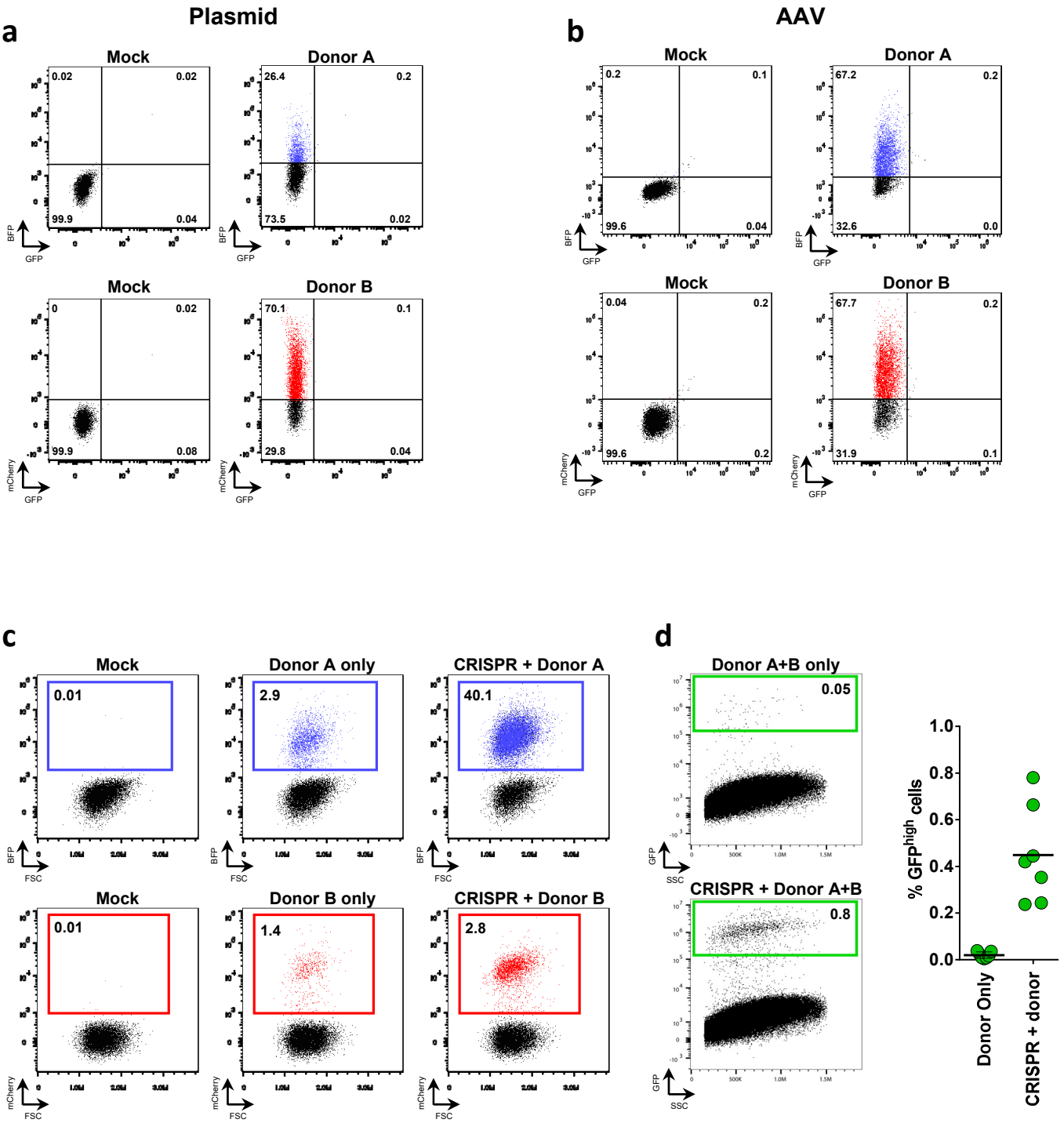

Figure S1. Neither of the two donors of the split *GFP* system express GFP on their own, only Donor A works as a homologous *CCR5* donor, and low levels of *GFP* reconstitution when using plasmid donors. Related to Figure 2. Donor A and Donor B of the split *GFP* system depicted in Figure 2a were delivered separately to K562 cells either by **(a)** plasmid electroporation or **(b)** AAV6 transduction. Representative FACS plots show BFP, mCherry, and GFP fluorescence as measured by flow cytometry four days after delivery and data show that neither donor alone expresses GFP. **(c)** Donor A or Donor B plasmids were electroporated into K562 cells with or without Cas9 mRNA and *CCR5* sgRNA (CRISPR). Representative FACS plots show BFP and mCherry expression measured by flow cytometry 16 days after electroporation when episomal plasmid DNA was diluted out. Data show that only Donor A can serve as donor template for homologous recombination at *CCR5* (40.1% of cells stably expressing BFP) while low integration rates of donor B was observed (2.8% of cells stably expressing mCherry), which are consistent with rates observed for random integration in K562 cells. **(d)** K562 cells were electroporated with Donor A and Donor B plasmids with or without Cas9 mRNA and *CCR5* sgRNA (CRISPR). Representative FACS plots show GFP expression measured 14 days after electroporation by flow cytometry with the targeted population gated as the GFP^{high} population (see also Figure S2a). The graph shows data from different biological replicates, bars represent means, N = 7.

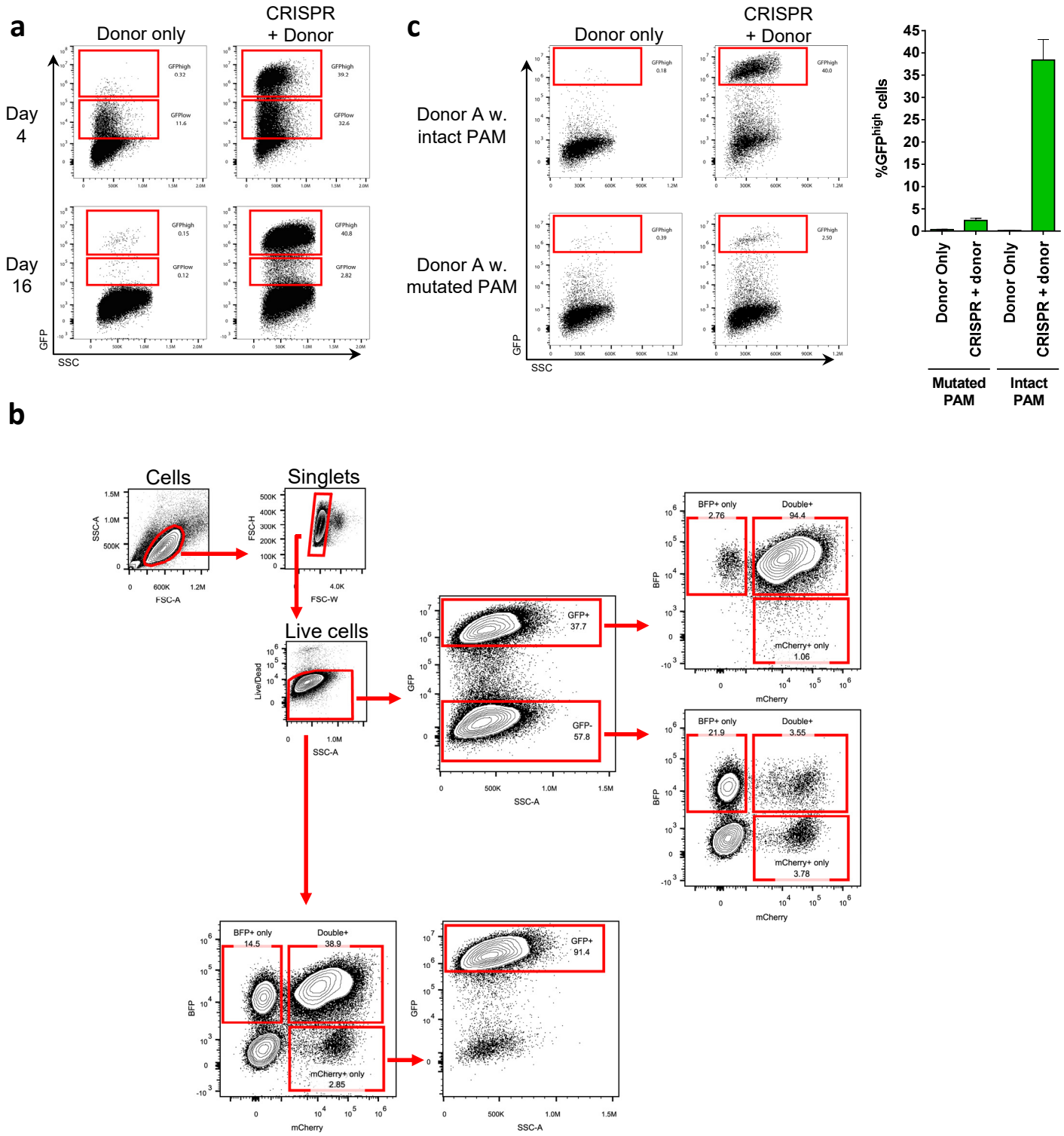

Figure S2. Targeting with the split *GFP* donor pair leads to transient episomal GFP expression and a stable GFP^{high} population, which is also mCherry^{+} and BFP^{+}. Related to Figure 2. (a) K562 cells were mock-electroporated or electroporated with Cas9 mRNA and *CCR5*-targeting sgRNA (CRISPR) and then transduced with the split *GFP* AAV6 donors. GFP levels were measured by flow cytometry 4 and 16 days after electroporation. **(b)** K562 cells were treated as in (a) and analyzed by flow cytometry 16 days after electroporation. Representative FACS plots are from one of the samples shown in Figure 2c. Almost all GFP^{+} cells (94.4%) are also double-positive for BFP and mCherry expression. Analogously, out of all cells double-positive for BFP and mCherry, 91.4% also express GFP. This supports the intended design, that reconstitution of the *GFP* cassette requires targeting of both Donor A (BFP) and Donor B (mCherry). Among all cells targeted with Donor A (all BFP^{+} cells: 14.5% + 38.9% = 53.4%), approximately 27% do not get targeted by donor B (14.5% / 53.4%). **(c)** K562 cells were treated as in (a) using a Donor A wither either an intact PAM (NGG) or a mutated PAM (NTA). Representative FACS plots are shown from flow cytometric analysis 8 days after electroporation (*left panel*), and data from independent replicate experiments are shown in the bar graph (*right panel*), columns represent mean +SD, N = 3.

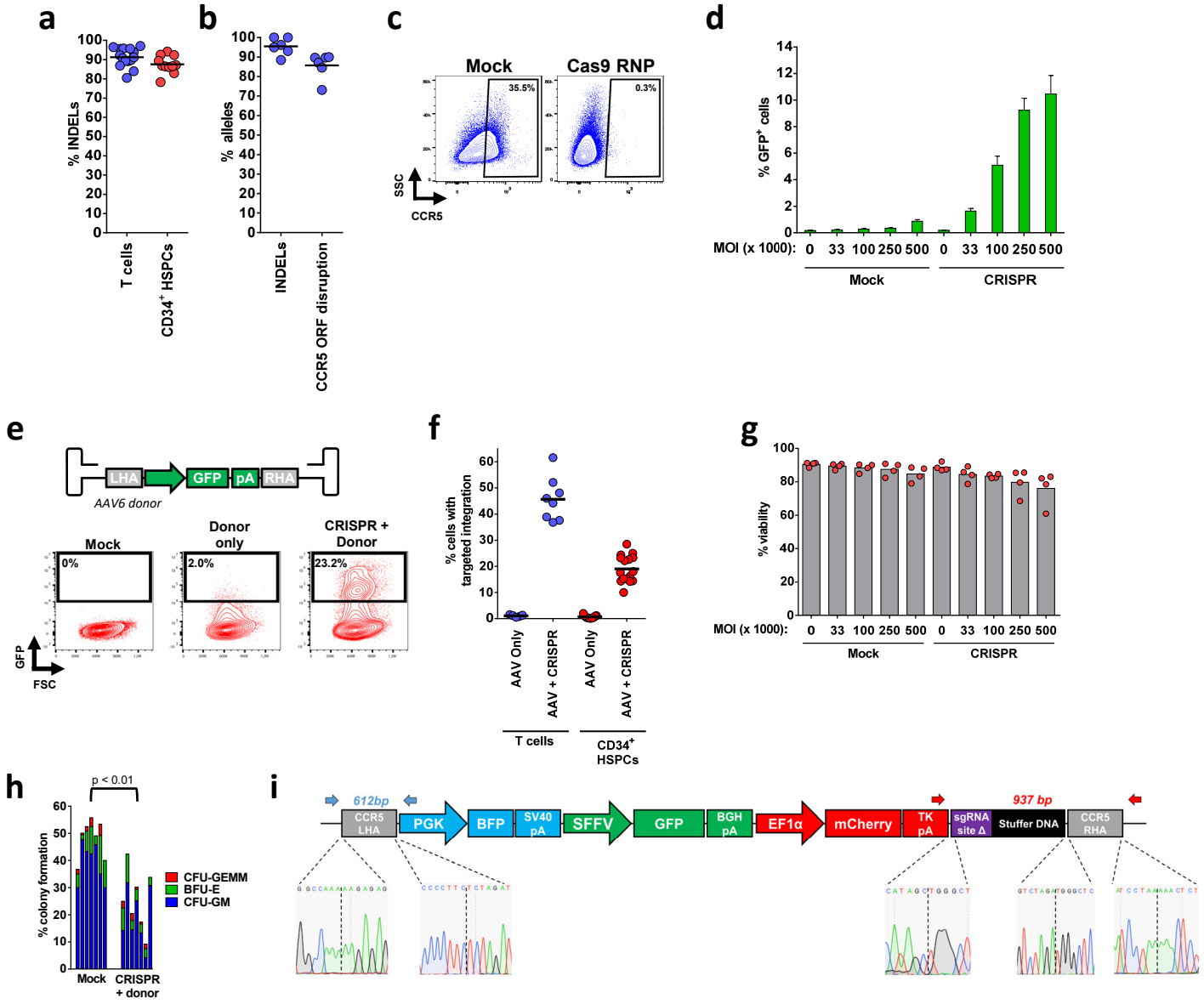

Figure S3. Testing the split *GFP* AAV6 donor pair in primary human T cells and CD34⁺ HSPCs. Related to Figure 3. (a) Stimulated T cells and CD34⁺ HSPCs were electroporated with *CCR5* Cas9 RNP (no AAV transduction) and genomic DNA was extracted after four days. The targeted *CCR5* locus was PCR-amplified, amplicons were Sanger-sequenced, and INDEL rates were analyzed using TIDE (Tracking of Indels by Decomposition). Bars represent means, N = 15 (T cells from 15 different buffy coat donors) and N = 11 (CD34⁺ HSPCs from 11 different umbilical cords). **(b)** The PCR amplicons derived from genomic DNA extracted from RNP-electroporated T cells were TOPO-cloned, transformed into E.coli, and a total of 160 individual colonies (representing different *CCR5* alleles) were sequenced. The sequences were aligned to the *CCR5* sequence of unedited cells and frequencies of total INDELs and open reading frame-disruptive INDELs are plotted. Bars represent means, N = 6 (T cells from 6 different buffy coat donors). **(c)** Representative FACS plots from Mock and RNP-electroporated T cells stained for *CCR5* surface expression four days after electroporation. The CCR5⁺ gate was set based on an isotype antibody control. **(d)** Primary human T cells were stimulated for three days and then electroporated with *CCR5* Cas9 RNP (CRISPR) or mock-electroporated, and then transduced with increasing MOIs of AAV6 split *GFP* donors (MOI is per donor). GFP expression was analyzed by flow cytometry after four days. Bars represent mean +SEM, N = 4 (T cells from 4 different buffy coat donors). **(e)** T cells and CD34⁺ HSPCs were electroporated with *CCR5* Cas9 RNP and transduced with a single AAV6 donor vector encoding GFP. *Upper panel*, Schematic representation of the single *CCR5* AAV6 donor used to assess targeted integration into the *CCR5* locus. The donor contains left and right homology arms (LHA and RHA), which flank the expression cassette with either an SFFV or EF1α promoter, the *GFP* gene, and the *BGH* polyadenylation signal. *Lower panel*, Representative FACS plots 4 days after *CCR5* Cas9 RNP electroporation (CRISPR) and *CCR5* AAV6 donor transduction of CD34⁺ HSPCs. The GFP^{high} population is gated, which when using a single GFP-encoding donor is the population with targeted integration. **(f)** Targeted integration in T cells and CD34⁺ HSPCs using the single *GFP* donor depicted in (e). Bars represent means, N = 8 (T cells from 8 different buffy coat donors) and N = 16 (CD34⁺ HSPCs from 16 different umbilical cords). **(g)** Primary human T cells were stimulated for three days and then electroporated with *CCR5* Cas9 RNP (CRISPR) or mock-electroporated, and then transduced with increasing MOIs of AAV6 split *GFP* donors (MOI is per donor). Viable cells were quantified by flow cytometry three days after electroporation and transduction as negative for an amine reactive viability dye and annexin V stain. N = 4 (T cells from four different buffy coat donors). **(h)** Cord blood-derived CD34⁺ HSPCs were cultured for two days and then electroporated with Cas9 RNP or mock-electroporated. The split *GFP* AAV6 donor pair was added at an MOI of 2 x 500,000 and after four days, mock-electroporated or GFP⁺ cells were single-cell sorted into 96-well plates containing methylcellulose. Formed colonies were counted and scored 14 days after seeding. Colony type distribution showed no difference between the two groups ($p \geq 0.11$; student's paired T test) while a 1.9-fold difference in total colony formation was observed ($p < 0.01$; student's paired T test), although with great donor-donor variability. Each bar represents results from a unique cord blood donor, N = 7. **(i)** PCR primers were designed to confirm targeted integration of Donor A (blue primers) and Donor B (red primers) of the split *GFP* donor system. Both PCRs are In-Out PCRs where one primer is located in the *CCR5* gene outside the region of the homology arm and the other primer is located inside the donor vector insert. PCR fragments from four GFP+ HSPC-derived colonies were gel-purified and sequencing showed seamless HR at all chromosomal junctions (shown with dashed lines) in all four colonies. Representative sequencing chromatograms are shown. See also Figure 3b.

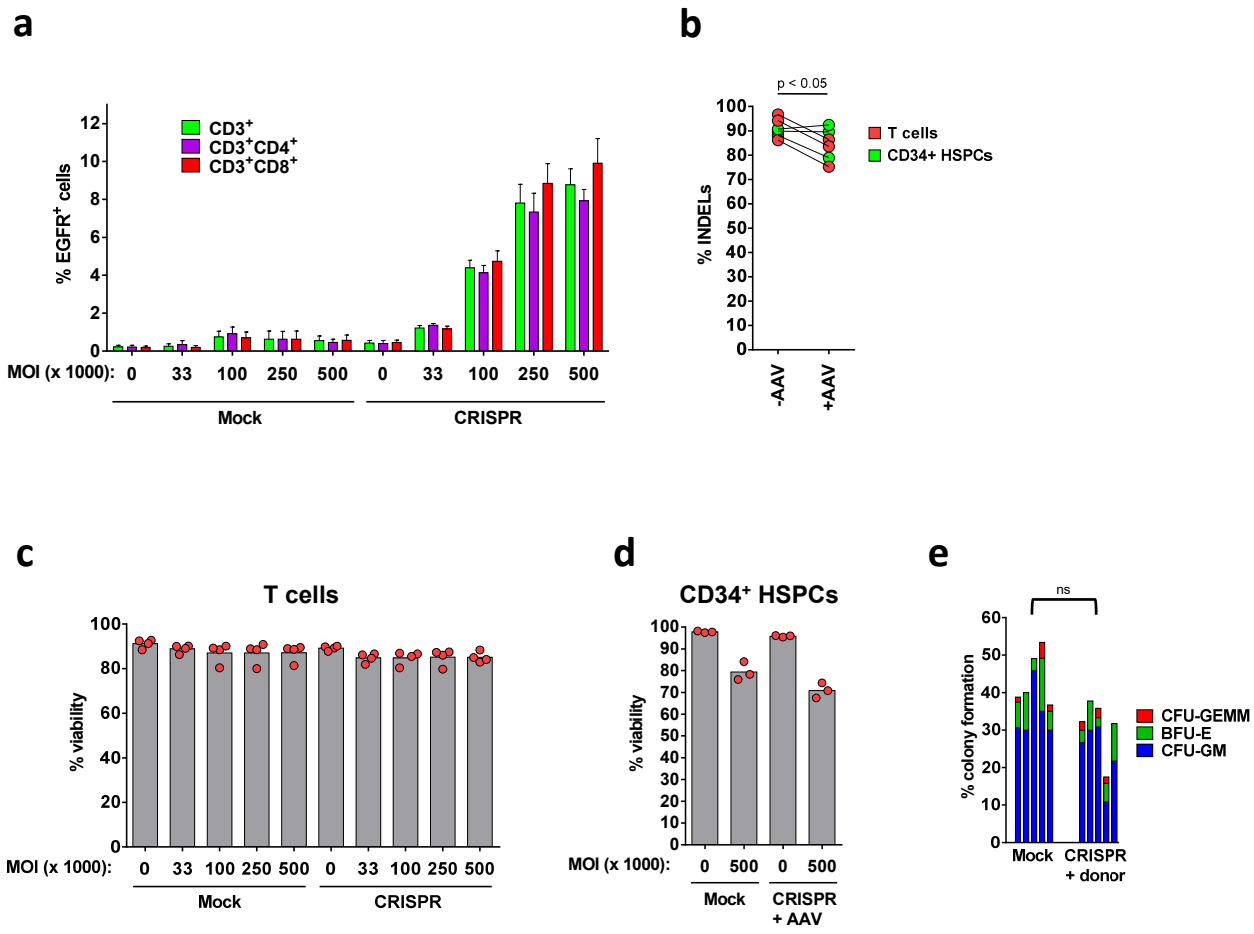

Figure S4. AAV6 dose response, INDEL rates, viabilities, and CFU assay using the split *EGFR* system in primary human T cells and CD34⁺ HSPCs. Related to Figure 4. (a) Primary human T cells were stimulated for three days, electroporated with Cas9 RNP (CRISPR) or mock-electroporated, and then transduced with increasing MOIs of AAV6 split *EGFR* donors (MOI is per donor). EGFR expression was analyzed by flow cytometry after four days in the total $CD3^+$ population and in the $CD4^+$ and $CD8^+$ subpopulations. Bars represent mean +SEM, N = 4 (T cells from four different buffy coat donors). **(b)** Stimulated T cells or $CD34^+$ HSPCs were electroporated with Cas9 RNP and split into two populations that were either left untransduced or transduced with the two AAV6 donors of the *EGFR* system. Four days after, genomic DNA was extracted and the targeted *CCR5* locus was PCR-amplified, amplicons were Sanger-sequenced, and INDEL rates analyzed using TIDE (Tracking of Indels by Decomposition). Note that the PCR only amplifies alleles that have not undergone HR, i.e. WT alleles or alleles with INDELs. N = 3 (T cells from 3 different buffy coat donors) and N = 3 ($CD34^+$ HSPCs from 3 different umbilical cords), $p < 0.05$; student's paired T test. **(c)** Stimulated primary human T cells or **(d)** $CD34^+$ HSPCs were electroporated with Cas9 RNP (CRISPR) or mock-electroporated and then transduced with different MOIs of the AAV6 split *EGFR* donors (MOI is per donor). Viabilities were assessed by flow cytometry three days after electroporation and transduction, and live cells were discriminated as negative for an amine reactive viability dye and annexin V stain. N = 4 for T cells (from four different buffy coat donors) and N = 3 for $CD34^+$ HSPCs (3 different cord blood donors). **(e)** Cord blood-derived $CD34^+$ HSPCs were cultured for two days and then electroporated with Cas9 RNP or mock-electroporated. The split *EGFR* AAV6 donor pair was added at an MOI of 2 x 500,000 and after four days, $EGFR^+$ or mock-electroporated cells were single-cell sorted into 96-well plates containing methylcellulose. Formed colonies were counted and scored 14 days after seeding. Colony type distribution showed no difference between the two groups ($p \geq 0.39$; student's paired T test) while a non-statistical significant 1.4-fold difference in total colony formation was observed ($p = 0.11$; student's paired T test). Each bar represents results from a unique cord blood donor, N=5.