Article

Base editing of haematopoietic stem cells rescues sickle cell disease in mice

https://doi.org/10.1038/s41586-021-03609-w

Received: 10 January 2021

Accepted: 4 May 2021

Published online: 02 June 2021

Check for updates

Gregory A. Newby[1,2,3,12], Jonathan S. Yen[4,12]✉, Kaitly J. Woodard[4,12], Thiyagaraj Mayuranathan[4,12], Cicera R. Lazzarotto[4], Yichao Li[4], Heather Sheppard-Tillman[5], Shaina N. Porter[6], Yu Yao[4], Kalin Mayberry[4], Kelcee A. Everette[1,2,3], Yoonjeong Jang[4], Christopher J. Podracky[1,2,3], Elizabeth Thaman[7], Christophe Lechauve[4], Akshay Sharma[8], Jordana M. Henderson[9], Michelle F. Richter[1,2,3], Kevin T. Zhao[1,2,3], Shannon M. Miller[1,2,3], Tina Wang[1,2,3], Luke W. Koblan[1,2,3], Anton P. McCaffrey[9], John F. Tisdale[10], Theodosia A. Kalfa[7,11], Shondra M. Pruett-Miller[6], Shengdar Q. Tsai[4], Mitchell J. Weiss[4]✉ & David R. Liu[1,2,3]✉

Sickle cell disease (SCD) is caused by a mutation in the β-globin gene *HBB*[1]. We used a custom adenine base editor (ABE8e-NRCH)[2,3] to convert the SCD allele (*HBB*^{S}) into Makassar β-globin (*HBB*^{G}), a non-pathogenic variant[4,5]. Ex vivo delivery of mRNA encoding the base editor with a targeting guide RNA into haematopoietic stem and progenitor cells (HSPCs) from patients with SCD resulted in 80% conversion of *HBB*^{S} to *HBB*^{G}. Sixteen weeks after transplantation of edited human HSPCs into immunodeficient mice, the frequency of *HBB*^{G} was 68% and hypoxia-induced sickling of bone marrow reticulocytes had decreased fivefold, indicating durable gene editing. To assess the physiological effects of *HBB*^{S} base editing, we delivered ABE8e-NRCH and guide RNA into HSPCs from a humanized SCD mouse[6] and then transplanted these cells into irradiated mice. After sixteen weeks, Makassar β-globin represented 79% of β-globin protein in blood, and hypoxia-induced sickling was reduced threefold. Mice that received base-edited HSPCs showed near-normal haematological parameters and reduced splenic pathology compared to mice that received unedited cells. Secondary transplantation of edited bone marrow confirmed that the gene editing was durable in long-term haematopoietic stem cells and showed that *HBB*^{S}-to-*HBB*^{G} editing of 20% or more is sufficient for phenotypic rescue. Base editing of human HSPCs avoided the p53 activation and larger deletions that have been observed following Cas9 nuclease treatment. These findings point towards a one-time autologous treatment for SCD that eliminates pathogenic *HBB*^{S}, generates benign *HBB*^{G}, and minimizes the undesired consequences of double-strand DNA breaks.

Sickle-cell disease is an autosomal recessive disorder caused by mutation of *HBB*, which normally encodes adult β-globin (β^A) (Fig. 1a). At low oxygen concentrations, the mutant β-globin (β^S) causes haemoglobin polymerization within red blood cells (RBCs), which results in characteristic sickle-shaped RBCs and a cascade of haemolysis, inflammation, and microvascular occlusions. Symptoms include anaemia, severe acute and chronic pain, immunodeficiency, multi-organ failure and early death[1]. Although allogeneic haematopoietic stem cell (HSC) transplantation can cure SCD, optimally matched donors are usually not available and the procedure can result in graft rejection or graft-versus-host disease (Supplementary References).

Ex vivo modification of autologous HSCs to circumvent the deleterious effects of the SCD mutation underlies several experimental therapies[7–9] (Supplementary References). Approaches that have shown early clinical promise include ectopic expression of an anti-sickling β-like globin gene by lentiviral vectors[10] and induction of fetal haemoglobin (HbF) by suppression[11] or Cas9-mediated disruption[12] of *BCL11A*. Lentiviral vectors carry risks of insertional mutagenesis, however, and may not effectively suppress the expression of pathological β^S. Genetic manipulation to induce expression of HbF does not eliminate β^S, and when mediated by double-stranded DNA breaks (DSBs), carries risks associated with uncontrolled mixtures of indels (insertions

[1]Merkin Institute of Transformative Technologies in Healthcare, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [2]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA. [3]Howard Hughes Medical Institute, Harvard University, Cambridge, MA, USA. [4]Department of Hematology, St. Jude Children's Research Hospital, Memphis, TN, USA. [5]Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, USA. [6]Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN, USA. [7]Division of Hematology, Cancer and Blood Diseases Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. [8]Department of Bone Marrow Transplantation and Cellular Therapy, St Jude Children's Research Hospital, Memphis, TN, USA. [9]TriLink BioTechnologies, San Diego, CA, USA. [10]Cellular and Molecular Therapeutics Branch, National Heart, Lung, and Blood Institute and National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, USA. [11]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA. [12]These authors contributed equally: Gregory A. Newby, Jonathan S. Yen, Kaitly J. Woodard, Thiyagaraj Mayuranathan. ✉e-mail: jonathan.yen@stjude.org; mitch.weiss@stjude.org; drliu@fas.harvard.edu

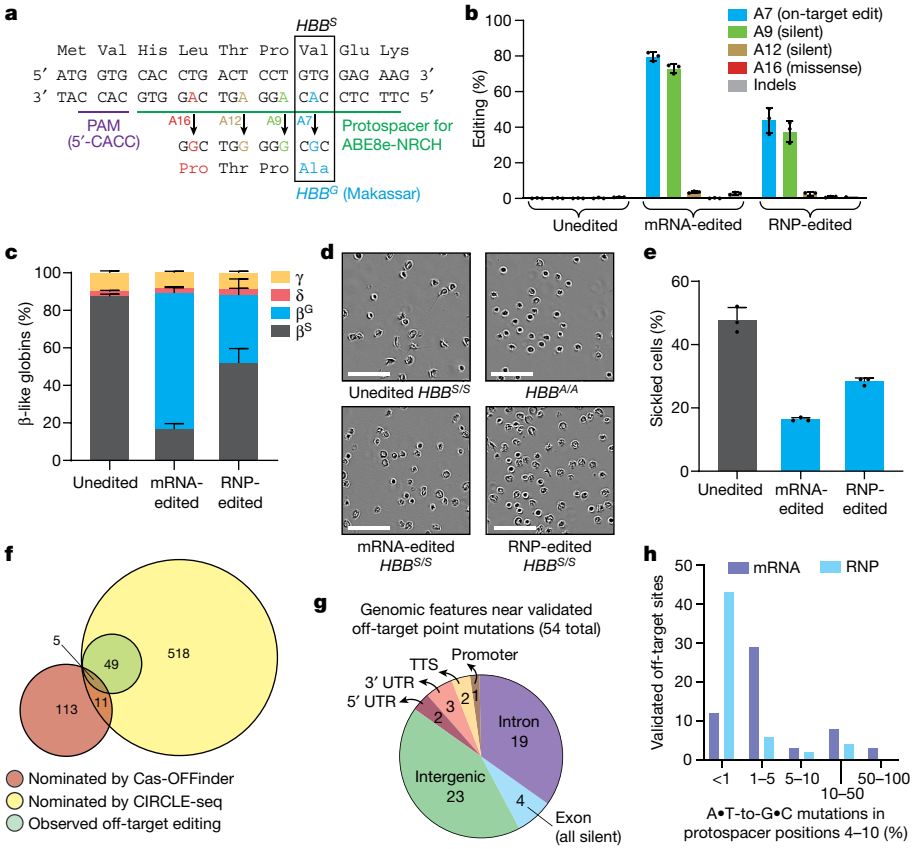


Fig. 1 | Adenine base editing converts SCD β-globin gene (*HBB*^{S}) to benign Makassar β-globin gene (*HBB*^{G}) in patient CD34$^+$ HSPCs. CD34$^+$ cells from three donors with SCD were electroporated with ABE8e-NRCH mRNA or RNP using an sgRNA targeting the SCD mutant *HBB* codon. **a**, The edited region of *HBB* with the target A at protospacer position 7 shown in blue along with potential bystander edits in green (silent), brown (silent), and red (non-silent). **b**, Editing efficiencies determined by HTS at target and bystander adenines and indels after 6 d in stem-cell culture medium after electroporation. **c**, Proportion of β-like globin proteins determined by HPLC of reticulocyte lysates after 18 d in differentiation medium after electroporation. **d**, Representative phase-contrast images of reticulocytes derived from unedited or edited donor HSPCs incubated for 8 h with 2% oxygen. Nine images of more than 50 cells each were collected per sample. Scale bars, 50 µm. **e**, Quantification of sickled reticulocytes from images as in **d** (more than 300 randomly selected cells counted by a blinded observer for each condition). **f**, Venn diagram of candidate off-target sites nominated by Cas-OFFinder and CIRCLE-seq, and nominated sites for which off-target editing was observed by targeted DNA sequencing in SCD CD34$^+$ cells electroporated with ABE8e-NRCH mRNA. **g**, Predicted genomic features of validated off-target sites. TTS, 1 kb or less from the transcription termination site; UTR, untranslated region. **h**, ABE8e-NRCH-treated HSPCs from two donors with SCD were sequenced at 697 potential off-target sites. The histogram shows the number of validated off-target base editing sites binned by average percentage of sequencing reads for each site with any A•T-to-G•C mutations in protospacer nucleotides 4–10. **b**, **c**, **e**, Data shown as mean ± s.d. of three independent biological replicates, with individual values shown as dots in **b**, **e**.

and deletion), translocations, loss of large chromosomal segments, chromothripsis, and p53 activation[13–18] (Supplementary References). Cas9 nuclease-mediated homology-directed repair can correct *HBB*^{S} but is difficult to achieve efficiently in repopulating HSCs[19,20] and also requires DSBs. Elimination of the root cause of SCD by converting the *HBB*^{S} allele to a benign variant without introducing DSBs could overcome these limitations.

Adenine base editors (ABEs) convert targeted A•T base pairs to G•C in living cells without requiring DSBs or donor DNA templates, and with minimal formation of indels[2]. In SCD, the GAG (Glu) codon that encodes amino acid 6 of β-globin is mutated to GTG (Val). Although adenine base editing cannot revert this mutation, it can convert the pathogenic codon to GCG (Ala), which produces a naturally occurring, non-pathogenic variant termed Hb-Makassar (*HBB*^{G})[3–5,21,22] (Fig. 1a).

We generated an ABE (ABE8e-NRCH[2,3]) that converts the SCD allele to the non-pathogenic *HBB*^{G} Makassar allele with minimal non-silent bystander edits in CD34$^+$ HSPCs from patients with SCD. Edited HSPCs were durable after engraftment in mice, with an *HBB*^{G} frequency of 68% sixteen weeks after transplantation and markedly reduced sickling in the derived erythroid cells. To assess phenotypic rescue, we edited HSPCs from a mouse model of SCD[6] in which endogenous β-globin genes are replaced by human *HBB*^{S}, and transplanted the edited HSPCs into irradiated adult recipient mice. Primary and secondary transplantation of edited mouse HSPCs confirmed editing in long-term HSCs and restored haematological parameters to near-normal levels. These findings show that autologous ex vivo base editing and transplantation of HSCs is a potential one-time treatment for SCD.

HBB^S base editing in HSPCs

The *HBB*^{S} mutation can be targeted by base editing using a phage-assisted continuous evolution (PACE)-generated Cas9-NRCH[3] that recognizes a CACC protospacer-adjacent motif (PAM; Fig. 1a). Separately, we used PACE to evolve TadA-8e, a deoxyadenosine deaminase that supports highly efficient base editing[2]. We combined TadA-8e with Cas9-NRCH nickase to generate ABE8e-NRCH. Co-delivery of ABE8e-NRCH and the *HBB*^{S}-targeting single guide RNA (sgRNA) to homozygous *HBB*^{S} HEK293T cells by plasmid lipofection achieved 58% conversion of *HBB*^{S} to *HBB*^{G} (Extended Data Fig. 1a).

Next, we used ABE8e-NRCH to edit human HSPCs ex vivo through electroporation of either the ABE8e-NRCH + sgRNA ribonucleoprotein (RNP) or ABE8e-NRCH mRNA with sgRNA. ABE8e-NRCH

RNP electroporated into plerixafor-mobilized peripheral blood CD34$^+$ HSPCs cells from three donors with SCD resulted in 44 ± 5.9% (mean ± s.d.) editing of *HBB^S* to *HBB^G*, 1.2 ± 0.33% indels, and less than 0.5% other missense alleles after 6 days. Electroporation of ABE8e-NRCH mRNA and sgRNA into the same cells resulted in 80 ± 2.1% conversion of *HBB^S* to *HBB^G*, 2.8 ± 0.50% indels, and less than 2% other missense bystander alleles (Fig. 1b, Extended Data Fig. 1b–c). Thus, introduction of ABE8e-NRCH RNP or mRNA using a clinically relevant delivery method can convert *HBB^S* to *HBB^G* in HSPCs efficiently and with few byproducts.

HSPCs from patients with SCD (referred to as SCD cells) that had been edited with ABE8e-NRCH mRNA or RNP were differentiated ex vivo into late-stage erythroid precursors (Extended Data Fig. 2, Supplementary Table 1). Quantification of β-like globin proteins by high-performance liquid chromatography (HPLC) showed that unedited SCD cells contained 87 ± 1.3% β^S and no detectable β^G (Extended Data Fig. 3), whereas base-edited cells contained 72 ± 3.0% β^G and 17 ± 3.0% β^S, a 5.1-fold decrease in the pathogenic β^S protein (Fig. 1c). These findings show that ABE8e-NRCH-mediated editing of *HBB^S* results in substantial production of β^G and concomitant loss of β^S protein.

To assess the effects of editing on sickling, we incubated purified reticulocytes from ex vivo differentiation of unedited or ABE8e-NRCH-edited SCD CD34$^+$ cells in 2% oxygen. Editing reduced sickling frequency from 47.7% to 16.3% (Fig. 1d, e), confirming that *HBB^S*-to-*HBB^G* conversion reduced sickling. Reticulocytes that were differentiated from cells treated with ABE8e-NRCH RNP showed similar results, but with lower efficiencies (Fig. 1c–e).

To determine whether base editing alters erythropoiesis, we used flow cytometry to track the expression of the cell-surface maturation markers CD49d, CD235a and BAND3. We found no differences in the expression of these markers between edited and unedited cells (Extended Data Fig. 2), suggesting that editing with ABE8e-NRCH does not alter erythropoiesis.

Genome-wide off-target analyses

We used both computational and experimental methods to extensively characterize off-target editing as a result of treatment with ABE8e-NRCH and sgRNA (Fig. 1f, Supplementary Discussion). The Cas-OFFinder algorithm[23] identified 140 NRCH PAM-containing human genomic sites with three or fewer mismatches to the target protospacer. We also performed CIRCLE-seq[24], a highly sensitive experimental off-target identification method, to identify where Cas9-NRCH, complexed with the *HBB^S*-targeting sgRNA, cleaved purified human genomic DNA in vitro. CIRCLE-seq identified 601 candidate off-target sites (Supplementary Table 2). The 140 sites nominated by Cas-OFFinder and the 601 sites nominated by CIRCLE-seq shared only 16 sites in common.

Of the 725 candidate off-target sites, 697 were amenable to multiplex-targeted DNA sequencing in SCD CD34$^+$ HSPCs treated with ABE8e-NRCH (Supplementary Discussion). We detected point mutations consistent with adenine base editing at 7.8% (54/697) of the sequenced sites. All 54 verified sites were candidates identified by CIRCLE-seq; five were also identified by Cas-OFFinder (Fig. 1f, Extended Data Figs. 4, 5, Supplementary Table 2), highlighting the importance of experimental identification of off-target sites. Off-target activity occurred predominantly in intergenic and intronic regions. One off-target site was in the promoter region of *CCDC85B* and four were in exons (Fig. 1g); activity at all of these sites led to silent mutations (Supplementary Table 2). Off-target sites in untranslated regions (UTRs) were all more than 500 bp from any coding region.

As anticipated, off-target editing by the RNP was lower than that resulting from the mRNA, probably because of the shorter duration of exposure or lower editing activity[2] (Fig. 1h, Extended Data Figs. 4, 5). Indel frequencies were lower than 2% at all off-target sites (Extended Data Fig. 5). Collectively, our extensive genome-wide off-target analyses of ABE8e-NRCH base-edited SCD HSPCs did not identify off-target mutations with anticipated clinical relevance.

Transplantation of human HSPCs into mice

We next investigated whether delivery of ABE8e-NRCH into SCD CD34$^+$ cells can convert *HBB^S* to *HBB^G* in HSCs that are used to repopulate bone marrow in an animal. CD34$^+$ HSPCs from three patients with SCD were edited by electroporation of ABE8e-NRCH and sgRNA in the RNA or RNP forms. After 24 h, the resulting six sets of edited HSPCs and a set of unedited control cells from each donor were each transplanted via tail-vein injection into 3–5 immunodeficient NOD B6.SCID *Il2rγ$^{-/-}$Kit$^{W41/W41}$* (NBSGW) mice[25]. Sixteen weeks after infusion, when persisting human cells are thought to be generated from bone marrow-repopulating HSCs capable of sustaining a haematopoietic system[25], we extracted bone marrow from the mice for analysis (Fig. 2a).

The disruption of targeted genes through DSBs or deletion can alter the engraftment and maintenance of certain lineages (Supplementary References). To determine how base editing of *HBB^S* affected differentiation potential and lineage survival, we assessed the human haematopoietic lineages that were present in recipient mouse bone marrow after transplantation. Flow cytometry using an anti-human CD45 antibody showed that human cells made up about 70% of bone marrow in all mice (Fig. 2b). Flow cytometry to quantify the relative abundances of human B cells (CD19$^+$), myeloid cells (CD33$^+$), T cells (CD3$^+$), and erythroid cells (CD235a$^+$) showed that the proportions of each lineage were equivalent in mice that received unedited or edited cells (Fig. 2c, d, Supplementary Fig. 1), indicating that the engraftment and differentiation potential of CD34$^+$ cells was not altered by base editing.

To examine the possibility that base editing caused skewed haematopoiesis, we used human lineage-specific antibodies to purify donor-derived mononuclear cells ('total bone marrow'; CD45$^+$), B cells (CD19$^+$), myeloid cells (CD33$^+$), HSPCs (CD34$^+$) and erythroblasts (CD235a$^+$) from mouse bone marrow (Supplementary Fig. 1). High-throughput sequencing (HTS) of the targeted genomic region showed that all isolated populations contained the *HBB^S*-to-*HBB^G* edit at similar frequencies (68 ± 6.6% to 69 ± 5.7%; Fig. 2e), which suggests that this allele proportion was maintained in HSCs and in their differentiated progeny. The modest decrease from 80% editing observed before transplantation could reflect slightly higher editing efficiency of non-repopulating cells within the complex mixture of CD34$^+$ HSPCs, with 68% editing efficiency achieved in repopulating HSCs. Collectively, these results indicate that ABE8e-NRCH-mediated conversion of *HBB^S* to *HBB^G* in repopulating HSCs does not impede their engraftment or multipotency.

Human CD235a$^+$ erythroblasts and reticulocytes isolated from the bone marrow of two mice transplanted with ABE-treated or untreated SCD cells were purified by magnetic-activated cell sorting (MACS; Extended Data Fig. 6) and subjected to single-cell RNA sequencing (RNA-seq) to determine the outcomes of clonal editing. An average of 46.5% of cells were edited in only one *HBB^S* allele, 40.6% were edited in both alleles, and 12.9% of cells were unedited (Fig. 2f). Base editing decreased the fraction of β^S from 96 ± 0.28% of total β-like globin protein to 40 ± 2.3% in CD235a$^+$ cells. β^G was undetectable in unedited cells but accounted for 58 ± 2.8% of β-like globin after base editing (Fig. 2g). Human erythroid cells derived from edited HSPCs showed a fivefold reduction in sickling compared to unedited control cells (Fig. 2h, i). Editing using RNP had similar effects but was less efficient (Extended Data Fig. 7). These data indicate that base-edited HSCs from donors with SCD can repopulate the haematopoietic system and generate erythroid cells with a greatly reduced propensity for hypoxic sickling.

Transplantation of mouse HSPCs into mice

Studying the physiological rescue of SCD phenotypes by transplantation of human cells into mice is difficult owing to the short lifetime of

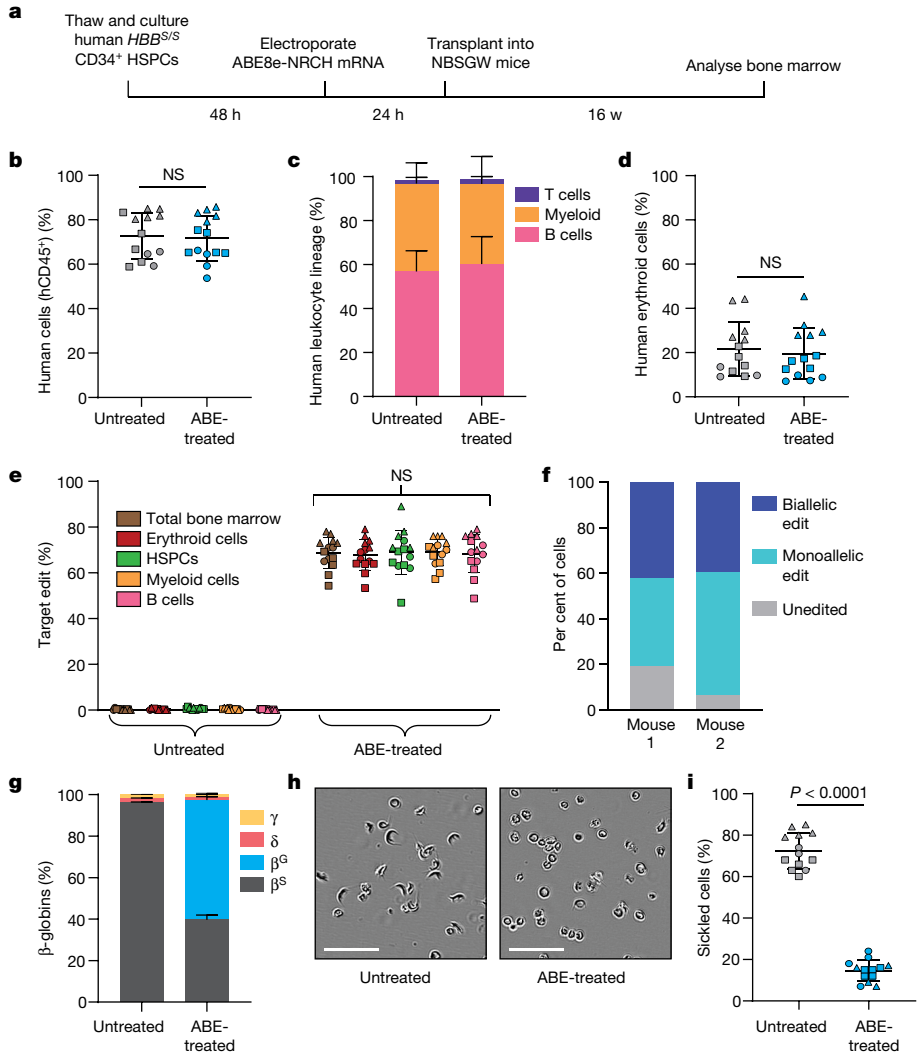


Fig. 2 | Engraftment of ABE8e-NRCH mRNA-treated SCD CD34$^+$ HSPCs after transplantation into immunodeficient mice. CD34$^+$ HSPCs from three donors with *HBB*$^{S/S}$ SCD were electroporated with ABE8e-NRCH mRNA and sgRNA targeting the SCD mutant *HBB* codon. We transplanted 2–5 × 10^5 treated cells into NBSGW mice via tail-vein injection. Mouse bone marrow was collected and analysed 16 weeks after transplantation. **a**, Experimental workflow. **b**, Engraftment measured by percentage of human CD45$^+$ (hCD45$^+$) cells in recipient mouse bone marrow. **c**, Human B cells (hCD19$^+$), myeloid cells (hCD33$^+$), and T cells (hCD3$^+$) as percentages of the hCD45$^+$ population in recipient mouse bone marrow. **d**, Human erythroid precursors (hCD235a$^+$) as percentage of human and mouse CD45$^-$ cells in recipient mouse bone marrow. **e**, *HBB*^{S}-to-*HBB*^{G} editing efficiency in human CD34$^+$ cell-derived lineages from recipient bone marrow. Erythroid, myeloid, B cell, and HSPC human lineages were collected using antibodies against hCD235a, hCD33, hCD19, and hCD34, respectively. **f**, Clonal editing outcomes determined by single-cell 5′ RNA-seq in CD235a$^+$ cells from the bone marrow of two recipient mice. **g**, Proportions of β-like globin proteins determined by HPLC of human donor-derived reticulocytes isolated from recipient mouse bone marrow. **h**, Representative phase-contrast images of human reticulocytes from bone marrow incubated for 8 h with 2% oxygen. Nine images of more than 50 cells each were collected per sample. Scale bars, 50 µm. **i**, Quantification of sickled cells as in Fig. 1e. $n = 14$ mice that received edited cells and $n = 13$ mice that received unedited cells in **b**–**e**, **g**, **i**. Triangle, square, and circle symbols represent HSPCs from three different donors with SCD. Data shown as mean ± s.d.; one-way ANOVA (**i**), two-tailed Student's *t*-test (elsewhere); NS, not significant.

circulating human RBCs in mice[25]. To evaluate physiological phenotypes, we edited lineage-negative (Lin$^-$) HSPCs from the Townes SCD mouse model, in which endogenous adult α- and β-like globin genes are replaced by human globin genes, resulting in SCD phenotypes[6]. Mice harbouring one normal and one SCD *HBB* allele (*HBB*$^{A/S}$) model a heterozygous 'sickle-cell trait', which is largely asymptomatic in this mouse model (Supplementary Table 3) and in humans.

We electroporated ABE8e-NRCH RNP into *HBB*$^{S/S}$ HSPCs from Townes mice and then transplanted the cells into irradiated adult recipient mice 24 h later. Unedited *HBB*$^{S/S}$ and *HBB*$^{A/S}$ HSPCs were used as disease and healthy controls, respectively. As donor mouse cells express CD45.2, whereas recipient cells express CD45.1, they can be distinguished using allele-specific antibodies[6] (Fig. 3a, Supplementary Table 1). We collected blood 6, 10, 14, and 16 weeks after transplantation to track engraftment and β-globin content.

CD45.2 expression showed that donor engraftment was above 90% in all mice ten weeks after transplantation (Extended Data Fig. 8a). Engraftment of edited and control donor HSPCs progressed similarly, suggesting that editing did not alter the fitness of transplanted HSCs. *HBB*^{S}-to-*HBB*^{G} editing efficiency measured 3 days after electroporation (that is, before transplantation) was 53 ± 4.5% (Fig. 3b). Editing levels in genomic DNA from whole blood in mice 16 weeks after transplantation showed an *HBB*^{G} allele frequency of 44 ± 11%. As we observed with human HSPCs (Fig. 2e), the modest decrease in *HBB*^{G} allele frequency after 16 weeks of engraftment could arise if repopulating HSCs are less amenable to electroporation or base editing than other cell types

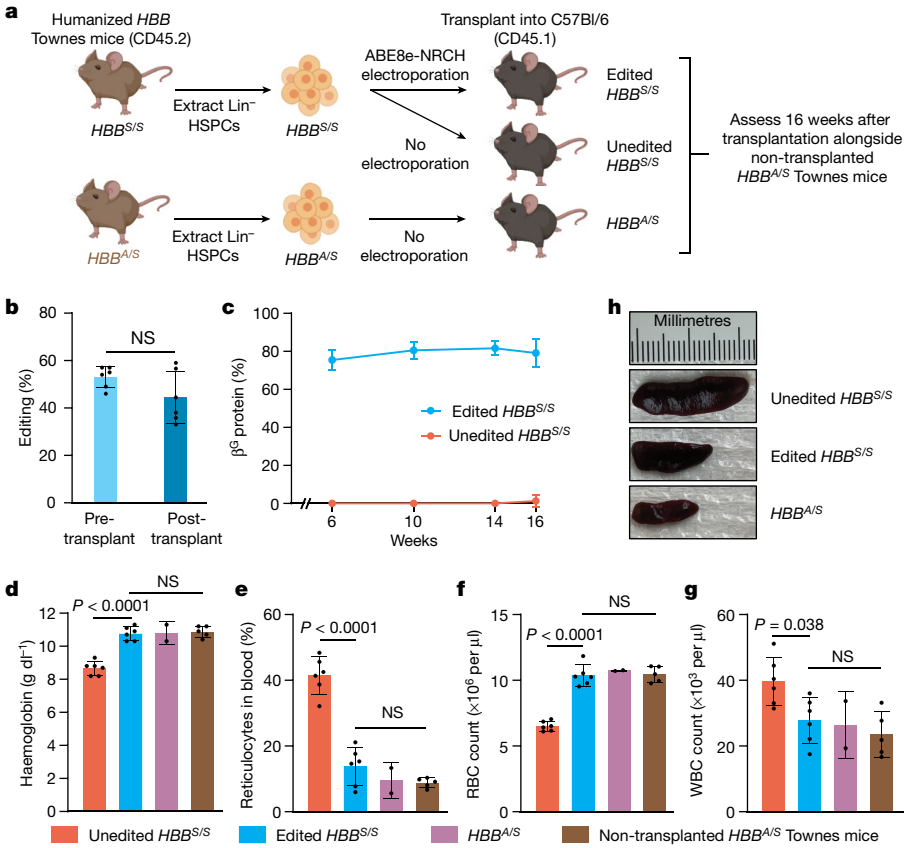


Fig. 3 | *HBB*^{S}-to-*HBB*^{G} base editing alleviates pathology in a mouse model of SCD. a, Lin$^-$ HSPCs from the bone marrow of Townes SCD mice (CD45.2, human *HBB*$^{S/S}$) were electroporated with ABE8e-NRCH and sgRNA RNP or not electroporated, then transplanted into irradiated CD45.1 C57Bl/6 recipient mice. Unedited *HBB*$^{A/S}$ HSPCs from Townes sickle-cell trait mice transplanted into irradiated CD45.1 C57Bl/6 mice and non-transplanted *HBB*$^{A/S}$ Townes mice were used as healthy controls. **b**, *HBB*^{S}-to-*HBB*^{G} editing efficiency in cells cultured 3 days after electroporation (before transplant) or in PBMCs collected 16 weeks after transplant. NS, not significant by two-tailed Student's *t*-test. **c**, Percentage of β^G among β-like globin proteins determined by reverse phase (RP)-HPLC analysis of blood. **d**–**g**, Haematological indices 16 weeks after transplantation. Statistical significance assessed using one-way ANOVA with Šidák's multiple comparisons test to calculate *P* values. Differences among edited *HBB*$^{S/S}$, transplanted *HBB*$^{A/S}$, and non-transplanted *HBB*$^{A/S}$ mice were not significant. Data shown as mean ± s.d. of $n = 6$ mice (unedited *HBB*$^{S/S}$, edited *HBB*$^{S/S}$), $n = 2$ mice (*HBB*$^{A/S}$), or $n = 5$ mice (non-transplanted *HBB*$^{A/S}$). **h**, Spleens were imaged from each mouse 16 weeks after transplantation with Townes mouse HSPCs. Representative images are shown.

within the HSPC population. A clonal analysis of colonies from bone marrow cells after mouse euthanasia showed that 40 ± 15% of cells were edited in both *HBB*^{S} alleles, 36 ± 12% in only one allele, and 24 ± 3.2% of cells in neither (Extended Data Fig. 8b).

To measure the effects of base editing on haemoglobin composition in circulating RBCs, we analysed peripheral blood cell lysates from each time point. On average, β^G made up 75–82% of total β-like globin protein in mice that had received edited *HBB*$^{S/S}$ HSPCs, with little fluctuation throughout the experiment (Fig. 3c). The enrichment of β^G over the observed editing efficiency (79% versus 44% at 16 weeks) is likely to reflect the increased lifetime of β^G-containing RBCs. We also found no difference in oxygen binding in blood from mice that had received unedited *HBB*$^{S/S}$ cells, *HBB*$^{A/S}$ cells, or edited *HBB*$^{S/S}$ cells 14 weeks after transplantation, which suggests that β^G-containing haemoglobin binds oxygen normally (Extended Data Fig. 8c).

Rescue of SCD in transplanted mice

We performed complete blood counts on mice transplanted with edited ($n = 6$) or unedited ($n = 6$) mouse *HBB*$^{S/S}$ HSPCs, and on mice transplanted with unedited *HBB*$^{A/S}$ cells ($n = 2$) or non-transplanted mice with an *HBB*$^{A/S}$ genotype ($n = 5$); the latter two groups served as two types of healthy control (Fig. 3d–g, Supplementary Table 3). Compared to healthy controls, mice that received unedited mouse *HBB*$^{S/S}$ HSPCs showed disruptions in total haemoglobin concentration and cell counts of reticulocytes, RBCs, and white blood cells (WBCs; Fig. 3d–g)—abnormalities that are consistent with haemolytic anaemia and inflammation in patients with SCD[6]. Notably, transplantation of base-edited *HBB*$^{S/S}$ HSPCs rescued the haematological defects in these mice, restoring all tested blood parameters to levels similar to those of healthy controls (Fig. 3d–g).

To assess the consequences of base editing on circulating RBC morphology, we analysed blood from mice 16 weeks after transplantation (Extended Data Fig. 9a). We found the expected morphological abnormalities in RBCs from mice that received unedited *HBB*$^{S/S}$ cells, including abundant oblong sickle forms, polychromasia (reflecting reticulocytosis), and fragmentation. RBCs from mice transplanted with *HBB*^{S}-to-*HBB*^{G}-edited HSPCs showed a reduction in all pathological morphologies and were more similar to RBCs from healthy *HBB*$^{A/S}$ controls. Separately, we incubated blood in 2% oxygen to induce sickling (Extended Data Fig. 9b, c). RBCs from mice transplanted with unedited *HBB*$^{S/S}$ HSPCs showed 86.3 ± 3.0% sickling, compared with 29.8 ± 6.5% sickling in RBCs from mice transplanted with base-edited *HBB*$^{S/S}$ HSPCs, a 2.9-fold decrease. These data show that transplantation of edited *HBB*$^{S/S}$ HSPCs leads to durable production of RBCs that are resistant to sickling both in vivo and in vitro after exposure to hypoxia.

An enlarged spleen is a hallmark symptom of young patients with SCD and of mouse models[6]. The average spleen mass of mice that received

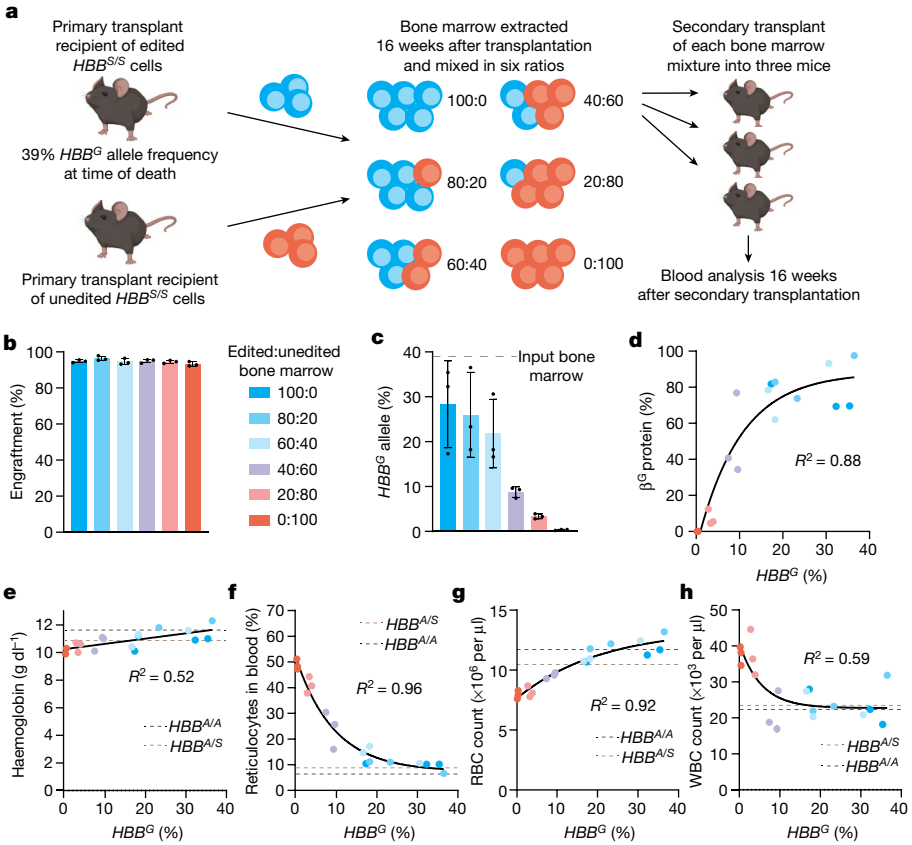


Fig. 4 | Secondary transplantation reveals *HBB*^{S}-to-*HBB*^{G} base editing requirements for haematological correction. a, Bone marrow from a CD45.1 C57Bl/6 mouse 16 weeks after primary transplantation with ABE8e-NRCH RNP-edited Lin^{-} HSPCs from Townes SCD mice (CD45.2, *HBB*$^{S/S}$) was mixed in varying proportions with bone marrow from a C57Bl/6 mouse 16 weeks after transplantation with unedited *HBB*$^{S/S}$ HSPCs from a Townes SCD mouse. For each of six bone marrow mixtures, secondary transplantations of 2×10^6 cells were performed into three irradiated CD45.1 C57Bl/6 recipient mice. Peripheral blood was analysed after 16 weeks. **b**, Engraftment measured by percentage of PBMCs with CD45.1. **c**, *HBB*^{S}-to-*HBB*^{G} editing efficiency in PBMCs. **d**, Percentage of β^{G} among β-like globin proteins determined by HPLC. **e**–**h**, Haematological indices plotted against *HBB*^{G} allele frequency measured for each mouse. Parameters from non-transplanted *HBB*$^{A/S}$ (brown dashed line) and *HBB*$^{A/A}$ (black dashed line) Townes mice were assessed as healthy controls. **d**, **f**–**h**, One-phase decay fits; **e**, linear fit. Data shown as mean ± s.d. of $n = 3$ mice. Dots represent different mice. **c**–**h**, Coloured as in **b**.

edited *HBB*$^{S/S}$ cells was 0.22 ± 0.043 g, compared to 0.39 ± 0.016 g in mice that received unedited *HBB*$^{S/S}$ cells and 0.11 ± 0.007 g in mice that received *HBB*$^{A/S}$ cells (Extended Data Fig. 9d). Average spleen mass was thus restored by 61% towards that of healthy controls in mice that received base-edited HSPCs (Fig. 3h). RBC pooling and extramedullary erythropoiesis in the spleen were also largely corrected in mice that received edited *HBB*$^{S/S}$ cells (Extended Data Fig. 9e). Thus, the persistence of ABE8e-NRCH-mediated *HBB*^{S}-to-*HBB*^{G} editing in bone marrow-repopulating HSCs and the partial or complete rescue of every examined SCD phenotype suggest that ex vivo base editing of *HBB*^{S} in HSPCs followed by transplantation can alleviate SCD.

Secondary transplant dose-dependent rescue

We performed secondary transplantations to confirm editing of long-term repopulating HSCs and determine the level of *HBB*^{S}-to-*HBB*^{G} base editing required to rescue SCD-associated haematological abnormalities. Following primary transplantation of mouse HSPCs as described above, bone marrow extracted from one recipient mouse had an *HBB*^{G} allele frequency of 39%. This bone marrow was mixed at ratios of 0:100, 20:80, 40:60, 60:40, 80:20, and 100:0 with bone marrow from a mouse that had been transplanted 16 weeks earlier with unedited *HBB*$^{S/S}$ Lin^{-} HSPCs. Each mixture was transplanted separately into three irradiated C57Bl/6 mice (Fig. 4a). Sixteen weeks after secondary transplantation, we collected peripheral blood to assess engraftment (Fig. 4b), *HBB*^{G} allele frequency and haematological phenotypes.

Mice that received mixtures containing at least 60% marrow from the recipients of edited HSCs maintained *HBB*^{G} allele frequencies of more than 20% following secondary transplant (Fig. 4c). In these mice, β^{G} protein represented more than 70% of all β-like globins in blood (Fig. 4d), and haematological parameters were similar to those of healthy *HBB*$^{A/S}$ and *HBB*$^{A/A}$ mice (Fig. 4e–h). Together, these results demonstrate durable base editing of long-term repopulating mouse HSCs and show that an *HBB*^{G} allele frequency of about 20% in engrafted cells—a threshold that was substantially exceeded by our base editing strategy—is sufficient to rescue haematological phenotypes.

Effects of nuclease and ABE treatment

Induction of HbF through Cas9-nuclease-mediated disruption of an erythroid *BCL11A* enhancer is in clinical trials for the treatment of SCD and β-thalassemia[12]. Nuclease-mediated DSBs have been reported to stimulate DNA damage responses that can enrich oncogenic cells[16,18], and to cause large DNA deletions or rearrangements that are difficult to detect by standard amplicon sequencing[14,15]. To compare DNA damage responses in HSPCs treated with either ABE8e-NRCH or Cas9, we performed reverse transcription and droplet digital PCR (ddPCR) to measure expression of *CDKN1* (also known as *P21*)—a readout of the p53-mediated DNA damage response (Extended Data Fig. 10a).

We assessed untreated cells, cells that had been electroporated with ABE8e-NRCH mRNA and sgRNA, cells that had been electroporated with Cas9 nuclease RNP targeting the *BCL11A* enhancer, and cells that had been electroporated with no cargo. We used HSPCs from a healthy donor and therefore we altered the sgRNA delivered with ABE8e-NRCH by one nucleotide to match the wild-type *HBB* locus.

Nuclease-treated cells showed 2.7-fold higher average *CDKN1* transcript levels 6 h after treatment, and 4.2-fold higher levels after 48 h, compared with control cells that had been electroporated with no cargo (Extended Data Fig. 10a). By contrast, cells treated with the base editor did not show an increase in *CDKN1*. Six days after electroporation, DNA sequencing revealed 84 ± 1.8% indels at the *BCL11A* locus in Cas9 RNP-treated cells and 64 ± 5.2% adenine base editing at *HBB* protospacer position 9 in ABE8e-NRCH treated cells (Extended Data Fig. 10b).

To detect long deletions or rearrangements at the targeted loci that could be missed by standard amplicon sequencing[14,15], we conducted ddPCR to quantify the amount of each target genomic locus six days after electroporation. We designed primers and probes for *BCL11A* to hybridize outside the range of deletions previously described for Cas9-mediated editing at this locus, so only longer deletions or DNA rearrangements should cause apparent loss of the target locus. We found no change in *HBB* allele quantification following base editor treatment (Extended Data Fig. 10c), but *BCL11A* allele quantification decreased by 14% relative to non-targeted *ACTB* following Cas9 nuclease treatment (Extended Data Fig. 10d), which suggests that long deletions or rearrangements occurred in approximately 14% of *BCL11A* alleles. This frequency is consistent with recent findings in edited human embryos[26]. Together, these results are concordant with previous findings[14] and suggest that base editor treatment of human HSPCs causes less stimulation of the p53 pathway and fewer large target site perturbations than Cas9 nuclease treatment.

Discussion

We describe a bespoke ABE that directly converts the major SCD allele to a β-globin variant that is non-pathogenic, even in homozygous[4] or hemizygous[21] form. This base editing strategy was efficient (up to 80% editing in HSPCs and 68% in bone marrow-repopulating HSCs), with minimal bystander edits or indels, and yielded non-sickling RBCs without disruption of globin gene regulation or haematopoiesis.

Several approaches for autologous therapies to treat SCD are being tested in clinical trials[7,8,10–12]. It is not yet known which strategy is safest or most effective. However, the base editing approach demonstrated here offers several potential advantages. First, elimination of the disease-causing mutation by precise HBB^S-to-HBB^G editing may reduce the concentration of sickle haemoglobin in RBCs (the primary determinant of pathogenic haemoglobin polymerization) more effectively than lentiviral expression of β-like globin or induction of HbF, both of which leave HBB^S alleles intact. Although the latter approaches can decrease the fraction of β^S in erythroid progeny by 30–70%[27,28], we achieved even greater β^S reduction in erythroid populations by base editing HBB^S.

Second, base editing largely avoids DSBs generated by nucleases, which lead to uncontrolled mixtures of indels at the target site as well as large deletions, translocations, chromosomal loss, chromothripsis, and activation of the p53 DNA damage response[13–18] (Supplementary References). Treatment of HSPCs with ABE8e-NRCH did not lead to a detected p53 response or large deletions, in contrast to treatment with Cas9 nuclease (Extended Data Fig. 10).

Third, base editing does not require DNA delivery, which is a requirement for gene therapy or homology-directed repair. The introduction of DNA can lead to toxicity and insertional mutagenesis (Supplementary References). By contrast, base editing using mRNA or RNP directly converts pathogenic HBB^S into a non-pathogenic allele with no requirement for exogenous DNA.

We examined potential undesired consequences of base editing HSPCs. Base editors can cause bystander editing of nearby nucleotides. In this study, we observed minimal (less than 2%) non-synonymous bystander edits as a result of careful positioning of the bespoke ABE at a CACC PAM[3] (Fig. 1a, Extended Data Fig. 1c). Spurious editing of RNA can occur[29] but is short-lived when base editor mRNA or RNP is used, and did not appear to affect repopulation, viability, or differentiation of HSCs. Off-target base editing can also occur[29], although on-target and off-target base editing in the same cell result in multiple point mutations that are less likely to be genotoxic than multiple DSBs from on- and off-target nuclease activity[29]. Of the 54 identified sites with observed off-target ABE8e-NRCH base editing in an extensive analysis of 697 computationally and experimentally nominated sites, we found no missense mutations or other off-target edits of anticipated consequence. Nevertheless, the safety and therapeutic potential of this approach might be further improved by testing alternative deaminase and Cas9 variants that have been shown to minimize Cas-dependent and Cas-independent off-target base editing, optimizing the dosage of the editing agent, or optimizing delivery methods[2] (Supplementary References). Although HBB^G is a naturally occurring benign variant, further studies are required to better understand the effects of this allele in combination with HBB^S (Supplementary Discussion).

The ex vivo delivery procedure used here resembles methods currently used for HSC editing in clinical trials[12]. The ABEs were electroporated as mRNA or RNP to minimize the duration of exposure to the editing agent, which reduces off-target editing compared to DNA delivery[2]. HSCs were edited using a single electroporation and transplanted into adult mice after 24 h to minimize the duration of in vitro culture and any associated loss of multipotency. The phenotypic rescue observed following secondary transplantation of varying proportions of edited and unedited bone marrow establishes that the observed base editing efficiencies substantially exceed the gene correction threshold needed for therapeutic benefit. Base-edited patient-derived $CD34^+$ cells thus provide a promising basis for a one-time autologous treatment for SCD.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03609-w.

1. Piel, F. B., Steinberg, M. H. & Rees, D. C. Sickle cell disease. *N. Engl. J. Med.* **377**, 305 (2017).
2. Richter, M. F. et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **38**, 883–891 (2020).
3. Miller, S. M. et al. Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nat. Biotechnol.* **38**, 471–481 (2020).
4. Sangkitporn, S., Rerkamnuaychoke, B., Sangkitporn, S., Mitrakul, C. & Sutivigit, Y. Hb G Makassar (β6:Glu-Ala) in a Thai family. *J. Med. Assoc. Thai.* **85**, 577–582 (2002).
5. Blackwell, R. Q., Oemijati, S., Pribadi, W., Weng, M. I. & Liu, C. S. Hemoglobin G Makassar: β6 Glu→Ala. *Biochim. Biophys. Acta* **214**, 396–401 (1970).
6. Wu, L. C. et al. Correction of sickle cell disease by homologous recombination in embryonic stem cells. *Blood* **108**, 1183–1188 (2006).
7. Leonard, A., Tisdale, J. & Abraham, A. Curative options for sickle cell disease: haploidentical stem cell transplantation or gene therapy? *Br. J. Haematol.* **189**, 408–423 (2020).
8. Magrin, E., Miccio, A. & Cavazzana, M. Lentiviral and genome-editing strategies for the treatment of β-hemoglobinopathies. *Blood* **134**, 1203–1213 (2019).
9. Zeng, J. et al. Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* **26**, 535–541 (2020).
10. Ribeil, J. A. et al. Gene therapy in a patient with sickle cell disease. *N. Engl. J. Med.* **376**, 848–855 (2017).
11. Esrick, E. B. et al. Post-transcriptional genetic silencing of *BCL11A* to treat sickle cell disease. *N. Engl. J. Med.* **384**, 205–215 (2021).
12. Frangoul, H. et al. CRISPR-Cas9 gene editing for sickle cell disease and β-thalassemia. *N. Engl. J. Med.* **384**, 252–260 (2021).

13. Zuccaro, M. V. et al. Allele-specific chromosome removal after Cas9 cleavage in human embryos. *Cell* **183**, 1650–1664.e15 (2020).
14. Song, Y. et al. Large-fragment deletions induced by Cas9 cleavage while not in the BEs system. *Mol. Ther. Nucleic Acids* **21**, 523–526 (2020).
15. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).
16. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **24**, 927–930 (2018).
17. Ihry, R. J. et al. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **24**, 939–946 (2018).
18. Enache, O. M. et al. Cas9 activates the p53 pathway and selects for p53-inactivating mutations. *Nat. Genet.* **52**, 662–668 (2020).
19. Wilkinson, A. C. et al. Cas9-AAV6 gene correction of beta-globin in autologous HSCs improves sickle cell disease erythropoiesis in mice. *Nat. Commun.* **12**, 686 (2021).
20. Dever, D. P. et al. CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature* **539**, 384–389 (2016).
21. Viprakasit, V., Wiriyasateinkul, A., Sattayasevana, B., Miles, K. L. & Laosombat, V. Hb G-Makassar [β6(A3)Glu→Ala; codon 6 (GAG→GCG)]: molecular characterization, clinical, and hematological effects. *Hemoglobin* **26**, 245–253 (2002).
22. Chu, S. H. et al. Rationally designed base editors for precise editing of the sickle cell disease mutation. *CRISPR J.* **4**, 169–177 (2021).
23. Bae, S., Park, J. & Kim, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).
24. Tsai, S. Q. et al. CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* **14**, 607–614 (2017).
25. McIntosh, B. E. et al. Nonirradiated NOD,B6.SCID *Il2rγ*$^{-/-}$ *Kit*$^{W41/W41}$ (NBSGW) mice support multilineage engraftment of human hematopoietic cells. *Stem Cell Reports* **4**, 171–180 (2015).
26. .Alanis-Lobato, G. et al. Frequent loss-of-heterozygosity in CRISPR-Cas9–edited early human embryos. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.2004832117 (2021).
27. Demirci, S. et al. βT87Q-globin gene therapy reduces sickle hemoglobin production, allowing for *ex vivo* anti-sickling activity in human erythroid cells. *Mol. Ther. Methods Clin. Dev.* **17**, 912–921 (2020).
28. Wu, Y. et al. Highly efficient therapeutic gene editing of human hematopoietic stem cells. *Nat. Med.* **25**, 776–783 (2019).
29. Anzalone, A. V., Koblan, L. W. & Liu, D. R. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **38**, 824–844 (2020).

Methods

HEK293T cell culture and editing

HEK293T cells (ATCC CRL-3216) modified to contain the sickle cell allele[3] were cultured in Dulbecco's modified Eagle's medium (Corning) supplemented with 10% fetal bovine serum (FBS) (ThermoFisher Scientific) and maintained at 37 °C with 5% CO_2. Cells were verified to be free of mycoplasma by PCR test in growth medium and confirmed to contain the SCD allele by HTS. Plasmid transfection of base editors into HEK293T cells has been previously described[30–32]. HEK293T cells were seeded for plasmid transfection at 20,000 cells per well on 96-well poly-D-lysine plates (Corning) in the same culture medium. Cells were transfected 24–30 h after plating with 0.5 µl Lipofectamine 2000 (ThermoFisher Scientific) using 200 ng base editor plasmid and 66 ng guide RNA plasmid according to the manufacturer's instructions. The plasmid encoding the new ABE8e-NRCH generated for this study has been deposited in AddGene (ID 165416). Cells were cultured for 3 days after lipofection, then washed with PBS (ThermoFisher Scientific). Genomic DNA was extracted after removal of PBS by addition of 50 µl freshly prepared lysis buffer (10 mM Tris-HCl, pH 7.5, 0.05% SDS, 25 µg/ml proteinase K (ThermoFisher Scientific)) directly into each transfected well. The mixture was incubated at 37 °C for 1 h then heat-inactivated at 80 °C for 30 min. One microlitre of this lysate was used as a PCR template for high-throughput sequencing.

High-throughput sequencing of the HBB SCD locus in HEK293T cells

High-throughput sequencing of genomic DNA was performed as previously described[30,31]. Primers for amplification of the *HBB* SCD locus in HEK293T cells were: GAN162F: 5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAGGGTTGGCCAATCTACTCCC-3′; GAN163R: 5′-TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTTCTCTGTCTCCACATGCC-3′. Underlined sequences represent adapters for Illumina sequencing. Following Illumina barcoding, PCR products were pooled and purified by electrophoresis with a 2% agarose gel using a Monarch DNA Gel Extraction Kit (New England Biolabs), eluting with 30 µl H_2O. DNA concentration was quantified using a Qubit dsDNA High Sensitivity Assay Kit (ThermoFisher Scientific) and sequenced on an Illumina MiSeq instrument (single-end read, 250–300 cycles) according to the manufacturer's protocols. Alignment of fastq files and quantification of editing frequency were performed using CRISPResso2[33] in batch mode with a window width of 34 nucleotides.

ABE8e-NRCH mRNA

ABE8e-NRCH mRNA was transcribed in vitro from the PCR product using full substitution of N1-methylpseudouridine for uridine. mRNA was capped co-transcriptionally using CleanCap AG analogue (TriLink Biotechnologies) resulting in a 5′ Cap 1 structure. The in vitro transcription reaction was performed as previously described[34] with the following changes; 16.5 mM magnesium acetate and 4 mM CleanCap AG were used as the final concentrations during transcription, and mRNAs were purified using RNeasy kit (QIAgen). Mammalian-optimized UTR sequences (TriLink) and a 120-base poly A tail were included in the transcribed PCR product.

ABE8e-NRCH protein

RNP delivery of genome editing agents has been previously described and established to decrease off-target editing activity compared to DNA delivery[2,35–39]. ABE8e-NRCH protein was codon optimized for bacterial expression and cloned into the protein expression plasmid pD881-SR (Atum, Cat. No. FPB-27E-269). This plasmid has been deposited on AddGene (ID# 165417). The expression plasmid was transformed into BL21 Star DE3 competent cells (ThermoFisher, Cat. No. C601003). Colonies were picked for overnight growth in terrific broth (TB) + 25 µg/ml kanamycin at 37 °C. The next day, 2 l pre-warmed TB was inoculated with overnight culture at a starting OD_{600} of 0.05. Cells were shaken at 37 °C for about 2.5 h until the OD_{600} was ~1.5. Cultures were cold-shocked in an ice-water slurry for 1 h, following which L-rhamnose was added to a final concentration of 0.8% to induce expression. Cultures were then incubated at 18 °C with shaking for 24 h to express protein. Following induction, cells were pelleted and flash-frozen in liquid nitrogen and stored at −80 °C. The next day, cells were resuspended in 30 ml cold lysis buffer (1 M NaCl, 100 mM Tris-HCl pH 7.0, 5 mM TCEP, 10% glycerol, with 5 tablets of cOmplete, EDTA-free protease inhibitor cocktail (Millipore Sigma, Cat. No. 4693132001). Cells were passed three times through a homogenizer (Avestin Emulsiflex-C3) at ~18,000 psi to lyse. Cell debris was pelleted for 20 min using 20,000*g* centrifugation at 4 °C. Supernatant was collected and spiked with 40 mM imidazole, followed by a 1 h incubation at 4 °C with 1 ml Ni-NTA resin slurry (G Bioscience Cat. No. 786-940, prewashed once with lysis buffer). Protein-bound resin was washed twice with 12 ml lysis buffer in a gravity column at 4 °C. Protein was eluted in 3 ml elution buffer (300 mM imidazole, 500 mM NaCl, 100 mM Tris-HCl pH 7.0, 5 mM TCEP, 10% glycerol). Eluted protein was diluted in 40 ml low-salt buffer (100 mM Tris-HCl, pH 7.0, 5 mM TCEP, 10% glycerol) just before loading into a 50-ml Akta Superloop for ion exchange purification on an Akta Pure25 FPLC. Ion exchange chromatography was conducted on a 5-ml GE Healthcare HiTrap SP HP pre-packed column (Cat. No. 17115201). After washing the column with low-salt buffer, we flowed the diluted protein through the column to bind. The column was then washed in 15 ml low-salt buffer before being subjected to an increasing gradient to a maximum of 80% high-salt buffer (1 M NaCl, 100 mM Tris-HCl, pH 7.0, 5 mM TCEP, 10% glycerol) over the course of 50 ml, at a flow rate of 5 ml per minute. One-millilitre fractions were collected during this ramp to high-salt buffer. Peaks were assessed using SDS–PAGE to identify fractions that containd the desired protein, which were concentrated first using an Amicon Ultra 15-ml centrifugal filter (100-kDa cutoff, Cat. No. UFC910024), followed by a 0.5-ml 100-kDa cutoff Pierce concentrator (Cat. No. 88503). Concentrated protein was quantified using a BCA assay (ThermoFisher, Cat. No. 23227).

Isolation and culture of $CD34^+$ human HSPCs

Circulating G-CSF-mobilized human mononuclear cells were obtained from de-identified healthy adult donors (Key Biologics, Lifeblood). Plerixafor-mobilized $CD34^+$ cells from patients with SCD were collected according to the protocol 'Peripheral Blood Stem Cell Collection for Sickle Cell Disease Patients' (ClinicalTrials.gov identifier NCT03226691), which was approved by the human subject research institutional review boards at the National Institutes of Health and St. Jude Children's Research Hospital. We complied with all relevant ethical regulations and all participants provided informed consent. $CD34^+$ cells were enriched by immunomagnetic bead selection using a CliniMACS Plus or AutoMACS instrument (Miltenyi Biotec). $CD34^+$ cells were maintained in stem cell culture medium: X-VIVO-10 (Lonza, BEBP02-055Q) medium supplemented with 100 ng/µl human SCF (R&D systems, 255-SC/CF), 100 ng/µl human TPO (R&D systems, 288-TP/CF) and 100 ng/µl human FLT-3 ligand (R&D systems, 308-FK/CF). Cells were seeded and maintained at a density of $0.5–1 \times 10^6$ cells per ml.

RNP and mRNA electroporation in human HSPCs

Electroporation was performed with an ATX MaxCyte electroporator using electroporation program HSC3. The modified synthetic sgRNA contained 2′-*O*-methyl modifications in the first three and last three nucleotides, and phosphorothioate bonds between the first three and last three nucleotides[40], and was purchased from BioSpring. $CD34^+$ HSPCs were thawed 48 h before electroporation. mRNA and sgRNA were mixed at a 1:1 weight ratio before electroporation. RNPs were formed at a 1:1.5 ratio of ABE and Makassar sgRNA, and incubated for 20 min at room temperature before electroporation. mRNA + sgRNA were electroporated at 200 µg/ml of mRNA; and RNP was electroporated

at a final concentration of 9 μM protein per reaction. We electroporated 20–40 million cells per ml in 100 μl Maxcyte Buffer in OC-100 cartridges for transplantation into NBSGW animals. Electroporated cells were recovered in stem cell culture medium composed of X-VIVO 10 medium including cytokines (FLT-3 ligand, SCF, and TPO). Cells were maintained in culture at a density of $0.5–1 \times 10^6$ per ml. Genomic DNA was extracted on culture day 6 using QuickExtract buffer (Lucigen Cat. No. QE09050) then analysed by HTS for editing efficiency.

High-throughput sequencing of the HBB SCD locus in blood cells

After editing, the *HBB* SCD locus was amplified from genomic DNA with oligonucleotide primers: Forward.LF: 5′-CTACACGACGCTCTTCCGATCTTGGCCAATCTACTCCCAGGAGCAGG-3′ and Reverse.LR: 5′-CAGACGTGTGCTCTTCCGATCTTCAAAGAACCTCTGGGTCCAAGGGT-3′. Underlined sequences represent adapters for Illumina sequencing. Following Illumina barcoding, PCR products were pooled and HTS was conducted using a MiSeq or MiniSeq (Illumina). Sequences were analysed by joining paired reads and analysing amplicons for indels or the desired test sequence using CRIS.py[41]. Indels were reported as the number of reads without the wild-type amplicon length.

Erythroid cell culture

Erythroid differentiation of $CD34^+$ cells was performed using a three-phase protocol[42,43]. Phase 1 (days 1–7): Iscove's modified Dulbecco's medium (IMDM; Thermo Fisher Scientific, 12440061) with 2% human blood type AB plasma (SeraCare, 1810-0001), 3% human AB serum (Atlanta Biologicals, S40110) 1% penicillin/streptomycin (Thermo Fisher Scientific, 15070063), 3 units/ml heparin (Sagent Pharmaceuticals, NDC 25021-401-02), 3 units/ml EPO (Amgen, EPOGEN NDC 55513-144-01), 200 μg/ml holo-transferrin (Millipore Sigma, T0665, 10 ng/ml human SCF (R&D systems, 255-SC/CF), and 1 ng/ml human interleukin IL-3 (R&D systems, 203-IL/CF). Phase 2 (days 8–14): phase 1 medium without IL-3. Phase 3 (days 15–21): phase 2 medium without SCF and with holo-transferrin concentration increased to 1 mg/ml. Cells were maintained daily at a density of 0.1×10^6 per ml (phase 1), 0.2×10^6 per ml (phase 2) and 1.0×10^6 per ml (phase 3)

Erythroblast maturation was monitored by immuno-flow cytometry for the cell surface markers CD235a (BD Pharmingen Cat. No. 559943, 1:100 dilution), CD49d (BioLegend Cat. No. 304304, 1:20 dilution), and BAND3 (gift from X. An, 1:100 dilution) (Supplementary Table 1).

Haemoglobin quantification

HPLC quantification of individual globin chains was performed using reverse-phase columns on a Prominence HPLC System (Shimadzu Corporation). The eluted proteins were identified by light absorbance at 220 nm using a diode array detector. For quantification of globin content from erythroid cells derived from in vitro differentiation of human $CD34^+$ cells, the relative amounts of different β-like globin proteins were calculated from the area under the 220-nm peak and normalized according to the DMSO control. They are expressed as a fraction of the total β-like globins including normal β (β^A), sickle β (β^S), Makassar β (β^G), γ, and δ-globin.

In vitro sickling assay

Erythroid cells were incubated with Hoechst 33342 (Millipore Sigma Cat. No. B2261, 1:1,000 dilution) for 20 min at 37 °C and the Hoechst-negative population was sorted using a SH800 (Sony Biotechnologies). Sorted cells ($0.5–1.0 \times 10^5$ cells) were seeded into 12- or 96-well plates with 1 ml or 0.1 ml of phase 3 ED medium under hypoxic conditions (2% oxygen) for 24 h. The IncuCyte S3 Live-Cell Analysis System (Sartorius) with a 20× objective was used to monitor cell sickling, with images captured after 8 h. The percentage of sickling was measured by manual counting of sickled cells versus normal cells on the basis of morphology. For each sickling assay, more than 300 cells per condition were counted by researchers blinded to that condition. For mouse transplantation studies, mouse blood was diluted (1:5,000) in RPMI medium and seeded in a 6-well plate with 3 ml RPMI medium before imaging in IncuCyte S3.

CIRCLE-seq off-target editing analysis

Genomic DNA from HEK293T cells was isolated using Gentra Puregene Kit (QIAGEN) according to the manufacturer's instructions. CIRCLE-seq was performed as previously described[24,44]. In brief, purified genomic DNA was sheared with a Covaris S2 instrument to an average length of 300 bp. The fragmented DNA was end repaired, A-tailed and ligated to a uracil-containing stem-loop adaptor, using the KAPA HTP Library Preparation Kit, PCR Free (KAPA Biosystems). Adaptor-ligated DNA was treated with Lambda Exonuclease (NEB) and *Escherichia coli* Exonuclease I (NEB) and then with USER enzyme (NEB) and T4 polynucleotide kinase (NEB). Intramolecular circularization of the DNA was performed with T4 DNA ligase (NEB) and residual linear DNA was degraded by Plasmid-Safe ATP-dependent DNase (Lucigen). In vitro cleavage reactions were performed with 250 ng Plasmid-Safe-treated circularized DNA, 90 nM Cas9-NRCH protein (purified using the method described above for ABE8e-NRCH), Cas9 nuclease buffer (NEB) and 90 nM synthetic chemically modified sgRNA (BioSpring), in a 100-μl volume. Cleaved products were A-tailed, ligated with a hairpin adaptor (NEB), treated with USER enzyme (NEB) and amplified by PCR with barcoded universal primers (NEBNext Multiplex Oligos for Illumina (NEB)), using Kapa HiFi Polymerase (KAPA Biosystems). Libraries were sequenced with 150-bp paired-end reads on an Illumina MiSeq instrument. CIRCLE-seq data analyses were performed using open-source CIRCLE-seq analysis software and default recommended parameters (https://github.com/tsailabSJ/circleseq).

CasOFFinder off-target editing analysis

Computational prediction of NRCH PAM-containing potential off-target sites with minimal mismatches relative to the intended target site (three or fewer mismatches overall, or two or fewer mismatches allowing G:U wobble base pairings with the guide RNA) was performed using CasOFFinder[23,45].

Targeted amplicon sequencing by rhAmpSeq

On- and off-target sites identified by CIRCLE-seq and CasOFFinder were amplified from genomic DNA from edited $CD34^+$ SCD cells or unedited control SCD donor cells using rhAMPSeq system (IDT), with primers listed in Supplementary Table 2. Sequencing libraries were generated according to the manufacturer's instructions and sequenced with 151-bp paired-end reads on an Illumina NextSeq instrument.

Quantification of base-editing efficiency at evaluated off-target sites

The A•T-to-G•C editing frequency for each position in the protospacer was quantified using CRISPRessoPooled (v2.0.41) with quantification_window_size 10, quantification_window_centre -10, base_editor_output, conversion_nuc_from A, conversion_nuc_to G. The genomic features of all off-target sites were initially annotated using HOMER (v4.10)[46]. HOMER does not offer high resolution at regions near junctions between annotations, so confirmed off-target sites were inspected individually and annotated using the NCBI Genome Data Viewer. Both HOMER annotations and annotation by inspection are shown in Supplementary Table 2. The editing frequency for each site was calculated as the ratio between the number of reads containing the edited base (that is, G) in a window from position 4 to 10 of each protospacer (where the NRCH PAM is positions 21–24) within which adenine base editing efficiency is typically maximal[29], and the total number of reads. To calculate the statistical significance of off-target editing for the ABE8e-NRCH mRNA or RNP treatments compared to control samples, we applied a χ^2 test for each of four samples (two donors, each with two replicates). The 2×2 contingency table was constructed using the number of edited reads and the number of unedited reads in treated and

control groups. The false discovery rate (FDR) was calculated using the Benjamini–Hochberg method. The 54 reported significant off-targets were called on the basis of: (1) FDR < 0.05 and (2) difference in editing frequency between treated and control >0.5% for at least one treatment. The custom code used to conduct off-target quantification and the statistical analysis is available to download at https://github.com/tsailabSJ/MKSR_off_targets.

Ethical approval for studies involving mice

All studies using mice were approved by the St. Jude Children's Research Hospital Institutional Animal Care and Use Committee under Protocol 579 entitled 'Genetic Models for the Study of Hematopoiesis'. Mice were maintained in the St. Jude Children's Research Hospital Animal Resource Center according to recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health.

Mouse experiments

No statistical methods were used to predetermine sample size. Recipient mice were randomly selected for transplantation cohorts. An animal facility staff member blinded to cohort identity determined which mouse would receive which cells. Mice were housed, fed and handled identically. Identification numbers were used to blind investigators to which conditions were assigned to each mouse. Assays were performed in reference to each ID number before then being identified as part of each experimental group.

Transplantation of gene-edited CD34$^+$ HSPCs into *NOD.Cg-Kit^{W-41J} Tyr+ Prkdc^{scid} Il2rg^{tm1Wjl}/ThomJ* (NBSGW) mice

NBSGW mice were purchased from The Jackson Laboratory (stock no. 026622). Cells were cultured 48 h after thawing before electroporation, then cultured for an additional 24 h before xenotransplanation. Minimization of time in culture is helpful to maintain repopulating stem cells[47]. Base-edited or control CD34$^+$ cells from donors with SCD were administered at a dose of 0.2×10^6 per mouse with intraperitoneal (IP) injection of 10 mg/kg busulfan (Busulfex; PDL BioPharm) 48 h before infusion[48] or at 0.5×10^6 per mouse with no busulfan preconditioning by tail-vein injection in female mice aged 7–9 weeks. Chimerism after transplantation was evaluated at 16–17 weeks in the bone marrow at the time of death. Cell lineage composition was determined in the bone marrow by using cell-type-specific antibodies (Supplementary Table 1), and lineages were analysed using the Attune NxT flow cytometer (ThermoFisher) and sorted using an Aria III cell sorter (BD Biosciences). The antibodies used were: anti-mouse CD45 (BD Pharmingen Cat. No. 561088, 1:40 dilution), anti-human CD45 (BD Horizon Cat. No. 564047, 1:20 dilution), anti-human CD33 (BD Biosciences Cat. No. 333946, 1:20 dilution), anti-human CD3 (BD Pharmingen Cat. No. 557832, 1:20 dilution), anti-human CD19 (BD Biosciences Cat. No. 349209, 1:20 dilution), anti-human CD34 (BD Pharmingen Cat. No. 561440, 1:20 dilution), anti-human CD235a (BD Pharmingen Cat. No. 551336, 1:20 dilution). CD34$^+$ HSPCs or CD235a$^+$ erythroblasts were isolated with magnetic beads, using the human-specific CD34 MicroBead Kit UltraPure (Miltenyi Biotec Inc., catalogue 130-100-453) and CD235a (glycophorin A) MicroBeads, human (Miltenyi Biotec Inc., catalogue 130-050-501).

Single cell RNA-seq to determine clonal editing outcomes in human CD235a$^+$ cells

CD235a$^+$ cells were sorted from the bone marrow of mice 16 weeks after xenotransplantation of patient-derived CD34$^+$ HSPCs into NBSGW mice using FACS. Sorted cells were than analysed using the Chromium Next GEM Single Cell 5′ Reagent Kit V2 (dual index) (10x Genomics 1000263) and sequenced on an Illumina NovoSeq according to the manufacturer's protocol. Reads were mapped to hg38 using cellranger (v5.0.1). Allelic editing was analysed using the 10x Genomics' Vartrix tool (v1.1.19) to identify the genotypes (A/A, A/G, or G/G) at the disease-causing nucleotide, chr11:5227002, with the following parameters: --out-variants --primary-alignments --umi --out-barcodes --ref-matrix -s coverage. Cells were filtered out if fewer than 100 reads mapped to the first exon of *HBB*. A cell was assigned to A/A or G/G if all reads contained the A allele or the G allele, respectively. A cell was assigned to A/G if at least 100 reads contained the A allele and 100 reads contained the G allele. The average number of reads per cell mapped to *HBB* per cell was 3,385. We used 339 cells from mouse 1 and 274 cells from mouse 2 in the analysis.

Base editing and transplantation of Townes Mouse SCD HSPCs

Townes SCD mice were purchased from The Jackson Laboratory (Stock no. 013071). This strain harbours the human α-globin locus (*HBA1*) in place of the orthologous mouse loci *Hba1* and *Hba2*, and the human γ-globin (*HBG1*) and β-globin (*HBB^S* or *HBB^A*) loci in place of the endogenous mouse loci *Hbb-b1* and *Hbb-b2*. Bone marrow mononuclear cells were obtained by flushing femurs, tibias, hip bones and humeri with IMDM (10% FBS) followed by RBC lysis (ACK Lysing Buffer). Mixed male and female lineage marker negative (Lin$^-$) cells enriched for HSPCs were purified by immuno-magnetic bead selection using the Mouse Lineage Cell Depletion Kit (Miltenyi, 130-090-858). Lin$^-$ cells were cultured in lineage negative medium: StemSpan SFEM supplemented with mSCF (100 ng/μl), mIL-3 (10 ng/μl), mIL-11 (100 ng/μl), hFLT3 ligand (100 ng/μl) and PenStrep (1×) for 24 h before base editing. Pilot studies indicated that mouse HSPCs were edited more efficiently with RNP than with base editor mRNA. Further optimization may reveal the reason for this or yield improved electroporation methods for mRNA. The ribonucleoprotein complex was generated by incubating ABE8e-NRCH with targeting sgRNA at concentrations of 2.25 μM base editor and 6.75 μM gRNA (1:3 ratio) in T buffer (total volume 50 μl) for 30 min at room temperature. Electroporation was performed using the ThermoFisher Neon Transfection System with 100-μl tips in buffer E2 at 1,700 pulse voltage, 20 pulse width, 1 pulse.

Following electroporation, cells were cultured overnight in lineage negative medium, followed by transplantation via tail-vein injection of 10^6 cells into lethally irradiated (1,125 cGy delivered as a single dose), 8- to 12-week-old female C57Bl/6 PepBoy (CD45.1) recipients. For analysis following transplantation, peripheral blood was collected from the retro-orbital sinus using heparinized micro-haematocrit capillary tubes at 6, 10, 14, and 16 weeks after transplantation to determine the fraction of engrafted donor cells and haemoglobin content. Mice were euthanized for necropsy 16 weeks after transplantation.

Complete blood counts (CBCs) were performed using a FORCYTE veterinary haematology analyser. CBC measurements were collected from untransplanted *HBB$^{A/S}$* mice at 4–6 months of age. Blood smears were prepared using modified Romanowsky methanolic staining and eosin and thiazin methods. Engraftment was determined by flow cytometry for mouse anti-CD45.1-PE (BD Pharmingen Cat. No. 553776, 1:50 dilution) and mouse anti-CD45.2-FITC (BE Pharmingen Cat. No. 561874, 1:50 dilution). Mouse illustrations were adapted from BioRender.com.

Colony-forming assay and analysis of clonal editing outcomes

Lin$^-$ cells were purified from the bone marrow of three mice 16 weeks after transplantation using immuno-magnetic bead selection with the Mouse Lineage Cell Depletion Kit (Miltenyi, 130-090-858). For each animal, 500 Lin$^-$ cells were plated in triplicate in methylcellulose (Stemcell Technologies, Methocult GF M3434) and incubated at 37 °C. After 12 days, 30–35 colonies per mouse were picked and washed in PBS before lysis. HTS analysis was performed on all colony lysates and allelic editing for each colony was classified on the basis of editing percentages.

Oxygen binding measurements

Haemoglobin–oxygen equilibrium curves (OECs), to determine the oxygen binding affinity of HbA, HbS and HbG, were obtained using the Hemox Analyzer (TCS Scientific, New Hope, PA). EDTA-treated mouse blood was added to the analysis buffer containing Hemox solution (pH 7.4 at 37 °C), Additive-A (BSA-20), and anti-foaming agent (AFA-25)

according to the manufacturer's protocol. Each sample was oxygenated at 37 °C using compressed air and then deoxygenated with compressed N_2, while being subjected to continuous dual-wavelength spectrophotometry to determine the oxyhaemoglobin:deoxyhaemoglobin ratio along with continuous measurement of the oxygen partial pressure. OECs and p50 values were generated by the TCS Hemox Analysis Software.

Secondary transplantation of Townes mouse SCD HSPCs

Whole bone marrow from a single female mouse that had received base-edited $HBB^{S/S}$ cells and from a female mouse that had received unedited $HBB^{S/S}$ cells were collected and RBCs were lysed using ACK Lysing Buffer. Varying ratios of cells from the two mice (0:100, 20:80, 40:60, 60:40, 80:20, and 100:0) were mixed and resuspended in PBS. Two million cells per recipient were injected into irradiated (1,125 cGy) female C57Bl/6 PepBoy (CD45.1) recipients. Three 8- to 12-week-old mice were injected for each ratio of bone marrow cells. For analysis following transplantation, peripheral blood was collected from the retro-orbital sinus using heparinized micro-haematocrit capillary tubes. CBCs were performed using a FORCYTE veterinary haematology analyser. Mouse illustrations were adapted from BioRender.com.

Cas9 nuclease purification

We transformed 3×NLS-SpCas9 plasmid[28] into BL21 (DE3) competent cells (MilliporeSigma, 702353) and grew the cells in TB medium at 37 °C until the density reached OD_{600} = 2.4–2.8. Cells were induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) per litre for 20 h at 20 °C. Cell pellets were lysed in 25 mM Tris, pH 7.6, 500 mM NaCl, 5% glycerol by homogenization and centrifuged at 20,000 rpm for 1 h at 4 °C. Cas9 was purified with Nickel-NTA resin and treated with TEV protease (1 mg TEV per 40 mg of protein) and benzonase (100 units/ml, Novagen 70664-3) overnight at 4 °C. Subsequently, Cas9 was purified using a size-exclusion column (Amersham Biosciences HiLoad 26/60 Superdex 200 17-1071-01) followed by a 5-ml SP–HP ion exchange column (GE 17-1151-01) according to the manufacturer's instructions. Proteins were dialysed in 20 mM Hepes buffer pH 7.5 containing 400 mM KCl, 10% glycerol, and 1 mM TCEP buffer. Contaminants were removed using a Toxin Sensor Chromogenic LAL Endotoxin Assay Kit (GenScript, L00350). Purified proteins were concentrated and filtered using Amicon ultrafiltration units with a 30-kDa MWCO (MilliporeSigma, UFC903008) and an Ultrafree-MC centrifugal filter (MilliporeSigma, UFC30GV0S). Protein fractions were further assessed using TGX stain-free 4–20% SDS–PAGE (Biorad, 5678093) and quantified by BCA assay.

Reverse transcription ddPCR to assess CDKN1 expression

Healthy donor $CD34^+$ HSPCs were electroporated with ABE8-NRCH mRNA + sgRNA targeting the *HBB* locus. In parallel, 3×NLS Cas9 nuclease RNP complexed with sgRNA targeting the BCL11A erythroid-specific enhancer was used to compare base editor to nuclease strategies. Cells electroporated without ABE8e-NRCH or 3×NLS Cas9 were used as a control (electroporation, no cargo), and as a separate control, cells were cultured without electroporation (not electroporated). mRNA and sgRNA were mixed at a 1:1 mass ratio before electroporation. RNPs were formed at a 1:1.5 ratio of Cas9 and sgRNA, and incubated for 20 min at room temperature to complex before electroporation. mRNA + sgRNA complexes were electroporated at 200 μg/ml of mRNA. Cas9 RNP was electroporated at a final concentration of 4 μM protein.

Guide RNA sequences were: *BCL11A*-targeting guide sequence, 5′-CUAACAGUUGCUUUUAUCAC-3′; healthy *HBB*-targeting guide sequence: 5′-UUCUCCUCAGGAGUCAGGUG-3′. After electroporation, RNA was extracted from 100K cells at several time points (0, 6, 12, 24, and 48 h) using the RNeasy Plus Mini Kit (QIAGEN; Cat. No. 74136) and RNA concentrations were determined using Nanodrop (Thermo Scientific). Genomic DNA was extracted on day 6 after electroporation using QuickExtract buffer and then analysed by HTS to determine editing efficiency.

One-step reverse transcription digital droplet PCR (RT-ddPCR) on extracted RNA was used to determine *CDKN1* (p21) mRNA expression levels, upregulation of which indicates an active p53-mediated DNA damage response[49]. Ribonuclease P/MRP subunit p30 (RPP30) was used as reference control. We mixed 3 ng RNA with reverse transcriptase, 300 mM DTT, and Supermix in a One-Step RT-ddPCR advanced kit for probes (Bio-Rad, Hercules, CA, USA), *CDKN1* primers/probe (Bio-Rad, 10031252; assay ID: dHsaCPE5052298), and RPP30 primers/probe (Bio-Rad, 10031255; assay ID: dHsaCPE5038241) according to the manufacturer's protocol. After making droplets using an Automated Droplet Generator (Bio-Rad), thermocycling was performed as follows: 50 °C for 60 min, 95 °C for 10 min, then 40 cycles consisting of 95 °C for 30 s followed by 55 °C for 1 min. Following the cycles, a final incubation was conducted at 98 °C for 10 min. Droplets were read using QX200 (Bio-Rad, 1864001) and data were analysed using QuantaSoft (Bio-Rad).

Assessing target site disruption using ddPCR

DNA from 100,000 $CD34^+$ HPSCs were extracted using an Agencourt DNAdvance kit (Beckman Coulter Cat. No. A48705) according to the manufacturer's instructions. DNA concentrations were determined by Nanodrop (Thermo Scientific). Digital droplet PCR was used to determine the change in abundance of target loci. We added 50–150 ng DNA to a reaction mixture containing ddPCR Supermix for Probes (Bio-Rad, 1863026), HindIII-HF (0.25 units/μl, New England BioLabs, R3104L), *ACTB* primers and probes (900 nM each primer, 250 nM probe; primers: ACTB-Forward: 5′-ACACTGTGCCCATCTAC-3′; ACTB-Reverse: 5′-AATGTCACGCACGATTTC-3′; probe: 5′-/5HEX/CGGGACCTG/ZEN/ACTGACTACCTCAT/3IABkFQ/-3′), and either *HBB* primers and probes (900 nM each primer, 250 nM probe, primers: HBB-Forward: 5′-GCCACACCCTAGGGTTG-3′; HBB-Reverse: 5′-GGGAAAATAGACCAATAGGCAG-3′; probe: 5′-AGGGC TGGGCATAAAAGTCAG-3′) or *BCL11A* primers and probes (900 nM each primer, 250 nM probe, primers: BCL11A-Forward: 5′-TCTTAGACATAACACACCAGG-3′; BCL11a-Reverse: 5′-GTCTGCCAGTCCTCTTC-3′; probe: 5′-TCAATACAACTTTGAAGCT AGTCTAGTG-3′) according to the manufacturer's protocol.

Forward primers for amplification of *HBB* and *BCL11A* also contained the Illumina adaptor at the 5′ end: ACACTCTTTCCCTACAC-GACGCTCTTCCGATCTNNNN. Reverse primers for amplification of *HBB* and *BCL11A* also contained the Illumina adaptor at the 5′ end: TGGAGTTCAGACGTGTGCTCTTCCGATCT. Primers and probes for *HBB* were positioned to distinguish the target from the homologous *HBD* gene. Droplets were generated using a QX200 Manual Droplet Generator (Bio-Rad, 186-4002). Digital droplet PCR was performed as follows: 95 °C for 10 min, then 50 cycles of 94 °C for 30 s, 59.5 °C for 2 min. Following the cycles, a final incubation was conducted at 98 °C for 10 min. Droplets were read by a QX200 Droplet Reader (Bio-Rad, 1864001) and data were analysed using QuantaSoft (Bio-Rad).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

HTS sequencing files can be accessed using the NCBI Sequence Read Archive (PRJNA725249).

Code availability

The code used to conduct off-target quantification and the statistical analysis is available at https://github.com/tsailabSJ/MKSR_off_targets.

30. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
31. Gaudelli, N. M. et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).
32. Huang, T. P., Newby, G. A. & Liu, D. R. Precision genome editing using cytosine and adenine base editors in mammalian cells. *Nat. Protoc.* **16**, 1089–1128 (2021).
33. Clement, K. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
34. Vaidyanathan, S. et al. Uridine depletion and chemical modification increase Cas9 mRNA activity and reduce immunogenicity without HPLC purification. *Mol. Ther. Nucleic Acids* **12**, 530–542 (2018).
35. Rees, H. A. et al. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat. Commun.* **8**, 15790 (2017).
36. Doman, J. L., Raguram, A., Newby, G. A. & Liu, D. R. Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* **38**, 620–628 (2020).
37. Zuris, J. A. et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* **33**, 73–80 (2015).
38. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* **19**, 770–788 (2018).
39. Vakulskas, C. A. et al. 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 (2018).
40. Hendel, A. et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* **33**, 985–989 (2015).
41. Connelly, J. P. & Pruett-Miller, S. M. CRIS.py: a versatile and high-throughput analysis program for CRISPR-based genome editing. *Sci. Rep.* **9**, 4194 (2019).
42. Hu, J. et al. Isolation and functional characterization of human erythroblasts at distinct stages: implications for understanding of normal and disordered erythropoiesis in vivo. *Blood* **121**, 3246–3253 (2013).
43. Traxler, E. A. et al. A genome-editing strategy to treat β-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat. Med.* **22**, 987–990 (2016).
44. Lazzarotto, C. R. et al. Defining CRISPR-Cas9 genome-wide nuclease activities with CIRCLE-seq. *Nat. Protoc.* **13**, 2615–2642 (2018).
45. Hwang, G. H., Kim, J. S. & Bae, S. Web-based CRISPR toolkits: Cas-OFFinder, Cas-Designer, and Cas-Analyzer. *Methods Mol. Biol.* **2162**, 23–33 (2021).
46. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
47. Kumar, S. & Geiger, H. HSC niche biology and HSC expansion ex vivo. *Trends Mol. Med.* **23**, 799–819 (2017).
48. Leonard, A. et al. Low-dose busulfan reduces human CD34$^+$ cell doses required for engraftment in c-kit mutant immunodeficient mice. *Mol. Ther. Methods Clin. Dev.* **15**, 430–437 (2019).
49. Karimian, A., Ahmadi, Y. & Yousefi, B. Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage. *DNA Repair* **42**, 63–71 (2016).
50. Kim, H. S., Jeong, Y. K., Hur, J. K., Kim, J. S. & Bae, S. Adenine base editors catalyze cytosine conversions in human cells. *Nat. Biotechnol.* **37**, 1145–1148 (2019).

Acknowledgements We thank the SCD patients who contributed samples for this study; D. Gao and other members of our laboratories for discussions; and M. O'Reilly for assistance with Adobe Illustrator. This work was supported by US National Institutes of Health awards U01 AI142756, RM1 HG009490, R01 EB031172, R35 GM118062 (D.R.L.), R01R01HL156647 (M.J.W. and D.R.L), U01AI157189 (S.Q.T.), and P01 HL053749 (M.J.W. and S.Q.T.), the Bill and Melinda Gates Foundation (D.R.L.), the Howard Hughes Medical Institute (D.R.L.), the St. Jude Collaborative Research Consortium (D.R.L, S.M.P.-M., S.Q.T and M.J.W.), the Doris Duke Foundation (for aspects of this study that did not use mice; M.J.W., S.Q.T. and A.S.), the American Lebanese Syrian Associated Charities (A.S. and M.J.W.) and the Assisi Foundation of Memphis (M.J.W.). G.A.N. was supported by a Helen Hay Whitney Postdoctoral Fellowship and the HHMI. A.S. was supported by a Scholar Award from the American Society of Hematology. L.W.K. and K.A.E. acknowledge NSF GRFP fellowships. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions G.A.N., J.S.Y., K.J.W., T.M., C.R.L., Y.L., H.S.-T., S.N.P., Y.Y., K.M., K.A.E., Y.J., C.J.P., E.T., C.L., and A.S. conducted experiments and analysed data. J.M.H., M.F.R., K.T.Z., S.M.M., T.W., L.W.K., and J.F.T. prepared materials and provided conceptual assistance, G.A.N., J.S.Y., K.J.W., M.J.W., and D.R.L. wrote the manuscript with input from all authors. J.S.Y., A.P.M., T.A.K., S.Q.T., S.M.P.-M., M.J.W., and D.R.L. supervised this study.

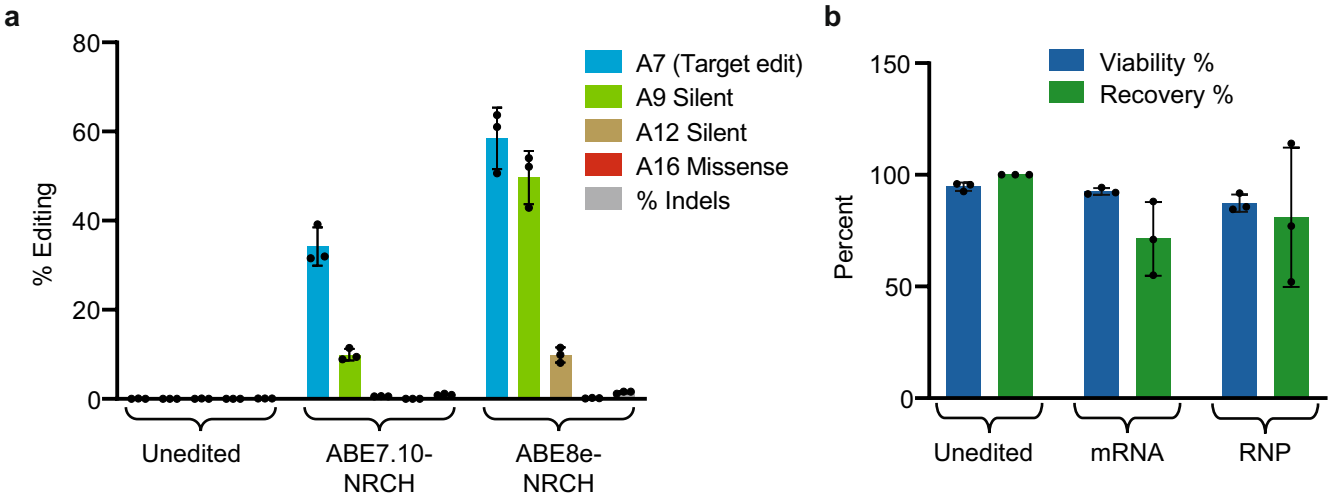
Competing interests G.A.N., K.A.E., M.F.R., K.T.Z., S.M.M., T.W., L.W.K., and D.R.L. have filed patent applications on aspects of base editing through the Broad Institute. D.R.L. is a consultant and equity owner of Beam Therapeutics, Prime Medicine, and Pairwise Plants (companies that use genome editing). M.J.W. is on advisory boards for Cellarity Inc., Novartis, and Forma Therapeutics, and is an equity owner of Beam Therapeutics. A.S. is a consultant for Spotlight Therapeutics and his institution receives clinical trial support for the conduct of sickle cell disease gene editing trials from Vertex Pharmaceuticals, CRISPR Therapeutics, and Novartis. J.S.Y. is an equity owner of Beam Therapeutics. The authors declare no non-financial competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03609-w.

Correspondence and requests for materials should be addressed to J.S.Y., M.J.W. or D.R.L.

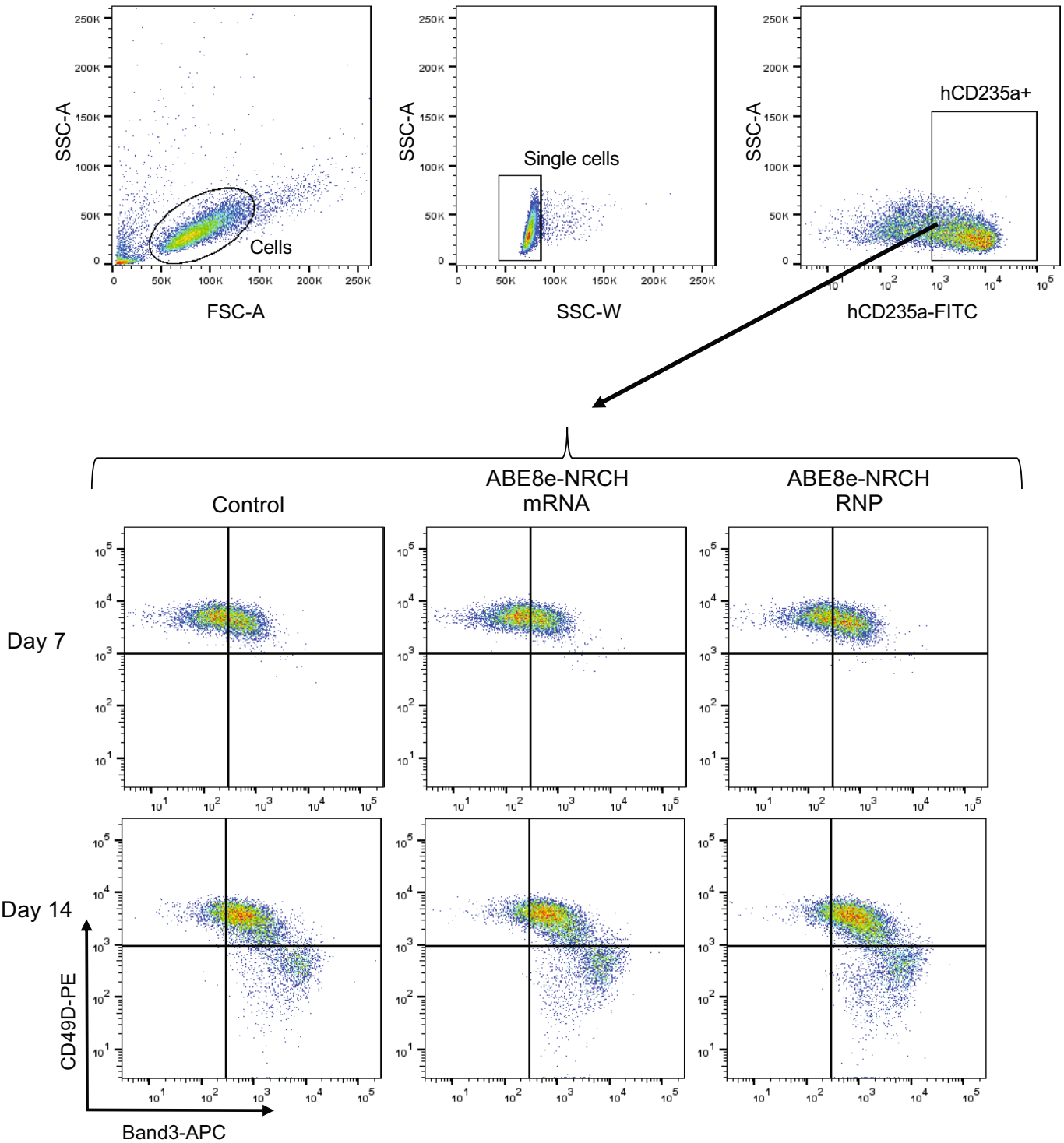
Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.



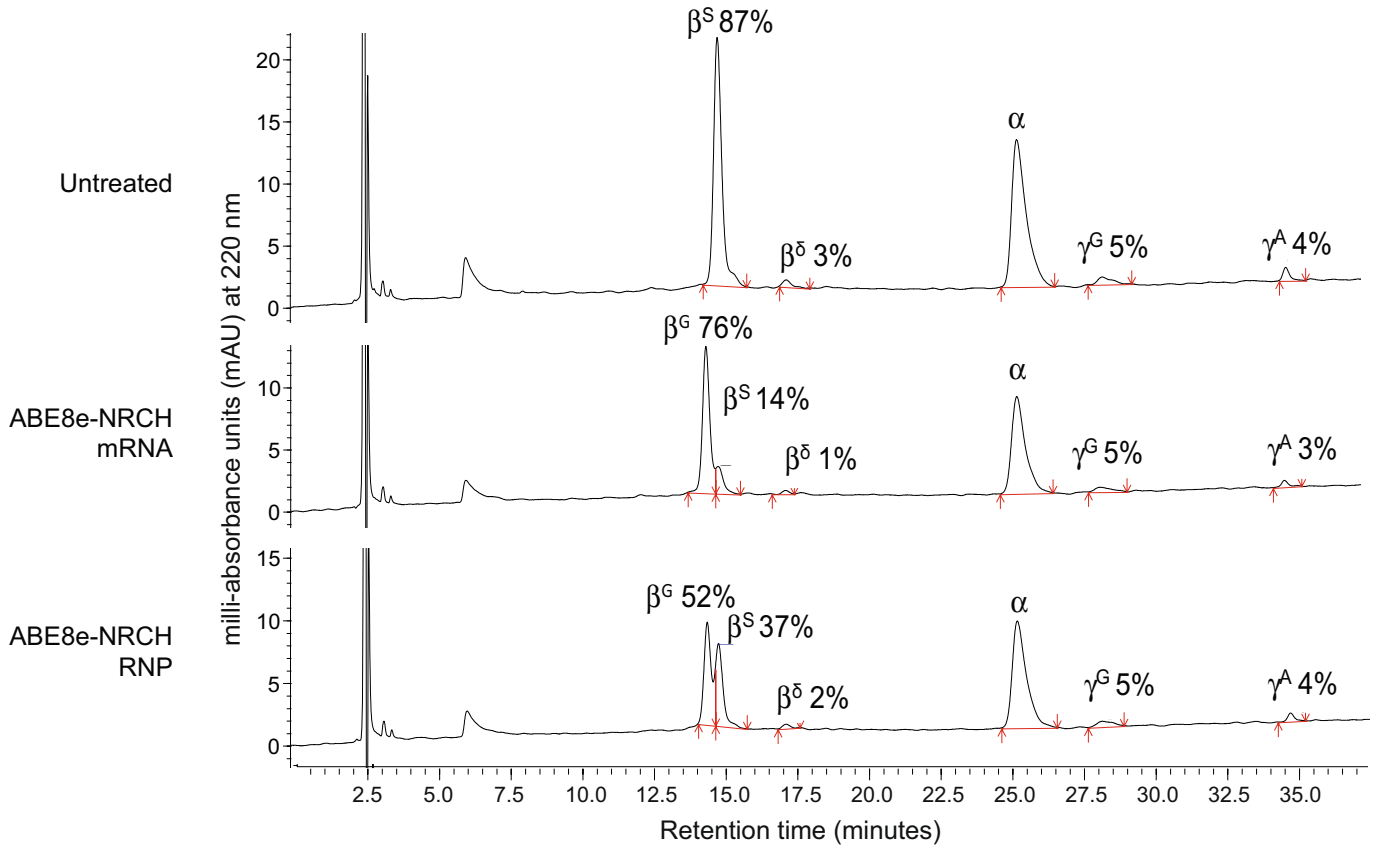
c

Allelic outcome (Protospacer shown)	Coding change	ABE8e-NRCH mRNA	ABE8e-NRCH RNP	Untreated
TTCTCCGCGGGAGTCAGGTG	V6A	66.92% ±1.82%	34.10% ±3.84%	0.00% ±0.00%
TTCTCCACAGGAGTCAGGTG	None	19.47% ±1.88%	54.19% ±5.42%	96.07% ±0.47%
TTCTCCGCAGGAGTCAGGTG	V6A	6.54% ±0.19%	6.13% ±0.97%	0.22% ±0.02%
TTCTCCGCGGGGTCAGGTG	V6A	3.21% ±0.10%	2.03% ±0.84%	0.00% ±0.00%
TTCTTCGCGGAGTCAGGTG	V6A, E7K	0.80% ±0.19%	0.22% ±0.09%	0.00% ±0.00%
TTCTCTGCGGAGTCAGGTG	V6A	0.57% ±0.09%	0.29% ±0.03%	0.00% ±0.00%
TTCTCCGCGGAGTCAGGTG (position -2 A>G)	V6A, S9P	0.46% ±0.03%	0.15% ±0.11%	0.00% ±0.00%
TTCTCAGCGGAGTCAGGTG	V6A	0.34% ±0.09%	0.00% ±0.00%	0.00% ±0.00%
TTCTCCGCAGGGGTCAGGTG	V6A	0.27% ±0.07%	0.08% ±0.05%	0.00% ±0.00%
TTCTCCACGGGAGTCAGGTG	None	0.34% ±0.08%	0.46% ±0.24%	0.25% ±0.04%
TTCTGCGCGGGAGTCAGGTG	V6A, E7Q	0.26% ±0.05%	0.04% ±0.06%	0.00% ±0.00%

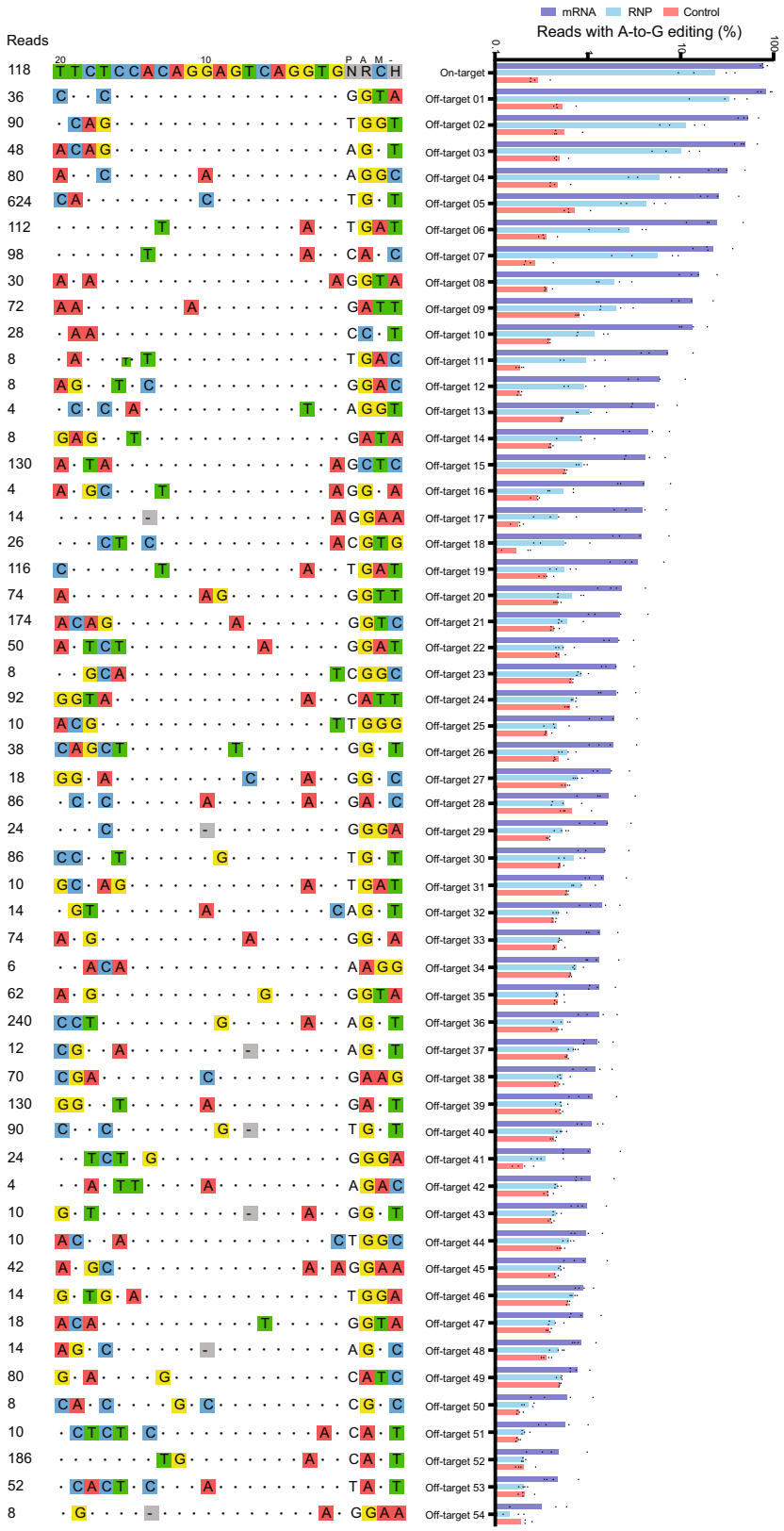
Extended Data Fig. 1 | Optimization in HEK293T cells, viability and recovery following human SCD HSPC editing, and allelic editing outcomes. **a**, Plasmids encoding the *HBB*^{S}-targeting sgRNA and either ABE7.10-NRCH or ABE8e-NRCH were transfected by lipofection into HEK293T cells. Editing efficiency was measured after 3 days by HTS. Unedited cells were not lipofected. **b**, Two days after electroporation into human SCD HSPCs of base editor mRNA and sgRNA, or electroporation of ribonucleoprotein (RNP), cell number and viability were measured using a Chemometec Nucleocounter-3000. Acridine orange was used to stain the total cell number and DAPI was used to stain dead, permeabilized cells. The percentage viability was calculated as the DAPI-stained cells divided by the acridine orange cells within each sample. The percentage recovery was normalized to the cell count of the unedited sample. Unedited cells were not electroporated. **c**, Six days after electroporation of SCD HSPCs, genomic DNA was extracted and the target *HBB* locus was PCR amplified and sequenced using an Illumina instrument. The sequencing analysis program CRIS.py was used to identify and quantify the resulting alleles. All alleles above a threshold of 0.2% frequency are shown. Below this threshold, variant alleles appeared with greatest frequency in the untreated control sample, suggesting they do not arise from base editor treatment. Nucleotides altered from the endogenous sequence are shown in blue. Rare cytosine base editing was observed at a frequency of less than 1%, as has been previously described as a possible outcome from adenine base editing[50]. Data shown as mean ± s.d., $n = 3$.



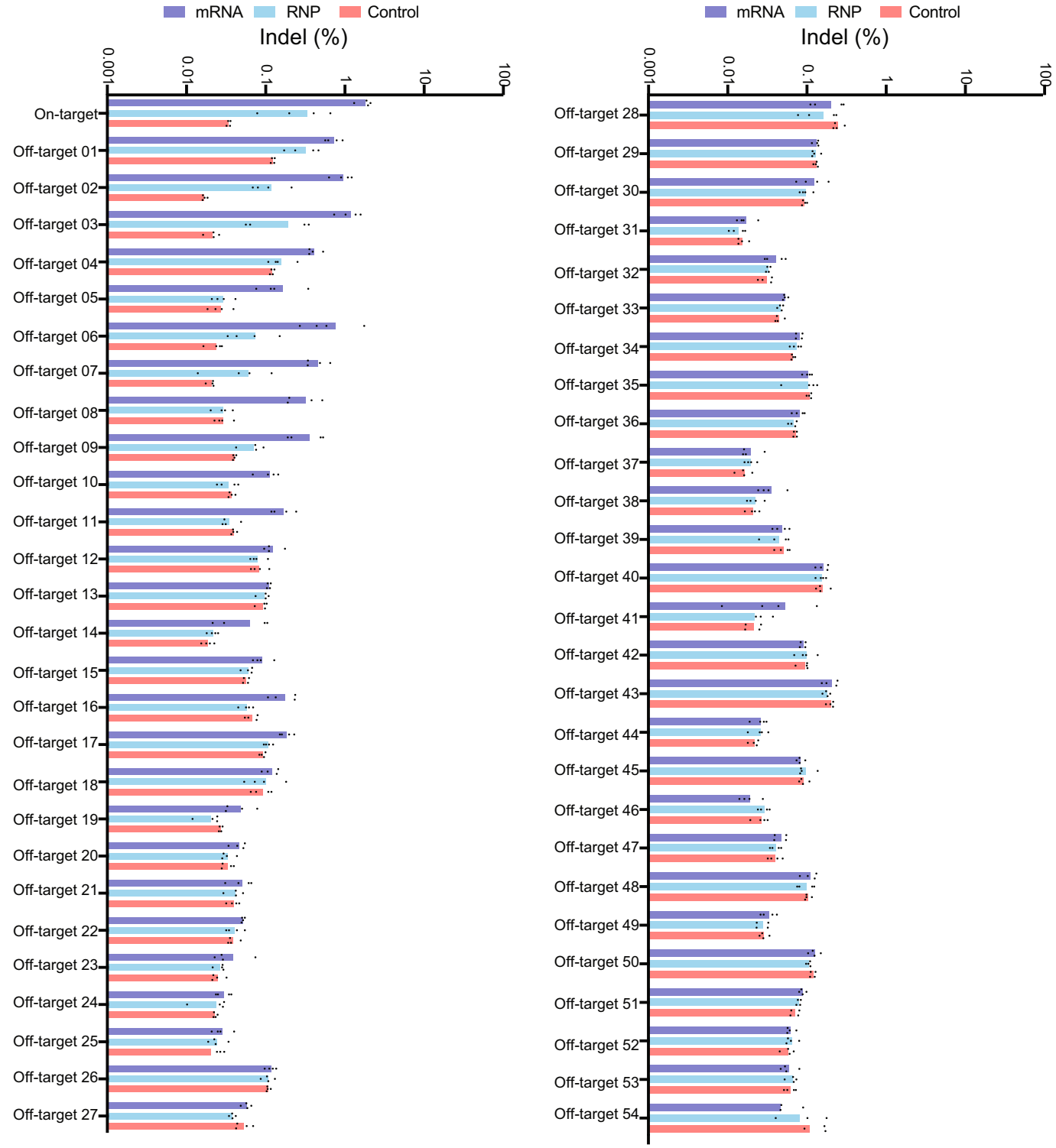
Extended Data Fig. 2 | Erythroid differentiation of edited SCD CD34$^+$ HSPCs. Representative, immuno-flow cytometry for erythroid maturation stage markers[42,43] at culture days 7 and 14. Top, gating strategy to identify single cells expressing the erythroid marker hCD235a. Bottom, gating strategy to track the progress of erythroid maturation based on expression of CD49D and BAND3 in hCD235a$^+$ cells. SSC-A, side scatter area; SSC-W, side scatter width; FSC-A, forward scatter area.



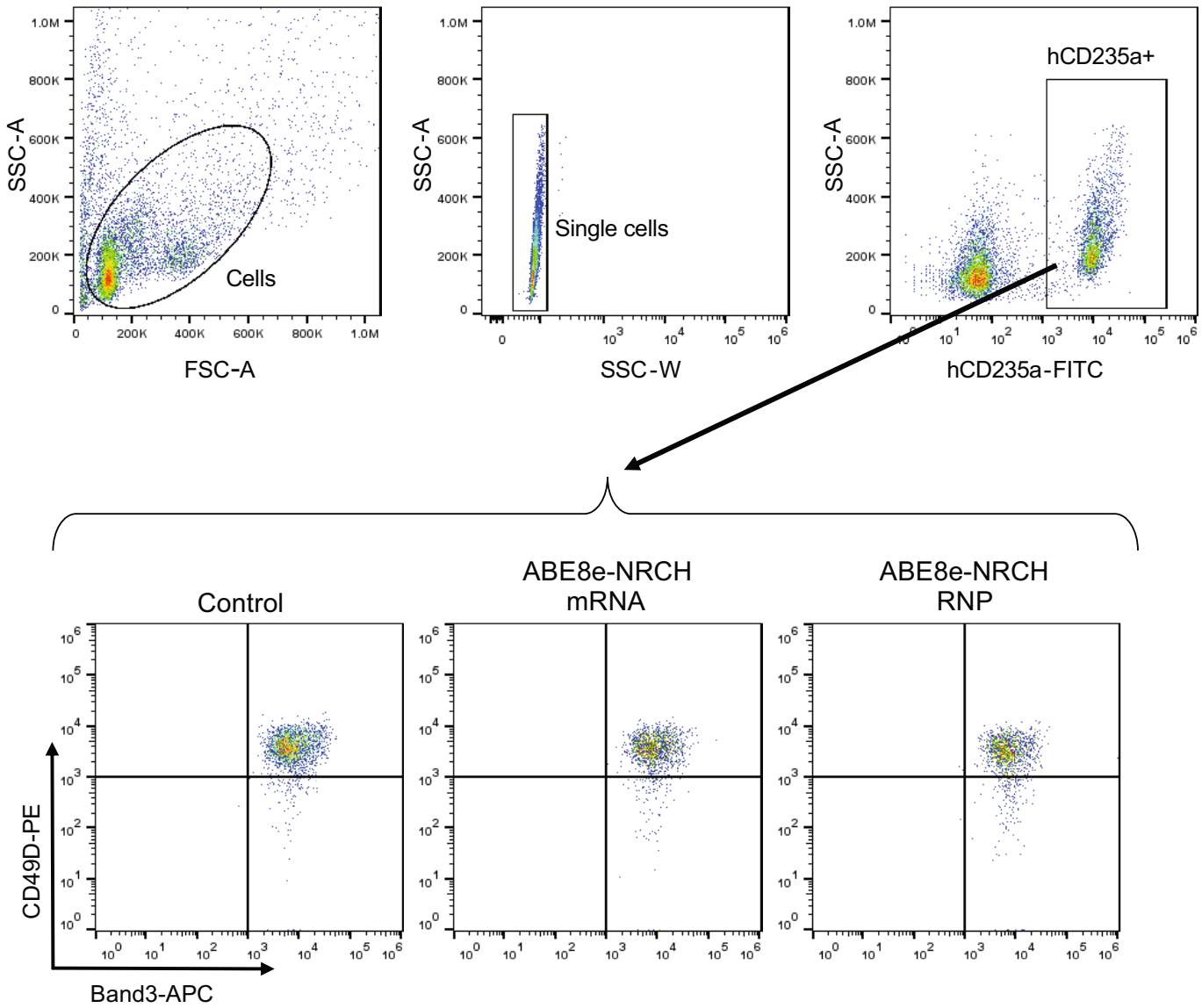
Extended Data Fig. 3 | Reverse-phase HPLC analysis of erythroid cells derived from in vitro differentiation of edited SCD CD34⁺ HSPCs. Reverse-phase HPLC chromatograms of erythroid cell lysates at culture day 18, with β-like globins and their associated fractions marked near the associated peak. Data from the most efficiently edited donor cells are shown. Red arrows indicate the start and end of globin chain peaks.



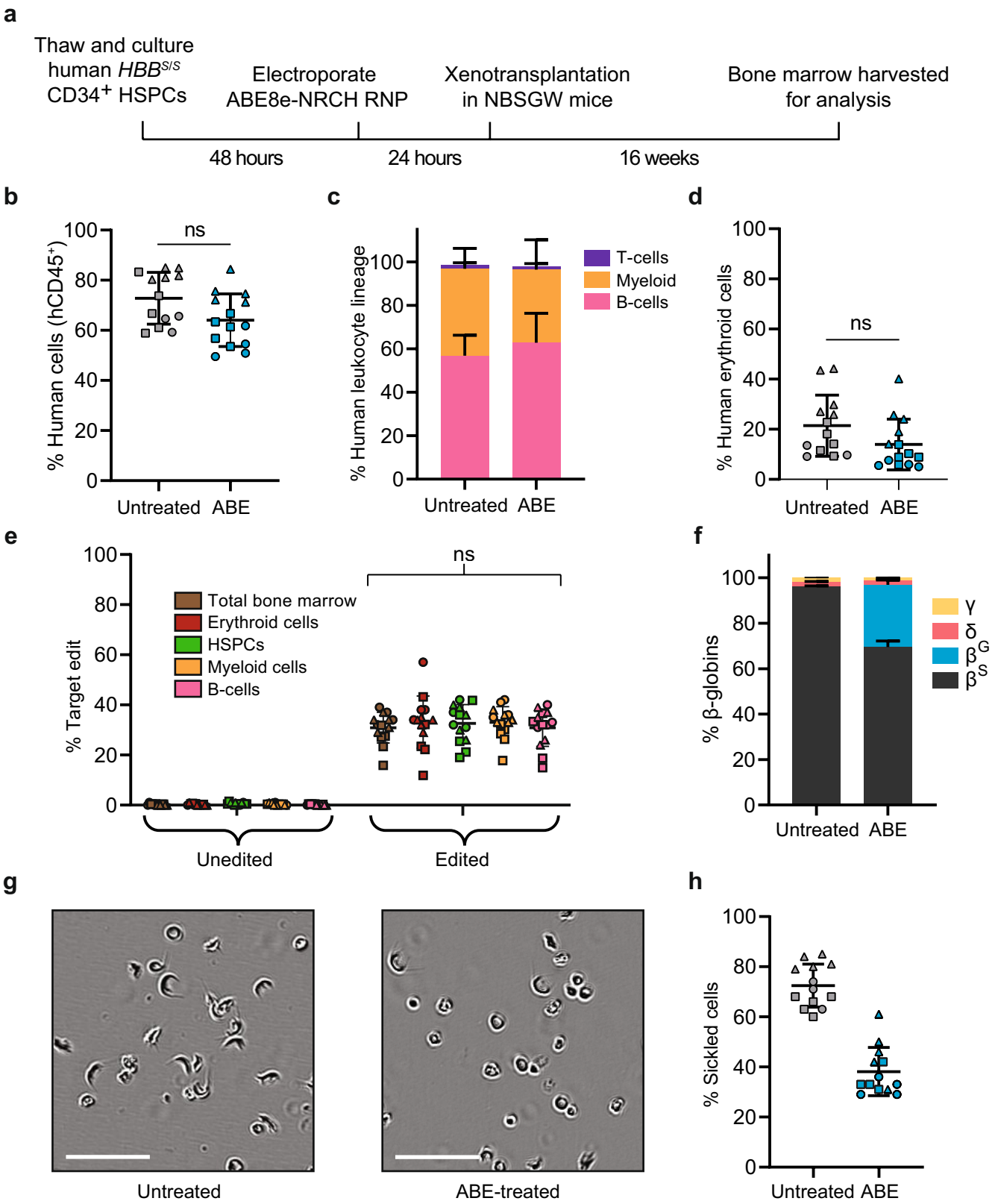
Extended Data Fig. 4 | Off-target base editing associated with ABE8e-NRCH conversion of *HBB*^{S} to *HBB*^{G} Makassar in SCD CD34^{+} HSPCs. CIRCLE-seq read counts obtained for each verified off-target site and the alignment of each site to the guide sequence are shown. Bar graphs show the percentage of sequencing reads containing A•T-to-G•C mutations within protospacer positions 4–10 at on- and off-target sites in genomic DNA samples from patient CD34^{+} HSPCs treated with ABE8e-NRCH mRNA, protein, or untreated controls ($n = 4$). Note that the mutation frequency shown is summed across all reads with one or more A•T-to-G•C mutations in this window. Sequencing errors therefore accumulate in control samples compared to standard sequencing error frequencies for a single nucleotide. Data shown as mean ± s.d.



Extended Data Fig. 5 | Off-target indel formation associated with ABE8e-NRCH conversion of *HBB*^{S} to *HBB*^{G} Makassar in SCD CD34^{+} HSPCs. Bar graph showing the percentage of sequencing reads containing alleles harbouring indels at on- and off-target sites in genomic DNA samples from patient CD34^{+} HSPCs treated with ABE8e-NRCH mRNA, protein, or untreated controls ($n = 4$). Data shown as mean ± s.d.

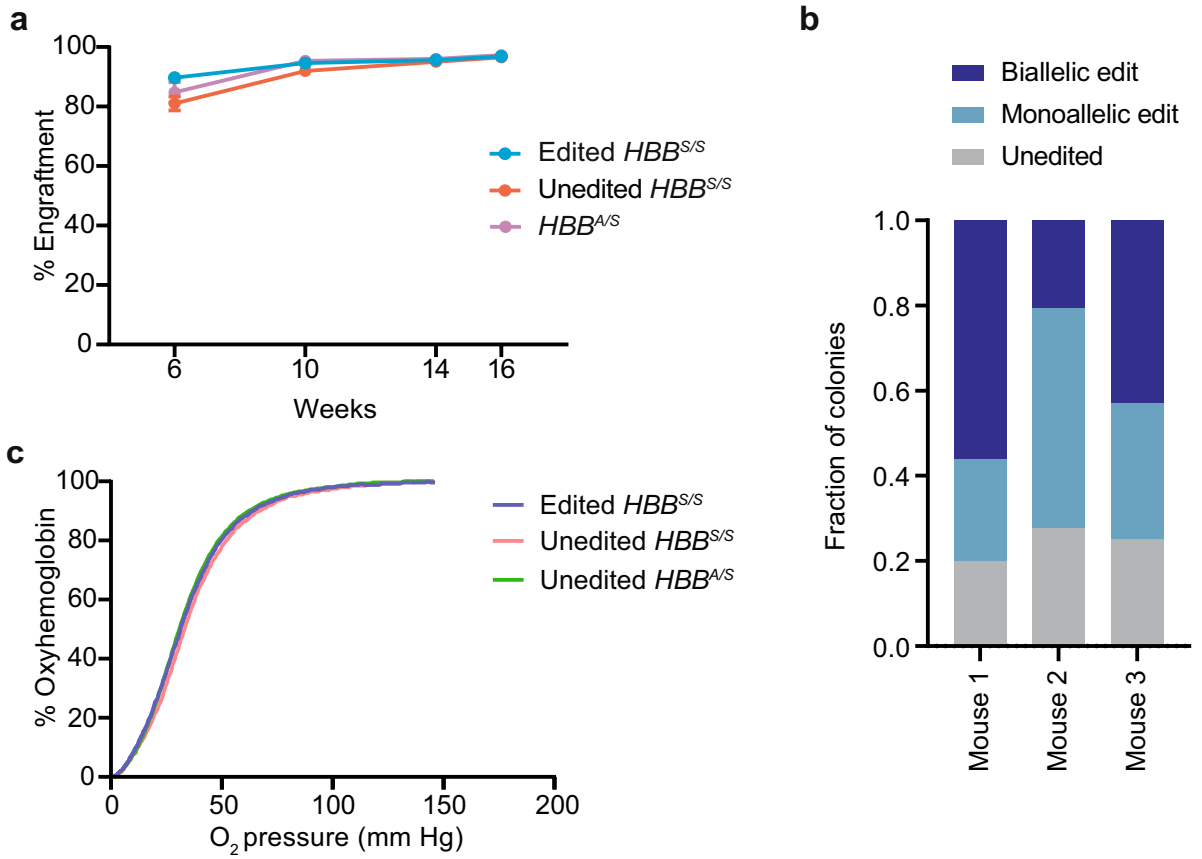


Extended Data Fig. 6 | Flow cytometry analysis of human donor-derived erythroid CD235a$^+$ cells after transplantation. Human CD235a$^+$ erythroid cells were purified by immuno-magnetic bead selection and analysed by flow cytometry for the indicated erythroid maturation markers[42,43].

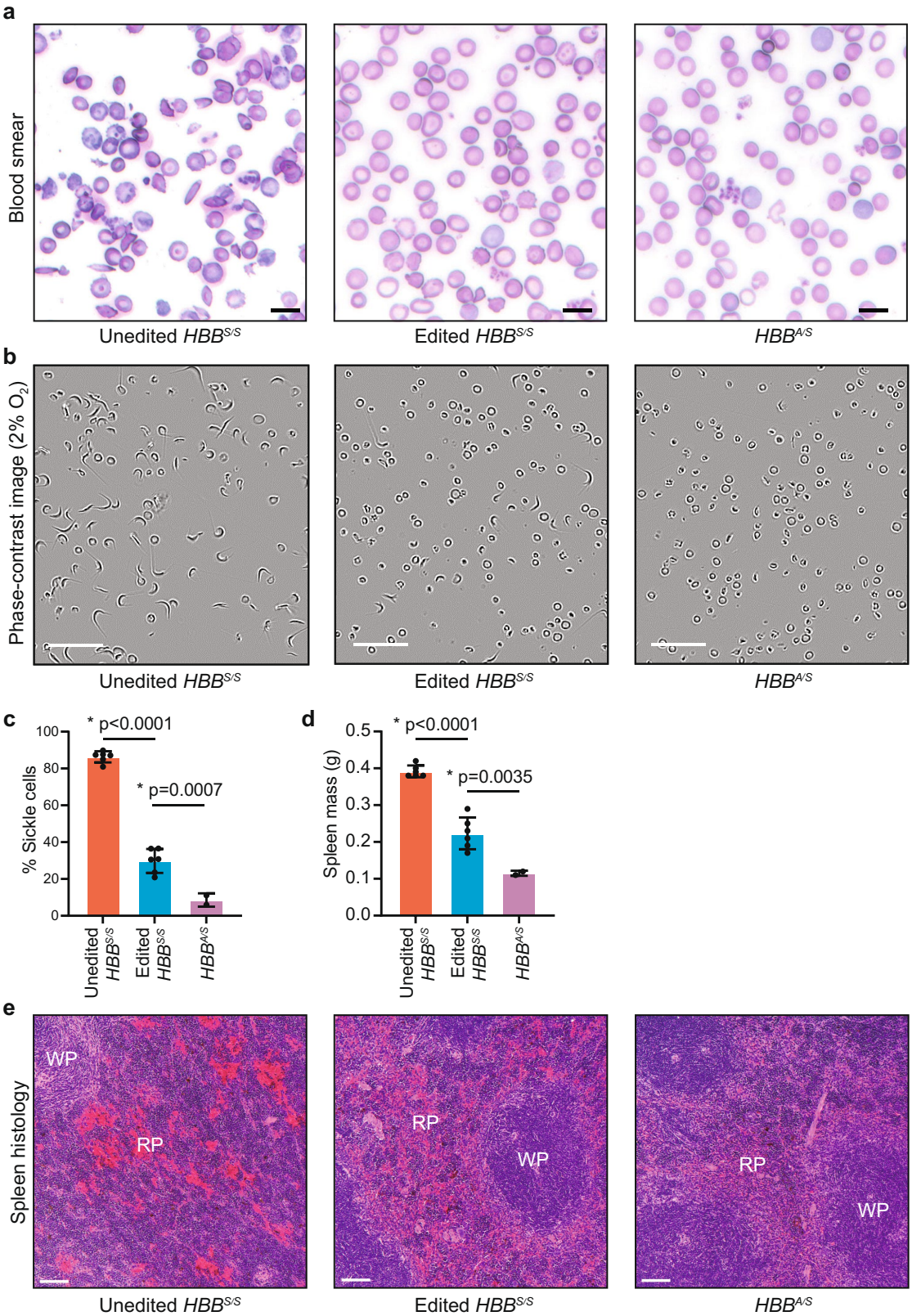


Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Engraftment of ABE8e-NRCH RNP-treated SCD CD34$^+$ HSPCs after transplantation into immunodeficient mice. CD34$^+$ HSPCs from three *HBB*$^{S/S}$ patients with SCD were electroporated with ABE8e-NRCH RNP using an sgRNA targeting the SCD mutant codon, followed by transplantation of 2–5 × 10^5 treated cells into NBSGW mice via tail-vein injection. Mice were euthanized and analysed 16 weeks after transplantation. **a**, Experimental workflow. **b**, Engraftment measured by the percentage of human donor CD45$^+$ cells (hCD45$^+$ cells) in recipient mouse bone marrow. **c**, Human B cells (hCD19$^+$), myeloid cells (hCD33$^+$), and T cells (hCD3$^+$) in recipient mouse bone marrow, shown as percentages of the total hCD45$^+$ population. **d**, Human erythroid precursors (hCD235a$^+$) in recipient mouse bone marrow, shown as a percentage of total human and mouse CD45$^-$ cells. **e**, On-target (A7, Fig. 1a) editing efficiencies in human donor CD34$^+$ cell-derived lineages purified from recipient bone marrow by FACS. Erythroid, myeloid, B cell, and HSPC human lineages were collected using antibodies against hCD235a, hCD33, hCD19, and hCD34, respectively. Statistical significance was assessed by one-way ANOVA to compare groups; ns, not significant. **f**, Percentages of β-like globin proteins determined by reverse-phase HPLC analysis of human donor-derived reticulocytes isolated from recipient mouse bone marrow. **g**, Representative phase contrast images of human reticulocytes purified from bone marrow and incubated for 8 h with 2% oxygen. Nine images of more than 50 cells per image were collected per sample. Scale bars, 50 µm. **h**, Quantification of sickled cells calculated by counting images after incubation for 8 h in 2% oxygen as in **g**. More than 300 randomly selected cells per sample were counted by a blinded observer. $n = 14$ total mice analysed (**b**–**f**); triangle, square, and circle symbols represent samples from three different donors with SCD. Negative control data are shared with Fig. 2. Data shown as mean ± s.d. Statistical significance between treated and untreated samples was assessed using two-tailed Student's *t*-test.

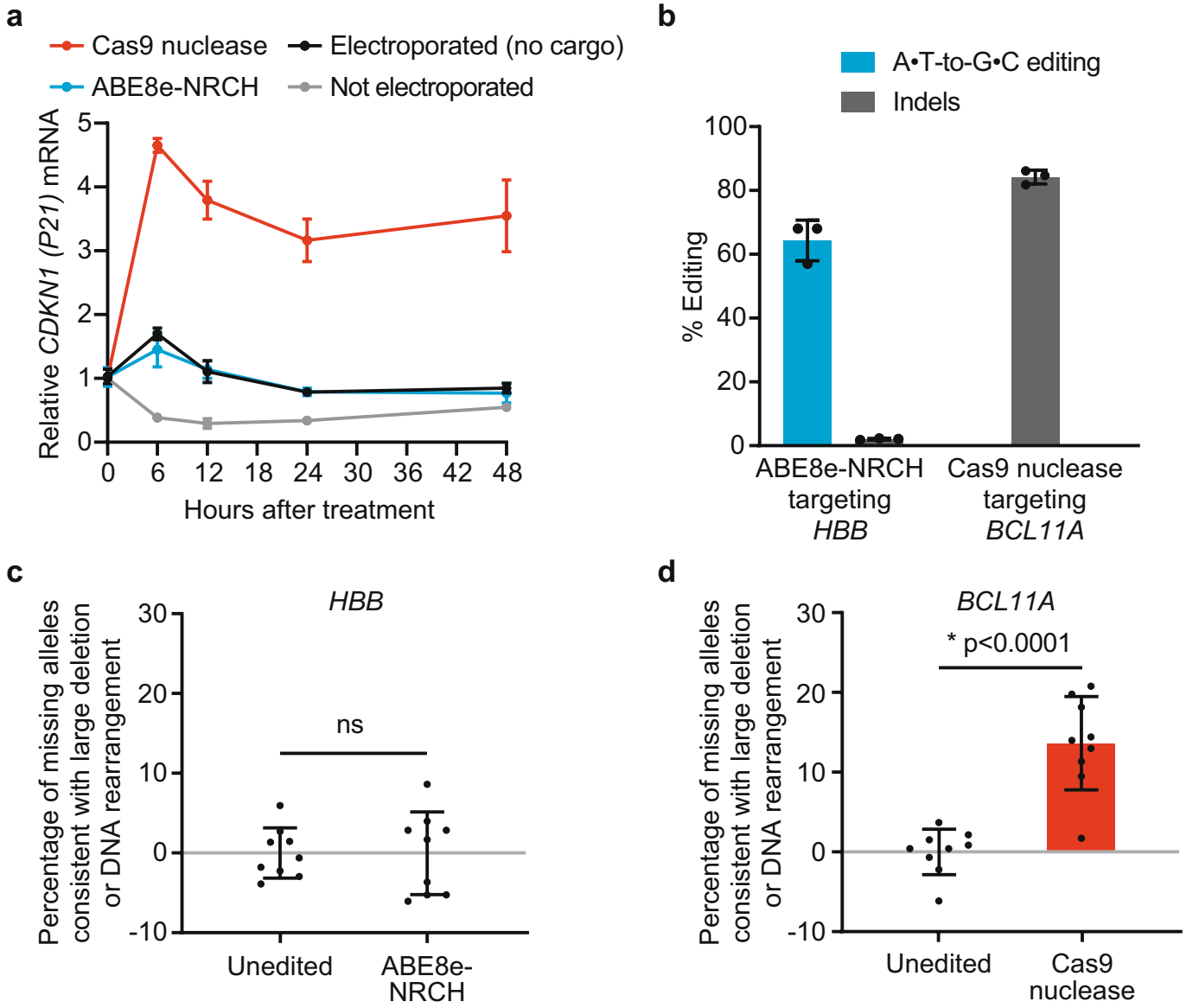


Extended Data Fig. 8 | Engraftment of transplanted Townes mouse HSPCs, clonality of editing outcomes, and oxygen binding affinity of blood. **a**, Donor cell engraftment measured by flow cytometry assessing the percentage of CD45.2$^+$ cells among PBMCs. **b**, Bone marrow from three mice transplanted with edited Townes mouse HSPCs was plated at low density in methylcellulose. After 12 days of culture, 30 to 35 individual colonies per mouse were picked into cell lysis buffer and the edited locus was amplified by PCR and sequenced by HTS. Colonies were categorized by whether they contained no editing, a monoallelic edit, or a biallelic edit. **c**, Blood was drawn from mice at week 14 after transplantation. Haemoglobin oxygenation was measured using a Hemox Analyzer (TCS Scientific) across a continuous declining gradient of oxygen pressure to assess whether *HBB*^{S}-to-*HBB*^{G} editing led to altered haemoglobin–oxygen binding. Data shown as mean ± s.d.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Adenine base editing of the SCD β-globin allele (*HBB*^{S}) to the Makassar variant (*HBB*^{G}) reduces erythrocyte sickling and splenic pathology in mice. Mice were treated as described in Fig. 3a. Blood and spleen were analysed 16 weeks after transplantation of Lin^{-} mouse HSPCs containing human *HBB* alleles. **a**, Representative images of blood smears. One blood smear image was collected per mouse. Scale bars, 25 µm. **b**, Representative phase contrast images of peripheral blood incubated for 8 h with 2% oxygen. Nine images of more than 50 cells per image were collected per sample. Scale bars, 50 µm. **c**, Quantification of sickled cells. More than 300 randomly selected cells per condition were counted by a blinded observer. **d**, Mass of dissected spleens. **e**, Histological sections of spleens of recipient mice 16 weeks after transplantation. Splenic pathologies in mice that received unedited donor *HBB*$^{S/S}$ HSCs include excessive extramedullary erythropoiesis and vascular congestion indicated by RBC pooling (bright red colour) resulting in expansion of red pulp (RP), reduction in white pulp (WP), and splenomegaly. Images were taken at 10× magnification and were processed, stained and photographed at the same time under identical conditions. Three images of each spleen were collected from different parts of the organ for each mouse. Scale bars, 100 µm. Unedited *HBB*$^{S/S}$, $n = 6$ mice; edited *HBB*$^{S/S}$, $n = 6$ mice; *HBB*$^{A/S}$, $n = 2$ mice. Data shown as mean ± s.d., with individual values as dots. Statistical significance was assessed using one-way ANOVA with Šidák's multiple comparisons test of the edited *HBB*$^{S/S}$ values compared to each other group to calculate *P* values.



Extended Data Fig. 10 | Comparison of DNA damage response and loss of target allele amplification consistent with large deletion or DNA rearrangement in HSPCs following treatment with Cas9 nuclease or with ABE. HSPCs from a healthy human donor were electroporated in triplicate with Cas9 nuclease RNP targeting the *BCL11A* erythroid-specific enhancer, ABE8e-NRCH mRNA and an sgRNA targeting the wild-type *HBB* locus, or no cargo as a control. An additional set of control cells was not electroporated. **a**, *CDKN1* transcription levels, a measure of the p53-mediated DNA damage response[49], were quantified by ddPCR after reverse transcription, and were normalized to *CDKN1* levels before electroporation ($n = 3$). **b**, Editing efficiencies at the targeted genomic loci in HSPCs were measured by HTS 6 days after electroporation. Adenine base editing at the synonymous bystander position 9 of the *HBB* protospacer is shown for ABE8e-NRCH. **c**, **d**, The indicated target sites were amplified and quantified by ddPCR to measure the fraction of missing alleles consistent with larger deletions, translocations, or other chromosomal rearrangements that result in loss of the ability to be amplified by PCR. PCR amplification of a non-targeted *ACTB* site was used to normalize each sample. Each DNA sample was assessed in triplicate ($n = 9$). Data shown as mean ± s.d., with individual values in bar graphs shown as dots. Statistical significance between edited and unedited samples was assessed by a two-tailed Student's *t*-test; ns, not significant.

nature research

Corresponding author(s): Jonathan S. Yen, Mitchell J. Weiss, David R. Liu

Last updated by author(s): Apr 27, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Illumina Miseq (v3.1), NovoSeq (v1.6) , and NextSeq (v2.2) Control software was used on the Illumina Miseq, NovoSeq, and NextSeq sequencers to collect the high-throughput DNA sequencing data. CellRanger (v5.0.1) was used to map RNA-seq data to the human genome. The 10x Genomics' Vartrix tool (v1.1.19) was used to assess allelic editing outcomes.

Data analysis

Illumina sequencing of HSPCs were analyzed by joining paired reads and analyzing amplicons for allelic outcomes using CRIS.py (https://pubmed.ncbi.nlm.nih.gov/30862905/). Indels were reported as the number of reads without the WT amplicon length.
Illumina sequencing of HEK293T cells were conducted by single end reads analyzed by CRISPResso2 (v2.0.40) (https://pubmed.ncbi.nlm.nih.gov/30809026/).
CIRCLE-seq data analyses were performed using open-source CIRCLE-seq analysis software (v1.1) and default recommended parameters (https://github.com/tsailabSJ/circleseq).

CasOFFinder off-target editing analysis
Computational prediction of NRCH PAM-containing potential off-target sites with minimal mismatches relative to the intended target site (three or fewer mismatches overall, or two or fewer mismatches allowing G:U wobble base pairings with the guide RNA) was performed using CasOFFinder (v2.4).

Quantification of base editing efficiency at evaluated off-target sites
The A•T-to G•C editing frequency for each position in the protospacer was quantified using CRISPRessoPooled (v2.0.41) with quantification_window_size 10, quantification_window_center -10, base_editor_output, conversion_nuc_from A, conversion_nuc_to G. The genomic features of the confirmed off-target sites were annotated using HOMER (v4.10). The editing frequency for each site was calculated as the ratio between the number of reads containing the edited base (i.e., G) and the total number of reads. To calculate statistical significance of off-target editing for the ABE8e-NRCH mRNA or RNP treatments compared to control samples, we applied a Chi-square test for each of four samples (two donors, each with two replicates). The 2x2 contingency table was constructed based on the number of edited reads and the number of unedited reads in treated and control groups. FDR was calculated using the Benjamini/Hochberg method. The 24 reported significant off-targets were called based on: (1) FDR < 0.05 and (2) difference in editing frequency

between treated and control > 1% for at least one treatment. The software used to conduct this analysis is available to download using this website: https://github.com/tsailabSJ/MKSR_off_targets.

CellRanger (v5.0.1) was used to map RNA-seq data to the human genome. The 10x Genomics' Vartrix tool (v1.1.19) was used to assess allelic editing outcomes in RNA-seq data..

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

High-throughput sequencing data are being deposited in the NCBI Sequence Read Archive database under accession code PRJNA627465.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Statistical analysis were not used to predetermine samples size for each experiment. Instead sample sizes were based on the feasibility of the experiment and availability of cells and mice. For Townes mouse transplantations a sample size of 6 edited and 6 unedited mice were used. Due to mouse availability, only 2 positive control A/S sickle trait mice could be transplanted. For Townes mouse secondary transplantations, 3 mice of each ratio (6 ratios) were used to sufficiently cover a range of editing with a manageable number of mice. Each experiment yielded sufficient differences between samples and controls that statistical significance could be assessed.
Data exclusions	No data were exlcuded
Replication	All attempts at replication were successful - experimental results matched expectations based on pilot experiments.
Randomization	Recipient mice were randomly selected for transplantation cohorts, an unbiased third party determined which mouse would receive which cells. Human CD34+ cells were de-identified before selection, so the chosen cells were random. For each donor used in an editing experiment, pools of the same cells were kept unedited and used as controls, so randomization is not relevant for such samples.
Blinding	Mice were housed, fed, and handled identically by staff at the animal facility without knowledge of the treatment groups. Mouse identification numbers were used to blind investigators from which conditions were assigned to each mouse during analysis. Analysis and quantification of NGS of editing were conducted by an unbiased automated system, so they were not subjected to qualitative judgment of the analysis. Sickling quantification was performed by a researcher blinded to sample identity.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Anti-Human CD235a FITC, clone GA-R2 (HIR2), BD Pharmingen™ 559943 (Dilution 1:100)
Anti-Human CD49d PE, clone 9F10, BioLegend 304304 (Dilution 1:20)
Anti-Human Band3 APC, custom clone, New York Blood Center Gift from X. An (Dilution 1:100)
Anti-Mouse CD45 FITC/BV786, clone 30-F11/30-F11, BD Pharmingen™/BD HorizonTM 561088/564225 (Dilution 1:40)
Anti-Human CD45 BV605, clone HI30, BD Horizon™ 564047 (Dilution 1:20)
Anti-Human CD33 PE-Cy™7, clone P67.6, BD Biosciences 333946 (Dilution 1:20)
Anti-Human CD3 APC-Cy™7, clone SK7 (Leu-4), BD Pharmingen™ 557832 (Dilution 1:20)
Anti-Human CD19 (Leu™-12) PE/FITC, clone 4G7/HIB19, BD Biosciences/BD PharmingenTM 349209/555412 (Dilution 1:20)
Anti-Human CD34 Alexa Flour 700/PE, clone 581/581, BD Pharmingen™/BD PharmingenTM 561440/555822 (Dilution 1:20)
Anti-Human CD235a APC, clone GA-R2 (HIR2), BD Pharmingen™ 551336 (Dilution 1:20)
Anti-mouse CD45.1 PE, clone A20, BD Biosciences 553776 (1:50 for IF) (Dilution 1:50)
Anti-mouse CD45.2 FITC, clone 104, BD Biosciences 561874 (1:50 for IF) (Dilution 1:50)

Validation

Anti-Human CD235a FITC, clone GA-R2 (HIR2), BD Pharmingen™ 559943 (1:100 for FACS) Metais et al, Blood Adv, 2019
Anti-Human CD49d PE, clone 9F10, BioLegend 304304 (1:20 for FACS) Validation: Metais et al, Blood Adv, 2019
Anti-Human Band3 APC, clone custom, New York Blood Center Gift from X. An (1:100 for FACS) Validation: Metais et al, Blood Adv, 2019
Anti-Mouse CD45 FITC/BV786, clone 30-F11/30-F11, BD Pharmingen™/BD HorizonTM 561088/564225 (1:40 for FACS) Validation: Laggase et al, Nature Med, 2020 / Metais et al, Blood Adv, 2019
Anti-Human CD45 BV605, clone HI30, BD Horizon™ 564047 (1:20 for FACS) Validation: Metais et al, Blood Adv, 2019
Anti-Human CD33 PE-Cy™7, clone P67.6, BD Biosciences 333946 (1:20 for FACS) Validation: Metais et al, Blood Adv, 2019
Anti-Human CD3 APC-Cy™7, clone SK7 (Leu-4), BD Pharmingen™ 557832 (1:20 for FACS) Validation: Metais et al, Blood Adv, 2019
Anti-Human CD19 (Leu™-12) PE/FITC, clone 4G7/HIB19, BD Biosciences/BD PharmingenTM 349209/555412 (1:20 for FACS) Validation: Metais et al, Blood Adv, 2019 / Bradbury et al, J Immunol, 1993
Anti-Human CD34 Alexa Flour 700/PE, clone 581/581, BD Pharmingen™/BD PharmingenTM 561440/555822 (1:20 for FACS) Validation: Egeland et al, Transplant Proc, 1993
Anti-Human CD235a APC, clone GA-R2 (HIR2), BD Pharmingen™ 551336 (1:20 for FACS) Validation: Metais et al, Blood Adv, 2019
Anti-mouse CD45.1 PE, clone A20, BD Biosciences 553776 (1:50 for FACS) Validation: Jafri et al, Scientific Reports, 2017
Anti-mouse CD45.2 FITC, clone 104, BD Biosciences 561874 (1:50 for FACS)Validation: Jafri et al, Scientific Reports, 2017

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s): HEK293T (ATCC)

Authentication: Cells from ATCC were authenticated by the supplier by STR analysis.

Mycoplasma contamination: All cells tested negative for mycoplasma.

Commonly misidentified lines (See ICLAC register): None used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Mouse, NOD.Cg-KitW-41J Tyr + Prkdcscid Il2rgtm1Wjl/ThomJ "NBSGW", Female, 7-9 weeks;
Mouse, B6;129-Hbbtm2(HBG1,HBB*)Tow/Hbbtm3(HBG1,HBB)Tow Hbatm1(HBA)Tow/J "Townes", Male and Female, 8-14 weeks;
Mouse, B6.SJL-Ptprca Pepcb/BoyJ, Female, 8-12 weeks.
Mice were housed in individually ventilated cages, 5 animals per cage, with free access to food and water. Animal rooms observe a 12-hour light and dark cycle (light 6am-6pm), and maintains an ambient temperature of 21C, and approx 30-70% humidity.

Wild animals: The study did not involve wild animals.

Field-collected samples: The study did not involve samples collected from the field.

Ethics oversight

Ethics oversight: The St. Jude Institutional Animal Care and Use Committee approved the use of mice in transplantation experiments and the animal studies were performed according to relevant ethical regulations.
All studies utilizing mice were approved by the St. Jude Children's Research Hospital Institutional Animal Care and Use Committee under Protocol 579 entitled "Genetic Models for the Study of Hematopoiesis". Mice were maintained in the St. Jude Children's Research Hospital Animal Resource Center according to recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Patients with HBSS Sickle cell disease. Thirteen participants with SCD (HbSS) were enrolled at SJCRH or NIH (n=11) between July 2017 and February 2019. Median age was 29 years (20-50 years) and 47% were male (n=7). In the clinical trials we isolated CD34 + cells from three donors. One donor from SJCRH and from the NIH were received de-identified, without the researchers knowing age or sex of the donors.
Recruitment	Participants were recruited from the respective hematology and sickle cell clinics at NIH and SJCRH. These participants were representative of the general clientele of the clinics except that all of them were adults. No minor/pediatric participants were recruited to these studies. Participants were volunteers with no immediate direct benefit from participation in this study. We do not expect any self-selection bias or other bias in recruitment would impact the results of this study, as samples of the same HBB genotype were considered identical. To assess the impact of genome editing we assessed the same cell populations with and without base editor treatment, so no direct comparisons between different participants were conducted.
Ethics oversight	Plerixafor-mobilized CD34+ cells from patients with SCD were collected according to the protocol "Peripheral Blood Stem Cell Collection for Sickle Cell Disease Patients" (ClinicalTrials.gov identifier NCT03226691), which was approved by the human subject research institutional review boards at the National Institutes of Health and St. Jude Children's Research Hospital. All participants provided informed consent.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NCT03226691
Study protocol	Available at clinicaltrials.gov with NCT # "NCT03226691"
Data collection	Participants were enrolled at St. Jude Children's Research Hospital and the NIH between July 2017 and February 2019
Outcomes	The current study used specimens collected and banked during the course of clinical trail NCT03226691, but does not report the outcomes of that clinical trial. Because we are not reporting the outcomes of the clinical trial in this publication, the CONSORT checklist does not apply.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Bone marrow, peripheral blood, and in-vitro cultured cells were resuspended in PBS with 0.1% BSA. Cells were filtered using 40-um filter before flow.
Instrument	Attune NxT Flow Cytometer, BD FACSAria III, BD LSRFortessa
Software	FACS Diva for Data Collection, FlowJo for data analysis
Cell population abundance	FACS machine cell sorting efficiency was confirmed by flow cytometric analysis of post-sorted cells
Gating strategy	FSC-A/SSC-A for mononuclear cells, followed by SSC-A/SSC-W for singlets, DAPI for DAPI- live cells. Human/mouse chimerism and lineages were analyzed using: Anti-Mouse CD45 FITC/BV786 Anti-Human CD45 BV605 Anti-Human CD33 PE-Cy™7 Anti-Human CD3 APC-Cy™7 Anti-Human CD19 (Leu™-12) PE/FITC Anti-Human CD34 Alexa Flour 700/PE

Anti-Human CD235a APC

Erythroid maturation were gated by
Anti-Human CD49d PE
Anti-Human Band3 APC
Anti-Human CD235a FITC
while mouse chimerism the cells were ggated by
Anti-mouse CD45.1 PE
Anti-mouse CD45.2 FITC

See Extended Data Figure 3, 8 and 9 for details.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.