

SICKLE CELL DISEASE

Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells

Mark A. DeWitt,^{1,2} Wendy Magis,³ Nicolas L. Bray,^{1,2} Tianjiao Wang,^{1,2} Jennifer R. Berman,⁴ Fabrizia Urbinati,⁵ Seok-Jin Heo,³ Therese Mitros,² Denise P. Muñoz,³ Dario Boffelli,³ Donald B. Kohn,⁵ Mark C. Walters,^{3,6} Dana Carroll,^{1,7*} David I. K. Martin,^{3*} Jacob E. Corn^{1,2*}

2016 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science.

Genetic diseases of blood cells are prime candidates for treatment through ex vivo gene editing of CD34⁺ hematopoietic stem/progenitor cells (HSPCs), and a variety of technologies have been proposed to treat these disorders. Sickle cell disease (SCD) is a recessive genetic disorder caused by a single-nucleotide polymorphism in the β -globin gene (*HBB*). Sickle hemoglobin damages erythrocytes, causing vasoocclusion, severe pain, progressive organ damage, and premature death. We optimize design and delivery parameters of a ribonucleoprotein (RNP) complex comprising Cas9 protein and unmodified single guide RNA, together with a single-stranded DNA oligonucleotide donor (ssODN), to enable efficient replacement of the SCD mutation in human HSPCs. Corrected HSPCs from SCD patients produced less sickle hemoglobin RNA and protein and correspondingly increased wild-type hemoglobin when differentiated into erythroblasts. When engrafted into immunocompromised mice, ex vivo treated human HSPCs maintain SCD gene edits throughout 16 weeks at a level likely to have clinical benefit. These results demonstrate that an accessible approach combining Cas9 RNP with an ssODN can mediate efficient HSPC genome editing, enables investigator-led exploration of gene editing reagents in primary hematopoietic stem cells, and suggests a path toward the development of new gene editing treatments for SCD and other hematopoietic diseases.

INTRODUCTION

Sickle cell disease (SCD) is a recessive genetic disorder that affects at least 90,000 predominantly African-American individuals in the United States and hundreds of thousands worldwide (1, 2). The genetic and molecular basis of SCD have been understood for nearly 70 years, but curative treatments have lagged (3, 4). SCD is caused by a single-nucleotide polymorphism (SNP) in the seventh codon of the gene for β -globin (*HBB*), one of two globins that make up the major adult form of hemoglobin. The resulting glutamate-to-valine substitution renders hemoglobin prone to polymerization under hypoxic conditions, producing characteristic “sickle”-shaped red blood cells (RBCs). Sickle RBCs have a markedly reduced life span in the bloodstream, damage the vasculature, and cause vaso-occlusion. Major clinical manifestations of SCD are chronic anemia, severe pain episodes, and progressive damage to vital organs such as the brain, lung, and kidney. In the United States, the disease causes a 30-year decrement in life span and a greatly diminished quality of life (2, 5–7).

RBCs are produced from repopulating hematopoietic stem cells (HSCs) in the bone marrow (BM), and allogeneic hematopoietic cell transplantation (HCT) from an unaffected human lymphocyte antigen (HLA)-matched donor is currently the only lasting cure for SCD (8). However, HCT has been used sparingly because of the difficulty in identifying donors, risks associated with the toxicity of the transplant regimen (requiring preparation with chemotherapy and immune suppression),

and potentially fatal graft-versus-host disease (9, 10). Recent transplant advances have reduced these risks in children (11) and have extended treatment to selected adults (12) and individuals for whom only a haplo-identical HLA donor is available (13). Still, the vast majority of individuals with SCD do not pursue allogeneic HCT because of an unfavorable risk-reward profile, especially during early childhood. A curative treatment for SCD that can be safely applied to more people remains an urgent need.

Gene editing has recently emerged as a promising avenue to treat genetic diseases affecting hematopoietic cells (14–16). Ex vivo editing of autologous hematopoietic stem/progenitor cells (HSPCs) could be followed by reimplantation of edited cells, bypassing donor requirements and eliminating the risks of graft-versus-host disease and postgrafting immunosuppression. Because sickle RBCs have a markedly shorter life span in circulation compared to wild-type (WT) RBCs, even low levels of genotypic correction are predicted to generate a clinical benefit (17). Observations in patients after allogeneic HCT suggest that clinical improvement may occur when as few as 2 to 5% of long-term engrafted cells carry a normal *HBB* allele (18–20). An ideal gene editing treatment would exceed this modest target, but to date, even this level of gene editing has not been achieved (14, 21).

During gene editing, a targeted nuclease creates a double-strand break (DSB) that can be repaired by one of two mechanisms: error-prone non-homologous end joining (NHEJ) that results in genomic insertions and deletions (indels), or templated homology-directed repair (HDR) to precisely insert, delete, or replace a genomic sequence (22). The recent development of CRISPR-Cas9, a programmable RNA-guided DNA endonuclease, has ignited an explosion of interest in gene editing to cure many genetic disorders, including SCD (23, 24). Guided by a single guide RNA (sgRNA), the Cas9 nuclease can be programmed to cut a target locus within the genome, allowing rapid iteration and optimization not possible with other gene editing approaches (23, 25).

Optimized methods for efficient ex vivo gene editing of human HSPCs are required to enable a CRISPR/Cas9-based treatment for blood disorders such as SCD. Recent work has demonstrated that Cas9 can be used for

¹Innovative Genomics Initiative, University of California, Berkeley, Berkeley, CA 94720, USA. ²Department of Molecular and Cellular Biology, University of California, Berkeley, Berkeley, CA 94720, USA. ³Children's Hospital Oakland Research Institute, University of California San Francisco (UCSF) Benioff Children's Hospital, Oakland, CA 94609, USA. ⁴Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA 94588, USA. ⁵Departments of Microbiology, Immunology, and Molecular Genetics; Pediatrics; and Molecular and Medical Pharmacology, University of California, Los Angeles, Los Angeles, CA 90095, USA. ⁶Blood and Marrow Transplant Program, Division of Hematology, UCSF Benioff Children's Hospital, Oakland, CA 94609, USA. ⁷Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84112, USA. *Corresponding author. Email: jcorn@berkeley.edu (J.E.C.); dimartin@chori.org (D.I.M.); dana@biochem.utah.edu (D.C.)

in vitro reversion of the SCD mutation in laboratory cell lines (26, 27) and induced pluripotent stem cells (iPSCs) (28), as well as efficient knockout of an erythroid enhancer in an immortalized cell line (29). Cas9-mediated in vitro editing at the *HBB* locus in HSPCs has recently been reported as well, using plasmid and mRNA delivery of Cas9 (27, 30). Correction of the SCD mutation in iPSCs has been reported using transcription activator-like effector nucleases (31, 32), and most recently, zinc finger nucleases (ZFNs) have been used to correct the SCD mutation in HSPCs, albeit at levels less than 1% in the long-term repopulating stem cell population (21). To date, gene editing has yet to achieve long-term correction of the SCD mutation in these cells at levels predicted to confer clinical benefit, based on engraftment in immunocompromised mice (21). Previous approaches to HSPC gene editing have furthermore required specialized technologies that are not broadly available, including highly engineered and proprietary ZFNs, chemically protected Cas9 guide RNAs, and viral HDR donors (14, 15, 21, 27). Although these factors are not inherently a barrier to clinical translation, they limit the discovery of efficient gene editing reagents for disease mutations, including SCD.

Electroporation of a preassembled ribonucleoprotein (RNP) complex, composed of recombinant Cas9 protein and unmodified in vitro transcribed sgRNA, can be used for gene knockout in a variety of cell types. We have recently extended this approach to enable extremely efficient HDR sequence replacement through rationally designed single-stranded DNA (ssDNA) donors in laboratory cell lines (33–36). Given the efficiency and speed of Cas9 RNP-based editing, we reasoned that a Cas9 RNP-based approach to gene editing could form the basis for an accessible protocol to correct mutations in human HSPCs. We paired the Cas9 RNP with single-stranded DNA oligonucleotide donors (ssODNs), which are generally available, easily designed, and able to mediate efficient sequence replacement in immortalized cell lines (36). Our goal is not only to develop an approach to edit HSPCs as a potential treatment for SCD but also to use reagents that are inexpensive and accessible to a wide variety of researchers, enabling investigator-led studies and rapid optimization to address a multitude of genetic hematopoietic diseases, including polygenic diseases.

We developed a pipeline to enable efficient, RNP/ssODN-based correction of the SCD mutation without introducing a selective marker (Fig. 1A). We first used an erythroleukemia cell line to explore a panel of Cas9 RNPs that cut near the SCD mutation. We then used the most effective RNPs to develop ex vivo editing methods in human HSPCs, achieving up to 33% sequence replacement. We demonstrate efficient correction of the sickle mutation in SCD HSPCs, with corresponding production of WT adult hemoglobin (HbA) RNA and protein in edited, differentiated erythroblasts. After the edited human HSPCs are engrafted in immunocompromised mice, sequence replacement at the SCD locus is retained 4 months after engraftment at levels likely to have clinical benefit.

RESULTS

Prioritizing SCD editing reagents in a model cell line

Pairing Cas9 with various sgRNAs can lead to different activities on the same gene target (37). For HDR-mediated editing, each sgRNA must in turn be paired with an HDR donor template that encodes the desired nucleotide changes. Coupled with subtleties of donor template design and reagent delivery, the complexity of this problem expands rapidly. We chose to use inexpensive, commercially available ssODNs and co-delivered these with Cas9 RNP by electroporation. Using these

components, we were able to quickly iterate combinations of sgRNA, HDR donor, and editing conditions. The rules governing the choice of sgRNA and ssODN are still unclear, although we have developed guidelines based on the mechanism of Cas9's cleavage activity (36).

We searched for maximally active sgRNAs and ssODNs in K562 cells, a human erythroleukemia line that resembles early committed hematopoietic progenitors. Two key restrictions when designing sgRNAs and ssODNs for HDR experiments are the distance between the sgRNA recognition site and the mutation, and the ability to silently ablate the sgRNA protospacer adjacent motif (PAM). This latter constraint ensures that Cas9 cannot recut corrected alleles, thus preventing the introduction of indels into the corrected allele.

To find maximally active sgRNAs compatible with SCD editing, we identified targets in the first exon of the *HBB* gene and tested six of them for which the PAM could be silently mutated (G3, G5, G10, G11, G17, and G18) and one for which it could not (G7) (Fig. 1A). Some of these sgRNAs have been evaluated previously (26, 27), but to our knowledge, they have not been used in a Cas9 RNP format. We also tested truncated versions of three guides (trG3, trG10, and trG11) (table S1) because truncation has been reported to reduce off-target cleavage without compromising on-target efficiency (37). We designed an initial ssODN (T1) that contained silent mutations in the PAMs of all sgRNAs and a WT-to-SCD edit (K562 cells are WT at *HBB*) (fig. S1A). The combination of the WT-to-SCD SNP mutation and removal of the G5 PAM generates a silent Sfc I restriction site, allowing easy tracking of HDR-mediated editing (fig. S1A).

We assembled Cas9 RNPs from each candidate sgRNA and individually delivered them to K562 cells together with the ssODN by electroporation. We used a dose of 100 pmol of each RNP and template per 150,000 K562 cells, based on a previous work (33). A T7 endonuclease I digest of polymerase chain reaction (PCR) amplicons from edited cell pools, which, in this case, detects mismatches arising from both NHEJ- and HDR-mediated repair of Cas9 DSBs, revealed sequence modifications with all but two sgRNAs (G3 and trG3) (Fig. 1B). Sfc I digest showed appreciable HDR-mediated editing of the SCD SNP with three sgRNAs: G5, G10, and trG10 (truncated G10) (Fig. 1C).

We quantified the frequency of HDR and indel formation at the SCD SNP by next-generation sequencing (NGS) of PCR amplicons derived from genomic DNA extracted from pools of edited cells (Fig. 1D) and by droplet digital PCR (ddPCR) (fig. S1B). When the trG10 Cas9 RNP was provided without an ssODN, 90% of NGS reads contained indels, indicating excellent delivery of the RNP and efficacious gene knockout (fig. S1C). We performed experiments using the T1 ssODN and a derivative of T1 bearing only the WT-to-SCD edit and the G5 and G10 PAM mutations (T2) (fig. S1A). Most candidate sgRNAs induced substantial quantities of indels (35 to 40% of reads) (Fig. 1, B and D), and G5, G10, and trG10 also yielded high levels of WT-to-SCD HDR (11% of reads for G5, 41% of reads for G10) from multiple donor designs (Fig. 1 and fig. S1). A phosphorothioate-protected ssODN did not improve HDR frequencies (fig. S1D). The G5 guide targets a sequence very similar to one found in the closely related δ -globin (*HBD*) gene, and we experimentally verified that this guide induces indels in *HBD* (fig. S1E). Hence, we selected the truncated trG10 guide for further testing.

To optimize an ssODN for HDR, we drew guidance from recent work in our laboratory showing that, upon binding its target, Cas9 releases the PAM-distal nontarget strand and that asymmetric homology arms taking advantage of this property can increase HDR efficiency (38). For sgRNAs that target the sense strand, such as G10, the best template to use thus matches the sense strand (annealing to the antisense strand) and

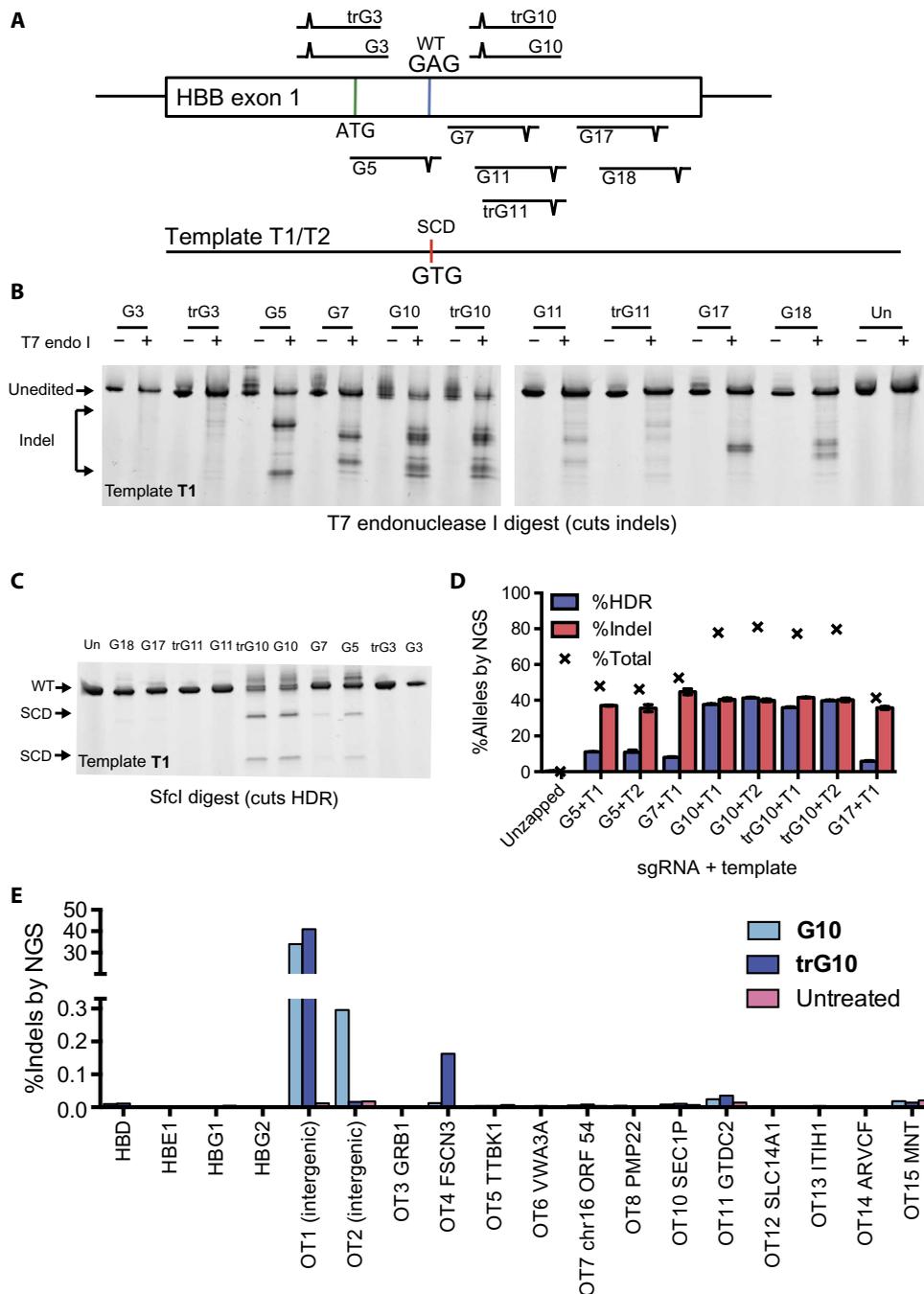


Fig. 1. Editing the SCD SNP in K562 cells. (A) Schematic depicting the experimental approach to editing in K562 cells. A panel of 10 sgRNAs that cut within 100 base pairs (bp) of the SCD SNP was selected. A WT-to-SCD edit was programmed by an ssDNA template (T1) bearing silent PAM mutations for the sgRNAs, which also introduces an Sfc I restriction site. (B) T7 endonuclease I assay depicting indel formation in pools of cells edited by candidate RNP. (C) Editing of candidate sgRNAs detected by Sfc I digestion. G5, G10, and a truncated variant, trG10, efficiently edit in K562 cells. (D) Gene modification of select sgRNAs and templates at the SCD SNP, assessed by NGS. See fig. S1 for definitions of donors T1 and T2. Data are the average of three biological replicates, and error bars indicate SD. (E) Analysis of off-target cutting by the G10 RNP at sites predicted by the online CRISPR Design Tool, in K562 cells, determined by NGS, of a single replicate.

has a long 5' homology arm and a shorter 3' annealing arm (fig. S1F). On the basis of these principles, we henceforth identify templates by the lengths of their 5' and 3' homology arms relative to the G10 cut site. For example, the initial template T2 is termed T88-107. Unless noted,

all templates bear only the G5 and G10 PAM mutations. We tested the effects of asymmetric templates containing the SCD-to-WT edit on HDR-mediated editing in K562 cells (fig. S1C). We found that a template with a 111-nt 5' arm and a 57-bp 3' arm (T111-57) yielded an HDR frequency modestly higher than the original template T88-107 (33% versus 28.5% in the same experiment). We also found low but measurable rates of conversion (0.5 to 1%) between the HBB coding sequence and the homologous region in HBD, but this was only observed in samples treated with G7, G10, and trG10 (fig. S1G).

The possibility that Cas9 may cut at off-target sites is a concern for the development of Cas9-based therapies (39, 40). This tendency may be reduced by the use of Cas9 RNP delivery, truncated sgRNAs, and other emerging techniques (33, 41, 42). We selected off-target candidates by sequence similarity using a popular off-target prediction tool (40) and used NGS of PCR amplicons to analyze the off-target activity of G10 and trG10 Cas9 RNPs in K562 cells (Fig. 1E). We assessed editing at the two top-scoring off-targets (which are both intergenic), the top 13 exonic off-targets, and predicted off-target sites within the four globin genes (*HBD*, *HBE1*, *HBG1*, and *HBG2*) (fig. S2) (43).

We observed very little exonic off-target activity, with most off-target sites showing no Cas9-dependent indel formation within the limit of detection (~0.001%) (Fig. 1E). Substantial off-target activity was observed for both G10 and trG10 at a top-scoring intergenic site (OT1, chr9:104,595,865, 45% indel formation), which lies ~3 kb from the nearest annotated gene sequence and is nearly identical to the G10 target site. High off-target activity by G10 at this site has been observed previously (26). Off-target activity was detected with G10 at the other intergenic site (OT2, chr17:66,624,238, 0.30%), but this was absent with the truncated guide trG10. Low but detectable off-target activity was observed with both the G10 and trG10 RNPs at two exonic off-targets, FSCN3 and GTDC2. Indel formation was lower for the full-length G10 RNP (FSCN3: 0.013 and 0.16% indel for G10 and trG10, respectively; GTDC2: 0.025 and 0.035% indel for G10 and trG10, respectively). Very few indels were observed at HBD (0.010 and 0.011% indel for G10 and trG10, respectively), and indels at other β -globin genes (*HBE1*, *HBG1*, and *HBG2*) were not detected over background (Fig. 1E).

Optimizing candidate Cas9 RNPs and ssODNs to edit CD34⁺ HSPCs

The results of RNP delivery to K562 cells encouraged us to investigate using this approach with HSPCs, recognizing that editing efficiency can differ between cell types. We used the trG10 sgRNA, which showed robust HDR and few off-target effects in K562, and reoptimized HDR in adult mobilized peripheral blood HSPCs, which are the cells most relevant for therapeutic gene editing to address SCD. Because these cells were obtained from healthy WT donors, our initial experiments used ssODNs bearing a WT-to-SCD mutation.

We first explored optimization of electroporation conditions, template design, and RNP/ssODN dose in HSPCs (fig. S3, A and B). To selectively interrogate editing in viable cells, we used NGS to assay HSPCs cultured under erythroid expansion conditions for 7 days after editing (erythroid-expanded), along with edited HSPCs cultured for only 2 days (unexpanded). During initial treatments, we found that treatment with 100 pmol of RNP led to a decline in viability (fig. S3C), and used 75 pmol of RNP as a lower dose in subsequent *in vitro* experiments. We also tested the effects of the small molecule SCR7 on editing in HSPCs. This NHEJ inhibitor has been reported to increase HDR-mediated editing in some cell types, but we did not observe an improvement in HDR in HSPCs (fig. S3A) (44, 45).

We observed appreciable initial levels of editing at the sickle SNP in HSPCs, with HDR rates between 6 and 11% (Fig. 2, A and B). After 5 days of erythroid expansion, HDR rates increased, with up to 33% editing at both high and low doses of RNP (150 and 75 pmol, respectively). Total editing (%HDR + %NHEJ) in expanded HSPCs was between 66 and 72%, indicating good delivery of the trG10 RNP to HSPCs. In general, higher editing was accompanied by some reduction in viability, as measured by fewer cells remaining after treatment, particularly at a high, 150 pmol dose of RNP (fig. S3C). The asymmetric ssODN was most effective in K562 cells (T111-57) drove HDR more efficiently at a lower Cas9 dose of 75 pmol RNP per 150,000 HSPCs, whereas a shorter template (T111-27) was more efficient at a higher dose of Cas9 (150 pmol RNP per 150,000 HSPCs) (Fig. 2A). In a separate experiment, we confirmed high rates of HDR by ddPCR (fig. S3D). These experiments demonstrate efficient *in vitro* editing of CD34⁺ HSPCs using the Cas9 RNP, including HDR-mediated

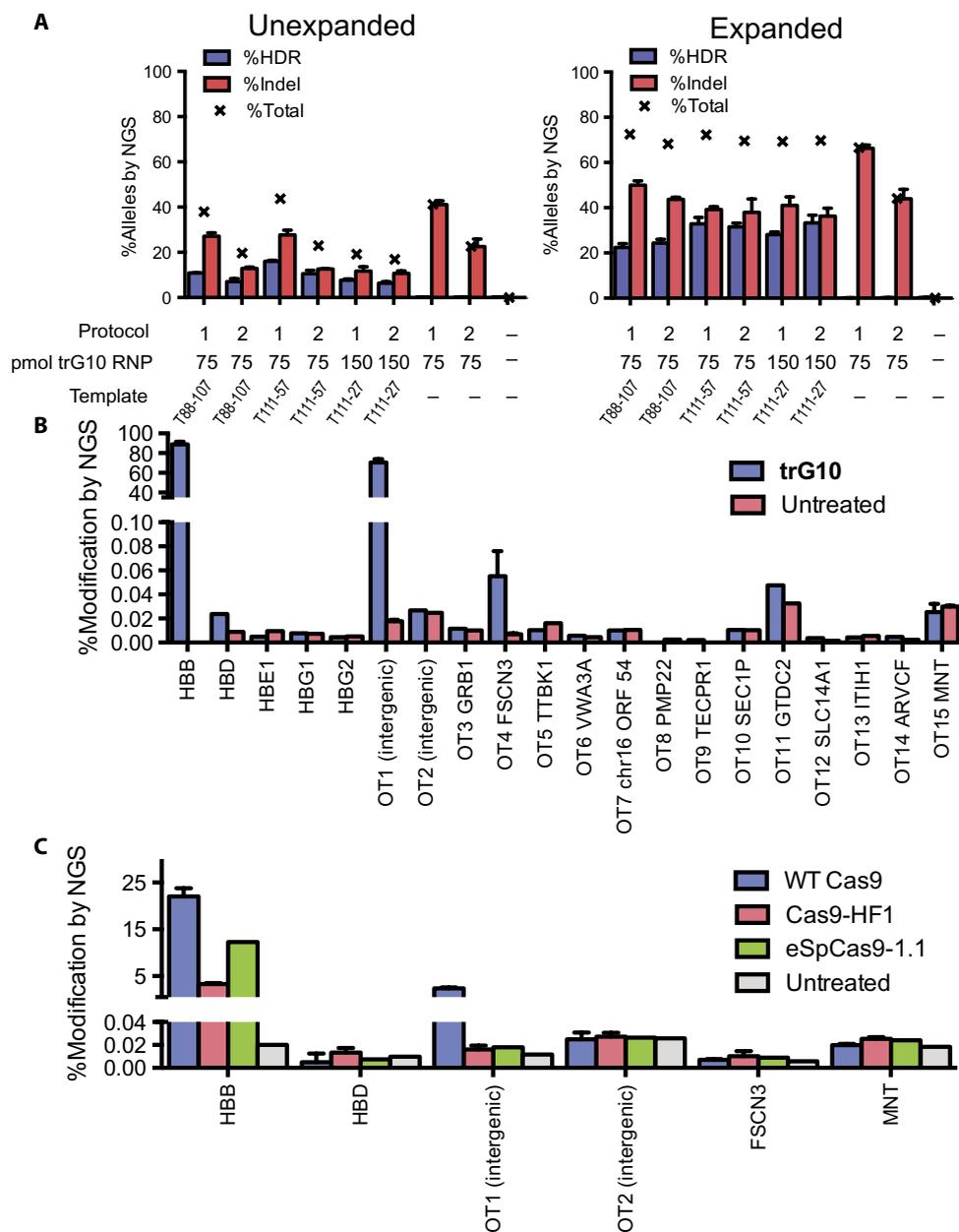


Fig. 2. Editing of WT human CD34⁺ HSPCs by the Cas9 RNP. (A) Analysis of editing in unexpanded HSPCs (left) and erythroid-expanded HSPCs (right), using trG10 RNP and under conditions as indicated. Templates, which are asymmetric about the G10 cut site, were designed as described in the text. All graphs show three biological replicates, and error bars are \pm SD. (B) Modification (HDR+Indel) at off-target sites in HSPCs edited with the trG10 RNP and template T88-107, as compared to untreated cells. The bars for HBB, OT1, FSCN3, and MNT represent three biological replicates, with error bars (\pm SD); the others are each a single replicate. Targets were selected using the online CRISPR Design Tool. (C) Indel formation at on- and off-target sites by Cas9 mutants with increased specificity (HF1 and eSpCas9-1.1), in HSPCs, as compared to WT Cas9, all complexed to the G10 sgRNA and no ssODN. The absence of an ssODN reduced indel formation overall. All data are shown as $n = 3$ biological replicates, with error bars (\pm SD).

sequence replacement using ssODNs without a selection marker. As was the case with K562 cells, in HSPCs we found low but measurable conversion of HBB coding sequence to HBD (0.2 to 2%), and rates of conversion increased after expansion of edited cells (fig. S3E).

To analyze how allele frequencies in the HSPC population translate to the editing of alleles in individual HSPCs, we repeated our best editing condition in CD34⁺ HSPCs (75 pmol trG10 RNP and 100 pmol T111-57),

recovered the cells in HSPC expansion medium for 2 days, and plated single edited HSPCs by limiting dilution in erythroid expansion medium. After 14 days of growth, 96 edited clones were individually genotyped by multiplexed NGS (Table 1). In this experiment, 21% of alleles were HDR. However, these alleles were spread among 32% of the cells. This increased prevalence of edited cells relative to edited alleles is predicted by the independent assortment of alleles within a population (although we observed that homozygous genotypes were still overrepresented relative to a prediction based on random assortment) (46).

We analyzed off-target activity of the trG10 RNP in HSPCs using the target selection criteria described above for K562 cells (Fig. 2B). Most predicted genic off-targets showed no detectable indel formation, although cutting at the previously observed intergenic site remained high (OT1, ~80% indel). The rates of off-target cleavage observed in HSPCs generally corresponded to observations in K562 cells, although the rates were often reduced (for example, rate of FSCN3 is ~0.05% in HSPCs versus up to 0.16% in K562 cells).

As an additional test of the effects of off-target activity of the trG10 RNP, we used overamplification PCR to detect the presence or absence of translocations between the on-target site at HBB and selected off-target sites (OT1, HBD, FSCN3, and MNT) in both K562 cells and HSPCs edited with trG10 (fig. S4). Translocations between OT1 and HBD with HBB were observed in K562 cells. Most sites showed no translocations in HSPCs; one translocation may have occurred between OT1 and HBB in HSPCs, but the frequency of this event is not known.

We also tested whether trG10 has off-target activity at cancer-associated genes. Although these sites bear little similarity to the trG10 protospacer, even low levels of activity in such locations would be a concern. We used a capture library (Illumina TruSight Cancer) to sequence 94 genes and 290 cancer-associated SNPs to ~8000-fold coverage each. Relative to unedited cells, we found a small number of indel mutations enriched in K562 cells edited with the trG10 RNP, typically at less than 1% of alleles (table S2). Almost all of these mutations were present in unedited cells.

In contrast, no indels were detected in similarly edited HSPCs. These results highlight the importance of performing rare off-target event detection in the target primary cell type and suggest that the trG10 RNP may not have substantial genotoxic liabilities, although this has not been tested at clinical scale.

Two mutant variants of the Cas9 protein have been reported to reduce off-target effects in cell lines, even at highly similar off-target sites such as OT1 for the G10 RNP (47, 48). We expressed and purified the eSpCas9-1.1 and HF1 variants of Cas9, paired them with the G10 sgRNA to form RNPs, and used them to edit WT CD34⁺ HSPCs in the absence of an ssODN. We used NGS to determine the on- and off-target editing frequencies (Fig. 2C). As expected based on our recent observation of enhancement of indel formation by oligonucleotides, the on-target indel formation was reduced in the absence of an ssODN (Fig. 2B versus Fig. 2C) (38). Compared to WT Cas9, both HF1 and eSpCas9-1.1 showed even further decreased on-target indel formation at the HBB locus, for example, an almost fivefold decrease in indels for HF1. However, both modified enzymes also completely eliminated all previously observed off-target events, including the prevalent OT1 intergenic off-target (Fig. 2C).

Correction of the SCD mutation in SCD HSPCs to produce WT hemoglobin

Our success in WT-to-SCD editing in HSPCs implies that the same method could be used to edit SCD to WT in HSPCs derived from SCD patients. Because human erythropoiesis does not occur when human HSPCs are xenografted into mice, and the availability of SCD HSPCs is limited, we evaluated the effects of correcting the SCD mutation in HSPCs in vitro, by carrying out erythroid differentiation of edited HSPCs. We obtained CD34⁺ HSPCs from whole blood discarded after exchange transfusion of SCD patients. Because the HF1 and eSpCas9-1.1 proteins yielded reduced levels of on-target editing and the predominant off-target from WT Cas9 lies in an intergenic region with no known function, we focused on experiments using more efficacious WT Cas9. We corrected the SCD mutation using the trG10 RNP and ssODNs carrying an SCD-to-WT edit. These SCD-to-WT templates, denoted by the suffix “S,” encode the same number of mutations as the WT-to-SCD templates, with the base identity different only at the SCD SNP. Measuring editing by both NGS and ddPCR, we found that SCD HSPCs were edited at levels similar to those observed in WT HSPCs from mobilized blood, with up to 25% of alleles corrected to WT at high RNP dose and 18% corrected at low RNP dose (Fig. 3A and fig. S5).

To analyze the hemoglobin production potential of corrected HSPCs, we differentiated pools of treated HSPCs into enucleated erythrocytes and late-stage erythroblasts and measured hemoglobin by high-performance liquid chromatography (HPLC) (21, 28, 49). We found that corrected HSPC pools produce substantial amounts of WT HbA, with a concomitant decrease in sickle hemoglobin (HbS) (22.2 to 22.4% HbA, 48.0 to 50.6% HbS at low-dose RNP, 29.3% HbA, and 38.7% HbS at high-dose RNP) (Fig. 3, B and C). We also observed a substantial increase in fetal hemoglobin (HbF) in edited cell pools (16.3 to 17.4% HbF in edited cells versus 2.0% HbF in unedited cells).

We used RNA sequencing (RNA-seq) to measure globin transcript abundance in pools of edited SCD HSPCs differentiated to erythrocytes (50). Globin transcript levels showed a trend similar to protein levels after editing, with sickle HBB transcripts decreasing from 56.7% of all transcripts to ~9% and WT HBB transcripts increasing from 0.1 to 13% across all three editing conditions (Fig. 3D). Consistent with the increase in HbF protein, we observed about threefold increase in the expression

Table 1. Zygosity of clonal colonies of CD34⁺ HSPCs edited with the trG10 RNP. HSPCs were edited with 75 pmol of the trG10 RNP (similar to Fig. 2) and cloned by limiting dilution. Of the resulting clones, 96 were then genotyped by NGS. The fraction of all three alleles, the frequency of clones with at least one copy of each of the three alleles, and the frequency of all six genotypes are indicated.

n = 96 colonies						
	%WT	%Indel	%HDR			
All alleles	46	33	21			
Clones with one allele %WT/_ %Indel/_ %HDR/_						
Actual	60	48	32			
Predicted	71	55	37			
Clones by genotype						
	%WT/ WT	%WT/ Indel	%WT/ HDR	%Indel/ Indel	%Indel/ HDR	%HDR/ HDR
Actual	32	18	10	18	13	9.4
Predicted	21	30	19	11	14	4.3

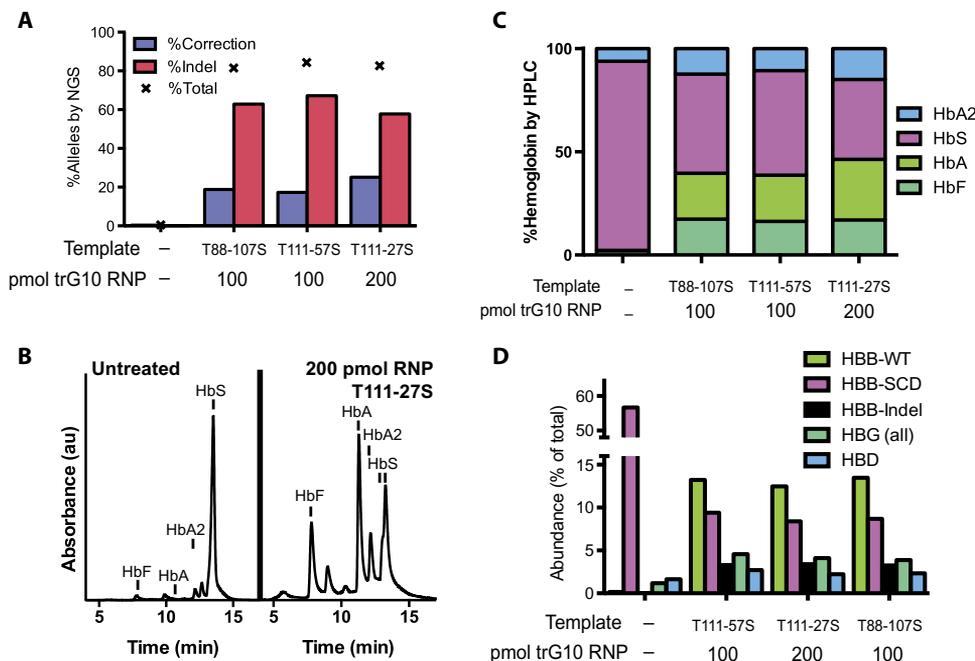


Fig. 3. Correction of the SCD mutation in SCD HSPCs. (A) Editing the SCD mutation in unexpanded CD34⁺ HSPCs from the whole blood of SCD patients, assessed by NGS. (B) HPLC trace depicting hemoglobin production in SCD HSPCs edited with 200 pmol of the trG10 RNP and the T111-27S donor, as compared to untreated HSPCs, after differentiation into late-stage erythroblasts. Increases in HbA, HbF, and HbA2 are apparent. au, arbitrary units. (C) Stacked bars showing HPLC results, with HSPCs edited as indicated before differentiation into erythroblasts. (D) Globin gene expression in SCD HSPCs edited as in (C), determined by RNA-seq. All experiments include a single replicate for each of the four treatment conditions.

of γ -globin (HBG1 and HBG2) mRNA. Thus, corrected pools of cells produce decreased sickle β -globin, increased adult WT β -globin, and increased fetal (γ) globin.

Repopulation by edited HSPCs in vivo

The trG10 Cas9 RNP and an ssODN donor template can efficiently edit the SCD mutation in CD34⁺ HSPCs, and erythrocytes derived from these cells have altered hemoglobin levels consistent with substantial gene correction. However, in order for gene correction to manifest in vivo, edited repopulating stem cells must engraft and repopulate within a recipient (51). A powerful method of assaying repopulating stem cells is through long-term xenografting in an immunodeficient mouse model, such as the NOD/SCID/IL-2 γ^{null} (NSG) mouse (52). In this model, edited human stem cells must engraft in the mouse and persist over several months for them to be observed in subsequent experiments. After 16 weeks, progenitor cells should be lost from the system, and human cells in the BM should be derived from long-term repopulating stem cells within the initial HSPC population (21, 52). Because SCD HSPCs are difficult to obtain at the necessary scale, we tested engraftment by editing WT HSPCs to SCD and implanting them in NSG mice.

We injected NSG mice with pools of WT CD34⁺ HSPCs edited with the trG10 RNP and the T88-107 template (seven mice over three treatments, 1×10^5 cells per mouse), along with two uninjected NSG mice (52). Engraftment was monitored by fluorescence-activated cell sorting (FACS) analysis of blood draws at 5 and 8 weeks after injection. Final engraftment was assessed at 16 weeks after injection, when mice were sacrificed and BM cells were harvested and subjected to FACS-based lineage analysis (six mice) (Fig. 4A and fig. S6). Substantial numbers of hematopoietic (CD45⁺) cells were detectable in BM of all injected mice at 16 weeks ($37 \pm 21\%$

human CD45, mean \pm SD), demonstrating that edited cells maintain long-term repopulating potential. Within the BM, engrafted human CD45⁺ cells were primarily B cells (CD19⁺; $57 \pm 18\%$, mean \pm SD) and myeloid cells (CD33⁺; $25 \pm 18\%$, mean \pm SD) (fig. S6B).

Genotyping of edited HSPCs was performed by NGS immediately after editing and before injection (Fig. 4B) and from mice at weeks 8 (blood) and 16 (BM and spleen) (Fig. 4, C and D). The input populations had $55 \pm 19\%$ indel alleles and $11.8 \pm 3.7\%$ HDR alleles (mean \pm SD) (Fig. 4B). NGS analysis revealed consistently high levels of indel alleles after 16 weeks ($46 \pm 6\%$ in BM, mean \pm SD), along with maintenance of HDR-mediated editing at the SCD SNP throughout the lifetime of the xenograft in both BM and spleen ($2.3 \pm 1.8\%$ in BM, $3.7 \pm 1.4\%$ in spleen; mean \pm SD). Human cells engrafted in one mouse maintained HDR-mediated editing in BM at a markedly high level (6.2%), although the reasons for this difference are unclear. To ensure that editing was present in the progenitor (CD34⁺) cells and not overrepresented in the B lymphocyte (CD19⁺) populations, marrow from two mice was sorted for these markers and

genotyped by NGS. Editing was maintained within both populations, indicating durable editing of the long-term stem cell (HSC) population (fig. S6). Despite the decreased HDR rate relative to input, this level of sequence replacement using Cas9 RNP and ssODNs is more than ninefold greater than the previously reported editing of the SCD mutation in HSCs using ZFN mRNA electroporation and ssODNs (21).

DISCUSSION

The treatment of genetic diseases by gene editing to replace an endogenous sequence is a long-standing goal of regenerative medicine. Unlike gene therapy using integrating viral vectors, wherein regulation of the introduced gene may be compromised and endogenous genes may be disrupted, gene editing corrects the disease mutation at the endogenous locus. In the case of most genetic blood disorders, lasting correction requires HDR-mediated editing of endogenous genes in repopulating HSCs. Ideally, these edits should be efficient enough to operate without the introduction of a nonnative selection marker. This is technically challenging, particularly because the edits do not typically confer a selective advantage in the target cell compartment, as is the case for correction of the HBB sickle mutation in BM stem cells (14, 21).

Here, we used the Cas9 RNP and ssODNs to develop a rapid and extensible gene editing pipeline to introduce SNPs into human adult HSPCs, focusing on the SCD mutation. We used K562 cells to iterate combinations of Cas9 RNPs and ssODNs that edit the HBB gene and then applied these reagents to human HSPCs. These reagents efficiently induce HDR-mediated editing of the SCD mutation in HSPCs with minimal genic off-target activity. Previous reports of gene knockout in HSPCs by Cas9 mRNA delivery required chemical protection of the

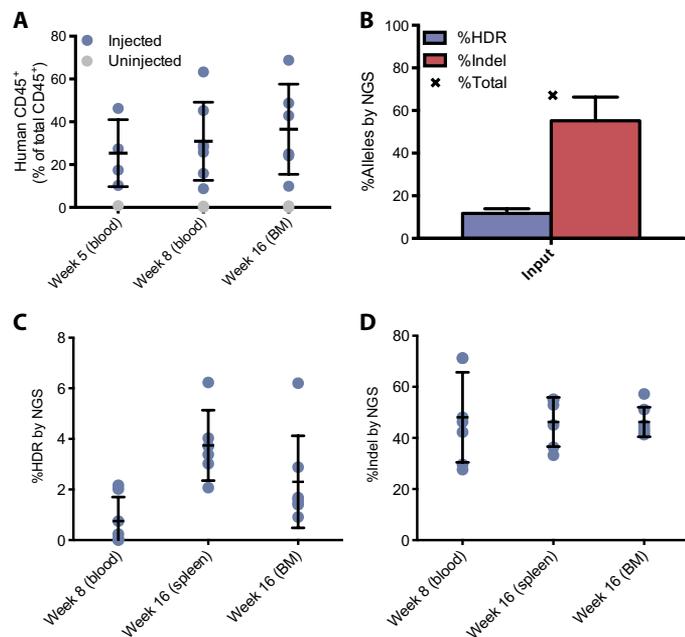


Fig. 4. Engraftment of edited HSPCs into NSG mice. (A) Engraftment of human CD45⁺ cells in NSG mice injected with edited HSPCs, compared to two uninjected mice. (B) Analysis by NGS of editing at the SCD SNP in cells before engraftment. Error bars indicate mean \pm SD of three separate experiments with cells from two healthy donors. (C) HDR-mediated editing and (D) indel formation at the SCD SNP in human cells engrafted in mouse blood, spleen, and BM at 5 and 16 weeks after injection. Error bars indicate mean \pm SD over either six mice (spleen) or seven mice (BM and blood).

sgRNA (27), but we found that RNP delivery yields efficient knockout and HDR even with unmodified sgRNAs and ssODNs. Although we used a truncated guide (trG10) here, we found no apparent decrease in off-target activity compared to the full-length guide, possibly due to the extremely high homology between the on- and off-target sites. RNP delivery has been associated with reduced off-target effects (41), suggesting that this modality could allow efficient editing of disease mutations in HSPCs using full-length unmodified sgRNAs such as G10.

We performed extensive off-target analysis using the trG10 RNP in HSPCs. We found few genic off-target events and one major intergenic event that had been described previously (26, 27). We did not find clear evidence of frequent translocation events between targeted loci in HSPCs. Similarly, we found low-frequency mutations at a few cancer-associated loci in edited K562 cells, typically at sites with measurable mutations in unedited cells, but did not observe any mutations in HSPCs. This suggests that off-target analysis should be performed in the target primary cell type and not cancer cell lines. To reduce off-target activity, we tested several recently characterized high-fidelity Cas9 variants as RNPs, but chose to pursue WT Cas9 due to their reduced on-target activity (47, 48).

Recent gene editing efforts in hematopoietic cells have focused on the use of viral HDR donors that are often delivered at very high multiplicity of infection (15, 16, 21). Advances in viral delivery technology have greatly improved safety, and viral donors are a valid option for gene correction. However, genomic insertion of even nonintegrating vector sequences has been observed during genome editing, and high levels of off-target integration have been observed in HSPCs (53, 54). Furthermore, the intensive engineering associated with viral donor design restricts optimization. By contrast, the co-delivered RNP/ssODN approach described here is nonviral, modular, and readily optimized.

Here, we describe levels of sequence replacement in long-term stem cells that may be clinically relevant (21). However, HDR was still diminished relative to the input CD34⁺ populations. Reductions in HDR frequency during long-term engraftment have been observed previously and remain a major impediment to bringing HDR-based therapies to the clinic (14, 15, 21). It may be that HDR-edited HSCs do not engraft as well as unedited cells, that the HDR donors are themselves toxic, or that HSCs are intrinsically more difficult to edit than other CD34⁺ cells. We found that relatively high doses of ssDNA donor do not affect indel formation even during long-term engraftment, implying that donor delivery is not in itself toxic, but that the stem cells may instead be intrinsically refractory to HDR-based editing or that HDR causes HSCs to lose their long-term regenerative potential.

Corrected SCD HSPCs produced greatly reduced sickle and increased WT hemoglobin protein and β -globin mRNA. Edited HSPCs also produced more γ -globin mRNA and HbF. Although the cause of this increase is unclear, it is possible that indels within HBB cause selective expansion of HbF-expressing cells during differentiation or that indels within HBB stimulate a cell-intrinsic up-regulation of HbF. The molecular basis by which alterations at β -globin cause changes in HbF abundance warrants further investigation.

Observations of transient mixed chimerism (2 to 5%) after allogeneic HCT for SCD suggest that correcting the sickle allele in relatively few BM stem cells could translate into a clinically meaningful increase in the numbers of nonsickling circulating RBCs (see Supplementary Discussion) (18–20). The abrogation of ineffective erythropoiesis intrinsic to sickle cell anemia and the greatly increased life span in circulation of WT RBCs carrying a WT β -globin allele imply that correction of a small fraction of HSCs can substantially increase HbA level and lower the HbS fraction in the blood. Furthermore, we found that alleles assort largely independently within the edited population, such that a given percentage of allele correction of alleles in a population corresponds to correction of nearly twice as many cells. This should be taken into account when comparing the allelic correction data presented here with observations of mixed chimerism (for example, correction of 1 to 3% of alleles may correct 2 to 6% of cells) (table S3 and Supplementary Discussion).

An approach to SCD treatment that integrates a gene editing approach could have several limitations. First, it is possible that the allelic correction frequency we observed in the HSPC population will ameliorate but not eliminate the clinical expression of SCD. Second, although autologous HSCT should be safer than allogeneic HSCT, it is still an intensive procedure, requiring myeloablative conditioning, which can have serious side effects. Finally, clinical translation of an HSCT-based therapy is not feasible in developing countries, where SCD is most prevalent. We anticipate that further developments will address these remaining issues.

The approach described here allows researchers to edit SNPs at endogenous loci in human adult HSPCs using readily available reagents that are conducive to rapid iteration and optimization. Given the low barrier to entry, we anticipate that these democratizing methods will enable investigator-led gene editing studies in a wide variety of disease areas. We predict that the methods outlined can be improved and scaled up for a gene editing treatment for SCD.

MATERIALS AND METHODS

For detailed Materials and Methods, please see the Supplementary Materials.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/8/360/360ra134/DC1

Discussion

Materials and Methods

Fig. S1. Additional data showing editing of K562 cells with the Cas9 RNP.

Fig. S2. Genomic context of predicted off-target cut sites for the G10 RNP.

Fig. S3. Additional data showing editing of CD34⁺ HSPCs with the Cas9 RNP.

Fig. S4. Qualitative detection of chromosomal translocations by overamplification PCR.

Fig. S5. Confirmation of efficient correction of SCD HSPCs by ddPCR.

Fig. S6. Additional data on engraftment and editing of HSPCs in NSG mice.

Table S1. Oligonucleotides used in this study.

Table S2. Indel mutations at cancer-associated genes in K562 cells and HSPCs edited with the trG10 RNP, as compared to unedited cells.

Table S3. Estimates of cellular editing in mice engrafted with edited HSPCs.

Table S4. Tabulated data from Figs. 1 to 4 (provided as an Excel file).

References (55–57)

REFERENCES AND NOTES

- J. A. Panepinto, M. Bonner, Health-related quality of life in sickle cell disease: Past, present, and future. *Pediatr. Blood Cancer* **59**, 377–385 (2012).
- A. S. Adewoyin, Management of sickle cell disease: A review for physician education in Nigeria (sub-Saharan Africa). *Anemia* **2015**, 791498 (2015).
- L. Pauling, H. A. Itano, S. J. Singer, I. C. Wells, Sickle cell anemia, a molecular disease. *Science* **110**, 543–548 (1949).
- J. V. Neel, The inheritance of sickle cell anemia. *Science* **110**, 64–66 (1949).
- D. K. McClish, L. T. Penberthy, V. E. Bovbjerg, J. D. Roberts, I. P. Aisiku, J. L. Levenson, S. D. Roseff, W. R. Smith, Health related quality of life in sickle cell patients: The PISCES project. *Health Qual. Life Outcomes* **3**, 50 (2005).
- O. S. Platt, D. J. Brambilla, W. F. Rosse, P. F. Milner, O. Castro, M. H. Steinberg, P. P. Klug, Mortality in sickle cell disease—Life expectancy and risk factors for early death. *N. Engl. J. Med.* **330**, 1639–1644 (1994).
- D. R. Powars, L. S. Chan, A. Hiti, E. Ramicone, C. Johnson, Outcome of sickle cell anemia: A 4-decade observational study of 1056 patients. *Medicine* **84**, 363–376 (2005).
- M. C. Walters, L. M. De Castro, K. M. Sullivan, L. Krishnamurti, N. Kamani, C. Bredeson, D. Neuberg, K. L. Hassell, S. Farnia, A. Campbell, E. Petersdorf, Indications and results of HLA-identical sibling hematopoietic cell transplantation for sickle cell disease. *Biol. Blood Marrow Transplant.* **22**, 207–211 (2016).
- B. P. Yawn, G. R. Buchanan, A. N. Afenyi-Annan, S. K. Ballas, K. L. Hassell, A. H. James, L. Jordan, S. M. Lanzkron, R. Lottenberg, W. J. Savage, P. J. Tanabe, R. E. Ware, M. H. Murad, J. C. Goldsmith, E. Ortiz, R. Fulwood, A. Horton, J. John-Sowah, Management of sickle cell disease: Summary of the 2014 evidence-based report by expert panel members. *JAMA* **312**, 1033–1048 (2014).
- F. R. Appelbaum, S. J. Forman, R. S. Negrin, K. G. Blume, *Thomas' Hematopoietic Cell Transplantation* (Wiley-Blackwell), 2009.
- F. Bernaudin, G. Socie, M. Kuentz, S. Chevret, M. Duval, Y. Bertrand, J.-P. Vannier, K. Yakouben, I. Thuret, P. Bordignon, A. Fischer, P. Lutz, J.-L. Stephan, N. Dhedin, E. Plouvier, G. Marguerite, D. Bories, S. Verlhac, H. Esperou, L. Coic, J.-P. Vernant, E. Gluckman, Long-term results of related myeloablative stem-cell transplantation to cure sickle cell disease. *Blood* **110**, 2749–2756 (2007).
- M. M. Hsieh, C. D. Fitzhugh, R. P. Weitzel, M. E. Link, W. A. Coles, X. Zhao, G. P. Rodgers, J. D. Powell, J. F. Tisdale, Nonmyeloablative HLA-matched sibling allogeneic hematopoietic stem cell transplantation for severe sickle cell phenotype. *JAMA* **312**, 48–56 (2014).
- J. Bolaños-Meade, E. J. Fuchs, L. Luznik, S. M. Lanzkron, C. J. Gamper, R. J. Jones, R. A. Brodsky, HLA-haploidentical bone marrow transplantation with posttransplant cyclophosphamide expands the donor pool for patients with sickle cell disease. *Blood* **120**, 4285–4291 (2012).
- P. Genovese, G. Schirolli, G. Escobar, T. Di Tomaso, C. Firrito, A. Calabria, D. Moi, R. Mazzieri, C. Bonini, M. C. Holmes, P. D. Gregory, M. van der Burg, B. Gentner, E. Montini, A. Lombardo, L. Naldini, Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235–240 (2014).
- J. Wang, C. M. Exline, J. J. DeClercq, G. N. Llewellyn, S. B. Hayward, P. W.-L. Li, D. A. Shivak, R. T. Surosky, P. D. Gregory, M. C. Holmes, P. M. Cannon, Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat. Biotechnol.* **33**, 1256–1263 (2015).
- B. D. Sather, G. S. Romano Ibarra, K. Sommer, G. Curinga, M. Hale, I. F. Khan, S. Singh, Y. Song, K. Gwiazda, J. Sahni, J. Jarjour, A. Astrakhan, T. A. Wagner, A. M. Scharenberg, D. J. Rawlings, Efficient modification of *CCR5* in primary human hematopoietic cells using a megaTAL nuclease and AAV donor template. *Sci. Transl. Med.* **7**, 307ra156 (2015).
- C. J. Wu, L. Krishnamurti, J. L. Kutok, M. Biernacki, S. Rogers, W. Zhang, J. H. Antin, J. Ritz, Evidence for ineffective erythropoiesis in severe sickle cell disease. *Blood* **106**, 3639–3645 (2005).
- M. C. Walters, M. Patience, W. Leisenring, Z. R. Rogers, V. M. Aquino, G. R. Buchanan, I. A. G. Roberts, A. M. Yeager, L. Hsu, T. Adamkiewicz, J. Kurtzberg, E. Vichinsky, B. Storer, R. Storb, K. M. Sullivan, Stable mixed hematopoietic chimerism after bone marrow transplantation for sickle cell anemia. *Biol. Blood Marrow Transplant.* **7**, 665–673 (2001).
- C. J. Wu, M. Gladwin, J. Tisdale, M. Hsieh, T. Law, M. Biernacki, S. Rogers, X. Wang, M. Walters, D. Zahrhieh, J. H. Antin, J. Ritz, L. Krishnamurti, Mixed haematopoietic chimerism for sickle cell disease prevents intravascular haemolysis. *Br. J. Haematol.* **139**, 504–507 (2007).
- R. Iannone, J. F. Casella, E. J. Fuchs, A. R. Chen, R. J. Jones, A. Woolfrey, M. Amylon, K. M. Sullivan, R. F. Storb, M. C. Walters, Results of minimally toxic nonmyeloablative transplantation in patients with sickle cell anemia and β -thalassemia. *Biol. Blood Marrow Transplant.* **9**, 519–528 (2003).
- M. D. Hoban, G. J. Cost, M. C. Mendel, Z. Romero, M. L. Kaufman, A. V. Joglekar, M. Ho, D. Lumaquin, D. Gray, G. R. Lill, A. R. Cooper, F. Urbinati, S. Senadheera, A. Zhu, P.-Q. Liu, D. E. Paschon, L. Zhang, E. J. Rebar, A. Wilber, X. Wang, P. D. Gregory, M. C. Holmes, A. Reik, R. P. Hollis, D. B. Kohn, Correction of the sickle-cell disease mutation in human hematopoietic stem/progenitor cells. *Blood* **125**, 2597–2604 (2015).
- H. Kim, J.-S. Kim, A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.* **15**, 321–334 (2014).
- M. Jinek, A. East, A. Cheng, S. Lin, E. Ma, J. Doudna, RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013).
- J. A. Doudna, E. Charpentier, The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
- M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- T. J. Cradick, E. J. Fine, C. J. Antico, G. Bao, CRISPR/Cas9 systems targeting β -globin and *CCR5* genes have substantial off-target activity. *Nucleic Acids Res.* **41**, 9584–9592 (2013).
- A. Hendel, R. O. Bak, J. T. Clark, A. B. Kennedy, D. E. Ryan, S. Roy, I. Steinfeld, B. D. Lunstad, R. J. Kaiser, A. B. Wilkens, R. Bacchetta, A. Tsalenko, D. Dellinger, L. Bruhn, M. H. Porteus, Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* **33**, 985–989 (2015).
- X. Huang, Y. Wang, W. Yan, C. Smith, Z. Ye, J. Wang, Y. Gao, L. Mendelsohn, L. Cheng, Production of gene-corrected adult β globin protein in human erythrocytes differentiated from patient iPSCs after genome editing of the sickle point mutation. *Stem Cells* **33**, 1470–1479 (2015).
- M. C. Canver, E. C. Smith, F. Sher, L. Pinello, N. E. Sanjana, O. Shalem, D. D. Chen, P. G. Schupp, D. S. Vinjamur, S. P. Garcia, S. Luc, R. Kurita, Y. Nakamura, Y. Fujiwara, T. Maeda, G.-C. Yuan, F. Zhang, S. H. Orkin, D. E. Bauer, *BCL11A* enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527**, 192–197 (2015).
- M. D. Hoban, D. Lumaquin, C. Y. Kuo, Z. Romero, J. Long, M. Ho, C. S. Young, M. Mojaddidi, S. Fitz-Gibbon, A. R. Cooper, G. R. Lill, F. Urbinati, B. Campo-Fernandez, C. F. Bjurstrom, M. Pellegrini, R. P. Hollis, D. B. Kohn, CRISPR/Cas9-mediated correction of the sickle mutation in human CD34⁺ cells. *Mol. Ther.* 10.1038/mt.2016.148 (2016).
- N. Sun, H. Zhao, Seamless correction of the sickle cell disease mutation of the *HBB* gene in human induced pluripotent stem cells using TALENs. *Biotechnol. Bioeng.* **111**, 1048–1053 (2014).
- S. Ramalingam, N. Annaluru, K. Kandavelou, S. Chandrasegaran, TALEN-mediated generation and genetic correction of disease-specific human induced pluripotent stem cells. *Curr. Gene Ther.* **14**, 461–472 (2014).
- S. Lin, B. T. Staahl, R. K. Alla, J. A. Doudna, Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* **3**, e04766 (2014).
- K. Schumann, S. Lin, E. Boyer, D. R. Simeonov, M. Subramanian, R. E. Gate, G. E. Haliburton, C. J. Ye, J. A. Bluestone, A. Marson, Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10437–10442 (2015).
- X. Liang, J. Potter, S. Kumar, Y. Zou, R. Quintanilla, M. Sridharan, J. Carte, W. Chen, N. Roark, S. Ranganathan, N. Ravinder, J. D. Chesnut, Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J. Biotechnol.* **208**, 44–53 (2015).
- C. D. Richardson, G. J. Ray, M. A. DeWitt, G. L. Curie, J. E. Corn, Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016).
- J. G. Doench, E. Hartenian, D. B. Graham, Z. Tothova, M. Hegde, I. Smith, M. Sullender, B. L. Ebert, R. J. Xavier, D. E. Root, Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* **32**, 1262–1267 (2014).
- C. D. Richardson, G. J. Ray, N. L. Bray, J. E. Corn, Non-homologous DNA increases gene disruption efficiency by altering DNA repair outcomes. *Nat. Commun.* **7**, 12463 (2016).

39. Y. Fu, J. A. Foden, C. Khayter, M. L. Maeder, D. Reyon, J. K. Joung, J. D. Sander, High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826 (2013).
40. P. D. Hsu, D. A. Scott, J. A. Weinstein, F. A. Ran, S. Konermann, V. Agarwala, Y. Li, E. J. Fine, X. Wu, O. Shalem, T. J. Cradick, L. A. Marraffini, G. Bao, F. Zhang, DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
41. S. Kim, D. Kim, S. W. Cho, J. Kim, J.-S. Kim, Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **24**, 1012–1019 (2014).
42. Y. Fu, J. D. Sander, D. Reyon, V. M. Cascio, J. K. Joung, Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
43. W. J. Kent, C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, D. Haussler, The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
44. V. T. Chu, T. Weber, B. Wefers, W. Wurst, S. Sander, K. Rajewsky, R. Kühn, Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**, 543–548 (2015).
45. T. Maruyama, S. K. Dougan, M. C. Truttmann, A. M. Bilate, J. R. Ingram, H. L. Ploegh, Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538–542 (2015).
46. G. H. Hardy, Mendelian proportions in a mixed population. *Science* **28**, 49–50 (1908).
47. I. M. Slaymaker, L. Gao, B. Zetsche, D. A. Scott, W. X. Yan, F. Zhang, Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2015).
48. B. P. Kleinstiver, V. Pattanayak, M. S. Prew, S. Q. Tsai, N. T. Nguyen, Z. Zheng, J. K. Joung, High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
49. M.-C. Giarratana, L. Kobari, H. Lapillonne, D. Chalmers, L. Kiger, T. Cynober, M. C. Marden, H. Wajcman, L. Douay, Ex vivo generation of fully mature human red blood cells from hematopoietic stem cells. *Nat. Biotechnol.* **23**, 69–74 (2005).
50. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
51. F. Notta, S. Doulatov, E. Laurenti, A. Poepl, I. Jurisica, J. E. Dick, Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science* **333**, 218–221 (2011).
52. A. C. Drake, Q. Chen, J. Chen, Engineering humanized mice for improved hematopoietic reconstitution. *Cell. Mol. Immunol.* **9**, 215–224 (2012).
53. J. Corrigan-Curay, M. O'Reilly, D. B. Kohn, P. M. Cannon, G. Bao, F. D. Bushman, D. Carroll, T. Cathomen, J. K. Joung, D. Roth, M. Sadelain, A. M. Scharenberg, C. von Kalle, F. Zhang, R. Jambou, E. Rosenthal, M. Hassani, A. Singh, M. H. Porteus, Genome editing technologies: Defining a path to clinic. *Mol. Ther.* **23**, 796–806 (2015).
54. H. Li, V. Haurigot, Y. Doyon, T. Li, S. Y. Wong, A. S. Bhagwat, N. Malani, X. M. Anguela, R. Sharma, L. Ivanciu, S. L. Murphy, J. D. Finn, F. R. Khazi, S. Zhou, D. E. Paschon, E. J. Rebar, F. D. Bushman, P. D. Gregory, M. C. Holmes, K. A. High, In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**, 217–221 (2011).
55. R. S. Franco, Z. Yasín, M. B. Palascak, P. Ciralo, C. H. Joiner, D. L. Rucknagel, The effect of fetal hemoglobin on the survival characteristics of sickle cells. *Blood* **108**, 1073–1076 (2006).
56. R. Iannone, L. Luznik, L. W. Engstrom, S. L. Tennessee, F. B. Askin, J. F. Casella, T. S. Kickler, S. N. Goodman, A. L. Hawkins, C. A. Griffin, L. Noffsinger, E. J. Fuchs, Effects of mixed hematopoietic chimerism in a mouse model of bone marrow transplantation for sickle cell anemia. *Blood* **97**, 3960–3965 (2001).
57. B. Li, C. N. Dewey, RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

Acknowledgments: We thank C. Jeans of the QB3 MacroLab at University of California (UC), Berkeley, for expression and purification of the Cas9 protein variants used in this study. We thank S. McDevitt and the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, which is supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303. We thank H. Turner and M. J. Kim [Children's Hospital Oakland Research Institute (CHORI)] for advice and assistance with mouse work and L. Parker (CHORI) for advice and assistance with flow cytometry. We thank J. Doudna and M. Botchan (UC Berkeley), and B. Lubin and E. Penhoet (UCSF Benioff Children's Hospital Oakland) for their contributions to the initiation of this project. **Funding:** J.E.C., M.A.D., D.C., N.L.B., T.M., and T.W. were supported by the Li Ka Shing Foundation. M.A.D. was a California Institute for Regenerative Medicine postdoctoral fellow under training program TG2-01164. D.C. is supported by NIH grant R01 GM078571 and was supported by the Siebel Scholars Fund at Berkeley. D.I.K.M., D.B., and S.-J.H. were supported by NIH grants 5R21AA022753 and 1R56ES022377-01; D.I.K.M., W.M., M.C.W., and D.P.M. were supported by the Jordan Family Fund at CHORI. D.B.K. and F.U. were supported by a research grant from the Doris Duke Charitable Foundation to D.B.K. (Innovations in Clinical Research Award #2013158). **Author contributions:** J.E.C., D.C., M.C.W., and D.I.K.M. conceived the project. J.E.C., D.I.K.M., M.A.D., D.C., D.P.M., D.B., W.M., D.B.K., and J.R.B. designed the experiments. W.M., M.A.D., D.C., T.W., S.-J.H., and F.U. performed the experiments. N.L.B., T.M., S.-J.H., M.A.D., W.M., and D.C. analyzed the data. M.A.D., J.E.C., D.I.K.M., M.C.W., and D.C. wrote the manuscript. **Competing interests:** J.R.B. is an employee of Bio-Rad Inc. J.E.C. is an inventor on a patent submitted by UC Berkeley that covers asymmetric HDR donors, which were used during this work. All other authors declare that they have no competing interests. **Data and materials availability:** All NGS data have been deposited at the National Center for Biotechnology Information Sequence Read Archive under accession no. SRP090223. Tabulated data for Figs. 1 to 4, including references to the original NGS data, are presented in table S4.

Submitted 21 April 2016
Accepted 22 September 2016
Published 12 October 2016
10.1126/scitranslmed.aaf9336

Citation: M. A. DeWitt, W. Magis, N. L. Bray, T. Wang, J. R. Berman, F. Urbinati, S.-J. Heo, T. Mitros, D. P. Muñoz, D. Boffelli, D. B. Kohn, M. C. Walters, D. Carroll, D. I. K. Martin, J. E. Corn, Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci. Transl. Med.* **8**, 360ra134 (2016).



Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells

Mark A. DeWitt, Wendy Magis, Nicolas L. Bray, Tianjiao Wang, Jennifer R. Berman, Fabrizia Urbinati, Seok-Jin Heo, Therese Mitros, Denise P. Muñoz, Dario Boffelli, Donald B. Kohn, Mark C. Walters, Dana Carroll, David I. K. Martin and Jacob E. Corn (October 12, 2016)

Science Translational Medicine **8** (360), 360ra134. [doi: 10.1126/scitranslmed.aaf9336]

Editor's Summary

Hammering out the sickle cell mutation

Sickle cell disease is a genetic disorder caused by a mutation in one of the hemoglobin genes, which causes deformation of red blood cells and results in occlusion of blood vessels, severe pain crises, and progressive organ injury. To correct the mutation that causes this disease, DeWitt *et al.* modified hematopoietic stem cells from sickle cell disease patients using a CRISPR/Cas9 gene editing approach. The authors showed that the corrected cells successfully engrafted in a mouse model and produced enough normal hemoglobin to have a potential clinical benefit in the setting of sickle cell disease.

The following resources related to this article are available online at <http://stm.sciencemag.org>. This information is current as of October 20, 2016.

Article Tools	Visit the online version of this article to access the personalization and article tools: http://stm.sciencemag.org/content/8/360/360ra134
Supplemental Materials	" <i>Supplementary Materials</i> " http://stm.sciencemag.org/content/suppl/2016/10/07/8.360.360ra134.DC1
Related Content	The editors suggest related resources on <i>Science's</i> sites: http://stm.sciencemag.org/content/scitransmed/6/267/267ra175.full http://stm.sciencemag.org/content/scitransmed/4/123/123ra26.full http://science.sciencemag.org/content/sci/354/6309/189.full
Permissions	Obtain information about reproducing this article: http://www.sciencemag.org/about/permissions.dtl

Science Translational Medicine (print ISSN 1946-6234; online ISSN 1946-6242) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Copyright 2016 by the American Association for the Advancement of Science; all rights reserved. The title *Science Translational Medicine* is a registered trademark of AAAS.

Supplementary Materials for
**Selection-free genome editing of the sickle mutation in human adult
hematopoietic stem/progenitor cells**

Mark A. DeWitt, Wendy Magis, Nicolas L. Bray, Tianjiao Wang, Jennifer R. Berman,
Fabrizia Urbinati, Seok-Jin Heo, Therese Mitros, Denise P. Muñoz, Dario Boffelli,
Donald B. Kohn, Mark C. Walters, Dana Carroll*, David I. K. Martin*, Jacob E. Corn*

*Corresponding author. Email: jcorn@berkeley.edu (J.E.C.); dimartin@chori.org (D.I.M.);
dana@biochem.utah.edu (D.C.)

Published 12 October 2016, *Sci. Transl. Med.* **8**, 360ra134 (2016)
DOI: 10.1126/scitranslmed.aaf9336

The PDF file includes:

Discussion

Materials and Methods

Fig. S1. Additional data showing editing of K562 cells with the Cas9 RNP.

Fig. S2. Genomic context of predicted off-target cut sites for the G10 RNP.

Fig. S3. Additional data showing editing of CD34⁺ HSPCs with the Cas9 RNP.

Fig. S4. Qualitative detection of chromosomal translocations by overamplification PCR.

Fig. S5. Confirmation of efficient correction of SCD HSPCs by ddPCR.

Fig. S6. Additional data on engraftment and editing of HSPCs in NSG mice.

Table S1. Oligonucleotides used in this study.

Table S2. Indel mutations at cancer-associated genes in K562 cells and HSPCs edited with the trG10 RNP, as compared to unedited cells.

Table S3. Estimates of cellular editing in mice engrafted with edited HSPCs.

References (55–57)

Other Supplementary Material for this manuscript includes the following:
(available at

www.sciencetranslationalmedicine.org/cgi/content/full/8/360/360ra134/DC1)

Table S4. Tabulated data from Figs. 1 to 4 (provided as an Excel file).

Supplementary Discussion

How much gene correction is sufficient for detectable clinical benefit in SCD patients? It has long been known that sickle RBCs have a 5-10-fold reduced lifespan in circulating blood (depending on fetal hemoglobin level) (55). Furthermore, sickle RBCs mature ineffectively compared to wild-type RBCs, as evidenced by selective loss of recipient (SCD) erythroblasts in the bone marrow of SCD patients with donor chimerism after allogeneic HSCT (17). These two factors imply that only a small minority of corrected HSCs will generate the preponderance of the circulating RBCs. Indeed, observations of SCD HSCT recipients with donor chimerism after undergoing nonmyeloablative conditioning indicate that when only 2-5% of bone marrow cells are from a healthy donor, >50% of hemoglobin is HbA, and these patients experience a marked decrease in clinical events (20). In parallel studies of lethally-irradiated mice transplanted with congenic mixtures of sickle and healthy marrow, similarly low levels of chimerism (~3-7%) result in substantial improvement in pathology and hematology (56).

Although we edited 2.3% of alleles, this may translate into nearly 2-fold more corrected cells, if the alleles assort independently as they do in edited CD34+ cells *in vitro* (Table 1). Consider the mixture of genotypes produced by editing a population of HSCs with a β^S/β^S genotype. Some alleles will not be targeted (remaining β^S), some will be targeted and receive an indel (β^{indel}), and some will be subjected to HDR and corrected to WT (β^{WT}). The frequency of each allele ($\beta^S, \beta^{indel}, \beta^{WT}$) in the population is established by NGS genotyping as we have shown, and can be used to calculate the frequency of cells with each biallelic genotype. We can use the Hardy-Weinberg principle to calculate the proportions of genotypes at various frequencies of HDR. It can be readily shown that when the edited allele is relatively rare (approximately 5%), roughly 2-fold as many cells have one edited allele (approximately 10 %) (46). Table S3 presents the calculated percentages of cells with at least one HDR allele using the Hardy-Weinberg equations, based on our NGS analysis of the marrow of edited WT HSPCs engrafted in NSG mice (Figure 4C). Because there is only a single nucleotide difference between WT and SCD HSPCs, we anticipate a similar result for correction of SCD HSPCs using our method.

The bar for clinical relevance for an autologous gene correction approach to SCD treatment may be as low as 3-10% of cells, or 1.5-5% of alleles. Establishing a clear pre-clinical target for SCD correction by gene editing within this range remains an urgent need for the gene therapy community.

Materials and Methods

Study Design. The goal of this study was to quantify editing of the SCD mutation in CD34+ HSPCs and HSCs in a variety of contexts, and in particular engraftment of

edited cells in immunocompromised mice. Editing of the long-term population of cells was measured by engraftment of 3 pools of cells from 2 healthy donors in 7 NSG mice. All mice that were engrafted with edited cells were used in the analysis, excluding 4 mice that died before the termination of the experiment, 2-10 weeks after injection. No blinding was used.

Synthesis of Cas9 RNPs. Cas9 RNP component synthesis and assembly was carried out based on published work (33). Cas9 was prepared by the UC Berkeley Macro Lab using a published protocol (33). Cas9 was stored and diluted in sterile-filtered Cas9 Buffer (20 mM HEPES pH 7.5, 150 mM KCl, 1 mM MgCl₂, 10% glycerol, 1 mM TCEP). TCEP was added to storage buffer only. sgRNA was synthesized by assembly PCR and *in vitro* transcription. A T7 RNA polymerase substrate template was assembled by PCR from a variable 57-59 nt primer containing T7 promoter, variable sgRNA guide sequence, and the first 15 nt of the non-variable region of the sgRNA (T7FwdVar primers, 10 nM, table S1), and an 83 nt primer containing the reverse complement of the invariant region of the sgRNA (T7RevLong, 10 nM), along with amplification primers (T7FwdAmp, T7RevAmp, 200 nM each). These primers anneal in the first cycle of PCR and are amplified in subsequent cycles. Phusion high-fidelity DNA polymerase was used for assembly (New England Biolabs, Inc.). Assembled template was used without purification as a substrate for *in vitro* transcription by T7 RNA polymerase using the HiScribe T7 High Yield RNA Synthesis kit (New England Biolabs, Inc.). Resulting transcription reactions were treated with DNase I, and RNA was purified either by treatment with a 5X volume of homemade solid phase reversible immobilization (SPRI) beads (comparable to

Beckman-Coulter AMPure beads) or the Qiagen RNeasy purification kit, and elution in DEPC-treated water. sgRNA concentrations were determined by fluorescence using the Qubit dsDNA HS assay kit (Life Technologies, Inc). Cas9 RNP was assembled immediately before electroporation of target cells (see below). To electroporate a 20 μ L cell suspension (150,000-200,000 cells) with Cas9 RNP, a 5 μ L solution containing a 1.2-1.3X molar excess of sgRNA in Cas9 buffer was prepared. A 5 μ L solution containing 75-200 pmol purified Cas9 in Cas9 buffer was prepared and added to the sgRNA solution slowly over \sim 30 seconds, and incubated at room temperature for $>$ 5 minutes before mixing with target cells.

Editing HBB in K562 cells. K562 cells were obtained from the UC Berkeley Tissue Culture facility, and cultured in IMDM with 10% FCS, penicillin-streptomycin (100 units/mL and 100 μ g/mL), plus 2 mM GlutaMax. K562 cells were edited by electroporation using the Lonza 4d nucleofector and manufacturer's protocols (Lonza, Inc.). For each electroporation, 150,000-200,000 late log-phase K562 cells were pelleted (100 x g, 5 minutes) and re-suspended in 20 μ L Lonza SF solution. 20 μ L cells, 10 μ L Cas9 RNP containing the desired guide (see above, and table S1), and 1 μ L 100 μ M ssDNA template carrying the desired edit at the SCD SNP were mixed and electroporated using the recommended protocol for K562 cells. After electroporation, K562 cells were incubated for 10 minutes in the cuvette, transferred to 1 mL of the above medium, and cultured for 48-72 h before genomic DNA extraction and genotyping.

Editing HBB in primary human CD34+ HSCs. Cryopreserved WT human mobilized peripheral blood CD34+ HSPCs were purchased from Allcells, Inc. SCD CD34+ HSPCs

were prepared by Allcells Inc. from whole blood discarded during exchange transfusion of SCD patients at Benioff Children's Hospital Oakland, and cryopreserved. To edit HSCs, ~1 million HSPCs were thawed and cultured in StemSpan SFEM medium supplemented with StemSpan CC110 cocktail (StemCell Technologies) for 24 h before electroporation with Cas9 RNP. To electroporate HSPCs, 100,000-200,000 were pelleted (200 x g, 10 minutes) and resuspended in 20 μ L Lonza P3 solution, and mixed with 10 μ L Cas9 RNP and 1 μ L 100 μ M ssDNA template programming the desired edit. This mixture was electroporated using the Lonza 4d nucleofector and either of two protocols ("1" : DO100, "2": ER100). Electroporated cells were recovered in the cuvette with 200 μ L StemSpan SFEM/CC110 for 10-15 minutes and transferred to culture in 1 mL StemSpan SFEM/CC110 for 48 hours after electroporation. Half of the cells were removed for genotyping ("un-expanded HSPCs"), and the remaining cells were transferred to erythroid expansion medium (StemSpan SFEM II with StemSpan Erythroid Expansion supplement, [StemCell Technologies]) for 5 additional days before genotyping of expanded cells. To determine the zygosity of edits in HSPCs (Table 1), a culture of edited HSPCs (2 days after electroporation) in SFEM/CC110 was diluted to 10 cells/mL in SFEM with erythroid expansion medium, and 100 μ L per well was plated into 96-well plates. Clones were grown for 14 days, then 96 clones were pelleted and their genomic DNA extracted for genotyping by NGS.

Editing HBB with newly-developed Cas9 Variants in CD34+ HSPCs. Plasmids encoding Cas9 variants HF1 and espCas91.1 were generated from the previously-published wild-type Cas9 construct (33, 47, 48), and expressed and purified from *E. coli* using

an identical protocol by the UC Berkeley Macro Lab. HSPCs were prepared and edited essentially identically as with wild-type Cas9: 75 pmol Cas9 variant complexed with guide **G10** was delivered to 150,000 CD34+ HSPCs with no HDR donor. Edited HSPCs were cultured for 5 days in expansion conditions before genomic DNA extraction and subsequent analysis of indel formation at selected targets by NGS.

Editing HSPCs before injection in NSG mice. For each mouse, 750,000-1,000,000 CD34+ HSPCs were edited. Editing was performed using a scaled-up reaction volume from *in vitro* experiments above. 750,000-1,000,000 HSPCs were thawed and recovered for 24 h before editing. Both stimulated and unstimulated conditions used StemSpan SFEM medium supplemented with StemSpan CC110 cocktail; stimulation was for 3 days, and unstimulated were treated as above. Before editing, HSPCs were pelleted and resuspended in 100 μ L P3. 500 pmol Cas9 RNP was prepared in 50 μ L Cas9 buffer. RNP, cells, and 5 μ L **T88-107** 100 μ M donor template were mixed in a large-sized cuvette and electroporated using the Lonza 4d Nucleofector and protocol "2" (ER100). Cells were recovered by addition of 400 μ L StemSpan SFEM/CC110 to the cuvette, then cultured in recovery medium for 24 hours at a density <1 million cells/mL. For the first and third mouse experiment (2 mice and 3 mice, respectively), cells were cultured in SFEM/CC110 for 1 day before editing. For the second experiment (2 mice), cells were cultured for 3 days before editing.

Xenografting of human CD34+ HSPCs into NSG mice. NSG mice (JAX) were maintained in clean conditions. 7-week-old female mice were subjected to 2.5Gy X-irradiation 4

hours before tail vein injection of edited cells under isoflurane anesthesia. At 5 weeks and 8 weeks after injection, 200 μ L blood was obtained from the submandibular vein under isoflurane anesthesia. 16 weeks after injection, mice were euthanized, and bone marrow and spleen were recovered for analysis.

Flow cytometry. Cells were prepared from peripheral blood of NSG mice by lysis in Qiagen buffer EL for 40 seconds (200 μ L blood plus 2 mL EL) followed by quenching with 5-6 volumes of cold PBS before staining for flow cytometry. Cells were prepared from bone marrow and spleen of NSG mice by resuspension of cells in 10 mL PBS, pelleting, lysis in 2 mL Qiagen buffer EL for 40 seconds, and quenching of lysis with 5-6 volumes of cold PBS. Pelleted lysed cells were stained with antibodies to the indicated cell surface markers, and analyzed on a BD FACS Fortessa flow cytometer. Flow cytometry data were analyzed using the FlowJo software package. For genotyping of the sorted pools, mouse bone marrow was sorted by FACS for the indicated cell surface marker. Sorted cells were pelleted and genotyped by NGS as described above. The following antibodies were used, all from BD Pharmingen: APC Rat anti-Mouse CD45,(561018, clone 30-F11), FITC Mouse anti-Human CD45 (555482, clone HI30), V450 Mouse anti-Human CD45 (560368, clone HI30), BV421 Mouse anti-Human CD3 (563797 clone SK7), BV421 Mouse anti-Human CD56 (562752, clone NCAM16.2), FITC Mouse anti-Human CD19 (340409, clone SJ25C1), FITC Mouse anti-Human CD33 (561818, clone HIM3-4), BV421 Mouse anti-Human CD34 (562577, clone 581).

Differentiation of HSCs into erythroblasts. After electroporation, cells were recovered and placed in StemSpan SFEM/CC110 for 24 hours. They were then transferred to

StemSpan SFEM II with StemSpan Erythroid Expansion supplement and grown for 7 days with maintenance of optimal density (200,000-1,000,000 cells/mL). The resulting erythroid progenitors were transferred to StemSpan SFEM II with 3 U/ml erythropoietin (Life Technologies), 3% normal human AB serum (Sigma), and 1 μ M mifepristone (Sigma). They were then cultured for a further 5 days with daily monitoring of cell morphology by Wright-Giemsa staining; at the conclusion, the majority of the cells were enucleated. Cells were then lysed in hemolysate reagent (Helena Laboratories) for preparation of hemoglobin for HPLC, or RNA extracted with the Direct-zol RNA Kit (Zymo Research).

Genotyping of edited cells. Pools of edited cells (K562 cells, CD34+ HSPCs, and nucleated cells from mouse blood) were lysed and their genomic DNA extracted using QuickExtract solution (Epicentre Inc.) to a final concentration of \sim 5,000 haploid genomes/ μ L. A 286 bp region around the SCD SNP and Cas9 cut site was amplified by PCR using Q5 DNA polymerase (New England Biolabs, Inc), and primers 1F and 1R (table S1). NHEJ-mediated indel formation within each pool was estimated by T7 endonuclease digestion using manufacturer's protocols (New England Biolabs, Inc). HDR-mediated editing was assessed by restriction digest with either *SfcI* (for WT->SCD edits) or *Hpy188III* (for SCD->WT edits). For analysis of editing by next-generation sequencing (NGS), initial PCR products were cleaned with SPRI beads (1.8X), followed by amplification of 20-50 ng product in a second 6-8 cycle PCR using primers 4F and 1R, generating amplicons of an appropriate length for NGS, followed by Illumina TruSeq adaptor ligation and purification (Bio Scientific DNaseq kit or Illumina TruSeq Nano HT kit). To avoid contamination of

ssDNA donor sequence, 1F and 1R amplify outside the genomic region matching donor, and 4F does not anneal to the ssDNA donor. Libraries from 12-96 pools of edited cells were pooled and run on a single MiSeq lane, using a paired-end 150 cycle read. HDR-mediated editing of the SCD SNP was also assessed by droplet digital PCR (ddPCR, QX200, Bio-Rad Laboratories, Inc.). ddPCR assays (1x assay: 900 nM primers, and 250 nM each probe) used TaqMan probes specific for unedited and HDR-edited alleles, with one primer positioned outside of the template matching region of HBB to prevent amplification of donor template. Assays were run using ddPCR supermix for probes (no dUTP) with the following thermal cycling protocol: 1) 95°C 10 min; 2), 94°C 30 s; 3), 55°C 1 min; 4), 72°C 2 min; 5) repeat steps 2-4 39 times; 6), 98°C 10 min, with all the steps ramped by 2°C/s.

NGS data analysis. 20 million MiSeq reads were de-multiplexed and analyzed using a custom analysis workflow written in Python. Each sample contained >75,000 reads, generally much more. For each sample, reads were called as “indel”, “HDR”, or “unedited”. Any read containing an indel within a window of 12-16 bases around the predicted cut site was called as “indel”, and remaining reads were called as “unedited” or “HDR” based whether they matched either the unedited sequence or the ssDNA donor sequence at the SCD SNP. To assess incorporation of HBD coding sequence into HBB, four single-nucleotide differences between HBB and HBD were used. If all four HBD differences were found, that allele was called as HBD. Indel alleles were excluded from this analysis.

HPLC analysis of edited SCD HSPCs. HPLC analysis was performed as previously described(21). Briefly, edited SCD HSPCs differentiated into erythroblasts were

harvested and lysed in Hemolysate reagent (Helena Laboratories). Cell lysates were characterized by HPLC (Infinity 1260, Agilent) using a weak cation-exchange column (PolyCAT A, PolyLC, Inc.). Analysis and peak integration was performed using OpenLAB CDS Chemstation software. HbF, HbA, HbS, HbC (FASC) Reference Material (Trinity Biotech) was used to define the elution time of common hemoglobins.

RNA-seq analysis of edited SCD HSPCs. Total RNA from SCD erythroblasts ($\sim 5 \times 10^6$, differentiated in vitro as described above) was isolated with the Direct-zol RNA Kit (Zymo Research). RNA integrity was checked on an Agilent Bioanalyzer; cDNA was synthesized from this RNA following the Smart-seq2 method, and fragmented with the Covaris apparatus. From the Covaris fragments, indexed libraries were constructed with the ThruPlex-FD prep kit (Rubicon Genomics) and sequenced on an Illumina HiSeq 2500 sequencer for 100 cycles (single read) at the Berkeley GSL. Resulting RNA-seq reads for each sample were quantified against version 80 of the human Ensembl annotation using the program kallisto (50) with default parameters, yielding the relative abundance of each mRNA in transcripts-per-million (57), which was then normalized for comparison between samples.

Analysis of indel formation in cancer-associated genes. To look for indel formation at cancer-associated genes in edited cells, large pools of K562 cells and HSPCs were edited with **trG10** (1000 pmol RNP, 1 million cells) as above, and cultured for 3 days (K562 cells) or 5 days (HSPCs, in SFEM/CC110). Genomic DNA was purified from these cells and untreated cells using the Qiagen Blood and Tissue DNA Extraction Kit (4 genomic DNA samples). Cancer-associated exons and SNPs were enriched and

prepared for Illumina sequencing using the Illumina TruSight Cancer capture kit according to the manufacturer's instructions, then sequenced on an Illumina HiSeq 4000 sequencer, 2x150 paired-end read. Resulting reads were analyzed for indel mutations compared to reference (human genome hg19) sequence using the MuTect2 algorithm. Indels in either sample that passed filters for sequencing errors and significance are listed in table S2.

Detection of translocation events by over-amplification PCR. Genomic DNA extracts of K562 cells and HSPCs edited with the **trG10** RNP were amplified for 35 cycles using primers for the on- and off-target editing sites indicated. All four combinations of primers for each site (two on- and two off-target) were used, and the products were visualized on a 2% agarose gel. The predicted translocation amplicon lengths were estimated from the predicted cut sites.

Statistical Analysis. All errors indicated are expressed as the mean \pm standard deviation of at least three replicates, except for experiments with a single replicate. All data from Figures 1D-E, 2, 3A, 3C, 3D, 4A, 4C, and 4D are provided in tabular form in table S4.

Supplementary Figures

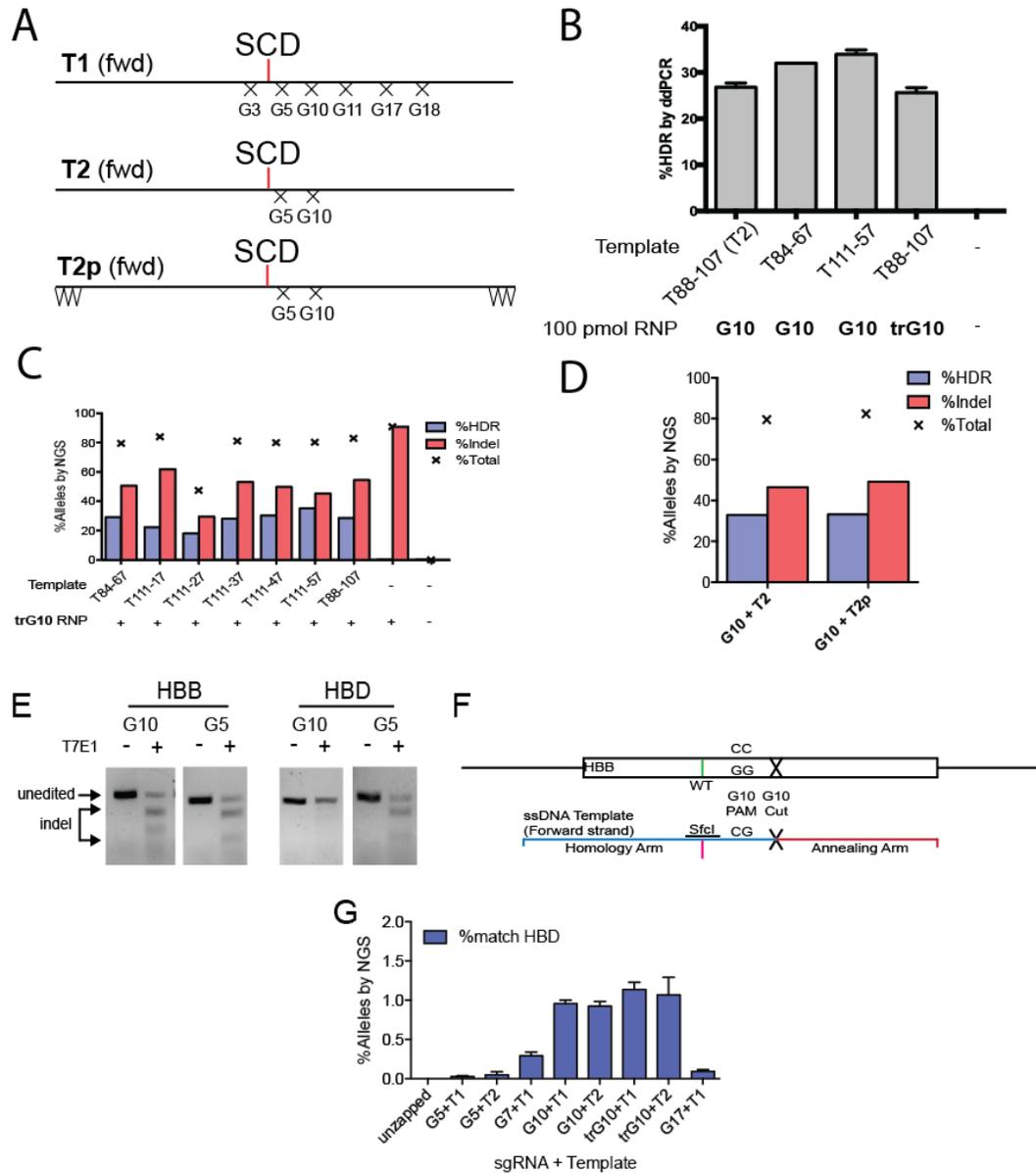


Figure S1. Additional data showing editing of K562 cells with the Cas9 RNP. A) Initial templates are 195 nt long, and edit the WT SNP to SCD. To prevent re-cutting, template 1 encodes silent mutations for the PAMs of all the sgRNAs tested. Templates **T2** and **T2p** have PAM mutations only for **G10** and **G5**. Template **T2p** has three 3' and three 5' phosphorothioate linkages to prevent degradation by

endogenous exonucleases. B) Genomic DNA from K562 cells edited with indicated RNPs and templates were analyzed by ddPCR using probes and primers as indicated in Materials and Methods. Error bars are SEM of 3 independent biological replicates, each consisting of 3 technical replicates. Template **T88-107** is identical to **T2**. See text for details. C) Editing outcomes for K562 cells with edited **trG10** sgRNA as indicated, along with the indicated asymmetric template, defined by the lengths of the homology and annealing arms, respectively (homology length-annealing length, see panel F). Editing outcomes were assessed by NGS. D) Editing outcomes of K562 cells edited with unprotected template **T2** and phosphorothioate-protected **T2p**, analyzed by NGS. E) T7 endonuclease 1 assay of PCR amplicons from the HBB SCD region or the corresponding region in HBD, from K562 cells edited with indicated RNPs, analyzed on a 2% agarose gel. Both the **G5** and **G10** RNPs cut efficiently at HBB, and **G5** also cuts at HBD. F) Design of ssDNA templates, relative to the **G10** cut site. Templates consist of a shorter “annealing” arm that anneals to the strand liberated by RNP binding, and a longer “homology” arm that drives incorporation of the desired edit. G) Conversion of HBB coding sequence to HBD at the HBB locus, assessed by NGS.

G10 Predicted Off-Target

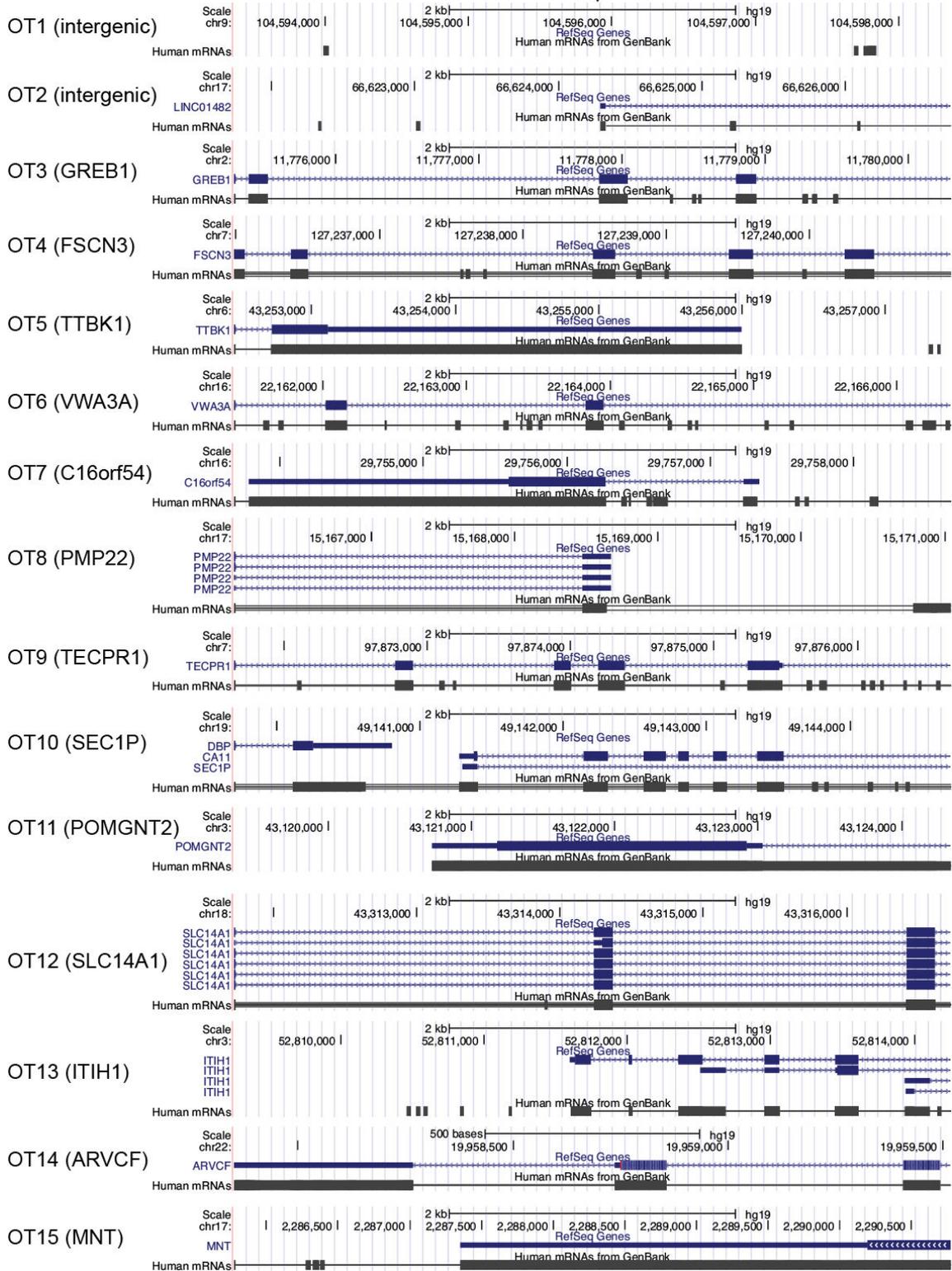


Figure S2. Genomic context of predicted off-target cut sites for the G10 RNP.

Sites were selected using the online CRISPR-design tool, according to criteria discussed in the text. Screenshots of the 5 kb region near the off-target site were prepared using the UCSC genome browser.

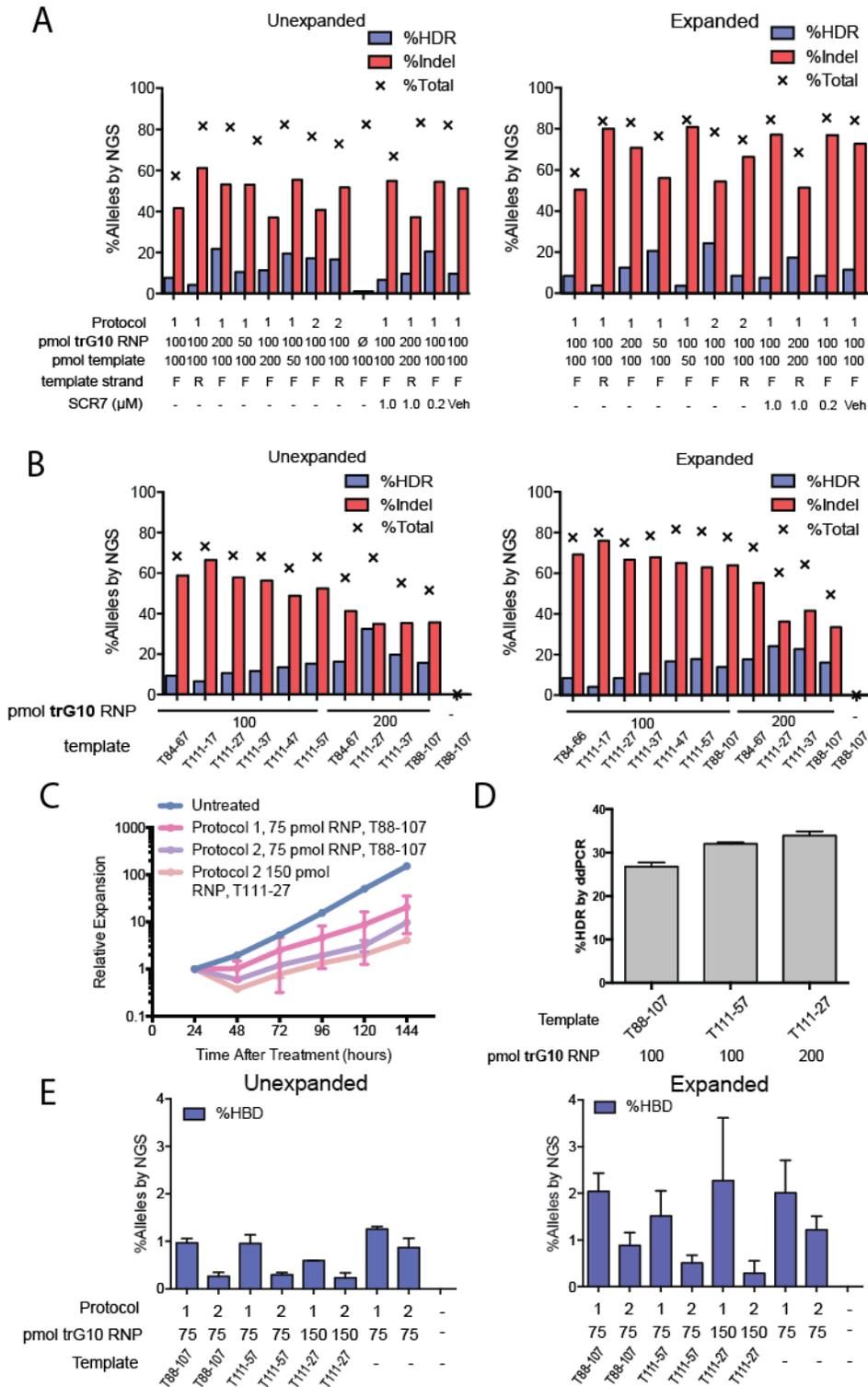


Figure S3. Additional data showing editing of CD34⁺ HSPCs with the Cas9 RNP.

A) Editing outcomes of un-expanded HSPCs (cultured for 2 days after editing in HSC

expansion medium) and erythroid-expanded (cultured for 5 additional days in erythroid expansion medium) HSPCs after electroporation with the indicated doses of templates, **trG10** RNP, and Scr7 (NHEJ inhibitor), and either protocol 1 (Lonza DO100) or protocol 2 (Lonza ER100). Either forward-strand matching (F) or reverse strand-matching (R) templates (template **T88-107**, and its reverse complement) were provided as indicated. B) Editing outcomes of HSPCs cultured in non-expanding and erythroid-expanding conditions, assessed by NGS. CD34+ HSPCs were edited with indicated templates and RNPs, and protocol 2. C) Viable cell counts of HPSCs 24-144 hours after treatment with the indicated electroporation protocol and dose of RNP (except untreated cells). Error bars indicate s.d. of 3 biological replicates for each condition. D) Genotyping of HSPCs by ddPCR. HSPCs (3 biological replicates) were edited with the indicated doses of **trG10** RNP, along with the indicated asymmetric template, and cultured in non-expanding conditions for 2 days before genotyping by ddPCR. Error bars indicate the SEM of the means from three technical replicates of three biological replicates, from a single healthy donor. E) Conversion of coding sequence of HBB to HBD, for HSPCs edited and cultured as in Figure 2A, assessed by NGS.

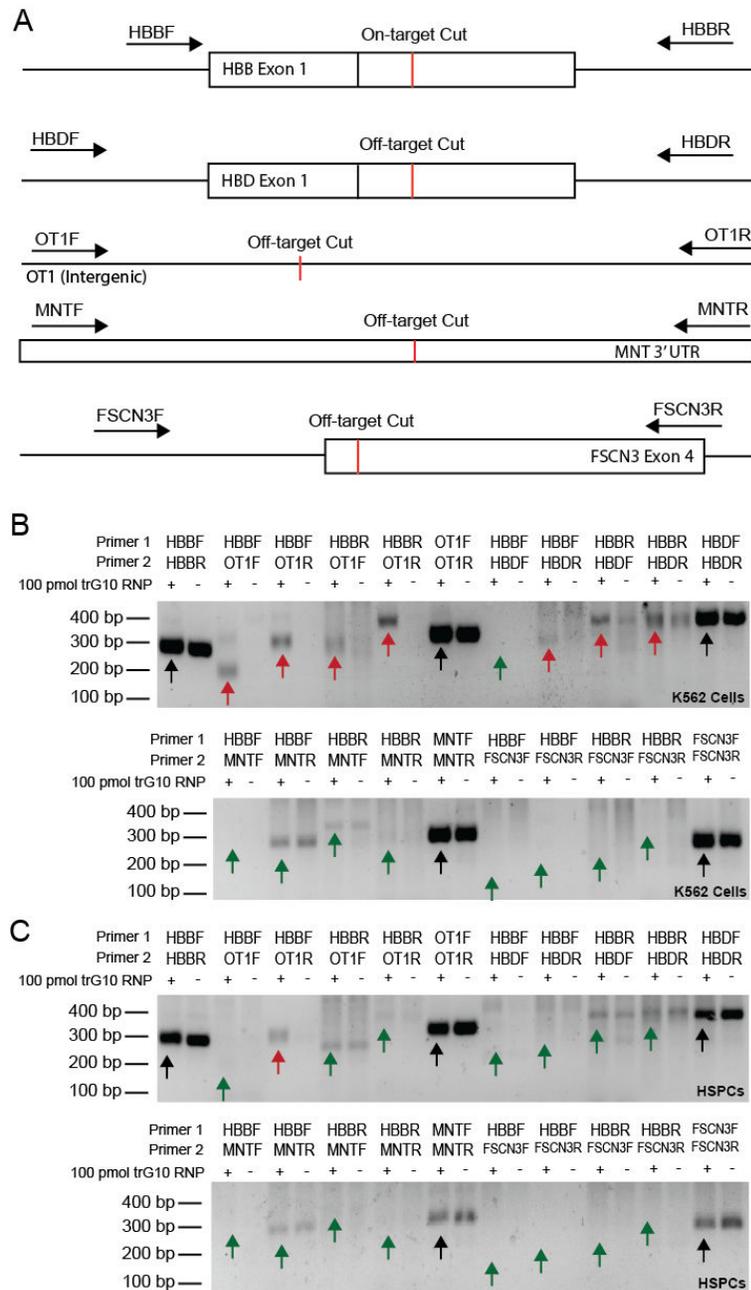


Figure S4. Qualitative detection of chromosomal translocations by overamplification PCR. A) Schematic depicting primers used to detect chromosomal translocations between HBB (on-target), HBD, OT1 (intergenic), FSCN3, and MNT. B-C) 2% agarose gel depicting translocations between on- and off-target sites in K562 cells (B) and HSPCs (C) edited with the **trG10** RNP, compared to

untreated cells. Arrows indicate trG10 presence (red arrows) or absence (green arrows) of **trG10**-dependent translocations, and positive controls (black arrows). Translocations between HBB and both OT1 and HBD are observed in K562 cells. One translocation event between OT1 and HBB is observed in HSPCs, but not between HBB and the other sites.

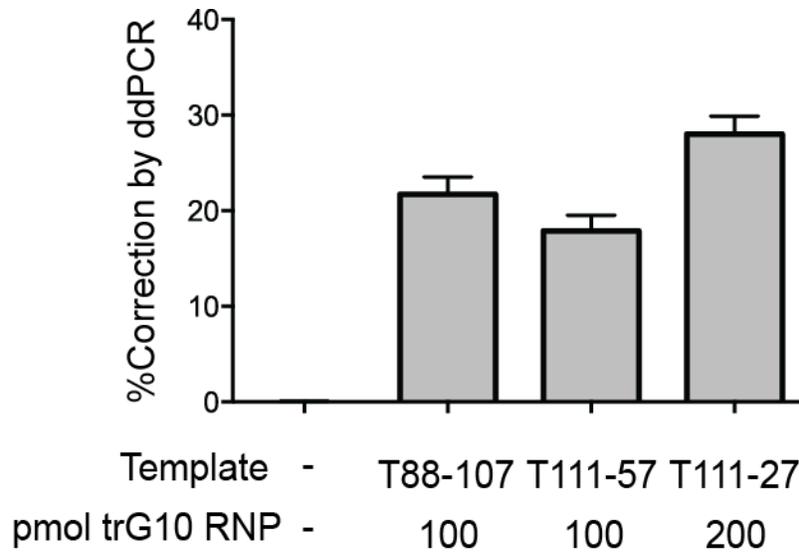


Figure S5. Confirmation of efficient correction of SCD HSPCs by ddPCR. HSPCs from SCD patients were edited with indicated templates and the **trG10** RNP, and cultured for 2 days in non-expanding conditions (identical to conditions in Fig. 3). Editing was then assessed by ddPCR of genomic DNA extracted from edited cells, using TaqMan probes for the SCD (unedited) and WT (edited) alleles. Data are from n = 3 technical replicates (mean ± s.d.).

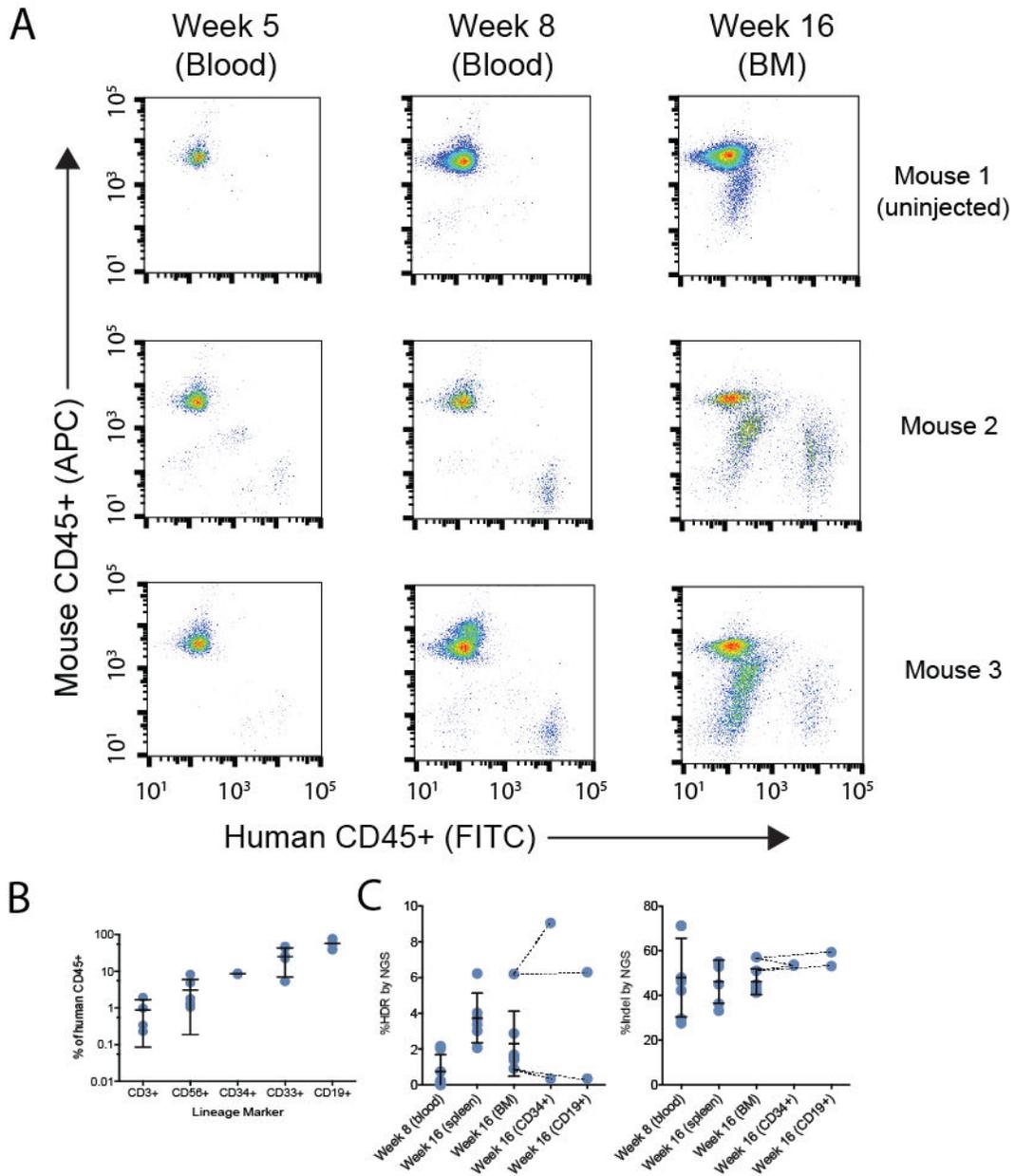


Figure S6. Additional data on engraftment and editing of HSPCs in NSG mice. A)

FACS plots depicting engraftment of edited HSPCs in NSG mice. Mice were engrafted with either no cells (Mouse 1), or edited HSPCs (Mice 2-3). At indicated times, engraftment in either blood or bone marrow (BM) was determined by immunostaining and FACS analysis with human anti-CD45-FITC and mouse anti-CD45-APC. B) Lineage characterization of engrafted human cells. Bone marrow from

Week 8 (blood) and Week 16 (spleen, BM, CD34+, CD19+) were analyzed by NGS for HDR and indel percentages.

mice was harvested 16 weeks after injection and stained with antibodies against indicated human-specific markers along with human CD45. Error bars represent mean \pm s.d. of four mice. C) Editing data as in Figure 4C-D, including indel and HDR within the CD34+ (progenitor) and CD19+ (B lymphocyte) compartments of the marrow of two mice, 16 weeks after injection.

Supplementary Tables

Table S1. Oligonucleotides used in this study.

Type	Name	Sequence
Amplification of SCD SNP region in human HBB	1F	gggcagagccatctattgctta
	1R	tgggaaaaatagaccaataggcagag
	4F	actgtgttcactagcaacctcaa
Assembly of T7 polymerase substrates for sgRNA synthesis by in vitro transcription (UPPERCASE = guide)	FwdVar-G3	ggatcctaatacgcactcactatagCAGACTTCT CCACAGGAGTCgttttagagctagaa
	FwdVar-trG3	ggatcctaatacgcactcactataGACTTCTCCA CAGGAGTCgttttagagctagaa
	FwdVar-G5	ggatcctaatacgcactcactatagCATGGTGCA CCTGACTCCTGgttttagagctagaa
	FwdVar-G7	ggatcctaatacgcactcactataGTGTGCCGTT ACTGCCCTGgttttagagctagaa
	FwdVar-G10	ggatcctaatacgcactcactatagCTTGCCCCA CAGGGCAGTAAgttttagagctagaa
	FwdVar-trG10	ggatcctaatacgcactcactatagTGCCCCACA GGGCAGTAAgttttagagctagaa
	FwdVar-G11	ggatcctaatacgcactcactatagCGTTACTGC CCTGTGGGGCgttttagagctagaa
	FwdVar-trG11	ggatcctaatacgcactcactataGTTACTGCCC TGTGGGGCgttttagagctagaa
	FwdVar-G17	ggatcctaatacgcactcactatagCGTGGATGA AGTTGGTGGTGgttttagagctagaa
	FwdVar-G18	ggatcctaatacgcactcactatagTGAAGTTGG TGGTGAGGCCgttttagagctagaa
	T7RevLong	AAAAAAGCACCGACTCGGTGCCACTTTTTCAAG TTGATAACGGACTAGCCTTATTTTAACTTGCTA TTTCTAGCTCTAAAAC
	T7FwdAmp	GGATCCTAATACGACTCACTATAG

	T7RevAmp	AAAAAAGCACCGACTCGG
Primers and probes for ddPCR analysis	ddPCR forward	CATAAAAGTCAGGGCAGAG
	ddPCR reverse	GTCTCCTTAAACCTGTCTTG
	WT->SCD HDR probe	FAM-CTCCTgtaGAGAAGTCTGC-Quencher
	WT-> SCD unedited probe	HEX-TGACTCCTgagGAGAAGT-Quencher
	SCD->WT HDR probe	FAM-CTCCTgaaGAGAAGTCTGC-Quencher
	SCD->WT unedited probe	HEX-CCTgtgGAGAAGTCTGC-Quencher
ddDNA donor templates for editing WT to SCD (Val7[gta] codon is lowercase)	T1: 195 bp template (Fig. S1)	TACATTTGCTTCTGACACAACCTGTGTTCACTAG CAACCTCAAACAGACACCATGGTGCATCTGACT CCTgtaGAGAAGTCTGCGGTTACTGCCCTGTGG GGCAAAGTGAACGTGGATGAAGTTGGTGGTGAA GCCCTCGGCAGGTTGGTATCAAGGTTACAAGAC AGGTTTAAGGAGACCAATAGAAACTGGGCA
	T2 (T88-107): 195 bp template (Fig. S1)	TACATTTGCTTCTGACACAACCTGTGTTCACTAG CAACCTCAAACAGACACCATGGTGCACCTGACT CCTgtaGAGAAGTCTGCGGTTA CTGCCCTGTGGGGCAAGGTGAACGTGGATGAAG TTGGTGGTGAGGCCCTGGGCAGGTTGGTATCAA GGTTACAAGACAGGTTTAAGGAGACCAATAGAA ACTGGGCA
	T2p: 195 bp template (Fig. S1) “*^*” indicates	T^A^C^ATTTGCTTCTGACACAACCTGTGTTCAC TAGCAACCTCAAACAGACACCATGGTGCACCTG ACTCCTgtaGAGAAGTCTGCGGTTACTGCCCTG TGGGGCAAGGTGAACGTGGATGAAGTTGGTGGT

	phosphorothioate protection	GAGGCCCTGGGCAGGTTGGTATCAAGGTTACAA GACAGGTTTAAGGAGACCAATAGAAACTGG^G^ C^A
	T84-67	TTTGCTTCTGACACAACCTGTGTTCACTAGCAAC CTCAAACAGACACCATGGTGCACCTGACTCCTg taGAGAAGTCTGCGGTTACTGCCCTGTGGGGCA AGGTGAACGTGGATGAAGTTGGTGGTGAGGCC TGGGCAGGTTGGTATCAAG
	T84-37	TTTGCTTCTGACACAACCTGTGTTCACTAGCAAC CTCAAACAGACACCATGGTGCACCTGACTCCTg taGAGAAGTCTGCGGTTACTGCCCTGTGGGGCA AGGTGAACGTGGATGAAGTTGG
	T84-27	TTTGCTTCTGACACAACCTGTGTTCACTAGCAAC CTCAAACAGACACCATGGTGCACCTGACTCCTg taGAGAAGTCTGCGGTTACTGCCCTGTGGGGCA AGGTGAACGTGG
	T84-17	TTTGCTTCTGACACAACCTGTGTTCACTAGCAAC CTCAAACAGACACCATGGTGCACCTGACTCCTg taGAGAAGTCTGCGGTTACTGCCCTGTGGGGCA AG
	T111-57	TCAGGGCAGAGCCATCTATTGCTTACATTTGCT TCTGACACAACCTGTGTTCACTAGCAACCTCAAA CAGACACCATGGTGCACCTGACTCCTgtaGAGA AGTCTGCGGTTACTGCCCTGTGGGGCAAGGTGA ACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCA GGT
	T111-47	TCAGGGCAGAGCCATCTATTGCTTACATTTGCT TCTGACACAACCTGTGTTCACTAGCAACCTCAAA CAGACACCATGGTGCACCTGACTCCTgtaGAGA AGTCTGCGGTTACTGCCCTGTGGGGCAAGGTGA ACGTGGATGAAGTTGGTGGTGAGGCC
	T111-37	TCAGGGCAGAGCCATCTATTGCTTACATTTGCT

		TCTGACACAACCTGTGTTCACTAGCAACCTCAAA CAGACACCATGGTGCACCTGACTCCTgtaGAGA AGTCTGCGGTTACTGCCCTGTGGGGCAAGGTGA ACGTGGATGAAGTTGG
	T111-27	TCAGGGCAGAGCCATCTATTGCTTACATTTGCT TCTGACACAACCTGTGTTCACTAGCAACCTCAAA CAGACACCATGGTGCACCTGACTCCTgtaGAGA AGTCTGCGGTTACTGCCCTGTGGGGCAAGGTGA ACGTGG
	T111-17	TCAGGGCAGAGCCATCTATTGCTTACATTTGCT TCTGACACAACCTGTGTTCACTAGCAACCTCAAA CAGACACCATGGTGCACCTGACTCCTgtaGAGA AGTCTGCGGTTACTGCCCTGTGGGGCAAG
ssDNA donor templates for editing WT to SCD (Glu7[gaa] codon is lowercase)	T88-107S (similar to T2, edits WT-to- SCD)	TACATTTGCTTCTGACACAACCTGTGTTCACTAG CAACCTCAAACAGACACCATGGTGCACCTGACT CCTgaaGAGAAGTCTGCGGTTACTGCCCTGTGG GGCAAGGTGAACGTGGATGAAGTTGGTGGTGAG GCCCTGGGCAGGTTGGTATCAAGGTTACAAGAC AGGTTTAAGGAGACCAATAGAAACTGGGCA
	T111-57S	TCAGGGCAGAGCCATCTATTGCTTACATTTGCT TCTGACACAACCTGTGTTCACTAGCAACCTCAAA CAGACACCATGGTGCACCTGACTCCTgaaGAGA AGTCTGCGGTTACTGCCCTGTGGGGCAAGGTGA ACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCA GGT
	T111-37S	TCAGGGCAGAGCCATCTATTGCTTACATTTGCT TCTGACACAACCTGTGTTCACTAGCAACCTCAAA CAGACACCATGGTGCACCTGACTCCTgaaGAGA AGTCTGCGGTTACTGCCCTGTGGGGCAAGGTGA ACGTGGATGAAGTTGG
	T111-27S	TCAGGGCAGAGCCATCTATTGCTTACATTTGCT TCTGACACAACCTGTGTTCACTAGCAACCTCAAA

		CAGACACCATGGTGCACCTGACTCCTgaaGAGA AGTCTGCGGTTACTGCCCTGTGGGGCAAGGTGA ACGTGG
--	--	--

Table S2. Indel mutations at cancer-associated genes in K562 cells and HSPCs edited with the trG10 RNP, as compared to unedited cells. Cancer-associated genes were captured from genomic DNA using the Illumina TruSight Cancer panel and mutations were detected using the open source Mutect2 algorithm (see methods).

I. K562 Cells

Chr	Location	Gene	Type	Reference	Mutation	% Reads ¹ (edited)	% Reads ¹ (unedited)
22	29130813	CHEK2 (intron)	Deletion	GAA	G	9.56	5.96
7	148543731	EZH2 (intron)	Insertion	C	CT	0.67	0.02
7	148543732	EZH2 (intron)	Insertion	A	AACAATGAACAAT TTCTCCTTCTCTC CTTCATTTTT	0.67	0.21
10	43601785	RET (intron)	Insertion	A	AGGTGTCCCGGG GAGCAGCGTGCTT GTGTACCGCTCAC CAGTCCCCTGATG CAGGTACCACG	0.74	0.18
10	72357821	PRF1	Insertion	C	CTGTCTCTTATACA CATCTCCGAG	0.55	0.11
10	104353831	SUFU (Intron)	Insertion	T	TGAGTGAGGAAAA CCCACTCAAGACAT ACTTGCAGGTGTG GATCGATCTC	0.76	0.33
17	25541278	Intergenic	Insertion	A	AGTGT	2.2	2.32
17	29557005	NF1 (Intron)	Insertion	C	CATTCAG	1.22	0

II. HSPCs

No indel mutations detected.

¹Percent of reads is percent of unique sequencing reads, based on read/insert length.

Table S3. Estimates of cellular editing in mice engrafted with edited HSPCs.

The estimates were derived from measurement of allelic editing in bone marrow by NGS, 16 weeks after injection. Estimates were calculated using the Hardy-Weinberg equation: $(p + q + r)^2 = 1$ (46).

Mouse	Week 16 bone marrow (NGS)			Week 16 cellular prevalence estimate		
	%HDR (p)	% indel (q)	% unedited (r)	%HDR+ (p ₋)	% indel+ (q ₋)	% unedited (r ₋)
1	0.91	57.20	41.88	1.81	81.68	66.23
2	6.20	51.07	42.73	12.02	76.06	67.20
3	1.60	45.04	53.36	3.17	69.80	78.24
4	1.46	43.05	55.49	2.90	67.57	80.19
5	1.40	42.08	56.52	2.78	66.45	81.09
6	1.68	43.91	54.41	3.33	68.54	79.22
7	2.88	41.31	55.81	5.68	65.55	80.47
Mean	2.30	46.24	51.46	4.53	70.81	76.09
SD	1.82	5.81	6.34	3.51	5.90	6.48