OMTN, Volume *35*

Supplemental information

Context base editing for splice correction of IVSI-110 β-thalassemia

Basma Naiisseh, Panayiota L. Papasavva, Nikoletta Y. Papaioannou, Marios Tomazou, Lola Koniali, Xenia Felekis, Constantina G. Constantinou, Maria Sitarou, Soteroula Christou, Marina Kleanthous, Carsten W. Lederer, and Petros Patsali

Supplemental Tables

Table S1. Primer identities for DNA and RNA expression analyses

Primer Name	Application	Sequence
wtHBB_Probe_ZNA	RT-qPCR/ddPCR	6-FAM-TGGG(PDC)AGG(PDC)TG(PDC)TG-ZNA-3-BHQ-1 VIC-TAAGGGTGGGAAAATAGA-MGB
IVSI-110_MGB_Probe	RT-qPCR/ddPCR	VIC-TAAGGGTGGGAAAATAGA-MGB
hHBB_FW	RT-qPCR	GCCCTGGCCCACAAGTATC
hHBB_RV	RT-qPCR	GCCCTTCATAATATCCCCCAGTT
hHBA_FW	RT-qPCR	GGTCAACTTCAAGCTCCTAAGC
hHBA_RV	RT-qPCR	GCTCACAGAAGCCAGGAACTTG
hHBB_EX1_FW_3	RT-qPCR/ddPCR	GGGCAAGGTGAACGTG
hHBB_EX2_RV_1	RT-qPCR/ddPCR	GGACAGATCCCCAAAGGAC
HBB_CD34_FW	PCR/Sequencing	TGAGGAGAAGTCTGCCGTTAC
HBB_CD34_RV	PCR/Sequencing	CAGCTCACTCAGTGTGGC
SpG7_OFF1_RP11-275O18.1-RP11-362A1.1_FW	PCR/Sequencing	AATACTCTCCTTTCATGTGCCATG
SpG7_OFF1_RP11-275O18.1-RP11-362A1.1_RV	PCR/Sequencing	ATGGCTTCTCTCTTGTGAGGAG
SpG7_OFF2/SpRY6(NNN)_OFF1_ YWHAEP4/AC097467.2-RP11-27G13.4_FW	PCR/Sequencing	CCCAGGAGAGCATGGAAAGG
SpG7_OFF2/SpRY6(NNN)_OFF1_ YWHAEP4/AC097467.2-RP11-27G13.4_RV1	PCR/Sequencing	AGGAACTCTCAAGAACTGCAAGC
SpG7_OFF2/SpRY6(NNN)_OFF1_ YWHAEP4/AC097467.2-RP11-27G13.4_RV2	PCR/Sequencing	GCAATAGGCGAGACATCACCG
SpRY6(NNN)_OFF2_ RP11-781M16.2_FW	PCR/Sequencing	ACCATACCAGTACCTTGGC
SpRY6(NNN)_OFF2_ RP11-781M16.2_RV	PCR/Sequencing	CTGACTGGCTGGAAGGAG
SpRY6(NYN)_OFF1_ KIAA1328_FW	PCR/Sequencing	AGACACAAAGGTTGAAAGTTAAAAGG
SpRY6(NYN)_OFF1_ KIAA1328_RV	PCR/Sequencing	TCATATATTTTGAGGTGCTGTTTGG
SpRY6(NYN)_OFF2_ RPS23P3-RNU6-699P_FW	PCR/Sequencing	ACAAGTGCCCTCACTGTCA
SpRY6(NYN)_OFF2_ RPS23P3-RNU6-699P_RV	PCR/Sequencing	ACCTGAGCATCCCCTGAGTC
SpRY20(NNN)_OFF1_ NCOA7-AS1-NCOA7_FW1	PCR/Sequencing	ACCACATTTCATTATTTTGTTGAC
SpRY20(NNN)/(NRN)_OFF1_ NCOA7-AS1-NCOA7_FW2	PCR/Sequencing	AATTATGCTCATAGTGAGGG
SpRY20(NNN)/(NRN)_OFF1_ NCOA7-AS1-NCOA7_RV	PCR/Sequencing	CAATCTAGGGAGGGCAG
SpRY20(NNN)_OFF2_ SNRPGP2-RP11-61D1.2_FW	PCR/Sequencing	ATGGGTGAGTTTATGACTCTTGAGAG
SpRY20(NNN)_OFF2_ SNRPGP2-RP11-61D1.2_RV	PCR/Sequencing	TGTTTCTGGGGCAGCAGTC
SpRY20(NRN)_OFF2_ AC009541.1-hsa-mir-490_FW	PCR/Sequencing	GCTCAGCCATTCTACTTCTAGGCA
SpRY20(NRN)_OFF2_ AC009541.1-hsa-mir-490_RV	PCR/Sequencing	AGCGGTATTTTCTCAGGATGGTGG

Table S2. Top 10 off-target predictions for SpG7 (ACTAATAGGCAGAGAGAGTCAGT) as ranked by CRISPOR.[1] PAM = NGN.

Off-target Sequence	Mismatch Positions	Mismatch Count	MIT Off-target	CFD Off-target	Chromo-some	Strand	Locus Description
ATTAATAGGAAAAGAGAGTCAGG	.*.....*.*.....	3	1.275483	0.735354	Chr12	+	intergenic:RP11-275O18.1-RP11-362A1.1
ACTAAAAGGCAGAGAGAGACAGG	.....*.....*.	2	2.105189	0.247619	Chr4	-	intergenic:YWHAEP4/AC097467.2-RP11-27G13.4
GCTGATAGGCAGAGAGAGACAGG	*..*.....*.	3	1.127119	0.178571	Chr16	-	intergenic:RP11-53L24.1-RPSAP56
GCAAATAAGAAGAGAGAGTCAGA	*.*....*.*.....	4	1.299248	0.042989	Chr13	-	intergenic:LINC00333-LINC00375
ATTAATTGGAAGAGAGAGTCAGA	.*....*..*.....	3	1.680987	0.023937	Chr14	+	intron:EML5
AGTATTAGGCAGAGAGAGTCAGA	.*..*.....	2	5.722892	0.017567	Chr4	+	exon:RP11-781M16.2
ATTTATAGGCTGAGAGAGTCTGA	.*.*.....*.....	3	1.483122	0.012361	Chr9	-	intergenic:RNA5SP279-RP11-443B9.1
TCTCCTAGGAAGAGAGAGTCTGA	*..**....*.....	4	1.317696	0.010621	Chr10	-	intergenic:LINC00841-AL512640.1
AGAAATAGGCAGAGAGAGACAGA	.**.....*.	3	1.040778	0.009859	Chr16	-	intergenic:AC136932.2-LINC00273/AC136932.1
AGAGATAGGGAGAGAGAGTCAGA	.***.....*.....	4	1.239513	0.008387	Chr19	+	intergenic:TUBB4A-TNFSF9

Yellow highlight identifies genes sequenced for off-target activity.

Table S3.1. Top 10 off-target predictions for SpRY6 (CTAATAGGCAGAGAGAGTCA) as ranked by CRISPOR.[1] PAM = NNN

Off-target Sequence	Mismatch Positions	Mismatch Count	MIT Off-target	CFD Off-target	Chromosome	Strand	Locus Description
CTAAAAGGCAGAGAGAGACAGGG	....*.....*..	2	2.165116	0.333333	Chr4	-	intergenic:YWHAEP4/AC097467.2-RP11-27G13.4
GTATTAGGCAGAGAGAGTCAGAG	*..*.....	2	5.722892	0.150637	Chr4	+	exon:RP11-781M16.2
CTCCTAGGCTGAGAGAGTCAAGG	..**.....*.....	3	2.309774	0.133465	Chr6	-	intergenic:RP3-404K8.2-HDGFL1
GTCACAGGCAGAGAGAGTCACAG	*.*.*.....	3	2.392593	0.101449	Chr11	-	exon:CWF19L2
ACCATAGGCAGAGAGAGTCATAG	***.....	3	2.287424	0.095238	Chr15	+	intron:GRAMD2
GTAATAGGCAGAAAGAGTCAAGA	*.....*.....	2	3.91117	0.058528	Chr8	+	intergenic:RP11-192P9.1-TRPS1
CTAGTAGACCGAGAGAGTCACAG	...*...*.*.....	3	2.34257	0.054012	Chr5	-	intergenic:IPO11/KIF2A-IPO11
CTACAAGGCTGAGAGAGTCAAAG	...**....*.....	3	2.34257	0.040369	Chr5	+	intergenic:CTC-505O3.3-CTC-505O3.2
CCAATAGGCAGAGAGAGTCAGTG	.*.....	1	100	0.033395	Chr11	+	exon:HBB
CTGATAGGCAGAGAGAGACAGGA	..*.....*..	2	2.62276	0.028292	Chr16	-	intergenic:RP11-53L24.1-RPSAP56

Yellow highlight identifies genes sequenced for off-target activity.

Table S3.2. Top 7[†] off-target predictions for SpRY6 (CTAATAGGCAGAGAGAGTCA) as ranked by CRISPOR.[1] PAM = NYN

Off-target Sequence	Mismatch Positions	Mismatch Count	MIT Off-target	CFD Off-target	Chromosome	Strand	Locus Description
CCAATAGGCAGAGAGAGTCAGTG	.*.....	1	100	0.033395	Chr11	+	exon:HBB
ACAATAGGCAGAGAGAGTCAATT	**.....	2	5.21978	0	Chr18	-	intron:KIAA1328
CTAATAGGGTGAGAGAGTCAGCA	.....**.....	2	2.937332	0	Chr4	+	intergenic:RPS23P3-RNU6-699P
CTAAGAGGCAGAGAGAATCAACC	....*.....*...	2	3.890957	0	Chr7	+	intergenic:AC011288.2-AC005019.3
CTGATAGGCTGAGAGAGTCAACA	..*.....*.....	2	6.438065	0	Chr8	+	intergenic:NIPAL2-KB-1458E12.1
CTGGGAGGCAGAGAGAGTCAACC	..***.....	3	2.287424	0	Chr10	-	intergenic:RP11-810B23.1-MTND5P17
CTCTCAGGCAGAGAGAGTCAGTT	..***.....	3	2.287424	0	Chr2	+	intergenic:FAM84A/AC011897.1-AC068286.1

Yellow highlight identifies genes sequenced for off-target activity.

[†]*For this PAM, Off-targets 8-10 are not given by the software. Additionally, the top ranked predicted gene (HBB) was considered to be on-target and therefore skipped in our targeted sequencing assay.*

Table S4.1. Top 10 off-target predictions for SpRY20 (TGCCTATTAGTCTATTTTCCCAC) as ranked by CRISPOR.[1] PAM = NNN

Off-target Sequence	Mismatch Positions	Mismatch Count	MIT Off-target	CFD Off-target	Chromosome	Strand	Locus Description
TGTCTATTATTCTATTTTCCAGG	..*.....*.....	2	6.438065	0.55859375	Chr6	+	intergenic:NCOA7-AS1-NCOA7
TACCTATTAGTCAATTTTCCCTG	.*.....*.....	2	3.604411765	0.022823	Chr18	-	intergenic:SNRPGP2-RP11-61D1.2
TTCCTATTAGTCTATTTTTCAGT	.*.....*.	2	5.541666667	0.004737	Chr13	-	intergenic:RPL7P45-DAOA-AS1
TGCCTATTGGTCTATTTTCCCAC	.....*.....	1	61.1	0	Chr11	-	exon:HBB
TGCCTATTTGTCTATTTTCCTCT	.....*.....	1	61.1	0	Chr2	+	intergenic:ERMN-FAM133DP
AGCCTATTAGTCTATTTCCCTCT	*.....*..	2	3.448148148	0	Chr9	+	intron:RAPGEF1
TGCCTAGGAGTCTATTTTCCCCA	.....**.....	2	3.56510989	0	Chr6	-	intergenic:PRIM2-GAPDHP41
TGCCTTTTACTCTATTTTCCAAT	.....*...*.....	2	3.350283228	0	Chr4	+	intergenic:FAT4-RP11-399F2.2
TGCCCATTAGTCTATTTTCAACA	....*.....*	2	5.659285714	0	Chr10	+	intergenic:YWHAZP5-RP11-56I23.1
TGCCTATCAGTCTATTTTCATTC	.....*.....*	2	4.214361702	0	Chr1	-	intergenic:TMA16P2-MAST2

Yellow highlight identifies genes sequenced for off-target activity.

Table S4.2. Top 10 off-target predictions for SpRY20 (TGCCTATTAGTCTATTTTCCCAC) as ranked by CRISPOR.[1] PAM = NRN

Off-target Sequence	Mismatch Positions	Mismatch Count	MIT Off-target	CFD Off-target	Chromosome	Strand	Locus Description
TGTCTATTATTCTATTTTCCAGG	..*.....*.....	2	6.438065	0.55859375	Chr6	+	intergenic:NCOA7-AS1-NCOA7
GGTTTATTAGTCTATTTTCCTGG	*.**.....	3	2.287424	0.526086956	Chr7	+	intergenic:AC009541.1-hsa-mir-490
TGCCTATTTTTCTATTTTCCCAG	.....**.....	2	2.937332	0.126388889	Chr7	-	intergenic:AC006007.1-RN7SL207P
TACATATTAATCTATTTTCCAGC	.*.*.....*.....	3	2.461181	0.014778827	Chr21	+	intergenic:RNU1-139P-RPL37P4
ATCATATTAGTCTATTTTCCTGC	**.*.....	3	2.319902	0.011908559	Chr1	+	intergenic:RP11-445J9.1-RP11-113I24.1
GGTGTATTAGTCTATTTTCCTGC	*.**.....	3	2.287424	0.007306763	ChrX	-	intergenic:RPS6KA3-RN7SKP183
TTCCTATTAGTCTATTTTTCAGT	.*.....*.	2	5.541667	0.004737198	Chr13	-	intergenic:RPL7P45-DAOA-AS1
TGCCTATTGGTCTATTTTCCCAC	.....*.....	1	61.1	0	Chr11	-	exon:HBB
TGCCTTTTACTCTATTTTCCAAT	.....*...*.....	2	3.350283	0	Chr4	+	intergenic:FAT4-RP11-399F2.2
CTCCCATTAGTCTATTTTCCCAT	**..*.....	3	2.426564	0	Chr13	+	intergenic:SRGNP1-RNY3P3

Yellow highlight identifies genes sequenced for off-target activity.

Supplemental Figures

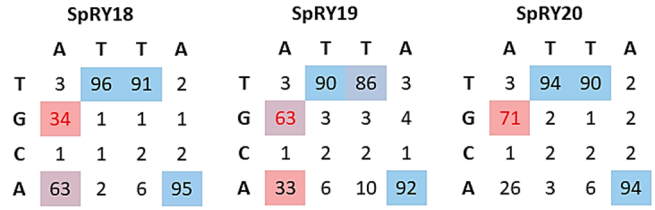


Figure S1. Output of EditR software showing efficiency of induced IVSI-107(A>G) base changes. *Editing efficiency (A>G base change) indicated in red numbers. gRNAs differ in the length of their PAM-distal, 5' sequence, as indicated in their names (gRNA for SpRY18 with 18 nt total length, up to 20 nt length for SpRY20).*

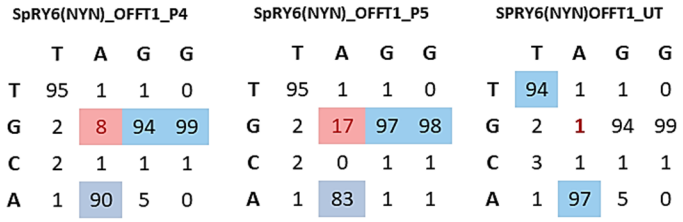


Figure S2. Output of EditR software showing off-target activity of SpRY6 editor in top predicted gene for NYN (KIAA1328, Table S3). *Editing efficiency (A>G base change) indicated in red numbers. Data shown for two patients (Patients 4 and 5, indicated as P4 and P5).*

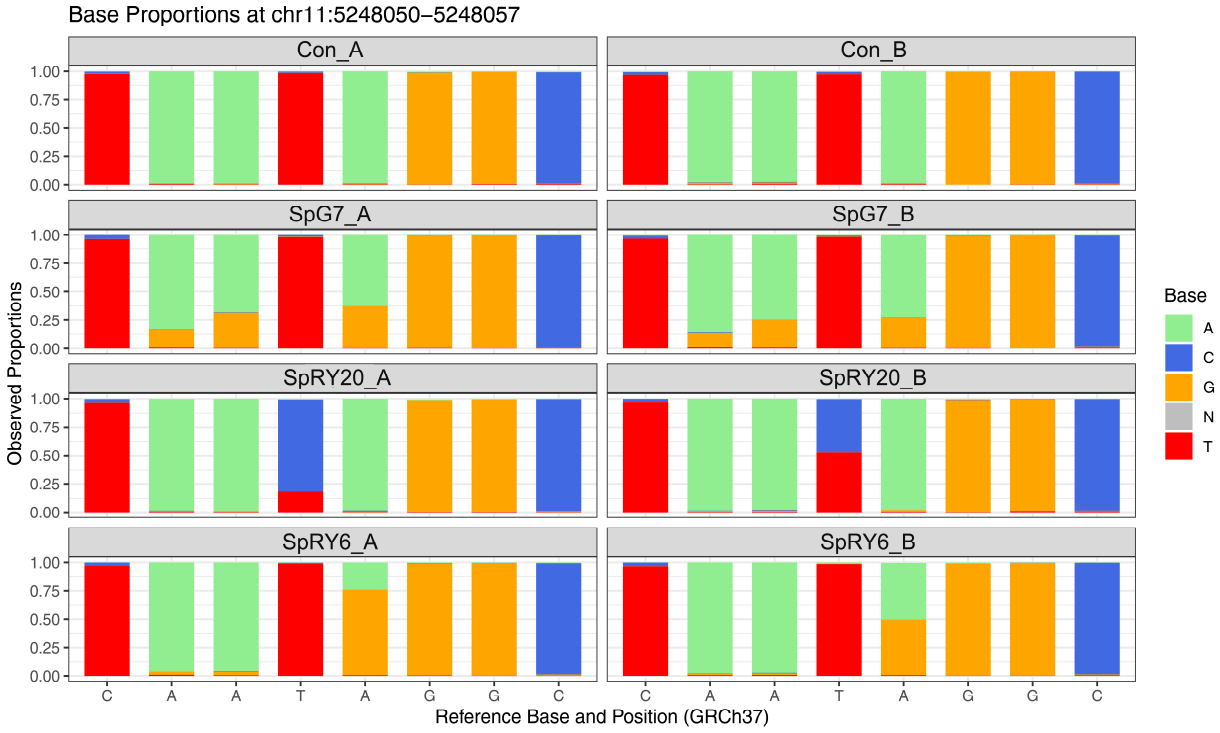


Figure S3. Plot illustrating the base composition of positions in the target region *indicating comparable editing distribution for two independent samples.*

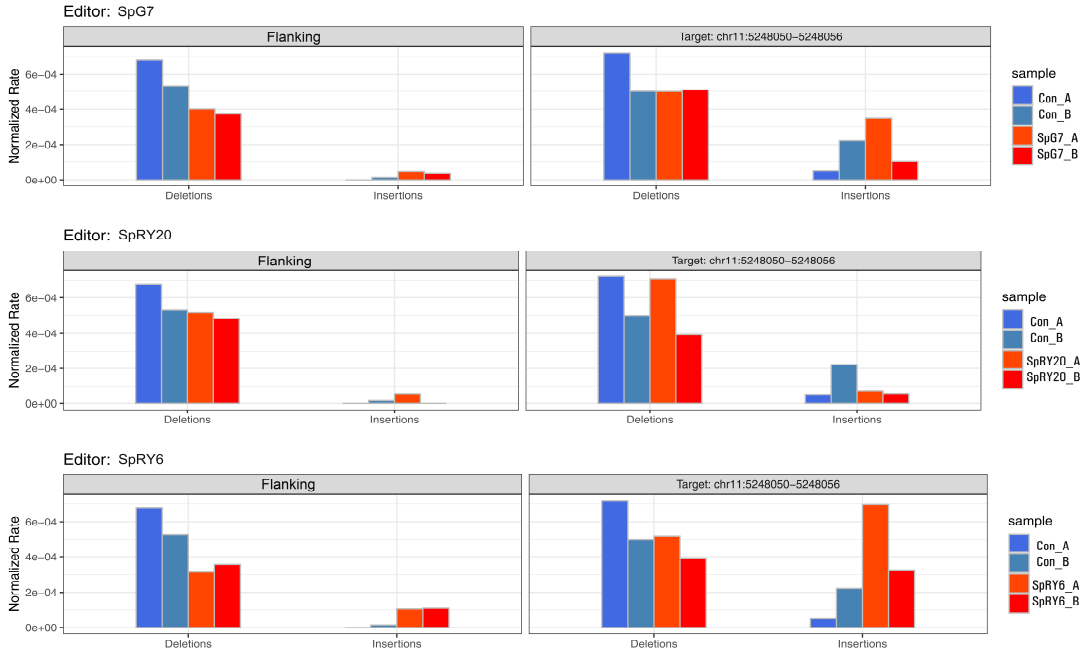


Figure S4. Bar plots illustrating normalized rate of indels *for reads overlapping target and flanking sequences of the region of interest (ROI) of two patients (A=Patient 5, B=Patient 4).*

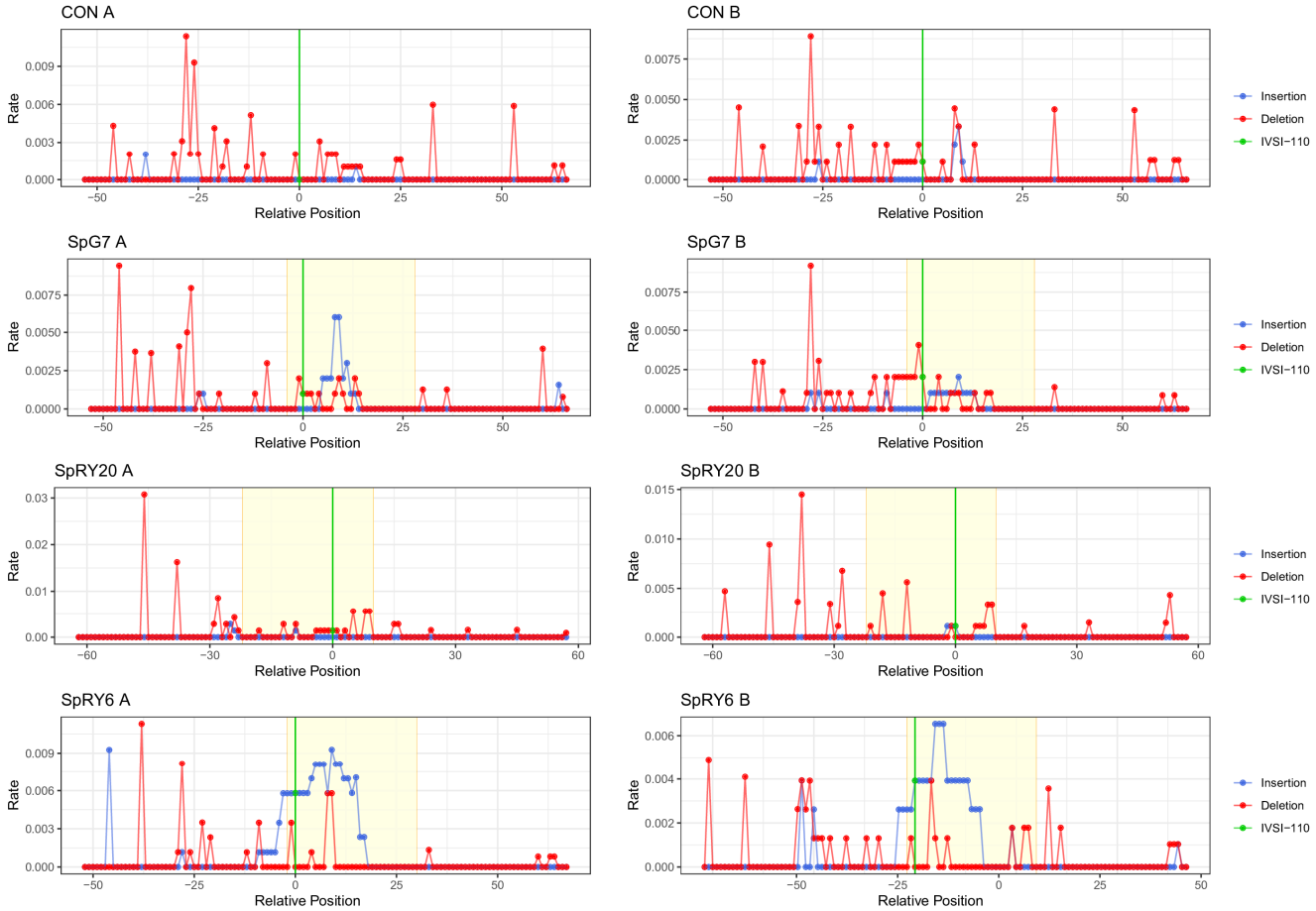


Figure S5. Illustration of insertion (blue) and deletion (red) rates *at each position of the targeted region (shaded yellow) with the IVS1-110 indicated in green.*

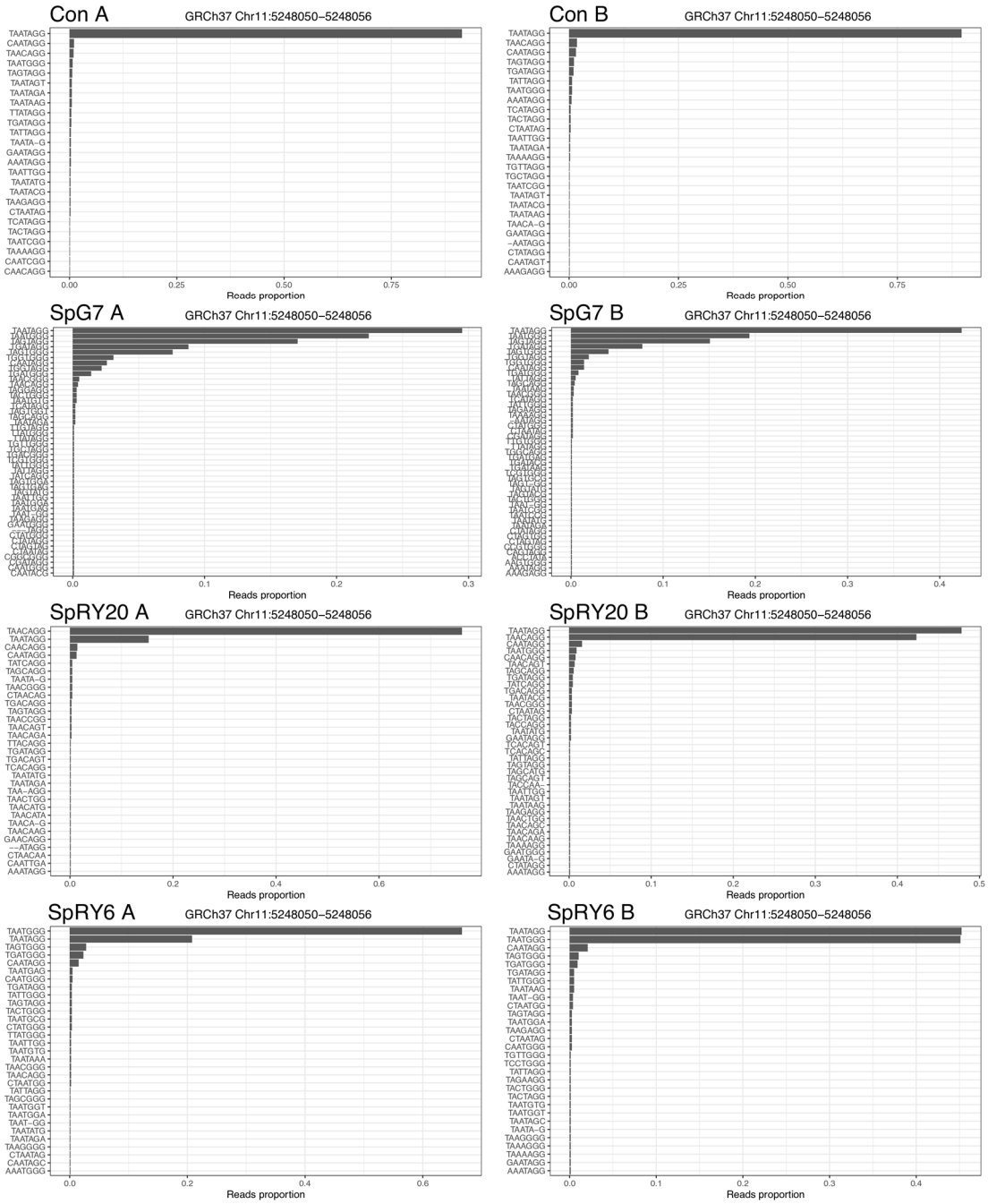


Figure S6. Plot showing relative proportion of 7-base oligonucleotides overlapping the target site as extracted from single reads. *(Chr11:5248050 = IVSI-110; Chr11:5248051 = IVSI-109; Chr11:5248052 = IVSI-108; Chr11:5248053 = IVSI-107; Chr11:5248054 = IVSI-106).*

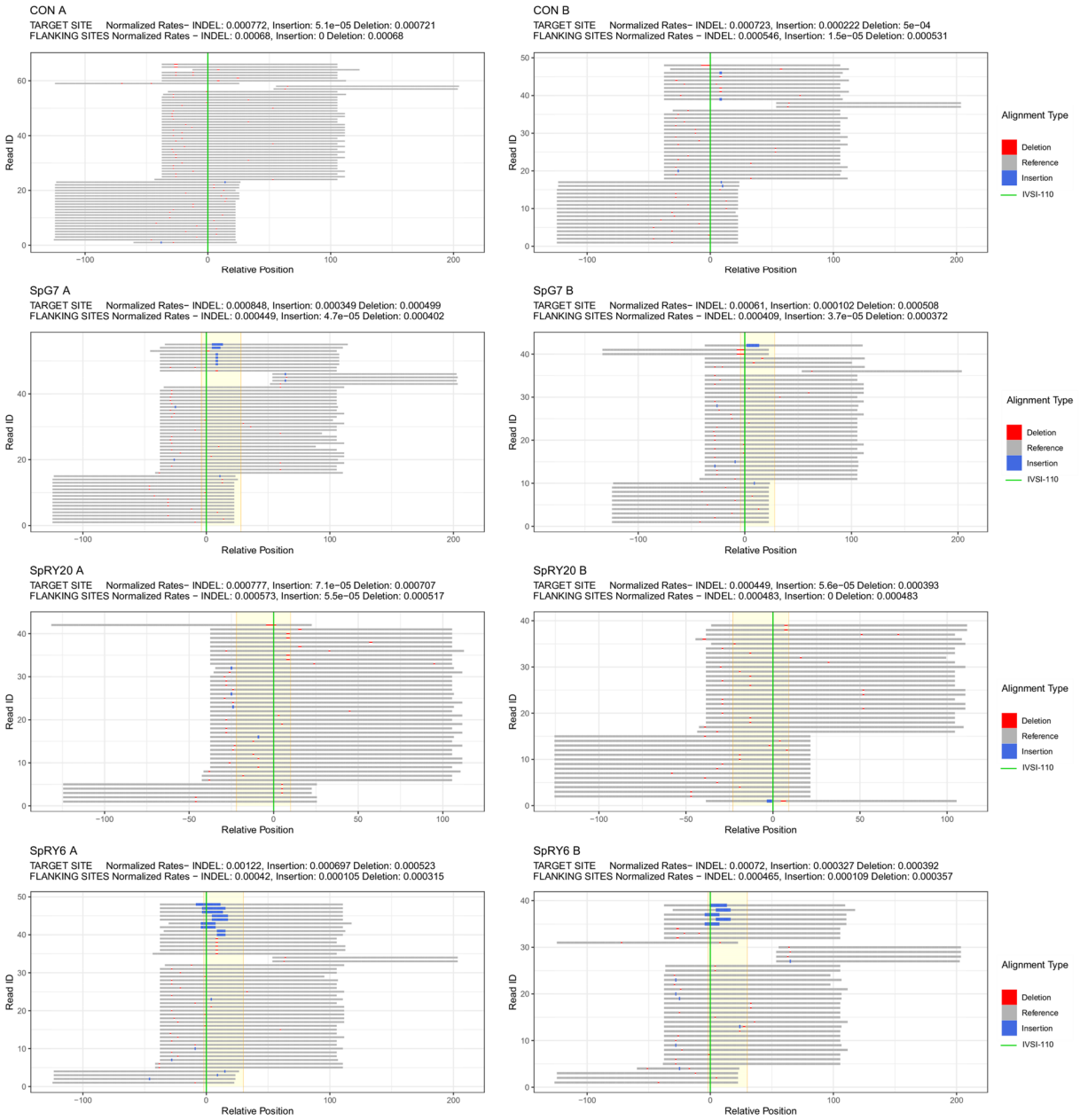


Figure S7. Representative alignment of reads overlapping the region of interest (ROI) ranked based on the size of the indel (large to small) in the target region. Grey bars indicate read segments matching the reference genome, and the yellow shaded region marks the binding loci of the gRNA. The green line indicates the IVSI-110(G>A) mutation site.

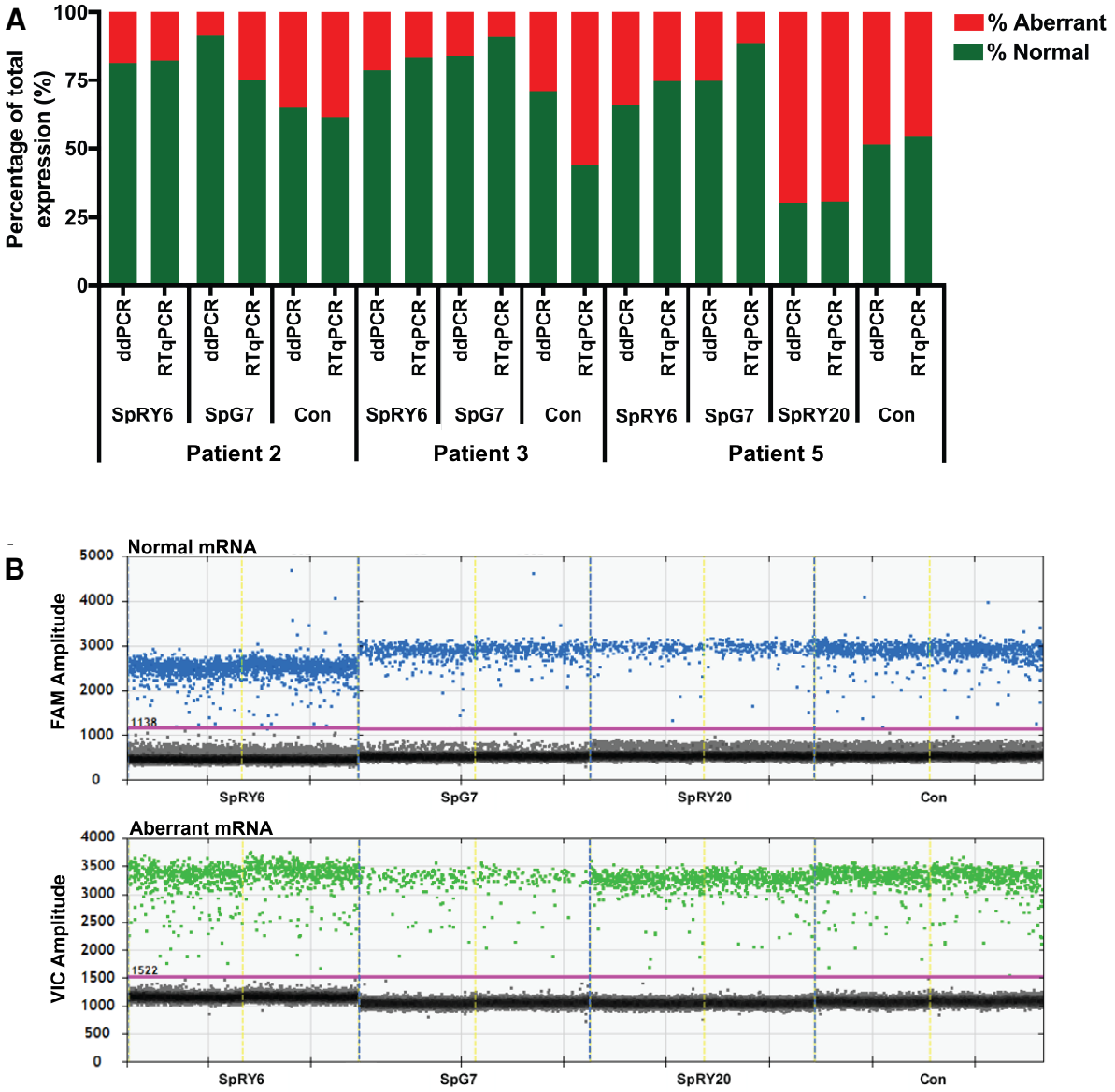
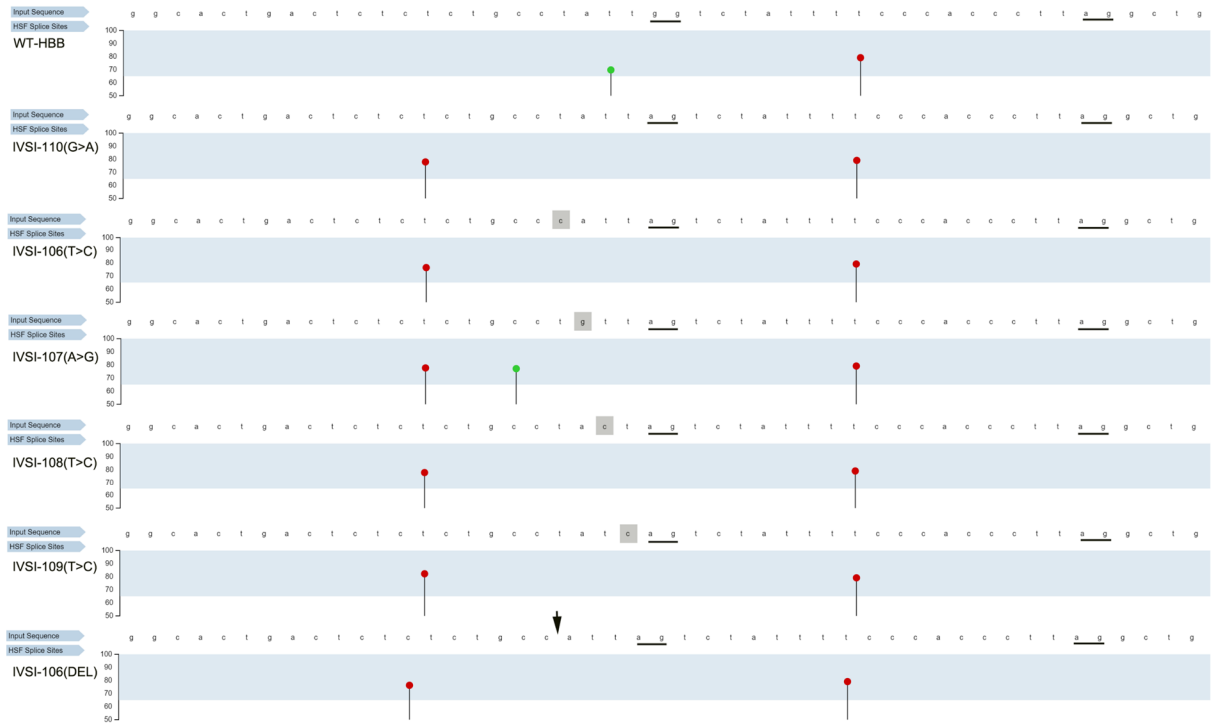
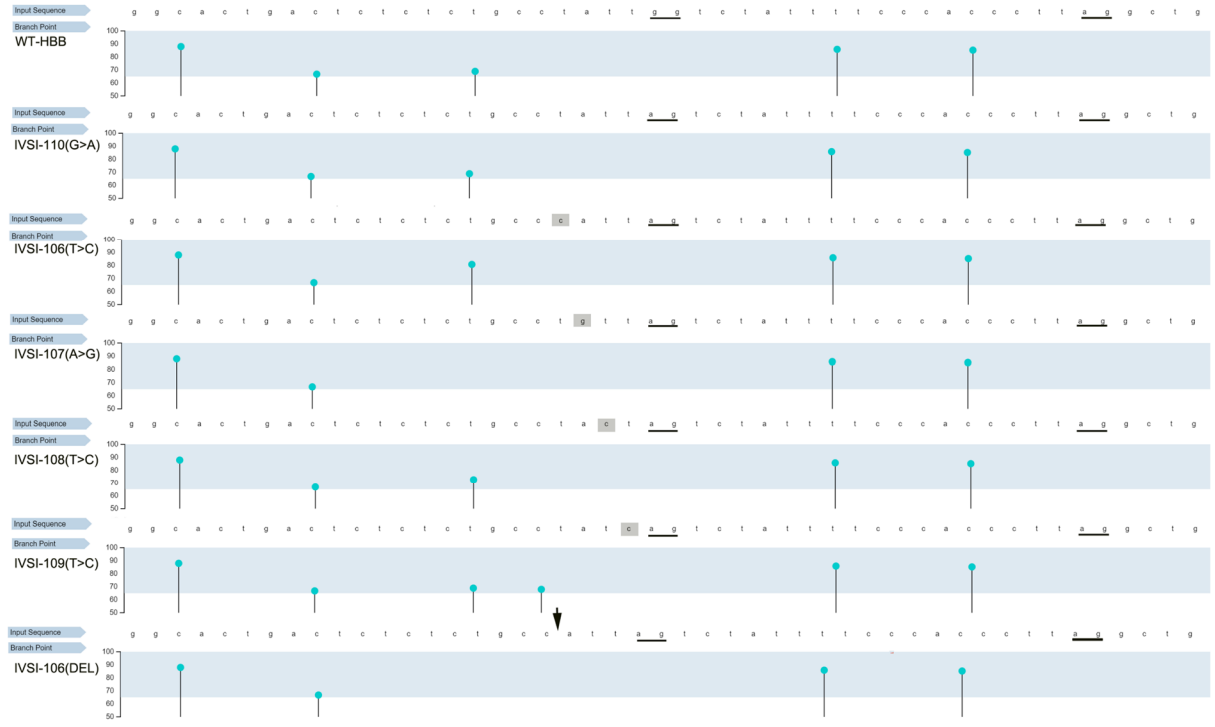


Figure S8. Assessment of splicing at the transcript level in patient-derived CD34$^+$ cells on day 11 of erythroid differentiation via multiplex RT-qPCR and droplet digital PCR. (A) Percentage of normal and aberrant HBB mRNAs of three patients (2, 3, and 5), as measured by two different quantitative technologies, RT-qPCR and (RT-)ddPCR. Of the samples shown, SpRY20 was only available for Patient 5. (B) FAM and VIC droplet amplitudes of Patient 5, as exported from ddPCR QuantaSoft software. The pink line represents the threshold that separates the positive and negative droplet clusters. The vertical yellow line separates two same-plate technical replicates for the same sample. Blue droplets are positive for the normal mRNA and green for the aberrant mRNA. Grey droplets do not contain the specific target. For ddPCR data not amplitude height but the ratio of droplets classified as positive and negative are the basis for quantification. Statistical analysis of the percentage of normal HBB mRNA in SpG7- (n=3) and SpRY6- (n=3) base edited samples relative to untreated control (UT; n=3) using parametric paired repeated measures one-way ANOVA with an assumption of sphericity and followed by Dunnett's multiple comparisons test. Statistical analysis showed a significant increase in the percentage of normal HBB mRNA measured via RTqPCR in SpG7 (84.73±8.67%; *p = 0.0242), SpRY6 (79.97±4.76%; *p = 0.0412) relative to UT (53.23±8.65%). Similarly, statistical analysis showed a significant increase in the percentage of normal HBB mRNA measured via RT-ddPCR in SpG7 (83.27±8.46%; ** p = 0.0040), SpRY6 (75.3±8.18%; *p = 0.0223) relative to UT (62.53±10.1%).

A


Input Sequence
HSF Splice Sites
WT-HBB
IVSI-110(G>A)
IVSI-106(T>C)
IVSI-107(A>G)
IVSI-108(T>C)
IVSI-109(T>C)
IVSI-106(DEL)

B


Input Sequence
Branch Point
WT-HBB
IVSI-110(G>A)
IVSI-106(T>C)
IVSI-107(A>G)
IVSI-108(T>C)
IVSI-109(T>C)
IVSI-106(DEL)

C

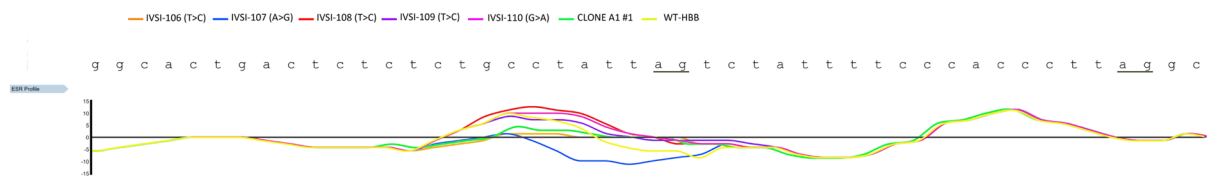

IVSI-106 (T>C)
IVSI-107 (A>G)
IVSI-108 (T>C)
IVSI-109 (T>C)
IVSI-110 (G>A)
CLONE A1 #1
WT-HBB
ESR Profile

Figure S9. *In silico* prediction of corrective effects on the splicing machinery. We employed the Human Splice Finder (HSF) software,[2] in order to elucidate the mechanism underlying the successful restoration of patient cell phenotype upon manipulation of the upstream intronic sequence using our ABEs for IVSI-106|108|109 (T>C) substitutions, as well as to explain the cause of deterioration of the IVSI-110 thalassemia phenotype upon IVSI-107(A>G) substitution, (A) *Representation of predicted donor and acceptor sites* across the sequence input, which is that of the HBB exon-intron junction, wherein the IVSI-110 mutation resides. Splice acceptor sites are indicated in red, donor sites in green. Base changes are highlighted in gray, acceptor sites are underlined in black, and an arrow is used to indicate the missing base in Clone A1 #2. The HSF software accurately predicted the gain of an acceptor site upon insertion of the IVSI-110(G>A) mutation. In the case of IVSI-107(A>G) (SpRY20 base editing), the formation of a novel donor site between the aberrant and normal splice acceptor sites is predicted in comparison to all other listed edits and the normal HBB locus. (B) *Representation of predicted branchpoint sites* across the sequence input, which is that of the HBB exon-intron junction wherein the IVSI-110 mutation resides. Base changes are highlighted in gray, acceptor sites are underlined in black, and an arrow is used to indicate the missing base in Clone A1 #2. The IVSI-107(A>G) base transition (SpRY20 base editing) and the IVSI-106 deletion (corresponding to the functionally corrected clone A1 #2 from our previous work)[3] are predicted to lose the same branch point site, indicating functional inutility of this site for normal HBB splicing. IVSI-109, on the other hand, appears to gain an extra branchpoint site. (C) *ESR profile of the HBB locus input sequence*, showing similar profiles for SpRY6 and corrective clone A1 #2. ESR = exonic splice enhancers ÷ exonic splicing silencers. Aberrant and normal splice acceptor sites are underlined. A change in trend is indicated when a base is edited, or deleted, as with our restorative clone A1 #2.[3] Compared to the curve for A1 #2 as reference for correction of splicing, a similar lowering to baseline is noticed upon IVSI-106 mutation (SpRY6), whereas the curve dips well below the baseline upon IVSI-107 mutation (SpRY20).

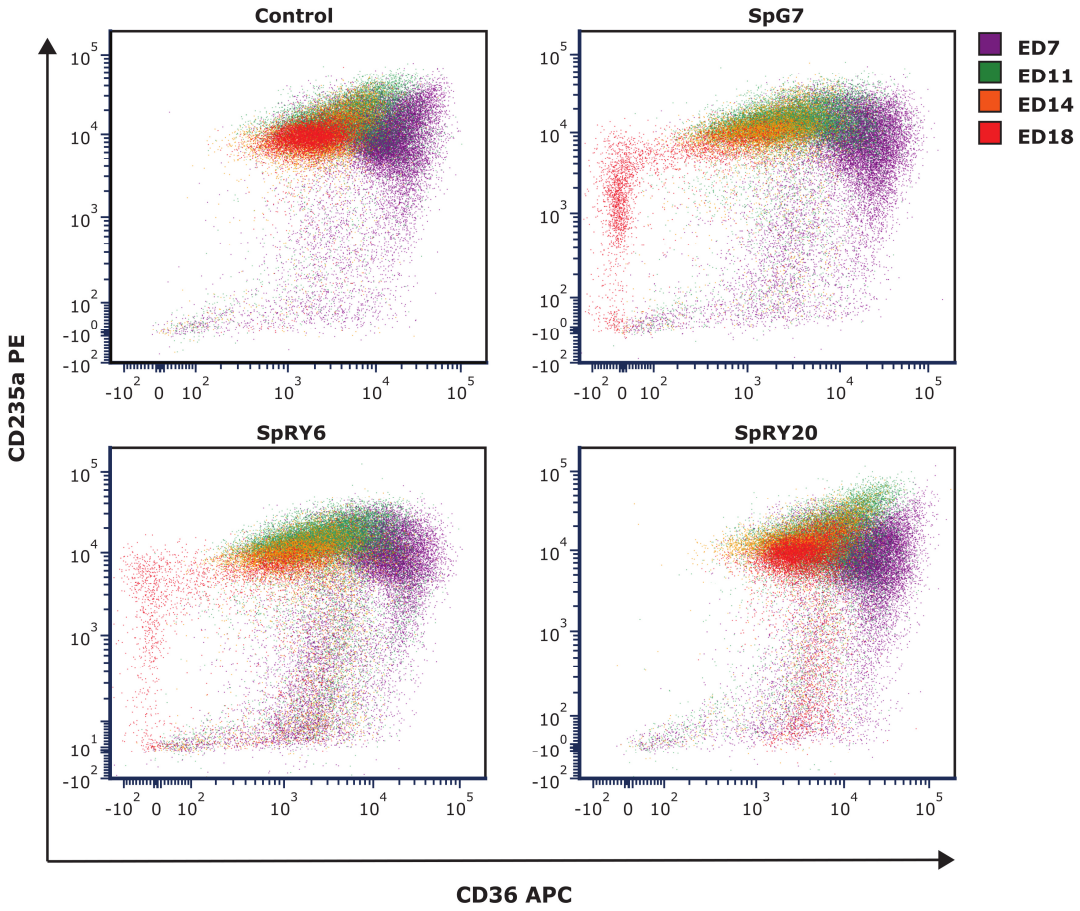


Figure S10. Representative flow cytometry color dot plots illustrating genome-edited and control cells stained for differentiation markers CD235a and CD36. Colors indicate stages of differentiation as different collection timepoints (ED7: purple; ED11: green; ED14: orange; ED18: red).

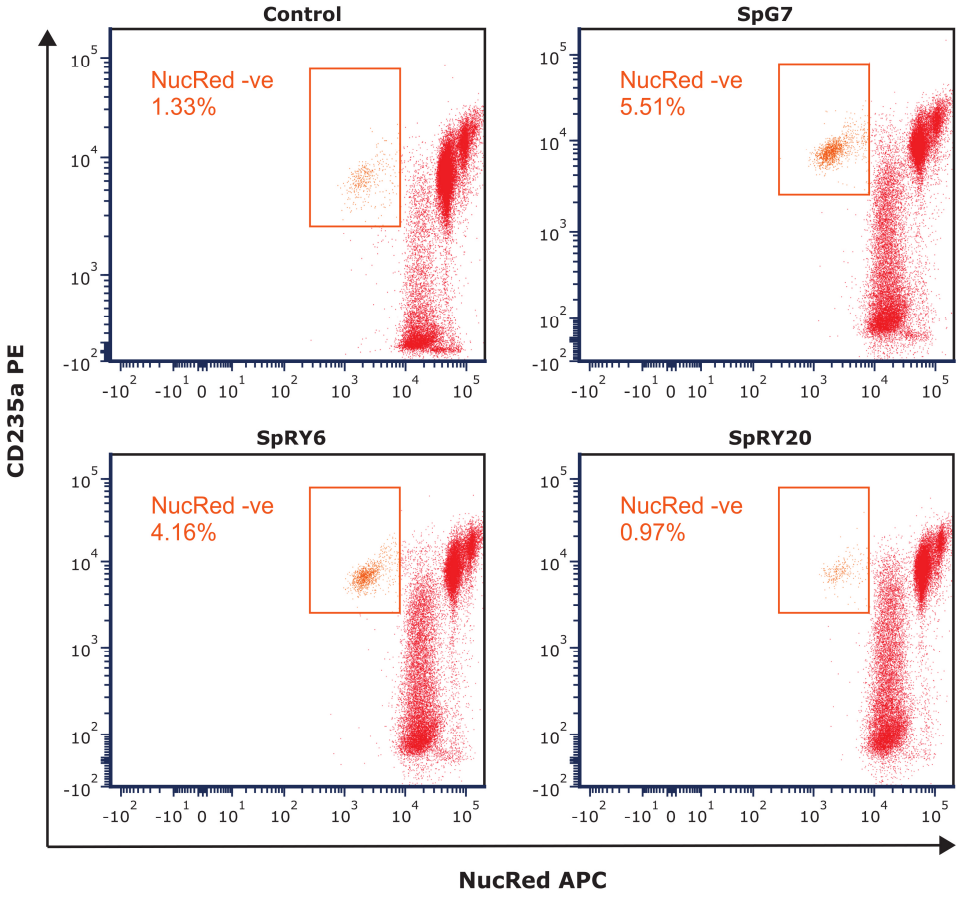


Figure S11. Representative flow cytometry color dot plots illustrating genome-edited and control cells stained for differentiation marker CD235a and nucleation marker NucRed on ED18. Percentage of enucleated erythroid cells were gated as CD235a^{+}/NucRed^{-} (in orange).

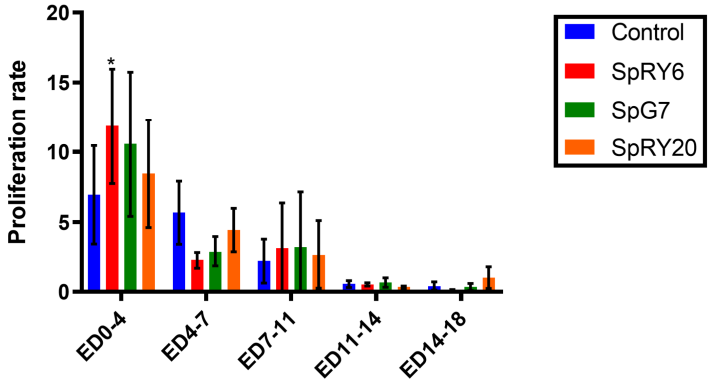


Figure S12. Proliferation rate of cells during erythroid differentiation at different time points (ED0-4, ED4-7, ED7-11, ED11-14 and ED14-18). Cell counts indicate a proliferation advantage in cells treated with corrective SpRY6 and SpG7, as compared to control, in the initial stage of differentiation. Contrarily, the final stage shows higher proliferation and corresponding delayed differentiation of the SpRY20-treated cells, in line with the treatment's worsening of the thalassemic disease phenotype. Groupwise comparisons with Control were performed in triplicate and significance obtained using two-way ANOVA with Dunnett's multiple comparison test. SpRY6 for ED0-4: *p = 0.0431.

Supplemental References

1. Concordet JP, Haeussler M. CRISPOR: Intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res. 2018;46(W1):W242–5.

2. Desmet FO, Hamroun D, Lalande M, Collod-Bëroud G, Claustres M, Béroud C. Human Splicing Finder: An online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009;37(9):1–14.

3. Patsali P, Mussolino C, Ladas P, Floga A, Kolnagou A, Christou S, Sitarou M, Antoniou MN, Cathomen T, Lederer CW, et al. The Scope for Thalassemia Gene Therapy by Disruption of Aberrant Regulatory Elements. J Clin Med. 2019 Nov;8(11).