Supplementary Appendix to:

CRISPR-Cas9 Editing of the *HBG1*/*HBG2* Promoters to Treat Sickle Cell Disease

Akshay Sharma, MBBS,[1] Jaap-Jan Boelens, MD, PhD,[2] Maria Cancio, MD,[2] Jane S. Hankins, MD, MS,[1] Prafulla Bhad, MSc,[3] Marjohn Azizy, PharmD,[3] Andrew Lewandowski, PhD,[3] Xiaojun Zhao, PhD,[3] Shripad Chitnis, PhD,[3] Radhika Peddinti, MD,[4] Yan Zheng, MD, PhD,[1] Neena Kapoor, MD,[5] Fabio Ciceri, MD,[6] Timothy Maclachlan, PhD,[3] Yi Yang, PhD,[3] Yi Liu, PhD,[3] Jianping Yuan, PhD,[3] Ulrike Naumann, PhD,[7] Vionnie W.C. Yu, PhD,[3] Susan C. Stevenson, PhD,[3] Serena De Vita, MD, PhD,[3] James L. LaBelle, MD, PhD.[4]

[1]St. Jude Children's Research Hospital, Memphis, TN, USA; [2]Memorial Sloan Kettering Cancer Center, New York, NY, USA; [3]Novartis Institutes for BioMedical Research, Cambridge, MA, USA; [4]The University of Chicago Medicine, Chicago, IL, USA; [5]Children's Hospital Los Angeles, Los Angeles, CA, USA; [6]IRCCS San Raffaele Hospital, Milan, Italy, [7]Novartis Institutes for BioMedical Research, Basel, Switzerland.

Contents

Methods

1. Preclinical screening using CRISPR-Cas9 tiling of the *HBG1* and *HBG2* promoter regions to identify the guide RNA that led to the highest F-cell induction

An unbiased guide RNA (gRNA) screen of the *HBG1* and *HBG2* promoter regions was performed to identify regulatory regions within the *HBG1* and *HBG2* promoter regions, disruption of which led to potent induction of fetal hemoglobin (HbF)-immunostaining erythroblasts (F-cells). Seventy-two gRNAs were designed to target the *HBG1* and *HBG2* promoter regions (Table S1). They were screened for F-cell induction in erythroid progeny after electroporation of the preformed RNP complex into human CD34+ cells followed by erythroid differentiation.

Initial gRNA selection was performed computationally, using the human reference genome (build GRCh38), by identifying the 5'-NGG-3' protospacer adjacent motifs (PAM) in the *HBG1* and *HBG2* upstream region and targeting the 20-nucleotide protospacer at the 5' of each site. For the gRNAs chosen, there were no other target sites with 100% sequence homology and a viable PAM elsewhere in the genome. The designed gRNAs were first prepared in a dual-guide RNA (dgRNA) format for the screening assay, then selected guides were prepared in a single-guide RNA (sgRNA) format for follow-up validation. Each gRNA was chemically synthesized and purified by high-performance liquid chromatography (HPLC).

The RNP complex was prepared by adding CRISPR RNA (crRNA) and transactivating CRISPR RNA (tracrRNA) to the Cas9 protein. On day 1, ribonucleic acid–protein (RNP) complexes consisting of Cas9 complexed with each dual-guide RNA (crRNA + tracrRNA) were electroporated into healthy-donor CD34+ hematopoietic stem and progenitor cells (HSPCs), then erythroid differentiation was induced. Briefly, CD34+ cells were thawed and placed in culture overnight in StemSpan SFEM medium (STEMCELL Technologies) with IL-12, stem cell factor (SCF), thrombopoietin (TPO), FLT3L, and penicillin–streptomycin added. For each RNP delivery reaction, 90,000 CD34+ cells from healthy adult donors were concentrated by centrifugation, resuspended in 63 µL of P3 nucleofection buffer (Lonza), to which active RNP was subsequently added, and electroporated with a Lonza Nucleofector, using program CA-137. Each crRNA was examined in triplicate and compared to the mock-edited control that delivered the RNP containing only Cas9 and tracrRNA. Positive controls included gRNA-08 (G8), which targets the BCL11A exon 2, and gRNA-57 or gRNA-63, which target the BCL11A binding motif in the *HBG1/2* promoters.[1]

After electroporation, the cells were immediately transferred into 250 µL of pre-warmed erythroid differentiation medium (EDM), consisting of Iscove's Modified Dulbecco's Medium (IMDM) (GE Life Sciences, cat. no. SH30228.01), 330 µg/mL human holotransferrin (Invitria, cat. no. 777TRF029), 10 µg/mL recombinant human insulin (Gibco, cat. no. Al 1382 11), 2 IU/mL heparin (Sigma, cat. no. H3393), 5% human AB serum (Sigma, cat. no. H4522), 2.5 U/mL human erythropoietin (PeproTech, cat. no. 100-064), and 1× penicillin–streptomycin–glutamine (Gibco, cat. no. 10378-016). During the initial culture period up to day 7, EDM was further supplemented with 1.38 µM hydrocortisone (Sigma, cat. no. H8672), 100 ng/mL human SCF (Life Technologies, cat. no. PHC21 13), and 5 ng/mL human IL-3 (PeproTech, cat. no. 10779-598) to make EDM-I. After 4 days, the cell cultures were diluted in fresh medium. Cultures were maintained for a total of 7 days in the culture conditions described above, after which time half of the cells were analyzed by intracellular staining for HbF expression. Briefly, the cells were washed once with PBS, resuspended in LIVE/DEAD® Fixable Violet Dead Cell Stain (Thermo Fisher Scientific, cat. no. L34963; diluted 1:1000 in PBS), and incubated for 30 min. Cells were then washed and stained with 1:50 dilutions of anti–CD71-BV711 (BD Biosciences, cat. no. BDB563767) and anti–CD235a-APC (BD Biosciences, cat. no. BDB551336) antibodies for 30 min. Next, the cells were washed, fixed with fixation buffer (BioLegend, cat. no. 420801), and permeabilized with 1× intracellular staining permeabilization wash buffer (BioLegend, cat. no.

421002) according to the manufacturer's instructions. The cells were then incubated with a 1:40 dilution of anti–HbF-PE antibody (Life Technologies, cat. no. MHFH04) in 50 µL of 1× intracellular staining permeabilization wash buffer (BioLegend, cat. no. 421002) for 20 min at room temperature. The cells were washed twice with 0.2 mL of 1× intracellular staining Perm/Wash buffer, resuspended in staining buffer, and analyzed for HbF expression on an LSRFortessa flow cytometer (BD Biosciences). The results were analyzed using FlowJo, and the data were presented as the percentage of F-cells in the viable CD71-positive erythroid cell population.

After 7 days, the efficiency of on-target editing was measured by next-generation sequencing (NGS). The crRNAs that consistently led to the highest F-cell induction in the erythroid progeny (an increase of more than 20% F-cells above the mock-edited control background on day 7) were chosen for further characterization and development. These selected gRNAs were then prepared in an sgRNA format for follow-up validation and were electroporated into CD34+ HSPCs with Cas9 protein as RNPs. Erythroid differentiation was then induced as described above. For these experiments, a gRNA targeting the *BCL11A* +58 erythroid-specific enhancer was used as a positive control. The single protospacer sequence of the *BCL11A* +58 erythroid-specific enhancer gRNA is located on the positive strand of chromosome 2 at base position 60,495,236-60,495,255 (genome reference build GRCh38) and is 5′-ATCAGAGGCCAAACCCTTCC-3′.

2. Isolation of CD34+ cells from healthy donors and from individuals with SCD for preclinical evaluation

Cryopreserved mobilized peripheral blood–derived CD34+ cells from healthy donors were purchased from Lonza (cat. no. 2M-101D). Cryopreserved mobilized peripheral blood–derived CD34+ cells from patients with SCD were supplied by Dr. John Tisdale (National Heart, Lung, and Blood Institute, Bethesda, MD). Bone marrow aspirates from patients with SCD were received from Dr. Edmund Waller (Emory University Hospital, Atlanta, GA). Blood samples from patients with SCD were obtained from Conversant Bio (order no. 4892-A, 5521-A). Peripheral blood mononuclear cells (PBMCs) of patients with SCD were obtained by density centrifugation on Ficoll-Paque PLUS (GE Healthcare Bio-Sciences, cat. no. 17144002), and CD34+ cells were enriched from them by using a Miltenyi Biotec CD34 MicroBead Kit (cat. no. 130-046-702) and the AutoMACS apparatus in accordance with the manufacturer's instructions. CD34+ cells were then either cryopreserved and stored in liquid nitrogen or used for expansion in SEM, which contained StemSpan-ACF medium (STEMCELL Technologies, cat. no. 09855) supplemented with 1× penicillin–streptomycin, 500 µM LHD221, and 50 ng/mL of SCF, TPO, and FLT3L.

3. On-target allelic editing frequency in preclinical samples

For preclinical studies, the percentages of on-target Cas9–gRNA-68 indels were identified by NGS and quantitative polymerase chain reaction (qPCR), as described previously.[1,2] Each indel created at each on-target site of gRNA-68 was sequenced by NGS. The primers for NGS were: *HBG1* forward: 5′-CGCTGAAACTGTGGCTTTATAGAAATT; *HBG2* forward: 5′-GCACTGAAACTGTTGCTTTATAGGAT; *HBG1/2* reverse: 5′-GGCGTCTGGACTAGGAGCTTATTG. qPCR was used to confirm the deletion of approximately 5 kb between the two on-target sites. A primer and probe set targeting the *HBG1/2* loci was obtained from Thermo Fisher Scientific (cat. no. 4400294). The primer sequences were as follows: *HBG* forward: ACGGATAAGTAGATATTGAGGTAAGC; *HBG* reverse: GTCTCTTTCAGTTAGCAGTGG; TaqMan probe (FAM): ACTGCGCTGAAACTGTGGTCTTTATAG. TaqMan qPCR was performed on genomic DNA samples, using the TaqMan Fast Advanced Master Mix (Applied Biosystems, cat. no. 4444557) to quantify each sample in triplicate. ΔΔCt values were calculated based on the amplification of a RPPH1 VIC PL reference probe mix (Thermo Fisher Scientific, Hs04930436_g1). The editing efficiency was calculated using the following equation: total % editing efficiency = % large deletion + (1 − % large deletion) × (% indel of *HBG1* + % indel of *HBG2*)/2.

The percentages of on-target Cas9–gRNA-*BCL11A* +58 erythroid-specific enhancer indels were identified by NGS. The primers used for NGS were: *BCL11A* +58 forward: 5′-AGCTCCAAACTCTCAAACCACAGGG; *BCL11A* +58 reverse: 5′-TACAATTTTGGGAGTCCACACGGCA.

4. Erythroid differentiation and HbF induction

Genome-edited CD34+ cells were subjected to a three-stage erythroid differentiation in liquid culture as previously described.[3] The percentage of F-cells was determined at day 14 of erythroid culture. As described previously for the screening assay, cells were stained intracellularly for HbF, using an anti–human γ-globin antibody directly conjugated with a fluorophore (Life Technologies, cat. no. MHFH014).

HbF protein levels were also determined by ion-exchange HPLC. The percentage of HbF protein was determined at day 18 of erythroid culture by ion-exchange HPLC on a D-10 Hemoglobin testing system (Bio-Rad Laboratories Inc., Hercules, CA, USA). A minimum of 1 million cells were collected then snap frozen. Cell pellets were lysed in water at 50,000 cells per microliter for 5 minutes, followed by 3 freeze thaw cycles. Cellular debris was cleared by centrifugation at 17,400 relative centrifugal field (rcf) for 5 minutes. The hemolysate was analyzed for HbF protein levels and the percentage of HbF protein determined in each sample.

5. In vivo engraftment studies in mice

Approximately 4 h before injection, NSG (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ) mice were irradiated with 175 cGy of whole-body cesium137 irradiation. Mice were then injected with human CD34+ cells edited using gRNA-68 (i.e., modified in the promoter regions of the duplicated γ-globin genes on chromosome 11), human CD34+ cells edited at the BCL11A +58 erythroid-specific enhancer, or mock-edited human CD34+ cells from three donors via a single tail vein injection.

Clinical observations (once daily during weeks 1–2 and at least twice weekly during weeks 3–26/27) and measurements of body weight (on day 1, twice weekly during weeks 1–4, and weekly thereafter) were performed for all groups of mice. Blood samples were collected at 4, 8, 12, 17, and 20 weeks post injection and at necropsy (26/27 weeks post injection) for flow cytometry. In addition, the right hind limb of each mouse was collected at necropsy and bone marrow was extracted from it for flow cytometry. The following mouse (m) and human (h) markers were assessed: mCD45, hCD45, hCD33, hCD3, and hCD19. Bone marrow was collected from selected animals for DNA isolation and NGS analysis of the edited genomic sites.

6. Flow cytometric analysis of engrafted human cells in transplantation experiments

Bone marrow cells were collected from hind limbs and vertebrae of mice at 26 weeks post transplant. RBCs were lysed, washed, and filtered to remove dead cell aggregates. A small fraction of the whole bone marrow cells was kept and stained with antibodies to hCD45, mCD45, hCD33, hCD3, and hCD19 to study human and lineage chimerism in the bone marrow. The following antibodies were used for flow cytometry: PE-conjugated anti–human CD45 (BD Pharmingen, cat. no. 561866), APC-cy7–conjugated anti–mouse CD45 (BD Pharmingen, cat. no. 557659), APC–conjugated anti–human CD33 (BioLegend, cat. no. 366606), BV421-conjugated anti–human CD3 (BioLegend, cat. no. 317344), and FITC-conjugated anti–human CD19 (BioLegend, cat. no. 302256). The percentage of human chimerism in the bone marrow was defined as % hCD45+ cells/(% hCD45+ cells +% mCD45+ cells).

7. Identification of candidate off-target sites by computational methods

Genomic loci homologous to the gRNA protospacer sequence were identified using the CasOFFinder algorithm.[4] Based on recent studies,[5-7] analyses focused on identifying and characterizing the

following locus types: 1) sites with 0–4 mismatches adjacent to a 5′-NGG-3′ PAM; 2) sites with five mismatches within RefSeq[8] exons adjacent to a 5′-NGG-3′ PAM; 3) sites with 0 mismatches adjacent to a noncanonical 5′-NAG-3′ PAM; and 4) sites with up to three mismatches with one- or two-nucleotide "bulges" in the gRNA sequence relative to the genomic sequence or adjacent to a 5′-NGG-3′ PAM. The computational method described here identified genomic loci homologous to the sgRNA targeting spacer regions that could represent potential off-target editing sites.

8. Identification of candidate off-target sites by using the Pan-Heme cancer panel

We have developed a custom-content hybrid capture panel–based targeted NGS assay (Pan-Heme v1.0, with the SureSelect panel from Agilent Technologies) targeting 352 genes that are frequently mutated in hematologic malignancies (Table S5), identified by mining mutations in the COSMIC database,[9] plus 83 malignant gene fusion/translocation events (Table S6). The panel was applied to unedited and gRNA-68–edited research materials derived from three healthy donors. Variant analysis, using the MuTect algorithm[10] for single nucleotide variant (SNV) detection and the PINDEL algorithm[11] for indel detection (with an empirically determined detection threshold for variants of at least 1%), identified no editing.

9. Identification of candidate off-target sites by using SITE-Seq and validation by rhAmpSeq

The SITE-Seq cell-free, unbiased, genome-wide biochemical genome-editing NGS assay enables the selective enrichment and identification of adapter-tagged genomic DNA ends after genome editing.[12] The assay was performed as previously described,[12] with the modification of using PBMC-derived genomic DNA from two individuals and from a pool of multiple individuals at one RNP concentration (64 nM, with a 3:1 molar ratio of gRNA to Cas9), using the gRNA-68 and the Cas9 protein used to manufacture OTQ923. In addition to the two on-target events, the analysis identified 156 potential off-target sites that were further characterized in DNA from unedited and gRNA-68–edited CD34+ cells derived from three healthy donors (with the latter cells being generated using gRNA-68 and the Cas9 protein). Hits situated in annotated repeats[13] were excluded. Technical replicates were analyzed using a multiplexed rhAmpSeq panel from Integrated DNA Technologies before NovaSeq (Illumina) sequencing was performed. BWA-MEM alignment[7] to the human genome reference sequence (build GRCh38) was then performed, followed by variant analysis using the VarDict algorithm[14] (with an empirically determined detection threshold for variants of 1%).

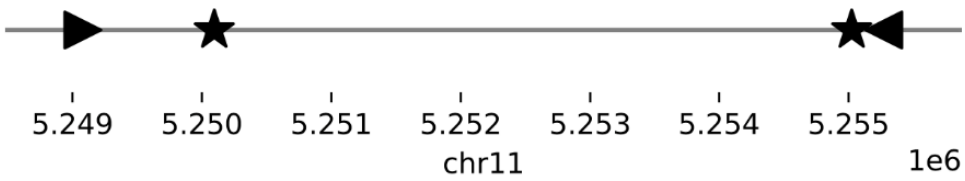
10. Additional preclinical characterization

UnIT assay: The Unique Identifier Tagmentation (UnIT) NGS assay is an adaptation of the UDiTaS assay,[15] which identifies on-target centric genome rearrangements in addition to the typical CRISPR-Cas9 on-target indels. Like UDiTaS, the UnIT assay uses a Tn5 transposase to fragment and tag double-stranded DNA with oligonucleotides containing unique molecular identifiers, which are used to remove PCR duplicates during analysis, and sequences to enable NGS library construction.

Karyotyping: Standard analysis of 300 cells in metaphase, scoring for chromosomal number changes and gross chromosomal aberrations, and 50 metaphases for karyotype for more granular analysis of chromosomal integrity, was performed in genome edited and unedited CD34+ cells.

11. Analysis of gRNA-68 on-target genome editing using long-read sequencing

Genomic DNA was extracted from gRNA-68–edited CD34+ cells derived from three donors (donors 1 2, and 3) or from unedited cells (no sgRNA or Cas9) from the same donors. A PCR amplicon spanning the *HBG1* and *HBG2* loci edited by gRNA-68 at genomic coordinates chr11:5,250,097-5,250,116 and chr11:5,255,025-5,255,044 (in human genome reference build GRCh38), respectively, was designed with Primer3 (version 2.3.6), using the default parameters.[16] For unedited samples, the PCR primers should amplify an amplicon approximately 6.2 kb in length, as depicted in the figure below.



The schematic diagram shows the region on chromosome 11 surrounding the genomic coordinates of the two gRNA-68 homology sites (chr11:5,250,097-5,250,116 and chr11:5,255,025-5,255,044, marked with asterisks), as well as the primer binding sites of the forward (right-pointing arrow) and reverse (left-pointing arrow) primers used for long-range amplicon sequencing. Primer sequences used for long-range amplification and genomic coordinates of the primer binding sites are shown in the table below.

Primer name	**Primer sequence 5' – 3'**	**Genomic coordinates**
HBG_LA_FWD5	GAAAGATGAAAGTTCTTTCTACTGG	chr11:5,249,081-5,249,105
GCR-001 off rev	TGAGACTAAGACGTGTCCCA	chr11:5,255,253-5,255,272

Genome editing with gRNA-68 should produce a product approximately 1.2 kb in length, resulting from the excision of the ~5-kb fragment between the two gRNA-68 targeting sites in *HBG1* and *HBG2*. Gel-purified PCR amplicon pools were subjected to NGS library construction with the SMRTbell Express Template Preparation Kit v2 (Pacific Biosciences, cat. no. 100-938-900), used in accordance with the manufacturer's recommendations, and the products were loaded onto a Sequel sequencer. Raw sequencing data were converted to sequencing reads in the FASTQ format (Cocket al., 2010) with SMRT Link (v.7.0) (Pacific Biosciences), running the circular consensus (CCS) pipeline with the CCSDMLT settings. Reads were aligned with pbmm2 (PacBio/Bioconda version 1.0.0) and screened for structural variants by using pbsv (PacBio/Bioconda version 2.2.1) and SVIM.[17] De novo motifs were detected using the HOMER algorithm.[18]

12. Mobilization and apheresis

Red blood cell (RBC) exchange transfusions were performed for at least 2 months before the first mobilization of CD34+ cells with plerixafor. Before starting plerixafor mobilization, participants were assessed by the study investigator to confirm that they were still eligible to proceed with apheresis (according to local institutional guidelines). Participants received an RBC exchange transfusion within 7 days before the start of the mobilization/apheresis cycle. Each participant underwent stem cell mobilization with plerixafor, and PBMCs were collected by apheresis, as described previously.[19]

13. OTQ923 manufacturing procedures

A schematic of OTQ923 manufacturing is shown in Figure S1. OTQ923 is an autologous hematopoietic stem and progenitor cell product generated in accordance with current Good Manufacturing Practice (cGMP) by a continuous manufacturing process that uses cryopreserved mobilized peripheral apheresis (mAph) product from the patient. Upon receipt at the manufacturing facility, the mAph product is thawed and washed, and CD34+ cells are isolated using a closed, automated, sterile, micro-bead system (CliniMACS Prodigy System, Miltenyi Biotec). CD34+-enriched cells are incubated in cell culture medium containing Rh TPO, rh SCF, rh Ft-3L, and human serum

albumin (HSA). Genome editing is then performed by electroporation in a buffer containing the RNP complex prepared by mixing Cas9 and gRNA-68. After electroporation, the cells are expanded in the presence of cytokines and an aryl hydrocarbon receptor antagonist, LHD221 (which is similar to, but more potent than SR1[20]). After culture, the cells are washed, resuspended in cryopreservation medium (HSA and CryoStor®CS10 solution), and cryopreserved. Quality control release assays are performed on cryopreserved material before batch release.

14. Conditioning

Before OTQ923 infusion, participants received conditioning with busulfan administered IV every 6 hours, with pharmacokinetic monitoring being used to achieve a total cumulative exposure of 80–100 mg*h/L over 4 days (Table S2).

15. Hemoglobin fractionation and HbF assessment in clinical samples

Blood samples were collected to assess the hemoglobin fractionation by capillary electrophoresis, based on the protocol-defined schedule of assessment. Samples for hemoglobin fractionation and HbF assessment were analyzed locally.

16. F-cell analysis using flow cytometry in clinical samples

Whole-blood samples were fixed with 0.05% glutaraldehyde, and the cells were permeabilized with detergent (0.1% Triton X100). The permeabilized cells were stained with a phycoerythrin (PE)-conjugated monoclonal antibody specific for fetal hemoglobin (Life Technologies MHFH04; clone HBF-1). Data were acquired on a FACSCantoII flow cytometer running BD FACSDIVA™ software, version 9.0.1, and were analyzed with DeNovo FCS Express 6. Scatter gates were set to collect singlet erythroid events, and HbF was estimated by fluorescence in the PE channel.

17. On-target editing efficiency in clinical samples

Genomic DNA was extracted from peripheral blood nucleated cells and from bone marrow–derived CD34+ cells at various time points. CD34+ cells were isolated from bone marrow aspirate samples (from Participant 1 at month 12) with a Ficoll-Paque PREMIUM mononuclear cell isolation kit (GE Healthcare, USA) followed by a CD34 MicroBead Kit (Miltenyi Biotec, Bergisch Gladbach, Germany).

The on-target allelic editing of *HBG1* and *HBG2* was assessed using the Illumina amplicon sequencing method. Specifically, each sample was sequenced in duplicate, together with its corresponding baseline sample as the control. A sequencing library was generated for each sample with two rounds of PCR amplification. The target-specific primer sequences for both *HBG1* and *HBG2* are GGTCAAGTTTGCCTTGTCAAGGCT and TACAGGCCTCACTGGAGCTA. Amplicon libraries were sequenced on an Illumina MiSeq sequencer with v3 chemistry and 2 × 300-bp paired-end sequencing. Demultiplexed FASTQ files were processed by the following steps: quality clipping, mapping (GRCh38), filtering, and downsampling to roughly 200,000 if more than 200,000 reads were aligned to a target region. Reads aligning outside the expected target regions were removed, as were reads with ambiguous alignments (having a mapping quality of 0). Aligned reads were used to determine the target allelic editing frequency and to identify unique indel changes by using a CAS9-aware version of VarDict.[14] Reported variants were also filtered using replicates and baseline controls to reduce false-positive variants. The threshold for allele frequency was set to >0.01%; deletions and insertions with a length of 1 bp and a frequency less than 0.02% are not considered. Data for *HBG1* and *HBG2* are presented as the total percentage edit, using the formula "total % editing efficiency = % large deletion + (1 − % large deletion) × (% indel of *HBG1* + % indel of *HBG2*)/2" (Figs. 2E and 2F). To calculate the percentage edit for *HBG1* and *HBG2*, individual fractions were calculated separately for *HBG1* and *HBG2* and applied as "(1 − % large deletion) × (% indel of *HBG1* or % indel of *HBG2*)."

A ddPCR method of detection by amplification in individual water-in-oil droplets was implemented to detect the approximately 5-kb genomic deletion in patient whole-blood or bone marrow samples. Droplets were analyzed with a Bio-Rad QX200 droplet reader, and data were processed with QuantaSoft_v1.7.4 software that included automatic Poisson correction. In this assay, specific primers for *HBG2* were used to measure the relative copy number of the gene. Genomic DNA extracted from patient samples collected post OTQ923 infusion had a decreased signal for the gene of interest (*HBG2*, on chromosome 11) relative to that for a stable, unedited reference albumin gene (*ALB*, on chromosome 4). The ratio of the concentrations (in copies/20 µL) of the *HBG2* gene and the *ALB* reference gene was reported as the editing percentage (%). As the assay detects the wildtype sequence (the unedited target gene), the percentage of edited cells is the inverse (1 − ratio) of the relative ratio between the two targets. Results were reported as the percentage edit of the inverse (1 − ratio). The ddPCR assay used for clinical samples replaced the assessment by qPCR of the approximately 5-kb genomic deletion in preclinical samples. The correlation between the two methods is represented in Figure S6.

Results

1. Selection of gRNA-68 for use in manufacturing OTQ923

An unbiased CRISPR-Cas9 tiled gRNA screen of the *HBG1* and *HBG2* promoter regions identified several gRNAs that resulted in the induction of HbF-positive erythroid cells or F-cells (Fig. S2A). Erythroblasts derived from genome-edited or unedited control CD34+ cells were analyzed by flow cytometry for HbF expression. Genome-edited CD34+ cells and mock-edited cells (i.e., cells treated with RNP containing only Cas9 and tracrRNA as a background control) generated similar fractions of CD71+ erythroblasts, indicating that editing did not adversely affect erythroid differentiation. Delivery of these gRNA RNPs to CD34+ cells resulted in an increased percentage of erythroid cells containing HbF (up to 62.8%), as compared to the percentage in mock electroporated cells (16.9%), at day 7 after electroporation. A gRNA (G8) targeting exon 2 of *BCL11A* was included as a positive control for this initial screening. Several of the dgRNAs targeting the *HBG1* and *HBG2* promoter regions resulted in higher F-cell levels than were obtained with G8 as well as with gRNAs targeting the known BCL11A target motif in the *HBG1/HBG2* promoters[1] (Fig. S2A). Editing at the *HBG1* and *HBG2* promoter regions was determined on day 7 after electroporation by NGS to determine the percentage of edited alleles in the cell population. Genome editing percentages for the *HBG1* and *HBG2* promoter regions were high in many of the cell cultures electroporated with RNPs containing Cas9, the crRNA of the given targeting domain, and tracrRNA, but not in mock-edited control cells (in which the RNP contained only Cas9 and tracrRNA). In particular, the dgRNAs that resulted in an increase in F-cells of more than 20% above the mock-edited control background on day 7 had a range of 5.92% to 77.84% edited alleles for the *HBG1* or *HBG2* target loci. Based on their ability to induce F-cells to levels more than 20% higher than those in mock-edited cells, which correlated with their genome editing efficiency, six gRNAs were selected for further characterization: gRNA-01, gRNA-08, gRNA-10, gRNA-22, gRNA-54, and gRNA-68. The location of these gRNAs relative to the *HBG1/HBG2* genes and the known transcription factor binding motifs in the *HBG1/HBG2* promoter regions are shown in Figure S2B. During further validation experiments, a gRNA targeting the BCL11A +58 erythroid-specific enhancer was included as a positive control in addition to gRNA-01, gRNA-08, gRNA-10, gRNA-22, gRNA-54, and gRNA-68. The editing efficiency and F-cell induction obtained after editing with these gRNAs and Cas9, followed by erythroid differentiation are shown in Figures S2C and S2D. Erythroid progeny of CD34+ cells edited with gRNA-

68 consistently produced the highest percentage of F-cells; hence, gRNA-68 was chosen for further validation in the clinical trial.

2. Identification of gRNA-68

The 100-nucleotide sequence of gRNA-68 is as follows:

5′-
ACUGAAUCGGAACAAGGCAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCA
ACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU-3′

The 20-nucleotide spacer sequence of gRNA-68 is as follows:

5′- ACUGAAUCGGAACAAGGCAA-3′

The 20-nucleotide protospacer sequence for gRNA-68 is as follows:

5′- ACTGAATCGGAACAAGGCAA-3′

The three-nucleotide canonical 5′-NGG-3′ PAM sequence 3′ to the protospacer sequence for gRNA-68 is as follows:

5′-AGG-3′

3. Genome location for gRNA-68 within the *HBG1/2* promoter regions

We mapped the binding site of gRNA-68 in the promoter regions of *HBG2* and *HBG1* and found that it is not near any known transcription factor binding site (Fig. S2B). As a result, none of the small indels observed around the cutting site at the *HBG2* or *HBG1* locus with a frequency higher than 0.1% are likely to destroy nearby transcription factor binding sites. The ~5-kb deletion was the most dominant indel in gRNA-68–edited cells from healthy donors (Fig. 1C) and from patients with SCD in preclinical studies (Fig. 1E); this is similar to observations with reference gRNA-57 and gRNA-63, which disrupt the known BCL11A binding motif in the *HBG1/HBG2* promoters, as reported in the literature.[1] Although we have been unable to characterize fully the fusion site of the newly reconstituted *HBG2*–*HBG1* fusion gene that is formed after the ~5-kb deletion, in which the *HBG2* promoter sequence is fused to the downstream *HBG1*, the reconstituted fusion gene is likely to have various insertions or deletions. Likewise, we observe various destructive edits at the gRNA binding site in 60% to 80% of sequencing reads, depending on the sample. In most cases, one of two edits is observed: either a six-nucleotide deletion at positions 15–21 of the binding site (AAGGCA deleted, 45% to 60% of destructive edits, 31% to 47% of reads) or a five-nucleotide deletion at positions 14–19 of the binding site (CAAGG deleted, 16% to 22% of destructive edits, 11% to 17% of reads). Known transcription factor binding sites (Fig. S2B) remain intact in the resulting allele in >97% of sequencing reads.

4. Validation of selected gRNAs

The lead gRNA (gRNA-68) was selected from among the six gRNAs chosen for further characterization (gRNA-01, gRNA-08, gRNA-10, gRNA-22, gRNA-54, and gRNA-68) based on several factors, including an initial characterization of potential off-target sites using computational methods, the percentage of F-cells obtained in an erythroid differentiation assay, and an *in vivo* assessment of the engraftment and multi-lineage differentiation potential of the genome-edited CD34+ HSPCs based on their long-term engraftment. Figure S2C illustrates the percentage of genome editing for each selected guide and Figure S2D shows the percentage of F-cells determined at day 14 of erythroid culture. In this assay, mock-electroporated cells served as F-cell background controls and the F-cell percentage of the controls was subtracted from the F-cell value for the test gRNA. A positive control targeting the *BCL11A* +58 erythroid-specific enhancer region was included in this experiment. gRNA-68 resulted in

the highest F-cell induction, with the level being 37.6% ± 4.1% (mean ± SD) above background, as compared to 16.1% ± 3.3% (mean ± SD) above background for the cells in the positive control group that were treated with the *BCL11A* +58 erythroid-specific enhancer editing gRNA (Fig. S2D). From these analyses, gRNA-68 was selected as the lead guide for clinical evaluation based on its ability to not only induce HbF upon erythroid differentiation but also maintain hematopoietic stem cell (HSC) function and stability of the editing. As shown in Figure S2B, gRNA-68 has two discrete expected on-target sites 246 bp upstream of the transcriptional start in the nearly identical *HBG1* and *HBG2* genes (GRCh38 chr11:5,255,025–5,255,044 and GRCh38 chr11:5,250,097–5,250,116, respectively), which are approximately 5 kb apart.

As an additional confirmation of gRNA-68 ability to induce HbF, HbF protein was determined in the erythroid differentiation assay using edited CD34+ HSPCs from both health donors (n=2) and sickle cell disease patients (n=3). Fig. 1 shows the percentage of HbF protein determined by ion-exchange HPLC analysis at day 18 of erythroid culture for the gRNA-68 edited samples compared to the mock electroporated controls. gRNA-68 resulted in the induction of HbF protein up to 20.9% ± 0.2% (mean ± SD) compared with 12.0% ± 0.3% in the mock-edited control samples derived from edited healthy donor HPSCs and in SCD patient derived samples, HbF protein levels increased up to 26.2% ± 2.9% compared to 16.1% ± 2.3% in mock-edited controls. These results corroborate the F-cell data obtained by flow cytometry with gRNA-68 edited samples.

5. gRNA-68–edited HSPCs engraft and reconstitute all hematologic lineages in a preclinical transplant model

Flow cytometry analysis of the bone marrow of NSG mice at 26 weeks post transplant showed high engraftment levels of 30–50% in the gRNA-68–edited group and 35–60% in the *BCL11A* +58 erythroid-specific enhancer–edited and mock-edited groups. Figure S3A shows the average percentage of human CD45 cells (huCD45) found in isolated bone marrow cells for each treatment group at week 26 (the terminal end point). Lineage reconstitution was similar in the mock-edited, *BCL11A* +58 erythroid-specific enhancer–edited, and gRNA-68–edited groups for each donor (Fig. S3B). In both the gRNA-68–edited and *BCL11A* +58 erythroid specific-enhancer–edited groups, an average of 78% genome editing was achieved in CD34+ cells from the three donors before xenotransplantation (Fig. S3C). When the proportions of genome-edited cells before xenotransplantation and after 26 weeks of reconstitution were compared, the percentage of edited cells was 71.1% and 80.9% after 26 weeks of reconstitution for the gRNA-68–edited and *BCL11A* +58 erythroid-specific enhancer–edited groups, respectively, suggesting that the gRNA-68–edited CD34+ cells were indeed long-term repopulating HSCs. Additionally, the frequency of the ~5-kb deletion was maintained throughout the reconstitution period until 26 weeks post transplant. These data demonstrate that gRNA-68–edited CD34+ cells persist in transplant-recipient animals for at least 26 weeks post transplantation.

To determine whether there had been any clonal evolution or expansion events, genomic DNA from the bone marrow of a subset of animals was sequenced at the on-target site to analyze the editing pattern after 26 weeks of engraftment, which was compared to the editing pattern before dosing. The indels generated by CRISPR-Cas9 editing in each cell serve as a unique barcode that enables clonal tracking of each genome-edited cell and its respective progenies. In this experiment, the representation of each genome-edited clone was determined in gRNA-68 genome-edited CD34+ cells before transplantation and then again in the hematopoietic system after 26 weeks of engraftment (Fig. S4). NGS data demonstrated clonal heterogeneity comprising major and minor

clones in gRNA-68 cells, as would be expected in any hematopoietic system. By tracking the edited clones, several key consistencies were observed: (1) CD34+ cells from different donors generated similar clonal patterns; (2) recipients of transplanted CD34+ cells derived from the same donor exhibited similar clonal patterns; and (3) the clonal patterns before transplantation and after 26 weeks of reconstitution were very similar. The heterogeneity of the editing pattern in the cells analyzed before dosing was mostly maintained in the cells harvested after 26 weeks of engraftment, indicating that no clonal expansion had occurred. We acknowledge the limitation of this method to track clonal expansion, given that the most common and predominant editing outcome was the 5-kb deletion. However, even the cells with the 5-kb deletion had diverse destructive edits at the fusion junction. Together, these data strongly demonstrate the reproducibility and persistence of the editing patterns generated by gRNA-68. Selection of stem cells for homing and engraftment was expected and was manifested as a reduction in the number of clones post transplant; however, the hematopoietic system reconstituted from gRNA-68–edited CD34+ cells remained polyclonal and diverse upon undergoing transplantation and reconstitutive hematopoietic stress. In animals dosed with gRNA-68–edited CD34+ HSCs, despite the larger clonal variation observed post transplantation, there was no reproducible pattern of clonal dominance, nor was any tumor with a human cell origin detected. Such clonal variation was probably the result of stochastic fluctuation, a phenomenon that occurs in any given normal hematopoietic system.[21] In summary, a high level of human cell engraftment with multi-lineage reconstitution was observed in NSG mice throughout the 26 weeks of the post-transplant monitoring period.

6. Computational, SITE-Seq, and Pan-Heme panel off-target site results and potential translocation events

The computational screen identified 163 potential off-target sites, of which 162 were characterized in unedited and gRNA-68–edited materials derived from three donors by single-plex PCR-based targeted NGS analysis and showed no off-target editing (Table S3). One site with five nucleotide mismatches relative to the gRNA spacer sequence, which is represented five times on chromosome 1 in both intergenic and pseudogene regions, could not be amplified because of its non-unique and repetitive genomic context. However, this site represents one of the lowest-risk sites, with only 15 of 20 nucleotides matching the gRNA spacer sequence. SITE-Seq analysis identified 156 hits, excluding sites with an edit distance larger than 6 or more than two bulges and sites in repetitive regions. Characterization in unedited and gRNA-68–edited materials derived from three donors by using rhAmpSeq targeted NGS analysis showed only editing of the two on-target sites on chromosome 11. Using the Pan-Heme panel of common hematologic oncogenes and tumor suppressors, analysis of the coding regions (and some intronic regions) of 352 genes that are frequently mutated in hematologic malignancies and 83 malignant gene fusion/translocation events (Tables S5 and S6) identified no bona fide candidate off-target editing indels or SNVs in gRNA-68–edited CD34+ cells derived from three healthy donors. Table S4 lists the 40 unique sites that are common to the computational and SITE-Seq analyses; none of these sites are present on the Pan-Heme panel. Using these complementary unbiased genome-wide methodologies, we detected no potential off-target editing or unintended genomic alterations of OTQ923 (Table S3).

Analysis of research material generated from three donors, unedited and gRNA-68–edited, identified dicentric fusions between edited sister chromosomes at a rate of approximately 1–2%. As expected, these fusions were not detected in mice after xenotransplantation, and no translocations to other genomic loci were detected. HSCs from individuals with SCD were not evaluated for this analysis as no off-target editing was identified in healthy donor CD34+ cells after xenotransplantation; and this is expected to be a strictly “on-target” effect. The dicentric sister chromosome fusions identified

here are consistent with previous results[15] and will probably be detected at similar rates with any high-efficiency gRNA.

Standard analysis of 300 cells in metaphase, scoring for chromosomal number changes and gross chromosomal aberrations, and 50 metaphases for karyotype for more granular analysis of chromosomal integrity, was performed in genome edited and unedited CD34+ cells. No alterations in chromosome number or structure in gRNA-68–edited research material were identified in this assay.

7. On-target genome-editing using long-read sequencing highlights a complex editing pattern at the targeted region

Long-range sequencing analysis of PCR amplicons spanning the *HBG1* and *HBG2* loci in the OTQ923 cell product was performed to characterize the on-target loci for large indels, changes in or creation of regulatory elements, and the formation of novel fusion genes after genome editing. Analysis of OTQ923 produced from multiple donors (n = 8) confirmed the creation of a ~5-kb deletion between the two gRNA-68 on-target cut sites, which was the most abundant allele. Analysis of gRNA-68 on-target genome editing by using long-read sequencing identified no other large structural variations.[17,18]

It is possible that the ~5kb deletion actually represents a broader array of larger deletions. Alleles with deletions spanning larger regions outside our primer binding sites would not be detected with our assay. Our primers were designed to bind 1 Kb upstream and 250 bp downstream of the gRNA-68 homology sites. The length of the sequences we observed was narrowly distributed around 1.2 kb in edited samples, in addition to the wildtype peak at 6.2 kb.

We searched for novel open reading frames (ORFs) in the gRNA-68–edited sequences by using ORFfinder.[22] The vast majority of the filtered ORFs were supported by fewer than 10 reads. The most abundant putative ORF was 57 amino acids long and was supported by only 36 reads, which accounted for only 0.005% of the total edited reads and were probably the result of technical noise.

We screened for regulatory elements by comparing the motifs resulting from gRNA-68 editing to position weight matrices for human transcription factors. Neither the most abundant nor the best matches for any motif enriched in sequences from gRNA-68–edited samples had perfect matches with known transcription factor binding sites, and many lacked multiple key residues. Therefore, the ability of those sites to recruit a transcription factor is assumed to be compromised.

The location of γ-globin–repressive BCL11A binding sites identified by Liu et al.[23,24] and the ~5-kb deletion created by gRNA-68 were also compared. Our data suggest that one copy of the BCL11A binding site is removed, along with the *HBG2* locus, by the ~5-kb excision, whereas the second copy remains intact with the *HBG1* locus, which implies a novel mechanism for HbF induction by gRNA-68 editing.

Figures and Tables

Figure S1. Schematic of OTQ923 manufacture and treatment approach

After a participant agrees to participate in the study and signs the consent form, screening evaluations are performed to assess their eligibility. Once their eligibility is confirmed, the participant is enrolled on the study, discontinues hydroxyurea (if taking it previously), and begins regular exchange transfusions. After at least 2 months of receiving exchange transfusions, the participant undergoes CD34+ cell mobilization using plerixafor and apheresis to collect the cells. The apheresis product is cryopreserved and shipped to the central manufacturing site, and a rescue aliquot containing at least 2 million CD34+ cells per kilogram of body weight of the recipient is cryopreserved and kept on site. This process of cryopreservation and apheresis is repeated until sufficient cells have been collected and an adequate dose of OTQ923 has been manufactured. After cryopreservation, the cellular product undergoes release testing and quality assurance and, upon release, it is shipped to the clinical site. The participant then undergoes myeloablative conditioning with busulfan and receives the OTQ923 product infusion. After the infusion, the participant remains hospitalized until engraftment. They are then followed as an outpatient on study for at least 2 years and for an additional 13 years on a long-term follow-up protocol.

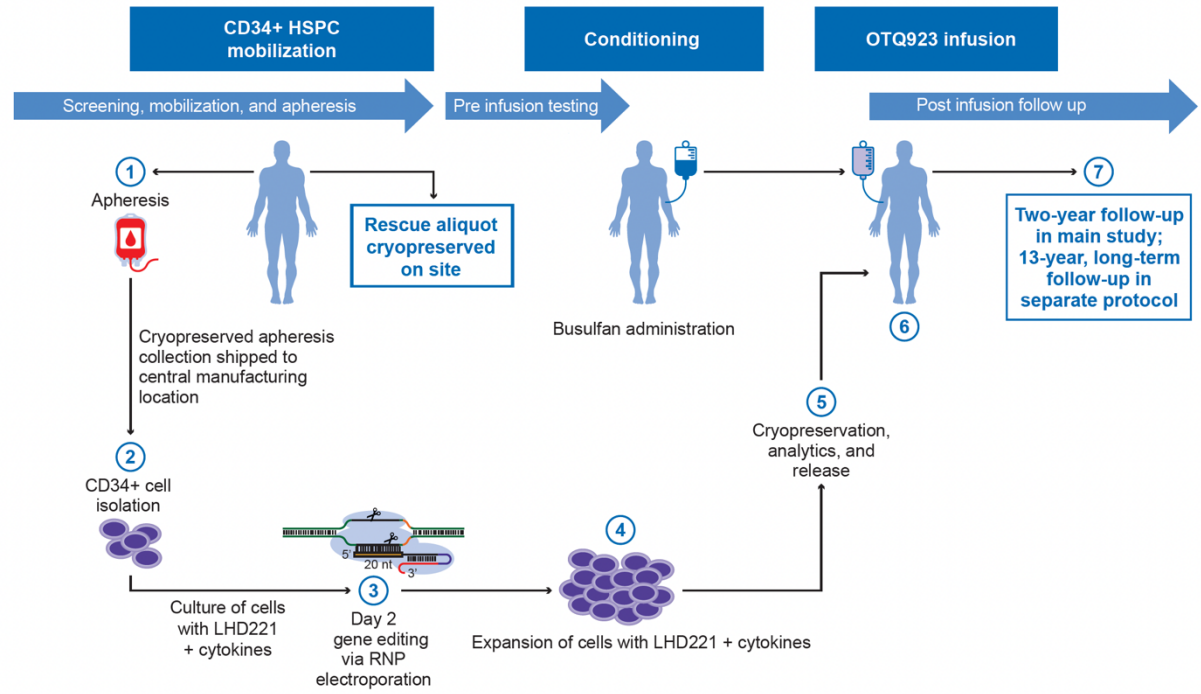


Figure S2: Selection of gRNA-68 to use for manufacturing OTQ923

A tiled CRISPR/Cas9 gRNA screening of *HBG1* and *HBG2* promoter regions for *cis*-regulatory elements that repress transcription was performed. An unbiased panel of 72 tiled gRNAs was designed to interrogate the *HBG1* and *HBG2* promoter regions (Table S1). Guide RNA selection was performed computationally, using the human reference genome (build GRCh38), by identifying the 5'-NGG-3' protospacer adjacent motifs (PAMs) in the *HBG1* and *HBG2* upstream region and targeting the 20-nucleotide protospacer at the 5' of each motif. All gRNAs represented unique sequences in the genome. The designed gRNAs were first prepared in a dual-guide RNA format for the screening assay. On day 1, RNP complexes consisting of Cas9 complexed with each dual-guide RNA (crRNA + tracrRNA) were electroporated into healthy-donor CD34+ hematopoietic stem and progenitor cells (HSPCs), after which erythroid differentiation was induced. **(A)** Percentage of HbF-immunostaining cells (F-cells) according to the gRNA used, measured by immunoflow cytometry after 7 days of erythroid differentiation. Mock-edited cells (as a negative control) received RNP containing Cas9 and tracrRNA only. Positive controls included cells treated with gRNA-G8 (solid black bar), which targets the *BCL11A* exon 2, and gRNA-57 or gRNA-63 (purple colored bars), which target a known BCL11A binding motif.[1] The teal colored bar corresponds to induction of F-cells by gRNA-68. The results are shown as the mean ± standard deviation (n = 3 replicates). **(B)** Diagram of the *HBG1* and *HBG2* promoter sequences with known regulatory protein binding sites shown in green. The numbering indicates the distance from the transcription start site (TSS) at position +1. Guide RNAs that generated the highest percentages of F-cells (gRNAs 08, 10, 22, 54, and 68) are indicated as red arrows. **(C)** Indel types for the indicated guide RNAs are shown. CD34+ HSPCs were electroporated with RNPs consisting of Cas9 complexed with the indicated single gRNAs (sgRNAs), induced to undergo erythroid differentiation, and analyzed for indel frequency via NGS after 7 days. The graph shows the mean ± standard deviation (SD) (n=4 experiments) using CD34+ HSPCs from two different donors. **(D)** F-cell induction after editing by the indicated sgRNAs and erythroid differentiation. CD34+ HSPCs were electroporated with RNPs as described for panel C and were analyzed by immuno-flow cytometry for the F-cell percentage after 14 days of erythroid differentiation. The bar chart shows the results for each sgRNA as the mean ± SD minus the average background level in negative control cells electroporated with Cas9 only (n = 4) using CD34+ HSPCs from two different donors. Note that sgRNA-68 produced the highest percentage of F-cells. One-way ANOVA with Dunnett's multiple comparisons test was performed: ns, not significant; **P* < 0.05; ***P* < 0.01; ****P* < 0.0001.

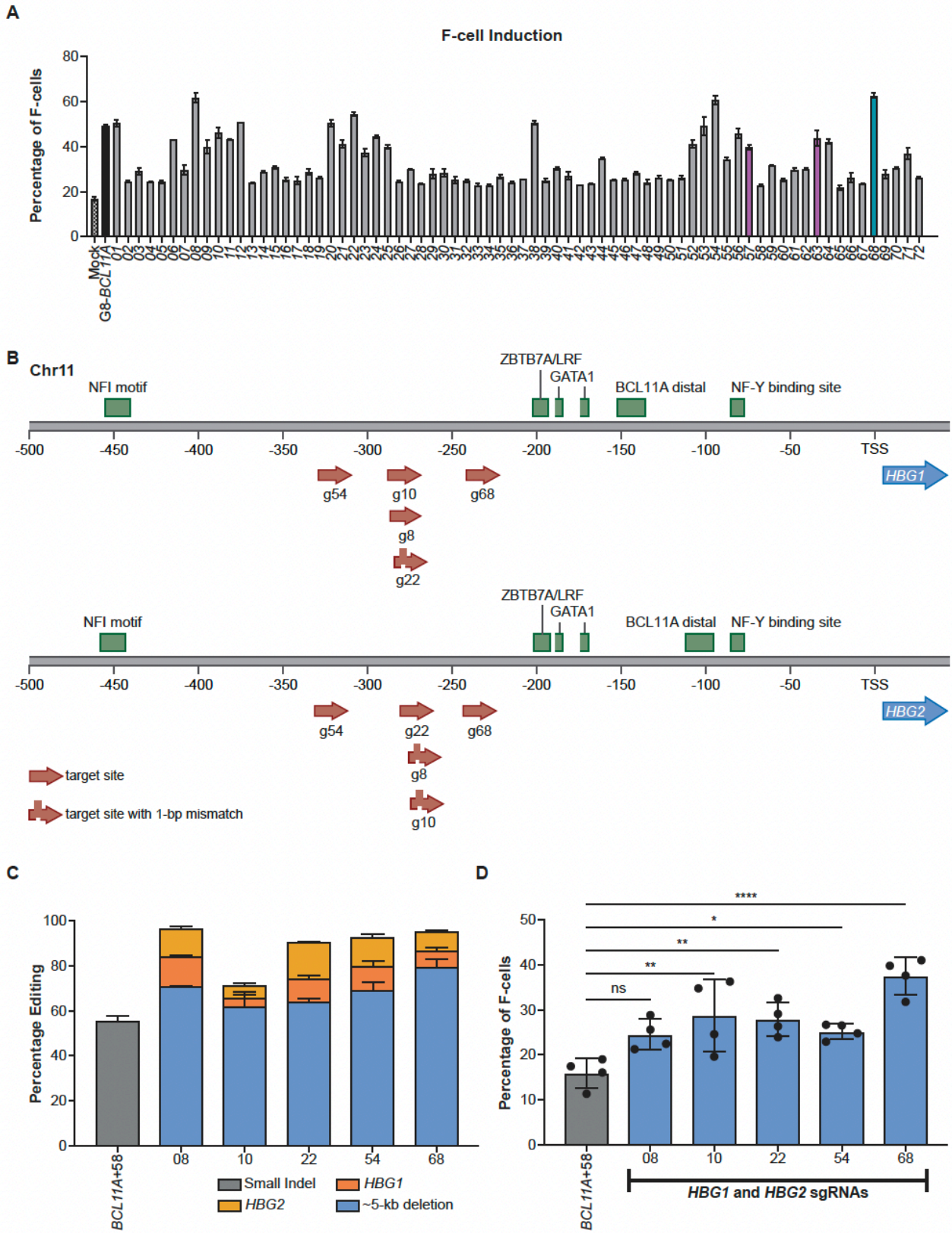

A
F-cell Induction
Percentage of F-cells
Mock
G8-BCL11A
B
Chr11
NFI motif
ZBTB7A/LRF
GATA1
BCL11A distal
NF-Y binding site
TSS
HBG1
HBG2
g54
g10
g68
g8
g22
target site
target site with 1-bp mismatch
C
Percentage Editing
BCL11A+58
Small Indel
HBG1
HBG2
~5-kb deletion
D
Percentage of F-cells
ns
HBG1 and HBG2 sgRNAs

Figure S3. Xenotransplantation of gRNA-68–edited HSPCs.
Healthy-donor CD34+ HSPCs were electroporated with RNP consisting of Cas9 complexed with sgRNA-68 or sgRNA targeting the *BCL11A* +58 erythroid-specific enhancer as a positive control. As a negative control, HSPCs were electroporated with Cas9 only (i.e., mock edited). Cells were transplanted into lethally irradiated NSG mice and analyzed after 26 weeks. **(A)** Human chimerism in recipient mouse bone marrow, as evaluated by the expression of human-specific CD45 (hCD45). Symbols represent results from the three different CD34+ cell donors, designated on the x-axis as 1, 2, and 3. **(B)** Multi-lineage reconstitution of bone marrow by edited HSPCs. The bar charts show the mean ± SD from three biological replicate experiments for each CD34+ cell donor (n = 3–40 mice per group). **(C)** Indel frequencies of input cells pre-transplantation and of donor-derived hCD45+ cells 26 weeks post transplantation (n = 3–6 mice per group). Data are shown as the mean ± SD.

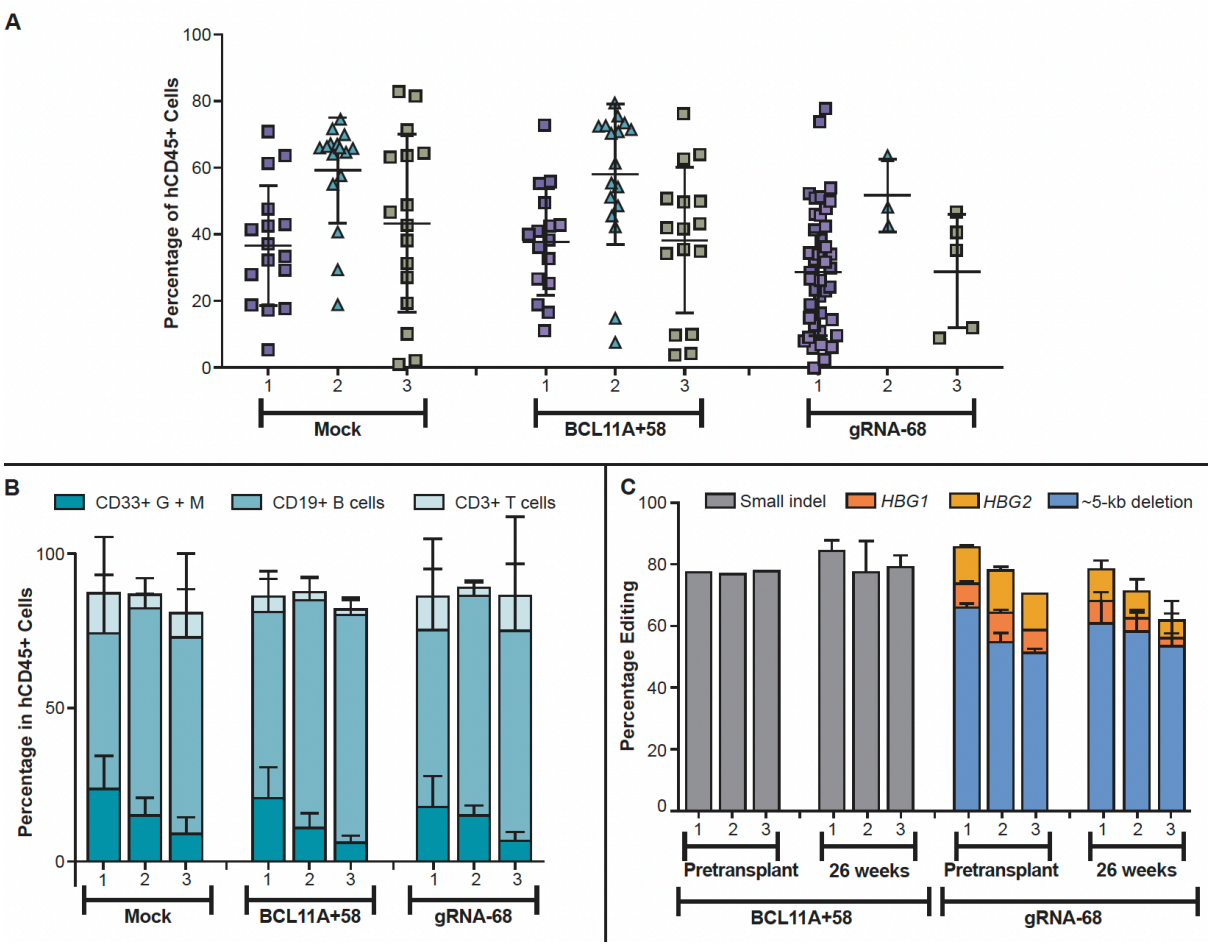


Figure S4. Editing pattern at the *HBG1* and *HBG2* promoter regions in cells before transplantation and in recipient NSG mice 26 weeks after transplantation.
On-target edits in gRNA-68–edited cells included small indels at the RNP cleavage sites in the *HBG1* and *HBG2* promoters or a ~5-kb deletion resulting from simultaneous RNP cleavage at *HBG1* and *HBG2* followed by the loss of the intervening DNA. The ~5-kb deletion resulted in a single chimeric *HBG2–HBG1* fusion gene with the *HBG2* promoter sequence fused to the downstream *HBG1* gene. The edited cells that were xenotransplanted into NSG mice are shown as the percentage of the total engrafted human CD45+ cells. The allele frequencies determined for each donor group before (Pre) and after (Post) xenotransplantation are shown for three independent donors. The percentages of the ~5-kb deletion are shown in yellow, and the indel distributions obtained at *HBG1* and *HBG2* are shown in different colors. The major indels at these sites are indicated, with each color representing an allele with a specific on-target indel. The light-green boxes collectively represent the edited alleles with low-frequency indels.

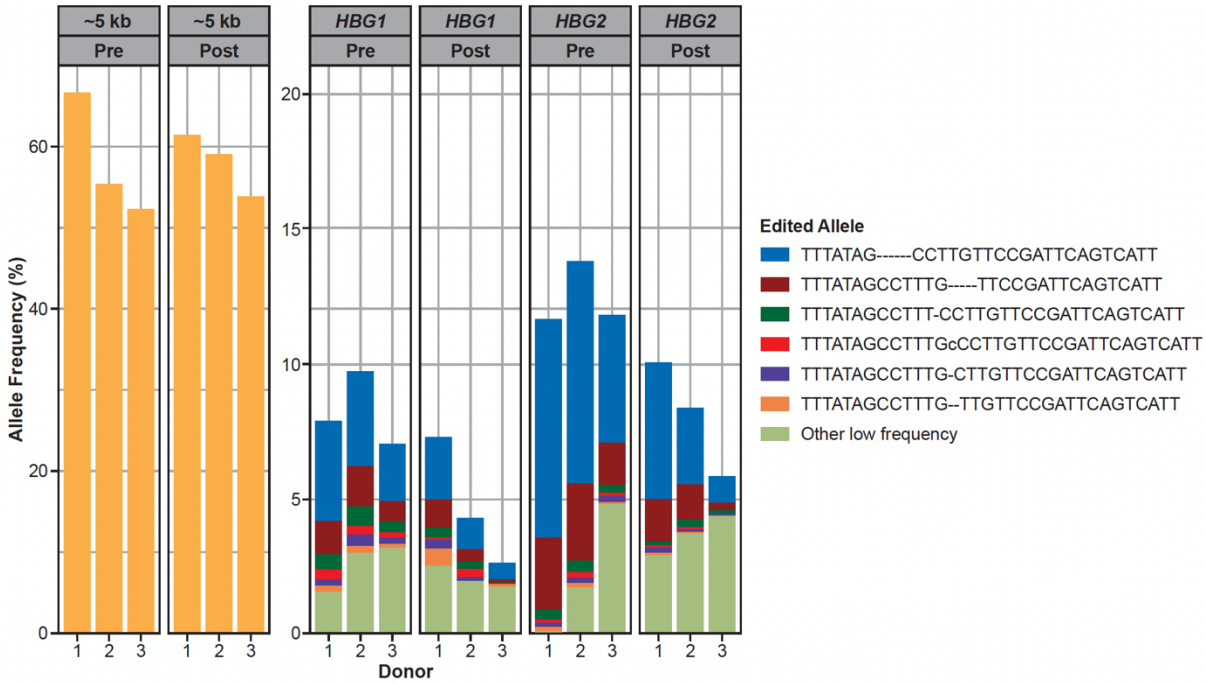
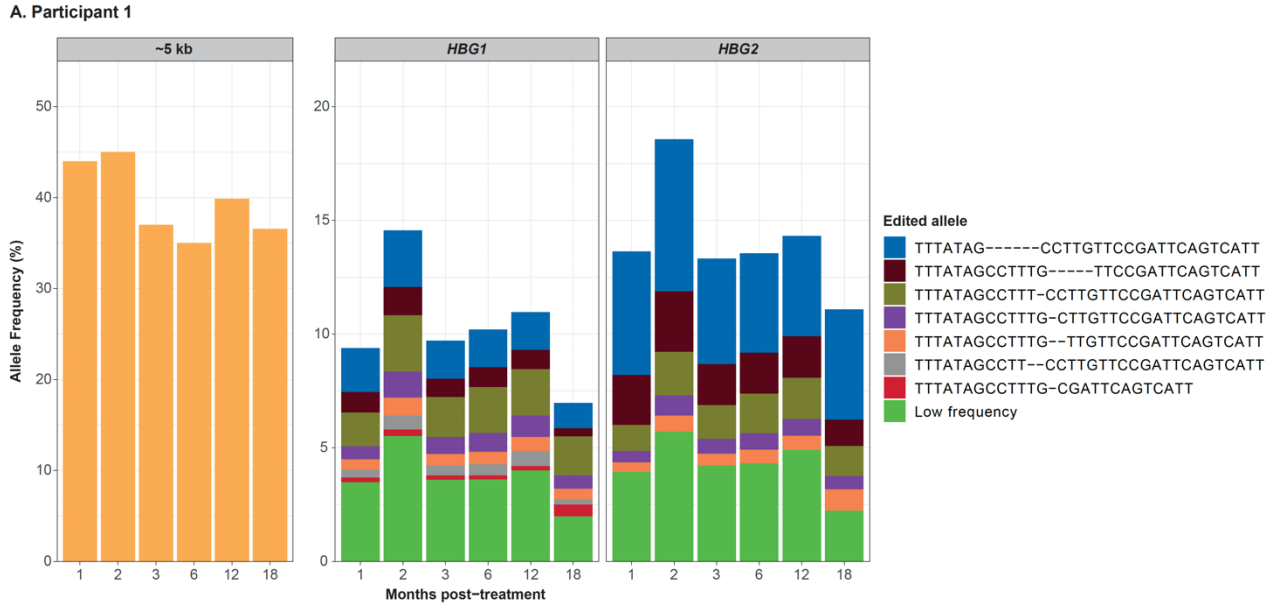
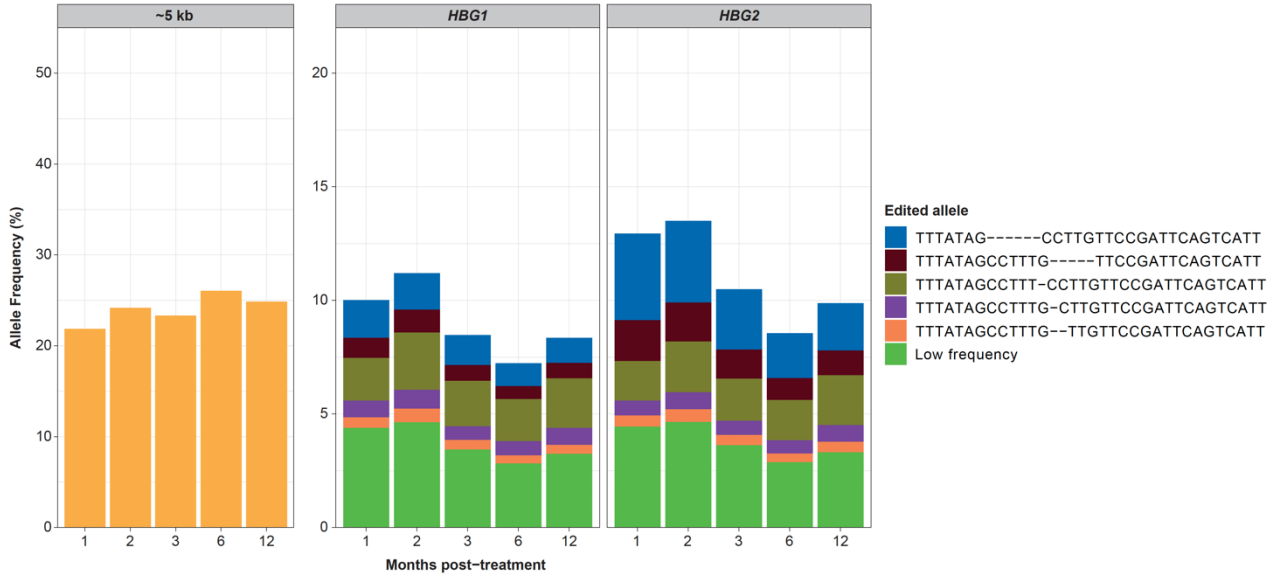


Figure S5. Clonal dynamics in post-infusion samples from trial participants treated with OTQ923
The fractions of cells with the ~5-kb deletion and various indels were estimated in the cells collected from the three clinical trial participants at various time points. The bar graphs show the frequency of the ~5-kb deletion and the indel patterns at the *HBG1* and *HBG2* promoters in DNA isolated from peripheral blood nucleated cells from Participant 1 up to month 18 **(A),** from Participant 2 up to month 12 **(B)**, and from Participant 3 up to month 6 **(C)**. The percentages of the ~5-kb deletion are shown in yellow, and the indel distributions obtained at *HBG1* and *HBG2* are shown in different colors. The major indels at these sites are indicated, with each color representing an allele with a specific on-target indel. The light-green boxes collectively represent the edited alleles with low-frequency indels. Indels in the sorted neutrophils, T cells and B cells collected from the participants are shown in panel **D**. Orange bars represent the percentages of indels detected by next-generation sequencing (bright orange indicates indels at the *HBG1* locus, pale orange indicates indels at the *HBG2* locus). Blue bars represent the ~5-kb deletion that was detected in the infused product by quantitative PCR or by droplet digital PCR. Indel frequencies were consistent between BM cells (Figs. 2C, 2F and 2I) representative of the hematopoietic stem cells as well as sorted B-cells and myeloid cells/neutrophils in each participant (Shown below in supplemental Fig. 5D). Indel frequencies fluctuated in the T-cells during the first few months after infusion likely due to differential expansion of unedited T-cells and those derived from the edited HSCs after infusion.

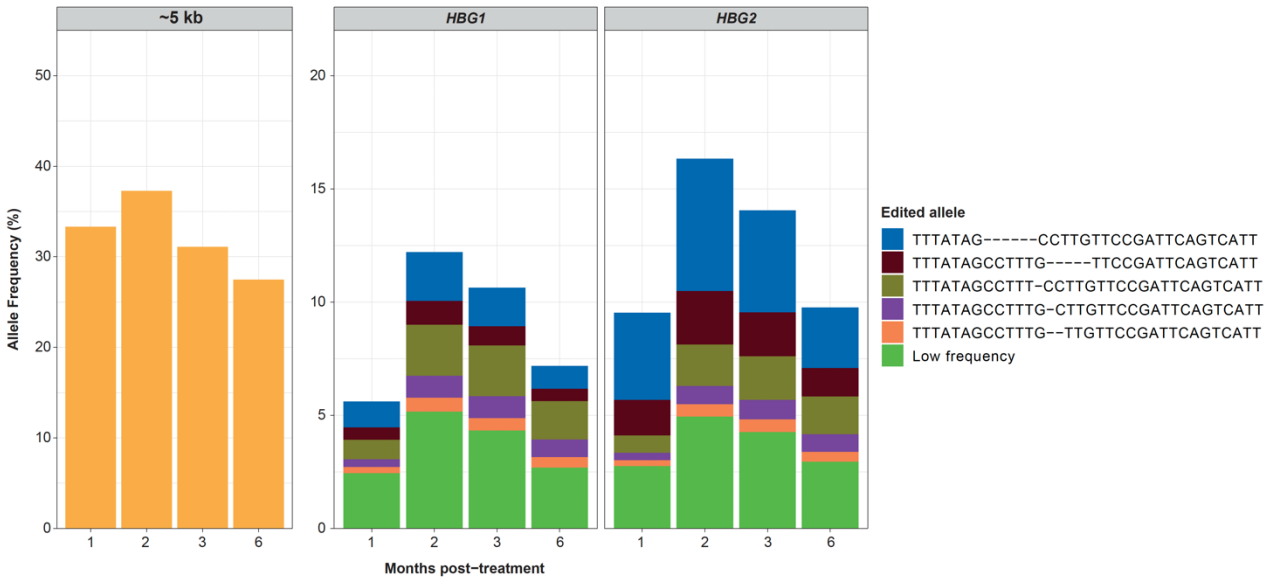


Figures continued on the next page

B. Participant 2



C. Participant 3



Figures continued on the next page

D. Indels in sorted neutrophils, T-cells, and B-cells.

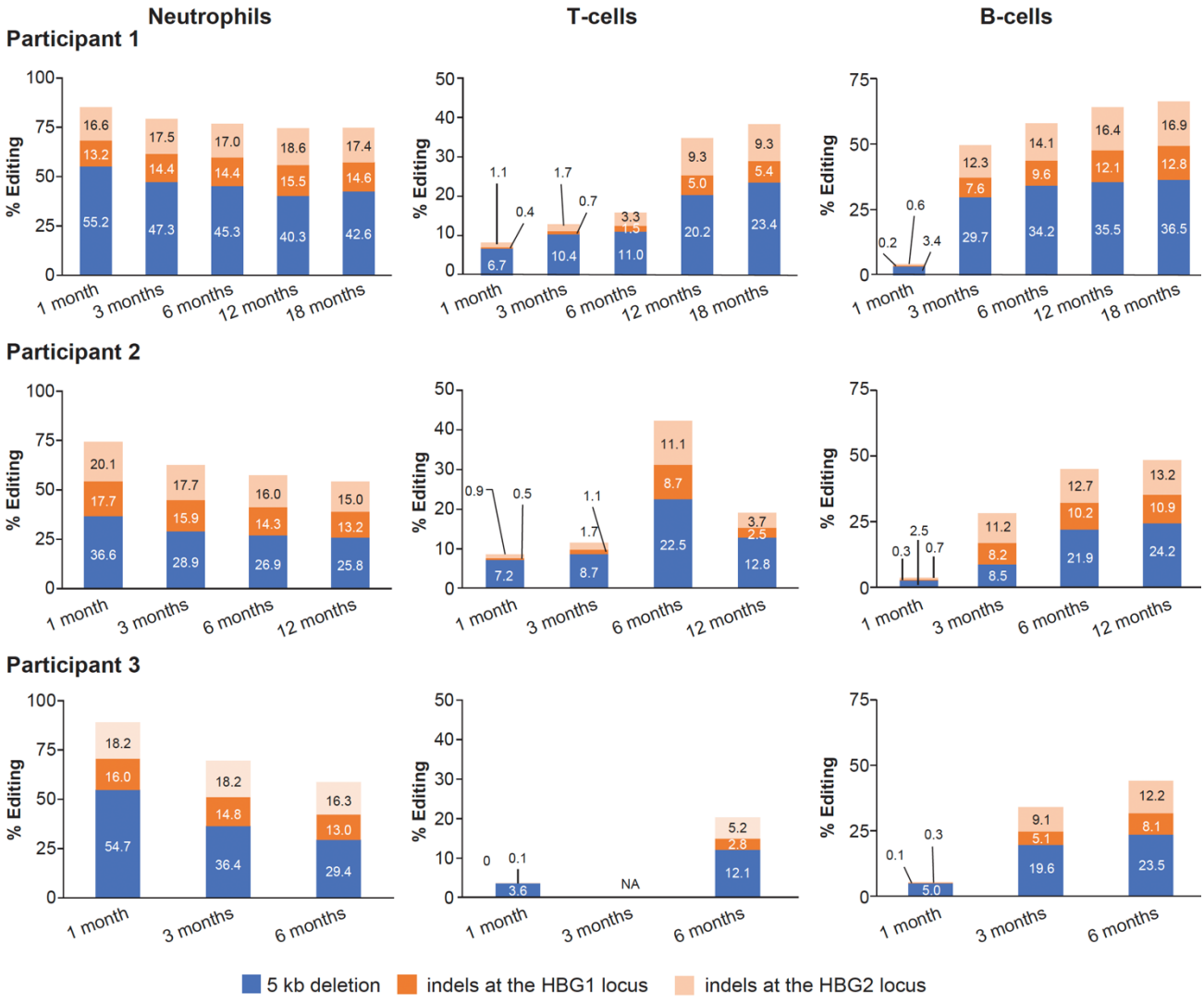


Figure S6. Comparison of ddPCR and qPCR

The correlation between the percentages of the ~5-kb deletion data generated using the ddPCR and qPCR techniques is shown in the plot below. The percentage of cells with the ~5-kb deletion in cryopreserved gRNA-68–edited CD34+ cells from healthy donors (n = 3) was determined using both the ddPCR and the qPCR techniques. The correlation between the percentages of cells with the ~5-kb deletions as calculated using ddPCR (x-axis) and qPCR (y-axis) revealed good agreement between the data, with a coefficient of determination (R2) of 0.9.

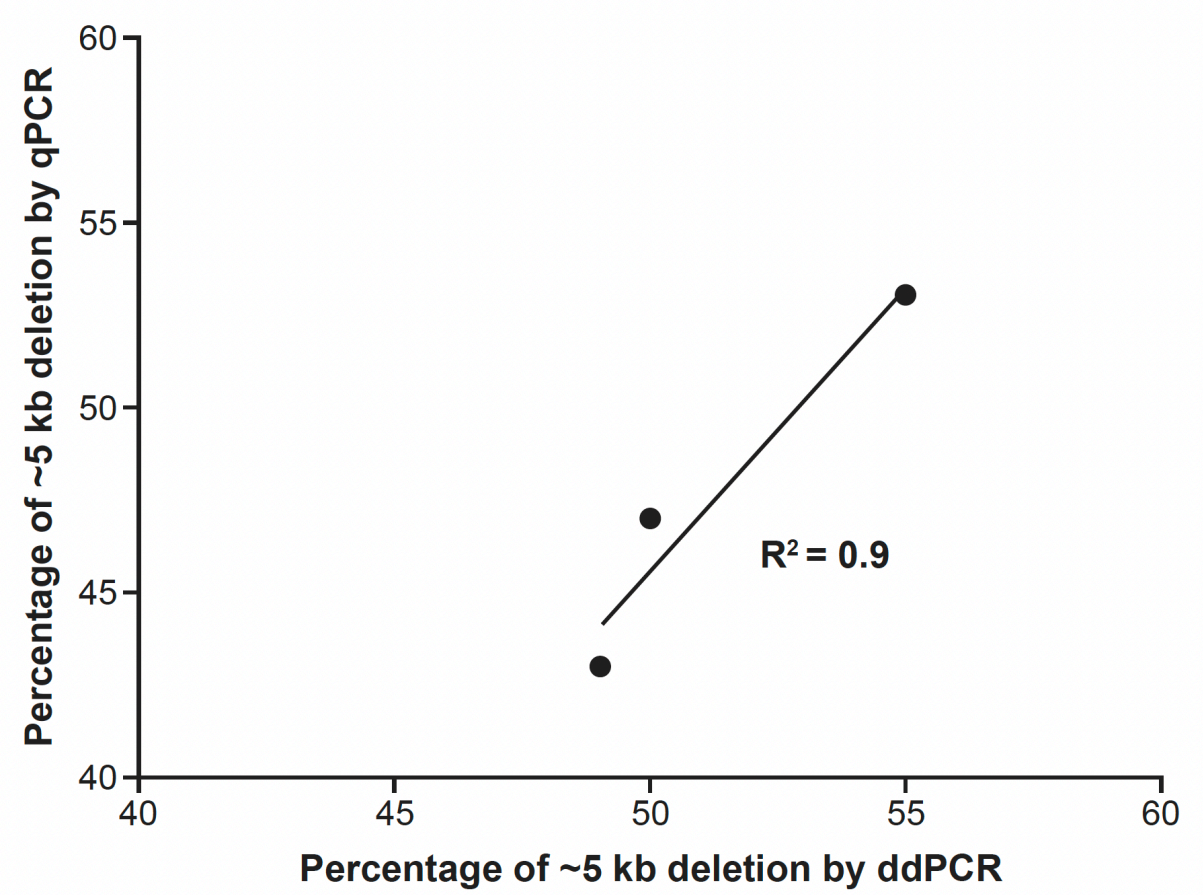


Table S1. gRNA targeting domains directed to the *HBG1* and *HBG2* promoter regions on chromosome 11

Targeting Domain ID	Target Promoter Region	gRNA Targeting Domain Sequence	Genomic Location (hg38) 1	Strand	Genomic Location (hg38) 2 (If present)	Strand
01	*HBG1*	AGUCCUGGUAUCCUCUAUGA	chr11:5250170-5250189	–		
02	*HBG1*	AAUUAGCAGUAUCCUCUUGG	chr11:5250064-5250083	–		
03	*HBG1*	AGAAUAAAUUAGAGAAAAAC	chr11:5250124-5250143	–		
04	*HBG1*	AAAAAUUAGCAGUAUCCUCU	chr11:5250067-5250086	–		
05	*HBG1*	AAAAUUAGCAGUAUCCUCUU	chr11:5250066-5250085	–		
06	*HBG1*	AAAAACUGGAAUGACUGAAU	chr11:5250110-5250129	–		
07	*HBG1*	CUCCCAUCAUAGAGGAUACC	chr11:5250164-5250183	+		
08	*HBG1*	GGAGAAGGAAACUAGCUAAA	chr11:5250148-5250167	–		
09	*HBG1*	GUUUCCUUCUCCCAUCAUAG	chr11:5250156-5250175	+		
10	*HBG1*	GGGAGAAGGAAACUAGCUAA	chr11:5250149-5250168	–		
11	*HBG1*	CACUGGAGCUAGAGACAAGA	chr11:5250214-5250233	–		
12	*HBG1*	AGAGACAAGAAGGUAAAAAA	chr11:5250204-5250223	–		
13	*HBG1*	AAAUUAGCAGUAUCCUCUUG	chr11:5250065-5250084	–		
14	*HBG1*	GUCCUGGUAUCCUCUAUGAU	chr11:5250169-5250188	–		
15	*HBG1*	GUAUCCUCUAUGAUGGGAGA	chr11:5250163-5250182	–		
16	*HBG2*	AUUAAGCAGCAGUAUCCUCU	chr11:5254991-5255010	–		
17	*HBG2*	AGAAUAAAUUAGAGAAAAAU	chr11:5255052-5255071	–		
18	*HBG2*	AGAAGUCCUGGUAUCUUCUA	chr11:5255101-5255120	–		
19	*HBG2*	UUAAGCAGCAGUAUCCUCUU	chr11:5254990-5255009	–		
20	*HBG2*	AAAAAUUGGAAUGACUGAAU	chr11:5255038-5255057	–		
21	*HBG2*	GGGAGAAGAAAACUAGCUAA	chr11:5255077-5255096	–		

22	*HBG2*	GGAGAAGAAAAC UAGCUAAA	chr11:5255076-5255095	–		
23	*HBG2*	CUCCCACCAUAGA AGAUACC	chr11:5255092-5255111	+		
24	*HBG2*	AGUCCUGGUAUC UUCUAUGG	chr11:5255098-5255117	–		
25	*HBG2*	GUCCUGGUAUCU UCUAUGGU	chr11:5255097-5255116	–		
26	*HBG2*	UAAGCAGCAGUA UCCUCUUG	chr11:5254989-5255008	–		
27	*HBG2*	AAGCAGCAGUAU CCUCUUGG	chr11:5254988-5255007	–		
28	*HBG1/HBG2*	CCUAGCCAGCCGC CGGCCCC	chr11:5249896-5249915	+	chr11:5254820-5254839	+
29	*HBG1/HBG2*	UAUCCAGUGAGG CCAGGGGC	chr11:5249911-5249930	–	chr11:5254835-5254854	–
30	*HBG1/HBG2*	CAUUGAGAUAGU GUGGGGAA	chr11:5250037-5250056	+	chr11:5254961-5254980	+
31	*HBG1/HBG2*	CCAGUGAGGCCA GGGGCCGG	chr11:5249908-5249927	–	chr11:5254832-5254851	–
32	*HBG1/HBG2*	GUGGGGAAGGGG CCCCCAAG	chr11:5250049-5250068	+	chr11:5254973-5254992	+
33	*HBG1/HBG2*	CCAGGGGCCGGC GGCUGGCU	chr11:5249899-5249918	–	chr11:5254823-5254842	–
34	*HBG1/HBG2*	UGAGGCCAGGGG CCGGCGGC	chr11:5249904-5249923	–	chr11:5254828-5254847	–
35	*HBG1/HBG2*	CAGUUCCACACAC UCGCUUC	chr11:5249847-5249866	–	chr11:5254771-5254790	–
36	*HBG1/HBG2*	CCGCCGGCCCCUG GCCUCAC	chr11:5249905-5249924	+	chr11:5254829-5254848	+
37	*HBG1/HBG2*	GUUUGCCUUGUC AAGGCUAU	chr11:5249950-5249969	+	chr11:5254874-5254893	+
38	*HBG1/HBG2*	GGCUAGGGAUGA AGAAUAAA	chr11:5249883-5249902	–	chr11:5254807-5254826	–
39	*HBG1/HBG2*	CAGGGGCCGGCG GCUGGCUA	chr11:5249898-5249917	–	chr11:5254822-5254841	–
40	*HBG1/HBG2*	ACUGGAUACUCU AAGACUAU	chr11:5249923-5249942	+	chr11:5254847-5254866	+
41	*HBG1/HBG2*	CCCUGGCUAAAC UCCACCCA	chr11:5249996-5250015	–	chr11:5254920-5254939	–
42	*HBG1/HBG2*	UUAGAGUAUCCA GUGAGGCC	chr11:5249917-5249936	–	chr11:5254841-5254860	–
43	*HBG1/HBG2*	CCCAUGGGUGGA GUUUAGCC	chr11:5249992-5250011	+	chr11:5254916-5254935	+
44	*HBG1/HBG2*	AGGCAAGGCUGG CCAACCCA	chr11:5249976-5249995	+	chr11:5254900-5254919	+

45	*HBG1/HBG2*	UAGAGUAUCCAGUGAGGCCA	chr11:5249916-5249935	–	chr11:5254840-5254859	–
46	*HBG1/HBG2*	UAUCUGUCUGAAACGGUCCC	chr11:5250013-5250032	–	chr11:5254937-5254956	–
47	*HBG1/HBG2*	AUUGAGAUAGUGUGGGGAAG	chr11:5250038-5250057	+	chr11:5254962-5254981	+
48	*HBG1/HBG2*	CUUCAUCCCUAGCCAGCCGC	chr11:5249889-5249908	+	chr11:5254813-5254832	+
49	*HBG1/HBG2*	GCUAUUGGUCAAGGCAAGGC	chr11:5249965-5249984	+	chr11:5254889-5254908	+
50	*HBG1/HBG2*	AUGCAAAUAUCUGUCUGAAA	chr11:5250020-5250039	–	chr11:5254944-5254963	–
51	*HBG1/HBG2*	GCAUUGAGAUAGUGUGGGGA	chr11:5250036-5250055	+	chr11:5254960-5254979	+
52	*HBG1/HBG2*	UGGUCAAGUUUGCCUUGUCA	chr11:5249943-5249962	+	chr11:5254867-5254886	+
53	*HBG1/HBG2*	GGCAAGGCUGGCCAACCCAU	chr11:5249977-5249996	+	chr11:5254901-5254920	+
54	*HBG1/HBG2*	ACGGCUGACAAAAGAAGUCC	chr11:5250185-5250204	–	chr11:5255113-5255132	–
55	*HBG1/HBG2*	CGAGUGUGUGGAACUGCUGA	chr11:5249851-5249870	+	chr11:5254775-5254794	+
56	*HBG1/HBG2*	CCUGGCUAAACUCCACCCAU	chr11:5249995-5250014	–	chr11:5254919-5254938	–
57	*HBG1/HBG2*	CUUGUCAAGGCUAUUGGUCA	chr11:5249956-5249975	+	chr11:5254880-5254899	+
58	*HBG1/HBG2*	AUAUUUGCAUUGAGAUAGUG	chr11:5250030-5250049	+	chr11:5254954-5254973	+
59	*HBG1/HBG2*	GCUAAACUCCACCCAUGGGU	chr11:5249991-5250010	–	chr11:5254915-5254934	–
60	*HBG1/HBG2*	ACGUUCCAGAAGCGAGUGUG	chr11:5249839-5249858	+	chr11:5254763-5254782	+
61	*HBG1/HBG2*	UAUUUGCAUUGAGAUAGUGU	chr11:5250031-5250050	+	chr11:5254955-5254974	+
62	*HBG1/HBG2*	GGAAUGACUGAAUCGGAACA	chr11:5250103-5250122	–	chr11:5255031-5255050	–
63	*HBG1/HBG2*	CUUGACCAAUAGCCUUGACA	chr11:5249958-5249977	–	chr11:5254882-5254901	–
64	*HBG1/HBG2*	CAAGGCUAUUGGUCAAGGCA	chr11:5249961-5249980	+	chr11:5254885-5254904	+
65	*HBG1/HBG2*	AAGGCUGGCCAACCCAUGGG	chr11:5249980-5249999	+	chr11:5254904-5254923	+
66	*HBG1/HBG2*	ACUCGCUUCUGGAACGUCUG	chr11:5249836-5249855	–	chr11:5254760-5254779	–

67	*HBG1/HBG2*	AUUUGCAUUGAGAUAGUGUG	chr11:5250032-5250051	+	chr11:5254956-5254975	+
68	*HBG1/HBG2*	ACUGAAUCGGAACAAGGCAA	chr11:5250097-5250116	–	chr11:5255025-5255044	–
69	*HBG1/HBG2*	CCAUGGGUGGAGUUUAGCCA	chr11:5249993-5250012	+	chr11:5254917-5254936	+
70	*HBG1/HBG2*	AGAGUAUCCAGUGAGGCCAG	chr11:5249915-5249934	–	chr11:5254839-5254858	–
71	*HBG1/HBG2*	GAGUGUGUGGAACUGCUGAA	chr11:5249852-5249871	+	chr11:5254776-5254795	+
72	*HBG1/HBG2*	UAGUCUUAGAGUAUCCAGUG	chr11:5249922-5249941	–	chr11:5254846-5254865	–

Guides were designated as *HBG1* or *HBG2* selective based on 100% sequence homology within the targeted region. Coordinates are provided in the UCSC genome browser position format. The selected guides (gRNA-8, gRNA-10, and gRNA-22) were found to have a one-nucleotide mismatch relative to the gRNA targeting domain sequence within the homologous *HBG1* (gRNA-22) or *HBG2* target site (gRNA-8 and gRNA-10).

Table S2. Busulfan administration

Study Day	Busulfan Total Daily Dose	Observed AUC
Participant 1		
Day −6	3.67 mg/kg	16 mg*h/L
Day −5	4.1 mg/kg	20.6 mg*h/L
Day −4	4.1 mg/kg	22.1 mg*h/L
Day −3	4.1 mg/kg	Not done
Participant 2		
Day −6	2.86 mg/kg	27.744 mg*h/L
Day −5	2.55 mg/kg	21.82 mg*h/L
Day −4	2.55 mg/kg	22.656 mg*h/L
Day −3	2.55 mg/kg	Not done
Participant 3		
Day −6	3.2 mg/kg	14.48 mg*h/L
Day −5	3.2 mg/kg	14.48 mg*h/L
Day −4	4.1 mg/kg	23.16 mg*h/L
Day −3	3.1 mg/kg	17.38 mg*h/L

Table S3. Off-target characterization of gRNA-68.

Method	Number of Sites Identified	Number of Sites Characterized	Confirmed Off-Targets
Computational	163	162[+]	0
Pan-Heme panel	352*	352	0
SITE-Seq	156	156[#]	0

*The number here refers to genes; [+]one 5-bp mismatch site was not characterized because of its repetitive nature. [#]Of these, 40 overlapped with the computationally derived sites. There were a total of 279 unique potential off-target sites.

Table S4. Potential off-target sites identified by both computational and SITE-Seq analysis for gRNA-68

	Sequence coordinate	Annotation
1	chr1:33477303-33477322	*ZSCAN20* intron
2	chr10:43645793-43645812	*ZNF32, ZNF32-AS3* intron
3	chr10:89056619-89056638	Intergenic
4	chr11:5250097-5250116	Intergenic
5	chr11:5255025-5255044	Intergenic
6	chr11:43428997-43429016	*TTC17* intron
7	chr11:129789715-129789734	Intergenic
8	chr12:190045-190064	Intergenic
9	chr13:98183649-98183669	*FARP1* intron
10	chr15:79413569-79413588	*TMED3* exon
11	chr17:10910316-10910335	Intergenic
12	chr17:30901624-30901643	*TEFM* intron
13	chr17:50766891-50766910	*ANKRD40CL* exon
14	chr18:21605369-21605388	Intergenic
15	chr18:41071436-41071455	Intergenic
16	chr18:43588394-43588413	Intergenic
17	chr18:75304356-75304375	Intergenic
18	chr19:55429168-55429187	*SHISA7* exon
19	chr2:7214577-7214596	Intergenic
20	chr2:85300579-85300598	*TCF7L1* intron
21	chr2:125044861-125044880	Intergenic
22	chr2:151189575-151189594	Intergenic
23	chr2:223639575-223639594	Intergenic
24	chr20:32930693-32930712	*EFCAB8* intron
25	chr22:17638029-17638046	Intergenic
26	chr22:31839366-31839385	*DEPDC5* intron
27	chr22:39514672-39514690	*MIEF1* exon
28	chr3:28154988-28155007	Intergenic
29	chr3:134478687-134478706	*ANAPC13* exon
30	chr3:153598641-153598660	*LINC02006* intron
31	chr4:11272770-11272790	Intergenic
32	chr4:52096951-52096970	*SPATA18* exon
33	chr4:103647034-103647053	*TACR3* intron
34	chr7:47656429-47656448	*C7orf65* intron
35	chr7:64192351-64192370	Intergenic
36	chr7:100214235-100214254	*STAG3* exon, *CASTOR3* intron
37	chr9:20865997-20866016	*FOCAD* intron
38	chr9:129933497-129933516	*FNBP1* intron
39	chrX:104749978-104749997	*IL1RAPL2* intron
40	chrX:147537296-147537315	Intergenic

Table S5. Genes targeted by the Pan-Heme v1.0 panel

RefSeq Gene Identifiers	Gene Coordinates (Human Genome Build hg19)	# Hybrid Capture Baits	Size of Targeted Region (bp)	Databases Used to Determine Coordinates	Bait Coverage of the Intended Targeted Regions
ABCG1	chr21:43621776-43720724	22	3420	Gencode, RefSeq, VEGA	98.450294
ABL1	chr9:133589697-133761080	14	3964	Gencode, RefSeq, VEGA	100
ACSL6	chr5:131283298-131347623	25	2847	Gencode, RefSeq, VEGA	100
ADAMTSL3	chr15:84324504-84706568	29	5656	Gencode, RefSeq, VEGA	100
ADARB2	chr10:1229123-1779354	10	2629	Gencode, RefSeq, VEGA	100
AKAP8	chr19:15465716-15490552	14	2400	Gencode, RefSeq	100
AKT1	chr14:105236668-105258990	13	1703	Gencode, RefSeq	100
ALPK2	chr18:56149045-56279039	12	6753	Gencode, RefSeq, VEGA	100
ANKLE2	chr12:133303818-133338394	14	3153	Gencode, RefSeq	100
ANO1	chr11:69924703-70034120	29	3910	Gencode, RefSeq	100
APC	chr5:112043405-112198243	19	9166	Gencode, RefSeq, VEGA	99.454506
ARFRP1	chr20:62331785-62338453	9	1139	Gencode, RefSeq, VEGA	100
ARHGEF12	chr11:120207943-120355796	41	5455	Gencode, RefSeq	100
ARID1A	chr1:27022885-27107257	20	7260	Gencode, RefSeq, VEGA	99.476585
ASMTL	chrY:1472152-1521743	13	2126	Gencode, RefSeq	100
ASMTL	chrX:1522152-1571743	13	2126	Gencode, RefSeq, VEGA	100

ASXL1	chr20:30946569-31025151	17	5054	Gencode, RefSeq, VEGA	99.48556
ATM	chr11:108098342-108236245	62	10411	Gencode, RefSeq	100
ATR	chr3:142168261-142297556	49	9120	Gencode, RefSeq, VEGA	99.08991
ATRX	chrX:76763819-77041497	37	8299	Gencode, RefSeq, VEGA	100
AURKA	chr20:54945204-54963263	8	1387	Gencode, RefSeq, VEGA	100
AXIN1	chr16:338112-397035	10	2847	Gencode, RefSeq, VEGA	100
B2M	chr15:45003735-45008550	3	443	Gencode, RefSeq, VEGA	100
B3GAT1	chr11:134251822-134257751	5	1144	Gencode, RefSeq	100
BAP1	chr3:52436294-52443904	17	2599	Gencode, RefSeq, VEGA	100
BCL2	chr18:60795848-60985909	2	793	Gencode, RefSeq, VEGA	100
BCL2L2	chr14:23776967-23778184	2	622	Gencode, RefSeq, VEGA	100
BCOR	chrX:39909159-39937192	16	5707	Gencode, RefSeq, VEGA	100
BIRC2	chr11:102220576-102248924	8	2032	Gencode, RefSeq	100
BIRC3	chr11:102195231-102207843	8	1975	Gencode, RefSeq	100
BIRC7	chr20:61867439-61870967	6	1205	Gencode, RefSeq, VEGA	100
BRAF	chr7:140426284-140624513	21	2799	Gencode, RefSeq, VEGA	100
BRCA1	chr17:41197685-41276123	24	6184	Gencode, RefSeq, VEGA	98.93273
BRCA2	chr13:32890588-32972917	26	10777	Gencode, RefSeq, VEGA	100
BTG1	chr12:92537846-92539321	2	556	Gencode, RefSeq	100

BTK	chrX:100604863-100630282	19	2366	Gencode, RefSeq, VEGA	100
C12ORF35	chr12:32133880-32145479	3	5304	Gencode, VEGA	100
CAMK4	chr5:110560172-110820174	11	1642	Gencode, RefSeq, VEGA	100
CARD11	chr7:2946262-2998150	24	3945	Gencode, RefSeq, VEGA	100
CARS	chr11:3022334-3078607	24	3095	Gencode, RefSeq, VEGA	100
CBL	chr11:119077118-119170501	16	3041	Gencode, RefSeq	100
CCND1	chr11:69456072-69466060	5	988	Gencode, RefSeq	100
CCND3	chr6:41903579-41909397	6	1151	Gencode, RefSeq, VEGA	100
CD36	chr7:80276047-80303473	12	1659	Gencode, RefSeq, VEGA	100
CDH5	chr16:66413231-66437082	11	2575	Gencode, RefSeq, VEGA	100
CDKN2A	chr9:21968218-21994463	5	1135	Gencode, RefSeq, VEGA	100
CEBPA	chr19:33792234-33793330	1	1097	Gencode, RefSeq, VEGA	100
CIT	chr12:120126027-120313982	49	7238	Gencode, RefSeq, VEGA	100
CREBBP	chr16:3777709-3930131	31	8153	Gencode, RefSeq, VEGA	98.94517
CRLF2	chrX:1314877-1331537	7	1133	Gencode, RefSeq, VEGA	78.905556
CRLF2	chrY:1264877-1281537	7	1133	Gencode, RefSeq	78.905556
CSF1	chr1:110453636-110467821	8	1825	Gencode, RefSeq, VEGA	96.76712
CSF1R	chr5:149433622-149466000	22	3449	Gencode, RefSeq, VEGA	100
CSNK1A1	chr5:148876406-148930537	13	1410	Gencode, RefSeq, VEGA	100
CTCF	chr16:67644726-67671785	10	2384	Gencode, RefSeq, VEGA	100
CTNNB1	chr3:41265550-41280843	14	2626	Gencode, RefSeq, VEGA	100

CUL4A	chr13:113863932-113917906	20	2704	Gencode, RefSeq, VEGA	100
CUX1	chr7:101459301-101926392	34	6043	Gencode, RefSeq, VEGA	100
CYLD	chr16:50783600-50830429	18	3245	Gencode, RefSeq	100
DAPK1	chr9:90113983-90322289	26	4888	Gencode, RefSeq, VEGA	100
DCHS1	chr11:6642714-6662854	20	10583	Gencode, RefSeq, VEGA	100
DNM2	chr19:10824133-10941733	24	3750	Gencode, RefSeq	100
DNMT3A	chr2:25457138-25536863	25	3388	Gencode, RefSeq, VEGA	100
DPYD	chr1:97544522-98386488	25	3716	Gencode, RefSeq, VEGA	100
DSC3	chr18:28574131-28622636	17	3058	Gencode, RefSeq	100
DUSP2	chr2:96809552-96811103	4	1025	Gencode, RefSeq, VEGA	100
DUSP27	chr1:167064077-167097855	5	3577	Gencode, RefSeq, VEGA	100
EIF2AK2	chr2:37334406-37374959	16	1997	Gencode, RefSeq, VEGA	97.946915
ELN	chr7:73442508-73483040	35	3177	Gencode, RefSeq, VEGA	100
EP300	chr22:41488999-41574970	31	7865	Gencode, RefSeq, VEGA	100
EPHA3	chr3:89156889-89528662	17	3318	Gencode, RefSeq, VEGA	100
EPHA7	chr6:93953134-94129069	17	3345	Gencode, RefSeq, VEGA	100
EPHB3	chr3:184280014-184299420	16	3317	Gencode, RefSeq, VEGA	100
ERBB2	chr17:37855803-37884307	28	4360	Gencode, RefSeq	100
ETV1	chr7:13935481-14028687	15	1951	Gencode, RefSeq, VEGA	100
ETV4	chr17:41605877-41622995	12	1695	Gencode, RefSeq	100
ETV6	chr12:11803052-12043990	8	1519	Gencode, RefSeq	100
EZH2	chr7:148504728-148544400	21	2876	Gencode, RefSeq, VEGA	100

FAF1	chr1:50907102-51425493	21	2415	Gencode, RefSeq, VEGA	100
FANCC	chr9:97863979-98011583	15	2127	Gencode, RefSeq, VEGA	100
FBXO31	chr16:87364884-87417360	9	1800	Gencode, RefSeq	100
FBXW7	chr4:153244023-153332965	14	2898	Gencode, RefSeq	100
FGF19	chr11:69514020-69518654	3	711	Gencode, RefSeq	100
FGF4	chr11:69588067-69589862	3	681	Gencode, RefSeq	100
FGFR1OP	chr6:167412891-167453658	14	1593	Gencode, RefSeq, VEGA	100
FGFR3	chr4:1795652-1809424	18	3360	Gencode, RefSeq, VEGA	100
FLT1	chr13:28877294-29068990	34	5019	Gencode, RefSeq, VEGA	97.88802
FLT3	chr13:28578179-28674657	25	3504	Gencode, RefSeq, VEGA	100
FLYWCH1	chr16:2979677-2998738	11	2871	Gencode, RefSeq	100
GATA1	chrX:48649507-48652685	5	1446	Gencode, RefSeq, VEGA	100
GATA2	chr3:128199852-128205884	5	1543	Gencode, RefSeq, VEGA	100
GNA13	chr17:63010365-63052721	4	1214	Gencode, RefSeq	100
GRK7	chr3:141497117-141535902	4	1742	Gencode, RefSeq, VEGA	100
H3F3A	chr1:226252043-226259190	3	561	Gencode, RefSeq, VEGA	100
HMGA1	chr6:34208548-34213402	4	1135	Gencode, RefSeq, VEGA	100
HNF1B	chr17:36047273-36104885	9	1947	Gencode, RefSeq, VEGA	100
HOXA3	chr7:27147524-27150269	2	1372	Gencode, RefSeq, VEGA	100
HOXA9	chr7:27203212-27219284	4	1266	Gencode, RefSeq, VEGA	100
HRAS	chr11:532626-534332	5	733	Gencode, RefSeq, VEGA	100
ICK	chr6:52869937-52906044	13	2159	Gencode, RefSeq, VEGA	100

ID3	chr1:23885441-23885927	2	400	Gencode, RefSeq, VEGA	100
IDH1	chr2:209101793-209116285	8	1408	Gencode, RefSeq, VEGA	100
IDH2	chr15:90627488-90645632	11	1579	Gencode, RefSeq, VEGA	100
IGF2	chr11:2154207-2161536	4	800	Gencode, RefSeq, VEGA	100
IGFBP3	chr7:45954409-45961308	5	1197	Gencode, RefSeq, VEGA	100
IKZF1	chr7:50358648-50468335	8	2085	Gencode, RefSeq, VEGA	100
IKZF3	chr17:37922033-38020389	8	1690	Gencode, RefSeq, VEGA	100
IL7R	chr5:35857070-35876598	8	1716	Gencode, RefSeq, VEGA	100
IMP4	chr2:131100496-131104051	10	1167	Gencode, RefSeq, VEGA	100
IRAK1	chrX:153276381-153285273	14	2825	Gencode, RefSeq, VEGA	100
IRF8	chr16:85936612-85954898	8	1441	Gencode, RefSeq, VEGA	100
JAK1	chr1:65300235-65351957	24	3945	Gencode, RefSeq, VEGA	100
JAK2	chr9:5021978-5126801	23	3859	Gencode, RefSeq, VEGA	100
JAK3	chr19:17937542-17955236	23	3987	Gencode, RefSeq	100
KDM4C	chr9:6720939-7174739	25	4032	Gencode, RefSeq, VEGA	96.37897
KDM5A	chr12:394612-498267	29	5698	Gencode, RefSeq	100
KDM6A	chrX:44732788-44970666	31	5090	Gencode, RefSeq, VEGA	100
KIT	chr4:55524172-55604733	22	3379	Gencode, RefSeq, VEGA	100
KLF2	chr19:16435725-16437852	3	1128	Gencode, RefSeq	99.645386
KRAS	chr12:25362719-25398328	5	787	Gencode, RefSeq	100
LMO1	chr11:8246153-8290005	5	593	Gencode, RefSeq	100
LRRK2	chr12:40618924-40761577	52	8746	Gencode, RefSeq, VEGA	100

MAGEC3	chrX:140926092-140985628	8	2705	Gencode, RefSeq, VEGA	100
MALT1	chr18:56338866-56415084	17	2815	Gencode, RefSeq, VEGA	100
MAP3K4	chr6:161412954-161551927	29	5469	Gencode, RefSeq, VEGA	100
MATK	chr19:3778171-3789355	14	1879	Gencode, RefSeq	100
MED12	chrX:70338595-70362078	45	7498	Gencode, RefSeq, VEGA	99.54655
MED12L	chr3:150804704-151150602	45	7494	Gencode, RefSeq, VEGA	100
MEF2B	chr19:19256493-19267955	9	1565	Gencode, RefSeq, VEGA	91.5655
MLL2	chr12:49415553-49449117	54	17694	Gencode	99.72872
MLL3	chr7:151833907-152132881	62	16276	Gencode, VEGA	99.637505
MLLT4	chr6:168227803-168370625	36	6563	Gencode, RefSeq, VEGA	100
MPL	chr1:43803510-43818453	12	2233	Gencode, RefSeq, VEGA	100
MSH2	chr2:47630321-47739583	18	3344	Gencode, RefSeq, VEGA	97.996414
MSH6	chr2:48010363-48034009	10	4318	Gencode, RefSeq, VEGA	100
MTMR8	chrX:63444205-63615253	17	3457	Gencode, RefSeq, VEGA	98.813995
MUC4	chr3:195474037-195538698	26	16784	Gencode, RefSeq, VEGA	100
MYC	chr8:128748830-128753214	3	1425	Gencode, RefSeq, VEGA	100
MYD88	chr3:38180143-38182787	5	1054	Gencode, RefSeq, VEGA	100
MYO3B	chr2:171034788-171509641	38	4987	Gencode, RefSeq, VEGA	97.152596
MYOM2	chr8:1998871-2092915	36	5118	Gencode, RefSeq, VEGA	100
NCOR2	chr12:124809938-124979807	50	8713	Gencode, RefSeq, VEGA	100
NEK8	chr17:27055822-27069015	14	2519	Gencode, RefSeq	100

NF1	chr17:29422318-29705959	60	9902	Gencode, RefSeq, VEGA	100
NF2	chr22:29999978-30090808	17	2221	Gencode, RefSeq, VEGA	100
NOTCH1	chr9:139390513-139440248	34	8348	Gencode, RefSeq, VEGA	100
NOTCH2	chr1:120457919-120612085	38	9195	Gencode, RefSeq, VEGA	100
NPM1	chr5:170814943-170837579	12	1134	Gencode, RefSeq, VEGA	93.1217
NR3C1	chr5:142658919-142780414	9	2576	Gencode, RefSeq, VEGA	100
NRAS	chr1:115251146-115258791	4	650	Gencode, RefSeq, VEGA	100
NTRK1	chr1:156785612-156851444	19	2902	Gencode, RefSeq, VEGA	100
PARP1	chr1:224102731-226595640	25	3828	Gencode, RefSeq, VEGA	96.31661
PAX5	chr9:36840516-37034038	11	1484	Gencode, RefSeq, VEGA	100
PCDH15	chr10:55566329-56424032	45	8504	Gencode, RefSeq, VEGA	98.97695
PDGFRA	chr4:55106210-55161449	24	3930	Gencode, RefSeq, VEGA	97.88804
PDGFRB	chr5:149495316-149516620	22	3795	Gencode, RefSeq, VEGA	100
PDPK1	chr16:2587990-2647778	14	2069	Gencode, RefSeq, VEGA	100
PHF6	chrX:133511638-133559370	9	1387	Gencode, RefSeq, VEGA	100
PHLPP2	chr16:71674793-71748708	20	4571	Gencode, RefSeq	100
PIK3CA	chr3:178916604-178952162	20	3609	Gencode, RefSeq, VEGA	100
PIM1	chr6:37138069-37141877	6	1335	Gencode, RefSeq, VEGA	100
PKN2	chr1:89150254-89299141	24	3593	Gencode, RefSeq, VEGA	100
PMS1	chr2:190656526-190742172	16	3397	Gencode, RefSeq, VEGA	97.05623

POLD1	chr19:50902096-50921214	26	3847	Gencode, RefSeq	100
PRDM1	chr6:106534419-106555371	9	2715	Gencode, RefSeq, VEGA	100
PRDM16	chr1:2985814-3350385	19	4238	Gencode, RefSeq, VEGA	99.45729
PRDM9	chr5:23509133-23527892	10	2885	Gencode, RefSeq, VEGA	100
PRKACG	chr9:71627943-71629018	1	1076	Gencode, RefSeq, VEGA	100
PRKCZ	chr1:1982060-2116458	21	2392	Gencode, RefSeq, VEGA	100
PROZ	chr13:113812965-113826429	9	1449	Gencode, RefSeq, VEGA	100
PTEN	chr10:89624217-89725239	9	1392	Gencode, RefSeq, VEGA	100
PTPN11	chr12:112856906-112942578	16	2112	Gencode, RefSeq, VEGA	100
PTPRD	chr9:8317864-8733853	42	6775	Gencode, RefSeq, VEGA	100
RB1	chr13:48878039-49054217	27	3327	Gencode, RefSeq, VEGA	100
RBM15	chr1:110882018-110888986	3	3041	Gencode, RefSeq, VEGA	100
RET	chr10:43572697-43623727	20	3777	Gencode, RefSeq, VEGA	100
RNF213	chr17:78237471-78367308	69	17473	Gencode, RefSeq, VEGA	100
ROCK2	chr2:11323553-11484272	33	4836	Gencode, RefSeq, VEGA	100
RPS2	chr16:2012089-2014636	6	1002	Gencode, RefSeq, VEGA	100
RPTOR	chr17:78519420-78938140	34	4724	Gencode, RefSeq	100
RSPO4	chr20:940990-982817	5	805	Gencode, RefSeq, VEGA	100
RUNX1	chr21:36164266-36421206	12	1922	Gencode, RefSeq, VEGA	100
SDHA	chr5:218461-256545	16	2449	Gencode, RefSeq, VEGA	100
SETBP1	chr18:42281302-42643673	6	5100	Gencode, RefSeq, VEGA	100

SF3B1	chr2:198257017-198299733	27	4585	Gencode, RefSeq, VEGA	100
SGK1	chr6:134491396-134638608	17	2498	Gencode, RefSeq, VEGA	100
SH2B3	chr12:111855940-111886116	8	2014	Gencode, RefSeq	100
SLC19A1	chr21:46916239-46957883	9	2514	Gencode, RefSeq, VEGA	100
SLC1A2	chr11:35282431-35441497	12	1970	Gencode, RefSeq, VEGA	100
SMAD2	chr18:45368188-45423137	10	1604	Gencode, RefSeq, VEGA	100
SMAD4	chr18:48573407-48604847	13	1999	Gencode, RefSeq, VEGA	100
SMARCB1	chr22:24129347-24176377	9	1392	Gencode, RefSeq, VEGA	100
SMO	chr7:128828983-128852302	13	2647	Gencode, RefSeq, VEGA	100
SNTG2	chr2:946673-1371256	18	2018	Gencode, RefSeq, VEGA	97.12587
SOCS1	chr16:11348690-11349345	1	656	Gencode, RefSeq, VEGA	100
SRSF2	chr17:74731930-74733333	3	832	Gencode, RefSeq	100
STAT3	chr17:40467753-40500544	23	2773	Gencode, RefSeq, VEGA	100
STAT5B	chr17:40353746-40384155	18	2724	Gencode, RefSeq, VEGA	100
SULT1A1	chr16:28617132-28634528	9	1206	Gencode, RefSeq, VEGA	100
TAF1L	chr9:32630087-32635587	1	5501	Gencode, RefSeq, VEGA	100
TBL1XR1	chr3:176743276-176782775	14	1825	Gencode, RefSeq, VEGA	100
TERT	chr5:1253833-1295114	16	3719	Gencode, RefSeq, VEGA	100
TET2	chr4:106111617-106197686	10	6365	Gencode, RefSeq, VEGA	100

TLL2	chr10:98127835-98273452	21	3468	Gencode, RefSeq, VEGA	100
TMEM30A	chr6:75965808-75994364	7	1226	Gencode, RefSeq, VEGA	100
TMPRSS2	chr21:42838059-42879941	14	1870	Gencode, RefSeq, VEGA	100
TNFAIP3	chr6:138192355-138202466	9	2558	Gencode, RefSeq, VEGA	100
TNFRSF14	chr1:2488094-2494722	11	1321	Gencode, RefSeq, VEGA	100
TNK2	chr3:195591042-195623107	23	4483	Gencode, RefSeq, VEGA	100
TP53	chr17:7565247-7579922	14	1697	Gencode, RefSeq, VEGA	94.34296
TP73	chr1:3598920-3649653	14	2230	Gencode, RefSeq, VEGA	100
TRAF2	chr9:139780940-139820363	11	1954	Gencode, RefSeq, VEGA	96.72466
TRAF3	chr14:103336529-103372131	10	1907	Gencode, RefSeq	100
TRAF7	chr16:2213912-2226585	20	2413	Gencode, RefSeq, VEGA	100
TRIP11	chr14:92436007-92506039	21	6360	Gencode, RefSeq	100
TSC2	chr16:2098261-2138621	42	6297	Gencode, RefSeq, VEGA	100
TSC22D1	chr13:45008091-45150220	6	3605	Gencode, RefSeq, VEGA	100
TYK2	chr19:10461500-10489092	23	4324	Gencode, RefSeq	100
TYMS	chr18:657733-673007	7	1082	Gencode, RefSeq, VEGA	100
U2AF1	chr21:44513202-44527614	9	970	Gencode, RefSeq, VEGA	93.40206
VPREB1	chr22:22599205-22599759	2	489	Gencode, RefSeq, VEGA	100
WDFY4	chr10:49917741-50190906	64	11323	Gencode, RefSeq, VEGA	100
WDR90	chr16:699407-717599	41	6175	Gencode, RefSeq, VEGA	100
WHSC1	chr4:1902372-1980646	23	4873	Gencode, RefSeq, VEGA	100

WT1	chr11:32410594-32456901	11	1784	Gencode, RefSeq, VEGA	100
WWOX	chr16:78133666-79246041	11	1937	Gencode, RefSeq	100
ZRSR2	chrX:15808609-15841375	12	1930	Gencode, RefSeq, VEGA	86.476685
ABI3BP	chr3:100469329-100712259	67	6727	Gencode, RefSeq, VEGA	100
ACAD9	chr3:128598525-128631678	18	2612	Gencode, RefSeq, VEGA	100
ACE	chr17:61554446-61574931	27	4688	Gencode, RefSeq, VEGA	99.744026
ALOX12B	chr17:7976079-7990770	15	2406	Gencode, RefSeq, VEGA	100
ATP2C2	chr16:84402212-84497348	28	3488	Gencode, RefSeq	100
BAI3	chr6:69348558-70098793	31	5287	Gencode, RefSeq, VEGA	100
BRD4	chr19:15349178-15383920	21	4747	Gencode, RefSeq	100
BTG2	chr1:203274725-203276576	2	517	Gencode, RefSeq, VEGA	100
BTLA	chr3:112184945-112218215	5	970	Gencode, RefSeq, VEGA	100
C6ORF103	chr6:146920167-147136363	39	6060	Gencode, VEGA	100
C6ORF27	chr6:31733361-31744566	16	3024	Gencode, VEGA	100
CAD	chr2:27440410-27466396	44	7726	Gencode, RefSeq, VEGA	100
CALR	chr19:13049484-13054737	9	1434	Gencode, RefSeq	100
CAMTA1	chr1:6845581-7826592	27	5735	Gencode, RefSeq, VEGA	98.86661
CCT6B	chr17:33255057-33288422	14	1873	Gencode, RefSeq	100
CD200	chr3:112052050-112080407	7	1025	Gencode, RefSeq, VEGA	100
CD274	chr9:5456104-5467872	6	993	Gencode, RefSeq, VEGA	100
CD70	chr19:6583301-6591023	4	841	Gencode, RefSeq	92.98454

CD79A	chr19:42381365-42385057	5	781	Gencode, RefSeq	100
CD79B	chr17:62006576-62009631	6	813	Gencode, RefSeq	100
CD83	chr6:14118033-14135477	5	718	Gencode, RefSeq, VEGA	100
CDC73	chr1:193091321-193219852	17	1936	Gencode, RefSeq, VEGA	100
CDH17	chr8:95140458-95206923	17	2839	Gencode, RefSeq, VEGA	100
CDKN2B	chr9:22005976-22008962	2	538	Gencode, RefSeq, VEGA	100
CNOT3	chr19:54646705-54659155	20	3099	Gencode, RefSeq, VEGA	100
CNTNAP5	chr2:124783218-125671875	24	4401	Gencode, RefSeq, VEGA	100
COL4A2	chr13:110960241-111164548	52	6296	Gencode, RefSeq, VEGA	100
CSF3R	chr1:36931687-36945107	16	3017	Gencode, RefSeq, VEGA	100
CXCR4	chr2:136872429-136875640	2	1126	Gencode, RefSeq, VEGA	100
CYP4F22	chr19:15636138-15662292	12	1836	Gencode, RefSeq	100
DDX3X	chrX:41193496-41206982	18	2526	Gencode, RefSeq, VEGA	100
DIS3	chr13:73333923-73355980	22	3385	Gencode, RefSeq, VEGA	100
DLGAP1	chr18:3499173-3880078	14	3535	Gencode, RefSeq, VEGA	100
DNMT3A	chr2:25457138-25536863	25	3388	Gencode, RefSeq, VEGA	100
DPP10	chr2:115200346-116599931	30	3175	Gencode, RefSeq, VEGA	100
DTX1	chr12:113495988-113534754	9	2043	Gencode, RefSeq	99.216835
EBF1	chr5:158126109-158526496	17	2130	Gencode, RefSeq, VEGA	100
ECT2L	chr6:139134402-139223774	20	3115	Gencode, RefSeq, VEGA	100
EED	chr11:85956262-85989577	13	1665	Gencode, RefSeq	94.29429

EHD1	chr11:64621795-64647112	6	1767	Gencode, RefSeq, VEGA	100
EIF2C4	chr1:36274008-36319216	18	2946	Gencode, VEGA	98.67617
EML2	chr19:46112911-46148712	26	3230	Gencode, RefSeq	98.26626
ENTPD3	chr3:40429539-40469009	12	1959	Gencode, RefSeq, VEGA	99.795815
EPHB2	chr1:23037466-23240373	19	3628	Gencode, RefSeq, VEGA	96.63727
ETS1	chr11:128332246-128443035	10	1744	Gencode, RefSeq	100
FAM129B	chr9:130269114-130341226	15	2557	Gencode, RefSeq, VEGA	100
FAM46C	chr1:118165481-118166676	1	1196	Gencode, RefSeq, VEGA	100
FAM5C	chr1:190067138-190424030	7	2441	Gencode, RefSeq, VEGA	100
FAT1	chr4:187509736-187630991	29	14602	Gencode, RefSeq, VEGA	100
FAT4	chr4:126237557-126412933	18	15434	Gencode, RefSeq, VEGA	100
FGL2	chr7:76825586-76829120	2	1360	Gencode, RefSeq, VEGA	100
FHIT	chr3:59737942-60522705	5	544	Gencode, RefSeq, VEGA	100
FTCD	chr21:47556358-47575447	15	2145	Gencode, RefSeq, VEGA	100
GATA3	chr10:8097609-8115996	5	1437	Gencode, RefSeq, VEGA	100
GNA12	chr7:2770805-2883805	7	1715	Gencode, RefSeq, VEGA	100
GRIK5	chr19:42503013-42569932	20	3775	Gencode, RefSeq, VEGA	100
HIST1H1E	chr6:26156609-26157288	1	680	Gencode, RefSeq, VEGA	100
HLA-A	chr6:29910321-29913242	8	1276	Gencode, RefSeq, VEGA	100
HNRNPK	chr9:86584255-86593177	16	1817	Gencode, RefSeq, VEGA	100
IKZF2	chr2:213872074-214014917	12	2065	Gencode, RefSeq, VEGA	100
IL9R	chrX:155227399-155240084	10	1989	Gencode, RefSeq, VEGA	100

IL9R	chrY:59330405-59343090	10	1989	Gencode, RefSeq	100
KIFC3	chr16:57792786-57836934	22	3163	Gencode, RefSeq, VEGA	100
KLHL6	chr3:183209705-183273451	7	2006	Gencode, RefSeq, VEGA	100
LEF1	chr4:108969818-109088933	13	1635	Gencode, RefSeq, VEGA	100
MAGED1	chrX:51637391-51645036	12	2745	Gencode, RefSeq, VEGA	100
MAP3K6	chr1:27681799-27693098	30	4514	Gencode, RefSeq, VEGA	100
MAPK1	chr22:22123483-22221740	8	1243	Gencode, RefSeq, VEGA	100
MYB	chr6:135502642-135539128	19	3013	Gencode, RefSeq, VEGA	90.8397
MYH10	chr17:8379112-8526574	42	6882	Gencode, RefSeq, VEGA	100
NBEAL1	chr2:203881098-204090883	55	9219	Gencode, RefSeq, VEGA	98.850204
NRXN2	chr11:64374658-64481181	24	5945	Gencode, RefSeq, VEGA	100
NT5C2	chr10:104849419-104934725	21	2269	Gencode, RefSeq, VEGA	100
P2RX5	chr17:3575738-3599309	13	1573	Gencode, RefSeq, VEGA	100
P2RY2	chr11:72945195-72946348	1	1154	Gencode, RefSeq	100
PC	chr11:66616360-66639640	21	4179	Gencode, RefSeq	100
PCBP1	chr2:70314866-70315956	1	1091	Gencode, RefSeq, VEGA	100
PCLO	chr7:82387881-82791918	29	16306	Gencode, RefSeq, VEGA	100
PDGFRA	chr4:55106210-55161449	24	3930	Gencode, RefSeq, VEGA	97.88804
PER3	chr1:7844928-7902825	21	4057	Gencode, RefSeq, VEGA	100
PIK3R1	chr5:67522494-67593439	19	2779	Gencode, RefSeq, VEGA	100
POLRMT	chr19:617264-633522	21	4134	Gencode, RefSeq	100
POR	chr7:75583301-75615809	17	2784	Gencode, RefSeq, VEGA	100

POT1	chr7:124462572-124540896	20	2450	Gencode, RefSeq, VEGA	93.83674
PTPRN	chr2:220154938-220174064	25	3500	Gencode, RefSeq, VEGA	100
RAD21	chr8:117859729-117878978	13	2156	Gencode, RefSeq, VEGA	100
RANBP6	chr9:6012280-6015617	1	3338	Gencode, RefSeq, VEGA	100
RELN	chr7:103113249-103629813	66	11853	Gencode, RefSeq, VEGA	100
RHOA	chr3:49397632-49413032	6	944	Gencode, RefSeq, VEGA	70.12712
RIPK1	chr6:3077048-3113583	10	2216	Gencode, RefSeq, VEGA	100
RPL10	chrX:153626851-153637498	9	1250	Gencode, RefSeq, VEGA	100
RPL5	chr1:93297662-93307432	8	1058	Gencode, RefSeq, VEGA	100
RUNX1T1	chr8:92972460-93115122	20	2750	Gencode, RefSeq, VEGA	98.07272
SALL3	chr18:76740265-76757332	3	3963	Gencode, RefSeq, VEGA	99.1673
SAMHD1	chr20:35521325-35580056	16	2205	Gencode, RefSeq, VEGA	100
SAPS2	chr22:50832328-50882686	22	3347	VEGA	100
SBF1	chr22:50885561-50913451	42	6568	Gencode, RefSeq, VEGA	97.8989
SETD1B	chr12:122242634-122268155	17	6112	Gencode, RefSeq	100
SETD2	chr3:47058573-47205424	27	8312	Gencode, RefSeq, VEGA	98.97738
SLC29A2	chr11:66130897-66139072	12	1611	Gencode, RefSeq	100
SMARCA4	chr19:11094818-11172502	39	6021	Gencode, RefSeq	99.850525
SMC1A	chrX:53407014-53449559	26	4403	Gencode, RefSeq, VEGA	100
SMC3	chr10:112327565-112364070	29	4234	Gencode, RefSeq, VEGA	100
SPEN	chr1:16174553-16265932	16	11596	Gencode, RefSeq, VEGA	100

SRC	chr20:36012547-36031792	12	1869	Gencode, RefSeq, VEGA	100
STAG2	chrX:123156468-123234457	34	4541	Gencode, RefSeq, VEGA	100
STAT5B	chr17:40353746-40384155	18	2724	Gencode, RefSeq, VEGA	100
SUZ12	chr17:30264256-30326032	16	2540	Gencode, RefSeq, VEGA	100
SYNE1	chr6:152443561-152949476	151	30352	Gencode, RefSeq, VEGA	100
SYNGAP1	chr6:33388032-33419693	21	4598	Gencode, RefSeq, VEGA	98.10787
SYPL1	chr7:105732241-105752985	7	1004	Gencode, RefSeq, VEGA	89.64143
TAL1	chr1:47685382-47691570	3	1062	Gencode, RefSeq, VEGA	100
TBC1D9B	chr5:179289532-179334831	23	4406	Gencode, RefSeq, VEGA	100
TLR2	chr4:154624050-154626424	1	2375	Gencode, RefSeq, VEGA	100
TNFAIP3	chr6:138192355-138202466	9	2558	Gencode, RefSeq, VEGA	100
TOX	chr8:59720296-60031556	9	1761	Gencode, RefSeq, VEGA	100
TP63	chr3:189349295-189612301	17	2756	Gencode, RefSeq, VEGA	91.43686
TPST2	chr22:26924177-26937606	4	1439	Gencode, RefSeq, VEGA	100
U2AF2	chr19:56166461-56185444	12	1668	Gencode, RefSeq	100
UBR5	chr8:103266520-103424472	59	9580	Gencode, RefSeq, VEGA	100
XBP1	chr22:29191153-29196522	6	1369	Gencode, RefSeq, VEGA	100
XPO1	chr2:61705945-61761042	24	3717	Gencode, RefSeq, VEGA	100
ZFHX4	chr8:77616314-77776811	13	11216	Gencode, RefSeq, VEGA	99.643364
ZMYM3	chrX:70460756-70473115	24	4617	Gencode, RefSeq, VEGA	100

ZNF229	chr19:44932468-44947041	4	2558	Gencode, RefSeq	100
CCND1	chr11:69456072-69466060	5	988	Gencode, RefSeq	100
CYP3A4	chr7:99355746-99381714	13	1811	Gencode, RefSeq, VEGA	100
PRPF8	chr17:1554086-1587875	44	7909	Gencode, RefSeq, VEGA	99.60804
SF1	chr11:64532851-64545874	14	2816	Gencode, RefSeq, VEGA	94.140625

Table S6. Fusions/translocations targeted by the Pan-Heme v1.0 panel

RefSeq Gene Identifiers (>> Translocation from Gene A to Gene B, C, etc.)	Interval Coordinates (Human Genome Build hg19)	# Hybrid Capture Baits	Size of Targeted Region (bp)	Databases Used to Determine Coordinates	Bait Coverage of the Intended Targeted Regions
ABL1 >> NUP214; BCR ; ETV6	chr9:133710862-133729500	1	18639	CustomRegion	85.45523
ALK >> ATIC; TFG; TPM4; TPM3; MSN; CLTC; NPM1	chr2:29446344-29448376	1	2033	CustomRegion	100
ATF1 >> EWSR1; FUS	chr12:51203322-51207842	1	4521	CustomRegion	74.25348
ATIC >> ALK	chr2:216191651-216197154	1	5504	CustomRegion	80.19622
BCR >> ABL1; FGFR1	chr22:23603677-23610644	1	6968	CustomRegion	88.719864
BCR >> ABL1; JAK2	chr22:23631758-23632575	1	818	CustomRegion	100
CBFB >> MYH11	chr16:67116192-67132662	1	16471	CustomRegion	75.10169
CLTC >> ALK	chr17:57768022-57771138	1	3117	CustomRegion	89.18832
ERG >> EWSR1; FUS	chr21:39755795-39762966	1	7172	CustomRegion	99.49805
ETV6 >> PDGFRB; NTRK3	chr12:12006445-12022407	1	15963	CustomRegion	88.91813
ETV6 >> SYK; ABL1; RUNX1; ABL2	chr12:12022853-12037428	1	14576	CustomRegion	92.48765
EWSR1 >> ETV1; WT1; ETV4; FLI1; ERG; ATF1	chr22:29683073-29684644	1	1572	CustomRegion	94.783714
EWSR1 >> WT1; ERG	chr22:29678496-29682961	1	4466	CustomRegion	56.09046
FGFR1OP >> FGFR1	chr6:167424331-167427045	1	2715	CustomRegion	96.05893
FIP1L1 >> PDGFRA	chr4:54292082-54294243	1	2162	CustomRegion	94.07955
FUS >> CREB3; L2DD; IT3; ATF1	chr16:31195667-31196309	1	643	CustomRegion	76.82737

FUS >> DDIT3; CREB3L1; ERG	chr16:31198107-31199695	1	1589	CustomRegion	100
HIP1 >> PDGFRB	chr7:75167495-75168692	1	1198	CustomRegion	92.40401
HOXA9 >> MSI2; NUP98	chr7:27203410-27204546	1	1137	CustomRegion	100
ITK >> SYK	chr5:156641280-156644926	1	3647	CustomRegion	100
MLL >> ARHGEF12; ELL; MLLT4; MLLT6; MLLT10; MLLT3; EP300; GMPS	chr11:118352757-118353186	1	430	CustomRegion	100
MLL >> EPS15; ARHGEF12; ELL; MLLT4; MLLT6; MLLT3; EP300; SH3; GL1; GAS7; GMPS	chr11:118353160-118354947	1	1788	CustomRegion	57.997765
MLL >> GAS7; SEPT6; MLLT4; MLLT1; LPP; MLLT3; EP300; CREBBP; GMPS	chr11:118355640-118359378	1	3739	CustomRegion	79.673706
MN1 >> ETV6	chr22:28147034-28192800	1	45767	CustomRegion	92.35475
MSI2 >> HOXA9	chr17:55693395-55704639	1	11245	CustomRegion	91.13384
NPM1 >> MLF1	chr5:170819770-170827206	1	7437	CustomRegion	67.16418
NUP98 >> DDX10; HOXA13; HOXD13; RAP1; GDS1; NSD1; HOXA9	chr11:3756504-3765788	1	9285	CustomRegion	70.2531
PAX5 >> ETV6	chr9:37002723-37006519	1	3797	CustomRegion	96.91862
PCM1 >> JAK2	chr8:17830146-17838149	1	8004	CustomRegion	98.95052

PCM1 >> RET	chr8:17851078-17863821	1	12744	CustomRegion	53.64093
PDGFRB >> TRIP11; HIP1; ETV6	chr5:149506127-149509369	1	3243	CustomRegion	94.38791
PICALM >> MLLT10	chr11:85692221-85692837	1	617	CustomRegion	90.11346
RARA >> PML; NUMA1	chr17:38499069-38504617	1	5549	CustomRegion	100
RBM15 >> MKL1	chr1:110884840-110888210	1	3371	CustomRegion	100
RUNX1 >> CBFA2T3	chr21:36206848-36231820	1	24973	CustomRegion	93.90141
SYK >> ITK; ETV6	chr9:93627329-93636535	1	9207	CustomRegion	95.145
TCF3 >> HLF	chr19:1615770-1619159	1	3390	CustomRegion	80.82596
TFG >> NTRK1; ALK	chr3:100451466-100455469	1	4004	CustomRegion	87.26274
TPM4 >> ALK	chr19:16204358-16204543	1	186	CustomRegion	100
CCND1	chr11:69455845-69469252	1	13408	Gencode, RefSeq	100
CDK6	chr7:92234225-92465951	1	231727	Gencode, RefSeq, VEGA	90.39128
FGFR1	chr8:38268646-38326362	1	57717	Gencode, RefSeq, VEGA	86.74048
JAK2	chr9:4985023-5128193	1	143171	Gencode, RefSeq, VEGA	72.427376
TYK2	chr19:10461194-10491362	1	30169	Gencode, RefSeq	58.44078
WAC	chr10:28821412-28912051	1	90640	Gencode, RefSeq, VEGA	88.51059

Table S7. Post-infusion adverse events

Participant 1			
Pre-conditioning/Screening Period			
Adverse Event	**Study Day**	**AE outcome**	**Severity**
Epstein-Barr virus infection/reactivation	-376	Recovered/Resolved	Mild
Human herpes virus infection/reactivation	-376	Recovered/Resolved	Mild
Anxiety	-360	Recovered/Resolved	Mild
Hyperphosphatemia	-360	Recovered/Resolved	Mild
Catheter site pain	-357	Recovered/Resolved	Moderate
Cough	-328	Recovered/Resolved	Mild
Migraine	-291	Recovered/Resolved	Mild
Limb injury	-276	Recovered/Resolved	Moderate
Catheter site pain	-234	Recovered/Resolved	Moderate
Musculoskeletal chest pain	-232	Recovered/Resolved	Mild
Hidradenitis	-232	Recovered/Resolved	Severe
Dry eyes	-231	Recovered/Resolved	Mild
Insomnia	-231	Recovered/Resolved	Mild
Dyspepsia	-230	Recovered/Resolved	Moderate
Anemia	-229	Recovered/Resolved	Severe
Constipation	-229	Recovered/Resolved	Mild
Catheter site pain	-161	Recovered/Resolved	Moderate
Nausea	-159	Recovered/Resolved	Mild
Hypoxia	-158	Recovered/Resolved	Moderate
Anemia	-158	Recovered/Resolved	Severe
Soft palate lesions	-158	Recovered/Resolved	Mild
Soft tissue infection	-81	Recovered/Resolved	Moderate
Dental caries	-19	Recovered/Resolved	Mild
Blood bilirubin increased	-13	Recovered/Resolved	Mild
Pain in extremity	-7	Recovered/Resolved	Mild
Conditioning Period			
Protein total decreased	-6	Recovered/Resolved	Mild
Nausea	-5	Recovered/Resolved	Mild
Anxiety	-5	Recovered/Resolved	Moderate
Aspartate aminotransferase increased	-1	Recovered/Resolved	Mild
Alanine aminotransferase increased	-1	Recovered/Resolved	Mild
Post-infusion Period			
Venous thrombosis limb	+1	Recovered/Resolved	Moderate
Serum ferritin increased	+1	Recovering/Resolving	Mild
Anemia	+2	Recovered/Resolved	Moderate

Insomnia	+2	Recovered/Resolved	Moderate
Hyperphosphatemia	+3	Recovered/Resolved	Mild
Stomatitis	+3	Recovered/Resolved	Severe
Hyponatremia	+4	Recovered/Resolved	Mild
Constipation	+4	Recovered/Resolved	Mild
C-reactive protein increased	+8	Recovered/Resolved	Moderate
Cytomegalovirus infection	+8	Recovered/Resolved	Mild
Blood alkaline phosphatase increased	+10	Recovered/Resolved	Mild
Neutrophil count decreased	+10	Recovered/Resolved	Severe
Febrile neutropenia	+11	Recovered/Resolved	Severe
Platelet count decreased	+11	Recovered/Resolved	Severe
Drug overdose	+13	Recovered/Resolved	Severe
Hypertension	+15	Recovered/Resolved	Moderate
Clostridium difficile	+22	Recovered/Resolved	Moderate
Vitamin D deficiency	+30	Not Recovered/Not Resolved	Moderate
Proctalgia	+37	Recovered/Resolved	Moderate
Hyperpigmentation	+44	Recovered/Resolved	Moderate
Alanine aminotransferase increased	+51	Recovered/Resolved	Mild
Aspartate aminotransferase increased	+51	Recovered/Resolved	Moderate
Blood lactate dehydrogenase increased	+51	Recovered/Resolved	Mild
COVID-19 infection	+110	Recovered/Resolved	Severe
Sinusitis	+121	Recovered/Resolved	Moderate
Aspartate aminotransferase increased	+135	Recovered/Resolved	Mild
Arthralgia	+169	Recovered/Resolved	Moderate
Alanine aminotransferase increased	+169	Recovered/Resolved	Mild
Arthralgia	+234	Recovered/Resolved	Moderate
Blood alkaline phosphatase increased	+240	Recovered/Resolved	Mild
Epstein-Barr virus	+275	Recovered/Resolved	Mild
Lower respiratory tract infection (viral)	+275	Recovered/Resolved	Moderate
Anemia	+275	Recovered/Resolved	Moderate
Flank pain	+314	Recovered/Resolved	Moderate
Hypertension	+329	Recovered/Resolved	Moderate
Pain in extremity	+329	Recovered/Resolved	Severe

Priapism	+477	Recovered/Resolved	Moderate
Blood lactate dehydrogenase	+484	Not Recovered/Not Resolved	Mild
Alanine aminotransferase increased	+501	Recovered/Resolved	Mild
Leukocytosis	+501	Recovered/Resolved	Mild
Aspartate aminotransferase increased	+501	Recovered/Resolved	Mild
Sickle cell anemia with crisis	+501	Recovered/Resolved	Severe
Hidradenitis	+501	Recovering/Resolving	Moderate
Hypertension	+501	Not Recovered/Not Resolved	Severe
Acute chest syndrome	+502	Recovered/Resolved	Severe
Blood bilirubin increased	+503	Recovered/Resolved	Mild
Hypoalbuminemia	+504	Recovered/Resolved	Mild
Blood alkaline phosphatase increased	+504	Recovered/Resolved	Mild
Anemia	+506	Recovered/Resolved	Severe
Headache	+507	Recovered/Resolved	Severe
Road traffic accident	+547	Recovered/Resolved	Severe
Participant 2			
Pre-conditioning/Screening Period			
Pain	-379	Recovered/Resolved	Mild
Nausea	-377	Recovered/Resolved	Mild
Pain	-293	Recovered/Resolved	Mild
Edema peripheral	-292	Recovered/Resolved	Mild
Hypomagnesaemia	-291	Recovered/Resolved	Mild
Pain	-279	Recovered/Resolved	Mild
COVID-19 infection	-91	Recovered/Resolved	Mild
Hyperphosphatemia	-63	Recovered/Resolved	Moderate
Chronic sinusitis	-41	Recovered/Resolved	Severe
Hyperglycemia	-7	Recovered/Resolved	Mild
Conditioning Period			
Aspartate aminotransferase increase	-5	Recovered/Resolved	Mild
Alanine aminotransferase increase	-5	Recovered/Resolved	Moderate
Insomnia	-5	Recovered/Resolved	Mild
Nausea	-3	Recovered/Resolved	Severe
Post-infusion Period			
Anemia	+3	Recovered/Resolved	Moderate
Stomatitis	+5	Recovered/Resolved	Severe
Neutrophil count decrease	+8	Recovered/Resolved	Severe
Febrile neutropenia	+11	Recovered/Resolved	Severe

Platelet count decrease	+11	Recovered/Resolved	Severe
Constipation	+11	Recovered/Resolved	Mild
Hypokalemia	+18	Recovered/Resolved	Severe
Peripheral sensory neuropathy	+20	Recovered/Resolved	Moderate
Arthralgia	+36	Recovered/Resolved	Moderate
Blood alkaline phosphatase increased	+85	Recovered/Resolved	Mild
Aspartate aminotransferase increase	+85	Recovered/Resolved	Moderate
Alanine aminotransferase increase	+85	Recovered/Resolved	Moderate
Clostridium difficile test positive	+106	Recovered/Resolved	Moderate
Iron overload	+108	Not Recovered/Not Resolved	Severe
Cough	+118	Recovered/Resolved	Moderate
Anemia	+120	Recovered/Resolved	Mild
Cholelithiasis	+176	Unknown	Mild
Dermatitis	+176	Not Recovered/Not Resolved	Mild
Osteonecrosis	+196	Recovered/Resolved	Severe
Anemia	+197	Recovered/Resolved	Mild
Gastroesophageal reflux disease	+226	Recovered/Resolved	Moderate
Sickle cell anemia with crisis	+356	Recovered/Resolved	Severe
Participant 3			
Pre-conditioning/Screening Period			
Sickle cell anemia with crisis	-399	Recovering/Resolving	Mild
Abdominal pain	-385	Recovering/Resolving	Mild
Sickle cell anemia with crisis	-385	Recovering/Resolving	Mild
Sickle cell anemia with crisis	-381	Recovering/Resolving	Mild
Anxiety	-371	Recovered/Resolved	Mild
Vertigo	-371	Recovered/Resolved	Mild
Constipation	-371	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-371	Recovered/Resolved	Mild
Nausea	-371	Recovered/Resolved	Mild
Headaches	-368	Recovered/Resolved	Mild
Osteonecrosis	-368	Recovered/Resolved	Mild
Abdominal pain	-365	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-350	Recovered/Resolved	Mild
Dyspnea	-341	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-341	Recovered/Resolved	Mild

Ocular icterus	-331	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-328	Recovered/Resolved	Mild
Abdominal pain	-325	Recovered/Resolved	Mild
Diarrhea	-325	Recovered/Resolved	Mild
Nausea	-325	Recovered/Resolved	Mild
Upper-airway cough syndrome	-325	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-325	Recovered/Resolved	Mild
Vomiting	-325	Recovered/Resolved	Mild
Vomiting	-321	Recovered/Resolved	Mild
Non-cardiac chest pain	-321	Recovered/Resolved	Mild
Nausea	-321	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-321	Recovered/Resolved	Mild
Fatigue	-320	Recovered/Resolved	Mild
Non-cardiac chest pain	-314	Recovered/Resolved	Mild
Constipation	-301	Recovered/Resolved	Mild
Headaches	-301	Recovered/Resolved	Mild
Sickle cell anemia	-301	Recovered/Resolved	Mild
Fatigue	-301	Recovered/Resolved	Mild
Temperature intolerance	-301	Recovered/Resolved	Mild
Prolonged menses	-301	Recovered/Resolved	Mild
Paranasal sinus discomfort	-301	Recovered/Resolved	Mild
Anxiety	-301	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-294	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-284	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-283	Recovered/Resolved	Mild
Congestion	-264	Recovered/Resolved	Mild
Nausea	-259	Recovered/Resolved	Mild
Headache	-259	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-259	Recovered/Resolved	Mild
Constipation	-245	Recovered/Resolved	Mild
Post-nasal drip	-245	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-245	Recovered/Resolved	Mild
Congestion	-236	Recovered/Resolved	Mild
COVID-19 infection	-236	Recovered/Resolved	Mild
Post-nasal drip	-236	Recovered/Resolved	Mild
Oropharyngeal pain	-236	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-235	Recovered/Resolved	Mild

Sickle cell anemia with crisis	-220	Recovered/Resolved	Mild
Headaches	-219	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-206	Recovered/Resolved	Mild
Vomiting	-189	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-185	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-182	Recovered/Resolved	Mild
Constipation	-182	Recovered/Resolved	Mild
Nausea	-173	Recovered/Resolved	Mild
Vertigo	-171	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-168	Recovered/Resolved	Mild
Vertigo	-168	Recovered/Resolved	Mild
Anxiety	-161	Recovered/Resolved	Mild
Urinary tract infection	-160	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-157	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-153	Recovered/Resolved	Mild
Vertigo	-153	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-143	Recovered/Resolved	Mild
Post-nasal drip	-143	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-140	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-125	Recovered/Resolved	Mild
Abdominal pain	-119	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-117	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-101	Recovered/Resolved	Mild
Sore throat	-101	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-98	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-81	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-71	Recovered/Resolved	Mild
Congestion	-68	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-46	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-35	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-28	Recovered/Resolved	Mild

Sickle cell anemia with crisis	-23	Recovered/Resolved	Mild
Sickle cell anemia with crisis	-14	Recovered/Resolved	Mild
Osteonecrosis	-14	Recovering/Resolving	Moderate
Fatigue	-12	Recovered/Resolved	Mild
Non-cardiac chest pain	-10	Recovered/Resolved	Mild
Conditioning Period			
Headache	-6	Recovered/Resolved	Mild
Immunodeficiency	-6	Recovered/Resolved	Mild
Nausea	-4	Recovered/Resolved	Mild
Post-infusion Period			
Non-cardiac chest pain	+2	Recovered/Resolved	Mild
Constipation	+3	Recovered/Resolved	Mild
Diarrhea	+4	Recovered/Resolved	Mild
Weight increase	+4	Recovered/Resolved	Mild
Skin hyperpigmentation	+6	Recovering/Resolving	Mild
Constipation	+7	Recovered/Resolved	Mild
Abdominal pain	+8	Recovered/Resolved	Mild
Stomatitis	+8	Recovered/Resolved	Mild
Febrile neutropenia	+9	Recovered/Resolved	Mild
Diarrhea	+10	Recovered/Resolved	Mild
Klebsiella infection	+10	Recovered/Resolved	Mild
Bacteremia	+10	Recovered/Resolved	Moderate
Dysuria	+11	Recovered/Resolved	Mild
Vomiting	+11	Recovered/Resolved	Mild
Colitis	+11	Recovered/Resolved	Mild
Diarrhea	+12	Recovered/Resolved	Mild
Vertigo	+15	Recovered/Resolved	Mild
Diarrhea	+15	Recovered/Resolved	Mild
Alopecia	+15	Recovering/Resolving	Mild
Anal abscess	+16	Recovered/Resolved	Mild
Catheter site erythema	+21	Recovered/Resolved	Mild
Catheter site pruritus	+21	Recovered/Resolved	Mild
Catheter site papule	+21	Recovered/Resolved	Mild
Headache	+24	Recovered/Resolved	Mild
Rhinorrhea	+24	Recovered/Resolved	Mild
Rash erythematous	+29	Recovered/Resolved	Mild
Nausea	+29	Recovered/Resolved	Mild
Hematochezia	+43	Recovered/Resolved	Mild
Constipation	+43	Recovered/Resolved	Mild
Pain in jaw	+57	Recovered/Resolved	Mild
Non-cardiac chest pain	+92	Recovered/Resolved	Mild
Non-cardiac chest pain	+99	Recovered/Resolved	Mild
Vomiting	+99	Recovered/Resolved	Mild
Nausea	+99	Recovered/Resolved	Mild
Headache	+99	Recovered/Resolved	Mild
Constipation	+118	Recovered/Resolved	Mild
Vomiting	+141	Recovered/Resolved	Mild

Peripheral swelling	+141	Recovered/Resolved	Mild
Abdominal pain	+141	Recovering/Resolving	Mild
Nausea	+143	Recovering/Resolving	Mild
Vomiting	+152	Recovered/Resolved	Mild
Vomiting	+161	Recovered/Resolved	Mild
Constipation	+162	Recovering/Resolving	Mild
Vomiting	+177	Recovering/Resolving	Mild

Table S8. Summary of laboratory values for the trial participants at various timepoints after OTQ923 infusion

		Participant 1					Participant 2				Participant 3		
Lab Test	**Normal Range**	**Screening**	**Month 3**	**Month 6**	**Month 12**	**Month 18**	**Screening**	**Month 3**	**Month 6**	**Month 12**	**Screening**	**Month 3**	**Month 6**
Chemistry													
Creatinine (µmol/L)	40−110	69	61	69	60	57	54	40	-	54	62	49	46
Aspartate aminotransferase (U/L)	8−40	39	36	30	24	**206**	**51**	**88**	**191**	**45**	**43**	**66**	33
Alanine aminotransferase (U/L)	5−48	23	45	29	23	**176**	33	**112**	**295**	**91**	35	**71**	29
Total bilirubin (µmol/L)	3−21	**26**	**24**	**48**	**43**	**56**	**60**	**32**	**65**	**72**	**67**	**31**	**31**
Alkaline phosphatase (U/L)	40−129	72	112	102	117	119	80	**133**	**168**	**130**	107	102	105
Hematology													
Hemoglobin (g/dL)	14−17	**10.0**	**12.6**	**11.4**	**10.3**	**11.9**	**7.6**	**11.9**	**11.2**	**11.5**	**10**	**10.5**	**10.1**
Hematocrit (%)	39−54	**31**	**36**	**34**	**29**	**32**	**22**	**34**	**33**	**32.1**	**29.9**	**30.7**	**28.7**
Mean corpuscular volume (MCV) (fL)	79−96	81	**99**	**102**	**98**	**100.9**	91	94	**98**	95.3	89.5	92.5	98.3
Mean corpuscular hemoglobin (MCH) (pg)	26−34	26	34	**36**	33	**37.5**	32	32	34	**34.1**	29.9	31.6	**34.6**
Mean corpuscular hemoglobin concentration (MCHC) (g/L)	310−380	320	350	350	340	372	350	340	340	358	334	342	352
Leukocytes (10^9/L)	3.8−10.7	10.49	9.81	3.94	5.86	7.78	**8.74**	8.5	7.96	**16.6**	**15.5**	5.8	8.0
Absolute Neutrophil Count (10^6/L)	2093-6072	**12224**	**6718**	2517	3500	4894	**7266**	3183	3713	**10936**	5800	3940	4540

Absolute Lymphocyte Count (10^6/L)	940-3075	**3242**	2565	1929	2207	2069	2460	2909	2055	2842	1600	1020	2020
Platelets (10^9/L)	140−400	**538**	179	316	371	308	**447**	166	300	261	327	200	226

Abnormal values are **bolded**

References

1. Traxler EA, Yao Y, Wang Y-D, et al. A genome-editing strategy to treat β-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. Nat Med 2016;22(9):987-90. DOI: 10.1038/nm.4170.
2. Yen J, Fiorino M, Liu Y, et al. TRIAMF: A new method for delivery of Cas9 ribonucleoprotein complex to human hematopoietic stem cells. Sci Rep 2018;8(1):16304. DOI: 10.1038/s41598-018-34601-6.
3. Canver MC, Smith EC, Sher F, et al. *BCL11A* enhancer dissection by Cas9-mediated *in situ* saturating mutagenesis. Nature 2015;527(7577):192-7. DOI: 10.1038/nature15521.
4. 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 2014;30(10):1473-5. DOI: 10.1093/bioinformatics/btu048.
5. Cho SW, Kim S, Kim Y, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res 2014;24(1):132-41.DOI: 10.1101/gr.162339.113.
6. Wang X, Wang Y, Wu X, et al. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. Nat Biotechnol 2015;33(2):175-8. DOI: 10.1038/nbt.3127.
7. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009;25(14):1754-60. DOI: 10.1093/bioinformatics/btp324.
8. Pruitt KD, Brown GR, Hiatt SM, et al. RefSeq: an update on mammalian reference sequences. Nucleic Acids Res 2014;42(Database issue):D756-D763. DOI: 10.1093/nar/gkt1114.
9. Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 2017;45(D1):D777-D783. DOI: 10.1093/nar/gkw1121.
10. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013;31(3):213-9. DOI: 10.1038/nbt.2514.
11. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics 2009;25(21):2865-71. DOI: 10.1093/bioinformatics/btp394.
12. Cameron P, Fuller CK, Donohoue PD, et al. Mapping the genomic landscape of CRISPR-Cas9 cleavage. Nat Methods 2017;14(6):600-6. DOI: 10.1038/nmeth.4284.
13. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. Genome Res 2002;12(6):996-1006. DOI: 10.1101/gr.229102.
14. Lai Z, Markovets A, Ahdesmaki M, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res 2016;44(11):e108. DOI: 10.1093/nar/gkw227.
15. Giannoukos G, Ciulla DM, Marco E, et al. UDiTaS, a genome editing detection method for indels and genome rearrangements. BMC Genomics 2018;19(1):212. DOI: 10.1186/s12864-018-4561-9.
16. Untergasser A, Cutcutache I, Koressaar T, et al. Primer3—new capabilities and interfaces. Nucleic Acids Res 2012;40(15):e115. DOI: 10.1093/nar/gks596.
17. Heller D, Vingron M. SVIM: structural variant identification using mapped long reads. Bioinformatics 2019;35(17):2907-15. DOI: 10.1093/bioinformatics/btz041.
18. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. Mol Cell 2010;38(4):576-89. DOI: 10.1016/j.molcel.2010.05.004.
19. Sharma A, Leonard A, West K, et al. Optimizing haematopoietic stem and progenitor cell apheresis collection from plerixafor-mobilized patients with sickle cell disease. Br J Haematol 2022;198(4):740-4. DOI: 10.1111/bjh.18311.

20. Boitano AE, Wang J, Romeo R, et al. Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. Science 2010;329(5997):1345-8. DOI: 10.1126/science.1191536.
21. Koelle SJ, Espinoza DA, Wu C, et al. Quantitative stability of hematopoietic stem and progenitor cell clonal output in rhesus macaques receiving transplants. Blood 2017;129(11):1448-57. DOI: 10.1182/blood-2016-07-728691.
22. Sayers EW, Agarwala R, Bolton EE, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 2019;47(D1):D23-D28. DOI: 10.1093/nar/gky1069.
23. Liu N, Xu S, Yao Q, et al. Transcription factor competition at the γ-globin promoters controls hemoglobin switching. Nat Genet 2021;53(4):511-20. DOI: 10.1038/s41588-021-00798-y.
24. Doerfler PA, Feng R, Li Y, et al. Activation of γ-globin gene expression by GATA1 and NF-Y in hereditary persistence of fetal hemoglobin. Nat Genet 2021;53(8):1177-86. DOI: 10.1038/s41588-021-00904-0.