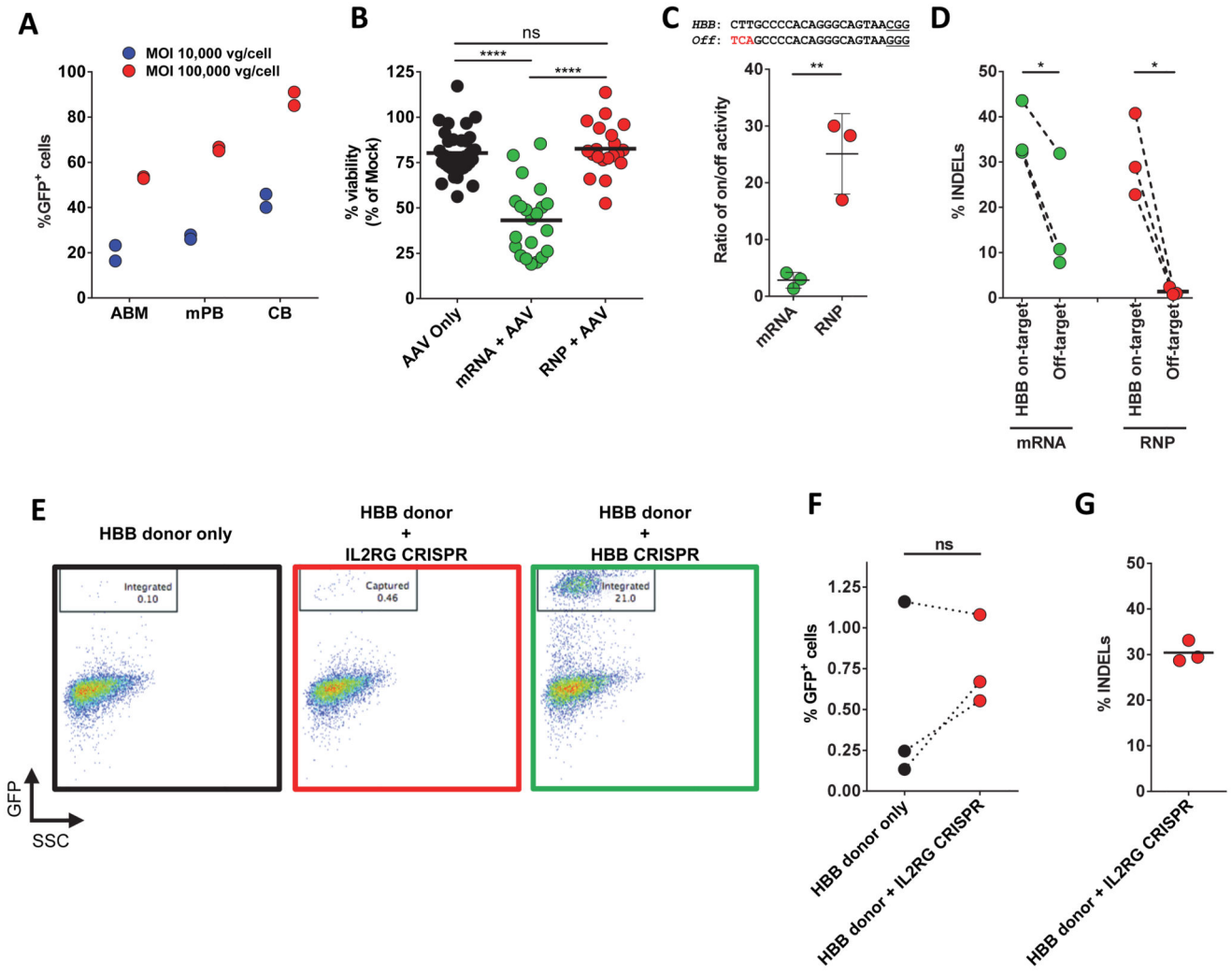
SCD patient-derived HSPCs were cultured in three phases following targeting at 37°C and 5% CO_2 in SFEM II media according to previously established protocols40, 41. Media was supplemented with 100 U/mL penicillin/streptomycin, 2mM L glutamine, 40 μg/mL lipids, 100 ng/mL SCF, 10 ng/mL IL-3 (PeproTech), 0.5 U/mL erythropoietin (eBiosciences), and 200 μg/mL transferrin (Sigma Aldrich). In the first phase, corresponding to days 0-7 (Day 0 being Day 4 post-electroporation), cells were cultured at 10^5 cells/mL. In the second phase, corresponding to days 7-11, cells were maintained at 10^5 cells/mL and erythropoietin was increased to 3 U/mL. In the third and final phase, days 11-21, cells were cultured at 10^6 cells/mL with 3 U/mL of erythropoietin and 1 mg/mL of transferrin. Erythrocyte differentiation of edited and non-edited HSPCs was assessed by flow cytometry using the following antibodies: hCD45 V450 (HI30, BD Biosciences), CD34 FITC (8G12, BD Biosciences), CD71 PE-Cy7 (OKT9, Affymetrix), and CD235a PE (GPA) (GA-R2, BD Biosciences).

Assessment of mRNA levels in differentiated erythrocytes

RNA was extracted from 100,000-250,000 differentiated erythrocytes between days 16-21 of erythroid differentiation using the RNeasy Mini Kit (Qiagen, Hilden, Germany) and was DNase-treated with RNase-Free DNase Set (Qiagen). cDNA was made from 100ng RNA using the iScript Reverse Transcription Supermix for RT-qPCR (Bio-Rad, Hercules, CA). Levels of HbS, HbA (from corrective SNP donor), and HbA-AA (anti-sickling *HBB* cDNA donor) were quantified by qPCR using the following primers and FAM/ZEN/IBFQ-labeled hydrolysis probes purchased as custom-designed PrimeTime qPCR Assays from IDT (San Jose, CA): HbS primer (fw): 5’ TCACTAGCAACCTCAAACAGAC 3’, HbS primer (rv): 5’ ATCCACGTTCACCTTGCC 3’, HbS probe: 5’ TAACGGCAGACTTCTCCACAGGAGTCA 3’, HbA primer (fw): 5’ TCACTAGCAACCTCAAACAGAC 3’, HbA primer (rv): 5’ ATCCACGTTCACCTTGCC 3’, HbA probe: 5’ TGACTGCGGATTTTTCCTCAGGAGTCA 3’, HbAS3 primer fw: 5’ GTGTATCCCTGGACACAAAGAT 3’, HbAS3 primer (rv): 5’ GGGCTTTGACTTTGGGATTTC 3’, HbAS3 probe: 5’ TTCGAAAGCTTCGGCGACCTCA 3’. Primers for HbA and HbS are identical, but probes differ by six nucleotides, and therefore it was experimentally confirmed that these two assays do not cross-react with targets. To normalize for RNA input, levels of the reference gene RPLP0 was determined in each sample using IDT’s predesigned RPLP0 assay (Hs.PT. 58.20222060). qPCR reactions were carried out on a LightCycler 480 II (Roche, Basel, Schwitzerland) using the SsoAdvanced Universal Probes Supermix (BioRad) following manufacturer’s protocol and PCR conditions of 10 minutes at 95°C, 50 cycles of 15 seconds at 95°C and 60 seconds at 58°C. Relative mRNA levels were determined using the relative standard curve method, in which a standard curve for each gene was made from serial dilutions of the cDNA. The standard curve was used to calculate relative amounts of target mRNA in the samples relative to levels of RPLP0.

Europe PMC Funders Author Manuscripts

Extended Data



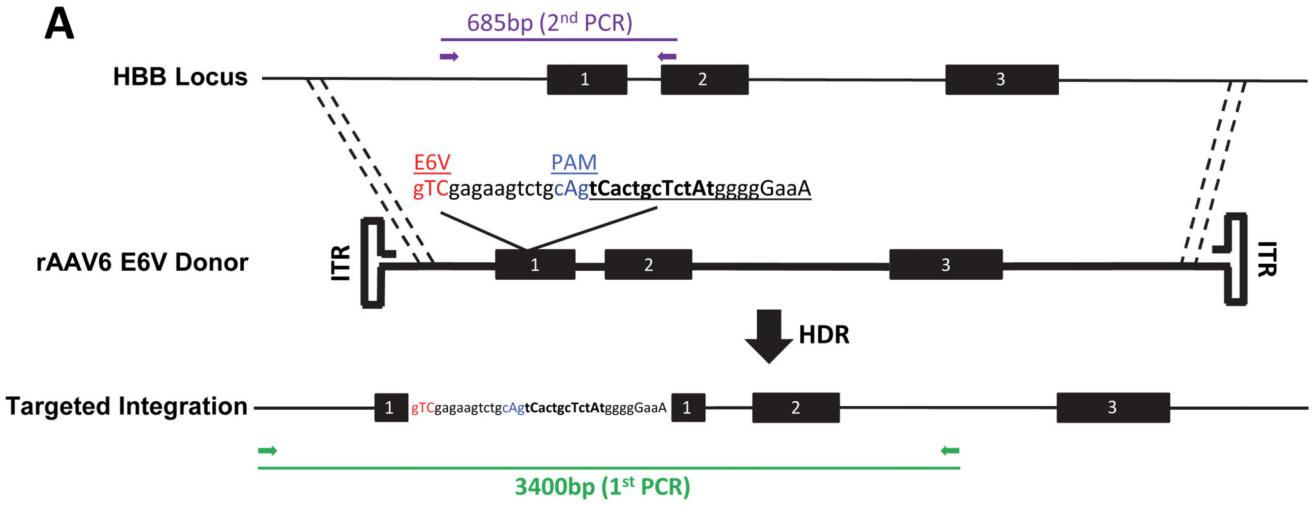
Extended Data Figure 1. High tropism of rAAV6 for CD34$^+$ HSPCs, and viability and specificity assessment of gene editing in CD34$^+$ HSPCs.

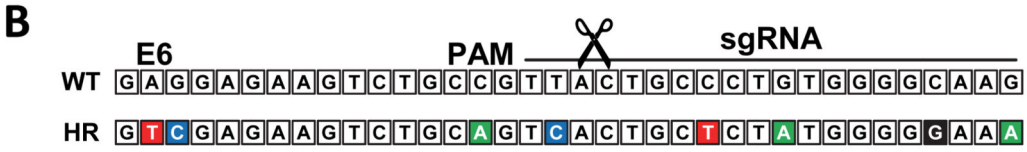
(A) CD34$^+$ HSPCs were transduced with a scAAV6 expressing GFP from an SFFV promoter at multiplicities of infections (MOIs) of 10,000 vg/cell or 100,000 vg/cell for 48hrs and then analyzed for percent GFP expression by flow cytometry using a non-transduced sample to set the GFP$^+$ gate at <0.1% GFP+ cells. scAAV was used because it eliminates second strand synthesis as a confounder of actual transduction. Results are from two independent experiments from at least two donors and error bars represent S.D. ABM: Adult Bone Marrow; mPB: Mobilized Peripheral Blood; CB: Cord Blood. **(B)** CD34$^+$ HSPCs were electroporated with the *HBB* CRISPR system (mRNA or RNP delivery) or without (AAV only), and then transduced with *HBB* rAAV6 donor vectors at an MOI of 100,000. Day 4 post electroporation, cells were analyzed by flow cytometry and live cells were gated in high forward scatter (FSC) and low side scatter (SSC). Percent cells in FSC/SSC gate is shown relative to that of Mock-electroporated cells. Each dot represents a unique CD34$^+$ HSPC

donor. **(C)** *Upper panel*: sgRNA target sequences at the *HBB* on-target site and a highly complementary off-target site (Chr9:101833584-101833606) are shown. PAM sequences are underlined and red sequence highlights the 3 mismatches of the off-target site. *Lower panel*: HSPCs were electroporated with either the "All RNA" or RNP-based CRISPR system, and 4 days post electroporation gDNA was extracted and analyzed for INDEL frequencies using TIDE at the on-target *HBB* and the off-target site. Results are graphed as a ratio of on to off-target activity highlighting the increased specificity of the RNP system. Averages from three different $CD34^+$ HSPC donors are shown and error bars represent S.E.M. ** $p < 0.01$, **** $p < 0.0001$, ns = $p \geq 0.05$, unpaired Student's t-test. **(D)** INDEL frequencies for the data presented in (C) * $p < 0.05$, paired Student's t-test. **(E)** Representative FACS plots showing stable GFP rates at Day 18 post-electroporation in donor-nuclease mismatch experiments. Mismatching nuclease and donor (red box) leads to infrequent end-capture events compared to on-target HR events observed with matched nuclease and homologous rAAV6 donor (green box). HSPCs were electroporated with 15μg Cas9 mRNA and either *HBB* MS sgRNA or *IL2RG* MS sgRNA, then transduced with *HBB*-GFP rAAV6 donor followed by 18 days of culture. **(F)** End-capture experiments were performed in three replicate experiments each in three unique CD34+ HSPC donors. ns = $p \geq 0.05$, paired Student's t-test. **(G)** Activity of the *IL2RG* CRISPR was confirmed by quantification of INDELs at the *IL2RG* target site using TIDE analysis

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts





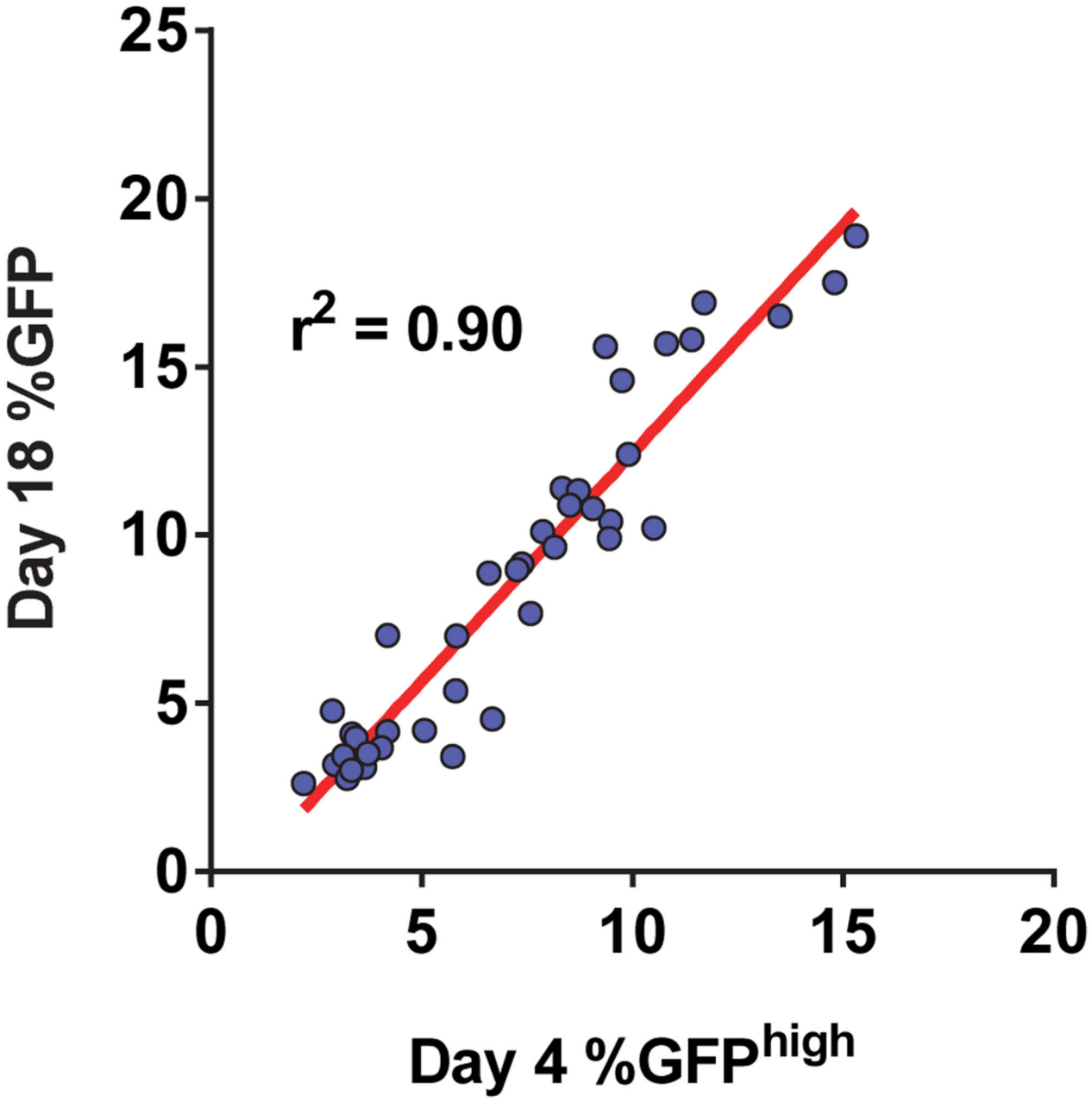
C

```
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTA---CCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -3
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCC---------CTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -9
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTA---CCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -3
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGT----GCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -4
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT---GCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -3
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGC--------CCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -8
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTA-------------CAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -13
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTA-TGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -1
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGT-ACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -1
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGT----------GGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -11
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCC---------CTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -9
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCT--------TGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -8
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTA--GCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -2
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTA------TGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -6
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCAA-------TGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -7
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT-CTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -1
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGC------TGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -6
TCAAACAGACACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGG-CTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT -2
```

Extended Data Figure 2. Schematic of targeting rAAV6 E6V homologous donor to the *HBB* locus.

(A) The human *HBB* locus on chromosome 11 is depicted at the top of the schematic and consists of three exons (black boxes) and two introns. The rAAV6 E6V donor includes the glutamic acid (E) to valine (V) mutation at codon 6, which is the amino acid change causing sickle cell disease. Other SNPs (all SNPs are capitalized) were introduced to PAM site (blue) and sgRNA binding site (bold) to prevent recutting following HR in HSPCs. To analyze targeted integration frequencies in HSPCs, a 2-step PCR was performed. First, a 3400bp In-Out PCR (green) was performed followed by a nested 685bp PCR (purple) on a gel-purified fragment from the first PCR. This 2nd PCR fragment was cloned into TOPO

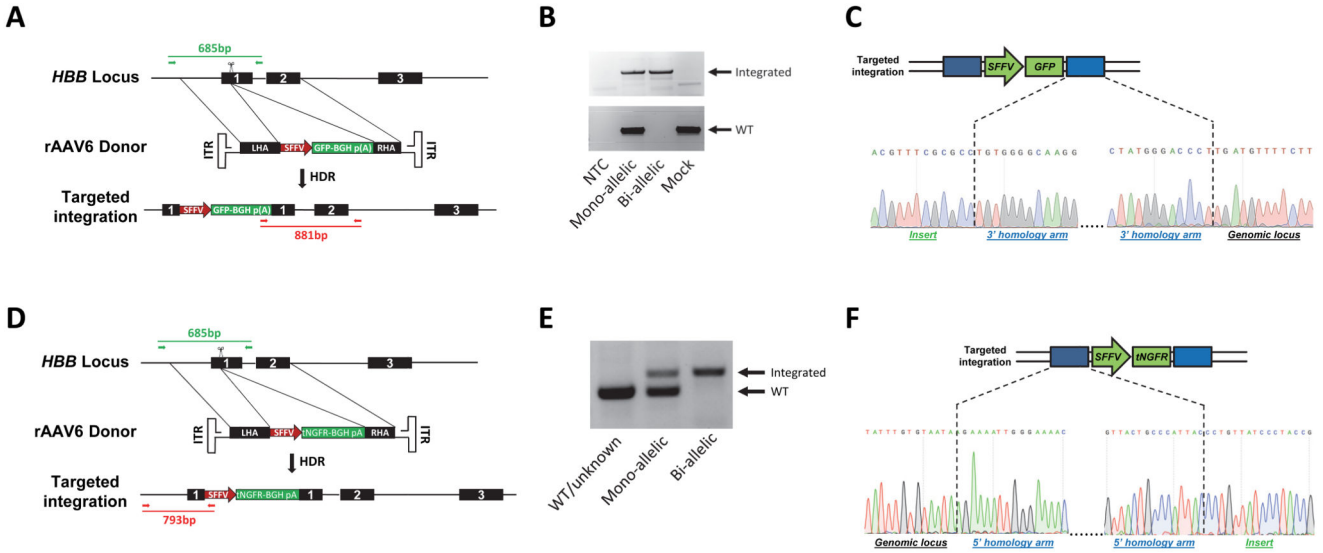
vectors, which were sequenced to determine the allele genotype (WT, INDEL, or HR). **(B)** The sequence of a wild-type *HBB* allele aligned with the sequence of an allele that has undergone HR. **(C)** Representative INDELs from the data represented in Figure 1d. The *HBB* reference sequence is shown in green.



Extended Data Figure 3. Linear regression model shows that the Day 4 GFP^{high} population is a reliable predictor of targeting frequencies.

Day 4 GFP^{high} percentages (x-axis) were plotted against Day 18 total GFP^{+} percentages (y-axis), and linear regression was performed. Data was generated from experiments including

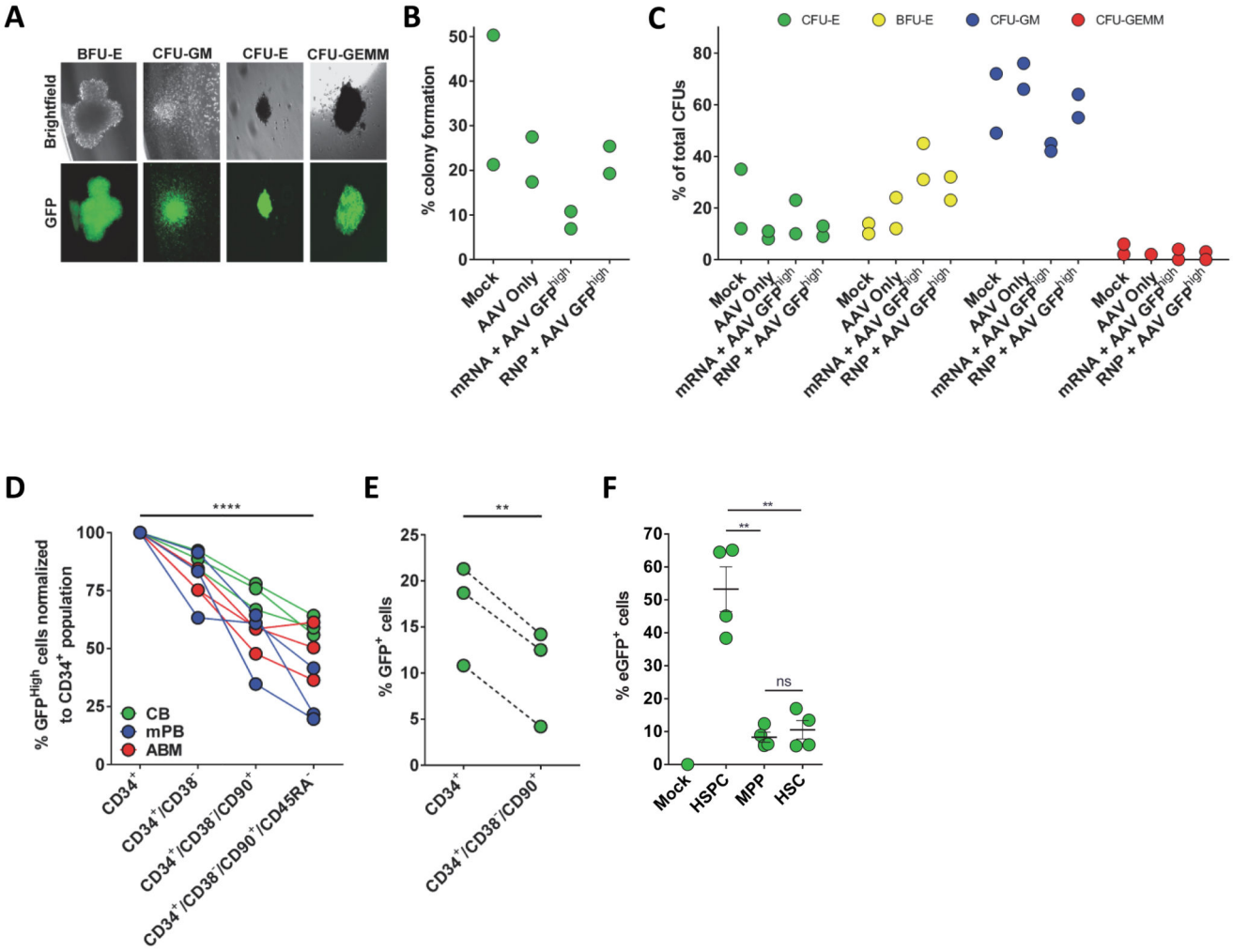
a total of 38 different CD34$^+$ HSPC donors, treated with either 15μg or 30μg Cas9 RNP to generate data points with a wider distribution of targeting frequencies.



Extended Data Figure 4. Overview of PCR genotyping of methylcellulose colonies with HR of the GFP and tNGFR donor at the *HBB* locus.

(A) The *HBB* locus was targeted by creating a DSB in exon 1 via Cas9 (scissors) and supplying a rAAV6 GFP donor template. Alleles with integrations were identified by PCR (red, 881bp) on methylcellulose-derived colonies using an In-Out primer set. Wildtype alleles were identified by PCR (green, 685bp) using primers flanking the sgRNA target site. **(B)** Representative genotyping PCRs showing mono- and bi-allelic clones as well as a clone derived from Mock-treated cells. NTC = non-template control (see Suppl. Fig. 1a for uncropped gel). **(C)** Representative Sanger sequence chromatograms for junctions between right homology arm (in blue) and insert (in green) or genomic locus (in white) highlighting seamless homologous recombination. **(D)** The *HBB* locus was targeted by creating a DSB in exon 1 via Cas9 (scissors) and supplying a rAAV6 tNGFR donor template. Genotypes were assessed by a 3-primer genotyping PCR on methylcellulose-derived colonies using an In-Out primer set (red, 793bp) and a primer set flanking the sgRNA target site (green, 685bp). Note that the two forward primers are the same. **(E)** Representative genotyping PCRs showing a WT/unknown, mono-, and bi-allelic clone (see Suppl. Fig. 1b for uncropped gel). **(F)** Representative Sanger sequence chromatograms for junctions between left homology arm (in blue) and insert (in green) or genomic locus (in white) highlighting seamless homologous recombination.

Europe PMC Funders Author Manuscripts

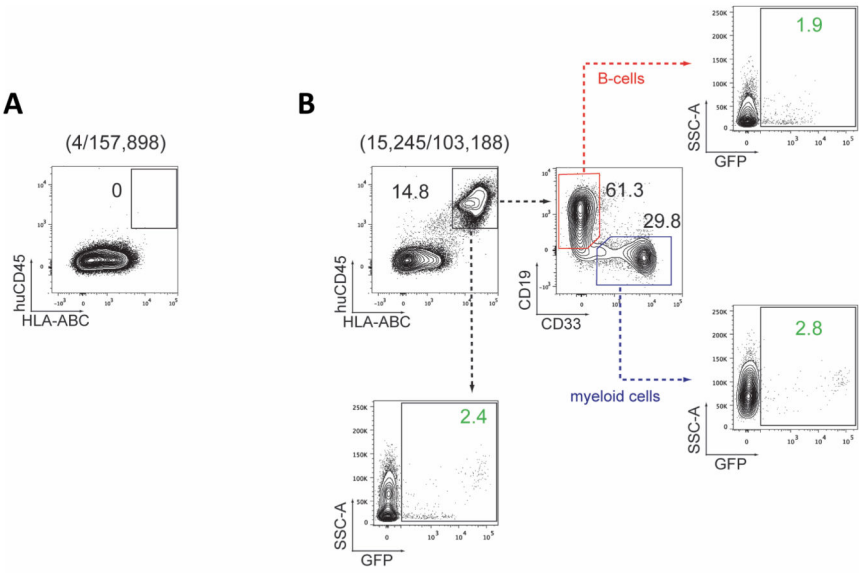


Extended Data Figure 5. Hematopoietic progenitor colony-forming unit (CFU) assay and targeting in different HSPC subpopulations.

(A) GFP^{high} HSPCs were single cell-sorted into 96-well plates containing methylcellulose. Representative images from fluorescence microscopy show lineage-restricted progenitors (BFU-E, CFU-E, CFU-GM) and multipotent progenitors (CFU-GEMM) with GFP expression. **(B)** Colony forming units (CFUs) derived as described above were counted 14 days post sort and shown relative to the total number of cells sorted (% cloning efficiency) (N=2 different HSPC donors). **(C)** Colonies from above were scored according to their morphology: 1) CFU-Erythroid (CFU-E), 2) Burst Forming Unit-Erythroid (BFU-E), 3) CFU-Granulocyte/Macrophage (CFU-GM), and 4) CFU-Granulocyte/Erythrocyte/Macrophage/Megakaryocyte (CFU-GEMM). (N=2 different HSPC donors) **(D)** 500,000 HSPCs isolated from mobilized peripheral blood (mPB), adult bone marrow (ABM), or cord blood (CB) were electroporated with RNP and transduced with GFP rAAV6 donor. At Day 4 post-electroporation, cells were phenotyped by flow cytometry for the cell surface markers CD34, CD38, CD90, and CD45RA (Supplemental Figure 2). Percent GFP^{high} cells in the indicated subpopulations are shown (data points represent unique donors, N=3 per HSPC source), **** $p < 0.0001$, paired Student's t-test. **(E)** CD34+ or CD34+/CD38-/CD90+ cells

Europe PMC Funders Author Manuscripts

were sorted directly from freshly isolated cord blood CD34+ HSPCs, cultured overnight, and then electroporated with RNP and transduced with GFP rAAV6. Bars show average percent GFP+ cells at Day 18 post-electroporation. Error bars represent S.E.M. (N=3 from different HSPC donors), ** $p < 0.01$, paired Student's t-test. **(F)** MPP ($CD34^+/CD38^-/CD90^-/CD45RA^-$) and HSC ($CD34^+/CD38^-/CD90^+/CD45RA^-$) populations were sorted from fresh cord-blood-derived CD34+ HSPCs and immediately after sorting, cells were transduced with scAAV6-SFFV-eGFP at an MOI of 100,000 vg/cell along the bulk HSPC population. scAAV6 was used because it eliminates second strand synthesis as a confounder of actual transduction, though the activity of the SFFV promoter may not be equivalent in each population thus potentially underestimating the degree of transduction of MPPs and HSCs. Two days later, transduction efficiencies were measured by flow cytometric analysis of eGFP expression using non-transduced cells (Mock) to set the GFP^+ gate. Error bars represent S.E.M., N=4, two different HSPC donors, ns = $p \geq 0.05$, ** $p < 0.01$, unpaired t test with Welch's correction.



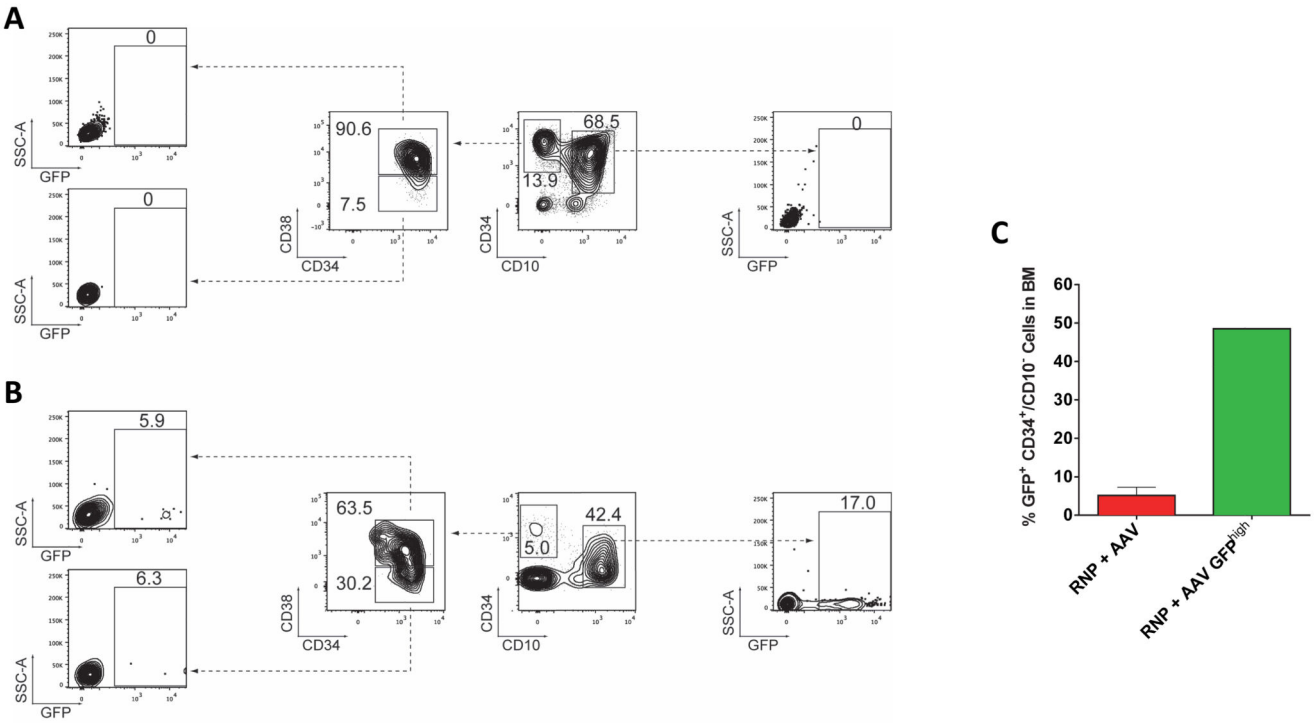
C

Sample Name	% engraftment (average per mouse)	% GFP^+ cells among human cells (average per mouse)	Total live cells injected per mouse (average per mouse)	Average total HSCs injected per mouse ($CD34^+/CD38^-/CD90^+/CD45RA^-$)	Estimated total modified human cells in BM per mouse at Week 16
RNP + AAV	6.99 (+/-1.8)	3.46 (+/-0.82)	500,000	24,000 (+/-1100)	270,000 (+/-33,000)
RNP + AAV GFP^{high}	1.67 (+/-1.0)	78.8 (+/-10.6)	350,000 100,000 x1 250,000 x2 500,000 x3	2,900 (+/-770)	1,460,000 (+/-390,000)

Extended Data Figure 6. Analysis of human engraftment.

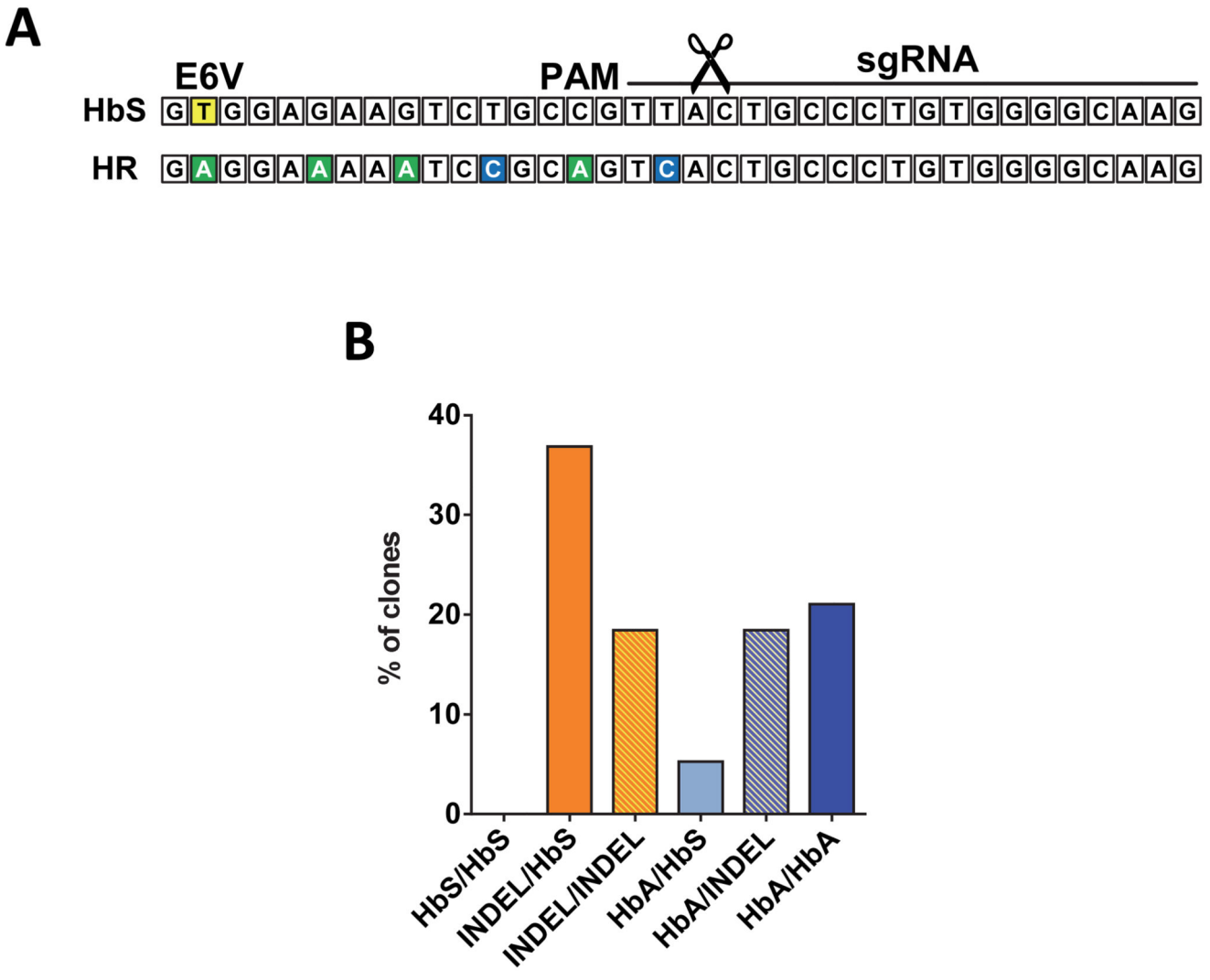
(A) Representative FACS plot from the analysis of the bone marrow of a control mouse not transplanted with human cells. Mice were sacrificed and bone marrow was harvested from femur, tibia, hips, humerus, sternum, and vertebrae. Cells were subject to Ficoll density gradient to isolate mononuclear cells, which were analyzed for human engraftment by flow

cytometry. Human engraftment was delineated as huCD45/HLA-ABC double positive. 4 out of 157,898 cells were found within the human cell gate. **(B)** Representative FACS plots showing gating scheme for analyses of NSG mice transplanted with human cells and analyzed as described in (A). Representative plots are from one mouse from the RNP + rAAV6 experimental group. As above, human engraftment was delineated as huCD45/HLA-ABC double positive. B cells were marked by CD19 expression, and myeloid cells identified by CD33 expression. GFP expression was analyzed in total human cells (2.4%), B-cells (1.9%) and myeloid cells (2.8%). The GFP brightness in B cells is lower than in myeloid cells suggesting that the SFFV promoter is not as active in the B-cell lineage compared to the myeloid lineage (see also Fig 4a). **(C)** Overview of engraftment for RNP+AAV and RNP +AAV GFP^{high} experimental groups. Average engraftment frequencies and percent GFP^{+} human cells +/-S.E.M are shown. Total number of cells transplanted was the same (500,000) for all mice in the RNP group, whereas in the GFP^{high} group, one mouse was transplanted with 100,000 cells, two mice with 250,000 cells, and three mice with 500,000 cells. The total number of HSCs transplanted per mouse (+/-S.E.M.) was calculated based on the frequencies of GFP^{+} cells in the CD34^{+}/CD38^{-}/CD90^{+}/CD45RA^{-} subset analyzed by flow cytometry (see Fig 3d) directly before injection. The total number of modified human cells in the bone marrow (BM) at Week 16 post transplant per mouse (+/-S.E.M.) was estimated based on calculations presented in the materials and methods. This shows that the enrichment not only resulted in a higher percentage of edited cells (column 3) but also resulted in an absolute higher number (column 6) of edited cells as well.



Extended Data Figure 7. Genome-edited human HSPCs in the bone marrow of NSG mice at Week 16 post transplantation.

Representative FACS plots from the analysis of NSG mice from the **(A)** Mock or **(B)** RNP +AAV experimental group at Week 16 post transplantation. Mice were sacrificed and bone marrow was harvested, PBMCs were isolated via Ficoll density gradient, after which human $CD34^+$ cells were enriched by magnetic-activated cell sorting (MACS), and finally cells were stained with antiCD34, antiCD38, and antiCD10 antibodies to identify human GFP^+ cells in the $CD34^+/CD10^-$ and $CD34^+/CD10^-/CD38^-$ populations (note that CD10 was included as a negative discriminator for immature B cells) **(C)** Collective data from the analysis of GFP^+ cells in the human $CD34^+/CD10^-$ population from the RNP+AAV (N=11) and RNP+AAV GFP^{high} (N=6) experimental groups. For the RNP+AAV GFP^{high} group, cells from all six mice were pooled before analysis and thus, no error bar is available. Error bar on RNP group represents S.E.M.

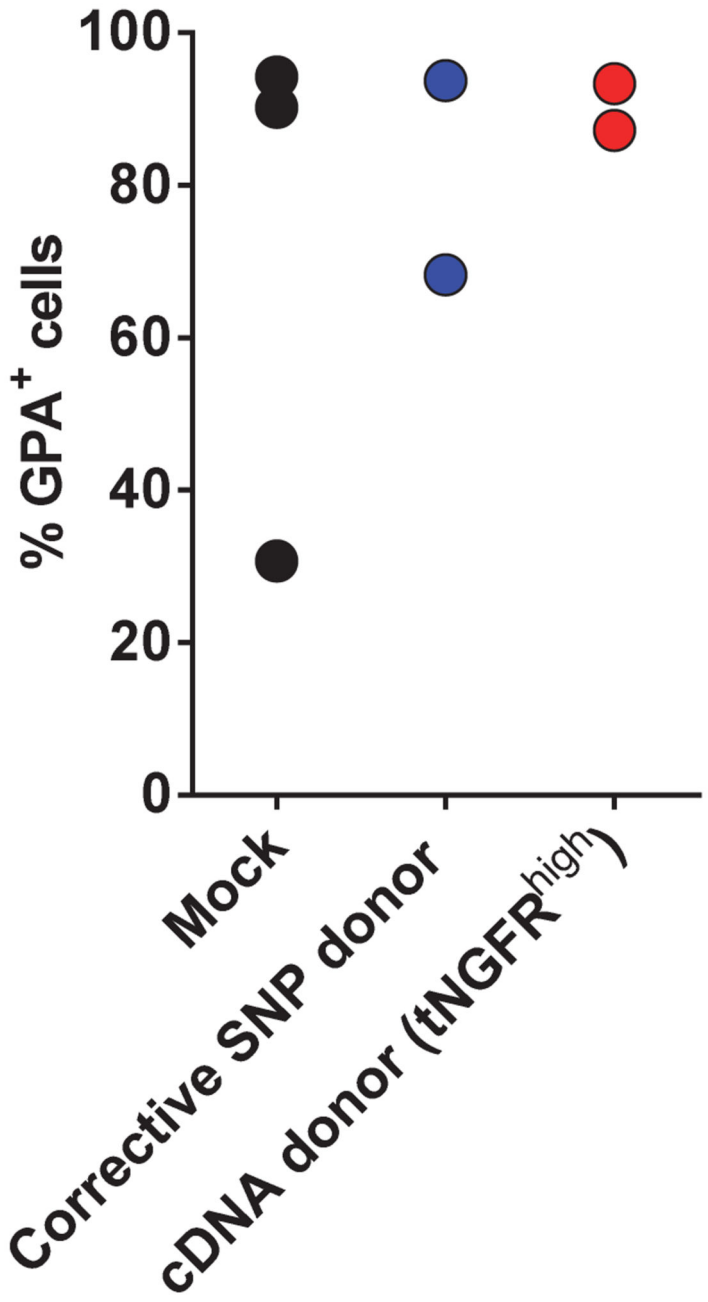


Extended Data Figure 8. Correction of the sickle cell mutation in patient-derived CD34+ HSPCs.
(A) Schematic overview of the sequence of the sickle allele aligned with the sequence of an allele that has undergone HR using the corrective SNP donor. The E6V mutation in sickle

cell patients (A->T) is highlighted in yellow. The sgRNA recognition sequence, the PAM site, and the cut site (scissors) are shown. The donor carries synonymous nucleotide changes between the sickle nucleotide and the cut site to avoid premature cross-over during HR. Synonymous changes are also added to the PAM and an early nucleotide in the sgRNA target site to avoid subsequent re-cutting and potential inactivation of the corrected allele. **(B)** HSPCs from two different sickle cell patients were targeted with the corrective SNP donor and seeded in methylcellulose. After 14 days, In-Out PCR amplicons from a total of 38 clones were sequenced and genotypes were extracted from sequence chromatograms.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts



Extended Data Figure 9. Edited HSPCs from sickle cell patients differentiate into erythrocytes that express Glycophorin A.

$CD34^+$ HSPCs derived from sickle cell patients were edited with HBB Cas9 RNP and either the corrective SNP donor or the cDNA donor. 4 days post-electroporation, cells edited with the cDNA donor were sorted for $tNGFR^+$ cells. This population as well as the populations edited with the corrective SNP donor and Mock-electroporated cells were subjected to a 21-day erythrocyte differentiation protocol, followed by staining for Glycophorin A (GPA). All data points within experimental groups are derived from experiments in cells from different sickle cell patients, N=3 (Mock) and N=2 (SNP and cDNA donor).

Europe PMC Funders Author Manuscripts

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

D.P.D was supported through the Stanford Child Health Research Institute (CHRI) Grant and Postdoctoral Award. R.O.B. was supported through an Individual Postdoctoral grant (DFF–1333-00106B) and a Sapere Aude, Research Talent grant (DFF–1331-00735B) both from the Danish Council for Independent Research, Medical Sciences. M.H.P. gratefully acknowledges the support of the Amon Carter Foundation, the Laurie Krauss Lacob Faculty Scholar Award in Pediatric Translational Research and NIH grant support PN2EY018244, R01-AI097320, and R01-AI120766. We thank David Russell (University of Washington) for the pDGM6 plasmid, Hans-Peter Kiem (Fred Hutchinson Cancer Research Center) for scAAV6, Gustavo de Alencastro and Mark Kay (Stanford University) for help with AAV production, the Binns Program for Cord Blood Research (Stanford University) for cord blood-derived $CD34^+$ HSPCs. We also thank Lonza for donating the LV unit for performing large-scale genome editing studies. We further thank members of the Porteus lab, David DiGiusto (Stanford University), and Maria Grazia Roncarolo (Stanford University) for helpful input, comments, and discussion.

References

1. Naldini L. Ex vivo gene transfer and correction for cell-based therapies. Nature reviews. Genetics. 2011; 12:301–315.
2. Porteus M. Genome Editing: A New Approach to Human Therapeutics. Annual review of pharmacology and toxicology. 2016; 56:163–190.
3. Smithies O, Gregg RG, Boggs SS, Koralewski MA, Kucherlapati RS. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. Nature. 1985; 317:230–234. [PubMed: 2995814]
4. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. Science. 2003; 300:763. [PubMed: 12730593]
5. Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America. 1994; 91:6064–6068. [PubMed: 8016116]
6. Hoban MD, et al. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. Blood. 2015; 125:2597–2604. [PubMed: 25733580]
7. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014; 157:1262–1278. [PubMed: 24906146]
8. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014; 346 1258096.
9. Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337:816–821. [PubMed: 22745249]
10. Kass EM, Jasin M. Collaboration and competition between DNA double-strand break repair pathways. FEBS letters. 2010; 584:3703–3708. [PubMed: 20691183]
11. Porteus MH. Towards a new era in medicine: therapeutic genome editing. Genome biology. 2015; 16:286. [PubMed: 26694713]
12. Woods NB, Bottero V, Schmidt M, von Kalle C, Verma IM. Gene therapy: therapeutic gene causing lymphoma. Nature. 2006; 440:1123. [PubMed: 16641981]
13. Hubbard N, et al. Targeted gene editing restores regulated CD40L expression and function in X-HIGM T cells. Blood. 2016
14. Voit RA, Hendel A, Pruett-Miller SM, Porteus MH. Nuclease-mediated gene editing by homologous recombination of the human globin locus. Nucleic acids research. 2014; 42:1365–1378. [PubMed: 24157834]
15. Hendel A, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. Nature biotechnology. 2015; 33:985–989.

16. Baum CM, Weissman IL, Tsukamoto AS, Buckle AM, Peault B. Isolation of a candidate human hematopoietic stem-cell population. Proceedings of the National Academy of Sciences of the United States of America. 1992; 89:2804–2808. [PubMed: 1372992]
17. Mukherjee S, Thrasher AJ. Gene therapy for PIDs: progress, pitfalls and prospects. Gene. 2013; 525:174–181. [PubMed: 23566838]
18. Cavazzana-Calvo M, et al. Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. Nature. 2010; 467:318–322. [PubMed: 20844535]
19. Naldini L. Gene therapy returns to centre stage. Nature. 2015; 526:351–360. [PubMed: 26469046]
20. Jenq RR, van den Brink MR. Allogeneic haematopoietic stem cell transplantation: individualized stem cell and immune therapy of cancer. Nature reviews Cancer. 2010; 10:213–221. [PubMed: 20168320]
21. Genovese P, et al. Targeted genome editing in human repopulating haematopoietic stem cells. Nature. 2014; 510:235–240. [PubMed: 24870228]
22. Wang J, et al. Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. Nature biotechnology. 2015; 33:1256–1263.
23. De Ravin SS, et al. Targeted gene addition in human CD34(+) hematopoietic cells for correction of X-linked chronic granulomatous disease. Nature biotechnology. 2016; 34:424–429.
24. Sather BD, et al. Efficient modification of CCR5 in primary human hematopoietic cells using a megaTAL nuclease and AAV donor template. Science translational medicine. 2015; 7:307ra156.
25. Wang J, et al. Highly efficient homology-driven genome editing in human T cells by combining zinc-finger nuclease mRNA and AAV6 donor delivery. Nucleic acids research. 2016; 44:e30. [PubMed: 26527725]
26. Miller DG, Petek LM, Russell DW. Adeno-associated virus vectors integrate at chromosome breakage sites. Nature genetics. 2004; 36:767–773. [PubMed: 15208627]
27. Russell DW, Hirata RK. Human gene targeting by viral vectors. Nature genetics. 1998; 18:325–330. [PubMed: 9537413]
28. Barzel A, et al. Promoterless gene targeting without nucleases ameliorates haemophilia B in mice. Nature. 2015; 517:360–364. [PubMed: 25363772]
29. Hoban MD, Orkin SH, Bauer DE. Genetic treatment of a molecular disorder: gene therapy approaches to sickle cell disease. Blood. 2016
30. Bonini C, et al. HSV-TK gene transfer into donor lymphocytes for control of allogeneic graft-versus-leukemia. Science. 1997; 276:1719–1724. [PubMed: 9180086]
31. Ciceri F, et al. Infusion of suicide-gene-engineered donor lymphocytes after family haploidentical haemopoietic stem-cell transplantation for leukaemia (the TK007 trial): a non-randomised phase I-II study. The Lancet Oncology. 2009; 10:489–500. [PubMed: 19345145]
32. Oliveira G, et al. Tracking genetically engineered lymphocytes long-term reveals the dynamics of T cell immunological memory. Science translational medicine. 2015; 7:317ra198.
33. Bonini C, et al. Safety of retroviral gene marking with a truncated NGF receptor. Nature medicine. 2003; 9:367–369.
34. Seita J, Weissman IL. Hematopoietic stem cell: self-renewal versus differentiation. Wiley interdisciplinary reviews. Systems biology and medicine. 2010; 2:640–653. [PubMed: 20890962]
35. Majeti R, Park CY, Weissman IL. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell stem cell. 2007; 1:635–645. [PubMed: 18371405]
36. Doulatov S, Notta F, Laurenti E, Dick JE. Hematopoiesis: a human perspective. Cell stem cell. 2012; 10:120–136. [PubMed: 22305562]
37. Weidner CI, et al. Hematopoietic stem and progenitor cells acquire distinct DNA-hypermethylation during in vitro culture. Scientific reports. 2013; 3:3372. [PubMed: 24284763]
38. Gu A, et al. Engraftment and lineage potential of adult hematopoietic stem and progenitor cells is compromised following short-term culture in the presence of an aryl hydrocarbon receptor antagonist. Human gene therapy methods. 2014; 25:221–231. [PubMed: 25003230]
39. Levasseur DN, et al. A recombinant human hemoglobin with anti-sickling properties greater than fetal hemoglobin. The Journal of biological chemistry. 2004; 279:27518–27524. [PubMed: 15084588]

40. Dulmovits BM, et al. Pomalidomide reverses gamma-globin silencing through the transcriptional reprogramming of adult hematopoietic progenitors. Blood. 2016; 127:1481–1492. [PubMed: 26679864]
41. Hu J, et al. Isolation and functional characterization of human erythroblasts at distinct stages: implications for understanding of normal and disordered erythropoiesis in vivo. Blood. 2013; 121:3246–3253. [PubMed: 23422750]
42. Romero Z, et al. beta-globin gene transfer to human bone marrow for sickle cell disease. The Journal of clinical investigation. 2013
43. Fares I, et al. Cord blood expansion. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. Science. 2014; 345:1509–1512. [PubMed: 25237102]
44. Khan IF, Hirata RK, Russell DW. AAV-mediated gene targeting methods for human cells. Nature protocols. 2011; 6:482–501. [PubMed: 21455185]
45. 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. 2012; 23:18–28. [PubMed: 22428977]
46. Hendel A, et al. Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing. Cell reports. 2014; 7:293–305. [PubMed: 24685129]
47. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. Nucleic acids research. 2013; 41:9584–9592. [PubMed: 23939622]
48. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic acids research. 2014; 42:e168. [PubMed: 25300484]
49. Boggs DR. The total marrow mass of the mouse: a simplified method of measurement. American journal of hematology. 1984; 16:277–286. [PubMed: 6711557]

Europe PMC Funders Author Manuscripts

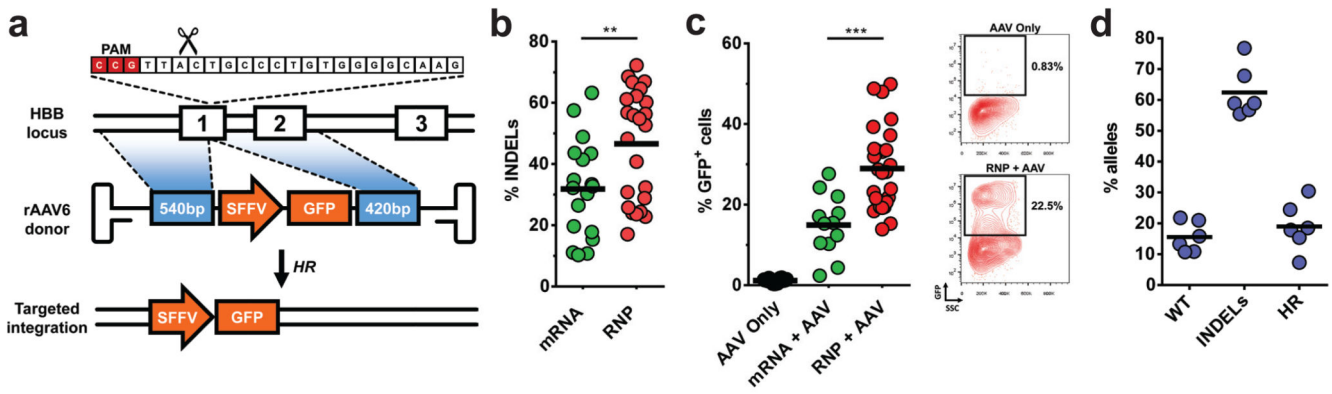


Fig 1. CRISPR/Cas9 and rAAV6-mediated targeted integration at the *HBB* locus in human CD34$^+$ hematopoietic stem and progenitor cells (HSPCs).

a) Schematic of targeted genome editing at the *HBB* locus using CRISPR/Cas9 and rAAV6. Site-specific double strand breaks (DSBs) are created by Cas9 (scissors) mainly between nucleotide 17-18 of the 20bp target site, which is followed by the 'NGG' PAM (red). A DSB stimulates homologous recombination (HR) using rAAV6 homologous donor as repair template. White boxes: *HBB* exons, blue boxes: homology arms, orange boxes: SFFV-GFP-polyA expression cassette **b)** HSPCs were electroporated with all RNA or RNP CRISPR system and INDELs were analyzed via TIDE software (N=number of data points within group, all from different mPB or CB donors). **c)** HSPCs electroporated as above and transduced with *HBB*-specific rAAV6s were analyzed by flow cytometry 18-21 days post-electroporation when GFP levels were found to be constant. Left panel shows percentage of HSPCs. Right panel shows representative FACS plots (N=number of data points within group, all from different donors). **d)** HSPCs were treated as above but targeted with rAAV6 E6V donor. Frequencies of allele types were quantified by sequencing of a total of 600 clones from TOPO-cloned In-Out PCRs (N=6, all from different CB or BM donors)

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

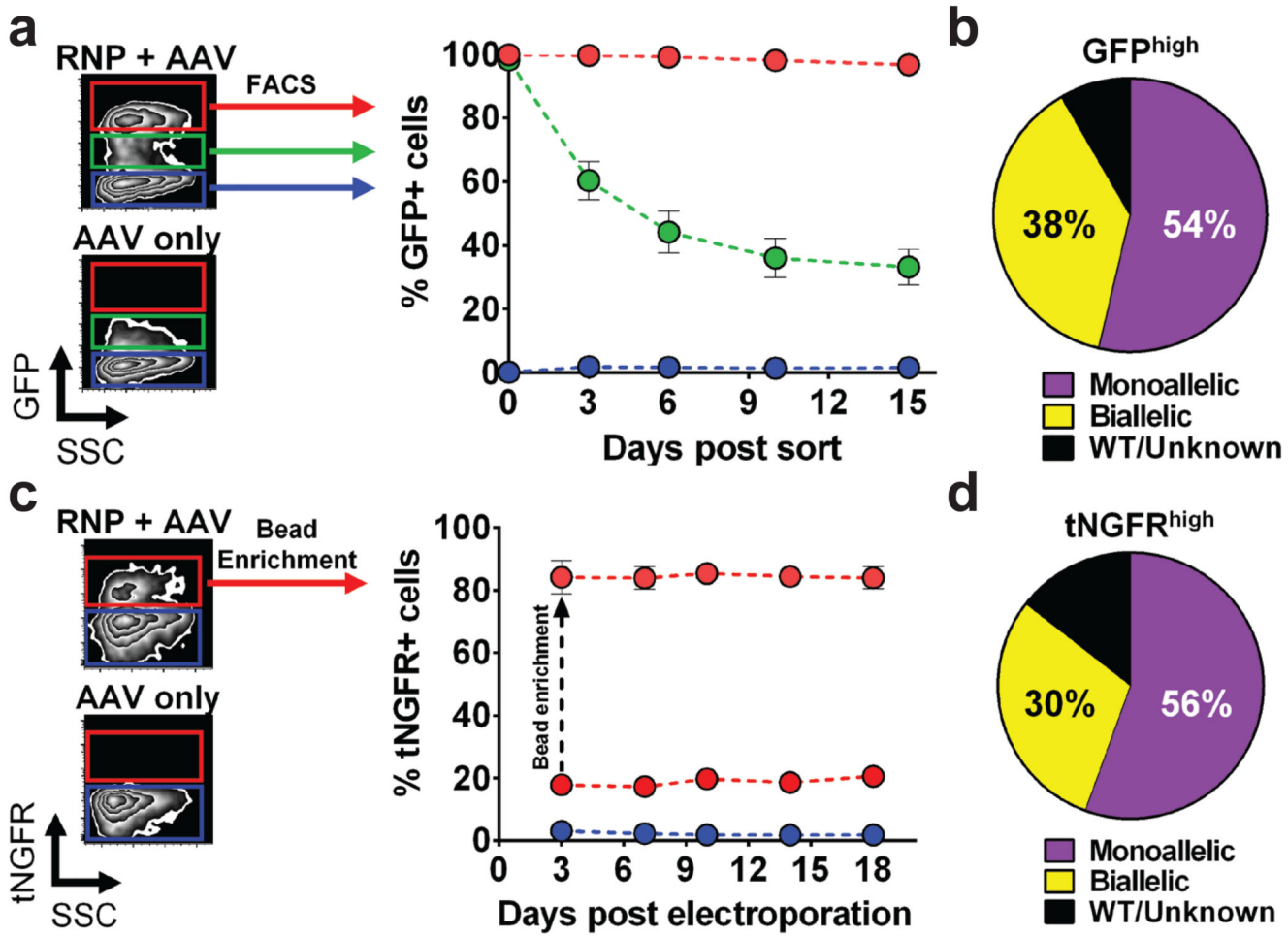


Fig 2. Enrichment of *HBB*-targeted HSPCs using FACS and magnetic bead-based technologies.

a) *Left panel*: representative FACS plots highlight the GFP^{high} population (red gate) generated by the addition of Cas9 RNP. *Right panel*: *HBB*-targeted HSPCs from GFP^{high} (red), GFP^{low} (green), and GFP^{neg} (blue) fractions were sorted and monitored for GFP expression. Error bars represent S.E.M. (N=11, all from unique mPB or CB donors) **b)** PCR was performed on methylcellulose colonies from GFP^{high} HSPCs to detect targeted integration at the 3' end.**c)** *Left panel*: representative FACS plots highlight the tNGFR^{high} population (red gate) generated by the addition of Cas9 RNP. *Right panel*: tNGFR^{high} (red) HSPCs were enriched using anti-CD271 (LNGFR) magnetic microbeads and cultured for 18 days while monitoring tNGFR expression. Error bars represent S.E.M. (N=5, all from unique CB donors). **d)** PCR was performed on tNGFR^{high}-derived methylcellulose colonies to detect targeted integrations at the 5' end.

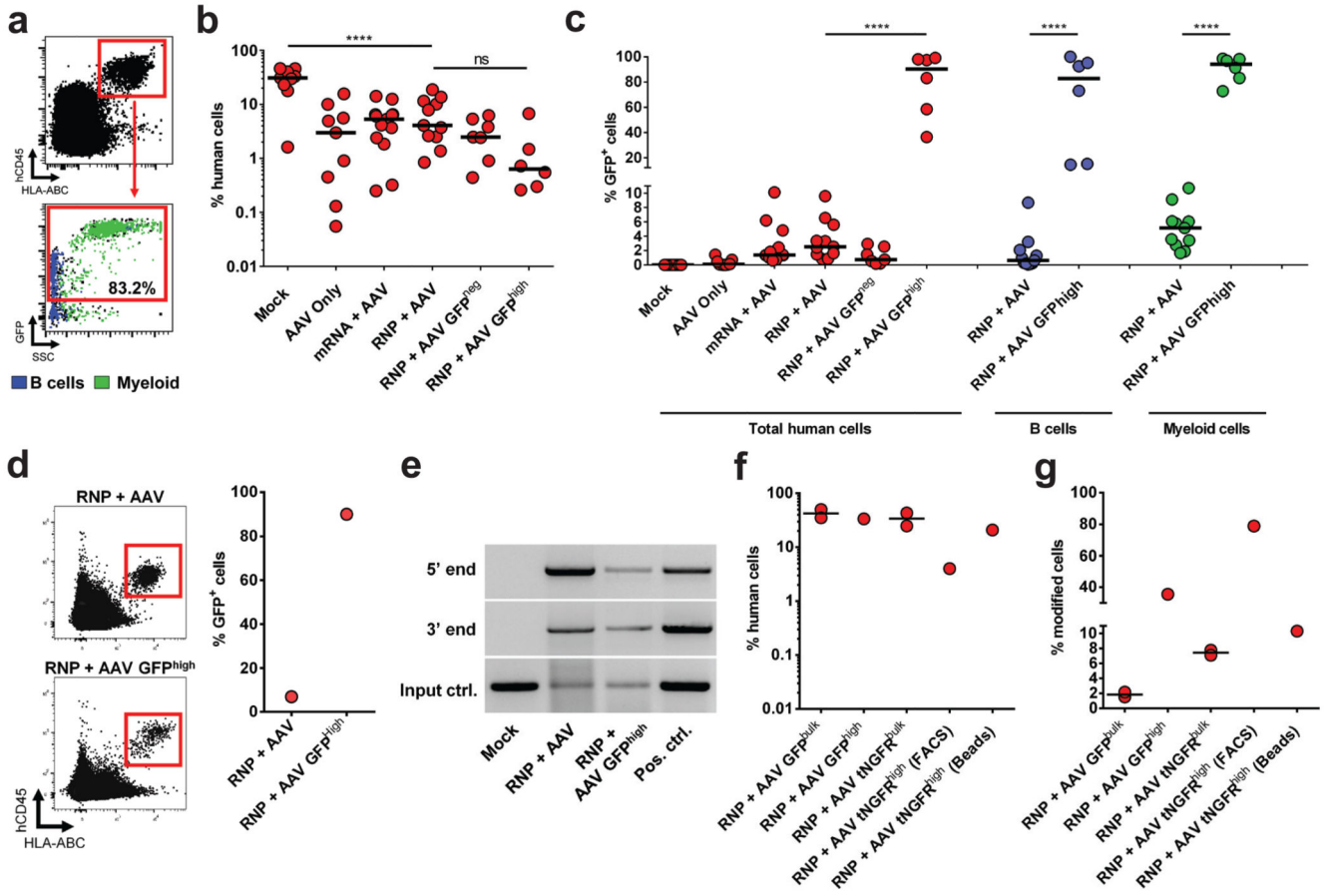


Fig 3. *HBB* gene targeted HSPCs display long-term and multi-lineage reconstitution in NSG mice.

a) 16 weeks post-transplantation, mouse bone marrow was analyzed for human cell chimerism and GFP expression by flow cytometry. *Top panel*: Representative FACS plot from a mouse transplanted with RNP+AAV GFP^{high} HSPCs showing engrafted human cells in the red gate. *Bottom panel:* Representative FACS plot showing GFP-expressing human cells (red gate). CD19^{+} B cells and CD33^{+} myeloid cells are backgated and shown in blue and green, respectively. **b)** Human engraftment in NSG mice from all experimental groups. Three different HSPC donors were used for engraftment studies (N=number of data points within group), **** $p < 0.0001$, ns = $p \geq 0.05$, one-way ANOVA and Tukey's multiple comparison test. Bars represent median. **c)** Percent GFP^{+} cells in the total human population (red), CD19^{+} B cells (blue), and CD33^{+} myeloid cells (green) (N=number of data points within group), * $p < 0.05$, **** $p < 0.0001$, one-way ANOVA and Tukey's multiple comparison test for total human cells and unpaired t test with Welch's correction for B and Myeloid cells. Bars represent median. **d)** 12-14 weeks post-secondary transplantation human cell chimerism and GFP expression was analyzed by flow cytometry. *Left panel*: Representative FACS plot from a secondary mouse transplanted with RNP+AAV (top) or RNP+AAV GFP^{high} (bottom) cells showing engrafted human cells in the red gate. *Right panel:* Percent GFP^{+} human cells in the BM of secondary recipients. **e)** Gel images of In-Out PCRs on sorted human cells from secondary recipients to analyze on-target integrations at

Europe PMC Funders Author Manuscripts

the 5’ and 3’ ends. Input control PCR was performed on the human *CCR5* gene. Positive control is an HSPC sample targeted at *HBB* with SFFV-GFP-PolyA. **f)** 80 million mPB-derived $CD34^+$ cells were electroporated with *HBB*-RNPs and transduced with *HBB* AAV6s Bulk HSPCs or HSPCs enriched for targeting (by FACS or bead-enrichment) were transplanted into the tail vein of sublethally irradiated mice. 16 weeks post transplant human cell chimerism was analyzed by flow cytometry (N=number of data points within group). **g)** Percent GFP^+ and $tNGFR^+$ cells in the human population was analyzed by flow cytometry (N=number of data points within group), bars represent median.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

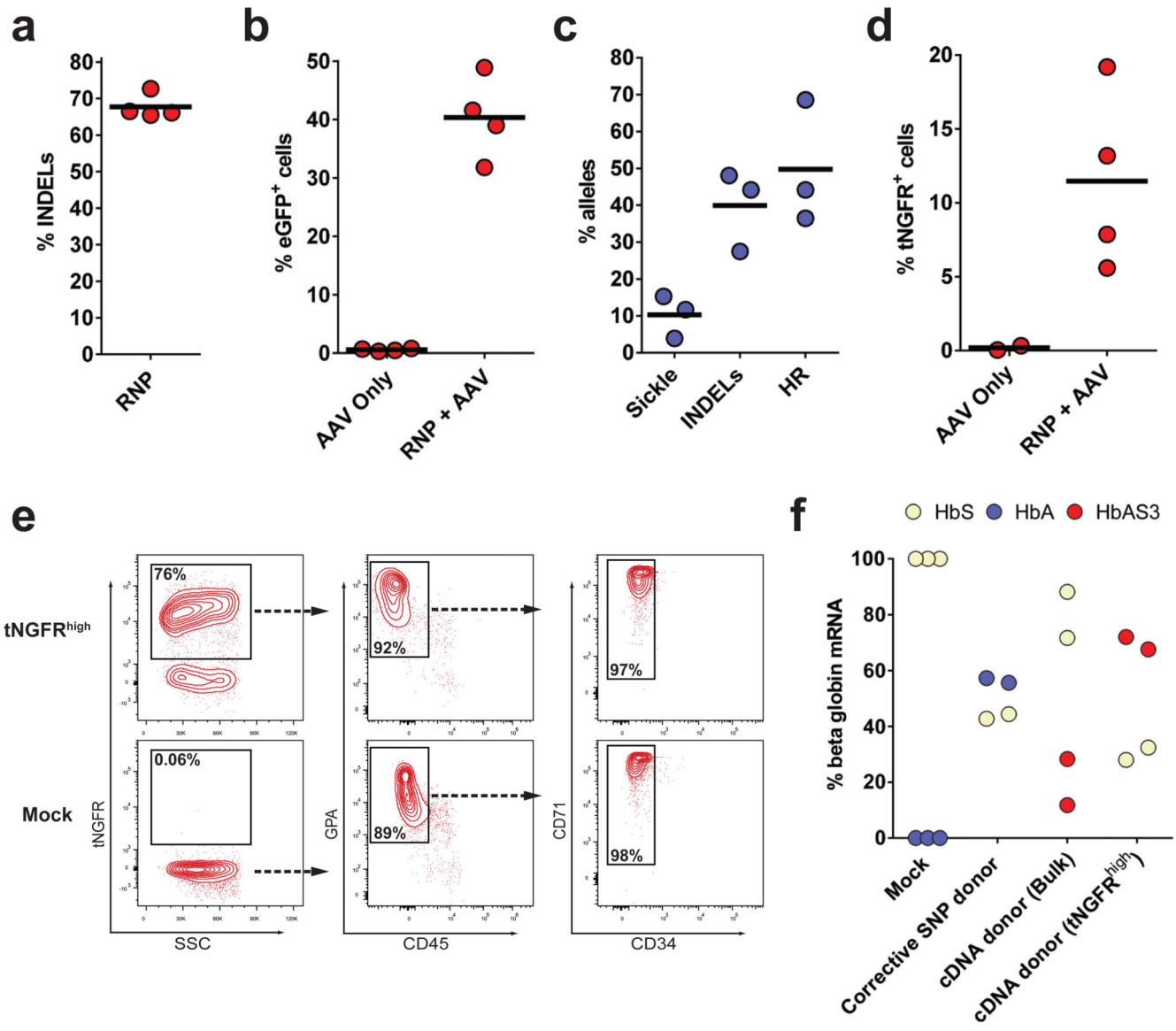


Fig 4. Correction of the E6V mutation in SCD patient-derived HSPCs.

a) gDNA from *HBB* RNP-treated SCD-HSPCs was harvested and INDELs were analyzed via TIDE software (N=4 different SCD patient donors). **b)** *HBB*-targeted SCD-HSPCs were analyzed for GFP expression by flow cytometry (N=4 different SCD patient donors) **c)** SCD HSPCs were targeted with rAAV6 corrective SNP donor. *HBB allele types* were analyzed by sequencing of TOPO-cloned PCR fragments derived from In-Out PCR. 50-100 TOPO clones were analyzed from each of three different HSPC donors (N=3 different SCD patient donors). **d)** SCD HSPCs were targeted with anti-sickling *HBB* cDNA-EF1-tNGFR correction donor. Frequencies of $tNGFR^+$ cells were analyzed by flow cytometry (N=3 different SCD patient donors). **e)** SCD Mock HSPCs and sorted SCD $tNGFR^{high}$ HSPCs were differentiated into erythrocytes *in vitro.* Representative FACS plots from Day 21 of differentiation show cell surface markers associated with erythrocytes ($GPA^+/CD45^-/CD71^+/CD34^-$) **f)** HbS, HbA, and HbAS3 mRNA expression was quantified by RT-qPCR in

erythrocytes differentiated from *HBB*-edited or Mock SCD HSPCs. All mRNA transcript levels were normalized to the *RPLP0* input control (N=2-3 different SCD patient donors).

Europe PMC Funders Author Manuscripts