A long noncoding RNA from the HBS1L-MYB intergenic region on chr6q23 regulates human fetal hemoglobin expression

Tasha A. Morrison, Ibifiri Wilcox, Hong-Yuan Luo, John J. Farrell, Ryo Kurita, Yukio Nakamura, George J. Murphy, Shuaiying Cui, Martin H. Steinberg, David H.K. Chui

A long noncoding RNA from the HBS1L-MYB intergenic region on chr6q23 regulates human fetal hemoglobin expression. An enhancer element within this locus, highlighted by a 3-bp deletion polymorphism (rs66650371), is known to interact with the promoter of the neighboring gene, MYB, to increase its expression, thereby regulating HbF production. RNA polymerase II binding and a 50-bp transcript from this enhancer region reported in ENCODE datasets suggested the presence of a long noncoding RNA (lncRNA). We characterized a novel 1283 bp transcript (HMI-LNCRNA; chr6:135,096,362–135,097,644; hg38) that was transcribed from the enhancer region of MYB. Within erythroid cells, HMI-LNCRNA was almost exclusively present in nucleus, and was much less abundant than the mRNA for MYB. HMI-LNCRNA expression was significantly higher in erythroblasts derived from cultured adult peripheral blood CD34+ cells which expressed more HbF, compared to erythroblasts from cultured cord blood CD34+ cells which expressed much more HbB. Down-regulation of HMI-LNCRNA in HUDEP-2 cells, which expressed mostly HbB, significantly upregulated HBG expression both at the mRNA (200-fold) and protein levels, and promoted erythroid maturation. No change was found in the expression of BCL11A and other key transcription factors known to modulate HBG expression. HMI-LNCRNA plays an important role in regulating HBG expression, and its downregulation can result in a significant increase in HbF. HMI-LNCRNA might be a potential therapeutic target for HbF induction treatment in sickle cell disease and β-thalassemia.

1. Introduction

Sickle cell disease (SCD) and β-thalassemia major are the most prevalent hemoglobinopathies worldwide. Once found primarily in regions of the world where malaria was and may still be endemic, these diseases are now widespread due to human migration and are increasingly important to global health. Fetal hemoglobin (HbF; α2β2) can inhibit polymerization of deoxy-sickle hemoglobin, and can also compensate for the lack of adult hemoglobin (HbA; α2ε2) in β-thalassemia major. Therefore, HbF is the major modifier of disease severity for both SCD and β-thalassemia major [1,2].

Genome-wide association studies (GWAS) have found multiple single nucleotide polymorphisms (SNPs) marking three major quantitative trait loci (QTL) associated with HbF levels — chr11p15 (HBB gene cluster), chr2p16 (BCL11A) and chr6q23 (HBS1L-MYB intergenic polymorphisms or HMIP) [3]. Together they account for 20–45% of HbF variance in different populations. In addition, other cis-acting elements such as the HBD-HBBF1 intergenic region and transcription factors including MYB, KLF1, BCL11A, ZBTB7A, CHD4, NR2C1/NR2C2 and KDM1α, play important roles in regulating HBG expression [4–7]. Nevertheless, significant gaps of knowledge on the regulation of HBG still remain.

The 126 kb HBS1L-MYB intergenic region on chr6q23 is between the genes HBS1L which is a member of the GTP-binding elongation factor family with no known association with erythroid-specific traits, and MYB which encodes for the transcription factor c-MYB. c-MYB regulates proliferation and maturation of erythroid cells, and modulates gene expression within the HBB gene cluster [8,9]. A distal enhancer
located at ~84 kb upstream of MYB has been shown by GWAS, insertional mutagenesis, long-range interaction demonstrable by chromosome conformation capture (3C) analysis, and gene editing with Cas9 nucleases [10-13]. This enhancer encompasses a 3-bp deletion polymorphism (rs66650371), which is surrounded by binding sites for erythroid-specific transcription factors TAL1/E47, GATA, RUNX1, LDB1 and KLF1, and is likely the functional motif to account for most of the effect upon HbF level by this QTL [10,12,13]. Alteration of this enhancer by polymorphisms such as rs66650371 reduced its interaction with the MYB promoter, which led to downregulation of MYB and upregulation of HBG expression. Furthermore, ENCODE datasets annotated RNA polymerase II occupancy and a 50-bp RNA transcript adjacent to rs66650371. This led us to hypothesize that this transcript is part of a long noncoding RNA (IncRNA) [10]. IncRNAs are usually >200 nucleotides long, are transcribed throughout the genome, and have broad functionality.

We now report the characterization of a novel 1283 bp lncRNA, herein named the HBS1L-MYB intergenic long noncoding RNA (HMI-LNCRNA). HMI-LNCRNA is transcribed from the enhancer for MYB, and its downregulation significantly increased HBG expression at both the mRNA and protein levels in human adult-like erythroid cells. These observations suggest that HMI-LNCRNA has an important role in silencing HBG expression in adults, and could become a therapeutic target for increasing HbF in patients with SCD and β-thalassemia major.

2. Materials and methods

2.1. K562 cells

K562 cells were cultured at 37 °C in RPMI medium containing 10% FBS and 2% penicillin/streptomycin.

2.2. RNA extraction

Total RNA was extracted using RNeasy Mini Kit (Qiagen), treated with DNase (RNase-Free DNase Set, Qiagen), followed by RNA cleanup using RNeasy Mini Kit. For tissue-specificity experiment, multiple human organ RNA panels (Invitrogen and Clontech) were also treated with DNase, followed by RNA cleanup.

2.3. Reverse transcription polymerase chain reaction RT-PCR

cDNA was synthesized from total RNA using SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen). PCR reactions were done using the Multiplex PCR kit (Qiagen). The following primers were used to amplify the 1180 bp product: 5′-CTGGACAGCAGATGTT-3′ (forward) and 5′-GGAAACGCCCCTGATAACATT-3′ (reverse).

2.4. Rapid amplification of cDNA ends (RACE)

5′- and 3′-RACE were done using the FirstChoice RLM-RACE Kit (Ambion), following the manufacturer’s instructions, using SuperTag Plus Polymerase (Life Technologies) for PCR reactions. The following gene-specific primers were used: 5′-ATCGCTCATGAGAAATG (forward) and 5′-GGACGCCCCTGATAACATT-3′ (reverse).

2.5. DNA sequencing

PCR products were purified using AccuPrep PCR Purification Kit, and prepared for sequencing using ABI Big Dye Terminator v3.1 Cycle Sequencing Kit. Sequence data was analyzed on FinchTV version 1.5.0. NCBI BLAST was used to determine length and location of sequence.

2.6. Human umbilical cord blood-derived erythroid progenitor (HUDEP) cells

HUDEP cells are immortalized erythroid cell lines derived from cord blood CD34+ mononuclear cells [14]. HUDEP-1 and HUDEP-2 cells were maintained in expansion medium—StemSpan SFEM medium (StemCell Technologies) supplemented with SCF (50 ng/ml, Invitrogen), EPO (3 U/ml, Invitrogen), dexamethasone (1 μg/ml, Sigma), doxycycline (1 μg/ml, Clontech), 1-glutamine (1%, Life Technologies) and penicillin/streptomycin (2%, Life Technologies). For erythroid maturation, cells were cultured in differentiation medium—IMDM medium (Invitrogen) supplemented with heat inactivated human serum from human male AB plasma (5%, Sigma), EPO (3 U/ml, Invitrogen), insulin (10 μg/ml, Sigma), doxycycline (1 μg/ml, Clontech), holo-transferrin (500 μg/ml, Sigma), heparin (3 U/ml, Sigma), SCF (100 ng/ml, Invitrogen), 1-glutamine (1%, Life Technologies) and penicillin/streptomycin (2%, Life Technologies)—for 5 days. For further erythroid maturation, doxycycline was removed and cells were cultured for 2 more days.

2.7. Primary CD34+ mononuclear cells

Primary CD34+ mononuclear cells derived from cord blood and peripheral blood (StemCell Technologies) were expanded for up to 7 days in StemSpan SFEM II medium (StemCell Technologies) supplemented with StemSpan CC100 (1× StemCell Technologies) and penicillin/streptomycin (2%, Life Technologies). To induce erythroid differentiation, cells were cultured for up to 10 days in StemSpan SFEM II medium (StemCell Technologies) supplemented with SCF (10 ng/ml, Invitrogen), EPO (5 U/ml, Invitrogen), IL-6 (10 ng/ml, Sigma) and penicillin/streptomycin (2%, Life Technologies).

2.8. Nuclear and cytoplasmic fractionation

Nuclear/Cytosol Fractionation Kit (BioVision) was used following manufacturer’s instructions.

2.9. Quantitative PCR (qPCR)

RNA was used for qPCR using TaqMan RNA-to-cDNA 1-Step Kit (Applied Biosystems) and the following TaqMan gene expression assays (Applied Biosystems): HBG1/2 (Hs00361313_g1), HBB (Hs00758889_s1), MYB (Hs00920556_m1), HBS1L (Hs00188641_g1), HMI-LNCRNA (custom TaqMan assay designed by Applied Biosystems to target genome position chr6: 135,096,354-135,097,644, hg38; assay ID number AJI1MTQO), BCL11A (Hs01093197_m1), CHD4 (Hs00172349_m1), KLF1 (Hs00610592_m1), ZBTB7A (Hs00252415_s1), NRZ1 (Hs00915957_m1), NRZ2 (Hs00991824_m1), KDM1a (Hs01002741_m1) and ACTB (Hs01060665_g1). qPCR reactions were run on a StepOne Plus qPCR machine (Applied Biosystems). ACTB was used as the endogenous control.

2.10. Western blot analysis

Cell pellets were suspended in Roche lysis buffer (protease inhibitor, 0.3% NP40, 10% glycine, 2 mM EDTA, 246 mM NaCl, 10% phosphatase inhibitor, PBS and water), placed on ice for 1 h and centrifuged at 14,500 rpm for 15 min at 4 °C to extract protein. Standard methodology was used for Western blot analysis. The following antibodies were used: c-MYB (ab109127, Abcam), HBB (16216-1-AP, Proteintech),
heminoglobin γ (sc-21756, Santa Cruz Biotechnology) and GAPDH (sc-47724, Santa Cruz Biotechnology).

2.11. Plasmids

The following and bottom strands were synthetically made and annealed. To make the shRNA template for HMI-LNCRNA: top strand 5'-GATCGGCTAG TATGGAAGCATTAGGCTCTGAGACTAATGGTCT CACATACACTTTTGG-3', bottom strand 5'-AATTCGAAAAGCTAGTAGTGTGAAGCACTTAGCTTCCTGTCAGACTAAGTGCTT-3'. Next, the shRNA template for HMI-LNCRNA was ligated to the lentivirus vector pGreenPuro (System Biosciences), upstream the EFl promoter, and labeled HMI-LNCRNA shRNA. pGreenPuro scrambled shRNA (System Biosciences) was used as the negative control. Both lentivectors co-express green fluorescent protein (GFP).

2.12. Lentivirus transduction

293T cells were transfected with scrambled shRNA and HMI-LNCRNA shRNA using the LentiStarter 2.0 kit (System Biosciences), following manufacturer's instructions, to generate lentiviral particles. HUDEP-2 cells were transduced at an MOI 50 with 5 μg/ml Polybrene, and centrifuged at room temperature for 30 min at 1250 × g. After 48 h, transduction was repeated. 48 h post-transduction, cells were cultured in expansion medium with 1 μg/ml Puromycin for up to 2 weeks.

2.13. Cell surface staining for FACS

Cells were harvested and washed in FACS buffer, and were resuspended in 100 μl of FACS buffer with antibody, and incubated on ice for 30 min. Next, cells were washed twice in FACS buffer by centrifugation at 300g for 5 min. Finally, cells were fixed in IC Fixation buffer (eBioscience) for 30 min before analyzing the cells using the BD FACScan. The following antibodies were used: CD71-PE (334105, BioLegend) and CD235-PerCP/Cy5.5 (306613, BioLegend).

2.14. Immunofluorescence (IF)

Fixed cells were washed in PBS, blocked in 5% BSA (in PBS) and incubated in primary antibody for 1 h. Next, cells were washed, blocked in 5% normal goat serum and incubated in secondary fluorescent-labeled antibody for 30 min. Finally, cells were washed and covered with mounting medium containing DAPI. Images were taken with the Eclipse Ci (Nikon) microscope with a DS-QiMc-03 mono-type (Nikon) camera, using NIS-Elements Br Imaging software (Nikon). The following antibodies were used: Hemoglobin γ (sc-21756, Santa Cruz), Monoclonal Mouse Anti-human β-globin (a gift from IsoLab), Anti-human alpha globin (Wallac Inc.), and Alexa Fluor-594-conjugated AffiniPure F(ab)_2 Fragment Goat Anti-mouse IgG (H + L) (115-586-062, Jackson ImmunoResearch).

3. Results

3.1. Determining the full length of HBS1L-MYB intergenic long noncoding RNA

To determine if the 50-bp transcript reported by ENCODE is part of a putative lncRNA, we used the sequence of this transcript as a guide to develop primers consecutively upstream and downstream of this region. cDNA from K562 cells, an immortalized human erythroid cell line, was used for PCR reactions until amplification failed. The longest PCR product amplified was 1180 bp (chr6:135,096,355–135,097,534; hg38) (Fig. 1A).

To obtain the full-length sequence of this RNA transcript, rapid amplification of cDNA ends (RACE) was employed to ascertain its 5'- and 3'-termini (Fig. 1B; Supplemental Figs. 1 and 2). RNA from K562 cells was used for the 3'-RACE, and RNA extracted from the thymus, which was more abundant for this lncRNA (Fig. 2), was used for the 5'-RACE. Moreover, both 3'- and 5'-RACE was repeated using RNA extracted from HUDEP-2 cells, and yielded identical results (Supplemental Fig. 3).

Fig. 1C illustrates the coordinates of the full-length lncRNA (chr6:135,096,362–135,097,644; hg38), and its nucleotide sequences are shown in Supplementary Table 1. It was 1283 bp in length, transcribed from a genomic site ~84.9 kb upstream from MYB and ~41.5 kb upstream from HBS1L, and herein called the HBS1L-MYB intergenic long noncoding RNA (HMI-LNCRNA). HMI-LNCRNA was transcribed from the enhancer for MYB and encompassed binding sites for erythroid transcription factors TAL1/E47, GATA, RUNX1, and within the transcript was the site of the HbF-associated 3-bp deletion polymorphism (rs66650371). HMI-LNCRNA was intron-less, had no 3'-end polyadenylation signal, and could be presumed to have a 5'-cap as 5'-RACE selected for only RNAs with a 5'-cap. It also contained no open reading frames longer than 300 nucleotides (data not shown); their presumed peptides were not catalogued in protein databases; and therefore they were unlikely to be translated [15].

3.2. Tissue-Specificity of HMI-LNCRNA

LncRNAs often are expressed in either one or a few cell types and are considered more tissue-specific than most mRNA [16]. The expression pattern of HMI-LNCRNA was analyzed by qPCR in 25 different human tissues and cells (Fig. 2). The thymus and testis had by far the highest expression. Erythroblasts derived from cord blood CD34+ mononuclear cells had high expression of HMI-LNCRNA, followed by other hematopoietic tissues and cells, such as the fetal liver, bone marrow, spleen and K562 cells. Non-hematopoietic tissues, except for the testis where lncRNAs are especially common[17], had mostly minimal expression pattern of HMI-LNCRNA and MYB in erythroid cells

Within erythroid cells, HMI-LNCRNA was expressed almost entirely in the nucleus (Supplemental Fig. 5), suggesting that it may function as a gene-regulating IncRNA. HMI-LNCRNA was much less abundant than MYB (Supplemental Fig. 6), which is a common characteristic of many IncRNAs compared to their protein-coding counterparts [18]. MYB is critical to erythropoiesis; high levels of expression favor proliferation of erythroid progenitor cells, while reduced levels promote erythroid maturation [19]. Additionally, downregulation of MYB in erythroid precursors was shown to increase HbF [8,12,13,20]. Therefore, it was important to determine if HMI-LNCRNA and MYB had similar expression patterns in erythroid cells and if this similarity was associated with differences in HBG and HBB expression.

Both HMI-LNCRNA and MYB transcript levels were compared in erythroblasts derived from culture of cord blood CD34+ cells, which expressed more HBG, to erythroblasts derived from culture of adult peripheral blood CD34+ cells, which expressed more HBB (Fig. 3A). Both HMI-LNCRNA and MYB were significantly higher in erythroblasts derived from adult peripheral blood CD34+ cells than in erythroblasts derived from cord blood CD34+ cells (Fig. 3A). With erythroid differentiation of erythroblasts derived from peripheral blood CD34+ cells, there was a decline in HMI-LNCRNA and in MYB expression (Fig. 3B; Supplemental Fig. 7). Next, the same comparison was made in HUDEP-1 cells, which mostly express HBG, and HUDEP-2 cells, which exclusively express HBB (Fig. 3C). HUDEP-2 cells had significantly higher expression of HMI-LNCRNA than HUDEP-1 cells; and there was no
Fig. 1. Determining the full length of the transcript found within the HBS1L-MYB intergenic region. (A) PCR was done to amplify a 1180 bp region within the HBS1L-MYB intergenic region using cDNA from K652 cells. Samples were run on agarose gels. Lane 1: 100-bp DNA ladder; lane 2: negative control, PCR mix without DNA input; lane 3: positive control, PCR mix with genomic DNA; lane 4: PCR mix with cDNA generated without reverse transcriptase; and Lane 5: PCR mix with cDNA generated with reverse transcriptase. (B) Agarose gels show PCR products from the 5′-RACE reactions (using RNA from thymus) and 3′-RACE reactions (using RNA from K562 cells). For 5′-RACE, lane 1: 100-bp DNA ladder; lane 2: PCR mix with cDNA generated without Tobacco Acid Pyrophosphatase (TAP) treatment; and lane 3: PCR mix with cDNA generated with TAP treatment. For 3′-RACE, lane 1: 100-bp DNA ladder; lane 2: PCR mix with cDNA generated without reverse transcriptase; and lane 3: PCR mix with cDNA generated with reverse transcriptase. (C) Illustration of genomic region between chr6: 134,950,000–135,250,000 (hg38), showing approximate locations of the 3-bp deletion polymorphism rs66650371 and rs9399137 a SNP in LD with rs66650371, and HBS1L and MYB (arrows represent transcription direction and approximate length of genes). Based on DNA sequencing of RACE products, the 5′- and 3′-ends of the transcript were revealed to determine the full length of the transcript, which is 1283 bp in length and named the HBS1L-MYB Intergenic Long Noncoding RNA (HMI-LNCRNA). Located in the genomic sequence for HMI-LNCRNA are binding sites for erythroid-specific transcription factors TAL1/E47, GATA and RUNX1, and rs66650371. HMI-LNCRNA does not include rs9399137.

Fig. 2. Expression pattern of HMI-LNCRNA among various human cells and tissue. Relative quantity of HMI-LNCRNA was measured by qPCR in 25 different samples. CD34+ cord blood day 11 is erythroblasts harvested at Day 11 of two-phase expansion and differentiation culture of CD34+ mononuclear cells derived from cord blood (see Supplemental Fig. 6). Means are shown for samples with two or more independent samples. ACTB was used as the endogenous control.
difference in MYB expression (Fig. 3C). During erythroid maturation of HUDEP-2 cells, both HMI-LNCRNA and MYB expression decreased similar to erythroblasts derived from peripheral blood CD34+ cells (Fig. 3D). These data show that HMI-LNCRNA expression is higher in erythroid cells which express more HBB, when compared to erythroblast cells which express more HBG. With erythroblast cell differentiation, the expression of both HMI-LNCRNA and MYB decreases.

3.4. HMI-LNCRNA regulates hemoglobin expression in HUDEP-2 cells

To examine the possible role of HMI-LNCRNA in regulating MYB and hemoglobin expression, a lentiviral vector co-expressing GFP and a shRNA template targeting HMI-LNCRNA (HMI-InRNA shRNA) was employed to downregulate the HMI-LNCRNA in HUDEP-2 cells. Fig. 4A displays the timeline of the shRNA transduction experiments. After transduction, cells were cultured in expansion medium containing puromycin to select for lentivirus-positive cells. Transduction of HUDEP-2 cells with HMI-InRNA shRNA led to a 50% knockdown of HMI-LNCRNA when compared to naïve (non-transduced) cells and cells expressing scrambled shRNA (Fig. 4B). Reduction in HMI-LNCRNA led to a 200-fold increase in HBG mRNA, a 30% reduction in MYB mRNA, and no change in HBB or HBS1L mRNA when compared to naïve cells and cells transduced with scrambled shRNA (Fig. 4C). Percent HBG increased from < 1% in naïve and scrambled shRNA cells to > 20% of total HBB and HBG expression in cells with HMI-InRNA shRNA (Fig. 4D). Furthermore, changes in HBG and MYB mRNA levels were associated with corresponding changes in HBG and c-MYB protein levels (Fig. 4E) shown by Western blot analyses. The up-regulation of HBG production was also confirmed by immunofluorescence staining using anti-HBG antibody (Fig. 4F). This immunocytological technique showed that high HBG expression was found in approximately 10% of more mature HUDEP-2 cells with knockdown of HMI-LNCRNA. There was no change in HBB or HBA protein expression (Supplemental Figs. 8 and 9). Down-regulation of MYB expression in HUDEP-2 cells transduced with HMI-InRNA shRNA was inconsistent (Supplemental Fig. 10).

The expression of key hemoglobin-regulating transcription factor genes (KLF1, BCL11A, ZBTB7A, CHD4, NRZ2C1, NRZ2C2 and KDM1a), which are all known to downregulate HBG, were either unchanged or slightly upregulated in HUDEP-2 cells transduced with HMI-InRNA shRNA compared to controls (Fig. 5).

3.5. Hemoglobin expression during erythroid differentiation of HUDEP-2 cells with knockdown of HMI-LNCRNA

In addition to analyzing the effect of downregulating HMI-LNCRNA in non-differentiated HUDEP-2 cells, after the two-week selection for transduced cells with puromycin treatment, the cells underwent erythroid maturation for up to seven days as shown in Fig. 6A. HBG expression was significantly higher on days 0, 5 and 7 in HUDEP-2 cells with HMI-InRNA shRNA (Fig. 6B), and HBG expression remained at about 20% of total HBG and HBB transcripts at each time point (Fig. 6C). The increased HBG expression was also corroborated at the protein level (Supplementary Figs. 10 and 11).

3.6. Role of HMI-LNCRNA in modulating erythroid maturation in HUDEP-2 cells

HUDEP-2 cells resemble basophilic erythroblasts expressing both transferrin receptor (CD71), a cell surface marker for immature
Fig. 4. The effects of knocking down HMI-LNCRNA on hemoglobin, MYB and HBS1L in HUDEP-2 cells. (A) Illustration of timeline for transduction and culture of HUDEP-2 cells. Cells were maintained in expansion medium. (B) Expression level of HMI-LNCRNA was determined by qPCR in HUDEP-2 cells that were not transduced (naïve), and transduced with either scrambled shRNA or HMI-lncRNA shRNA lentiviruses by qPCR. (C) HBG, HBB, MYB and HBS1L transcript levels, and (D) percent HBG and HBB out of the total of both transcripts were measured in these same samples. (E) Protein expression of c-MYB, HBG and HBB were analyzed by Western blot in naïve HUDEP-2 cells, and cells transduced with scrambled shRNA and HMI-lncRNA shRNA (HUDEP-1 cells were used as control). GAPDH was used as loading control. (F) Naïve HUDEP-2 cells, and cells transduced with scrambled and HMI-lncRNA shRNAs were stained with anti-HBG antibody, followed by secondary antibody labeled with Alexa Fluor-594 (in red) and DAPI to stain nuclei (in blue), and imaged with a fluorescent microscope at 40 × magnification. For all qPCR analyses, ACTB was used as the endogenous control. p-Values: * < 0.05; ** < 0.005; *** < 0.0005; NS (not significant). p-Values were obtained by Student t-test. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 5. The effects of knocking down HMI-LNCRNA on erythroid-regulating transcription factors in HUDEP-2 cells. Relative expression of KLF1, BCL11A, ZBTB7A, CHD4, NR2C1, NR2C2 and KDM1a transcripts were analyzed by qPCR in naïve HUDEP-2 cells, and cells infected with scrambled shRNA and HMI-lncRNA shRNA. Beta actin was used as the endogenous control. p-Values: * < 0.05; ** < 0.005; *** < 0.0005; NS (not significant). p-Values were obtained by Student t-test.
erythroblasts, and glycophorin A (CD235), a cell surface marker for more mature erythroblasts [6]. The percentage of cells positive for these markers were examined by flow cytometry in naïve cells, cells expressing scrambled shRNA and cells expressing HMI-lncRNA shRNA. Within both control groups, approximately 55–65% of the cells were positive for both CD71 and CD235 (Fig. 7). In contrast, about 85% of cells transduced with HMI-lncRNA shRNA were positive for both markers. These observations suggest that downregulation of HMI-LNCRNA promotes maturation of HUDEP-2 cells.

4. Discussion

A novel 1283 bp lncRNA in erythroid cells is transcribed from the distal enhancer of MYB on chr6q23 highlighted by the 3-bp deletion polymorphism (rs66650371). Knockdown of HMI-LNCRNA to half of its normal level increased HBG expression by 200-fold and this is reflected by a marked increase in HBG production.

Long non-coding RNAs are important regulators of multiple cellular processes, biomarkers for disease prognosis and possible therapeutic targets [21–24]. Many tissue specific lncRNAs are expressed during hematopoiesis and erythropoiesis [25–29]. To date, long non-coding RNAs that can modulate globin gene expression have not been described.

HUDEP-2 cells, which do not have the 3-bp deletion at rs66650371 (data not shown) and synthesize almost exclusively HBB have been used to study the regulation of HBG expression [6,13,14]. It will be necessary to replicate the present experiments in cultured erythroblasts derived from primary CD34+ progenitor cells from healthy individuals and SCD patients. To determine if upregulation of HMI-LNCRNA could reduce HBG expression, we overexpressed this lncRNA in HUDEP-1 cells that express primarily HBB but saw no change (Supplemental Fig. 12).
The expression of HBG in HUDEP-2 cells was heterogeneous. With erythroid maturation, the proportion of these HUDEP-2 cells that expressed HBG increased from approximately 10% to 25% (Fig. 4F; Supplementary Fig. 11). It remains to be determined if additional down-regulation of HMI-LNCRNA or allowing these HUDEP-2 cells to become fully mature erythroblasts might lead to more cells expressing HBG. It is possible that transduction of HUDEP-2 cells with lentiviral particles was not homogeneous, thus accounting for the heterocellular expression of HBG. The genomic sites for the lentiviral particle integration might also be another factor that could modulate HBG expression [30]. Furthermore, HBG expression in cultured erythroid cells derived from single adult erythroid progenitors is stochastic [31].

The mechanism whereby HMI-LNCRNA augments HBG expression requires further study. LncRNAs may regulate gene expression by providing the necessary scaffold which transcription factor and other protein complexes can be associated with and be guided to cognate promoters [21,22]. They may recruit chromatin modification enzymes for epigenetic remodeling processes [22,32]. They may also act by promoting intra-chromosomal looping to foster interaction between enhancers and promoters [33,34]. Since HMI-LNCRNA is transcribed from the distal enhancer for MYB, it is highly probable that HMI-LNCRNA is an enhancer RNA (eRNA), and functions by facilitating the interaction between the distal enhancer and the MYB promoter. We hypothesized that downregulation of HMI-LNCRNA would reduce interaction of the distal enhancer with the promoter of MYB and reduce its transcription, which would ultimately lead to increase in HBG expression. However, decreased expression of MYB was not consistently observed in HUDEP-2 cells transduced with HMI-LnRNA shRNA to downregulate HMI-LNCRNA to 50% of normal level (Fig. 4E, Supplementary Fig. 10). Conceivably, additional down-regulation of HMI-LNCRNA may result in decreased consistent MYB expression. Whether or not interaction between distal enhancer and promoter of MYB is attenuated with HMI-LNCRNA down-regulation should be informative. Alternatively, HMI-LNCRNA might function through a yet undescribed molecular mechanism to modulate HBG expression. An important future study is to determine genomic or protein binding sites for HMI-LNCRNA [35].

Increasing HbF levels in SCD and β-thalassemia major is a proven approach to treatment. Hydroxyurea, the first FDA-approved drug for SCD, can induce high HbF level. Unfortunately, not all treated patients responded to this agent with sufficient increase in HbF to effect substantial clinical improvement. There has been remarkable progress on gene therapy by different approaches [36–38], and many different agents are also being investigated for their efficacy to induce HbF expression [39–44]. Down-regulating HMI-LNCRNA can possibly be an additional therapeutic target for HbF induction. If the increment in HbF increase found in vitro can be recapitulated in vivo in SCD and β-thalassemia major, it should be clinically beneficial [45].

Acknowledgements

We thank Dr. Richard Myers and Dr. Wenqian Hu for the helpful discussions, comments and guidance, especially at the earlier stages of this project. This work was supported by the National Institutes of Health, National Heart, Lung, and Blood Institute (U01HL07443-01 and 1F31HL126462, and 5T32HL007501-31).

Authorship

T.A.M. performed all experiments; I.W. carried out initial experiments; H-Y.L. did the immunofluorescence experiments and provided expertise in tissue cultures; R.K. and Y.N. provided HUDEP-2 and HUDEP-1 cells; J.J.F., M.H.S. and D.H.K.C. conceived of the idea, and together with T.A.M., H-Y.L., G.J.M., and S.C. designed experiments. T.A.M., M.H.S. and D.H.K.C. wrote the manuscript, and all authors have read and approved of the manuscript.

All authors have no conflict of interest disclosure, nor financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bcmd.2017.11.003.

References
