Title

Highly Efficient and Marker-free Genome Editing of Human Pluripotent Stem Cells by CRISPR-Cas9 RNP and AAV6 Donor-Mediated Homologous Recombination

Journal

Cell Stem Cell, 24(5)

ISSN

1934-5909

Authors

Martin, Renata M
Ikeda, Kazuya
Cromer, M Kyle
et al.

Publication Date

2019-05-01

DOI

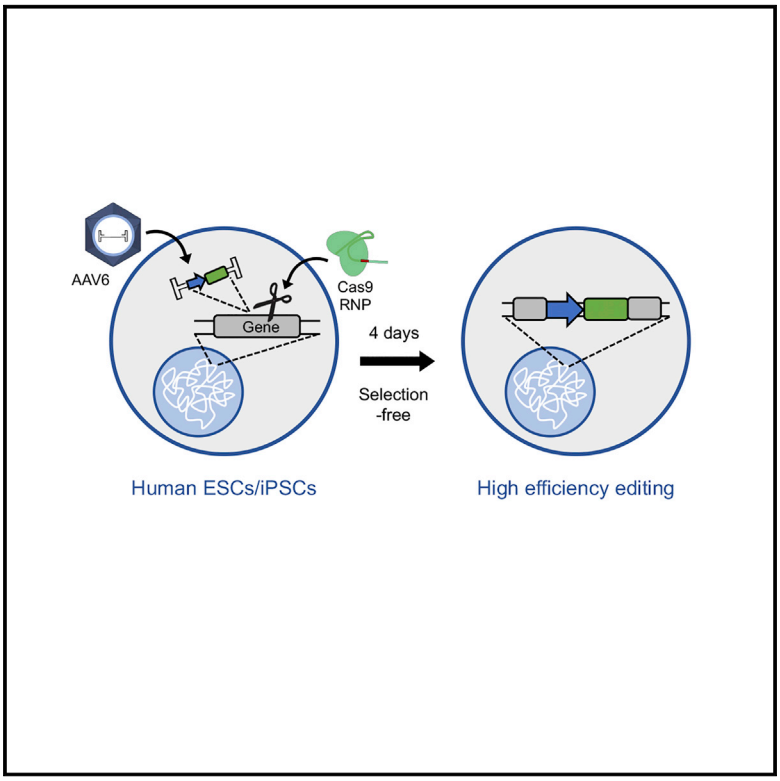
10.1016/j.stem.2019.04.001

Resource

Cell Stem Cell

Highly Efficient and Marker-free Genome Editing of Human Pluripotent Stem Cells by CRISPR-Cas9 RNP and AAV6 Donor-Mediated Homologous Recombination

Graphical Abstract



Authors

Renata M. Martin, Kazuya Ikeda, M. Kyle Cromer, ..., Vittorio Sebastiano, Hiromitsu Nakauchi, Matthew H. Porteus

Correspondence

mporteus@stanford.edu

In Brief

Matthew Porteus and colleagues report a method for high-frequency genome editing in human pluripotent stem cells, enabling introduction of site-specific modifications ranging from single-base-pair changes to 3-kb integrations without the use of fluorescent markers or antibiotic selection.

Highlights

- AAV6 is an effective donor delivery vector for genome editing in hPSCs
- Electroporation of Cas9 RNP prior to AAV6 transduction yields editing up to 90%
- The Cas9 RNP/AAV6 method allows for specific modifications ranging from 1 to >3,000 bp
- This method yields highly edited cells without selection markers or antibiotics

Martin et al., 2019, Cell Stem Cell *24*, 821–828
May 2, 2019
https://doi.org/10.1016/j.stem.2019.04.001

CellPress

Cell Stem Cell
Resource

Highly Efficient and Marker-free Genome Editing of Human Pluripotent Stem Cells by CRISPR-Cas9 RNP and AAV6 Donor-Mediated Homologous Recombination

Renata M. Martin,[1,7] Kazuya Ikeda,[1,7] M. Kyle Cromer,[1,7] Nobuko Uchida,[2] Toshinobu Nishimura,[3] Rosa Romano,[1] Andrew J. Tong,[1] Viktor T. Lemgart,[1] Joab Camarena,[1] Mara Pavel-Dinu,[1] Camille Sindhu,[1] Volker Wiebking,[1] Sriram Vaidyanathan,[1] Daniel P. Dever,[1] Rasmus O. Bak,[1] Anders Laustsen,[4,5] Benjamin J. Lesch,[1] Martin R. Jakobsen,[4,5] Vittorio Sebastiano,[6] Hiromitsu Nakauchi,[3] and Matthew H. Porteus[1,8,*]
[1]Department of Pediatrics, Stanford University, Stanford, CA 94305, USA
[2]ReGen Med Division, BOCO Silicon Valley, Palo Alto, CA 94303, USA
[3]Department of Genetics, Stanford University, Stanford, CA 94305, USA
[4]Department of Biomedicine, Aarhus University, Wilhelm Meyers Alle 4, 8000 Aarhus C, Denmark
[5]Aarhus Research Centre of Innate Immunology, Aarhus University, Wilhelm Meyers Alle 4, 8000 Aarhus C, Denmark
[6]Department of Obstetrics & Gynecology, Stanford University, Stanford, CA 94305, USA
[7]These authors contributed equally
[8]Lead Contact
*Correspondence: mporteus@stanford.edu
https://doi.org/10.1016/j.stem.2019.04.001

SUMMARY

Genome editing of human pluripotent stem cells (hPSCs) provides powerful opportunities for *in vitro* disease modeling, drug discovery, and personalized stem cell-based therapeutics. Currently, only small edits can be engineered with high frequency, while larger modifications suffer from low efficiency and a resultant need for selection markers. Here, we describe marker-free genome editing in hPSCs using Cas9 ribonucleoproteins (RNPs) in combination with AAV6-mediated DNA repair template delivery. We report highly efficient and bi-allelic integration frequencies across multiple loci and hPSC lines, achieving mono-allelic editing frequencies of up to 94% at the *HBB* locus. Using this method, we show robust bi-allelic correction of homozygous sickle cell mutations in a patient-derived induced PSC (iPSC) line. Thus, this strategy shows significant utility for generating hPSCs with large gene integrations and/or single-nucleotide changes at high frequency and without the need for introducing selection genes, enhancing the applicability of hPSC editing for research and translational uses.

INTRODUCTION

With the maturation of zinc-finger nuclease (ZFN) and transcription activator-like effector nuclease (TALEN) technology, as well as the advent of CRISPR, we now have an unprecedented ability to create DNA double-strand breaks (DSBs) at virtually any location in the genome. Following initiation of a site-specific DSB, we can then provide a DNA donor template, which the cell can use via endogenous damage repair pathways to introduce modifications ranging from single-base-pair substitutions to large insertions through homology-directed repair (HDR). Although the application of this genome editing technology in human pluripotent stem cells (hPSCs) holds great promise for *in vitro* disease modeling as well as personalized cell-based therapies, current genome editing protocols for hPSCs remain limited due to low editing frequencies. This makes it difficult to identify mono-allelic edited clones for expansion, much less those that have undergone bi-allelic editing events.

To overcome these limitations, previous studies have used fluorescent-tagged donors for the identification and selection of edited cells (Byrne and Church, 2015; Arias-Fuenzalida et al., 2017). Although this method was able to successfully monitor recombination frequencies, it not only introduced a selection marker in the genome that could possibly interfere with regional transcriptional regulation but also would be incompatible with clinical translation. To avoid the possible influence of a selection marker, piggyBac or Cre/*loxP* systems have been used to excise this marker. However, the Cre/*loxP* leaves a large *loxP* sequence behind after editing, while the piggyBac system requires a nearby TTAA site to avoid introduction of additional sequences (Li et al., 2013). While prior methods such as single-stranded oligodeoxynucleotides (ssODNs) for donor template delivery and iCRISPR have been able to achieve high-frequency editing without use of selection markers, these technologies only apply to small base pair edits and are severely limited when it comes to the introduction of large insertions (Yang et al., 2013; Wang et al., 2017). These methods are useful in modeling diseases caused by single or several base pair mutations, but inapplicable to research applications that require targeted integration of large gene cassettes. For example, diseases such as Huntington's and myotonic dystrophy are caused by genetic changes spanning a large region of the genome; thus, there is an unmet need for an efficient and safe method for introducing changes larger than a few nucleotides. Furthermore, the introduction of large gene fragments (>1 kb) is also required for basic science applications, such as the integration of a reporter gene

as a marker for differentiation, which is important for biological understanding and clinical application of hPSCs.

Recently, a technology called targeted integration with linearized double-strand DNA (TILD)-CRISPR was described to introduce large insertions (Yao et al., 2018). Microinjection of genome editing components is one strategy capable of delivering large amounts of donor into the nucleus, a key factor in generating high frequencies of homologous-recombination-mediated genome editing (Hendel et al., 2014). Since microinjection is impractical for editing large populations of hPSCs, an alternative method was required to achieve high donor concentrations in populations of cells in the nucleus. We previously demonstrated the donor could be effectively delivered to primary human $CD34^+$ hematopoietic stem and progenitor cells (HSPCs) using electroporation-aided transduction of AAV6 (Charlesworth et al., 2018), resulting in high frequencies of HDR in HSPCs and primary human T cells (Dever et al., 2016; Bak and Porteus, 2017; Eyquem et al., 2017; Bak et al., 2018). In these experiments, we used electroporation to deliver Cas9 protein complexed to a chemically modified guide RNA (a ribonucleoprotein [RNP] complex), which has been shown to be well tolerated by the target cell (Hendel et al., 2015; Cromer et al., 2018). We hypothesized that a similar strategy might also effectively generate high frequencies of HDR in hPSCs.

In this study, we demonstrate that Cas9 protein complexed with chemically modified single guide RNA (sgRNA) followed by transduction of recombinant AAV6 (rAAV6) vector for donor delivery can be used to introduce large transgenes in hPSCs at high frequencies without marker selection.

RESULTS

Initiation of a DSB and delivery of high concentrations of donor are critical to achieving high frequencies of homologous-recombination-mediated genome editing (Hendel et al., 2014). Therefore, we believe that the low frequencies of HDR previously reported in hPSCs are due to inefficiencies in one or both events. Prior work demonstrated that AAV6 is able to effectively edit HSPCs in combination with electroporation of Cas9/sgRNA RNP (Dever et al., 2016). The biology of hPSCs is different than human HSPCs in terms of their proliferation parameters, gene expression profiles, and epigenetic states; nonetheless, we hypothesized that a similar strategy might also generate high frequencies of HDR in hPSCs. To test this hypothesis, we first measured the ability of a panel of AAV serotypes to transduce hPSCs following electroporation. We found that among all AAV serotypes, AAV6 is most effective on H9 embryonic stem cells (ESCs), transducing 73% of cells (Figure S1A). Therefore, we used AAV6 to deliver the donor cassette as well as electroporation of Cas9 RNP in H9 ESCs at the *HBB* locus to integrate a 2.2-kb cassette (GFP driven by a ubiquitin C [UbC] promoter), yielding ~45% of cells that stably express GFP compared to <1% when AAV6 was used alone (Figures 1A and 1B). When we substituted the delivery of RNP by lipofectamine rather than electroporation, however, the HDR efficiency dropped significantly and we were unable to achieve targeted integration frequencies >3% even at an AAV6 MOI of 100,000 (Figure S1B). As further confirmation, we quantified the ability of electroporation to deliver payloads to hPSCs by electroporating a GFP expression plasmid, yielding >96% GFP+ hPSCs (Figure S1C).

We then explored the timing of AAV6 transduction with respect to electroporation. We found that transducing hPSCs within 5 min following electroporation yielded substantially higher integration frequencies (>80%) compared to transducing AAV6 24 h prior to electroporation (~15%) (Figure S1D). These findings—that electroporation immediately prior to AAV transduction most effectively mediates HDR—are in agreement with those described previously in human HSPCs (Charlesworth et al., 2018).

To determine the time course by which episomal expression of the transgene is lost in hPSCs after transduction with AAV6, we monitored GFP expression from day 2 onward in "RNP+AAV6"-treated populations and "AAV6 only" populations. By day 4, we detected <1% of GFP+ cells in the AAV6-only group, whereas GFP expression remained stable in the RNP+AAV6 group (Figure 1B), indicating that episomal expression from nonintegrated vectors had been lost by this time point and that GFP was driven solely by integrated vectors. We then demonstrated the importance of AAV6 in generating high frequencies of HDR by comparing AAV6 versus plasmid versus TILD as a method of donor delivery. We found that delivery by AAV6 generated significantly higher targeted integration frequencies than the other two methods (AAV6, 53.7%; plasmid, 0.13%; and TILD, 0.03%) (Figure 1C; TILD data not shown). We also compared donor delivery using AAV6 versus ssODN in their ability to modify the *HTT* locus at exon 1 by inserting a short CAG repeat. Using ssODN as the donor, none of the clones picked had the desired allele modification, while in two separate experiments using our AAV6/Cas9 protocol, 45% and 37% of cells had the desired HDR allele modification (data not shown).

We next determined the robustness of the method by determining the targeted integration frequency across induced pluripotent stem cell (iPSC) lines. When targeting the *HBB* locus with a UbC-GFP AAV6 donor vector in five different iPSC lines, we measured an average targeted integration frequency of 51% (Figure 1D). To determine if we could improve these editing frequencies, we explored different parameters within the protocol. Because delivery of Cas9 and subsequent formation of insertion-deletions (INDELs) is often a rate-limiting step to achieving high-frequency editing (Hendel et al., 2014, 2015), we hypothesized that increasing the amount of Cas9 delivered by electroporation could improve editing rates. Across a wide range of concentrations of Cas9 RNP complexed with sgRNA (0–150μg/mL), we found that increased amounts of Cas9 increased editing frequencies at both the *HBB* and *MYD88* loci in a dose-dependent manner (Figure 1E). We also tested different electroporation programs and found that CA137 protocol of the Lonza 4-D system more than doubled the editing frequency at the *HBB* locus compared to the previously used CB150 (84% versus 41%, respectively; Figure 1F).

Combining the fully optimized conditions (150 μg/mL Cas9 with CA137 electroporation program), we were able to achieve 91% targeted integration at *HBB* and 59% at *MYD88* in ESCs when analyzed 4 days post-editing (Figures 1G and 2A). To analyze whether cells possessed mono- or bi-allelic integrations, we performed single-cell cloning of populations targeted at either *MYD88* and *HBB*—a process that took 2–3 weeks

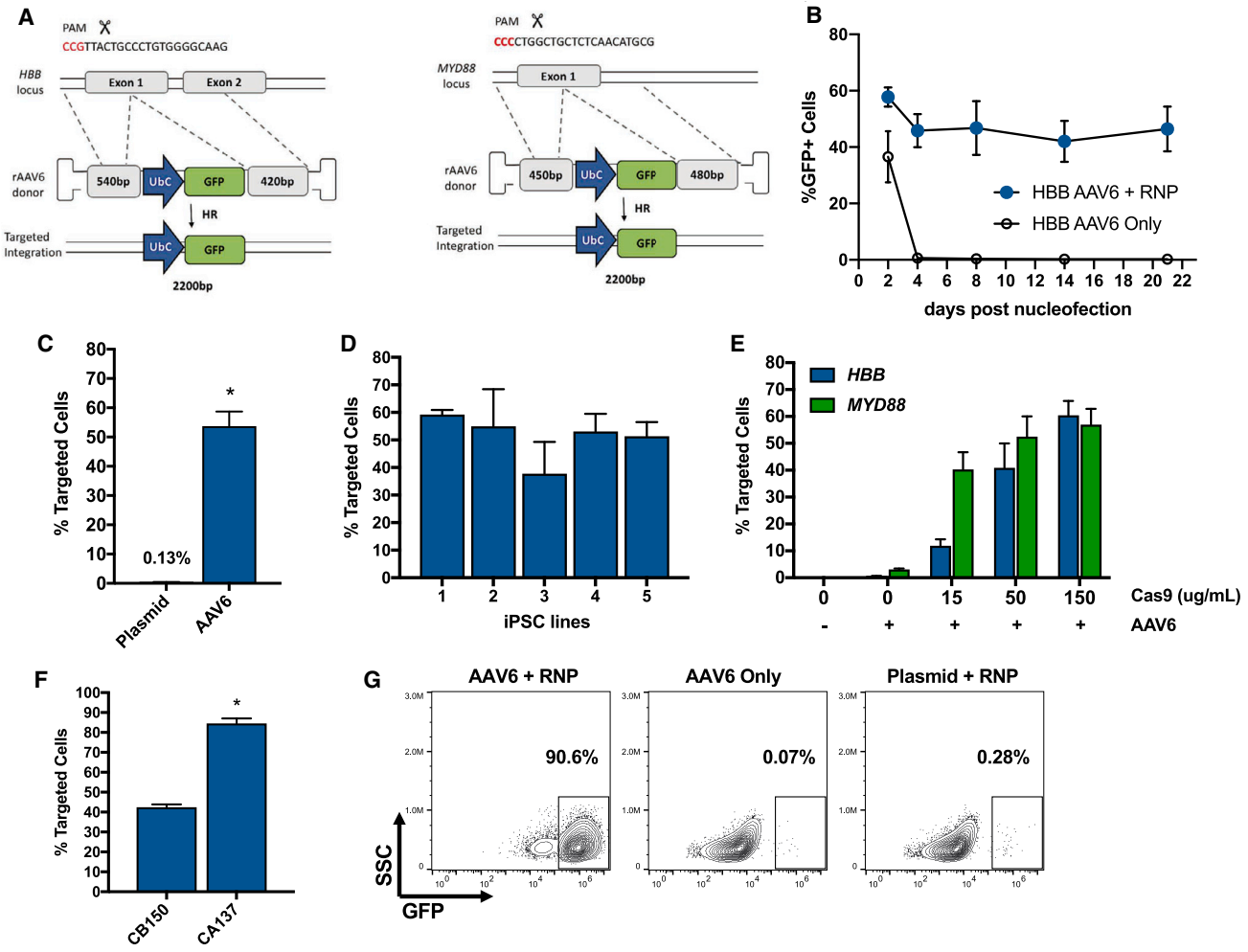


Figure 1. Optimization of the AAV6/Cas9 System in hPSCs

(A) Schematic of genome editing at *HBB* (left) and *MYD88* (right) loci using CRISPR/Cas9 RNP and AAV6.

(B) Time course of targeting frequencies of ESCs at the *HBB* locus using AAV6 and Cas9 RNP (n = 5) as determined by flow cytometry. Data are shown as mean ± SEM.

(C) Targeting frequencies of ESCs at *HBB* (using CB-150 electroporation protocol) comparing AAV6 (n = 5) to plasmid-based donor delivery (n = 3). The percentage of GFP+ cells was quantified 8 days post-electroporation by flow cytometry. $p < 0.0005$ based on unpaired t test. Data are shown as mean ± SEM.

(D) Targeting frequencies at the *HBB* locus in five different iPSC lines using electroporation protocol CA-137. The percentage of GFP+ cells was quantified 8 days post-electroporation by flow cytometry (n = 2). ANOVA detected no significant differences among cell lines. Data are shown as mean ± SEM.

(E) Targeting frequencies at the *HBB* and *MYD88* loci across a range of Cas9/sgRNA concentrations at a molar ratio of 1/3, respectively (n = 2 for *MYD88*; n = 3 for *HBB* samples). The percentage of GFP+ cells was quantified 8 days post-electroporation by flow cytometry. Data are shown as mean ± SEM.

(F) Targeting frequencies at the *HBB* locus in ESCs when using CB-150 and CA-137 electroporation protocols to deliver Cas9 RNP followed by AAV6 transduction. The percentage of GFP+ cells was quantified 8 days post-electroporation by flow cytometry (n = 4). If not stated otherwise, 150 μg/mL Cas9 and the CB-150 electroporation program were used. $p < 0.00001$ based on unpaired t test. Data are shown as mean ± SEM.

(G) A representative flow cytometry plot from a single replicate after optimizations comparing plasmid versus AAV6+RNP. The percentage of GFP+ cells was quantified 8 days post-electroporation.

(Figure 2B). PCR analysis using primers annealing outside each homology arm allowed us to determine that 17 out of 18 clones (94%) for *HBB* and 13 out of 21 clones (62%) for *MYD88* possessed at least one integration (Figure 2C). Editing occurred with such frequency that we obtained a high percentage of bi-allelic integration events for *HBB* (74% of clones screened) (Figures 2C and 2D). For *MYD88*, we obtained 52% mono-allelic clones and 10% bi-allelic clones (Figures 2D and 2D).

We then tested the AAV6/Cas9 system across five additional loci and found the system to be highly robust, with >40% of cells positive for the respective fluorescent markers across all loci targeted as assessed by flow cytometry (Figures S2A–S2C). This was further confirmed by droplet digital PCR (ddPCR), which measured targeted integration frequencies at >20% of alleles across all loci tested (Figure S2D), which corresponds to the ~40% of cells that were shown to have at least one editing event via flow cytometry.

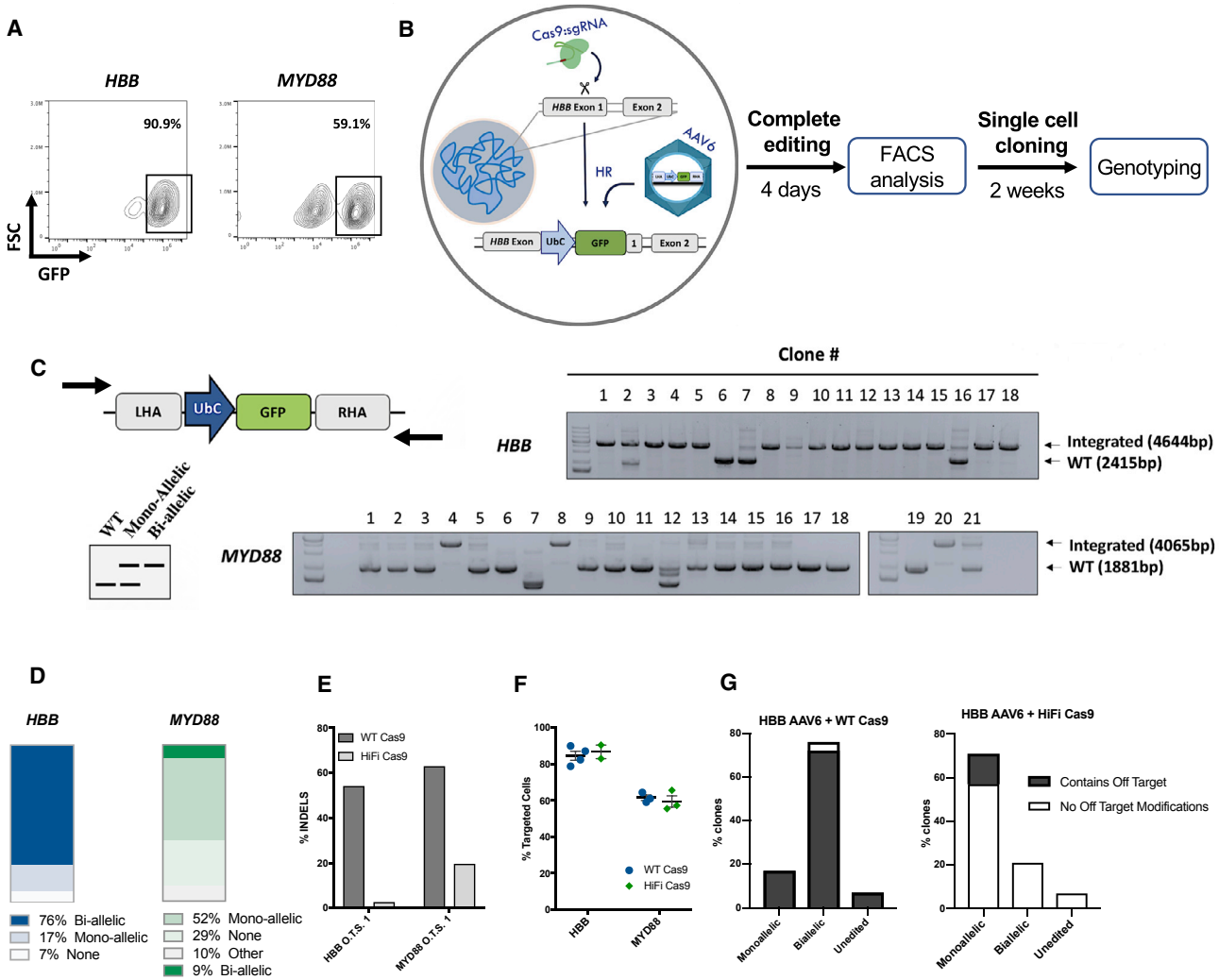


Figure 2. Analysis of AAV6-Edited Clones at *HBB* and *MYD88* Loci

(A) Targeted integration at the *HBB* and *MYD88* loci using optimized conditions (150 μg/mL Cas9 and the CA-137 electroporation program). The percentage of GFP+ cells was quantified 8 days post-electroporation by flow cytometry.

(B) Schematic depicting the genome editing process using Cas9 RNP and AAV6-mediated donor delivery followed by flow cytometry analysis 4 days post-targeting, allowing genotyping of clones within 2 weeks.

(C) Schematic depicting primer locations for clonal genotyping of the *HBB* and *MYD88* loci (top left), and a schematic representation of an agarose gel with all three possible genotypes (bottom left). Agarose gel images show genotypes of analyzed clones targeted at the *HBB* and *MYD88* loci (top and bottom right, respectively).

(D) Distribution of bi-allelic, mono-allelic, and WT clones from *HBB*- and *MYD88*-targeted populations.

(E) Off-target analysis for *MYD88*- and *HBB*-edited samples with WT and HiFi Cas9. Off-target rates were measured by the percentage of INDELS using TIDE analysis. No activity was observed at the second predicted off-target site for *HBB* or *MYD88*.

(F) Targeting frequencies at *HBB* and *MYD88* with WT and HiFi Cas9 as measured by the percentage of GFP+ cells by flow cytometry. Bars indicate mean ± SEM.

(G) On- and off-target analysis of cells edited at *HBB* using WT Cas9 (18 clones) and HiFi Cas9 (14 clones).

A major advantage of our combined AAV6/Cas9 approach is that it facilitates integration of a large cassette into a specific site in the genome. To confirm the specificity of integration only at the target locus, we performed ddPCR on three edited clones with bi-allelic integrations in *HBB* to analyze copy number. We found that all edited clones had exactly four copies of the human UbC promoter (two from the bi-allelic editing and two endogenous copies; Figure S3B). To determine if the integrations were seamless (without scars or residual AAV inverted terminal repeats (ITRs) at the edge of the homology arm junctions) at the desired locus, we sequenced the junctions and found that they were indeed precise.

To further confirm that our editing strategy does not induce widespread genomic changes, we performed karyotype analysis on a *MYD88* and *HBB* bi-allelic edited clone and observed no gross chromosomal abnormalities (Figure S3D). In further confirmation of these results, we performed SNP array analysis capable of detecting genomic insertions or deletions and found no changes between our unedited population and *MYD88* and *HBB* bi-allelic edited clones.

To further confirm that our editing strategy does not induce aberrant genomic changes or enrich for specific mutations, we performed Sanger sequencing of the *TP53* gene for *HBB* and *MYD88* clones and found a synonymous mutation located at position Chr17:7674237. However, the unedited cell line also contained the identical mutation, indicating that this had occurred prior to editing. We then checked our edited and unedited SCD iPSC lines and found no *TP53* mutations, indicating that mutations in the *TP53* gene are not required in order to edit iPSCs at high frequencies. These findings are contrary to recent studies that suggest CRISPR-mediated editing itself could induce and enrich for *TP53* mutations in a population of edited cells or that mutations in *TP53* are necessary to achieve high frequencies of targeted integration in hPSCs when using the AAV6/Cas9 system described here (Haapaniemi et al., 2018; Ihry et al., 2018).

Furthermore, in order for our strategy to be effective, we tested whether the editing process reduced pluripotency of our stem cell population. To do so, we compared two edited clones each for *HBB* and *MYD88* using the hPSC Scorecard panel to assess expression levels of markers for pluripotency as well as the three germ layers (Figure S3C; Fergus et al., 2015). We also performed antibody staining for OCT4 and NANOG on edited clones for *HBB* to confirm expression of canonical pluripotency markers (Figure S3E).

We performed off-target analysis at the two most probable sites for both *HBB* and *MYD88* sgRNAs identified by the COSMID target prediction tool (Cradick et al., 2013, 2014; Hendel et al., 2015; Dever et al., 2016; DeWitt et al., 2016; Table S1). This revealed off-target activity for both sgRNAs with 54% INDELs at off-target site 1 for *HBB* and 63% INDELs at off-target site 1 for *MYD88* (Figure 2E). Importantly, these off-target INDELs occurred at intergenic sites of no known biologic significance and thus are unlikely to have important functional consequences. Nonetheless, we found that use of a high-fidelity (HiFi) variant of Cas9 was able to reduce the occurrence of off-target INDELs (2.6% and 19.7% for *HBB* and *MYD88*, respectively) without sacrificing on-target editing frequencies (Figure 2F; Vakulskas et al., 2018). We then confirmed these results through clonal analysis for cells edited with *HBB* AAV6 and wild type (WT) or HiFi Cas9 (Figure 2G). In the HiFi Cas9-edited group, none of the bi-allelic clones contained any off-target modifications or AAV6 ITR sequences outside the homology arm.

We next tested our system in a disease-relevant iPSC line homozygous for the Glu6Val sickle cell disease (SCD)-causing mutation (Sebastiano et al., 2011). Toward this end, we used a previously described AAV6 donor vector that corrects the disease-causing SNP (T→A) and introduces five silent variants to prevent Cas9 from re-cutting and disrupting the gene (Figure 3A; Dever et al., 2016). Combining this SCD-correction donor and Cas9 RNP, we targeted the SCD iPSC line using the optimized protocol, yielding 63% allele correction by ddPCR (Figure 3B). As before, we confirmed these results using single-cell cloning followed by sequencing, which confirmed a high correction frequency—63% of clones possessed bi-allelic correction, and 17% had mono-allelic correction accompanied by an INDEL on the other allele (Figure 3C). In addition, we observed a small percentage of clones (4%) with very short-track homologous recombination events in which only the first three synonymous changes proximal to the break were introduced into the genome but the more distal changes in the donor were not. We have not observed these events in human somatic cells, and this finding perhaps highlights a subtle difference in how human pluripotent cells repair breaks by homologous recombination. The off-target analysis identified a corrected clone that did not contain INDELs at the two off-target sites for the *HBB* sgRNA (Figure S3A), thus demonstrating that using the AAV6/Cas9 system makes it feasible to identify clones with on-target HDR edits without off-target INDELs.

In order to demonstrate that the editing of the SCD-causing mutation in the patient-derived iPSC line resulted in functional expression of the corrected allele, we used a previously reported qPCR assay that is able to distinguish between the sickle hemoglobin allele (HbS) and the WT adult hemoglobin allele (HbA) at the *HBB* locus (Dever et al., 2016). To do so, we obtained a bi-allelically edited clone and confirmed the edited sequence. Because the *HBB* locus is only activated when cells are differentiated into red blood cells (RBCs), following targeting, we differentiated this iPSC clone into erythrocytes using an established protocol (Fujita et al., 2016). We then used fluorescence-activated cell sorting (FACS) to sort out erythrocytes from unedited and edited populations and isolated mRNA upon which we performed our qPCR assay. As expected, we found our unedited SCD cells to have high levels of HbS expression and virtually no WT HbA (Figure 3D). In contrast, we found that the correction of the SCD-causing mutation virtually eliminates HbS expression and leads to high levels of HbA.

DISCUSSION

In summary, our data establish that the Cas9 RNP and AAV6 system is able to mediate targeted integration of large donor cassettes in ESCs and iPSCs across many loci as well as precisely correct a disease-causing mutation in a patient-derived iPSC line, all at high frequencies. The correction of a disease-causing mutation provides a proof of concept for therapeutic editing strategies and generating isogenic lines with SNPs—a key research approach to understanding variants of unknown significance (VUSs) that are now being described through sequencing efforts. Moreover, these high editing frequencies enable population-based studies without the requirement for selection markers or the need to screen large numbers of clones to find one with the desired HDR event. While a previous study combined DSB initiation by zinc-finger nucleases and an AAV1.9 repair donor in human iPSCs, the maximum editing frequency observed using this method was less than 2% for a 35-bp deletion (Asuri et al., 2012). Here, we use Cas9 complexed with a chemically modified guide RNA (making the guide resistant to exonuclease degradation) rather than a zinc-finger nuclease in order to initiate DSBs. We deliver Cas9/sgRNA as RNP, which allows for immediate cutting to occur once electroporated into the cell (Liang et al., 2015) and decreases the cellular response to the nuclease (Cromer et al., 2018). Immediate cutting is critical, since our data indicate that the donor template is quickly diluted out in hPSCs because of their high proliferation rate. Furthermore, we found that the timing of AAV6 donor delivery is crucial to achieving high editing frequencies and that transduction of AAV6 immediately following electroporation of Cas9/sgRNA ensures the donor template is present at high concentrations at the

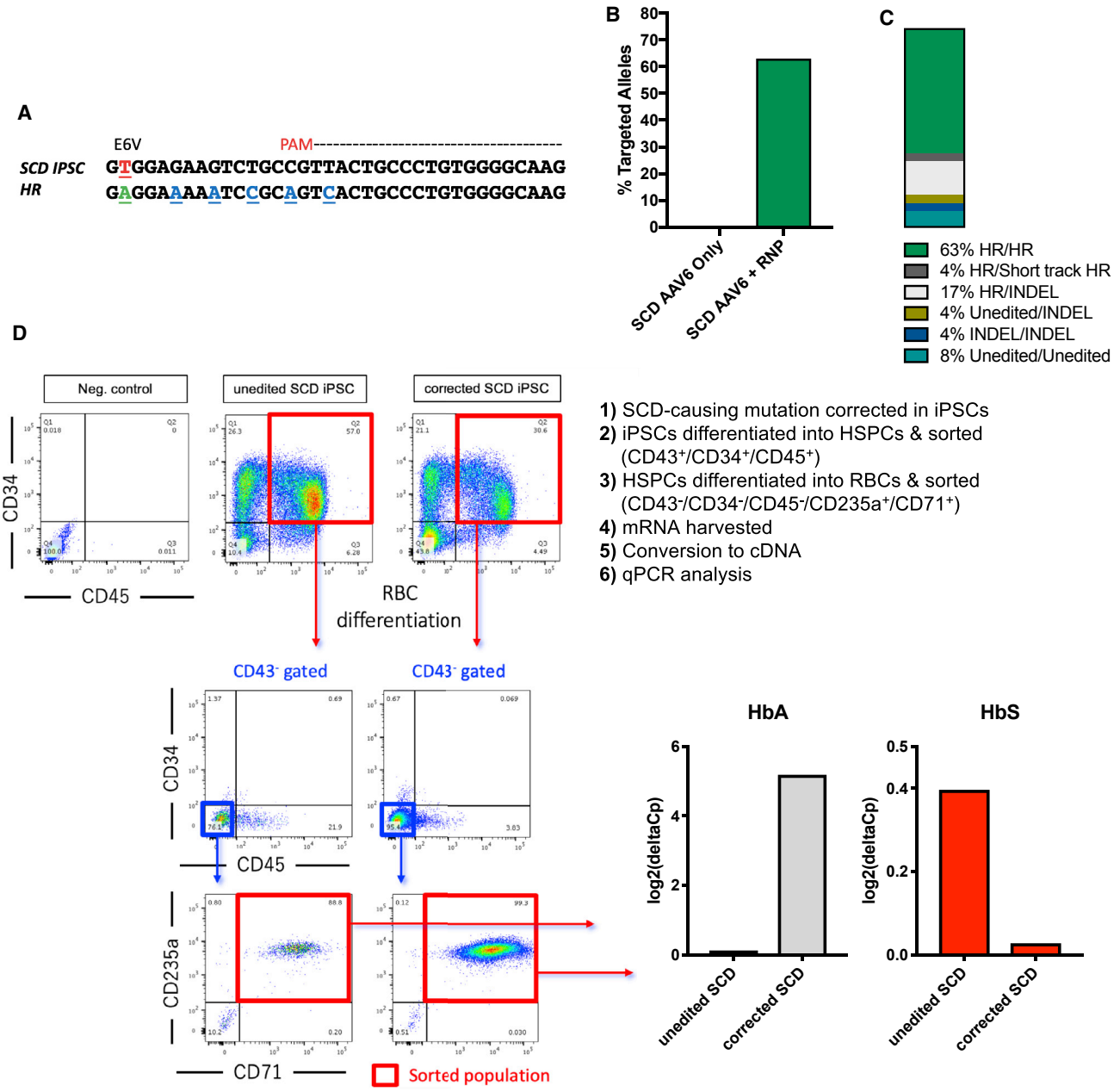


Figure 3. Sickle Cell iPSC Correction with AAV6 Donor

(A) Sequence of the SCD-causing SNP (E6V) with *HBB* sgRNA target site depicted with dashes and protospacer adjacent motif (PAM) in red. Below is the sequence after corrective HR, which introduces silent mutations in the sgRNA target site to prevent Cas9 re-cutting.

(B) Targeting frequencies of SCD-iPSCs targeted with corrective donor as determined by ddPCR.

(C) Distribution of bi-allelic, mono-allelic, and WT edits in SCD-iPSCs.

(D) Schematic depicting the workflow used to edit SCD-derived iPSC, isolate HSPCs, differentiate into RBCs, isolate RBCs, and analyze mRNA levels of HbA and HbS alleles in edited and unedited RBC populations by qPCR.

time of cutting. Therefore, achieving maximum editing efficiency in hPSCs requires optimization of both donor vector delivery as well as nuclease-mediated DSB initiation.

In addition to previous attempts to combine AAV delivery with nuclease-mediated DSB formation mentioned above, prior studies have attempted to improve low homologous recombination (HR)-based genome editing frequencies in hPSCs using small-molecule HR enhancers, modulation of non-homologous end joining (NHEJ) and HR balance by gene transduction, and cold shock (Yu et al., 2015; Paulsen et al., 2017; Guo et al., 2018). However, all of these methods used ssODNs to deliver the donor template. While combining Cas9/sgRNA-mediated DSB formation with ssODN-mediated donor delivery is able to achieve high editing frequencies, the effectiveness of this

approach is inversely correlated to the size of the edit that is being introduced (Yang et al., 2013). Furthermore, because cDNA knockin is a common treatment strategy for correction of diseases with loss-of-function mutations scattered throughout a particular gene (Hubbard et al., 2016; Schiroli et al., 2017; Pavel-Dinu et al., 2019), our ability to introduce insertions of over 2 kb into patient-derived hPSCs has great clinical applicability. Our approach could also be applied to generate genetically engineered human pluripotent cells for a variety of research purposes, such as creating stable transgenic or reporter lines. In conclusion, we present a selection-free, one-step editing protocol that addresses limitations of insert size and editing efficiency that have hindered the development of pluripotent stem cell technologies as disease models and gene-correction therapies.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human ESC and iPSCs
- METHOD DETAILS
  - AAV6/Cas9 Genome Editing
  - AAV6 Cloning and Production
  - Flow cytometry
  - Genotyping and sequence analysis
  - Sickle cell clones edited with the anti-sickling donor (Figure 3A) were amplified with the following primers to check for integration
  - ddPCR analysis
  - Off-target analysis
  - ddPCR-based copy number analysis
  - SCD allele correction analysis by nested ddPCR
  - Erythrocyte differentiation of hiPSCs
  - Assessment of mRNA levels in erythrocytes
  - Immunofluorescence staining to determine pluripotency
  - Karyotype analysis
  - High-density SNP array analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stem.2019.04.001.

ACKNOWLEDGMENTS

R.M.M. was funded through an NRSA (NIH National Research Services Award) institutional predoctoral training (T32) grant. M.K.C. was funded through an institutional NIH postdoctoral T32 hematology training grant. R.O.B. was supported through an individual postdoctoral grant (DFF–1333-00106B) and a Sapere Aude, Research Talent grant (DFF–1331-00735B), both from the Danish Council for Independent Research, Medical Sciences. M.H.P. thanks the Laurie Kraus Lacob Translational Scholar Fund, the Amon Carter Foundation, and the Tad Taube Program in Neurodegenerative Research for their support of this work. M.H.P. is a Chan-Zuckerberg Investigator through the Chan-Zuckerberg Biohub. D.P.D. thanks the Children's Health Research Institute at Stanford for their support. V.W. is supported by a research fellowship from Deutsche Forschungsgemeinschaft (DFG).

AUTHOR CONTRIBUTIONS

R.M.M., K.I., M.K.C., and M.H.P. conceived and designed the experiments and wrote the manuscript. R.M.M. and K.I. performed the experiments. M.K.C., N.U., T.N., R.R., A.J.T., V.T.L., J.C., M.P.-D., C.S., V.W., S.V., D.P.D., R.O.B., A.L., B.J.L., M.R.J., V.S., and H.N. all participated in the design and conception of experiments and provided editorial feedback on the manuscript.

DECLARATION OF INTERESTS

M.H.P. has equity and serves on the scientific advisory board of CRISPR Therapeutics. Nobuko Uchida is a current employee of ReGen Med Division, BOCO Silicon Valley, and Kazuya Ikeda is a current employee of Daiichi-Sankyo Co., Ltd. However, none of these companies had input into the design, execution, interpretation, or publication of the work in this manuscript.

Received: June 3, 2018
Revised: October 8, 2018
Accepted: March 29, 2019
Published: May 2, 2019

REFERENCES

Ando, M., Nishimura, T., Yamazaki, S., Yamaguchi, T., Kawana-Tachikawa, A., Hayama, T., Nakauchi, Y., Ando, J., Ota, Y., Takahashi, S., et al. (2015). A safeguard system for induced pluripotent stem cell-derived rejuvenated T cell therapy. Stem Cell Reports *5*, 597–608.

Arias-Fuenzalida, J., Jarazo, J., Qing, X., Walter, J., Gomez-Giro, G., Nickels, S.L., Zaehres, H., Schöler, H.R., and Schwamborn, J.C. (2017). FACS-assisted CRISPR-Cas9 genome editing facilitates Parkinson's disease modeling. Stem Cell Reports *9*, 1423–1431.

Asuri, P., Bartel, M.A., Vazin, T., Jang, J.H., Wong, T.B., and Schaffer, D.V. (2012). Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. Mol. Ther. *20*, 329–338.

Aurnhammer, C., Haase, M., Muether, N., Hausl, M., Rauschhuber, C., Huber, I., Nitschko, H., Busch, U., Sing, A., Ehrhardt, A., and Baiker, A. (2012). Universal real-time PCR for the detection and quantification of adeno-associated virus serotype 2-derived inverted terminal repeat sequences. Hum. Gene Ther. Methods *23*, 18–28.

Bak, R.O., and Porteus, M.H. (2017). CRISPR-mediated integration of large gene cassettes using AAV donor vectors. Cell Rep. *20*, 750–756.

Bak, R.O., Dever, D.P., and Porteus, M.H. (2018). CRISPR/Cas9 genome editing in human hematopoietic stem cells. Nat. Protoc. *13*, 358–376.

Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Research *42*, e168.

Byrne, S.M., and Church, G.M. (2015). CRISPR-mediated gene targeting of human induced pluripotent stem cells. Curr. Protoc. Stem Cell Biol. *35*, 1–22.

Charlesworth, C.T., Camarena, J., Cromer, M.K., Vaidyanathan, S., Bak, R.O., Carte, J.M., Potter, J., Dever, D.P., and Porteus, M.H. (2018). Priming human repopulating hematopoietic stem and progenitor cells for Cas9/sgRNA gene targeting. Mol. Ther. Nucleic Acids *12*, 89–104.

Cradick, T.J., Fine, E.J., Antico, C.J., and Bao, G. (2013). CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. Nucleic Acids Res. *41*, 9584–9592.

Cradick, T.J., Qiu, P., Lee, C.M., Fine, E.J., and Bao, G. (2014). COSMID: A web-based tool for identifying and validating CRISPR/Cas off-target sites. Mol. Ther. Nucleic Acids *3*, e214.

Cromer, M.K., Vaidyanathan, S., Ryan, D.E., Curry, B., Lucas, A.B., Camarena, J., Kaushik, M., Hay, S.R., Martin, R.M., Steinfeld, I., et al. (2018). Global transcriptional response to CRISPR/Cas9-AAV6-based genome editing in $CD34^{+}$ hematopoietic stem and progenitor cells. Mol. Ther. *26*, 2431–2442.

Dever, D.P., Bak, R.O., Reinisch, A., Camarena, J., Washington, G., Nicolas, C.E., Pavel-Dinu, M., Saxena, N., Wilkens, A.B., Mantri, S., et al. (2016). CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. Nature *539*, 384–389.

DeWitt, M.A., Magis, W., Bray, N.L., Wang, T., Berman, J.R., Urbinati, F., Heo, S.J., Mitros, T., Muñoz, D.P., Boffelli, D., et al. (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. Sci. Transl. Med. *8*, 360ra134.

Eyquem, J., Mansilla-Soto, J., Giavridis, T., van der Stegen, S.J., Hamieh, M., Cunanan, K.M., Odak, A., Gönen, M., and Sadelain, M. (2017). Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. Nature *543*, 113–117.

Fergus, J., Quintanilla, R., and Lakshmipathy, U. (2015). Characterizing pluripotent stem cells using the TaqMan® hPSC scorecard™ panel. Methods Mol. Biol. *1307*, 25–37.

Fujita, A., Uchida, N., Haro-Mora, J.J., Winkler, T., and Tisdale, J. (2016). β-Globin-expressing definitive erythroid progenitor cells generated from embryonic and induced pluripotent stem cell-derived sacs. Stem Cells *34*, 1541–1552.

Guo, Q., Mintier, G., Ma-Edmonds, M., Storton, D., Wang, X., Xiao, X., Kienzle, B., Zhao, D., and Feder, J.N. (2018). 'Cold shock' increases the frequency of homology directed repair gene editing in induced pluripotent stem cells. Sci. Rep. *8*, 2080.

Haapaniemi, E., Botla, S., Persson, J., Schmierer, B., and Taipale, J. (2018). CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. Nat. Med. *24*, 927–930.

Hendel, A., Kildebeck, E.J., Fine, E.J., Clark, J., Punjya, N., Sebastiano, V., Bao, G., and Porteus, M.H. (2014). Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing. Cell Rep. *7*, 293–305.

Hendel, A., Bak, R.O., Clark, J.T., Kennedy, A.B., Ryan, D.E., Roy, S., Steinfeld, I., Lunstad, B.D., Kaiser, R.J., Wilkens, A.B., et al. (2015). Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. Nat. Biotechnol. *33*, 985–989.

Hubbard, N., Hagin, D., Sommer, K., Song, Y., Khan, I., Clough, C., Ochs, H.D., Rawlings, D.J., Scharenberg, A.M., and Torgerson, T.R. (2016). Targeted gene editing restores regulated CD40L function in X-linked hyper-IgM syndrome. Blood. *127*, 2513–2522.

Ihry, R.J., Worringer, K.A., Salick, M.R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J., Sondey, M., Ye, C., et al. (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. Nat. Med. *24*, 939–946.

Li, X., Burnight, E.R., Cooney, A.L., Malani, N., Brady, T., Sander, J.D., Staber, J., Wheelan, S.J., Joung, J.K., McCray, P.B., Jr., et al. (2013). piggyBac transposase tools for genome engineering. Proc. Natl. Acad. Sci. USA *110*, E2279–E2287.

Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R., Sridharan, M., Carte, J., Chen, W., Roark, N., Ranganathan, S., et al. (2015). Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J. Biotechnol. *208*, 44–53.

Nishimura, T., Kaneko, S., Kawana-Tachikawa, A., Tajima, Y., Goto, H., Zhu, D., Nakayama-Hosoya, K., Iriguchi, S., Uemura, Y., Shimizu, T., et al. (2013). Generation of rejuvenated antigen-specific T cells by reprogramming to pluripotency and redifferentiation. Cell Stem Cell *12*, 114–126.

Ochi, K., Takayama, N., Hirose, S., Nakahata, T., Nakauchi, H., and Eto, K. (2014). Multicolor Staining of Globin Subtypes Reveals Impaired Globin Switching During Erythropoiesis in Human Pluripotent Stem Cells. Stem Cells Trans. Med. *3*, 792–800.

Paulsen, B.S., Mandal, P.K., Frock, R.L., Boyraz, B., Yadav, R., Upadhyayula, S., Gutierrez-Martinez, P., Ebina, W., Fasth, A., Kirchhausen, T., et al. (2017). Ectopic expression of RAD52 and dn53BP1 improves homology-directed repair during CRISPR-Cas9 genome editing. Nat. Biomed. Eng. *1*, 878–888.

Pavel-Dinu, M., Wiebking, V., Dejene, B.T., Srifa, W., Mantri, S., Nicolas, C.E., Lee, C., Bao, G., Kildebeck, E.J., Punjya, N., et al. (2019). Gene correction for SCID-X1 in long-term hematopoietic. Nat. Comm. Published online April 9, 2019. https://doi.org/10.1038/s41467-019-09614-y.

Schiroli, G., Ferrari, S., Conway, A., Jacob, A., Capo, V., Albano, L., Plati, T., Castiello, M.C., Sanvito, F., Gennery, A.R., et al. (2017). Preclinical modeling highlights the therapeutic potential of hematopoietic stem cell gene editing for correction of SCID-X1. Sci. Transl. Med. *9*, eaan0820.

Sebastiano, V., Maeder, M.L., Angstman, J.F., Haddad, B., Khayter, C., Yeo, D.T., Goodwin, M.J., Hawkins, J.S., Ramirez, C.L., Batista, L.F., et al. (2011). In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. Stem Cells *29*, 1717–1726.

Takayama, N., Nishimura, S., Nakamura, S., Shimizu, T., Ohnishi, R., Endo, H., Yamaguchi, T., Otsu, M., Nishimura, K., Nakanishi, M., et al. (2010). Transient activation of *c-MYC* expression is critical for efficient platelet generation from human induced pluripotent stem cells. J. Exp. Med. *207*, 2817–2830.

Vakulskas, C.A., Dever, D.P., Rettig, G.R., Turk, R., Jacobi, A.M., Collingwood, M.A., Bode, N.M., McNeill, M.S., Yan, S., Camarena, J., et al. (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. Nat. Med. *24*, 1216–1224.

Wang, G., Yang, L., Grishin, D., Rios, X., Ye, L.Y., Hu, Y., Li, K., Zhang, D., Church, G.M., and Pu, W.T. (2017). Efficient, footprint-free human iPSC genome editing by consolidation of Cas9/CRISPR and piggyBac technologies. Nat. Protoc. *12*, 88–103.

Yang, L., Guell, M., Byrne, S., Yang, J.L., De Los Angeles, A., Mali, P., Aach, J., Kim-Kiselak, C., Briggs, A.W., Rios, X., et al. (2013). Optimization of scarless human stem cell genome editing. Nucleic Acids Res. *41*, 9049–9061.

Yao, X., Zhang, M., Wang, X., Ying, W., Hu, X., Dai, P., Meng, F., Shi, L., Sun, Y., Yao, N., et al. (2018). Tild-CRISPR allows for efficient and precise gene knockin in mouse and human cells. Dev. Cell *45*, 526–536.e5.

Yu, C., Liu, Y., Ma, T., Liu, K., Xu, S., Zhang, Y., Liu, H., La Russa, M., Xie, M., Ding, S., and Qi, L.S. (2015). Small molecules enhance CRISPR genome editing in pluripotent stem cells. Cell Stem Cell *16*, 142–147.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-human Nanog (D73G4) XP Rabbit mAb	Cell Signaling Technology	RRID: AB_10559205
Cy5 AffiniPure Donkey Anti-Rabbit IgG (H+L)	Jackson Immuno Research Labs	RRID: AB_2340607
Anti-Human Oct4 Goat mAb	Abcam	RRID: AB_776898
Cy3 AffiniPure Donkey Anti-Goat IgG (H+L)	Jackson Immuno Research Labs	RRID: AB_2307351
DAPI	Thermo Fisher Scientific	RRID: AB_2629482
Anti-human CD34 APC (Clone: 4H11)	eBioscience	Cat#: 17-0349-42; RRID: AB_2016672
Anti-human CD43 PE (Clone: 10G7)	BioLegend	Cat#: 343204; RRID: AB_2255209
Anti-human CD45 BV421 (Clone: HI30)	BD Biosciences	Cat#: 563879; RRID: AB_2744402
Anti-human CD71 PE-Cy7 (Clone: CY1G4)	BioLegend	Cat#: 334112; RRID: AB_2563119
Anti-human CD235a FITC (Clone: HI246)	BioLegend	Cat#: 349104, RRID: AB_10613463
Experimental Models: Cell Lines		
Human ES H9	WiCell	N/A
H9 *HBB* KO (H9-HBB-UBC-GFP)	This paper	N/A
H9 *MYD88* KO (H9-HBB-UBC-GFP)	This paper	N/A
H9 *HBA1* KO (H9-HBA1-UBC-GFP)	This paper	N/A
TkDA3-4 iPSC line established from human dermal fibroblasts	Takayama et al., 2010	N/A
iAM9-M2 iPSC line established from human T cells	Ando et al., 2015	N/A
iLC13-F1 iPSC line established from human T cells	Ando et al., 2015	N/A
1759 iPSC line established from human T cells	This paper	N/A
iSB7-M3 iPSC line established from human T cells	Ando et al., 2015	N/A
TkDA3-4 *HBB* KO (TkDA3-4-HBB-UBC-GFP)	This paper	N/A
iAM9 *HBB* KO (iAM9-HBB-UBC-GFP)	This paper	N/A
iLC13 *HBB* KO (iLC13-HBB-UBC-GFP)	This paper	N/A
1759 *HBB* KO (1759-HBB-UBC-GFP)	This paper	N/A
iSB7 *HBB* KO (iSB7-HBB-UBC-GFP)	This paper	N/A
SCD iPSC line established from human dermal fibroblasts	Sebastiano Laboratory	N/A
SCD-corrected iPSC line	This paper	N/A
Chemicals, Peptides, and Recombinant Proteins		
Y-27632 (Rock Inhibitor)	Tocris	Cat# 1254
SsoAdvanced Universal Probes Supermix	Bio-Rad	Cat# 1725280
ddPCR Supermix for Probes (No dUTP)	Bio-Rad	Cat# 1863023
iScript Reverse Transcriptase	Bio-Rad	Cat# 1708840
Phusion Green High-Fidelity DNA Polymerase	Thermo Fisher Scientific	Cat# F534L
QuickExtract DNA Extraction Solution	Lucigen	Cat# QE09050
mTeSR1 maintenance medium for iPSCs	Stem Cell Technologies	Cat# 85850
ACCUTASE cell detachment solution	Stem Cell Technologies	Cat# 07920
Matrigel basement membrane matrix	Corning	Cat# 354234
Y-27632 (Rock Inhibitor)	Tocris	Cat#: 1254
Recombinant Human SCF	Peprotech	Cat#: 300-07
Recombinant Human TPO	Peprotech	Cat#: 300-18
Recombinant Human IL-3	Peprotech	Cat#: 200-03
Recombinant Human EPO	Peprotech	Cat#: 100-64
Human Holo-Transferrin	R&D Systems	Cat#: 2914-HT-100MG

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
MEM alpha	GIBCO	Cat#: 11900-024
FBS	HyClone	n/a
ITS-X	GIBCO	Cat#: 51500-056
Critical Commercial Assays		
CytoSNP-850K BeadChip	Illumina	
GeneJET genomic DNA purification	Thermo Fisher Scientific	Cat# K0721
TaqMan hPSC Scorecard Assay	Thermo Fisher Scientific	Cat# A15871
AAVpro Purification Kit (All Serotypes)	Takara Bio	Cat# 6666
RNeasy Mini Kit	QIAGEN	Cat# 74104
RPLPO Primer/Probe Assay	Thermo Fisher Scientific	Cat# 4333761F
PureLink Expi Endotoxin-Free Maxi Plasmid Purification Kit	Thermo Fisher Scientific	Cat# A33073
Other		
LightCycler 480 II machine	Roche	
QX200 Droplet Digital PCR System	Bio-Rad	
Oligonucleotides		
HBB sgRNA target sequence: 5′-CTTGCCCCACAGGG CAGTAA**CGG**-3′ *sgRNA was modified*	Synthego	CRISPRevolution sgRNA EZ Kit
MYD88 sgRNA target sequence: 5′-*CGCATGTTGAGA GCAGCCAG**GGG**-3′*. sgRNA was modified	Synthego	CRISPRevolution sgRNA EZ Kit
Primers for Off Target Analysis	Table S1	N/A
Primers for Integration Analysis	Table S2	N/A
Primer/probe assays for ddPCr/qPCR	Table S3	N/A
Recombinant DNA		
AAV6 backbone plasmid: pAAV-MCS plasmid	Agilent Technologies	N/A
HBB AAV6: UBC-GFP using AAV6 backbone plasmid	Dever et al., 2016	N/A
SCD Correction AAV6	Dever et al., 2016	N/A
MYD88 AAV6: UBC-GFP using AAV6 backbone plasmid		https://benchling.com/s/seq-3h5cpWXHPCkc1R69UNEt/edit
Software and Algorithms		
TIDE Well Tool		https://tide.nki.nl/
Primer3web version 4.1.0		http://bioinfo.ut.ee/primer3/
COSMID Off-Target Analysis Tool		https://crispr.bme.gatech.edu
PRISM v8 graphing and statistical software	GraphPad Software	N/A
Flowjo-v10	FlowJo LLC	N/A
GenomeStudio Software	Illumina	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Matthew Porteus (mporteus@stanford.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human ESC and iPSCs

Human ES H9 cells (WiCell) were used for *MYD88*, *HBB, CCR5*, *CGD*, *HBA1*, *RAG1*, and *WAS* editing, and the iPSC lines described in Figure 1D were the following: line 1 = 1759, line 2 = iLC13-F1, line 3 = iAM9-M2, line 4 = TkDA3-4, line 5 = iSB7-M3. TkDA3-4 iPSCs and SCD-patient derived iPSCs were established from human dermal fibroblasts (Cell Applications, Inc.) as described previously (Takayama et al., 2010; Sebastiano et al., 2011). Cell lines 1759, iLC13-F1, iAM9-M2, and iSB7-M3 were established from human T cells as described previously (Nishimura et al., 2013). Cells were maintained in mTeSR1 (STEMCELL Technologies) on feeder free Matrigel (Corning)-coated plates. Subculture was performed every 5-6 days by EDTA method and using media supplemented with 10 μM Y-27632 (Tocris). Normal karyotype was determined in post-edited cells.

METHOD DETAILS

AAV6/Cas9 Genome Editing

ESCs or iPSCs were treated with 10 μm ROCK inhibitor (Y-27632) 24 hours pre-electroporation. Cells at 70%–80% confluence were harvested with Accutase (Life Technologies). Prior to electroporation, a nucleofection solution was made by first mixing 150 μg/mL of SpCas9 (Integrated DNA Technologies) and 87.5 μg/mL of sgRNA at 1:3 molar ratio directly, incubating for 10 minutes at room temperature, then diluting with 20 uL of P3 Primary Cell solution (Lonza). For each reaction, 500,000 cells were mixed with the nucleofection solution. Nucleofection was performed using 16-well Nucleocuvette Strip with 4D Nucleofector system (Lonza) using CA137 electroporation code. Immediately after electroporation, cells were transferred into one well of a matrigel-coated 24-well plate containing 500 μl of mTeSR media with 10 μM Y-27632. AAV6 donor vector was added at 100K MOI (measured by qPCR as described previously) directly to cells after plating and incubated at 37°C for 24 hours. Media was changed 24 hours post-editing and 10 μm Y-27632 was removed 48 hours after.

The *HBB* and *MYD88* synthetic sgRNAs were purchased from Synthego with chemically-modified nucleotides at the three terminal positions at both the 5′ and 3′ ends. Modified nucleotides contained 2′-O-methyl 3′-phosphorothioate. The genomic sgRNA target sequences, with PAM in bold, are: *HBB*: 5′-CTTGCCCCACAGGGCAGTAA**CGG**-3′ and sgRNA *MYD88* 5′-*CGCATGTTGAGAG CAGCCAG**GGG**-3′*.

AAV6 Cloning and Production

All AAV integration donors in this paper were cloned into pAAV-MCS plasmid (Agilent Technologies) containing AAV2 ITRs. The *HBB-UbC-GFP* donor creates an insertion of 2.2 kb flanking the Cas9 cut site within exon 1 of *HBB* and contains left and right homology arms that are 540bp and 420 bp, respectively. The *MYD88-UbC-GFP* donor creates an insertion of 2.2 kb within exon 1 of *MYD88* and contains left and right homology arms that are 480 bp and 450 bp, respectively, and flank the Cas9 cut site. The SCD corrective donor contains a total of 2.4 kb sequence homology surrounding the Glu6Val mutation (Dever et al., 2016).

For AAV production, vector plasmids were grown in E.coli and harvested using Invitrogen's Endotoxin-Free Maxi Plasmid Purification Kit (Cat# A33073). Following DNA purification, each of five 15cm^2 dishes of 293FT cells (Life Technologies) were transfected using 120uL 1 mg/mL PEI (MW 25K)(Polysciences) along with 6 μg ITR-containing plasmid and 22 μg pDGM6 (which carried AAV6 cap, AAV2 rep, and adenoviral helper genes)(gift from D. Russell). 72h post-transfection, cells were harvested and purified using a Takara AAVpro Purification Kit (Cat. 6666) using the manufacturer's protocol. Vector titer was determined by using qPCR to measure vector genome concentration as described previously (Aurnhammer et al., 2012).

Flow cytometry

For measuring targeted editing frequency, cells were harvested into single cell suspension with Accutase. FITC channel was used to determine percent GFP+ cells. Data was acquired using an Accuri C6 plus flow cytometer (BD Biosciences).

Genotyping and sequence analysis

HBB and *MYD88 ESC* clones were amplified with the following primers to determine integration (Table S2):
MYD88-FW 5′-GACAAGCTCTCTAACTGGAGAATGA-3′
MYD88-REV 5′-CCTAAGAATGGTACCAAGGTAGGTC-3′
HBB-FW 5′-TAGATGTCCCCAGTTAACCTCCTAT-3′
HBB-REV 5′-TTATTAGGCAGAATCCAGATGCTCA-3′.

Sickle cell clones edited with the anti-sickling donor (Figure 3A) were amplified with the following primers to check for integration

HBB-SCD-FW: 5′-GAAGATATGCTTAGAACCGAGG-3′
HBB-SCD-REV: 5′-CCACATGCCCAGTTTCTATTGG-3′.

PCR was done with Phusion Green HSII Master Mix (Thermo Fisher). Sanger sequencing was performed on gel-extracted PCR amplicons by MCLab (South San Francisco, CA, USA). Genomic DNA sample was prepared by QuickExtract DNA Extraction Solution (Epicenter) following the manufacturer's instructions.

ddPCR analysis

In order to determine HR frequency, genomic DNA was harvested from ESCs and iPSCs post-targeting using QuickExtract (Epicenter). Targeted gDNA was then digested with BamHI-HF following the manufacturer's instructions (New England Biolabs). 1-3 μL of gDNA sample was then used in 25 μL ddPCR Supermix for Probes (no dUTP)(Bio-Rad), following the manufacturer's instructions with integration-specific and reference primer and probes at a primer:probe ratio of 3.6:1. Amplification was performed using in-out PCR (one primer binding to transgene insert in the genomic locus and the other immediately outside the homology arm). Droplet samples were prepared using 20 μL of the PCR mix, 70 μL droplet generation oil, and 40 μL of the droplet sample. PCR cycling conditions were as follows: 98°C (10 minutes); 94°C (30 s), 60°C (30 s), 72 °C (2 minutes)(40 cycles); 98 °C

(10 minutes). Finally, droplets were analyzed according to the manufacturer's instructions using the QX200 Droplet Digital PCR System (Bio-Rad).

Off-target analysis

Gene edited ESC/iPSC lines were tested for off-target editing events predicted for each sgRNA by COSMID46 tool (https://crispr.bme.gatech.edu). COSMID is a website tool that considers guide RNA mismatch, insertions and deletions based on the target sequence. Based on these predictions, *HBB* and SCD-iPSC clones were analyzed for off-target effects at chromosome 9 and chromosome 17. *MYD88* clones were analyzed for off-target effects at chromosome 6 and chromosome 1 using primers detailed in Table S1. In order to determine frequency of INDELs at the above regions, we performed PCR using Phusion Green Master Mix according to manufacturer's instructions (Thermo Fisher). Template DNA for PCR was harvested from ESCs and iPSCs post-targeting using QuickExtract (Epicenter). TIDE software was used to estimate INDEL frequencies using Sanger sequences of the PCR amplicons of edited and unedited populations at the noted off-target loci (Brinkman et al., 2014).

ddPCR-based copy number analysis

Following isolation of gDNA from hPSCs using QuickExtract (Epicenter), ddPCR was used to quantify copy number of the UbC promoter used to drive expression of the selection marker was performed on clones confirmed by PCR to have bi-allelic integrations of UbC-GFP at the *HBB* locus. The QX200 Droplet Digital PCR System (Bio-Rad) and ddPCR Supermix for Probe (noUTP)(Bio-Rad) were used according to manufacturer's protocol. The primer and probes used are detailed in Table S3.

SCD allele correction analysis by nested ddPCR

SCD-iPSCs edited at *HBB* were harvested 7 days post-nucleofection and then analyzed for correction frequencies of the Glu6Val mutation by nested ddPCR. An in-out PCR approach was performed to amplify the *HBB* locus in order to exclude episomal AAV6 from analysis. The following primers were used to generate an *HBB*-specific 1.4 kb band: FW(out): 5′-AGGAAGCAGAACTCTG CACTTCA-3′ and REV(in): 5′-AGTCAGTGCCTATCAGAAACCCAAGAG-3′. The PCR product was purified and then diluted to 10 μg/μL in nuclease-free water and then ddPCR was carried out with the PCR product generated as the template. The protocol used a 2-probe set on the sample amplicon, where one probe (HEX) bound the integrated sequence and the other probe (FAM) is a reference sequence downstream of the Cas9 break site. The following primer/probes were used in the ddPCR reaction: HR probe (HEX)-5′-TGACTCCTGAGGAAAAATCCGCAGTCA-3′, reference probe (FAM)-5′-ACGTGGATGAAGTTGGTGGTGAGG-3′, nested-forward 5′-TCACTAGCAACCTCAAACAGAC-3′, and nested-reverse 5′-CCTGTCTTGTAACCTTGATACC-3′.

Erythrocyte differentiation of hiPSCs

Differentiation of iPSCs into HSPCs was done as reported previously, with slight modifications (Nishimura et al., 2013). In brief, small clumps of iPSCs (<100 cells) were transferred onto over-confluent C3H10T1/2 cells and co-cultured in EB medium (Iscove's modified Dulbecco's medium supplemented with 15% FBS and a cocktail of human insulin, human transferrin, sodium selenite (ITS-X), 2 mM L-glutamine, 0.45 μM monothioglycerol, and 50 μg/ml ascorbic acid) in the presence of 20 ng/ml VEGF, 20 ng/ml SCF, 20 ng/ml TPO, 20 ng/ml IL-3, and 20 ng/ml IL-6. Media was changed every 3 days. After 14 days of cultivation, cells were harvested by treatment with TrypLE Select (Thermo Fisher Scientific) for 5 min at 37°C, then CD34+ cells were isolated by FACS.

Sorted CD34+ cells were seeded on over-confluent C3H10T1/2 cells at ~2.0 × 10^5 cell/cm^2 density and cultured in Alpha MEM supplemented with 15% FBS, 50 ng/ml SCF, 10 ng/ml IL-3, 1 U/ml EPO, and 5.5 μg/ml Human Holo-Transferrin for 1 week. Then 5 U/ml EPO and 5.5 μg/ml Human Holo-Transferrin was added to media for the following 2 weeks. Media was changed every 3 days, and cells were transferred onto new feeder layers every 6 days. All floating cells were collected, filtrated through 40 μm cell strainers, and subjected to FACS sorting of CD71+/CD235a+ cells.

Assessment of mRNA levels in erythrocytes

Edited and unedited populations of SCD patient-derived iPSCs were differentiated into red blood cells (RBCs) using an established protocol (Ochi et al., 2014). Following the differentiation, immunostaining followed by FACS was performed in order to identify and isolate RBCs (identified as CD43-/CD34-/CD45-/CD235a+/CD71+). RNA was then extracted from differentiated RBCs using the RNeasy Mini Plus Kit (QIAGEN). cDNA was made from 100 ng RNA using the SuperScript III First-Strand Synthesis (Invitrogen). Levels of HbS and HbA were quantified by qPCR using the following primers and FAM/ZEN/IBFQ-labeled hydrolysis probes as custom-designed PrimeTime qPCR Assays (IDT): HbS primer (fw): 5′-TCACTAGCAACCTC AAACAGAC-3′, HbS primer (rv): 5′-ATCCACGTTCACCTTGCC-3′, HbS probe: 5′-TAACGG CAGACTTCTCCACAGGAGTCA-3′, HbA primer (fw): 5′-TCACTAGCAAC CTCAAACAGA C-3′, HbA primer (rv): 5′-ATCCACGTTCACCTTGCC-3′, HbA probe: 5′-TGACTGCGGATTTT TCCTCAGGAGTCA-3′. qPCR was then performed on a LightCycler 480 II machine (Roche) using SsoAdvanced Universal Probes Supermix (Bio-Rad) according to manufacturer's instructions. Cycling times were as follows: (1) 98°C, 2 min initial denaturation; (2) 98°C 5 s denaturation; (3) 60°C 20 s annealing and extension; (4) return to step 2 39x. Relative expression levels of individual genes were determined within each sample by comparison to *RPLP0* control probe (Cat# 4333761F; Thermo-Fisher).

Immunofluorescence staining to determine pluripotency

An edited clone for *HBB* was stained for NANOG, OCT4, and DAPI. Cells were fixed in 4% paraformaldehyde, permeabilized in 0.2% Triton X-100 in PBS, blocked with blocking buffer (0.1% Triton-X and 2% FBS in PBS). Anti-NANOG (RRID: AB_10559205) and Anti-OCT4 (RRID: AB_776898) antibodies were diluted 200-fold with blocking buffer and used to stain an *HBB* edited clone at 4°C overnight. Cells were then washed three times and stained with 1:500 diluted Cy5 Donkey Anti-Rabbit IgG (RRID: AB_2340607) and Cy3 Donkey Anti-Goat IgG (RRID: AB_2307351) for 1 hour. Cells were washed again and DAPI (RRID: AB_2629482) staining was used on the third wash. The fluorescence images were acquired using an EVOS FL cell imaging system.

Karyotype analysis

Karyotype analysis was performed by the Cytogenetics lab at Stanford University. Cells were harvested and chromosomes were analyzed using the GTW banding method. Twenty metaphase cells were analyzed, all of which were concluded to have a normal karyotype (46, XY).

High-density SNP array analysis

Genomic DNA of edited clones (200 ng) was prepared by GeneJET genomic DNA purification kit (Thermo-Fisher). SNP analysis was performed by the Stanford Functional Genomics Facility using the Infinium CytoSNP-850k Assay (Illumina). Samples were scanned using the iScan system (Illumina). Data was analyzed using GenomeStudio software with cnvPartition algorithm (Illumina).

QUANTIFICATION AND STATISTICAL ANALYSIS

All data are presented as mean ± SEM. The statistical significance of the observed differences was determined using ANOVA. Statistics were performed using Excel (Microsoft) and Prism (GraphPad software) programs.