CRISPR-Cas9 gene edited Hematopoietic Stem and Progenitor cells for the gene therapy of β-hemoglobinopathies

Introduction:

β-Hemoglobinopathies - β-thalassemia and sickle cell disease (SCD) - are highly prevalent inherited globin chain disorders that are autosomal recessive. They account for 3.4% of mortalities in children younger than 5 years[1,2]. β-Thalassemia is caused by more than 300 different mutations in the β-globin gene or its flanking nucleotides; these mutations impair the synthesis of the β-globin chain, affecting the tightly coordinated equilibrium of adult haemoglobin (HbA/$\alpha_2\beta_2$) chains [3]. The excess free α-globin precipitates in erythroblasts and induces apoptosis, resulting in ineffective erythropoiesis[4]. SCD is caused by the E6V (rs334) missense mutation in the β-globin gene. This mutation causes polymerization of deoxygenated sickle hemoglobin (HbS) tetramers, which severely reduces the circulating lifespan of red blood cells (RBCs) and eventually causes vascular damage and progressive multiorgan damage[5].

Morbidity in β-thalassemia and SCD patients is inversely correlated with the levels of fetal hemoglobin (HbF) in adulthood[6,7]. Expression of γ-globin, the fetal β-like globin component of HbF, improves the globin chain equilibrium and thus prevents apoptosis of erythroid cells in β-thalassemia. Similarly, in SCD, γ-globin competes with the sickle β-globin chains (βs) to form HbF tetramers ($\alpha_2\gamma_2$), thereby reducing the production of sickle RBCs. Hence, several studies are focused on identifying and manipulating genetic factors involved in HbF regulation[8–10]. Two recent clinical studies involving short hairpin RNA (shRNA)- mediated erythroid-specific downregulation of BCL11A and gene editing-mediated disruption of its erythroid-specific enhancer have demonstrated reactivated HbF levels sufficient to reach transfusion independence[11,12]. However, up to 50% of hemoglobin remained as HbS in the SCD patients; thus, strategies that reduce or eliminate defective β-globin production are worth further exploration.

Mutations causing hereditary persistence of fetal haemoglobin (HPFH) are documented to produce varying levels of HbF in healthy individuals without any deleterious effects[13]. Importantly, the HPFH mutations are beneficial in alleviating disease severity when co-inherited with β-hemoglobinopathies[14]. Among the genetic variants that induce HbF expression, deletional HPFH mutations produce higher levels of HbF and are highly prevalent[15]. β-Globin production is ablated in HPFH deletions, distinguishing them from other HbF reactivating mutations. HPFH deletions range in size from 12.9 to 84.9 kb, encompassing *HBG1*, *HBBP1*, *HBD,* and *HBB* genes in the β-globin cluster, and result in pancellular HbF production[16]. The introduction of HPFH deletions in adult hematopoietic stem and progenitor cells (HSPCs) results in activation of γ-globin with subsequent amelioration of the sickle phenotype[17,18]. However, much is still unknown, such as the minimal genomic deletion required for therapeutically relevant γ-globin activation, genome integrity, engraftment, and repopulation potential of HSPCs harboring such genomic deletions and their efficacy in

reversing the disease phenotype. In a very recent study, Topfer *et al*. showed that the deletion of the proximal promoter of HBB, excised in HPFH and δβ- thalassemia deletions is sufficient for γ-globin activation [19].

To investigate the translational potential of HSPCs with deletional HPFH mutations, we used CRISPR-Cas9 to screen multiple HPFH deletions and identified an 11 kb core-regulatory region from putative repressor region (PRR) to β-globin exon-1 (βE1) - (PRR-βE1). Gene editing of PRR-βE1 repressed the β-globin and activated γ-globin to levels greater than known candidates targeting the BCL11A enhancer and HBG promoter region, reversing the SCD and β-thalassemia phenotypes. We also demonstrated long-term hematopoiesis of the edited HSPCs and achieved efficient editing without cytokine pre-stimulation and genotoxicity.

Objectives

1. CRISPR-Cas9 screening of β-globin locus to identify novel regulatory regions of fetal haemoglobin re-activation.
2. To Validate the HbF re-activation targets in the HSPCs
3. Transplantation of gene edited HSPCs and characterizing the engraftment and differentiation potential of gene modified HSPCs in NBSGW mice.
4. Characterizing the disease reversal effects by gene editing HSPCs from the patients of SCD and Thalassemia.

Materials and Methods

Purification and culture of CD34^{+ve} HSPCs

The unused G-CSF mobilised peripheral blood collected for allogenic stem cell transplantation and Plerixafor mobilised peripheral blood from SCD, or β-thalassemia patients were collected from transplantation unit of Christian Medical College, Vellore with prior IRB (Institutional Review Board) approval. The CD34^{+ve} HSPCs were purified as described in our previous studies[20–22].

Electroporation of RNP complex in HUDEP-2, CD34^{+ve} cells and β-hemoglobinopathies patient HSPCs

SgRNAs were designed using CRISPR Design Tool (Synthego) and CRISPR-Cas9 guide RNA design checker (IDT), and the efficient gRNAs with least off-target sites were selected. For nucleofection of HUDEP-2 cell lines, 100pmols of Cas9 (Takara) was incubated at room temperature for 10 minutes with 200pmols of sgRNA (Synthego). For dual sgRNA gene editing 100 pmols of Cas9 RNP with cut site A sgRNA and 100 pmols of Cas9 RNP with cut site B sgRNA were nucleofected (Lonza 4D nucleofector) with CA137 pulse code. For electroporation of CD34^{+ve} HSPCs, 50pmol of Cas9 RNP with sgRNA against PRR and 50pmol of Cas9 RNP with sgRNA against βE1 were used. 2x10^5 cells were electroporated using P3 primary cell solution

and supplement and were electroporated using Lonza 4D nucleofector with DZ100 pulse code. For nucleofection of SCD and β-thalassemia patient HSPCs, 100pmols of Cas9 (Takara) was incubated at room temperature for 10 minutes with 200pmols of sgRNA (Synthego). For dual sgRNA gene editing 100 pmols of Cas9 RNP with cut site A sgRNA and 100 pmols of Cas9 RNP with cut site B sgRNA were nucleofected (Lonza 4D nucleofector) with DZ100 pulse code.

HUDEP-2 expansion and differentiation:

The HUDEP-2 cells were cultured in StemSpan SFEM-II media containing SCF (50ng/ml), EPO (3U/ml), Dexamethasone (1μM), Doxycycline (1μg/ml), and Glutamine (1x) at $2x10^5$ cells/ml confluency with media change on alternative days. For erythroid differentiation, previously reported protocol with minor modifications was used [23]. The cells were seeded at a density of $2x10^5$ cells/ml in IMDM GlutaMAX Supplement media containing 3% AB serum, 2% FBS, Insulin (10μg/ml), Heparin (3U/ml), EPO (3U/ml), Holotransferrin (200 μg/ml), SCF (100ng/ml), IL3 (10ng/ml) and Doxycycline (1μg/ml). On day 2, cells were seeded at a cell density of $3.5x10^5$ cells/ml. On day 4, the cells were seeded at a cell density of $5x10^5$ cells/ml in the media containing the above-mentioned cytokine except doxycycline. On day 6, the cells were seeded at a cell density of $1x10^6$ cells/ml in the media with all the components of day 4 media along with increased concentration of Holotransferrin (500 μg/ml). The cells were analysed for HbF^{+ve} cells, differentiation profile and globin chains using HPLC

Erythroid differentiation of $CD34^{+ve}$ HSPCs

The protocol for erythroid differentiation from $CD34^{+ve}$ HSPCs was adopted from the literature with minor modifications [24]. The three-phase erythroid differentiation protocol involves culturing the $CD34^{+ve}$ cells at a seeding density of $5x10^4$ cells/ml in phase I from day 0 – day 8 with a media change on day 4. The phase I media is prepared using IMDM GlutaMAX Supplement media containing 5% AB serum, Insulin (20μg/ml), Heparin (2U/ml), EPO (3U/ml), Holotransferrin (330 μg/ml), SCF (100ng/ml), IL3 (50ng/ml) and Hydrocortisone (1μg/ml). In phase II, the cells were seeded at a density of $2x10^5$ cells from day 8 – day 12 in media containing all the components of phase I except Hydrocortisone and IL3. In phase III, the cells were seeded at a density of $5x10^5$ cells from day 12 – day 20 in media containing all the components of phase II except SCF with a media change on day 16. On day 20, the cells were collected for F^{+ve} cells analysis, differentiation marker analysis and for haemoglobin and globin chain HPLC.

For differentiation of healthy donor and thalassemia patient HSPCs, the cells were cultured at 37°C, 5% CO_2 and in normoxia conditions (21% O_2). For the erythroid differentiation of SCD patient HSPCs, the above protocol was followed except the oxygen levels where we have cultured the SCD patient cells under hypoxic conditions (5% O_2) till day 20 to promote robust sickling of the erythroid differentiated cells [25].

Flow cytometry

For HbF^{+ve} cell analysis, $1x10^5$ erythroid differentiated cells were briefly washed with PBS and fixed with 0.05% glutaraldehyde for 10 minutes and permeabilized with 0.1% Triton-X-100 for 5 minutes. The cells were stained with anti-HbF APC antibody (dilution 1:50) and was acquired and analysed using Cytoflex LX Flow Cytometer (Beckmann Coulter) or AriaIII flow cytometer (BD Biosciences) and analysed using FlowJo (BD Biosciences). For erythroid differentiation analysis, $1x10^5$ cells from the terminal day of erythroid differentiation were stained for erythroid differentiation markers anti-CD71-FITC (dilution 1:33), anti-CD235a PE-Cy7 (dilution 1:50) and Hoechst 33342 (dilution 1:1000). After 20 minutes of incubation in dark, the cells were washed with PBS followed by the analysis using Cytoflex LX Flow Cytometer (Beckmann Coulter) or AriaIII flow cytometer (BD Biosciences).

In vivo engraftment analysis

All the *in vivo* experiments in NBSGW mice models were conducted with approval form IAEC (Institutional Animal Ethics Committee) of Christian Medical College, Vellore, India. The NBSGW were bred in-house and were conditioned with busulfan at a concentration of 12.5mg/kg of body weight, 48hrs prior to the infusion.

$CD34^+$ HSPCs were pre-stimulated for 36hrs – 40hrs with culture media containing appropriate cytokines and RUS (Resveratrol UM729 SR1) cocktail [20,21]. $5x10^5$ -$6x10^5$ cells of control edited and PRR-βE1 edited were infused into NBSGW mice, immediately post electroporation. 16-18 weeks post infusion, the mice were euthanised and peripheral blood, bone marrow and spleen were collected. After RBC lysis buffer incubation, the harvested cells were incubated with mouse Fc block and stained with hCD45 and mCD45 antibody. The % of engraftment is calculated using the formula (% hCD45/% hCD45+% mCD45) x 100. In addition, the multilineage markers including CD19, CD3, CD33, CD13, and CD235a in bone marrow hCD45+ cells were also analysed. For *ex vivo* erythroid differentiation, $3x10^6$ cells were harvested from mouse bone marrow, seeded in erythroid differentiation media and at the end of phase III of differentiation, the % of F+ve cells, the differentiation profile and globin chains were analysed. For *in-vivo* HbF^{+ve} cell analysis in NBSGW, $1x10^6$ bone marrow cells were stained with 10µL of CD235a antibody and sorted based on the presence of immunophenotypic marker CD235a and was followed by F^{+ve} cell analysis. For secondary infusion, $4x10^6$ cells from the pooled fraction harvested from primary recipient bone marrow were infused to secondary recipients 48 hrs post busulfan conditioning. After 14 weeks, the mice were euthanised and the harvested cells were stained with hCD45 and mCD45 antibody for calculating the % of engraftment.

In vitro sickling assay

Sickling assay protocol was adopted from the literature with minor modifications[25,26]. The gene edited SCD patient HSPCs were differentiated till day 20 of erythroid differentiation under hypoxia (5% O_2). On day 20 of erythroid differentiation, enucleated cells (reticulocytes) marked by $Hoechst^{-ve}$ were flow sorted. The flow sorted

cells were resuspended with phase III erythroid differentiation medium and seeded in 24 well plates. Freshly prepared 1.5% sodium metabisulfite in 1x PBS were mixed with phase III media containing the reticulocytes in 1:1 ratio and incubated at 37°C for 1 hour under hypoxia (5% O_2). After the incubation, cover the sides of the 24 well plate with parafilm. Live cells images were acquired using EVOS FL Auto microscope. The percentage of sickle cells were calculated as no. of sickle cells divided by the total number of cells.

Quantitative Real-Time PCR Analysis

$3x10^6$ cells from the day 8 of $CD34^+$ HSPC (Hematopoietic Stem and Progenitor Cells) and day 6 of HUDEP-2 erythroid differentiation were used for total RNA using RNeasy Mini Kit (Qiagen). For reverse transcription using Primescript RT reagent kit (Takara Bio Inc.), 1 µg of extracted RNA was used according to manufacturer's instruction. For quantitative PCR, the SYBR Premix Ex Taq II (Takara Bio) was used for quantifying the specific transcripts and analyzed with QuantStudio 6 Flex (Life Technologies).

Colony formation assay

48hrs post electroporation, $5x10^2$ HSPCs were seeded in 1.5ml of Methocult Optimum (STEMCELL Technologies) and after 14 days, the colonies were scored based on the morphology as CFU-GM, CFU-GEMM, BFU-E, and CFU-E.

Digital droplet PCR (ddPCR)

The frequency of large genomic deletions were quantified using EvaGreen based ddPCR assay. The reaction mixture includes 20ng of genomic DNA, 1x QX200 ddPCR EvaGreen supermix and 100nM primers for 20µL reaction. For absolute measure of deletions, we designed primers that amplify the sequences flanking the cut sites after targeted deletion. Control primers amplifying embryonic globin gene were used as loading control (EG). The percentage of deletion was calculated using the formula.

$$\left(\frac{EAB}{EE}\right) * 100$$

Where,

EAB – DNA copies/µl from primers flanking the cut sites of edited samples.

EG - DNA copies/µL from primers amplifying embryonic globin gene of edited samples.

The second approach involves the quantification of the individual cut sites of the deletion, normalised with the read outs from the unedited control samples.

$$\left(100 - \left(\frac{EA * \left(\frac{UE}{UA}\right)}{EG}\right) + \left(\frac{EB * \left(\frac{UE}{UB}\right)}{EG}\right)\right) * 100$$

Where,

EA – DNA copies/µL from primers flanking the cut site A of edited samples.

EB – DNA copies/µL from primers flanking the cut site B of edited samples.

EG - DNA copies/µL from primers amplifying embryonic globin gene of edited samples.

UA – DNA copies/µL from primers flanking the cut site A of unedited samples.

UB – DNA copies/µL from primers flanking the cut site B of unedited samples.

UE - DNA copies/µL from primers amplifying embryonic globin gene of unedited samples.

Haemoglobin and globin chain analysis using High performance liquid chromatography.

The gene edited HUDEP-2 cell lines and $CD34^{+ve}$ HSPCs were collected on day 8 and day 20 of erythroid differentiation, respectively. The cells were sonicated for 60 seconds with 50% AMP in ice using ultrasonicator (Vibra-Cell) and centrifuged at 14000 rpm for 5 minutes at 4ºC. For haemoglobin HPLC, the protein lysate was analysed for haemoglobin tetramer using G8 HPLC Analyzer (Tosoh). The globin chain analysis was performed using HPLC equipment with UV detector (Shimadzu) and the analysis was performed using LC SolutionsTM software (Shimadzu) using previously reported method[27]. Aeris Widepore 3.6 lm XB-C18 25cm 4.6mm column behind a Security Guard UHPLC Widepore C18 4.6mm guard column (PhenomenexTM) is used for chromatographic separation of the analytes. HPLC conditions include 0.1% trifluoroacetic acid (TFA), pH 3.0 (solvent A), mobile phase - 0.1% TFA in acetonitrile (solvent B) with gradient elution at a flow rate of 1.0 ml/min and column temperature maintained at 70º C with runtime around 8 minutes and UV detection range of 190nm was set for globin chain detection.

Western Blot Analysis

Approximately $6x10^6$ erythroblasts were collected on day 8 of erythroid differentiation. The lysates were prepared sonicating the cell pellets resuspended in RIPA buffer supplemented 1x protease and phosphatase inhibitor cocktail. 20µg of protein lysates resuspended in 1x Lamelli buffer were loaded to the wells of SDS PAGE. The western blots were performed using the primary antibodies, anti-hemoglobin α (1:1000 dilution), anti-hemoglobin β (1:1000 dilution), anti-hemoglobin γ (1:1000 dilution) and anti-actin (1:5000 dilution) along with anti-mouse IgG HRP secondary antibodies. Densitometric analysis of the globin bands were performed by normalising with β-actin.

Transcriptome analysis

Total RNA was extracted using Qiagen RNA isolation kit, quantified using Qubit RNA Assay HS, purity checked using QIAxpert and RNA integrity was assessed on

TapeStation using RNA HS ScreenTapes (Agilent, Cat# 5067-5579). NEB Ultra II Directional RNA-Seq Library Prep kit protocol was used to prepare libraries for total RNA sequencing. Prepared libraries were quantified using Qubit High Sensitivity Assay (Invitrogen, Cat# Q32852). A cluster flow cell is loaded on Illumina HiSeq 4000 instrument to generate 60M, 100bp paired end reads. Read Counts from mapped reads were obtained using Feature Counts. Differential expression analysis was performed using DESEQ2. Gene set enrichment analysis was by GSEA (Gene Set Enrichment Analysis) software from the Broad Institute. A ranked list of differentially expressed genes from RNAseq data was loaded into GSEA and tested against a list of genes documented from published reports. Heat map for differentially regulated genes were generated using Morpheus (broad institute).

CAST-Seq analysis

CAST-Seq library preparation was performed as previously described[28]. Two technical replicates derived from two independent editing experiments were prepared and analysed against a corresponding untreated control sample. We considered as relevant findings all sites identified in at least two out of four replicates with a read-to-CAST-Seq hit ratio of >10 to eliminate unspecific reads. CAST-Seq libraries were sequenced by NGS (Next Generation Sequencing) service provider Genewiz (part of Azenta Life Sciences). At Genewiz, sequencing was performed on an Illumina NovaSeq device collecting 2x150bp paired-end reads. The bioinformatics analysis was adapted to allow for concomitant input of more than one target site/guide RNA sequence.

Micronucleus assay

Micronucleus assay was performed as previously described[29] following the published guidelines [30]. Briefly, the gene edited HSPCs were incubated with 10 µM cytochalasin B for 23 hours and fixed with methanol and stained with giemsa stain. The images were captured using 40x magnification using Olympus upright microscope BX43.

KaryoStat assay

Gene edited HSPCs were collected for genomic DNA isolation using PureLink Genomic DNA Mini Kit (catalog # K182000) and quantified using Qubit dsDNA assay. After digestion of 250ng of genomic DNA using Nsp I restriction enzyme, the DNA were ligated with adapter and amplified. The DNA were fragmented followed by labelling with biotin and the labelled DNA was hybridised onto GeneChip arrays. GeneChip Fluidics Station 450 were used for washing and staining of Chips simultaneously scanned using GeneChip Scanner 3000 7 G. Data were analyzed using ChAS 3.2. The raw data were processed using Genotyping Console v4.0 and Chromosome Analysis Suite 3.2 with NetAffx na33.1 (UCSC GRCh37/hg19), and the output data were interpreted with the UCSC Genome Browser (https://genome.ucsc.edu/; GRCh37/hg19 assembly).

Results

Genomic deletion encompassing PRR to βE1 is sufficient to reproduce deletional HPFH phenotype

To identify an HPFH deletion suitable for therapeutic gene editing, we introduced deletional HPFH mutations of <30 kb in size, mirroring the Algerian[31] (24 kb), French[31] (20 kb), South-East (SE) Asian[32] (27 kb), and Sicilian[33] (12.9 kb) genotypes, by CRISPR-Cas9 dual guide RNA (gRNA) gene editing in the HUDEP-2 cell line (Fig. 1A). The 7.2-kb Corfu deletion, which is now considered as δβ-thalassemia and requires homozygous deletion to activate therapeutic HbF levels [34,35], was excluded from our screening. The efficiency of gene editing was assessed using droplet digital polymerase chain reaction (ddPCR) and Sanger sequencing in conjunction with Inference of CRISPR Edits (ICE) analysis.

All candidates exhibited >60% editing efficiency (Fig. 1B). Erythroid differentiation of gene-edited HUDEP-2 cells showed an increased percentage of HbF^{+ve} cells (Fig. 1C) and activation of γ-globin chains (Fig. 1D) in all HPFH deletions. Sicilian HPFH produced a marginally higher level of γ-globin chains than the other targets, and is the central region among the HPFH deletions.

To decipher the HbF regulatory region in the Sicilian HPFH deletion, we excised different regions spanning the deletion, such as putative repressor region (PRR) to the upstream region of β-globin promoter (PRR-HBB(-250)), promoter to exon-1 (HBB(-250)-βE1), PRR to exon-1 (PRR-βE1), PRR to intron-1 (PRR-βI1), and PRR to exon-1 CD27 (PRR-βE2) of the β-globin gene (Fig. 1A, 1E and 1F). Among these candidates, HBB(-250)-βE1 and PRR-βE1 showed a 1.8-fold higher percentage of HbF^{+ve} cells than the Sicilian HPFH (Fig. 1G, H). Both candidates showed a substantial increase in the γ-globin activation (Fig.1I, 1J). Variant High Performance liquid Chromatography (HPLC) analysis confirmed the functional HbF tetramer in both these samples, and they had a greater than four-fold higher proportion of HbF tetramers than the Sicilian HPFH (Fig. 1K).

The PRR-βE1 region is excised in all the deletional HPFH mutations. The sequence spanning PRR is completely or partially intact in δβ-thalassemia and β-thalassemia deletions, suggesting it as a region that distinguishes HPFH and thalassemia phenotypes. Whereas, HBB(-250)-βE1 deletion resembles the British black and Croatian β-thalassemia genotypes[36,37] and was also reported recently as a target for HbF reactivation [19]. Therefore PRR-βE1 was considered for further studies.

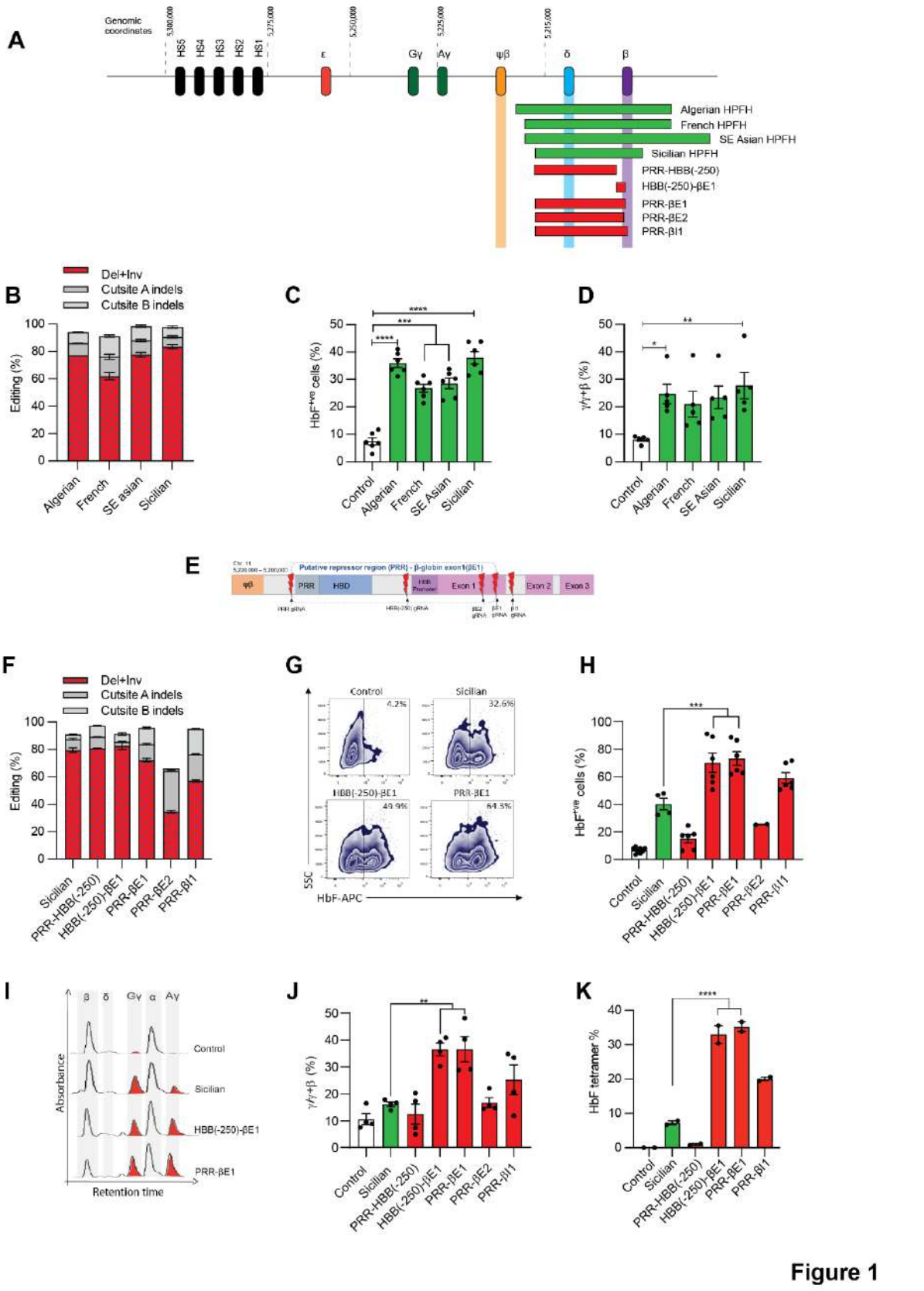

Figure 1. Genomic deletion encompassing PRR to βE1 is sufficient to reproduce deletional HPFH phenotype

(A) Diagrammatic representation of β-globin cluster and the break points of naturally occurring HPFH deletions (green). These deletions are introduced in the experiments shown in figure 1B-D. Break points of deletions (red) introduced in the experiment shown in figure 1E-J. All these deletions were generated in HUDEP-2 cell lines by

CRISPR/Cas9 dual gRNA approach. (B) Percentage of gene editing in HUDEP-2 cell lines, gene edited for HPFH deletions. Type of HPFH deletions were indicated at the X-axis. Indels (cut site A and cut site B) measured by sanger sequencing and ICE analysis. Deletion + Inversion (Del+Inv) (red checker box) quantified by ddPCR. n = 2. (C) Percentage of HbF^{+ve} cells upon introducing HPFH deletions indicated on the x-axis. The edited cells were differentiated into erythroblasts and analysed for HbF by flow cytometry. n = 6. (D) γ-Globin chain synthesis as measured by γ/γ+β ratio in the HUDEP-2 erythroblasts as measured by HPLC chain analysis. n = 5. (E) Magnified image of β-globin locus showing the binding sites of the key gRNA employed in this study to create various deletions mentioned in Fig. 1F-K. (F) Percentage of gene editing in HUDEP-2 cell lines gene edited for various deletions as indicated in the X-axis. Indels (cut site A and cut site B) measured by sanger sequencing and ICE analysis. Deletion + Inversion (Del+Inv) (red checker box) quantified by ddPCR. n=2. (G) Representative flow cytometry plot of HbF^{+ve} cells . HUDEP-2 cell lines gene edited for Sicilian HPFH deletion and deletion of its encompassing region in the β-globin cluster, were differentiated into erythroblasts and analysed for HbF^{+ve} cells. Inset shows percentage of HbF^{+ve} cells. (H) Percentage of HbF^{+ve} cells upon introducing deletions in the region encompassing Sicilian HPFH. n = 4. (I) Representative globin chain HPLC chromatograms. (J) γ-Globin chain synthesis as measured by γ/γ+β ratio in the HUDEP-2 cell lines as measured by HPLC chain analysis. n = 4. (K) Percentage of HbF tetramer in erythroid differentiated HUDEP-2 cells gene edited for introducing deletions in the region encompassing Sicilian HPFH as measured by Variant HPLC analysis. n = 2. Error bars represent mean ± SEM, $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, $^{****}p \leq 0.0001$ (one-way ANOVA followed by Dunnett's multiple comparisons test).

Robust γ-globin induction and β-globin silencing in the erythroblasts differentiated from PRR-βE1 -edited HSPCs

To investigate the effect of PRR-βE1 editing in therapeutically relevant cells, granulocyte colony-stimulating factor (G-CSF)- mobilized HSPCs from five healthy donors were electroporated with Cas9 ribonucleoproteins (RNPs) targeting cut site A - PRR and cut site B - βE1 sites individually and in combination. The total gene-editing efficiency in PRR-βE1 was 92±4%, among which PRR-βE1 deletion and inversion (del+inv) comprised 72.1±2% (Fig. 2A). Upon differentiation of HSPCs into erythroblasts using a three- phase *in vitro* erythroid differentiation protocol, a significant increase in the percentage of HbF^{+ve} cells was observed in PRR-βE1 (74.2±3%) and βE1 -edited cells (68.0 ± 3%) relative to the *AAVS1* control (24.7±3%) (Fig. 2B). Variant HPLC analysis showed up to 13-fold higher levels of HbF tetramers upon βE1 and PRR-βE1 editing (Fig. 2C). Consistent with all aforementioned analyses, Western blot confirmed the increased γ-globin and decreased β-globin expression (Fig. 2D), quantitatively confirming the absolute levels of γ-globin produced on PRR-βE1 gene editing. The erythroid maturation analysis using CD235a and Hoechst, showed that PRR-βE1 edited cells had comparable levels of reticulocytes with the control, whereas the percentage of reticulocytes was significantly lower in βE1-edited cells (Fig. 2E).

Similarly, in PRR-βE1 editing, erythroid colony forming potential as assessed by the ratio of erythroid (Burst-forming unit (BFU-E) + colony forming unit (CFU-E)) to granulocyte-monocyte (GM) generation, remained equivalent to *AAVS1* but significantly decreased in βE1 editing (Fig. 2F). All these analyses indicate normal erythropoiesis in PRR-βE1 editing but not in βE1 editing.

We next compared the HbF induction by PRR-βE1 and Sicilian HPFH with well-characterized targets: the BCL11A erythroid specific enhancer and BCL11A -binding site in the *HBG* promoter that have advanced into clinical studies[12,38]. Gene editing efficiencies and

the ratio of erythroid to GM colonies were comparable for all these targets (Fig. 2G). All four targets produced HbF^{+ve} cells, with PRR-βE1 cells producing the highest proportion of HbF^{+ve} cells (Fig. 2H). PRR-βE1-edited cells produced HbF tetramers that were two-fold higher than the other targets and HbA tetramers were three-fold lower (Fig. 2I, J, and K). Western blot analysis further confirmed that PRR-βE1 editing increased γ-globin chains relative to other targets tested (Fig. 2L).

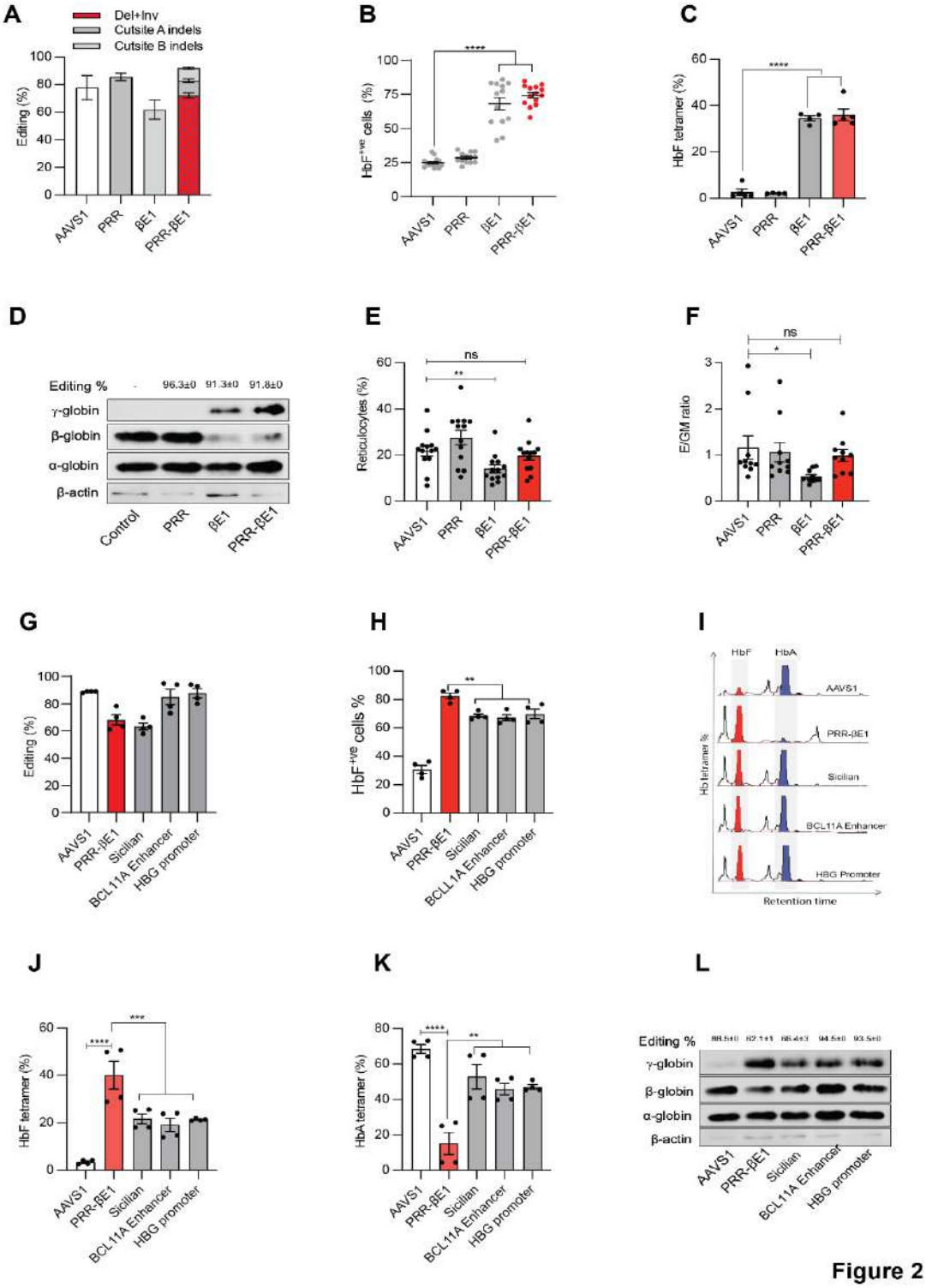

Figure 2

Figure 2. Robust γ-globin induction and β-globin silencing in the erythroblasts differentiated from PRR-βE1 -edited HSPCs

(A) Percentage of gene editing in PRR, βE1, and PRR-βE1 gene edited healthy donor HSPCs. Indels measured by sanger sequencing and ICE analysis. Deletion + Inversion (Del+Inv) (red checker box) in PRR-βE1 quantified by ddPCR. The PRR-βE1 edited cells had deletion, indels at PRR region and βE1. Donor = 5, n = 11. (B) FACS analysis of percentage of HbF^{+ve} cells in erythroblasts generated from gene edited HSPCs. Gene editing targets are indicated at the bottom. Control refers to unedited cells. Each dot indicates individual experiment. Donor = 5, n = 11. (C) Percentage of fetal hemoglobin (HbF) tetramer as measured by variant HPLC for HSPCs gene edited for PRR, βE1 and PRR-βE1 and differentiated into erythroblasts. Donor = 3, n = 4. (D) Representative western blot image showing the band intensity of globin chains for erythroblasts derived from control, PRR, βE1, and PRR-βE1 gene edited HSPCs. The editing in PRR and βE1 indicates the percentage of Indels by ICE analysis and for PRR- βE1 edited, the percentage of editing includes the deletion + inversion quantified by ddPCR, cut site A and cut site B indels by ICE analysis. Donor = 1, n = 3. (E) Percentage of reticulocytes generated on erythroid differentiation of HSPCs gene edited for PRR, βE1, and PRR-βE1. Flowcytometric analysis of reticulocytes percentage were quantified on day 20 of three phase erythroid differentiation. Donor = 5, n = 11. (F) Ratio of Erythroid (E) to granulocyte-monocyte (GM) CFU (Colony Forming Units) colonies. HSPCs were gene edited for *AAVS1*, PRR, βE1 and PRR-βE1 and plated in methocult medium. Both BFU-E and CFU-E colonies were considered as erythroid (E) colonies. Donor = 2, n = 10. (G) Percentage of gene manipulation as measured by ddPCR for quantifying deletions in PRR-βE1 and Sicilian HPFH. Indel analysis of *AAVS1*, BCL11A enhancer, and HBG promoter by ICE analysis. Donor =2, n =4. (H) FACS analysis of percentage of HbF^{+ve} cells in erythroblasts generated from for PRR-βE1, Sicilian HPFH, BCL11A enhancer, and HBG promoter, Donor = 2, n = 4. (I) Representative hemoglobin variant HPLC chromatograms showing HbF and HbA tetramers in gene edited cells. (J) Percentage of HbF tetramers. HSPCs were gene edited for PRR-βE1, Sicilian HPFH, BCL11A enhancer, and HBG promoter, differentiated into erythroblasts and analysed by variant HPLC. Donor = 2, n = 4. (K) Percentage of HbA tetramers. HSPCs were gene edited for PRR-βE1, Sicilian HPFH, BCL11A enhancer, and HBG promoter, differentiated into erythroblasts and analysed by variant HPLC. Donor = 2, n = 4. (L) Representative western blot image showing the band intensity of globin chains for erythroblasts derived from PRR-βE1, Sicilian HPFH, BCL11A enhancer, and HBG promoter gene edited HSPCs. The editing in *AAVS1*, BCL11A enhancer, and HBG promoter indicates the Indels quantified by ICE analysis. For PRR-βE1 and Sicilian HPFH, editing indicates the percentage of deletion and inversion quantified by ddPCR excluding the cut site indels. Donor = 1, n = 2. Error bars represent mean ± SEM, $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$, $****p \leq 0.0001$ (one-way ANOVA followed by Dunnett's multiple comparisons test).

PRR-βE1 gene-edited HSPCs repopulate for long-term and generate HbF^{+ve} cells *in vivo*

To characterize the *in vivo* reconstitution capability of PRR-βE1 gene-edited cells, we used NBSGW mice, which support robust human cell engraftment and erythropoiesis[39]. Gene editing was performed on HSPCs from two healthy donors using PRR-βE1 and CRISPR RNA (crRNA) less RNP (control). The crRNA-free RNPs does not induce DNA double-strand breaks and therefore serve as an ideal control for assessing engraftment defects associated with Cas9 gene editing. The edited cells were transplanted into NBSGW mice as two cohorts, each infused with different donor cells and analyzed 16 weeks post-transplantation.

The engraftment of PRR-βE1-edited cells in the bone marrow, peripheral blood, and spleen of the mice was comparable with that of the control group (Fig. 3A and S5A). The multilineage repopulation potential of engrafted cells was also comparable among the groups (Fig. 3B). The percentage of $CD235a^{+ve}$ erythroblasts were also similar, confirming the intact erythropoiesis *in vivo* from PRR-βE1-edited HSPCs. Importantly, genotyping of the long-term

repopulating cells in all the mice revealed the retention of PRR-βE1 editing, and the percentage of editing was comparable with that of infused cells in 12 of the 13 animals tested (Fig. 3C). Nextly, human CD235a^{+ve} erythroblasts were sorted from mouse bone marrow and was found to be increased in the proportion of HbF^{+ve} cells *in vivo* following PRR-βE1 editing (Fig. 3D). Furthermore, *in vitro* erythroid differentiation of cells retrieved from mouse bone marrow showed a significant increase in HbF^{+ve} cells (Fig. 3E), γ/(γ+β) ratio (Fig. 3F).

To assess the serial repopulation potential of HSCs (Hematopoietic Stem Cells) harbouring PRR-βE1 editing, we infused bone marrow cells of primary recipients (cohort 2) to secondary recipients and analysed the bone marrow 14 weeks post infusion. The analysis showed similar frequencies of engraftment of PRR-βE1-edited cells and control edited in the secondary recipients. (Fig. 3G).

PRR-βE1 gene editing without *ex vivo* culturing of HSPCs

Unlike lentiviral transduction or HDR-based gene editing, cytokine pre-stimulation of HSPCs may not be necessary for NHEJ-mediated gene editing. To examine whether PRR-βE1 gene editing is feasible without culture and cytokine prestimulation, HSPCs were electroporated immediately following purification and infused into NBSGW mice. The strategy was compared with the standard protocol, which consists of 48 hours of cytokine stimulation prior to electroporation. On analysis after 16 weeks post transplantation, both groups exhibited comparable levels of bone marrow engraftment and PRR-βE1 gene editing (Fig. 3H). Functionally, PRR-βE1 -edited HSPCs from both uncultured and cultured HSPCs produced significantly more HbF^{+} cells *in vivo* (Fig. 3I), corroborating with earlier findings. Furthermore, *in vitro* erythroid differentiation of cells from mouse bone marrow showed a significant increase in γ/(γ+β) ratio (Fig. 3J) in PRR-βE1- edited cells compared with the *AAVS1* control.

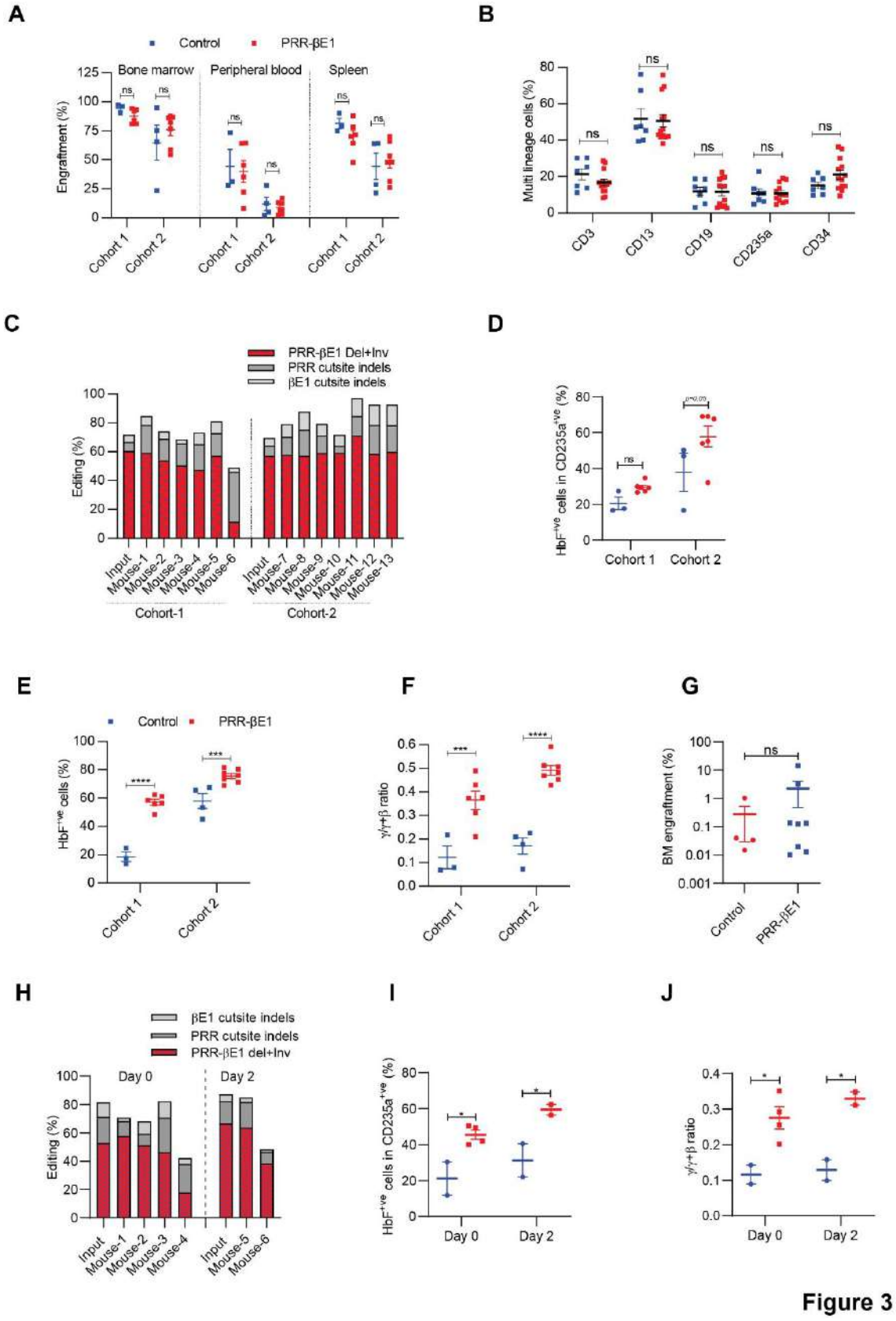

Figure 3

Figure 3. PRR-βE1 gene-edited HSPCs repopulate for long-term and generate HbF^{+ve} cells *in vivo*

Control and PRR-βE1 gene edited healthy donor HSPCs were transplanted into NBSGW mice and analysed 16 weeks post transplantation (Fig. 3A-G). Each dot indicates a single mouse. Donor = 2. Each cohort indicates an independent experiment infused with HSPCs gene edited for PRR-βE1. Error bars represent mean ± SEM, ns; non-significant. $^{***}p \leq 0.001$, $^{****}p \leq 0.0001$ (two-way ANOVA followed by Dunnett's test). (A) Percentage of engraftment in the bone marrow, peripheral blood, and spleen calculated flow cytometrically using hCD45 and mCD45.1 markers. (B) Percentage of HSPC and lineage markers in BM (Bone Marrow) – CD3 (T-cells), CD13 (monocyte), CD19 (B-cells), CD235a (Erythroid), and CD34 (HSPCs) in engrafted cells. $CD235a^{+}$ cells were analyzed form $CD45^{-}$ cells. (C) Percentage of PRR-βE1 deletion+inversion (Del+Inv), PRR cut site indels, and βE1 cut site indels in PRR-βE1 gene edited HSPCs in infused fraction and in engrafted cells.
(D) Percentage of HbF^{+ve} cells in $hCD235a^{+ve}$ cells obtained from mouse BM. (E) Percentage of HbF^{+ve} cells generated by erythroid differentiation of engrafted cells in the BM. (F) Ratio of γ/γ+β chains. Mouse BM was collected, *in vitro* differentiated into erythroblasts and analyzed by chain HPLC. (G) Percentage of engraftment in bone marrow (BM) of secondary recipients analysed 14 weeks post transplantation.

AAVS1 and PRR-βE1 gene edited healthy donor HSPCs were gene edited immediately after CD34 purification (day 0) and 48hrs post CD34 purification (day 2) and transplanted into NBSGW mice and analysed 16 weeks post transplantation (Fig. 3G-J). Each dot indicates a single mouse. Donor = 1. Error bars represent mean ± SEM, ns; non-significant. $^{*}p \leq 0.05$ (two-way ANOVA followed by Dunnett's test). (H) Percentage of PRR-βE1 deletion+inversion (Del+Inv), PRR cut site indels, and βE1 cut site indels in PRR-βE1 gene edited HSPCs in Day 0 and Day 2 edited input fraction and in engrafted cells. (I) Percentage of HbF^{+ve} cells in $hCD235a^{+ve}$ cells obtained from mouse BM. (J) Ratio of γ/γ+β chains. Mouse BM was collected, *in vitro* differentiated into erythroblasts and analyzed by chain HPLC.

PRR-βE1 gene-edited patient HSPCs reverses SCD phenotype

To test the potential of our PRR-βE1 gene-editing strategy in the reversal of the sickle phenotype, the plerixafor-mobilized HSPCs from two SCD patients of compound heterozygous genotype HbS/CD41/CD42(-TCTT) and HbS/IVS1-5 were gene edited with Cas9-RNPs targeting *AAVS1*, PRR, βE1 and PRR-βE1. The gene-editing frequency in each condition was >80%, with a PRR-βE1 deletion frequency of >56% (Fig. 4A). The gene edited cells were *in vitro* differentiated into erythroblasts under hypoxia (5% O_2)[25] and the erythroblasts derived from PRR-βE1 and βE1 gene edited HSPCs showed a significant increase in the γ-globin mRNA expression (Fig. 4B) and the percentage of HbF^{+ve} cells (Fig. 4C). Further, we performed a sickling assay by treating erythroblasts with sodium metabisulfite. Upon treatment, control and PRR edited cells underwent sickling, whereas the βE1 and PRR-βE1-edited groups had a 12-fold (HbS/CD41/CD42(-TCTT)) and 30-fold (HbS/IVS1-5) reduction in sickling, respectively (Fig. 4D, E). Variant HPLC analysis further showed that all the hemoglobin in the βE1 and PRR-βE1 edited cells was composed of HbF tetramers with nearly complete reduction of HbS (Fig. 4F-G).

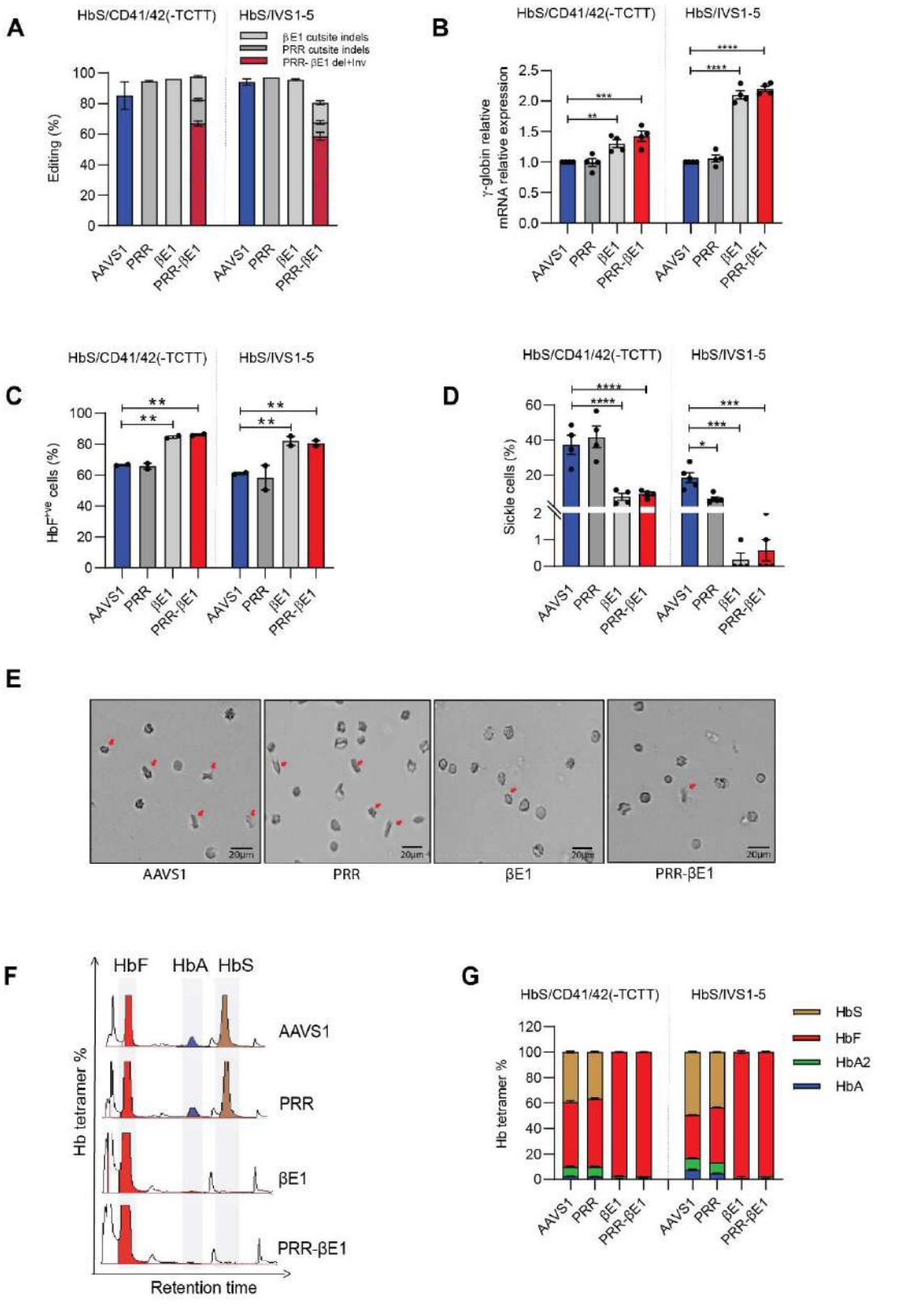

Figure 4

Figure 4. PRR-βE1 gene edited patient HSPCs reverses sickle cell disease phenotype

Plerixafor mobilized HSPCs from sickle cell patients of genotype HbS/CD41/42(-TCTT) and HbS/IVS1-5 were gene edited for *AAVS1*, PRR, βE1, and PRR-βE1. Error bars represent mean ± SEM. $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, $^{****}p \leq 0.0001$ (two-way ANOVA followed by Dunnett's test). (A) Percentage of gene editing. Indels

measured by sanger sequencing and ICE analysis. Deletion/inversion (Del+Inv) (red checker box) in PRR-βE1 quantified by ddPCR. The indels in the PRR-βE1 edited cells (grey checker box) were assessed using ICE analysis. Donor = 2, n = 4.

(B) Relative globin mRNA expression. The patient HSPCs were gene edited for PRR, βE1, and PRR-βE1 and differentiated into erythroblasts. Real-time PCR analysis was used for mRNA quantification and the globin chain expression was normalised with β-actin. The patient genotype is indicated at the bottom. Donor = 2, n = 4. (C) Percentage of HbF^{+ve} cells. The gene edited patient HSPCs were differentiated into erythroblasts and intracellular HbF positive cells were analyzed by FACS. Donor = 2, n = 4. (D) Percentage of sickle cells. Gene edited patient HSPCs were differentiated into erythroblasts in hypoxia (5% O2) and treated with 1.5% sodium metabisulfite. Cells were scored from random fields using EVOS FL Auto Imaging System microscope. Atleast 8 fields were analysed. Each field contained a minimum of 150 cells. Donor = 2, n = 4. (E) Representative image of sickle cells (red arrow) and non-sickled cells. (F) Representative variant HPLC chromatogram showing HbA, HbF and HbS. Donor = 2, n = 4. (G) Proportion of hemoglobin tetramer. The gene edited patient HSPCs were differentiated into erythroblasts and the hemoglobin tetramers were analyzed by variant HPLC. Donor = 2, n = 4.

PRR-βE1 gene-edited patient HSPCs reverse β-thalassemia phenotype

To test the therapeutic potential of our gene-editing strategy in reversing β-thalassemia defects, we edited HSPCs obtained from β-thalassemia patients of three different β^0/β^0 genotypes: CD26 (G>A)/IVS1-5 (G>C), IVS1-5 (G>C), and CD30 (G>A). These β-thalassemia mutations are highly prevalent in India and Southeast Asian countries [40,41]. Due to poor peripheral blood mononuclear cell (PBMNC) yield, CD26 (G>A)/IVS1-5 (G>C) PBMNCs (Peripheral blood mononuclear cells) were differentiated into erythroblasts and edited on day 8 of erythroid differentiation. The PRR-βE1 gene editing efficiency remained >80% in all the genotypes (Fig. 5A). In *in vitro* erythropoiesis, the ratio of α to non-α-globin chains were also observed to be reduced (Fig. 5B, C), suggesting the reduction of free α-globin chains. Western blot analysis of CD30 (G>A) gene-edited cells further confirmed that PRR-βE1 editing resulted in enhanced induction of γ-globin chains (Fig. 5D).

Ineffective erythropoiesis, the classical phenotype of β-thalassemia, results from increased reactive oxygen species (ROS) levels, apoptosis of erythroid progenitors, and reticulocyte maturation arrest[4,42]. Erythroblasts originated from the PRR-βE1 gene-edited group showed a modest decrease in the proportion of apoptotic erythroblasts (stained by Annexin V) (Fig. 5E-F), and importantly, a three-fold increase in reticulocyte generation (Fig. 5G-H) compared with the control. All these findings suggest that PRR-βE1 gene editing functionally rescues erythropoiesis in β-thalassemia by robust activation of γ-globin and silencing of defective β-globin.

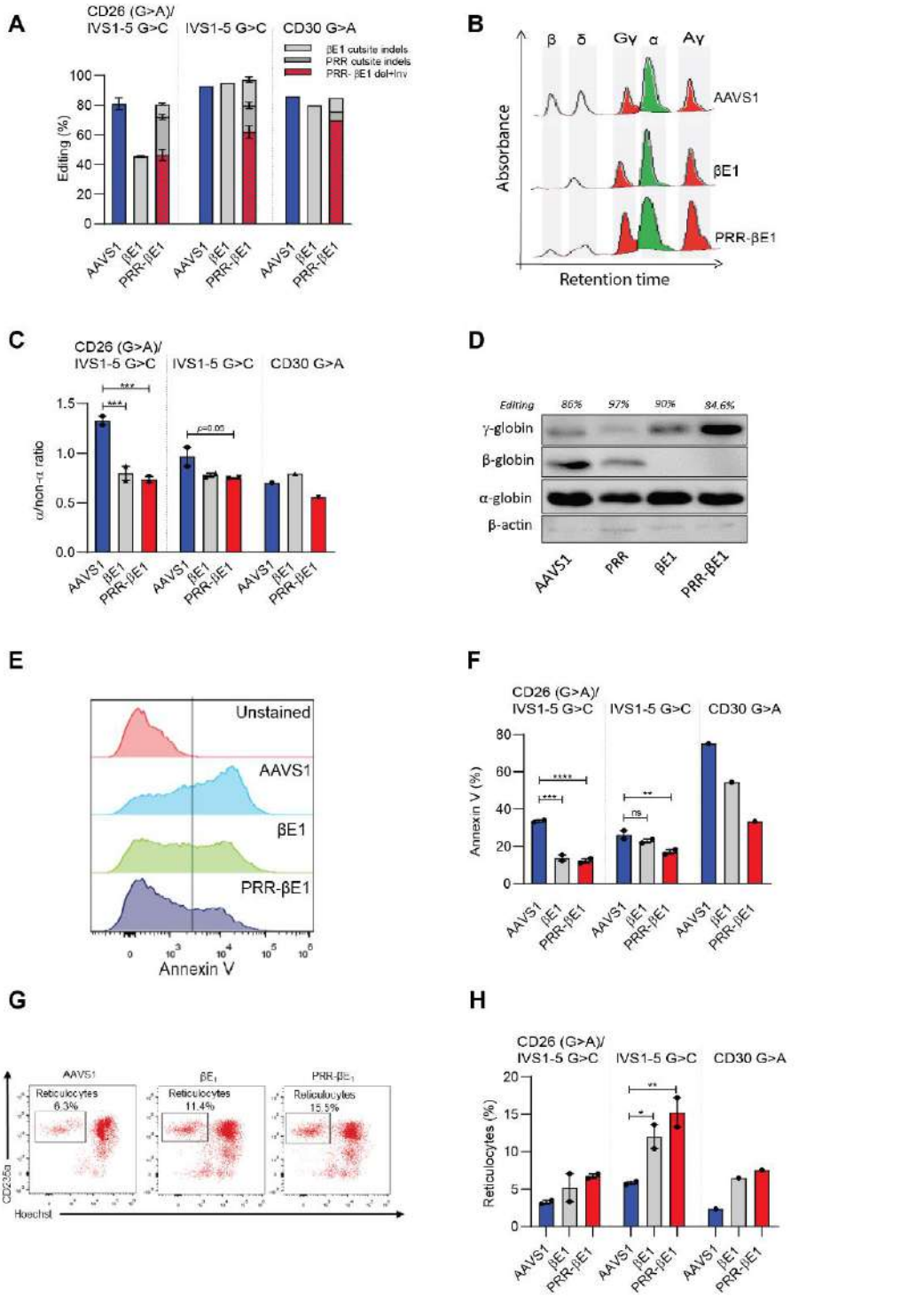

Figure 5

Figure 5. PRR- βE1 gene edited patient HSPCs reverse β-thalassemia phenotype

The G-CSF mobilized HSPCs from β- thalassemia patients of genotype IVS1-5 (G>C), and CD30 (G>A) were gene edited for *AAVS1*, βE1, and PRR-βE1. For HbE (G>A)/IVS1-5 (G>C), the PBMNCs were differentiated into erythroblasts and gene edited for *AAVS1*, βE1, and PRR-βE1. Error bars represent mean ± SEM. $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$ (two-way ANOVA followed by Dunnett's test). Donor = 3, n = 5. (A) Percentage of gene editing

in HSPCs. Deletion/Inversion (Del+Inv) in PRR-βE1 as quantified by ddPCR. Indels of the cut sites PRR and βE1 were measured by ICE analysis. (B) Representative globin chain HPLC chromatograms. (C) α/non-α ratio in the erythroblasts generated from gene edited HSPCs. (D) Representative western blot image showing the band intensity of globin chains erythroblasts derived from *AAVS1*, PRR, βE1, and PRR-βE1 gene edited CD30 (G>A) patient HSPCs. Donor = 1, n = 1. The editing in PRR and βE1 indicates the percentage of Indels in ICE analysis and for PRR-βE1 edited, the percentage of editing includes the deletion + inversion quantified by ddPCR, cut site A and cut site B indels by ICE analysis. Donor = 1, n = 1. (E) Representative flow cytometry image of Annexin V staining. (F) Percentage of Annexin V in the erythroblasts generated from gene edited HSPCs. (G) Representative flow cytometry plots of reticulocytes marked by $CD235a^{+}/Hoechst^{-}$. (H) Percentage of reticulocytes generated from gene edited HSPCs.

PRR-βE1 gene -edited HSPCs have intact genome integrity

Reportedly, Cas9-generated DNA double strand breaks pose a risk of genome-wide effects, such as genomic rearrangements [28,43]. Using HSPCs with micronuclei as a readout for cells with genomic instability, we microscopically evaluated individual HSPCs. In these experiments, we edited HSPCs using HiFi-Cas9, which has been demonstrated to minimize off-target editing and off-target mediated translocation [28,44]. HiFi-Cas9 retained the same frequency of on-target gene editing obtained in our earlier experiments with wild-type Cas9 (Fig 6A). Mitomycin C, an interstrand crosslinker that induces chromosomal rearrangements was used as positive control. The frequency of micronuclei-positive HSPCs in PRR-βE1 edited HSPCs was not significantly greater than in unedited control HSPCs (Fig. 6B).

As a second method, we employed array-based KaryoStat analysis to identify chromosomal abnormalities. The edited HSPCs were expanded for 7 days to magnify any potential defect. The whole-genome coverage analysis with a resolution of >1 Mb revealed that the PRR-βE1 gene-edited HSPCs exhibited neither loss nor gain of chromosomal copy number (Fig. 6C).

To analyze genomic integrity with the highest resolution possible, we performed chromosomal aberrations analysis by single-targeted ligation-mediated PCR sequencing (CAST-Seq), which is sensitive enough to detect a single translocation event in 10,000 cells and can classify the type of structural variation[28]. CAST-Seq detects chromosomal abnormalities caused by on-target editing as well as the fusion of off-target edited sites to the on-target region. Upon gene editing HSPCs with HiFi-Cas9, CAST-Seq identified a single off-target- mediated translocation event between chr16: 684705-685217, which codes for the 3'-untranslated region (UTR) of WD repeat domain 24 (WDR24), and the on-target site (Fig. 6D). This translocation was identified in two of our four CAST-Seq runs. The number of unique footprints (CAST-Seq hits), which testify to this translocation, is very low (13 hits) compared with the cumulated 58,897 on-target hits. This indicates that this particular translocation is an ultra-rare event, happening at a frequency close to the lower limit of detection of CAST-Seq (i.e., 0.01%). No homology-mediated translocations event were identified in the modified HSPCs. All these experiments indicate that chromosomal abnormalities occurred at very low frequency and the PRR-βE1 gene editing does not majorly compromise the integrity of the genome.

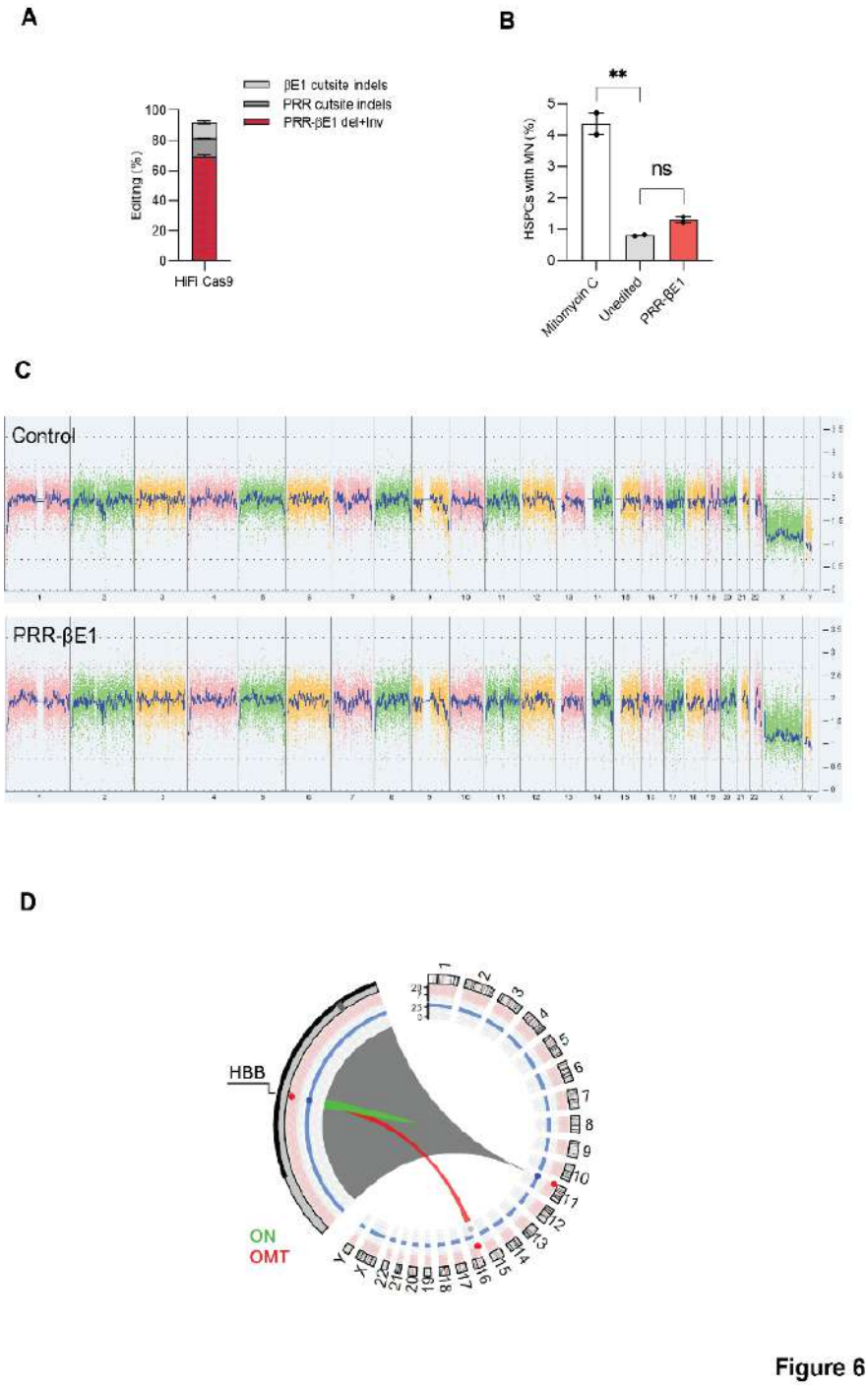

Figure 6

Figure 6 PRR-βE1 gene edited HSPCs have intact genome integrity

(A) Percentage of gene editing on PRR-βE1 editing in healthy donor HSPCs using HiFi Cas9. Indels measured by sanger sequencing and ICE analysis. Deletion/inversion (Del+Inv) (red checker box) in PRR-βE1 quantified by ddPCR. Donor = 1, n = 2. (B) Percentage of micronucleus (MN) in Mitomycin C treated, unedited, and PRR-βE1 gene edited HSPCs scored 48 hrs post nucleofection after staining with Giemsa. Donor =1, n = 2. (C) KaryoStat analysis of healthy donor HSPCs gene edited for PRR-βE1 deletion. Donor = 1, n = 2. (D) Circos plot showing off-target mediated translocation between the PRR-βE1 on-target site and βE1 off-target site in PRR-βE1 edited samples present in chr16 identified by CAST-Seq. Donor = 1, n = 4. Error bars represent mean ± SEM, ns – nonsignificant, $**p \leq 0.01$ (one-way ANOVA followed by Dunnett's multiple comparisons test).

PRR-βE1 gene editing reconfigures chromosome looping and alters globin expression

Long-range chromatin interaction of the locus control region (LCR) and the promoters in the β-globin cluster regulate developmental stage-specific expression of globin genes [45]. To test the potential impact of PRR-βE1 gene editing on the configuration of the β-globin cluster, we employed a circular chromosome conformation capture (4C) assay. An interaction between hypersensitive site 1 (HS1) within the LCR and the *HBG2* promoter was observed in HUDEP-2 control cells. However, this interaction was enhanced in HUDEP-2 clones harbouring a PRR-βE1 biallelic deletion. Furthermore, the interaction between other HS sites and *HBG2* promoters was newly gained in PRR-βE1 deleted cells. (Fig. 7A). These data suggests that genomic proximity between the LCR and the *HBG* gene increases upon PRR-βE1 deletions and thus reactivates γ -globin in edited cells.

Thereafter, to understand the trans-acting factors involved in β-globin reactivation in the PRR-βE1 gene-edited cells, transcriptome analysis was carried out using the erythroblasts generated *in vitro* from gene-edited HSPCs. This analysis confirmed the overexpression of *HBG1* and *HBG2* with simultaneous downregulation of *HBB*. HBG was not among the top 20 significantly upregulated candidates in βE1 and showed a distinct set of upregulated genes than PRR-βE1 (Fig. 7B). This indicates that different pathways are involved in PRR-βE1 and βE1 editing for γ-globin activation, with βE1 editing resulting in a weaker level of γ-activation on comparison. Cluster per million (cpm) values for the globin transcripts obtained from RNA sequencing further supports higher γ-globin induction on PRR-βE1 editing (Fig. 7C).

The transcriptome analysis and the followed-up Real-time PCR analysis indicated the overexpression of HBBP1 in PRR-βE1 gene-edited cells (Fig. 7D). HBBP1 was recently implicated in γ-globin activation [46]. BGLT-3, which is reported to promote transcriptional assembly at the γ-globin promoter was seen to be abundant in edited cells by RT-PCR (Reverse transcription polymerase chain reaction) (Fig. 7E)[47].

All these findings suggest that the γ-globin activation in PRR-βE1 gene-edited cells occurs through altered chromatin looping mediated by promoter competition for the LCR.

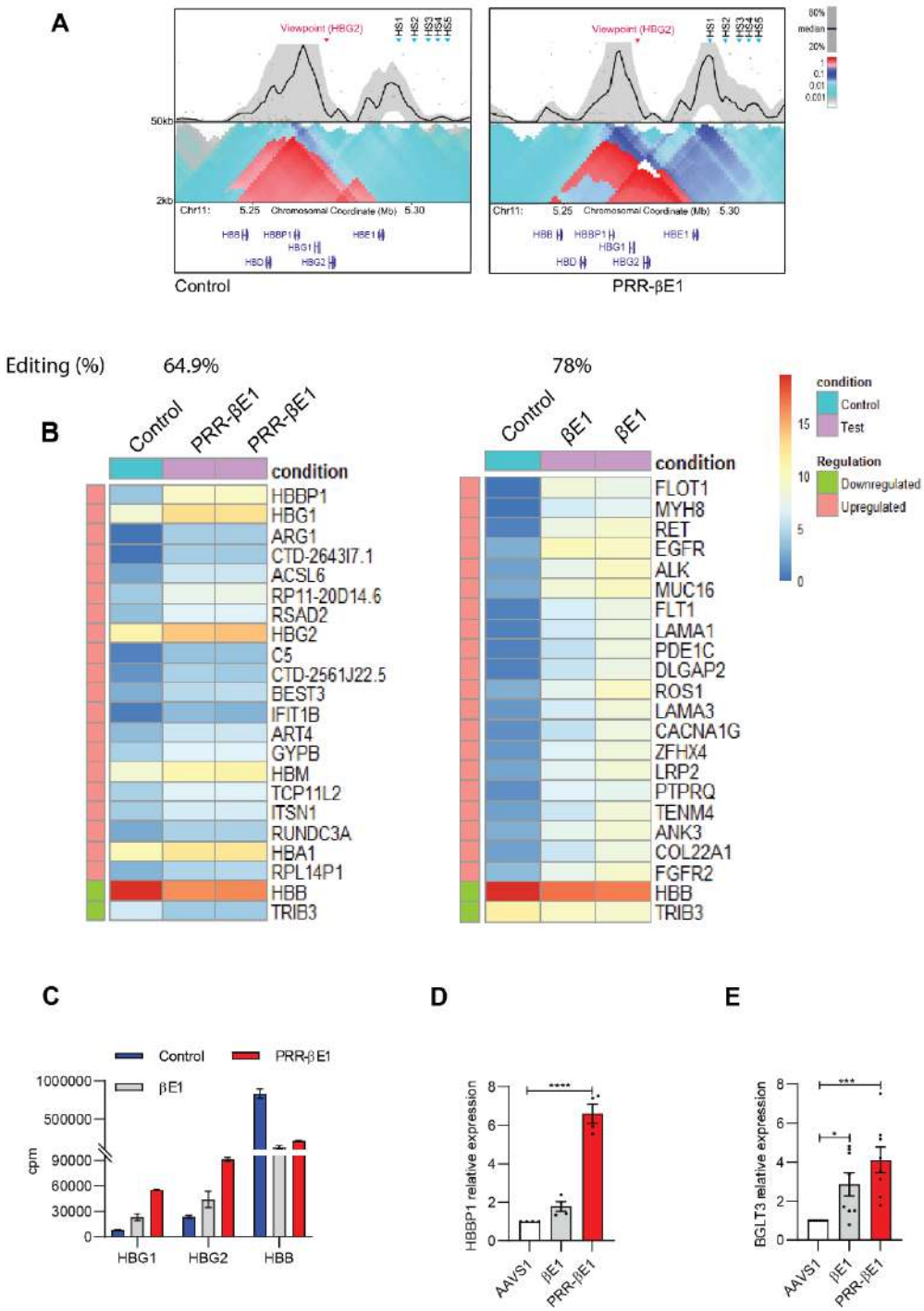

Figure 7

Figure 7. PRR-βE1 gene editing reconfigures chromosome looping and alters globin expression

(A) 4C analysis of single cell sorted control and two PRR-βE1 bi-allelic gene edited HUDEP-2 clones using *HBG2* promoter as a viewpoint. n = 4. (B) Heat-map of the top differentially expressed genes of erythroblasts derived from PRR-βE1 and βE1 gene edited HSPC indicating the relative gene expression pattern of genes up and downregulated compared to control. Donor = 1, n = 2. (C) Cluster per million (cpm) values for the globin transcripts obtained from RNA sequencing. (D) Relative HBBP1 mRNA expression in erythroblasts derived from βE1 and PRR-βE1 gene edited HSPCs compared to *AAVS1*. The globin chain expression is normalised with β-actin. Donor =1, n = 4. (E) Relative BGLT3 mRNA expression in erythroblasts derived from βE1, an PRR-βE1

gene edited HSPCs compared to *AAVS1*. The globin chain expression is normalised with β-actin. Donor =2, n = 8. Error bars represent mean ± SEM.

Statistical Analysis

All the statistical analysis was performed using PRISM GraphPad 8 package (GraphPad Software Inc., San Diego, CA, USA). Data analysis was done using unpaired t-test, one-way and two-way ANOVA followed by Dunnett's multiple comparisons test as indicated in figure legends. Error bars denotes ± SEM. Number of independent experimental replicates (n), number of donors used are indicated in the figure legend. P value < 0.05 is considered as statistically significant.

Discussion

Genetic reactivation of developmentally silenced HbF has gained considerable attention as a potential therapy for the broad spectrum of β-hemoglobinopathies. In this study, we have identified the PRR-βE1 sequence as a core HbF regulatory region present in all the deletional HPFH mutations. When present, PRR-βE1 effectively reverses the cellular phenotype of both SCD and β-thalassemia major by disrupting the production of defective β-globin and concurrently inducing robust HbF production through LCR switching mechanism. We specifically showed that PRR-βE1 gene-edited HSPCs have sustained engraftment, repopulation fitness, and genome integrity, highlighting the potential of this approach for future clinical studies.

Among the naturally existing mutations that produce pancellular HbF, deletional HPFH mutations are highly prevalent and are shown to generate a high frequency of HbF^{+ve} RBCs [13]. Even a heterozygous deletion can result in an HbF level of 65.6% with 8.9 g/dL of hemoglobin on co-inheritance with β-thalassemia [32]. Identifying the core region in HPFH deletions will enable us to recreate the HPFH phenotype by gene editing only the core region. PRR region is conserved in δβ- thalassemia but excised in HPFH deletions [48]. However, deletion of the PRR site alone did not activate the HbF in our studies, consistent with earlier observations [48]. Even a deletion of 10.5 kb spanning PRR region to the region located before the β-globin promoter had little effect on γ-globin production. In contrast, disruption of βE1 alone induced γ-globin production. Shen *et al*. showed that the improved γ-globin levels obtained by disrupting the HBB gene and its regulatory region is not sufficient to compensate for the loss of β-globin [49]. Our results provide compelling evidence that simultaneous disruption of PRR region and β-globin reactivates γ-globin robustly without negatively impacting the erythroid maturation. While PRR disruption ensures that δβ-thalassemia-like phenotype is not created and the βE1 cut site confirms that even in the case of inversion or indels, the β-globin gene expression gets ablated and is associated with γ-globin production: thus PRR-βE1 is a more potent target than the original Sicilian HPFH.

We observed a new genomic interaction between *HBG2* and a region downstream of *HBB* in PRR-βE1- edited cells (Fig. 7A). Interestingly, this interaction site is deleted in Sicilian HPFH.

Whether this region has a regulatory role on *HBG* expression and contributes to increased HbF in PRR-βE1 over Sicilian HPFH is to be explored.

While our manuscript was under preparation, two new articles provided deeper insight into the PRR-βE1 regulatory region. Topfer *et al*., analyzed both HPFH and δβ-thalassemia deletions and identified that the disruption of the β-globin promoter is sufficient for HbF reactivation[19]. βE1 gene- editing also disrupts the β-globin and reactivates the γ-globin, supporting the earlier findings[19,50]. The loss of β-globin reduces the levels of ATF4 which in turn decreases MYB and BCL11A to upregulate γ- globin [50]. This approach also resulted in decreased γ-globin activation than PRR-βE1 editing (Fig 2D, 5D and 7C). Secondly, through a single-cell functional assay, Shen *et al*. demonstrated globin chain imbalance in erythroid colonies with β-globin (HBB-HBD$^{-/-}$) disruption but not in colonies with an *HBB-3.5kb* deletion that encompasses PRR 3.5kb region, HBD and HBB.[49]. A long-range distal regulatory role has been proposed for the region upstream of HBD and this merges with the functional role of BCL11A in the HbF reactivation[49].

Ramdier *et al*. reported a strategy of combining lentiviral transduction of anti-sickling β-globin and gene editing to disrupt endogenous β-globin and enhance the proportion of anti-sickling haemoglobin [51]. This clearly depicts the competition between defective endogenous β-globin chains and exogenously supplemented globin chains for haemoglobin tetramer formation. In the ongoing clinical trial CLIMB SCD-121, HSPCs from SCD patient were gene edited for the BCL11A erythroid-specific enhancer; the gene editing efficiency was up to 82.6%. HbF levels were 43.2% and the presence of HbS tetramers was up to 52.3% [12]. This indicates that, irrespective of the gene-editing efficiency and γ-globin activation efficacy, an intact β-globin regulatory region allows production of mutated β-globin chains at reduced levels. Similarly, in the BCL11A shRNA clinical trial, HbS constitutes up to 70% of hemoglobin tetramers[11]. The PRR-βE1-editing strategy directly excises the promoter and coding regions of β-globin, resulting in a major reduction in the concentration of HbS, which will prevent the sickling of RBCs. This strategy will also be applicable for β-thalassemia, where the intact β-globin promoter drives production of truncated β-globin chains. Whether such approach results in any free alpha globin levels is yet to be determined.

For a while, HPFH deletions were considered to be a potential gene-editing targets. However, there were no reports on the genome integrity of the HSPCs and their ability to engraft post editing. HSPCs with a 4.9-kb deletion in the *HBG* promoter were shown to be lost post transplantation, and it was hypothesized that HSPCs having larger deletions are transplantation incompetent [52,53]. To our knowledge, this is the first study to demonstrate that HSPCs with large deletions can engraft and repopulate in both primary and secondary recipients. Our study also suggests that gene editing for large HPFH deletions is feasible without chromosomal aberrations when HiFi-Cas9 is used in conjunction with a carefully chosen sgRNA. The gene-editing approach and the cytokine pre-stimulation that we described can potentially simplify the manufacturing process and reduce the cost associated with HSPC gene therapy.

Impact of the research in the advancement of knowledge or benefit to mankind

β- thalassemia and sickle cell disorders pose a significant health burden in India with an incidence rate of 1.2 per 1000 births[2]. As gene therapy is the only curative treatment option, two recent clinical studies focusing on HbF reactivation achieved transfusion independence[11,12]. However, up to 70% of hemoglobin remained as defective sickle β-globin (HbS) and 30% being the beneficial fetal hemoglobin (HbF) bringing the therapeutic outcomes in the SCD patients. The presence of excess of defective β-globin affects the therapeutic outcome thus, strategies that reduce or eliminate defective β-globin production are to be explored. On the other hand, is the unaffordability of gene therapy product. The currently approved Bluebird Bio gene therapy product available in the US market is 2.8 million USD and are unaffordable by majority of population worldwide.

To address these issues, we screened the β-globin cluster and identified a novel gene editing strategy that activates γ-globin and completely ablates defective sickle β-globin (Fig. 4G). Moreover, our novel target induces two folds higher γ-globin and threefold decreased β-globin than the well characterised gene editing targets that are currently being extensively studied in the clinical trials (Fig. 2J-L). As the identified PRR-βE1 gene editing retained HSPCs genome integrity (Fig. 6D), engraftment potential (Fig. 3) and ameliorates both SCD (Fig. 4) and β-thalassemia (Fig. 5), we are currently working on scaling up the gene editing process by optimising the editing conditions in clinical grade MaxCyte electroporator. Moreover, our team will be initiating the clinical studies in the near future for developing an affordable indigenous gene therapy product.

- Our study is the first one to demonstrate that HSPCs gene edited for large deletions have intact engraftment and repopulation potential *in vivo.*
- This is the only study to show the near complete absence of defective β-globin in SCD patients on gene editing with our novel approach.
- On careful designing of the gene editing reagents, we showed that the translocation events and the chromosomal abnormalities can be limited and occurs at very low frequency, not majorly compromise the genomic integrity (Fig. 6).
- We further showed that the minimal deletion of 11kb region in the β-globin cluster is adequate enough to create HPFH phenotype (naturally occurring HPFH deletion ranges from 84.9kb to 12.9kb) and the activation is mediated by the conventional LCR looping (Fig. 7A).
- We further showed that the identified PRR-βE1 deletion results in three folds γ-globin induction than the naturally occurring variants of HPFH deletions (Fig. 1H and K).

Literature references

1 Sanctis V De, Kattamis C, Canatan D, Soliman AT, Elsedfy H, Karimi M *et al.* β-thalassemia distribution in the old world: An ancient disease seen from a historical standpoint. *Mediterr J Hematol Infect Dis* 2017; **9**. doi:10.4084/mjhid.2017.018.

2 Colah R, Italia K, Gorakshakar A. Burden of thalassemia in India: The road map for control. *Pediatr Hematol Oncol J* 2017; **2**: 79–84.

3 Boonyawat B, Monsereenusorn C, Traivaree C. Molecular analysis of beta-globin gene mutations among Thai beta-thalassemia children: results from a single center study. *Appl Clin Genet* 2014; **7**: 253–258.

4 Origa R. β-Thalassemia. *Genet Med* 2017; **19**: 609–619.

5 Sundd P, Gladwin MT, Novelli EM. Pathophysiology of Sickle Cell Disease. *Annu Rev Pathol* 2019; **14**: 263–292.

6 Fitzhugh CD, Hsieh MM, Allen D, Coles WA, Seamon C, Ring M *et al.* Hydroxyurea-Increased Fetal Hemoglobin Is Associated with Less Organ Damage and Longer Survival in Adults with Sickle Cell Anemia. *PLoS One* 2015; **10**: e0141706.

7 Musallam KM, Sankaran VG, Cappellini MD, Duca L, Nathan DG, Taher AT. Fetal hemoglobin levels and morbidity in untransfused patients with β-thalassemia intermedia. *Blood* 2012; **119**: 364–367.

8 Nuinoon M, Makarasara W, Mushiroda T, Setianingsih I, Wahidiyat PA, Sripichai O *et al.* A genome-wide association identified the common genetic variants influence disease severity in β0-thalassemia/hemoglobin e. *Hum Genet* 2010; **127**: 303–314.

9 Masuda T, Wang X, Maeda M, Canver MC, Sher F, Funnell APW *et al.* Gene regulation: Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science (80-)* 2016; **351**: 285–289.

10 Canver MC, Smith EC, Sher F, Pinello L, Sanjana NE, Shalem O *et al.* BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* 2015; **527**: 192–197.

11 Esrick EB, Lehmann LE, Biffi A, Achebe M, Brendel C, Ciuculescu MF *et al.* Post-Transcriptional Genetic Silencing of BCL11A to Treat Sickle Cell Disease. *N Engl J Med* 2020. doi:10.1056/nejmoa2029392.

12 Frangoul H, Altshuler D, Cappellini MD, Chen Y-S, Domm J, Eustace BK *et al.* CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β-Thalassemia. *N Engl J Med* 2021; **384**: 252–260.

13 Ringelhann B, Acquaye CTA, Oldham JH, Konotey-Ahulu FID, Yawson G, Sukumaran PK *et al.* Homozygotes for the hereditary persistence of fetal hemoglobin: The ratio of Gγ to Aγ chains and biosynthetic studies. *Biochem Genet* 1977; **15**: 1083–1096.

14 Cianetti L, Care A, Spos1 NM, Giampaolo A, Calandrini M, Petrini M *et al.* Association of heterocellular HPFH,-i-thalassaemia, and a 0-thalassaemia: haematological and molecular aspects. *J Med Genet* 1984; **21**: 263–267.

15 Steinberg MH. Fetal Hemoglobin in Sickle Hemoglobinopathies: High HbF Genotypes and Phenotypes. *J Clin Med* 2020; **9**: 1–10.

16 Venkatesan V, Srinivasan S, Babu P, Thangavel S. Manipulation of Developmental Gamma-Globin Gene Expression: an Approach for Healing Hemoglobinopathies. *Mol Cell Biol* 2020; **41**: 1–18.

17 Ye L, Wang J, Tan Y, Beyer AI, Xie F, Muench MO *et al.* Genome editing using CRISPR-Cas9 to create the HPFH genotype in HSPCs: An approach for treating sickle cell disease and β-thalassemia. *Proc Natl Acad Sci U S A* 2016; **113**: 10661–10665.

18 Antoniani C, Meneghini V, Lattanzi A, Felix T, Romano O, Magrin E *et al.* Induction of fetal hemoglobin synthesis by CRISPR/Cas9-mediated editing of the human b-globin locus. *Blood* 2018; **131**: 1960–1973.

19 Topfer SK, Feng R, Huang P, Ly LC, Martyn GE, Blobel GA *et al.* Disrupting the adult globin promoter alleviates promoter competition and reactivates fetal globin gene expression. *Blood* 2022; **139**: 2107–2118.

20 Karuppusamy K V, Demosthenes JP, Venkatesan V, Christopher AC, Babu P, Azhagiri MK *et al.* The CCR5 Gene Edited CD34 + CD90 + Hematopoietic Stem Cell Population Serves as an Optimal Graft Source for HIV Gene Therapy. 2022; **13**: 1–15.

21 Christopher AC, Venkatesan V, Karuppusamy K V, Srinivasan S, Babu P, Azhagiri MKK *et al.* Preferential expansion of human CD34+CD133+CD90+ hematopoietic stem cells enhances gene-modified cell frequency for gene therapy. *Hum Gene Ther* 2021; : 1–33.

22 Venkatesan V, Christopher AC, Karuppusamy K V, Babu P, Kumar M, Alagiri K *et al.* CRISPR / Cas9 Gene Editing of Hematopoietic Stem and Progenitor Cells for Gene Therapy Applications. 2022; : 1–20.

23 Trakarnsanga K, Griffiths RE, Wilson MC, Blair A, Satchwell TJ, Meinders M *et al.* An immortalized adult human erythroid line facilitates sustainable and scalable generation of functional red cells. *Nat Commun 2017 81* 2017; **8**: 1–7.

24 Chang KH, Smith SE, Sullivan T, Chen K, Zhou Q, West JA *et al.* Long-Term Engraftment and Fetal Globin Induction upon BCL11A Gene Editing in Bone-Marrow-Derived CD34+ Hematopoietic Stem and Progenitor Cells. *Mol Ther - Methods Clin Dev* 2017; **4**: 137–148.

25 El Hoss S, Cochet S, Godard A, Yan H, Dussiot M, Frati G *et al.* Fetal hemoglobin rescues ineffective erythropoiesis in sickle cell disease. *Haematologica* 2020; : haematol.2020.265462.

26 Zeng J, Wu Y, Ren C, Bonanno J, Shen AH, Shea D *et al.* Therapeutic base editing of

human hematopoietic stem cells. Nat. Med. 2020; **26**: 535–541.

27 Loucari CC, Patsali P, Van Dijk TB, Stephanou C, Papasavva P, Zanti M *et al.* Rapid and Sensitive Assessment of Globin Chains for Gene and Cell Therapy of Hemoglobinopathies. *Hum Gene Ther Methods* 2018; **29**: 60–74.

28 Turchiano G, Andrieux G, Klermund J, Blattner G, Pennucci V, el Gaz M *et al.* Quantitative evaluation of chromosomal rearrangements in gene-edited human stem cells by CAST-Seq. *Cell Stem Cell* 2021; **28**: 1136-1147.e5.

29 Mosesso P, Cinelli S. In Vitro Cytogenetic Assays: Chromosomal Aberrations and Micronucleus Tests. *Methods Mol Biol* 2019; **2031**: 79–104.

30 Fenech M, Chang WP, Kirsch-Volders M, Holland N, Bonassi S, Zeiger E. HUMN project: detailed description of the scoring criteria for the cytokinesis-block micronucleus assay using isolated human lymphocyte cultures. *Mutat Res Toxicol Environ Mutagen* 2003; **534**: 65–75.

31 Joly P, Lacan P, Garcia C, Couprie N, Francina A. Identification and molecular characterization of four new large deletions in the β-globin gene cluster. *Blood Cells, Mol Dis* 2009; **43**: 53–57.

32 Changsri K, Akkarapathumwong V, Jamsai D, Winichagoon P, Fucharoen S. Molecular mechanism of high hemoglobin F production in Southeast Asian-type hereditary persistence of fetal hemoglobin. *Int J Hematol* 2006; **83**: 229–237.

33 Camaschella C, Serra A, Gottardi E, Alfarano A, Revello D, Mazza U *et al.* A new hereditary persistence of fetal hemoglobin deletion has the breakpoint within the 3' beta-globin gene enhancer. *Blood* 1990; **75**: 1000–1005.

34 Kulozik AE, Yarwood N, Jones RW. The Corfu αβ° Thalassemia: A Small Deletion Acts at a Distance to Selectively Abolish β Globin Gene Expression. *Blood* 1988; **71**: 457–462.

35 Ye L, Wang J, Tan Y, Beyer AI, Xie F, Muench MO *et al.* Genome editing using CRISPR-Cas9 to create the HPFH genotype in HSPCs: An approach for treating sickle cell disease and β-thalassemia. *Proc Natl Acad Sci U S A* 2016; **113**: 10661–10665.

36 Anand R, Boehm C, Kazazian HJ, Vanin E. Molecular Characterization of a β0 - Thalassemia Resulting From a 1.4 Kilobase Deletion. *Blood* 1988; **72**: 636–641.

37 Dimovski AJ, Efremov DG, Jankovic L, Juricic DPD, Efremov GD. A β°-thalassaernia due to a 1605 bp deletion of the 5'β-globin gene region. *Br J Haematol* 1993; **85**: 143–147.

38 Métais JY, Doerfler PA, Mayuranathan T, Bauer DE, Fowler SC, Hsieh MM *et al.* Genome editing of HBG1 and HBG2 to induce fetal hemoglobin. *Blood Adv* 2019; **3**: 3379–3392.

39 McIntosh BE, Brown ME, Duffin BM, Maufort JP, Vereide DT, Slukvin II *et al.*

Nonirradiated NOD,B6.SCID Il2rγ-/- Kit(W41/W41) (NBSGW) mice support multilineage engraftment of human hematopoietic cells. *Stem cell reports* 2015; **4**: 171–180.

40 Yang Z, Cui Q, Zhou W, Qiu L, Han B. Comparison of gene mutation spectrum of thalassemia in different regions of China and Southeast Asia. *Mol Genet Genomic Med* 2019; **7**: 680.

41 Sinha S, Black ML, Agarwal S, Colah R, Das R, Ryan K *et al.* Profiling β-thalassaemia mutations in India at state and regional levels: implications for genetic education, screening and counselling programmes. *Hugo J* 2009; **3**: 51.

42 Arlet J-B, Ribeil J-A, Guillem F, Negre O, Hazoume A, Marcion G *et al.* HSP70 sequestration by free α-globin promotes ineffective erythropoiesis in β-thalassaemia. *Nature* 2014; **514**: 242–246.

43 Leibowitz ML, Papathanasiou S, Doerfler PA, Blaine LJ, Sun L, Yao Y *et al.* Chromothripsis as an on-target consequence of CRISPR–Cas9 genome editing. *Nat Genet* 2021; **53**: 895–905.

44 Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z *et al.* High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nat 2015 5297587* 2016; **529**: 490–495.

45 Vinjamur DS, Bauer DE, Orkin SH. Recent progress in understanding and manipulating haemoglobin switching for the haemoglobinopathies. *Br J Haematol* 2018; **180**: 630–643.

46 Ma SP, Xi HR, Gao XX, Yang JM, Kurita R, Nakamura Y *et al.* Long noncoding RNA HBBP1 enhances γ-globin expression through the ETS transcription factor ELK1. *Biochem Biophys Res Commun* 2021; **552**: 157–163.

47 Ivaldi MS, Diaz LF, Chakalova L, Lee J, Krivega I, Dean A. Fetal γ-globin genes are regulated by the BGLT3 long noncoding RNA locus. *Blood* 2018; **132**: 1963–1973.

48 Chung JE, Magis W, Vu J, Heo S-J, Wartiovaara K, Walters MC *et al.* CRISPR-Cas9 interrogation of a putative fetal globin repressor in human erythroid cells. *PLoS One* 2019; **14**: e0208237.

49 Shen Y, Verboon JM, Zhang Y, Liu N, Kim YJ, Marglous S *et al.* A unified model of human hemoglobin switching through single-cell genome editing. *Nat Commun 2021 121* 2021; **12**: 1–12.

50 Boontanrart MY, Schröder MS, Stehli GM, Banović M, Wyman SK, Lew RJ *et al.* ATF4 Regulates MYB to Increase γ-Globin in Response to Loss of β-Globin. *Cell Rep* 2020; **32**: 107993.

51 Ramadier S, Chalumeau A, Felix T, Othman N, Aknoun S, Casini A *et al.* Combination of lentiviral and genome editing technologies for the treatment of sickle cell disease. *Mol Ther* 2022; **30**: 145–163.

52 Métais JY, Doerfler PA, Mayuranathan T, Bauer DE, Fowler SC, Hsieh MM *et al.* Genome editing of HBG1 and HBG2 to induce fetal hemoglobin. *Blood Adv* 2019; **3**: 3379–3392.

53 Humbert O, Radtke S, Samuelson C, Carrillo RR, Perez AM, Reddy SS *et al.* Therapeutically relevant engraftment of a CRISPR-Cas9-edited HSC-enriched population with HbF reactivation in nonhuman primates. *Sci Transl Med* 2019; **11**. doi:10.1126/SCITRANSLMED.AAW3768.