

High level of fetal-globin reactivation by designed transcriptional activator-like effector

Jun Zhan,^{1,*} Maria Johnson Irudayam,^{1,*} Yukio Nakamura,² Ryo Kurita,³ and Arthur W. Nienhuis¹

¹Division of Experimental Hematology, Department of Hematology, St. Jude Children's Research Hospital, Memphis, TN; ²Cell Engineering Division, RIKEN BioResource Research Center, National Research and Development Agency, University of Tsukuba, Tsukuba, Japan; and ³Department of Research and Development, Central Blood Institute, Japanese Red Cross Society, Tokyo, Japan

Key Points

- TALEs targeted to the γ -globin gene promoters reactivated their mRNA expression more than 70-fold with a collateral reduction in β -globin mRNA.
- At day 19 of CD34 erythroid differentiation, TALEs increased γ -globin more than 40-fold in mRNA level and up to 70% of the total globin protein.

The fetal-to-adult hemoglobin switch has been a focus of a long-standing effort to potentially treat sickle cell disease and β thalassemia by induction of fetal hemoglobin. In a continuation of this effort, we designed specific transcriptional activator-like effectors (TALEs) to target both the $^G\gamma$ and $^A\gamma$ -globin promoters. We fused the TALEs to a LIM domain binding protein (Ldb1) dimerization domain, followed by a T2A green fluorescent protein (GFP) cassette, which were assembled into a lentiviral vector. To prevent deletions caused by the repeats of TALEs during the lentivirus packing process, we changed the TALE encoding DNA by codon optimization. Intriguingly, 5 of 14 TALEs showed forced reactivation of fetal-globin expression in human umbilical cord blood-derived erythroid progenitor (HUDEP-2) cells, with a significant increase in the γ -globin mRNA level by more than 70-fold. We also observed a more than 50% reduction of β -globin mRNA. High-performance liquid chromatography analysis revealed more than 30% fetal globin in TALE-induced cells compared with the control of 2%. Among several promoters studied, the β -globin gene promoter with the locus control region (LCR) enhancer showed the highest TALE expression during CD34 erythroid differentiation. At day 19 of differentiation, 2 TALEs increased fetal-globin expression more than 40-fold in the mRNA level and up to 70% of the total globin protein. These TALEs have potential for clinical translation.

Introduction

Sickle cell disease (SCD) and thalassemia are very severe inherited disorders of the red cell and are among the world's most common genetic diseases, with a high socioeconomic burden. For example, medical care of patients with SCD costs the United States more than a billion dollars annually. β thalassemia is caused by a severe deficiency in the synthesis of normal β -globin in red cells, whereas SCD reflects a mutant hemoglobin that produces protein aggregates.

A recent revolution in gene editing technology has provided new opportunities for highly efficient fetal-globin gene activation with a potential for therapeutic benefit. Earlier we demonstrated that an artificial zinc finger transcription factor, GG1-VP64, interacted with the proximal γ -globin gene promoters and induced significant amounts of fetal hemoglobin (HbF) in maturing adult erythroblasts.¹ Studies by others showed that reactivation of fetal-globin gene expression occurred through LCR-promoter looping formation by tethering the self-association domain of Ldb1 to the γ -globin promoter via artificial zinc fingers. Targeting the fetal γ -globin promoters in human adult erythroblasts enhanced the LCR-promoter looping mechanism, leading to increased γ -globin synthesis, reaching 85% of the total with a reciprocal reduction in adult β -globin gene expression. Thus, forced chromatin looping can override the γ -to- β

Submitted 21 May 2019; accepted 27 January 2020; published online 21 February 2020. DOI 10.1182/bloodadvances.2019000482.

*J.Z. and M.J.I. contributed equally to this study.

Data available by e-mail to either corresponding author, Arthur W. Nienhuis, at arthur.nienhuis@stjude.org, or Jun Zhan, at jun.zhan@stjude.org.

The full-text version of this article contains a data supplement.

© 2020 by The American Society of Hematology

globin gene switch.^{2,3} Synthetic zinc finger DNA-binding domains (ZF-DBDs) have also been used in genome editing and transcription modulation experiments; in particular, ZF-DBDs targeted to a known site associated with hereditary persistence of fetal hemoglobin (HPFH), a benign genetic condition in which mutations in a sequence located at 567 bp upstream of the γ -globin promoter attenuate γ -to- β switching. The ZF-DBD that interacted with the targeted upstream sequence of γ -globin promoter with high affinity increased the HbF level.⁴

In general, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 are considered to be powerful genome editing tools with achievable therapeutic potential for SCD and other genetic disorders,⁵ with some advantages and pitfalls of each. More recently, homologous recombination at the β -globin gene in hematopoietic stem cells was efficiently achieved using Cas9 ribonucleoproteins together with adeno-associated viral vector delivery of a homologous donor.⁶ This method has provided an efficient correction of the Glu6Val mutation responsible for SCD in patient-derived stem and progenitor cells after differentiation into erythrocytes that express adult β -globin messenger RNA. To achieve fetal-globin reactivation, researchers have also tried to recapitulate the HPFH activation using CRISPR/Cas9-mediated genome editing to mutate a 13-nt sequence that is present in the promoters of the *HBG1* and *HBG2* genes.⁷

In contrast, ZFNs are designed to target the sickle cell mutation to provide efficient targeted cleavage at the β -globin locus with minimal off-target modifications. A high level of fetal hemoglobin was achieved, resulting in decreased potential for sickling.⁸ Reactivation of γ -globin expression was also achieved with the same ZFN technology through a mutation in the binding domain of the *SOX6* gene.⁹ Although there are benefits in using these nucleases, the main downside is the risk for off-target cleavages. Several strategies have been developed, such as using high-fidelity Cas9¹⁰ or using a truncated gRNA¹¹ to reduce off-target effects. Methods such as droplet digital polymerase chain reaction (PCR) assay and high-throughput sequencing have been devised to monitor the frequency of rearrangements between the β - and δ -globin paralogs by targeting nucleases. This has resulted in the identification of intergenic β - δ -globin deletion and inversion of an intergenic fragment as high- and low-frequency rearrangements, respectively. Altogether, these results have proven the need to develop site-specific endonucleases with high specificity to avoid unwanted gene alterations.¹²

CRISPR/Cas9 has a significant rate of off-target mutations, and therefore, it might cause some adverse effects.¹³ ZFNs are very difficult to engineer and, in addition, produce greater cell toxicity by inhibiting cell growth. In comparison with ZFNs, TALENs have turned out to be much easier to design with less or comparable cytotoxicity.^{14,15} However, additional genotoxicity resulted from insertional mutagenesis when lentiviral vector delivery of CRISPR, TALE, or ZFN was tested. The use of a γ -retroviral vector can lead to oncogenesis by activating *LMO2* gene expression, as seen in the Europe X-linked severe combined immune deficiency trial.^{16,17} Using a lentiviral vector instead of a γ -retroviral vector would seem much safer, but still retain a concern for oncogenesis.¹⁸ In recent years, several methods have been developed to evaluate lentiviral vector safety,^{19,20} including reduction of its genotoxicity.²¹ All these

efforts have led to lentiviral vector being used widely for most gene therapy protocols for hematological diseases.²² In the current study, we have used a lentiviral vector backbone that was developed in our laboratory after rigorous safety testing and validation analyses.²³

Because TALEs have advantages, as noted here, we used this strategy to increase fetal-globin expression with the potential for treatment of β -thalassemia and SCD with the help of lentiviral delivery.¹⁵ We here show that targeting specific sites in the γ -globin gene promoters using TALEs fused with Ldb1 self-association domain reactivated significant production of γ -globin with concomitant reduction in β -globin gene expression. This could inhibit sickling by formation of mixed tetramers, including both the γ and β ^S chains. Hence, the induction of fetal-globin by TALEs holds significant promise for alleviation of the phenotypic features of β -thalassemia and SCD.

Materials and methods

Construction of the TALE destination vector backbone plasmid

The plasmid, pCL20cwEF1GFPD46, was cut by BstBI and BsrGI to remove the EF1 GFP cassette, followed by the insertion of a cassette containing the EF1-TALE N terminus, linker, and TALE C terminus that were fused to the Ldb1 dimerization domain and T2A GFP. The N terminus of the TALE had 139 aa, followed by a linker with a MluI/BsmBI site, whereas C-terminus TALE had 117 aa including NLS, and half of the TALE repeat was inserted before the C terminus. This plasmid was called pCL20cwEF1 link Ldb1GFP. For the backbone vector with Ldb2 dimerization domain, the above vectors were cut with PspOMI and *Dra*III, and the synthesized DNA fragment coding for the Ldb2 dimerization domain was inserted, yielding pCL20cwEF1-link Ldb2GFP. The TALE central domain (17mer) was synthesized and inserted into the linker at a MluI/BsmBI site yielding pCL20cwEF1TALE-Ldb1T2A GFP.

TALE design

Using web-based tool (<https://tale-nt.cac.cornell.edu/node/add/talen>), we designed TALEs targeting both γ -globin gene promoters. Off-target effects were assessed, using BLAST search and the TALE design tool as an off-target finder. A total of 14 TALEs were designed. First, the targeted DNA sequence was converted to repeat variable diresidue (RVD) according to the binding code. The RVD was then converted to 6 nucleotides, using genetic code; subsequently, every 6 nucleotides were inserted into TALE DNA template. The DNA template was a 17mer TALEs optimized for human codon usage with elimination of the repeat sequence.²⁴ The 17mer TALE was synthesized and inserted into the vector pCL20cwEF1 linkLdb1GFP and called pCL20cwEF1aTALE-Ldb1T2AGFP. The TALE destination vectors were constructed by modifying the TALE cloning backbones. We recoded the 0.5 RVD regions in the vectors and subsequently incorporated a MluI/BbsI cutting site for the newly synthesized 17mer TALE central domain. Further, this was inserted at the designated TALE cloning sites.

Human umbilical cord blood derived erythroid progenitor cell culture

Human umbilical cord blood derived erythroid progenitor (HUDEP) clone 2 (HUDEP-2) cells were cultured as previously described.²⁵ Briefly, cells were expanded in StemSpan SFEM

(Stem Cell Technologies) supplemented with 1 μM dexamethasone, 1 $\mu\text{g}/\text{mL}$ doxycycline, 50 ng/mL human stem cell factor (SCF), 3 U/mL erythropoietin, and 1% penicillin/streptomycin.

Vector copy number determination

For human vector copy number analyses, a 20X HIV GAG probe was prepared (forward primer: 18 μM 5'GGAGCTAGAACGATT CGCAGTTA3' and reverse primer: 18 μM 5'GGTTGTAGCTGT CCCAGTATTTGTC3'). Primers and probe (5 μM of 6FAM-CAGCCTTCTGATGTTTCTAACAGGCCAGG-NFQ-MGB) were synthesized and used with TaqMan fast universal PCR master mix (Thermo Fisher, Foster City, CA) in quantitative PCR (qPCR) reactions. Human RNase P probe (Thermo Fisher) was included in the qPCR reactions to quantitate the relative vector copy numbers between cell samples, using the $\Delta\Delta\text{Ct}$ method against a 1-copy lentiviral Jurkat cell control.

Quantitative reverse transcription PCR

RNA was extracted from between 0.5 and 1 million cells with RNeasy kits (Qiagen, Germantown, MD). Taqman qPCR probes for *HBG* (Hs03835607_m1), *HBB* (Hs00758889_s1), *HBE* (Hs00362216_m1), and *HBA* (Hs01109871_m1) were commercially purchased (Thermo Fisher). One-step quantitative reverse transcription PCR (qRT-PCR) was then performed and γ -globin, β -globin, and ϵ -globin mRNAs were quantified by Taqman qPCR (Thermo Fisher). The levels of γ -globin, β -globin, and ϵ -globin mRNAs were normalized against the levels of α -globin mRNAs.

Flow cytometry

Flow cytometry (BD Biosciences, FACS Diva 8.0.1 software) was performed for assessing the expression of designed TALEs using GFP as probe in the St. Jude Children's Research Hospital flow cytometry core facility.

HbF protein analysis

High-performance liquid chromatography (HPLC) quantification of HbF was performed using a cation-exchange column (Primus Diagnostics) according to the manufacturer's protocol.

Erythroid differentiation of human peripheral blood CD34⁺ cells

Circulating G-CSF-mobilized human CD34⁺ cells were obtained from de-identified healthy donors (Key Biologics). CD34⁺ cells were cultured in a 3-phase erythroid differentiation protocol consisting of Iscove modified Dulbecco medium (Thermo Fisher) supplemented with 2% human AB plasma, 3% human AB serum, 1% penicillin/streptomycin, 3 U/mL heparin, 10 $\mu\text{g}/\text{mL}$ insulin, and 3 U/mL erythropoietin (Amgen). Phase 1 (days 1-7) included 200 $\mu\text{g}/\text{mL}$ Holo-Transferrin (Sigma-Aldrich), 10 ng/mL SCF (PeproTech, Inc.), and 1 ng/mL interleukin 3 (PeproTech, Inc.). Phase 2 (days 8-12) included the same cytokines as phase 1, except without interleukin 3. During phase 3 (days 13 and beyond), Holo-Transferrin was increased to 1 mg/mL and SCF was removed.

Human CD34⁺ cell transduction and methylcellulose colony assays

Human CD34⁺ cells were thawed and prestimulated in stem-spam supplemented with 100 ng/mL Flt-3, 100 ng/mL SCF, and 100 ng/mL thrombopoietin overnight. Various lenti-TALE vectors were transduced

into human CD34⁺ cells in the presence of protamine sulfate (0.4 $\mu\text{g}/\text{mL}$) and seeded onto retronectin-coated plates. After overnight transduction, CD34⁺ cells were seeded at 300 cells/mL into cytokine-free human methylcellulose (Stem Cell Technologies; 300 cells/3 cm dish) supplemented with 2 U/mL erythropoietin, 10 ng/mL SCF, 1 ng/mL interleukin 3, and 1% penicillin/streptomycin. Individual BFU-E colonies were picked up after 2 weeks and processed for genomic DNA and colony PCR assays. The transduction efficiency was determined on the basis of the positivity of PCR.

Statistics

Statistical differences between experimental groups were determined by 1-way analysis of variance (ANOVA), followed by Dunnett's multiple comparisons test using statistical software within GraphPad Prism, as indicated (GraphPad Software, La Jolla, CA). All experiments were performed in triplicate unless otherwise indicated. Data are shown as mean \pm standard error of the mean (SEM; or standard deviation). $P < .05$ was considered statistically significant.

Results

TALE design

We used TALE to reactivate fetal globin to evaluate the potential for clinical therapy for SCD and β thalassemia. TALEs (supplemental Figure 1A) can be engineered to bind to DNA sites, including specific sites on globin promoters. TALEs were fused to the Ldb1 dimerization domain that binds to the LCR, which in turn brings the distal LCR enhancer to the γ -globin gene promoters, thereby activating gene expression (supplemental Figure 1B). As shown in supplemental Figure 1C, we designed TALE using the pCL20 vector backbone with an EF1 α short promoter (called EF1) to drive TALE Ldb1 dimerization domain fusion protein expression, including the T2A GFP cassette. Fourteen TALEs were designed to target both the $\text{G}\gamma$ and $\text{A}\gamma$ -globin promoters (Figure 1A). Several (TALE2/5/6) were designed to bind on or around fetal-globin negative regulatory sites (supplemental Figure 2A) overlapping the fetal-globin negative regulator BCL11A binding site²⁶ or some of the HPFH sites. Our hypothesis is that targeting these regions will replace negative regulators with positive regulators, leading to fetal-globin reactivation. Details of the TALE design (supplemental Figure 2B) are provided in the Materials and Methods. The integrity of the TALE in the genome was confirmed by PCR and Southern blot analyses (supplemental Figure 3A-B). The expression of TALE was validated at the mRNA level by qRT-PCR, using a GFP probe (supplemental Figure 4A). Flow cytometry has confirmed GFP expression at the protein level (supplemental Figure 4B). These data strongly support that the designed TALE is capable of expressing at both levels with almost comparable mRNA-protein correlations.

TALE Function in HUDEP-2 cells

We employed HUDEP-2 cells to examine the function of TALEs with respect to fetal-globin gene reactivation. qRT-PCR was performed to determine the relative mRNA level for fetal-globin after normalization to the α -globin mRNA content. Among 14 TALEs (Figure 1A), 5 TALEs (TALE2, TALE6 TALE7 TALE7b, and TALE11) showed significant reactivation of fetal-globin expression with a substantial increase in mRNA level. TALE6 showed more than

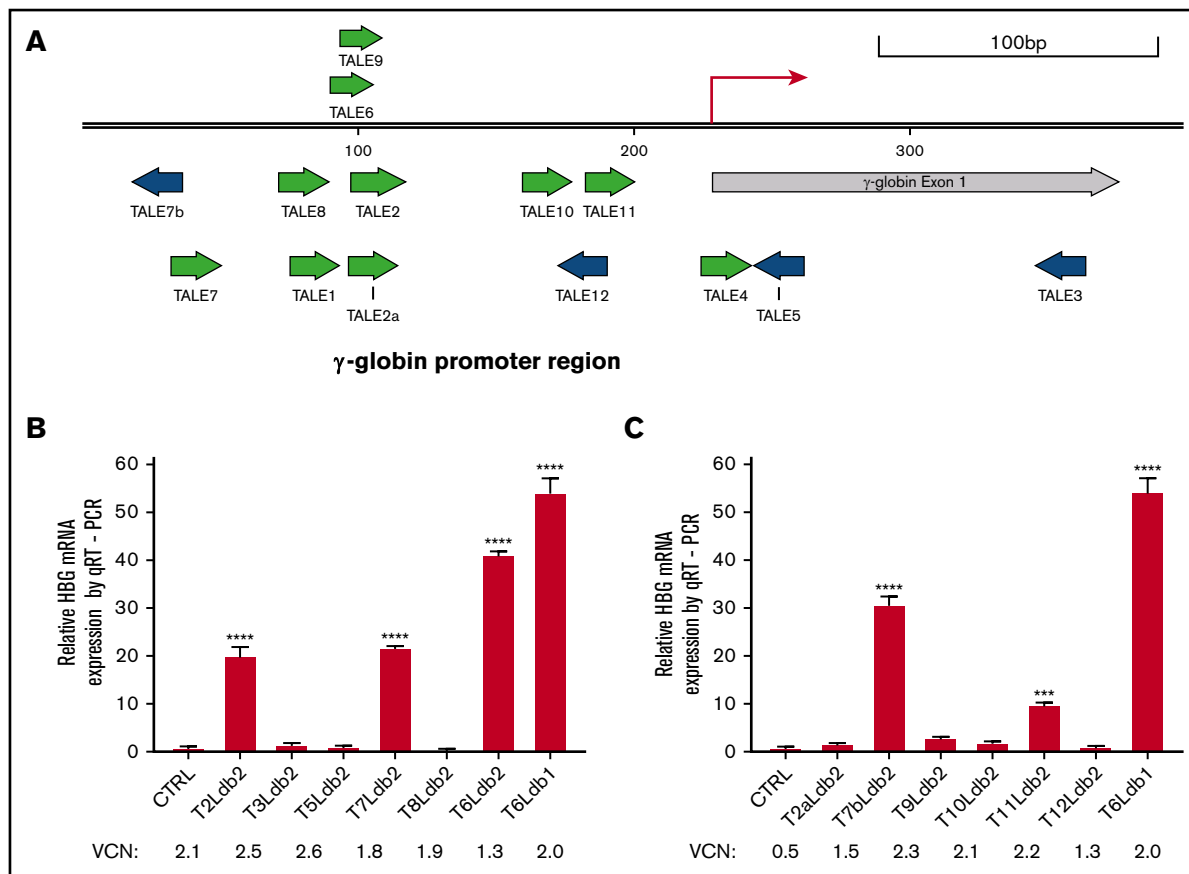


Figure 1. TALEs activate fetal-globin expression in HUDEP-2 cells. (A) Diagram of 14 designed TALEs' relative binding locations on γ -globin gene promoter. (B-C) HUDEP-2 cells were infected with TALE expressing lentivirus. RNA was prepared after 7 days. qRT-PCR was performed using *HBG* as probe and normalized by *HBA* amount. The vector copy number was determined for these analyses as indicated. The statistical significance was calculated by 1-way ANOVA followed by Dunnett's multiple comparisons test. **** $P < .0001$. Values mean \pm SEM.

a 40-fold (Figures 1B-C) increase, indicating the high functionality of TALE design. TALE also slightly activated the expression of ϵ -globin (supplemental Figure 5). This is mainly because of its spatial proximity to the γ -globin promoter, indicating that the newly imposed spatial constraint of the LCR can also increase the likelihood of interaction with a nearby gene such as the ϵ -globin. Among the 5 TALEs, 2 TALEs (TALE2, TALE6) were designed to target the direct repeat and BCL11A sites^{27,28} (supplemental Figure 2A). Two other TALEs (TALE7, TALE7b) were targeted toward the upstream of the direct repeat site and close to the LRF1 binding site.^{26,29} Thus, this targeting encompasses the vital regions of negative regulatory sites for the enhancement of γ -globin expression.

Next, we compared the level of TALE activity with various activation domains. In addition to Ldb1, Ldb2 can also dimerize via the self-association domain with high affinity,³⁰ and hence more strong gene activation. We also considered VP64, a different gene activation domain that could possibly increase the transcription activity. We also compared the TALE-Ldb1 and TALE-Ldb2 fusion proteins in the presence of VP64 (Figure 2A). Intriguingly, all these designs strongly activated fetal-globin gene expression with a concomitant repression of β -globin expression (Figures 2B-C). The reactivation of fetal-globin genes resulted in a 50% reduction in

β -globin mRNA. The level of fetal-globin protein, determined by HPLC, was more than 30% of total globin compared with 2% in control cells (Figure 2D), in both the presence and absence of VP64. As predicted, the addition of VP64 showed a bolster effect and increased fetal-globin RNA and repressed β -globin mRNA to a lesser degree (Figure 2C). TALE-Ldb1 seemed to slightly enhance the percentage at the protein level (Figure 2D). This substantiates that TALE design with the Ldb1 self-association domain is more efficient in the reactivation of fetal-globin, regardless of the presence or absence of VP64.

TALEs Function in Human CD34-Differentiated Erythroid Cells

To evaluate the potential for clinical translation, we transduced human CD34⁺ cells with the TALE-expressing vectors (TALE7Ldb2, TALE6vpLdb2, TALE6Ldb2, TALE6vpLdb1, T6Ldb1). The transduced human CD34⁺ cells were grown in MethoCult-based medium for erythroid differentiation. The transduction efficiency of these vectors is a critical index for potential clinical use. PCR assay performed using genomic DNA isolated from MethoCult colonies indicated that more than 30% of CD34 cells could be transduced with the above vectors (supplemental Figure 6).

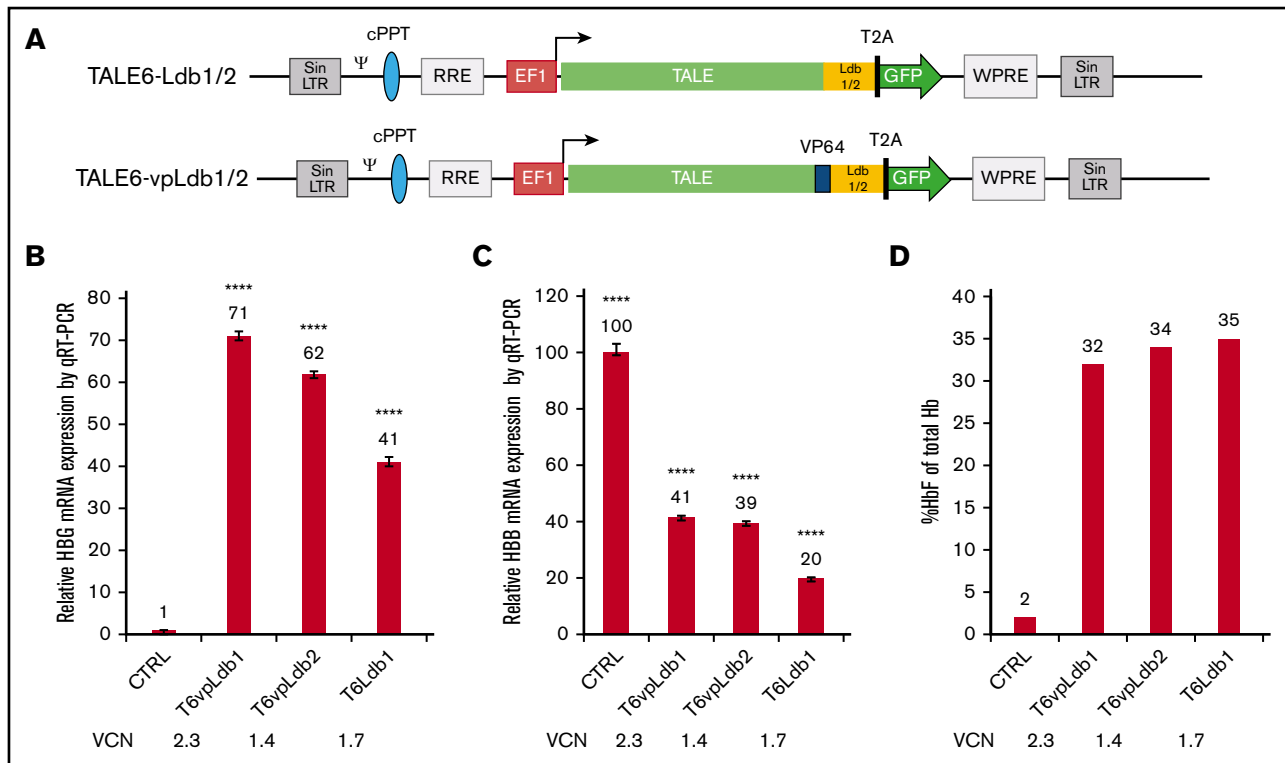


Figure 2. Comparison of TALEs with different activation domain in HUDEP-2 cells. (A) TALE vector design. The EF1 short promoter was used to drive the TALE expression. The TALEs were fused with a Ldb1 or Ldb2 dimerization domain, followed by a T2A GFP cassette. The activation domain of VP64 was inserted before the Ldb1/2 dimerization domain to test whether the gene activation could be bolstered. To prevent deletions caused by repeats of TALEs during the lentivirus packing process, we changed the TALE's DNA codons by using codon-optimized DNA for their expression. (B) HUDEP-2 cells were infected with TALE-expressing lentivirus. RNA was prepared after 7 days. qRT-PCR was performed using *HGB* as probe, normalized by *HBA* amount. (C) Relative β -globin mRNA level by qRT-PCR, using *HBB* as probe, normalized by *HBA* amount. The vector copy number (VCN) was determined and depicted in the graph. The statistical significance was calculated by 1-way ANOVA followed by Dunnett's multiple comparisons test. **** $P < .0001$. Values mean \pm SEM. (D) Bar graphs show proportions of the fetal-globin protein relative to total globin amount level determined by HPLC. Analyses are also indicated with their relevant VCN.

Subsequently, we also checked the functionality of TALE-Ldb1 in erythroid cells. We used an in vitro CD34 erythroid differentiation system that recapitulates the developmental pattern of Hb expression associated with the developmental stage of the originating primary cell source.¹ Initially we identified that the erythroid differentiation gradually shut down the EF1 promoter and failed to drive TALE expression in erythroid cells. We introduced an erythroid-specific promoter to drive TALE expression. Among several promoters tested (supplemental Figure 7), the β -globin promoter with shortened LCR, called the dLCR, a 1.7-kb fragment containing the cores of HSs (HS2, HS3, and HS4), was found to be the best driver for TALE expression in erythroid cells.

We observed that the β -globin promoter plus dLCR enhancer driving the expression of TALE7b and TALE11 with Ldb1 (Figure 3A) resulted in more than a 40-fold increase in the γ -globin mRNA level with the fetal-globin level above 70% of total globin (Figure 3B-C). However, T6 constructs failed to activate fetal-globin gene in human CD34 differentiated erythroid cells (data not shown). These results substantiate that the use of dLCR enhancer and β -globin promoter significantly enhanced TALE expression at both the mRNA and protein levels. The LRF/ZBTB7A transcription factor induces the fetal-globin genes by silencing the NuRD repressor complex, which is independent of the fetal globin

repressor, BCL11A,³¹ whereas transcription factors LRF and BCL11A are the independent repressors of fetal hemoglobin expression. Knockdown of *LRF/ZBTB7A* gene by shRNA or by CRISPR/gRNA led to a very high level of fetal-globin expression.³¹ On the basis of this result, we chose CRISPR-LRF as a positive control in our system to compare the efficacy of different TALEs with different promoter combinations. As predicted, we noted that CRISPR-LRF activated the fetal-globin expression with a 40-fold increase in the γ -globin mRNA level with activation of fetal-globin to more than 70% of total globin, suggesting that the effect of the fetal-globin repressor was deactivated through CRISPR cleavage of *LRF1* (Figure 3A-C). Furthermore, we tested the efficacy of a Lenti- γ globin vector, 400v5m3,^{32,33} that was previously developed and tested in our laboratory for gene therapy of SCD.³² The efficacy of 400V5m3, a traditional γ -globin gene therapy vector, showed low induction of fetal-globin at both mRNA and protein levels as compared with that of the dLCR-T7b/11 or CRISPR-LRF1, suggesting that Lenti-TALE-Ldb1 is a promising candidate to mitigate SCD and thalassemia.

Discussion

TALEs are a proven powerful easy design tool for targeting the genome in cell lines and organisms. Numerous strategies have been

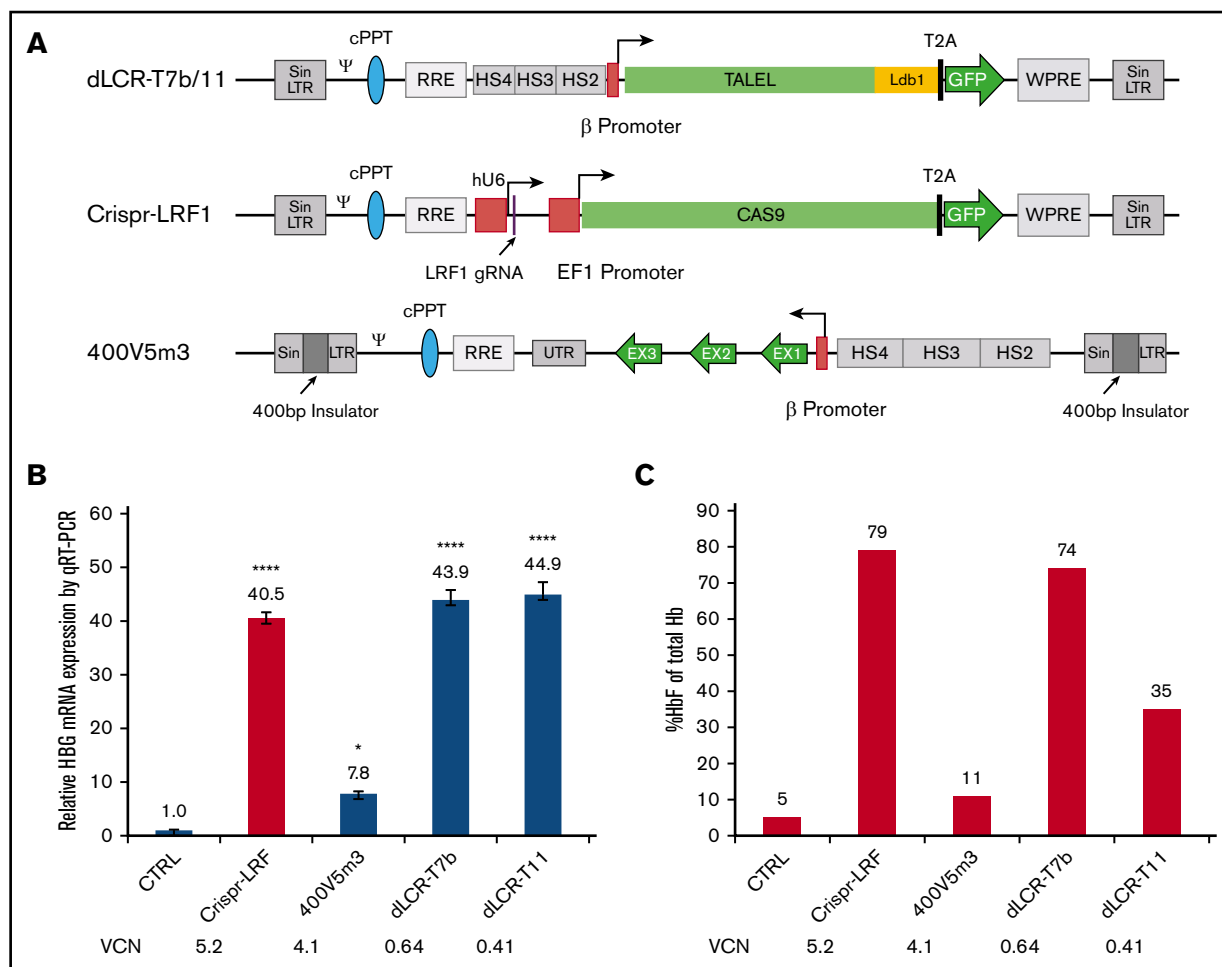


Figure 3. TALEs activate fetal-globin expression in human CD34 differentiated erythroid cells. (A) New vector design for fetal-globin gene action. The 1.7-kb shortened LCR enhancer (called dLCR) and β -globin promoter were used to drive TALE7b and TALE11 expression. 400v5m3 is the lenti- γ globin vector previously developed in this laboratory for the gene therapy of SCD. (B) Human CD34 cells subjected to 3-phase erythroid differentiation at day 19, and harvested cells for RNA. Relative fetal-globin mRNA level was determined by qPCR, using *HBG* as probe, normalized by *HBA* amount. The vector copy number (VCN) was determined and depicted in the graph. The statistical significance was calculated by 1-way ANOVA followed by Dunnett's multiple comparisons test. * $P = .01$; **** $P < .0001$. Values are mean \pm SEM. (C) Bar graphs show proportions of the fetal-globin protein relative to total globin amount level determined by HPLC.

developed in the past to assemble the repetitive TALE RVD sequence.³⁴⁻³⁶ Because of the presence of repeats in the central domain, TALE expression is unstable with the potential for deletion of some of the repeats during the process of lentiviral packaging and expression. Therefore, generally it is very hard to express the full-length TALE using lentiviral-based vectors. This eventually led us to adapt and modify Church's re-TALE strategy,²⁴ which uses codon optimization to eliminate the repeats in the central domain. With this design, PCR and Southern blot have confirmed the integrity of lenti-TALE in the genome.

Among 14 TALEs tested, 5 were able to activate fetal-globin gene expression: 2 (TALE2, TALE6) were designed to bind close to the BCL11A binding sites and 2 (TALE7, TALE7b) were designed to bind close to the LRF/ZBTB7A binding sites to prevent the inhibition of γ -globin expression consistent with the 2 HPFH mutation clusters at -115 and -200 base pairs upstream of the γ -globin gene promoters. Two major γ -globin repressors, BCL11A^{27,28} and ZBTB7A,³¹ directly bind to -115 and -200

sites, respectively, and repress γ -globin gene expression. HPFH mutations at these sites disrupt the binding of these repressors.^{26,29} BCL11A binding was disrupted by $-117G>A$, $-114C>A$, $-114C>T$, $-114C>G$, and a 13-bp deletion ($\Delta 13$ bp). Similarly, LRF1/ZBTB7A binding at -200 bp upstream of the γ -globin gene promoter was disrupted by $-195C>G$, $-196C>T$, $-197C>T$, $-201C>T$, $-202C>T$, and $-202C>G$.³⁷⁻³⁹ Thus, both BCL11A and ZBTB7A account for most of the repression of the γ -globin gene in adult cells.³¹ As the above-mentioned 4 TALEs were designed to span the repressor sites, we deduced that mechanically repressor displacement might have played a pivotal role; however, the potential involvement of chromatin looping of the enhancer to increase the fetal globin levels needs to be elucidated.

In addition to these 4 TALEs, the fifth TALE (TALE11), designed to bind close to the TATA box, showed significant fetal globin expression (Figures 1A and 3B-C). No repressor elements have been identified around this site. This may be another relevant

site, along with the other known negative regulators of fetal globin expression, which invites further study to determine whether it is a repressor site for the interference of fetal globin expression.

Two TALEs designed to target negative regulatory elements did not show fetal-globin gene activation. The TALE2a targeting sequence is almost the same, except 2 nucleotides shorter than that of TALE2. Because there is no significant difference in the expression of these TALEs, it is presumed that TALE2a does not appear to have enough binding affinity to its target. TALE5 is designed to target the LYAR⁴⁰ site that is about 20 bp downstream of fetal-globin gene transcription start site, as mutation of LYAR often results in HPFH.⁴¹ However, TALE5 may not have enough binding affinity even if it binds the LYAR site to displace the transcriptional machinery, and therefore downstream targeting will not be beneficial. Another 2 nonfunctional TALEs (TALE1/8) were designed to target KLF11 sites that are required for the activation of fetal-globin gene expression. As TALE can be designed to target almost any sequence in the genome, our goals of globin gene editing will span deeper into these 2 important repressor sites in the future.

We used an in vitro CD34 erythroid differentiation system that could recapitulate the developmental pattern of Hb expression associated with the developmental stage of the originating primary cell source.¹ Initially, we used the EF1 α short promoter to drive TALE-Ldb1 expression. This short promoter was able to drive the gene activation in HUDEP-2 cells. However, the activity of TALE was gradually lost during erythroid differentiation as a result of shutting down of EF1 α promoter. To find the best promoter that can show its functional ability to withstand the final stage of erythroid differentiation and maturation, we tested various promoters. The ankyrin promoter with a GC mutation showed a low activity in HUDEP-2 cells. Only the LCR enhancer with β -globin promoter drove TALE expression effectively in differentiating CD34 erythroid cells.

Although the current work needs to address the underlying mechanism involved in the TALEs reactivation of fetal globin expression, as far as the fetal-globin gene activation (in the human CD34 erythroid differentiated cells) is concerned, TALE-Ldb1 is highly comparable to GG1-Ldb1.^{3,42} GG1-Ldb1 has been more extensively evaluated, even analyzing its efficacy using CD34 cells of patients with sickle cell.² However, TALE (as presented here) or the GG1-Ldb1 need further testing with the help of animal models

to determine whether they merit a clinical trial. The current design of TALEs, similar to the presence of T2A GFP, leads to incomplete cleavage of T2A peptide and interferes with TALEs expression. Efforts are already underway to eliminate this drawback, and our preliminary data (not shown) indicate that gene activation increases sevenfold on deletion of the T2A GFP cassette. A poly A signal greatly increases the transgene expression, but unfortunately it reduces titers.⁴³ This could be corrected by putting the poly A signal in the inverted vector design to minimize titer reduction. Modifying the WPRE may enhance the transgene expression. Instead of using VSVG pseudotype to pack lentiviral vector, pseudotyping with the baboon endogenous virus envelope glycoprotein BaEVRles efficiently and specifically targets human CD34⁺ cells.⁴⁴ Thus, the vector modification may provide a better clinical approach to develop Lenti-TALE-Ldb1 as a promising candidate for effective therapeutic intervention in sickle cell and thalassemia diseases.

Acknowledgments

The authors acknowledge the excellent administrative assistance provided by Pat Streich.

The funds to support this work were provided by American Lebanese Syrian Associated Charities.

Authorship

Contribution: J.Z designed and performed the research; M.J.I. assisted in the research and performed statistical analyses of the data, formatting and graphing figures with Prism; J.Z and M.J.I. wrote the original draft of the manuscript; Y. N. and R.K. provided the HUDEP-2 cell line; and A.W.N. developed the research plan, supervised the research, and participated in the preparation of the manuscript and funding acquisition.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profile: J.Z., 0000-0002-1974-0098.

Correspondence: Arthur W. Nienhuis; e-mail: arthur.nienhuis@stjude.org; Jun Zhan, St. Jude Children's Research Hospital, 262 Danny Thomas Pl, Memphis, TN 38105; e-mail: jun.zhan@stjude.org.

References

1. Wilber A, Tschulena U, Hargrove PW, et al. A zinc-finger transcriptional activator designed to interact with the gamma-globin gene promoters enhances fetal hemoglobin production in primary human adult erythroblasts. *Blood*. 2010;115(15):3033-3041.
2. Breda L, Motta I, Lourenco S, et al. Forced chromatin looping raises fetal hemoglobin in adult sickle cells to higher levels than pharmacologic inducers. *Blood*. 2016;128(8):1139-1143.
3. Deng W, Rupon JW, Krivega I, et al. Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell*. 2014;158(4):849-860.
4. Hossain MA, Shen Y, Knudson I, et al. Activation of fetal γ -globin gene expression via direct protein delivery of synthetic zinc-finger DNA-binding domains. *Mol Ther Nucleic Acids*. 2016;5(10):e378.
5. Hoban MD, Lumaquin D, Kuo CY, et al. CRISPR/Cas9-mediated correction of the sickle mutation in human CD34⁺ cells. *Mol Ther*. 2016;24(9):1561-1569.
6. Dever DP, Bak RO, Reinisch A, et al. CRISPR/Cas9 β -globin gene targeting in human haematopoietic stem cells. *Nature*. 2016;539(7629):384-389.
7. Traxler EA, Yao Y, Wang YD, et al. A genome-editing strategy to treat β -hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat Med*. 2016;22(9):987-990.

8. Hoban MD, Cost GJ, Mendel MC, et al. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood*. 2015; 125(17):2597-2604.
9. Modares Sadeghi M, Shariati L, Hejazi Z, Shahbazi M, Tabatabaiefar MA, Khanahmad H. Inducing indel mutation in the SOX6 gene by zinc finger nuclease for gamma reactivation: An approach towards gene therapy of beta thalassemia. *J Cell Biochem*. 2018;119(3):2512-2519.
10. Vakulskas CA, Dever DP, Rettig GR, et al. A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat Med*. 2018;24(8):1216-1224.
11. Zhang JP, Li XL, Neises A, et al. Different effects of sgRNA length on CRISPR-mediated gene knockout efficiency. *Sci Rep*. 2016;6(1):28566.
12. Long J, Hoban MD, Cooper AR, et al. Characterization of gene alterations following editing of the β -globin gene locus in hematopoietic stem/progenitor cells. *Mol Ther*. 2018;26(2):468-479.
13. Cullot G, Boutin J, Toutain J, et al. CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat Commun*. 2019;10(1):1136.
14. Gaj T, Gersbach CA, Barbas CF III. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*. 2013;31(7):397-405.
15. Gupta RM, Musunuru K. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *J Clin Invest*. 2014;124(10):4154-4161.
16. Hacein-Bey-Abina S, Von Kalle C, Schmidt M, et al. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*. 2003;302(5644):415-419.
17. Howe SJ, Mansour MR, Schwarzwaelder K, et al. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest*. 2008;118(9):3143-3150.
18. Cesana D, Ranzani M, Volpin M, et al. Uncovering and dissecting the genotoxicity of self-inactivating lentiviral vectors in vivo. *Mol Ther*. 2014;22(4):774-785.
19. Modlich U, Bohne J, Schmidt M, et al. Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood*. 2006; 108(8):2545-2553.
20. Ryu BY, Evans-Galea MV, Gray JT, Bodine DM, Persons DA, Nienhuis AW. An experimental system for the evaluation of retroviral vector design to diminish the risk for proto-oncogene activation. *Blood*. 2008;111(4):1866-1875.
21. David RM, Doherty AT. Viral vectors: the road to reducing genotoxicity. *Toxicol Sci*. 2017;155(2):315-325.
22. Milone MC, O'Doherty U. Clinical use of lentiviral vectors. *Leukemia*. 2018;32(7):1529-1541.
23. Wielgosz MM, Kim YS, Carney GG, et al. Generation of a lentiviral vector producer cell clone for human Wiskott-Aldrich syndrome gene therapy. *Mol Ther Methods Clin Dev*. 2015;2:14063.
24. Yang L, Guell M, Byrne S, et al. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res*. 2013;41(19):9049-9061.
25. Kurita R, Suda N, Sudo K, et al. Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS One*. 2013;8(3):e59890.
26. Martyn GE, Wienert B, Yang L, et al. Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. *Nat Genet*. 2018;50(4):498-503.
27. Menzel S, Garner C, Gut I, et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat Genet*. 2007;39(10):1197-1199.
28. Sankaran VG, Menne TF, Xu J, et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science*. 2008;322(5909):1839-1842.
29. Liu N, Hargreaves VV, Zhu Q, et al. Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell*. 2018;173(2):430-442.
30. Cross AJ, Jeffries CM, Trewheella J, Matthews JM. LIM domain binding proteins 1 and 2 have different oligomeric states. *J Mol Biol*. 2010;399(1):133-144.
31. Masuda T, Wang X, Maeda M, et al. Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science*. 2016; 351(6270):285-289.
32. Pestina TI, Hargrove PW, Zhao H, et al. Amelioration of murine sickle cell disease by nonablative conditioning and γ -globin gene-corrected bone marrow cells. *Mol Ther Methods Clin Dev*. 2015;2:15045.
33. Urbinati F, Hargrove PW, Geiger S, et al. Potentially therapeutic levels of anti-sickling globin gene expression following lentivirus-mediated gene transfer in sickle cell disease bone marrow CD34+ cells. *Exp Hematol*. 2015;43(5):346-351.
34. Briggs AW, Rios X, Chari R, et al. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic Acids Res*. 2012;40(15):e117.
35. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*. 2011;29(2):149-153.
36. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol*. 2012; 30(5):460-465.
37. Costa FF, Zago MA, Cheng G, Nechtman JF, Stoming TA, Huisman TH. The Brazilian type of nondeletional A gamma-fetal hemoglobin has a C \rightarrow G substitution at nucleotide -195 of the A gamma-globin gene. *Blood*. 1990;76(9):1896-1897.
38. Tasiopoulou M, Boussiou M, Sinopoulou K, Moraitis G, Loutradi-Anagnostou A, Karababa P. G gamma-196 C \rightarrow T, A gamma-201 C \rightarrow T: two novel mutations in the promoter region of the gamma-globin genes associated with nondeletional hereditary persistence of fetal hemoglobin in Greece. *Blood Cells Mol Dis*. 2008;40(3):320-322.

39. Amato A, Cappabianca MP, Perri M, et al. Interpreting elevated fetal hemoglobin in pathology and health at the basic laboratory level: new and known γ - gene mutations associated with hereditary persistence of fetal hemoglobin. *Int J Lab Hematol*. 2014;36(1):13-19.
40. Ju J, Wang Y, Liu R, et al. Human fetal globin gene expression is regulated by LYAR. *Nucleic Acids Res*. 2014;42(15):9740-9752.
41. Breveglieri G, Bianchi N, Cosenza LC, et al. An A γ -globin G->A gene polymorphism associated with $\beta^0\gamma^39$ thalassemia globin gene and high fetal hemoglobin production. *BMC Med Genet*. 2017;18(1):93.
42. Deng W, Lee J, Wang H, et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell*. 2012;149(6):1233-1244.
43. Hager S, Frame FM, Collins AT, Burns JE, Maitland NJ. An internal polyadenylation signal substantially increases expression levels of lentivirus-delivered transgenes but has the potential to reduce viral titer in a promoter-dependent manner. *Hum Gene Ther*. 2008;19(8):840-850.
44. Drakopoulou E, Georgomanoli M, Lederer CW, et al. A novel BaEVRless-pseudotyped γ -globin lentiviral vector drives high and stable fetal hemoglobin expression and improves thalassemic erythropoiesis *in vitro*. *Hum Gene Ther*. 2019;30(5):601-617.