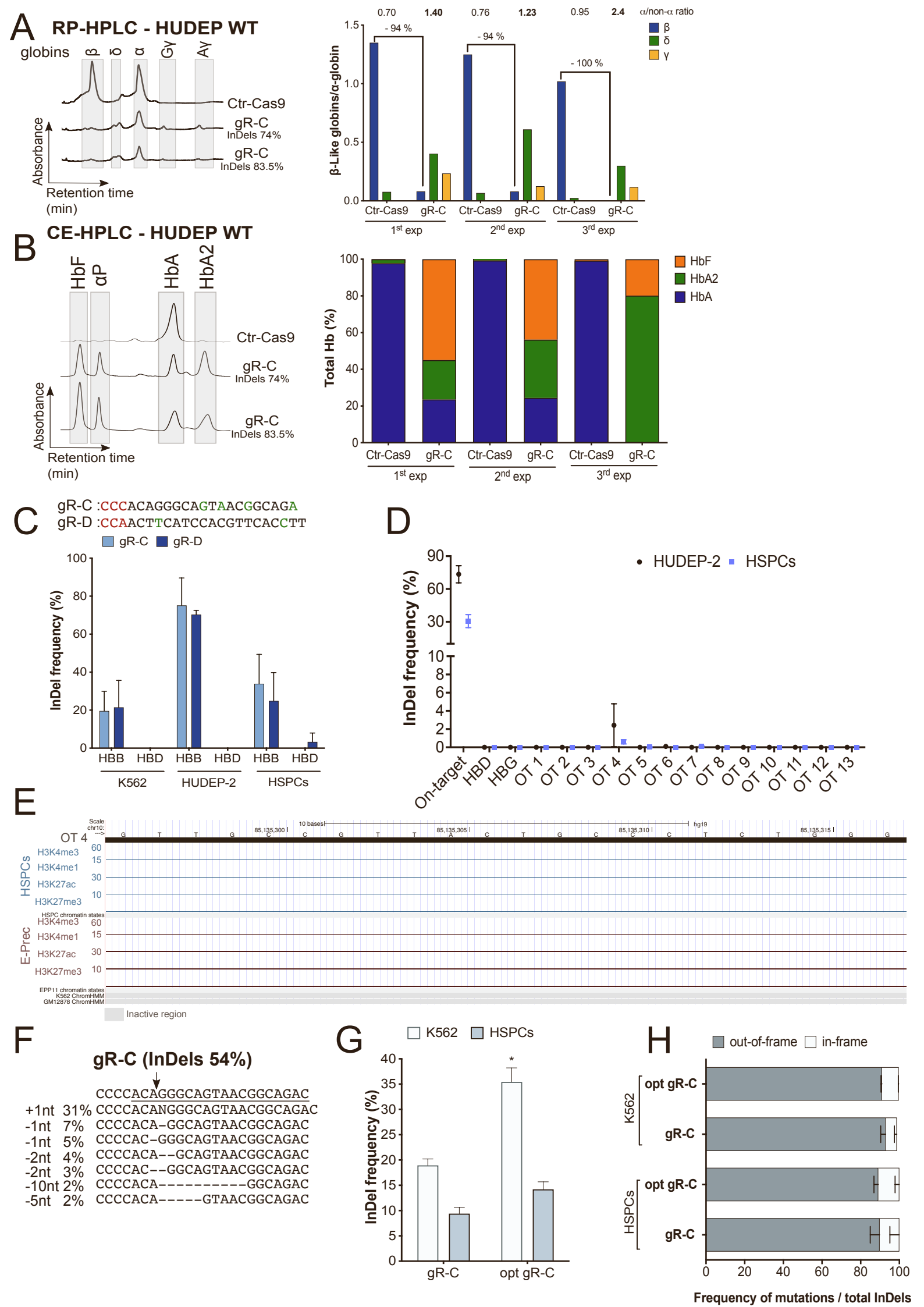
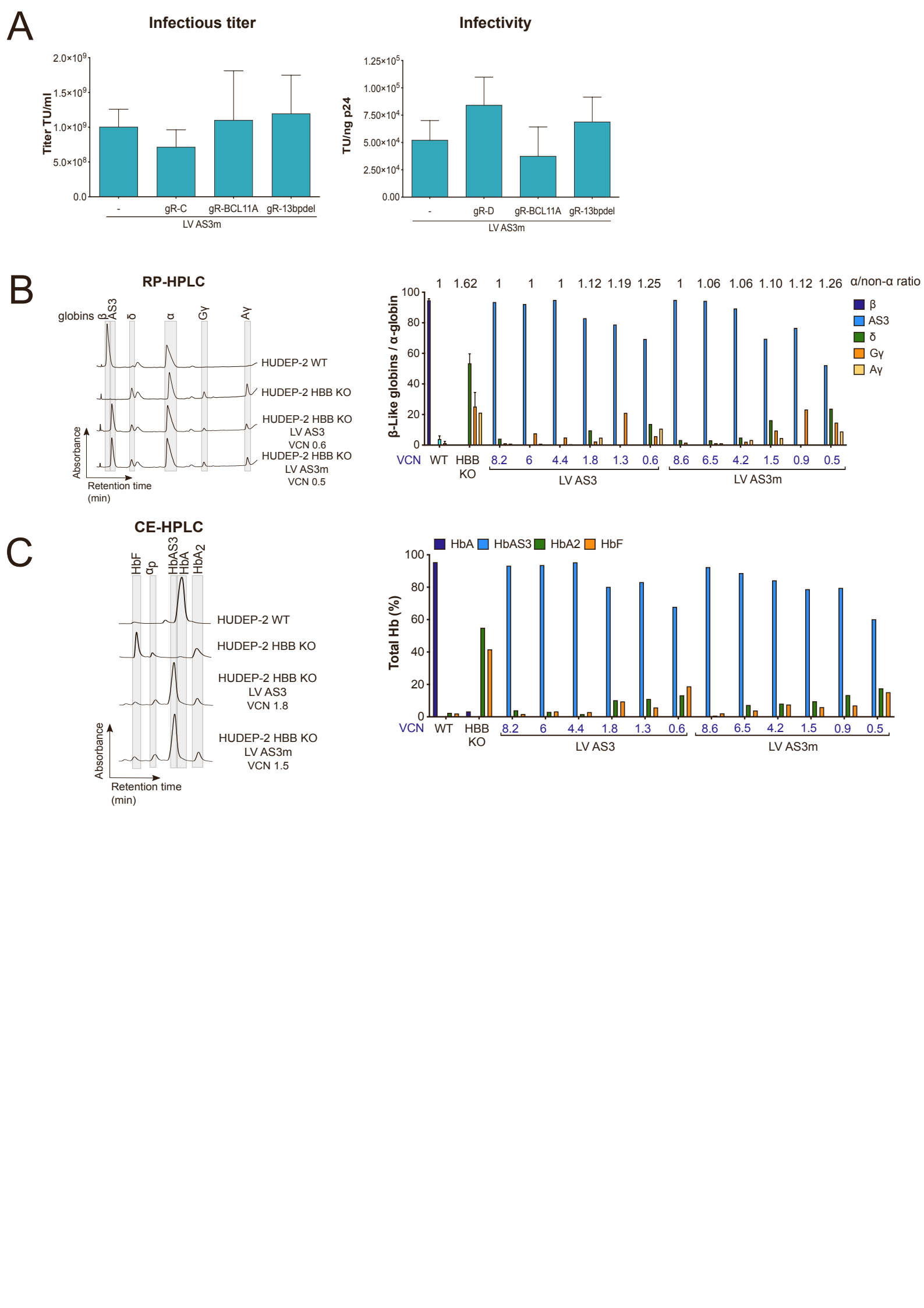
YMTHE, Volume *30*

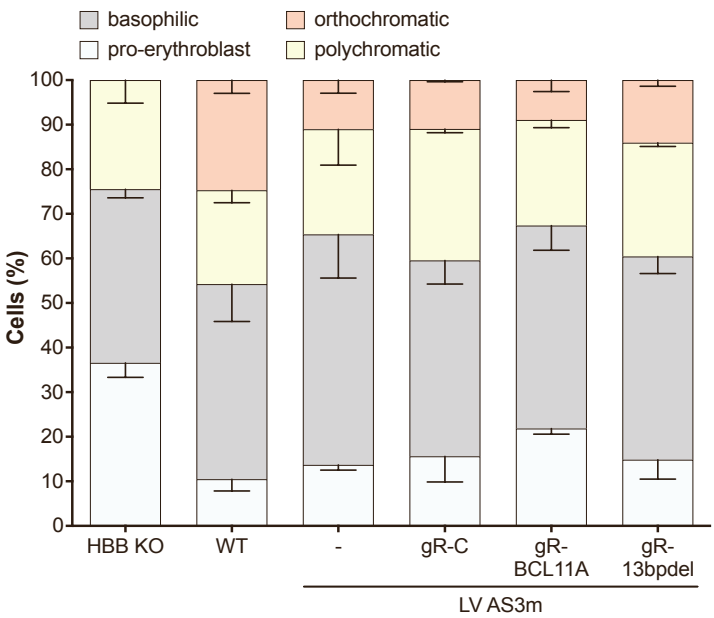
Supplemental Information

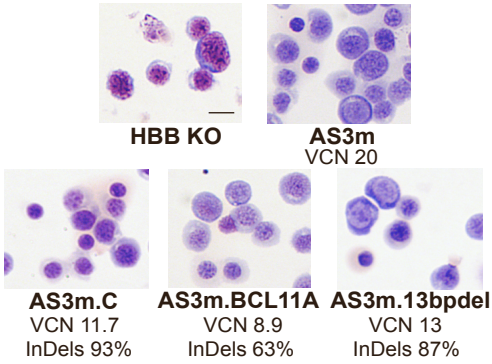
Combination of lentiviral and genome editing technologies for the treatment of sickle cell disease

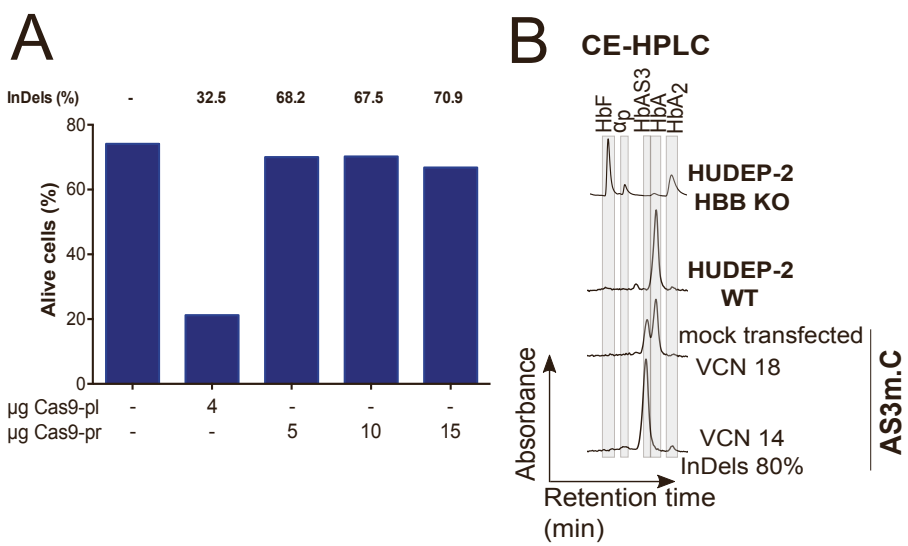
Sophie Ramadier, Anne Chalumeau, Tristan Felix, Nadia Othman, Sherazade Aknoun, Antonio Casini, Giulia Maule, Cecile Masson, Anne De Cian, Giacomo Frati, Megane Brusson, Jean-Paul Concordet, Marina Cavazzana, Anna Cereseto, Wassim El Nemer, Mario Amendola, Benoit Wattellier, Vasco Meneghini, and Annarita Miccio

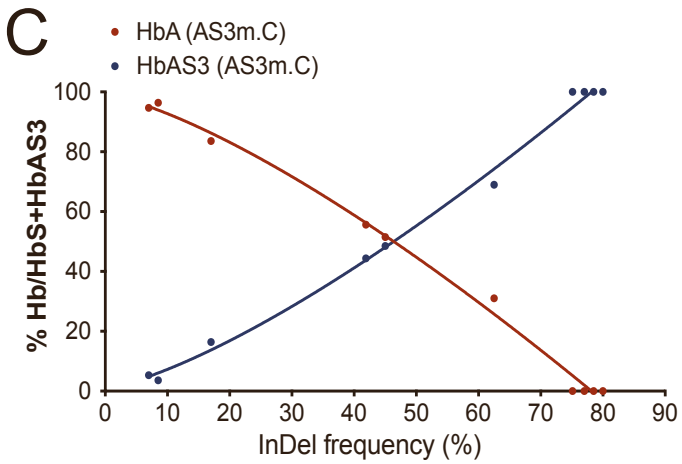

A
RP-HPLC - HUDEP WT
globins
β
δ
α
Gγ
Aγ
Ctr-Cas9
gR-C
InDels 74%
gR-C
InDels 83.5%
Absorbance
Retention time
(min)
0.70
1.40
0.76
1.23
0.95
2.4
α/non-α ratio
β
δ
γ
- 94 %
- 94 %
- 100 %
β-Like globins/α-globin
Ctr-Cas9
gR-C
1st exp
2nd exp
3rd exp
B
CE-HPLC - HUDEP WT
HbF
αP
HbA
HbA2
Ctr-Cas9
gR-C
InDels 74%
gR-C
InDels 83.5%
Absorbance
Retention time
(min)
HbF
HbA2
HbA
Total Hb (%)
C
gR-C :CCCACAGGGCAGTAACGGCAGA
gR-D :CCAACTTCATCCACGTTCACCTT
gR-C
gR-D
InDel frequency (%)
HBB
HBD
K562
HUDEP-2
HSPCs
D
HUDEP-2
HSPCs
InDel frequency (%)
On-target
HBD
HBG
OT 1
OT 2
OT 3
OT 4
OT 5
OT 6
OT 7
OT 8
OT 9
OT 10
OT 11
OT 12
OT 13
E
OT 4
HSPCs
H3K4me3
H3K4me1
H3K27ac
H3K27me3
HSPC chromatin states
E-Prec
EPP11 chromatin states
K562 ChromHMM
GM12878 ChromHMM
Inactive region
F
gR-C (InDels 54%)
CCCCACAGGGCAGTAACGGCAGAC
+1nt 31% CCCCACANGGGCAGTAACGGCAGAC
-1nt 7% CCCCACA-GGCAGTAACGGCAGAC
-1nt 5% CCCCAC-GGGCAGTAACGGCAGAC
-2nt 4% CCCCACA--GCAGTAACGGCAGAC
-2nt 3% CCCCAC--GGCAGTAACGGCAGAC
-10nt 2% CCCCACA----------GGCAGAC
-5nt 2% CCCCACA-----GTAACGGCAGAC
G
K562
HSPCs
InDel frequency (%)
gR-C
opt gR-C
H
out-of-frame
in-frame
K562
opt gR-C
gR-C
HSPCs
opt gR-C
gR-C
Frequency of mutations / total InDels

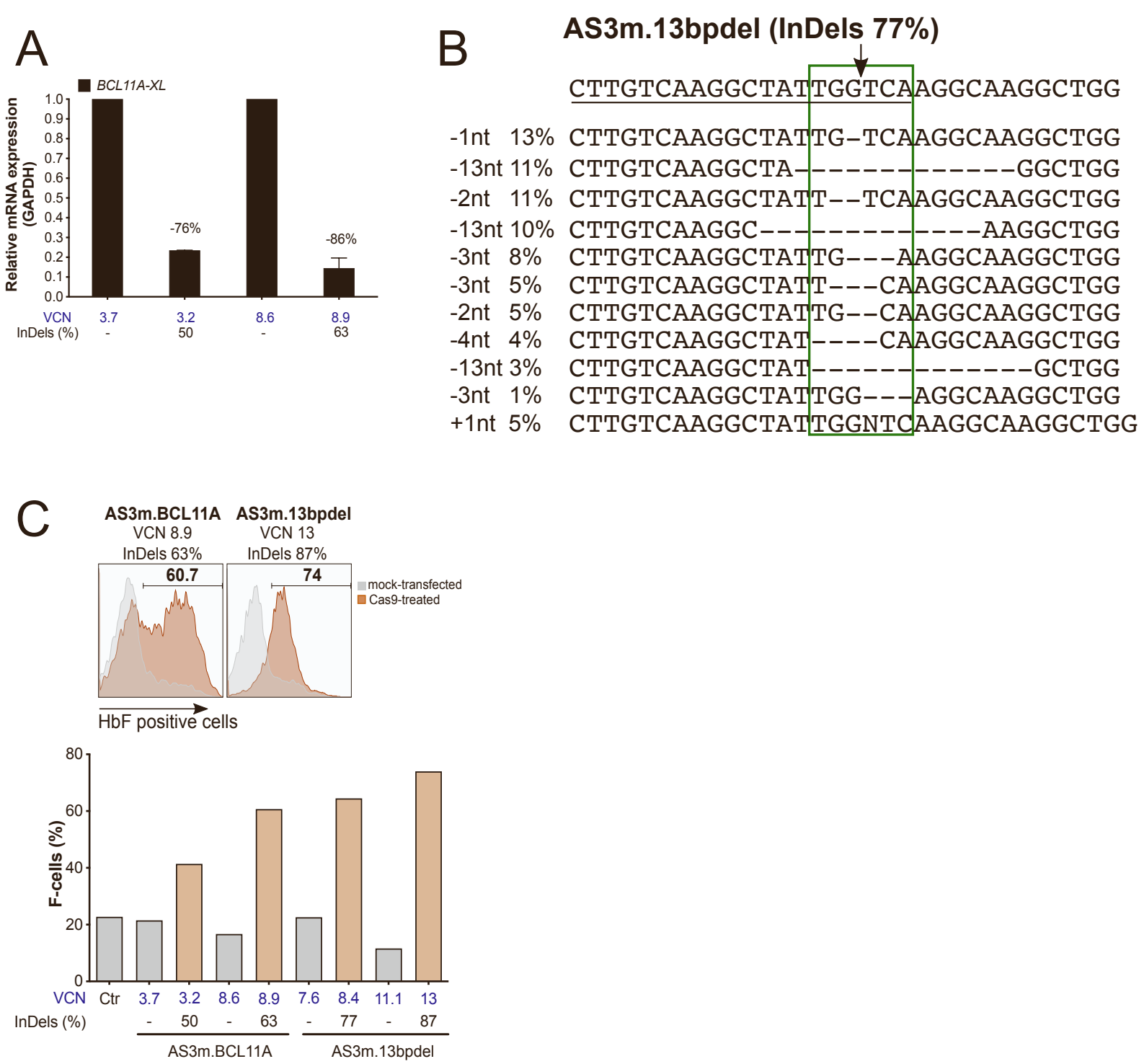

A
Infectious titer
Titer TU/ml
2.0×10^9
1.5×10^9
1.0×10^9
5.0×10^8
0.0
-
gR-C
gR-BCL11A
gR-13bpdel
LV AS3m
Infectivity
TU/ng p24
1.25×10^5
1.00×10^5
7.50×10^4
5.00×10^4
2.50×10^4
0.00
-
gR-D
gR-BCL11A
gR-13bpdel
LV AS3m
B
RP-HPLC
globins
β AS3
δ
α
Gγ
Aγ
HUDEP-2 WT
HUDEP-2 HBB KO
HUDEP-2 HBB KO LV AS3 VCN 0.6
HUDEP-2 HBB KO LV AS3m VCN 0.5
Absorbance
Retention time (min)
β-Like globins / α-globin
α/non-α ratio
1 1.62 1 1 1 1.12 1.19 1.25 1 1.06 1.06 1.10 1.12 1.26
β
AS3
δ
Gγ
Aγ
VCN WT HBB KO 8.2 6 4.4 1.8 1.3 0.6 8.6 6.5 4.2 1.5 0.9 0.5
LV AS3
LV AS3m
C
CE-HPLC
HbF
αp
HbAS3
HbA
HbA2
HUDEP-2 WT
HUDEP-2 HBB KO
HUDEP-2 HBB KO LV AS3 VCN 1.8
HUDEP-2 HBB KO LV AS3m VCN 1.5
Absorbance
Retention time (min)
Total Hb (%)
HbA HbAS3 HbA2 HbF
VCN WT HBB KO 8.2 6 4.4 1.8 1.3 0.6 8.6 6.5 4.2 1.5 0.9 0.5
LV AS3
LV AS3m


basophilic
orthochromatic
pro-erythroblast
polychromatic
Cells (%)
100
90
80
70
60
50
40
30
20
10
0
HBB KO
WT
-
gR-C
gR-BCL11A
gR-13bpdel
LV AS3m

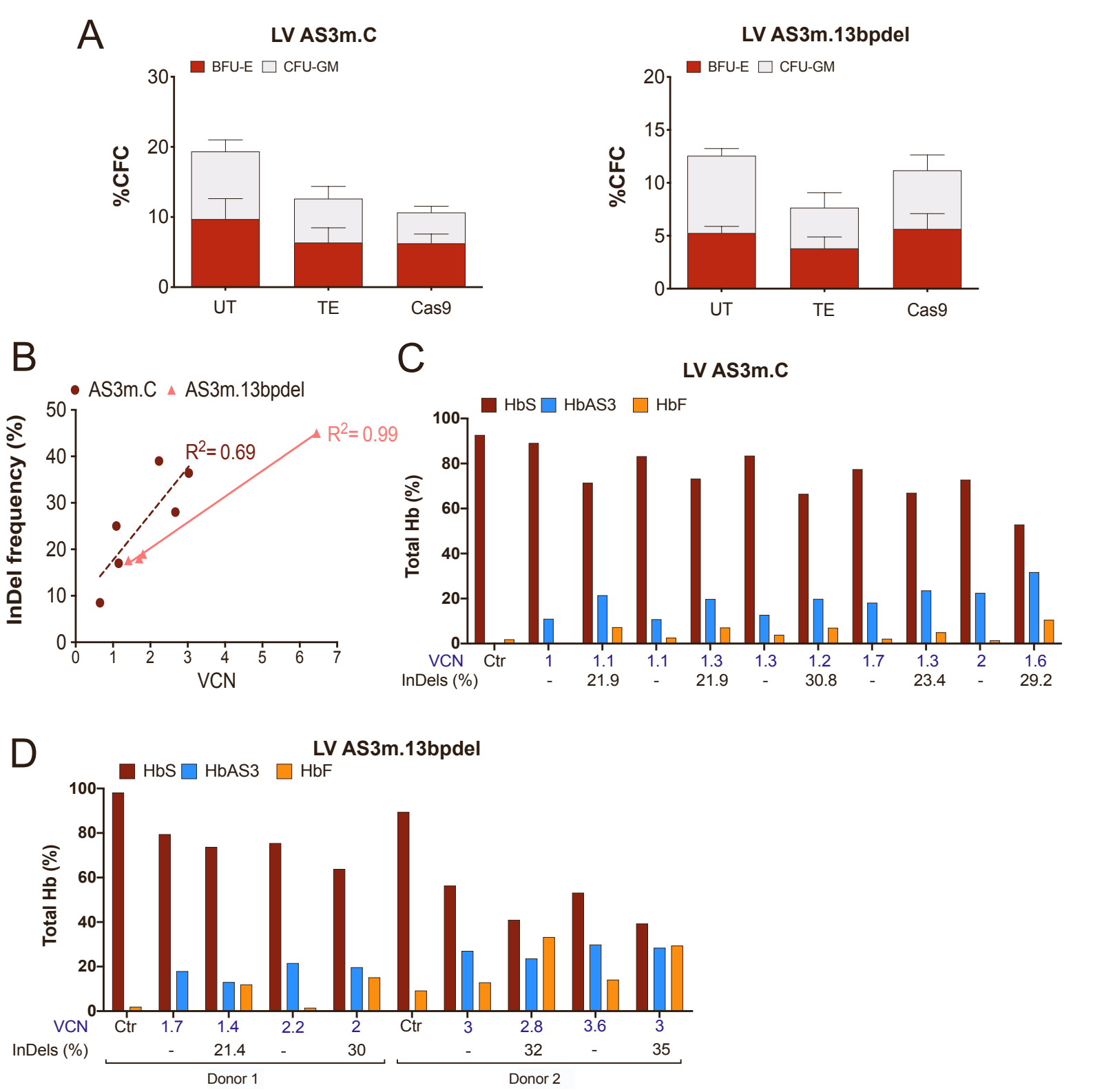

HBB KO
AS3m
VCN 20
AS3m.C
VCN 11.7
InDels 93%
AS3m.BCL11A
VCN 8.9
InDels 63%
AS3m.13bpdel
VCN 13
InDels 87%

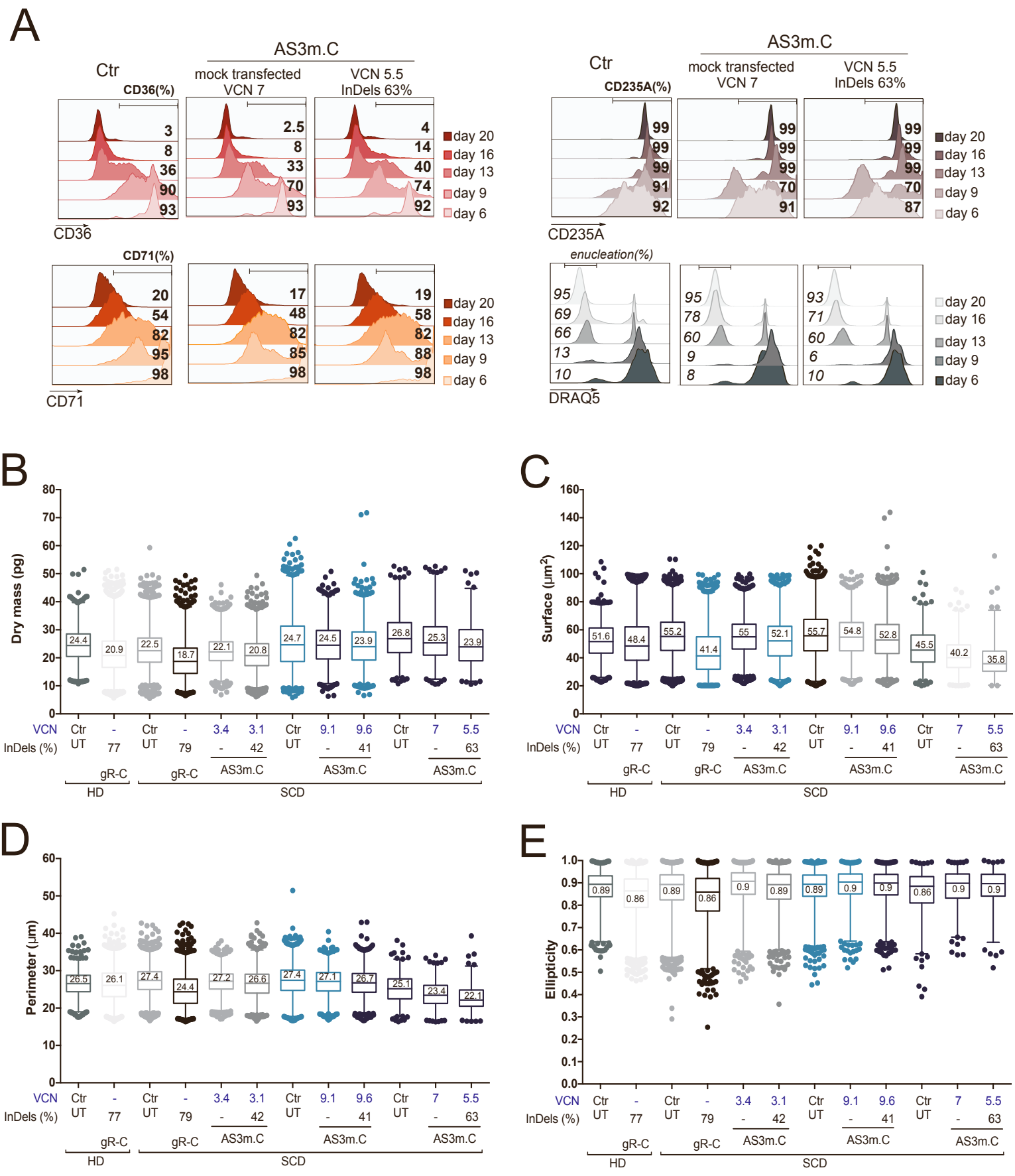

A
InDels (%) - 32.5 68.2 67.5 70.9
Alive cells (%)
0
20
40
60
80
µg Cas9-pl - 4 - - -
µg Cas9-pr - - 5 10 15
B
CE-HPLC
HbF
αp
HbAS3
HbA
HbA2
HUDEP-2
HBB KO
HUDEP-2
WT
mock transfected
VCN 18
VCN 14
InDels 80%
Absorbance
Retention time
(min)
AS3m.C
HbA (mock transfected)
HbAS3 (mock transfected)
HbAS3 (Cas9 transfected)
HbA (Cas9 transfected)
% Hb/HbS+HbAS3
R2= 0.95
R2= 0.88
R2= 0.88
R2= 0.95
VCN

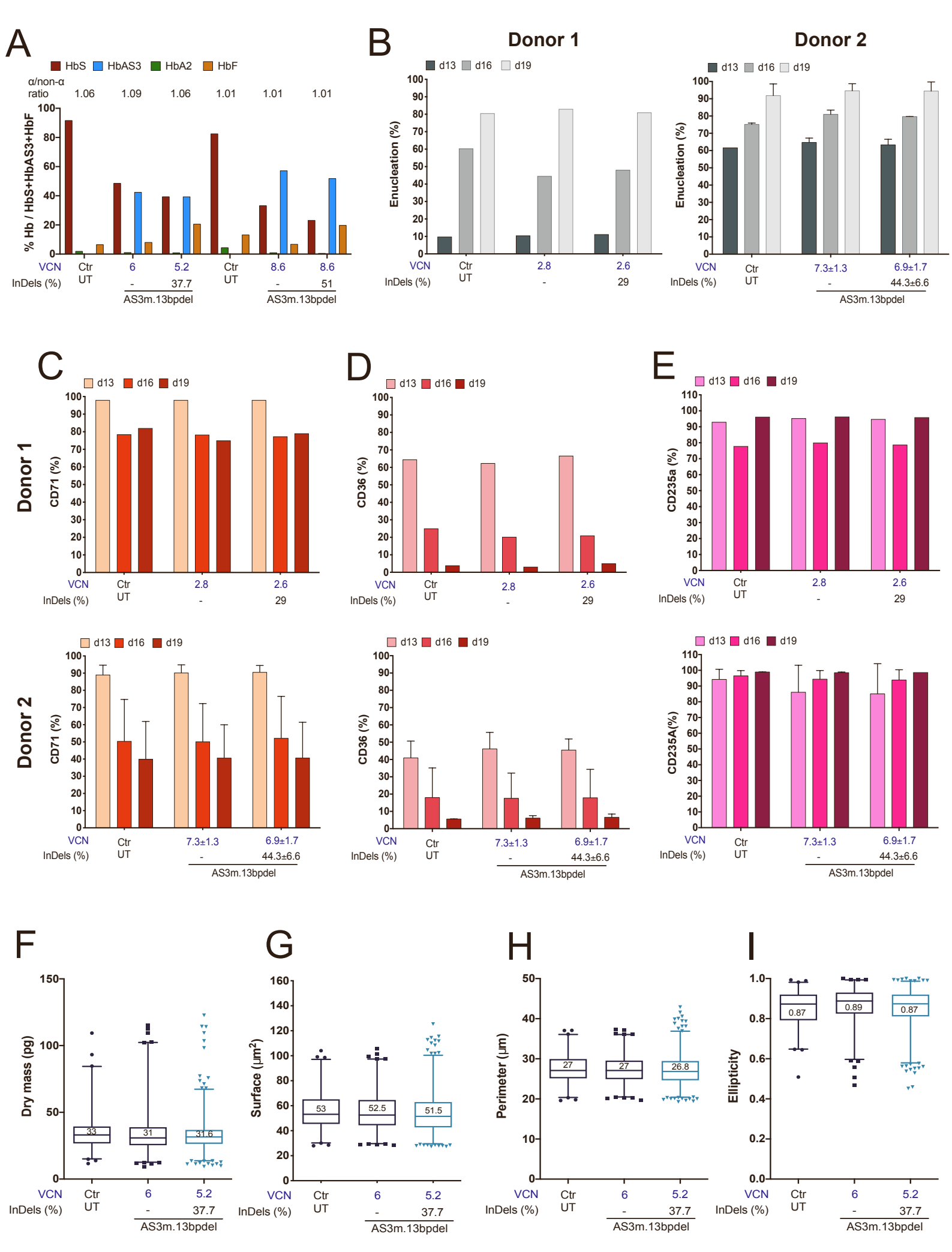

C
HbA (AS3m.C)
HbAS3 (AS3m.C)
% Hb/HbS+HbAS3
InDel frequency (%)


A
BCL11A-XL
Relative mRNA expression (GAPDH)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-76%
-86%
VCN 3.7 3.2 8.6 8.9
InDels (%) - 50 - 63
B
AS3m.13bpdel (InDels 77%)
CTTGTCAAGGCTATTGGTCAAGGCAAGGCTGG
-1nt 13% CTTGTCAAGGCTATTG-TCAAGGCAAGGCTGG
-13nt 11% CTTGTCAAGGCTA----------GGCTGG
-2nt 11% CTTGTCAAGGCTATT--TCAAGGCAAGGCTGG
-13nt 10% CTTGTCAAGGC----------AAGGCTGG
-3nt 8% CTTGTCAAGGCTATTG---AAGGCAAGGCTGG
-3nt 5% CTTGTCAAGGCTATT---CAAGGCAAGGCTGG
-2nt 5% CTTGTCAAGGCTATTG--CAAGGCAAGGCTGG
-4nt 4% CTTGTCAAGGCTAT----CAAGGCAAGGCTGG
-13nt 3% CTTGTCAAGGCTAT----------GCTGG
-3nt 1% CTTGTCAAGGCTATTGG---AGGCAAGGCTGG
+1nt 5% CTTGTCAAGGCTATTGGNTCAAGGCAAGGCTGG

C
AS3m.BCL11A
VCN 8.9
InDels 63%
AS3m.13bpdel
VCN 13
InDels 87%
60.7
74
mock-transfected
Cas9-treated
HbF positive cells
F-cells (%)
80
60
40
20
0
VCN Ctr 3.7 3.2 8.6 8.9 7.6 8.4 11.1 13
InDels (%) - 50 - 63 - 77 - 87
AS3m.BCL11A
AS3m.13bpdel


A
LV AS3m.C
BFU-E
CFU-GM
%CFC
30
20
10
0
UT
TE
Cas9
LV AS3m.13bpdel
BFU-E
CFU-GM
%CFC
20
15
10
5
0
UT
TE
Cas9
B
AS3m.C
AS3m.13bpdel
InDel frequency (%)
50
40
30
20
10
0
R2= 0.69
R2= 0.99
0
1
2
3
4
5
6
7
VCN
C
LV AS3m.C
HbS
HbAS3
HbF
Total Hb (%)
100
80
60
40
20
0
VCN
Ctr
1
1.1
1.1
1.3
1.3
1.2
1.7
1.3
2
1.6
InDels (%)
-
21.9
-
21.9
-
30.8
-
23.4
-
29.2
D
LV AS3m.13bpdel
HbS
HbAS3
HbF
Total Hb (%)
100
80
60
40
20
0
VCN
Ctr
1.7
1.4
2.2
2
Ctr
3
2.8
3.6
3
InDels (%)
-
21.4
-
30
-
32
-
35
Donor 1
Donor 2


A
AS3m.C
Ctr
mock transfected VCN 7
VCN 5.5 InDels 63%
CD36(%)
3
8
36
90
93
2.5
8
33
70
93
4
14
40
74
92
day 20
day 16
day 13
day 9
day 6
CD36
CD71(%)
20
54
82
95
98
17
48
82
85
98
19
58
82
88
98
CD71
CD235A(%)
99
99
99
91
92
99
99
99
70
91
99
99
99
70
87
CD235A
enucleation(%)
95
69
66
13
10
95
78
60
9
8
93
71
60
6
10
DRAQ5
B
Dry mass (pg)
C
Surface (µm²)
D
Perimeter (µm)
E
Ellipticity
VCN
InDels (%)
Ctr UT
77
79
3.4
3.1
42
9.1
9.6
41
7
5.5
63
gR-C
AS3m.C
HD
SCD


A
HbS
HbAS3
HbA2
HbF
α/non-α ratio
1.06
1.09
1.06
1.01
1.01
1.01
% Hb / HbS+HbAS3+HbF
VCN
Ctr UT
6
5.2
Ctr UT
8.6
8.6
InDels (%)
-
37.7
-
51
AS3m.13bpdel
AS3m.13bpdel
B
Donor 1
Donor 2
d13
d16
d19
Enucleation (%)
VCN
Ctr UT
2.8
2.6
InDels (%)
-
29
7.3±1.3
6.9±1.7
-
44.3±6.6
AS3m.13bpdel
C
D
E
Donor 1
Donor 2
CD71 (%)
CD36 (%)
CD235a (%)
CD235A(%)
VCN
Ctr UT
2.8
2.6
InDels (%)
-
29
7.3±1.3
6.9±1.7
-
44.3±6.6
AS3m.13bpdel
F
G
H
I
Dry mass (pg)
Surface (µm²)
Perimeter (µm)
Ellipticity
33
31
31.6
53
52.5
51.5
27
27
26.8
0.87
0.89
0.87
VCN
Ctr UT
6
5.2
InDels (%)
-
37.7
AS3m.13bpdel

Supplementary Figure Legends

Supplementary figure 1: Efficiency and safety analysis of sgRNAs knocking down the *HBB* gene.

(A) RP-HPLC analysis of globin chains in differentiated HUDEP-2 cells transfected with Cas9-GFP plasmid only (Ctr-Cas9) or with both Cas9-GFP and gR-C plasmids (gR-C). Representative chromatograms are reported in the left panel. Quantification is reported in the right panel. β-like globin chains were normalized to the α-globin chains. The percentage of β-globin decrease is reported above the histogram bars. α-/non-α ratios are reported on top of each sample (n=3; genome editing efficiency of 80.1±3.0 (mean±SEM)). **(B)** CE-HPLC quantification of Hb tetramers in HUDEP-2 cells transfected with Cas9-GFP plasmid only (Ctr-Cas9) or with both Cas9-GFP and gR-C plasmids (gR-C). Representative chromatograms are reported in the left panel. Quantification is reported in the right panel. We calculated the percentage of each Hb type over the total Hb tetramers. αP=α-precipitates **(C)** On- and off-target activity evaluated by TIDE in *HBB* and *HBD* in K562 (n=6 for gR-C and n=5 for gR-D in *HBB;* n=6 for gR-C and gR-D in *HBD*), HUDEP-2 (n=4 for gR-C and gR-D in *HBB* and *HBD*) and cord blood (CB) HSPCs edited with gR-C (n=3) or D (n=2) . Sequences of gR-C and D are reported above the graph; mismatches between the sgRNA target sequence and the potential off-target in *HBD* are indicated in green. PAM is indicated in red. Data are expressed as mean±SEM. **(D)** Editing efficiency of the *HBB* (On-target), *HBD*, *HBG* genes and the top 13 off-target sites (OT) identified by COSMID in HUDEP-2 (n=4 for *HBB* and *HBD*; n=3 for *HBG*, OT 1. OT 3, OT 4, OT 6, OT 8, OT 9, OT 11 and OT 12; n= 2 for OT 2, OT 5 , OT 7, OT 10 and OT 13) and CB HSPCs edited with gR-C (n=5), as measured by TIDE. Data are expressed as mean±SEM. Off-target editing was detected only for OT 4. Its closest genes are Neuregulin 3 (NRG3) (located 300 kb far from off-target 4), which is mainly expressed in CNS tissues, and AL356140.1 miRNA (located 68 kb far from the off-target site) with unknown functions and potentially involved in angiogenesis. **(E)** Analysis of the chromatin states at the OT 4 sequence in primary human HSPCs

and HSPC-derived erythroid precursors (E-Prec)[1], and in erythroid (K562) and granulo-monocytic (GM12878) cell lines (UCSC datasets). OT 4 maps to an inactive region in all the cell types. **(F)** Representative Sanger sequencing analysis of edited *HBB* alleles in G-CSF mPB HD HSPCs (54% of total InDels) using ICE Analysis[2]. The top line shows the unmodified *HBB* gene sequence. gR-C-targeted sequence is underlined. The arrow indicates the cleavage site. Dashes and "N" indicate deleted and inserted nucleotides, respectively. InDel type, length and frequency are indicated on the left. **(G)** Editing efficiency and **(H)** out-of-frame mutations evaluated by TIDE analysis after PCR amplification of the target region and Sanger sequencing in K562 (n=2; unpaired t test *$p<0.05$) and mobilized peripheral blood HD HSPCs edited with gR-C harboring the original (gR-C; n=3) or the optimized (opt gR-C; n=4) scaffold. Data are expressed as mean±SEM.

Supplementary figure 2: Bifunctional LVs expressing an *HBB* modified transgene.

(A) Histograms showing: (i) the infectious titer (left panel) of LV AS3m "-" (n=6) and LV AS3m containing gR-C (n=6), gR-BCL11A (n=5) and gR-13bpdel (n=5); (ii) infectivity (right panel) of LV AS3m "-" (n=4) and LV AS3m containing gR-C (n=6), gR-BCL11A (n=3) and gR-13bpdel (n=3). Titer and infectivity were evaluated in HCT116 cells. The infectivity was measured as a ratio between the infectious (TU/ml) and the physical titer (ng p24/ml). Data are expressed as mean±SEM. **(B)** RP-HPLC chromatograms (left panel) and quantification (right panel) of globin chains in LV AS3- and LV AS3m-treated cells. WT HUDEP-2 and HBB KO HUDEP-2 served as controls. β-like globin chains are normalized to α-globin chains. VCN are reported in blue below the graph. α/non-α ratios are reported above the histogram bars. **(C)** CE-HPLC chromatograms (left panel) and quantification (right panel) of Hb tetramers in LV AS3- and LV AS3m-treated cells. WT HUDEP-2 and HBB KO HUDEP-2 served as controls. We plotted the percentage of each Hb type over the total Hbs. VCN are reported in blue below the graph.

Supplementary figure 3: LV-treated-HUDEP-2 cells can differentiate into mature erythroblasts.

Quantification (left panel) and representative photomicrographs (right panel) of HUDEP-2 erythroid precursors obtained at day 9 of the erythroid differentiation and stained with May-Grünwald Giemsa. HBB KO, WT and mock-transfected HUDEP-2 cells served as controls. HUDEP-2 HBB KO cells display a higher percentage of immature pro-erythroblasts and absence of more mature orthochromatic cells compared to WT cells. Data are expressed as mean±SEM. Scale bar, 40 µm (upper left).

Supplementary figure 4: Superior efficiency of Cas9-GFP protein compared to Cas9-GFP plasmid in LV AS3m.C-transduced HUDEP-2 cells.

(A) A representative experiment showing the percentage of alive HUDEP-2 cells measured by flow cytometry 24 hours after transfection with either Cas9-GFP plasmid (4 µg) or Cas9-GFP protein (5-15 µg). InDel frequency is reported above the histograms. The proportion of alive cells was significantly lower in plasmid-transfected samples (37±6%) compared to untransfected cells (68±4%; $p<0.005$) or to samples transfected with 15 µg of Cas9-GFP protein (58±4%; $p<0.05$). No significative differences were observed between untransfected and Cas9-GFP protein-treated samples (n=3-4). **(B)** CE-HPLC chromatograms (left panel) and correlation between VCN and percentage of HbA or HbAS3 over the total adult tetramers (HbA+HbAS3) (right panel) in LV AS3m.C-transduced HUDEP-2 cells transfected with 15 µg of Cas9-GFP protein or mock-transfected. R^2 and line-of-best-fit equation are indicated. **(C)** Correlation between InDel frequency and percentage of HbA or HbAS3 over the total adult tetramers (HbA+HbAS3) in LV AS3m.C-transduced HUDEP-2 cells transfected with Cas9-GFP protein or mock-transfected. R^2 and line-of-best-fit equation are indicated.

Supplementary figure 5: Bifunctional LVs reactivate HbF expression.

(A) Relative expression of *BCL11A-XL* normalized to GAPDH, as measured by qRT-PCR in LV AS3m.BCL11A-transduced HUDEP-2 cells treated with Cas9-GFP plasmid or mock-transfected control ("-"). VCN are reported in blue and InDels in black below the graph. Data are expressed as mean ± SEM. **(B)** Representative Sanger sequencing analysis of edited *HBG* promoters in HUDEP-2 cells (77% of total InDels) using ICE Analysis[2]. The top line shows the unmodified *HBG1/2* promoter sequence. sgRNA 13bpdel-targeted sequence is underlined. The arrow indicates the cleavage site. The BCL11A binding site is highlighted with a green rectangle. Dashes and "N" indicate deleted and inserted nucleotides, respectively. InDel type, length and frequency are indicated on the left. **(C)** Representative flow cytometry plots (upper panel) and quantification (bottom panel) of HbF-expressing cells in untreated HUDEP-2 cells (Ctr), Cas9-treated and mock-transfected "-" HUDEP-2 cells transduced either with LV AS3m.BCL11A or LVAS3m.13bpdel. VCN are reported in blue and InDels in black below the graph.

Supplementary figure 6: Correlation between VCN and editing frequency in BFU-E transduced with bifunctional LVs

(A) Frequency of CFC in edited SCD HSPCs (Cas9). LV AS3m.C and LV AS3m.13bpdel-transduced SCD HSPCs transfected only with TE (TE) or untransfected (UT) were used as control. Data are expressed as mean±SEM (n=4-9; 2 mobilized SCD donors). **(B)** VCN and InDel frequency were determined in burst forming unit-erythroid (BFUE) derived from SCD HSPCs transduced either with LV AS3m.C or LV AS3m.13bpdel and transfected with Cas9-GFP protein. R^2 and line-of-best-fit equation are indicated. **(C)** CE-HPLC quantification of Hb tetramers in RBCs derived from untreated SCD HSPCs (Ctr) and LV AS3m.C-treated SCD HSPCs (1 donor) that were mock-transfected ("-") or transfected with Cas9-GFP protein. We plotted the percentage of each Hb type over the total Hb tetramers. VCN are reported in blue and InDels in black below the graph. **(D)** CE-HPLC quantification of Hb tetramers in RBCs obtained from untreated SCD HSPCs (Ctr) and LV AS3m.13bpdel-treated SCD HSPCs (2 donors) that were mock-transfected ("-") or transfected with

Cas9-GFP protein. We plotted the percentage of each Hb type over the total Hb tetramers. VCN are reported in blue and InDels in black below the graph.

Supplementary figure 7: Erythroid differentiation and RBC parameters in cells derived from LV AS3m.C-edited SCD HSPCs

(A) Representative flow cytometry analysis of the early (CD36 and CD71) and late (CD235A) erythroid markers at day 6, 9, 13, 16 and 20 of erythroid differentiation of untreated SCD cells (Ctr) and LV AS3m.C-treated SCD HSPCs that were either mock-transfected or transfected with Cas9-GFP protein. The enucleation rate was measured using DRAQ5 nuclear staining. **(B-E)** RBC parameters extracted using the BIO-Data software. RBCs were obtained after 19 days of differentiation from SCD HSPCs transduced with LV AS3m.C and either mock- or Cas9-transfected. As controls, we used untreated SCD RBCs (Ctr UT), and RBCs obtained from SCD/HD HSPCs transfected with RNPs containing gR-C (gR-C). For each population, data were normalized to the total number of RBCs and are reported as boxplots showing quartiles, median and outliers. Median value is indicated inside the boxplot. VCN are reported in blue and InDels in black below the graph. (**B**) Dry mass (pg). (**C**) Surface (μm^2). (**D**) Perimeter (µm). (**E**) Ellipticity.

Supplementary figure 8: Erythroid markers and RBC parameters were not impaired in cells derived from LV AS3m.13bpdel-edited SCD HSPCs.

(A) CE-HPLC quantification of Hb tetramers in untreated SCD cells (Ctr UT), and LV AS3m.13bpdel-treated SCD HSPCs that were either mock-transfected "-" or transfected with Cas9-GFP protein. We plotted the percentage of each Hb type over the total Hbs. VCN are reported in blue and InDels in black below the graph. The α-/non-α-globin ratios evaluated by RP-HPLC are reported on top of the histograms. **(B-E)** Flow cytometry analysis of the enucleation rate and of the early (CD71 and CD36) and late (CD235A) erythroid markers at day 13, 16 and 19 of erythroid differentiation of untreated SCD HSPCs (Ctr UT) and LV AS3m.13bpdel-treated SCD HSPCs that

were either mock-transfected or transfected with Cas9-GFP protein (n=3 biological replicates. 2 donors). VCN and InDels values are reported below the graph as mean±SEM. (**B**) Enucleation rate measured using DRAQ5 nuclear staining. (**C-E**) Proportion of $CD71^+$, $CD36^+$ and $CD235A^+$ cells during erythroid differentiation. **(F-I)** RBC parameters extracted using the BIO-Data software. RBCs were obtained after 19 days of differentiation from SCD HSPCs transduced with LV AS3m.13bpdel and either mock- or Cas9-transfected. As control, we used untreated SCD RBCs (Ctr UT). For each population, data were normalized to the total number of RBCs belonging and are reported as boxplots showing quartiles, median and outliers. Median value is indicated inside the boxplot. VCN are reported in blue and InDels in black below the graph. (**F**) Dry mass (pg). (**G**) Surface (μm^2). (**H**) Perimeter (µm). (**I**) Ellipticity.

Supplementary Methods

Plasmid construction

Plasmids expressing a Cas9-GFP fusion protein (pMJ920) and sgRNA (MLM3636) were purchased from Addgene (plasmids #42234 and #43860). The MA128 plasmid, containing the optimized sgRNA scaffold[3], was kindly provided by Dr. Amendola. The list of the sgRNA target sequences is provided in **Supplementary table 1.**

sgRNA name	Target sequence + PAM (5' to 3')	Strand	Hg19 genomic location
gR-A	gCTTGCCCCACAGGGCAGTAA**CGG**	-	chr11:5,226,968-5,226,987
gR-B	GTAACGGCAGACTTCTCCTC**AGG**	-	chr11:5,226,980-5,227,009

gR-C	gTCTGCCGTTACTGCCCTGT**GGG**	+	chr11:5,226,973-5,226,994
gR-D	gAAGGTGAACGTGGATGAAGT**TGG**	+	chr11:5,226,948-5,226,970

Supplementary Table 1: List of the sgRNAs targeting *HBB* exon 1. When the first nucleotide of the target sequence was different from a guanosine, we inserted a guanosine (indicated in lowercase) to allow U6-driven sgRNA expression. PAM sequences are highlighted in bold. For each sgRNA, we reported the hg19 genomic coordinates. gR-A sequence was retrieved from [4], gR-B sequence from [5], and gR-C and gR-D sequenced from [6].

HUDEP-2 and K562 cultures and differentiation

K562 were maintained in RPMI 1640 medium (Lonza) containing 2 mM glutamine and supplemented with 10% fetal bovine serum (FBS, BioWhittaker, Lonza), HEPES (20 mM, LifeTechnologies), sodium pyruvate (1 mM, LifeTechnologies) and penicillin/streptomycin (100U/ml each, LifeTechnologies).

HUDEP-2 cells[7] were cultured and differentiated for 8-9 days as previously described[8]. A standard May-Grumwald Giemsa staining was performed to evaluate the cell morphology during HUDEP-2 differentiation. Two fields per condition and a total of around 280 cells were counted. Erythroid differentiation was monitored during the culture by flow cytometry analysis.

HUDEP-2 thalassemic cells (HBB KO) were generated by transfecting HUDEP-2 WT cells with plasmids harboring the gR-D targeting *HBB* exon 1 and the Cas9-GFP. FACS-sorted GFP^+ cells were cloned by limiting dilution. After 14 days, single clones were isolated, expanded and screened by PCR and Sanger sequencing of the target sequence, followed by TIDE analysis. We selected a HUDEP-2 clone harboring bi-allelic frameshift mutations in the exon 1 of *HBB* gene and characterized by the absence of β-globin chain expression.

Lentiviral transduction of erythroid HUDEP-2 cell lines

Cells were transduced ($1x10^6$ cells/ml) for 24 h in the proliferation medium[8] supplemented with protamine sulfate (4 µg/ml, Sigma-Aldrich or APP Pharmaceuticals).

Plasmid transfection of erythroid cell lines and HSPCs

One million cells were transfected with 4 ug of the Cas9-GFP plasmid (pMJ920) and 0.8-1.6 µg of each sgRNA plasmid (MLM3636 or MA128) in a 100 µl volume using Nucleofector I (Lonza). We used AMAXA Cell Line Nucleofector Kit V (VCA-1003) for K562 and HUDEP-2 (T-16/U-16 and L-29 programs, respectively), and AMAXA Human CD34 Cell Nucleofector Kit (VPA-1003; U-08 program) for HD HSPCs. GFP^+ HUDEP-2 cells and GFP^+ HSPCs were FACS-sorted using SH800 Cell Sorter (Sony Biotechnology). As control, we used untreated cells, cell transfected only with Cas9-GFP plasmid or cells transfected only with nuclease-free TE buffer (10mM Tris and 1mM EDTA, pH 8).

CFC Assay

In the Colony Forming Cell (CFC) assay, SCD HSPCs were plated at a concentration of $2x10^3$ cells/ml in a methylcellulose-based medium (GFH4435, Stem Cell Technologies, Vancouver, BC, USA). BFU-E and CFU-GM colonies were counted 14 days after plating. BFU-E colonies were randomly picked and collected as bulk populations (containing at least 25 colonies) to evaluate the LV transduction and editing efficiencies.

Vector copy number analysis

HUDEP-2 cells and HSPC-derived mature erythroblasts were collected at day 9 and 13 of erythroid differentiation, respectively. BFU-E pools were collected 14 days after HSPC plating in the methylcellulose medium. Genomic DNA (gDNA) was extracted using PURE LINK Genomic DNA Mini kit (Life Technologies) following manufacturer's instructions. VCN per diploid genome was determined by digital droplet polymerase chain reaction (ddPCR), using the droplet reader QX200

droplet reader (Biorad), as previously described[9]. Fifty ng of gDNA was digested with 20 unit of DraI enzyme (New England BioLabs) in a total volume of 6 µL for 30 min at 37°C. We used 6 µL of the restriction mixture for the subsequent ddPCR analysis. We used primers and probes specific for: (i) the viral Ψ (PSI) packaging signal (HIV1-PSI FOR 5'-TCCCCCGCTTAATACTGACG-3', HIV1-PSI REV 5'-CAGGACTCGGCTTGCTGAAG-3', HIV1-PSI PROBE FAM 5'-CGCACGGCAAGAGGCGAGG-3'); (ii) the human albumin gene (*ALB*), as an internal reference standard (ALB FOR 5'-GCTGTCATCTCTTGTGGGCTGT-3', ALB REV 5'-ACTCATGGGAGCTGCTGGTTC-3', ALB PROBE VIC 5'-CCTGTCATGCCCACACAAATCTCTCC-3'). The VCN was determined with the QuantaSoft software by calculating the ratio between the target molecule concentration and the reference molecule concentration multiplied by the number of copies of reference species in the genome.

PCR-based assays for detection of genome editing events

Genomic DNA was extracted using PURE LINK Genomic DNA Mini kit (Life Technologies) following manufacturer's instructions. To evaluate editing efficiency at sgRNA on-target and off-target sites (predicted using COSMID; OT 1 to OT 13, **Supplementary Table 2**), we performed PCR followed by Sanger sequencing and TIDE (Tracking of InDels by Decomposition)[10,11] or Synthego Performance ICE Analysis[2]. Primer sequences used for PCR analysis are listed in **Supplementary Table 3**.

Name	Target sequence	Mismatches	Hg19 genomic location	Strand	COSMID score	Type	Gene
HBB	TCTGCCGTTACTGCCCTGT	-	chr11:5,226,973-5,226,994	-	N/A	exon 1	HBB
HBD	ACTGCTGTCAATGCCCTGT	4	chr11:5,234,385-5,234,407	-	N/A	exon 1	HBD
HBG	GGCTACTATCACAAGCCTGT	8	chr11:5,249,756-5,249,778	+	N/A	exon 1	HBG1/2
OT 1	TCTGCCATTCTGCCCTGT	2	Chr8:38,506,485-38,506,505	-	1.48	intergenic	
OT 2	TCTGCTGTTATGCCCTGT	2	Chr5:121558627-121558647	+	2.24	intergenic	
OT 3	TCTGCCCTTACTGTCCTGT	2	Chr19:1379110-1379131	+	2.17	intergenic	
OT 4	GTTGCCGTTACTGCCCTCT	3	Chr10:83,375,539-83,375,561	+	5.28	intergenic	
OT 5	TCTGCCGTTTACTGCCCTGT	1	Chr1:162859324-162859346	+	21.05	intergenic	
OT 6	TCTGCCGTTCTGCCCTGT	1	Chr8:10621755-10621775	-	21.21	intron 5	SOX7
OT 7	TCAGCAGATACTGCCCTGT	3	Chr1:160398097-160398118	-	0.75	exon 8	VANGL2
OT 8	CCTGCCTGTTACTGCCCTGT	2	Chr11:28827844-28827866	+	1.1	intergenic	
OT 9	TCTGCCTTTCTGCCCTGT	2	Chr7:101091535-101091555	+	1.48	intron 3	COL26A1
OT 10	TCTGCCCTTCCTGCCCTGC	3	Chr12:48176342-48176363	-	6.97	exon 3	SLC48A1
OT 11	TCAGCCATTACTGCCCTGT	2	Chr7:127238493-127238514	-	20.44	exon 4	FSCN3
OT 12	TCTGCCGTTACTTCCCTGC	2	Chr14:89076092-89076113	+	27.3	exon 15	ZC3H14
OT 13	TCTGCCGATAGTGCCCTGT	2	Chr17:17636353-17636374	-	1.15	intron 2	RAI1

Supplementary Table 2: List of the potential off-target loci for gR-C. Putative off-target sequences for gR-C in *HBD* and *HBG* genes and top-13 off-target loci predicted by COSMID[12]. Mismatches between the on-target *HBB* sequence and off-targets are underlined.

Name	Orientation	Primer sequence (5' to 3')
HBB	Fwd	CAGTGCAGCTCACTCAGGTGT
	Rev	ACTCCTAAGCCAGTGCCAGA
HBAS3	Fwd	CAGTGCAGCTCACTCAGCTG
	Rev	ACTCCTAAGCCAGTGCCAGA
HBD	Fwd	TGAGCCAGGCCATCACTAAAGG
	Rev	CAGGGTTTCTGAGTCAAGACACAC
HBG	Fwd	CCTCTGGGTCCATGGGTAGA
	Rev	GCAGTATCCTCTTGGGGGCC
OT 1	Fwd	CAAGCCGTAGATGGAATCTCTTGG
	Rev	CCCAGGGAGAAAGGGAGAAAG

OT 2	Fwd	CAGTTTGAGCTGCTGAGGCAC
	Rev	GAGAACTGCTTTTGCTGCATCACG
OT 3	Fwd	TGGCCACACAGTGAGACTCC
	Rev	AAGCGTGCAGGCTTCTGAGG
OT 4	Fwd	TGCAAGATAGGGACAGAAGAAGCC
	Rev	GCCAGGAACATGGTAGACATTACG
OT 5	Fwd	TCCCCTCCCTGGTTTCACCAT
	Rev	ATCTGGCTAAGACATCCTGGCTC
OT 6	Fwd	GTCACCGAATTGGGGGCAAG
	Rev	AGGAGGTCCTCAGAAGGCTTAAG
OT 7	Fwd	CCCTCCCATTTCAGCCCTTAAC
	Rev	CAACAGAATGCCCCCACAAAAGTC
OT 8	Fwd	TCTTCCTGGCCCAGAACTGTTCC
	Rev	CAGAATAGAGGTCGGGGATTGAG
OT 9	Fwd	CACCGAGGCAGGTCCTAGTT
	Rev	AGGAGTTCGAGGCTGCAGTG
OT 10	Fwd	AGCGTGTGTGTGAGTGAGGC
	Rev	GGGAAGGAGCCATCAACAGTG
OT 11	Fwd	TCTCTCCCTATGATATCCTGGCG
	Rev	TGGGGTTTCACGTGGTCAGG
OT 12	Fwd	TGAACAGTTTCAGACCACTTGGCC
	Rev	CTCCAAAACCAGGTGAGTGAGTG
OT 13	Fwd	GTCCTGCTGCCCAGCTTGTT
	Rev	CTCCTTATCTTTCCCCACGCAG
BCL11A	Fwd	TGGACAGCCCGACAGATGAA

	Rev	AAAAGCGATACAGGGCTGGC
13bpdel	Fwd	AAAAACGGCTGACAAAAGAAGTCCTGGTAT
	Rev	ATAACCTCAGACGTTCCAGAAGCGAGTGTG

Supplementary Table 3: List of primers for on-target and off-target analyses. Primers used to amplify the on-target loci of gR-A, gR-B, gR-C, gR-D, gR-BCL11A and gR-13bpdel, and the top-predicted off-target loci of gR-C. HBD primers were also used to evaluate the potential off-target activity of gR-D at the *HBD* gene. Fwd: forward and Rev: reverse.

GUIDE-seq and deep-sequencing analysis of off-target loci

GUIDE-seq was performed in HEK293 cells as previously described [13,14]. Off-targets identified by GUIDE-seq were PCR-amplified using the Phusion High-Fidelity Taq polymerase with GC Buffer (New England BioLabs). Primer sequences used to amplify off-targets identified by GUIDE-seq are listed in **Supplementary Table 4.** We performed PCR followed by deep sequencing analyses as previously described[15]. Briefly, Illumina compatible barcoded DNA amplicon libraries were prepared using the TruSeq DNA PCR-Free kit (Illumina). Libraries were pooled and sequenced using Illumina HiSeq2500 (paired-end sequencing 130 130 bases). A total of 0.59 to 1.12 million passing filter reads per sample were produced. Targeted deep-sequencing data were analyzed using CRISPRESSO [16]. The GUIDE-seq datasets are available in the BioProject repository under the accession number PRJNA734605.

Name	Orientation	Primer sequence (5' to 3')
chr 9	Fwd	CTCCCAAATTGAAAGCACAGCCAG
	Rev	TTTCCCGTTCTCCACCCAATAGC
chr 10	Fwd	TGGAGAAAGACAATGGCAGTGAGG
	Rev	CTAGCACTGCCCCACAATAGTAC

chr 6	Fwd	CCAGCATCACTACCAAGTCTCC
	Rev	AAAATCCCCCCACGGATGCC
chr 3	Fwd	CCCTAAGATTGTGGTTCCTTAGCC
	Rev	CAGGATTACTTGGGCAGAGACTAC
chr 17	Fwd	GGAATGACTGAATCGGAACAAGG
	Rev	CTGGCCTCACTGGATACTCT

Supplementary Table 4. Primers used to amplify putative off-targets of gR-C (chr 9, 10, 6 and 3) or gR-13bpdel (chr 17). Fwd: forward and Rev: reverse.

Quantitative RT-PCR (qRT-PCR)

Total RNA was extracted using RNeasy micro kit (QIAGEN) following manufacturer's instructions. Mature transcripts were reverse-transcribed using SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen) with oligo(dT) primer. qRT-PCR was performed using an iTaq Universal SYBR green master mix (Bio-Rad) and Viia7 Real-Time PCR system (Thermo Fisher Scientific). Primer sequences are listed in **Supplementary Table 5**.

Gene	Orientation	Sequence (5' to 3')
HBA	Fwd	CGGTCAACTTCAAGCTCCTAA
	Rev	ACAGAAGCCAGGAACTTGTC
HBB	Fwd	AAGGGCACCTTTGCCACA
	Rev	GCCACCACTTTCTGATAGGCAG
HBAS3	Fwd	AAGGGCACCTTTGCCCAG
	Rev	GCCACCACTTTCTGATAGGCAG
HBG1/2	Fwd	CCTGTCCTCTGCCTCTGCC

	Rev	GGATTGCCAAAACGGTCAC
GAPDH	Fwd	GAAGGTGAAGGTCGGAGT
	Rev	GAAGATGGTGATGGGATTTC
BCL11A-XL	Fwd	ATGCGAGCTGTGCAACTATG
	Rev	GTAAACGTCCTTCCCCACCT

Supplementary Table 5. List of primers used for qRT-PCR.

Flow cytometry analysis

We labeled HUDEP-2 cells and HSPC-derived erythroblasts with antibodies against CD36 (CD36-V450, BD Horizon), CD71 (CD71-FITC, BD Pharmingen), CD235a (CD235a-APC, BD Pharmingen; CD235a-PECY7, BD Pharmingen), CD49d (CD49d-APC, BD Bioscience) and Band3 (Band3-PE, Bristol Institute for Transfusion Sciences) surface markers. We used the nuclear dye DRAQ5 (eBioscience, 0.1/100 dilution) to evaluate the proportion of enucleated RBCs.

Differentiated cells were fixed and permeabilized using BD Cytofix/Cytoperm solution (BD Pharmingen) for HUDEP-2 cells and glutaraldehyde and Triton 1X for HSPCs-derived erythroblasts. After permeabilization, cells were stained with antibodies recognizing HbF (HbF-APC, MHF05, Life Technologies and HbF-FITC, 552829, BD Pharmingen, 1/100 dilution).

We performed flow cytometry analyses using Fortessa X20 flow cytometer (BD Biosciences) and Gallios (Beckman Coulter).

Western blot

Differentiated HUDEP-2 cells (9 days of differentiation; 2-3x10^6 cells) were collected and resuspended for 30 min at 4°C in a lysis buffer containing: 10mM Tris, 1 mM EDTA, 0.5mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxicholate, 140mM NaCl (Sigma-Aldrich) and protease inhibitor cocktail (Roche-Diagnostics). Cell lysates were sonicated twice (50% amplitude, 10 sec per cycle, pulse 9 sec on/1 sec off) and underwent 3 cycles of freezing/thawing (3 min at -80°C/3 min at

37°C). Lysates were centrifuged at 12.000 x g for 10 min at 4°C, and supernatants were used as protein extracts for biochemical analysis. Protein concentration was measured using the Pierce™ BCA Protein Assay Kit (ThermoScientific). 25 µg of samples were loaded on a 15% (wt/vol) Sodium Dodecyl Sulphate (SDS)-PAGE gel and transferred onto PVDF membranes (Millipore) pre-activated in ethanol, using a Mini Blot Transfer Module (Thermo). After transfer, membranes were incubated for 1 hour in Tris-buffered saline (TBS)-Tween buffer (50 mM Tris-HCl pH8, 150 mM NaCl, 0.1% Tween 20 in water) completed with 5% skim milk. Immunoblots were carried out over-night at 4°C in a solution containing 3% (wt/vol) BSA in TBS with a mouse antibody recognizing β-globin (1:200; sc-21757, SantaCruz) or a goat antibody against α-globin (1:200; sc-31110, SantaCruz). Blots were washed and incubated with peroxidase-conjugated donkey anti-mouse (1:10,000) or donkey anti-goat (1:5,000) antibody for 60 min at RT. After washing, blots were incubated with enhanced chemiluminescent (ECL) substrate (Pierce) and exposed to film (Amersham) that was developed following manufacturer's instructions. Images were quantified by using Gel-Pro analyzer 4.0 (Media Cybernetics).

Keywords

Lentiviral vectors, genome editing, CRISPR/Cas9 nuclease, sickle cell disease.

References

1. Romano O, Petiti L, Felix T, et al. GATA Factor-Mediated Gene Regulation in Human Erythropoiesis. *iScience*. 2020;23(4):.

296 2. Hsiau T, Conant D, Rossi N, et al. Inference of CRISPR Edits from Sanger Trace Data.
297 *BioRxiv*. 2018;
298 3. Dang Y, Jia G, Choi J, et al. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout
299 efficiency. *Genome Biology*. 2015;15;16:280.
300 4. Cottle RN, Lee CM, Archer D, Bao G. Controlled delivery of β-globin-targeting TALENs and
301 CRISPR/Cas9 into mammalian cells for genome editing using microinjection. *Sci Rep*. 2015;5:.
302 5. Liang P, Xu Y, Zhang X, et al. CRISPR/Cas9-mediated gene editing in human tripronuclear
303 zygotes. *Protein Cell*. 2015;6(5):363–372.
304 6. Huang X, Wang Y, Yan W, et al. Production of Gene-Corrected Adult Beta Globin Protein in
305 Human Erythrocytes Differentiated from Patient iPSCs After Genome Editing of the Sickle Point
306 Mutation: Corrected β Globin in Genome-Edited Human iPSCs. *Ste*. 2015;33(5):1470–1479.
307 7. Kurita R, Suda N, Sudo K, et al. Establishment of Immortalized Human Erythroid Progenitor
308 Cell Lines Able to Produce Enucleated Red Blood Cells. *PLoS One*. 2013;8(3):.
309 8. Antoniani C, Meneghini V, Lattanzi A, et al. Induction of fetal hemoglobin synthesis by
310 CRISPR/Cas9-mediated editing of the human β-globin locus. *Blood*. 2018;131(17):1960–1973.
311 9. Lin H-T, Okumura T, Yatsuda Y, et al. Application of Droplet Digital PCR for Estimating
312 Vector Copy Number States in Stem Cell Gene Therapy. *Hum Gene Ther Methods*. 2016;27(5):197–
313 208.
314 10. Brinkman EK, van Steensel B. Rapid Quantitative Evaluation of CRISPR Genome Editing by
315 TIDE and TIDER. *Methods Mol. Biol.* 2019;1961:29–44.
316 11. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome
317 editing by sequence trace decomposition. *Nucleic Acids Res*. 2014;42(22):e168.
318 12. Cradick TJ, Qiu P, Lee CM, Fine EJ, Bao G. COSMID: A Web-based Tool for Identifying
319 and Validating CRISPR/Cas Off-target Sites. *Mol Ther Nucleic Acids*. 2014;3(12):e214.
320 13. Tsai SQ, Zheng Z, Nguyen NT, et al. GUIDE-seq enables genome-wide profiling of off-target
321 cleavage by CRISPR-Cas nucleases. *Nature Biotechnology*. 2015;33(2):187–197.

14. Weber L, Frati G, Felix T, et al. Editing a γ-globin repressor binding site restores fetal hemoglobin synthesis and corrects the sickle cell disease phenotype. *Sci. Adv.* 2020;6(7):eaay9392.

15. Lattanzi A, Meneghini V, Pavani G, et al. Optimization of CRISPR/Cas9 Delivery to Human Hematopoietic Stem and Progenitor Cells for Therapeutic Genomic Rearrangements. *Molecular Therapy*. 2018;

16. Pinello L, Canver MC, Hoban MD, et al. Analyzing CRISPR genome-editing experiments with CRISPResso. *Nature Biotechnology*. 2016;34(7):695–697.