2017

THE ROLE OF KLF1 IN REGULATING γ-GLOBIN GENE REPRESSORS

Anna P. Kovilakath
Virginia Commonwealth University

THE ROLE OF KLF1 IN REGULATING γ-GLOBIN GENE REPRESSORS

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

By

ANNA PHILIP KOVILAKATH

Bachelor of Science, Manipal University, India, 2015

Director: Dr. Joyce A. Lloyd

Professor, Vice Chair of Education

Department of Human and Molecular Genetics

School of Medicine

Virginia Commonwealth University

Richmond, Virginia

July 2017

ACKNOWLEDGMENTS

This journey would not have been as wonderful and eventful without the support, the lessons and encouragement received from my parents, siblings, friends and mentors. Words cannot describe how grateful I am to my parents for allowing me to embark on this journey and the constant stream of support they have shown me throughout these two years. My younger siblings Molly and James – you two are the reasons I do this. I feel very fortunate to have worked under the guidance and tutelage of Dr. Joyce Lloyd. As the mentee of Dr. Lloyd I have learned to become an independent thinker and problem solver which will serve me well in my next stage of life as a doctorate student. Aside from this, Dr. Lloyd has taught me to be a better critical reader of scientific papers and has taught me to never be afraid or lazy to ask questions whenever in doubt.

I would like to thank my thesis committee members: Dr. Joseph Landry from the department of Human and Molecular Genetics and Dr. Gordon Ginder from the department of Internal Medicine. They have been invaluable in setting me on the right path and providing the knowledge and insight to ensure that all questions I initially set out to answer have been answered. In addition to sharing their technical and critical expertise, they have been a wonderful source of inspiration to me. Drs. Ginder and Landry have labs adjacent to the Lloyd lab, they are always willing to help and explain any queries or doubts I might have. Special thanks are also due to Dr. Javeria Aijaz, Dr. Xiaofei Yu, Alexander Azzo, and Dr. Kimberly Mayes in the Ginder and Landry labs for sharing their vast research expertise and resources. I would be remiss if I did not mention Joshna Seelam, Farrah Hermes and Mark Hairston for teaching me the basics of Lloyd lab techniques and helping me along the way.

DEDICATION

I would like to dedicate my thesis to my parents.

Mrs. Sheryn Fast Kovilakath and Dr. Philip Thomas Kovilakath

Table of Contents:

List of Figures:

List of Tables:

List of Abbreviations:

β-YAC – Entire β-globin locus yeast artificial chromosome

BCL11A – B-Cell lymphoma/leukemia 11a

BFU-E – Burst-forming unit erythroid

CBP – CREB-binding protein

CD235a – Glycophorin A

CD71 – Transferrin receptor

cDNA – Complementary DNA

CFU-E – Colony-forming unit erythroid

CHD4 – Chromodomain-helicase-DNA-binding protein 4

ChIP – Chromatin immunoprecipitation

CO_2 – Carbon dioxide

Cre – Cyclization recombinase

DRED – Direct repeat erythroid definitive

E – Embryonic day

EDTA – Ethylenediaminetetraacetic acid

Epo – Erythropoietin

EryP-CFC – Primitive erythroid colony forming cells

FBS – Fetal bovine serum

HEL – Human erythroleukemia

HbA – Adult hemoglobin

HbF – Fetal hemoglobin

HbS – Sickle hemoglobin

HPFH – Hereditary persistence of fetal hemoglobin

HS – DNase I hypersensitive site

HSC – Hematopoietic stem cell

HUDEP – Human umbilical cord blood-derived erythroid progenitor

IMDM – Iscove's modified Dulbecco's medium

KDM1A – Lysine-specific histone demethylase 1A

KLF – Krüppel-like factor

KO – Knockout

LCR – Locus control region

LMO2 – LIM domain only 2

LRF – Leukemia/lymphoma-related factor

LSD1/KDM1A – Lysine-specific histone demethylase 1A

MBD2 – Methyl binding domain 2

MEL – Mouse erythroleukemia

MEP – Megakaryocyte / erythroid progenitor

ng – Nanogram

NuRD – Nucleosome remodeling deacetylase

P/CAF – p300/CBP-associated factor

PBS – Phosphate buffered saline

PCR – Polymerase chain reaction

PDS – Fetal bovine plasma derived serum

qPCR – Quantitative PCR

qRT-PCR – Quantitative reverse transcriptase PCR

SCL/Tal1 – Stem cell leukemia / T-Cell acute lymphocytic leukemia 1

SFEM – Serum free expansion medium

shRNA – Short hairpin RNA

YAC – Yeast artificial chromosome

µg - Microgram

µl – Microliter

ZBTB7A – Zinc finger and BTB-domain containing 7a

ABSTRACT

THE ROLE OF KLF1 IN REGULATING γ-GLOBIN GENE REPRESSORS

By Anna P. Kovilakath, M.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2017.

Major Director: Joyce A. Lloyd, Professor, Vice Chair of Education
Department of Human and Molecular Genetics, School of Medicine

Sickle cell disease and β-thalassemia affect millions of people worldwide. γ-globin is the fetal counterpart to the adult β-globin. Research has shown that affected patients with higher than normal γ-globin show less severe symptoms. Therefore, reversing or preventing the hemoglobin switch from γ- to β- globin is a promising avenue of research for treating these diseases.

KLF1 is an erythroid transcription factor involved in hemoglobin switching. Herein, we show that KLF1 directly regulates the γ-globin repressor gene LRF in both the mouse and human systems. KLF1 may also directly activate γ-globin expression by binding the promoter. In human HUDEP-2 cells, an increase in γ-globin expression is seen upon modest knockdown (~50%) of KLF1, whereas normal amounts of KLF1 are observed upon robust knockdown (>75%) of KLF1. The data suggest that KLF1 plays both a positive and negative role in γ-globin expression.

Chapter 1: Introduction

The production of all blood cells and platelets in the body is termed hematopoiesis. It is one of the first processes established in developing mammalian embryos because it involves the synthesis of red and white blood cells (Figure 1.1). In both mice and humans, hemangioblast precursors are synthesized during gastrulation, prior to the formation of hematopoietic progenitors, including primitive erythroid progenitors (Sabin, 1920; Choi et al., 1998; Zambidis et al., 2005; Kennedy et al., 2007). After this, the process of hematopoiesis occurs in three waves throughout mammalian development in various anatomical sites of the body. The first occurs in the yolk sac to produce megakaryocytes, macrophages and primitive erythroid cells and is primitive. The second wave also occurs in the yolk sac to produce megakaryocytes, several myeloid lineages and definitive erythroid cells (Tavian M, 2005; Tober Koniski, 2007; James Palis, 1999), but is definitive. The third wave is also definitive and involves hematopoietic stem cells (HSCs) which are produced within the major arteries of the embryo, yolk sac, and placenta (Lee LK, 2010). These HSCs expand within the fetal liver during fetal development and then move to the bone marrow after birth (Kumaravelu P, 2002).

1.1 Erythropoiesis

The synthesis of erythrocytes (red blood cells) from earlier hematopoietic and progenitor cells is termed erythropoiesis. Erythroid cells are the first cells to be specified and developed in the post-implantation mammalian embryo. Primitive and definitive erythropoiesis are the two types of erythropoiesis in mammals (Figure 1.2) due to the presence of two morphologically and developmentally distinct cell populations in the bloodstream of vertebrates (Gulliver, 1875; Baron and Fraser, 2005; McGrath and Palis, 2005; Fraser et al., 2007; Palis J, 2008).

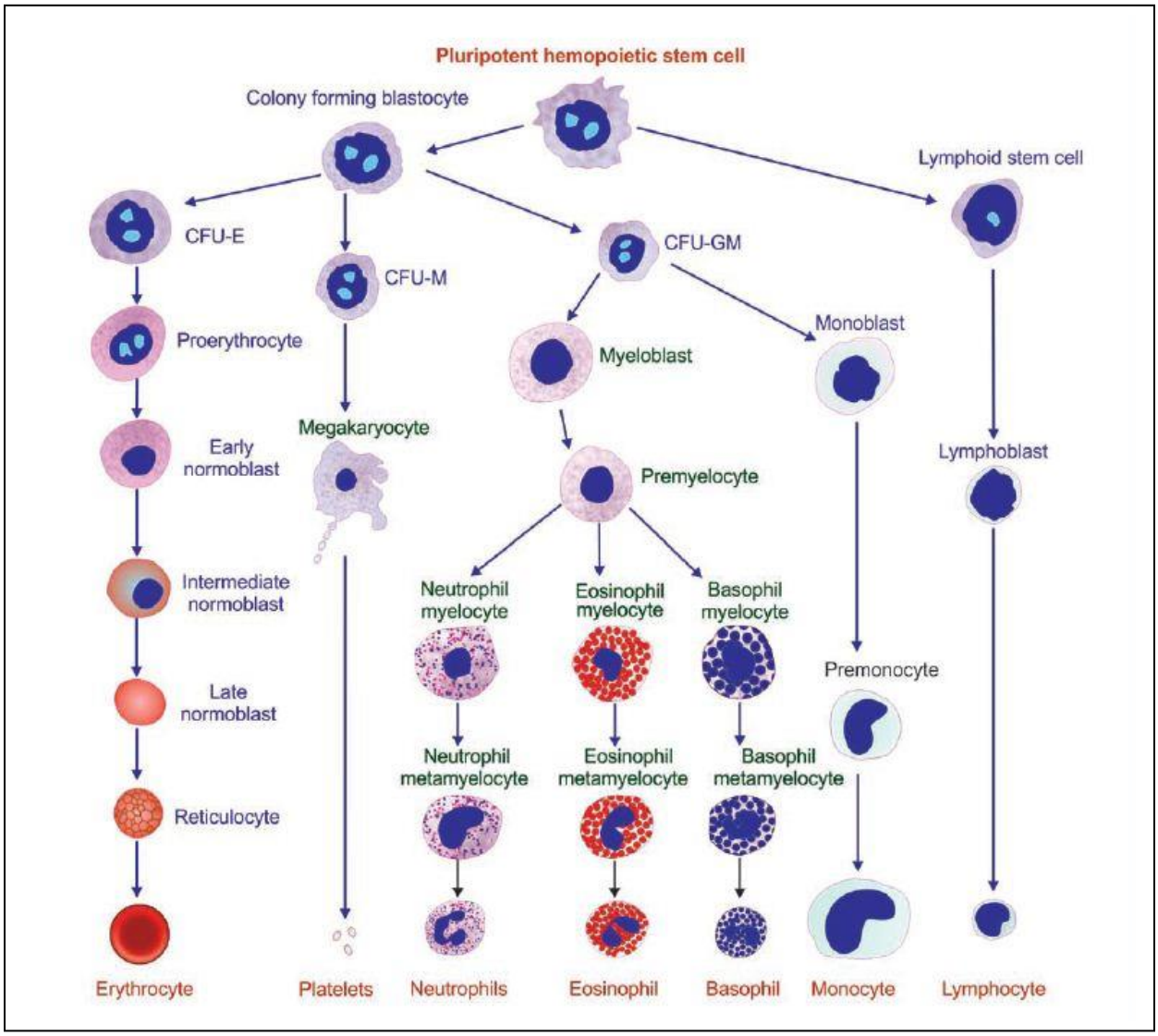


Figure 1.1: Progenitor Cell Lineages of Pluripotent Hematopoietic Stem Cells
Pluripotent hematopoietic stem cells (HSCs) have the ability to differentiate into the progenitors of all types of mature blood cells. CFU-E/CFU-M/CFU-GM = Colony forming Unit-Erythroid/Megakaryocyte /Granulocyte-Monocyte. This study focuses on the lineage from the CFU-E lineage cells and all samples collected in both mice and human systems are at varying erythroblastic stages. Figure adapted from http://nhealthcare.blogspot.com/ 2013/04/red-blood-cell-formation-erythropoiesis.html, accession date: June 19th, 2017

1.1.1 Primitive Erythropoiesis

Primitive erythropoiesis, also termed embryonic erythropoiesis, produces erythroid progenitor colony forming cells (EryP-CFCs) in the blood islands of the yolk sac of the embryo (Wong, 1986; Palis J, 1999; Isern J, 2011) starting from embryonic day 7.5 (E7.5) in mice and two weeks in humans after conception (McGrath and Palis, 2005; Qiu et al., 2008). Primitive erythropoiesis results in the formation of embryonic hemoglobin in vertebrates (Papayannopoulou T et al., 1980; Karlsson and Nienhuis, 1985), further discussion on types of hemoglobin is in section 1.2.1. Blood circulation starts upon initiation of the first heartbeat around E8.25 in mice and in the 8^{th} week of development in humans (Ji et al., 2003; Palis, 2008). The EryP-CFCs differentiate from the mesoderm just after gastrulation and are capable of forming colonies of large, mature erythrocytes which become enucleate. Their differentiated progeny remain in circulation even after birth but are outnumbered by the definitive erythroid cells formed in the growing fetal liver of the fetus (Palis J, 1999; Kingsley PD, 2004).

The maturation of EryPs during development remains poorly understood, but they give rise to primitive erythroid precursors which gradually mature to become smaller, accumulate hemoglobin and undergo nuclear pyknosis and decreased RNA content. Only recently was it discovered that primitive erythroblasts enucleate as a process of maturation into primitive erythrocytes between E12.5 – E16.5 in mice and 18-20 days in humans (Kingsley 2004). Primitive erythrocytes are still present in the bloodstream several days after birth and have an unknown lifespan (Kingsley, 2004; Fraser, 2007).

1.1.2 Definitive Erythropoiesis

Definitive erythroid cells originate from the pool of HSCs in the fetal liver *in utero* and are detected at E11.5 and six-weeks in mouse and human embryos respectively (Kingsley et al., 2006; Qiu et al., 2008). After birth definitive erythroid cells emerge from the HSCs present in bone marrow throughout postnatal life (Orkin SH, 2008; Dzierzak, 2008). Therefore, definitive erythropoiesis results in the formation of both fetal and adult hemoglobin in humans and solely adult hemoglobin in other vertebrates including mice (Maniatis et al. 1980; Karlsson and Nienhuis 1985), further discussion on types of hemoglobin is in section 1.2.1. The earliest definitive progenitor cells termed burst-forming unit erythroid (BFU-E) are small with no particular histology characteristic and express the cell surface marker CD34 which all early hematopoietic progenitors express. They could be considered the counterpart to the EryP-CFCs seen in primitive erythropoiesis. The BFU-Es give rise to the morphologically larger colony-forming unit erythroid (CFU-E) progenitor cells. Both of these definitive erythroid progenitor cells are lineage-committed and are capable of forming colonies of mature erythrocytes (Frame, J.M., 2013).

CFU-E give rise to the first immature precursor cells proerythroblasts which produce hemoglobin. They mature in the sequential order from proerythroblasts (ProE), basophiliic erythroblasts (BasoE), polychromatophilic erythroblasts (PolyE), and finally to orthochromatic erythroblasts (OrthoE). The orthochromatic erythroblasts are smaller, have higher accumulation of hemoglobin, irreversible chromatin condensation and decrease in RNA content similar to the maturation process in primitive erythropoiesis. Maturation of erythroblasts occurs in the fetal liver *in utero* and bone marrow after birth within erythroblasts attached to a central macrophage termed erythroblastic islands (Chasis and Mohandas, 2008).

After nuclear pyknosis, the mature erythroid precursors completely expel their condensed chromatin to become enucleate and form immature erythrocytes, reticulocytes. The final stage of definitive erythropoiesis contains circulating reticulocytes and mature erythrocytes. Reticulocytes are released and circulate in the blood to balance the amount of erythrocytes present in the bloodstream due to removal of senescent erythrocytes after 120 days by splenic macrophages into the spleen (Bennett and Kay, 1981).

To determine whether the primitive or definitive cells are maturing erythroid cells (proerythroblasts), two cell surface marker antigens are used via flow cytometry (Zhang et al. 2003). The two markers used are the transferrin receptor (CD71) and glycophorin A (CD235A) in humans and the transferrin receptor (CD71) and Ter119 in mice.

1.1.3 Regulation of Erythropoiesis

The activity of erythroid-specific and non-erythroid specific transcription factors and cytokines in the overall process of erythropoiesis has been elucidated in studies of their activation in leukemic translocations and in the effect seen in embryonic stem cells upon their disruption (Shivdasani RA, 1996). This is of clinical importance because β-hemoglobinopathies like sickle cell disease and β-thalassemias, (discussed further in Section 1.3.1), affect many people. But patients with one or more of these diseases show ameliorated symptoms if they have naturally occurring mutations that cause HPFH (discussed further in Section 1.3.2). This is why a lot of past and ongoing research has specifically worked on identifying and targeting γ-globin gene repressors to reverse the hemoglobin switch or reactivate fetal globin gene expression (Sankaran VG et al., 2013; Philipsen, 2013).

Tal-1/SCL is a basic helix-loop-helix transcription factor expressed in both primitive and definitive erythropoietic cells, mast cells, megakaryocytes, endothelial cells and T-cell acute lymphoblastic leukemia (Begley CG et al., 1989; Finger LR et al., 1989; Mouthon MA et al., 1993). Its primary role is with erythroid and megakaryocyte lineages differentiation (Mikkola HK, 2003). Knockout of the TAL-1 gene in *in vitro* and *in vivo* leads to *in utero* death from absence of myeloid colony and blood formation (Shivdasani RA, 1995; Robb L, 1995).

LIM-only domain nuclear protein rhombotin 2 (Rbtn2/LMO2) is another transcription factor involved in erythropoiesis (Boehm T, 1991; Warren AJ, 1994). It has been shown that LMO2 complexes with Tal-1/SCL for primitive erythropoiesis to occur (Valge-Archer VE, 1994; Wadman I, 1994). Again, the same phenotype is seen in mice and cells with knockout of this gene (Warren AJ, 1994).

GATA-1 is a zinc-finger transcription factor that binds to the β-globin locus and encodes the GATA-1 protein (Pevny L, 1991). Its expression is limited to erythroid, eosinophil, mast, megakaryocyte lineages and multipotent progenitors (Martin DI, 1990; Romeo PH, 1990; Zon LI, 1993; Zheng J, 2006). The phenotype seen in adult mice without GATA-1 resembles that of aplastic red blood cell crisis in humans whereas in mouse embryos this leads to *in utero* death (Gutierrez, 2008). Constitutive GATA-1 mutations that prevent the interaction between its cofactor FOG-1 result in different types of anemia (Arnaud L, 2010). Germline mutations that truncate N-terminal GATA-1 protein have been associated with inheritance of bone marrow failure, by impairing erythroid differentiation (Sankaran VG, 2012).

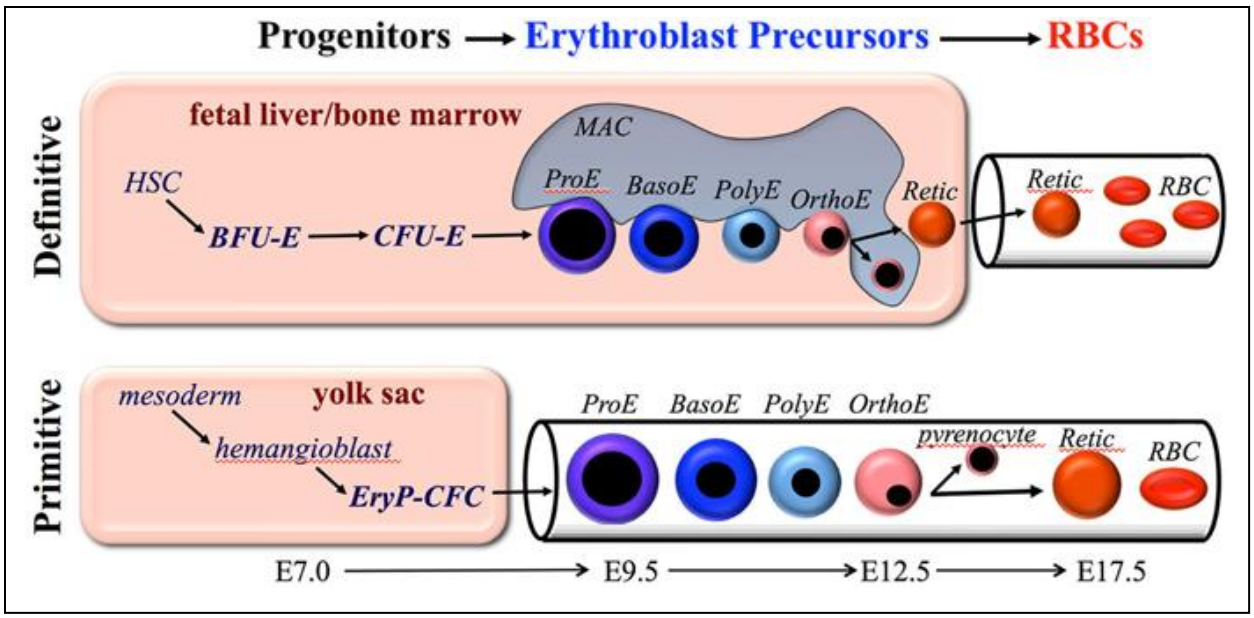


Figure 1.2: Primitive and Definitive Erythropoietic Pathways
HSC = Hematopoietic Stem Cell, BFU-E = Burst-Forming Unit Erythroid, CFU-E = Colony Forming Unit-Erythroid, ProE = Proerythroblast, BasoE = Basophilic erythroblast, PolyE = Polychromatophilic erythroblast, OrthoE = Orthochromatic erythroblast, Retic = Reticulocyte, EryP-CFC = Erythroid Progenitor Colony Forming Cells. Figure adapted from Palis J. Front Physiol. 2014

Unlike GATA-1, GATA-2 is expressed in the yolk sac, fetal liver, bone marrow, spleen and in progenitor cells (Dorfman DM, 1992; Leonard M, 1993; Visvader J, 1993). For this reason GATA-2 regulates genes controlling both stem cell/early progenitor cell proliferation and response to hematopoietic growth factors (Tsai FY, 1994). Heterozygous mutations in the GATA-2 gene lead to a susceptible immune system (Spinner MA, 2014). CD34+ cells isolated from aplastic anemic patients had reduced GATA-2 mRNA expression (Fujimaki S, 2001).

A family of transcription factors known as Krüppel-like factors plays important roles in regulating erythropoiesis. They are further discussed in section 1.4 with the emphasis on KLF1 as this transcription factor is integral to the research presented in the thesis.

1.2 Hemoglobin of Mice and Men

Hemoglobin is the most abundant protein found in erythrocytes responsible for binding and transportation of oxygen or carbon dioxide in the bloodstream of vertebrates. Each subunit of hemoglobin is a globular protein with an embedded heme group. Different kinds of hemoglobin are present in the blood due to composition of different subunits at different stages of development.

1.2.1 The α-globin and β-globin Loci

In humans, The α-family globin genes – ζ2 (embryonic), ζ1 (embryonic), α2 (fetal/adult) and α1 (fetal/adult) reside in the α-globin locus on chromosome 16 (Figure 1.3 A) (Weatherall, 2001; Zhang et al., 2002). The β-family globin genes are arranged in the order of their developmental expression pattern 5'- ε (embryonic) - Gγ - Aγ (fetal) - δ - β (adult) – 3', with the β-globin gene being expressed last, on chromosome 11 (Figure 1.3 A) in humans (S. Philipsen, 2013; Bank et al., 1980; Weatherall, 2001; Crossley M, 1993).

Mice do not express a distinct fetal hemoglobin (McConnell SC et al., 2011) as seen in humans, that is, there is only embryonic and adult hemoglobin synthesized. However, in recent years, many transgenic mouse models have been generated carrying the human β-globin gene locus to mimic human hemoglobins for scientific study and are further discussed in section 1.6.3. The orthologous (to humans) mouse hemoglobins that appear are embryonic $\zeta_2\varepsilon y_2$ and $\zeta_2\beta h1_2$ and adult $\alpha_2\beta_{min}$ and $\alpha_2\beta_{maj}$. The four β-family globin genes – εy (embryonic), βh1 (embryonic), β_{maj} (adult) and β_{min} (adult) reside in the mouse β-globin locus on chromosome 7 (Noordermeer and de Laat, 2008). The same α-family globin genes – ζ2 (embryonic), ζ1 (embryonic), α2 (adult)

and α1 (adult) are present in mice as in humans but on chromosome 11 (Popp, 1981; Tufarelli, 2004).

In humans, embryonic hemoglobin Gower-1 is the first hemoglobin to appear in the very early stages of embryonic development and consists of two ζ-globin chains and two ε-globin chains to form the $\zeta_2\varepsilon_2$ tetramer. Embryonic hemoglobin Gower-2 is the next hemoglobin to appear and is a tetramer of two α-globin and two ε-globin chains to form $\alpha_2\varepsilon_2$ and is followed by embryonic hemoglobin Portland which instead of ε-globin has two chains of γ-globin to form the tetramer $\zeta_2\gamma_2$. These three embryonic globins are derived from the embryonic yolk sac and are expressed in primitive erythropoietic cells. Fetal hemoglobin (HbF) is made in the fetal liver and consists of two α-globin chains and two γ-globin chains to form the $\alpha_2\gamma_2$ tetramer. HbF replaces the embryonic globins and predominates the blood hemoglobin for the remainder of fetal development. Adult β-globin chains arise around the 12th week of development in the bone marrow leading to the progressive increase of adult hemoglobin (HbA) levels $\alpha_2\beta_2$ and the accompanying decrease of HBF. At about six months after birth the hemoglobin in human blood is comprised of 97% HbA, 2% HbA_2 and 1% HbF (Figure 1.3 B) (Bank et al., 1980, Weatherall, 2001, Qiu et al., 2008), this shift from predominantly embryonic globin to HbF and then from predominant HbF to HbA in the bloodstream of humans is termed hemoglobin switching (Figure 1.3 B) and is discussed in more detail under section 1.2.3.

Definitive erythroid cells express the adult β1-and β2-globin genes in mice, and fetal γ- and adult β-globin genes in humans from the β-globin gene cluster. Whereas, primitive erythroid cells express the mouse embryonic εy- and βh1-globin genes and the human embryonic ε- and fetal γ-globin genes from the β-globin gene cluster. The embryonic ζ- and adult α-globin genes of the α-globin gene cluster is expressed in both types of cells.

1.2.2 Regulation of β-Globin Genes

In humans, the β-globin genes are regulated by regulatory elements in promoters of each β-globin gene, i.e., CAAT, TATA and CACCC boxes or sequences (Cao and Moi, 2002) and the locus control regions (LCRs). LCRs are DNA regulatory elements that are responsible for cis-linked genes being highly expressed regardless of LCR location within the locus via stabilization of open chromatin conformation (Grosveld, 1987; Forrester, 1990; Minnie, 1992). The human β-globin locus has an LCR spanning 15 kb located upstream of the ε-globin gene containing 7 DNAse I hypersensitive sites (HS) and controls the differential expression exhibited by the β-globin genes during different stages of ontogeny (Epner, 1992 ; Enver, 1990; Behringer, 1990; Dillon, 1991; Hanscombe 1991). The active sequences of the βLCR are contained within 7 DNase I hypersensitive sites four designated – 5' HS1, HS2, HS3 and HS4 and the remaining three – 3'HS5, HS6 and HS7. Each HS contains DNA-binding elements such as NF-E2, KLF1, and GATA-1 binding sites which are DNA regulatory elements (Tuan, 1985; Forrester, 1986; Wijgerde et al., 1995; Noordermeer and de Laat, 2008). Only one β-globin gene can be activated at one time via joint action between the LCR and promoter regulatory elements. In mice, the two βLCRs can activate either the two embryonic genes or the two adult genes simultaneously (Trimborn et al., 1999; Noordermeer and de Laat, 2008). Reasoning for βLCR action as not just a classical enhancer but as a novel type of regulatory element that controls the entire β-globin locus's genetic activity (Gerardo Jiminez, 1992) was explained by the ability of the βLCR to loop or spatially interact with β-globin genes (Tolhuis et al., 2002).

1.2.3 Hemoglobin Switching

A phenomenon unique to vertebrates termed hemoglobin switching is studied with the ultimate goal of treating hemoglobinopathies. Hemoglobin switching in humans is defined as the change in expression from the predominant embryonic hemoglobin to fetal hemoglobin to adult hemoglobin during development, i.e., two switches (Figure 1.3 B). In mice and other mammals there is only one hemoglobin switch, i.e., from the predominant embryonic to adult hemoglobin during development (Sankaran, 2010; Stamatoyannopoulos, 2005; Kingsley, 2006; Weatherall, 2001). The α-globin chains replace the ζ-globin chains at E12.5 and week 12 *in utero* of mice and humans, respectively (Palis et al. 2010). Transcriptional regulators GATA1, KLF1, SOX6 and BCL11A suppress expression of embryonic globin genes (Sankaran, 2009).

1.3 β-globin Gene Cluster Anomalies

1.3.1 β-Hemoglobinopathies

These diseases become manifested following the hemoglobin switch from HbF to HbA because of decreased synthesis of, absence of, or ineffective β-globin protein production. The most prevalent β-hemoglobinopathies are sickle cell disease (SCD) and β-thalassemia. Currently the number of people living with SCD is approximated to be 100,000 in the USA and 200 to 300 million worldwide (CDC, 2017; Piel FB et al., 2013). SCD is an autosomal recessive disorder caused by a point mutation where glutamic acid (charged amino acid) is substituted by valine (non-polar amino acid) in the 6^{th} codon of the β-globin gene to form the βs mutant chain. This forms abnormal sickle shaped erythrocytes in the blood stream under hypoxic conditions which have difficulty moving through blood capillaries and also reduce oxygen carrying capacity. β-

thalassemia is one of the most common genetic diseases in the world with approximately 15,000 afflicted in the USA and 72 million people worldwide having the major form of the disease (CDC, 2017). There are three types of β-thalassemia based on residual amount of β-globin: β-thalassemia minor with β^+/β or β/β^0; intermedia with β^+/β^+ or β^0/β^+; and major with β^0/β^0, where β^+ is decreased and β^0 is absent β-globin protein.

Phenotypes vary significantly in severity with respect to anemia, organ failure and life span. Due to these reasons, In the 1980s and 90s median lifespan of SCD and thalassemic patients was between 40 to 50 years (Platt OS, 1994). Current standard medical care for thalassemia and SCD patients is aimed at alleviating these symptoms through blood transfusions, iron chelation therapy, treatment with hydroxyurea and/or a combination of these treatments (Agrawal RK, 2014). Life expectancy of sickle cell patients has only increase to ~50 years (Chaturvedi, S et al., 2016), thalassemic patients has increased, and now approaches that of healthy individuals (Diamantidis, M.D. et al., 2016).

Blood transfusions are performed every 3-4 months throughout an affected patient's life. This involves a higher expense, risk of acquiring a transmitted infection and/or severe iron overload. To overcome these issues, iron chelation therapy was developed to remove excess iron from vital organs, but it is not economical and requires trained, professional personnel, which is difficult to acquire in those countries where SCD rates are highest. Hydroxyurea is the only effective drug proven to reduce the frequency of painful episodes seen in SCD patients by 50%. The known mechanism of action in increasing HbF is inhibition of the enzyme involved in the reaction transforming ribonucleosides into deoxyribonucleosides, ribonucleotide reductase, which reduces DNA synthesis overall (Agrawal RK et al., 2014). Inhibited DNA synthesis is a good thing for a short period of time as in the case of hydroxyurea. One-daily dose of

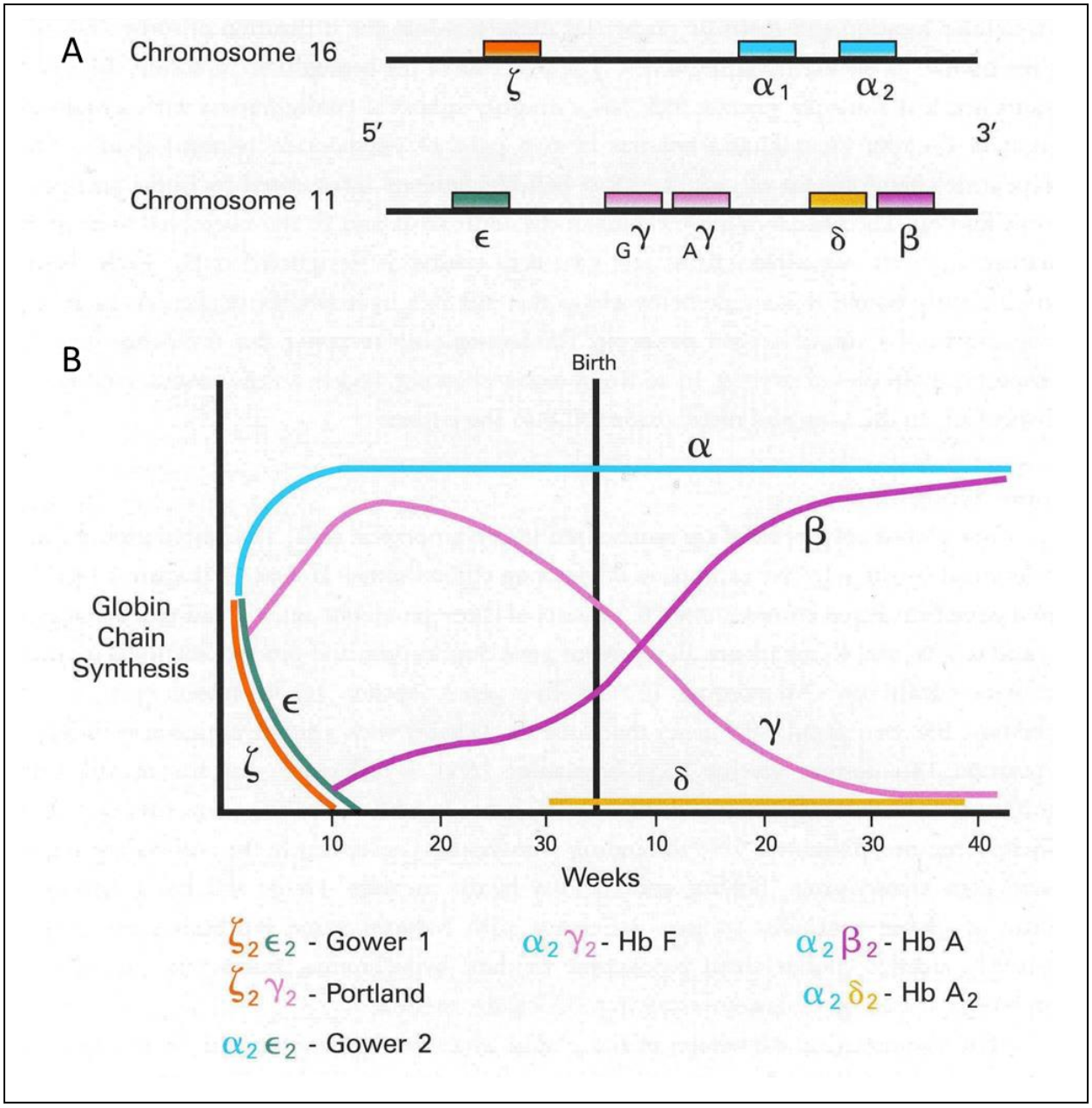


Figure 1.3: Human α- and β-Globin Loci and Hemoglobin Switching
In 1.3 A, the α-globin locus on chromosome 16 in humans and the β-globin locus on chromosome 11 in humans is depicted. In B, The synthesis levels of each individual globin protein is shown from time of conception to adult life, the hemoglobin switch between β- and γ-globin's can be easily seen. The corresponding predominant hemoglobin prevalent is shown below the respective time-period of life. Figure adapted from https://www.studyblue.com/#flashcard/flip/hemoglobintypes/3334284, accession date: June 19th, 2017.

hydroxyurea in patients causes periodic erythroid progenitors and erythroid cell stress signals to be suppressed, which recruits erythroid progenitors that generate erythrocytes with more HbF than normal (Agrawal RK et al., 2014). It is conjectured that side effects of this drug are due to its interference with DNA synthesis causing DNA to be susceptible to chromosome breakage (Mazouzi A et al., 2014). These side effects include but are not limited to neutropenia, bone marrow suppression, elevation of hepatic enzymes and infertility (Brawley OW et al., 2008; Agrawal RK et al., 2014).

Currently, the only viable cure for β-hemoglobinopathies is bone marrow transplantation. Not only is it difficult to find a matching bone marrow donor but it can be complicated by debilitating graft versus host disease (GVHD) or host versus graft disease (HVGD) and is costly (Sankaran and Nathan 2010). GVHD occurs when the donor's bone marrow T-cells reject all of the recipient's cells as "non-self" and destroy the recipient's body. HVGD occurs when the recipient's immune system identifies the donor bone marrow as "non-self" and destroys the transplant tissue. Most patients affected with a β-hemoglobinopathy have normal γ-globin genes (Gγ, Aγ) and it has been shown that reversing the switch from HbF to HbA ameliorates symptoms seen (Weatherall BJ et al., 2001). Patients with greater than 20% γ-globin levels of total β-globin chains require little to no hospital care and suffer no "attacks". This is why research to discover pharmacologic induction of HbF is the next step in a potential cure to treat β-hemoglobinopathies (Trompeter S, 2009; Hankins J, 2009).

1.3.2 Hereditary Persistence of Fetal Hemoglobin

The sustained expression of fetal hemoglobin into adult life is termed hereditary persistence of fetal hemoglobin (HPFH). There has been no direct evidence showing adverse effects related to individuals with HPFH, rather it has been shown to ameliorate symptoms seen in patients affected with β-hemoglobinopathies (Galanello R et al., 1998). First it was thought that HPFH was caused by naturally occurring mutations only in the γ-globin gene on the β-globin locus, but now >50% of all naturally occurring mutations leading to HPFH are due to variants in the β-globin gene, the γ-globin repressor gene BCL11A, the HBSIL-MYB loci and the transcription factor KLF1 (Thein SL et al., 2009; Borg J et al., 2010).

Variants in the β-globin gene promoter cause β-thalassemia minor along with elevated HbF (Huisman TH, 1997; Felber BK, 1982). The more common γ-globin mutations causing HPFH involve small deletions in the 3' sequences of the γ-globin genes by removing suppressor elements, like the deletion in the intergenic region between the Aγ and δ-globin genes (Huisman TH et al., 1974; Gazouli M et al., 2009). Less common are large deletions which bring novel enhancer elements closer to the γ-globin gene promoters (Anagnou NP et al., 1995) and point mutations. The *XmnI* polymorphism is a single nucleotide polymorphism (SNP) C>T at -158 position of the γ-globin promoter with population frequency ranging from 0.32 to 0.35 (Garner C et al., 2000). A 20% increase in HbF is observed in these heterozygous sickle cell patients as compared to the heterozygous general population (Bhagat S, 2012). A rarer naturally occurring point mutation G>A at -117 position of the Aγ promoter displays a 10-20% HbF increase in heterozygotes (Siegel W et al., 1970).

The human population also has less frequent but naturally occurring HPFH –associated mutations caused by deletions ranging from 13 nucleotides to 106 kb (EA Traxler et al., 2016;

Ye L et al., 2016) and by point mutations in the β-globin locus. The overall increase in HbF for total hemoglobin is quite variable ranging from 2% to 41% (Figure 1.4). Reproducing these HPFH mutations to reactivate γ-globin gene expression is a relatively new potential therapy being researched, now possible by *in vitro* by gene editing (Traxler EA et al., 2016; Ye L et al., 2016; Hossain MA et al., 2016), though not yet available as therapy. One of the major caveats of gene editing is off-target mutation induction,

The first case of a KLF1 mutation causing HPFH was studied in a Maltese family and revealed a nonsense mutation that ablated the DNA binding domain of KLF1 on chromosome 19, termed the p.K288X mutation (Siatecka M, 2011). Only carriers or heterozygotes of the K288X mutation showed increased HbF levels in the range between 3.3% – 19.5%. Other KLF1 mutations involved in HPFH were associated with variable but significantly increased ranges of HbF (Borg J et al., 2010; Borg J et al., 2011). Compound heterozygotes of reduced function and truncating loss of function display up to 40% HbF of total hemoglobin (Perkins AC et al., 1996; Perkins AC et al., 2016). The presence of a heterozygous dominant KLF1 mutation c.973G>A was associated with a profound dysregulation of globin gene expression termed congenital dyserythropoietic anemia (CDA) as well as HPFH of ~35% HbF of hemoglobin (Arnaud L et al., 2010).

BCL11A, c-MYB and other γ-globin gene repressors were once considered as potential targets for reactivation of γ-globin gene expression. However, these repressors have been linked to malignancies in humans (Lahortiga I et al., 2007; Satterwhite E et al., 2001, S. Philipsen, 2013). Also, there is a distinct advantage in identifying direct or indirect erythroid-specific repressors (like KLF1) because only erythrocytes would be affected and any other cell types would remain undisturbed (Funnell AP et al., 2015).

1.4 Krüppel-like Factor 1 (KLF1)

1.4.1 KLF Family

KLF1, discovered in 1992, is the first of 17 transcription factors in the Krüppel-like family (Miller IJ, 1993; Turner J, 1999). All KLFs are involved in embryonic and cellular processes and have three characteristic Cys2-His2 zinc finger DNA-binding proteins at their C-terminus (Figure 1.5) which share homology to the fruit fly regulatory patterning protein – Krüppel (McConnell BB et al., 2010). The C-termini are similar in all KLFs and the N-termini vary allowing different

Table 1.1: Single Point and Deletion Mutations in the β-globin Locus Associated with HPFH (Adapted from Carrocini, G.C.D.S et al., 2011)

Mutation Type	Mutation	% of HbF
Single Point	-110 (A -> C) γ^G (Czech)	0.6% – 0.8%
	-114 (C -> G) γ^G (Australian)	8.6%
	-114 (C -> T) γ^G (Japanese)	11.0% - 14.0%
	-175 (T -> C) γ^G	22.1%
	-202 (C -> G) γ^G	18.0% - 23.5%
	-114 (C -> T) γ^A (Georgia)	4.7% - 4.8%
	-117 (G -> A) γ^A (Black-Greek)	10.9% - 15.9%
	-117 (G -> A) γ^A (Greek-Italian)	7.1% - 19.0%
	-158 (C -> T) γ^A (Greek)	2.9% - 5.1%
	-175 (T -> C) γ^A(Black)	36.7% - 38.5%
	-195 (C -> G) γ^A (Brazilian)	4.5% - 7.0%
	-196 (C -> T) γ^A (Italian)	12.0% - 16.0%
	-198 (T -> C) γ^A (British)	1.8% - 13.0%
	-202 (C -> T) γ^A	1.6% - 3.9%
Deletional	HPFH-1 (Black)	21.4% - 28.2%
	HPFH-2 (Ghanian)	21.6% - 27.2%
	HPFH-3 (Indian	21.6% - 23.6%
	HPFH-4 (Italian)	21.0% - 30.0%
	HPFH-5 (Italian	16.0% - 20.0%
	HPFH-6 (Thai)	17.2% - 20.0%

KLFs to act on other genes to repress, activate or both. KLF2, KLF3 and KLF8 are other KLFs that play roles in erythropoiesis. Activity of both KLF3 and KLF8 is mainly transcriptional repression of γ-globin gene expression and other genes due to recruitment of C-terminal Binding Protein co-repressors CtBP1 and CtBP2 (Turner J, 1998; Turner J, 2003; Eaton SA, 2008). The mild phenotype observed in both KLF3-/- and KLF8-/- mice shows that KLF3 and KLF8 are not integral to erythropoiesis (Funnell et al. 2012; Sue et al. 2008; Funnell et al. 2013). KLF1 positively regulates KLF3 and KLF8 in erythroid cells and KLF3 represses KLF8, forming a transcriptional network that may be important for the normal progression of erythropoiesis (Eaton et al. 2008; Funnell et al. 2007). KLF2 directly binds the βLCR, the βYAC transgenic mouse ε-globin and γ-globin gene promoters, and the mouse embryonic εy- and βh1-globin gene promoters suggesting it could have a similar role to KLF1 in regulating embryonic and fetal erythropoiesis (Alhashem YN et al., 2011; Vinjamur DS et al., 2014). KLF2 is also essential for embryonic erythropoiesis, erythroid precursor maintenance and progression of erythroblasts from G to S1 phase of the cell cycle (Vinjamur DS et al., 2014).

1.4.2 Structure, Expression and Functions

KLF1 resides on chromosome 19 and 8 in humans and mice, respectively. The KLF1 protein has three zinc finger domain proteins – ZF1, ZF2 and ZF3 present at the C-terminus and two short transactivation domains – TAD1 and TAD2 present at the N-terminus (Mas C et al., 2011; Raiola L, 2013). The protein has three exons of which parts of exons 2 and 3 encode the zinc finger domain proteins.

Within the bloodstream, expression of KLF1 mRNA is first seen in E7.5 of mouse extraembryonic mesoderm of the yolk sac (Southwood, 1996), is upregulated by E8.5 and then gets downregulated by E12.5 (Isern J, 2010). These phases of KLF1 mRNA expression levels are verified by observation of KLF1 ablated (KLF1-/-) mouse embryos; they become anemic around E11 and ultimately die *in utero* at E16 (Nuez et al., 1995; Perkins AC, 1995; Magor GW et al., 2015; Hodge D et al., 2006; Pilon AM, 2008). Definitive and primitive KLF1-/- erythroid cells at E15 show abnormal morphology, reduced β-globin chain expression and dysregulation of embryonic βh1- and εy-globin genes (Perkins AC, 1996; Basu et al., 2007; Hodge D et al., 2006). This parallels the dependence on the hemoglobin switch from embryonic hemoglobin in the yolk sac to fetal hemoglobin in the fetal liver (Perkins AC, 1995). Expression is seen in both types of erythropoietic cells during embryonic development yet is contained to only the fetal liver, spleen and bone marrow in definitive erythropoietic cells during the later fetal and adult stages of development (Southwood 1996; Miller IJ, 1993). A significant amount of KLF1 is found in the cytoplasm of cells throughout erythropoiesis (Quadrini KJ, 2008; Schoenfelder S, 2010). The first cells to express KLF1 are uncommitted erythroid hematopoietic cells (Frontelo et al. 2007), then at low levels multipotent hematopoietic progenitors (MPP) and common myeloid-erythroid progenitor (CMP). KLF1 is expressed at higher levels in CMPs than in the common megakaryocyte-erythroid progenitor (MEP) (Frontelo et al. 2007; Lohmann and Bieker 2008).

The multitudes of functions performed by KLF1 are quite vast but are all ultimately erythropoiesis-related. Hence, Erythroid KLF (EKLF) was the previous name given to KLF1 and therefore, mainly functions to transcribe globin and non-globin erythroid-expressed genes (Myers RM et al., 1986; Miller IJ et al., 1993; Hartzog GA et al., 1993). KLF1 is responsible for

proliferation of erythroid cells in both primitive and definitive erythropoiesis, though it is present in a lower amount in primitive erythropoietic cells; it is needed for embryonic erythropoiesis (Alhashem et al. 2011; Zhou et al. 2006). These functions include activating the PKLR gene which produces Pyruvate Kinase protein and coordinating the expression of cytoskeletal and membrane proteins, iron processing proteins and heme synthesis enzymes (Drissen et al. 2005; Hodge et al. 2006; Nilson et al. 2006; Pang et al. 2012). The iron processing proteins like TFR2, and ABCB10 (Magor GW, 2015; Tallack MR, 2010) and heme synthesis enzymes such as ALAS2, ALAD, HMBS (Drissen R et al., 2005; Pilon Am et al., 2008; Desgardin AD et al., 2012). It is also required for normal histone modifications and to generate DNAseI hypersensitivity in the promoters of β-globin locus genes to maintain normal chromatin structure. This is done by recruiting histone acetyl transferases (HATs) (Mas C, 2011; Zhang W, 1998), a H3.3 chaperone (Soni S, 2014) and a chromatin remodeling complex (Zhang W, 1998; Bottardi et al., 2006; Armstrong JA, 1998; Kadam S, 2000). It is also responsible for chromatin looping which activates adult β-globin gene expression by bringing the βLCR in closer proximity to the β-globin gene promoter (Nuez B, 1995; Perkins AC, 1995; Drissen R, 2004).

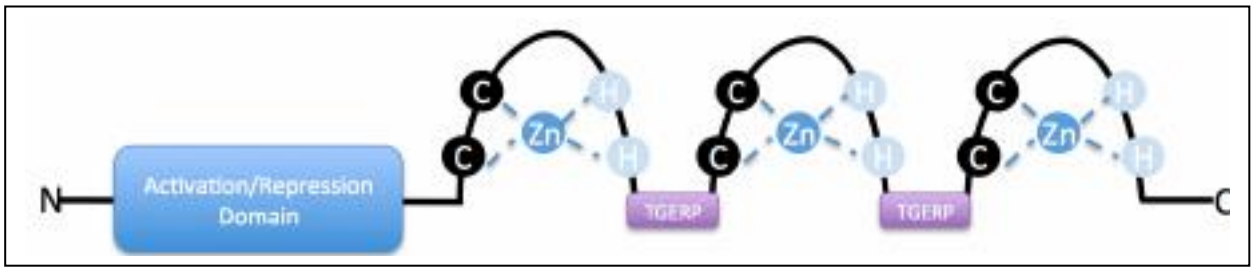


Figure 1.4: Structure of the KLF-Family Members
Adapted from Vinjamur DS Dissertation, 2014

KLF1 regulates the hemoglobin switch, it was seen that KLF1-/- mice carrying the human β-globin locus transgene affected KLF1 binding in the promoter region of the γ- and β-globin

genes (Feng WC, 1994; Donze D, 1995, Wijgerde M, 1996; Perkins AC, 1996). While β-globin gene expression is positively regulated by KLF1, regulation of γ-globin gene expression is more complicated. KLF1 positively regulates γ-globin expression while also negatively regulating it by upregulating γ-globin gene repressors BCL11A and SOX6 (Borg J, 2010; Zhou D, 2010; Tallack MR, 2013).

It was shown in 2007, 2008 and 2009 that KLF1 -/- erythroid cells cannot move from the G1- to S- phases of the cell cycle as KLF1 also regulates cell cycle genes E2f2 and p18INK4c (Tallack, 2007; Pilon et al. 2008; Tallack et al. 2009). Another significant function KLF1 makes is its requirement in co-association of KLF1-regulated genes in nuclear transcription factories, making KLF1 a global regulator in gene expression (Schoenfelder S, 2010).

The transcriptome pattern observed in KLF1 -/- mice showed that KLF1 rarely acts as a transcriptional repressor *in vivo* (Hodge D et al, 2006; Tallack MR et al., 2012), and is indirectly responsible for erythrocyte formation by coordinating expression of a multitude of genes involved in erythrocyte production (Magor GW et al., 2015). KLF1 regulates cell surface proteins such as aquaporins that are involved in erythrocyte membrane integrity. These surface proteins also interact with developing erythroid cells and macrophages in the erythroblastic islands during definitive erythropoiesis (Arnaud L et al., 2010; Singleton BK et al., 2008; Xue L et al., 2014). More recently, KLF1 has been implicated in regulating >20 genes associated with cell signaling and autophagy. KLF1 also indirectly represses γ-globin gene expression and is further discussed in section 1.5.1

In summary, over 120 genes and their respective proteins involved in cell division, cytokinesis, cytoskeleton formation, autophagy, cell signaling, heme and globin formation, etc. are KLF1-dependent (Magor GW et al., 2015).

1.4.3 Variants

KLF1 is vital for normal erythropoiesis to occur and as discussed earlier, total ablation results in *in utero* death. It is therefore important to understand the types of KLF1 variants as they are responsible for a wide range of red cell phenotypes like dysregulated globin expression, hydrops fetalis, non-spherocytic anemia, congenital dyserythropoietic anemia (CDA) and pyruvate kinase deficiency. These variants that alter the protein-coding sequence of KLF1can be categorized into four groups (Figure 1.6).

Variants in class 1 are neutral polymorphism missense mutations occurring outside of the DNA-binding domain. In total there are 11 of these variants with a few occurring in tranactivation domain 1 and 2 (TAD1 and TAD2).

Class 2 variants are hypomorphic missense mutations or small in-frame deletions within the DNA-binding domain with reduced function. The potential mechanisms of class 2 variants include reduced affinity to all if not most *in vivo* DNA-binding sites, off-target interactions and altered protein-protein interaction (Perkins AC et al., 2016). Variants in class 3 are stop codon or frame shift variants resulting in truncated loss-of-function KLF1 protein (Perkins AC et al., 2016). These classes have the most number of total variants because compound heterozygotes carry at least one or the other on one allele or both classes on two alleles. Compound heterozygotes of class 2 and 3 variants are characterized by HPFH with approximately 30% HbF of total hemoglobin in the blood of β-thalassemic patients (Borg J, 2010; Satta S, 2011; Satta S et al., 2012; Huang J et al., 2015).

When KLF1 expression in the blood is low it causes the phenotype of the individual to have the blood group In(Lu) or inhibitor of Lutheran, patients that present with this blood type normally have one normal allele and a class 2 or class 3 variant (Singleton BK, 2008; Helias V,

2013). The KLF1 variant K288X causing HPFH in the Maltese family discussed in section 1.2.3 is in fact a class 3 variant (Borg J et al., 2010). Therefore carriers of class 2 or 3 variants express the In(Lu) and the HPFH phenotypes. Screening for these variants is done by detecting high levels of HbA_2 or $\alpha_2\delta_2$ (Liu D et al., 2014; Perseu L et al., 2011) or elevated zinc protoporphyrin (ZnPP), an inhibitor of the enzyme that degrades heme, in the presence of normal iron (Satta S et al., 2011; Huang J et al., 2015), i.e., when iron levels are normal but is not incorporated into heme. Class 2 and 3 Compound heterozygotes exhibit low pyruvate kinase levels and exhibit premature erythrocyte destruction (Viprakasit V et al., 2014; Huang J, 2015). An example is when Class 2 variant p.A298P is inherited along with a class 3 variant to give the severe NSHA phenotype.

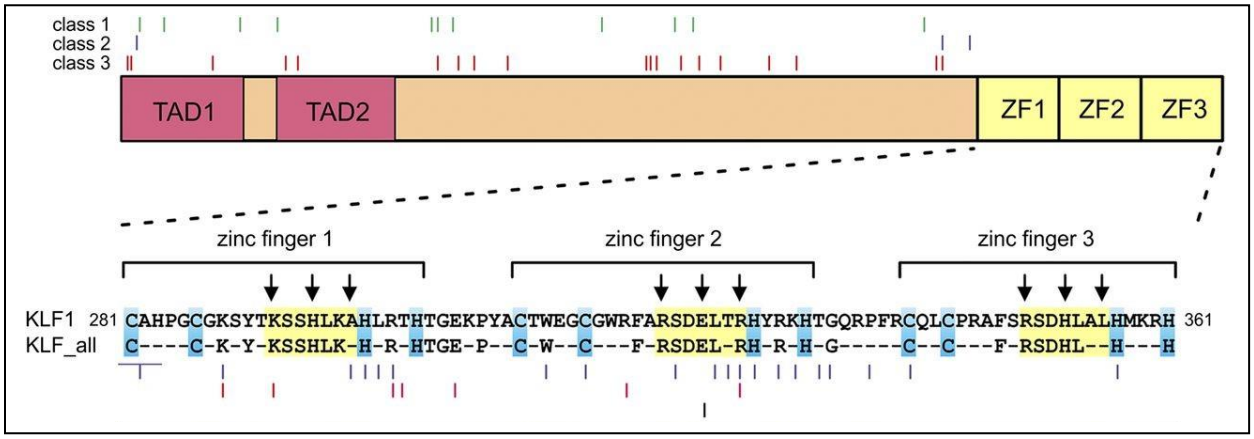


Figure 1.5: KLF1 Functional Domains and Variants
The two transactivation domains (TAD) located at the N-terminus and the three zinc fingers (ZF) located at the C-terminus from the DNA –binding domain are shown. Variants are color-coded: Class 1 is green, class 2 is blue and class 3 is red. The blue highlights are histidine and cysteine residues involved in zinc coordination, yellow highlights are residues that are KLF1-binding sites. Adapted from Perkins AC, 2016, Blood.

Only two variants are categorized under Class 4, which are dominant and the most phenotypically severe of all KLF1 variants. This human KLF1 variant p.E235K causes CDA IV by altering the ZF2 of KLF1 and is generally appears *de novo* (Arnaud L et al., 2010; Jaffray JA, 2013; Singleton BK, 2009). The second variant also expresses a similar phenotype to CDA IV

patients and is seen in mice as a missense mutation p.E339D in ZF2 of KLF1that causes *in vivo* degenerate DNA binding specificity of KLF1 and is further discussed under the neonatal mutation (Nan) mouse model in section 1.6.3. (Huang S, et al., 2016; Planutis A et al., 2017; Heruth DP, 2010; Siatecka M, 2010). Therefore, most monoallelic KLF1 mutations are associated with benign phenotypes which include increased levels of HbF and HbA2 (Liu D, 2014).

The first human KLF1-null neonate ever observed was reported in 2015 which is contradictory to the fetal lethality observed in KLF1 -/- mouse embryos (Perkins AC et al., 1995; Singleton BK et al., 2009; Magor GW et al., 2015). The child in question was born with >70% HbF, brain damage, jaundice and fetal distress all of which required blood transfusions from time of birth. This phenotype was even more severe than phenotype seen in CDA IV from the class 4 variants and termed hydrops fetalis (Magor GW et al. 2015). The parents who were phenotypically normal carried class 3 KLF1 variants – a frameshift p.R319Efs34X variant and a stop codon p.W30X variant. However, the altered gene expression throughout the transcriptome in the human was similar to the expression seen in KLF1 null mice (Magor GW et al., 2015; Hodge D et al., 2006; Pilon AM et al., 2008). It is suggested that persistence of HbF (HPFH) was the only explanation for survival of the KLF1-null neonate.

It makes sense that rates of KLF1 variants are higher than previously thought since KLF1 is the positive regulator of hemoglobin switching (Borg J et al., 2010; Wijgerde M et al., 1996) and prevention or reversal of hemoglobin switching is beneficial to those affected with a form of β-hemoglobinopathy (Weatherall DJ et al., 2001). Another reason for high prevalence of class 2 and 3 variants is the fact that they change the basic morphology and activity of erythrocytes making it more difficult for malarial parasites to propagate. For example, southern Chinese

populations have a very high incidence rate of class 2 and 3 variants combined at ~1.3%. Co-inheritance of KLF1 variants and hemoglobinopathies is common (Yu LH et al., 2015; Liu D et al., 2014). Although DNA sequencing has not been performed in Mediterranean, south east Asia and Africa, it can be assumed that incidence rates of class 2 and 3 variants are high as these are areas where hemoglobinopathies are endemic (Viprakasit V et al., 2014; Perseu L et al., 2011; Satta S et al., 2011; Yu LH et al., 2015; Tepakhan W et al., 2015).

1.4.4 Regulation and Binding

Mouse genome-wide analysis indicates that the core 9 base pair DNA consensus sequence that KLF1 binds to is 5' CCM CRC CCN 3' (R: A/G, M: A/C) (Siatecka et al., 2010; Siatecka M, 2011; Feng et al., 1994; Miller and Bieker 1993; Tallack et al.,2010; Pilon Am et al., 2011). KLF1 regulates β- and γ-globin genes by binding to the sequences CCACACCCT at -90 in the β-globin promoter, and CTCCACCCA in the γ-globin promoter, with greatest affinity for the β-globin promoter (Miller and Bieker, 1993; Donze et al., 1995; Zhang and Bieker, 1998; Dang et al., 2000; Perkins et al., 1999). In mice, KLF1 occupies less than 0.5% of potential KLF1-binding motifs in progenitor and erythroblast chromatin, similar to that seen in GATA1 and TAL1 (Cheng Y et al., 2009; Fujiwara T et al., 2009; Kassouf MT et al., 2010). Of the 945 progenitor- and erythroblast-specific sites occupied by endogenous KLF1, the majority are distant from any known gene which may be explained by association with organization of chromosomal neighborhoods and transcriptional elongation (Chien R et al., 2011; Lee HY et al., 2011).

Binding of KLF1 in the β-globin locus of mice at E10.5 and E11.5 showed that KLF1 is enriched at the εy-, βh1- globin gene promoters and enriched at the mouse 5'HS2 in the βLCR.

In transgenic βYAC mice with the human β-globin locus, KLF1 was enriched at human 5'HS3 and 5'HS2 sites in the βLCR and also enriched at the γ-globin gene promoter. KLF1 binds to all mouse and human embryonic and fetal gene promoters and to the βLCR as shown by ChIP assays (Alhashem YN et al., 2011; Alhashem YN, Dissertation, 2012.). This evidence suggests KLF1 acts directly as a positive regulator of human and mouse embryonic and fetal genes during embryonic erythropoiesis, which is in contrast with research that supports an indirect negative role for KLF1 in regulating γ-globin gene expression (Alhashem et al., 2011; Perkins et al., 1996; Wijerde et al., 1996) (Figure 1.7).

Examining how KLF1 affects histone modifications in wildtype and βYAC transgenic mice on the globin genes we see that, at E10.5, the mouse εy- and βh1 genes and human ε- and γ-globin mRNAs were enriched with H3K9Ac and H3K4me3. Specifically at the βLCR the amount of H3K9Ac and H3K4me3 is generally less than at the globin genes, but is enriched compared to IgG. This is validated when there is a significant reduction in enrichment of H3K9Ac and H3K4me3 in mouse and human embryonic and fetal genes upon ablation of KLF (Alhashem YN et al., 2011; Alhashem YN. 2012). This shows that H3K9Ac is only associated with actively transcribed globin genes in native embryonic erythroid cells. Based on this data, it can be said that KLF1 acts positively on γ-globin gene expression (Alhashem et al., 2011).

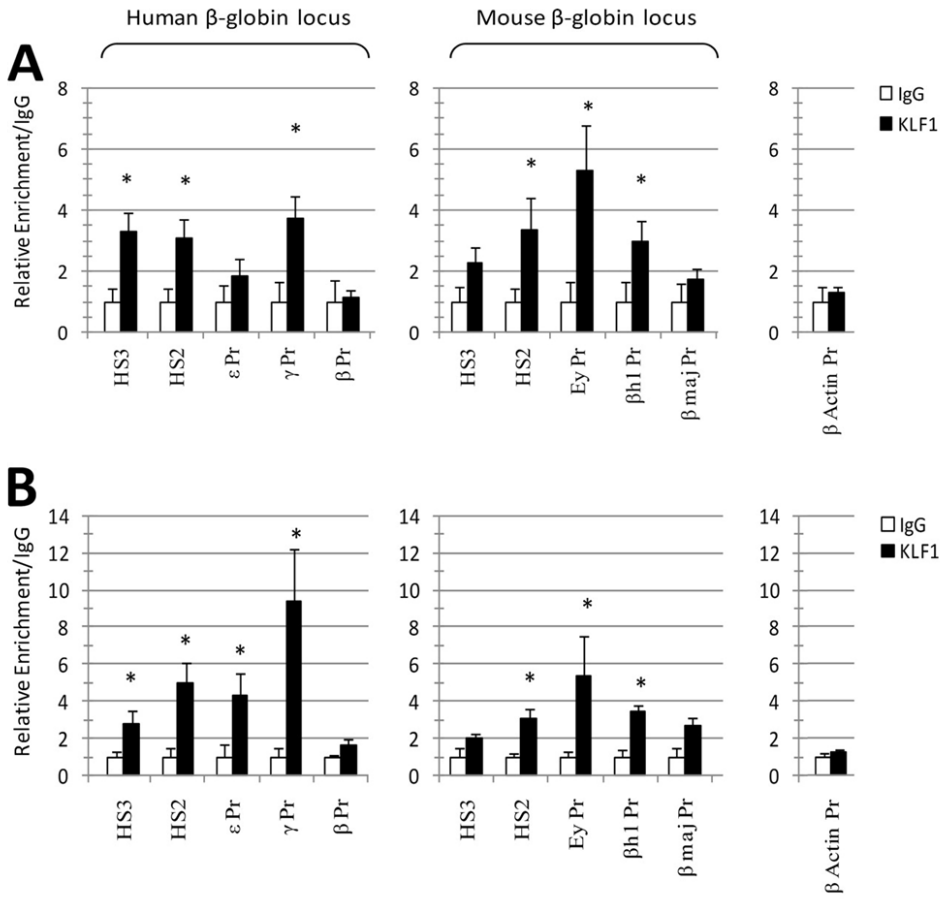


Figure 1.6: KLF1 Binds the Mouse and Human β-globin Loci in Primitive Erythroid Cells

ChIP assays were performed on E10.5 A) or E11.5 B) primitive erythroid cells collected from peripheral blood of normal mice or βYAC transgenic mice. A) In transgenic mice, KLF1 binds HS3 site, HS2 site and γ-globin promoter region in significantly. B) In transgenic mice, KLF1 binds HS3 site, HS2 site, ε-globin and γ-globin-promoters significantly. A) & B) Binding of KLF1 vs. control IgG shows that KLF1 binds HS2 site, Eγ promoter and the βh1 promoter regions in normal mice significantly when compared to IgG. Adapted from Alhashem YN et al., 2011 J. Biol. Chem. (Alhashem YN et al., 2011)

1.5 The γ-Globin Gene Suppression Network

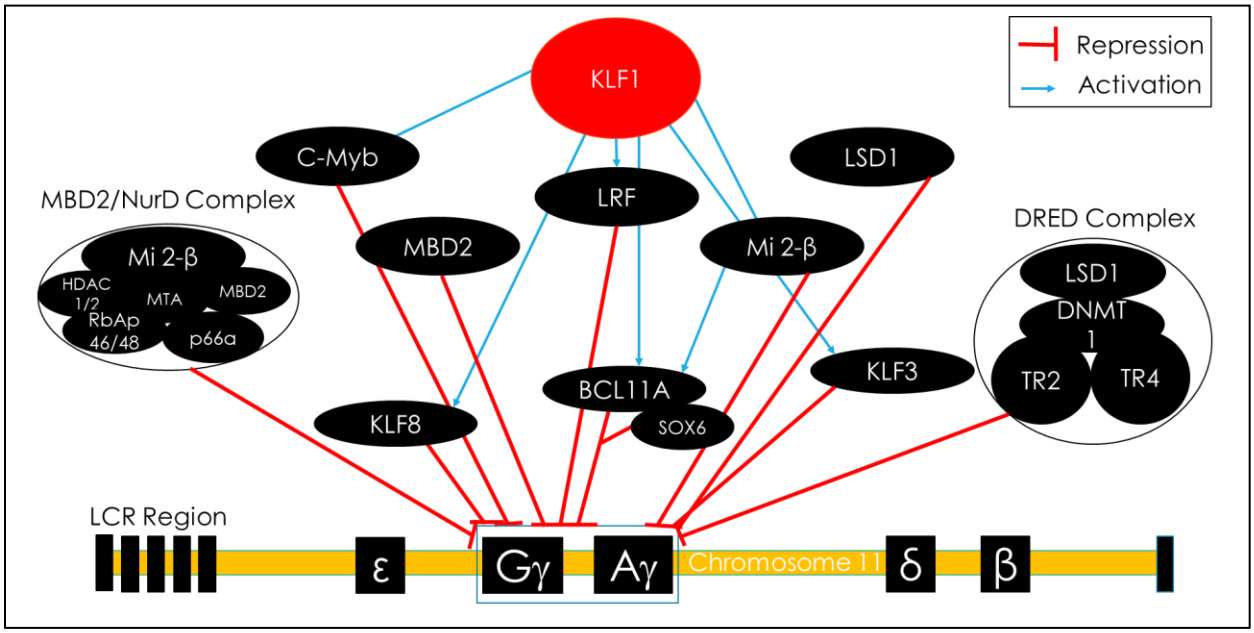


Figure 1.7: The γ-globin Suppression Network

1.5.1 KLF1

The role of KLF1 can interestingly be both negative and positive on γ -globin gene regulation in human and mouse fetal erythroblasts. Previous studies have shown that lentiviral knockdown of KLF1 increases the amount of γ-globin while decreasing the amount of BCL11A (Miller IJ et al., 1993; Pilon AM et al., 2008). KLF1 directly regulates direct γ-globin repressor genes KLF3, KLF8 (Funnell et al., 2013) and BCL11A (Nuez Beatrice et al., 1995; Zhou D et al., 2010). This might indicate that KLF1 directly regulates other γ-globin repressor genes. Human CD34+ precursor cells with a 40-70% KLF1 knockdown show elevated γ-globin transcript levels, but with a > 80% KLF1 knockdown the γ -globin mRNA levels remain normal (Vinjamur DS, 2014; Vinjamur DS et al., 2016). This is termed the threshold amount, which is a residual amount of KLF1 needed by the cell to induce γ-globin above normal amounts (Figure 1.9) (Vinjamur DS, 2014; Vinjamur DS et al., 2016).

1.5.2 BCL11A

One of the major γ-globin repressor genes discovered in primary human erythroid progenitor cells and transgenic βYAC mice was B-cell Lymphoma/Leukemia 11A (BCL11A) (Sankaran VG, 2008; Zhou D, 2010; Borg J, 2010; Vinjamur DS, 2014). The mechanistic action of BCL11A is as follows: It interacts with lysine-specific demethylase 1 (LSD1) and the corepressor element-1 silencing transcription factor (CoREST) complex to silence γ-globin gene in adult erythroid cells (Xu J et al., 2013). The BCL11A-associated protein, DNA methyltransferase 1 (DMNT1), silences HbF in primary human adult erythroid cells (Xu J et al., 2013). Knockdown of KLF1 in both human and mouse adult erythroid progenitors significantly decreases BCL11A levels with an accompanied increase in ratio of γ-globin/β-globin (Zhou D et al., 2010). This suggests KLF1 indirectly represses γ-globin and directly upregulates β-globin. Adult KLF1 KO/WT mice showed a higher percentage of reticulocytes as compared to BCL11A CKO/CKO which has mild anemia. However the percentage of reticulocytes increased and anemia was seen to be more pronounced in compound heterozygotes. The γ-globin expression levels were high during fetal development but decreased in adult KLF1::BCL11A mutant mice, though these levels were still higher than adult wildtype mice (Esteghamat F et al., 2016). Targeting both KLF1 and BCL11A leads to severe anemia and increased percentage of reticulocytes. This is why targeting BCL11A along the KLF1 axis as a therapeutic approach to increase HbF levels in β-hemoglobinopathic patients is not feasible (Sankaran VG et a;, 2009; Sankaran VG et al., 2008; Borg J et al., 2010; Zhou D et al., 2010). Knockdown or knockout of BCL11A causes some concern as BCL11A is expressed during embryonic development and after birth in the central nervous system for neuronal differentiation and morphogenesis and is an

unlikely target for novel therapeutics in treating hemoglobinopathies (Bauer et al., 2012; Leid M et al., 2004; Kuo TY et al., 2010; John A et al., 2012).

1.5.3 LRF

The Leukemia/Lymphoma-related factor or zinc finger and BTB domain containing 7A (ZBTB7A) acts as a proto-oncogene via repression of the p19Arf tumor suppressor (Maeda T et al., 2005). Only recently was LRF found to be another potent silencer of γ-globin gene expression (Masuda T et al., 2016). Transgenic βYAC mice and HUDEP-2 cells with the LRF KO showed upregulated γ-globin gene expression ranging from 23.6 – 35.9% as LRF binds the β-globin gene and also binds immediately downstream of the human HBG1 gene region. It is assumed that loss of LRF from this region facilitates long-range interaction between the γ-globin loci and the βLCR leading to γ-globin reactivation and β-globin downregulation (Masuda T et al., 2016; Tolhuis B et al., 2002). The HUDEP-2 cells lacking either LRF or BCL11A express a 50–60% increase in HbF protein levels determined by HbF over sum of HbF and HbA present as compared to control. But HbF amounts of > 95% are seen in double knockout cells, suggesting that the majority of γ-globin silencing activity is comprised of BCL11A and LRF in adult erythroid cells (Smith EC et al., 2016; Masuda T et al., 2016). LRF has a vast set of functions beyond erythropoietic development which makes this gene an unlikely target for treating β-hemoglobinopathies (Maeda T et al., 2007; Maeda T et al., 2009).

1.5.4 The DRED Complex

The direct repeat erythroid-definitive (DRED) complex is assumed to be a repressor complex that binds with high affinity to direct repeat elements of the human embryonic ε- and fetal γ-globin promoters, as well as to the murine embryonic εy- and βh1-globin promoters (Cui S. et al., 2011; Tanabe O et al., 2002). The complex is comprised of four subunits: the nuclear orphan receptors TR2 and TR4, the DNA methyltransferases (DNMT1) and the lysine-specific demethylase KDM1A or LSD1. This thesis focuses on the LSD1 subunits which will be discussed in detail.

Histone demethylase lysine-specific demethylase 1 (LSD1) is shown to interact with BCL11A in proteomic analysis as mentioned in section 1.5.2 and the TR2/TR4 subunits of the DRED complex (Xu J et al., 2013). Two specific LSD1 inhibitors namely, Tranylcypromine (TCP) and RN-1, induced γ-globin gene expression in human erythroid progenitors and a sickle cell model (Shi L et al., 2013; Rivers A et al., 2015; Cui S et al., 2015). Findings indicate that induction of HbF upon LSD1 knockdown or knockout is associated with impaired erythroid differentiation and induction is lower than induction achieved by knockdown or knockout of BCL11A (Xu J et al., 2013).

1.5.5 The NuRD Complex

The nucleosome remodeling and deacetylase (NuRD) complex is the only known group of proteins coupling both histone deacetylase and chromatin remodeling ATPase activities to mainly be involved in transcriptional repression (Xue Y, 1998; Zhang Y, 2010; Denslow SA, 2007). The complex is comprised of seven subunits: the methyl-CpG-binding domain protein MBD2 or MBD3, chromodomain-helicase-DNA-binding protein CHD3 (Mi2α) or CHD4

(Mi2β), histone deacetylase core proteins HDAC1 and HDAC2, MTA1 or MTA2 or MTA3 which ae the metastasis-associated proteins and the histone-binding proteins RbAp46 and RbAp48 (Feng Q et al., 2001). This thesis focuses on the two NuRD subunits Mi2β and MBD2 which will be discussed in detail.

MBD2 is part of the larger MCBP family responsible for recognizing and directly binding to DNA sequences with methylated cytosines in CpG islands leading to transcriptional repression, including that of γ-globin (Rupon JW et al., 2006; Gnanapragasam MN et al., 2011; Rupon JW et al., 2011; Nan X, 2001; Wade PA, 1999; Feng Q 2001). Transgenic βYAC MBD2 null mice show higher γ-globin expression levels at E14.5 and E16.5 compared to wildtype, but MBD2 has not been shown to bind the γ-globin gene promoter (Rupon JW et al., 2006). This indicates that silencing of the γ-globin gene in adult erythroid cells in mice and human CD34+ cells isolated from umbilical cord blood (discussed in section 1.6.1) is indirectly mediated by the MBD2-NuRD complex via methylation (Rupon JW et al., 2006; Gnanapragasam MN et al., 2011). P66α and MBD2 coiled-coil interaction, found in the MBD2-NuRD complex, facilitate the recruitment of CHD4/Mi2β to competitively silence the globin genes through DNA methylation in both mammalian and chicken erythroid cells (Gnanapragasam MN et al., 2011; Amaya MD et al., 2013).

CHD4/Mi2β is expressed in hematopoietic stem cells and precursors of myeloid, erythroid and lymphoid lineages (Kim J et al., 1999). It acts to repress globin genes as a subunit in the NuRD complex or to specifically repress γ-globin gene expression as a member of the GATA1-FOG-1-Mi2 complex (Williams CJ et al., 2004; Costa Fc et al., 2012). Mi2β is the largest NuRD protein and is solely responsible for chromatin remodeling function of NuRD (Amaya M, 2013; Xu J et al., 2013). It is the major functional component of the γ-globin gene repressor

MBD2/NuRD complex and disruption between p66α and MBD2 interaction displaces both MBD2 and Mi2β from the complex, causing γ-globin gene expression (Gnanapragasam et al, 2011). Knockdown of Mi2β induces higher expression of γ-globin gene than do knockdowns of the MBD2 or GATA-1-FOG-1-Mi2 complexes. Mi2β knockdown is also associated with reduced expression of KLF1 and BCL11A in murine CID cells, human primary erythroid cells, HUDEP-2 cells and βYAC mice independent of the MBD2/NuRD complex (Borg J et al., 2010; Zhou D et al., 2010; Sankaran VG et al., 2009; Amaya M et al., 2013).

1.5.6 c-MYB

The transcription factor c-MYB is highly expressed in immature hematopoietic cells and downregulated during differentiation (Bianchi E et al., 2010). Analysis of transcription factor c-MYB transfected K652 clones demonstrated that c-MYB negatively regulates production of HbF (Jiang J et al., 2006). C-MYB knockout in mice and human CD34+ erythroid precursor cells result in *in utero* death and presence of only mature megakaryocytes and macrophages, whereas, knockdown results in perturbed erythropoiesis differentiation (Emambokus N et al., 2003. Mucenski ML et al., 1991). It was shown that c-MYB binds to the promoter and transactivates expression of KLF1 through chromatin immunoprecipitation and luciferase assays (Bianchi E et al., 2010). Cells overexpressing KLF1 had partially rescued erythropoiesis caused by silencing of c-MYB. This transcription factor plays a key role in regulating stem cells of the bone marrow, colonic crypts and neuronic region in adults. But tinkering with c-MYB is likely to adversely affect hematopoiesis and erythroid differentiation as c-MYB potentially have been implicated in human malignancy, and several animal species transplanted with c-MYB-deficient cells died from different forms of leukemia (Ramsay RG et al., 2008).

1.6 Model Systems Used in Studying γ-globin Gene Regulation

Over the decades many different *in vitro*, *in vivo* and *ex vivo* models have been used to study erythropoiesis.

1.6.1 Immortalized Human Cell Lines

Cell lines in the erythropoietic field are often used to study biological processes over their primary cell counterparts because primary erythropoietic precursor cells are difficult to culture, maintain and manipulate. Therefore experiments can also be performed using hematopoietic progenitor cell lines such as the human erythroleukemic K652 and HEL, and mouse MEL cell lines and the more recent immortalized erythroid progenitor cell lines HUDEP and HiDEP (Lozzio and Lozzio et al., 1975; Friend et al., 1957; Kurita Ryo et al., 2013). The main disadvantages to using the erythroleukemic cell lines is that any results seen may be due to the abnormal chromosomes present inherently in the cell line. It must always be kept in mid that when interpreting results from immortalized cell samples you must be careful as cell lines do not always accurately replicate the primary cells (Kaur G, 2012).

The first erythroid progenitor cell line termed murine erythroleukemia (MEL) was derived from the spleens of Friend virus complex infected mice immortalized by infecting mouse splenic cells with a virus (Friend et al., 1957; Antoniou et al., 1991). Less than 1% of the MEL cells spontaneously differentiate and so induction of differentiation to express mouse adult hemoglobin $\alpha_2\beta_{maj2}$ was performed by treatment with dimethyl sulfoxide (DMSO) (Friend et al., 1971; Antoniou et al., 1991).

K652 cells were established from an erythroleukemic patient in a blast crisis (Lozzio and Lozzio, 1975). These cells are non-adherent, rounded and have two reciprocal translocations the

bcr:abl fusion gene (Philadelphia chromosome) present on chromosome 22 and the translocation between long arm of chromosome 15 with 17 (Lozzio et al., 1975). Deacytylase activity induces cell differentiation into mature erythrocytes, monocytes or macrophages (Duncan MT, 2016). The erythroid cells formed upon induction of differentiation exhibit an embryonic/fetal pattern of globin gene expression, rather than expressing adult hemoglobin, which is one of the major disadvantages of using these cell lines.

The human erythroleukemia (HEL) cell line was first established in 1982. These cells have induced and spontaneous production of mainly G- and A-γ chains. Expression of the β- and ε-globin chains is negligible. For some time, this cell line was useful in research studying erythroid cell differentiation and differential globin gene expression. HEL cells were derived from the peripheral blood of a Hodgkin's patient with erythroleukemia. Cells resemble large abnormal proerythroblasts with increased frequency of polyploidy (Martin P et al., 1982).

The most recent cell line to be established is the human umbilical cord blood-derived erythroid progenitor (HUDEP) cell line from three different cord blood samples termed HUDEP-1, HUDEP-2 and HUDEP-3. A tet-inducible HPV16-E6/E7 transgene lentivirus was used to immortalize umbilical cord blood CD34+ cells, by inhibiting p53 and retinoblastoma tumor suppressor proteins (Kurita R, 2013; Münger K, 2002). The differences between the three HUDEP cell lines arise from variation in expression of β-globin genes. HUDEP-1 and -3 are characteristically more fetal-like by predominantly expressing γ-globin. Whereas, the HUDEP-2 cell line most closely resemble adult erythroid cell pattern based on β-globin gene and γ-globin repressor BCL11A expression. Development of this system has made studies on adult erythropoiesis using immortalized cell lines possible for the first time (Canver MC et al., 2015; Masuda T et al., 2016; Traxler EA et al; 2016).

1.6.2 Primary Human Cells

Many dependable, refined methods have been developed over the years for isolation of fetal or adult hematopoietic $CD34^+$ precursor cells from human umbilical cord blood, peripheral blood and bone marrow. They have become powerful tools in studying γ-globin gene repression. The main idea is to expand these few progenitor cells to obtain a large number of erythroid cells (Migliaccio et al., 2002). These methods have become very popular because the $CD34^+$ cells can be easily *in vitro* differentiated along an erythroid pathway, while simultaneously performing knockdown of a gene of choice. The use of lentiviral vectors that express small hairpin RNA (shRNA) is an efficient method to downregulate genes and hence gives the researcher more flexibility to perform their experiments. The more recent methods to isolate $CD34^+$ cells from human umbilical cord blood, bone marrow or peripheral blood are considered the best available cellular model of erythroid development (Migliaccio et al., 2002; Giarratana et al., 2011). The expression of the active form of BCL11A is negligible in these cells and (Sankaran VG, 2008). The drawbacks lie in their relatively limited lifespans and terminal erythroid differentiation while also maintaining cell viability (Ginder GD, 2015).

1.6.3 Mouse Models with modified KLF1

The processes of primitive and definitive erythropoiesis are very similar in humans and mice which make using mice to perform *in vivo* studies on research in β-hemoglobinopathies, the global role of KLF1, membrane skeleton protein functions and regulation of human β-globin genes plausible (Peters LL et al., 2001). For this reason mutant, transgenic mouse and knockout mouse models have been developed.

In 1992, the first mouse model carrying the entire human β-globin locus was developed with ligation of two cosmid constructs (Strouboulis J et al., 1992), which is a vector that can only carry DNA fragments ranging from 37 to 52 kb. A year later with the advent of yeast artificial chromosomes (YAC), which is a vector that can carry DNA fragments ranging from 100 to 1000 kb, the βYAC transgenic mice were developed (Gaensler KM et al., 1993; Peterson KR et al., 1993). These models express human globin genes in a manner that resembles the expression if endogenous mouse genes. Since mice do not produce fetal hemoglobin as seen in humans, the human embryonic ε-globin gene is expressed in the cells derived from the yolk sac. Fetal γ-globin genes and the adult β- and δ-globin genes are expressed in the cells derived from the fetal liver and bone marrow, respectively. A method for how to overcome this major drawback to performing *in vivo* studies on βYAC transgenic mice is yet to be discovered.

A naturally occurring mutation in mice called neonatal anemia (Nan) is a semi-dominant hemolytic anemic phenotype present on chromosome 8 first described in 1985 (Lyon MF, 1986). At E11 Nan homozygotes die due to lack of hematopoiesis. To further study this mutation Nan mutant models were created to pinpoint the location of the mutation which was determined to be a single amino acid change E339D within the second zinc finger of KLF1 (Siatecka et al., 2010).

The first KLF1-/- knockout mouse models were created in 1995 by either inserting a) the neomycin antibiotic resistance gene into the second of three exons of KLF1 to disrupt the zinc finger binding domain (Perkins AC et al., 1995) or b) a LacZ reporter gene and neomycin antibiotic resistance gene cassette in the ATG codon, which would prevent initiation of translation of KLF1 protein (Nuez et al., 1995). This study is using the KLF1 knockout mouse model developed by Perkins AC et al. in 1995. Only through the construction of these knockout

mice have researchers been able to discover what role KLF1 plays in the body as previously discussed in section 1.4.1.

1.7 Rationale

The β-hemoglobinopathies, sickle cell disease and β-thalassemia, are major sources of morbidity and mortality worldwide. Understanding the underlying mechanisms involved in the switch from fetal to adult hemoglobin and reactivation of γ-globin gene expression in adults is vital in discovering a cure for these diseases.

Based on previous research, multiple sometimes seemingly contradictory roles for KLF1 in γ-globin gene regulation have been put forth. KLF1 directly upregulates (Alhashem et al., 2011; Perkins AC et al., 1995; Perkins AC et al., 1996; Wijgerde et al., 1996), indirectly downregulates (Perkins AC et al., 1995; Funnell et al., 2013; Nuez B et al., 1995; Zhou D et al., 2010) or directly downregulates (Miller IJ et al., 1993; Pilon AM et al., 2008) γ-globin gene expression. Elucidating the positive and negative roles of KLF1 in regulating γ-globin gene expression as a first step, will help in clarifying this ongoing debate. This will be done by examining five γ -globin gene repressors LSD-1, LRF, c-MYB and the NuRD complex genes Mi2β and MBD2. As discussed in section 1.5, knockdown of c-MYB (Ramsay RG et al., 2008), LSD1 (Xu J et al., 2013), BCL11A (Bauer et al., 2012; Leid M et al., 2004; Kuo TY et al., 2010; John A et al., 2012) and Mi2-β (Larsen DH et al., 2010) with the end goal of increasing HbF has been shown to adversely affect processes other than erythropoiesis. LSD-1 has a minor role in repressing γ-globin and is therefore alone is not an efficient target of therapy. Testing knockdowns of KLF1, which is expressed in erythroid cells, will only affect processes involved in erythropoiesis and have less chance of unwarranted adverse effects. Utilizing the ~40-70% knockdown of KLF1

which has no known adverse effects on model organisms would help us understand the comprehensive role it plays in regulating γ-globin gene expression via these repressors. For example, a patient with sickle cell disease had ameliorated symptoms and 20.3% HbF, due to a heterozygous mutation that resulted in the genotype: c[914-1_914-4 del CTAG] that deleted 4 base pairs of one of his KLF1 genes (Gallienne AE et., 2012).

The intent of this research is to decode the indirect mechanism of action of KLF1on γ-globin repression. Therefore, ~40% to 70% knockdown of KLF1 will be utilized while performing experiments in mice and HUDEP-2 cells on the five prominent repressors genes of γ-globin gene expression: CHD4/Mi2β, LRF/ZBTB7A, LSD1/KDM1A, MBD2 and c-MYB. It will be of interest to look at the role of KLF1 in regulating these γ-globin gene repressors and to study the evolutionary conservation of this mechanism between mice and humans for more insight into the regulation of hemoglobin switching.

Therefore, we hypothesize that KLF1 actively regulates transcription of some of these repressors by directly binding to their promoter regions in both mouse fetal livers and the human HUDEP-2 cell line.

It was previously shown that γ-globin mRNA levels were significantly elevated with ~40-70% knockdown of KLF1 when compared to the control. But, the γ-globin mRNA levels were not significantly increased or decreased with >80% knockdown of KLF1 when compared to the control. This phenomenon is termed the threshold amount of KLF1, i.e. a certain amount of KLF1 is required by cells to produce γ-globin levels significantly greater than normal. We also hypothesize that the KLF1 knockdowns will result in the same relative amounts of γ- and β-globin protein as that of mRNA.

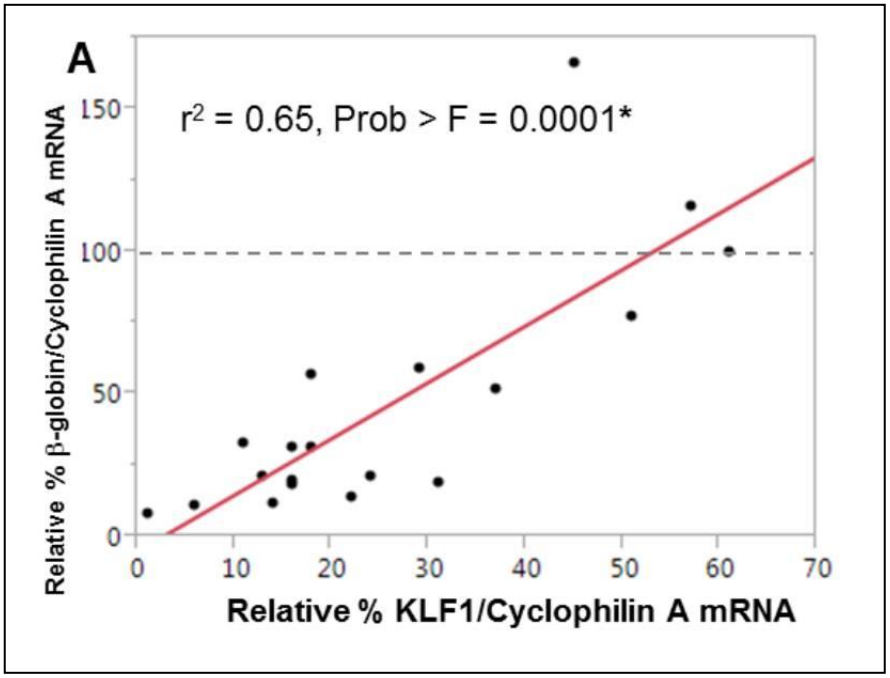


Figure 1.8: Regulation of β-Globin Expression by KLF1 in Erythroid Cells Derived from Human Umbilical Cord Blood CD34+ Hematopoietic Progenitors Follows a Linear Pattern

There is a positive correlation observed between the amount of KLF1 mRNA present on the Y-axis and the corresponding amount of β-globin expression by linear regression analysis. Adapted from Vinjamur DS, 2016, PLoS ONE.

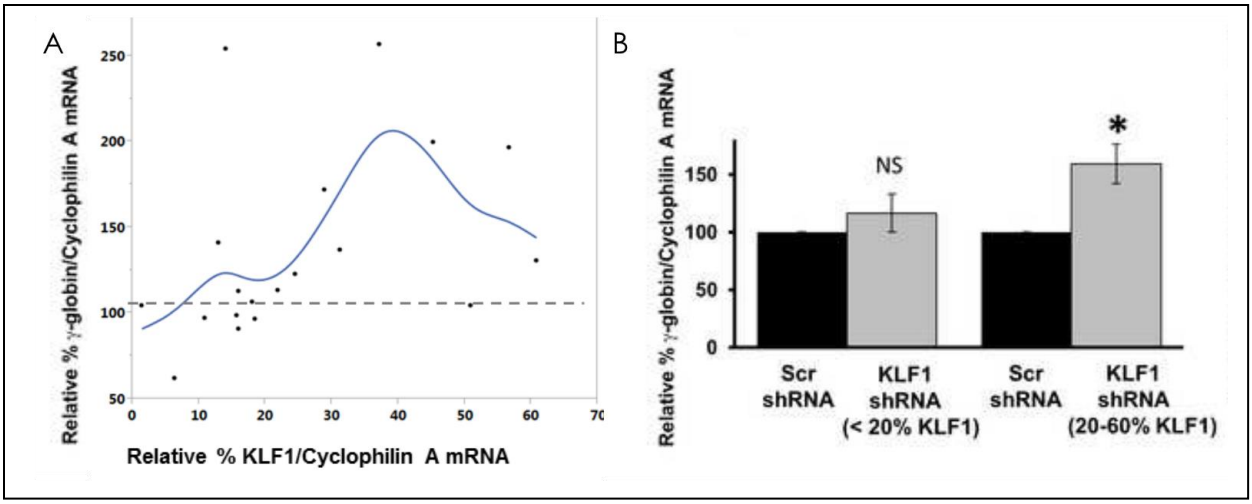


Figure 1.9: Regulation of γ-globin Expression by KLF1 in Erythroid Cells Derived from Human Umbilical Cord Blood CD34+ Hematopoietic Progenitors Follows a Non-Linear Pattern.

A) No correlation was observed between the amount of KLF1 mRNA present in cells and the corresponding amount of γ-globin expression by linear regression analysis. A non-linear model made the smooth curve shown in the figure. B) A statistically significant increase in γ-globin mRNA levels was observed when compared to control. Adapted from Vinjamur DS, 2016, PLoS ONE.

Chapter 2: Materials and Methods

2.1 Studies in the Mouse and HUDEP-2 Cell Line

2.1.1 Generation of mouse models

The KLF1 knockout (KO) mouse model was generated by targeted insertion of the neomycin resistance gene (Figure 2.1) (Perkins, Sharpe, Orkin 1995; Vinjamur DS, 2014).

2.1.2 Mouse Dissection and Collection of Fetal Livers

For the qRT-PCR, non-transgenic wildtype mice were bred with KLF1+/- mice to obtain KLF1 +/+ or KLF1 +/- offspring. For ChIP assays, non-transgenic wildtype mice were bred with each other to obtain KLF1 +/+ offspring. The fetal liver, brain and placenta were collected from each fetus at E12.5 and E14.5 as described (Alhashem et al., 2011). Females in the timed mating were checked every morning for the presence of a vaginal plug, indicating a successful mating, after which the plugged females were separated from the males and that day was considered E0.5. At E12.5 or E14.5 for qRT-PCR or ChIP assays respectively plugged females were dissected after being anesthetized using isofluorine, followed by cervical dislocation. The uterine horns were dissected out and whole mouse fetuses were placed in Petri dishes containing 1X PBS to prevent dehydration of the fetuses and contamination with maternal tissue. One fetus was processed at a time by separating maternal tissue from the yolk sac and care was taken to ensure that the yolk sac with the fetus inside did not rupture and lose blood. To collect fetal liver, E12.5 or E14.5 fetus is dissected out of their yolk sac before fetal liver is dissected away from other embryonic tissues. The fetal liver was then transferred into a clean well filled with 1X PBS and

rinsed thoroughly before being stored in a vial in liquid nitrogen. The dissected brain and placenta of each mouse fetus was stored at 4°C.

2.1.3 Genotyping

The ear punches obtained from by ear-punching the mice at 3-4 weeks of age was used to genotype adult mice. The fetal brain or placenta collected during dissection as mentioned in section 2.1.2 were used to genotype E12.5 and E14.5 fetuses. These tissues were digested overnight in a water bath at 55-60°C in 50 µl digestion buffer for ear punches and 500 µl digestion buffer for the fetal brain or placenta (10 mM Tris HCL, pH 8.5; 50 mM KCl, 40 mM $MgCl_2$, 0.45% Tween 20 and 0.45% NP40) containing 1µg/µl proteinase K (Roche, PCR grade). The following morning after digestion, proteinase K was inactivated by 2 cycles of alternate boiling (> 95°C) and cooling (4°C) for ten minutes each. A Polymerase Chain Reaction (PCR) was performed using genotyping primers designed by Yousef Alhashem (Alhashem Y et al., 2011) seen in (Table 2.1) and 2 - 5 µl of the resulting tissue lysate was run on a 2% agarose gel at 100 V for 45 minutes to one hour. A 100 bp ladder was run along with the samples. The wildtype mice have one band that is 170 bp and the KLF1 +/- mice have two bands one of 170 bp and one of 220 bp for each allele.

2.1.4 RNA extraction and cDNA synthesis

After genotyping was performed, the mouse fetal livers collected at E12.5 and stored in liquid nitrogen were removed and resuspended in 400 µl denaturation solution (Totally RNA total RNA isolation kit, Ambion) and thawed on ice. $1/10^{th}$ volume of the denaturation solution

of 3M sodium acetate (Totally RNA kit, Ambion) was added, followed by 400µl of acid-phenol: chloroform (pH 4.5, Totally RNA 42 kit, Ambion). The samples were inverted and then

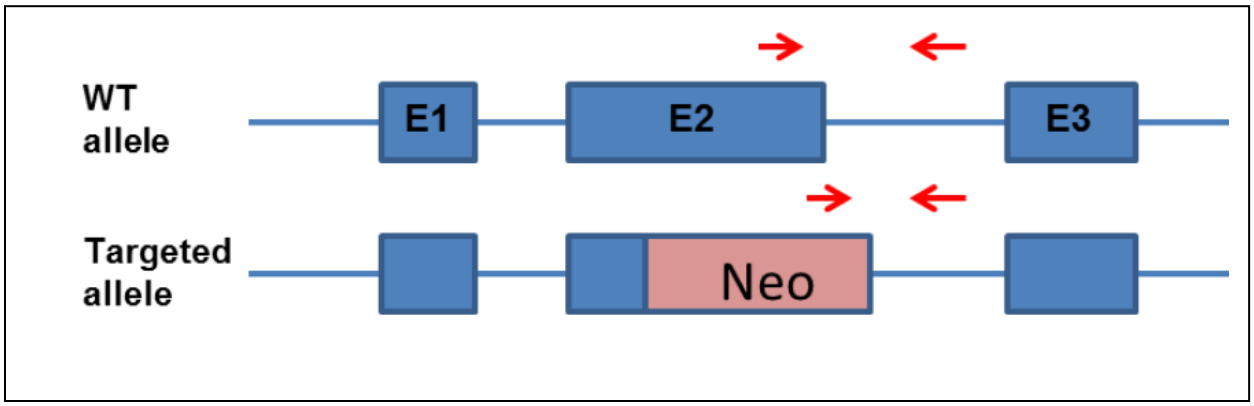


Figure 2.1: Graphical Representation of Knockout and wildtype KLF1 Alleles
Insertion of neomycin resistance gene to form the knockout allele

Table 2.1: Mouse KLF1 Genotyping Primers

Gene & Allele	Primer	Primer Sequence 5' to 3'	Amplicon Length
KLF1 KO	Forward Primer	GCCAGAGGCCACTTGTGTAG	220
	Reverse Primer	CTGGGACCTCTGTCAGTTGC	
KLF1 WT	Forward Primer	GGTGAACCCGAAAGGTACAA	170
	Reverse Primer	CTGGGACCTCTGTCAGTTGC	

vortexed for one minute and then centrifuged at 12,000 X g for 10 minutes at 4°C. After centrifugation, the aqueous phase was transferred to a fresh 1.5 ml tube again 400 µl of acid-phenol: chloroform (pH 4.5, Totally RNA 42 kit, Ambion) was added to each tube. The samples were inverted and then vortexed for one minute and then centrifuged at 12,000 X g for 10 minutes at 4°C. After centrifugation, the aqueous phase was transferred to a fresh 1.5ml tube and 4 µl of glycogen (5mg/ml stock, Ambion) and an equal volume of cold 100% isopropanol was added to the sample and mixed by inverting the tubes 4-5 times. RNA was allowed to precipitate overnight at -80°C. The following day, samples were thawed on ice and centrifuged at 12,000 X g for 30 minutes at 4°C. The supernatant was discarded and the pellet was washed with 1 ml of chilled 80% ethanol. The ethanol was discarded and the pellet was air-dried for 2-5 minutes on ice. The pellet was then resuspended in 20 µl of RNase-free water (USB Corporation) containing SuperaseIn (1:20 dilution, Ambion). 2 µl of the resuspended pellet was run on a 1% agarose gel to determine if the 28S and 18S rRNA subunits are intact to assess the quality of total RNA purified. The concentration of the remaining 16 µl of RNA of each mouse fetal liver sample was tested using a NanoDrop (Thermo Scientific NanoDrop 3300). Based on these concentrations the samples were further diluted to determine RNA integrity and concentration using the Agilent Bioanalyzer and the Agilent RNA 6000 Pico kit (Agilent Technologies), following the manufacturers protocol. For each fetal liver sample, a total of 1 µg of RNA was treated with DNaseI (Life Technologies) and used to prepare cDNA in 20 µl of solution following the instructions of the iScript cDNA synthesis kit (BioRad) to give a stock cDNA concentration 50 ng/µl.

For RNA preparations from the human HUDEP-2 samples, the cells were spun down after three days of differentiation and every ~2-3 x 10^6 cells were resuspended in 750 µl TRIzol

reagent (Thermofisher Scientific) and incubated for 5 minutes. 200 µl of chloroform was added and cells again incubated for 3 minutes and then centrifuged at 12,000 rpm at 4°C for 15 minutes. The aqueous phase was transferred to a new tube and 5 µl of glycogen and 500 µl chilled isopropanol were added to each sample and then incubated for 10 minutes on ice after which the tubes were stores at -80ºC overnight. After incubation samples were centrifuged for 15 minutes at 4°C at 12,000 rpm. The supernatant was discarded and the pellet was washed with 1 ml of chilled 80% ethanol. The ethanol was discarded and the pellet was air-dried for 2-5 minutes on ice. The pellet was then resuspended in 20 µl of RNase-free water (USB Corporation) containing SuperaseIn (1:20 dilution, Ambion).

2.1.5 Primer Design

The primer pairs were designed for qRT-PCR utilizing SYBR green reagent. The most recent human or mouse genome was selected in the UCSC Genome and then an exon common to all transcripts of the gene of interest was selected to design the primer around. Then the link of the RefSeq number from the UCSC genome to NCBI database was followed. The FASTA sequence was opened and that particular exon was picked out and 400 bases were copied into the Integrated DNA Technologies (iDT) design a primer program (200 bases before and 200 bases after the exon start). The primer pair where one of the primers extended the exon-exon junction was selected and BLAST was run to ensure specificity of the pair. It should be noted that two primer pairs were designed for each gene and tested afterward on qRT-PCR as amplicon sizes < 100 bp and so running an agarose gel was not feasible. The primer pair with higher efficiency, i.e., lower Ct value with amplification of one product, was selected.

The Primer pairs utilized in the ChIP assay qPCR experiments were designed differently. Those genes which showed a statistically significant difference in expression between the WT and KLF +/- mice and the HUDEP-2 scramble and K1V1 or K1V2 samples had primers designed around the CACCC regions present in the promoter regions of the genes, i.e., the promoter regions of the genes from -600 base pairs of the transcription start site (TSS) to the TSS, designated as 0, were downloaded using Ensembl. Then CACCC and complementary regions were identified and listed in order of which site was most similar to the designated 10 base pair consensus sequence (Tallack et al., 2011). Then the primers were designed around these sites using the same software as above, however the only difference was the primers were designed such that the CACCC sites were found in the middle of the amplicon (or as close to the middle as possible). This was to ensure that after sonication a maximum number of chromatin fragments held the protein.

2.1.6 Quantitative reverse transcriptase PCR (qRT-PCR)

Standard curves were made by serially diluting the 50 ng/µl stock cDNA to 25 ng/µl. Then serial dilutions in 1:5 ratio were obtained, i.e., 25 ng/µl to 5 ng/µl to 1 ng/µl to 0.2 ng/µl and to 0.04 ng/µl. Since Cyclophilin A is an abundant mRNA only 1 ng of each cDNA sample was loaded into each well whereas for mRNAs that are less abundant such as MBD2, 5 ng of cDNA were loaded into each well. Using the standard curves, known quantities of cDNA were used to perform absolute quantification (AQ) of the samples to determine mRNA amounts. A qRT-PCR was carried out using the SYBR Green reagent (Applied Biosystems) and plates were run on an ABI Prism 7900HT Fast Real-Time PCR System (Applied Biosystems). The various quality control measures applied to all results were: The R-value of the standard curve must be > 0.95, the amount of cDNA in all samples must fall within the limits of the standard curve, each sample

was run in quadruplet, at least three out of the four readings obtained for each sample must be within 20% of the mean quantity for all four readings. That is between 80% and 120% of the quantity mean, otherwise that reading was discarded and not used for further analysis. Cyclophilin A mRNA was used as the internal standard for normalization. Primer sequences used for qRT-PCR analyses of the mouse and human genes are listed in Table 2.2.

Three female mice were dissected to obtain 8 E12.5 KLF1 +/- and 8 E12.5 KLF1 +/+ fetuses which were compared to each other. The primer efficiency was calculated for each primer set and then the average Ct value for each sample of both the test gene and Cyclophilin A were calculated using: original Ct + Log(Primer efficiency, ng of sample added in each well). Next, the number of copies of Cyclophilin A and the test gene were calculated for each sample using the formula: Primer Efficiency^ (-1*Corrected Ct). To get the final fold change of the test gene normalized to Cyclophilin A the number of copies of test gene was divided by the number of copies of Cyclophilin A. An X by Y analysis was done in JMP and the mean of KLF1 +/+ vs. KLF1 +/- obtained from JMP was plotted in the bar graphs seen in the results section.

Ten knockdown samples (five K1V1 and five K1V2) were compared to five scramble control samples in the human HUDEP-2 cell line. The data analysis was similar to mice; primer efficiency, copy number and fold change were calculated using the same formulas. Additionally, the expression of each gene was set to 100 for scramble-treated cells and fold change over scramble was calculated for KLF1-shRNA infected cells, to minimize variations observed due to differences between each HUDEP-2 sample.

The cDNA was obtained and serially diluted similar to how the mouse samples were prepared. Absolute quantification (AQ) of the samples performed quantitative reverse transcriptase PCR (qRT-PCR) to analyze changes in gene expression in knockdown samples

(K1V1 and K1V2) compared to scramble controls. Additionally, the expression of each gene was set to 100 for scramble-treated cells and fold change over scramble was calculated for KLF1-shRNA infected cells, to minimize variations observed due to differences between each HUDEP-2 sample. The same quality control measures as in mice were also applied here.

Table 2.2: Human and Mouse qRT-PCR Primer Sequences

Gene	Primer Sequence 5' to 3'	Primer Length	Product Size
mLRF			
Forward Primer	CTGTAAAGTTCGATTCACCAGAC	23	110
Reverse Primer	CACTGCTGGCACAGGTA	17	
mLSD1			
Forward Primer	GAAGAGCCGTCTGGTGTG	18	102
Reverse Primer	CTGCTTCCTGAGAGGTCATTC	21	
mMBD2			
Forward Primer	GAAGGAGGAAGTGATCCGAAA	21	69
Reverse Primer	CTTCTTACCACTTGGACTGAAGTA	24	
mc-MYB			
Forward Primer	TCCATCTCAGCTCTCTCCAA	20	134
Reverse Primer	GGAACGTGACTGGAGATGTTT	21	
mMi2-β			
Forward Primer	CAGTGAGGAGGAGGATATGGA	21	127
Reverse Primer	GATAAATCCTCGTCTGGGTCTTC	23	
mCyclophilin A			

Forward Primer	ATTTCTTTTGACTTGCGGGC	20	112
Reverse Primer	AGACTTGAAGGGGAATG	17	
hLRF			
Forward Primer	TGCAAGGTCCGCTTCAC	17	74
Reverse Primer	TCGTAGTTGTGGGCAAAGG	19	
hLSD1			
Forward Primer	CTGACATTTGAGGCTACTCTCC	22	78
Reverse Primer	AGATGCCGAAGTTGATAAGACC	22	
hMBD2			
Forward Primer	CAGAGACGTGGCAAGAGGG	19	97
Reverse Primer	GTTTCTCCGAGGGCAAGAGG	20	
hc-MYB			
Forward Primer	GATGGGCAGAAATCGCAAAG	20	122
Reverse Primer	TGCAGATAACCTTCCTGTTCG	21	
hMi2-β			
Forward Primer	ATAAAGATAAGCCATTGCCTCC	22	101
Reverse Primer	GACATATGCCTTGAACTCTTTCTC	24	
hCyclophilin A			
Forward Primer	CCGAGGAAAACCGTGTACTATTAG	24	103
Reverse Primer	TGCTGTCTTTGGGACCTTG	19	
hβ-globin			
Forward Primer	GCAAGGTGAACGTGGATGAAGT	22	97

Reverse Primer	TAACAGCATCAGGAGTGGACAGA	23	
Probe	CAGGCTGCTGGTGTACCCTTGGACCC	26	
hγ-globin			
Forward Primer	GTGGAAGATGCTGGAGGAGAAA	22	81
Reverse Primer	TGCCATGTGCCTTGACTTTG	20	
Probe	AGGCTCCTGGTTGCTACCCATGGACC	26	
hKLF1			
Forward Primer	GCAAGAGCTACACCAAGAG	19	96
Reverse Primer	GTGTTTCCGGTAGTGGC	17	

2.1.7 Chromatin Immunoprecipitation (ChIP)

ChIP assays were performed as previously described (Hou et al., 2008). 1% formaldehyde was added to each biological replicate of ~5 x 10^6 blood cells to cross-link for 10 minutes at room temperature. In mice three E12.5 fetal livers were mixed to obtain one sample and in the human HUDEP-2 cell line ~10 x 10^6 cells were used for one sample. A total of three samples were used for ChIP assays in both mice and human models. Then, 125 mM glycine was added to inhibit crosslinking and replicates were incubated at room temperature for 5 minutes. The cells were centrifuged at 300 X g for 5 minutes and the supernatant was discarded. Cell lysis buffer (10 mM Tris, 10 mM NaCl, 0.2 % NP-40 [pH 8.0], and three protease inhibitors – Pepstatin, AEBSF and Aprotinin) was then added to lyse the cells and was then centrifuged at 500 X g for 5 minutes. To release the chromatin, the nuclei which is present as a pellet, is lysed with 0.5 ml Nuclei lysis buffer (50 mM Tris, 10 mM EDTA, 1% SDS, and three protease inhibitors) for 10 minutes on ice. 100 µl of IP dilution buffer (20 mM Tris, 150 mM NaCl, 2 mM EDTA, 0.01% SDS, 1% Triton X-100, and three protease inhibitors) is added for every 2 x 10^6 cells and

distributed to 300 µl per 1.5 ml tube. A bioruptor sonicator (Diagenode, Sparta, NJ) was utilized to shear the chromatin using the following settings: 3 cycles of: "H" level, 30 sec ON, 30 sec OFF, for 10 min (Total time 30 mins, 15 mins sonication time). While sonicating the chromatin, the Protein G Agarose (Millipore, Cat. 16-266) was washed three times using IP dilution buffer then kept at 50 µl per IP in the same buffer. After sonification ~10 µl of the chromatin was run on a 1% agarose gel to look at the size of the sheared chromatin fragments. The sheared chromatin was then centrifuged for 10 minutes at 10,000 X g at 4°C. The supernatant which contains the soluble chromatin is moved to a new tube and diluted to 1 ml per antibody using the IP dilution buffer. 10% of the chromatin is aliquoted into a new tube and stored at -20°C for the input DNA. Chromatin is distributed to one tube per antibody with 1 ml chromatin per tube. 5 µg of each antibody is added to the chromatin and 50 µl of washed protein G agarose is added to each tube which is incubated at 4°C overnight. The following day, the sheared chromatin is washed twice with 500 µl of IP wash buffer 1 (20 mM Tris, 50 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100 [pH 8.1]), once with IP wash buffer 2 (10 mM Tris, 0.25 M LiCl, 1 mM EDTA, 1% Nonidet P-40, 1% Desoxycholate [pH 8.1]) and twice with TE (pH 8.0). 400 µl freshly prepared IP elution buffer (0.1 M NaHCO3, 1% SDS (10ml: 1M NaHCO3 1ml, 20% SDS 0.5ml, water 8.5 ml) is added to each tube to elute the chromatin. Eluted chromatin is centrifuged to remove the agarose beads and is incubated at 65°C for 6 hours after adding NaCl (final concentration 0.3 M) and proteinase K (0.25 mg/ml). The proteinase K is deactivated by incubating the chromatin at 95°C for 10 minutes. An extraction with phenol chloroform is performed. Purified DNA is quantitated using quantitative PCR (qPCR) and SYBR Green chemistry. Fold enrichment was calculated as 2^ (Ct.input – Ct.test) and expressed relative to the IgG control. Antibodies used for mouse-ChIP: Anti-KLF1 (Abcam, ab2483); non-specific Rabbit

IgG (Abcam, ab46540). Antibodies used for human-ChIP: Anti-KLF1 (Abcam, ab2483); Anti-Human IgG (Abcam, ab2410). Primer sequences used for qPCR analyses of the mouse and human HUDEP-2 ChIP assay genes are listed in Table 2.3.

Table 2.3: Mouse and Human qPCR Primer Sequences for ChIP Assays

Gene	Primer Sequence 5' to 3'	Primer Length	Product Size
mLRF-1 ChIP			
Forward Primer	TAGCCTCGAACTCACAGAGA	20	95
Reverse Primer	CCCGCATTTACCAACACATTT	21	
mLRF-2 ChIP			
Forward Primer	CCAGGCGCTGGTTCTAACTCCAGGC	20	106
Reverse Primer	GCTGGCCACCAGACGACCTGACTGT	20	
mLSD1-1 ChIP			
Forward Primer	CCCTGTCTCAATTAGGTAGTAAGC	24	135
Reverse Primer	CACTTCAACTAATAACATTCACTCTCC	27	
mLSD1-2 ChIP			
Forward Primer	CCTGCAGGGAGAGTGAATGTTAT	23	87
Reverse Primer	GCCTGACTTCGGGTTCTGA	19	
mβmaj-promoter ChIP			
Forward Primer	GCTTCTGACATAGTTGTGTTGACTCA	26	104
Reverse Primer	CAGCAGCCTTCTCAGCATCA	20	
mβActin ChIP			
Forward Primer	ACCCCATTGAACATGGCATT	20	73

Reverse Primer	TGTAGAAGGTGTGGTGCCAGAT	22	
hLRF-1 ChIP			
Forward Primer	ACTGAGGCCCGGGGAG	16	141
Reverse Primer	CACAGCGAGGAGCGAC	16	
hLRF-2 ChIP			
Forward Primer	GTCGCTCCTCGCTGTGC	17	99
Reverse Primer	TAAACAGGCCCCCAGCCT	18	
hLRF-3 ChIP			
Forward Primer	CAGAGACGTGGCAAGAGGG	19	115
Reverse Primer	GTTTCTCCGAGGGCAAGAGG	20	
hLSD1-1 ChIP			
Forward Primer	GCTAGAACTTCATAATGTAACCCTTG	26	83
Reverse Primer	CCATTTCCTCCAGGGACTTT	20	
hLSD1-2 ChIP			
Forward Primer	CGAGAACGTCCGCTCTATGG	20	80
Reverse Primer	AGGGACAAAAAGGGTCGGAG	20	
hMBD2-1 ChIP			
Forward Primer	GGACACTTTGCTTCATTCCTTC	22	111
Reverse Primer	ACATACCAACTCGGCTATTCAT	22	
hMBD2-2 ChIP			
Forward Primer	CTAACCACAGTCACTATGACACC	23	119
Reverse Primer	CATCTTGTCTCCGCCTTCTC	20	
hc-MYB1 ChIP			

Forward Primer	CCAGAGGGCACAGTTGTAAA	20	126
Reverse Primer	GGACCAGAGCTCCTCCA	17	
hc-MYB2 ChIP			
Forward Primer	GGTTTGCTCAGGAAAAGGCG	20	144
Reverse Primer	CGCGGTGCGCCTAGC	15	
hγ-promoter ChIP			
Forward Primer	CAAATATCTGTCTGAAACGGTCCCT	24	86
Reverse Primer	TGCCTTGTCAAGGCTATTGGT	21	
hβ-promoter ChIP			
Forward Primer	GAGGGTTTGAAGTCCAACTCCTAA	24	116
Reverse Primer	CAGGGTGAGGTCTAAGTGATGACA	24	
hNecdin ChIP			
Forward Primer	TTCGTCCAGCAGAATTACCTGAAG	24	134
Reverse Primer	GGACCCCCAGAAGAACTCGTA	21	

2.1.8 Statistical Analysis

The Student's one-sample, one-tailed t-test was used for statistical analyses. P values < 0.05 for qPCR and P values < 0.025 for qRT-PCR and western blot were considered significant.

2.2 Studies in the Human-Umbilical Cord Blood Derived Erythroid Progenitor (HUDEP-2) Cell Line

2.2.1 Expansion of HUDEP-2 Cell Line

The HUDEP-2 cells were obtained from Dr. Yukio Nakamura (RIKEN Bio Resource Center, Tsukuba, Ibaraki, Japan). They were established by immortalization of umbilical cord blood CD34+ cells using lentiviral delivery of a Tet-inducible HPV16-E6/E7 transgene (Nakamura Y et al., 2013). The human papillomavirus 16 (HPV16) E6 and E7 proteins induce immortalization of cells by inhibiting the function of the p53 and retinoblastoma (pRb) tumor suppressor proteins respectively (Münger K, Howley PM, 2002). The HUDEP-2 cells most closely resemble adult erythroid cells since they express predominantly β-globin upon culture in erythroid differentiation media, whereas the HUDEP-1 and HUDEP-3 lines express predominantly γ-globin. Additionally the HUDEP-2 (but not HUDEP-1 or HUDEP-3) line expresses BCL11A, a known repressor of γ-globin.

HUDEP-2 cells were stored at 1-2 x 10^6 cells/ml at -196°C in FBS with 10% DMSO in cryovials. To begin expansion one vial of HUDEP-2 cells is removed from liquid nitrogen and warmed at 37°C until thawed. Contents of the vial are added to a conical 15 ml tube with 14 ml of Serum-Free Expansion Medium (SFEM). The tube is centrifuged at 300 rcf for 3 minutes and the pellet is then dissolved in 1 ml of HUDEP-2 expansion medium and placed in one T25 flask containing 5 ml of expansion medium. The HUDEP-2 cells are then grown in expansion medium for 10 days in T25 flasks at a cell density of 1-2 million cells/ml. Pelleted HUDEP-2 cells are a light pink in color after 10 days in expansion medium. Cells were counted and medium changed every 2-3 days. Expansion medium: SFEM medium (Stem Cell Technologies) containing 2% penicillin/streptomycin (Gibco), 1 µg/ml doxycycline (BioBasic), 1 x 10^{-6} M dexamethasone (Sigma-Aldrich), 3 U/ml erythropoietin (Amgen) and 50 ng/ml human stem cell factor (Stem Cell Technologies). Doxycycline inhibits differentiation of the HUDEP-2 cells as it is used as a

substitute for Tet, and the presence of doxycycline induces HPV16-E6/E7 expression. (Nakamura Y et al., 2013).

2.2.2 Calcium phosphate transfection of 293T Cell Line

293T stock cells were stored at -196°C in FBS with 10% DMSO in cryovials. To begin transfection one vial of 293T cells are removed from liquid nitrogen and warmed at 37ºC until thawed. Contents of the vial are added to a conical 15 ml tube with 14 ml of Dulbecco's modified Eagle's medium-complete (DMEM with 10% FBS and 1% Penicillin/Streptomycin). The tube is centrifuged at 300 rcf for 3 minutes and the pellet is then dissolved in 1 ml of DMEM-complete and placed in one T75 flask containing 14 ml DMEM-complete. The 293T cells were transferred to round tissue culture plates at ~5 million cells per plate. The 293T cells were divided evenly into three categories for the calcium phosphate transfection: mock, scramble, K1V1 (~40-70% KD) and K1V2 (>75% KD). Lentiviral vectors carrying KLF1-targeted shRNA coding sequences were acquired from other labs. Two KLF1-targeted shRNA vectors (pRRL K1V1 and pRRL K1V2) and a scrambled shRNA vector (pRRL Scramble) were acquired from Dr. François Morlé (Université Claude Bernard Lyon1, Lyon, France) and Dr. Fawzia Louache (INSERM, Villejuif, France). These vectors were constructed on a pRRL backbone (Bouilloux, Blood 2008) with the shRNA-coding sequence under the control of the H1 promoter and also carry a GFP-coding sequence. The oligonucleotide sequences are listed in Table 2.3. Bacterial culture with plasmid was plated on LB agar plates containing 100μg / ml Ampicillin (Thermo Fisher Scientific). One ampicillin resistant colony of each plasmid was picked from each transformation reaction and cultured for mini preps (VCU Molecular Biology Core). The lentivirus is produced when the packaging plasmid pCMV-dR8.74, the envelope plasmid pMD2G and one of the three pRRL plasmids (K1V1, K1V2 or scramble) are co-

transfected into 293T cells (semi-confluent in 10 cm tissue culture dishes). The packaging and envelope plasmids were a gift from Dr. Gordon Ginder's laboratory. After being grown in a culture dish for 24 hours, the 293T cells were transfected. The transfection solution was prepared separately in a single 5 ml falcon tube for each of the categories, mock, scramble and knockdown. Each tube had 1 ml and is used to transfect one round tissue culture dish. For all categories, we combined 5.6 µg of pCMVRd8.74 and 2.3 µg of pMD2.G and sterile water up to 437 µl in the falcon tube. The tube was tilted to add 63 µl of 2 M $CaCl_2$ down the side of the tube. Next, 500 µl HBS Solution was added in the same manner to make a total volume of 1 ml. For scramble, 7.5 µg of pRRL encoding scramble shRNA was additionally added. For K1V1 and K1V2 knockdown, 7.5 µg of pRRL-pK1V1 or pRRL-K1V2 was additionally added respectively. The 1 ml of transfection solution was added dropwise to the respective 293T plate. Starting from the center of the plate and continuing in a spiral pattern of approximately 6 concentric circles towards the edge of the plate, distributing the drops evenly. Each plate was incubated at room temperature for 5 minutes. The plates were grown for 16-18 hours in a 37°C incubator at 3% CO_2. After 18 hours the cells begin to die due to $CaPO_4$ formation, which disrupts the cell membrane. So, the medium was removed and 15 ml of DMEM with 2% FBS and 1% penicillin/streptomycin warmed to 37ºC was added to each plate then incubated at 37°C, 10% CO_2 overnight. The following day, the medium containing viral particles was collected and first filtered through a 0.45µm filter, supplemented with polybrene, making the lentivirus ready for infection of HUDEP-2 cells.

2.2.3 Lentiviral infection of HUDEP-2 Cell Line

The HUDEP-2 cells were then collected, spun down and resuspended in HUDEP-2 expansion medium at ~5 x 10^6 cells/ml. 800 µl to 1 ml of lentiviral medium is added to 2 x 10^5 HUDEP-2 cells in 100 µl expansion medium in each well of a 12 well tissue culture plate and are incubated at 37°C at 5% CO_2. The following afternoon, the cells were collected and spun down to remove the viral medium. The infected HUDEP-2cells are then resuspended in 650µl of expansion medium and returned to the incubator for another two days.

2.2.4 Differentiation of HUDEP-2 Cell Line

Three days after infection the HUDEP-2 cells are resuspended in differentiation medium. differentiation medium: IMDM medium (Sigma-Aldrich) with 2% penicillin/streptomycin (Gibco), 1µg/ml doxycycline (BioBasic), 3 U/ml erythropoietin (Amgen) and 50 ng/ml human stem cell factor (Stem Cell Technologies), 5% human serum albumin (Thermo Fisher Scientific), 500 µg/ml human holo-transferrin (Gemini), 1% L-glutamine (Thermo Fisher Scientific), 10 µg/ml human insulin (Gemini) and 3 U/ml heparin (Gemini). The infected HUDEP-2 cells are differentiated for three days (DD3) for qRT-PCR, ChIP assays and for protein analysis. The pelleted cells appear red after three days of differentiation.

2.2.5 Western Blot

Protein extraction began by spinning down one HUDEP-2 sample of ~1 x 10^6 cells after 3 days of differentiation. The number of samples tested for β-globin protein detection was three K1V1, three K1V2 and four Scramble controls. The number of samples tested for γ-globin protein detection was three K1V1, three K1V2 and three Scramble controls. The cells were then

resuspended in 1 ml PBS and transferred to a 1.7 ml micro centrifuge tube and again spun down and resuspended in PBS then centrifuged at 18,000 X g for 1 minute and all PBS was removed. Every 1 x 10^6 cells were resuspended in 100 µl of RIPA buffer (1 M Tris-HCl, 0.5 M EDTA, 5M NaCl, 20% glycerol and NP-40) containing 2 µl of 50X protease inhibitor cocktail (Promega). The cells were then incubated in this buffer on ice for 45 minutes and then centrifuged at 18,000 X g for 15 minutes at 4°C. The supernatant contains the isolated protein and was collected and moved to a new micro centrifuge tube and kept on ice. To calculate the concentration of protein we took the absorbance readings of 6 standards containing 1 ml BioRad protein reagent with no bovine serum albumin (BSA), 2µl, 4 µl, 8 µl, 16 µl or 32 µl of BSA using a spectrophotometer. Using these values a standard curve was plotted for calculation of protein concentrations of the samples. The test samples were set up and contained a mixture of 1 ml protein reagent and 2 µl protein sample. After obtaining the protein concentrations of the samples from this standard curve we correct the value for the 5 X Laemmli buffer that was later used in SDS-polyacrylamide gel (SDS-PAGE) and for the 2 µl of protein added: Final protein concentration = (protein concentration x 0.8)/2.

After analyzing the protein concentration each sample's protein was diluted so as to load 30 µg of protein per well in a 10 cm x 8 cm x 5 mm gel. This was done by preparing the 30 µg in 4 µl of sample with 10 µl of Laemmli buffer and heating in a water bath at 95°C for 5 minutes and then spun down and loaded on the 8% SDS-PAGE. The resolving gel components were 40% acrylamide, 10% APS, 10% SDS, TEMED and 1.5 M Tris pH 8.8. The stacking gel components were 40% acrylamide, 10% APS, 10% SDS, TEMED and 0.5 M Tris pH 6.8. After loading the wells of the gel with the samples and the pre-stained low molecular weight range protein ladder, (Thermo Fisher) the gel was run for the first 30 minutes at 70 V and later at 100 V until the

loading dye reached the end of the gel. After this the gel was removed and placed in transfer buffer (10% SDS, methanol, Tris Base, Glycine) for one hour on a rotator. The molecular weights of β-Actin, β-globin and γ-globin are 42 kDa, 16 kDa and18 kDa respectively.

The transfer sandwich was made by placing these things in order from the black side of the cassette: a wet sponge, wet blot paper, the gel from above, nitrocellulose membrane, blot paper and sponge (white side of cassette), at each stage the bubbles were removed with a roller. The cassette was closed and placed in the holder with an ice pack and run at 100 V for 90 minutes. After transfer, the nitrocellulose membrane was removed and rinsed in PBS-T (PBS with 0.05% Tween-20) and subsequently stained with Ponceau Red stain for 10 seconds on a rocker to check that transfer occurred, here the membrane was cut above the 60 kDa mark on the ladder to conserve antibodies. To remove the stain, the membrane was again washed with PBS-T and then incubated for 30 minutes on a rocker with 5% dry milk solution as a blocking solution. After blocking the membrane, PBS-T was used to rinse once again. The membrane was incubated in 1X TBS, 0.1% Tween-20, 1% BSA, 1% calf serum dilution with a 1:1000 dilution of either the γ-globin (sc-21756) or β-globin (sc-21757) primary antibody overnight at 4°C on a rotating shaker. The following morning, the membrane was washed with PBS-T and then 1-2 ml of chemiluminescence (ECL) was uniformly pipetted over the membrane and observed in the dark room. The membrane was then washed with PBS-T after viewing and immersed in stripping buffer on the rocker for 15 minutes. The stripping buffer was removed and blocking agent was added. The control antibody β-Actin (C4) (sc-47778) was then added at a 1:1000 dilution in 1X TBS, 0.1% Tween-20, 1% BSA, 1% calf serum dilution overnight at 4°C to the membrane and the process was repeated to view the control bands on the membrane.

The relative densities of each sample were calculated using ImageJ. Each of the sample bands was normalized to β-actin bands of the 16 kDa β-globin and 18 kDa γ-globin bands. The various quality control measures applied to the western blot results were: the density of the bands in per category must fall within the limits of the average density of that category, each sample was run in quadruplet, at least three out of the four readings obtained for each sample must be within 20% of the mean density for all four readings. That is between 80% and 120% of the quantity mean, otherwise that reading was discarded and not used for further analysis.

2.2.6 Statistical Analysis

The Student's one-sample, one-tailed t-test was used for statistical analyses. P values < 0.05 for qPCR and P values < 0.025 for qRT-PCR and western blot were considered significant.

Chapter 3: Results

Regulation of Known γ-globin Gene Repressors by KLF1 in the Mouse and Human Model Systems

---

3.1 Introduction

KLF1 is a transcription factor expressed in erythroid cells and is a master regulator of the hemoglobin switch from fetal (HbF) to adult hemoglobin (HbA) in humans. and embryonic to adult hemoglobin in mice (Feng WC et al., 1994; Donze D et al., 1995; Wijgerde M et al., 1996; Perkins AC et al., 1996; Bieker JJ, 1996; Bieker JJ, 2001). KLF1 directly upregulates both β- and γ-globin gene expression (Alhashem et al., 2011; Borg J et al., 2010; Zhou D et al., 2012; Tallack MR et al., 2013). In primary cells there is a direct positive linear correlation between KLF1 and β-globin mRNA levels, but the regulatory role of KLF1 in γ-globin gene expression seem to be more complex. It was discovered that a 40-70% knockdown of KLF1 in differentiated primary human cells showed a two-fold increase in γ-globin when compared to the scramble control samples (Vinjamur D et al., 2016). The primary cells with a robust knockdown of KLF1, i.e., greater than 80%, showed an almost identical γ-globin level as that of the scramble cells (Vinjamur D et al., 2016). This suggests that KLF1 is responsible for repressing γ-globin gene expression, but introduced the concept of a threshold amount of KLF1 which is required by erythroid cells to increase γ-globin mRNA over normal levels. That is why in this research, we utilized the threshold amount of KLF1- in the mouse as the genotype KLF1+/-; and in the human as the K1V1 lentivirus. The K1V1 lentivirus gives a KLF1 knockdown range between 40% and

70%, also referred to as modest knockdown. The other lentivirus, K1V2, gives a KLF1 knockdown efficiency of >75% and is referred to as robust knockdown.

This phenomenon becomes important when trying to identify potential therapeutic targets to cure β-hemoglobinopathies like sickle cell disease and β-thalassemia's. It should also be stated that KLF1, unlike other γ-globin gene repressors, only affects erythropoietic processes, giving it another advantage over other target genes.

An evolutionary conservation study was performed on three master regulators of erythropoiesis, GATA1, NFE2 and KLF1 to examine their occupancy sites in proerythroblasts over 75 million years, i.e., between mouse and human (Ulirsch JC et al., 2014). GATA1 and NFE2 showed approximately 25% of master regulatory sites were conserved. Interestingly, >60% of KLF1 occupancy peaks were conserved between species along the proerythroblast differentiation lineage (Ulirsch JC et al., 2014; Pishesha N et al., 2014). This is in keeping with research showing that specific master transcriptional regulators are conserved in both of these species (Orkin SH et al., 2008; Cantor AB et al., 2002; Wu W et al., 2011). It is of interest to observe the KLF1 binding in the γ-globin gene repressors to determine if there is any evolutionary conservation of function and binding of KLF1 between mouse and human. KLF1 is a transcription factor that binds a specific 9 base pair consensus sequence 5'- CCM-CRC-CCN - 3', (where R represents A or G and M represents A or C) (Feng, Southwood, Bieker 1994; Miller and Bieker 1993; Tallack et al. 2010). These sites are found throughout the mouse and human genome (Tallack et al., 2010). If a protein directly regulates expression of another gene binding may occur within the promoter region of that gene. This has been shown to be the case for binding of KLF1 to the mouse βh1-globin promoter and human γ- and β-globin promoters (Alhashem Y et al., 2011). Thus, the first two aims of this research project further assess how

KLF1 indirectly downregulates γ-globin gene expression by studying transcription levels. Then, KLF1 binding by identifying the specific 9 base pair sequence in the promoter regions of these five known γ-globin gene repressors. This was done in both the mouse and human systems using the aforementioned threshold amount of KLF1.

HUDEP-2 cells are immortalized cells that were established by delivering a Tet-inducible HPV16-E6/E7 lentiviral transgene in CD34+ cells isolated from human umbilical cord blood (Nakamura Y et al., 2013). These cells are much easier to expand and differentiate than primary human cells. The erythroblasts isolated from HUDEP-2 cells are physiologically more similar to adult erythroblasts than the erythroblasts isolated from primary human cells which tend to be more fetal-like. It was of interest to see if the phenomenon of KLF1 threshold amount could be observed in the HUDEP-2 cells as was previously seen in the primary human cells (Vinjamur D et al., 2016), i.e., they could be characterized similarly with respect to γ-globin and β-globin mRNA levels when KLF1 knockdown is induced in the cells. After this characterization it was of importance for us to look at the protein levels of the globins to ensure that protein plot is similar to the mRNA plot of percent γ-or β-globin vs. residual KLF1. Thus we focused on characterization of the γ-globin and β-globin mRNA and protein levels in HUDEP-2 cells with various knockdowns of KLF1.

3.2 LSD1 and LRF are Regulated by KLF1 in Mice

As stated before, KLF1 positively regulates β-globin gene expression whereas it both directly positively and indirectly negatively regulates γ-globin gene expression. Five known γ-globin gene repressors namely, Mi2-β, MBD2, c-MYB, LRF and LSD1 were examined, to determine whether or not KLF1 regulates them. To do this we mated KLF1 +/+ and KLF1 +/- mice to

obtain KLF1 +/+ and KLF1 +/- fetuses at Embryonic day 12.5 (E12.5) and E14.5. In doing so, we were able to make use of the modest knockdown of KLF1 that increased γ-globin in CD34+ cells, i.e. the KLF1 +/- mice are comparable genotypically to cells with the modest KLF1 knockdown. E12.5 mouse fetal livers were used as a majority of cells present in the fetal liver are erythroid and by E14.5 more of the cells are hepatocytes (Crawford LW et al., 2010; Dzierzak, E et al., 2013).

3.2.1 KLF1 Upregulates LRF and LSD1 Genes in Mouse E12.5 Fetal Livers

After analyzing the qRT-PCR data, the average mRNA expression level normalized to Cyclophilin A was compared between 8 KLF1 +/+ and 8 KLF1 +/- E12.5 samples, where the KLF1 +/+ values were set to 100% (Figure 3.1). It was observed that in KLF1 +/-, the mRNA amount of LRF was 62±6% and the LSD1 mRNA was 31±7%, compared to 100% in KLF1 +/+ (Figure 3.1 B & C), which was significant. A single sample, one-paired, one-tailed t-test was performed to test whether KLF1 +/- values were statistically significant from KLF1 +/+ values. However, there was no statistically significant difference between the expression levels of MBD2, Mi2-β and c-MYB mRNA in KLF1 +/+ and KLF1 +/- fetal livers (Figure 3.1 A, D & E). Based on our data it can be concluded that KLF1 positively regulates the LSD1 and LRF genes in the mouse fetal liver. KLF1 does not appear to have any measurable effect on the MBD2, Mi2-β or c-MYB genes in E12.5 mouse fetal livers.

3.2.2 KLF1 Binds the Promoters of LRF and LSD1 Genes in Mouse E12.5 Fetal Livers

To determine whether the LRF and LSD1 genes are directly regulated by KLF1, KLF1-ChIP was performed using mouse fetal liver cells. First, all 9 base pair and complementary consensus

sites were identified in the promoter regions of the genes – LRF and LSD1 (Table 3.1). After performing the ChIP assay on the chromatin, the average size of the fragments ranges from 100 bp to 300 bp, seen in the 2nd to 4th lanes of the gel (Figure 3.2).

Out of all consensus binding sites present in the promoter regions of LRF and LSD1 from Table 3.1 we selected three sites for LRF (Figure 3.3 A) and two sites for LSD1 (Figure 3.3 B) genes which are highlighted in red.

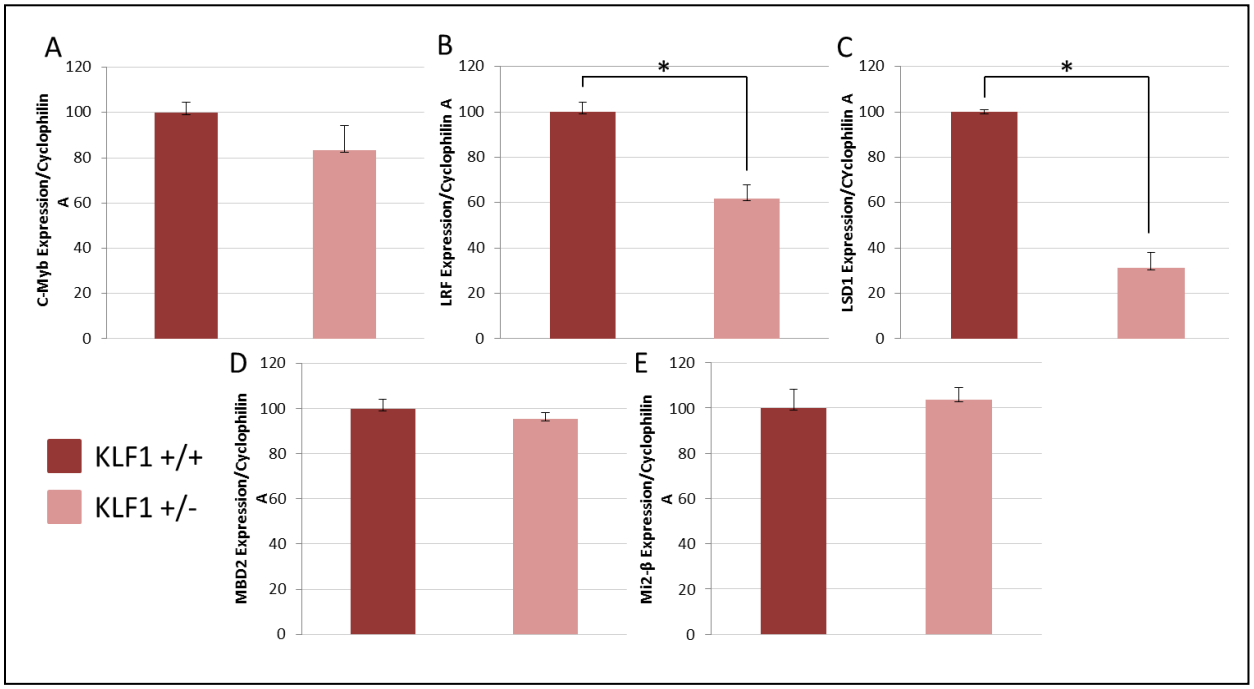


Figure 3.1: KLF1 Upregulates LRF and LSD1 Genes in Mouse E12.5 Fetal Livers
A qRT-PCR was performed on γ-globin gene repressors A) c-MYB, B) LRF, C) LSD1, D) MBD2 and E) Mi2-β. Paired mouse fetal livers were dissected from E12.5 KLF1 +/+ and KLF1 +/- fetuses, for RNA extraction. The amount of KLF1 mRNA was measured using qRT-PCR methods and expressed using the EdCT method to allow for direct comparison of mRNAs from paired fetuses. Cyclophilin A was used as a normalization control. N=8 per genotype; Error bars represent SD; *, $P < 0.05$ (Paired-Student's one-tailed test)

There was a 1.8 fold increase in relative enrichment of the -511 to -492 promoter region of mouse LRF gene when compared to mouse IgG, which is statistically significant. A 1.6 fold

increase in relative enrichment of the -214 to -206 promoter region of mouse LSD1 gene when compared to mouse IgG that was statistically significant (Figure 3.4 A). The mouse β-actin promoter was used as a negative control and as expected showed no statistically significant binding of KLF1. The β-major promoter was used as the positive control and a statistically significant 1.8 fold increase in relative enrichment was obtained when compared to IgG. Based on our data it can be concluded that KLF1 binds to the mouse LSD1 and LRF promoters at positions -214 to -206 and -511 to -492, respectively.

Table 3.1: The Mouse Consensus Sequences Located in the Promoter Regions of LSD1 and LRF Genes

Mouse Gene	Position of CACCC from 5' UTR of TSS	Position of CGCCC from 5' UTR of TSS	Position of GGGTG from 5' UTR of TSS	Position of GGGCG from 5' UTR of TSS
LSD-1 (KDM1A)	-263 to -271 ATCCACCCC	-	-206 to -214 GGGGTGAAT +17 to +24 GGGGTGTT	+52 to +59 GGGGCGGA
LRF (ZBTB7A)	-	-491 to -500 CCACGCCCG -167 to -175 CTCCGCCCC	-503 to -511 TGGGTGTGC -494 to -501 TGGGTGTG	-

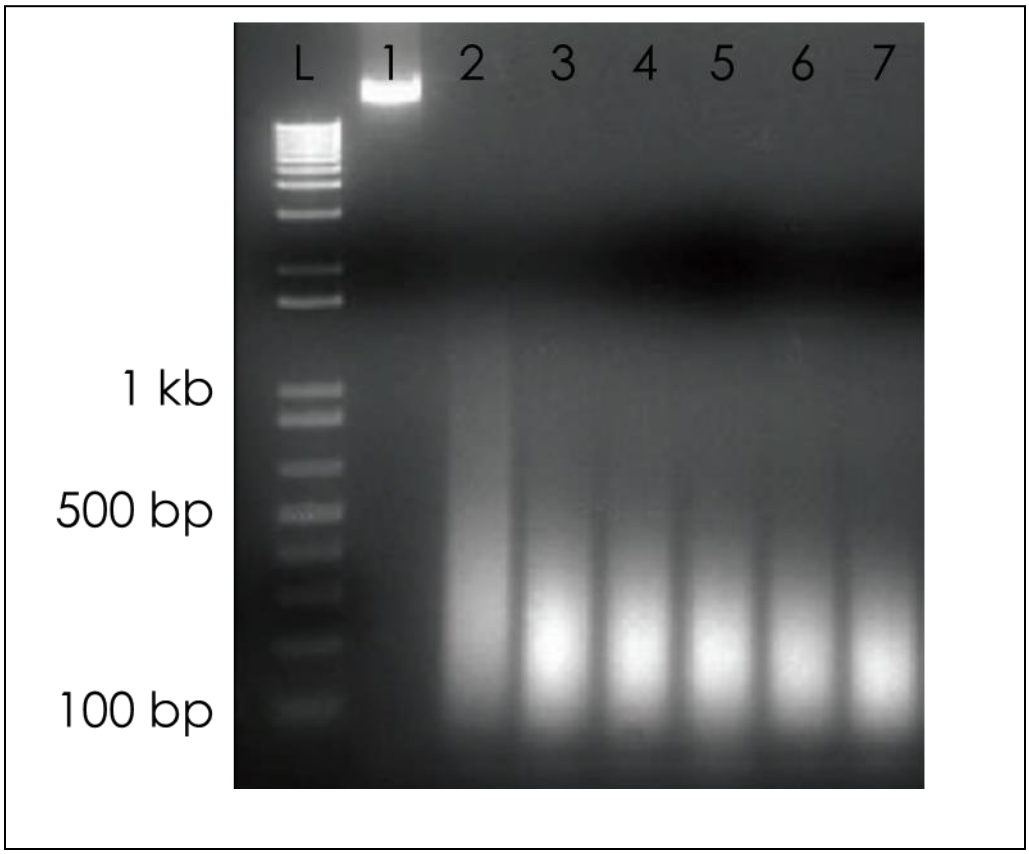


Figure 3.2: Average Chromatin Fragment Size after Sonication of Mouse Fetal Livers
A 1.5% agarose gel was run with 100 base pair ladder. 1st lane is chromatin before sonication. 2nd- 4th lanes are chromatin fragments of E12.5 mouse fetal liver cells, 5th – 7th lanes are chromatin fragments of E14.5 mouse fetal liver cells.

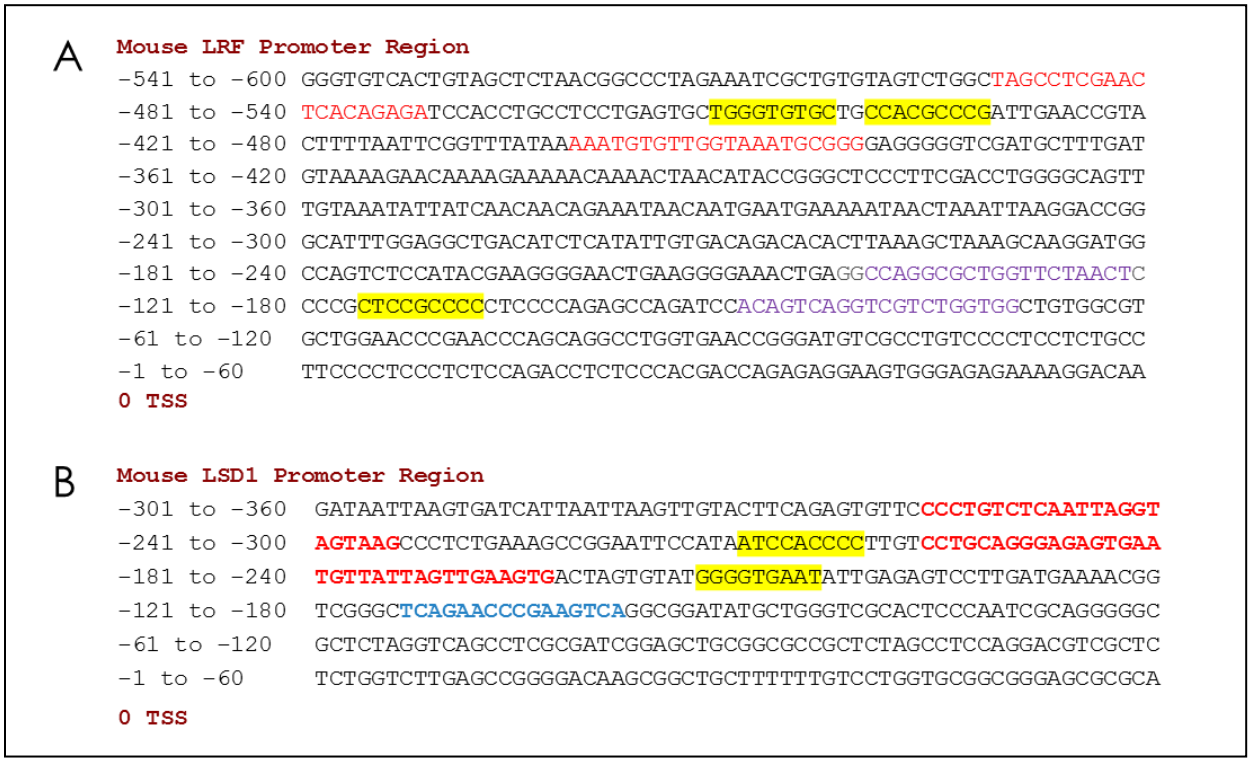


Figure 3.3: Promoter Regions of Mouse LRF and LSD1 Genes
Promoter regions were acquired from Ensembl. The consensus sites are highlighted in yellow and primers in red or blue. A) LRF Promoter Region with LRF1 and LRF2 consensus sites and B) LSD1 Promoter Region with LSD1-1 and LSD1-2 consensus sites highlighted in yellow.

3.2.3 No Evidence that KLF1 Binds LRF and LSD1 Promoters in Mouse E14.5 Fetal Livers

A mouse KLF1-ChIP assay was performed on E14.5 mouse fetal livers because previous data analyzed from two KLF1 ChIP-seqs was performed on E14.5 in 2010 and 2011 (Tallack MR et al., 2010; Pilon AM et al., 2011). This data showed that KLF1 does not bind mouse LRF or LSD1 in their promoters (Figure 3.4 B).

Therefore, to negate the effect of biological differences between our E12.5 ChIP assay results and the E14.5 ChIP-seq results, a ChIP assay was also performed on E14.5 mouse fetal liver. The same three consensus sites in the promoter region of the LRF gene were examined:

After performing the ChIP assay on the chromatin, the average size of the fragments range from 100 bp to 400 bp, seen in the 5^{th} to 7^{th} lanes of the gel (Figure 3.2). There was no significant increase or decrease in relative fold enrichment of the mouse LSD1 and LRF genes when compared to mouse IgG (Figure 3.4 C). The negative control, mouse β-actin promoter showed no significant increase or decrease in relative fold enrichment when compared to the control IgG enrichment. The positive control, β-major promoter showed a statistically significant 6-fold enrichment compared to IgG. Based on our data there is no evidence that KLF1 binds the mouse LSD1 and LRF genes at positions -214 to -206 and -511 to -492 respectively in E14.5 mouse fetal livers. This is in agreement with the aforementioned research on E14.5 mouse fetal livers (Tallack MR et al., 2010; Pilon AM et al., 2011).

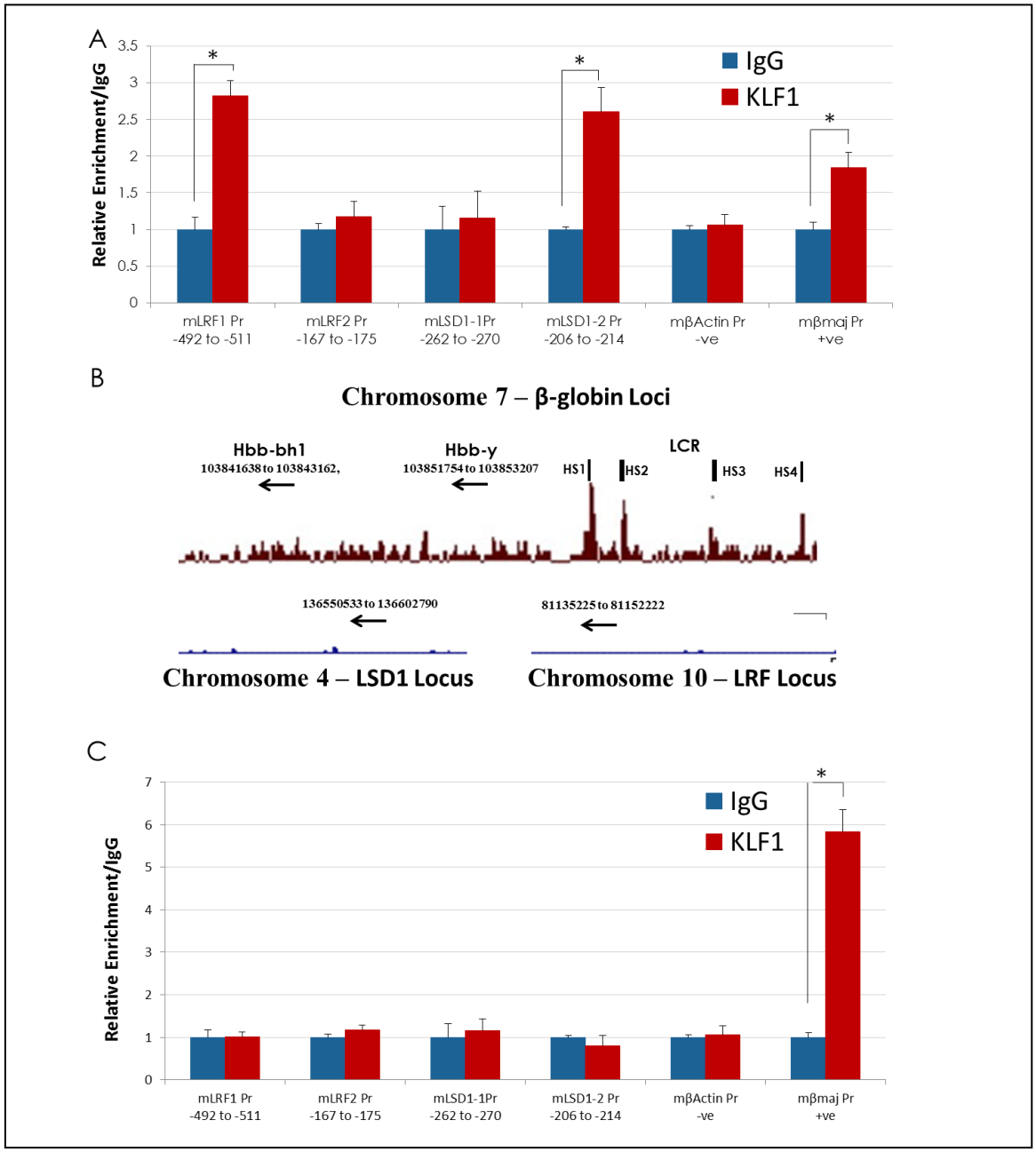


Figure 3.4: KLF1 Binding in E12.5 and E14.5 Mouse Fetal Livers

ChIP assays were performed on A) E12.5 and C) E14.5 fetal livers of normal, non-transgenic mice. Polyclonal antibody specific for KLF1 and non-specific IgG control antibody were used. The Y-axis represents the relative fold enrichment. The mean IgG enrichment was set as 1.0 and the enrichment of KLF1 was scaled appropriately. The X- axis shows the location of the primers used for qPCR (Pr: promoter). The error bars indicate standard error mean (SEM); N = 3; *, $P < 0.05$ (Paired-Student's one-tailed test). B) Data analysis of mouse E14.5 fetal livers KLF1 ChIP-seq data showing the peaks of KLF1 at the β-globin gene promoters and the LSD1 and LRF promoter regions.

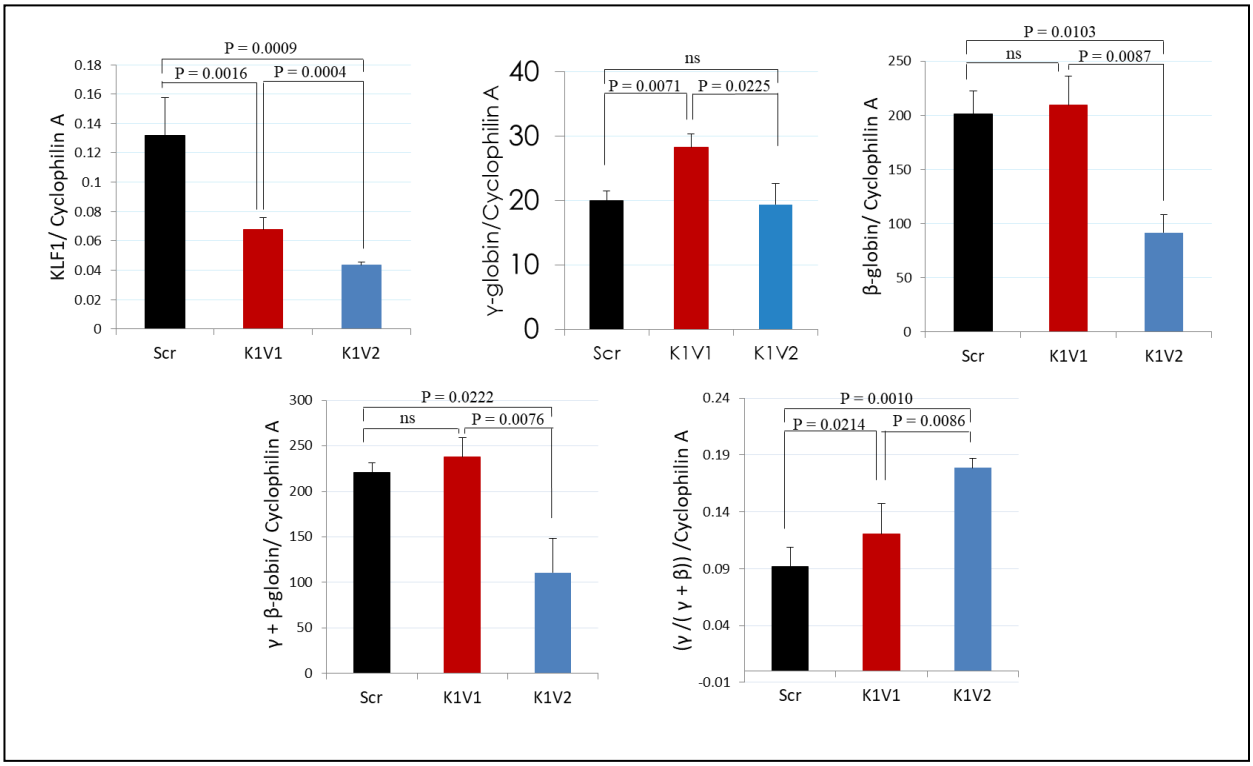
3.3 Characterization of γ-globin and β-globin mRNA Amounts with Corresponding KLF1 mRNA Amounts in HUDEP-2 Cells

Currently, there is no data on the characterization of the γ-globin and β-globin mRNA and protein levels in the HUDEP-2 cells with respect to KLF1 knockdown. Similar to the experiments performed using erythroblasts derived from umbilical cord blood CD34+ cells, we performed lentiviral infection of the various KLF1 knockdowns in HUDEP-2 cells. As stated before all experiments using HUDEP-2 cells utilize the threshold amount of KLF1. That is the KLF1 modest knockdown, designated as K1V1 giving a range between 40% and 70%. The KLF1 robust knockdown, K1V2, gives a KLF1 knockdown efficiency of >75%. Prior to examining the role of KLF1 in regulating the five γ-globin repressor genes, the KLF1 mRNA levels were tested to show that these knockdowns worked.

3.3.1 Raw KLF1, γ-globin and β-globin mRNA Levels Sorted by Knockdown Category in HUDEP-2 Cells

The raw values of KLF1, γ-globin and β-globin mRNA levels normalized to Cyclophilin A sorted by category are depicted (Figure 3.5). To verify that our knockdowns of KLF1 worked we looked at expression levels of KLF1 in the three various categories (Figure 3.5 A). We then were interested in analyzing the γ-globin (Figure 3.5 B) and β-globin (Figure 3.5 C) expression levels with respect to KLF1 knockdowns. The γ-globin expression levels show a significant increase in the K1V1 group as compared to scramble and K1V2 levels. No significant increase or decrease in γ-globin expression levels is noted in the K1V2 group when comparted to the scramble control. When comparing the scramble control values of γ-globin expression to the control

values of β-globin expression we observe that the total γ-globin i.e. γ/ (γ +β) = 20/220 is 11% in control cells with the scramble lentivirus. A significant increase in γ-globin of the HUDEP-2



samples is seen between scramble control and K1V1 (modest knockdown). A significant

Figure 3.5: Raw Values of KLF1, β-globin and γ-globin mRNA Expression Levels by Category in HUDEP-2 Cells.
Raw values of A) KLF1, B) γ-globin, C) β-globin, D) Total globin and E) γ-globin over total globin mRNA sorted by category. The error bars indicate SD; N = 5 per category; P-values as indicated, ns = not significant (Paired-Student's one-tailed test, P < 0.025)

increase is also seen in γ-globin of the HUDEP-2 samples when comparing K1V1 and K1V2 (robust knockdown). We see that there is no significant difference in total globin when comparing the modest knockdown category to the scramble control (Figure 3.5 D). But we do see a significant decrease in total globin when comparing the robust knockdown to both modest knockdown and scramble control groups. The γ-globin over total globin is significantly increased in both K1V1 and K1V2 groups when compared to the control group (Figure 3.5 E). Overall the results show that the K1V1 and K1V2 shRNAs give KLF1 knockdowns as previously shown. A

significant decrease in β-globin levels is observed with the K1V2 knockdown samples when compared to both the scramble control and K1V1 knockdown samples. It should be noted that protein amounts of KLF1 in the categories was calculated in human CD34+ cells isolated from fetal umbilical cord blood and showed corresponding protein amounts to their mRNA amounts (Vinjamur DS et al., 2016). The HUDEP-2 cell line is derived from umbilical cord blood and so we would expect the protein levels of KLF1 in each category to correspond to mRNA levels here.

3.3.2 γ-globin and β-globin Protein Levels Correlate with mRNA Levels in HUDEP-2 Cells

It was of interest to determine if the γ- and β-globin protein amounts correspond to their mRNA amounts with respect to KLF1 knockdown. Relative densities normalized to β-actin bands of the 16 kDa β-globin (Figure 3.6 A) and 18 kDa γ-globin (Figure 3.6 B) bands were calculated. There is a significant reduction in expression of β-globin protein by ~4.8-fold upon knockdown of KLF1 with K1V2 shRNA when compared to both K1V1 and scramble categories (Figure 3.6 C). This is greater than the ~2-fold reduction of β-globin mRNA at the K1V2 category when compared to scramble control and K1V1 (Figure 3.5 C).

There is a significant increase in expression of γ-globin protein by ~2.0-fold upon knockdown of KLF1 with K1V1 shRNA when compared to both K1V2 and scramble categories (Figure 3.6 D). This is similar to the ~2-fold increase of γ -globin mRNA at the K1V1 category when compared to control and K1V2 (Figure 3.5 B). This shows that only the modest knockdown of KLF1 results in significant increase in γ-globin protein.

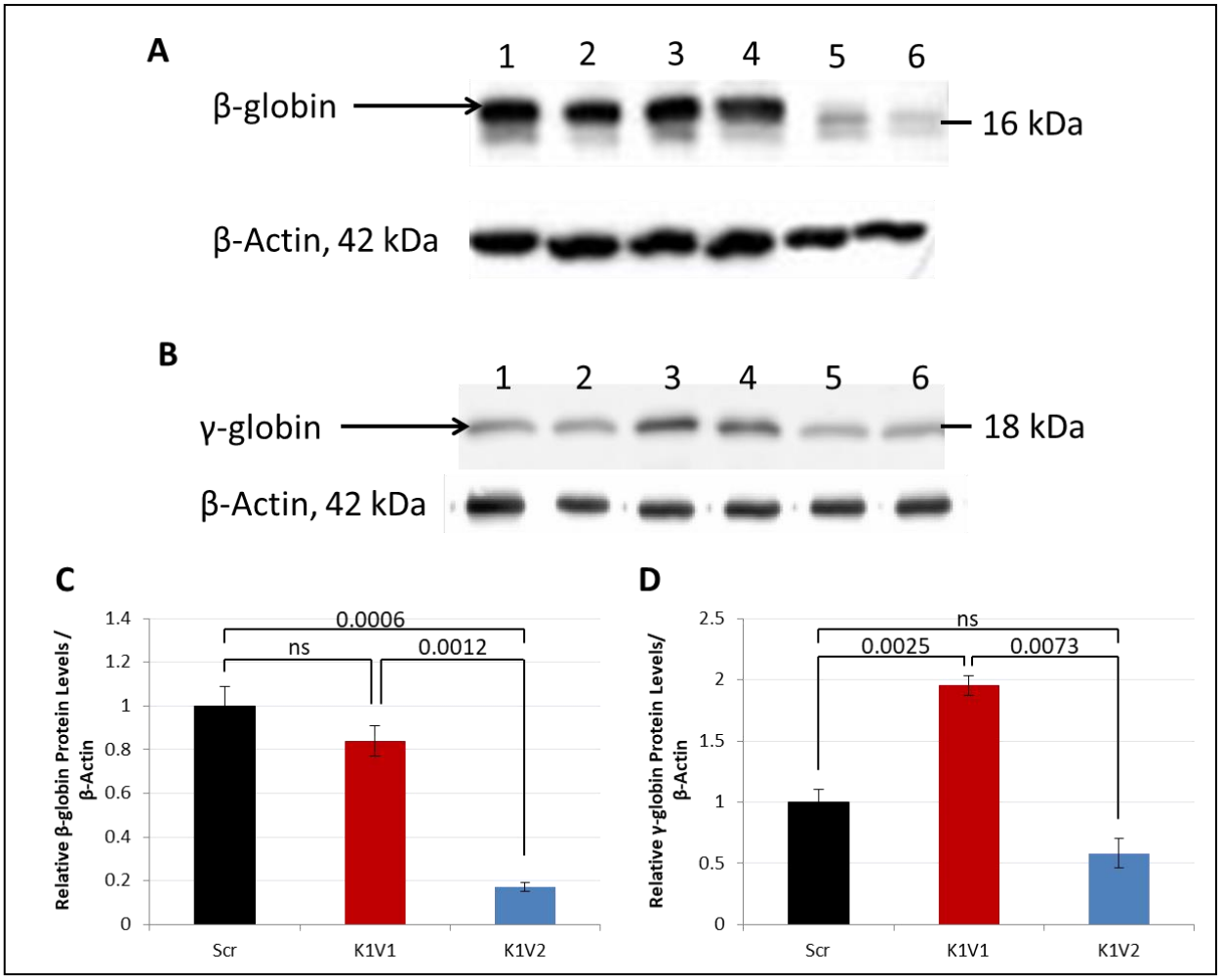


Figure 3.6: γ-globin and β-globin Protein Levels Correlate with mRNA Levels in HUDEP-2 Cells

A) β-globin and B) γ-globin protein levels with respect to corresponding KLF1 protein levels. 1st – 2nd lanes are scramble controls, 3rd – 4th lanes are the K1V1 (KLF1 modest KD) samples and 5th – 6th lanes are the K1V2 (KLF1 robust KD) samples. Relative C) β-globin and D) γ-globin protein levels sorted by category. The error bars indicate SD; N = 3 or 4 per category; P-values as indicated, ns = not significant (Paired-Student's one-tailed test, $P < 0.025$)

3.3.3 β-globin and KLF1 mRNA Amounts Show a Linear Correlation in HUDEP-2 Cells

After analyzing the qRT-PCR data, the average mRNA expression level normalized to Cyclophilin A was compared between 5 scramble (control), 5 K1V1 and 5 K1V2 samples, where the scramble values were set to 100% (Figure 3.7). To test whether our KLF1-shRNAs worked

the average KLF1 mRNA levels were measured in each group and compared to the scramble set to 100 (Figure 3.7 A). The K1V1 category had relative KLF1 mRNA amounts of a mean of 57% and the K1V2 had a mean 25% compared to scramble set to 100%. Similar to the CD34+ erythroid cells, there is an increase in γ-globin mRNA expression in the K1V1 category of 163% as compared to the scramble control samples set at 100%, which is statistically significant. The K1V2 knockdown category shows a mean of 98% γ-globin mRNA, which is not statistically significant to the scramble control samples. Next, the β-globin mRNA levels were observed in each category (Figure 3.7 B). There was no statistically significant decrease in relative expression of β-globin mRNA when K1V1 and scramble control categories were compared, however there was a significant difference between K1V2 and scramble control (Figure 3.7 C). This was been seen previously in primary cells (Vinjamur DS et al., 2016). When all the samples were plotted with the scramble control set to 100%, the results are interesting (Figure 3.8). A linear regression was performed to examine the correlation between β-globin (Figure 3.8 A). There is a direct positive correlation with an R^2 value = 0.92 that is statistically significant for β-globin and KLF1 mRNA levels. Therefore, it can be said that regulation of β-globin by KLF1 is linear.

3.3.4 γ-globin and KLF1 mRNA Amounts Have a Non-Linear Relationship in HUDEP-2 Cells

A linear regression was performed to examine the correlation between γ-globin (Figure 3.8 C) mRNA levels to the residual percent of KLF1 present. There is no direct positive correlation between γ-globin and KLF1 mRNA levels. A smooth curve was fit to the graph, to determine if the threshold amount of KLF1 with respect to γ-globin mRNA levels is seen (Figure 3.8 D). In fact, the results obtained in the HUDEP-2 cells were much better than expected, in that, the γ-globin mRNA to KLF1 mRNA curve is very tight. Therefore, it can be said that KLF1 regulation of γ-globin is non-linear. A threshold amount of KLF1 is required for induced g-globin mRNA. It is clear in figure 3.8 D that when residual KLF1 is between 50% to 60% we observe γ-globin mRNA increases nearly 2-fold. With knockdown less or greater than this, baseline mRNA levels of γ-globin amounts are observed.

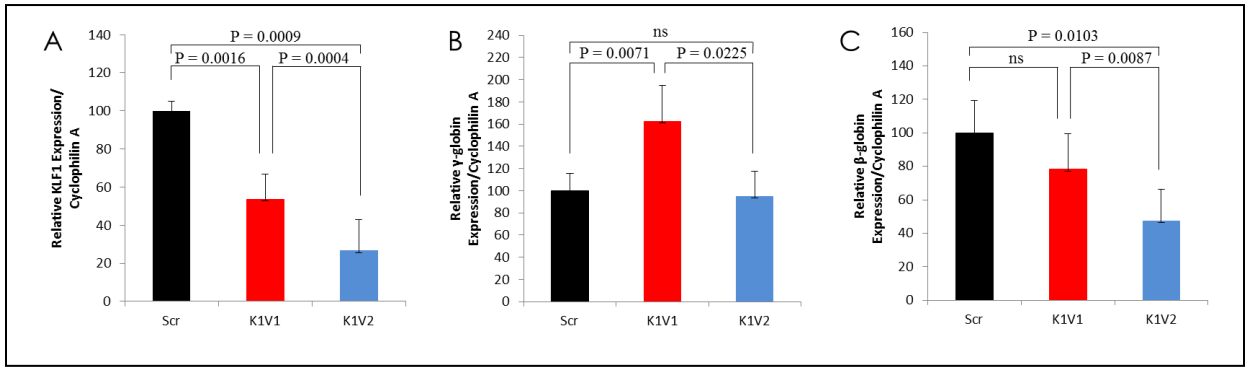


Figure 3.7: γ-globin and β-globin mRNA Levels in HUDEP-2 Cells Mimic Those Seen in CD34+ Cells Isolated from Umbilical Cord Blood
The amount of A) KLF1 B) γ-globin and C) β-globin mRNA was measured by qRT-PCR and normalized to Cyclophilin A. mRNA in K1V1 and K1V2 categories as compared to scramble controls set to 100. The error bars indicate SD; N = 5 per category; P-values as indicated, ns = not significant (Paired-Student's one-tailed test).

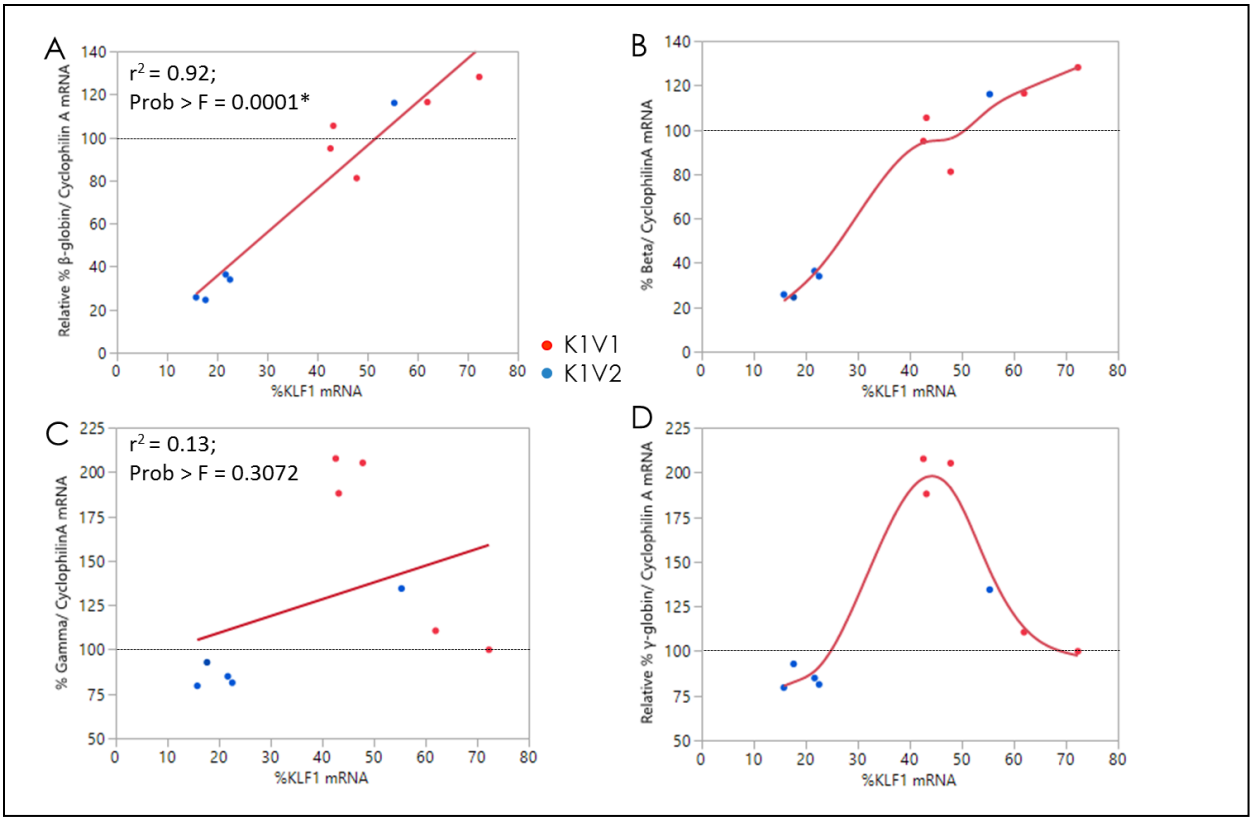


Figure 3.8: Regulation of β-globin and γ-globin Expression by KLF1 in HUDEP-2 Cells Follows Both a Linear and Non-Linear Pattern
The amount of KLF1 mRNA and β-globin or γ-globin mRNA was measured by qRT-PCR and normalized to Cyclophilin A mRNA. The fold change in KLF1 expression and γ-globin expression in each sample was calculated by setting the value for each gene in scramble controls to 100. A) A linear correlation that is statistically significant is observed between amount if KLF1 mRNA and corresponding amount of β-globin mRNA by linear regression analysis. C) No correlation was observed between the amount of KLF1 mRNA present in cells and the corresponding amount of γ-globin expression by linear regression analysis. B & D) when the data was analyzed by allowing the statistical program JMP to fit a non-linear model to the available data in an unbiased manner, the smooth curves shown in the figures were obtained. The non-linear model supports our interpretation of the data with regard to sensitivity of the regulation of γ-globin gene expression to the amount of KLF1 available in cells. n = 5 per category. Each dot on the scatter-plot represents a biological sample, i.e., cells obtained from one HUDEP-2 cell sample.

3.4 KLF1 Indirectly Downregulates γ-globin Gene Expression

Similar to the qRT-PCR experiments performed in mice, the same five known γ-globin gene repressors were examined in HUDEP-2 cells. That is the Mi2-β, MBD2, c-MYB, LRF and LSD1

genes, and whether they are regulated by KLF1. This was performed by lentiviral infection of the various KLF1 knockdowns in HUDEP-2 cells.

3.4.1 KLF1 Regulates Human LRF, LSD1 and c-MYB Genes in HUDEP-2 cells

After analyzing the qRT-PCR data, the average mRNA expression level normalized to Cyclophilin A was compared between 5 scramble (control), 5 K1V1 and 5 K1V2 samples, where the scramble values were set to 100 (Figure 3.9). LRF was the only gene to be significantly downregulated with modest knockdown of KLF1, i.e., the LRF mRNA amount with the modest knockdown was 89± 4 (Figure 3.9 A). The human c-MYB, LRF and LSD1 mRNAs showed a statistically significant difference between the robust knockdown and scramble (Figure 3.9 A – C). The gene expression levels with robust knockdown were 90±7, 70±6 and 89±10 for c-MYB, LRF and LSD1 respectively, when compared to the scramble. MBD2 and Mi2-β showed no statistically significant change in gene expression levels at any knockdown level (Figure 3.9 D & E). A single sample, one-paired, one-tailed t-test was performed to show the values that are statistically significant. Based on our data from the HUDEP-2 cells it can be concluded that human KLF1 positively regulates mouse LRF with the modest knockdown and the human c-MYB, LRF, LSD1 and MBD2 genes with the robust knockdown and does not regulate Mi2-β at all. Interestingly, these same four genes showed statistically significant change in relative gene expression when comparing the modest and robust knockdowns. Based on our data it can be concluded that human KLF1 positively regulates human LRF with modest knockdown, positively regulates c-MYB, LRF and LSD1 with robust knockdown and negatively regulates MBD2 with a robust knockdown in HUDEP-2 cells. KLF1 does not have any detectable effect on the Mi2-β gene expression in HUDEP-2 cells.

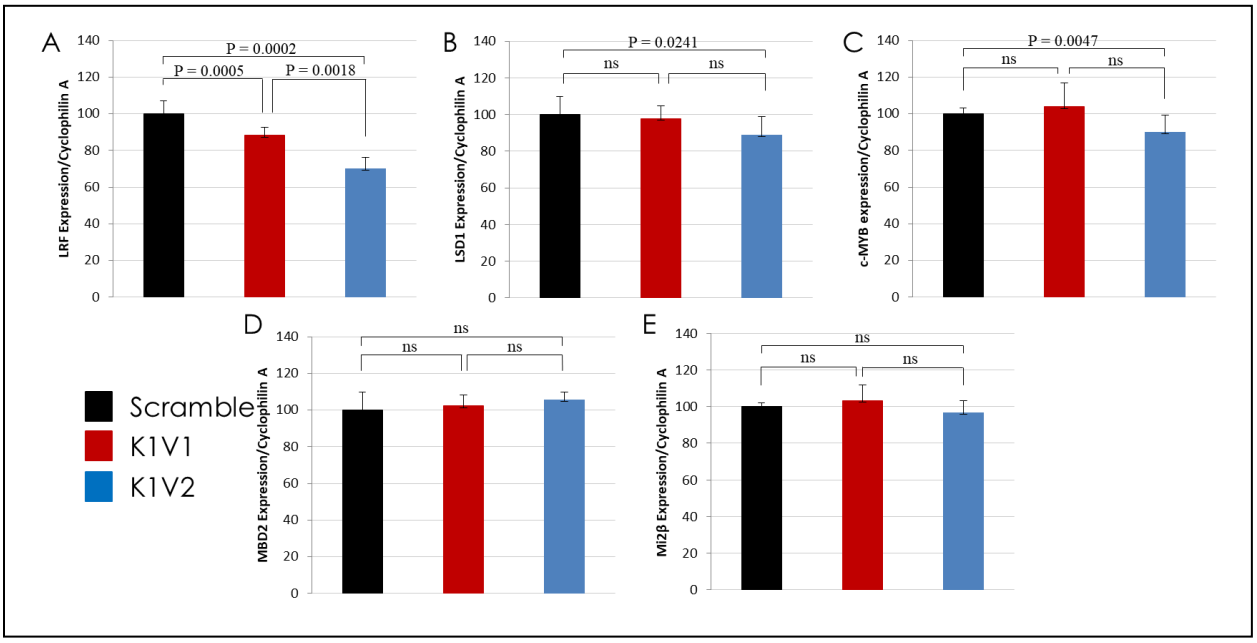


Figure 3.9: KLF1 Regulates Human LRF, LSD1 and c-MYB Genes in HUDEP-2 Cells
A qRT-PCR was performed on γ-globin gene repressors A) c-MYB, B) LRF, C) LSD1, D) MBD2 and E) Mi2-β. 5 samples each of HUDEP-2 cells infected with scramble, K1V1 or K1V2 were collected, for RNA extraction. The amount of KLF1 mRNA was measured using qRT-PCR methods and expressed using the EdCT method to allow for direct comparison of mRNAs to scramble (set to 100). Cyclophilin A was used as a normalization control. N=5 per category; Error bars represent SD; P-values as indicated, ns = not significant (Paired-Student's one-tailed test, $P < 0.025$).

3.4.2 KLF1 Binds the Promoter Region of the Human LRF Gene in HUDEP-2 Cells

Next we wanted to determine whether the c-MYB, LRF, LSD1 and MBD2 genes are bound by KLF1 by performing a human KLF1-ChIP. Before beginning the human KLF1-ChIP assays the 9 base pair and complementary consensus sites were taken from Table 3.2, for these four genes. Out of all consensuses binding sites present in the promoter regions of c-MYB, LRF, LSD1and MBD2 we selected 6 sites for the LRF (Figure 3.10 A), three for the LSD1 (Figure 3.10 B), three for the c-MYB (Figure 3.10 C) and four for the MBD2 (Figure 3.10 D) genes which are highlighted in red. After performing the ChIP assay the chromatin was run on a qPCR and the data was analyzed, the average size of the fragments are on average between 400 bp to

800 bp, seen in the 1st to 3rd lanes of the gel (Figure 3.11). There was a 3.5 fold increase of the -478 to -470, a 4.2 fold increase of the -409 to -351 and a 4.5 fold increase of the -25 to -9 relative fold enrichment in the promoter region of the human LRF gene when compared to human IgG (Figure 3.12). The human β-globin and γ-globin gene promoters, used as controls, showed an 8.1 and 6.4 increase in fold enrichment for KLF1, respectively. The negative control, Necdin, showed no statistically significant change in relative fold change with respect to IgG. To summarize, from our data it appears that KLF1upregulates c-MYB, LRF and LSD1 while simultaneously downregulating MBD2. KLF1 directly binds the promoter sites of the LRF gene. There is no evidence that KLF1 binds the promoter regions of the other three genes. Interestingly, it is observed that the γ-globin promoter is bound by KLF1 more in HUDEP-2 cells than in erythroblasts derived from umbilical cord blood CD34+ cells (Alhashem Y et al., 2011). KLF1 binding γ-globin: β-globin is ~4:3 in HUDEP-2 samples and ~1:1 in primary erythroid cells derived from $CD34^+$ umbilical cord blood cells.

Table 3.2: The Human Consensus Sequences Located in the Promoter Regions of LSD1, LRF, MBD2 and c-MYB Genes.

Human Gene	Position of CACCC from 5' UTR of TSS	Position of CGCCC from 5' UTR of TSS	Position of GGGTG from 5' UTR of TSS	Position of GGGCG from 5' UTR of TSS
LSD-1 (KDM1A)	-	-	-244 to -252 TGGGGTGAG -18 to -26 TGGGTGTCT	-31 to -39 GGGGCGGGG
LRF (ZBTB7A)	-470 to -478 ACCCACCCC -294 to -301 CCCACCCC -283 to -290 TCCACCCC	-400 to -408 CCCGCCCCC -379 to -387 CCACGCCCC -360 to -352 CCGCGCCCC -336 to -343 CCCGCCCC	-	-200 to -207 GGGGCGGA -166 to -173 GGGGCGGG -17 to -25 GGGGCGGGC -8 to -16 TGGGCGTGG
MBD2	-221 to -213 TATCACCCC -77 to -69 TGACACCCC	-237 to -230 GGGCGGGA	-26 to – 18 CCAGTGGGG	-238 to -230 GGGGCGGGA +01 to +08 GGGGCGTG
C-Myb	-454 to -446 CCCCACCGC -258 to -250 GGCCACCCC	-468 to -460 TGTGCGGGA	-	-

Figure 3.10: Promoter Regions of Human LRF, LSD1, c-MYB and MBD2 Genes
Promoter regions were acquired from Ensembl. The consensus sites are highlighted in yellow and primers in red or blue. A) LRF promoter region with LRF1, LRF2, LRF3 and LRF4 consensus sites, B) LSD1 promoter region with LSD1-1, LSD1-2 and LSD1-3 consensus sites, C) c-MYB promoter region with c-MYB1, c-MYB2 and c-MYB3 consensus sites and D) MBD2 promoter region with MBD2-1, MBD2-2, MBD2-3 and MBD2-4 consensus sites.

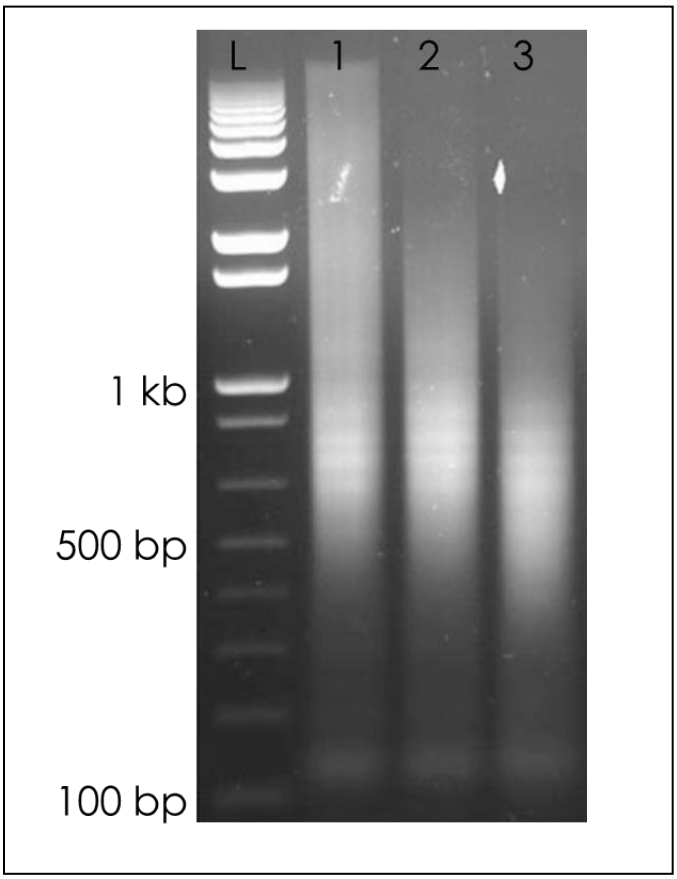


Figure 3.11: Average Chromatin Fragment Size after Sonication of HUDEP-2 Cells
A 1.5% agarose gel was run with 100 base pair ladder. 1st – 3rd lanes are chromatin fragments of HUDEP-2 cells.

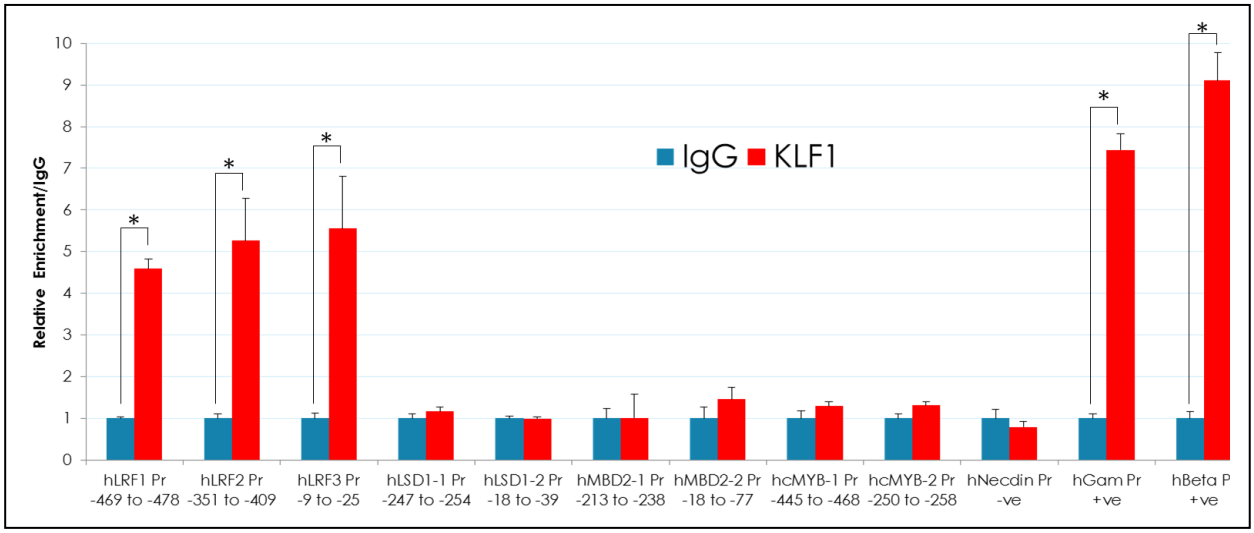


Figure 3.12: KLF1 Binds the Promoter Regions of the Human LRF Gene in HUDEP-2 Cells

ChIP assays were performed on, non-transgenic mice. Polyclonal antibody specific for human KLF1 and non-specific IgG control antibody were used. Sequences highlighted in red and blue are the primers. A) LRF Promoter Region with LRF1, LRF2 and LRF3 consensus sites highlighted in yellow and B) LSD1 Promoter Region with LSD1-1 and LSD1-2 consensus sites highlighted in yellow. C) c-MYB promoter region with c-MYB-1 and c-MYB-2 consensus sites highlighted in yellow. D) MBD2 Promoter region with MBD2-1 and MBD2-2 consensus sites highlighted in yellow. The Y-axis represents the relative fold enrichment. The mean IgG enrichment was set as 1.0 and the enrichment of KLF1 was scaled appropriately. The X- axis shows the location of the primers used for qPCR (Pr: promoter). The error bars indicate standard error mean (SEM); N = 3; *, $P < 0.05$ (Paired-Student's one-tailed test)

Chapter 4: Discussion and Future Directions

4.1 Discussion

A main focus of this research was to understand the intricate mechanisms by which KLF1 regulates γ-globin in both the mouse and human systems. Based on the data in this thesis, mouse KLF1 binding to γ-globin repressor genes showed major differences between E12.5 and E14.5 mouse fetal livers. In E14.5 fetal livers, KLF1 binding is in agreement with ChIP-seq data from 2011, which shows no evidence that KLF1, binds to the promoter regions of the LSD1 or LRF genes (Tallack et al., 2010; Pilon AM et al., 2011). On the other hand, KLF1 in E12.5 mouse fetal livers directly binds the LSD1 and LRF promoter regions, which goes along with a ChIP-seq performed in the murine cell line, K1-ER cells (Norton LJ et al., 2017) which identified the LRF promoter region as having one of the most highly enriched KLF1 peaks (Funnell AP et al., 2012; Hodge D et al., 2006). Another paper showed no binding to either of these genes in the K1-ER cell line (Tallack MR et al., 2010). The K1-ER cell lines are used to study the role of KLF1 in erythropoiesis. They were created from KLF1-/- mouse fetal liver progenitor cells that carry one copy of the human β-globin locus (Coghill E et al., 2001). Also E14.5 KLF1 -/- fetal livers showed a significant reduction in expression of LRF when compared to controls (Norton LJ et al., 2017).

There are multiple possible reasons why we see differences in KLF1 binding at different developmental time points in the mouse LRF gene. First, Norton et al. put forward the theory that KLF1 and GATA-1 both act to upregulate LRF since GATA-1 activates expression of

both LRF and KLF1 (Crossley M et al., 1994; Martin DI et al., 1990; Norton LJ et al., 2017). Second, mouse KLF1 plays a direct role in repressing βh1-globin in earlier embryonic stage E12.5 and gradually moves to a more indirect role of repressing γ-globin by E14.5. LRF is a downstream target of GATA-1 (Maeda T et al., 2009), a transcription factor required for terminal erythroid maturation as mentioned before. What is proposed is LRF is regulated by KLF1 at E12.5 when terminal erythroid maturation is not in full swing, then at E14.5 GATA-1 binds and activates LRF for this purpose of terminal maturation of erythroblasts.

LRF-/- mouse embryos at E12.5 have no morphological abnormalities, by E14.5 gross morphological abnormalities can be seen (Maeda T et al., 2009). LRF-/- fetuses die *in utero* at E15.5 to E16.5 due to lack of differentiation from erythroblast stage to mature erythrocytes, so it seems logical that by E14.5 GATA-1 would be responsible for all binding to LRF in E14.5 mice.

KLF1 also binds to the LSD1 promoter at E12.5 but not E14.5. LSD1 removes H3K4me2 an activation methylation mark (Nicholson TB et al., 2009). Therefore we propose that in E12.5 mouse fetal livers, LSD1 is activated by KLF1, to repress the mouse embryonic-globin genes. It is possible that this histone modification is then kept as an epigenetic marker during the M phase of the cell cycle. The next generation of cells could inherit this epigenetic mark and so by E14.5 the mouse embryonic-globin gene is still being repressed. Therefore, the need for KLF1 to repress the embryonic-globin genes via activation of LSD1 is no longer required by E14.5 fetuses.

The recently developed HUDEP-2 cell line has allowed researchers to study regulation of γ- and β-globin genes in a more regulated environment, i.e. in a human erythroid cell line. Unlike the other human erythroid cell lines, K562 and HEL, the HUDEP-2 cell line is

representative of an adult erythroid system and these cells are not chromosomally aberrant (Nakamura Y et al., 2013). The erythroid HUDEP-2 cells have 97% β-globin and < 2% γ-globin content which is considered to be an 'adult-like' system. This is in comparison to primary human erythroid cells isolated from umbilical cord blood that have higher fetal-like globin levels, i.e. ~50% γ-globin mRNA (Alhashem YN, 2012). From the raw values of γ- and β-globin in our transfected scramble control HUDEP-2 cells, it was observed that γ / (γ + β) globin is equivalent to 11%, but our non-transfected mock control γ/ (γ+β) globin is negligible. Current research shows that HUDEP-2 lentiviral treated control samples have negligible amounts of γ-globin with differentiation (Masuda T et al., 2016) and also has approximately 3% to 4% amount of γ-globin (Traxler E et al., 2016). To date, our data shows the highest percent of γ-globin induced in the HUDEP-2 control samples. A logical explanation for the relatively high amount of γ-globin observed in scramble control is the three days of differentiation versus 7 days in the Masuda and Traxler papers. Many cite the fact that infection of control CD34+ cells with a lentivirus leads to increased HbF (Masuda T et al., 2016; Vinjamur DS et al., 2016).

When the scramble value is set to 100 and the K1V1 and K1V2 values are adjusted accordingly we observe that γ-globin gene expression is sensitive to KLF1 dosage. The robust knockdown of KLF1 (<25% residual KLF1) in samples showed no significant increase or decrease in γ-globin mRNA compared to the control. This is the same phenomenon observed by our lab previously, termed the threshold amount (Vinjamur DS et al., 2016). That is, the cells have inadequate amounts of KLF1 to increase γ-globin with robust KLF1 knockdown. Samples with residual KLF1 between 40% and 50% (K1V1), showed 2-fold higher γ-globin expression than in control samples. This confirms the same

observations made by our lab in the human primary erythroid cells isolated from umbilical cord blood (Vinjamur DS et al., 2016). The relative density or intensity of the 16 kDa β-globin and 18 kDa γ-globin protein bands further confirm that what is occurring at the mRNA level upon various KLF1 knockdowns is also seen at the protein level. It should be noted that other research shows that there is a 6- to 10-fold increase of γ-globin expression in adult erythroid cells where KLF1 is depleted (Borg J et al., 2010; Zhou D et al., 2010). Reasoning for this higher induction of γ-globin than we have observed could be due to the fact that these samples were differentiated for 7-10 days when γ-globin expression levels are negligible. Therefore, any increase in γ-globin expression would be multi-fold higher. Previous research has shown that γ-globin expression decreases when ≤ 10% residual KLF1 is present in primary erythroid cells and upon observing KLF1 -/- βYAC mice (Vinjamur DS Dissertation, 2014; Alhashem YN et al., 2011). In agreement with this research, we also noted a slight decrease in γ-globin gene expression with robust knockdowns of KLF1when compared to scramble control γ-globin expression levels.

Elevated γ-globin expression observed in HUDEP-2 cells with between 40 - 50% residual KLF1 can be explained with data from this research: First, the dysregulation of γ-globin gene repressor LRF by being directly bound by KLF1 in the promoter region. Second, indirect upregulation of Mi2β and LSD1 by KLF1, which are known γ-globin repressor genes.

The threshold amount of KLF1 defined as having at least 40% KLF1 required to see induction of increased γ-globin expression is seen in the HUDEP-2 cells. Reasoning for this could be: Reduced competition of the γ-globin promoter for the LCR due to reduced KLF1 binding to its promoter. The probable hypothesis as to why there is a normal amount of γ-globin expression and a major decrease in β-globin expression upon depletion of residual

KLF1 to ≤ 25% may be: A result of multiple processes that require KLF1such as recruitment of β- and γ-globin genes to transcription factories as well as active chromatin hub formation at the β-globin locus and interaction between the LCR and the globin gene promoters (Drissen et al. 2004; Schoenfelder et al. 2010).

The protein levels of β -globin were respectively ~4.8 fold lower in the K1V2 category as compared to the scramble and K1V1 categories. This is in agreement with mRNA obtained in human primary cells where a significant decrease in β-globin mRNA is noted when compared to the other categories, though that showed a 2-fold decrease in expression. γ-globin protein levels were respectively ~2.0-fold higher in the K1V1 category as compared to the scramble and K1V2 categories. This is also in agreement with mRNA obtained in human primary cells where a similar increase in γ-globin mRNA is noted when compared to the other categories. This verifies that there is a threshold amount of ~50% KLF1 required to increase γ-globin protein expression to a significant extent and that the robust knockdown shows baseline levels of γ-globin protein. Therefore, it can be stated that knockdown of 50% KLF1 required to increase γ-globin protein expression would be beneficial in treating SCD or β-thalassemic patients.

Expression levels of γ-globin mRNA do not vary significantly when comparing different HUDEP-2 samples with relatively similar amounts of residual KLF1. This is in contrast to γ-globin expression levels among different samples of primary erythroid cells isolated from umbilical cord blood. In the primary cell model, each sample is taken from an individual and hence varies greatly (Vinjamur DS Dissertation, 2014), i.e. a cell line derived from one patient (HUDEP-2 cells) vs. primary cells from different patients. Overall we conclude that the HUDEP-2 cells mimic the CD34+ Cells isolated from umbilical cord blood with respect

to regulation of γ-globin expression by KLF1. Based on evidence from this study, it is suggested HUDEP-2 cells can be used as an alternative to erythroid primary cells and cell lines. This is because they more closely resemble the adult erythroid system than do primary cells and they are not chromosomally aberrant as are other popular cell lines.

This research shows that human KLF1 directly binds the LRF gene promoter, which is in agreement with recent literature (Norton LJ et al., 2017). Though fragment sizes were too large to discern between different consensus binding sites in our data, this research showed that KLF1 binds multiple binding sites in the LRF promoter (Norton LJ et al., 2017). This data thus gives more merit to our hypothesis that elevated γ-globin expression observed in cells with 40% to 50% residual KLF1 is caused by the dysregulation of γ-globin repressors which are regulated by KLF1. KLF1 downregulates MBD2 and upregulate LSD1 and c-MYB, but does not bind them in their promoter regions. This could indicate that KLF1 with many binding sites found at large distances from any of these known genes in mice, could act at distant enhancers (Tallack et al., 2010) or might function by chromatin looping as seen particularly with the β-globin locus (Drissen et al., 2004) or directly binds regions of these genes other than the promoter. Another possibility is KLF1 indirectly regulates these genes via another repressor in the case of MBD2 or an activator in the case of LSD1 and c-MYB. The γ-globin gene repressor c-MYB binds the KLF1 promoter and transactivates its expression (Bianchi E et al., 2010). Based on this research we also see that KLF1 activates c-MYB though it does not bind its promoter. KLF1 and c-MYB could act as activators of each other to ensure repression of γ-globin just as BCL11A and KLF1 do by both activating BCL11A.

Also of note, KLF1 shows more binding to the γ-globin gene promoter region than previously seen in primary erythroid cells (Vinjamur DS et al., 2016). KLF1 may have an indirect role in γ-globin gene regulation during adult erythropoiesis which is supported by evidence from previous research (Borg J et al., 2010; Zhou D et al., 2010; Arnaud et al., 2010; Tallack MR et al., 2013). The γ- to β –globin ratio in human primary cells derived from umbilical cord blood is ~ 1:1 (Alhashem YN et al., 2011) whereas in HUDEP-2 cells based on our findings it is 1:9. When we look at evidence on the ratio of KLF1 binding to γ-globin to the β-globin promoter we see that: E11.5 βYAC mouse primitive erythroid cells show an 5:1 ratio (Alhashem YN et al., 2011); HUDEP-2 cells show a 1:1 (Weinert B et al., 2017) and in our data, HUDEP-2 cells after 3 days of differentiation show a 7.5:9 ratio. This suggests that KLF1 directly binds and activates the γ-globin gene in its promoter and with increased binding for γ-globin over β-globin promoter. Our own evidence suggests KLF1 indirectly represses γ-globin during adult erythropoiesis by activating LSD1 and LRF in E12.5 mouse fetal livers and activating LRF, LSD1 and c-MYB in HUDEP-2 cells. Therefore, we conclude that KLF1 appears to have a direct role in activating and an indirect repressive role on γ-globin gene expression in humans.

The human and mouse LRF genes in HUDEP-2 cells (Norton LJ et al., 2017) and E12.5 mouse fetal livers are bound by KLF1 based on our research. This is accordance with other research that has shown that >60% of KLF1 occupancy peaks along the proerythroblast lineage are conserved between species (Ulirsch JC et al., 2014; Pishesha N et al., 2014 Therefore, it can be concluded that KLF1 binding the LRF gene is evolutionarily conserved between mice and humans.

Based on current research and this research, the proposed human model on how KLF1 regulates γ-globin gene expression can be explained as such: KLF1 directly binds γ-globin gene to activate expression (Weinert B et al., 2017; Alhashem YN et al., 2011). This is because the γ-globin gene precludes expression of the β-globin gene in early transgenic mouse development due to the major γ-globin promoter element, CACCC, which suppresses β-globin gene expression in the fetus (Sargent TG et al., 1999; Sargent TG et al, 2001). It was concluded that a protein(s) such as KLF1 binds to the γ-globin CACCC element to coordinate stage-specific gene expression (Sargent TG et al., 2001). KLF1 directly binds and activates LRF (Norton LJ et al., 2017) which in-turn represses γ-globin expression (Masuda T et al., 2016). KLF1 indirectly activates c-MYB and LSD1 which in turn repress γ-globin expression (Jiang J et al., 2006; Shi L et al., 2013; Rivers A et al., 2015; Cui S et al., 2015). c-MYB and LSD1 are shown to have indirect activation arrows. This is because this study did not find any indication of KLF1 binding of these genes in their promoters (Figure 4.1). Based on this research, it is more likely that KLF1 will make an excellent therapeutic target for increasing γ-globin protein levels in sickle cell and β-thalassemic patients.

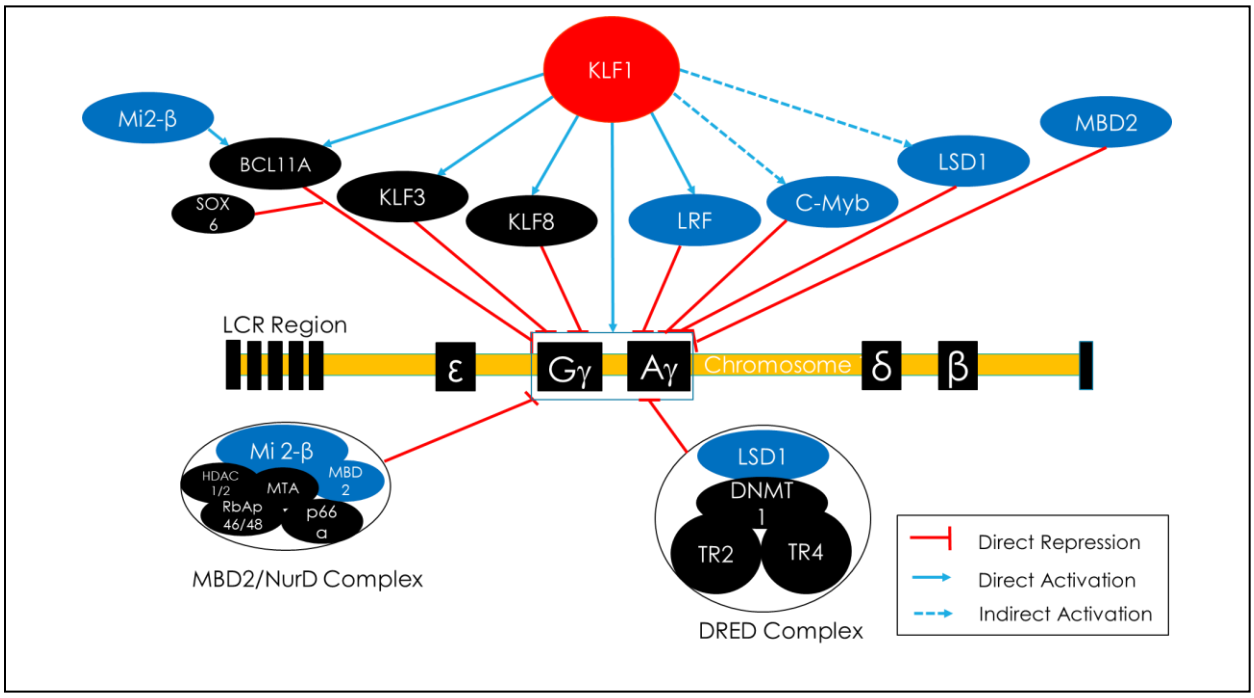


Figure 4.1: Proposed Model of the Human γ-globin Gene Suppression Network

4.2 Future Directions

This work further elucidates the dichotomous role of KLF1 in regulating γ-globin expression in the human and mouse systems. Our results suggest that KLF1 indirectly represses γ-globin gene expression via upregulating γ-globin repressor genes LRF (Masuda T et al., 2016), LSD1 (Shi L et al., 2013) and c-MYB (Jiang J et al., 2006) and in other studies BCL11A (Sankaran VG, 2008). This research also suggests a direct activation of γ-globin gene expression by KLF1. Both these aspects of γ-globin gene regulation by KLF1 can be broadened into further studies explained in the following paragraphs.

To expand on the indirect repressive role of KLF1 on γ-globin expression in the human system, the TR2 and TR4 subunits of the DRED complex, a known γ-globin repressor (Tanabe O et al., 2002), will be studied. This will be done to observe whether KLF1 regulates and binds these genes directly and our lab has already designed primer sets for each of these

genes. The next step in elucidating the extent to which KLF1 indirectly represses γ-globin expression is to perform an unbiased approach of γ-globin repressor genes differentially expressed genes in HUDEP-2 samples with modest knockdown of KLF1 when compared to control samples. A list of genes differentially expressed in the knockdown samples compared to control samples and knockdown samples compared to each other will be compiled. Each of these listed genes will be ranked on criteria that include but are not limited to: fold decrease, function in erythropoiesis/hemoglobin switching, potential role as a transcriptional regulator, etc. Genes selected will then have the same experiments performed as was done on LRF, LSD1, c-MYB, MBD2 and Mi2-β genes in this study.

To validate the direct binding and activation of γ-globin expression by KLF1 a luciferase assay will be performed in HUDEP-2 cells. The luciferase gene will be inserted downstream of the γ-globin gene promoter which includes the 9 base pair consensus sequence required for KLF1 binding. Light emitted by the mutated consensus sequence and the normal site will be compared. If the mutated binding site shows decreased light emitted as compared to the control sample, then it will be concluded that KLF1 directly binds and activates γ-globin gene expression. Data analysis of KLF1 ChIP-seq data on the HUDEP-2 cells would also be of interest to identify direct binding of KLF1 to γ-globin repressor genes (Norton LJ et al., 2017).

The fragment sizes of HUDEP-2 cell chromatin looking at KLF1 binding in LRF were too large to discern which of the consensus sequences KLF1 binds. Therefore, another experiment would identify which of these consensus sequences by KLF1 is most efficacious in activating human LRF. This will also be done by performing a luciferase assay with each site being mutated and then compared to the control sample. The LRF promoter region from

+1 to -500 of the LRF TSS will be cloned upstream of the luciferase gene in an expression vector, each clone will have one consensus site mutated compared to a control with no mutations. The DNA vector will be introduced into the HUDEP-2 cells (or other sample).

The HUDEP-2 cell line is a relatively new human adult-like system used in this field. So it is assumed that the cells at differentiation day 3 resemble those human primary cells at differentiation day 8 with regard to erythroblast maturity. To verify that this is the case, first, cytospin slides at HUDEP-2 differentiation day 3are prepared to score the level of erythroid maturity as previously done in human primary cells (Alhashem YN, 2012). The scoring is based on the number of cells out of 100 at each erythroblast stage, i.e. basophilic, polychromatic or orthochromatic erythroid cells. We would expect to see no significant difference in the number of cells at each stage in the K1V1 and K1V2 samples when compared to the control. Second, fluorescent-activated cell sorting (FACS) will be performed to identify the percentage of cells with transferrin (CD71) and glycophorin A (CD235a) receptors in K1V1 and K1V2 knockdown compared to the scramble control.

The ultimate aim of a study like this is to design an ideal targeted therapy strategy to increase γ-globin expression and alleviate symptoms associated with sickle cell disease. That is why, we would like to compare levels of five to 6 of the most efficacious γ-globin repressor genes in sickle cell disease patients with low HbF and high HbF. ShRNAs will be created to knockdown each gene individually for testing which gene or combination of two genes gives the highest induction of HbF.

Finally, our work confirms the role of KLF1 in regulating γ-globin gene expression and shows that modest knockdown of KLF1 is a viable option for induction of γ-globin gene

expression. In addition it establishes the conserved binding by KLF1 in both the mouse and human systems to the γ-globin repressor gene LRF.

References

Agrawal, R.K., Patel, R.K., Nainiwal, L. and Trivedi, B., 2014. Hydroxyurea in sickle cell disease: drug review. *Indian Journal of Hematology and Blood Transfusion*, *30*(2), pp.91-96.

Alhashem, Y.N., Vinjamur, D.S., Basu, M., Klingmüller, U., Gaensler, K.M. and Lloyd, J.A., 2011. Transcription factors KLF1 and KLF2 positively regulate embryonic and fetal β-globin genes through direct promoter binding. *Journal of Biological Chemistry*, *286*(28), pp.24819-24827.

Alhashem YN. 2012. Regulation of the mouse and human β-globin genes by Krüppel-like transcription factors KLF1 and KLF2. Doctoral thesis. Virginia Commonwealth University.

Amaya, M., Desai, M., Gnanapragasam, M.N., Wang, S.Z., Zu Zhu, S., Williams, D.C. and Ginder, G.D., 2013. Mi2β-mediated silencing of the fetal γ-globin gene in adult erythroid cells. *Blood*, *121*(17), pp.3493-3501.

Anagnou, N.P., Perez-Stable, C., Gelinas, G., Costantini, C., Liapaki, L., Constantopoulou, C., Kosteas, K., Moschonas, N.K. and Stamatoyannopoulos, G., 1995. Sequences located 3′ to the breakpoint of the hereditary persistence of fetal hemoglobin-3 deletion exhibit enhancer activity and can modify the developmental expression of the human fetal Aγ-globin gene in transgenic mice. *Journal of Biological Chemistry*, *270*(17), pp.10256-10263.

Antoniou, M., 1991. Induction of erythroid-specific expression in murine erythroleukemia (MEL) cell lines. *Gene Transfer and Expression Protocols*, pp.421-434.

Armstrong, J.A., Bieker, J.J. and Emerson, B.M., 1998. A SWI/SNF–related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. *Cell*, *95*(1), pp.93-104.

Arnaud, L., Saison, C., Helias, V., Lucien, N., Steschenko, D., Giarratana, M.C., Prehu, C., Foliguet, B., Montout, L., de Brevern, A.G. and Francina, A., 2010. A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. *The American Journal of Human Genetics*, *87*(5), pp.721-727.

Bank A, Mears JG, Ramirez F. 1980. Disorders of human hemoglobin. Science 207:486-493.

Baron, M.H. and Fraser, S.T., 2005. The specification of early hematopoiesis in the mammal. *Current opinion in hematology*, *12*(3), pp.217-221.

Basu P, Lung TK, Lemsaddek W, Sargent TG, Williams DC,Jr, Basu M, Redmond LC, Lingrel JB, Haar JL, Lloyd JA. 2007. EKLF and KLF2 have compensatory roles in embryonic beta-globin gene expression and primitive erythropoiesis. Blood 110(9):3417-25.

Bauer, D.E., Kamran, S.C. and Orkin, S.H., 2012. Reawakening fetal hemoglobin: prospects for new therapies for the β-globin disorders. *Blood*, *120*(15), pp.2945-2953.

Begley, C.G., Aplan, P.D., Denning, S.M., Haynes, B.F., Waldmann, T.A. and Kirsch, I.R., 1989. The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. *Proceedings of the National Academy of Sciences*, *86*(24), pp.10128-10132.

Behringer, R.R., Ryan, T.M., Palmiter, R.D., Brinster, R.L. and Townes, T.M., 1990. Human gamma-to beta-globin gene switching in transgenic mice. *Genes & Development*, *4*(3), pp.380-389.

Bennett, G.D. and Kay, M.M., 1981. Homeostatic removal of senescent murine erythrocytes by splenic macrophages. *Experimental hematology*, *9*(3), pp.297-307.

Bhagat, S., Patra, P.K. and Thakur, A.S., 2012. Association between XmnI polymorphism and HbF level in sickle cell disease patients from Chhattisgarh. *Int J Biomed Sci*, *1*, pp.36-39.

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*, *116*(22), pp.e99-e110.

Boehm, T., Foroni, L., Kaneko, Y., Perutz, M.F. and Rabbitts, T.H., 1991. The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13. *Proceedings of the National Academy of Sciences*, *88*(10), pp.4367-4371.

Borg, J., Papadopoulos, P., Georgitsi, M., Gutiérrez, L., Grech, G., Fanis, P., Phylactides, M., Verkerk, A.J., van der Spek, P.J., Scerri, C.A. and Cassar, W., 2010. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nature genetics*, *42*(9), pp.801-805.

Borg, J., Patrinos, G.P., Felice, A.E. and Philipsen, S., 2011. Erythroid phenotypes associated with KLF1 mutations.

Bottardi, S., Ross, J., Pierre-Charles, N., Blank, V. and Milot, E., 2006. Lineage-specific activators affect β-globin locus chromatin in multipotent hematopoietic progenitors. *The EMBO Journal*, *25*(15), pp.3586-3595.

Brawley, O.W., Cornelius, L.J., Edwards, L.R., Gamble, V.N., Green, B.L., Inturrisi, C., James, A.H., Laraque, D., Mendez, M., Montoya, C.J. and Pollock, B.H., 2008. National Institutes of Health Consensus Development Conference statement: hydroxyurea treatment for sickle cell disease. *Annals of internal medicine*, *148*(12), pp.932-938.

Cantor AB, Orkin SH (2002) Transcriptional regulation of erythropoiesis: An affair involving multiple partners. Oncogene 21(21):3368–3376.

Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P. and Luc, S., 2015. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature*.

Cao, A. and Moi, P., 2002. Regulation of the globin genes. *Pediatric research*, *51*(4), pp.415-421.

Carrocini, G.C.D.S., Zamaro, P.J.A. and Bonini-Domingos, C.R., 2011. What influences Hb fetal production in adulthood? *Revista brasileira de hematologia e hemoterapia*, *33*(3), pp.231-236.

Chasis, J.A. and Mohandas, N., 2008. Erythroblastic islands: niches for erythropoiesis. *Blood*, *112*(3), pp.470-478.

Chaturvedi, S. and DeBaun, M.R., 2016. Evolution of sickle cell disease from a life-threatening disease of children to a chronic disease of adults: The last 40 years. *American journal of hematology*, *91*(1), pp.5-14.

Cheng, Y., Wu, W., Kumar, S.A., Yu, D., Deng, W., Tripic, T., King, D.C., Chen, K.B., Zhang, Y., Drautz, D. and Giardine, B., 2009. Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. *Genome research*, *19*(12), pp.2172-2184.

Chien, R., Zeng, W., Kawauchi, S., Bender, M.A., Santos, R., Gregson, H.C., Schmiesing, J.A., Newkirk, D.A., Kong, X., Ball, A.R. and Calof, A.L., 2011. Cohesin mediates chromatin interactions that regulate mammalian β-globin expression. *Journal of Biological Chemistry*, *286*(20), pp.17870-17878.

Choi, K., 1998. Hemangioblast development and regulation. *Biochemistry and cell biology*, *76*(6), pp.947-956.

Coghill, E., Eccleston, S., Fox, V., Cerruti, L., Brown, C., Cunningham, J., Jane, S. and Perkins, A., 2001. Erythroid Kruppel-like factor (EKLF) coordinates erythroid cell proliferation and hemoglobinization in cell lines derived from EKLF null mice. *Blood*, *97*(6), pp.1861-1868.

Costa, F.C., Fedosyuk, H., Chazelle, A.M., Neades, R.Y. and Peterson, K.R., 2012. Mi2β is required for γ-globin gene silencing: temporal assembly of a GATA-1-FOG-1-Mi2 repressor complex in β-YAC transgenic mice. *PLoS Genet*, *8*(12), p.e1003155.

Crawford, L.W., Foley, J.F. and Elmore, S.A., 2010. Histology atlas of the developing mouse hepatobiliary system with emphasis on embryonic days 9.5-18.5. *Toxicologic pathology*, *38*(6), pp.872-906.

Crossley, M. and Orkin, S.H., 1993. Regulation of the β-globin locus. *Current opinion in genetics & development*, *3*(2), pp.232-237.

Cui, S., Kolodziej, K.E., Obara, N., Amaral-Psarris, A., Demmers, J., Shi, L., Engel, J.D., Grosveld, F., Strouboulis, J. and Tanabe, O., 2011. Nuclear receptors TR2 and TR4 recruit multiple epigenetic transcriptional corepressors that associate specifically with the embryonic β-type globin promoters in differentiated adult erythroid cells. *Molecular and cellular biology*, *31*(16), pp.3298-3311.

Cui, S., Lim, K.C., Shi, L., Lee, M., Jearawiriyapaisarn, N., Myers, G., Campbell, A., Harro, D., Iwase, S., Trievel, R.C. and Rivers, A., 2015. The LSD1 inhibitor RN-1 induces fetal hemoglobin synthesis and reduces disease pathology in sickle cell mice. *Blood*, *126*(3), pp.386-396.

Dang, D.T., Pevsner, J. and Yang, V.W., 2000. The biology of the mammalian Krüppel-like family of transcription factors. *The international journal of biochemistry & cell biology*, *32*(11), pp.1103-1121.

Denslow, S.A. and Wade, P.A., 2007. The human Mi-2/NuRD complex and gene regulation. *Oncogene*, *26*(37), pp.5433-5438.

Desgardin, A.D., Abramova, T., Rosanwo, T.O., Kartha, S., Shim, E.H., Jane, S.M. and Cunningham, J.M., 2012. Regulation of delta-aminolevulinic acid dehydratase by krüppel-like factor 1. *PloS one*, *7*(10), p.e46482.

Diamantidis, M.D., Neokleous, N., Agapidou, A., Vetsiou, E., Manafas, A., Fotiou, P. and Vlachaki, E., 2016. Iron chelation therapy of transfusion-dependent [beta]-thalassemia during pregnancy in the era of novel drugs: is deferasirox toxic?. *International journal of hematology*, *103*(5), p.537.

Dillon, N. and Grosveld, F., 1991. Human (gamma)-Globin Genes Silenced Independently of Other Genes in the (beta)-Locus. *Nature*, *350*(6315), p.252.

Donze, D., Townes, T.M. and Bieker, J.J., 1995. Role of erythroid Kruppel-like factor in human-to-globin gene switching. *Journal of Biological Chemistry*, *270*(4), pp.1955-1959.

Dorfman, D.M., Wilson, D.B., Bruns, G.A. and Orkin, S.H., 1992. Human transcription factor GATA-2. Evidence for regulation of preproendothelin-1 gene expression in endothelial cells. *Journal of Biological Chemistry*, *267*(2), pp.1279-1285.

Drissen, R., Palstra, R.J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S. and de Laat, W., 2004. The active spatial organization of the β-globin locus requires the transcription factor EKLF. *Genes & development*, *18*(20), pp.2485-2490.

Drissen, R., von Lindern, M., Kolbus, A., Driegen, S., Steinlein, P., Beug, H., Grosveld, F. and Philipsen, S., 2005. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. *Molecular and cellular biology*, *25*(12), pp.5205-5214.

Dulmovits, B.M., Appiah-Kubi, A.O., Papoin, J., Hale, J., He, M., Al-Abed, Y., Didier, S., Gould, M., Husain-Krautter, S., Singh, S.A. and Chan, K.W., 2016. Pomalidomide reverses γ-globin silencing through the transcriptional reprogramming of adult hematopoietic progenitors. *Blood*, *127*(11), pp.1481-1492.

Duncan, M.T., DeLuca, T.A., Kuo, H.Y., Yi, M., Mrksich, M. and Miller, W.M., 2016. SIRT1 is a critical regulator of K562 cell growth, survival, and differentiation. *Experimental cell research*, *344*(1), pp.40-52.

Dzierzak, E. and Speck, N.A., 2008. Of lineage and legacy: the development of mammalian hematopoietic stem cells. *Nature immunology*, *9*(2), pp.129-136.

Dzierzak, E. and Philipsen, S., 2013. Erythropoiesis: development and differentiation. *Cold Spring Harbor perspectives in medicine*, *3*(4), p.a011601.

Eaton, S.A., Funnell, A.P., Sue, N., Nicholas, H., Pearson, R.C. and Crossley, M., 2008. A network of Krüppel-like factors (Klfs) Klf8 is repressed by Klf3 and activated by Klf1 in vivo. *Journal of Biological Chemistry*, *283*(40), pp.26937-26947.

Emambokus, N., Vegiopoulos, A., Harman, B., Jenkinson, E., Anderson, G. and Frampton, J., 2003. Progression through key stages of haemopoiesis is dependent on distinct threshold levels of c-Myb. *The EMBO Journal*, *22*(17), pp.4478-4488.

Enver, T., Raich, N., Ebens, A.J., Papayannopoulou, T. and Constantini, F., 1990. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature*, *344*(6264), p.309.

Epner, E., Kim, C.G. and Groudine, M., 1992. What does the locus control region control?. *Current Biology*, *2*(5), pp.262-264.

Esteghamat, F., Gillemans, N., Bilic, I., van den Akker, E., Cantù, I., van Gent, T., Klingmüller, U., van Lom, K., von Lindern, M., Grosveld, F. and van Dijk, T.B., 2013. Erythropoiesis and globin switching in compound Klf1:: Bcl11a mutant mice. *Blood*, *121*(13), pp.2553-2562.

Felber, B.K., Orkin, S.H. and Hamer, D.H., 1982. Abnormal RNA splicing causes one form of α thalassemia. *Cell*, *29*(3), pp.895-902.

Feng, W.C., Southwood, C.M. and Bieker, J.J., 1994. Analyses of beta-thalassemia mutant DNA interactions with erythroid Krüppel-like factor (EKLF), an erythroid cell-specific transcription factor. *Journal of Biological Chemistry*, *269*(2), pp.1493-1500.

Feng, Q. and Zhang, Y., 2001. The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. *Genes & development*, *15*(7), pp.827-832.

Finger, L.R., Kagan, J., Christopher, G., Kurtzberg, J., Hershfield, M.S., Nowell, P.C. and Croce, C.M., 1989. Involvement of the TCL5 gene on human chromosome 1 in T-cell leukemia and melanoma. *Proceedings of the National Academy of Sciences*, *86*(13), pp.5039-5043.

Forrester, W.C., Epner, E., Driscoll, M.C., Enver, T., Brice, M., Papayannopoulou, T. and Groudine, M., 1990. A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. *Genes & development*, *4*(10), pp.1637-1649.

Forrester, W.C., Thompson, C., Elder, J.T. and Groudine, M., 1986. A developmentally stable chromatin structure in the human beta-globin gene cluster. *Proceedings of the National Academy of Sciences*, *83*(5), pp.1359-1363.

Frame, J.M., McGrath, K.E. and Palis, J., 2013. Erythro-myeloid progenitors:"definitive" hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. *Blood Cells, Molecules, and Diseases*, *51*(4), pp.220-225.

Fraser, S.T., Isern, J. and Baron, M.H., 2007. Maturation and enucleation of primitive erythroblasts during mouse embryogenesis is accompanied by changes in cell-surface antigen expression. *Blood*, *109*(1), pp.343-352.

Friend, C., 1957. Cell-free transmission in adult Swiss mice of a disease having the character of a leukemia. *Journal of Experimental Medicine*, *105*(4), pp.307-318.

Friend, C., Scher, W., Holland, J.G. and Sato, T., 1971. Hemoglobin synthesis in murine virus-induced leukemic cells in vitro: stimulation of erythroid differentiation by dimethyl sulfoxide. *Proceedings of the National Academy of Sciences*, *68*(2), pp.378-382.

Frontelo, P., Manwani, D., Galdass, M., Karsunky, H., Lohmann, F., Gallagher, P.G. and Bieker, J.J., 2007. Novel role for EKLF in megakaryocyte lineage commitment. *Blood*, *110*(12), pp.3871-3880.

Fujimaki, S.I., Harigae, H., Sugawara, T., Takasawa, N., Sasaki, T. and Kaku, M., 2001. Decreased expression of transcription factor GATA-2 in haematopoietic stem cells in patients with aplastic anaemia. *British journal of haematology*, *113*(1), pp.52-57.

Fujiwara, T., O'Geen, H., Keles, S., Blahnik, K., Linnemann, A.K., Kang, Y.A., Choi, K., Farnham, P.J. and Bresnick, E.H., 2009. Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. *Molecular cell*, *36*(4), pp.667-681.

Funnell, A.P., Mak, K.S., Twine, N.A., Pelka, G.J., Norton, L.J., Radziewic, T., Power, M., Wilkins, M.R., Bell-Anderson, K.S., Fraser, S.T. and Perkins, A.C., 2013. Generation of mice deficient in both KLF3/BKLF and KLF8 reveals a genetic interaction and a role for these factors in embryonic globin gene silencing. *Molecular and cellular biology*, *33*(15), pp.2976-2987.

Funnell, A.P., Maloney, C.A., Thompson, L.J., Keys, J., Tallack, M., Perkins, A.C. and Crossley, M., 2007. Erythroid Krüppel-like factor directly activates the basic Krüppel-like factor gene in erythroid cells. *Molecular and cellular biology*, *27*(7), pp.2777-2790.

Funnell, A.P., Norton, L.J., Mak, K.S., Burdach, J., Artuz, C.M., Twine, N.A., Wilkins, M.R., Power, C.A., Hung, T.T., Perdomo, J. and Koh, P., 2012. The CACCC-binding protein KLF3/BKLF represses a subset of KLF1/EKLF target genes and is required for proper erythroid maturation in vivo. *Molecular and cellular biology*, *32*(16), pp.3281-3292.

Funnell, A.P., Prontera, P., Ottaviani, V., Piccione, M., Giambona, A., Maggio, A., Ciaffoni, F., Stehling-Sun, S., Marra, M., Masiello, F. and Varricchio, L., 2015. 2p15-p16. 1 microdeletions encompassing and proximal to BCL11A are associated with elevated HbF in addition to neurologic impairment. *Blood*, *126*(1), pp.89-93.

Gaensler, K.M., Kitamura, M. and Kan, Y.W., 1993. Germ-line transmission and developmental regulation of a 150-kb yeast artificial chromosome containing the human beta-globin locus in transgenic mice. *Proceedings of the National Academy of Sciences*, *90*(23), pp.11381-11385.

Galanello, R. and Cao, A., 1998. Relationship between Genotype and Phenotype: Thalassemia Intermediaa. *Annals of the New York Academy of Sciences*, *850*(1), pp.325-333.

Garner, C., Tatu, T., Reittie, J.E., Littlewood, T., Darley, J., Cervino, S., Farrall, M., Kelly, P., Spector, T.D. and Thein, S.L., 2000. Genetic influences on F cells and other hematologic variables: a twin heritability study. *Blood*, *95*(1), pp.342-346.

Gazouli, M., Katsantoni, E., Kosteas, T. and Anagnou, N.P., 2009. Persistent fetal gamma-globin expression in adult transgenic mice following deletion of two silencer elements located 3′ to the human Agamma-globin gene. *Mol Med*, *15*(11-12), pp.415-424.

Ginder, G.D., 2015. Epigenetic regulation of fetal globin gene expression in adult erythroid cells. *Translational Research*, *165*(1), pp.115-125.

Ginder, G.D., Gnanapragasam, M.N. and Mian, O.Y., 2008. The role of the epigenetic signal, DNA methylation, in gene regulation during erythroid development. *Current topics in developmental biology*, *82*, pp.85-116.

Giarratana, M.C., Rouard, H., Dumont, A., Kiger, L., Safeukui, I., Le Pennec, P.Y., François, S., Trugnan, G., Peyrard, T., Marie, T. and Jolly, S., 2011. Proof of principle for transfusion of in vitro–generated red blood cells. *Blood*, *118*(19), pp.5071-5079.

Gnanapragasam, M.N., Scarsdale, J.N., Amaya, M.L., Webb, H.D., Desai, M.A., Walavalkar, N.M., Wang, S.Z., Zu Zhu, S., Ginder, G.D. and Williams, D.C., 2011. p66α–MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2–NuRD complex. *Proceedings of the National Academy of Sciences*, *108*(18), pp.7487-7492.

Grosveld, F., van Assendelft, G.B., Greaves, D.R. and Kollias, G., 1987. Position-independent, high-level expression of the human β-globin gene in transgenic mice. *Cell*, *51*(6), pp.975-985.

Gulliver, G., 1875. *Observations on the Sizes and Shapes of the Red Corpuscles of the Blood of Vertebrates...*

Gutiérrez, L., Tsukamoto, S., Suzuki, M., Yamamoto-Mukai, H., Yamamoto, M., Philipsen, S. and Ohneda, K., 2008. Ablation of Gata1 in adult mice results in aplastic crisis, revealing its essential role in steady-state and stress erythropoiesis. *Blood*, *111*(8), pp.4375-4385.

Hankins, J. and Aygun, B., 2009. Pharmacotherapy in sickle cell disease–state of the art and future prospects. *British journal of haematology*, *145*(3), pp.296-308.

Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N. and Grosveld, F., 1991. Importance of globin gene order for correct developmental expression. *Genes & Development*, *5*(8), pp.1387-1394.

Hartzog, G.A. and Myers, R.M., 1993. Discrimination among potential activators of the beta-globin CACCC element by correlation of binding and transcriptional properties. *Molecular and Cellular Biology*, *13*(1), pp.44-56.

Helias, V., Saison, C., Peyrard, T., Vera, E., Prehu, C., Cartron, J.P. and Arnaud, L., 2013. Molecular analysis of the rare in (Lu) blood type: toward decoding the phenotypic outcome of haploinsufficiency for the transcription factor KLF1. *Human mutation*, *34*(1), pp.221-228.

Heruth, D.P., Hawkins, T., Logsdon, D.P., Gibson, M.I., Sokolovsky, I.V., Nsumu, N.N., Major, S.L., Fegley, B., Woods, G.M., Lewing, K.B. and Neville, K.A., 2010. Mutation in erythroid specific transcription factor KLF1 causes Hereditary Spherocytosis in the Nan hemolytic anemia mouse model. *Genomics*, *96*(5), pp.303-307.

Hodge, D., Coghill, E., Keys, J., Maguire, T., Hartmann, B., McDowall, A., Weiss, M., Grimmond, S. and Perkins, A., 2006. A global role for EKLF in definitive and primitive erythropoiesis. *Blood*, *107*(8), pp.3359-3370.

Hossain, M.A., Shen, Y., Knudson, I., Thakur, S., Stees, J.R., Qiu, Y., Pace, B.S., Peterson, K.R. and Bungert, J., 2016. Activation of Fetal γ-globin Gene Expression via Direct Protein Delivery of Synthetic Zinc-finger DNA-Binding Domains. *Molecular Therapy—Nucleic Acids*, *5*(10), p.e378.

Huang, J., Zhang, X., Liu, D., Wei, X., Shang, X., Xiong, F., Yu, L., Yin, X. and Xu, X., 2015. Compound heterozygosity for KLF1 mutations is associated with microcytic hypochromic anemia and increased fetal hemoglobin. *European Journal of Human Genetics*, *23*(10), pp.1341-1348.

Huang, S., Gillinder, K.R., Sorolla, A., Whitelaw, E. and Perkins, A.C., 2016. Mutations in the second linker of KLF1 cause congenital non-spherocytic hemolytic anemia Due to global reduction of in vivo DNA-binding affinity.

Huisman, T.H.J., 1997. Levels of Hb A2 in heterozygotes and homozygotes for beta-thalassemia mutations: influence of mutations in the CACCC and ATAAA motifs of the beta-globin gene promoter. *Acta haematologica*, *98*(4), pp.187-194.

Huisman, T.H.J., Schroeder, W.A., Efremov, G.D., Duma, H., Mladenovski, B., Hyman, C.B., Rachmilewitz, E.A., Bouver, N., Miller, A., Brodie, A. and Shelton, J.R., 1974. The present status of the heterogeneity of fetal hemoglobin in β-thalassemia: an attempt to unify some observations in thalassemia and related conditions. *Annals of the New York Academy of Sciences*, *232*(1), pp.107-124.

Isern J, Fraser ST, He Z, Zhang H, Baron MH. 2010. Dose-dependent regulation of primitive erythroid maturation and identity by the transcription factor eklf. Blood 116(19):3972-80.

Isern, J., He, Z., Fraser, S.T., Nowotschin, S., Ferrer-Vaquer, A., Moore, R., Hadjantonakis, A.K., Schulz, V., Tuck, D., Gallagher, P.G. and Baron, M.H., 2011. Single-lineage transcriptome analysis reveals key regulatory pathways in primitive erythroid progenitors in the mouse embryo. *Blood*, *117*(18), pp.4924-4934.

Jaffray, J.A., Mitchell, W.B., Gnanapragasam, M.N., Seshan, S.V., Guo, X., Westhoff, C.M., Bieker, J.J. and Manwani, D., 2013. Erythroid transcription factor EKLF/KLF1 mutation causing congenital dyserythropoietic anemia type IV in a patient of Taiwanese origin: review of all reported cases and development of a clinical diagnostic paradigm. *Blood Cells, Molecules, and Diseases*, *51*(2), pp.71-75.

Ji RP, Phoon CK, Aristizabal O, McGrath KE, Palis J, Turnbull DH. 2003. Onset of cardiac function during early mouse embryogenesis coincides with entry of primitive erythroblasts into the embryo proper. Circ Res 92:133-135.

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*, *108*(3), pp.1077-1083.

Jimenez, G., Gale, K.B. and Enver, T., 1992. The mouse β-globin locus control region: hypersensitive sites 3 and 4. *Nucleic acids research*, *20*(21), pp.5797-5803.

John A, Brylka H, Wiegreffe C, et al. Bcl11a is required for neuronal morphogenesis and sensory circuit formation in dorsal spinal cord development. *Development.* 2012;139(10):1831-1841.

Kadam, S., McAlpine, G.S., Phelan, M.L., Kingston, R.E., Jones, K.A. and Emerson, B.M., 2000. Functional selectivity of recombinant mammalian SWI/SNF subunits. *Genes & development*, *14*(19), pp.2441-2451.

Karlsson, S. and Nienhuis, A.W., 1985. Developmental regulation of human globin genes. *Annual review of biochemistry*, *54*(1), pp.1071-1108.

Kassouf, M.T., Hughes, J.R., Taylor, S., McGowan, S.J., Soneji, S., Green, A.L., Vyas, P. and Porcher, C., 2010. Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells. *Genome research*, *20*(8), pp.1064-1083.

Kaur, G. and Dufour, J.M., 2012. Cell lines: Valuable tools or useless artifacts.

Kazazian, HH Jr, and Cornine D. Boehm. "Molecular basis and prenatal diagnosis of beta-thalassemia." *Blood* 72.4 (1988): 1107-1116.

Kennedy, M., D'Souza, S.L., Lynch-Kattman, M., Schwantz, S. and Keller, G., 2007. Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. *Blood*, *109*(7), pp.2679-2687.

Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T. and Kingston, R., 1999. Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity*, *10*(3), pp.345-355.

Kingsley, P.D., Malik, J., Emerson, R.L., Bushnell, T.P., McGrath, K.E., Bloedorn, L.A., Bulger, M. and Palis, J., 2006. "Maturational" globin switching in primary primitive erythroid cells. *Blood*, *107*(4), pp.1665-1672.

Kingsley, P.D., Malik, J., Fantauzzo, K.A. and Palis, J., 2004. Yolk sac–derived primitive erythroblasts enucleate during mammalian embryogenesis. *Blood*, *104*(1), pp.19-25.

Kransdorf, E.P., Wang, S.Z., Zu Zhu, S., Langston, T.B., Rupon, J.W. and Ginder, G.D., 2006. MBD2 is a critical component of a methyl cytosine-binding protein complex isolated from primary erythroid cells. *Blood*, *108*(8), pp.2836-2845.

Kumaravelu, P., Hook, L., Morrison, A.M., Ure, J., Zhao, S., Zuyev, S., Ansell, J. and Medvinsky, A., 2002. Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. *Development*, *129*(21), pp.4891-4899.

Kuo TY, Chen CY, Hsueh YP. Bcl11A/CTIP1 mediates the effect of the glutamate receptor on axon branching and dendrite outgrowth. *J Neurochem.* 2010;114(5):1381-1392.

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*(3), p.e59890.

Lahortiga, I., De Keersmaecker, K., Van Vlierberghe, P., Graux, C., Cauwelier, B., Lambert, F., Mentens, N., Beverloo, H.B., Pieters, R., Speleman, F. and Odero, M.D., 2007. Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. *Nature genetics*, *39*(5), pp.593-595.

Larsen, D.H., Poinsignon, C., Gudjonsson, T., Dinant, C., Payne, M.R., Hari, F.J., Danielsen, J.M.R., Menard, P., Sand, J.C., Stucki, M. and Lukas, C., 2010. The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage. *The Journal of cell biology*, *190*(5), pp.731-740.

Lee HY, Johnson KD, Boyer ME, Bresnick EH. Relocalizing genetic loci into specific subnuclear neighborhoods. Journal of Biological Chemistry. 2011 May 27;286(21):18834-44.

Lee, J.S., Ngo, H., Kim, D. and Chung, J.H., 2000. Erythroid Krüppel-like factor is recruited to the CACCC box in the β-globin promoter but not to the CACCC box in the γ-globin promoter: The role of the neighboring promoter elements. *Proceedings of the National Academy of Sciences*, *97*(6), pp.2468-2473.

Lee, L.K., Ueno, M., Van Handel, B. and Mikkola, H.K., 2010. Placenta as a newly identified source of hematopoietic stem cells. *Current opinion in hematology*, *17*(4), p.313.

Leid M, Ishmael JE, Avram D, Shepherd D, Fraulob V, Dolle P. CTIP1 and CTIP2 are differentially expressed during mouse embryogenesis.*Gene Expr Patterns.* 2004;4(6):733-739.

Leonard, M., Brice, M., Engel, J.D. and Papayannopoulou, T., 1993. Dynamics of GATA transcription factor expression during erythroid differentiation. *Blood*, *82*(4), pp.1071-1079.

Liu, D., Zhang, X., Yu, L., Cai, R., Ma, X., Zheng, C., Zhou, Y., Liu, Q., Wei, X., Lin, L. and Yan, T., 2014. Erythroid Krüppel-like factor mutations are relatively more common in a thalassemia endemic region and ameliorate the clinical and hematological severity of β-thalassemia. *Blood*, pp.blood-2014.

Lohmann, F. and Bieker, J.J., 2008. Activation of Eklf expression during hematopoiesis by Gata2 and Smad5 prior to erythroid commitment. *Development*, *135*(12), pp.2071-2082.

Lozzio, C.B. and Lozzio, B.B., 1975. Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood*, *45*(3), pp.321-334.

Lyon, M.F. and Glenister, P.G., 1986. Position of neonatal anemia (Nan) on chromosome 8. *Mouse News Lett*, *74*, p.95.

Maeda, T., Hobbs, R.M., Merghoub, T., Guernah, I., Zelent, A., Cordon-Cardo, C., Teruya-Feldstein, J. and Pandolfi, P.P., 2005. Role of the proto-oncogene Pokemon in cellular transformation and ARF repression. *Nature*, *433*(7023), pp.278-285.

Maeda, T., Ito, K., Merghoub, T., Poliseno, L., Hobbs, R.M., Wang, G., Dong, L., Maeda, M., Dore, L.C., Zelent, A. and Luzzatto, L., 2009. LRF is an essential downstream target of GATA1

in erythroid development and regulates BIM-dependent apoptosis. *Developmental cell*, *17*(4), pp.527-540.

Magor, G.W., Tallack, M.R., Gillinder, K.R., Bell, C.C., McCallum, N., Williams, B. and Perkins, A.C., 2015. KLF1 null neonates display hydrops fetalis and a deranged erythroid transcriptome. *Blood*, pp.blood-2014.

Martin, D.I. and Orkin, S.H., 1990. Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1. *Genes & development*, *4*(11), pp.1886-1898.

Martin, P. and Papayannopoulou, T., 1982. HEL cells: a new human erythroleukemia cell line with spontaneous and induced globin expression. *Science*, *216*(4551), pp.1233-1235.

Mas, C., Lussier-Price, M., Soni, S., Morse, T., Arseneault, G., Di Lello, P., Lafrance-Vanasse, J., Bieker, J.J. and Omichinski, J.G., 2011. Structural and functional characterization of an atypical activation domain in erythroid Krüppel-like factor (EKLF). *Proceedings of the National Academy of Sciences*, *108*(26), pp.10484-10489.

Masuda, T., Wang, X., Maeda, M., Canver, M.C., Sher, F., Funnell, A.P., Fisher, C., Suciu, M., Martyn, G.E., Norton, L.J. and Zhu, C., 2016. Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science*, *351*(6270), pp.285-289.

Mazouzi, A., Velimezi, G. and Loizou, J.I., 2014. DNA replication stress: causes, resolution and disease. *Experimental cell research*, *329*(1), pp.85-93.

McConnell, B.B. and Yang, V.W., 2010. Mammalian Krüppel-like factors in health and diseases. *Physiological reviews*, *90*(4), pp.1337-1381.

McConnell, S.C., Huo, Y., Liu, S. and Ryan, T.M., 2011. Human globin knock-in mice complete fetal-to-adult hemoglobin switching in postnatal development. *Molecular and cellular biology*, *31*(4), pp.876-883.

McGrath, K.E. and Palis, J., 2005. Hematopoiesis in the yolk sac: more than meets the eye. *Experimental hematology*, *33*(9), pp.1021-1028.

Migliaccio, G., Di Pietro, R., Di Giacomo, V., Di Baldassarre, A., Migliaccio, A.R., Maccioni, L., Galanello, R. and Papayannopoulou, T., 2002. In vitro mass production of human erythroid cells from the blood of normal donors and of thalassemic patients. *Blood Cells, Molecules, and Diseases*, *28*(2), pp.169-180.

Mikkola, H.K., Klintman, J., Yang, H., Hock, H., Schlaeger, T.M., Fujiwara, Y. and Orkin, S.H., 2003. Haematopoietic stem cells retain long-term repopulating activity and multipotency in the absence of stem-cell leukaemia SCL/tal-1 gene. *Nature*, *421*(6922), pp.547-551.

Miller, I.J. and Bieker, J.J., 1993. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear proteins. *Molecular and cellular biology*, *13*(5), pp.2776-2786.

Minie, M., Clark, D., Trainor, C., Evans, T., Reitman, M., Hannon, R., Gould, H. and Felsenfeld, G., 1992. Developmental regulation of globin gene expression. *J Cell Sci*, *1992*(Supplement 16), pp.15-20.

Mouthon, M.A., Bernard, O., Mitjavila, M.T., Romeo, P.H., Vainchenker, W. and Mathieu-Mahul, D., 1993. Expression of tal-1 and GATA-binding proteins during human hematopoiesis. *Blood*, *81*(3), pp.647-655.

Münger, K. and Howley, P.M., 2002. Human papillomavirus immortalization and transformation functions. *Virus research*, *89*(2), pp.213-228.

Mucenski, M.L., McLain, K., Kier, A.B., Swerdlow, S.H., Schreiner, C.M., Miller, T.A., Pietryga, D.W., Scott, W.J. and Potter, S.S., 1991. A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell*, *65*(4), pp.677-689.

Münger K, Howley PM. Human papillomavirus immortalization and transformation functions. Virus Res. 2002 Nov;89(2):213-28. Review. PubMed PMID: 12445661.

Myers, R.M., Tilly, K. and Maniatis, T., 1986. Fine structure genetic analysis of a beta-globin promoter. *Science*, *232*, pp.613-619.

Nan, X. and Bird, A., 2001. The biological functions of the methyl-CpG-binding protein MeCP2 and its implication in Rett syndrome. *Brain and Development*, *23*, pp.S32-S37.

Nicholson, T.B. and Chen, T., 2009. LSD1 demethylates histone and non-histone proteins. *Epigenetics*, *4*(3), pp.129-132.

Nilson, D.G., Sabatino, D.E., Bodine, D.M. and Gallagher, P.G., 2006. Major erythrocyte membrane protein genes in EKLF-deficient mice. *Experimental hematology*, *34*(6), pp.705-712.

Noordermeer, D. and de Laat, W., 2008. Joining the loops: β-Globin gene regulation. *IUBMB life*, *60*(12), pp.824-833.

Norton, L.J., Funnell, A.P., Burdach, J., Wienert, B., Kurita, R., Nakamura, Y., Philipsen, S., Pearson, R.C., Quinlan, K.G. and Crossley, M., 2017. KLF1 directly activates expression of the novel fetal globin repressor ZBTB7A/LRF in erythroid cells. *Blood Advances*, *1*(11), pp.685-692.

Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. and Grosveld, F., 1995. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature: international weekly journal of science*, *375*(6529), pp.316-318.

Orkin, S.H. and Zon, L.I., 2008. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell*, *132*(4), pp.631-644.

Palis, J., 2008. Ontogeny of erythropoiesis. *Current opinion in hematology*, *15*(3), pp.155-161.

Palis, J., Malik, J., Mcgrath, K.E. and Kingsley, P.D., 2010. Primitive erythropoiesis in the mammalian embryo. *International Journal of Developmental Biology*, *54*(6-7), pp.1011-1018.

Palis, J., Robertson, S., Kennedy, M., Wall, C. and Keller, G., 1999. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development*, *126*(22), pp.5073-5084.

Palladini, G., Milani, P., Foli, A., Basset, M., Russo, F., Perlini, S. and Merlini, G., 2017. A phase 2 trial of pomalidomide and dexamethasone rescue treatment in patients with AL amyloidosis. *Blood*, *129*(15), pp.2120-2123.

Pang, C.J., Lemsaddek, W., Alhashem, Y.N., Bondzi, C., Redmond, L.C., Ah-Son, N., Dumur, C.I., Archer, K.J., Haar, J.L., Lloyd, J.A. and Trudel, M., 2012. Krüppel-like factor 1 (KLF1), KLF2, and myc control a regulatory network essential for embryonic erythropoiesis. *Molecular and cellular biology*, *32*(13), pp.2628-2644.

Papayannopoulou, T., Chen, P., Maniatis, A. and Stamatoyannopoulos, G., 1980. Simultaneous assessment of i-antigenic expression and fetal hemoglobin. *Blood*, *55*(2), p.221.

Perkins, A., 1999. Erythroid Kruppel like factor: from fishing expedition to gourmet meal. *The international journal of biochemistry & cell biology*, *31*(10), pp.1175-1192.

Perkins, A., Xu, X., Higgs, D.R., Patrinos, G.P., Arnaud, L., Bieker, J.J. and Philipsen, S., 2016. Krüppeling erythropoiesis: an unexpected broad spectrum of human red blood cell disorders due to KLF1 variants. *Blood*, *127*(15), pp.1856-1862.

Perkins, A.C., Gaensler, K.M. and Orkin, S.H., 1996. Silencing of human fetal globin expression is impaired in the absence of the adult beta-globin gene activator protein EKLF. *Proceedings of the National Academy of Sciences*, *93*(22), pp.12267-12271.

Perkins, A.C., Sharpe, A.H. and Orkin, S.H., 1995. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature*, *375*(6529), p.318.

Perseu, L., Satta, S., Moi, P., Demartis, F.R., Manunza, L., Sollaino, M.C., Barella, S., Cao, A. and Galanello, R., 2011. KLF1 gene mutations cause borderline HbA2. *Blood*, *118*(16), pp.4454-4458.

Peters, L.L., Lane, P.W., Andersen, S.G., Gwynn, B., Barker, J.E. and Beutler, E., 2001. Downeast anemia (dea), a new mouse model of severe nonspherocytic hemolytic anemia caused by hexokinase (HK I) deficiency. *Blood Cells, Molecules, and Diseases*, *27*(5), pp.850-860.

Peterson, K.R., Clegg, C.H., Huxley, C., Josephson, B.M., Haugen, H.S., Furukawa, T. and Stamatoyannopoulos, G., 1993. Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human beta-globin locus display proper developmental control of human globin genes. *Proceedings of the National Academy of Sciences*, *90*(16), pp.7593-7597.

Pevny, L. and Simon, M.C., 1991. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*, *349*(6306), p.257.

Philipsen, S., Molecular control of hemoglobin switching.

Piel, F.B., Patil, A.P., Howes, R.E., Nyangiri, O.A., Gething, P.W., Dewi, M., Temperley, W.H., Williams, T.N., Weatherall, D.J. and Hay, S.I., 2013. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *The Lancet*, *381*(9861), pp.142-151.

Pilon, A.M., Ajay, S.S., Kumar, S.A., Steiner, L.A., Cherukuri, P.F., Wincovitch, S., Anderson, S.M., Center, N.C.S., Mullikin, J.C., Gallagher, P.G. and Hardison, R.C., 2011. Genome-wide ChIP-Seq reveals a dramatic shift in the binding of the transcription factor erythroid Kruppel-like factor during erythrocyte differentiation. *Blood*, *118*(17), pp.e139-e148.

Pilon, A.M., Arcasoy, M.O., Dressman, H.K., Vayda, S.E., Maksimova, Y.D., Sangerman, J.I., Gallagher, P.G. and Bodine, D.M., 2008. Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. *Molecular and cellular biology*, *28*(24), pp.7394-7401.

Pishesha, N., Thiru, P., Shi, J., Eng, J.C., Sankaran, V.G. and Lodish, H.F., 2014. Transcriptional divergence and conservation of human and mouse erythropoiesis. *Proceedings of the National Academy of Sciences*, *111*(11), pp.4103-4108.

Planutis, A., Xue, L., Trainor, C.D., Dangeti, M., Gillinder, K., Siatecka, M., Nebor, D., Peters, L.L., Perkins, A.C. and Bieker, J.J., 2017. Neomorphic effects of the neonatal anemia (Nan-Eklf) mutation contribute to deficits throughout development. *Development*, *144*(3), pp.430-440.

Platt, O.S., Brambilla, D.J., Rosse, W.F., Milner, P.F., Castro, O., Steinberg, M.H. and Klug, P.P., 1994. Mortality in sickle cell disease--life expectancy and risk factors for early death. *New England Journal of Medicine*, *330*(23), pp.1639-1644.

Popp, R.A., Lalley, P.A., Whitney, J.B. and Anderson, W.F., 1981. Mouse alpha-globin genes and alpha-globin-like pseudogenes are not syntenic. *Proceedings of the National Academy of Sciences*, *78*(10), pp.6362-6366.

Qiu, C., Olivier, E.N., Velho, M. and Bouhassira, E.E., 2008. Globin switches in yolk sac–like primitive and fetal-like definitive red blood cells produced from human embryonic stem cells. *Blood*, *111*(4), pp.2400-2408.

Quadrini, K.J., Gruzglin, E. and Bieker, J.J., 2008. Non-random subcellular distribution of variant EKLF in erythroid cells. *Experimental cell research*, *314*(7), pp.1595-1604.

Raiola, L., Lussier-Price, M., Gagnon, D., Lafrance-Vanasse, J., Mascle, X., Arseneault, G., Legault, P., Archambault, J. and Omichinski, J.G., 2013. Structural characterization of a noncovalent complex between ubiquitin and the transactivation domain of the erythroid-specific factor EKLF. *Structure*, *21*(11), pp.2014-2024.

Ramsay, R.G. and Gonda, T.J., 2008. MYB function in normal and cancer cells. *Nature Reviews Cancer*, *8*(7), pp.523-534.

Rivers, A., Vaitkus, K., Ruiz, M.A., Ibanez, V., Jagadeeswaran, R., Kouznetsova, T., DeSimone, J. and Lavelle, D., 2015. RN-1, a potent and selective lysine-specific demethylase 1 inhibitor, increases γ-globin expression, F reticulocytes, and F cells in a sickle cell disease mouse model. *Experimental hematology*, *43*(7), pp.546-553.

Robb, L., Lyons, I., Li, R., Hartley, L., Köntgen, F., Harvey, R.P., Metcalf, D. and Begley, C.G., 1995. Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proceedings of the National Academy of Sciences*, *92*(15), pp.7075-7079.

Romeo, P.H., Prandini, M.H., Joulin, V., Mignotte, V. and Prenant, M., 1990. Megakaryocytic and erythrocytic lineages share specific transcription factors. *Nature*, *344*(6265), p.447.

Rupon, J.W., Wang, S.Z., Gaensler, K., Lloyd, J. and Ginder, G.D., 2006. Methyl binding domain protein 2 mediates γ-globin gene silencing in adult human βYAC transgenic mice. *Proceedings of the National Academy of Sciences*, *103*(17), pp.6617-6622.

Rupon, J.W., Wang, S.Z., Gnanapragasam, M., Labropoulos, S. and Ginder, G.D., 2011. MBD2 contributes to developmental silencing of the human ε-globin gene. *Blood Cells, Molecules, and Diseases*, *46*(3), pp.212-219.

Sabin, F.R., 1920. Studies on the origin of blood vessels and of red corpuscles as seen in the living blastoderm of the chick during the second day of incubation. *Contributions to embryology*, *9*, pp.213-262.

Sankaran, V.G., 2009. *Molecular mechanisms of erythropoiesis and globin gene regulation*. Harvard University.

Sankaran, V.G., Ghazvinian, R., Do, R., Thiru, P., Vergilio, J.A., Beggs, A.H., Sieff, C.A., Orkin, S.H., Nathan, D.G., Lander, E.S. and Gazda, H.T., 2012. Exome sequencing identifies GATA1 mutations resulting in Diamond-Blackfan anemia. *The Journal of clinical investigation*, *122*(7), pp.2439-2443.

Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K., Hirschhorn, J.N., Cantor, A.B. and Orkin, S.H., 2008. Human fetal hemoglobin expression is

regulated by the developmental stage-specific repressor BCL11A. *Science*, *322*(5909), pp.1839-1842.

Sankaran, V.G. and Nathan, D.G., 2010. Thalassemia: an overview of 50 years of clinical research. *Hematology/oncology clinics of North America*, *24*(6), pp.1005-1020.

Sankaran, V.G. and Orkin, S.H., 2013. The switch from fetal to adult hemoglobin. *Cold Spring Harbor perspectives in medicine*, *3*(1), p.a011643.

Sankaran, V.G., Xu, J. and Orkin, S.H., 2010. Advances in the understanding of haemoglobin switching. *British journal of haematology*, *149*(2), pp.181-194.

Sargent, T.G. and Lloyd, J.A., 2001. The human γ-globin TATA and CACCC elements have key, distinct roles in suppressing β-globin gene expression in embryonic/fetal development. *Journal of Biological Chemistry*, *276*(45), pp.41817-41824.

Sargent, T.G., DuBois, C.C., Buller, A.M. and Lloyd, J.A., 1999. The roles of 5′-HS2, 5′-HS3, and the γ-globin TATA, CACCC, and stage selector elements in suppression of β-globin expression in early development. *Journal of Biological Chemistry*, *274*(16), pp.11229-11236.

Satta, S., Perseu, L., Moi, P., Asunis, I., Cabriolu, A., Maccioni, L., Demartis, F.R., Manunza, L., Cao, A. and Galanello, R., 2011. Compound heterozygosity for KLF1 mutations associated with remarkable increase of fetal hemoglobin and red cell protoporphyrin. *haematologica*, *96*(5), pp.767-770.

Satta, S., Perseu, L., Maccioni, L., Giagu, N. and Galanello, R., 2012. Delayed fetal hemoglobin switching in subjects with KLF1 gene mutation. *Blood Cells, Molecules, and Diseases*, *48*(1), pp.22-24.

Satterwhite, E., Sonoki, T., Willis, T.G., Harder, L., Nowak, R., Arriola, E.L., Liu, H., Price, H.P., Gesk, S., Steinemann, D. and Schlegelberger, B., 2001. The BCL11 gene family: involvement of BCL11A in lymphoid malignancies. *Blood*, *98*(12), pp.3413-3420.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S. and Eskiw, C.H., 2010. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nature genetics*, *42*(1), pp.53-61.

Shi, L., Cui, S., Engel, J.D. and Tanabe, O., 2013. Lysine-specific demethylase 1 is a therapeutic target for fetal hemoglobin induction. *Nature medicine*, *19*(3), pp.291-294.

Shivdasani, R.A. and Orkin, S.H., 1996. The transcriptional control of hematopoiesis [see comments]. *Blood*, *87*(10), pp.4025-4039.

Shivdasani, R.A., Mayer, E.L. and Orkin, S.H., 1995. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature*, *373*(6513), p.432.

Siatecka, M. and Bieker, J.J., 2011. The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood, 118*(8), pp.2044-2054.

Siatecka, M., Sahr, K.E., Andersen, S.G., Mezei, M., Bieker, J.J. and Peters, L.L., 2010. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Krüppel-like factor. *Proceedings of the National Academy of Sciences, 107*(34), pp.15151-15156.

Siegel, W., Cox, R., Schroeder, W., Huisman, T.H.J., Penner, O. and Rowley, P.T., 1970. An adult homozygous for persistent fetal hemoglobin. *Annals of internal medicine, 72*(4), pp.533-536.

Singleton, B.K., Burton, N.M., Green, C., Brady, R.L. and Anstee, D.J., 2008. Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In (Lu) phenotype. *Blood, 112*(5), pp.2081-2088.

Singleton, B.K., Fairweather, V.S., Lau, W., Parsons, S.F., Burton, N.M., Frayne, J., Brady, R.L. and Anstee, D.J., 2009. A novel EKLF mutation in a patient with Dyserythropoietic anemia: the first association of EKLF with disease in man. *Blood, 114*(22), pp.162-162.

Smith, E.C. and Orkin, S.H., 2016. Hemoglobin genetics: recent contributions of GWAS and gene editing. *Human Molecular Genetics*, p.ddw170.

Soni, S., Pchelintsev, N., Adams, P.D. and Bieker, J.J., 2014. Transcription factor EKLF (KLF1) recruitment of the histone chaperone HIRA is essential for β-globin gene expression. *Proceedings of the National Academy of Sciences, 111*(37), pp.13337-13342.

Southwood, C.M., Downs, K.M. and Bieker, J.J., 1996. Erythroid Krüppel-like factor exhibits an early and sequentially localized pattern of expression during mammalian erythroid ontogeny. *Developmental dynamics, 206*(3), pp.248-259.

Spinner, M.A., Sanchez, L.A., Hsu, A.P., Shaw, P.A., Zerbe, C.S., Calvo, K.R., Arthur, D.C., Gu, W., Gould, C.M., Brewer, C.C. and Cowen, E.W., 2014. GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics, and immunity. *Blood, 123*(6), pp.809-821.

Stamatoyannopoulos, G., 2005. Control of globin gene expression during development and erythroid differentiation. *Experimental hematology, 33*(3), pp.259-271.

Strouboulis, J., Dillon, N. and Grosveld, F., 1992. Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. *Genes & Development, 6*(10), pp.1857-1864.

Sue, N., Jack, B.H., Eaton, S.A., Pearson, R.C., Funnell, A.P., Turner, J., Czolij, R., Denyer, G., Bao, S., Molero-Navajas, J.C. and Perkins, A., 2008. Targeted disruption of the basic Krüppel-like factor gene (Klf3) reveals a role in adipogenesis. *Molecular and cellular biology, 28*(12), pp.3967-3978.

Tallack, M.R., Keys, J.R., Humbert, P.O. and Perkins, A.C., 2009. EKLF/KLF1 controls cell cycle entry via direct regulation of E2f2. *Journal of Biological Chemistry*, *284*(31), pp.20966-20974.

Tallack, M.R., Keys, J.R. and Perkins, A.C., 2007. Erythroid Kruppel-like Factor Regulates the G1 Cyclin Dependent Kinase Inhibitor p18 INK4c. *Journal of molecular biology*, *369*(2), pp.313-321.

Tallack, M.R., Magor, G.W., Dartigues, B., Sun, L., Huang, S., Fittock, J.M., Fry, S.V., Glazov, E.A., Bailey, T.L. and Perkins, A.C., 2012. Novel roles for KLF1 in erythropoiesis revealed by mRNA-seq. *Genome research*, *22*(12), pp.2385-2398.

Tallack, M.R. and Perkins, A.C., 2013. Three fingers on the switch: Krüppel-like factor 1 regulation of γ-globin to β-globin gene switching. *Current opinion in hematology*, *20*(3), pp.193-200.

Tallack, M.R., Whitington, T., Yuen, W.S., Wainwright, E.N., Keys, J.R., Gardiner, B.B., Nourbakhsh, E., Cloonan, N., Grimmond, S.M., Bailey, T.L. and Perkins, A.C., 2010. A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells. *Genome research*, *20*(8), pp.1052-1063.

Tanabe, O., Katsuoka, F., Campbell, A.D., Song, W., Yamamoto, M., Tanimoto, K. and Engel, J.D., 2002. An embryonic/fetal β-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. *The EMBO journal*, *21*(13), pp.3434-3442.

Tavian, M., Zheng, B., Oberlin, E., Crisan, M., Sun, B., Huard, J. and Peault, B., 2005. The vascular wall as a source of stem cells. *Annals of the New York Academy of Sciences*, *1044*(1), pp.41-50.

Tepakhan, W., Yamsri, S., Fucharoen, G., Sanchaisuriya, K. and Fucharoen, S., 2015. Krüppel-like factor 1 mutations and expression of hemoglobins F and A2 in homozygous hemoglobin E syndrome. *Annals of hematology*, *94*(7), pp.1093-1098.

Thein, S.L., Menzel, S., Lathrop, M. and Garner, C., 2009. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. *Human molecular genetics*, *18*(R2), pp.R216-R223.

Tober, J., Koniski, A., McGrath, K.E., Vemishetti, R., Emerson, R., de Mesy-Bentley, K.K., Waugh, R. and Palis, J., 2007. The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. *Blood*, *109*(4), pp.1433-1441.

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. *Molecular cell*, *10*(6), pp.1453-1465.

Traxler, E.A., Yao, Y., Wang, Y.D., Woodard, K.J., Kurita, R., Nakamura, Y., Hughes, J.R., Hardison, R.C., Blobel, G.A., Li, C. and Weiss, M.J., 2016. A genome-editing strategy to treat [beta]-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nature medicine*.

Traxler, E.A., 2016. *Insights into terminal erythropoiesis influenced by human genetic variation* (Doctoral dissertation, University of Pennsylvania).

Trimborn, T., Gribnau, J., Grosveld, F. and Fraser, P., 1999. Mechanisms of developmental control of transcription in the murine α-and β-globin loci. *Genes & Development*, *13*(1), pp.112-124.

Trompeter, S. and Roberts, I., 2009. Haemoglobin F modulation in childhood sickle cell disease. *British journal of haematology*, *144*(3), pp.308-316.

Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W. and Orkin, S.H., 1994. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, *371*(6494), p.221.

Tuan, D., Solomon, W., Li, Q. and London, I.M., 1985. The" beta-like-globin" gene domain in human erythroid cells. *Proceedings of the National Academy of Sciences*, *82*(19), pp.6384-6388.

Tufarelli, C., Hardison, R., Miller, W., Hughes, J., Clark, K., Ventress, N., Frischauf, A.M. and Higgs, D.R., 2004. Comparative analysis of the α-like globin clusters in mouse, rat, and human chromosomes indicates a mechanism underlying breaks in conserved synteny. *Genome research*, *14*(4), pp.623-630.

Turner, J. and Crossley, M., 1998. Cloning and characterization of mCtBP2, a co-repressor that associates with basic Krüppel-like factor and other mammalian transcriptional regulators. *The EMBO journal*, *17*(17), pp.5129-5140.

Turner, J. and Crossley, M., 1999. Mammalian Krüppel-like transcription factors: more than just a pretty finger. *Trends in biochemical sciences*, *24*(6), pp.236-240.

Turner, J., Nicholas, H., Bishop, D., Matthews, J.M. and Crossley, M., 2003. The LIM protein FHL3 binds basic Krüppel-like factor/Krüppel-like factor 3 and its co-repressor C-terminal-binding protein 2. *Journal of Biological Chemistry*, *278*(15), pp.12786-12795.

Ulirsch, J.C., Lacy, J.N., An, X., Mohandas, N., Mikkelsen, T.S. and Sankaran, V.G., 2014. Altered chromatin occupancy of master regulators underlies evolutionary divergence in the transcriptional landscape of erythroid differentiation. *PLoS genetics*, *10*(12), p.e1004890.

Valge-Archer, V.E., Osada, H., Warren, A.J., Forster, A., Li, J., Baer, R. and Rabbitts, T.H., 1994. The LIM protein RBTN2 and the basic helix-loop-helix protein TAL1 are present in a complex in erythroid cells. *Proceedings of the National Academy of Sciences*, *91*(18), pp.8617-8621.

Vinjamur, D.S., Wade, K.J., Mohamad, S.F., Haar, J.L., Sawyer, S.T. and Lloyd, J.A., 2014. Krüppel-like transcription factors KLF1 and KLF2 have unique and coordinate roles in regulating embryonic erythroid precursor maturation. *Haematologica*, *99*(10), pp.1565-1573.

Vinjamur, D.S., Alhashem, Y.N., Mohamad, S.F., Amin, P., Williams Jr, D.C. and Lloyd, J.A., 2016. Krüppel-Like Transcription Factor KLF1 Is Required for Optimal γ-and β-Globin Expression in Human Fetal Erythroblasts. *PloS one*, *11*(2), p.e0146802.

Vinjamur DS. 2014. The Roles of Krüppel-like Transcription Factors KLF1 and KLF2 in Mouse Embryonic and Human Fetal Erythropoiesis. Doctoral thesis. Virginia Commonwealth University.

Viprakasit, V., Ekwattanakit, S., Riolueang, S., Chalaow, N., Fisher, C., Lower, K., Kanno, H., Tachavanich, K., Bejrachandra, S., Saipin, J. and Juntharaniyom, M., 2014. Mutations in Krüppel-like factor 1 cause transfusion-dependent hemolytic anemia and persistence of embryonic globin gene expression. *Blood*, *123*(10), pp.1586-1595.

Visvader, J. and Adams, J.M., 1993. Megakaryocytic differentiation induced in 416B myeloid cells by GATA-2 and GATA-3 transgenes or 5-azacytidine is tightly coupled to GATA-1 expression. *Blood*, *82*(5), pp.1493-1501.

Wade, P.A., Gegonne, A., Jones, P.L., Ballestar, E., Aubry, F. and Wolffe, A.P., 1999. Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nature genetics*, *23*(1), pp.62-66.

Wadman, I., Li, J., Bash, R.O., Forster, A., Osada, H., Rabbitts, T.H. and Baer, R., 1994. Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. *The EMBO Journal*, *13*(20), p.4831.

Warren, A.J., Colledge, W.H., Carlton, M.B., Evans, M.J., Smith, A.J. and Rabbitts, T.H., 1994. The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. *Cell*, *78*(1), pp.45-57.

Weatherall, D.J., 2001. Phenotype—genotype relationships in monogenic disease: lessons from the thalassaemias. *Nature Reviews Genetics*, *2*(4), pp.245-255.

Weinert, 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*,

Wijgerde, M., Gribnau, J., Trimborn, T., Nuez, B., Philipsen, S., Grosveld, F. and Fraser, P., 1996. The role of EKLF in human beta-globin gene competition. *Genes & Development*, *10*(22), pp.2894-2902.

Wijgerde, M., Grosveld, F. and Fraser, P., 1995. Transcription complex stability and chromatin dynamics in vivo. *Nature*, *377*(6546), p.209.

Williams, C.J., Naito, T., Gómez-del Arco, P., Seavitt, J.R., Cashman, S.M., De Souza, B., Qi, X., Keables, P., Von Andrian, U.H. and Georgopoulos, K., 2004. The chromatin remodeler Mi-2β is required for CD4 expression and T cell development. *Immunity*, *20*(6), pp.719-733.

Wong, P.M., Chung, S.W., Chui, D.H. and Eaves, C.J., 1986. Properties of the earliest clonogenic hemopoietic precursors to appear in the developing murine yolk sac. *Proceedings of the National Academy of Sciences*, *83*(11), pp.3851-3854.

Wood, K.H. and Zhou, Z., 2016. Emerging molecular and biological functions of MBD2, a reader of DNA methylation. *Frontiers in Genetics*, *7*.

Wu W, et al. (2011) Dynamics of the epigenetic landscape during erythroid differentiation after GATA1 restoration. Genome Res 21(10):1659–1671.

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. *Proceedings of the National Academy of Sciences*, *110*(16), pp.6518-6523.

Xue, L., Galdass, M., Gnanapragasam, M.N., Manwani, D. and Bieker, J.J., 2014. Extrinsic and intrinsic control by EKLF (KLF1) within a specialized erythroid niche. *Development*, *141*(11), pp.2245-2254.

Xue, Y., Wong, J., Moreno, G.T., Young, M.K., Côté, J. and Wang, W., 1998. NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Molecular cell*, *2*(6), pp.851-861.

Ye, L., Wang, J., Tan, Y., Beyer, A.I., Xie, F., Muench, M.O. and Kan, Y.W., 2016. Genome editing using CRISPR-Cas9 to create the HPFH genotype in HSPCs: An approach for treating sickle cell disease and β-thalassemia. *Proceedings of the National Academy of Sciences*, p.201612075.

Yu, L.H., Liu, D., Cai, R., Shang, X., Zhang, X.H., Ma, X.X., Yan, S.H., Fang, P., Zheng, C.G., Wei, X.F. and Liu, Y.H., 2015. Changes in hematological parameters in α-thalassemia individuals co-inherited with erythroid Krüppel-like factor mutations. *Clinical genetics*, *88*(1), pp.56-61.

Zambidis, E.T., Peault, B., Park, T.S., Bunz, F. and Civin, C.I., 2005. Hematopoietic differentiation of human embryonic stem cells progresses through sequential hematoendothelial, primitive, and definitive stages resembling human yolk sac development. *Blood*, *106*(3), pp.860-870.

Zhang HB, Liu DP, Liang CC. 2002. The control of expression of the alpha-globin gene cluster. Int J Hematol 76:420-426.

Zhang J, Socolovsky M, Gross AW, Lodish HF. 2003. Role of ras signaling in erythroid differentiation of mouse fetal liver cells: Functional analysis by a flow cytometry-based novel culture system. Blood 102(12):3938-46.

Zhang, W. and Bieker, J.J., 1998. Acetylation and modulation of erythroid Krüppel-like factor (EKLF) activity by interaction with histone acetyltransferases. *Proceedings of the National Academy of Sciences*, *95*(17), pp.9855-9860.

Zhang, Y. and Li, Y., 2010. The expanding Mi-2/NuRD complexes: a schematic glance. *Proteomics Insights*, *3*, p.79.

Zheng, J., Kitajima, K., Sakai, E., Kimura, T., Minegishi, N., Yamamoto, M. and Nakano, T., 2006. Differential effects of GATA-1 on proliferation and differentiation of erythroid lineage cells. *Blood*, *107*(2), pp.520-527.

Zhou, D., Liu, K., Sun, C.W., Pawlik, K.M. and Townes, T.M., 2010. KLF1 regulates BCL11A expression and [gamma]-to [beta]-globin gene switching. *Nature genetics*, *42*(9), pp.742-744.

Zhou D, Pawlik KM, Ren J, Sun CW, Townes TM. 2006. Differential binding of erythroid krupple-like factor to embryonic/fetal globin gene promoters during development. J Biol Chem 281(23):16052-7.

Zon, L.I., Yamaguchi, Y., Yee, K., Albee, E.A., Kimura, A., Bennett, J.C., Orkin, S.H. and Ackerman, S.J., 1993. Expression of mRNA for the GATA-binding proteins in human eosinophils and basophils: potential role in gene transcription. *Blood*, *81*(12), pp.3234-3241.

Vita

Anna Philip Kovilakath was born on September 18th, 1993 in Fairfax, Virginia and is an American citizen. She received her Bachelor of Science in Biotechnology from the School of Life Sciences, Manipal University in 2015. She was intrigued by the field of Genetics and pursued the field in Richmond, Virginia in 2015. She received her Master of Science in Human and Molecular Genetics, under the guidance of Dr. Joyce Lloyd, from Virginia Commonwealth University in 2017.