Using human genetics to develop strategies to increase erythropoietic output from genome-edited hematopoietic stem and progenitor cells

Joab Camarena[1,2]*, Sofia E. Luna[1,2]*, Jessica P. Hampton[1,2], Kiran R. Majeti[1,2], Carsten T. Charlesworth[1,2], Eric Soupene[3], Vivien A. Sheehan[4], M. Kyle Cromer[5,6,7]†, & Matthew H. Porteus[1,2]†

[1]Department of Pediatrics, Stanford University, Stanford, CA.

[2]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University Stanford, CA.

[3]Department of Pediatrics, University of California, San Francisco, Oakland, CA.

[4]Aflac Cancer Center and Blood Disorders, Department of Pediatrics, Emory University School of Medicine, Atlanta, GA.

[5]Department of Surgery, University of California, San Francisco, San Francisco, CA.

[6]Department of Bioengineering & Therapeutic Sciences, University of California, San Francisco, San Francisco, CA, USA.

[7]Eli & Edythe Broad Center for Regeneration Medicine, University of California, San Francisco, San Francisco, CA, USA.

*Contributed equally

†Correspondence: kyle.cromer@uscf.edu, mporteus@stanford.edu

Abstract

Human genetic polymorphisms result in a diversity of phenotypes. Some sequences are pathologic and lead to monogenic diseases, while others may confer beneficial traits. Genome editing is a powerful tool to recreate genotypes found in the population, including the ability to correct pathologic mutations. One interesting naturally occurring mutation is a truncation in the erythropoietin receptor (tEPOR) causing congenital erythrocytosis—characterized by non-pathogenic hyper-production of red blood cells (RBCs). Using the precision of CRISPR/Cas9 genome editing, we have recreated tEPOR and studied the effect of variations of the genotype on RBC development. We then combined tEPOR with a correction strategy developed for β-thalassemia and demonstrated that coupling the two genome editing events gave RBCs a significant selective advantage. This demonstrates the potential of combining human genetics with the precision of genome editing to enable safer and more effective genome editing therapies for patients with serious genetic diseases.

39 **Main**

40 Many of the initial applications of clinical genome editing have aimed to correct or
41 compensate for disease-causing mutations of monogenic diseases. However, human
42 genetic variation is more nuanced than monogenic diseases, as there are variants that
43 appear to confer positive health benefits as well. Examples are people that have bi-allelic
44 deletions in *CCR5* causing resistance to HIV infection, a variety of polymorphisms
45 causing upregulation of fetal hemoglobin, and mutations in *PCSK9* causing low
46 cholesterol levels[1–3]. It is important, however, to broadly assess variations that occur in
47 only small numbers of people for both their potential risks as well as for their potential
48 benefits.

49 Congenital erythrocytosis (CE) is a rare phenotype in which people have higher
50 than normal levels of red blood cells and consequently elevated hemoglobin. The genetic
51 cause of this phenotype was identified by studying the family of a Finnish Olympic gold
52 medal-winning cross-country skier who was found to have levels of hemoglobin >50%
53 higher than normal[4]. It is caused by truncations in the erythropoietin receptor (tEPOR) in
54 which the intracellular inhibitory domain to erythropoietin (EPO) signaling is eliminated[4,5].
55 This domain contains binding sites for SHP1 which normally leads to downregulation of
56 EPO-dependent JAK2/STAT5 signaling (Extended Data Figure 1a). Further studies have
57 shown that tEPOR does not create a constitutively active EPOR signaling cascade, but
58 rather imparts hypersensitivity to EPO[5]. As a consequence, these kindreds with tEPOR
59 typically present with abnormally low levels of EPO, indicating a new homeostasis is
60 attained to prevent CE from becoming pathogenic. Finally, studies of families with CE
61 show that it is not pre-disposing to cancer and is thus considered a benign trait[6].

62 Genome editing is a powerful method that enables the precise changing of
63 nucleotides in the DNA of a cell. There are multiple genome-editing strategies including
64 nuclease-based insertion/deletion (indel) formation, base editing, and prime editing, but
65 the most versatile approach to genome-editing is homology-directed repair (HDR). In
66 HDR, a nuclease-induced double-strand break (DSB) is repaired using a donor template.
67 The natural donor template for HDR is the sister chromatid and the natural repair pathway
68 is homologous recombination. By providing a donor template that resembles a sister
69 chromatid with large homology arms flanking the intended cut site, the homologous
70 recombination machinery can use this "substitute" sister chromatid to repair the DSB.
71 HDR editing is the most flexible because it can create single nucleotide changes,
72 precisely insert large gene cassettes, and even swap out large genomic regions for other
73 gene sequences[7–10]. We utilize all of these applications of HDR in this work.
74 One of the challenges in hematopoietic stem cell gene therapy is to achieve
75 sufficient engraftment of the genetically engineered cells to have a beneficial clinical effect
76 without increasing risk. To make this possible, effort is exerted to maximize editing
77 frequencies in hematopoietic stem and progenitor cells (HSPCs)[11–13]. Even if clinically
78 relevant editing frequencies are achieved, high-morbidity chemo-therapeutic regimens
79 are currently required to create niche space in the bone marrow (BM) for these edited
80 HSPCs, which can create toxicities, including oncogenic risk[14–16]. In this work, we aimed
81 to give edited cells a selective advantage such that low levels of engraftment might still
82 result in a clinical benefit, perhaps enabling less toxic conditioning, through use of
83 genome editing to recreate the CE phenotype by engineering tEPOR into HSPCs in
84 different ways. We find that when tEPOR is engineered into human HSPCs using

genome-editing, there is a significant selective advantage to the derived red blood cells (RBCs). We then show that this selective advantage can be coupled to a therapeutic gene edit to give the cells with the therapeutic edit a selective advantage in RBC development, without affecting the stem and progenitor cells. In this way, we demonstrate the power of combining human genetics with precision genome editing to potentially enable safer and more effective genome editing therapies for patients with serious genetic diseases, particularly those of red blood cells.

Results

Cas9-guided EPOR truncation in HSPCs enhances erythroid proliferation

Truncating mutations in the *EPOR* gene that cause clinically benign erythrocytosis[4,6] provide a potentially safe avenue to increase erythropoietic output from genome-edited HSPCs. In this study we designed Cas9 single-guide RNAs (sgRNAs)[11] (termed *EPOR*-sg1 and *EPOR*-sg2) that overlap the location of the originally identified nonsense mutation, *EPOR* c.1316G>A (p.W439X) ("Mäntyranta variant").[4] (Figure 1a, Extended Data Figure 2a). Our hypothesis was that targeting this site in exon 8 with Cas9 would create a spectrum of indels, a subset of which would result in a frameshift of the reading frame and yield premature downstream stop codons in the *EPOR* gene.

To test this hypothesis, we pre-complexed each sgRNA with high-fidelity Cas9 protein[17] and delivered these ribonucleoprotein (RNP) complexes to human $CD34^+$ HSPCs. At 2-3 days post-editing, we transferred cells into culture media that promotes erythroid differentiation over the course of two weeks[18] (Figure 1a). To determine whether edited HSPCs have a proliferative advantage compared to unedited cells, we harvested

108 genomic DNA at days 0, 4, 7, 11, and 14 of RBC differentiation. We then quantified indel
109 frequency by PCR amplification followed by Sanger sequencing and decomposition
110 analysis using TIDE[19]. In the absence of a selective advantage or disadvantage, percent
111 edited alleles in cells at the beginning and end of RBC differentiation will be roughly
112 equivalent—which is what we observe for the editing frequency of the *HBB* sgRNA used
113 for correction of sickle cell disease. However, we observe that the editing frequency of
114 *EPOR*-targeting sgRNAs increases significantly over the course of erythroid
115 differentiation, to a greater extent in *EPOR*-sg1 than *EPOR*-sg2 (Figure 1b, Extended
116 Data Figure 2b) (P=0.0016 for *EPOR*-sg1 from day 0 to day 14). Additionally, we show
117 that the increase in indels for both *EPOR*-sg1 and *EPOR*-sg2 are predominantly driven
118 by indels that yield downstream stop codons (Extended Data Figure 2c-d, Extended Data
119 Figure 3a-b). This indicates that edited cells, particularly those with premature stop
120 codons, are outcompeting unedited cells because of the EPO hypersensitivity of tEPOR-
121 expressing cells in culture.

122 Because not all indels created by the sgRNAs cause truncations in *EPOR*, we
123 speculated that we could increase this proliferative effect by using HDR to insert a stop
124 codon at the exact location of the original variant (c.1316G>A). To accomplish this, we
125 designed an AAV6 repair template vector that introduces a stop codon into *EPOR* at the
126 439th amino acid (W439X) followed by a BGH-polyA tail to terminate transcription. We
127 also included a downstream GFP marker driven by the constitutive human UbC promoter
128 to ensure that each GFP^+ allele harbors the intended *EPOR*-truncating mutation. The
129 entire integration cassette was flanked by 950-base pair homology arms that
130 corresponded to the genomic DNA immediately upstream and downstream of the

131 intended Cas9 cut site created by the more effective *EPOR*-sg1 (Figure 1c). To determine
132 whether this editing strategy was also able to drive enrichment of genome-edited RBCs,
133 we complexed *EPOR*-sg1 with Cas9 protein and delivered this by electroporation to
134 human $CD34^+$ HSPCs followed by transduction with AAV6 DNA repair template. Two to
135 three days post-editing we either maintained cells in HSPC media or began erythroid
136 differentiation with 3 U/mL of EPO (+EPO), as has been previously described[8], or with 0
137 U/mL of EPO (-EPO) in order to determine whether *tEPOR*-expressing cells retain EPO
138 sensitivity or became EPO independent during their differentiation. At day 14 of erythroid
139 differentiation, we stained for established RBC markers[8] and analyzed cells using flow
140 cytometry. We observed no differentiation when cells were kept in HSPC media and
141 efficient RBC differentiation in all treatments with EPO. In the edited conditions in the
142 absence of EPO, we observed moderate differentiation which may indicate the
143 hypersensitivity of tEPOR-expressing cells to trace amounts of EPO in the media, as has
144 been previously observed[4] (Figure 1d). By analyzing GFP^+ cells over the course of RBC
145 differentiation, in the *EPOR*^{W439X}-edited conditions we observed a significant increase in
146 the frequency of edited cells in only the +EPO conditions (P=0.0006 comparing day 0 to
147 day 14) (Figure 1e). At the end of differentiation, we estimate that almost all of the RBCs
148 are derived from cells with either an INDEL or UBC-GFP knock-in. In addition to the
149 competitive advantage that *tEPOR* expression gives to edited cells over the course of
150 RBC differentiation, we also observed increased RBC production in both *EPOR*-sg1 and
151 *EPOR*-sg1 + UbC-GFP-BGH conditions compared to Mock control in the +EPO condition
152 (average 1.45×10^3 total fold increase in mock edited cells versus 4.71×10^3 in *EPOR*-sg1
153 + UbC-GFP-BGH over the 14-day RBC differentiation) (Figure 1f). These increased cell

154 counts were not observed in the -EPO condition or when cells were maintained in HSPC
155 media, indicating an EPO-driven increase in erythroid proliferation in cells expressing
156 *tEPOR* (Extended Data Figure 4a).

157 In terms of the safety of this editing strategy, we show that introduction of these
158 *EPOR* indels yields RBCs with production of both fetal and adult hemoglobin tetramers
159 following hemoglobin tetramer high-performance liquid chromatography (HPLC)
160 (Extended Data Figure 4b). We also found similar colony number and lineage distribution
161 from $CD34^+$ HSPCs plated into wells containing methylcellulose media either with or
162 without EPO that were scored for colony-formation ability after 14 days (Extended Data
163 Figure 4c). As expected, there was a marked decrease in the ability to form BFU-E
164 colonies in the absence of EPO even if the cells contained the tEPOR. These data
165 reinforce the idea that truncation of EPOR does not alter the lineage bias of HSPCs, but
166 rather has an effect only after commitment to the erythroid lineage.

167

168 **Integration of *tEPOR* cDNA at safe-harbor locus recapitulates the proliferative**
169 **effect**

170 Given the therapeutic utility of transgene integration at safe harbor sites[20], we
171 hypothesized that integration of a *tEPOR* cDNA at a safe harbor site may also enable
172 increased erythroid production from edited HSPCs while leaving the endogenous *EPOR*
173 locus intact. Given the fact that integration at the *CCR5* locus is an established method
174 for delivery of therapeutic transgenes in HSPCs[21], we developed a custom AAV6-
175 packaged DNA repair template that would facilitate integration of an exogenous human
176 UbC promoter driving expression of *tEPOR* cDNA followed by a T2A-YFP-BGH reporter

177 (Figure 2a). Given the strong, constitutive expression of the UbC promoter, this method
178 of insertion is expected to express *tEPOR* ubiquitously in all hematopoietic cell types,
179 regardless of lineage.

180 To test this strategy for *tEPOR* expression, we edited HSPCs with Cas9
181 complexed with an established sgRNA targeting exon 2 of *CCR5*[21] (*CCR5*-sg3),
182 immediately followed by transduction with our custom DNA repair template. We then
183 performed RBC differentiation post-editing and analyzed the kinetics of editing frequency,
184 YFP expression, and erythroid differentiation using droplet digital PCR (ddPCR) and flow
185 cytometry. As with endogenous *EPOR* truncation strategies, we observed more efficient
186 erythroid differentiation in all treatments with EPO compared to the minus EPO conditions
187 (Figure 2b). While flow cytometry confirmed the ubiquitous expression of YFP in edited
188 cells, regardless of presence of CD71 and GPA erythroid markers during differentiation
189 (Figure 2c), we did observe significant enrichment of YFP-expressing RBCs in the
190 presence of EPO ($P<0.0001$ when comparing day 0 to day 14 of RBC differentiation
191 (Figure 2d)). This enrichment was confirmed at the genomic level by ddPCR which
192 showed an increase in percent edited alleles when *tEPOR*-expressing cells were
193 subjected to erythroid differentiation, increasing by an average of 7.3-fold in the presence
194 of EPO, as well as enrichment to a limited degree enrichment in the -EPO condition
195 (Extended Data Figure 5a). With this editing strategy, we observed increased RBC
196 production in the CCR5-sg3 + *tEPOR* condition and no increase in proliferation of CCR5-
197 sg3 alone compared to mock control cells in the +EPO condition. We saw no increased
198 proliferation in the cells maintained in HSPC media or in the -EPO condition (Extended
199 Data Figure 5b).

200 We once again observed that edited and unedited cells had no noticeable
201 difference in lineage commitment or colony-forming ability following a colony-forming unit
202 assay (Extended Data Figure 5c). In addition, we analyzed RBCs post-differentiation
203 using hemoglobin tetramer HPLC and found that UbC-mediated expression of *tEPOR*
204 resulted in both HgbF and HgbA expression but a relative increase in HgbF expression
205 (Extended Data Figure 5d). These results indicate that expression of *tEPOR* from a safe
206 harbor site is an effective means of driving increased RBC production from genome-
207 edited HSPCs.
208

209 **Integration of *tEPOR* cDNA at *HBA1* demonstrates an erythroid-specific**
210 **proliferative effect**

211 While integration at a safe harbor locus effectively increased erythropoietic output from
212 edited HSPCs, there is the concern that constitutive expression in all cell types could
213 disrupt stem-ness or lead to other unintended effects. As an alternative, we can introduce
214 promoter-less transgenes into endogenous genes in order for the integration cassette to
215 be regulated by endogenous expression machinery. For instance, in prior work we
216 designed a genome editing strategy to fully replace the *HBA1* gene with an *HBB*
217 transgene in order to correct β-thalassemia[8]. We found that because α-globin is produced
218 by duplicate genes, *HBA1* may serve as a safe harbor site to deliver custom payloads
219 with strong erythroid-specific expression.

220 We therefore hypothesized that integration of the *tEPOR* cDNA at the *HBA1* locus
221 could further enhance production of edited RBCs while avoiding potential complications
222 with ubiquitous transgene expression because of the specificity of expression of *HBA1* in

223 the RBC lineage. To test this, we designed a custom integration cassette (also packaged
224 in AAV6) to use with an established sgRNA (*HBA1*-sg4) to introduce a promoter-less
225 *tEPOR* cDNA followed by a T2A-YFP reporter under expression of endogenous *HBA1*
226 regulatory machinery (Figure 2e). Following editing, we observed efficient erythroid
227 differentiation in all treatments with EPO and little to no differentiation in the absence of
228 EPO (Figure 2f). We found that this integration strategy indeed yielded RBC-specific
229 expression of YFP, which was only detectable as cells gained CD71 and GPA erythroid
230 markers (Figure 2g). Over the course of RBC differentiation, we observed dramatic
231 enrichment of GFP^+ cells exclusively in the +EPO condition (Figure 2h). While we
232 observed a mild degree of enrichment of edited alleles using ddPCR in the absence of
233 EPO, this effect was more pronounced in the presence of EPO, eliciting an average 4.4-
234 fold increase in percentage of edited alleles over the course of RBC differentiation
235 (Extended Data Figure 6a). Once again, we found that cells edited with *tEPOR* at *HBA1*
236 demonstrated similar colony-forming ability as mock edited cells (Extended Data Figure
237 6b). Additionally, we found that *tEPOR*-expressing cells produce ratios of human
238 hemoglobin that are similar to unedited RBCs (Extended Data Figure 6c).
239

240 **Combined *HBB-tEPOR* bicistronic cassettes simultaneously correct β-thalassemia**
241 **and increase production of edited RBCs**

242 Because the above results demonstrate that *tEPOR* expression yields increased RBC
243 production from edited HSPCs, we then sought to couple this selective advantage with a
244 therapeutic gene edit. We chose to combine tEPOR with our prior β-thalassemia
245 correction approach to simultaneously correct the disease and increase production of

246 clinically meaningful RBCs from these corrected HSPCs. One way this can be
247 accomplished is by creating a bicistronic cassette that links expression of the therapeutic
248 full-length *HBB* transgene with a *tEPOR* cDNA. Prior studies have found that the type of
249 linker domain used can have a great bearing on transgene expression and protein
250 function—particularly when function is dependent on formation of protein complexes, as
251 is the case with the globin genes[8]. Therefore, we designed and tested a variety of AAV6
252 repair template vectors linking the two genes using standard T2A peptides, optimized T2A
253 peptides with Furin cleavage sites[22] (referred to as "FuT2A"), internal ribosome entry sites
254 (referred to as "IRES"), as well by driving *tEPOR* expression from a separate exogenous
255 promoter—human PGK1 (referred to as "PGK"). To evaluate the different vectors, we
256 edited healthy donor HSPCs as previously described at *HBA1* using the bicistronic AAV6
257 repair templates and evaluated their ability to differentiate and enrich for edited RBCs
258 over the course of erythroid differentiation (Figure 3a). While efficient RBC differentiation
259 was achieved in all editing conditions (Figure 3b), we found that all four bicistronic
260 cassettes drove >2-fold enrichment of edited alleles (range: 2.1-fold to 3.5-fold) (P=0.003
261 for PGK-tEPOR, P=0.0055 for T2A-tEPOR, P=0.0259 for F2A-tEPOR, and P=0.0003 for
262 IRES-tEPOR when comparing day 0 to day 14) (Figure 3c-d). We note that allele targeting
263 frequencies of 60% translates into >80% of the cells having at least one allele targeted
264 (cell targeting frequency) and thus would not expect to see much more enrichment than
265 we observed.

266 To ensure this strategy was also effective in patient-derived cells, we tested these
267 bicistronic vectors in HSPCs derived from sickle cell disease (SCD) patients, this time
268 comparing them to a therapeutic full-length *HBB* transgene[8]. All of the constructs are

269 knocked-in to the *HBA1* locus without disrupting the endogenous *HBB* locus expressing
270 HgbS. We found that vectors did not disrupt erythroid differentiation when compared to
271 mock edited cells in the same donor, although their ability to differentiate was likely
272 impacted by the variable quality of the frozen patient samples (Extended Data Figure 7a).
273 Again, we observed >2-fold enrichment of edited alleles for all four bicistronic vectors
274 (range: 2.0-fold to 3.1-fold), but no change in editing frequency for the original β-
275 thalassemia correction vector (P=0.0061 for PGK-tEPOR, P=0.011 for T2A-tEPOR,
276 P=0.0016 for Fu2A-tEPOR, P=0.0153 for IRES-tEPOR when comparing day 0 to day 14)
277 (Figure 3e-f). We estimate that at the end of the differentiation almost all of the cells have
278 at least one allele with the HBB-tEPOR knock-in and thus no biologic drive for further
279 enrichment. When we analyzed differentiated RBCs for hemoglobin tetramers by HPLC,
280 we found that the 2A vectors showed almost no HgbA expression (consistent with what
281 has been seen previously in which addition of a 2A can disrupt HBB protein function[23]).
282 In contrast, we found that PGK-tEPOR and IRES-tEPOR vectors showed an improvement
283 in HgbA production relative to the *HBB*-only edited cells (Extended Data Figure 7b). There
284 was no change in the HgbF expression in these samples.
285

286 **Multiplexed genome editing to introduce *tEPOR* and *HBB* leads to robust increase**
287 **in *β-globin* mRNA in edited RBCs**

288 In lieu of coupling the *tEPOR* cDNA and therapeutic edit at the same locus, an alternative
289 strategy would be to multiplex two editing events at different loci to simultaneously
290 truncate the endogenous *EPOR* and introduce the original β-thalassemia correction
291 vector at *HBA1*. This may have the additional advantage that the endogenous *EPOR*

292 truncation will more reliably recapitulate congenital erythrocytosis. We hypothesized that
293 we could simultaneously deliver Cas9 separately pre-complexed with *EPOR*-truncating
294 *EPOR*-sg1 and *HBA1*-sg4 gRNA and then transduce HSPCs with both the β-thalassemia
295 correction vector as well as the W439X *EPOR*-targeting vector. Because homology arms
296 of each vector are specific for each site—*HBB* for the *HBA1* locus and W439X for the
297 *EPOR* locus—integration for each vector will occur only at the intended locus. This
298 strategy may allow simultaneous correction of β-thalassemia and increased erythropoietic
299 output from corrected HSPCs. For this to be maximally effective, the two editing events
300 must be present in the same cell. Therefore, during editing we used a DNA-PKcs inhibitor
301 to increase the frequency of template integration at each locus[24,25]. Following editing, we
302 determined whether this multiplexed editing strategy increases the frequency of corrected
303 RBCs over the course of erythroid differentiation.

304 To model the clinical setting where both edited and unedited HSPCs would occupy
305 the patient's bone marrow, we introduced unedited cells at various concentrations at the
306 start of erythroid differentiation (Figure 4a). Importantly, none of the multiplexed
307 conditions disrupted erythroid differentiation compared to single-edited *HBB* or mock
308 conditions (Figure 4b). In all of the multiplexed conditions, we observed an increase in
309 the frequency of corrected alleles over the course of RBC differentiation (P=0.0332 for
310 *HBB*+*tEPOR* 100%, P=0.0086 for *HBB*+*tEPOR* 30%, P=0.0122 for *HBB*+*tEPOR* 10%
311 when comparing day 0 to day 14) (Figure 4c). In fact, in both the 30% and 10%
312 multiplexed conditions, we achieved a higher frequency of edited alleles by the conclusion
313 of erythroid differentiation compared to single-edited conditions with unedited cells
314 introduced at an equivalent concentration (P=0.0113 for *HBB* versus *HBB*+*tEPOR* 30%

315 at day 14 and P=0.008 for *HBB* versus *HBB+tEPOR* 10% at day 14) (Figure 4d). We
316 confirmed truncation at the *EPOR* locus measured by GFP^+ cells at day 14 of RBC
317 differentiation (Extended Data Figure 8a). We also observed a corresponding increase in
318 *HBB* mRNA expression from the *HBA1* locus in the 30% and 10% multiplexed edited
319 conditions on day 14 of RBC differentiation compared to single-edited conditions (Figure
320 4e), indicating an improved ability for multiplexed editing to increase therapeutic potential
321 of this β-thalassemia correction strategy. When we measured the colony-forming ability
322 of the edited cells and found that differentiation into the various lineages was similar
323 between mock cells and cells edited with *HBB* alone or *HBB+tEPOR*, although we did
324 observe a decrease in total colonies produced in both edited conditions as would be
325 expected from the increased amount of AAV used to target two loci[26,27] (Extended Data
326 Figure 8b-c).
327

328 **Discussion**

329 While insights from clinical genetics have typically implicated novel genes and
330 pathways in disease, here we sought to use human genetics to develop novel strategies
331 to *treat* disease. We used the precision of genome editing to capitalize on a previously
332 characterized disorder called congenital erythrocytosis, which leads to EPO
333 hypersensitivity and hyperproduction of erythrocytes, without causing pathology[4]. As
334 demonstrated previously with variants identified in human genetics, such as those found
335 in *CCR5*, *PCSK9*, and the gamma-globin promoter region, we hypothesized that we could
336 use CRISPR to introduce this natural *EPOR* variant (*tEPOR*) to increase erythropoietic
337 output from edited HSPCs. Prior work has highlighted the challenge of achieving long-

338 term correction of disease following delivery of gene therapy or genome editing correction
339 strategies[2,28,29]. While many efforts are underway to improve editing and engraftment
340 frequencies, we hypothesized that we could develop a strategy to increase production of
341 the clinically relevant cell type—the red blood cell—from edited HSPCs. If successful,
342 then lower editing and engraftment frequencies could yield sufficient production of RBCs
343 to achieve therapeutic benefit, and thus be curative for patients. Uchida *et al*. have shown
344 that introducing the *tEPOR* variant using lentiviral delivery enhanced the efficacy of
345 shRNA knockdown of *BCL11A* in upregulating HgbF, confirming the beneficial function of
346 this variant[30]. However, our work deploys the precision of genome editing to generate
347 tEPOR which may have broad utility across a spectrum of blood disorders.

348 We explored multiple genome-editing strategies to create the *tEPOR* variant, either
349 through truncation of the endogenous *EPOR* gene or integration of a *tEPOR* cDNA at
350 safe harbor loci. We found that HSPCs expressing *tEPOR* consistently demonstrated
351 increased erythropoiesis but otherwise normal, EPO-dependent production of
352 hemoglobin. In order to increase RBC production of genome-edited HSPCs in the context
353 of disease correction, we combined the *tEPOR* cassette with a previously described β-
354 thalassemia strategy in which the *HBB* gene replaces the *HBA1* gene using homology-
355 directed repair-based genome editing[8]. This allowed us to simultaneously introduce an
356 *HBB* transgene to restore normal hemoglobin production and to increase erythropoietic
357 output from edited HSPCs. To demonstrate the flexibility of the various *tEPOR*-
358 introduction strategies, we also developed an alternative multiplexed, site-specific
359 genome editing strategy to pair the original β-thalassemia correction strategy with
360 introduction of the *EPOR* truncation at the endogenous locus. Both strategies led to

361 enrichment of genome-edited RBCs over the course of erythroid differentiation compared
362 to the traditional β-thalassemia correction strategy. Because we found the effects to be
363 EPO-dependent, it is possible the selective advantage *in vivo* may be even more
364 pronounced because patients suffering from the hemoglobinopathies display elevated
365 EPO levels due to their severe anemia[31].

366 In terms of safety, the strategy to use CRISPR to introduce natural variants has
367 the benefit of having been already "tested" *in vivo* in humans. However, it must be noted
368 that several of the genome editing strategies introduce *tEPOR* under non-native
369 regulation which could alter the normal function of tEPOR. In considering this possibility,
370 we note while every cell in patients with CE harbor an *EPOR* truncation, therapeutic
371 deployment of the genome editing strategy will result in the introduction of *tEPOR* in only
372 a subset of HSPCs resident in the BM. Therefore, any aberrant effects of non-native
373 *tEPOR* expression (e.g., bias away from lymphoid or other cell types) are unlikely to lead
374 to cytopenia given the large number of unedited HSPCs remaining in the bone marrow
375 post-transplant. Furthermore, by increasing erythropoietic output from edited HSPCs, we
376 believe this work could enable the reduction or elimination of high-morbidity myeloablation
377 regimens that are currently required to attain therapeutic levels of edited HSPCs.
378 Expression of *tEPOR* could therefore be integrated into any treatment for blood disorders
379 that involve transplantation of HSPCs. For example, even in an allogeneic hematopoietic
380 stem cell transplant for red blood cell disorders, a truncation in the natural *EPOR* could
381 be created using indel-based genome editing to give the derived transplanted erythroid
382 progenitors a selective advantage. This strategy could thereby enable less toxic

383 myeloablative conditioning to be effectively utilized where mixed chimerism might be the
384 result.

385 Taken together, these results demonstrate the power of combining knowledge
386 from human genetics with the precision of CRISPR genome editing technology to
387 introduce clinically meaningful variants. As human genome sequencing becomes more
388 commonplace and clinically routine[32,33], it is likely that a greater number of variants of
389 unknown significance will be discovered and characterized. We therefore believe that the
390 strategy defined in this work—using CRISPR to introduce natural human variants—may
391 be deployed to amplify the therapeutic potential of current and future cell therapies.

392 **Methods**

393 **AAV6 vector design, production, and purification.** Adeno-associated virus, serotype
394 6 (AAV6) vector plasmids were cloned into the pAAV-MCS plasmid (Agilent
395 Technologies, Santa Clara, CA, USA), comprised of inverted terminal repeats (ITRs)
396 derived from AAV2. Gibson Assembly Mastermix (New England Biolabs, Ipswich, MA,
397 USA) was used for the creation of all DNA repair vectors as per manufacturer's
398 instructions. AAV6 vector was produced and purified with little variation from previously
399 described processes[34]. 293T cells (Life Technologies, Carlsbad, CA, USA) were seeded
400 in five 15 cm^2 dishes with 13-15×10^6 cells per plate 24 hours pre-transfection. Then, each
401 dish was transfected with a standard polyethylenimine (PEI) transfection of 6µg ITR-
402 containing plasmid and 22µg pDGM6 (gift from David Russell, University of Washington,
403 Seattle, WA, USA), which holds the AAV6 cap, AAV2 rep, and Ad5 helper genes.
404 Following a 48-72h incubation, cells were harvested and vectors were purified using the
405 AAVpro purification kit (cat.: 6666; Takara Bio, Kusatsu, Shiga, Japan) as per
406 manufacturer's instructions and stored at -80 °C until further use. AAV6 vectors were
407 titered using ddPCR to measure the number of vector genomes as previously described[35].
408

409 ***In vitro* culture of CD34$^+$ HSPCs.** Human CD34$^+$ HSPCs were cultured in conditions as
410 previously described[7,36–39]. CD34$^+$ HSPCs were isolated from cord blood (provided by
411 Stanford Binns Program for Cord Blood Research) or sourced from Plerixafor- and/or G-
412 CSF-mobilized peripheral blood (AllCells, Alameda, CA, USA and STEMCELL
413 Technologies, Vancouver, Canada). Frozen Plerixafor- and/or G-CSF-mobilized
414 peripheral blood of patients with sickle cell disease were provided by Dr. Vivien Sheehan,

415 Emory University. $CD34^+$ HSPCs were cultured at 1×10^5–5×10^5 cells/mL in StemSpan
416 Serum-Free Expansion Medium II (STEMCELL Technologies, Vancouver, Canada) or
417 Good Manufacturing Practice Stem Cell Growth Medium (SCGM, CellGenix, Freiburg,
418 Germany) supplemented with a human cytokine (PeproTech, Rocky Hill, NJ, USA)
419 cocktail: stem cell factor (100 ng/mL), thrombopoietin (100 ng/mL), Fms-like tyrosine
420 kinase 3 ligand (100 ng/ml), interleukin 6 (100 ng/mL), streptomycin (20 mg/mL), and
421 penicillin (20 U/mL), and 35 nM of UM171 (cat.: A89505; APExBIO, Houston, TX, USA).
422 The cell incubator conditions were 37°C, 5% CO_2, and 5% O_2.
423
424 **Electroporation-aided transduction of cells.** The synthetic chemically modified
425 sgRNAs used to edit $CD34^+$ HSPCs were purchased from Synthego (Menlo Park, CA,
426 USA) or TriLink Biotechnologies (San Diego, CA, USA) and were purified by HPLC.
427 These modifications are comprised of 2′-O-methyl-3′-phosphorothioate at the three
428 terminal nucleotides of the 5′ and 3′ ends described previously[11]. The target sequence for
429 the gRNAs were as follows:
430
431 *EPOR* gRNA (*EPOR*-sg1)
432 5′-AGCTCAGGGCACAGTGTCCA-3′
433
434 *EPOR* gRNA (*EPOR*-sg2)
435 5′-GCTCCCAGCTCTTGCGTCCA-3′
436
437 *CCR5* gRNA (*CCR5*-sg3)

438 5′-GCAGCATAGTGAGCCCAGAA-3′

439

440 *HBA1* gRNA (*HBA1*-sg4)

441 5′-GGCAAGAAGCATGGCCACCG-3′

442

443 The HiFi Cas9 protein was purchased from Integrated DNA Technologies (IDT)
444 (Coralville, Iowa, USA) or Aldevron (Fargo, ND, USA). Preceding electroporation, RNPs
445 were complexed at a Cas9:sgRNA molar ratio of 1:2.5 at 25°C for 10-20 min. Next, $CD34^+$
446 cells were resuspended in P3 buffer (Lonza, Basel, Switzerland) with complexed RNPs
447 and subsequently electroporated using the Lonza 4D Nucleofector and 4D-Nucleofector
448 X Unit (program DZ-100). Electroporated cells were then plated at 1×10^5-5×10^5 cells/mL
449 in the previously described cytokine-supplemented media. Immediately following
450 electroporation, AAV6 was dispensed onto cells at 2.5×10^3-1×10^4 vector genomes/cell
451 based on titers determined by ddPCR. For multiplex editing experiments, in addition to
452 the steps described above, cells were incubated with 0.5 μM of a DNA-PKcs inhibitor,
453 AZD7648 (cat.: S8843; Selleck Chemicals, Houston, TX) for 24 hours, as previously
454 described[24,25].

455

456 **Allelic modification analysis using ddPCR.** Edited HSPCs were harvested within 2-3
457 days post-electroporation and at each media change throughout erythrocyte
458 differentiation and then analyzed for modification frequencies of the alleles of interest. To
459 quantify editing frequencies, we created custom ddPCR primers and probes to quantify
460 HDR alleles (using in-out PCR and probe corresponding to the expected integration

461 event) compared to an established genomic DNA reference (REF) at the *CCRL2* locus[8].
462 QuickExtract DNA extraction solution (Epicentre, Madison, WI, USA) was used to collect
463 gDNA input, which was then digested using BamHI-HF or HindIII-HF as per the
464 manufacturer's instructions (New England Biolabs, Ipswich, MA, USA). The percentage
465 of targeted alleles within a cell population was measured with a Bio-Rad QX200 ddPCR
466 machine and QuantaSoft software (v.1.7; Bio-Rad, Hercules, CA, USA) using the
467 following reaction mixture: 1-4 µL of digested gDNA input, 10 µL of ddPCR SuperMix for
468 Probes (no dUTP) (Bio-Rad), primer/probes (1:3.6 ratio; Integrated DNA Technologies),
469 and volume up to 20 µL with H_2O. ddPCR droplets were then generated following the
470 manufacturer's instructions (Bio-Rad): 20 µL of ddPCR reaction, 70 µL of droplet
471 generation oil, and 40 µL of droplet sample. Thermocycler (Bio-Rad) settings were as
472 follows: 98°C (10 min), 94°C (30 s), 55.7-60°C (30 s), 72°C (2 min), return to step 2 × 40–
473 50 cycles, and 98°C (10 min). ddPCR analysis was carried out as described above with
474 the PCR product generated as the template (20 fg total amplicon). Analysis of droplet
475 samples was then performed using the QX200 Droplet Digital PCR System (Bio-Rad).
476 We next divided the copies/µL for HDR (%): HDR (FAM) / REF (HEX). The following
477 primers and probes were used in the ddPCR reaction:
478
479 *CCR5 (for tEPOR-YFP construct)*
480 Forward Primer (FP): 5'-GGGAGGATTGGGAAGACA-3'
481 Reverse Primer (RP): 5'-AGGTGTTCAGGAGAAGGACA-3'
482 Probe: 5'-6-FAM/AGCAGGCATGCTGGGGATGCGGTGG/3IABkFQ-3'
483

484 *HBA1 (for tEPOR-YFP construct)*

485 FP: 5'-AGTCCAAGCTGAGCAAAGA-3'

486 RP: 5'-ATCACAAACGCAGGCAGAG-3'

487 Probe: 5'-CGAGAAGCGCGATCACATGGTCCTGC-3'

488

489 *HBA1 (for HBB construct and tEPOR-HBB constructs)*

490 FP: 5'-GTGGCTGGTGTGGCTAATG-3'

491 RP: 5'-CAGAAAGCCAGCCAGTTCTT-3'

492 Probe: 5'-6-FAM/CCTGGCCCACAAGTATCACT/3IABkFQ-3'

493

494 *HBA1 (for HBB-tEPOR constructs)*

495 FP: 5'-TCTGCTGCCAGCTTTGAGTA-3'

496 RP: 5'-GCTGGAGTGGGACTTCTCTG-3'

497 Probe: 5'-6-FAM/ACTATCCTGGACCCCAGCTC/3IABkFQ-3'

498

499 *CCRL2 (reference)*

500 FP: 5'- GCTGTATGAATCCAGGTCC-3',

501 RP: 5'- CCTCCTGGCTGAGAAAAAG -3'

502 Probe: 5'- HEX/TGTTTCCTC/ZEN/CAGGATAAGGCAGCTGT/3IABkFQ -3'

503

504 **Indel analysis using TIDE software.** Within 2-4 days post-electroporation, HSPCs were

505 harvested with QuickExtract DNA extraction solution (Epicentre, Madison, WI, USA) to

506 collect gDNA. The following primer sequences were used to amplify the respective cut
507 site at the *EPOR* locus*:*
508 FP: 5'-CAGCTGTGGCTGTACCAGAA-3'
509 RP: 5'-CAGCCTGGTGTCCTAAGAGC-3'
510 Sanger sequencing of respective samples was then used as input for indel frequency
511 analysis relative to a mock, unedited sample using TIDE as previously described[19].
512
513 ***In vitro* differentiation of CD34$^+$ HSPCs into erythrocytes.** Following editing, HSPCs
514 derived from healthy or sickle cell disease patients were cultured 2-3 days as described
515 above. Subsequently, a 14-day *in vitro* differentiation was performed in supplemented
516 SFEMII medium as previously described[18,40]. SFEMII base medium was supplemented
517 with 100 U/mL penicillin–streptomycin, 10 ng/mL SCF (PeproTech, Rocky Hill, NJ, USA),
518 1 ng/mL IL-3 (PeproTech, Rocky Hill, NJ, USA), 3 U/mL EPO (eBiosciences, San Diego,
519 CA, USA), 200 µg/mL transferrin (Sigma-Aldrich, St. Louis, MO, USA), 3% human serum
520 (heat-inactivated from Sigma-Aldrich or ThermoFisherScientific), 2% human plasma
521 (isolated from umbilical cord blood provided by Stanford Binns Cord Blood Program), 10
522 µg/mL insulin (Sigma-Aldrich, St. Louis, MO, USA), and 3 U/mL heparin (Sigma-Aldrich,
523 St. Louis, MO, USA). Cells were cultured in the first phase of medium for seven days at
524 1×10^5 cells/mL. In the second phase of medium, day 7-10, cells were maintained at 1×10^5
525 cells/mL, and IL-3 was removed from the culture. In the third phase of medium, day 11-
526 14, cells were cultured at 1×10^6 cells/mL, with a transferrin increase to 1 mg/mL.
527

528 **Immunophenotyping of differentiated erythrocytes.** Differentiated erythrocytes were
529 analyzed by flow cytometry on day 14 for erythrocyte lineage-specific markers using a
530 FACS Aria II (BD Biosciences, San Jose, CA, USA). Edited and unedited cells were
531 analyzed using the following antibodies: hCD45-V450 (HI30; BD Biosciences, San Jose,
532 CA, USA), CD34-APC (561; BioLegend, San Diego, CA, USA), CD71-PE-Cy7 (OKT9;
533 Affymetrix, Santa Clara, CA, USA), and CD235a-PE (GPA) (GA-R2; BD Biosciences, San
534 Jose, CA, USA). In addition to cell-specific markers, cells were also stained with Ghost
535 Dye Red 780 (Tonbo Biosciences, San Diego, CA, USA) to measure viability.
536

537 **Hemoglobin tetramer analysis.** Frozen pellets of approximately 1×10^6 cells *in vitro*-
538 differentiated erythrocytes were thawed and lysed in 30 µL of RIPA buffer with 1x Halt™
539 Protease Inhibitor Cocktail (Thermo Fisher Scientific, Waltham, MA, USA) for 5 minutes
540 on ice. The mixture was vigorously vortexed and cell debris were removed by
541 centrifugation at 13,000 RPM for 10 minutes at 4°C. HPLC analysis of hemoglobins in
542 their native form was performed on a cation-exchange PolyCAT A column (35 × 4.6 mm^2,
543 3 µm, 1,500 Å) (PolyLC Inc., Columbia, MD, USA) using a Perkin-Elmer Flexar HPLC
544 system (Perkin-Elmer, Waltham, MA, USA) at room temperature and detection at 415 nm.
545 Mobile phase A consisted of 20 mM Bis-tris and 2 mM KCN at pH 6.94, adjusted with HCl.
546 Mobile phase B consisted of 20 mM Bis-tris, 2 mM KCN, and 200 mM NaCl at pH 6.55.
547 Hemolysate was diluted in buffer A prior to injection of 20 µL onto the column with 8%
548 buffer B and eluted at a flow rate of 2 mL/min with a gradient made to 40% B in 6 min,
549 increased to 100% B in 1.5 min, and return to 8% B in 1 min and equilibrated for 3.5 min.
550 Quantification of the area under the curve of the peaks was performed with TotalChrom

551 software (Perkin-Elmer, Waltham, MA, USA) and raw values were exported to GraphPad
552 Prism 9 for plotting and further analysis.
553
554 **mRNA analysis.** Following differentiation of HSPCs into erythrocytes, cells were
555 harvested, and RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, Hilden,
556 Germany). Subsequently, complementary DNA was made from approximately 100 ng of
557 RNA using the iScript Reverse Transcription Supermix for quantitative PCR with reverse
558 transcription (Bio-Rad, Hercules, CA, USA). Expression levels of β-globin transgene and
559 α-globin mRNA were quantified with a Bio-Rad QX200 ddPCR machine and QuantaSoft
560 software (v.1.7; Bio-Rad, Hercules, CA, USA) using the following primers and 6-FAM/
561 ZEN/IBFQ-labeled hydrolysis probes, purchased as custom-designed PrimeTime qPCR
562 Assays from IDT:
563 *HBB and HBB-tEPOR into HBA1:*
564 FP: 5'-GGTCCCCACAGACTCAGAGA-3'
565 RP: 5'-CAGCATCAGGAGTGGACAGA
566 Probe: 5'-6-FAM/AACCCACCATGGTGCATCTG/3IABkFQ -3'
567 To normalize for RNA input, levels of the RBC-specific reference gene GPA were
568 determined in each sample using the following primers and HEX/ZEN/IBFQ-labeled
569 hydrolysis probes, purchased as custom-designed PrimeTime qPCR Assays from IDT:
570 *GPA (reference):*
571 FP: 5′-ATATGCAGCCACTCCTAGAGCTC-3′
572 RP: 5′-CTGGTTCAGAGAAATGATGGGCA-3′
573 Probe: 5'- HEX/AGGAAACCGGAGAAAGGGTA/3IABkFQ -3'

574 ddPCR reactions were created using the respective primers and probes, and droplets
575 were generated as described above. Thermocycler (Bio-Rad; settings were as follows:
576 98°C (10 min), 94°C (30 s), 54°C (30 s), 72°C (30 s), return to step 2 × 50 cycles, and
577 98°C (10 min). Analysis of droplet samples was done using the QX200 Droplet Digital
578 PCR System (Bio-Rad). To determine relative expression levels, the numbers of *HBB*
579 transgene copies/mL were divided by the numbers of *GPA* copies/mL.
580
581 **Methylcellulose colony forming unit (CFU) assay.** 2-3 days post-electroporation
582 HSPCs were plated in SmartDish 6-well plates (cat.: 27370; STEMCELL Technologies,
583 Vancouver, Canada) containing MethoCult H4434 Classic or MethoCult H4434 Classic
584 without EPO (cat.: 04444; cat: 04544.; STEMCELL Technologies, Vancouver, Canada).
585 After 14 days, the wells were imaged using the STEMvision Hematopoietic Colony
586 Counter (STEMCELL Technologies, Vancouver, Canada). Colonies were counted and
587 scored to determine the number of BFU-E, CFU-E, CFU-GM, and CFU-GEMM colonies.
588
589 **Statistical analysis.** GraphPad Prism 9 software was used for all statistical analysis.
590
591 **References**

592 1. Hütter G, Nowak D, Mossner M, et al. Long-Term Control of HIV by CCR5
593 Delta32/Delta32 Stem-Cell Transplantation. *N Engl J Med*. 2009;360(7):692-698.
594 doi:10.1056/NEJMoa0802905

595 2. Frangoul H, Altshuler D, Cappellini MD, et al. CRISPR-Cas9 Gene Editing for Sickle
596 Cell Disease and β-Thalassemia. *N Engl J Med*. 2021;384(3):252-260.
597 doi:10.1056/NEJMoa2031054

598 3. Fitzgerald K, White S, Borodovsky A, et al. A Highly Durable RNAi Therapeutic
599 Inhibitor of PCSK9. *N Engl J Med*. 2017;376(1):41-51. doi:10.1056/NEJMoa1609243

600 4. de la Chapelle A, Träskelin AL, Juvonen E. Truncated erythropoietin receptor
601 causes dominantly inherited benign human erythrocytosis. *Proc Natl Acad Sci U S*
602 *A*. 1993;90(10):4495-4499.

603 5. Sokol L, Luhovy M, Guan Y, Prchal J, Semenza G, Prchal J. Primary familial
604 polycythemia: a frameshift mutation in the erythropoietin receptor gene and
605 increased sensitivity of erythroid progenitors to erythropoietin. *Blood*. 1995;86(1):15-
606 22. doi:10.1182/blood.V86.1.15.bloodjournal86115

607 6. Arcasoy MO, Degar BA, Harris KW, Forget BG. Familial erythrocytosis associated
608 with a short deletion in the erythropoietin receptor gene. *Blood*. 1997;89(12):4628-
609 4635.

610 7. Dever DP, Bak RO, Reinisch A, et al. CRISPR/Cas9 β-globin gene targeting in
611 human haematopoietic stem cells. *Nature*. 2016;539(7629):384-389.
612 doi:10.1038/nature20134

613 8. Cromer MK, Camarena J, Martin RM, et al. Gene replacement of α-globin with β-
614 globin restores hemoglobin balance in β-thalassemia-derived hematopoietic stem
615 and progenitor cells. *Nat Med*. 2021;27(4):677-687. doi:10.1038/s41591-021-01284-
616 y

617 9. Vaidyanathan S, Salahudeen AA, Sellers ZM, et al. High-Efficiency, Selection-free
618 Gene Repair in Airway Stem Cells from Cystic Fibrosis Patients Rescues CFTR
619 Function in Differentiated Epithelia. *Cell Stem Cell*. 2020;26(2):161-171.e4.
620 doi:10.1016/j.stem.2019.11.002

621 10. Choi E, Koo T. CRISPR technologies for the treatment of Duchenne muscular
622 dystrophy. *Mol Ther*. 2021;29(11):3179-3191. doi:10.1016/j.ymthe.2021.04.002

623 11. Hendel A, Bak RO, Clark JT, et al. Chemically modified guide RNAs enhance
624 CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol*.
625 2015;33(9):985-989. doi:10.1038/nbt.3290

626 12. Shin JJ, Schröder MS, Caiado F, et al. Controlled Cycling and Quiescence Enables
627 Efficient HDR in Engraftment-Enriched Adult Hematopoietic Stem and Progenitor
628 Cells. *Cell Rep*. 2020;32(9):108093. doi:10.1016/j.celrep.2020.108093

629 13. Transient Inhibition of 53BP1 Increases the Frequency of Targeted Integration in
630 Human Hematopoietic Stem and Progenitor Cells. doi:10.21203/rs.3.rs-2621625/v1

631 14. Chakraborty NG, Bilgrami S, Maness LJ, et al. Myeloablative chemotherapy with
632 autologous peripheral blood stem cell transplantation for metastatic breast cancer:
633 immunologic consequences affecting clinical outcome. *Bone Marrow Transplant*.
634 1999;24(8):837-843. doi:10.1038/sj.bmt.1701999

15. Bhatia S. Long-term health impacts of hematopoietic stem cell transplantation inform recommendations for follow-up. *Expert Rev Hematol*. 2011;4(4):437-454. doi:10.1586/ehm.11.39

16. Inamoto Y, Shah N, Savani B, et al. Secondary solid cancer screening following hematopoietic cell transplantation. *Bone Marrow Transplant*. 2015;50(8):1013-1023. doi:10.1038/bmt.2015.63

17. Vakulskas CA, Dever DP, Rettig GR, et al. A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat Med*. 2018;24(8):1216-1224. doi:10.1038/s41591-018-0137-0

18. Dulmovits BM, Appiah-Kubi AO, Papoin J, et al. Pomalidomide reverses γ-globin silencing through the transcriptional reprogramming of adult hematopoietic progenitors. *Blood*. 2016;127(11):1481-1492. doi:10.1182/blood-2015-09-667923

19. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res*. 2014;42(22):e168. doi:10.1093/nar/gku936

20. Papapetrou EP, Schambach A. Gene Insertion Into Genomic Safe Harbors for Human Gene Therapy. *Mol Ther J Am Soc Gene Ther*. 2016;24(4):678-684. doi:10.1038/mt.2016.38

21. Gomez-Ospina N, Scharenberg SG, Mostrel N, et al. Human genome-edited hematopoietic stem cells phenotypically correct Mucopolysaccharidosis type I. *Nat Commun*. 2019;10(1):4045. doi:10.1038/s41467-019-11962-8

22. Fang J, Yi S, Simmons A, et al. An antibody delivery system for regulated expression of therapeutic levels of monoclonal antibodies in vivo. *Mol Ther J Am Soc Gene Ther*. 2007;15(6):1153-1159. doi:10.1038/sj.mt.6300142

23. Reinhardt A, Kagawa H, Woltjen K. N-Terminal Amino Acids Determine KLF4 Protein Stability in 2A Peptide-Linked Polycistronic Reprogramming Constructs. *Stem Cell Rep*. 2020;14(3):520-527. doi:10.1016/j.stemcr.2020.01.014

24. Selvaraj S, Feist WN, Viel S, et al. High-efficiency transgene integration by homology-directed repair in human primary cells using DNA-PKcs inhibition. *Press Nat Biotechnol*.

25. Wimberger S, Akrap N, Firth M, et al. Simultaneous inhibition of DNA-PK and PolΘ improves integration efficiency and precision of genome editing. Published online December 15, 2022:2022.12.15.520396. doi:10.1101/2022.12.15.520396

26. Schiroli G, Conti A, Ferrari S, et al. Precise Gene Editing Preserves Hematopoietic Stem Cell Function following Transient p53-Mediated DNA Damage Response. *Cell Stem Cell*. 2019;24(4):551-565.e8. doi:10.1016/j.stem.2019.02.019

672 27. Ferrari S, Jacob A, Cesana D, et al. Choice of template delivery mitigates the
673 genotoxic risk and adverse impact of editing in human hematopoietic stem cells. *Cell*
674 *Stem Cell*. 2022;29(10):1428-1444.e9. doi:10.1016/j.stem.2022.09.001

675 28. Kanter J, Walters MC, Krishnamurti L, et al. Biologic and Clinical Efficacy of
676 LentiGlobin for Sickle Cell Disease. *N Engl J Med*. 2022;386(7):617-628.
677 doi:10.1056/NEJMoa2117175

678 29. George LA, Sullivan SK, Giermasz A, et al. Hemophilia B Gene Therapy with a
679 High-Specific-Activity Factor IX Variant. *N Engl J Med*. 2017;377(23):2215-2227.
680 doi:10.1056/NEJMoa1708538

681 30. Uchida N, Ferrara F, Drysdale CM, et al. Sustained fetal hemoglobin induction in
682 vivo is achieved by BCL11A interference and coexpressed truncated erythropoietin
683 receptor. *Sci Transl Med*. 2021;13(591):eabb0411.
684 doi:10.1126/scitranslmed.abb0411

685 31. Nişli G, Kavakli K, Aydinok Y, Oztop S, Cetingül N. Serum erythropoietin levels in
686 patients with beta thalassemia major and intermedia. *Pediatr Hematol Oncol*.
687 1997;14(2):161-167. doi:10.3109/08880019709030902

688 32. Halldorsson BV, Eggertsson HP, Moore KHS, et al. The sequences of 150,119
689 genomes in the UK Biobank. *Nature*. 2022;607(7920):732-740. doi:10.1038/s41586-
690 022-04965-x

691 33. Gudbjartsson DF, Helgason H, Gudjonsson SA, et al. Large-scale whole-genome
692 sequencing of the Icelandic population. *Nat Genet*. 2015;47(5):435-444.
693 doi:10.1038/ng.3247

694 34. Khan IF, Hirata RK, Russell DW. AAV-mediated gene targeting methods for human
695 cells. *Nat Protoc*. 2011;6(4):482-501. doi:10.1038/nprot.2011.301

696 35. Aurnhammer C, Haase M, Muether N, et al. Universal real-time PCR for the
697 detection and quantification of adeno-associated virus serotype 2-derived inverted
698 terminal repeat sequences. *Hum Gene Ther Methods*. 2012;23(1):18-28.
699 doi:10.1089/hgtb.2011.034

700 36. Bak RO, Porteus MH. CRISPR-Mediated Integration of Large Gene Cassettes Using
701 AAV Donor Vectors. *Cell Rep*. 2017;20(3):750-756.
702 doi:10.1016/j.celrep.2017.06.064

703 37. Bak RO, Dever DP, Porteus MH. CRISPR/Cas9 genome editing in human
704 hematopoietic stem cells. *Nat Protoc*. 2018;13(2):358-376.
705 doi:10.1038/nprot.2017.143

706 38. Charlesworth CT, Camarena J, Cromer MK, et al. Priming Human Repopulating
707 Hematopoietic Stem and Progenitor Cells for Cas9/sgRNA Gene Targeting. *Mol*
708 *Ther Nucleic Acids*. 2018;12:89-104. doi:10.1016/j.omtn.2018.04.017

709 39. Cromer MK, Vaidyanathan S, Ryan DE, et al. Global Transcriptional Response to
710 CRISPR/Cas9-AAV6-Based Genome Editing in CD34+ Hematopoietic Stem and
711 Progenitor Cells. *Mol Ther*. 2018;26(10):2431-2442.
712 doi:10.1016/j.ymthe.2018.06.002

713 40. Hu J, Liu J, Xue F, et al. Isolation and functional characterization of human
714 erythroblasts at distinct stages: implications for understanding of normal and
715 disordered erythropoiesis in vivo. *Blood*. 2013;121(16):3246-3253.
716 doi:10.1182/blood-2013-01-476390

717

Figure 1

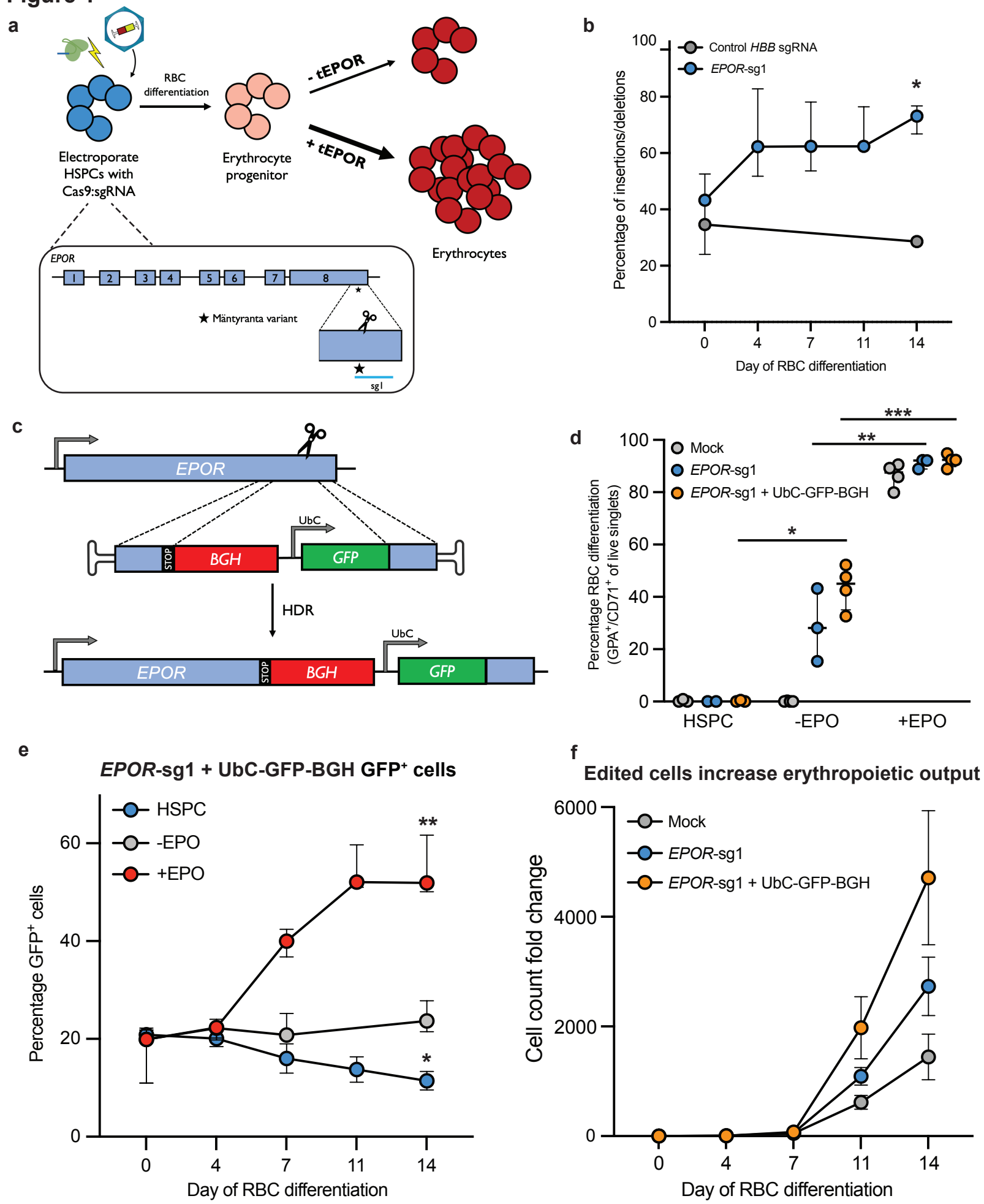

a
RBC differentiation
Electroporate HSPCs with Cas9:sgRNA
Erythrocyte progenitor
- tEPOR
+ tEPOR
Erythrocytes
EPOR
1 2 3 4 5 6 7 8
★ Mäntyranta variant
sgI
b
Control HBB sgRNA
EPOR-sg1
Percentage of insertions/deletions
Day of RBC differentiation
c
EPOR
UbC
STOP
BGH
GFP
HDR
d
Mock
EPOR-sg1
EPOR-sg1 + UbC-GFP-BGH
Percentage RBC differentiation (GPA+/CD71+ of live singlets)
HSPC
-EPO
+EPO
e
EPOR-sg1 + UbC-GFP-BGH GFP+ cells
HSPC
-EPO
+EPO
Percentage GFP+ cells
Day of RBC differentiation
f
Edited cells increase erythropoietic output
Mock
EPOR-sg1
EPOR-sg1 + UbC-GFP-BGH
Cell count fold change
Day of RBC differentiation

Figure 1: Cas9-guided *EPOR* truncation in HSPCs enhances erythroid proliferation. **a.** Schematic of HSPC editing and model of tEPOR effect. Representation of *EPOR* gene and location of the candidate sgRNA (*EPOR*-sg1) indicated by a line. Location of c.1316G>A mutation is denoted by the star. **b.** Frequency of indels created by sg1 in primary human $CD34^+$ HSPCs over the course of erythroid differentiation compared to control *HBB* sgRNA. Points represent median ± interquartile range. Values represent biologically independent HSPC donors: N=5 for *EPOR*-sg1, N=1 for control *HBB* sgRNA. *P=0.0016 of day 1 vs. day 14 by unpaired two-tailed *t*-test. **c.** Genome editing strategy when using an AAV6 DNA repair template to introduce the *EPOR* c.1316G>A mutation followed by a BGH-polyA region and UbC-driven GFP reporter. **d.** Percentage of $GPA^+/CD71^+$ of live single cells on day 14 of differentiation. Bars represent median ± interquartile range. Values represent biologically independent HSPC donors: N=2-3 for HSPC, N=3-4 for -EPO and +EPO conditions. *P=0.0016 of -EPO vs. HSPC conditions; **P=0.003 and ***P=0.0001 of -EPO vs. +EPO conditions by unpaired two-tailed *t*-test. **e.** Percentage of GFP^+ cells of live single cells maintained in RBC media +/- EPO or HSPC media as determined by flow cytometry. Points represent median ± interquartile range. Values represent biologically independent HSPC donors N=2 for HSPC condition and N=3-4 for -EPO and +EPO conditions. *P=0.04, **P=0.0006 of day 0 vs. day 14 by unpaired two-tailed *t*-test. **f.** Fold change in cell count throughout RBC differentiation (e.g., if at day 0 starting cell numbers were 1E5 cells total, therefore a fold count change of 1000 would yield a total cell number of 1E8 at day 14). Points represent mean ± SEM. Values represent biologically independent HSPC donors N=3 for Mock and *EPOR*-sg1 + BGH and N=2 for *EPOR*-sg1.

Figure 2

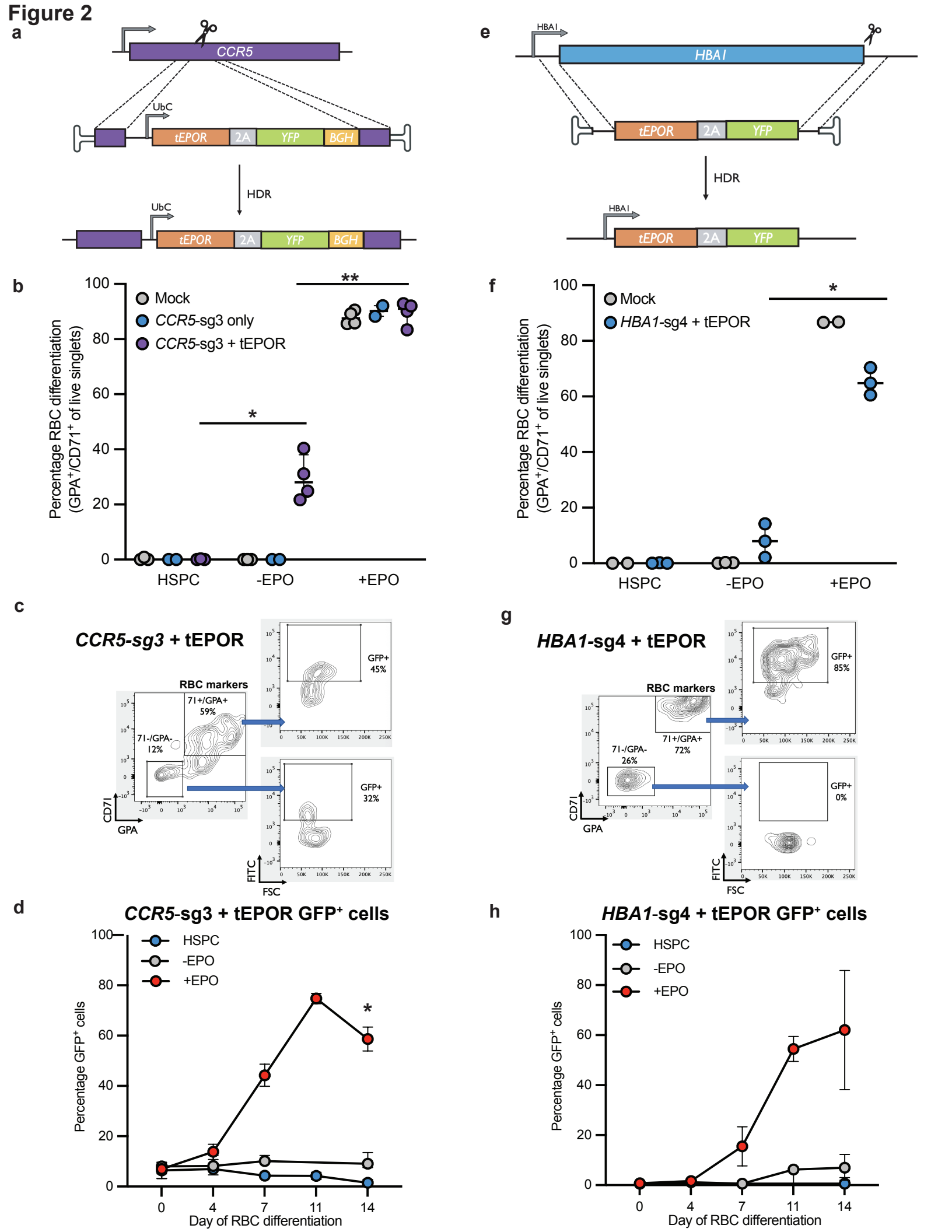

a
CCR5
UbC
tEPOR
2A
YFP
BGH
HDR
UbC
tEPOR
2A
YFP
BGH
e
HBA1
HBA1
tEPOR
2A
YFP
HDR
HBA1
tEPOR
2A
YFP
b
Mock
CCR5-sg3 only
CCR5-sg3 + tEPOR
Percentage RBC differentiation (GPA+/CD71+ of live singlets)
100
80
60
40
20
0
*
**
HSPC
-EPO
+EPO
f
Mock
HBA1-sg4 + tEPOR
Percentage RBC differentiation (GPA+/CD71+ of live singlets)
100
80
60
40
20
0
*
HSPC
-EPO
+EPO
c
CCR5-sg3 + tEPOR
RBC markers
71+/GPA+
59%
71-/GPA-
12%
CD71
GPA
GFP+
45%
GFP+
32%
FITC
FSC
g
HBA1-sg4 + tEPOR
RBC markers
71+/GPA+
72%
71-/GPA-
26%
CD71
GPA
GFP+
85%
GFP+
0%
FITC
FSC
d
CCR5-sg3 + tEPOR GFP+ cells
HSPC
-EPO
+EPO
Percentage GFP+ cells
100
80
60
40
20
0
*
0
4
7
11
14
Day of RBC differentiation
h
HBA1-sg4 + tEPOR GFP+ cells
HSPC
-EPO
+EPO
Percentage GFP+ cells
100
80
60
40
20
0
0
4
7
11
14
Day of RBC differentiation

Figure 2: Integration of *tEPOR* cDNA demonstrates an erythroid-specific proliferative effect.

a. Genome editing strategy to introduce *tEPOR*-T2A-YFP-BGH-polyA cDNA at the *CCR5* locus with expression driven by a ubiquitous UbC promoter. **b.** Percentage of $GPA^+/CD71^+$ of live single cells on day 14 of differentiation following introduction of *tEPOR* at *CCR5* locus. Bars represent median ± interquartile range. Values represent biologically independent HSPC donors: N=2-3 for HSPC condition, N=2-4 for -EPO and +EPO conditions. *P=0.0018 of -EPO vs. HSPC conditions; **P<0.0001 of -EPO vs. +EPO conditions by unpaired two-tailed *t*-test. **c.** Representative flow cytometry plots of *CCR5*-edited HSPCs following differentiation into RBCs. **d.** Percentage of GFP^+ cells of live single cells as determined by flow cytometry. Points represent mean ± SEM. Values represent biologically independent HSPC donors: N=2 for HSPC condition, N=3 for -EPO condition, N=3-4 for +EPO condition. *P = 0.0003 of day 0 vs. day 14 by unpaired two-tailed *t*-test. **e.** Genome editing strategy to introduce *tEPOR*-T2A-YFP cDNA at the *HBA1* locus by whole gene replacement to place integration cassette under regulation of the endogenous *HBA1* promoter. **f.** Percentage of $GPA^+/CD71^+$ of live single cells on day 14 of differentiation following introduction of *tEPOR* cDNA at the *HBA1* locus. Bars represent median ± 95% confidence interval. Values represent biologically independent HSPC donors: N=2 for HSPC condition, N=2-3 for -EPO and +EPO condition. *P=0.0002 of -EPO to +EPO condition by unpaired two-tailed *t*-test. **g.** Representative flow cytometry plot of *HBA1*-edited HSPCs following differentiation into RBCs. **h.** Percentage of GFP^+ cells of live single cells as determined by flow cytometry. Points represent mean ± SEM.

Values represent biologically independent HSPC donors: N=2 for HSPC condition, N=3 for -EPO and +EPO condition.

Figure 3

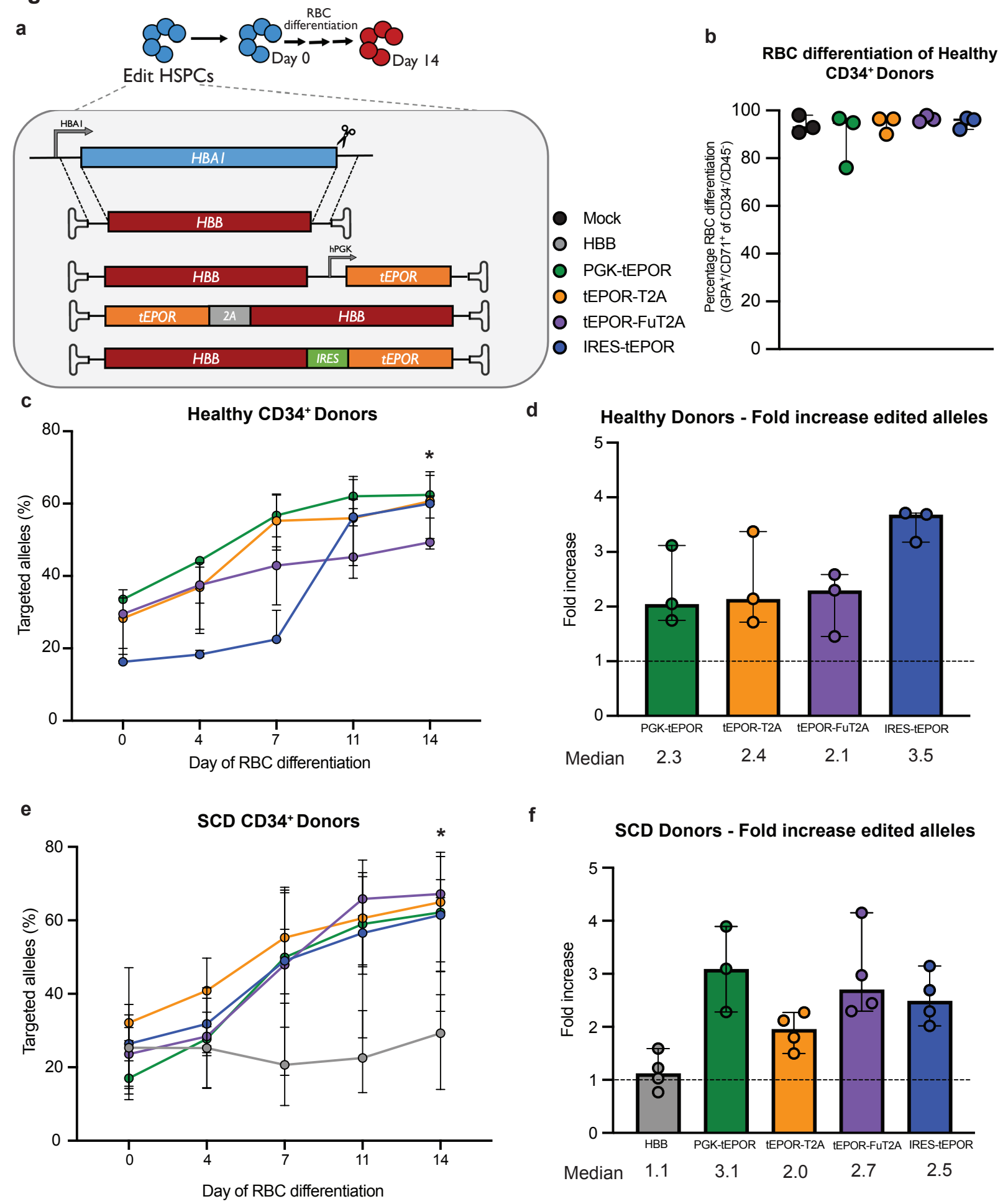

a
RBC differentiation
Edit HSPCs
Day 0
Day 14
HBA1
HBA1
HBB
hPGK
HBB
tEPOR
tEPOR
2A
HBB
HBB
IRES
tEPOR
Mock
HBB
PGK-tEPOR
tEPOR-T2A
tEPOR-FuT2A
IRES-tEPOR
b
RBC differentiation of Healthy CD34+ Donors
Percentage RBC differentiation (GPA+/CD71+ of CD34-/CD45-)
100
80
60
40
20
0
c
Healthy CD34+ Donors
Targeted alleles (%)
Day of RBC differentiation
0
4
7
11
14
*
d
Healthy Donors - Fold increase edited alleles
Fold increase
PGK-tEPOR
tEPOR-T2A
tEPOR-FuT2A
IRES-tEPOR
Median
2.3
2.4
2.1
3.5
e
SCD CD34+ Donors
Targeted alleles (%)
Day of RBC differentiation
*
f
SCD Donors - Fold increase edited alleles
Fold increase
HBB
PGK-tEPOR
tEPOR-T2A
tEPOR-FuT2A
IRES-tEPOR
Median
1.1
3.1
2.0
2.7
2.5

Figure 3: Therapeutic editing frequencies are achieved using bicistronic *HBB-tEPOR* cassette.

a. Design of *HBB* (control) and *HBB-tEPOR* (bicistronic) AAV6 donor cassettes targeted to the *HBA1* locus by whole gene replacement. **b.** Percentage of $GPA^+/CD71^+$ of $CD34^-/CD45^-$ cells on day 14 as determined by flow cytometry. Points shown as median ± 95% confidence interval. Values represent N=3 biologically independent HSPC donors. **c.** Percentage of edited alleles for each bicistronic *tEPOR* vector in cord-blood derived $CD34^+$ cells over the course of RBC differentiation. Points shown as median ± 95% confidence interval. Values represent N=3 biologically independent HSPC donors. *****P=0.003 for PGK-*tEPOR*, P=0.0055 for T2A-*tEPOR*, P=0.0259 for Fu2A-*tEPOR*, and P=0.0003 for IRES-*tEPOR* (day 0 vs. day 14) by unpaired two-tailed *t*-test. **d.** Fold change in edited alleles from beginning (day 0) to end (day 14) of RBC differentiation. Bars represent median ± 95% confidence interval. **e.** Percentage of edited alleles for control (*HBB*) and bicistronic *HBB-tEPOR* vectors in sickle cell disease patient cells over the course of RBC differentiation. Points shown as median ± 95% confidence interval. Values represent biologically independent HSPC donors: N=3 for PGK-*tEPOR*, N=4 for all other vectors. *****P=0.0061 for PGK-*tEPOR*, P=0.011 for T2A-*tEPOR*, P=0.0016 for Fu2A-*tEPOR*, P=0.0153 for IRES-*tEPOR* (day 0 vs. day 14) by unpaired two-tailed *t*-test. **f.** Fold change in edited alleles from beginning (day 0) to end (day 14) of RBC differentiation. Bars represent median ± standard deviation.

Figure 4

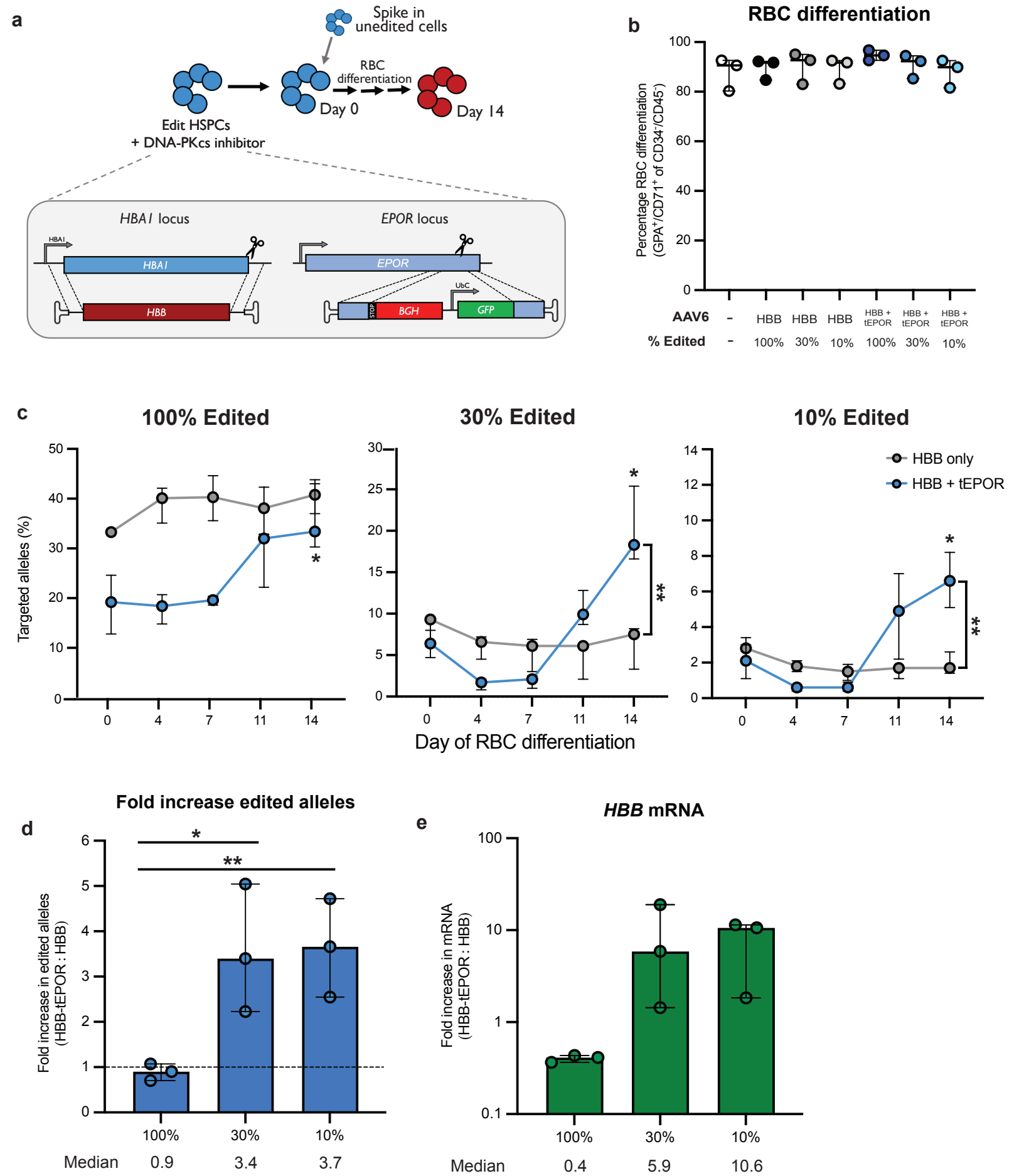
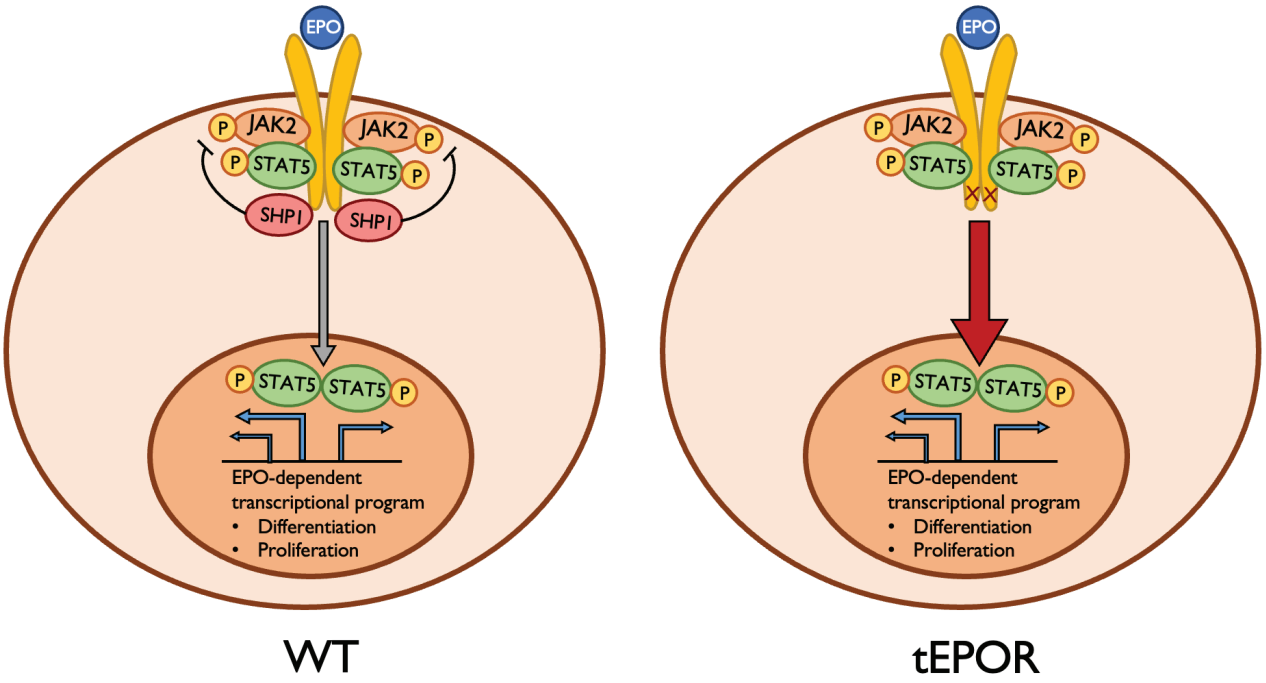

a
Spike in unedited cells
RBC differentiation
Day 0
Day 14
Edit HSPCs
+ DNA-PKcs inhibitor
HBA1 locus
EPOR locus
HBA1
HBB
EPOR
UbC
STOP
BGH
GFP
b
RBC differentiation
Percentage RBC differentiation (GPA+/CD71+ of CD34-/CD45-)
AAV6
% Edited
HBB
HBB + tEPOR
100%
30%
10%
c
100% Edited
30% Edited
10% Edited
Targeted alleles (%)
HBB only
HBB + tEPOR
Day of RBC differentiation
d
Fold increase edited alleles
Fold increase in edited alleles (HBB-tEPOR : HBB)
Median
0.9
3.4
3.7
e
HBB mRNA
Fold increase in mRNA (HBB-tEPOR : HBB)
Median
0.4
5.9
10.6

Figure 4: Multiplexed editing of *EPOR* and *HBA1* leads to robust increase in *HBB* mRNA within editing HSPCs.

a. Schematic of multiplexed editing strategy with spike-in of unedited cells at start of erythroid differentiation to model HSPC transplantation. **b.** Percentage of $GPA^+/CD71^+$ of $CD34^-/CD45^-$ cells on day 14 of RBC differentiation as determined by flow cytometry. Points shown as median ± interquartile range. N=3 biologically independent HSPC donors. **c.** Percentage of edited alleles at *HBA1* in all multiplexed editing/spike-in conditions throughout RBC differentiation. Points shown as median ± 95% confidence interval. N=3 biologically independent HSPC donors. *P=0.0332 for *HBB*+*tEPOR* 100%, P=0.0086 for *HBB*+*tEPOR* 30%, P=0.0122 for *HBB*+*tEPOR* 10% (day 0 vs. day 14) by unpaired two-tailed *t*-test. **P=0.0113 for *HBB* versus *HBB*+*tEPOR* 30% at day 14 and P=0.008 for HBB versus *HBB*+*tEPOR* 10% at day 14 by unpaired two-tailed *t*-test. **d.** Fold increase in edited alleles on day 14 of differentiation of multiplexed conditions versus *HBB* only. Bars represent median ± 95% confidence interval. *P=0.0315 and **P=0.0123 by unpaired two-tailed *t*-test. **e.** mRNA expression of integrated *HBB* at *HBA1* locus normalized to *HBB* expression from mock. *GPA* mRNA expression was used as a reference. N=3 biologically independent HSPC donors. Bars represent median ± 95% confidence interval.

Extended Data Figure 1

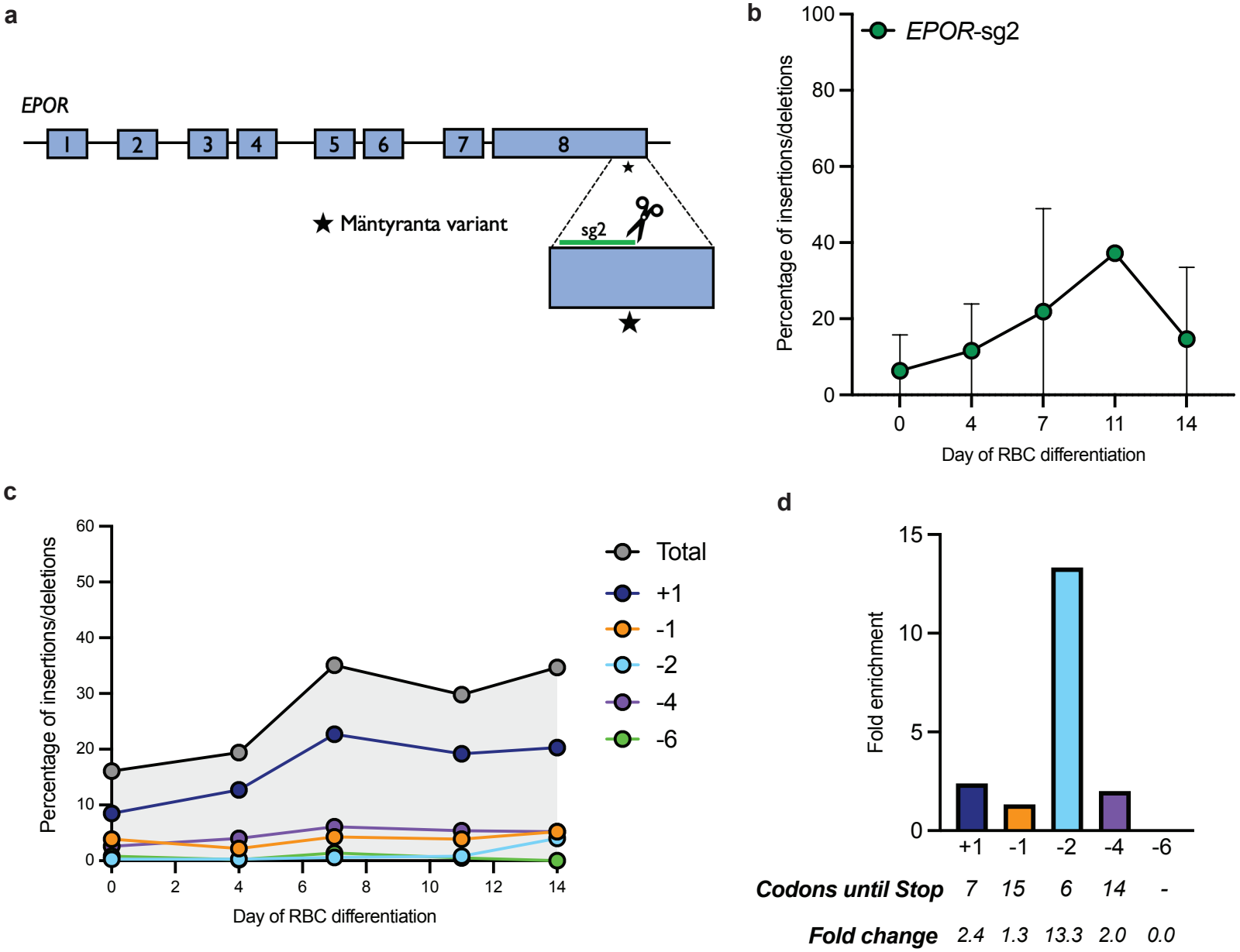
a



Extended Data Figure 1. Schematic of wild-type (WT) and tEPOR signaling cascade.

a. Visual representation of EPOR signaling cascade. SHP1 binds to intracellular inhibitory domain in WT cells, downregulating EPOR signaling. This domain is truncated in tEPOR cells making them hypersensitive to EPO.

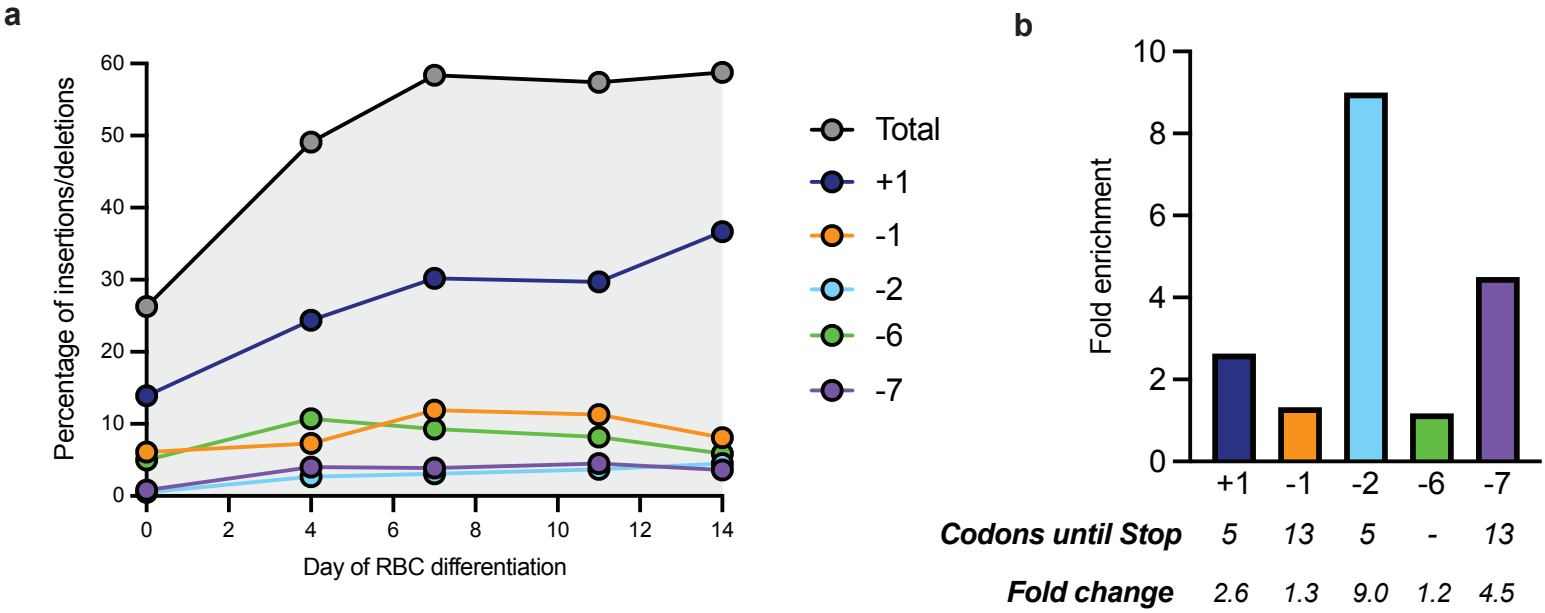
Extended Data Figure 2



Extended Data Figure 2. Indel analysis of HSPCs edited with second candidate *EPOR* sgRNA (sg2) over course of erythroid differentiation.

a. Schematic of *EPOR* gene and location of the second candidate sgRNA (*EPOR*-sg2) indicated by a line. Location of c.1316G>A mutation is denoted by the star **b.** Frequency of indels created by *EPOR*-sg2 in primary human $CD34^+$ HSPCs over the course of erythroid differentiation. Points represent median ± interquartile range. Values represent N=4 biologically independent HSPC donors. **c.** Frequency of five most common indels found in one HSPC donor targeted with *EPOR*-sg2 over the course of RBC differentiation. **d.** Fold enrichment of five most common indels over course of RBC differentiation in one HSPC donor targeted with *EPOR*-sg2.

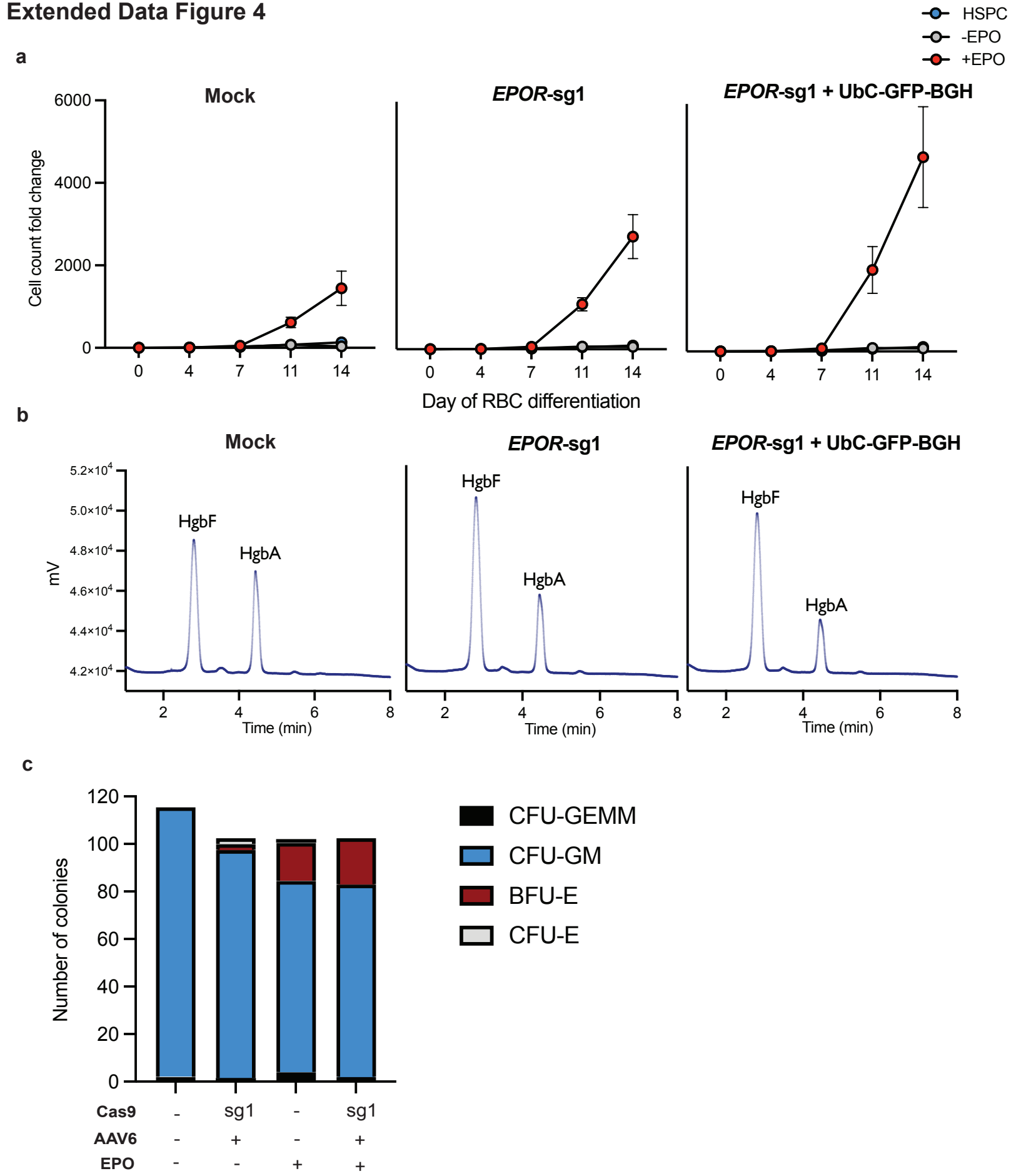
Extended Data Figure 3



Extended Data Figure 3. Indel analysis of HSPCs edited with most effective *EPOR* sgRNA (sg1) over course of erythroid differentiation.

a. Frequency of five most common indels found in one HSPC donor targeted with *EPOR*-sg1over the course of RBC differentiation. **b.** Fold enrichment of five most common indels over course of RBC differentiation in one HSPC donor targeted with *EPOR*-sg1.

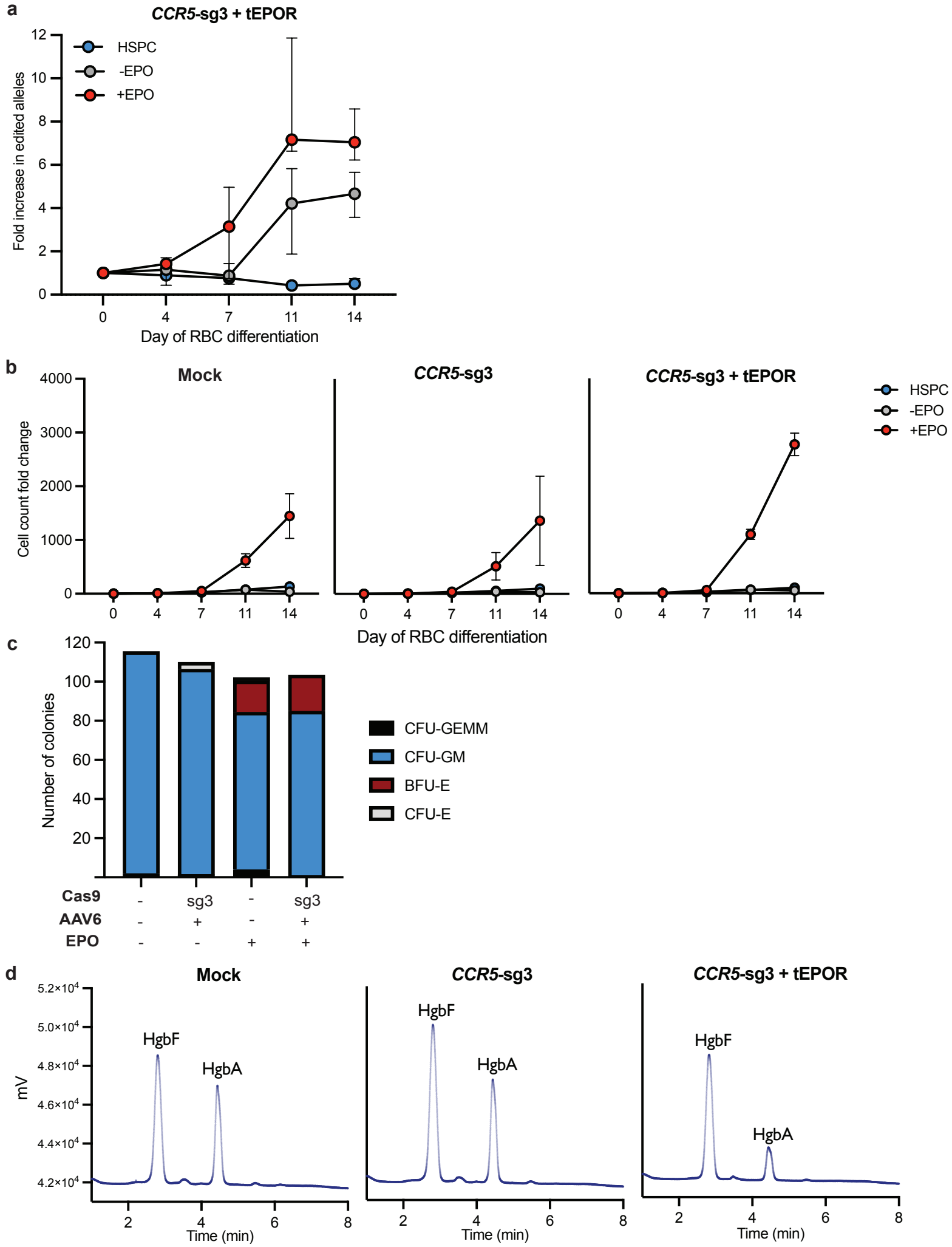
Extended Data Figure 4



Extended Data Figure 4. HSPCs edited at *EPOR* locus show increased erythroid proliferation and maintain lineage formation.

a. Cell count fold change in mock and edited cells maintained in RBC media +/- EPO or HSPC media. Points represent mean ± SEM. Values represent biologically independent HSPC donors N=3 for Mock and *EPOR*-sg1+ BGH and N=2 for sg1. **b.** Representative HPLC plots of cells targeted with *EPOR*-sg1at day 14 of RBC differentiation. HgbF=fetal hemoglobin, HgbA=adult hemoglobin. **c.** Colony formatting unit (CFU) assay of mock edited cells versus cells edited with *EPOR*-sg1 + BGH. Bars represent total number of colonies of each type: CFU-GEMM (multi-potential granulocyte, erythroid, macrophage, megakaryocyte progenitor cells), CFU-GM (colony forming unit-granulocytes and monocytes), BFU-E (erythroid burst forming units), CFU-E (colony forming unit-erythroid) colonies. N=1.

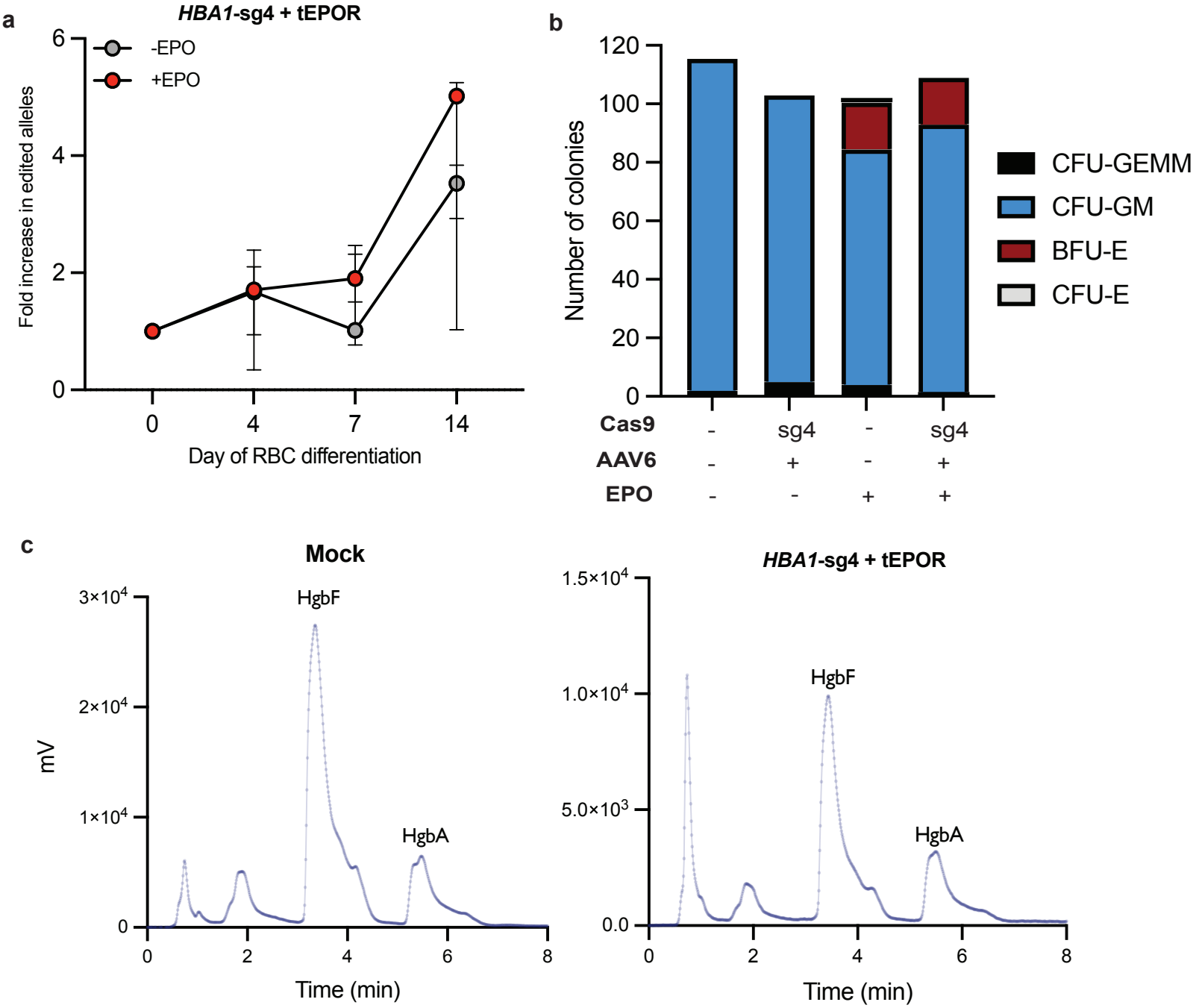
Extended Data Figure 5



Extended Data Figure 5. Safe harbor integration of *tEPOR* leads to increased erythroid proliferation and editing frequencies without affecting HSPC lineage or hemoglobin formation.

a. Fold change in edited alleles at *CCR5* over course of RBC differentiation +/- EPO or maintained in HSPC media measured by ddPCR. Points shown as median ± interquartile range. N=3-4 biologically independent HSPC donors. **b.** Cell count fold change in mock and edited cells maintained in RBC media +/- EPO or HSPC media. Points represent mean ± SEM. Values represent biologically independent HSPC donors N=3 for Mock and *CCR5*-sg3 + *tEPOR* and N=2 for *CCR5*-sg3. **c.** CFU assay of mock edited cells versus cells edited with *CCR5*-sg3 + *tEPOR.* Bars represent total number of colonies of each type: CFU-GEMM (multi-potential granulocyte, erythroid, macrophage, megakaryocyte progenitor cells), CFU-GM (colony forming unit-granulocytes and monocytes), BFU-E (erythroid burst forming units), CFU-E (colony forming unit-erythroid) colonies. N=1. Mock condition from Extended Data 3c shown here for comparison. **d.** Representative HPLC plots of cells targeted with *CCR5*-sg3 or *CCR5*-sg3 + *tEPOR* at day 14 of RBC differentiation. HgbF=fetal hemoglobin, HgbA=adult hemoglobin.

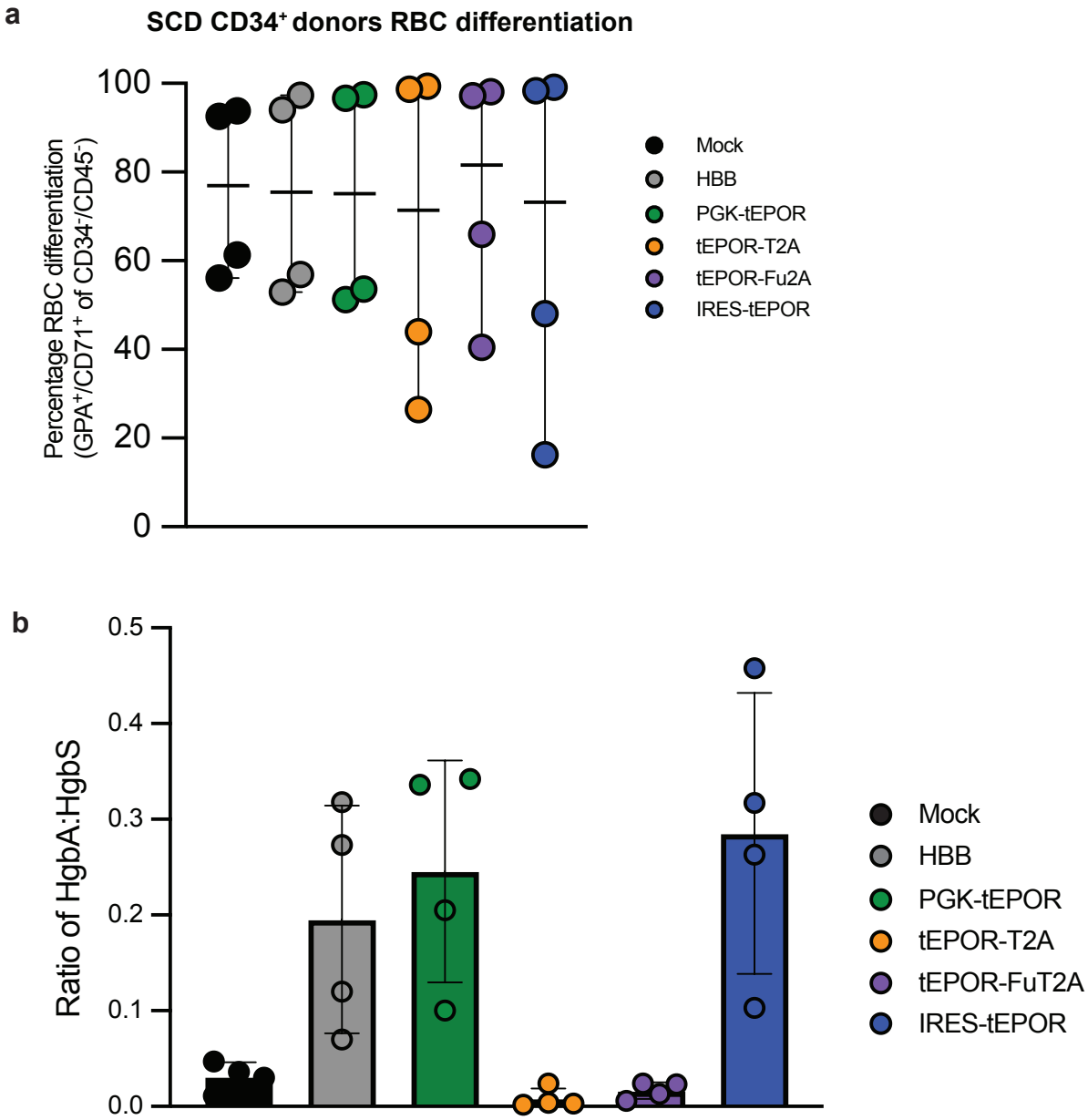
Extended Data Figure 6



Extended Data Figure 6. Erythroid specific expression of *tEPOR* leads to increased editing frequencies without affecting HSPC lineage or hemoglobin formation.

a. Fold change in edited alleles at *HBA1* over course of RBC differentiation +/- EPO measured by ddPCR. Points shown as median ± 95% CI. N=3 biologically independent HSPC donors. **b.** CFU assay of mock edited cells versus cells edited with *HBA1*-sg4 + *tEPOR.* Bars represent total number of colonies of each type: CFU-GEMM (multi-potential granulocyte, erythroid, macrophage, megakaryocyte progenitor cells), CFU-GM (colony forming unit-granulocytes and monocytes), BFU-E (erythroid burst forming units), CFU-E (colony forming unit-erythroid) colonies. N=1. Mock condition from Extended Data 3c shown here for comparison. **c.** Representative HPLC plots of cells targeted with *HBA1*-sg4 + *tEPOR* at day 14 of RBC differentiation. HgbF=fetal hemoglobin, HgbA=adult hemoglobin.

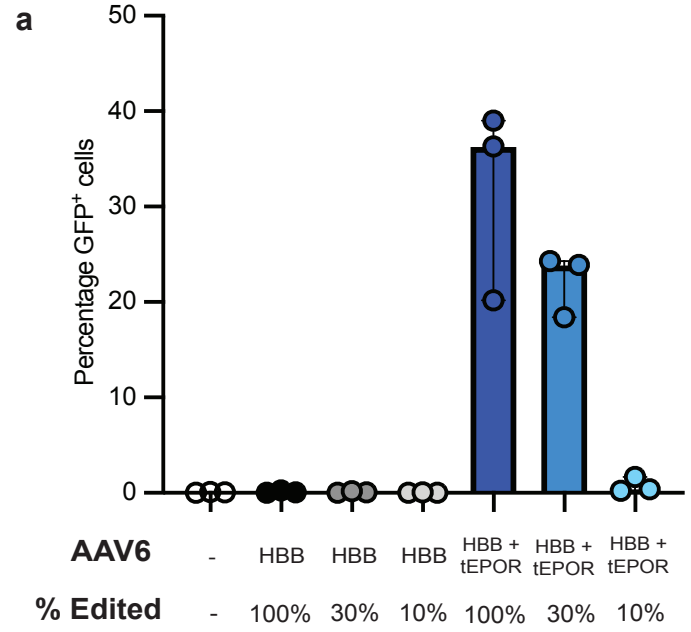
Extended Data Figure 7

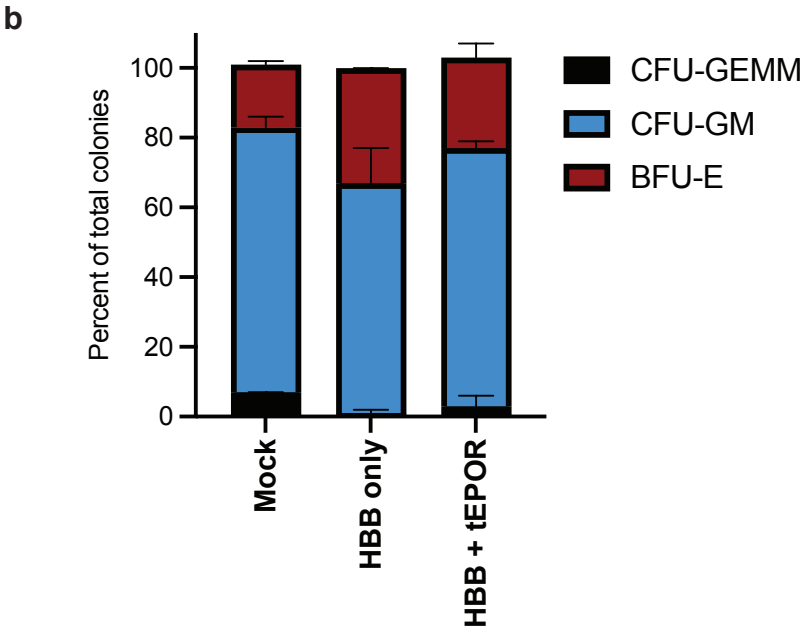


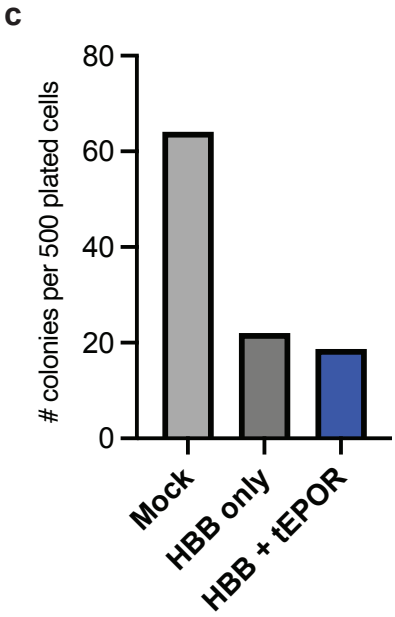
Extended Data Figure 7. Hemoglobin tetramer analysis of *tEPOR* edited SCD patient HSPCs following RBC differentiation.

a. Percentage of $GPA^+/CD71^+$ of $CD34^-/CD45^-$ cells on day 14 of RBC differentiation as determined by flow cytometry. Points shown as median ± 95% confidence interval. N=4 biologically independent HSPC donors. **b.** Ratio of adult hemoglobin to sickle hemoglobin from HPLC analysis of hemoglobin tetramers from differentiated SCD patient HSPCs. Bars shown as mean ± SD. N=4 biologically independent HSPC donors.

Extended Data Figure 8







Extended Data Figure 8. Multiplexed edited strategy shows maintenance of *EPOR* truncation at day 14 of RBC differentiation and does not disrupt HSPC lineage formation.

a. Percentage of GFP^+ cells of live single cells on day 14 of RBC differentiation as determined by flow cytometry. Points represent median ± interquartile range. Values represent N=3 biologically independent HSPC donors. **b.** CFU assay of mock, single edited, and multiplex edited HSPCs. Bars represent percent of total colonies: CFU-GEMM (multi-potential granulocyte, erythroid, macrophage, megakaryocyte progenitor cells), CFU-GM (colony forming unit-granulocytes and monocytes), and BFU-E (erythroid burst forming units). N=1. **c.** Number of total colonies in CFU assay produced per 500 plated cells in each condition.