iScience, Volume *25*

Supplemental information

High-level correction of the sickle mutation is amplified *in vivo* during erythroid differentiation

Wendy Magis, Mark A. DeWitt, Stacia K. Wyman, Jonathan T. Vu, Seok-Jin Heo, Shirley J. Shao, Finn Hennig, Zulema G. Romero, Beatriz Campo-Fernandez, Suzanne Said, Matthew S. McNeill, Garrett R. Rettig, Yongming Sun, Yu Wang, Mark A. Behlke, Donald B. Kohn, Dario Boffelli, Mark C. Walters, Jacob E. Corn, and David I.K. Martin

SUPPLEMENTARY MATERIALS:

Figures S1 – S13

Tables S1, S2, and S4

FIGURES

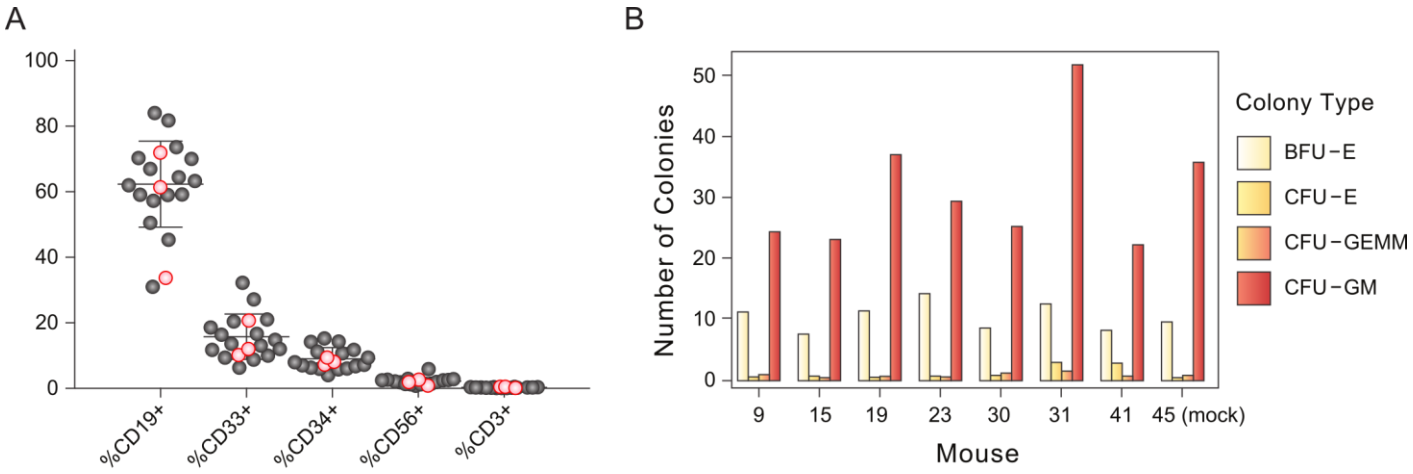


Figure S1, related to Figure 2. Multilineage differentiation of long-term xenografted human hematopoietic cells in a subset of mice from each cohort. A) Immunophenotyping of human cells in mouse marrow by flow cytometry for CD19 (B cells), CD33 (myeloid cells), CD34 (progenitors), CD56 (NK cells), and CD3 (T cells). Lineages are displayed as percent of human CD45+ cells. Cells from mice engrafted with unedited HSPCs are shown as pink circles. B) Colony assays: colony types counted after culture (in semisolid medium) of human CD34+ cells isolated from 16-20 week engrafted mouse marrows.

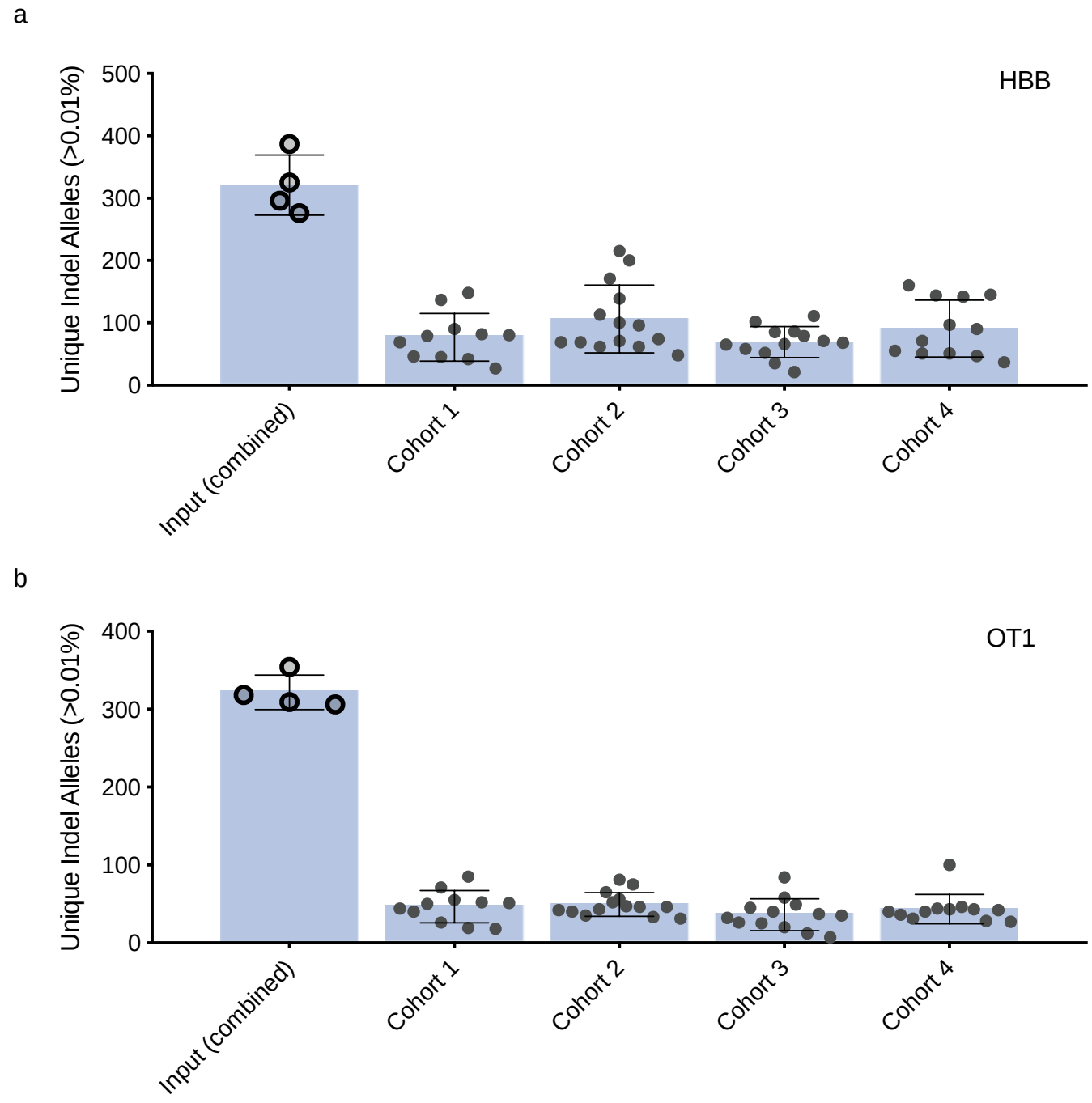
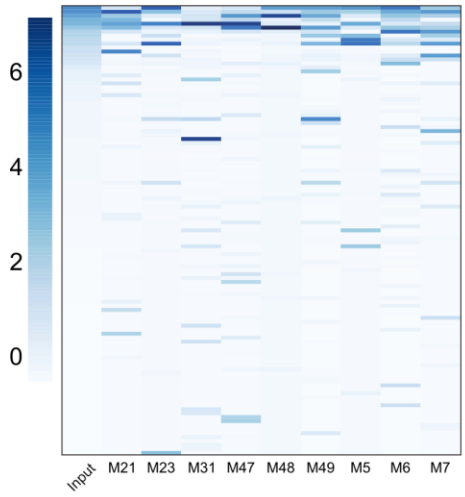
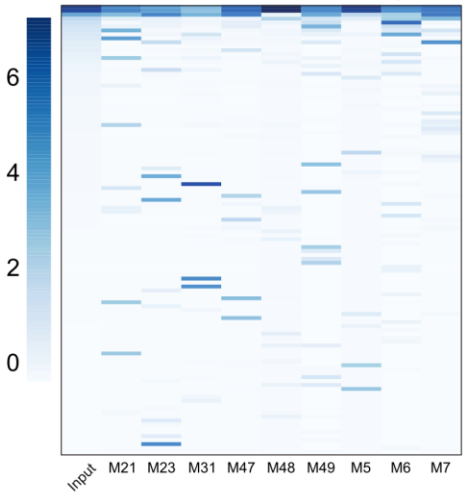


Figure S2, related to Figure 3. Estimation of the number of indel alleles at *HBB* (a) and the principal off-target site OT1 (b) in input human CD34+ HSPCs vs. xenografted marrows in this study (related to Figure 3). An indel is counted as "present" in a sample if it is found in >0.01% of alleles. Error bars represent mean ± s.d. for each set of samples.

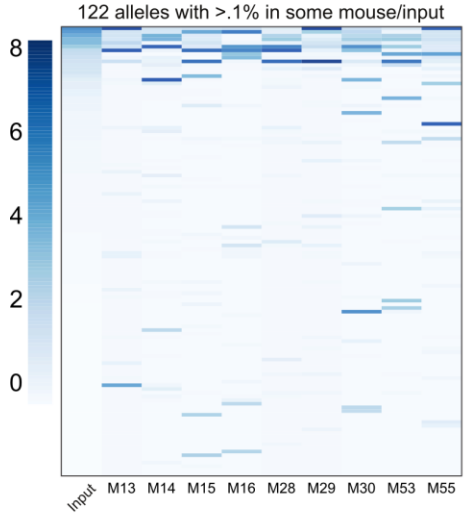
Marrow Cohort 1 HBB
113 alleles with >.1% in some mouse/input



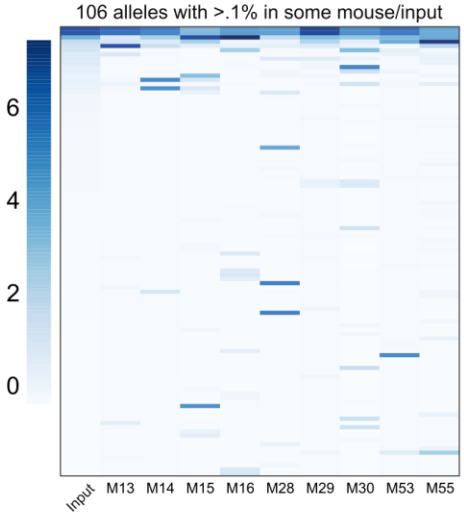
Marrow Cohort 1 OT1
114 alleles >.1% in some mouse/input



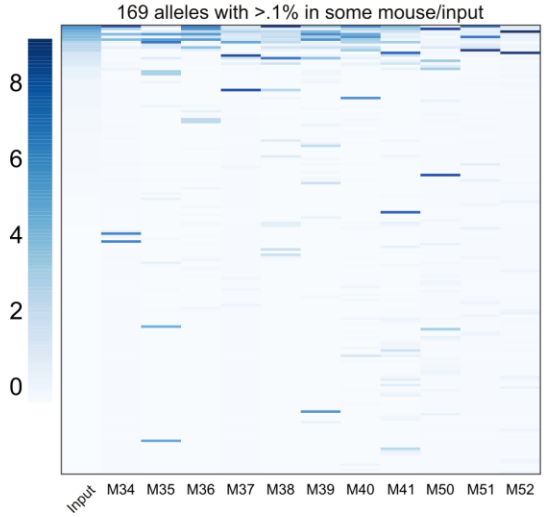
Marrow Cohort 3 HBB
122 alleles with >.1% in some mouse/input



Marrow Cohort 3 OT1
106 alleles with >.1% in some mouse/input



Marrow Cohort 4 HBB
169 alleles with >.1% in some mouse/input



Marrow Cohort 4 OT1
98 alleles with >.1% in some mouse/input

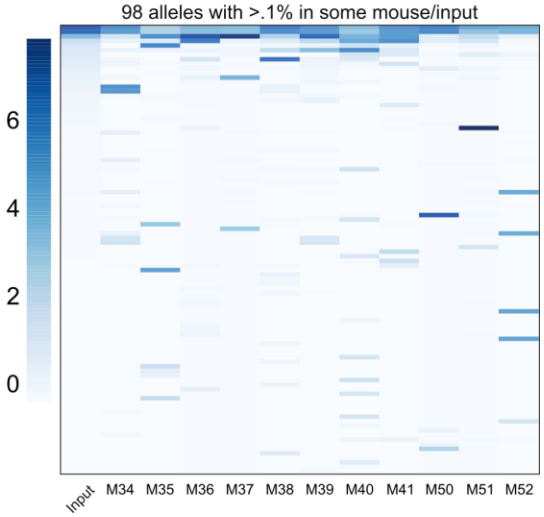
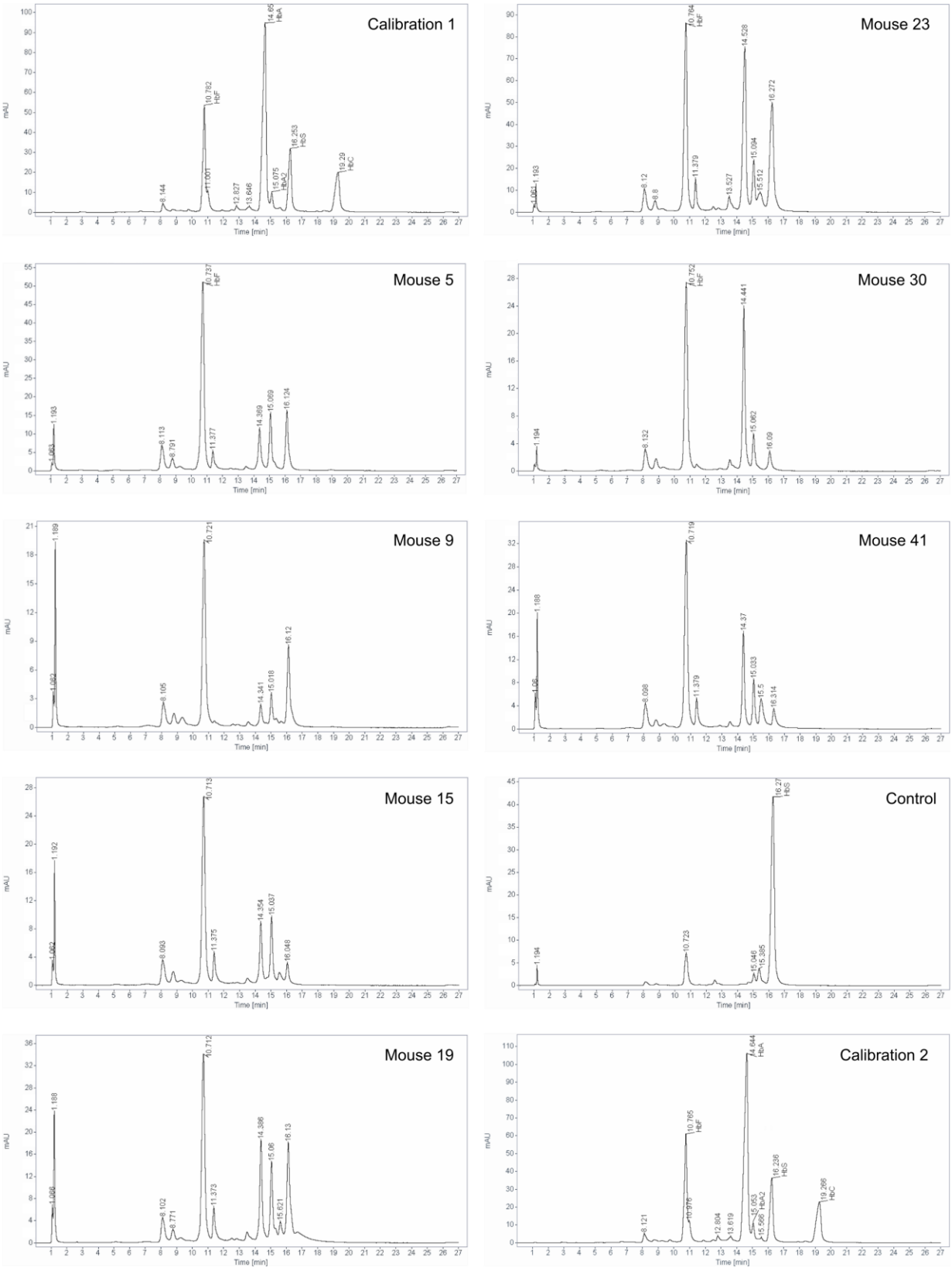


Figure S3, related to Figure 3. Analysis of allelic diversity in edited and xenografted hematopoietic cells: cohorts 1, 3, and 4, related to Figure 3A. Heatmaps of the 140 (*HBB*) and 135 (OT1) indel alleles whose abundance is >0.1% in at least one marrow or in the input cells. Alleles are sorted vertically by their relative frequency in the input population, and normalized in each column by frequency. Note that the most common alleles in the input population are of minor abundance in some mice, while alleles that are rare in the input may be common in one or more mice.

Figure S4, related to Figure 4. HPLC traces from which graphs in Figure 4A were derived, with calibration runs. Table S4 contains data derived from these traces.



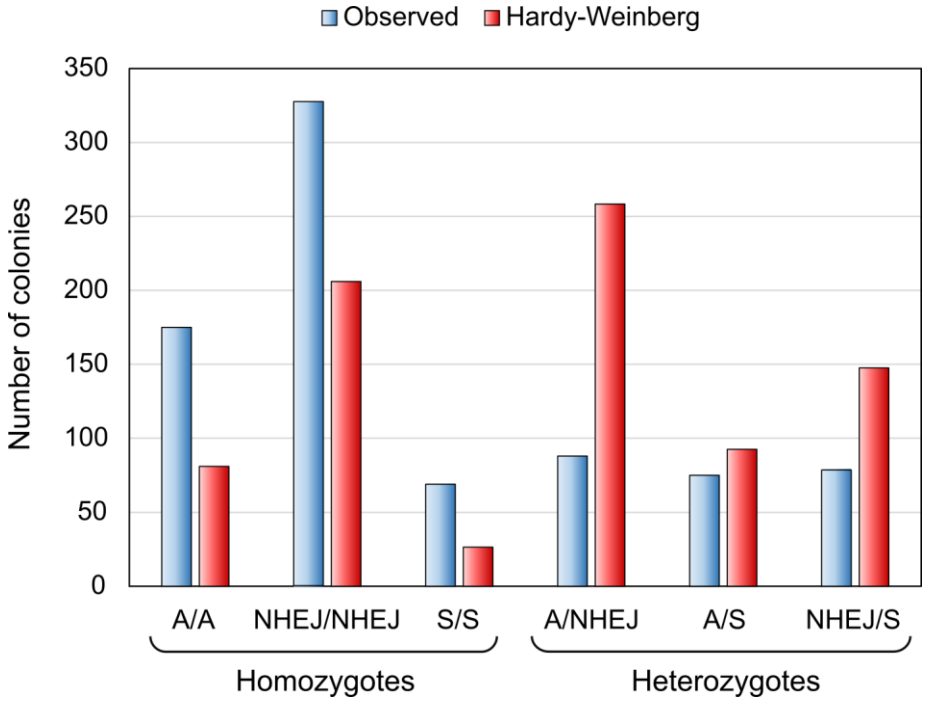


Figure S5, related to Figure 4. *HBB* genotype frequencies in 812 clonal erythroid colonies derived from edited human sickle HSPCs. The two alleles present in each colony were inferred from RNA-Seq data. The y-axis shows the number of colonies with a given genotype. Alleles are identified as follows – A: homology-directed repair (HDR) extends to the site of sickle mutation, producing wild-type ß-globin; NHEJ: any allele containing an insertion or a deletion near the Cas9 cleavage site; S: the sickle mutation is unaffected (this includes alleles in which HDR extended only to the PAM motif – see Figure 1A). Blue columns represent counts of observed genotypes, and red columns represent counts expected for allelic assortment according to the Hardy-Weinberg equilibrium (based on allele frequencies in the 812 colonies). In comparison to the distribution predicted by the Hardy-Weinberg principle, the colonies exhibit an excess of homozygous genotypes ($Chi^2=10^{-37}$), with a corresponding deficit in heterozygous genotypes. This finding might reflect hemizygosity for the edited region in some apparently homozygous cells, since alleles with deletions larger than the genotyped region would not be detected by our genotyping assay. Thus some colonies that are apparently homozygous for gene correction (A/A) might actually be heterozygotes carrying a corrected allele and a large indel.

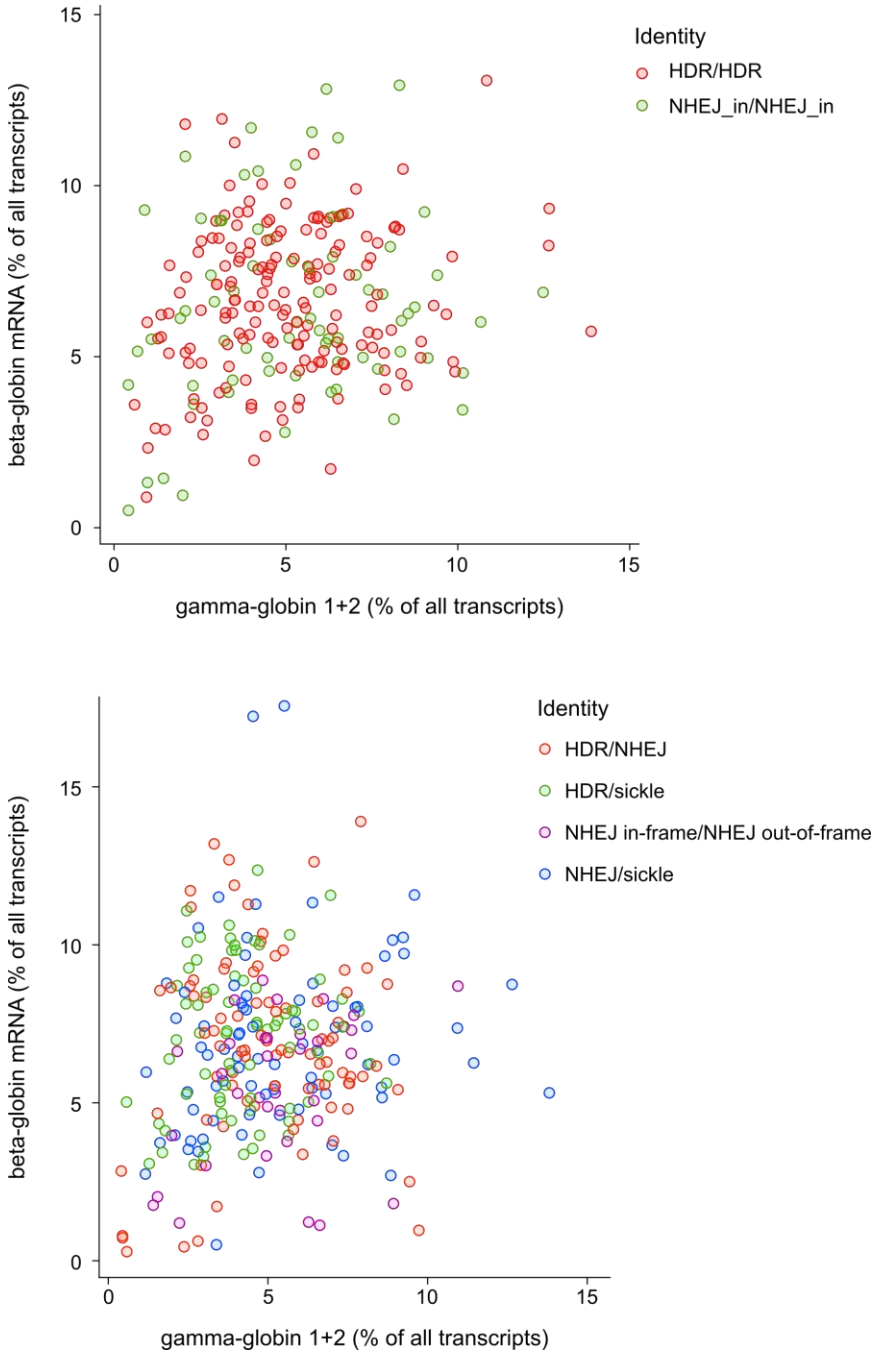


Figure S6, related to Figure 4. **Relationship between the expression of γ-globin and β-globin in erythroid colonies.** Dots represent individual colonies; the x-axis represents the cumulative expression of *HBG1* and *HBG2*; the y-axis represents expression of *HBB*. Expression is reported as the percentage of the expression of all transcripts within a colony. **A**. Comparison of colonies whose *HBB* genotypes are homozygous for in-frame indels with colonies homozygous for the corrected genotype (wild type, also shown in Figure 4C). Colonies with in-frame indels at *HBB* have a similar ratio of *HBB* to *HBG* mRNA as corrected colonies: contrast with the out-of-frame indel colonies shown in Figure 4C. **B**. Comparison of colonies with heterozygous HBB genotypes.

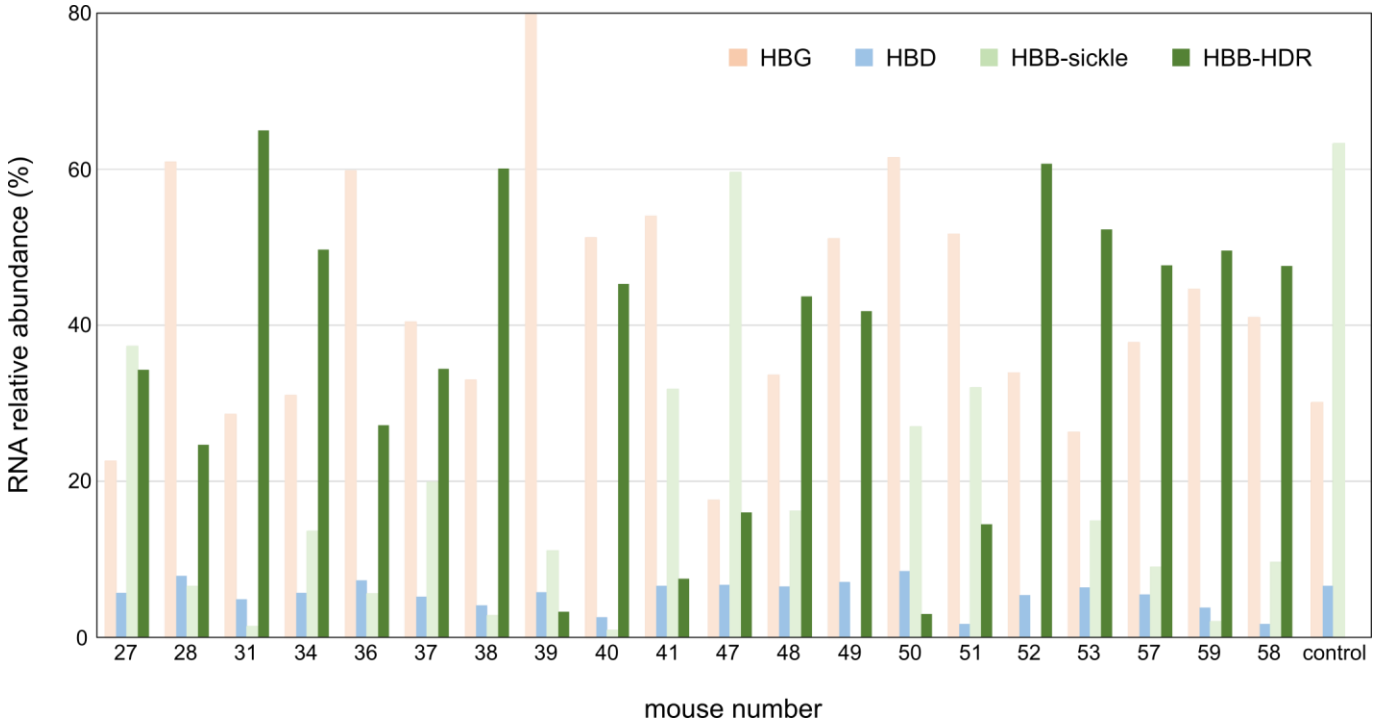


Figure S7, related to Figure 4. Analysis of globin expression in CD235+ erythroid cells. RNA-Seq on CD235+ erythroid cells isolated from xenografted mice (mouse numbers are on the x-axis). The control was engrafted with unedited cells. Only the expression of ß-like globins is shown, with each individual globin reported as abundance relative to total ß-like globin expression.

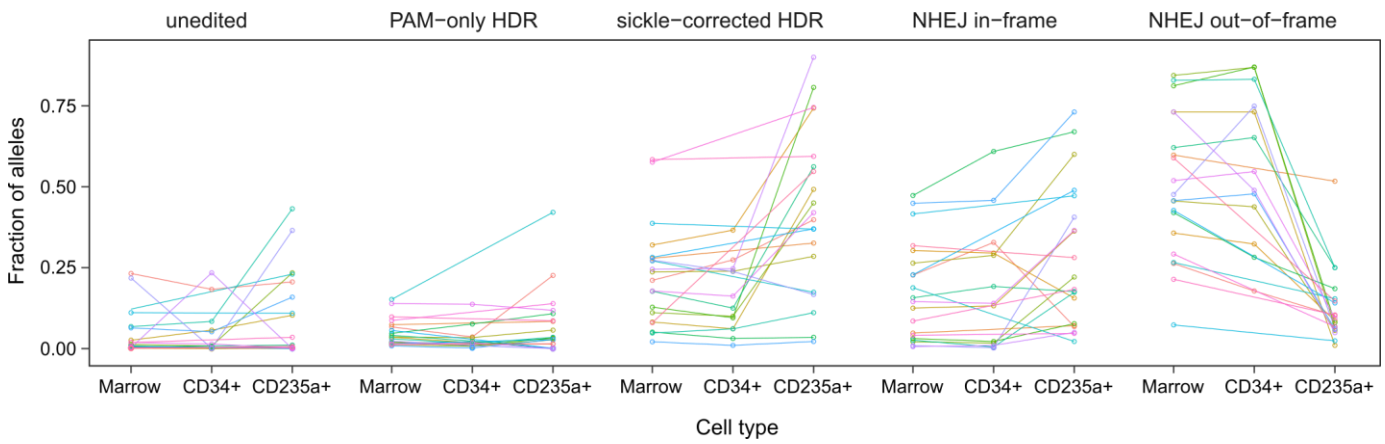


Figure S8, related to Figure 4. Fraction of alleles that have undergone homology-directed repair (HDR) in different populations of hematopoietic cells isolated from mouse marrow: enrichment of corrected alleles in CD235+ erythroid cells. This figure is derived from the same data as Figure 4D, but allows comparison of genotypes in different cell types from the same mouse. PAM-only HDR shows the fraction of haplotypes in which HDR extended to the PAM motif but not to the site of the sickle mutation. Sickle-corrected HDR shows the fraction of haplotypes in which HDR extended to the site of the sickle mutation, correcting the sequence to wild type. NHEJ in-frame shows the fraction of haplotypes with an indel that maintains the β-globin reading frame, while NHEJ out-of-frame shows the fraction of haplotypes in which an indel disrupts it. Haplotype frequencies were determined in cells from whole marrow (n= 40), CD34+ cells (n=26), and CD235+ cells (n=40). Cell populations isolated from a given mouse are shown in the same color and are connected by a line.

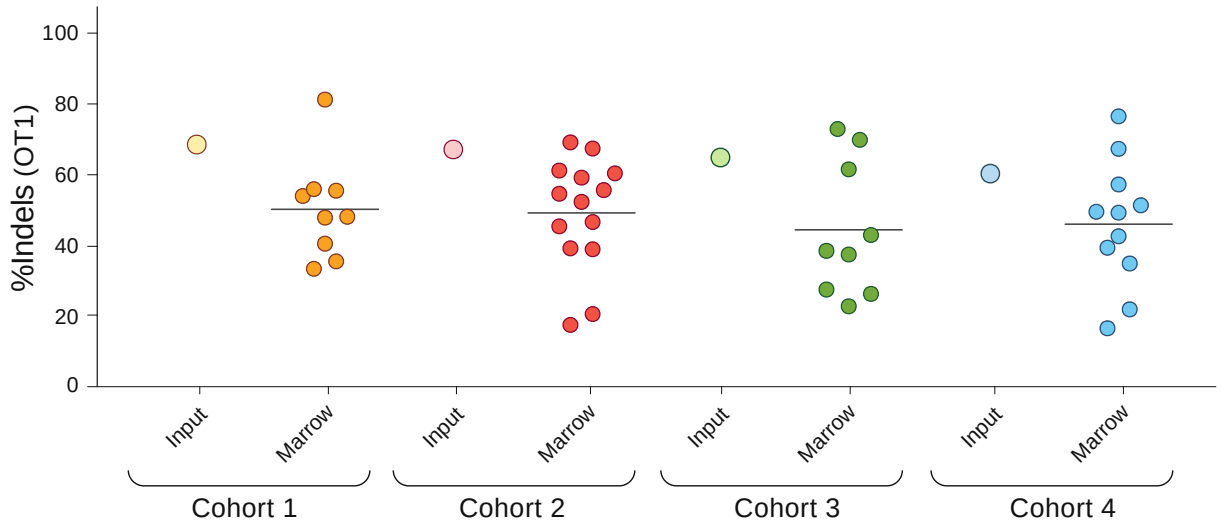


Figure S9, related to Figure 2. NHEJ creates indels at the off-target OT1 site. Indels at intergenic off-target OT1 (chr9:101833584-101833606 in human genome assembly hg38) in xenografted marrow cells at 16-20 weeks post-injection, expressed as percent of all alleles; related to Figure 2C and Table S1. The horizontal line denotes the average indel percent in each group; Input (large open circles) denotes the percent indels in the pool of edited cells injected into each cohort.

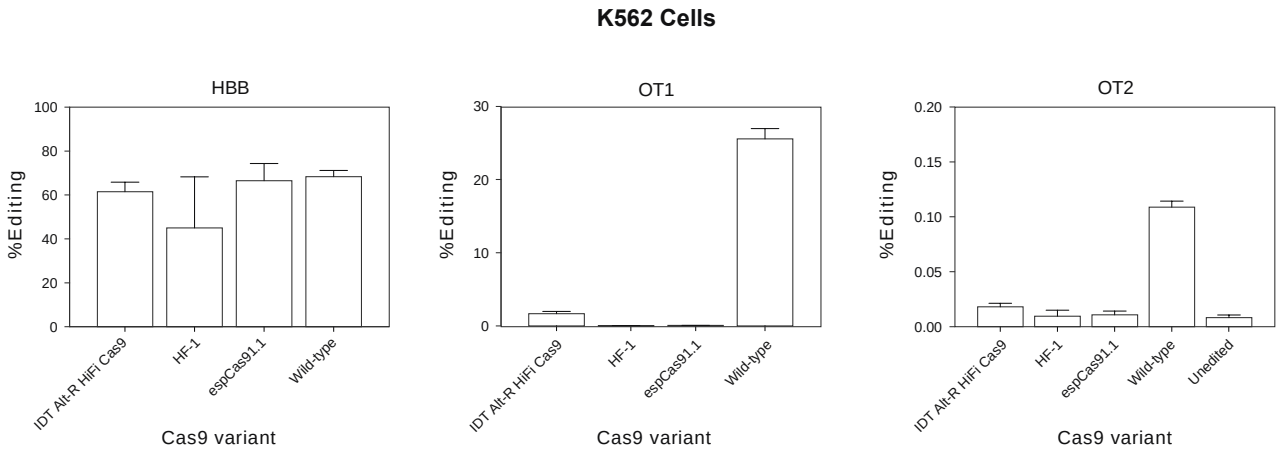


Figure S10, related to Figure 5. Tests of high-fidelity Cas9 RNPs in K562 cells at HBB, OT1, and OT2. Purified IDT AltR HiFi Cas9 (Vakulskas et al., 2018), Cas9 HF-1(Kleinstiver et al., 2016), espCas9-1.1 (Slaymaker et al., 2016) were complexed with guide 3xMS-G10 and delivered to K562 cells by electroporation using published methods (DeWitt et al., 2016), before analysis of editing at off-target sites OT1 and OT2 by targeted amplification (Table S3).

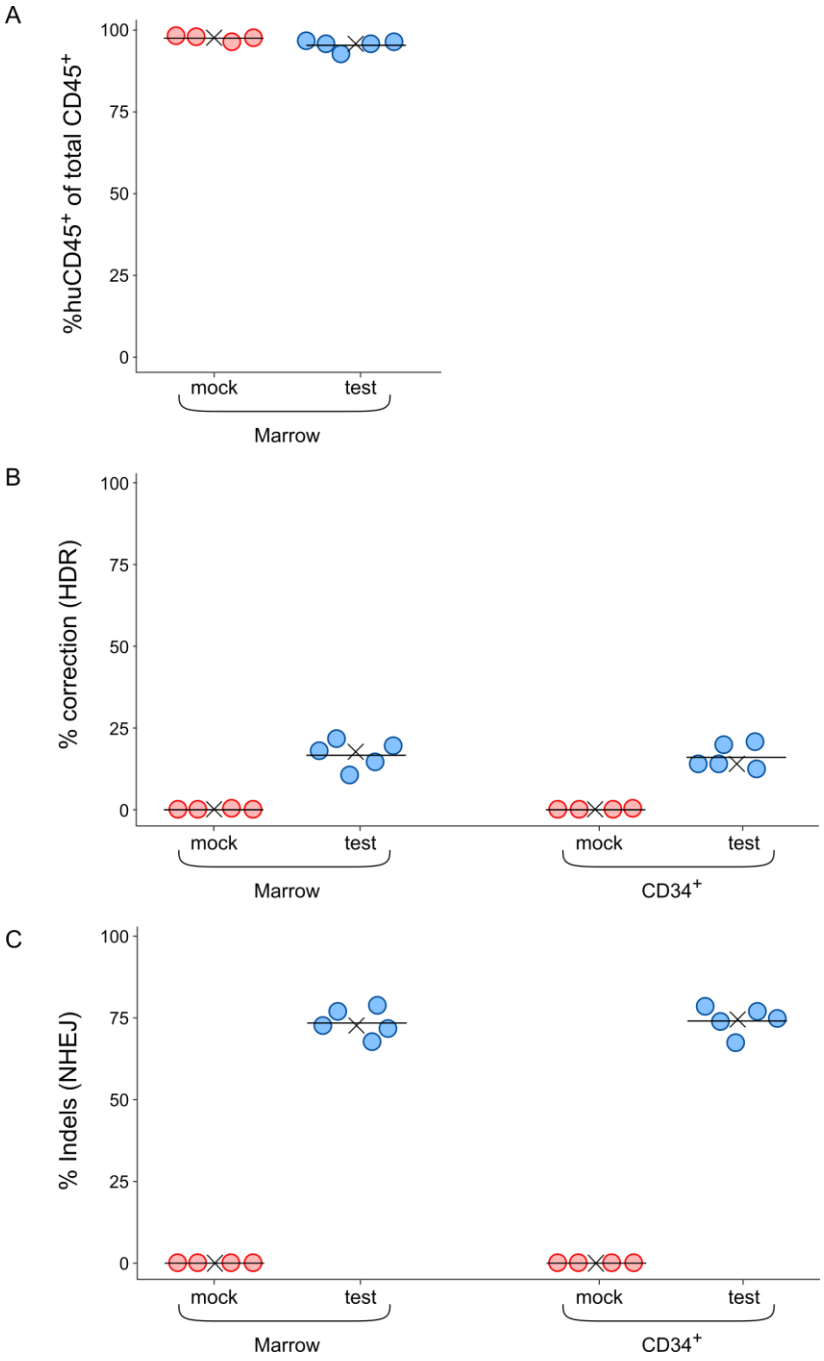


Figure S11, related to Figure 5. **Engraftment and editing of human hematopoietic cells using AltR HiFi Cas9**. The horizontal line and the X symbol denote the mean and the median of each group, respectively. **A**) Engraftment of human hematopoietic cells in mouse bone marrow at 16-20 weeks post-injection, assessed by flow cytometry for human CD 45+ cells, and displayed as percent human CD 45+ in total marrow hematopoietic cells (human CD45+ and mouse CD45+). **B**) Correction of the sickle mutation in xenografted marrow cells of individual mice, and marrow human CD34+ cells, at 16-20 weeks post-injection, assayed by amplicon sequencing at the *HBB* site and expressed as percent of *HBB* alleles. **C.** Indels at the *HBB* target site in xenografted marrow cells, and marrow CD34+ cells, at 16-20 weeks post-injection, assayed and displayed as for B.

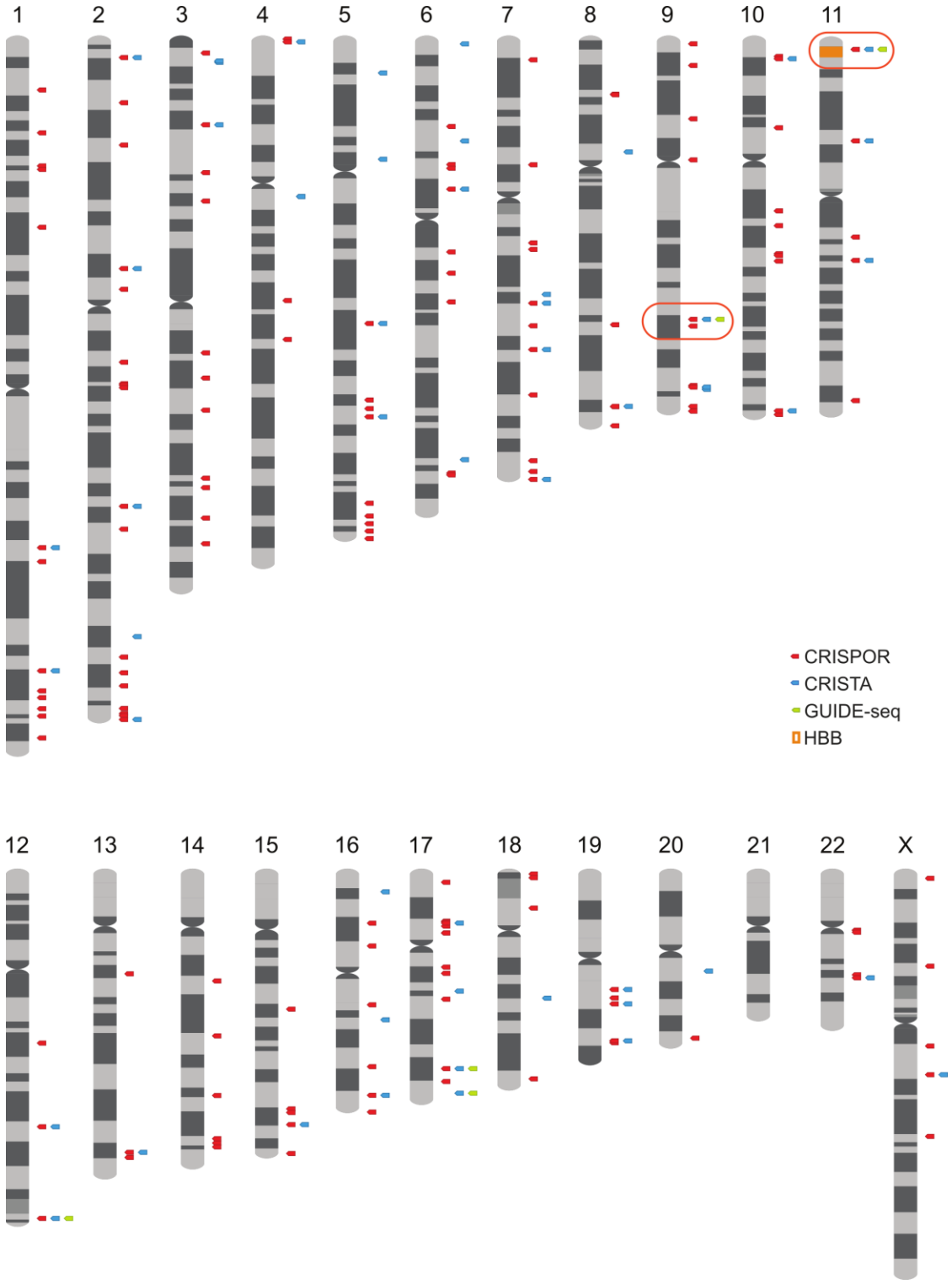


Figure S12, related to Figure 5. Off-target sites identified by GUIDE-seq, CRISTA, and CRISPor. Sites where editing was detected in >0.2% of alleles using high-fidelity Cas9 are circled in red (*HBB*, OT1).

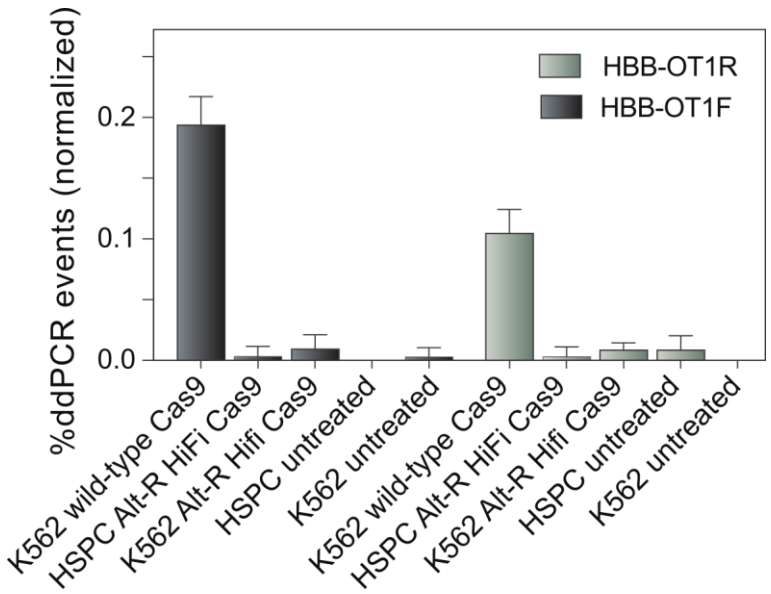


Figure S13, related to Figure 5. Droplet digital PCR (ddPCR) detects translocation linking the region upstream of *HBB* with regions either upstream (OT1F) or downstream (OT1R) of OT1, when *HBB* is targeted with wild-type Cas9, but not with Alt-R HiFi Cas9. Translocations involving the HBB target site have the potential to remodel distal chromatin and thereby affect gene expression. To assess the most likely translocation, between the HBB on-target and the OT1 sites, we devised a droplet digital PCR assay based upon our previous work with HBB-HBD rearrangements (Long et al., 2018). This assay compares the frequency of translocated and non-translocated alleles counted in two separate ddPCRs (normalized to a reference gene).

Table S1, related to Figure 2. Characteristics and key data from NBSGW mice injected for this study (and presented in Figure 2). All mice were ordered from Jackson Labs and were 7-8 weeks old at time of injection.

			Marrow			CD34+		
Cohort	Mouse	% hCD45+	% HDR	% NHEJ	%Un*	% HDR	% NHEJ	%Un*
1	(Input)	N/A	36.48	41.27	2 21.95	N/A	N/A	N/A
1	5 (F)	89.3	6.26	79.47	14.27	2.53	79.91	17.56
1	6 (F)	10 8	29.23	66.91	3.86	21.88	74.69	3.43
1	7 (F)	9 6	27.02	68.42	4.56	22.11	75.41	2.48
1	21 (F)	47.2	19.35	74.06	6.59	15.61	78.41	5.98
1	23 (F)	60.2	14.58	80.35	5.07	9.43	87.32	3.25
1	31 (F)	64.1	28.62	70.55	0.83	34.16	65.30	0.54
1	47 (F)	70 8	45.96	32.43	21.61	nd	nd	nd
1	48 (F)	80.1	39.81	47.34	12.85	nd	nd	nd
1	49 (F)	28.2	26.99	63.09	9.92	nd	nd	nd
2	(Input)	N/A	35.40	40.55	2 24.05	N/A	N/A	N/A
2	9 (F)	85.4	24.18	73.41	2.41	23.74	75.39	0.87
2	10 (F)	75.9	36.33	56.40	7.27	34.32	61.33	4.35
2	11(F)	76.4	9.68	89.99	0.33	6.99	92.82	0.19
2	12 (F)	65.0	16.48	76.37	7.15	10.04	82.97	6.99
2	17 (F)	79.7	35.30	56.76	7.94	35.28	58.28	6.44
2	18 (F)	58.2	14.85	81.49	3.66	15.03	81.45	3.52
2	19 (F)	92.3	37.32	49.94	12.74	49.85	42.00	8.15
2	20 (F)	85.0	29.55	62.37	8.08	24.33	65.38	10.29
2	24 (F)	76.7	37.01	40.74	22.25	43.08	32.12	24.8
2	25 (F)	44.9	49.54	41.57	8.89	47.91	43.30	8.79
2	27 (F)	65.0	22.93	54.69	22.38	27.80	52.64	19.56
2	57 (F)	92.5	52.36	31.89	15.75	nd	nd	nd
2	58 (F)	50.7	51.92	30.47	17.61	nd	nd	nd
2	59 (F)	82.4	8.27	79.95	11.78	nd	nd	nd
3	(Input)	N/A	31.7	47.8	2 20.50	N/A	N/A	N/A
3	13 (M)	25.9	8.99	89.79	1.22	3.48	95.97	0.55
3	14 (M)	21.2	47.46	46.97	5.57	55.28	39.64	5.08
3	15 (M)	63.8	12.40	86.10	1.5	4.56	94.93	0.51
3	16 (M)	20.7	8.46	89.23	2.31	3.34	93.83	2.83
3	28 (F)	76.4	46.09	46.67	7.24	nd	nd	nd
3	29 (M)	26.9	22.32	73.79	3.89	24.71	72.35	2.94
3	30 (M)	15.0	14.68	65.36	19.96	8.54	72.73	18.73
3	53 (M)	35.9	28.52	45.17	26.31	30.02	42.57	27.41
3	55 (Ml	14.7	6.25	68.67	25.08	1.13	75.66	23.21
4	(Input)	N/A	30.38	51.04	18.58	N/A	N/A	N/A
4	34 (M)	23.0	11.62	87.53	0.85	9.23	90.16	0.61
4	35 (M)	44.6	9.68	75.20	15.12	5.86	80.96	13.18
4	36 (M)	58.7	22.52	76.69	0.79	23.30	76.10	0.6
4	37 (M)	68.4	10.23	87.06	2.71	9.27	88.90	1.83
4	38 (M)	32.8	12.37	70.88	16.75	9.69	71.89	18.42
4	39 (M)	19.4	4.41	88.49	7.1	3.43	92.28	4.29
4	40 (M)	10.8	16.42	58.44	25.14	11.86	60.82	27.32
4	41 (M)	69.2	3.14	65.01	31.85	3.63	63.95	32.42
4	50 (M)	17.4	3.06	88.71	8.23	1.66	92.09	6.25
4	51 (M)	41.7	27.41	35.92	36.67	23.59	50.33	26.08
4	52 (M)	47.5	24.69	49.78	25.53	25.38	38.38	36.24

4(control	43(M)	21.9	nd	nd	nd	nd	nd	nd
4(control	44(M)	26.5	nd	nd	nd	nd	nd	nd
4(control	C-3(M)	87.7	nd	nd	nd	nd	nd	nd

*Unedited alleles are defined as encoding a an unedited (i.e. sickle cell) *HBB* gene. These are either completely unedited or include a silent "PAM-only" HDR mutation (discussed at more length in Figure 4D and associated text).

Table S2, related to Figure 5. GUIDE-seq results in K562 cells edited with either in-house produced wild-type Cas9 RNP or Alt-R HiFi Cas9 RNP, purchased from IDT, Inc. or from Aldevron Inc (sold as SpyFi Cas9).

Alt-R HiFi Cas9 (IDT, Inc.)								
Replicate 1			*Replicate 2*			*Replicate 3*		
Cleavage site	Alias	# unique reads	Cleavage site	Alias	# unique reads	Cleavage site		
chr11:5226814	HBB	16891	chr11:5226984	HBB	2583	chr11:5226814	HBB	3789
chr9:101833600	OT1	425	chr9:101833601	OT1	6686	chr9:101833600	OT1	214
Alt-R HiFi Cas9 (Aldevron, Inc.)								
Replicate 1			*Replicate 2*			*Replicate 3*		
Cleavage site	Alias	# unique reads	Cleavage site	Alias	# unique reads	Cleavage site	Alias	GS reads
chr11:5226984	HBB	44802	chr11:5226814	HBB	5035	chr11:5226814	HBB	35221
chr12:124319288	OT2	5	chr9:101833600	OT1	665	chr9:101833600	OT1	568
chr9:101833600	OT1	11305						
WT Cas9 (Berkeley)								
Replicate 1			*Replicate 2*			*Replicate 3*		
Cleavage site	Alias	# unique reads	Cleavage site	Alias	# unique reads	Cleavage site	Alias	GS reads
chr11:5226984	HBB	80313	chr11:5226984	HBB	45556	chr11:5226984	HBB	899800
chr12:93155425		171	chr12:93155426		88	chr12:93155425		3
chr12:124319288	OT2	1834	chr12:124319298	OT2	1718	chr12:124319288	OT2	1012
chr17:78243561		220	chr17:68628105		36	chr12:132288220		56
chr19:33890031		1532	chr17:78243561		207	chr17:68628104		56
chr9:101833600	OT1	49755	chr9:101833600	OT1	27211	chr19:33890031		34
						chr9:101833600	OT1	58685

Table S4. HPLC and RNA-seq data, related to Figure 4 (derived from traces in Figure S4).

	mouse 5		mouse 9		mouse 15		mouse 19		mouse 23		mouse 30		mouse 41		mouse 45	
Hemoglobin	**RT [min]**	**Area %**	**RT [min]**	**Area %**	**RT [min]**	**Area %**	**RT [min]**	**Area %**	**RT [min]**	**Area %**	**RT [min]**	**Area %**	**RT [min]**	**Area %**	**RT [min]**	**Area %**
HbA	14.37	9.08	14.34	3.75	14.35	11.74	14.39	13.35	14.53	26.01	14.44	32.97	14.37	15.60		
HbA2	15.07	12.55	15.02	5.37	15.04	13.81	15.06	10.89	15.09	5.36	15.06	7.62	15.03	7.89	15.05	3.12
HbF	10.74	46.76	10.72	47.58	10.71	45.73	10.71	33.54	10.76	30.39	10.75	47.24	10.72	41.38	10.72	12.19
HbS	16.12	13.60	16.12	19.45	16.05	4.77	16.13	17.40	16.27	21.38	16.09	4.65	16.31	4.66	16.27	76.19
RT_1.1	1.06	0.60	1.06	2.76	1.06	1.88	1.07	2.08	1.06	0.39			1.06	2.48		
RT_1.2	1.19	3.91	1.19	15.11	1.19	9.22	1.19	8.40	1.19	1.64	1.19	1.89	1.19	8.79	1.19	2.13
RT_8.8	8.79	2.91					8.77	2.18	8.80	1.82						
RT_11.4	11.38	3.97			11.38	6.60	11.37	4.86	11.38	3.84			11.38	5.36		
RT_13.5									13.53	2.13						
RT_15.6							15.62	2.49	15.51	3.10			15.50	8.14	15.39	6.36
cal_8.1	8.11	6.62	8.11	5.99	8.09	6.24	8.10	4.80	8.12	3.94	8.13	5.63	8.10	5.71		

globin RNA-seq	**mouse 5 Abundance (% of globin)**	**mouse 9 Abundance (% of globin)**	**mouse 15 Abundance (% of globin)**	**mouse 19 Abundance (% of globin)**	**mouse 23 Abundance (% of globin)**	**mouse 30 Abundance (% of globin)**	**mouse 41 Abundance (% of globin)**	**mouse 45 Abundance (% of globin)**
HBA	56.57	44.83	40.15	49.95	44.02	46.32	47.92	41.79
HBG	18.71	29.98	28.12	20.51	21.63	23.98	24.86	14.50
HBD	4.37	2.45	5.61	3.53	4.64	3.33	4.02	1.39
HBB-sickle	5.11	13.63	2.87	5.54	11.44	1.98	0.00	42.16
HBB-HDR	2.38	4.55	5.04	4.76	11.83	16.32	0.09	0.00
HBB-indel	12.90	4.60	18.15	15.77	6.36	7.96	23.10	0.08
TOTAL	100.04	100.04	99.94	100.06	99.92	99.89	99.99	99.92
BETA-LIKE	43.47	55.21	59.79	50.11	55.90	53.57	52.07	58.13