T. Michael Creed
TMCreed09@gmail.com
502-558-3091

Education

PhD Molecular Medicine — 2016-2020
Degree Conferred December 2020
University of Maryland Baltimore
Baltimore, MD

MS Molecular Biology — 2009-2011
Degree Conferred May 2011
Murray State University
Murray, KY

BA German Studies. Minor in Chemistry — 2005-2009
Degree Conferred May 2009
Transylvania University
Lexington, KY

Research experience

Graduate research (PhD) — 2016-2020
Center for Stem Cell Biology, University of Maryland Baltimore.
PI: Tami Kingsbury, PhD, and Curt Civin, MD

- Discovered a novel regulator of erythropoiesis, SIX1, and its interaction with the GATA:FOG network
- Applied CRISPR/Cas9 technologies to generate multiple knockout/knockdown cell lines in -TF1 and primary cell cultures
- Optimized *in vitro* differentiation of the TF1 erythroleukemia cell model
- Utilized BioID technology to identify novel SIX1 interactors and designed a CRISPR/Cas9-based screen to determine their necessity during SIX1-induced erythropoiesis
- Designed and produced lentiviral-based delivery system for protein expression and CRISPR/Cas9 genome editing in cell cultures

Guest Researcher, National Institute on Aging, Johns Hopkins University
PI: Supriyo De, PhD

- Developed R-based strategies to analyze ChIP- and RNA-seq datasets from SIX1 cell cultures

Senior Research Assistant — 2011-2015
Kidney Disease Program, School of Medicine, University of Louisville.
PI: Kenneth McLeish, MD

- Determined key molecular signaling pathways of TNFα-stimulation in neutrophil priming
- Established kinetic assays for the prolyl isomerase, PIN1
- Optimized electroporation protocols for cell cultures and patient-derived neutrophils to perturb molecular networks necessary for endo/exocytosis

Graduate Research Assistantship — 2009-2011
Department of Biology, Murray State University.
PI: Alexey Arkov, PhD

- Determined the interactome of *tudor* proteins and their function in fly germ-cell development

Theses

PhD thesis (December 2020)
Creed TM. “SIX1 transcription factor enhances human erythropoiesis via GATA1.” University of Maryland, Baltimore. Baltimore, MD.

MS thesis
Creed TM, (2011) “Analysis of the *tudor* interactome in drosophila.” Murray State University. Murray, KY.

Publications

Creed TM, Baldeosingh R, Eberly CL, Schlee CS, Kim M, Cutler JA, Pandey A, Civin CI, Fossett NG, Kingsbury TJ. “The PAX-SIX-EYA-DACH network modulates GATA-FOG function in fly hematopoiesis and human erythropoiesis.” Development. Jan 2020: 147(1)

Caster D, Korte E, Tan M, Barati M, Tandon S, **Creed TM**, Salant D, Hata J, Epstein P, Huang H, Powell D, McLeish K. “Neutrophil exocytosis induces podocyte cytoskeletal reorganization and proteinuria glomerulonephritis” AJPRenal. Sep 1, 2018; 315(3):F595-F606

McLeish KR, Merchant ML, **Creed TM**, Tandon S, Barati MT, Uriarte SM, Ward RA. “Frontline Science: Tumor necrosis factor-α stimulation and priming of human neutrophil granule exocytosis.” J. Leukoc Biol. 2017 Jul; 102(1):19-29

Creed TM, Tandon S, Ward RA, McLeish KR. “Endocytosis is required for exocytosis and priming of respiratory burst activity in human neutrophils.” Inflamm. Res. 2017; 66(10):891-899

Mahmood J, Shamah AA, **Creed TM**, Pavlovic R, Matsui H, Kimura M, Molitoris J, Shukla H, Jackson I, Vujaskovic Z. “Radiation-induced erectile dysfunction: Recent advances and future directions” Adv. Radiat Oncol. 2016; 1(3)161-169

Gao M, Thomson TC, **Creed TM**, Tu S, Loganathan SN, Jackson CA, McCluskey P, Lin Y, Collier SE, Weng Z, Lasko P, Ohi MD, Arkov AA. (2014) “Glycolytic enzymes localize to ribonucleoprotein granules in Drosophila germ cells, bind *tudor* and protect from transposable elements.” EMBO Reports. 2015 Mar; 16(3):379-386

Krzyzanki W, Brier ME, **Creed TM**, Gaweda AE, “Reticulocyte-based estimation of red blood cell lifespan.” Experimental Hematology. 2013: 41(9):817-22

McLeish KR, Uriarte SM, Tandon S, **Creed TM**, Le J, Ward RA “Exocytosis of neutrophil granule subsets and activation of prolyl isomerase 1 are required for respiratory burst priming.” Journal of Innate Immunity. 2013;5(3):277-89

Creed TM, Loganathan SN, Varonin D, Jackson CA, Arkov AL. “Novel role of specific tudor domains in *tudor-aubergine* protein complex assembly and distribution during drosophila oogenesis.” Biochemica Biophysica Research Communications. 2010; 402(2):384-9

Poster and conference presentations

Creed TM (presenter), Eberly CE, Kim MJ, Cutler JA, Pandey A, Civin CI, Kingsbury TJ. “SIX1 transcription factor enhances human erythropoiesis via GATA1.” The American Society of Hematology (ASH) Annual Meeting. Orlando, FL. Dec. 2019.

Creed TM (presenter), Eberly CE, Kim MJ, Civin CI, Kingsbury TJ. “SIX1 associates with GATA complexes and drives human erythropoiesis.” Red Cell Club Meeting. Rochester, NY. Oct. 2019.

Creed TM (presenter) Eberly C, Shlee C, Raj, Cutler J, Pandey A, Kim MJ, Civin C, Fossett N, Kingsbury T. “SIX2 interacts with GATA1 and drives erythroid differentiation in the TF1 model.” University of Maryland, Baltimore Annual Cancer Retreat. April 2018.

Creed TM (presenter), Eberly C, Shlee C, Raj, Cutler J, Pandey A, Kim MJ, Civin C, Fossett N, Kingsbury T. “PSEDN member SIX terminally differentiates erythroleukemia and interacts with GATA network” University of Maryland Baltimore Annual Cancer Retreat. April 2017.

Creed TM (presenter), Loganathan S, Jackson C, Varonin D, Arkov A. “Studying germline stem cells using the fruit fly as a model system.” NSF KY EPSCoR Conference. Lexington, KY. May 2010.

Advanced skills

Expertise highlights: CRISPR/Cas9 (KO, KD, activation) sgRNA and cell line generation (lentiviral and RNP electroporation), flow cytometry, western blotting, qPCR, cell staining, bacterial cloning, lentivirus generation, mammalian cell culture, protein-protein interaction assays.

Computer program/language skills: R Studio (highly experienced), Python, ImageJ, FlowLogic, FlowJo, GraphPad Prism, QuantStudio, ImageLab, Adobe Illustrator.

Mentor/supervisory experience: Taught and managed seven research assistants, medical students, and rotation students.

Collaboration highlights: With Feyruz Rassool, PhD, designed and generated CRISPR/Cas9 KO cell lines in MOLM14 and A549 cell cultures. With Javed Mahmood, PhD, generated CRISPR/Cas9 knockdown and overexpression lines of Caveolin-1 for radiation-induced injury studies.

Awards and honors

Full-tuition and research stipends:

Graduate Research Stipend, University of Maryland, Baltimore.	2015-2020
Innovation Research Assistantship Award, Murray State University.	2010

Other:

Graduate Poster Competition Award, Sigma Xi. Murray State University.	2010
Outstanding Student Research Award, Sigma Xi. Murray State University.	2010

Abstract

Title of Dissertation:

The PAX-SIX-EYA-DACH Network regulates GATA factors and drives human erythropoiesis

T. Michael Creed, Doctor

of Philosophy, 2020

Dissertation Directed by:

Tami Kingsbury, PhD,
Assistant Professor
Department of
Physiology
University of Maryland, Baltimore

Erythropoiesis is orchestrated by the coordinated action of multiple transcription factors. The master erythropoietic-regulator GATA1 is itself modulated via interactions with multiple co-regulatory factors. Though the PAX-SIX-EYA-DACH network (PSEDN) of conserved transcription factors has been well characterized in human organogenesis, a role for PSEDN members in hematopoietic systems has only recently been recognized (Liu *et al.*, Nature 2019). Here we studied the PSEDN member SIX and discovered its ability to drive erythropoiesis of human hematopoietic cells.

Overexpression of SIX1 or SIX2 in human TF1 erythroleukemia or primary $CD34^+$ hematopoietic stem- progenitor cells (HSPCs) stimulated the generation of erythroid cells, as measured by flow cytometry, qPCR, and Western blot. Conversely, SIX1 knockout in TF1 cells or primary HSPCs reduced erythroid cell generation in response to erythropoietin. By gene set enrichment analysis of RNA-seq data, SIX1/SIX2 overexpression stimulated heme metabolism genes as well as genes

known to be regulated by GATA1. SIX1/SIX2 overexpression reduced GATA2 and increased GATA1 expression, resembling GATA switching downstream of EPO signaling. To determine whether GATA1 was necessary for SIX1 to stimulate erythropoiesis, we generated GATA1-knockout cells using CRISPR/Cas9. In contrast to control cells, SIX1 OE in GATA1-knockout cells failed to stimulate erythropoiesis, indicating that SIX1 stimulation of erythropoiesis requires GATA1.

To gain further insight into the mechanism by which SIX1/SIX2 stimulates erythropoiesis we generated a BirA*-SIX fusion protein to determine its interactome. Streptavidin-enrichment of biotinylated proteins in BirA*-SIX overexpression lysates revealed GATA1 and FOG1 as proximal interactors of BirA*-SIX. When co-expressed in HEK293T cells GATA1 and SIX1 were found to co- immunoprecipitate, suggesting the two proteins can physically interact in a complex. We demonstrated the functional consequence of the SIX1/SIX2 and GATA1 using a GATA1-dependent luciferase reporter construct. Cells in which SIX1/SIX2 and GATA1 were co-expressed exhibited significantly higher levels of luciferase expression compared to cells expressing only GATA1, suggesting SIX1 could stimulate GATA1-dependent transcription. Together our results suggest that SIX1 can stimulate erythropoiesis via multiple mechanisms. This thesis provides the first demonstration of a role for the PSEDN in erythropoiesis and reveal physical and functional interactions between two central developmental transcriptional networks (GATA/FOG and PSEDN).

The PAX-SIX-EYA-DACH Network interacts with
GATA factors in human erythropoiesis

by
Timothy Michael Creed

Dissertation submitted to the faculty of the Graduate School of the
University of Maryland, Baltimore in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Acknowledgements

First and foremost, I would like to thank my mentors, Dr. Tami Kingsbury and Dr. Curt Civin, for their support and guidance throughout my graduate studies.

Tami, you taught me to always trust my data, even when it infuriatingly told the opposite story than we had expected. You kept me focused and driven even when the data and project scope was overwhelming. I have learned invaluable aspects of what it takes to be great scientist while working for you. Your incredible work ethic and unwavering enthusiasm for science kept me constantly engaged with my research and made my time in your lab truly incredible. I will forever be grateful for your mentorship.

Curt, through our daily discussions and monthly meetings, you helped me hone my critical thinking skills which will undoubtedly fuel my entire scientific career. Thank you to my committee: Dr. Feyruz Rassool, Dr. Stephen Baylin, Dr. Chozha Rathinam, and Dr. Ricardo Feldman for so generously sharing your time and knowledge. Additionally, I owe a great deal of gratitude to Dr. Nancy Fossett, whose discoveries helped jumpstart my project and direct experiments which led to our publication in *Development.* Next, I would like to thank the members of the Civin/Kingsbury lab, both current and past: MJ, Christian, Blake, Caroline, Tristan, and Miya. This thesis has my name on it, but I truly could not have completed this work without your expertise and friendship. I want to specially thank Christian for help in figure making, experiment work, and the all-important series of bucketball. To my family in Kentucky and Indiana, thank you for encouraging me to move across the country and pursue my dreams. And last but certainly not least, thank you, Annie. You have given me more than you can imagine. It's along your side that I have confidence to dream big and leap. Thank you for your patience and

encouragement through the long hours and weekends in lab. I could have not done this without you (or our two pups Molly and Pancake), and I wouldn't have wanted to.

Table of Contents

List of Tables

List of Figures

Chapter 1 figures

Chapter 2 figures

Chapter 3 figures

Chapter 4 figures

List of Abbreviations

BFU-E	Burst forming unit - erythroid
CD34	Hematopoietic progenitor cell marker
CFU-E	Colony forming unit - erythroid
CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
CoIP	Coimmunoprecipitation
DC	Dendritic cell
EPO	Erythropoietin
EPOR	Erythropoietin receptor
ETS	Erythroblast transformation specific
EV	Empty vector
FOG1	Friend of GATA1
FPKM	Fragments per kilobase of exon per million reads mapped
GFP	Green fluorescent protein
GM-CSF	Granulocyte macrophage colony-stimulating factor
GMP	Granulocyte macrophage progenitor
GOF	Gain of function
GYPA/CD235a	Glycophorin A
HBB	Hemoglobin Beta
HD	Homeobox domain
HIF-2α	Hypoxia-inducible factor 2α

HRP	Horseradish peroxidase
HSC	Hematopoietic stem cell
HSPC	Hematopoietic stem-progenitor cell
KD	Knockdown
KO	Knockout
LT-HSC	Long term hematopoietic stem cell
MEL	Mouse erythroleukemia
MEP	Megakaryocyte erythroid progenitors
MLLr	Mixed lineage leukemia rearrangement
MLL-WT	Mixed lineage leukemia Wild Type
NFκB	Nuclear factor kappa B
NK	Natural killer cell
OE	Overexpression
PCA	principle component analysis
PHD	Prolyl hydroxylase
PLA	Proximity Ligation Assay
PSEDN	PAX-SIX-EYA-DACH Network
PSM	Peptide spectral match
RBC	Red blood cell
ROS	Reactive oxygen species
SCF	Stem cell factor
SD	Six domain
so	*since oculis* (fly homologue of SIX1/SIX2)

SQCC	Squamous cell carcinoma
srp	*serpent* (fly homologue of GATA1)
SSNC	Second-site noncomplementation
ST-HSC	Short term hematopoietic stem cell
TEABC	Trierthylammonium bicarbonate
TFBS	Transcription factor binding site
TFR/CD71	Transferrin receptor
TPO	Thrombopoietin
TSS	Transcription start site
ush	*u-shaped* (fly homologue of FOG1)
VHL	Von-Hippel Lindau
WT	Wild type

Chapter 1: Introduction

1.1 Erythropoiesis overview

Erythropoiesis is the tightly regulated process responsible for daily replenishment of approximately $2x10^{11}$ terminally differentiated, functional erythrocytes from hematopoietic stem-progenitor cells (HSPCs) (Muckenthaler *et al*, Cell 2017). Briefly, hematopoietic stem cells (HSCs) generate multipotent progenitor cells (MPPs), which in turn generate erythroid-committed progenitor cells that are then signaled by extrinsic cytokines, including stem cell factor (SCF) and erythropoietin (EPO) to induce survival, proliferation and differentiation into mature erythrocytes (Dzierzak, CSH 2013; Sauvageau, Oncogene 2004). Because HSCs are responsible for the generation of the entire blood-immune system (not solely red blood cells), their differentiation and proliferation phenotypes must be regulated selectively in response to diverse conditions, such as the stimulation of erythropoiesis in response to anemia, or the stimulation of granulopoiesis in response to infection (Paulson *et al*, Curr Opin in Hematology 2011). Defining the molecular mechanisms regulating the development and maturation of terminally differentiated erythrocytes is paramount to understanding and treating clinical pathologies associated with red blood cells including sickle cell disease, anemias, hemoglobinopathies, and other diseases that lead to decreased oxygen-carrying capacities of erythrocytes.

This work has identified the PAX-SIX-EYA-DACH network (PSEDN) member, SIX1, as a novel transcriptional regulator of erythropoietic functions in conjunction with GATA1, a core regulator of erythropoiesis, to stimulate human erythropoiesis *in vitro*.

Defining how SIX1 functions in erythrocyte development may aid in treatment of red-cell malignancies and pathologies and elucidate novel mechanisms by which SIX1functions during development or malignancies in other developmental disorders.

1.2 Molecular control of erythropoiesis

Erythropoiesis encompasses/refers to the developmental process by which vast quantities of erythrocytes are generated from hematopoietic stem progenitor cells (HSPCs). At the apex of human hematopoietic hierarchy lies the hematopoietic stem cell (HSC). HSCs are exceeding rare in the bone marrow (comprising 1 per 10^5 cells) compared to red blood cells, which populate greater than 99% of total circulating blood cells (Metcalf, Immunity 2007; Manwani and Bieker, Curr Top Dev Biol 2011). The defining characteristics of HSCs are their ability to divide and ultimately give rise to all the cells of each hematopoietic lineage (multipotency) and their ability to self-renew. In addition to these two properties HSCs also lie dormant as they must populate the blood organ for an entire lifespan, leaving HSPCs that are generated from HSCs to yield the huge number of functional hematopoietic ells. HSCs are classified by immunophenotypic analysis and functionally by sustained blood repopulation in serial transplant assays (Mayani, Fac Review 2016). Using fluorescence-assisted cell sorting (FACS) to characterize the dynamic expression of external surface-markers has long been a mainstay for characterizing stages of hematopoiesis (Bender *et al*, Clin Immunol. Immuopath. 1994; Terstappen *et al*, Blood 1991, Civin *et al*, Int J Cell Cloning 1987). For example, human long-term hematopoietic stem cell (LT-HSC) populations can be defined by predictable antigen expression ($CD34^+$ $CD45RA^-$ $CD49f^+$ $CD90^+$ $CD38^-$ Lin^-) (Notta *et al*, Science 2011). Careful consideration of surface-marker expression is crucial for defining cell developmental stage and functionality considering that loss of CD49f and CD90 yields a population of cells that are temporary multipotent progenitors, incapable of reconstituting blood cells in serial transplantation assays (Laurenti, Nature 2018).

HSC viability, self-renewal, proliferation, and differentiation are controlled by both extrinsic and intrinsic elements. Extrinsically HSCs are regulated by their microenvironment and external cytokines such as SCF, which is essential for HSC quiescence, survival, and self-renewal (Rojas-Sutterlin *et al*, Curr Opin Hematol 2014). Intrinsically HSCs are regulated by dynamic signaling cascades vital for maintaining HSC populations, such as the variable expression of cyclins and cyclin-dependent kinases like CDK6. Normally CDK6 is absent in LT-HSCs keeping them quiescent. However, it is potently upregulated in short-term hematopoietic stem cells (ST-HSCs) resulting in entry into the cell cycle and cellular division (Laurenti *et al*, Cell Stem Cell 2015). As HSCs generate daughter cells yielding multipotent- and oligopotent progenitors, their lineage becomes increasingly restricted.

Classically, lineage restriction and differentiation of HSCs during hematopoiesis has been thought of as a hierarchical tree. At the top of the hierarchy lie the HSCs, which then generate MPP cell lineages that branch off to populate either myeloid or lymphoid cells (Figure 1). The oligopotent common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs) are multipotent cell lineages responsible for the vast expansion of cells capable of generating mature blood cells. CLPs give rise to cells of the lymphoid lineage that populate the body’s T-lymphocytes, B-lymphocytes, natural killer (NK) cells, and dendritic cells (DC) (Seita and Weissman, WIRS Bio Med 2011). The rest of the blood cells, including the erythrocytes, megakaryocytes, and granulocytes are generated from the CMPs, all of which express well-defined cell-surface markers. Though the traditional classic hematopoietic hierarchy suggests that MPPs arise from HSPCs and are committed early towards either myeloid or lymphoid lineages, new data stemming from single cell sequencing and improved cell-sorting schemes suggests that MPPs commit very early to a megakaryocytic fate and follow step-wise unipotent commitments to either erythroid, myeloid, or lymphoid lineages (Figure 1) (Notta *et al*, Science 2016; Pellin *et al*, Nature Comm 2019). These new and exciting data are changing the field of hematopoiesis and understanding fully how HSPCs are controlled molecularly might be crucial to correcting hematologic malignancies.

Artificial systems utilizing forced protein expression have produced data that suggest commitment into particular lineages can be influenced by specific nuclear transcription factors. Overexpression of the retinoic acid receptor alpha (RARα) leads to monocyte formation and a decrease in granulopoiesis (Kastner *et al*, Blood 2001). Similarly, underexpression of PU.1 can increase granulopoiesis at the expense of monocyte formation (Dahl *et al*, Nat Immunol 2003). Here we are investigating whether forced overexpression or knockout of PAX-SIX-EYA-DACH network (PSEDN) members influence hematopoiesis in human cell cultures, particularly during erythropoiesis.

Steady-state erythropoiesis occurs in the bone marrow largely at niches called erythroblastic islands. These islands comprise a central macrophage surrounded by terminally differentiating erythrocytes and are the predominant location at which hemoglobins are synthesized (Manwani and Bieker, Curr Top Dev Biol 2008). The central macrophage is essential to the full development of mature red bloods cells as they provide nutrients and proliferative and survival signals to the developing erythrocyte, and are crucial to the enucleation that occurs before the red blood cell enters circulation (Hanspal *et al*, Blood 1998). The processes that must occur between the progenitor and mature erythrocyte stages of erythropoiesis proceed in a predictable manner. Cellular morphology, cell-surface marker expression, gene regulation, and protein production all allow erythropoiesis to be reliably observed and measured (Figure 2).

The first erythroid-committed stage of human erythropoiesis is known as the burst forming unit-erythroid (BFU-E) (Figure 2). BFU-E cells are responsive to EPO and can be classified according to their early expression of the EPO receptor (EPOR) and concomitant loss of the HSPC marker CD34 (Li *et al*, Blood 2014; Tirelli *et al*, Stem Cell Int 2011). These BFU-E cells are an immature cell that divides less frequent than subsequent stages of erythroid cells and in addition to EPO, are responsive to granulocyte/macrophage colony stimulating factor (GM-CSF), SCF, and IL-3, which are necessary to develop grouped clusters of erythroblasts or larger colonies (bursts) of more than 500 erythroblasts (Wu *et al*, Cell 1995). The progression from BFU-E to the more mature colony forming unit-erythroid (CFU-E) is a continuous process that yields cells highly dependent on EPO for their survival but have less, although still retain a slight, self-renewal capacity (Koury, Blood Rev 2014). CFU-Es are more highly proliferative than their BFU-E predecessors and, unlike their predecessors, can survive and proliferate solely in response EPO (Li *et al*, Blood 2014). In mouse cell lines EPOR dependence coincides with the expression of the transferrin receptor (TFR/CD71) and defines erythroid cells that are highly proliferative (Pop *et al*, PLos Bio 2010). The transition between BFU-E and CFU-E is also dependent on differential dependence on GATA factors (Li *et al*, Blood 2014). GATA2 is highly expressed in early progenitor cells, but after erythroid commitment and the subsequent transcriptional initiation of GATA1, it is quickly reduced and replaced by expression of GATA1, a necessary phenomenon known as the “GATA switch” (Dore and Crispino, Blood 2011). The switch refers to the dynamic change required during erythroid development where early progenitor expression of GATA2 is rapidly downregulated upon activation and reliance on GATA1. Chromatin-IP (ChIP) experiments suggest that chromatin occupancy and repression of genes by GATA2 in progenitor cells is replaced by GATA1 during erythropoiesis at several key erythroid genes including the β-hemoglobin locus, resulting in increased globin expression (Grass *et al*, PNAS 2003).

GATA1 activity in developing erythroid cells is modulated through interactions with several known factors including Friend of GATA1 (FOG1), LMO2, PU.1, TAL1, and KLF1, though the complete interactome is incompletely defined (Rodriguez *et al*, EMBO 2005). FOG1 is a cofactor of GATA1 that is necessary for both erythroid and megakaryocytic development. Mutations in the FOG1-binding N-terminal zinc finger of GATA1 leads to several hematologic disorders including severe thrombocytopenia with or without pronounced anemias (Campbell *et al*, Blood 2013).

Following the CFU-E stage is the second phase of erythropoiesis known as erythroid terminal differentiation (ETD) where fully enucleated red bloods are formed and exit the bone marrow into circulation. This terminal differentiation phase is highly dependent on several molecular ontologies including cell cycle regulation, chromatin modifications, and global DNA methylation (Shearstone *et al*, Science 2011). In fact, CFU-E cells dependent on TFR/CD71 expression are synchronized in S-phase of the cell cycle, which is dependent on downregulation of the cell-cycle inhibitor $p57^{KIP2}$ (Pop *et al*, PLoS Bio

2010). Dividing erythroblasts in ETD mature through several phases of erythroid cell divisions from proerythroblasts to basophilic, polychromatophilic, and orthochromatic erythroblasts. As shown in Figure 2, these differentiated phases can be determined morphologically or by measurement of erythroid protein and gene expression, and will be done frequently in this thesis.

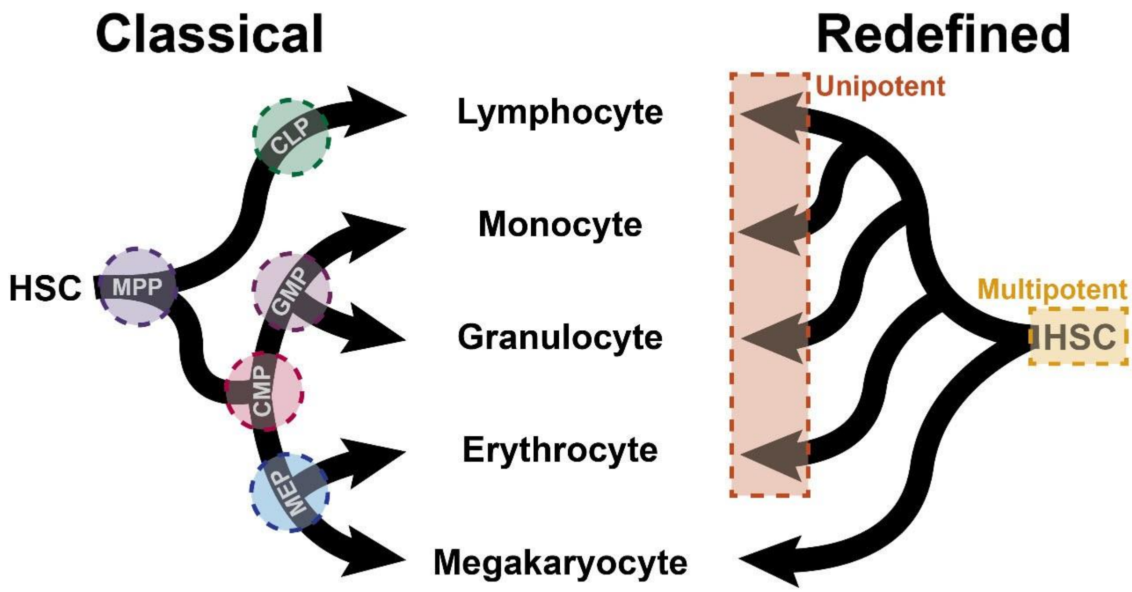

Figure 1.1. Outline of 'Classical' vs 'Redefined' hematopoiesis. Adapted from Notta et al, Science 2016

The classical model exhibits hematopoiesis as a hierarchy where oligopotent hematopoietic stem cells (HSCs) are at the apex and give rise to multipotent progenitors (MPPs) that population either lymphoid or myeloid lineages. Further lineage restriction of MPP produce common myeloid progenitors (CMP) or common lymphoid progenitors (CLP). These progenitor cells populate the further function blood system (MEP = megakaryocyte erythrocyte progenitor, CMP = common myeloid progenitor, GMP = granulocyte macrophage progenitor). The redefined model was reimagined following updated single-cell sequencing and improved immunophenotypical analysis of multipotent HSCs. They found that the megakaryocyte lineage could directly differentiate from HSCs, possibly to regulate the blood stem cell niche. While the stem cell compartment was found to be multipotent, the progenitors instead were unipotent, they did not require intermediate CMP or CLP stages.

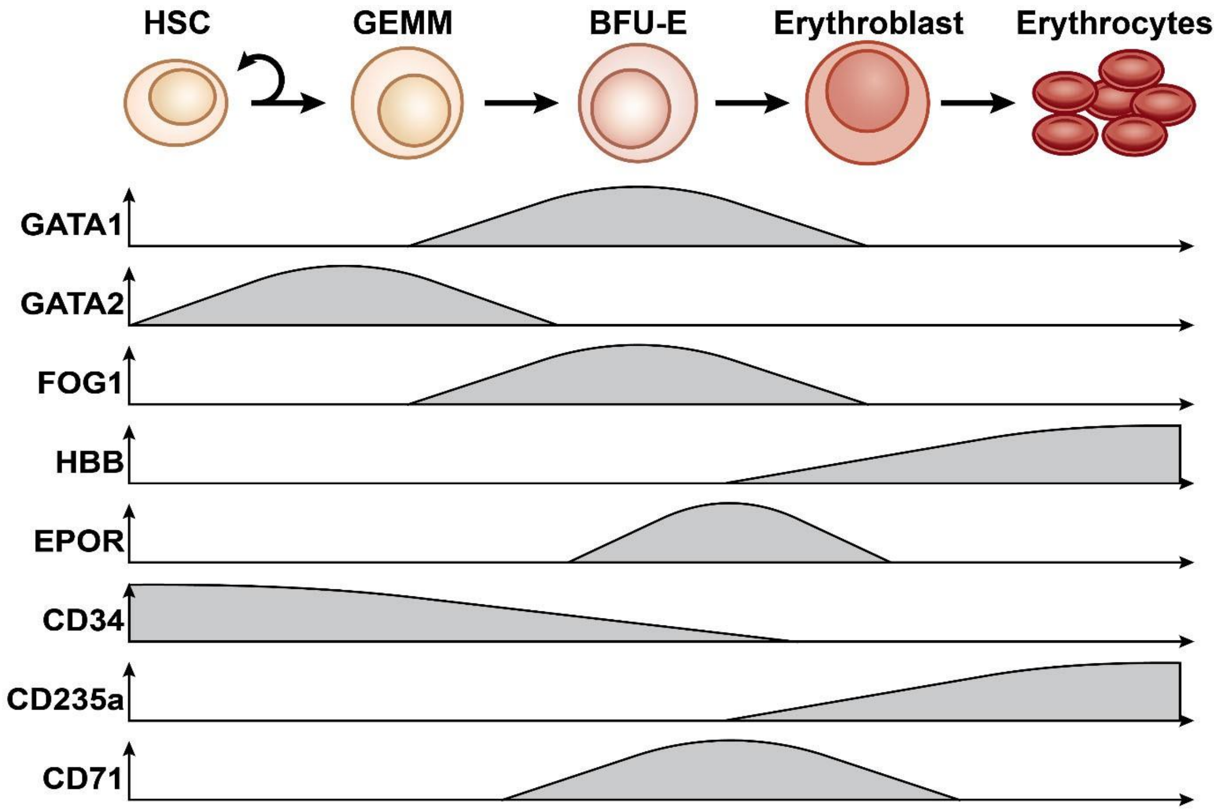

Figure 1.2. Gene expression during erythropoiesis. Adapted from Okumura *et al*, Blood 1991

As hematopoietic stem cells (HSCs) self-renew, proliferate, and differentiate down the erythroid lineage they have a precise and measurable gene-expression pattern. Listed here are some of the well-known stem cell and erythroid genes. The X axis indicates the stage of erythropoiesis from HSC to mature erythrocytes (red blood cells). The Y axis represents expression level of the individual gene. The level of gene expression can be evaluated for each gene during erythropoiesis and will be referred to often in this thesis. (HSC = hematopoietic stem cell, GEMM = granulocyte erythrocyte monocyte megakaryocyte, BFU-E = burst-forming unit erythroid)

1.3 **GATA family in erythropoiesis**

The GATA family of zinc-finger transcription factors comprise six members, GATA1-6. GATA1, GATA2, and GATA3 are expressed predominantly in cells of the hematopoietic lineage and rarely in other adult tissues (with the exception of certain neural tissues and adipocytes) (Jack and Crossley, J Biol Chem 2010; Marjolein *et al*, ERMM 2016). GATA2 is essential for formation and maintenance of HSPC populations (Ling *et al*, J Exp Med 2004). GATA3 is required for lymphocyte development and is involved in regulation of hematopoietic stem cells (Frelin *et al*, Nat Immun 2013; Zaidan and Ottersbach, Open Bio 2018). Both Gata2 and Gata3 knockout mice die embryonically due to disruption of definitive hematopoiesis as well as central nervous system defects in the case of Gata3 (Tsai *et al*, Nature 1994; Pandolfi et al, Nat Genet 1995; Lim et al, Nat Gen 2000). GATA4, GATA5, and GATA6 regulate development of non-hematopoietic tissues, especially heart, intestine, and lung (Simon, Nat. Gen. 1995). These proteins comprise a family due to their similar structure and function (Chlon and Crispino, Development 2012). The characteristic two zinc-finger structure of the genes bind the eponymous sequence [A/T]GATA[A/G] of gene promoters and enhancers. The sequence of the zinc-fingers has a high degree of homology between GATA family members while the remaining domains differ significantly, suggesting that their structure and function may differ due to their contextual expression and interactions with other proteins. To that effect, the zinc fingers are responsible for protein binding as well as their well-known DNA binding capabilities.

The founding member of the GATA family, GATA1, is known as the master erythroid regulator and was originally cloned by the lab of Stu Orkin in 1989 from the mouse erythroleukemia (MEL) cell line (Tsai *et al*, Nature 1989). Previous experiments had identified cis-regulatory modules (CRMs) upstream to the transcription start sites (TSS) of globin genes (Kollias *et al*, Nuc Acids Res 1987; Bodine and Ley, EMBO 1987). Dr. Orkin's group cleverly dissected and isolated major protein bands from heparin-sepharose gels containing K562 nuclear lysates. Radiolabeled-oligonucleotides of the α-globin promoter in conjunction with the isolated nuclear proteins revealed a 50 kDa peptide that resulted in a band shift, suggesting this was the factor that could bind the globin gene promoter (Tsai *et al*, Nature 1989). Lysates from MEL cell lines were also shown to slow the migration of the radiolabeled globin gene promoter, suggesting a conserved function of GATA1 from mouse to humans. cDNA libraries of MEL cells were transfected into monkey kidney COS cell lines, which were then clonally selected and screened in their radiolabeled gel-shift assays until just one clone was found to retain binding activity of the globin gene promoter. DNA sequencing of the isolated clone revealed that the translation of their protein contained two zinc fingers, an N- and C-terminal zinc finger, which are domains capable of protein-protein interactions and binding DNA.

In erythroid development the N-terminal zinc finger is crucial for binding FOG1, an indispensable protein for erythropoiesis (Liew *et al*, PNAS 2005). The C-terminal zinc finger is also capable of binding proteins during erythropoiesis as it binds the transactivation domain (TAD) of p53 (Trainor *et al*, Blood 2009). The process of erythroid differentiation requires enucleation and nuclear condensation thus producing DNA breakage, however the p53 pathway is largely inactive. GATA1 overexpression in cell culture experiments prevents p53-responsive promoters, suggesting that the C-terminal domain of GATA1 may be necessary to prevent p53-induced apoptosis during erythropoiesis.

Though it is expressed in multiple hematopoietic cells (eosinophils, mast cells, megakaryocytes (Crispino, Semin Cell Dev Biol, 2005), GATA1 is indispensable for both primitive and definitive erythropoiesis. GATA1 is responsible for commitment to the erythroid lineage from the HSPCs, where expression correlates to commitment into BFU-Es and progression through terminal differentiation (Zheng *et al*, Blood 2006). Mice lacking Gata1 were able to form all non-hematopoietic tissues, however lacked mature red blood cells and died embryonically due to severe anemia (Pevny *et al*, Nature 1991). Careful analysis revealed that the erythroblasts never differentiated past the proerythroblast stage. Lack of Gata1 had noticeable effects on platelet formation, eosinophil development, and mast cell maturation, however they were not lethal and less severe than erythroid effects (Tsai *et al*, Blood 1997). Gata2, similar to Gata1, is also critical for erythropoiesis and is embryonic lethal in mice. Gata2 haploinsufficient (Gata2$^{+/-}$) mice have a reduced abundance of immunophenotypically defined HSCs which are a result of decreased HSC proliferation, enhanced quiescence, and apoptosis due to the downregulation of anti-apoptotic BCL-XL (Ling *et al*, J Exp Med 2004). Gata2$^{+/-}$ mice have reduced granulocyte macrophage progenitor (GMP) populations while the other myeloid lineages remain intact. These data from GATA1 and GATA2 studies suggest that our studies should likely focus on the involvement of GATA1 in our erythroid experiments.

During erythroid commitment from MPPs and differentiation to mature red blood cells a critical switch known as the “GATA switch,” must occur (Anguita *et al*, EMBO 2004;

Romano *et al*, iScience 2020). This switch is conserved across species and refers the switch in occupancy of GATA-binding sites from GATA2 binding to GATA1. As previously mentioned, GATA2 drives HSPC expansion and proliferation. Studies in mice expressing a Gata2-GFP reporter showed high levels of GFP in isolated HSPC cells (Suzuki *et al*, Genes to Cells 2013). Early on during erythroid commitment, coinciding with the expression of Gata1, the levels of GFP rapidly decreased suggesting a negative-feedback loop and a resulting loss of Gata2. Indeed, there are well-established Gata1-binding sites that lie -1.8kb upstream and 9.5kb downstream of the Gata2 transcription start site (TSS) that are required to maintain Gata2 repression by Gata1 (Wlodarski *et al*, Semin Hematol 2017). Additionally, Gata1 contains cis-regulatory elements upstream of its TSS where it acts as a positive-feedback loop, thus increasing its own expression levels – and simultaneously driving erythropoiesis. Erythroid-committed cells that lack GATA1, and thus are incapable of GATA-switching, are incapable of undergoing erythropoiesis (Kitajima *et al*, Genes Dev 2006). The downregulation of GATA2 and the increase of GATA1 leads to altered occupancy at GATA-binding sites specific for erythroid differentiation which can be measured *in vitro* by assays such as Western blot, flow cytometry, and qRT-PCR to measure gene expression.

1.4 <u>The PAX-SIX-EYA-DACH network (PSEDN)</u>

The PSEDN is comprised of four families of highly conserved developmental transcription factors that play a pivotal role in development across species. The PSEDN was originally discovered in the fly where it is referred to as the Retinal Gene Determination Network (RDGN) where, among other tissues, it is essential for eye formation (Silver and Rebay, Dev Suppl 2005). Ectopic expression of *eyes absent* (*eya*) or *sine oculus* (*so*) (Drosophila homologues of mammalian EYA and SIX), can force expression of the fly eye on places such as the fly leg (Bonini *et al*, Development 1997). In contrast, the absence of RGDN members blocks eye formation in the fly. The functional conservation of PSEDN members is evidenced by the ability of mouse Eya2 to fully restore eye formation in flies lacking eya (Bonini *et al*, Development 1997). Dr. Nancy Fossett (University of Maryland, Baltimore) discovered that members of the fly RDGN were necessary to maintain hematopoietic progenitors in Drosophila through genetic interactions with the fly homologues of the GATA/FOG network, specifically with the FOG1 homologue *u-shaped* (*ush*). Furthermore, *so* and *eya* were shown through gain- and loss-of-function studies to both control functional lamellocyte (a rare blood type found in flies) differentiation and to alter the expression of the GATA homologue, *serpent* (*srp*). This prompted the question if the well-conserved PSEDN members would translate to human hematopoietic regulation, perhaps through similar interactions with the GATA/FOG network.

The PSEDN has been studied extensively in mammalian development where it is vital for formation of sensory organs, branchial arches, skeletal muscle, lung, and a multitude of additional systems (Fortunato, EvoDevo 2014). The expression of PSEDN-member sine oculis homeobox-1 (SIX1) is believed to be lost in adult tissues, but that is changing as SIX1 is shown to be expressed in muscle satellite cells where it plays a role in muscle regeneration and in macrophages where it exhibits immunomodulatory functions in conjunction with NFκB signaling (Le Grand *et al*, JCB 2012; Zhang *et al*, J Cell Mol Med 2019). However, the role of the PSEDN has only recently been recognized in hematopoiesis. We expanded this to erythropoiesis and provided evidence for unexpected physical and functional interactions between SIX1 and GATA1.

1.5 <u>Roles of SIX1 in mammalian development</u>

The SIX family of proteins are homeobox-containing transcriptional regulators required during human development of numerous organs including craniofacial structures, auditory system, brain, lung, muscle, gonads, kidney, and retina (Blevins *et al*., Expert Opin Ther Targets 2015). The first SIX family member, *sine oculis* (*so*), was discovered in the fly as a requirement for normal visual system development (Cheyette *et al*, Neuron 1994). Loss of function studies showed that development of both the compound eye as well as the entire visual system was disrupted in *so* mutants demonstrating its regulatory role in eye development (Weasner *et al*, Dev Bio 2007). Soon after *so* discovery in the fly additional homologues were discovered in the mouse (Oliver *et al*, Development 1995). The existence of six individual Six family members allows for diversification and differences in expression pattern and function. RNA *in situ* hybridization studies mapped them to distinct layers of the developing mouse embryo (Wu *et al*. Int J Biol Sci, 2014). More specifically Six3 functions in neural plate, eye, brain, and sensory organ development whereas Six1 and Six2 showed broader expression throughout mesodermal tissues. Humans have six individual members of the SIX family of proteins (SIX1-6), which can be subdivided into three separate groups that correspond to a specific Drosophila ancestor; SIX1/2 (*so*), SIX3/6 (*optix*), SIX4/5 (*Dsix4*) (Figure 1). SIX1 is the most extensively studied of the SIX family of proteins as it has been shown to regulate development of multiple organ systems and is reactivated and plays prominent roles in carcinogenesis (discussed below).

Given its broad importance in organogenesis it is unsurprising that Six1 deletion in mice results in a host of defects and is lethal in neonatal mice. The Six1 knockout mouse dies shortly after birth due to a hypoplastic skeletal muscle of the diaphragm, which inhibits the ability to breathe (Laclef *et al*, Development 2003). Defects in other organs include hypoplastic kidneys (Li *et al*, Nature 2003), craniofacial structure abnormalities (Laclef *et al*, Development 2003), and flawed olfactory epithelia (Zou *et al*, Development 2004). Deletion of hematopoetic Six1 in $Six1^{fl/fl}$ Scl-Cre-ER^{T} mice showed normal blood cell formation with the potential exception of megakaryocyte erythroid progenitors (MEPs), which did not reach significance in Six1 knockout mice, but the difference appears obvious from their data (Zhang *et al*, EBioMedicine 2019). These data suggest that Six1 may play a role in MEP development. It is important to note that in light of our findings in human erythropoiesis, it would be interesting to look at what happens in Six1-null mice during stress erythropoietic conditions and not homeostatic conditions. Erythroid pathways rely on steady-state and stress erythropoiesis to maintain homeostatic levels of red blood cells and have evolved specialized pathways in each condition. For example, BMP4-dependent stress erythropoiesis is connected to inflammatory signals but is otherwise dispensable for steady-state erythropoiesis (Paulson *et al*, Exp Hem 2020).

SIX1 is known to be vital in muscle development through transcriptional regulation of myogenic regulatory factors (MRFs). During myogenesis in the mouse embryo Six1 controls expression of the muscle-specific genes myogenic factor 5 (Myf5) and MyoD (Kumar, Cell Mol Life Sci 2009). Mice lacking Six1 die at birth due to malformations of the ribs and severe muscle hypoplasia throughout the body. Interestingly, Six1 is only thought to be expressed in normal adult cells in muscle satellite cells, a specific type of muscle progenitor cell necessary for replenishment of damaged muscle tissue. SIX1 null satellite cells were shown to be inefficient at regenerating muscle tissue due to their inability to activate quiescent tissues, demonstrating their control of progenitor cells leads to downstream muscle regeneration (Le Grand *et al*, JCB 2012). SIX1 has additionally been shown to reprogram slow-twitch fibers to fast-twitch in response to increased load, suggesting multiple regulatory roles in cellular differentiation (Wu *et al*, Int J Biol Sci 2014).

The development of the kidney is also dependent on functional Six1. Normal mammalian kidney development requires interaction between the ureteric bud and metanephric mesenchyme. Critical to this process is the glial cell-line-derived neurotrophic factor (GDNF), which regulates ureteric bud outgrowth (Costantini and Shakya, Bioessays 2006). Through its interaction with Eya1, Six1 has been shown to regulate transcription of GDNF. Furthermore, loss of Six1 results in failure of ureteric bud invasion leading to apoptosis of the metanephric mesenchyme and disrupted kidney development (Wu *et al*, Int J Cancer 2015).

Six2, an isoform of the SIX1 family most similar to Six1, has also been shown to maintain pools of progenitor cells, though not much has been shown in the cancer-forming capabilities of Six2 (Combes *et al*, Kidney Int 2017). In the developing mouse metanephric mesenchyme Six2 is expressed in a small subset of cells though expression is lost as mice progress into adulthood (Kumar, Cell Mol Life Sci 2009). In Six2 null mice it was found that the developing kidneys form ectopic renal vesicles and nephrogenesis terminates after development of only a few nephrons, leading to hypoplastic kidneys similar to Six1 knockout mice (Combes *et al*, Kidney Int 2017).

The degree of conservation of Six proteins between lower invertebrates and mammals throughout evolution signifies the importance of these proteins in normal system development. Indeed, mutations of Six family members has been implicated in developmental diseases such as sensory, auditory, and craniofacial disorders (Kumar, Cell Mol Life Sci 2009).

1.6 <u>Structure of SIX</u>

Of the SIX family of proteins SIX1 is the most extensively studied in the human as it is vital to the development of multiple organ systems (Wu *et al*, Int J Cancer 2015). SIX1, like all SIX family members, contains two highly conserved functional domains (Figure 3). The most N-terminal six domain (SD) shares >95% homology between family pairs (i.e. SIX1/2, SIX4/5, SIX3/6) and is responsible for protein-protein interactions such as with its multifunctional cofactor, EYA (Preethi *et al*, Frontiers in Biology 2015). Mutations in these domains have been discovered in patients with branchio-oto-renal (BOR) syndrome, a developmental disorder characterized by renal abnormalities as well as hearing loss (Ruf *et al*, PNAS 2004). Patrick *et al*. showed that the SD mutation V17E was sufficient to disrupt binding of SIX to EYA thus failing to localize EYA to the nucleus, though SIX still retained DNA-binding capabilities (Patrick *et al*, J Biol Chem 2009). The SIX1 DNA-binding homeodomain (HD) is primarily responsible for DNA motif recognition, which has been shown to be responsible for cranial structure development in vertebrates (Christophorou *et al*, Developmental Biology 2009). The canonical sequence bound by SIX1 is the MEF3-like sequence (TCAGGTTTC). Originally this sequence was found in the myogenin (MyoG) promoter region as a target of SIX1 binding during embryogenesis, however recent ChIP-on-chip analysis of SIX1 suggests that there is a certain level of tolerance for variability in the MEF3 sequence (Liu *et al*, Nuc Acis Res 2012; Parmacek *et al*, Mol Cell Biol 1994). This leaves open the possibility that there could be a range of undiscovered genes regulated by SIX1, elucidation of which may be critical to understanding how SIX1 functions in the cellular environment.

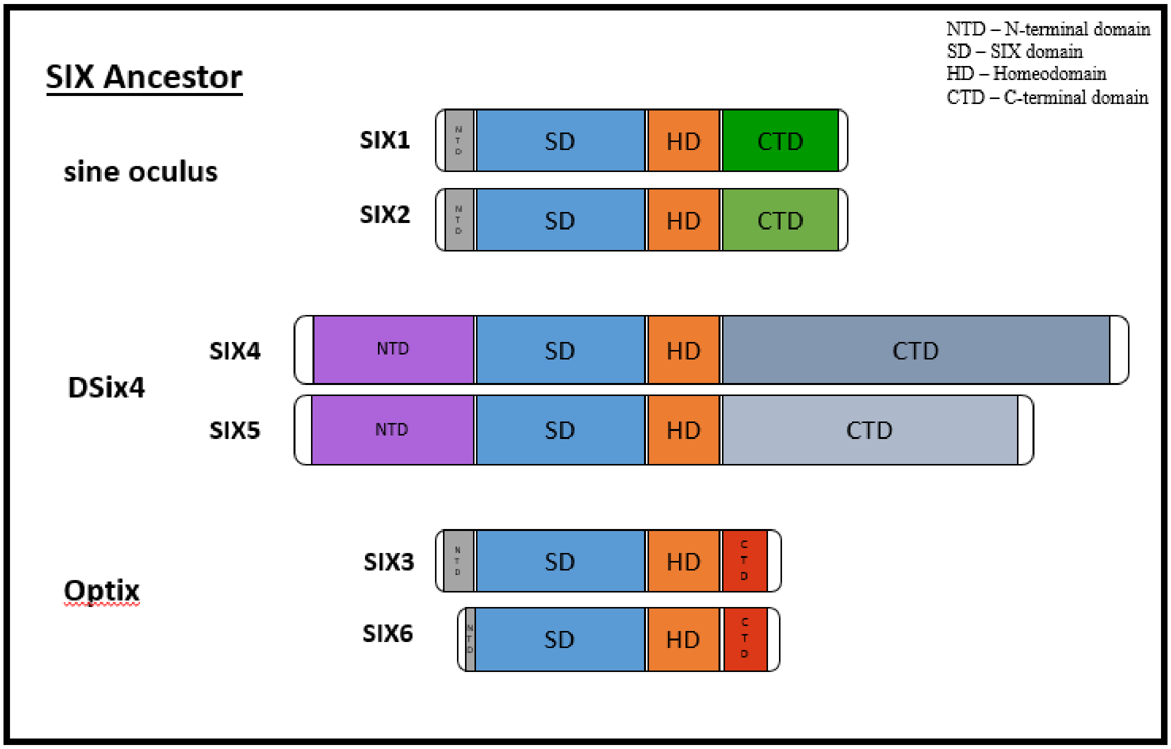

Figure 1.3. Structural domains of the SIX family protein members

Six family members consist of a highly conserved protein-interaction capable six domain (SD) and a DNA binding homeodomain (HD). Homology between Six proteins varies depending on their evolutionary ancestor, N-terminal domain (NTD) structure, and C-terminal domain (CTD) structures.

1.7 <u>Roles in SIX Cancer</u>

Homeobox genes play key roles in the development of certain cancers where they may be either upregulated or downregulated. For example, HOXA gene expression is increased in acute myeloid leukemias (AML) that harbor NPM1c mutations, which account for approximately 35% of all AML patients (Li *et al*, Cancers Basel 2019). NPM1c mutations delete the N-terminal nucleolar localization signal and produces a novel export signal, resulting in aberrant cytoplasmic expression and a subsequent increase of several HOXA genes (Vassiliou *et al*, Nat Genet 2011). This increase of HOXA genes is due to epigenetic regulation at H3K27 on HOX gene super-enhancers and results in increased leukemic burden (Brunetti *et al*, Cancer Cell 2018).

In the context of the SIX family of proteins, SIX1 has been most extensively studied and characterized as a potential driver of cancer development (Table 1). For example, SIX1 is overexpressed in approximately 25% of cervical cancers (Zheng *et al*, Int J Gynecol cancer 2010). Though new studies suggest that other members, such as SIX3, might potentially provide protective benefits in certain breast cancer patients (Xu *et al*, Medicine 2016). SIX1 has been shown to be involved the development of various human cancers including rhabdomyosarcoma, cervical cancer, colon cancer, breast cancer, and certain types of AMLs (Kumar J Cell Mol Life Sci 2009; Zheng *et al*, Int J Gynecol 2010; Xu *et al*, Medicine 2016; Wang *et al*, Blood 2011). Unsurprisingly SIX2 has been implicated in similar types of cancer as SIX1, though this may be a compensatory mechanism given the high degree of homology that exists between SIX1 and SIX2.

SIX1, like many homeobox genes, is expressed during embryogenesis and not during adulthood. Reactivation of such developmental protein expression in mature tissues has been shown to be involved in cancer development. The embryonic protein receptor tyrosine kinase-like orphan receptor 2 (ROR2) is a WNT5A receptor that regulates cellular polarity and is required for cartilage and growth plate development (Roarty *et al*, Oncogene 2017). When aberrantly expressed in adulthood ROR2 signals through WNT/c-Jun N-terminal kinase (JNK) pathways to aid in ovarian carcinogenesis. Similarly, adult expression of the developmental DNA-binding and cell differentiation Id-family genes that encode transcription factors has been linked to multiple cancer types including cervical, thyroid, breast, and pancreatic cancers (Roschger and Cabrele, Cell Comm Signal 2017). In the same manner breast cancer patients exhibit upregulation of SIX1 and have an increased risk of metastatic disease (Xu *et al*, Medicine 2016).

Expression of SIX proteins is limited to early embryogenesis and is restricted in adult tissues making it an interesting treatment candidate for cancers expressing SIX1 protein. Currently there are no clinical therapies targeting SIX1, though drugs are being developed to inhibit the catalytic function of the SIX1-interacting phosphatase EYA1 (Blevins *et al*, Expert Opin Ther Targets 2015). Small molecule targeted therapies would ideally disrupt protein-protein interactions between SIX1 and known interactors such as EYA by targeting the SD. Importantly the SD specifically present in the SIX family of proteins has little to no known homology to other proteins, including homeobox-containing transcription factors. New more sensitive methods have been developed that have overcome many hurdles to developing small molecules capable of inhibiting protein interactions, currently there are small molecules in clinical trials targeting these types of interactions (Nero *et al*, Nat Rev Cancer 2014). A promising example is the use of ABT-

263 in multiple types of cancer to target the BH3 domain in BAX, which disrupts its interaction with the antiapoptotic BCL-XL protein (Tse *et al*, Exp Therapeutics 2008).

If a treatment were developed targeting SIX1 given the recent discoveries of its function in humans, it may impact developing cells going against what is commonly thought in the literature. My data suggests that lack of a functioning SIX1 or SIX2 will have a potentially detrimental erythropoietic effect. Despite the initial belief that SIX1 and SIX2 proteins are lost in adult tissues, recent studies in macrophages suggest that SIX1 and SIX2 suppress immune responses through interaction with NFκB, perhaps leaving people more susceptible to bacterial and viral infection (Liu *et al*, Nature 2019). However, inhibition of SIX1 via small molecule inhibitor designed to inhibit the SD domain would likely be well tolerated in normal tissues due to the lack of SIX1 protein, implying that even long-term tumor maintenance would theoretically be well-tolerated.

SIX1 has only recently been a gene of interest in hematologic malignancies including Hodgkins lymphomas and AMLs that harbor a genetic rearrange of the mixed-lineage leukemia (MLLr) (Nagel *et al*, Oncotarget 2015. Chu *et al*, Cancer Sci 2019). In a study of 23 MLLr AML patients compared to 56 MLL wild-type MLL, SIX1 was preferentially upregulated in the MLLr cohort (Wang *et al*, Blood 2011). Further studies in mice show that Wnt signaling was a target of the MLL-AF9 oncogene (a subset of MLL-rearrangements), and that signaling by Wnt altered the accessibility of Six1 DNA to the transcription factor Tcfl2 (Zhang *et al*, EBioMedicine, 2019). Attenuating either Wnt signaling pathways by small molecule inhibition or reduction of Six1 gene expression levels by conditional knockouts reduced the leukemic burden and increased survival in these mouse models. Approximately 12% of Hodgkin's lymphoma patients had aberrant expression of SIX1, suggesting an additional role in pathogenesis (Nagel *et al*, Oncotarget 2015). In cell culture models of Hodgkin's lymphoma knockdown of SIX1 led to altered expression of both GATA2 and GATA3, suggesting that perhaps SIX1 regulation of GATA transcription factors is a conserved mechanism in hematopoiesis. This switch between GATA2 and GATA3 is another example of a developmentally important switch that occurs between GATA-family members (similar to GATA2 and GATA1 in erythropoiesis). In our erythroid cell cultures we see that GATA switching occurs in response to forced SIX1 and SIX2 overexpression perhaps highlighting a conserved link between SIX and GATA protein families.

Tissue	Cancer	Analysis	Result from Analysis	Source
Bone	Osteosarcoma	Stage	Higher in Enneking Stage	Chao et al. 2015
	Osteosarcoma	Tumor Size	Higher in Tumors >5cm	Chao et al. 2015
	Osteosarcoma	Patient Prognosis	Higher in Shorter Overall Survival	Chao et al. 2015
Breast	Breast Carcinoma	Lymph Node Involvement	Higher in Lymph Node Positive Samples	Xu et al, 2016
	Breast Carcinoma	Five-year Survival	Higher in Dead vs Alive	Micalizzi et al. 2009
Cervix	Cervical	Differentiation Status	Highest in Poorly Differentiated	Micalizzi et al. 2009
Colon	Colon Carcinoma	Metastasis	Higher in Metastatic Cancer	Micalizzi et al. 2009
Kidney	Renal Cell Carcinoma	Stage	Higher in Stage IV vs Stage I	Micalizzi et al. 2009
Pancreatic	Pancreatic Ductal Adenocarcinoma	Expression in Tumor Sample	Higher in Tumors vs Normal Tissue	Lerbs et al. 2017
Prostate	Prostate Cancer N Stage	Lymph Node Involvement	Higher in Lymph Node Positive Samples	Micalizzi et al. 2009
	Prostate Cancer	Metastasis	Higher in Metastatic Cancer	Micalizzi et al. 2009

Table 1.1 Analysis of SIX1 expression in multiple human cancers.
SIX1 expression has been analyzed in primary patient samples from multiple cancer types. Various analyses have confirmed that in primary patient data SIX1 expression correlates with metastatic disease, lymph node involvement, tumor grade, size, and overall poor reduced survival. *Adapted from Micalizzi et al, Cancer Res 2010.*

1.8 **Goals and Hypothesis**

The PSEDN is responsible for mammalian development of multiple organ systems including eyes, inner ear, lung, skeletal muscles, and kidneys. Mutations in PSEDN members EYA and SIX have been linked to Branchio-oto-renal (BOR) syndrome, characterized by branchial fistulae, hypoplastic kidneys, and mild to severe hearing loss (Smith, GeneReviews 1999). Recent studies have highlighted the role of the PSEDN in neoplastic and metastatic carcinomas (reviewed in Wu *et al*, Int J Cancer 2015). Due to its lack of adult tissue expression, SIX1 has become a protein of interest as a potential druggable target where its interaction with cofactors such as EYA1 could reduce metastasis of primary cancers. However, the role of the PSEDN is incompletely understood in adult tissues. Given the central role of GATA/FOG network in human hematopoiesis and erythropoiesis we tested the role of PSEDN members SIX1, SIX2, and EYA1 in both primary hematopoietic stem and progenitor cells (HSPCs) as well as in established erythroleukemic cell lines used to model aspects of human erythropoiesis *in vitro*.

The **purposes** of this study were to 1) examine if PSEDN members are sufficient or necessary to regulate human hematopoiesis in HSPCs and TF1 cell cultures, then 2) determine if PSEDN members functionally regulate GATA/FOG members as discovered in the fly, and 3) identify novel interactions of SIX using proximity ligation assays to determine other crucial cofactors of the PSEDN to order to identify potential mechanisms and protein associations.

The **hypotheses** tested are 1) *PSEDN members SIX1, SIX2, and EYA1, regulate human erythropoiesis through coordinated interactions with GATA/FOG members, and* 2) *SIX* *function is dependent on GATA1.* Additionally, determining the SIX interactome will identify novel interactors of PSEDN and provide insights into its function in erythropoiesis.

The overall goals of this study are to test if the PSEDN and GATA/FOG interactions are conserved across species. Linking these two evolutionarily conserved networks suggests fundamental developmental processes that allow crosstalk between two central pathways. Our data – as well as data from Drosophila – suggests regulation of GATA/FOG by PSEDN at multiple levels including its expression levels and its activity.

Therefore, the functional relevance of PSEDN and GATA/FOG described in this thesis will provide the framework for future studies beyond erythroid development and may provide novel druggable targets in human diseases.

<u>Chapter 2: SIX stimulates human erythroid phenotypes *in vitro*</u>

2.1 <u>Introduction</u>

2.1.1 <u>Role of transcription factors in hematopoiesis, a case for PSEDN</u>

Although the mechanisms regulating erythropoiesis have been studied for decades, molecular details required to support efforts to develop methods for efficient production of transfusable RBCs remain an important topic of hematopoietic research. The differentiation of hematopoietic stem and progenitor cells (HSPCs) into mature erythrocytes is orchestrated by the coordinated action of transcriptional networks that control cellular proliferation and differentiation. Bipotent progenitor cells, known as megakaryocyte erythroid progenitors (MEPs), capable of populating both megakaryocytic and erythroid lineages, dynamically express transcription factors that control fate specification. Several factors have been discovered that play a role in MEP fate decisions including the erythroblast transformation specific (ETS) transcription factor family members FLI-1 and GABPα, as well as the GATA family factors GATA1 and GATA2 (Dore and Crispino, Blood 2011). Once either a megakaryocyte or erythroid lineage is specified its transcriptional factor and microRNA networks activate appropriate and inhibit competing linages to drive differentiation (Reviewed in Kim *et al*, Wiley Int Rev 2019). For example, KLF1 promotes erythroid-fate specification and is highly expressed in MEPs and erythroid cells. However, KLF1 is downregulated during megakaryopoiesis by the megakaryocyte-specific gene, RUNX1, and increased by GATA1 during erythropoiesis, thus highlighting the importance of balance between transcription factor levels during megakaryocyte/erythroid decisions (Kuvardina *et al*, Blood 2015; Tallack *et al*, Genome Res 2012).

In vivo, HSPCs are signaled by extracellular cytokines including stem cell factor (SCF), IL-3, thrombopoietin (TPO), and especially erythropoietin (EPO) to drive the daily generation of billions of red blood cells (RBC) (Kaushansky, NEJM 2006). Approximately 90% of EPO in the blood is produced in the juxtatubular interstitial cells of the renal cortex in response to low oxygen levels. The production of EPO is controlled by an oxygen-dependent prolyl hydroxylase (PHD) that is responsible for the stability of the transcription factor hypoxia-inducible factor 2α (HIF-2α). Low oxygen-levels disable the protein-protein interaction between HIF-2α and E3 ubiquitin ligase von Hippel-Lindau (VHL) thus stabilizing HIF-2α and leading to EPO production to signal the generation of additional RBCs. Once released into the bloodstream, EPO promotes the survival and proliferation of erythroid committed progenitors through upregulation of antiapoptotic proteins such as BCL-XL and by reducing level of cell-cycle inhibitors (reviewed in Krantz, Blood 1991).

Several transcription factors have been established as key controllers of erythroid development in mammalian hematopoiesis, including the founder member of the GATA family, GATA1. GATA1 expression is highly regulated in hematopoiesis, being largely restricted to erythroid, eosinophils, and megakaryocyte lineages, where GATA1 interactions with other proteins control cell-fate specifications (Amigo *et al*, Blood 2009). For example, interaction of GATA1 with Friend-of-GATA1 (FOG1) functions both to activate and repress certain factors in erythropoiesis as well as megakaryopoiesis (Cantor and Orkin, Sem in Cell & Dev Bio 2005). Mutations that disrupt this GATA1/FOG1 interaction cause severe anemia in mice and X-linked thrombocytopenia in humans (Nichols *et al*, Nat Gen 2000; Shimizu *et al*, Blood 2004).

During mammalian development PAX-SIX-EYA-DACH network (PSEDN) members such as SIX1 are involved in regulation of survival, proliferation, and differentiation in a wide array of progenitor cells in various tissues. However, PSEDN members are often aberrantly expressed in solid cancers where their re-expression increases cell growth and drives epithelial-mesenchymal transitions (EMT) phenotypes (Reviewed in Kingsbury *et al*, 2019). SIX1 and EYA1 have also been shown to be upregulated in leukemias. *In silico* analysis of leukemia patients harboring mixed-linage leukemia rearrangements (MLLr) show preferential upregulation of PSEDN members SIX1, EYA1, and DACH1 when compared to MLL-wild type (MLL-WT) patients (Wang *et al*, Blood 2011). Patients who harbored both MLLr as well as high-SIX1 expression had a worse prognosis than patients with low-SIX1 expression. In mouse MLLr, specifically MLL-AF9 rearrangements, knockdown (KD) of SIX1 prolonged survival and decreased tumor burden (Chu *et al*, Cancer Sci 2019). Recent studies of mice leukemias driven by MLL-AF9 rearrangements also showed that SIX1 expression was driven by WNT signaling pathways (Zhang *et al*, EBioMedicine 2019). Increased WNT signaling and subsequent CTTNB1 as well as TCF expression increased chromatin accessibility at the SIX locus, leading to increased SIX1 expression. SIX1 conditional knockout mice in their study showed a decrease (nearing statistical significance) in the megakaryocyte-erythroid progenitor (MEP) population while their other blood counts remained normal. This suggests that SIX1 may have a role in maintenance or development of MEPs in mammalian hematopoietic cells. However, the role of the PSEDN in hematopoiesis remains largely unknown.

Our collaborator, Dr. Nancy Fossett (University of Maryland, Baltimore) showed that the fly homologue of EYA1 (*eya*) regulated hematopoietic progenitors in coordination with the FOG1-homologue, *u-shaped* (*ush*) in Drosophila. Additionally, she showed that PSEDN members regulated expression of the fly homologue of GATA1 (*serpent, srp).* Dr. Fossett's data that PSEDN members had genetic interactions with GATA/FOG complexes led to my hypothesis that this cross-regulation between two central developmental complexes would be conserved from flies to humans.

We made the novel discovery that human PSEDN members, SIX1, SIX2, and EYA1 can regulate and drive erythropoiesis in human cell lines. Conversely, CRISPR/Cas9-mediated knockout of SIX1 and SIX2 in primary HSPCs as well as erythroleukemic cell lines reduced the ability to generate erythroid markers *in vitro* in response to exogenous EPO. Additionally, forced overexpression of SIX1, SIX2, or EYA1, promoted the expression of erythroid markers hemoglobin, GATA1, ALAS2, SLC4a1, and EPOR, in the absence of added EPO. We also made the novel discovery that GATA1 is a necessary cofactor for SIX1 to drive erythropoiesis as genetic knockout of GATA1 blocks the ability of SIX1 to drive erythroid phenotypes. Taken together we show that the interaction between the PSEDN and GATA/FOG networks is conserved across species. We discovered that in human erythropoiesis the PSEDN, specifically SIX1, SIX2, and EYA1, regulate erythroid cell production in conjunction with the master erythroid regulator GATA1.

2.1.2 GATA/Fog and PSEDN regulate *Drosophila* hematopoiesis

Drosophila melanogaster provides a powerful model system to investigate protein function *in vivo*, enabling the identification of conserved co-regulators derived from ancient core hematopoietic networks (Fossett, Biochim Biophy Acta 2013). The comparatively simple Drosophila blood system shares important commonalities with vertebrate hematopoietic systems, including distinct spatial and temporal regulation and multipotent progenitors (prohemocytes) that give rise to all three mature blood cell types. These blood lineages (plasmatocytes, crystal cells and lamellocytes) carry out functions associated with the vertebrate myeloid lineages (Dearolf, Biochim Biophy Acta 1998; Lanot *et al*, Dev Biol 2001). Plasmatocytes are operational macrophages; crystal cells are involved in wound healing; and lamellocytes, which are rarely observed under steady-state conditions, differentiate to combat wasp parasitization and in response to other forms of stress including increased levels of reactive oxygen species (ROS) and nutritional deprivation (Frandsen *et al*, Natl Acad Sci 2008; Gold *et al*, Semin Immunol 2015; Owusu-Ansah *et al*, Nature 2009; Rizki *et al*, Dev Comp Immunol 1992; Vlisidou *et al*, FEBS J 2015).

Prior to our work in human erythropoiesis our collaborator, Dr. Fossett provided the first evidence of coordinated hematopoietic regulation by both the PSEDN and GATA/FOG network. The Drosophila hematopoietic GATA/FOG network exerts functions during fly embryogenesis analogous to all three vertebrate hematopoietic GATAs ranging from prohemocyte specification to differentiation of blood lineages (Mandal *et al*, Nat Genet 2004). Like mammalian GATAs, *serpent* (*srp*) interacts with a variety of conserved hematopoietic transcriptional regulators (*ush*/FOG, Runx, NuRD complex) and regulatory pathways (BMP, Hh, JAK/STAT) that direct specific hematopoietic *srp*/GATA activities including cell fate and differentiation (Fosset, Biochim Biophy Acta 2013; Frandsen *et al*, Nat Acad Sci 2008; Gao *et al*, Mol Cell Biol 2009). Utilizing a second site non-complementation (SSNC) screen to identify novel conserved *in vivo* interactors GATA network members based on the role of GATA in lamellocyte differentiation (Baldeosingh *et al*, Dev Biol 2018), Dr. Fossett linked for the first time members of the PSEDN and GATA/FOG networks (Table 1). SSNC refers to the occurrence of a mutant phenotype in double heterozygotes harboring heterozygous recessive mutations in distinct genes. SSNC occurs when the functional output of cellular pathways drops below the threshold required to maintain the wild type phenotype, thereby enabling SSNC screens to identify proximal or distal upstream regulators, co-regulators, downstream effectors and parallel pathways (Hawley *et al*, Genetics 2006). The SSNC screen was designed to identify genetic interacting factors along with GATA/FOG complexes. This screen utilized heterozygous mutants of *ush* (FOG homologue) as *srp* (GATA homologue) mutants were not viable. By crossing heterozygous mutants of *ush* with a library of mutants, including members of the PSEDN, she found that a cross with heterozygous mutants of PSEDN-member *eya* resulted in increased lamellocyte production (Figure 1). This suggests that the double mutants impeded the ability of *srp* to block lamellocyte production and maintain progenitor cell populations.

Beyond demonstrating that heterozygous mutants of PSEDN and GATA/FOG members were instrumental to lamellocyte production, her lab also demonstrated that loss- and gain-of-function of *sine oculus (so,* SIX1 homologue*)* either reduced or increased protein expression levels of *serpent (srp*, GATA homologue) (Figure 2A-D). This shows that *srp* levels in the developing fly lymph glands are dependent upon *so*. Her lab has previously demonstrated that *srp* overexpression in prohemocytes reduces prohemocytes and increased lamellocytes (Gao *et al*, Development 2013). Consistent with increased levels of *srp*/GATA observed in *so* overexpression, she observed a significant reduction of *odd*-positive progenitor cells (Figure 2E-G) and increased numbers of differentiated lamellocytes (Figre 2H-J) in so^{GOF} lymph glands. Thus, gain of *so* function increases lamellocytes and reduces prohemocytes by likely by increasing the level of *srp* expression. This data provides the framework for the major findings of this thesis by establishing the first link between two well-conserved and crucial developmental networks. While this link has been demonstrated in Drosophila, it is still unknown whether there is regulation between these two complexes in human development.

PSEDN member	*Drosophila* allele	ush/PSEDN member (trans-heterozygotes) # positives*/ # scored	% positive	PSEDN member/+ (PSEDN single heterozygotes) # positives*/ # scored	% positive	P value***
NA	**Control**	2/30	6.7	1/67	1.6	0.2252
EYA	***eyaCli-IID***	16/29	55	0/22		>0.0001
SIX	***so3*** **/+**	18/20	90	7/61	11.5	>0.0001
	so2 **/+**	7/34	21	2/51	4	>0.02
	so1 ******	11/29	38	1/45	2.2	>0.0001
PAX	***ey2*****	16/16	100	11/47	23.4	>0.0001
	ey1**	37/38	97			NA
DACH	***dac1*** **/+**	11/21	52	8/70	11.4	>0.0003
	dac9 **/+**	15/21	71	6/82	7.3	>0.0001
	dace462/+	13/20	65	7/73	5.5	>0.0001

Table 2.1. PAX-SIX-EYA-DACH transcriptional network members exhibit second-site noncomplentation with *ush* (FOG)

Members of the fly PSEDN complexes were crossed with either control heterozygotes (+) or with *ush* mutants. Lymph gland lobes were stained with the lamellocyte marker, attilla, and scored for number of lobes positive as having aberrant lamellocyte differentiation.

*Estimated > 10x increase in lms #;
**Homozygotes
***Fisher's exact test
****Tests if *ush* heterozygotes produce significant increase compared to control

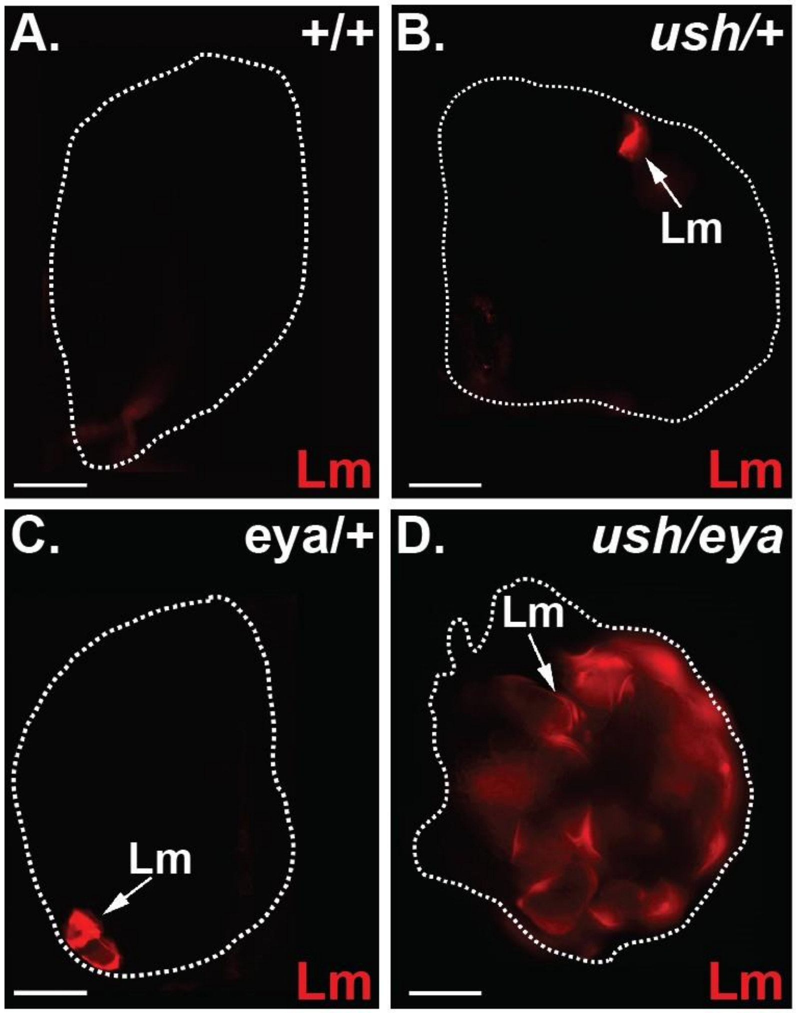

Figure 2.1. PSEDN member *eya* exhibits SSNC with *ush* in *Drosophila*

Lamellocytes were detected in late third larval instar using the *MSN-C* reporter (in red) and are marked with arrows. White dotted lines delineate the entire lymph gland. Scale bars: 50 µm. (A-C) Very few lamellocytes (Lm) were observed in wild-type (+), *ush/+* or *eya/+* single heterozygotes. (D) In contrast, *ush/eya* double heterozygotes showed increased lm numbers. (Adapted from Creed *et al*, Development 2020)

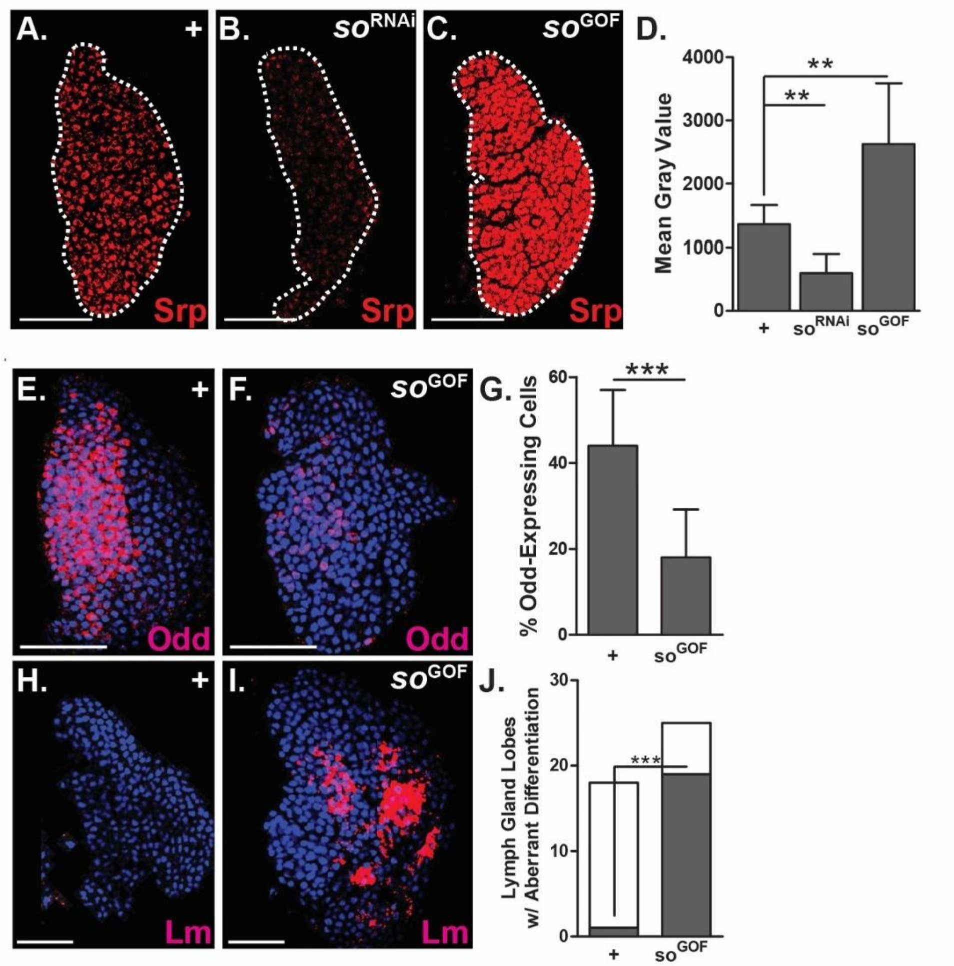

Figure 2.2. ***So*** **proteins regulate fly** ***srp*** **expression levels and control hematopoiesis**

A) control (+), (B) *so*^{RNAi}, (C) *so*^{GOF}. (D) Histogram showing that *so*^{RNAi} significantly reduced *srp* levels by >50% (n=10; ANOVA, **P<0.01), whereas *so*^{GOF} significantly increased *srp* levels approximately twofold (n=10; ANOVA, **P<0.01). Data are mean gray value/lobe for each genotype±s.d.(E-J) The tissue-restricted *Tep*- Gal4 driver was used to overexpress wild-type *so* (*so*^{GOF}) in prohemocytes. Lymph glands were dissected from late third instar larvae derived from Tep-Gal4 driver flies crossed to wild type (+) or *so*^{GOF} and lymph glands were counterstained with DAPI (blue). (E-G) The percentage of *odd*-positive cells was significantly reduced in *so*^{GOF} compared with controls. (G) Histogram showing quantitation of percentage of *odd*-positive cells (n=10). Data are mean±s.d. Student's t-test; ***P<0001. (H-J) Lamellocytes were identified using the specific marker L1. (J) Histogram showing lamellocytes significantly increased in *so*^{GOF} compared with controls. Gray region indicates number of lymph gland lobes with aberrant lamellocyte differentiation and white indicates number of lobes with the wild- type phenotype. *Te*p-Gal4 control (n=18); *Tep* driving UAS-*so*^{GOF} (n=25). Fisher's exact test; ***P<0001. Scale bars: 50 μm. n=number of lymph gland lobes.

2.2 Results

2.2.1 SIX proteins control GATA levels across species

To test whether PSEDN family members controlled GATA-levels across species, we overexpressed PSEDN-members SIX1 and SIX2 in human TF1 erythroleukemia cell lines. Cells were transduced to an MOI ~2 (90% GFP^{+}) using lentivirus and cultured in GM-CSF-containing growth media (RPMI 10% FBS). Transduction efficiencies were confirmed using flow cytometry. Forced overexpression of SIX1 or SIX2 resulted in an ~50% increase in GATA1 and ≥50% decrease in GATA2 mRNA levels six days post-transduction as measured by RT-qPCR (Figure 3A). Likewise, SIX1 or SIX2 overexpression reduced GATA2 protein levels by >80% and elevated GATA1 protein levels ~50% and 70%, using Western blot and normalizing to internal actin controls (Figure 3B-D). The observed changes in GATA expression levels were similar to those observed in TF1 cells following EPO stimulation (Figure 4). Thus, in the absence of exogenous EPO, enforced expression of SIX1 and SIX2 triggers changes in GATA1 and GATA2 expression consistent with the GATA switch and confirms that regulation of GATA expression levels by SIX proteins is conserved across species.

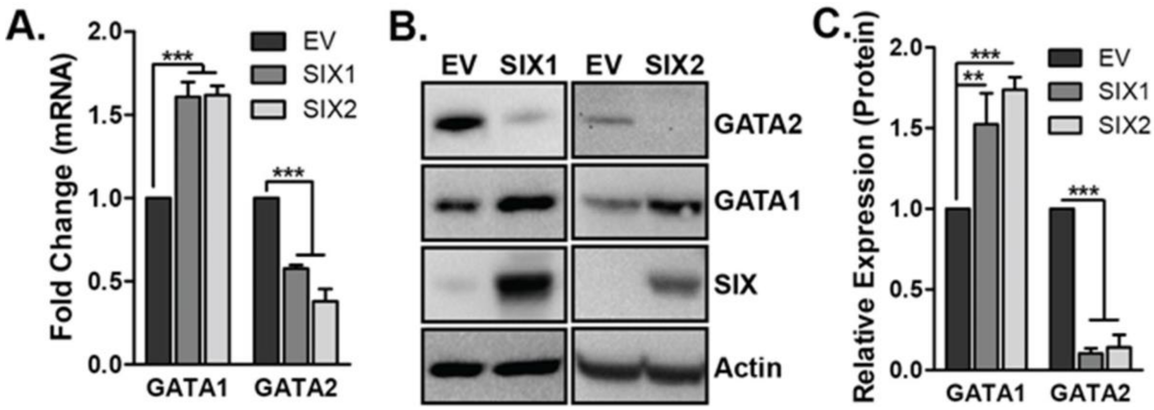

Figure 2.3. SIX proteins regulate GATA expression levels

(A) Analysis of GATA1 and GATA2 RNA levels following SIX overexpression. Total RNA was harvested from SIX1 and from SIX2 versus EV-transduced TF1 cells and analyzed by qRT-PCR. GATA levels were normalized to actin and fold change calculated relative to EV control samples. Data are mean+/-s.e.m., n=3, all values were significantly different compared with EV. ***$P<0.005$. (B) Representative blots from SIX1- or SIX2- transduced TF1 cells after 6 days in standard growth media culture post-transduction. Protein lysates were harvested and analyzed by western blot for GATA2, GATA1, SIX1, SIX2 and Actin. (C) Data are mean fold change of densitometric mean above EV after normalization to Actin+/-s.e.m. (**$P<0.01$, ***$P<0.005$).

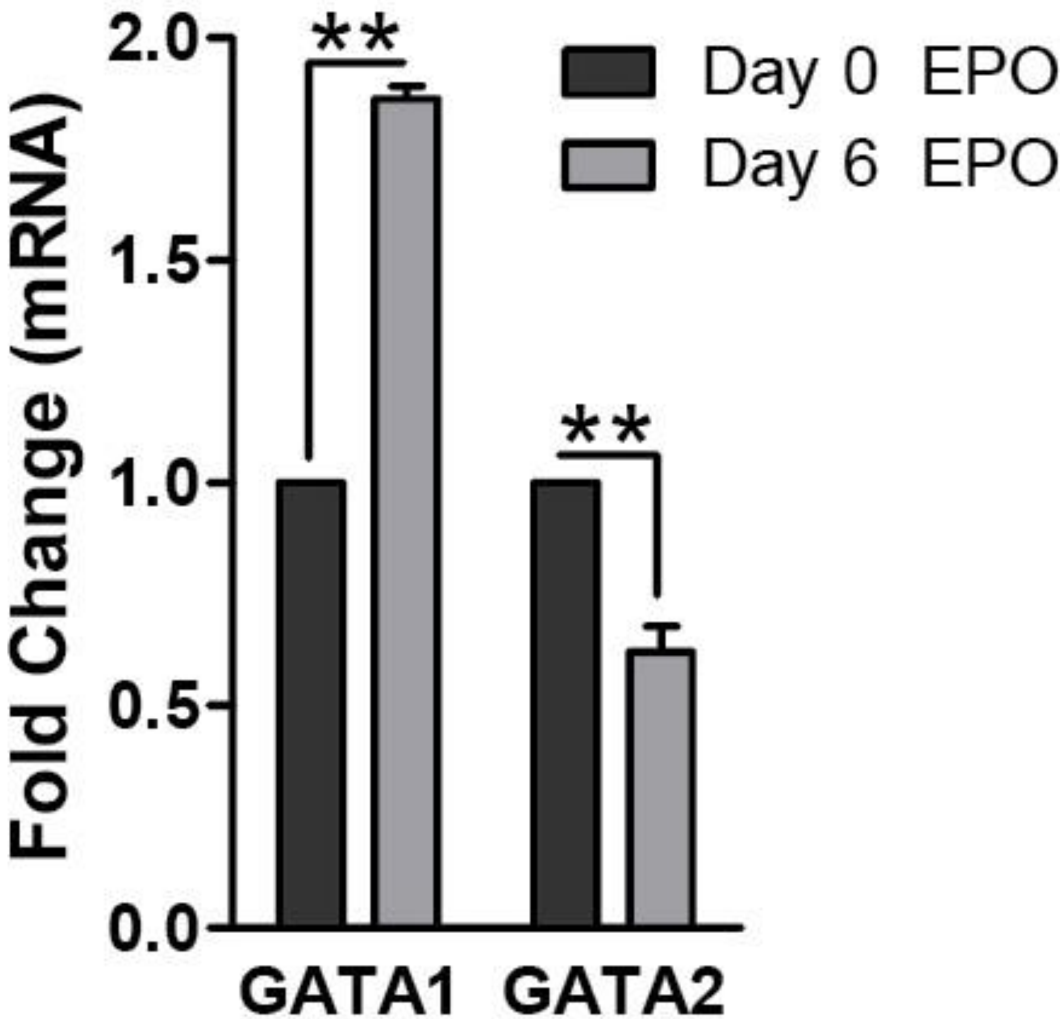

Figure 2.4. GATA expression levels following EPO-stimulated erythropoiesis

TF1 cell cultures were depleted from GM-CSF overnight and incubated six days in the presence of EPO. After EPO stimulation cells were washed and RNA was harvested for RT-qPCR. Bar graphs represent three independent experiments. ** $p < 0.01$ +/- s.e.m.

2.2.2 SIX1 overexpression phenocopies GATA1 overexpression in human erythroleukemia cells

To determine if cross-talk between PSEDN members and the GATA/FOG network is conserved across taxa, we tested whether human SIX proteins could influence GATA1/FOG1 function in TF1 erythroleukemia cells, which can recapitulate early GATA1-dependent erythroid differentiation upon stimulation with erythropoietin (EPO) (Chretien *et al*, EMBO J 1996; Kitamura *et al*, J Cell Physiol 1989). TF1 cell lines model early HSPCs *in vitro* as they express high levels of the HSPC-marker CD34, which is downregulated during differentiation mirroring the loss of CD34 that accompanies *in vivo* HSPC differentiation (Alexaki *et al*, MOJ Immunol 2014). Our lab has established *in vitro* assays sufficient to measure erythropoiesis in response to exogenous EPO stimulation. We utilize both primary HSPCs and the erythroleukemia cell line, TF1, which recapitulate erythropoiesis in cell cultures (Kim *et al*, Brit J Hemat 2014). TF1 cell culture proliferate *in vitro* with GM-CSF incubation, however, depletion of GM-CSF and addition of EPO causes erythroid development that can be measured *in vitro* using RT-qPCR, Western blot, and flow cytometry (Figure 5).

Overexpression of GATA1 in erythroid progenitors produces a slow growth phenotype and promotes erythroid differentiation (Briegel *et al*, Development 1996). We similarly observed that lentivirus-mediated GATA1 expression, which are marked with GFP allowing for measurement of transduction efficiencies, reduced TF1 cell growth and triggered erythroid differentiation in the absence of exogenous EPO (Figure 6A-D). Of the six human SIX genes, SIX1 and SIX2 are most closely related to each other and to *Drosophila* So (Kumar, Cell Mol Life Sci 2009). TF1 cells transduced with either

GATA1, SIX1 or SIX2 expressing GFP labeled lentivirus were monitored in long-term growth competition assays. As observed for GATA1, TF1 cells expressing SIX1 or SIX2 were rapidly outcompeted in partially-transduced cell populations (~50% GFP^{+}), consistent with reduced growth. In contrast, control GFP-expressing cells were maintained at a constant frequency (Figure 6A) during prolonged culture.

We next tested whether SIX1 or SIX2 overexpression reduced growth was associated with erythroid differentiation in the absence of EPO as we observed in cells overexpressing GATA1. Six days post transduction GATA1-, SIX1- or SIX2-expressing cell populations each exhibited an increased percentage of the erythroid cells as assessed by transferrin receptor (CD71) and glycophorin A (CD235a) cell surface expression ($CD71^{hi}CD235a^{hi}$) (Gautier *et al*, Cell Rep 2016; Loken *et al*, Blood 1987; van Lochem *et al*, Cytom B Clin Ctyom 2004), a decreased percentage of cells expressing the hematopoietic stem-progenitor marker CD34 (Figure 6B-C), increased mRNA levels of key erythroid genes including HBB, ALAS2, SLC4A1, ANK1, KLF1, and SPTA1 (Figure 6D) and elevated beta-hemoglobin protein levels (Figure 6E). Thus, like GATA1, enforced expression of SIX1 or SIX2 produced phenotypes indicative of erythroid differentiation. Conversely, CRISPR engineered SIX1 knockout TF1 cells exhibited impaired erythroid differentiation in response to EPO stimulation, compared to controls (Figure 6F-H), suggesting that SIX1 contributes to EPO-driven TF1 cell differentiation. No compensatory changes in the RNA levels of remaining SIX family members (SIX2-6) were observed in SIX1 knockout cell lines (unpublished observation).

We further tested whether EYA1, a human homolog of Drosophila *eya*, could stimulate TF1 erythropoiesis. EYA1 overexpression resulted in increased numbers of

$CD71^{hi}CD235a^{hi}$ cells, reduced numbers of $CD34^{+}$ cells (Figure 7A-B), and increased hemoglobin protein levels (Figures 7C, lane 1 vs 3) 10 days post-transduction in culture without EPO. Since EYA proteins function as SIX co-regulators, we further tested whether SIX1 was necessary for EYA1 to stimulate TF1 erythropoiesis. Inducible CRISPRi-mediated knockdown of SIX1, which yielded >80% reduction in SIX1 protein levels, reduced EYA1-mediated induction of hemoglobin (Figure 7C, lane 3 vs 4). Consistent with previous observations that EYA stabilizes SIX1 protein levels in the MCF7 breast cancer cell line (Patrick *et al*, J Biol Chem 2009), we routinely observed slightly elevated levels of SIX1 in cells overexpressing EYA1 (Figure 7C, lane 1 vs 3). Collectively, these data suggest EYA1 stimulates erythropoiesis and that EYA1 function is mediated in part via endogenous SIX1 protein.

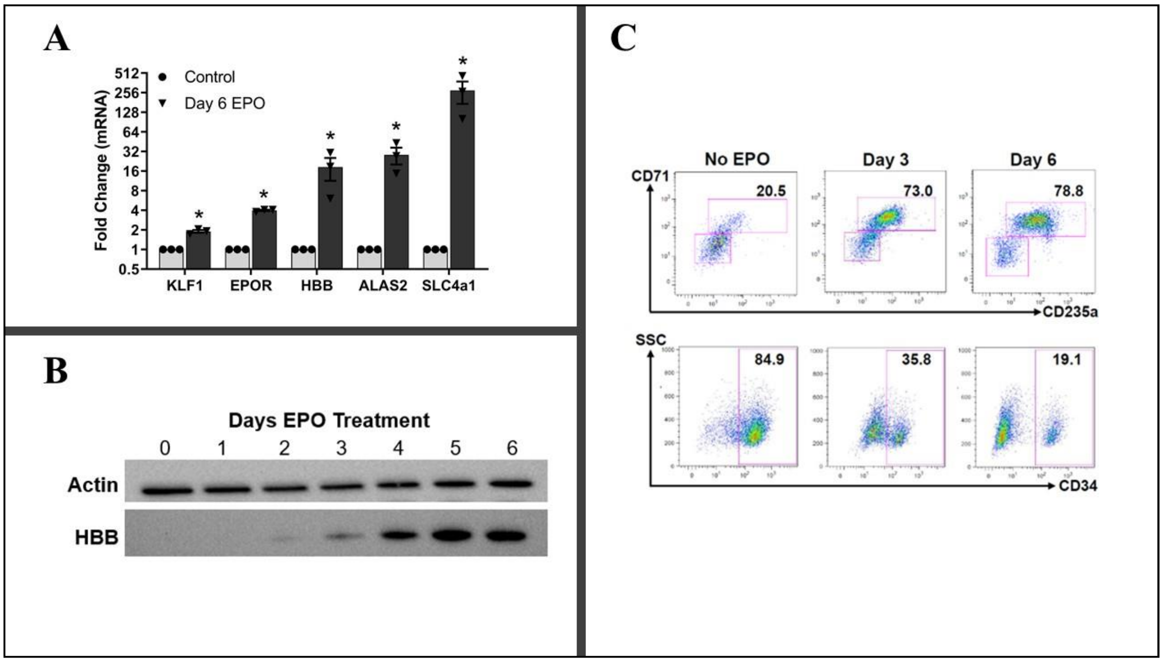

Figure 2.5. Measuring erythropoiesis *in vitro*

A) RT-qPCR measures key erythroid gene expression in TF1 cell cultures. After a six- day incubation in EPO, erythroid gene expression is significantly increased. Bars represent the mean of three independent experiments +/- SEM. B) Western blot for beta- hemoglobin (HBB) reveals an EPO-dependent increase in protein expression, actin serves as an endogenous control. C). Flow cytometry measures cell-surface expression of erythroid genes three- or six-days post EPO incubation. CD71 – transferrin receptor, CD235a – glycophorin A, CD34 – HSPC marker.

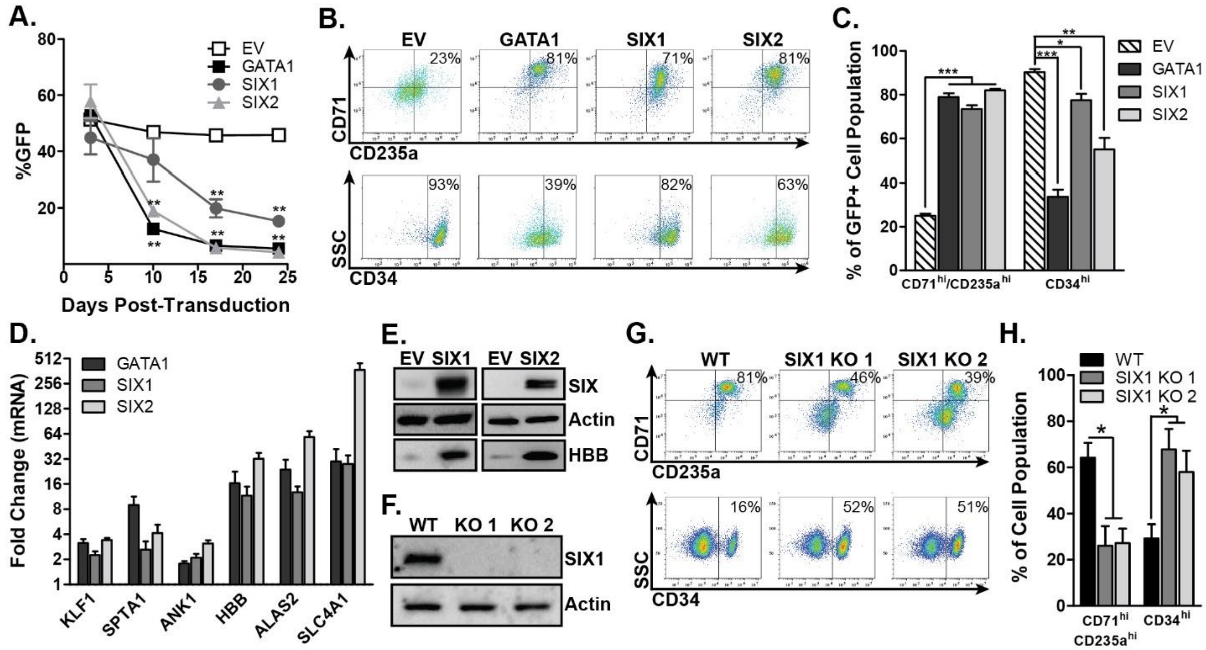

Figure 2.6. SIX overexpression phenocopies GATA1 and drives TF1 cell erythroid differentiation.

(A) GFP competition assays conducted on TF1 cell populations transduced to MOI=0.5 (~50% GFP$^+$) with GFP-expressing empty vector (EV), GATA1, SIX1, or SIX2 lentiviruses. (SEM, n=3, **p < 0.01 compared to EV). (B-D) TF1 cells were transduced with EV, GATA1, SIX1, vs SIX2 MOI=2 (~90% GFP$^+$) and cultured in standard growth media supplemented with GM-CSF. Cells were analyzed 6 days post- transduction. (B) Representative flow cytometric analysis for erythroid (CD71^{hi}CD235a^{hi}) and HSPC (CD34$^+$) markers. (C) Quantitation of flow analysis performed in Flowlogic software. (Bars represent mean ±SEM of GFP$^+$ cells, n=3, *p<0.05, **p<0.01, ***p < 0.005). (D) Total RNA was harvested from cells and levels of indicated erythroid gene RNA quantified by qRT-PCR (Bars represent mean ±SEM, n=3, all values were statistically significant compared to EV, p < 0.05). (E) Western blot from SIX1 or SIX2 transduced TF1 cells after 6 days in GM-CSF culture. Total cellular lysates were harvested and analyzed with Western blot for hemoglobin β (HBB), actin, and either SIX1 or SIX2. (F) Western blot of CRISPR/Cas9-engineered SIX1 knockout cell lines. (G) Representative flow cytometric analysis for erythroid cells (CD71^{hi}CD235a^{hi}) and CD34 in WT vs SIX1 KO TF1 cells 3 days after EPO stimulation (H) Quantitation of flow analysis performed in Flowlogic software. (Bars represent mean ±SEM, n=3, *p< 0.05).

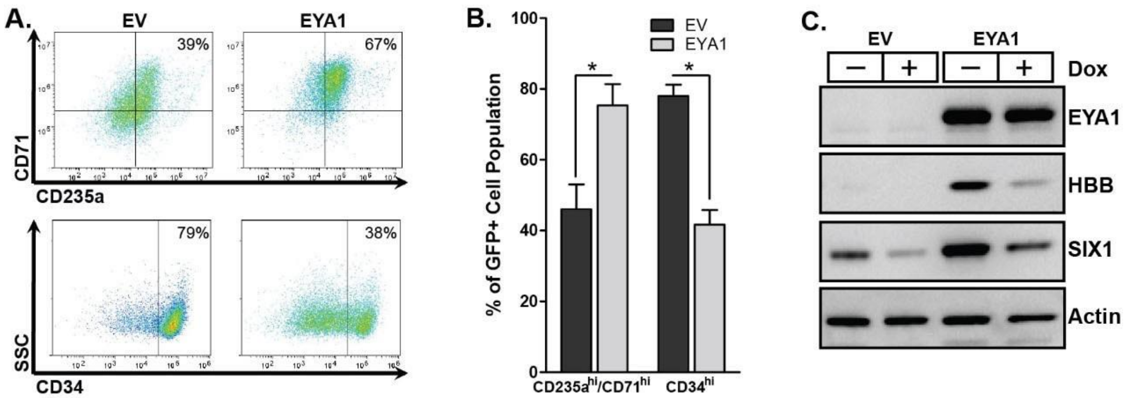

Figure 2.7. EYA stimulates erythroid differentiation via SIX1

WT TF1 cells were transduced to an MOI=2 (~90% GFP$^+$) with EV or EYA1-expressing lentiviruses cultured for 10 days in GM-CSF. (A) Representative flow cytometry plots of transduced TF1s stained for erythroid markers (CD71^{hi}/CD235a^{hi}) and CD34. (B) Quantification of three independent experiments performed in Flowlogic (Bars represent mean ±SEM, n=3, *$p < 0.05$). (C) TF1 cells harboring doxycycline (DOX) inducible *TRE*-dCas9-KRAB cells and SIX1 targeting sgRNA were cultured in the presence or absence of DOX for 4 days prior to transduction with EV vs EYA1-expressing lentiviruses at MOI=2 (~90% GFP$^+$). Cells were cultured an additional 10 days in GM-CSF media supplemented with DOX every other day and then total protein lysates were harvested and analyzed via Western blot for EYA1, SIX1, HBB, and actin.

2.2.3 SIX proteins modulate erythropoiesis in human hematopoietc stem-progenitor cells (HSPCs)

Since our previous experiments were conducted solely in TF1 erythroleukemia cell lines, we wanted to know if this effect extended to primary donor HSPCs. Therefore, we utilized primary human stem-progenitor cells (HPSCs) and tested whether SIX1 influences erythropoiesis. HSPCs transduced with control vs SIX1 expressing lentiviruses were analyzed by flow cytometry after three days culture in media supplemented with stem cell factor (SCF) and EPO. Overexpression of SIX1 in primary HSPCs resulted in an increased percentage of cells expressing erythroid surface markers ($CD71^{hi}$ $CD235a^{hi}$) (Figure 8A-B). Similar results were also observed for SIX2 overexpression (Figure 9). In contrast, HSPCs transduced with SIX1 knockout CRISPR/Cas9 lentiviruses exhibited reduced erythroid cell generation compared to control transduced HSPCs (Figure 8C-D). Thus, SIX1 and SIX2 can enhance erythropoiesis in human cell lines as well as in primary donor HSPCs.

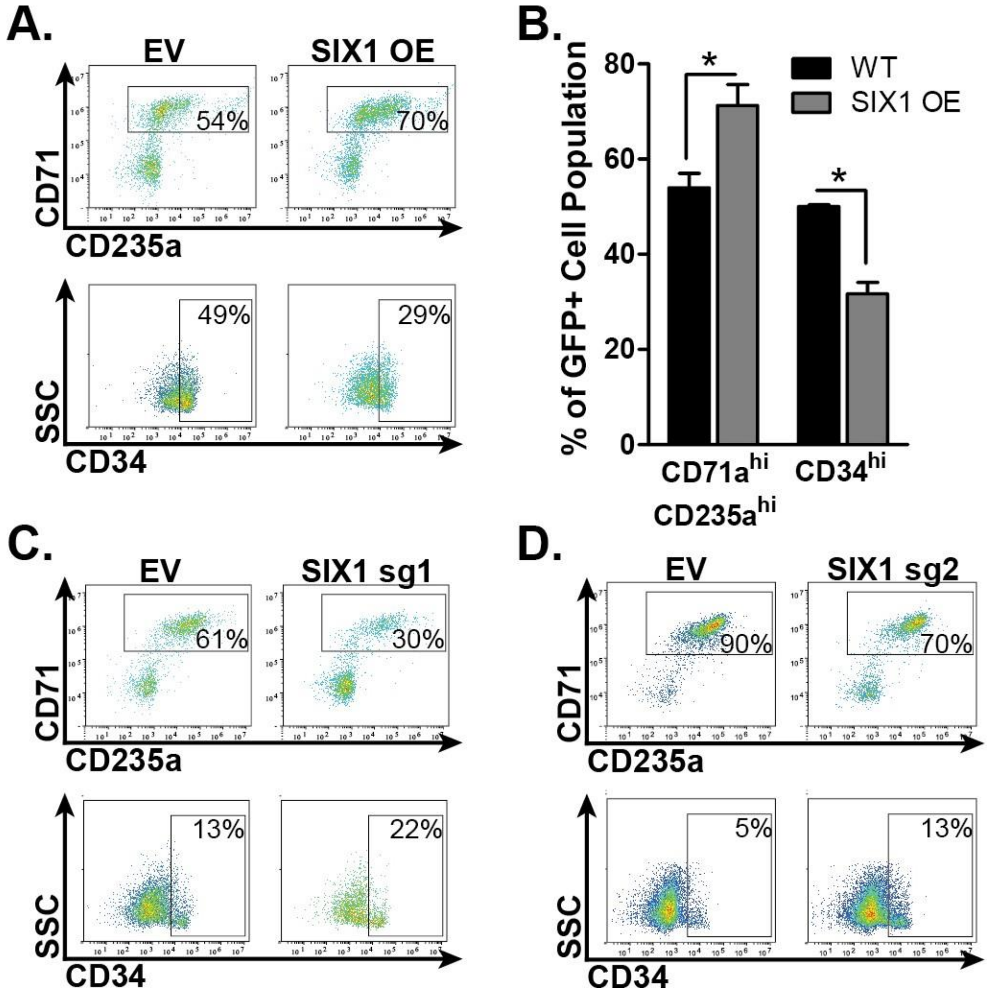

Figure 2.8. SIX1 overexpression drives erythroid differentiation, whereas CRISPR/Cas9-mediated SIX1 knockout inhibits erythroid differentiation in primary human CD34^{+} HSPCs.

(A) Representative flow cytometric analysis for erythroid (CD71^{hi}CD235a^{hi}) cells. Three days after transduction with lentiviruses, EV- vs SIX1 OE-transduced HSPCs were cultured in the presence of SCF and EPO for 3 days. (B) Quantification of three independent experiments. Symbols represent mean ±SEM. Each value is statistically different from EV ($p<0.05$). (C) Flow cytometric analysis for erythroid (CD71^{hi}CD235a^{hi}) cells. Three days after transduction with lentiviruses, EV- vs SIX1 KO (sgRNA1)-transduced HSPCs were cultured in the presence of SCF and EPO for 6 days.
(D) Flow cytometric analysis for erythroid (CD71^{hi}CD235a^{hi}) cells. Three days after transduction with lentiviruses, EV- vs SIX1 KO (sgRNA2)-transduced HSPCs were cultured in the presence of SCF and EPO for 6 days.

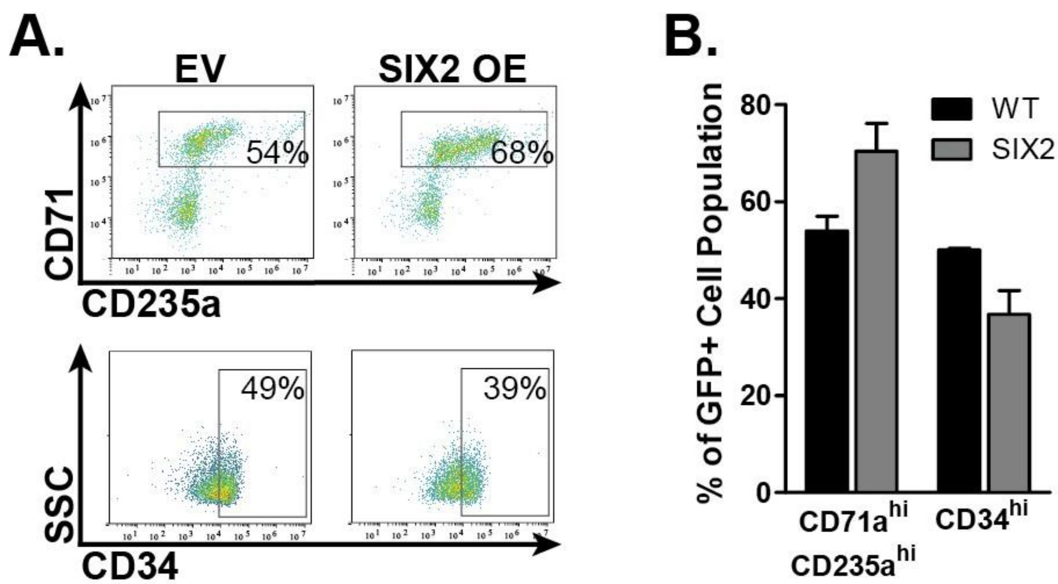

Figure 2.9. SIX2 overexpression stimulates erythropoiesis in HSPCs.

HSPCs transduced with control vs SIX2 expressing lentiviruses were analyzed by flow cytometry after 3 days culture in media supplemented with stem cell factor (SCF) and EPO. Overexpression of SIX2 in primary HSPCs resulted in an increased percentage of cells expressing erythroid surface markers (CD71^{hi} CD235a^{hi}).

2.2.4 SIX1-mediated TF1 erythroid differentiation requires GATA1

We next tested whether stimulation of erythropoiesis by SIX1 was dependent on GATA1. Overexpression of SIX1 in wild type vs GATA1 knockout cells revealed that GATA1 was required for SIX1-driven generation of CD71^{hi}CD235a^{hi} erythroid cells, reduction of the CD34^{+} cell population, increased hemoglobin protein expression, and increased KLF1, SPTA1, ANK1, HBB, ALAS2 and SLC4A1 mRNA expression (Figure 10A-D). These findings are consistent with the SIX1-driven erythroid phenotype being mediated by GATA1. In contrast, the ability of SIX1 to increase CD235a surface expression appeared to be largely independent of GATA1 as shown in both GATA1 knockout clones in the CD71^{hi}CD235a^{hi} plots along the X-axis (Figure 10A), suggesting SIX1 may regulate CD235a expression via multiple mechanisms.

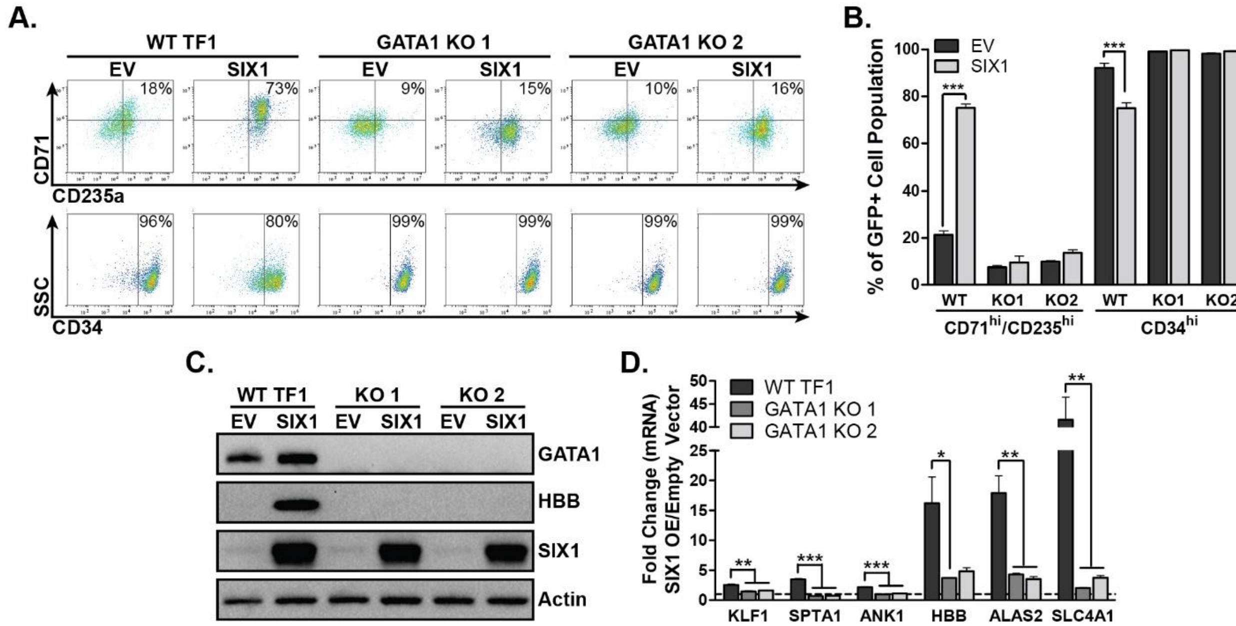

Figure 2.10. SIX1 stimulates erythroid differentiation via GATA1.

TF1 cells cultured in GM-CSF were transduced to at MOI=2 (~90% GFP^+) with empty vector control (EV) vs SIX1 expressing lentiviruses and analyzed 6 days post- transduction. (A) Representative flow cytometry plots of wild type (WT) vs 2 independent GATA1 knockout (KO). (B) Analysis of 3 independent experiments analyzed by Flowlogic software (Bars represent mean ±SEM, ***p < 0.005). (C) Representative Western blot of protein extracts generated from WT TF1s vs GATA1 KOs probed with anti-HBB, anti-GATA1, anti-SIX1 and anti-actin. (D) qRT-PCR analysis of erythroid gene panel conducted on total RNA isolated from three independent experiments. (Bars represent mean fold change ±SEM relative to EV (dotted line), n=3,*p<0.05, **p<0.01, ***p < 0.005).

2.3 Materials and Methods

2.3.1 Plasmids

All plasmids and primers used in this study are outlined in Figure 11. Gene expression constructs were generated by PCR or G-blocks synthesis and cloned into GFP-marked lentivector pWCC43 (Kim *et al*, Br J Haematol 2015). BirA* was PCR amplified from pcDNA3.1 mycBioID plasmid (gift from Kyle Roux, Addgene plasmid #35700). SIX1 sgRNAs were synthesized as primers (Integrated DNA Technologies) and cloned into pSpCas9(BB)-2A-Puro (PX459) V2.0, gift from Feng Zhang (Addgene plasmid 62988) or plenti SpBsmBI sgRNA Puro vector, gift from Rene Maehr (Addgene plasmid # 62207)(Pham *et al*, Methods Mol Biol 2016). pHAGE TRE dCas9-KRAB was a gift from Rene Maehr and Scot Wolfe (Addgene plasmid #50917)(Kearns *et al*, Development 2014). plentiCRISPRv2GFP was a gift from David Feldser (Addgene #82416) (Walter *et al*, Cancer Res 2017).

2.3.2 Cell culture and transduction

TF1 human erythroleukemia cells (CRL-2003; American type Culture Collection) were cultured as previously described (Kim *et al*, Br J Haematol 2015) except recombinant human GM-CSF was from TONBO Bioscience. HEK293T cells (American Type Culture Collection) were cultured in DMEM media (CellGro) containing 10% FBS. Lentivirus transductions were conducted in the presence of 8ng/ml polybrene.

2.3.3 Lentivirus transduction and erythroid differentiation of human primary $CD34^+$ HSPCs

Human $CD34^+$ HSPCs from normal adult donors were obtained from the Cellular Therapy and Cell Processing Facilities (Fred Hutchinson Cancer Center, Seattle, WA, USA). To overexpress or knockout SIX1 in primary human $CD34^+$ HSPCs, lentivirus transductions were conducted with 1 mg/ml Kolliphor P407 (Sigma, St. Louis, MO, USA) and 10uM 16,16-dimethyl-postaglandin E2 (dmPGE2; Abcam, Cambridge, MA, USA) in StemSpanSFEM medium (Stemcell Technologies, Vancouver, Canada) containing 100 ng/ml human stem cell factor (SCF; PeproTech), 20 ng/ml thrombopoietin (TPO; PeproTech), and 100 ng/ml FMS-like tyrosine kinase receptor 3 ligand (FLT3L; Peprotech) (Kim *et al*, Br J Haematol 2015; Masiuk *et al*, Mol Ther Methods Clin Dev, 2019). $CD34^+$ HSPCs were transduced with lentiviruses at MOI=100, which is calculated based on the transduction units in HEK293T cells. Transduced HSPCs were cultured in StemSpanSFEM medium (Stemcell Technologies, Vancouver, Canada) containing 100 ng/ml human stem cell factor (SCF; PeproTech) and 1U/ml erythropoietin (EPO; TONBO Bioscience) to induce erythropoiesis. At specific time points, total cell numbers were counted by trypan blue dye exclusion method and erythroid differentiation was quantified by flow cytometry analysis for cell surface marker expression or by qRT-PCR analysis for mRNA expression. SIX1 knockout in CD34 cells was conducted by cloning SIX1 targeting guide RNA sequences into plentiCRISPRv2GFP, a gift from David Feldser (Addgene #82416) (Walter *et al*, Cancer Res 2017)

2.3.4 Flow cytometry

TF1 cells were stained with monoclonal antibodies against CD235a (551336, BD Biosciences), CD71 (25-0719-42, eBioscience), and CD34 (555824, BD Biosciences), assessed by flow cytometry (Accuri C6, BD Biosciences) and data analyzed using FlowLogic Software (Milltenyi Biotec) as previously described (Kim *et al*, Br J Haematol 2015). A representative gating strategy is shown in Supplemental Figure 6. For GFP competition assays, lentivirus-transduced (MOI=0.5) TF1 cells containing ~50% GFP^+ cells on day 3 post-transduction, were assayed weekly by flow cytometry to determine % GFP^+.

2.3.5 Protein isolation and Western blot analysis

Protein lysates from transduced TF1 cells were quantitated and analyzed by Western blot as previously described (119) using anti-SIX1 (12891, Cell Signaling), SIX2 (ab111827 Abcam; detects SIX1 and SIX2), HBB (sc-21757, Santa Cruz), GATA1 (sc-266, Santa Cruz), GATA2 (sc-267. Santa Cruz), EYA1 (H00002138-A01, Abnova), Beta-Actin (8457S, Cell Signaling), with HRP-conjugated secondary antibodies from Jackson Immunoresearch. Signal was detected using enhanced chemiluminescence (ECL; ThermoFisher Scientific) and quantitated using ImageJ (Chen *et al*, Cell 2013).

2.3.6 RNA analysis

Total RNA was extracted and analyzed as previously described (Kim *et al*, Br J Haematol 2015). Reverse transcription was conducted with 1.0-1.5 μg RNA using a High Capacity RNA-to-cDNA kit (ThermoFisher Scientific). qPCR was performed using Power SYBR Green PCR Master Mix (ThermoFisher Scientific) on an Applied Biosystems Quant-

Studio 6 Flex according to manufacturers' protocol and analyzed using QuantStudio Real-Time PCR Software (ThermoFisher Scientific). RNA levels were normalized using levels of internal GAPDH. Sequences are outlined in Table 2.

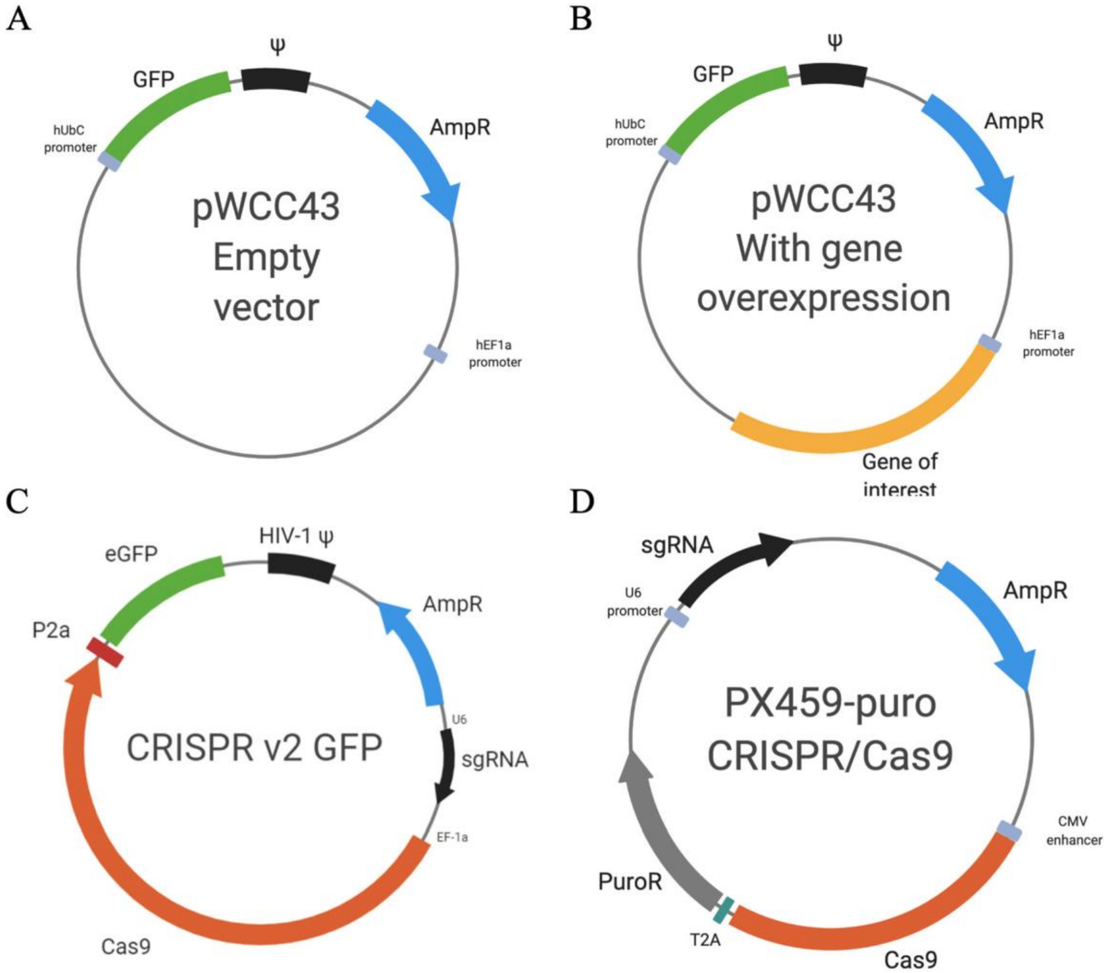

Figure 2.11. Plasmid construct maps

(A-B) pWCC43 constructs were used to overexpress proteins of interest including SIX1, SIX2, GATA1, EYA1, and BirA* constructs. Viruses were made using packaging and envelope plasmids in HEK293T cells. Vectors contain GFP to measure transduction efficiency. (C) Lentiviral construct used to transduce TF1 cell cultures as well as primary HSPCs. Viruses were created in HEK293T cells similarly to pWCC43 viruses. GFP markers were used to measure transduction efficiencies. (D) PX459 CRISPR/Cas9 constructs were used to electroporate TF1 cell cultures instead of lentiviral transductions. Cells were selected with puromycin and subjected to single cell cloning to obtain SIX1 knockouts.

PCR primers:

Name	Sequences (5’ → 3’)
FOG1 Forward	CTCCCAGCGCAGATGTT
FOG1 Reverse	GGTTCTGGTTCCTGTCCTTC
GATA1 Forward	ACACTGTGGCGGAGAAATG
GATA1 Reverse	GTTCACCTGGTGTAGCTTGTAG
GATA2 Forward	GACGACAACCACCACCTTAT
GATA2 Reverse	AGTCTGGATCCCTTCCTTCT
HBB Forward	TCCACTCCTGATGCTGTTATG
HBB Reverse	GAGCCAGGCCATCACTAAA
SLC4A1 Forward	CTGTTCAAGCCACCCAAGTA
SLC4A1 Reverse	GATCTGGATGCCCGTGAATAA
ALAS2 Forward	CAACATCTCAGGCACCAGTAAG
ALAS2 Reverse	CAACAAAGCAGGAGGAGAAGAG
KLF1 Forward	CCTGTTTGGTGGTCTCTTCA
KLF1 Reverse	TGTCTATGGGTCCGTGTTTG
SPTA1 Forward	GAGTCCTTGGTATGCCGATTT
SPTA1 Reverse	TCCTCATCCTCTGTGTCTCTAC
ANK1 Forward	CGCCTCCACTGAATCTACAC
ANK1 Reverse	GGCAGAGAGAGAGACAACATAAA
GAPDH Forward	GGTGTGAACCATGAGAAGTATGA
GAPDH Reverse	GAGTCCTTCCACGATACCAAAG

Table 2.2 SYBR green primer sequences used in RT-qPCR experiments

2.3.7 <u>CRISPR/Cas9-engineered cell lines</u>

To generate GATA1 and SIX1 knockout cell lines, corresponding sgRNAs targeting exon 3 and exon 1, respectively, were cloned into PX459 and DNA electroporated using Amaxa nucleofector 4D system. Following puromycin selection, single cell clones were generated via plating at 0.5 cells/well in 96 well plates. Knockout cell lines were confirmed by sequence and Western blot. For inducible dCas9-KRAB parental cell line, TF1 cells were transduced with pHAGE TRE dCas9-KRAB lentivirus (Kearns *et al*, Development 2014), G418 selected and single cell clones generated. sgRNA13, which spans the SIX1 transcriptional start site was cloned into plenti SpBsmBI sgRNA Puro vector (pTJK301) or a modified version of plenti SpBsmBI sgRNA Puro vector (pTJK466) engineered to contain U-A flip and extended hairpin reported to improve targeting (Chen *et al*, Cell 2013). Corresponding lentiviruses were transduced into TF1 dCas9-KRAB cells to generate inducible SIX1 knockdown cell lines. SIX1 knockdown efficiency was assessed by Western blot analysis following 3-10 days doxycycline treatment. Oligo sequences used for CRISPR/Cas9-engineered cell lines are outlined in Table 3..

Name	Purpose	Sequences (5'-3')
SIX1sgRNA13 TOP	CRISPRi	ACACCGTGGATGCTGCTAGTTGCCGG
SIX1sgRNA13 BOT	CRISPRi	AAAACCGGCAACTAGCAGCATCCACG
SIX1sgRNA13* BOT	CRISPRi	TAAACCGGCAACTAGCAGCATCCACG
SIX1sgRNA1 TOP	KO	CACCGTGCTTCAAGGAGAAGTCGAG
SIX1sgRNA1 BOT	KO	AAACCTCGACTTCTCCTTGAAGCAC
SIX1sgRNA2 TOP	KO	CACCGGCTGAAGGCGCATTACGTGG
SIX1sgRNA2 BOT	KO	AAACCCACGTAATGCGCCTTCAGCC
GATA1 sgRNA	KO	CACCGGCTGGGCCTACGGCAAGACG
GATA1sgRNA	KO	AAACCGTCTTGCCGTAGGCCCAGCC

*modified for cloning into pTJK466

Table 2.3. sgRNA sequences used in CRISPR/Cas9 experiments

2.4 <u>Discussion</u>

The discovery by Dr. Fossett that PSEDN homologues regulated hematopoiesis in the fly was an exciting and novel discovery (Creed et al, Development 2020). Heterozygous mutants of PSEDN member *eya* and GATA/FOG member *u-shaped* (FOG) reduced the amount of *oddskipped*-positive prohemocytes. In fact, other members of the PSEDN including *so* (SIX) showed aberrant expression of lamelocytes (Table 1), suggesting that the PSEDN is required for *srp* (GATA) to maintain prohemocyte numbers. Given that loss- and gain-of-function studies of the SIX1 homologue (*so)* showed differential expression levels of the GATA1 homologue (*srp*), the data are consistent that *so* gain-of-function increases lamellocyte production at the expense of prohemocyte numbers (Figure 2).

Our collaboration with Dr. Fossett elucidated a conserved transcriptional regulation between PSEDN and GATA/FOG network from flies to humans. Excitingly, we found that SIX1 and SIX2 were novel regulators of human erythroid development. Using erythropoiesis, a robust model for mammalian GATA function where GATA1/FOG1 drives central features of erythroid differentiation, we demonstrate that human PSEDN members SIX and EYA enhance GATA1-dependent cell phenotypes *in vitro*. Overexpression of SIX1 or SIX2 in TF1 erythroleukemia cells regulated expression levels of GATA1 and GATA2, reminiscent of levels seen during EPO-induced erythropoiesis. Additionally, SIX1 and SIX2 overexpression phenocopied GATA1 overexpression, triggering erythropoiesis in the absence of exogenous EPO. In primary human HSPCs, enforced SIX expression enhanced EPO-stimulated erythropoiesis. In contrast, SIX1 knockout impaired erythropoiesis, suggesting endogenous SIX1 contributes to human erythropoiesis *in vitro*. Our observation that SIX overexpression could not stimulate erythropoiesis in the absence of GATA1 indicates that SIX proteins function in conjunction with, or upstream of, GATA1. This conclusion is supported by our finding that SIX overexpression increases GATA1 and decreases GATA2 RNA and protein levels, as occurs during EPO-stimulated erythropoiesis (Kaneko, Curr. Opin. Hematol 2010).

This may have potential consequences outside of erythroid differentiation that warrant further study. For example, both SIX1 and SIX2 and GATA1 increase metastasis in breast cancers (Iwanaga *et al*, Breast Cancer Research 2012; Li *et al*, Oncotarget 2015; Xie *et al* Int J Clin, Exp Pathol 2016). Future studies could examine the expression of GATA1 with and without SIX1 to see if clinical outcomes varied. We could imagine a scenario where SIX1 breast cancers are unable to induce EMT in the absence of GATA1. This could be a particularly powerful diagnostic marker for breast cancer patients. In the next chapter we will examine the effect that SIX1 has on GATA1-transcriptional output and the formation of SIX/GATA1 complexes.

Chapter 3: RNA seq reveals common targets of SIX1 and GATA1

3.1 Introduction

Overexpression of either SIX1 or SIX2 enhanced erythropoiesis in CD34$^+$ primary hematopoietic stem-progenitor cells (HSPCs) and TF1 erythroleukemia cell cultures in the presence of exogenous erythropoietin (EPO). Consistent with the overexpression effects in erythropoiesis, CRISPR/Cas9-mediated knockout of SIX1 reduced the generation of erythroid cells in response to EPO. This could be due to several potential mechanisms including cell survival, proliferation, or sensitivity to EPO. Here we focus on SIX1 given the current availability of suitable antibodies for protein detection and quantitation as well as downstream immunologic assays such as immunoprecipitation and proximity ligation (discussed in chapter 4). We were unable to test the effect of SIX2 knockout in our system due to the lack of functional antibodies available to detect endogenous SIX2 proteins without forced overexpression and therefore prevented screening for CRISPR-generated knockouts. The RNA signals of SIX2 in our SIX1 knockout cells do not change suggesting that a compensatory mechanism between SIX proteins is unlikely. Though there exists the possibility that while levels of other SIX members do not change, their presence may act in a compensatory manner. For example, during mouse myogenesis Six1 and Six4 are expressed in overlapping areas including neural placodes, limb bud mesenchyme, and myogenic migrating precursors, where they act in concert to control myogenesis (Esteve and Bovolenta, Mech Dev. 1999; Oliver *et al*, 1995). However, although Six1$^{-/-}$ null mice die embryonically and Six4$^{-/-}$ mice show no obvious skeletal defects, the double Six1$^{-/-}$ Six4$^{-/-}$ knockout mice, show more severe muscle defects than did Six1$^{-/-}$ alone (Grifone *et al*, Development 2005). Future studies should test whether knockout of multiple SIX proteins have a more severe hematopoietic deficiencies than SIX1 alone.

Though second site noncomplementation (SSNC) assays may, but not necessarily, identify physically interacting proteins, we utilized a global RNA sequencing approach to detect not only SIX1-induced gene expression, but also to identify target-genes of additionally active transcriptional regulators. We further showed that GATA1 is necessary for SIX1 to stimulate erythroid differentiation as evidenced by GATA1 loss-of-function studies in TF1 cell cultures (Chapter 2, Figure 10). These findings provide the first evidence of PAX-SIX-EYA-DACH Network (PSEDN) members as regulators of erythropoiesis. While SIX1 knockout mice do not have any obvious hematopoietic phenotypes (Zhang *et al*, EBioMedicine 2019), new research suggests that SIX1 and SIX2 control expression of inflammatory response genes through interactions with NF□B in macrophages and regulate GATA2 and GATA3 levels in lymphomas, suggesting they have crucial roles in hematopoietic cells (Liu *et al*, Nature 2019; Nagel *et al*, Oncotarget 2015).

In both mouse primary macrophage cells as well as in human cancer cell lines cell lines SIX1 and SIX2 negatively regulated inflammatory signals through interactions with RELA and RELB members of the NFκB signaling pathways (Liu *et al*, Nature 2019). In mice this interaction promoted recovery of mice from LPS-induced endotoxic shock. However, the cofactors and downstream effectors of SIX1 and SIX2 in hematopoiesis remain incompletely understood.

We used RNA seq to determine the global transcriptome of SIX1 overexpression TF1 cell cultures. This provides the opportunity to determine the enrichment of genes known to be involved in signaling pathways or as targets of known transcription factors. In thyroid carcinoma SIX1 activates STAT3 signaling in conjunction with EYA1 through increasing the levels of phosphor-Try107 (Kong *et al*, Front. Oncol. 2019). Increased STAT3 signaling lead to a downstream increase of c-MYC and BCL-XL levels to promote cellular proliferation and protection from apoptosis. This is a similar mechanism observed in rhabdomyosarcoma, a rare form of skeletal muscle cancer, where Six1 expression promotes protumorigenic genes such as c-Myc and cyclin D1 in mice (Yu *et al*, Cancer Res. 2006).

The function of the SIX family of proteins is regulated by cofactors that operate in a cell-dependent manner. In developing mammalian auditory organs Six1 function is dependent upon binding Eya and it has been shown *in vivo* that Six1 or Six2 can interact with Eya1 (Ahmed *et al*, Dev Cell 2012; Buller *et al*, Hum Mol Genet 2001). This interaction couples the protein phosphatase activity of Eya1 to the Six1 or Six2 DNA-binding function to control the genetic transcription or repression of Six-target genes such as Fgf3, Fgf10, Bmp4, and Gata3. Furthermore, during myogenesis Six1 and Six4 coordinate with MyoD on cis-regulatory models (CRMs) to control synergistically the expression of approximately 96 target genes capable of reprogramming differentiated cells towards a myogenic fate (Santolini *et al*, Nucl Acids Res 2016). This was determined using an *in silico* approach utilizing microarray analysis coupled with *de novo* gene motif finding algorithms. We adopted a similar approach using RNA sequencing datasets in combination with gene set enrichment analysis (GSEA) to identify putative

SIX1-interacting genes that may coordinate to control erythroid gene expression in our TF1 cell model.

RNA sequencing of three independent SIX1 overexpression experiments in TF1 erythroleukemia cell lines showed specific enrichment of genes linked with erythropoiesis and heme metabolism and suppression of genes that are actively involved in inflammatory signaling pathways. Further GSEA designed to identify transcription factor binding sites (TFBS) within differentially regulated genes revealed that genes harboring GATA-binding sites are specifically enriched. Looking at specific-target genes of GATA1 filtered through our RNA sequencing dataset reveals that SIX1 overexpression can regulate both, positively and negatively, many of these same genes. Though this could be due to GATA-factor switching downstream of SIX1 overexpression, *in silico* analysis of several erythroid gene promoters show putative binding sites of SIX1 and GATA1 near one another. Table 1 shows a fraction of differentially regulated genes that are altered in response to SIX1 overepxression. Though these are only a handful of genes regulated by GATA1 and SIX1, they show potential binding sites relative to the transcription start sites that could be identified in future studies using chromatin immunoprecipitation (ChIP). Taken together these data suggests that perhaps SIX1 and GATA1 cooperate in erythroid cell cultures to promote erythroid differentiation.

Upregulated	# GATA1 TFBS	Location Relative to TSS	# SIX1 TFBS	Location Relative to TSS
CA1	1	**-181**	5	**-183**, -237, -487, -719, -837
SLC4A1	2	-66, -146	2	-607, -842
HBZ	2	-228, -111	2	-621, -327
PRG2	2	-79, -786	2	-360, -565
PRDX2	2	-521, -666	3	-425, -577, -728
KLF1	2	**-5**, -41	1	-158
RHAG	4	-23, **-115**, -605, -640	2	**-162**, -258
LTBP1	3	-835, **-637**, -142	2	-776, -763
HOXB9	2	-157, **-284**	4	**-308**, -677, -901, -933

Downregulated	# GATA1 TFBS	Location Relative to TSS	# SIX1 TFBS	Location Relative to TSS
CD34	1	**-964**	4	-22, -255, -832, **-968**
TFEC	3	-529, -672, -991	5	-123, -335, -584, -772, -889
ARHGAP15	3	-744, -530, -16	2	-692, -366
SERPINB6	6	-37, -369, -413, -524, -589, -853	7	-108, -197, -260, -283, -498, -836, -916
CLCN5	2	-706, **-55**	2	-399, **-27**
ERG	1	**-138**	2	**-85**, -455

Table 3.1. Genes showing altered expression in response to SIX1 OE have potential SIX1 and GATA1 binding sites

In silico analysis of known GATA1-target genes that are differentially regulated in SIX1 overexpressing TF1 cell lines were examined for potential transcription factor binding sites (TFBS) in their promoters using the Eukaryotic Promoter Database (Swiss Insitute of Bioinformatics). Noted here are several of the genes that harbor neighboring binding sites of the GATA-specific and SIX-specific DNA-binding sequences.

3.2 Results

3.2.1 RNA sequencing analysis of SIX1 OE TF1 cells

We utilized a GFP-marked lentiviral-based overexpression system in TF1 erythroleukemia cell lines. TF1 cells were transduced to an MOI ~ 2 (90% GFP) with either empty vector (EV) or SIX1 overexpressing lentiviruses. Six days post-transduction RNA libraries were generated from three independent experiments using a RNeasy kit according to the manufacturer's protocol (Qiagen). These libraries were purified and sequenced using an Illumina HiSeq 4000. Global analysis of mRNA profiles in SIX1 overexpression TF1 cells reveals genes that are differentially regulated in response to enforced SIX1 expression. Principal component analysis (PCA) analysis of the top 500 differentially regulated genes with the largest coefficient variation based on fragments per kilobase of transcript per million mapped reads (FPKM) abundance show a distinguishable gene expression profile among the samples. Clustering of the EV samples (blue bars) from the SIX1 overexpression samples (green bars) is consistent with a differential expression that results from forced SIX1 expression (Figure 1A). Hierarchical cluster analysis of significantly differentially expressed genes in SIX1 vs. EV organizes the gene expression levels from highest upregulated genes (red) to the genes that are most downregulated (green). The result from the SIX1 overexpression hierarchical cluster analysis shows a distinct gene expression profile among samples (Figure 1B). Volcano plot analysis allows for the identification and visualization of genes displaying large magnitude changes that are also statistically significant. Volcano plot analysis of SIX1 vs EV TF1 cell cultures shows the significantly regulated genes in response to SIX1 overexpression (Figure 1C). The genes listed in Tables 2-3 highlight the genes whose differential expression reaches a significance of $p<0.05$ between the three independent replicates and had a greater than 1.5-fold change. The highest upregulated gene in Table 2 is SIX1 and is a consequence of forced SIX1 overexpression and was used as a control to ensure data analyses were functioning properly. We then applied GSEA to our SIX1 overexpressing TF1 RNA sequencing data to highlight any enriched biochemical pathways, signaling pathways, or genes enriched that harbor known binding sites in their promoters (Subramanian *et al*, PNAS 2005). Only two hallmark gene sets were significantly upregulated ($p<0.05$) in response to SIX1 overexpression according to GSEA with "Heme Metabolism" at the top of the list (Figure 2A, Table 4). This mirrors what we see in response to SIX1 overexpression when we measure erythroid-specific markers using Western blot, flow cytometry, and RT-qPCR (Chapter 2, Figures 6-7, 9). Additionally, GATA1 is known to drive genes involved in the heme biosynthetic pathway (Tanimura et al, EMBO Rep 2016). Conversely there were 13 different hallmark gene sets that were downregulated in response to SIX1 overexpression (Table 6). Interestingly, the pathway most negatively regulated by SIX1 was the IL6/JAK signaling pathway that is activated by inflammatory signaling (Figure 2B). It has been previously shown that IL6 is capable of suppressing hemoglobin production as well as additional erythroid markers, suggesting that SIX1 may enhance erythropoiesis by downregulating IL6/inflammatory signaling, consistent with data concerning SIX1 expression in macrophages (McCranor *et al*, Blood Cell Mol Dis 2014; Liu *et al*, Nature 2019).

Enrichment analysis designed to group genes that harbor distinct transcription factor binding site was conducted using GSEA and found 19 statistically significant enrichments (Table 6) (Subramanian *et al*, PNAS 2005). Of those groups that were significant, 5 of the 19 groups were identified as having GATA binding sites within their promoters (Table 6, Figure 3). Many genes known to be regulated by GATA1, including both up- and downregulated genes, are also differentially regulated in our SIX1 overexpression dataset (Figure 4).

The largest statistical enrichment of genes harboring specific TFBS in their promoter sequences contained genes that are targets of Prohibitin-2 (PBH2), a negative regulator of transcription, largely of ribosomal RNA transcription, in the nucleolus (Zhou et al, Sci Reports 2018). Loss of PBH2 results in reduced differentiation, proliferation, and survival of differentiating erythroblasts (Liu et al, Nature Cell Bio 2017). Perhaps this suggests that not only SIX1 is regulating genes in conjunction with GATA1 proteins (given that SIX1 overexpression modulates the expression of many known GATA1 genes), but it conceivably also plays several roles with multiple cofactors during erythroid maturation.

Research has shown that GATA1 does not occupy all GATA-binding sites during erythropoiesis, rather it selectively regulates gene expression through modulation and interactions with distinct cofactors (Chlon et al, Haematologica 2015; Lowry and Mackay, IJBCB 2006). Given that SIX1 regulates many genes known to be regulated by GATA1, and that SIX1 and GATA1 binding sites appear in close proximity in a host of erythroid-specific genes (Table 1), future studies in this thesis will test the ability for SIX1 and GATA1 to physically interact.

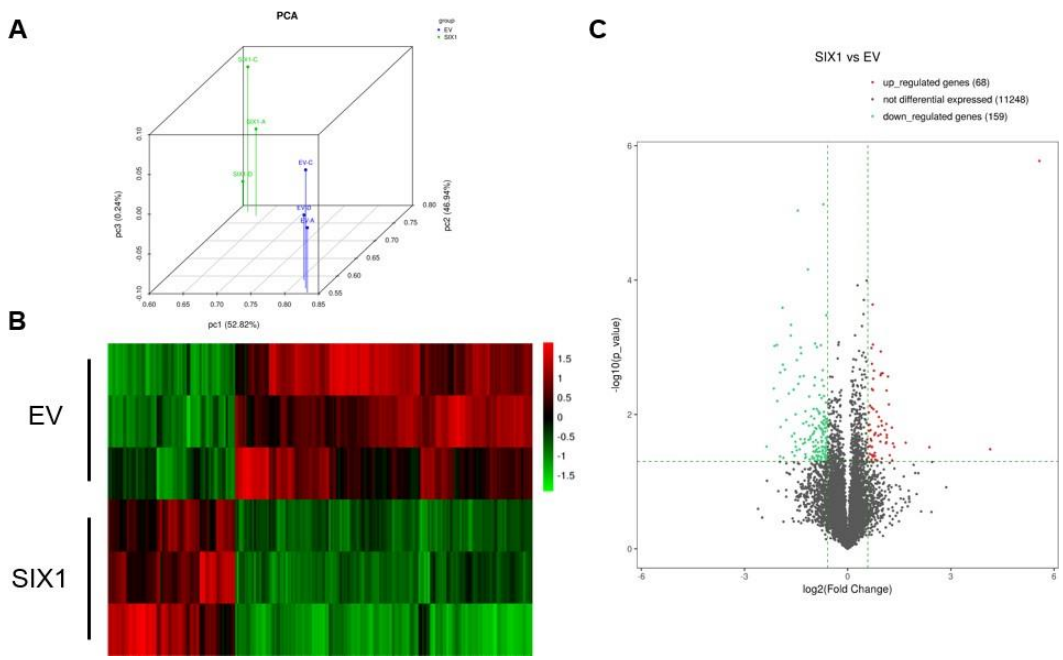

Figure 3.1. RNA sequencing analysis of SIX1 OE in TF1 cell cultures

(A) Principal component analysis (PCA) plot for all samples (EV A,C,D & SIX1 A,C,D). Each circle represents a sample. The largest component (pc1) in the variation is plotted along the X-axis, the second largest (pc2) is plotted on the Y-axis, the smallest (pc3) is plotted on the Z-axis. (B) An unsupervised hierarchical clustering heatmap using the genes of the set that were significantly altered in response to SIX1 OE. Each column represents a gene and each row represent an independently analyzed sample. The color represents log2-scaled FPKM (red indicates higher values, green indicates lower values). (C) Volcano plot for SIX1 vs EV. Red or green circles indicates significant differentially expressed genes with a fold change greater than or equal to 1.5 and a p-val less that 0.05.

Gene Name	FC	p-Value
SIX1	47.6	0.000
AQP1	17.7	0.033
CA1	5.2	0.031
MAFA	3.2	0.026
PPP1R14C	2.6	0.030
LSR	2.5	0.027
SNTA1	2.5	0.048
HIST1H2BK	2.4	0.016
ALPL	2.3	0.040
PTPRF	2.3	0.007
JDP2	2.3	0.003
RHOBTB3	2.2	0.022
HIST1H1C	2.2	0.020
PIK3AP1	2.2	0.004
GALNT5	2.2	0.026
BCL2L2-PABPN1	2.2	0.014
CA2	2.2	0.027
ZNF302	2.1	0.017
RP11-80H18.3	2.0	0.002
KRT13	2.0	0.045
AMMECR1	2.0	0.015
SLC37A1	2.0	0.013
HIST2H2AA3	2.0	0.002
TAX1BP3	1.9	0.001
SLC35D3	1.9	0.003
HIST1H2AC	1.9	0.020
ALOX5	1.9	0.025
HIST2H2AA4	1.9	0.011
SPNS2	1.9	0.015
SMPX	1.9	0.021
ZNF260	1.8	0.049
TUBB2A	1.8	0.048
PPP2R5D	1.8	0.009
BNIP3	1.7	0.019

Gene Name	FC	p-Value
PPARA	1.7	0.047
ADGRG1	1.7	0.019
CYSTM1	1.7	0.041
SELPLG	1.7	0.037
FAM222A	1.7	0.020
EPS8L3	1.7	0.002
HIST2H4B	1.7	0.008
SAT1	1.7	0.003
YBX2	1.7	0.040
USP49	1.7	0.001
LRP10	1.7	0.043
CXCR1	1.7	0.004
AQP3	1.7	0.042
ODAM	1.7	0.000
TFPI	1.6	0.031
CYTIP	1.6	0.004
AL356356.1	1.6	0.002
ATP11C	1.6	0.040
CREB3L3	1.6	0.008
TEAD3	1.6	0.001
TRAF4	1.6	0.047
ASIC4	1.6	0.025
C9orf91	1.6	0.023
ARHGAP18	1.6	0.017
PKP3	1.6	0.030
CORO1C	1.6	0.007
TSPAN33	1.6	0.039
IL1R1	1.6	0.014
USP44	1.5	0.030
AGPAT1	1.5	0.018
FAM24B	1.5	0.009
HSPA6	1.5	0.014
SLC9A7	1.5	0.043
SEPT8	1.5	0.031

Table 3.2. Significantly differentially upregulated genes in SIX1 OE TF1 cells

Significant differentially expressed gene transcripts analysis was performed using the R package ‘ballgown.’ Included genes meet three criteria: fold change (FC) (≥ 1.5), p-val ≤ 0.05, fragments per kilobase million (FPKM) (difference ≥ to 0.05). Genes are ranked by level of expression in response to SIX1 overexpression.

Gene Name	FC	p-Value
IFITM1	0.19	0.030
VSTM1	0.23	0.004
IFITM3	0.23	0.001
COL2A1	0.24	0.001
IL6R	0.24	0.006
CD44	0.26	0.042
KDR	0.26	0.002
FBXO34	0.26	0.015
RP11-343C2.9	0.27	0.000
ERG	0.27	0.002
FAM171B	0.27	0.044
CD38	0.28	0.002
STAB1	0.29	0.012
PLA2G6	0.31	0.005
NLRC3	0.32	0.001
CLCN5	0.32	0.000
SLA2	0.33	0.021
FOXB1	0.33	0.036
S100A10	0.34	0.004
HLA-DPA1	0.34	0.019
METTL7A	0.34	0.041
HLA-DRB1	0.35	0.025
HLA-DPB1	0.35	0.010
HLA-DRB5	0.35	0.030
CD74	0.36	0.034
ARPC5	0.36	0.002
DCP2	0.37	0.000
NPDC1	0.37	0.003
CFP	0.38	0.033
JAK3	0.38	0.003
SHANK3	0.39	0.023
SMIM24	0.39	0.001
MMRN1	0.39	0.001
SERPINB1	0.41	0.027
PCSK9	0.41	0.014
MERTK	0.41	0.003
TMEM150C	0.42	0.023
IRAK3	0.42	0.026
TEK	0.42	0.044
STYK1	0.43	0.009
CALCRL	0.43	0.039
RASAL3	0.43	0.043
TCTEX1D1	0.43	0.013
HLA-A	0.45	0.000
ITGB3	0.46	0.024
FAM171A1	0.46	0.015
CIITA	0.46	0.005
RAB37	0.47	0.047
EVA1B	0.47	0.023
CD226	0.47	0.021
SLC12A6	0.48	0.011
VAMP5	0.49	0.042
SRC	0.49	0.045

Gene Name	FC	p-Value
SCARF1	0.49	0.020
BCAT1	0.49	0.014
ARHGAP42	0.50	0.026
ADRA2A	0.50	0.036
FST	0.50	0.030
PDE4DIP	0.50	0.008
TIE1	0.50	0.019
FAM124B	0.51	0.012
SH3RF3	0.51	0.045
ANTXR1	0.51	0.004
CRISP3	0.51	0.018
CYP4F22	0.51	0.036
PIK3C2B	0.51	0.049
ARAP3	0.52	0.003
MGLL	0.52	0.001
F2RL3	0.52	0.050
FAM43A	0.52	0.036
HLA-DQB1	0.52	0.016
ALDH1A1	0.52	0.029
CD200R1	0.52	0.035
CMTM5	0.53	0.024
LFNG	0.53	0.016
PARVG	0.53	0.041
NOTCH1	0.53	0.033
DOK3	0.54	0.008
PPM1F	0.54	0.001
HLA-DMA	0.54	0.041
PKIG	0.54	0.047
CD40	0.55	0.036
PLA2G16	0.55	0.043
ARRB2	0.55	0.026
EPSTI1	0.55	0.009
ABCB1	0.56	0.041
LPAR2	0.56	0.048
NMU	0.56	0.031
KLRG1	0.56	0.015
MMP23B	0.56	0.010
SLC27A2	0.57	0.005
PLEKHO1	0.57	0.009
ORAI2	0.57	0.036
HIC1	0.57	0.016
SPATS2	0.58	0.042
ZNF48	0.58	0.001
ARHGAP9	0.58	0.049
CDH5	0.59	0.045
MTF2	0.59	0.017
HLA-DMB	0.59	0.022
SESN3	0.59	0.016
POLR3F	0.59	0.050
RCAN3	0.59	0.043
AKNA	0.59	0.013
PLXDC2	0.59	0.006
IL3RA	0.59	0.025

Gene Name	FC	p-Value
DSTYK	0.59	0.030
ANOS1	0.59	0.026
MYO10	0.60	0.003
LEF1	0.60	0.025
SSC4D	0.60	0.035
ZBTB46	0.60	0.017
CDKL1	0.60	0.013
ZYX	0.60	0.034
KRT72	0.61	0.018
IL16	0.61	0.019
ZNF516	0.61	0.002
DENND6B	0.61	0.019
PIK3R5	0.61	0.038
GXYLT1	0.61	0.014
GPR161	0.61	0.020
RAB11FIP3	0.61	0.000
RBM20	0.61	0.030
S1PR1	0.62	0.004
MYOF	0.62	0.011
SAPCD1	0.62	0.039
SH3KBP1	0.62	0.012
GAB2	0.62	0.005
TMEM37	0.62	0.009
RBMS2	0.63	0.014
RGL2	0.63	0.044
CIB3	0.63	0.031
CTSD	0.63	0.011
PLXND1	0.63	0.016
PIK3R6	0.64	0.018
ISPD	0.64	0.042
PLXNA4	0.64	0.040
FHOD1	0.64	0.016
TGFB1	0.64	0.035
SLC22A17	0.64	0.050
COPS8	0.64	0.022
CPNE8	0.64	0.044
KDELC1	0.64	0.047
VANGL1	0.64	0.012
C3orf70	0.64	0.007
SIGLEC5	0.65	0.013
GP5	0.65	0.030
RARRES3	0.65	0.027
CSF2RA	0.65	0.012
SCN9A	0.65	0.000
TMEM170B	0.65	0.041
RP11-949J7.8	0.65	0.034
FNDC11	0.65	0.013
DLL1	0.66	0.024
CSRP1	0.66	0.016
FARP1	0.66	0.022
ATP13A2	0.66	0.031
SLC25A29	0.66	0.025
NCKAP1L	0.66	0.013

Table 3.3. Significantly differentially downregulated genes in SIX1 OE TF1 cells

Significant differentially expressed gene transcripts analysis was performed using the R package ‘ballgown.’ Included genes meet three criteria: fold change (FC) greater than 1.5 fold downregulated (≤ 0.66), p-val ≤ 0.05, fragments per kilobase million (FPKM) (difference ≥ 0.05). Genes are ranked from the largest level of downregulation to least while remaining statistically significant.

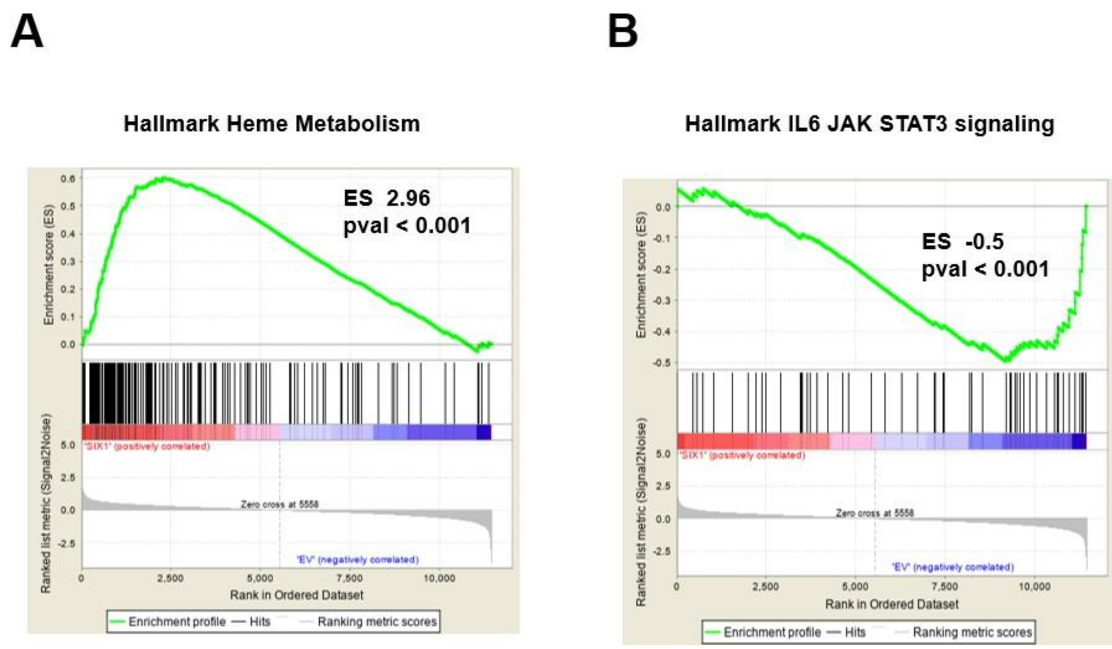

Figure 3.2. Enrichment plots of significantly regulated gene sets in SIX1 OE TF1 cells

Normalized enrichment plots of the most upregulated and downregulated hallmark gene set as determined by GSEA. The ranked gene list of SIX1 OE RNA seq is represented by the red-blue X-axis. (A) Black lines representing individual genes show genes from the 'heme metabolism' set are enriched in SIX1 OE. (B) Genes clustered to the right in 'IL6 STAT3' signaling suggest that genes involved are downregulated in response to SIX1 OE.

Upregulated Hallmark Gene Set	NES	p-val
Heme Metabolism	2.96	0.00
G2M Checkpoint	1.35	0.03
TNFA Signaling Via Nfκb	1.23	0.07
Hedgehog Signaling	1.37	0.10
ESTROGEN Response Late	1.15	0.14
E2F Targets	1.06	0.25
Spermatogenesis	1.04	0.38
UV Response Up	0.98	0.52
UNFOLDED Protein Response	0.96	0.57
Angiogenesis	0.86	0.65
MYC Targets V2	0.78	0.89
Xenobiotic Metabolism	0.81	0.91
Protein Secretion	0.76	0.96
MYC Targets V1	0.78	0.98
DNA Repair	0.66	1.00

Table 3.4. Upregulated gene sets in response to SIX1 OE

Hallmark gene set enrichment analysis (GSEA) was performed against hallmark gene sets using GSEA software. The top 15 unbiased results from the upregulated gene sets are reported here. Only ‘heme metabolism’ and ‘G2M checkpoints’ were considered statistically significant. (NES = normalized enrichment score)

Downregulated Hallmark Gene Set	NES	p-val
IL6 JAK STAT3 Signaling	-1.72	0.00
IFN Gamma Response	-1.71	0.00
Inflammatory Response	-1.62	0.00
UV Response	-1.61	0.00
Complement	-1.56	0.00
Notch Signaling	-1.56	0.02
Epithelial Mesenchymal	-1.55	0.00
Coagulation	-1.51	0.02
IL2 STAT5 Signaling	-1.47	0.01
Myogenesis	-1.38	0.05
KRAS Signaling	-1.35	0.05
Apical Junction	-1.35	0.04
Apoptosis	-1.30	0.06
Glycolysis	-1.29	0.06

Table 3.5. Downregulated gene sets in response to SIX1 OE

Hallmark gene set enrichment analysis (GSEA) was performed against hallmark gene sets using GSEA software. The top 15 unbiased results from the downregulated gene sets are reported here. Only ‘apoptosis’ and ‘glycolysis’ failed to reach statistical significance. (NES = normalized enrichment score)

GSEA Transcription Factor Targets	NES	p-val
PHB2_TARGET_GENES	2.319	<0.001
GATA_C	1.622	<0.001
GATAAGR_GATA_C	1.571	0.003
CREBP1_Q2	1.324	0.013
ZNF589_TARGET_GENES	1.227	0.014
ZNF704_TARGET_GENES	1.55	0.017
ZNF354B_TARGET_GENES	1.314	0.017
CCTNTMAGA_UNKNOWN	1.434	0.02
LMO2COM_02	1.352	0.022
GATA1_04	1.375	0.026
GTF2E2_TARGET_GENES	1.244	0.026
YBX3_TARGET_GENES	1.286	0.028
YAATNANRNNNCAG_UNKNOWN	1.584	0.029
ZNF507_TARGET_GENES	1.187	0.031
GTTGNYNNRGNAAC_UNKNOWN	1.404	0.035
CAGNWMCNNNGAC_UNKNOWN	1.497	0.043
FOXJ2_TARGET_GENES	1.182	0.045
GATA1_03	1.285	0.047
GATA_Q6	1.336	0.049

Table 3.6. Enriched transcription factor binding sites in SIX1 OE-regulated genes

GSEA analysis of the top 19 transcription factor binding sites (TFBS) in enrich SIX1 differentially regulated genes. 5 of the 19 show an enrichment of genes that harbor GATA binding sites in their promoter (red highlighted gene sets). (NES = normalized enrichment score)

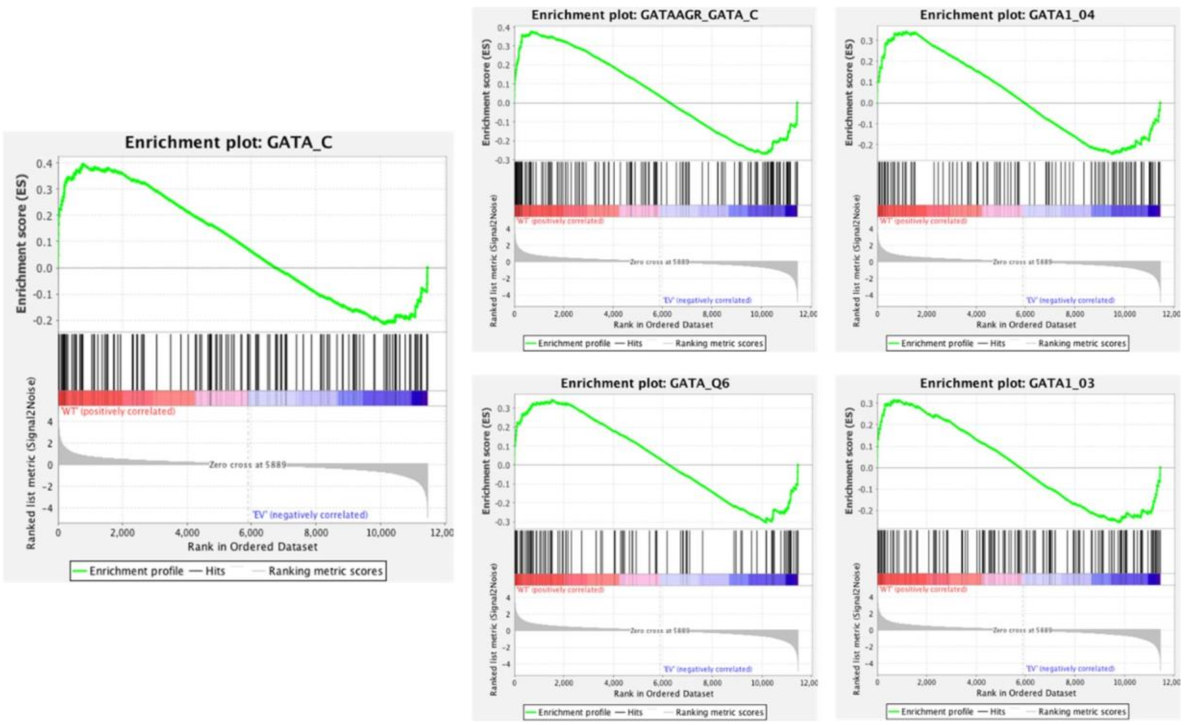

Figure 3.3. Enrichment plots for transcription factor binding sites of GATA in SIX1 OE cells

Enrichment plots for all GATA TFBS discovered via GSEA analysis and are outlined in Table 6. For all 5 enrichment plots you see genes that are both up- and downregulated in response to SIX1.

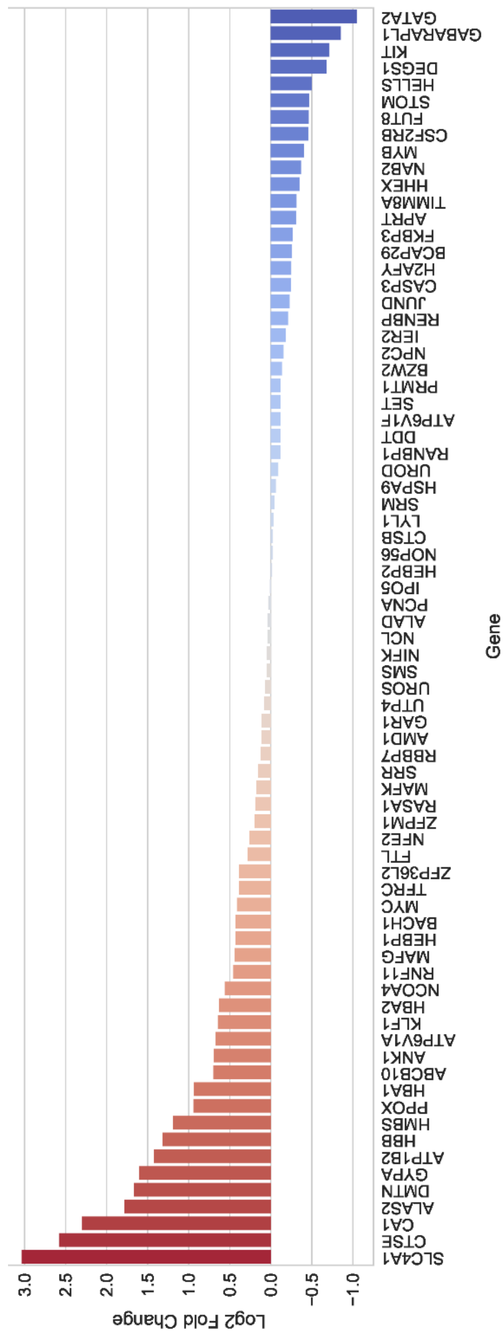

Figure 3.4. Differential regulation of known GATA1-target genes in SIX1 OE cells

Ribbon plot of genes in the SIX1 OE RNA seq dataset that are known GATA1-target genes are differentially regulated. These genes harbor GATA1-binding sites and change, both upregulated and downregulated, in response to SIX1 OE in TF1 erythroleukemia cell cultures.

3.3 <u>Materials and Methods</u>

3.3.1 <u>Cell culture and transduction</u>

TF1 human erythroleukemia cells (CRL-2003; American type Culture Collection) were cultured as previously described (Kim *et al*, Br J Haematol 2015) except recombinant human GM-CSF was from TONBO Bioscience. HEK293T cells (American Type Culture Collection) were cultured in DMEM media (CellGro) containing 10% FBS. Lentivirus transductions were conducted in the presence of 8ng/ml polybrene.

3.3.2 RNA sequencing of TF1 cell cultures

Three independent TF1 erythroleukemia cell lines were transduced to an MOI=2 (~90% GFP^{+}) with WT SIX1 or EV controls. After 6 days in GM-CSF cultures cells were harvested and washed twice in PBS and lysed in 1ml of Qiazol (Qiagen). RNA was isolated using a RNeasy Mini Kit according to the manufacturer's protocol and quantified using a Nanodrop. Approximately 1-2ug of RNA per sample was shipped on dry ice to Arraystar for RNA sequencing (Arraystar Inc, Rockville, MD). After RNA integrity checks via agarose gel electrophoresis, mRNA was isolated from rRNA with NEBNext® Poly(A) mRNA Magnetic Isolation. The enriched mRNA was used for RNA-seq library preparation using KAPA Stranded RNA-Seq Library Prep Kit (Illumina). The samples were then fragmented and reverse transcribed into cDNA and ultimately adapted to undergo sequencing using Illumina-compatible sequencing adapters. PCR amplification was used to expand the sample and purified to create the final RNA-seq library. The completed libraries were qualified on Agilent 2100 Bioanalyzer for concentration, fragment size distribution between 400~600 bp, and adapter dimer contamination. The amount was determined by absolute quantification qPCR method. The barcoded libraries were mixed in equal amounts and used for sequencing on the instrument. Data was analyzed a previously reported using FPKM using the R packages StringTie and Ballgown (Pertea *et al*, Nat Biotech. 2015; Frazee *et al*, Nat. Biotech 2015). PCA and correlation analysis were based on gene expression level, Hierarchical Clustering, Gene Ontology, Pathway analysis, and volcano plots were performed with the differentially expressed genes in R. GSEA enrichment analysis was performed using the GSEA software (Broad Institute, v4.0.3).

3.4 <u>Discussion</u>

We have recently shown that SIX1 overexpression in primary HSPCs as well as TF1 cell cultures is capable of enhancing *in vitro* erythropoiesis as measured by erythroid-specific markers. Conversely, CRISPR/Cas9 knockout of SIX1 reduced their ability to undergo EPO-stimulated differentiation. Stimulation of erythropoiesis by SIX1 was found to be dependent on the presence of the master erythroid regulator, GATA1. RNA sequencing coupled with GSEA was used to determine the transcriptome of SIX1 in TF1 erythroid cell cultures. Genes known to be involved in and regulated by erythropoiesis were among the most significantly differentially expressed genes. This was not surprising given that the assays were performed in an erythroleukemic cell culture and our previous experiments have shown the effect of SIX1 overexpression in TF1 cell cultures (Chapter 2).

However, the specific enrichment of GATA1 binding sequences [A/TGATAA/T] as opposed to other erythroid transcription factors, such as KLF1 or SCL, suggests that there are selective functional interactions between SIX1 and GATA1, which when combined with our previous fly data highlights a potential conserved interaction. Interestingly these enriched genes include those known to be upregulated as well as those that are repressed (Ferreira et al, Mol Cell Bio 2005). And while many GATA1 genes were also regulated by SIX1, it is hard to determine whether these genes were directly related to the transcriptional action of SIX1 or if they change passively due to erythropoietic nature of the cells. In fact, after six days in culture with SIX1 overexpression the TF1 cells were substantially differentiated down the erythroid lineage. GSEA on only the significantly regulated genes showed an increase in heme metabolism, and cell cycle changes, which should be expected in cells undergoing erythropoiesis (Chuch *et al*, Curr Opin, Hematol 2012; Pop *et al*, PLoS Biol. 2010). It would be worthwhile in future studies to sequence additional samples in GATA1 knockout lines to see if SIX1 can regulate erythroid gene expression independently of GATA1, as is suggesting in GATA1 knockout studies with SIX1 overexpression regulating expression of CD235a/GYPA (Chapter 2, Figure 10). Or perhaps utilizing CRISPR/Cas9 to edit SIX1-binding sites in gene promoters to determine more easily those genes that are regulated by SIX1, and conversely mutating the GATA1 sites neighboring SIX1 binding sites to examine if SIX1 does indeed alter the transcriptional regulation of GATA1.

Overall, our results suggest that TF1 cell cultures differentiate down the erythroid cell lineage and produce erythroid-specific markers such as HBB, SLC4A1, and ALAS2 in response to SIX1 overexpression. Our findings that many GATA1-regulated proteins are differentially regulated in response to SIX1 supports our original hypothesis that SIX1 increases GATA1 levels and stimulates erythropoiesis. The fact that many key erythroid genes harbor both SIX1 and GATA1 sites, often in close proximity, strengthens our hypothesis that SIX1 and GATA1 act in concert to drive erythroid gene expression or repression. The overlap and proximity of these binding sites suggests and directs future experiments outlined in chapter 4, that there may be additional regulation of SIX1 expression that go beyond simply increasing GATA1 levels and decreasing GATA2 levels.

Chapter 4: SIX1 and SIX2 bind GATA1 and increase its transcriptional output

4.1 Introduction

SIX1 overexpression can enhance erythropoiesis in both $CD34^+$ hematopoietic stem and progenitor cells (HSPCs) as well as erythroleukemia cell lines. Though we have shown that GATA1 is a necessary factor for SIX1 to undergo erythropoiesis, our data do not reveal whether SIX1 and GATA1 functional interactions reflect physical interactions or are limited to indirect regulatory interactions, such as SIX1 stimulation of GATA1 levels. To begin to assess the mechanisms by which SIX1 may regulate erythropoiesis, we undertook two approaches simultaneously. First, we conducted global analysis to identify the proximal interactome of SIX to gain unbiased insight into how SIX1 is exerting its function in TF1 cell erythropoiesis. Secondly, we directly tested whether SIX1 could associate with GATA1 using proximal and coimmunoprecipitation assays.

Regulation of transcription factors is a potent mechanism to direct cell proliferation, survival, and differentiation (Reviewed in Spitz et al, Nat Rev Gen 2012). For example SIX1 itself modulates the expression of NFkB-targeted inflammatory genes (IL-1ß, IL-8, and CCL3) by binding RELA and RELB proteins in human fibroblasts, however in the presence of E3-ligase complexes SIX1 and SIX2 are degraded and fail to repress inflammatory signals (Liu *et al*, Nature 2019). To further understand the mechanism of SIX in erythropoiesis we need to identify its interacting members.

Gaining mechanistic insight is possible by determining proteins partners and associations, therefore we took advantage of the BioID technology to globally assess SIX interactions.

BioID relies on the activity of a mutant form promiscuous biotin-ligase, BirA*, which specifically ligates lysine residues of adjacent proteins within a 10nM radius (Figure 1). Proteins that reside within 10nM of the BirA* tag, including direct interactors, indirect interactors, as well as neighboring proteins will be covalently labeled with bioavailable biotin within the cell (free biotin denoted as gray circles and ligated biotin as black circles in Figure 1). The mutant form of BirA* (R118G) is preferentially utilized over the wild type due to its inability to bind DNA, self-associate, as well reducing its affinity for biotinyl-5'-AMP allowing for promiscuous labeling of neighboring peptides (Roux *et al*, J Cell Bio 2012; Kwon and Beckett, Protein Sci 2000). BioID has been widely used to identify interacting members of numerous proteins including IFGR1, c-MYC, and SOX2 in various cell types (Bareja et al, JBC 2018; Dingar et al, J. Proteomics 2014; Kim et al, Mol Cell Prot. 2017).

Here we chose to determine the proteins associated with SIX2 as opposed to the SIX1 interactome due to its consistently higher and robust erythroid phenotype in overexpression experiments using TF1 cells compared to SIX1 overexpression (Chapter 2 Fig1A-D). We created a fusion protein comprising BirA* fused to the amino-terminus of SIX2, creating a functional BirA*-SIX2 protein (referred to in this chapter as BioID-SIX2).

BioID identified GATA1 as a proximal interactor of SIX2 in TF1 erythroid cell cultures, extending our findings from previous chapters and hinting that GATA1 is a necessary cofactor of SIX in erythropoiesis. In collaboration with the Pandey lab at Johns Hopkins University, we employed the same BirA*-SIX2 construct in TF1 cell cultures to identify the global interactome of SIX2 by mass spectrometry analysis. Gene ontological analysis highlight diverse cell processes affected by SIX overexpression, suggesting that SIX exerts its function in erythropoiesis through multiple routes. We further show that both SIX1 and SIX2 are capable of binding GATA1 as demonstrated by coimmunoprecipitation experiments using MYC-tagged constructs in HEK293T cell lines. Additional interaction studies using a proximity ligation assay (PLA) designed to identify *in situ* interactions show that SIX1 and GATA1 are co-located in TF1 erythroid cell lines. These results coupled with our previous findings involving the necessity of GATA1 in SIX1-stimulated erythropoiesis (Chapter 2, Figure 10) suggest that SIX1 or SIX2 may stimulate erythroid differentiation by binding and stimulating GATA1 transcriptional output but may require additional cofactors to completely exert its function. Furthermore, these results are the first to highlight a physical and functional interaction between the two central developmental networks of PSEDN and GATA/FOG. The functional relevance of this interaction was revealed using a luciferase transcription assay in HEK293T cells where the co-transfection of either SIX1 or SIX2 along with GATA1 constructs increased luciferase expression by approximately 40-fold.

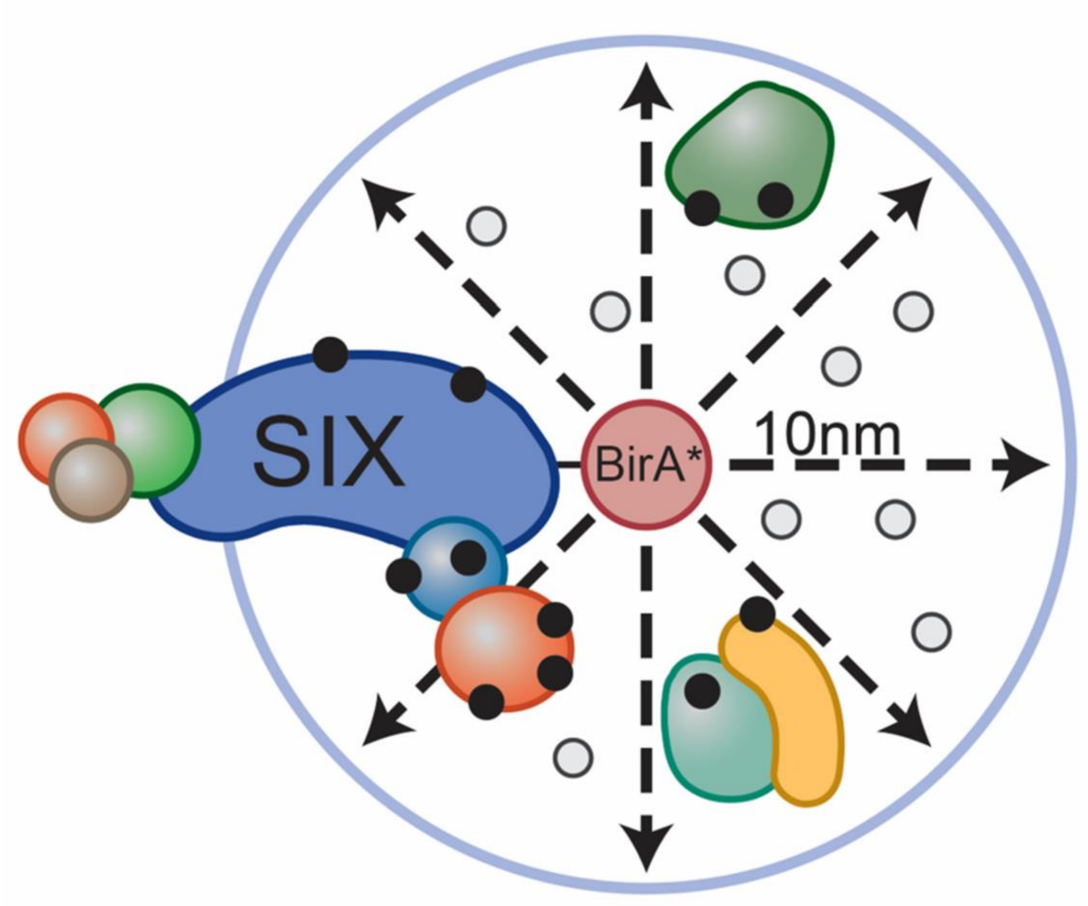

Figure 4.1. Schematic of the proximity labeling BioID system

The promiscuous biotin ligase, BirA*, is fused to SIX to create a fusion gene to determine interacting or neighboring proteins of SIX *in vitro*. With the addition of exogenous biotin and 24 hours incubation proteins that are within a 10nM diameter will be covalently labeled with the addition of biotin by BirA*. The proteins include those that are directly interacting, indirectly interacting, self-labeled, and proteins that fall within the distance limit of BirA*. Any proteins that fall outside of the range, though they may still be interacting, will not be biotinylated. Downstream analysis require that biotinylated proteins are enriched, usually with the use of streptavidin-coated agarose beads. Interacting proteins can then be identified by multiple methods such as Western blot or mass spectrometry. *Adapted from Varnaite and MacNeill, Proteomics 2016)*

4.2 Results

4.2.1 BioID identifies members of the SIX interactome

BioID can identify interacting protein interactions that occur within a 10nM range by biotinylating lysine residues on neighboring proteins (Roux *et al*, Curr Protoc Protein Sci, 2018). We fused the proximity biotin ligase, BirA*, in-frame to the C-terminus of our SIX2 protein (BirA*-SIX2) and cloned into our overexpression vector with a GFP marker to track levels of viral transduction. We first confirmed that the BirA* tag did not hinder the ability of SIX2 overexpression to drive expression of erythroid markers. Forced overexpression of BirA*-SIX2 for 6 days increased the expression of the erythroid-specific markers CD235a and CD71 and reduced the HSPC marker CD34 similar to wild-type SIX2 (Figure 2). TF1 cells were transduced to an MOI=2 (~90% GFP^+) for 5 days and subjected to overnight treatment with exogenous biotin. Cell pellets were collected the following day and sent to the Pandey lab at Johns Hopkins University for mass spectrometric identification. Table 1-2 list the labeled proteins discovered through BioID that were deemed to be the highest confidence peptides labeled in BirA*-SIX2 cell lysates when compared to BirA* controls. Proteins were deemed high confidence if they had three or more SIX2-BioID peptide spectral matches (PSMs) and had a two-fold increase in SIX2-BioID compared to BirA* controls. PSM refers to the total number of identified peptide sequences, including both unique and identical sequences, for any specific protein as identified by mass-spectrometry. Table 1 highlights the proteins that were only identified in the BirA*-SIX2 conditions and not in the BirA* controls, whereas Table 2 ranks PSMs by fold change (BirA*-SIX2 / BirA*). The cutoff criteria for Table 2 are that the proteins that have a minimum two-fold change of PSMs over BirA* controls with at least four total PSMs in BirA*-SIX2.

The interacting proteins include well-known SIX interactors such as EYA3, TLE3, TLE4, and our recently discovered SIX-interacting partner GATA1, which has been highlighted in yellow (Brugmann *et al*, Development 2004, Patrick *et al*, Nat. Struct. Mol. Biol. 2013). Preliminary experiments using a BirA*-SIX1 construct also showed that GATA1 was biotinylated more than BirA* controls (unpublished observation). Given that EYA3 was among the most-labeled proteins in the BirA*-SIX2 interactome we tested whether EYA3 interaction with SIX2 was necessary for erythroid gene expression in TF1 cells. Overexpression constructs of a mutant form of SIX that is known to disrupt EYA interaction (V17E) (Patrick et al, JBC 2009) reduced the ability of SIX2 to drive erythroid cell-surface marker expression in our TF1 cell cultures, suggesting that SIX drives erythropoiesis through interactions with several cofactors (Figure 3).

Interestingly, proteins that were proximal to SIX were enriched in DNA-binding transcription factors, and proteins involved in chromatin remodeling (SWI/SNF, NuRD complexes). Similarly, gene ontology (GO) analysis revealed nuclear protein complexes as well as nucleolus-specific proteins as SIX proximal interactors. Examining the enriched proteins in biological processes revealed regulation of nucleic acid biosynthetic processes (Table 3).

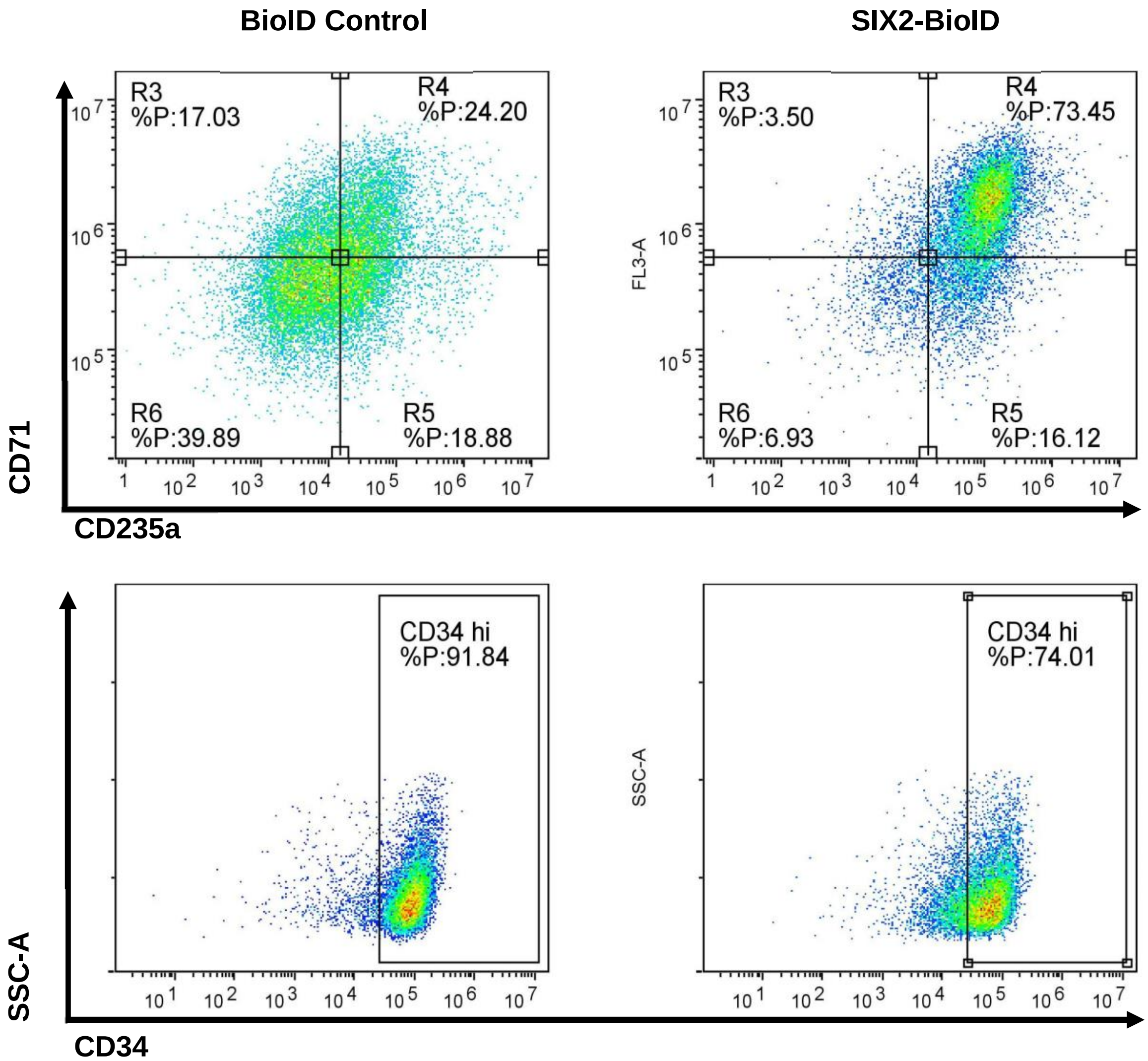

Figure 4.2. BioID tag does not inhibit SIX2-driven erythroid gene expression

Flow cytometry analysis of TF1 cell cultures 6 days post-transduction with BioID control or SIX2-BioID. Samples that were transduced with SIX2-BioID had increased CD71 and CD235a expression as compared to BioID controls. Additionally, the SIX2-BioID samples had reduced expression of the HSPC-marker CD34 compared to control cells.

Gene symbol	# PSMs BirA*	#PSMs BirA*-SIX2
SIX2	0	74
EYA3	0	22
ZNF529	0	18
ZNF430	0	18
ZNF644	0	17
HMG20B	0	15
BRD7	0	15
ZNF292	0	15
CDCA5	0	13
NSD2	0	13
ZNF552	0	12
FLI1	0	11
MSL1	0	10
ZNF544	0	10
FBRS	0	10
ZNF614	0	10
NFIX	0	10
MYBL2	0	9
CEBPB	0	8
SKI	0	8
TLE4	0	8
CUX1	0	8
RBM14	0	8
HMGB2	0	8
ATN1	0	7
ZNF670	0	7
ZNF28	0	7
RPP38	0	7
NFIA	0	7
CBX2	0	7
ZNF792	0	7
RPL13	0	7
HHEX	0	6
FBRSL1	0	6
ELF2	0	6
HELLS	0	6
ZNF507	0	6
GAR1	0	6
ZNF121	0	5
RNMT	0	5
RPP30	0	5
ZEB1	0	5
TRERF1	0	5
ZHX2	0	5

Gene symbol	# PSMs BirA*	#PSMs BirA*-SIX2
TET2	0	5
ZNF107	0	5
UTP11	0	5
TAF1B	0	5
INCENP	0	5
POGK	0	5
BRIX1	0	5
USP3	0	5
TCF3	0	4
FOXM1	0	4
MYB	0	4
DNTTIP1	0	4
PHF2	0	4
AES	0	4
NKAP	0	4
TCF3	0	4
HEMGN	0	4
GNL3L	0	4
MSL2	0	4
ZNF286A	0	4
ZNF675	0	4
ZNF567	0	4
INO80E	0	4
RBM15B	0	4
POP1	0	4
DDX24	0	4
ACTR5	0	4
C1orf131	0	4
HIRA	0	4
CBX4	0	4
BRD3	0	4
KMT2C	0	4
ZNF652	0	4
LENG1	0	4
ARID4A	0	4
RPL36	0	4
ANKRD11	0	4
PRPF6	0	4
YY1AP1	0	3
NFIC	0	3
CXXC5	0	3
ZKSCAN8	0	3
CHD6	0	3
ZNF24	0	3

Gene symbol	# PSMs BirA*	#PSMs BirA*-SIX2
VIRMA	0	3
ZNF628	0	3
ZNF765	0	3
FOXJ2	0	3
FOXJ3	0	3
ZNF91	0	3
ZNF570	0	3
ZNF136	0	3
C5orf24	0	3
SPTY2D1	0	3
ZNF432	0	3
ZNF668	0	3
RCOR3	0	3
XPC	0	3
HOMEZ	0	3
DPF1	0	3
DPF3	0	3
MAZ	0	3
KIF22	0	3
ZBED4	0	3
ZNF518A	0	3
BAZ1A	0	3
ZNF518B	0	3
ZNF217	0	3
LRRC42	0	3
RBBP8	0	3
CREBBP	0	3
PHC2	0	3
NAB1	0	3
PIAS2	0	3
NOL7	0	3
BRD9	0	3
KAT8	0	3
RBM4	0	3
POLL	0	3
MECOM	0	3
HJURP	0	3
EMG1	0	3
CASZ1	0	3
PELP1	0	3
SRSF5	0	3
NCOA6	0	3
HSF1	0	3
KIF20B	0	3

Table 4.1. BioID results of BirA*-SIX2 experiments with no BirA* control labeled peptides

Listed are proteins labeled by SIX2-BioID from most to least identified peptide specral matches (PSMs). In order to be considered high confidence samples had to have at a minimum 3 PSMs and contain no identified peptides in BirA* controls.

Gene symbol	# PSMs BirA*	#PSMs BirA*-SIX2
TLE3	3	154
WIZ	1	20
UBN2	1	17
NKRF	1	17
ELMSAN1	1	16
KIF23	1	16
ZNF100	1	16
ARID3A	2	31
ZNF267	1	15
LIN54	1	14
BEND3	1	13
RREB1	9	102
NSD3	1	10
BCORL1	1	10
UTP14A	3	29
RUNX1	2	19
MGA	5	46
TCF7L2	2	18
ZNF850	1	9
RAVER1	1	9
ZNF512B	2	17
KMT2D	3	24
PDS5B	2	16
VEZF1	1	8
HMG20A	1	8
PRR12	1	8
ELF1	1	8
ZNF254	1	8
TOP2B	1	8
ARID1A	3	23
BLM	2	15
IKZF3	3	21
TAF1	2	14
NCOR2	1	7
ZNF761	1	7
ZNF141	1	7
PPIL2	1	7
PINX1	1	7
POGZ	1	7
BBX	1	7
JMJD1C	4	27
IKZF1	8	53
ZFPM1	4	26
BCOR	5	31
ZEB2	3	18
KMT2A	3	18
TCP1	3	18
ZNF384	2	12
PBRM1	2	12
PHF8	1	6
CCDC71L	1	6
TCF12	1	6
ZNF480	1	6
DDX27	1	6
ZNF486	1	6
RAI1	1	6
CTNNBL1	6	33
PHIP	4	22
EHMT1	5	27
LRIF1	3	16
TFAP4	5	26
TAL1	6	30
SMARCC1	5	25
PHF14	1	5
ZNF8	1	5
CXXC1	1	5
TRIM33	1	5
ELAVL1	1	5
NSD1	4	18
SMARCC2	2	9
GPATCH4	4	17
GSE1	9	38
GNL2	5	21
TCF20	14	58
ZNF680	2	8
SAP130	2	8
HOXB4	1	4
ZNF585B	1	4
SMCHD1	1	4
FOXK2	1	4
ZNF574	1	4

Gene symbol	# PSMs BirA*	#PSMs BirA*-SIX2
ZNF146	1	4
RRP7A	1	4
DMAP1	1	4
ZNF492	1	4
DPF2	1	4
NEPRO	1	4
SUMO3	1	4
MTHFS	1	4
SSRP1	1	4
BRD4	3	11
PPHLN1	3	11
ZNF148	8	29
ARID3B	10	36
NR2C2	4	14
PDCD11	4	14
ZNF92	2	7
WRN	2	7
SYF2	2	7
RAD18	2	7
MORC2	2	7
ZNF280D	2	7
MSH6	2	7
BAZ1B	32	110
ZNF629	3	10
USP36	3	10
XRCC6	11	36
UBTF	8	26
ATF1	5	16
MECP2	6	19
TOP2A	8	25
ZNF609	13	39
UBTF	9	27
ZNF592	9	27
EBNA1BP2	5	15
GATA1	4	12
NOC3L	3	9
MORC3	3	9
SRFBP1	2	6
TDP2	2	6
RLF	2	6
NGDN	2	6
ZNF736	2	6
STAG2	2	6
CBX8	2	6
SMARCD1	2	6
GLTSCR1	2	6
CIR1	2	6
IRF2BP2	2	6
ATRX	2	6
ZNF174	1	3
RTFDC1	1	3
CBFA2T3	1	3
INO80D	1	3
SENP5	1	3
ZNF66	1	3
CENPV	1	3
JADE3	1	3
ZNF468	1	3
QSER1	1	3
CCNT2	1	3
KDM1A	1	3
TPX2	25	73
NAT10	29	84
NFRKB	7	20
CHD7	55	157
NCOR1	25	70
ZNF281	8	22
RRP12	4	11
ZHX3	4	11
SAP30BP	18	49
TMPO	15	40
DIDO1	6	16
ZNF845	3	8
COIL	3	8
WDR3	3	8
BUD13	3	8
YY1	8	21
RRS1	29	76
ZNF724	10	26
EMSY	5	13
NOL8	5	13

Gene symbol	# PSMs BirA*	#PSMs BirA*-SIX2
PARP1	24	62
CHD1	7	18
NIPBL	14	35
ZMYM2	6	15
JUNB	4	10
ZNF273	4	10
ZMYM3	2	5
HBB	2	5
RNF2	2	5
HIVEP1	2	5
MCM10	2	5
ZBTB21	2	5
SUPT5H	2	5
INTS12	2	5
FBL	2	5
INTS1	2	5
BRF1	2	5
WBP4	2	5
SIX4	2	5
LDB1	5	12
RNF169	5	12
GATAD2A	16	38
CHD8	8	19
RPS25	9	21
GABPA	3	7
ZNF12	3	7
ZNF787	3	7
CCDC94	3	7
MKI67	134	311
ZNF512	10	23
KIF4A	10	23
GTF2I	14	32
PHF6	7	16
RNPS1	7	16
ZMYND8	31	70
ARID2	16	36
SMARCA2	4	9
NCOA2	4	9
MIS18BP1	4	9
IWS1	4	9
CHD4	24	53
PHF10	5	11
HIST1H1C	5	11
CDCA2	6	13
REXO4	6	13
SUGP2	13	28
CCDC174	8	17
ORC2	16	32
RALY	8	16
KDM2A	7	14
PPM1G	6	12
NCBP3	6	12
HMGB3	5	10
ZMAT2	5	10
BMS1	5	10
SBNO1	5	10
ZNF195	4	8
BPTF	4	8
TAF3	4	8
GNL3	4	8
SMARCE1	4	8
RNF113A	4	8
MTA2	4	8
DDX5	4	8
ZBTB7A	3	6
NR2C1	3	6
GATAD2B	3	6
SENP6	3	6
HEXIM1	3	6
CCNK	3	6
NOP16	3	6
ETV6	2	4
UNG	2	4
ZNF184	2	4
PARN	2	4
USP22	2	4
NOC2L	2	4
RBM19	2	4
NCOA3	2	4
MBIP	2	4
XAB2	2	4
RPL15	2	4

Table 4.2. BioID results of BirA*-SIX2 experiments with some BirA* control labeled peptides

Listed are proteins labeled by SIX2-BioID from most to least identified peptide specral matches (PSMs). In order to be considered high confidence samples had to have at a minimum 4 PSMs and a greater than 2-fold change (BirA*-SIX2 / BirA*).

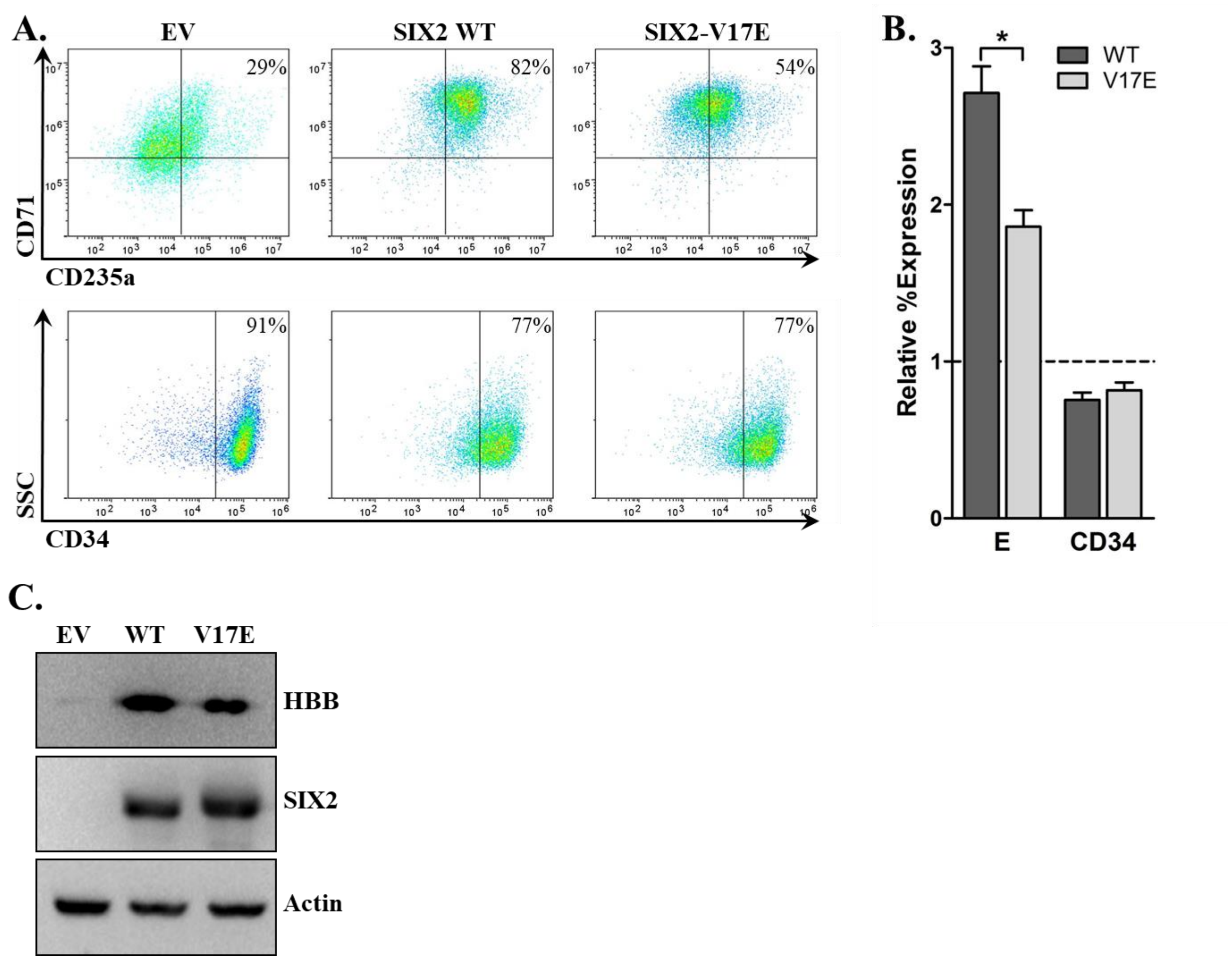

Figure 4.3. SIX2 mutant (V17E) OE has reduced erythroid gene expression

(A) Representative flow cytometry analysis of TF1 cells transduced to an MOI=2 (GFP+ 90%) six days post-transduction. (EV = empty vector). (B) Bar graph showing results of three independent experiments. Bars represent mean fluorescence intensity (MFI) +/- s.e.m. (* $p < 0.05$). (C) Representative Western blot of EV, SIX2, or SIX2 V17E expression TF1 cell lines probed for hemoglobin (HBB), SIX2, or SIX2 V17E mutant

Category	Name	BioID list	GO Term	
Molecular Function	DNA-binding transcription factor activity, RNA polymerase II-specific	59	1615	2.98E-29
	chromatin binding	26	629	1.10E-13
	regulatory region nucleic acid binding	31	1036	1.82E-12
	transcription coregulator activity	23	597	1.44E-11
	transcription factor binding	25	750	3.77E-11
	sequence-specific DNA binding	30	1199	3.60E-10
	histone binding	13	203	1.33E-09
	transcription regulatory region sequence-specific DNA binding	25	898	1.62E-09
	double-stranded DNA binding	27	1048	1.74E-09
	sequence-specific double-stranded DNA binding	25	948	4.84E-09
Biological Process	negative regulation of nucleic acid-templated transcription	47	1370	2.33E-22
	negative regulation of RNA biosynthetic process	47	1372	2.47E-22
	negative regulation of transcription, DNA-templated	45	1316	2.80E-21
	negative regulation of RNA metabolic process	47	1467	4.00E-21
	negative regulation of nucleobase-containing compound metabolic process	48	1573	1.01E-20
	negative regulation of macromolecule biosynthetic process	49	1700	3.83E-20
	negative regulation of cellular macromolecule biosynthetic process	47	1605	1.59E-19
	negative regulation of cellular biosynthetic process	49	1761	1.70E-19
	negative regulation of biosynthetic process	49	1797	3.96E-19
	chromosome organization	41	1287	3.54E-18
Cellular Component	nuclear chromatin	72	1924	9.23E-40
	nucleolus	32	1495	7.74E-10
	catalytic complex	31	1536	6.17E-09
	nuclear body	21	799	3.34E-08
	SWI/SNF superfamily-type complex	8	79	3.87E-08
	chromosome, centromeric region	11	202	7.04E-08
	ATPase complex	8	111	5.53E-07
	methyltransferase complex	8	121	1.07E-06
	MSL complex	3	5	2.50E-06
	chromosomal region	12	362	3.55E-06

Table 4.3. Gene ontology analysis of SIX2-interacting proteins

Gene ontology analysis of SIX2-interacting proteins were identified using ToppFun (Chen *et al*, Nucl Acids Res 2009). The three major categories (molecular function, cellular component, and biological process) suggest that SIX2 may be involved in regulating DNA. Additionally, SIX2 may be involved in regulating chromatin structure through its interactions with chromatin-modifiers and regulators.

4.2.2 <u>SIX-BioID biotinylates GATA1 in TF1 cell cultures</u>

Given that GATA1 was identified in our SIX2-interactome (as well as our preliminary experiments using BirA*-SIX1) we tested whether we could replicate this in our cell cultures. Using a BirA*-SIX2 fusion protein we identified GATA1 as a member of the SIX proximal interactome. TF1 cells were transduced to an MOI=2 with either BirA* control or BirA*-SIX2 constructs and incubated for 6 days before lysis. Biotin-labeled proteins were enriched using streptavidin-coated agarose beads. Both input controls and enriched biotin samples show biotinylated proteins in the whole cell lysates of transduced TF1 cells (Figure 4A). BirA* alone samples show a heavy enrichment of non-specific biotin-labeled proteins compared to BirA*-SIX2 constructs (Figure 4A, lanes 3 vs. 4). These streptavidin-enriched proteins were probed for the presence of GATA1 (Figure 4B). Both BirA* controls and BirA*-SIX2 samples showed expression of GATA1, however only GATA1 was enriched in the BirA*-SIX2 conditions (Figure 4B. Lanes 3 vs. 4). Using a BirA*-specific antibody we observed that both BirA* and BirA*-SIX2 constructs were self-biotinylated as expected (Figure 4B).

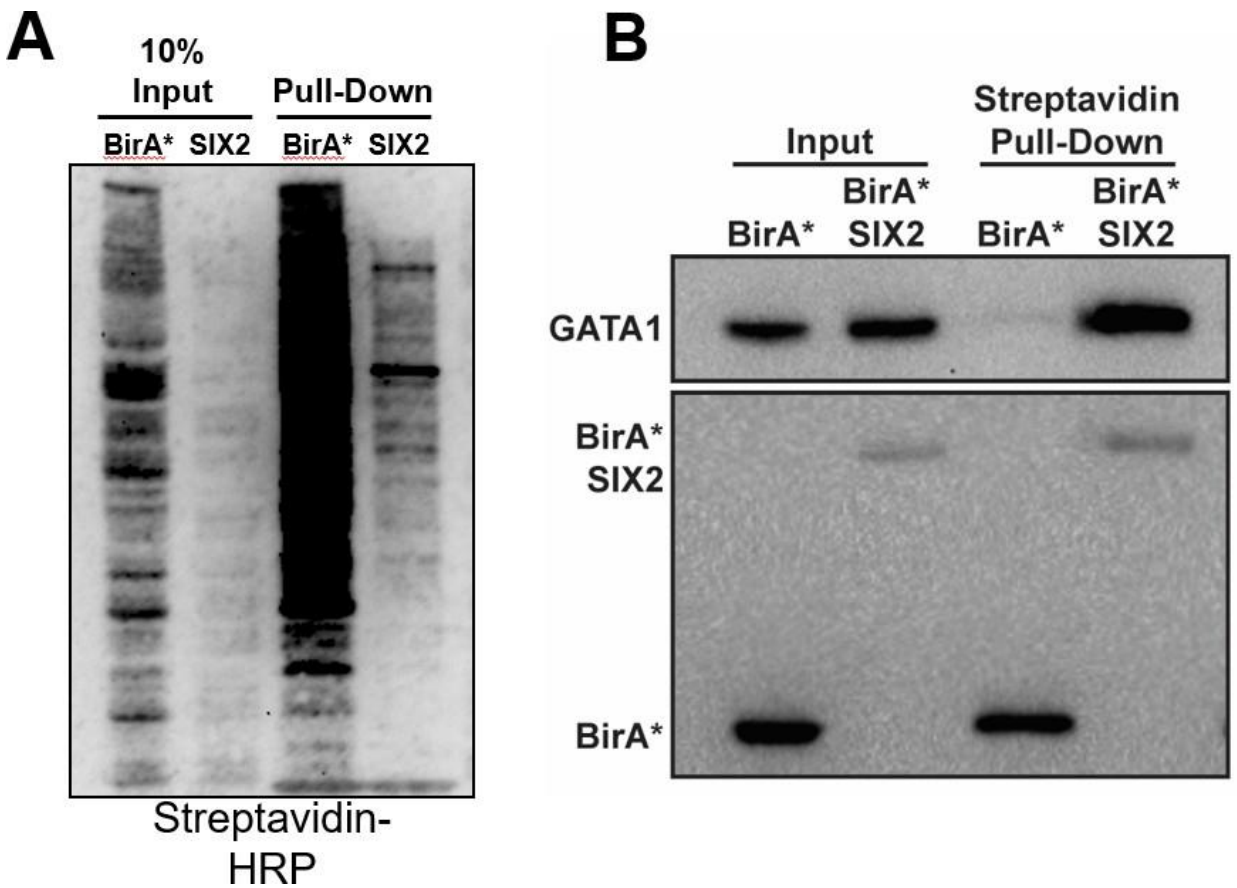

Figure 4.4. GATA1 is a proximal interaction of SIX2

(A) TF1 cells were transduced to an MOI =2 (~90% GFP^{+}) with either BirA* control or BirA*-SIX2 constructs before lysis and streptavidin enrichments. Western blot using streptavidin-HRP antibodies shows biotinylated proteins in control or SIX2-BioID TF1 cell lines. Lanes 1-2 are whole lysates (10% of total input) while lanes 3-4 show enrichment of labeled peptides post-streptavidin enrichment. (B) Western blot analysis of total cell extracts (input) vs streptavidin pull-downs generated from biotin labeled TF1 cells transduced with BirA* vs SIX2-BioID probed with anti-GATA1, or anti-BirA as indicated.

4.2.3 SIX1/2 coimmunoprecipitates GATA1 in HEK293T cells

Our first attempts to determine if SIX could bind GATA1 were done using TF1 cell lysates, as we wanted to confirm they interact in the context of an erythroid cell. However, despite several rounds of nuclear enrichment, various buffer compositions, and different incubation times, we were unable to precipitate any endogenous protein complexes, including the well-studied GATA1/FOG1 interaction (Rodriguez *et al*, EMBO J 2005). Therefore, we moved our experiment to the easily transfectable HEK293T cells lines. We fused an N-terminal MYC tag to both SIX1 and SIX2 proteins to test the ability of both isoforms to immunoprecipitate GATA1. We took advantage of commercially available anti-MYC coated magnetic beads that reduce both non-specific binding and sample loss during critical wash and purification steps.

MYC-SIX1 or MYC-SIX2 overexpressing constructs were co-transfected with either GATA1 or an empty vector (EV) control using the same amount of DNA in all conditions. Cells were then lysed in coimmunoprecipitation buffer and incubated with MYC-tagged magnetic beads for one hour. After enrichment of MYC-tagged, samples were washed, boiled, and subjected to Western blot. Anti-GATA1 antibodies detected enrichment of GATA1 specifically in MYC-SIX1 or MYC-SIX2 conditions and not in the EV control (Figure 5). Anti-MYC antibodies were used to confirm pulldown of MYC-tagged proteins.

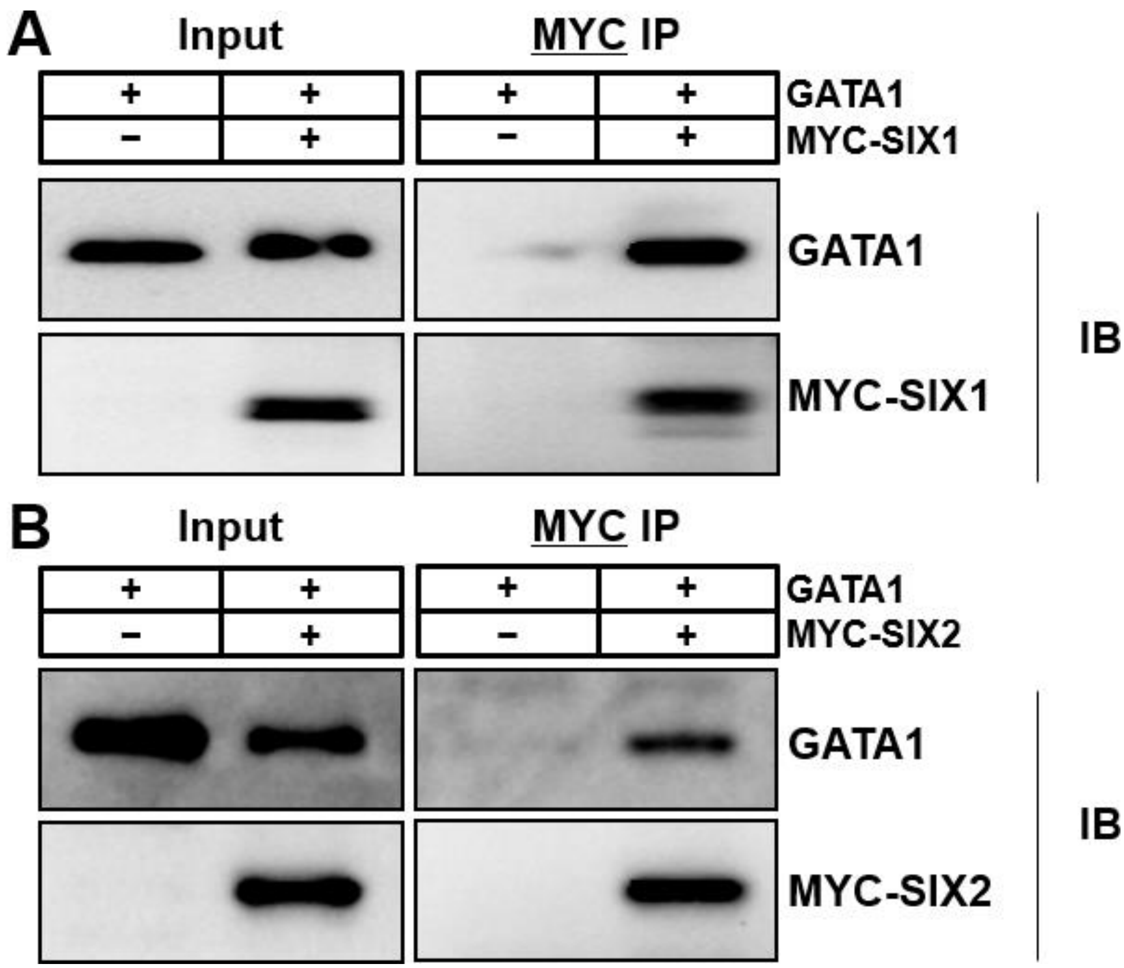

Figure 4.5. SIX1 and SIX2 coimmunoprecipitate GATA1 in HEK293T cells

(A-B) Western blot analysis of anti-MYC immunoprecipitation assays conducted on total cell lysates generated from HEK293T cells co-transfected with GATA1 and MYC- SIX1/MYC-SIX2 vs empty vector. Total cell extract (input) vs anti-MYC-IP samples probed with anti-GATA1 or anti-MYC antibodies.

4.2.4 Proximity ligation assay shows endogenous interaction between GATA1 and SIX1

We have demonstrated that SIX1 and GAT1 can bind in cell cultures but have not shown their association in erythroid TF1 cells. To determine their interaction in endogenous TF1 erythroid cultures we used the Duolink® proximity ligation assay. This technique is used to visualize, by either microscopy or flow cytometry, the proximal interaction between adjacent proteins within a theoretical range of 40nM. The proximal interaction allows for proteins of interest to be targeted using primary antibodies produced from two different species (i.e. rabbit and mouse). Oligonucleotides coupled to secondary antibodies against the host target-protein antibodies are ligated via addition of DNA ligation and amplified intracellularly by rolling circle amplification by a DNA polymerase. This new DNA structure is then visualized with the aid of a fluorescently labeled antisense oligonucleotide probe. Either wild-type or SIX1 knockout TF1 cells were fixed to glass slides using a cytospin and fixed and permeabilized according to the manufacturer's instructions (Sigma-Aldrich). Overnight incubation with rabbit anti-SIX1and mouse anti-GATA1 antibodies showed a positive interaction signal in the wild-type (WT) TF1 cells in two independent experiments. This interaction was lost in the SIX1 knockout (KO) TF1 cells (Figure 6).

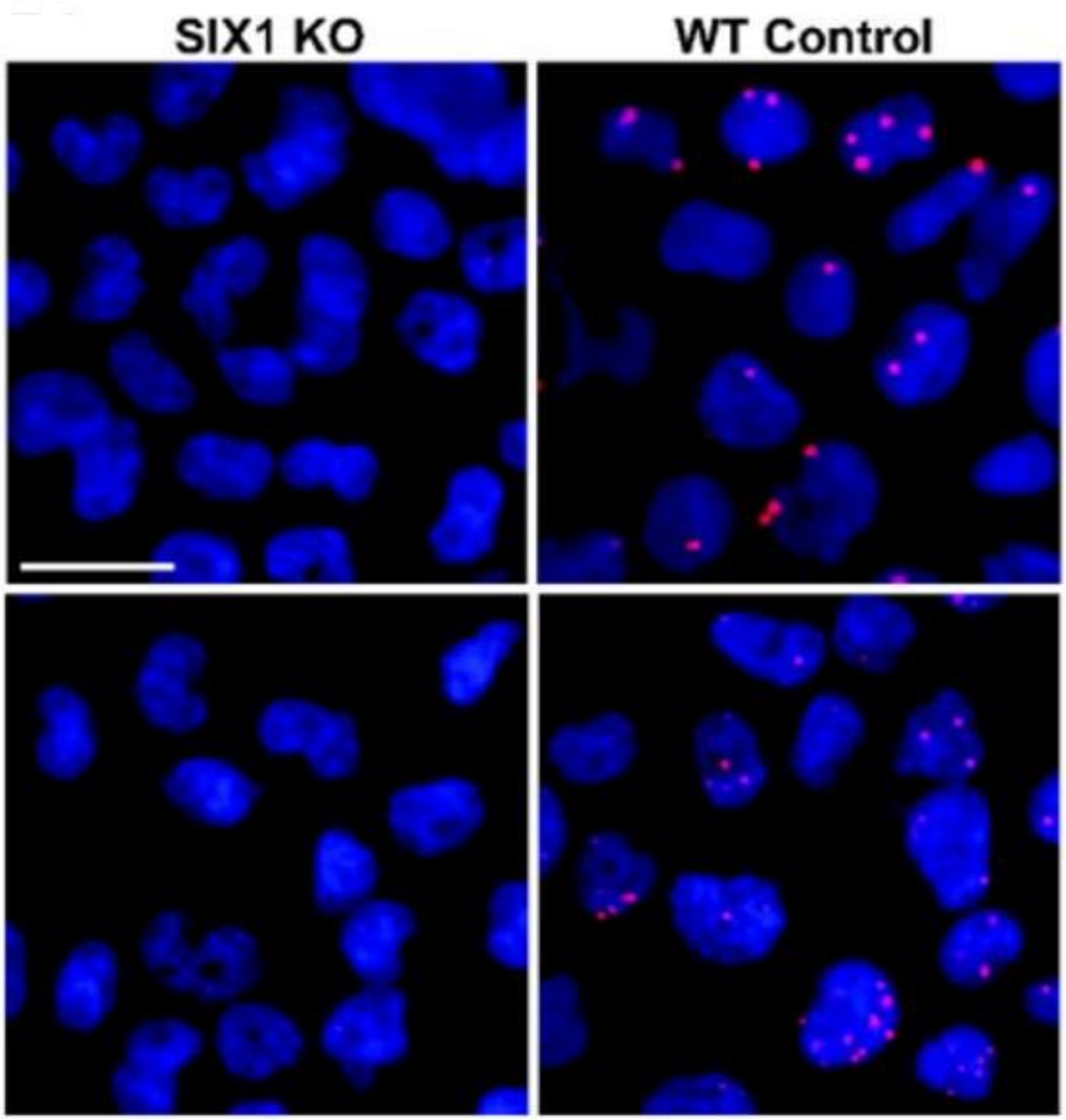

Figure 4.6. Proximal ligation assays reveal a SIX1 and GATA1 interaction

PLA assays testing intracellular colocation of SIX1 and GATA1. Primary antibodies targets SIX1 and GATA1 detect signal in wild-type (WT) TF1 cell lines (red dots indicate proximal association). No signals were observed in SIX1 knockout (KO) compared to control cells with endogenous SIX1. Scale bar: 25 µm. Images from two independent experiments shown.

4.2.5 SIX stimulates GATA transcriptional output

We have shown an interaction between SIX and GATA, but it is unknown what the functional outcome of this physical interaction is. To determine the consequences of SIX-GATA1 interaction on GATA1 transcriptional output, HEK293T cells were co-transfected with a minimal GATA-dependent reporter gene containing 3 copies of the GATA binding site ([A/T]GATA[A/G]) (Hasegawa *et al*, Mol Cell Biol. 2016) in the presence or absence of GATA1 and SIX1 or SIX2 expression constructs. Enforced expression of GATA1 alone resulted in a 9-fold stimulation of reporter gene activity, whereas enforced expression of either SIX1 or SIX2 alone resulted in a 2-3-fold stimulation of reporter activity. Co-expression of GATA1 with either SIX1 or SIX2, however, resulted in an approximate 40-fold stimulation of activity (Figure 7). Collectively, these findings suggest that SIX1 and SIX2 can bind to GATA1 protein complexes and increase GATA1-mediated transcription (Figure 7).

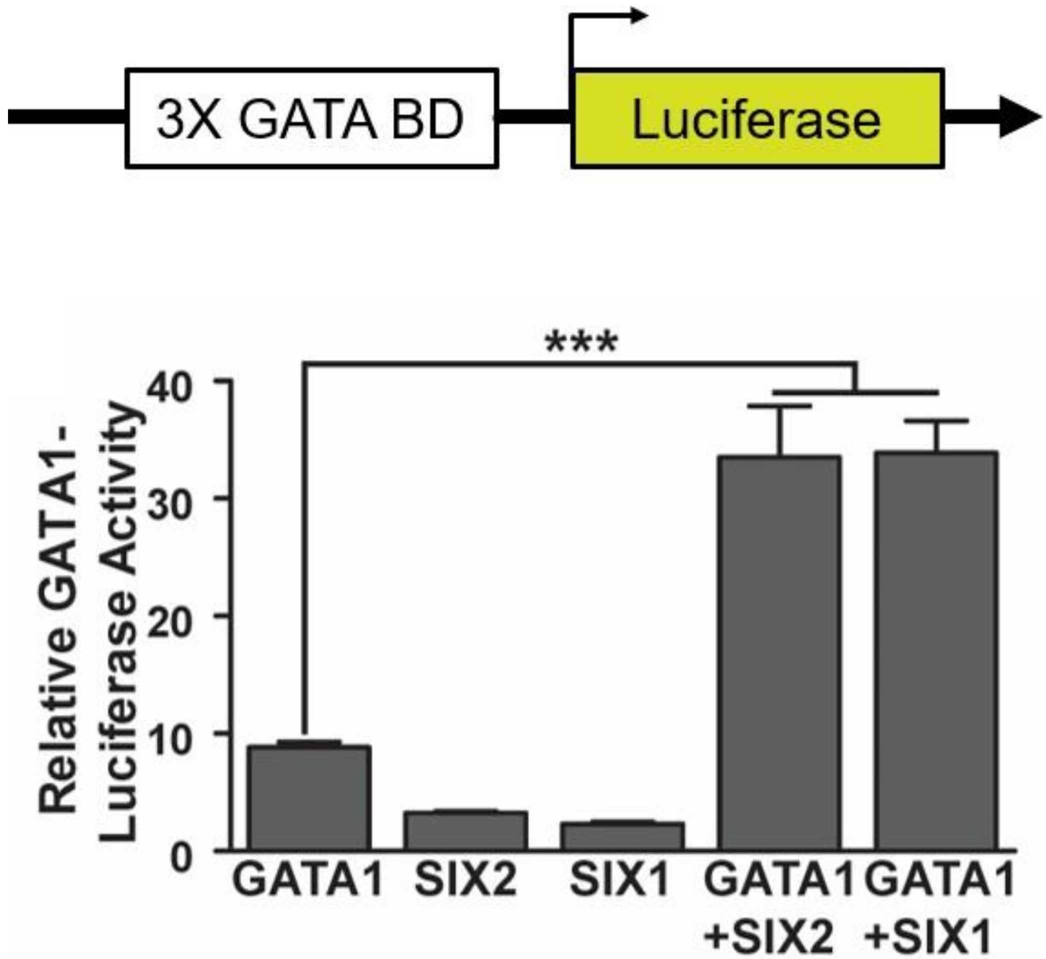

Figure 4.7. SIX1 stimulates GATA1 transcriptional output in luciferase assays

HEK293T cells were co-transfected with GATA1-luciferase reporter plasmid, SV40- Renilla luciferase and empty vector, SIX1, SIX2, or GATA1 expression plasmids. Firefly luciferase activity was normalized to Renilla luciferase activity. EV transfected cultures were set to 1. Data plotted are representative set of quadruplicates with SD indicated (***$p < 0.001$).

4.3 <u>Materials and Methods</u>

4.3.1 Cell culture and transduction

TF1 human erythroleukemia cells (CRL-2003; American type Culture Collection) were cultured as previously described (119) except recombinant human GM-CSF was from TONBO Bioscience. HEK293T cells (American Type Culture Collection) were cultured in DMEM media (CellGro) containing 10% FBS. Lentivirus transductions were conducted in the presence of 8ng/ml polybrene.

4.3.2 BioID and mass spectrometry

Biotin ligase BirA* versus SIX2-BioID lentivirus was transduced into TF1 cells (MOI=2; ~90% GFP+). Cultures were expanded for 3 days, then incubated in 50 µM biotin 24 h prior to harvesting total cell lysate in RIPA buffer containing protease inhibitors. Lysate (200 µg) was incubated with streptavidin-coated agarose beads (Sigma-Aldrich) at 4°C, washed three times prior to elution by boiling in Laemmli buffer (Sigma-Aldrich) followed by western blot analysis with streptaviding-HRP (Sigma-Aldrich), anti-BirA (59938, Novus Biologicals; 1:500) or anti-GATA1 (sc-266, Santa Cruz Biotechnology; 1:200) antibodies.

To identify the interactome of SIX2-BioID*, TF1 cells were prepared as previously described. However, cell pellets were frozen at -80° Celsius and sent to the former lab of Dr. Akhilesh Pandey for analysis in collaboration with Johns Hopkins University. Cell pellets were lysed in BioID lysis buffer (50mM Tris, pH 7.5, 500mM NaCl, 0.4% SDS, 2% Triton X-100, with Halt protease inhibitors (Thermo Fisher), followed by three rounds of sonication. After lysis, samples were cleared at 16,000*g* and protein concentrations were quantified using a bicinchoninic acid (BCA) assay. Ten milligrams of protein was then incubated with 200uls of high capacity biotin-specific beads (NeutrAvidin, Thermo Fisher) overnight at 4° Celsius. After incubation, the beads were washing three times with 2% SDS/50mM Tris, three times with BioID lysis buffer, three times with 50mM Tris, and three times with 50mM triethylammonium bicarbonate (TEABC). After biotin-enrichment and several rounds of washing, peptides were digested "on-bead" to generate peptide fragments using sequencing-grade 3ug trypsin overnight at 37° Celsius. The next day the supernatant was collected and the beads once again washed with 50uL of 50mM TEACB until a collected volume reached 450uLs. This collection was acidified to 1% trifluoroacetic acid, centrifuged at 18,000*g*, and desalted using a reverse-phase C_{18} column. Specific mass-spectrometry settings and analysis are outlined in the publication from our collaborators entitled "BioSITe: A Method for Direct Detection and Quantitation of Site-specific Biotinylation" (Kim *et al*, J Prot Res 2012).

4.3.3 <u>Coimmunoprecipitation Assay</u>

HEK293T cells cultured in 6-well dishes were transfected with a total of 2μg DNA consisting of 1μg pTJK482 (GATA1) with 1μg pTJK438 (3xMYC-SIX1), pTJK477 (3xMYC-SIX2) or pWCC43 empty vector (EV) to maintain equal total DNA quantity in all transfections. Cells were lysed and protein quantitated by Bradford assay. For immunoprecipitation, 500μl protein lysates (1mg/ml) were incubated overnight in the presence of 20μl pre-washed anti-Myc magnetic beads (88842, ThermoFisher Scientific). Samples were washed three times with Co-IP buffer (25uM Tris-HCl pH 7.4, 150uM NaCl, 1mM EDTA, 1% NP40, 3% Glycerol, Protease Inhibitors) and once with PBS before elution in Laemmli buffer (SigmaAldrich) and analysis by Western blot.

4.3.4 Proximity Ligation Assay (PLA)

WT or SIX1 KO TF1 were resuspended in PBS at a concentration of $2e^6$/ml, 100uls were centrifuged in a Cytospin3 at 200rpm for 5 minutes. Cells were dried at room temperature prior to fixation in 4% paraformaldehyde for 10 minutes. Slides were then washed 3 times in PBS before a 10-minute incubation in permeabilization buffer (50mM NaCl, 3mM $MgCl_2$, 10mM HEPES, 200mM sucrose, 0.5% Triton X-100). After washing cells were incubated in a humidity chamber at 37° for 1 hour in PLA blocking buffer. The PLA assay (Millipore Sigma) was performed according to manufacturer's instructions using primary antibodies against SIX1 (1:50, 12891S, Cell Signaling) and GATA1 (1:50, 60011-1, ProteinTech). Cells were imaged on using a (Zeiss Axioplan microscope) with appropriate filters for DAPI and PLA signal using 40X magnification and processed using ImageJ software.

4.3.5 Luciferase assays

HEK293T cells cultured in 24-well dishes were transfected using Lipofectamine 2000 (ThermoFisher Scientific). All wells were transfected with equal amounts of total DNA, comprised of pGL3 GATA Luc, gift from Licio Collavin and Giannino Del Sal (Addgene plasmid #85695)(102), pRL-SV40 (SV40-Renilla luciferase; Promega), pTJK482 (GATA1), pTJK422 (SIX2), pTJK299 (SIX1) or WCC43 (EV). For each condition, 3-4 independent wells were transfected. Two days post-transfection, cells were harvested and luciferase quantitated using dual luciferase reporter assays (Promega) according to manufacturer's instructions on a Perkin Elmer Victor™ X3 Multilabel Reader. Firefly luciferase was normalized to Renilla luciferase values with entire experiment conducted in triplicate.

4.4 Discussion

Here we have shown novel interactions between two well-conserved developmental complexes, namely the PSEDN and GATA/FOG complexes mediated by SIX1, SIX2, and GATA1. In addition to their ability to bind one another – as shown via multiple interaction experiments including proximal labeling assays, coimmunoprecipitations, and proximal ligation assays – the physical interaction results in stimulation of GATA1-mediated transcription of a minimal GATA-dependent reporter gene. The question still remains whether or not this interaction is direct or is mediated through binding partners orchestrated via a multiprotein complex. Our BioID identified proteins that are known to interact with GATA1 such as FOG1 (ZFPM1), LDB1, FLI1, and SPI1 (PU.1) (Eisbacher et al, Mol Cell Bio 2003; Wilkinson-White et al, Protein Sci 2015; Krivega et al, Genes Dev 2014). Future studies should test whether the interaction between SIX and GATA factors rely on the presence of these additional cofactors. CRISPR/Cas9 knockout lines for GATA1-interacting factors could further test if SIX1 or SIX2 could function in the absence of these newly identified, and erythropoietically important, factors.

We can hypothesize which domain might be responsible for interaction with GATA1 given what is known about the conservation of two key structural domains located with the SIX-family of proteins. SIX1 and SIX2 proteins contain two conserved functional domains, the homeobox domain (HD) and six-domain (SD) (Chapter 1, Figure 3). The HD mediates DNA-binding at MEF3 sites on chromatin and the SD is required for protein-protein interactions (Patrick et al, Nat Struct Mol Biol 2013). Function variation between the different members of the SIX familes lies within the N- and C-terminal domains. Given their conservation, we would expect the necessary domain for SIX-

GATA1 interaction lies within the conserved HD or SD. It is entirely possible that since the SD and HD are moderately-well conserved in all six SIX protein family members that additional SIX proteins will interact with GATA1 (or even other members of the GATA family if the conserved zinc finger domains are needed for GATA interactions). In fact, studies from our lab have shown that multiple members of the SIX family are capable of stimulating transcriptional output of multiple GATA family members, not solely GATA1, using our 3X GATA-binding domain luciferase construct (unpublished observation). This suggests that the ability of SIX and GATA family members to physically and functionally interact may be a common aspect between these complexes and are likely to be important during development. Further mutational analysis is needed to identify the necessary domains and residues required.

This is interaction would be interesting to study not only in hematopoiesis, but also in development of other cell lineages. Expression of SIX proteins is implicated in cancer metastasis as well as conferring poor prognosis to patients (Kingsbury *et al*, Adv Cancer Res 2019). Additionally, GATA proteins have been implicated in many of the same phenotypes seen with SIX. For example GATA family of proteins expressed in breast cancer, neuroblastoma, and pancreatic cancers, leads to increased metastasis and cancer cell proliferation similar to the phenotype seen with cancers that express SIX proteins (Li *et al*, Oncotarget 2015; Gulbinas *et al*, J Histochem Cytochem 2006, Molenaar *et al*, Oncogene 2010). Identifying the repertoire of protein interactions between these two families will reveal key mechanisms that drive normal lineage-specific development and malignancies, as specific SIX-GATA interactions may fine-tune GATA output.

Additionally, we sought to identify further interactions necessary for SIX to impart its erythroid-driving phenotype using BioID technology. BioID has successfully been used to identify novel interacting factors necessary for protein interactions. For example, in squamous cell carcinoma (SQCC) lung cancers—the second most common type of lung cancer—copy numbers gains of SOX2 promote cancer stem cell growth and increase proliferation (Kim *et al*, Mol Cell Prot. 2017). The mechanism of SOX2 was found to be dependent on its interaction with EP300 and hypothesize that targeting this interaction may prove to be a novel treatment for SQCC. In our system we identified a novel SIX-erythroid interactome. While the list contains known SIX-interactions such as EYA and TLE, it contains a list of potential interactions that may be necessary for SIX to function in erythropoiesis. Future studies should aim to CRISPR-engineer cells lines that lack genes of interest to see if SIX overexpression continues to function in the absence of BioID-identified interactors. Additional experiments should be performed to test the ability of SIX to interact with well-known cancer drivers that were identified via BioID. Several well-known drivers of cancers (not strictly limited to leukemia) including FLI1, RUNX1, NPM1, and ERG were identified as potential SIX-interactors (Ben-David *et al*, Genes Dev 1991; Sood *et al*, Blood 2017; Brunetti *et al*, Cancer Cell 2018; Adamo and Ladomery, Oncogene 2015). Future studies should examine whether the interaction between SIX and these oncogenes is necessary for carcinogenesis and may provide future druggable targets. BioID in SIX-erythroid cultures also identified key pathways enriched via gene ontology analysis. Interestingly among them were several pathways identified as chromatin organizers including the SWI/SNF superfamily-type complexes. SWI/SNF chromatin remodeling is critical in normal blood cell development and mice that lack members of SWI/SNF complexes die at a midgestational stage due to anemia (Griffin *et al*, Development 2008). In mammalian inner-ear development SIX1 along with EYA1 recruit the SWI/SNF chromatin-remodeling complex to mediate gene transcription of Neurog1 and Neurod1 (Ahmed *et al*, Development 2012). Future studies will determine if SIX is required for SWI/SNF remodeling complexes to exert their function during mammalian erythropoiesis.

In summation we have identified a novel collision between two central developmental networks, the PSEDN and GATA complexes. Beyond our physical and functional interaction studies outlined in this chapter, SIX proteins stimulate GATA1 levels and drive transcription of TF1 erythroleukemia cells and primary hematopoietic stem-progenitor cells (HSPCs). SIX1 knockout cells had reduced differentiation capacity in response to erythropoietin (EPO). This functional outcome of SIX1 overexpression was shown to be dependent on the presence of GATA1 and demonstrated by CRISPR/Cas9 knockout studies. Figure 8 shows the working model thus far for our studies in erythropoiesis, though studies in other model systems should be pursued. This has implications not only in development and homeostasis but may also serve for future therapeutic targets in cancers and diverse pathologies.

This thesis was a culmination of several years of difficult yet rewarding work, including late nights, long weekends, and countless failed experiments. All of this has helped me understand what science truly is and that at the end of each project you have created and discovered something that was once completely unknown. I hope that this work will have a long-lasting impact on the understanding human development and on the treatment of diseases.

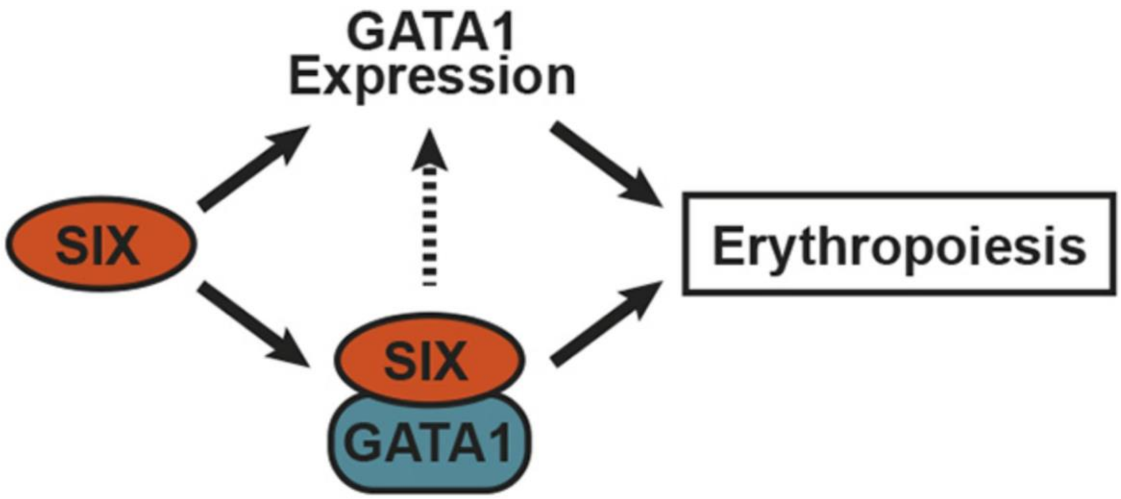

Figure 4.8. Overall model for the SIX-stimulation of erythroid differentiation

This thesis has outlined the ability of SIX-family members to modulate GATA1 expression levels and activity. SIX overexpression drives expression of GATA1 and effects downstream erythroid genes such as HBB, GATA2, CD235a, CD71, and heme- synthesis genes. SIX proteins are capable of binding GATA1 which has a functional outcome of stimulating GATA1-transcriptional output. Taken together the resulting phenotype is development of differentiated erythroid cells in TF1 erythroleukemia cells as well as in primary human hematopoietic stem-progenitor cells.

References

Adamo, P., & Ladomery, M. R. (2016, January 28). The oncogene ERG: A key factor in prostate cancer. *Oncogene*. Nature Publishing Group. https://doi.org/10.1038/onc.2015.109

Ahmed, M., Wong, E. Y.-M., Sun, J., Xu, J., Wang, F., & Xu, P.-X. (2011). Eya1-Six1 interaction is sufficient to induce hair cell fate in the cochlea by activating Atoh1 expression in cooperation with Sox2. https://doi.org/10.1016/j.devcel.2011.12.006

Ahmed, M., Xu, J., & Xu, P. X. (2012). EYA1 and SIX1 drive the neuronal developmental program in cooperation with the SWI/SNF chromatin-remodeling complex and SOX2 in the mammalian inner ear. *Development*, *139*(11), 1965–1977. https://doi.org/10.1242/dev.071670

Amigo, J. D., Ackermann, G. E., Cope, J. J., Yu, M., Cooney, J. D., Ma, D., … Paw, B. H. (2009). The role and regulation of friend of GATA-1 (FOG-1) during blood development in the zebrafish. *Blood*, *114*(21), 4654–4663. https://doi.org/10.1182/blood-2008-12-189910

Anguita, E., Hughes, J., Heyworth, C., Blobel, G. A., Wood, W. G., & Higgs, D. R. (2004). Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *EMBO Journal*, *23*(14), 2841–2852. https://doi.org/10.1038/sj.emboj.7600274

Ben-David, Y., Giddens, E. B., Letwin, K., & Bernstein, A. (1991). Erythroleukemia induction by Friend murine leukemia virus: Insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. *Genes and Development*, *5*(6), 908–918. https://doi.org/10.1101/gad.5.6.908

Bender, J. G., Unverzagt, K., Walker, D. E., Lee, W., Smith, S., Williams, S., & Van Epps, D. E. (1994). Phenotypic analysis and characterization of CD34+ cells from normal human bone marrow, cord blood, peripheral blood, and mobilized peripheral blood from patients undergoing autologous stem cell transplantation. *Clinical Immunology and Immunopathology*, *70*(1), 10–18. https://doi.org/10.1006/clin.1994.1003

Blevins, M. A., Towers, C. G., Patrick, A. N., Zhao, R., & Ford, H. L. (2015, February 1). The SIX1-EYA transcriptional complex as a therapeutic target in cancer. *Expert Opinion on Therapeutic Targets*. Informa Healthcare. https://doi.org/10.1517/14728222.2014.978860

Bodine, D. M., & Ley, T. J. (1987). An enhancer element lies 3′ to the human A gamma globin gene. *The EMBO Journal*, *6*(10), 2997–3004. https://doi.org/10.1002/j.1460-2075.1987.tb02605.x

Bonini, N. M., Bui, Q. T., Gray-Board, G. L., & Warrick, J. M. (1997). The Drosophila eyes absent gene directs ectopic eye formation in a pathway conserved between flies and vertebrates. *Development*, *124*(23), 4819–4826. Retrieved from https://pubmed.ncbi.nlm.nih.gov/9428418/

Brugmann, S. A., Pandur, P. D., Kenyon, K. L., Pignoni, F., & Moody, S. A. (2004). Six1 promotes a placodal fate within the lateral neurogenic ectoderm by functioning as both a transcriptional activator and repressor. *Development*, *131*(23), 5871–5881. https://doi.org/10.1242/dev.01516

Brunetti, L., Gundry, M. C., Sorcini, D., Guzman, A. G., Huang, Y. H., Ramabadran, R., … Goodell, M. A. (2018). Mutant NPM1 Maintains the Leukemic State through HOX Expression. *Cancer Cell*, *34*(3), 499-512.e9. https://doi.org/10.1016/j.ccell.2018.08.005

Buller, C., Xu, X., Marquis, V., Schwanke, R., & Xu, P.-X. (2001). *Molecular effects of Eya1 domain mutations causing organ defects in BOR syndrome*. *Human Molecular Genetics* (Vol. 10).

Cantor, A. B., & Orkin, S. H. (2005). Coregulation of GATA factors by the Friend of GATA (FOG) family of multitype zinc finger proteins. *Seminars in Cell and Developmental Biology*, *16*(1), 117–128. https://doi.org/10.1016/j.semcdb.2004.10.006

Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G. W., … Huang, B. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, *155*(7), 1479–1491. https://doi.org/10.1016/j.cell.2013.12.001

Cheyette, B. N. R., Green, P. J., Martin, K., Garren, H., Hartenstein, V., & Zipursky, S. L. (1994). The drosophila sine oculis locus encodes a homeodomain-containing protein required for the development of the entire visual system. *Neuron*, *12*(5), 977–996. https://doi.org/10.1016/0896-6273(94)90308-5

Chlon, T. M., & Crispino, J. D. (2012, November 1). Combinatorial regulation of tissue specification by GATA and FOG factors. *Development (Cambridge)*. Oxford University Press for The Company of Biologists Limited. https://doi.org/10.1242/dev.080440

Chlon, T. M., McNulty, M., Goldenson, B., Rosinski, A., & Crispino, J. D. (2015). Global transcriptome and chromatin occupancy analysis reveal the short isoform of

GATA1 is deficient for erythroid specification and gene expression. *Haematologica*, *100*(5), 575–584. https://doi.org/10.3324/haematol.2014.112714

Christophorou, N. A. D., Bailey, A. P., Hanson, S., & Streit, A. (2009). Activation of Six1 target genes is required for sensory placode formation. *Developmental Biology*, *336*(2), 327–336. https://doi.org/10.1016/j.ydbio.2009.09.025

Chu, Y., Chen, Y., Li, M., Shi, D., Wang, B., Lian, Y., … Yuan, W. (2019). Six1 regulates leukemia stem cell maintenance in acute myeloid leukemia. *Cancer Science*, *110*(7), 2200–2210. https://doi.org/10.1111/cas.14033

Civin, C. I., & Loken, M. R. (1987). Cell surface antigens on human marrow cells: dissection of hematopoietic development using monoclonal antibodies and multiparameter flow cytometry. *International Journal of Cell Cloning*, *5*(4), 267–288. https://doi.org/10.1002/stem.5530050403

Comb, C. L., Vmrtmt, P., Vardler, F., Gobart, S., Cartran, J. P., Lbraeh, S., & Umveu, P. (1996). Erythropoietin-induced erythroid differentiation of the human erythroleukemia cell line TF-1 correlates with impaired stats activation. *Experimental Hematology*, *24*(9), 1107. Retrieved from /pmc/articles/PMC452141/?report=abstract

Costantini, F., & Shakya, R. (2006, February). GDNF/Ret signaling and the development of the kidney. *BioEssays*. Bioessays. https://doi.org/10.1002/bies.20357

Crispino, J. D. (2005). GATA1 in normal and malignant hematopoiesis. *Seminars in Cell and Developmental Biology*, *16*(1), 137–147. https://doi.org/10.1016/j.semcdb.2004.11.002

Dahl, R., Walsh, J. C., Lancki, D., Laslo, P., Iyer, S. R., Singh, H., & Celeste Simon, M. (2003). Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBPα ratio and granulocyte colony-stimulating factor. *Nature Immunology*, *4*(10), 1029–1036. https://doi.org/10.1038/ni973

Dearolf, C. R. (1998, February 20). Fruit fly "Leukemia." *Biochimica et Biophysica Acta - Reviews on Cancer*. Elsevier B.V. https://doi.org/10.1016/S0304-419X(97)00031-0

Dingar, D., Kalkat, M., Chan, P. K., Srikumar, T., Bailey, S. D., Tu, W. B., … Raught, B. (2015). BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. *Journal of Proteomics*, *118*, 95–111. https://doi.org/10.1016/j.jprot.2014.09.029

Doré, L. C., & Crispino, J. D. (2011, July 14). Transcription factor networks in erythroid cell and megakaryocyte development. *Blood*. Blood. https://doi.org/10.1182/blood-2011-04-285981

Dzierzak, E., & Philipsen, S. (2013). Erythropoiesis: Development and differentiation. *Cold Spring Harbor Perspectives in Medicine*, *3*(4). https://doi.org/10.1101/cshperspect.a011601

Eisbacher, M., Holmes, M. L., Newton, A., Hogg, P. J., Khachigian, L. M., Crossley, M., & Chong, B. H. (2003). Protein-Protein Interaction between Fli-1 and GATA-1 Mediates Synergistic Expression of Megakaryocyte-Specific Genes through Cooperative DNA Binding. *Molecular and Cellular Biology*, *23*(10), 3427–3441. https://doi.org/10.1128/mcb.23.10.3427-3441.2003

Esteve, P., & Bovolenta, P. (1999). cSix4, a member of the six gene family of transcription factors, is expressed during placode and somite development. *Mechanisms of Development*, *85*(1–2), 161–165. https://doi.org/10.1016/S0925-4773(99)00079-9

Fossett, N. (2013, February 1). Signal transduction pathways, intrinsic regulators, and the control of cell fate choice. *Biochimica et Biophysica Acta - General Subjects*. Elsevier. https://doi.org/10.1016/j.bbagen.2012.06.005

Frandsen, J. L., Gunn, B., Muratoglu, S., Fossett, N., & Newfeld, S. J. (2008). Salmonella pathogenesis reveals that BMP signaling regulates blood cell homeostasis and immune responses in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(39), 14952–14957. https://doi.org/10.1073/pnas.0808208105

Frazee, A. C., Pertea, G., Jaffe, A. E., Langmead, B., Salzberg, S. L., & Leek, J. T. (2015). Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nature Biotechnology*. Nature Publishing Group. https://doi.org/10.1038/nbt.3172

Frelin, C., Herrington, R., Janmohamed, S., Barbara, M., Tran, G., Paige, C. J., … Iscove, N. N. (2013). GATA-3 regulates the self-renewal of long-term hematopoietic stem cells. *Nature Immunology*, *14*(10), 1037–1044. https://doi.org/10.1038/ni.2692

Gold, K. S., & Brückner, K. (2015, December 1). Macrophages and cellular immunity in Drosophila melanogaster. *Seminars in Immunology*. Academic Press. https://doi.org/10.1016/j.smim.2016.03.010

Grass, J. A., Boyer, M. E., Pal, S., Wu, J., Weiss, M. J., & Bresnick, E. H. (n.d.). *GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling*. Retrieved from www.pnas.orgcgidoi10.1073pnas.1432147100

Grifone, R., Demignon, J., Houbron, C., Souil, E., Niro, C., Seller, M. J., … Maire, P. (2005). Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. *Development*, *132*(9), 2235–2249. https://doi.org/10.1242/dev.01773

Gulbinas, A., Berberat, P. O., Dambrauskas, Z., Giese, T., Giese, N., Autschbach, F., … Friess, H. (2006). Aberrant gata-3 expression in human pancreatic cancer. *Journal of Histochemistry and Cytochemistry*, *54*(2), 161–169. https://doi.org/10.1369/jhc.5A6626.2005

Hanspal, M., Smockova, Y., & Uong, Q. (1998). Molecular Identification and Functional Characterization of a Novel Protein That Mediates the Attachment of Erythroblasts to Macrophages. *Blood*, *92*(8), 2940–2950. https://doi.org/10.1182/blood.V92.8.2940

Hasegawa, A., Kaneko, H., Ishihara, D., Nakamura, M., Watanabe, A., Yamamoto, M., … Shimizu, R. (2016). GATA1 Binding Kinetics on Conformation-Specific Binding Sites Elicit Differential Transcriptional Regulation. *Molecular and Cellular Biology*, *36*(16), 2151–2167. https://doi.org/10.1128/mcb.00017-16

Hawley, R. S., & Gilliland, W. D. (2006). Sometimes the result is not the answer: The truths and the lies that come from using the complementation test. *Genetics*. Genetics Society of America. https://doi.org/10.1534/genetics.106.064550

Jack, B. H. A., & Crossley, M. (2010). GATA proteins work together with Friend of GATA (FOG) and C-terminal Binding Protein (CTBP) co-regulators to control adipogenesis. *Journal of Biological Chemistry*, *285*(42), 32405–32414. https://doi.org/10.1074/jbc.M110.141317

Kastner, P., Lawrence, H. J., Waltzinger, C., Ghyselinck, N. B., Chambon, P., & Chan, S. (2001). Positive and negative regulation of granulopoiesis by endogenous RARα. *Blood*, *97*(5), 1314–1320. https://doi.org/10.1182/blood.V97.5.1314

Kaushansky, K., & St, D. (2006). *Lineage-Specific Hematopoietic Growth Factors*. *N Engl J Med* (Vol. 354). Retrieved from www.nejm.org

Kearns, N. A., Genga, R. M. J., Enuameh, M. S., Garber, M., Wolfe, S. A., & Maehr, R. (2014). Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Development (Cambridge)*, *141*(1), 219–223. https://doi.org/10.1242/dev.103341

Kim, B. R., Coyaud, E., Laurent, E. M. N., St-Germain, J., Van De Laar, E., Tsao, M. S., … Moghal, N. (2017). Identification of the SOX2 interactome by BioID reveals EP300 as a mediator of SOX2-dependent squamous differentiation and lung squamous cell carcinoma growth. *Molecular and Cellular Proteomics*, *16*(10), 1864–1888. https://doi.org/10.1074/mcp.M116.064451

Kim, D. I., Cutler, J. A., Na, C. H., Reckel, S., Renuse, S., Madugundu, A. K., … Pandey, A. (2018). BioSITe: A Method for Direct Detection and Quantitation of Site-Specific Biotinylation. *Journal of Proteome Research*, *17*(2), 759–769. https://doi.org/10.1021/acs.jproteome.7b00775

Kim, M., Civin, C. I., & Kingsbury, T. J. (2019). MicroRNAs as regulators and effectors of hematopoietic transcription factors. *Wiley Interdisciplinary Reviews: RNA*, *10*(5), e1537. https://doi.org/10.1002/wrna.1537

Kim, M., Tan, Y. S., Cheng, W.-C., Kingsbury, T. J., Heimfeld, S., & Civin, C. I. (2015). MIR144 and MIR451 regulate human erythropoiesis via RAB14. *British Journal of Haematology*, *168*(4), 583–597. https://doi.org/10.1111/bjh.13164

Kingsbury, T. J., Kim, M. J., & Civin, C. I. (2019). Regulation of cancer stem cell properties by SIX1, a member of the PAX-SIX-EYA-DACH network. In *Advances in Cancer Research* (Vol. 141, pp. 1–42). Academic Press Inc. https://doi.org/10.1016/bs.acr.2018.12.001

Kitajima, K., Zheng, J., Yen, H., Sugiyama, D., & Nakano, T. (2006). Multipotential differentiation ability of GATA-1-null erythroid-committed cells. *Genes and Development*, *20*(6), 654–659. https://doi.org/10.1101/gad.1378206

Kitamura, T., Tange, T., Terasawa, T., Chiba, S., Kuwaki, T., Miyagawa, K., … Takaku, F. (1989). Establishment and characterization of a unique human cell line that proliferates dependently on GM·CSF, IL·3, or erythropoietin. *Journal of Cellular Physiology*, *140*(2), 323–334. https://doi.org/10.1002/jcp.1041400219

Kong, D., Li, A., Liu, Y., Cui, Q., Wang, K., Zhang, D., … Wu, K. (2019). SIX1 Activates STAT3 Signaling to Promote the Proliferation of Thyroid Carcinoma via EYA1. *Frontiers in Oncology*, *9*, 1450. https://doi.org/10.3389/fonc.2019.01450

Koury, M. J. (n.d.). and Veterans Affairs Tennessee Valley Healthcare System, *777*. https://doi.org/10.1016/j.blre.2014.01.002

Krivega, I., Dale, R. K., & Dean, A. (2014). Role of LDB1 in the transition from chromatin looping to transcription activation. *Genes and Development*, *28*(12), 1278–1290. https://doi.org/10.1101/gad.239749.114

Kuhrt, D., & Wojchowski, D. M. (2015, June 4). Emerging EPO and EPO receptor regulators and signal transducers. *Blood*. American Society of Hematology. https://doi.org/10.1182/blood-2014-11-575357

Kumar, J. P. (2009, February). The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease. *Cellular and Molecular Life Sciences*. NIH Public Access. https://doi.org/10.1007/s00018-008-8335-4

Kuvardina, O. N., Herglotz, J., Kolodziej, S., Kohrs, N., Herkt, S., Wojcik, B., … Lausen, J. (2015). RUNX1 represses the erythroid gene expression program during megakaryocytic differentiation. *Blood*, *125*(23), 3570–3579. https://doi.org/10.1182/blood-2014-11-610519

Kwon, K., & Beckett, D. (2000). Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Science*, *9*(8), 1530–1539. https://doi.org/10.1110/ps.9.8.1530

Lanot, R., Zachary, D., Holder, F., & Meister, M. (2001). Postembryonic hematopoiesis in Drosophila. *Developmental Biology*, *230*(2), 243–257. https://doi.org/10.1006/dbio.2000.0123

Laurenti, E., & Göttgens, B. (2018, January 24). From haematopoietic stem cells to complex differentiation landscapes. *Nature*. Nature Publishing Group. https://doi.org/10.1038/nature25022

Laurenti, E., Frelin, C., Xie, S., Ferrari, R., Dunant, C. F., Zandi, S., … Dick, J. E. (2015). CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell*, *16*(3), 302–313. https://doi.org/10.1016/j.stem.2015.01.017

Le Grand, F., Grifone, R., Mourikis, P., Houbron, C., Gigaud, C., Pujol, J., … Maire, P. (2012). Six1 regulates stem cell repair potential and self-renewal during skeletal muscle regeneration. *Journal of Cell Biology*, *198*(5), 815–832. https://doi.org/10.1083/jcb.201201050

Lentjes, M. H. F. M., Niessen, H. E. C., Akiyama, Y., Bruïne, D. A. P., Melotte, V., & Engeland, M. V. A. N. (2016, March 8). The emerging role of GATA transcription factors in development and disease. *Expert Reviews in Molecular Medicine*. Cambridge University Press. https://doi.org/10.1017/erm.2016.2

Li, B., Huang, Q., & Wei, G. H. (2019). The role of hox transcription factors in cancer predisposition and progression. *Cancers*, *11*(4). https://doi.org/10.3390/cancers11040528

Li, J., Hale, J., Bhagia, P., Xue, F., Chen, L., Jaffray, J., … An, X. (2014). Isolation and transcriptome analyses of human erythroid progenitors: BFU-E and CFU-E. *Blood*, *124*(24), 3636–3645. https://doi.org/10.1182/blood-2014-07-588806

Li, X., Oghi, K. A., Zhang, J., Krones, A., Bush, K. T., Glass, C. K., … Rosenfeld, M. G. (2003). Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. *Nature*, *426*(6964), 247–254. https://doi.org/10.1038/nature02083

Li, Y., Ke, Q., Shao, Y., Zhu, G., Li, Y., Geng, N., … Li, F. (2015). GATA1 induces epithelial-mesenchymal transition in breast cancer cells through PAK5 oncogenic signaling. *Oncotarget*, *6*(6), 4345–4356. https://doi.org/10.18632/oncotarget.2999

Liew, C. K., Simpson, R. J. Y., Kwan, A. H. Y., Crofts, L. A., Loughlin, F. E., Matthews, J. M., … Mackay, J. P. (2005). Zinc fingers as protein recognition motifs: Structural basis for the GATA-1/friend of GATA interaction. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(3), 583–588. https://doi.org/10.1073/pnas.0407511102

Ling, K. W., Ottersbach, K., Van Hamburg, J. P., Oziemlak, A., Tsai, F. Y., Orkin, S. H., … Dzierzak, E. (2004). GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *Journal of Experimental Medicine*, *200*(7), 871–882. https://doi.org/10.1084/jem.20031556

Liu, X., Zhang, Y., Ni, M., Cao, H., Signer, R. A. J., Li, D., … Xu, J. (2017). Regulation of mitochondrial biogenesis in erythropoiesis by mTORC1-mediated protein translation. *Nature Cell Biology*, *19*(6), 626–638. https://doi.org/10.1038/ncb3527

Mandal, L., Banerjee, U., & Hartenstein, V. (2004). Evidence for a fruit fly hemangioblast and similarities between lymph-gland hematopoiesis in fruit fly and mammal aorta-gonadal-mesonephros mesoderm. *Nature Genetics*, *36*(9), 1019–1023. https://doi.org/10.1038/ng1404

Manwani, D., & Bieker, J. J. (2008). The Erythroblastic Island. *Current Topics in Developmental Biology*. NIH Public Access. https://doi.org/10.1016/S0070-2153(07)00002-6

Masiuk, K. E., Zhang, R., Osborne, K., Hollis, R. P., Campo-Fernandez, B., & Kohn, D. B. (2019). PGE2 and Poloxamer Synperonic F108 Enhance Transduction of Human HSPCs with a β-Globin Lentiviral Vector. *Molecular Therapy - Methods and Clinical Development*, *13*, 390–398. https://doi.org/10.1016/j.omtm.2019.03.005

Mayani, H. (2016). The regulation of hematopoietic stem cell populations [version 1; referees: 2 approved]. *F1000Research*. Faculty of 1000 Ltd. https://doi.org/10.12688/F1000RESEARCH.8532.1

McCranor, B. J., Kim, M. J., Cruz, N. M., Xue, Q. L., Berger, A. E., Walston, J. D., … Roy, C. N. (2014). Interleukin-6 directly impairs the erythroid development of human TF-1 erythroleukemic cells. *Blood Cells, Molecules, and Diseases*, *52*(2–3), 126–133. https://doi.org/10.1016/j.bcmd.2013.09.004

Merika, M., & Orkin, S. H. (1995). Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Krüppel family proteins Sp1 and EKLF. *Molecular and Cellular Biology*, *15*(5), 2437–2447. https://doi.org/10.1128/mcb.15.5.2437

Metcalf, D. (2007, June 22). On Hematopoietic Stem Cell Fate. *Immunity*. Cell Press. https://doi.org/10.1016/j.immuni.2007.05.012

Molenaar, J. J., Ebus, M. E., Koster, J., Santo, E., Geerts, D., Versteeg, R., & Caron, H. N. (2010). Cyclin D1 is a direct transcriptional target of GATA3 in neuroblastoma tumor cells. *Oncogene*, *29*(18), 2739–2745. https://doi.org/10.1038/onc.2010.21

Muckenthaler, M. U., Rivella, S., Hentze, M. W., & Galy, B. (2017, January 26). A Red Carpet for Iron Metabolism. *Cell*. Cell Press. https://doi.org/10.1016/j.cell.2016.12.034

Nagel, S., Meyer, C., Kaufmann, M., Drexler, H. G., & MacLeod, R. A. F. (2015). Aberrant expression of homeobox gene SIX1 in Hodgkin lymphoma. *Oncotarget*, *6*(37), 40112–40126. https://doi.org/10.18632/oncotarget.5556

Nakahata, T., & Okumura, N. (1994). Cell surface antigen expression in human erythroid progenitors: erythroid and megakaryocytic markers. *Leukemia & Lymphoma*, *13*(5–6), 401–409. https://doi.org/10.3109/10428199409049629

Nero, T. L., Morton, C. J., Holien, J. K., Wielens, J., & Parker, M. W. (2014). Oncogenic protein interfaces: Small molecules, big challenges. *Nature Reviews Cancer*. Nature Publishing Group. https://doi.org/10.1038/nrc3690

Nichols, K. E., Crispino, J. D., Poncz, M., White, J. G., Orkin, S. H., Maris, J. M., & Weiss, M. J. (2000). Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA 1. *Nature Genetics*, *24*(3), 266–270. https://doi.org/10.1038/73480

Notta, F., Doulatov, S., Laurenti, E., Poeppl, A., Jurisica, I., & Dick, J. E. (2011). Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science*, *333*(6039), 218–221. https://doi.org/10.1126/science.1201219

Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O. I., Wilson, G., … Dick, J. E. (2016). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*, *351*(6269). https://doi.org/10.1126/science.aab2116

O'Brien, L. L., Guo, Q., Lee, Y. J., Tran, T., Benazet, J. D., Whitney, P. H., … McMahon, A. P. (2016). Differential regulation of mouse and human nephron progenitors by the six family of transcriptional regulators. *Development (Cambridge)*, *143*(4), 595–608. https://doi.org/10.1242/dev.127175

Oliver, G., Mailhos, A., Wehr, R., Copeland, N. G., Jenkins, N. A., & Gruss, P. (1995). Six3, a murine homologue of the sine oculis gene, demarcates the most anterior border of the developing neural plate and is expressed during eye development. *Development*, *121*(12).

Oliver, G., Wehr, R., Jenkins, N. A., Copeland, N. G., Cheyette, B. N., Hartenstein, V., … Gruss, P. (1995). Homeobox genes and connective tissue patterning. *Development*, *121*(3).

Parmacek, M. S., Ip, H. S., Jung, F., Shen, T., Martin, J. F., Vora, A. J., … Leiden, J. M. (1994). A novel myogenic regulatory circuit controls slow/cardiac troponin C gene transcription in skeletal muscle. *Molecular and Cellular Biology*, *14*(3), 1870–1885. https://doi.org/10.1128/mcb.14.3.1870

Patrick, A. N., Cabrera, J. H., Smith, A. L., Chen, X. S., Ford, H. L., & Zhao, R. (2013). Structure-function analyses of the human SIX1-EYA2 complex reveal insights into metastasis and BOR syndrome. *Nature Structural and Molecular Biology*, *20*(4), 447–453. https://doi.org/10.1038/nsmb.2505

Patrick, A. N., Schiemann, B. J., Yang, K., Zhao, R., & Ford, H. L. (2009). Biochemical and functional characterization of six SIX1 Branchio-oto-renal syndrome mutations. *Journal of Biological Chemistry*, *284*(31), 20781–20790. https://doi.org/10.1074/jbc.M109.016832

Paulson, R. F., Hariharan, S., & Little, J. (2020). Stress erythropoiesis: definitions and models for its study. *Experimental Hematology*, *0*(0). https://doi.org/10.1016/j.exphem.2020.07.011

Paulson, R. F., Shi, L., & Wu, D.-C. (2011). Stress erythropoiesis: new signals and new stress progenitor cells. *Current Opinion in Hematology*, *18*(3), 139–145. https://doi.org/10.1097/MOH.0b013e32834521c8

Pellin, D., Loperfido, M., Baricordi, C., Wolock, S. L., Montepeloso, A., Weinberg, O. K., … Biasco, L. (2019). A comprehensive single cell transcriptional landscape of human hematopoietic progenitors. *Nature Communications*, *10*(1), 1–15. https://doi.org/10.1038/s41467-019-10291-0

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, *33*(3), 290–295. https://doi.org/10.1038/nbt.3122

Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., … Costantini, F. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*, *349*(6306), 257–260. https://doi.org/10.1038/349257a0

Pham, H., Kearns, N. A., & Maehr, R. (2016). Transcriptional regulation with CRISPR/cas9 effectors in mammalian cells. In *Methods in Molecular Biology* (Vol. 1358, pp. 43–57). Humana Press Inc. https://doi.org/10.1007/978-1-4939-3067-8_3

Ponka, P., & Lok, C. N. (1999). The transferrin receptor: Role in health and disease. *International Journal of Biochemistry and Cell Biology*, *31*(10), 1111–1137. https://doi.org/10.1016/S1357-2725(99)00070-9

Pop, R., Shearstone, J. R., Shen, Q., Liu, Y., Hallstrom, K., Koulnis, M., … Socolovsky, M. (2010). A key commitment step in erythropoiesis is synchronized with the cell cycle clock through mutual inhibition between PU.1 and S-phase progression. *PLoS Biology*, *8*(9). https://doi.org/10.1371/journal.pbio.1000484

Preethi, B., Shanthi, V., & Ramanathan, K. (2015). Reckoning the SIX1 mutation's effects in branchio-oto-renal syndrome — A bioinformatics approach. *Frontiers in Biology*, *10*(5), 448–457. https://doi.org/10.1007/s11515-015-1370-2

Rizki, T. M., & Rizki, R. M. (1992). Lamellocyte differentiation in Drosophila larvae parasitized by Leptopilina. *Developmental and Comparative Immunology*, *16*(2–3), 103–110. https://doi.org/10.1016/0145-305X(92)90011-Z

Rodriguez, P., Bonte, E., Krijgsveld, J., Kolodziej, K. E., Guyot, B., Heck, A. J. R., … Strouboulis, J. (2005). GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO Journal*, *24*(13), 2354–2366. https://doi.org/10.1038/sj.emboj.7600702

Rojas-Sutterlin, S., Lecuyer, E., & Hoang, T. (2014). Kit and Scl regulation of hematopoietic stem cells. *Current Opinion in Hematology*. Lippincott Williams and Wilkins. https://doi.org/10.1097/MOH.0000000000000052

Romano, O., Petiti, L., Felix, T., Meneghini, V., Portafax, M., Antoniani, C., … Miccio, A. (2020). GATA Factor-Mediated Gene Regulation in Human Erythropoiesis. *IScience*, *23*(4), 101018. https://doi.org/10.1016/j.isci.2020.101018

Roux, K. J., Kim, D. I., & Burke, B. (2018). BioID: A screen for protein-protein interactions. *Current Protocols in Protein Science*, *2013*(1), 19.23.1-19.23.14. https://doi.org/10.1002/0471140864.ps1923s74

Roux, K. J., Kim, D. I., Raida, M., & Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of Cell Biology*, *196*(6), 801–810. https://doi.org/10.1083/jcb.201112098

Ruf, R. G., Xu, P. X., Silvius, D., Otto, E. A., Beekmann, F., Muerb, U. T., … Hildebrandt, F. (2004). SIX1 mutations cause branchio-oto-renal syndrome by disruption of EYA1-SIX1-DNA complexes. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(21), 8090–8095. https://doi.org/10.1073/pnas.0308475101

Santolini, M., Sakakibara, I., Gauthier, M., Ribas-Aulinas, F., Takahashi, H., Sawasaki, T., … Maire, P. (2016). MyoD reprogramming requires Six1 and Six4 homeoproteins: genome-wide cis-regulatory module analysis. *Nucleic Acids Research*, *44*(18), 8621–8640. https://doi.org/10.1093/nar/gkw512

Sauvageau, G., Iscove, N. N., & Humphries, R. K. (2004, September 20). In vitro and in vivo expansion of hematopoietic stem cells. *Oncogene*. Oncogene. https://doi.org/10.1038/sj.onc.1207942

Schechter, A. N. (2008). Hemoglobin research and the origins of molecular medicine. *Blood*, *112*(10), 3927–3938. https://doi.org/10.1182/blood-2008-04-078188

Seita, J., & Weissman, I. L. (2010, November). Hematopoietic stem cell: Self-renewal versus differentiation. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*. NIH Public Access. https://doi.org/10.1002/wsbm.86

Shimizu, R., Ohneda, K., Engel, J. D., Trainor, C. D., & Yamamoto, M. (2004). Transgenic rescue of GATA-1-deficient mice with GATA-1 lacking a FOG-1 association site phenocopies patients with X-linked thrombocytopenia. *Blood*, *103*(7), 2560–2567. https://doi.org/10.1182/blood-2003-07-2514

Simon, M. C. (1995). Gotta have GATA. *Nature Genetics*, *11*(1), 9–11. https://doi.org/10.1038/ng0995-9

Smith, R. J. (1993). *Branchiootorenal Spectrum Disorder*. *GeneReviews®*. University of Washington, Seattle. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20301554

Sood, R., Kamikubo, Y., & Liu, P. (2017, April 13). Role of RUNX1 in hematological malignancies. *Blood*. American Society of Hematology. https://doi.org/10.1182/blood-2016-10-687830

Spitz, F., & Furlong, E. E. M. (2012, September 7). Transcription factors: From enhancer binding to developmental control. *Nature Reviews Genetics*. Nature Publishing Group. https://doi.org/10.1038/nrg3207

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., … Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Suzuki, M., Kobayashi-Osaki, M., Tsutsumi, S., Pan, X., Ohmori, S., Takai, J., … Yamamoto, M. (2013). GATA factor switching from GATA2 to GATA1 contributes to erythroid differentiation. *Genes to Cells*, *18*(11), 921–933. https://doi.org/10.1111/gtc.12086

Tanimura, N., Miller, E., Igarashi, K., Yang, D., Burstyn, J. N., Dewey, C. N., & Bresnick, E. H. (2016). Mechanism governing heme synthesis reveals a GATA factor/heme circuit that controls differentiation. *EMBO Reports*, *17*(2), 249–265. https://doi.org/10.15252/embr.201541465

Terstappen LW, H. S. S. M. L. M. L. M. (1991). Sequential generations of hematopoietic colonies derived from single nonlineage-committed CD34+CD38- progenitor cells - PubMed. *Blood*, *77*(6), 1218–1227. Retrieved from https://pubmed.ncbi.nlm.nih.gov/1705833/

Tirelli, V., Ghinassi, B., Migliaccio, A. R., Whitsett, C., Masiello, F., Sanchez, M., & Migliaccio, G. (2011). Access to. *Research Stem Cells International*, *2011*. https://doi.org/10.4061/2011/602483

Trainor, C. D., Mas, C., Archambault, P., Di Lello, P., & Omichinski, J. G. (2009). GATA-1 associates with and inhibits p53. *Blood*, *114*(1), 165–173. https://doi.org/10.1182/blood-2008-10-180489

Tsai, F. Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., … Orkin, S. H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, *371*(6494), 221–226. https://doi.org/10.1038/371221a0

Tsai, S. F., Martin, D. I. K., Zon, L. I., D’Andrea, A. D., Wong, G. G., & Orkin, S. H. (1989). Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature*, *339*(6224), 446–451. https://doi.org/10.1038/339446a0

Tse, C., Shoemaker, A. R., Adickes, J., Anderson, M. G., Chen, J., Jin, S., … Elmore, S. W. (2008). ABT-263: A potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Research*, *68*(9), 3421–3428. https://doi.org/10.1158/0008-5472.CAN-07-5836

Van Lochem, E. G., Van Der Velden, V. H. J., Wind, H. K., Te Marvelde, J. G., Westerdaal, N. A. C., & Van Dongen, J. J. M. (2004). Immunophenotypic differentiation patterns of normal hematopoiesis in human bone marrow: Reference patterns for age-related changes and disease-induced shifts. *Cytometry Part B - Clinical Cytometry*, *60*(1), 1–13. https://doi.org/10.1002/cyto.b.20008

Vassiliou, G. S., Cooper, J. L., Rad, R., Li, J., Rice, S., Uren, A., … Bradley, A. (2011). Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nature Genetics*, *43*(5), 470–476. https://doi.org/10.1038/ng.796

Vlisidou, I., & Wood, W. (2015). Drosophila blood cells and their role in immune responses. *FEBS Journal*, *282*(8), 1368–1382. https://doi.org/10.1111/febs.13235

Walter, D. M., Venancio, O. S., Buza, E. L., Tobias, J. W., Deshpande, C., Gudiel, A., … Feldser, D. M. (2017). Systematic in vivo inactivation of chromatin- regulating enzymes identifies Setd2 as a potent tumor suppressor in lung adenocarcinoma. *Cancer Research*, *77*(7), 1719–1729. https://doi.org/10.1158/0008-5472.CAN-16-2159

Wang, Q. F., Wu, G., Mi, S., He, F., Wu, J., Dong, J., … Thirman, M. J. (2011). MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. *Blood*, *117*(25), 6895–6905. https://doi.org/10.1182/blood-2010-12-324699

Weasner, B., Salzer, C., & Kumar, J. P. (2007). Sine oculis, a member of the SIX family of transcription factors, directs eye formation. *Developmental Biology*, *303*(2), 756–771. https://doi.org/10.1016/j.ydbio.2006.10.040

Wilkinson-White, L., Lester, K. L., Ripin, N., Jacques, D. A., Mitchell Guss, J., & Matthews, J. M. (2015). GATA1 directly mediates interactions with closely spaced pseudopalindromic but not distantly spaced double GATA sites on DNA. *Protein Science*, *24*(10), 1649–1659. https://doi.org/10.1002/pro.2760

Wlodarski, M. W., Collin, M., & Horwitz, M. S. (2017). GATA2 deficiency and related myeloid neoplasms. *Seminars in Hematology*, *54*(2), 81–86. https://doi.org/10.1053/j.seminhematol.2017.05.002

Wu, H., Liu, X., Jaenisch, R., & Lodish, H. F. (1995). Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell*, *83*(1), 59–67. https://doi.org/10.1016/0092-8674(95)90234-1

Wu, W., Huang, R., Wu, Q., Li, P., Chen, J., Li, B., & Liu, H. (2014, September 5). The role of Six1 in the genesis of muscle cell and skeletal muscle development. *International Journal of Biological Sciences*. Ivyspring International Publisher. https://doi.org/10.7150/ijbs.9442

Wu, W., Ren, Z., Li, P., Yu, D., Chen, J., Huang, R., & Liu, H. (2015). Six1: A critical transcription factor in tumorigenesis. *International Journal of Cancer*, *136*(6), 1245–1253. https://doi.org/10.1002/ijc.28755

Xie W, Q. W. Q. A. Z. Z. W. L. K. Y. Z. Z. O. Y. (2016). GATA1 promotes tumorigenesis and metastasis in breast cancer by cooperating with ZEB2. *Int J Clin Exp Path*, *9*(4), 4167–4178.

Xu, H. X., Wu, K. J., Tian, Y. J., Liu, Q., Han, N., He, X. L., … Wu, K. M. (2016, July 1). Expression profile of SIX family members correlates with clinic-pathological features and prognosis of breast cancer. *Medicine (United States)*. Lippincott Williams and Wilkins. https://doi.org/10.1097/MD.0000000000004085

Yu, Y., Davicioni, E., Triche, T. J., & Merlino, G. (2006). The homeoprotein Six1 transcriptionally activates multiple protumorigenic genes but requires ezrin to promote metastasis. *Cancer Research*, *66*(4), 1982–1989. https://doi.org/10.1158/0008-5472.CAN-05-2360

Zaidan, N., & Ottersbach, K. (2018, November 1). The multi-faceted role of Gata3 in developmental haematopoiesis. *Open Biology*. Royal Society Publishing. https://doi.org/10.1098/rsob.180152

Zhang, L. shu, Kang, X., Lu, J., Zhang, Y., Wu, X., Wu, G., … Lum, L. (2019). Installation of a cancer promoting WNT/SIX1 signaling axis by the oncofusion protein MLL-AF9. *EBioMedicine*, *39*, 145–158. https://doi.org/10.1016/j.ebiom.2018.11.039

Zheng, J., Kitajima, K., Sakai, E., Kimura, T., Minegishi, N., Yamamoto, M., & Nakano, T. (2006). Differential effects of GATA-1 on proliferation and differentiation of erythroid lineage cells. https://doi.org/10.1182/blood-2005-04

Zheng, X. hui, Liang, P. he, Guo, J. xin, Zheng, Y. ru, Han, J., Yu, L. li, … Li, L. (2010). Expression and clinical implications of homeobox gene six1 in cervical cancer cell lines and cervical epithelial tissues. *International Journal of Gynecological Cancer*, *20*(9). https://doi.org/10.1111/IGC.0b013e3181f9defb

Zhou, Z., Ai, H., Li, K., Yao, X., Zhu, W., Liu, L., … Li, Y. (2018). Prohibitin 2 localizes in nucleolus to regulate ribosomal RNA transcription and facilitate cell proliferation in RD cells. *Scientific Reports*, *8*(1), 1479. https://doi.org/10.1038/s41598-018-19917-7

Zou, D., Silvius, D., Fritzsch, B., & Xu, P. X. (2004). Eya1 and Six1 are essential for early steps of sensory neurogenesis in mammalian cranial placodes. *Development*, *131*(22), 5561–5572. https://doi.org/10.1242/dev.01437