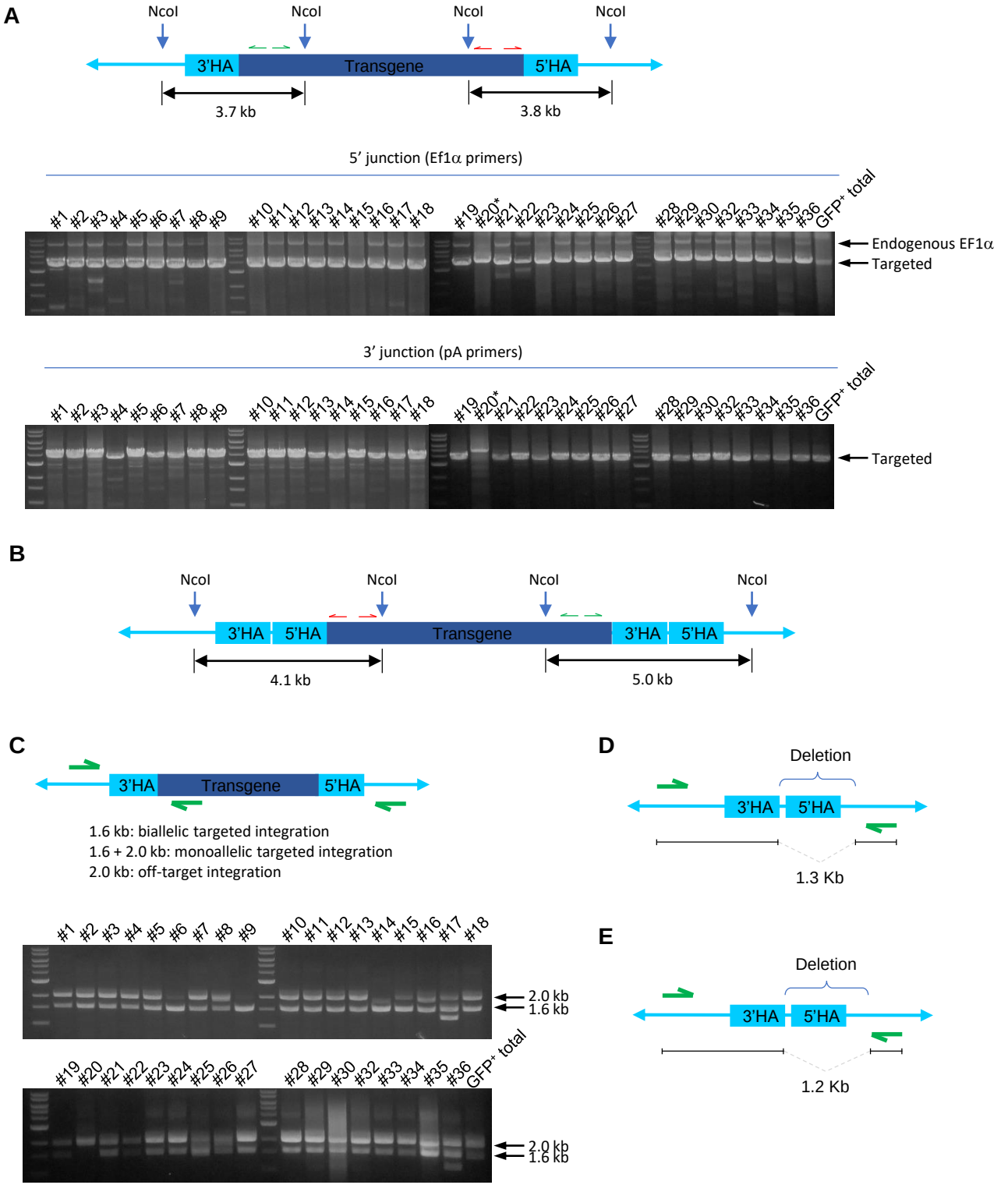
YMTHE, Volume *27*

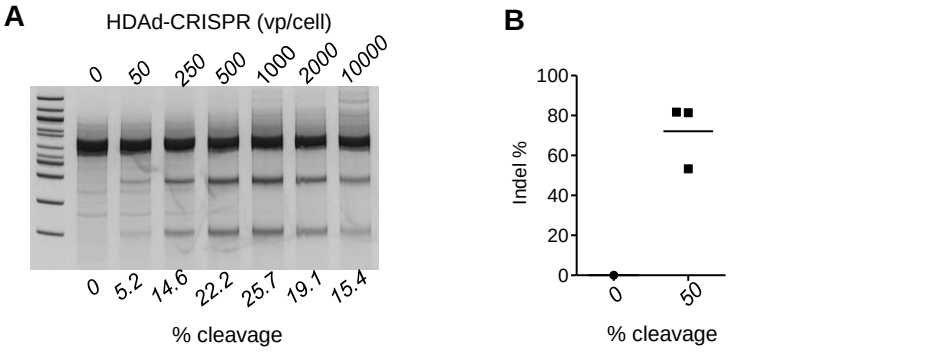
Supplemental Information

Targeted Integration and High-Level Transgene Expression in AAVS1 Transgenic Mice after *In Vivo* HSC Transduction with HDAd5/35++ Vectors

Chang Li, Arpit Suresh Mishra, Sucheol Gil, Meng Wang, Aphrodite Georgakopoulou, Thalia Papayannopoulou, R. David Hawkins, and André Lieber


A
NcoI
NcoI
NcoI
NcoI
3'HA
Transgene
5'HA
3.7 kb
3.8 kb
5' junction (Ef1α primers)
#1 #2 #3 #4 #5 #6 #7 #8 #9 #10 #11 #12 #13 #14 #15 #16 #17 #18 #19 #20* #21 #22 #23 #24 #25 #26 #27 #28 #29 #30 #32 #33 #34 #35 #36 GFP+ total
Endogenous EF1α
Targeted
3' junction (pA primers)
#1 #2 #3 #4 #5 #6 #7 #8 #9 #10 #11 #12 #13 #14 #15 #16 #17 #18 #19 #20* #21 #22 #23 #24 #25 #26 #27 #28 #29 #30 #32 #33 #34 #35 #36 GFP+ total
Targeted
B
NcoI
NcoI
NcoI
NcoI
3'HA
5'HA
Transgene
3'HA
5'HA
4.1 kb
5.0 kb
C
3'HA
Transgene
5'HA
1.6 kb: biallelic targeted integration
1.6 + 2.0 kb: monoallelic targeted integration
2.0 kb: off-target integration
#1 #2 #3 #4 #5 #6 #7 #8 #9 #10 #11 #12 #13 #14 #15 #16 #17 #18
2.0 kb
1.6 kb
#19 #20 #21 #22 #23 #24 #25 #26 #27 #28 #29 #30 #32 #33 #34 #35 #36 GFP+ total
2.0 kb
1.6 kb
D
Deletion
3'HA
5'HA
1.3 Kb
E
Deletion
3'HA
5'HA
1.2 Kb

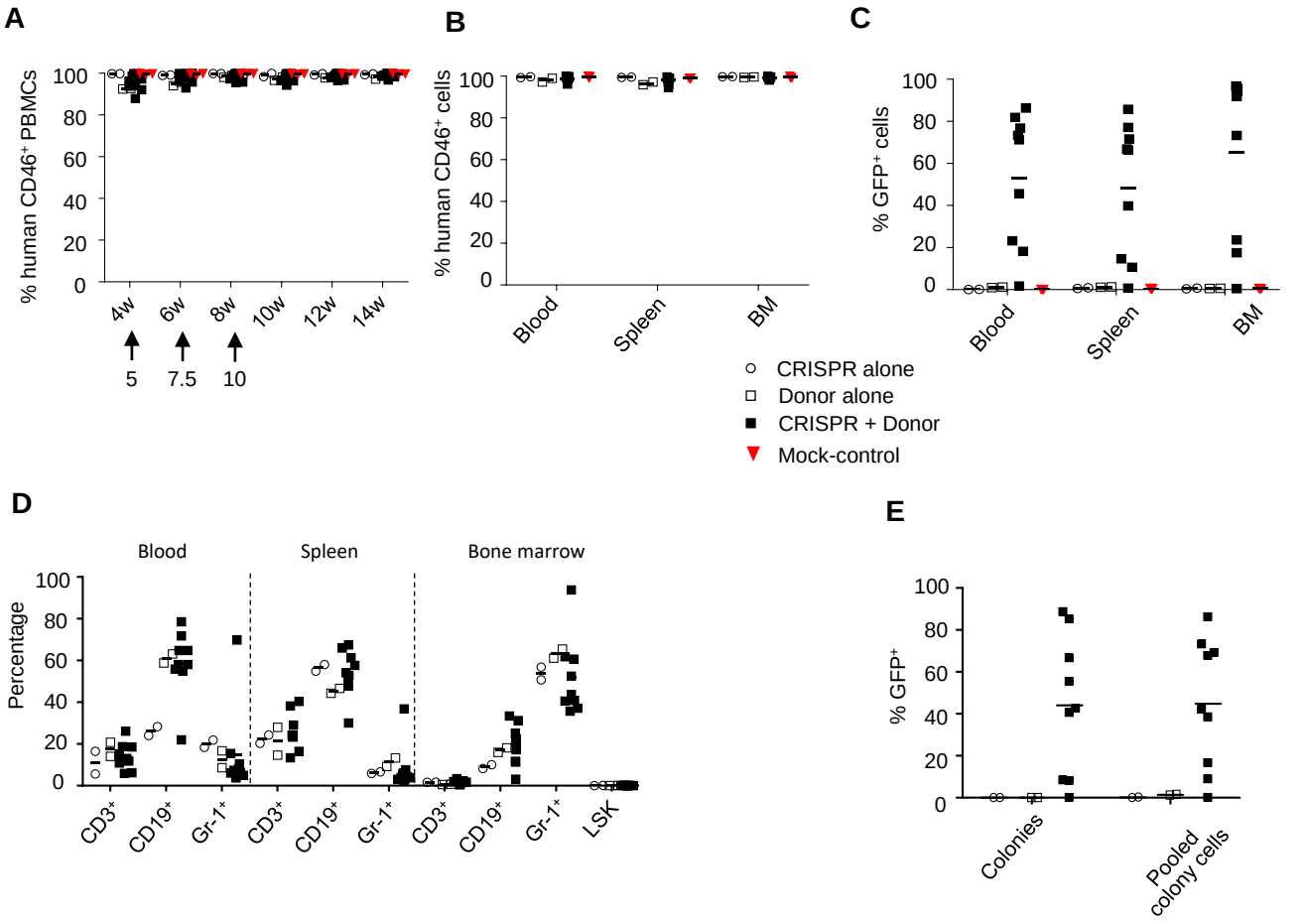
Supplementary Figure S1. Integration analysis of HUDEP-2 clones transduced with vectors for targeted integration. A) Integration site analysis by inverse PCR. The upper diagram shows the locations of *NcoI* sites and primers (half arrows. Red: EF1α primers for 5'- junctions; green: pA primers for 3' junctions). HA: homology arms. The expected amplicon size at each side for targeted integration is indicated. The lower gel pictures show iPCR results. Each lane represents one cell clone. An extra band derived from the endogenous EF1α promoter was detected. For clone #20, although the amplicon size is different from prediction, cloning and sequencing revealed that it is a clone with targeted integration. **B)** Diagram showing the integration pattern in clone #20. **C)** In-/out-PCR analysis. The upper diagram shows the location of primers. Expected product sizes for various integration patterns are listed. The lower gel pictures demonstrate that most clones had monoallelic targeted integration. **D-E)** Diagram explaining the origin of the smaller bands in #17 and #36 in C), respectively. Both smaller bands resulted from large fragment deletions.



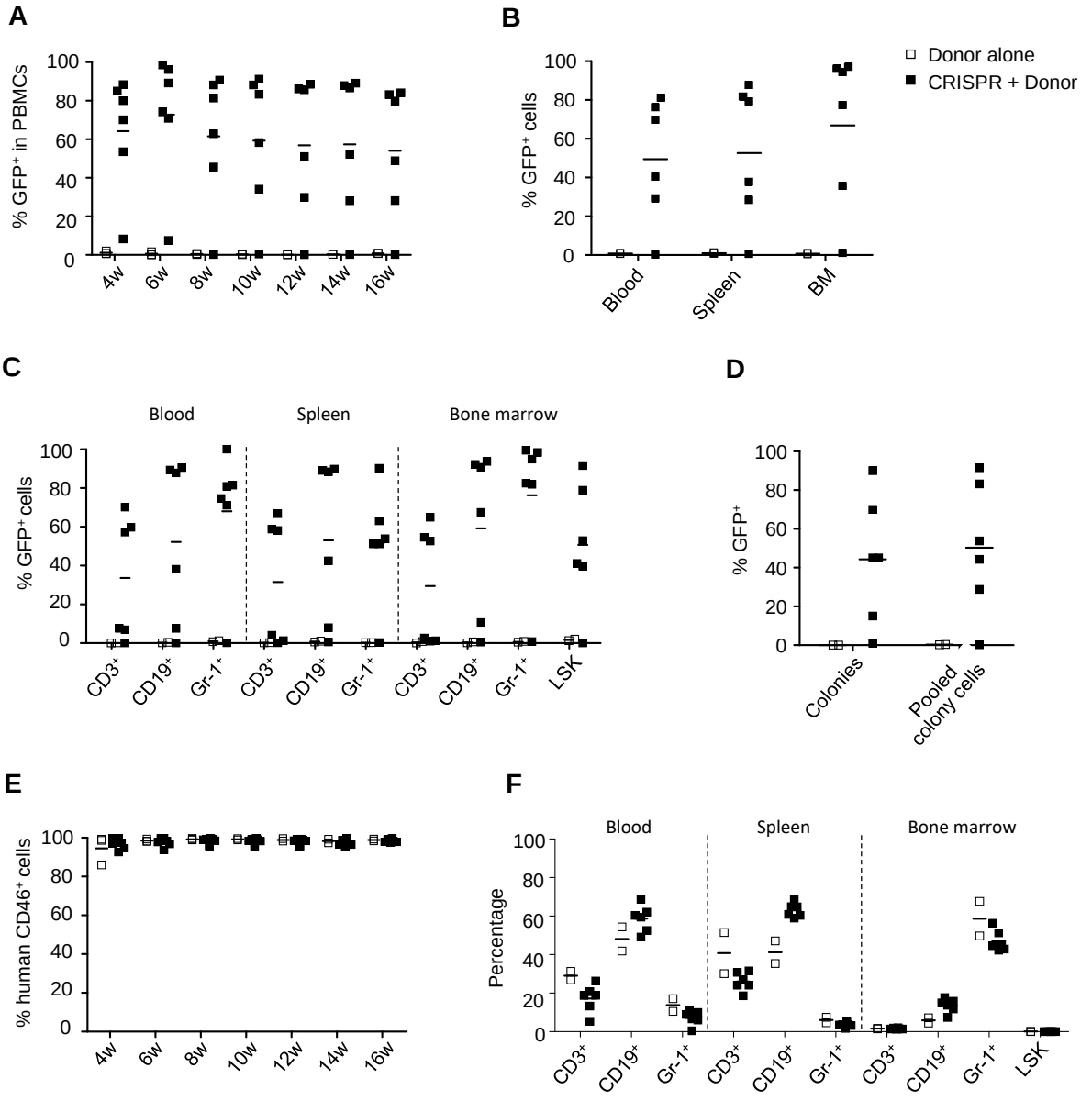
C

			# of reads	% of reads
	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	0	19505	15.8%
1	TGTCCCCTCCACCCCACAGTGGGGCC-------ACAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-7	348	0.3%
2	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACA-----GGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-5	10016	7.5%
3	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACA-----GGTGACAGAAAATCCCCATCCTTAGGCCTCCTCCTTCCT	-5	899	0.7%
4	TGTCCCCTCCACCCCACAGTGGGGCCACTAG-----GGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-5	703	0.5%
5	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACA-----GGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-5	699	0.5%
6	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACA-----GGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-5	300	0.2%
7	TGTCCCCTCCACCCCACAGTGGGGCCACTAGG----GGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-4	418	0.3%
8	TGTCCCCTCCACCCCACAGTGGGGCCACTAGG---AGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	13995	10.5%
9	TGTCCCCTCCACCCCACAGTGGGGCCACTAG---CAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	10311	7.7%
10	TGTCCCCTCCACCCCACAGTGGGGCCACTAGG---CGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	1586	1.2%
11	TGTCCCCTCCACCCCACAGTGGGGCCACTAG---CCGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	1113	0.8%
12	TGTCCCCTCCACCCCACAGTGGGGCCACTAGG---AGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	713	0.5%
13	TGTCCCCTCCACCCCACAGTGGGGCCACTAGG---AGTATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	524	0.4%
14	TGTCCCCTCCACCCCACAGTGGGGCCACTAG---CAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	503	0.4%
15	TGTCCCCTCCACCCCACAGTGGGGCCACTAGG---AGGATTGGTGACAGAACAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	361	0.3%
16	TGTCCCCTCCACCCCACAGTGGGGCCACTAG---CAGTATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	335	0.3%
17	TGTCCCCTCCACCCCACAGTGGGGCCACTAG---CAGGATTGGTGACAGAACAGCCCCATCCTTAGGCCTCCTCCTTCCT	-3	290	0.2%
18	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGAC--GATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	17441	13.1%
19	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGG--AGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	16863	12.7%
20	TGTCCCCTCCACCCCACAGTGGGGCCACTAGG--CAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	6392	4.8%
21	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGG--ATGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	600	0.5%
22	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGAC--GCTTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	566	0.4%
23	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGG--AGGCTTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	559	0.4%
24	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGAC--GATTGGTGACAGACAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	452	0.3%
25	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGG--AGGATTGGTGACAGACAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	445	0.3%
26	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGAC--GATTGGTGCCAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	378	0.3%
27	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGG--AGGATTGGTGCCAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-2	318	0.2%
28	TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACA-GATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCT	-1	544	0.4%

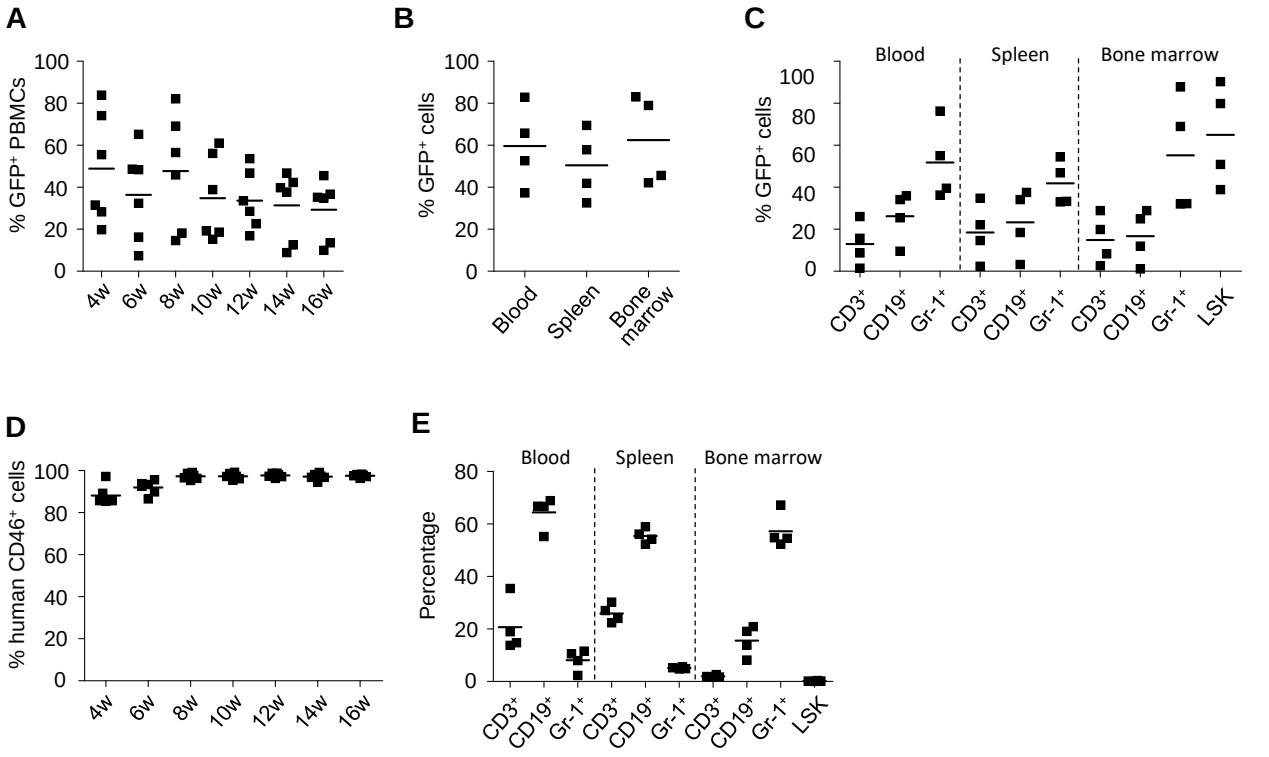
Supplementary Figure S2. Cleavage of AAVS1 target site in AAVS1/CD46tg mice. A) *In vitro* analysis. Target site cleavage frequency in bone marrow lineage-negative cells from AAVS1/CD46tg mice measured three days after HDAd-CRISPR transduction at the indicated MOIs. **B)** Percentage of total AAVS1 indels obtained by deep sequencing of DNA from total bone marrow mononuclear cells at week 16 after transplantation (see Fig.3A). Each symbol is an individual animal. **C)** Top 30 most frequent indels found in a mouse. Representative data are shown. The yellow sequence shows the target of the guide RNA with the PAM sequence marked in blue. The CRISPR/Cas9 cleavage site is marked by a vertical arrow.



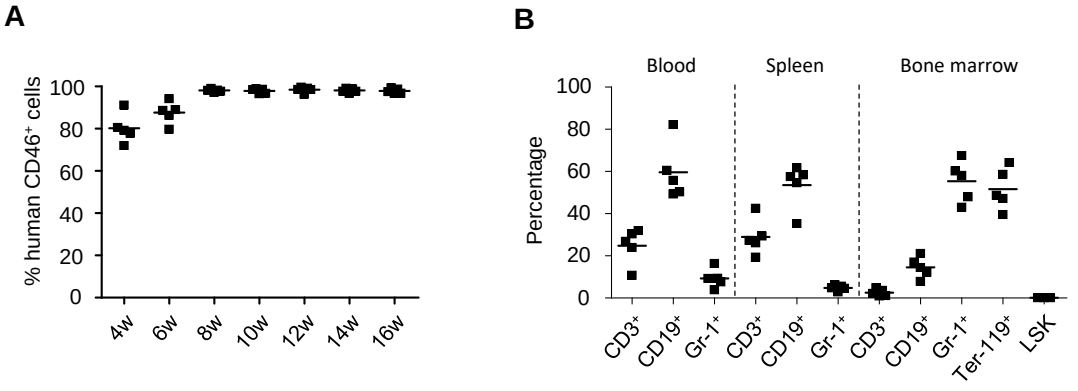
Supplementary Figure S3. Analysis of engraftment of *ex vivo* transduced Lin$^-$ cells. A) Engraftment of transplanted cells based on human CD46 expression in PBMCs measured by flow cytometry. Each symbol is an individual animal. Notably, transduced donor cells expressed CD46, while recipient C57Bl/6 mice did not. N=3 for "CRISPR alone"; N=3 for "Donor alone"; N=10 for "CRISPR+Donor". "Mock-control" are cells that were incubated with the same volume of virus storage buffer (PBS/10% glycerol). **B)** Percentage of CD46-positive cells in PBMCs (blood), spleen, and bone marrow at week 16. **C)** Percentage of GFP-positive cells in PBMCs, spleen and bone marrow, at week 16. **D)** Percentage of LSK and lineage-positive cells. The difference between the three groups is not significant. **E)** Analysis of GFP$^+$ colonies. Total bone marrow Lin$^-$ cells from week 16 mice were plated and GFP expression in colonies was analyzed 12 days later. Each symbol is the average GFP$^+$ colony number for an individual mouse (left panels). Cells from all colonies were pooled and analyzed by flow cytometry (right panels).



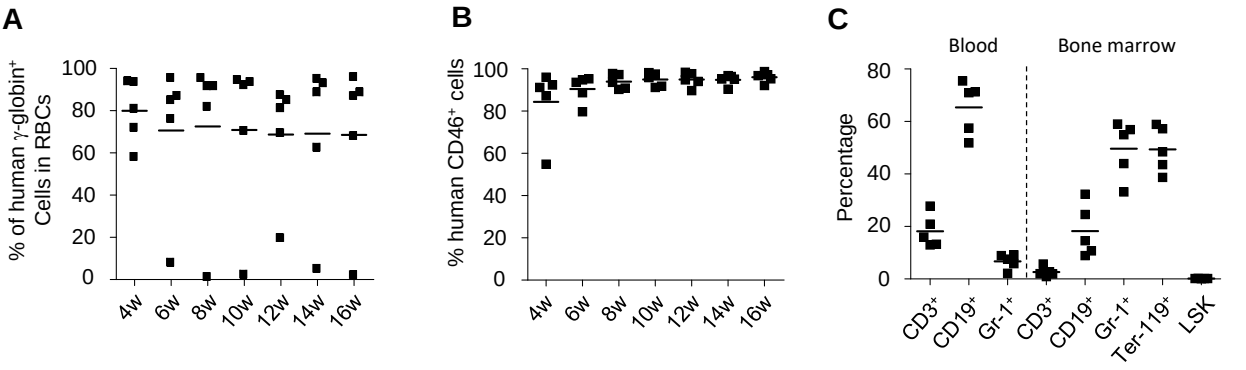
Supplementary Figure S4. Analysis of GFP marking in secondary recipients in the *ex vivo* transduction study (see Fig.3). HDAd-GFP-donor or HDAd-CRISPR + HDAd-GFP-donor transduced Lin$^-$ cells were harvested at week 16 after transplantation, depleted for lineage-positive cells, and transplanted into lethally irradiated C57Bl/6 mice. **A)** GFP expression in PBMCs in recipient mice measured by flow cytometry. N=3 for "Donor alone", N=6 for "CRISPR+Donor". $p<0.01$ **B)** Percentage of GFP-positive cells in PBMCs, spleen and bone marrow at week 16. $p<0.01$. **C)** GFP flow analysis of lineage-positive and LSK cells in recipients 16 weeks after transplantation. $p<0.01$. **D)** Analysis of GFP$^+$ colonies. Total bone marrow Lin$^-$ cells from week 16 mice were plated and GFP expression in colonies was analyzed 12 days later. Each symbol is the average GFP$^+$ colony number for an individual mouse (left panels). Cells from all colonies were pooled and analyzed by flow cytometry (right panels). **E)** Engraftment of transplanted cells based on human CD46 expression on PBMCs measured by flow cytometry. **F)** Percentage of lineage-positive and LSK cells in Donor alone and Donor+CRISPR groups. The difference between the two groups is not significant.



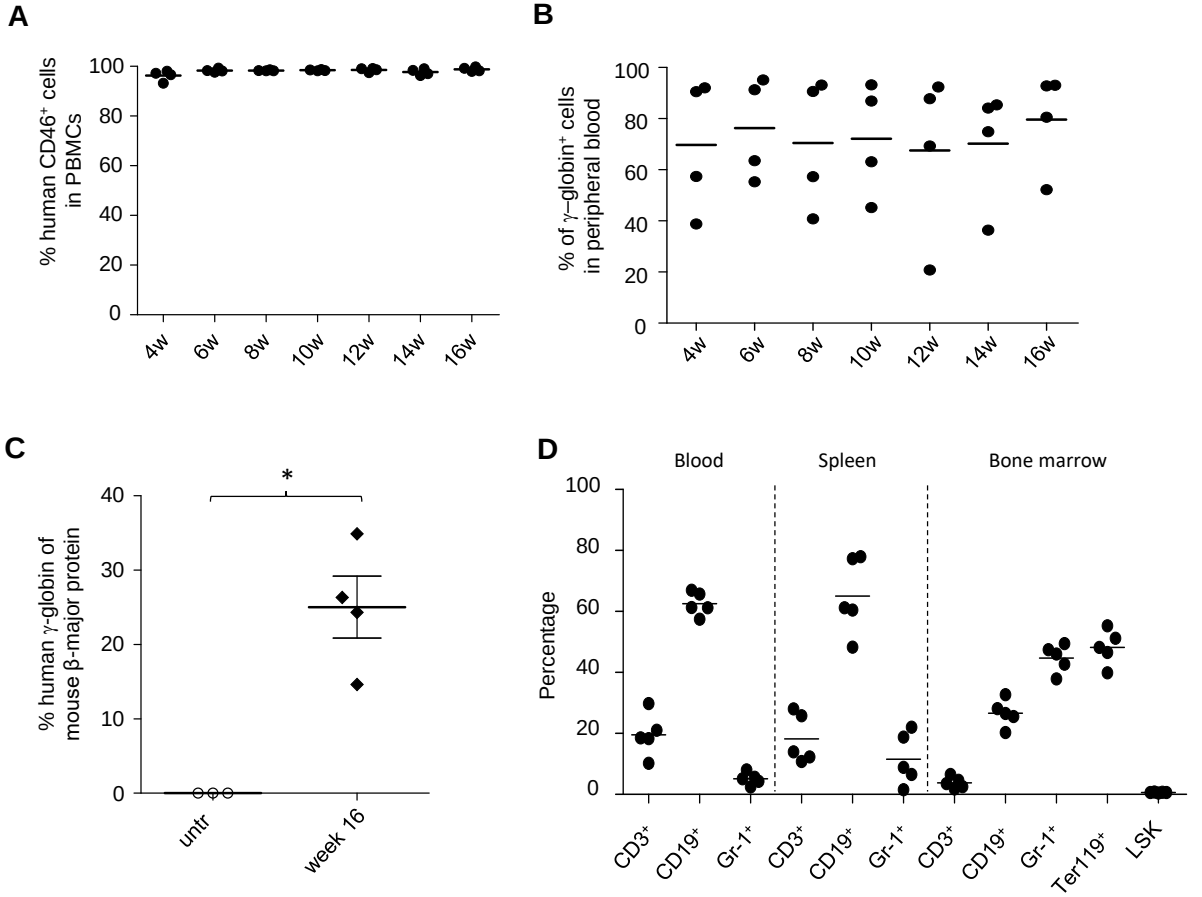
Supplementary Figure S5. Analysis of secondary recipients from Fig.4. At week 14, bone marrow Lin$^-$ cells from *in vivo* transduced AAVS1/hCD46tg mice were transplanted into lethally irradiated C57Bl/6 recipients. **A)** GFP-flow cytometry of PBMCs. N=6. **B)** GFP expression in mononuclear cells in blood, spleen and bone marrow. N=4. **C)** GFP flow analysis of lineage-positive and -negative cells in recipients 16 weeks after transplantation. N=4. **D)** Engraftment of transplanted cells based on human CD46 expression in PBMCs measured by flow cytometry. N=6. **E)** Percentage of lineage-positive and LSK cells at week 16.



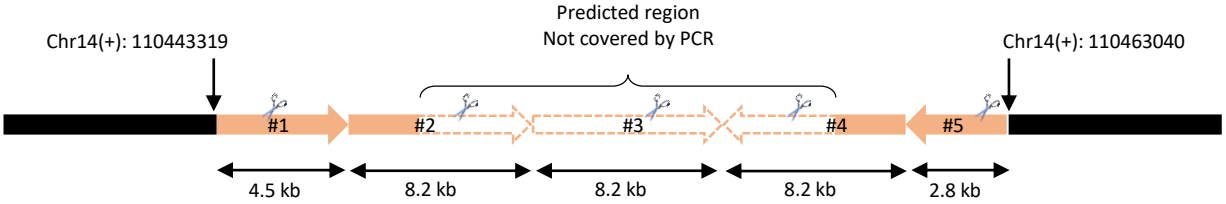
Supplementary Figure S6. Engraftment of AAVS1/CD46 Lin$^-$ cells transduced with HDAd-CRISPR and HDAd-globin-donor vectors. A) Engraftment of transplanted cells based on human CD46 expression on PBMCs measured by flow cytometry. N=5. **B)** Percentage of lineage-positive and LSK cells at week 16. Cells from blood, spleen, and bone marrow were analyzed. N=5.



Supplementary Figure S7. Analysis of secondary recipients from Fig. 5. Bone marrow cells from mice that were transplanted with HDAd-CRISPR + HDAd-globin-donor transduced Lin$^-$ cells were harvested at week 16 after transplantation, depleted for lineage-positive cells, and transplanted into lethally irradiated C57Bl/6 mice. N=5. **A)** γ-globin flow cytometry of RBCs in five recipient mice. **B)** Percentage of CD46-positive cells in lineage-positive PBMCs. **C)** Cellular composition in blood (PBMCs) and bone marrow at week 16 after transplantation into secondary recipients. N=5.



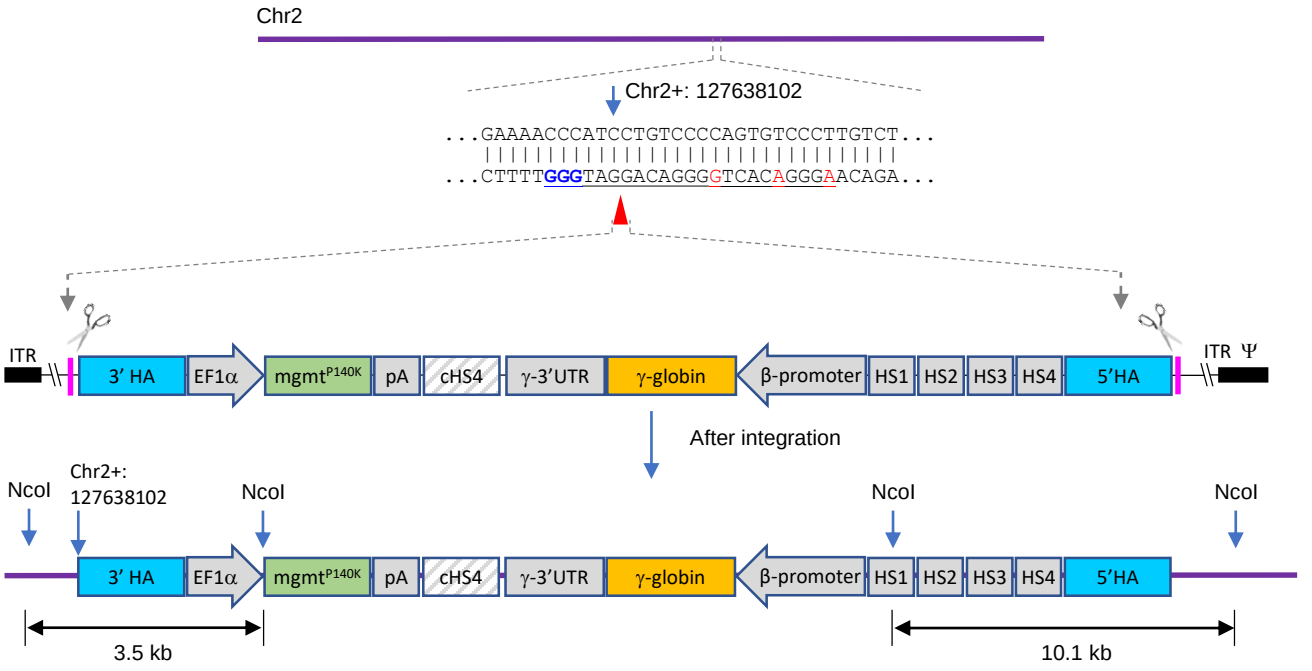
Supplementary Figure S8. Analysis of secondary recipients from Fig.6. At week 16, bone marrow Lin- cells from AAVS1/hCD46tg mice *in vivo* transduced with HDAd-CRISPR + HDAd-globin-donor were transplanted into lethally irradiated C57Bl/6 recipients. N=4. **A)** Engraftment of transplanted cells based on human CD46 expression on PBMCs measured by flow cytometry. **B)** γ-globin expression in RBCs. **C)** Percentage of γ-globin chains relative to mouse beta-major chains measured in RBCs of secondary recipients at week 16 by HPLC. * $p<0.05$. **D)** Lineage-positive and LSK cell composition in blood, spleen and bone marrow at week 16.



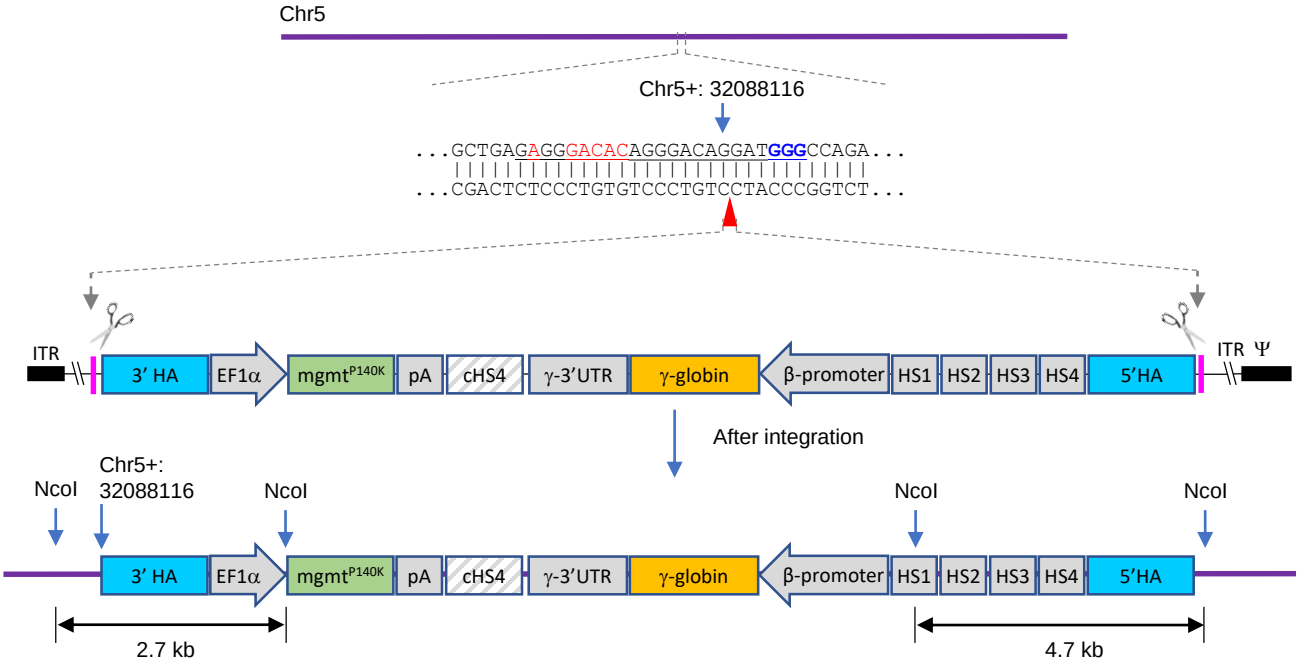
a) One single cut in repeat #1 to #4: preferred
b) One single cut in repeat #5: reduced preference due to incomplete left homology arm
c) Two cuts in two oppositely oriented repeats (e.g. #1 and #4):
no HDR-mediated targeted integration due to missing right homology arm
d) Two cuts in two repeats facing the same direction (e.g. #1 and #2):
preferred
e) For more than 2 cuts, only consider the one proximal to mouse gDNA sequence at each side
Apply rule c) or d) accordingly
f) Cuts in repeats #1 and 5 and deletion of the central region

In addition, if HDR-mediated targeted integration occurred in repeat #2 to #4, continuous cutting in flanking repeats, for example, #1 and #5, by CRISPR may result in loss of the already integrated transgene.

Supplementary Figure S9. Outcome depending on CRISPR/Cas9 cleavage of the multicopy AAVS1 locus present in AAVS1tg mice.



Supplementary Figure S10. Diagram showing CRISPR/Cas9-mediated HDAd-globin-donor integration into an off-target site on chromosome 2. This off-target integration was detected by iPCR in mouse #816. The off-target site Chr2(-): 127638097-127638119 bears 3bp mismatches with the AAVS1 guide sequence. The globin donor cassette (from 5'-HA to 3'-HA) released by AAVS1 CRISPR was reversely integrated into the CRISPR cleavage off-target site (Chr2: 127638102). The location of *NcoI* sites and iPCR amplicon size are indicated.



Supplementary Figure S11. Diagram showing CRISPR/Cas9-mediated HDAd-globin-donor integration into an off-target site on chromosome 5. Mice #857 and #945 displayed this off-target integration detected by iPCR. The off-target site Chr5(+): 32088100-32088122 bears 6 bp mismatches with the AAVS1 guide sequence at the PAM-distal side. The globin donor cassette (from 5'-HA to 3'-HA) released by AAVS1 CRISPR was reversely integrated into CRISPR cleavage site (Chr5: 32088116). The location of NcoI sites and iPCR amplicon size are indicated.

mouse #	GFP-donor *(ex vivo)*	γ-globin-donor *(ex vivo)*	GFP-donor *(in vivo)*	γ-globin-donor *(in vivo)*	targeted	off-target	complete HDAd
009	X					X	
023	X				X		
943	X				X		
944	X				X	X	
946	X				X		
147			X		X		
304			X		X		
467			X		X		X
321		X			X		X
322		X			X		
856		X			X		
858		X			X		
945		X			X	X	
504				X	X		
816				X	X	X	
869				X	X		
898				X	X		

Supplementary Figure S12. Summary of integration analyses by iPCR

Supplementary Methods

Generation of HDAd vectors. For the cloning of the HDAd-CRISPR vector, sgRNA (5'-GGGGCCACTAGGGACAGGAT-3') (1) targeting the human AAVS1 locus was synthesized, annealed and inserted into the *BbsI* site of pSPgRNA (Addgene, Cambridge, MA), generating pSP-sgAAVS1. A Cas9 coding sequence amplified from pLentiCRISPRv2 (Addgene), U6sgAAVS1 fragments released by *BamHI* digestion of pSP-sgAAVS1, and a previously described microRNA targeting region (miR-183/218) (2) were sequentially cloned into the *EcoRV-NotI*, *BamHI* and *NotI* sites of pBS-T-EF1α (2), forming pBST-sgAAVS1-miR. To obtain the recombinant adenoviral plasmids, an 8 kb cassette starting from the U6 promoter to the SV40 polyA signal sequence was amplified from pBST-sgAAVS1-miR and ligated with *NheI-XmaI* digested pHCA (3) by Gibson assembly (New England Biolabs), generating the corresponding pHCA-sgAAVS1-miR plasmid.

For the construction of the HDAd-GFP-donor vector, two 0.8kb homology arms (HA) immediately flanking the AAVS1 CRISPR cutting site were synthesized as gBlocks (IDT, San Jose, CA). One 23bp sgAAVS1 with PAM sequence was included upstream of the 5'HA and downstream of 3'HA, respectively, to mediate the release of the donor cassette. A EF1α-mgmt-2A-GFP-pA fragment was synthesized by Genscript (Nanjing, China), and ligated with the two 5'HAs by overlap PCR, forming sgAAVS1-5'HA-Ef1α-mgmt-2A-GFP-pA-3'HA-sgAAVS1 which was subsequently inserted into the *XmaI* site of pHCA (3), generating GFP donor vector pHCA-AAVS1-GFP-mgmt.

The cloning of the HDAd-globin-donor vector involved 3 steps. Step 1) The 11.8kb LCR-globin-mgmt cassette was released from pHM5-FR-IR-LCR-globin-mgmt (4) by *EcoRV-KpnI* digestion and ligated with a 2.8kb plasmid backbone amplified from pBS-Z (2), resulting in pBS-LCR-globin-mgmt. Two 1.8kb HAs immediately adjacent to the AAVS1 CRISPR cutting site were PCR amplified from genomic DNA isolated from bone marrow cells of AAVS1-tg mice using primers containing the 23bp sgAAVS1 with PAM sequence. The 5' and 3' side HAs were sequentially inserted into the *EcoRV* and *KpnI* sites, respectively, of pBS-LCR-globin-mgmt, generating pBS-AAVS1-globin-mgmt. Step 2) The nt1588-12121 region of pHCA was deleted by *EcoRI* digestion and self-ligation, generating pHCAS1. The original *PacI* site in pHCAS1 was destroyed by inserting two annealed oligo sequences. A new *PacI* cloning site was created at the *BstBI* site, getting pHCAS1-MCS. This cloning site was designed in such a way that two 15bp homologous regions will be exposed upon *PacI* digestion. The size of pHCAS1-MCS was further reduced by removing the 1.5kb *NheI* fragment, resulting in pHCAS1S-MCS. Step 3) Following *PacI* digestion of the two final constructs from the above two steps, the products were recombined by Gibson Assembly, generating the globin donor vector pHCA-AAVS1-globin-mgmt.

Next generation sequencing. For deep sequencing of insertion/deletions (indels), we PCR-amplified a ~250-bp region surrounding the predicted AAVS1 cleavage site and sequenced the products using an Illumina system. Genomic DNA was isolated as previously described (2). A 249 bp genomic region encompassing the AAVS1 target site was amplified using the following primers: AAVS1 forward, 5'-CGGTTAATGTGGCTCTGGTT-3'; reverse, 5'-CCTCTCTGGCTCCATCGTAA-3'. After cleaning-up the amplicon using AMPure XP Beads (Beckman Coulter, Indianapolis, IN), dA-tailing was performed using Klenow fragment. Illumina-compatible adaptors were ligated with the product by T4 ligase (New England Biolabs). A unique barcode sequence was introduced by PCR to allow sequencing multiple samples on the same sequencing run. Each step was followed by purification with AMPure XP Beads. The final libraries were quantified by Qubit (Invitrogen) and tested on an Agilent 2100 Bioanalyzer to determine average size of the amplicons. The amplicons were pooled at equal molarity and deep sequenced on an Illumina MiSeq system. Approximately 10^5 reads per amplicon were generated to adequately probe the types of mutations. Sequencing data were aligned to the AAVS1 reference sequence using the Cas-Analyzer online

tool (http://www.rgenome.net/cas-analyzer/#!) (5), a JavaScript-based implementation for NGS data analysis.

Flow cytometry: Cells were resuspended at 1×10^6 cells/100 μL in FACS buffer (PBS supplemented with 1 % heat-inactivated FBS) and incubated with FcR blocking reagent (Miltenyi Biotech, Auburn CA) for ten minutes on ice. Next the staining antibody solution was added in 100 μL per 10^6 cells and incubated on ice for 30 minutes in the dark. After incubation, cells were washed once in FACS buffer. For secondary staining the staining step was repeated with a secondary staining solution. After the wash, cells were resuspended in FACS buffer and analyzed using a LSRII flow cytometer (BD Biosciences, San Jose, CA). Debris was excluded using a forward scatter-area and sideward scatter-area gate. Single cells were then gated using a forward scatter-height and forward scatter-width gate. Flow cytometry data were then analyzed using FlowJo (version 10.0.8, FlowJo, LLC). For flow analysis of LSK cells, cells were stained with biotin-conjugated lineage detection cocktail (Miltenyi Biotec, San Diego, CA) and antibodies against c-Kit and Sca-1 as well as APC-conjugated streptavidin. Other antibodies from eBioscience (San Diego, CA) included anti-mouse LY-6A/E (Sca-1)-PE-Cyanine7 (clone D7), anti-mouse CD117 (c-Kit)-PE (Clone 2B8), anti-mouse CD3-APC (clone 17A2), anti-mouse CD19-PE-Cyanine7 (clone eBio1D3), and anti-mouse Ly-66 (Gr-1)-PE, (clone RB6-8C5). Other antibodies from Miltenyi Biotec included anti-human CD46-APC (clone: REA312). Anti-mouse Ter-119-APC (clone: Ter-119) was from Biolegend (San Diego, CA).
Intracellular staining of human γ-globin was performed using PE-conjugated anti-human γ-globin antibody from Santa Cruz (clone 51.7). The Fix & Perm cell permeabilization kit from Invitrogen was used according to manufacturer's instructions.

Inverse PCR*:* Junctions in total bone marrow cells, single colonies, HUDEP-2 cell mixture or clones were analyzed by inverse PCR as described elsewhere with modifications (6). Briefly, genomic DNA was isolated by incubating with genomic DNA lysis buffer (100 mM Tris-Cl (pH 8.0), 50 mM EDTA, 1% (w/v) SDS, and 400 μg/mL Proteinase K) at 55°C overnight with shaking, followed by phenol-chloroform extraction, precipitation with isopropanol, and wash with 70% ethanol. The DNA samples were dissolved in 10 mM Tris/HCL buffer (pH 8.5). 5 μg of DNA was digested with 30 U *NcoI* in 50 μL reaction at 37°C for 5 hours. After heat-inactivation and clean-up, the digested DNA was treated with 2.5 μL T4 ligase (New England Biolabs, M0202L) in 500 μL reaction buffer at 16°C overnight for intramolecular ligation. Following heat-inactivation and clean-up, the religated product was used for inverse PCR using KOD Hot Start DNA Polymerase. The following primers were used: EF1α forward, 5'-AACAAAAGCTGGTTAATTAAATCGGACGGGGGTAGTCTCAAG-3', and reverse, 5'-TATTGTACCATCTTAATTAAAGGGGCGAGTCCTTTTGTATG-3'; pA forward, 5'-AACAAAAGCTGGTTAATTAATCATTTTATGTTTCAGGTTCAGGGGG-3', and reverse, 5'-TATTGTACCATCTTAATTAAATGGTTACAAATAAAGCAATAGCATCAC-3'; HS4 forward, 5'-AACAAAAGCTGGTTAATTAAGTTTTTGTATTCTGTTTCGTGAGGCA-3'; and reverse, 5'-TATTGTACCATCTTAATTAAGCATTTGCCTAAGGTCGGACAT-3'. The underlined bases are *PacI* sites used for downstream cloning. The Ef1α and pA primer pairs were used for analyzing 5' and 3' junctions of GFP donor vector-treated samples, respectively. The HS4 and EF1α primer pairs were used for analyzing 5' and 3' junctions of globin donor vector-treated samples, respectively. PCR amplicons were gel purified, cloned, sequenced and aligned to identify the integration sites.
In-Out PCR: Genomic DNA was extracted as described in the section of Inverse PCR. 5 ng genomic DNA was directly used as template for In-Out PCR by KOD Hot Start DNA Polymerase in a 25 μl of reaction. The following PCR program was used: 94°C 2 min; 5 cycles of 98°C 10 sec, 66°C 30 sec and 68°C 1.5 min; 5 cycles of 98°C 10 sec, 63°C 30 sec and 68°C 1.5 min; 15 cycles of 98°C 10 sec, 60°C 30 sec and 68°C 1.5 min; 68°C 5min. Primers used are In-Out P1, 5'-CCACACCCAGACCTGACCCAAACC-3', In-Out P2, 5'-CGGGAACCACACACGGCACTTACC-3', and In-Out P3, 5'-TCTAACGCTGCCGTCTCTCTCCTG-3'. The products

were resolved in a 1% Agarose gel. One single 1.6 kb band indicates biallelic targeted integration; one 1.6 kb plus one 2.0 kb band indicates monoallelic targeted integration; one single 2.0 kb band indicates potential off-target integration.

References:

1. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-826.
2. Saydaminova, K., Ye, X., Wang, H., Richter, M., Ho, M., Chen, H., Xu, N., Kim, J.S., Papapetrou, E., Holmes, M.C. *et al.* (2015) Efficient genome editing in hematopoietic stem cells with helper-dependent Ad5/35 vectors expressing site-specific endonucleases under microRNA regulation. *Mol Ther Methods Clin Dev*, **1**, 14057.
3. Sandig, V., Youil, R., Bett, A.J., Franlin, L.L., Oshima, M., Maione, D., Wang, F., Metzker, M.L., Savino, R. and Caskey, C.T. (2000) Optimization of the helper-dependent adenovirus system for production and potency in vivo. *Proc Natl Acad Sci U S A*, **97**, 1002-1007.
4. Li, C., Psatha, N., Wang, H., Singh, M., Samal, H.B., Zhang, W., Ehrhardt, A., Izsvak, Z., Papayannopoulou, T. and Lieber, A. (2018) Integrating HDAd5/35++ Vectors as a New Platform for HSC Gene Therapy of Hemoglobinopathies. *Mol Ther Methods Clin Dev*, **9**, 142-152.
5. Park, J., Lim, K., Kim, J.S. and Bae, S. (2017) Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics*, **33**, 286-288.
6. Wang, H., Shayakhmetov, D.M., Leege, T., Harkey, M., Li, Q., Papayannopoulou, T., Stamatoyannopolous, G. and Lieber, A. (2005) A capsid-modified helper-dependent adenovirus vector containing the beta-globin locus control region displays a nonrandom integration pattern and allows stable, erythroid-specific gene expression. *J Virol*, **79**, 10999-11013.