THE UNIVERSITY OF CHICAGO

ERYTHROPOIESIS IS CONTROLLED BY EXTERNAL STIMULI THROUGH COVALENT EPIGENETIC MODIFICATIONS AND CHROMATIN LOOPING

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

COMMITTEE ON CANCER BIOLOGY

BY

ZHONGYANG CAO

CHICAGO, ILLINOIS

MARCH 2021

Dedication

To my parents, who inspired me to aim for excellence and work hard to reach it.

And to my extended families, who cared for and encouraged me throughout my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2-HG	2-hydroxyglutarate
2-OG	2-oxoglutarate
3C	Chromatin Conformation Capture
5-caC	5-carboxylcytosine
5-fC	5-formylcytosine
5-hmC	5-hydroxymethylcytosine
5-mC	5-methylcytosine
AhR	Aryl hydrocarbon receptor
AML	Acute myeloid leukemia
BER	Base-excision repair
BET	Bromodomain and extra-terminal protein
CMML	Chronic myelomonocytic leukemia
ChIP	Chromatin immunoprecipitation
DEF	Diethyl fumarate
DNMT	DNA methyl transferases
DMS	Dimethyl succinate
DSBH	Double-stranded β-helix
ES cell	Embryonic stem cell
EPO	Erythropoietin
EPOR	Erythropoietin receptor
FBS	Fetal bovine serum
FH	Fumarate hydratase
FPKM	Fragments per kilobase per million reads
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GO	Gene ontology
HDAC	Histone deacetylase
HIF	Hypoxia-inducible factor
hMe-SEAL	5-hmC chemical labeling and pull-down
HPFH	Hereditary persistence of fetal hemoglobin
HSPC	Hematopoietic stem and progenitor cells

HU	Hydroxyurea
IDH	Isocitrate dehydrogenase
IMF	Idiopathic myelofibrosis
JAK	Janus-activated kinase
JBP	J-binding proteins
KO	Knock-out
LCR	Locus control region
MDS	Myelodysplastic syndrome
MnSOD	Manganese-dependent superoxide demutase
MPN	myeloproliferative neoplasm
NEIL1	Nei-like DNA glycosylase 1
PHD	Prolyl hydroxylase
qPCR	Quantitative real-time PCR
$O_2^{\bullet-}$	Superoxide
PTM	Post-translational modification
RNA-seq	RNA-sequencing
ROS	Reactive oxygen species
SAM	S-adenosylmethionine
SCA	Sickle cell anemia
SCF	Stem cell factor
SDH	Succinate dehydrogenase
TCA	Tricarboxylic acid
TDG	Thymine-DNA glucosylase
TET	Ten-eleven translocation methylcytosine dioxygenase
TF	Transcription factor
VHL	Von Hippel-Lindau E3 ubiquitin ligase
WT	Wild-type

ACKNOWLEDGEMENT

I thank my advisor, Dr. Lucy Godley, for the guidance and support she provided throughout my PhD career. Lucy has been a constant inspiration for me to reach for excellence, and she never runs out of ideas to keep me moving forward. She trained me to think rigorously and methodically in scientific research, and for that I am grateful. I am also thankful for her support in my career choices and help in finding my postdoctoral position, which would not have been possible without her support.

I also thank my thesis committee members, Drs. Amittha Wickrema, Alex Ruthenburg, Jill de Jong, and Robert Grossman, for their time, advice, and critiques throughout my PhD career. Your advice and suggestions along the way guided me toward my goal and kept me away from distractions. I would like to thank Dr. Amittha Wickrema especially, who served as my committee Chair and collaborated extensively with me on my projects. His expertise on erythropoiesis as well as the valuable input he provided when I was designing my studies made my projects feasible.

I would like to thank all of the great people I worked with over the years, particularly the current and past members of the Godley and Wickrema laboratories, including: Dr. Jozef Madzo, Dr. Christopher Mariani, Dr. Sakshi Uppal, Dr. Aparna Vasanthakumar, Dr. Kristina Bigelow, Dr. Jeong Jong, Hui Liu, Janet Lepore, Art Wolin, Maya Lewinsohn, Matthew Pozsgai, Aelin Kim, Shawn Albert, Anastasia Hains, Stephen Arnovitz, Connie Phung, Dr. Simone Feurstein, and Dr. Michael Drazer. I am especially thankful for Dr. Jozef Madzo and Dr. Christopher Mariani, who introduced me to bioinformatics and taught me many techniques in both wet and dry lab. I have greatly enjoyed my time spent with Anastasia Hains, Stephen Arnovitz, Connie Phung, and

Maya Lewinsohn. We have become close friends through our shared experiences in and outside of lab, and I will always be grateful for their help and support.

I have made many friends during my PhD career, and I would like to thank them for their emotional and physical support. I especially thank my friends and classmates in my CCB cohort: Lindsey Ludwig, Ashley Sample, Maya Zafrir Springer, and Yuanyuan Li, as well as friends in other PhD programs: Anna Chen, Beth Pollard, and James Goodman. We have bonded through our shared experiences, and I would not have been as happy as I am without their support.

I cannot thank my parents enough for all they have done to get me to where I am. They taught me to work hard and strive for excellence by their own actions, and their moral, financial, and emotional support is the only reason I can achieve as much as I have now. They are the source of inspiration for me to pursue higher education and a PhD degree, and they will always serve as my role models in life.

ABSTRACT

In humans, physiological hypoxia is the main driver for erythropoiesis. Hypoxia is sensed by kidney interstitial fibroblasts, which then secrete erythropoietin (EPO) to stimulate erythropoiesis in the bone marrow. In recent years, epigenetic mechanisms controlling erythropoiesis have attracted increasing attention. Recent advances in the field elucidated how epigenetic regulation contributes to maintenance of normal erythropoiesis, and how deregulation of the epigenetic machinery leads to abnormal differentiation of hematopoietic stem and progenitor cells, leading to hematopoietic malignancies. In particular, the covalent cytosine modification 5-hydroxymethylcytosine (5-hmC) is crucial for normal erythropoiesis. The enzymes that convert 5-methylcytosine (5-mC) into 5-hmC are the Ten-eleven translocation methylcytosine dioxygenases (TETs), which use oxygen and 2-oxoglutarate (2-OG) as additional substrates for the reaction and Fe(II) as a cofactor. However, it was not known how a hypoxic environment like the bone marrow regulates the *TET* genes and their transcripts. In my doctoral work, I investigated the effects of hypoxia on 5-hmC levels and distribution during erythropoiesis. My results showed that hypoxia induces 5-hmC accumulation despite low oxygen availability, and part of this effect is mediated through the direct upregulation of *TET3* by HIF-1. In addition, I tested whether altering chromatin structure by inhibiting the bromodomain and extra-terminal (BET) proteins leads to re-expression of fetal and/or embryonic *β-globin* genes. My results showed that BET inhibition leads to downregulation of genes encoding suppressors of fetal or embryonic *β-globin* and disrupts the association of the enhancers with adult *β-globin* genes, resulting in re-activation of *embryonic globins*. This observation suggests that BET inhibitors could be

useful in treating β-globinopathies like sickle cell anemia. Overall, my results regarding the epigenetic regulation of erythropoiesis have revealed previously unknown molecular mechanisms and provide novel strategies for the treatment of blood disorders.

CHAPTER I

Introduction

The majority of this chapter is adapted from a review co-authored by me and Anastasia Hains under the supervision of Dr. Lucy Godley, published as a chapter in *The DNA, RNA, and Histone Methylomes*:

Cao JZ, Hains AE, Godley LA. Regulation of 5-Hydroxymethylcytosine Distribution by the TET Enzymes. In: Jurga SB, Jan, editor. *The DNA, RNA, and Histone Methylomes*: Springer; 2019. p. 229-63. ISBN 978-3-030-14792-1

In addition, figures ,tables, and texts in several sections of this chapter were adapted from published papers for which I served as co-authors:

(i) Data presented in the section of this chapter, “TET activity is influenced by mitochondrial metabolic products” have been published in: Laukka, T., Mariani, C.J., Ihantola, T., **Cao, J.Z.,** Hokkanen, J., Kaelin, W.G., Godley, L.A., and Koivunen, P. (2015). Fumarate and Succinate Regulate Expression of Hypoxia-Inducible Genes via TET Enzymes. *J Biol Chem*. 291(8):4256-65. PMID: 26703470.

(ii) Data presented in the section of this chapter, “Mitochondrial superoxide disrupts 5-mC/5-hmC balance in activated T-cells” have been published in: Moshfegh, C.M., Collins, C.W., Gunda, V., Vasanthakumar, A., **Cao, J.Z.,** Singh, P.K., Godley, L.A., and Case, A.J. (2019). Mitochondrial superoxide disrupts the metabolic and epigenetic landscape of CD4+ and CD8+ T-lymphocytes. *Redox Biol*, 27:101141. PMID: 30819616.

(iii) Data presented in the section of this chapter, “MYC-dependent D2HGDH and L2HGDH transcription influences TET activity and the epigenome” have been published in: Qiu, Z., Lin, A-P., Jiang, S., Elkashef, S.M., Myers, J., Srikantan, S., Sasi, B., **Cao, J.Z.,** Godley, L.A., Rakheja, D., Lyu, Y., Zheng, S., Madesh, M., Shiio, Y., Dahia, P.L.M., Aguiar, R.C.T. (2020). MYC Regulation of D2HGDH and L2HGDH Influences the Epigenome and Epitranscriptome. *Cell Chem Biol.* 27(5):538–550.e7. PMID: 32101699.

(iv) Data presented in the section of this chapter, “Post-translational modifications (PTMs) affects TET activity and stability” have been published in: Jeong J.J., Gu X., Nie J., Sundaravel S., Liu H., Kuo W-L., Bhagat T.D., Pradhan K., **Cao J.,** Nischal S., McGraw K.L., Bhattacharyya S., Bishop M.R., Artz A., Thirman M.J., Moliterno A., Ji P., Levine R.L., Godley L.A., Steidl U., Bieker J.J., List A.F., Saunthararajah Y., He C., Verma A., Wickrema A. (2019). Cytokine-Regulated Phosphorylation and Activation of TET2 by JAK2 in Hematopoiesis. *Cancer Discov.* 9(6):778-95. PMID: 30944118.

In all of the studies listed above, I performed mass spectrometry to measure the abundance of DNA cytosine covalent modifications. In study (iv), I also performed hMe-SEAL to pull down 5-hmC containing DNA for next-generation sequencing, as well as analyzed and interpreted the data from the sequencing result.

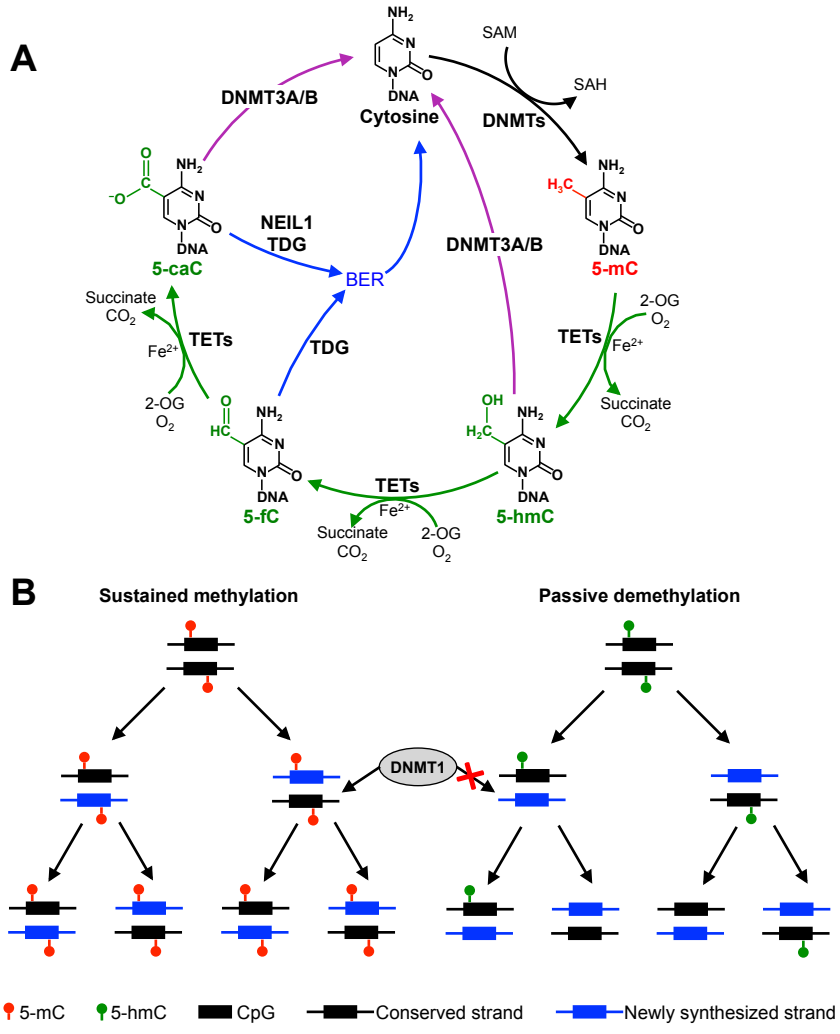
Cytosine modifications in mammalian DNA

Cells in multicellular organisms are exposed to various signals in their microenvironment, including small molecules such as oxygen or glucose, as well as macromolecules like cytokines, which induce changes in cellular behavior. Cellular response to these stimuli often results in change in transcription and translation activities as well as metabolic shifts, and in the case of stem cells, differentiation toward specific lineages. These processes all require expression of new gene programs that involve changes to the covalent DNA cytosine and histone modifications that allow those genes to be expressed (Atlasi and Stunnenberg, 2017; Geng et al., 2019; Turinetto and Giachino, 2015; Völker-Albert et al., 2020; Wu and Sun, 2006). These epigenetic regulations do not change the primary DNA sequence and are generally inherited by daughter cells after mitosis.

DNA cytosine modifications have been known to have important roles in maintaining normal functions of cells since the early 1980s (Baylin et al., 1986; Feinberg and Vogelstein, 1983; Gama-Sosa et al., 1983). In mammalian genomes, cytosines within 5'-CpG-3' dinucleotides (CpG) can be methylated to 5-methylcytosine (5-mC) by DNA methyl transferases (DNMTs) using S-adenosylmethionine (SAM) as the methyl group donor (Figure 1.1A, Figure 1 of original publication). DNMT1 has optimal activity for methylating hemimethylated cytosines and thus functions mainly to maintain DNA methylation patterns after DNA replication (Lee et al., 2014). In contrast, DNMT3A and DNMT3B function mainly as *de novo* methyltransferases (Lee et al., 2014). When found within gene bodies, 5-mC is associated with gene expression (Jjingo et al., 2012; Jones, 2012; Varley et al., 2013), whereas when the modified base is concentrated in

CpG islands that overlap gene promoters, 5-mC represses gene transcription, a phenomenon commonly observed in cancer cells (Baylin et al., 1986; Feinberg and Tycko, 2004; Feinberg and Vogelstein, 1983; Gama-Sosa et al., 1983).

Figure 1.1 Cytosine modifications and active/passive demethylation



(**A**) Cytosine is methylated by DNMTs to 5-mC. TETs convert 5-mC to 5-hmC in an oxygen and 2-OG dependent reaction. Further oxidation of 5-hmC by TETs produces 5-fC and 5-caC, which can be excised and replaced by unmodified cytosine via base excision repair (BER). In addition, DNMT3A/B are shown to directly remove the modification from 5-hmC and 5-caC. (**B**) 5-hmC promotes passive demethylation by blocking CpG recognition of DNMT1. As methylation is not maintained by DNMT1 after each cell division, the cytosine eventually becomes unmodified after consecutive cell divisions.

The first oxidized product of 5-mC, 5-hydroxymethylcytosine (5-hmC), was originally identified as a specialized base in T2, T4, and T6 bacteriophages (Wyatt and Cohen, 1953). Although the presence of 5-hmC was also observed in mammalian DNA, it was believed that this base was the result of oxidative damage (Tardy-Planechaud et al., 1997; Valinluck and Sowers, 2007). 5-hmC did not receive major attention as a chromatin-regulating base until it was discovered that 5-hmC is the oxidized product of TET dioxygenase activity using 5-mC as a substrate (Iyer et al., 2009; Tahiliani et al., 2009). TET and other dioxygenases also require oxygen and 2-OG as additional substrates, and use Fe(II) as a cofactor. Since then, studies have identified *TET2* as one of the most commonly mutated genes in acute myeloid leukemia (AML), myelodysplastic syndrome (MDS), and chronic myelomonocytic leukemia (CMML), establishing 5-hmC as an important epigenetic modification in the hematopoietic system (Abdel-Wahab et al., 2009; Bowman and Levine, 2017; Cheng et al., 2013; Delhommeau et al., 2009; Figueroa et al., 2010; Ko et al., 2010; Langemeijer et al., 2009; Pan et al., 2017; Yan et al., 2017). Studies in other normal systems, including embryonic/pluripotent stem cells and neuronal tissues, likewise showed that TETs and 5-hmC are indispensible for normal function of the cells, and deregulation of TETs contribute to malignant progression including neuroblastoma, glioma, melanoma, squamous cell carcinoma, and colorectal cancer (Chapman et al., 2015; Ficz et al., 2011; Guo et al., 2011; Jäwert et al., 2013; Koh et al., 2011; Kumar et al., 2015; Lian et al., 2012; Mariani et al., 2014; Pastor et al., 2013; Rasmussen and Helin, 2016; Verma et al., 2018; Xu et al., 2012; Zhang et al., 2013).

The functions and physiological roles of 5-hmC

5-hmC is a relatively rare base in mammalian genomes, and its abundance is highly tissue dependent. For example, 5-hmC consists of 0.01-0.3% of all cytosine species in myeloid cells (Ko et al., 2010; Madzo et al., 2014), constitutes around 1% of all cytosines in cerebral tissue (Field et al., 2015; Kriaucionis and Heintz, 2009; Wen et al., 2014), and is present at levels as high as 5% of all cytosines in undifferentiated embryonic stem (ES) cells (Tahiliani et al., 2009). 5-hmC can be converted further to 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC), but the level of these two bases are 2-3 orders of magnitude lower than those of 5-hmC (Ito et al., 2011). Currently, two major functions of 5-hmC are recognized: the base is a component of the demethylation process and is also a stable epigenetic mark that promotes gene transcription.

5-hmC mediates passive and active DNA demethylation

5-hmC facilitates demethylation both within passive and active DNA demethylation pathways (Figure 1.1, Figure 1 of original publication). When DNA is replicated, a newly inserted cytosine on the daughter strand is unmodified at first. If the cytosine base within the parent 5'-CpG-3' dinucleotide is methylated, then DNMT1 methylates the cytosine within the newly synthesized strand since its highest activity is on hemi-methylated CpG dinucleotides (Lee et al., 2014). DNMT1 sits at the replication fork to catalyze this reaction as DNA is synthesized, thereby maintaining the DNA methylation pattern in the chromatin of daughter cells (Figure 1.1B). However, because DNMT1 does not recognize 5-hmC as a substrate, a hydroxymethylated cytosine base within the parent DNA strand will result in an unmodified cytosine in the daughter strand

(Seiler et al., 2018; Valinluck and Sowers, 2007). This process is referred to as passive demethylation (Figure 1.1B). Passive demethylation is observed in cells undergoing rapid division, for example in early erythropoiesis as reported by Madzo *et al.* (Madzo et al., 2014), and leads to decreased 5-hmC as well due to decreased substrate availability. However, some regions of the genome gain 5-hmC despite rapid division, suggesting a two-step process in which the newly synthesized DNA strand is first methylated by DNMTs and then hydroxymethylated by TETs. The function(s) of 5-hmC a stable epigenetic mark will be discussed in the next section.

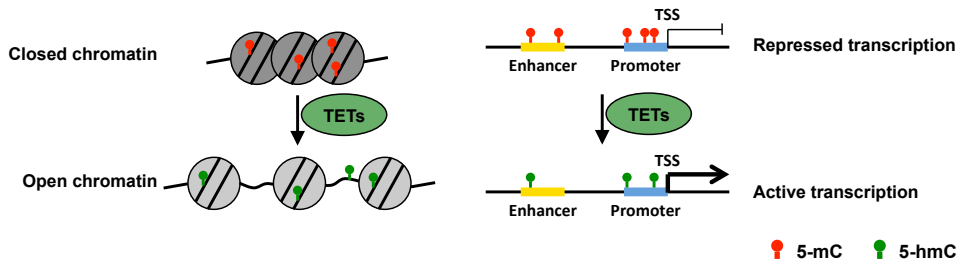
A more active and cell-cycle independent demethylation pathway involves further oxidation of 5-hmC by TETs to 5-fC and 5-caC (Ito et al., 2011), both of which can be replaced by unmodified cytosines through base-excision repair (BER) by thymine-DNA glucosylase (TDG) (Coey et al., 2016; He et al., 2011; Maiti and Drohat, 2011; Pidugu et al., 2016) (Figure 1.1A). 5-caC, but not 5-fC, can also undergo BER by Nei-like 1 (NEIL1) excision (Slyvka et al., 2017) (Figure 1.1A). In addition to BER-mediated demethylation, some studies have also found that DNMT3A/B have the ability to remove the 5-modification from 5-hmC and 5-caC directly, but not from 5-mC or 5-fC, resulting in unmodified cytosine (Chen et al., 2012; Liutkevičiūtė et al., 2014) (Figure 1.1A).

5-hmC is an activating epigenetic mark and is involved in DNA damage repair

5-hmC can also function as a stable and distinct epigenetic mark associated with active gene transcription. Many reports have found that 5-hmC is strongly associated with promoters, enhancers, and transcription factor (TF) binding sites (Madzo et al., 2014; Mariani et al., 2014; Stroud et al., 2011; Vasanthakumar and Godley, 2015)

(Figure 1.2). Similarly, 5-hmC is associated with active transcription and histone modifications that mark with open chromatin (Ficz et al., 2011; Lin et al., 2017; Madzo et al., 2014; Szulwach et al., 2011) (Figure 1.2). Evidence for a role of 5-hmC as a stable and actively maintained epigenetic mark is most apparent in the context of cell differentiation, where gene expression and chromatin structure undergo major changes. For example, in the course of erythropoiesis, 5-hmC at certain genomic loci remains highly enriched despite a global decrease in total 5-hmC level and multiple rounds of DNA replication (Madzo et al., 2014). To maintain 5-hmC during consecutive rounds of DNA replication, both DNMTs and TETs must be involved, indicating a functional role of 5-hmC at these loci.

Figure 1.2 5-hmC is an activator of gene expression



5-hmC has been shown to associate with open chromatin and active gene transcription. 5-hmC increases frequently co-occur with activating histone marks (left), and contribute to the remodeling of chromatin structure. 5-hmC is also enriched in promoters and enhancers of upregulated genes (right) and serves as a stable epigenetic mark for gene transcription.

In addition to its role in regulating chromatin structure and consequent gene expression, several studies have revealed potential involvement of 5-hmC in DNA damage repair. During murine embryonic development, oocytes with Tet1 loss-of-function have a much higher rate of defective meiosis and unresolved double strand

breaks compared to those with wild-type (WT) Tet1 (Yamaguchi et al., 2012). In addition, DNA damage in mouse Purkinje cells leads to increased 5-hmC levels (Jiang et al., 2015). Further studies in mouse ES cells showed that knock-out (KO) of all three Tet genes led to a high level of mitotic defects, as measured by the presence of chromatin fragments in anaphase, which resulted from a higher sensitivity to replication stress and delayed DNA repair (Kafer et al., 2016). In human cancer cell lines, 5-hmC foci co-localize with those associated with DNA damage marked by γH2AX and 53BP1 (Kafer et al., 2016). When *TET2* is knocked-down with shRNA, the 5-hmC foci no longer form, suggesting that TET2 is directly involved in this process (Kafer et al., 2016). It is possible that 5-hmC promotes local chromatin remodeling or serves as an epigenetic mark to recruit additional DNA damage repair machinery (Kafer et al., 2016).

TET enzyme structure and functions

TET genes are present in all metazoan genomes. In mammals and other gnathostomata, the *TET* gene underwent triplication to generate *TET1*, *TET2*, and *TET3* (Liu et al., 2020a). During this process, *TET2* likely underwent a local chromosome inversion causing the 5' end of the inversion to create a distinct gene, *CXXC4* (Iyer et al., 2009; Ko et al., 2013). All three *TET* genes are transcribed and translated into catalytically functional TET enzymes.

<u>TET protein structure and DNA binding mechanism</u>

All three TET proteins share a C-terminal catalytic domain, whereas TET1 and TET3 share an N-terminal CXXC domain (Figure 1.3A). The catalytic domains consist of

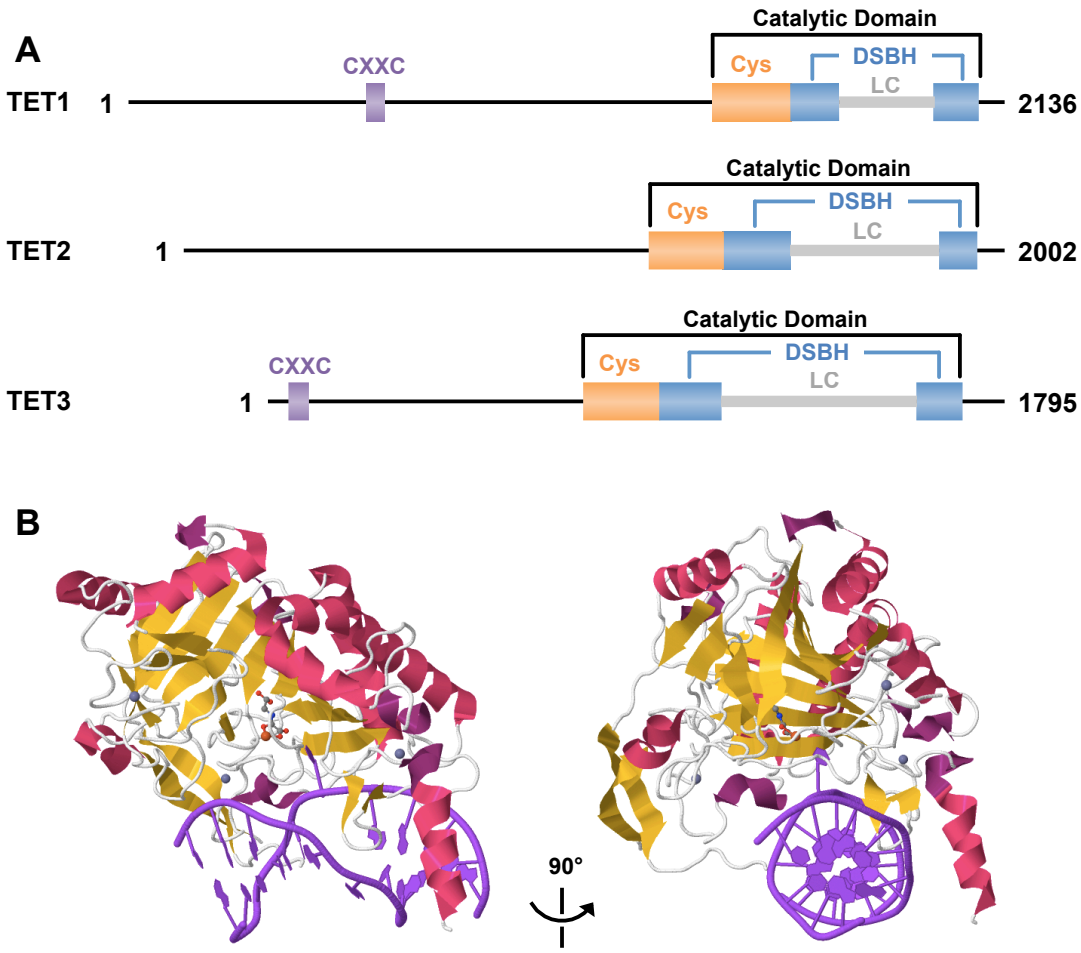
a cysteine-rich domain and a double-stranded β-helix (DSBH) domain, together forming a globular structure with the core DNA-binding DSBH in the center (Hu et al., 2013; Hu et al., 2015) (Figure 1.3). TET catalytic domains specifically recognize and bind CpG dinucleotides and flip the modified cytosine base into the catalytic cavity, with no preference for flanking sequences (Hu et al., 2013) (Figure 1.3B). This base flipping mechanism appears to be conserved evolutionally, as the same mechanism is found in the *Naegleria gruberi* Tet enzyme (Hashimoto et al., 2015; Hashimoto et al., 2014).

Substrates and cofactors of TET

TET proteins are 2-oxoglutarate (2-OG, aka α-ketoglutarate) and oxygen (O_2) dependent dioxygenases that use Fe(II) as a co-factor. TET catalytic domains share high homology to the previously discovered J-binding proteins (JBP1/2) (Iyer et al., 2009). JBP1/2 are dioxygenases in kinetoplastids that catalyze the oxidation of thymine at the methyl group to carboxyuracil, which is the first catalytic step in forming base J (Iyer et al., 2009). Based on this homology, the catalytic function of oxidizing 5-mC to 5-hmC was identified for the TETs (Tahiliani et al., 2009). Additional studies found that TETs can oxidize 5-hmC further to 5-fC and 5-caC in a stepwise fashion (He et al., 2011; Ito et al., 2011) (Figure 1.1A). At a minimum, the TETs are 2-3 fold more efficient at oxidizing 5-mC compared to 5-hmC or 5-fC due to differences in the conformation of the active sites (Hu et al., 2015), which leads to faster 5-hmC creation than its removal through further oxidation. This difference in catalytic efficiency allows the accumulation of 5-hmC at some loci rather than further oxidation to 5-fC or 5-caC. The catalytic

differences, together with TDG/NEIL1 mediated removal of 5-fC and 5-caC, may explain the relative abundance of 5-hmC over 5-fC and 5-caC.

Figure 1.3 Structural domains of the TET proteins



(**A**) All three TETs share a evolutionarily conserved catalytic domain, consisting of a cystein-rich domain (Cys) and a double-stranded beta-helix (DSBH) domain. A non-conserved low-complexity (LC) insert of variable length is present in the DSBH domain in all three TETs. TET1 and TET3 also possess a CXXC domain near the N-terminus, whereas TET2 lost the CXXC domain during evolution. (**B**) Crystal structure of the catalytic domain of TET2 binding to 5'-CpG-3' containing DNA (Protein Data Bank (PDB) ID: 4NM6). Note that the cytosine base is flipped and inserted into the active site of the DSBH domain.

In a study I participated in, we show that the K_m values of TETs for oxygen, Fe(II), and 2-OG are about 30μM, 4μM, and 60μM, respectively (Table 1.1, adapted

from Table 2 of the original publication) (Laukka et al., 2015). Compared to the prolyl hydroxylases (PHD) that utilize the same reaction to hydroxylate the hypoxia-inducible factor (HIF) proteins, the oxygen affinity of TETs is almost 10 times higher, allowing TET1/2 to be functional in an environment much lower oxygen availability (Laukka et al., 2015). This study also showed that the TETs bind modified CpG dinucleotides (5-mC, 5-hmC, or 5-fC) at around 100nM, much stronger than oxygen, Fe(II), or 2-OG, making oxygen the most likely rate-limiting substrate (Laukka et al., 2015). Additionally, this study showed that a number of *TET2* missense mutations found in AML lead to reduced affinity to 2-OG or Fe(II) and greatly reduced catalytic efficiency (Table 1.2, adapted from Table 3 of the original publication).

Table 1.1 K_m values of TET1 and TET2 in comparison to HIF/collagen prolyl hydroxylases.

Compound (K_m unit)	TET1	TET2	HIF PHD	Collagen PHD
Methylated DNA (nM)	75 ± 55	125 ± 85	-	-
Fe(II) (μM)	4.8 ± 4	3.6 ± 3	0.03 ± 0.004	2
2-oxoglutarate (μM)	55 ± 20	60 ± 15	60	20
Oxygen (μM)	30 ± 10	30 ± 3	250	40

Table 1.2 K_m and V_{max} values of AML-associated TET2 mutations.

Variable (unit)	WT TET2	H1302Y	D1304A	H1802R	R1817M	R1817S
K_m of Fe(II) (μM)	3.6 ± 3	110 ± 4	165 ± 90	200 ± 185	ND	ND
K_m of 2-OG (μM)	60 ± 15	ND	ND	ND	>5000	>5000
V_{max} (% WT TET2)	100	45	55	50	ND	ND

ND: not determined.

2-OG is produced from isocitrate by the isocitrate dehydrogenases (IDH1/2) in the TCA cycle and is typically abundant in cells and therefore not limiting for TET-mediated catalysis. However, 2-OG analogs can accumulate in pathological conditions as a result of mutations in genes encoding TCA cycle enzymes (Dang et al., 2009; Pollard et al., 2005). The IDH1 R100/R132 and IDH2 R140/R172 amino acid substitutions occur in the active sites of the enzymes and instead of producing 2-OG,

the mutated enzymes convert isocitrate into 2-hydroxyglutarate (2-HG), which is a potent inhibitor of dioxygenases including TETs (Dang et al., 2009; Losman and Kaelin, 2013; Xu et al., 2011). As a result, tumors with *IDH1 or IDH2* mutations often display DNA hypermethylation similar to those caused by *TET* loss-of-function mutations (Figueroa et al., 2010; Turcan et al., 2012). It is worth noting that *IDH* mutations and *TET* mutations are almost always mutually exclusive, highlighting the close relationship between IDH and TET in the same molecular pathway (Hou and Tien, 2020; Patel et al., 2012).

TET activity is enhanced by ascorbate (vitamin C) by maintaining the Fe co-factor in a reduced (II) state (Kuiper and Vissers, 2014). Ascorbate is essential for the function of collagen prolyl-4-hydroxylase, another 2-OG/O2 dependent dioxygenase similar to the TETs (Kuiper and Vissers, 2014). Ascorbate deficiency causes failure of prolyl hydroxylation of collagen that is required to stabilize its secondary structure, which leads to scurvy (Gorres and Raines, 2010; Kuiper and Vissers, 2014). The addition of ascorbate to cultured mouse ES cells or embryonic fibroblasts leads to a rapid increase in 5-hmC levels that is Tet dependent (Blaschke et al., 2013; Minor et al., 2013; Yin et al., 2013). Demethylation is also demonstrable in ascorbate-treated human ES cells and mouse embryonic fibroblasts as well as both up- and downregulation of numerous genes (Blaschke et al., 2013; Chen et al., 2013; Chung et al., 2010).

TET activity is influenced by mitochondrial metabolic products

In most eukaryotes, mitochondria are responsible for generating almost all of the cell's ATP through oxidative glycolysis to fulfill its energy needs. Glycolysis starts in the

cytosol, where glucose is gradually broken down to pyruvate, generating a small amount of energy in the process. Pyruvate is then oxidized to acetyl-CoA, which together with oxaloacetate, is used to synthesize citrate, where the carbons enter the tricarboxylic acid (TCA) cycle in mitochondria. Through successive steps in the TCA cycle, citrate is first converted to isocitrate, and then oxidized to 2-OG by IDH1/2. 2-OG is oxidized in subsequent steps to succinate, then fumarate, and eventually converted back to oxaloacetate, completing the cycle. Each molecule of acetyl-CoA generates 3 NADH and hydrogen ions (H^+) through the TCA cycle, which are then used in the oxygen-dependent electron transport chain to generate 2.5 ATPs per NADH+H^+.

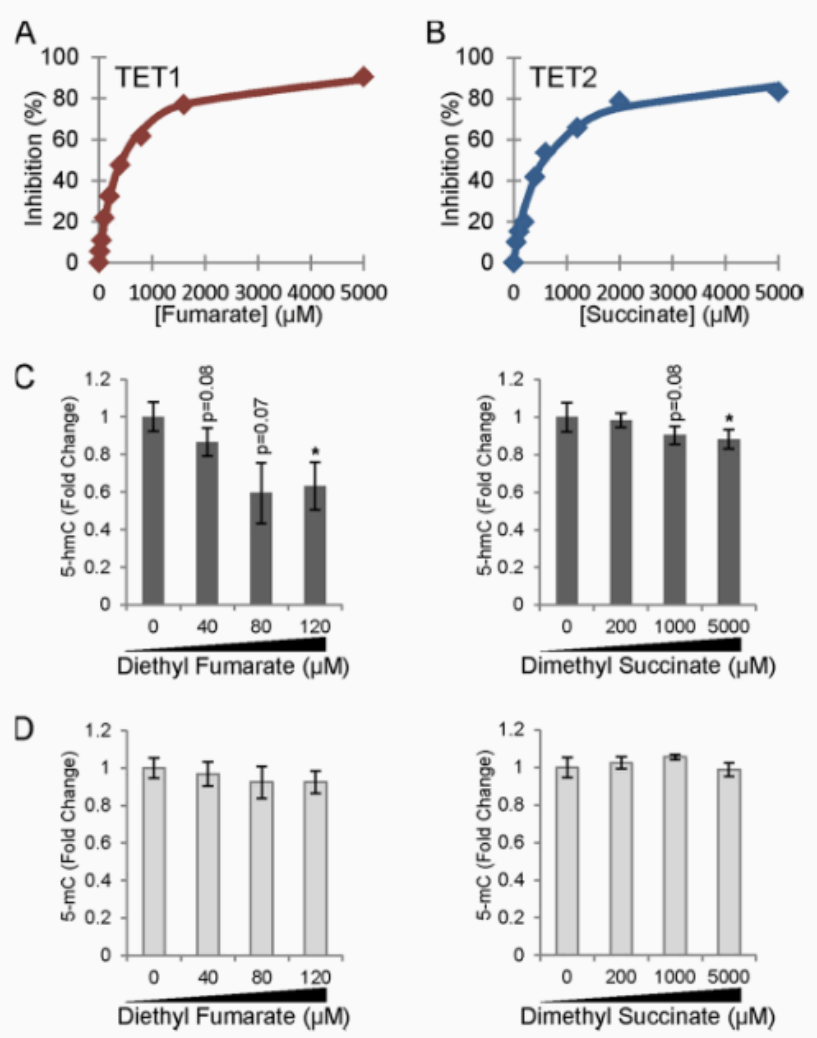
As mentioned earlier, the TET enzymes require 2-OG and oxygen as co-substrates to carry out their enzymatic reactions. Since 2-OG and oxygen are both crucial parts of the mitochondrial metabolic processes, TET activity is tied intimately to the metabolic state of mitochondria (Dang et al., 2009; Losman and Kaelin, 2013). Besides the 2-HG discussed above, two additional TCA cycle metabolites, fumarate and succinate, accumulate in hereditary cancers with *fumarate hydratase* (*FH*) or *succinate dehydrogenase* (*SDH*) mutations (Pollard et al., 2005). In particular, *SDH* mutations predispose individuals to hereditary paraganglioma and pheochromocytomas, and *FH* mutations predispose patients to hereditary leiomyoma and renal cell cancer (Pollard et al., 2005). In the Laukka *et al.* (2015) study, we measured the inhibitory effect of succinate and fumarate on TET1 and TET2. The IC_{50} for succinate and fumarate are around 550μM and 400μM, respectively, for both TET1 and TET2, making them much stronger inhibitors in comparison to (R)- or (S)-2-HG (Table 1.3, adapted from Table 4 of the original publication, and Figure 1.4 A-B, Figure 3 of the original publication)

(Laukka et al., 2015). Here, I performed mass spectrometry to quantify global 5-mC and 5-hmC in SK-N-BE(2) cells treated with diethyl fumarate (DEF) and dimethyl succinate (DMS), which are cell-permeable versions of fumarate and succinate, respectively. My measurements showed that increased fumarate or succinate levels induced a loss of global 5-hmC (Figure 1.4C) (Laukka et al., 2015), supporting the roles of succinate and fumarate as inhibitors of TETs. I did not find significant changes in 5-mC levels in the treated cells (Figure 1.4D) (Laukka et al., 2015), which could be in part due to the treatment duration (48 hours) of DEF and DMS on these cells was too short to exert an measurable effect on global 5-mC level.

Table 1.3 IC_{50} values of succinate and fumarate for TET1 and TET2.

Enzyme (unit)	Succinate	Fumarate	(R)-2-HG	(S)-2-HG
TET1 (μM)	540 ± 100	390 ± 160	4000	1000
TET2 (μM)	570 ± 190	400 ± 70	5000	1600

Figure 1.4 TET inhibition by succinate or fumarate led to global 5-hmC decrease in SK-(N)-BE2 cells.



(**A-B**) IC_{50} curves of Tet1 and Tet2 for fumarate and succinate, respectively. (**C-D**) HPLC-MS/MS quantification of global (**C**) 5-hmC and (**D**) 5-mC levels in SK-N-BE(2) cells incubated with increasing concentrations of diethyl fumarate or dimethyl succinate for 48 h (n≥3). 5-hmC and 5-mC quantitation graphs represent mean ± S.E. *, p<0.05.

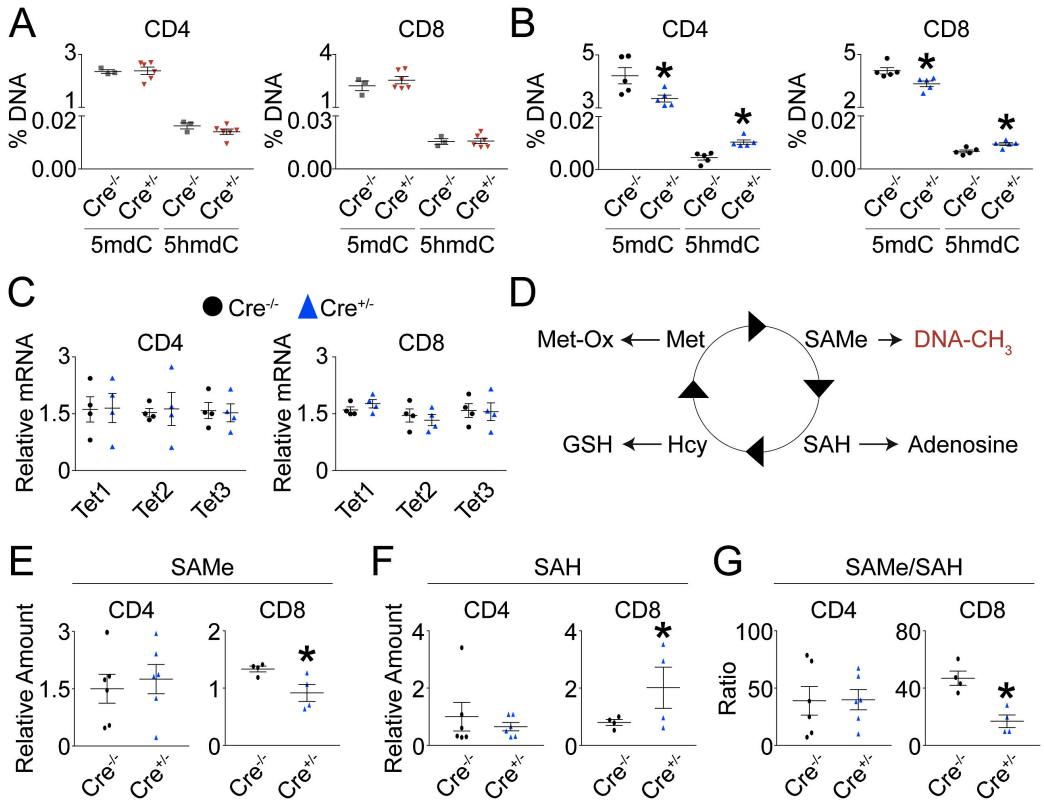
Mitochondrial superoxide disrupts 5-mC/5-hmC balance in activated T-cells

Mitochondria are the primary organelle where reactive oxygen species (ROS) are generated, mainly through the electron transport chain (Murphy, 2009). Eukaryotes that have evolved in an oxidative environment have developed a highly conserved system to

counter the excessive production of ROS (Case, 2017). A central component to this system is the manganese-dependent superoxide demutase (MnSOD), which converts the highly reactive and toxic superoxide ($O_2^{\bullet -}$) into more stable hydrogen peroxide and molecular oxygen (Case, 2017). The importance of MnSOD is highlighted by the early post-natal lethality of constitutive KO animals (Case et al., 2011). In another study I contributed to, we showed that loss of MnSOD and increase of $O_2^{\bullet -}$ caused profound metabolic changes in both activated CD4+ and CD8+ T-cells, including decreased oxygen consumption rates, decreased ATP production, decreased glycolytic capacity, and cytokine/chemokine dysregulation (Moshfegh et al., 2019). I performed mass spectrometry to measure the 5-mC and 5-hmC levels in both naïve and activated CD4+ and CD8+ T-cells with or without MnSOD deletion (Figure 1.5 A-B, Figure 6 of original publication) (Moshfegh et al., 2019). Although there was no appreciable change in global 5-mC and 5-hmC levels in naïve CD4+ and CD8+ T-cells (Figure 1.5A), loss of MnSOD resulted in less 5-mC and more 5-hmC in activated T-cells (Figure 1.5B) (Moshfegh et al., 2019). This change in cytosine modification was not caused by expression changes in any of the *TET* genes (Figure 1.5C) (Moshfegh et al., 2019). The study then showed metabolic perturbation in the methionine cycle (Figure 1.5 D-G), suggesting that the decreased SAM required for cytosine methylation as a main cause for the hypomethylation (Moshfegh et al., 2019). However, as the SAM decrease was only observed in CD8+ T-cells (Figure 1.5E), the methionine cycle alone could not explain the epigenetic changes completely (Moshfegh et al., 2019). It is also possible that the superoxide increases TET activity, so that more 5-mC is converted to 5-hmC,

causing hypomethylation in these cells, although this hypothesis was not tested in this study and remains an open question.

Figure 1.5 Elevated mitochondrial $O_2^{\bullet-}$ disrupts the T-cell metabolic-epigenetic axis.



Tamoxifen was administered to MnSOD knock-out (*Cre*$^{+/-}$) and control (*Cre*$^{-/-}$) mice for five consecutive days followed by two weeks to allow full recombination and loss of MnSOD. Splenic T-cells were isolated from the mice and analyzed either in their naïve or activated (cultured with anti-CD3 and anti-CD28 beads for 96 h) states. (**A**) Global DNA cytosine methylation (5mdC) and hydroxymethylation (5hmdC) shown as percent of total cytosines in naïve T-cells. (**B**) Global 5mdC and 5hmdC shown as a percent of total cytosines in activated T-cells. (**C**) *TET* gene mRNA levels as assessed by quantitative real-time RT-PCR. (**D**) Schematic of the methionine cycle in reference to DNA methylation. Targeted intracellular metabolomics analysis of (**E**) S-adenosyl methionine (SAM), (**F**) S-adenosyl homocysteine (SAH), and (**G**) SAMe/SAH ratio. *p < 0.01 versus activated *Cre*$^{-/-}$ by parametric Student's t-test, nonparametric Mann-Whitney U test, or 2-way ANOVA with Bonferroni post-hoc analysis where appropriate.

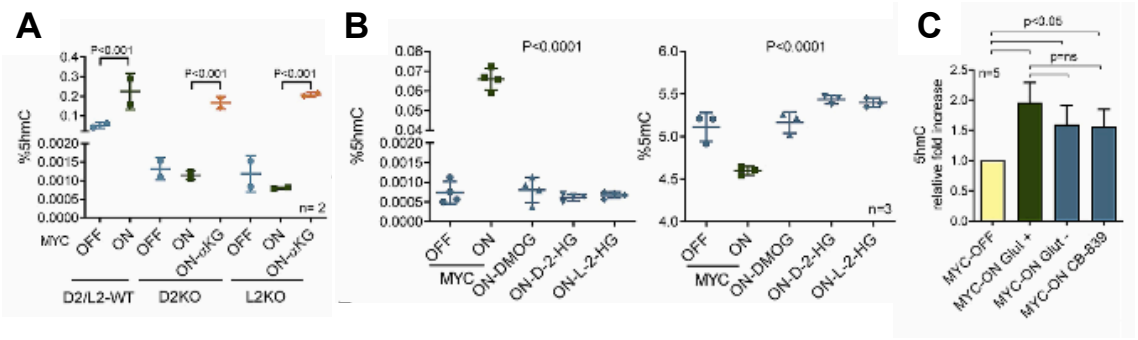
MYC-dependent *D2HGDH* and *L2HGDH* transcription influences TET activity and the epigenome

In addition to mutated IDH proteins, 2-HG is also produced as a byproduct of a number of normal enzymes in various metabolic processes (Ye et al., 2018). In non-pathological conditions, the naturally occurring D- and L-2-HG enantiomers are converted back to 2-OG by D-2-HG dehydrogenase (D2HGDH) and L-2-HG dehydrogenase (L2HGDH), respectively (Ye et al., 2018). In Qiu *et al.* (2020), we identified MYC as a TF that binds the promoters of *D2HGDH* and *L2HGDH*, and their expression levels are MYC-dependent (Qiu et al., 2020). The effects of MYC on the epigenome were tested in two cell models: the human B cell line P493-6, where *MYC* expression is under the control of tetracycline responsive element, and mouse B cells from WT or Eμ-Myc mice that have higher *Myc* expression (Qiu et al., 2020). Through mass spectrometry, I found that higher MYC/Myc levels are associated with higher global 5-hmC and lower 5-mC levels (Figure 2 of the original publication) (Qiu et al., 2020). Notably, TET protein levels remained unchanged in all the samples, suggesting that the effects were only mediated through TET activity (Qiu et al., 2020). I found that MYC-dependent 5-mC and 5-hmC changes was severely attenuated when either D2HGDH or L2HGDH was knocked-out, consistent with the reduced TET activity changes (Figure 3 of the original publication) (Qiu et al., 2020).

In order to show that D2HGDH and L2HGDH affect the epigenome via their metabolic activities, P493-6 cells were treated with compounds that either increase 2-OG (dimethyl-α-ketoglutarate, DM-α-KG) or 2-HG (octyl-D-2HG or octyl-L-2HG), as well as the 2-OG analog DMOG that competitively inhibits the dioxygenases. Using mass

spectrometry, I found that DM-α-KG restored 5-hmC in *D2HGDH* or *L2HGDH* KO cells to a level comparable to WT cells (Figure 1.6, Figure 4 of the original publication) (Qiu et al., 2020). Conversely, introduction of D/L-2-HG or DMOG reduced 5-hmC and increased 5-mC to levels similar to that seen in MYC-OFF P493-6 cells (Figure 1.6B). Lastly, because MYC also regulates the enzymes involved in glutaminolysis, where glutamine is converted first to glutamate and then 2-OG (Hsieh et al., 2015), P492-6 cells were cultured in glutamine-free medium or treated with the glutaminase inhibitor CB-839 (Figure 1.6C). I found a trend in 5-hmC levels, with a slightly decrease in cells grown in glutamine-free medium or treated with CB-839, although the result was not statistically significant (Figure 1.6C). Overall, results of this study show that TET activity is enhanced by MYC through the clearing of 2-HG, providing a link between MYC protein levels and the 5-hmC landscape.

Figure 1.6 MYC-D2HGDH-L2HGDH Axis, Intermediate Metabolites, and TET Activity.



(**A**) 5-hmC level in P493-6 cells (MYC OFF or ON; D2HGDH and L2HGDH WT or KO), with or without DM-α-KG rescue (5 mM); p values are from two-tailed Student's t test. (**B**) 5-hmC and 5-mC levels in P493-6 cells (MYC OFF or ON) exposed to DMOG (1 mM), octyl-D-2-HG or L-2-HG (100 mM); p values are from one-way ANOVA. (**C**) Fold increase in 5-hmC upon MYC induction in cells grown in full media (Glut+), in glutamine-free media (Glut-) or in the presence of CB-839 (1 mM) (3 biological replicates for both assays, 5 technical replicates for 5-hmC); p values are from two-tailed Student's t test. Data are means ± SD.

Post-translational modifications (PTMs) affects TET activity and stability

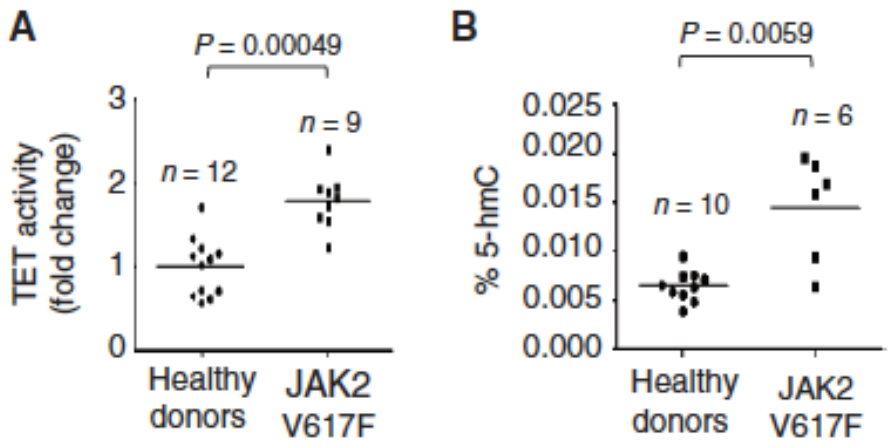
As discussed earlier, the proper maintenance of the epigenetic landscape during hematopoiesis is crucial for successful differentiation, including erythropoiesis (Madzo et al., 2014). The Janus-activated kinases (JAKs) are receptor-associated tyrosine kinases activated by cytokine receptors. The JAK family consists of JAK1, JAK2, JAK3, as well as tyrosine kinase 2. In erythropoiesis, erythropoietin (EPO), stem cell factor (SCF), and FLT3 ligand (FLT3L) are key regulators of the differentiation program that utilize the JAK2 signaling pathway. However, the pathways linking cytokine response and epigenetic changes have not been delineated well. In addition, Madzo *et al.* (2014) observed a maximal level of TET2 protein at day 3, which was inconsistent with *TET2* transcription that is highest at day 0 and substantially decreased by day 3 (Madzo et al., 2014). This change in *TET2* transcription was validated by my RNA-seq data (Chapter II) (Cao et al., 2020). The disconnect between *TET2* mRNA expression and protein level in erythropoiesis is likely due to PTMs that increase TET2 stability. A known PTM that enhances TET2 stability is acetylation, which inhibits polyubiquitination and protect against proteasome degradation (Zhang et al., 2017b). Other studies have shown numerous PTMs that modulate TET activities, including O-GlcNAcylation (Hrit et al., 2018; Shi et al., 2013; Vella et al., 2013; Zhang et al., 2014), phosphorylation (Bauer et al., 2015; Wu et al., 2018), and monoubiquitination (Nakagawa et al., 2015).

More recently, I took part in a study showing that TET2 is phosphorylated by JAK2, leading to increased activity of the enzyme (Jeong et al., 2019). In this study, I showed that the total 5-hmC level increased under FLT3L, SCF, and EPO treatment, which is consistent with increased TET2 activity (Figure 1 of original publication) (Jeong

et al., 2019). We identified that two tyrosine residues on TET2, Y1939 and Y1964, were phosphorylated by EPO treatment, and showed that JAK2 is directly responsible for phosphorylating these sites (Jeong et al., 2019). Interestingly, these two tyrosine residues are unique to TET2 and are not present in TET1 or TET3. We further showed that mutation of both residues on TET2 impaired its binding to KLF1, and cells with the double mutation exhibited reduced colony forming capability (Jeong et al., 2019).

We then investigated how *JAK2*^{V617F}, a common activating mutation of JAK2 found in myeloid malignancies, affects the distribution of covalently-modified cytosines across the genome. I measured 5-hmC levels in DNA obtained from hematopoietic stem/progenitor cells in healthy donors versus patients with *JAK2*^{V617F} myeloproliferative neoplasms (MPN). 5-hmC levels in the MPN patients were higher than in healthy donors (Figure 1.7B, Figure 6 of original publication), consistent with TET2 activity in the same samples (Figure 1.7A) (Jeong et al., 2019). Comparing the *JAK2*^{V617F} to *JAK2*^{WT}, we found that this mutation also caused hypomethylation in cultured murine bone marrow Factor Dependent Continuous cell line, Paterson Laboratories (FDCP) and idiopathic myelofibrosis (IMF) samples. In the IMF samples, *JAK2*^{V617F} mutation also led to a greater number of genes overexpressed than ones that were underexpressed, which is consistent with global hypomethylation and the increase of 5-hmC (Figure 1.7B) (Jeong et al., 2019).

Figure 1.7 *JAK2*^{V617F} activating mutation leads to higher TET2 activity and higher 5-hmC



(**A**) TET2 activity assay in HSPC samples obtained from healthy donors vs. MPN patients with *JAK2*^{V617F} mutation. (**B**) Global 5-hmC levels in HSPC DNA from healthy donors vs. MPN patients with *JAK2*^{V617F} mutation. HSPC: hematopoietic stem/progenitor cells.

The physiological roles of TET enzymes in hematopoiesis

TET enzymes are indispensible at various stages of development and differentiation. In normal hematopoiesis, only *TET2* and *TET3* are expressed (Uhlén et al., 2015; Uhlen et al., 2017; Yan et al., 2017). Abnormal *TET1* expression, such as in the case of *TET1-MLL* fusion AML, is a key driver of the malignancy (Lee et al., 2013; Lorsbach et al., 2003; Meyer et al., 2013; Ono et al., 2002). *TET2* is one of the most well studied genes in leukemia, given the abundance of *TET2* mutations found in various types of leukemia: *TET2* mutations are found in 20-30% of AML, 20-30% of MDS or myeloproliferative neoplasm, and up to 58% of CMML (Delhommeau et al., 2009; Itzykson et al., 2013; Jankowska et al., 2009; Langemeijer et al., 2009; Yamazaki et al., 2015; Yamazaki et al., 2012). Copy number alterations of *TET2* are found in

about 5.6% of various hematologic malignancies, 70% of which also have cytogenetic alterations, most of which involve complex karyotypes (Bacher et al., 2012). Notably, most in-frame deletions and missense mutations occur in exons that code for the C-terminal catalytic domain, whereas nonsense and frameshift deletions can occur along the entire gene (Euba et al., 2012; Langemeijer et al., 2009). Most of the known mutations result in truncated proteins or those with impaired catalytic function (Euba et al., 2012; Langemeijer et al., 2009). Many of the AML associated missense mutations encode changes in the catalytic domain of TET2 reducing the affinity of the enzyme for Fe(II) and 2-OG and greatly reducing its activity (Laukka et al., 2015). *TET3* mutations are rare compared to *TET2* mutations in myeloid leukemias. An early study with a cohort of 408 myeloid leukemia patients did not find any *TET3* mutations (Abdel-Wahab et al., 2009). More recently, two loss-of-function mutations were identified in *TET3* in the bone marrow of 28 CMML patients, both co-existing with *TET2* mutations (Merlevede et al., 2016). In addition, another study identified seven *TET3* mutations in a study of 83 leukemia patients (Lasho et al., 2018). None of the newly identified mutations co-occurred with *TET2* mutations. Six of these mutations were found in CMML patients (16% of all CMML patients), two of whom were found in the same patient (Lasho et al., 2018). These findings suggest that *TET3* mutations preferentially occur in CMML, although the significance and consequences of *TET3* mutations remain to be studied.

TET2 mutated AML cells have globally decreased 5-hmC (Ko et al., 2010; Madzo et al., 2014), but the mechanism behind this loss is not understood. Early studies of *TET2*-mutated myeloid leukemia patient blood or bone marrow samples found no global difference in 5-mC, but showed site-specific hyper- and hypomethylation compared to

healthy controls (Ko et al., 2010). In contrast, other studies have shown global hypermethylation in CMML cells (Madzo et al., 2014). It should be noted that Ko et al. (2010) measured global methylation using the Illumina Infinium 27K array that probes 27,578 selected CpG sites, whereas Madzo *et al.* (2014) used mass spectrometry that measure all nucleotides in the genome, which may explain the discrepancy in these findings. Curiously, in this study, distribution of 5-hmC in *TET2*-mutated CMML was dramatically different from normal, which is not explained solely by the global decrease of 5-hmC. Although most 5-hmC peaks decrease correspondingly with the global decrease of 5-hmC, certain sites gain a high density of 5-hmC not seen in normal controls. This redistribution process suggests that *TET1* or *TET3* may compensate for the loss of *TET2,* resulting in skewed 5-hmC distribution as well as increased propensity to differentiate toward the myeloid, but not erythroid, lineage (Madzo et al., 2014). Lastly, in CMML cells, *TET2* mutation-induced cytosine modification changes, especially hyper-modifications, are enriched in enhancer regions and negatively correlate with gene expression changes (Meldi et al., 2015; Yamazaki et al., 2015). Note that these studies used bisulfite sequencing based methods to characterize genome-wide cytosine modification changes, which cannot distinguish 5-mC from 5-hmC. Using the modification status from as few as 14 sites, a support vector machine model was able to achieve 79% accuracy in predicting patient response to decitabine, a hypomethylating agent commonly used to treat CMML (Meldi et al., 2015). This result suggests that the changes in cytosine modifications have prognostic value to risk-stratify patients at diagnosis.

Erythropoiesis as a model to study the role and regulation of 5-hmC

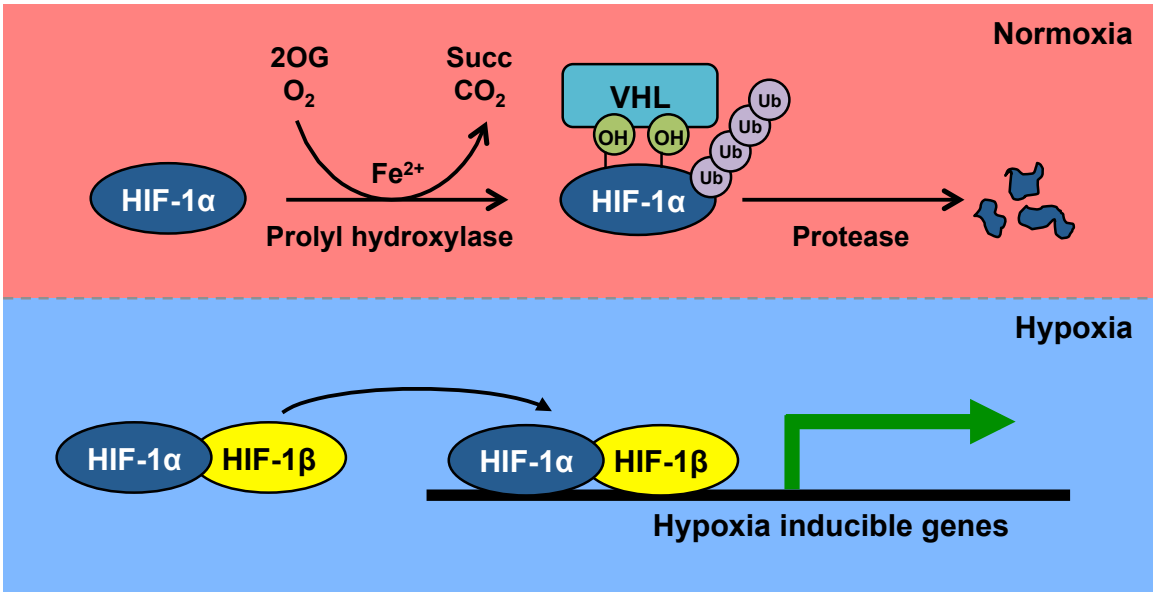
In order to study how 5-hmC and other epigenetic changes regulate chromatin structure and consequent gene expression programs, we reasoned that we would need to examine a system in which cells underwent a functional change that required the expression of a new gene expression program. For example, we studied how 5-hmC distribution changes as neuroblastoma cells grow in normoxia versus hypoxia (Mariani et al., 2014) and as colon cancer cells grow as single cells or linked by tight junctions (Chapman et al., 2015). The third system that we studied was one of the most well-understood differentiation programs- that of human erythropoiesis, the process by which hematopoietic stem and progenitor cells (HSPCs) differentiate into mature red blood cells. By studying the roles and regulation of 5-hmC in normal erythropoiesis, it may be possible to discover novel molecular mechanisms governing the normal differentiation process and identify druggable targets to enhance erythropoiesis in hematopoietic malignancies and other erythropoiesis-related diseases.

Role of hypoxia in erythropoiesis

In mammals, the primary driver for erythropoiesis is hypoxia. The cellular hypoxia response is primarily mediated through oxygen-sensitive HIF-1/2 proteins, which are master TFs that direct the expression of hypoxic gene programs. HIF-1/2 are dimers consisting of one α-subunit (HIF-1α or HIF-2α) and one β-subunit (ARNT/HIF-1β, shared between both HIF-1/2 proteins). Although both the α and β subunits are constitutively produced, the α subunits are quickly degraded under normoxic oxygen levels (Ivan and Kaelin, 2017). This is achieved through the oxygen-dependent

hydroxylation of two proline residues on the α subunits by PHDs. The hydroxyproline residues are recognized by Von Hippel-Lindau (VHL), an E3 ubiquitin ligase, which ubiquitinates the α subunit, tagging it for degradation (Figure 1.8) (Ivan and Kaelin, 2017). At lower oxygen levels, PHDs can no longer catalyze the hydroxylation reaction, and HIF-1/2α subunits quickly accumulate and dimerize with the β subunits, to form the active HIF TFs that then bind hypoxia responsive regions (HREs) in the DNA to drive expression changes in response to hypoxia (Figure 1.8) (Ivan and Kaelin, 2017).

Figure 1.8 Mechanism of oxygen sensing by HIF-1



In normoxia (red background), two prolines on HIF-1α are hydroxylated by prolyl hydroxylases (PHDs). VHL binds to the hydroxyprolines (green circles) and poly-ubiquitinates HIF-1α , targeting it for degradation. In hypoxia (blue background), the oxygen level is too low for PHDs to function efficiently, leading to HIF-1α stabilization, which dimerizes with HIF-1β to activate hypoxia inducible genes.

When an organism experiences hypoxia, (*e.g.,* moving to higher altitude or acute blood loss), HIF-2 is stabilized in kidney interstitial fibroblasts and drives the production

of the hormone EPO (Paliege et al., 2010; Safran and Kaelin, 2003). EPO is released into the bloodstream, where it travels to the bone marrow to stimulate the production of red blood cells. EPO-stimulated HSPCs then differentiate along the erythroid lineage through a series of intermediate stages, in which cells accumulate hemoglobin and eventually enucleate to become reticulocytes and eventually mature erythrocytes (Figure 1.8).

Despite this well-established connection between physiological hypoxia and erythropoiesis, little was known about the effects of hypoxia on the HSPCs themselves. The bone marrow is one of the most hypoxic microenvironments in mammals, where the HSPC niche outside the vasculature experiences about 1% oxygen (Nombela-Arrieta et al., 2013; Spencer et al., 2014). Previous studies by the Godley Laboratory have shown the importance of TET2 and 5-hmC in maintaining normal erythropoiesis, but how hypoxia affects the 5-hmC landscape remain unknown (Madzo et al., 2014). Considering the physiological environment in which the HSPCs reside, it is important to incorporate hypoxia to describe accurately the evolution of the epigenomic landscape in erythropoiesis.

Developmental globin switching in erythropoiesis

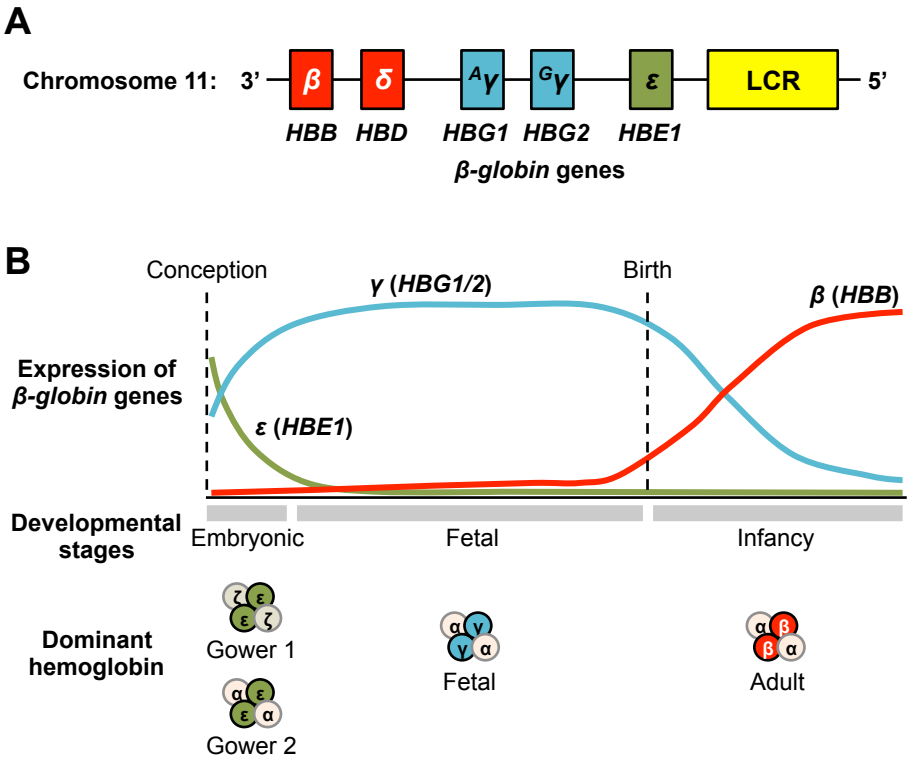
Mammalian hemoglobin is a tetramer consisting of two α-like globin and two β-like globin subunits. The genes encoding the β-like globins are located in a gene cluster on chromosome 11p15.4 (*β-globin* locus), consisting of five genes: 5'-*HBE1* (*ε*)-*HBG2* (*Gγ*)-*HBG1* (*Aγ*)-*HBD* (*δ*)-*HBB* (*β*)-3' (Figure 1.9A). These genes are expressed in distinct developmental stages, with transitions controlled by a series of epigenetic

switches. The *ε-globin* gene *HBE1* is expressed in early embryogenesis during primitive erythropoiesis, and ε-globin forms the embryonic hemoglobins Gower 1 ($\zeta_2\varepsilon_2$) and Gower 2 ($\alpha_2\varepsilon_2$) with two α-like globins (Paikari and Sheehan, 2018; Sankaran and Orkin, 2013; Wilber et al., 2011) (Figure 1.9B). The expression of *γ-globin* genes (*HBG1* and *HBG2*, abbreviated *HBG1/2*) surpasses *HBE1* expression during the primitive to definitive erythropoiesis switch at the early fetal stage and is maintained throughout gestation (Ley et al., 1989; Peschle et al., 1985; Sankaran and Orkin, 2013; Sankaran et al., 2010) (Figure 1.9B). The resulting γ-globins are subunits of the fetal hemoglobin (HbF, $\alpha_2\gamma_2$) (Sankaran and Orkin, 2013; Sankaran et al., 2010). The fetal to adult hemoglobin switch starts around birth and is complete typically within six months after birth, when *HBB* becomes the predominant transcript from the *β-globin* locus, leading to adult hemoglobin (HbA, $\alpha_2\beta_2$; Figure 1.9B) (Sankaran and Orkin, 2013; Sankaran et al., 2010; Stamatoyannopoulos, 2005). The adult *δ-globin* gene *HBD* is expressed at a low level compared to *HBB*, such that HbA_2 ($\alpha_2\delta_2$) accounts for only a small fraction of adult hemoglobin (Sankaran et al., 2010). Importantly, the fetal *γ-globin* genes remain expressed at a low level throughout life (Sankaran and Orkin, 2013; Sankaran et al., 2010; Thein and Menzel, 2009), prompting exploration of therapeutic strategies to re-activate their full expression as a means of treating β-globinopathies, such as sickle cell anemia (SCA) and β-thalassemia.

Expression of genes in the *β-globin* locus is controlled by the interplay between local chromatin structure and erythroid-specific TFs, which are themselves regulated at both transcriptional and post-translational levels (Wilber et al., 2011; Zheng et al., 2014). Locally, expression of *β-globin* genes is influenced by the spatial proximity of

their promoters to the locus control region (LCR), an enhancer-rich region 5' of the *β-globin* gene cluster (Bender et al., 2000; Forrester et al., 1986; Sankaran et al., 2010) (Figure 1.9A). Previous studies have shown that TF binding in the LCR promotes looping and brings the enhancer in contact with promoters of *β-globin* genes (Wilber et al., 2011). In particular, the erythroid-lineage TFs GATA1, KLF1, and IKZF1/IKAROS promote looping of the LCR to the promoter of *HBB* (Kang et al., 2017; Kang et al., 2015b; Keys et al., 2008; Vakoc et al., 2005). Notably, BET bromodomain proteins BRD2, BRD3, and BRD4 bind to acetylated GATA1 and facilitate its binding to erythroid genes (Gamsjaeger et al., 2011; Lamonica et al., 2011; Stonestrom et al., 2015). Moreover, expression of *HBG1/2* is promoted by the non-coding *BGLT3* lncRNA, which facilitates the LCR-*HBG1/2* contact (Ivaldi et al., 2018). Finally, forced looping of the LCR to *HBG1/2* via artificial zinc finger proteins leads to reactivation of HbF in adult CD34+ hematopoietic stem cells (HSCs) during erythropoiesis (Deng et al., 2012; Deng et al., 2014), emphasizing that chromatin structure is a major regulator in *β-globin* gene switches.

Figure 1.9 *β-globin* gene cluster and developmental expression of *β-globins*



(**A**) Structure of the *β-globin* gene cluster on chromosome 11. Labels underneath the genes are the official gene symbols. Greek letters overlaying each gene is the name of the globin it encodes. All genes in the cluster are on the Crick strand as indicated by the 5' ➔ 3' direction. Transcription of genes are controlled by enhancers in the locus control region (LCR, yellow bar) (**B**) Expression switches of the *β-globin* genes during embryonic and fetal development. *HBE1* is expressed at embryonic stage, and is replaced by *HBG1/2* at fetal stage. After birth, *HBB* expression overtakes *HBG1/2*. *HBD* expression is low throughout development and not represented on the graph.

In addition to chromatin looping, the expression of *fetal-* and *embryonic-globin* genes is suppressed by the direct binding of a number of repressive TFs. In addition to its role in activating adult *β-globin* transcription, GATA1 also exhibits inhibitory effects, forming an inhibitory complex with BCL11A, COUP-TF2, and NuRD to inhibit fetal and embryonic *β-globin* gene expression (Aerbajinai et al., 2009; Avram et al., 2000; Chan et al., 2013;

Sankaran et al., 2010; Xu et al., 2013). *BCL11A* expression is activated by KLF1, such that lower levels of KLF1 lead to decreased BCL11A levels and hereditary persistence of fetal hemoglobin (HPFH) (Borg et al., 2010; Zhou et al., 2010). Furthermore, various genetics studies of HPFH have established MYB as a fetal β-globin inhibitor (Sankaran and Orkin, 2013; Stadhouders et al., 2014; Wang et al., 2018), with *MYB* mRNA targeted for degradation directly by the microRNAs (miRs) *miR-15A* and *miR-16-1* (Sankaran et al., 2011).

Hypothesis and specific aims

To investigate the how erythropoiesis is regulation via 5-hmC and chromatin looping, I proposed two Specific Aims:

The first Specific Aim was to characterize 5-hmC differences under hypoxic versus normoxic erythropoiesis. I hypothesized that due to the lack of available oxygen as substrate, TET enzymes would be less efficient and result in less global 5-hmC. I also expected a shift in 5-hmC distribution across the genome as a result of the TET activity change, similar to the *TET2*-mutated cells previously reported (Madzo et al., 2014). I was also interested in characterizing HIF-1α binding sites in erythroid cell lines and determining whether known epigenetic regulators are regulated directly by HIF-1.

The second Specific Aim was to investigate how epigenetic modifications can influence *β-globin* gene switches. I hypothesized that by targeting the molecular machinery mediating the looping of the LCR to the *β-globin* genes, fetal and embryonic *β*-globin genes could be reactivated. Specifically, since BET proteins had been shown to facilitate LCR looping (Gamsjaeger et al., 2011; Lamonica et al., 2011; Stonestrom et

al., 2015), I expected that BET inhibitors would decrease *HBB* expression and increase *HBG1/2* and/or *HBE1* expression. Given the abundance of BET inhibitors in clinical trials for various hematological malignancies (Alqahtani et al., 2019; Stathis and Bertoni, 2018), I expect that BET inhibitors have the potential to become a new and affordable option for treatment of β-globinopathies, like SCA and/or β-thalassemia.

CHAPTER II

Methods

Cell culture and *in vitro* human erythroid differentiation

K562, HEL, and HL-60 cells were cultured in RPMI 1640 medium supplemented with 10% FBS. TF-1 cells were cultured in RPMI 1640 medium supplemented with 10% FBS and 4ng/mL GM-CSF. HEK 293T cells were cultured in DMEM medium with 10% FBS. K562 erythroid differentiation was performed by treating the cells with 1 mM sodium butyrate for up to 5 days. TF-1 erythroid differentiation was performed by first depriving the cells of GM-CSF for 24 hours, followed by stimulating with 3 units/mL EPO.

The *in vitro* erythroid differentiation of primary human HSPCs was described in Kang et al (Kang et al., 2008) and Madzo et al (Madzo et al., 2014) and was performed by Hui Liu from the Wickrema laboratory. Primary human erythroblasts were generated by culturing CD34+ early HSPCs initially isolated from growth factor-mobilized peripheral blood (purchased from ALL Cells, Inc.) using an Isolex 300i cell selection device. The culture contained 15% fetal bovine serum, 15% human serum, Iscove modified Dulbecco's medium (IMDM), 10 ng/mL interleukin-3, 2 units/mL EPO, and 50 ng/mL SCF. During the initial 8 days of culture, cells were fed on days 3 and 6 by adding equal volumes of fresh culture media supplemented with growth factors. However, no new interleukin-3 was added after the initial addition on day 0, and the amount of SCF added to the fresh media was gradually decreased at each feeding (day 3, 25 ng/mL; day 6, 10 ng/mL; day 8, 2 ng/mL). The amount of EPO added was 2 units/mL during each feeding until day 8 of culture. Cells were fed one more time on day

10 of culture by adding equal volumes of fresh media with only EPO (2 units/mL) during this final feeding. Cells were collected at various time points during the culture and stained with benzidine and hematoxylin, and flow cytometry was done with antibodies directed against CD71 and Glycophorin A to monitor the differentiation program.

Hypoxia and BETi treatment

Hypoxia samples were cultured with 1% O_2 and 5% CO_2 in a hypoxia chamber, maintained at 37°C, whereas normoxia samples were cultured with 21% O_2 and 5% CO_2 at 37°C. All additional treatments or collection of hypoxic cells were performed within a hypoxic glove box to avoid exposure to oxygen.

BETi concentrations were determined based on published data and a preliminary dose response curve. JQ1 (Sigma-Aldrich, SML1524), CPI-0610 (Selleck Chemicals, S7853), and PLX51107 (MedChemExpress, HY-111422) were dissolved in DMSO to 10mM and stored at -80°C for future use. For K562, HEL, and HL-60 cells, a final concentration of 200nM was used for the treatment groups, with equal amount of DMSO (0.002%) added to the control group. TF-1 cells were first washed four times in RPMI 1640 without FBS and cultured for 24 hours in RPMI 1640 with 10% FBS, but without GM-CSF. The cells were then separated into four treatment groups: GM-CSF (4ng/mL) + DMSO (0.002%, control), GM-CSF (4ng/mL) + BETi (200nM), EPO (3 units/mL) + DMSO (0.002%), and EPO (3 units/mL) + BETi (200nM). JQ1 and PLX51107 were added to the cell cultures on day 0 without additional supplementation up to day 5. CPI-0610 was re-applied to the cells on a daily basis.

Mass spectrometry quantification of 5-mC and 5-hmC

To measure total levels of 5-mC and 5-hmC as a fraction to total cytosine, DNA was purified from samples using phenol-chloroform extraction. Two micrograms of the purified DNA was then hydrolyzed to single nucleosides by nuclease P1 (Sigma-Aldrich N8630), phosphodiesterase I (Sigma-Aldrich P3243), and calf intestine alkaline phosphatase (Thermo Fisher Scientific 18009027). Hydrolyzed nucleosides were separated by an Acquility UPLC Oligonucleotide BEH C18 Column (Waters, 186003950) in an Agilent 1290 liquid chromatography system, before being measured by Agilent 6460 Triple-Quadrupole tandem mass spectrometry. The mass spectrometry protocol was optimized to detect deoxyadenosine (dA), thymidine (T), deoxyguanosine (dG), deoxycytidine (dC), 5-methyl-deoxycytidine (5mdC), and 5-hydroxymethyl-deoxycytidine (5hmdC). Standards containing mixtures of all detection targets in different known ratios were run alongside the samples to correct for detection bias. The total cytosine level was estimated by the total dG signal, since cytosine and guanine bases exist in equal amounts in the genome. 5-mC and 5-hmC levels were calculated as a percentage of total cytosine from the output signals, which were then normalized to the known standards.

5-hmC chemical labeling and pull-down (hMe-SEAL)

5-hmC pull-down and sequencing were performed as previously described (Song et al., 2011). Briefly, genomic DNA was sheared to ~300bp, and T4 β-glucosyltransferase was used to label 5-hmC with azide-modified 6-UDP-glucose. Biotin-S-S-DBCO was then covalently attached to the modified glucose via a Click

reaction with the azide group. Labeled DNA fragments were then pulled-down using streptavidin beads and purified for next-generation sequencing along with input DNA controls.

HIF-1α chromatin immunoprecipitation (ChIP)

K562 cells cultured in normoxia or hypoxia (1% O_2) for 72 hours were fixed by 1% formaldehyde for 10 minutes. Nuclear extraction and sonication of the cross-linked chromatin were performed using the Covaris truChIP kit (Covaris, 520154) and the Covaris S220 sonicator, respectively. Sonication of samples was done with 140W peak incidence power using 12 million cells/mL, in 20×45s cycles with 30s break between each cycle. Sonication duty factors were optimized separately for normoxic and hypoxic samples to achieve comparable DNA fragment sizes: 14% for normoxic samples and 18% for hypoxic samples. HIF-1α bound DNA was enriched with rabbit anti HIF-1α antibody (Abcam, ab2185) at 5 µg antibody per 1 mL sample and purified using Genelute PCR clean-up kit (Sigma-Aldrich, NA1020-1KT) for next-generation sequencing. The success of HIF-1α ChIP was validated by quantitative PCR (qPCR) of a known HIF-1 binding site in the *ENO1* promoter (Semenza et al., 1996). An amplicon in *VEGFA* intron without HIF-1α motif was selected to serve as the negative control. See Table 2.1 for primers used for in qPCR evaluations.

Table 2.1 ChIP primers

Name	Sequence (5'--->3')
ENO1 promoter F (positive control)	GTGAGCCGAACTGGGGTG
ENO1 promoter R (positive control)	ACTCGGAGTACGTGACGGAG
VEGFA intron F (negative control)	ACTGTCTCTACCCTGGTCTCC
VEGFA intron R (negative control)	CCCTGCCAGCCACTGATAAC
TET3 site 1 F	GGTAGGAGGTGTGGGTGATG
TET3 site 1 R	GAGGCAGGGCATTCTCACAA
TET3 site 2 F	TGCCCTCTCTGTGTTTAGTGTC
TET3 site 2 R	TACGTGACTCTGATGGAACCCT

hMe-SEAL and ChIP sequencing read mapping and peak calling

The next generation sequencing (single-end, 50 bp) for both hMe-SEAL and ChIPed DNA was performed on an Illumina HiSeq 4000 instrument by the University of Chicago Functional Genomics Facility. Raw sequences in fastq format were aligned to the hg19 reference genome using the BWA-MEM algorithm (Li and Durbin, 2009). 5-hmC peaks were called using MACS2 (Zhang et al., 2008) with input alignment as control.

Differential peak calling and genomic localization analysis

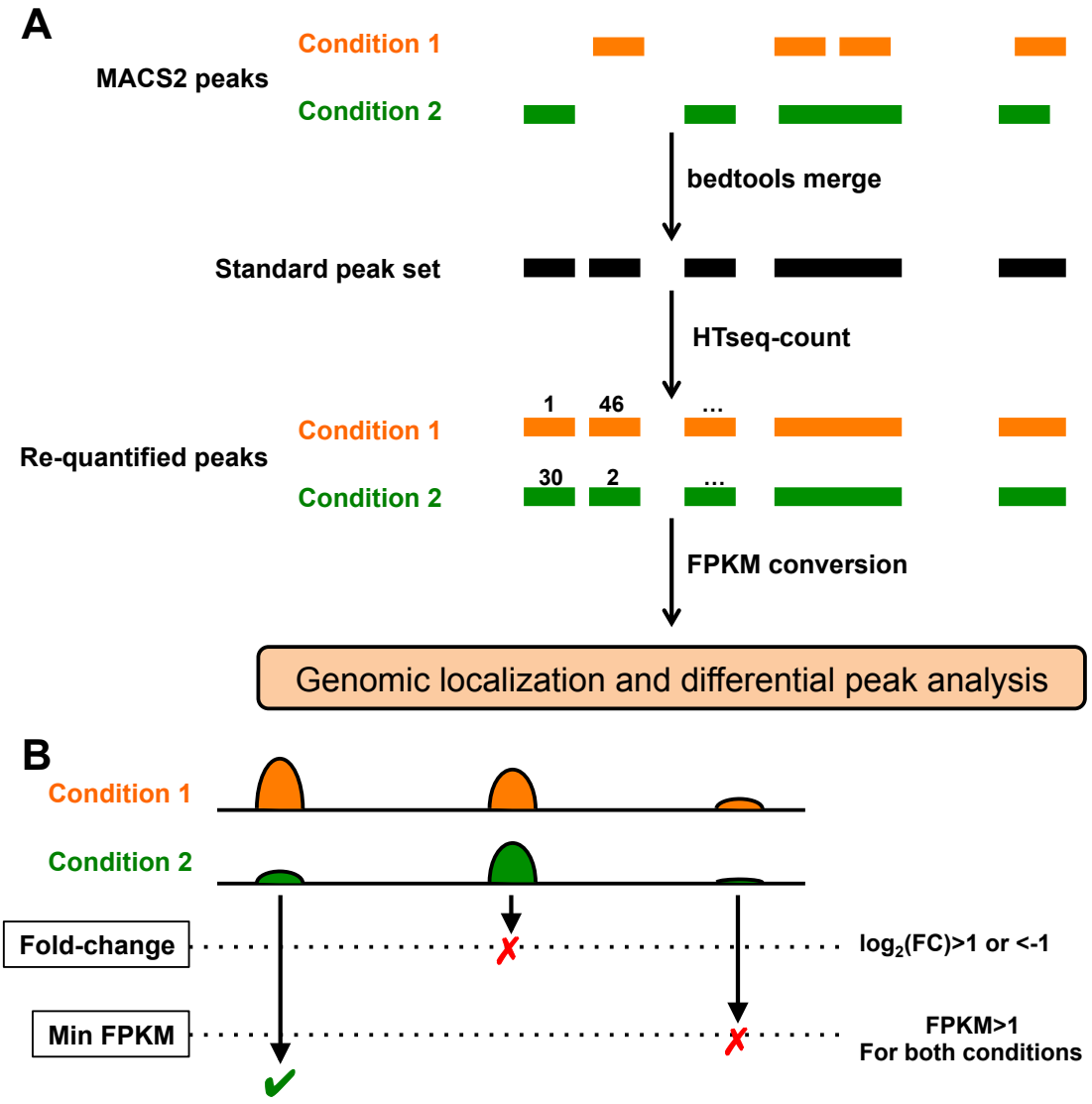
In order to compare 5-hmC peaks in different conditions, a unified peak set was made by merging peaks in different conditions with bedtools merge (Quinlan and Hall, 2010). 5-hmC densities in peak regions were then quantified by counting aligned reads in each peak region with HTseq-count (Anders et al., 2015) and converting to FPKM (Figure 2.1A). The required .gtf annotation for HTseq-count was generated from the merged .bed file using the following Unix command:

```
awk 'BEGIN{OFS="\t "};{print $1,"gene","exon",$2,$3,".", ".",".","gene_id
\x22"NR"\x22"}' merged_peaks.bed > peak_annotation.gtf
```

Once a unified peak set was established and quantified for all samples, differential peaks were called using the following filters: (1) at least one of the samples has FPKM greater than 1, and (2) the fold-change in either direction must be greater than 2 (Figure 2.1B).

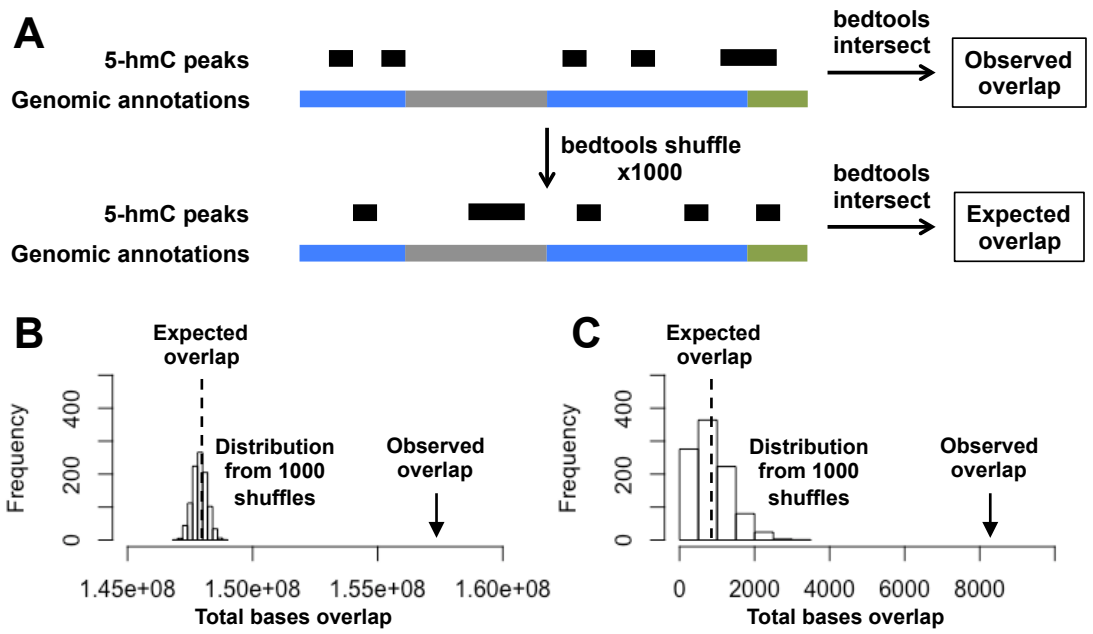
Genomic distribution of 5-hmC peaks were examined by calculating enrichment of 5-hmC peaks in particular genomic elements. Enrichment was defined as *(Observed overlap)/(Expected overlap)*. Expected overlaps were estimated by using bedtools shuffle to redistribute the peaks randomly across the genome 1000 times, each time quantifying the total overlaps with each genomic element (Figure 2.2A). The option "–noOverlapping" was used to ensure the shuffled peaks do not occupy the same genomic space. Assuming that the distribution of the shuffled results are roughly normal, the mean value from the 1000 bedtools shuffle defines the expected overlap, whereas the standard deviation of the results were used to estimate the statistical significance of the enrichment using the *Z*-test. If the distribution was not normal, the shuffled value closest to the observed value was used to estimate the p-value by ranking, e.g. $p \approx 0.005$ if the observed value is closest to the 5^{th} highest shuffled value (Figure 2.2B).

Figure 2.1 Generating standard peak set and differential peaks calling.



(**A**) Peaks from two conditions are merged to generate the standard peak set. Numbers of reads within each standard peak region are quantified with HTseq-count, using respective alignment files as input. The read counts are then normalized by total aligned reads and the length of each peak to FPKM values for further analysis. (**B**) Filters used to call differential peaks. Fold-change filter removes peaks that don't experience much increase or decrease in 5-hmC density. FPKM filter removes peaks with very low density, which are easily affected by noises in sequencing.

Figure 2.2 Localization analysis for peak enrichment



(**A**) The observed overlap between 5-hmC peaks and different genomic elements (represented as different color segments) are calculated as total base pairs overlap across the genome. To generate an expected overlap, the 5-hmC peaks are randomly shuffled 1000 times producing 1000 randomly overlaps for for each genomic element. Enrichment is calculated by *Observed overlap/ Expected overlap*. (**B**) Example of a shuffle distribution that is roughly normal. P value of the observed overlap is calculated with Z-test using mean (expected overlap) and standard deviation from the 1000 shuffled values. (**C**) Example of a non-normal distribution from shuffles. P value is estimated by the relative ranking of the shuffled value closest to the observed overlap. In this example, P<0.001 since the observed value lies outside the range of the shuffled values.

RNA-sequencing (RNA-seq) and Library Preparation

RNA was extracted from cell pellets with RNAzol (Sigma-Aldrich, R4533) following the manufacturer's instructions. Sample RNA libraries were prepared using the KAPA mRNA HyperPrep Kit (Roche, KK8580) with 1μg input RNA and 2 μL of 1:100 diluted ERCC spike-in control (Fisher Scientific, 4456740), and a mixture of all sample libraries at 1nM each was submitted to the University of Chicago Functional Genomics Facility for sequencing. Single-end sequencing (50bp) on HiSeq 4000 (Illumina) was

performed for the hypoxia project (Chapter III) and K562 JQ1 treated samples (Chapter IV). Paired-end sequencing (100bp) on NovaSeq (Illumina) was performed for the TF-1 JQ1 treated samples (Chapter IV).

RNA-seq Alignment and Quantification

For RNA-seq done in the hypoxia project (Chapter III), raw reads in fastq format were aligned to the hg19 reference genome using Tophat2 (Kim et al., 2013) with GENCODE gene and transcript annotation (GRCh37.p13) (Frankish et al., 2019). Raw reads from the BETi project (Chapter IV) were aligned to hg19 reference genome using HISAT2 (Kim et al., 2019) with the same GENCODE annotation (GRCh37.p13) (Frankish et al., 2019). The alignment tool change in between these two projects was due to the significantly faster alignment speed of HISAT2 compared to Tophat2. Gene expression was quantified and compared by tools in the Cufflinks package (Trapnell et al., 2012).

Genomic data visualization

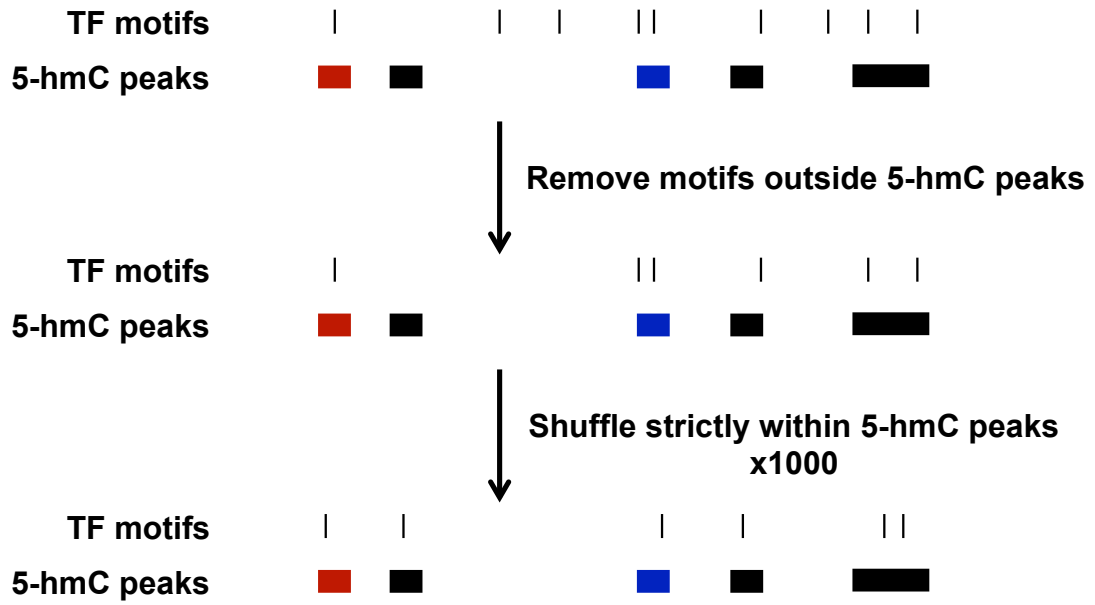
Visualization of 5-hmC distribution, ChIP-seq peaks, and RNA-seq results was done using Integrative Genomics Viewer (IGV) (Robinson et al., 2011). Read distribution visualization (tiled data file, .tdf) was created from the .bam alignment files using igvtools count. Peak calling results (.bed files) were used to indicate location of peaks.

Transcription factor motif enrichment analysis

Known TF motifs were obtained from the HOMER database (Heinz et al., 2010). The list of TFs was filtered by expression data from RNA-seq, and the ones with low expression (FPKM<1) were dropped for subsequent analysis. Potential TF binding sites based on motifs were predicted using scanMotifGenomeWide.pl from the HOMER package.

To test if certain TFs were enriched or depleted in particular genomic regions, sets of background and target regions were defined. For example, to calculate enrichment in differential 5-hmC peaks, the differential 5-hmC peaks were used as target regions, while all 5-hmC peaks were used as background. All TF motifs outside of the background regions were dropped for the analysis (Figure 2.3). Motif enrichment was defined as *(Observed # of motifs)/(Expected # of motifs)*. Expected # of motifs were estimated by using bedtools shuffle to redistribute the motifs randomly within the background regions 1000 times, each time quantifying number of motifs within the target regions (Figure 2.3). The standard deviations of the results were used to estimate the statistical significance of the enrichment.

Figure 2.3 TF- motif enrichment in differential 5-hmC peaks



Modified shuffle procedure to analyze TF motif enrichment in differential 5-hmC peaks relative to all 5-hmC peaks. TF motifs outside 5-hmC peaks are ignored since they are irrelevant to the goal of the analysis. With the remaining peaks, shuffling was restricted within all 5-hmC peaks. This is to avoid losing peaks to the much larger non-peak regions. The numbers of motifs within increased/ decreased 5-hmC peaks are counted with each shuffle. The resulting distribution of the 1000 shuffles are used to generate enrichment score and P values as described previously. Red and blue segments represent differential 5-hmC peaks.

Bioinformatics script depository

Example scripts used in the analyses mentioned above are deposited at my GitHub Thesis Library: https://github.com/johnzcao/Thesis-library.

Quantitative real-time PCR (qPCR)

cDNA libraries were made from 1μg input RNA using High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher, 4368814). Gene expressions were quantified

using Power SYBR Green PCR Master Mix (Thermo Fisher, 4367659) using 18S rRNA as the housekeeping gene (Table 2.2).

Table 2.2 qPCR primers

Name	Sequence (5'--->3')
TET3 F	CGTCGAACAAATAGTGGAGA
TET3 R	CTTTCCCCTTCTCTCCATAC
VEGFA F	AGGAGGAGGGCAGAATCATCA
VEGFA R	ATGTCCACCAGGGTCTCGATTG
HBB F	GGCTCACCTGGACAACCTCA
HBB R	AAAGTGATGGGCCAGCACACAG
*HBG*1/2 F	AAAGCACCTGGATGATCTCA
*HBG*1/2 R	AAAACGGTCACCAGCACATTT
*HBE*1 F	CAAGCCCGCCTTTGCTAAGCT
*HBE*1 R	CTCCTTGCCAAAGTGAGTAGCCAG
MYB F	TCCCAAGTCTGGAAAGCGTC
MYB R	GCACATCTGTTCGATTCGGG
NR2F2 F	GCCATAGTCCTGTTCACCTCA
NR2F2 R	GAATCTCGTCGGCTGGTTG
IKZF1 F	CACCCGAGGATCAGTCTTGG
IKZF1 R	CATGTCTTGACCCTCATCAGCAT
KLF1 F	GGACACACAGGATGACTTCC
KLF1 R	GCTGGTCCTCAGACTTCAC
GATA1 F	GATGAATGGGCAGAACAGGC
GATA1 R	CAGTGTCGTGGTGGTCGT
BCL11A F	TGCCCCAAACAGGAACACAT
BCL11A R	ATTCTGCACTCATCCCAGGC
BGLT3 F	TCACTGGTACGCAGGGTTTT
BGLT3 R	TATTGAGTTGTGGGGACTGGC
18S F	GAGGGAGCCTGAGAAACGG
18S R	GTCGGGAGTGGGTAATTTGC

CRISPR-Cas9 guided deletion of HIF-1α binding sites

CRISPR guides were designed using the online design tool at crispr.mit.edu so that the cut site was within 10 bps of the HIF-1α binding sites. CRISPR guide sequences were inserted to the lentiCRISPR v2 plasmid (Addgene #52961) according to the associated protocol (Sanjana et al., 2014). Lentivirus was produced in HEK 293T

cells to transduce K562 cells. Cells that were successfully transduced were selected via puromycin resistance to the drug at 12μg/mL after selection for 3 days. To obtain single-cell clones from the mixed population, cells post-selection were diluted to 2.5 cells/mL and deposited into 96-well plates at 100 μL per well to achieve a cell dose of 0.25 cells/well. After clones were identified in the plates, each clone was expanded for DNA extraction. Most clones were expected to be grown from a single starting cell, as the chances for more than one cell in a single well is 6.25% (25%×25%) or less. PCR primers were designed to amplify regions containing the binding sites (Table 2.3). PCR products of the targeted region from each clone were run on agarose gels to detect shift in sizes, and the clones with smaller PCR product sizes were sent for Sanger sequencing. Remaining mixed clones were identified at this step if any bases show signals for more than 2 nucleotides. Clones that lost the core "ACGT" motif on both alleles of *TET3* were expanded further and frozen viably for future use. Site 1/Site 2 double deletion clones were made from validated single deletion clones by targeting the intact binding site using CRISPR-Cas9. See Figure 3.9 for the genomic positions of the two sites.

Table 2.3 CRISPR clone screening primers

Amplicon	Primer name	Sequence (5'--->3')
TET3 Site 1	TET3_S1_F	TTTTGAGGGATTGGGGGCTT
TET3 Site 1	TET3_S1_R	CAGCACACAAGAACCAGGTC
TET3 Site 2	TET3_S2_F	TGCTGTTCATGCTTGAGGGA
TET3 Site 2	TET3_S2_R	GCAAATCTGTCAGTGGCTGG

Cytospin and benzidine-hematoxylin staining of hemoglobin producing cells

Cells were counted, and 150,000 cells were collected by spinning at 400g for 5 minutes. The culture media was discarded, and the cells were re-suspended in 100 μL

of PBS and spun onto glass slides at 600 rpm. The slides were dried for ~20 minutes before staining with a 0.2% benzidine solution in 0.5 M acetic acid for 4 minutes, followed by 2 minutes in 2% H_2O_2 in 70% ethanol, and 3 minutes washing in running tap water. After the wash, the slides were further stained in 1x hematoxylin solution (Astral Diagnostics, 7012) for 1 minute, followed by another 3 minutes wash in running tap water. The slides were then air-dried and mounted with cover slips.

Microscopy and quantification of benzidine staining

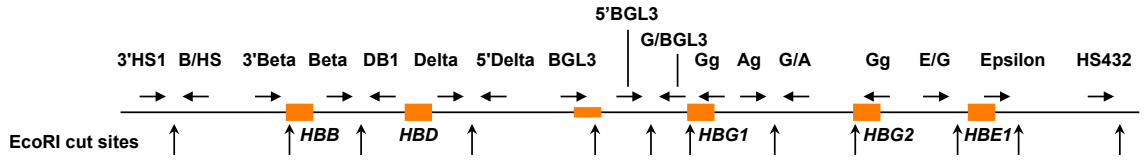
At least three randomly selected fields of stained cells were imaged at 40x magnification. Images were processed using the Color Deconvolution plugin in ImageJ software to separate hematoxylin and benzidine staining into different channels (Ruifrok and Johnston, 2001). The numbers of benzidine-positive cells and the intensity of benzidine staining were quantified in the benzidine channel, whereas the total cell number and abnormal nuclear morphology were quantified in the hematoxylin channel. Results of all fields from each slide were summed for the final quantification.

Chromatin Conformation Capture (3C)-qPCR quantification of Locus Control Region (LCR) interaction frequency

The method for 3C-qPCR was adapted from Hagège et al (Hagège et al., 2007). Briefly, 10 million cells were collected and fixed with 1% formaldehyde for 5 minutes in PBS with 10% FBS, then quenched by 124 mM glycine on ice. The fixed cells were lysed and digested by 400 units of EcoRI (NEB, R3101). 20 µL of the samples were taken before and after the digestion as digestion controls. Following the digestion, the

samples were diluted and ligated by T4 DNA ligase (NEB, M0202). The samples were then reverse-crosslinked by overnight proteinase K digestion, and the DNA was purified by phenol-chloroform extraction. The digestion efficiency, sample purity, and DNA concentration of each sample were assessed by qPCR to exclude low-quality samples prior to the 3C quantification. We used previously published 3C primers (Kiefer et al., 2011; Morgan et al., 2017), listed in Table 2.4 (Figure 2.4). Primer pairings for all processes in 3C-qPCR are listed in Table 2.5. The calculated quantity of each primer pair was normalized as number of templates per million B13 control site template. Additionally, because the Gg primer anneals to both *HBG1* and *HBG2* fragments, whereas the Ag primer anneals only to the *HBG2* fragment, we subtracted the Gg primer signal by the Ag primer signal to quantify the *HBG1* interaction frequency. All 3C-qPCR was performed using Power SYBR Green PCR Master Mix (ThermoFisher, 4367659).

Figure 2.4 Schematic of 3C primers and EcoRI cut sites



Positions of the 3C primers relative to the *β-globin* genes and EcoRI cut sites (not to scale). Orange blocks represent the *β-globin* genes. Horizontal arrows represent 3C primers with their respective orientation on the template. Vertical arrows denote the locations of EcoRI restriction sites. All primers are within 50bp of and amplify toward an EcoRI restriction site. Several EcoRI sites with no associated primers are not shown for clarity purpose.

Table 2.4 3C-qPCR primers

Name	Sequence (5'--->3')
3'HS1	ATTCCCGTTTTTATGAAATCAACTTT
B/HS	TCTTAGAAAGCCTTTACAATTTCCTTTATC
3'Beta	AGCTTAGTGATACTTGTGGGCCA
Beta	GCTCGGCACATGTCCCATCCAG
DB1	GTCAGTGAGTCTAGGCAAGATGTTGGCC
Delta	AAAAAATGTGGAATTAGACCCAGGAATG
5'Delta	GGGTGTGTATTTGTCTGCCA
BGL3	TTGCCATACCTCATATCCTTAG
5'BGL3	CTTAGGCATCCACAAGGG
G/BGL3	AGCAAGGATGGTTCTTAAGGAAGGG
Ag	ATCCATGATCTCTAACCTTGC
G/A	AATTTGAAGATACAGCTTGCCTCCGATAAG
Gg	GGGTTCATCTTTATTGTCTCCT
E/G	CCACCCCGATAAAGATTTTTCTCCATCA
Epsilon	ATTAACCAATGGTATCTTTCTGAGCA
HS432	CCAAATGGGTGACTGTAGGGTTGAGA
B13 F	CGTGAGAGCATACTTCCTGGTTC
B13 R	ACACCAGAGAGGTCTTGCCCT

Table 2.5 3C-qPCR primer pairs

Name	Forward	Reverse	Procedures Used
No digestion control	B13 F	B13 R	Digestion efficiency, purity test, concentration test, 3C-qPCR
3'HS site	3'HS1	B/HS	Digestion efficiency
HBB site	Beta	DB1	Digestion efficiency
HBD site	Delta	5'Delta	Digestion efficiency
BGLT3 site	5'BGL3	G/BGL3	Digestion efficiency
HBG1/2 site	Ag	G/A	Digestion efficiency
3'HS	HS432	3'HS1	3C-qPCR
3'Beta	HS432	3'Beta	3C-qPCR
Beta	HS432	Beta	3C-qPCR
Beta/Delta	HS432	DB1	3C-qPCR
Delta	HS432	Delta	3C-qPCR
5’Delta	HS432	5'Delta	3C-qPCR
BGLT3	HS432	BGL3	3C-qPCR
5'BGLT3	HS432	5'BGL3	3C-qPCR
3'Gamma1	HS432	G/BGL3	3C-qPCR
Gamma1	HS432	Ag	3C-qPCR
3'Gamma2	HS432	G/A	3C-qPCR
Gamma1/2	HS432	Gg	3C-qPCR
3'Epsilon	HS432	E/G	3C-qPCR
Epsilon	HS432	Epsilon	3C-qPCR
BGLT3-3’HS	5’BGL3	3’HS1	3C-qPCR
BGLT3-Gamma1	5’BGL3	Ag	3C-qPCR
BGLT3-3'Gamma2	5’BGL3	G/A	3C-qPCR
BGLT3-Gamma1/2	5’BGL3	Gg	3C-qPCR
BGLT3-3'Epsilon	5’BGL3	E/G	3C-qPCR
BGLT3-Epsilon	5’BGL3	Epsilon	3C-qPCR

CHAPTER III

HIF-1 directly induces TET3 expression to enhance 5-hmC density in erythropoiesis and induce erythroid gene expression in hypoxia

The work presented in this chapter has been published: Cao JZ, Liu H, Wickrema A, Godley LA. HIF-1 directly induces TET3 expression to enhance 5-hmC density and induce erythroid gene expression in hypoxia. Blood advances. 2020;4(13):3053-62. PMID: 32634239

In this study, I designed and performed the experiments (except *in vitro* erythroid differentiation), analyzed the data and wrote the manuscript. Other author contributions include: H. L. performed the *in vitro* erythroid differentiation experiment; L.A.G. conceived of the study and provided insights in experimental design and data interpretation; and A.W. provided primary HSPCs for the differentiation experiments and additional input for experimental design and data interpretation.

Summary

In mammalian cells, cytosines found within CpG dinucleotides can be methylated to 5-mC by DNA methyltransferases and further oxidized by TET enzymes to 5-hmC. The Godley Laboratory has previously shown that HSPCs with *TET2* mutations have aberrant 5-hmC distribution and less erythroid differentiation potential. However, these experiments were performed under standard tissue culture conditions with 21% O_2, whereas HSPCs residing in human bone marrow reside in ~1% O_2. Therefore to model human erythropoiesis more accurately, I compared 5-hmC distribution and gene expression in hypoxia versus normoxic conditions. Despite TET enzymes having limited O_2 as a substrate in hypoxia, 5-hmC peaks were more numerous and pronounced than in normoxia. Among the *TET* genes, *TET3* was upregulated specifically in hypoxia. I identified two HIF-1 binding sites in *TET3* by ChIP of HIF-1α followed by sequencing, and *TET3* upregulation was abrogated with deletion of both of them, indicating that *TET3* is a direct HIF-1 target. Finally, I demonstrated that loss of one or both of these HIF-1 binding sites in K562 cells disrupted erythroid differentiation and lowered cell viability under hypoxia. This work provides a molecular link between oxygen availability, epigenetic modification of chromatin, and erythroid differentiation.

Introduction

5-hmC is an epigenetic mark that regulates chromosome structure and promotes transcription (Cao et al., 2019; Pastor et al., 2013; Wu and Zhang, 2017). The Ten-eleven translocation dioxygenases (TETs) convert 5-mC to 5-hmC in a reaction that requires O_2, Fe(II), and α-ketoglutarate and is facilitated by ascorbate as a co-factor.

The human genome contains three *TET* genes (*TET1-3*) expressed at different tissue and developmental stages. In human hematopoietic cells, both *TET2* and *TET3* are expressed (Yan et al., 2017). Moreover, *TET2* is one of the most frequently somatically mutated genes in a condition now commonly referred to as clonal hematopoiesis, as well as myeloid malignancies, T-cell lymphomas, melanomas, and gliomas (Abdel-Wahab et al., 2009; Bowman and Levine, 2017; Cao et al., 2019; Delhommeau et al., 2009; Langemeijer et al., 2009; Lasho et al., 2018). Previously, we have reported that TET2 is the predominant TET enzyme during erythropoiesis under normoxic conditions, and its activity is augmented by JAK2-mediated phosphorylation (see Chapter V for detailed discussion) (Jeong et al., 2019; Madzo et al., 2014). These studies highlight the importance of 5-hmC in regulating the differentiation of HSPCs down the erythroid lineage.

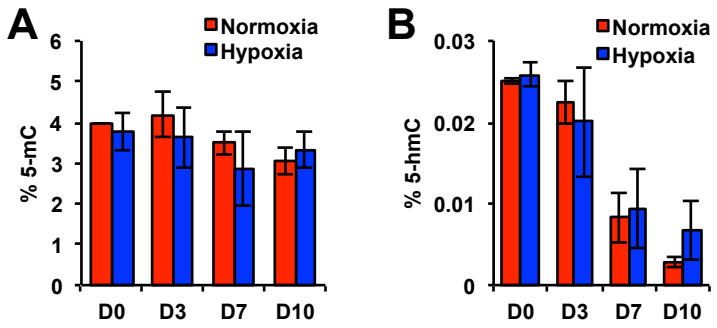
HSPCs reside within the bone marrow niche, which is poorly oxygenated (Nombela-Arrieta et al., 2013; Spencer et al., 2014). In addition, environmental hypoxia is a strong driver of erythropoiesis through stimulating EPO production in kidney interstitial fibroblasts, which then stimulate erythroid differentiation of HSPCs (Ivan and Kaelin, 2017). Therefore, I undertook a study to understand how hypoxia affects 5-hmC distribution and gene expression during erythropoiesis in HSPCs. We expected that hypoxia would lead to decreased global 5-hmC levels and attenuation of 5-hmC peaks compared to normoxia. I expected that changes in 5-hmC distribution together with gene expression changes directed by HIF would promote erythroid differentiation of HSPCs.

Results

Hypoxia promotes 5-hmC accumulation during erythropoiesis

To investigate the effects of hypoxia on 5-hmC distribution and gene expression during erythropoiesis, the Wickrema Laboratory performed their erythroid differentiation protocol on normal human CD34+ HSPCs under parallel normoxic (21% O_2) versus hypoxic (1% O_2) conditions. Samples were collected for DNA and RNA extraction at days 0, 3, 7, and 10 of the differentiation assay. I measured total levels of 5-mC and 5-hmC in genomic DNA using mass spectrometry. No significant differences were found between normoxic and hypoxic samples in 5-mC or 5-hmC levels (Figure 3.1). This observation was contrary to our expectation that a lack of O_2 would lower substrate availability for the 5-mC to 5-hmC conversion and result in a decrease in total 5-hmC levels.

Figure 3.1 Hypoxia does not change global 5-mC and 5-hmC levels



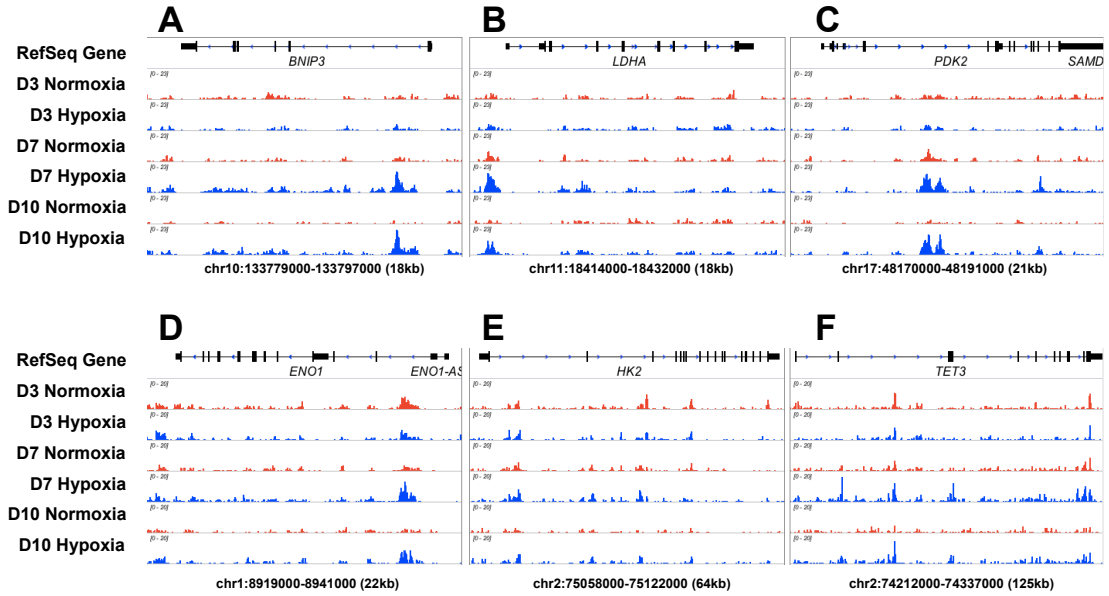
(**A**) Mass spectrometry quantification of 5-mC relative to all cytosine species during erythroid differentiation. (**B**) Mass spectrometry quantification of 5-hmC relative to all cytosine species during erythroid differentiation.

I next examined the genomic distribution of 5-hmC during erythroid differentiation under normoxic versus hypoxic conditions using hMe-SEAL (Song et al., 2011) (GSE40243 [as published in (Madzo et al., 2014)], GSE142870 [Hypoxia data from the

current study (Cao et al., 2020)]). I found that 5-hmC peaks appeared more prominent in hypoxia than in normoxia in multiple hypoxia-responsive genes (Figure 3.2). I then used MACS2 to identify 5-hmC peaks in all samples. Consistent with our previous observations (Madzo et al., 2014), 5-hmC distribution in all samples was enriched in promoters, enhancers, and gene bodies, but was depleted in intergenic regions (Figure 3.3). I found more peaks in hypoxic samples than in normoxic samples, especially at days 7 and 10 (Figure 3.4A). Similarly, peak coverage across the genome, defined as total base pairs covered by peaks in a dataset, was higher in hypoxia by day 10 (Figure 3.4B). I then examined the normalized counts across the genome (Figure 3.4C), and found that although normoxic 5-hmC peaks gradually diminished throughout differentiation, hypoxic 5-hmC peaks remained prominent up to day 10. This trend was observed across the entire genome (Figure 3.4C), resulting in a higher percentage of 5-hmC reads located inside peaks in hypoxic samples at later time points compared to those in normoxic samples (Figure 3.4D). Consistently, overall 5-hmC peaks density (FPKM) at days 7 and 10 were higher in hypoxia compared to normoxia (Figure 3.4E), suggesting enhanced maintenance and *de novo* synthesis of 5-hmC. Lastly, I quantified the percentage of 5-hmC peaks that were gained or lost in hypoxia compared to those in normoxia (Figure 3.4 F-G). To exclude noise in the data, I classified a peak as a differential peak if it had FPKM>1 in both normoxia and hypoxia, and if the difference in FPKM between normoxia and hypoxia was greater than 2-fold (See Figure 2.1B). With this definition, I found that nearly 40% of all 5-hmC peaks had higher density in hypoxia (5-hmC gain), whereas fewer than 10% of the peaks had higher 5-hmC density in normoxia (5-hmC loss) (Figure 3.4 F-G). 5-hmC gains were enriched in promoters, gene

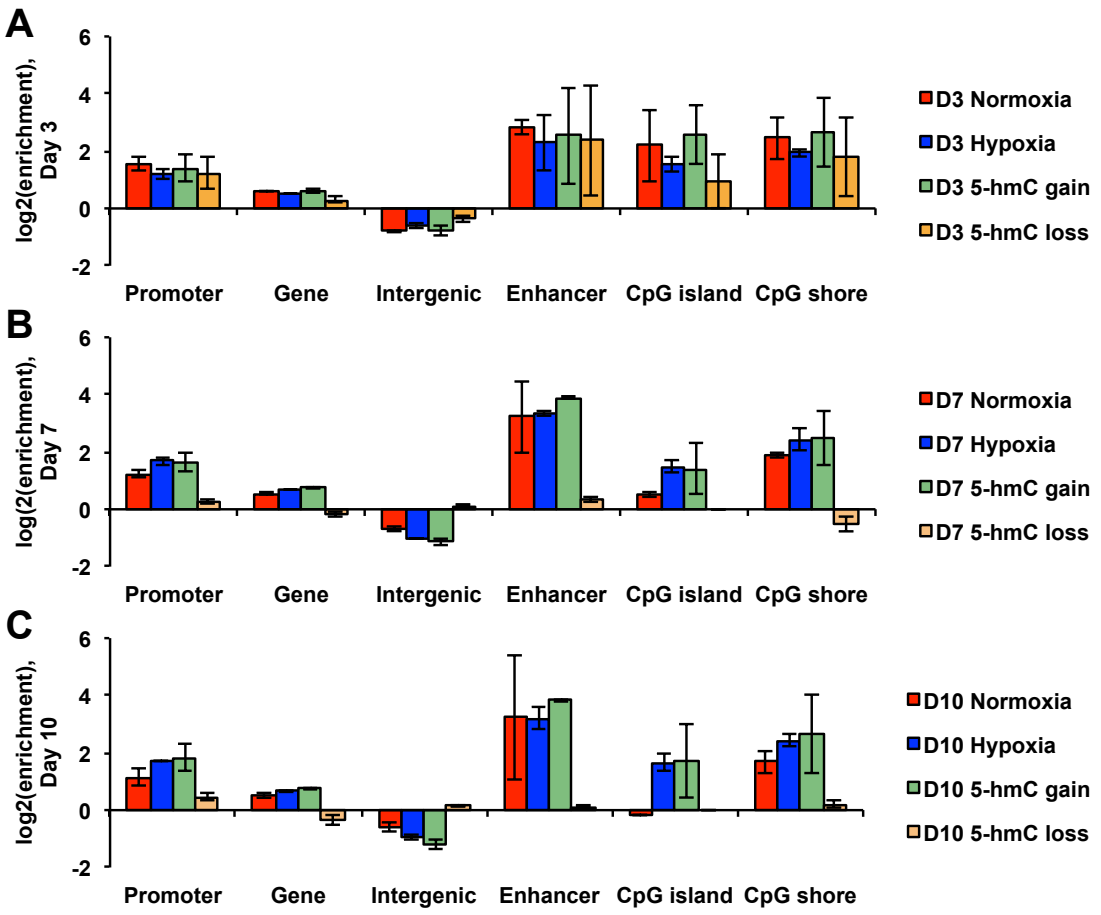
bodies, CpG islands and shores, and especially in enhancers throughout differentiation, whereas 5-hmC losses were not enriched in any genomic elements in days 7 and 10 (Figure 3.3). Based on these observations, I hypothesized that TET activity in hypoxia increases to re-organize 5-hmC distribution across the genome to promote hypoxic gene expression.

Figure 3.2 Representative 5-hmC distribution in hypoxia vs normoxia.



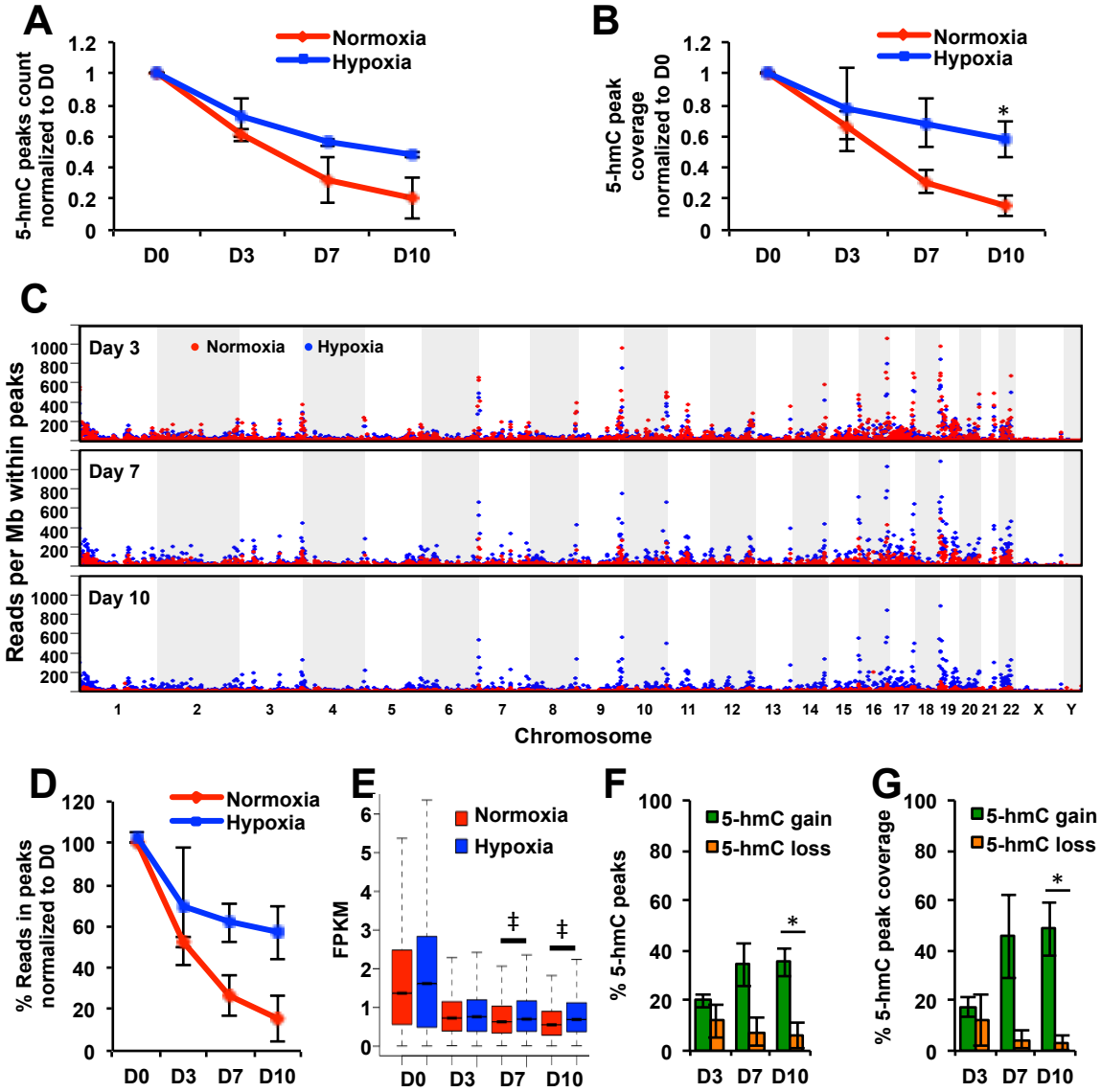
5-hmC distribution in and around several genes: (**A**) *BNIP3*, (**B**) *LDHA*, (**C**) *PDK2*, (**D**) *ENO1*, (**E**) *HK2*, and (**F**) *TET3*, showing normoxic and hypoxic conditions at days 3, 7, and 10. Chromosome coordinates (hg19) and sizes of window for each panel are shown below the tracks in the format chrN:start-end (size in kb).

Figure 3.3 Genomic distribution of 5-hmC peaks



(**A**) Day 3, (**B**) Day 7, and (**C**) Day 10 enrichment analysis of total and differential 5-hmC peaks in genomic regions, normoxia versus hypoxia. Red bars denote total peaks in normoxia; blue bars denote total peaks in hypoxia; green bars denote peaks with 5-hmC gain; and orange bars denote peaks with 5-hmC loss. N=2.

Figure 3.4 Hypoxia increases overall 5-hmC density during *in vitro* erythroid differentiation.

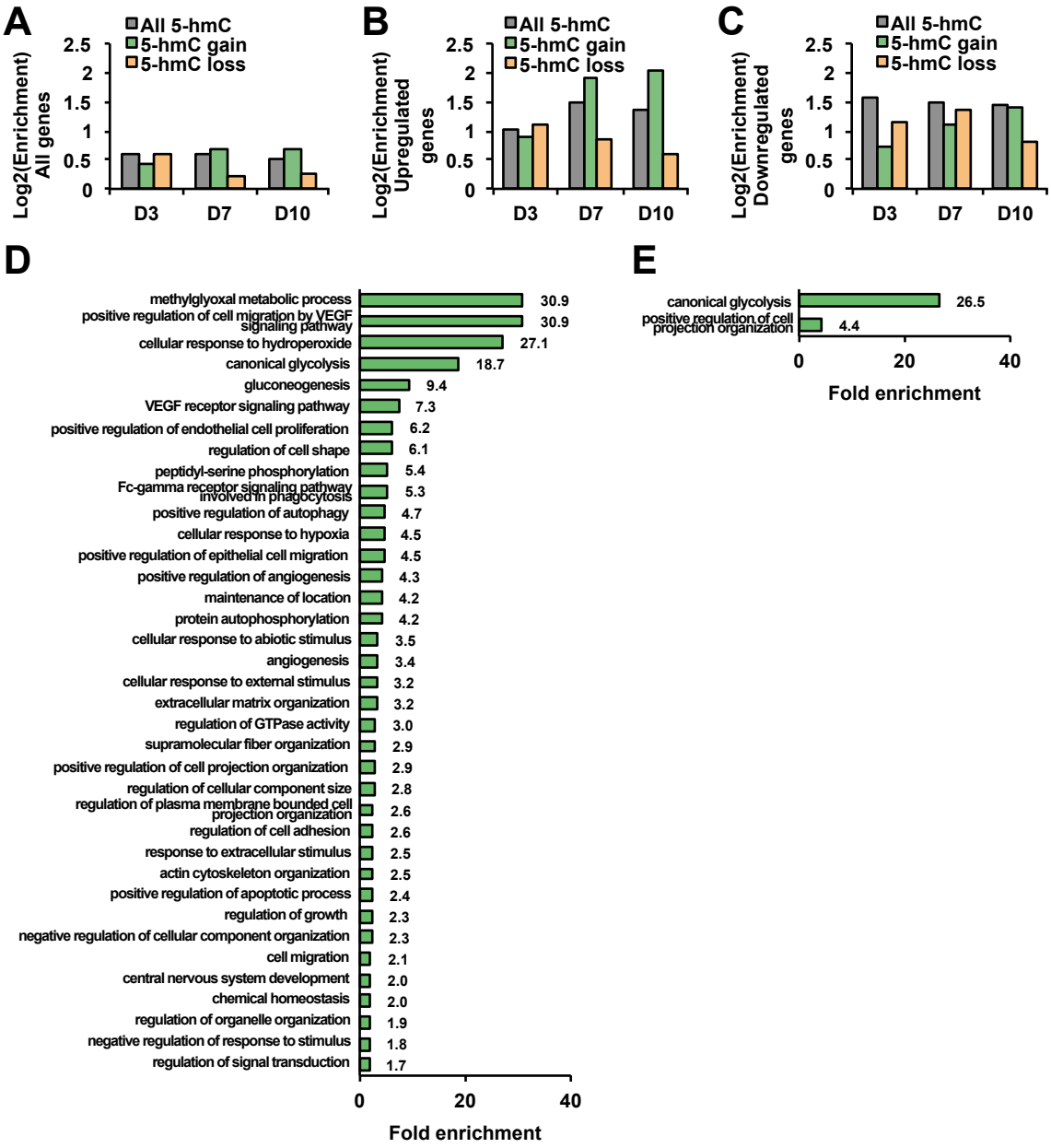


(**A-B**) Quantification of 5-hmC peaks by (**A**) peak counts and by (**B**) total bases covered by peaks in normoxia versus hypoxia at days 3, 7, and 10, normalized to corresponding measurements at day 0. Day 0 peak count was 473,206 for replicate 1 and 185,665 for replicate 2. (**C**) Dot plots of numbers of sequencing reads within peaks called by MACS2 per million total aligned reads. Each dot represents a genomic region of 1Mb. (**D**) Global measurement of sequencing reads within peaks, normalized to day 0. (**E**) Box plot of all 5-hmC peaks FPKM in all time points in normoxia vs. hypoxia. Wilcoxon rank-sum test was used to evaluate statistical significance. (**F-G**) Quantification of peaks that gained or lost 5-hmC in hypoxia (**F**) by peak counts and (**G**) by total bases covered by peaks as percentages of total peaks. N=2. ‡: $p<10^{-8}$. *: $p<0.05$

TET3 is upregulated in hypoxia in erythropoietic cells

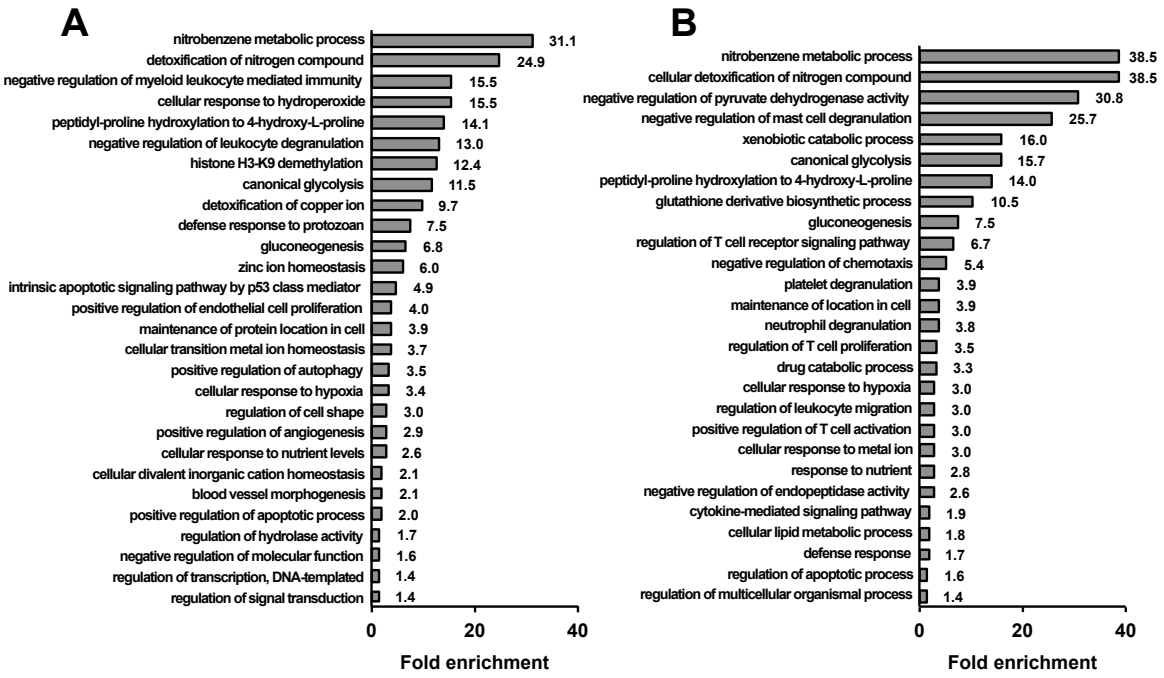
Next, I performed RNA-sequencing in normoxic versus hypoxic CD34+ HSPCs during erythroid differentiation (GSE40243 [as published in (Madzo et al., 2014)], GSE142870 [Hypoxia data from the current study (Cao et al., 2020)]). I examined whether 5-hmC changes correlated with changes in gene expression by calculating the enrichment of 5-hmC gain or loss in or near upregulated or downregulated genes. I defined 5-hmC peaks to be near a gene if they locate within 2kb flanking the gene. Although 5-hmC peaks were enriched in genes generally, I found a particularly strong correlation between 5-hmC gains and upregulated genes at days 7 and 10 (Figure 3.5B). I used gene ontology (GO) enrichment analysis (Mi et al., 2019a; Mi et al., 2019b) to show that the upregulated genes containing 5-hmC gains at day 7 were involved in glycolysis, hypoxic response, signal transduction, and cytoskeleton remodeling (Figure 3.5D). The upregulated genes containing 5-hmC gains regions at day 10 were enriched in glycolysis (26.5-fold) and positive regulation of cell projection organization (Figure 3.5E). In contrast, GO analysis on all upregulated genes at day 7 and 10, regardless of associated 5-hmC changes, returned notably different lists (Figure 3.6). Glycolysis and hypoxic response genes were less enriched, while more terms related to ion homeostasis and small molecule metabolism were more enriched (Figure 3.6). These results suggest that 5-hmC may have important function in promoting the expression of hypoxia inducible genes and may contribute to cellular response to hypoxia.

Figure 3.5 5-hmC gain is enriched in hypoxia induced genes.



(**A-C**) Enrichment of 5-hmC peaks in (**A**) all genes, (**B**) genes upregulated by hypoxia, and (**C**) genes downregulated by hypoxia. Total 5-hmC enrichment is represented by grey bars; 5-hmC gain is represented by green bars; and 5-hmC loss is represented by orange bars. (**D-E**) GO term enrichment of (**D**) day 7 and (**E**) day 10 upregulated genes containing regions with 5-hmC gain (green bars in **B**). Fold enrichment scores with FDR<0.05 are given to the right of each GO term. N=2 for RNA-seq experiments.

Figure 3.6 GO enrichment analysis of genes upregulated in hypoxia.

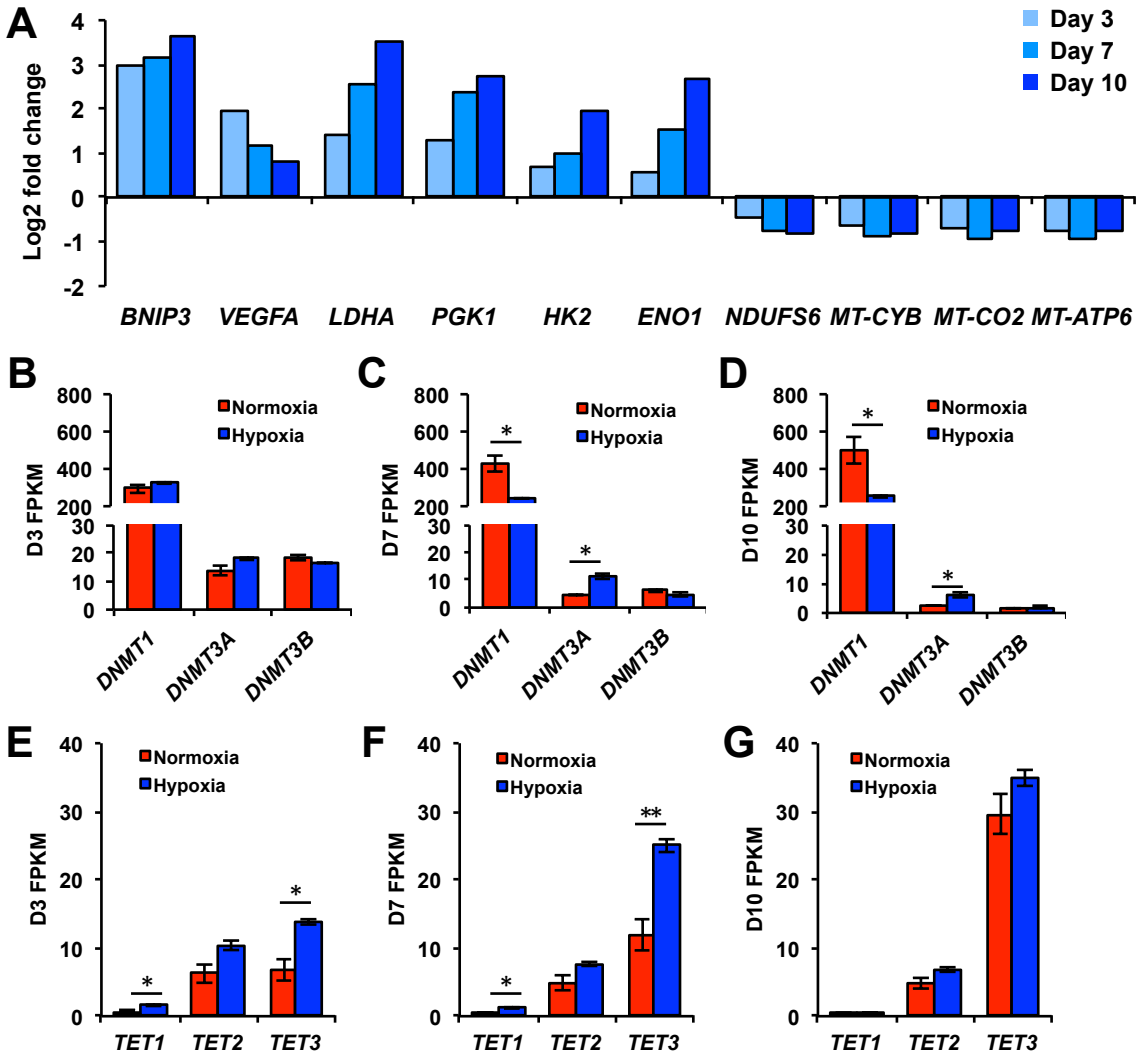


(**A-B**) GO term enrichment for all genes upregulated in hypoxia at (**A**) day 7 and (**B**) day 10. Fold enrichment scores with FDR<0.05 are given to the right of each GO term.

I examined the RNA-seq data further to find candidate genes that could explain the increased 5-hmC in hypoxia. Hypoxic conditions were confirmed by observing the upregulation of canonical hypoxia responsive genes as well as the downregulation of mitochondrial genes (Figure 3.7A). I found that *DNMT1*, which encodes the main maintenance methyltransferase, was suppressed at days 7 and 10, whereas expression of the *de novo* methyltransferase *DNMT3A* was increased at these time points (Figure 3.7 B-D). I found that both *TET2* and *TET3* were expressed throughout erythropoiesis (Figure 3.7 E-G). In particular, *TET3* expression was upregulated as erythropoiesis progressed, whereas *TET2* expression remained relatively stable. *TET1* expression was

very low compared to *TET2* or *TET3* (Figure 3.7 E-G). The expression pattern of *TET3* found here was measured by RNA-seq, which I consider to be more reliable than previous methods (Madzo et al., 2014) and are validated by independent work (Yan et al., 2017) as well as consensus bone marrow expression data from Human Protein Atlas (Uhlén et al., 2015; Uhlen et al., 2017) (Figure 3.9A). Notably, I found that *TET3*, but not *TET2*, was upregulated during erythroid differentiation up to day 7 (Figure 3.7 F-G). This indicated that *TET3* could be upregulated by hypoxia and facilitate 5-hmC accumulation and gene expression.

Figure 3.7 Hypoxia upregulates TET3 expression.

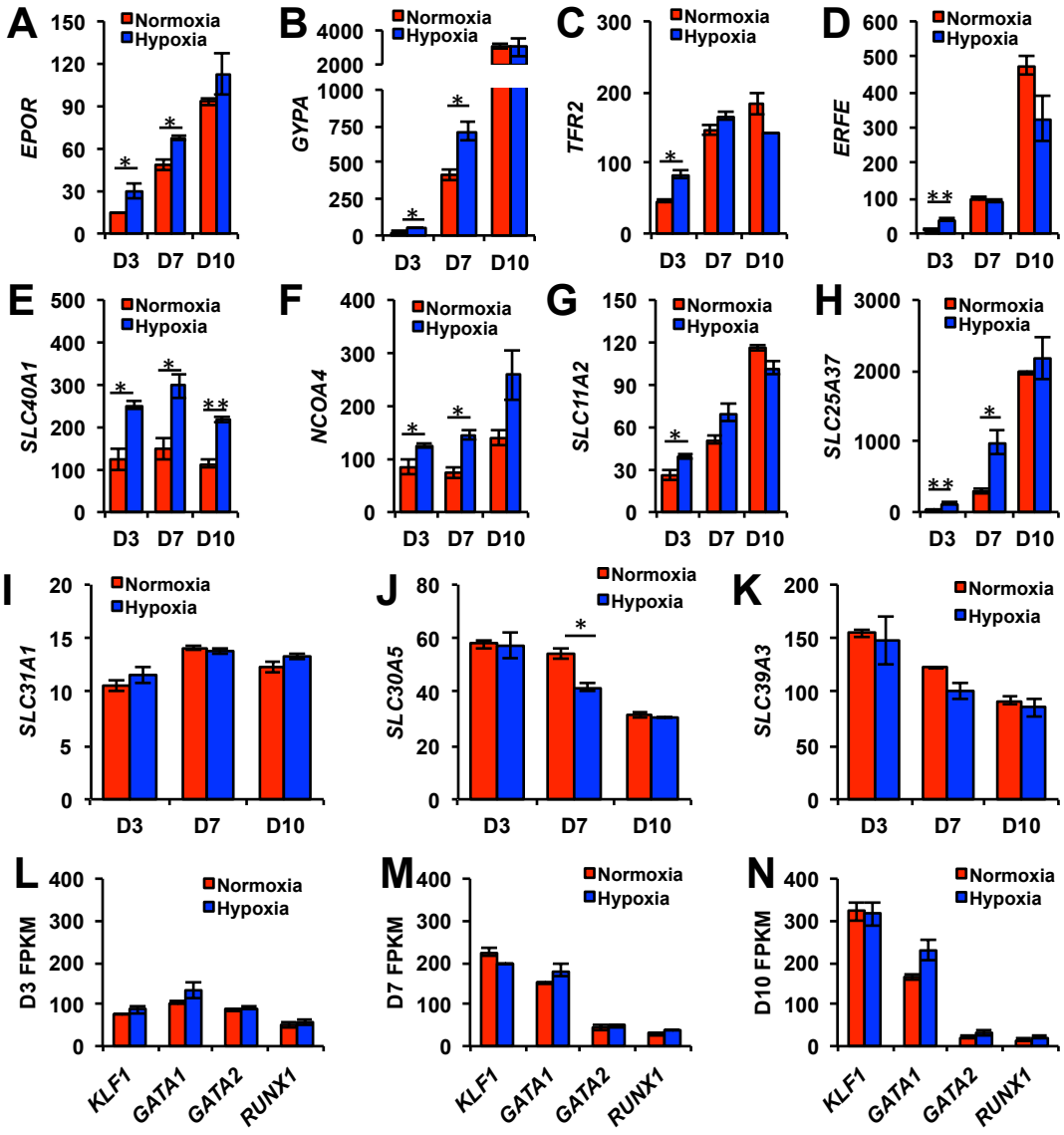


(**A**) Hypoxia treatment upregulates HIF-1 target genes and genes encoding glycolysis enzymes, while downregulating genes encoding proteins in the electron transport chain. (**B-D**) *DNMT1*, *DNMT3A*, and *DNMT3B* expression in normoxia versus hypoxia at (**B**) day 3, (**C**) day 7, and (**D**) day 10 of *in vitro* erythropoiesis. (**E-G**) *TET1-3* expression in normoxia versus hypoxia at (**E**) day 3, (**F**) day 7, and (**G**) day 10 of *in vitro* erythropoiesis. N=2 for all panels shown. *: p<0.05, **: p<0.01.

Lastly, I examined the expression of genes involved in erythropoiesis in normoxia versus hypoxia. I found that hypoxic cells express more *erythropoietin receptor* (*EPOR*), *glycophorin A* (*GYPA*), and *transferrin receptor 2* (*TFR2*) at days 3 and/or 7 (Figure 3.8

A-C), suggesting a higher sensitivity to EPO stimulation and a more robust early erythroid phenotype. *Erythroferrone* (*ERFE*), which encodes a secreted protein that inhibits hepcidin production by the liver to increase blood iron levels, was upregulated at day 3 (Figure 3.8D). We then examined the genes involved in iron homeostasis in erythrocytes: *SLC40A1* encodes the only known iron exporter (*ferroportin*) (Andrews, 2008); *NCOA4* encodes a receptor for ferritin endocytosis (Mancias et al., 2015; Mancias et al., 2014); *SLC11A2* encodes the proton-coupled divalent metal ion transporter responsible for transporting Fe(II) from endosomes to the cytosol (Andrews, 2008); and *SLC25A37* encodes the erythroid specific mitochondrial iron transporter (*Mitoferrin1*) that is crucial for heme biosynthesis (Andrews, 2008; Richardson et al., 2010). All four genes were significantly upregulated by hypoxia at day 3, and all but *SLC11A2* were significantly upregulated by hypoxia (up to 3-fold) at day 7 (Figure 3.8 E-H). By day 10, ferroportin remained significantly upregulated in hypoxia (1.9-fold) (Figure 3.8E). The genes encoding the copper transporter *SLC31A1* and zinc transporters *SLC30A5 and SLC39A3* were not upregulated by hypoxia (Figure 3.8 I-K). Together, the RNA-seq data suggest that HSPCs in hypoxia are more sensitive to EPO stimulation, exhibit a more robust activation of erythroid lineage genes, and have increased iron homeostatic activities, consistent with previous studies (Chin et al., 2000; Manalo et al., 2005; Yoon et al., 2006). Finally, expression of erythroid TFs *KLF1*, *GATA1/2*, and *RUNX1* was not different in normoxia versus hypoxia (Figure 3.8 L-N), suggesting that the expression differences described above were not mediated by transcriptional changes of these erythroid TF genes.

Figure 3.8 Hypoxia enhances early erythroid differentiation and iron transport.

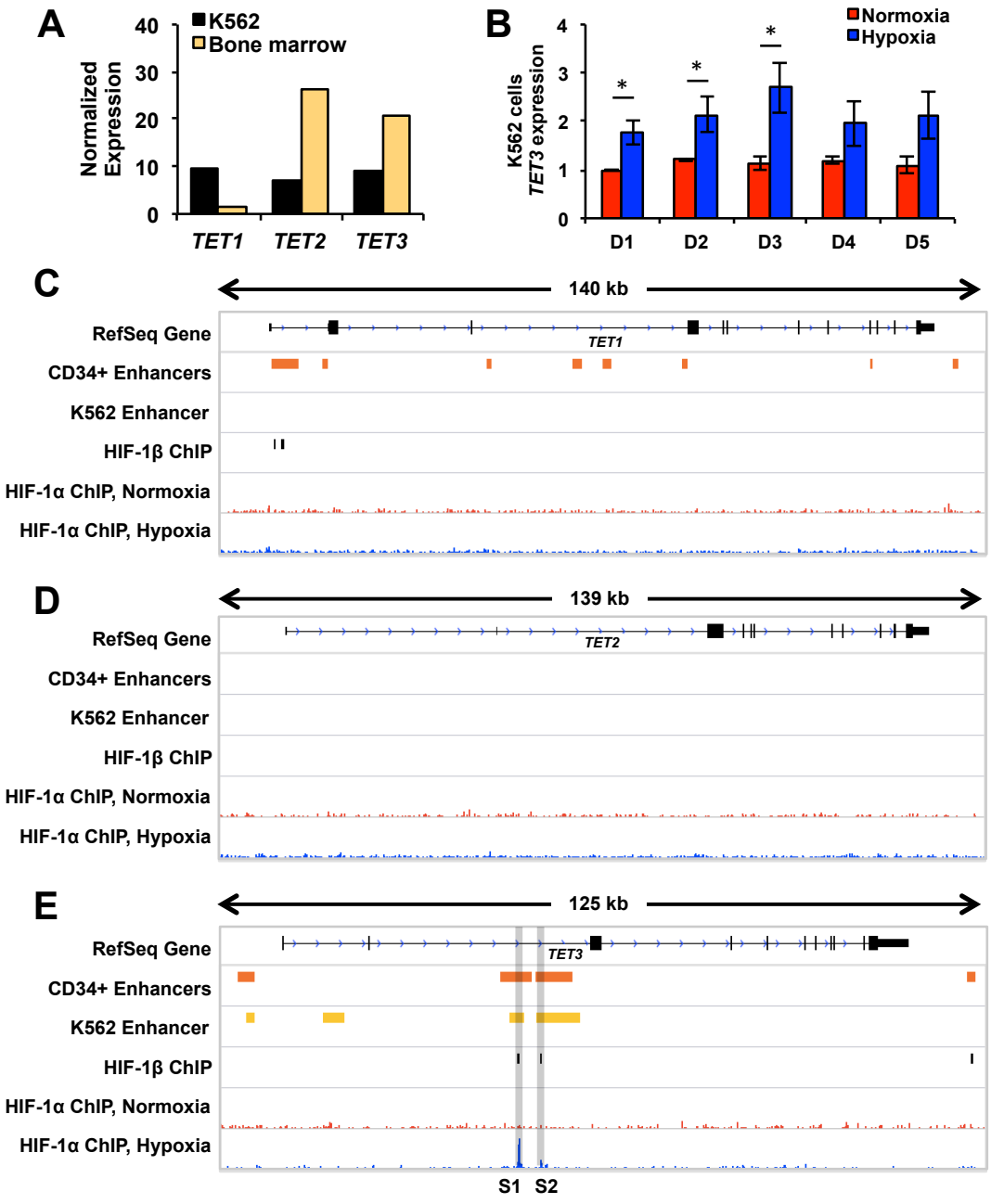


(**A-K**) Expression of (**A**) *EPOR*, (**B**) *GYPA*, (**C**) *TFR2*, (**D**) *ERFE*, (**E**) *SLC40A1*, (**F**) *NCOA4*, (**G**) *SLC11A2*, (**H**) *SLC25A37*, (**I**) *SLC31A1*, (**J**) *SLC30A5*, and (**K**) *SLC39A3* in normoxia (red) versus hypoxia (blue) at day 3, 7, and 10 of *in vitro* erythropoiesis. (**L-N**) Expression of genes encoding erythroid transcription factors: *KLF1*, *GATA1*, *GATA2*, and *RUNX1* at (**L**) day 3, (**M**) day 7, and (**N**) day 10 in normoxia (red) versus hypoxia (blue). N=2. *: p<0.05, **: p<0.01.

HIF-1α binds to enhancers within *TET3* intron 2

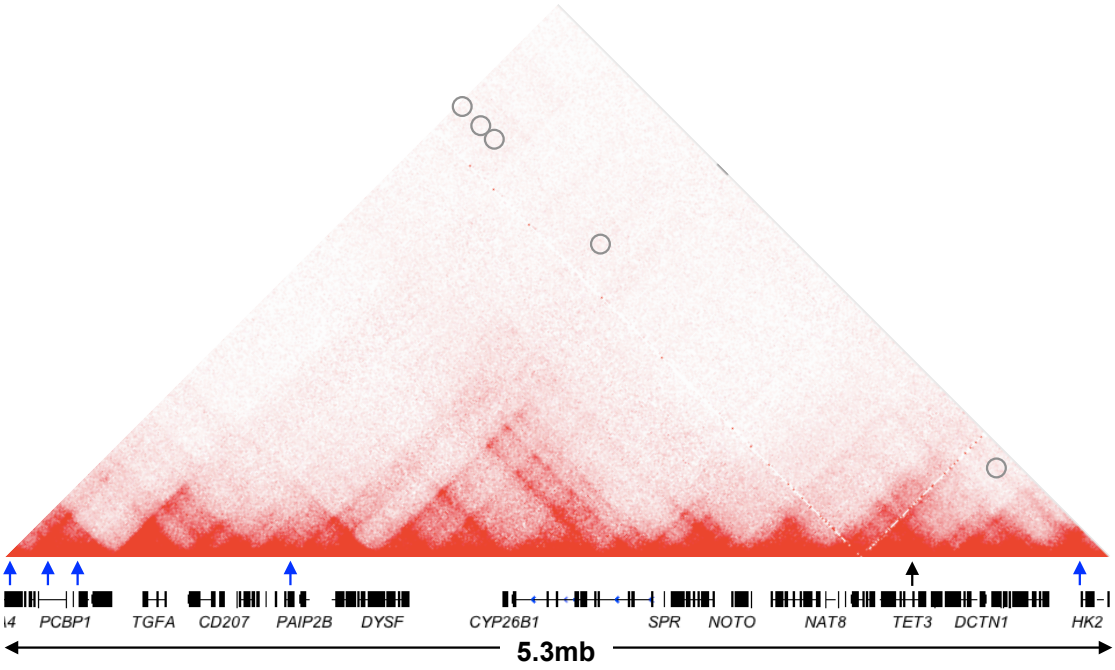
I then investigated whether *TET3* is a direct target of HIF-1. I found that K562 cells tolerate hypoxia treatment well, and that *TET3* is upregulated by hypoxia in these cells (Uhlén et al., 2015; Uhlen et al., 2017) (Figure 3.9 A-B). Therefore, I chose to use this cell line as a model to investigate the molecular mechanisms controlling the hypoxia-induced *TET3* upregulation. I performed HIF-1α ChIP-seq in K562 cells cultured under normoxic versus hypoxic conditions (GSE142870). I found HIF-1α binding at 2 sites (Sites 1 and 2) within intron 2 (Figure 3.9E). In contrast, HIF-1α did not bind either *TET1* or *TET2* (Figure 3.9C-E). K562 Hi-C data showed no strong interactions between *TET3* and other nearby HIF-1 binding sites, suggesting that the Sites 1 and 2 are the main binding sites responsible for *TET3* upregulation (Rao et al., 2014; Robinson et al., 2018) (Figure 3.10).

Figure 3.9 HIF-1 only binds *TET3*, but not *TET1* and *TET2*, in K562 cells.



(**A**) Comparison of the expression of *TET* genes in K562 cells versus human bone marrow (data from Human Protein Atlas (Uhlén et al., 2015; Uhlen et al., 2017)). (**B**) *TET3* expression in K562 cells in normoxia (red) versus hypoxia (blue), up to 5 days. N=3. *: p<0.05. (**C-E**) K562 HIF-1α ChIP-seq revealed no binding site at (**C**) *TET1* or (**D**) *TET2*. (**E**) *TET3* is the only *TET* genes that was bound by HIF-1α in K562 cells. CD34+ and K562 enhancer tracks were downloaded from Enhancer Atlas {Gao, 2016 #247}. ENCODE accession number for HIF-1β ChIP track: ENCFF507MGL. S1: HIF1 binding Site 1. S2: HIF1 binding Site 2.

Figure 3.10 Hi-C heatmap of the genomic region near *TET3*.



K562 Hi-C heatmap showing 5.3 megabases (chr2:70,000,000-75,300,000) containing *TET3* and the closest HIF1 binding sites (blue arrows). Grey circles indicate parts of the heatmap that quantify *TET3*:HIF1 binding site interactions. Hi-C data from Rao et al., 2014; Visualization by Juicebox (Robinson et al., 2018)

HIF-1α binding is required for *TET3* upregulation in hypoxia

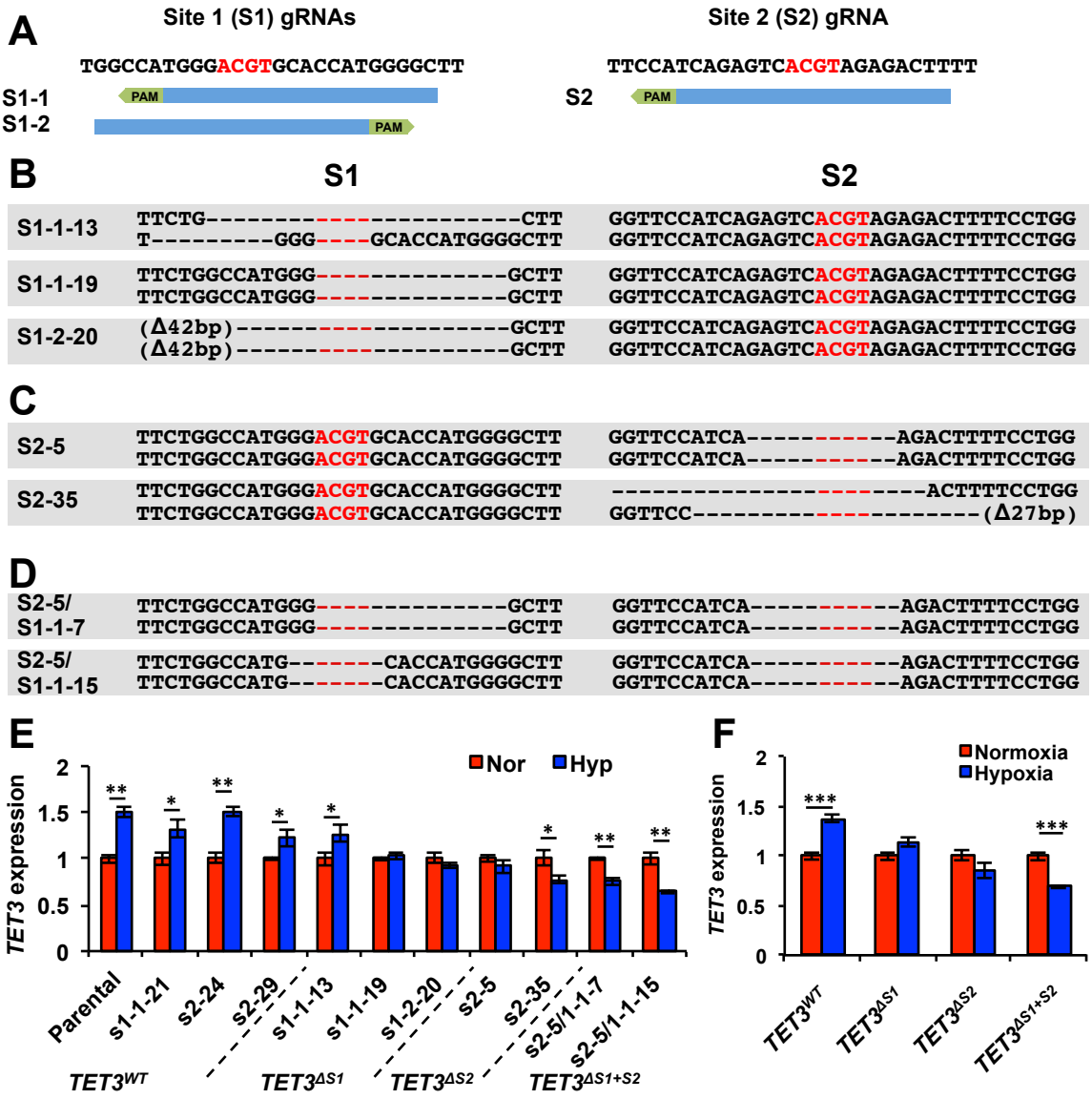
To assess the effects of HIF-1α binding at these sites, I designed CRISPR/Cas9 guides to delete one or both of the binding sites (Figure 3.11A). In total, I identified three clones with Site 1 deletions ($TET3^{\Delta S1}$), two with Site 2 deletions ($TET3^{\Delta S2}$), two with double-deletions ($TET3^{\Delta S1+S2}$), as well as three WT clones from vector only transfection ($TET3^{WT}$) (Figure 3.11 B-D). I then treated these clones in normoxia versus hypoxia for 72 hours and measured *TET3* expression (Figure 3.11 E-F). $TET3^{WT}$ cells exhibited the expected upregulation of *TET3* in hypoxia (~1.4 fold), whereas neither $TET3^{\Delta S1}$ nor

TET3$^{\Delta S2}$ cells upregulated *TET3* (Figure 3.11 E-F). Strikingly, *TET3*$^{\Delta S1+S2}$ cells showed *TET3* suppression (0.7-fold) in hypoxia (Figure 3.11 F), demonstrating an additive effect of these sites. These results suggested that both HIF binding sites are key sites that dictate *TET3* expression in response to hypoxia in erythropoiesis.

Loss of *TET3* induction impairs K562 cell differentiation potential and survival in hypoxia

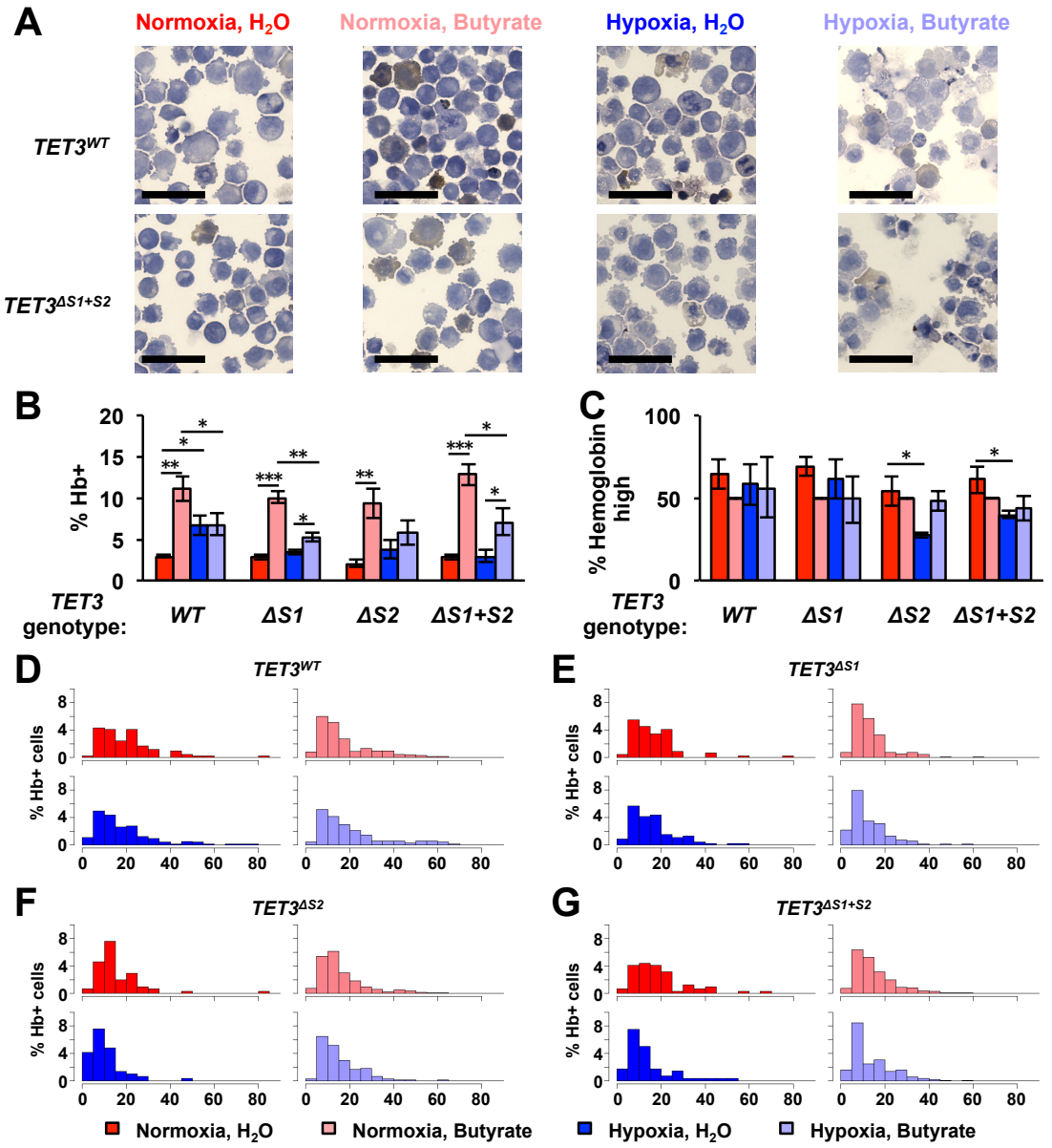
To investigate the physiological significance of the two HIF-1α binding sites in erythropoiesis, I induced erythroid differentiation in the parental and CRISPR ΔS1/ΔS2 cell lines in normoxia versus hypoxia. I used 1mM sodium butyrate to drive erythropoiesis in K562 cells under normoxic or hypoxic conditions (Shariati et al., 2016). After 3 days of treatment, samples were collected to make cytospin slides, which were stained with hematoxylin and benzidine to quantify hemoglobin production in each cell (Figure 3.12A). Interestingly, parental K562 cells spontaneously differentiated in hypoxia without sodium butyrate treatment, resulting in a higher percentage of hemoglobin producing cells (Figure 3.12B). However, this phenomenon was not observed in any of the cell lines containing deletions of the HIF-1 binding sites (Figure 3.12B). I quantified the percentage of hemoglobin producing cells that accumulated high levels of hemoglobin (Figure 3.12C). I found that cells with Site 2 deleted were less capable of accumulating hemoglobin in hypoxia (Figure 3.12C), which can be seen as a shift in benzidine staining intensity from normoxia to hypoxia (Figure 3.12D-G, left panels). Together, my results show that loss of either Site 1 or Site 2 impairs the initiation of erythropoiesis in respond to hypoxia, and the loss of Site 2 further inhibits hemoglobin production.

Figure 3.11 HIF-1 binding is required for *TET3* upregulation in hypoxia.



(**A**) Sequences of binding site 1 and site 2, as well as positions of CRISPR guides used to delete the binding sites. The core E-box motif (ACGT) is indicated by red letters. PAM sequences (NGG) are indicated by green boxes. (**B-D**) Sequences of deletion clones used in subsequent experiments. The core E-box motif is highlighted in red. Deletions are represented by dashes. The total numbers of bases deleted are indicated if the deletion extends beyond the sequences shown. (**B**) Sequences of 3 *ΔS1* clones. (**C**) Sequences of 2 *ΔS2* clones. (**D**) Sequences of 2 *ΔS1+S2* clones generated from the S2-5 clone (**C**). (**E-F**) *TET3* expression in (**E**) each individual clone and (**F**) data combined from all clones for each genotype in normoxia (red) versus hypoxia (blue). N=3. *: p<0.05, **: p<0.01, ***: p<0.001.

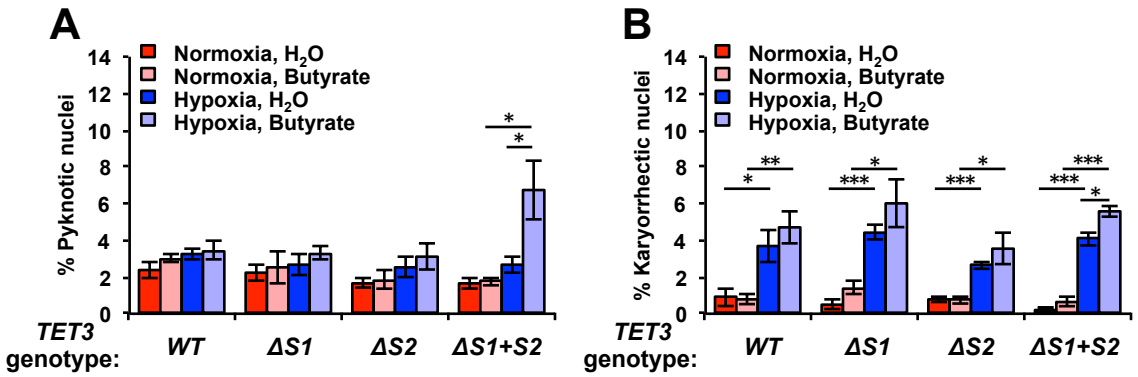
Figure 3.12 Loss of TET3 negatively impact K562 erythroid differentiation potential.



(**A**) Representative images of benzidine-hematoxylin staining of K562 *TET3*^{WT} and *TET3*$^{\Delta S1+S2}$ cells. Benzidine staining (brown) differentiates hemoglobin positive (Hb+) and negative (Hb-) cells. Scale bars indicate 50 µm. (**B-C**) Quantification of cells (**B**) producing hemoglobin (% all cells) and (**C**) with high hemoglobin accumulation (% all hemoglobin+ cells). (**D-G**) Distribution of benzidine staining intensity in (**D**) *TET3*^{WT}, (**E**) *TET3*$^{\Delta S1}$, (**F**) *TET3*$^{\Delta S2}$, and (**G**) *TET3*$^{\Delta S1+S2}$ K562 cell. Cells were treated for 3 days with water in normoxia (control, red), 1mM sodium butyrate in normoxia (light red), water in hypoxia (blue), or 1mM sodium butyrate in hypoxia (light blue). A minimum of 500 cells were counted in each replicate for each treatment condition. N=3. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$.

In addition to hemoglobin production, I also examined the nuclear morphology of the non-hemoglobin-producing cells under all treatment conditions. Notably, compared to any other cells in any treatment conditions, the *TET3*$^{\Delta S1+\Delta S2}$ double deletion cells contained more pyknotic nuclei after sodium butyrate treatment and under hypoxic conditions (Figure 3.13A), suggesting higher rates of cell death. Likewise, *TET3*$^{\Delta S1+\Delta S2}$ cells were the only ones with higher rates of karyorrhectic (fragmented) nuclei after sodium butyrate treatment and under hypoxic conditions (Figure 3.13B), indicating that loss of both binding sites and the subsequent loss of *TET3* expression (Figure 3.11 E-F) decreases cell survivability in hypoxia, especially under differentiating conditions.

Figure 3.13 Loss of *TET3* negatively impact K562 chromatin stability.



(**A**) Quantification of K562 cells with pyknotic nuclei. (**B**) Quantification of K562 cells with karyorrhectic nuclei. Cells were treated for 3 days with water in normoxia (control, red), 1mM sodium butyrate in normoxia (light red), water in hypoxia (blue), or 1mM sodium butyrate in hypoxia (light blue). A minimum of 500 cells was counted in each replicate for each treatment condition. N=3. *: p<0.05, **: p<0.01, ***: p<0.001.

Discussion

Erythropoiesis in human is intricately regulated by oxygen availability. Low oxygen tension, as occurs at high altitude, leads to HIF-α stabilization in renal cells,

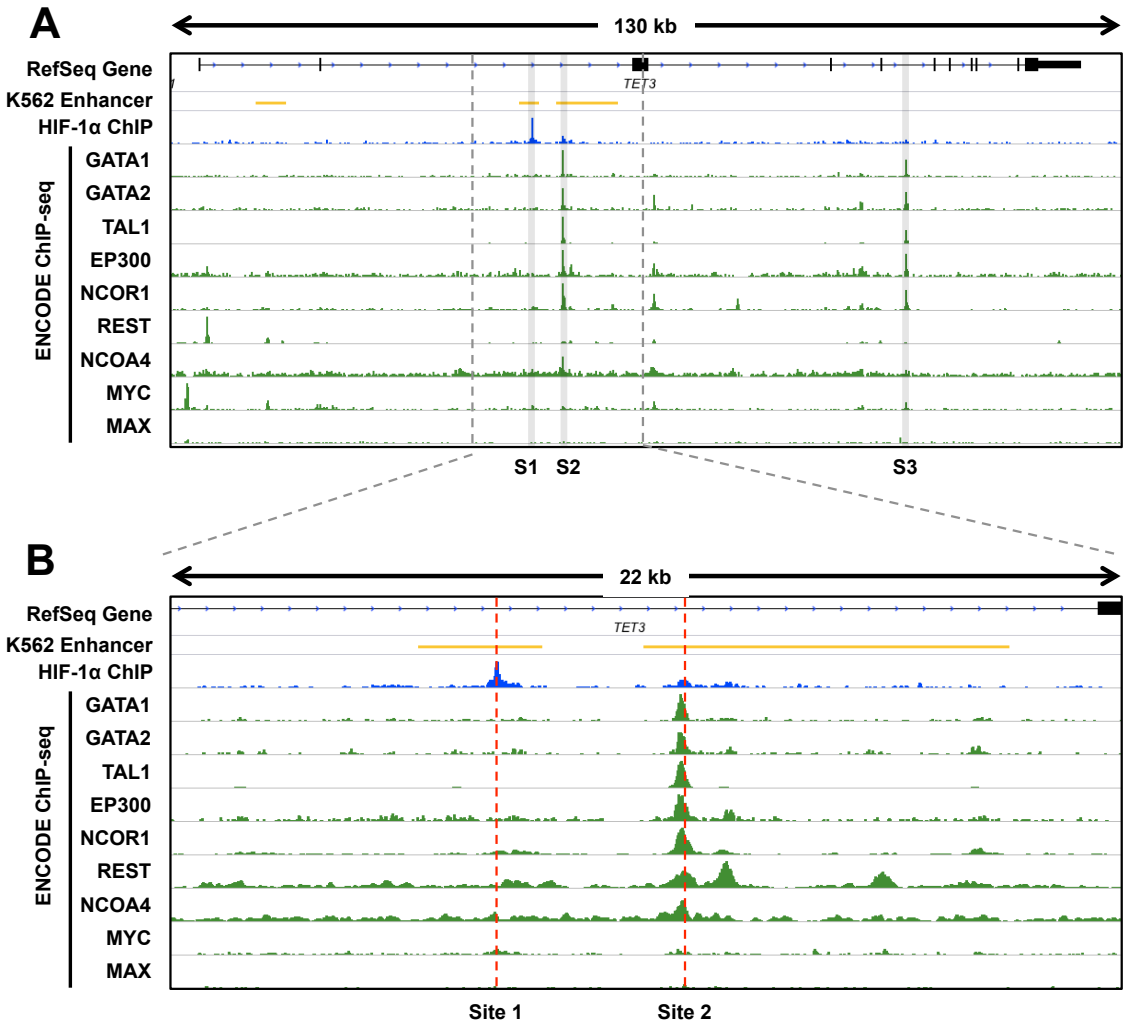
which then induces EPO production to stimulate erythropoiesis in the bone marrow (Ivan and Kaelin, 2017). Changes in the HIF-EPO axis have been demonstrated to have major effects on erythropoiesis. For example, patients with congenital *VHL* mutations or *VHL*-mutated renal cell carcinoma cells do not ubiquitinate HIF-α subunits and constitutively upregulate the *EPO* gene, leading to polycythemia (Ang et al., 2002; Bond et al., 2011; Pastore et al., 2003; Ricketts and Linehan, 2017; Yao et al., 2017). Genetic adaptations, *e.g.,* a gain-of-function variant of the PHD2 protein that increases oxygen affinity, have been identified in Tibetans that attenuate erythropoiesis in hypoxia to avoid polycythemia (Jeong et al., 2014; Jeong et al., 2018; Lorenzo et al., 2014; Tashi et al., 2017). These studies highlight the oxygen sensitive control of EPO production in renal cells, but did not examine the effects of hypoxia on HSPCs.

HIF stabilization alters histone modifications by recruiting p300 histone acetyltransferase in *VHL*-deficient clear cell renal cell carcinoma, which in turn activates gene transcription (Ricketts and Linehan, 2017; Yao et al., 2017). In addition, HIF binding promotes chromatin accessibility by remodeling local histone occupancy (Suzuki et al., 2018). Our study suggests that hypoxia sensitizes and promotes erythroid differentiation of HSPCs by inducing *EPOR* expression as well as by changing the epigenome via TET3 and possibly DNMT3A. Zhang *et al* have shown that DNMT3A and TET2 exhibit both competitive and cooperative activities in HSPCs to suppress the expression of lineage-specific genes (Zhang et al., 2016). It is possible that DNMT3A and TET3 have similar interactions in the erythroid lineage under hypoxic conditions. In addition, we have also shown that hypoxia increases the expression of genes encoding proteins involved in transporting iron ions. The iron homeostasis is important as Fe(II) is

both required for heme biosynthesis and a crucial cofactor for dioxygenases like TETs. The effects of hypoxia on HSPCs could amplify the renal cell response and contribute to rapid acclimation to hypoxic environments.

Available ENCODE ChIP data in K562 revealed that Site 1 and Site 2 in *TET3* are very different with respect to TF binding in general (Figure 3.14). Although the two HIF-1 binding sites have similar effects on *TET3* expression under hypoxia, Site 2, but not Site 1, is bound by a number of erythroid TFs or complexes, including GATA1, GATA2, TAL1, EP300, NCOR1, REST, and NCOA4 (Figure 3.14). Two binding motifs of GATA1/2 (GATA-box) are upstream (-60 bp and -105 bp) of Site 2, suggesting that this site is likely also important in regulating erythropoiesis-specific *TET3* expression. This hypothesis is supported, in part, by the observation that K562 cells with Site 2 deletion were less capable of accumulating hemoglobin in hypoxic conditions (Figure 3.12C). Notably, despite the motif similarity of MYC and MAX to HIF-1, no strong MYC/MAX binding was found at either Site 1 or Site 2 (Figure 3.14), which implies that these sites are HIF-specific sites. Based on these observations, I concluded that Sites 1 and 2 are the main regulatory sites of *TET3* in erythropoiesis.

Figure 3.14 Binding of other transcription factors in *TET3*.



(**A**) ENCODE ChIP-seq data overlaid on HIF-1α ChIP-seq results at *TET3*. Site 1 is largely unique to HIF-1, whereas Site 2 has multiple transcription factors binding sites very close to HIF-1 binding site. Site 3 lacks HIF-1 binding, but is occupied by multiple other transcription factors. (**B**) Zoomed in view of Sites 1 and 2. Dotted red lines indicate the positions of the ACGT motifs that were deleted by CRISPR. Note that the summits of GATA1/2, TAL1, and EP300 are slightly upstream of the HIF-1 binding site, which is likely due to the two GATA-boxes (GATAAG) being upstream (-60 bp and -105 bp) of the HIF-1 binding site. ENCODE accession ID for each track: GATA1- ENCFF942NUX; GATA2- ENCFF990JRT; TAL1- ENCFF389WLJ; EP300- ENCFF703ULA; NCOR1- ENCFF246WTN; REST- ENCFF686NWE; NCOA4- ENCFF662JKA; MYC- ENCFF058VAU; MAX- ENCFF000YTL.

Despite having similar core catalytic domains, TET3 harbors a CXXC domain that is not present in TET2 (Cao et al., 2019; Pastor et al., 2013). In contrast, TET2 is uniquely phosphorylated by JAK2 at two tyrosine residues near the C-terminus, which increases its enzymatic activity (Jeong et al., 2019). *TET3*-deficient HSPCs have slower growth, undergo more apoptosis, and fail to enucleate properly, whereas *TET2*-deficient HSPCs have increased growth and delayed differentiation (Yan et al., 2017). Adding to the differences between *TET2* and *TET3*, our data show that *TET3* expression is responsive to hypoxia, whereas *TET2* expression is insensitive to oxygen availability (Figure 3.7 E-G). These results demonstrate that TET2 and TET3 respond to different environmental signals and have distinct functions during erythropoiesis.

Further differences between the two *TET* genes are observed in clinical studies. Although *TET2* mutations are common in patients with hematopoietic malignancies (Abdel-Wahab et al., 2009; Bowman and Levine, 2017; Cao et al., 2019; Delhommeau et al., 2009; Langemeijer et al., 2009; Lasho et al., 2018), *TET3* mutations are relatively rare. Two studies have described several *TET3* somatic exonic mutations in CMML (Lasho et al., 2018; Merlevede et al., 2016). One of these mutations resulted in a truncated protein ($TET3^{Y473}$) lacking the catalytic domain, and another missense mutation ($TET3^{R1548H}$) was shown to impair TET3 catalytic function severely (Lasho et al., 2018; Merlevede et al., 2016). Another study on chronic myeloid leukemia also identified one *TET3* mutated patient out of a cohort of 24 patients (Togasaki et al., 2017), although the significance of this mutation ($TET3^{A128T}$) is unknown (Togasaki et al., 2017). The results from these studies suggest that *TET3* mutations in hematopoietic cells may predispose patients to chronic leukemias. Considering the remarkable effects

of hypoxia on the epigenome, future studies regarding erythropoiesis and especially relating to TET functions should incorporate hypoxia to better simulate the physiological environment of the HSPCs.

CHAPTER IV

BET inhibitors enhance the expression of fetal and embryonic *β-globin* genes

The data shown in this chapter is adapted from the following manuscript: Cao, J. Z., Bigelow, K., Wickrema, A., and Godley, L. A. BET inhibitors enhance the expression of fetal and embryonic globin genes in erythroleukemia cell lines. In preparation for submission.

In this study, I designed and performed the experiments, analyzed the data, and wrote the manuscript. Other author contributions include: K.B. performed additional experiments. L.A.G. conceived of the study and provided insights in experimental design and data interpretation. A.W. provided additional input for experimental design and data interpretation.

Summary

The human *β-globin* gene cluster comprises genes expressed at distinct developmental stages: embryonic (*HBE1*), fetal (*HBG2/1*), and adult (*HBD* and *HBB*). The mechanisms underlying the embryonic-fetal-adult β-globin switches involve epigenetic regulation of gene silencing and activation mediated by numerous protein complexes, miRNA control of gene expression, and chromatin looping. In inherited blood disorders caused by *HBB* mutations, such as SCA and β-thalassemia, a major treatment strategy is to reactivate expression of fetal or embryonic β-globin genes to dilute or replace the defective adult β-globin. However, current options for pharmacological epigenetic regulators are limited. In this chapter, I show that BET inhibitors induce specific upregulation of *HBE1* and *HBG1/2 in vitro*. I have shown that, in a cell lines expressing *HBB*, treatment with JQ1 or similar compounds reduces the expression of *HBB* in favor of *HBE1* and *HBG1/2* by downregulating the expression of TFs that mediate the fetal-adult switch as well as decreasing the interaction between the LCR and *HBB/HBG1/2,* while maintaining the interaction with *HBE1*. This work suggests that BET inhibitors could have clinical utility in the treatment of *HBB*-related globinopathies by upregulating the expression of embryonic and fetal *β-globin* genes.

Introduction

In mammals, the bromodomain and extra-terminal domain (BET) family of proteins are histone acetyl-lysine readers crucial for epigenetic regulation of gene expression through recruitment of the transcription machinery (Dhalluin et al., 1999; Doroshow et al., 2017; Filippakopoulos and Knapp, 2014; Stathis and Bertoni, 2018).

The BET family consists of the ubiquitously expressed BRD2, BRD3, BRD4, and the germ cell-specific BRDT, each of which contains two acetyl-lysine binding bromodomains (BD1 and BD2) (Doroshow et al., 2017; Stathis and Bertoni, 2018). BET proteins upregulate oncogenes such as *MYC* in various cancers, including in hematopoietic malignancies (Reyes-Garau et al., 2019), and consequently, various BET inhibitors are being tested in clinical trials for those patients (Reyes-Garau et al., 2019; Roe et al., 2015). However, the effects of BET inhibitors on hematopoiesis are unclear, with some data suggesting that BET inhibition promotes erythropoiesis specifically (Goupille et al., 2012).

With detailed understanding of the transcriptional and epigenetic control of the *β-globin* locus, it may be possible to design therapeutic interventions to re-express the epigenetically silenced embryonic and fetal *β-globins*. As described in Chapter I, the expression of genes in the *β-globin* locus is controlled by the interplay between local chromatin structure and erythroid-specific TFs. Prior studies have shown that BET family proteins are involved in the looping of LCR to the adult *β-globin* gene via binding to acetylated GATA1 (Gamsjaeger et al., 2011; Lamonica et al., 2011). Others have established that BET inhibitors disrupt GATA1-mediated activation in mouse G1E cells (Stonestrom et al., 2015) and induce differentiation of the UT7 human erythroid cell line (Goupille et al., 2012). However, it is not clear how BET inhibitors affect the transcription of genes encoding erythroid TFs, and the effects of BET inhibitors on the transcription *β-globin* genes have not been evaluated. Therefore, I hypothesized that BET inhibitors would promote expression of erythropoietic genes and suppress genes involved in myeloid differentiation. In addition, I hypothesized that BET inhibitors would disrupt the

interaction between the LCR and the adult globin genes as well as suppress the expression of TFs known to inhibit expression of fetal and embryonic *β-globins*. To this end, I tested whether BET inhibitors affect the interaction between the LCR and its target *β-globin* genes as well as the expression of additional genes regulating erythropoiesis.

Results

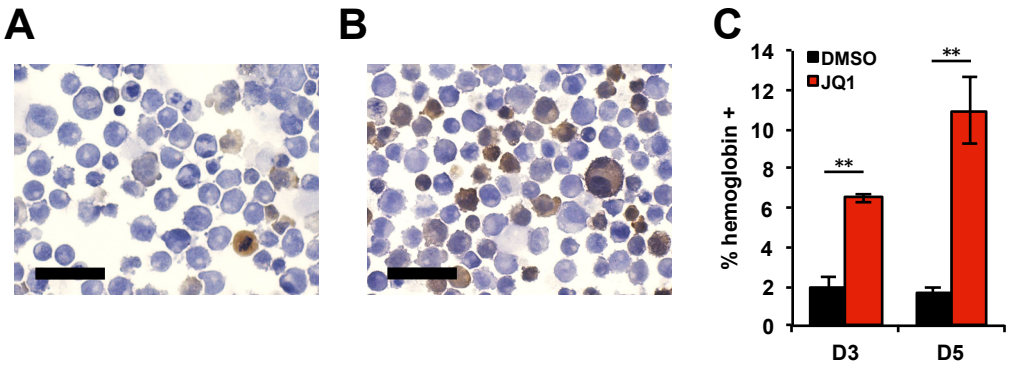
JQ1 Promotes Erythroid Differentiation

To examine the effects of BET inhibitors in erythroid cell lines, I first treated K562 cells with JQ1 for up to 5 days. I quantified the percentage of cells that accumulated hemoglobin by benzidine-hematoxylin staining (Figure 4.1). By day 3, JQ1-treated K562 cells (Figure 4.1B) had more hemoglobin-containing cells (~6%) than control cells (~2%, Figure 4.1C). By day 5, over 10% of JQ1-treated K562 cells contained hemoglobin (Figure 4.1C), demonstrating that JQ1 induces erythropoiesis as has been shown for another human erythroid cell line, UT-7 (Goupille et al., 2012).

Next, I performed RNA-sequencing (RNA-seq) on K562 and TF-1 cells treated with JQ1 in order to examine the systematic gene expression changes induced by JQ1. Unlike K562 cells, TF-1 cells are dependent on granulocyte-macrophage colony-stimulating factor (GM-CSF) or interleukin 3 (IL-3) to survive, and upon cytokine deprivation, the cells become responsive to EPO and differentiate down the erythroid lineage. As a control, I compared K562 transcriptomes between untreated (D0) versus DMSO-treated cells and found that DMSO alone did not alter the K562 transcriptome (Figure 4.2 A-B). Using GO term clustering (Reimand et al., 2019), I found that JQ1

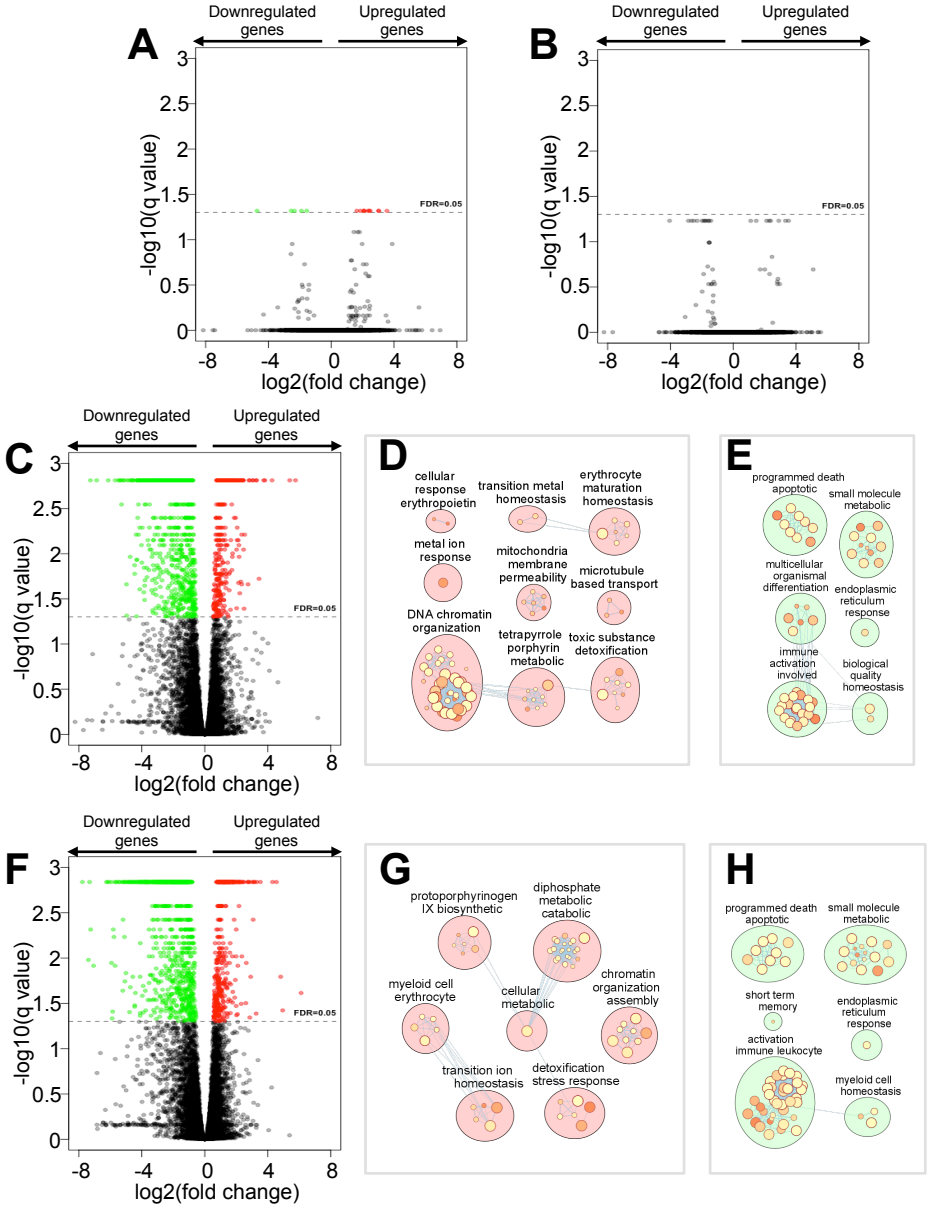
treatment upregulated genes involved in erythroid differentiation as well as chromatin organization, with downregulation of genes involved in immune activation and myeloid cell homeostasis (Figure 4.2 C-H). In addition, JQ1 upregulated *ALAS2* (Figure 4.3 A-B), which encodes the erythroid specific 5-aminolevulinate synthase 2, the enzyme that catalyzes the rate-limiting step of heme biosynthesis. Furthermore, I examined the expression of several lineage specific genes in TF-1 cells with or without EPO stimulation (Figure 4.3 C-H) and found that JQ1 treatment resulted in upregulation of erythroid-specific *transferrin receptor 1* (*TFRC*) while downregulating the gene encoding the hematopoietic stem and progenitor cell marker *CD34*, the myeloid TF gene *SPI1* (encoding PU.1), the macrophage marker *ITGAM* (encoding CD11b), and the megakaryocyte markers *ITGA2B* (encoding CD41) and *GP1BA* (encoding CD42), suggesting that JQ1 promotes erythropoiesis at the expense of other hematopoietic lineages.

Figure 4.1 JQ1 induces hemoglobin accumulation.



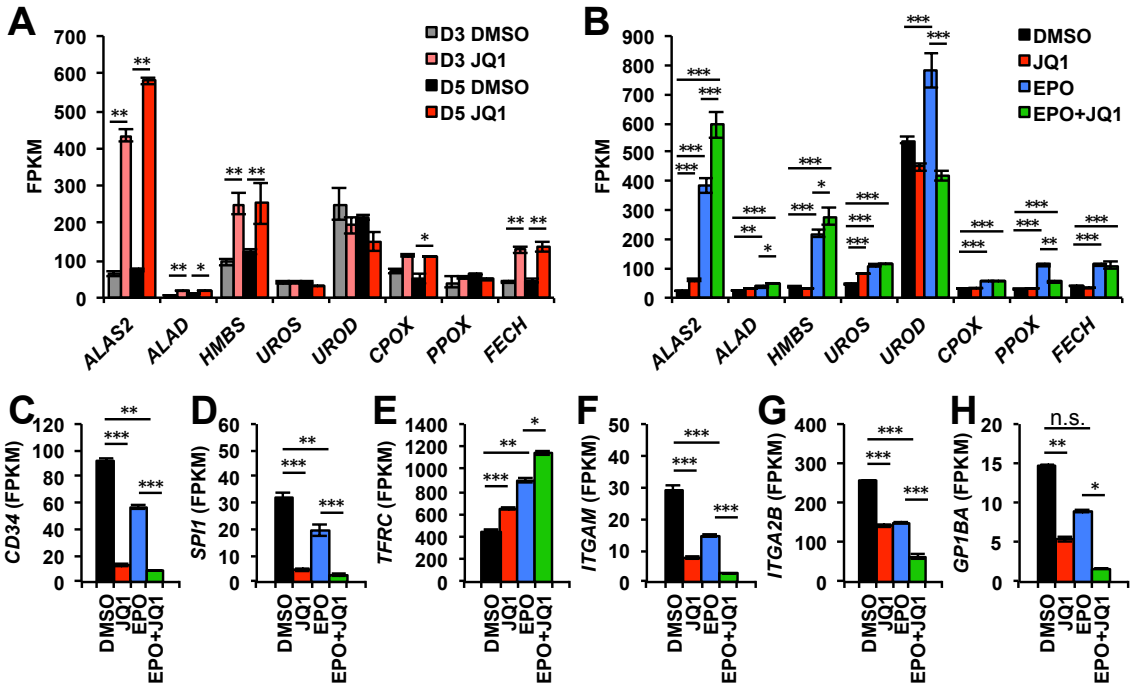
(**A-B**) Representative image of day 3 (**A**) DMSO or (**B**) JQ1-treated K562 cells. Cells were stained with benzidine (brown) to label hemoglobin, and hematoxylin (blue) to label DNA. (**C**) Quantification of percent hemoglobin+ cells at day 3 (D3) and 5 (D5) of K562 cells treated with DMSO (control, black) versus JQ1 (red) (N=3). **: $p<0.01$ (Student's t-test).

Figure 4.2 JQ1 induces erythropoietic expression profile.



(**A-B**) Volcano plot of gene expression in (**A**) day 0 (no treatment) versus day 3 DMSO (control); (**B**) day 0 (no treatment) versus day 5 DMSO (control). (**C**) Volcano plot of gene expression in K562 cells at day 3 after treatment with JQ1 versus DMSO (control). (**D-E**) GO term clustering of (**D**) upregulated or (**E**) downregulated genes from panel (**C**). (**F**) Volcano plot of gene expression in K562 cells at day 5 after treatment with JQ1 versus DMSO (control). (**G-H**) GO term clustering of (**G**) upregulated or (**H**) downregulated genes from panel (**F**). For volcano plots, red and green dots represent genes that are upregulated or downregulated, respectively. Black dots represent genes with no change in gene expression. N=2.

Figure 4.3 JQ1 upregulates heme-biosynthesis genes and downregulates genes encoding markers of myeloid differentiation.



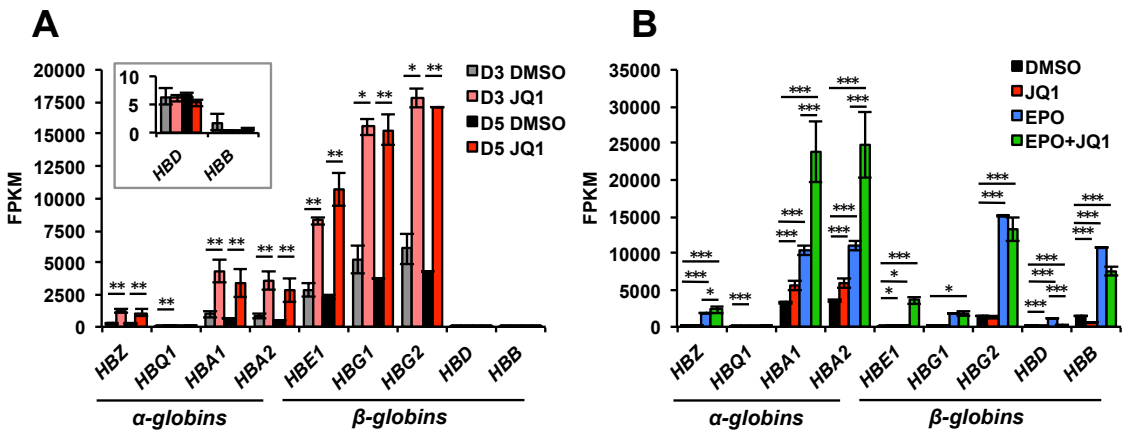
(**A**) Expression of genes involved in heme biosynthesis in K562 cells at day 3 (D3) and day 5 (D5) after treatment with JQ1 (red shades) versus DMSO (control, gray and black) (N=2). (**B**) Expression of genes involved in heme biosynthesis in TF-1 cells at day 3 after treatment with JQ1 (red), EPO (blue), EPO+JQ1 (green) versus DMSO (black) (N=2). (**C-H**) RNA-seq quantification of (**C**) *CD34*, which encodes a hematopoietic stem and progenitor cell marker; (**D**) *SPI1*, which encodes the myeloid lineage transcription factor PU.1; (**E**) *TFRC*, which encodes the erythroid lineage transferrin receptor; (**F**) *ITGAM*, which encodes the macrophage marker CD11b; (**G**) *ITGA2B*, which encodes the megakaryocyte marker CD41; and (**H**) *GP1BA,* which encodes the megakaryocyte marker CD42. Gene expression was measured in TF-1 cells after 3 days of treatment with JQ1 (red), EPO (blue), EPO+JQ1 (green) versus DMSO (control, black) (N=2). *: FDR<0.05, **: FDR<0.01, ***: FDR<0.001.

BET Inhibitors Upregulate Fetal and Embryonic *β-Globin* Expression in Erythroid Cell Lines

Next, I examined the expression of *β-globin* genes. Interestingly, in both K562 and TF-1 cells, JQ1 upregulated the embryonic and fetal *ε/γ-globin* genes, although it did not increase the expression of the adult *β-globin* gene *HBB* (Figure 4.4). To validate

these results, I designed qPCR primers to distinguish adult (*HBB*), fetal (*HBG1/2*), and embryonic (*HBE1*) *β-globin* gene expression. I excluded the adult *δ-globin* gene *HBD,* since its expression is very low compared to the other genes in this locus (Figure 4.4). I treated three different erythroid cell lines (K562, TF-1, and HEL) and one myeloid cell line (HL-60) with JQ1 for up to 5 days and measured the expression of each *β-like globin* gene every 24 hours (Figure 4.5). In K562 and HEL cells, both *HBE1* and *HBG1/2* were upregulated under JQ1 treatment, whereas *HBB* expression remained unchanged (Figure 4.5 A-C,G-I). In TF-1 cells, EPO stimulation led to upregulation of all three types of *β-like globin* genes, whereas JQ1 treatment only upregulated the embryonic *ε-globin HBE1* (Figure 4.5 D-F). Strikingly, when TF-1 cells were treated with both EPO and JQ1, *HBE1* expression increased over 200-fold by day 5 compared to DMSO control-treated cells, and *HBB* expression was decreased compared to those treated only with EPO (Figure 4.5F). In contrast, JQ1 treatment in the myeloid cell line HL-60 did not change expression levels of any of the *β-like globin* genes (Figure 4.5 J-L). I then calculated the composition of *β-like globin* transcripts in K562 and TF-1 cells from the qPCR results. K562 expresses very little *HBB*, and JQ1 treatment did not change the transcription ratio despite upregulating *HBG1/2* and *HBE1* (Figure 4.5M). In contrast, *HBB* transcripts constitute 40% of all *β-globin* cluster transcripts in TF-1 cells at baseline, and JQ1 treatment reduced the proportion of *HBB* transcripts while increasing *HBE1* transcripts with or without EPO treatment (Figure 4.5N).

Figure 4.4 JQ1 upregulates adult *α-globin* genes and fetal/embryonic *β-globin* genes.

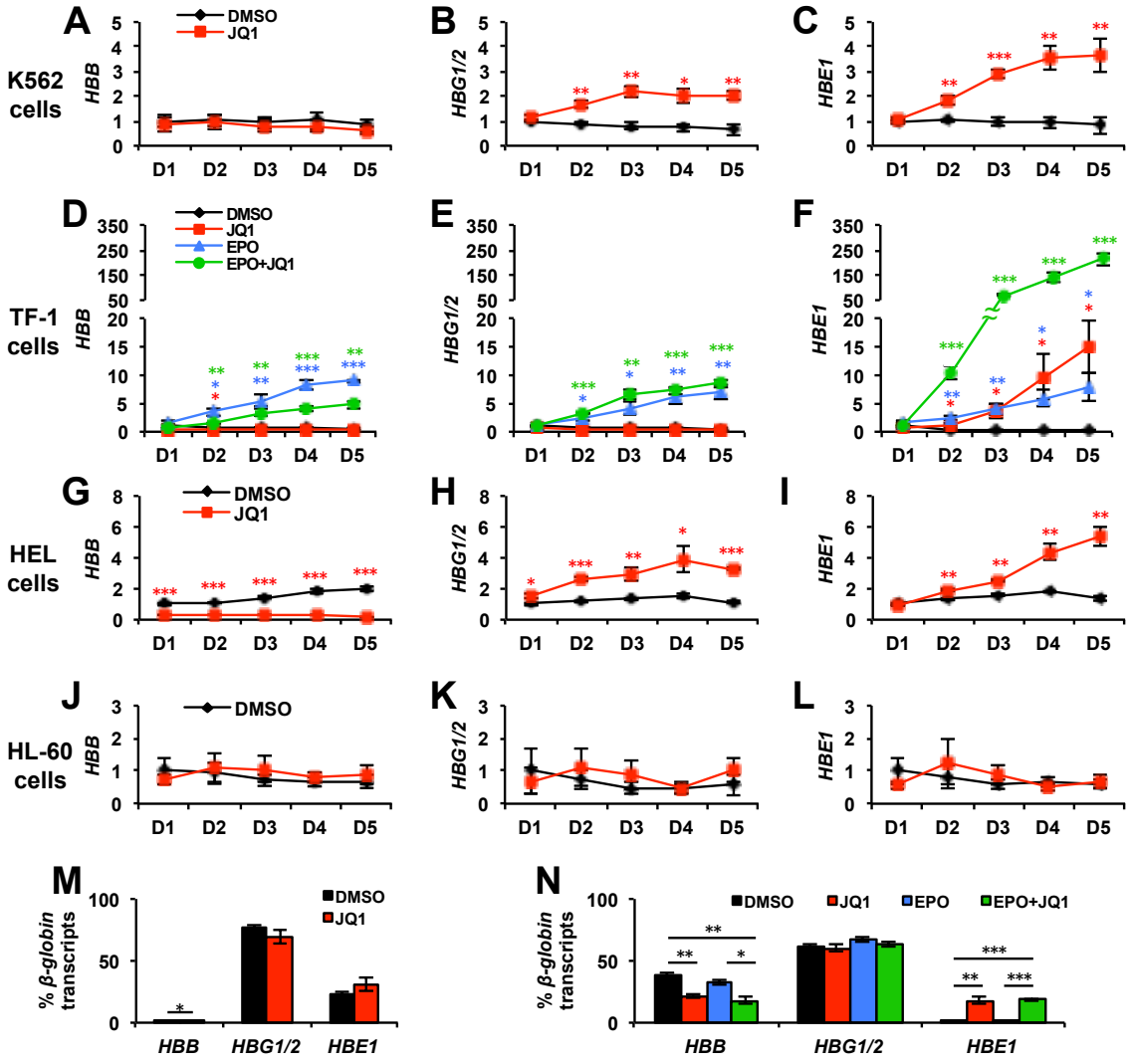


(**A**) Expression of genes in the α-*globin* or β-*globin* clusters in K562 cells at day 3 (D3) and day 5 (D5) after treatment with JQ1 (red shades) versus DMSO (control, gray and black). Inset: *HBD* and *HBG* expression, adjusted y-axis scale (N=2). *: FDR<0.05, **: FDR<0.01, ***: FDR<0.001. (**B**) Expression of genes in the α-*globin* or β-*globin* clusters in TF-1 cells at day 3 after treatment with JQ1 (red), EPO (blue), EPO+JQ1 (green) versus DMSO (black) (N=2). *: FDR<0.05, **: FDR<0.01, ***: FDR<0.001.

To determine if the observed effects were specific to JQ1 or were a class effect of BET inhibitors, I repeated the treatment of K562 and TF-1 cells using two other BET inhibitors: CPI-0610 and PLX51107. CPI-0610 is currently in phase II clinical trials for those with myelofibrosis or MDS in combination with ruxolitinib, and PLX51107 is in phase I clinical trials for individuals with AML or MDS in combination with azacitidine. Unlike JQ1, which has a higher affinity for BRD3 and BRD4 (Filippakopoulos et al., 2010), CPI-0610 binds to BRD2, BRD3, BRD4, and BRDT with comparable affinities, with a preference for binding the BD2 of the BET proteins (Albrecht et al., 2016). PLX51107 has a somewhat lower affinity for BRDT and preferentially binds to BD1 (Ozer et al., 2018). I found that both CPI-0610 and PLX51107 mimic the effects of JQ1 in K562 and TF-1 cells (Figure 4.6). In particular, PLX51107 suppressed *HBB* expression in TF-1 cells even without EPO treatment (Figure 4.6J), which was not observed with JQ1. These results suggest that, as a class, BET inhibitors can de-

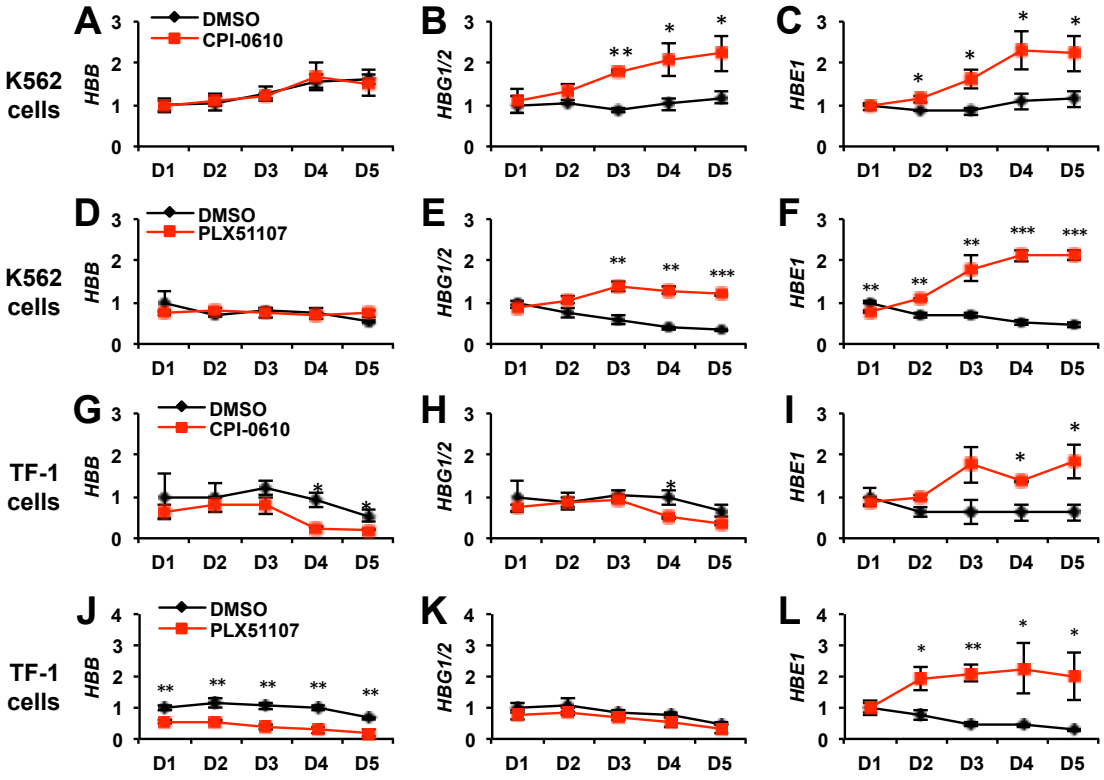
repress the fetal and/or embryonic *β-like globin* genes in cells poised for erythroid differentiation.

Figure 4.5 JQ1 induces fetal and embryonic *β-globin* genes in erythroid, but not myeloid, cell lines.



(**A-C**) RT-qPCR quantification of (**A**) *HBB*, (**B**) *HBG1/2*, and (**C**) *HBE1* expression in K562 cells. (**D-F**) RT-qPCR quantification of (**D**) *HBB*, (**E**) *HBG1/2*, and (**F**) *HBE1* expression in TF-1 cells. (**G-I**) RT-qPCR quantification of (**G**) *HBB*, (**H**) *HBG1/2*, and (**I**) *HBE1* expression in HEL cells. (**J-L**) RT-qPCR quantification of (**J**) *HBB*, (**K**) *HBG1/2*, and (**L**) *HBE1* expression in HL-60 cells. (**M**) Percentage of *HBB*, *HBG1/2*, and *HBE1* transcripts relative to total *β-globin* transcripts in K562 cells after 3 days of treatment. (**J**) Percentage of *HBB*, *HBG1/2*, and *HBE1* transcripts relative to total *β-globin* transcripts in TF-1 cells after 3 days of treatment. N=3 (all panels). Color representations In all panels: DMSO (black), JQ1 (red), EPO (blue), and EPO+JQ1 (green). *: p<0.05, **: p<0.01, ***: p<0.001 (Student's t-test). D1-D5: Days 1-5.

Figure 4.6 Several BET inhibitors exhibit similar effects to JQ1 in erythroleukemia cell lines.



(**A-C**) The expression of (**A**) *HBB,* (**B**) *HBG1/2*, and (**C**) *HBE1* in K562 cells treated with CPI-0610 (red) versus DMSO (control, black) (N=3). (**D-F**) The expression of (**D**) *HBB*, (**E**) *HBG1/2*, and (**F**) *HBE1* in K562 cells treated with PLX51107 (red) versus DMSO (control, black) (N=3). (**G-I**) The expression of (**G**) *HBB,* (**H**) *HBG1/2*, and (**I**) *HBE1* in TF-1 cells treated with CPI-0610 (red) versus DMSO (control, black) (N=3). (**J-L**) The expression of (**J**) *HBB,* (**K**) *HBG1/2*, and (**L**) *HBE1* in TF-1 cells treated with PLX51107 (red) versus DMSO (control, black) (N=3). D, days of treatment. *: p<0.05, **: p<0.01, ***: p<0.001 (Student's t-test).

α-Globin Genes were Upregulated by JQ1

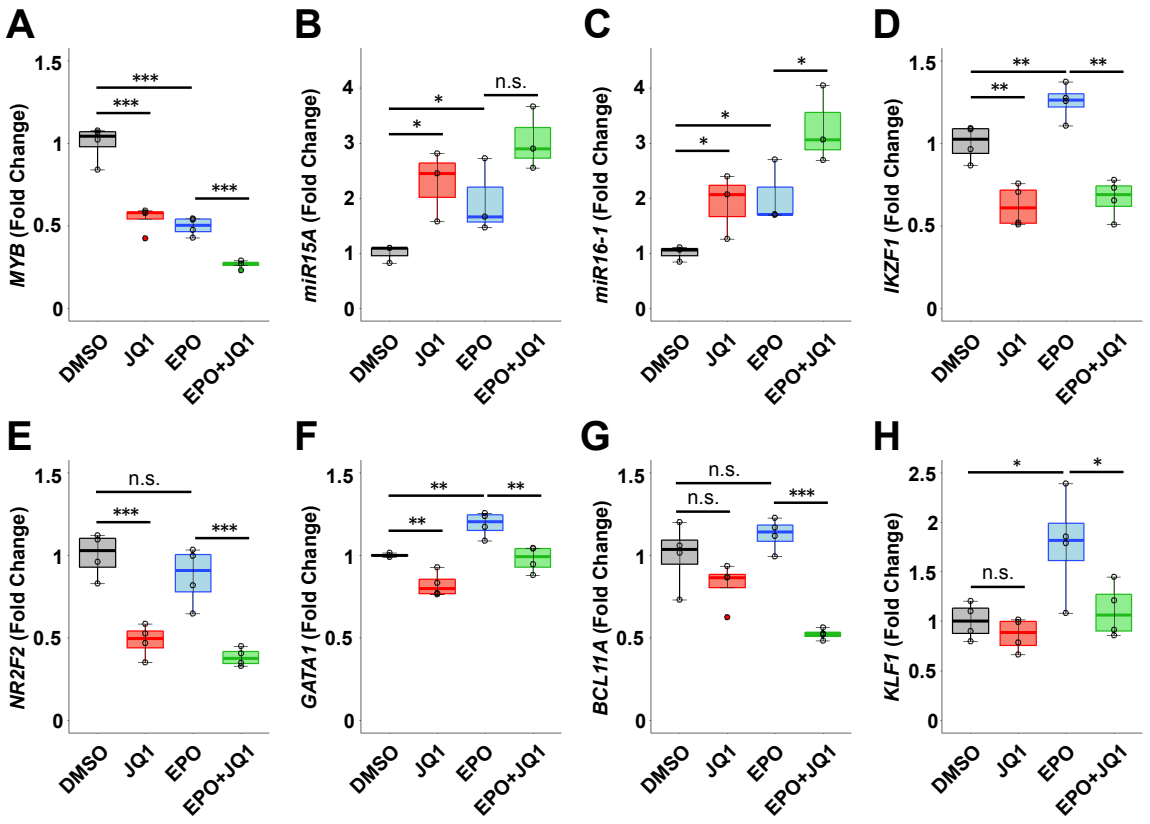
Balanced production of α– and β–globin chains is required to avoid inducing a thalassemia-like condition (Mettananda and Higgs, 2018; Taher et al., 2018). Because up-regulation of *β-globins* without a similar effect on *α-globin* gene expression could

result in an α-thalassemia-like syndrome, I examined the RNA-seq data to determine if JQ1 exerted a similar de-repressive effect on the *α-globin* gene cluster (Vernimmen, 2014). In both K562 and TF-1 cells, JQ1 upregulated the adult *α-globin* genes (*HBA1/2*) (Figure 4.4). Thus, because JQ1 results in de-repression of *β- and α-globin* genes in parallel, treatment of patients with β-globinopathies like SCA and β-thalassemia is unlikely to generate an α-thalassemia-like condition.

JQ1 Downregulates Known Inhibitors of Embryonic and Fetal Globin Genes

To investigate the molecular mechanisms leading to the specific upregulation of fetal/embryonic *β-like globin* genes, I examined the expression of genes encoding proteins known to inhibit ε- and γ-globin production. Because TF-1 cells express adult *β-globin* and thus provide a better model of human erythropoiesis than other erythroid cell lines, I used TF-1 cells for these subsequent mechanistic studies. I found that *MYB* expression was decreased by JQ1 and EPO, with an additive effect seen when the two treatments were combined (Figures 4.7A, 4.8A). Since *MYB* is a target of *miR-15A* and *miR-16-1*, I then measured the levels of these microRNAs and found that both were upregulated by either JQ1 and EPO in an additive fashion (Figure 4.7 B-C), showing an inverse correlation compared to *MYB* expression as expected. Similar expression changes were also observed in K562 cells, consistent with our RNA-seq data (Figure 4.8 B-D). I then examined the expression of known *miR-15A/16-1* targets and found that a majority were downregulated by treatment with JQ1 and/or EPO (Table 4.1), as expected from the *miR* upregulation.

Figure 4.7 JQ1 downregulates known inhibitors of fetal globin.

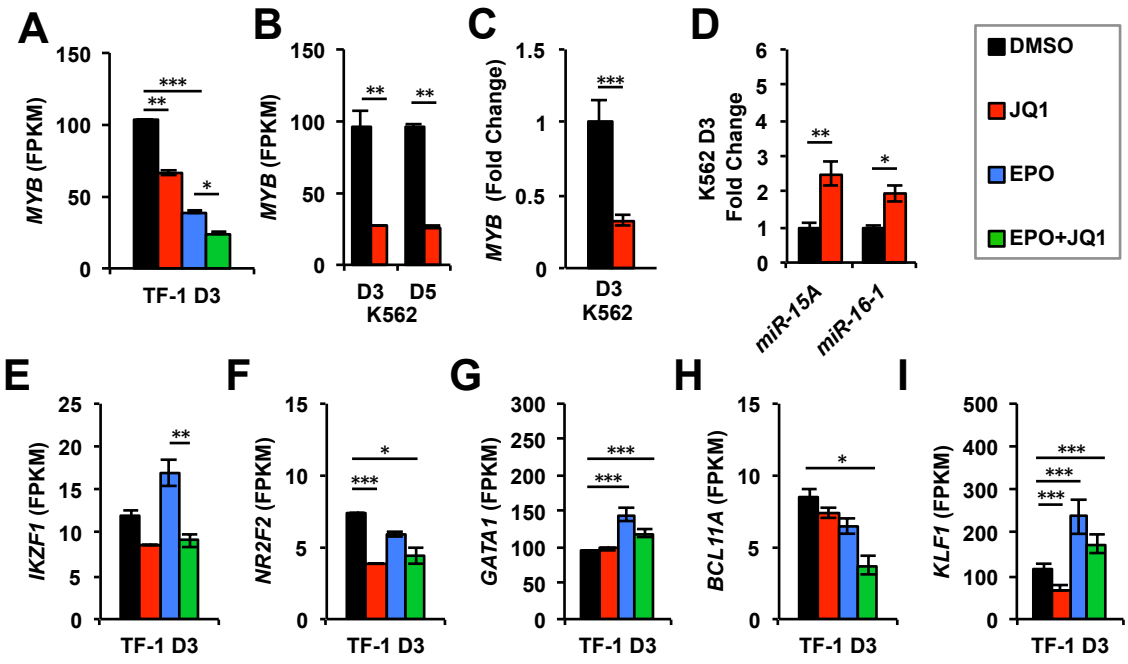


qPCR quantification of (**A**) *MYB*, (**B**) *miR-15A* , (**C**) *miR-16-1*, (**D**) *IKZF1*, (**E**) *NR2F2*, (**F**) *GATA1*, (**G**) *BCL11A*, and (**H**) *KLF1* in TF-1 cells after treatment with JQ1 (red), EPO (blue), EPO+JQ1 (green) versus DMSO (control, black) for 3 days. N=3 for panels (**B**) and (**C**). N=4 for all other panels. *: p<0.05, **: p<0.01, ***: p<0.001, n.s.: not statistically significant (Student's t-test).

Next, I examined the genes encoding several well-established *HBE1* and *HBG1/2* inhibitors: *IKZF1 (*which encodes IKZF1/IKAROS*)*, *NR2F2* (which encodes COUP-TF2), and *GATA1*. Each of these genes was downregulated by JQ1 with or without EPO (Figure 4.7 D-F). KLF1 promotes the expression of *BCL11A* (Zhou et al., 2010), and therefore as expected, I observed downregulation of *BCL11A* only when we also saw *KLF1* repression in the setting of joint JQ1 and EPO treatment (Figure 4.7 G-

H). RNA-seq data from TF-1 cells were consistent with our qPCR quantifications (Figure 4.8 E-I).

Figure 4.8 JQ1 suppresses the expression of genes encoding fetal/ embryonic *β-globin* inhibitors.



(**A**) RNA-seq quantification of *MYB* in TF-1 cells after 3 days of treatment (N=2). (**B**) RNA-seq quantification of *MYB* expression in K562 cells after 3- (D3) or 5-day (D5) treatment (N=2). (**C**) qPCR quantification of *MYB* expression in K562 cells after 3 days of treatment (N=4). (**D**) qPCR quantification of *miR-15A* and *miR-16-1* in K562 cells after 3 days (N=3). (**E-I**) RNA-seq quantification of (**E**) *IKZF1*, (**F**) *NR2F2*, (**G**) *GATA1*, (**H**) *BCL11A*, and (**I**) *KLF1* in TF-1 cells after 3 days of treatment (N=2). Color representations in all panels: DMSO (black), JQ1 (red), EPO (blue), and EPO+JQ1 (green). (**A, E-I**) *: FDR<0.05, **: FDR<0.01, ***: FDR<0.001. (**C-D**) *: p<0.05, **: p<0.01, ***: p<0.001 (Student's t-test).

Table 4.1. Expression changes of known miR-15A and miR-16-1 targets in K562 and TF-1 cells.

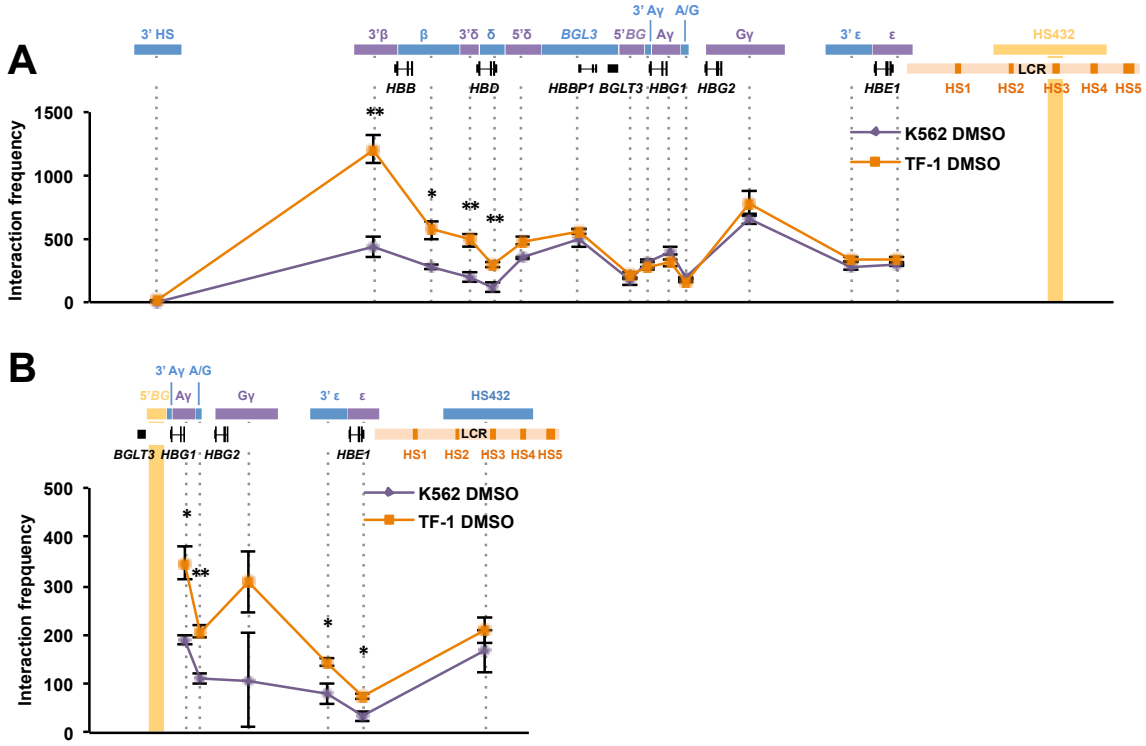
N=2 for all data shown.

	K562	TF-1		
Gene	JQ1 vs. DMSO	JQ1 vs. DMSO	EPO+JQ1 vs. EPO	EPO vs. DMSO
MYB	-1.80	-0.64	-0.68	-1.40
BCL2	-1.89	-2.09	Low FPKM	-3.08
MCL1	0.02	0.85	0.91	0.17
BMI1	0.14	0.33	0.06	-0.16
CCND1	-4.21	-1.71	-2.26	-0.26
CCND2	-2.68	-1.02	-0.37	-0.98
CCNE1	0.04	0.11	-0.63	0.47
SOX5	-3.96	Low FPKM	Low FPKM	Low FPKM
TGFB3	-1.51	-1.75	-1.98	-0.28
SMAD2	0.12	-0.01	0.00	-0.32
SMAD3	-0.82	-1.23	-0.72	-1.47
VEGFA	-1.39	-2.85	-1.08	-2.13
GP6	-1.05	-2.76	-2.68	-1.01
FYN	-0.15	2.34	2.38	0.57
SRGN	-1.16	-1.03	-1.35	-0.44
FCER1G	-0.23	0.03	-1.89	0.21
PRKCQ	-0.77	-0.25	-0.59	-0.44
WT1	-0.94	-0.76	-0.81	-0.59
PDCD4	-0.67	0.00	-0.16	0.08
RAB21	0.20	0.48	0.43	-0.10
SKAP2	-0.81	0.16	-0.25	-0.06
AKT3	-0.40	Low FPKM	Low FPKM	Low FPKM

JQ1 Modulates Chromatin Spatial Organization at the *β-Globin* Locus.

Previous studies showed that looping between the promoter of each *β-globin* gene and the LCR enhancer is mediated by the bromodomain proteins (Lamonica et al., 2011; Stonestrom et al., 2015). I thus reasoned that BET inhibitors would decrease interactions between the LCR and the *β-globin* genes and performed chromatin conformation capture followed by qPCR (3C-qPCR) (Hagège et al., 2007) to examine how the spatial structure of this locus changes in response to JQ1 treatment. Comparing K562 and TF-1 baseline interaction frequencies, I found that the LCR interaction with *HBB* and *HBG* in TF-1 cells was higher than that in K562 cells (Figure 4.9).

Figure 4.9 Baseline interaction between LCR/*BGLT3* with *β-globin* loci.



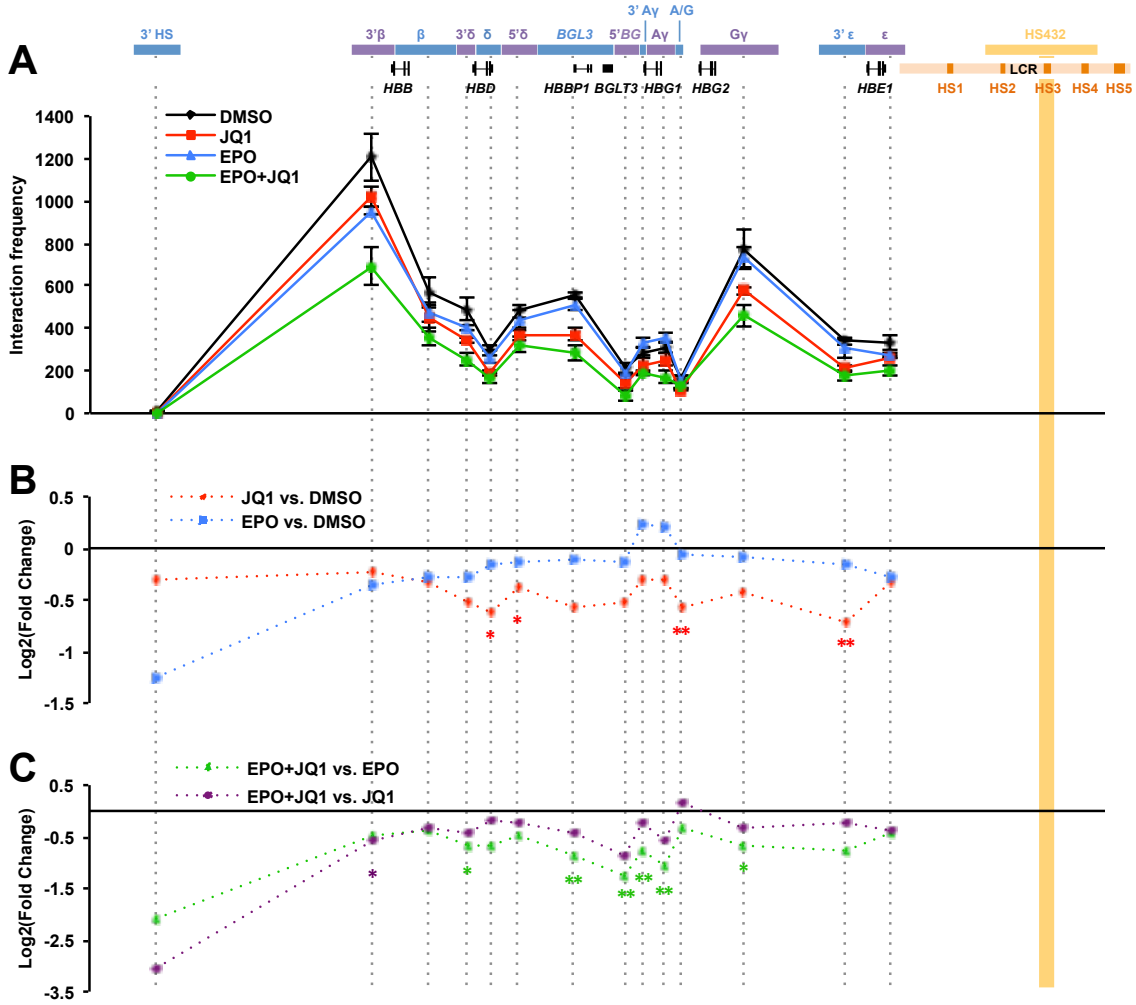
(**A**) 3C-qPCR quantification of LCR interaction frequency with various *β-globin* loci in untreated K562 (purple) and TF-1 cells (orange) (N=3). (**B**) 3C-qPCR quantification of *BGLT3* locus interaction frequency with the fetal/embryonic *β-globin* genes in untreated K562 (purple) and TF-1 (orange) cells (N=3). *: p<0.05, **: p<0.01 (Student's t-test).

Again, I focused our analysis on TF-1 cells since they better resemble adult erythroid lineage cells. I quantified the LCR interaction with the *β-globin* genes in all treatment groups (Figure 4.10). Interestingly, comparing the effect of EPO or JQ1 to DMSO control-treated cells, I found that although EPO did not change the interaction frequency, JQ1 induced less interaction at multiple loci near or in the fetal and adult *β-globin* genes (Figure 4.10 B-C). The same differences were observed with combined EPO and JQ1 treatment compared to either EPO or JQ1 single treatments, with more substantial decreases seen with JQ1 (Figure 4.10 B-C). Notably, neither JQ1 nor EPO decreased the interaction between the LCR and the *HBE1* promoter (Figure 4.10 B-C,

right most data point), suggesting that JQ1 treatment relaxes the looping between the LCR and the adult/fetal *β-globin* genes, which biases LCR contacts in favor of the promoter of embryonic *HBE1*.

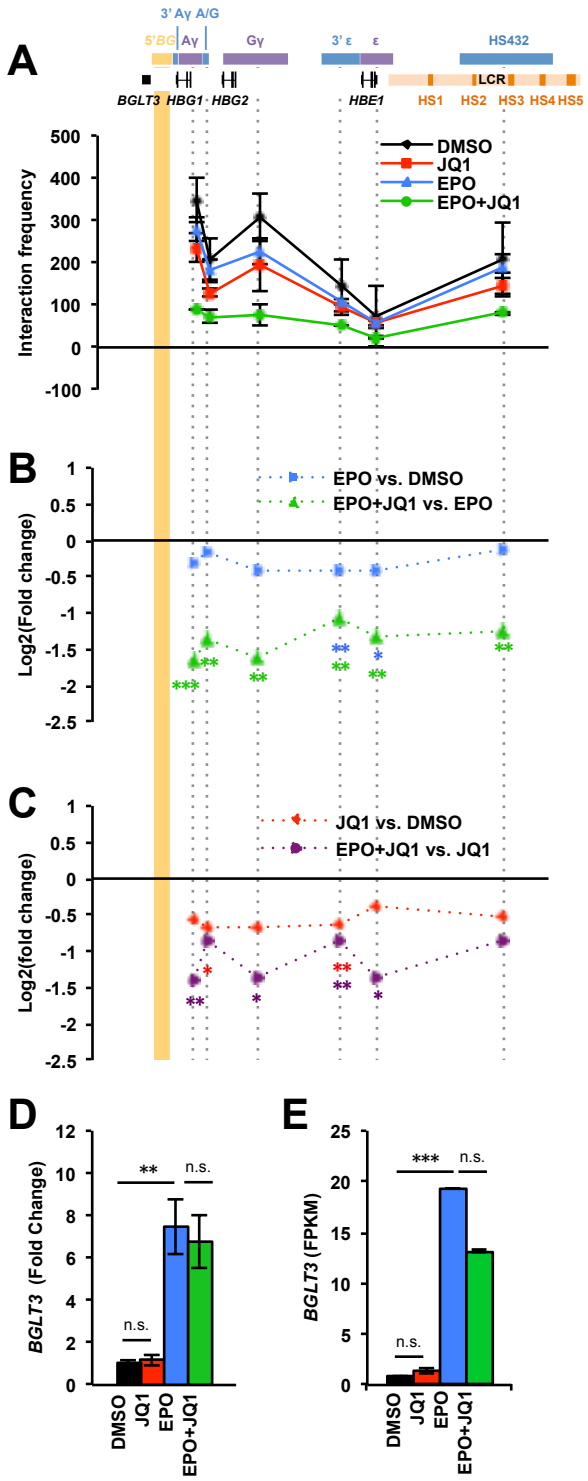
The *BGLT3* lncRNA promotes looping between *HBG1/2* and the LCR (Ivaldi et al., 2018), and therefore, I next examined the expression of the *BGLT3* lncRNA as well as the interaction between the *5'BGLT3* (5'BG) locus and its upstream loci. In TF-1 cells, JQ1 or EPO alone caused slight decreases in the interactions among 5'BG and the other *β-globin* genes (Figure 4.11 A-C). However, combined treatment drastically decreased the interactions within this chromatin loop (Figure 4.11 A-C), suggesting cooperative effects between JQ1 and EPO. Interestingly, although JQ1 did not change *BGLT3* expression (Figure 4.11 D-E), consistent with the effect of JQ1 on the loop structure (Figure 4.11 A-C), EPO increased *BGLT3* expression by 7-fold (Figure 4.11 D-E). The lack of an increase of interaction within this chromatin loop under EPO treatment suggests that there are other EPO-related factors counteracting the effects of *BGLT3*.

Figure 4.10 JQ1 changes interaction frequencies between the LCR and the *β-globin* genes.



(**A**) 3C-qPCR quantification of interaction frequencies between the LCR and segments of the *β-globin* locus in TF-1 cells after treatment with JQ1 (red), EPO (blue), EPO+JQ1 (green) versus DMSO (control, black) for 3 days. (N=3). (**B**) log2 fold-change and statistical significance comparing JQ1 versus DMSO control (red) and EPO versus DMSO control (blue). (**C**) log2 fold-change and statistical significance comparing EPO+JQ1 versus EPO (green) and EPO+JQ1 versus JQ1 (purple). Track on top indicates chromosomal gene positions and the DNA fragments generated from an EcoRI digestion. LCR: locus control region. HS: DNase hypersensitivity site. The EcoRI-generated fragment containing the bait primer is shown as a yellow bar. *: p<0.05, **: p<0.01, ***: p<0.001 (Student's t-test).

Figure 4.11 JQ1 changes the interaction frequencies between *BGLT3* and the fetal globin genes.



(**A**) 3C-qPCR quantification of the interaction frequencies between *BGLT3* and the fetal globin genes in TF-1 cells after treatment with JQ1 (red), EPO (blue), EPO+JQ1 (green) versus DMSO (control, black) for 3 days. (N=3). (**B**) log2 fold-change and statistical significance comparing EPO versus DMSO (blue) and EPO +JQ1 versus EPO (green). (**C**) log2 fold-change and statistical significance comparing JQ1 versus DMSO control (red) and EPO+JQ1 versus JQ1 (purple). Track on top indicates chromosomal gene positions and the DNA fragments generated from an EcoRI digestion. LCR: locus control region. HS: DNase hypersensitivity site. The EcoRI-generated fragment containing the bait primer is shown as a yellow bar. *: p<0.05, **: p<0.01, ***: p<0.001 (Student's t-test). (**D**) qPCR quantification of *BGLT3* expression in TF-1 cells after treatment with JQ1 (red), EPO (blue), EPO+JQ1 (green) versus DMSO (control, black) for 3 days (N=4). **: p<0.01 (**E**) RNA-seq quantification of *BGLT3* expression in TF-1 cells after treatment with JQ1 (red), EPO (blue), EPO+JQ1 (green) versus DMSO (control, black) for 3 days. (N=2). ***: FDR<0.001.

Thus, the *β-globin* gene expression changes in BET inhibitor-treated cells likely resulted from the combined effects of chromatin changes at the *β-globin* locus and downregulation of genes encoding *γ/ε-globin* gene inhibitors. Overall, the shifts in LCR interactions favoring *HBE1* and the decreased *γ/ε-globin* inhibition result in decreased *β-globin* transcripts and increased *ε-globin* transcripts (Figure 4.5N). JQ1 in combination with EPO also disrupts the *BGLT3*-mediated *γ-globin* looping, which likely further shifts the LCR interaction towards *HBE1* and contributes to the synergistic upregulation of *HBE1* (Figure 4.5F).

DISCUSSION

Hemoglobin switching in human development has fascinated biologists and physicians for many decades given its immense clinical implications (Sankaran and Orkin, 2013; Sankaran et al., 2010; Stamatoyannopoulos, 2005). *β-globin* gene switching is a prime example of epigenetic regulation of gene expression, and the understanding of local chromatin structure and TF binding makes it possible to devise new strategies to reactivate gene expression for positive clinical effects (Ali et al., 2020; Houwing et al., 2019). In this study, I demonstrate that BET inhibitors specifically upregulate *HBE1* and *HBG1/2*, but have no effect on or even suppress *HBB* in multiple erythroid cell lines. My finding suggests that specific inhibition of bromodomain-containing proteins could be used to treat β-globinopathies like SCA and β-thalassemia.

SCA is the most common inherited disorder in the United States and is especially prevalent in many parts of Africa where malaria is common. The most common treatment for SCA is administration of hydroxyurea, which elevates fetal hemoglobin

production to alleviate symptoms (Ali et al., 2020; Houwing et al., 2019). In late 2019, two additional drugs were FDA-approved to treat SCA: voxelotor and crizanlizumab. Voxelotor enhances the affinity of hemoglobin to oxygen, reducing the risk of sickle hemoglobin polymerization in hypoxia (Ali et al., 2020). Crizanlizumab is a monoclonal antibody against P-selectin, which plays a major part in initiating vaso-occlusive crises, a common clinical manifestation of SCA (Ali et al., 2020). Neither of these drugs reduces the amounts of the mutant hemoglobin itself. In recent years, multiple gene therapy strategies for SCA have emerged, including *β-globin* gene addition and nuclease-assisted *β-globin* gene modification/repair (Hoban et al., 2016; Ikawa et al., 2019; Orkin and Bauer, 2019). Although these recent advances in gene therapy offer the potential for cure, they are extremely expensive, making them unaffordable for low-income patients and unattainable for less affluent countries or health systems. My findings suggest that BET inhibitor treatment could provide a reasonable alternative and/or adjunct to the treatment of SCA and other disorders caused by mutations of the *HBB* gene.

My data in TF-1 cells show specific upregulation of the embryonic hemoglobin gene *HBE1*. An important clinical correlate is whether the *HBE1* encoded ε-globin could function as a reasonable substitute for the adult β-globin chain. Biochemical analysis of the embryonic hemoglobin Hb-Gower 2 ($\alpha_2\varepsilon_2$) show that its P_{50} for oxygen, affinity to 2,3-BPG, Bohr Coefficient, and Hill Coefficient are comparable to those of HbA (He and Russell, 2001) (Table 4.2). Hb-Gower 2 also has a comparable tetramer-dimer disassociation constant to that of HbA (Manning et al., 2007). A study in transgenic α/β-thalassemia mice found that human embryonic hemoglobins consist of ζ-globin and ε-

globin rescue the lethal phenotype of α/β-thalassemia (Russell and Liebhaber, 1998). Similarly, a study in sickle cell mice found that the presence of human Hb-Gower 2 ($\alpha_2\varepsilon_2$) greatly alleviated sickle cell phenotypes, and Hb-Gower 2 inhibits sickle cell hemoglobin (HbS) polymerization (He and Russell, 2002). Taken together, these studies indicate that reactivation of ε- and/or γ-globin could be beneficial in treating SCA and β-thalassemia.

Table 4.2 Comparisons of oxygen affinity and stability in adult and embryonic hemoglobins.

Name	Composition	P_{50} for O_2 (torr)[a]	Hill Coefficient[a]	Tetramer-dimer K_d (μM)[b]
HbA	$\alpha_2\beta_2$	3.2	2.9	0.68
Hb-Gower 1	$\zeta_2\varepsilon_2$	1.4	1.7	2.14
Hb-Gower 2	$\alpha_2\varepsilon_2$	2.7	2.3	0.17

[a] He Z, and Russell JE. Expression, purification, and characterization of human hemoglobins Gower-1 (zeta(2)epsilon(2)), Gower-2 (alpha(2)epsilon(2)), and Portland-2 (zeta(2)beta(2)) assembled in complex transgenic-knockout mice. *Blood.* 2001;97(4):1099-105.

[b] Manning LR, Russell JE, Padovan JC, et al. Human embryonic, fetal, and adult hemoglobins have different subunit interface strengths. Correlation with lifespan in the red cell. *Protein Sci.* 2007;16(8):1641-58.

DECLARATION OF INTERESTS

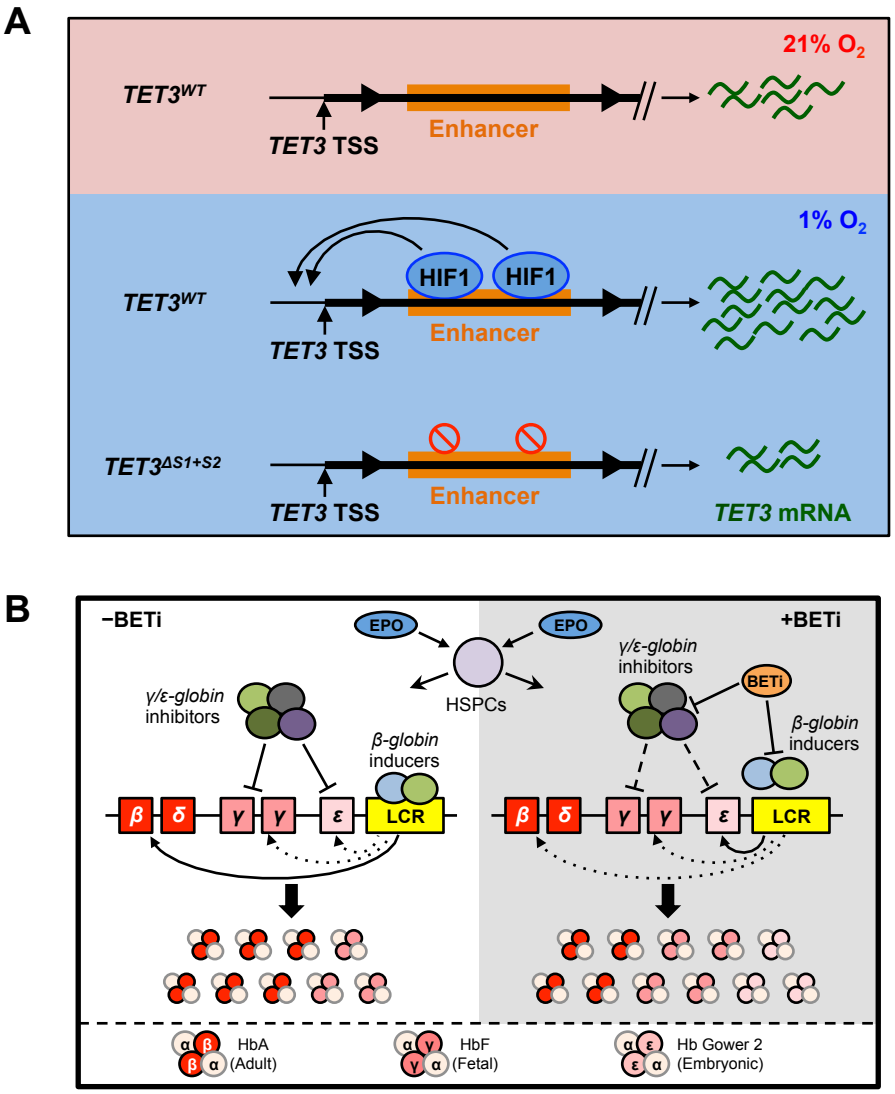
L.A.G, A.W., and J.Z.C are owners of the patent PCT/US20/52842 titled "METHODS AND COMPOSITIONS FOR TREATING SICKLE CELL DISEASE AND THALASSEMIA", filed on September 25, 2020.

CHAPTER V

Discussion

In the previous chapters, I presented results that support a model that the erythropoietic epigenetic landscape is intricately linked with cellular metabolism and signaling, and that the elements regulating the epigenome can be targeted for favorable clinical outcomes (Figure 5.1). In recent years, other studies have produced results that support my observations. These studies have highlighted that: (i) TET2 and TET3 have distinct roles in regulating erythropoiesis and are controlled by unique regulatory mechanisms (Li et al., 2015; Sohni et al., 2015; Wu et al., 2018; Yan et al., 2017; Zhang et al., 2014; Zhao et al., 2015); (ii) TET activity is closely linked to energy metabolism, particularly related to mitochondrial oxidative glycolysis (Agathocleous et al., 2017; Laukka et al., 2015; Minor et al., 2013; Moshfegh et al., 2019; Raffel et al., 2017; Xu et al., 2011); and (iii) targeting epigenetic regulators is a valid strategy in treating hematopoietic malignancies and other diseases influenced by epigenomic modifications (Chaidos et al., 2015; Cimmino et al., 2017; Lamonica et al., 2011; Ozer et al., 2018; Roe et al., 2015). Many of these findings were corroborated by studies in non-hematopoietic tissues (Chen et al., 2017; Dawlaty et al., 2014; Jäwert et al., 2013; Kang et al., 2015a; Kumar et al., 2015; Mariani et al., 2014; Misawa et al., 2018; Pan et al., 2017; Rudenko et al., 2013; Vincent et al., 2013), demonstrating the universal roles of TETs and 5-hmC in maintaining proper differentiation and normal cellular functions. Thus, understanding the molecular mechanisms controlling covalent cytosine modifications provides insights into cellular transcriptional regulation and clinical opportunities to target relevant epigenetic pathways to treat various diseases.

Figure 5.1 The epigenetic regulations in erythropoiesis.

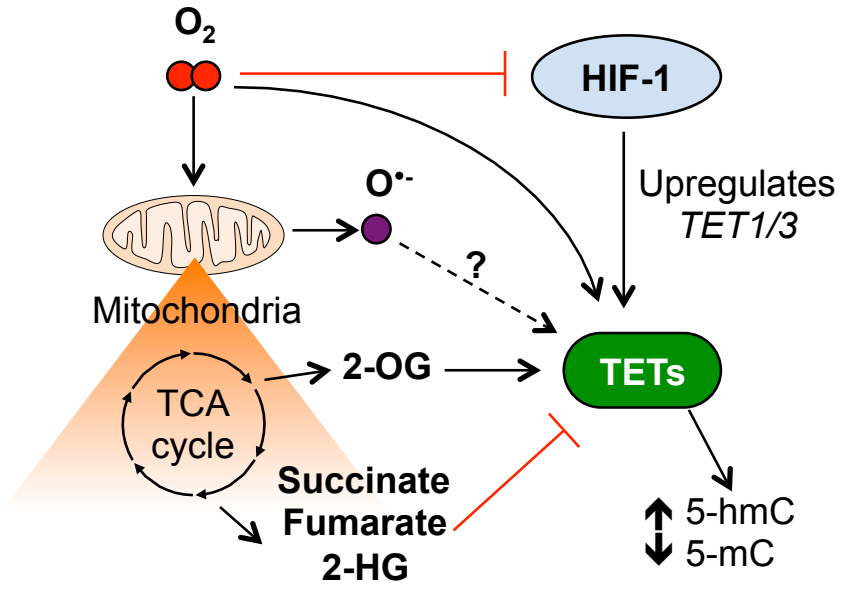


(**A**) Adapted from (Cao et al, 2020). Hypoxia leads to *TET3* upregulation by HIF-1α. Two binding sites within a *TET3* intronic enhancer are bound by HIF-1 in hypoxia. Deletion of either site abolishes *TET3* upregulation, and double deletion leads to a decrease in *TET3* expression. (**B**) BET inhibitors lead to re-expression of fetal/embryonic *β-globins*. BET inhibitors downregulate genes encoding fetal/embryonic *β-globin* suppressors and disrupt chromatin looping at the *β-globin* locus, leading to re-expression of silenced *β-globin* genes, particularly embryonic *HBE1* (*ε*).

Oxygen availability, mitochondrial energy production, and TET activities

Dioxygenase activity is inherently linked to the availability of molecular oxygen and 2-OG as substrates. Therefore, dioxygenase is related to cellular metabolic pathways that generate or consume 2-OG, such as the TCA cycle and the electron transport chain in mitochondria (Figure 5.2). The TCA cycle also produces molecules that inhibit dioxygenase activity: succinate, fumarate, and 2-HG (Figure 5.2). Specific to the *TET* genes, low-oxygen availability leads to upregulation of *TET1* or *TET3* via HIF-1 depending on the cell types (Cao et al., 2020) (and unpublished data from Anastasia Hains, Godley Laboratory), forming a feedback loop in which cells compensate for low oxygen concentration by increasing TET enzyme production and/or stability (Figure 5.2). In addition, TET2 phosphorylation by AMPK stabilizes the protein, which is disrupted in hyperglycemic conditions due to suppression of AMPK activity, leading to decreased overall TET availability (Wu et al., 2018).

Figure 5.2 Oxygen and mitochondrial metabolism are intimately tied to TET activities.



Schematic model of how TET activity is influenced by oxygen-dependent pathways. Oxygen is consumed directly by the TETs in their enzymatic reaction. Low oxygen levels lead to HIF-1α stabilization, allowing HIF-1 to accumulate and thereby upregulate *TET1* or *TET3* depending on the cell type. The oxygen-dependent TCA cycle produces 2-OG, which is also directly used in TET reactions, as well as succinate, fumarate, and 2-HG, which inhibit TET enzymatic activities. The oxygen free radical ($O^{\bullet-}$) increases global 5-hmC level, although a direct link between $O^{\bullet-}$ and TET activity has not been established.

The link between TET catalytic activity and oxygen-related metabolic processes has profound implications in systems beyond hematopoiesis. In normal human physiology, most cells reside in varying degrees of hypoxia, except for skin and lung epithelial cells (Keeley and Mann, 2019). Most cells inside the body experience levels of oxygen referred to as “physiological normoxia”, which range from 3 to 12% oxygen (Keeley and Mann, 2019). Oxygen levels below physiological normoxia are also found in certain cell populations, notably HSPCs in the bone marrow despite its highly vascularized environment (Nombela-Arrieta et al., 2013; Spencer et al., 2014). Hypoxia

has been shown to maintain the stem cell state of HSPCs, facilitate HSPC expansion, and promote their mobilization (Bisht et al., 2019; Jackson et al., 2017). Hypoxia and inhibition of aryl hydrocarbon receptor (AhR) produces similar effects in HSPCs, which is due to HIF-1α competing with AhR for HIF-1β (ARNT) binding (Jackson et al., 2017). Regions of solid tumors also experience hypoxia due to a lack of organized vasculature for efficient oxygen transport (Nakazawa et al., 2016). Importantly, most cell culture experiments are performed under atmospheric oxygen, which does not replicate the native environment of the cells and likely biases the transcriptome due to a shifted epigenetic landscape. To study normal and abnormal epigenomes under more accurate cell physiology, *in vitro* experiments, especially ones involving primary cells, should be performed, under oxygen levels that reflect their native niches whenever possible.

Finally, the multitude of substrate and cofactors necessary for TET catalytic activity provides a direct link between nutrient levels and epigenetic changes. Specifically, Fe(II) and ascorbate (vitamin C) levels directly affect TET activities (Chen et al., 2013; Chung et al., 2010; Minor et al., 2013). In particular, ascorbate has been shown to enhance TET catalytic activities in multiple studies (Agathocleous et al., 2017; Blaschke et al., 2013; Chen et al., 2013; Chung et al., 2010; Cimmino et al., 2017; Minor et al., 2013). For example, treating *Tet2*-deficient mouse HSPCs with ascorbate resulted in increased 5-hmC and decreased leukemic colony forming capability in these cells, mimicking the effect of restoring normal *Tet2* (Cimmino et al., 2017). This suggests that activities of the other Tets can compensate for the loss of Tet2 when supplemented with ascorbate. Additional research in the connection between these

substrates/cofactors and TET catalytic activities may shed light on how nutrient deficiency affects the epigenome in various cell types.

Modulation of TET activity at translational and post-translational levels

The results presented in Chapter III, as well as previous research by the Godley laboratory, have shown that *TET1* and *TET3* are transcriptionally upregulated in response to hypoxia (Cao et al., 2020; Mariani et al., 2014)(unpublished data from Anastasia Hains, Godley Laboratory). A recent study showed that hypoxia-induced glucose-6-phosphate dehydrogenase is involved in downregulating mouse *Tet2* in lungs in the context of hypoxia-induced pulmonary hypertension (Joshi et al., 2020). However, our understanding of additional factors regulating *TET* expression in normal and malignant cells remains poor, and the exact relationships between transcription levels and protein levels remain to be defined. For example, my analysis showed that Site 2 of *TET3* is in the proximity of multiple TF binding sites, including GATA1, GATA2, TAL1, EP300, and REST complex (Figure 3.14). The co-localization of these TFs within the *TET3* enhancer suggests the presence of a protein complex controlling normal transcription of *TET3* at this region. These TF binding sites in *TET3* likely determine its lineage-specific transcription. Further studies of the transcriptional regulation of *TET* genes may reveal more insights in how the 5-hmC epigenetic landscape is modified and maintained in various cell types depending on the combination of *TET* genes expressed.

TET transcripts are also regulated via microRNA. A study in 2013 showed that microRNAs *miR-7*, *miR-26*, *miR-29*, *miR-101*, and *miR-129* target *TET2* transcripts, and

are often upregulated in *TET2* WT AML (Cheng et al., 2013). Among these microRNAs, *miR-26* and *miR-29* also target *TET1* and *TET3* transcripts (Cheng et al., 2013). Three other studies reported that *miR-22*, *let-7* family, and *miR-210* target *TET2* in hematopoietic cells, macrophages, and brain cells, respectively (Jiang et al., 2019; Ma et al., 2021; Song et al., 2013). More recently, various studies reported a multitude of microRNAs targeting *TET1* and *TET3*, including the *TET1*-related *miR-21*, *miR-129*, *miR-191*, and *miR-4284*, as well as the *TET3*-related *miR-27*, *miR-150*, and *miR-200* (Li et al., 2018; Lu et al., 2018; Ma et al., 2018; Selimoglu-Buet et al., 2018; Yang et al., 2020a; Yang et al., 2020b; Zhong et al., 2019). These results reveal the intricate regulatory networks controlling *TET* mRNA levels in various tissues.

We have shown that TET2 is phosphorylated by JAK2 in response to EPO signaling during erythropoiesis (Jeong et al., 2019). In addition to this study, many others have shown various PTMs on all three TET enzymes that affect enzymatic activity, protein stability, and binding partners. Known modifications include phosphorylation (Bauer et al., 2015; Wu et al., 2018), acetylation (Zhang et al., 2017b), ubiquitination (Nakagawa et al., 2015), and O-GlcNAcylation (Bauer et al., 2015; Hrit et al., 2018; Shi et al., 2013; Vella et al., 2013; Zhang et al., 2014). Beyond these validated modifications, there are many more PTMs identified by proteomic studies on all three TETs with unknown functional significance (Hornbeck et al., 2019). Further studies are needed to better understand the regulatory networks that control TET protein stability, activity, and binding partners.

As discussed in Chapter III, hypoxia results in stabilized HIF-1, which directly upregulates *TET3,* and results in increased 5-hmC (Cao et al., 2020). However, similar

to how JAK2 phosphorylates TET2 to increase its activity (Jeong et al., 2019), it is highly likely that hypoxia induced 5-hmC increases are also due to hypoxia-associated PTMs of TET2 or TET3 that augment their activities. One possibility is that one or more of the TET proteins could undergo hydroxylation changes in a similar fashion to HIF-1α to increase protein stability or activity. Additional study is required to fully elucidate how PTM changes impact TET catalytic activities and its downstream effects on the epigenome.

Another area of interest is how TET PTMs affect TET binding partners and modify their target regions in the genome. As two studies of TET catalytic domain showed, TET enzymes do not recognize DNA flanking sequence of a CpG dinucleotide (Hu et al., 2013; Hu et al., 2015). However, the regions that gain 5-hmC are not random and are associated with enhancers, promoters, and gene bodies of upregulated genes (Hon et al., 2014; Madzo et al., 2014; Mariani et al., 2014; Stroud et al., 2011; Verma et al., 2018). These results indicate that TETs are guided to these regions by other proteins with sequence binding preferences. Some of the proteins or protein complexes that have been shown to bind TETs include MBD1 (Zhang et al., 2017a), the REST complex (Perera et al., 2015), OGT (Hrit et al., 2018), TDG (Hassan et al., 2017), CRL4 E3 ubiquitin ligase (Nakagawa et al., 2015), and the aforementioned JAK2 and KLF1 (Jeong et al., 2019). The plethora of PTMs likely contribute to the overlapping but distinct functions of the three TETs, as well as the distinct responses to different stimulation.

Targeting epigenetic modulators as a treatment strategy for β-globinopathies

Epigenetic modulators are broadly divided into three categories: epigenetic writers, readers, and erasers (Biswas and Rao, 2018; Villaseñor and Baubec, 2020). Epigenetic writers are enzymes that add new modifications, such as histone methyl transferases, DNMTs, and TETs. Epigenetic readers are proteins that recognize existing modification and recruit additional factors to activate or suppress transcription. These include the BET family proteins and the methyl-CpG binding domain proteins (MBDs). Epigenetic erasers are enzymes that remove existing modifications, such as histone deacetylases (HDAC). In recent years, as we gain a greater understanding of the role of epigenetic modifications in health and disease, an increasing number of pharmaceutical inhibitors of epigenetic modulators have been tested in clinical trials for treatment of various diverse diseases.

In Chapter IV, I presented my results that suggest BET inhibitors may be effective in treating β-globinopathies by reactivating the embryonic globin gene *HBE1* while maintaining expression of fetal globin genes and suppressing adult globin genes. The effect was mediated through suppressing genes that inhibit fetal and embryonic *β-globin* expression and disrupting the chromatin structure at the *β-globin* locus. BET inhibitors are already widely being tested in various clinical trials prior to my project. These clinical trials included solid tumors and hematological malignancies as well as non-malignant conditions like inflammatory diseases and cardiovascular diseases (Cochran et al., 2019). Another class of inhibitors receiving great clinical attention is the HDAC inhibitors. To date, over 30 different HDAC inhibitors have been or are being tested in clinical trials, and 5 have been approved in the US for treatment of various

malignant and non-malignant conditions (Ho et al., 2020). Of note, a few studies have reported that HDAC inhibitors can elevate fetal hemoglobin levels in human erythroid cells, making them ideal candidates for treating SCA (Hebbel et al., 2010; Johnson et al., 2005). More recently, various dual-inhibitors based on existing BET or HDAC inhibitors have been developed, including BET/HDAC dual inhibitors that target both classes of proteins (Liu et al., 2020b; Zhang et al., 2020). Based on my observations presented in Chapter IV and reports showing fetal *β-globin* upregulation by HDAC inhibition, it is possible that the dual inhibitors have a stronger effect reactivating fetal/embryonic *β-globin* genes compared to BET- or HDAC-specific inhibitors alone, although the actual effects have yet to be tested/elucidated. A third class of drugs that are currently used to treat SCA are the hypomethylating agents including decitabine and 5-azacitidine (Saunthararajah, 2019). Decitabine is highly effective in reactivating the fetal *β-globin* genes, inducing >10% increase of HbF at 0.2mg/kg twice a week in patients who had no respond to high doses (20mg/kg/day) of HU (Saunthararajah, 2019).

A common theme among the drugs discussed so far is that they all target different layers of the epigenetic regulatory mechanisms: the hypomethylating agents target epigenetic writers (DNMTs), BET inhibitors target epigenetic readers (BRD2/3/4/T), and HDAC inhibitors target epigenetic erasers (HDACs). The fact that the inhibition of different components of the epigenetic machinery produce similar effects at the *β-globin* locus highlights the importance of epigenetic control of the *β-globin* switching process in development (Sankaran and Orkin, 2013). The binding of TFs that promote adult and inhibit fetal/embryonic *β-globin* expression, as well as the looping of

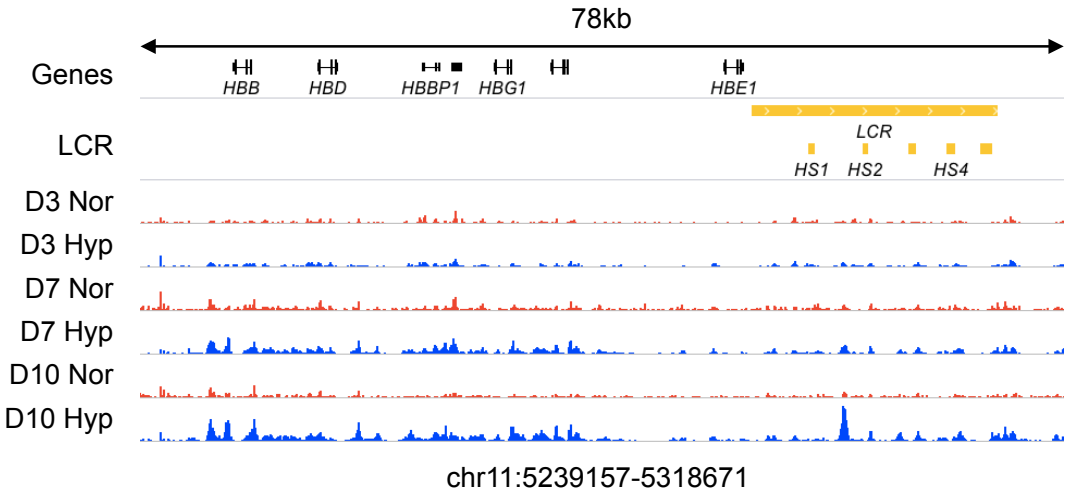
LCR are all dependent on proper epigenetic modifications across the locus. The development of these drugs in treating SCA or other β-globinopathies demonstrate that in-depth understanding of epigenetic regulation of gene expression can lead to great advances in clinical strategies to treat multiple diseases.

Possible interplay between hypoxia and the BRD proteins

As discussed earlier, hematopoiesis occurs in a highly hypoxic environment, yet the experiments presented in Chapter VI regarding BET inhibitors and their effects were conducted entirely in 21% oxygen. This means that changes observed at the molecular level in cell lines may be very different from what occurs *in vivo*. Interestingly, through examination of expression data from CD34+ cells, I found that *BRD2* is downregulated in hypoxia. Considering that K562 cells spontaneously produce hemoglobin under both hypoxic conditions (Figure 3.12B) and JQ1 treatment (Figure 4.1C), it is possible that *BRD2* downregulation or BRD2 inhibition is at least partially responsible for driving erythropoiesis in both processes. Furthermore, it is unclear how hypoxia-induced 5-hmC changes affect TF binding at the *β-globin* genes. However, I found a unique 5-hmC peak between HS1 and HS2 of the LCR in hypoxia day 10 that is not present in earlier hypoxic time points or any of the time points of normoxic differentiation (Figure 5.3). This observation suggests that hypoxia indeed has an effect on the chromatin structure at the *β-globin* locus, but the mechanism has yet to be investigated. In addition, the late appearance of the 5-hmC peak is consistent with the role of TET3 in later stages of erythropoiesis (Yan et al., 2017), suggesting that this is likely an TET3-mediated event. The 5-hmC changes in hypoxia are likely accompanied by changes in histone

modifications that affect BRD and other TF binding, which have not yet been characterized.

Figure 5.3 5-hmC distribution at the *β-globin* locus in normoxic versus hypoxic erythropoiesis.



5-hmC distribution across the *β-globin* locus in normoxic versus hypoxic erythroid differentiation in CD34+ HSPCs. The yellow locus control region (LCR) is given above the locations of five DNase I hypersensitivity sites (HS1-5). Normoxic 5-hmC tracks are colored red, and hypoxic 5-hmC tracks are colored blue. 5-hmC tracks are scaled to normalize the number of total aligned reads.

Future Directions

My results have raised additional questions regarding the role and regulation of the epigenetic landscape in erythropoiesis, some of which were briefly addressed in previous sections. These questions can be summarized as: (i) What other erythropoietic genes are targets of HIF-1, and how do 5-hmC changes in hypoxia affect their transcription? (ii) Do TET PTMs in hypoxia influence their catalytic activity? And (iii) Are BET inhibitors effective in primary human erythroid progenitor cells *in vitro* and *in vivo*?

i. What other erythropoietic genes are targets of HIF-1, and how do 5-hmC changes in hypoxia affect their transcription?

Hypoxia likely induces many different genes in HSPCs that drives erythropoiesis. Although I showed that *TET3* is directly induced by HIF-1 (Chapter III) (Cao et al., 2020), there remain other potential direct HIF-1 target genes. With RNA-seq expression data generated by my project, genes that are up- or downregulated in hypoxia can be identified. Potential HIF-1 bindings sites can be identified from a combination of HIF-1α/β ChIP-seq data, HOMER TF binding site predictions, and enhancer annotations. In particular, it is interesting to see if any other erythroid genes are directly upregulated by HIF-1. Once identified, the binding sites can be deleted using CRISPR-Cas9 to test if HIF-1 binding is necessary for the regulation of these genes.

It is also interesting to characterize the covalent cytosine modification status of the HIF-1 binding sites that regulate the hypoxia response in erythropoiesis. The consensus HIF-1 binding site contains a central CpG dinucleotide, and HIF-1 cannot bind to the DNA if the cytosine in this dinucleotide on either strand is modified in any way (Mariani et al., 2014). Therefore, the 5-hmC status of any given HIF-1 binding site could dictate the HIF-1-dependent expression changes around the binding site. This would further elucidate the interplay between hypoxia and 5-hmC regulation. This can be done by TET-assisted bisulfite sequencing, which allows the quantification of 5-hmC/5-mC at single-base resolution within the sequenced region, as was performed in the previous study from our laboratory (Mariani et al., 2014).

Lastly, it would be interesting to investigate whether any of the HIF-1 binding sites that regulate erythropoietic gene targets contain polymorphisms across different

populations, particularly within ones that are adapted to high-altitude hypoxic environment like the Tibetans (Jeong et al., 2014; Tashi et al., 2017). Potential polymorphisms at these locations can modify the behavior of HIF-1 transcriptional activity in erythropoiesis, which may contribute to the evolutionary adaptation to chronic hypoxic environments.

ii. Do TET PTMs in hypoxia influence their catalytic activity?

Although my results have shown a regulatory mechanism at the transcriptional level for *TET3*, it is equally likely that both TET2 and TET3 undergo changes in their PTMs in response to hypoxia, as discussed earlier in this chapter. The increased 5-hmC observed in hypoxia is likely the combined result of increased *TET3* transcription and increased TET2/3 stability/activity. Potential PTMs can be identified by predictions based on peptide motifs. To validate the predictions, Western blots could be used for larger PTMs like ubiquitination, while smaller PTMs like phosphorylation and acetylation are best identified through mass spectrometry. Once a PTM is validated, site-specific mutagenesis can be used to specifically remove the amino acid residue with the PTM, and the resulting mutated enzyme can be compared to the WT enzymes for functional characterization. With the abundance of predicted PTMs of unknown significance (Hornbeck et al., 2019), an in-depth study on TET PTMs in erythropoiesis could provide further insights into the regulation of TET protein and subsequently the epigenome.

(iii) Are BET inhibitors effective in primary human erythroid progenitor cells in vitro *and* in vivo*?*

My results presented in Chapter IV were done entirely in immortalized erythroid leukemia cell lines. These results serve well as a proof of concept for the potential of BET inhibitors to treat SCA or other β-globinopathies, but fall short of modeling normal erythroid cells. In order to test whether BET inhibitors can reactivate fetal/embryonic *α/β-globin* genes, additional cell lines that better represent normal erythroid progenitor cells could be employed. These include the human umbilical cord blood-derived erythroid progenitor (HUDEP-2) cells and BEL-A cells (Trakarnsanga et al., 2017; Vinjamur and Bauer, 2018), as well as normal primary CD34+ cells undergoing erythroid differentiation *in vitro*. In addition, CD34+ cells from sickle cell patients could provide another model in which to test the effectiveness of BET inhibitors in reactivating fetal and embryonic *α/β-globins*. My results focused primarily on the re-expression of fetal and embryonic *β-globin* genes, whereas future studies should additionally quantify the composition of hemoglobins through hemoglobin electrophoresis.

Finally, as mentioned previously, erythroid differentiation occurs primarily under hypoxia. Additionally, sickle hemoglobin polymerization and sickling also occurs under low oxygen conditions. Thus, it is essential to test the effects of BET inhibitors under hypoxia to best model the *in vivo* environment the erythroid progenitor cells experience.

In conclusion, my results have revealed a novel mechanism regulating *TET3* expression in hypoxia, as well as the potential of BET inhibitors in treating β-globinopathies. Both of my projects highlighted the importance of epigenetic regulation in erythropoiesis. I contributed to multiple other studies by mass spectrometry and genome-wide profiling of 5-hmC, demonstrating the significance of these techniques as tools to study epigenetic changes. Future studies should continue to investigate the

effect of hypoxia on the epigenome in erythropoiesis and in other cellular processes as well as further scrutinize the clinical potential of BET inhibitors in reactivating fetal/embryonic *α/β-globins* and the mechanisms driving these changes.

REFERENCES

Abdel-Wahab, O., Mullally, A., Hedvat, C., Garcia-Manero, G., Patel, J., Wadleigh, M., Malinge, S., Yao, J., Kilpivaara, O., Bhat, R.*, et al.* (2009). Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. Blood *114*, 144-147.

Aerbajinai, W., Zhu, J., Kumkhaek, C., Chin, K., and Rodgers, G.P. (2009). SCF induces gamma-globin gene expression by regulating downstream transcription factor COUP-TFII. Blood *114*, 187-194.

Agathocleous, M., Meacham, C.E., Burgess, R.J., Piskounova, E., Zhao, Z., Crane, G.M., Cowin, B.L., Bruner, E., Murphy, M.M., Chen, W.*, et al.* (2017). Ascorbate regulates haematopoietic stem cell function and leukaemogenesis. Nature *549*, 476-481.

Albrecht, B.K., Gehling, V.S., Hewitt, M.C., Vaswani, R.G., Côté, A., Leblanc, Y., Nasveschuk, C.G., Bellon, S., Bergeron, L., Campbell, R.*, et al.* (2016). Identification of a Benzoisoxazoloazepine Inhibitor (CPI-0610) of the Bromodomain and Extra-Terminal (BET) Family as a Candidate for Human Clinical Trials. J Med Chem *59*, 1330-1339.

Ali, M.A., Ahmad, A., Chaudry, H., Aiman, W., Aamir, S., Anwar, M.Y., and Khan, A. (2020). Efficacy and safety of recently approved drugs for sickle cell disease: a review of clinical trials. Exp Hematol *92*, 11-18.e11.

Alqahtani, A., Choucair, K., Ashraf, M., Hammouda, D.M., Alloghbi, A., Khan, T., Senzer, N., and Nemunaitis, J. (2019). Bromodomain and extra-terminal motif inhibitors: a review of preclinical and clinical advances in cancer therapy. Future Sci OA *5*, FSO372.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Andrews, N.C. (2008). Forging a field: the golden age of iron biology. Blood *112*, 219-230.

Ang, S.O., Chen, H., Hirota, K., Gordeuk, V.R., Jelinek, J., Guan, Y., Liu, E., Sergueeva, A.I., Miasnikova, G.Y., Mole, D.*, et al.* (2002). Disruption of oxygen homeostasis underlies congenital Chuvash polycythemia. Nat Genet *32*, 614-621.

Atlasi, Y., and Stunnenberg, H.G. (2017). The interplay of epigenetic marks during stem cell differentiation and development. Nat Rev Genet *18*, 643-658.

Avram, D., Fields, A., Pretty On Top, K., Nevrivy, D.J., Ishmael, J.E., and Leid, M. (2000). Isolation of a novel family of C(2)H(2) zinc finger proteins implicated in transcriptional repression mediated by chicken ovalbumin upstream promoter transcription factor (COUP-TF) orphan nuclear receptors. J Biol Chem *275*, 10315-10322.

Bacher, U., Weissmann, S., Kohlmann, A., Schindela, S., Alpermann, T., Schnittger, S., Kern, W., Haferlach, T., and Haferlach, C. (2012). TET2 deletions are a recurrent but rare phenomenon in myeloid malignancies and are frequently accompanied by TET2 mutations on the remaining allele. British journal of haematology *156*, 67-75.

Bauer, C., Göbel, K., Nagaraj, N., Colantuoni, C., Wang, M., Müller, U., Kremmer, E., Rottach, A., and Leonhardt, H. (2015). Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). J Biol Chem *290*, 4801-4812.

Baylin, S.B., Höppener, J.W., de Bustros, A., Steenbergh, P.H., Lips, C.J., and Nelkin, B.D. (1986). DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. Cancer research *46*, 2917-2922.

Bender, M.A., Bulger, M., Close, J., and Groudine, M. (2000). Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. Mol Cell *5*, 387-393.

Bisht, K., Brunck, M.E., Matsumoto, T., McGirr, C., Nowlan, B., Fleming, W., Keech, T., Magor, G., Perkins, A.C., Davies, J.*, et al.* (2019). HIF prolyl hydroxylase inhibitor FG-4497 enhances mouse hematopoietic stem cell mobilization via VEGFR2/KDR. Blood Adv *3*, 406-418.

Biswas, S., and Rao, C.M. (2018). Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. Eur J Pharmacol *837*, 8-24.

Blaschke, K., Ebata, K.T., Karimi, M.M., Zepeda-Martínez, J.A., Goyal, P., Mahapatra, S., Tam, A., Laird, D.J., Hirst, M., Rao, A.*, et al.* (2013). Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature.

Bond, J., Gale, D.P., Connor, T., Adams, S., de Boer, J., Gascoyne, D.M., Williams, O., Maxwell, P.H., and Ancliff, P.J. (2011). Dysregulation of the HIF pathway due to VHL mutation causing severe erythrocytosis and pulmonary arterial hypertension. Blood *117*, 3699-3701.

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.*, et al.* (2010). Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. Nat Genet *42*, 801-805.

Bowman, R.L., and Levine, R.L. (2017). TET2 in Normal and Malignant Hematopoiesis. Cold Spring Harb Perspect Med *7*, pii: a026518.

Cao, J.Z., Hains, A.E., and Godley, L.A. (2019). Regulation of 5-Hydroxymethylcytosine Distribution by the TET Enzymes. In The DNA, RNA, and Histone Methylomes, S.B. Jurga, Jan, ed. (Springer), pp. 229-263.

Cao, J.Z., Liu, H., Wickrema, A., and Godley, L.A. (2020). HIF-1 directly induces TET3 expression to enhance 5-hmC density and induce erythroid gene expression in hypoxia. Blood Adv *4*, 3053-3062.

Case, A.J. (2017). On the Origin of Superoxide Dismutase: An Evolutionary Perspective of Superoxide-Mediated Redox Signaling. Antioxidants (Basel, Switzerland) *6*.

Case, A.J., McGill, J.L., Tygrett, L.T., Shirasawa, T., Spitz, D.R., Waldschmidt, T.J., Legge, K.L., and Domann, F.E. (2011). Elevated mitochondrial superoxide disrupts normal T cell development, impairing adaptive immune responses to an influenza challenge. Free Radic Biol Med *50*, 448-458.

Chaidos, A., Caputo, V., and Karadimitris, A. (2015). Inhibition of bromodomain and extra-terminal proteins (BET) as a potential therapeutic approach in haematological malignancies: emerging preclinical and clinical evidence. Ther Adv Hematol *6*, 128-141.

Chan, C.M., Fulton, J., Montiel-Duarte, C., Collins, H.M., Bharti, N., Wadelin, F.R., Moran, P.M., Mongan, N.P., and Heery, D.M. (2013). A signature motif mediating selective interactions of BCL11A with the NR2E/F subfamily of orphan nuclear receptors. Nucleic Acids Res *41*, 9663-9679.

Chapman, C.G., Mariani, C.J., Wu, F., Meckel, K., Butun, F., Chuang, A., Madzo, J., Bissonnette, M.B., Kwon, J.H., and Godley, L.A. (2015). TET-catalyzed 5-hydroxymethylcytosine regulates gene expression in differentiating colonocytes and colon cancer. Sci Rep *5*, 17568.

Chen, C.-C., Wang, K.-Y., and Shen, C.-K.J. (2012). The mammalian de novo DNA methyltransferases DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases. J Biol Chem *287*, 33116-33121.

Chen, D., Maruschke, M., Hakenberg, O., Zimmermann, W., Stief, C.G., and Buchner, A. (2017). TOP2A, HELLS, ATAD2, and TET3 Are Novel Prognostic Markers in Renal Cell Carcinoma. Urology *102*, 2650.

Chen, J., Guo, L., Zhang, L., Wu, H., Yang, J., Liu, H., Wang, X., Hu, X., Gu, T., Zhou, Z.*, et al.* (2013). Vitamin C modulates TET1 function during somatic cell reprogramming. Nat Genet.

Cheng, J., Guo, S., Chen, S., Mastriano, S.J., Liu, C., D'Alessio, A.C., Hysolli, E., Guo, Y., Yao, H., Megyola, C.M.*, et al.* (2013). An extensive network of TET2-targeting MicroRNAs regulates malignant hematopoiesis. Cell Rep *5*, 471-481.

Chin, K., Yu, X., Beleslin-Cokic, B., Liu, C., Shen, K., Mohrenweiser, H.W., and Noguchi, C.T. (2000). Production and processing of erythropoietin receptor transcripts in brain. Mol Brain Res *81*, 29-42.

Chung, T.L., Brena, R.M., Kolle, G., Grimmond, S.M., Berman, B.P., Laird, P.W., Pera, M.F., and Wolvetang, E.J. (2010). Vitamin C promotes widespread yet specific DNA demethylation of the epigenome in human embryonic stem cells. Stem Cells.

Cimmino, L., Dolgalev, I., Wang, Y., Yoshimi, A., Martin, G.H.H., Wang, J., Ng, V., Xia, B., Witkowski, M.T., Mitchell-Flack, M.*, et al.* (2017). Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression. Cell *170*, 1079.

Cochran, A.G., Conery, A.R., and Sims, R.J. (2019). Bromodomains: a new target class for drug development. Nat Rev Drug Discov *18*, 609-628.

Coey, C.T., Malik, S.S., Pidugu, L.S., Varney, K.M., Pozharski, E., and Drohat, A.C. (2016). Structural basis of damage recognition by thymine DNA glycosylase: Key roles for N-terminal residues. Nucleic Acids Res *44*, 10248-10258.

Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittinger, M.A., Driggers, E.M., Fantin, V.R., Jang, H.G., Jin, S., Keenan, M.C.*, et al.* (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature *462*, 739-744.

Dawlaty, M.M., Breiling, A., Le, T., Barrasa, M.I., Raddatz, G., Gao, Q., Powell, B.E., Cheng, A.W., Faull, K.F., Lyko, F.*, et al.* (2014). Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. Dev Cell *29*, 102-111.

Delhommeau, F., Dupont, S., Della Valle, V., James, C., Trannoy, S., Massé, A., Kosmider, O., Le Couedic, J.-P., Robert, F., Alberdi, A.*, et al.* (2009). Mutation in TET2 in myeloid cancers. N Engl J Med *360*, 2289-2301.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell *149*, 1233-1244.

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

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. Nature *399*, 491-496.

Doroshow, D.B., Eder, J.P., and LoRusso, P.M. (2017). BET inhibitors: a novel epigenetic approach. Ann Oncol *28*, 1776-1787.

Euba, B., Vizmanos, J.L.L., García-Granero, M., Aranaz, P., Hurtado, C., Migueliz, I., Novo, F.J., and García-Delgado, M. (2012). A meta-analysis of TET2 mutations shows a distinct distribution pattern in de novo acute myeloid leukemia and chronic myelomonocytic leukemia. Leuk Lymphoma *53*, 1230-1233.

Feinberg, A.P., and Tycko, B. (2004). The history of cancer epigenetics. Nat Rev Cancer *4*, 143-153.

Feinberg, A.P., and Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature *301*, 89-92.

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature *473*, 398-402.

Field, S.F., Beraldi, D., Bachman, M., Stewart, S.K., Beck, S., and Balasubramanian, S. (2015). Accurate measurement of 5-methylcytosine and 5-hydroxymethylcytosine in human cerebellum DNA by oxidative bisulfite on an array (OxBS-array). PloS One *10*.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F.*, et al.* (2010). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer cell *18*, 553-567.

Filippakopoulos, P., and Knapp, S. (2014). Targeting bromodomains: epigenetic readers of lysine acetylation. Nat Rev Drug Discov *13*, 337-356.

Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W.B., Fedorov, O., Morse, E.M., Keates, T., Hickman, T.T., Felletar, I.*, et al.* (2010). Selective inhibition of BET bromodomains. Nature *468*, 1067-1073.

Forrester, W.C., Thompson, C., Elder, J.T., and Groudine, M. (1986). A developmentally stable chromatin structure in the human beta-globin gene cluster. Proc Natl Acad Sci U S A *83*, 1359-1363.

Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J.*, et al.* (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res *47*, D766-D773.

Gama-Sosa, M.A., Slagel, V.A., Trewyn, R.W., Oxenhandler, R., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983). The 5-methylcytosine content of DNA from human tumors. Nucleic Acids Res *11*, 6883-6894.

Gamsjaeger, R., Webb, S.R., Lamonica, J.M., Billin, A., Blobel, G.A., and Mackay, J.P. (2011). Structural basis and specificity of acetylated transcription factor GATA1 recognition by BET family bromodomain protein Brd3. Mol Cell Biol *31*, 2632-2640.

Geng, T., Zhang, D., and Jiang, W. (2019). Epigenetic Regulation of Transition Among Different Pluripotent States: Concise Review. Stem Cells *37*, 1372-1380.

Gorres, K.L., and Raines, R.T. (2010). Prolyl 4-hydroxylase. Crit Rev Biochem Mol Biol *45*, 106-124.

Goupille, O., Penglong, T., Lefèvre, C., Granger, M., Kadri, Z., Fucharoen, S., Maouche-Chrétien, L., Leboulch, P., and Chrétien, S. (2012). BET bromodomain inhibition rescues erythropoietin differentiation of human erythroleukemia cell line UT7. Biochem Biophys Res Commun *429*, 1-5.

Guo, J.U., Su, Y., Zhong, C., Ming, G.-l., and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423-434.

Hagège, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forné, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat Protoc *2*, 1722-1733.

Hashimoto, H., Pais, J.E., Dai, N., Corrêa, I.R., Zhang, X., Zheng, Y., and Cheng, X. (2015). Structure of Naegleria Tet-like dioxygenase (NgTet1) in complexes with a reaction intermediate 5-hydroxymethylcytosine DNA. Nucleic Acids Res *43*, 10713-10721.

Hashimoto, H., Pais, J.E., Zhang, X., Saleh, L., Fu, Z.-Q., Dai, N., Corrêa, I.R., Zheng, Y., and Cheng, X. (2014). Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. Nature *506*, 391-395.

Hassan, H.M., Kolendowski, B., Isovic, M., Bose, K., Dranse, H.J., Sampaio, A.V., Underhill, T.M., and Torchia, J. (2017). Regulation of Active DNA Demethylation through RAR-Mediated Recruitment of a TET/TDG Complex. Cell Rep *19*, 1685-1697.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., *et al.* (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science *333*, 1303-1307.

He, Z., and Russell, J.E. (2001). Expression, purification, and characterization of human hemoglobins Gower-1 (zeta(2)epsilon(2)), Gower-2 (alpha(2)epsilon(2)), and Portland-2 (zeta(2)beta(2)) assembled in complex transgenic-knockout mice. Blood *97*, 1099-1105.

He, Z., and Russell, J.E. (2002). A human embryonic hemoglobin inhibits Hb S polymerization in vitro and restores a normal phenotype to mouse models of sickle cell disease. Proc Natl Acad Sci U S A *99*, 10635-10640.

Hebbel, R.P., Vercellotti, G.M., Pace, B.S., Solovey, A.N., Kollander, R., Abanonu, C.F., Nguyen, J., Vineyard, J.V., Belcher, J.D., Abdulla, F.*, et al.* (2010). The HDAC inhibitors trichostatin A and suberoylanilide hydroxamic acid exhibit multiple modalities of benefit for the vascular pathobiology of sickle transgenic mice. Blood *115*, 2483-2490.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Ho, T.C.S., Chan, A.H.Y., and Ganesan, A. (2020). Thirty Years of HDAC Inhibitors: 2020 Insight and Hindsight. J Med Chem *63*, 12460-12484.

Hoban, M.D., Orkin, S.H., and Bauer, D.E. (2016). Genetic treatment of a molecular disorder: gene therapy approaches to sickle cell disease. Blood *127*, 839-848.

Hon, G.C., Song, C.-X., Du, T., Jin, F., Selvaraj, S., Lee, A.Y., Yen, C.-A., Ye, Z., Mao, S.-Q., Wang, B.-A.*, et al.* (2014). 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. Mol Cell *56*, 286-297.

Hornbeck, P.V., Kornhauser, J.M., Latham, V., Murray, B., Nandhikonda, V., Nord, A., Skrzypek, E., Wheeler, T., Zhang, B., and Gnad, F. (2019). 15 years of PhosphoSitePlus®: integrating post-translationally modified sites, disease variants and isoforms. Nucleic Acids Res *47*.

Hou, H.-A., and Tien, H.-F. (2020). Genomic landscape in acute myeloid leukemia and its implications in risk classification and targeted therapies. J Biomed Sci *27*, 81.

Houwing, M.E., de Pagter, P.J., van Beers, E.J., Biemond, B.J., Rettenbacher, E., Rijneveld, A.W., Schols, E.M., Philipsen, J.N.J., Tamminga, R.Y.J., van Draat, K.F.*, et al.* (2019). Sickle cell disease: Clinical presentation and management of a global health challenge. Blood Rev *37*, 100580.

Hrit, J., Goodrich, L., Li, C., Wang, B.-A., Nie, J., Cui, X., Martin, E.A., Simental, E., Fernandez, J., Liu, M.Y.*, et al.* (2018). OGT binds a conserved C-terminal domain of TET1 to regulate TET1 activity and function in development. eLife *7*.

Hsieh, A.L., Walton, Z.E., Altman, B.J., Stine, Z.E., and Dang, C.V. (2015). MYC and metabolism on the path to cancer. Semin Cell Dev Biol *43*, 11-21.

Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y.G., Zhu, J., Wang, P., and Xu, Y. (2013). Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. Cell *155*, 1545-1555.

Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., Gong, W.*, et al.* (2015). Structural insight into substrate preference for TET-mediated oxidation. Nature *527*, 118-122.

Ikawa, Y., Miccio, A., Magrin, E., Kwiatkowski, J.L., Rivella, S., and Cavazzana, M. (2019). Gene therapy of hemoglobinopathies: progress and future challenges. Hum Mol Genet *28*, R24-R30.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300-1303.

Itzykson, R., Kosmider, O., Renneville, A., Gelsi-Boyer, V., Meggendorfer, M., Morabito, M., Berthon, C., Adès, L., Fenaux, P., Beyne-Rauzy, O.*, et al.* (2013). Prognostic score including gene mutations in chronic myelomonocytic leukemia. J Clin Oncol *31*, 2428-2436.

Ivaldi, M.S., Diaz, L.F., Chakalova, L., Lee, J., Krivega, I., and Dean, A. (2018). Fetal γ-globin genes are regulated by the BGLT3 long noncoding RNA locus. Blood *132*, 1963-1973.

Ivan, M., and Kaelin, W.G. (2017). The EGLN-HIF O2-Sensing System: Multiple Inputs and Feedbacks. Mol Cell *66*, 772-779.

Iyer, L.M., Tahiliani, M., Rao, A., and Aravind, L. (2009). Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle *8*, 1698-1710.

Jackson, C.S., Durandt, C., Janse van Rensburg, I., Praloran, V., Brunet de la Grange, P., and Pepper, M.S. (2017). Targeting the aryl hydrocarbon receptor nuclear translocator complex with DMOG and Stemregenin 1 improves primitive hematopoietic stem cell expansion. Stem Cell Res *21*, 124-131.

Jankowska, A.M., Szpurka, H., Tiu, R.V., and Blood, M.-H. (2009). Loss of heterozygosity 4q24 and TET2 mutations associated with myelodysplastic/myeloproliferative neoplasms. Blood.

Jäwert, F., Hasséus, B., Kjeller, G., Magnusson, B., Sand, L., and Larsson, L. (2013). Loss of 5-hydroxymethylcytosine and TET2 in oral squamous cell carcinoma. Anticancer research *33*, 4325-4328.

Jeong, C., Alkorta-Aranburu, G., Basnyat, B., Neupane, M., Witonsky, D.B., Pritchard, J.K., Beall, C.M., and Di Rienzo, A. (2014). Admixture facilitates genetic adaptations to high altitude in Tibet. Nat Commun *5*, 3281.

Jeong, C., Witonsky, D.B., Basnyat, B., Neupane, M., Beall, C.M., Childs, G., Craig, S.R., Novembre, J., and Di Rienzo, A. (2018). Detecting past and ongoing natural selection among ethnically Tibetan women at high altitude in Nepal. PLoS Genet *14*, e1007650.

Jeong, J.J., Gu, X., Nie, J., Sundaravel, S., Liu, H., Kuo, W.-L., Bhagat, T.D., Pradhan, K., Cao, J., Nischal, S.*, et al.* (2019). Cytokine-Regulated Phosphorylation and Activation of TET2 by JAK2 in Hematopoiesis. Cancer Discov *9*, 778-795.

Jiang, D., Zhang, Y., Hart, R.P., Chen, J., Herrup, K., and Li, J. (2015). Alteration in 5-hydroxymethylcytosine-mediated epigenetic regulation leads to Purkinje cell vulnerability in ATM deficiency. Brain *138*, 3520-3536.

Jiang, S., Yan, W., Wang, S.E., and Baltimore, D. (2019). Dual mechanisms of posttranscriptional regulation of Tet2 by Let-7 microRNA in macrophages. Proc Natl Acad Sci U S A *116*, 12416-12421.

Jjingo, D., Conley, A.B., Yi, S.V., Lunyak, V.V., and Jordan, I.K. (2012). On the presence and role of human gene-body DNA methylation. Oncotarget *3*, 462-474.

Johnson, J., Hunter, R., McElveen, R., Qian, X.H., Baliga, B.S., and Pace, B.S. (2005). Fetal hemoglobin induction by the histone deacetylase inhibitor, scriptaid. Cell Mol Biol (Noisy-le-Grand) *51*, 229-238.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet *13*, 484-492.

Joshi, S.R., Kitagawa, A., Jacob, C., Hashimoto, R., Dhagia, V., Ramesh, A., Zheng, C., Zhang, H., Jordan, A., Waddell, I.*, et al.* (2020). Hypoxic activation of glucose-6-phosphate dehydrogenase controls the expression of genes involved in the pathogenesis of pulmonary hypertension through the regulation of DNA methylation. American journal of physiology Lung cellular and molecular physiology *318*.

Kafer, G.R., Li, X., Horii, T., Suetake, I., Tajima, S., Hatada, I., and Carlton, P.M. (2016). 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability. Cell Rep *14*, 1283-1292.

Kang, J., Lienhard, M., Pastor, W.A., Chawla, A., Novotny, M., Tsagaratou, A., Lasken, R.S., Thompson, E.C., Surani, M.A., Koralov, S.B.*, et al.* (2015a). Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. Proc Natl Acad Sci U S A *112*, 45.

Kang, J.-A., Zhou, Y., Weis, T.L., Liu, H., Ulaszek, J., Satgurunathan, N., Zhou, L., van Besien, K., Crispino, J., Verma, A.*, et al.* (2008). Osteopontin regulates actin cytoskeleton and contributes to cell proliferation in primary erythroblasts. J Biol Chem *283*, 6997-7006.

Kang, Y., Kim, Y.W., Kang, J., Yun, W.J., and Kim, A. (2017). Erythroid specific activator GATA-1-dependent interactions between CTCF sites around the β-globin locus. Biochim Biophys Acta Gene Regul Mech *1860*, 416-426.

Kang, Y., Kim, Y.W., Yun, J., Shin, J., and Kim, A. (2015b). KLF1 stabilizes GATA-1 and TAL1 occupancy in the human β-globin locus. Biochim Biophys Acta *1849*, 282-289.

Keeley, T.P., and Mann, G.E. (2019). Defining Physiological Normoxia for Improved Translation of Cell Physiology to Animal Models and Humans. Physiol Rev *99*, 161-234.

Keys, J.R., Tallack, M.R., Zhan, Y., Papathanasiou, P., Goodnow, C.C., Gaensler, K.M., Crossley, M., Dekker, J., and Perkins, A.C. (2008). A mechanism for Ikaros regulation of human globin gene switching. Br J Haematol *141*, 398-406.

Kiefer, C.M., Lee, J., Hou, C., Dale, R.K., Lee, Y.T., Meier, E.R., Miller, J.L., and Dean, A. (2011). Distinct Ldb1/NLI complexes orchestrate γ-globin repression and reactivation through ETO2 in human adult erythroid cells. Blood *118*, 6200-6208.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol *37*, 907-915.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Ko, M., An, J., Bandukwala, H.S., Chavez, L., Aijö, T., Pastor, W.A., Segal, M.F., Li, H., Koh, K.P., Lähdesmäki, H.*, et al.* (2013). Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. Nature *497*, 122-126.

Ko, M., Huang, Y., Jankowska, A.M., Pape, U.J., Tahiliani, M., Bandukwala, H.S., An, J., Lamperti, E.D., Koh, K.P., Ganetzky, R.*, et al.* (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature *468*, 839-843.

Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C.A., Mostoslavsky, G.*, et al.* (2011). Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell *8*, 200-213.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929-930.

Kuiper, C., and Vissers, M.C. (2014). Ascorbate as a co-factor for fe- and 2-oxoglutarate dependent dioxygenases: physiological activity in tumor growth and progression. Front Oncol *4*, 359.

Kumar, D., Aggarwal, M., Kaas, G.A., Lewis, J., Wang, J., Ross, D.L., Zhong, C., Kennedy, A., Song, H., and Sweatt, J.D. (2015). Tet1 Oxidase Regulates Neuronal Gene Transcription, Active DNA Hydroxy-methylation, Object Location Memory, and Threat Recognition Memory. Neuroepigenetics *4*, 12-27.

Lamonica, J.M., Deng, W., Kadauke, S., Campbell, A.E., Gamsjaeger, R., Wang, H., Cheng, Y., Billin, A.N., Hardison, R.C., Mackay, J.P.*, et al.* (2011). Bromodomain

protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. Proc Natl Acad Sci U S A *108*, 68.

Langemeijer, S.M., Kuiper, R.P., Berends, M., Knops, R., Aslanyan, M.G., Massop, M., Stevens-Linders, E., van Hoogen, P., van Kessel, A.G., Raymakers, R.A.*, et al.* (2009). Acquired mutations in TET2 are common in myelodysplastic syndromes. Nat Genet *41*, 838-842.

Lasho, T.L., Vallapureddy, R., Finke, C.M., Mangaonkar, A., Gangat, N., Ketterling, R., Tefferi, A., and Patnaik, M.M. (2018). Infrequent occurrence of TET1, TET3, and ASXL2 mutations in myelodysplastic/myeloproliferative neoplasms. Blood Cancer J *8*, 32.

Laukka, T., Mariani, C.J., Ihantola, T., Cao, J.Z., Hokkanen, J., Kaelin, W.G., Godley, L.A., and Koivunen, P. (2015). Fumarate and Succinate Regulate Expression of Hypoxia-Inducible Genes via TET Enzymes. J Biol Chem.

Lee, H.J., Hore, T.A., and Reik, W. (2014). Reprogramming the methylome: erasing memory and creating diversity. Cell Stem Cell *14*, 710-719.

Lee, S.-G., Cho, S.Y., Kim, M.J., Oh, S.H., Cho, E.H., Lee, S., Baek, E.J., Choi, J.H., Bohlander, S.K., Lode, L.*, et al.* (2013). Genomic breakpoints and clinical features of MLL-TET1 rearrangement in acute leukemias. Haematologica *98*, 7.

Ley, T.J., Maloney, K.A., Gordon, J.I., and Schwartz, A.L. (1989). Globin gene expression in erythroid human fetal liver cells. J Clin Invest *83*, 1032-1038.

Li, C., Lan, Y., Schwartz-Orbach, L., Korol, E., Tahiliani, M., Evans, T., and Goll, M.G. (2015). Overlapping Requirements for Tet2 and Tet3 in Normal Development and Hematopoietic Stem Cell Emergence. Cell Rep *12*, 1133-1143.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, Y., Shen, Z., Jiang, H., Lai, Z., Wang, Z., Jiang, K., Ye, Y., and Wang, S. (2018). MicroRNA□4284 promotes gastric cancer tumorigenicity by targeting ten-eleven translocation 1. Mol Med Rep *17*, 6569-6575.

Lian, C.G., Xu, Y., Ceol, C., Wu, F., Larson, A., Dresser, K., Xu, W., Tan, L., Hu, Y., Zhan, Q.*, et al.* (2012). Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. Cell *150*, 1135-1146.

Lin, I.-H., Chen, Y.-F., and Hsu, M.-T. (2017). Correlated 5-Hydroxymethylcytosine (5hmC) and Gene Expression Profiles Underpin Gene and Organ-Specific Epigenetic Regulation in Adult Mouse Brain and Liver. PloS One *12*.

Liu, J., Hu, H., Panserat, S., and Marandel, L. (2020a). Evolutionary history of DNA methylation related genes in chordates: new insights from multiple whole genome duplications. Sci Rep *10*, 970.

Liu, T., Wan, Y., Xiao, Y., Xia, C., and Duan, G. (2020b). Dual-Target Inhibitors Based on HDACs: Novel Antitumor Agents for Cancer Therapy. J Med Chem *63*, 8977-9002.

Liutkevičiūtė, Z., Kriukienė, E., Ličytė, J., Rudytė, M., Urbanavičiūtė, G., and Klimašauskas, S. (2014). Direct decarboxylation of 5-carboxylcytosine by DNA C5-methyltransferases. J Am Chem Soc *136*, 5884-5887.

Lorenzo, F.R., Huff, C., Myllymäki, M., Olenchock, B., Swierczek, S., Tashi, T., Gordeuk, V., Wuren, T., Ri-Li, G., McClain, D.A.*, et al.* (2014). A genetic mechanism for Tibetan high-altitude adaptation. Nat Genet *46*, 951-956.

Lorsbach, R.B., Moore, J., Mathew, S., Raimondi, S.C., Mukatira, S.T., and Downing, J.R. (2003). TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia *17*, 637-641.

Losman, J.-A., and Kaelin, W.G. (2013). What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. Genes Dev *27*, 836-852.

Lu, J.L., Zhao, L., Han, S.C., Bi, J.L., Liu, H.X., Yue, C., and Lin, L. (2018). MiR-129 is involved in the occurrence of uterine fibroid through inhibiting TET1. Eur Rev Med Pharmacol Sci *22*, 4419-4426.

Ma, Q., Dasgupta, C., Shen, G., Li, Y., and Zhang, L. (2021). MicroRNA-210 downregulates TET2 and contributes to inflammatory response in neonatal hypoxic-ischemic brain injury. J Neuroinflammation *18*, 6.

Ma, Y., Qin, C., Li, L., Miao, R., Jing, C., and Cui, X. (2018). MicroRNA-21 promotes cell proliferation by targeting tumor suppressor TET1 in colorectal cancer. Int J Clin Exp Pathol *11*, 1439-1445.

Madzo, J., Liu, H., Rodriguez, A., Vasanthakumar, A., Sundaravel, S., Caces, D.B., Looney, T.J., Zhang, L., Lepore, J.B., Macrae, T.*, et al.* (2014). Hydroxymethylation at gene regulatory regions directs stem/early progenitor cell commitment during erythropoiesis. Cell Rep *6*, 231-244.

Maiti, A., and Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. J Biol Chem *286*, 35334-35338.

Manalo, D.J., Rowan, A., Lavoie, T., Natarajan, L., Kelly, B.D., Ye, S.Q., Garcia, J.G.N., and Semenza, G.L. (2005). Transcriptional regulation of vascular endothelial cell responses to hypoxia by HIF-1. Blood *105*, 659-669.

Mancias, J.D., Pontano Vaites, L., Nissim, S., Biancur, D.E., Kim, A.J., Wang, X., Liu, Y., Goessling, W., Kimmelman, A.C., and Harper, J.W. (2015). Ferritinophagy via NCOA4 is required for erythropoiesis and is regulated by iron dependent HERC2-mediated proteolysis. eLife *4*.

Mancias, J.D., Wang, X., Gygi, S.P., Harper, J.W., and Kimmelman, A.C. (2014). Quantitative proteomics identifies NCOA4 as the cargo receptor mediating ferritinophagy. Nature *509*, 105-109.

Manning, L.R., Russell, J.E., Padovan, J.C., Chait, B.T., Popowicz, A., Manning, R.S., and Manning, J.M. (2007). Human embryonic, fetal, and adult hemoglobins have different subunit interface strengths. Correlation with lifespan in the red cell. Protein Sci *16*, 1641-1658.

Mariani, C.J., Vasanthakumar, A., Madzo, J., Yesilkanal, A., Bhagat, T., Yu, Y., Bhattacharyya, S., Wenger, R.H., Cohn, S.L., Nanduri, J.*, et al.* (2014). TET1-mediated hydroxymethylation facilitates hypoxic gene induction in neuroblastoma. Cell Rep *7*, 1343-1352.

Meldi, K., Qin, T., Buchi, F., Droin, N., Sotzen, J., Micol, J.-B., Selimoglu-Buet, D., Masala, E., Allione, B., Gioia, D.*, et al.* (2015). Specific molecular signatures predict decitabine response in chronic myelomonocytic leukemia. J Clin Invest *125*, 1857-1872.

Merlevede, J., Droin, N., Qin, T., Meldi, K., Yoshida, K., Morabito, M., Chautard, E., Auboeuf, D., Fenaux, P., Braun, T.*, et al.* (2016). Mutation allele burden remains unchanged in chronic myelomonocytic leukaemia responding to hypomethylating agents. Nat Commun *7*, 10767.

Mettananda, S., and Higgs, D.R. (2018). Molecular Basis and Genetic Modifiers of Thalassemia. Hematol Oncol Clin North Am *32*, 177-191.

Meyer, C., Hofmann, J., Burmeister, T., Gröger, D., Park, T.S., Emerenciano, M., Pombo de Oliveira, M., Renneville, A., Villarese, P., Macintyre, E.*, et al.* (2013). The MLL recombinome of acute leukemias in 2013. Leukemia *27*, 2165-2176.

Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P.D. (2019a). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res *47*.

Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X., and Thomas, P.D. (2019b). Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). Nat Protoc *14*, 703-721.

Minor, E.A., Court, B.L., Young, J.I., and Wang, G. (2013). Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. J Biol Chem.

Misawa, K., Imai, A., Mochizuki, D., Mima, M., Endo, S., Misawa, Y., Kanazawa, T., and Mineta, H. (2018). Association of TET3 epigenetic inactivation with head and neck cancer. Oncotarget *9*, 24480-24493.

Morgan, S.L., Mariano, N.C., Bermudez, A., Arruda, N.L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.-F.*, et al.* (2017). Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. Nat Commun *8*, 15993.

Moshfegh, C.M., Collins, C.W., Gunda, V., Vasanthakumar, A., Cao, J.Z., Singh, P.K., Godley, L.A., and Case, A.J. (2019). Mitochondrial superoxide disrupts the metabolic and epigenetic landscape of CD4+ and CD8+ T-lymphocytes. Redox Biol, 101141.

Murphy, M.P. (2009). How mitochondria produce reactive oxygen species. The Biochemical journal *417*, 1-13.

Nakagawa, T., Lv, L., Nakagawa, M., Yu, Y., Yu, C., D'Alessio, A.C., Nakayama, K., Fan, H.-Y., Chen, X., and Xiong, Y. (2015). CRL4(VprBP) E3 ligase promotes monoubiquitylation and chromatin binding of TET dioxygenases. Mol Cell *57*, 247-260.

Nakazawa, M.S., Keith, B., and Simon, M.C. (2016). Oxygen availability and metabolic adaptations. Nat Rev Cancer *16*, 663-673.

Nombela-Arrieta, C., Pivarnik, G., Winkel, B., Canty, K.J., Harley, B., Mahoney, J.E., Park, S.-Y., Lu, J., Protopopov, A., and Silberstein, L.E. (2013). Quantitative imaging of haematopoietic stem and progenitor cell localization and hypoxic status in the bone marrow microenvironment. Nat Cell Biol *15*, 533-543.

Ono, R., Taki, T., Taketani, T., Taniwaki, M., Kobayashi, H., and Hayashi, Y. (2002). LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). Cancer research *62*, 4075-4080.

Orkin, S.H., and Bauer, D.E. (2019). Emerging Genetic Therapy for Sickle Cell Disease. Annu Rev Med *70*, 257-271.

Ozer, H.G., El-Gamal, D., Powell, B., Hing, Z.A., Blachly, J.S., Harrington, B., Mitchell, S., Grieselhuber, N.R., Williams, K., Lai, T.-H.*, et al.* (2018). BRD4 Profiling Identifies Critical Chronic Lymphocytic Leukemia Oncogenic Circuits and Reveals Sensitivity to PLX51107, a Novel Structurally Distinct BET Inhibitor. Cancer Discov *8*, 458-477.

Paikari, A., and Sheehan, V.A. (2018). Fetal haemoglobin induction in sickle cell disease. Br J Haematol *180*, 189-200.

Paliege, A., Rosenberger, C., Bondke, A., Sciesielski, L., Shina, A., Heyman, S.N., Flippin, L.A., Arend, M., Klaus, S.J., and Bachmann, S. (2010). Hypoxia-inducible

factor-2alpha-expressing interstitial fibroblasts are the only renal cells that express erythropoietin under hypoxia-inducible factor stabilization. Kidney Int *77*, 312-318.

Pan, F., Wingo, T.S., Zhao, Z., Gao, R., Makishima, H., Qu, G., Lin, L., Yu, M., Ortega, J.R., Wang, J.*, et al.* (2017). Tet2 loss leads to hypermutagenicity in haematopoietic stem/progenitor cells. Nat Commun *8*, 15102.

Pastor, W.A., Aravind, L., and Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nat Rev Mol Cell Biol *14*, 341-356.

Pastore, Y.D., Jelinek, J., Ang, S., Guan, Y., Liu, E., Jedlickova, K., Krishnamurti, L., and Prchal, J.T. (2003). Mutations in the VHL gene in sporadic apparently congenital polycythemia. Blood *101*, 1591-1595.

Patel, J.P., Gönen, M., Figueroa, M.E., Fernandez, H., Sun, Z., Racevskis, J., Van Vlierberghe, P., Dolgalev, I., Thomas, S., Aminova, O.*, et al.* (2012). Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. N Engl J Med *366*, 1079-1089.

Perera, A., Eisen, D., Wagner, M., Laube, S.K., Künzel, A.F., Koch, S., Steinbacher, J., Schulze, E., Splith, V., Mittermeier, N.*, et al.* (2015). TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression. Cell Rep *11*, 283-294.

Peschle, C., Mavilio, F., Carè, A., Migliaccio, G., Migliaccio, A.R., Salvo, G., Samoggia, P., Petti, S., Guerriero, R., and Marinucci, M. (1985). Haemoglobin switching in human embryos: asynchrony of zeta----alpha and epsilon----gamma-globin switches in primitive and definite erythropoietic lineage. Nature *313*, 235-238.

Pidugu, L.S., Flowers, J.W., Coey, C.T., and Biochemistry, P.-E. (2016). Structural basis for excision of 5-formylcytosine by thymine DNA glycosylase. Biochemistry.

Pollard, P.J., Brière, J.J., Alam, N.A., Barwell, J., Barclay, E., Wortham, N.C., Hunt, T., Mitchell, M., Olpin, S., Moat, S.J.*, et al.* (2005). Accumulation of Krebs cycle intermediates and over-expression of HIF1alpha in tumours which result from germline FH and SDH mutations. Hum Mol Genet *14*, 2231-2239.

Qiu, Z., Lin, A.-P., Jiang, S., Elkashef, S.M., Myers, J., Srikantan, S., Sasi, B., Cao, J.Z., Godley, L.A., Rakheja, D.*, et al.* (2020). MYC Regulation of D2HGDH and L2HGDH Influences the Epigenome and Epitranscriptome. Cell Chem Biol *27*, 538-1205032704.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Raffel, S., Falcone, M., Kneisel, N., Hansson, J., Wang, W., Lutz, C., Bullinger, L., Poschet, G., Nonnenmacher, Y., Barnert, A.*, et al.* (2017). BCAT1 restricts αKG

levels in AML stem cells leading to IDHmut-like DNA hypermethylation. Nature *551*, 384-388.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

Rasmussen, K.D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. Genes Dev *30*, 733-750.

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., Wadi, L., Meyer, M., Wong, J., Xu, C.*, et al.* (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc *14*, 482-517.

Reyes-Garau, D., Ribeiro, M.L., and Roué, G. (2019). Pharmacological Targeting of BET Bromodomain Proteins in Acute Myeloid Leukemia and Malignant Lymphomas: From Molecular Characterization to Clinical Applications. Cancers (Basel) *11*, 1483.

Richardson, D.R., Lane, D.J., Becker, E.M., Huang, M.L., Whitnall, M., Suryo Rahmanto, Y., Sheftel, A.D., and Ponka, P. (2010). Mitochondrial iron trafficking and the integration of iron metabolism between the mitochondrion and cytosol. Proc Natl Acad Sci U S A *107*, 10775-10782.

Ricketts, C.J., and Linehan, W.M. (2017). Insights into Epigenetic Remodeling in VHL-Deficient Clear Cell Renal Cell Carcinoma. Cancer Discov *7*, 1221-1223.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.

Robinson, J.T., Turner, D., Durand, N.C., Thorvaldsdóttir, H., Mesirov, J.P., and Aiden, E.L. (2018). Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data. Cell Syst *6*, 256-258.

Roe, J.-S., Mercan, F., Rivera, K., Pappin, D.J., and Vakoc, C.R. (2015). BET Bromodomain Inhibition Suppresses the Function of Hematopoietic Transcription Factors in Acute Myeloid Leukemia. Mol Cell *58*, 1028-1039.

Rudenko, A., Dawlaty, M.M., Seo, J., Cheng, A.W., Meng, J., Le, T., Faull, K.F., Jaenisch, R., and Tsai, L.-H. (2013). Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. Neuron *79*, 1109-1122.

Ruifrok, A.C., and Johnston, D.A. (2001). Quantification of histochemical staining by color deconvolution. Anal Quant Cytol Histol *23*, 291-299.

Russell, J.E., and Liebhaber, S.A. (1998). Reversal of lethal alpha- and beta-thalassemias in mice by expression of human embryonic globins. Blood *92*, 3057-3063.

Safran, M., and Kaelin, W.G. (2003). HIF hydroxylation and the mammalian oxygen-sensing pathway. J Clin invest *111*, 779-783.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods *11*, 783-784.

Sankaran, V.G., Menne, T.F., Šćepanović, D., Vergilio, J.-A., Ji, P., Kim, J., Thiru, P., Orkin, S.H., Lander, E.S., and Lodish, H.F. (2011). MicroRNA-15a and -16-1 act via MYB to elevate fetal hemoglobin expression in human trisomy 13. Proc Natl Acad Sci U S A *108*, 1519-1524.

Sankaran, V.G., and Orkin, S.H. (2013). The switch from fetal to adult hemoglobin. Cold Spring Harb Perspect Med *3*, a011643.

Sankaran, V.G., Xu, J., and Orkin, S.H. (2010). Advances in the understanding of haemoglobin switching. Br J Haematol *149*, 181-194.

Sauntharajah, Y. (2019). Targeting sickle cell disease root-cause pathophysiology with small molecules. Haematologica *104*, 1720-1730.

Seiler, C.L., Fernandez, J., Koerperich, Z., Andersen, M.P., Kotandeniya, D., Nguyen, M.E., Sham, Y.Y., and Tretyakova, N.Y. (2018). Maintenance DNA Methyltransferase Activity in the Presence of Oxidized Forms of 5-Methylcytosine: Structural Basis for Ten Eleven Translocation-Mediated DNA Demethylation. Biochemistry *57*, 6061-6069.

Selimoglu-Buet, D., Rivière, J., Ghamlouch, H., Bencheikh, L., Lacout, C., Morabito, M., Diop, M.b., Meurice, G., Breckler, M., Chauveau, A.*, et al.* (2018). A miR-150/TET3 pathway regulates the generation of mouse and human non-classical monocyte subset. Nat Commun *9*, 5455.

Semenza, G.L., Jiang, B.H., Leung, S.W., Passantino, R., Concordet, J.P., Maire, P., and Giallongo, A. (1996). Hypoxia response elements in the aldolase A, enolase 1, and lactate dehydrogenase A gene promoters contain essential binding sites for hypoxia-inducible factor 1. J Biol Chem *271*, 32529-32537.

Shariati, L., Modaress, M., Khanahmad, H., Hejazi, Z., Tabatabaiefar, M.A., Salehi, M., and Modarressi, M.H. (2016). Comparison of different methods for erythroid differentiation in the K562 cell line. Biotechnol Lett *38*, 1243-1250.

Shi, F.-T., Kim, H., Lu, W., He, Q., Liu, D., Goodell, M.A., Wan, M., and Songyang, Z. (2013). Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. J Biol Chem *288*, 20776-20784.

Slyvka, A., Mierzejewska, K., and Bochtler, M. (2017). Nei-like 1 (NEIL1) excises 5-carboxylcytosine directly and stimulates TDG-mediated 5-formyl and 5-carboxylcytosine excision. Sci Rep *7*, 9001.

Sohni, A., Bartoccetti, M., Khoueiry, R., Spans, L., Vande Velde, J., De Troyer, L., Pulakanti, K., Claessens, F., Rao, S., and Koh, K.P. (2015). Dynamic switching of active promoter and enhancer domains regulates Tet1 and Tet2 expression during cell state transitions between pluripotency and differentiation. Mol Cell Biol *35*, 1026-1042.

Song, C.-X., Szulwach, K.E., Fu, Y., Dai, Q., Yi, C., Li, X., Li, Y., Chen, C.-H., Zhang, W., Jian, X.*, et al.* (2011). Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat Biotechnol *29*, 68-72.

Song, S.J., Ito, K., Ala, U., Kats, L., Webster, K., Sun, S.M., Jongen-Lavrencic, M., Manova-Todorova, K., Teruya-Feldstein, J., Avigan, D.E.*, et al.* (2013). The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation. Cell Stem Cell *13*, 87-101.

Spencer, J.A., Ferraro, F., Roussakis, E., Klein, A., Wu, J., Runnels, J.M., Zaher, W., Mortensen, L.J., Alt, C., Turcotte, R.*, et al.* (2014). Direct measurement of local oxygen concentration in the bone marrow of live animals. Nature *508*, 269-273.

Stadhouders, R., Aktuna, S., Thongjuea, S., Aghajanirefah, A., Pourfarzad, F., van Ijcken, W., Lenhard, B., Rooks, H., Best, S., Menzel, S.*, et al.* (2014). HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. J Biol Chem *124*, 1699-1710.

Stamatoyannopoulos, G. (2005). Control of globin gene expression during development and erythroid differentiation. Exp Hematol *33*, 259-271.

Stathis, A., and Bertoni, F. (2018). BET Proteins as Targets for Anticancer Treatment. Cancer Discov *8*, 24-36.

Stonestrom, A.J., Hsu, S.C., Jahn, K.S., Huang, P., Keller, C.A., Giardine, B.M., Kadauke, S., Campbell, A.E., Evans, P., Hardison, R.C.*, et al.* (2015). Functions of BET proteins in erythroid gene expression. Blood *125*, 2825-2834.

Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S., and Jacobsen, S.E. (2011). 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome Biol *12*.

Suzuki, N., Vojnovic, N., Lee, K.-L., Yang, H., Gradin, K., and Poellinger, L. (2018). HIF-dependent and reversible nucleosome disassembly in hypoxia-inducible gene promoters. Exp Cell Res *366*, 181-191.

Szulwach, K.E., Li, X., Li, Y., Song, C.-X., Han, J.W., Kim, S., Namburi, S., Hermetz, K., Kim, J.J., Rudd, M.K.*, et al.* (2011). Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. PLoS Genet *7*.

Taher, A.T., Weatherall, D.J., and Cappellini, M.D. (2018). Thalassaemia. Lancet *391*, 155-167.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L.*, et al.* (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Tardy-Planechaud, S., Fujimoto, J., Lin, S.S., and Sowers, L.C. (1997). Solid phase synthesis and restriction endonuclease cleavage of oligodeoxynucleotides containing 5-(hydroxymethyl)-cytosine. Nucleic Acids Res *25*, 553-559.

Tashi, T., Scott Reading, N., Wuren, T., Zhang, X., Moore, L.G., Hu, H., Tang, F., Shestakova, A., Lorenzo, F., Burjanivova, T.*, et al.* (2017). Gain-of-function EGLN1 prolyl hydroxylase (PHD2 D4E:C127S) in combination with EPAS1 (HIF-2α) polymorphism lowers hemoglobin concentration in Tibetan highlanders. J Mol Med (Berl) *95*, 665-670.

Thein, S., and Menzel, S. (2009). Discovering the genetics underlying foetal haemoglobin production in adults. Br J Haematol *145*, 455-467.

Togasaki, E., Takeda, J., Yoshida, K., Shiozawa, Y., Takeuchi, M., Oshima, M., Saraya, A., Iwama, A., Yokote, K., Sakaida, E.*, et al.* (2017). Frequent somatic mutations in epigenetic regulators in newly diagnosed chronic myeloid leukemia. Blood Cancer J *7*, e559.

Trakarnsanga, K., Griffiths, R.E., Wilson, M.C., Blair, A., Satchwell, T.J., Meinders, M., Cogan, N., Kupzig, S., Kurita, R., Nakamura, Y.*, et al.* (2017). An immortalized adult human erythroid line facilitates sustainable and scalable generation of functional red cells. Nat Commun *8*, 14750.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Turcan, S., Rohle, D., Goenka, A., Walsh, L.A., Fang, F., Yilmaz, E., Campos, C., Fabius, A.W., Lu, C., Ward, P.S.*, et al.* (2012). IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature *483*, 479-483.

Turinetto, V., and Giachino, C. (2015). Histone variants as emerging regulators of embryonic stem cell identity. Epigenetics *10*, 563-573.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A.*, et al.* (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419.

Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F.*, et al.* (2017). A pathology atlas of the human cancer transcriptome. Science *357*.

Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. (2005). Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. Mol Cell *17*, 453-462.

Valinluck, V., and Sowers, L.C. (2007). Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. Cancer research *67*, 946-950.

Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E.*, et al.* (2013). Dynamic DNA methylation across diverse human cell lines and tissues. Genome Res *23*, 555-567.

Vasanthakumar, A., and Godley, L.A. (2015). 5-hydroxymethylcytosine in cancer: significance in diagnosis and therapy. Cancer genetics *208*, 167-177.

Vella, P., Scelfo, A., Jammula, S., Chiacchiera, F., Williams, K., Cuomo, A., Roberto, A., Christensen, J., Bonaldi, T., Helin, K.*, et al.* (2013). Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. Mol Cell *49*, 645-656.

Verma, N., Pan, H., Doré, L.C., Shukla, A., Li, Q.V., Pelham-Webb, B., Teijeiro, V., González, F., Krivtsov, A., Chang, C.-J.J.*, et al.* (2018). TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. Nat Genet *50*, 83-95.

Vernimmen, D. (2014). Uncovering enhancer functions using the α-globin locus. PLoS Genet *10*, e1004668.

Villaseñor, R., and Baubec, T. (2020). Regulatory mechanisms governing chromatin organization and function. Curr Opin Cell Biol *70*, 10-17.

Vincent, J.J., Huang, Y., Chen, P.-Y., Feng, S., Calvopiña, J.H., Nee, K., Lee, S.A., Le, T., Yoon, A.J., Faull, K.*, et al.* (2013). Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells. Cell Stem Cell *12*, 470-478.

Vinjamur, D.S., and Bauer, D.E. (2018). Growing and Genetically Manipulating Human Umbilical Cord Blood-Derived Erythroid Progenitor (HUDEP) Cell Lines. Methods Mol Biol *1698*, 275-284.

Völker-Albert, M., Bronkhorst, A., Holdenrieder, S., and Imhof, A. (2020). Histone Modifications in Stem Cell Development and Their Clinical Implications. Stem Cell Reports *15*, 1196-1205.

Wang, X., Angelis, N., and Thein, S.L. (2018). MYB - A regulatory factor in hematopoiesis. Gene *665*, 6-17.

Wen, L., Li, X., Yan, L., Tan, Y., Li, R., Zhao, Y., Wang, Y., Xie, J., Zhang, Y., Song, C.*, et al.* (2014). Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. Genome Biol *15*.

Wilber, A., Nienhuis, A.W., and Persons, D.A. (2011). Transcriptional regulation of fetal to adult hemoglobin switching: new therapeutic opportunities. Blood *117*, 3945-3953.

Wu, D., Hu, D., Chen, H., Shi, G., Fetahu, I.S., Wu, F., Rabidou, K., Fang, R., Tan, L., Xu, S.*, et al.* (2018). Glucose-regulated phosphorylation of TET2 by AMPK reveals a pathway linking diabetes to cancer. Nature *559*, 637-641.

Wu, H., and Sun, Y.E. (2006). Epigenetic regulation of stem cell differentiation. Pediatr Res *59*.

Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. Nat Rev Genet *18*, 517-534.

Wyatt, G.R., and Cohen, S.S. (1953). The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. The Biochemical journal *55*, 774-782.

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

Xu, W., Yang, H., Liu, Y., Yang, Y., Wang, P., Kim, S.-H.H., Ito, S., Yang, C., Wang, P., Xiao, M.-T.T.*, et al.* (2011). Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer cell *19*, 17-30.

Xu, Y., Xu, C., Kato, A., Tempel, W., Abreu, J.G., Bian, C., Hu, Y., Hu, D., Zhao, B., Cerovina, T.*, et al.* (2012). Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. Cell *151*, 1200-1213.

Yamaguchi, S., Hong, K., Liu, R., Shen, L., Inoue, A., Diep, D., Zhang, K., and Zhang, Y. (2012). Tet1 controls meiosis by regulating meiotic gene expression. Nature *492*, 443-447.

Yamazaki, J., Jelinek, J., Lu, Y., Cesaroni, M., Madzo, J., Neumann, F., He, R., Taby, R., Vasanthakumar, A., Macrae, T.*, et al.* (2015). TET2 Mutations Affect Non-CpG

Island DNA Methylation at Enhancers and Transcription Factor-Binding Sites in Chronic Myelomonocytic Leukemia. Cancer research *75*, 2833-2843.

Yamazaki, J., Taby, R., Vasanthakumar, A., Macrae, T., Ostler, K.R., Shen, L., Kantarjian, H.M., Estecio, M.R., Jelinek, J., Godley, L.A.*, et al.* (2012). Effects of TET2 mutations on DNA methylation in chronic myelomonocytic leukemia. Epigenetics *7*, 201-207.

Yan, H., Wang, Y., Qu, X., Li, J., Hale, J., Huang, Y., An, C., Papoin, J., Guo, X., Chen, L.*, et al.* (2017). Distinct roles for TET family proteins in regulating human erythropoiesis. Blood.

Yang, C., Ota-Kurogi, N., Ikeda, K., Okumura, T., Horie-Inoue, K., Takeda, S., and Inoue, S. (2020a). MicroRNA-191 regulates endometrial cancer cell growth via TET1-mediated epigenetic modulation of APC. J Biochem *168*, 7-14.

Yang, D., Wu, X., Zhou, Y., Wang, W., and Wang, Z. (2020b). The microRNA/TET3/REST axis is required for olfactory globose basal cell proliferation and male behavior. EMBO Rep *21*.

Yao, X., Tan, J., Lim, K.J., Koh, J., Ooi, W.F., Li, Z., Huang, D., Xing, M., Chan, Y.S., Qu, J.Z.*, et al.* (2017). VHL Deficiency Drives Enhancer Activation of Oncogenes in Clear Cell Renal Cell Carcinoma. Cancer Discov *7*, 1284-1305.

Ye, D., Guan, K.-L., and Xiong, Y. (2018). Metabolism, Activity, and Targeting of D- and L-2-Hydroxyglutarates. Trends Cancer *4*, 151-165.

Yin, R., Mao, S.-Q., Zhao, B., Chong, Z., Yang Ying, Zhao, C., Zhang, D., Huang, H., Gao, J., Li, Z.*, et al.* (2013). Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. J Am Chem Soc.

Yoon, D., Pastore, Y.D., Divoky, V., Liu, E., Mlodnicka, A.E., Rainey, K., Ponka, P., Semenza, G.L., Schumacher, A., and Prchal, J.T. (2006). Hypoxia-inducible factor-1 deficiency results in dysregulated erythropoiesis signaling and iron homeostasis in mouse development. J Biol Chem *281*, 25703-25711.

Zhang, P., Rausch, C., Hastert, F.D., Boneva, B., Filatova, A., Patil, S.J., Nuber, U.A., Gao, Y., Zhao, X., and Cardoso, M.C. (2017a). Methyl-CpG binding domain protein 1 regulates localization and activity of Tet1 in a CXXC3 domain-dependent manner. Nucleic Acids Res.

Zhang, Q., Liu, X., Gao, W., Li, P., Hou, J., Li, J., and Wong, J. (2014). Differential regulation of the ten-eleven translocation (TET) family of dioxygenases by O-linked β-N-acetylglucosamine transferase (OGT). J Biol Chem *289*, 5986-5996.

Zhang, R.-R., Cui, Q.-Y., Murai, K., Lim, Y.C., Smith, Z.D., Jin, S., Ye, P., Rosa, L., Lee, Y.K., Wu, H.-P.*, et al.* (2013). Tet1 regulates adult hippocampal neurogenesis and cognition. Cell Stem Cell *13*, 237-245.

Zhang, S., Chen, Y., Tian, C., He, Y., Tian, Z., Wan, Y., and Liu, T. (2020). Dual-target inhibitors based on BRD4: Novel therapeutic approaches for cancer. Current Med Chem.

Zhang, X., Su, J., Jeong, M., Ko, M., Huang, Y., Park, H.J., Guzman, A., Lei, Y., Huang, Y.-H., Rao, A.*, et al.* (2016). DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. Nat Genet *48*, 1014-1023.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhang, Y.W., Wang, Z., Xie, W., Cai, Y., Xia, L., Easwaran, H., Luo, J., Yen, R.-W.C., Li, Y., and Baylin, S.B. (2017b). Acetylation Enhances TET2 Function in Protecting against Abnormal DNA Methylation during Oxidative Stress. Mol Cell *65*, 323-335.

Zhao, Z., Chen, L., Dawlaty, M.M., Pan, F., Weeks, O., Zhou, Y., Cao, Z., Shi, H., Wang, J., Lin, L.*, et al.* (2015). Combined Loss of Tet1 and Tet2 Promotes B Cell, but Not Myeloid Malignancies, in Mice. Cell Rep *13*, 1692-1704.

Zheng, W., Dong, X., Yin, R., Xu, F., Ning, H., Zhang, M., Xu, C., Yang, Y., Ding, Y., Wang, Z.*, et al.* (2014). EDAG positively regulates erythroid differentiation and modifies GATA1 acetylation through recruiting p300. Stem Cells *32*, 2278-2289.

Zhong, J.-Y.Y., Cui, R.-R.R., Lin, X., Xu, F., Zhu, T., Li, F., Wu, F., Zhou, E., Yi, L., and Yuan, L.-Q.Q. (2019). Aberrant DNA methylation of synaptophysin is involved in adrenal cortisol-producing adenoma. Aging *11*, 5232-5245.

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