Multiplex base editing of *BCL11A* regulatory elements to treat sickle cell disease

Fontana L.[1], Martinucci P.[1], Amistadi S.[1], Felix T.[1], Mombled M.[2,3], Tachtsidi A.[2,3], Corre G.[2,3], Chalumeau A.[1], Hardouin G.[1], Martin J.[1], Romano O. [4,**], Amendola M.[2,3**], Antoniou P.[1*], Miccio A[1*].

[1]Université Paris Cité, Imagine Institute, Laboratory of chromatin and gene regulation during development, INSERM UMR 1163, 75015 Paris, France.

[2]Genethon, 91000 Evry, France.

[3]Université Paris-Saclay, Univ Evry, Inserm, Genethon, Integrare research unit UMR_S951, 91000 Evry, France.

[4] University of Padova, Department of Molecular Medicine, Padova, Italy

[*]Corresponding authors.

**equally contributed

Abstract

Sickle cell disease (SCD) is a genetic anemia caused by the production of an abnormal adult hemoglobin. The clinical severity is lessened by elevated fetal hemoglobin (HbF) production in adulthood. A promising therapy is the transplantation of autologous, hematopoietic stem/progenitor cells (HSPCs) treated with CRISPR/Cas9 to downregulate the HbF repressor BCL11A via generation of double strand breaks (DSBs) in the +58-kb erythroid-specific enhancer. Here, to further enhance HbF production without increasing the mutagenic load, we targeted both +58-kb and +55-kb *BCL11A* erythroid-specific enhancers using base editors. We systematically dissected DNA motifs recognized by the key transcriptional activators within these regions and identified the critical nucleotides required for activator binding. Multiplex base editing of these residues was efficient and safe and generated no or little DSBs and genomic rearrangements. We observed substantial HbF reactivation, exceeding the levels achieved using the CRISPR/Cas9 nuclease-based strategy, thus efficiently rescuing the sickling phenotype. Multiplex base editing was efficient in long-term repopulating HSPCs and resulted in potent HbF reactivation *in vivo*. In summary, these results show that multiplex base editing of *BCL11A* erythroid-specific enhancers is a safe and potent strategy for treating sickle cell disease.

Introduction

SCD is a highly prevalent recessive disorder caused by a single point mutation in the β-globin (*HBB*) gene, which leads to an amino acid substitution (Glu to Val) at position 6 of the β-globin chain. This sickle β-globin variant (β^s) combines with α-chains to form sickle hemoglobin (HbS), which tends to polymerize in low oxygen conditions, causing red blood cells (RBCs) to adopt a sickle shape and lose flexibility. Clinical manifestations are due to the short lifespan of sickle RBCs (leading to anemia) and to the obstruction of small blood vessels causing multi-organ damage Therefore, patients with SCD have a poor quality of life and reduced life expectancy (Kato et al. 2018; Fernandes 2017)(Kato *et al.,* 2018, Nat Rev Dis Primers, Fernandes *et al.,* 2017, Life Sci). Beta-thalassemia is another β-hemoglobinopathy caused by mutations that reduce the production of the adult β-globin, causing severe anemia. Patients with β-hemoglobinopathies who lack a compatible donor for allogeneic HSPC transplantation, can benefit from gene therapy approaches based on the transplantation of autologous, genetically modified HSPCs (L. Li e Mandal 2024).

The severity of SCD and β-thalassemia is alleviated by the production of γ-globin chains, which compose the HbF. Gamma-globin exerts a critical anti-sickling effect and competes with β^s-globin for incorporation into the hemoglobin tetramer, thereby reducing the formation of HbS (Steinberg et al., 2014, Blood). Thus, transplantation of autologous HSPCs genetically modified to re-express HbF in their erythroid progeny is a treatment option for patients with β-hemoglobinopathies. Currently, different approaches aiming to reactivate HbF have been developed. CRISPR/Cas9 nuclease has been used to efficiently disrupt binding sites (BS) for transcriptional repressors in the γ-globin genes (*HBG1/HBG2)* promoters leading to restoration of HbF expression and rescue of the sickle phenotype (Traxler et al. 2016; Métais et al. 2019; Weber et al. 2020). Furthermore, CRISPR-Cas9 nuclease-mediated down-regulation of *BCL11A,* a major γ-globin repressor (Y. Wu et al. 2019), was recently approved as therapy for the treatment of β-hemoglobinopathies (Frangoul et al. 2024; Locatelli et al. 2024). This approach targets the GATA1 activator BS within the +58-kb *BCL11A* erythroid-specific enhancer to reduce *BCL11A* expression exclusively in erythroid cells, thereby reactivating HbF production. Alternatively, targeting the ATF4 activator BS within the +55-kb *BCL11A* erythroid-specific enhancer represents another strategy to induce HbF expression in adult cells (P. Huang et al. 2020)

However, the clinical study targeting the +58-kb enhancer showed variability in the extent of HbF reactivation among individuals, relatively low levels of Hb with HbS still accounting for a large proportion of the total Hb and modest correction of ineffective erythropoiesis (Frangoul et al. 2024; Locatelli et al. 2024). Therefore, we hypothesize that simultaneous editing of the +58-kb and +55-kb enhancers could maximize *BCL11A* downregulation, resulting in higher and more consistent HbF levels. To avoid DSB-induced toxicity and the generation of large genomic rearrangements associated with the use of CRISPR/Cas9 nuclease (Cullot et al. 2019; Fiumara et al. 2023), we used base editing approaches to precisely target the GATA1 and ATF4 activator BS in the +58-kb and +55-kb *BCL11A* enhancers. In particular, we exploited cytosine BE (CBEs), adenine BE (ABEs), and dual BE (DBEs) to dissect the GATA1 and ATF4 BSs in SCD HSPCs and identify the critical base conversions that induce changes in enhancers' activity, *BCL11A* downregulation, and consecutively, HbF reactivation.

Materials and Methods

HSPC purification and culture

We obtained adult human non-mobilized CD34+ HSPCs from SCD patients and fetal human CD34+ from feta liver of a healthy donor (HD). Adult SCD and HD fetal liver samples eligible for research purposes were obtained from the "Hôpital Necker-Enfants malades" Hospital (Paris, France). Written informed consent was obtained from all adult subjects. All experiments were performed in accordance with the Declaration of Helsinki. The study was approved by the regional investigational review board (reference: DC 2022-5364, CPP Ile-de-France II "Hôpital Necker-Enfants malades"). HSPCs were purified by immunomagnetic selection with manual magnetic cell separator (Miltenyi Biotec) after immunostaining with the CD34 MicroBead Kit (Miltenyi Biotec). Forty-eight hours before transfection, CD34+ cells were thawed and cultured at a concentration of 5x105 cells/ml in the "HSPC medium" containing StemSpan (STEMCELL Technologies) supplemented with penicillin/streptomycin (Gibco), 250 nM StemRegenin1 (STEMCELL Technologies), and the following recombinant human cytokines (PeproTech): human stem cell factor (SCF) (300 ng/ml), Flt-3L (300 ng/ml), thrombopoietin (TPO) (100 ng/ml), and interleukin-3 (IL-3) (60 ng/ml).

Base editor-expressing plasmids

Plasmids used in this study include: pCMV_ABEmax_P2A_GFP (Addgene #112101), pCMV_AncBE4max_P2A_GFP (Addgene #112100), NG-ABEmax (Addgene #124163), ABE8e (Addgene #138489), SpCas9 TadDE (Addgene #193837), CBE-SpRY-OPT1 (Antoniou et al. 2022), ABE-SpRY-OPT and SpRY-ABE8e.
A DNA fragment (3'UTR + poly-A) containing two copies of the 3' untranslated region (UTR) of the *HBB* gene and a poly-A sequence of 96 adenines were purchased by Genscript. ABE-SpRY-OPT plasmid was created by inserting the 3'UTR+poly-A fragment in the pCMV-T7-SpRY-P2A-EGFP (RTW4830) (Addgene #139989) plasmid. SpRY-ABE8e plasmid was created by replacing the Cas9 sequence of the ABE8e plasmid with the SpRY-P2A-EGFP from the pCMV-T7-ABEmax(7.10)-SpRY-P2A-EGFP (RTW5025) (Addgene #140003) plasmid. Plasmids are available upon request.

sgRNA design and production

We manually designed single guide RNAs (sgRNAs) targeting the +58-kb and +55-kb regions of *BCL11A* (**Table 1**). We used chemically modified synthetic gRNAs harboring 2′-O-methyl analogs and 3′-phosphorothioate non-hydrolyzable linkages at the first three 5′ and 3′ nucleotides (Synthego).

Table 1. sgRNA target sequences.

sgRNA	Target sequence (5'to 3')	Position (hg38)	Strand
GATA_bs_1	TTTATCACAGGCTCCAGGAA	Chr2: 60495252-60495271	-
GATA_bs_2	GATAAAAGCAACTGTTAGCT	Chr2: 60495266-60495285	+
GATA_bs_3	TGATAAAAGCAACTGTTAGC	Chr2: 60495265-60495284	+
GATA_bs_4	CTAACAGTTGCTTTTATCAC	Chr2: 60495264-60495283	-
ATF4_bs_1	CATTGCATCATCCTGGTACC	Chr2: 60498482-60498501	+
ATF4_bs_2	GCATCATCCTGGTACCAGGA	Chr2: 60498486-60498505	+
ATF4_bs_3	TTGCATCATCCTGGTACCAG	Chr2: 60498484-60498503	+
ATF4_bs_4	CTCCAAGCATTGCATCATCC	Chr2: 60498469-60498488	-

mRNA in vitro transcription

10 µg of BE-expressing plasmids were digested overnight with 20 Units of a restriction enzyme that cleaves once after the poly-A tail. The linearized plasmids were purified using a PCR purification kit (QIAGEN) and were eluted in 30 µl of DNase/RNase-free water. 1 µg of linearized plasmid was used as a template for the *in vitro* transcription (IVT) reaction (MEGAscript, Ambion). The IVT protocol was modified as follows. The GTP nucleotide solution was used at a final concentration of 3.0 mM instead of 7.5 mM and the anti-reverse cap analog N7-Methyl-3'-O-Methyl-Guanosine-5'-Triphosphate-5'-Guanosine (ARCA, Trilink) was used at a final concentration of 12.0 mM resulting in a final ratio of Cap:GTP of 4:1 that allows efficient capping of the mRNA. The incubation time for the IVT reaction was reduced to 30 minutes. mRNA was precipitated using lithium chloride and resuspended in TE buffer in a final volume that allowed to achieve a concentration of >1 µg/µl. The mRNA quality was evaluated using TapeStation 2200 (Agilent).

RNA transfection

$1x10^5$ to $2x10^5$ or $3x10^6$ $CD34^+$ HSPCs per condition were transfected with 3.0 µg or 15.0 µg of the enzyme encoding mRNA, respectively, and a synthetic sgRNA at a final concentration of 2.3 µM. We used the P3 Primary Cell 4D-Nucleofector X Kit S or L (Lonza) and the CA137 program (Nucleofector 4D). Untransfected cells or cells transfected with TE buffer or with the enzyme-encoding mRNA only, or with the enzyme-encoding mRNA and a sgRNA targeting the *AAVS1* locus, served as negative controls.

Ribonucleoprotein (RNP) transfection

RNP complexes were assembled at room temperature using a 90 µM Cas9-GFP protein and a 180 µM synthetic sgRNA (ratio Cas9:sgRNA of 1:2). $CD34^+$ HSPCs ($2x10^5$ cells/condition) were transfected with RNP complexes using the P3 Primary Cell 4D-Nucleofector X Kit S (Lonza) and the CA137 program (Nucleofector 4D) in the presence of a transfection enhancer (IDT). Untransfected cells or cells transfected with TE buffer served as negative controls.

HSPC differentiation

Transfected $CD34^+$ HSPCs were differentiated into mature RBCs using a three-phase erythroid differentiation protocol, as previously described (Giarratana et al. 2005; Weber et al. 2020). During the first phase (day 0 to day 6), cells were cultured in a basal erythroid medium supplemented with 100 ng/ml recombinant human SCF (PeproTech), 5 ng/ml recombinant human IL-3 (PeproTech), 3 IU/ml EPO Eprex (Janssen-Cilag) and 10^{-6} M hydrocortisone (Sigma). During the second phase (day 6 to day 9), cells were co-cultured with MS-5 stromal cells in the basal erythroid medium supplemented with 3 IU/ml EPO Eprex (Janssen-Cilag). During the third phase (day 9 to day 20), cells were co-cultured with stromal MS-5 cells in a basal erythroid medium without cytokines. Erythroid differentiation was monitored by flow cytometry analysis of CD36, CD71, CD235, BAND3 and CD49d erythroid surface markers and of enucleated cells using the DRAQ5 double-stranded DNA dye. 7AAD was used to identify live cells.

174 ***Colony-forming cell (CFC) assay***
175 CD34$^+$ HSPCs were plated at 1x10^3 cells/mL concentration in a methylcellulose-based
176 medium (Stem Cell Technologies) under conditions supporting erythroid and granulo-
177 monocytic differentiation. BFU-E and CFU-GM colonies were counted after 14 days.
178 Colonies were randomly picked and collected as bulk populations (containing at least
179 25 colonies) to evaluate base editing efficiency, globin expression by RT-qPCR and
180 RP-HPLC, and hemoglobin expression by CE-HPLC. BFU-Es were randomly picked
181 and collected as single colonies (around 30 colonies per sample) to evaluate base
182 editing efficiency and globin expression by RT-qPCR.
183
184 ***Genome-wide, unbiased identification of DSBs enabled by sequencing (GUIDE-***
185 ***seq)***
186 Human erythroleukemia K562 cells (2.5 × 10^5) were transfected with 500 ng of Cas9-
187 expressing plasmid, together with 250 ng of each sgRNA–coding plasmid or an empty
188 pUC19 vector (background control), 10 pmol of the bait double-stranded
189 oligodeoxynucleotide (dsODN) (designed according to the original GUIDE-seq
190 protocol; Tsai et al. 2015), and 50 ng of a pEGFP-IRES-Puro plasmid, expressing both
191 enhanced GFP (EGFP) and the puromycin resistance genes. One day after
192 transfection, cells were selected with puromycin (1 µg/ml) for 48 hours to enrich for
193 transfected cells. Cells were then collected, and genomic DNA was extracted using the
194 DNeasy Blood and Tissue Kit (Qiagen) and sheared using the Covaris S200 sonicator
195 to an average length of 500 bp. Library preparation was performed using the original
196 adapters and primers according to previous work (Tsai et al. 2015)
197 End-repair reaction was performed using NEBNext Ultra End Repair/dA Tailing Module
198 and adaptor ligation using NEBNext® UltraTM Ligation Module, as previously
199 described. Amplification steps were then performed following the GUIDEseq protocol
200 previously described (Tsai et al. 2015)
201 Libraries were sequenced with a MiSeq sequencing system (Illumina) using the
202 Illumina MiSeq Reagent kit V2-300 cycles (paired-end sequencing; 2×150-bp). Raw
203 sequencing data (FASTQ files) were analyzed using the GUIDE-seq computational
204 pipeline (Tsai et al. 2016) Identified sites were considered *bona fide* off-targets if a
205 maximum of six mismatches against the on-target were present and if they were
206 absent in the background control.
207
208 ***Evaluation of editing efficiency***
209 Base editing efficiency and InDel frequency were evaluated in HSPC-derived erythroid
210 cells at the end of the first phase of differentiation and in BFU-E and CFU-GM 14 days
211 after plating. Genomic DNA was extracted from control and edited cells using PURE
212 LINK Genomic DNA Mini kit (LifeTechnologies), or Quick-DNA/RNA Miniprep (ZYMO
213 Research), following manufacturers' instructions. To evaluate base editing efficiency
214 at sgRNA target sites, we performed PCR using primers listed in **Table 2**, followed by
215 Sanger sequencing and EditR analysis (Kluesner et al. 2018). TIDE analysis (Tracking
216 of InDels by Decomposition) was also performed to evaluate the percentage of InDels
217 in edited samples (Brinkman et al. 2014).
218 On- and off-target regions were also PCR-amplified and subjected to next-generation
219 sequencing (NGS). We selected the top 10 predicted off-targets (top 5 nominated by
220 GUIDe-seq and top five nominated by COSMID *in silico* analysis) and assessed editing
221 at day 6 of culture. On-target and off-target sites were PCR-amplified using the
222 Phusion High-Fidelity polymerase (M0530; NEB, Ipswich, MA) and primers containing
223 specific DNA stretches (MR3 for forward primers and MR4 for reverse primers; **Table**

2) located 5' to the sequence recognizing the off-target. Amplicons were purified using Ampure XP beads (A63881; Beckman Coulter, Brea, CA). Illumina- compatible barcoded DNA amplicon libraries were prepared by a second PCR step using the Phusion High-Fidelity polymerase (M0530; NEB) and primers containing Unique Dual Index barcodes and annealing to MR3 and MR4 sequences. Libraries were pooled, purified using the High Pure PCR product purification kit (11732676001; Sigma-Aldrich, Saint Louis, MO), and sequenced using Illumina NovaSeq 6000 system (paired-end sequencing; 2×100-bp) to obtain a minimum of 100,000 reads per amplicon. Targeted NGS data were analyzed using CRISPResso2 (Clement et al. 2019)

Table 2. Primers used to detect on- and off-target base editing and InDels events

Amplified region	F/R	Sequence (5' to 3')	Sequencing
+58-kb region	F	TGGACAGCCCGACAGATGAA	Sanger
	R	AAAAGCGATACAGGGCTGGC	Sanger
+55-kb region	F	CCCTATCAGTGCCGACCAAG	Sanger
	R	GTGAGTAGGTAGAGGGGTCAGA	Sanger
+58-kb on target	F	GCAGCGTCAGATGTGTATAAGAGACAGACCGATGGAGAGGTCTGCCAGT	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGTCTACTCTTAGACATAACACACCAGGGTCA	NGS
+58-OT1	F	GCAGCGTCAGATGTGTATAAGAGACAGTGACAGGCACAGAGTAGTGCTCAA	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGGAAAGTGAAGACACAAAGAAAACACAAGGTGAG	NGS
+58-OT2	F	GCAGCGTCAGATGTGTATAAGAGACAGTTGTGTTTTATCCTGAGTGTGTCAGGAAGC	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGATTGCACTGTAGGTCAAGCTACCG	NGS
+58-OT3	F	GCAGCGTCAGATGTGTATAAGAGACAGCTGAATCAACATGAAAAGACTGGCTTAGTCTG	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGGACAGAAACTGGATTTCATTTCTTTATGTCGGC	NGS
+58-OT4	F	GCAGCGTCAGATGTGTATAAGAGACAGTTTTACCTCTGACACCACATGGGG	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGGGTTCACAGTGCCTTTTCTCAGGA	NGS
+58-OT5	F	GCAGCGTCAGATGTGTATAAGAGACAGGTCTCCTAACCGCAAGGCAG	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGAGATGCAGCATGGTTCAGCTTGG	NGS
+58-OT6	F	GCAGCGTCAGATGTGTATAAGAGACAGAATGTGCCAGGGGAGCAGAG	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGTCGGCCTCACAGAATCGGCT	NGS
+58-OT7	F	GCAGCGTCAGATGTGTATAAGAGACAGCTCATCAGCCTGGCAACGTAC	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGCCATCACCTTGACCTGCTCAGA	NGS
+58-OT8	F	GCAGCGTCAGATGTGTATAAGAGACAGAGGAGCTGGACCTCCTGTCT	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGCCAGGATAACCTCCCCATCTC	NGS
+58-OT9	F	GCAGCGTCAGATGTGTATAAGAGACAGGTAGAGCAGGTAGGCAGGGT	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGAACTCCAGGTTGAGCCCCGT	NGS
+58-OT10	F	GCAGCGTCAGATGTGTATAAGAGACAGCAGGCTGAACTGAGTCTTCTGC	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGGCCAGCAAGAAATATGGCCTCAG	NGS
+55-kb on target	F	GCAGCGTCAGATGTGTATAAGAGACAGATGCCATTCTCCCAGCAACACC	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGGTATTGCTGGGAGAAGAAGGGC	NGS

+55-OT2	F	GCAGCGTCAGATGTGTATAAGAGACAGACCCAGGAGAGAT TGGATGATGAG	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGCTATCATGGGCTA TCTACCCTCTG	NGS
+55-OT3	F	GCAGCGTCAGATGTGTATAAGAGACAGTCAGGTCATCTCAC ATGCCAGAGT	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGTCTCAATCCCTTT CTCTCTCCCC	NGS
+55-OT5	F	GCAGCGTCAGATGTGTATAAGAGACAGTGTAGCTGGGGAT GGTACAGG	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGAGAGTAGTCCTCT CTCAGCAGC	NGS
+55-OT6	F	GCAGCGTCAGATGTGTATAAGAGACAGAGCAAGGCATAAAT TTTCCATGA	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGGATTCATCTTGGC ATTGACTCCT	NGS
+55-OT7	F	GCAGCGTCAGATGTGTATAAGAGACAGAGTCTCTCCACATC GCCTTT	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGCAGGATGTGGGC CCAAGG	NGS
+55-OT8	F	GCAGCGTCAGATGTGTATAAGAGACAGCAGCCTGGAGCTG GTATCCAT	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGCAGGCTATAGATC ATTCCCAGCAC	NGS
+55-OT9	F	GCAGCGTCAGATGTGTATAAGAGACAGGGCAAGGTGAGGC CTAGGTA	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGGTTACTTCAGGGA TGACACAGGAC	NGS
+55-OT10	F	GCAGCGTCAGATGTGTATAAGAGACAGAGGTTCAGAGAGC TGAACATGTGG	NGS
	R	TGGGCTCGGAGATGTGTATAAGAGACAGCATAGGGTGAGG TCTAGAAGGATC	NGS

F, forward primers. R, reverse primer.

RT-qPCR

Total RNA was extracted from SCD HSPCs differentiated towards the erythroid lineage (day 13) using the RNeasy micro kit (QIAGEN), and from pool BFU-E and single colonies using the Quick-DNA/RNA Miniprep kit (ZYMO Research). RNA was treated with DNase using the DNase I kit (Invitrogen), following the manufacturer's instructions. Mature transcripts were reverse-transcribed using SuperScript First-Strand Synthesis System for RT-qPCR (Invitrogen) with oligo (dT) primers. RT-qPCR was performed using primers listed in **Table 3**, the iTaq universal SYBR Green master mix (Biorad), and the CFX384 Touch Real-Time PCR Detection System (Biorad).

Table 3 Primers used for RT-qPCR

Amplified region	F/R	Sequence (5' to 3')
HBA	F	CGGTCAACTTCAAGCTCCTAA
	R	ACAGAAGCCAGGAACTTGTC
HBB	F	GCAAGGTGAACGTGGATGAAGT
	R	TAAAGCATCAGGAGTGGACAGA
HBG1+HBG2	F	CCTGTCCTCTGCCTCTGCC
	R	GGA TTGCCAAAACGGTCAC
BCL11A-XL	F	ATGCGAGCTGTGCAACTATG
	R	GTAAACGTCCTTCCCCACCT
GAPDH	F	ACCACAGTCCATGCCATCACT
	R	CCATCACGCCACAGTTTCC

Flow cytometry analysis

HSPC-derived erythroid cells were fixed with 0.05 % cold glutaraldehyde and permeabilized with 0.1 % TRITON X-100. After fixation and permeabilization, cells

were stained with an antibody recognizing CD235 erythroid surface marker (PE-Cy7-conjugated anti- CD235 antibody, 563666, BD Pharmingen) and either an antibody recognizing HbF (FITC-conjugated anti-HbF antibody, clone 2D12 552829 BD), or an antibody recognizing HbS (anti-HbS antibody, H04181601, BioMedomics) followed by the staining with a secondary antibody recognizing rabbit IgG (BV421-conjugated anti-rabbit IgG, 565014, BD). Flow cytometry analysis of CD36, CD71, CD235, BAND3 and CD49d erythroid surface markers was performed using a V450-conjugated anti-CD36 antibody (561535, BD Horizon), a FITC-conjugated anti-CD71 antibody (555536, BD Pharmingen), a PE-Cy7-conjugated anti- CD235 antibody (563666, BD Pharmingen), a PE-conjugated anti-BAND3 antibody (9439, IBGRL) and an APC-conjugated anti-CD49d antibody (559881, BD). Flow cytometry analysis of enucleated or viable cells was performed using double-stranded DNA dyes (DRAQ5, 65-0880- 96, Invitrogen and 7AAD, 559925, BD, respectively). Flow cytometry analyses were performed using Gallios (Beckman coulter) flow cytometers. Data were analyzed using the FlowJo (BD Biosciences) software.

RP-HPLC analysis of globin chains

RP-HPLC analysis was performed using a NexeraX2 SIL-30AC chromatograph and the LC Solution software (Shimadzu). A 250x4.6 mm, 3.6 µm Aeris Widepore column (Phenomenex) was used to separate globin chains by HPLC. Samples were eluted with a gradient mixture of solution A (water/acetonitrile/trifluoroacetic acid, 95:5:0.1) and solution B (water/acetonitrile/trifluoroacetic acid, 5:95:0.1). The absorbance was measured at 220 nm.

CE-HPLC analysis of hemoglobin tetramers

Cation-exchange HPLC analysis was performed using a NexeraX2 SIL-30AC chromatograph and the LC Solution software (Shimadzu). A 2 cation-exchange column (PolyCAT A, PolyLC, Columbia, MD) was used to separate hemoglobin tetramers by HPLC. Samples were eluted with a gradient mixture of solution A (20mM bis Tris, 2mM KCN, pH=6.5) and solution B (20mM bis Tris, 2mM KCN, 250mM NaCl, pH=6.8). The absorbance was measured at 415 nm.

Digital droplet PCR

The 3.2-kb deletion and inversion frequencies were measured by Digital Droplet PCR (ddPCR) using a primer/probe mix (Bio-rad) containing the primers and probes listed in **Table 4**. Control primers annealing to hALB (located on chr 4) were used as DNA loading control. Synthetic double-stranded DNA fragments recapitulating the sequence generated after the occurrence of the 3.2-kb deletion or inversion served as positive controls.

Abundance of BCL11A-XL mRNA in control and edited samples was evaluated by ddPCR. cDNA was prepared as previously described. A primer/probe mix (Bio-rad) containing the primers and probes listed in **Table 5** were used to detect BCL11A-XL. Synthetic double-stranded fragments containing BCL11A-XL sequences were used as positive controls. To detect GAPDH we used the TaqMan Gene Expression assays, VIC (Thermo)

Data were acquired through QX200 analyzer (Bio-Rad) and results were analyzed with QuantaSoftTM Analysis Pro (Bio-Rad). A positive droplet count threshold was set at 30 to allow proper calculation of copy/µL concentration through the application of the Poisson distribution.

304 **Table 4. Primers used for ddPCR**

Amplicon	F/R/Probe	Sequence (5' to 3')
3.2-Kb-Deletion	F	CCACCGATGGAGAGGTCT
	R	AGGAAGGGAAGTGGGTATGG
	Probe	[FAM]CCAGTCCTCTTCTACCCCACC [BHQ1]
3.2-Kb-Inversion	F	TGGCATCTACTCTTAGACATAACAC
	R	AGGAAGGGAAGTGGGTATGG
	Probe	[FAM]ACCAGGGTCAATACAACTTTGAAGC[BHQ1]
BCL11A-XL	F	AGCTCACCAGGCACATGAAA
	R	CACTCGATCACTGTGCCATT
	Probe	[FAM]CAGGTGGGGAAGGACGTTTA [BHQ1]

305
306 ***Sickling assay***
307 180,000 in vitro differentiated RBCs (at day 19 or 20) were resuspended in CellStab
308 (BioRad), placed in an 8-well m-Slide (Ibidi), and incubated under gradual hypoxic
309 conditions (20% O2 for 20 min; 10% O2 for 20 min; 5% O2 for 20 min; 0% O2 for 60-
310 180min). A time course analysis of sickling was performed in real time by video
311 microscopy. Images were captured every 20 min using Zeiss Spinning Disk
312 microscope and a 40x objective. Throughout the time course, images were captured
313 and then processed with ImageJ to determine the percentage of non-sickle RBCs per
314 field of acquisition in the total RBC population. More than 400 cells were counted per
315 condition.
316
317 ***sgRNA design for Cas9-enrichment library preparation***
318 sgRNAs were designed to introduce cuts on complementary strands flanking the
319 region of interest (ROI), using CHOPCHOP online design tool
320 (https://chopchop.cbu.uib.no/) and selected for the highest predicted on-target
321 efficiency and minimal off-target activity. The sgRNAs were assembled as a duplex
322 from synthetic CRISPR RNAs (crRNAs) (Integrated DNA Technologies-IDT, custom
323 designed) and trans-activating crRNAs (tracrRNAs) (IDT, catalog no. 1072532).
324 sgRNA sequences are provided in **Table 5.** The ROI was centered on the expected
325 cut sites of the gene editing approaches on BCL11A gene and has a size of 14.3 kb.
326
327 **Table 5. sgRNA for Cas9-enrichment library**

sgRNA	Target sequence (5'to 3')
BCL11A_1	TGCCCACGTCAAAACACCCT
BCL11A_2	GGTAGGACCCAACACTACGC
BCL11A_3	GCACCCGCTTGTTCCCTGTG
BCL11A_4	CACTATCTTCTCGCACCATC

328 ***Cas9-enrichment library preparation for nanopore sequencing***
329 High molecule weight DNA was extracted from patient's HSPC with the Nanobind CBB
330 kit (PacBio, catalog no. 102-301-900) according to manufacturer's instructions. DNA
331 was size selected using Short Read Eliminator XS kit (PacBio, catalog no. 102-208-
332 200) and quantified using the Qubit fluorometer (Thermo Fisher Scientific). 5 µg of
333 DNA were used for the library preparation using the Cas9 Sequencing Kit (Oxford
334 Nanopore Technologies-ONT, SQK-CS9109) and following the Cas9-mediated PCR-
335 free enrichment protocol (version: CAS_9106_v109_revC_16Sep2020) available

through ONT. Libraries were loaded onto MinION flow cells with R9.4.1 nanopores (ONT, catalog no. FLO-MIN106D). One flow cell was used per biological condition and run on GridION using MinKNOW software for 72 h.

Bioinformatic analysis of Cas9-enrichment library from nanopore sequencing

Raw reads from Nanopore sequencing were preprocessed using cutadapt (v4.4) to remove low quality bases (q=5 threshold at 5' and 3' ends) and to select reads longer that 4 kb. Adaptors were then trimmed using porechop (v 0.2.4) and reads aligned on the human reference genome (GRCh38.p13) using minimap2 (v 2.26-r1175; arguments*: -x map-ont -a -Y --secondary=no*). Reads that align at the enrichment locus (11:5220519-55121185) were re-aligned on this region using minimap2 *(*arguments*: -x map-ont -n 20 -I 1K -r 400,1000 -k 15 -w 15 -a -Y --secondary=no).* Finally, the depth of coverage at each position was measured using samtools (v1.17), the base composition extracted using pysamstats (v1.1.2) and the presence of INDELS detected using the variant caller sniffles (v 1.0.12; arguments*: -n -1 -r 1000 -s 1 -d 1 -l 4*). Additional processing was performed using R (v4.4.1) to prepare figures and tables. INDELS were classified as small, intermediate and large according to their size ([4-50bp], [51-200bp], >200bp). The co-editing at GATA1 and ATF4 loci was assessed using mpileup output from samtools (1.21) using a 5bp window around editing positions of the respective gRNA (GATA_bs_3 and ATF4_bs_2).
Script & pipeline available available on demand. Raw data were deposited on SRA under accession number PRJNA1192026.

RNA-seq

Total RNA was isolated from HD HSPCs 48 h after RNA transfection using the RNeasy Kit (QIAGEN) including a DNAse treatment step. Read quality was assessed using FastQC [version 0.11.9; https://www.bioinformatics.babraham.ac.uk/projects/fastqc/]. Adapter sequences and low-quality bases (Q < 20) were trimmed from raw reads with BBDuk [version 38.92; https://sourceforge.net/projects/bbmap/]; moreover, the first 10 nucleotides were force-trimmed for low quality. Reads shorter than 35 bp post-trimming were discarded. Trimmed reads were aligned to the human reference genome (hg38) using STAR [version 2.7.9a (Dobin et al. 2013)]. Raw gene counts were generated in R-4.1.1 using the featureCounts function of the Rsubread package [version 2.8.1 (Y. Liao, Smyth, e Shi 2014)] and the GENCODE 44 basic gene annotation for hg38 reference genome. Raw gene counts were normalized to counts per million mapped reads (CPM) and to fragments per kilobase of exon per million mapped reads (FPKM) using the edgeR R package [version 3.36.0 (Robinson, McCarthy, e Smyth 2010)]; only genes with a CPM greater than 1 in at least 3 samples were retained for differential analysis. Differential gene expression analysis was performed using the glmQLFTest function of the edgeR R package, using donor as a blocking variable. Genes with FDR < 0.05 and absolute log2 fold change ≥ 1 were defined as differentially expressed. Functional enrichment analysis was performed with the clusterProfiler R package [version 4.12.5 (T. Wu et al. 2021)].

HSPC xenotransplantation in NBSGW mice

NOD.Cg-KitW-41JTyr +PrkdcscidIl2rgtm1Wjl/ThomJ (NBSGW) mice were housed in a pathogen-free facility. Control or edited mobilized $CD34^+$ cells (3.5×10^5 cells per mouse) were transplanted into non-irradiated NBSGW female mice of 6 to 9 weeks of age via retro-orbital sinus injection. NBSGW female mice were conditioned with busulfan (Sigma-Aldrich) injected intraperitoneally (15 mg/kg body weight) 24 hours

before transplantation. Sixteen weeks after transplantation, NBSGW primary recipients were euthanized. Cells were harvested from bone marrow (BM), thymus, spleen, and blood, and stained with antibodies against the following murine and human surface markers: murine CD45 (1/50 mCD45-VioBlue; Miltenyi Biotec), human CD45 (1/50 hCD45-APCvio770; Miltenyi Biotec). BM cells were also stained with human CD3 (1/50 CD3-APC; Miltenyi Biotec), human CD14 (1/50 CD14- PECy7; BD Biosciences), human CD15 (1/50 CD15-PE; Miltenyi Biotec), human CD11b (1/100 CD11b-APC; Miltenyi Biotec), human CD19 (1/100 CD19-BV510; BD Biosciences), human CD235a (1/50 CD235a-PE; BD Biosciences), human CD71 (1/10 CD71-APC; BD Biosciences), CD36 (1/50 CD36-FITC; BD Bio- sciences), and CD34 (1/100 CD34-PE-Vio770; Miltenyi Biotec). Cells were analyzed by flow cytometry using the Novocyte analyzer (Agilent) and the FlowJo software (BD Biosciences).

Human BM CD45+ cells were sorted by immunomagnetic selection with CD45 MicroBeads (Miltenyi Biotec). Furthermore, BM cells were subjected to immunostaining with biotinylated antibodies that recognized the following surface markers: CD3 (dilution 1/25, clone HIT3a; BD), CD19 (dilution 1/25, clone HIB19; BD), B220 (dilution 1/50, clone RA3-6B2; BD), Ter119 (dilution 1/50, clone TER-119; BD), and mCD117 (clone 2B8; BD). BM cells were washed and incubated with 20 µL of Anti-Biotin beads (Miltenyi Biotec). After washing, the cells were magnetically purified using an LS column (Miltenyi Biotec) according to the manufacturer's instructions. Cells from the positive fraction were immuno-stained with the following antibodies: CD19-PE (dilution 1/20, BD), streptavidin (SA)-APC (dilution 1/20 BD) and hCD45-BV510 (dilution 1/100, BD). The hCD45^{high}/CD19^{high} were sorted using MA900 cell sorter and subjected to RT-qPCR analysis. Cells from the negative fraction were immuno-stained with the following antibodies: CD235a-PE (dilution 1/5000, BD) and hCD45-BV510 (dilution 1/100, BD). The hCD45$^{low/-}$/CD235a^{high} cells were sorted using the MA900 cell sorter (Sony Biotechnology, San Jose, CA) and subjected to flow cytometry, reverse phase high-performance liquid chromatography (RP-HPLC) and RT-qPCR analysis.

All experiments and procedures were performed in compliance with the French Ministry of Agriculture's regulations on animal experiments and were approved by the regional Animal Care and Use Committee (APAFIS#2019061312202425_v4). Mice were housed in a temperature- (20◦C-22◦C) and humidity (40%-50%)-controlled environment with 12:12 hour light-dark cycle and fed *ad libitum* a standard diet.

Results

Base editing in disrupts the +58-kb and +55-kb BCL11A enhancers without affecting SCD HSPC viability and differentiation

To disrupt the *BCL11A* erythroid enhancers, we targeted the +58-kb and +55-kb regions in human SCD HSPCs by inserting point mutations at the GATA1 and ATF4 BS, respectively (**Figure 1A**). SCD HSPCs were transfected with various combinations of *in vitro* transcribed BE mRNAs and chemically modified sgRNAs, and afterwards subjected to erythroid differentiation (**Figure 1B**). We evaluated the base editing efficiency in erythroblasts derived from transfected HSPCs. At both the +58-kb and +55-kb regions, different combinations of CBEs/ABEs and sgRNAs led to the generation of several editing profiles. In the +58-kb region, CBEs created the +58 CBEI profile (GTGATAAA; targeted nucleotides are underlined; 55.6% ± 4.7) and ABEs created the +58 ABEI (GTGATAAA; 28.4% ± 6.1), +58 ABEII (GTGATAAA; 19.0% ± 3.3) and +58 ABEIII profiles (GTGATAAA; 14% ± 2.0; **Figure 1C**). In the +55-kb region, CBEs led to the +55 CBEI (TTGCATCATCC; 47.0% ± 7.0) and +55 CBEII profiles (TTGCATCATCC; 46.7% ± 3.7), while ABEs generated the +55 ABEI (TTGCATCATCC; 71% ± 3.5) and the +55 ABEII profiles (TTGCATCATCC; 87.0% ± 5.1; **Figure 1D**). To compare base editing and Cas9 nuclease strategies, we transfected Cas9 nuclease RNP complexes targeting either the GATA1 or ATF4 BS in the +58-kb or the +55-kb region (InDels at the +58 -kb: 75.8% ± 3.0; InDels at the +55-kb region: 82.3% ± 4.1 **Figure 1E-F**). Low or no InDels were detected in base edited samples, confirming the DSB-low nature of BEs (**Figure 1E-F**). Transfected SCD HSPCs were plated on a semi-solid medium that allows erythroid and granulocyte/monocyte differentiation at clonal level (CFC assay). No differences in the frequency of erythroid (BFU-E) and granulocyte/monocyte (CFU-GM) colonies were detected between control and edited samples, demonstrating that our therapeutic strategy does not affect progenitors' viability and differentiation towards the erythroid and granulomonocytic lineages (**Figure 1G**). The base editing efficiency and InDel profiles in pool BFU-E and CFU-GM population were similar to those measured in the liquid erythroid cultures (**Supplementary Figure 1A-H)**. In conclusion, we efficiently targeted the GATA1 and ATF4 BSs at the +58-kb and +55-kb BCL11A enhancer regions with base editing strategies without affecting HSPCs viability and differentiation potential. This approach allowed for precise modifications of alternative bases, resulting in significantly fewer DSBs compared to the Cas9 nuclease.

HbF reactivation after base editing of BCL11A enhancers in SCD HSPCs

We next differentiated transfected SCD HSPCs in liquid culture towards the erythroid lineage to produce mature RBCs. Flow cytometry analysis of enucleation and early and late erythroid markers showed no difference between control and edited groups, indicating no impairment of SCD HSPC erythroid differentiation upon targeting of the GATA1 and the ATF4 BS in the *BCL11A* erythroid enhancers (**Supplementary Figure 2A-E**). We assessed *HBG* reactivation at RNA and protein levels by RT-qPCR, RP-HPLC and CE-HPLC (**Figure 2A-C**). Regarding the +58-kb region, CBE-treated cells bearing the +58 CBEI profile, expressed the highest HbF level among base edited groups and compared to the control groups. Notably, the levels of HbF obtained in cells carrying the +58 CBEI profile were only modestly lower compared to the +58 Cas9 profile, despite the significantly lower editing efficiency (56.0% ± 4.7 and 76% ± 3.0 in +58 CBEI and +58 Cas9, respectively). ABE-treated samples exhibited low HbF levels that were consistent with limited editing efficiency (**Figure 2A-C**). Interestingly, cells

472 carrying the +58 ABEIII profile showed HbF levels similar to those achieved with ABEI
473 and ABEII, despite having the lowest editing, suggesting that this base conversion is
474 potentially highly productive in terms of HbF reactivation (**Figure 2A-C**). To increase
475 the editing efficiency in the +58 ABEIII profile, we combined the sgRNA used to create
476 this profile (GATA_bs_3 sgRNA) with the more efficient ABE8e (T. P. Huang et al.
477 2023), instead of the first generation ABEmax. However, due to a wider editing window,
478 we were not able to generate the +58 ABEIII profile, but instead, we recreated the +58
479 ABEII profile with an editing efficiency of 80.5% ± 0.7 that was associated with a higher
480 Indel frequency of 4.4% ± 2.8 **(Supplementary Figure 3A and B)**.
481 Concerning the +55-kb region, CBE- and ABE-treated samples showed HbF
482 reactivation, with an extent similar to Cas9-treated samples even though the CBE
483 editing frequency was around two-fold lower compared to ABE and Cas9 (**Figure 2A-**
484 **C**).
485 Similar results in term of HbF reactivation were observed also in pool of BFU-E
486 (**Supplementary Figure 1I**). Flow cytometry analysis of F- and S-cells evidenced
487 similar results in terms of HbF-expressing cells with a concomitant slight reduction of
488 HbS-expressing cells (**Figure 2D**). Finally, a sickling assay was performed in control
489 and edited samples. Edited samples showed decreased frequencies of sickle cells,
490 that were in accordance with the levels of HbF (**Figure 2E**), with +58 CBEI, +58 Cas9,
491 + and +55 Cas9 profiles causing the highest reduction in the frequency of sickle cells
492 (**Figure 2E**). Overall, these data showed that base editing-mediated disruption of the
493 GATA1 or ATF4 activator BS in the +58-kb or +55-kb regions, respectively, led to HbF
494 reactivation, which, however, was at best only approaching that achieved with the
495 approved strategy using Cas9 to target the +58-kb region. Moreover, phenotypic
496 rescue (evaluated by performing a sickling assay) was incomplete and inconsistent
497 across different donors for all editing profiles, including +58 Cas9, indicating the need
498 for further optimization of the approach. We hypothesize that this can be accomplished
499 by further increasing HbF expression.
500
501 **Dissecting GATA1 and ATF4 binding motifs to achieve robust HbF reactivation**
502 To precisely evaluate the potency of the different editing profiles, we correlated γ-
503 globin expression and editing efficiency at the clonal level in BFU-Es derived from
504 edited SCD HSPCs. Single colonies were analyzed for *HBG1/2* mRNA expression,
505 base editing efficiency or InDel frequency (**Figure 3A and B**). We observed a positive
506 correlation between editing efficiency and γ-globin reactivation in all the groups. For
507 the +58-kb enhancer, the generation of the +58 ABEIII profile was significantly the most
508 potent profile in terms of γ-globin reactivation, outperforming even the +58 Cas9 profile
509 (**Figure 3A**). However, generating this profile with high editing efficiency proved to be
510 challenging with current BEs (**Supplementary Figure 3A**). Conversely, higher editing
511 efficiency obtained for the +58 ABEI and +58 ABEII profiles led to the generation of bi-
512 allelic colonies, making these profiles more suitable for future experiments (**Figure**
513 **3A**). Given the higher number of bi-allelic colonies and the better correlation between
514 editing frequency and *HBG1/2* levels, the +58 ABEII profile was selected for further
515 testing. The results emphasize the importance of specific bases in the GATA1 BS motif
516 for GATA1 binding and γ-globin regulation. Targeting the A6 position (by generating
517 the +58 ABEIII profile: GTGAT<u>A</u>AA) significantly increased γ-globin expression,
518 indicating its pivotal role. However, simultaneous targeting of A6 to A8 (+58 ABEII)
519 reduced γ-globin expression, suggesting that modifications at A7 and A8 modulate
520 GATA1 binding or recruit other transcriptional activators, leading to decreased γ-globin
521 reactivation (Behera et al. 2018). Additionally, T>C conversions at T2 and T5 (+58

522 ABEI) were less effective in disrupting GATA1 BS, indicating a lower impact on GATA1
523 binding capacity, as previously reported (Behera et al. 2018) (**Figure 3A**). Finally, the
524 activity of the +58 CBEI profile mirrored that of the +58 Cas9 profile, leading to the
525 highest γ-globin reactivation amongst the different editing profiles. This confirmed
526 previously reported observations of the importance of G in position 3 for GATA1
527 binding (Behera et al. 2018); thus, the + 58 CBEI profile was also chosen for further
528 testing (**Figure 3A**).
529 With regards to the +55-kb enhancer, all the profiles exhibited comparable γ-globin
530 reactivation (with the exception of the +55 ABE I showing the lowest *HBG1/2* levels)
531 but we selected +55 CBEII and +55 ABEII profiles that showed a better correlation
532 between editing frequency and *HBG1/2* levels **(Figure 3B**). These data suggest that
533 insertion of C>T conversion at C4 of the binding motif to generate +55 CBEI profile
534 (TTGCATCATCC) does not further impact the binding capacity of ATF4 compared to
535 targeting only the last 3 Cs in the +55 CBEI profile (TTGCATCATCC). On the other
536 hand, reduced γ-globin levels were associated with the +55 ABEI and +55 ABEII
537 profiles, suggesting that A>G mutations (and in particular targeting of only 1 A of the
538 ATF4 motif in the +55 ABEI profile) does not strongly impact ATF4 binding.
539
540 ***Enhanced γ-Globin Reactivation Through Multiplex Editing of +58 and +55***
541 ***Enhancers***
542 We hypothesized that simultaneous editing of the +58-kb and +55-kb enhancers of
543 SCD HSPCs could lead to a stronger reduction of *BCL11A* expression, and
544 consequently greater HbF induction compared to the editing of a single enhancer
545 alone. We transfected SCD HSPCs with the base editors and sgRNAs, aiming at
546 creating the most potent editing profiles (**Figure 3A and B**). In erythroid liquid cultures,
547 modestly reduced editing efficiencies were observed when inserting the combination
548 of +58 ABEII and +55 ABEII profiles compared to samples transfected with the
549 individual ABE/sgRNA combinations (70% ± 18.1 and 52.0% ± 11.5 achieved for +58
550 ABEII/+55 ABEII profiles at GATA1 and ATF4 BS respectively, compared to 77.7 % ±
551 7.6 and 60.0% ± 5.7 obtained for +58 ABEII and +55 ABEII single editing profiles,
552 respectively; **Figure 4A and B**). We then used CBEs to insert the most potent CBE
553 profiles, +58 CBEI and +55 CBEII, reaching editing efficiencies similar to those
554 achieved by single editing (61.3% ± 15.5 and 54.3% ± 17.0, respectively, for the
555 combined +58 CBEI/+55 CBEII profile compared to 60.7% ± 14.5 and 45.3% ± 22.8
556 for +58 CBEI and +55 CBEII individual profiles, respectively; **Figure 4A and B**). We
557 also used a recently developed, dual function, cytosine and adenine BE (TadDE;
558 Neugebauer et al., 2022), combined with the same sgRNAs used with CBE
559 (GATA1_bs_1 and ATF4_bs_2) to simultaneously introduce C>T and A>G point
560 mutations at the GATA1 and ATF4 BSs, generating two additional profiles (+58 DBEI:
561 GTGATAAA; +55 DBEI: TTGCATCATCC), reaching high editing efficiencies upon
562 single or multiplex editing (65.3% ± 22.1 and 57.7% ± 5.0 for +58 DBEI and +55 DBEI
563 individual profiles, and 67.0% ± 13.0 and 49% ± 7.2 for the combined +58 DBEI/+55
564 DBEI profile; **Figure 4A and B**). Importantly, base-edited samples showed no or few
565 InDels (**Figure 4C and D**). In parallel, we used Cas9 nuclease to simultaneously target
566 both enhancers. The efficiency of concomitant generation of the +58 and + 55 Cas9
567 profiles was similar to that obtained upon single Cas9 editing (41.9% ± 9.6 and 40.6%
568 ± 15.4, when targeting the GATA1 and ATF4 BS simultaneously and 49.6% ± 9.6 and
569 44.7% ± 21.1 for individual +58 Cas9 and +55 Cas9 profiles, respectively; **Figure 4C**
570 **and D**). This approach potentially generates a 3.2-kb deletion or inversion that disrupts
571 the targeted GATA1 and ATF4 BSs and removes additional BSs for transcriptional

activators (**Figure 4E**). ddPCR showed 33.1% ± 8.1 of deletion and 10.4% ± 2.3 of inversion for the +58/+55 Cas9 profile (**Figure 4F**).). On the other hand, no deletion or inversion was detected with the +58 ABEII/+55 ABEII profile, and only 0.5% ± 0.4 and 1.1% ± 0.9 of deletion were detected for +58 CBEI/+55 CBEII and +58 DBEI/+55 DBEI profiles, respectively (**Figure 4F**). A CFC assay demonstrated no difference in terms of viability and clonogenic potential between control, single edited and multiplex edited SCD HSPCs (**Figure 4G**). The frequency of base editing, InDel and 3.2-kb deletion/inversion in the pool of BFU-E populations were similar to those measured in the liquid erythroid cultures (**Supplementary Figure 4A-E**).

Overall, in erythroid liquid cultures and pool of BFU-E multiplex editing led to greater down-regulation of *BCL11A-XL* (encoding the isoform responsible for HbF silencing) and γ-globin production compared to single editing, regardless of the specific editor used (**Figure 4H-I** and **Supplementary Figure 4F**). To circumvent the disparity in editing efficiency between the samples, we correlated HbF reactivation and editing efficiency obtained in erythroid liquid cultures and pool of BFU-Es (**Figure 4J** and **Supplementary Figure 6A-D**). This analysis confirmed that multiplex editing correlated with stronger HbF reactivation compared to single editing (**Supplementary Figure 6A-D**). A comparison of multiplex edited samples showed a significantly higher potency of Cas9, followed by DBEs, CBEs and lastly ABEs (**Figure 4J**).

CBE- and DBE-mediated multiplex targeting of BCL11A efficiently rescues the sickling phenotype

We selected CBE and DBE for further efficacy and safety analyses, given the higher HbF reactivation achieved with the +58 CBEI/+55 CBEII and +58 DBEI/+55 DBEI profiles compared to the ABE profiles. SCD HSPCs were co-transfected with the AncBE4max CBE- (Koblan et al. 2018) or the TadDE DBE- (Neugebauer et al. 2023)encoding mRNAs and the same sgRNAs. Of note, while most of the enzymes used in the screening phase (**Figure 1 and 2 and Supplementary Figure 1**) are nearly PAM-less, both these base editors recognize a protospacer followed by an NGG PAM, to reduce the occurrence of off-targets (T. P. Huang et al. 2023). Control and edited HSPCs were differentiated towards the erythroid lineage. Flow cytometry analysis of enucleation and early and late erythroid markers, showed no differences between control and edited samples (**Supplementary Figure 7**).

Editing efficiency in erythroid liquid cultures was similar with CBE and DBE at both GATA1 and ATF4 BSs. In particular, in the +58-kb enhancer 64.6% ± 5.3 and 67.0% ± 9.6 of editing efficiency were obtained with CBE and DBE, respectively (**Figure 5A**). Regarding the +55-kb enhancer, an editing efficiency of 67.8% ± 5.4 and 58.4% ± 12.0 was observed for CBE and DBE, respectively (**Figure 5B**). NGS analysis confirmed high editing efficiency with the generation of predicted base modification for all the profiles (**Supplementary Figure 8A-D**). Rare C>G and InDel events were observed at similar frequency when generating the +58 CBEI/+55 CBEII and the +58 DBEI/+55 DBEI profiles (**Figure 5C-D; Supplementary Figure 8A-E**). In particular, low frequency of Indels was detected when using CBE or DBE at both GATA1 (0.4% ± 0.2 and 0.4% ± 1.0 obtained with CBE and DBE, respectively,) and ATF4 BSs (1.8% ± 0.3 and 2.1% ± 0.8 with CBE and DBE, respectively) (**Supplementary Figure 8A-E**). As previously observed, a higher frequency of 3.2-kb deletion was observed when generating the +58 DBEI/ +55 DBEI profile compared to +58 CBEI/+55 CBEII profile (1.2% ± 0.5 with DBE and 0.3% ± 0.2 with CBE, **Figure 5E)**.

Following a marked reduction in *BCL11A-XL* mRNA observed in the edited samples compared to the controls (**Figure 5F**), a significant and consistent reactivation of γ-

622 globin at the mRNA level was detected in the edited samples. Specifically, 39.5% ±
623 16.7% and 38.0% ± 18.4% of mRNAs encoding γ-globin were detected in HSPCs-
624 derived cells carrying the +58 CBEI/+55 CBEII and +58 DBEI/+55 DBEI profiles,
625 respectively (**Figure 5G**). These data were also confirmed at protein level in mature
626 RBCs, by analyzing both globin chains and hemoglobin tetramers with HbF accounting
627 for 29.1% ± 3.6 and 28.3% ± 9.8 of total hemoglobin in the +58 CBEI/+55 CBEII and
628 +58 DBEI/+55 DBEI profile, respectively (**Figure 5H-I**). Importantly, an increase of
629 HbF-expressing cells was observed by generating the +58 CBEI/+55 CBEII and +58
630 DBEI/+55 DBEI profiles (**Figure 5J**). Concurrently, a significant and homogeneous
631 reduction in the frequency of HbS-expressing cells was observed in the edited samples
632 compared to controls (**Figure 5J**). Finally, only 43.0% ± 15.0 and 38.7% ± 10.5 of sickle
633 cells were detected in CBE- and DBE-edited samples, respectively. These frequencies
634 are similar to those observed in heterozygous, asymptomatic SCD carriers (**Figure**
635 **5K**).
636 Overall, these data showed the efficiency of our strategy in reactivating HbF and
637 rescuing the sickle phenotype.
638
639 ***CBE- and DBE-mediated multiplex targeting of BCL11A: genotoxicity***
640 ***assessment***
641 To evaluate the integrity of the locus, we performed long-read sequencing of the region
642 encompassing the GATA1 and ATF4 BSs. We generated the +58/ +55 Cas9 profile
643 and compared it with our base editing profiles. Editing efficiencies between 27% and
644 43% were achieved in both BSs when generating the +58 CBEI/+55 CBEII, +58
645 DBEI/+55 DBEI and +58/+55 Cas9 profiles (**Figure 6A and B**). Moreover, long-read
646 sequencing showed a higher frequency of alleles simultaneously edited at GATA1 and
647 ATF4 BSs when generating the +58 DBEI/+55 DBEI compared to the +58 CBEI/+55
648 CBEII profile (**Figure 6D**).
649 As expected, lower frequencies of alleles with 3.2-kb deletion/inversion were observed
650 when using BEs compared to Cas9 nuclease (**Figure 6C**). For the remaining alleles,
651 InDels associated with the generation of DSBs at the target sites were more frequent
652 for Cas9 nuclease than BEs (**Figure 6C**). In addition, when using BE, the majority of
653 InDels were small (<50 bp) except for the +58 DBEI/+55 DBEI profile that showed 3.2%
654 of large deletion (>200bp) at the GATA1 BS, (**Figure 6C**).
655 To identify sgRNA-dependent potential off-target sites, we performed an unbiased
656 genome wide analysis, GUIDE-Seq, in K562 cells. Interestingly, the top ten predicted
657 off-target sites mapped to non-coding sequences and harbored more than 3
658 mismatches (**Figure 6E**). NGS of the top 5 GUIDE-seq predicted off target and of the
659 top 5 *in silico* predicted off-targets **(Supplementary Table 1**) in erythroblasts derived
660 from edited SCD HSPCs validated few off-targets, which were limited to intronic and
661 intergenic regions, suggesting no impact at protein level (**Figure 6F**). Additionally, off-
662 target activity did not affect the expression of targeted genes in HSPCs
663 (**Supplementary Table 2)**. Notably, although on-target activity was similar, CBE
664 showed higher off-target activity compared to DBE, suggesting a higher specificity of
665 the latter enzyme (**Figure 6E**). Importantly, no InDels were detected at the different
666 off-target sites (**Figure 6F**).
667 Finally, to examine the effect of base editing on the overall gene expression profile, we
668 performed RNA-seq of control HSPCs and SCD HSPCs edited with DBE, CBE or Cas9
669 nucleases targeting both +58-kb and +55-kb enhancers. Overall, we observed few or
670 no differentially expressed genes (DEGs) in DBE- and CBE-treated samples,
671 respectively (**Supplementary Table 2**). On the contrary, 49 genes were dysregulated

672 in Cas9 -treated samples and belonged to the P53 pathway (**Supplementary Table**
673 **2**).
674 Overall, these data showed that CBE- and DBE-mediated multiplex targeting of
675 *BCL11A* is a safe approach compared to the CRISPR/Cas9 treatment.
676
677
678 **Efficient multiplex base editing of *BCL11A* enhancers in repopulating HSCs**
679 To evaluate the ability of BEs to simultaneously target the *BCL11A* enhancers in
680 repopulating HSCs, we xenotransplanted control and CBE- or DBE-edited HD HSPCs
681 into immunodeficient NBSGW mice (**Figure 7A**). Sixteen to seventeen weeks post-
682 transplantation, no differences were observed between edited and control HSPCs with
683 regard to engraftment and differentiation potential, as measured by the frequency of
684 human CD45+ cells in hematopoietic tissues and the proportion of the different
685 lineages (**Figure 7B and Supplementary Figure 9A**). The base editing efficiency in
686 human BM cells (33.7% ± 4.6/59.2% ± 13.3 for +58 CBEI/+55 CBEII, and 68.8% ±
687 15/63.1% ± 9.4 for +58 DBEI/+55 DBEII) was similar to the frequency observed in the
688 input cells and in the other hematopoietic compartments (blood, erythroid precursors,
689 thymus, spleen and B-cells; **Figure 7C-F**). However, CBE editing efficiency at the +58
690 enhancer tends to be reduced in vivo compared to the input HSPCs (**Figure 7C**).
691 InDels were detected at low frequencies in engrafted cells (**Supplementary Figure**
692 **9B**).
693 Human CD235+ erythroid cells sorted from BM of mice transplanted with edited cells,
694 showed increased HbF expression at the RNA and protein levels compared to control
695 groups (**Figure 7G, and Supplementary Figure 9C and D**). Notably, higher HbF
696 reactivation was observed in samples carrying the +58 DBEI/+55 DBEI profile, which
697 showed the highest base editing efficiency (**Figure 7G, and Supplementary Figure**
698 **9C and D**). Correlation analysis between HbF expression and editing efficiency in BM
699 CD235a+ sorted cells also confirmed a more potent HbF reactivation associated with
700 the +58 DBEI/+55 DBEII profile *in vivo* compared to the +58 CBEI/+55 CBEII profile
701 (**Supplementary Figure 9E**). Importantly, *BCL11A* down-regulation was observed in
702 edited human BM CD235a+ cells (**Figure 7H**), while no differences were detected
703 between control and edited samples in BM human CD19+ B-cells (**Figure 7I**).
704 Human CD45+ BM cells were isolated and subjected to a CFC assay. Control and
705 edited samples showed a similar number of erythroid and granulo-monocytic colonies,
706 demonstrating no impact of base editing on the clonogenic potential of engrafted
707 human cells (**Figure 7J**). Editing efficiency in BFU-E and CFU-GM was similar to that
708 measured in the input populations except for the GATA1 BS in the +58-kb enhancer in
709 CBE-treated samples (**Supplementary Figure 10A-D**). *BCL11A-XL* downregulation
710 was observed in BFU-E (**Supplementary Figure 10E**), resulting also in globin
711 reactivation at RNA and protein levels (**Supplementary Figure 10G to I**).
712 Overall, these results suggest high efficiency and a good safety profile of our multiplex
713 base editing approach in HSCs and their differentiated progeny.
714
715
716
717
718

Discussion

Since the discovery of elements controlling the expression of HbF, a variety of genome editing strategies based on the disruption of cis- or trans-regulatory elements of *HBG1/2* to induce HbF reactivation, have been developed for the treatment of β-hemoglobinopathies (Y. Wu et al. 2019; P. Huang et al. 2020; Weber et al. 2020; Antoniou et al. 2022).

Importantly, Cas9 nuclease-mediate disruption of the +58-kb erythroid-specific *BCL11A* enhancer has been recently approved as a clinical therapy for patients with SCD and β-thalassemia (Frangoul et al. 2024; Locatelli et al. 2024). However, although significant HbF reactivation was achieved, the clinical study showed variability between the patients in HbF levels. Furthermore, SCD patients retained high HbS levels and showed a modest correction of ineffective erythropoiesis. Thus, further optimization of the strategy is advisable to reach the complete rescue of the phenotype by reaching higher levels of HbF. In addition, the generation of Cas9 nuclease-induced DSBs or large genomic rearrangements can be detrimental for HSCs (Kosicki, Tomberg, e Bradley 2018; Cullot et al. 2019; Schiroli et al. 2019; Blattner et al. 2020; Leibowitz et al. 2021). Therefore, it has been lately proposed to use the base editing technology to disrupt the GATA1 activator BS in the erythroid-specific *BCL11A* enhancer without creating InDels, and reactivate HbF in the progeny of SCD and β-thalassemia patient HSPCs (Zeng et al. 2020; C. Li et al. 2021; J. Liao et al. 2023).

Here, we used SCD HSPCs to perform a screening of sgRNAs, combined with different BEs that allowed us to modify the GATA1 and ATF4 activator BS in the erythroid-specific *BCL11A* enhancer regions (+58-kb and +55-kb, respectively) in alternative ways, based on the specific targeted nucleotides. Indeed, modifying the GATA1 motif by targeting different nucleotides would differentially decrease the ability of GATA1 to bind to the *BCL11A* enhancer, and thus modulate gene expression (Behera et al. 2018). Importantly, we were able to create different editing profiles for both motifs, by combining different sgRNAs and ABEs or CBEs to generate A>G or C>T base conversions. Additionally, the use of TadDE BEs allowed us to extend the number of edited bases and then further disrupt the GATA1 and ATF4 BSs.

When targeting individually the GATA1 and ATF4 BSs, only the generation of the +58 CBEI profile at the GATA1 BS led to HbF expression and consequent rescue of the sickle phenotype, similar to the approved Cas9 nuclease-based approach. Thus, to further increase HbF levels, we hypothesized that combined editing of the +58-kb and +55-kb enhancers could lead to a stronger reduction of *BCL11A* expression, and consequently greater HbF induction compared to editing a single enhancer alone. Of note, multiplex targeting of the +58-kb and +55-kb *BCL11A* enhancers using CRISPR/Cas9 (although targeting different binding site locus in the +55-kb enhancer) was recently reported to efficiently reactivate HbF expression (Demirci et al. 2021; Zeng et al. 2023). However, the use of CRISPR/Cas9 nucleases cause potentially dangerous DSBs and large genomic rearrangements.

Correlation analysis on single cell-derived colonies allowed us to identify the critical bases involved in the binding of GATA1 and ATF4 at the *BCL11A* enhancers and the most potent profiles in terms of HbF reactivation that we can combine to further improve the pathological phenotype.

Regarding the +58-kb region, we confirmed previous studies that reported an important role of the G_3 of the GATA binding motif in regulating GATA1 binding (Behera et al. 2018). Thus, amongst the different editing profiles, the +58 CBEI (<u>G</u>T<u>G</u>ATAAA) was associated with the highest γ-globin levels, mirroring the effect obtained with the

769 approved Cas9 nuclease strategy. In addition, the A_6 position in the GATA1 BS
770 (GTGATAAA; +58 ABEIII profile) plays an important role for GATA1 binding, as already
771 observed in other works (Behera et al. 2018) and evidenced by the strong increase of
772 γ-globin re-activation. In contrast, targeting the positions A_6 to A_8 in the GATA1 binding
773 motif (GTGATAAA) to generate the +58 ABEII profile resulted in reduced γ-globin
774 expression compared with the +58 ABEIII profile. This reduction suggests that the A>G
775 conversion at the positions A_7 and A_8 in addition to the A_6 base conversion might
776 restores a binding motif for transcriptional activator, which affects *BCL11A* and,
777 therefore, γ-globin expression.
778 Similarly, T>C conversions generated at position T_2 and T_5 (GTGATAAA) to
779 create the +58 ABEI profile led to less efficient disruption of GATA1 BS compared to
780 the generation of the +58 ABEIII profile, although previous studies reported high
781 GATA1 displacement when T>C mutation was occurring at position 5 of the GATA
782 binding motif compared to the targeting of A_6 (Behera et al. 2018) . These results
783 suggest that a site-specific requirements for GATA1 binding at the +58-kb enhancer,
784 likely because the surrounding sequence could also influence GATA1 occupancy
785 (Behera et al. 2018).
786 Regarding the +55-kb region, the generation of the +55 CBEI profile led to C>T
787 conversion of all the Cs (C_4, C_7, C_{10} and C_{11}) of the ATF4 binding motif
788 (TTGCATCATCC), while with the generation of +55 CBEII profile only C_7, C_{10} and C_{11}
789 were converted (TTGCATCATCC). According to published ChIP-seq data, C>T
790 conversion of C_{10} and C_{11} are not implicated in ATF4 binding (P. Huang et al. 2020).
791 Our data suggest that the insertion of two points mutations at C_4 and C_7 (+55 CBEI
792 profile) does not further displace ATF4 as compared to the generation of C_7 mutation
793 alone (+55 CBEII profile), indicating that the primary factor affecting ATF4 binding (and
794 subsequently BCL11A expression and γ-globin reactivation), is the C>T conversion at
795 C_7. Conversions of the A_5 and A_8 of the ATF4 binding motif (TTGCATCATCC, + 55
796 ABEI and TTGCATCATCC, + 55 ABEII) led to low *HBG* levels, suggesting a modest
797 impact on ATF4 binding. However, the generation of the +55 ABEII profile
798 (TTGCATCATCC) showed a higher effect in terms of γ-globin reactivation compared
799 to the +55 ABEI profile (TTGCATCATCC). This result can either indicate a stronger
800 impact of A_8 mutation in the displacement of ATF4 or that the simultaneous generation
801 of A_5 and A_8 mutation enhanced the disruption of the ATF4 binding motif.
802 Finally, we used DBE to further disrupt the GATA1 and ATF4 BSs; however, *in*
803 *vitro* we observed only a modest increase in HbF reactivation compared to the use of
804 CBE alone, suggesting a minor role of As in the binding of these factors to the *BCL11A*
805 enhancers.
806 Multiplex targeting of the +58-kb and +55-kb regions using CBE or DBE led to
807 a significant reduction of *BCL11A* expression and consequently to a potent HbF
808 reactivation. Importantly, HbF reactivation for most of the patient samples reached the
809 levels needed to ameliorate the clinical manifestations of SCD, which have been
810 defined as 70% of HbF-expressing cells and HbF accounting for 30% of the total
811 hemoglobin (Steinberg et al., 2014; Magnani et al., 2020).
812 Importantly, fewer genomic rearrangements were detected in base-edited
813 samples compared to cells treated with Cas9 nuclease, highlighting the better safety
814 profile of our approach. Studies on large genomic rearrangements occurring in the 3.2-
815 kb region between the GATA1 and ATF4 showed a higher frequency of deletions when
816 using DBE compared to the CBE, while no large deletions were observed with ABE.
817 While higher rearrangements with DBE can be explained simply by the additive effect
818 of CBE and ABE activity, the reduced deletion frequency observed with ABE compared

819 to CBE is probably due to an inefficient excision of inosines (in the case of ABEs)
820 compared to the excision of the uracil base (in the case of CBEs), which reduces the
821 generation of DSBs. In addition, the C>G base conversions that was observed with
822 CBE and DBE have been previously associated with DSB generation, deletions and
823 other large genomic rearrangements (M. E. Huang et al. 2024). As for the 3.2-kb
824 deletion, also on-target unintended large deletions (> 200 bp) were more frequent in
825 DBE-treated samples (+58 DBEI/+55 DBEI profile) compared to CBE-treated samples
826 (+58 CBEI/ +55 CBEII profile).
827 Targeted deep sequencing of the top 5 GUIDE-seq and *in* silico predicted off-
828 target in base-edited HSPCs showed a low sgRNA-dependent off-target activity,
829 occurring in non-exonic regions, with no predicted effect at protein level, and no
830 consequences at RNA level, as confirmed by targeted RNA-seq. In addition,
831 comparison between DBE and CBE revealed a higher specificity of DBE. Importantly,
832 no indels were detected at the off-target sites, suggesting that base editors minimize
833 the possibility of generating DSB-induced genomic rearrangements such as
834 translocations. Future WES and total RNA-seq analyses in base-edited HSPCs will be
835 performed to evaluate the potential sgRNA-independent DNA and RNA off-target
836 activity of our strategies.
837 Xenotransplantation experiments of multiplex base-edited HSPCs in
838 immunodeficient mice confirmed high editing efficiencies in long-term repopulating
839 HSCs and their progeny, demonstrating that multiplex base editing does not affect
840 engraftment and multilineage differentiation of HSCs. Of note, we achieved higher
841 editing frequency and HbF reactivation with DBE compared to CBE *in vivo*. Correlation
842 analysis between editing efficiency and HbF levels suggested that DBE-induced
843 multiple mutations in the *BCL11A* enhancers better evict GATA1 and ATF4 and down-
844 regulate *BCL11A in vivo*. These results also suggest that CBEs might be less efficient
845 in *bona fide* HSCs. In addition, CBEs might have induced some toxicity *in vivo* which
846 could have affect cell viability and self-renewal capacity of HSCs, as previously
847 proposed (Antoniou et al. 2022).
848 Of note, complete inactivation of *BCL11A,* adversely affects lymphoid
849 development and vital HSC functions (Liu et al., 2003; Yu et al., 2012; Guda et al.,
850 2015). However, *in vivo* experiments confirmed previous observations that the
851 downregulation of *BCL11A* achieved by targeting the GATA1 BS in the +58-kb
852 erythroid-specific enhancer does not impair *BCL11A* expression in other lineages
853 lineage (Bauer et al. 2013) . In addition, here we demonstrated that also the targeting
854 of ATF4 BS in the +55-kb enhancer led to specific downregulation of *BCL11A* in the
855 erythroid lineage without impacting its expression in B cells. Nonetheless, complete
856 knock out of *BCL11A* can also harm the erythroid lineage by affecting the human RBC
857 enucleation (Chang et al., 2017), but no alteration in the enucleation rate were
858 observed with our base editing strategy, likely because of the partial down-regulation
859 of *BCL11A* expression.
860 In conclusion, this study provides insights into the applicability of multiplexed
861 base editing strategies to treat SCD and potentially β-thalassemia patients, through
862 disruption of the erythroid-specific *BCL11A* enhancers in the +58-kb and +55-kb
863 regions. It is worth-noted that, our approaches can serve as universal therapeutic
864 strategies for both SCD and β-thalassemia patients, as they do not require the design
865 of mutation-specific CRISPR-Cas9-based tools, as proposed in previous studies (Zeng
866 et al. 2020; Newby et al. 2021; Hardouin et al. 2023). The levels of HbF achieved with
867 multiplex base editing were sufficient to ameliorate the sickling phenotype. The clinical
868 translation of our approach would require a full assessment of off-target activity of our

base editing systems to ensure the safety of our approaches, the establishment of a large-scale transfection protocol with clinical-grade reagents as well as biodistribution studies in mice.

907 **Figure legends**
908
909 **Figure 1**. **Base editing of the erythroid-specific *BCL11A* enhancers in SCD HPSC-**
910 **derived erythroblasts.**
911 **A**. Schematic representation of part of the *BCL11A* gene on chromosome 2, depicting
912 exons 2, 3 and 4 (Ex 2, 3 and 4) and the DNaseI hypersensitive sites (HS) +62-kb,
913 +58-kb and +55-kb. The sequences of the +58-kb and +55-kb regions of the erythroid-
914 specific *BCL11A* enhancers are depicted. GATA1 and ATF4 transcriptional activator
915 binding sites (BSs) are highlighted in green and bold. Target sgRNAs used in
916 combination with base editing enzymes are indicated with arrows and aligned with the
917 DNA sequence that they bind to in a stranded-oriented way. Green ovals indicate the
918 transcriptional activators GATA1 and ATF4.
919 **B**. Experimental protocol used for base editing experiments in SCD HSPCs. A BE
920 mRNA and a sgRNA were co-transfected in SCD HSPCs. Cells were differentiated into
921 mature RBCs using a three-phase erythroid differentiation protocol or plated in
922 methylcellulose-containing medium under conditions supporting erythroid (BFU-E) and
923 granulo-monocytic (CFU-GM) differentiation.
924 **C-D**. C-G to T-A or A-T to G-C base-editing efficiency calculated by the EditR software
925 in samples subjected to Sanger sequencing in erythroblasts derived from SCD HSPCs
926 edited in the +58-kb (**C**) or +55-kb (**D**) regions. Different editing profile were generated
927 through the combination of different BE and sgRNA. GATA_bs_1 sgRNA was used in
928 combination with CBE-SpRY or AncBE4max and ABEmax-SpRY to generate the +58
929 CBEI and +58 ABEI profile, respectively. GATA_bs_2 sgRNA was coupled with NG-
930 ABEmax and ABEmax-SpRY to generate +58ABEII profile while GATA_bs_3 sgRNA
931 was combined with ABEmax-SpRY generating the +58 ABEIII profile. CBE-SpRY and
932 AncBE4max coupled with ATF4_bs_1 or ATF4_bs_2 sgRNAs led to +55 CBEI and
933 +55 CBEII profiles, respectively. ATF4_bs_1 sgRNA or 55KB_ATF4_3 combined with
934 ABEmax-SpRY led to +55 ABEI profile and ATF4_bs_2 sgRNA combined with ABE8e-
935 SpRY led to +55 ABEII profile. In the results, we reported the highest editing efficiency
936 achieved in the editing window. Data are expressed as mean (n = 2 to 6 biologically
937 independent experiments, 2 to 5 donors).
938 **E-F**. Frequency of InDels, measured by TIDE analysis, in samples subjected to Sanger
939 sequencing in erythroblasts derived from SCD HSPCs edited in the +58-kb (**E**) or +55-
940 kb (**F**) regions. Data are expressed as mean ± SEM (n = 2 to 6 biologically independent
941 experiments, 2 to 5 donors). ****P≤ 0.0001 (One-way ANOVA. Comparison of controls
942 vs edited samples. No asterisk (*) = not significant.
943 **G**. CFC frequency for controls (untreated, or transfected with TE buffer, or transfected
944 with a BE mRNA only, or transfected with a BE mRNA and a sgRNA targeting the
945 unrelated *AAVS1* locus) and edited samples. Data are expressed as mean ± SEM (n
946 = 7 biologically independent experiments, 7 donors) (Two-way ANOVA with Dunnet
947 correction for multiple comparison; not significant).
948
949 **Figure 2. HbF reactivation and correction of the sickle phenotype in edited SCD**
950 **HPSC-derived erythroblasts.**
951 **A**. RT-qPCR analysis of γ (Gγ + Aγ)- and β^S-globin mRNA in control and edited SCD
952 patient erythroblasts at day 13 of erythroid differentiation. γ- and β^S-globin mRNA
953 expression was normalized to α-globin mRNA and expressed as a percentage of the
954 γ- + β^S- globin mRNA. The editing efficiency (EE) ± SEM is indicated for each sample
955 in the lower part of the panel. Data are expressed as mean ± SEM (n = 2 to 6
956 biologically independent experiments, 2 to 5 donors). *P ≤ 0.05; **P ≤ 0.01; ***P ≤

957 0.001; ****P ≤ 0.0001 (two-way ANOVA with Dunnett correction for multiple
958 comparisons. Comparison of mock vs edited samples).
959 **B**. Expression of γ (Gγ + Aγ)-globin chains measured by RP-HPLC in SCD RBCs
960 derived from control and edited SCD HSPCs. γ-globin expression was normalized to
961 α-globin. The EE ± SEM is indicated for each sample in the lower part of the panel.
962 Data are expressed as mean ± SEM (n = 2 to 6 biologically independent experiments,
963 2 to 5 donors). *P ≤ 0.01 (one-way ANOVA with Dunnett correction for multiple
964 comparisons. Comparison of mock vs edited samples).
965 **C.** Analysis of HbF and HbS expression measured by cation-exchange HPLC in RBCs
966 derived from control and edited SCD HSPCs. We calculated the percentage of each
967 Hb type over the total Hb tetramer. The EE ± SEM is indicated for each sample in the
968 lower part of the panel. Data are expressed as mean ± SEM (n = 2 to 6 biologically
969 independent experiments, 2 to 5 donors). *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤
970 0.0001 (two-way ANOVA with Dunnett correction for multiple comparisons.
971 Comparison of mock vs edited samples).
972 **D**. Representative flow cytometry histograms showing the percentage of HbF-
973 expressing cells (F-cells) and HbS-expressing cells (S-cells) in CD235+ population for
974 unstained (CD235 stained only), control and edited samples. Data expressed as mean
975 ± SEM of F-cells and S-cells (n = 2 to 4 biologically independent experiments, 2 to 5
976 donors).
977 **E**. Frequency of sickling cells upon O_2 deprivation in control and edited samples after
978 normalization to the mock. The EE ± SEM is indicated for each sample in the lower
979 part of the panel. Representative photomicrographs of SCD patient RBCs under
980 hypoxia conditions are shown. Data are expressed mean ± SEM (n = 2 to 6 biologically
981 independent experiments, 2 to 5 donors). *P ≤ 0.05; **P ≤ 0.01 (one-way ANOVA wit
982 Dunnett correction for multiple comparison, Comparison of ctrls vs edited samples).
983
984 **Figure 3**. **Gamma-globin reactivation in single BFU-E colonies derived from**
985 **edited SCD HSPCs.**
986 **A-B** Correlation between *HBG* mRNA relative expression and editing efficiency in
987 single BFU-E colonies differentiated from SCD HSPCs edited in the +58-kb (**A**) or +55-
988 kb (**B**) regions (1 donor). Different editing profiles were generated through the
989 combination of different BE and sgRNA. GATA_bs_1 sgRNA was used in combination
990 with CBE-SpRY or AncBE4max and ABEmax-SpRY or ABE8e-SpRY to generate the
991 +58 CBEI and +58 ABEI profile, respectively. GATA_bs_2 sgRNA was coupled with
992 NG-ABEmax, ABEmax-SpRY and NG-ABE8e to generate +58ABEII profile, while
993 GATA_bs_3 sgRNA was combined with ABEmax-SpRY generating the +58 ABEIII
994 profile. CBE-SpRY and AncBE4max coupled with ATF4_bs_1 or ATF4_bs_2 sgRNAs
995 led to +55 CBEI and +55 CBEII profiles, respectively. ATF4_bs_1 sgRNA or
996 55KB_ATF4_3 combined with ABEmax-SpRY led to +55 ABEI profile and ATF4_bs_2
997 sgRNA combined with ABE8e-SpRY and ABE8e-SpRY led to +55 ABEII profile. mRNA
998 expression was measured by RT-qPCR analysis of γ (Gγ + Aγ)- and β^S-globin mRNA
999 in control and edited SCD patient BFU-E. γ- and β^S-globin mRNA expression was
1000 normalized to α-globin mRNA and expressed as a percentage of the γ- + β^S- globin
1001 mRNA. Base-editing efficiency was calculated using EditR and Cas9-editing efficiency
1002 (InDels) was calculated using TIDE in samples subjected to Sanger sequencing. BFU-
1003 Es transfected with TE buffer, or transfected with a BE mRNA only, or transfected with
1004 a BE mRNA and a sgRNA targeting the unrelated AAVS1 locus were used as negative
1005 controls. *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001 (Multiple t test).
1006

1007 **Figure 4**. **Multiplex base editing of the erythroid-specific *BCL11A* enhancers in**
1008 **SCD HPSC-derived erythroblasts.**
1009 **A-B**. C-G to T-A or/and A-T to G-C base-editing efficiency in the +58-kb (**A**) or +55-kb
1010 (**B**) regions calculated by the EditR software in sample subjected to Sanger
1011 sequencing in erythroblasts derived from SCD HSPCs edited at either the +58-kb or
1012 the +55-kb region, or simultaneously edited at both the +58-kb and +55-kb regions.
1013 GATA_bs_2 or/and ATF4_bs_2 sgRNAs were combined with NG-ABE8e to generate
1014 +58ABEII, +55ABEII, and +58ABEII/+55ABEII profiles. GATA_bs_1 or/and
1015 ATF4_bs_2 sgRNAs combined with AncBE4max, generated the +58CBEI, +55CBEII,
1016 and +58CBEI/+55 CBEII profiles. Finally, GATA_bs_1 or/and ATF4_bs_2 sgRNAs
1017 combined with TadDE, generated the +58DBEI, +55DBEI, and +58DBEI/+55 DBEI
1018 profiles. Data are expressed as mean (n = 3 biologically independent experiments, 3
1019 donors).
1020 **C-D**. Frequency of InDels in the +58-kb (**C**) or +55-kb (**D**) regions measured by TIDE
1021 analysis, in samples subjected to Sanger sequencing in erythroblasts derived from
1022 SCD HSPCs edited at either the +58-kb or the +55-kb regions, or simultaneously
1023 edited at both the +58-kb and the +55-kb regions. Data are expressed as mean ± SEM
1024 (n = 3 biologically independent experiments, 3 donors. Comparison of ctrls vs edited
1025 samples). ****P ≤ 0.0001; *** ≤ 0.001 (One-way ANOVA. Comparison of controls vs
1026 edited samples)
1027 **E**. Schematic representation of the generation of a 3.2-kb deletion/inversion following
1028 the simultaneous targeting of GATA1 BS in the +58-kb region and ATF4 BS in the +55-
1029 kb region (highlighted with violet arrows). Black arrows indicate the location of BSs for
1030 additional transcriptional activators.
1031 **F**. Frequency of the 3.2-kb deletion/inversion, measured by ddPCR, for samples
1032 simultaneously edited at the +58-kb and +55-kb regions. Data are expressed as mean
1033 ± SEM (n = 3 biologically independent experiments, 3 donors). **P ≤ 0.01; ****P ≤
1034 0.0001 (two-way ANOVA with Tukey's correction for multiple comparison. Comparison
1035 of mock vs edited samples).
1036 **G**. CFC frequency for control and edited samples. Data are expressed as mean ± SEM
1037 (n = 2 biologically independent experiments, 2 donors) (two-way ANOVA with Dunnet
1038 correction for multiple comparison; not significant).
1039 **H**. ddPCR analysis of BCL11A expression in erythroblasts derived from edited SCD
1040 HSPCs at day 13 of erythroid differentiation. BCL11A mRNA expression was
1041 normalized to GAPDH mRNA. Data are expressed as mean ± SEM (n = 3 biologically
1042 independent experiment, 3 donors) *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001
1043 (one-way ANOVA with Dunnett correction for multiple comparison. Comparison of
1044 controls vs edited samples).
1045 **I**. RT-qPCR analysis of γ (Gγ + Aγ)- and β^S-globin mRNA in control and edited SCD
1046 erythroblast at day 13 of erythroid differentiation. γ- and β^S-globin mRNA expression
1047 was normalized to α-globin mRNA and expressed as a percentage of the γ- + β^S- globin
1048 mRNA. The EE ± SEM is indicated for each sample in the lower part of the panel. Data
1049 are expressed as mean ± SEM (n = 3 biologically independent experiment, 3 donors).
1050 *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001 (two-way ANOVA with Dunnett
1051 correction for multiple comparisons. Comparison of mock vs edited samples).
1052 **J**. Correlation between HbF expression and editing efficiency in erythroid samples
1053 derived from edited SCD HSPCs (erythroblasts and pool of BFU-E; n = 3 biological
1054 independent experiments, 3 donors). HbF expression was measured by cation-
1055 exchange HPLC and calculated over the total Hb tetramers. The background level
1056 measured in the control was subtracted from all the conditions. Base-editing efficiency

1057 was calculated using EditR software and Cas9-editing efficiency (InDels) was
1058 calculated using TIDE in samples subjected to Sanger sequencing. *P ≤ 0.05; **P ≤
1059 0.01; ***P ≤ 0.001; ****P ≤ 0.0001 (Multiple t test).
1060
1061 **Figure 5. Multiplex CBE and DBE base editing in SCD HPSC-derived**
1062 **erythroblasts.**
1063 **A-B**. C-G to T-A or/and A-T to G-C base-editing efficiency in the +58-kb (**A**) or +55-kb
1064 (**B**) regions calculated by EditR software in samples subjected to Sanger sequencing
1065 in erythroblasts derived from edited SCD HSPCs. GATA_bs_1 and ATF4_bs_2 were
1066 used in combination with AncBE4max to generate the +58 CBE/ +55 CBEII profile and
1067 in combination with TadDE to generate the +58 DBEI/ +55 DBEI profile. Data are
1068 expressed as mean (n = 1 biologically independent experiment, 5 donors).
1069 **C-D**. Frequency of InDels in the +58-kb (**C**) or +55-kb (**D**) regions measured by TIDE
1070 analysis, in samples subjected to Sanger sequencing in erythroblasts derived from
1071 edited SCD HSPCs. Data are expressed as mean ± SEM (n = 1 biologically
1072 independent experiment, 5 donors) (one-way ANOVA; not significant).
1073 **E**. Frequency of the 3.2-kb deletion/inversion, measured by ddPCR, in erythroblasts
1074 derived from edited SCD HSPCs. Data are expressed as mean ± SEM (n = 1
1075 biologically independent experiment, 5 donors). ****P ≤ 0.0001 (one-way ANOVA with
1076 Dunnett correction for multiple comparison. Comparison of controls vs edited
1077 samples).
1078 **F**. ddPCR analysis of BCL11A mRNA in erythroblasts derived from edited SCD HSPCs
1079 at day 13 of in vitro erythroid differentiation. BCL11A mRNA expression was
1080 normalized to GAPDH mRNA. Data are expressed as mean ± SEM (n = 1 biologically
1081 independent experiment, 5 donors) (one-way ANOVA; not significant).
1082 **G**. RT-qPCR analysis of γ (Gγ + Aγ)- and β^S-globin mRNA in control and edited SCD
1083 patient's erythroblasts at day 13 of erythroid differentiation. γ- and β^S-globin mRNA
1084 expression was normalized to α-globin mRNA and expressed as a percentage of the
1085 γ- + β^S- globin mRNA. Data are expressed as mean ± SEM (n = 1 biologically
1086 independent experiment, 5 donors). **P ≤ 0.01 (two-way ANOVA with Dunnett
1087 correction for multiple comparisons. Comparison of mock vs edited samples).
1088 **H**. Expression of γ (Gγ + Aγ)-globin chains measured by RP-HPLC in RBCs derived
1089 from SCD HSPCs. γ-globin expression was normalized to α-globin. Data are
1090 expressed as mean ± SEM (n=1 biologically independent experiment, 5 donors). **P
1091 ≤ 0.01 (one-way ANOVA with Dunnett correction for multiple comparisons.
1092 Comparison of mock vs edited samples).
1093 **I**. Analysis of HbF and HbS by cation-exchange HPLC in RBCs derived from edited
1094 SCD HSPCs. We calculated the percentage of each Hb type over the total Hb
1095 tetramers. Data are expressed as mean ± SEM (n = 1 biologically independent
1096 experiment, 5 donors). ****P ≤ 0.0001 for +58 CBEI/ +55 CBEII, ****P ≤ 0.0001 (two-
1097 way ANOVA with Dunnett correction for multiple comparisons. Comparison of mock vs
1098 edited samples).
1099 **J**. Representative flow cytometry histograms showing the percentage of HbF-
1100 expressing cells (F-cells) and HbS-expressing cells (S-cells) in CD235+ population for
1101 unstained (CD235 stained only), control and edited samples. Data expressed as mean
1102 ± SEM is reported (n = 1 biologically independent experiment, 5 donors). ***P ≤ 0.001;
1103 ****P ≤ 0.0001 (two-way ANOVA with Dunnett correction for multiple comparisons.
1104 Comparison of mock vs edited samples).
1105 **K**. Frequency of sickling cells upon O2 deprivation in mock and edited samples after
1106 normalization to the mock. Data are expressed mean ± SEM (n = 1 biologically

1107 independent experiment, 5 donors). ****P ≤ 0.0001 (one-way ANOVA with Dunnett
1108 correction for multiple comparisons. Comparison of mock vs edited samples).
1109
1110 **Figure 6**. **Evaluation of on-target aberrations and off-target activity of multiplex**
1111 **base editing of +58-kb and +55-kb regions.**
1112 **A-B**. Logo plots of the +58-kb (**A**) and +55-kb (**B**) target region obtained with the
1113 generation of +58 CBEI/+55 CBEII profile, +58 DBEI/+55 DBEII and +58/+55 Cas9
1114 profile. Editing efficiency in SCD HSPC-derived erythroblasts was evaluated by Sanger
1115 sequencing reaching 39% at GATA1 BS and ATF4 BS when generating the +58
1116 CBEI/+55 CBEII profile and 50% and 38% at GATA1 and ATF4 BS, respectively when
1117 generating the +58 DBEI/ +55 DBEI profile. 25% and 13% of InDels were obtained at
1118 GATA1 and ATF4 BS, respectively when using Cas9 nuclease.
1119 **C**. Comprehensive allele frequencies of diverse types of genomic rearrangements and
1120 InDel types in SCD HSPC-derived erythroblasts. We reported frequencies of events
1121 observed at individual GATA1 and ATF4 BSs (in the absence of the 3.2-kb
1122 deletion/inversion) and the frequencies of 3.2-kb deletion/inversion (GATA1-ATF4)..
1123 InDels were classified as small, intermediate and large according to their size ([4-
1124 50bp], [51-200bp], >200bp). Del, deletion; Ins, insertion; Inv, inversion. Unmodified, no
1125 deletion, inversion or insertion observed.
1126 **D.** We reported the frequency of alleles edited at the GATA1 and ATF4 BSs, either
1127 simultaneously (GATA1 and ATF4) or at only one of the sites (GATA1 or ATF4). The
1128 reported data were adjusted for background error typical of this analysis by subtracting
1129 the frequency of edited alleles in the mock condition from the fraction of edited alleles
1130 in the edited conditions.
1131 **E**. sgRNA-dependent off-target sites of the GATA_bs_1 (top panel) and ATF4_bs _2
1132 (low panel) sgRNAs in K562 cells, as evaluated by GUIDE-seq analysis. sgRNAs were
1133 coupled with a Cas9-nuclease corresponding to the Cas9 nickase included in base
1134 editor (Cas9-NGG). The protospacer targeted by each sgRNA and the PAM are
1135 reported on top of each panel, followed by the off-target sites and their mismatches
1136 with the on-target (highlighted in color). The number of sequencing reads (abundance),
1137 the chromosomal coordinates (hg38), the type of mapped site, the symbol of the gene,
1138 the DNA strand and the number of mismatches of each off-target site are reported.
1139 Cutoff was set to a level of detection of at least 3 reads. In the symbol column, a star
1140 (*) indicates that the site is within the transcription unit of the gene.
1141 **F**. Base editing and InDel frequency at on-target and off-target (OT) sites, for control,
1142 +58 CBEI/+ 55 CBEII and +58 DBEI/+55 DBEI profiles as measured by targeted NGS
1143 sequencing. Data are expressed as individual values and median (n=3 biologically
1144 independent experiments, 3 donors) *P≤0.05, ***P ≤ 0.001 and ****P≤0.0001 (two-way
1145 ANOVA with Dunnet's correction for multiple comparisons).
1146
1147 **Figure 7. Simultaneous base editing of BCL11A enhancers in repopulating**
1148 **HSCs.**
1149 **A**. Experimental protocol used for HD HSPC xenotransplantation experiments. BE
1150 mRNAs and sgRNAs were co-transfected in HD HSPCs, and control and edited cells
1151 were xenotransplanted into NBSGW immunodeficient mice 1 day after transfection.
1152 Mice were euthanized 16/17 weeks after transplantation and their hematopoietic
1153 tissues and organs were collected and analyzed.
1154 **B**. Human hematopoietic progenitor content in BM human CD45+ cells derived from
1155 mice transplanted with control and edited HSPCs. We plotted the percentage of human

1156 CD45+ cells giving rise to BFU-E and CFU-GM. Data are expressed as mean ± SEM
1157 (two-way ANOVA; not significant).
1158 **C-F**. C-G to T-A or/and A-T to G-C base-editing efficiency, calculated using EditR for
1159 +58 CBEI/ +55 CBEII (**C and D**) and +58 DBEI/ +55 DBEI (**E and F**) profiles at the
1160 +58-kb (**C and E**) and +55-kb (**D and F**) regions in input, blood-, bone marrow (BM)-,
1161 thymus-, spleen-derived human samples subjected to Sanger sequencing. BM
1162 samples include CD45+ cells, B cells and erythroid cells (Ery). Data are expressed as
1163 mean ± SEM [n = 3 biologically independent experiments (Input), n = 7 mice per group].
1164 Each data point represents an individual mouse. The frequency of base editing in input
1165 was calculated in cells cultured in the HSPC medium, in liquid erythroid cultures and
1166 in pool of BFU-E.
1167 **G**. Analysis of HbF and HbS by cation-exchange HPLC in human CD235+ BM-sorted
1168 erythroid cells in mice transplanted with control and edited HSPCs, 16/17 weeks post-
1169 transplantation. We calculated the percentage of each Hb type over the total Hb
1170 tetramers. Data are expressed as mean ± SEM (n = 7 mice per group). Each data point
1171 represents an individual mouse. ****P ≤ 0.0001 (two-way ANOVA with Dunnett
1172 correction for multiple comparisons. Comparison of mock vs edited samples).
1173 **H-I**. RT-qPCR analysis of BCL11A in BM-sorted CD235+ (**I**) or CD19+ (**J**) cells in mice
1174 transplanted with control and edited HSPCs, 16/17 weeks post-transplantation.
1175 BCL11A mRNA expression was normalized to GAPDH mRNA. Data are expressed as
1176 mean ± SEM (n = 5 to 7 mice per group) Each data point represents an individual
1177 mouse. **P ≤ 0.01; ***P ≤ 0.001 (One-way ANOVA with Tukey's correction for multiple
1178 comparisons. Comparison of mock vs edited samples).
1179 **J**. Engraftment of human cells in NBSGW mice transplanted with control (mock-
1180 transfected) or edited HSPCs 16/17 weeks post-transplantation (n = 7 mice per group).
1181 Engraftment is represented as the percentage of human CD45+ cells in the total
1182 murine and human CD45+ cell population in BM, spleen, thymus, and peripheral blood.
1183 Each data point represents an individual mouse. Data are expressed as mean ± SEM.
1184 (two-way ANOVA; not significant).
1185
1186 **Supplementary Figure legends**
1187
1188 **Supplementary Figure 1. Base editing of the erythroid-specific *BCL11A***
1189 **enhancers in SCD HPSC-derived BFU-E and CFU-GM colonies.**
1190 **A-B**. C-G to T-A or A-T to G-C base-editing efficiency calculated by the EditR software
1191 in samples subjected to Sanger sequencing in pool of BFU-E colonies derived from
1192 SCD HSPCs edited in the +58-kb (**A**) or +55-kb (**B**) regions. Data are expressed as
1193 mean (n = 2 to 4 biologically independent experiments, 2 to 5 donors).
1194 **C-D**. Frequency of InDels, measured by TIDE analysis, in samples subjected to Sanger
1195 sequencing in pool of BFU-E colonies derived from SCD HSPCs edited in the +58-kb
1196 (**C**) or +55-kb (**D**) regions. Data are expressed as mean ± SEM (n = 2 to 4 biologically
1197 independent experiments, 2 to 5 donors). ****P≤ 0.0001 (One-way ANOVA.
1198 Comparison of controls vs edited samples)
1199 **E-F**. C-G to T-A or A-T to G-C base-editing efficiency calculated by the EditR software
1200 in samples subjected to Sanger sequencing in pool of CFU-GM colonies derived from
1201 SCD HSPCs edited in the +58-kb (**E**) or +55-kb (**F**) regions. Data are expressed as
1202 mean (n = 2 to 4 biologically independent experiments, 2 to 5 donors).
1203 **G-H**. Frequency of InDels, measured by TIDE analysis, in samples subjected to
1204 Sanger sequencing in pool of CFU-GM colonies derived from SCD HSPCs edited in
1205 the +58-kb (**G**) or +55-kb (**H**) regions. Data are expressed as mean ± SEM (n = 2 to 4

biologically independent experiments, 2 to 5 donors). ****P≤ 0.0001 (One-way ANOVA. Comparison of controls vs edited samples)

I. Analysis of HbF and HbS by cation-exchange HPLC in pool of BFU-E colonies. We calculated the percentage of each Hb type over the total Hb tetramers. Data are expressed as mean ± SEM (n = 3 to 4 biologically independent experiments, 3 to 4 donors). *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001 (two-way ANOVA with Dunnett correction for multiple comparisons. Comparison of mock vs edited samples).

Supplementary Figure 2. Erythroid differentiation of SCD HSPCs upon base editing of the erythroid-specific *BCL11A* enhancers.

A. Frequency of enucleated cells at day 6, 13, 16 and 19 of erythroid differentiation, as measured by flow cytometry analysis of DRAQ5 nuclear staining in control and edited samples. Data are expressed as mean ± SEM (n = 2 to 6 biologically independent experiments, 2 to 5 donors). Representative flow cytometry histograms showing the DRAQ5- cell population for control and edited samples are reported.

B-D. Frequency of $CD36^+$ (**B**), $CD71^+$ (**C**) and $CD235^+$ (**D**) cells at day 6, 13 and 19 of erythroid differentiation, as measured by flow cytometry analysis of CD36, CD71 and CD235 erythroid markers. Data are expressed as mean ± SEM (n = 2 to 6 biologically independent experiments, 2 to 5 donors). Representative flow cytometry histograms showing the $CD36^+$ (**B**), $CD71^+$ (**C**) and $CD235^+$ (**D**) cell population for control and edited samples are reported.

E. Frequency of $CD49d^+$, $BAND3^+$ and $CD49d^+/BAND3^+$ in $7AAD^-/CD235^+$ cells at day 6, 13 and 19 of erythroid differentiation, as measured by flow cytometry analysis of CD49d and BAND3 erythroid markers. Data are expressed as mean ± SEM (n = 2 to 6 biologically independent experiments, 2 to 5 donors). Representative flow cytometry contour plots showing the $CD49d^+$, $BAND3^+$ and $CD49d^+/BAND3^+$ cell population for control and edited samples are reported.

Supplementary Figure 3. Highly processive ABEs efficiently create the +58 ABEII profile.

A. A-T to G-C base-editing efficiency calculated by the EditR software in samples subjected to Sanger sequencing in pool of BFU-E derived from SCD HSPCs edited in the +58-kb region. Data are expressed as mean ± SEM (n = 2 to 3 biologically independent experiments, 1 to 4 donors).

B. Frequency of InDels, measured by TIDE analysis, in samples subjected to Sanger sequencing in pool of BFU-E colonies derived from SCD HSPCs edited in the +58-kb. The GATA_bs_2 sgRNA was combined with NG-ABE-max and NG-ABE8e. Data are expressed as mean ± SEM (n = 2 to 3 biologically independent experiments, 2 to 4 donors). **P ≤ 0.01 (One-way ANOVA comparison. Comparison of controls vs edited samples).

Supplementary Figure 4. **Multiplex base editing of the erythroid-specific *BCL11A* enhancers in pool of BFU-E.**

A-B. C-G to T-A or/and A-T to G-C base-editing efficiency in the +58-kb (**A**) or +55-kb (**B**) regions calculated using EditR in samples subjected to Sanger sequencing in BFU-E colonies derived from SCD HSPCs edited at either the +58-kb or the +55-kb region, or simultaneously edited at both the +58-kb and the +55-kb regions. Data are expressed as mean (n = 2 biologically independent experiments, 2 donors).

C-D. Frequency of InDels in the +58-kb (**C**) or +55-kb (**D**) regions measured by TIDE analysis, in samples subjected to Sanger sequencing in pool BFU-E colonies

differentiated from SCD HSPCs edited at either the +58-kb or the +55-kb regions, or simultaneously edited at both the +58-kb and the +55-kb regions. Data are expressed as mean ± SEM (n = 2 biologically independent experiments, 2 donors). **P ≤ 0.01; ****P ≤ 0.0001 (one-way ANOVA with Dunnett correction for multiple comparisons. Comparison of controls and edited samples).

E. Frequency of the 3.2-kb deletion/inversion, measured by ddPCR, in samples simultaneously edited at the +58-kb and +55-kb regions. Data are expressed as mean ± SEM (n = 2 biologically independent experiments, 2 donors). **P ≤ 0.01; ****P ≤ 0.0001 (two-way ANOVA with Sidak correction for multiple comparison. Comparison of mock vs edited samples).

F. Analysis of HbF and HbS by cation-exchange HPLC in pool of BFU-E colonies derived from edited SCD HSPCs. We calculated the percentage of each Hb type over the total Hb tetramers. Data are expressed as mean ± SEM (n = 2 biologically independent experiments, 2 donors). *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001 (two-way ANOVA with Dunnett correction for multiple comparisons. Comparison mock vs edited samples).

Supplementary Figure 5. HbF reactivation in RBCs derived from edited SCD HSPCs.

A. Expression of γ (Gγ + Aγ)-globin chains measured by RP-HPLC in RBCs derived from SCD HSPCs. γ-globin expression was normalized to α-globin. The EE ± SEM is indicated for each sample in the lower part of the panel. Data are expressed as mean ± SEM (n=3 biologically independent experiments, 3 donors). *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001 (one-way ANOVA with Dunnett correction for multiple comparisons. Comparison mock vs edited samples). **B**. Analysis of HbF and HbS by cation-exchange HPLC in RBCs derived from edited SCD HSPCs. We calculated the percentage of each Hb type over the total Hb tetramers. Data are expressed as mean ± SEM (n = 3 biologically independent experiments, 3 donors). *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001 (two-way ANOVA with Dunnett correction for multiple comparisons. Comparison mock vs edited samples).

Supplementary Figure 6. **Correlation between HbF expression and editing efficiency obtained with ABE, CBE, DBE and Cas9 nuclease.**

Correlation between HbF expression and editing efficiency in erythroid cells derived from SCD HSPCs edited with either ABE (**A**), CBE (**B**), DBE (**C**) or Cas9-nuclease (**D**) (erythroblasts and pool of BFU-E; n = 3 biological independent experiments, 3 donors). HbF expression was measured by cation-exchange HPLC and calculated over the total Hb tetramers. Base-editing efficiency was calculated using EditR and Cas9-editing efficiency (InDels) was calculated using TIDE in samples subjected to Sanger sequencing. **P* ≤ 0.05; ***P* ≤ 0.01; ****P* ≤ 0.001; *****P* ≤ 0.0001 (Multiple t test).

Supplementary Figure 7. **Erythroid differentiation of SCD HSPCs upon multiplex base editing of the +58-kb and +55-kb regions.**

A. Frequency of enucleated cells at day 6, 13, 16 and 20 of erythroid differentiation, as measured by flow cytometry analysis of DRAQ5 nuclear staining in control and edited samples. Data are expressed as mean ± SEM (n = 5 donors).

B-D. Frequency of $CD36^+$ (**B**), $CD71^+$ (**C**) and $CD235^+$ (**D**) cells at day 6, 13 and 20 of erythroid differentiation, as measured by flow cytometry analysis of CD36, CD71 and CD235 erythroid markers. Data are expressed as mean ± SEM (n = 1 biologically independent experiment, 5 donors).

E. Frequency of CD49d$^+$, BAND3$^+$ and CD49d$^+$/BAND3$^+$ in 7AAD$^-$/CD235$^+$ cells at day 6, 13 and 20 of erythroid differentiation, as measured by flow cytometry analysis of CD49d and BAND3 erythroid markers. Data are expressed as mean ± SEM (n = 1 biologically independent experiment, 5 donors).

Supplementary Figure 8. **NGS analysis of on-target editing in erythroid cells derived from SCD HSPCs upon multiplex base editing of the +58-kb and +55-kb regions.**

A-D. Representative frequency and sequence of modified and unmodified alleles in edited SCD samples, for +58 CBEI/ +55 CBEII (**A and B**) and +58 DBEI/ +55 DBEI (**C and D**) profiles at the +58-kb (**A and C**) and +55-kb (**B and D**) regions, as measured by targeted NGS sequencing. Target base positions are highlighted with a blue box. Red box indicates inserted bases. Grey squares indicate deletions.

E. InDel frequency in edited SCD samples, for +58 CBEI/ +55 CBEII and +58 DBEI/ +55 DBEI profiles at the +58-kb (left panel) and +55-kb (right panel) regions, as measured by targeted NGS sequencing.

Supplementary Figure 9. Simultaneous base editing of *BCL11A* enhancers in repopulating HSCs.

A. Frequency of human T (CD3) and B (CD19) lymphoid, myeloid (CD11b, CD14, and CD15), HSPC (CD34), and erythroid (CD235, CD36 and CD71) cells in BM in mice transplanted with control and edited HSPCs, 16/17 weeks after the transplantation (n = 7 mice per condition). Each data point represents an individual mouse. Data are expressed as mean ± SEM. (two-way ANOVA; not significant).

B. Frequency of InDels calculated by TIDE for the +58 CBEI/+55 CBEII and +58 DBEI/+55 DBEI profiles at the +58-kb and +55-kb regions in input, blood-, BM-, thymus-, and spleen-derived human samples subjected to Sanger sequencing. BM samples include CD45+ cells, B cells and erythroid cells (Ery). Data are expressed as mean ± SEM [n = 3 biologically independent experiments (Input), n = 4 to 7 mice per group]. Each data point represents an individual mouse (two-way ANOVA; not significant). The frequency of Indels in input was calculated in cells cultured in the HSPC medium, in liquid erythroid cultures and in pool BFU-E.

C. RT-qPCR analysis of γ (^{G}γ + ^{A}γ)- and β-globin mRNA in BM sorted human CD235$^+$ erythroid cells in mice transplanted with control and edited HSPCs, 16/17 weeks post-transplantation. γ- and β^S-globin mRNA expression was normalized to α-globin mRNA and expressed as a percentage of the γ- + β^S- globin mRNA. Each data point represents an individual mouse. Data are expressed as mean ± SEM (n = 7 mice per group). ****P ≤ 0.0001 (two-way ANOVA with Dunnett correction for multiple comparisons. Comparison mock vs edited samples).

D. Expression of γ (^{G}γ + ^{A}γ)-globin chains measured by RP-HPLC in BM sorted human CD235$^+$ erythroid cells in mice transplanted with control and edited HSPCs, 16/17 weeks post-transplantation. γ-globin expression was normalized to α-globin. Each data point represents an individual mouse. Data are expressed as mean ± SEM (n = 7 mice per group). ****P ≤ 0.0001 (one-way ANOVA with Dunnett correction for multiple comparisons. Comparison mock vs edited samples).

E. Correlation between HbF expression and editing efficiency in human CD235+ BM-sorted erythroid cells obtained from mice transplanted with control and edited HSPCs. HbF was measured by cation-exchange HPLC and calculated over the total Hb tetramers. Base-editing efficiency was calculated using EditR software. ****P ≤ 0.0001 (Multiple t test).

Supplementary Figure 10. Simultaneous base editing of *BCL11A* enhancers in erythroid and granulo-monocytic progenitors derived from engrafting HSPCs

A-D. C-G to T-A or/and A-T to G-C base-editing efficiency, calculated using EditR for the +58 CBEI/ +55 CBEII (**A and B**) and +58 DBEI/ +55 DBEI (**C and D**) profiles at the +58-kb (**A and C**) and +55-kb (**B and D**) regions in input and pooled BFU-E and CFU-GM derived from BM CD45+ cells and subjected to Sanger sequencing. Data are expressed as mean ± SEM [n = 3 biologically independent experiments (Input), n = 7 mice per group]. Each data point represents an individual mouse. The frequency of base editing in input was calculated in cells cultured in the HSPC medium, in liquid erythroid cultures and in pool of BFU-E.

E. RT-qPCR analysis of *BCL11A-XL* expression in pooled BFU-E derived from BM CD45+ cells. *BCL11A-XL* mRNA expression was normalized to *GAPDH*. Data are expressed as mean ± SEM (n = 7 mice per group). Each data point represents an individual mouse. *P ≤ 0.05 (One-way ANOVA with Dunnett's correction for multiple comparison. Comparison of mock vs edited samples).

F. RT-qPCR analysis of γ (^{G}γ + ^{A}γ)- and β-globin mRNA in pooled BFU-E derived from BM CD45+ cells. γ- and β^{-}globin mRNA expression was normalized to α-globin mRNA and expressed as a percentage of the γ- + β- globin mRNA. Each data point represents an individual mouse. Data are expressed as mean ± SEM (n = 7 mice per group). ****P ≤ 0.0001 (two-way ANOVA with Dunnett)

G Expression of γ (^{G}γ + ^{A}γ)-globin chains measured by RP-HPLC in pooled BFU-E derived from BM CD45+ cells. γ-globin expression was normalized to α-globin. Each data point represents an individual mouse. Data are expressed as mean ± SEM (n = 7 mice per group). ****P ≤ 0.0001 (one-way ANOVA with Dunnett correction for multiple comparisons. Comparison mock vs edited samples).

H. Analysis of HbF and HbS by CE-HPLC in pooled BFU-E derived from BM CD45+ cells. We calculated the percentage of each Hb type over the total Hb tetramers. Data are expressed as mean ± SEM (n = 7 mice per group). Each data point represents an individual mouse. ****P ≤ 0.0001 (two-way ANOVA with Dunnett correction for multiple comparisons. Comparison of mock vs edited samples).

Acknowledgements

This work was supported by state funding from the French National Research Agency (Agence Nationale de la Recherche) as part of the Investissements d'Avenir program (ANR-10-IAHU-01) and by the European Research Council (865797 DITSB), the European Commission (HORIZON-RIA EDITSCD grant no. 101057659), the Fondation pour la Recherche Médicale [FRM PLP202110014595] and the COST (European Cooperation in Science and Technology (the COST Action Gene Editing for the treatment of Human Diseases, CA21113).

We thank Carine Giovannangeli for the production of Cas9 nuclease RNP and all the patients for their contribution to this work.

Author contributions

L.F. designed, conducted experiments, analyzed the data, and wrote the paper. P.M., S.A., T.F., M.M., G.C., A.T., A.C., G.H., M.J. conducted experiments and analyzed data. O.R. analyzed NGS data. M.A. contributed to the design of the experimental strategy. P.A. conceived the study, designed, conducted experiments, analyzed the data, and wrote the paper. A.M. conceived the study, designed experiments, and wrote the paper.

Competing interests

P.A. and A. M. are the inventors of a patent describing base editing approaches for hemoglobinopathies (PCT/EP2022/083904: Methods for increasing fetal hemoglobin content by editing the +55-kb region of the erythroid-specific bcl11a enhancer). Panagiotis Antoniou is an employee of Novo Nordisk. The remaining authors declare no competing interests.

References

Antoniou, Panagiotis, Giulia Hardouin, Pierre Martinucci, Giacomo Frati, Tristan Felix, Anne Chalumeau, Letizia Fontana, et al. 2022. «Base-Editing-Mediated Dissection of a γ-Globin Cis-Regulatory Element for the Therapeutic Reactivation of Fetal Hemoglobin Expression». *Nature Communications* 13 (1): 6618. https://doi.org/10.1038/s41467-022-34493-1.

Bauer, Daniel E., Sophia C. Kamran, Samuel Lessard, Jian Xu, Yuko Fujiwara, Carrie Lin, Zhen Shao, et al. 2013. «An Erythroid Enhancer of *BCL11A* Subject to Genetic Variation Determines Fetal Hemoglobin Level». *Science* 342 (6155): 253–57. https://doi.org/10.1126/science.1242088.

Behera, Vivek, Perry Evans, Carolyne J. Face, Nicole Hamagami, Laavanya Sankaranarayanan, Cheryl A. Keller, Belinda Giardine, et al. 2018. «Exploiting Genetic Variation to Uncover Rules of Transcription Factor Binding and Chromatin Accessibility». *Nature Communications* 9 (1): 782. https://doi.org/10.1038/s41467-018-03082-6.

Blattner, Georges, Alessia Cavazza, Adrian J. Thrasher, e Giandomenico Turchiano. 2020. «Gene Editing and Genotoxicity: Targeting the Off-Targets». *Frontiers in Genome Editing* 2 (dicembre):613252. https://doi.org/10.3389/fgeed.2020.613252.

Brinkman, Eva K., Tao Chen, Mario Amendola, e Bas van Steensel. 2014. «Easy Quantitative Assessment of Genome Editing by Sequence Trace Decomposition». *Nucleic Acids Research* 42 (22): e168–e168. https://doi.org/10.1093/nar/gku936.

Clement, Kendell, Holly Rees, Matthew C. Canver, Jason M. Gehrke, Rick Farouni, Jonathan Y. Hsu, Mitchel A. Cole, et al. 2019. «CRISPResso2 Provides Accurate and Rapid Genome Editing Sequence Analysis». *Nature Biotechnology* 37 (3): 224–26. https://doi.org/10.1038/s41587-019-0032-3.

Cullot, Grégoire, Julian Boutin, Jérôme Toutain, Florence Prat, Perrine Pennamen, Caroline Rooryck, Martin Teichmann, et al. 2019. «CRISPR-Cas9 Genome Editing Induces Megabase-Scale Chromosomal Truncations». *Nature Communications* 10 (1): 1136. https://doi.org/10.1038/s41467-019-09006-2.

Demirci, Selami, Alexis Leonard, Khaled Essawi, e John F. Tisdale. 2021. «CRISPR-Cas9 to Induce Fetal Hemoglobin for the Treatment of Sickle Cell Disease». *Molecular Therapy - Methods & Clinical Development* 23 (dicembre):276–85. https://doi.org/10.1016/j.omtm.2021.09.010.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, e Thomas R. Gingeras. 2013. «STAR: Ultrafast Universal RNA-Seq Aligner». *Bioinformatics* 29 (1): 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Fernandes, Queenie. 2017. «Therapeutic Strategies in Sickle Cell Anemia: The Past Present and Future». *Life Sciences* 178 (giugno):100–108. https://doi.org/10.1016/j.lfs.2017.03.025.

Fiumara, Martina, Samuele Ferrari, Attya Omer-Javed, Stefano Beretta, Luisa Albano, Daniele Canarutto, Angelica Varesi, et al. 2023. «Genotoxic Effects of Base and Prime Editing in Human Hematopoietic Stem Cells». *Nature Biotechnology*, settembre. https://doi.org/10.1038/s41587-023-01915-4.

Frangoul, Haydar, Franco Locatelli, Akshay Sharma, Monica Bhatia, Markus Mapara, Lyndsay Molinari, Donna Wall, et al. 2024. «Exagamglogene Autotemcel for Severe Sickle Cell Disease». *New England Journal of Medicine* 390 (18): 1649–62. https://doi.org/10.1056/NEJMoa2309676.

Giarratana, Marie-Catherine, Ladan Kobari, Hélène Lapillonne, David Chalmers, Laurent Kiger, Thérèse Cynober, Michael C Marden, Henri Wajcman, e Luc Douay. 2005. «Ex Vivo Generation of Fully Mature Human Red Blood Cells from Hematopoietic Stem Cells». *Nature Biotechnology* 23 (1): 69–74. https://doi.org/10.1038/nbt1047.

Hardouin, Giulia, Panagiotis Antoniou, Pierre Martinucci, Tristan Felix, Sandra Manceau, Laure Joseph, Cécile Masson, et al. 2023. «Adenine Base Editor–Mediated Correction of the Common and Severe IVS1-110 (G>A) β-Thalassemia Mutation». *Blood* 141 (10): 1169–79. https://doi.org/10.1182/blood.2022016629.

Huang, Peng, Scott A. Peslak, Xianjiang Lan, Eugene Khandros, Jennifer A. Yano, Malini Sharma, Cheryl A. Keller, et al. 2020. «The HRI-Regulated Transcription Factor ATF4 Activates BCL11A Transcription to Silence Fetal Hemoglobin Expression». *Blood* 135 (24): 2121–32. https://doi.org/10.1182/blood.2020005301.

Huang, Tony P., Zachary J. Heins, Shannon M. Miller, Brandon G. Wong, Pallavi A. Balivada, Tina Wang, Ahmad S. Khalil, e David R. Liu. 2023. «High-Throughput Continuous Evolution of Compact Cas9 Variants Targeting Single-Nucleotide-Pyrimidine PAMs». *Nature Biotechnology* 41 (1): 96–107. https://doi.org/10.1038/s41587-022-01410-2.

Kato, Gregory J., Frédéric B. Piel, Clarice D. Reid, Marilyn H. Gaston, Kwaku Ohene-Frempong, Lakshmanan Krishnamurti, Wally R. Smith, et al. 2018. «Sickle Cell Disease». *Nature Reviews. Disease Primers* 4 (marzo):18010. https://doi.org/10.1038/nrdp.2018.10.

Kluesner, Mitchell G., Derek A. Nedveck, Walker S. Lahr, John R. Garbe, Juan E. Abrahante, Beau R. Webber, e Branden S. Moriarity. 2018. «EditR: A Method to Quantify Base Editing from Sanger Sequencing». *The CRISPR Journal* 1 (3): 239–50. https://doi.org/10.1089/crispr.2018.0014.

Koblan, Luke W, Jordan L Doman, Christopher Wilson, Jonathan M Levy, Tristan Tay, Gregory A Newby, Juan Pablo Maianti, Aditya Raguram, e David R Liu. 2018. «Improving Cytidine and Adenine Base Editors by Expression Optimization and Ancestral Reconstruction». *Nature Biotechnology* 36 (9): 843–46. https://doi.org/10.1038/nbt.4172.

Kosicki, Michael, Kärt Tomberg, e Allan Bradley. 2018. «Repair of Double-Strand Breaks Induced by CRISPR–Cas9 Leads to Large Deletions and Complex Rearrangements». *Nature Biotechnology* 36 (8): 765–71. https://doi.org/10.1038/nbt.4192.

Leibowitz, Mitchell L., Stamatis Papathanasiou, Phillip A. Doerfler, Logan J. Blaine, Lili Sun, Yu Yao, Cheng-Zhong Zhang, Mitchell J. Weiss, e David Pellman. 2021. «Chromothripsis as an On-Target Consequence of CRISPR–Cas9 Genome Editing». *Nature Genetics* 53 (6): 895–905. https://doi.org/10.1038/s41588-021-00838-7.

Li, Chang, Aphrodite Georgakopoulou, Arpit Mishra, Sucheol Gil, R. David Hawkins, Evangelia Yannaki, e André Lieber. 2021. «In Vivo HSPC Gene Therapy with Base Editors Allows for Efficient Reactivation of Fetal γ-Globin in β-YAC Mice». *Blood Advances* 5 (4): 1122–35. https://doi.org/10.1182/bloodadvances.2020003702.

Li, Lushen, e Pankaj K. Mandal. 2024. «Recent advancements in gene therapy for sickle cell disease and β-thalassemia». *Frontiers in Hematology* 3 (settembre):1468952. https://doi.org/10.3389/frhem.2024.1468952.

Liao, Jiaoyang, Shuanghong Chen, Shenlin Hsiao, Yanhong Jiang, Yang Yang, Yuanjin Zhang, Xin Wang, Yongrong Lai, Daniel E. Bauer, e Yuxuan Wu.

2023. «Therapeutic Adenine Base Editing of Human Hematopoietic Stem Cells». *Nature Communications* 14 (1): 207. https://doi.org/10.1038/s41467-022-35508-7.

Liao, Yang, Gordon K. Smyth, e Wei Shi. 2014. «featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features». *Bioinformatics* 30 (7): 923–30. https://doi.org/10.1093/bioinformatics/btt656.

Locatelli, Franco, Peter Lang, Donna Wall, Roland Meisel, Selim Corbacioglu, Amanda M. Li, Josu De La Fuente, et al. 2024. «Exagamglogene Autotemcel for Transfusion-Dependent β-Thalassemia». *New England Journal of Medicine* 390 (18): 1663–76. https://doi.org/10.1056/NEJMoa2309673.

Métais, Jean-Yves, Phillip A. Doerfler, Thiyagaraj Mayuranathan, Daniel E. Bauer, Stephanie C. Fowler, Matthew M. Hsieh, Varun Katta, et al. 2019. «Genome Editing of HBG1 and HBG2 to Induce Fetal Hemoglobin». *Blood Advances* 3 (21): 3379–92. https://doi.org/10.1182/bloodadvances.2019000820.

Neugebauer, Monica E., Alvin Hsu, Mandana Arbab, Nicholas A. Krasnow, Amber N. McElroy, Smriti Pandey, Jordan L. Doman, et al. 2023. «Evolution of an Adenine Base Editor into a Small, Efficient Cytosine Base Editor with Low off-Target Activity». *Nature Biotechnology* 41 (5): 673–85. https://doi.org/10.1038/s41587-022-01533-6.

Newby, Gregory A., Jonathan S. Yen, Kaitly J. Woodard, Thiyagaraj Mayuranathan, Cicera R. Lazzarotto, Yichao Li, Heather Sheppard-Tillman, et al. 2021. «Base Editing of Haematopoietic Stem Cells Rescues Sickle Cell Disease in Mice». *Nature* 595 (7866): 295–302. https://doi.org/10.1038/s41586-021-03609-w.

Robinson, Mark D., Davis J. McCarthy, e Gordon K. Smyth. 2010. «edgeR : A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data». *Bioinformatics* 26 (1): 139–40. https://doi.org/10.1093/bioinformatics/btp616.

Schiroli, Giulia, Anastasia Conti, Samuele Ferrari, Lucrezia Della Volpe, Aurelien Jacob, Luisa Albano, Stefano Beretta, et al. 2019. «Precise Gene Editing Preserves Hematopoietic Stem Cell Function Following Transient P53-Mediated DNA Damage Response». *Cell Stem Cell* 24 (4): 551-565.e8. https://doi.org/10.1016/j.stem.2019.02.019.

Traxler, Elizabeth A, Yu Yao, Yong-Dong Wang, Kaitly J Woodard, Ryo Kurita, Yukio Nakamura, Jim R Hughes, et al. 2016. «A Genome-Editing Strategy to Treat β-Hemoglobinopathies That Recapitulates a Mutation Associated with a Benign Genetic Condition». *Nature Medicine* 22 (9): 987–90. https://doi.org/10.1038/nm.4170.

Tsai, Shengdar Q, Ved V Topkar, J Keith Joung, e Martin J Aryee. 2016. «Open-Source Guideseq Software for Analysis of GUIDE-Seq Data». *Nature Biotechnology* 34 (5): 483–483. https://doi.org/10.1038/nbt.3534.

Tsai, Shengdar Q., Zongli Zheng, Nhu T. Nguyen, Matthew Liebers, Ved V. Topkar, Vishal Thapar, Nicolas Wyvekens, et al. 2015. «GUIDE-Seq Enables Genome-Wide Profiling of off-Target Cleavage by CRISPR-Cas Nucleases». *Nature Biotechnology* 33 (2): 187–97. https://doi.org/10.1038/nbt.3117.

Weber, Leslie, Giacomo Frati, Tristan Felix, Giulia Hardouin, Antonio Casini, Clara Wollenschlaeger, Vasco Meneghini, et al. 2020. «Editing a γ-Globin Repressor Binding Site Restores Fetal Hemoglobin Synthesis and Corrects the Sickle Cell Disease Phenotype». *Science Advances* 6 (7): eaay9392. https://doi.org/10.1126/sciadv.aay9392.

Wu, Tianzhi, Erqiang Hu, Shuangbin Xu, Meijun Chen, Pingfan Guo, Zehan Dai, Tingze Feng, et al. 2021. «clusterProfiler 4.0: A Universal Enrichment Tool for

Interpreting Omics Data». *The Innovation* 2 (3): 100141. https://doi.org/10.1016/j.xinn.2021.100141.

Wu, Yuxuan, Jing Zeng, Benjamin P. Roscoe, Pengpeng Liu, Qiuming Yao, Cicera R. Lazzarotto, Kendell Clement, et al. 2019. «Highly Efficient Therapeutic Gene Editing of Human Hematopoietic Stem Cells». *Nature Medicine* 25 (5): 776–83. https://doi.org/10.1038/s41591-019-0401-y.

Zeng, Jing, My Anh Nguyen, Pengpeng Liu, Lucas Ferreira Da Silva, Linda Y. Lin, David G. Justus, Karl Petri, et al. 2023. «Gene Editing without *Ex Vivo* Culture Evades Genotoxicity in Human Hematopoietic Stem Cells». https://doi.org/10.1101/2023.05.27.542323.

Zeng, Jing, Yuxuan Wu, Chunyan Ren, Jasmine Bonanno, Anne H. Shen, Devlin Shea, Jason M. Gehrke, et al. 2020. «Therapeutic Base Editing of Human Hematopoietic Stem Cells». *Nature Medicine* 26 (4): 535–41. https://doi.org/10.1038/s41591-020-0790-y.

Figure 1

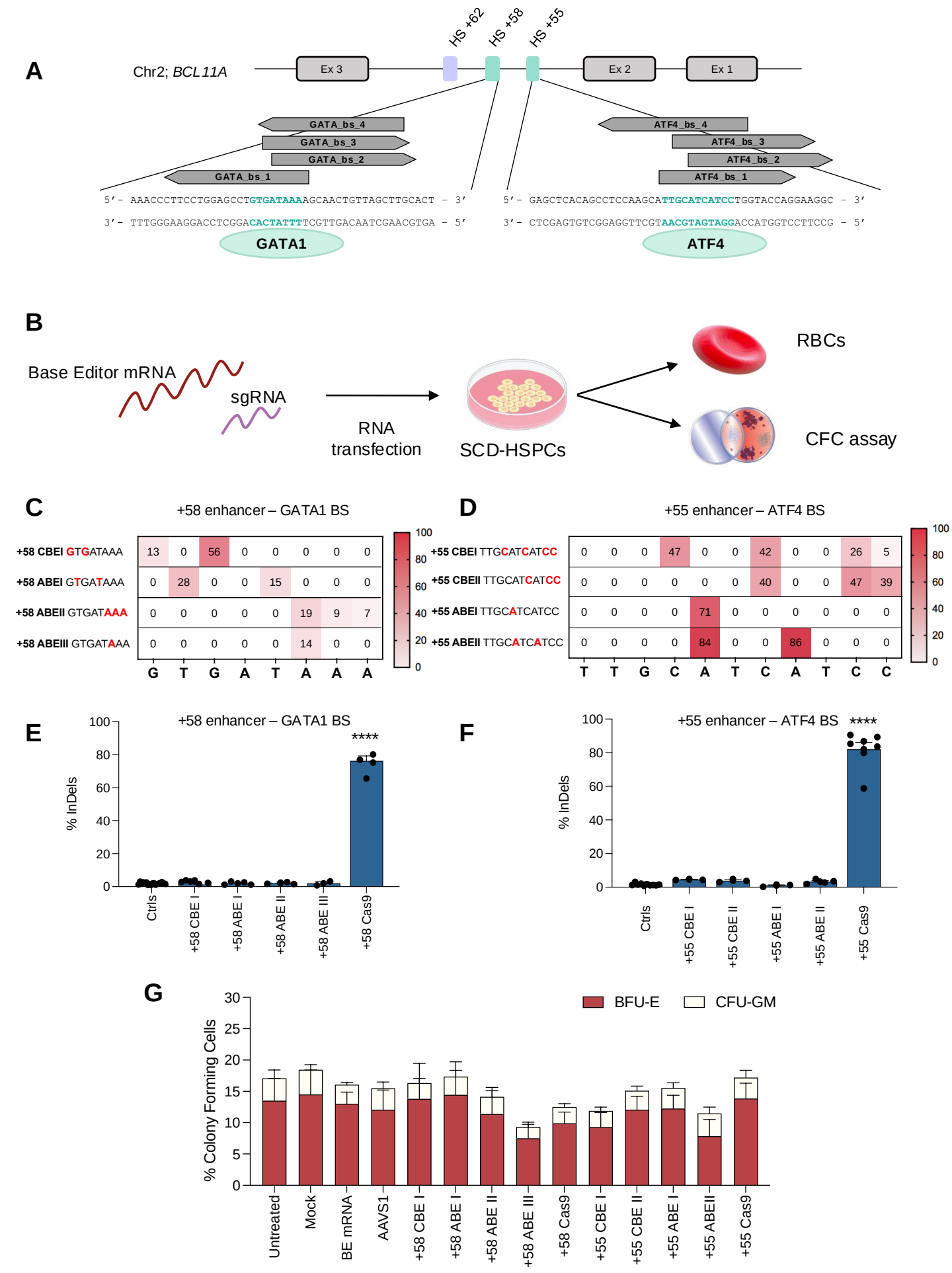
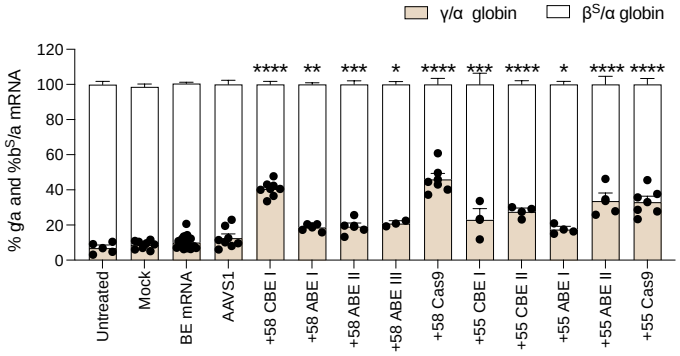

A
Chr2; BCL11A
Ex 3
HS +62
HS +58
HS +55
Ex 2
Ex 1
GATA_bs_4
GATA_bs_3
GATA_bs_2
GATA_bs_1
ATF4_bs_4
ATF4_bs_3
ATF4_bs_2
ATF4_bs_1
5'- AAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTTAGCTTGCACT - 3'
3'- TTTGGGAAGGACCTCGGACACTATTTTCGTTGACAATCGAACGTGA - 5'
5'- GAGCTCACAGCCTCCAAGCATTGCATCATCCTGGTACCAGGAAGGC - 3'
3'- CTCGAGTGTCGGAGGTTCGTAACGTAGTAGGACCATGGTCCTTCCG - 5'
GATA1
ATF4
B
Base Editor mRNA
sgRNA
RNA transfection
SCD-HSPCs
RBCs
CFC assay
C
+58 enhancer – GATA1 BS
+58 CBEI GTGATAAA 13 0 56 0 0 0 0 0
+58 ABEI GTGATAAA 0 28 0 0 15 0 0 0
+58 ABEII GTGATAAA 0 0 0 0 0 19 9 7
+58 ABEIII GTGATAAA 0 0 0 0 0 14 0 0
G T G A T A A A
D
+55 enhancer – ATF4 BS
+55 CBEI TTGCATCATCC 0 0 0 47 0 0 42 0 0 26 5
+55 CBEII TTGCATCATCC 0 0 0 0 0 0 40 0 0 47 39
+55 ABEI TTGCATCATCC 0 0 0 0 71 0 0 0 0 0 0
+55 ABEII TTGCATCATCC 0 0 0 0 84 0 0 86 0 0 0
T T G C A T C A T C C
E
+58 enhancer – GATA1 BS
% InDels
****
Ctrls
+58 CBE I
+58 ABE I
+58 ABE II
+58 ABE III
+58 Cas9
F
+55 enhancer – ATF4 BS
% InDels
****
Ctrls
+55 CBE I
+55 CBE II
+55 ABE I
+55 ABE II
+55 Cas9
G
% Colony Forming Cells
BFU-E
CFU-GM
Untreated
Mock
BE mRNA
AAVS1
+58 CBE I
+58 ABE I
+58 ABE II
+58 ABE III
+58 Cas9
+55 CBE I
+55 CBE II
+55 ABE I
+55 ABEII
+55 Cas9

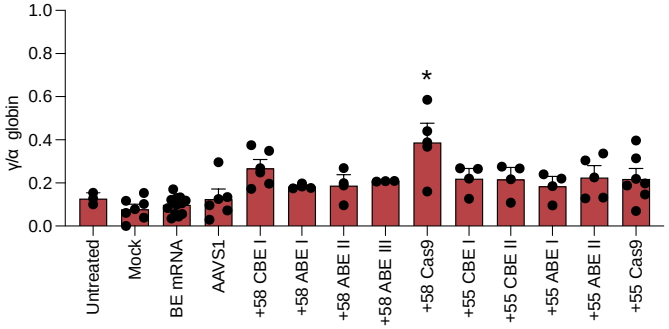
Figure 2

A



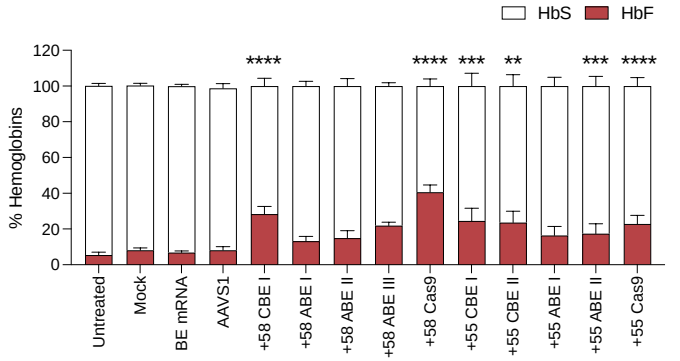
	Untreated	Mock	BE mRNA	AAVS1	+58 CBE I	+58 ABE I	+58 ABE II	+58 ABE III	+58 Cas9	+55 CBE I	+55 CBE II	+55 ABE I	+55 ABE II	+55 Cas9
%EE	-	-	-	21	56	28	19	14	76	47	47	71	86	82
SEM	-	-	-	1.5	4.7	3.4	3.3	2.0	3.0	7.0	3.7	3.5	5.1	4.1

B



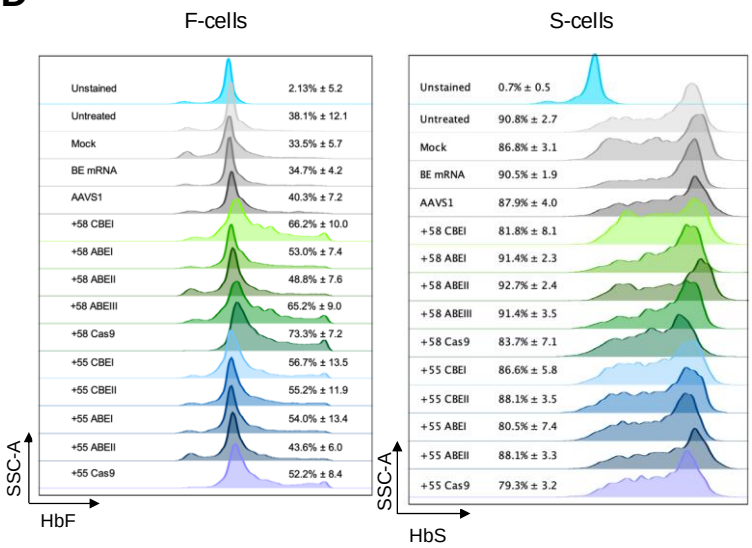
	Untreated	Mock	BE mRNA	AAVS1	+58 CBE I	+58 ABE I	+58 ABE II	+58 ABE III	+58 Cas9	+55 CBE I	+55 CBE II	+55 ABE I	+55 ABE II	+55 Cas9
%EE	-	-	-	21	56	28	19	14	76	47	47	71	86	82
SEM	-	-	-	1.5	4.7	3.4	3.3	2.0	3.0	7.0	3.7	3.5	5.1	4.1

C

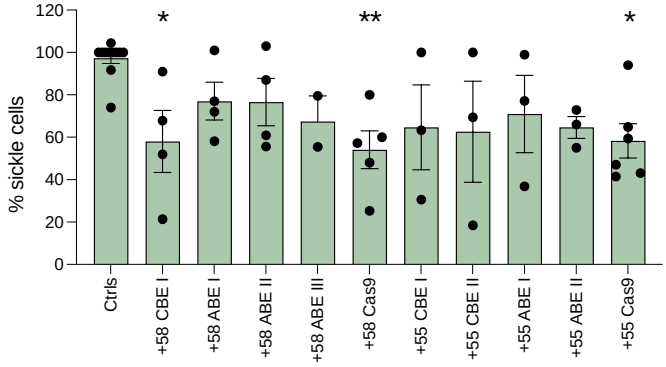


	Untreated	Mock	BE mRNA	AAVS1	+58 CBE I	+58 ABE I	+58 ABE II	+58 ABE III	+58 Cas9	+55 CBE I	+55 CBE II	+55 ABE I	+55 ABE II	+55 Cas9
%EE	-	-	-	21	56	28	19	14	76	47	47	71	86	82
SEM	-	-	-	1.5	4.7	3.4	3.3	2.0	3.0	7.0	3.7	3.5	5.1	4.1

D

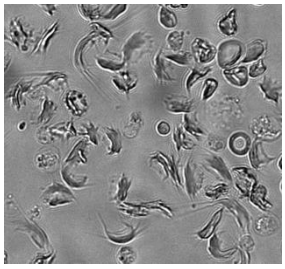


E



	Ctrls	+58 CBE I	+58 ABE I	+58 ABE II	+58 ABE III	+58 Cas9	+55 CBE I	+55 CBE II	+55 ABE I	+55 ABE II	+55 Cas9
%EE	-	56	28	19	14	76	47	47	71	91	82
SEM	-	4.7	3.4	3.3	2.0	3.0	7.0	3.7	3.5	5.1	4.1

Control RBCs



Corrected RBCs

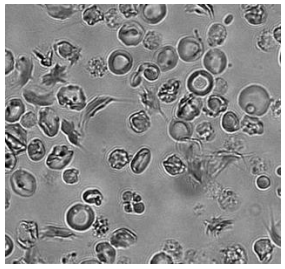
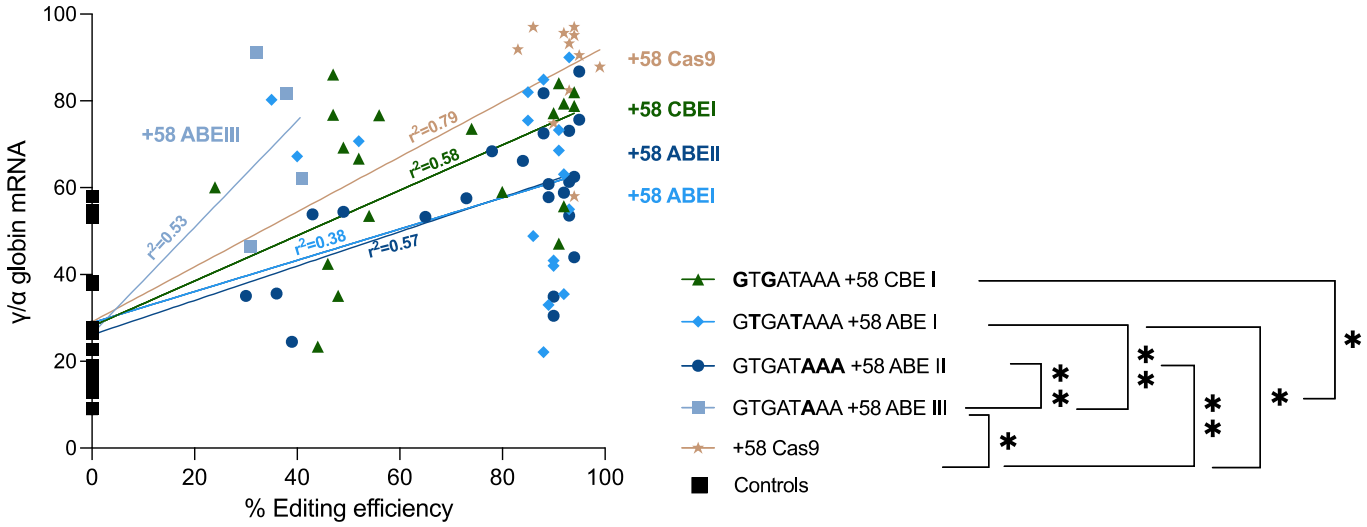


Figure 3

A


γ/α globin mRNA
% Editing efficiency
+58 ABEIII
+58 Cas9
+58 CBEI
+58 ABEII
+58 ABEI
r²=0.79
r²=0.58
r²=0.53
r²=0.38
r²=0.57
GTGATAAA +58 CBE I
GTGATAAA +58 ABE I
GTGATAAA +58 ABE II
GTGATAAA +58 ABE III
+58 Cas9
Controls

B

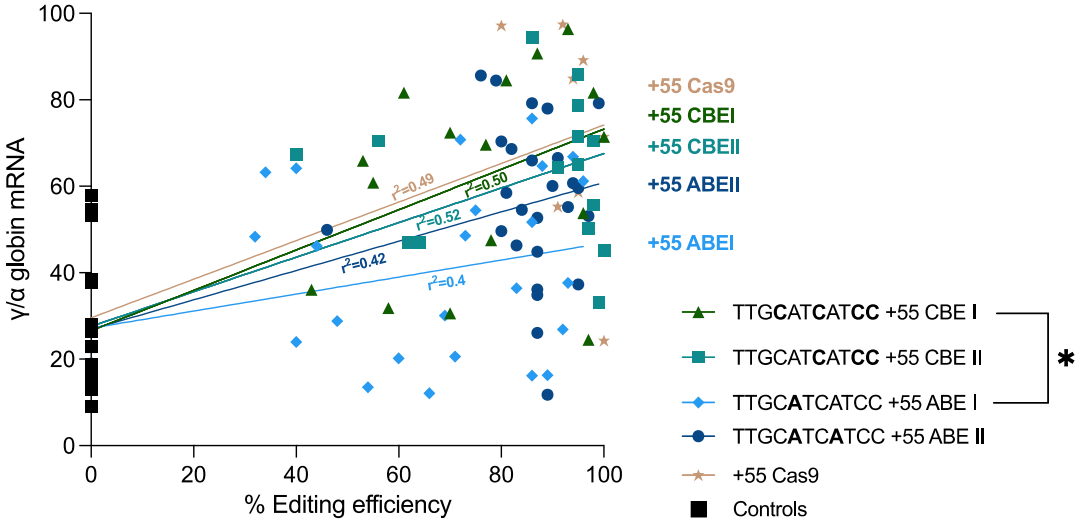

γ/α globin mRNA
% Editing efficiency
+55 Cas9
+55 CBEI
+55 CBEII
+55 ABEII
+55 ABEI
r²=0.49
r²=0.50
r²=0.52
r²=0.42
r²=0.4
TTGCATCATCC +55 CBE I
TTGCATCATCC +55 CBE II
TTGCATCATCC +55 ABE I
TTGCATCATCC +55 ABE II
+55 Cas9
Controls

Figure 4

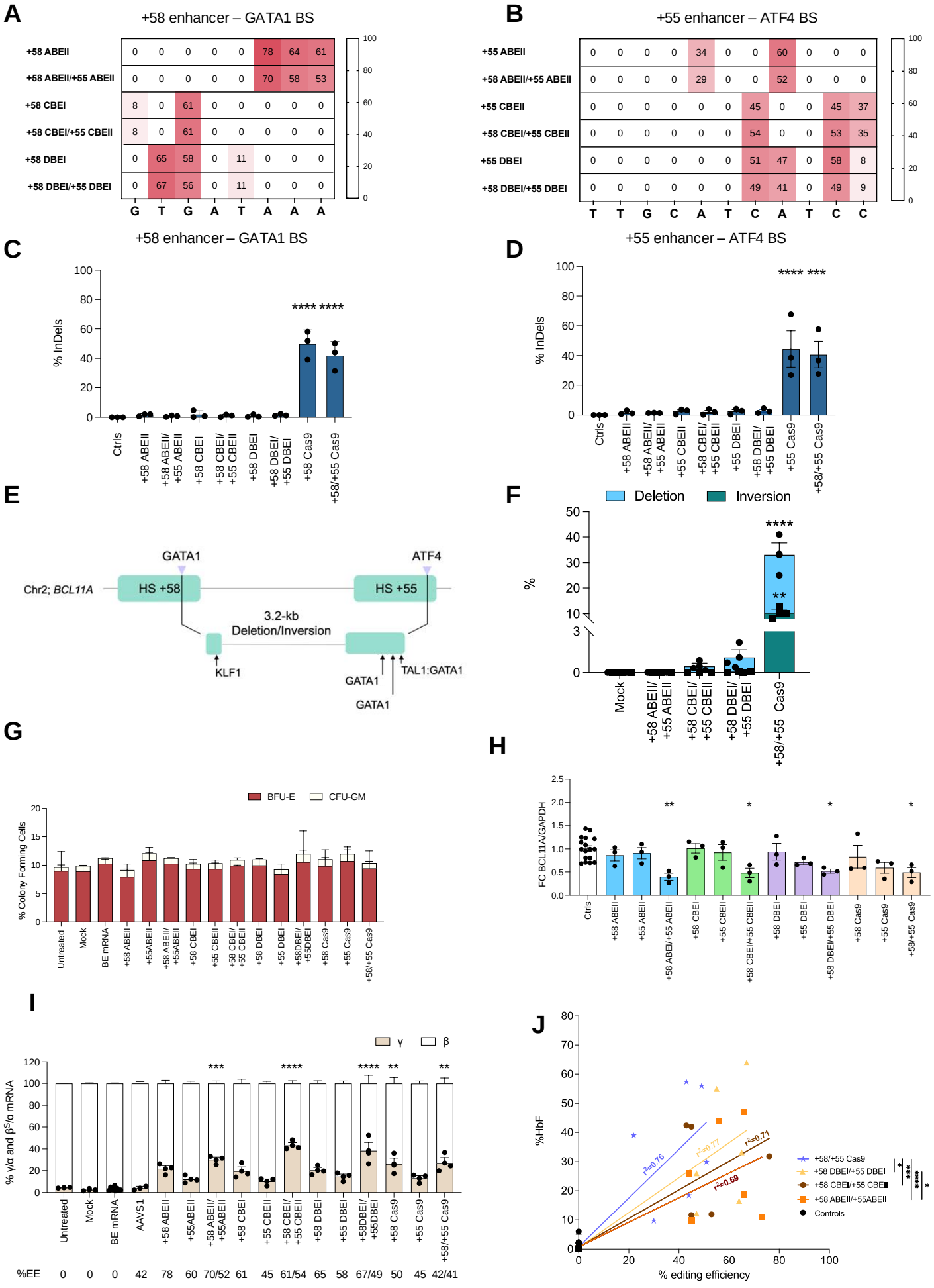

A
+58 enhancer – GATA1 BS
+58 ABEII
+58 ABEII/+55 ABEII
+58 CBEI
+58 CBEI/+55 CBEII
+58 DBEI
+58 DBEI/+55 DBEI
G T G A T A A A
B
+55 enhancer – ATF4 BS
+55 ABEII
+58 ABEII/+55 ABEII
+55 CBEII
+58 CBEI/+55 CBEII
+55 DBEI
+58 DBEI/+55 DBEI
T T G C A T C A T C C
C
+58 enhancer – GATA1 BS
% InDels
Ctrls
+58 ABEII
+58 ABEII/+55 ABEII
+58 CBEI
+58 CBEI/+55 CBEII
+58 DBEI
+58 DBEI/+55 DBEI
+58 Cas9
+58/+55 Cas9
D
+55 enhancer – ATF4 BS
+55 ABEII
+55 CBEII
+55 DBEI
+55 Cas9
E
GATA1
ATF4
Chr2; BCL11A
HS +58
HS +55
3.2-kb
Deletion/Inversion
KLF1
GATA1
TAL1:GATA1
GATA1
F
Deletion
Inversion
%
Mock
G
BFU-E
CFU-GM
% Colony Forming Cells
Untreated
Mock
BE mRNA
H
FC BCL11A/GAPDH
I
γ
β
% γ/α and β^S/α mRNA
AAVS1
%EE 0 0 0 42 78 60 70/52 61 45 61/54 65 58 67/49 50 45 42/41
SEM - - - 6 4.4 3.2 10/6.6 8.4 12 9/10 5 2.9 7.5/3.5 5.5 12 5.5/8.9
J
%HbF
% editing efficiency
+58/+55 Cas9
+58 DBEI/+55 DBEI
+58 CBEI/+55 CBEII
+58 ABEII/+55ABEII
Controls

Figure 5

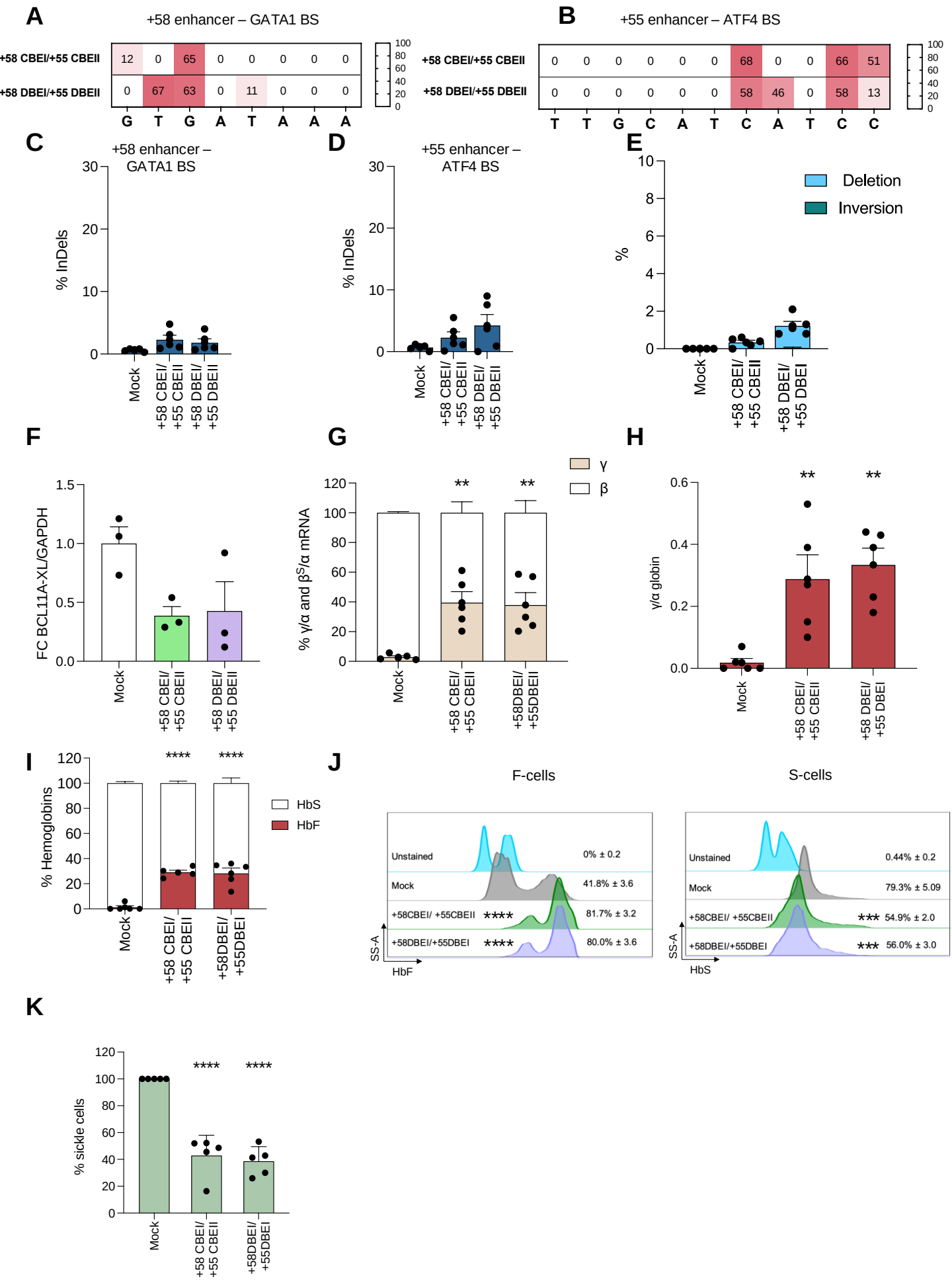

A
+58 enhancer – GATA1 BS
+58 CBEI/+55 CBEII
+58 DBEI/+55 DBEII
G T G A T A A A
B
+55 enhancer – ATF4 BS
T T G C A T C A T C C
C
+58 enhancer – GATA1 BS
% InDels
D
+55 enhancer – ATF4 BS
E
Deletion
Inversion
%
Mock
+58 CBEI/ +55 CBEII
+58 DBEI/ +55 DBEII
F
FC BCL11A-XL/GAPDH
G
% γ/α and β^S/α mRNA
γ
β
H
γ/α globin
I
% Hemoglobins
HbS
HbF
J
F-cells
S-cells
Unstained
Mock
+58CBEI/ +55CBEII
+58DBEI/+55DBEI
0% ± 0.2
41.8% ± 3.6
81.7% ± 3.2
80.0% ± 3.6
0.44% ± 0.2
79.3% ± 5.09
54.9% ± 2.0
56.0% ± 3.0
SS-A
HbF
HbS
K
% sickle cells

Figure 6

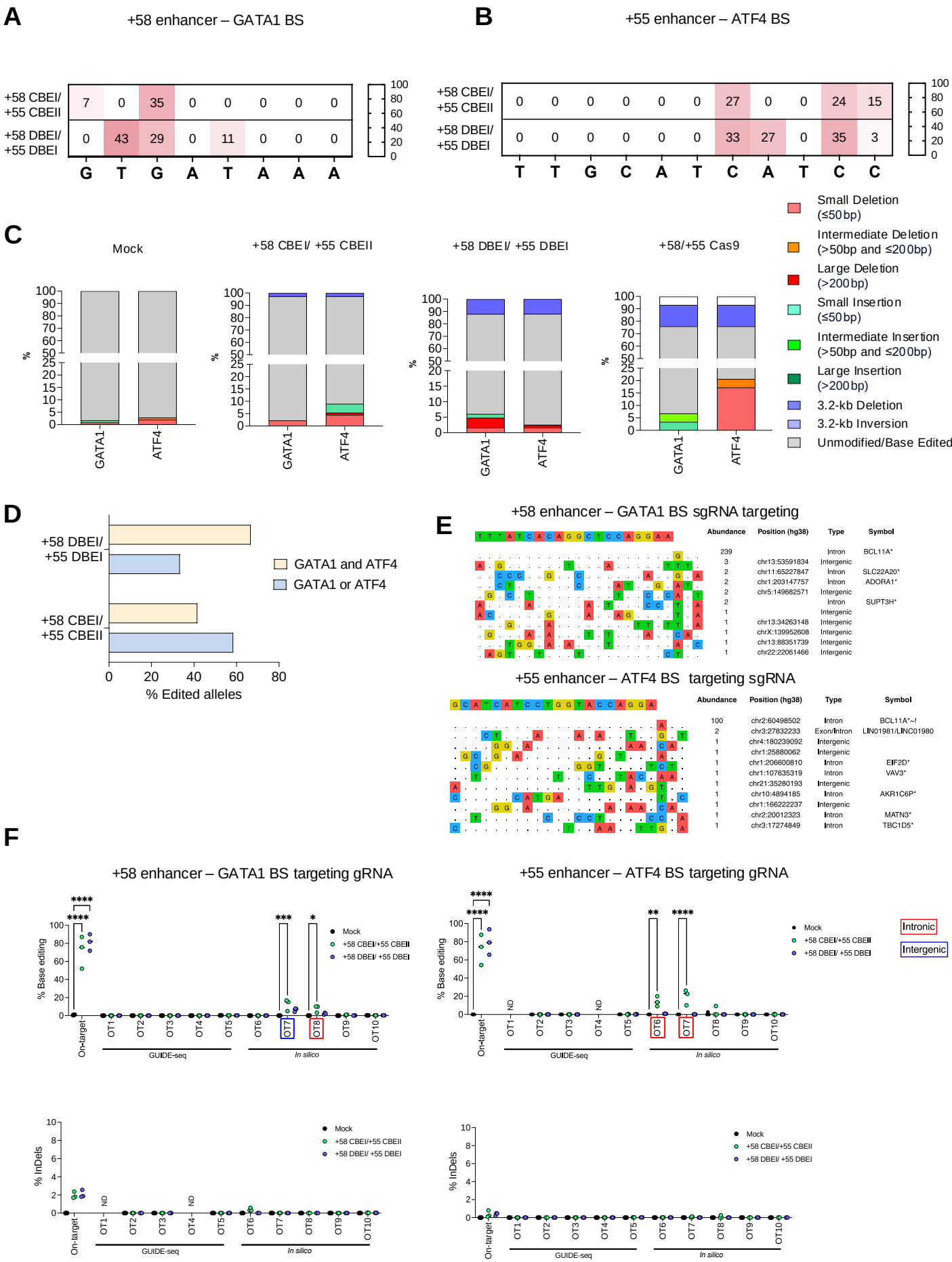
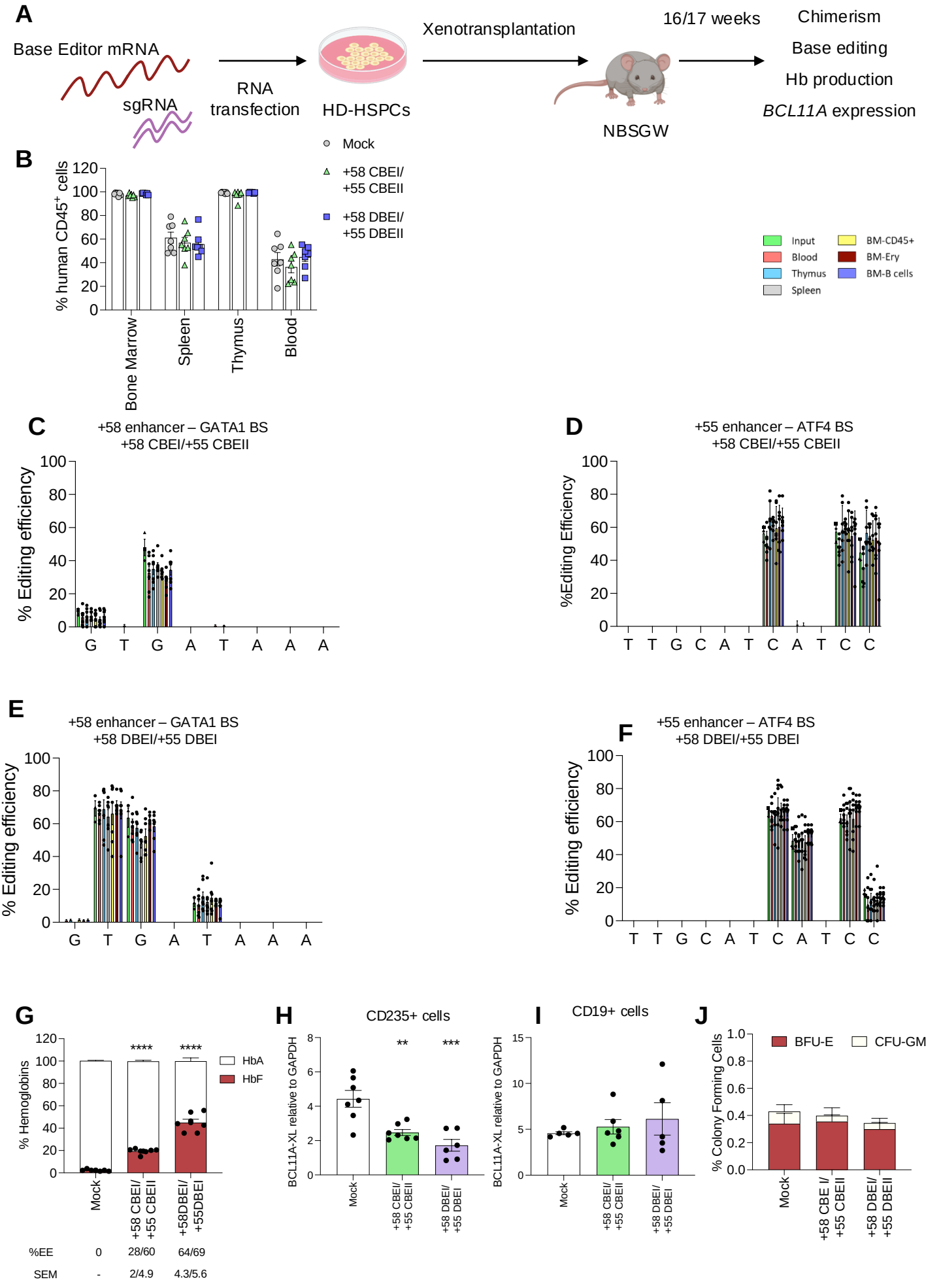
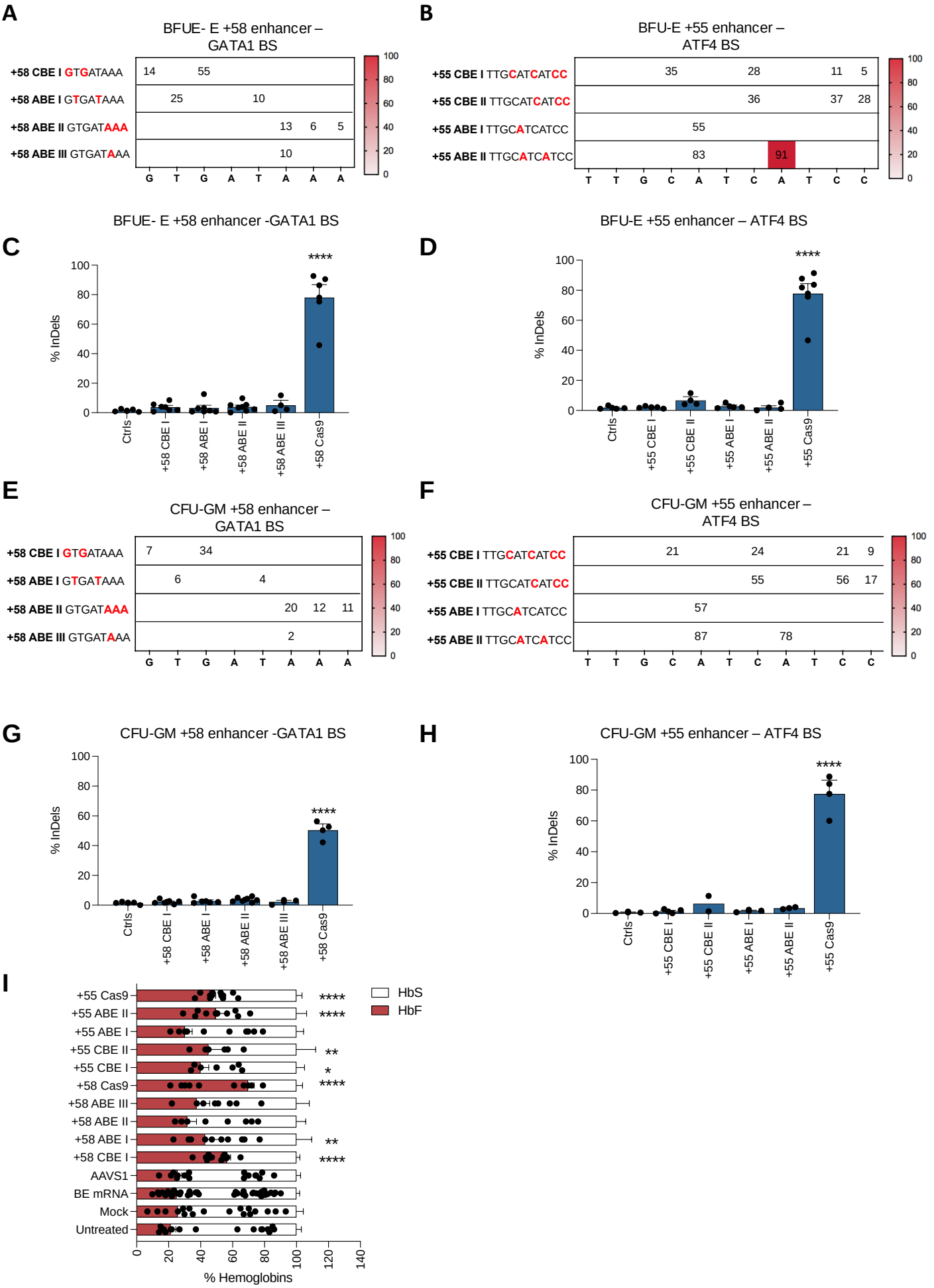

A
+58 enhancer – GATA1 BS
+58 CBEI/ +55 CBEII
+58 DBEI/ +55 DBEI
G T G A T A A A
B
+55 enhancer – ATF4 BS
T T G C A T C A T C C
C
Mock
+58 CBEI/ +55 CBEII
+58 DBEI/ +55 DBEI
+58/+55 Cas9
GATA1
ATF4
Small Deletion (≤50bp)
Intermediate Deletion (>50bp and ≤200bp)
Large Deletion (>200bp)
Small Insertion (≤50bp)
Intermediate Insertion (>50bp and ≤200bp)
Large Insertion (>200bp)
3.2-kb Deletion
3.2-kb Inversion
Unmodified/Base Edited
D
GATA1 and ATF4
GATA1 or ATF4
% Edited alleles
E
+58 enhancer – GATA1 BS sgRNA targeting
Abundance
Position (hg38)
Type
Symbol
+55 enhancer – ATF4 BS targeting sgRNA
F
+58 enhancer – GATA1 BS targeting gRNA
+55 enhancer – ATF4 BS targeting gRNA
% Base editing
% InDels
Mock
Intronic
Intergenic
On-target
GUIDE-seq
In silico
ND

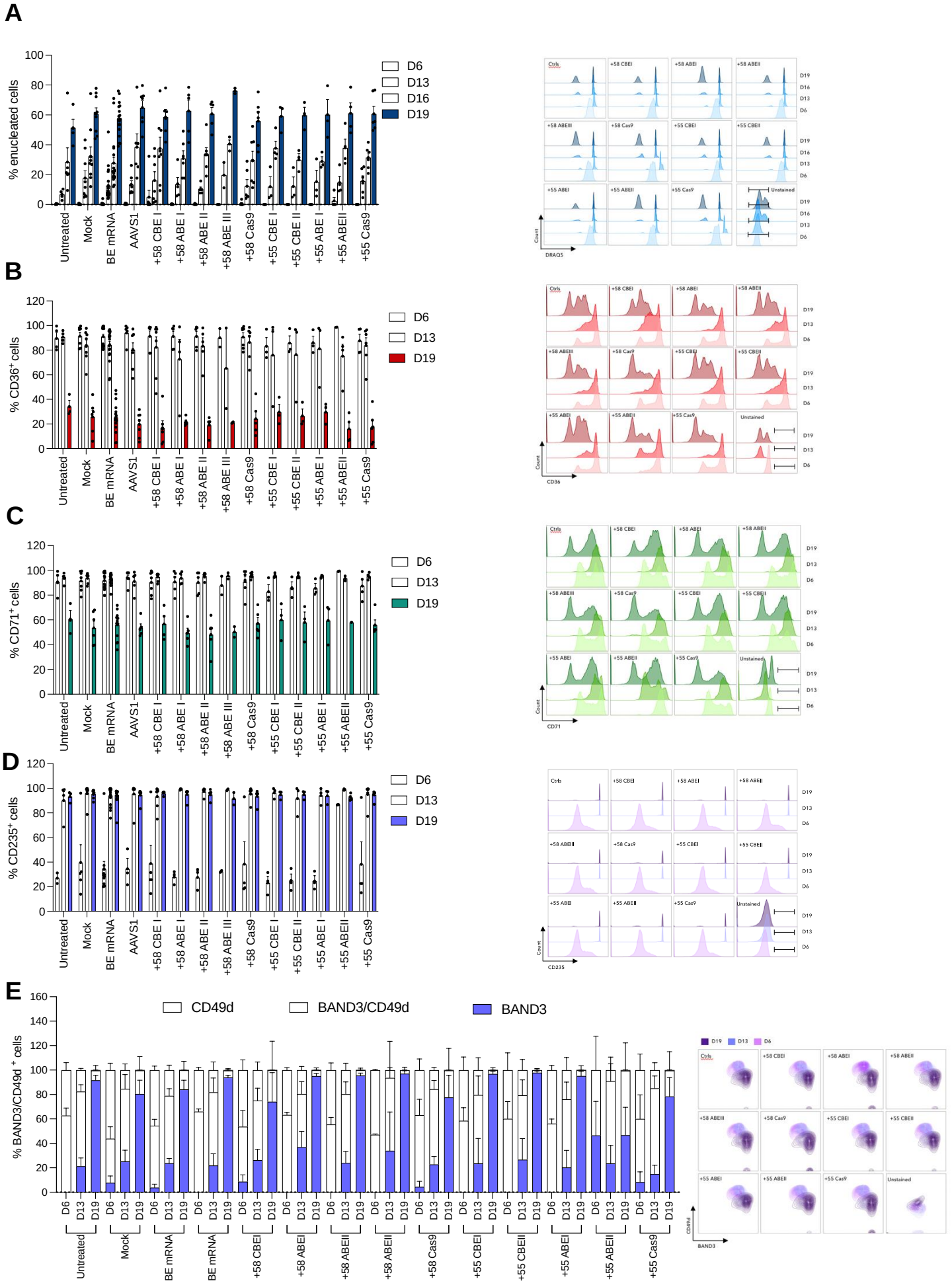
Figure 7


A
Base Editor mRNA
sgRNA
RNA transfection
HD-HSPCs
Xenotransplantation
NBSGW
16/17 weeks
Chimerism
Base editing
Hb production
BCL11A expression
B
% human CD45+ cells
Mock
+58 CBEI/ +55 CBEII
+58 DBEI/ +55 DBEII
Bone Marrow
Spleen
Thymus
Blood
Input
Blood
Thymus
Spleen
BM-CD45+
BM-Ery
BM-B cells
C
+58 enhancer – GATA1 BS
+58 CBEI/+55 CBEII
% Editing efficiency
G T G A T A A A
D
+55 enhancer – ATF4 BS
+58 CBEI/+55 CBEII
%Editing Efficiency
T T G C A T C A T C C
E
+58 enhancer – GATA1 BS
+58 DBEI/+55 DBEI
% Editing efficiency
G T G A T A A A
F
+55 enhancer – ATF4 BS
+58 DBEI/+55 DBEI
% Editing efficiency
T T G C A T C A T C C
G
% Hemoglobins
HbA
HbF
Mock
+58 CBEI/ +55 CBEII
+58DBEI/ +55DBEI
%EE 0 28/60 64/69
SEM - 2/4.9 4.3/5.6
H
CD235+ cells
BCL11A-XL relative to GAPDH
I
CD19+ cells
J
% Colony Forming Cells
BFU-E
CFU-GM

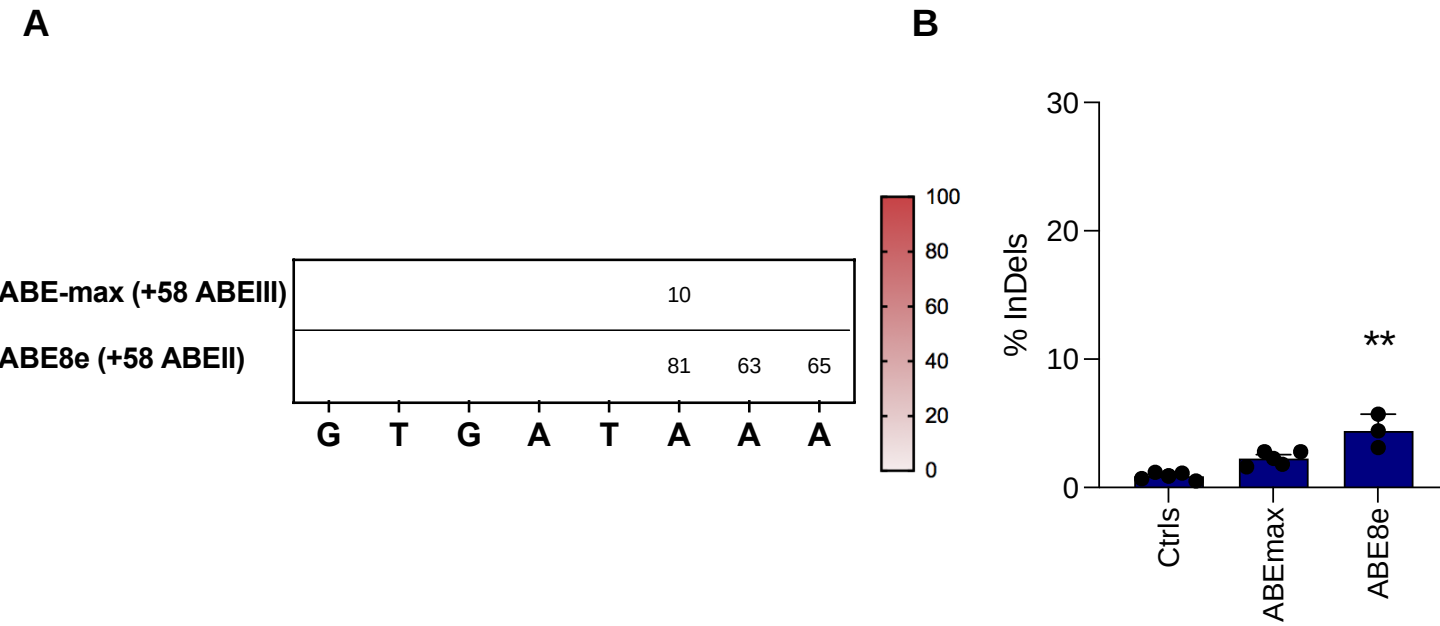
Supplementary Figure 1



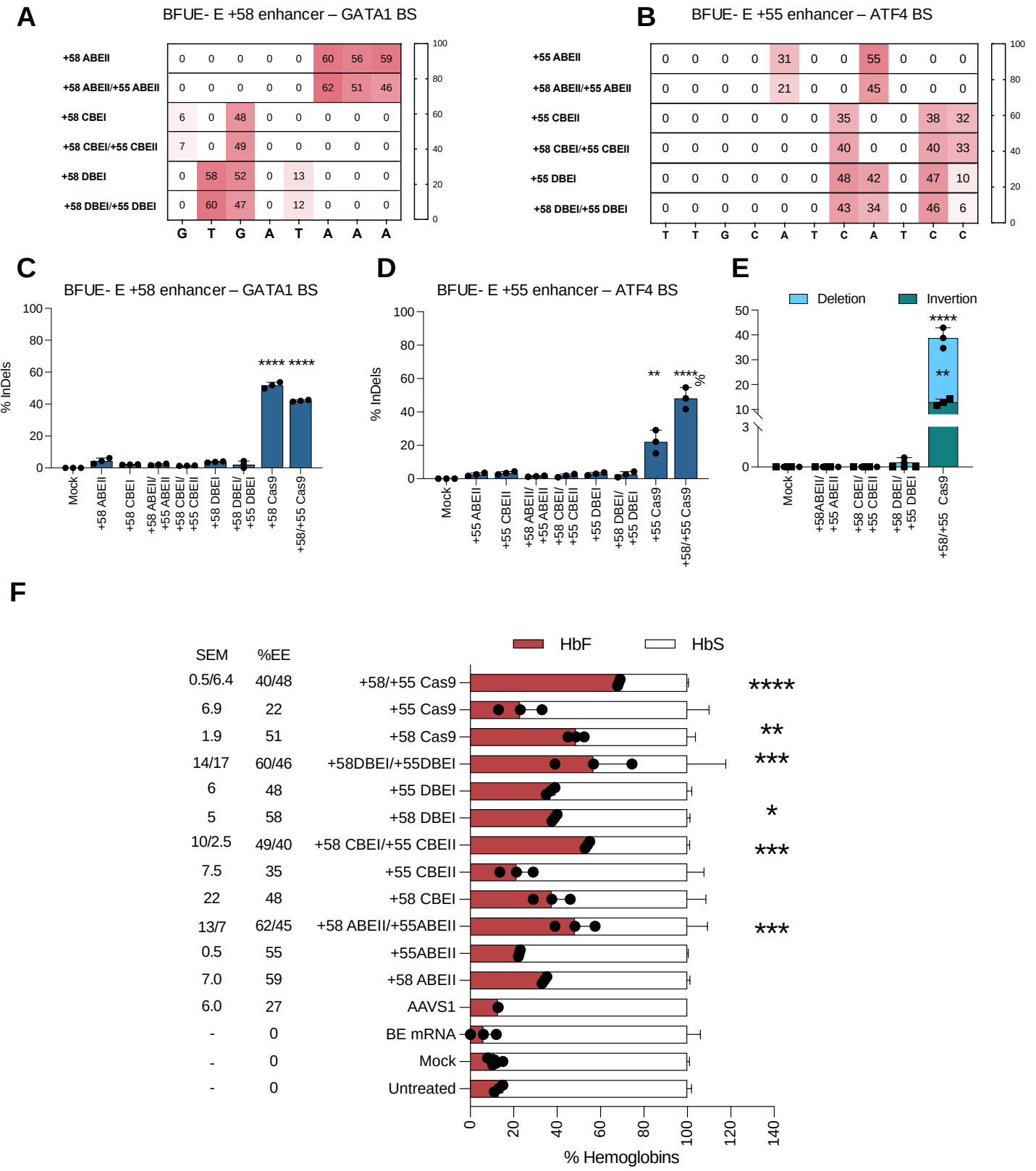
Supplementary Figure 2



Supplementary Figure 3

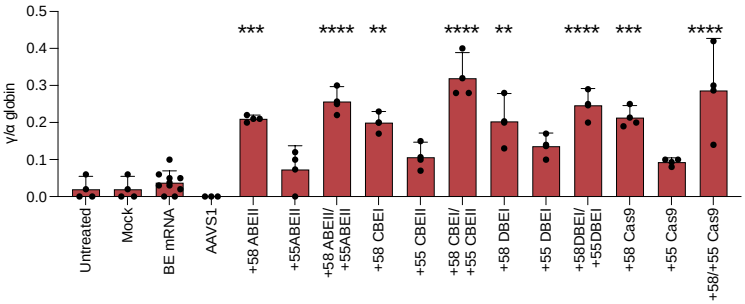


Supplementary Figure 4



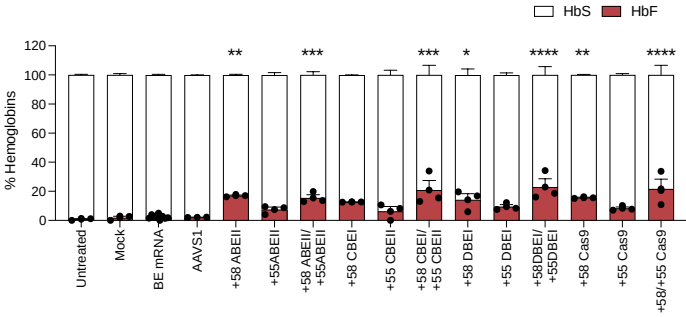
Supplementary Figure 5

A



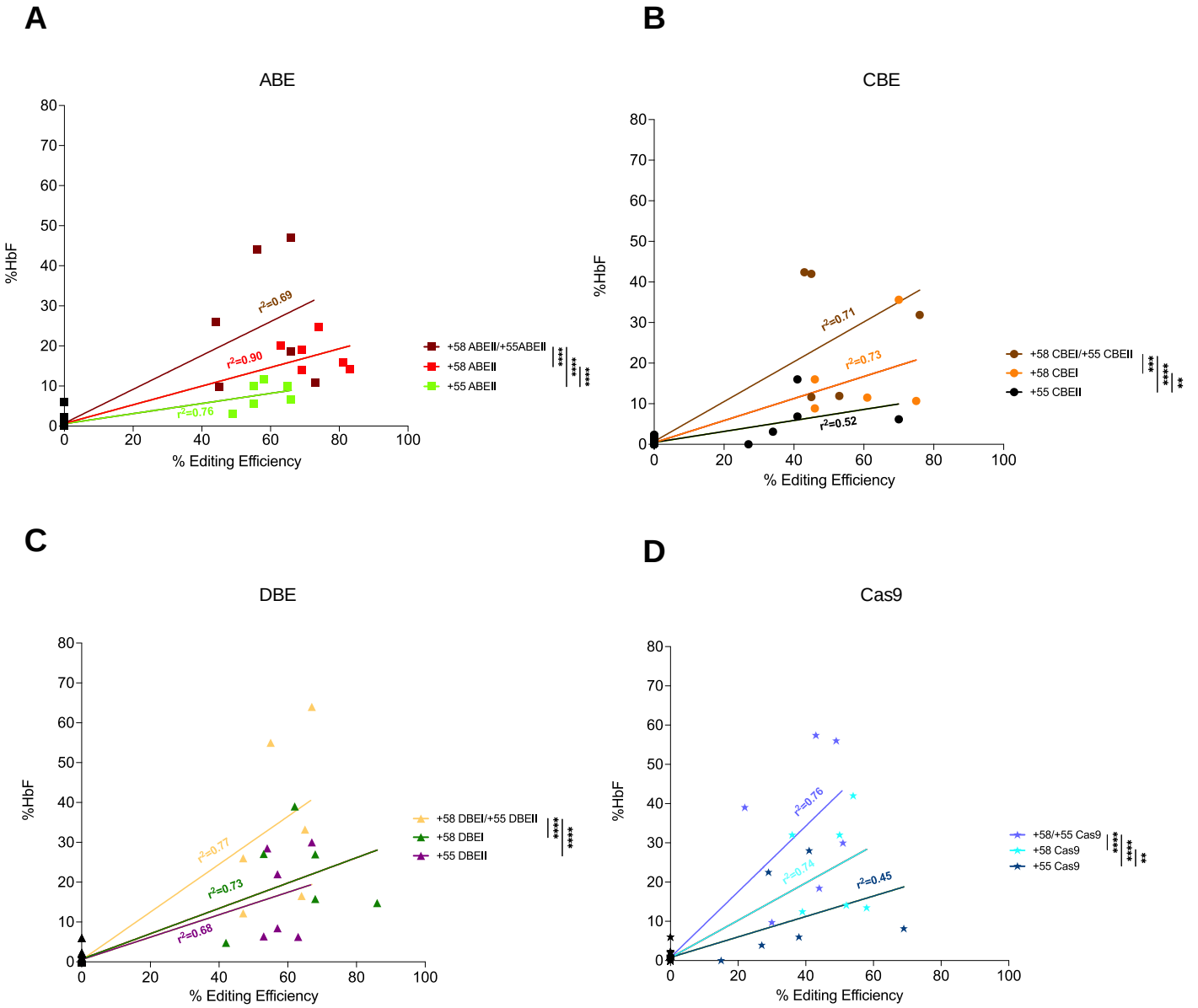
	Untreated	Mock	BE mRNA	AAVS1	+58 ABEII	+55ABEII	+58 ABEII/ +55ABEII	+58 CBEI	+55 CBEII	+58 CBEI/ +55 CBEII	+58 DBEI	+55 DBEI	+58DBEI/ +55DBEI	+58 Cas9	+55 Cas9	+58/+55 Cas9
%EE	0	0	0	42	78	60	70/52	61	45	61/54	65	58	67/49	50	45	42/41
SEM	-	-	-	6	4.4	3.2	10/6.6	8.4	12	9/10	5	2.9	7.5/3.5	5.5	12	5.5/8.9

B

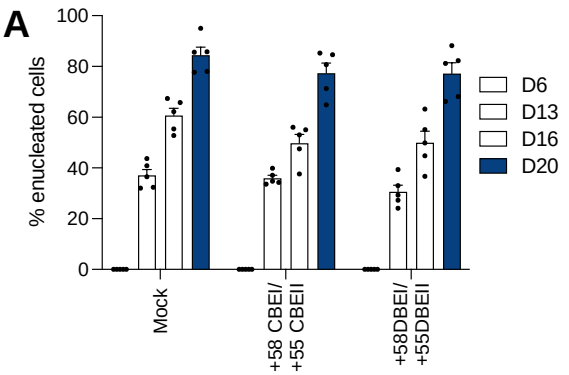


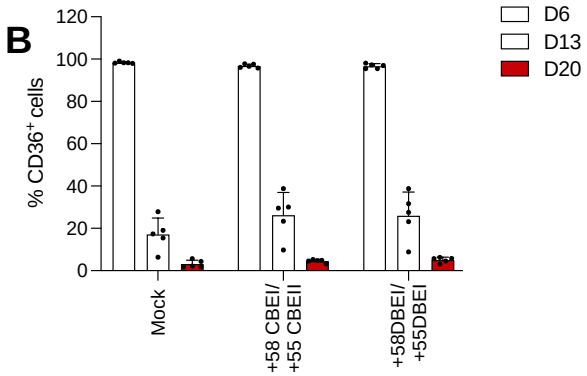
	Untreated	Mock	BE mRNA	AAVS1	+58 ABEII	+55ABEII	+58 ABEII/ +55ABEII	+58 CBEI	+55 CBEII	+58 CBEI/ +55 CBEII	+58 DBEI	+55 DBEI	+58DBEI/ +55DBEI	+58 Cas9	+55 Cas9	+58/+55 Cas9
%EE	0	0	0	42	78	60	70/52	61	45	61/54	65	58	67/49	50	45	42/41
SEM	-	-	-	6.5	4.4	3.2	10/6.6	8.4	12	9/10	5	2.9	7.5/3.5	5.5	12	5.5/8.9

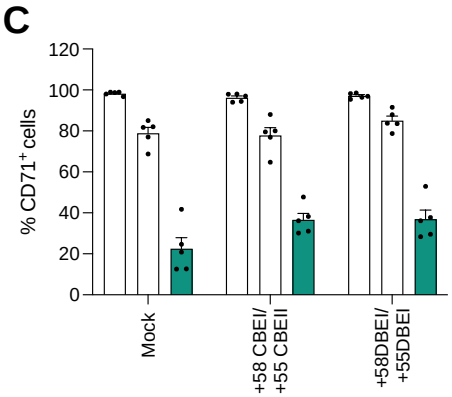
Supplementary Figure 6

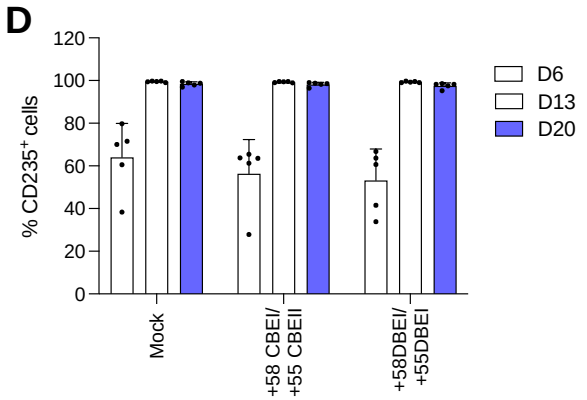


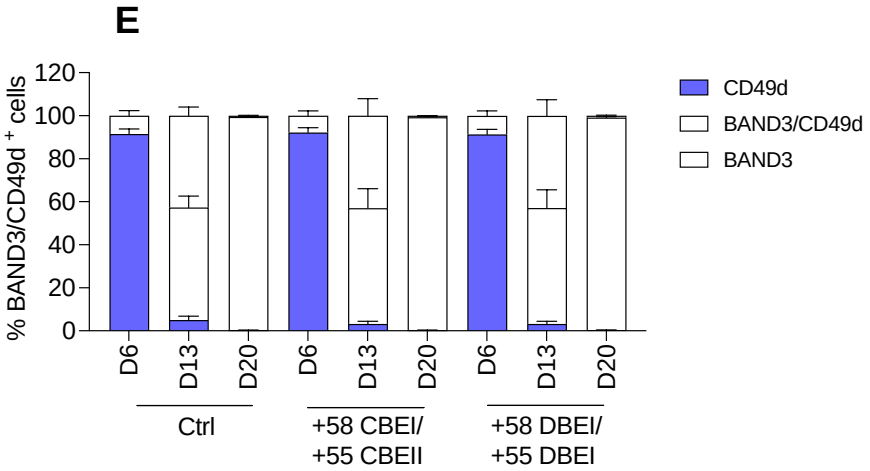
Supplementary Figure 7











Supplementary Figure 8

A

+58 CBEI/ +55 CBEII
+58 enhancer – GATA1 BS

Donor1

```
GGCCAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTT-Reference
        sgRNA

GGCCAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTT-43.96% (128486 reads)
GGCCAAACCCTTCCTGGAGCCTGTAATAAAAGCAACTGTT-39.98% (116879 reads)
GGCCAAACCCTTCCTGGAGCCTATATAAAAGCAACTGTT-5.14% (15038 reads)
GGCCAAACCCTTCCTGGAGCCTGTCATAAAAGCAACTGTT-4.20% (12288 reads)
GGCCAAACCCTTCCTGAAGCCTGTATAAAAGCAACTGTT-0.79% (2305 reads)
GGCCAAACCCTTCCTGGAGCCTGTTATAAAAGCAACTGTT-0.37% (1075 reads)
TGCCAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTT-0.35% (1034 reads)
TGCCAAACCCTTCCTGGAGCCTGTATAAAAGCAACTGTT-0.33% (966 reads)
GGCCAAACCCTTCCTGGAGCCTATGATAAAAGCAACTGTT-0.32% (950 reads)
GGCCAAACCCTTCCTGGAGCCTATCATAAAAGCAACTGTT-0.25% (718 reads)
GGCCAAACCCTTCCTGAAGCCTGTGATAAAAGCAACTGTT-0.24% (711 reads)
```

Donor2

```
GGCCAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTT-Reference
        sgRNA

GGCCAAACCCTTCCTGGAGCCTGTATAATAAAAGCAACTGTT-60.53% (50039 reads)
GGCCAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTT-21.50% (17773 reads)
GGCCAAACCCTTCCTGGAGCCTATATAATAAAAGCAACTGTT-9.98% (8247 reads)
GGCCAAACCCTTCCTGAAGCCTGTATAATAAAAGCAACTGTT-1.38% (1138 reads)
GGCCAAACCCTTCCTGGAGCCTGTCATAAAAGCAACTGTT-1.37% (1129 reads)
GGCCAAACCCTTCCTGGAGCCTATCATAAAAGCAACTGTT-0.79% (654 reads)
GGCCAAACCCTTCCTAAAGCCTGTATAATAAAAGCAACTGTT-0.58% (480 reads)
GGCCAAACCCTTCCTGCAGCCTGTGATAAAAGCAACTGTT-0.39% (319 reads)
GGCCAAACCCTTCCTGAAGCCTATATAATAAAAGCAACTGTT-0.24% (196 reads)
TGCCAAACCCTTCCTGGAGCCTGTATAATAAAAGCAACTGTT-0.23% (189 reads)
GGCCAAACCCTTCCTGGAACCTGTATAATAAAAGCAACTGTT-0.22% (183 reads)
```

Donor3

```
GGCCAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTT-Reference
        sgRNA

GGCCAAACCCTTCCTGGAGCCTGTATAATAAAAGCAACTGTT-67.39% (83155 reads)
GGCCAAACCCTTCCTGGAGCCTATATAATAAAAGCAACTGTT-11.84% (14615 reads)
GGCCAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTT-9.36% (11552 reads)
GGCCAAACCCTTCCTGGAGCCTGTCATAAAAGCAACTGTT-3.75% (4621 reads)
GGCCAAACCCTTCCTGAAGCCTGTATAATAAAAGCAACTGTT-1.53% (1894 reads)
GGCCAAACCCTTCCTGGAGCCTATCATAATAAAAGCAACTGTT-0.97% (1195 reads)
TGCCAAACCCTTCCTGGAGCCTGTATAATAAAAGCAACTGTT-0.35% (430 reads)
GGCCAAACCCTTCCTGAAGCCTATATAATAAAAGCAACTGTT-0.33% (413 reads)
GACCAAACCCTTCCTGGAGCCTGTATAATAAAAGCAACTGTT-0.27% (338 reads)
GGCCAAACCCTTCCTGGAGCATGTATAATAAAAGCAACTGTT-0.23% (289 reads)
GGCCAAACCCTTCCTGAAGCCTGTGATAATAAAAGCAACTGTT-0.20% (248 reads)
GGCCAAACCCTTCCTGGAGCCTGTTATAATAAAAGCAACTGTT-0.20% (247 reads)
```

B

+58 CBEI/ +55 CBEII
+55 enhancer – ATF4 BS

Donor1

```
CCCATTGCCTTCCTGGTACCAGGGATGATGCAATGCTTGGA-Reference
        sgRNA

CCCATTGCCTTCCTGGTACCAGGGATGATGCAATGCTTGGA-40.77% (239267 reads)
CCCATTGCCTTCCTGGTACCAAAATAATGCAATGCTTGGA-22.56% (132428 reads)
CCCATTGCCTTCCTGGTACCAGAATAATGCAATGCTTGGA-17.20% (100964 reads)
CCCATTGCCTTCCTGGTACCAGGATATGCAATGCTTGGA-6.75% (39622 reads)
CCCATTGCCTTCCTGGTACCAGAATATGCAATGCTTGGA-3.28% (19230 reads)
CCCATTGCCTTCCTGGTACCAAAATGATGCAATGCTTGGA-0.92% (5411 reads)
CCCATTGCCTTCCTGGTACCAGCATATGCAATGCTTGGA-0.77% (4494 reads)
CCCATTGCCTTCCTGGTACCAGCATGATGCAATGCTTGGA-0.65% (3840 reads)
CCCATTGCCTTCCTGGTACCAGGATTATGCAATGCTTGGA-0.52% (3045 reads)
CCCATTGCCTTCCTGGTACCAGTATATGCAATGCTTGGA-0.44% (2588 reads)
CCCATTGCCTTCCTGGTACCACAATAATGCAATGCTTGGA-0.35% (2039 reads)
CCCATTGCCTTCCTGGTACCAGTATGATGCAATGCTTGGA-0.33% (1932 reads)
CCCATTGCCTTCCTGGTACC--AAAATGCAATGCTTGGA-0.31% (1838 reads)
CCCATTGCCTTCCTGGTACCAGAATTATGCAATGCTTGGA-0.29% (1686 reads)
CCCATTGCCTTCCTGGTACCAGGATCATGCAATGCTTGGA-0.23% (1321 reads)
```

Donor2

```
CCCATTGCCTTCCTGGTACCAGGATGATGCAATGCTTGGA-Reference
        sgRNA

CCCATTGCCTTCCTGGTACCAAAATAATGCAATGCTTGGA-36.21% (109037 reads)
CCCATTGCCTTCCTGGTACCAGAATGATGCAATGCTTGGA-24.14% (72678 reads)
CCCATTGCCTTCCTGGTACCAGGATGATGCAATGCTTGGA-22.71% (68386 reads)
CCCATTGCCTTCCTGGTACCAGGATATGCAATGCTTGGA-5.50% (16548 reads)
CCCATTGCCTTCCTGGTACCAAAATGATGCAATGCTTGGA-1.66% (5003 reads)
CCCATTGCCTTCCTGGTACCAGCATGATGCAATGCTTGGA-0.92% (2774 reads)
CCCATTGCCTTCCTGGTACCAAGATAATGCAATGCTTGGA-0.84% (2516 reads)
CCCATTGCCTTCCTGGTACCAGAATGATGCAATGCTTGGA-0.63% (1905 reads)
CCCATTGCCTTCCTGGTACCAGCATAATGCAATGCTTGGA-0.60% (1798 reads)
CCCATTGCCTTTCTGGTACCAGGATGATGCAATGCTTGGA-0.30% (912 reads)
CCCATTGCCTTCCTGGTACCACAATGATGCAATGCTTGGA-0.30% (895 reads)
CCCATTGCCTTCCTGGTACCAATATAATGCAATGCTTGGA-0.29% (888 reads)
CCCATTGCCTTCCTGGTACCAGTATAATGCAATGCTTGGA-0.29% (884 reads)
CCCATTGCCTTCCTGGTACCACAATAATGCAATGCTTGGA-0.29% (879 reads)
CCCATTGCCTTCCTGGTACCAAAATTATGCAATGCTTGGA-0.29% (872 reads)
CCCATTGCCTTCCTGGTACCAGGATTATGCAATGCTTGGA-0.28% (853 reads)
CCCATTGCCTTCCTGGTACCAGTATGATGCAATGCTTGGA-0.28% (839 reads)
CCCATTGCCTTCCTGGTAC-AGGATGATGCAATGCTTGGA-0.27% (828 reads)
CCCATTGCCTTCCTGGTACCAGAATGATGCAATGCTTGGA-0.27% (825 reads)
CCCATTGCCTT--TGGTACCAGGGATGATGCAATGCTTGGA-0.27% (809 reads)
CCCATTGCCTTCCTAGTACCAAAATTATGCAATGCTTGGA-0.27% (808 reads)
CCCATTGCCTTCCTGGTACCAAAATAATGCAATGCTTGAA-0.26% (780 reads)
CCCATTGCCTTCCTGGTACCAGGAATAATGCAATGCTTGG-0.24% (732 reads)
CCCATTGCCTTCCTGGTACC---ATGATGCAATGCTTGGA-0.22% (668 reads)
CCCATTGCCTTCCTGGTACC---AAAATGCAATGCTTGGA-0.22% (663 reads)
CCCATTGCCTTCCTGGTACC---ATAATGCAATGCTTGGA-0.21% (622 reads)
```

Donor3

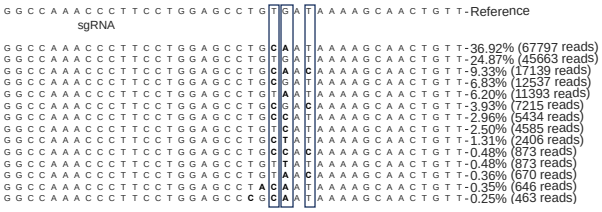
```
CCCATTGCCTTCCTGGTACCAGGATGATGCAATGCTTGGA-Reference
        sgRNA

CCCATTGCCTTCCTGGTACCAAAATAATGCAATGCTTGGA-41.63% (277340 reads)
CCCATTGCCTTCCTGGTACCAGAATAATGCAATGCTTGGA-32.55% (216862 reads)
CCCATTGCCTTCCTGGTACCAGGATGATGCAATGCTTGGA-8.08% (53810 reads)
CCCATTGCCTTCCTGGTACCAGGATAATGCAATGCTTGGA-6.81% (45379 reads)
CCCATTGCCTTCCTGGTACCAAAATGATGCAATGCTTGGA-1.47% (9774 reads)
CCCATTGCCTTCCTGGTACCAGCATAATGCAATGCTTGGA-1.24% (8271 reads)
CCCATTGCCTTCCTGGTACCAGAATGATGCAATGCTTGGA-0.72% (4801 reads)
CCCATTGCCTTCCTGGTACCACAATAATGCAATGCTTGGA-0.47% (3145 reads)
CCCATTGCCTTCCTGGTACCAAAATAATGCAATGCTTGAA-0.41% (2740 reads)
CCCATTGCCTTCCTGGTACCAGTATAATGCAATGCTTGGA-0.32% (2125 reads)
CCCATTGCCTTCCTGGTACCAATTATAATGCAATGCTTGGA-0.30% (1966 reads)
CCCATTGCCTTCCTGGTACCAGGATTATGCAATGCTTGGA-0.29% (1965 reads)
CCCATTGCCTTCCTGGTACCAGCATGATGCAATGCTTGGA-0.24% (1597 reads)
CCCATTGCCTTCCTGGTACCAGAATCATGCAATGCTTGGA-0.24% (1568 reads)
CCCATTGCCTTCCTGGTACCACAATGATGCAATGCTTGGA-0.23% (1502 reads)
CCCATTGCCTTCCTGGTACCAAAATCATGCAATGCTTGGA-0.22% (1477 reads)
CCCATTGCCTTCCTGGTACCAAAATTATGCAATGCTTGGA-0.20% (1364 reads)
CCCATTGCCTTCCTGGTACCAGAATTATGCAATGCTTGGA-0.20% (1348 reads)
```

Supplementary Figure 8

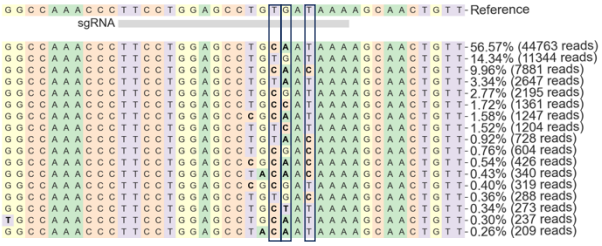
C

+58 DBEI/ +55 DBEI
+58 enhancer – GATA1 BS

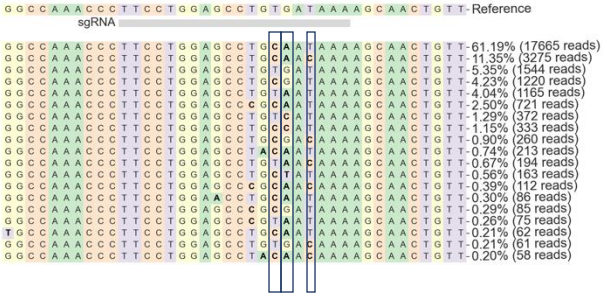
Donor1



Donor2



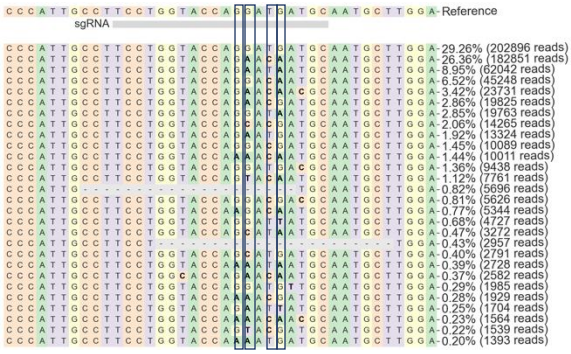
Donor3



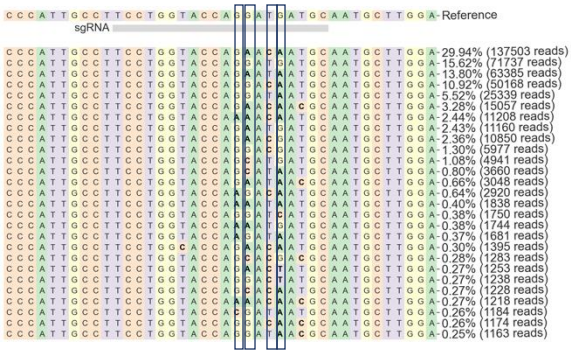
D

+58 DBEI/ +55 DBEI
+55 enhancer – ATF4 BS

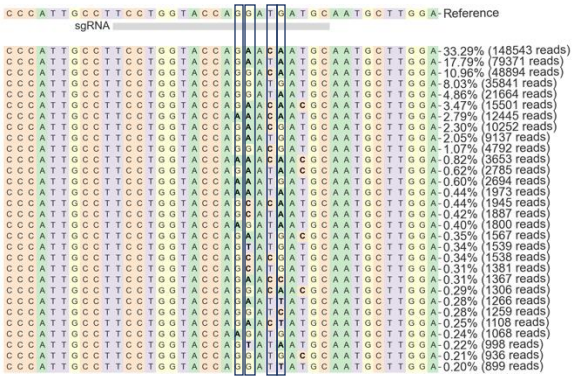
Donor1



Donor2

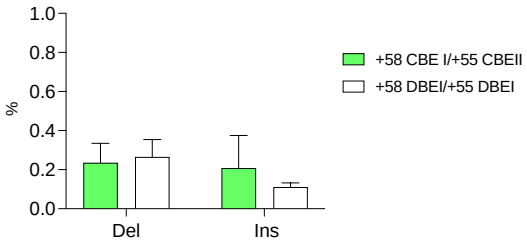


Donor3

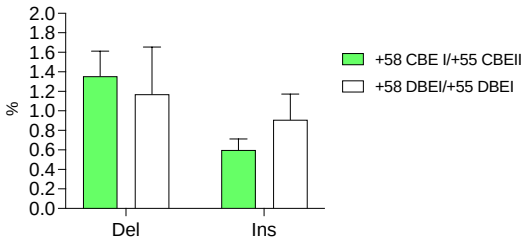


E

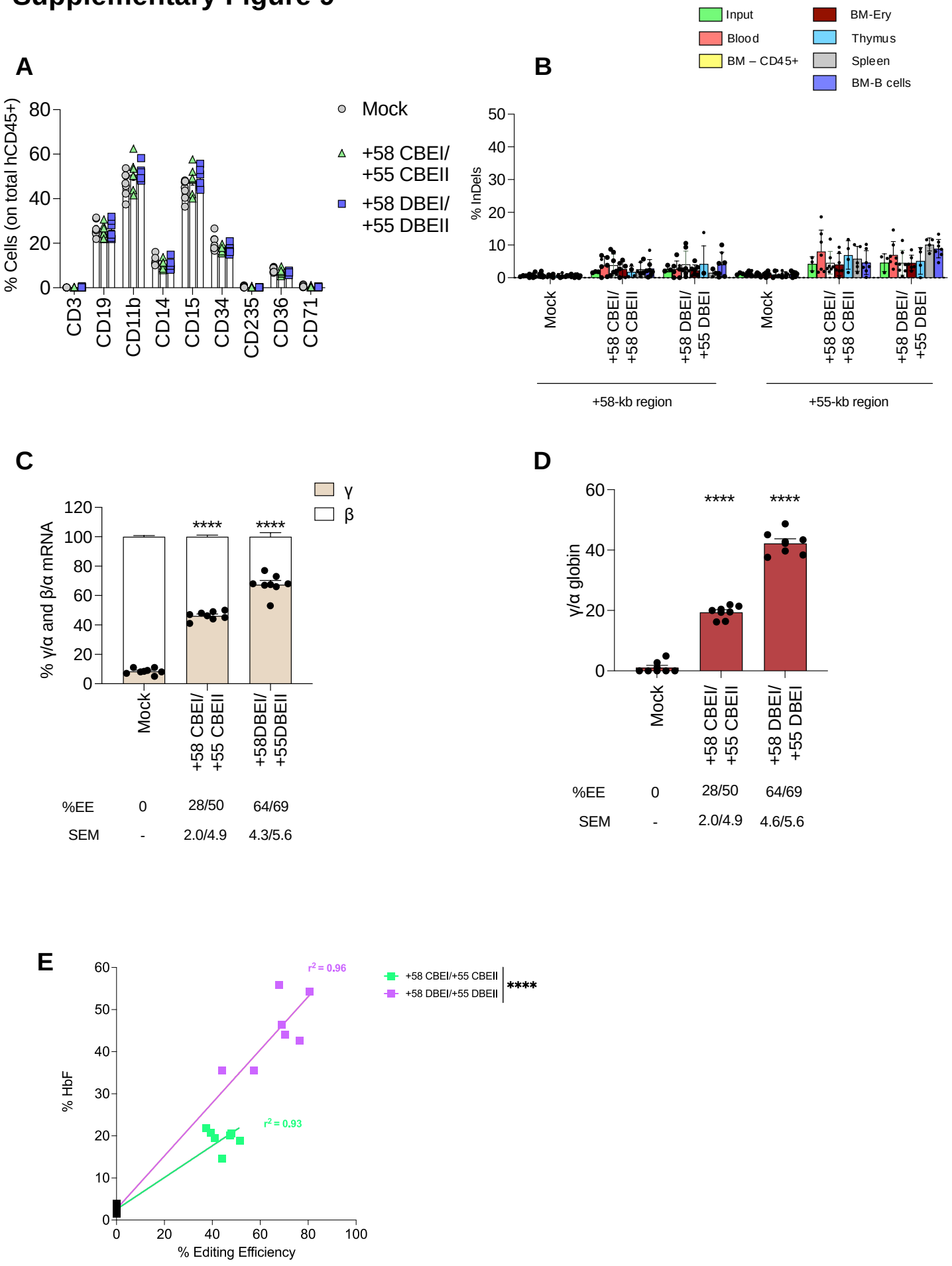
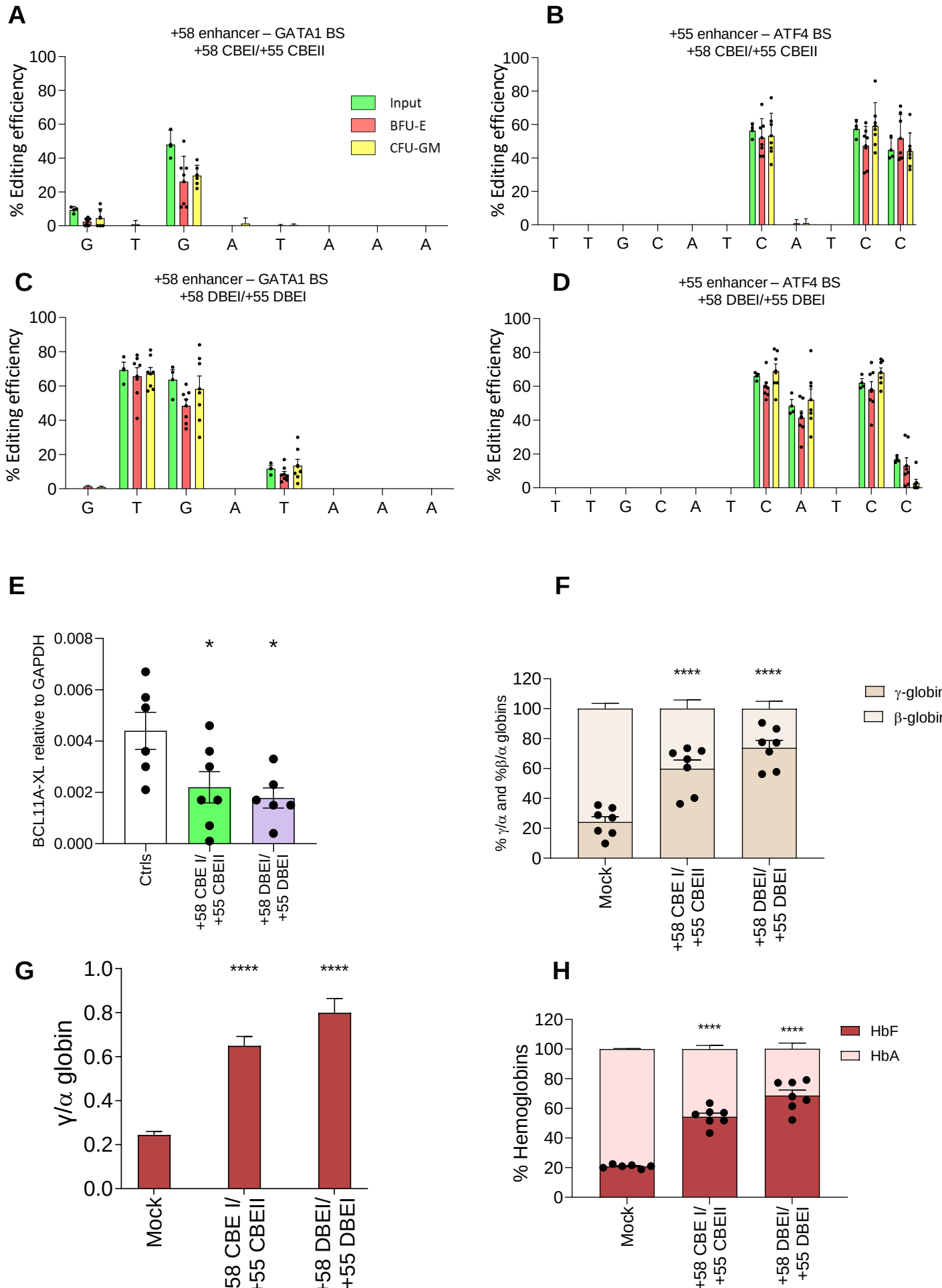
+58 enhancer – GATA1 BS



+55 enhancer – ATF4 BS



Supplementary Figure 9



Supplementary Figure 10