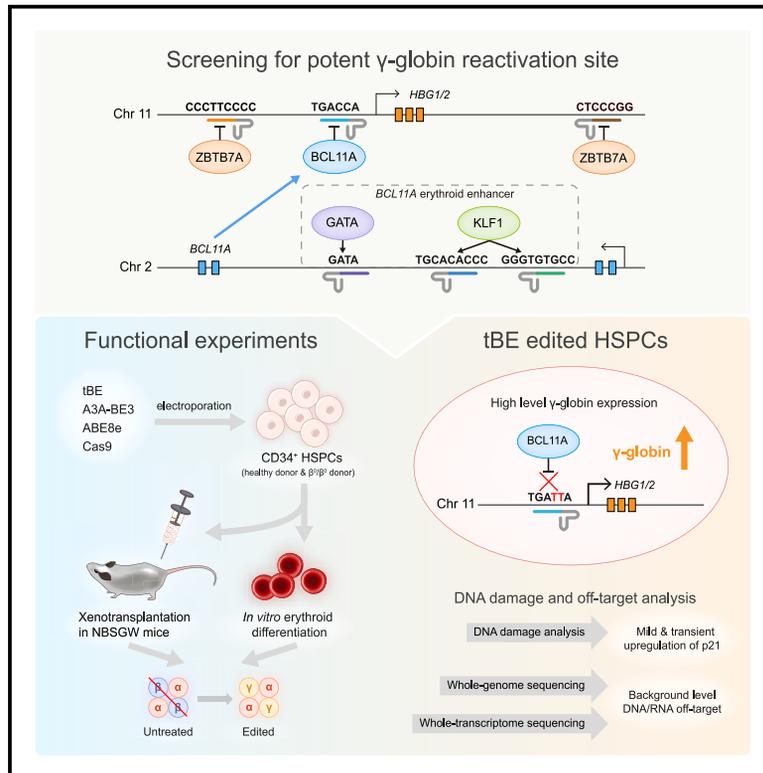


Base editing of the HBG promoter induces potent fetal hemoglobin expression with no detectable off-target mutations in human HSCs

Graphical abstract



Authors

Wenyan Han, Hou-Yuan Qiu, Shangwu Sun, ..., Jia Chen, Bei Yang, Ying Zhang

Correspondence

liyang_fudan@fudan.edu.cn (L.Y.), chenjia@shanghaitech.edu.cn (J.C.), yangbei@shanghaitech.edu.cn (B.Y.), ying.zhang84@whu.edu.cn (Y.Z.)

In brief

Han et al. test gene editing approaches for inducing γ -globin expression in HSPCs as a therapeutic strategy for β -hemoglobinopathies. They find that transformer base editor-mediated disruption of the BCL11A binding motif at *HBG1/2* promoter triggered the highest fetal hemoglobin levels in healthy and patient HSPCs with no detectable off-target mutations.

Highlights

- Enhanced γ -globin level via tBE-edited BCL11A-binding motif at *HBG1/2* promoter
- tBE-edited HSCs had repopulation ability and potent efficacy in patient cells
- tBE-edited HSCs displayed the background level of DNA/RNA off-target activities

Article

Base editing of the HBG promoter induces potent fetal hemoglobin expression with no detectable off-target mutations in human HSCs

Wenyan Han,^{1,13} Hou-Yuan Qiu,^{2,13} Shangwu Sun,^{1,3,13} Zhi-Can Fu,^{4,5,13} Guo-Quan Wang,^{2,13} Xiaowen Qian,^{6,13} Lijie Wang,⁷ Xiaowen Zhai,⁶ Jia Wei,⁵ Yichuan Wang,⁷ Yi-Lin Guo,⁵ Guo-Hua Cao,² Rui-Jin Ji,² Yi-Zhou Zhang,² Hongxia Ma,⁷ Hongsheng Wang,⁶ Mingli Zhao,¹ Jing Wu,¹ Lili Bi,² Qiu-Bing Chen,² Zifeng Li,⁶ Ling Yu,⁶ Xiaodun Mou,⁷ Hao Yin,^{2,8,9,12} Li Yang,^{5,*} Jia Chen,^{1,10,11,*} Bei Yang,^{1,3,10,11,*} and Ying Zhang^{2,12,14,*}

¹Gene Editing Center, School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China

²Department of Rheumatology and Immunology, Medical Research Institute, Frontier Science Center for Immunology and Metabolism, Zhongnan Hospital of Wuhan University, Wuhan University, Wuhan 430071, China

³Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China

⁴Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China

⁵Center for Molecular Medicine, Children's Hospital of Fudan University and Shanghai Key Laboratory of Medical Epigenetics, International Laboratory of Medical Epigenetics and Metabolism, Ministry of Science and Technology, Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China

⁶Department of Hematology and Oncology, Children's Hospital of Fudan University, Shanghai 201102, China

⁷CorrectSequence Therapeutics, Shanghai 201210, China

⁸Department of Pathology and Department of Urology, Zhongnan Hospital of Wuhan University, Wuhan 430071, China

⁹TaiKang Centre for Life and Medical Sciences, TaiKang Medical School, Wuhan University, Wuhan, China

¹⁰Shanghai Clinical Research and Trial Center, Shanghai 201210, China

¹¹Shanghai Frontiers Science Center for Biomacromolecules and Precision Medicine, ShanghaiTech University, Shanghai 200031, China

¹²State Key Laboratory of Virology, Wuhan University, Wuhan, China

¹³These authors contributed equally

¹⁴Lead contact

*Correspondence: liyang_fudan@fudan.edu.cn (L.Y.), chenjia@shanghaitech.edu.cn (J.C.), yangbei@shanghaitech.edu.cn (B.Y.),

ying.zhang84@whu.edu.cn (Y.Z.)

<https://doi.org/10.1016/j.stem.2023.10.007>

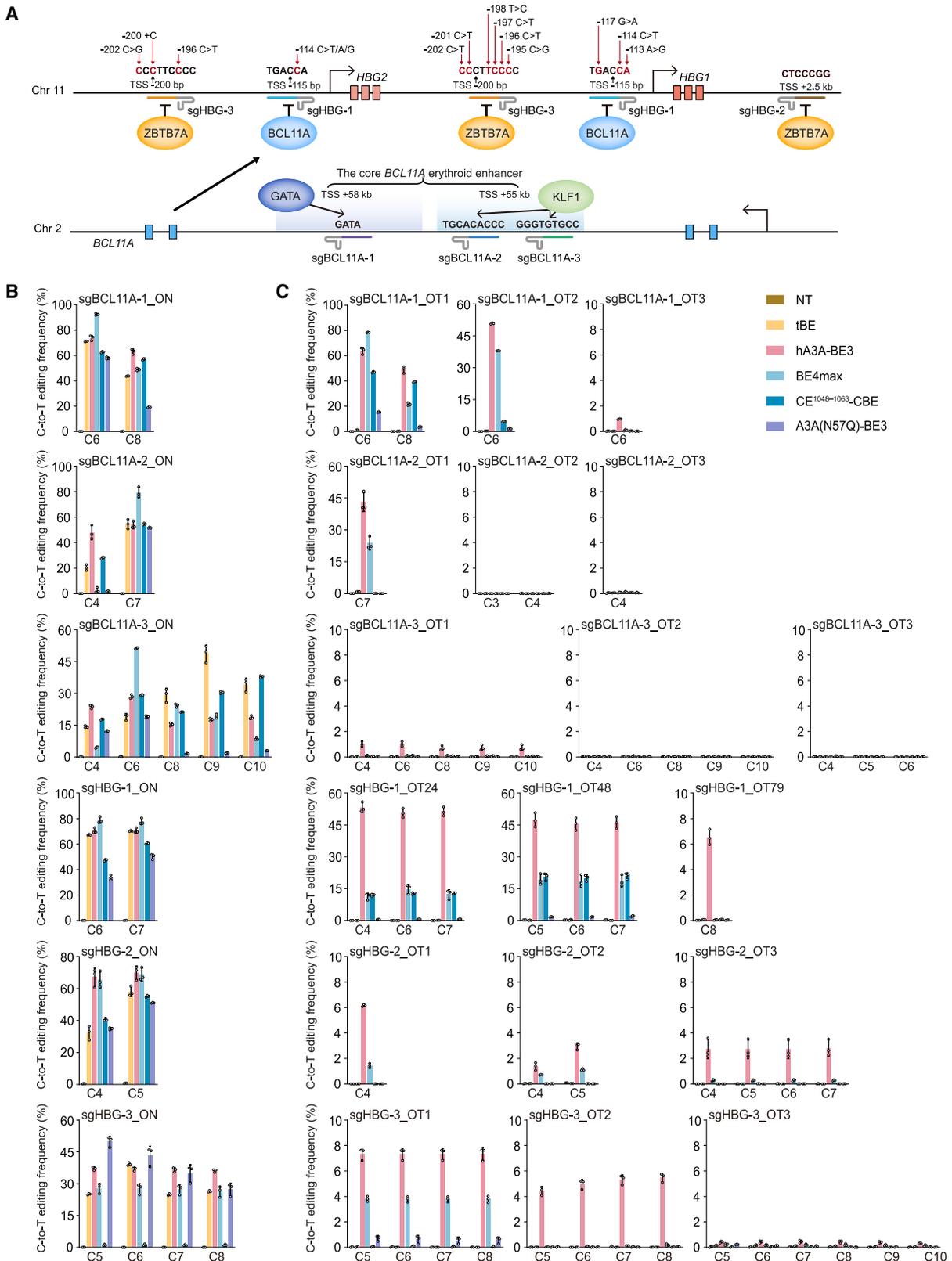
SUMMARY

Reactivating silenced γ -globin expression through the disruption of repressive regulatory domains offers a therapeutic strategy for treating β -hemoglobinopathies. Here, we used transformer base editor (tBE), a recently developed cytosine base editor with no detectable off-target mutations, to disrupt transcription-factor-binding motifs in hematopoietic stem cells. By performing functional screening of six motifs with tBE, we found that directly disrupting the BCL11A-binding motif in *HBG1/2* promoters triggered the highest γ -globin expression. Via a side-by-side comparison with other clinical and preclinical strategies using Cas9 nuclease or conventional BEs (ABE8e and hA3A-BE3), we found that tBE-mediated disruption of the BCL11A-binding motif at the *HBG1/2* promoters triggered the highest fetal hemoglobin in healthy and β -thalassemia patient hematopoietic stem/progenitor cells while exhibiting no detectable DNA or RNA off-target mutations. Durable therapeutic editing by tBE persisted in repopulating hematopoietic stem cells, demonstrating that tBE-mediated editing in *HBG1/2* promoters is a safe and effective strategy for treating β -hemoglobinopathies.

INTRODUCTION

β -hemoglobinopathies, including β -thalassemia and sickle cell disease (SCD), are the common genetic diseases caused by mutations in the hemoglobin subunit beta (*HBB*) gene locus. Mutations in the *HBB* locus lead to impaired β -globin production, which can be compensated via direct restoration of the pathogenic mutations in the *HBB* locus or reactivation of γ -globin expression.^{1–3} Although direct correction of pathogenic muta-

tions requires individual drug development for each mutation, gene replacement of α -globin with β -globin can be an alternative. However, due to the low efficiency of large gene replacement,⁴ reactivation of γ -globin expression is considered a universal therapeutic strategy that covers β -hemoglobinopathy patients.^{5–7} γ -Globin expression is tightly regulated throughout development and is silenced shortly after birth by repressors that bind to its promoter regions.⁸ Genome-wide association studies (GWASs) have identified several transcription repressors



(legend on next page)

and their binding motifs, disruption of which has shown elevated γ -globin expression.^{9,10} BCL11 transcription factor A (BCL11A) and zinc finger and BTB (Broad-Complex, Tramtrack, and Bric a brac) domain containing 7A (ZBTB7A, also known as leukemia/lymphoma-related factor [LRF]) are the two major repressors of γ -globin gene expression, responsible for its silencing.^{11–18} Knocking out the BCL11A or ZBTB7A proteins is not tolerated due to their crucial roles in other biological processes.^{19–21} Alternatively, disrupting the erythroid-specific *BCL11A* enhancer located on chromosome 2 greatly reduces BCL11A expression in erythroid cells without affecting other lineage development.²² The BCL11A enhancer has two binding motifs that can be recognized by GATA binding protein (GATA) or Kruppel like factor 1 (KLF1) transcription factors, and mutating either motifs could downregulate BCL11A expression.¹³ Mutating the BCL11A or ZBTB7A binding motif located in the regulatory region of the *HBG* (Hemoglobin subunit gamma) locus is another way to reactivate γ -globin expression.^{23,24} One key question associated with γ -globin reactivation strategy is how much HbF ($\alpha 2\gamma 2$, fetal hemoglobin) is required to achieve therapeutic benefit. Studies of β -thalassemia patients with naturally reactivated γ -globin have indicated that the HbF level correlates negatively with the number of morbidities.^{25,26} Therefore, it is important to evaluate the efficacy of different reactivation strategies, ensuring a potent protective level of HbF.

Hematopoietic stem and progenitor cells (HSPCs) are multilineage precursor cells that can self-renew and reconstitute the entire blood system. Establishing a safe modification of HSPCs is particularly important because edited cells can persist throughout a lifetime to regenerate the blood systems. HSPCs are sensitive to genomic damages, such as DNA double-strand breaks (DSBs).^{27–30} The application of Cas9 nuclease to disturb repressor binding motifs in HSPCs relies on generating DSBs,^{5,22} thereby posing a risk of p53-dependent DNA damage response and cell toxicity in HSPCs. In the *HBG* locus, owing to gene duplication of *HBG1* and *HBG2*, Cas9 nuclease-mediated editing often leads to a complex mixture of genomic deletions and insertions or deletions (indels).³¹ Cytosine or adenine base editors (CBEs or ABEs) can convert C-to-T or A-to-G base changes without generating DSBs.^{32–35} Previous studies using base editors targeting regulatory motifs successfully induced γ -globin expression.^{23,24} However, when analyzing off-target (OT) events, both CBE or ABE triggered a high level of OT activity,^{36–40} raising strong safety concerns. Hence, the development of a safe and efficacious treatment strategy for β -hemoglobinopathies is highly desirable.

The transformer base editor (tBE), whose activity is tightly regulated through a controlled-release deaminase inhibitor, was recently developed to significantly reduce OT activity.⁴¹ In tBE, the cytidine deaminase is fused with a cleavable deoxycytidine deaminase inhibitor. By introducing a truncated helper single guide RNA (hsgRNA) designed to target a nearby

sequence of on-target site, the hsgRNA localizes the protease in close proximity to the on-target site. This, in turn, cleaves the deaminase inhibitor, allowing on-target editing while significantly reducing OT editing.⁴¹ Here, we evaluated the feasibility of tBE editing in HSPCs for HbF induction. By comparing the aforementioned regulatory motifs, we found that tBE editing of the BCL11A binding motif at transcription start site (TSS) –114/115 within the *HBG1/2* promoters region triggered the highest level of γ -globin expression. More importantly, through direct comparisons of tBE-mediated editing with other clinical or preclinical editing strategies using Cas9 nuclease, hA3A-CBEs, or ABE8e, we observed that tBE-modified BCL11A binding motif (TGAC_{–115}C_{–114}A to TGAT_{–115}T_{–114}A) induced a higher HbF level than the others, and the editing persisted in the repopulated stem cells. *In vitro* and cellular binding assays revealed that the tBE-generated TGATTA motif completely abolished its interaction with BCL11A. In comparison, biochemical analysis of BCL11A binding affinity against the mutated motifs showed that ABE8e-generated TGGCCG motif only partially reduced BCL11A binding, and the residual binding could still contribute to the inhibition of γ -globin expression. Lastly, we examined gRNA-dependent and gRNA-independent DNA and RNA OT profiles and found no detectable OT mutations in tBE-edited cells. Collectively, our study presents a proof-of-concept for a potentially safer and more potent treatment strategy for β -hemoglobinopathies.

RESULTS

Design of tBE gRNAs for efficient base editing across six different regulatory motifs

γ -Globin expression is a highly orchestrated and complicated process, involving *cis*- or *trans*-regulatory motifs that precisely control its expression.^{42–44} To activate γ -globin expression, several preclinical and clinical studies have been carried out by either disrupting the erythroid enhancer of BCL11A,²² a transcription repressor for *HBG1/2*, or the repressor binding sites on the promoter region of *HBG1/2*.^{31,45} In the *BCL11A* enhancer region located on chromosome 2, there are three transcription-factor-binding motifs^{11,13} (two motifs for KLF1 and one motif for GATA) (TSS +55 kb and TSS +58 kb; **Figure 1A**), the binding of which could substantially increase BCL11A expression in erythroid lineage. In the *HBG1/2* locus, there are three motifs located in the promoter region or in the 3' enhancer of *HBG1* (TSS –200 bp, TSS –115 bp, and TSS +2.5 kb; **Figure 1A**),^{12,13,15} which can be recognized by BCL11A or ZBTB7A.

Initially, we designed sgRNAs (single guide RNAs) targeting these six regulatory motifs and screened for the optimal hsgRNA in 293FT cells⁴⁶ (**Figures S1A** and **S1B**). For each sgRNA, three hsgRNAs were tested based on C-to-T editing efficiency. The optimal sgRNA-hsgRNA pair was further compared with other CBEs,^{47–51} the latter of which have been engineered to have

Figure 1. Base editing at transcription-factor-binding motifs associated with γ -globin regulation

(A) Schematic representation of the transcription-factor-binding motifs associated with γ -globin regulation in chromosomes 2 and 11. The sgRNAs designed for each motif are indicated below. TSS, transcription start site. Naturally occurring mutations in the HBG promoter are indicated with red arrows.

(B) C-to-T editing efficiency of different base editors at the sgRNA on-target (ON) sites shown in (A) in 293FT cells. NT, no treatment control.

(C) Off-target (OT) editing efficiency induced by different base editors with the indicated sgRNAs at three gRNA-dependent OT sites in 293FT cells.

Data represent mean \pm SD. n = 3.

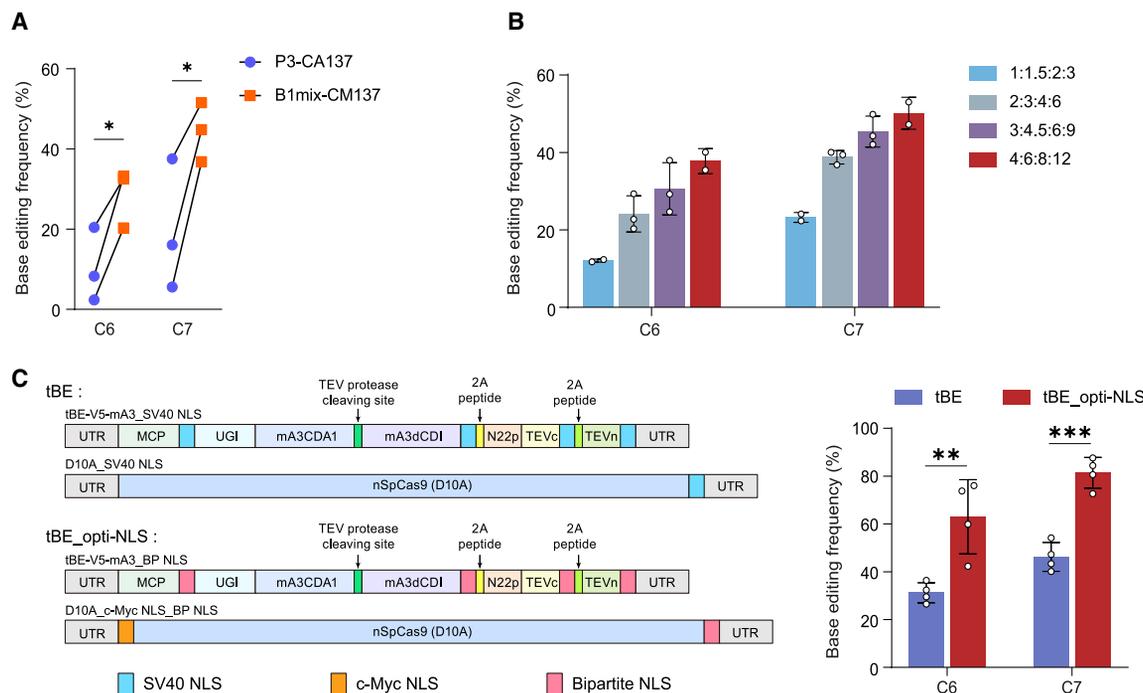


Figure 2. Optimization of RNA delivery of tBE

(A) Optimization of the electroporation conditions for RNA delivery of tBE in HSPCs.

(B) Comparison of the ratio and dose of each RNA component of tBE. The relative ratios of RNA components are represented in keys from left to right: nCas9 (D10A) mRNA, tBE-V5-mA3 mRNA, sgRNA, hsgRNA.

(C) (Left) Schematics of the original and optimized mRNA constructs of tBE. BP NLS, bipartite nuclear localization signal. (Right) Base-editing efficiency (sum of C-to-T, C-to-G, and C-to-A editing) in HSPCs induced by the original and NLS-optimized tBEs.

Data represent mean \pm SD. $n = 3$. In (A) and (C), p values were calculated by t test; * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

robust editing potency or reduced OT activities. When examining on-target activity, despite the tBE group having slightly less expression than the other editors, its editing efficiency showed a comparable level to other CBEs across six regulatory motifs (Figures 1B and S1C). The *HBG1* and *HBG2* genes share high homology with more than 96% sequence similarity with each other. The precise editing efficiency at individual *HBG1* or *HBG2* locus was very similar to each other, indicating unbiased editing at both loci (Figure S1D). When comparing gRNA-dependent OT editing with other CBEs, we found tBE displayed no detectable OT mutations across all six sgRNAs. By contrast, other base editors, i.e., hA3A-BE3, BE4max, CE¹⁰⁴⁸⁻¹⁰⁶³-CBE, or A3A(N57Q)-BE3, using the same sgRNA, all induced multiple OT editing with high frequency (up to 78.8%) (Figure 1C). A close examination of hsgRNA binding sites revealed that only 5% of editing events occurred with two out of six hsgRNAs, namely hsgHBG-1_1 or hsgBCL11A-3_2, respectively (Figure S1E), which were not reported to be within any functional elements, to our knowledge.

RNA delivery of tBE

To develop a clinically relevant system for delivering the tBE machinery in HSPCs without enrichment, we first codon-optimized the RNA expression construct and used uridine depletion to increase base-editing efficiency while reducing genome toxicity. Previously, we found that the osmolality of the electroporation buffer could greatly affect transfection efficiency and cell health.⁵²

The application of isotonic buffer, referred to as B1mix, resulted in enhanced mRNA delivery in primary cells.⁵³ Consistently, when applied in tBE-mediated sgHBG-1 editing, we observed that the isotonic B1mix buffer significantly improved editing efficiency by 2.8- and 2.3-fold at C6 and C7 compared with the hypertonic commercial P3 buffer (Figure 2A). Increasing the mRNA dosage of tBE and its associated sgRNA/hsgRNA ratios could also enhance editing efficiency (Figure 2B). The tBE system is composed of two messenger RNAs, one encoding a SpCas9 nickase and the other encoding a controlled-release cytosine deaminase, named as tBE-V5-mA3 (Figure 2C). When translating into protein, both SpCas9 and tBE-V5-mA3 are required to simultaneously enter the nucleus for successful genome editing. We next compared several nuclear localization signals (NLSs) and found that when SpCas9, bearing an N-terminal c-Myc NLS and a C-terminal bipartite NLS, is co-delivered with tBE-V5-mA3, which bears 4x bipartite NLS, the editing efficiency substantially increased from 31.2% to 63.0% at C6 and from 46.2% to 81.3% at C7 site, respectively (Figure 2C). So far, we have optimized an RNA system to successfully deliver tBE into CD34⁺ HSPCs with high efficiency.

Identification of tBE-editing sites for inducing the highest γ -globin expression

To evaluate the therapeutic potential of tBE-based editing strategies, we electroporated mRNA encoding tBE and chemically modified sgRNA/hsgRNA pairs into the HUDEP-2 (Human Umbilical cord blood-Derived Erythroid Progenitor-2) cell line,

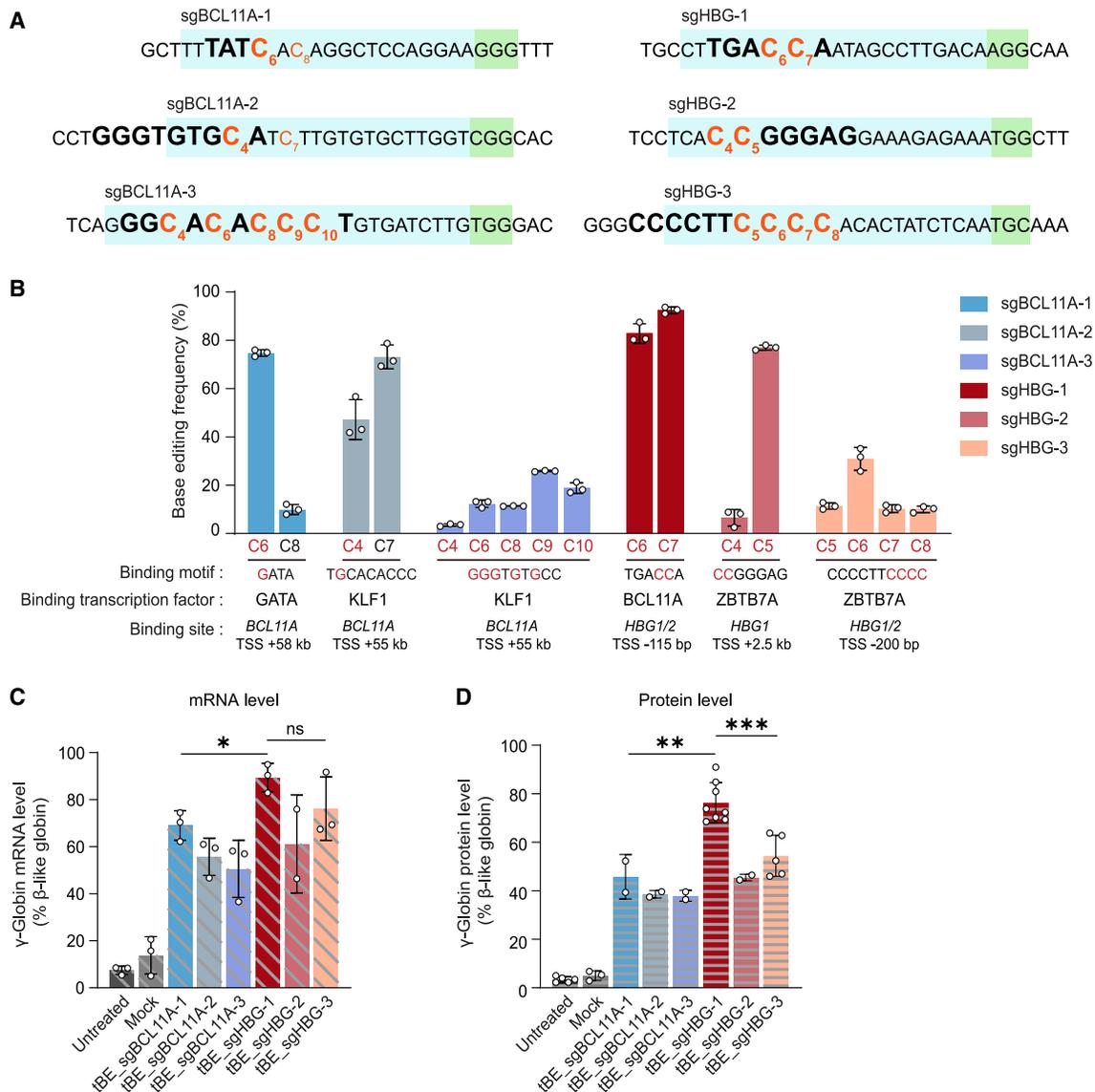


Figure 3. Reactivated γ -globin expression by tBE-mediated disruption of transcription-factor-binding motifs in HUDEP-2

(A) Illustration of sgRNA design for the targeted transcription-factor-binding motifs. Green box, PAM; light blue box, sgRNA-targeted protospacer sequence; bold letter, transcription-factor-binding motif; orange letter, edited cytosine.

(B) Base-editing efficiency (sum of C-to-T, C-to-G, and C-to-A editing) of tBE at the on-target sites shown in (A) in HUDEP-2 cells.

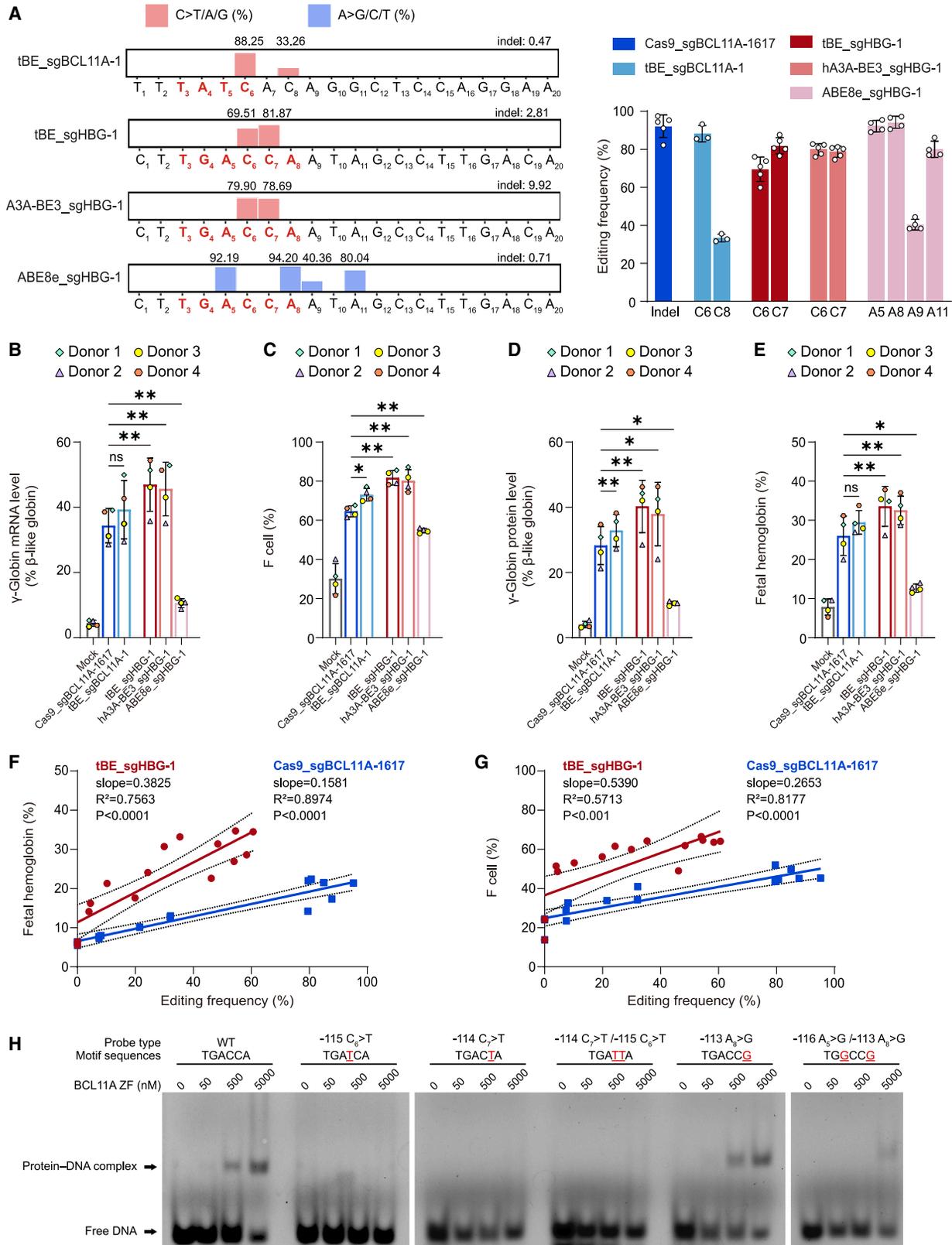
(C) Ratios of γ -globin mRNA to total β -like globin mRNA triggered by tBE with different sgRNAs in HUDEP-2 cells after *in vitro* erythroid differentiation.

(D) Ratios of γ -globin protein to total β -like globin protein triggered by tBE with different sgRNAs in HUDEP-2 cells after *in vitro* erythroid differentiation.

All data represent mean \pm SD. $n \geq 2$. In (C) and (D), p values were calculated by t test; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and ns, non-significant.

which is a CD34-derived immortalized cell line capable of differentiation into erythroid cells, allowing us to measure γ -globin levels.⁵⁴ In the *BCL11A* enhancer region, editing with sgBCL11A-1 directly converted C at protospacer positions 6 and 8, with a mean frequency of 74.8% and 9.9%, respectively (Figures 3A and 3B). Although editing at C8 had relatively low efficiency, it's worth noting that this site was not within the GATA binding motif. Editing against one KLF1 binding site by sgBCL11A-2 yielded 47.2% and 73.2% base conversions at C4 and C7 sites, respectively. Lower editing efficiency was observed in another KLF1 binding site targeted by sgBCL11A-3 (Figures 3A and 3B). In the *HBG1/2* promoters re-

gion, tBE-based sgHBG-1 (sgRNA targeting *HBG1/2* promoter No. 1) effectively converted C6 and C7 (82.8% and 92.4%, respectively), and both sites were within the *BCL11A* binding motif (TGACCA box) (Figures 3A and 3B). At the TSS +2.5 kb ZBTB7A binding site, sgHBG-2 yielded an average of 77.0% conversion at the C5 site, which is within the motif. In the TSS -200 bp ZBTB7A binding site, due to the absence of an NGG (the N represents any nucleotide [A, C, G, or T] followed by two guanine [G] nucleotide bases) PAM (Protospacer adjacent motif), a SpCas9 variant, SpG, was used in conjunction with sgHBG-3 to convert four tandem Cs within the motif at an efficiency ranging from 10.0% to 30.9% (Figures 3A and



(legend on next page)

3B). When measuring reactivated γ -globin mRNA expression by real-time quantitative PCR (qPCR), we found that mutating the BCL11A binding motif located at the *HBG1/2* promoters produced the highest γ -globin mRNA levels (Figure 3C). Similarly, the protein level of γ -globin monomer exhibited a similar trend, with disruption of the *cis*-BCL11A motif by sgHBG-1 resulting in up to a 1.76-fold increase compared with targeting the *trans*-erythroid enhancer motif (Figure 3D).

Since HbF levels negatively correlate with the number of morbidities,²⁵ we next asked whether mutating two motifs simultaneously could promote more robust γ -globin activation than single editing. Although disrupting the *BCL11A* enhancer or directly mutating the BCL11A binding site in the *HBG1/2* promoters functions in the same pathway, we chose to simultaneously mutate the binding motifs recognized by BCL11A and ZBTB7A in the *HBG1/2* promoters. Though co-editing by sgHBG-1 and sgHBG-3 showed comparable editing efficiency to the single editing control by sgHBG-1 (Figure S2A), the dual editing strategy did not result in more γ -globin production than single editing, as evidenced by comparable mRNA and protein levels (Figures S2B–S2D). Notable, the editing efficiency at the sgHBG-3 targeting site was low, which may compromise its role in restoring γ -globin levels. To address this, we isolated single clones from HUDEP-2 cells that were edited by sgHBG-1 and sgHBG-3, individually or in combination. We observed that mutating the BCL11A binding motif at TSS –115 showed about 78% γ -globin, whereas mutating the ZBTB7A binding motif at TSS –200 showed an average γ -globin level of 52%. The co-edited clones of sgHBG-1/sgHBG-3 exhibited a higher γ -globin level than the single clones edited by sgHBG-1 or sgHBG-3, suggesting that BCL11A and ZBTB7A function through distinct mechanisms to suppress γ -globin expression (Figure S2E).

Comparison of tBE-based editing efficiency and γ -globin reactivation with other CRISPR tools in CD34⁺ HSPCs

Several strategies have been developed to reactivate γ -globin, including disrupting *BCL11A* erythroid enhancer or *HBG* promoter regions by SpCas9 nuclease^{22,31,45} or base editors.^{23,24,55} To compare our strategy with other preclinical and clinical method, we performed a side-by-side experiment including a therapeutic lead gRNA targeting *BCL11A* erythroid enhancer that has been tested in clinical trials, i.e., sgBCL11A-1617.^{6,22} After the examination of Cas9 expression, we found that tBE exhibited comparable expression to other base editors (hA3A-BE3 and ABE8e) formulated in mRNA (Figure S3A). When targeting the *BCL11A* erythroid enhancer, tBE-based editing exhibited

88.3% on the C6 site within the GATA motif, a level comparable to SpCas9 nuclease-induced total indel frequency but with relative purer product (Figures 4A and S3B). When targeting the BCL11A binding motif in the promoter region of *HBG1/2*, tBE-based editing achieved similar efficiency compared with another CBE, hA3A-BE3, reaching ~80% on C6 and C7 sites but with a 3.5-fold lower indel rate (Figures 4A and S3C). Single-clone analysis of tBE_sgHBG-1-edited HSPCs also confirmed effective editing at four copies of the BCL11A binding motifs in the *HBG1/2* promoters region (Figure S3D).

Edited HSPCs were then *in vitro* differentiated into erythroid cells to measure the γ -globin mRNA and protein expression levels. By using sgHBG-1, CBEs (i.e., tBE and hA3A-BE3)-modified BCL11A binding motif produced significantly more γ -globin mRNA and protein monomer than SpCas9 nuclease-modified *BCL11A* erythroid enhancer in all four donors tested (Figures 4B–4D, S3E, and S3F). We then performed a dose-response comparison of tBE-edited BCL11A binding motif with Cas9 nuclease-edited BCL11A enhancer. The percentage of HbF and F cells correlated with the editing frequencies (Figures 4F, 4G, and S3G). The steeper slope of tBE-edited BCL11A binding motif (–114C > T –115C > T) also suggests more potent globin reactivation by tBE-edited *HBG1/2* promoter than the nuclease-based strategy (Figures 4F and 4G).

It was noted that ABE8e-mediated editing of BCL11A binding motif, although achieved up to 94.2% editing efficiency, produced the lowest γ -globin expression. The different levels of γ -globin expression triggered by CBE and ABE suggest that the altered motif sequence may impact the binding affinity to its repressor. To test this hypothesis, we performed electrophoretic mobility shift assay (EMSA) using various mutated sequences that can be generated by tBE- or ABE8e-mediated C-to-T or A-to-G conversion, respectively. Quite interestingly, we found that tBE-mediated conversion of TGACTA or TGATTA completely abolished the BCL11A binding, whereas the A-to-G conversion at positions –116 and –113, the two main edited bases by ABE8e, only partially affected the binding (Figure 4H). Consistently, chromatin-immunoprecipitation qPCR (ChIP-qPCR) in HUDEP-2 cells showed that the C-to-T conversions at positions –114 and –115 significantly abolished the interaction (Figure S3H). These results together explained why ABE-mediated disruption of the BCL11A motif only partially reactivated γ -globin.

Persistent editing of HSCs with high induction of HbF in erythroid progeny

To determine whether tBE-edited HSPCs retained repopulation ability and whether the editing persisted over the differentiation

Figure 4. tBE-mediated disruption of BCL11A binding motif in HSPCs triggers high γ -globin reactivation

- (A) Comparison of editing efficiency targeting the BCL11A binding motif in the promoter of *HBG1/2* or the *BCL11A* erythroid enhancer GATA motif by different genome-editing tools. (Left) The sequences of sgRNA spacer regions covering the binding motifs. Red bold letter, transcription-factor-binding motif. (Right) Base editing or indel frequencies induced by different genome-editing tools with the indicated sgRNAs.
- (B) Real-time qPCR analysis of γ -globin mRNA level relative to β -like globin after erythroid differentiation of edited HSPCs with indicated genome editors.
- (C) Fluorescence-activated cell sorting (FACS) analysis of F cell percentage in edited HSPCs after erythroid differentiation.
- (D) High performance liquid chromatography (HPLC) analysis of globin chains in edited HSPCs after erythroid differentiation.
- (E) HPLC analysis of hemoglobin tetramer in edited HSPCs after erythroid differentiation.
- (F) Correlation analysis of HbF percentage and on-target editing efficiency.
- (G) Correlation analysis of F cell percentage and on-target editing efficiency.
- (H) EMSA analysis of BCL11A binding affinity with the motifs bearing indicated mutations generated by tBE or ABE8e.
- Data represent mean \pm SD. $n \geq 3$. p values were calculated by t test; * $p < 0.05$, ** $p < 0.01$, and ns, non-significant.

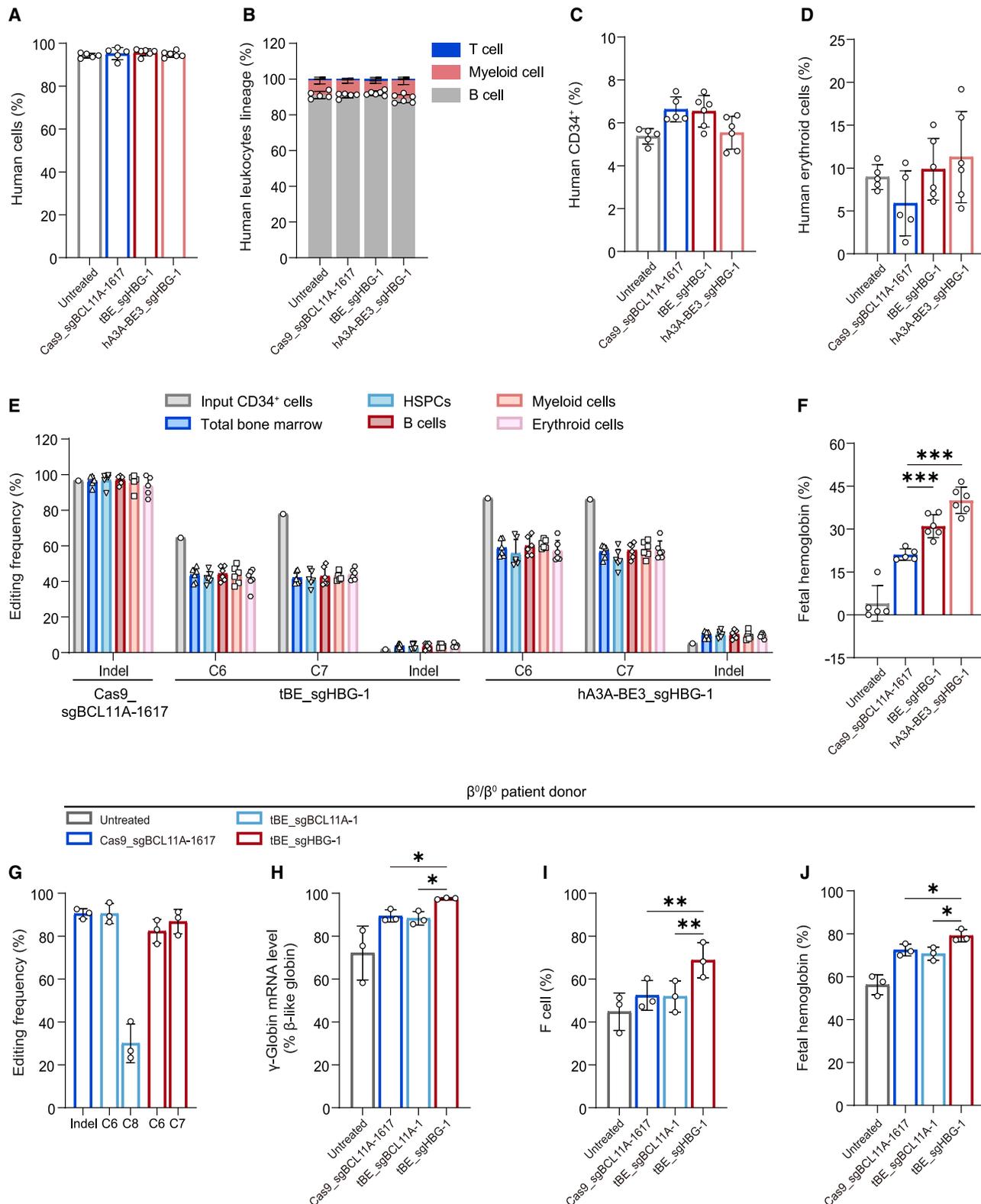


Figure 5. Durable editing of BCL11A-binding motif in BM-repopulating HSCs with enhanced HbF induction in erythroid progeny

(A) Engraftment efficiency of control and edited HSPCs with indicated editors.

(B) Percentage of engrafted human B cells (hCD45⁺hCD19⁺), myeloid cells (hCD45⁺hCD33⁺), and T cells (hCD45⁺CD19⁻CD33⁻CD3⁺) in the bone marrow 16 weeks after transplantation.

(legend continued on next page)

process, we xenotransplanted tBE-edited HSPCs into immunodeficient NBSGW (Nonirradiated NOD.B6.SCID II2r $\gamma^{-/-}$ Kit^{W41/W41}) mice and compared them with mock, hA3A-BE3-edited or Cas9-nuclease-edited HSPCs. After 16 weeks, no significant differences were observed between the edited and the mock-treated HSPCs in terms of engraftment and differentiation potential (Figures 5A–5D and S4A). To examine whether tBE editing causes skewed hematopoiesis, we purified human T cells (CD3⁺), myeloid cells (CD33⁺), B cells (CD19⁺), HSPCs (CD34⁺), and erythroblasts (CD235a⁺) by lineage-specific antibodies from mouse bone marrow (Figure S4A). Deep sequencing of the targeted genomic locus showed that all isolated populations had a similar tBE-editing frequencies of 43.6% and 42.0% at C6 and C7, respectively (Figures 5E, S4B, and S4C), which suggests that the edited alleles were maintained in hematopoietic stem cells (HSCs) and in their progenies. Similar to hA3A-BE3-edited HSPCs, we observed a modest decrease of editing efficiency from input cells, possibly indicating higher editing efficiency among non-repopulating cells within the HSPC mixture. Consistent with previous studies,²² Cas9-ribonucleoprotein (RNP)-edited HSPCs had a similar editing efficiency of 96% in all populations. In human CD235a⁺ erythroblasts isolated from individual mouse bone marrow, the average HbF percentage was 31.0% after editing with tBE_sgHBG-1, compared with 21.08% in Cas9-nuclease-edited BCL11A enhancer (Figure 5F), showing the potent efficacy of tBE-mediated editing at *HBG1/2* promoter region.

We further compared the effects of treating β -thalathemia patient CD34⁺ HSPCs with tBE mRNA to disrupt BCL11A binding site or with Cas9 RNP or tBE mRNA to disrupt BCL11A erythroid enhancer. The on-target editing efficiencies for Cas9, tBE_sgBCL11A, and tBE_sgHBG-1 were 92.3% \pm 2.1%, 86.8% \pm 3.4%, and 87.0% \pm 4.3%, respectively (Figure 5G). When measuring γ -globin mRNA, F cell, and HbF levels in the edited patient's cells, we observed similar trend of globin reactivation: disruption of the BCL11A binding site at *HBG1/2* promoter by tBE_sgHBG-1 exhibited the highest γ -globin reactivation (Figures 5H–5J). Collectively, tBE disruption of the BCL11A binding site in the *HBG1/2* promoter causes particularly potent induction of erythroid HbF *in vivo*.

Evaluation of DNA damage response and DNA OT mutation by tBE

HSCs are highly sensitive to DNA damage such as DSB triggered by SpCas9 nuclease.^{27,56} In response, edited cells initiate p53 pathway as a safeguard to protect against DSBs. As a consequence, cells may undergo cell cycle arrest and apoptosis.⁵⁷ When measuring the expression of p21 mRNA, a readout of p53-induced DNA damage response,⁵⁸ we found that SpCas9

nuclease-edited cells induced 6.7-fold upregulation of p21, indicating a strong activation of DNA damage response (Figure 6A). By contrast, tBE-edited HSPCs exhibited a temporal and mild upregulation of p21 at 6 h and dropped down to control level at 24 h post-electroporation (Figure 6A). Previous reports have indicated that in addition to its canonical role of tumor suppressor, p53 is also an interferon stimulative gene (ISG), induced by interferon upon viral infection.^{59,60} The tBE mRNA system is composed of two constructs, with lengths of 4.7 and 3.5 kb, respectively. The secondary structure of mRNA could stimulate a mild and transient upregulation of ISGs including p53. Upon investigating the RNA sensing pathway, we observed a temporal and mild upregulation of RIG-I (Retinoic acid-inducible gene-1) and MAD5 (Melanoma differentiation associated gene 5) sensing pathway at 6 h post-electroporation (Figure S5B), which may indirectly activate p53. A significant consequence of persistent p53 activation is cell cycle arrest.⁵⁷ To determine whether the temporal upregulation of p53 in tBE editing could affect cell growth, we monitored cell growth over time. The tBE-edited HSPCs exhibited a comparable cell growth rate to the single-variable controls, whereas Cas9 nuclease-edited HSPCs or hA3A-BE3-edited cells displayed reduced cell growth over time (Figures 6B and S5B).

In addition, when using SpCas9 nuclease to edit BCL11A binding site within the *HBG1/2* promoter region, a significant proportion of 4.9-kb genomic deletion was generated due to the dual cutting at the duplicated *HBG1* and *HBG2* genomic loci (Figures S6A and S6B). Quantification of deletion frequency by droplet digital PCR revealed 50.4% \pm 1.0% frequency using Cas9 nuclease compared with 6.4% \pm 4.5% frequency using tBE (Figure S6B).

We next determined whether tBE caused OT editing in genomic DNA. DNA OT activity can happen on loci that share sequence similarity to gRNA (known as gRNA-dependent OT), or loci whose sequence has no relation to gRNA (known as gRNA-independent OT).^{36,38,40,61} We first determined gRNA-dependent OT loci by using the CasOFFinder tool.⁶² We identified 81 potential OT sites that have four or less mismatches to sgHBG-1 spacer sequence. Deep sequencing of each site in edited HSPCs, we observed little detectable OT editing in tBE-edited HSPCs (Figure 6C). By contrast, three OT sites with relatively high editing frequency at 16.6%, 12.1%, and 1.9% were identified in hA3A-BE3-edited cells and one OT site with editing frequency at 2.1% was identified in ABE8e-edited cells. Examination of engrafted mouse bone marrow, the OT editing at OT24, OT48, and OT79 persisted in hA3A-BE3-edited group, but no OT events were identified in tBE-edited group (Figure 6D).

Next, we measured gRNA-independent OT activities by whole-genome sequencing.^{41,63} In tBE-edited single-cell colonies, we

(C) Percentage of engrafted HSPCs (hCD45⁺ hCD34⁺).

(D) Percentage of erythroid progeny (hCD45⁺ mCD45⁺ hCD235a⁺).

(E) Comparison of editing efficiency of engrafted cells in different lineages isolated at 16 weeks post-transplantation over input cells.

(F) Percentage of HbF in the engrafted erythroid progeny.

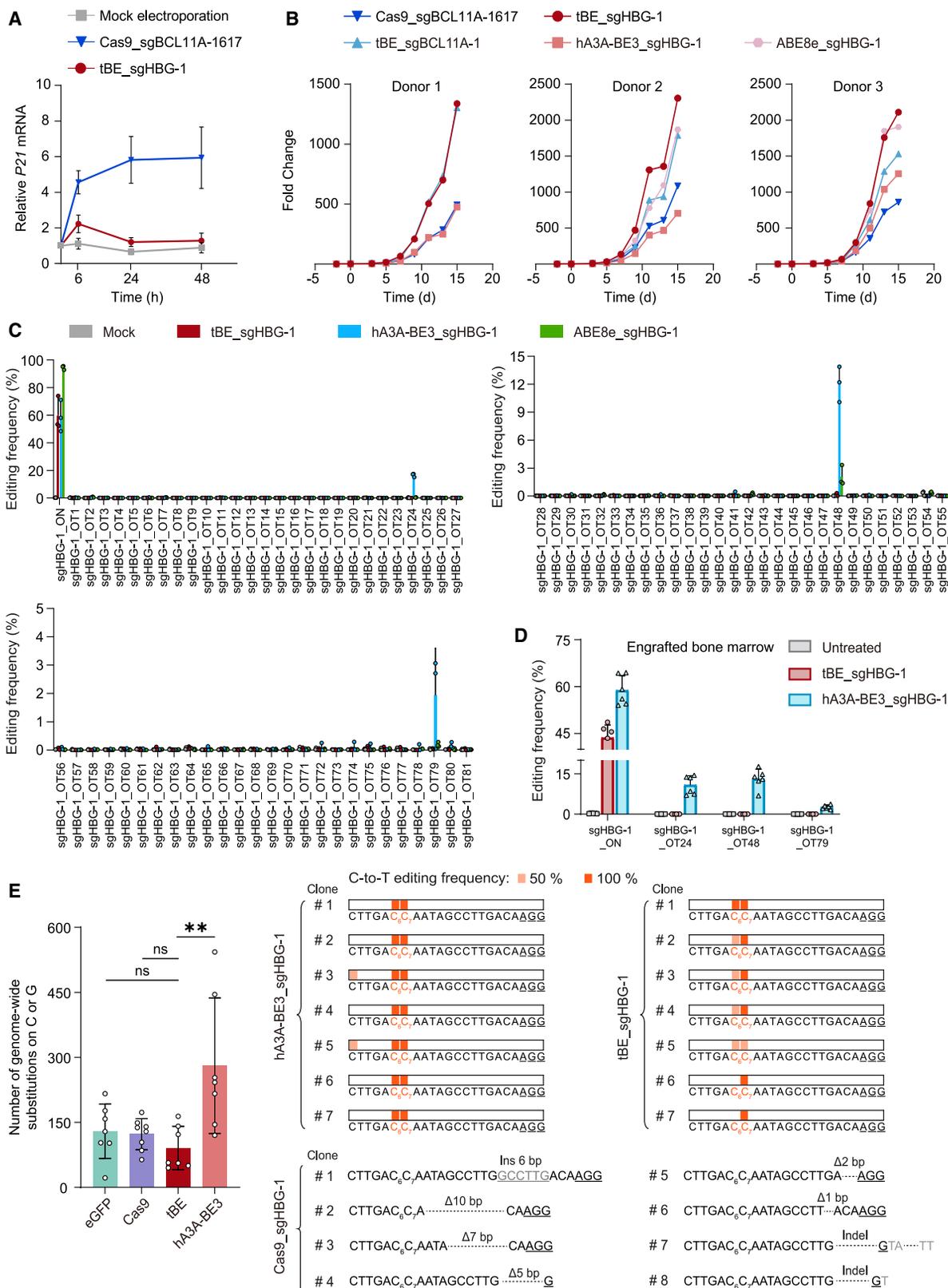
(G) Comparison of editing efficiency in patient CD34⁺ cells.

(H) Real-time qPCR analysis of γ -globin mRNA level relative to β -like globin after erythroid differentiation of edited patient HSPCs with indicated editors.

(I) FACS analysis of F cell percentage in edited patient HSPCs after erythroid differentiation.

(J) HPLC analysis of hemoglobin tetramer in edited patient HSPCs after erythroid differentiation.

In (A)–(F), data represent median \pm SEM. Each dot represents one animal. In (G)–(J), data represent mean \pm SD, and each dot represents one patient donor. p values were calculated by t test; *p < 0.05, **p < 0.01, and ***p < 0.001.



(legend on next page)

only found background levels of cytosine or guanine mutations, similar to those in unedited or Cas9-edited single-cell colonies, demonstrating that tBE did not cause gRNA-independent OT mutations (Figure 6E). By contrast, hA3A-BE3 induced significantly more gRNA-independent OT mutations in the genomes of edited cells (Figure 6E). Together, the above data demonstrated that tBE did not trigger detectable DNA damage response or DNA OT mutations in HSPCs.

Evaluation of RNA OT mutation by tBE

In addition to DNA OT mutations, recent studies found that base editors can also trigger OT mutations in transcriptomic RNA.^{33,34} To analyze whether tBE caused RNA OT mutations, we performed transcriptome sequencing to determine RNA OT events in HSPCs using sgHBG-1. Since tBE mRNA expression peaked at 12 h post-electroporation (Figure S3A), we harvested cells at 12 h for RNA transcriptome analysis. We found that tBE induced efficient base editing at on-target site (Figure 7A), whereas tBE-edited HSPCs displayed only a background level of RNA OT mutations, similar to mock control or SpCas9 nuclease-edited group (Figures 7B and S7A). Similar results were observed in HUDEP-2 cells as well (Figures S7B–S7D). By contrast, when formulated in plasmid, we observed a large number of transcriptomic OT events in hA3A-BE3- or BE3-edited cells (Figure 7C), consistent to previous reports.⁴¹

Thus, these data indicate that tBE-edited HSPCs has little detectable transcriptome-wide OT mutations, highlighting tBE as a safe and powerful tool to reactivate high level of γ -globin for treating β -hemoglobinopathy.

DISCUSSION

Several CRISPR-based genome-editing strategies have been developed to reactivate γ -globin expression for treating β -hemoglobinopathies.^{22–24,31,45,55,64} Among them, the earliest developed SpCas9 nuclease-mediated disruption of GATA motif in *BCL11A* erythroid enhancer has accomplished clinical trials (NCT03655678,⁶⁵ NCT03745287⁶⁶), providing proof of the reactivation strategy in treating the diseases. Although significant progress has been made, safety concerns related to the occurrence of harmful double-stranded breaks and large genomic rearrangements in HSCs remain a crucial issue to address.^{27,28,30} In addition, a few patients still required blood transfusion after the treatment,⁶⁷ suggesting that the efficacy of SpCas9 nuclease-mediated therapy needs to be further improved. In this study, we used a recently developed CBE (tBE),⁴¹ which minimized the generation of DSBs and is designed to eliminate

OT activity, to systemically dissect the *cis*- and *trans*-regulatory motifs for γ -globin reactivation.

Firstly, we screened pairs of sgRNA and hsgRNA and performed functional screen in an erythroid cell line to identify the most efficient regulatory motifs for γ -globin reactivation. In particular, we found that when directly mutating the *cis*-*BCL11A* binding motif in the *HBG1/2* promoter region, tBE-edited cells exhibited higher level of γ -globin reactivation than mutating the GATA motif within the *BCL11A* erythroid enhancer. When using SpCas9 nuclease targeting the GATA motif within *BCL11A* erythroid enhancer in HSPCs, we also observed lower level of HbF than tBE-mediated disruption of the *BCL11A* binding motif. The difference is likely attributed to the residual *BCL11A* expression driven by its native promoters or other motifs within the enhancer, as highlighted by GWAS, showing the existence of several loci in TSS +55, +58, and +62 of *BCL11A* gene.^{11,64} Quite surprising, when using the same sgHBG-1 targeting the *cis*-motif but edited by ABE8e, though the editing efficiency reached over 90%, it resulted much less γ -globin than CBE- or SpCas9 nuclease-based editing strategies. Similar observation is reported in a recent study comparing the efficacy of targeting *cis*- and *trans*-elements by ABE8e.⁵⁵ In the core motif TGACCA, ABE8e-generated sequence TGGCCG still retained binding affinity with *BCL11A*, whereas tBE-generated TGATTA completely abolished its interaction in *in vitro* EMSA experiment. In support, direct measurement of the binding affinity of *BCL11A* to the core motif revealed that mutating into –116A or –113A (A5 and A8 for sgHBG-1; Figure 4A) only reduced the binding affinity by ~3.7- or 1.6-fold,⁶⁸ whereas mutating into –115C or –114C (C6 and C7 for sgHBG-1; Figure 4A) greatly decreased the affinity by ~31- or ~55-fold, respectively.⁶⁸ Together, these data suggest that not only editing efficiency is critical, but the diverse editing types by different base editors can cause phenotypic variation that together determine the final therapeutic outcome.

ZBTB7A is a recently identified repressor to regulate globin silencing.¹⁶ Clonal analysis in HUDEP-2 cells showed that simultaneously knocking out *ZBTB7A* and *BCL11A* genes exhibited a significantly greater HbF than did the single knockout.¹⁶ Consistently, simultaneously disrupting *ZBTB7A* binding motif (TSS –200) and *BCL11A* binding motif (TSS –115) further potentiate γ -globin reactivation than the single editing in HUDEP-2 clonal cell line. Unfortunately, in HSPCs, tBE-based dual editing of the *ZBTB7A* and *BCL11A* binding motif did not induce more γ -globin than the single editing. Due to the lack of an NGG PAM, the SpG variant⁵¹ was used to mutate the *ZBTB7A* binding motif, resulting in low editing efficiency (~30%). It is possible that a threshold level of editing at the *ZBTB7A* binding motif is required to observe additive effect.

Figure 6. DNA damage and DNA off-target analysis

(A) Real-time qPCR measurement of *CDKN1A* (*P21*) transcription levels triggered by tBE and Cas9 nuclease in HSPCs.

(B) Fold change of edited HSPCs cell growth after the treatment of different genome editors. Counting of cell numbers was performed on day 0 after electroporation.

(C) Comparison of off-target (OT) editing frequency induced by indicated genome editors at potential genomic loci with four or fewer mismatches to the on-target (ON) site. For tBE and hA3A-BE3, data represent mean C-to-T frequency at each potential OT site; for ABE8e, data represent mean A-to-G frequency at each potential OT site.

(D) Off-target analysis in engrafted bone marrow cells.

(E) (Left) Whole-genome analysis of substitutions of C or G in the single-cell clones treated with indicated genome editors or eGFP control. (Right) On-target genotype of the single-cell clones used for whole-genome analysis. Underlined letter, PAM.

Data represent mean \pm SD. $n \geq 3$. In (D), p values were calculated by t test; **p < 0.01 and ns, non-significant.

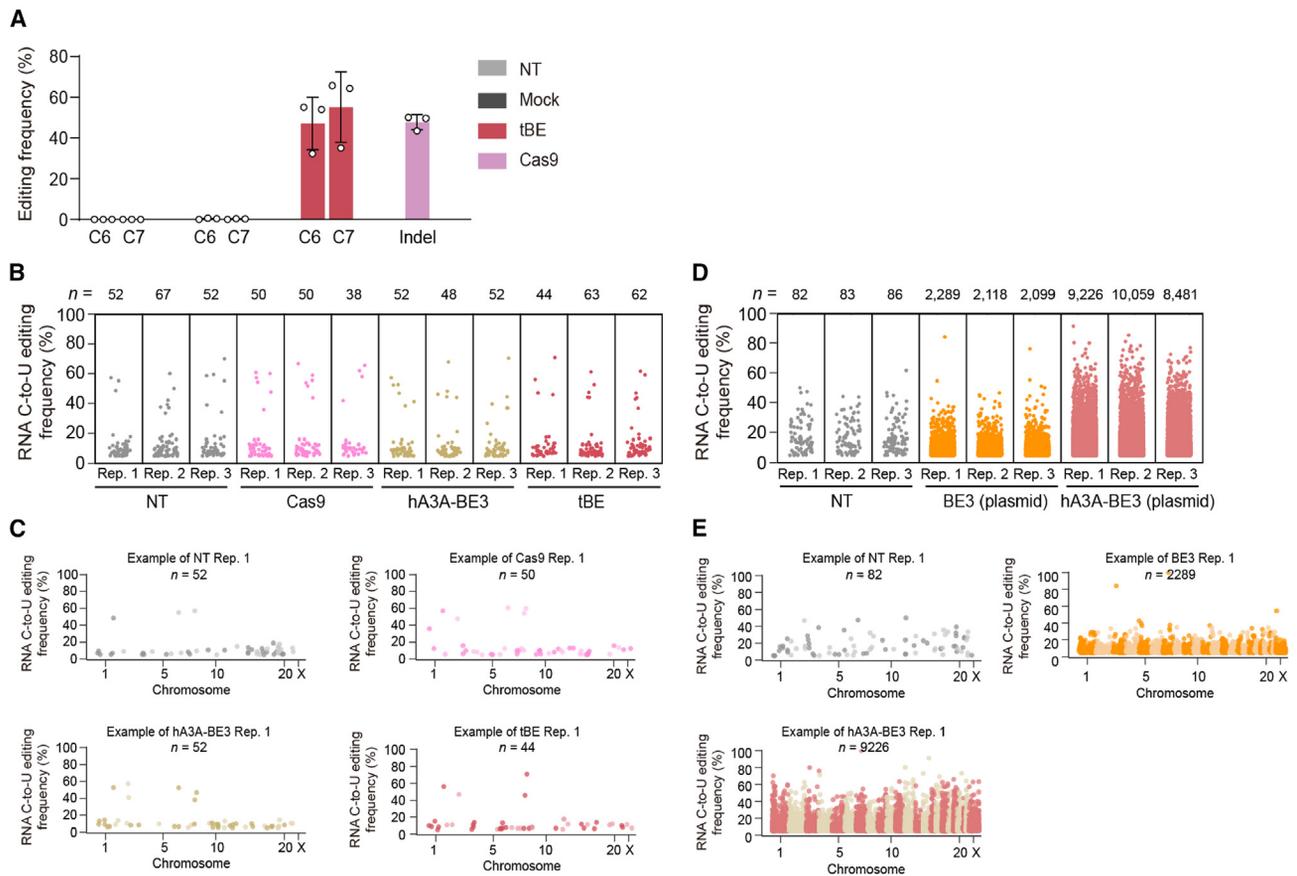


Figure 7. Transcriptome-wide RNA off-target analysis

(A) Base editing or indel frequency induced by tBE or Cas9 nuclease at the sgHBG-1 on-target site in HSPCs.

(B) Manhattan plot of transcriptome-wide C-to-U editing frequency in HSPCs using the indicated editors that are formulated in mRNA or RNP. Data were collected at 12 h post-electroporation.

(C) Frequency and distribution of RNA C-to-U editing in replicate 1 (Rep. 1) of HSPCs.

(D) Manhattan plot of transcriptome-wide C-to-U editing frequency in 293FT cells using the indicated editors that are formulated in plasmid. sgHBG-1 was used.

(E) Frequency and distribution of RNA C-to-U editing in replicate 1 (Rep. 1) of 293FT.

Data represent mean \pm SD. $n = 3$.

HSPCs are lineage-committed stem cells and can repopulate the entire blood system. OT editing or unintended side effects from DSBs could invoke unexpected consequences in stem cells such as reduced HSC clonal repertoire.^{30,69} Thus, it is important to evaluate the OT effect of genome-edited HSPCs. Conventional base editors are often associated with high OT editing.^{23,55,70} To address this, we performed a comprehensive OT analysis on tBE, including gRNA-dependent OT, genome-wide DNA OT, and transcriptome-wide RNA OT evaluation. Compared with hA3A-BE3 or ABE8e, which displayed different levels of OT activities, tBE showed little detectable DNA or RNA OT events, similar to mock-treated or untreated samples.

In conclusion, our study provided proof that tBE-based editing targeting the BCL11A binding site within the *HBG1/2* promoter is a potent and highly efficient strategy to reactivate HbF in human HSCs. The feasibility of mRNA delivery with little detectable OT activity provides a strong basis for advancing the clinical development of tBE in autologous HSCs as a potentially curative therapy for β -hemoglobinopathies.

Limitations of the study

The tBE system is composed of a sgRNA and a hsgRNA. Although all nicks in tBE or hA3A-BE3 were generated on the same strand, we still observed a small percentage of 4.9-kb deletion in the duplicated *HBG1* and *HBG2* region, indicating the generation of DSBs. Similarly, ABE8e-mediated editing has been reported to induce the 4.9-kb deletion.²⁴ The DSBs generated in the base editor can be attributed to the DNA repair mechanism triggered by the nucleobase deamination process⁷¹. Deamination causes a C-to-U change in the non-nicking strand. Despite the fusion of the uracil DNA glycosylase inhibitor to the editor, there is a chance of U cleavage by uracil DNA glycosylase to generate an abasic site. Consequently, cleavage of abasic site in the non-nicking strand can lead to a DSB,⁷² which accounts for the low but detectable DSBs observed in base editor-mediated editing. In applications that require preventing DSB generation as much as possible, BEs derived from catalytically dead Cas proteins can be tested.

In the tBE mRNA preparation, we included uridine depletion and cellulose-based double-strand RNA removal methods,

which have been shown to reduce immunogenicity.⁷³ However, we still observed a temporal and mild upregulation of the RNA sensing pathway at 6 h after the delivery of mRNA-encoded editors, including tBE, ABE8e, and hA3A-BE3. The mRNA system of tBE is composed of two constructs, with lengths of 4.7 and 3.5 kb, respectively. The naturally formed secondary structure of mRNA could stimulate the immune response. Future studies using sequence optimization to reduce the secondary structure of RNA can be tested to further avoid immune stimulation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead contact
 - Materials availability
 - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
 - Patient and healthy donor cells
 - Mice
- **METHOD DETAILS**
 - Cell line culture and plasmid transfection
 - CD34⁺ HSPC culture and *in vitro* erythroid differentiation
 - gRNA, mRNA preparation and electroporation
 - Real-time qPCR of globin expression
 - HPLC analysis of globin chains and hemoglobin tetramers
 - Xenotransplantation in NBSGW mice
 - Flow cytometry for F-cell
 - Protein expression and purification
 - Electrophoretic mobility shift assay (EMSA)
 - Quantification of Cas9 expression level
 - Chromatin-immunoprecipitation quantitative PCR (ChIP-qPCR)
 - Droplet Digital PCR (ddPCR)
 - DNA library preparation and amplicon sequencing
 - Libraries preparation for genome-wide and transcriptome-wide off-target analysis
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Amplicon sequencing data analysis
 - Analysis of gRNA-dependent off-target mutation
 - Whole-genome sequencing data analysis
 - Whole-transcriptome sequencing data analysis
 - Statistical tests

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.stem.2023.10.007>.

ACKNOWLEDGMENTS

This work is kindly supported by the National Key Research and Development Program National Key R&D Program of China (2018YFA0801401 to J.C. and H.Y., 2022YFF1002801 to Y.Z., 2019YFA0802804 to B.Y. and L. Yang, and 2019YFA0802801 to H.Y.), the National Natural Science Foundation of China

(31972936 to Y.Z., 31925011 to L. Yang, 32371514 to J.C., and 32070170 to B.Y.), Ministry of Agriculture and Rural Affairs of the People's Republic of China (NK2022010207), the Major Scientific and Technological Project of Hubei Province (2022ACA005 to Y.Z. and 2022BCA089 to H.Y.), the Medical Science Advancement Program (Basic Medical Sciences) of Wuhan University (TFJC2018005), the Fundamental Research Funds for the Central Universities (2042022dx0003 and 2042022kf1190), Shanghai Municipal Science and Technology Commission (21JC1404600 to J.C. and 23ZR1442500 to B.Y.), and Program of Shanghai Academic/Technology Research Leader (23XD1422500 to J.C.). We thank the Shanghai Frontiers Science Center for Biomacromolecules and Precision Medicine, ShanghaiTech University, Molecular and Cell Biology Core Facility, School of Life Science and Technology, ShanghaiTech University and the core facility of Medical Research Institute at Wuhan University for their technical support.

AUTHOR CONTRIBUTIONS

gRNA and functional screens, W.H., S.S., M.Z., and J. Wu; HSPCs experiments, H.-Y.Q., G.-Q.W. (mRNA preparation), Y.W., and L.W.; xenotransplant experiments, H.-Y.Q., G.-H.C., R.-J.J., L.B., and Q.-B.C.; globin mRNA quantification, W.H. and H.-Y.Q.; globin monomer and tetramer quantification, S.S. and H.M.; OT analysis, W.H., Z.-C.F., and S.S.; patient cell collecting, X.Q., X.Z., H.W., Z.L., and L. Yu; patient cell editing and analysis, L.W., Y.W., and H.M.; library preparation for amplicon sequencing, whole-genome sequencing (WGS), and RNA sequencing (RNA-seq), J. Wei and Y.-L.G.; bioinformatics analysis, Z.-C.F. and Y.-Z.Z.; project design and supervision, Y.Z., B.Y., J.C., L. Yang, H.Y., and X.M. Y.Z., J.C., W.H., and H.-Y.Q. wrote the manuscript with input from all authors.

DECLARATION OF INTERESTS

J.C., L. Yang, B.Y., and H.Y. are the scientific cofounders of CorrectSequence Therapeutics, a company that uses gene-editing technologies. X.M., L.W., Y.W., and H.M. are the employees of CorrectSequence Therapeutics. J.C., B.Y., L. Yang, W.H., S.S., and Y.Z. filed a provisional patent application related to this work.

Received: April 18, 2023

Revised: September 13, 2023

Accepted: October 24, 2023

Published: November 20, 2023

REFERENCES

1. Qiu, H.-Y., Ji, R.-J., and Zhang, Y. (2022). Current advances of CRISPR-Cas technology in cell therapy. *Cell Insight* 7, 100067. <https://doi.org/10.1016/j.cellin.2022.100067>.
2. Ferrari, G., Thrasher, A.J., and Aiuti, A. (2021). Gene therapy using haematopoietic stem and progenitor cells. *Nat. Rev. Genet.* 22, 216–234. <https://doi.org/10.1038/s41576-020-00298-5>.
3. Ernst, M.P.T., Broeders, M., Herrero-Hernandez, P., Oussoren, E., van der Ploeg, A.T., and Pijnappel, W.W.M.P. (2020). Ready for repair? Gene editing enters the clinic for the treatment of human disease. *Mol. Ther. Methods Clin. Dev.* 18, 532–557. <https://doi.org/10.1016/j.omtm.2020.06.022>.
4. Cromer, M.K., Camarena, J., Martin, R.M., Lesch, B.J., Vakulskas, C.A., Bode, N.M., Kurgan, G., Collingwood, M.A., Rettig, G.R., Behlke, M.A., et al. (2021). Gene replacement of alpha-globin with beta-globin restores hemoglobin balance in beta-thalassemia-derived hematopoietic stem and progenitor cells. *Nat. Med.* 27, 677–687. <https://doi.org/10.1038/s41591-021-01284-y>.
5. Wilber, A., Nienhuis, A.W., and Persons, D.A. (2011). Transcriptional regulation of fetal to adult hemoglobin switching: new therapeutic opportunities. *Blood* 117, 3945–3953. <https://doi.org/10.1182/blood-2010-11-316893>.
6. Frangoul, H., Altshuler, D., Cappellini, M.D., Chen, Y.S., Domm, J., Eustace, B.K., Foell, J., de la Fuente, J., Grupp, S., Handgretinger, R.,

- et al. (2021). CRISPR-Cas9 gene editing for sickle cell disease and beta-thalassemia. *N. Engl. J. Med.* 384, 252–260. <https://doi.org/10.1056/NEJMoa2031054>.
7. Bauer, D.E., Kamran, S.C., and Orkin, S.H. (2012). Reawakening fetal hemoglobin: prospects for new therapies for the beta-globin disorders. *Blood* 120, 2945–2953. <https://doi.org/10.1182/blood-2012-06-292078>.
8. Hoban, M.D., Orkin, S.H., and Bauer, D.E. (2016). Genetic treatment of a molecular disorder: gene therapy approaches to sickle cell disease. *Blood* 127, 839–848. <https://doi.org/10.1182/blood-2015-09-618587>.
9. Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G., et al. (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. *Proc. Natl. Acad. Sci. USA* 105, 1620–1625. <https://doi.org/10.1073/pnas.0711566105>.
10. Smith, E.C., and Orkin, S.H. (2016). Hemoglobin genetics: recent contributions of GWAS and gene editing. *Hum. Mol. Genet.* 25, R99–R105. <https://doi.org/10.1093/hmg/ddw170>.
11. Bauer, D.E., Kamran, S.C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M.C., Smith, E.C., Pinello, L., et al. (2013). An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* 342, 253–257. <https://doi.org/10.1126/science.1242088>.
12. Martyn, G.E., Wienert, B., Yang, L., Shah, M., Norton, L.J., Burdach, J., Kurita, R., Nakamura, Y., Pearson, R.C.M., Funnell, A.P.W., et al. (2018). Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. *Nat. Genet.* 50, 498–503. <https://doi.org/10.1038/s41588-018-0085-0>.
13. Cheng, L., Li, Y., Qi, Q., Xu, P., Feng, R., Palmer, L., Chen, J., Wu, R., Yee, T., Zhang, J., et al. (2021). Single-nucleotide-level mapping of DNA regulatory elements that control fetal hemoglobin expression. *Nat. Genet.* 53, 869–880. <https://doi.org/10.1038/s41588-021-00861-8>.
14. Doerfler, P.A., Feng, R., Li, Y., Palmer, L.E., Porter, S.N., Bell, H.W., Crossley, M., Pruett-Miller, S.M., Cheng, Y., and Weiss, M.J. (2021). Activation of gamma-globin gene expression by GATA1 and NF- κ B in hereditary persistence of fetal hemoglobin. *Nat. Genet.* 53, 1177–1186. <https://doi.org/10.1038/s41588-021-00904-0>.
15. Liu, N., Hargreaves, V.V., Zhu, Q., Kurland, J.V., Hong, J., Kim, W., Sher, F., Macias-Trevino, C., Rogers, J.M., Kurita, R., et al. (2018). Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell* 173, 430–442.e17. <https://doi.org/10.1016/j.cell.2018.03.016>.
16. Masuda, T., Wang, X., Maeda, M., Canver, M.C., Sher, F., Funnell, A.P., Fisher, C., Suciu, M., Martyn, G.E., Norton, L.J., et al. (2016). Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science* 351, 285–289. <https://doi.org/10.1126/science.aad3312>.
17. Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008). Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* 322, 1839–1842. <https://doi.org/10.1126/science.1165409>.
18. Liu, N., Xu, S., Yao, Q., Zhu, Q., Kai, Y., Hsu, J.Y., Sakon, P., Pinello, L., Yuan, G.C., Bauer, D.E., and Orkin, S.H. (2021). Transcription factor competition at the gamma-globin promoters controls hemoglobin switching. *Nat. Genet.* 53, 511–520. <https://doi.org/10.1038/s41588-021-00798-y>.
19. Liu, P., Keller, J.R., Ortiz, M., Tessarollo, L., Rachel, R.A., Nakamura, T., Jenkins, N.A., and Copeland, N.G. (2003). Bcl11a is essential for normal lymphoid development. *Nat. Immunol.* 4, 525–532. <https://doi.org/10.1038/ni925>.
20. Yu, Y., Wang, J., Khaled, W., Burke, S., Li, P., Chen, X., Yang, W., Jenkins, N.A., Copeland, N.G., Zhang, S., and Liu, P. (2012). Bcl11a is essential for lymphoid development and negatively regulates p53. *J. Exp. Med.* 209, 2467–2483. <https://doi.org/10.1084/jem.20121846>.
21. Maeda, T., Ito, K., Merghoub, T., Polisenio, L., Hobbs, R.M., Wang, G., Dong, L., Maeda, M., Dore, L.C., Zelent, A., et al. (2009). LRF is an essential downstream target of GATA1 in erythroid development and regulates BIM-dependent apoptosis. *Dev. Cell* 17, 527–540. <https://doi.org/10.1016/j.devcel.2009.09.005>.
22. Wu, Y., Zeng, J., Roscoe, B.P., Liu, P., Yao, Q., Lazzarotto, C.R., Clement, K., Cole, M.A., Luk, K., Baricordi, C., et al. (2019). Highly efficient therapeutic gene editing of human hematopoietic stem cells. *Nat. Med.* 25, 776–783. <https://doi.org/10.1038/s41591-019-0401-y>.
23. Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A.H., Shea, D., Gehrke, J.M., Clement, K., Luk, K., Yao, Q., et al. (2020). Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* 26, 535–541. <https://doi.org/10.1038/s41591-020-0790-y>.
24. Antoniou, P., Hardouin, G., Martinucci, P., Frati, G., Felix, T., Chalumeau, A., Fontana, L., Martin, J., Masson, C., Brusson, M., et al. (2022). Base-editing-mediated dissection of a gamma-globin cis-regulatory element for the therapeutic reactivation of fetal hemoglobin expression. *Nat. Commun.* 13, 6618. <https://doi.org/10.1038/s41467-022-34493-1>.
25. Musallam, K.M., Sankaran, V.G., Cappellini, M.D., Duca, L., Nathan, D.G., and Taher, A.T. (2012). Fetal hemoglobin levels and morbidity in untransfused patients with beta-thalassemia intermedia. *Blood* 119, 364–367. <https://doi.org/10.1182/blood-2011-09-382408>.
26. Urbinati, F., Hargrove, P.W., Geiger, S., Romero, Z., Wherley, J., Kaufman, M.L., Hollis, R.P., Chambers, C.B., Persons, D.A., Kohn, D.B., and Wilber, A. (2015). Potentially therapeutic levels of anti-sickling globin gene expression following lentivirus-mediated gene transfer in sickle cell disease bone marrow CD34+ cells. *Exp. Hematol.* 43, 346–351. <https://doi.org/10.1016/j.exphem.2015.01.009>.
27. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B., and Taipale, J. (2018). CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* 24, 927–930. <https://doi.org/10.1038/s41591-018-0049-z>.
28. Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* 36, 765–771. <https://doi.org/10.1038/nbt.4192>.
29. Schirotti, G., Conti, A., Ferrari, S., Della Volpe, L., Jacob, A., Albano, L., Beretta, S., Calabria, A., Vavassori, V., Gasparini, P., et al. (2019). Precise Gene Editing Preserves hematopoietic Stem Cell Function following Transient p53-Mediated DNA Damage Response. *Cell Stem Cell* 24, 551–565.e8. <https://doi.org/10.1016/j.stem.2019.02.019>.
30. Boutin, J., Rosier, J., Cappellen, D., Prat, F., Toutain, J., Pennamen, P., Bouron, J., Rooryck, C., Merlio, J.P., Lamrissi-Garcia, I., et al. (2021). CRISPR-Cas9 globin editing can induce megabase-scale copy-neutral losses of heterozygosity in hematopoietic cells. *Nat. Commun.* 12, 4922. <https://doi.org/10.1038/s41467-021-25190-6>.
31. Métais, J.Y., Doerfler, P.A., Mayuranathan, T., Bauer, D.E., Fowler, S.C., Hsieh, M.M., Katta, V., Keriwala, S., Lazzarotto, C.R., Luk, K., et al. (2019). Genome editing of HBG1 and HBG2 to induce fetal hemoglobin. *Blood Adv.* 3, 3379–3392. <https://doi.org/10.1182/bloodadvances.2019000820>.
32. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. <https://doi.org/10.1038/nature17946>.
33. Li, X., Wang, Y., Liu, Y., Yang, B., Wang, X., Wei, J., Lu, Z., Zhang, Y., Wu, J., Huang, X., et al. (2018). Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.* 36, 324–327. <https://doi.org/10.1038/nbt.4102>.
34. Wang, X., Ding, C., Yu, W., Wang, Y., He, S., Yang, B., Xiong, Y.C., Wei, J., Li, J., Liang, J., et al. (2020). Cas12a base editors induce efficient and specific editing with low DNA damage response. *Cell Rep.* 31, 107723. <https://doi.org/10.1016/j.celrep.2020.107723>.
35. Yang, L., and Chen, J. (2020). A tale of two moieties: rapidly evolving CRISPR/Cas-based genome editing. *Trends Biochem. Sci.* 45, 874–888. <https://doi.org/10.1016/j.tibs.2020.06.003>.

36. Chen, J., Yang, B., and Yang, L. (2019). To be or not to be, that is the question. *Nat. Biotechnol.* 37, 520–522. <https://doi.org/10.1038/s41587-019-0119-x>.
37. Grūnewald, J., Zhou, R., Garcia, S.P., Iyer, S., Lareau, C.A., Aryee, M.J., and Joung, J.K. (2019). Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* 569, 433–437. <https://doi.org/10.1038/s41586-019-1161-z>.
38. Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C., Wang, D., Qiu, J.L., Zhang, F., and Gao, C. (2019). Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* 364, 292–295. <https://doi.org/10.1126/science.aaw7166>.
39. Zhou, C., Sun, Y., Yan, R., Liu, Y., Zuo, E., Gu, C., Han, L., Wei, Y., Hu, X., Zeng, R., et al. (2019). Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* 571, 275–278. <https://doi.org/10.1038/s41586-019-1314-0>.
40. Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., Yuan, L., Steinmetz, L.M., Li, Y., and Yang, H. (2019). Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364, 289–292. <https://doi.org/10.1126/science.aav9973>.
41. Wang, L., Xue, W., Zhang, H., Gao, R., Qiu, H., Wei, J., Zhou, L., Lei, Y.N., Wu, X., Li, X., et al. (2021). Eliminating base-editor-induced genome-wide and transcriptome-wide off-target mutations. *Nat. Cell Biol.* 23, 552–563. <https://doi.org/10.1038/s41556-021-00671-4>.
42. Sankaran, V.G., Xu, J., Byron, R., Greisman, H.A., Fisher, C., Weatherall, D.J., Sabath, D.E., Groudine, M., Orkin, S.H., Premawardhana, A., and Bender, M.A. (2011). A functional element necessary for fetal hemoglobin silencing. *N. Engl. J. Med.* 365, 807–814. <https://doi.org/10.1056/NEJMoa1103070>.
43. Ghedira, E.S., Lecerf, L., Faubert, E., Costes, B., Moradkhani, K., Bachir, D., Galactéros, F., and Pissard, S. (2013). Estimation of the difference in HbF expression due to loss of the 5' delta-globin BCL11A binding region. *Haematologica* 98, 305–308. <https://doi.org/10.3324/haematol.2012.061994>.
44. Ginder, G.D. (2015). Epigenetic regulation of fetal globin gene expression in adult erythroid cells. *Transl. Res.* 165, 115–125. <https://doi.org/10.1016/j.trsl.2014.05.002>.
45. Traxler, E.A., Yao, Y., Wang, Y.D., Woodard, K.J., Kurita, R., Nakamura, Y., Hughes, J.R., Hardison, R.C., Blobel, G.A., Li, C., and Weiss, M.J. (2016). A genome-editing strategy to treat beta-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat. Med.* 22, 987–990. <https://doi.org/10.1038/nm.4170>.
46. Han, W., Gao, B.Q., Zhu, J., He, Z., Li, J., Yang, L., and Chen, J. (2023). Design and application of the transformer base editor in mammalian cells and mice. *Nat. Protoc.* <https://doi.org/10.1038/s41596-023-00877-w>.
47. Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* 36, 843–846. <https://doi.org/10.1038/nbt.4172>.
48. Wang, X., Li, J., Wang, Y., Yang, B., Wei, J., Wu, J., Wang, R., Huang, X., Chen, J., and Yang, L. (2018). Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. *Nat. Biotechnol.* 36, 946–949. <https://doi.org/10.1038/nbt.4198>.
49. Grūnewald, J., Zhou, R., Iyer, S., Lareau, C.A., Garcia, S.P., Aryee, M.J., and Joung, J.K. (2019). CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.* 37, 1041–1048. <https://doi.org/10.1038/s41587-019-0236-6>.
50. Liu, Y., Zhou, C., Huang, S., Dang, L., Wei, Y., He, J., Zhou, Y., Mao, S., Tao, W., Zhang, Y., et al. (2020). A Cas-embedding strategy for minimizing off-target effects of DNA base editors. *Nat. Commun.* 11, 6073. <https://doi.org/10.1038/s41467-020-19690-0>.
51. Walton, R.T., Christie, K.A., Whittaker, M.N., and Kleinstiver, B.P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368, 290–296. <https://doi.org/10.1126/science.aba8853>.
52. An, J., Zhang, C.P., Qiu, H.Y., Zhang, H.X., Chen, Q.B., Zhang, Y.M., Lei, X.L., Zhang, C.X., Yin, H., and Zhang, Y. (2023). Enhancement of the viability of T cells electroporated with DNA via osmotic dampening of the DNA-sensing cGAS-STING pathway. *Nat. Biomed. Eng.* <https://doi.org/10.1038/s41551-023-01073-7>.
53. Zhang, C.P., Qiu, H.Y., Zhang, C.X., Zhang, Y.M., Zhang, Y.Z., Yin, H., Zhang, K.Q., and Zhang, Y. (2023). Efficient non-viral delivery of macromolecules in human hematopoietic stem cells and lymphocytes. *J. Mol. Cell Biol.* 15, mjad018. <https://doi.org/10.1093/jmcb/mjad018>.
54. Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroyama, T., Miyoshi, H., Tani, K., and Nakamura, Y. (2013). Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS One* 8, e59890. <https://doi.org/10.1371/journal.pone.0059890>.
55. Liao, J., Chen, S., Hsiao, S., Jiang, Y., Yang, Y., Zhang, Y., Wang, X., Lai, Y., Bauer, D.E., and Wu, Y. (2023). Therapeutic adenine base editing of human hematopoietic stem cells. *Nat. Commun.* 14, 207. <https://doi.org/10.1038/s41467-022-35508-7>.
56. Ihry, R.J., Worringer, K.A., Salick, M.R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J., Sondey, M., Ye, C., et al. (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* 24, 939–946. <https://doi.org/10.1038/s41591-018-0050-6>.
57. Vousden, K.H., and Prives, C. (2009). Blinded by the light: the growing complexity of p53. *Cell* 137, 413–431. <https://doi.org/10.1016/j.cell.2009.04.037>.
58. Karimian, A., Ahmadi, Y., and Yousefi, B. (2016). Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage. *DNA Repair (Amst)* 42, 63–71. <https://doi.org/10.1016/j.dnarep.2016.04.008>.
59. Takaoka, A., Hayakawa, S., Yanai, H., Stoiber, D., Negishi, H., Kikuchi, H., Sasaki, S., Imai, K., Shibue, T., Honda, K., and Taniguchi, T. (2003). Integration of interferon-alpha/beta signalling to p53 responses in tumour suppression and antiviral defence. *Nature* 424, 516–523. <https://doi.org/10.1038/nature01850>.
60. Wiatrek, D.M., Candela, M.E., Sedmik, J., Oppelt, J., Keegan, L.P., and O'Connell, M.A. (2019). Activation of innate immunity by mitochondrial dsRNA in mouse cells lacking p53 protein. *RNA* 25, 713–726. <https://doi.org/10.1261/ra.069625.118>.
61. Kim, D., Lim, K., Kim, S.T., Yoon, S.H., Kim, K., Ryu, S.M., and Kim, J.S. (2017). Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat. Biotechnol.* 35, 475–480. <https://doi.org/10.1038/nbt.3852>.
62. Bae, S., Park, J., and Kim, J.S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475. <https://doi.org/10.1093/bioinformatics/btu048>.
63. Gao, R., Fu, Z.C., Li, X., Wang, Y., Wei, J., Li, G., Wang, L., Wu, J., Huang, X., Yang, L., and Chen, J. (2022). Genomic and transcriptomic analyses of prime editing guide RNA-independent off-target effects by prime Editors. *CRISPR J.* 5, 276–293. <https://doi.org/10.1089/crispr.2021.0080>.
64. Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P., et al. (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* 527, 192–197. <https://doi.org/10.1038/nature15521>.
65. U.S. National Library of Medicine. A safety and efficacy study evaluating CTX001 in subjects with transfusion-dependent β -thalassemia. <https://ClinicalTrials.gov/show/NCT03655678>.
66. U.S. National Library of Medicine. A safety and efficacy study evaluating CTX001 in subjects with severe sickle cell disease. <https://ClinicalTrials.gov/show/NCT03745287>.
67. CRISPR Therapeutics. Efficacy and safety of a single dose of exagamglogene autotemcel for transfusion-dependent β -thalassemia and severe sickle cell disease. <http://ir.crisprtx.com/static-files/c85b519-0d4a-43e6-9b38-1176f547364b>.

68. Yang, Y., Xu, Z., He, C., Zhang, B., Shi, Y., and Li, F. (2019). Structural insights into the recognition of gamma-globin gene promoter by BCL11A. *Cell Res.* *29*, 960–963. <https://doi.org/10.1038/s41422-019-0221-0>.
69. Ferrari, S., Jacob, A., Beretta, S., Unali, G., Albano, L., Vavassori, V., Cittaro, D., Lazarevic, D., Brombin, C., Cugnata, F., et al. (2020). Efficient gene editing of human long-term hematopoietic stem cells validated by clonal tracking. *Nat. Biotechnol.* *38*, 1298–1308. <https://doi.org/10.1038/s41587-020-0551-y>.
70. Newby, G.A., Yen, J.S., Woodard, K.J., Mayuranathan, T., Lazzarotto, C.R., Li, Y., Sheppard-Tillman, H., Porter, S.N., Yao, Y., Mayberry, K., et al. (2021). Base editing of haematopoietic stem cells rescues sickle cell disease in mice. *Nature* *595*, 295–302. <https://doi.org/10.1038/s41586-021-03609-w>.
71. Lei, L., Chen, H., Xue, W., Yang, B., Hu, B., Wei, J., Wang, L., Cui, Y., Li, W., Wang, J., et al. (2018). APOBEC3 induces mutations during repair of CRISPR-Cas9-generated DNA breaks. *Nat. Struct. Mol. Biol.* *25*, 45–52. <https://doi.org/10.1038/s41594-017-0004-6>.
72. Krokan, H.E., and Bjørås, M. (2013). Base excision repair. *Cold Spring Harb. Perspect. Biol.* *5*, a012583. <https://doi.org/10.1101/cshperspect.a012583>.
73. Vaidyanathan, S., Azizian, K.T., Haque, A.K.M.A., Henderson, J.M., Hendel, A., Shore, S., Antony, J.S., Hogrefe, R.I., Kormann, M.S.D., Porteus, M.H., and McCaffrey, A.P. (2018). Uridine depletion and chemical modification increase Cas9 mRNA activity and reduce immunogenicity without HPLC purification. *Mol. Ther. Nucleic Acids* *12*, 530–542. <https://doi.org/10.1016/j.omtn.2018.06.010>.
74. Wang, G., Cheng, R., Chen, Q., Xu, Y., Yu, B., Zhu, B., Yin, H., and Xia, H. (2022). mRNA produced by VSW-3 RNAP has high-level translation efficiency with low inflammatory stimulation. *Cell Insight* *1*, 100056. <https://doi.org/10.1016/j.cellin.2022.100056>.
75. Yang, L.Z., Gao, B.Q., Huang, Y., Wang, Y., Yang, L., and Chen, L.L. (2022). Multi-color RNA imaging with CRISPR-Cas13b systems in living cells. *Cell Insight* *1*, 100044. <https://doi.org/10.1016/j.cellin.2022.100044>.
76. Zhang, H.X., Zhang, C., Lu, S., Tong, X., Zhang, K., Yin, H., and Zhang, Y. (2023). Cas12a-based one-pot SNP detection with high accuracy. *Cell Insight* *2*, 100080. <https://doi.org/10.1016/j.cellin.2023.100080>.
77. Chalumeau, A., Frati, G., Magrin, E., and Miccio, A. (2021). Reverse Phase-high-performance Liquid Chromatography (RP-HPLC) analysis of globin chains from human erythroid cells. *Bio Protoc.* *11*, e3899. <https://doi.org/10.21769/BioProtoc.3899>.
78. Shi, Y.J., Duan, M., Ding, J.M., Wang, F.Q., Bi, L.L., Zhang, C.X., Zhang, Y.Z., Duan, J.Y., Huang, A.H., Lei, X.L., et al. (2022). DNA topology regulates PAM-Cas9 interaction and DNA unwinding to enable near-PAMless cleavage by thermophilic Cas9. *Mol. Cell* *82*, 4160–4175.e6. <https://doi.org/10.1016/j.molcel.2022.09.032>.
79. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* *20*, 1297–1303. <https://doi.org/10.1101/gr.107524.110>.
80. Wilm, A., Aw, P.P., Bertrand, D., Yeo, G.H., Ong, S.H., Wong, C.H., Khor, C.C., Petric, R., Hibberd, M.L., and Nagarajan, N. (2012). LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res.* *40*, 11189–11201. <https://doi.org/10.1093/nar/gks918>.
81. Kim, S., Scheffler, K., Halpern, A.L., Bekritsky, M.A., Noh, E., Källberg, M., Chen, X., Kim, Y., Beyter, D., Krusche, P., and Saunders, C.T. (2018). Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods* *15*, 591–594. <https://doi.org/10.1038/s41592-018-0051-x>.
82. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* *30*, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
83. Narzisi, G., O’Rawe, J.A., Iossifov, I., Fang, H., Lee, Y.H., Wang, Z., Wu, Y., Lyon, G.J., Wigler, M., and Schatz, M.C. (2014). Accurate de novo and transmitted indel detection in exome-capture data using microassembly. *Nat. Methods* *11*, 1033–1036. <https://doi.org/10.1038/nmeth.3069>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
BL21 (DE3) Chemically Competent Cell	TransGen	Cat# CD601-02
Antibodies		
Pacific Blue™ anti-human CD34 Antibody	Biolegend	Cat# 343512; RRID:AB_1877197
FITC anti-human CD235ab Antibody	Biolegend	Cat# 306609; RRID:AB_756045
APC/Cyanine7 anti-human CD45 Antibody	Biolegend	Cat# 368516; RRID:AB_2566376
APC anti-mouse CD45 Antibody	Biolegend	Cat# 103112; RRID:AB_312977
PerCP/Cyanine5.5 anti-human CD33 Antibody	Biolegend	Cat# 366616; RRID:AB_2566418
PE/Cyanine7 anti-human CD19 Antibody	Biolegend	Cat# 302216; RRID:AB_314246
PE anti-human CD3 Antibody	Biolegend	Cat# 300408; RRID:AB_314062
TruStain FcX™ (anti-mouse CD16/32) Antibody	Biolegend	Cat# 101320; RRID:AB_1574975
Human TruStain FcX™ (Fc Receptor Blocking Solution)	Biolegend	Cat# 422302; RRID:AB_2818986
Fetal Hemoglobin Monoclonal Antibody (HBF-1), FITC	Thermo Fisher Scientific	Cat# MHFH01; RRID: AB_10375007
Cas9 (<i>S. pyogenes</i>) (7A9-3A3) Mouse mAb	Cell Signaling Technology	Cat# 14697S; RRID:AB_2750916
Anti- α -Tubulin antibody	Sigma	Cat# T6199-100UL; RRID:AB_477583
Histone H3 (D2B12) XP® Rabbit mAb	Cell Signaling Technology	Cat# 4620; RRID:AB_1904005
Normal Rabbit IgG	Cell Signaling Technology	Cat# 2729; RRID:AB_1031062
Anti-HA tag antibody	Abcam	Cat# ab91110; RRID:AB_307019
Biological samples		
HSPC	Milestone® Biotechnologies	Cat# CB34-4C-P
Chemicals, peptides, and recombinant proteins		
DMEM	Thermo Fisher Scientific	Cat# 10566
SFEM	Stemcell	Cat# 09650-500MI
IMDM	Thermo Fisher Scientific	Cat# 12440053-500MI
FBS	Thermo Fisher Scientific	Cat# 16000-044
penicillin-streptomycin	Invitrogen	Cat# 15140122
puromycin	InvivoGen	Cat# ant-pr-1
dexamethasone	Sigma	Cat# 50-02-2
doxycycline	Sigma	Cat# D9891
Recombinant human stem cell factor (SCF)	peprotech	Cat# 300-07
Recombinant human thrombopoietin (TPO)	peprotech	Cat# 300-18
Recombinant human Flt3-ligand (Flt3-L)	peprotech	Cat# 300-19
GemCell™ U.S. Human Serum AB	Gemini	Cat# 100-512
Recombinant human IL-3	peprotech	Cat# 200-03
Holo-transferrin human	Sigma-Aldrich	Cat# T0665
Heparin	Sigma-Aldrich	Cat# H3149
Recombinant human Erythropoietin (EPO)	Peprtech	Cat# 100-64
Hydrocortisone	R&D	Cat# 4093-50mg
Recombinant human insulin	Solarbio Life Science	Cat# I8830
Glutaraldehyde	Sigma	Cat# G5882
Triton X-100	BBI Life Sciences	Cat# A600198-0500
T7 RNA polymerase	New England Biolabs	Cat# M0251L
mRNA Cap 2'-O-Methyltransferase	New England Biolabs	Cat# M0366S
RQ1 RNase-Free DNase	Promega	Cat# M6101

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cellulose	Sigma	Cat# C6288-250G
isopropylthio- β -D-galactoside (IPTG)	Sinopharm Chemical ReagentCo., Ltd	Cat# 367-93-1
1M Tris-HCl Solution, pH 7.5, Sterile	Sangon Biotech	Cat# B548124-0500
Sodium chloride (NaCl)	Sangon Biotech	Cat# 7647-14-5
TCEP	Sangon Biotech	Cat# 51805-45-9
HEPES	Sangon Biotech	Cat# 7365-45-9
KCl	Sangon Biotech	Cat# 7447-40-7
glycerol	ABCONE	Cat# G46055-100ML
Imidazole	Sangon Biotech	Cat# 288-32-4
ZnSO ₄	Sangon Biotech	Cat# 7446-20-0
EDTA	Sangon Biotech	Cat# E0322-500g
Exonuclease I	Thermo Fisher Scientific	Cat# EN0581
MgCl ₂	Sangon Biotech	Cat# 7791-18-6
DTT	Sangon Biotech	Cat# 3483-12-3
PrimeSTAR® HS DNA Polymerase	Takara	Cat# R010B
Poly(I:C) HMW	invivogen	Cat# tlr1-piclv
PBS	meilunbio	Cat# MA0015-1
CRISPR-Cas9 Protein, <i>Streptococcus pyogenes</i> , Recombinant	SinoBiological	Cat# 40572-A08B
Tris-MOPS-SDS Running Buffer Powder	GenScript	Cat# M00138
Transfer Buffer (10 \times)	Epizyme Biotech	Cat# PS109S
EZ-Buffers H 10 \times TBST Buffer	Sangon Biotech	Cat# C520009
Micrococcal Nuclease	Cell SignalingTechnology	Cat# 10011
Critical commercial assays		
Human CD235a (Glycophorin A) MicroBeads	Miltenyi Biotec Inc.	Cat# 130-050-501
Human CD34 MicroBeads kit	Miltenyi Biotec Inc.	Cat# 130-046-702
FastScan™ Cas9 (<i>S. pyogenes</i>) ELISA Kit	Cell Signaling Technology	Cat# 29666C
Lipofectamine® LTX & PLUS™ Reagent	Thermo Fisher Scientific	Cat# 15338100
Lipofectamine™ MessengerMAX™ Reagent	Thermo Fisher Scientific	Cat# LMRNA003
QuickExtract™ DNA Extraction Solution	Epicentre	Cat# QE09050
Vaccinia Capping System	New England Biolabs	Cat# M2080S
Monarch® RNA Cleanup Columns	New England Biolabs	Cat# T2047L
Murine RNase inhibitor	Vazyme	Cat# R301-01
FastPure Gel DNA Extraction Mini Kit	Vazyme	Cat# DC301-01
P3 Primary Cell 4D-Nucleofector X Kit S	Lonza	Cat# V4XP-3032
TransZol Up Plus RNA Kit	TransGen	Cat# ER501-01-V2
HiScript® III All-in-one RT SuperMix Perfect for qPCR	Vazyme	Cat# R333-01
ChamQ Universal SYBR qPCR Master Mix	Vazyme	Cat# Q711-03
NEBNext Ultra II FS DNA Library Prep Kit for Illumina	NEB	Cat# E7103L
iQuant™ dsDNA HS Assay Kit	ABP Biosciences	Cat# N021-ABP
FastPure Cell/Tissue DNA Isolation Mini Kit	Vazyme	Cat# DC102-01
PureLink™ RNA Micro Kit	Thermo Fisher Scientific	Cat# 12183016
TruSeq® Stranded Total RNA Library Prep Globin (96 Samples)	Illumina	Cat# 20020613
KAPA RiboErase Kit	HMR	Cat# KK8483
KAPA RNA HyperPrep Kit	HMR	Cat# KK8504
ExpressPlus™ PAGE Gel, 10 \times 8, 4-20%, 15 wells	GenScript	Cat# M42015C
Pierce™ ECL Plus Western Blotting Substrate	Thermo Fisher Scientific	Cat# 32134

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SimpleChIP® Enzymatic Chromatin IP Kit (Magnetic Beads)	Cell Signaling Technology	Cat# 9003
Deposited data		
DNA-amplicon-seq data	This paper	SRA: PRJNA961909 SRA: PRJNA1013694
Whole-genome-seq (WGS) data	This paper	SRA: PRJNA960629
RNA-Seq data	This paper	SRA: PRJNA1013210
Raw FACS data	This paper (Mendeley Data)	https://doi.org/10.17632/sf7zywbswy.4
Experimental models: Cell lines		
293FT	Thermo Fisher Scientific	Cat# R70007
HUDEP-2	Riken	Cat# RCB4557
293FT ^{A3-/-} cells	Gao et al. ⁶³	N/A
Oligonucleotides		
Chemically modified sgRNA sequences, See Table S1	This paper	N/A
Primers for real-time PCR, See Table S1	This paper	N/A
Primers for amplicon deep sequencing, See Table S1	This paper	N/A
Plasmids		
pUC57-sgRNA-MS2-U6	Addgene	171,694
pU6-ccdB-boxB-tBE-V5-mA3	Addgene	171,693
pGEX-4T1 vector	Cytiva	Cat# 28954549
Software and algorithms		
FlowJo 10.7.1	BD Biosciences	https://www.flowjo.com
GraphPad Prism (9.0)	GraphPad	https://www.graphpad.com/
CFBI	Wang et al. ⁴¹	https://github.com/YangLab/CFBI/releases/tag/V1.0.0
BEIDOU	Wang et al. ⁴¹	https://github.com/YangLab/BEIDOU/releases/tag/V1.0.0
RADAR	Wang et al. ³⁴	https://github.com/YangLab/RADAR/releases/tag/V1.0.0
Other		
Lonza 4D-Nucleofector	Lonza	N/A
QuantStudio 7 Flex Real-Time PCR System	Thermo Fisher Scientific	N/A
FACSAria III	BD Biosciences	N/A
Novocyte	Agilent	N/A
Typhoon FLA 9500 biomolecular imager	GE Healthcare	N/A
ÅKTA pure™ chromatography system	cytiva	N/A
NexeraX2 HPLC System	Shimadzu	N/A
D-10™ Hemoglobin Analyzer	Bio-Rad	N/A
AutoMACSpro	Miltenyi Biotec Inc.	N/A
SpectraMax® M5e multimode plate reader	Molecular Devices	N/A
Amersham Imager 680 Chemiluminescent Imaging System	GE Healthcare	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ying Zhang (ying.zhang84@whu.edu.cn)

Materials availability

All unique reagents generated in this study are available from the [lead contact](#) with a completed Materials Transfer Agreement.

Data and code availability

The original DNA amplicon sequencing data can be accessed in the NCBI Sequence Read Archive (accession number: PRJNA961909, PRJNA1013694) and is included in the [key resources table](#). The original whole-genome sequencing (WGS) data can be accessed in the NCBI Sequence Read Archive (accession number: PRJNA960629) and RNA sequencing data from this study can be accessed in the NCBI Sequence Read Archive (accession number: PRJNA1013210) and is included in the [key resources table](#).

The code used in this study has been previously published and included in the [key resources table](#). The CFBI is the custom Perl and Shell scripts for calculating frequencies of base substitutions and indels. The BEIDOU toolkit, which is used to call high-confidence base substitution or indel events from WGS data. The RADAR pipeline is used to detect and visualize all possible twelve types of RNA editing events from RNA-seq data.

Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Patient and healthy donor cells

Approval to collect CD34⁺ cells from patients with β -thalassemia was approved by the human subject institutional review boards at the Children's Hospital of Fudan University. All patients provided informed consent. CD34⁺ cells from all three patients were mobilized by G-CSF, with plerixafor added for the 3rd patient on the 4th and 5th day of mobilization. The apheresis materials were processed according to the protocol of Miltenyi CD34 Microbead kit (Miltenyi Biotec). The genotypes of the three β^0/β^0 donors as shown: the first donor: $\alpha\alpha^{-4.2}/\alpha\alpha$, $\beta^{CD41-42(-TTCT)}/\beta^{CD41-42(-TTCT)}$; the second donor: $\alpha\alpha^{-3.7}/\alpha\alpha$, $\beta^{CD41-42(-TTCT)}/\beta^{CD41-42(-TTCT)}$; the third donor: $\alpha\alpha/\alpha^{-3.7}\alpha^{Hb\ Westmead}$, $\beta^{CD41-42(-TTCT)}/\beta^{CD41-42(-TTCT)}$.

Frozen healthy CD34⁺ cells were purchased from Allcells and SAILY BIO (China).

Mice

NOD.Cg-Kit^{W-41J} Tyr⁺ Prkdc^{scid} Il2rg^{tm1Wjl} (NBSGW) mice were purchased from Jackson Laboratory (Stock 026622). 7-8 week old female mice were used in this study. All animal experiments were performed according to the protocol approved by the Animal Care and Ethical Committee at Wuhan University.

METHOD DETAILS

Cell line culture and plasmid transfection

293FT cell line was maintained in DMEM supplemented with 10% FBS and 1% penicillin-streptomycin at 37 °C with 5% CO₂. Cells are regularly tested to exclude mycoplasma contamination. HUDEP-2 cells was cultured in SFEM supplemented with 1 mM dexamethasone, 1 μ g mL⁻¹ doxycycline, 3 U mL⁻¹ EPO and 50 ng mL⁻¹ SCF at a density of 0.5 \times 10⁶ cells per mL.

To transfect plasmids into 293FT cells, 1 \times 10⁵ cells per well was seeded in a 24-well-plate 24 h prior to transfection. For trans-former base editor (tBE), 0.5 μ g tBE-V5-mA3 and 0.5 μ g pEFS-nSpCas9 or 0.5 μ g pEFS-nSpCas9-spG expression plasmids were co-transfected into 293FT cells using Lipofectamine® LTX & PLUS™ Reagent. For canonical BEs, 0.5 μ g of pCMV-hA3A-BE3, pCMV_BE4max, CE¹⁰⁴⁸⁻¹⁰⁶³-CBE, pCMV-A3A(N57Q)-BE3, pCMV-hA3A-BE3-spG, pCMV_BE4max-spG, CE¹⁰⁴⁸⁻¹⁰⁶³-CBE-spG or pCMV-A3A(N57Q)-BE3-spG expression vectors were co-transfected with 0.25 μ g sgRNA expression vector for each well. After 24 h, 4 μ g mL⁻¹ puromycin was added into the medium. Genomic DNA was isolated at 72 h post-transfection using QuickExtract™ DNA Extraction Solution for subsequent sequencing analysis.

CD34⁺ HSPC culture and *in vitro* erythroid differentiation

Purified human CD34⁺ HSPCs from mobilized peripheral blood were thawed and cultured in SFEM supplemented with 100 ng mL⁻¹ human SCF, 100 ng mL⁻¹ human TPO and 100 ng mL⁻¹ human Flt3-L. HSPCs were seeded and maintained at the density of 5 \times 10⁵ mL⁻¹. *In vitro* erythroid differentiation HSPCs was performed at 24 h post-electroporation following the three-step culture protocol.²³ In brief, cells were transferred to erythroid differentiation medium composed of IMDM supplemented with 5% human solvent detergent pooled plasma AB, 1% penicillin-streptomycin, 10 μ g mL⁻¹ recombinant human insulin, 3 IU mL⁻¹ heparin, 3 IU mL⁻¹ EPO. In the first step (days 1–7), additional supplements including 200 μ g mL⁻¹ holo-transferrin human, 10 ng mL⁻¹ human SCF, 1 ng mL⁻¹ human IL-3 and 10⁻⁶ M hydrocortisone was added to the differentiation medium. In the second step (days 7–11), the same supplements as in step 1 were added, except that human IL-3 was withdrawn. In the third step (days 11–21), the concentration of human holo-transferrin was increased to 1 mg mL⁻¹ and human SCF was removed compared with the medium in step 2. The density of cell culture varies with different stages of erythroid differentiation from 1 \times 10⁵ to 5 \times 10⁶ cells per milliliter.

For HUDEP-2 cell, the erythroid differentiation was conducted as described.¹³ In brief, HUDEP-2 cells were cultured in IMDM with 50 ng mL⁻¹ human SCF, 3 U mL⁻¹ EPO, 2.5% FBS, 250 μ g mL⁻¹ holo-transferrin, 1% penicillin-streptomycin, 10 ng mL⁻¹ heparin, 10 μ g mL⁻¹ insulin and 1 μ g mL⁻¹ doxycycline. Cells were maintained in the differentiation medium for five days prior to downstream globin measurement. To isolate a single clone of HUDEP-2 cell, the edited HUDEP-2 cells were sorted into 96-well plates with one cell per well and cultured for 2 weeks. Single clones were genotyped using sanger sequencing.

gRNA, mRNA preparation and electroporation

Chemically modified sgRNA (2'-O-methyl 3' phosphorothioate modifications in the first and last three nucleotides) was synthesized from GenScript. The gDNA sequences used in this study were listed in [Table S1](#). All mRNAs were transcribed *in vitro* using T7 RNA polymerase. The transcription template was amplified by PCR and a 110-base poly A tail was included in the transcribed PCR product. RQ1 RNase-Free DNase was used to remove the DNA template, and the resulting mRNA was purified using Monarch RNA Cleanup Kit. The mRNA was heat-denatured at 65 °C for 5 min before capping with a Cap-1 structure using vaccinia capping system and mRNA Cap 2'-O-Methyltransferase. The mRNA was purified by cellulose purification, as described.^{74–76} Purified mRNA was analyzed by agarose gel electrophoresis and was stored at –80°C.

CD34⁺ HSPCs were electroporated with mRNA or RNP 24 h post-thaw. Electroporation was performed using Lonza 4D Nucleofector. For 20 µL Nucleocuvette Strips, 0.2 million HSPCs were resuspended in 20 µL B1mix buffer and RNP complex (6 µg SpCas9 protein and 6 µg sgRNA) were added. For mRNA electroporation, unless elsewhere mentioned, mRNA and sgRNA were mixed at the following ratios to the cell suspensions: tBE (3 µg D10A, 4.5 µg tBE-V5-mA3, 6 µg sgRNA, 9 µg hsgRNA), hA3A-BE3 (4.4 µg hA3A-BE3, 4.8 µg sgRNA) and ABE8e (4.4 µg ABE8e, 4.8 µg sgRNA). Electroporation program CM137 and buffer B1mix was used for both RNA and RNP delivery as described.^{52,53} Buffer B1mix is composed of 3.6mM KCl, 10.8 mM MgCl₂, 62.5 mM Na₂HPO₄, 24.3 mM NaH₂PO₄, 10 mM sodium succinate, 18 mM mannitol, 2.8 mM i-inositol, 0.88 mM Ca(NO₃)₂, 0.55 mM sodium pyruvate, 21.8 mM D-glucose, 2.19% GlutaMAX, pH 7.2. The buffer can be stored at 4 °C for two weeks.

HUDEP-2 cell were electroporated with mRNA using Lonza 4D Nucleofector, 0.2 million HUDEP-2 cells were resuspend in 20 µL Nucleofector™ Solutions (P3 Primary Cell 4D-Nucleofector X Kit S), mRNA and sgRNA were mixed at the following ratios to the cell suspensions: tBE (2 µg D10A, 3 µg tBE-V5-mA3, 4 µg sgRNA, 6 µg hsgRNA), other canonical BEs used the same dose as HSPCs used.

Real-time qPCR of globin expression

Total RNA was isolated using *TransZol* Up Plus RNA Kit and reverse transcribed using HiScript III All-in-one RT SuperMix Perfect for qPCR. The qPCR reactions were performed in QuantStudio 7 Flex Real-Time PCR System with ChamQ Universal SYBR qPCR Master Mix, cDNA and individual primers. *CAT* was employed as an internal reference as described previously.²³ *GAPDH* was used as a reference transcript when measuring p21 mRNA and RNA immune pathway. Primer sequences are provided in [Table S1](#).

HPLC analysis of globin chains and hemoglobin tetramers

Globin chains of *in vitro* differentiated erythroid cells were analyzed via Reverse Phase-High Performance Liquid Chromatography (RP-HPLC) method following established protocol.⁷⁷ Briefly, at least 250,000 *in vitro* differentiated erythroid cells were lysed in ice-cold MilliQ deionized water, and HPLC samples prepared from the lysate were separated with a 250 × 4.6-mm, 3.6-µm Aeris Widepore column (Phenomenex) over a 95 min gradient on a NexeraX2 HPLC System. The HPLC system was extensively cleaned with blank injections between samples to minimize any carryover. The relative abundance of individual globin chains was monitored by light absorbance at 220 nm and calculated as the area under the curve (AUC). Hemoglobin tetramers were separated and analyzed with D-10™ Hemoglobin Analyzer following manufacturer's manual.

Xenotransplantation in NBSGW mice

NOD.Cg-Kit^{W-41J} Tyr⁺ Prkdc^{scid} Il2rg^{tm1Wjl} (NBSGW) mice were purchased from Jackson Laboratory (Stock 026622). Female NBSGW mice (7–8 weeks) were intravenously injected with 0.5 million CD34⁺ HSPCs. After 16 weeks of engraftment, bone marrow was collected from NBSGW mice. For flow cytometry analysis, bone marrow cells were incubated with Human TruStain FcX™ (Fc Receptor Blocking Solution) (BioLegend,1:20) and TruStain FcX™ (anti-mouse CD16/32,1:50) antibody for 10 min on ice. After Fc receptor blocking, bone marrow cells were stained with APC/Cyanine7 anti-human CD45 Antibody (Biolegend,1:20), APC anti-mouse CD45 Antibody (Biolegend,1:100), Pacific Blue™ anti-human CD34 Antibody (Biolegend,1:50), FITC anti-human CD235ab Antibody (Biolegend,1:50), PerCP/Cyanine5.5 anti-human CD33 Antibody (Biolegend,1:200), PE/Cyanine7 anti-human CD19 Antibody (Biolegend,1:100) and PE anti-human CD3 Antibody (Biolegend,1:50) at room temperature for 15 minutes in the dark. After washing with 1% BSA dPBS, cells were analyzed by Novocyte. Transplantation efficiency was calculated as hCD45⁺/(mCD45⁺ add hCD45⁺). B cells, myeloid cells, T cells, HSPCs and erythroid cells were gated respectively on the following markers: hCD45⁺hCD19⁺, hCD45⁺hCD33⁺, hCD45⁺hCD3⁺, hCD45⁺hCD34⁺ and hCD45⁺mCD45⁺hCD235a⁺. B cells and myeloid cells were sorted using BD FACSAria III following the same gating strategy. hCD235a⁺ erythroid cells and hCD34⁺ HSPCs were isolated by Miltenyi autoMACSpro with human CD235a microBeads (Miltenyi Biotec Inc.) and human CD34 microBeads (Miltenyi Biotec Inc.) according to the manufactory protocol.

Flow cytometry for F-cell

During *in vitro* erythroid differentiation, cells collected on day 11 were analyzed for their expression of fetal hemoglobin by FACS. Cells were fixed in 0.05% Glutaraldehyde for 10 min at room temperature (16–25°C) and then permeabilized in 0.1% Triton X-100 for 5 min at room temperature. Cells were stained with anti-human Fetal Hemoglobin Monoclonal Antibody (1:20) and analyzed by Novocyte.

Protein expression and purification

The recombinant plasmid pET28a-SpCas9⁷⁸ was transformed into *Escherichia coli* BL21 (DE3) cells for protein expression. The cells were cultured in Luria-Bertani (LB) medium at 37 °C until the OD₆₀₀ reached 0.6~0.8. The cultures were then incubated for 16–20 h at 16 °C in the presence of 0.2 mM IPTG before being harvested by centrifugation at 5,000g for 5 min. The cells were then re-suspended in buffer A (20 mM Tris, pH 8.0, 500 mM NaCl, 1 mM TCEP) and lysed by sonication. After centrifugation, the clarified cell lysate was loaded onto a Ni-NTA column (Cytiva) pre-equilibrated with buffer A. The column was then extensively washed with buffer A and eluted with buffer B (20 mM HEPES, pH 7.5, 100 mM KCl, 1 mM TCEP, 10% glycerol, 200 mM Imidazole). Elution fractions were then further purified on a HiTrap SP Sepharose column (Cytiva) before being loaded onto a Superdex 200 increase 10/300 GL column (Cytiva) in buffer C (20 mM HEPES, pH 7.5, 150 mM KCl and 1 mM TCEP). The purified protein was then concentrated, aliquoted and stored at -80 °C until further use.

The DNA fragment encoding Znf4-6 (residues 731–835) of B-cell lymphoma/leukemia 11A (BCL11A) (UniProt ID: Q9H165) was synthesized (GenScript) and cloned into a modified pGEX-4T1 vector (Cytiva) wherein the thrombin protease site was replaced with a HRV 3C protease site. The recombinant plasmid was then transformed into *Escherichia coli* BL21 (DE3) cells for protein expression. The cells were cultured in LB medium supplemented with 100 μM ZnSO₄ at 37 °C until the OD₆₀₀ reached 1.2. The cultures were then incubated for another 24 h at 16 °C in the presence of 0.2 mM IPTG before being harvested by centrifugation at 5,000g for 5 min. The cells were then re-suspended in buffer D (20 mM Tris, pH 7.5, 1 M NaCl) and lysed with a NANO homogenizer (ATS Engineering Limited) at 800 bar, 4 °C. After centrifugation, the clarified cell lysate was incubated with glutathione Sepharose resin (Cytiva) pre-equilibrated with buffer D, and the GST-tagged proteins were eluted by buffer D supplemented with 30 mM reduced glutathione. Elution fractions were then pooled and treated with HRV 3C protease to remove the N-terminal GST tag. The cleaved protein was further purified via size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column (GE Healthcare) in buffer D. The purified protein was concentrated, dialyzed against buffer E (20 mM Tris, pH 7.5, 150 mM NaCl), and stored at -80 °C until further use.

Electrophoretic mobility shift assay (EMSA)

Oligonucleotides used in the EMSA were listed in [Table S1](#). FAM-labeled oligos were incubated with their complimentary strand at 1:10 molar ratio in annealing buffer (10 mM Tris, pH 7.5, 50 mM NaCl, 1 mM EDTA) at 95 °C for 3 min and then slowly cooled to 25 °C to generate DNA duplex. Exonuclease I was used to digest single strand DNA. 5 nM FAM-labeled DNA duplex was then incubated with BCL11A Znf4-6 protein at indicated concentrations ranging from 0 μM to 5 μM for 1 h on ice in binding buffer (10 mM Tris, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM MgCl₂, 1 mM DTT, 5% glycerol). Samples were resolved on 6% native PAGE gel through electrophoresis at 120 V for 3 h at 4 °C. Electrophoretic bands were detected using a biomolecular imager (GE Healthcare).

Quantification of Cas9 expression level

HSPCs electroporated with indicated editors were harvested at different time points. Cas9 expression was determined using the FastScan™ Cas9 ELISA kit (Cell Signaling technology) according to the protocol. Recombinant Cas9 protein (SinoBiological) was used as quantitative reference standards. Data was recorded using SpectraMax® M5e multimode plate reader.

To quantify Cas9 expression of different base editors in 293FT cells, transfected cells were harvested at 72 h post-transfection and lysed in protein loading buffer. The denatured sample was loaded into SDS-PAGE gel and transferred onto nitrocellulose membrane. Antibody against Cas9 (Cell Signaling technology), and alpha-tubulin (Sigma) were used to quantify the expression level.

Chromatin-immunoprecipitation quantitative PCR (ChIP-qPCR)

ChIP experiments were performed in HUDEP-2 cells using isolated single clone (-114T/-115T). Cells were electroporated with PC3.1 3×HA-BCL11A plasmid and collected at 48 h post-electroporation for ChIP analysis using SimpleChIP Enzymatic Chromatin IP Kit (Cell Signaling Technology). In brief, 5 × 10⁶ cells were cross-linked with 1% formaldehyde for 10 min at room temperature, and the reaction was quenched with glycine at a final concentration of 125 mM. DNA was fragmented using Micrococcal Nuclease (Cell Signaling Technology) and ultrasonication. Immunoprecipitation was performed using antibodies against Histone H3 (positive control), rabbit IgG (negative control) and HA-tag (experiment) DNA was eluted, purified and analyzed for real-time qPCR using the SYBR™Green PCR Master Mix (Vazyme) as described previously.²⁴ Primer sequences used for ChIP-qPCR are provided in [Table S1](#).

Droplet Digital PCR (ddPCR)

The quantification of the 4.9-kb deletion was carried out using ddPCR (Bio-rad). Reference primers/probe targeting the hVEGFA gene was used as a DNA loading control. Wildtype primers/probe targeting the HBG1-HBG2 deletion region was designed to quantify non-deleted frequency. The frequency of deletion was calculated by normalizing the ratio of the wildtype probe with reference probe against the average ratio observed in control samples. Primers and probe sequences used in ddPCR are listed in [Table S1](#).

DNA library preparation and amplicon sequencing

To quantify the base editing efficiency, PrimeSTAR® HS DNA Polymerase was used to amplify the sequence of target site from the genomic DNA. The PCR primers used are listed in [Table S1](#). PCR product was gel purified using FastPure Gel DNA Extraction Mini Kit. DNA library was prepared following the manufacturer's instruction manual (NEBNext Ultra II FS DNA Library Prep Kit, NEB). The amplicon-containing library was quantified using the iQuant™ dsDNA HS Assay Kit, and the library was sequenced with the Illumina NovaSeq 6000 Sequencing System (2 × 150).

Libraries preparation for genome-wide and transcriptome-wide off-target analysis

To determine gRNA-independent genome-wide DNA off-target mutations, 293FT^{A3-/-} cells⁶³ were co-transfected with indicated editors and eGFP mRNA using Lipofectamine™ MessengerMAX™. Cells were sorted into single cell after 48 h post-transfection. Cells were expanded in culture for approximately three weeks. Genomic DNA was extracted from single clone using FastPure Cell/Tissue DNA Isolation Mini Kit, and the genotype for each clone was analyzed by Sanger sequencing. Indexed DNA libraries were prepared using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina. The Illumina NovaSeq 6000 Sequencing System (2 × 150) was used to generate a total of 1 Tb whole-genome sequencing (WGS) data. The average coverage of sequencing data generated for each transfected 293FT^{A3-/-} single cell clone sample was 14x.

To determine transcriptome-wide off-target mutations, CD34⁺ HSPCs were thawed and electroporated with indicated editors. Total RNA was isolated using PureLink™ RNA Micro Kit at 48 h post-electroporation. RNA library was prepared using TruSeq® Stranded Total RNA Library Prep Globin (96 Samples). After rRNA depletion using the KAPA RiboErase Kit, RNA-seq libraries were prepared using the KAPA RNA HyperPrep Kit. Size-selected libraries were sequenced with the Illumina NovaSeq 6000 Sequencing System (2 × 150).

QUANTIFICATION AND STATISTICAL ANALYSIS

Amplicon sequencing data analysis

FastQC (v0.11.8, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, parameters: default) was used to evaluate the raw read qualities. For paired-end sequencing, only R1 reads were used. Adaptor sequences and read sequences with Phred quality score lower than 30 were trimmed. Trimmed reads were then mapped to target sequences using the BWA-MEM algorithm (BWA v0.7.17). Base substitution and indel frequencies at on-target sites were calculated using methods reported in previously published literature.^{33,41,48} In brief, base substitutions at every position of the target sites were counted using at least 1000 independent reads. Base substitution frequencies were calculated using the published CFBI pipeline as: count of reads with substitution at the target base / count of reads covering the target base. In Figure 4, base substitutions were measured by CRISPResso2 (--quantification_window_size 50 --quantification_window_center -10 -q 30 -s 25). The windows used to analysis indels was set as ±50bp of the cleavage site. The bar plot of base editing efficiency was draw by ggplot2.

To measure unintended indel frequency, reads aligned in the region spanning from 5 nucleotides upstream and downstream to the nicking site (total 10 nt) were selected as the estimated region. Unintended indel frequencies for base substitution were calculated as: count of reads containing at least one unintended inserted and/or deleted nucleotide / count of total reads aligned in the estimated region.

Analysis of gRNA-dependent off-target mutation

To determine gRNA-dependent off-target mutations, CasOFFinder⁶² was used to predict potential target sites with the following parameters: 1) NGG/NGC PAM, 2) four or fewer overall mismatches, or three or fewer mismatches that allow G:U wobble base pairings with guide RNA. Candidate sites were PCR amplified and followed by amplicon sequencing to validate. Information of off-target sites are listed in Table S2.

Whole-genome sequencing data analysis

The BEIDOU⁴¹ toolkit was employed to call high-confidence base substitution or indel events that could be identified by all three different callers, GATK,⁷⁹ Lofreq,⁸⁰ and Strelka2.⁸¹ Briefly, to reduce the impact of varying sequence depth among samples, 120M reads were randomly sampled by Seqtk (v1.3, <https://github.com/lh3/seqtk>, parameters: sample -s100 120000000) from raw data for further analyses. After quality control by FastQC (parameters: default), WGS DNA-seq reads were trimmed by Trimmomatic⁸² (v0.38, parameters: ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) to remove low-quality read sequence. BWA-MEM algorithm (v0.7.17, parameters: default) was used to map clean reads to the human reference genome (hg38). Samtools (v1.9, parameters: -bh -F 4 -q 30) was used to select reads with a mapping quality score ≥ 30 and convert SAM files to sorted BAM files. After marking duplicate reads by Picard in the BAM file, GATK (v4.1.3.0) was employed to correct systematic bias by a two-stage process (BaseRecalibrator and ApplyBQSR, parameters: default).

Single nucleotide variants were individually computed by the BEIDOU toolkit with three algorithms GATK, Lofreq (v2.1.3.1, parameters: default) and Strelka2 (v2.9.10, parameters: default) with workflows for the germline variant calling. Genome-wide indels were also detected by the BEIDOU toolkit with GATK, Strelka2 (parameters: default) and Scalpel⁸³ (v0.5.4, parameters: --single --window-dow 600).

For GATK, genome-wide de novo variants were determined by three GATK commands, HaplotypeCaller (parameters: default), VariantRecalibrator (parameters: "--resource: hapmap, known=false, training=true, truth=true, prior=15.0 hapmap_3.3.hg38.vcf.gz --resource: omni, known=false, training=true, truth=false, prior=12.0 1000G_omni2.5.hg38.vcf --resource: 1000G, known=false, training=true, truth=false, prior=10.0 1000G_phase1.snps.high_confidence.hg38.vcf --resource: dbsnp, known=true, training=false, truth=false, prior=2.0 dbsnp_146.hg38.vcf -an QD -an MQ -an MQRankSum -an ReadPosRankSum -an FS -an SOR -an DP --max-gaussians 4" for SNVs; "--resource:mills,known=true,training=true,truth=true,prior=12.0 Mills_and_1000G_gold_standard.indels.hg38.vcf.gz -an QD -an MQRankSum -an ReadPosRankSum -an FS -an SOR -an DP --max-gaussians 4 -mode INDEL" for indels) and ApplyVQSR (parameters: "--mode SNP -ts-filter-level 95" for SNVs; "--mode INDEL -ts-filter-level 95" for indels). VCF files

used for VariantRecalibrator were downloaded from <https://ftp.ncbi.nih.gov/snp/> and <https://console.cloud.google.com/storage/browser/genomics-public-data/resources/broad/hg38/v0>.

Of note, overlaps of SNVs/indels called by these three algorithms were considered reliable variants by the BEIDOU toolkit. To further obtain *de novo* SNVs/indels, we filtered out the background variants, including: (1) SNVs/indels in non-transfected cells of this study and dbSNP (v151, <http://www.ncbi.nlm.nih.gov/SNP/>) database; (2) SNVs/indels with allele frequencies less than 10% or depth less than 10 reads; (3) SNVs/indels overlapped with the UCSC repeat regions. Analyses were only focused on SNVs/indels from canonical (chr 1–22, X, Y and M) chromosomes.

Whole-transcriptome sequencing data analysis

RNA-seq reads were trimmed using Trimmomatic (v0.38, parameters: ILLUMINACLIP: TruSeq3-PE-2.fa: 2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) to remove low-quality read sequences, and read qualities were evaluated using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). RNA editing sites were called using the published RADAR pipeline.³⁴

Statistical tests

Statistical analyses were done with GraphPad Prism (GraphPad Software, Inc.). In Figures 2 and 6, two-tailed Student's t-tests was used to analyze statistical significance. In Figures 3 and 4, p values were calculated by paired t test. p value < 0.05 was considered statistically significant.