Cell Stem Cell, Volume *29*

Supplemental Information

Choice of template delivery mitigates the genotoxic risk and adverse impact of editing in human hematopoietic stem cells

Samuele Ferrari, Aurelien Jacob, Daniela Cesana, Marianne Laugel, Stefano Beretta, Angelica Varesi, Giulia Unali, Anastasia Conti, Daniele Canarutto, Luisa Albano, Andrea Calabria, Valentina Vavassori, Carlo Cipriani, Maria Carmina Castiello, Simona Esposito, Chiara Brombin, Federica Cugnata, Oumeya Adjali, Eduard Ayuso, Ivan Merelli, Anna Villa, Raffaella Di Micco, Anna Kajaste-Rudnitski, Eugenio Montini, Magalie Penaud-Budloo, and Luigi Naldini

SUPPLEMENTAL INFORMATION

SUPPLEMENTAL FIGURES

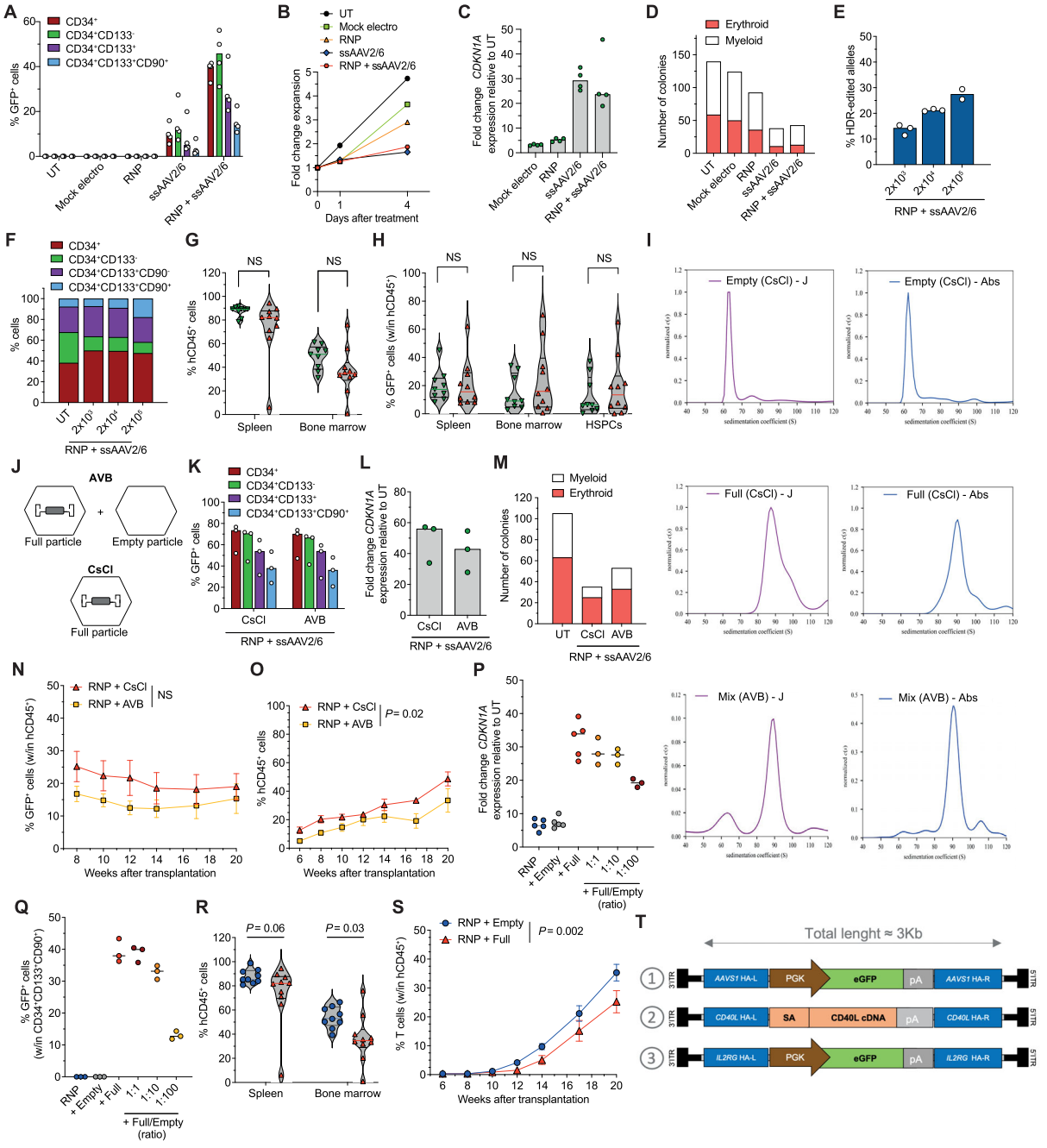

Figure S1. Related to Figure 1.

Extent of p53 activation correlates with AAV dose. A-D) Percentage of GFP$^+$ cells within mobilized peripheral blood (mPB) HSPC subpopulations (**A**), fold cell expansion (**B**), fold change expression of *CDKN1A* relative untreated (UT) cells at day 1 (**C**) and number of colonies grown from HSPCs (**D**) upon indicated treatments (n = 4). Median. **E-F)** Percentage of HDR-edited alleles measured by 3' targeted integration (TI) ddPCR (**E**, n = 3, 3, 2; median) and culture composition (**F**, n = 1, 3, 3, 3; mean) in CB HSPCs from 1B. **G-H)** Percentage of hCD45$^+$ (**G**) and GFP$^+$ cells within the human graft (**H**) at the end of the experiments within spleen and bone marrow (BM) of

transplanted mice from CB experiments in 1E (n = 9, 10). Median with quartiles, LME. **I)** Sedimentation distribution plots from analytical ultracentrifugation (AUC) representing sedimentation coefficients "S" with interference fringes "J" (left) or absorbance "Abs" (right) for ssAAV2/6 fractionated by CsCl gradient with "empty" (top) and "full" (middle) particles or AVB purified (bottom). **J)** Illustration of the ssAAV2/6 vectors used for experiments in Figure 1G-K and S1K-S. **K-M)** Percentage of GFP^+ cells within CB HSPC subpopulations (**K**), fold change expression of *CDKN1A* relative UT at day 1 (**L**), number of colonies grown from HSPCs (**M**) after *AAVS1* editing with AAV purified by CsCl gradient or AVB affinity chromatography (n = 3). Median. **N-O)** Percentage of circulating $hCD45^+$ (**N**) and GFP^+ cells within human graft (**O**) in mice transplanted with the outgrown progeny of starting-matched limiting cell doses of CB HSPCs edited as indicated (n = 5). Mean ± s.e.m., LME followed by post-hoc analysis; results are shown for the last timepoint. **P-Q)** Fold change expression of *CDKN1A* relative to UT at 1 d after editing (**P;** n = 5, 5, 5, 3, 3, 3) and percentage of GFP^+ cells (**Q**; n = 3) within CB HSPCs in presence of "full" ($5x10^5$ viral capsids/cell, equivalent to $2x10^4$ vg/cell), "empty" ($5x10^7$ viral capsids/cell) or "full+empty" (admixed at the indicated viral capsid ratios) AAV particles (n = 5, 5, 5, 3, 3, 3). Adding increasing amounts of empty particles to full ones progressively decreased p21 induction and editing efficiency, likely due to competition for cellular transduction. Median. **R)** Percentage of $hCD45^+$ cells within organs of transplanted mice from 1J (n = 9, 10). Median with quartiles, LME. **S)** Percentage of circulating $hCD3^+$ cells in transplanted mice from 1J (n = 9, 10). Mean ± s.e.m. LME followed by post-hoc analysis; results are shown for the last timepoint. **T)** Schematic representation of the ssAAV2/6 constructs targeting *AAVS1* (top), *CD40L* (middle) or *IL2RG* (bottom) used in 1L. HA: homology arm. pA: polyadenylation signal.

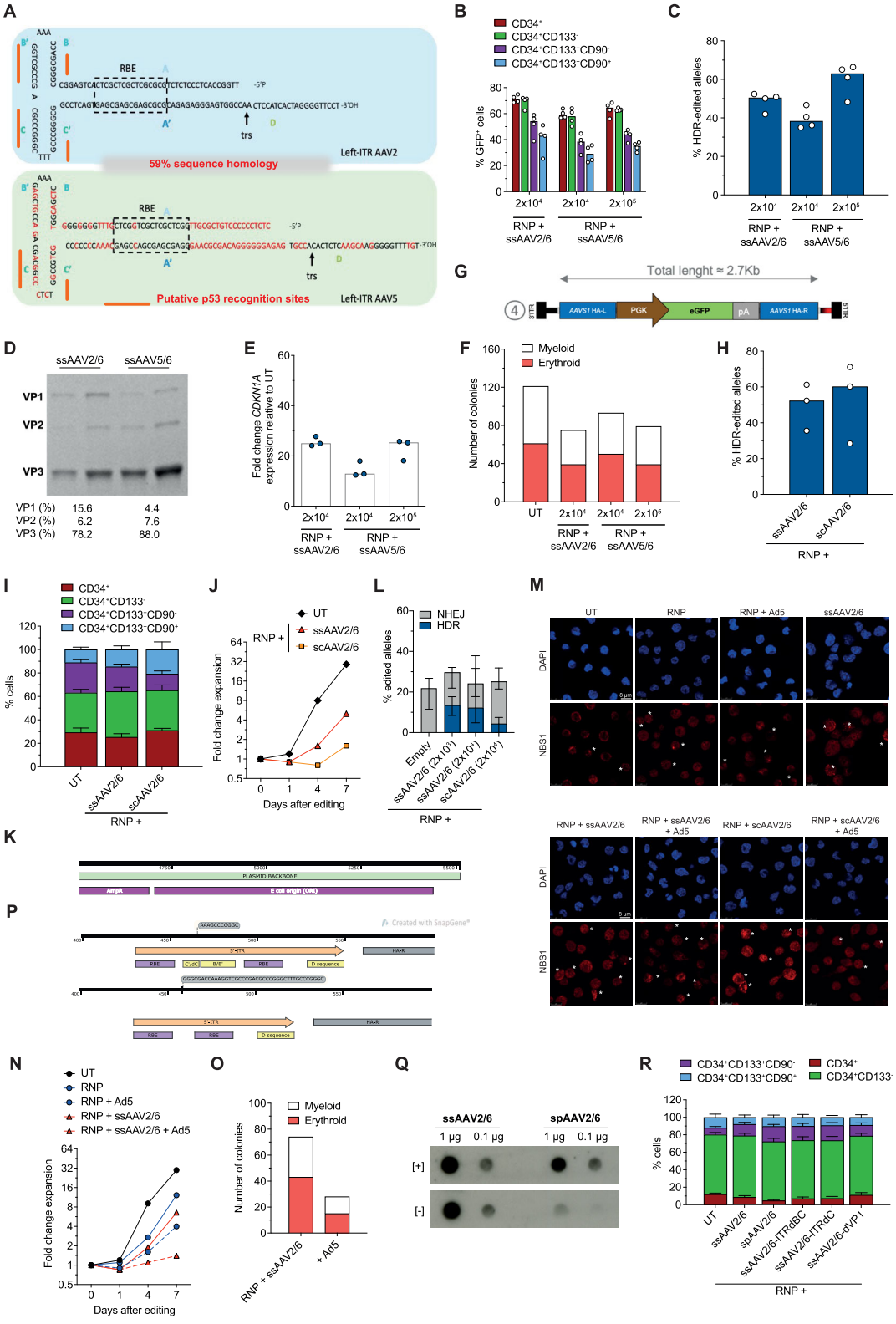

Figure S2. Related to Figure 2.

AAV ITRs engineering does not mitigate MRN-mediated p53 activation.

A) Sequences of the 5' AAV2 (blue box) and AAV5 (green box) ITRs. The different portions are indicated: A-A', B-B', C-C' and D sequences, the Rep Binding Element (RBE) and the terminal

resolution site (trs). Divergent nucleotides are in red. Putative p53 binding sites identified by the Alggen-Promo web tool using dissimilarity margin ≤ 10% (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3) are underlined in orange. **B-D)** Percentage of GFP^+ cells within CB HSPC subpopulations (**B**) and of HDR-edited alleles in bulk HSPCs (**C**) after editing with either ssAAV2/6 or ssAAV5/6 at indicated doses (n = 4). Median. ssAAV5/6 allowed lower HDR editing than ssAAV2/6 at matching dose due to the decreased particle infectivity resulting from lower content of AAV Cap viral protein 1 (VP1), as shown from the silver stained SDS PAGE (**D**) revealing AAV Cap proteins (VP1, VP2 and VP3) loaded with purified $2x10^{12}$ vg or $4x10^{12}$ vg of ssAAV2/6 and ssAAV5/6. Quantifications are shown. **E-F)** Fold change expression of *CDKN1A* relative to UT at day 1 (**E**) and number of colonies (**F**) grown from CB HSPCs from experiments in S2B-C (n = 3). Median. **G)** Schematic representation of the scAAV2/6 construct. **H-J)** Percentage of HDR-edited alleles (**H**; n = 3, median), HSPC culture composition (**I**; n = 6, mean ± s.e.m.) and fold change expansion over time (**J**; n = 3, median) of HSPCs from experiments in 2A. **K)** Schematics of the ddPCR probe amplifying the origin of replication sequence ("ORI") of the transfer plasmids for ss- and sc-AAV2/6 production. **L)** Percentage of HDR- and NHEJ-edited alleles, quantified by ddPCR and *AAVS1* deep sequencing, respectively, in human splenocytes of mice in 1E, 1J and 2F (n = 9, 9, 10, 5). **M)** Representative images from 2I (day 1). **N)** Fold change expansion of CB HSPCs over time after *IL2RG* editing with indicated treatments (n = 1, 1, 1, 2, 2). Median. **O)** Number of colonies grown from CB HSPCs in S2N (n = 3, 3). Median. **P)** Schematic representation of the engineered 5' AAV ITR (dC and dBC) sequences. The same deletions were applied to the respective 3' AAV ITR. AAVs were used at matched doses. **Q)** ssAAV2/6 and spAAV2/6 genomes characterization by dot blot for the detection of positive and negative viral genome strands. **R)** HSPC culture composition from 2N (n = 3). Mean.

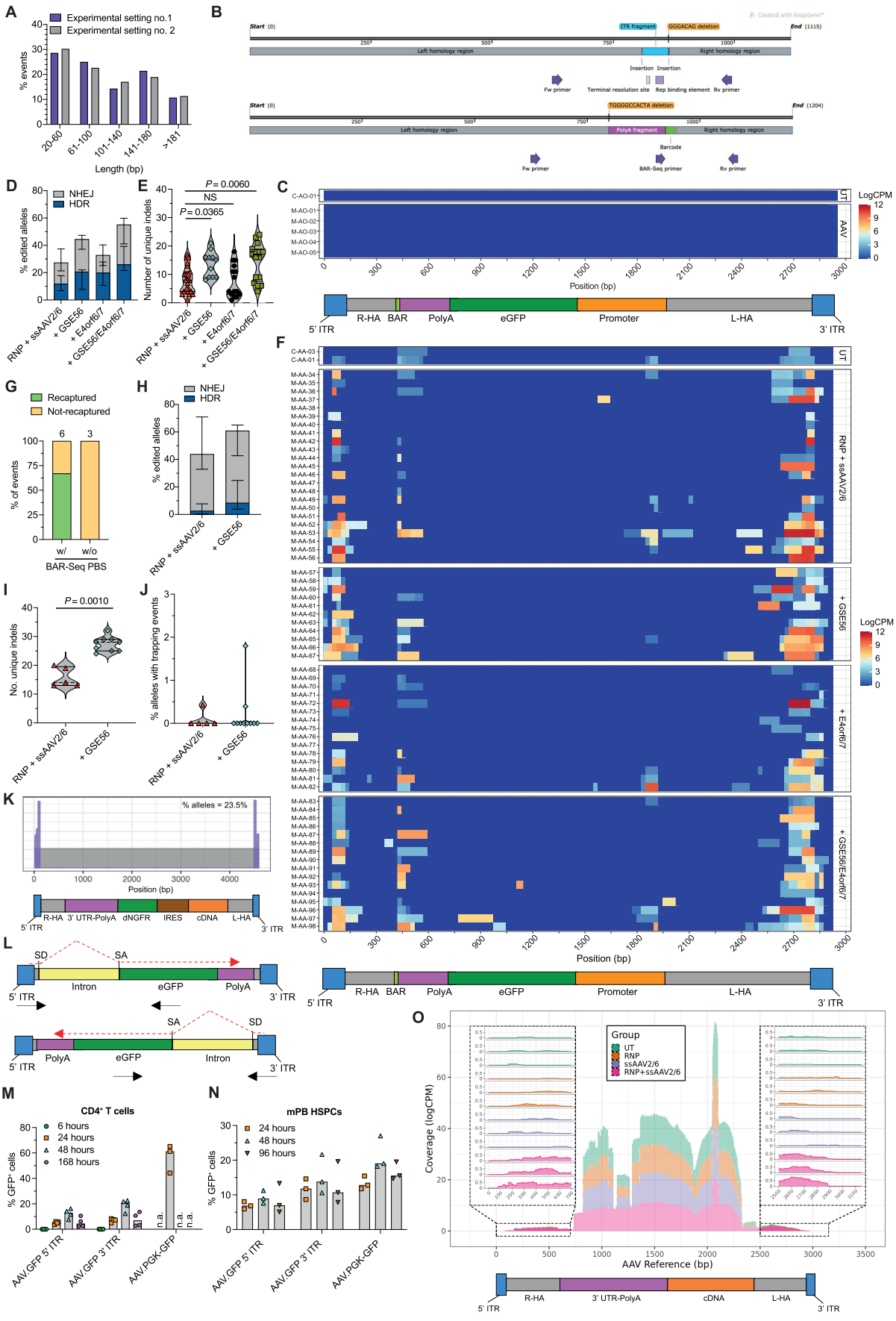

Figure S3. Related to Figure 3.

AAV2 ITRs harbor transcriptional activity in human hematopoietic cells. A) Fraction of DNA trapping events categorized by length ranges in the experimental setting of 3A (no. 1, n = 28 events) and S3F (no. 2, n = 53 events). **B)** Allele structure for two representative trapping events (top: ITR fragment trapping with indels; bottom: aborted HDR integrating the molecular barcode). Fw: forward primer. Rv: reverse primer; BAR-Seq primer: primer used for barcode (BAR) retrieval in previous analyses[14]. ITR integrations often comprised the D sequence, the trs, the RBE and a portion of the T-shaped palindromic sequences. **C)** Heatmap as in 3A for mice transplanted with HSPCs only transduced with the *AAVS1* ssAAV2/6 (n = 5). One UT sample is shown. **D-E)** Percentage of HDR- and NHEJ-edited alleles (**D,** n = 23, 10, 14, 16; median with 95% CI) and number of unique indels (**E**, n = 23, 11, 15, 16; median with quartiles, Kruskal-Wallis with Dunn's multiple comparisons) in human splenocytes of mice from previously published experiments[14]. The frequency and number of unique indels-bearing alleles in the human graft were higher in groups treated with the p53 inhibitor GSE56, confirming the previously observed increased clonal composition of the edited graft upon transient p53 inhibition[14]. **F)** Heatmap as in 3A for mice from 3E (n = 23, 11, 15, 16). **G)** Percentage of BARs retrieved from Figure S3F and recaptured (fragment length ≥ 45 bp, thus including the BAR-Seq primer binding site) or not in the BAR-Seq dataset[14] when analyzing the same samples. The number of events is shown. **H-I)** Percentage of HDR- and NHEJ-edited alleles, quantified by ddPCR and deep sequencing, respectively, (**H,** median with 95% CI) and number of unique indels (**I;** median with quartiles, Mann-Whitney test.) in human splenocytes of mice from previously published experiments[13] (n = 5, 9). **J)** Percentage of alleles in human splenocytes from S3I carrying integrated DNA fragments as in 3B (n = 5, 9). Median with quartiles. **K)** Alignment on the AAV genome of trapped DNA fragments in human $CD4^+$ T cells from one xenograft from previously published experiments[13]. **L)** Schematics of AAV constructs used to address putative AAV2 ITRs transcriptional activity in human hematopoietic cells. SD: artificial splice donor site; SA: artificial splice acceptor site. Dashed red arrows indicate putative transcripts. Black arrows indicate primer binding sites used to amplify the spliced transcript (expected length = 500 bp). **M-N)** Percentage of GFP^+ cells over time within $CD4^+$ T cells (**M,** n = 4, 4, 4, 3) or mPB HSPCs (**N,** n = 3) from healthy-donor cells transduced with the promoter-less AAV depicted in Figure S3L. Cells transduced with the PGK-GFP ssAAV2/6 are shown as reference. Median. **O)** Coverage of the AAV genome by sequencing reads obtained from an RNA-Seq experiment performed on three independent $CD4^+$ T cell donors treated as indicated (n = 3). A promoter-less ssAAV2/6 containing the *CD40LG* coding sequence was used in this experiment.

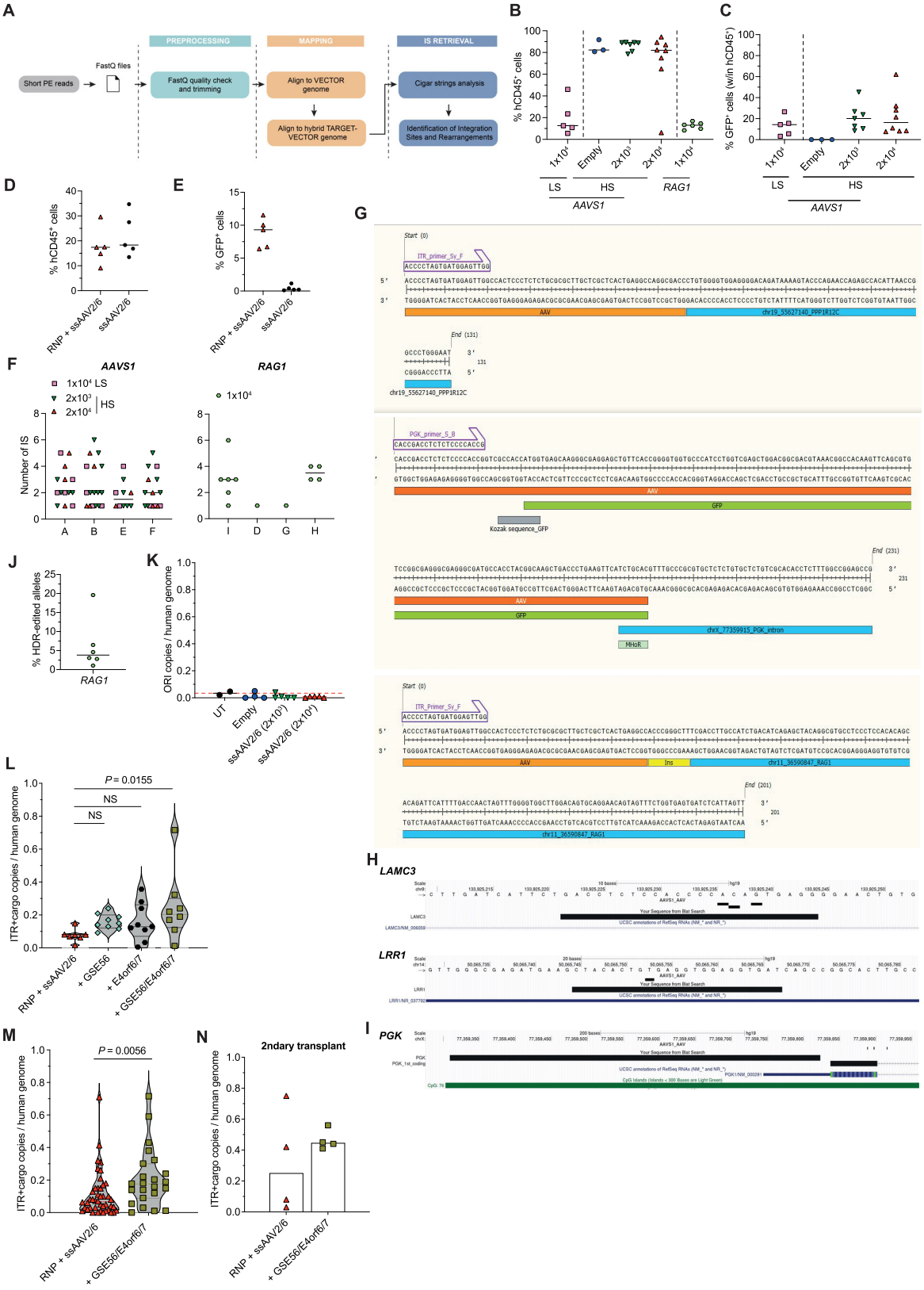

Figure S4. Related to Figure 4.

Characterization and quantification of AAV IS in edited LT-HSCps. A) Schematic workflow of the RAAVIoli bioinformatic pipeline. **B)** Percentage of hCD45$^+$ cells in the hematopoietic organs of mice transplanted with CB HSPCs edited as indicated at *AAVS1* (n = 5, 3, 7, 8) or *RAG1* (n = 8).

Median. Note that mice from the HS groups belong to S1G, R. **C)** Percentage of GFP^+ cells in *AAVS1* experiments from S4B (n = 5, 3, 7, 8). Median. **D-E)** Percentage of $hCD45^+$ cells (**D**) and GFP^+ cells within the human graft (E) in the hematopoietic organs of mice transplanted with mPB HSPCs treated as indicated (n = 5). Median. **F)** Number of AAV IS retrieved by the PCR systems shown in 4A in mice transplanted with *AAVS1*- (left) or *RAG1*- (right) edited HSPCs (n = 14, 17, 10, 15). Median. **G)** Examples of chimeric reads including AAV IS. The read portion aligning to the AAV genome is indicated in orange, while that aligning on the human genome is indicated in blue. AAV IS reads having a precise homology breakpoint between the vector and the host chromosomal sequences are referred as "exact breakpoint" (top); AAV IS reads with a micro-homology sequence between the vector and the host chromosomal sequences are referred as "Micro-Homology Region" (MHoR, light green rectangle); AAV IS reads with random nucleotide insertion between the vector and the host chromosomal sequences are referred as "insertion breakpoint" (Ins, yellow rectangle). **H-I)** Genomic views as in 4D of AAV IS within two predicted off-target sites of the LS gRNA (**H**), *LAMC3* (top) and *LRR1* (bottom), and within the *PGK1* gene (**I**). Of note, these latter IS were amplified using the PCR set whose primers anneal to a vector-specific sequence at the 3' end of the *PGK* promoter and the amplified downstream cellular sequences were not comprised within the vector, making unlikely that this sequence originated from a technical artifact in their retrieval (see also S4G). Genomic coordinates and scale are indicated. Black rectangles indicate AAV IS, black horizontal bars indicate the off targeted protospacer sequence targeted by the gRNA and the *PGK* promoter sequence comprised within the AAV vector. **J)** Percentage of HDR-edited alleles in the BM of mice transplanted with human BM HSPCs edited at *RAG1* (n = 6). Median. **K)** "ORI" CG within human splenocytes of mice from 1E, 1J, 2F and UT samples (n = 2, 4, 5, 5). Median. The dashed red line indicates the background noise threshold set by the median of UT samples. Only background noise signal was detected in the human long-term grafts. **L-N)** "ITR+cargo" CG within human splenocytes of mice from previously published mPB HSPC experiments[14] (**L**, n = 7, 9, 9, 8; median with quartiles, Kruskal-Wallis with Dunn's multiple comparisons), human splenocytes of mice from CB HSPC experiments (**M**, n = 43, 23; median with quartiles, Mann-Whitney test), BM-purified $hCD45^+CD34^+$ HSPCs in serially transplanted mice from previously published experiments[14] (**N**, n = 4; median).

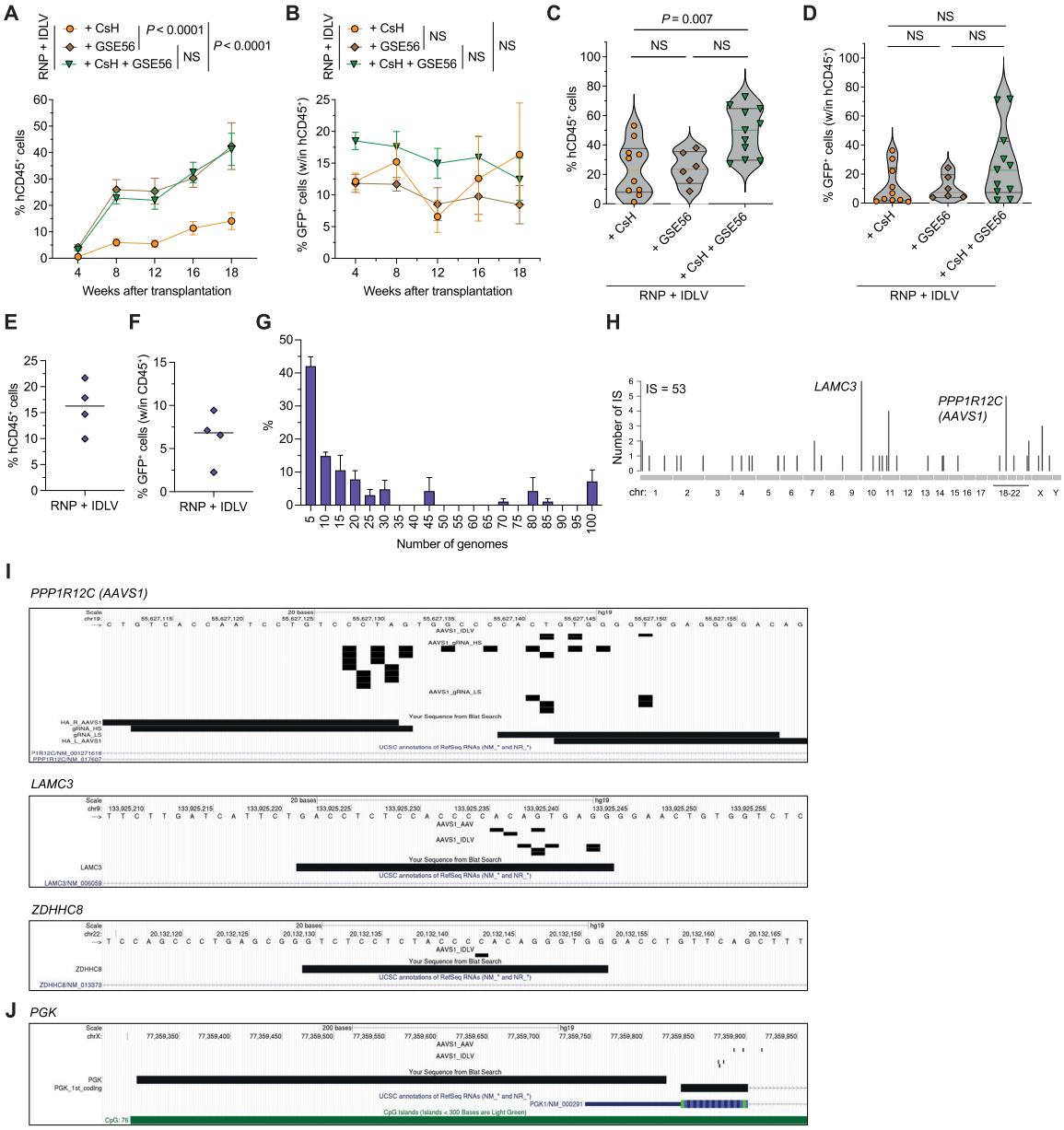

Figure S5. Related to Figure 5.

Retrieval of IDLV IS in the long-term xenografts. A-B) Percentage of circulating hCD45$^+$ (**A**) and GFP$^+$ cells within the human graft (**B**) of mice transplanted with the outgrown progeny of starting-matched limiting cell doses of CB HSPCs edited as indicated (n = 10, 6, 11). Mean ± s.e.m., LME followed by post-hoc analysis; results are shown for the last timepoint. **C-D)** Percentage of hCD45$^+$ (**C**) and GFP$^+$ cells within human BM cells (**D**) of transplanted mice from S5A (n = 10, 6, 11). Mean ± s.e.m., LME. Statistics are shown for the last timepoint. **E-F)** Percentage of hCD45$^+$ (**E**) and GFP$^+$ cells within human BM cells (**F**) of mice transplanted with CB HSPCs edited with the IDLV-based protocol and the LS *AAVS1* gRNA (n = 4). Median. **G)** Percentage of IDLV IS represented by the indicated number of genomes from mice in S5E (n = 4). Mean ± s.e.m. **H)** Genome-wide distribution of IDLV integrations retrieved from mice in S5E. **I-J)** Genomic views as in 4D of IDLV IS within

AAVS1 (top) and two predicted off-target sites, *LAMC3* (middle) and *ZDHH8* (bottom) (**I**) and within the *PGK1* gene (**J**). For comparison AAV IS in the same loci were also reported.

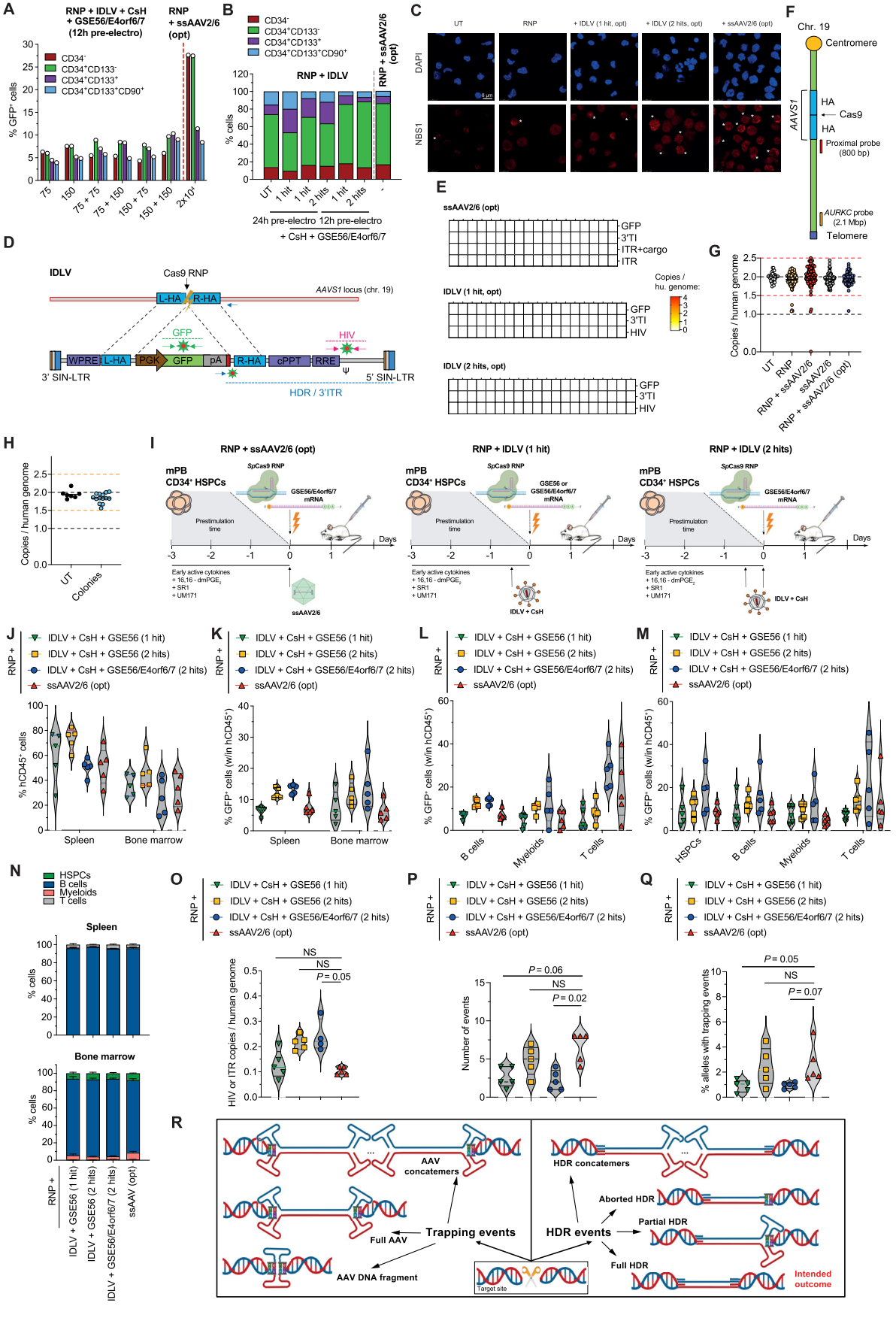

Figure S6. Related to Figure 6.

Improved editing efficiency and mitigated cyto- and geno-toxicity by optimized IDLV template delivery. A) Percentage of GFP^+ cells within mPB HSPC subpopulations after editing with different IDLV doses for each transduction hit (n = 1). **B)** HSPC culture composition in experiments from 6A (n = 5, 1, 2, 2, 5, 4, 5). Mean. No differences in the proportion of most primitive HSPCs were found after editing among the optimized IDLV or AAV protocols. **C)** Representative images from 6E (day 7). **D)** Panel of ddPCR probes tiling the *AAVS1* IDLV genome ("HIV", "GFP") and mapping on the 3'*AAVS1* vector-genome junction ("3'TI" or "3'HDR"). **E)** Heatmap as in 6G for colonies plated from bulk HSPCs transduced as indicated (n = 18, 19, 20). **F)** Schematic representation of ddPCR probes mapping nearby *AAVS1* and at the *AURKC* locus. **G)** CG as in Figure 6I within colonies grown from HSPCs treated as indicated (n = 39, 95, 95, 91, 69). Mean. **H)** CG measured by ddPCR probes mapping on *AURKC* (n = 7, 14) in "red" colonies from 6I and UT. Mean. Orange lines indicate cut-off values for colonies carrying long-range deletions or duplications. **I)** Experimental workflows of editing procedures used in 6J. **J-K)** Percentage of $hCD45^+$ cells (**J**) and GFP^+ cells within $hCD45^+$ cells (**K**) in organs of transplanted mice from 6J (n = 5). Median with quartiles. **L-M)** Percentage of GFP^+ cells within total $hCD45^+$ cells (**L**) and hematopoietic lineages (**M**) of transplanted mice from 6J (n = 5). Median with quartiles. **N**) Percentage of cells for each hematopoietic lineage within $hCD45^+$ cells from spleen (top) and bone marrow (bottom) of transplanted mice from 6J (n = 5). Mean ± s.e.m. **O)** CG measured by HIV or ITR probes within human splenocytes from mice in S6J (n = 5, 5, 4, 5). Median with quartiles. Kruskal-Wallis test with Dunn's multiple comparisons. **P-Q)** Number (**P**) and percentage (**Q**) of *AAVS1* alleles in human splenocytes carrying integrated fragments (length ≥ 20 bp) (n = 5). Median with quartiles. Kruskal-Wallis test with Dunn's multiple comparisons. **R)** Graphical schematic representation of HDR and NHEJ- or MMEJ-mediated integration at nuclease-induced DNA DSBs.

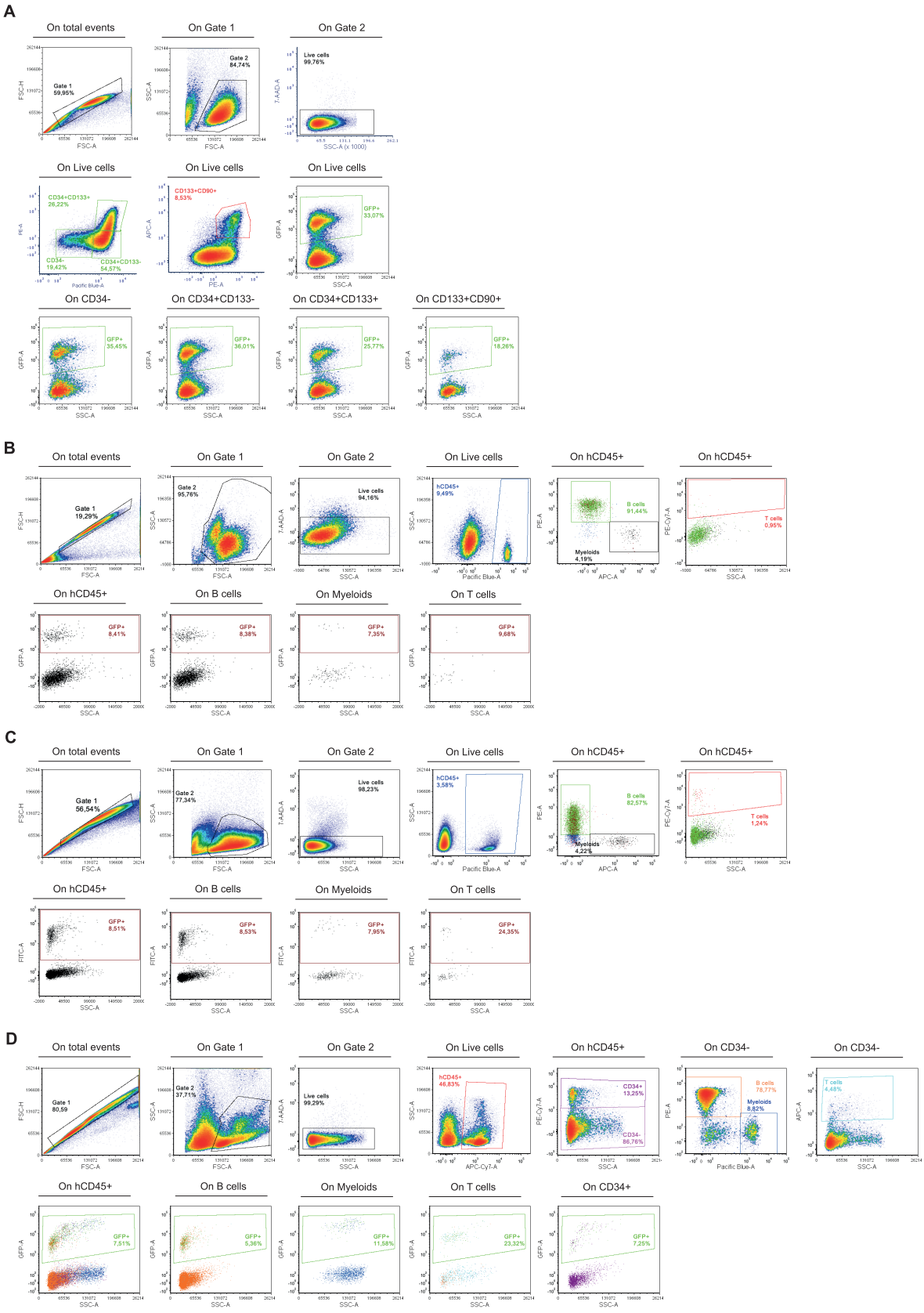

Figure S7. Related to STAR Methods.

Gating strategies for flow cytometry analyses. Gating strategies for the analysis of **A)** HSPC phenotype and **B-D)** circulating human peripheral blood mononuclear cells (**B**), splenocytes (**C**) and BM cells (**D**) of xenotransplanted mice.

SUPPLEMENTAL TABLES

		Interference (J)					OD 260 nm				Ratio (OD260/J)	Expected values
Purification method		*from*	*to*	*weighted s-value*	*area*	*%*	*from*	*to*	*mean*	*area*		
CsCl	**ssAAV2/6 empty**											
	empty	58.7	69.6	63.2	0.1	**79.3**	57.6	69.3	62.8	0.06	0.7	*[0.3-0.7]*
	intermediate	69.4	84.1	76.0	0.01	**12.7**	69.2	87.5	77.3	0.01	0.8	*[0.8-1.2]*
	full	84.1	104.8	93.3	0.01	**7.9**	90.5	106.0	98.0	0.01	0.5	*[1.7-2.9]*
	ssAAV2/6 full											
	empty	-	-	-	-	-	-	-	-	-	-	-
	intermediate	-	-	-	-	-	-	-	-	-	-	-
	full	76.2	109.2	90.8	0.032	**100**	73.827	106.948	90.665	0.06	2.0	*[1.7-2.9]*
AVB	**ssAAV2/6**											
	empty	40.3	74.9	60.1	0.01	**23.2**	47.769	79.794	68.862	0.01	0.8	*[0.3-0.7]*
	full	74.3	101.9	88.7	0.03	**69.8**	80.281	100.008	90.542	0.06	2.1	*[1.7-2.9]*
	non-determined	102.1	119.8	111.9	0.003	**7.0**	99.886	119.612	107.544	0.01	3.3	

Table S1. Related to Figure 1 and STAR Methods. Raw data of the analytical ultracentrifugation (AUC) analysis.