Supplementary information

Supplemental Table 1 : antibodies for flow cytometry

Name	Fluorochrome	Clone	Company	Catalog Number
human Fc Receptor binding inhibitor polyclonal antibody			eBioscience	16-9161-73
anti-human Fetal Hemoglobin	APC		Life Technologies	MHF05
anti-human Hemoglobin β	PE	37-8	Santa Cruz Biotechnology	sc-21757
anti-human HLA-ABC	PE	G46-2.6	BD Bioscience	555553
anti-human CD36	FITC	C836	BD Bioscience	555454
anti-human CD235a (GYPA)	PE-Cy7	GA-R2	BD Bioscience	563666
anti-human CD71	APC	M-A712	BD Bioscience	551374
anti-human CD45	APC	HI30	BD Bioscience	555485
anti-human CD33	BV421	P67.6	BD Bioscience	744761
anti-human CD19	BV605	HIB19	BD Bioscience	740394
anti-human CD3	BB700	SK7	BD Bioscience	566575
anti-human CD271 (NGFR)	APC	ME20.4-1.H4	Miltenyi Biotec	130-113-418
anti-human CD34	PE	AC136	Miltenyi Biotec	130-113-179
IgG1, κ isotype control	APC		Life Technologies	MG105
IgG1, κ isotype control	APC		BD Bioscience	550854
IgG1, κ isotype control	FITC		BD Bioscience	555748
IgG1, κ isotype control	BV421		BD Bioscience	562438
IgG1, κ isotype control	BV605	X40	BD Bioscience	562652
IgG1, κ isotype control	BB700	X40	BD Bioscience	566404
IgG2a, κ isotype control	APC	G155-178	BD Bioscience	551414
IgG2a, κ isotype control	PE		BD Bioscience	555574
IgG2a, κ isotype control	PE	S43.10	Miltenyi Biotec	130-113-272
IgG2b, κ isotype control	PE-Cy7		BD Bioscience	560542

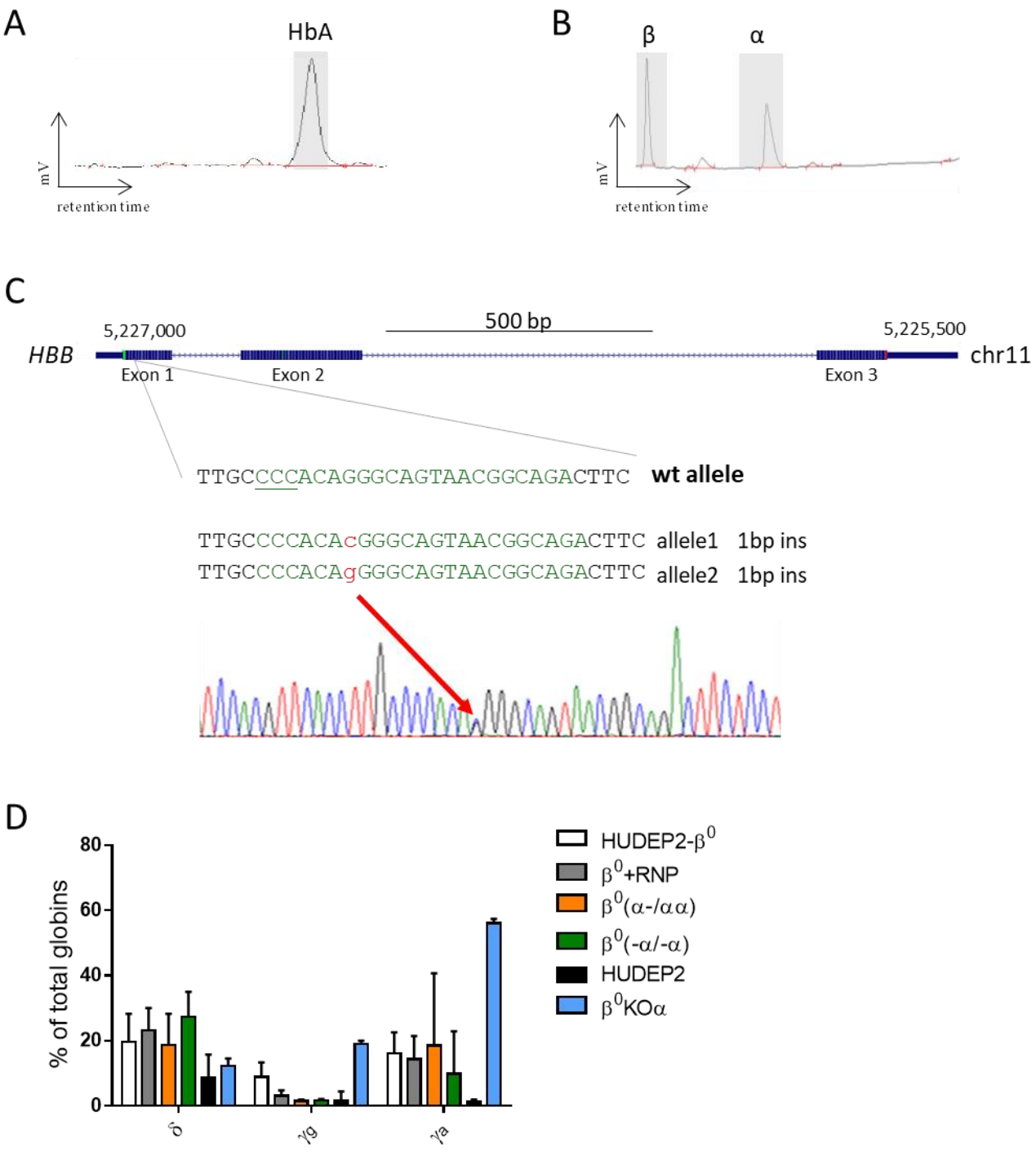
All antibodies were mouse monoclonal antibodies, unless differently specified.

Supplemental Table 2: primer list

PCR		
Gene/amplicon	**name**	**sequence**
HBA 1/2	3820 F	TATCGCCAGAGGGAAAGGGA
	4870 R	CTTGAAGTTGACCGGGTCCA
HBA1/2 sequencing	894 R	TAGGTCTTGGTGGTGGGGAA
HBA2	611 F	GCACTCTTCTGGTCCCCAC
	1512R	GCAGAGAGGTCCTTGGTCTG
HBA1	HBA EX1F	CCGACAAGACCAACGTCAA
	5614 R	CTCTAGGGTCCAGCGTTTTTCC
AAVS1	MA359	CAGCTCAGGTTCTGGGAGAG
	MA360	CTTGTAGGCCTGCATCATCA
HBA1/del	3820 F	TATCGCCAGAGGGAAAGGGA
	5614 R	CTCTAGGGTCCAGCGTTTTTCC
HBB	F7895	AGGCCATCACTAAAGGCACC
	R8327	AGTCAGGGCAGAGCCATCTA
HBB promoter	HBB9 F	CTGTCTCCACATGCCCAGTT
	HBB9 R	GGAGACGCAGGAAGAGATCC
Off Target 1	OT1-F	CCTCTGCTGGCAGGATTT
	OT1-R	AGCAAGGGTAAGAATCTGTCTC
Off-Target 2	OT2-F	TCTCACAATTCAAGAGGCCAGA
	OT2-R	TCAGACACACACAGGGAGAG
Off-Target 3	OT3-F	GTTCTCCAGCAGCCTTTCA
	OT3-R	AGCACTTCCAAAGGGTTGT
ddPCR		
On-target 3'junction	3'HBA INT F	TGGACAAACCACAACTAGAATGC
	3'HBA INT 3 R	AAGTGCGGGAAGTAGGTCTT
	3'HBA INT PRB	56-FAM/CTGTCTCCTGCCGACAAGACCAAC
HBA2	HBA2 3' F	GCCCTTCCTGGTCTTTGAATA
	HBA2 3' R	ACCTCCATTGTTGGCACAT
	HBA2 3' PRB	56-FAM/TGTGTGTGCCTGGGTTCTCTCTAT
ALB	ALB F	GCTGTCATCTCTTGTGGGCTGT
	ALB R	ACTCATGGGAGCTGCTGGTTC
	ALB P	CCTGTCATGCCCACACAAATCTCTCC
qPCR		
GAPDH	GAPDH F	CTTCATTGACCTCAACTACATGGTTT
	GAPDH R	TGGGATTTCCATTGATGACAAG
HBA1/HBA2	HBA F	CGGTCAACTTCAAGCTCCTAA
	HBA R	ACAGAAGCCAGGAACTTGTC
HBB	HBB F	GCAAGGTGAACGTGGATGAAGT
	HBB R	TAACAGCATCAGGAGTGGACAGA
HBB AS3	HBBAS3 F	AAGGGCACCTTTGCCCAG
	HBBAS3 R	GCCACCACTTTCTGATAGGCAG
HBG1/HBG2	HBG1 2 F	CCTGTCCTCTGCCTCTGCC
	HBG1 2 R	GGATTGCCAAAACGGTCAC
HBD	HBD F	CAAGGGCACTTTTTCTCAG
	HBD R	AATTCCTTGCCAAAGTTGC
GYPA	hGYPA F	GGTGGCAATGCACACTTCAA
	hGYPA R	ACCCTTTCTCCGGTTTCCTC

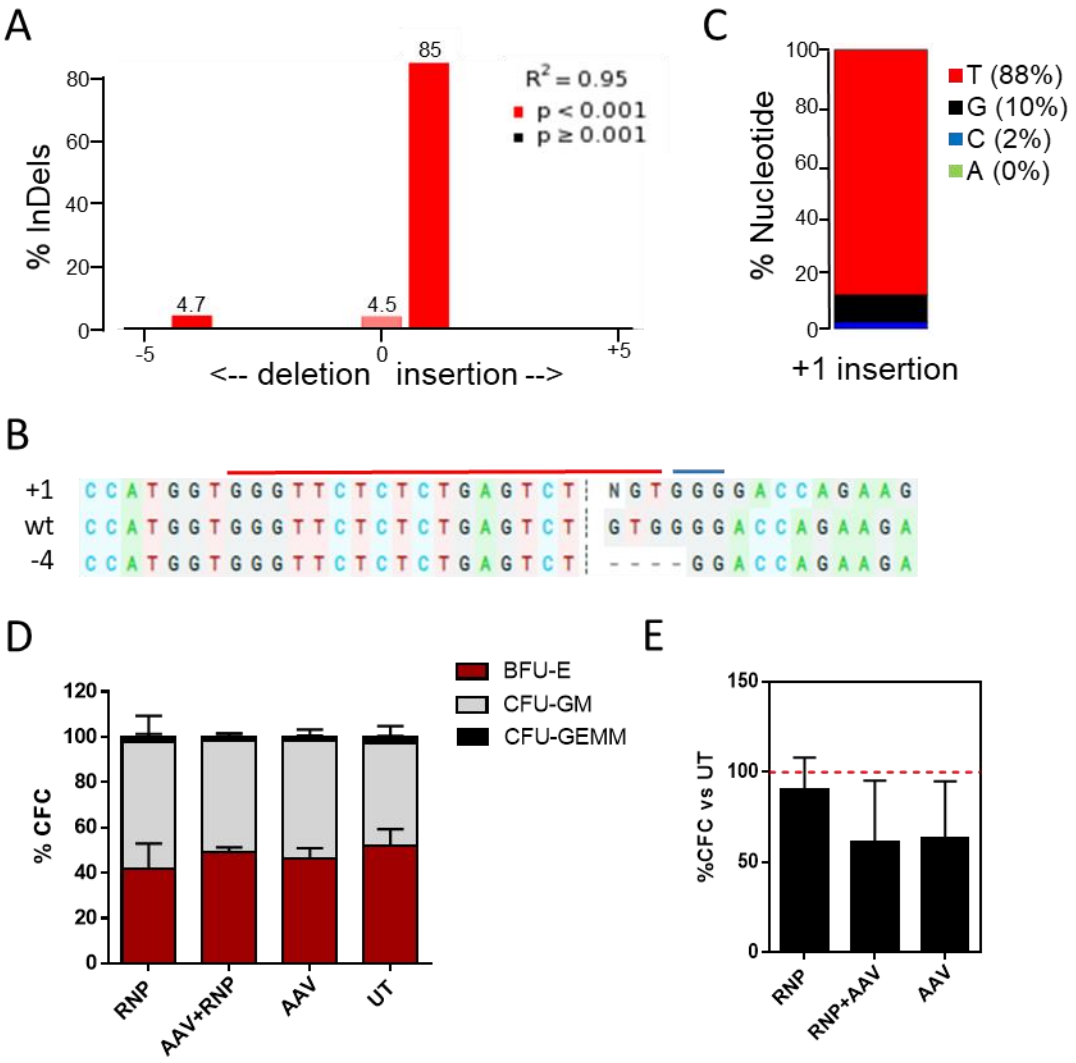
F=forward; R=reverse, PRB=probe.

Supplementary data



Supplementary figure 1. Genotype of HUDEP-2 β^0

(**A-B**) Representative HPLC chromatograms of globin tetramer (**A**) and b-like subunits (**B**) analysis. HbA, adult hemoglobin ($\alpha_2\beta_2$). (**C**) Schematic representation of the β-globin gene (*HBB*; hg38). Reference allele (wt) is shown above HUDEP-2 β^0 Sanger sequencing profiles. Frameshift disruption occur at leucine 15. gRNA target site is shown in green, PAM sequence is underlined and +1 insertions are shown in red. (**D**) RT-qPCR quantification of β-like globin mRNA ratios in HUDEP-2 β^0 (n=3-5) and in single cell clones with monoallelic (α-/αα; n=3) or biallelic (α-/α-; n=3) *HBA2* deletions. β^0 + RNP = HUDEP2 b^0 bulk population edited with gRNA HBA15/Cas9. β^0 KOa= HUDEP2 β^0 bulk population edited with gRNA HBB/Cas9 to induce a-globin KO. Bars represent mean±SD (n=3; ** $p<0.01$, ANOVA, Tukey's test).

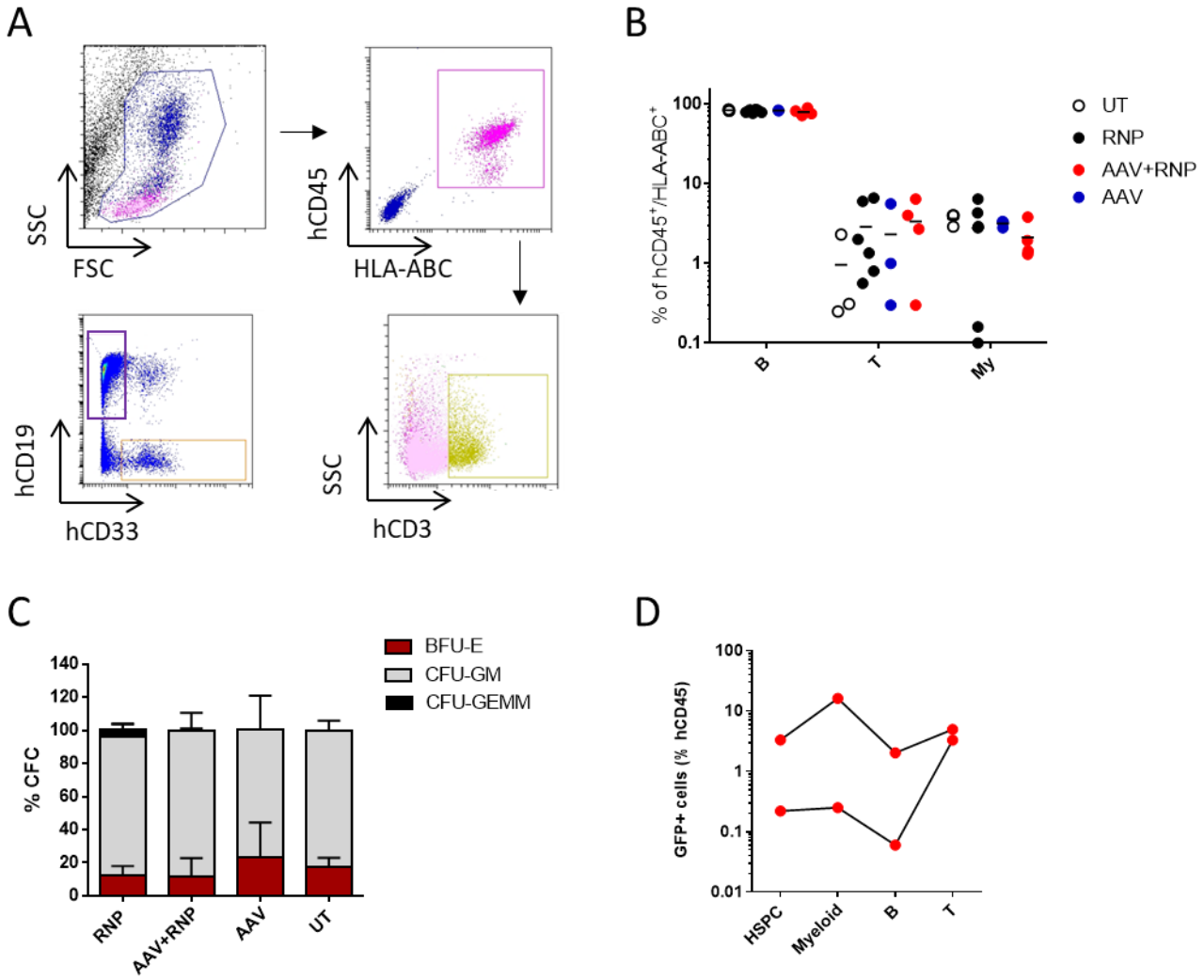


F

	wt	1 β^{AS3}	2 β^{AS3}	**Tot**
wt	20	7	1	28
1 del	6	17	2	25
2 del	0	5	0	5
Tot	26	29	3	

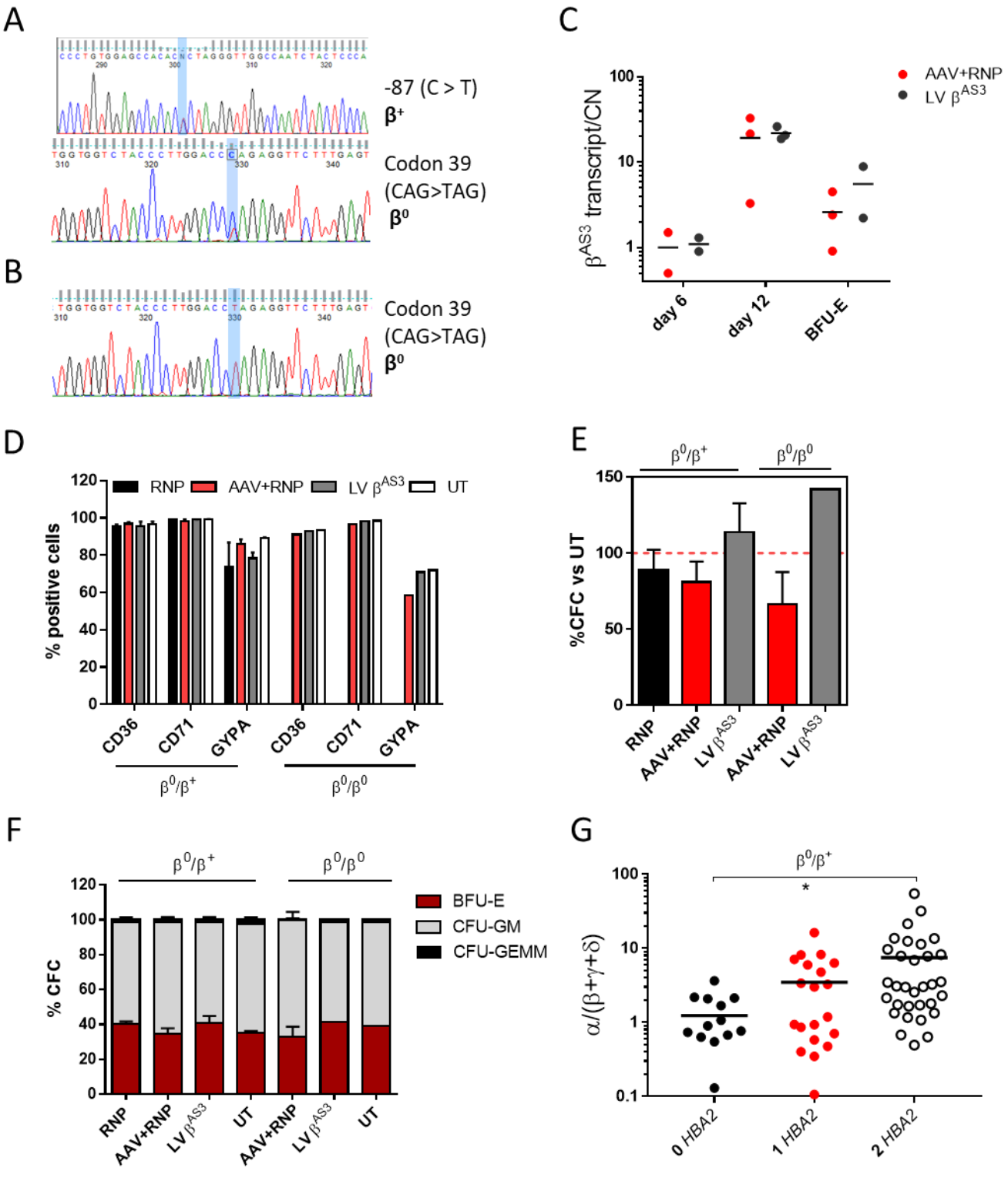
Supplemental Figure 2. *Ex vivo* characterization of edited HSPCs.

(**A-C**) Representative InDel spectrum (A), sequence (B) and 1+ nucleotide insertion probability (C) for gRNA HBA15. In (B) Red line: gRNA sequence; blue line: PAM sequence. (**D**) Colony formation unit (CFU) frequency in edited HSPCs. CFU-GEMM, granulocyte, erythroid, macrophage, megakaryocyte; BFU-E, burst-forming unit-erythroid; CFU-GM, granulocyte, macrophage. Bars represent mean ±SD (n=3-5). (**E**) CFC number expressed as percentage of untreated control (UT). Bars represent mean ±SD (n=3-5), red dashed line indicates 100%. (**F**) Colony genotypes shown in Figure 3I.



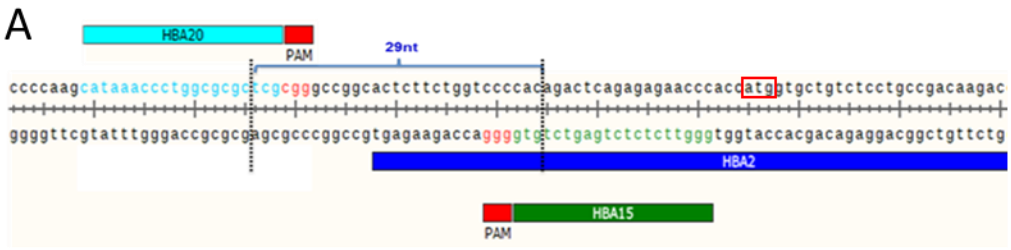
Supplemental Figure 3. *In vivo* characterization of edited HSPCs.

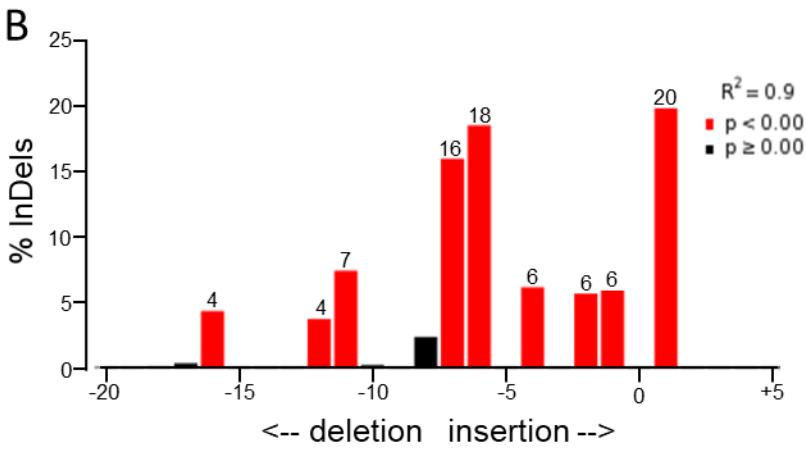
(**A**) Gating strategy to analyze engraftment levels and cell lineages of human xenografts in NSG mice. Representative plots for quantification of human hematopoietic (CD45+/HLA-ABC+), human B (CD19+), human myeloid (CD33+) and human T (CD3+) cells. (**B**) Immunophenotyping of human xenografts in bone marrow of NSG mice. Each dot represents one animal, lines indicate mean. (**C**) Colony formation unit (CFU) frequency in bone marrow-derived CD34. CFU-GEMM, granulocyte, erythroid, macrophage, megakaryocyte; BFU-E, burst-forming unit-erythroid; CFU-GM, granulocyte, macrophage. Bars represent mean ±SD (n=2-4). (**D**) GFP positive cells in HSPCs (CD34), myeloid (CD33), B (CD19) and T (CD3) cells in BM of transplanted mice. Each line represents one animal.

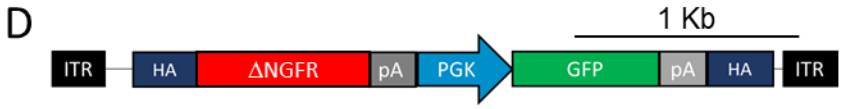

A
-87 (C > T)
β+
Codon 39
(CAG>TAG)
β0
B
Codon 39
(CAG>TAG)
β0
C
AAV+RNP
LV βAS3
βAS3 transcript/CN
day 6
day 12
BFU-E
D
RNP
AAV+RNP
LV βAS3
UT
% positive cells
CD36
CD71
GYPA
β0/β+
β0/β0
E
%CFC vs UT
RNP
AAV+RNP
LV βAS3
F
BFU-E
CFU-GM
CFU-GEMM
% CFC
G
α/(β+γ+δ)
0 HBA2
1 HBA2
2 HBA2

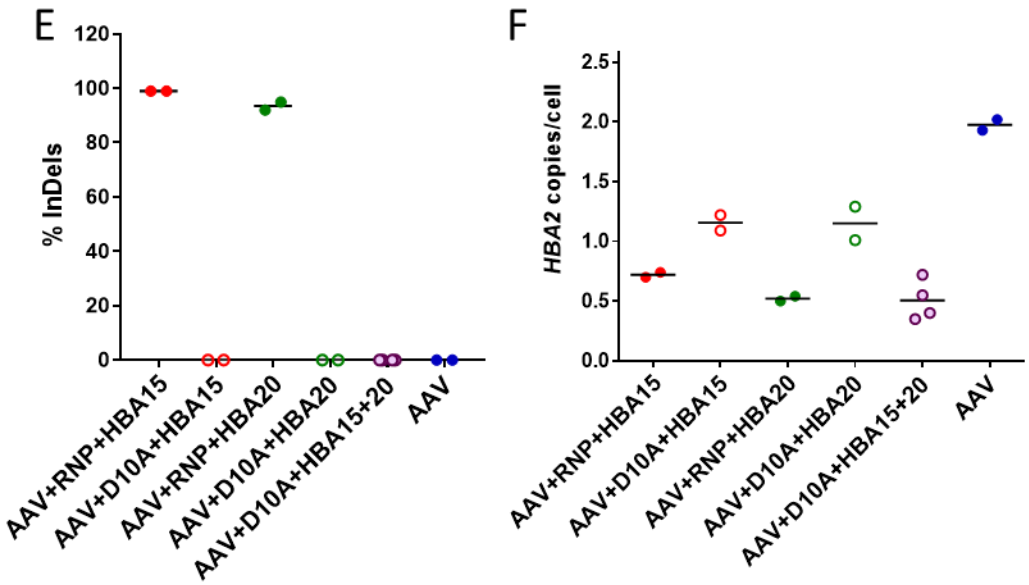
Supplemental figure 4. Characterization of edited thalassemic HPSCs.

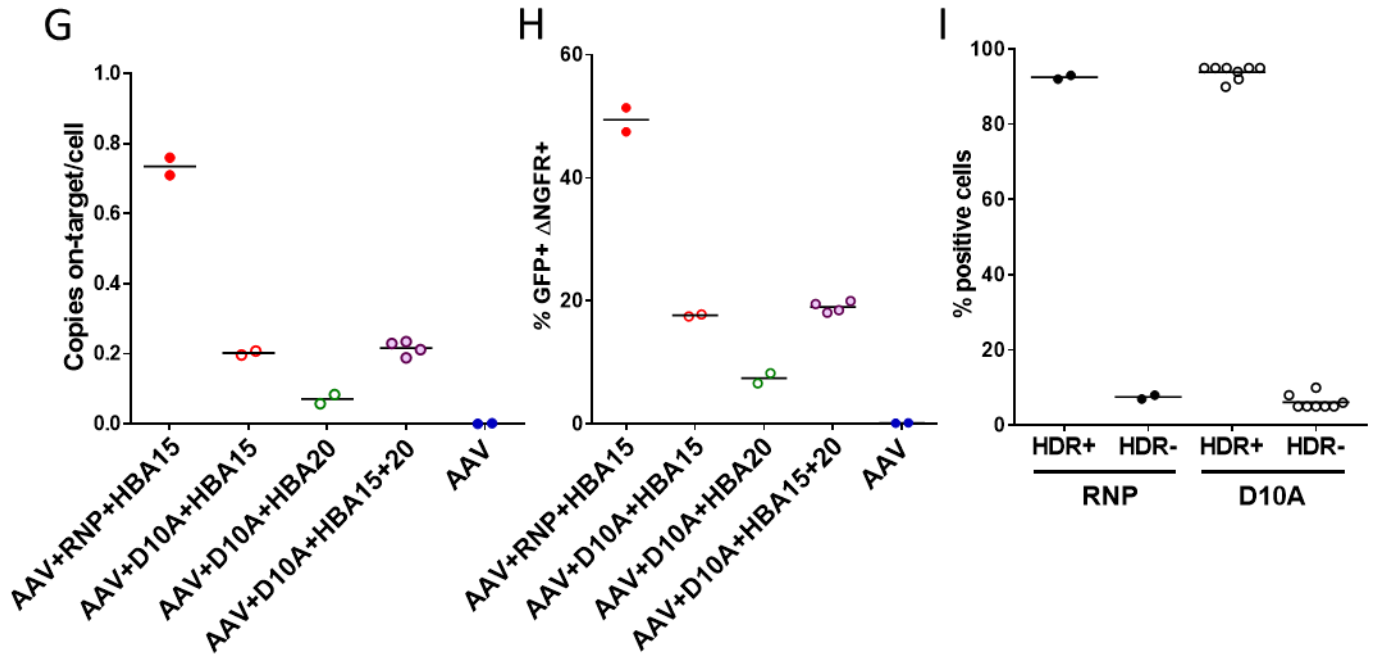
(**A**-**B**) Sanger sequencing of *HBB* gene in patients' HSPCs. Patient 1 carries the -87 C>T β^+ mutation in heterozygosis with the codon 39 (CAG>TAG) β^0 mutation (A). Patient 2 is homozygous for the codon 39 (CAG>TAG) β^0 mutation (B). Blue boxes highlight the mutated nucleotides. (**C**) β^{AS3} RNA expression during erythroid differentiation in β^+ erythroblasts. Values are normalized per copy number (CN). Black lines indicate mean. (**D**) Flow-cytometry analysis at day 12 of erythroid differentiation of edited HSPCs. Results are shown as mean ±SD (n=1-3). (**E**) CFC numbers expressed as percentage of untreated control (UT). Red dotted line indicates 100%; bars represent mean ±SD (n=1-3). (**F**) Colony-forming cell (CFC) frequency in edited HSPCs (mean ±SD). CFU-GEMM, granulocyte, erythroid, macrophage, megakaryocyte; BFU-E, burst-forming unit-erythroid; CFU-GM, granulocyte, macrophage. (**G**) α/β-like globin mRNA ratios in edited thalassemic BFU-E (n=67) are plotted. 2 HBA2 colonies come from untreated and wt edited BFUE. Each dot represents a single colony. Black lines indicate mean (n= ;*, $p<0.01$; 1 way ANOVA, Dunn's test).


A
HBA20
PAM
29nt
HBA2
HBA15
PAM


B
% InDels
$R^2 = 0.9$
p < 0.001
p ≥ 0.001
<-- deletion insertion -->
C
% Nucleotide
T (30.3%)
G (49.5%)
C (7.6%)
A (12.5%)
+1 insertion


D
1 Kb
ITR
HA
ΔNGFR
pA
PGK
GFP
pA
HA
ITR


E
% InDels
AAV+RNP+HBA15
AAV+D10A+HBA15
AAV+RNP+HBA20
AAV+D10A+HBA20
AAV+D10A+HBA15+20
AAV
F
HBA2 copies/cell
AAV+RNP+HBA15
AAV+D10A+HBA15
AAV+RNP+HBA20
AAV+D10A+HBA20
AAV+D10A+HBA15+20
AAV



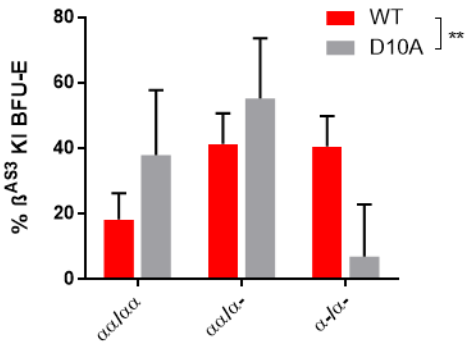
Supplemental figure 5. Single vs dual nick editing of K562

(**A**) Genomic location of HBA15 and HBA20 gRNAs. The distance between the two cutting sites is indicated. Red rectangle: HBA start codon. (**B**-**C**) Representative InDel spectrum (A) and 1+ nucleotide insertion probability (B) for HBA20. (**D**) AAV6 donor used for KI experiments. Vector contains a promoterless transgene encoding for the truncated inactive form of the human low affinity nerve growth factor receptor (ΔNGFR), followed by a phosphoglycerate kinase (PGK) promoter with a GFP reporter and simian virus polyA (pA). This cassette is flanked by 250 bp homology arms (homology) to gRNA genomic target. ITR, inverted terminal repeats. 1KB scale bar is indicated on top. (**E**-**F**) InDels (E) and *HBA2* copies (F) quantification in edited K562 cells. Black lines represent mean. (**G**-**H**) AAV knock-in efficiency in term of on-target DNA integration (G) and NGFR/GFP expression (H). Black lines represent mean. (I) Percentage of cells with HDR or not HDR mediated DNA targeted integration (same data of H). HDR+ = (GFP+ DNGFR+ / GFP+) * 100; HDR- = 100 - HDR+. Black lines represent mean.

A

	wt	1 β^{AS3}	2 β^{AS3}	**Tot**
wt	105	11	0	116
1 del	58	16	0	74
2 del	10	2	0	12
Tot	173	29	0	

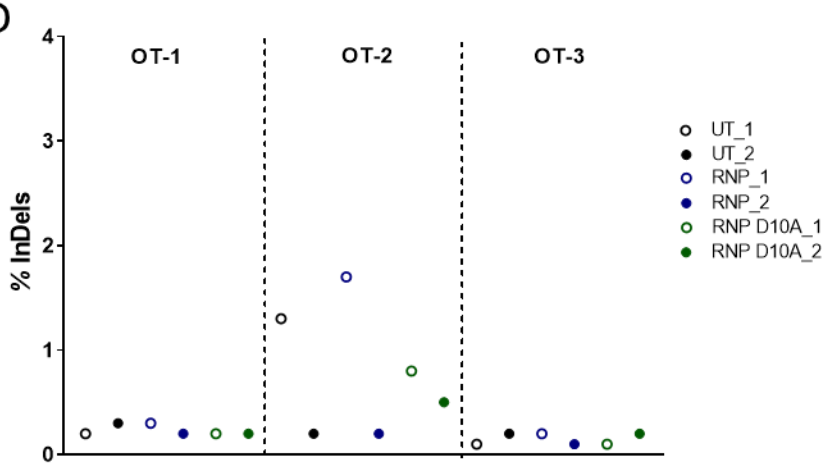
B



C

		Reads	Position (hg19)	Strand	Type	Gene
ON TARGET	G G G T T C T C T C T G A G T C T G T G N G G	329	Chr16_222886/226690	-	Exon	HBA2
OT-1	· A · · · · C · · · G · · · · · · · · · · ·	320	Chr3_141179078	-	Intergenic	
OT-2	· · · · · · C · · · · · · · G · · · · · · ·	135	Chr9_77688205	-	Intron	NMRK1
OT-3	· · · C · · · C · · · · · A · · · · A · · ·	75	Chr8_143095162	+	Intergenic	
OT-4	· · · · · · · C · C A T · · · · · · · · · ·	29	Chr4_66205583	-	Intron	EPHA5
OT-5	A · A · · · · C · · · · · A · · · · · · · ·	28	Chr5_2859208	-	Intergenic	
OT-6	· · · C · · · · · · · · · G · · · · A · · ·	20	Chr20_38984675	-	Intergenic	
OT-7	A · · G · · · · · · · · · · · C T · · · · ·	16	Chr10_43101518	-	Intron	ZNF33B
OT-8	· · · · A · · · C · A · · · · · A · · · · ·	13	Chr12_117293884	+	Intergenic	
OT-9	A · · G · · · · A · · · · · · · · G · · · ·	11	Chr17_26110678	+	Intron	NOS2
OT-10	· · · G · · · G · · · · T · · · · · · · · ·	10	Chr22_22683405	+	Intergenic	
OT-11	A A · · · · · G · · · · · · A · · · · · · ·	8	Chr2_217147479	+	Intron	MARCH4
OT-12	· · · · · · C · · · · A · · · G · · · · · ·	5	Chr11_70355141	-	Intron	SHANK2
OT-13	· · · · · · · · · · · C · · · · · · · · · T	4	Chr6_24657192	-	Intron	TDP2
OT-14	C · · · · · · · · · · · · · · · · A · · · A ·	4	Chr6_169440547	-	Intergenic	
OT-15	A · · C · · C · · · · · · · · · · · · · · A	3	Chr2_227599376	-	Exon	IRS1

D



Supplemental figure 6. CFC and off-target analysis

(**A**) Colony genotypes shown in Figure 6H. (**B**) Genotype of CFC with β^{AS3} transgene integration using Cas9 (WT, gray; n=121) or Cas9 nickase (D10A, red; n=29). Bar represent frequency ±confidence intervals (**, $p < 0.01$; Chi-square test). (**C**) GUIDE-Seq analysis. The on-target sequence and PAM motif are shown in the top line. Off-targets and their mismatches to the on-target site (highlighted in color), sequencing read counts, chromosomal position (hg19), strand and type of genomic region are reported. The majority of off-targets map to intergenic regions or introns except OT-15, which lays in exon 2 of the IRS1 gene that is not expressed in the hematopoietic system. (**D**) InDel quantification by deep-sequencing of top 3 off-target sites identified by GUIDE-Seq in erythroblasts derived from control and edited HSPCs (n=2 different healthy donors). On-target InDels were 92 and 98% (TIDE) for Cas9 wt. The InDels measured in OT-2 come from amplification/sequencing errors, as mutations in the amplicons are not specific to the gRNA matching sequence.