

# UCSF

## UC San Francisco Previously Published Works

### Title

Dual  $\alpha$ -globin-truncated erythropoietin receptor knockin restores hemoglobin production in  $\alpha$ -thalassemia-derived erythroid cells

### Permalink

<https://escholarship.org/uc/item/3p53m80j>

### Journal

Cell Reports, 44(1)

### ISSN

2639-1856

### Authors

Chu, Simon N  
Soupene, Eric  
Sharma, Devesh  
[et al.](#)

### Publication Date

2025

### DOI

10.1016/j.celrep.2024.115141

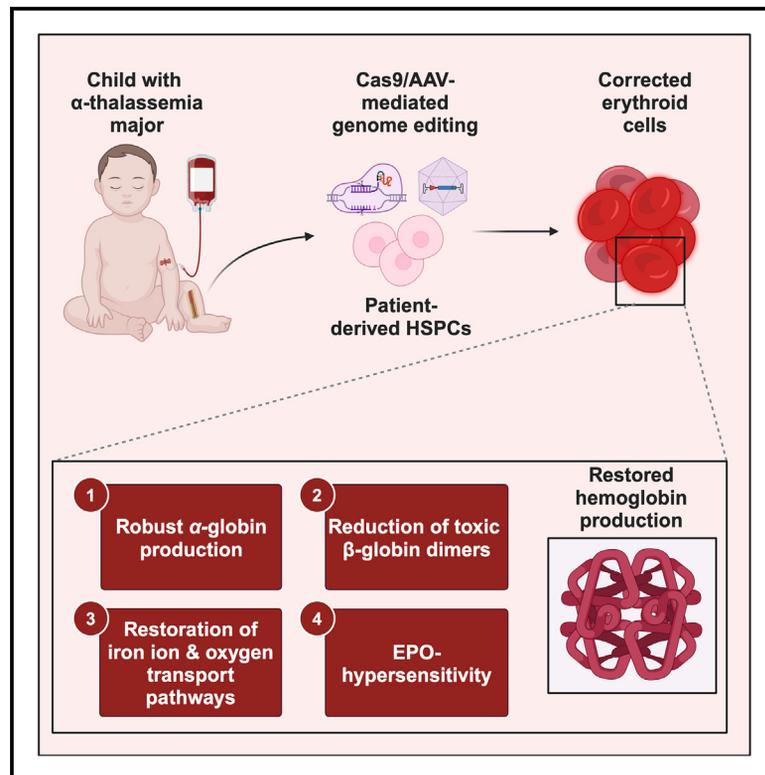
### Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at <https://creativecommons.org/licenses/by-nc/4.0/>

Peer reviewed

# Dual $\alpha$ -globin-truncated erythropoietin receptor knockin restores hemoglobin production in $\alpha$ -thalassemia-derived erythroid cells

## Graphical abstract



## Authors

Simon N. Chu, Eric Soupene, Devesh Sharma, ..., Vivien A. Sheehan, Matthew H. Porteus, M. Kyle Cromer

## Correspondence

kyle.cromer@uscf.edu

## In brief

Chu et al. present a universal correction strategy for  $\alpha$ -thalassemia major using Cas9/AAV6 to integrate a functional  $\alpha$ -globin gene and hypermorphic erythropoietin receptor into the  $\beta$ -globin locus in hematopoietic stem and progenitor cells. The approach restores the hemoglobin production and enhances erythropoietic output, offering a potential curative therapy for  $\alpha$ -thalassemia.

## Highlights

- Cas9/AAV6 genome editing effectively integrates  $\alpha$ -globin gene into the  $\beta$ -globin locus
- Genome editing restores hemoglobin production to  $\alpha$ -thalassemia patient-derived HSPCs
- Addition of hypermorphic EPO receptor enhances erythropoietic output of edited HSPCs
- Edited HSPCs engraft long-term in xenotransplantation models



## Article

# Dual $\alpha$ -globin-truncated erythropoietin receptor knockin restores hemoglobin production in $\alpha$ -thalassemia-derived erythroid cells

Simon N. Chu,<sup>1,2,3</sup> Eric Soupene,<sup>4</sup> Devesh Sharma,<sup>1,2</sup> Roshani Sinha,<sup>1,2</sup> Travis McCreary,<sup>1,2</sup> Britney Hernandez,<sup>5</sup> Huifeng Shen,<sup>5</sup> Beeke Wienert,<sup>6</sup> Chance Bowman,<sup>1,2,7</sup> Han Yin,<sup>1,3</sup> Benjamin J. Lesch,<sup>1,2,7</sup> Kun Jia,<sup>1,2</sup> Kathleen A. Romero,<sup>5</sup> Zachary Kostamo,<sup>5</sup> Yankai Zhang,<sup>5</sup> Tammy Tran,<sup>1,2</sup> Marco Cordero,<sup>1,2</sup> Shota Homma,<sup>8</sup> Jessica P. Hampton,<sup>9</sup> James M. Gardner,<sup>1,3</sup> Bruce R. Conklin,<sup>6,10</sup> Tippi C. MacKenzie,<sup>1,2</sup> Vivien A. Sheehan,<sup>5</sup> Matthew H. Porteus,<sup>9</sup> and M. Kyle Cromer<sup>1,2,7,11,\*</sup>

<sup>1</sup>Department of Surgery, University of California, San Francisco, San Francisco, CA 94143, USA

<sup>2</sup>Eli & Edythe Broad Center for Regeneration Medicine, University of California, San Francisco, San Francisco, CA 94143, USA

<sup>3</sup>Diabetes Center, University of California, San Francisco, San Francisco, CA 94143, USA

<sup>4</sup>Department of Pediatrics, University of California, San Francisco, Oakland, CA 94609, USA

<sup>5</sup>Aflac Cancer and Blood Disorders Center, Children's Healthcare of Atlanta, Emory University School of Medicine, Atlanta, GA 30329, USA

<sup>6</sup>Gladstone Institutes, San Francisco, CA 94158, USA

<sup>7</sup>Department of Bioengineering & Therapeutic Sciences, University of California, San Francisco, San Francisco, CA 94158, USA

<sup>8</sup>Department of Genetics, Stanford University, Stanford, CA 94305, USA

<sup>9</sup>Department of Pediatrics, Stanford University, Stanford, CA 94305, USA

<sup>10</sup>Department of Medicine, University of California, San Francisco, San Francisco, CA 94158, USA

<sup>11</sup>Lead contact

\*Correspondence: [kyle.cromer@uscf.edu](mailto:kyle.cromer@uscf.edu)

<https://doi.org/10.1016/j.celrep.2024.115141>

## SUMMARY

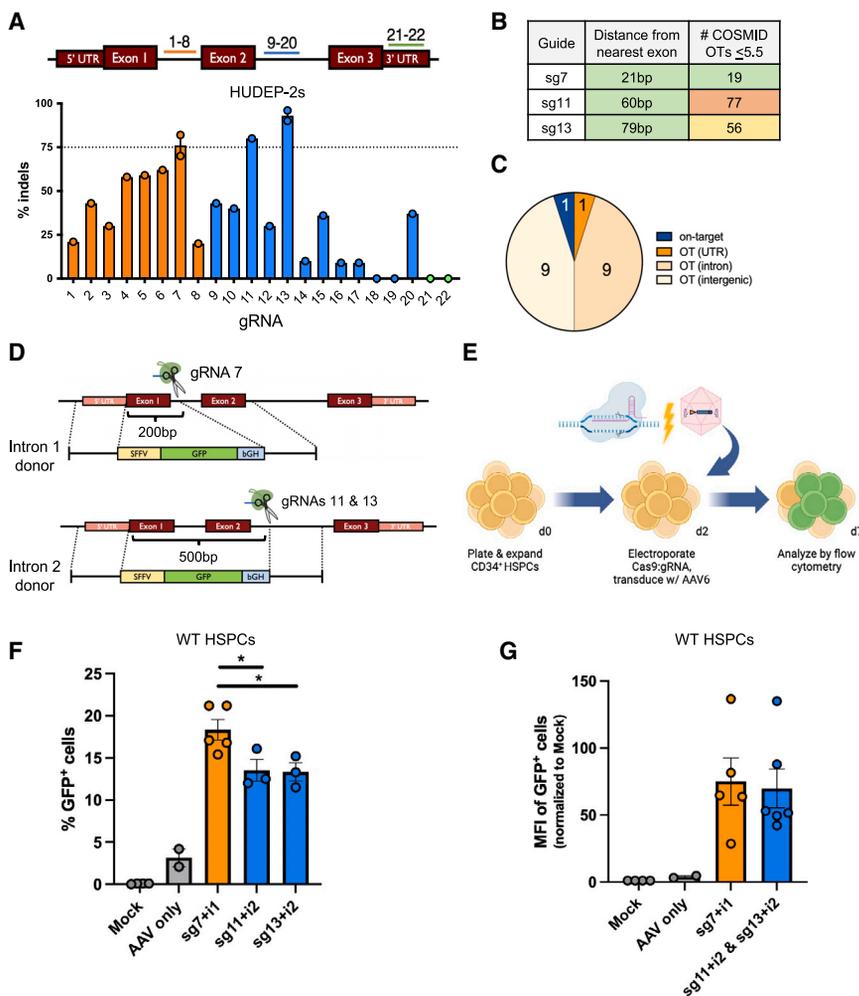
The most severe form of  $\alpha$ -thalassemia results from loss of all four copies of  $\alpha$ -globin. Postnatally, patients face challenges similar to  $\beta$ -thalassemia, including severe anemia and erythrotoxicity due to the imbalance of  $\beta$ -globin and  $\alpha$ -globin chains. Despite progress in genome editing treatments for  $\beta$ -thalassemia, there is no analogous curative option for  $\alpha$ -thalassemia. To address this, we designed a Cas9/AAV6-mediated genome editing strategy that integrates a functional  $\alpha$ -globin gene into the  $\beta$ -globin locus in  $\alpha$ -thalassemia patient-derived hematopoietic stem and progenitor cells (HSPCs). Incorporation of a truncated erythropoietin receptor transgene into the  $\alpha$ -globin integration cassette significantly increased erythropoietic output from edited HSPCs and led to the most robust production of  $\alpha$ -globin, and consequently hemoglobin tetramers. By directing edited HSPCs toward increased production of clinically relevant erythroid cells, this approach has the potential to mitigate the limitations of current treatments for the hemoglobinopathies, including low genome editing and low engraftment rates.

## INTRODUCTION

$\alpha$ -Thalassemia is one of the most common monogenic diseases in the world. It is currently estimated that 5% of the world's population harbors an  $\alpha$ -thalassemia variant.<sup>1</sup> While the carrier frequency is highest in those with southeast Asian heritage, data suggests that these rates are rising due to population growth, human migration, and advances in treating milder forms of the disease.<sup>2</sup> Humans harbor four copies of  $\alpha$ -globin genes (a direct tandem repeat on chromosome 16), and disease severity directly correlates with the number of mutated or deleted alleles. The most severe form of the disease, in which all four copies are disrupted, is called  $\alpha$ -thalassemia major ( $\alpha$ TM). While this disease has historically been lethal *in utero*, patients may now survive to birth after *in utero* blood transfusions, often with excellent neurological outcomes.<sup>3,4</sup> Postnatally, these patients present with a dis-

ease similar to  $\beta$ -thalassemia, with the absence of  $\alpha$ -globin leading to an inability to form hemoglobin heterotetramers and subsequent severe anemia. Moreover, the accumulation of orphan  $\beta$ -globin chains, which form oxidized, covalently linked dimers and an unstable homotetramer (hemoglobin H), leads to erythrotoxicity and hemolysis.<sup>5</sup> As a consequence, these patients require chronic transfusions, which often results in iron overload and need for iron chelation therapy.<sup>6</sup> Although allogeneic-hematopoietic stem cell transplantation (HSCT) may provide a cure, suitable matched donors are only available in a minority of cases and carry a risk of immune rejection and graft-versus-host disease (GvHD).<sup>7</sup> Furthermore, while numerous gene therapy and genome editing strategies have been developed for patients with  $\beta$ -thalassemia,<sup>8–10</sup> there are no such therapies for the most severe forms of  $\alpha$ -thalassemia, indicating a major unmet medical need for this patient population.





**Figure 1. Efficient integration of HDR templates at *HBB* locus**

(A) Schematic of *HBB*-targeting gRNAs and resulting indel frequencies in HUDEP-2 cells following electroporation-mediated delivery of Cas9 protein:gRNA complexes. The dotted line indicates indel frequencies at 75% threshold. Bars represent mean  $\pm$  standard error of the mean (SEM).

(B) *In silico* off-target analysis of gRNA 7, 11, and 13 using COSMID.

(C) Pie chart displaying genomic features at 19 predicted off-target sites for sg7.

(D) Schematic of custom AAV6 DNA repair donors designed to mediate integration at intron 1 using sg7 or intron 2 using sg11 or 13. The bracket indicates the approximate distance between left and right homology arms.

(E) Schematic of Cas9/AAV6 genome editing workflow in primary HSPCs.

(F) Percentage of GFP<sup>+</sup> cells following editing in WT HSPCs was determined at day 5 post-editing using flow cytometry. “sg7+i1” shorthand refers to gRNA 7 (shown in A) paired with intron 1 donor (shown in D). Bars represent mean  $\pm$  SEM. \* $p < 0.05$  by unpaired two-tailed t test.

(G) Mean fluorescence intensity (MFI) of GFP<sup>+</sup> cells (from F) was determined by flow cytometry. Bars represent mean  $\pm$  SEM.

with  $\beta$ -thalassemia,<sup>8</sup> this strategy could be used to develop an autologous-HSCT treatment for  $\alpha$ TM, overcoming the shortage of matched donors and ameliorating the risk of immune rejection and GvHD for patients suffering from this disease.

In this work, we developed a Cas9/adenovirus type 6 (AAV6)-mediated genome editing strategy to integrate a functional copy of the  $\alpha$ -globin gene into the  $\beta$ -globin locus in  $\alpha$ TM patient-derived hematopoietic stem and progenitor cells (HSPCs) as a universal treatment strategy for the disorder—an approach that could be curative for any  $\alpha$ TM patient regardless of the specific causative mutations or deletions. This approach allows both correction of the underlying disease and restoration of the  $\beta$ -globin: $\alpha$ -globin imbalance by placing the  $\alpha$ -globin transgene under erythroid-specific expression of the  $\beta$ -globin locus. Taking a cue from human genetics, we also incorporated a naturally occurring truncated erythropoietin receptor (*tEPOR*) cDNA into the  $\alpha$ -globin integration cassette, allowing for simultaneous correction of  $\alpha$ -thalassemia and increased erythropoietic output from edited HSPCs.<sup>11</sup> By directing edited HSPCs toward increased production of clinically relevant erythroid cells instead of other divergent cell types, this approach has the potential to overcome many of the clinical challenges of HSCT for the treatment of the hemoglobinopathies, including low editing and engraftment rates and high morbidity from prerequisite myeloablative regimens.<sup>12–14</sup> Similar to our prior work

## RESULTS

### Efficient Cas9 cleavage with *HBB* intron-targeting gRNAs

Due to the high editing frequencies achieved in primary HSPCs in prior work,<sup>8,15</sup> we sought to use a Cas9/AAV6-mediated genome editing strategy to knock an  $\alpha$ -globin (*HBA*) transgene into the  $\beta$ -globin (*HBB*) locus in patient-derived HSPCs. To do so, we first designed and screened Cas9 guide RNAs (gRNAs) at the *HBB* locus. So that cleavage alone without homology-directed repair (HDR) would not disrupt  $\beta$ -globin production, we chose 22 candidate gRNAs located in intron 1 or intron 2 as well as the 3' UTR (Figure 1A; Table S1). To determine cleavage frequencies, we delivered candidate gRNAs pre-complexed with Cas9 protein to the human HUDEP-2 cell line<sup>16</sup> via transfection. Genomic DNA was harvested several days post-editing and subjected to PCR amplification of the region surrounding the expected cleavage site. Insertion and deletion (indel) frequencies of the corresponding Sanger sequences were then quantified using inference of CRISPR edits (ICE) analysis.<sup>17</sup> We found that editing the 5' and 3' UTR regions of *HBB* proved not to be feasible, as the former had significant homology to  $\delta$ -globin and the latter had no

gRNAs with detectable cleavage. We found that the most efficient gRNAs corresponded to sg7 in intron 1 and to sg11 and 13 in intron 2 (Figures 1A and S1A).<sup>18</sup> *In silico* off-target analysis of these guides using COSMID (CRISPR Off-Target Sites with Mismatches, Insertions, and Deletions)<sup>19,20</sup> revealed that sg7 had the most favorable predicted off-target profile (Figure 1B). On further evaluation of the 19 predicted off-target sites for sg7, all were found to reside in non-coding regions of the genome (Figure 1C).

### Efficient integration of HDR templates at *HBB* locus

Following the identification of effective *HBB* intron-targeting gRNAs, we developed DNA repair templates packaged in AAV6 delivery vectors that could effectively mediate HDR at these cleavage sites (termed intron 1 and intron 2 donors; Figure 1D). Each integration cassette comprised a spleen focus-forming virus (SFFV) promoter driving the expression of a GFP reporter to allow rapid readout of integration frequencies via flow cytometry. While the right homology arm of each DNA repair template corresponded to the ~900 bp immediately downstream of the intron 1 or intron 2 Cas9 cut site, the left homology arm was split away from the cleavage sites to correspond to the ~900 bp immediately upstream of the start codon of the endogenous *HBB* gene. As in a prior study,<sup>8</sup> this split homology arm strategy is expected to allow promoterless integration cassettes—ultimately, an  $\alpha$ -globin transgene in this work—to be driven by the regulatory machinery of the endogenous locus.

Once DNA repair templates were assembled and packaged into AAV6 vectors, we tested each integration strategy by delivering the most effective gRNAs (sg7, sg11, or sg13) complexed with high-fidelity Cas9 protein<sup>21</sup> to wild-type (WT) human primary CD34<sup>+</sup>-enriched HSPCs via electroporation. Immediately following electroporation, we transduced cells with AAV6 vectors corresponding to either intron 1 or intron 2 integration schemes. Five days later, we analyzed editing frequencies by flow cytometry (Figures 1E and S2A). We aimed to determine the effectiveness of transgene knockin with various gRNA and homology arm combinations using the most effective gRNAs defined in Figure 1A. We found that the intron 1 integration strategy was significantly more efficient (median of 17.1% GFP<sup>+</sup> cells) compared to intron 2 integration (median of 12.5% for sg11 and 13.3% for sg13;  $p < 0.05$ ) (Figure 1F), perhaps due to the shorter distance that the left homology arm was split away from the Cas9 cleavage site. In addition, both integration strategies yielded a high mean fluorescence intensity (MFI) per edited cell (Figure 1G). Given the high cleavage and integration frequency, as well as the favorable off-target profile of sg7, we proceeded with further testing in primary WT HSPCs using the intron 1 integration strategy.

### Integration at *HBB* drives erythroid-specific expression of transgenes

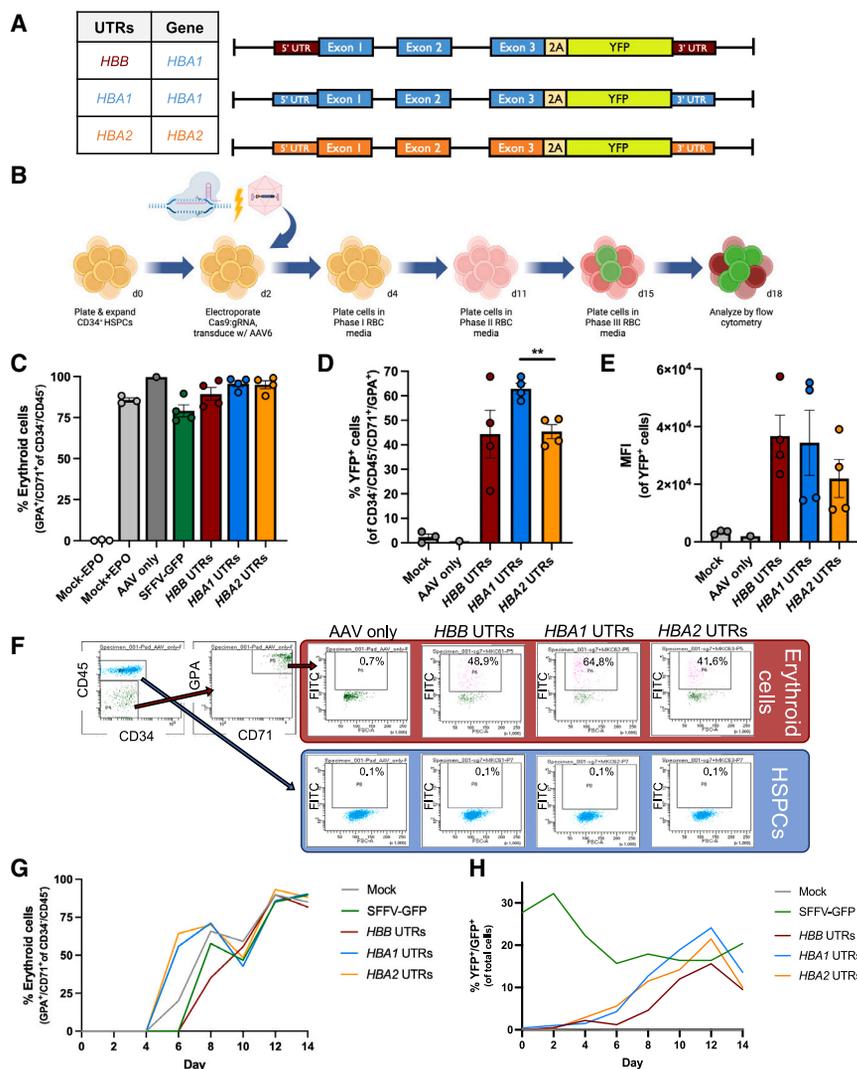
A key goal of this work is to achieve erythroid-specific expression of an  $\alpha$ -globin transgene. To test whether the *HBB* intron 1 integration strategy is able to achieve this, we designed custom promoterless integration cassettes that included either the full-length *HBA1* or *HBA2* transgene (i.e., including all UTRs, exons, and introns) linked to a 2A cleavage peptide

derived from *Thosea asigna* virus (T2A)-yellow fluorescent protein (YFP) reporter to allow quantification of HDR frequency by flow cytometry and to serve as a surrogate for  $\alpha$ -globin protein production (Figure 2A). Because the specific UTRs flanking an integrated transgene can have a great bearing on expression,<sup>8</sup> we also designed integration strategies that would allow our *HBA* transgene to be flanked by *HBB*, *HBA1*, and *HBA2* 5' and 3' UTRs. Importantly, the left homology arm does not contain the *HBB* 5' UTR, resulting in integrations with only a single 5' UTR corresponding to *HBB*, *HBA1*, or *HBA2* for each respective vector.

To determine whether integration of these promoterless  $\alpha$ -globin cassettes at *HBB* yielded erythroid-specific expression, we edited WT HSPCs using electroporation-mediated delivery of Cas9 protein:sg7 gRNA ribonucleoprotein (RNP) followed by transduction with AAV6 vectors containing each repair template. We then performed an established, 14-day erythroid differentiation protocol<sup>22</sup> and quantified YFP expression and erythroid differentiation (staining with antibodies for CD34, CD45, CD71, and GPA) via flow cytometry (Figures 2B and S2B). We examined the percentage of CD34<sup>+</sup>/CD45<sup>-</sup> cells that acquired the erythroid cell surface markers GPA and CD71 on day 14 of erythroid differentiation, when erythroid cells are expected to be at their most mature stage. Our results demonstrated efficient acquisition of erythroid markers across all experimental conditions, with the exception of unedited (mock) cells cultured in the absence of EPO (Figure 2C). However, we did observe differences among the integration vectors in terms of HDR frequency and MFI of edited cells. We found that the vector integrating the *HBA1* transgene and *HBA1* UTRs achieved the most highly HDR-edited population of erythroid cells (median of 63% YFP<sup>+</sup> cells; Figure 2D) and the second highest MFI of the HDR-edited cell population (Figure 2E), although these changes did not reach statistical significance. As confirmation that these integration strategies achieved erythroid-specific transgene expression, we observed high YFP fluorescence in cells that acquired erythroid cell markers (CD71<sup>+</sup>/GPA<sup>+</sup>) toward the end of erythroid differentiation in all editing conditions but no fluorescence above background in cells retaining HSPC markers (CD34<sup>+</sup>/CD45<sup>+</sup>) (Figure 2F). We also tracked differentiation and YFP expression kinetics and found that all editing conditions proceed through erythroid differentiation at a similar rate (Figure 2G). When monitoring the expression of fluorescent reporters, we find that cells edited with the constitutive SFFV-GFP integration vector maintain high expression over the course of erythroid differentiation, whereas all three conditions edited with *HBA*-2A-YFP constructs display increasing fluorescence that dovetails with erythroid differentiation (Figure 2H). Collectively, these results indicate that integration of the *HBA* transgene at *HBB* achieves high levels of erythroid-specific transgene expression.

### $\alpha$ -Globin integration in $\alpha$ TM HSPCs yields low levels of hemoglobin

While T2A-YFP reporters allow rapid readout of HDR-editing frequency, to generate clinically relevant constructs, we created *HBA1* integration vectors without T2A-YFP that were flanked by either *HBB* or *HBA1* UTRs (Figure 3A). To test these additional vectors, we edited WT HSPCs as before and found that both



**Figure 2. Integration at *HBB* drives erythroid-specific expression of transgenes**

(A) Schematic of custom AAV6 DNA repair donors designed to integrate promoterless *HBA*-2A-YFP transgenes at the start codon of *HBB*. The table to the left indicates whether cassette integrated *HBB*, *HBA1*, or *HBA2* UTRs, as well as *HBA1* or *HBA2* transgene.

(B) Schematic of Cas9/AAV6 genome editing workflow in primary HSPCs followed by *in vitro* erythroid differentiation.

(C) Percentage of  $CD34^+/CD45^-$  HSPCs acquiring erythroid cell surface markers GPA and CD71 as determined by flow cytometry. Bars represent mean  $\pm$  SEM.

(D) Percentage of YFP<sup>+</sup> cells among  $CD34^+/CD45^-/CD71^+/GPA^+$  cells was determined at day 14 of erythroid differentiation using flow cytometry. Bars represent mean  $\pm$  SEM. \*\* $p < 0.005$  by unpaired two-tailed t test.

(E) MFI of YFP<sup>+</sup> cells (from D) was determined by flow cytometry. Bars represent mean  $\pm$  SEM.

(F) At day 11 of erythroid cell differentiation, cells were stained for HSPC/erythroid markers and analyzed by flow cytometry. The percentage of YFP<sup>+</sup> cells are noted ( $n = 1$ ).

(G) The percentage of  $CD34^+/CD45^-$  cells that acquired CD71 and GPA erythroid cell markers are plotted over the course of erythroid differentiation ( $n = 1$ ).

(H) The percentage of YFP<sup>+</sup> or GFP<sup>+</sup> cells are plotted over the course of erythroid differentiation ( $n = 1$ ).

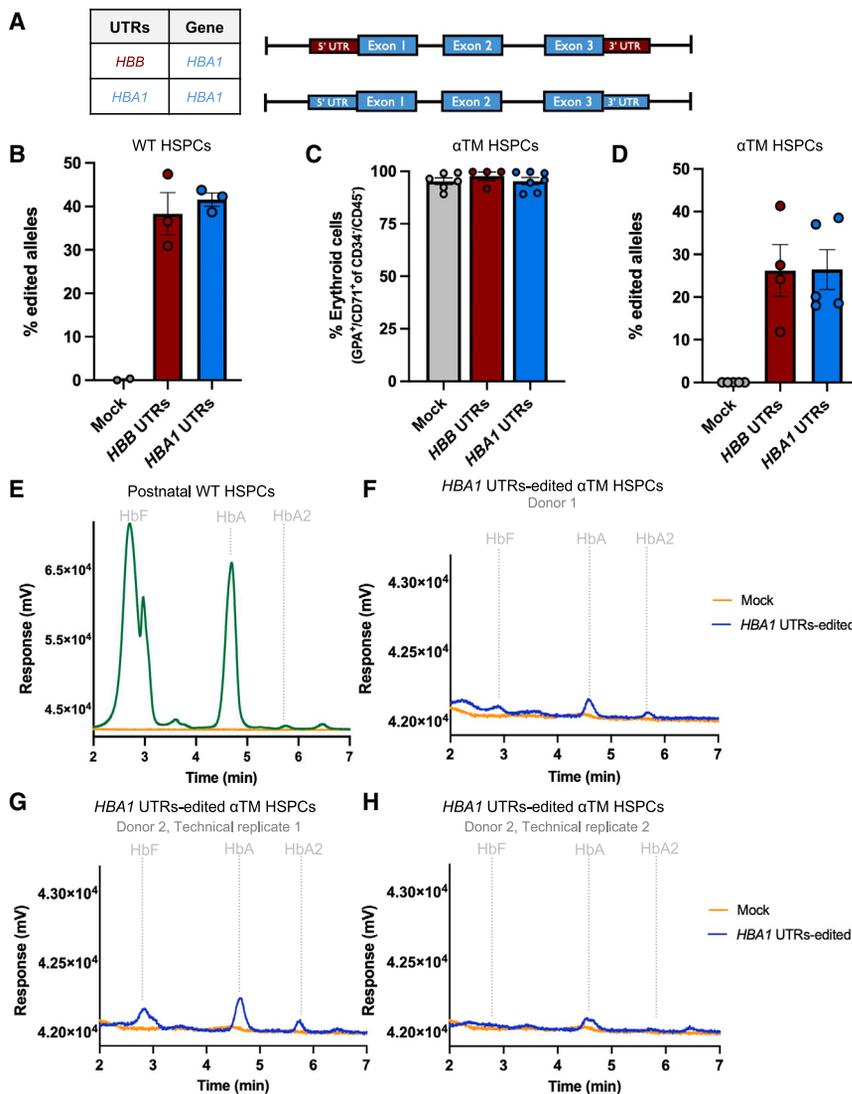
vectors mediate high frequencies of HDR (median of 36.6% and 42.3% HDR-edited alleles for *HBB* and *HBA1* UTR vectors, respectively), as measured by a custom droplet digital PCR (ddPCR) assay (Figure 3B).

We next sought to determine whether  $\alpha$ -globin integration can restore hemoglobin tetramer production in patient-derived HSPCs following erythroid differentiation. We isolated  $CD34^+$  HSPCs from a bone marrow aspirate that was taken from a  $\sim 1$ -year-old patient with  $\alpha$ TM. Clinical information regarding  $\alpha$ TM donors, including information regarding genotype and age of the patient at the time of bone marrow aspirate are provided in Table S2. We then edited and differentiated HSPCs as before and found that all conditions were able to effectively acquire erythroid markers by day 14 of erythroid differentiation (Figure 3C). Following ddPCR analysis, we found that cells were edited at a median of 25.8% and 20.1% HDR-edited alleles for *HBB* and *HBA1* UTR vectors, respectively (Figure 3D). We then performed hemoglobin tetramer high-performance liquid chromatography (HPLC) analysis on edited patient-derived  $\alpha$ TM cells

and S3A). While these results demonstrate for the first time that gene therapy or genome editing may be used to increase  $\alpha$ -globin production in  $\alpha$ TM patient-derived erythroid cells, it is unlikely that this low level of  $\alpha$ -globin production will yield significant clinical benefit.

### Dual integration of $\alpha$ -globin and tEPOR in $\alpha$ TM HSPCs increases hemoglobin production

Given that  $\alpha$ -globin integration into *HBB* does not substantially restore hemoglobin production, we sought alternative strategies to increase the therapeutic potential of this approach. As described in a cohort that included an Olympic gold medal-winning cross-country skier,<sup>23</sup> a premature stop codon in the EPO receptor (*EPOR*) was linked with a condition called congenital erythrocytosis, which is characterized by non-pathogenic hyper-production of red blood cells (RBCs). In subsequent work,<sup>11</sup> we demonstrated that genome editing-mediated integration of a truncated *EPOR* (*tEPOR*)-coding cDNA could increase erythropoietic output from primary human HSPCs.



**Figure 3.  $\alpha$ -Globin integration in  $\alpha$ TM HSPCs yields low levels of hemoglobin**

(A) Schematic of custom AAV6 DNA repair donors designed to integrate promoterless *HBA* transgenes at the start codon of *HBB*. Table to the left indicates whether the cassette integrated *HBB* or *HBA1* UTRs along with *HBA1* transgene.

(B) The percentage of edited alleles in WT HSPCs at day 14 of erythroid differentiation. Bars represent mean  $\pm$  SEM.

(C) Percentage of CD34<sup>+</sup>/CD45<sup>+</sup>  $\alpha$ TM HSPCs acquiring erythroid cell surface markers as determined by flow cytometry. Bars represent mean  $\pm$  SEM.

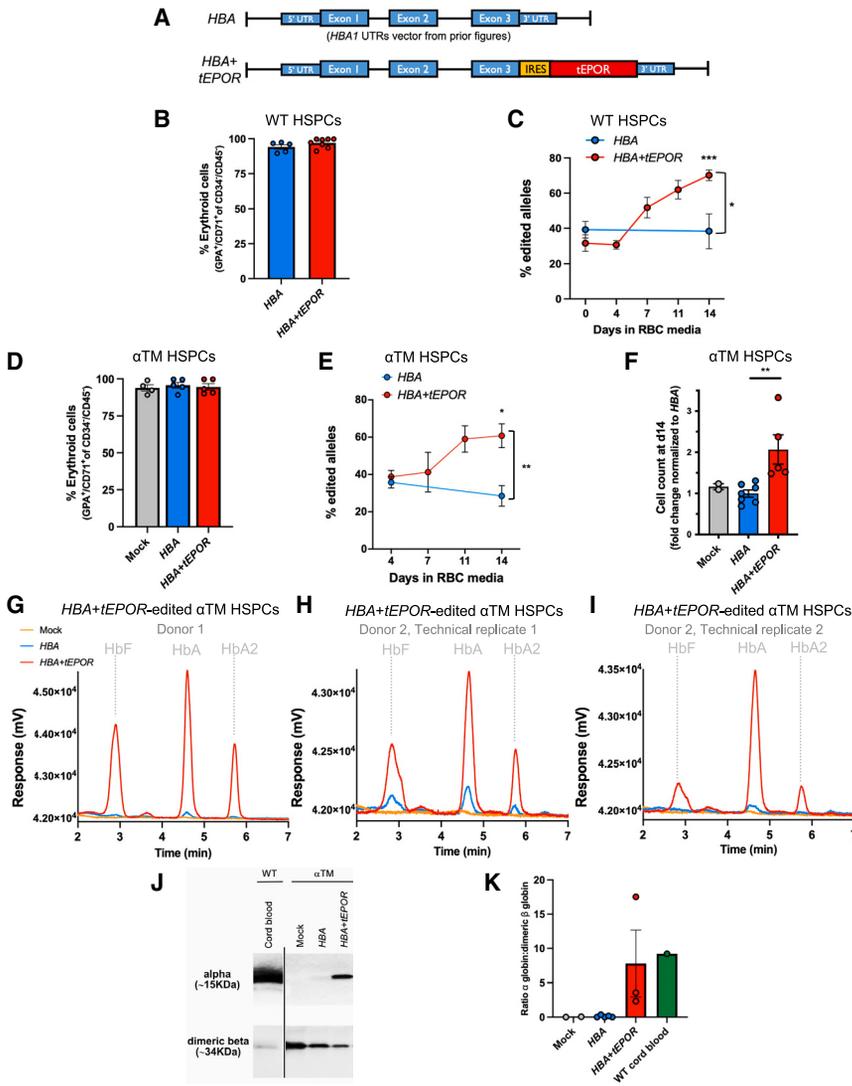
(D) The percentage of edited alleles in  $\alpha$ TM HSPCs at day 14 of erythroid differentiation. Bars represent mean  $\pm$  SEM.

(E) HPLC elution chromatogram displaying the hemoglobin tetramer profile from WT healthy control HSPCs following *in vitro* erythroid differentiation. Time displayed on the x axis represents retention time in minutes for each hemoglobin tetramer type to elute. Absorbance on the y axis indicates the concentration of a particular hemoglobin tetramer.

(F–H) HPLC elution chromatograms displaying the hemoglobin tetramer profile from  $\alpha$ TM HSPCs that have undergone editing and erythroid differentiation. Chromatograms represent two different donors (F and G) and a technical replicate from donor 2 (H) that was edited independently.

Thus, we hypothesized that co-expression of *tEPOR* cDNA from our  $\alpha$ -globin integration cassette could lend a selective advantage to genome-edited and, therefore,  $\alpha$ -globin-expressing erythroid cells. To test this, we added a *tEPOR* cDNA at the C terminus of the  $\alpha$ -globin integration cassette with *HBA1* UTRs, with expression linked by an internal ribosome entry site (Figure 4A). The *tEPOR* variant used is W439X, which was reported in the original publication.<sup>23</sup> We then performed editing via HDR in WT HSPCs with our *HBA1*-UTRs  $\alpha$ -globin cassette from Figure 3 and compared it against the bicistronic *HBA*+*tEPOR* cassette in WT HSPCs. Sanger sequencing of the *HBA*+*tEPOR* vector in edited cells at the 5' and 3' ends confirmed correct integration and the absence of inverted terminal repeats from the AAV vector (Figure S1B). In both instances, edited cells efficiently acquired erythroid markers by day 14 of erythroid differentiation (Figure 4B) and maintained viability comparable to that of unedited cells cultured with EPO (Figure S1C). While cells edited with the *HBA* integration cassette maintained a constant editing rate (median of 39.2% edited alleles at day 0 and 38.4% at day 14),

prior work, this suggests that cells expressing a *tEPOR* have a significant competitive advantage over unedited cells during erythroid differentiation. We next tested these vectors in  $\alpha$ TM patient-derived HSPCs and again observed that edited and unedited cells effectively acquired erythroid markers, as assessed by flow cytometry on day 14 of erythroid culture (Figure 4D). As with WT cells, we found that HDR-editing frequencies among cells edited with the *HBA* vector maintained a consistent editing rate over the course of erythroid differentiation (Figure 4E). However, we again observed a dramatic increase in HDR frequencies in cells edited with the *HBA*+*tEPOR* cassette, rising from a median of 40.5% HDR-edited alleles at day 4 to 60.5% at day 14 ( $*p = 0.02$ ). Similar to the increase in cell counts observed in WT HSPCs, we also observed a 2-fold increase in cell counts at the end of erythroid differentiation in conditions edited with the *HBA*+*tEPOR* cassette compared to cells edited with the  $\alpha$ -globin vector ( $**p = 0.007$ ) (Figure 4F), indicating that *tEPOR* expression is driving increased erythropoietic output from genome-edited HSPCs.



**Figure 4. Dual integration of  $\alpha$ -globin and tEPOR in  $\alpha$ TM HSPCs increases hemoglobin production**

(A) Schematic of custom AAV6 donors designed to integrate promoterless HBA and HBA+tEPOR transgenes at the start codon of HBB. Both vectors are flanked by HBA1 UTRs.

(B) The percentage of CD34<sup>+</sup>/CD45<sup>+</sup> WT HSPCs acquiring erythroid cell surface markers as determined by flow cytometry. Bars represent mean  $\pm$  SEM.

(C) The percentage of edited alleles in WT HSPCs over the course of erythroid differentiation. Bars represent mean  $\pm$  SEM. \* $p$  = 0.01 comparing editing frequencies of HBA-edited cells at day 14 vs. HBA+tEPOR-edited cells at day 14 of differentiation by unpaired two-tailed t test; \*\*\* $p$  = 0.0004 comparing editing frequencies of HBA+tEPOR-edited cells at day 0 vs. day 14 of differentiation by unpaired two-tailed t test.

(D) The percentage of CD34<sup>+</sup>/CD45<sup>+</sup>  $\alpha$ TM HSPCs acquiring GPA and CD71 as determined by flow cytometry. Bars represent mean  $\pm$  SEM.

(E) The percentage of edited alleles in  $\alpha$ TM HSPCs over the course of erythroid differentiation. Bars represent mean  $\pm$  SEM. \* $p$  = 0.02 comparing editing frequencies of HBA+tEPOR-edited cells at day 0 vs. day 14 of differentiation by unpaired two-tailed t test; \*\* $p$  = 0.009 comparing editing frequencies of HBA-edited cells at day 14 vs. HBA+tEPOR-edited cells at day 14 of differentiation by unpaired two-tailed t test.

(F) Cell count at day 14 of erythroid cell differentiation with fold change normalized to HBA. Bars represent mean  $\pm$  SEM. \*\* $p$  = 0.007 by unpaired two-tailed t test.

(G–I) Hemoglobin tetramer HPLC plots from two different donors (G and H) and a technical replicate from donor 2 (I) that was edited independently. For comparison purposes, the HPLC plots for HBA-edited cells, originally shown in Figures 3F–3H, are also presented here.

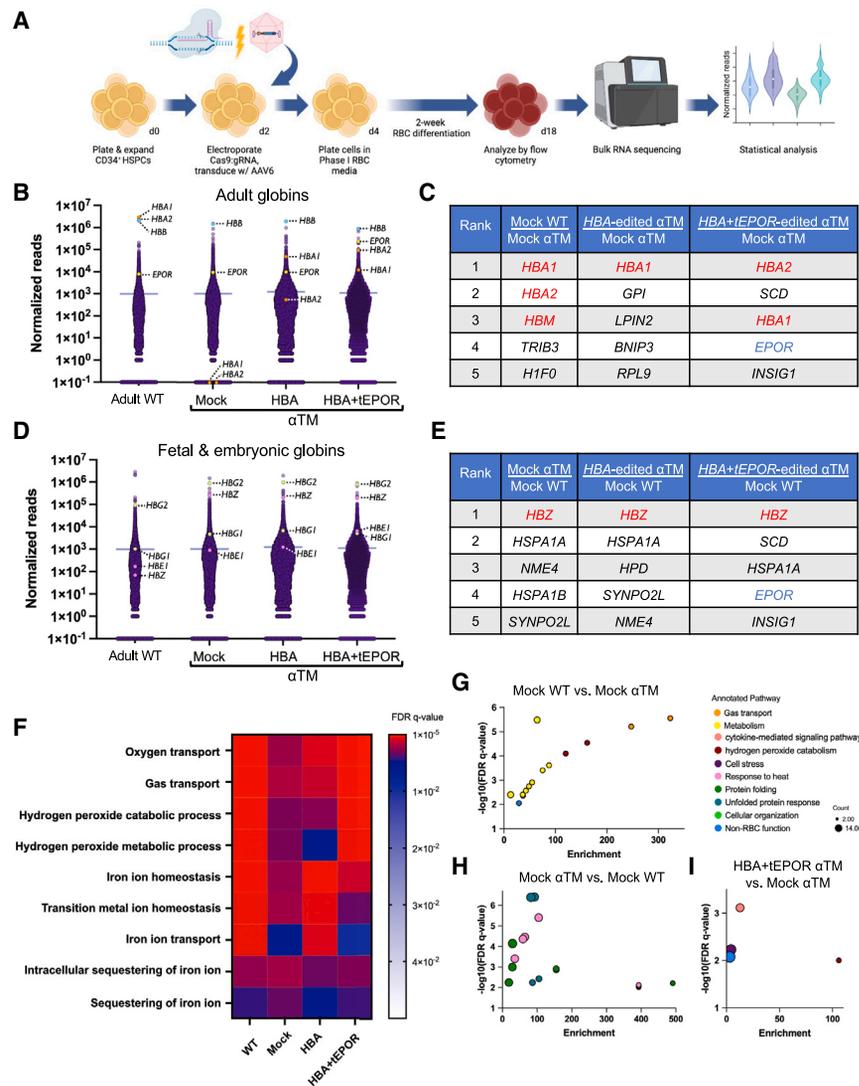
(J) Western blot of mock and edited  $\alpha$ TM HSPCs at the end of erythroid cell differentiation compared to WT umbilical cord blood-derived erythroid cells. The western blot image was taken from a single gel that was cropped to place the WT control next to edited conditions, as indicated by the black line. Loading was standardized by using the same number of cells for input.

(K) Ratio of  $\alpha$ -globin to dimeric  $\beta$ -globin quantification from western blot. Bars represent mean  $\pm$  SEM.

To further characterize cells edited with our two clinical editing strategies, HBA and HBA+tEPOR, we performed deeper characterization and morphological analysis of edited HSPCs at the end of erythroid differentiation. Staining for CD36, a marker of early erythroid differentiation, revealed no difference between unedited cells compared to those edited with the HBA vector (Figures S4A and S4B). However, we did observe a significant increase in the proportion of CD36<sup>+</sup> cells in the HBA+tEPOR editing condition. We also stained cells with DRAQ5, a marker of enucleated erythrocytes, which revealed that our *in vitro* differentiation is able to achieve terminal differentiation of mature erythroid cells (Figures S4C and S4D). We did find that the proportion of enucleated cells was significantly lower among HBA+tEPOR-edited cells compared to unedited and HBA-edited cells. However, when considering the total cell counts, since cell counts were significantly increased among HBA+tEPOR-

edited conditions (Figure S4E), we do not observe a significant decrease in the total number of enucleated cells within this condition. These data indicate that tEPOR is leading to increased proliferation among early erythroid progenitors (indicated by an increased proportion of CD36<sup>+</sup> cells and decreased proportion of DRAQ5<sup>+</sup> enucleated cells), without impeding terminal maturation as the total numbers of enucleated cells are comparable to numbers yielded by unedited and HBA-edited conditions. No discernible differences in cell morphology were observed by cresyl blue staining among the HBA-edited and HBA+tEPOR-edited cells compared to unedited cells cultured with EPO, indicating that the vast majority of cells in all conditions had reached terminal differentiation, becoming either normoblasts, reticulocytes, or enucleated erythrocytes (Figures S5A and S5B).

To determine whether this strategy was able to restore  $\alpha$ -globin production in the context of  $\alpha$ TM, we performed hemoglobin



**Figure 5. RNA sequencing of edited  $\alpha$ TM HSPCs demonstrates restoration of  $\alpha$ -globin gene expression and oxygen transport and iron ion metabolic processes**

(A) Schematic of bulk RNA sequencing workflow. (B) Normalized reads across all samples displayed, with specific genes *HBA1*, *HBA2*, *HBB*, and *EPOR* highlighted. A count of “0.1” was assigned to *HBA1* and *HBA2* for display purposes, although they had 0 detectable reads. (C) Table displaying top five genes with the greatest rank change comparing WT and edited  $\alpha$ TM cells to mock  $\alpha$ TM cells. (D) Normalized read counts for all genes across editing conditions; displayed with fetal and embryonic globin genes (*HBG1*, *HBG2*, *HBE1*, and *HBE2*) highlighted. (E) Table displaying top five genes with the greatest rank change comparing mock and edited  $\alpha$ TM cells to WT. (F) Heatmap of false discovery rate (FDR)  $q$  values across nine commonly enriched GO processes across conditions. Input for GO analysis was the top 0.001% most highly expressed genes within each condition. (G–I) Bubble plots depicting enriched GO processes across treatment comparisons; y axis displays significance by  $-\log_{10}(\text{FDR } q \text{ value})$  and x axis displays enrichment score. Bubble color pertains to GO process and bubble size pertains to number of genes represented in the respective GO process. Input for GO analysis was the top 0.001% of genes showing greatest rank change for each comparison.

tetramer HPLC on edited patient-derived HSPCs following erythroid differentiation. We found that the addition of *tEPOR* to the  $\alpha$ -globin integration cassette dramatically increased the formation of hemoglobin tetramers, with a substantial increase in HbF, HbA, and HbA2 (Figures 4G–4I). Single globin chain HPLC confirmed that  $\alpha$ -globin production was indeed elevated in HDR-edited samples (Figures S3B–S3D), likely driving the observed increases in hemoglobin tetramer formation. We next validated these results using western blot for  $\alpha$ -globin and  $\beta$ -globin and found that not only do cells edited with *HBA+tEPOR* show a dramatic increase in  $\alpha$ -globin compared to unedited and *HBA*-edited patient-derived erythroid cells but also that *HBA+tEPOR*-edited cells show a decrease in the formation of toxic  $\beta$ -globin dimers to nearly WT levels (Figures 4J, 4K, and S3E).

### RNA sequencing demonstrates restoration of $\alpha$ -globin expression in $\alpha$ TM HSPCs

We next used RNA sequencing (RNA-seq) to reveal the transcriptional profile of  $\alpha$ TM patient-derived HSPCs edited

control. Bulk RNA-seq was then performed, followed by quality control and statistical analysis (Figures 5A and S6A–S6E; STAR Methods). Total normalized read counts across samples ranged from 17.6 to 22.0 million and mapped to 17,776 unique genes. The high sequence similarity between *HBA1* and *HBA2* precluded clear differentiation; however, reads were exclusively assigned to one gene or another without overlap. For the *EPOR* gene, the alignment was unable to differentiate between the endogenous *EPOR* and our *tEPOR* transgene, leading to a combined set of reads from both. In analyzing the data, we observed consistent expression across erythroid-associated genes in both healthy control and patient-derived cells, which is expected given the efficient acquisition of erythroid cell surface markers in edited and unedited samples (Figure S6F). While the most highly expressed genes by WT erythroid cells are the adult globins (*HBA1*, *HBA2*, and *HBB*), unedited (mock)  $\alpha$ TM erythroid cells expressed no detectable *HBA1* or *HBA2* (Figure 5B) as expected due to confirmation of four-gene  $\alpha$ -globin deletion in this patient. When investigating the

genes with the greatest rank change between WT erythroid cells compared to  $\alpha$ TM erythroid cells, it is unsurprising that the top genes are *HBA1* and *HBA2* (Figure 5C). Despite the modest increase in  $\alpha$ -globin production by  $\alpha$ -globin integration alone according to HPLC, we observed substantial restoration of *HBA1* and *HBA2* expression compared to mock (*HBA1* and *HBA2* rising from undetectable in  $\alpha$ TM mock to the 99.9<sup>th</sup> and 71.7<sup>th</sup> percentiles of expressed genes by rank order, respectively) (Figure 5B). Comparing *HBA*-edited  $\alpha$ TM cells to mock, we found that *HBA1* is the gene that displays the greatest rank change (Figure 5C). Predictably, this restoration of  $\alpha$ -globin gene expression was further amplified in cells edited with the dual *HBA+tEPOR* vector, with *HBA1* and *HBA2* expression increasing to the 98.9<sup>th</sup> and 99.9<sup>th</sup> percentiles, respectively (Figure 5B). Because this editing strategy places *tEPOR* under the strong, erythroid-specific *HBB* promoter, we also found that *EPOR* expression was elevated compared to all other conditions. These elevated expression levels are confirmed by the fact that *HBA1*, *HBA2*, and *EPOR* are among the top four genes with the greatest rank change over unedited  $\alpha$ TM cells (Figure 5C).

In addition, we found that  $\gamma$ -globin genes (*HBG1* and *HBG2*) were more highly expressed in all  $\alpha$ TM samples, regardless of editing, compared to the WT control (Figure 5D). This was expected as the  $\alpha$ TM HSPCs were derived from an  $\sim$ 1-year-old patient, whereas the WT HSPCs were derived from an adult donor. Although  $\zeta$ -globin (*HBZ*) is normally expressed only during the first 3 months of gestation, we found that this gene was highly elevated in all  $\alpha$ TM cells (among the eight most highly expressed genes) compared to WT cells (Figure 5D).  $\zeta$ -Globin was also the gene with the greatest rank change in expression in all  $\alpha$ TM conditions compared to WT control cells (Figure 5E).

To identify large-scale changes in biological processes, Gene Ontology (GO) enrichment analysis was performed across the top 0.001% (18) most highly expressed genes across all samples. Compared to unedited mock  $\alpha$ TM cells, *HBA*-edited cells displayed an enrichment of various iron ion-associated pathways to levels equivalent to the WT control (Figure 5F; STAR Methods). Further improvements were observed in  $\alpha$ TM cells edited with the dual *HBA+tEPOR* integration strategy, showing enrichment of oxygen and gas transport as well as hydrogen peroxide ( $H_2O_2$ )-associated pathways to levels observed in the WT control (Figure 5F). In addition to assessing GO enrichment on the most highly expressed genes within a given sample, we also performed this analysis on the top 0.001% of genes with the greatest rank change across samples. Unsurprisingly, we found that gas transport and gas metabolism were the most significantly enriched pathways when comparing WT to unedited  $\alpha$ TM cells (Figure 5G). When performing the opposite comparison, for genes with the greatest rank change in unedited  $\alpha$ TM vs. WT cells, we found that the most significantly enriched pathways corresponded to the incorrect/unfolded protein response (Figure 5H), likely representing the cellular response to the formation of toxic  $\beta$ -globin aggregates in the absence of  $\alpha$ -globin. Interestingly, when comparing *HBA+tEPOR*-edited  $\alpha$ TM cells to WT cells, we saw no significantly enriched GO pathways, indicating that this editing strategy yielded cells that were not appreciably different from erythroid cells derived from a healthy donor.

Along with our western blot data (Figure 4J), these findings suggest that while unedited  $\alpha$ TM cells have highly elevated  $\beta$ -globin dimers, even a modest increase in  $\alpha$ -globin is able to mitigate the unfolded protein response that occurs in the context of  $\alpha$ TM pathology. As with WT cells, we found no significantly upregulated GO pathways when comparing *HBA*-edited  $\alpha$ TM cells to unedited  $\alpha$ TM cells. However, when comparing *HBA+tEPOR*-edited cells to unedited  $\alpha$ TM cells, we found a number of upregulated GO pathways, the most significantly enriched being associated with cytokine-mediated signaling (likely a result of elevated *EPOR* signaling) and  $H_2O_2$  catabolism (Figure 5I).

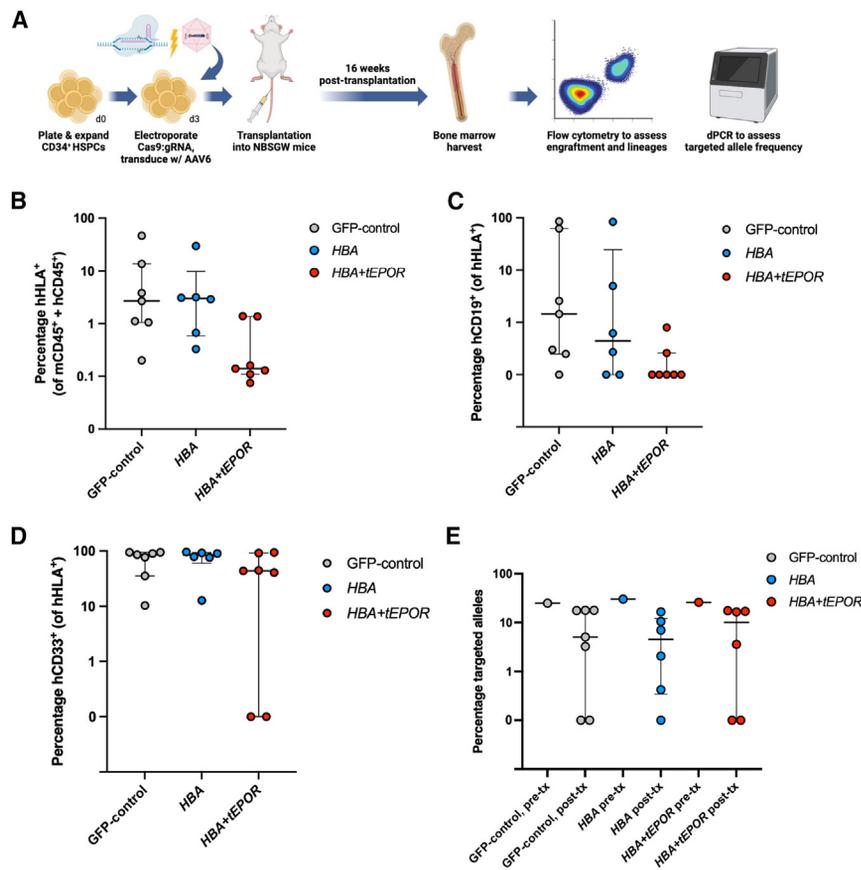
### Transplantation of edited HSPCs into NBSGW mice demonstrates engraftment and multilineage reconstitution

To determine the long-term impact of our editing strategy on the ability of edited HSPCs to engraft and recapitulate various lymphoid and myeloid lineages, we performed xenotransplantation experiments in immunodeficient NBSGW mice. Due to the limited availability of donor bone marrow, we performed editing and transplantation experiments in WT human HSPCs, obtained commercially from a healthy donor. At 16 weeks post-transplantation, we harvested bone marrow and determined engraftment and edited allele frequencies by flow cytometry and ddPCR, respectively (Figures 6A and S2C). We found that WT HSPCs edited with our clinical vectors were capable of engraftment and multilineage reconstitution in  $CD19^+$  lymphoid and  $CD33^+$  myeloid subpopulations (Figures 6B–6D). While all treatment groups were able to engraft into the bone marrow, we observed lower engraftment among *HBA+tEPOR*-edited HSPCs, although this trend did not reach statistical significance. However, this could indicate a limitation of our approach that requires further investigation prior to clinical translation. We next used ddPCR to quantify the percentage of edited alleles within our bulk population of edited HSPCs, which was consistent across all editing conditions pre-transplant (24.8% GFP-control, 30.2% *HBA*, and 25.9% *HBA+tEPOR*) and in engrafted HSPCs post-transplant (8.8%, 6.1%, and 9.1%, respectively) (Figure 6E).

## DISCUSSION

Given the increasing global burden of  $\alpha$ TM and the tremendous cost of disease management for patients and health care systems, there is an urgent need for curative treatments for this disease. To our knowledge, these results demonstrate for the first time that gene therapy or genome editing may be used to restore hemoglobin production to  $\alpha$ TM patient-derived erythroid cells. This approach addresses the two major elements contributing to molecular pathology of the disease in a single genome editing event: successfully increasing  $\alpha$ -globin production and reducing the formation of toxic  $\beta$ -globin dimers without noticeably disrupting  $\beta$ -globin production. While the cells in this work were derived from patients with four-gene deletion  $\alpha$ TM, our editing strategy may also be effective in other transfusion-dependent phenotypes of  $\alpha$ -thalassemia, such as in patients with inactivation of three  $\alpha$ -globin genes or hemoglobin H disease.<sup>24</sup>

In addition to these protein-level changes, deeper analysis reveals several interesting findings at a transcriptional level. The



**Figure 6. Engraftment of edited  $\alpha$ TM HSPCs into NBSGW mice**

(A) Schematic of editing, xenotransplantation, and mouse harvest workflow.  $n = 7$  for GFP-control group,  $n = 6$  for HBA-edited group, and  $n = 7$  for HBA+tEPOR-edited group.

(B) The percentages of hHLA<sup>+</sup> cells that were mCD45<sup>+</sup> and hCD45<sup>+</sup> among mTer119<sup>-</sup> cells. Bars represent median  $\pm$  interquartile range.

(C) The percentages of hCD19<sup>+</sup> cells of hHLA<sup>+</sup> cells. Bars represent median  $\pm$  interquartile range.

(D) The percentages of hCD33<sup>+</sup> cells of hHLA<sup>+</sup> cells. Bars represent median  $\pm$  interquartile range.

(E) Edited allele frequency determined by ddPCR among engrafted HSPCs. Bars represent median  $\pm$  interquartile range.

reside within the *HBB* introns.<sup>26</sup> Consequently, integration of *HBA1* with its introns at the *HBB* locus may not support high levels of transgene expression.

As a solution, we discovered that pairing the  $\alpha$ -globin transgene with a naturally occurring *EPOR* variant substantially increased erythropoietic output of  $\alpha$ -globin-expressing  $\alpha$ TM patient-derived cells. While we observed significantly higher production of HbA tetramers from our HBA+tEPOR edited conditions, it is clear from our HPLC analysis that HbF-to-HbA switching is still underway. Because we are knocking

enrichment of pathways associated with various oxidative pathways in HBA+tEPOR-edited cells suggests an improvement in the stability and function of edited erythroid cells. Most notably, H<sub>2</sub>O<sub>2</sub> catabolism and metabolism has been identified as a critical process in preventing cellular injury in  $\beta$ -thalassemic erythrocytes,<sup>25</sup> as the auto-oxidation of unpaired globin chains results in the generation and release of significant amounts of superoxide (O<sub>2</sub><sup>-</sup>) and H<sub>2</sub>O<sub>2</sub>. These observations suggest that our genome editing strategy not only restores  $\alpha$ -globin gene expression but also re-establishes essential antioxidant pathways that may be critical in addressing the pathophysiology of  $\alpha$ -thalassemia.

Interestingly, while our genome editing approach using the HBA vector alone achieved an editing frequency of  $\sim$ 25% of alleles, it did not yield high levels of HbF or HbA production. Given this editing rate, it is unlikely that a high frequency of bi-allelic integration is occurring to account for the low amounts of hemoglobin production. This is because high frequencies of bi-allelic editing would be expected to yield a substantial HbF peak (as restored  $\alpha$ -globin would pair with significant levels of  $\gamma$ -globin from this 1-year-old patient), which we do not observe in our HPLC data. Still, without comprehensive genotyping analysis, we cannot rule out the possibility that some bi-allelic editing is occurring. Nonetheless, we believe that the relatively low levels of HbF and HbA are more likely attributed to the complex regulatory landscape of the globin loci as crucial enhancer regions

our transgene cassette into the *HBB* locus that is only expressed later in development, we expect the results to be further amplified as hemoglobin switching is completed. This method has the potential to resolve clinical challenges previously encountered in HSCT for the hemoglobinopathies by lowering editing and engraftment frequencies needed to correct RBC disorders. In doing so, this strategy has the capacity to reduce or eliminate the need for high-morbidity myeloablation currently required to clear the bone marrow niche, which stands as a major barrier to safe and effective HSCT. Moreover, as patients with severe anemia, such as those with  $\alpha$ TM, have increased levels of circulating erythropoietin,<sup>27</sup> a tEPOR-based genome editing strategy may be especially potent in this context.

Finally, we believe this work stands as a compelling proof of concept. However, before this editing strategy is ready for translation into patients, further preclinical studies are needed. This may be difficult due to the limited availability of  $\alpha$ TM patient HSPCs, given the historical prenatal lethality of the disorder as well as the inability of existing mouse models to effectively model human erythropoiesis *in vivo*. To circumvent this, we performed transplantation experiments of edited WT HSPCs into NBSGW mice to evaluate the impact of our editing strategy on long-term engraftment and multilineage cell reconstitution of edited cells. While cells from all editing conditions persisted, we observed lower engraftment in HBA+tEPOR-edited cells, although this trend did not reach statistical significance. This

reduced engraftment may be attributed to elements of the editing process itself, such as variations in the purity or titer of separately prepared AAV vectors. Future clinical adoption of this editing approach will depend on expanded animal model studies and further investigation to elucidate the precise mechanisms impacting the engraftment of edited cells.

In summary, we present the results of a comprehensive genome editing approach for the treatment of  $\alpha$ -thalassemia that restores  $\alpha$ -globin production, reduces the formation of toxic unpaired  $\beta$ -globin aggregates, and restores the transcriptional profile akin to that of healthy erythroid cells. Thus, we believe these findings support the development of a definitive *ex vivo* autologous genome editing strategy that may be curative for patients with  $\alpha$ -thalassemia.

### Limitations of the study

This work defines a potentially curative universal genome editing strategy for the treatment of  $\alpha$ TM, regardless of the deletions or mutations responsible for the disorder. However, several limitations exist that would need to be addressed before this strategy is ready for clinical translation. For instance, we observed differences in erythroid development in our editing condition that incorporated a *tEPOR* transgene. While this is a naturally occurring variant, expression from a strong globin promoter departs native regulation and could lead to an abnormal erythroid phenotype. We also observed lower engraftment frequencies among HSPCs edited with the *HBA+tEPOR* cassette, which could be due to negative effects of Cas9-mediated DNA double-strand breaks, transduction with AAV vectors, and/or *tEPOR* expression. Future experiments will need to more deeply characterize function of *HBA+tEPOR*-edited HSPCs, and better understand their behavior *in vivo*. More broadly, there are concerns about the time- and cost-intensive nature of *ex vivo* genome editing. *In vivo* genome editing emerges as a potentially scalable and accessible solution, however, editing HSPCs *in vivo* at frequencies required to correct the hemoglobinopathies remains challenging. Future work is needed to improve delivery and expression of genome editing machinery to HSPCs *in vivo*, which may one day enable adaptation of potentially curative *ex vivo* editing strategies to edit HSPCs directly where they reside in the bone marrow.

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Kyle Cromer ([kyle.cromer@ucsf.edu](mailto:kyle.cromer@ucsf.edu)).

#### Materials availability

Plasmids used to generate DNA repair templates packaged in AAV will be deposited with Addgene. No other unique reagents were generated during the course of this work.

#### Data and code availability

- Data: RNA-seq data have been deposited in the NCBI Sequence Read Archive database (accession no. PRJNA1193132) and are publicly available as of the date of publication.
- Code: no novel code was generated during the course of this work.
- Other items: any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

### ACKNOWLEDGMENTS

The authors thank the following funding sources that made this work possible: S.N.C. was supported by the University of California, San Francisco NIH T32 Research Training in Transplant Surgery Fellowship; T.M. was supported by the California Institute for Regenerative Medicine Bridges Program; B.J.L. was supported by the National Science Foundation Graduate Research Fellowship Program; B.R.C. was supported by the National Institutes of Health (R01-AG072052, R01-HL130533, R01-HL13535801, and P01-HL146366), the California Institute for Regenerative Medicine (INFR6.2-15527), the Charcot-Marie-Tooth Association, and support through a gift from the Roddenberry Foundation and Pauline and Thomas Tusher; T.C.M. was supported by the UCSF Center for Maternal-Fetal Precision Medicine; M.K.C. was supported by the American Society of Gene & Cell Therapy Career Development Award; and T.C.M., M.H.P., and M.K.C. were supported by the National Institutes of Health (R01-HL161291). We thank Emma Canepa and Billie Lianoglou for their assistance with patient sample collection and our patients for their gracious participation in the research.

### AUTHOR CONTRIBUTIONS

M.H.P. and M.K.C. supervised the project. S.N.C., D.S., R.S., B.H., H.S., B.W., B.R.C., T.C.M., V.A.S., M.H.P., and M.K.C. designed the experiments. S.N.C., E.S., D.S., R.S., T.M., B.H., H.S., B.W., C.B., B.J.L., K.J., K.A.R., Z.K., Y.Z., T.T., M.C., S.H., J.P.H., and M.K.C. carried out the experiments. S.N.C., E.S., D.S., R.S., B.W., H.Y., J.M.G., M.H.P., and M.K.C. analyzed the data. S.N.C. and M.K.C. wrote the manuscript.

### DECLARATION OF INTERESTS

B.R.C. is a founder and holds equity in Tenaya Therapeutics. T.C.M. is on the scientific advisory board of Acrigen and receives grant funding from Novartis, BioMarin, and Biogen. M.H.P. is a member of the scientific advisory board of Allogene Therapeutics, is on the Board of Directors of Graphite Bio, and has equity in CRISPR Therapeutics. B.W., T.C.M., M.H.P., and M.K.C. hold patents related to this work (WO/2023/060059 and WO/2021/022189).

### DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT-4 by OpenAI and Claude Haiku by Anthropic PBC solely to improve language and readability. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

### STAR★METHODS

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

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
  - Healthy donor human primary HSPCs
  - $\alpha$ TM donor human primary HSPCs
  - NBSGW mice
- [METHOD DETAILS](#)
  - *In vitro* culture of primary HSPCs
  - Genome editing of HSPCs
  - *In vitro* differentiation of HSPCs into erythrocytes
  - HUDEP-2 cell culture and nucleofection
  - Indel frequency analysis by ICE
  - Genome editing analysis by flow cytometry
  - Editing analysis by ddPCR
  - AAV6 vector design, production, and purification
  - Immunophenotyping of differentiated erythroid cells
  - Hemoglobin tetramer analysis
  - Single globin chain analysis

- Western blot to quantify globin protein
- RNA-sequencing of edited erythroid cells
- Gene ontology enrichment analysis
- Mouse conditioning and transplant
- Bone marrow isolation and flow

● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2024.115141>.

Received: January 11, 2024

Revised: November 21, 2024

Accepted: December 11, 2024

REFERENCES

1. Vichinsky, E. (2010). Complexity of alpha thalassemia: growing health problem with new approaches to screening, diagnosis, and therapy. *Ann. N. Y. Acad. Sci.* *1202*, 180–187.
2. Piel, F.B., and Weatherall, D.J. (2014). The  $\alpha$ -thalassemias. *N. Engl. J. Med.* *371*, 1908–1916.
3. Kreger, E.M., Singer, S.T., Witt, R.G., Sweeters, N., Lianoglou, B., Lal, A., Mackenzie, T.C., and Vichinsky, E. (2016). Favorable outcomes after in utero transfusion in fetuses with alpha thalassemia major: a case series and review of the literature. *Prenat. Diagn.* *36*, 1242–1249.
4. Schwab, M.E., Lianoglou, B.R., Gano, D., Gonzalez Velez, J., Allen, I.E., Arvon, R., Baschat, A., Bianchi, D.W., Bitanga, M., Bourguignon, A., et al. (2023). The impact of in utero transfusions on perinatal outcomes in patients with alpha thalassemia major: The UCSF registry. *Blood Adv.* *7*, 269–279.
5. Schrier, S.L., Centis, F., Verneris, M., Ma, L., and Angelucci, E. (2003). The role of oxidant injury in the pathophysiology of human thalassemias. *Redox Rep.* *8*, 241–245.
6. Baird, D.C., Batten, S.H., and Sparks, S.K. (2022). Alpha- and beta-thalassemia: Rapid evidence review. *Am. Fam. Physician* *105*, 272–280.
7. Fleischhauer, K., Locatelli, F., Zecca, M., Orofino, M.G., Giardini, C., De Stefano, P., Pession, A., Iannone, A.M., Carcassi, C., Zino, E., and La Nasa, G. (2006). Graft rejection after unrelated donor hematopoietic stem cell transplantation for thalassemia is associated with nonpermissive HLA-DPB1 disparity in host-versus-graft direction. *Blood* *107*, 2984–2992.
8. Cromer, M.K., Camarena, J., Martin, R.M., Lesch, B.J., Vakulskas, C.A., Bode, N.M., Kurgan, G., Collingwood, M.A., Rettig, G.R., Behlke, M.A., et al. (2021). Gene replacement of  $\alpha$ -globin with  $\beta$ -globin restores hemoglobin balance in  $\beta$ -thalassemia-derived hematopoietic stem and progenitor cells. *Nat. Med.* *27*, 677–687.
9. Thompson, A.A., Walters, M.C., Kwiatkowski, J., Rasko, J.E.J., Ribeil, J.A., Hongeng, S., Magrin, E., Schiller, G.J., Payen, E., Semeraro, M., et al. (2018). Gene therapy in patients with transfusion-dependent  $\beta$ -thalassemia. *N. Engl. J. Med.* *378*, 1479–1493.
10. Frangoul, H., Altshuler, D., Cappellini, M.D., Chen, Y.S., Domm, J., Eustace, B.K., Foell, J., de la Fuente, J., Grupp, S., Handgretinger, R., et al. (2021). CRISPR-Cas9 gene editing for sickle cell disease and  $\beta$ -thalassemia. *N. Engl. J. Med.* *384*, 252–260.
11. Luna, S.E., Camarena, J., Hampton, J.P., Majeti, K.R., Charlesworth, C.T., Soupene, E., Selvaraj, S., Jia, K., Sheehan, V.A., Cromer, M.K., and Porteus, M.H. (2024). Enhancement of erythropoietic output by Cas9-mediated insertion of a natural variant in hematopoietic stem and progenitor cells. *Nat. Biomed. Eng.* *8(12)*:1540–1552. <https://doi.org/10.1038/s41551-024-01222-6>.
12. Chakraborty, N., Bilgrami, S., Maness, L., Guo, C., Perez-Diez, A., Mukherji, B., and Tutschka, P. (1999). Myeloablative chemotherapy with autologous peripheral blood stem cell transplantation for metastatic breast cancer: Immunologic consequences affecting clinical outcome. *Bone Marrow Transplant.* *24*, 837–843.
13. Bhatia, S. (2011). Long-term health impacts of hematopoietic stem cell transplantation inform recommendations for follow-up. *Expert Rev. Hematol.* *4*, 437–454.
14. Inamoto, Y., Shah, N.N., Savani, B.N., Shaw, B.E., Abraham, A.A., Ahmed, I.A., Akpek, G., Atsuta, Y., Baker, K.S., Basak, G.W., et al. (2015). Secondary solid cancer screening following hematopoietic cell transplantation. *Bone Marrow Transplant.* *50*, 1013–1023.
15. 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  $\beta$ -globin gene targeting in human hematopoietic stem cells. *Nature* *539*, 384–389.
16. Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroshima, T., Miyoshi, H., Tani, K., and Nakamura, Y. (2013). Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS One* *8*, e59890.
17. Conant, D., Hsiao, T., Rossi, N., Oki, J., Maures, T., Waite, K., Yang, J., Joshi, S., Kelso, R., Holden, K., et al. (2022). Inference of CRISPR edits from Sanger trace data. *CRISPR J.* *5*, 123–130.
18. Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* *42*, e168.
19. 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.
20. Cromer, M.K., Majeti, K.R., Rettig, G.R., Murugan, K., Kurgan, G.L., Bode, N.M., Hampton, J.P., Vakulskas, C.A., Behlke, M.A., and Porteus, M.H. (2023). Comparative analysis of CRISPR off-target discovery tools following ex vivo editing of CD34<sup>+</sup> hematopoietic stem and progenitor cells. *Mol. Ther.* *31*, 1074–1087.
21. 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.
22. Dulmovits, B.M., Appiah-Kubi, A.O., Papoin, J., Hale, J., He, M., Al-Abed, Y., Didier, S., Gould, M., Husain-Krautter, S., Singh, S.A., et al. (2016). Pomalidomide reverses  $\gamma$ -globin silencing through the transcriptional reprogramming of adult hematopoietic progenitors. *Blood* *127*, 1481–1492.
23. de la Chapelle, A., Träskelin, A.L., and Juvonen, E. (1993). Truncated erythropoietin receptor causes dominantly inherited benign human erythrocytosis. *Proc. Natl. Acad. Sci. USA* *90*, 4495–4499.
24. Vichinsky, E.P. (2013). Clinical manifestations of  $\alpha$ -thalassemia. *Cold Spring Harb. Perspect. Med.* *3*, a011742.
25. Scott, M.D. (2006). H<sub>2</sub>O<sub>2</sub> injury in beta thalassemic erythrocytes: protective role of catalase and the prooxidant effects of GSH. *Free Radic. Biol. Med.* *40*, 1264–1272.
26. Bartman, C.R., Hsu, S.C., Hsiung, C.C.S., Raj, A., and Blobel, G.A. (2016). Enhanced regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Mol. Cell* *62*, 237–247.
27. Nişli, G., Kavakli, K., Aydinok, Y., Oztop, S., and Cetingül, N. (1997). Serum erythropoietin levels in patients with beta thalassemia major and intermedia. *Pediatr. Hematol. Oncol.* *14*, 161–167.
28. Selvaraj, S., Feist, W.N., Viel, S., Vaidyanathan, S., Dudek, A.M., Gastou, M., Rockwood, S.J., Ekman, F.K., Oseghale, A.R., Xu, L., et al. (2024). High-efficiency transgene integration by homology-directed repair in human primary cells using DNA-PKcs inhibition. *Nat. Biotechnol.* *42*, 731–744.
29. 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.

30. Khan, I.F., Hirata, R.K., and Russell, D.W. (2011). AAV-mediated gene targeting methods for human cells. *Nat. Protoc.* 6, 482–501.
31. 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 adenovirus-associated virus serotype 2-derived inverted terminal repeat sequences. *Hum. Gene Ther. Methods* 23, 18–28.
32. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419.
33. Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinf.* 10, 48.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Human CD45 V450	BD Biosciences	Clone HI30; RRID: AB_1645573
Human CD34 APC	BioLegend	Clone 561; RRID: 343607
Human CD71 PE-Cy7	Invitrogen	Clone OKT9; RRID: AB_2573366
Human CD235a (GPA) PE	BD Biosciences	Clone GA-R2; RRID: AB_395949
Human CD36 APC	BioLegend	Clone 5-271; RRID: AB_1279224
Mouse anti-human $\alpha$ -globin	Santa Cruz Biotechnology	Cat. sc-514378; RRID: AB_2716828
Mouse anti-human $\beta$ -globin	Santa Cruz Biotechnology	Cat. sc-21757; RRID: AB_627713
Mouse anti-IgG <sub>1</sub>	Santa Cruz Biotechnology	Cat. sc-525408; RRID: N/A
TruStain FcX Plus (anti-mouse CD16/32)	BioLegend	Clone S17011E; RRID: AB_2783137
Mouse Cd45.1 PE-Cy7	BioLegend	Clone A20; RRID: AB_1134170
Mouse TER-119 BUV395	BD Biosciences	Cat. 563827; RRID: AB_2738438
Human HLA-ABC APC	BD Biosciences	Cat. 555555; RRID: AB_398603
Human CD19 PE	BD Biosciences	Cat. 555413; RRID: AB_395813
Human CD33 BV421	BD Biosciences	Cat. 562854; RRID: AB_2737405
Human CD45 PerCP	BioLegend	Cat. 368505; RRID: AB_400307
Human CD235ab (GPA) Pacific Blue	BioLegend	Cat. 306612; RRID: AB_2248153
<b>Bacterial and virus strains</b>		
Stable Competent E. coli	New England Biolabs	C30401
Adeno-associated virus, serotype 6	Gift from David Russell	N/A
<b>Biological samples</b>		
Heat-inactivated human serum	Sigma-Aldrich	H3667
Human plasma	Innovative Research	IPLAWBCATNAC
<b>Chemicals, peptides, and recombinant proteins</b>		
Human SCF	PeptoTech	300-07
Human TPO	PeptoTech	300-18
Human FLT3L	PeptoTech	300-19
Human IL6	PeptoTech	200-06
Human IL3	PeptoTech	200-03
Human EPO	PeptoTech	100-64
Human insulin	Sigma-Aldrich	91077C
Human holo-transferrin	Sigma-Aldrich	T0665
Heparin sodium salt	Sigma-Aldrich	H3393
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	67-68-5
Phosphate buffered saline (PBS)	Thermo-Fisher Scientific	10010001
Doxycycline hyclate	Sigma-Aldrich	PHR1145
Penicillin/streptomycin	Thermo-Fisher Scientific	15140148
L-glutamine	Thermo-Fisher Scientific	35050061
Polyethylenimine (PEI)	Polysciences	23966
KCN	Sigma-Aldrich	207810
NaCl	Sigma-Aldrich	433209
Bis-tris	Sigma-Aldrich	B9754
NaOH	Sigma-Aldrich	655104
Trifluoroacetic acid	Sigma-Aldrich	74564
Busulfan	Sigma-Aldrich	B2635

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Ethylenediaminetetraacetic acid (EDTA) disodium salt dihydrate	Sigma-Aldrich	E4884
UM171	MedChemExpress	HY-12878
HSPC base media	STEMCELL Technologies	09655
Cas9 protein	Aldevron, Integrated DNA Technologies	9214, 1081059
AZD-7648	MedChemExpress	HY-111783
<b>Deposited data</b>		
RNA-sequencing data	NCBI Sequence Read Archive	PRJNA1193132
<b>Experimental models: Cell lines</b>		
HUDEP-2 cell line	RIKEN Institute	RCB4557
293T cell line	ATCC	CRL-3216
<b>Experimental models: Organisms/strains</b>		
Human CD34 <sup>+</sup> -enriched HSPCs derived from Plerixafor and/or G-CSF-mobilized peripheral blood from healthy donors	STEMCELL Technologies, AllCells, Fred Hutchinson Cancer Center Hematology Core	70073.2
Human CD34 <sup>+</sup> -enriched HSPCs derived from bone marrow aspirates from $\alpha$ TM donors	IRB Protocol	16-21157
NBSGW mice	The Jackson Laboratory	026622
<b>Oligonucleotides</b>		
HBB gRNA sg1-22	Synthego	Synthetic sgRNA
ddPCR primers & probes	Integrated DNA Technologies	Custom probe-based qPCR assays
<b>Recombinant DNA</b>		
AAV6 DNA repair templates	Cloned in-house & purified using Takara Bio AAVpro Purification Kit	6675
<b>Software and algorithms</b>		
HPLC analysis software	PerkinElmer	TotalChrom
Data visualization and statistical analysis software	GraphPad	Prism v.9
Western blot analysis software	Bio-Rad	Quantity One 1-D
RNA-sequencing quality control software	Babraham Bioinformatics	FastQC
RNA-sequencing analysis software	Open source	Salmon
Flow cytometry visualization and analysis software	FlowJo	v10.10.0
DNA sequence visualization software	SnapGene	V8.0
<b>Other</b>		
Human CD34 <sup>+</sup> selection kit	STEMCELL Technologies	17856
Electroporation kit	Lonza Bioscience	V4XP-3032
Transfection kit	Invitrogen	MPK10096
DNA extraction solution	LGC Biosearch Technologies	SS000035-D2
Gibson assembly mix	New England Biolabs	E2611L
Deposited data	NCBI Sequence Read Archive	PRJNA1193132

**EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**

**Healthy donor human primary HSPCs**

HSPCs were isolated from healthy human donors either from umbilical cord blood (provided by Stanford Binns Program) or Plerixafor- and/or G-CSF-mobilized peripheral blood (AllCells, STEMCELL Technologies, and Fred Hutchinson Cancer Center Hematology Core). CD34<sup>+</sup> HSPCs were isolated from umbilical cord blood using a EasySep Human CD34 Positive Selection Kit II according to the manufacturer's protocol; mobilized peripheral blood HSPCs were purchased as CD34-enriched aliquots. All patients provided

informed consent and samples were de-identified after collection. HSPCs derived from both male and female donors were included in this study with no exclusion criteria applied regarding ethnicity.

### **$\alpha$ TM donor human primary HSPCs**

Bone marrow aspirates were taken from patients with  $\alpha$ TM (Table S2) under protocol no. 16–21157, which was approved by the NHLBI Institutional Review Board. CD34<sup>+</sup> HSPCs were isolated using a EasySep Human CD34 Positive Selection Kit II according to the manufacturer's protocol. Patients provided informed consent and samples were de-identified after collection. HSPCs derived from both male and female donors were included in this study with no exclusion criteria applied regarding ethnicity.

### **NBSGW mice**

Immuno-deficient NBSGW mice (strain # 026622, The Jackson Laboratory) were used for xenotransplantation experiments. Only female mice were used in order to determine differences in engraftment frequencies among editing conditions and to limit influences of sex in these experiments.

## **METHOD DETAILS**

### ***In vitro* culture of primary HSPCs**

CD34<sup>+</sup> HSPCs were cultured at 10<sup>5</sup> cells/mL in StemSpan SFEM II Medium supplemented with 100 ng/mL SCF, 100 ng/mL TPO, 100 ng/mL FLT3L, 100 ng/mL IL6, 20 mg/mL streptomycin, and 20 U/mL penicillin/streptomycin; 35nM UM171 was included in HSPC culture media during *ex vivo* culture of cells prior to mouse transplantation experiments. Incubator conditions were 37°C, 5% CO<sub>2</sub>, and 5% O<sub>2</sub>.

### **Genome editing of HSPCs**

Cas9 protein was purchased from Integrated DNA Technologies or Aldevron. Preceding electroporation, ribonucleoproteins (RNPs) were complexed with guide RNAs (gRNAs) at a Cas9:gRNA molar ratio of 1:2.5 (6  $\mu$ g Cas9:3.2  $\mu$ g gRNA per electroporation) at 25°C for 10–20mins. HSPCs were then resuspended in Lonza P3 buffer with Cas9 RNPs and subsequently electroporated using a Lonza 4D Nucleofector (program DZ-100). Electroporated cells were then plated at 10<sup>5</sup> cells/mL in HSPC media and AAV6 was added at 5  $\times$  10<sup>3</sup> vector genomes/cell based on titers determined by droplet digital PCR (ddPCR). The small molecule AZD-7648 was also added to cells to improve HDR frequencies as previously reported.<sup>28</sup>

### ***In vitro* differentiation of HSPCs into erythrocytes**

2–3d post-editing, HSPCs were transitioned into StemSpan SFEM II Medium supplemented with 100 U/mL penicillin–streptomycin, 10 ng/mL SCF, 1 ng/mL IL3, 3U/mL EPO, 200  $\mu$ g/mL transferrin, 3% antibody serum, 2% human plasma, 10  $\mu$ g/mL insulin, and 3 U/mL heparin and maintained at 10<sup>5</sup> cells/mL. At d7 of culture, cells were transitioned to the above media without IL3 and maintained at 10<sup>5</sup> cells/mL. At d11, cells were transitioned to the above media without IL3 and with transferrin increased to 1 mg/mL and maintained at 10<sup>6</sup> cells/mL.

### **HUDEP-2 cell culture and nucleofection**

HUDEP-2 cells were grown in StemSpan SFEM Medium supplemented with 50 ng/mL SCF, 3U/mL EPO, 10<sup>–6</sup>M DMSO, 1  $\mu$ g/mL doxycycline, 1  $\times$  penicillin/streptomycin, and L-glutamine. Cells were maintained at a concentration of 5  $\times$  10<sup>5</sup> cells/mL. Cells were transfected with 100pmol of Cas9 RNPs by nucleofection using a Neon Transfection System. 10<sup>5</sup> cells were resuspended in nucleofection buffer T (Neon Transfection Kit) and given one pulse of 1100V for 20ms. Cells were harvested 72h after nucleofection for downstream experiments. Synthetic chemically modified gRNAs used to edit HSPCs were purchased from Synthego and purified by HPLC. Modifications include 2'-O-methyl-3'-phosphorothioate at the three terminal nucleotides of the 5' and 3' ends described previously.<sup>29</sup>

### **Indel frequency analysis by ICE**

HUDEP-2 cells were washed with PBS and DNA extracted using QuickExtract DNA extraction solution (LGC Biosearch Technologies) following the manufacturer's instructions. To quantify genome editing outcomes, a PCR was performed across the *HBB* gene spanning the gRNA cut sites. Amplicons were Sanger sequenced and indels quantified using ICE software.<sup>17</sup>

### **Genome editing analysis by flow cytometry**

5d post-editing with fluorescent reporter vectors (Figures 1D–1G), CD34<sup>+</sup> HSPCs were harvested and the percentage of edited cells was determined by flow cytometry. This analysis was performed using either the Accuri C6 flow cytometer and software (v.9.4.11) or the FACS Aria II cytometer and FACS Diva software (v.8.0.3). The data were subsequently analyzed using FlowJo (v.10.6.1).

### **Editing analysis by ddPCR**

Genome-edited HSPCs were harvested within 2d–4d post-electroporation and analyzed for modification frequencies of the alleles of interest. Our custom ddPCR analysis utilized a two-probe set where one probe binds the WT sequence (WT) and another binds

unedited reference (REF) sequence outside of DNA repair template homology arms but residing on the same BamHI-digested gDNA fragment. Quantification of WT alleles, rather than HDR alleles, was necessary due to the large amplicons that would be required for in-out PCR given the long (~900bp) homology arms of our DNA repair templates. To calculate allele frequencies, we divided the WT copies/ $\mu$ L by the REF copies/ $\mu$ L to yield frequency of unedited WT alleles. The inverse of our WT allele frequencies yielded our HDR editing frequencies. The following primers and probes were used in the ddPCR reaction.

#### **HBB WT unedited alleles**

Forward primer: 5'-ACGTGGATGAAGTTGGTGGT-3'

Reverse primer: 5'-CAGCATCAGGAGTGGACAGA-3'

Probe: 5'-6-FAM/AGAAACTGGGCATGTGGAGA/3IABkFQ-3'

#### **HBB REF alleles**

Forward primer: 5'-CGATCACGTTGGGAAGCTAT-3'

Reverse primer: 5'-GAGTGTCTGGGGGAACAAAA-3'

Probe: 5'-6-FAM/CTTTTCCCTCCTACCCCTA/3IABkFQ-3'

#### **AAV6 vector design, production, and purification**

Adeno-associated virus, serotype 6 (AAV6) vector plasmids were cloned into the pAAV-MCS plasmid, comprised of inverted terminal repeats (ITRs) derived from AAV2. Gibson Assembly Mastermix (New England Biolabs, Ipswich, MA) was used for the creation of all DNA repair vectors according to manufacturer's instructions. Once cloned, plasmids were transformed into Stable Competent *E. coli* (New England Biolabs) and amplified using Plasmid Plus Midi Kits (QIAGEN N.V., Hilden, Germany). AAV6 vector was produced with little variation as previously described.<sup>30</sup> 293T cells (ATCC, Manassas, VA) were seeded in five 15cm<sup>2</sup> dishes with 1.3–1.5 × 10<sup>7</sup> cells per plate 24h pre-transfection. Each dish was then transfected with 112 $\mu$ g polyethylenimine, 6 $\mu$ g ITR-containing plasmid, and 22 $\mu$ g pDGM6 (gift from David Russell, University of Washington), which expresses AAV6 cap, AAV2 rep, and Ad5 helper genes. Following a 48–72h incubation, cells were harvested, and vectors were purified using the Takara Bio AAVpro purification kit as per manufacturer's instructions and stored at –80°C until further use. AAV6 vectors were titered using a Bio-Rad QX200 ddPCR machine and QuantaSoft software (v.1.7) to measure the number of vector genomes, as described previously.<sup>31</sup>

The length in base pairs of the exogenous fragments inserted into the AAV vector (not including homology arms) are:

*HBB* UTRs (with 2A YFP): 1,598bp.

*HBA1* UTRs (with 2A-YFP): 1,643bp.

*HBA2* UTRs (with 2A-YFP): 1,635bp.

*HBB* UTRs (without 2A-YFP): 827bp.

*HBA1* UTRs (without 2A-YFP; "HBA"): 872bp

*HBA+tEPOR*: 2,776bp.

#### **Immunophenotyping of differentiated erythroid cells**

Differentiated erythroid cells were analyzed by flow cytometry for lineage-specific markers using a FACS Aria II. Edited and unedited cells were analyzed using the following antibodies: hCD45 V450 (HI30; BD Biosciences), CD34 APC (561; BioLegend), CD71 PE-Cy7 (OKT9; Affymetrix), and CD235a PE (GPA) (GA-R2; BD Biosciences). To measure viability, cells were also stained with Ghost Dye Red 780 (Tonbo Biosciences).

#### **Hemoglobin tetramer analysis**

Frozen pellets of approximately 10<sup>6</sup> cells *in vitro*-differentiated erythrocytes were thawed and lysed in 30 $\mu$ L of RIPA buffer with 1x Halt Protease Inhibitor Cocktail for 5mins on ice. The mixture was vigorously vortexed and cell debris were removed by centrifugation at 13,000 RPM for 10mins at 4°C. HPLC analysis of hemoglobins in their native form was performed on a cation-exchange PolyCAT A column (35 × 4.6mm<sup>2</sup>, 3 $\mu$ m, 1,500Å; PolyLC Inc.) using a PerkinElmer Flexar HPLC system at room temperature with detection at 415nm. Mobile phase A consisted of 20mM Bis-tris and 2mM KCN at pH 6.94, adjusted with HCl. Mobile phase B consisted of 20mM Bis-tris, 2mM KCN, and 200mM NaCl at pH 6.55. Hemolysate was diluted in buffer A prior to injection of 20 $\mu$ L onto the column with 8% buffer B and eluted at a flow rate of 2 mL/min with a gradient made to 40% B in 6mins, increased to 100% B in 1.5mins, and returned to 8% B in 1min, and re-equilibrated for 3.5mins. Quantification of the area under the curve of the peaks were performed with TotalChrom software and raw values were exported to GraphPad Prism (v.9) for plotting and further analysis.

#### **Single globin chain analysis**

Frozen pellets of approximately 10<sup>6</sup> cells *in vitro*-differentiated erythrocytes were thaw and lysed in 30 $\mu$ L of RIPA buffer with 1x Halt Protease Inhibitor Cocktail for 5mins on ice. The mixture was vigorously vortexed and cell debris were removed by centrifugation at 13,000 RPM for 10 mins at 4°C. Analysis of globin chains was performed by reverse-phase PerkinElmer Flexar HPLC system with a Vydac 214TP C4 column (250 × 4.6mm<sup>2</sup>, 5 $\mu$ m, 300Å; Avantor, Inc.) at 20°C and detection at 280nm. Mobile phase A consisted of 10% methanol made in acetonitrile and mobile B of 0.5% trifluoroacetic acid in water adjusted at pH 2.9 with NaOH. Hemolysate was diluted in B prior injection at a flow rate of 1 mL/min in 49% A for 2mins, followed by 2mins in 50% A, and a 20min gradient

to 60% A. The column was then equilibrated to 49% A for 6mins. Quantification of the area under the curve of the peaks was performed with TotalChrom software and raw values were exported to GraphPad Prism (v.9) for plotting and further analysis.

### Western blot to quantify globin protein

Samples were subjected to electrophoresis on 14% SDS-polyacrylamide gel electrophoresis under denaturing conditions. A WT umbilical cord blood-derived sample was included as a positive control. Following electrophoresis, proteins were transferred to polyvinylidene difluoride membranes and blocked for 1h at room temperature using TBS-Tween 20 (0.1%) containing 5% non-fat dry milk to prevent non-specific binding. The blocked membranes were then incubated with primary antibodies diluted in TBS-Tween 20 (0.1%) containing 1% non-fat dry milk. The antibodies used were mouse anti-human  $\alpha$ -globin (sc-514378) and  $\beta$ -globin (sc-21757) (Santa Cruz Biotechnology). All primary antibodies were diluted at 1:1,000 and incubated with membranes for 4h at room temperature with gentle shaking. After primary antibody incubation, membranes were washed 3x with TBS-Tween 20 (0.1%) for 10mins each to remove any unbound primary antibodies. Membranes were then incubated with the secondary antibody, anti-mouse *m*-IgG<sub>1</sub> (sc-525408), diluted 1:2,000 in TBS-Tween 20 (0.1%) containing 1% non-fat dry milk. This incubation lasted for 1h at room temperature with gentle shaking. Post-secondary antibody incubation, membranes were washed 3x with TBS-Tween 20 (0.1%) for 10mins each to remove any unbound secondary antibodies. Protein bands were visualized using a classic ECL detection system and chemiluminescent signal was captured on X-ray film. The intensity of the protein bands was quantified by Quantity One 1-D Analysis Software.  $\beta$ -globin dimers were detected using the previously referred to anti- $\beta$ -globin antibody (sc-21757, Santa Cruz Biotechnology).  $\beta$ -globin monomers and dimers were detected based on the molecular weight of the band.

### RNA-sequencing of edited erythroid cells

Total RNA was extracted from frozen pellets of approximately  $10^6$  cells *in vitro*-differentiated erythrocytes using RNeasy Plus Micro Kit according to the manufacturers' instructions. Sequencing was provided by Novogene. Raw sequencing reads were analyzed with FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>) to ensure adequate data quality. Reads were mapped to the GRCh38.p14 human genome assembly using the pseudo-alignment tool Salmon (<https://github.com/COMBINE-lab/salmon>).<sup>32</sup> After pseudo-alignment, transcript levels count estimates were collapsed to the gene level, and then transformed using a regular log transformation. Relative expression levels of genes of interest were compared using these normalized values.

### Gene ontology enrichment analysis

Gene sets were submitted to the GOrilla gene ontology (GO) enrichment analysis tool (<https://cbl-gorilla.cs.technion.ac.il/>)<sup>33</sup> for pathway analysis. GOrilla reports the FDR q-value of each GO term following Benjamini-Hochberg correction for multiple testing, and those GO processes with FDR  $\leq 0.01$  were considered to be highly significant. For heatmap generation, the top 0.001% most highly expressed genes were analyzed for enriched GO processes across all samples. To generate the bubble plot, the top 0.001% of genes with the greatest rank change in expression were identified in each comparison group, using mock or WT as the denominator. Enrichment is defined by GOrilla using the formula  $(b/n)/(B/N)$  where N is the total number of genes, B is the total number of genes associated with a specific GO term, n is the number of genes in the top of the target gene set, and b is the number of genes in the intersection. Values were exported to GraphPad Prism (v.9) for plotting and further analysis.

### Mouse conditioning and transplant

Twenty 6-8week old female NBSGW mice (strain # 026622, The Jackson Laboratory) were conditioned with 10 mg/kg of Busulfan (B2635, Sigma Aldrich) 48h and 24h before engraftment with edited human CD34<sup>+</sup> HSPCs. Busulfan was dissolved in 100% DMSO, then diluted 1:2 in PBS and kept at 37°C until injection. Immediately prior to injection, Busulfan was diluted once more at 1:10 in warm PBS to bring the DMSO concentration to 5%. 24h after the second Busulfan injection,  $0.5 \times 10^6$  HSPCs were injected retro-orbitally per mouse. Mice were then fed moist chow for the first 2d post-engraftment and closely monitored for the 2wks following engraftment. Mice were sacrificed 16wks after engraftment.

### Bone marrow isolation and flow

Femurs and tibias were collected for bone marrow isolation. Bone marrow collection tubes were prepared by punching a hole at the bottom of a 0.5mL Eppendorf tube using an 18G needle and placed in a 1.5mL siliconized Eppendorf tube. 100 $\mu$ L of bone marrow harvest media (RPMI, 10% FBS ES, 33.3 U/mL DNase, 4 U/mL heparin) was added to the 1.5mL tube for collection. Epiphyses were removed from each side of the bone and the exposed shaft was placed in a prepared 0.5mL Eppendorf tube for each mouse. Bones were spun at 10,000 x g for 20-30s. Cells were immediately resuspended in 1mL of bone marrow harvest media and passed through a prewashed 30 $\mu$ M filter. The filter was then washed with 5mL of bone marrow harvest media. Samples were then spun at 500 x g for 5mins and supernatant removed. Cells were then resuspended in 2.5mL of bone marrow harvest media, followed by 2.5mL of dilution buffer (1x PBS, 2% FBS ES, 2mM EDTA) and spun at 500 x g for 5mins. The supernatant was removed, and cells were washed one final time in dilution buffer at 500 x g for 5mins. Cells were resuspended in dilution buffer for flow cytometry.  $1 \times 10^6$  cells/mouse were incubated with TruStain FcX Plus (anti-mouse CD16/32) for 10mins before staining. Without washing, the following antibodies were added and incubated in the dark at room temperature for 20mins: anti-mouse Cd45.1 PE-Cy7 (110730, BioLegend), anti-mouse

TER-119 BUV395 (563827, BD Biosciences), anti-human HLA-ABC APC (555555, BD Biosciences), anti-human CD19 PE (555413, BD Biosciences), anti-human CD33 BV421 (562854, BD Biosciences), anti-human CD45 PerCP (368505, BioLegend), and anti-human CD235a Pacific Blue (306612, BioLegend). Samples were run on a 5-laser Cytex Aurora and data was analyzed using FlowJo version 10.10.0.

#### QUANTIFICATION AND STATISTICAL ANALYSIS

Throughout the study, when comparing two treatment groups we used unpaired two-tailed t-tests to determine  $p$ -values and assess statistical significance.  $P$ -values generated from these statistical tests are reported in [Figures 1, 2, and 4](#). When comparing across multiple treatment groups, we used 2-way analysis of variance (ANOVA) for multiple comparisons to determine  $p$ -values and assess statistical significance.  $P$ -values generated from these statistical tests are reported in [Figure S6](#). GraphPad Prism 9 software was used for all statistical analysis.