Correction of IVS I-110(G>A) β-thalassemia by CRISPR/Cas- and TALEN-mediated disruption of aberrant regulatory elements in human hematopoietic stem and progenitor cells

Petros Patsali,[1,2,†] Giandomenico Turchiano,[3,4,†] Panayiota Papasavva,[1,5] Marianna Romito,[3,4] Constantinos C. Loucari,[1,5] Coralea Stephanou,[1,2] Soteroulla Christou,[6] Maria Sitarou,[6] Claudio Mussolino,[3,4] Tatjana I. Cornu,[3,4] Michael N. Antoniou,[2] Carsten W. Lederer,[1,5,‡,] Toni Cathomen[3,4,7,‡,*] and Marina Kleanthous[1,5,‡,*]*

[†]PP and GT contributed equally to this work.

[‡]CWL, TC and MK contributed equally to this work.

[1]Department of Molecular Genetics Thalassemia, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus; [2]Department of Medical and Molecular Genetics, King's College London, London, UK; [3]Institute for Transfusion Medicine and Gene Therapy, Medical Center, University of Freiburg, Freiburg, Germany; [4]Center for Chronic Immunodeficiency, Medical Center, University of Freiburg, Freiburg, Germany; [5]Cyprus School of Molecular Medicine, Nicosia, Cyprus; [6]Thalassemia Center, Cyprus Ministry of Health, Cyprus; and [7]Faculty of Medicine, University of Freiburg, Freiburg, Germany.

Correspondence: CARSTEN W. LEDERER - lederer@cing.ac.cy or
TONI CATHOMEN - toni.cathomen@uniklinik-freiburg.de or
MARINA KLEANTHOUS - marinakl@cing.ac.cy

doi:10.3324/haematol.2018.215178

Supplementary Data

Correction of IVS I-110(G>A) β-thalassemia by CRISPR/Cas- and TALEN-mediated disruption of aberrant regulatory elements in human hematopoietic stem and progenitor cells

Petros Patsali,[1,2,†] Giandomenico Turchiano,[3,4,†] Panayiota Papasavva,[1,5], Marianna Romito,[3,4] Constantinos C. Loucari,[1,5] Coralea Stephanou,[1,2] Soteroulla Christou,[6] Maria Sitarou,[6] Claudio Mussolino,[3,4] Tatjana I. Cornu,[3,4] Michael N. Antoniou,[2] Carsten W. Lederer,[1,5,‡,*] Toni Cathomen,[3,4,7,‡,*] and Marina Kleanthous,[1,5,‡,*]

[1] Department of Molecular Genetics Thalassemia, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus

[2] Department of Medical and Molecular Genetics, King's College London, London, United Kingdom

[3] Institute for Transfusion Medicine and Gene Therapy, Medical Center – University of Freiburg, Freiburg, Germany

[4] Center for Chronic Immunodeficiency, Medical Center – University of Freiburg, Freiburg, Germany

[5] Cyprus School of Molecular Medicine, Nicosia, Cyprus

[6] Thalassemia Centre, Cyprus Ministry of Health, Cyprus

[†] Shared first authors

[‡] Shared last authors

[*] Corresponding authors

Corresponding authors

E-mail: lederer@cing.ac.cy (C.W.L.), toni.cathomen@uniklinik-freiburg.de (T.C.), marinakl@cing.ac.cy (M.K.)

Running title

Disruption as β-thalassemia gene therapy

Article Summary

- Disrupting aberrant regulatory elements effectively restores normal splicing for the *HBB*$^{IVSI\text{-}110(G>A)}$ mutation.
- CRISPR/Cas9 ribonucleoprotein or TALEN mRNA delivery reaches high correction efficiencies in patient-derived CD34^{+} cells without enrichment and at minimal *HBD* off-target activity

Content

SUPPLEMENTARY MATERIALS AND METHODS

Study design

The goal of this study was establishment of a gene correction approach based on disruption of aberrant regulatory elements. For proof of principle we targeted *HBB*$^{IVSI-110(G>A)}$, using mutation-specific TALENs and an RGN nuclease. Therapeutic potential of designer nucleases was then tested on *HBB*$^{IVSI-110(G>A)}$-homozygous HSPCs of four different patients, based on the correction of erythroid terminal differentiation, correction of HBB/HBA ratios and hemoglobinization as key disease parameters. On-target and shortlisted off-target events where then characterized in modified *HBB*$^{IVSI-110(G>A)}$-homozygous HSPCs by targeted deep sequencing.

Design, cloning and synthesis of designer nucleases

Suitable for NHEJ-based repair are mutations positioned at least 5 bp from exon-intron borders. In particular for mutations closer than 50 bp to exon-intron borders, disruption of aberrant regulatory elements may affect additional regulatory regions, such as the branch point site (BPS) or the polypyrimidine tract (PPT).

Transcription activator like-effector nuclease (TALEN) expression constructs

TALEN constructs were assembled using Golden Gate assembly.[16] Two TALEN pairs, each consisting of one left and one right monomer, relative to the target site, were designed targeting the *HBB*$^{IVSI-110(G>A)}$-adjacent region, two left monomers (HBB TALEN L1 and L2) and a common right monomer (HBB TALEN R1), all holding the commonly used repeat-variable-diresidues (RVDs), with amino acid pairs NI, NG, HD and NN. Additional right monomers, R2, R3 and R4, targeted the same sequence as R1 but carried varying substitutions of NN with NK RVD modules.[17] The NN module can bind both guanine and adenine whereas NK preferentially binds only guanine but with lower overall binding affinity than NN [17]. TALEN RVD sequences are shown in table S2. After functional evaluation in HEK293T cells, R1, L1 and L2 monomers were cloned from pVAX CMV TAL to pPIX-K_CMV and confirmed by sequencing with primers SEQ TAL FW and SEQ TAL RV (table S3) before *in vitro* transcription.[18]

HBB RGN construct

Guide RNA (gRNA) inserts were produced by oligonucleotide annealing and cloned into the MLM3636 gRNA vector backbone (*Addgene*, UK) using standard cloning techniques. Positive clones were confirmed by sequencing with CMV FW primer (table S3). A schematic representation of the *HBB* RGN target sequence is shown in Figure 1 and supplementary Figure S1. Sequences of gRNAs are shown in table S5.

HEK 293T cells

HEK 293T cells were polyethylenimine-transfected with TALEN-monomer-encoding plasmids (left/right) or Cas9- and gRNA-encoding plasmid, including pUC118 control plasmid to level total plasmid amount, and pEGFP-N1 (*Clontech*, USA) or pLLS mOrange N1 (for GFP reporter assay; *Addgene*, UK) as reporters of transfection efficiencies. Targeted disruption efficiency on *HBB* and *HBD* genes was assessed by T7E1 assay.[15,16] An episomal disruption assay targeting the in-frame *HBB*$^{IVSI-110(G>A)}$ target sequence of a novel *HBB*$^{IVSI-110}$ GFP reporter was assessed by flow cytometry for disruption (*HBB*$^{IVSI-110}$ GFP), transfection efficiency (pLSS mOrange N1; *Addgene*, UK) and cell death (SYTOX Red; *Thermo Fisher Scientific*, USA). GFP- and CCR5-specific nuclease-expressing plasmids (a TALEN pair and an RGN each) served as positive and negative controls, respectively.

CD34^{+} HSPCs

Primary human CD34^{+} HSPCs from patients and controls were obtained after written informed consent and sequence confirmation of normal or homozygous *HBB*$^{IVSI-110(G>A)}$ (HGVS name: *HBB*:c.93-21G>A) status. Cells were handled and cultured as described,[6] with additional CD34^{+} magnetic-activated cell sorting (*Miltenyi Biotec*, Germany) after buffy-coat isolation and before expansion culture. Cells were collected during erythroid differentiation for RNA (day 3) and protein (day 7). For gDNA disruption, TALEN mRNA pairs or preassembled RGN RNP complex of gRNA (*Synthego Corporation*, California, USA) and Cas9 protein (*PNA*

bio), were nucleofected using buffer P3 and the CA-137 protocol of the 4D-Nucleofector (*Lonza, Inc*, Switzerland) before initial culture at hypothermic (32 °C) conditions and detection of transfection efficiency (GFP) and cell death (7-aminoactinomycin D, 7-AAD; *Thermo Fisher Scientific*, USA). Cell viability and functional correction after targeted disruption and erythroid differentiation were assessed microscopically by quantification of trypan-blue positive cells and by differential morphological scoring after staining with o-Dianisidine/May-Grünwald-Giemsa, respectively.

Assessment of disruption efficiency

T7 endonuclease I assay (T7E1 assay)

T7E1 assays were performed as described previously [15,16]. Briefly, genomic DNA was extracted using QIAamp mini blood kit (Qiagen) according to manufacturer's instructions. The genomic region encompassing the target sequence of the designer nucleases, *HBB, HBD* and *CCR5*, were PCR amplified with high-fidelity DNA polymerases (Phusion and Q5-DNA polymerases) using the appropriate primers and conditions (table S3 and S6). PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) according to manufacturer's instructions. 300 ng of purified PCR products were denatured and re-annealed by slow cooling to 35 °C at -0.1 °C /sec, for the formation of heteroduplexes. Re-annealed PCR products were split into two tubes (≈150 ng each) and one was treated with 6 units of T7 endonucleases I (New England Biolabs, MA, USA) at 37 °C for 20 min. Untreated and treated samples were resolved by electrophoresis on 2.5% agarose gel. Targeted disruption percentage was measured as the fraction of band intensities of the cleaved and uncleaved bands.

TIDE

Assessment of genome editing by designer nucleases on the target locus was measured by the online web tool TIDE (https://tide.nki.nl/) [19]. TIDE calculates the editing efficacy and identifies the predominant types of indels in the DNA of a targeted cell pool based on quantitative sequence trace data from standard capillary sequencing reactions. Purified PCR products were Sanger sequenced using the appropriate primers (table S3) and traces analyzed with TIDE. INDEL frequencies were measured relative to INDEL frequencies of the nuclease-free negative control sample.

RT-qPCR-based quantifications

1–5 x 10^6 erythroid precursors were collected by centrifugation at 300 x g for 5 min at 4 °C. Total RNA was extracted and isolated using TRIZOL™ kit (*Invitrogen*, *Thermo Fisher Scientific*) and dissolved in RNase-free water (*Sigma-Aldrich*, UK). 1 µg of RNA was treated with 0.5 units of DNase I (*Invitrogen*, *Thermo Fisher Scientific*) according to the manufacturer's instructions. Gene expression quantification was performed in a two-step reverse-transcription qPCR. First, 250 ng of DNase I-treated RNA were reverse-transcribed using the TaqMan Reverse transcription PCR kit and random hexamers (*Applied Biosystems*, MA, USA) following the manufacturer's instructions. Second, approximately 12.5 ng of cDNA/sample and triplicate non-template controls were used in the qPCR reactions on a 7900HT Fast Real-Time PCR System (*Applied Biosystems*).

SYBR Green I chemistry containing a ROX passive reference dye (*Applied Biosystems*) was used for the relative quantification of human *HBB* expression, in which HBA expression was used as reference to normalize for differential erythroid differentiation using the $2^{-\Delta(\Delta Ct)}$ analysis method. Non-edited samples were used as negative controls. Default cycle conditions were used, starting with activation of the polymerase at 95 °C for 10 min, followed by 40 cycles of two-step amplification by denaturation and annealing/extension, at 95 °C for 15 sec and 60 °C for 1 min, respectively. Contamination and formation of primer dimers was assessed by dissociation curve analysis at the last stage of the qPCR reaction.

Variant-specific quantification of samples was performed by duplex PCR with the Multiplex PCR Kit (*Qiagen*, Germany), in triplicate against a plasmid-based standard curve holding the aberrant and normal amplicons, as published elsewhere.[20] Briefly, duplex PCR reactions consisted of 6.25 µL of 2 x Qiagen multiplex master mix (Qiagen), 900 nM of each primer (HBB EX1_FW_3 and HBB EX2_RV_1, (table S3), 250 nM of each specific probe (IVSI-110_MGB_VIC and wtHBB_ZNA_FAM), (table S3) and as template 2 µL of plasmid DNA standard

curve dilutions or (≈6.25 ng/µL) 4x diluted sample cDNA, topped up with water to 12.5 µL final volume. The real-time PCR reaction started with activation of the polymerase at 95 °C for 15 min, followed by 40 cycles of denaturation at 95 °C for 30 sec and annealing / extension at 60 °C for 1 min, respectively. Based on the plasmid standard curve, absolute quantities of each variant mRNA were calculated in order to allow assessment of the percentage of each variant in the total population, as well as the percentage difference of each variant relative to non-treated negative controls.

In line with procedures for total HBB quantification, relative expression of each variant was normalized to HBA using the $2^{-\Delta(\Delta Ct)}$ analysis method, where non-edited samples served as negative controls. For sequences of primers and probes (*Metabion International AG*) and annealing/extension cycle conditions, see tables S3 and S7.

Globin chain analysis of patient-derived CD34$^+$ HSPCs by HPLC analysis

Globin expression in CD34$^+$-derived primary cells was analyzed by RP-HPLC analysis as described elsewhere.[15] Briefly, induced differentiated HSPCs were collected on day 7–8 of differentiation and were pellet at 300 x g for 5 min at 4 °C. At least 1 x 10^6 cells were processed for analysis. Cells were resuspended in HPLC water at cells densities of 0.3 x 10^6 cells/50 µL and were lysed with two rounds of freezing and thawing on ice, followed by centrifugation at 21 000 x g for 10 min at 4 °C in order to remove debris. Protein extracts were collected as supernatant and transferred to 250-µL HPLC micro-inserts or HPLC vials (Altmann Analytik, Germany) and analyzed with injection volumes of 30 µL per run on a Shimadzu Prominence system. Protein separation was performed with a linear gradient of 0.1% trifluoroacetic acid 0.033% NaOH in H_2O against 0.1% trifluoroacetic acid in acetonitrile (all *Sigma-Aldrich, UK*) on an Aeris Widepore 3.6 µm XB-C18 25 cm 4.6 mm reversed-phase column (*Phenomenex*) and for absorbance readout at 190 nm. After treatment-blinded manual correction of automatic peak detection, levels of HBB-like chains were calculated as ratios of the peak area of HBB-like globins/HBA. Changes in the expression of HBB-like chains in edited samples were analyzed as the percentage change of the corresponding ratios compared to UT controls. Control reference samples were used for the characterization of globin chain elution time (i.e. adult peripheral blood, cord blood and HbA2 (H0266, *Sigma-Aldrich*)), and cultured samples from four healthy controls served as standards for normal expression. The broad peak flanked by HBD and HBG2 in Figure 3 and supplementary Figure S4 is only observed in tissue culture samples and cannot be identified by control (or additional thalassemic blood) samples. Inverse proportion of its area with that of the β-globin peak for different treatments suggests contribution of excess α-globin chains to its formation.

Microscopy, histological staining and differential counts

Post-differentiation viability was measured based on trypan blue staining of unprocessed cells. Phenotypic characterization of HSPC-derived erythroid subpopulations in cultures was assessed by a trained hematologist in treatment-blinded manner according to criteria outlined in Wintrobe's Clinical Hematology,[21] based on cytocentrifugation of 0.5–1 x 10^4 cells on days 3 and 7 of erythroid differentiations using the Tharmac Cellspin II cytocentrifuge with an EASY rotor (*Tharmac/Hettich*, Germany; A320) and with o-Dianisidine staining (Sigma-Aldrich, UK) before standard May-Grünwald and Giemsa (all *Sigma-Aldrich*, UK) histological staining, as described previously [4]. Cells were preserved on slides under mounting medium (Entellan, *Merck*). Images were acquired using an IX73P1F inverted microscope, LED illumination, a 40x lens and averaging of seven frames per HDR image in CellSens 1.7 software (*Olympus Corporation*). Scoring of cellular differentiation stages was performed by a trained hematologist in treatment-blinded manner according to criteria outlined in Wintrobe's Clinical Hematology [21].

Flow cytometry

Flow cytometry was performed using a CyFlow Cube 8 6-channel instrument (Partec/Sysmex, Germany) and BD FACSCanto II (*BD Biosciences*, NY, USA) for the assessment of transfection and GFP-based targeted disruption efficiencies. Percentages of dead and apoptotic cells were measured by staining with SYTOX Red (*Life Technologies*, *Thermo Fisher Scientific*) and 7-AAD stains following the manufacturers' instructions. Transfected cells (5–10 x 10^5) were washed once with 1 mL PBS and recovered by centrifugation at 300 x g

for 5 min. Cell pellets were resuspended in 500 µL cold PBS, divided into two flow cytometry tubes for stained and unstained readout, and a minimum of 10^4 cellular events were recorded. Data analysis was performed with FCS Express 4 flow cytometry software (*De Novo Software*, CA, USA) and FACSDiva Software (*BD Biosciences*, NY, USA).

Prediction of off-target sites

TALENs – PROGNOS

Potential off-target sites of TALEN pairs were identified by employing the PROGNOS web tool (http://bao.rice.edu/cgi-bin/prognos/prognos.cgi) using the TALEN v2.0 algorithm on the Hg19 human genome [10]. Six maximum mismatches per half sites were set, as recommended, whereas spacer length was changed after initial employment of the default +63 C-terminus TALEN settings (10–30 nt), in order to include the 8-nt TALEN R1/L2 spacer on the *HBD* off-target locus.

RGN – MIT guide design tool

Potential off-target sites for the RGN were identified by the CRISPR design web tool (http://crispr.mit.edu/) on the Hg19 human genome [11]. The software was used to rank the potential off-target sites starting from the site with the highest score for off-target binding.

Sanger sequencing

Purified plasmids and PCR products were sequenced using the BigDye Terminator v1.1 Cycler sequencing kit (*Applied Biosystems, MA, USA*) and the reaction was performed on a Tgradient Thermocycler (*Biometra GmbH,* Germany) using the following reaction cycle conditions: 1 min at 96 °C followed by 25 cycles of, 10 sec at 96 °C, 5 sec at 50 °C, 4 min at 60 °C, and finally a hold step at 15 °C. Sequencing reaction master mix consisted of 4 µL BigDye v1.1, 4 µL 5 x BigDye buffer, 4 µL 5 x GC-RICH solution (*Roche*, Switzerland), 250 nM primer (table S3), and 50 ng and 800 ng of purified PCR product and plasmid as template, respectively, in a final reaction volume of 20 µL. DNA sequencing products were purified using Performa® DTR Gel Filtration Cartridges Performa® DTR Gel Filtration Cartridges (*Edge Biosystems*, Maryland, USA), according to manufacturer's instructions. Sequencing results were analyzed on a Hitachi 3031xl Genetic Analyzer with Sequence Detection Software version 5.2 (*Applied Biosystems*, MA, USA).

Targeted deep sequencing in genome-disrupted patient-derived HSPCs

According to prediction off-target sites for TALEN R1/L2 and RGN, we selected the respective top ten off-target sites plus the paralogous *HBB* and *HBD* sites for targeted deep sequencing. Specific primer pairs were designed for the amplification of 150–400 bp of each off-target site at optimized PCR condition (tables S8 and S9), which gave rise to a unique distinct PCR amplicon. PCR amplicons of all off-target sites for each designer nuclease were pooled together for deep sequence analysis on a MiSeq benchtop sequencer (*Illumina*) according to the manufacturer's instructions.

Statistical analysis

Statistical analyses were conducted using Prism 7.0 (GraphPad Software Inc., CA, USA). Samples were tested for normality by Shapiro-Wilk test, and group-wise comparisons performed by parametric analyses (one-way ANOVA with Dunnett's *post hoc* test) and non-parametric test (Kruskal-Wallis with Dunn's *post hoc* text, where at least one sample failed the normality test).

SUPPLEMENTARY DISCUSSION

In this study, we investigated the option of permanent NHEJ-based mutation-specific gene therapy at high efficiency and low toxicity and in the absence of selection markers, viral sequences and exogenous DNA. For proof of principle we chose the common, exon-proximal *HBB*$^{IVSI\text{-}110(G>A)}$ mutation and employed both, TALEN and RGN, platforms for NHEJ-mediated disruption of aberrant regulatory elements that are causative of missplicing and severe β-thalassemia phenotype in *HBB*$^{IVSI\text{-}110(G>A)}$ homozygotes.

NHEJ is the typical repair mechanism of DSBs in interphase and quiescent cells, and in the absence of specific conditions[22,23] is several times faster than HDR,[24] which though available in interphase[25] is considered the typical mode of repair during M phase.[26] Likely linked to this phenomenon and relating to the mostly quiescent cell state of rudimentary stem cells,[27] it appears that LT-HSPCs are particularly recalcitrant to HDR-based repair.[28–31] Adding to possible therapeutic options, genome modification without DSB induction has been established by chemical modification of the DNA-binding scaffold of designer nucleases. The resulting base editor molecules allow specific nucleotide changes in the form of transitions (i.e. changes of purines to purines and pyrimidines to pyrimidines),[32–38] and in the case of adenine base editors would thus be suitable for HDR-independent precise correction of the *HBB*$^{IVSI\text{-}110(G>A)}$ mutation. However, sequence constraints of currently available scaffolds would not allow targeting the *HBB*$^{IVSI\text{-}110(G>A)}$ for precise correction. Moreover, efficiency concerns, in part owing to base-editor size, and comparably incomplete characterization of genome-wide effect of base editors are presently still impediments to their clinical exploitation. In consequence, the exploitation of NHEJ for curative therapy of hematopoietic disorders in general and for *HBB*$^{IVSI\text{-}110(G>A)}$ in particular appears to be superior to HDR-based and base-editing-based approaches at present.

Proof of principle for therapeutic action of NHEJ in general has been established by several studies for major monogenic diseases. These include delivery of RGN pairs to reactivate γ-globin by remodeling of the β-globin locus,[39] removal of deep-intronic splice-site mutations linked to cystic fibrosis in a mini-gene splicing assay,[40] the 44-kb excision of genomic DNA containing the mutant huntingtin (*HTT*) gene,[41] and removal of the expanded trinucleotide repeat causative of myotonic dystrophy type 1.[42] Importantly, three independent studies recently employed NHEJ-based strategies to alter aberrant splicing and allow production of functional mRNAs for murine disease models. The first employed pairs of RGNs in order to recreate a functional SD site for congenital muscular dystrophy type 1A,[43] the second a single RGN to allow exon skipping by targeting a conserved splice enhancer for muscular dystrophy,[44] and the third zygote microinjection for plasmid-based TALEN delivery and mono-allelic repair of humanized *HBB*$^{IVSII\text{-}654(C>T)}$ mice.[45] Owing to the lower efficiency and a generally undesirable mixture of deletion and inversion events inherent to paired-DSB approaches, approaches using single RGNs are preferable. In this vein, removal of trinucleotide repeats has also been achieved using a single RGN,[42] and two studies based on permanent lentiviral delivery utilized single RGNs in order to knock out the *BCL11A* erythroid-specific enhancer in human HSPCs[46] and achieve a short γ-globin-inducing deletion in the γ-globin promoters,[47] respectively. None of these studies have demonstrated efficient correction when combining target cells and delivery methods suitable for therapeutic application. Moreover, presently all NHEJ-based studies in cells from β-hemoglobinopathy patients rely on reactivation of γ-globin for therapeutic effect, and it is hoped but not certain that the resulting activation of γ-globin may serve as universal cure for β-hemoglobinopathies. Conversely, mutation-specific therapies based on designer nucleases are likely the most efficient approach in suitably stratified patient populations, but to date rely on HDR.

Here we demonstrated performance of the NHEJ-based disruption approach in *HBB*$^{IVSI\text{-}110(G>A)}$ homozygote human CD34^{+} cells. We showed significant correction of key disease parameters, including HBB/HBA ratios and erythroid differentiation, at a high level of cell viability and minimal levels of *HBD* off-targeting. As a parameter specific to splice-site mutations like *HBB*$^{IVSI\text{-}110(G>A)}$, we moreover showed correction of missplicing of HBB mRNA to up to 100% in RGN- and up to 93.9% in TALEN-modified cells. Importantly, we employed the authentic substrate for clinical gene-therapy of β-thalassemia and other hematological disorders, human CD34^{+} HSPCs,[48,49] based on delivery methods compatible with clinical application. What is more, we achieved up to 95.4% on-target disruption efficiencies according to T7E1 assays and up to 88.3% on-target disruption

of alleles according to targeted deep sequencing analysis in bulk cell populations by TALEN and RGN, respectively. The latter data suggest that in the ideal case of binomial distribution of DSBs, 98% of cells would have at least one allele modified and thus normal or *HBB*$^{IVSI\text{-}110(G>A)}$ carrier status. Clinical data from allogeneic bone marrow (BM) transplantations for sickle cell disease and thalassemia indicate that 10–30% BM chimerism for healthy HSPCs is sufficient to achieve transfusion independence of patients, owing to *in vivo* selection of healthy RBCs.[50–55] NHEJ-based gene correction of *HBB*$^{IVSI\text{-}110(G>A)}$ thalassemia would therefore allow high correction efficiencies and potentially lowered conditioning requirements compared to current gene addition approaches in order to achieve transfusion independence, at lowered risk to the patient.

Of further note for clinical translation, NHEJ-mediated correction by TALEN pair R1/L2 in cells of patient A (n=3) gave consistently marginal to undetectable *HBD* off-targeting efficiency of (0.6±1.1)% (data not shown) and consistently high on-targeted *HBB* disruption efficiency of (85.7±12.8)%, whereas lower disruption frequencies were observed in cells of patients B at (58.9±10.9)% (n=2), C at 76.5% (n=1) or E at 55.6% (n=1), even though treatment was performed in parallel and under identical conditions (supplementary Figure S8). Consequently, we detected a significantly increased HBB/HBA protein ratio of 0.40±0.12 (*P*=.014) in treated bulk populations for patient A, relative to the UT control (0.19±0.08) (supplementary Figure S9) and higher than that seen across samples for different patients (Figure 3H). While correction efficiencies were thus high for all patients tested, this suggests that designer nuclease efficiencies and therefore NHEJ-based correction are nevertheless patient-dependent and, in agreement with findings for gene addition,[56] should first be evaluated in cells derived from each patient to gauge suitability before proceeding to clinical application.

Our observations inform future selection of target sites for NHEJ-based strategies. Investigation of indel patterns showed that NHEJ-based functional correction is suitable even for treatment of exon-proximal mutations. Further inferences can be drawn for NHEJ-specific target site selection and nuclease design for effective and safe functional correction. The observed distribution of indels tied in with target alignments of both nucleases on *HBB*$^{IVSI\text{-}110(G>A)}$ (Figure 1) and TIDE analyses (supplementary Figure S6). It also suggested that slightly higher correction efficiency with RGN compared to R1/L2 (Figure 3), despite lower RGN disruption efficiency, may be based on R1/L2 creating more distant deletions of context sequences, the effect of which may be size dependent. Regarding the target site selection, cleavage (here by RGN) immediately adjacent to the aSA had the highest effect on splice correction, while frequency of events and levels of correction for TALEN show that aberrant splicing can also be achieved by upstream deletions that leave the aSA intact. Targeting of context sequences by NHEJ would thus be a promising alternative strategy for functional correction in cases where the primary mutation is not a suitable target for disruption, for instance because of adjacent essential sequences or because of absence of suitable PAM sites for currently available RGNs.[35] Regarding nuclease design, the RGN employed here included the causative mutation and several other off-target mismatches in the gRNA recognition sequence as obvious design choices, while two different TALEN designs were evaluated experimentally. We observed differential *HBD* off-targeting by R1/L1 and R1/L2, which shared a common monomer but differed in their spacer length (Figure 1B and supplementary Figures S1&2). R1/L2 with its shorter 10-bp on-target spacer and a correspondingly suboptimal 8-bp spacer[57] on *HBD* had minimal *HBD* off-targeting in contrast to R1/L1 (Figure 3 and supplementary Figure S4), and was thus chosen for full functional analysis instead. Toward greater biosafety of paired designer nucleases in general, differential spacer sizes between on- and potential off-target sites can thus be exploited in order to allow effective on-target cleavage while excluding the off-target site.

We noted reduction in HBD/HBA ratios upon correction of HBB expression. The T7E1 assay and deep sequencing showed that off-targeting as contribution to this phenomenon was minimal for RGN and R1/L2. Instead, cells established an equilibrium of β-like globin expression that led to downregulation of alternative β-like globins when HBB expression was restored [9]. Importantly, three parameters inherent to the methods used here bring about an underestimation of the true level of correction at it may be expected in the human body, as has been detailed elsewhere.[4,15] First, RP-HPLC analysis does not measure insoluble α-globin cell wall aggregates that occur in β-thalassemia, so that the HPLC method alone underestimates the correction of HBB/HBA ratios achieved here. Second, our cell cultures have high basal γ-globin expression, which is in competition with upregulation of HBB and additionally leads to a higher level of effective erythropoiesis in

untreated thalassemic controls than would be expected in the peripheral blood of patients. Third, baseline HBB levels are much higher in our CD34+ differentiation cultures than they would be in the blood of *HBB*$^{IVSI-110(G>A)}$ homozygote patients, reducing the incremental effect seen by *HBB* splice correction across all analyzed parameters. Despite these three factors, bulk cultures of primary patient cells allowed the detection of statistically and biologically highly significant correction by disruption of regulatory elements at the protein and morphological level, besides exceptionally high correction at the level of mRNA splicing.

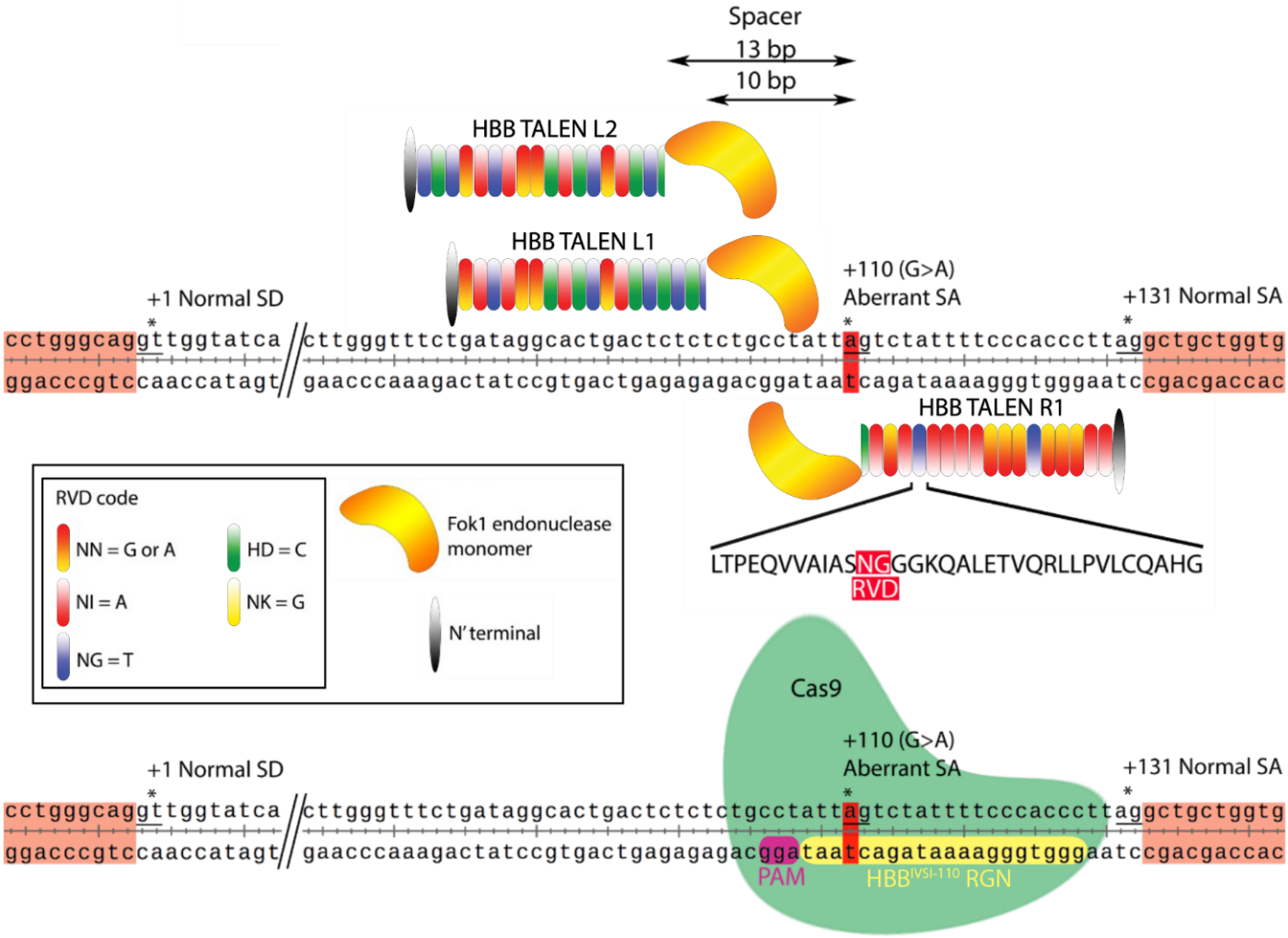
Analysis of the top ten off-target sites plus *HBD* showed a high level of safety for R1/L2 with minimal residual off-targeting of HBD. The latter may still be addressed by replacement of individual NN with NK RVDs, in order to exploit additional sequence differences between *HBB* and *HBD*, and by carefully balancing any gain in specificity against loss of on-target activity. Importantly, analysis of top ten sites for the RGN revealed off-targeting of three intronic sites, including effective intronic disruption of the large long non-coding RNA gene, *RNF219-AS1*, on chromosome 13. Even though *RNF219-AS1* may not be functional in hematopoietic lineages and has not been associated with any disease, the high level of off-targeting needs to be addressed in ordered to decrease the risks of chromosomal rearrangements after clinical translation.[58] This can be accomplished by using the CRISPR/Cpf1 system,[59] dose optimization or the use of regulated Cas9.[11,60] An improved balance of on-target vs off-target activity may also be achieved by application of high fidelity Cas9, such as SpCas9-HF1,[61] eSpCas[62] or Alt-R Cas9 HiFi,[63] truncated guide RNAs,[64] paired nickases[65] or minimal exposure of genomic DNA to the active nucleases, all of which likely also reduce on-target efficiency. A case in point is the application of normal instead of hypothermic culture conditions for *ex vivo* nuclease treatment, which concurrently reduces on- and off-target activity (supplementary Figure S10 for TALENs). For the TALEN platform specifically, residual monomer-dependent off-target sites could be avoided altogether without reduction in on-target efficiency by adopting an obligate heterodimeric architecture.[66] Regarding safety of disruption of regulatory elements beyond off-target risks, it would appear that highly transient *ex vivo* delivery will not be affected by pre-existing immunity to Cas9,[67,68] and may reduce the risk of selecting for P53-deficient cells posed by widespread DSB induction.[69,70]

Finally, analysis of long-term safety and efficiency of our approach is required in order to move the specific *HBB*$^{IVSI-110(G>A)}$ treatment forward as a therapeutic option. In this context it is doubtful whether assessment in transgenic humanized mouse models, such as the *HBB*$^{IVSI-110(G>A)}$ model developed by Vadolas *et al.*,[71] would be informative. Specific clinically relevant off-target sites would only be present in the human genome, and curative therapy will depend on modification and viability of human LT-HSPCs. Of note and while HSPCs are the authentic substrate of curative therapies for hemoglobinopathies,[48,49] this study relied on expansion of primary CD34+ cells from unmobilized peripheral blood samples.[6,72] Although this allowed the assessment of correction efficiency for primary cell material from an unprecedented number of independent HBB$^{IVSI-110(G>A)}$-homozygous individuals (n=4), the resulting cell material would not allow verification of high-level correction in LT-HSPCs, such as by transplantation into non-obese diabetic (NOD)-severe combined immunodeficient (SCID) IL2rγ^{null} (NSG) or derivative mouse models.[73] While existing studies have shown the suitability of NHEJ-mediated modification for LT-HSPCs, experimental confirmation of those observations by *in vivo* studies in NSG mice will be another critical step towards clinical translation of our approach for *HBB*$^{IVSI-110(G>A)}$ β-thalassemia. Combined with clonogenic assays this will allow long-term assessment of the level of correction, engraftment and multi-lineage potential of modified human CD34+ cells.

Disruption of aberrant regulatory elements as a therapeutic approach would be widely applicable. It is conceptually safer than LV-based gene addition, with its inherent risk of insertional mutagenesis and often suboptimal gene expression at moderate VCNs.[48] Also, disruption of aberrant regulatory elements is more efficient than gene correction approaches based on HDR, which may require means of enrichment after modification for clinical application.[74,75] It is suitable for the removal of any gain-of-function regulatory element at several nucleotides distance from open reading frames or other conserved elements. As demonstrated in this study, one typical application of the approach would be to disrupt aberrant splice acceptor or donor sites or inadvertently activated cryptic splice sites. The number of known disease-causing mutations of this type is rapidly increasing, and literature searches (supplementary Figure S11) and existing mutation databases dedicated to splice-site mutations readily reveal that there are already mutations suitable for this approach in over 180 genes responsible for many human diseases,[76,77] in which the same

approach could therefore have essential therapeutic effects. Besides its immediate relevance for *HBB*$^{IVSI-110(G>A)}$ therapy development, disruption of aberrant regulatory elements therefore represents a mutation-specific, effective gene therapy approach with potential for clinical application for a range of diseases.

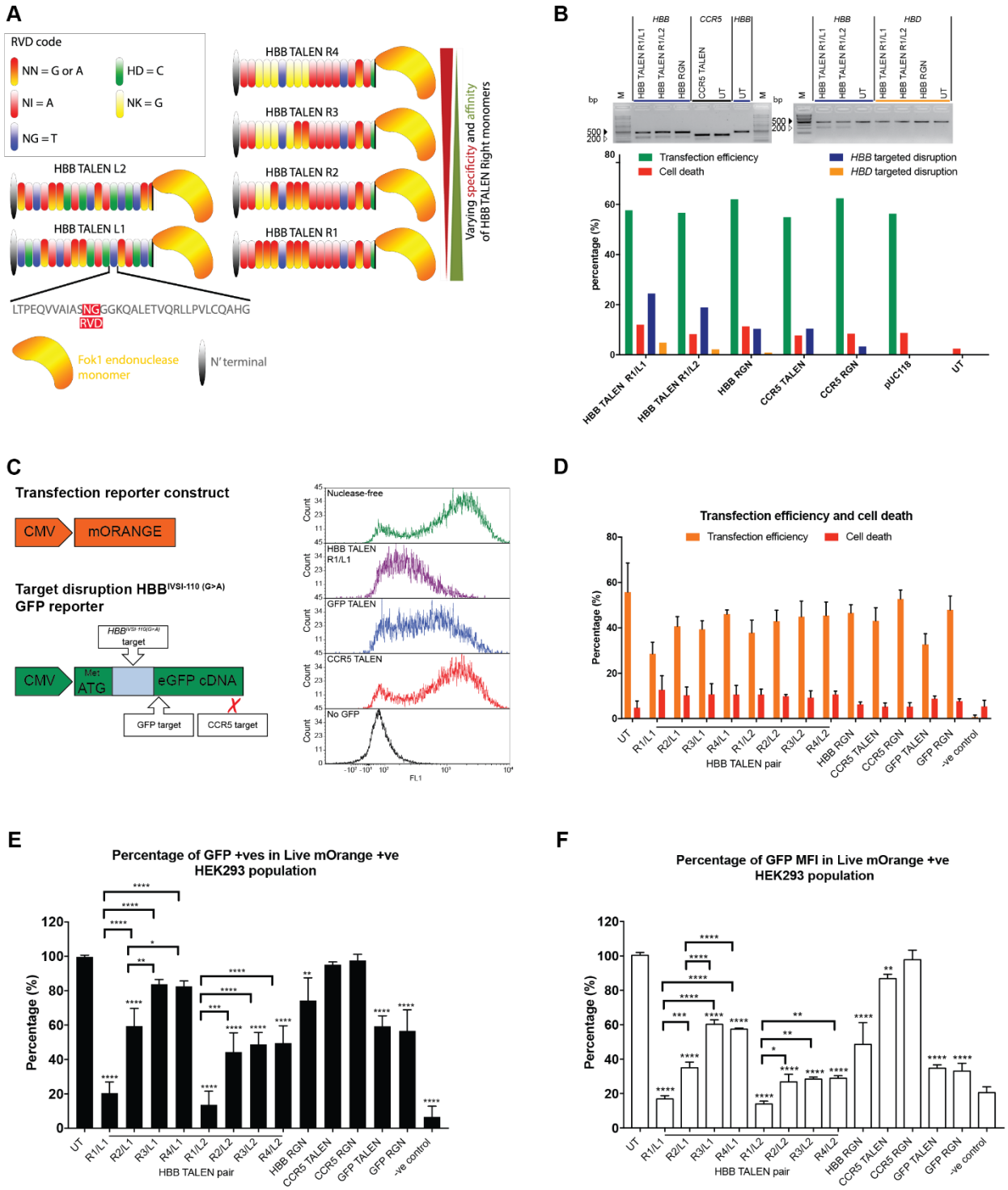
SUPPLEMENTARY FIGURES



Supplementary Figure S1. Basic TALEN and RGN nuclease design. TALEN pairs (top) and CRISPR/Cas9 RGN (bottom) are shown relative to the *HBB*$^{IVSI\text{-}110(G>A)}$ mutation (red highlight) within exon 1 of HBB (unshaded sequence area, with flanking exons as orange boxes).

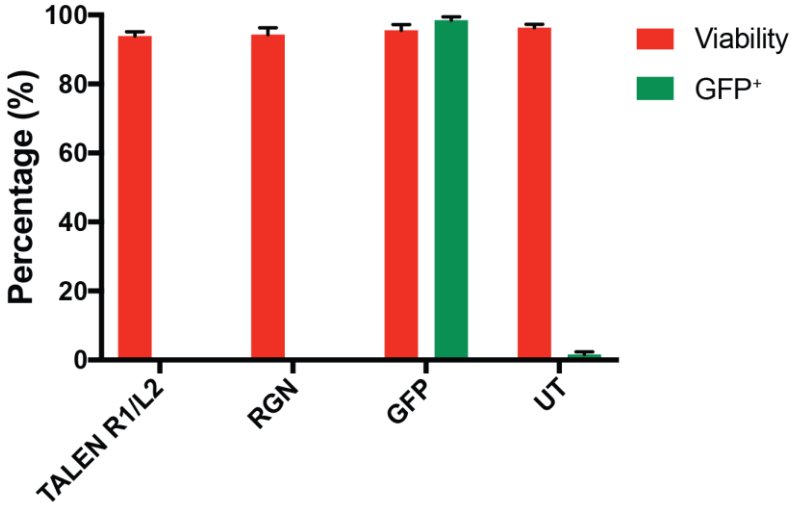
Three TALEN monomers (L1, L2 and R1) were used as differentially spaced active dimers, R1/L1 (13-bp spacer) and R1/L2 (10-bp spacer), to induce DSBs upstream of the *HBB*$^{IVSI\text{-}110(G>A)}$ mutation. Coloring of TALEN repeat elements indicates the specific RVDs used, as labelled in the inset.

The RGN guide RNA binding sequence (yellow highlight) encompasses the mutation close to its protospacer-adjacent motif (PAM, purple highlight), creating DSBs immediately adjacent to the +110 aSA.

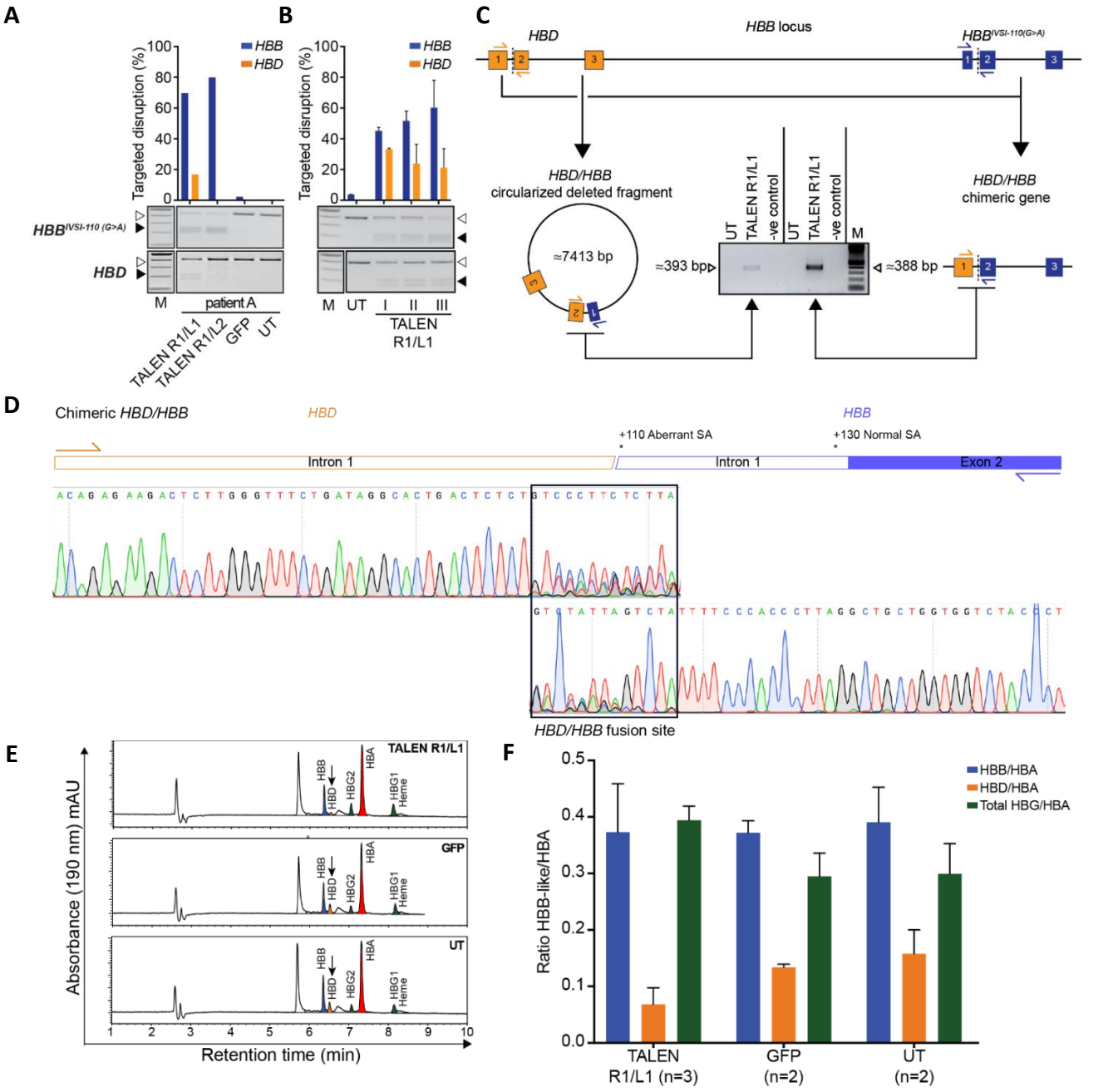


Supplementary Figure S2. Disruption efficiencies of all TALEN pair combinations and RGN in HEK293 cells. (**A**) Schematic illustration of all HBB-specific TALEN monomers employed in this study. Left 1 (L1), Left 2 (L2) and Right 1 – 4 (R1 – R4) monomers. Right monomers were modified by replacing NN (G or A) modules (red to yellow gradient) with the more specific but less efficient NK (G) module (yellow). The number of substitutions of up to 6 NN to NK modules in the modified versions of the Right TALEN monomer was 0 for R1, 2 for R2, 4 for R3 and 6 for R4. Each TALEN monomer consists of 17.5 TALE repeats, and the conserved amino acid sequence per TALE repeat is indicated with the RVD (12th and 13th amino acid) highlighted in red. The yellow freeform shape is the FokI endonuclease cleavage domain and the grey ellipse the TALE N-terminus. (**B**) (**Top**) Targeted disruption efficiency of designer nucleases on HEK 293T genomic DNA by T7E1 assay. Left panel: agarose gel electrophoresis analysis of PCR products after treatment with T7E1.

Lanes 1, 2, 3 and 6 are HBB PCR products derived from transfected HEK 293T cells with HBB TALEN R1/L1, HBB R1/L2, HBB RGN and pUC118, whereas lanes 4 and 5 are CCR5 PCR products from cells transfected with CCR5 TALEN and RGN, respectively. Right panel: targeted disruption efficiency of HBB TALEN R1/L1 (lane 1) and R1/L2 (lane 2) on the HBB locus relative to the negative pUC118-transfected cells (lane 3). Lanes 4, 5, 6 and 7 show the parallel off-target potential of both HBB TALEN pairs, R1/L1 and R1/L2, and RGN, relative to the nuclease-free negative control, pUC118, on the highly sequence-similar *HBD,* respectively. (**Bottom**) Genome editing in HEK 293 cells. Analysis of transient transfection efficiency (mCherry-positive, green bars) and cell death (trypan blue positives, red bars) of HEK 293T cells with designer nucleases at 48 h post-transfection. Dark blue bars: on-target disruption efficiencies. Orange bars: off-target potential of HBB specific designer nucleases TALEN R1/L1 and R1/L2 and HBB RGN on *HBB* and *HBD* loci, respectively. CCR5-specific TALEN and RGN are included as positive controls of the method and accordingly on-targeted disruption efficiencies on CCR5 are shown in red bars. pUC118 is used as nuclease-free negative control. (**C**) Schematic representations of the transfection reporter construct pCMV mOrange and targeted disruption of the HBB$^{IVSI-110}$GFP reporter. The HBB$^{IVSI-110}$GFP reporter was prepared by inserting the HBB$^{IVSI-110}$target sequence between the start codon (ATG) and the eGFP cDNA, so that the construct is targeted by *HBB$^{IVSI-110(G>A)}$*-specific designer nucleases and allows assessment of targeted disruption efficiencies as reduction of the percentage of GFP positive cells or mean fluorescence intensities in the mOrange live (SYTOX Red-negative) cell population (see histograms). GFP- and CCR5-specific designer nucleases were used as designer-nuclease positive and negative controls, pUC118 as nuclease-free negative control. (**D**) Transfection efficiencies measured as mOrange positive cells (orange bars) and cell death as SYTOX Red-positive cells (red bars) measured via flow cytometry. Experiments were conducted in triplicate by transient transfection of HEK 293 cells for the assessment of targeted disruption of the HBB$^{IVSI-110}$-GFP reporter gene construct. Cells were transfected with all HBB TALEN combination (R1–4/L1 and R1–4/L2), the HBB RGN, CCR5 TALEN and RGN, and GFP TALEN and RGN, with CCR5- and GFP-specific nucleases as designer-nuclease-negative and -positive controls, respectively. All cells were co-transfected with equal amounts of mOrange construct for the assessment of transfection efficiency. HBB$^{IVSI-110}$GFP reporter/mOrange sample served the nuclease-free negative control. (**E**) and (**F**) Assessment of targeted disruption efficiency of designer nucleases on the episomal HBB$^{IVSI-110}$GFP reporter via flow cytometry. Experiments were conducted in triplicate by transient transfection of HEK 293 cells for the assessment of targeted disruption of the HBB$^{IVSI-110}$GFP reporter gene construct. Cells were transfected with all HBB TALEN combinations (R1–4/L1 and R1–4/L2), the HBB RGN, CCR5 TALEN and RGN, and GFP TALEN and RGN, with CCR5- and GFP-specific nucleases as designer-nuclease-negative and -positive controls, respectively. (**E**) Targeted disruption efficiency of the HBB$^{IVSI-110}$GFP reporter gene construct quantified as a reduction of the percentage of GFP positive cells in the live (SYTOX Red negative) and mOrange positive transiently transfected HEK 293T cell population compared to the nuclease-free negative control (HBB$^{IVSI-110}$GFP reporter). Data obtained from duplicates in three independent experiments. (**F**) Targeted disruption efficiency was quantified as a reduction of the percentage of GFP MFI in live (SYTOX Red-negative) and mOrange-positive HEK 293T cells compared with the nuclease-free negative control (HBB$^{IVSI-110}$ GFP reporter). Data obtained from duplicates in three independent experiments. Statistical differences in the percentage of GFP-positive cells and GFP MFI for all samples was tested compared with the HBB$^{IVSI-110}$GFP reporter/mOrange negative control. Moreover, group-wise comparisons were performed for HBB TALEN combinations with R1–4/L1 and R1–4/L2, separately. Statistical analysis was performed by one-way ANOVA. * $P < .05$, ** $P < 0.01$*** $P < 0.001$, **** $P < 0.0001$. UT – untransfected control receiving no plasmids; -ve control – negative control receiving mOrange and eGFP reporter plasmids but no nuclease-encoding plasmids

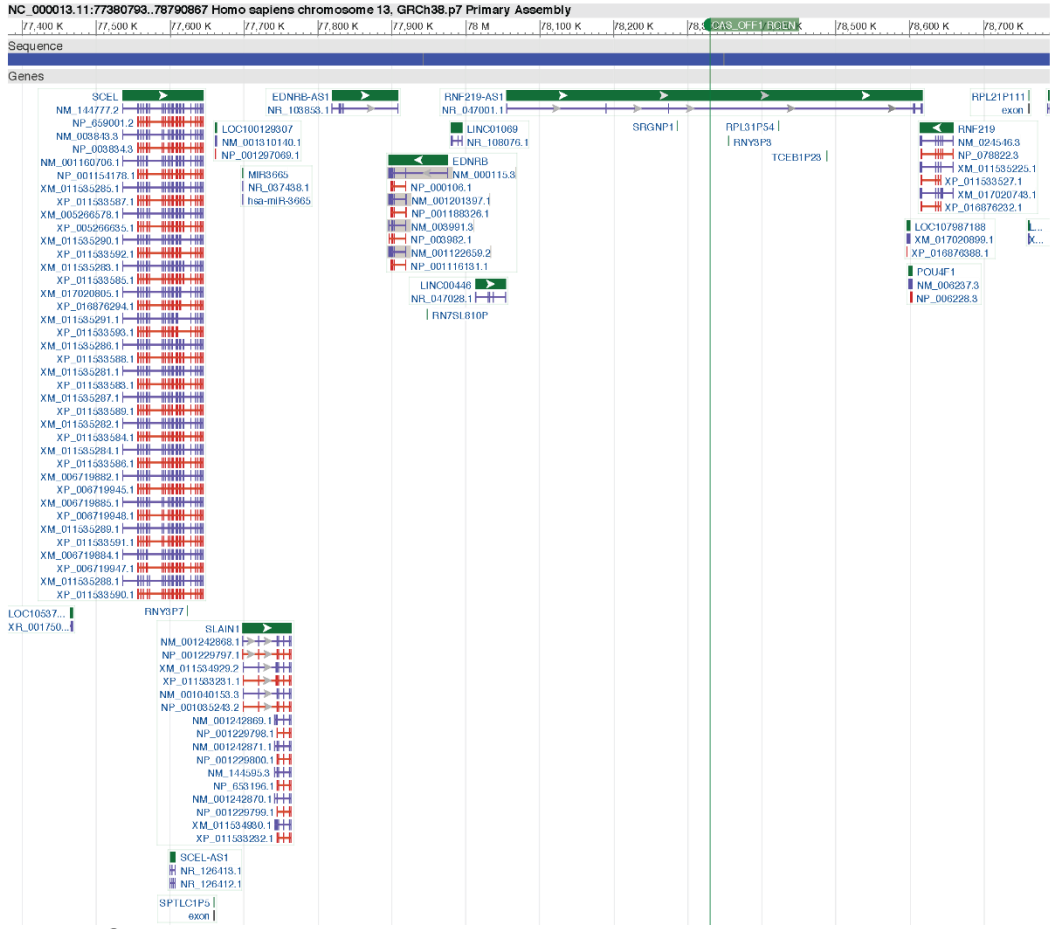


Supplementary Figure S3. Average percentage of cell viability and transfection efficiency in CD34⁺ cells. Analysis 24 h after DNA-free nucleofection of *HBB*$^{IVSI\text{-}110(G>A)}$-homozygous CD34⁺ cells with TALEN R1/L2 (mRNA), RGN (RNP) and GFP (mRNA) and untreated negative control under optimized conditions. Cell viability (red) of all samples was measured after P18 staining (cell death stain) and Nucleoview Count analysis. Average percentages of GFP positives (green) in samples nucleofected with GFP mRNA were analyzed by flow cytometry for the assessment of transfection efficiency. All displayed data comprised the average values of biological triplicates (n = 3; ±SD).

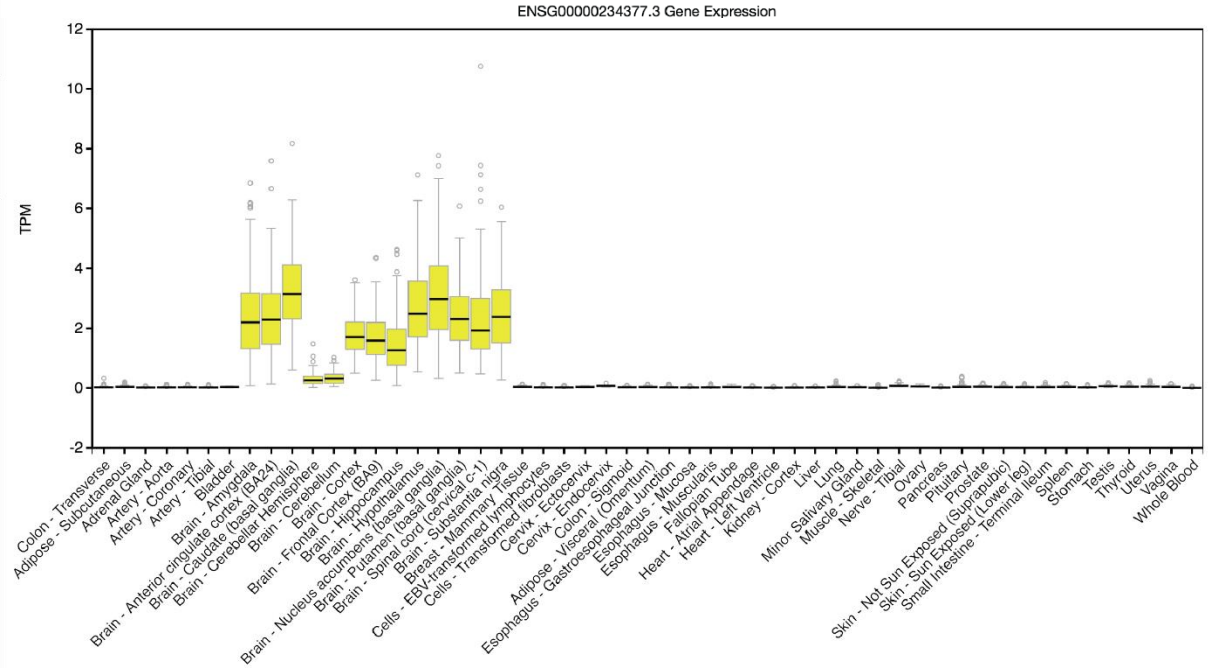


Supplementary Figure S4. *HBB* and *HBD* disruption by TALENs in *HBB*$^{IVSI\text{-}110(G>A)}$-homozygous HSPCs. R1/L1- and R1/L2-mediated modification of the *HBB* locus by mRNA-based delivery in patient-derived HSPCs, including GFP and untransfected (UT) negative controls. Analyses in (B)–(F) are based on same-patient biological triplicates I, II and III for patient D. For illustration, the single row of gel bands representing fragments for *HBB* and *HBD* has been separated in two and the size marker placed adjacent to lanes of interest (A&B) and duplicated (B). (**A**) T7E1-based assessment of targeted disruption of *HBB*$^{IVSI\text{-}110(G>A)}$ (blue bars) and *HBD* (orange bars) in cells of patient A for all four treatments, as indicated. (**B**) T7E1-based assessment of targeted disruption of *HBB*$^{IVSI\text{-}110(G>A)}$ (blue bars) and *HBD* (orange bars) in replicate samples of patient D (n=3) after R1/L1 and UT treatment and showing one representative assay. (**C**) Schematic illustration of *HBB* and *HBD*, indicating predicted L1/R1-induced DSBs (dashed lines), site-specific primer pairs and fusion products resulting from concurrent cleavage. Conventional PCRs with hybrid primer pairs detected a circularized epigenomic *HBB-HBD* fusion product (≈398 bp) and a novel genomic *HBD-HBB* fusion product (≈388 bp), respectively, as shown for replicate II. (**D**) Sanger sequencing of the chimeric *HBD/HBB* amplified PCR product with *HBD*- (orange arrow) and *HBB*- (blue arrow) specific primers. The black box indicates the *HBD/HBB* fusion site, which is identified as mixed sequencing traces. (**E**) Representative RP-HPLC-based detection of human globin chains in patient-derived HSPC cultures on day 7 of induced erythroid differentiation after R1/L1 and control treatments. (**F**) Quantification of mean (±SD) HBB-like/HBA globin chain ratios as determined in (E) across experiments (n=3).

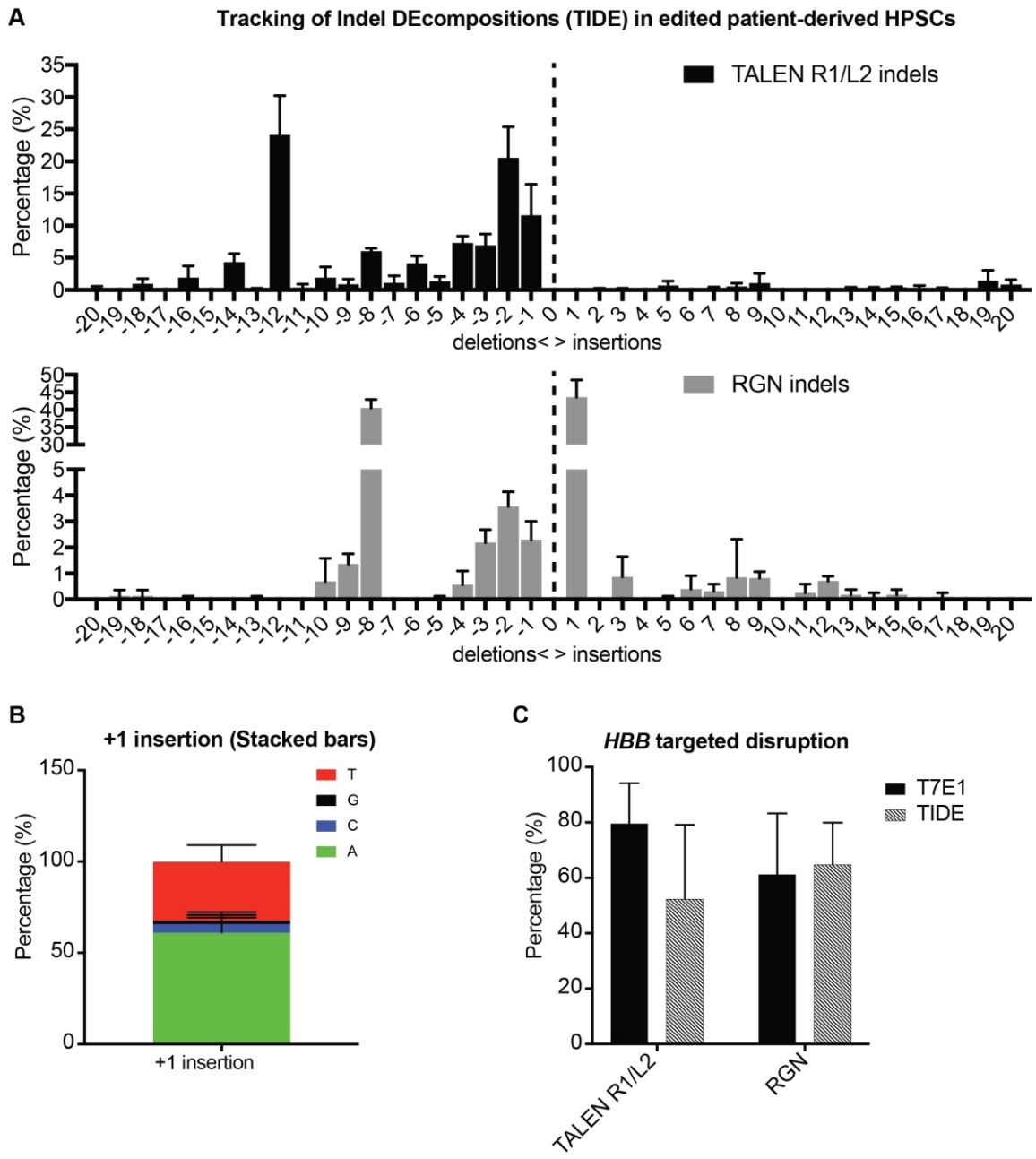
A **Cas_OFF1 target site (RNF219_AS1)**



B **RNF219_AS1 tissue expression**

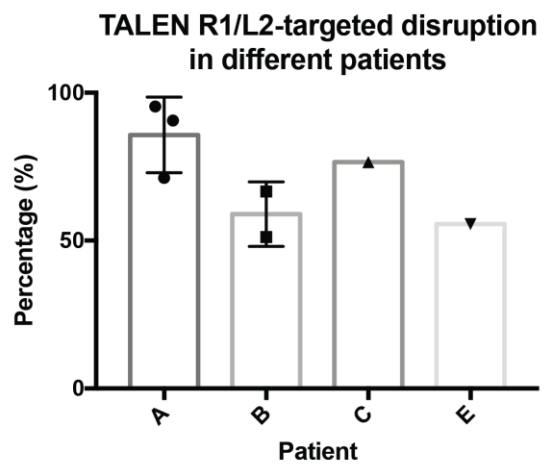


Supplementary Figure S5. Top RGN off-target site CAS_OFF1_RGN (*RNF219-AS1*). (**A**) Ensembl analysis of the top RGN off-target site (CAS_OFF1_RGN) on the human genome (GRCh38.P7). CAS_OFF1 lies in intron 3 of the *RNF219-AS1* (HGNC:42700) (long non-coding RNA) gene (size: 697 640 bp), between *SRGNP1* and *RNY3P3* pseudogenes, which is expressed as at least 14 splice transcript variants (Ensembl)/19 splice transcript variants (LNCpedia), mainly in brain. (**B**) *RNF219-AS1* gene expression in human tissues (www.gtexportal.org). TPM; Transcripts per million.

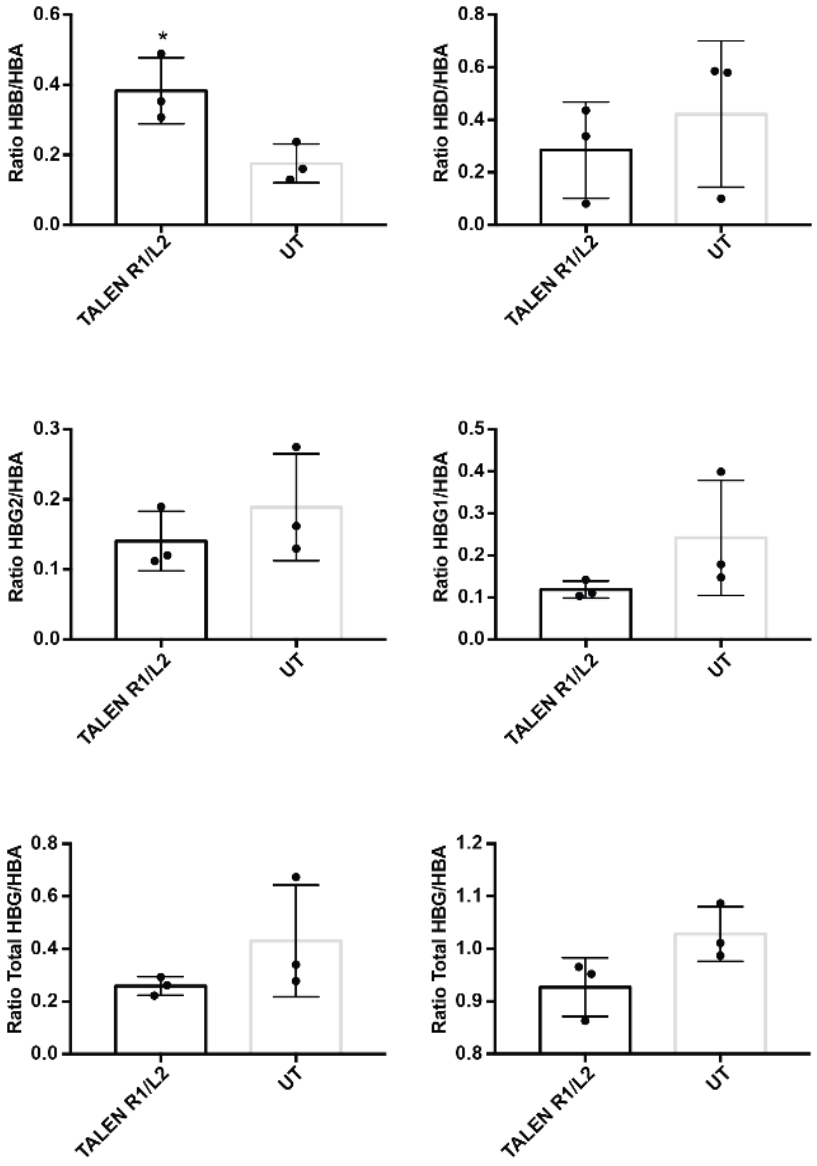


Supplementary Figure S6. Indels produced in TALEN R1/L2- and RGN-disrupted *HBB*$^{IVSI\text{-}110(G>A)}$-homozygous HSPCs. (**A**) Average percentages of different types of indels based on their type (insertion or deletion) and size in the total number of modifications after editing with TALEN R1/L2 (black bars) and RGN (grey bars) in samples from three different patients (n=3). Data derived from TIDE analysis of Sanger sequencing traces of edited bulk HSPC population using the untransfected control (UT) as reference. (**B**) Average percentages of the type of nucleotide (red: thymidine, black: guanine, blue: cytosine, green: adenosine) detected for single-base insertions at the DSB site in RGN-edited patient-derived HSPCs based on TIDE analysis. (**C**) Average percentages of *HBB* targeted disruption in TALEN R1/L2- and RGN-edited HSPCs from three different patients as measured by T7E1 assay (solid bars) and TIDE (striped bars). All displayed data comprised the average values of biological triplicates (n=3; ±SD). *Of note, TIDE was conceived for RGN analyses, which might have brought about that TIDE and T7E1 analyses match better for the RGN than for TALEN R1/L2.*

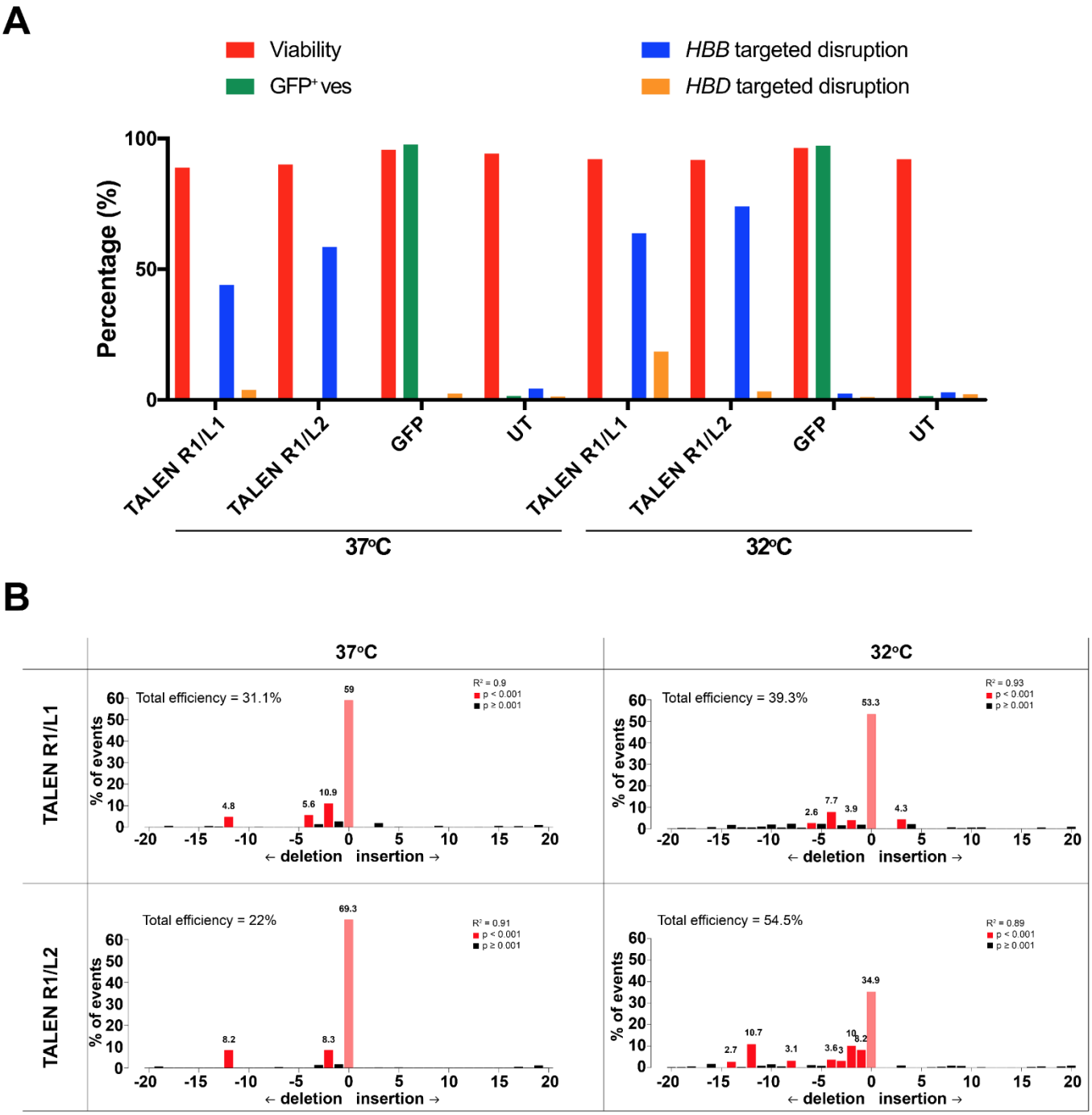
Supplementary Figure S7. Changes in splicing factor binding motifs in the most frequent CD34⁺ indel events. The exon-2-proximal region of *HBB* intron 1 is shown for splicing factor analyses using the Human Splice Finder (HSF; http://www.umd.be/HSF3/) online tool for the most frequent disruption events observed (12-bp deletion, 2-bp deletion, 3-bp deletion for TALEN R1/L2, and 1-bp insertion (A) or (T), 8-bp deletion (downstream) and 8-bp deletion (upstream) for the RGN), as detailed below. The SA (HSF relative motif strength: 83.58) remains unaffected for all events shown. As general points of reference, for the normal *HBB* locus the aSA motif has a strength of 48.42, whereas this is an elevated 77.36 for *HBB*$^{IVSI\text{-}110(G>A)}$. For TALEN R1/L2, 96.7% of indels were deletions of various length, the most frequent deletion being one of 12 bp (13.2% of all deletions), which removes the aSA altogether. Two additional deletions above 5% frequency were upstream deletions of 2 bp (6.1% of all deletions) and 3 bp (5.6%), respectively, both slightly increasing the aSA motif strength but reducing the balance of enhancing/silencing motifs. For the RGN, the indel pattern was more balanced between insertions (42.7%) and deletions (57.3%) (Fig 6D and 6F). The majority were 1-bp insertions, with the frequent insertion of adenine (91.8%) reducing the aSA motif strength almost to the level of the normal locus (53.56), and the insertion of thymidine (7.6%) reducing the aSA motif strength to half that of the normal locus (24.61). The most frequent RGN-induced deletions were two 8-bp deletions, the first removing the aSA and six nucleotides downstream (53.18% of all deletions) and the second removing the aSA and six nucleotides upstream (11.89% of all deletions). Changes of motifs compared with normal (*HBB*) (removal, creation or shift in prediction confidence) are indicated by rounded rectangles. Symbols above the x-axis refer to enhancing motifs, those below the x-axis to silencing motifs.



Supplementary Figure S8. Genome disruption efficiency appears patient-specific. Average percentages (±SD) of TALEN R1/L2-targeted disruption efficiency on the *HBB*$^{IVSI\text{-}110(G>A)}$ site in HSPCs derived from different patients (A, B, C and E), as measured by T7E1.

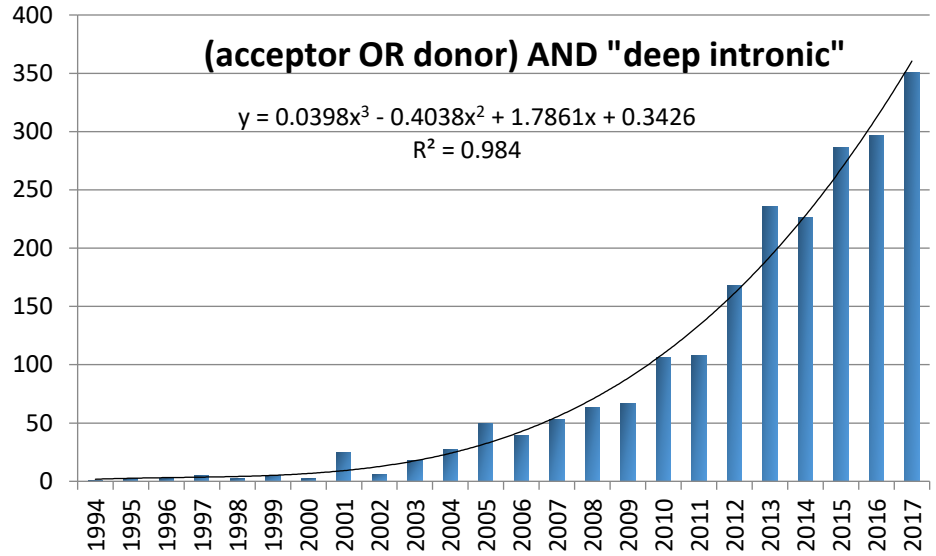


Supplementary Figure S9. NHEJ-based correction of HBB protein expression for patient A. HBB-like/HBA globin chain ratios from a triplicate gene-disruption experiment with TALEN R1/L2 in HSPCs derived from a single patient (patient A) on day 7 of induced differentiation. Statistically significant differences of HBB-like/HBA globin chain ratios were measured relative to the untransfected control (UT) by paired t-test. HBB/HBA * *P* =.0115. Displayed data are average values of biological triplicates (n = 3; ±SD).

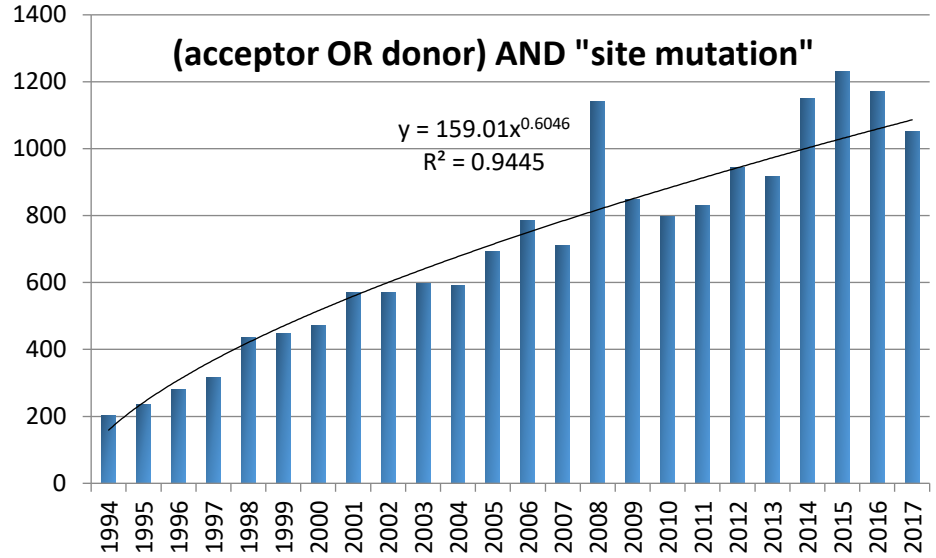


Supplementary Figure S10. Effect of post-nucleofection culture temperature on TALEN-mediated indel creation. (**A**) Percentages of cell viability (red), transfection (green) and targeted disruption efficiencies of *HBB* (blue) and *HBD* (orange) relative to the untransfected negative control (UT), post-nucleofection with TALEN (R1/L1 or R1/L2) and GFP mRNAs without (37 °C) and with incubation at hypothermic condition (32 °C). All samples were analyzed in parallel 48 h post-nucleofection by NucleoCounter NC-250 (*Chemotech*, Denmark), flow cytometry and T7EI assay. (**B**) TIDE analysis depicting the overall $HBB^{IVSI\text{-}110(G>A)}$-targeted disruption efficiencies of TALEN (R1/L1 or R1/L2) relative to nuclease-free negative controls. Same-size induced indels (-20 deletions to +20 insertions) are scored as a percentage of the total number of events. Significance cutoff was the TIDE default ($P < .001$).

A



B



Supplementary Figure S11. Long-term trends for splicing-related publications. The search strings indicated were used in Google Scholar (3 May 2018) without manual curation to gauge (**A**) likely publication of mutations suitable for disruption of aberrant regulatory elements against (**B**) likely publications of mutations in splice consensus sites. Publications for the latter are increasing at a diminishing rate, whereas with the advent of massively parallel sequencing a presently cubic growth has begun of publications for splicing-related deep intronic mutations as potential targets for disruption of aberrant regulatory elements. Formulae indicate the curve fit of the trend line, R^2 indicates the goodness of the fit.

SUPPLEMENTARY TABLES

Target name	Nuclease type	Score	Targeted gene (closest known gene)	Genomic region	Chromosome
HBB	CRISPR	38.50	HBB[1]	Intron	chr11
HBD	CRISPR	N/A	HBD[2]	Intron	chr11
CAS_OFF1	CRISPR	5.38	RNF219-AS1	Intron	chr13
CAS_OFF2	CRISPR	1.03	DGKK	Intron	chrX
CAS_OFF4	CRISPR	0.93	CDC42BPB	Intron	chr14
CAS_OFF5	CRISPR	0.92	Intergenic	Intergenic	chr15
CAS_OFF6	CRISPR	0.86	Intergenic	Intergenic	chr6
CAS_OFF7	CRISPR	0.63	GUCY1A2	Intron	chr11
CAS_OFF9	CRISPR	0.56	Intergenic	Intergenic	chr11
CAS_OFF8	CRISPR	0.56	Intergenic	Intergenic	chr2
CAS_OFF3	CRISPR	0.54	CDK8	Intron	chr13
CAS_OFF10	CRISPR	0.54	Intergenic	Intergenic	chr8
HBB	TALEN	100.00	HBB[1]	Intron	chr11
HBD	TALEN	82.35	HBD	Intron	chr11
TALEN_OFF1	TALEN	64.03	(SLC10A6)	Intergenic	chr4
TALEN_OFF2	TALEN	62.73	UQCC	Intron	chr20
TALEN_OFF3	TALEN	62.59	CNBD1	Intron	chr8
TALEN_OFF4	TALEN	62.53	(LSP1P3)	Intergenic	chr5
TALEN_OFF5	TALEN	62.47	(MAP3K)	Intergenic	chr6
TALEN_OFF6	TALEN	61.76	(C1D)	Intergenic	chr2
TALEN_OFF7	TALEN	61.51	KIAA1217	Intron	chr10
TALEN_OFF8	TALEN	61.31	TMEM64	Intron	chr8
TALEN_OFF9	TALEN	61.18	(LPHN3)	Intergenic	chr4
TALEN_OFF10	TALEN	61.12	OXR1	Intron	chr8

Supplementary Table S1. List of the top ten *in silico* predicted off-target sites. The table shows top ten off-targets for the RGN and for TALEN R1/L2 and corresponding target scores as calculated by the CRISPR design online tool and PROGNOS, respectively. Targets shown were selected for targeted deep sequencing.

[1] *HBB* refers only to the normal *HBB* allele. [2] *HBD* was included for deep sequencing analysis although it had not been predicted by the CRISPR design tool as an off-target site for the RGN.

Monomer name	RVD sequence	Binding sequence
TALEN R1	NI NI NN NN NN NG NN NN NN NI NI NI NI NG NI NN NI HD	TAAGGGTGGGAAAATAGAC
TALEN R2	NI NI NK NK NN NG NN NN NN NI NI NI NI NG NI NN NI HD	TAAGGGTGGGAAAATAGAC
TALEN R3	NI NI NK NK NK NG NK NN NN NI NI NI NI NG NI NN NI HD	TAAGGGTGGGAAAATAGAC
TALEN R4	NI NI NK NK NK NG NK NK NK NI NI NI NI NG NI NN NI HD	TAAGGGTGGGAAAATAGAC
TALEN L1	NG HD NG NN NI NG NI NN NN HD NI HD NG NN NI HD NG HD	TTCTGATAGGCACTGACTC
TALEN L2	NN NI NG NI NN NN HD NI HD NG NN NI HD NG HD NG HD NG	TGATAGGCACTGACTCTCT
RM98 CCR5 TAL L	NG NG NN NG NN NN NN HD NI NI HD NI NG NN HD NG NN NN	TTTGTGGGCAACATGCTGG
RM101 CCR5 TAL R	HD NI NN HD HD NG NG NG NG NN HD NI NN NG NG NG NI NG	TCAGCCTTTTGCAGTTTAT
GFP TAL Right	NN NG NN NN NG HD NN NN NN NN NG NI NN HD NN NN HD NG	TGTGGTCGGGGTAGCGGCT
GFP TAL Left	NN NI HD HD NI HD HD HD NG NN NI HD HD NG NI HD NN NK	TGACCACCCTGACCTACGG

Supplementary Table S2. Sequences of TALEN monomers. RVD sequences are given from the N-terminus to the C-terminus, and binding sequences 5' to 3' on the target gene. The [T] at position 0 is excluded from the RVD the sequence.

Primer name	Application	Sequence
Seq TAL FW	Sequencing	GCCGTGGAAGCCGTGC
Seq TAL RV	Sequencing	TCAGGGCGGCCAGAGC
M13 FW	Sequencing	GTAAAACGACGGCCAG
109 RV	Sequencing/T7E1 Assay	CCCTTCCTATGACATGAACTTAACCAT
CMV FW	Sequencing	GAGACTTGGAAATCCCCGTGA
GCB1 FW	T7E1 Assay	TTCACTAGCAACCTCAAACAGACACC
HBB EX2.2 RV	T7E1 Assay	CAGCTCACTCAGTGTGGCAA
HBD EX2.1 RV	T7E1 Assay	GCAGCTCACTCAGCTGAGAA
CCR5 FW	T7E1 Assay	AAGATGGATTATCAAGTGTCAAGTCC
CCR5 RV	T7E1 Assay	CAAAGTCCCACTGGGCG
hHBB_EX1_FW_3	RT-qPCR	GGGCAAGGTGAACGTG
hHBB_EX2_RV_1	RT-qPCR	GGACAGATCCCCAAAGGAC
wtHBB_Probe_ZNA	RT-qPCR	6-FAM-TGGG(PDC)AGG(PDC)TG(PDC)TG-ZNA-3-BHQ-1
IVSI-110_MGB_Probe	RT-qPCR	VIC-TAAGGGTGGGAAAATAGA-MGB
hHBB_EX1_RV_2_A	RT-qPCR	CACCACCAACTTCATCCAC
hHBB_EX1_FW_1	RT-qPCR	GGTGCATCTGACTCCTGAG
hHBB_FW_EX2_B	RT-qPCR	GGCAAGAAAGTGCTCGG
hHBB_EX2.3_RV_B	RT-qPCR	GTGCAGCTCACTCAGTG
hHBA_FW	RT-qPCR	GGTCAACTTCAAGCTCCTAAGC
hHBA_RV	RT-qPCR	GCTCACAGAAGCCAGGAACTTG

Supplementary Table S3. PCR primers and probes. Sequences of primers and probes categorized based on their applications (Sanger sequencing, PCR amplification for T7E1 assays, PCR amplification for sequencing, RT-qPCR amplification for RNA expression analysis), as indicated. All sequences are given 5' to 3'.

Oligonucleotide name	Application	Sequence
HBB-rep-FW	GFP reporter production	taaTCTTCTGATAGGCACTGACTCTCTCTGCCTATTAGTCTATTTTC CCACCCTTAgga
HBB-rep-RV	GFP reporter production	ccggtccTAAGGGTGGGAAAATAGACTAATAGGCAGAGAGAGTCA GTGCCTATCAGAAGAttaat
HBB-CR-FW	gRNA production	acaccGGGTGGGAAAATAGACTAATG
HBB-CR-RV	gRNA production	aaaaCATTAGTCTATTTTCCCACCCG

Supplementary Table S4. Oligonucleotides for oligonucleotide annealing. Upper cases sequence of interest (insert), lower case the extension for formation of the suitable overhangs post–oligo-annealing. All sequences are given 5' to 3'.

gRNA name	gRNA sequence (PAM)	Orientation on target gene
HBB RGN	GGGTGGGAAAATAGACT**A**AT (AGG)	- strand
CCR5 RGN	GTGAGTAGAGCGGAGGCAGG (AGG)	- strand

Supplementary Table S5. RGN guide RNA sequences. The bold underlined nucleotide is the site of the *HBB*$^{IVSI\text{-}110\ (G>A)}$ mutation on the HBB RGN guide RNA sequence and in parenthesis is the 3' NGG PAM sequence on each target site. All sequences are given 5' to 3'.

Primer pair	Target locus	Cell type	Annealing temperature (°C)	Amplicon size (bp)
GCB1 FW / HBB EX2.2 RV	*HBB*$^{IVSI\text{-}110}$	HEK 293T	68	432
GCB1 FW / HBD EX2.1 RV	*HBD*	HEK 293T	68	431
CCR5 FW/RV	*CCR5*	HEK 293T	60	292
HBB_CD34_FW/RV	*HBB*$^{IVSI\text{-}110}$	Human CD34^{+}	66	389
HBD_CD34_FW/RV	*HBD*	Human CD34^{+}	66	392

Supplementary Table S6. Primer pairs for T7E1 assays in detail. The primers shown were employed in order to produce target-site-specific PCR products by Phusion/Q5 PCR amplification for quantification of disruption efficiencies by T7EI assay. All sequences are given 5' to 3'.

Method	Primer pair	Transcript	Annealing temperature (°C)
SybrGreen	hHBB_EX1_FW_1 / hHBB_EX1_RV_2_A	total human HBB on Exon 1	65
	hHBB_FW_EX2_B hHBB_EX2.3_RV_B	total human HBB on Exon 2	65
Qiagen Multiplex RT-qPCR	hHBB_EX1_FW_3 / hHBB_EX2_RV_1 wtHBB_Probe_ZNA IVSI-110_MGB_Probe	normal and aberrantly spliced HBB variants	60

Supplementary Table S7. Primer and probe combinations and annealing temperatures for RT-qPCR.

Labeled	Oligo Name	Sequence	Annealing temperature (°C)	Amplicon size (bp)	Cleavage site
HBB	HBB_CD34_FW	TGAGGAGAAGTCTGCCGTTAC	66	389	180
	HBB_CD34_RV	CAGCTCACTCAGTGTGGC			
HBD	HBD_CD34_FW	TGAGGAGAAGACTGCTGTCAA	66	392	180
	HBD_CD34_RV	CAGTGCAGCTCACTCAGCT			
TALEN_OFF1	SLC10A6_FW	CAACCACTTAGGCATAGCTCAGG	65	235	78
	SLC10A6_RV	CCTGAGGAGTATCTTTCACCCC			
TALEN_OFF2	UQCC_FW	GTCTCGGTGTTCCTCTACCC	65	215	145
	UQCC_RV	TACGGCAGAGACTTCCTACTCAAC			
TALEN_OFF3	CNBD1_FW	CAGTGCTCTGAGAGAGATGAGAC	65	159	82
	CNBD1_RV	CCCCAGCACAGCACTATGTGA			
TALEN_OFF4	LSP1P3_FW	GGAAACTGGAGCAAGGACAG	68	200	106
	LSP1P3_RV	ATGCAGAACTGTGAGCCAAG			
TALEN_OFF5	MAP3K7_FW	GCTCTTCTTCGCATCATG	68	205	100
	MAP3K7_RV	CTGCACTGCTATTACCAA			
TALEN_OFF6	C1D_FW	GCTCCTCTACATCTCCAAAGGAAG	65	175	104
	C1D_RV	GGACTGGAGTGGAATATTGTAGGG			
TALEN_OFF7	KIAA1217_FW	ATGGCAGGTGGTGGCCAACT	65	185	103
	KIAA1217_RV	CTGCTCTCCTTCTAGTTTCCTGTC			
TALEN_OFF8	TMEM64_FW	AGCATTCACCACTTATTCCTTCTG	65	201	68
	TMEM64_RV	ACAGTATTAGAGCTCCAAATAAGC			
TALEN_OFF9	LPHN3_FW	GGCAAGTGGTGATAAGTGGATCAG	65	244	128
	LPHN3_RV	GCAACCCACCTTGCCAAACTTTC			
TALEN_OFF10	OXR1_FW	GTCCCAGTGCACTTCATTGTGTTC	65	178	71
	OXR1_RV	GGTAAACAGCTGGGAGCTCAATC			

Supplementary Table S8. Primers used for targeted deep sequencing of potential TALEN off-targets. The primers shown were employed for amplification of the top 10 *in silico* predicted off-target site for TALEN R1/L2 plus the *HBB* and *HBD*. PCR cycle conditions: 95 °C, 3 min; 95°C for 20 sec, optimized annealing temperatures for each primer pair (as shown) for 20 sec and 72 °C for 40 sec, repeat 35 times; and 72 °C for 7 min. All sequences are given 5' to 3'.

Labeled	Oligo Name	Sequence	Annealing Temp (°C)	Amplicon size (bp)	Cleavage site
HBB	HBB_CD34_FW	TGAGGAGAAGTCTGCCGTTAC	66	389	184
	HBB_CD34_RV	CAGCTCACTCAGTGTGGC			
HBD	HBD_CD34_FW	TGAGGAGAAGACTGCTGTCAA	66	392	180
	HBD_CD34_RV	CAGTGCAGCTCACTCAGCT			
	RNF219-AS1_FW	CTGCCCTAGACAGAGGAA	60	156	74
CAS_OFF1	RNF219-AS1_RV	ATGTCCTATGCCCTTTCC			
	DGKK_FW	TAGATATTTTCTGGTTCAAGGACAG	60	205	123
CAS_OFF2	DGKK_RV	TTTGACCCAAAAGGGCAA			
	CDC42BPB_FW	CCCATATGTGGAATGCTAAT	58	163	84
CAS_OFF3	CDC42BPB_RV	GCTCCTGTTCCTTTCAAG			
	chr15_FW	GGTCATTTATGCCACGTG	57	208	113
CAS_OFF4	chr15_RV	GGCACTGAAAAGCATAAG			
	chr6_FW	GTCAACAAGGGATATTTATG	53	200	114
CAS_OFF5	chr6_RV	TTTGAAGTAAGAAAGCAATA			
	GUCY1A2_FW	CCAACAGGGGATAATAGAC	55	220	88
CAS_OFF6	GUCY1A2_RV	TTGATCATGCCTTTTGCA			
	chr11_FW	CCATTGTTAGAGGTTTCACGTTATT	63	269	170
CAS_OFF7	chr11_RV	TGATCAAGGGTTGTGGGTATAAG			
	chr2_FW	CCTCCACATTGCTGGTAGAA	63	267	88
CAS_OFF8	chr2_RV	GGACATGCAGAAAGGAAGAATG			
	CDK8_FW	TAATTTGAAGAGAATGGAGC	57	196	127
CAS_OFF9	CDK8_RV	CCATGTCAGCTGTAAAATAA			
	chr8_FW	TGGCTTAAATCACACAACAGAAC	63	284	119
CAS_OFF10	chr8_RV	GACAAGAGCTAGACTTCATCTCAA			

Supplementary Table S9. Primers used for targeted deep sequencing of potential RGN off-targets.
The primers shown were employed for amplification of the top ten *in silico* predicted off-target sites plus the *HBB* and *HBD* for RGN. PCR cycle conditions: 95 C, 3 min; 95°C for 20 sec, optimized annealing temperatures for each primer pair (as shown) for 20 sec and 72°C for 40 sec, repeat 35 times; and 72 °C for 7 min. All sequences are given 5' to 3'.

SUPPLEMENTARY REFERENCES

1. Kountouris P, Lederer CW, Fanis P, et al. IthaGenes: An interactive database for haemoglobin variations and epidemiology. *PLoS One*. 2014;9(7):e103020 (10 pages).
2. Liu M, Rehman S, Tang X, et al. Methodologies for Improving HDR Efficiency. *Front. Genet.* 2019;9(January):691.
3. Chang K-H, Smith SE, Sullivan T, 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.
4. Patsali P, Papasavva P, Stephanou C, et al. Short-hairpin RNA against aberrant HBBIVSI-110(G>A) mRNA restores β-globin levels in a novel cell model and acts as mono- and combination therapy for β-thalassemia in primary hematopoietic stem cells. *Haematologica*. 2018;103(9):e419–e423.
5. Miller JC, Zhang L, Xia DF, et al. Improved specificity of TALE-based genome editing using an expanded RVD repertoire. *Nat. Methods*. 2015;12(5):465–71.
6. Stephanou C, Papasavva P, Zachariou M, et al. Suitability of small diagnostic peripheral-blood samples for cell-therapy studies. *Cytotherapy*. 2017;19(2):311–326.
7. Zhang Z, Chen Y, Sun X, et al. CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. *Protein Cell*. 2015;6(5):363–372.
8. Cencic R, Miura H, Malina A, et al. Protospacer adjacent motif (PAM)-distal sequences engage CRISPR Cas9 DNA target cleavage. *PLoS One*. 2014;9(10):1–13.
9. Thein SL. Molecular basis of β thalassemia and potential therapeutic targets. *Blood Cells, Mol. Dis.* 2017;(June):0–1.
10. Fine EJ, Cradick TJ, Zhao CL, Lin Y, Bao G. An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage. *Nucleic Acids Res.* 2014;42(6):e42.
11. Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 2013;31(9):827–832.
12. Zerbino DR, Achuthan P, Akanni W, et al. Ensembl 2018. *Nucleic Acids Res.* 2018;46(D1):D754–D761.
13. Aguet F, Brown AA, Castel SE, et al. Genetic effects on gene expression across human tissues. *Nature*. 2017;550(7675):204–213.
14. Desmet F-O, Hamroun D, Lalande M, et al. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009;37(9):e67–e67.
15. Loucari CC, Patsali P, van Dijk TB, et al. Rapid and Sensitive Assessment of Globin Chains for Gene and Cell Therapy of Hemoglobinopathies. *Hum. Gene Ther. Methods*. 2018;29(1):60–74.
16. Mussolino C, Alzubi J, Fine EJ, et al. TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res.* 2014;42(10):6762–73.
17. Streubel J, Blücher C, Landgraf A, Boch J. TAL effector RVD specificities and efficiencies. *Nat. Biotechnol.* 2012;30(7):593–595.
18. Dreyer A-K, Hoffmann D, Lachmann N, et al. TALEN-mediated functional correction of X-linked chronic granulomatous disease in patient-derived induced pluripotent stem cells. *Biomaterials*. 2015;69:191–200.
19. Brinkman EK, Chen T, Amendola M, Steensel B Van, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* 2014;42(22):e168.
20. Christodoulou I, Patsali P, Stephanou C, et al. Measurement of lentiviral vector titre and copy number by cross-species duplex quantitative PCR. *Gene Ther.* 2016;23(1):113–118.
21. Dessypris EN, Sawyer ST. Chapter 6: Erythropoiesis. *Wintrobe's Clin. Hematol. 12th Ed.* 2009;

22. Vartak S V., Raghavan SC. Inhibition of nonhomologous end joining to increase the specificity of CRISPR/Cas9 genome editing. *FEBS J.* 2015;282(22):4289–4294.
23. Maruyama T, Dougan SK, Truttmann MC, et al. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnolgy*. 2015;
24. Mao Z, Bozzella M, Seluanov A, Gorbunova V. Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair (Amst).* 2008;7(10):1765–71.
25. Zaboikin M, Zaboikina T, Freter C, Srinivasakumar N. Non-Homologous End Joining and Homology Directed DNA Repair Frequency of Double-Stranded Breaks Introduced by Genome Editing Reagents. *PLoS One*. 2017;12(1):e0169931.
26. Mohrin M, Bourke E, Alexander D, et al. Hematopoietic Stem Cell Quiescence Promotes Error-Prone DNA Repair and Mutagenesis. *Cell Stem Cell*. 2010;7(2):174–85.
27. Höfer T, Rodewald H-R. Output without input: the lifelong productivity of hematopoietic stem cells. *Curr. Opin. Cell Biol.* 2016;43:69–77.
28. Wang J, Exline CM, DeClercq JJ, et al. Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat. Biotechnol.* 2015;33(12):1256–1263.
29. Genovese P, Schiroli G, Escobar G, et al. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature*. 2014;510(7504):235–240.
30. Hoban MD, Cost GJ, Mendel MC, et al. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood*. 2015;125(17):2597–2604.
31. DeWitt MA, Magis W, Bray NL, et al. Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci. Transl. Med.* 2016;8(360):.
32. Komor AC, Kim YB, Packer MS, Zuris JA, David R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016;533(7603):420–424.
33. Kim YB, Komor AC, Levy JM, et al. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* 2017;35(4):371–376.
34. Gaudelli NM, Komor AC, Rees HA, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. 2017;551(7681):464–471.
35. Komor AC, Badran AH, Liu DR. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell*. 2017;168:20–36.
36. Rees HA, Komor AC, Yeh W-H, et al. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat. Commun.* 2017;8:15790.
37. Liu L, Li X, Ma J, et al. The Molecular Architecture for RNA-Guided RNA Cleavage by Cas13a. *Cell*. 2017;170(4):714–726.e10.
38. Li X, Wang Y, Liu Y, et al. Base editing with a Cpf1–cytidine deaminase fusion. *Nat. Biotechnol.* 2018;36(4):324–327.
39. Antoniani CA, Meneghini V, Lattanzi A, et al. Induction of Fetal Hemoglobin Synthesis By Crispr/Cas9-Mediated Disruption of the β-Globin Locus Architecture. *Am. Soc. Hematol. 58th Annu. Meet. Exhib. Abstr. #321, Oral Present. 4/12/2016.* 2016;128(22):.
40. Sanz DJ, Hollywood JA, Scallan MF, Harrison PT. Cas9/gRNA targeted excision of cystic fibrosis-causing deep-intronic splicing mutations restores normal splicing of CFTR mRNA. *PLoS One*. 2017;12(9):1–13.
41. Shin JW, Kim K-H, Chao MJ, et al. Permanent inactivation of Huntington's disease mutation by personalized allele-specific CRISPR/Cas9. *Hum. Mol. Genet.* 2016;25(20):ddw286.
42. van Agtmaal EL, André LM, Willemse M, et al. CRISPR/Cas9-Induced (CTG·CAG)n Repeat Instability in the Myotonic Dystrophy Type 1 Locus: Implications for Therapeutic Genome Editing. *Mol. Ther.* 2017;25(1):24–43.
43. Kemaladewi DU, Maino E, Hyatt E, et al. Correction of a splicing defect in a mouse model of congenital muscular dystrophy type 1A using a homology-directed-repair-independent

mechanism. *Nat. Med.* 2017;23(8):984.
44. Amoasii L, Long C, Li H, et al. Single-cut genome editing restores dystrophin expression in a new mouse model of muscular dystrophy. *Sci. Transl. Med.* 2017;9(418):eaan8081.
45. Fang Y, Cheng Y, Lu D, et al. Treatment of β ^{654} -thalassaemia by TALENs in a mouse model. *Cell Prolif.* 2018;(May):e12491.
46. Canver MC, Smith EC, Sher F, et al. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature*. 2015;527(7577):192–7.
47. Traxler EA, Yao Y, Wang YD, 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):987–990.
48. Finotti A, Breda L, Lederer CW, et al. Recent trends in the gene therapy of β-thalassemia. *J. Blood Med.* 2015;6:69–85.
49. Negre O, Eggimann A-V, Beuzard Y et al, et al. Gene Therapy of the β-Hemoglobinopathies by Lentiviral Transfer of the β(A(T87Q))-Globin Gene. *Hum. Gene Ther.* 2016;27(2):148–65.
50. Krishnamurti L, Kharbanda S, Biernacki MA, et al. Stable Long-Term Donor Engraftment following Reduced-Intensity Hematopoietic Cell Transplantation for Sickle Cell Disease. *Biol. Blood Marrow Transplant.* 2008;14(11):1270–1278.
51. Walters MC, Patience M, Leisenring W, et al. Stable mixed hematopoietic chimerism after bone marrow transplantation for sickle cell anemia. *Biol. Blood Marrow Transplant.* 2001;7(12):665–673.
52. Wu CJ, Gladwin M, Tisdale J, et al. Mixed haematopoietic chimerism for sickle cell disease prevents intravascular haemolysis [1]. *Br. J. Haematol.* 2007;139(3):504–507.
53. Mabaera R, West RJ, Conine SJ, et al. A cell stress signaling model of fetal hemoglobin induction: what doesn't kill red blood cells may make them stronger. *Exp. Hematol.* 2008;36(9):1057–1072.
54. Rouleau E, Lefol C, Bourdon V, et al. Quantitative PCR high-resolution melting (qPCR-HRM) curve analysis, a new approach to simultaneously screen point mutations and large rearrangements: Application to MLH1 germline mutations in lynch syndrome. *Hum. Mutat.* 2009;30(6):867–875.
55. Andreani M, Testi M, Gaziev J, et al. Quantitatively different red cell/nucleated cell chimerism in patients with long-term, persistent hematopoietic mixed chimerism after bone marrow transplantation for thalassemia major or sickle cell disease. *Haematologica*. 2011;96(1):128–133.
56. Breda L, Casu C, Gardenghi S, et al. Therapeutic hemoglobin levels after gene transfer in beta-thalassemia mice and in hematopoietic cells of beta-thalassemia and sickle cells disease patients. *PLoS One*. 2012;7(3):e32345.
57. Mussolino C, Morbitzer R, Lütge F, et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 2011;39(21):9283–9293.
58. Weinstock DM, Richardson CA, Elliott B, Jasin M. Modeling oncogenic translocations: Distinct roles for double-strand break repair pathways in translocation formation in mammalian cells. *DNA Repair (Amst).* 2006;5(9–10):1065–1074.
59. Kim D, Kim J, Hur JK, et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* 2016;34(8):863–868.
60. Guha TK, Wai A, Hausner G. Programmable Genome Editing Tools and their Regulation for Efficient Genome Engineering. *Comput. Struct. Biotechnol. J.* 2017;15:146–160.
61. Kleinstiver BP, Pattanayak V, Prew MS, et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 2016;529(7587):490–495.
62. Slaymaker IM, Gao L, Zetsche B, et al. Rationally engineered Cas9 nucleases with improved specificity. *Science*. 2016;351(6268):84–8.
63. Seki A, Rutz S. Optimized RNP transfection for highly efficient CRISPR/Cas9-mediated gene

knockout in primary T cells. *J. Exp. Med.* 2018;jem.20171626.
64. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 2014;32(3):279–284.
65. Ran FA, Hsu PD, Lin C-Y, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013;154(6):1380–9.
66. Doyon Y, Vo TD, Mendel MC, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods*. 2011;8(1):74–79.
67. Chew WL. Immunity to CRISPR Cas9 and Cas12a therapeutics. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 2018;10(1):1–23.
68. Charlesworth CT, Deshpande PS, Dever DP, et al. Identification of Pre-Existing Adaptive Immunity to Cas9 Proteins in Humans. *bioRxiv*. 2018;243345.
69. Ihry RJ, Worringer KA, Salick MR, et al. P53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* 2018;24(7):939–946.
70. Haapaniemi E, Botla S, Persson J, Schmierer B, Taipale J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* 2018;24(7):927–930.
71. Vadolas J, Nefedov M, Wardan H, et al. Humanized β-thalassemia mouse model containing the common IVSI-110 splicing mutation. *J. Biol. Chem.* 2006;281(11):7399–7405.
72. Cosenza LC, Breda L, Breveglieri G, et al. A validated cellular biobank for β-thalassemia. *J. Transl. Med.* 2016;14(1):255.
73. Karponi G, Psatha N, Lederer CWCW, et al. Plerixafor+G-CSF-mobilized CD34+ cells represent an optimal graft source for thalassemia gene therapy. *Blood*. 2015;126(5):616–619.
74. Lederer CW, Kleanthous M. Beta testing : preclinical genome editing in β-globin disorders. *Cell Gen Ther Insights*. 2015;1(2):231–242.
75. Dever DP, Bak RO, Reinisch A, et al. CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature*. 2016;539(7629):384–389.
76. Scotti MM, Swanson MS. RNA mis-splicing in disease. *Nat. Rev. Genet.* 2016;17(1):19–32.
77. Buratti E, Chivers M, Hwang G, Vorechovsky I. DBASS3 and DBASS5: databases of aberrant 3'- and 5'-splice sites. *Nucleic Acids Res.* 2011;39(Database issue):D86-91.