Title
Flipping the Hemoglobin Switch and Discovering Regulators Involved in Fetal Hemoglobin Reactivation

Author
Boontanrart, Mandy

Publication Date
2020

Peer reviewed|Thesis/dissertation

Flipping the Hemoglobin Switch and Discovering Regulators Involved in Fetal Hemoglobin Reactivation

By

Mandy Y Boontanrart

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jacob Corn, Chair
Professor Ellen Robey
Professor Nicholas Ingolia
Professor John Dueber

Summer 2020

Abstract

Flipping the Hemoglobin Switch and Discovering Regulators
Involved in Fetal Hemoglobin Reactivation

By

Mandy Y Boontanrart

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Jacob Corn, Chair

The fetal to adult hemoglobin switch is a developmental process by which fetal hemoglobin becomes silenced after birth and replaced by adult hemoglobin. Diseases caused by defective or missing adult hemoglobin, such as Sickle Cell Disease or β-Thalassemia, can be ameliorated by reactivating fetal hemoglobin. We discovered that knockdown or knockout of β-globin, a subunit of adult hemoglobin, led to robust upregulation of γ-globin, a subunit of fetal hemoglobin. This phenomenon suggested that red blood cells have an inherent ability to upregulate fetal hemoglobin in the event that adult hemoglobin is lacking.

We developed multiple gene-editing tools in an immortalized erythroid cell model to investigate the molecular mechanisms behind the increase in fetal hemoglobin. Time-course transcriptomics identified ATF4, a transcription factor, as a causal regulator of this response. Further analysis also converged upon downregulation of MYB and BCL11A, known repressors of γ-globin, described in detail in chapter 2. Further work in chapter 3 explores other possible fetal hemoglobin regulators as discovered by CRISPRi arrayed mediated knockdown experiments. This work furthers our understanding of fundamental mechanisms of gene regulation and how cellular and molecular events influence red blood cell differentiation.

Acknowledgements

I would like thank Jacob Corn for his guidance and mentorship during my PhD. Thank you for giving me the freedom to drive my scientific endeavors and also for extending me the opportunity to finish my PhD in Switzerland – working in your lab has allowed me to grow both professionally and personally.

My Corn Lab co-workers and friends made graduate school a truly memorable and positive experience. Alan Wang, Emily Lingeman, Ron Baik, Jenny Shin, Amos Liang, Charles Yeh, Marija Banovic, and Erman Kerasu, thank you for your friendship. Our support system came in the form of shared beers and coffee breaks and are some of my fondest memories from my PhD. Benjamin Gowen, Mark DeWitt, Beeke Wienert, and Markus Schroder, thank you for your guidance and collaborations. Thank you to Rachel Lew and Gautier Stehli, for allowing me the opportunity to mentor such brilliant young scientists.

I would like to acknowledge my friends, Deepti Rokkam and George Huang, who played a pivotal role in my deciding to pursue a PhD despite my initial hesitations.

A big thank you to my family - my brother, Michael Boontanrart, for always supporting me, and my parents, Suradeh Boontanrart and Yowvanij Boontanrart, for encouraging me to achieve the most I can. I would especially like to thank my father for being my biggest cheerleader since day one.

Most of all, I would like to thank my husband, Po Bunyapamai, for sticking this out with me. My PhD experience came with plenty of joy and freedom, but there were also many late nights and a few breakdowns in between. Midway through this journey, we also moved continents and I also got scooped. It was a truly complete PhD experience. Thank you, Po, I would not have been able to complete this journey without you by my side for every step of the way.

Table of Contents

List of Diagrams

List of Figures

List of Supplemental Figures

List of Tables

List of Supplemental Tables

1 Introduction

1.1 The Human Hemoglobin Genes and Developmental Switches

Red blood cells (RBCs), also known as erythrocytes, are packed with hemoglobin tetramers and circulate throughout the body to supply all tissues with oxygen. Hemoglobin was one of the first proteins discovered in 1840 (Hünefeld., 1840). Hemoglobin switching has been of great interest to the fields of developmental biology and genomics as the switching involves both stage and tissue specific processes. In humans, there are two developmental switches that occur. During early embryogenesis, red blood cells are produced in the yolk sac and express embryonic hemoglobin. The first hemoglobin switch occurs after the first trimester when erythropoiesis switches to the fetal liver and begins expressing fetal hemoglobin. The second developmental switch occurs after shortly after birth. Erythropoiesis relocates to the bone marrow and fetal hemoglobin becomes silenced and replaced by adult hemoglobin (Dame and Juul, 2000).

The globin genes are located on two different loci. The β-like globin genes are located on chromosome 11 and includes five genes: *HBE, HBG1, HBG2, HBB, and HBD*, coding for ε-globin, γ-globin 1 and 2, β-globin, and δ-globin, respectively. The β-like globin locus is also controlled by the locus control region (LCR), a super enhancer region characterized by 5 DNase hypersensitivity regions (Deng et al., 2014). The α-like globin genes are located on chromosome 16 and also includes five genes: HBZ, HBM, HBA1, HBA2, and HBQ1, coding for ζ-globin, μ-globin, α-globin 1 and 2, and τ-globin, respectively (**Diagram 1**). Interestingly, HBM and HBQ1 are still not very well-characterized – studies show that by qPCR of μ-globin and τ-globin transcripts are expressed, however, so far without detectable protein levels (Albitar et al., 1989; Goh et al., 2005).

All human hemoglobins are tetrameric proteins comprised of two copies of an α-like globins complexes with two copies of a β-like globin. There are multiple variants of embryonic, fetal, and adult hemoglobin. The adult hemoglobins include HbA, comprised of two copies of α-globin and two copies of β-globin (α2β2), and HbA_2, comprised of two copies of α-globin and two copies of δ-globin (α2δ2). The majority of the hemoglobin in adults is the HbA variant as HBD has a weak promoter and only low amounts of δ-globin are made. The one variant of fetal hemoglobin is HbF (α2γ2). The embryonic hemoglobin variants include Hb Portland-2 (ζ2β2), Hb Portland-1 (ζ2γ2), Hb Gower-1 (ζ2ε2), and Hb Gower-2 (α2ε2)(Manning et al., 2007) (**Diagram 1**).

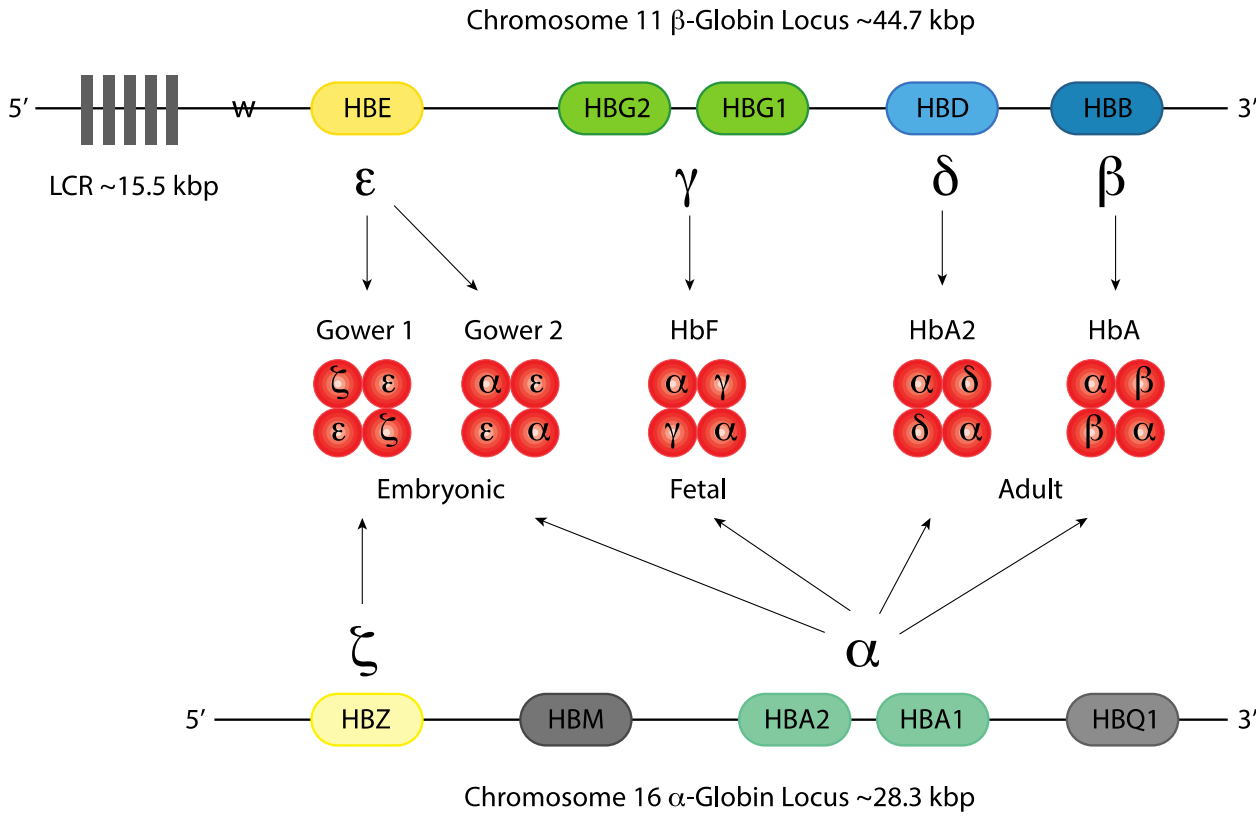


Diagram 1. The Human Hemoglobin Loci and Hemoglobin variants
The β-globin locus is located on chromosome 11 and consists of 5 β-like globin genes. These genes are regulated by an upstream locus control region (LCR). The α-globin locus is located on chromosome 6 and consists of 5 α-like globin genes. The β-like globins complex with α-like globins to form the hemoglobin variants.

1.2 The Molecular Factors of the Fetal to Adult Hemoglobin Switch

The β-globin locus is in part controlled by the locus control region (LCR). The LCR is located approximately 50kbp upstream of the β-like globin genes and is characterized by 5 DNase hypersensitivity regions. Chromosome conformational capture techniques have shown that the LCR makes direct contact with the downstream β-like globin genes (Tolhuis et al., 2002). Additionally, it has been shown that in transgenic mice, integrating LCR elements led to high levels of β-globin gene expression and that deletion of the LCR region led to globin gene repression (Wilber et al., 2011). Taken together, the numerous studies of the LCR and its effect on the globin gene expression has led to the current model of hemoglobin switching, which is a combination of chromosomal looping and transcriptional gene silencing. The LCR works upon the β-globin locus with the orchestration of multiple trans-acting factors. These factors include BCL11A, GATA1, SOX6, LRF, and KLF1 (Wilber et al., 2011).

Genetic association studies first identified that variants in *BCL11A,* a zinc-finger transcription factor, were associated to HbF expression. Molecular studies then determined that expression of the full length BCL11A protein was restrictred to adult

erythroid cells, consistent with BCL11A being a repressor of γ-globin. Indeed, affinity purification methods identified that BCL11A associates with components of the nucleosome remodelling and histone deacetylase (NuRD) complex. Chromosomal conformation experiments found that BCL11A interacted at the LCR at hypersensitivity site 3 and an intergenic region between *HBG* and *HBD* (Sankaran et al., 2008). Later studies showed with chromatin immunoprecipitation was able to pinpoint with nucleotide resolution that BCL11A directly bound to motifs in both *HBG1/2* promoters to repress transcription (Liu et al., 2018; Martyn et al., 2018).

Interestingly, another study also found that BCL11A could additionally be acting upon HBE (Xu et al., 2010) as well as HBZ in the α-globin locus (Liu et al., 2018). This results suggests that BCL11A could be a repressor for embryonic and fetal globin genes, and not specifically *HBG*. The current model is that the LCR makes contact with the β-like globin genes closest to it. During embryogenesis, the LCR is in contact with the fetal globin genes. After birth, the fetal to adult hemoglobin switch is initiated as BCL11A becomes expressed and binds to the γ-globin promoter, blocking the interaction with the LCR. The LCR then makes contact and promotes the transcription of the adult *HBB* and HBD genes.

SOX6, an HMG box transcription factor, accumulates during erythroblast maturation and is capable of downregulating its own transcription (Cantu' et al., 2011). Sox6 was found to be co-expressed and interact directly with BCL11A. A study showing knockdown of both factors in erythroblasts led to a synergistic increase in HbF. SOX6 has been found to bind to both *HBG1/2* promoters, but requires expression of BCL11A to do so. Knockout of SOX6 alone led to just a modest increase in HbF, but had a much stronger effect when BCL11A was knocked out as well. A chromosome conformation capture assay showed that BCL11A modulates the LCR interaction to the β-globin cluster via chromosomal looping. The same study also showed that BCL11A and SOX6 co-occupy the same regions along with GATA-1 (Xu et al., 2010). Taken together, the protein interactions between SOX6 and BCL11A facilitate strong repression of the γ-globin genes mainly.

GATA-1 and GATA-2 are essential transcription factors involed in multiple facets of hematopoiesis that ultimately lead to functional erythrocytes. GATA-2 is important for early hematopoiesis from CD34+ stem cells to differentiate and commit to an erythroid lineage. During erythopoiesis, there is a switch from GATA-2 to GATA-1. GATA-1 is a master transcription factor that is essential for the progression of erythroid-lineage differentiation. GATA-1 is part of several protein complexes that can either activate or repress gene expression (Wilber et al., 2011). At the β-globin locus, it has been shown that the ChIP profile is similar to that of SOX6, with binding sites within the LCR, and at the γ-globin genes. Analysis of BCL11A, SOX6, and GATA1 revealed a significant number of shared sites (Xu et al., 2010). Taken together, the data supports a model in which GATA1, SOX6, and BCL11A cooperate as a complex in repressing the fetal globin genes (Cavazzana et al., 2017).

Erythroid Krupple-like factor, KLF1, is a zinc finger protein that is essential for regulation of the β-globin locus (Wilber et al., 2011). KLF1 was first discovered as a globin gene regulator through studies of β-Thalasemmia, in which a mutation at a KLF1 binding site in *HBB*, led to the disease in adults (Sankaran et al., 2010). A genome-wide study found that haploinsufficiency of KLF1 harboring a mutation that abolished the

DNA binding domain led to an high persistence of fetal hemoglobin (HPFH) phenotype (Wilber et al., 2011). Most recently, is has been found that KLF1 bound and activated the BCL11A promoter in erythoblast cells as well as bound to the β-globin promoter (Zhou et al., 2010), proving to be an important transcription factor of adult erythropoiesis.

LRF/ZBTB7A, a ZBTB transcription factor, is one of the most recently discovered repressors of γ-globin expression. It was found that LRF binds to the γ-globin genes and associates with the NuRD repressive complex to inhibit transcription. The effect of LRF is independent of BCL11A though they both repress through the NuRD complex (Masuda et al., 2015). Similarly to BCL11A, LRF is also directly regulated by KLF1 (Norton et al., 2017).

So far, the regulators that have been discussed have included cis-elements, such as the LCR, and trans-elements, such as transcription factors. Another hemoglobin regulatory element that is farther removed is the HBS1L-MYB intergenic region, which is located on chromosome 6 (Antoniani et al., 2017). Prior to the discovery of BCL11A and other more recent discoveries of the β-globin cluster regulators, a genome-wide study found that mutations within the HBS1L-MYB intergenic region was highly correlated with high levels of fetal hemoglobin (Thein et al., 2007). In a study comparing individuals with normal and elevated levels of HbF, it was found that both MYB and HBS1L were down-regulated in persons with higher HbF levels. Notably, over-expressing MYB in K562 cells shows inhibited γ-globin expression, while over expression of HBS1L had no effect (Jiang et al., 2006). It was later shown that MYB supports erythropoiesis through the transactivation of KLF1 (Bianchi et al., 2010), which could account for the effect of γ-globin. A more recent study in 2019 further identified that a long noncoding RNA (lncRNA) from the HBS1L-MYB intergenic region regulates fetal hemoglobin expression (Morrison et al., 2018). The lncRNA was characterized as a 1283bp transcript (*HMI-LNCHRNA*), was highly expressed in erythroid lineages and was shown that down-regulation of this in an erythroid cell line significantly upregulated γ-globin (Morrison et al., 2018). Taken together, the HBS1L-MYB intergenic region regulates the β-globin locus through multiple levels.

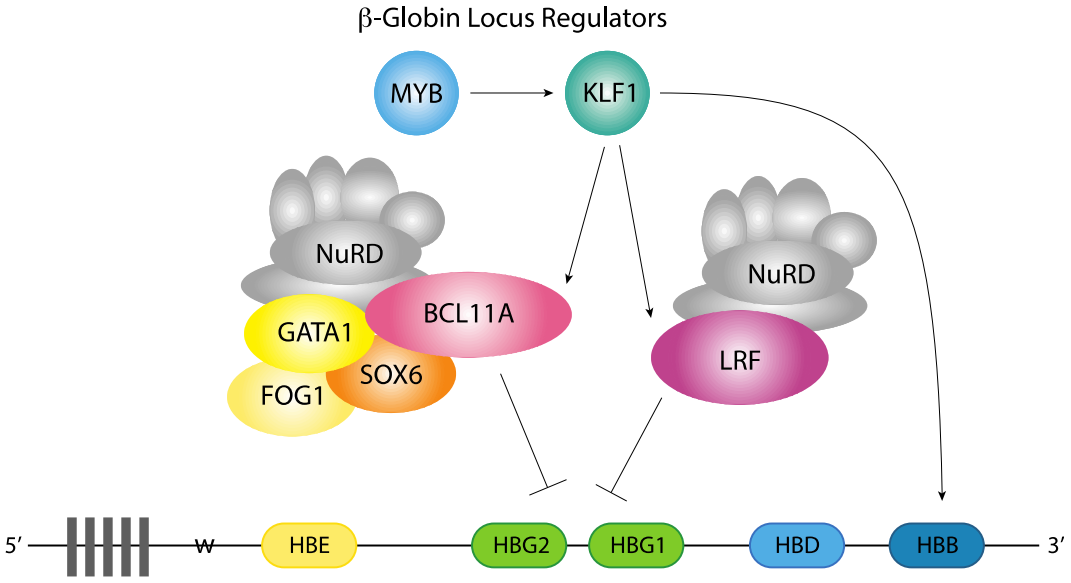


Diagram 2**.** The β-Like Globin Gene Regulators.
BCL11A and LRF are direct repressors of *HBG1/2*, binding promoter regions to block transcription through interaction with the NuRD complex as well as other transcription factors. BCL11A and LRF are both activated by KLF1, which also activates *HBB*. KLF1 is activated by MYB.

1.3 Heme-HRI-eIF2a-ATF4 Pathway.

Essential components of functional hemoglobin are heme and iron. Most of the iron in the human body is sequestered by the prosthetic group heme. Each subunit of hemoglobin binds to a heme, totaling 4 per hemoglobin complex. The balance between hemoglobin and heme is tightly controlled during erythropoiesis. Heme-regulated eIF2a kinase, also known as heme-regulated inhibitor (HRI), is a heme-binding protein that senses levels of uncomplexed intracellular heme, and modulates hemoglobin production through eIF2a phosphorylation signalling. Under normal conditions of sufficient heme, HRI binds to the free heme and is inactive, and so it does not phosphorylate eIF2a, and global translation continues. However, under conditions of iron deficiency, HRI will phosphorylate eIF2a and trigger a decrease in translation (Chen and Zhang, 2019). The purpose of this regulation is to maintain the cellular balance between hemoglobin and intracellular heme. Under conditions of sufficient intracellular heme, globins will be translated to complex with the free heme. Under conditions of insufficient heme, globin translation will decrease.

As mentioned, phosphorylation of eIF2a leads to decreased global translational within the cell. However, eIF2a phosphorylation also increases translation of ATF4, a master transcription factor, to induce stress response genes to mitigate cellular stress (Chen, 2007). A recent study found that loss of HRI led to increased HbF through an uncharacterized mechanism of BCL11A reduction (Grevet et al., 2018). While this was a novel discovery, it did conflict with the model of hemoglobin regulation through eIF2a at the time. Previous studies reported that HRI increased translation of γ-globin mRNA (Hahn and Lowrey, 2013, 2014a). However, this new study reported that loss of HRI led to increased transcription of *HBG* through loss of BCL11A by an uncharacterized mechanism. Taken together, HRI and the eIF2a pathway can have opposing effects on regulating γ-globin at the transcriptional and translational level.

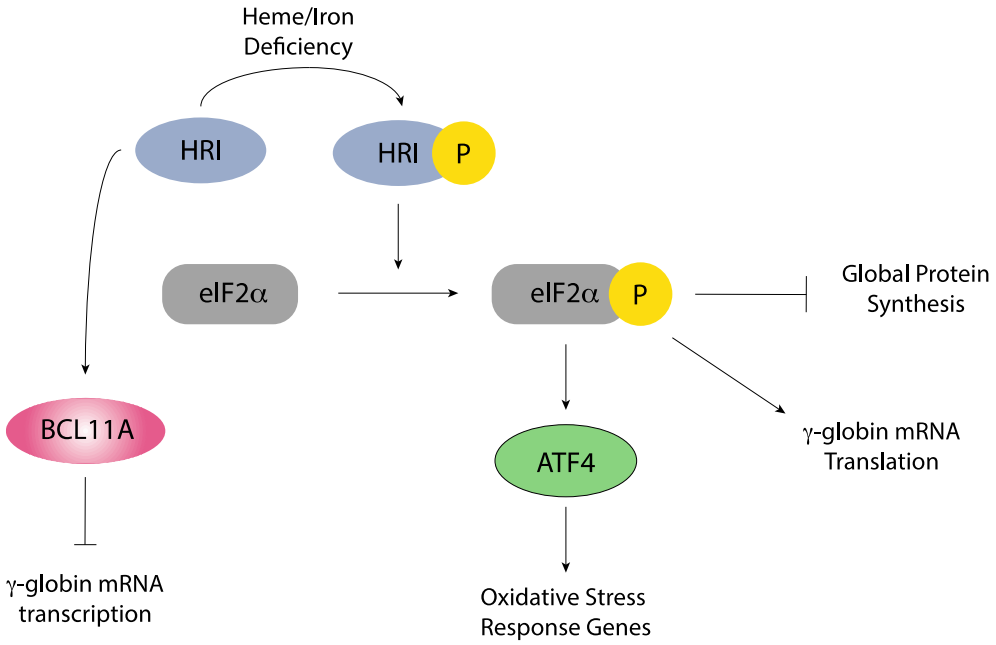


Diagram 3. The Heme-HRI-eIF2aP-ATF4 Pathway.
HRI senses levels of heme to phosphorylate eIF2a. HRI has been shown to indirectly affect levels of BCL11A. Phorsphorylation of eIF2a leads increased translation of ATF4 and the promotion of ATF4-mediated stress response.

1.4 Hemoglobinopathies

The hemoglobinopathies are some of the most common and longest studied genetic diseases worldwide. The hemoglobinopathies, such as Sickle Cell Disease (SCD) and Thalassemia, are genetic diseases caused by mutations to the globin genes of adult hemoglobin. Interestingly, hemoglobinopathy mutations are most commonly found in regions with high prominence of mosquito-borne malaria, *Plasmodium falciparum*. The mild versions of hemoglobinopathies do not manifest in debilitating clinical symptoms. Instead, studies have shown that some hemoglobinopathy mutations lead to less *P.falciparum* infection through the weakening of the cytoadherence interactions between RBCs that the parasite relies on (Taylor et al., 2013). Ultimately, heterozygous hemoglobinopathy mutations have been positively selected for in these populations.

SCD was first observed in 1910, when James Herrick observed sickled erythrocytes under a microscope (Wienert et al., 2018). It was not until 1977 that it was discovered that SCD is caused by an A→G basepair mutation in the first exon of *HBB*, resulting in a glutamic acid to valine (E6V) conversion. This pathological mutation causes β-globin to polymerize and leads to sickling of the erythrocyte. The sickled cells are rigid and cause vasoclusion of blood vessels, membrane damage, and hemolytic anemia. In severe cases this manifests as growth defects, organ damage and dramatically shortened lifespan (de Dreuzy et al., 2016).

Thalassemia is a disease caused by mutations to either the α-globin or β-globin gene and leads to an a/b chain imbalance. In severe cases, this imbalance of globin

chains leads to erythroid lineage damage and premature cell destruction (de Dreuzy et al., 2016). During hemoglobinopathies, the body compensates for the lack of functional RBCs by over-producing erythroid cells in a process called stress erythropoiesis. A chronic state of stress erythropoiesis results in ineffective erythropoiesis – having an excess pool of premature, erythroid progenitors with a/b chain imbalance, chronic hemolysis and the expansion of the bone marrow and leading to permanent bone deformities (Oikonomidou and Rivella, 2018).

It has been shown that a subset of β-hemoglobinopathy patients have high persistence of fetal hemoglobin into adulthood and that this expression ameliorates disease symptoms (Akinsheye et al., 2011). For β-hemoglobinopathy patients, RBCs that have fetal hemoglobin expression were longer-lived, had less hemolysis events, and had lowered chances of vasoocclusion of blood cells. In fact, patients that were compound heterozygote for SCD and HPFH were often asymptomatic with almost normal levels of hemoglobin (Akinsheye et al., 2011). Because of these studies, reactivation of fetal hemoglobin is a promising therapeutic strategy to treat β-hemoglobinopathies.

1.5 Current Methods of Reactivation of Fetal Hemoglobin

Since the discovery that fetal hemoglobin can ameliorate β-hemoglobinopathy diseases, there have been numerous therapeutic efforts aiming to reactivate fetal hemoglobin. A number of small molecules have been discovered that induce fetal globin, the most common, and FDA-approved, being hydroxyurea (Platt, 2008). While small molecule therapies have proven effective for some patients, they have had side-effects and therefore, better therapies to treat β-hemoglobinopathies, is an on-going effort.

A recently discovered gene-editing system that recognizes DNA sequences has given renewed hope to all genetic diseases. Clustered, Regularly Interspaced, Short Palindromic Repeat (CRISPR)-CRISPR-associated (Cas) genome editing has become a widespread tool in molecular biology. The system was first discovered in prokaryotes as a way for bacteria to combat viruses and foreign plasmids (Bhaya et al., 2011; Terns and Terns, 2011). *Streptococcus pyogenes* Cas9, one of the Cas family members, has been the most well-characterized and adapted for molecular biology use. Cas9 is two domains (HNH and RuvC-like domains) that enable it to make a double-stranded DNA cleavage (Jinek et al., 2012). The Cas9 cleavage is specific, requiring a target-specific CRISPR RNA (crRNA), which contains a 20-nucleotide protospacer that complexes to the targeted DNA region, and a trans-activating crRNA (tracrRNA), which functions to keep Cas9 catalytically active. For research purposes, the crRNA and tracrRNA are often combined to a single guide RNA (gRNA) for ease of use (Jinek et al., 2012).

The gRNA is interchangeable and can target any DNA sequence with the limitation that is adjascent to a 5' PAM sequence 'NGG.' Wildtype Cas9 only recognizes the NGG sequence, although now there have been engineered Cas9 variants and other Cas proteins that have different PAM requirements, ultimately expanding the range of DNA sequences that can be targeted for cleavage (Hu et al., 2018). The DNA cleavage is repaired by the cells own DNA-repair pathways. The cell either corrects the double-strand break through non-homologous-end-joining (NHEJ) or homologous DNA repair

(HDR). NHEJ results in insertions or deletions around the cut site, while HDR utilizes a homologous DNA region as a template for repair (Yeh et al., 2019). The DNA repair pathway favors NHEJ and can occur at any time in the cell cycle, while HDR is limited to G2 and S phase (Mohrin et al., 2010). Since rates of HDR can be low, there are on-going efforts to increase the rate HDR since there is a precedent to correct the DNA break to a desirable sequence (Liu et al., 2019). For cell lines, the CRISPR-Cas9 system can be introduced by plasmid or lenti-viral transfection. For primary cells, transfection is difficult and can lead to undesirable integrations. For these cells, methods of nucleofecting in purified Cas9 and gRNA as an RNP complex has proven successful (Lingeman et al., 2017).

Given the nature of the SCD and its single basepair mutation, it is no surprise that it is a textbook example for CRISPR-Cas9 editing. Multiple labs have been able to show editing and correction of the SCD mutation in erythroid cell lines, primary human cells, and SCD patient cells (DeWitt et al., 2016; Park et al., 2019a; Vakulskas et al., 2018a). However, while correcting the SCD is possible, this approach does not address the rest of the β-hemoglobinopathy diseases. Combined, there are hundreds of mutations and this calls for a more generalized therapeutic approach. This has led to focusing efforts on the fetal globin repressors.

As mentioned before, the direct repressors of γ-globin are BCL11A and LRF. Gene-editing strategies have found that knockout of these repressors in eyrthroid cells indeed increase fetal hemoglobin (Masuda et al., 2015; Sankaran et al., 2008). However, knockout of transcription factors tend to have pleiotropic effects and can lead cellular growth defects. In the case of BCL11A, it has been shown that systemic knockout of BCL11A led to dyregulation of hematopoietic stem cell biology (Luc et al., 2016). Instead, researchers have now focused on editing an erythroid-specific enhancer region of BCL11A (Bauer et al., 2013). This has shown promising results in erythroid cell lines and primary cells that targeting this region is now a clinical trial for treating β-hemoglobinopathies (Wu et al., 2019).

Other methods of reactivating γ-globin include introducing naturally-occuring mutations or deletions at the globin locus that cause High Persistence of Fetal Hemoglobin (HPFH). Most of the mutations have now been molecularly characterized and have been found to disrupt repressor-binding motifs at the γ-globin genes (Wienert et al., 2018).

1.6 Current CRISPR-Cas9 Technologies

The integration of CRISPR-Cas9 into molecular biology has not only provided translational research with a new therapeutic tool, but it has also spawned new and innovative uses of the technology for fundamental research. For example, discovery of a catalytically dead Cas9 (dCas9) has led to a whole generation of CRISPR-Cas tools that are able to bind to DNA without cleaving it. One of the tools that has developed from this is the fusing of dCas9 to a repressive KRAB domain, resulting in the transcriptional repression of targetted genes (Gilbert et al., 2014; Qi et al., 2013). This technique, termed CRISPR-interference (CRISPRi), is an appealing tool as it allows the study of phenotypes resulting from a partial knockdown instead of a complete knockout and is useful in cases where knockout of a gene may be lethal to the cell. CRISPRi has

proven to be a useful tool to uncover biological mechanisms and has since been used in multiple genome-wide screens (Kabir et al., 2019; Liang et al., 2020).

Another tool that has resulted from the discovery of dCas9 is the fusing of dCas9 to base editors, enzymes that catalyze a change of one nucleotide to another (Gaudelli et al., 2017). Different groups have now used the base-editor to introduce point mutations with higher efficiency than with classical CRISPR-Cas9 editing as it bypasses the need for high HDR editing efficiency (Rees and Liu, 2018). These base-editors have been used to target the BCL11A enhancer region to disrupt a GATA1 binding site and to achieve upregulation of fetal hemoglobin (Zeng et al., 2020).

As the toolbox of CRISPR-Cas technologies continues to expand in innovative ways, our understanding of hemoglobin regulation, and our methods of treating β-hemoglobinopathies, will only continue to advance along with it.

2 ATF4 regulates MYB to increase γ-globin in response to loss of β-globin

A version of the material in this chapter was previously reported as:
Mandy Y. Boontanrart, Markus S. Schröder, Gautier M. Stehli, Marija Banović, Stacia K. Wyman, Rachel J. Lew, Matteo Bordi, Benjamin G. Gowen, Mark A. DeWitt, Jacob E. Corn (2020). ATF4 regulates MYB to increase γ-globin in response to loss of β-globin. *Cell Reports*. 2020 Aug 4.

The work has been adapted here with permission from all co-authors.

2.1 Connection to overall dissertation

In this chapter, we explored the cellular response to loss of *HBB*. We were most interested in uncovering the pathways involved in the upregulation of fetal hemoglobin. For this, we established the use of an erythroid cell model as well as generated and established gene-editing tools within the cell model. We used these tools to systematically answer biological questions pertaining to what occurs upon loss of *HBB*.

2.2 Summary

β-hemoglobinopathies can trigger rapid production of red blood cells in a process known as stress erythropoiesis. Cellular stress prompts differentiating erythroid precursors to express high levels of fetal γ-globin. However, the mechanisms underlying γ-globin production during cellular stress are still poorly defined. Here we use CRISPR-Cas genome editing to model the stress caused by reduced levels of adult β-globin. We find that decreased β-globin is sufficient to induce robust re-expression of γ-globin, and RNA-seq of differentiating isogenic erythroid precursors implicates ATF4 as a causal regulator of this response. ATF4 binds within the HBS1L-MYB intergenic enhancer and regulates expression of MYB, a known γ-globin regulator. Overall, the reduction of ATF4 upon β-globin knockout decreases the levels of MYB and BCL11A. Identification of ATF4 as a key regulator of globin compensation adds mechanistic insight to the poorly understood phenomenon of stress-induced globin compensation and could inform strategies to treat hemoglobinopathies.

2.3 Introduction

Red blood cells (RBCs), also known as erythrocytes, are packed with hemoglobin tetramers and circulate throughout the body to supply all tissues with oxygen. Adult RBCs primarily contain adult hemoglobin (HbA), which consists of two copies of α-globin and two copies of β-globin. β-thalassemic genetic disorders are caused by disruption of β-globin expression, causing loss of HbA and resulting in severe anemia, poor growth, and dramatically shortened lifespan. β-thalassemia can be ameliorated by re-expression of γ-globin, which complexes with α-globin to form fetal hemoglobin (HbF). γ-globin is normally expressed during development and silenced soon after birth, in an inverse relationship with β-globin.

Severe anemias such as β-thalassemia can trigger rapid production of red blood cells in a process known as stress erythropoiesis. Stress erythropoiesis is induced by tissue hypoxia resulting from anemia and involves a distinct erythropoietic program that favors increased γ-globin. For example, recovery from bone marrow transplant or other treatments that greatly reduce erythroblast levels is often characterized by high levels of γ-globin and HbF (Alter, 1979; Meletis et al., 1994; Papayannopoulou et al., 1980; Weinberg et al., 1986). Homozygous β-thalassemia patients transplanted with bone marrow from heterozygous siblings express high levels of HbF that can be sustained for up to two years after transplant (Galanello et al., 1989). β-thalassemia patients can also re-express high levels of γ-globin without transplant (Manca and Masala, 2008; Rochette et al., 1994). Intrinsic cellular processes can also mimic stress erythropoiesis, and ex vivo cultured stress erythroid progenitors express high levels of HbF (Xiang et al., 2015). Overall, this has led to the suggestion that erythroid stress recapitulates fetal erythropoiesis, but the pathways involved in fetal globin expression during this process remain to be determined.

Fetal hemoglobin reactivation is being actively explored as a potential route to treat a variety of globinopathies, including β-thalassemia and sickle cell disease (Platt et al., 1984; Wienert et al., 2018). These approaches stem from the observation that naturally occurring mutations which promote HbF re-expression can ameliorate β-globinopathy symptoms (Berry et al., 1992; Jacob and Raper, 1958). A variety of strategies for HbF re-expression are being pursued, including large deletions within the β-globin locus, mutations of the *HBG*1/2 promoter, and cell-specific reduction of the BCL11A repressor by enhancer editing (Sankaran, 2011; Wienert et al., 2018).

Here, we sought to model the cellular erythroid stress occurring after disruption of β-globin, hereafter referred to as β_0-stress, in order to determine the mechanism underlying spontaneous re-expression of fetal hemoglobin. We and others previously found that fetal hemoglobin was upregulated after genome editing of *HBB* in CD34+ adult mobilized hematopoietic stem and progenitor cells (HSPCs) to reverse the causative allele of sickle cell disease. In vitro differentiation of bulk edited cell populations induced high *HBG*1/2 transcript levels as compared to unedited cells, which translated to high levels of HbF tetramers (DeWitt et al., 2016; Park et al., 2019b; Vakulskas et al., 2018b). This effect persisted even after lonγ-term xenotransplantation of edited cells to immunodeficient mice (https://www.biorxiv.org/content/10.1101/432716v6).

We now show that β_0-stress caused by reductions in β-globin using either CRISPR-Cas genome editing or CRISPRi transcriptional repression are sufficient to induce increased levels of γ-globin in immortalized hematopoietic progenitors and adult mobilized CD34+ HSPCs. Time-course transcriptomics of isogenic *HBB* knockout and wild type cells during erythroid differentiation reveal that loss of β-globin leads to a transcriptional program that induces very high levels of *HBG* and moderately activates the transcription of other globins such as *HBE* and *HBZ*. Pathway analysis indicates that β-globin knockout cells reduce ATF4 (activating transcription factor 4) activity, leading to reductions in transcripts of many ATF4 targets. Endogenous mutation of ATF4 leads to upregulation of several globins, especially γ-globin, much like *HBB* knockout or knockdown. BCL11A levels were reduced in cells with knockout of *HBB* or an *ATF4* mutation, leading us to first suspect BCL11A as a possible ATF4 target. However, ChIP-seq showed no evidence for ATF4 binding anywhere near *BCL11A* regardless of differentiation status or *HBB* genotype. Instead, we find evidence for ATF4 binding within the *HBS1L-MYB* intergenic enhancer region. *MYB is* a known regulator of HbF expression at multiple levels, including promotion of *BCL11A* expression. (Antoniani et al., 2017). *MYB* transcript levels were reduced in cells with *HBB* knockout or harboring an *ATF4* mutation. Lastly, we found that *HBB* knockout reduces ATF4 binding at multiple genes involved in the unfolded protein response, leading to their transcriptional repression. Overall, our data indicate that β_0-stress inhibits an ATF4-mediated transcriptional program, which paradoxically reduces the unfolded protein response despite the presence of free alpha chains. Reduced ATF4 lowers *MYB* and BCL11A expression to upregulate multiple globins, especially γ-globin. These data provide mechanistic insight into the lonγ-observed but poorly-understood phenomenon of fetal globin expression during cell-intrinsic erythroid stress.

2.4 Results

2.4.1 Loss of *HBB* leads to upregulation of γ-globin

We previously observed that CRISPR-Cas9 editing at *HBB* in CD34+ mobilized peripheral blood HSPCs induced γ-globin transcription relative to unedited cells, leading to the formation of HbF tetramers (DeWitt et al., 2016). To mechanistically investigate how β_0-stress caused by knockout of *HBB* upregulates *HBG*, we used the HUDEP-2 cord-blood derived erythroid progenitor cell line (Kurita et al., 2013). HUDEP-2 cells normally express high levels of *HBB* and low levels of *HBG*, making them a popular model for the study of adult globin regulation (Bauer et al., 2013; Grevet et al., 2018; Wienert et al., 2017).

We made a clonal HUDEP-2 line with homozygous knockout of *HBB* (HBBko) using electroporation of a CRISPR-Cas9 RNP complexed with a well-validated *HBB*-targeting guide “e66”, which targets the exonic (e) region 66 base pairs from the start of *HBB* (DeWitt et al., 2016)(**Fig S1A**). HBBko was derived from a wild type (WT) HUDEP-2 clone (previously published as H2.1) in order to control for clonal effects associated with the heterogeneous HUDEP-2 parental population (Chung et al., 2019a; Wienert et al., 2017). γ-globin levels were strikingly increased in the HBBko line by intracellular

flow cytometry, with the vast majority of edited cells expressing detectable levels of the protein (**Fig 1A**).

Both the WT and HBBko lines were equally well in vitro differentiated to erythroblasts under standard conditions, as measured by staining for CD235a (Glycophorin A). WT and HBBko cells exhibited no apparent difference in cell survival during bulk differentiation, suggesting that β_0-stress induces γ-globin expression in the majority of cells as opposed to conferring a fitness advantage towards cells that stochastically express more γ-globin (**Fig S1B**). WT expressed high levels of β-globin after differentiation, but as expected HBBko completely lost β-globin protein by intracellular FACS. *HBB* mRNA was still detectable by qRT-PCR in HBBko, but was slightly reduced relative to WT in undifferentiated cells and 95% reduced in differentiated cells (**Fig S1C**).

DNA damage associated with CRISPR-Cas9 genome editing has been linked to lonγ-lasting cellular phenotypes in HSPCs, including p53 activation (Schiroli et al., 2019). To distinguish non-specific effects of genome editing from specific intervention at the β-globin locus, we tested a panel of Cas9 RNPs complexed with *HBB* guide RNAs targeting exonic or neighboring intronic regions of *HBB* (**Fig. 1B**). After bulk editing and in vitro differentiating the WT line, we found that multiple RNPs targeting coding regions consistently reduced *HBB* levels and increased *HBG*, as measured by qRT-PCR (**Fig 1C,D**). RNPs targeting neighboring intronic regions neither decreased *HBB* nor increased *HBG* despite similar levels of editing (**Fig S1D**). We found similar results when editing and in vitro differentiating CD34+ adult mobilized HSPCs from multiple donors with a panel of exonic or intronic *HBB*-targeting Cas9 RNPs (**Fig S1E, S1F**). Bulk-editing *HBB* in CD34+ HSPC cells did not significantly alter cell numbers during expansion (**Fig S1G**). Additionally, total cell numbers after in vitro differentiation were comparable amongst exonic and intronic *HBB*-targeting guides (**Fig S1H**). As with isogenically edited HUDEP-2 cells, this implies that β_0-stress in HSPCs induces γ-globin expression rather than positively selecting for cells that already express γ-globin.

The ability of multiple *HBB*-targeting RNPs to induce *HBG* expression suggested that this effect was independent of the genomic change and was instead a response to loss of *HBB* transcript. To test this hypothesis, we made a stable WT HUDEP-2 subclone expressing dCas9-KRAB (CRISPRi) (**Fig S1I**). We found that stable CRISPRi of *HBB* using two different guide RNAs led to potent downregulation of β-globin and upregulation of γ-globin protein by intracellular flow cytometry (**Fig 1E**). This was also reflected by levels of *HBB* and *HBG* transcripts as measured by qRT-PCR (**Fig S1J**). Overall, these data indicate that upregulation of *HBG* is not a consequence of DNA damage or genomic targeting, but is induced by loss of *HBB*.

To better characterize the globin protein composition resulting from loss of *HBB*, we used high performance liquid chromatography (HPLC) coupled to mass spectrometry (MS) of in vitro differentiated HUDEP-2 clones. We separated intact globin peaks using standard HPLC conditions while also collecting fractions. We subjected each fraction to mass spectrometry in order to unambiguously assign the globin composition of each peak. HPLC of the WT clone resulted in a peak pattern suggesting a typically “adult” set of globins, including high levels of HbA (**Fig 1F)**. This peak assignment was confirmed on the molecular level by mass spectrometry (Table S1). HPLC of the HBBko clone instead revealed a very different pattern of peaks (**Fig 1F**).

Mass spectrometry and comparison to literature globin HPLC led us to assign these as uncomplexed globins, α-globin aggregates, and HbF tetramers (Lechauve et al., 2019) (Table S1).

Our data thus far indicate that low levels of *HBB* caused by either genome editing or CRISPRi are associated with high levels of *HBG*. The presence of α-globin aggregates in cells with β0-stress suggests that an imbalance of hemoglobin proteins might lead to an unfolded protein response.

We therefore asked whether the relationship between levels of *HBB* and *HBG* was specific for reductions in *HBB* or a reflected an issue of general globin balance. For this, we utilized HUDEP-1 cells, which are similar to HUDEP-2s but exhibit a mostly immature globin profile of low *HBB* and high *HBG*. Pooled knockout of *HBG* using a CRISPR-Cas RNP led to loss of *HBG* mRNA and upregulation of *HBB* (**Fig S1K**). Therefore, our data suggest that erythroid progenitors sense globin levels during differentiation and attempt to compensate for missing globins by upregulating what they can. This model is consistent with the phenomenon of increased γ-globin during cell-intrinsic erythroid stress, and indicates that β0-stress modeled by the HBBko line is useful to molecularly characterize this process.

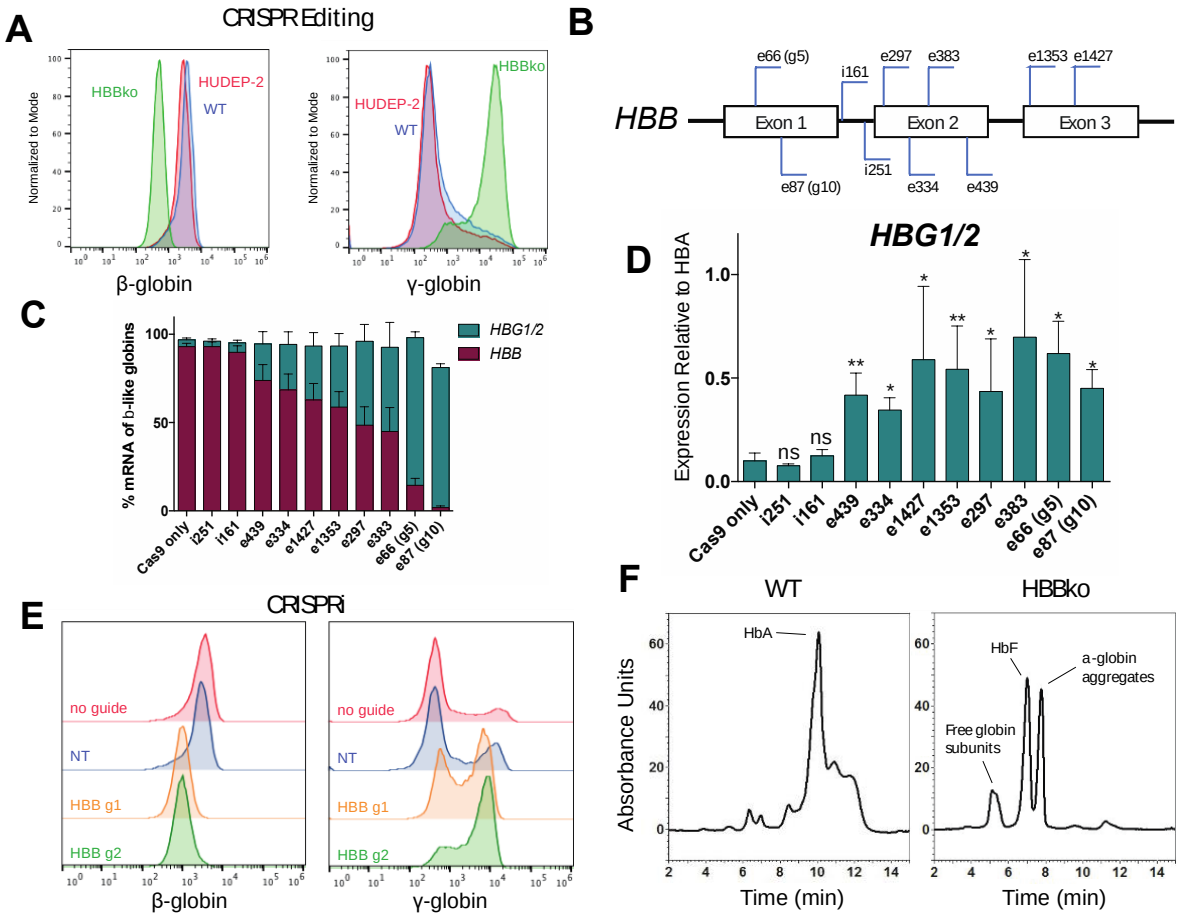


Figure 1. Loss of β-globin leads to increased γ-globin in HUDEP-2 cells and CD34+ HSPCs.

A) An *HBB* knockout (HBBko) line was generated from a wild type (WT) HUDEP-2 subclone though Cas9 ribonucleoprotein (RNP) targeting of exon 1 of *HBB*. Intracellular flow cytometry staining of β-globin and γ-globin of differentiated HUDEP-2 pool, WT clone, and HBBko clone.
B) Schematic of sgRNAs targeting *HBB* at either exonic (e) or intronic (i) positions. Numbers indicate the number of bases each RNP would cut after the transcription start site of the *HBB* gene. Previously published "G5" and "G10" sgRNAs are noted (DeWitt et al., 2016).
C) qRT-PCR of *HBB* and *HBG1/2* after pooled editing of the WT line with each sgRNA and 5 days of differentiation. Data is plotted as % of all β-like globins (*HBE, HBB, HBG1/2, HBD*).
D) qRT-PCR of *HBB* and *HBG1/2* plotted relative to *HBA* transcripts. Editing efficiency of each sgRNA is shown in **Fig S1D**. The data is presented as mean ± SD of three biological replicates. P value indicates paired, two-tailed student *t* test (ns, non-significant; *, P ≤ 0.05; **, P ≤ 0.01).
E) WT HUDEP-2 cells stably expressing dCas9-KRAB with guides targeting *HBB* were differentiated for 5 days and hemoglobin levels were measured by intracellular flow cytometry.
F) WT and HBBko cells were differentiated for 5 days and hemolysates were analyzed by HPLC. Peak compositions were identified through mass spectrometry (**Table S1**).

2.4.2 *HBB* knockout cells reduce an ATF4-mediated transcriptional response during differentiation

Because *HBB* and *HBG* reciprocally regulated one another, we asked if *HBB* loss affects the transcription of other globins. We performed biological triplicate in vitro differentiation of paired WT and HBBko clones. RNA-seq of terminally differentiated cells revealed widespread transcriptional alterations between WT and HBBko, with some of the largest fold changes being greatly reduced *HBB* and increased *HBG1*, *HBG2*, *HBE*, and *HBZ* (**Fig 2A**). There were relatively low total levels of *HBE* and *HBZ* in even HBBko cells after differentiation, such that *HBG1/2* comprised the vast majority of globin transcripts.

To capture the dynamics of differentiation, we performed triplicate time course RNA-seq by collecting samples in their undifferentiated state, after two days of differentiation, and after five days of differentiation (**Fig S2A-B**). Globally, the transcriptomes of WT and *HBB* knockout cells are similar throughout differentiation (**Fig S2C**). However, on top of the overall similarity, the knockout of *HBB* induces a distinct transcriptional response that emerged after two days of differentiation and intensified by five days (**Fig 2B**). This was reflected in PCA analysis, showing that WT and HBBko cells start quite similar, but have distinct differentiation trajectories (**Fig 2C**). Specifically examining the expression of globin genes, we found that as *HBB* expression declines during differentiation in HBBko cells as compared to WT, other globins become more highly expressed (**Fig 2D**).

Nonsense-Induced Transcriptional Compensation (NITC) is a response to nonsense-mediated decay (NMD) of an important transcript and has emerged as a new mechanism for transcriptional reprogramming to upregulate related genes (El-Brolosy et al., 2019; Ma et al., 2019; Rossi et al., 2015). However, we found that knocking down the mRNA decay exonuclease XRN1 or the single-stranded RNA binding protein SMG6 that are required for NITC had no effect on the upregulation of γ-globin in HBBko cells by intracellular FACS and qRT-PCR (**Fig S2D**).

We used pathway analysis to identify features in the RNA-seq data that could explain the transcriptional response in differentiated HBBko cells (Krämer et al., 2014a). We found that ATF4 (Activating Transcriptional Factor 4) and 37 out of 40 ATF4 target genes were downregulated in HBBko cells relative to WT cells (**Fig 2E, F**).

ATF4 is a bZIP family transcription factor associated with the integrated stress response (ISR) to unfolded proteins. ATF4 signaling decreases global protein synthesis while simultaneously inducing target genes. The downregulation of ATF4 and its ISR targets in HBBko cells was therefore surprising since these cells harbor free α-chains that might have conversely initiated the ISR (**Fig 1D**) (Chen and Zhang, 2019). However, a focused CRISPR screen found that knockout of the known ATF4 upstream regulator HRI (heme regulated eIF2a kinase) can induce moderate levels of HbF expression (Grevet et al., 2018). We hypothesized that both *HBB* and *HRI* knockout may induce globin synthesis by reducing ATF4 activity. This would mark ATF4 as a master regulator whose level maintains globin levels at homeostasis when a major globin becomes critically low.

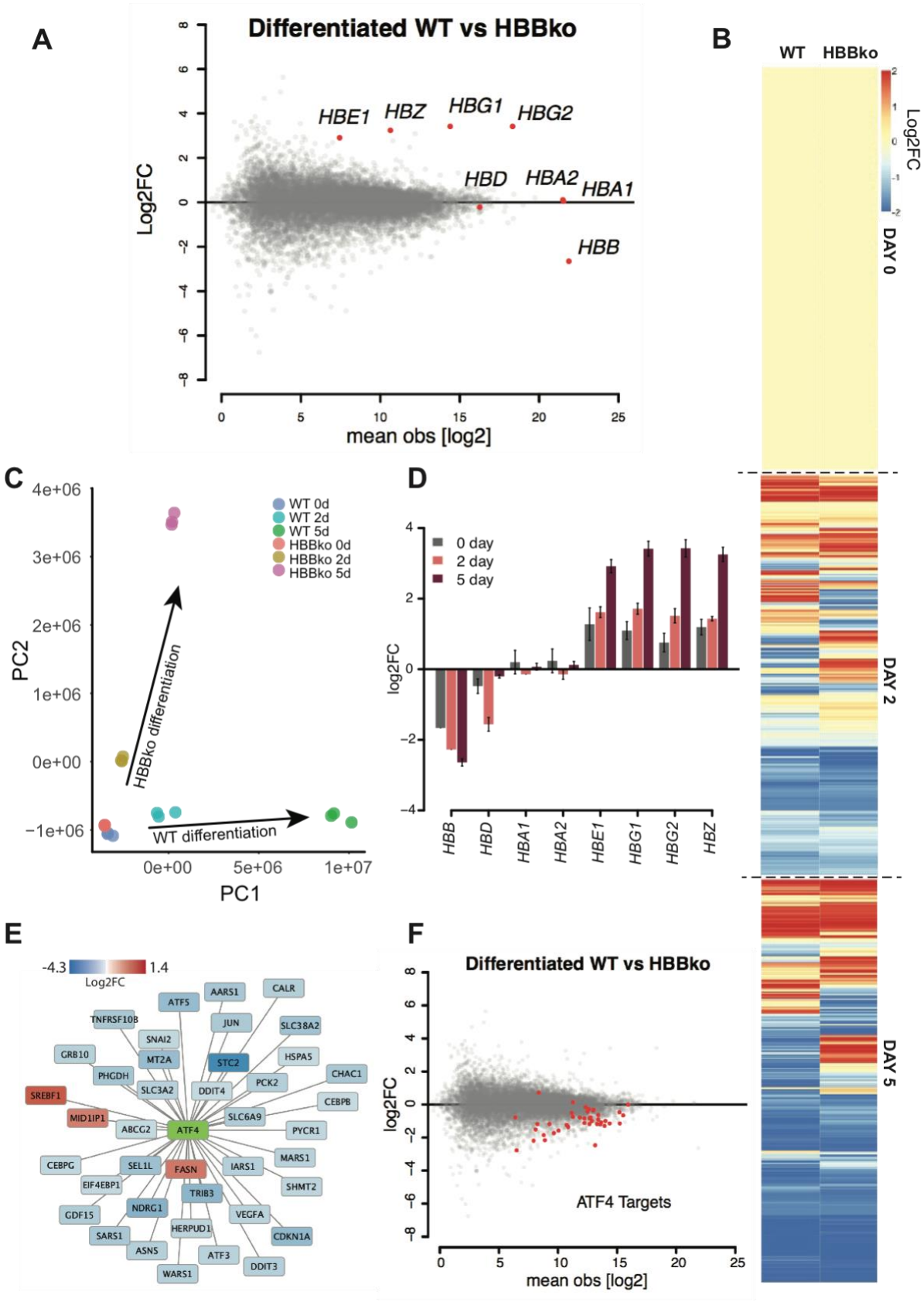

A
Differentiated WT vs HBBko
HBE1
HBZ
HBG1
HBG2
HBD
HBA2
HBA1
HBB
Log2FC
mean obs [log2]
B
WT
HBBko
Log2FC
DAY 0
DAY 2
DAY 5
C
WT 0d
WT 2d
WT 5d
HBBko 0d
HBBko 2d
HBBko 5d
HBBko differentiation
WT differentiation
PC1
PC2
D
0 day
2 day
5 day
log2FC
HBB
HBD
HBA1
HBA2
HBE1
HBG1
HBG2
HBZ
E
-4.3
1.4
Log2FC
ATF5
AARS1
CALR
TNFRSF10B
JUN
SLC38A2
SNAI2
GRB10
MT2A
STC2
HSPA5
PHGDH
PCK2
CHAC1
SREBF1
SLC3A2
DDIT4
CEBPB
MID1IP1
ABCG2
ATF4
SLC6A9
PYCR1
CEBPG
SEL1L
FASN
IARS1
MARS1
EIF4EBP1
TRIB3
SHMT2
GDF15
NDRG1
VEGFA
SARS1
HERPUD1
CDKN1A
ASNS
ATF3
WARS1
DDIT3
F
Differentiated WT vs HBBko
log2FC
ATF4 Targets
mean obs [log2]

Figure 2. Transcriptomics of time-course differentiation reveals an ATF4 signature in HBBko cells.

A) MA plot after 5 days of differentiation comparing HBBko and WT cells. Globin genes are highlighted in red. RNA-seq was performed in biological triplicates and data is shown as the mean.
B) Biological triplicate RNA-seq data of WT and HBBko cells undifferentiated (day 0), 2 days of differentiation, and 5 days of differentiation. Transcript levels are expressed as $\log_2$ fold-change normalized to their respective day 0 expression.
C) Principal component analysis of all RNA-seq samples shows agreement of replicates and alterations in differentiation trajectories between WT and HBBko. Replicated colors indicate n=3 biological replicates of each condition.
D) The expressions of embryonic, fetal, and adult globin genes are plotted for the 3 timepoints of differentiation. Data is shown as log2 fold change relative to WT cells and presented as mean ± SD of 3 biological replicates.
E) Pathway analysis identifies 38 out of 40 ATF4 targets as significantly affected in HBBko cells. Genes highlighted in blue and red are ATF4 targets that are downregulated and upregulated genes by log2 fold change on Day 5 of differentiation, respectively.
F) MA plot from 5 days of differentiation highlighting 40 known ATF4 targets (red). The mean of biological triplicates is shown.

2.4.3 ATF4 represses fetal hemoglobin expression

By Western blotting, we found that ATF4 levels are decreased in differentiated HBBko compared to WT cells (**Fig 3A, S3A**). ATF4 is itself positively regulated by phosphorylation of eIF2a via HRI. We also found decreased eIF2a phosphorylation in HBBko cells (**Fig 3A**). These data further support our hypothesis that ATF4 and the ISR pathway are not upregulated by free alpha chains in the context of HBBko cells. Instead, both *HBB* and HRI knockout may induce globin synthesis by reducing ATF4 activity.

To exclude the possibility of clonal effects during isolation of the HBBko clone, we tested whether ATF4 expression is decreased during bulk RNP knockout of *HBB* using a different guide RNA (e87, **Fig 1B**) than was used to generate HBBko. After bulk editing of *HBB* of HUDEP-2 cells using the e87 guide RNA and differentiation, we found decreased ATF4 expression and increased γ-globin by Western blot (**Fig S3B**), qRT-PCR (**Fig S3C**), and flow cytometry (**Fig S3D**). We further tested whether the decrease in ATF4 after knockout of *HBB* was specific to loss of *HBB* or whether it applied to loss of other globins as well. For this, we utilized HUDEP-1 cells which normally express high levels of *HBG* and no *HBB* (Kurita et al., 2013). We performed bulk *HBG1/2* RNP knockout and observed a compensatory increase in β-globin by Western blot (**Fig S3E**), but no difference in ATF4 expression. We note that basal ATF4 expression in untreated HUDEP-1 cells is low by Western blot (**Fig S3E**), which is consistent with a model that low levels of ATF4 yields higher levels of γ -globin. Overall, this suggests that decrease of ATF4 expression is specific to loss of *HBB*. Loss of *HBG* likely engages different pathways to upregulate *HBB* and whether these pathways are also involved in loss of *HBB* is currently unknown.

Taken together, our data suggest that reduction of β-globin leads to a decrease in ATF4 expression via a paradoxical reduction of eIF2a signaling. Reduced ATF4 leads to high levels of γ-globin in the context of *HBB* knockout.

Next, we developed two CRISPR-Cas9 genome editing strategies to further assess whether ATF4 was involved in globin homeostasis, and specifically in fetal globin regulation in cells that predominantly express adult globin. First, we used a dual-guide approach to excise the entirety of the *ATF4* gene, starting from the WT HUDEP-2 clone (**Fig S3F).** Two *ATF4* knockout clones resulting from total excision were named ATF4ko-1 and ATF4ko-2 (**Fig S3G**). The knockout clones showed successful knockout of ATF4 by Western blotting, with no protein detected even after treatment with cyclopiazonic acid (CPA), a drug that increases ATF4 expression through induction of endoplasmic reticulum stress (**Fig S3H).** Notably, ATF4ko-1 and ATF4ko-2 were unable to undergo in vitro differentiation, and instead exhibited extensive cell death after 5 days in differentiation media (**Fig S3I**). This suggests that complete loss of *ATF4* is not well tolerated and is consistent with a known role of ATF4 as essential in erythroid differentiation (Masuoka and Townes, 2002; Suragani et al., 2012; Zhang et al., 2019).

Second, we explored N-terminal truncation mutants of ATF4 with separation-of-function properties (Steinmüller and Thiel, 2003). We found that dual guides targeted to remove most of the first exon of *ATF4* (**Fig S3F**) lead to reinitiation of translation at the next ATG in the second exon. This produces a stable protein fragment of ~30 kDa lacking the N-terminal regulatory region but retaining the C-terminal dimerization and DNA binding domain (**Fig S3H**). We derived an edited clone from WT HUDEP-2s using the above strategy, naming it ATF4ΔN (**Fig S3G**). By ChIP-qPCR, we found that ATF4ΔN still binds the *ASNS* promoter but qRT-PCR indicated that ATF4ΔN does not support transcription of *ASNS* (**Fig S3J**). ATF4ΔN was more highly expressed than wild type ATF4 and no longer induced by ER stress (**Fig S3H**). However, this may be offset by the loss of ATF4ΔN's transcriptional activation activity. By contrast to ATF4ko-1 and ATF4ko-2, the ATF4ΔN clone was still able to successfully differentiate as measured by high live cell numbers and successful globin expression after 5 days of culture in differentiation media (**Fig S3I, Fig S3M**). ATF4 has been found to heterodimerize with other transcription factors, such as CEPB (Su and Kilberg, 2008). Therefore, while ATF4ΔN may have reduced transciptional activity, retention of ATF4's heterodimer partner activity could be sufficient for cell differentiation.

We tested the globin status of the ATF4ko and ATF4ΔN cell lines. Undifferentiated ATF4ko and ATF4ΔN cells expressed normal levels of β-globin and slightly increased γ-globin by intracellular FACS (**Fig 3B, Fig S3L**) and Western Blot (**Fig 3C**). Strikingly, in vitro differentiation of ATF4ΔN led to high levels of transcripts for *HBB, HBE, HBZ*, and *HBG* (**Fig S3M**). This was reflected in increased γ-globin by intracellular FACs (**Fig 3D, Fig S3K**). The global upregulation of globin expression with especially high levels of γ-globin upon differentiation of ATF4ΔN mirrors the response in differentiated *HBB* knockout cells. These data indicate that ATF4 is involved in repressing fetal hemoglobin in the undifferentiated state, and a lack of ATF4 signaling in the differentiated state leads to upregulation of multiple globins that mimics the response to lack of *HBB*.

The ability of *ATF4* knockout or ATF4ΔN clones to modestly upregulate HbF in the undifferentiated state prompted us to ask whether ATF4 is involved in basal HbF

repression in other cell types. We used CRISPR-Cas9 to remove the entire *ATF4* gene from K562 erythroleukemia cells, which normally express no β-globin and moderate levels of γ-globin. We generated two different *ATF4* knockout K562 cell lines that we termed ATF4ko.562-1 and ATF4ko.562-2 using an analogous protocol as for the HUDEP-2 ATF4 knockout lines (**Fig S3F**). We found that both K562 *ATF4* knockout lines became strikingly red and expressed very high levels of γ-globin but still failed to express β-globin (**Fig 3E**), even when they were maintained in undifferentiated culture conditions.

We next tested whether overexpression of *ATF4* is sufficient to prevent upregulation of fetal hemoglobin. We attempted to overexpress *ATF4* in HUDEP-2 cells, but were unsuccessful as high levels of ATF4 were not well tolerated and induced cell death in this model (data not shown). Instead, we stably expressed a doxycycline-inducible *ATF4* construct in *ATF4* knockout K562 cells, which basally express high levels of gamma globin. This construct could not be used in HUDEP-2 cells, which require doxycycline in their media. We found that 48 hours of doxycycline treatment led to high levels of ATF4 and decreased γ-globin and *HBG* transcript levels by Western blot (**Fig 3F**) and qRT-PCR, respectively (**Fig 3G**).

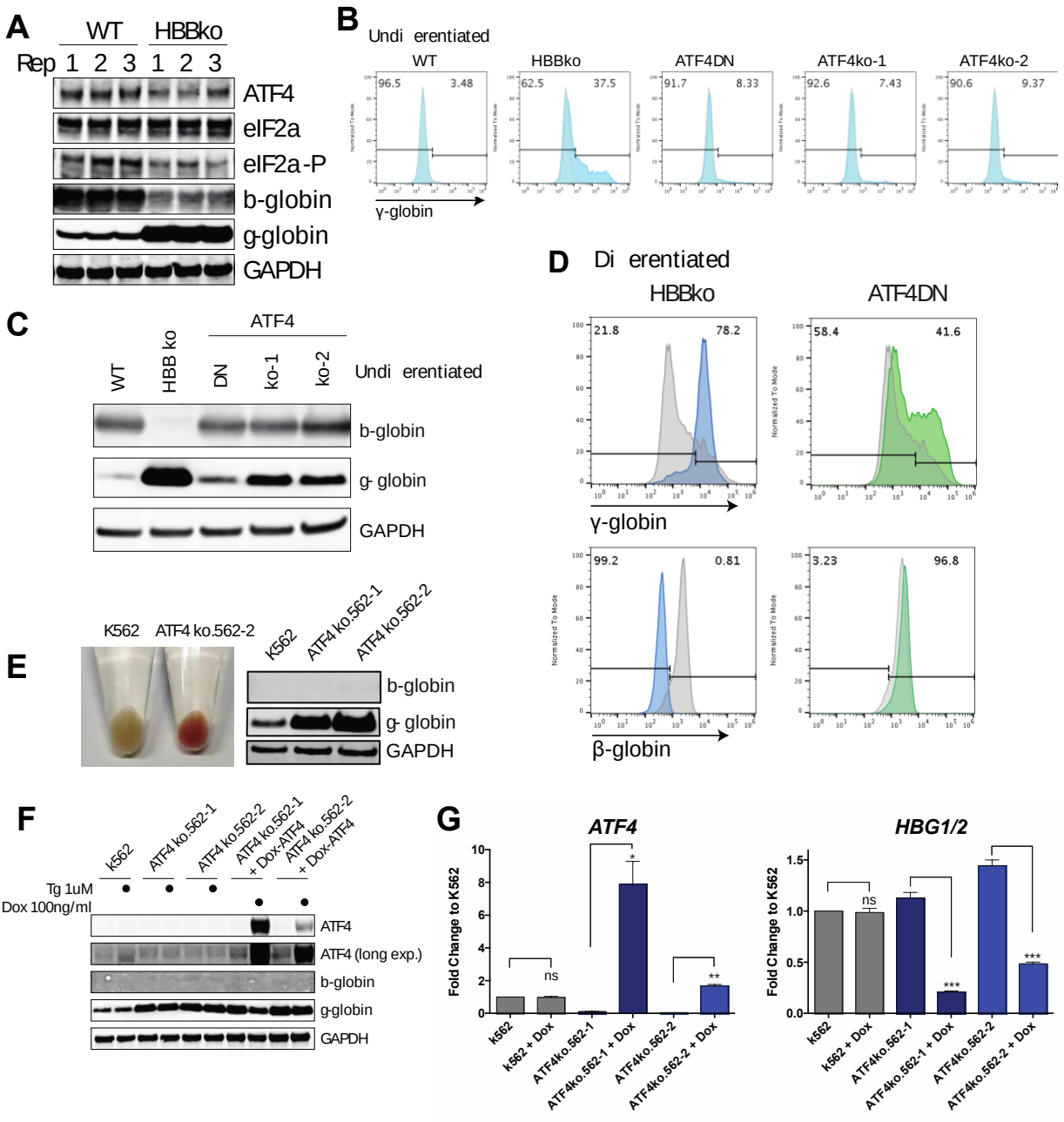


Figure 3. Endogenous mutation of ATF4 leads to increased γ-globin in HUDEP-2 and K562 cells.

A) Western blot of 5 day differentiated WT and HBBko showing ATF4, eIF2a, eIF2a-P, β-globin and γ-globin. Biological triplicates are shown. Quantification is shown in **Fig S3A**.

B) Intracellular FACS for γ-globin of undifferentiated HUDEP2 cells comparing WT, HBBko, and ATF4 clones. Biological quadruplicates are quantified and shown in **Fig S3L**.

C) Western blot of undifferentiated WT, HBBko, and ATF4 mutant cells showing β-globin and γ-globin levels.

D) Intracellular FACS for γ-globin of differentiated HBBko and ATF4ΔN cells as compared to WT (trace in gray). Biological triplicates are quantified in **Fig S3K**.
E) Left panel: cell pellets from wild type K562 and ATF4ko K562 cells. Right panel: Western blot for β-globin and γ-globin in WT and ATF4ko K562 clones. qRT-PCR confirmed loss of *ATF4* transcript in ATF4ko clones and increased *HBG1/2* (**Fig S3N**).
F) Western blot of k562 cells and ATF4ko.562-1,2 with doxycycline-inducible ATF4 constructs treated with 1uM Thapsigargin or 100ng/ml doxycycline for 48 hours showing ATF4, β-globin, and γ-globin.
G) qRT-PCR of *ATF4* and *HBG1/2* in k562 cells and ATF4ko.562-1/2 cells. ATF4ko.562-1/2 have stable expression of doxycycline-induced *ATF4* transgene. Cells were treated with 100ng/ml doxycycline for 48 hours. The data is presented as mean ± SD of three biological replicates. P value indicates paired, two-tailed student *t* test (ns, non-significant; *, P ≤ 0.05; **, P ≤ 0.01; ***, P ≤ 0.001).

2.4.4 Loss of ATF4 downregulates BCL11A

As a transcription factor, ATF4 could directly or indirectly regulate the expression of *HBG1/2* and other globins. BCL11A and ZBTB7A (also known as LRF) are two transcriptional repressors that directly repress fetal globin expression after birth. We asked whether BCL11A and ZBTB7A are misregulated in our *HBB* knockout model of β_0-stress or after loss of ATF4.

Examining the WT and HBBko time course differentiation RNA-seq data, we found that *ATF4* transcripts were much lower in HBBko in the undifferentiated state, but stabilized to WT levels after 5 days of differentiation. The expression of both *BCL11A* and *ZBTB7A* were reduced in the HBBko line, with BCL11A expression declining during differentiation and ZBTB7A remaining at a consistent reduced level (**Fig S4A**). This corresponded to reduced expression of BCL11A protein in HBBko during differentiation (**Fig 4A**). A decrease in BCL11A also provides a potential explanation for the observed increase in embryonic globin transcripts in the HBBko line, as BCL11A binds to the *HBE* and *HBZ* promoters (Liu et al., 2018; Xu et al., 2010).

Stable CRISPRi knockdown of BCL11A resulted in similar increases in γ-globin to knockdown of *HBB*, as measured by FACS (**Fig S4B**). Knockout or knockdown of *HBB* also both led to a decrease in BCL11A by Western Blot (**Figure 4B**). In the undifferentiated state, the ATF4ΔN and ATF4ko-1 and 2 expressed normal levels of BCL11A and ZBTB7A (**Fig S4C**). However, after differentiation the ATF4ΔN clone exhibited lower levels of BCL11A protein (**Fig 4C**). *HBB* knockout and reduced ATF4 signaling might therefore increase γ-globin expression by directly or indirectly reducing levels of BCL11A during differentiation. However, this is at odds with data from K562 cells, which do not express BCL11A (**Fig S4D**), yet dramatically increase *HBG* upon knockout of *ATF4* (**Fig 3D**).

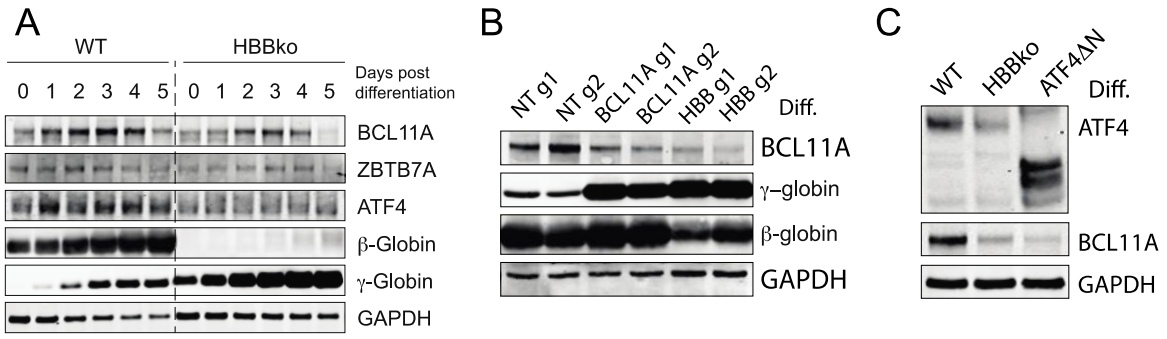


Figure 4. Loss of β-globin or ATF4 downregulates BCL11A

A) Western blot of HBBko and WT cells over the course of differentiation showing BCL11A, ZBTB7A, ATF4, β-globin, and γ-globin.
B) Western blot of 5 day differentiated HUDEP-2 cells showing CRISPRi knockdown efficiency of BCL11A and β-globin, as well as resulting increases in γ-globin protein.
C) Western blot of 5 day differentiated WT, HBBko, and ATF4ΔN comparing BCL11A levels.

2.4.5 ATF4 binding is attenuated in HBBko cells relative to WT

To test whether ATF4 is directly involved in transcriptional regulation of BCL11A, we performed endogenous ATF4 ChIP-seq in various HUDEP-2 genetic backgrounds and differentiation states to map ATF4 targets during hemoglobinization. In the undifferentiated state, we performed ATF4 ChIP-seq in WT, HBBko, and ATF4ko (negative control) clones. After five days of in vitro differentiation, we performed ATF4 ChIP-seq in WT and HBBko clones, and with an IgG isotype negative control in the WT clone (since the ATF4ko clone cannot be differentiated). We also performed ATF4 and IgG ChIP-seq in the differentiated ATF4ΔN clone, which verified that this ATF4 mutant retains a wild type binding profile on a genome-wide level (**Fig S5A)**. Lastly, we performed ATF4 ChIP-seq on human CD34+ derived early and late erythroblasts, with IgG as a negative control.

The overall number of ATF4 binding sites increased upon differentiation in both primary human erythroblasts as well as differentiated HUDEP-2 cells. (**Fig 5A, Fig S5B).** This increase in ATF4 binding further highlights the importance of ATF4 signaling during erythroid differentiation.

Between all four experimental conditions in isogenically controlled HUDEP-2 cells (WT undifferentiated, HBBko undifferentiated, WT differentiated, HBBko differentiated), we found 1245 unique ATF4 ChIP-seq peaks with an Irreproducible Discovery Rate (IDR) > 2% and a signal value > 30 over background (ATF4 knockout or IgG control). Bona fide ATF4 binding sites showed high read count and clear enrichment over background (e.g up to 255 reads and 85 fold enrichment at ASNS). Unbiased motif discovery on these ChIP-seq peaks identified a central ATF4 consensus

site (**Fig 5B, S5C**), verifying the quality of the datasets. Cross-comparison with our time-course RNA-seq data revealed that ATF4 binding was associated with a subset of well-expressed genes, and less ATF4 binding was present at lowly expressed genes. This further indicates that the identified ATF4 sites are involved in active transcription (**Fig S5D**).

We used the four isogenically-controlled HUDEP-2 ChIP-seq datasets to identify new ATF4 binding sites that varied depending on differentiation status and *HBB* genotype. Unbiased enrichment analysis against databases of transcription factor targets or transcription factor ChIP-seq (Huang et al., 2009; Kuleshov et al., 2016) revealed clear patterns for several distinct categories of ATF4 ChIP-seq peaks.

"Constitutive" ATF4 sites were defined as present regardless of differentiation status in both WT and HBBko. Constitutive peaks were dominated by an ATF4 signal, validating our set of ChIP-seq experiments against datasets in other cell contexts (**Figure 5C, Fig S5E**). "WT Undifferentiated" ATF4 sites were defined as at least two-fold enriched in WT cells in the undifferentiated state and not found in the differentiated state. WT Undifferentiated peaks were highly represented at targets and binding sites of ATF4 partners that either form bZIP heterodimers with ATF4 (e.g. CEBPs) or are part of the same family as ATF4 (e.g. ATF3) (**Figure 5D, Fig S5F**. "WT Differentiated" ATF4 sites were defined as at least two-fold enriched in WT cells in the differentiated state and not found in the undifferentiated state. WT Differentiated peaks were unexpectedly represented at targets and binding sites of transcription factors involved in hematopoietic differentiation (e.g. GATA1/2, and ZBTB7A) but with no known ATF4 partnership (**Figure 5E, Fig S5G**).

ATF4 binding to GATA1 target genes implies that ATF4 cooperates with hematopoietic differentiation transcription factors at their target genes, providing a rationale for ATF4's requirement during hematopoietic differentiation. We compared our ATF4 ChIP-seq data to GATA1 ChIP-seq in primary human erythroblasts (Dunham et al., 2012), and identified 753 cases in which ATF4 and GATA1 peaks are associated with the same gene and are within 10 kilobases of one another (**Fig S76A**. Some of these were cases of non-overlapping regulatory regions, in which the same genomic target was associated with distinct ATF4 and GATA1 peaks. A large proportion of putative co-regulatory sites involved overlap between the ATF4 and GATA1 ChIP-seq signal (371 of 2766 total ATF4 peaks, 13.41%)(**Fig S7A**). Performing motif analysis with the ATF4 ChIP-seq peaks as an anchor, we found that overlapping sites frequently harbor a GATA1 consensus motif either underneath or immediately adjacent to the ATF4 peak (**Fig S6B**). Taken together, these data suggest that ATF4 has unappreciated regulatory functions that overlap those of the GATA1 hematopoietic transcription factor, and highlight the importance of ATF4 in erythroid differentiation.

Returning to the role of ATF4 in loss of *HBB*, we found that ChIP-seq signal at ATF4 binding sites was attenuated in HBBko cells relative to WT in both the differentiated and undifferentiated states (**Fig 5F-G**). Binding of ATF4 was particularly reduced at several genes involved in the unfolded protein response and feedback regulators of ATF4 (e.g. XBP1, HSPA5/BiP, CREB3L1, EDEM1, ATF3, ATF6, RACK1, SEL1L, and TRIB3). Many of these factors are not canonical ATF4 targets, but their expression levels were also reduced in RNA-seq of differentiating HBBko cells (**Fig**

S6C). These data are consistent with downregulation of ATF4 levels and signaling in HBBko cells, despite increased alpha chain aggregates in HBBko cells.

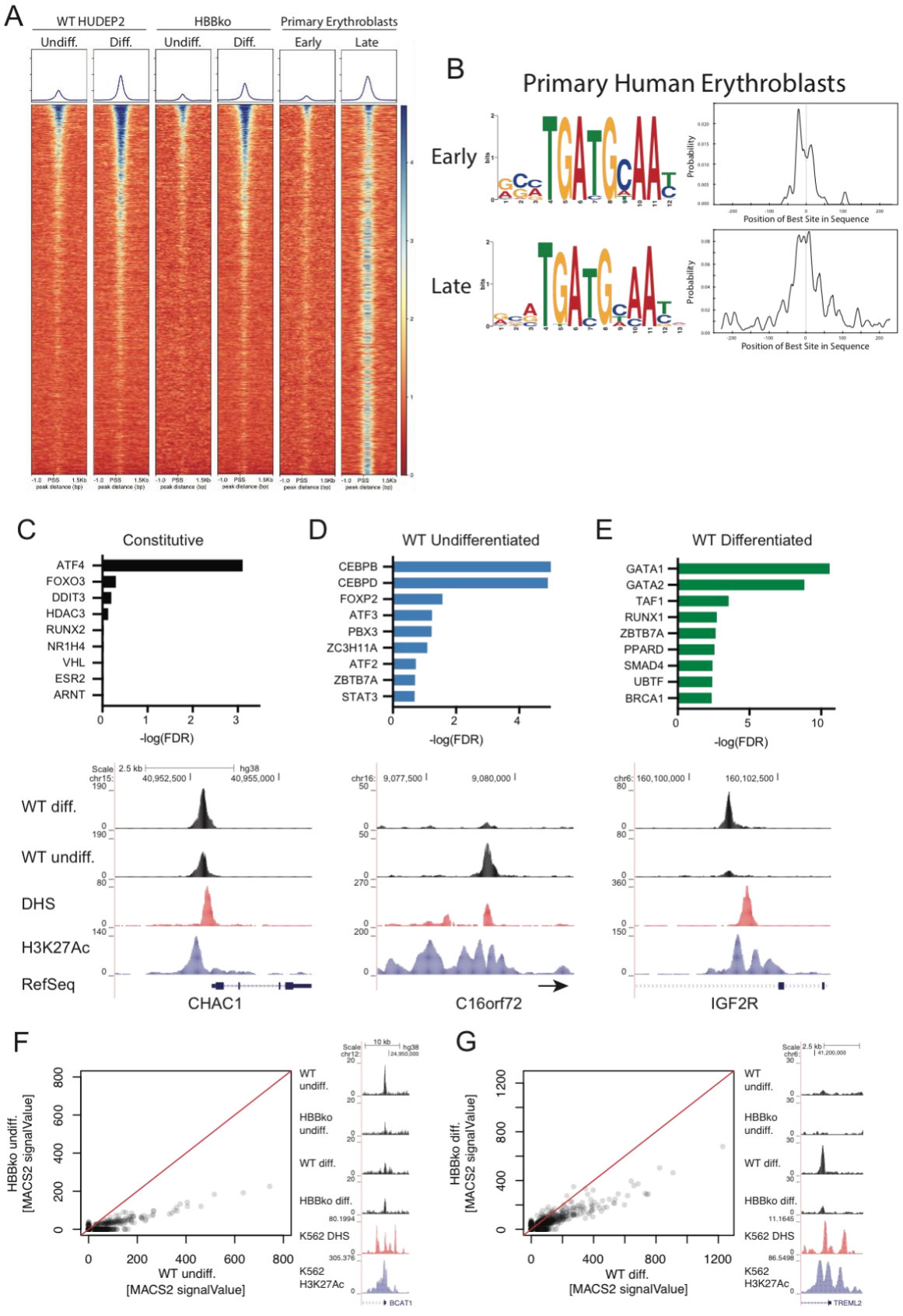

A
WT HUDEP2
HBBko
Primary Erythroblasts
Undiff.
Diff.
Undiff.
Diff.
Early
Late
B
Primary Human Erythroblasts
Early
Late
Probability
Position of Best Site in Sequence
C
Constitutive
ATF4
FOXO3
DDIT3
HDAC3
RUNX2
NR1H4
VHL
ESR2
ARNT
-log(FDR)
D
WT Undifferentiated
CEBPB
CEBPD
FOXP2
ATF3
PBX3
ZC3H11A
ATF2
ZBTB7A
STAT3
-log(FDR)
E
WT Differentiated
GATA1
GATA2
TAF1
RUNX1
ZBTB7A
PPARD
SMAD4
UBTF
BRCA1
-log(FDR)
WT diff.
WT undiff.
DHS
H3K27Ac
RefSeq
CHAC1
C16orf72
IGF2R
F
HBBko undiff. [MACS2 signalValue]
WT undiff. [MACS2 signalValue]
WT undiff.
HBBko undiff.
WT diff.
HBBko diff.
K562 DHS
K562 H3K27Ac
BCAT1
G
HBBko diff. [MACS2 signalValue]
WT diff. [MACS2 signalValue]
WT undiff.
HBBko undiff.
WT diff.
HBBko diff.
K562 DHS
K562 H3K27Ac
TREML2

Figure 5**.** ATF4 signaling is attenuated in HBBko cells

A) Heatmap of ATF4 ChIP-seq showing fold change over each respective control for all peaks across samples. Differentiation increases ATF4 occupancy at most identified binding sites.
B) Unbiased motif discovery on ATF4 ChIP-seq in primary human early erythroblasts identifies the ATF4 consensus sequence.
C) Unbiased enrichment analysis of ATF4 ChIP-seq peaks present in both the undifferentiated and differentiated states against databases of transcription factor targets. Example ATF4 ChIP-seq profiles are shown below, together with reference ENCODE DnaseI hypersensitivity and H3K27Ac data. Lack of ATF4 signal at the globin, *BCL11A*, and *ZBTB7A* loci is shown in **Fig S6D-F**.
D) Unbiased enrichment analysis of ATF4 ChIP-seq peaks present in only the undifferentiated state.
E) Unbiased enrichment analysis of ATF4 ChIP-seq peaks present in only the differentiated state.
F) MACS2 signal value for all ATF4 CHIP-seq peaks comparing WT to HBBko cells in the undifferentiated state. Example ATF4 ChIP-seq profiles are shown to the right of each global ChIP-seq comparison.
G) MACS2 signal value for all ATF4 CHIP-seq peaks comparing WT to HBBko cells in the differentiated state.

2.4.6 ATF4 binds the *HBS1L-MYB* intergenic enhancer region and regulates *MYB* expression

We were surprised to find no evidence of ATF4 binding anywhere in the vicinity of the globin locus, *BCL11A*, or *ZBTB7A* including all known promoters and enhancers (**Fig S6D-F**). This was true of all cells tested, regardless of genotype and whether they were HUDEP-2 cells or CD34+ derived erythroblasts. Upon further investigation, we instead found evidence of ATF4 binding at an ATF4 consensus motif within the *HBS1L-MYB* intergenic enhancer region (**Fig 6A**). This enhancer region is known to regulate fetal globin expression, and MYB is an essential transcription factor for erythropoiesis and hematopoiesis that regulates fetal globin, *BCL11A*, and *ZBTB7A* (Bianchi et al., 2010; Cavazzana et al., 2017; Jiang et al., 2006; Lettre et al., 2008; Stadhouders et al., 2014; Thein et al., 2007; Wang et al., 2018). The LDB1 erythroid transcription factor complex, which contains GATA1 (Soler et al., 2010), acts on the *HBS1L-MYB* intergenic region at multiple sites (Stadhouders et al., 2014). The ATF4 peak identified here overlaps a DNAseI hypersensitive site and is directly adjacent to the -92 LDB1 site (**Fig 6A**). This adjacency of an ATF4 peak to an LDB1 complex site in the *HBS1L-MYB* intergenic region provides specific context to the global co-occurrence of ATF4 and GATA1 peaks highlighted earlier (**Fig 5E, Fig S6A-B)**.

Having identified a previously unappreciated ATF4 binding site in a known *MYB* regulatory region, we asked whether ATF4 levels and the loss of *HBB* affect *MYB.* Indeed, *MYB* expression was greatly decreased after bulk Cas9 editing of *HBB* and differentiation of HUDEP-2 cells (**Fig 6B**). *MYB* expression was also decreased in differentiated ATF4DN HUDEP-2 cells as compared to WT (**Fig 6C**), and in ATF4 knockout K562 cells (**Fig 6E**). Significantly, doxycycline-induced expression of a stable *ATF4* transgene rescued *MYB* in the ATF4 knockout background (**Fig S6G**). By western blot, we observed that MYB was downregulated in ATF4 knockout cells and over-expression of ATF4 rescued MYB levels while decreasing γ-globin (**Fig 6D**). Since K562 cells do not express BCL11A, involvement of the upstream multifunctional regulator *MYB* potentially explains how loss of ATF4 leads to upregulation of fetal globin in this context. Taken together, our data indicate that reduced ATF4 upregulates fetal globin through direct regulation of *MYB* and indirect regulation of *BCL11A*.

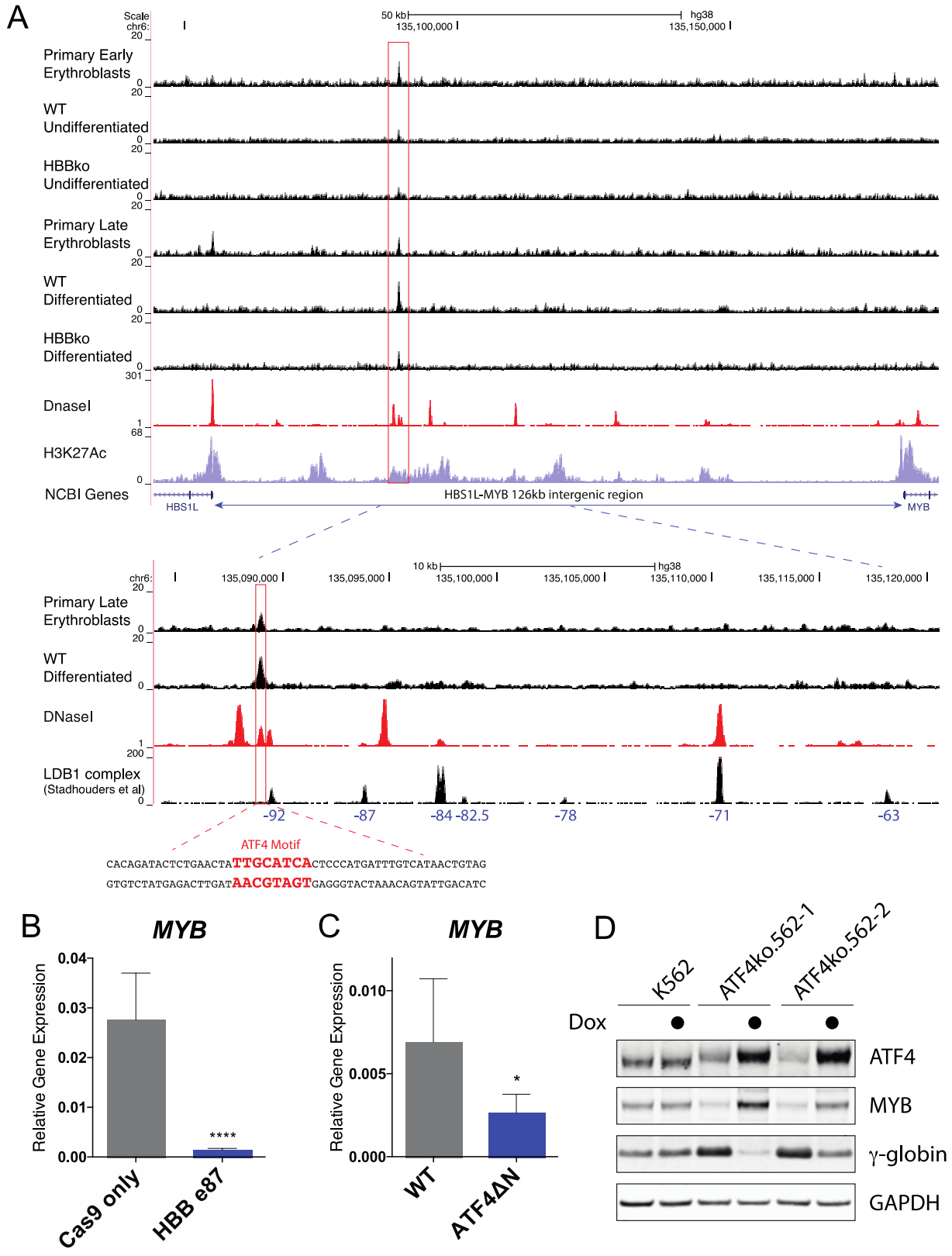

A
Scale
chr6:
50 kb
hg38
135,100,000
135,150,000
Primary Early Erythroblasts
WT Undifferentiated
HBBko Undifferentiated
Primary Late Erythroblasts
WT Differentiated
HBBko Differentiated
DnaseI
H3K27Ac
NCBI Genes
HBS1L
HBS1L-MYB 126kb intergenic region
MYB
10 kb
hg38
chr6:
135,090,000
135,095,000
135,100,000
135,105,000
135,110,000
135,115,000
135,120,000
Primary Late Erythroblasts
WT Differentiated
DNaseI
LDB1 complex (Stadhouders et al)
-92
-87
-84
-82.5
-78
-71
-63
ATF4 Motif
CACAGATACTCTGAACTATTGCATCACTCCCATGATTTGTCATAACTGTAG
GTGTCTATGAGACTTGATAACGTAGTGAGGGTACTAAACAGTATTGACATC
B
MYB
Relative Gene Expression
0.04
0.03
0.02
0.01
0.00
****
Cas9 only
HBB e87
C
MYB
Relative Gene Expression
0.010
0.005
0.000
*
WT
ATF4ΔN
D
K562
ATF4ko.562-1
ATF4ko.562-2
Dox
ATF4
MYB
γ-globin
GAPDH

Figure 6. ATF4 regulation within the *HBS1L-MYB* intergenic enhancer region

A) Top Panel: ATF4 ChIP-seq tracks at the *HBS1L-MYB* intergenic enhancer region. An ATF4 peak is boxed in red. DnaseI track is from previously published data in CD34+ primary cells (Neph et al., 2012). H3K27Ac track is from previously published primary erythroblasts (Xu et al., 2012). Bottom panel: zoomed in ATF4 ChIP-seq tracks along with previously published LDB1 complex ChIP-seq track (Stadhouders et al., 2014). Numbers in blue correspond to characterized LDB1 regulatory regions and indicate the number of kb upstream from *MYB* gene (Stadhouders et al., 2014). An ATF4 peak containing an ATF4 consensus motif is boxed in red.
B) qRT-PCR for *MYB* expression in bulk edited and differentiated HUDEP2 cells using guide *HBB* e87. The data is presented as mean ± SD of six biological replicates. P value indicates unpaired, two-tailed student *t* test (****, P ≤ 0.0001).
C) qRT-PCR for *MYB* expression in differentiated ATFDN compared to WT cells. The data is presented as mean ± SD of three biological replicates. P value indicates paired, two-tailed student *t* test (*, P ≤ 0.05).
D) Western blot of K562 and ATF4 knockout cells with a doxycycline-inducible ATF4 transgene showing ATF4, MYB, and γ-globin. Cells were treated with or without 50ng/ml of doxycycline for 5 days.
E) qRT-PCR for *MYB* expression in K562 and ATF4ko.562-1/2 with doxycycline-inducible ATF4 constructs treated with 100ng/ml doxycycline for 48 hours. The data is presented as mean ± SD of three biological replicates. P value indicates unpaired, two-tailed student *t* test (ns, non-significant; **, P 0.005).

2.5 Discussion

Stress erythropoiesis results from a high erythropoietic drive and is induced during anemias such as β-thalassemia. Cell-intrinsic erythroid stress leads to upregulation of γ-globin, but the pathways that mediate this phenomenon have remained unclear. We found that β_0-stress caused by loss of β-globin in erythroid precursors induces ATF4-mediated reprogramming that includes upregulation of several globins including very high levels of γ-globin. ATF4 signaling is required to maintain levels of BCL11A during differentiation, but ATF4 surprisingly does not bind *BCL11A* or *ZBTB7A*, nor the globin locus. Instead, our data support a model in which reduced ATF4 signaling during β_0-stress regulates γ-globin levels via *MYB*.

Translation of ATF4 mRNA is regulated by phosphorylation of eIF2-α, which is itself activated by several different kinases including HRI (Chen, 2014). Upon stress, the HRI pathway will inhibit protein synthesis, which then promotes efficient translation of ATF4 mRNA. Knockout of *HRI* was reported to decrease expression of BCL11A and increase expression of *HBG* (Grevet et al., 2018). However, the link between HRI and BCL11A remained undetermined. Our data suggest that a lack of HRI leads to decreased eIF2-α-phosphorylation and reduction of ATF4 activity, similar to what we observe for knockout of *HBB*. Deletions of *HRI*, *HBB*, and *ATF4* all lead to reduced *BCL11A*.

While our manuscript was in revision, a transcription factor focused CRISPR screen found that knockout of ATF4 upregulates gamma globin (Huang et al., 2020). Our own work reveals why ATF4 might be reduced in the first place, and together the two results independently highlight ATF4 as a regulator of globin expression. In contrast to our results, the other study found evidence for ATF4 binding within the *BCL11A* +55 enhancer. While we originally hypothesized that ATF4 directly regulates BCL11A, none of our carefully controlled ATF4 ChIP-seq datasets in multiple genetic backgrounds and cell types supported this hypothesis. Additionally, given that knockout of ATF4 increases γ-globin in K562 cells that do not express BCL11A, ATF4-mediated regulation of γ-globin must not be solely dependent upon BCL11A. Our data instead show that ATF4 binds the *HBS1L-MYB* intergenic enhancer region, adjacent to an LDB1 erythroid complex binding site. ATF4 binding at the *HBS1L-MYB* intergenic enhancer is also present but not commented upon in ChIP-seq data from the study that suggested direct regulation of *BCL11A* by ATF4 (Huang et al., 2020). Notably, ATF4 binds to the BCL11A intronic enhancer in both human and mouse cells, but ATF4 only regulates γ-globin in human cells (Huang et al., 2020). Mice harbor several mutations that disrupt the human *HBS1L-MYB* intergenic region ATF4 motif, and ATF4 does not bind the intergenic region in murine G1-ER4 cells by ChIP-seq (Huang et al., 2020).

In human cells, we found that levels of MYB are responsive to both *HBB* status and *ATF4* genotype. *MYB* is a known regulator of fetal globin on multiple levels, including *BCL11A* and *ZBTB7A* (Bianchi et al., 2010; Cavazzana et al., 2017; Stadhouders et al., 2014; Wang et al., 2018). We therefore suggest an alternative model in which ATF4 directly regulates *MYB* to indirectly regulate γ-globin. Follow-up experiments will be needed to further unravel the connections between ATF4 and MYB, as well as to test the differential activity of the ATF4 upon *MYB* in mouse models.

We found that knockout of *HBB* or *ATF4* leads to upregulation of multiple globins, with *HBG1/2* exhibiting the largest change, as well as reductions in almost all known ATF4 targets. Overall, this suggests that ATF4 signaling in response to loss of β-globin is part of a larger program to maintain high globin levels in differentiating erythroid cells. The response to loss of β-globin could include HRI-like signaling as well as additional facets. Knockout of *HRI* has been suggested as a strategy to ameliorate anemic disorders via re-expression of HbF (Grevet et al., 2018). Indeed, knockout or CRISPRi of *HBB* leads cells to express remarkable levels of *HBG1/2*. But the broad roles of ATF4 in processes other than erythropoiesis also suggest that this strategy should be pursued with caution.

The HRI-eIF2aP-ATF4 axis represses globin synthesis under low heme or iron conditions in order to keep excess globins in line with the availability of cofactors (Chen and Zhang, 2019; Crosby et al., 2000; Suragani et al., 2012). Unfolded proteins such as α-chain aggregates can also activate HRI-eIF2aP-ATF4, leading to the integrated stress response and repression of protein synthesis. ATF4 downstream of reduced *HBB* could therefore either be activated by an unfolded protein response or attenuated by reduced HRI activity caused by excess heme. While knockout of *HBB* leads to the formation of α-chain aggregates, we paradoxically found that ATF4 signaling is reduced. This suggests that excess heme remaining after loss of *HBB* and intact HbA plays a dominant role over alpha chain aggregates to reduce ATF4 activity. Indeed, supplementing K562 cells with hemin leads to a reduction in MYB (Fuglerud et al.,

2017). We found that lowered ATF4 activity reactivates globin synthesis, and the formation of HbF could reduce the burden of accumulating α-chain aggregates to allow coordination of cofactors and a return to homeostasis. Our findings demonstrate that the HRI-eIF2aP-ATF4 integrated stress response is involved in adaptation to loss of *HBB* and that down-regulation of the ISR induces embryonic and fetal globin gene transcription. Down-regulation of ISR by stress recovery has also been found to increase the translational efficiency of β-globin and γ-globin. (Hahn and Lowrey, 2014b). Taken together, the ISR works cooperatively to increase β-like globin production during β_0-stress.

The dramatic increase in γ-globin caused by *ex vivo* reduction in *HBB* raises the question of why β-thalassemia phenotypes are not always ameliorated by ATF4-mediated globin compensation. β-thalassemia patients can express high levels of HbF over long periods of time (Galanello et al., 1989; Manca and Masala, 2008; Rochette et al., 1994), but this phenomenon is not ubiquitous. Mouse models have indicated that the HRI-eIF2a-ATF4 axis is protective against some forms of β-thalassemia, at least in part by upregulating fetal globin (Han et al., 2005; Suragani et al., 2012). This is at odds with observations that knockout of *HRI* is protective against sickle cell anemia phenotypes by inducing HbF (Grevet et al., 2018), and our own data that knockout of ATF4 or reductions in ATF4 levels downstream of *HBB* knockout both induce high levels of HbF. Since HRI-eIF2a-ATF4 is involved in the integrated stress response in all cells, reductions in this signaling axis might be negatively selected for during development or otherwise compensated in individuals with β-thalassemia. In this case, developing a cell-type specific approach to disrupting the HRI-eIF2a-ATF4 axis after birth (e.g. during *ex vivo* editing of HSPCs) could be an important avenue for the treatment of hemoglobinopathies. The molecular pathways we find at work during *ex vivo* culture could also be quite different during long term function *in vivo*. We previously observed high levels of HbF in *HBB* edited LT-HSCs recovered after 4 months of xenotransplantation in NBSGW mice (https://www.biorxiv.org/content/10.1101/432716v6), but this does not conclusively address long term function in an individual with β-thalassemia. Determining the long term roles of the HRI-eIF2a-ATF4 axis in patient HSPCs will be an important future step to further unraveling the mechanisms of stress erythropoiesis.

2.6 Methods

HUDEP-2 cell culture and differentiation
All cell culture was performed at 37°C in a humidified atmosphere containing 5% CO2. HUDEP-2 cells were cultured in a base medium of SFEM (Stemcell Technologies 9650) containing to a final concentration of dexamethasone 1uM (Sigma D4902-100MG), doxycycline 1ug/ml (Sigma D9891-1G), human stem cell factor 50ng/ml (PeproTech 300-07), erythropoietin 50ng/ml (Peprotech 100-64), and penstrept 1%. Cells were cultured at a density of 2e5 – 1e6 cells/ml. For differentiation, HUDEP-2 cells are centrifuged at 500g for 5 minutes, media is removed and replaced with differentiation media. Differentiation media consists of a base media of IMDM+Glutamax (ThermoFisher 31980030) containing to a final concentration human serum 5% (Sigma H4522-100mL), heparin 2IU/ml (Sigma H3149-25KU), insulin 10ug/ml (Sigma I2643-25mg), erythropoietin 50ng/ml (Peprotech 100-64), holo-transferrin 500ug/ml (Sigma T0665-100mg), mifepristone 1uM (Sigma M8046-100MG), and doxycyline 1ug/ml (Sigma D9891-1G). Cells are differentiated for 5 days and then harvested for analysis.

K562 cell culture
All cell culture was performed at 37°C in a humidified atmosphere containing 5% CO2. K562 cells were grown in a base media of RPMI 1640 GlutaMAX (Gibco 61870010) supplemented with 10% fetal bovine serum, 10% sodium pyruvate, and 1% penstrept. Cells were cultured at a density of 2e5 – 1e6 cells/ml. Stable expression of doxycycline-inducible ATF4 in K562 cells was generated as described previously (Gowen et al., 2015).

HEK293T cell culture
All cell culture was performed at 37°C in a humidified atmosphere containing 5% CO2. HEK293T cells were grown in a base media of DMEM supplemented with 10% fetal bovine serum, 10% sodium pyruvate, and 1% penstrept. Cells were cultured at a density of 2e5 – 1e6 cells/ml.

Primary Cell Cultures

mPB-HSPCs cell culture and differentiation
All cell culture was performed at 37°C in a humidified atmosphere containing 5% CO2. For editing for human CD34+ cells, CD34+ mobilized peripheral blood HSPCs were thawed and cultured in SFEM containing CC110 supplement (Stemcell Technologies 02697) for 2 days. CD34+ cells were then electroporated and recovered for 24 hours in SFEM with CC110. After recovery, cells were transferred into erythroid expansion media containing SFEM and erythroid expansion supplement (Stemcell Technologies 02692) for 7 days and cultured at a density of 2e5-1e6 cells/ml. The resulting early erythroblasts were harvested for ChIP-seq or transferred to differentiation media containing SFEM with 50ng/ml erythropoietin, 3% normal human serum, and 1 µM mifepristone. The resulting late erythroblasts were harvested for analysis after 5 days of differentiation.

Generation of CRISPRi HUDEP-2 cells
WT-HUDEP-2 or *HBB* ko cells were transduced with lentivirus containing the construct EF1a-dCas9-HA-BFP-KRAB-NLS (Liang et al., 2018). The cells were FACs sorted for BFP expression and single-cell cloned. Individual clones were transduced with lentivirus containing guide RNAs targeting either *CD55, CD59, or HBB*. The clones were validated for successful knockdown by flow cytometry staining for extracellular markers CD55 and CD59 or intracellularly stained for β-globin.

Lenti-viral Packaging
Lenti-viral packaging of all constructs was performed using HEK-293T cells. TransIT®-LT1 Transfection Reagent (Mirus) was used following manufacturer's instructions. The plasmid mixture contained 50% construct plasmid, 40%DVPR, and 10% VSVG. Viral supernatant was harvested after 48 and 72 hours and filtered through 0.45 µM. For transduction of HUDEP-2 cells, cells were cultured in 50% HUDEP-2 media and 50% viral supernatant for 24 hours.

sgRNA Plasmid Cloning
sgRNA guide sequences for CRISPRi transcriptional repression were obtained from the Weissman CRISPRi-v2 library (Horlbeck et al., 2016). The chosen guides were cloned into pGL1-library vector (Addgene 84832). All guides used are listed in Supplementary Table 2.

Cas9 RNP Nucleofection
Cas9 RNP was performed as described previously (Lingeman et al., 2017). Briefly, IVT guides are purified and complexed with purified Cas9-NLS protein. The nucleofection was performed using Lonza 4D-Nucleofector and using the P3 Primary Cell 96-well NucleofectorTM Kit (V4SP-3096) following manufacturer's instructions. The HUDEP-2 nucleofector code used was DD-100 and for primary HSPCs ER-100.

IVT sgRNA
Guide RNAs were in vitro transcribed as described previously (Lingeman et al., 2017). Briefly, guide sequences were ordered as oligonucleotides and formed into duplexes using a PCR thermocycler. The DNA template was transcribed to RNA using HiScribe™ T7 High Yield RNA Synthesis Kit (E2040S) following manufacturer protocol. The resulting RNA was purified using RNeasy Mini kit (74104) and Rnase-Free DnaseI Kit (79254)

High Pressure Liquid Chromatography and Mass Spectrometry
WT-HUDEP-2 and HBBko cells were differentiated and harvested for lysis in hemolysate reagent containing 0.005M EDTA and 0.07% KCN at 10,000 cells per microliter. The lysis was incubated at room temperature for ten minutes and then centrifuged at max speed for 5 minutes. The supernatant was collected and run on Agilent 1260 Infinity II using a PolyCAT A column, 35x4.6mm (3µm;1500Å) Serial# B19916E; Lot# 16-133-3 3.54CT0315. The following Buffer compositions were used:

Mobile Phase A: 20mM Bis-tris, 2mM NaCN pH 6.8 and Mobile Phase B: 20mM Bis-tris, 2mM NaCN, 200mM NaCl, pH 6.9. The following flow settings were used: Gradient: 0-8' 2-25% Phase B, 8-18' 25-100% Phase B, 18-23' 100-2% Mobile Phase B using a Flow Rate: 1.5mL/min and measuring detection of 415nm Diode Array. Three fractions from the HBBko and one fraction from the WT samples were collected and were then processed by the Proteomics group of Functional Genomics Center Zurich (FGCZ) for Proteolytic digestion ZipTip and analysis by LC/MS. Briefly, 100ul of samples were digested with trypsin (5 ng/ul in 10 mM Tris/2 mM CaCl2, pH 8.2) and 2 ul buffer (10 mM Tris/2 mM CaCl2, pH 8.2). Samples were microwaved for 30 minutes at 60C. The samples were dried, dissolved in 20 ul 0.1% TFA and subjected to C18 ZipTip desalting. The eluted sample (10ul of 50% CAN, 0.1%TFA) was dried, dissolved in 20ul of 0.1% FA and transferred to autosampler vials for LC/MS/MS and 1ul was used for injection.

Intracellular FACs Staining

Staining was performed as adapted from the existing methods (Chung et al., 2019b). Briefly, undifferentiated or differentiated HUDEP-2 cells were centrifuged at 500g for 5 minutes. The cells were washed with PBS with 0.1% BSA, re-centrifuged, and fixed in 0.05% glutaraldehyde. The fixed cells were centrifuged and washed and re-suspended in PBS with 0.1% BSA and 0.1% Triton-X 100 for permeabilization. The fixed and permeabilized cells were then centrifuged and washed. Antibodies were diluted in PBS with 0.1% BSA and incubated with cells for 20 minutes. The following antibody dilutions were used: Human Fetal Hemoglobin APC (Thermo Fisher MHFH05) 1:10, Anti-Human Fetal Hemoglobin FITC 1:10 (BD Pharmigen 552829), and Anti-Hemoglobin B-(37-8) PE/FITC 1:100 (Santa Cruz Biotechnology SC-21757 PE, SC-21757 FITC). After staining, the cells were centrifuged and washed twice before analysis by flow cytometry.

RNA-Seq

WT-HUDEP-2 and HBBko cells were cultured and differentiated in triplicates. Cells were harvested and RNA was extracted using the RNeasy Mini kit 74104, Rnase-Free DnaseI Kit 79254 and Qiashredder 79654. RNA concentrations were quantified using Qubit™ RNA BR Assay (ThermoFisher) and 500ng were used for library preparation. RNA-seq library was made using Illumina TruSeq RNA Library Prep Kit v2 and following manufacturer's instructions. Paired-end 150bp reads were generated on a HiSeq4000. Reads for all eighteen samples (three replicates of HBBko and three WT HUDEP-2 samples at three time-points) were quantified using kallisto (Bray et al., 2016) and the hg38 index to assign reads to transcripts. Differential analysis of transcript abundance and consolidation of individual transcripts to gene-level abundance was calculated using sleuth (Pimentel et al., 2017).

qRT-PCR

RNA was harvested from cells using Qiagen RNeasy Mini Kit and Rnase-Free DnaseI Kit following manufacturer's instructions. RNA was reverse transcribed to cDNA using Iscript™ Reverse Transcription Supermix (BioRad) and qRT-PCR reactions were set up using SsoAdvanced Universal SYBR Green or SsoFast™ EvaGreen Supermix (BioRad). Reactions were run on the StepOne Plus Real-Time PCR System (Applied Biosystems) or the QuantStudio 6 Flex (Thermo Fisher). Samples were analyzed using

a two-step amplification and melt curves were obtained after 40 cycles. The Ct values for genes of interest were normalized to GAPDH, and expressions of genes are represented as 2-[ΔCt] or 2-[ΔΔCt] for fold change over control condition. All primers used for qRT-PCR are listed in Table S2.

ChIP-qPCR

ChIP was performed as described previously (Wienert et al., 2019). Briefly, 10 million cells per sample were harvested and cross-linked in 1% Formaldehyde. Cross-linking was quenched with the addition of 1.5M glycine. Samples were then lysed for 10 minutes at 4C in 50 mM Hepes–KOH, pH 7.5; 140 mM NaCl; 1 mM EDTA; 10% glycerol; 0.5% NP-40 or Igepal CA-630; 0.25% Triton X-100. Cells were then centrifuged at 1500g for 3 minutes and the supernatant was discarded. The pellet was resuspended in 10 mM Tris–HCl, pH8.0; 200 mM NaCl; 1 mM EDTA; 0.5 mM EGTA and incubated for 5 minutes at 4C. The cells were then centrifuged at 1500g for 3 minutes and the supernatant was discarded. The pellet was resuspended in 10 mM Tris–HCl, pH 8; 100 mM NaCl; 1 mM EDTA; 0.5 mM EGTA; 0.1% Na–Deoxycholate; 0.5% N-lauroylsarcosine and sonicated using the Covaris S220 following manufacturer's instructions. Protein A beads (ThermoFisher) were complexed with antibody and the antibody-bead complexes were incubated with cell lysates at 4C overnight with rotation. The antibodies used were rabbit anti-ATF4 (CST 11815S) and rabbit IgG (Novus Biologicals NBP2-24891). The beads were retrieved using a magnetic stand and rinsed with RIPA buffer. Elution buffer containing 50 mM Tris–HCl, pH 8; 10 mM EDTA; 1% SDS was added to the beads for reverse crosslinking at 65C overnight with shaking. After reverse crosslinking, the beads were removed. The eluted DNA was treated with RNaseA and Proteinase K and then purified using Qiagen MinElute PCR Purification Kit, following the manufacturer's instructions. Q-PCR reactions were set up using SsoAdvanced Universal SYBR Green or SsoFast™ EvaGreen Supermix (BioRad). Reactions were run on the StepOne Plus Real-Time PCR System (Applied Biosystems) or the QuantStudio 6 Flex (Thermo Fisher). The Ct values were analyzed by the enrichment compared to input method.

ChIP-seq

ChIP was performed as described in ChIP-qPCR method section. Sequencing library was prepared using NEBNext Ultra II DNA Library Prep Kit for Illumina (E7647) and NEBNext Multiplex Oligos for Illumina (Dual Index Primers Set 1) following manufacturer's instructions.

Paired-end 150bp reads were generated on an Illumina NextSeq500 at the Functional Genomics Center Zürich (FGCZ) and demultiplexed. FastQC (Andrews S. (2010) A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used for initial quality control of reads. All samples, WT diff and undiff, HBBko diff and undiff, HSPC diff and undiff, ATF4ΔN diff, ATF4KO3, and IgG and HSPC IgG diff and undiff, were processed according to ENCODE guidelines for unreplicated transcription factor ChIP-seq analysis (Landt et al., 2012). In detail, raw reads were aligned against GRCh38 using bowtie2. Duplicate reads were marked using Picard's MarkDuplicates and multimapping, low

quality, duplicated and non-properly paired reads were removed. Library complexity measures and flagstats were generated for each BAM file. BAM files were converted to tagAlign format and two subsampled pseudoreplicates were generated for each sample with half the total reads. Peak calling, fold change and p-value signal tracks were generated using MACS2 (Zhang et al., 2008). Irreproducible Discovery Rate (IDR) analysis was performed using self-pseudoreplicates and the main samples to obtain self-consistent sets of peaks. Final peak calls were filtered by ENCODE blacklist (Amemiya et al., 2019) and an IDR of 2% and a signal value > 30.
Sets of peaks for each comparison were analysed and associated to genes using the R package ChIPseeker (Yu et al., 2015) and Bioconductor hg38 TxDb (Team BC, Maintainer BP (2019). TxDb.Hsapiens.UCSC.hg38.knownGene: Annotation package for TxDb object(s). R package version 3.4.6.). ChIP-seq peaks and RNA-seq results were merged using HGNC symbols using custom scripts. Figures were generated using the Integrative Genome Viewer (Robinson et al., 2011) and the UCSC Genome Brower (Kent et al., 2002). MEME-ChIP v5.1.1 (Machanick and Bailey, 2011) was used for motif discovery using peak boundaries +-250bp. FIMO v5.1.1 (Grant et al., 2011) was used to identify motif occurences for ATF4 (JASPAR matrix ID: MA0833.1) and GATA1 (JASPAR matrix ID: MA0035.4) with a p-value less than 0.0001.

Western Blot

Cells were harvested and lysed using RIPA buffer (Millipore) supplemented with Halt Protease Inhibitor Cocktail (ThermoFisher) and Phosphatase Inhibitor (ThermoFisher). Cell lysate concentrations were measured using Bradford reagent (VWR) or BCA assay (ThermoFisher). Cell lysates were normalized using NuPage LDS 4x Sample Buffer (Invitrogen) and samples were run on NU-PAGE Noves Bis-Tris 4–12% gels (Invitrogen) in NuPAGE™ MES SDS Running Buffer at 180V for 40 minutes. Protein gels were transferred to a 0.4-µm nitrocellulose membranes at 1.3 A and 25 V for 15 to 20 minutes in a semi-dry apparatus (BioRad). After protein transfer, membranes were blocked in Tris-buffered saline with 1% Tween-20 (TBS-T) containing 5% nonfat milk for 30 minutes at room temperature with rocking. Primary antibodies were diluted in TBST containing 5%BSA and 0.1% sodium azide. The following antibody concentrations were used: rabbit anti-ATF4 (CST 11815S) 1:1000, mouse anti-BCL11A (Abcam ab19489) 1:1000, rabbit anti-LRF (ThermoFisher PA528144) 1:1000, mouse anti-GAPDH D4C6R (CST 97166S) 1:1000, Rabbit anti-GAPDH 14C10 (CST 2118S) 1:1000, Mouse anti-B-Globin (SCBT sc-21757) 1:500, Rabbit anti-Fetal-Globin (Abcam ab137096) 1:1000, Rabbit anti-MYB D2R4Y (CST 12319S) 1:1000. Membranes were rinsed in TBST and incubated with antibody for 2 hours at room temperature or 4C overnight with rocking. After incubation, membranes are rinsed twice with TBST for 5 minutes. Secondary antibodies are diluted in 1% Tween-20 (TBS-T) containing 5% nonfat milk and incubated at room temperature for 30 minutes with rocking. The secondary antibodies used for Western blotting were obtained from Li-Cor and are as follows: donkey anti-mouse IRDye 680CW (926-32222), donkey anti-mouse IRDye 800CW (926-32212), donkey anti-rabbit IRDye 680CW (926-32223), and donkey anti-rabbit IRDye 800CW (926-32213). Membranes are then rinsed twice with TBST for 5 minutes each. Membranes are rinsed with PBS for 5 minutes and then imaged using Li-Cor Odyssey CLx.

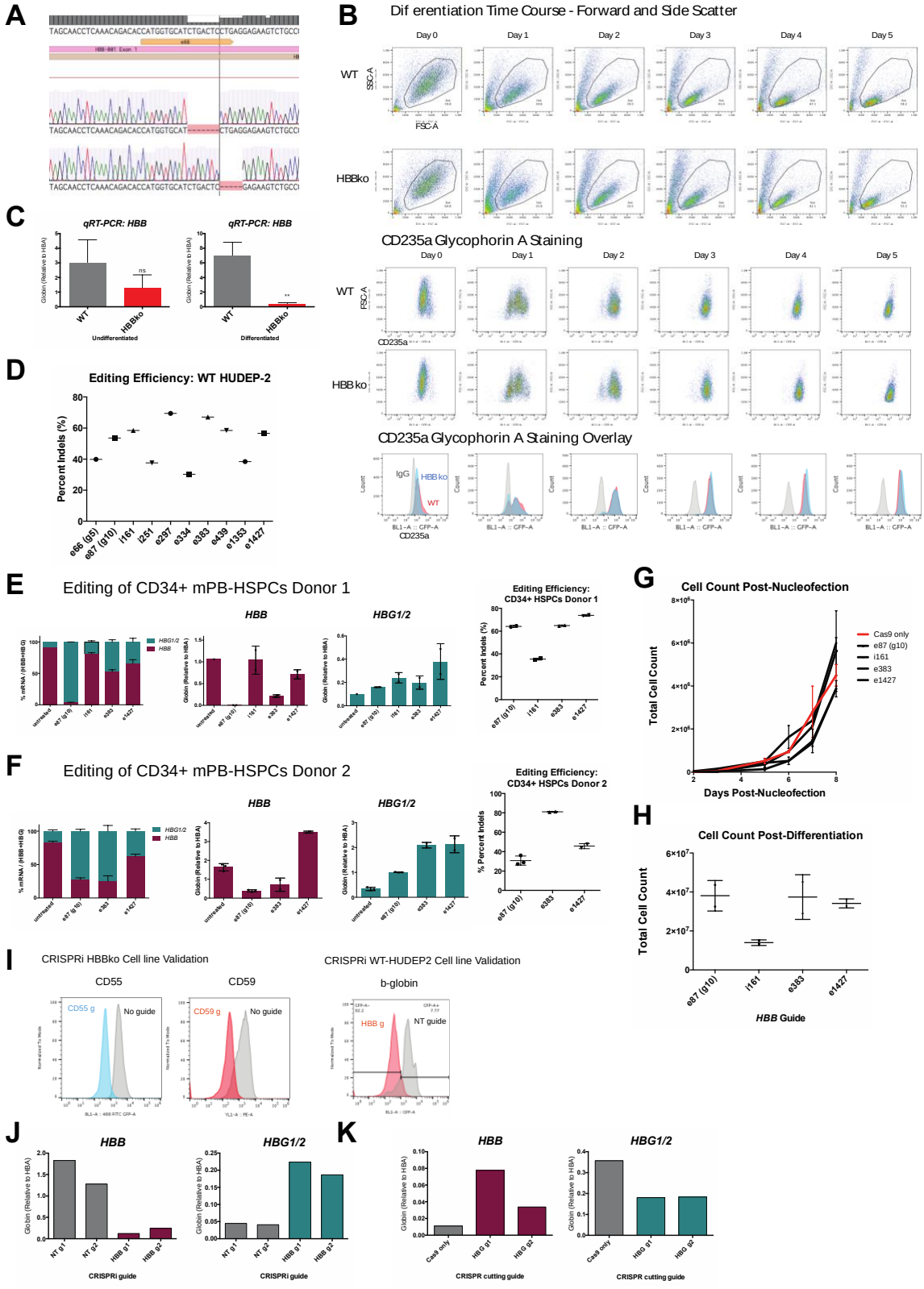
QUANTIFICATION AND STATISTICAL ANALYSIS

All analysis was performed using data from at least three independent biological replicates (exact number of replicates are stated in the figure legend). All statistical analyses were performed in PRISM6 software using paired Student's t test. P values are indicated as follow: * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001. The distribution of the data was assumed to be normal, but this was not formally tested.

2.7 Supplemental Materials

2.7.1 Supplemental Figures

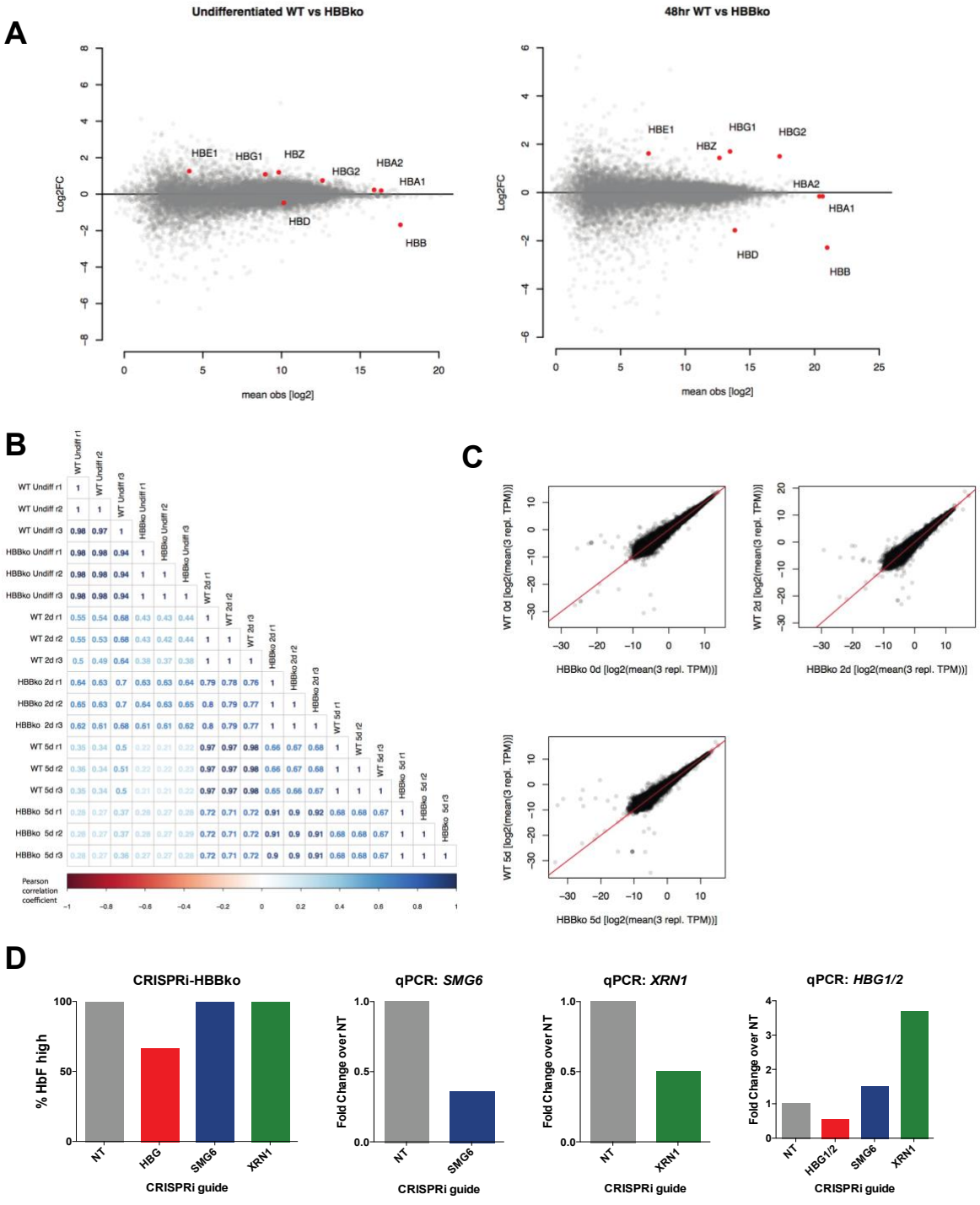
Supplement Fig 1



Supplemental Figure 1 (related to Figure 1)

A) Schematic showing the e66 (g5) guide location in *HBB*, with the cut site indicated by a black line. The HBBko clone was genotyped by TA cloning of the *HBB* locus followed by Sanger sequencing. Two alleles were found: a 5- and 7-basepair deletion.
B) Flow cytometry plots showing forward and side scattering of WT and HBBko over the course of 5 days in differentiation media. Cells were stained for CD235a (Glycophorin A) and shown as dot plots and overlaid histograms to compare WT to HBBko during differentiation in red and blue, respectively. IgG isotype control trace is in gray.
C) qRT-PCR of *HBB* in undifferentiated and differentiated cells comparing transcript expression levels between WT and HBBko cells. P value indicates paired, two-tailed student *t* test (ns, non-significant; *, $P \leq 0.005$).
D) Editing efficiency of the panel of *HBB* guides in HUDEP-2 cells as measured by TIDE analysis.
E) Mobilized peripheral blood CD34+ cells from a healthy donor were edited with a subset of the *HBB* guides, recovered in erythroid expansion media, differentiated for 5 days, and analyzed for qRT-PCR for *HBB* and *HBG 1/2* for (n=2) technical replicates for one healthy donor. Editing efficiency was measured by TIDE.
F) Data from healthy donor 2.
G) Total cell counts of edited CD34+ cells during recovery in erythroid expansion media. Data is shown as the mean ± SEM of (n=2) technical replicates for one healthy donor.
H) Total cell counts of edited CD34+ cells after 5 days of differentiation. Data is shown as the mean ± SD of (n=2) technical replicates for one healthy donor.
I) HBBko CRISPRi cell line was validated using guides targeting *CD55* and *CD59*. FACs staining was used for validating knockdown of the cell marker. WT-CRISPRi cell line was validated using an HBB guide and differentiated and intracellularly stained for B-globin protein and compared to non-targeting guide.
J) qRT-PCR of *HBB* and *HBG1/2* for CRISPRi knockdown of *HBB* in WT differentiated cells. Data is shown for one replicate.
K) qRT-PCR of *HBB* and HBG1/2 from pooled knockout of *HBG1/2* in HUDEP-1 cells. N=2 technical replicates are shown as mean.

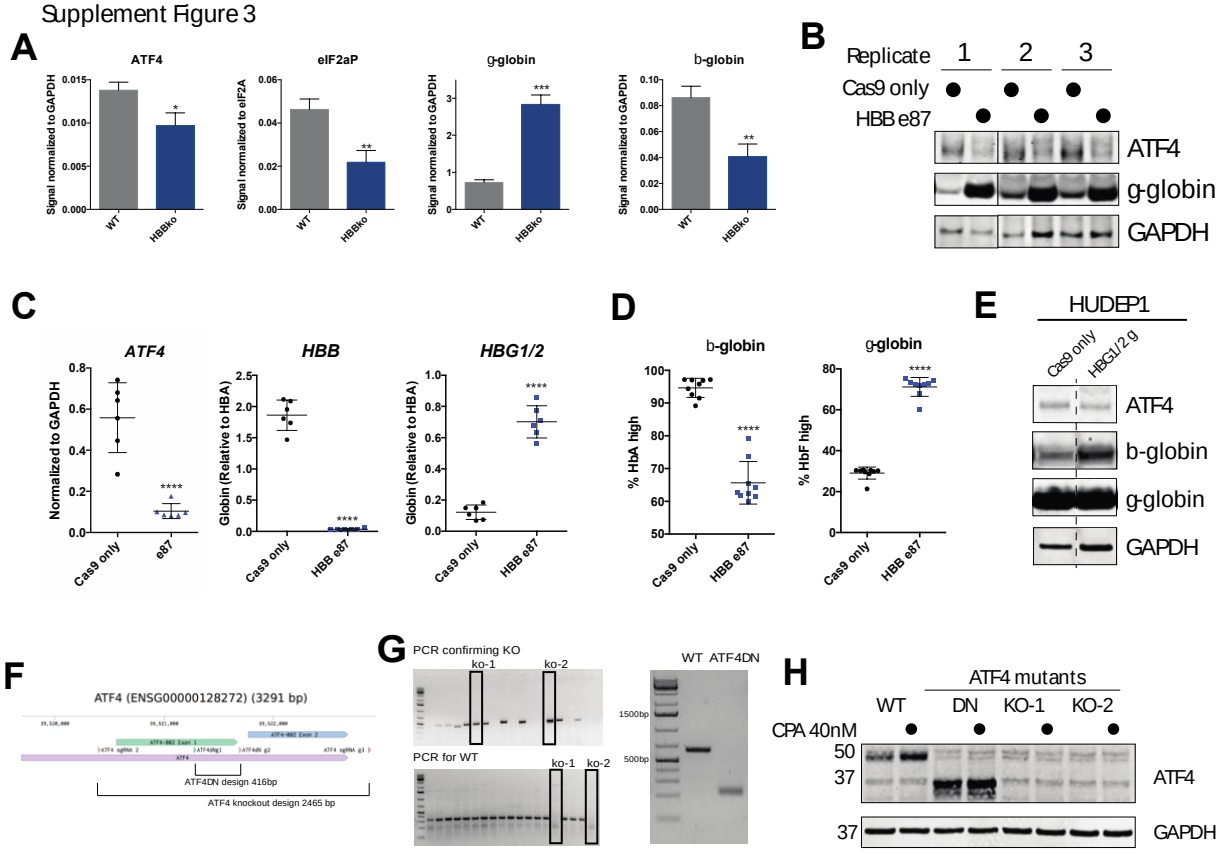
Supplemental Figure 2

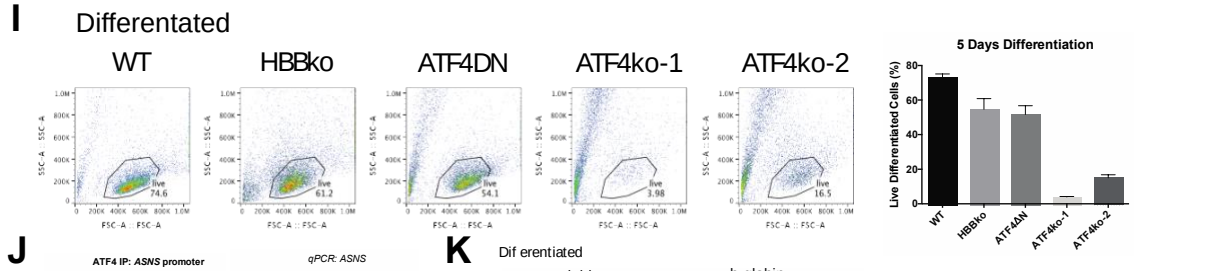


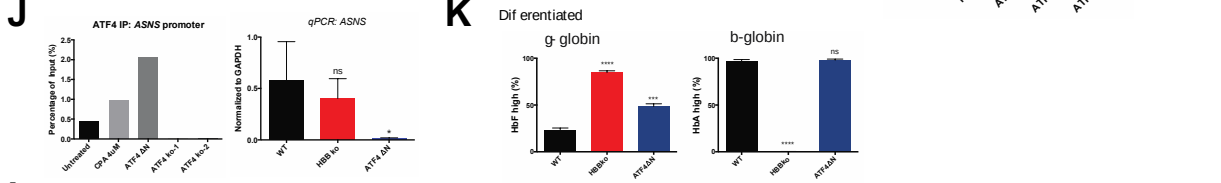
Supplemental Figure 2 (related to Figure 2)

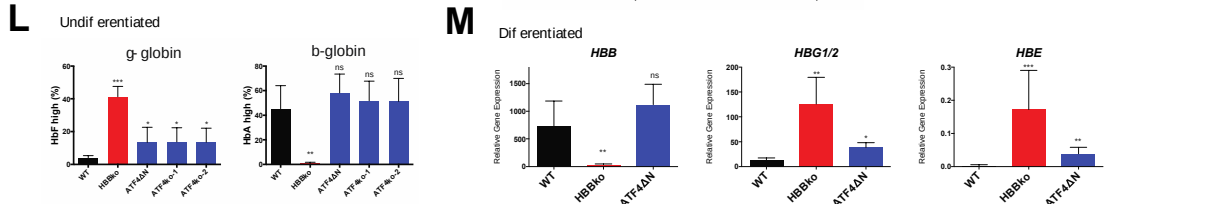
A) MA plots of RNAseq data for undifferentiated cells and 2 day differentiation, with globin genes highlighted.
B) Pearson correlation matrix of RNAseq samples, highlighting inter-condition reproducibility and differences between conditions.
C) Pearson correlation analysis of RNAseq data comparing WT to HBBko at three different time points of differentiation
D) Leftmost panel shows quantified intracellular flow staining of γ-globin for HBBko CRISPRi cells expressing either non-targeting, *HBG1/2*, *SMG6*, or *XRN1* guide. Right three panels show qRT-PCR for *SMG6*, *XRN1*, or *HBG1/2* for knockdown validation.

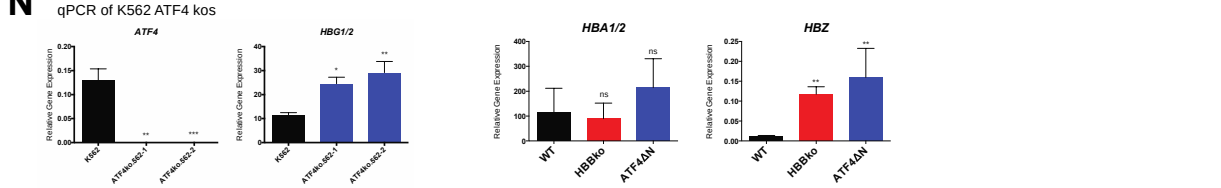
Supplement Figure 3











Supplemental Figure 3 (related to Figure 3)

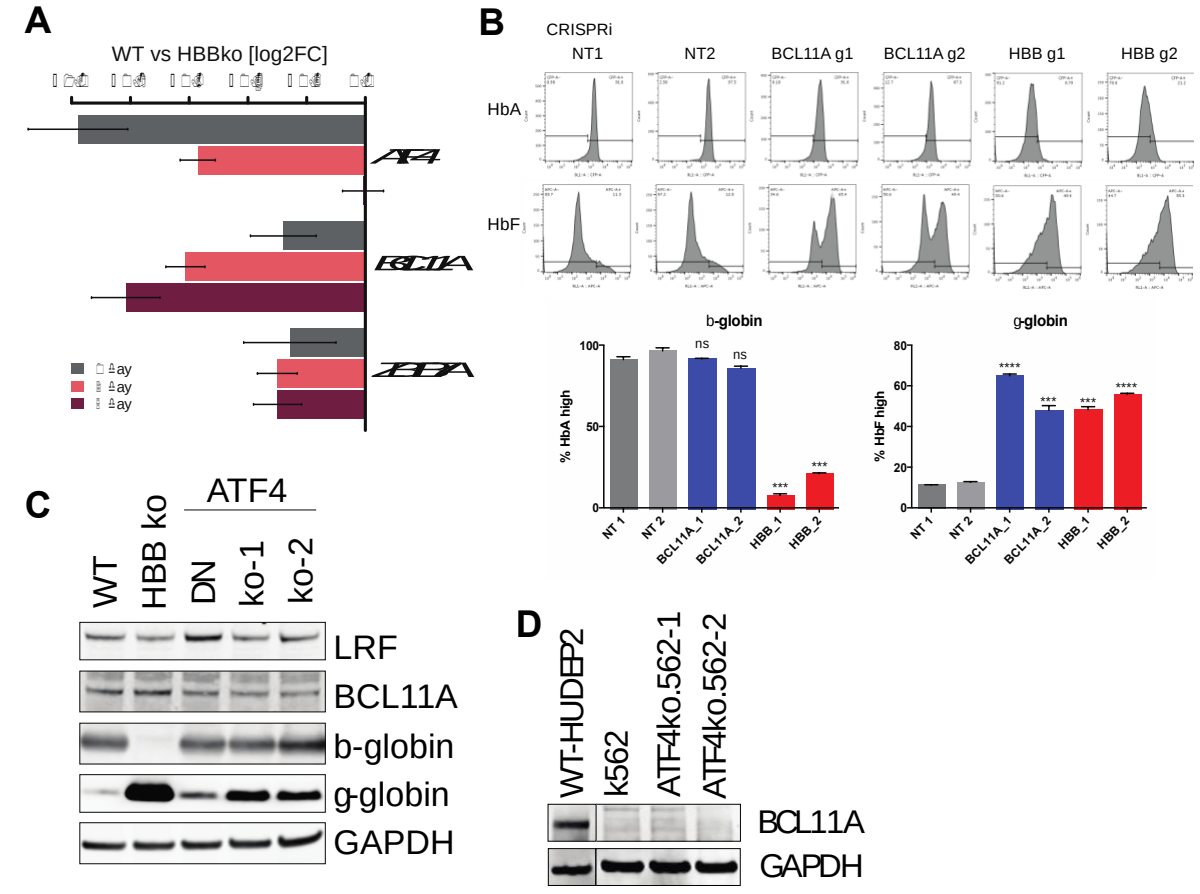
A) Quantification of western blot from **Fig 3A**. The data is shown as the mean ± SD of 3 biological replicates. P value indicates paired, two-tailed student *t* test (*, P ≤ 0.05; **, P ≤ 0.005; ***, P ≤ 0.001).
B) HUDEP-2 cells were pool-edited with *HBB* guide e87 (g10) and differentiated for 5 days. Western blot shows ATF4 and γ-globin. The image shows 3 biological replicates and is from the same blot with the solid line indicating irrelevant conditions removed.
C) qRT-PCR of *ATF4, HBB*, and *HBG1/2* from *HBB* pooled editing and 5 day differentiated HUDEP-2 cells from A). Data is shown as six biological replicates. P value indicates unpaired, two-tailed student *t* test (****, P ≤ 0.0001).
D) Intracellular Flow cytometry staining for HbA and HbF from pooled editing and 5 day differentiated HUDEP-2 cells from A). P value indicates unpaired, two-tailed student *t* test (****, P ≤ 0.0001). Data is shown as six biological replicates.
E) HUDEP-1 cells were edited with guide targeting *HBG1/2* and differentiated for 5 days. Western blot shows ATF4, β-globin, and γ-globin. The image is from the same blot with the dotted line indicating irrelevant conditions removed.
F) Schematic of ATF4 with sgRNAs used for generation of ATF4 knockouts and ATF4ΔN. See **Table S2** for all guide sequences.
G) PCR genotyping of ATF4ko-1 and ATF4ko-2 clones. Top panel shows PCR amplification using primers outside the cut region. Successful *ATF4* deletion is indicated by ~500 base pair fragment. Unsuccessful deletion results in failure of PCR amplification due to the large size. Lower panel shows PCR amplification of a region within the *ATF4* gene. Successful deletion of *ATF4* is indicated by no amplification. Right panel: PCR genotyping of ATF4ΔN using primers flanking the ATF4ΔN cut sites. TA-cloning and Sanger sequencing revealed one 398- and one 415-basepair deletion for the ATF4ΔN clone.
H) Western blotting for ATF4 of WT and ATF4 mutant clones treated with 40nM cyclopiazonic acid (CPA) for 16 hours. ATF4ΔN removes the amino terminal regulatory region but retains the C-terminal DNA binding domain.
I) Flow cytometry plots of forward and side scatter 5 days post differentiation of WT, HBBko, ATF4ΔN, ATF4ko-1, and ATF4ko-2. Right panel shows percentages of differentiated cells as the mean ± SD of 3 biological replicates.
J) ATF4 CHIP-qPCR for the *ASNS* promoter region shows ATF4 binding at WT and ATF4ΔN and loss of binding for ATF4ko-1 and ATF4ko-2 in undifferentiated cells. The data is shown as one replicate. Right panel shows qRT-PCR on differentiated cells for *ASNS*. The data is shown as the mean ± SD of 3 biological replicates. P value indicates paired, two-tailed student *t* test (ns, non-significant; *, P ≤ 0.05).
K) Quantification of intracellular FACs staining of differentiated WT, HBBko, and ATF4ΔN cells. Data is presented as mean ± SD of 3 biological replicates. P value indicates paired, two-tailed student *t* test (ns, non-significant; ***, P ≤ 0.001; ****, P ≤ 0.0001).
L) Quantification of intracellular FACs staining of undifferentiated WT, HBBko, ATF4ΔN, and ATF4ko cells. Data is presented as mean ± SD of 4 biological

replicates. P value indicates paired, two-tailed student *t* test (ns, non-significant; *, P ≤ 0.05; **, P ≤ 0.005).

M) qRT-PCR of globin gene expression relative to *GAPDH* for differentiated WT, HBBko, and ATF4ΔN. Data is presented as mean ± SD of 3 biological replicates. P value indicates paired, two-tailed student *t* test (ns, non-significant; *, P ≤ 0.05; **, P ≤ 0.005; ***, P ≤ 0.001)

N) qRT-PCR of *ATF4* and *HBG1/2* relative to *GAPDH* in K562 WT cells and two K562 ATF4 knockout clones. Data is presented as mean ± SD of 3 biological replicates. P value indicates paired, two-tailed student *t* test (*, P ≤ 0.05; **, P ≤ 0.005; ***, P ≤ 0.001).

Supplement Fig 4



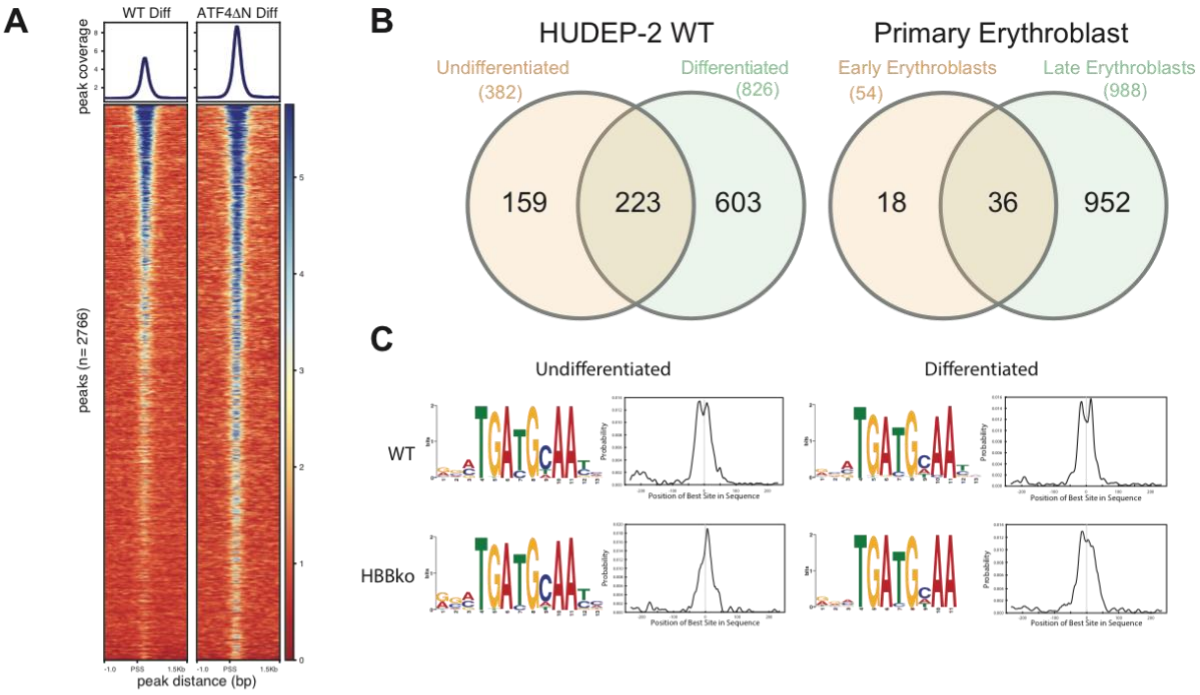
Supplemental Figure 4 (related to Figure 4 and 5)

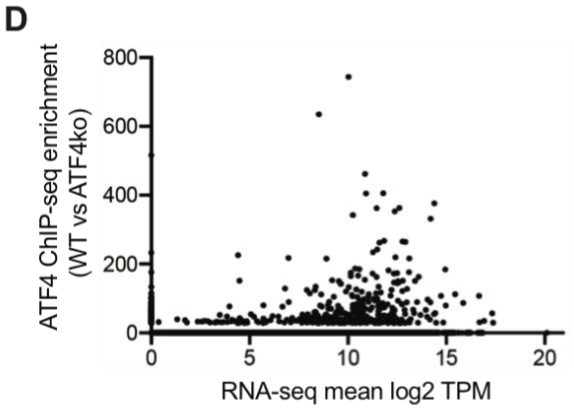
A) *ATF4, BCL11A*, and *ZBTB7A* transcript levels in HBBko HUDEP-2 cells are shown as log2 fold change relative to WT cells undifferentiated, after 2 days of differentiation, and after 5 days of differentiation.

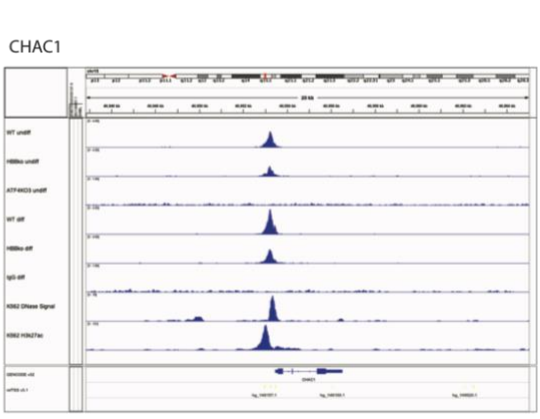
B) Top panel shows intracellular FACs staining of differentiated cells for HbA and HbF in WT CRISPRi cells with guides against non-targeting, *BCL11A*, and *HBB*. Lower panel shows quantified data as mean ± SD of 3 biological replicates. P value indicates paired, two-tailed student *t* test (ns, non-significant; ***, P ≤ 0.001; ****, P ≤ 0.0001).

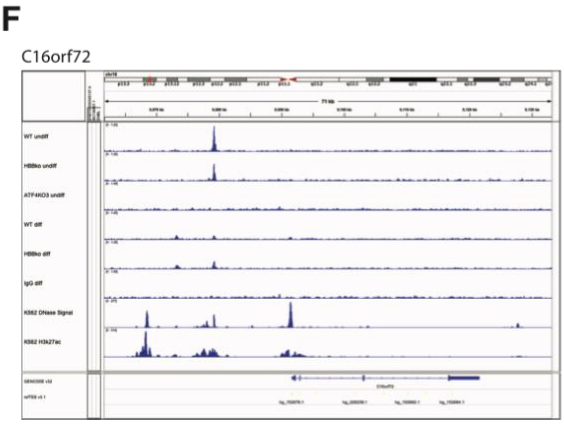
C) Western blot of undifferentiated WT, HBBko, ATF4ΔN, and ATF4ko cells.
D) Western blot of K562 WT, ATF4ko.562-1 and ATF4ko.562-2 for BCL11A expression. Image is from the same blot with the black line indicating irrelevant conditions removed.

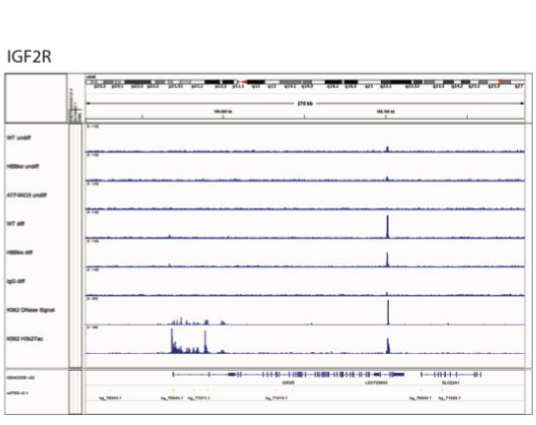
Supplement Fig 5







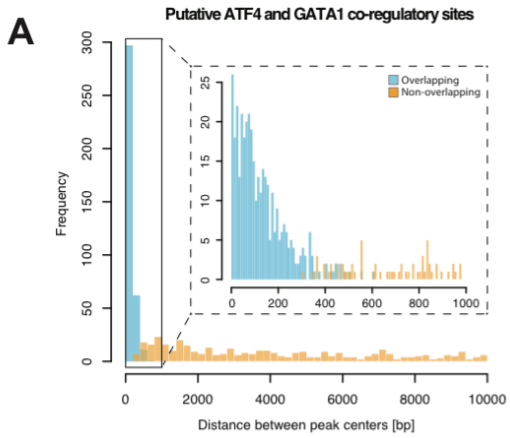


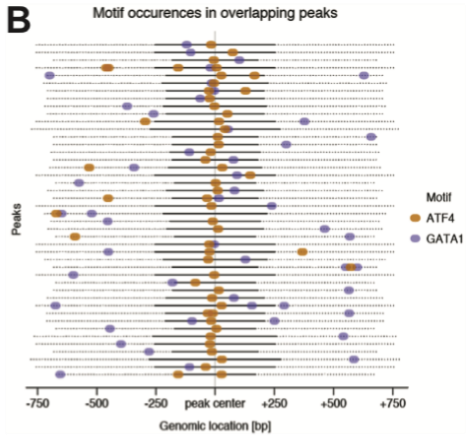


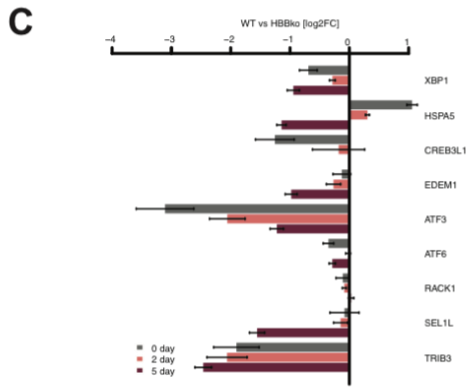
Supplemental Figure 5 (related to Figure 5)

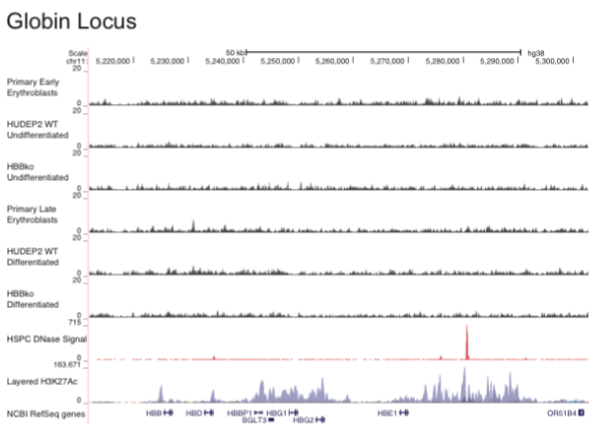
A) Heatmap showing ATF4 ChIP-seq fold change over control for all peaks across samples. ChIP-seq samples shown are differentiated WT and ATF4DN.
B) Venn diagrams showing overlap of high confidence ATF4 peaks between WT undifferentiated and differentiated (left) and between primary early erythroblasts and late erythroblasts (right).
C) Unbiased motif discovery (MEME) for ChIP-seq samples WT and HBBko undifferentiated and differentiated for identification of the consensus sequence.
D) ATF4 ChIP-seq signal enrichment in WT HUDEP-2 cells and corresponding RNA-seq log2 TPM values.
E) *CHAC1* raw read coverage for ATF4 ChIP-seq. Tracks show library size normalized tags per million (TPM) coverage of WT undiff and diff, HBBko undiff and diff, ATF4KO3 undiff and IgG data. DNase signal and H3k27ac tracks show reference ENCODE data from K562 cells. Annotation tracks show GENCODE v32 gene annotation and transcription start site annotation from refTSS v3.1
F) *C16orf72* raw read coverage for ATF4 ChIP-seq.
G) *IGF2R1* raw read coverage for ATF4 ChIP-seq.

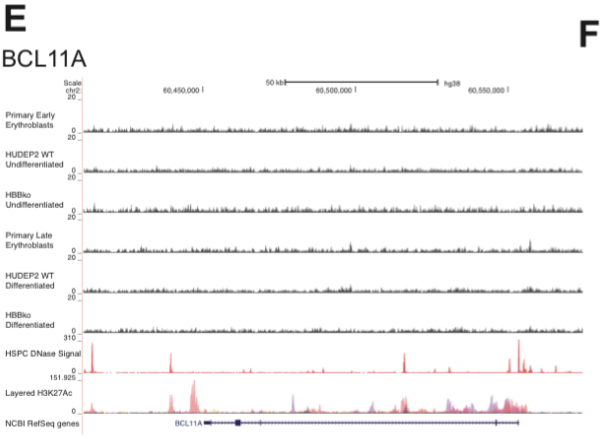
Supplement Fig 6

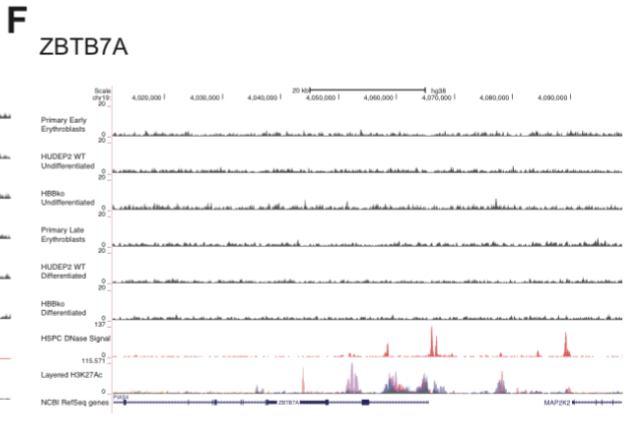


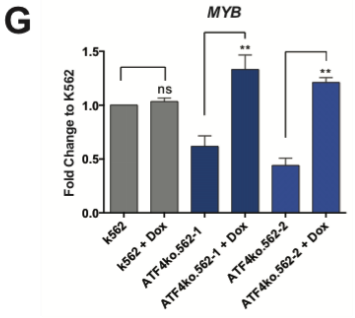












Supplemental Figure 6 (related to Figure 5 and 6)

A) Distances between ATF4 and GATA1 peaks annotated as binding at the same gene. ATF4 peaks with a GATA1 peak within 10kb (n=753, outer panel) or within 1kb (n=434, inner panel) between peak centers are shown. Distances for overlapping ATF4 and GATA1 peaks are shown in blue (n=371), non-overlapping peaks that are still adjacent to the same gene are shown in orange (n=382).
B) Motif occurrences in overlapping ATF4 and GATA1 peaks. Centered ATF4 peak regions are shown (black solid lines) with 500bp extensions on both sides (dotted lines) for 44 ATF4 peaks that overlapped with GATA1 peaks. FIMO v5.1.1 was used to identify motif occurrences for ATF4 (JASPAR matrix ID: MA0833.1) and GATA1 (JASPAR matrix ID: MA0035.4) with a p-value less than 0.0001. ATF4 motif occurrences are shown in brown, GATA1 motif occurrences in blue.
C) Transcript levels of UPR targets in HBBko HUDEP-2 cells are shown as log2 fold change relative to WT cells undifferentiated, after 2 days of differentiation, and after 5 days of differentiation.
D) ATF4 ChIP-seq enrichment tracks for the globin locus. DNAse signal is from previously published primary human CD34+ cells (Neph et al., 2012). H3k27ac tracks show reference ENCODE data from K562 cells.
E) ATF4 ChIP-seq fold enrichment tracks for the *BCL11A* locus.
F) ATF4 ChIP-seq fold enrichment tracks for the *ZBTB7A* locus.
G) qRT-PCR for *MYB* expression in K562 and ATF4ko.562-1/2 with doxycycline-inducible ATF4 constructs treated with 100ng/ml doxycycline for 48 hours. The data is presented as mean ± SD of three biological replicates. P value indicates unpaired, two-tailed student *t* test (ns, non-significant; **, P 0.005).

2.7.2 Supplemental Tables

Supplemental Table 1: Mass Spectrometry Spectrum Counts from HPLC elutions, Related to Figure 1. Peaks 1, 2, and 3 are elutions from differentiated HBBko cells at 5 minutes, 7 minutes, and 7.5 minutes, respectively. Peak 4 is the elution from differentiated WT cells at 10 minutes.

Alternate ID	Molecular Weight	Peak 1	Peak 2	Peak 3	Peak 4
HBA1	15 kDa	31	88	108	224
HBB	16 kDa	8	15	20	278
HBB	16 kDa	4	7	6	186
HBD	16 kDa	2	4	7	7
HBG2	16 kDa	43	146	79	10
HBG2	16 kDa	9	30	14	2
HBG1	16 kDa	8	18	10	1
GPI	63 kDa	0	0	12	39
GOT2	48 kDa	0	10	23	2
TUBA1A	50 kDa	13	10	6	3
RACK1	35 kDa	3	2	14	12
ENO1	47 kDa	0	5	9	17
ENO1	47 kDa	0	4	7	15
ENO2	47 kDa	0	0	0	1
GMPS	77 kDa	0	16	12	0
PPIA	18 kDa	14	12	2	0
CA2	29 kDa	13	5	4	5
PEBP1	21 kDa	19	5	3	0
FSCN1	55 kDa	0	2	12	12
RPL7	29 kDa	9	10	3	2
TUBB	50 kDa	8	7	5	4
TUBB	50 kDa	2	2	1	1
TUBB4B	50 kDa	1	1	1	0
EEF1A1	50 kDa	9	6	6	2
HSPA8	71 kDa	7	6	6	3
PSMA7	28 kDa	4	7	4	7
MDH1	36 kDa	13	6	2	0
CFL1	19 kDa	9	4	5	2
FH	55 kDa	7	10	2	0
RPL6	33 kDa	10	7	2	0
UROD	41 kDa	0	5	13	0
HIST1H4A	11 kDa	6	7	3	2
GARS	83 kDa	0	2	0	15
TKT	68 kDa	0	0	0	17

Supplemental Table 2: CRISPR/Cas9 and CRISPRi sgRNA sequences used in this study, Related to methods sections on sgRNA Plasmid Cloning and IVT sgRNA.

Construct	Protospacer	Source
pLG1-puro HBB g1 sgRNA	TAGACCACCAGCAGCCTAA	Horlbeck et al. 2016
pLG1-puro HBB g2 sgRNA	GAACTTCAGGGTGAGTCTA	Horlbeck et al. 2016
pLG1-puro Non-targeting 1 sgRNA	CGCCAAACGTGCCCTGACGG	Horlbeck et al. 2016
pLG1-puro Non-targeting 2 sgRNA	GCTCGGTCCCGCGTCGTCG	Horlbeck et al. 2016
pLG1-puro BCL11A g1 sgRNA	GCTTGCGGCGAGACATGGT	Horlbeck et al. 2016
pLG1-puro BCL11A g2 sgRNA	GAGAGCCGGGTTAGAAAGA	Horlbeck et al. 2016
pLG1-puro SMG6 sgRNA	ctcctccccgctcggccct	Horlbeck et al. 2016
pLG1-puro XRN1 sgRNA	CTCGAAAGCCCCAGCTCTA	Horlbeck et al. 2016
pLG1-puro CD55 sgRNA	GCTGCGACTCGGCGGAGTCC	Horlbeck et al. 2016
pLG1-puro CD59 sgRNA	GCGCAGAAGCGGCTCGAGGC	Horlbeck et al. 2016
e66 (g5)	CATGGTGCATCTGACTCCTG	DeWitt et al. 2016
e87 (g10)	CTTGCCCCACAGGGCAGTAA	DeWitt et al. 2016
i161	TGGTATCAAGGTTACAAGAC	This Study
i251	GGGTGGGAAAATAGACCAAT	This Study
e297	TGGTCTACCCTTGGACCCAG	This Study
e334	GCCCATAACAGCATCAGGAG	This Study
e383	GCTCATGGCAAGAAAGTGCT	This Study
e439	CAGCTCACTCAGTGTGGCAA	This Study
e1353	CACAGACCAGCACGTTGCCC	This Study
e1427	GCAGGCTGCCTATCAGAAAG	This Study
ATF4dN g1	AGTCCCGCCTCATAAGTGGA	This Study
ATF4dN g2	ATCCCACCAGGACGATG	This Study
ATF4 sgRNA 2	CTCGTCACAGCTACGCCCT	Gowen et al. 2016
ATF4 sgRNA 3	TGGCCAACTATACGGCTCCA	Gowen et al. 2016

Supplemental Table 3: Primers for qRT-PCR, ChIP-qPCR, and amplification for genotyping and knockout validations, Related to Methods sections on qRT-PCR, ChIP-qPCR, and Cas9 RNP Nucleofection.

Primers for qRT-PCR		
Target	Forward	Reverse
HBB	TGTCCACTCCTGATGCTGTTATG	GGCACCGAGCACTTTCTTG
HBG1/2	CCTGTCCTCTGCCTCTGCC	GGATTGCCAAAACGGTCAC
HBA	GGGTGGACCCGGTCAACTT	GAGGTGGGCGGCCAGGGT
HBE	TGCTGAGGAGAAGGCTGCCG	TGGGTCCAGGGGTAAACAACGAGG
HBZ	GAGGACCATCATTGTGTCCA	AGTGCGGGAAGTAGGTCTTG
GAPDH	CAACAGCGACACCCACTCCT	CACCCTGTTGCTGTAGCCAAA
SMG6	GCACGCATGGTGAACAGAAA	TGTTCAGCTCTCATTCTCGTCC
XRN1	TGGCTTACTGGTACATGGGC	TCATGCCTAAGGAGCCCAAC
ASNS	ACTGGCTGCTAGAAAGGTGG	GCCTGAATGCCTTCCTCAGA
MYB	CGCAGCCATTCAGAGACACT	GCTGCATGTGTGGTTCTGTG
Primers for ChIP-qPCR		
ASNS promoter	ATGATGAAACTTCCCGCACG	GGGATGTGGACAGCTTGACG
Primers for knockout validation and genotyping		
ATF4 WT for	CATTCCTCGATTCCAGCAAAGC	TGAGTGATGGGGCCAAGTGAG
ATF4 knockout for	CGTCCTCGGCCTTCACAATA	TCTTCAGGATGAGGCTTCTGC
ATF4DN	CCT GGT CTC CGT GAG CGT C	CATCCAATCTGTCCCGGAGAAGG
HBB exon 1-2	ATG CTT AGA ACC GAG GTA GAG TTT	CCT GAG ACT TCC ACA CTG ATG
HBB exon 3	AGGCAGAATCCAGATGCTCAAGGC	GCA CGT GGA TCC TGA GAA CTT CAG

3 Chapter 3 CRISPRi Arrayed Screens For Gene Candidates Regulating Fetal Hemoglobin

3.1 Connection to overall dissertation

The previous chapter focused on discovering candidate genes and pathways involved in the response of increasing fetal hemoglobin upon the loss of adult hemoglobin. We briefly introduced CRISPRi-HUDEP-2 cells in the previous chapter. This chapter fully utilizes the CRISPRi technology in a screen-based approach to identify other factors that increase fetal hemoglobin. To do this, we used CRISPRi in both wildtype and HBBko conditions in order to look for genes involved in upregulating fetal hemoglobin.

3.2 Summary

γ-globin, which comprises fetal hemoglobin, is silenced in humans shortly after birth. There few known regulators of fetal hemoglobin, such as BCL11A and LRF, that have been well-characterized. However, on-going research is still uncovering new regulators of γ-globin. Here, we use CRISPR interference (CRISPRi) to knockdown candidate genes in HUDEP-2 cells in order to uncover new regulators of γ-globin. This targeted screening approach reveals that in addition to ATF4, BACH1, HMOX1, and DENND4A, could also modulate levels of γ-globin. Our findings may inform future strategies of treating β-hemoglobinopathies.

3.3 Introduction

Clustered, Regularly Interspaced, Short Palindromic Repeat (CRISPR)-associated (Cas) genome editing has become a widespread tool in molecular biology. The system was first discovered in prokaryotes as a mechanism for bacteria to combat viruses and foreign plasmids (Bhaya et al., 2011; Terns and Terns, 2011) and was quickly adapted for research use (Jinek et al., 2012). The discovery of mutations that lead to a catalytically dead Cas9 (dCas9) (Qi et al., 2013) spawned the development a range of gene-editing tools. By fusing a repressive KRAB domain to dCas9, eukaryotic transcription of specific genes was able to be repressed in a technique termed CRISPR-interference (CRISPRi) (Gilbert et al., 2014; Qi et al., 2013). CRISPRi is an appealing tool as it enables the inhibition of expression of specific genes without inducing DNA double-strand breaks in the cell. This also allows for a partial knockdown phenotype instead of a knockout and is useful in cases where knockout of a gene is lethal to the cell. CRISPRi has proven to be a useful tool to uncover biological mechanisms and has since been used in multiple genome-wide screens (Kabir et al., 2019; Liang et al., 2020)

We utilized the CRISPRi technology to perform targeted screening of candidate genes involved in fetal hemoglobin regulation. We performed CRISPRi knockdown in the WT HUDEP-2 background as well as in the HBBko background. We tested about 130 genes which included genes that were differentially expressed by RNAseq comparing WT to HBBko cells, genes that were targeted by ATF4 as revealed by ATF4 ChIP-seq of HUDEP-2, and hypothesis-driven genes that may have been implicated or related to known globin regulation pathways.

Our targeted approach of knocking down candidate genes revealed that BACH1, HMOX1, LSD1, and DENND4A led to increases in fetal hemoglobin. ATF4 was explored in detail in the previous chapter and revealed to be a repressor of fetal hemoglobin by modulating levels of MYB and BCL11A. HMOX1 and LSD1 have been previously linked to fetal hemoglobin expression. Here, we show further evidence that BACH1 may also be a fetal hemoglobin repressor.

3.4 Results

3.4.1 Knockdown of differentially expressed genes between HBBko and WT HUDEP-2 cells reveals fetal hemoglobin regulators

We found that decreased β-globin is sufficient to induce robust re-expression of γ-globin. As described in the previous chapter, we generated a homozygous knockout of *HBB* (HBBko) from a wildtype HUDEP-2 clone (WT). We performed RNA-seq of these two isogenic erythroid precursors in order to compare the transcriptomic profiles at three different timepoints: day 0 (undifferentiated), day 2 (partially differentiated), and day 5 (differentiated). Using a pathway analysis tool (Krämer et al., 2014b), we discovered that ATF4 as a causal regulator of the response to increase γ-globin. Addtionally, the RNA-seq data showed many genes to be upregulated or down-regulated in the HBBko compared to WT (**Table 1 and 2**). We were most interested in the day 5 differentiated dataset as day 5 was when γ-globin was most upregulated.

We asked whether any of these differentially expressed genes could be contributing to the observed increased fetal hemoglobin in HBBko cells. To answer this, we first generated dCas9-KRAB stable cell lines in the WT background. For genes that were down-regulated in HBBko cells, we performed CRISPRi knockdown in WT cells and looked whether any of the gene knockdowns led to a phenocopy of HBBko. Our flow cytometry assay reports the fraction of cells staining positive for fetal hemoglobin. If any of the knockdowns led to an increase in number of fetal hemoglobin positive cells cells, this would suggest that the gene could be involved in the response of upregulating γ-globin. We performed knockdown using two different guides for each gene and differentiated the cells. By intracellular flow cytometry, there were several genes that led to modest upregulation of fetal hemoglobin (**Fig 7A**). Analysis of WT cells expressiong non-targetting guides revealed that the basal levels of fetal hemoglobin positive cells, were under 20%. Knockdown of *HBB* led to over 40% fetal hemoglobin, which was the highest level amongst all the gene knockdowns tested. HBBko cells had the highest level of fetal hemoglobin at 80%. Candidate gene knockdowns that led to modest increases in fetal hemoglobin include HMOX1, PYGB, TMEM123, VCL, and LSD1 (**Fig 7B**). Notably, only one guide against HMOX1 led to fetal hemoglobin levels as high as the HBB-targeting guides. However, the second guide for HMOX1 did not upregulate fetal hemoglobin levels above non-targeting control (data not shown) and therefore the results from guide one must be analyzed more carefully. Overall, no 2 guides for a singular gene was able to upregulate fetal hemoglobin levels as high as the HBB-targeting guides. This suggests that there could be mutiple pathways activated by loss of *HBB*, and not a singular gene is a causal regulator. Interestingly, HMOX1, a heme catabolism enzyme, and LSD1, a histone demethylase complex, have both previously been implicated in fetal hemoglobin regulation (Gil et al., 2013; Xu et al., 2013), strongly suggesting that loss of *HBB* may converge on these known regulatory pathways.

We next asked whether any of the upregulated genes in HBBko could be responsible for the increase in fetal hemoglobin. To test this, we generated a dCas9-KRAB stable cell line in the HBBko background. We performed knockdown of candidate genes that were upregulated in HBBko to test whether loss of any of these genes would

decrease the levels of fetal hemoglobin in HBBko. We differentiated these cells and by intracellualr flow staining, we did not observe that knockdown of any gene, except for *HBG1/2*, led to a decreased percentage of cells expressing fetal hemoglobin (**Fig 7C**).

Table 1. Genes most upregulated in 5 day differentiated HBBko cells compared to WT HUDEP-2

Number	Gene Name	p-Value	q-Value	ln fold change	Description
1	HBG2	1.3E-44	1.3E-42	2.4E+00	hemoglobin subunit gamma 2(HBG2)
2	HBG1	2.6E-59	5.4E-57	2.4E+00	hemoglobin subunit gamma 1(HBG1)
3	HBZ	6.1E-58	1.1E-55	2.3E+00	hemoglobin subunit zeta(HBZ)
4	HBBP1	1.1E-237	1.6E-233	2.1E+00	hemoglobin subunit beta pseudogene 1(HBBP1)
5	HBE1	3.6E-55	5.9E-53	2.0E+00	hemoglobin subunit epsilon 1(HBE1)
6	SSUH2	8.8E-58	1.6E-55	1.9E+00	ssu-2 homolog (C. elegans)(SSUH2)
7	HIST1H1C	2.2E-36	1.5E-34	1.7E+00	histone cluster 1 H1 family member c(HIST1H1C)
8	TMEM233	5.6E-15	8.5E-14	1.4E+00	transmembrane protein 233(TMEM233)
9	CRB1	2.6E-32	1.3E-30	1.4E+00	crumbs 1, cell polarity complex component(CRB1)
10	MYO5B	1.5E-07	9.4E-07	1.3E+00	myosin VB(MYO5B)
11	TSPAN32	1.8E-26	6.5E-25	1.3E+00	tetraspanin 32(TSPAN32)
12	LGR4	3.8E-49	4.6E-47	1.3E+00	leucine rich repeat containing G protein-coupled receptor 4(LGR4)
13	TMEM161B	1.6E-125	2.2E-122	1.3E+00	transmembrane protein 161B(TMEM161B)
14	FTCD	4.4E-30	1.9E-28	1.2E+00	formimidoyltransferase cyclodeaminase(FTCD)
15	FAM13A	3.4E-53	5.1E-51	1.2E+00	family with sequence similarity 13 member A(FAM13A)
16	FZD6	6.8E-23	2.0E-21	1.2E+00	frizzled class receptor 6(FZD6)

17	IQGAP2	1.4E-38	1.1E-36	1.2E+00	IQ motif containing GTPase activating protein 2(IQGAP2)
18	CNN3	5.9E-59	1.1E-56	1.2E+00	calponin 3(CNN3)
19	ASIC3	7.5E-33	3.9E-31	1.2E+00	acid sensing ion channel subunit 3(ASIC3)
20	MALT1	4.4E-111	4.3E-108	1.1E+00	MALT1 paracaspase(MALT1)
21	HIST1H2BK	2.8E-11	2.9E-10	1.1E+00	histone cluster 1 H2B family member k(HIST1H2BK)
22	TAF12	3.1E-78	1.1E-75	1.1E+00	TATA-box binding protein associated factor 12(TAF12)
23	B9D1	1.9E-30	8.9E-29	1.1E+00	B9 domain containing 1(B9D1)
24	SAPCD2	1.5E-22	4.2E-21	1.1E+00	suppressor APC domain containing 2(SAPCD2)
25	C11orf21	4.4E-26	1.5E-24	1.1E+00	chromosome 11 open reading frame 21(C11orf21)
26	HIST2H2AA4	8.8E-15	1.3E-13	1.0E+00	histone cluster 2 H2A family member a4(HIST2H2AA4)
27	MYL4	3.3E-36	2.2E-34	1.0E+00	myosin light chain 4(MYL4)
28	HIST1H2AC	6.5E-26	2.3E-24	1.0E+00	histone cluster 1 H2A family member c(HIST1H2AC)
29	ESPNP	7.4E-21	1.8E-19	1.0E+00	espin pseudogene(ESPNP)
30	BEX1	2.9E-52	4.0E-50	1.0E+00	brain expressed X-linked 1(BEX1)
31	FES	4.3E-42	4.0E-40	1.0E+00	FES proto-oncogene, tyrosine kinase(FES)
32	DHCR24	3.7E-14	5.1E-13	1.0E+00	24-dehydrocholesterol reductase(DHCR24)
33	IL12RB2	2.9E-13	3.7E-12	1.0E+00	interleukin 12 receptor subunit beta 2(IL12RB2)
34	MVK	7.4E-37	5.1E-35	1.0E+00	mevalonate kinase(MVK)
35	CAPN2	1.0E-30	4.8E-29	1.0E+00	calpain 2(CAPN2)
36	SQLE	2.5E-12	2.9E-11	1.0E+00	squalene epoxidase(SQLE)
37	DHCR7	5.5E-16	9.1E-15	1.0E+00	7-dehydrocholesterol reductase(DHCR7)

38	BAG3	2.2E-56	3.9E-54	9.9E-01	BCL2 associated athanogene 3(BAG3)
39	TMC8	4.7E-13	5.9E-12	9.7E-01	transmembrane channel like 8(TMC8)
40	DUSP6	3.5E-16	6.1E-15	9.6E-01	dual specificity phosphatase 6(DUSP6)
41	WWC2	5.8E-36	3.7E-34	9.6E-01	WW and C2 domain containing 2(WWC2)
42	PPT2-EGFL8	3.9E-03	1.1E-02	9.5E-01	PPT2-EGFL8 readthrough (NMD candidate)(PPT2-EGFL8)
43	GSTT1	1.1E-26	4.0E-25	9.3E-01	glutathione S-transferase theta 1(GSTT1)
44	TNFRSF25	4.5E-11	4.6E-10	9.3E-01	TNF receptor superfamily member 25(TNFRSF25)
45	TRIM10	8.0E-33	4.1E-31	9.2E-01	tripartite motif containing 10(TRIM10)
46	AIFM2	1.8E-152	3.8E-149	9.1E-01	apoptosis inducing factor, mitochondria associated 2(AIFM2)
47	PPM1K	5.3E-22	1.4E-20	9.1E-01	protein phosphatase, Mg2+/Mn2+ dependent 1K(PPM1K)
48	SLC30A10	1.5E-22	4.2E-21	9.0E-01	solute carrier family 30 member 10(SLC30A10)
49	LY6G5C	2.4E-05	1.1E-04	9.0E-01	lymphocyte antigen 6 complex, locus G5C(LY6G5C)
50	AC013474.1	6.3E-17	1.2E-15	9.0E-01	
51	SPINT2	2.7E-36	1.8E-34	9.0E-01	serine peptidase inhibitor, Kunitz type 2(SPINT2)
52	MTERF2	1.3E-14	1.9E-13	8.9E-01	mitochondrial transcription termination factor 2(MTERF2)
53	MTND1P23	3.5E-22	9.5E-21	8.9E-01	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 1 pseudogene 23(MTND1P23)
54	ATP9A	2.8E-13	3.6E-12	8.9E-01	ATPase phospholipid transporting 9A (putative)(ATP9A)
55	FAM98A	5.4E-05	2.3E-04	8.9E-01	family with sequence similarity 98 member A(FAM98A)
56	ATF7IP2	1.5E-17	2.9E-16	8.9E-01	activating transcription factor 7 interacting protein 2(ATF7IP2)

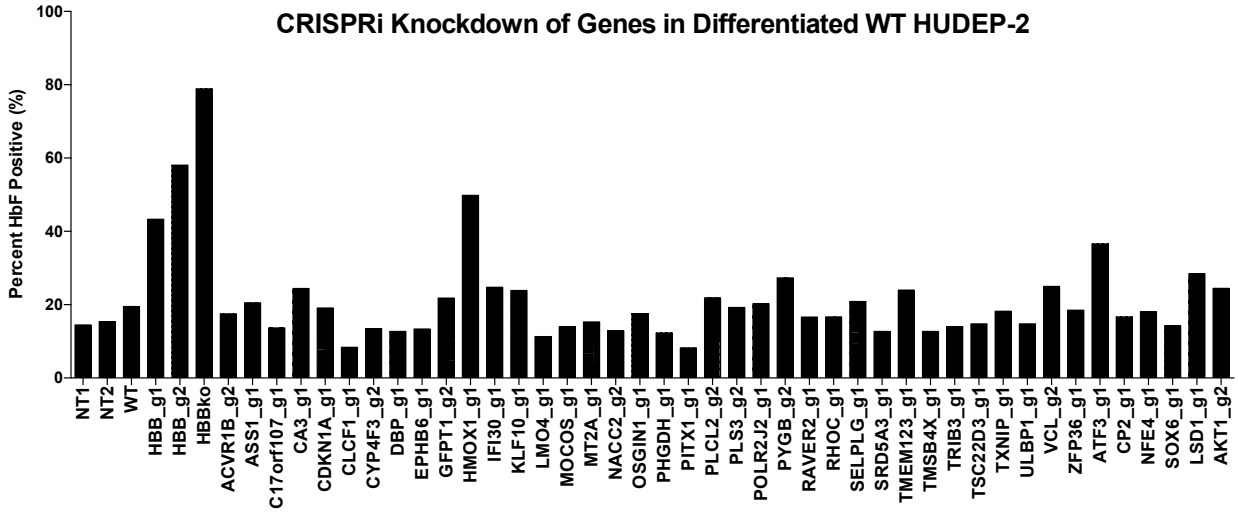
57	RIMS3	9.0E-14	1.2E-12	8.8E-01	regulating synaptic membrane exocytosis 3(RIMS3)
58	ACSBG1	3.4E-16	5.9E-15	8.7E-01	acyl-CoA synthetase bubblegum family member 1(ACSBG1)
59	AC006116.1	1.1E-11	1.2E-10	8.7E-01	
60	MC1R	4.0E-54	6.2E-52	8.7E-01	melanocortin 1 receptor(MC1R)
61	ZNF664	2.0E-21	5.2E-20	8.7E-01	zinc finger protein 664(ZNF664)
62	FEZ1	1.9E-11	2.0E-10	8.5E-01	fasciculation and elongation protein zeta 1(FEZ1)
63	ESPN	9.2E-85	4.2E-82	8.5E-01	espin(ESPN)
64	PIMREG	4.7E-24	1.4E-22	8.4E-01	PICALM Interacting Mitotic Regulator
65	APOBEC3G	1.1E-10	1.1E-09	8.4E-01	apolipoprotein B mRNA editing enzyme catalytic subunit 3G(APOBEC3G)
66	TTC8	1.6E-07	1.1E-06	8.4E-01	tetratricopeptide repeat domain 8(TTC8)
67	TDRKH	1.8E-07	1.2E-06	8.4E-01	tudor and KH domain containing(TDRKH)
68	FCHSD1	3.5E-10	3.2E-09	8.4E-01	FCH and double SH3 domains 1(FCHSD1)
69	AGL	4.0E-41	3.5E-39	8.3E-01	amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase(AGL)
70	PHTF1	1.7E-13	2.2E-12	8.3E-01	putative homeodomain transcription factor 1(PHTF1)
71	TMEM98	1.1E-14	1.7E-13	8.3E-01	transmembrane protein 98(TMEM98)
72	MED31	2.0E-16	3.5E-15	8.2E-01	mediator complex subunit 31(MED31)

Table 2. Genes most downregulated in 5 day differentiated HBBko cells compared to WT HUDEP-2

Number	Gene Name	p-Value	q-Value	ln Fold Change	Description
1	PLS3	6.14E-217	4.54E-213	-3.395	plastin 3(PLS3)
2	CLCF1	1.25E-32	6.39E-31	-2.285	cardiotrophin-like cytokine factor 1(CLCF1)
3	ASS1	2.21E-37	1.55E-35	-2.191	argininosuccinate synthase 1(ASS1)
4	HMOX1	7.65E-38	5.46E-36	-2.019	heme oxygenase 1(HMOX1)
5	POLR2J2	6.06E-13	7.55E-12	-1.877	RNA polymerase II subunit J2(POLR2J2)
6	HBB	2.68E-146	4.39E-143	-1.832	hemoglobin subunit beta(HBB)
7	ULBP1	6.81E-49	8.25E-47	-1.809	UL16 binding protein 1(ULBP1)
8	C17orf107	6.70E-67	1.83E-64	-1.781	chromosome 17 open reading frame 107(C17orf107)
9	TMSB4X	3.22E-50	4.13E-48	-1.737	thymosin beta 4, X-linked(TMSB4X)
10	TRIB3	1.78E-70	5.61E-68	-1.708	tribbles pseudokinase 3(TRIB3)
11	RAVER2	1.03E-13	1.38E-12	-1.705	ribonucleoprotein, PTB binding 2(RAVER2)
12	TXNIP	3.71E-33	1.99E-31	-1.698	thioredoxin interacting protein(TXNIP)
13	RHOC	1.04E-46	1.15E-44	-1.657	ras homolog family member C(RHOC)
14	KLF10	5.49E-90	3.01E-87	-1.633	Kruppel like factor 10(KLF10)
15	PITX1	3.47E-14	4.86E-13	-1.562	paired like homeodomain 1(PITX1)
16	CA3	1.71E-35	1.05E-33	-1.552	carbonic anhydrase 3(CA3)
17	IFI30	2.03E-162	4.99E-159	-1.549	IFI30, lysosomal thiol reductase(IFI30)
18	MT2A	6.46E-14	8.81E-13	-1.516	metallothionein 2A(MT2A)
19	CDKN1A	9.72E-22	2.56E-20	-1.493	cyclin dependent kinase inhibitor 1A(CDKN1A)
20	LMO4	1.82E-88	9.61E-86	-1.455	LIM domain only 4(LMO4)

21	EPHB6	3.29E-49	4.06E-47	-1.442	EPH receptor B6(EPHB6)
22	SELPLG	8.30E-125	1.02E-121	-1.428	selectin P ligand(SELPLG)
23	VCL	1.71E-97	1.26E-94	-1.397	vinculin(VCL)
24	NACC2	1.12E-30	5.20E-29	-1.381	NACC family member 2(NACC2)
25	DBP	3.15E-17	5.91E-16	-1.377	D-box binding PAR bZIP transcription factor(DBP)
26	PYGB	1.84E-25	6.21E-24	-1.336	phosphorylase, glycogen; brain(PYGB)
27	MOCOS	1.19E-46	1.31E-44	-1.334	molybdenum cofactor sulfurase(MOCOS)
28	TMEM123	4.21E-36	2.72E-34	-1.330	transmembrane protein 123(TMEM123)
29	PLCL2	1.14E-80	4.57E-78	-1.329	phospholipase C like 2(PLCL2)
30	GFPT1	3.09E-40	2.56E-38	-1.318	glutamine--fructose-6-phosphate transaminase 1(GFPT1)
31	ACVR1B	7.54E-26	2.60E-24	-1.315	activin A receptor type 1B(ACVR1B)
32	TSC22D3	8.06E-106	7.45E-103	-1.311	TSC22 domain family member 3(TSC22D3)
33	OSGIN1	2.01E-14	2.88E-13	-1.301	oxidative stress induced growth inhibitor 1(OSGIN1)
34	CYP4F3	3.49E-16	5.97E-15	-1.286	cytochrome P450 family 4 subfamily F member 3(CYP4F3)
35	SRD5A3	8.81E-54	1.36E-51	-1.281	steroid 5 alpha-reductase 3(SRD5A3)
36	ZFP36	4.52E-05	0.00019581	-1.277	ZFP36 ring finger protein(ZFP36)
37	PHGDH	5.19E-29	2.21E-27	-1.275	phosphoglycerate dehydrogenase(PHGDH)

A



B

Gene knockdown	% HbF-positive		Description of gene product
	Replicate 1	Replicate 2	
HMOX_g1	30.4	49.8	Heme oxygenase 1: essential enzyme in heme catabolism
PYGB_g2	21.6	27.3	Glycogen phosphorylase B: catalyzes the rate-determining step in glycogen degradation
TMEM123_g1	25.3	24.0	Transmembrane protein 123: cell surface receptor that mediates cell death
VCL_g2	26.2	25.0	Vinculin: cytoskeletal protein that helps anchor F-actin to cell membrane
LSD1_g1	29.5	28.5	Lysine-specific histone demethylase 1A: silences genes by demethylating histones
LSD1_g2	24.0	23.0	
NT1	15.7	15.4	N/A (negative control)

C

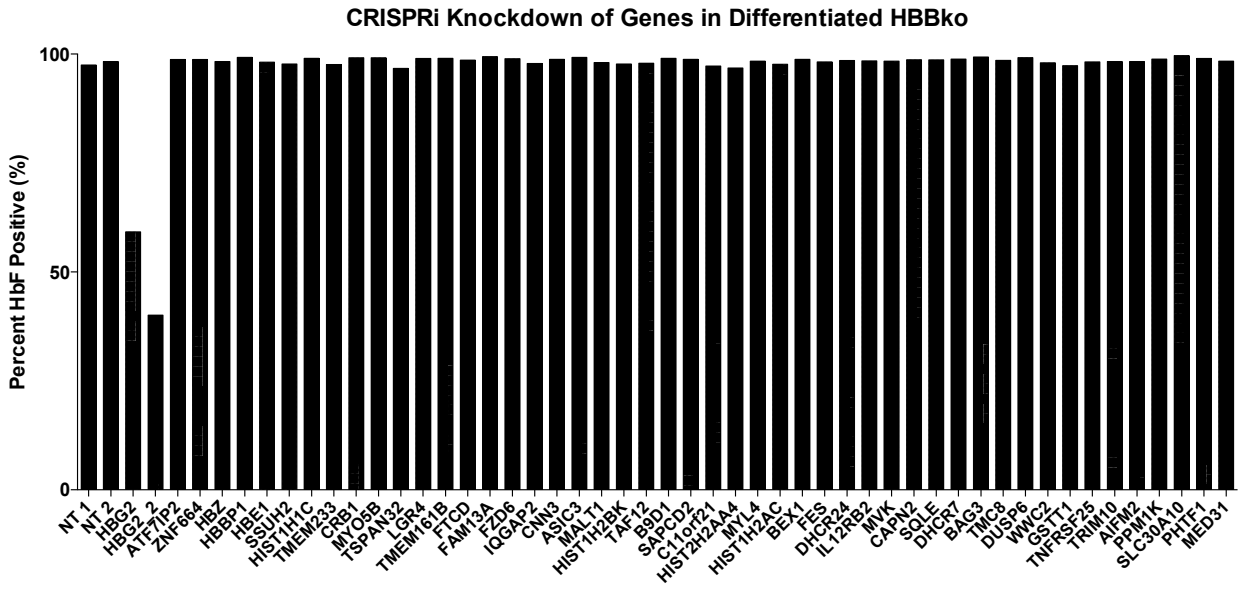


Figure 7**.** Knockdown of differentially expressed genes between HBBko and WT HUDEP-2 cells.

A) Intracellular flow cytometry staining for HbF in differentiated WT HUDEP-2 cells. CRISPRi-mediated knockdown of genes found to be down-regulated in HBBko cells. This experiment was performed in biological duplicate with data from one replicate shown. Two guides were used for each gene and the data shown is for guide 1. In cases where guide 1 led to cell death, data from guide 2 is shown. Guide sequences for CRISPRi knockdown can be found in **Table S4**.

B) Genes that led to increase of HbF upon knockdown. Percentages of fetal hemoglobin are shown for two biological replicates as well as the description of the genes.

C) Intracellular flow cytometry staining for HbF in differentiated HBBko HUDEP-2 cells. CRISPRi-mediated knockdown of genes found to be up-regulated in HBBko cells. This experiment was performed in biological duplicate with data from one replicate shown. Two guides were used for each gene and the data shown is for guide 1. In cases where guide 1 led to cell death, data from guide 2 is shown.

3.4.2 Knockdown of ATF4-targeted candidate genes reveals that DENND4A may regulate fetal hemoglobin levels.

In the previous chapter, we discovered ATF4 as a causal regulator in the response to loss of *HBB*. We performed ChIP-seq on ATF4 in undifferentiated and differentiated HUDEP-2 as well as primary human early and late erythroblasts. The ChIP-seq dataset revealed that ATF4 was binding within the HBS1L-MYB intergenic enhancer region. Additionally, the ChIP-seq dataset also revealed the rest of ATF4 binding sites in erythroid cells. From the ChIP-seq dataset, we curated a list of candidate ATF4 target genes that could be involved in fetal hemoglobin regulation.

Our ATF4 ChIP-seq dataset revealed 1245 ATF4 targets. We narrowed this list down to 24 genes for CRISPRi knockdown (**Table 3**). In curating this list, we prioritized genes that had high ATF4 peak signal compared to the IgG negative control. We favored transcription factors and genes implicated in erythroid phenotypes as reported by databases such as Genome Wide Association Studies (GWAS). Additionally, we categorized genes that were either targeted by ATF4 only in the differentiated state, undifferentiated state, or both undifferentiated and differentiated state. Lastly, we cross-referenced these genes to our RNA-seq database and confirmed that these genes were also downregulated in HBBko compared to WT cells.

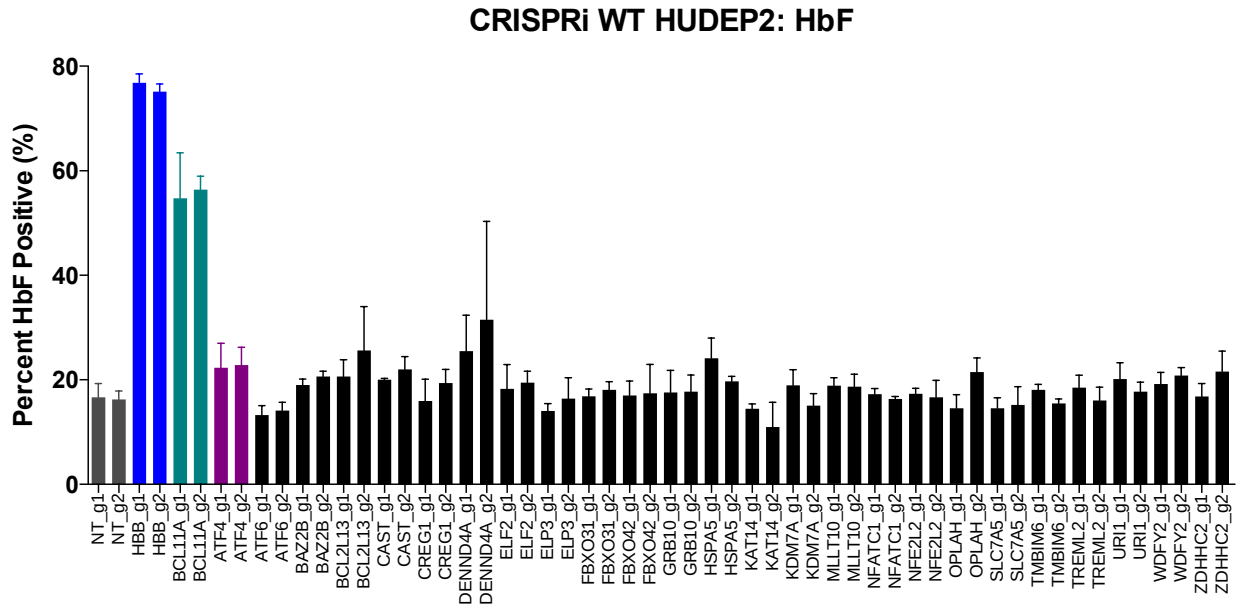
We performed CRISPRi knockdown of 24 candidate genes and controls of non-target guides and guides targeting *HBB*, BCL11A, and ATF4 (**Fig 8A**). We differentiated these cells and performed intracellular flow staining. We observed that knockdown of *HBB* led to the highest increase in fetal hemoglobin reaching almost 80%, compared to less than 20% for non-targeting guides. Knockdown of BCL11A led to fetal hemoglobin levels of 60%. Notably, knockdown of ATF4 led to only a modest increase fetal hemoglobin. Given that our previous chapter shows much evidence for ATF4 regulating fetal hemoglobin, this modest phenotype may be due to the fact that complete loss of ATF4 results in lethality during differentiation, and so this CRISPRi assay may select for cells that maintain normal levels of ATF4. ATF4 is also known to be translationally regulated (Blais et al., 2004)and so another possibility is that the translation of the reduced number of ATF4 transcripts could elevated. Additionally, we postulate that ATF4 is not the sole regulatory pathway in the HBBko response and that other pathways are likely involved.

Of the 24 candidate genes, we found that DENND4A, a C-Myc promoter-binding protein, resulted in the highest increase of fetal hemoglobin compared to the non-targeting guides (**Fig 8A**). DENND4A is highly expressed in erythroid cells and SNPs within this gene has been associated to hemoglobin concentrations within the cell (Ding et al., 2012). The other 23 candidate genes either did not appear to affect fetal hemoglobin levels. Other than knockdown of *HBB*, none of the other genetic knockdowns affected levels of adult hemoglobin (**Fig 8B**).

Table 3**.** ATF4-targeted candidate genes for CRISPRi knockdown in WT HUDEP-2 cells.

SYMBOL	HBBko vs IgG diff	WenDEP vs IgG diff	WenDEP vs HBBko diff	HBBko vs WenDEP diff	HBBko vs ATF4KO3 undiff	WenDEP vs ATF4KO3 undiff	WenDEP vs HBBko undiff	HBBko vs WenDEP undiff	GENENAME
MLLT10	148.3	450.3	305.0	0.0	0.0	0.0	0.0	0.0	MLLT10 histone lysine methyltransferase DOT1L cofa
ELP3	102.3	395.2	295.8	0.0	0.0	0.0	0.0	0.0	elongator acetyltransferase complex subunit 3
ATF6	136.6	367.9	235.6	0.0	0.0	0.0	0.0	0.0	activating transcription factor 6
BAZ2B	129.3	290.6	140.6	0.0	0.0	0.0	0.0	0.0	bromodomain adjacent to zinc finger domain 2B
TMBIM6	102.7	200.6	126.3	0.0	0.0	0.0	0.0	0.0	transmembrane BAX inhibitor motif containing 6
TREML2	0.0	169.0	126.2	0.0	0.0	0.0	0.0	0.0	triggering receptor expressed on myeloid cells like 2
ELF2	0.0	117.1	91.9	0.0	0.0	0.0	0.0	0.0	E74 like ETS transcription factor 2
KDM7A	0.0	82.5	57.9	0.0	0.0	0.0	0.0	0.0	lysine demethylase 7A
KAT14	0.0	87.0	57.8	0.0	0.0	0.0	0.0	0.0	lysine acetyltransferase 14
ZDHHC2	0.0	97.3	53.8	0.0	0.0	0.0	0.0	0.0	zinc finger DHHC-type containing 2
FBXO31	211.1	532.4	322.4	0.0	0.0	0.0	0.0	0.0	F-box protein 31
HSPA5	247.1	538.0	298.8	0.0	0.0	47.1	0.0	0.0	heat shock protein family A (Hsp70) member 5
SLC7A5	395.5	640.7	251.4	0.0	132.8	376.1	153.6	0.0	solute carrier family 7 member 5
FBXO42	245.9	500.6	243.2	0.0	44.3	174.9	96.3	0.0	F-box protein 42
NFATC1	162.3	394.9	231.0	0.0	44.0	134.0	61.7	0.0	nuclear factor of activated T cells 1
WDFY2	0.0	94.2	67.3	0.0	44.2	66.4	0.0	0.0	WD repeat and FYVE domain containing 2
URI1	0.0	0.0	0.0	0.0	0.0	69.5	47.9	0.0	URI1 prefoldin like chaperone
CAST	0.0	0.0	0.0	0.0	0.0	57.3	45.0	0.0	calpastatin
BCL2L13	0.0	0.0	0.0	0.0	0.0	74.5	0.0	0.0	BCL2 like 13
GRB10	51.2	0.0	0.0	0.0	0.0	67.3	0.0	0.0	growth factor receptor bound protein 10
NFE2L2	0.0	0.0	0.0	0.0	0.0	63.6	0.0	0.0	nuclear factor, erythroid 2 like 2
CREG1	0.0	0.0	0.0	0.0	0.0	51.8	0.0	0.0	cellular repressor of E1A stimulated genes 1
DENND4A	0.0	0.0	0.0	0.0	0.0	47.7	0.0	0.0	DENN domain containing 4A
OPLAH	0.0	53.6	40.5	0.0	0.0	0.0	0.0	0.0	5-oxoprolinase, ATP-hydrolysing

A



B

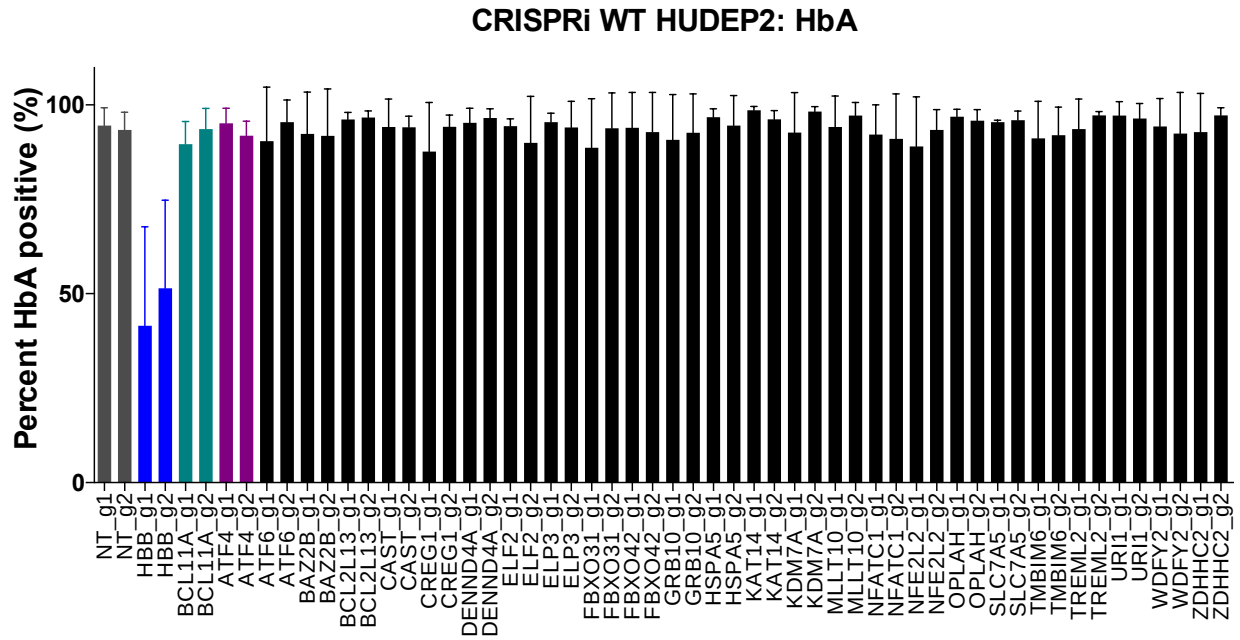


Figure 8. Knockdown of candidate ATF4-targeting genes in WT HUDEP-2 cells

A) Intracellular flow staining of HbF in differentiated WT HUDEP-2 cells. CRISPRi knockdown was performed for two guides for each gene. The data is shown as biological triplicates.

B) Intracellular flow staining of HbA in differentiated WT HUDEP-2 cells. CRISPRi knockdown was performed for two guides for each gene. The data is shown as biological triplicates.

3.4.3 Knockdown of hypothesis-driven candidate genes reveals that BACH1 could be a repressor of fetal hemoglobin in the NRF2 pathway.

We have shown that knockout of *HBB* leads to high upregulation of HBG. We have also shown that knockdown of *HBB* leads to upregulation of HBG, although not to the same extent as complete loss of *HBB*. Because of this, we postulate that there must be a mechanism sensing levels of hemoglobin in the cell, allowing the cell to compensate if there is insufficient hemoglobin present. We formed hypotheses on what these genes may be based on a few scenarios. Lack of *HBB* could result in the following: excess α-globin chains, excess heme, and unbound *HBB* mRNA interactors. Based on literature around these possibilities, we curated a list of genes to test for their role in fetal hemoglobin regulation (**Table 4**).

We performed CRISPRi knockdown of these genes in WT HUDEP-2 cells (**Fig 9A**). We differentiated these cells and performed intracellular flow staining for fetal hemoglobin. Notably, this experiment was not ideal as the non-targeting conditions cells died during puromycin selection. However, we proceeded given that the other control conditions (guides targeting HBG, BCL11A, and ZBTB7A) were still valid. Initially, we were excited to see high levels of fetal hemoglobin induction in NFKBIB, NF-kappa-B inhibitor, and YBX1, a protein implicated in binding the *HBB* mRNA at the 3'UTR (van Zalen et al., 2012). However, only one of two guides yielded this phenotype, and upon repeating the experiment focusing on knocking down just NFKBIB and YBX1, the results did not repeat (data not shown). The strong phenotypes are likely due to clonal effects

Knockdown of BACH1 for both guides led to upregulation of fetal hemoglobin to similar levels as knockdown of ZBTB7A. BACH1 is a transcription factor normally part of a repressor complex (Zenke-Kawasaki et al., 2007). Upon complexing with heme, BACH1 loses its DNA-binding ability and becomes exported from the nucleus and degraded (Zenke-Kawasaki et al., 2007). This leads to derepression of BACH1 target genes. We first asked whether HBG is one of the genes repressed by BACH1. For this, we generated a pool of BACH1 knockouts by using a dual cutting CRISPR-Cas9 approach to remove the start codon (**Fig 9B**). Editing efficiency was low and therefore, by intracellular staining, we were only able to observe a slight increase in fetal hemoglobin (**Fig 9C**).

Due to low editing efficiency of the BACH1 pooled knockout, we screened to find clones and was able to isolate heterozygote and a homozygote BACH1 deletion (**Fig 9D**). We differentiated these cells and performed a western blot on BACH1 and γ-globin and observed that BACH1 mutants had increased γ-globin. Interestingly, some of the BACH1 mutants expressed a truncated BACH1 protein product, most prominently observed in the BACH1 homozygote knockout. Similar to ATF4, BACH1 is a bZIP transcription factor and removing the first start codon potentially leads to a truncated version of BACH1 that we observed with ATF4ΔN in the previous chapter. We performed intracellular flow staining for fetal hemoglobin and observed that all BACH1 mutants had higher levels of fetal hemoglobin compared to WT cells (**Fig 9F**).

Table 4. Candidate genes involved in sensing loss of *HBB*.

Candidates	Description
Heme -Sensing	
HRI	Heme-regulated kinase
DNAJC3	Hsp40 homolog involved in unfolded protein response
NFKBIB	Nuclear factor for NFKB. Its phosphorylation governs the transport of NFKB into nucleus
NFE2L2/NRF2	NRF2 leucine zipper type TF, implicated in coordinating response to oxidative stress
KEAP1	Binds to NRF2 to promote cytosolic degradation via ubiquitin pathway
BACH1	Competes for Nrf2 sites
MAFK	basic leucine zipper type TF, binds to NFE2 complex, which binds to HS2 of LCR
MAFG	redundant to MAFK
MAFF	somewhat redundant to MAFK
DDIT3	TF in ER-stress and cell stress
Alpha-chain sensing	
AHSP	Alpha-hemoglobin stabilizing protein
ZC3H12A	AHSP interactor, KO leads to anemia, unknown function
HBB mRNA interactors	
AUF-1	binds HBB 3'UTR mRNA
YB-1	binds HBB 3'UTR mRNA

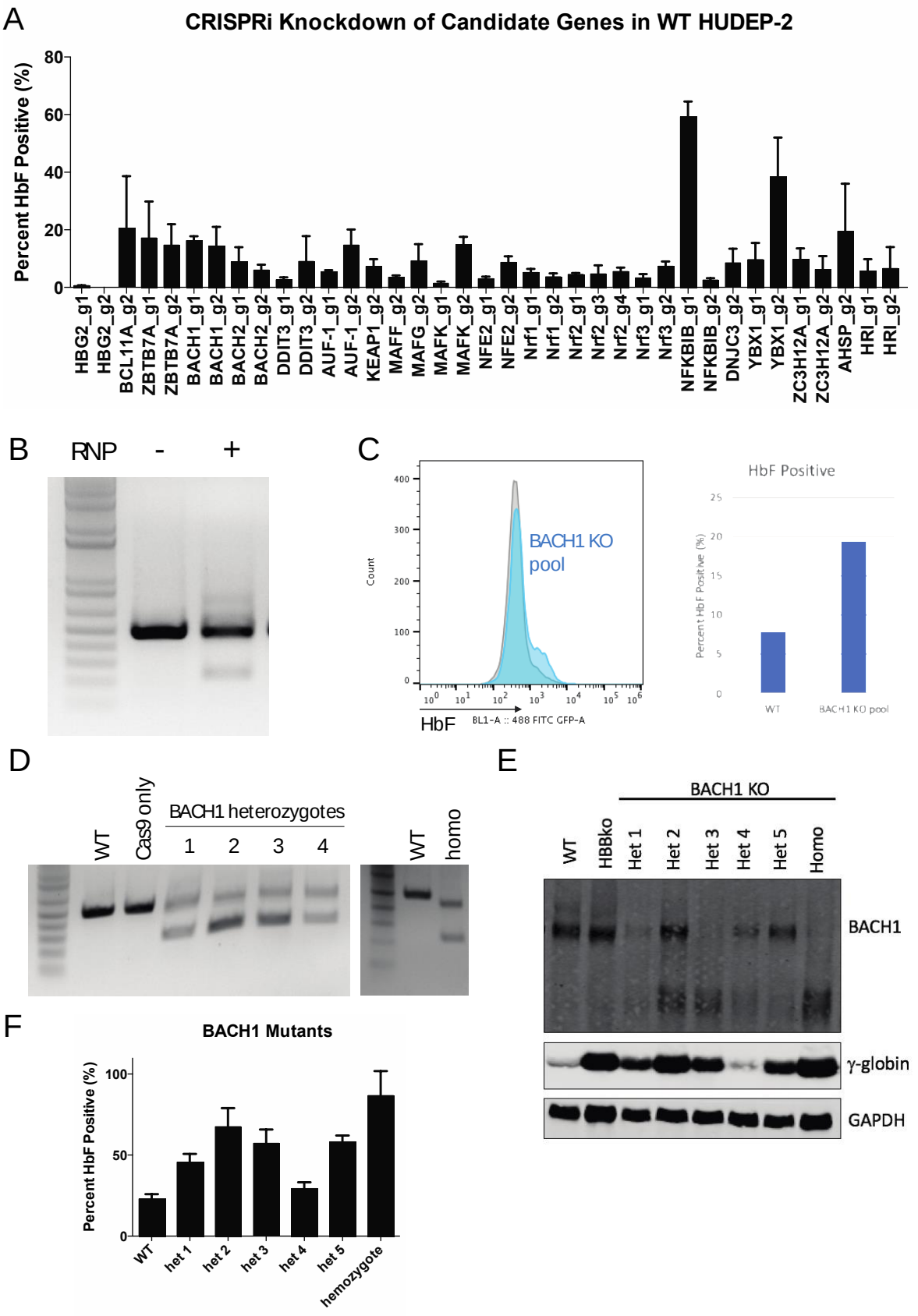

A
CRISPRi Knockdown of Candidate Genes in WT HUDEP-2
Percent HbF Positive (%)
80
60
40
20
0
HBG2_g1
HBG2_g2
BCL11A_g2
ZBTB7A_g1
ZBTB7A_g2
BACH1_g1
BACH1_g2
BACH2_g1
BACH2_g2
DDIT3_g1
DDIT3_g2
AUF-1_g1
AUF-1_g2
KEAP1_g2
MAFF_g2
MAFG_g2
MAFK_g1
MAFK_g2
NFE2_g1
NFE2_g2
Nrf1_g1
Nrf1_g2
Nrf2_g1
Nrf2_g3
Nrf2_g4
Nrf3_g1
Nrf3_g2
NFKBIB_g1
NFKBIB_g2
DNJC3_g2
YBX1_g1
YBX1_g2
ZC3H12A_g1
ZC3H12A_g2
AHSP_g2
HRI_g1
HRI_g2
B
RNP
-
+
C
Count
400
300
200
100
0
BACH1 KO pool
HbF
BL1-A :: 488 FITC GFP-A
HbF Positive
Percent HbF Positive (%)
25
20
15
10
5
0
WT
BACH1 KO pool
D
WT
Cas9 only
BACH1 heterozygotes
1
2
3
4
WT
homo
E
BACH1 KO
WT
HBBko
Het 1
Het 2
Het 3
Het 4
Het 5
Homo
BACH1
γ-globin
GAPDH
F
BACH1 Mutants
Percent HbF Positive (%)
100
50
0
WT
het 1
het 2
het 3
het 4
het 5
hemozygote

Figure 9. BACH1 mutants have upregulated fetal hemoglobin in HUDEP-2 cells

A) Intracellular flow staining of HbF in differentiated WT HUDEP-2 cells. CRISPRi knockdown was performed for two guides for each gene. The data is shown as biological triplicates.
B) PCR validation of BACH1 pooled knockouts
C) Left panel shows intracellular flow staining of HbF in BACH1 pooled knockout (trace in blue) overlayed with WT (trace in gray). Right panel shows quantification of percent HbF positive. This experiment was performed once.
D) PCR validation of BACH1 clones. Four heterozygotes are shown and one homozygote BACH1 knockout.
E) Western Blot showing BACH1 full length and truncation form and γ-globin in differentiated HUDEP-2 cells.

3.4.4 Loss of BACH1 upregulates fetal hemoglobin independent of NRF2.

BACH1 is known to interact within a corepressor complex to repress expression of the targeted gene (Oyake et al., 1996). Conversely, nuclear factor erythroid 2-related factor 2 (NRF2) is a transcription factor known to co-occupy BACH1 binding sites to activate the target gene. BACH1 and NRF2 compete for the same sites in order to control regulation of that gene (Dhakshinamoorthy et al., 2005). It has previously been shown that cellular treatment with tert-butylhydroquinone (tBHQ) leads to increased levels and nuclear translocation of NRF2 (Macari and Lowrey, 2011). NRF2 binds to the γ-globin promoter to promote transcription of *HBG1/2* and ultimately increases expression of fetal hemoglobin (Macari and Lowrey, 2011).

We postulate that in adult erythoid cells, BACH1 could be binding and repressing the *HBG1/2* genes. We asked whether loss of BACH1 could lead to increased cellular expression or increased nuclear translocation of NRF2 which could explain the rise in γ-globin expression. To answer this, we performed a fractionation western blot to cytoplasic lysates and nuclear lysates. We treated cells with tBHQ as a control to observe increased expression and nuclear translocation of NRF2. We found that increased γ-globin in HBBko and BACH1ko were independent of NRF2 expression and nuclear localization (**Fig 10**). Treatment with tBHQ led to increased NRF2 nuclear localization and increased γ-globin, consistent with previous reports. However, loss of BACH1 in BACH1ko cells, had no affect on expression of NRF2. Interestingly, it seems that HBBko cells have slightly decreased levels of BACH1 compared to WT. Taken together, our data suggests that loss of BACH1 is sufficient to increase fetal hemoglobin regardless of NRF2.

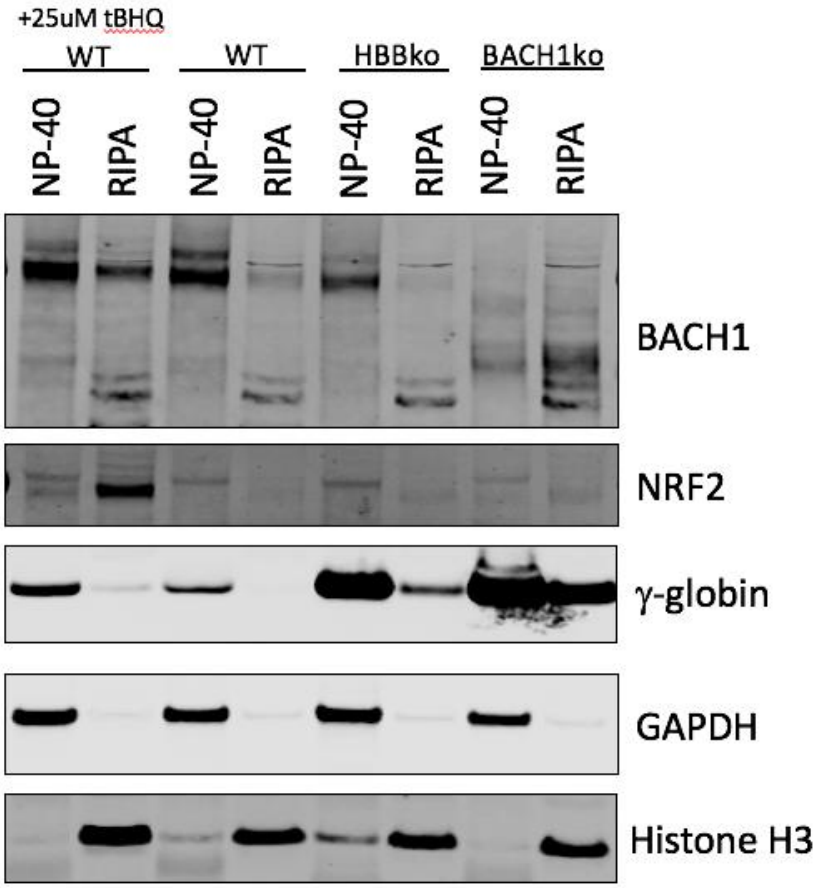


Figure 10**.** Loss of BACH1 leads to increased γ-globin and does not affect cellular levels or nuclear localization of NRF2.
Fractionation western blot showing BACH1, NRF2, and γ-globin. NP-40 refers to cytoplasmic lysates and RIPA refers to nuclear lysates. Cells were differentiated with or without tBHQ for 5 days and harvested for fractionation.

3.5 Discussion

Loss of *HBB* leads to robust reactivation of *HBG1/2* in erythroid cells. In our previous chapter, we showed that ATF4 is one causal regulator of this response. However, there are likely other pathways involved since loss of ATF4 is insufficient to upregulate fetal hemoglobin to levels observed in HBBko. This chapter explored other pathways by utilizing the CRISPRi to repidly screen many candidate genes.

For the RNA-seq candidates, we looked at both upregulated and downregulated genes in HBBko. Knockdown of HMOX1, PYGB, TMEM123, VCL, and LSD1, genes downregulated in HBBko, all led to modest increases in fetal hemoglobin in WT cells. Notably, HMOX1, a heme catabolism enzyme, and LSD1, a histone demethylase complex, have both previously been implicated in fetal hemoglobin regulation (Gil et al., 2013; Xu et al., 2013). Our data suggests that the response to loss of *HBB* could converge on these known regulatory pathways. Additionally, these positive results indicate that CRISPRi could be useful in an unbiased screening approach to uncover new or validate implicated regulators of fetal hemoglobin. It is important to note that we only explored genes that were differentially expressed at day 5 of differentiation. The assumption of looking at day 5 genes is that the differentially expressed genes were causative of the response, and not just an effect of the response. Additionally, we assumed that mRNA levels correlate with protein levels. It would be interesting to look at the differentially expressed genes from undifferentiated and 2 days differentiated datasets in the case that these genes are more of an upstream causative regulator that could explain the results of 5 day differentiation differences.

Our ATF4 ChIP-seq dataset allowed us to observe where ATF4 binds in the genome in differentiated erythroid cells. In our previous chapter, we discovered that ATF4 regulated MYB, a known regulator of fetal hemoglobin. We further knocked down 24 ATF4 target genes and found that DENND4A also leads to increases in fetal hemoglobin. Overall, little is known about DENND4A and further experiments would need to be done to validate its role in hemoglobin regulation. DENND4A has been associated with hemoglobin phenotypes in GWAS studies and is therefore a promising candidate for followup experiments.

In knocking down our curated list of candidate genes, we showed much evidence that BACH1 is repressor of fetal hemoglobin. Of note, this CRISPRi experiment was poorly controlled with a lack of non-targeting control parameters and therefore, our results for this section should be analyzed carefully, and ideally repeated with proper controls in place. Regardless, BACH1 is a very promising candidate given the reported literature around NRF2. Studies have shown that NRF2 is an activator of fetal hemoglobin by direct binding the the *HBG1/2* promoters (Macari and Lowrey, 2011). Further studies have shown that BACH1 and NRF2 compete for the same binding sites (Dhakshinamoorthy et al., 2005). However, it has not been shown that BACH1 directly binds and respresses *HBG1/2*.

It is possible that BACH1 is basally repressing *HBG1/2* at the same site that NRF2 has reported activation activity at. Followup studies regarding BACH1 would need to show BACH1 binding sites at the *HBG1/2* promoter and how excess heme leads to removal of BACH1 and subsequent *HBG1/2* derepression. With the removal and

degradation of BACH1 via heme-mediated ubiquitination (Zenke-Kawasaki et al., 2007), NRF2 may then bind to *HBG1/2* to activate transcription. Further studies would need to characterize the BACH1 truncation mutant to determine whether it has lost transcriptional activity, as was the case with ATF4DN. Finally, it would be interesting to explore whether BACH1 plays a role in the response to loss of *HBB*. Indeed, in one experiment, HBBko seemed to have less BACH1 compared to WT (**Fig 10**).

It is important to note that in all of our CRISPRi experiments, knockdown of *HBB* led to the highest upregulation of fetal hemoglobin. Knockdown of BCL11A and ZBTB7A, known regulators of *HBG1/2*, led to high increases of fetal hemoglobin, but still not to the same extent as knockdown of *HBB*. This strongly suggests that the response to loss of *HBB* is a culmination of multiple players and pathways. Followup studies should further explore some of the implicated regulators found in this chapter in hopes that more of the pathways involved in the loss of *HBB* can be elucidated.

3.6 Methods

HUDEP-2 cell culture and differentiation
All cell culture was performed at 37°C in a humidified atmosphere containing 5% CO2. HUDEP-2 cells were cultured in a base medium of SFEM (Stemcell Technologies 9650) containing to a final concentration of dexamethasone 1uM (Sigma D4902-100MG), doxycycline 1ug/ml (Sigma D9891-1G), human stem cell factor 50ng/ml (PeproTech 300-07), erythropoietin 50ng/ml (Peprotech 100-64), and penstrept 1%. Cells were cultured at a density of 2e5 – 1e6 cells/ml. For differentiation, HUDEP-2 cells are centrifuged at 500g for 5 minutes, media is removed and replaced with differentiation media. Differentiation media consists of a base media of IMDM+Glutamax (ThermoFisher 31980030) containing to a final concentration human serum 5% (Sigma H4522-100mL), heparin 2IU/ml (Sigma H3149-25KU), insulin 10ug/ml (Sigma I2643-25mg), erythropoietin 50ng/ml (Peprotech 100-64), holo-transferrin 500ug/ml (Sigma T0665-100mg), mifepristone 1uM (Sigma M8046-100MG), and doxycyline 1ug/ml (Sigma D9891-1G). Cells are differentiated for 5 days and then harvested for analysis.

HEK293T cell culture
All cell culture was performed at 37°C in a humidified atmosphere containing 5% CO2. HEK293T cells were grown in a base media of DMEM supplemented with 10% fetal bovine serum, 10% sodium pyruvate, and 1% penstrept. Cells were cultured at a density of 2e5 – 1e6 cells/ml.

Generation of CRISPRi HUDEP-2 cells
WT-HUDEP-2 or *HBB* ko cells were transduced with lentivirus containing the construct EF1a-dCas9-HA-BFP-KRAB-NLS (Liang et al., 2018). The cells were FACs sorted for BFP expression and single-cell cloned. Individual clones were transduced with lentivirus containing guide RNAs targeting either *CD55, CD59, or HBB*. The clones were validated for successful knockdown by flow cytometry staining for extracellular markers CD55 and CD59 or intracellularly stained for β-globin.

Lenti-viral Packaging
Lenti-viral packaging of all constructs was performed using HEK-293T cells. TransIT®-LT1 Transfection Reagent (Mirus) was used following manufacturer's instructions. The plasmid mixture contained 50% construct plasmid, 40%DVPR, and 10% VSVG. Viral supernatant was harvested after 48 and 72 hours and filtered through 0.45 µM. For transduction of HUDEP-2 cells, cells were cultured in 50% HUDEP-2 media and 50% viral supernatant for 24 hours.

sgRNA Plasmid Cloning
sgRNA guide sequences for CRISPRi transcriptional repression were obtained from the Weissman CRISPRi-v2 library (Horlbeck et al., 2016). The chosen guides were cloned into pGL1-library vector (Addgene 84832). All guides used are listed in Supplementary Table 2.

Cas9 RNP Nucleofection

Cas9 RNP was performed as described previously (Lingeman et al., 2017). Briefly, IVT guides are purified and complexed with purified Cas9-NLS protein. The nucleofection was performed using Lonza 4D-Nucleofector and using the P3 Primary Cell 96-well NucleofectorTM Kit (V4SP-3096) following manufacturer's instructions. The HUDEP-2 nucleofector code used was DD-100 and for primary HSPCs ER-100.

IVT sgRNA

Guide RNAs were in vitro transcribed as described previously (Lingeman et al., 2017). Briefly, guide sequences were ordered as oligonucleotides and formed into duplexes using a PCR thermocycler. The DNA template was transcribed to RNA using HiScribe™ T7 High Yield RNA Synthesis Kit (E2040S) following manufacturer protocol. The resulting RNA was purified using RNeasy Mini kit (74104) and Rnase-Free DnaseI Kit (79254)

Intracellular FACs Staining

Staining was performed as adapted from the existing methods (Chung et al., 2019b). Briefly, undifferentiated or differentiated HUDEP-2 cells were centrifuged at 500g for 5 minutes. The cells were washed with PBS with 0.1% BSA, re-centrifuged, and fixed in 0.05% glutaraldehyde. The fixed cells were centrifuged and washed and re-suspended in PBS with 0.1% BSA and 0.1% Triton-X 100 for permeabilization. The fixed and permeabilized cells were then centrifuged and washed. Antibodies were diluted in PBS with 0.1% BSA and incubated with cells for 20 minutes. The following antibody dilutions were used: Human Fetal Hemoglobin APC (Thermo Fisher MHFH05) 1:10, Anti-Human Fetal Hemoglobin FITC 1:10 (BD Pharmigen 552829), and Anti-Hemoglobin B-(37-8) PE/FITC 1:100 (Santa Cruz Biotechnology SC-21757 PE, SC-21757 FITC). After staining, the cells were centrifuged and washed twice before analysis by flow cytometry.

Fractionation Western Blot

Western Blot Fractionation was performed as described in a previously published protocol (Holden and Horton, 2009). Briefly, cells are harvested and washed in ice cold PBS. Cells are then resuspended in Digitonin buffer containing 150mM NaCl, 50mM HEPES, pH 7.4, and 25ug/ml digitonin and incubated on ice or 10 minutes. Cell mixture was then centrifuged at 2000 RCF. The remaining cell pellet was resuspended in NP40 buffer containing 150mM NaCl, 50mM HEPES, pH 7.4, and 1%NP40 and incubated on ice for 30 minutes. Cell mixture was centrifuged at 7000 RCF and supernatant was kept containing all cytosolic proteins. The remaining nuclei pellet was resuspended in RIPA buffer containing 150mM NaCl, 50mM HEPES, pH 7.4, 0.5% Sodium Deoxycholate, 0.1% Sodium Dodecyl Sulfate, and 1U/ml Benzonase. The mixture is incubated on ice for 1 hour and then centrifuged at 7000 RCF for 10 minutes. Supernatant containing nuclear proteins was kept.

3.7 Supplemental Materials

Supplemental Table 4. Guide sequences for CRISPRi Knockdowns

Gene	Guide Sequence
MB01__HBB 1	TAGACCACCAGCAGCCTAA
MB02__HBB 2	GAACTTCAGGGTGAGTCTA
MB725_ACVR1B_f1	GCGAGCAGGAGGACAACAA
MB726_ACVR1B_f2	gcggcggcggtggTTACTA
MB727_ASS1_f1	AGACGGCGAGTCCGAGAGA
MB728_ASS1_f2	AAGCGGGTGCTCTCGGCTA
MB729_C17orf107_f1	GCAGAGCAGTAGAGTAGCA
MB730_C17orf107_f2	TTAGAGGTGGGGTACTGGG
MB731_CA3_f1	GCACGGAGGAAGGCGACCA
MB732_CA3_f2	CGACCATGGCCAAGGAGTG
MB733_CDKN1A_f1	GCACCGAGGCACTCAGAGG
MB734_CDKN1A_f2	GCCCAGCTCCGGCTCCACA
MB735_CLCF1_f1	AACCTGCGAGTGGGCCTGG
MB736_CLCF1_f2	CCGGACGATCGCAGACCTG
MB737_CYP4F3_f1	GGTGGTGAGTGAGGTCCTG
MB738_CYP4F3_f2	AACAGGGATGGATAAACAG
MB739_DBP_f1	GGGCCGCCCAGCAGCAGAG
MB740_DBP_f2	CTGCAGCTCTCGCAACGGA
MB741_EPHB6_f1	TCCCTCAGGAAGCCACCTA
MB742_EPHB6_f2	TCCACTGCGAGGCCTCGGG
MB743_GFPT1_f1	GCGCCGACACGACTCCCTC
MB744_GFPT1_f2	ATCCCCGAGGGAGTCGTGT
MB745_HMOX1_f1	GAACCAAGCATGCTCCCCC
MB746_HMOX1_f2	GGTCCTCCGTAGCTAGTGA
MB747_IFI30_f1	GGAGCCGCGTGCCCCGACA
MB748_IFI30_f2	ACCCTTGGGCGGGACTAGG
MB749_KLF10_f1	GCGCCGGCCAATGGGCTCG
MB750_KLF10_f2	TGTGTGAGGCGCCGGCCAA
MB751_LMO4_f1	GAGCGAGCGAGCCGGCTAG
MB752_LMO4_f2	CGCGGCGCGGCTTCAGGCG
MB753_MOCOS_f1	AGGGGTCTCCCGCAAGAGA
MB754_MOCOS_f2	CCTAAGCCCCGCTCGCGGA
MB755_MT2A_f1	GGCTCCGAATAACCTAGAA
MB756_MT2A_f2	GGCATCCCCAGCCTCTTAC
MB757_NACC2_f1	GAGGTCGGAGAAAGTCCAT
MB758_NACC2_f2	TCGGCCAGTCGGGTCCCTG

MB759_OSGIN1_f1	CGCTCCTAGCAGGAAAGAG
MB760_OSGIN1_f2	TGTGCACTAGGAGGGTCAG
MB761_PHGDH_f1	CAGCCTGGACCCTTGCTGC
MB762_PHGDH_f2	CTGGCGCGCTAGAGTAACT
MB763_PITX1_f1	CGGTTTCTAAGGCTCCGGA
MB764_PITX1_f2	CCATGGACGCCTTCAAGGG
MB765_PLCL2_f1	TTTGTGCAGGCGGGTCTGC
MB766_PLCL2_f2	TTCGTCCCCGGACACGCCG
MB767_PLS3_f1	TAGGACTCCAGCCCAAGAC
MB768_PLS3_f2	AGCTGGAGAAGGACGACAG
MB769_POLR2J2_f1	GGGGCGATGAGCCGATCCT
MB770_POLR2J2_f2	GATCCTCCAGCCCAACGGG
MB771_PYGB_f1	CGAGCGCGACCCCAAGGTG
MB772_PYGB_f2	GCACACGCGAACGCCGGGA
MB773_RAVER2_f1	TCGCGCGGCGACCTCCTCA
MB774_RAVER2_f2	GTGTCCAGGCTGGGATCGG
MB775_RHOC_f1	GCGGAGCGGAAGGTACCCT
MB776_RHOC_f2	ACGCCCCAGCCCCCAGGGT
MB777_SELPLG_f1	CTCGTGGGCCCAGAAGAAG
MB778_SELPLG_f2	CATGGGACAGCTGCCTCGT
MB779_SRD5A3_f1	GCCATGGCTCCCTGGGCGG
MB780_SRD5A3_f2	CCGCTAGCCCGCGGAAGAG
MB781_TMEM123_f1	CCCCCCGAATGACCAAATA
MB782_TMEM123_f2	TTGTTCCGAGGGCAGGATG
MB783_TMSB4X_f1	ACCCCGCCCGACAACTCGG
MB784_TMSB4X_f2	CCGCCCGACAACTCGGTGG
MB785_TRIB3_f1	CCTGCCCGGGTCCTAGCAG
MB786_TRIB3_f2	AGACTCGCAGCGGAAGTGG
MB787_TSC22D3_f1	CGCGGGGAACAGGGACCGG
MB788_TSC22D3_f2	TGAAGGGAAGTGACTCGAG
MB789_TXNIP_f1	GCTGGCCGGGTGATAGTGG
MB790_TXNIP_f2	CTCCTTTGGAGAAAAAGAG
MB791_ULBP1_f1	CGGAGCTCCAGGTCTACAA
MB792_ULBP1_f2	CCGTGGTGTGAGCCTCGAA
MB793_VCL_f1	GATCCACGGGGCCGGGAAC
MB794_VCL_f2	GAGCGGGGCGGCGAACCGC
MB795_ZFP36_f1	TAGATGGCAGTCAGATCCA
MB796_ZFP36_f2	AGGGGTGTCGGCCGAGAGT
MB797_ATF3_f1	GCTGAAGGGTGCGCTCGGG
MB798_ATF3_f2	GGCCGGCGCGCGGGCTGAA

MB799_ATF4_f1	GACGAAGTCTATAAAGGGC
MB800_ATF4_f2	CATGGCGTGAGTACCGGGG
MB801_CP2_f1	GGAAGGGCCTGGCGAGTCT
MB802_CP2_f2	GAAGGAGAAAGTCCAGCAC
MB803_NFE4_f1	ggtggagggagacgatgaa
MB804_NFE4_f2	gcaagtacctggcacacgt
MB805_SOX6_f1	GCTGCACAGGCAAATGGAG
MB806_SOX6_f2	GCCTGTGCAGCTGATGGAG
MB807_LSD1_f1	GCGCGGGCAGCGTGAAGCG
MB808_LSD1_f2	GCGCGTGCGTACGCGACGG
MB809_AKT1_f1	CGCAGCGCGGCCCGAAGAC
MB810_AKT1_f2	GCCCGAAGACGGGAGCAGG
MB521_HBG2_f1	AGATCATCCAGGTGCTTTA
MB522_HBG2_f2	GCCATGGGTCATTTCACAG
MB523_HBG1_f1	GTAGCAAGGATGGTTCTTA
MB524_HBG1_f2	GGTGTAGGAGAAATCCGGT
MB525_HBZ_f1	tggggaggggaccatggag
MB526_HBZ_f2	acagtggggagggcactgt
MB527_HBBP1_f1	AGCTTGAAAGACTTGGTAA
MB528_HBBP1_f2	TGAGGAGTAAGAGCTGCCG
MB529_HBE1_f1	GCTCTCAGGCCTGGCATCA
MB530_HBE1_f2	CAATCACTAGCAAGCTCTC
MB531_SSUH2_f1	ATGGGAGGCCTAACATCAG
MB532_SSUH2_f2	CACTCACCCCTCCATCCGG
MB533_HIST1H1C_f1	TGCCCAGGCGCTGCTTCAT
MB534_HIST1H1C_f2	GCTGCCGCCGGCTATGATG
MB535_TMEM233_f1	TAGGGGCGTACTGAGACAT
MB536_TMEM233_f2	AGTACCACGAGGAGTGCAG
MB537_CRB1_f1	CTGTGAAGGAGCTGTAAGT
MB538_CRB1_f2	GCTCCTTCACAGAAAATAC
MB539_MYO5B_f1	GCCGCCGCGGCGGGAGTAA
MB540_MYO5B_f2	GGCGCTCGCAACTCCTGTC
MB541_TSPAN32_f1	GGGCCTTGGAGTCGAGTCA
MB542_TSPAN32_f2	CGACTCCAAGGCCCCATGA
MB543_LGR4_f1	AGAGCGTGCAACCCTAGAA
MB544_LGR4_f2	ACGGGAAGAGATTGAGCCG
MB545_TMEM161B_f1	GCTCGGCTCAGACAAAAGG
MB546_TMEM161B_f2	TGTCACCTTGAGGCCCTCG
MB547_FTCD_f1	GAGAGGAGCATCTGGATCA
MB548_FTCD_f2	CTGCCCAGAGAGGAGCATC

MB549_FAM13A_f1	ACGCTTTCTGAGAGAATGG
MB550_FAM13A_f2	TCCCAATGCAAAGGCCCCA
MB551_FZD6_f1	TCCCCCCAGGGACCTGAGA
MB552_FZD6_f2	GCCGTCTCAGGTCCCTGGG
MB553_IQGAP2_f1	CCGCGAGCCTGGCCAGCGA
MB554_IQGAP2_f2	TCTGGCGAGAGAGCACCGA
MB555_CNN3_f1	GCAGCCGGGTGCTTCGCAG
MB556_CNN3_f2	TGGGCGACTGGGTCCAACT
MB557_ASIC3_f1	TACGCCAAGGAGCTCTCCA
MB558_ASIC3_f2	CGGCTCGGGATCCGCACCA
MB559_MALT1_f1	AGCGCGGCACGGCCCTCAG
MB560_MALT1_f2	GCGGCACGGCCCTCAGAGG
MB561_HIST1H2BK_f1	TCTTAAAGCCCAACCCCAA
MB562_HIST1H2BK_f2	TAGTACTGAGTGAGGAATA
MB563_TAF12_f1	TGAGACGAACGCTTCACTG
MB564_TAF12_f2	GAGGGCCGCGCAGTCGGAC
MB565_B9D1_f1	GTGAACGGGCGCCGTTATA
MB566_B9D1_f2	GCGTTGCCCTAGAAACAGA
MB567_SAPCD2_f1	CCCAGCGCGGCCCCACGGA
MB568_SAPCD2_f2	CCGGGCCGGCCCCTCCGTG
MB569_C11orf21_f1	CTCTCAGCGCCGTCCTCAG
MB570_C11orf21_f2	AGCCCAGCCCAAGACACCA
MB571_HIST2H2AA4_f1	CTACTATCGCTGTCATGTC
MB572_HIST2H2AA4_f2	GCCAGGCAGGAGTTTCTCT
MB573_MYL4_f1	AAGCCTGAGCCTAAGAAGG
MB574_MYL4_f2	CAGCTAAGTTGGGAATGGG
MB575_HIST1H2AC_f1	CAAAAATCACCAAAACCAG
MB576_HIST1H2AC_f2	CAAGGGCAAGTGATTTGAC
MB577_BEX1_f1	GCCCCGCAGACCTGCAGAA
MB578_BEX1_f2	TGGTGCCTGCCGAGATCTA
MB579_FES_f1	TTTCAGTTGGCCCAGGCCC
MB580_FES_f2	GGGCTAGAGGTACCCGCAC
MB581_DHCR24_f1	TCCCCGGGCTGTGGGCTAC
MB582_DHCR24_f2	TGGGCTACAGGCGCAGAGC
MB583_IL12RB2_f1	GTGGTGGGCTCAAACGGAA
MB584_IL12RB2_f2	GCGGAGAGCGCGACACGTG
MB585_MVK_f1	GCGACGACACCACGATTCA
MB586_MVK_f2	cggcttcggcgcggagggg
MB587_CAPN2_f1	ACTGCGCAGAGAAAGGTCG
MB588_CAPN2_f2	GGCCGTACTGCGCAGAGAA

MB589_SQLE_f1	GGGCGTGCGACGGTTACTC
MB590_SQLE_f2	ATGCTGGTGAGGAAGCCGT
MB591_DHCR7_f1	CGACCCCTGGCTAGAGGGT
MB592_DHCR7_f2	AGCGAGGCCAGGGGAAGGT
MB593_BAG3_f1	CGCGGGTCGTTCCACGTAG
MB594_BAG3_f2	GGCGGCCCGGCCAGAGACT
MB595_TMC8_f1	CAGCTCACCTGGGGACCTA
MB596_TMC8_f2	TGAGAAGGGGACTCCTCCA
MB597_DUSP6_f1	GCGCGAGGCAGCTCCTCAA
MB598_DUSP6_f2	GCCTACCAGACGCCCCTCG
MB599_WWC2_f1	gggaggaggGTTCGCTCGC
MB600_WWC2_f2	CGGCGGCCTCGCCGGCGAG
MB601_GSTT1_f1	TACAGCTCCAGGCCCATAG
MB602_GSTT1_f2	GTCGGTCGGTCCCCACTAT
MB603_TNFRSF25_f1	cTGAAGGCGGAACCACGAC
MB604_TNFRSF25_f2	TCGGAGGGCTATGGAGCAG
MB605_TRIM10_f1	ATAGTGACCGGCTCCCTCA
MB606_TRIM10_f2	CATCTGTCAGGGTACCCTG
MB607_AIFM2_f1	CGCCTCGGGTCAGTAACTC
MB608_AIFM2_f2	GGCCGGGAAGGGCGAACCG
MB609_PPM1K_f1	CAGGCTTCCAGCTGCGCGG
MB610_PPM1K_f2	AGATCCGAGACGACGGAGA
MB611_SLC30A10_f1	GCCGGCGAGACAATCTGGG
MB612_SLC30A10_f2	CCCGCGCGCAGCCACAGGT
MB705_ATF7IP2_f1	TAACGGCTGCGGTTCGAAG
MB706_ATF7IP2_f2	GCCGGACCAGCATTGTTAA
MB707_ZNF664_f1	GCTCGGAGGCGCACCTGTG
MB708_ZNF664_f2	CACGCGCGCTGTCCCCCGG
MB709_PHTF1_f1	CGCGGGCGAGCGAAGGAGG
MB710_PHTF1_f2	CGGCTGAGTCACTGCATCC
MB711_MED31_f1	GAACTTCGCCCGCGTCTCT
MB712_MED31_f2	GTCTCCATAGCGACAGCAG
MB1126_ATF6_F1	TTAATATCTGGGACGGCGG
MB1127_ATF6_F2	TATTAATCACGGAGTTCCA
MB1128_BAZ2B_F1	GACCGGCGAGGATAAACAG
MB1129_BAZ2B_F2	GAGCCGCGAAAGAGAACGG
MB1130_BCL2L13_F1	TGGCGGCGGCGGTAGATTA
MB1131_BCL2L13_F2	TTCCGACAGGGTCCCCCAG
MB1132_CAST_F1	ACGCAGGGCAGGTAGCCGG
MB1133_CAST_F2	GGAGCGCCACTCACCTCAT

MB1134_CREG1_F1	CCATGGCCGGGCTATCCCG
MB1135_CREG1_F2	AGCGTCGACGCCAGCAGGG
MB1136_DENND4A_f1	CGCGGTAGCGGAACACGAG
MB1137_DENND4A_f2	GCGGATCGAGCTCGGAGAG
MB1138_ELF2_f1	CCAACCGCCTAGGCGACGG
MB1139_ELF2_f2	CCCCCAACCGCCTAGGCGA
MB1140_ELP3_f1	CTACGGCGGCGCAGAAATG
MB1141_ELP3_f2	ATCTCTGGTGGCTCTGCTA
MB1142_FBXO31_f1	TCTCCACCAGCAGCTCGGG
MB1143_FBXO31_f2	CCGAGCGCCCACAGACCCG
MB1144_FBXO42_f1	GGCGGCGGCGTTTGAGTGG
MB1145_FBXO42_f2	GGCGGCGGCGGCGTTTGAG
MB1146_GRB10_f1	CAGAGAGGGAAGCGAGCTG
MB1147_GRB10_f2	GGAGCGCCCAGTCCCTCGG
MB1148_HSPA5_f1	CACGAACCAGGCGAAGGGC
MB1149_HSPA5_f2	GGTCTAGAAATACAGGCCG
MB1150_KAT14_f1	GGCGGCGGTGCGCACTCTG
MB1151_KAT14_f2	TCGGGCGGGCGGGAGAGAG
MB1152_KDM7A_f1	AGCCCGCCCGGCGAACCCG
MB1153_KDM7A_f2	CGGCTGCGCTCGCTCCAGT
MB1154_MLLT10_f1	CTAACGCGGGCGCCCGGAG
MB1155_MLLT10_f2	TCCGCACAGGCGCACTGGG
MB1156_NFATC1_f1	GAGGCTCCGAACTCGCCGG
MB1157_NFATC1_f2	GCCGGGAGAACCGAACCCC
MB1158_NFE2L2_f1	AGGCGCGGCGCGGACAGGG
MB1159_NFE2L2_f2	AGACAAACAACTTTAAAGG
MB1160_OPLAH_f1	ACGGAGCCTGAGCACCGAC
MB1161_OPLAH_f2	CCCACCTGGCGCTCGGCTC
MB1162_SLC7A5_f1	AGCGCGGACGGCTCGGCGC
MB1163_SLC7A5_f2	CACACTGCTCGCTGGGCCG
MB1164_TMBIM6_f1	GCGCTGGTAGGCCTTGGAG
MB1165_TMBIM6_f2	TTGGAGAGGCGGGTTAGGA
MB1166_TREML2_f1	CTGCCTCACTTCCCTAAAT
MB1167_TREML2_f2	CTGCCCATTTAGGGAAGTG
MB1168_URI1_f1	GGCCCAGCCAGGCGTCGCG
MB1169_URI1_f2	CCGCACCGGAGAGGCGTCT
MB1170_WDFY2_f1	TAGCCTGGCCGCCTCTCTC
MB1171_WDFY2_f2	GCGGCGCGCAGCTGAACAC
MB1172_ZDHHC2_f1	GAGCGTCCCGGGACAGTCG
MB1173_ZDHHC2_f2	GTCCCGGGACAGTCGGGGT

MB305_BACH1_F1	CGGCGGACAGGTCCAGTGA
MB306_BACH1_F2	ACAGGTCCAGTGAAggcgc
MB307_BACH2_F1	ACCGCAGCCCGGGCGTGCA
MB308_BACH2_F2	GGCGAGGTGGGCAGTTCGT
MB309_DDIT3_F1	ACGGTCCCTAACTTCACTG
MB310_DDIT3_F2	GACCGTCCGAGAGAGGAAT
MB311_HNRNPD_F1	cgtctggaggcaccgaagg
MB312_HNRNPD_F2	CGGCGCCATTAAAGCGAGG
MB313_KEAP1_F1	GCCCTGGCCTCAGGCGGTA
MB314_KEAP1_F2	GCTGGGCGACAGCAGGCGT
MB315_KLF1_F1	GCGGTCAGTGTGCTGATGG
MB316_KLF1_F2	GCTGTGATAGCCCCTTCGA
MB317_MAFF_F1	GAGCGGAGGGGAGACTGAC
MB318_MAFF_F2	CCCGGTTCAGAGCGACCTG
MB319_MAFG_F1	cgggccgagcgcactggga
MB320_MAFG_F2	GCGCCCGGGTACTGGAGGT
MB321_MAFK_F1	GCCCGGGCGGCACGTGCAG
MB322_MAFK_F2	GCGGCAAGCATGGCCCGGG
MB323_NFE2_F1	GACTGAGAACTCAGGCCGA
MB324_NFE2_F2	GGTCCTCGCAGATTCTCAA
MB325_NFE2L1_F1	AAGCTCCGGCGCCGAGAGT
MB326_NFE2L1_F2	CCCTGGAGGCTAGAAGCTC
MB327_NFE2L2_F1	AGCCCGAGGGCGAACGGGT
MB328_NFE2L2_F2	CACGAGCCCGAGGGCGAAC
MB329_NFE2L2_F3	ACATATATAAAGTACTCAG
MB330_NFE2L2_F4	ATTCTCTTCTGTGCTGTCA
MB331_NFE2L3_F1	GCTGGTGCGAGGGACGAAG
MB332_NFE2L3_F2	TCCGCACGTGTCAccccgg
MB333_NFKBIB_F1	CCCGGCAAAGCCCAGCTAC
MB334_NFKBIB_F2	GCAAAGCCCAGCTACAGGC
MB335_DNJC3_F1	AAGCGCGGGGGGCCTGGTC
MB336_DNJC3_F2	CTGGGCTCCAGAGCCACTG
MB337_YBX1_F1	TCGAACTAGCGAGAATGGC
MB338_YBX1_F2	TCCCCGCGAGAGGGAGTGA
MB339_ZC3H12A_F1	GGCCGGCGGCGTCTCCATG
MB340_ZC3H12A_F2	CGAGTCCTGGGGGTAAGGA
MB115_BCL11A_F1	GCTTGCGGCGAGACATGGT
MB116_BCL11A_F2	GAGAGCCGGGTTAGAAAGA
MB117_ZBTB7A_F1	ccggcgcggcccccggcac
MB118_ZBTB7A_F2	CCGCGCGAGGGAGCGACCA

Conclusion

We have shown that loss of *HBB* leads to robust upregulation of *HBG1/2* in erythroid cells. By analyzing the transcriptomics of WT HUDEP-2 cells and comparing to that of HBBko, we discovered ATF4 to be a causal regulator of upregulating fetal hemoglobin. Knockout of *HBB* leads to decreased ATF4 and upregulation of fetal hemoglobin. ATF4 ChIP-seq in erythroid cells revealed that ATF4 binds and acts upon the intergenic enhancer region *HBS1L-MYB* to activate MYB, a known repressor of *HBG1/2*.

In our CRISPRi studies, we knocked down candidate genes to look for possible repressors of fetal hemoglobin. We discovered several genes to be involved in fetal hemoglobin regulation, and likely to be involved in the response to loss of *HBB*. From our CRISPRi arrayed approach, BACH1 is quite promising as a fetal hemoglobin repressor and further studies will be needed to validate the mechanism of repression.

The fetal to adult hemoglobin switch has been extensively researched, and for good reason. The β-hemoglobinopathies are a model genetic disease that can be ameliorated by the reactivation of fetal hemoglobin. With the developments in CRISPR-Cas9 technologies, new methods of curing or ameliorating β-hemoglobinopathies are now being explored.

This dissertation sheds light on the previously unknown cellular mechanism of how loss of adult hemoglobin leads to increased fetal hemoglobin. We discovered the ATF4 pathway to be involved as well as uncovered potentially new fetal hemoglobin repressors. We hope that our findings presented in this dissertation may help other researchers in their understanding of erythrocyte differentiation as well as help clinical researchers to flip the fetal to adult hemoglobin switch and reactivate fetal hemoglobin in β-hemoglobinopathy patients.

References

Akinsheye, I., Alsultan, A., Solovieff, N., Ngo, D., Baldwin, C.T., Sebastiani, P., Chui, D.H.K., and Steinberg, M.H. (2011). Fetal hemoglobin in sickle cell anemia. Blood.

Albitar, M., Peschle, C., and Liebhaber, S.A. (1989). Theta, zeta, and epsilon globin messenger RNAs are expressed in adults. Blood.

Alter, B.P. (1979). Fetal erythropoiesis in stress hematopoiesis. Exp. Hematol.

Amemiya, H.M., Kundaje, A., and Boyle, A.P. (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Sci. Rep.

Antoniani, C., Romano, O., and Miccio, A. (2017). Concise Review: Epigenetic Regulation of Hematopoiesis: Biological Insights and Therapeutic Applications. Stem Cells Transl. Med.

Bauer, D.E., Kamran, S.C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M.C., Smith, E.C., Pinello, L., et al. (2013). An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science.

Berry, M., Grosveld, F., and Dillon, N. (1992). A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. Nature.

Bhaya, D., Davison, M., and Barrangou, R. (2011). CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. Annu. Rev. Genet.

Bianchi, E., Zini, R., Salati, S., Tenedini, E., Norfo, R., Tagliafico, E., Manfredini, R., and Ferrari, S. (2010). c-myb supports erythropoiesis through the transactivation of KLF1 and LMO2 expression. Blood.

Blais, J.D., Filipenko, V., Bi, M., Harding, H.P., Ron, D., Koumenis, C., Wouters, B.G., and Bell, J.C. (2004). Activating Transcription Factor 4 Is Translationally Regulated by Hypoxic Stress. Mol. Cell. Biol. *24*, 7469–7482.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

Cantu', C., Grande, V., Alborelli, I., Cassinelli, L., Cantu', I., Colzani, M.T., Ierardi, R., Ronzoni, L., Cappellini, M.D., Ferrari, G., et al. (2011). A highly conserved SOX6 double binding site mediates SOX6 gene downregulation in erythroid cells. Nucleic Acids Res.

Cavazzana, M., Antoniani, C., and Miccio, A. (2017). Gene Therapy for β-Hemoglobinopathies. Mol. Ther.

Chen, J.J. (2007). Regulation of protein synthesis by the heme-regulated eIF2α kinase: Relevance to anemias. Blood.

Chen, J.J. (2014). Translational control by heme-regulated eIF2α kinase during erythropoiesis. Curr. Opin. Hematol.

Chen, J.J., and Zhang, S. (2019). Heme-regulated eIF2α kinase in erythropoiesis and hemoglobinopathies. Blood.

Chung, J.E., Magis, W., Vu, J., Heo, S.J., Wartiovaara, K., Walters, M.C., Kurita, R., Nakamura, Y., Boffelli, D., Martin, D.I.K., et al. (2019a). CRISPR-Cas9 interrogation of a putative fetal globin repressor in human erythroid cells. PLoS ONE.

Chung, J.E., Magis, W., Vu, J., Heo, S.J., Wartiovaara, K., Walters, M.C., Kurita, R., Nakamura, Y., Boffelli, D., Martin, D.I.K., et al. (2019b). CRISPR-Cas9 interrogation of a putative fetal globin repressor in human erythroid cells. PLoS ONE.

Crosby, J.S., Chefalo, P.J., Yeh, I., Ying, S., London, I.M., Leboulch, P., and Chen, J.J. (2000). Regulation of hemoglobin synthesis and proliferation of differentiating erythroid cells by heme-regulated eIF-2α kinase. Blood.

Dame, C., and Juul, S.E. (2000). The switch from fetal to adult erythropoiesis. Clin. Perinatol. *27*, 507–526.

Deng, W., Rupon, J.W., Krivega, I., Breda, L., Motta, I., Jahn, K.S., Reik, A., Gregory, P.D., Rivella, S., Dean, A., et al. (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. Cell *158*, 849–860.

DeWitt, M.A., Magis, W., Bray, N.L., Wang, T., Berman, J.R., Urbinati, F., Heo, S.J., Mitros, T., Muñoz, D.P., Boffelli, D., et al. (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. Sci. Transl. Med. *8*.

Dhakshinamoorthy, S., Jain, A.K., Bloom, D.A., and Jaiswal, A.K. (2005). Bach1 competes with Nrf2 leading to negative regulation of the antioxidant response element (ARE)-mediated NAD(P)H:quinone oxidoreductase 1 gene expression and induction in response to antioxidants. J. Biol. Chem. *280*, 16891–16900.

Ding, K., Shameer, K., Jouni, H., Masys, D.R., Jarvik, G.P., Kho, A.N., Ritchie, M.D., McCarty, C.A., Chute, C.G., Manolio, T.A., et al. (2012). Genetic Loci Implicated in Erythroid Differentiation and Cell Cycle Regulation Are Associated With Red Blood Cell Traits. Mayo Clin. Proc. *87*, 461–474.

de Dreuzy, E., Bhukhai, K., Leboulch, P., and Payen, E. (2016). Current and future alternative therapies for beta-thalassemia major. Biomed. J.

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

El-Brolosy, M.A., Kontarakis, Z., Rossi, A., Kuenne, C., Günther, S., Fukuda, N., Kikhi, K., Boezio, G.L.M., Takacs, C.M., Lai, S.L., et al. (2019). Genetic compensation triggered by mutant mRNA degradation. Nature.

Fuglerud, B.M., Lemma, R.B., Wanichawan, P., Sundaram, A.Y.M., Eskeland, R., and Gabrielsen, O.S. (2017). A c-Myb mutant causes deregulated differentiation due to impaired histone binding and abrogated pioneer factor function. Nucleic Acids Res.

Galanello, R., Barella, S., Maccioni, L., Paglietti, E., Melis, M.A., Rosatelli, M.C., Argiolu, F., and Cao, A. (1989). Erythropoiesis following bone marrow transplantation from donors heterozygous for β-thalassaemia. Br. J. Haematol.

Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of T to G C in genomic DNA without DNA cleavage. Nature.

Gil, G.P., Ananina, G., Oliveira, M.B., Costa, F.F., Silva, M.J., Santos, M.N.N., Bezerra, M.A.C., Hatzlhofer, B.L.D., Araujo, A.S., and Melo, M.B. (2013). Polymorphism in the HMOX1 gene is associated with high levels of fetal hemoglobin in Brazilian patients with sickle cell anemia. Hemoglobin *37*, 315–324.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell *159*, 647–661.

Goh, S.H., Lee, Y.T., Bhanu, N. V, Cam, M.C., Desper, R., Martin, B.M., Moharram, R., Gherman, R.B., and Miller, J.L. (2005). A newly discovered human alpha-globin gene. Blood.

Gowen, B.G., Chim, B., Marceau, C.D., Greene, T.T., Burr, P., Gonzalez, J.R., Hesser, C.R., Dietzen, P.A., Russell, T., Iannello, A., et al. (2015). A forward genetic screen reveals novel independent regulators of ULBP1, an activating ligand for natural killer cells. ELife *4*.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Grevet, J.D., Lan, X., Hamagami, N., Edwards, C.R., Sankaranarayanan, L., Ji, X., Bhardwaj, S.K., Face, C.J., Posocco, D.F., Abdulmalik, O., et al. (2018). Domain-focused CRISPR screen identifies HRI as a fetal hemoglobin regulator in human erythroid cells. Science.

Hahn, C.K., and Lowrey, C.H. (2013). Eukaryotic initiation factor 2α phosphorylation mediates fetal hemoglobin induction through a post-transcriptional mechanism. Blood.

Hahn, C.K., and Lowrey, C.H. (2014a). Induction of fetal hemoglobin through enhanced translation efficiency of γ-globin mRNA. Blood.

Hahn, C.K., and Lowrey, C.H. (2014b). Induction of fetal hemoglobin through enhanced translation efficiency of γ-globin mRNA. Blood.

Han, A.P., Fleming, M.D., and Chen, J.J. (2005). Heme-regulated eIF2α kinase modifies the phenotypic severity of murine models of erythropoietic protoporphyria and β-thalassemia. J. Clin. Invest.

Holden, P., and Horton, W.A. (2009). Crude subcellular fractionation of cultured mammalian cell lines. BMC Res. Notes *2*, 243.

Horlbeck, M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y., Fields, A.P., Park, C.Y., Corn, J.E., Kampmann, M., et al. (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. ELife.

Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc.

Huang, P., Peslak, S.A., Lan, X., Khandros, E., Yano, J.A., Sharma, M., Keller, C.A., Giardine, B.M., Qin, K., Abdulmalik, O., et al. (2020). HRI-regulated transcription factor ATF4 activates BCL11A transcription to silence fetal hemoglobin expression. Blood.

Jacob, G.F., and Raper, A.B. (1958). Hereditary Persistence of Foetal Haemoglobin Production, and its Interaction with the Sickle-Cell Trait. Br. J. Haematol.

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M.I., Surdulescu, G.L., Spector, T.D., and Thein, S.L. (2006). cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA – Guided DNA Endonuclease S figs. Science *337*, 816–822.

Kabir, S., Cidado, J., Andersen, C., Dick, C., Lin, P.-C., Mitros, T., Ma, H., Baik, S.H., Belmonte, M.A., Drew, L., et al. (2019). The CUL5 ubiquitin ligase complex mediates resistance to CDK9 and MCL1 inhibitors in lung cancer cells. ELife *8*, e44288.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, a. D. (2002). The Human Genome Browser at UCSC. Genome Res.

Krämer, A., Green, J., Pollard, J., and Tugendreich, S. (2014a). Causal analysis approaches in ingenuity pathway analysis. Bioinformatics.

Krämer, A., Green, J., Pollard, J., and Tugendreich, S. (2014b). Causal analysis approaches in ingenuity pathway analysis. Bioinformatics.

Kuleshov, M. V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res.

Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroyama, T., Miyoshi, H., Tani, K., and Nakamura, Y. (2013). Establishment of Immortalized Human Erythroid Progenitor Cell Lines Able to Produce Enucleated Red Blood Cells. PLoS ONE *8*.

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., et al. (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res.

Lechauve, C., Keith, J., Khandros, E., Fowler, S., Mayberry, K., Freiwan, A., Thom, C.S., Delbini, P., Romero, E.B., Zhang, J., et al. (2019). The autophagy-activating kinase ULK1 mediates clearance of free α-globin in β-thalassemia. Sci. Transl. Med.

Lettre, G., Sankaran, V.G., Bezerra, M.A.C., Araújo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D., Costa, F.F., Hirschhorn, J.N., et al. (2008). DNA polymorphisms at the BCL11A, HBS1L-MYB, and β-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. Proc. Natl. Acad. Sci. U. S. A.

Liang, J.R., Lingeman, E., Ahmed, S., and Corn, J.E. (2018). Atlastins remodel the endoplasmic reticulum for selective autophagy. J. Cell Biol. *217*, 3354–3367.

Liang, J.R., Lingeman, E., Luong, T., Ahmed, S., Muhar, M., Nguyen, T., Olzmann, J.A., and Corn, J.E. (2020). A Genome-wide ER-phagy Screen Highlights Key Roles of Mitochondrial Metabolism and ER-Resident UFMylation. Cell *180*, 1160-1177.e20.

Lingeman, E., Jeans, C., and Corn, J.E. (2017). Production of purified CasRNPs for efficacious genome editing. Curr. Protoc. Mol. Biol.

Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D., and Ma, W. (2019). Methodologies for improving HDR efficiency. Front. Genet.

Liu, N., Hargreaves, V. V., Zhu, Q., Kurland, J. V., Hong, J., Kim, W., Sher, F., Macias-Trevino, C., Rogers, J.M., Kurita, R., et al. (2018). Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell.

Luc, S., Huang, J., McEldoon, J.L., Somuncular, E., Li, D., Rhodes, C., Mamoor, S., Hou, S., Xu, J., and Orkin, S.H. (2016). Bcl11a Deficiency Leads to Hematopoietic Stem Cell Defects with an Aging-like Phenotype. Cell Rep.

Ma, Z., Zhu, P., Shi, H., Guo, L., Zhang, Q., Chen, Y., Chen, S., Zhang, Z., Peng, J., and Chen, J. (2019). PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components. Nature.

Macari, E.R., and Lowrey, C.H. (2011). Induction of human fetal hemoglobin via the NRF2 antioxidant response signaling pathway. Blood *117*, 5987–5997.

Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics *27*, 1696–1697.

Manca, L., and Masala, B. (2008). Disorders of the synthesis of human fetal hemoglobin. IUBMB Life.

Manning, L.R., Russell, J.E., Padovan, J.C., Chait, B.T., Popowicz, A., Manning, R.S., and Manning, J.M. (2007). Human embryonic, fetal, and adult hemoglobins have different subunit interface strengths. Correlation with lifespan in the red cell. Protein Sci.

Martyn, G.E., Wienert, B., Yang, L., Shah, M., Norton, L.J., Burdach, J., Kurita, R., Nakamura, Y., Pearson, R.C.M., Funnell, A.P.W., et al. (2018). Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. Nat. Genet.

Masuda, T., Wang, X., Maeda, M., Canver, M.C., Sher, F., Funnell, A.P.W., Fisher, C., Suciu, M., Martyn, G.E., Norton, L.J., et al. (2015). The LRF/ZBTB7A Transcription Factor Is a BCL11A-Independent Repressor of Fetal Hemoglobin. Blood.

Masuoka, H.C., and Townes, T.M. (2002). Targeted disruption of the activating transcription factor 4 gene results in severe fetal anemia in mice. Blood.

Meletis, J., Papavasiliou, S., Yataganas, X., Vavourakis, S., Konstantopoulos, K., Poziopoulos, C., Samarkos, M., Michali, E., Dalekou, M., Eliopoulos, G., et al. (1994). "Fetal" erythropoiesis following bone marrow transplantation as estimated by the number of F cells in the peripheral blood. Bone Marrow Transplant.

Mohrin, M., Bourke, E., Alexander, D., Warr, M.R., Barry-Holson, K., Le Beau, M.M., Morrison, C.G., and Passegué, E. (2010). Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. Cell Stem Cell.

Morrison, T.A., Wilcox, I., Luo, H.Y., Farrell, J.J., Kurita, R., Nakamura, Y., Murphy, G.J., Cui, S., Steinberg, M.H., and Chui, D.H.K. (2018). A long noncoding RNA from the HBS1L-MYB intergenic region on chr6q23 regulates human fetal hemoglobin expression. Blood Cells. Mol. Dis.

Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., et al. (2012). An expansive human regulatory lexicon encoded in transcription factor footprints. Nature.

Norton, L.J., Funnell, A.P.W., Burdach, J., Wienert, B., Kurita, R., Nakamura, Y., Philipsen, S., Pearson, R.C.M., Quinlan, K.G.R., and Crossley, M. (2017). KLF1 directly activates expression of the novel fetal globin repressor ZBTB7A/LRF in erythroid cells. Blood Adv.

Oikonomidou, P.R., and Rivella, S. (2018). What can we learn from ineffective erythropoiesis in thalassemia? Blood Rev.

Oyake, T., Itoh, K., Motohashi, H., Hayashi, N., Hoshino, H., Nishizawa, M., Yamamoto, M., and Igarashi, K. (1996). Bach proteins belong to a novel family of BTB-basic leucine zipper transcription factors that interact with MafK and regulate transcription through the NF-E2 site. Mol. Cell. Biol. *16*, 6083–6095.

Papayannopoulou, T., Vichinsky, E., and Stamatoyannopoulos, G. (1980). Fetal Hb Production during Acute Erythroid Expansion: I. OBSERVATIONS IN PATIENTS WITH TRANSIENT ERYTHROBLASTOPENIA AND POST-PHLEBOTOMY. Br. J. Haematol.

Park, S.H., Lee, C.M., Dever, D.P., Davis, T.H., Camarena, J., Srifa, W., Zhang, Y., Paikari, A., Chang, A.K., Porteus, M.H., et al. (2019a). Highly efficient editing of the β-globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. Nucleic Acids Res.

Park, S.H., Lee, C.M., Dever, D.P., Davis, T.H., Camarena, J., Srifa, W., Zhang, Y., Paikari, A., Chang, A.K., Porteus, M.H., et al. (2019b). Highly efficient editing of the β-globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. Nucleic Acids Res.

Pimentel, H., Bray, N.L., Puente, S., Melsted, P., and Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. Nat. Methods *14*, 687–690.

Platt, O.S. (2008). Hydroxyurea for the treatment of sickle cell anemia. N. Engl. J. Med.

Platt, O.S., Orkin, S.H., and Dover, G. (1984). Hydroxyurea enhanced fetal hemoglobin production in sickle cell anemia. J. Clin. Invest.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-γuided platform for sequence-specific control of gene expression. Cell.

Rees, H.A., and Liu, D.R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. Nat. Rev. Genet. *19*, 770–788.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol.

Rochette, J., Craig, J.E., Thein, S.L., and Rochette, J. (1994). Fetal hemoglobin levels in adults. Blood Rev.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M., and Stainier, D.Y.R. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature.

Sankaran, V.G. (2011). Targeted therapeutic strategies for fetal hemoglobin induction. Hematol. Educ. Program Am. Soc. Hematol. Am. Soc. Hematol. Educ. Program.

Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K.A., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008). Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science.

Sankaran, V.G., Xu, J., and Orkin, S.H. (2010). Advances in the understanding of haemoglobin switching: Review. Br. J. Haematol.

Schiroli, G., Conti, A., Ferrari, S., della Volpe, L., Jacob, A., Albano, L., Beretta, S., Calabria, A., Vavassori, V., Gasparini, P., et al. (2019). Precise Gene Editing Preserves Hematopoietic Stem Cell Function following Transient p53-Mediated DNA Damage Response. Cell Stem Cell.

Soler, E., Andrieu-Soler, C., De Boer, E., Bryne, J.C., Thongjuea, S., Stadhouders, R., Palstra, R.J., Stevens, M., Kockx, C., Van Ijcken, W., et al. (2010). The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. Genes Dev.

Stadhouders, R., Aktuna, S., Thongjuea, S., Aghajanirefah, A., Pourfarzad, F., Van IJcken, W., Lenhard, B., Rooks, H., Best, S., Menzel, S., et al. (2014). HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. J. Clin. Invest.

Steinmüller, L., and Thiel, G. (2003). Regulation of gene transcription by a constitutively active mutant of activating transcription factor 2 (ATF2). Biol. Chem.

Su, N., and Kilberg, M.S. (2008). C/EBP Homology Protein (CHOP) Interacts with Activating Transcription Factor 4 (ATF4) and Negatively Regulates the Stress-dependent Induction of the Asparagine Synthetase Gene. J. Biol. Chem. *283*, 35106–35117.

Suragani, R.N.V.S., Zachariah, R.S., Velazquez, J.G., Liu, S., Sun, C.W., Townes, T.M., and Chen, J.J. (2012). Heme-regulated eIF2α kinase activated Atf4 signaling pathway in oxidative stress and erythropoiesis. Blood.

Taylor, S.M., Cerami, C., and Fairhurst, R.M. (2013). Hemoglobinopathies: Slicing the Gordian Knot of Plasmodium falciparum Malaria Pathogenesis. PLoS Pathog.

Terns, M.P., and Terns, R.M. (2011). CRISPR-based adaptive immune systems. Curr. Opin. Microbiol.

Thein, S.L., Menzel, S., Peng, X., Best, S., Jiang, J., Close, J., Silver, N., Gerovasilli, A., Ping, C., Yamaguchi, M., et al. (2007). Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. Proc. Natl. Acad. Sci. U. S. A.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and De Laat, W. (2002). Looping and interaction between hypersensitive sites in the active β-globin locus. Mol. Cell.

Vakulskas, C.A., Dever, D.P., Rettig, G.R., Turk, R., Jacobi, A.M., Collingwood, M.A., Bode, N.M., McNeill, M.S., Yan, S., Camarena, J., et al. (2018a). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. Nat. Med.

Vakulskas, C.A., Dever, D.P., Rettig, G.R., Turk, R., Jacobi, A.M., Collingwood, M.A., Bode, N.M., McNeill, M.S., Yan, S., Camarena, J., et al. (2018b). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. Nat. Med.

Wang, X., Angelis, N., and Thein, S.L. (2018). MYB – A regulatory factor in hematopoiesis. Gene.

Weinberg, R.S., Schofield, J.M., Lenes, A.L., Brochstein, J., and Alter, B.P. (1986). Adult'fetal-like'erythropoiesis characterizes recovery from bone marrow transplantation. Br. J. Haematol.

Wienert, B., Martyn, G.E., Kurita, R., Nakamura, Y., Quinlan, K.G.R., and Crossley, M. (2017). KLF1 drives the expression of fetal hemoglobin in British HPFH. Blood.

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 β-Hemoglobinopathies. Trends Genet.

Wienert, B., Wyman, S.K., Richardson, C.D., Yeh, C.D., Akcakaya, P., Porritt, M.J., Morlock, M., Vu, J.T., Kazane, K.R., Watry, H.L., et al. (2019). Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. Science.

Wilber, A., Nienhuis, A.W., and Persons, D.A. (2011). Transcriptional regulation of fetal to adult hemoglobin switching: New therapeutic opportunities. Blood.

Wu, Y., Zeng, J., Roscoe, B.P., Liu, P., Yao, Q., Lazzarotto, C.R., Clement, K., Cole, M.A., Luk, K., Baricordi, C., et al. (2019). Highly efficient therapeutic gene editing of human hematopoietic stem cells. Nat. Med.

Xiang, J., Wu, D.C., Chen, Y., and Paulson, R.F. (2015). In vitro culture of stress erythroid progenitors identifies distinct progenitor populations and analogous human progenitors. Blood.

Xu, J., Sankaran, V.G., Ni, M., Menne, T.F., Puram, R. V., Kim, W., and Orkin, S.H. (2010). Transcriptional silencing of γ-globin by BCL11A involves long-range interactions and cooperation with SOX6. Genes Dev.

Xu, J., Shao, Z., Glass, K., Bauer, D.E., Pinello, L., Van Handel, B., Hou, S., Stamatoyannopoulos, J.A., Mikkola, H.K.A., Yuan, G.C., et al. (2012). Combinatorial Assembly of Developmental Stage-Specific Enhancers Controls Gene Expression Programs during Human Erythropoiesis. Dev. Cell.

Xu, J., Bauer, D.E., Kerenyi, M.A., Vo, T.D., Hou, S., Hsu, Y.-J., Yao, H., Trowbridge, J.J., Mandel, G., and Orkin, S.H. (2013). Corepressor-dependent silencing of fetal hemoglobin expression by BCL11A. Proc. Natl. Acad. Sci. *110*, 6518–6523.

Yeh, C.D., Richardson, C.D., and Corn, J.E. (2019). Advances in genome editing through control of DNA repair pathways. Nat. Cell Biol.

Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIP seeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics.

van Zalen, S., Jeschke, G.R., Hexner, E.O., and Russell, J.E. (2012). AUF-1 and YB-1 are critical determinants of beta-globin mRNA expression in erythroid cells. Blood *119*, 1045–1053.

Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A.H., Shea, D., Gehrke, J.M., Clement, K., Luk, K., Yao, Q., et al. (2020). Therapeutic base editing of human hematopoietic stem cells. Nat. Med.

Zenke-Kawasaki, Y., Dohi, Y., Katoh, Y., Ikura, T., Ikura, M., Asahara, T., Tokunaga, F., Iwai, K., and Igarashi, K. (2007). Heme Induces Ubiquitination and Degradation of the Transcription Factor Bach1. Mol. Cell. Biol. *27*, 6962–6971.

Zhang, S., Macias-Garcia, A., Ulirsch, J.C., Velazquez, J., Butty, V.L., Levine, S.S., Sankaran, V.G., and Chen, J.J. (2019). HRI coordinates translation necessary for protein homeostasis and mitochondrial function in erythropoiesis. ELife.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol.

Zhou, D., Liu, K., Sun, C.W., Pawlik, K.M., and Townes, T.M. (2010). KLF1 regulates BCL11A expression and γ- to β-globin gene switching. Nat. Genet.