



Transcriptome Analyses of β -Thalassemia –28(A>G) Mutation Using Isogenic Cell Models Generated by CRISPR/Cas9 and Asymmetric Single-Stranded Oligodeoxynucleotides (assODNs)

OPEN ACCESS

Edited by:

Xiao Li,
Texas Heart Institute, United States

Reviewed by:

Mo Li,
King Abdullah University of Science
and Technology, Saudi Arabia
Ling Li,
Beckman Research Institute, City
of Hope, United States
Liren Wang,
East China Normal University, China

*Correspondence:

Xiuqing Zhang
zhangxq@genomics.cn
Chao Liu
liuchao4@cngb.org
Ying Gu
guying@cngb.org

†These authors have contributed
equally to this work

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 28 June 2020

Accepted: 01 September 2020

Published: 08 October 2020

Citation:

Li J, Zhou Z, Sun H-X, Ouyang W,
Dong G, Liu T, Ge L, Zhang X, Liu C
and Gu Y (2020) Transcriptome
Analyses of β -Thalassemia –28(A>G)
Mutation Using Isogenic Cell Models
Generated by CRISPR/Cas9
and Asymmetric Single-Stranded
Oligodeoxynucleotides (assODNs).
Front. Genet. 11:577053.
doi: 10.3389/fgene.2020.577053

Jing Li^{1,2,3†}, Ziheng Zhou^{2,3†}, Hai-Xi Sun^{2,3,4}, Wenjie Ouyang^{2,3}, Guoyi Dong^{1,2,3},
Tianbin Liu^{1,2,3}, Lei Ge^{2,3}, Xiuqing Zhang^{2,3,5*}, Chao Liu^{2,3*} and Ying Gu^{2,3,6*}

¹ BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, China, ² BGI-Shenzhen, Shenzhen, China, ³ China National GeneBank, BGI-Shenzhen, Shenzhen, China, ⁴ Institute for Stem cell and Regeneration, Chinese Academy of Sciences, Beijing, China, ⁵ Guangdong Provincial Key Laboratory of Human Disease Genomics, Shenzhen Key Laboratory of Genomics, BGI-Shenzhen, Shenzhen, China, ⁶ Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen, China

β -thalassemia, caused by mutations in the human hemoglobin β (*HBB*) gene, is one of the most common genetic diseases in the world. The *HBB* –28(A>G) mutation is one of the five most common mutations in Chinese patients with β -thalassemia. However, few studies have been conducted to understand how this mutation affects the expression of pathogenesis-related genes, including globin genes, due to limited homozygote clinical materials. Therefore, we developed an efficient technique using CRISPR/Cas9 combined with asymmetric single-stranded oligodeoxynucleotides (assODNs) to generate a K562 cell model with *HBB* –28(A>G) named K562–28(A>G). Then, we systematically analyzed the differences between K562–28(A>G) and K562 at the transcriptome level by high-throughput RNA-seq before and after erythroid differentiation. We found that the *HBB* –28(A>G) mutation not only disturbed the transcription of *HBB*, but also decreased the expression of *HBB*, which may further aggravate the thalassemia phenotype and partially explain the more severe clinical outcome of β -thalassemia patients with the *HBB* –28(A>G) mutation. Moreover, we found that the K562–28(A>G) cell line is more sensitive to hypoxia and shows a defective erythrogenic program compared with K562 before differentiation. Importantly, all abovementioned abnormalities in K562–28(A>G) were reversed after correction of this mutation with CRISPR/Cas9 and assODNs, confirming the specificity of these phenotypes. Overall, this is the first time to analyze the effects of the *HBB* –28(A>G) mutation at the whole-transcriptome level based on isogenic cell lines, providing a landscape for further investigation of the mechanism of β -thalassemia with the *HBB* –28(A>G) mutation.

Keywords: β -thalassemia, *HBB* –28(A>G), isogenic cells, K562, CRISPR/Cas9, assODNs, RNA-seq

INTRODUCTION

β -thalassemia is one of most common autosomal recessive blood diseases spread worldwide, caused by either point mutations or deletions in the β -globin (*HBB*) gene (de Dreuzy et al., 2016; Mansilla-Soto et al., 2016). Approximately 80–90 million people carried β -thalassemia, and about 300,000 children with severe thalassemia are born every year (Galanello and Origa, 2010; Weatherall et al., 2010; Origa, 2017). Mutations or deletions of β -globin genes result in the reduction of hemoglobin β chain (β -globin), deformation of hemoglobin tetramer, and subsequent lysis of erythrocytes, finally causing oxygen shortage, bone deformity, organ dysfunction, and even organ failure in many parts of the human body (Cao and Galanello, 2010; Liang et al., 2017). As for the thalassemia major patients, life-long blood transfusion and iron chelation treatments are required for survival, but often accompanied by numerous complications, including arrhythmia, congestive heart failure, hypothyroidism, hypoparathyroidism, hypogonadism, diabetes, osteoporosis, liver cirrhosis, and recurrent infections. Thus, thalassemia has threatened millions of people's lives for decades and is still a major public health issue (Xu et al., 2015; Kumar et al., 2016).

β -thalassemia mutations are prevalent in the southern part of China and Southeast Asia, and *HBB* –28(A>G) mutation is one of the five most common *HBB* mutations carried by β -thalassemia patients in China (Lai et al., 2017; Liang et al., 2017). In the *HBB* –28(A>G) mutation, adenine (A) base located at 28 base pairs upstream from the cap site is mutated to guanine (G), disrupting the binding of transcription factor of ATA box and decreasing the RNA expression of *HBB* (Orkin et al., 1983). Patients with the homozygous or compound heterozygous –28(A>G) mutation may develop severe anemia or intermedia anemia (Orkin et al., 1983; Cao and Galanello, 2010). Although the description of severe thalassemia was first published more than 90 years ago and a considerable amount of work has been reported to refine the understanding of the pathophysiology of thalassemia syndromes in the past 50 years, the cellular and molecular basis of this group of diseases has still not been thoroughly investigated (Taher et al., 2018). In particular, few studies have been conducted on the *HBB* –28(A>G) mutation to understand how this mutation affects gene expression at the transcriptome level although correcting the *HBB* –28(A>G) mutation with the base editing (BE) system has also been reported in human iPSCs and reconstituted embryos (Liang et al., 2017). Without a full understanding of the defects at the molecular level, it will be difficult to comprehensively evaluate the rescue effect after correcting the mutation.

In a recent study, high-throughput RNA sequencing (RNA-seq) was used to compare control samples with patient samples carrying a novel *HBB* mutation (*HBB*: c.51C > T). It shows that hemopoiesis, heme biosynthesis, response to oxidative stress, and other cellular activity pathways were directly or indirectly enriched by differentially expressed genes related to β -thalassemia (Taghavifar et al., 2019), suggesting that genome-wide RNA-seq analysis is a useful approach to understand the mechanism of β -thalassemia with different mutations. However, control samples in this study are allogenic, and different genetic

backgrounds and a mixture of short- and long-term effects would prevent a deep understanding of the effect of this mutation. The cell disease models of hemoglobinopathies have also been reported. Induced pluripotent stem cells (iPSCs) were generated by reprogramming patient-derived cells and provided a good model toward the development of iPSC-based therapy (Li et al., 2011). However, differentiation of human iPSCs into mature and β -globin expressing erythrocytes remains a major challenge in the field. A human umbilical cord blood-derived erythroid progenitor (HUDEP) cell line was established and used as a disease model because of its differentiation ability to produce enucleated red blood cells (Kurita et al., 2013; Traxler et al., 2016; Grevet et al., 2018). However, this cell line is derived from primary hematopoietic stem cells with E6/E7 immortalized basophilic erythroblasts, which depends on erythropoietin (EPO) and cytokine stem cell factor (SCF) for survival, so it is expensive to maintain and hard to do gene editing (Couch et al., 2019). Therefore, we chose the K562 cell line to generate the thalassemia disease cell model for its ready availability, cost-effective maintenance, relatively mature differentiation conditions, and high efficiency of gene editing (Dean et al., 1983; Witt et al., 2000; Ma et al., 2013; DeWitt et al., 2016; Shariati et al., 2016; Ghosh et al., 2018).

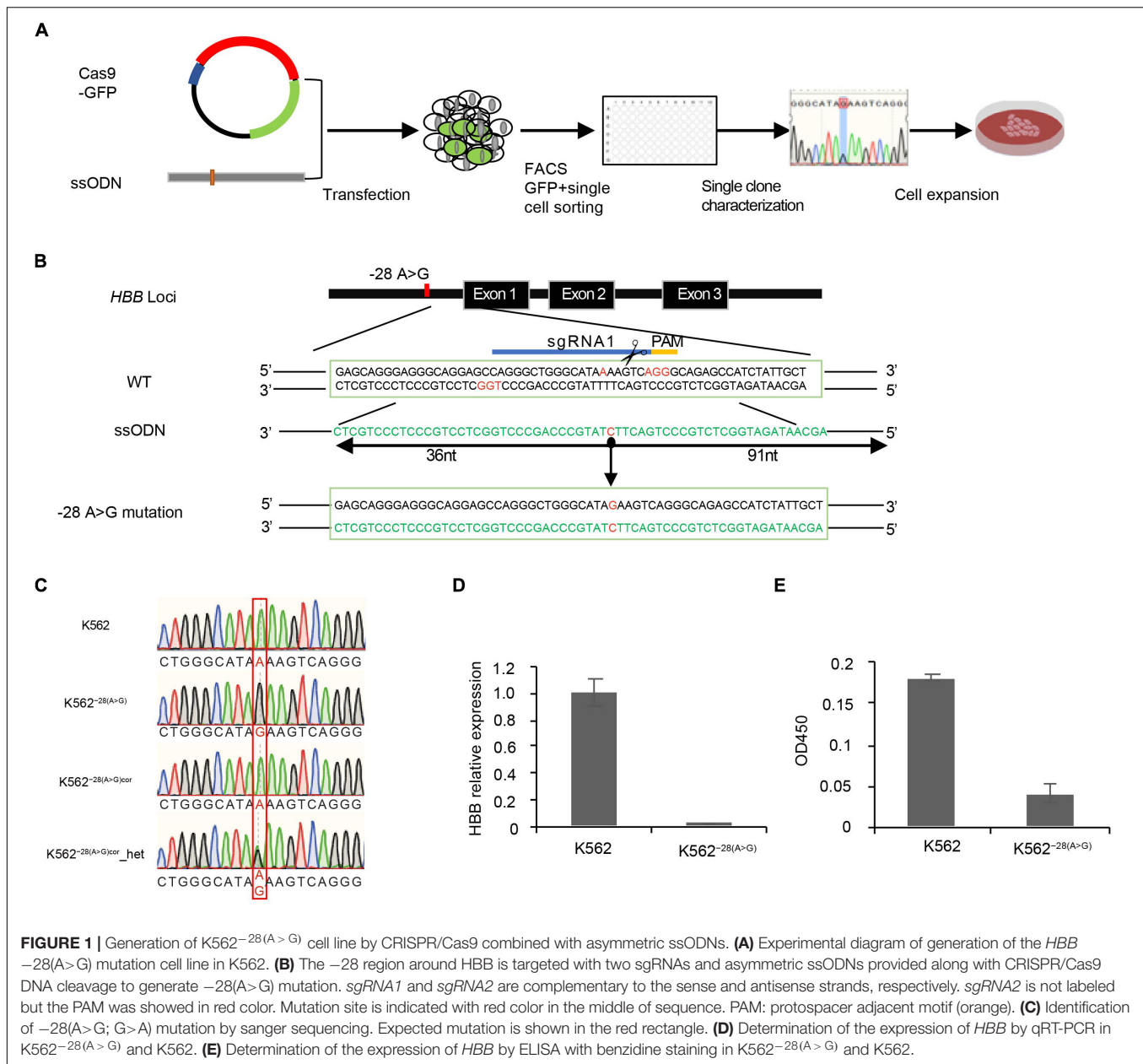
In order to explore the specific impact of the *HBB* –28(A>G) mutation on erythroid differentiation and how it affects genome-wide gene expression without confounding factors, such as different genetic background in allogenic samples, we used the CRISPR/Cas9 gene-editing system combined with asymmetric single-stranded oligodeoxynucleotides (ssODNs) to generate the disease model of isogenic K562 cell lines (Men et al., 2017; Zhang et al., 2017) and then conducted transcriptome analysis by high-throughput RNA sequencing. The mutant cell line K562^{–28(A>G)} showed no detectable *HBB* gene expression, enabling comparative studies in the same genetic background. We found that *HBB* was transcriptionally prevented by the mutation, and the K562^{–28(A>G)} cell line is more sensitive to hypoxia and present a defective erythrogenic program when compared with wild-type K562 before erythroid differentiation by Gene Ontology (GO) and KEGG analysis. Interestingly, *HBB* showed a lower rate of induction in K562^{–28(A>G)} when compared with wild-type K562 after erythroid differentiation. Taken together, our study is the first to analyze the effects of the *HBB* –28(A>G) mutation at the whole-transcriptome level based on isogenic cell lines. The unraveled molecular biomarkers and signaling pathways that affected in K562^{–28(A>G)} cell line may be further investigated to explore the study of the mechanism of β -thalassemia in the future.

MATERIALS AND METHODS

Generation of β -thalassemia cell line with –28(A>G) mutation

sgRNA Design and Construction

Two 20-bp sgRNAs were chosen containing the –28(A>G) site and the cutting sites were about 3 and 9 bp before the mutation site, respectively. The sequence shown in **Figure 1B**. The guiding RNA oligonucleotides were synthesized by BGI and inserted into



the gRNA cloning vector pSpCas9(BB)-2A-GFP (PX458) (Add gene 48138) according to the protocol provided by Zhang Feng's lab (Ran et al., 2013).

Design of ssODNs Repair Templates

The 127-nt asymmetric single-stranded oligodeoxynucleotides (ssODNs) templates were designed by overlapping the CRISPR/Cas9 cleavage site with 36 bp on the PAM-distal side and a 91-bp extension on the PAM-proximal side of the break (Niu et al., 2016; Paquet et al., 2016; **Figures 1A,B**) and were synthesized by BGI. Once the site of –28(A > G) was mutated by the ssODN template, the site for gRNA recognition was blocked, and Cas9 could not cut this site again. It is easily manipulated and seamless on the genome.

Gene Editing

To generate the *HBB* –28(A > G; G > A) mutation, 1×10^6 cells were collected and resuspended in 100 μ L electroporation solution (Buffer R) at a final density of 1×10^7 cells/mL, and then 3 μ g ssODNs and 5 μ g Cas9/*sgRNA* (Addgene 48138) were added into the 100 μ L resuspended cells. Electroporation was completed by the Neon Transfection System (Invitrogen). Electroporated cells were slowly loaded into growth medium without antibiotics and cultured in 5% CO₂ at 37°C. 48 h after transfection, flow cytometry was used to isolate the single cell into 96-well plates. 1 week later, we characterized the single cell clone by PCR and Sanger sequencing. The resistant colonies were picked and expanded until we generated the expected cell line.

Cell Culture and Differentiation

K562 cells were obtained from the American Tissue Culture Collection. The cells of K562 and K562^{-28(A>G)} were cultured in glutamine-minus RPMI 1640 medium (Gibco) with 10% FBS in the presence of 1 mM sodium butyrate for 7 days to induce the erythroid differentiation in humidified atmosphere of 5% CO₂ at 37°C (Shariati et al., 2016). The control cells were cultured in RPMI 1640 medium with 10% FBS and P/S antibiotics for 7 days in humidified atmosphere of 5% CO₂ at 37°C.

Quantitative Real-Time PCR (qRT-PCR) Analysis

Total RNA was isolated from the cells by the TRIzolTM Reagent (Invitrogen). Single-stranded cDNA was synthesized with the oligo(dT) primer using the PrimeScriptTM RT reagent Kit with gDNA Eraser (Takara), the obtained cDNAs were analyzed by real-time PCR, using the indicated primers. The *HBB* primers were 5'- GCTCGGTGCCTTTAGTGATG -3' (forward) and 5'- GCACACAGACCAGCACGTT -3' (reverse); for *HBG*, 5'- GGAAGATGCTGGAGGAGAAACC -3' (forward) and 5'- GTCAGCACCTTCTTGCCATGTG -3' (reverse). The cycling conditions were 95°C denaturation for 10 min, 95°C for 15 s, annealing and extension at 60°C for 40 s, 40 cycles on the ABI step one machine.

Comparison of the Hemoglobin Expression ELISA

To compare the hemoglobin expression of the uninduced cells [K562, K562^{-28(A>G)}, K562^{-28(A>G)cor}] and induced cells [K562-dif, K562^{-28(A>G)}-dif], we used the tetramethylbenzidine (TMB) ELISA Kit (Invitrogen): 1 × 10⁵ cells from each of the test and control groups were resuspended in phosphate buffered saline (PBS) and mixed with an equal volume of the TMB solution and then kept at room temperature for 30 minutes. Then add a double volume of TMB stop solution to the wells. Brown-blue wells were regarded as positive and read at 450 nm by Biotek epoch. Each treatment was performed in triplicate, and each experiment was repeated three times.

Flow Cytometry

FITC-conjugated CD235ab (BioLegend) was used as the erythroid-specific surface marker. About 5 × 10⁵ cells were harvested, counted, and resuspended in 100 mL PBS with Human BD Fc BlockTM (BD) for about 10 min at room temperature; then we added 100 mL FITC-conjugated CD235ab antibody diluted by 1:10 and the samples were incubated at 4°C for 30 min. Cells were washed with ice-cold PBS three times, and the fluorescence was measured by flow cytometry of BD Aria III. The unstained group was used as control for each sample [K562, K562^{-28(A>G)}, K562-dif, K562^{-28(A>G)}-dif].

RNA-Seq Analysis

There are more than 65 million 100 bp pair-ended reads per sample. Prior to assembly, raw reads were filtered by SOAPnuke (Chen et al., 2017) with the parameters "-l 15 -q 0.2 -n

0.05". Reads were mapped to the human reference genome using HISAT2 (Kim et al., 2015), which included both the genome sequences (GRCh38.p12) and known reference sequence (RefSeq) transcripts. Finally, 80% of reads were aligned to the human genome.

We used Samtools (Li et al., 2009) to sort and index the alignment files, and the Integrative Genomics Viewer (IGV) tool (Robinson et al., 2011) was used to visualize the reads. The sorted binary sequence analysis file (BAM files) was also used to generate UCSC browser tracks with a genome coverage bed from Bed Tools (Dennis et al., 2003; Quinlan and Hall, 2010). To this end, coverage files were normalized using the total signal for each sample.

The StringTie (Pertea et al., 2015) suite of tools was used to calculate and compare gene expression levels, which were normalized as fragments per kilobase of transcript per million (FPKM) mapped reads. Differential expression analysis was assessed at each sample by comparing the pre- and post-induction cells using the EdgeR Bioconductor package (Robinson et al., 2010). Differentially expressed genes (DEGs) were identified when the fold change (FC) > 2 and *p* < 0.05.

Gene Ontology (GO) Analysis

Gene ontology enrichment analysis of the gene sets (Yu et al., 2012) were performed on each sample using the cluster Profiler R package, including KEGG (Kanehisa and Goto, 2000) pathways and Gene Ontology (Ashburner et al., 2000)¹, which were collected in the molecular signatures database (MSigDB) (Liberzon et al., 2011). Meanwhile, DAVID² and Metascape were utilized to assess the enrichment of functional categories (GO and KEGG) of the DEGs (Dennis et al., 2003; Zhou et al., 2019).

Transcription Factor Prediction

To compare the expression levels of transcription regulator genes between pre- and post-induction K562 cells, we collected a comprehensive transcription factor annotation from AnimalTFDB 3.0 (Hu et al., 2018) and iRegulon (Janky et al., 2014) and the results were visualized using Cytoscape (Shannon et al., 2003).

RESULTS

Generation of *HBB* -28(A>G) Mutant Cell Line by CRISPR/Cas9 and Asymmetric Single-Stranded Oligodeoxynucleotides

To study how the *HBB* -28(A>G) mutation affects gene expression at the genome-wide level, an isogenic human cell line carrying this mutation was generated. The diagram of the generating mutant cell line is shown, including transfection, CRISPR/Cas9 editing, single cell sorting, cell characterization, and expansion (Figure 1A). As previous studies demonstrated

¹<http://www.geneontology.org/>

²<https://david.ncifcrf.gov/>

that using ssODNs (Niu et al., 2016) or asymmetric double-strand DNA (Paquet et al., 2016) as a repair template resulted in a higher efficiency of accurate replacement of target sequences through homology-directed repair (HDR), we developed a technology that combined CRISPR/Cas9 with asymmetric ssODNs (assODNs). To generate the *HBB* –28(A>G) mutation, sgRNA mediating DNA cleavage 3 bp aside from the –28 mutation site and assODNs with 36 bp on the PAM-distal side and 91 bp on the PAM-proximal side of the cutting site were used (Figure 1B). K562, a human erythroleukemia line that resembles undifferentiated erythrocytes, was transduced with these Cas/sgRNA and assODNs. We identified one clone with the *HBB* –28(A>G) mutation that showed a single peak at the mutation site by sanger sequencing and named it K562^{–28(A>G)} (Figure 1C). In order to exclude the possibility that the single peak of *HBB* –28(A>G) is generated by a large deletion of the other allele, we then performed CRISPR/Cas9-mediated gene correction on this K562^{–28(A>G)} cell clone and obtained heterozygosity on the mutation site in our heterozygous correction clone K562^{–28(A>G)cor}-het (Figure 1C), suggesting both alleles of *HBB* are present in the original mutant cell clone K562^{–28(A>G)} and, thus, confirming the homozygosity of K562^{–28(A>G)}. In order to reconfirm the functional effect of the *HBB* –28(A>G) mutation, we used qRT-PCR and ELISA to detect the expression of *HBB* mRNA and HBB protein, respectively. In agreement with the sequencing result, the expression of both *HBB* mRNA and HBB protein were undetectable in the K562^{–28(A>G)} cell line. In contrast, wild-type K562 shows a considerable expression of *HBB* mRNA and HBB protein even without erythroid differentiation (Figures 1D,E). These results suggest that the mutant cell line of *HBB* –28(A>G) was successfully generated in which the expression of *HBB* was eliminated.

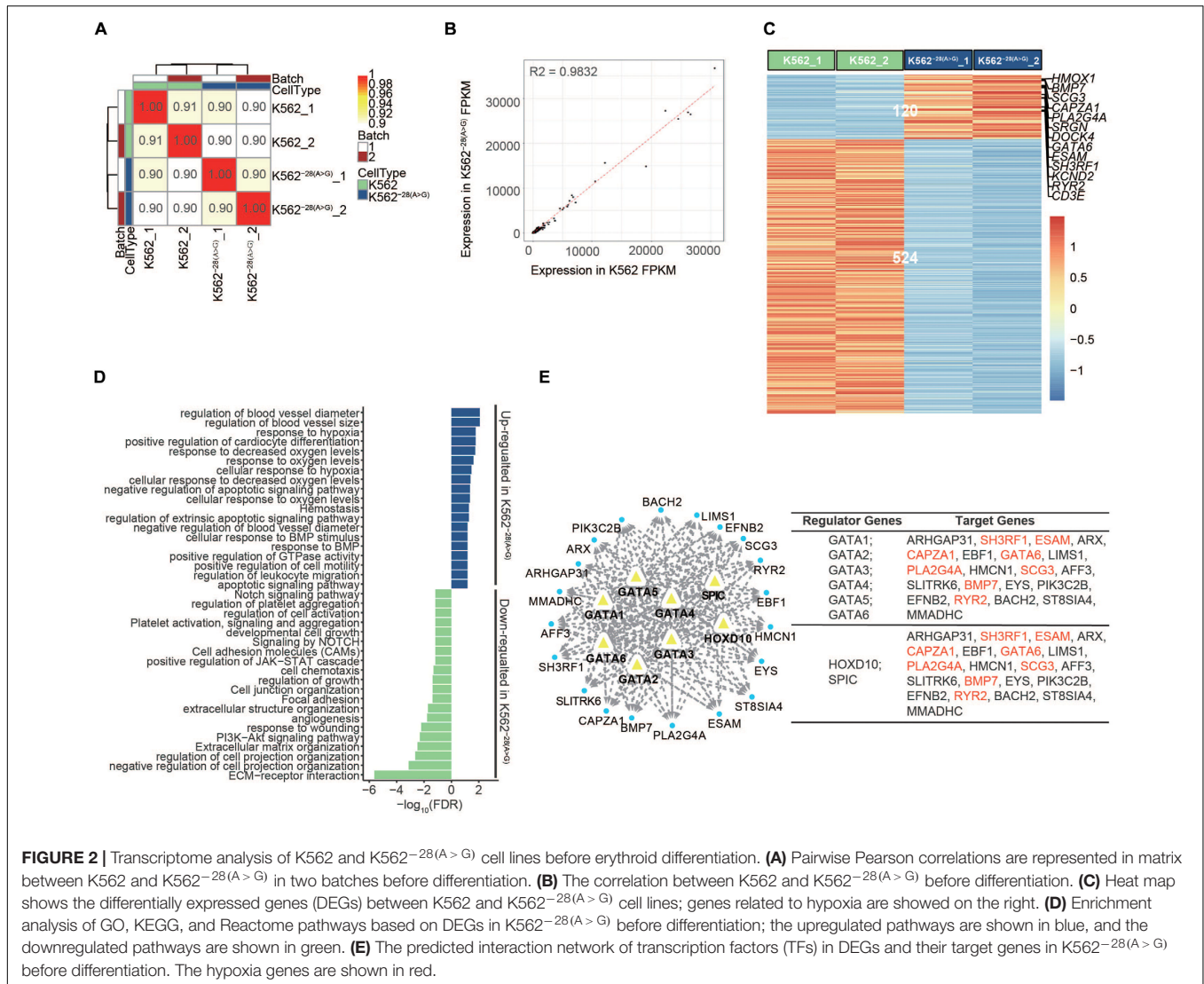
Transcriptome Analysis of K562^{–28(A>G)} and K562 Before Erythroid Differentiation

Many K562^{–28(A>G)} cells appear to have irregular morphology and spontaneous cell death compared to the wild-type counterpart, suggesting loss of *HBB* expression by –28(A>G) mutation may compromise cell viability in normal culture condition (Supplementary Figure S1). To understand how the mutation affects the molecular function of cells at the transcriptional level, we conducted RNA-seq to analyze the transcriptome differences between the isogenic cell line of K562^{–28(A>G)} and its control K562. Pairwise Pearson correlation analysis revealed high similarity between replicates from K562^{–28(A>G)} and K562 cells, indicating high reproducibility of our data (Figure 2A). Overall, the gene expression levels between K562^{–28(A>G)} and K562 cell are similar on a transcriptome level, suggesting that the mutation may not affect the cell identity (Figure 2B). We conducted analysis of differentially expressed genes (DEGs) and found 120 and 524 genes were consistently upregulated and downregulated, respectively, in K562^{–28(A>G)} compared to K562 (Figure 2C). Interestingly, K562^{–28(A>G)} had a higher expression of genes related to the mitochondrial electron transport chain, such as MT-CO1 and MT-ND,

implying that the mutant cell line may need more energy and be subject to oxidative stress (Supplementary Table S2). To further explore the affected underlying biological functions, we conducted GO term, KEGG, and Reactome analysis using those DEGs. Interestingly, the PI3K-Akt signaling pathway that is important for erythrocyte differentiation (Jafari et al., 2019) was downregulated in K562^{–28(A>G)} (Figure 2D). Furthermore, we found pathways of (cellular) response to hypoxia and response to (decreased) oxygen level were upregulated in K562^{–28(A>G)} mutant cell line (Figure 2D). Consistently, hypoxia-related genes, such as *HMOX1*, *BMP7*, *GATA6*, *ESAM*, and *RYR2* were upregulated in K562^{–28(A>G)} (Figure 2C). To further explore the core regulators of hypoxia, we performed an interaction assay to predict the interaction of hypoxia-related DEGs and transcription factors (TFs) that target upregulated genes in K562^{–28(A>G)}. In agreement with previous results, we observed the hypoxia-related genes, such as *HMOX1* and *SRGN* on the PPI network, and the GATA family, *HOXD10* and *SPIC*, which are well-known regulators of erythroid differentiation and hypoxia (Myers et al., 2002; Bennett et al., 2019) were the core regulators for upregulated genes in K562^{–28(A>G)} (Figure 2E and Supplementary Figure S3). The regulators and their corresponsive target genes were listed, and genes related to the hypoxia response are labeled in red (Figure 2E). Taken together, these data suggest the hypoxia response was upregulated in K562^{–28(A>G)} and the core regulators were the GATA family, *HOXD10* and *SPIC*.

Transcriptome Analysis of Coregulated Genes in K562^{–28(A>G)} and K562 Cell Lines After Erythroid Differentiation

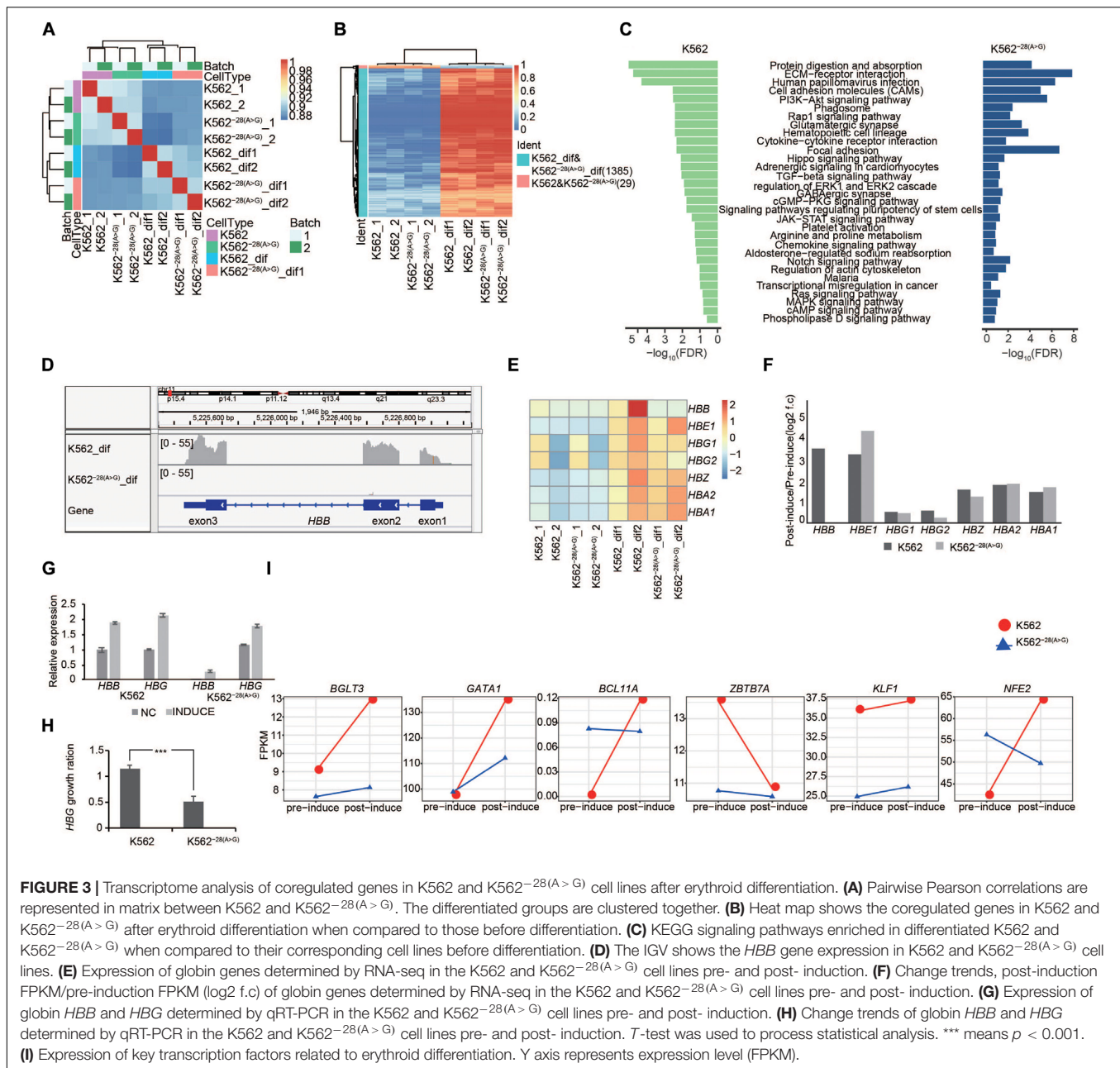
The K562 cell line is often used to study erythroid differentiation *in vitro*. In order to investigate the effect of the *HBB* –28(A>G) mutation during erythroid differentiation, we induced erythroid differentiation using sodium butyrate, confirmed the differentiation efficiency (Supplementary Figure S6), and then performed genome-wide RNA-Seq in the K562 and K562^{–28(A>G)} mutant cell line. Pairwise Pearson correlations represented in the matrix indicated that the isogenic cell lines showed high similarities up to 90%, and the samples were clustered together depending on their differentiation conditions (Figure 3A). The expression of 1,385 genes were consistently upregulated in differentiated K562^{–28(A>G)} and K562 when compared to undifferentiated samples (Figure 3B and Supplementary Figure S4A), which means the mutation couldn't prevent the process of erythroid differentiation, and these DEGs were used to conduct the GO term, KEGG, and Reactome analysis. Multiple signaling pathways were activated in both K562^{–28(A>G)} and K562 (Figure 3C), including PI3K-Akt, MAPK, and ERK pathways that have been reported to be activated during erythroid differentiation (Witt et al., 2000; Yang et al., 2001; Grass et al., 2003; Jafari et al., 2019). In addition, signal pathways, such as cell adhesion, pluripotency of stem cells, platelet activation, and Notch pathway, were also coactivated in differentiated samples (Figure 3C), suggesting our induction process was successful



in both cell lines. We also performed a comparative analysis of erythrocytes pre- and post- induction of K562^{-28(A>G)} and K562. GO analysis showed that erythrocytes take up carbon dioxide and release oxygen pathway and O₂/CO₂ exchange in the erythrocytes pathway were enriched in mutant cell lines (**Supplementary Figures S4B,C**).

HBB belongs to the globin gene family, mainly including *HBA* (α -globin), *HBG* (γ -globin), and *HBE* (ϵ -globin). *HBB*, *HBG*, and *HBE* are β -globin-like and all capable of forming a tetramer with α -globin. They are located within a gene cluster on chromosome 11, and their expression is coordinated by the same locus control region (LCR) as other regulatory DNA elements (Palstra et al., 2008). In β -thalassemia patients, *HBG* may be upregulated to compensate the loss of *HBB*. To study the compensatory gene expression of globin genes in K562^{-28(A>G)} after differentiation [K562^{-28(A>G)}-dif], the expression of *HBB*, *HBA*, and *HBG* was analyzed by RNA-seq in isogenic cell lines of K562^{-28(A>G)} and K562. As expected,

IGV analysis showed that *HBB* expression was induced but undetectable in K562^{-28(A>G)} after differentiation (**Figure 3D**). In contrast, expression of other globin genes was induced and detected in K562^{-28(A>G)} after differentiation (**Figure 3E**). We noticed that the fold-induction rates of *HBA1*, *HBA2*, and *HBE1* were similar or slightly increased while *HBZ* slightly decreased in K562^{-28(A>G)}-dif when compared to those in K562-dif. However, *HBG* expression in K562^{-28(A>G)}-dif was two-fold lower by FPKM than that in K562-dif (**Figure 3F**). The expression of *HBB* and *HBG* before and after differentiation were further confirmed by qRT-PCR. Consistently, expression of *HBB* was undetectable in undifferentiated K562^{-28(A>G)}, and its expression level in K562^{-28(A>G)}-dif after differentiation was negligible when compared to K562-dif control (**Figure 3G**). In agreement with RNA-seq results, the expression level of *HBG* was increased in both cell lines after differentiation, but the induction level of *HBG* caused by differentiation was nearly two-fold lower in K562^{-28(A>G)} than that in



K562 (Figures 3G,H). We further analyzed the expression of key TFs that are important for the regulation of globin genes during erythroid differentiation. *KLF1*, a previously reported positive regulator of *HBG* expression (Zhou et al., 2010), is found downregulated in K562^{-28(A>G)}. Expression of other positive regulators, including *GATA1*, *BGL3*, and *NFE2* (Grass et al., 2003; Stamatoyannopoulos, 2005; Zhou et al., 2010; Wienert et al., 2018), were dramatically upregulated in K562 after differentiation, and those increases were largely attenuated in K562^{-28(A>G)}. In contrast, negative regulators *BCL11A* and *ZBTB7A* were upregulated in K562^{-28(A>G)} compared to K562 before differentiation (Figure 3I and Supplementary Figure S5). The similar expression pattern of

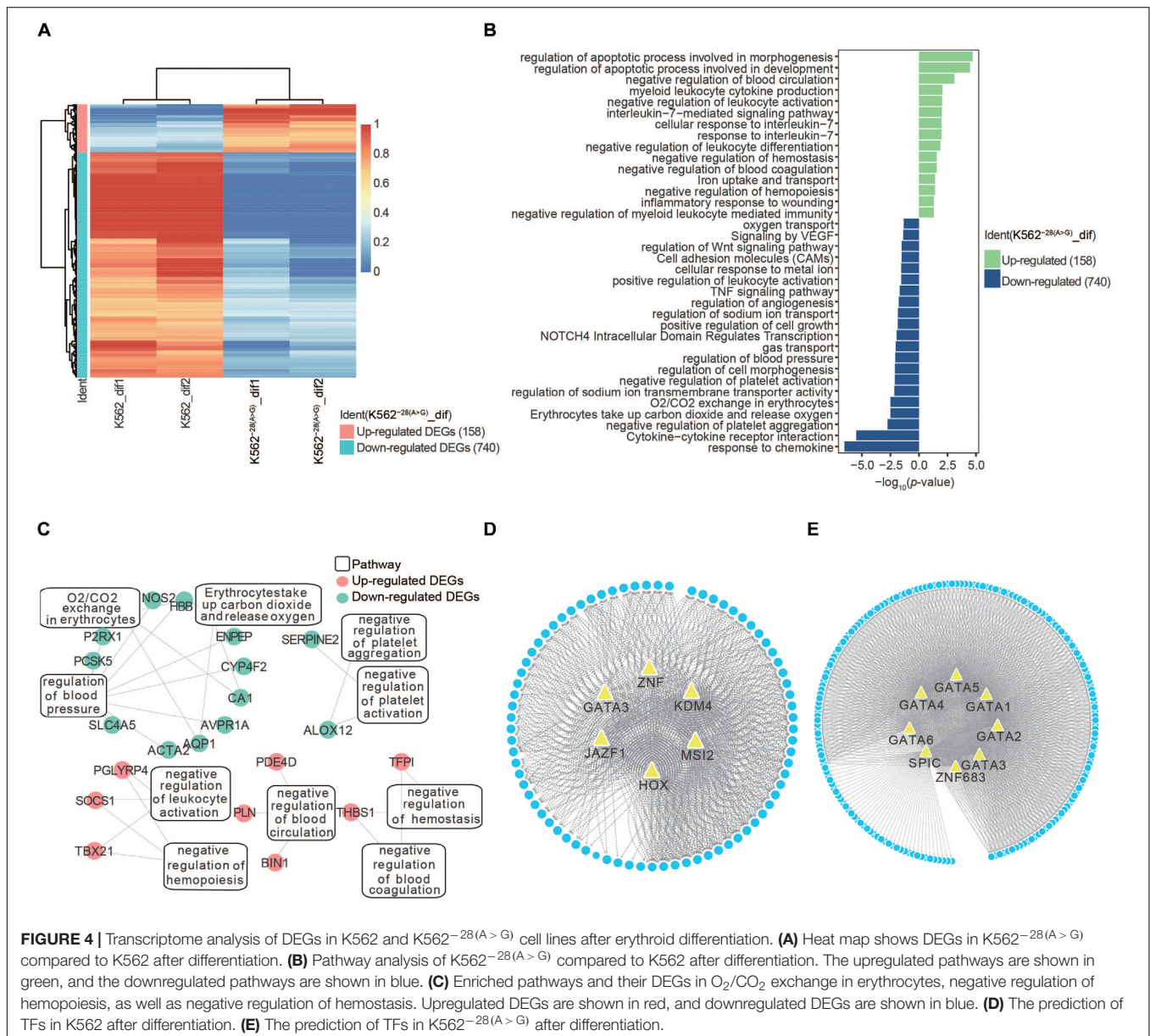
globin genes and key translation factors between K562^{-28(A>G)} and K562 suggest the erythroid differentiation is generally normal in K562^{-28(A>G)}, but the mutation of *HBB* with -28(A>G) may affect the *HBG* expression through a network of transcription factors.

Transcriptome Analysis of DEGs in K562 and K562^{-28(A>G)} Cell Lines After Erythroid Differentiation

To understand the effect of the *HBB* -28(A>G) mutation during erythroid differentiation, we identified the DEGs (*p* < 0.05, fold change [FC] > 2) within K562^{-28(A>G)}-dif

and K562-dif. The number of upregulated and downregulated DEGs in K562^{-28(A>G)} were 158 and 740, respectively (Figure 4A). With the GO term, KEGG, and Reactome analysis, we found that upregulated DEGs in K562^{-28(A>G)}-dif were enriched in pathways related to stress-response and hematopoiesis disorder, such as regulation of the apoptosis process, negative regulation of leukocyte activation, myeloid leukocyte cytokine production, negative regulation of blood coagulation, negative regulation of hemostasis, and negative regulation of hemopoiesis. Meanwhile, downregulated DEGs in K562^{-28(A>G)}-dif were enriched in oxygen-related pathways, including oxygen transport; erythrocytes take up carbon dioxide and release oxygen and O₂/CO₂ exchange in erythrocytes (Figure 4B). Pathway analysis of critical DEGs in K562^{-28(A>G)}-dif reveal the oxygen

transport and blood circulation were enriched (Figure 4C). And *PDE4D*, *TFPI*, *CA1*, and *AQP1*, which were related to hypoxia, were found to be critical targets (Figure 4C and Supplementary Table S1). Moreover, TF prediction revealed that JAZF1, MSI2, KDM4, HOX family, and ZNF family were core regulators for upregulated genes in K562^{-28(A>G)}-dif (Figure 4D) while the SPIC and GATA family are core regulators for downregulated genes in K562^{-28(A>G)}-dif (Figure 4E). Interestingly, GATA3 was found to be the core regulator for both upregulated and downregulated genes (Figures 4D,E). Taken together, dysregulated genes and pathways were present in K562^{-28(A>G)} after differentiation and transcription factors, such as GATA, HOX, and ZNF families, may play important roles as core regulators.



Reversion of Observed Abnormalities in Corrected K562^{-28(A>G)_{cor}} Cell Line

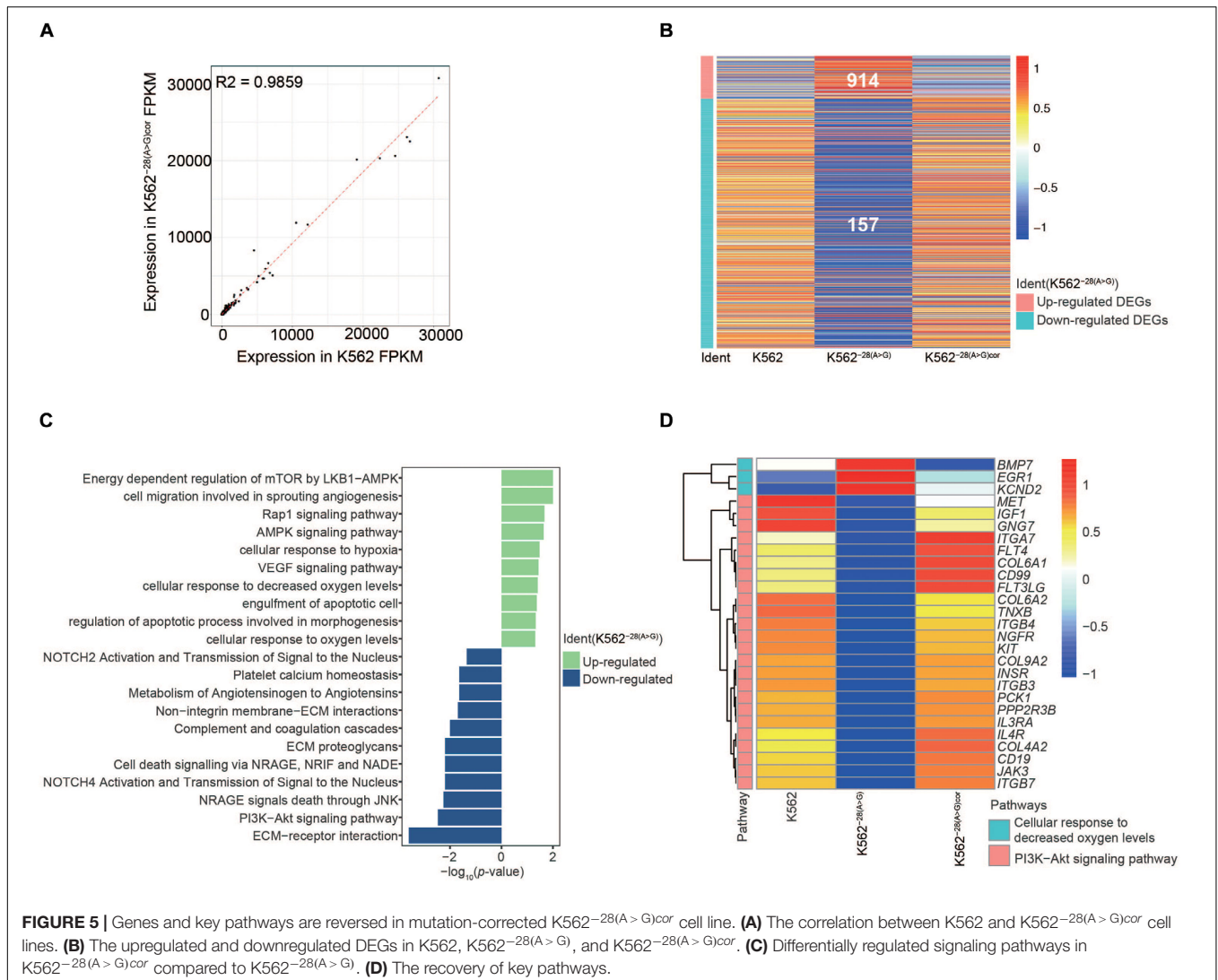
As it took weeks for the generation of K562 mutant cells by CRISPR/Cas9 and clone selection, it is possible that the differential gene expression is due to secondary effects of the editing process, similar to studies of patient samples (Taghavifar et al., 2019). In order to exclude the potential influence on secondary effects on our transcriptome analyses, we further corrected the mutation in K562^{-28(A>G)} as mentioned before and used a homozygous correction cell line K562^{-28(A>G)_{cor}} to perform transcriptome comparison (Figure 1C and Supplementary Figure S2). As expected, we found the expression profile of K562^{-28(A>G)_{cor}} had a high relationship with K562 (Figure 5A). Importantly, DEGs heatmap shows a highly similar pattern between K562^{-28(A>G)_{cor}} and K562 (Figure 5B), indicating that the editing process by CRISPR/Cas9 did not generate obvious effects on transcriptome. Through analysis of pathway enrichment, the PI3K pathway and cell response to the oxygen pathway were also recovered

in K562^{-28(A>G)_{cor}} (Figure 5C). The genes of key pathways were also recovered (Figure 5D). In summary, a large portion of abnormalities observed in K562^{-28(A>G)} are reversed in corrected K562^{-28(A>G)_{cor}}, suggesting these phenotypes are specifically caused by mutation of *HBB* -28(A>G) and not caused by editing process.

DISCUSSION

In our study, we successfully generated the cell line K562^{-28(A>G)} of *HBB* with the -28(A>G) mutation using CRISPR/Cas9 and a 127 nt assODN. This cell-based disease model was used to study how the *HBB* -28(A>G) mutation affected the cellular function on the transcriptome level before and after erythroid differentiation.

Our results showed the *HBB* -28(A>G) mutation prevented the transcription of *HBB* gene. Analysis of enriched pathways suggested the PI3K pathway as well as the JAK-STAT pathway,



which play important roles in erythroid differentiation, were disrupted in K562^{–28(A>G)} before erythroid differentiation. The PI3K-Akt signaling pathway is a significant pathway that controls many cellular processes known as cell division, autophagy, survival, and differentiation (Jafari et al., 2019). Moreover, the mutation activated the hypoxia pathway in undifferentiated K562^{–28(A>G)}. Many clinical manifestations observed in β -thalassemia are attributed to the chronic hypoxic environment due to pathologic erythrocyte production, and our data suggest hematopoietic precursors may also be subject to oxidative stress before differentiation.

To induce erythroid differentiation, we chose the glutamine-minus medium with sodium butyrate, as hemoglobin synthesis was markedly induced using this condition with a differentiation efficiency of 11–19% in K562 (Qun and Yuan, 2007; Bruchova-Votavova et al., 2010). Consistent with previous reports, the differentiation efficiency of K562 in our study was nearly 12% (Supplementary Figure S6), indicating that our erythroid differentiation is effective. In agreement, the MAPK and ERK pathway was activated in both K562 and K562^{–28(A>G)} (Figure 2), a finding consistent with observations in previous studies (Witt et al., 2000; Park et al., 2001). Interestingly, the PI3K-Akt signaling pathway was activated in K562^{–28(A>G)} after induction, suggesting the defective PI3K pathway may be caused by a lack of activators in undifferentiated K562^{–28(A>G)}. Other pathways, such as cell adhesion, pluripotency of stem cells, platelet activation, and Notch pathway, were also coactivated in differentiated K562 and K562^{–28(A>G)} samples, indicating mutation of *HBB* –28(A>G) did not block the pathways required for differentiation. Nevertheless, consistent with data from undifferentiated K562^{–28(A>G)}, oxygen-related pathways were downregulated in differentiated K562^{–28(A>G)} (Figure 4). In both undifferentiated and differentiated conditions, SPIC and GATA families are predicted as core regulators (Figures 2, 4). The GATA family of transcription factors (GATA1–6) are essential for normal hematopoiesis and a multitude of other developmental processes (Grass et al., 2003; Li et al., 2012; Alsayegh et al., 2019; Castano et al., 2019). *GATA-1* regulates terminal differentiation and the function of erythroid, which activates or represses the erythroid-specific gene, such as β -globin locus-binding protein, and it might regulate the switch of fetal to adult hemoglobin in human (Sankaran et al., 2010). Interestingly, increased expression of *GATA1* was largely attenuated in K562^{–28(A>G)} during erythroid differentiation (Figure 3), which may play a role for dysregulated oxygen-related pathways.

As improving the levels of *HBG* in adults could partially reverse the severity of symptoms in sickle disease and β -thalassemia, it is important to understand the coordinated regulation between *HBB* and *HBG* (Chen et al., 2009; Traxler et al., 2016; Li et al., 2018). In this study, we noticed that fold induction of *HBG* was decreased in K562^{–28(A>G)}. *ZBTB7A* and *BCL11A* were two major repressors of *HBG* by directly binding the *HBG* gene promoters (Masuda et al., 2016; Liu et al., 2018; Martyn et al., 2018; Wienert et al., 2018). Expression of *ZBTB7A* was decreased during differentiation in K562, consistent with increased expression of *HBG*. However, the overall expression

level was higher in K562 than that in K562^{–28(A>G)}, paradoxical with the results of decreased fold induction of *HBG* in a mutant cell line. Although expression of *BCL11A* was lower in K562 when compared to K562^{–28(A>G)} before differentiation, its expression increased dramatically during differentiation (Figure 3I and Supplementary Figure S5). Our uncoordinated expression data suggest that the regulation of *HBG* by *ZBTB7A* or *BCL11A* should be a comprehensive process, and the detailed interaction between factors on *HBG* regulation requires further investigation. In summary, our RNA-seq data from isogenic K562 cell models suggest multiple pathways, including PI3K-AKT, JAK-STAT, and oxygen-related pathways, and GATA families may be involved in thalassemia pathogenesis caused by *HBB* –28(A>G) mutation. However, these results may require further validation in clinical samples or alternative models.

Last, we not only used CRISPR/Cas9 to generate mutation, but also corrected the mutation with the same strategy. In one aspect, reversed abnormalities in corrected cell line confirmed the specificities of these phenotypes. In another aspect, the efficient editing results indicate our gene editing strategy using assODNs is powerful and provide means for gene editing treatment of *HBB* –28(A>G) mutation.

In summary, we show the isogenic K562^{–28(A>G)} cell line generated with CRISPR/Cas9 and assODNs is a valuable model to evaluate the β -thalassemia with homozygous mutation of *HBB* with –28(A>G). It provides us with the first transcriptome data for mechanistic study on the effects of *HBB* –28(A>G), and the mutant cell line generated in our study can serve as a cellular disease model to screen lentivirus vectors in potential gene therapies or to evaluate the gene-editing strategies targeting *HBB* –28(A>G) mutation.

DATA AVAILABILITY STATEMENT

The datasets for this study can be found in the CNSA. Please see the <https://db.cngb.org/cnsa/> of CNGBdb with accession code CNP0000981 for more details.

AUTHOR CONTRIBUTIONS

YG, CL, and XZ conceived and designed the study. JL performed experiment from LG, TL, GD, and WO. ZZ and H-XS performed bioinformatics analysis. JL, ZZ, YG, and CL wrote the manuscript with the inputs from all authors. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Natural Science Foundation of China (Nos. 81903159 and 31970816), the Guangdong Provincial Key Laboratory of Genome Read and Write (No. 2017B030301011), and the Shenzhen Municipal Government of China (No. 20170731162715261).

ACKNOWLEDGMENTS

This manuscript has been released as a pre-print at biorxiv (Li et al., 2020), <https://www.biorxiv.org/content/10.1101/2020.06.18.159004v1>. We would like to thank Yanmei Deng, Jiaying Zhang, and Wenwen Yao for helping single cell identification. We thank Keiichi Suzuki for the suggestion on the strategy of gene editing. We thank the Genome Synthesis and Editing Platform of the China National Gene Bank for providing support on gene

synthesis. We also thank BGI colleagues for helping to output the high-quality data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.577053/full#supplementary-material>

REFERENCES

- Alsayegh, K., Cortes-Medina, L. V., Ramos-Mandujano, G., Badraiq, H., and Li, M. (2019). Hematopoietic differentiation of human pluripotent stem cells: HOX and GATA transcription factors as master regulators. *Curr. Genom.* 20, 438–452. doi: 10.2174/1389202920666191017163837
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25:25.
- Bennett, L. F., Liao, C., Quickel, M. D., Yeoh, B. S., Vijay-Kumar, M., Hankey-Giblin, P., et al. (2019). Inflammation induces stress erythropoiesis through heme-dependent activation of SPI-C. *Sci. Signal.* 12:eaa7336. doi: 10.1126/scisignal.aap7336
- Bruchova-Votavova, H., Yoon, D., and Prchal, J. T. (2010). MiR-451 enhances erythroid differentiation in K562 cells. *Leuk. Lymphoma* 51, 686–693. doi: 10.3109/10428191003629362
- Cao, A., and Galanello, R. (2010). Beta-thalassemia. *Genet. Med.* 12, 61–76. doi: 10.1097/GIM.0b013e3181cd68ed
- Castano, J., Aranda, S., Bueno, C., Calero-Nieto, F. J., Mejia-Ramirez, E., Mosquera, J. L., et al. (2019). GATA2 promotes hematopoietic development and represses cardiac differentiation of human mesoderm. *Stem Cell Rep.* 13, 515–529. doi: 10.1016/j.stemcr.2019.07.009
- Chen, Y., Chen, Y., Shi, C., Huang, Z., Zhang, Y., Li, S., et al. (2017). SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* 7:gix120.
- Chen, Z., Luo, H. Y., Steinberg, M. H., and Chui, D. H. (2009). BCL11A represses HBG transcription in K562 cells. *Blood Cells Mol. Dis.* 42, 144–149. doi: 10.1016/j.bcmd.2008.12.003
- Couch, T., Murphy, Z., Getman, M., Kurita, R., Nakamura, Y., and Steiner, L. A. (2019). Human erythroblasts with c-Kit activating mutations have reduced cell culture costs and remain capable of terminal maturation. *Exp. Hematol.* 74, 19–24.e14. doi: 10.1016/j.exphem.2019.04.001
- de Dreuzy, E., Bhukhai, K., Leboulch, P., and Payen, E. (2016). Current and future alternative therapies for beta-thalassemia major. *Biomed. J.* 39, 24–38. doi: 10.1016/j.bj.2015.10.001
- Dean, A., Ley, T. J., Humphries, R. K., Fordis, M., and Schechter, A. N. (1983). Inducible transcription of five globin genes in K562 human leukemia cells. *Proc. Natl. Acad. Sci. U.S.A.* 80, 5515–5519. doi: 10.1073/pnas.80.18.5515
- Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., et al. (2003). DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* 4:R60.
- DeWitt, M. A., Magis, W., Bray, N. L., Wang, T., Berman, J. R., Urbinati, F., et al. (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci. Transl. Med.* 8:360ra134. doi: 10.1126/scitranslmed.aaf9336
- Galanello, R., and Origa, R. (2010). Beta-thalassemia. *Orphanet. J. Rare Dis.* 5:11. doi: 10.1186/1750-1172-5-11
- Ghosh, A., Garee, G., Sweeny, E. A., Nakamura, Y., and Stuehr, D. J. (2018). Hsp90 chaperones hemoglobin maturation in erythroid and nonerythroid cells. *Proc. Natl. Acad. Sci. U.S.A.* 115, E1117–E1126. doi: 10.1073/pnas.1717993115
- Grass, J. A., Boyer, M. E., Pal, S., Wu, J., Weiss, M. J., and Bresnick, E. H. (2003). GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8811–8816. doi: 10.1073/pnas.1432147100
- Grevet, J. D., Lan, X., Hamagami, N., Edwards, C. R., Sankaranarayanan, L., Ji, X., et al. (2018). Domain-focused CRISPR screen identifies HRI as a fetal hemoglobin regulator in human erythroid cells. *Science* 361, 285–290. doi: 10.1126/science.aao0932
- Hu, H., Miao, Y.-R., Jia, L.-H., Yu, Q.-Y., Zhang, Q., and Guo, A.-Y. (2018). AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res.* 47, D33–D38.
- Jafari, M., Ghadami, E., Dadkhah, T., and Akhavan-Niaki, H. (2019). PI3k/AKT signaling pathway: Erythropoiesis and beyond. *J. Cell Physiol.* 234, 2373–2385. doi: 10.1002/jcp.27262
- Janky, R., Verfaillie, A., Imrichova, H., Van De Sande, B., Standaert, L., Christiaens, V., et al. (2014). iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput. Biol.* 10:e1003731. doi: 10.1371/journal.pcbi.1003731
- Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30.
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12:357. doi: 10.1038/nmeth.3317
- Kumar, S. R., Markusic, D. M., Biswas, M., High, K. A., and Herzog, R. W. (2016). Clinical development of gene therapy: results and lessons from recent successes. *Mol. Ther. Methods Clin. Dev.* 3:16034. doi: 10.1038/mtm.2016.34
- Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroyama, T., Miyoshi, H., et al. (2013). Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS One* 8:e59890. doi: 10.1371/journal.pone.0059890
- Lai, K., Huang, G., Su, L., and He, Y. (2017). The prevalence of thalassemia in mainland China: evidence from epidemiological surveys. *Sci. Rep.* 7:920. doi: 10.1038/s41598-017-00967-2
- Li, B., Ding, L., Li, W., Story, M. D., and Pace, B. S. (2012). Characterization of the transcriptome profiles related to globin gene switching during in vitro erythroid maturation. *BMC Genom.* 13:153. doi: 10.1186/1471-2164-13-153
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, J., Lai, Y., and Shi, L. (2018). BCL11A down-regulation induces gamma-globin in human beta-thalassemia major erythroid cells. *Hemoglobin* 42, 225–230. doi: 10.1080/03630269.2018.1515774
- Li, J., Zhou, Z., Sun, H.-X., Ouyang, W., Dong, G., Liu, T., et al. (2020). Transcriptome analyses of β -thalassemia –28 (A>G) mutation using isogenic cell models generated by CRISPR/Cas9 and asymmetric single-stranded oligodeoxynucleotides (assODN). *bioRxiv* [Preprint],
- Li, M., Suzuki, K., Qu, J., Saini, P., Dubova, I., Yi, F., et al. (2011). Efficient correction of hemoglobinopathy-causing mutations by homologous recombination in integration-free patient iPSCs. *Cell Res.* 21, 1740–1744. doi: 10.1038/cr.2011.186
- Liang, P., Ding, C., Sun, H., Xie, X., Xu, Y., Zhang, X., et al. (2017). Correction of beta-thalassemia mutant by base editor in human embryos. *Protein Cell* 8, 811–822. doi: 10.1007/s13238-017-0475-6
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. doi: 10.1093/bioinformatics/btr260
- Liu, N., Hargreaves, V. V., Zhu, Q., Kurland, J. V., Hong, J., Kim, W., et al. (2018). Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell* 173, 430–442.e417. doi: 10.1016/j.cell.2018.03.016

- Ma, Y. N., Chen, M. T., Wu, Z. K., Zhao, H. L., Yu, H. C., Yu, J., et al. (2013). Emodin can induce K562 cells to erythroid differentiation and improve the expression of globin genes. *Mol. Cell Biochem.* 382, 127–136. doi: 10.1007/s11010-013-1726-3
- Mansilla-Soto, J., Riviere, I., Boulad, F., and Sadelain, M. (2016). Cell and gene therapy for the beta-thalassemias: advances and prospects. *Hum. Gene Ther.* 27, 295–304. doi: 10.1089/hum.2016.037
- Martyn, G. E., Wienert, B., Yang, L., Shah, M., Norton, L. J., Burdach, J., et al. (2018). Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. *Nat. Genet.* 50, 498–503. doi: 10.1038/s41588-018-0085-0
- Masuda, T., Wang, X., Maeda, M., Canver, M. C., Sher, F., Funnell, A. P. W., et al. (2016). Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science* 351, 285–289. doi: 10.1126/science.aad3312
- Men, K., Duan, X., He, Z., Yang, Y., Yao, S., and Wei, Y. (2017). CRISPR/Cas9-mediated correction of human genetic disease. *Sci. China Life Sci.* 60, 447–457. doi: 10.1007/s11427-017-9032-4
- Myers, C., Charboneau, A., Cheung, L., Hanks, D., and Boudreau, N. (2002). Sustained expression of homeobox D10 inhibits angiogenesis. *Am. J. Pathol.* 161, 2099–2109. doi: 10.1016/S0002-9440(10)64488-4
- Niu, X., He, W., Song, B., Ou, Z., Fan, D., Chen, Y., et al. (2016). Combining Single Strand oligodeoxynucleotides and CRISPR/Cas9 to correct gene mutations in beta-thalassemia-induced pluripotent stem cells. *J. Biol. Chem.* 291, 16576–16585. doi: 10.1074/jbc.M116.719237
- Origa, R. (2017). Beta-Thalassemia. *Genet. Med.* 19, 609–619. doi: 10.1038/gim.2016.173
- Orkin, S. H., Sexton, J. P., Cheng, T. C., Goff, S. C., Giardina, P. J., Lee, J. I., et al. (1983). ATA box transcription mutation in beta-thalassemia. *Nucleic Acids Res.* 11, 4727–4734. doi: 10.1093/nar/11.14.4727
- Palstra, R. J., De Laat, W., and Grosveld, F. (2008). Beta-globin regulation and long-range interactions. *Adv. Genet.* 61, 107–142. doi: 10.1016/S0065-2660(07)00004-1
- Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., et al. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533, 125–129. doi: 10.1038/nature17664
- Park, J. I., Choi, H. S., Jeong, J. S., Han, J. Y., and Kim, I. H. J. C. G. D. (2001). Involvement of p38 kinase in hydroxyurea-induced differentiation of K562 cells. *Cell Growth Differ.* 12, 481–486.
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33:290. doi: 10.1038/nbt.3122
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Qun, H. E., and Yuan, L. B. (2007). Dopamine inhibits proliferation, induces differentiation and apoptosis of K562 leukaemia cells. *Chin. Med. J.* 120, 970–974. doi: 10.1097/00029330-200706010-00006
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308. doi: 10.1038/nprot.2013.143
- Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26. doi: 10.1038/nbt.1754
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Sankaran, V. G., Xu, J., and Orkin, S. H. (2010). Advances in the understanding of haemoglobin switching. *Br. J. Haematol.* 149, 181–194. doi: 10.1111/j.1365-2141.2010.08105.x
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Shariati, L., Modares, M., Khanahmad, H., Hejazi, Z., Tabatabaiefar, M. A., Salehi, M., et al. (2016). Comparison of different methods for erythroid differentiation in the K562 cell line. *Biotechnol. Lett.* 38, 1243–1250. doi: 10.1007/s10529-016-2101-8
- Stamatoyannopoulos, G. (2005). Control of globin gene expression during development and erythroid differentiation. *Exper. Hematol.* 33, 259–271. doi: 10.1016/j.exphem.2004.11.007
- Taghavifar, F., Hamid, M., and Shariati, G. (2019). Gene expression in blood from an individual with beta-thalassemia: An RNA sequence analysis. *Mol. Genet. Genom. Med.* 7:e00740. doi: 10.1002/mgg3.740
- Taher, A. T., Weatherall, D. J., and Cappellini, M. D. (2018). Thalassaemia. *Lancet* 391, 155–167. doi: 10.1016/s0140-6736(17)31822-6
- Traxler, E. A., Yao, Y., Wang, Y. D., Woodard, K. J., Kurita, R., Nakamura, Y., et al. (2016). A genome-editing strategy to treat beta-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat. Med.* 22, 987–990. doi: 10.1038/nm.4170
- Weatherall, D. J., Williams, T. N., Allen, S. J., and O'donnell, A. (2010). The population genetics and dynamics of the thalassemias. *Hematol. Oncol. Clin. North Am.* 24, 1021–1031. doi: 10.1016/j.hoc.2010.08.010
- Wienert, B., Martyn, G. E., Funnell, A. P. W., Quinlan, K. G. R., and Crossley, M. (2018). Wake-up Sleepy gene: reactivating fetal globin for beta-hemoglobinopathies. *Trends Genet.* 34, 927–940. doi: 10.1016/j.tig.2018.09.004
- Witt, O., Sand, K., and Pekrun, A. (2000). Butyrate induced erythroid differentiation of human K562 cells involves inhibition of ERK and activation of p38 MAP kinase pathways. *Blood* 95, 2391–2396. doi: 10.1182/blood.v95.7.2391.007k21_2391_2396
- Xu, P., Tong, Y., Liu, X. Z., Wang, T. T., Cheng, L., Wang, B. Y., et al. (2015). Both TALENs and CRISPR/Cas9 directly target the HBB IVS2-654 (C > T) mutation in beta-thalassemia-derived iPSCs. *Sci. Rep.* 5:12065. doi: 10.1038/srep12065
- Yang, J., Kawai, Y., Hanson, R. W., and Arinze, I. J. (2001). Sodium butyrate induces transcription from the $G\alpha i2$ gene promoter through multiple Sp1 sites in the promoter and by activating the MEK-ERK signal transduction pathway. *J. Biol. Chem.* 276:25742. doi: 10.1074/jbc.m102821200
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS J. Integr. Biol.* 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhang, X., Wang, L., Liu, M., and Li, D. (2017). CRISPR/Cas9 system: a powerful technology for in vivo and ex vivo gene therapy. *Sci. China Life Sci.* 60, 468–475. doi: 10.1007/s11427-017-9057-2
- Zhou, D., Liu, K., Sun, C. W., Pawlik, K. M., and Townes, T. M. (2010). KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. *Nat. Genet.* 42, 742–744. doi: 10.1038/ng.637
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10:1523. doi: 10.1038/s41467-019-09234-6

Conflict of Interest: All authors were employed by BGI-Shenzhen which is a non-profit organization. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Li, Zhou, Sun, Ouyang, Dong, Liu, Ge, Zhang, Liu and Gu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.