

**Epigenome-targeted CRISPR screen reveals stage-specific dependencies in
human erythropoiesis**

By

Evrin Goksel

A Dissertation Submitted to the
Graduate School of Health Sciences in Partial
Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
in
Cellular and Molecular Medicine



KOÇ ÜNİVERSİTESİ

Dec 24, 2025

**Epigenome-targeted CRISPR screen reveals stage-specific dependencies in
human erythropoiesis**

Koc University
Graduate School of Health Sciences

This is to certify that I have examined this copy of a doctoral dissertation by

Evrin Goksel

and have found that it is complete and satisfactory in all respects,
and that all revisions required by the final
examining committee has been made.

Committee Members:

Prof. Dr. Ozlem Yalcin (Advisor)

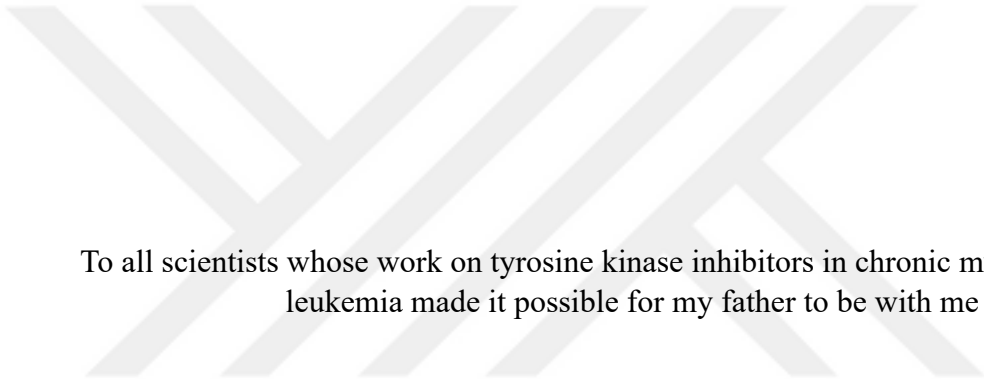
Prof. Dr. T. Tamer Onder

Asst. Prof. Dr. Sercin Karahuseyinoglu

Prof. Dr. Soner Dogan

Prof. Dr. Dilek Telci

Date: 24.12.2025



To all scientists whose work on tyrosine kinase inhibitors in chronic myeloid
leukemia made it possible for my father to be with me today.

ABSTRACT

Epigenome-targeted CRISPR screen reveals stage-specific dependencies in human erythropoiesis

Evrin Goksel

Doctor of Philosophy in Cellular and Molecular Medicine, Dec 24, 2025

Erythropoiesis is a tightly regulated differentiation process that proceeds through discrete cellular states and culminates in terminal maturation and enucleation. While lineage-defining transcription factors are essential to erythroid gene regulation, including the master erythroid transcription factor GATA1, the mechanism of how epigenetic regulators shape cell state-specific transcriptional programs during erythroid maturation remains poorly understood.

In this thesis, we systematically interrogated the epigenetic regulatory landscape of human erythropoiesis through CRISPR/Cas9-based functional screens. This work revealed two principal chromatin regulatory pathways that play central roles in erythroid differentiation: the histone acetyltransferase EP300 and the SMARCB1-containing cBAF (SWI/SNF) chromatin remodeling complex. Using synchronized human erythroid differentiation models, including BEL-A cells and primary CD34⁺ progenitors, we demonstrate that EP300 and SMARCB1 regulate erythropoiesis through distinct, cell state-dependent mechanisms.

We show that EP300 is required at early stages of erythroid differentiation. Loss of EP300 disrupts early erythroid cell state transitions, leading to a broad failure in erythroid gene induction and impaired commitment to the erythroid lineage. In contrast, SMARCB1 is essential during terminal maturation. In primary CD34⁺ hematopoietic stem/progenitor cells, EP300 loss blocked progression at the basophilic–polychromatic transition, while SMARCB1 deficiency significantly reduced erythrocyte output and enucleation efficiency. Furthermore, using single-cell and bulk RNA sequencing, we further validated that EP300 governs differentiation-associated transcriptional programs.

To understand the role of SMARCB1 in terminal erythropoiesis, we assessed chromatin accessibility across the genome and found that SMARCB1 loss does not alter global chromatin accessibility. We also found that GATA1 protein expression does not change with SMARCB1 perturbation. These findings indicated that SMARCB1 does not regulate erythroid maturation through changes in chromatin accessibility or transcription factor expression. To investigate the protein–DNA interactions underlying SMARCB1-dependent erythroid transcriptional programs, we performed Docking & Deamination-Seq (D&D-Seq), a deaminase-based footprinting approach that enables high-resolution mapping of transcription factor occupancy on chromatin. The results indicated that SMARCB1 knockout alters GATA1 occupancy at a subset of erythroid regulatory loci, including reduced binding at the *SNCA* locus and increased binding at *TFR2*. As *SNCA* is robustly induced during normal late-stage erythroid differentiation, these results demonstrate that SMARCB1 is required for proper GATA1-mediated activation of terminal maturation gene programs.

Together, these findings establish EP300 as an early differentiation dependency and SMARCB1/cBAF as a late enucleation dependency, uncovering distinct chromatin-based mechanisms that coordinate sequential stages of human erythropoiesis.

ÖZETÇE

CRISPR taraması ile eritropoezde farklılaşma aşamasına özgün epigenetik faktörlerin belirlenmesi

Evrin Göksel

Hücresel ve Moleküler Tıp, Doktora, 24 Aralık 2025

Eritropoez, farklılaşmanın belirli hücresel evreler boyunca ilerlediği ve terminal olgunlaşma ile enükleasyonla sonuçlanan, sıkı biçimde düzenlenen bir süreçtir. Soy belirleyici transkripsiyon faktörleri eritroid gen düzenlenmesi için vazgeçilmez olsa da, epigenetik düzenleyicilerin eritroid olgunlaşma sırasında hücre evresine özgü transkripsiyonel programları nasıl yönlendirdiği hâlâ yeterince anlaşılmamıştır.

Bu tez çalışmasında, insan eritropoezinin epigenetik düzenleyici yapısı CRISPR/Cas9 tabanlı fonksiyonel taramalar kullanılarak sistematik biçimde incelenmiştir. Çalışmalar, eritroid farklılaşmada merkezi rol oynayan iki temel kromatin düzenleyici yolak ortaya koymuştur: histon asetiltransferaz EP300 ve SMARCB1 içeren cBAF (SWI/SNF) kromatin yeniden düzenleme kompleksi. BEL-A hücreleri ve primer CD34⁺ progenitörler dahil olmak üzere senkronize insan eritroid farklılaşma modelleri kullanılarak, EP300 ve SMARCB1'in eritropoezi birbirinden farklı ve hücre evresine bağlı mekanizmalarla düzenlediği gösterilmiştir.

EP300'ün eritroid farklılaşmanın erken aşamalarında gerekli olduğu gösterilmiştir. EP300 kaybı, erken eritroid evre geçişlerini bozarak eritroid genlerin geniş ölçekte aktive edilememesine ve eritroid soya yönelimin zayıflamasına yol açmaktadır. Buna karşılık, SMARCB1 erken eritroid farklılaşma için daha az kritik olmakla birlikte, terminal olgunlaşma aşamasında vazgeçilmez hâle gelmektedir. Primer CD34⁺ hematopoietik kök/progenitör hücrelerde EP300 kaybı bazofilik–polikromatik geçişte ilerlemeyi durdururken, SMARCB1 eksikliği eritrosit üretimini ve enükleasyon verimliliğini belirgin biçimde azaltmıştır. Ayrıca, tek hücreli ve toplu RNA dizileme analizleri, EP300'ün farklılaşmaya eşlik eden transkripsiyonel programları yönettiğini doğrulamıştır.

SMARCB1'in terminal eritropoezdeki rolünü daha iyi anlamak amacıyla genom genelinde kromatin erişilebilirliği incelenmiş ve SMARCB1 kaybının global kromatin erişilebilirliğini değiştirmedeği görülmüştür. Benzer şekilde, SMARCB1 perturbasyonunun GATA1 protein düzeylerini etkilemediği saptanmıştır. Bu bulgular, SMARCB1'in eritroid olgunlaşmayı kromatin erişilebilirliği ya da transkripsiyon faktörü ekspresyonundaki değişiklikler yoluyla düzenlemediğini göstermektedir. SMARCB1'e bağlı eritroid transkripsiyonel programların temelindeki protein-DNA etkileşimlerini incelemek için, kromatin üzerinde transkripsiyon faktörü bağlanmasını yüksek çözünürlükte haritalamaya olanak tanıyan, deaminaz temelli bir footprinting yöntemi olan Docking & Deamination-Seq (D&D-Seq) kullanılmıştır. Elde edilen sonuçlar, SMARCB1 knockout'unun eritroid düzenleyici bölgelerin bir kısmında GATA1 bağlanmasını değiştirdiğini; özellikle SNCA lokusunda bağlanmanın azaldığını, TFR2 lokusunda ise arttığını göstermiştir. SNCA'nın normal geç evre eritroid farklılaşma sırasında güçlü biçimde aktive edildiği göz önüne alındığında, bu bulgular SMARCB1'in terminal olgunlaşma gen programlarının GATA1 aracılığıyla doğru şekilde aktive edilmesi için gerekli olduğunu ortaya koymaktadır.

Sonuç olarak bu çalışma, EP300'ü erken eritroid farklılaşma için gerekli bir faktör, SMARCB1/cBAF kompleksini ise terminal olgunlaşma ve enükleasyon için kritik bir bağımlılık olarak tanımlamakta ve insan eritropoezinin ardışık evrelerini koordine eden, birbirinden farklı kromatin temelli mekanizmaları ortaya koymaktadır.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis advisor, Prof. Dr. Özlem Yalçın Seyhan, for her guidance, support, and encouragement throughout my doctoral studies. Her scientific insight, patience, and trust have been invaluable to my development as an independent researcher. I am deeply thankful for her mentorship and for the opportunities she provided, which shaped both my scientific perspective and my academic journey.

I would like to thank the members of my thesis committee—Prof. Dr. Tamer Önder, Assoc. Prof. Dr. Serçin Karahüseyinoğlu, for their constructive feedback, insightful discussions, and thoughtful questions, which greatly strengthened this thesis.

I am sincerely grateful to Prof. Dr. Dan Landau for hosting me during my research period in the United States and for providing an intellectually stimulating and supportive scientific environment. Our scientific conversations and his support in exploring my own research questions have been invaluable to both this thesis and my development as a scientist. I am also especially thankful to Dr. Ivan Raimondi, Dr. Catherine Potenski and Dr. Nir Ben Chetrit who served as both a mentor and a close colleague. His guidance consistently helped broaden my scientific perspective, and his support was instrumental in shaping experimental design, troubleshooting, and data analysis throughout this work.

I would like to sincerely thank Dr. Selahattin Can Özcan for their important contributions to this project. Their belief in the idea, their help in shaping the project, and their willingness to think, learn, and explore together with me were essential to this work. Also thank you, for making me laugh every single day. Thank you for every beautiful day we shared together.

When we started to think about this project, there was a night when we stood in front of a whiteboard at midnight, quietly thinking about whether this idea could work. That moment has a special place in my academic journey, as it marked the real beginning of this project. Working closely together, their influence on my doctoral training and on

how I approach science in the future has been lasting, and I truly appreciate their support and mentorship. Thank you, Dr. Selahattin Can Özcan and Dr. Alişan Kayabölen!

I would also like to thank Dr. Jane Frayne and her team for generously providing the BEL-A cell line and for their valuable support. I am grateful to Dr. Tuğba Bağcı Önder for the collaboration on the EPIKOL library, which was essential for the completion of this work. I sincerely appreciate their willingness to collaborate and their contributions to this project.

During my doctoral studies, my education was supported by two fellowships provided by the TÜBİTAK Scientific and Technological Research Council of Turkey, Directorate of Science Fellowships and Grant Programs (BİDEB), under the 2211-A Domestic Graduate Scholarship Program and the TÜBİTAK 2214-A International Research Fellowship Program for PhD Students. This study was conducted using the services and research infrastructure of Koç University and the Koç University Research Center for Translational Medicine (KUTTAM). In addition, this study was supported by the TÜBİTAK 1001 Research Project (Project No. 122Z388).

Another thanks go to Dr. İpek Bulut. Your scientific perspective has always meant so much to me, and in the lab you were truly my extra pair of hands. But more than anything, thank you for sharing every emotion of this PhD journey with me. Above all, İpek, thank you for being exactly who you are.

I would also like to thank Elif Uğurel for guidance and share and willingness to share their experience. Always being a cleaver and calm person in the lab. Thank you.

I would like to sincerely thank Uğur Çerez for his invaluable support with laboratory instrument calibration and technical issues throughout my PhD. I am grateful to Güven Ayvaz for his assistance with consumable purchases and laboratory organization, and to Ergün Başoğlu for his practical solutions and dedication to maintaining laboratory equipment. I also thank Birsu Kölemen for her support in purchasing processes and administrative coordination, and İmren Karaca for her calm and reliable administrative assistance. Finally, I would like to thank Fatma Abla for her kindness and care, and for always making the laboratory a warmer place as much as her warmer tea and coffee. I

am grateful to all of you for being a supportive team and for helping us overcome countless challenges during this long journey.

Also, I extend my heartfelt thanks to my friends Deniz Gülfem Öztürk, Bizim Deniz, Barış Sergi, Hande Özkan, İlayda Yıldız, Elit Altun, Neslihan Cilek and Fatma Öz for their constant encouragement, understanding, and friendship throughout this journey. Their support and presence made the challenging moments more manageable and the successes more meaningful. I would also like to thank Wei-Yu Chi, Sang-Ho Yoon, Ramon Massoni, Jesus Sotelo, Yiyun Lin, Manu Singh, Joe Pelt, Saravanan Ganesan, Foteini Fotopoulou, Levan Mekerishvili, Dennis Yuan, Sri Rajagopalan, Tamara Prieto and Jane Park for their scientific input, generosity in sharing expertise, and for the collaborative and motivating atmosphere they contributed to during this period.

Son olarak, en büyük teşekkürü canım aileme, anneme ve babama, Evren Göksel ve Naci Göksel'e sunmak isterim. Bu tezin var olmasında ve bugün bağımsız bir bilim insanı olarak burada durabilmemde en büyük pay sizlere aittir. Doktora eğitiminin yalnızca bir akademik unvan değil, gerçek anlamıyla bir Doctor of Philosophy olduğunu bana çocukluğumdan itibaren öğrettiniz. Beni bağımsız bir kadın olarak yetiştirdiniz; keşfetmeyi, merak etmeyi, sormaktan ve sorgulamaktan korkmamayı öğrettiniz. Düşüncelerime alan açtınız, hayallerime sınır koymadınız. En zor anlarımda bile sessizce yanımda durdunuz ve bana her zaman güvendiniz. Beni sevgiyle büyüttüğünüz, emeğinizi esirgemediğiniz, yolumu bulmama izin verdiğiniz ve kim olduğuma inandığınız için size teşekkür ederim. Bugün attığım her adımda, keşfettiğim ve keşfedeceğim her yeni bilgide iziniz var. Teşekkür ederim.

TABLE OF CONTENTS

LIST OF TABLES	xiv
LIST OF FIGURES.....	xv
ABBREVIATIONS.....	xvii
INTRODUCTION	1
1.1. Hematopoiesis.....	1
1.1.1. Hematopoiesis and Erythropoiesis	1
1.1.2 Regulation of the Erythropoiesis.....	2
1.1.3 Impaired Erythropoiesis and Enucleation	3
1.2. Epigenetics and Epigenetic Regulation	5
1.2.1 Epigenetics and Lineage Commitment	5
1.2.2 Chromatin Dynamics during Erythroid Differentiation and Enucleation	6
1.3. EP300 and histone acetyltransferases	8
1.3.1 Histone Acetylation and Acetyltransferases.....	8
1.3.2 Role of EP300 in Hematopoietic Stem Cell Biology.....	9
1.4. SWI/SNF (BAF) Complex.....	10
1.4.1 Functions of BAF complex and subunits	10
1.4.2 Roles of BAF complexes in HSC differentiation and erythropoiesis	12
MATERIALS and METHODS	15
2.1. Cell culture experiments and erythroid differentiation.....	15
2.1.1 BEL-A cell culture conditions.....	15
2.1.2 BEL-A cell differentiation.....	15
2.1.3 CD34 ⁺ hematopoietic stem cell culture conditions	16
2.1.4 CD34 ⁺ hematopoietic stem cell differentiation	16
2.2. Chemicals and antibodies	17
2.3. gRNA sequences and gRNA cloning.....	17
2.4. HSC Electroporation	18
2.5. Lentivirus generation and BEL-A cell spinfection	19
2.6. Epigenome targeted CRISPR screen	20
2.6.1 Production of EPIKOL gRNA library	20

2.6.2 Plasmid library sequencing	21
2.6.3 MOI calculation and library transduction of BEL-A-Cas9 cells.....	21
2.6.4 Sample collection and extraction of genomic DNA.....	22
2.6.5 Library preparation and NGS sequencing.....	22
2.6.6 CRISPR screen results analysis with MageCK.....	23
2.7. Flow cytometry experiments	24
2.8. May-Grunwald Giemsa Staining.....	25
2.9. RNA sequencing (RNAseq)	25
2.10. Single Cell RNA sequencing (sc-RNAseq).....	25
2.11. Sanger sequencing of CRISPR edits in HSCs	26
2.12. Bulk Docking & Deamination – Sequencing (D&D-Seq) analysis.....	27
2.13. Single cell Combined D&D-Seq and Multiome analysis	29
2.14. Computational analysis of D&D-Seq	29
2.14. Statistical Analyses	30
RESULTS	32
3.1. Validation of erythroid differentiation conditions required for CRISPR screening	32
3.2. CRISPR screen identifies the epigenetic dependencies of erythroid differentiation	36
3.3. CRISPR screen identifies stage specific epigenetic dependencies of erythropoiesis	39
3.4. Dual-color competition experiments validate CRISPR screen results	44
3.5. EP300 is required for differentiation of BEL-A cells	46
3.6. cBAF/pBAF complex members represent a late-stage erythroid dependency.....	48
3.7. SMARCB1 is required for enucleation stage of erythropoiesis.....	49
3.8. Targeted experiments validate CRISPR screen results in primary HSCs....	50
3.9. scRNA-seq reveals stage specific expression of CRISPR screen hits.....	55
3.10. Six days of BEL-A differentiation reflects a transcriptomic signature of orthochromatic erythroblasts.....	60
3.11. EP300 knock-out results with changes in differentiation associated gene expressions	62
3.12. D&D-Seq infers GATA1 binding events in HSC cells.....	64
3.13. SMARCB1 knock-out affects GATA1 binding in BEL-A cells	65
DISCUSSION.....	73

CONCLUSION	81
BIBLIOGRAPHY.....	83



LIST OF TABLES

Table 1 – Chemicals and antibodies used in the study.....	17
Table 2 – sgRNA sequences for pLentiGuide-puro and pLentiGuide-GFP backbone	18
Table 3 – sgRNA sequences for electroporation of HSC cells	19
Table 4 – Lentiviral plasmids used in the study.....	20
Table 5 – Primer sequences used in CRISPR screen gRNA NGS sequencing libraries	23
Table 6 – Total reads of sequencing results	24
Table 7 – Reagents used in flow cytometry experiments	24
Table 8 – Primer sequences for Sanger sequencing.....	26
Table 9 – Tn5 tagmentation oligonucleotide sequences	29

LIST OF FIGURES

Figure 1 – Overview of Erythropoiesis and the Regulation of the Erythropoiesis.....	3
Figure 2 – Chromatin regulators essential for key hematopoietic lineage decisions.....	6
Figure 3 – Members of three different BAF complexes	12
Figure 4 – Distribution of gRNA counts of produced library and source library	21
Figure 5 – D&D-Seq workflow	27
Figure 6 – Characterization of different differentiation stages of BEL-A cells.....	33
Figure 7 – Cas9 expression does not affect differentiation and enucleation of BEL-A cells	34
Figure 8 – Cas9 activity assay reveals the effectiveness of Cas9 in BEL-A cells	35
Figure 9 – H2B-GFP expression is maintained in pyrenocytes	36
Figure 10 – Experimental design of the epigenome targeted CRISPR screen	37
Figure 11 – CRISPR screen experiments in three biological repeats	37
Figure 12 – Growth and differentiation associated comparisons.....	38
Figure 13 – Volcano plots of stage specific comparisons	40
Figure 14 – Pairwise comparisons of depletions in different cell types	41
Figure 15 – Combined results of top20 hits of all comparisons	42
Figure 16 – Functional characterization of CRISPR screen hits	43
Figure 17 – GO enrichment analysis of CRISPR screen hits	44
Figure 18 – Overview of competition experiments.	44
Figure 19 – EP300 is required for survival in early differentiation stages	45
Figure 20 – Dual color competition experiments for enucleation	46
Figure 21 – EP300 knock-out inhibits differentiation of BEL-A cells.	47
Figure 22 – EP300 knock-out inhibits hemoglobin production in differentiating BEL-A cells	48
Figure 23 – Results of SWI/SNF complex member genes in CRISPR screen	49
Figure 24 – SMARCB1 knock-out inhibits enucleation.....	50
Figure 25 – Erythroid differentiation of primary CD34+ HSCs.....	51
Figure 26 – Experimental flow of HSC experiments.....	52
Figure 27 – Electroporation of HSCs with RNPs and Cas9-GFP.....	52
Figure 28 – TIDE analysis results of sgSMARCB1 induced DNA edits	53
Figure 29 – sgEP300 reduces the differentiation of HSCs	53

Figure 30 – Erythrocyte production of sgNT and sgSMARCB1 transfected differentiated HSCs	54
Figure 31 – Histological assessment of SMARCB1 and EP300 knock-out on HSC differentiation.....	55
Figure 32 – Integrated scRNA-seq analysis of erythroid differentiation induced HSCs	56
Figure 33 – Cell stage specific markers of erythroid cells.....	57
Figure 34 – Stage specific expression patterns of CRISPR screen hits in different cell clusters	58
Figure 35 – Expression of CRISPR screen hits in different cell clusters	58
Figure 36 – Expression of key erythroid transcription factors in different stages	59
Figure 37 – Gene expression changes in erythroid differentiation stages	60
Figure 38 – Differentiation induced gene expression changes in BEL-A cells	61
Figure 39 – Expression of polychromatic and orthochromatic erythroblast markers from scRNAseq in differentiated BEL-A cells	62
Figure 40 – Effect of EP300 knock-out on the transcriptome of undifferentiated BEL-A cells	63
Figure 41 – sgEP300 associated gene expression changes in differentiating BEL-A cells	64
Figure 42 – D&D-Seq identifies GATA1 binding in differentiating HSCs	65
Figure 43 – Bulk D&D-Seq captures GATA1 binding in BEL-A cells	66
Figure 44 – SMARCB1-KO does not alter the GATA1 expression and global gene accessibility profile of the BEL-A cells	67
Figure 45 – SMARCB1 KO associated changes of GATA1 binding	68
Figure 46 – SMARCB1 associated GATA1 binding changes of known GATA1 target genes	69
Figure 47 – Enrichment analysis of SMARCB1-KO associated D&D-Seq hits	69
Figure 48 – Expression patterns of genes that have SMARCB1 associated GATA1 binding changes in different single cell RNA-seq clusters	70
Figure 49 – SNCA expression increases with erythroid differentiation	70
Figure 50 – SMARCB1 associated GATA1 binding events are enriched in erythroid differentiated cells.....	71
Figure 51 – GATA1 D&D edits and SNCA expression are increased with erythroid differentiation.....	72
Figure 52 – Regulation of erythropoiesis by EP300 and SMARCB1	79
Figure 53 – Regulation of enucleation by SMARCB1	80

ABBREVIATIONS

BFU-U	Burst Forming Unit Erythroblast
CFU-U	Colony Forming Unit Erythroblast
CRISPR	Clustered regulatory interspaced palindromic repeats
D&D-Seq	Docking & deamination followed by sequencing
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco's modified Eagle medium
EPO	erythropoietin
EPOR	erythropoietin receptor
GATA	GATA binding protein
GFP	Green fluorescent protein
H2B	Histone 2B
HbF	Hemoglobin F
HbS	Hemoglobin S
HDAC	Histone Deacetylase
HSC	Hematopoietic stem cell
IL-3	Interleukin-3
JAK	Janus kinase
MAPK	Mitogen activated protein kinase
MDS	Myelodysplastic syndrome
MOI	Multiplicity of infection
PCR	Polymerase chain reaction
PI3K	Phosphoinositid-3-kinase
RBC	Red blood cell

SCD	Sickle cell disease
SCF	Stem cell factor
sgRNA	small guide RNA
STAT	Signal transducer and activator of transcription
TU	Transducing units



Chapter 1

INTRODUCTION**1.1. Hematopoiesis***1.1.1. Hematopoiesis and Erythropoiesis*

Hematopoiesis is the production of all blood and immune cells, including erythrocytes, leukocytes (white blood cells), and platelets. Erythropoiesis is specifically defined as the production of red blood cells. It starts with a tiny contingent of pluripotent hematopoietic stem cells (HSCs) with self-renewing capacity, which execute a system of lineage commitment and differentiation. Following the erythroid lineage fate determination, common myeloid progenitors give rise to bipotent megakaryocyte-erythroid progenitors, which can ultimately form megakaryocytes (which produce platelets) or erythrocytes. The erythroid cells downstream of megakaryocyte-erythroid progenitors are erythroid committed and slow-proliferating Burst Forming Unit Erythroblasts (BFU-E) mature to rapidly dividing Colony Forming Unit Erythroblasts (CFU-E), which then undergo terminal differentiation sequentially becoming proerythroblasts, basophilic erythroblasts, polychromatic erythroblasts, and orthochromatic erythroblasts, which enucleate to form reticulocytes (**Figure 1**). Reticulocytes then egress from the bone marrow and mature into erythrocytes in the bloodstream. During terminal differentiation, cells undergo substantial changes that include progressive reduction in cell volume, increase in hemoglobin content, loss of RNA, expulsion of organelles, induction of hemoglobin synthesis, nuclear condensation, and enucleation (**Figure 1**) (Lee et al., 2004).

Erythrocytes, also known as red blood cells (RBCs), are the most abundant cells in the human body, comprising 84% of all cells (Pretini et al., 2019). RBCs are produced in a process, known as erythropoiesis and this process starts in the bone marrow and continues in the blood stream. Immature erythrocytes (reticulocytes) pass through the bloodstream at a rate of 2 million per second (Palis, 2014) and mature into erythrocytes in circulation

over approximately 48 hours (Gifford, Derganc, Shevkoplyas, Yoshida, & Bitensky, 2006). Erythrocytes do not have organelles such as nucleus, ribosomes, mitochondria, which are critical for the survival of cells. Due to this highly specialized structure of erythrocytes, the primary functions of these cells in circulation can be implemented as; (i) hemoglobin builds up in the cell and oxygen is transported to the peripheral tissues, (ii) shapes become biconical, so the cells are more flexible and can live for 120 days in circulation. After a long lifespan in the circulation, damaged RBCs are detected and expelled by the cells of the reticuloendothelial system, primarily by macrophages in the spleen (Schroit, Madsen, & Tanaka, 1985).

1.1.2 Regulation of the Erythropoiesis

Erythropoiesis is regulated by a number of extrinsic and intrinsic factors. Several extrinsic factors regulate proliferation, differentiation, and the survival of erythroid cells by activating several downstream pathways (Caulier & Sankaran, 2022; Sui, Krantz, & Zhao, 2000). IL-3 supports the expansion of early erythroid progenitors such as BFU-E. Stem cell factor (SCF), acting through its receptor c-KIT (CD117), promotes the growth and survival of BFU-Es, CFU-Es, and early erythroblasts (Lewis et al., 1998). Erythropoietin (EPO) is the central cytokine that drives the erythroid maturation, ensuring the survival and proliferation of CFU-Es and later erythroid stages (Nocka et al., 1989). While erythropoietin receptor (EPOR) expression was previously known to get decreased as the cells mature, recent evidence suggests that EPOR may be more broadly expressed in hematopoietic progenitors than previously believed, implying erythroid lineage commitment might occur earlier in development (Caulier & Sankaran, 2022).

Cytokines such as EPO, SCF, and IL-3 converge on several major signaling cascades, including the JAK/STAT, MAPK, and PI3K pathways, which are essential for cell survival and proliferation (Hattangadi, Wong, Zhang, Flygare, & Lodish, 2011). EPO binding induces EPOR dimerization and activates JAK2, leading to phosphorylation and nuclear entry of STAT5, a key regulator of the erythroid transcriptional program (Capellera-Garcia et al., 2016). Excessive activation of these pathways can cause erythrocythemia, typically through somatic JAK2 mutations or less commonly via germline EPOR mutations (Gaspersic, Kristan, Kunej, Zupan, & Debeljak, 2021). Earlier

simplified models suggested these cytokines triggered fixed, linear pathways, but recent work using engineered ligands and natural genetic variation shows that receptor activity and downstream signaling are more dynamic and context-dependent than previously appreciated. Understanding these complexities is crucial, as dysregulation of these pathways contributes to disease and may be targeted therapeutically.

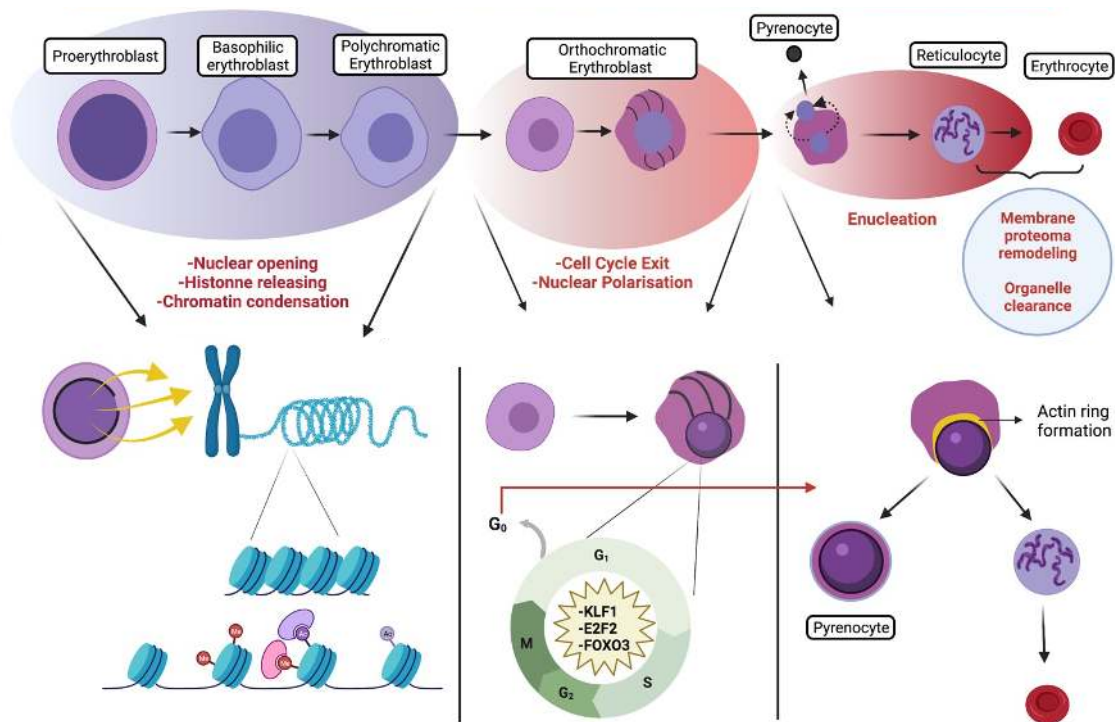


Figure 1 – Overview of Erythropoiesis and the Regulation of the Erythropoiesis
Created in BioRender. ID: SZ2961FT7K

1.1.3 Impaired Erythropoiesis and Enucleation

Anemia is a condition that lacks enough healthy red blood cells to carry adequate oxygen to peripheral tissues. Having anemia, also referred as to low hemoglobin concentrations, causes people to feel fatigue and weak. Anemia affects almost one-third of the population of which children, pregnant women, and older people are the most vulnerable (Chaparro & Suchdev, 2019). Decreased production (ineffective erythropoiesis), increased loss of RBCs (bleeding), and increased destruction (hemolysis) of RBCs are the main factors causing anemia.

Myelodysplastic syndrome (MDS) is a disease of unknown cause that can occur with disorders at different stages of hematopoiesis and has different subtypes (Hasserjian, 2019). Among these subtypes, it is known that the disruptions in the erythroid lineage are more common than the disorders in the myeloid lineage (Flores-Figueroa, Gutierrez-Espindola, Guerrero-Rivera, Pizzuto-Chavez, & Mayani, 1999). It has been shown that the erythropoiesis disorders seen in MDS are caused by disruptions in chromatin condensation and histone release dynamics (Boyault, Sadoul, Pabion, & Khochbin, 2007) and that the increase in GATA1 expression causes erythropoiesis defects in MDS (Fadlallah et al., 2022; Hofer et al., 2021). In recent years, it has been discussed that the disorders in the erythroid lineage in MDS are associated with mutations of various epigenetic factors (Sperling, Gibson, & Ebert, 2017) and that epigenetic arrangements are potentially important in the pathophysiology of MDS (Bond, Lee, & Enjeti, 2020).

β -thalassemia is a hemoglobinopathy caused by mutations in the β -globin gene that results in an imbalance between α - and β -globin chains which unstabilizes α -tetramers within erythroid cells leading to premature cell death within erythropoiesis and causes severe anemia (Centis et al., 2000). Histone deacetylase inhibitors and hypomethylating agents are clinically trialed to increase HbF production in β -thalassemia. The most promising and tolerable inhibitor is Givinostat up to date, which induces HbA and HbF production in erythroid cells at a level comparable to HU and butyrate (Esrick, McConkey, Lin, Frisbee, & Ebert, 2015; Shearstone et al., 2016). While insufficient erythropoiesis has long been recognized as a key factor to anemia in thalassemia due to an accumulation of free α -globin chains (Arlet, Dussiot, Moura, Hermine, & Courtois, 2016), insufficient erythropoiesis in anemia related to sickle cell disease (SCD) has not been completely explored. SCD is a hereditary recessive disorder caused by a point mutation in the β globin gene resulting in a Glu-to-Val substitution at the 6th position of the β globin protein resulting HbS polymerization under hypoxic conditions driving RBC sickling (Pauling, Itano, & et al., 1949). Because of HbS polymerization, the lifespan of sickled RBCs is short in blood circulation, which is a major determinant of chronic anemia in SCD. Although it is generally believed that insufficient erythropoiesis in SCD do not have a remarkable contribution in anemia, hypoxic condition of the bone marrow environment alters the sickling stages and hemoglobin polymerizations (0.1 to 6% O₂) (Mantel et al., 2015). Therefore, cell death occurs in hypoxic conditions because of HbS

polymer formation in the late stages of erythropoiesis characterized by high intracellular hemoglobin concentration (El Hoss et al., 2021). Increased HbF concentration is preferred for SCD treatment due to high survival rate of F-cells in the circulation. Even though the beneficial effect of HbF is not denied in SCD patients, the enucleation rate of HbF cells in bone marrow under hypoxic condition is limited (Bauer, Kamran, & Orkin, 2012). Therefore, better understanding of the regulation of enucleation, will help to elucidate new targets that may improve the lifespan of erythroid cells in diseases.

1.2. Epigenetics and Epigenetic Regulation

1.2.1 Epigenetics and Lineage Commitment

Hematopoietic cell differentiation is regulated by series of cascades such as; histone modifications, transcription factors, vesicle trafficking, noncoding RNAs, cytoskeleton remodeling, autophagic organelle clearance, and protein proteasome-mediated proteolysis (Mei, Liu, & Ji, 2021). Epigenetic complexes determine the HSC fate by shaping the chromatin landscape that lineage defining transcription factors can access, thereby regulating the gene expression programs. As illustrated in **Figure 2**, murine hematopoiesis studies have revealed that distinct epigenetic regulators operate at specific stages of differentiation to control cell fate decisions. For instance, the cBAF complex containing SMARCD2 opens early myeloid enhancers that are rich in CEBP and PU.1 motifs, allowing these transcription factors to engage and initiate the myeloid program (Priam et al., 2017). Without this chromatin opening, CEBP and PU.1 simply cannot reach their targets, leading to skewing of myeloid lineage. Additionally, MLL4 (KMT2D) adds H3K4 marks on early myeloid enhancers to promote myeloid fate specification. These events pinpoint that the myeloid cell fate is tightly regulated by epigenetic programming. Moreover, SET1 complex (SETD1A/B) reinforces the erythroid identity by stabilizing H3K4me3 marks at key erythroid TF such as GATA1 and KLF1 (Strang, Golde, & Giasson, 2019). As a result, GATA1 and KLF1 expression increase, promoting the erythroid differentiation. Meanwhile, repressive complexes like NuRD, ISWI, and SETDB1/ATF7IP actively suppress the inflammatory AP-1–driven enhancer modules that would otherwise push HSCs toward excessive granulopoiesis (Ding et al., 2023; Lara-Astiaso et al., 2023; M. Li et al., 2025; Subudhi, Subudhi, Patro, & Sitaram, 2022). Altogether, these examples show that epigenetic mechanisms control the lineage

commitment by directing transcription factor binding, enabling specific genes to be activated and determining the fate an HSC commits.

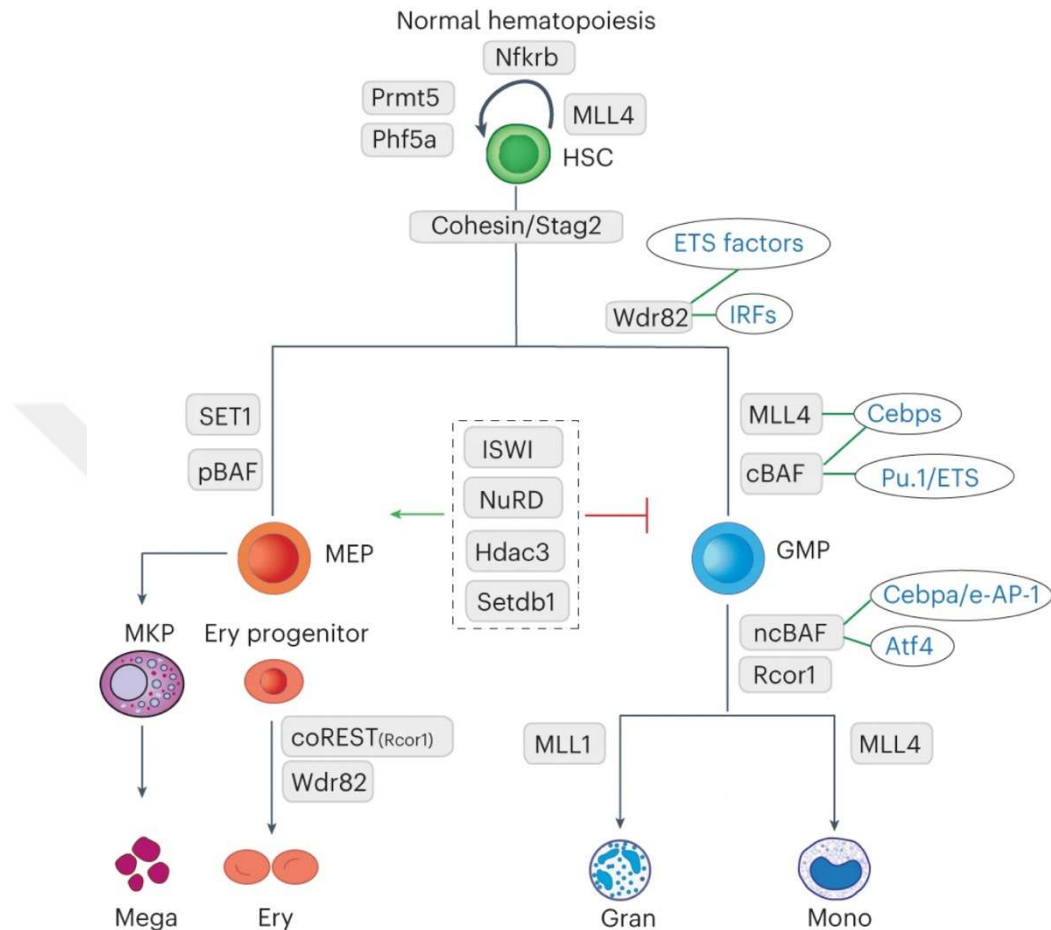


Figure 2 – Chromatin regulators essential for key hematopoietic lineage decisions

(Adopted from Lara-Astiaso et al., 2023)

1.2.2 Chromatin Dynamics during Erythroid Differentiation and Enucleation

Terminal erythropoiesis requires tightly coordinated chromatin condensation, cytoskeletal remodeling, and nuclear extrusion (Zhao, Yang, & Ji, 2016). That shows that this differentiation process would be under control by strong epigenetic modifications. Several histone deacetylases as HDAC2, HDAC5, and HDAC6 have been previously shown to regulate contractile actin ring formation, mDia2 recruitment, and chromatin compaction, all of which are essential for efficient enucleation (Ji, Yeh, Ramirez, Murata-Hori, & Lodish, 2010; X. Li et al., 2017; Y. Wang et al., 2021). Additional chromatin modifiers, including DOT1L (H3K79 methylation) and Setd8 (H4K20me1), are critical

for chromatin structure in erythroblasts; loss of these enzymes results in impaired nuclear condensation, disrupted the nuclear membranes, and reported to cause embryonic anemia in mice (Feng et al., 2010; Malik, Lillis, Couch, Getman, & Steiner, 2017). N6-methyladenosine (m6A) mRNA methyltransferase complex has been reported to regulate erythropoiesis in erythroleukemia cells (Kuppers, 2019). These studies collectively demonstrate that chromatin architecture must be dynamically remodeled for erythroid maturation to proceed.

Despite these recent advances, current knowledge remains incoherent. Most studies have focused on single enzymes or relied on chemical inhibitors, which cannot distinguish direct chromatin-dependent effects from indirect consequences such as apoptosis or cell-cycle arrest. Moreover, although transcription factors such as GATA1 and KLF1 are known to drive erythroid gene expression, it is still unclear which epigenetic complexes shape the chromatin landscape that enables their binding during terminal maturation, particularly during the final transition from condensed nucleus to enucleation.

A recent genome-wide CRISPR screening study, which was conducted in human erythroid progenitors, identified hundreds of genes required for terminal erythroid differentiation (Myers et al., 2025). Although this study provides a broad information for essential regulators of erythropoiesis, it does not specifically address how chromatin regulators coordinate the transcription factor binding, control chromatin accessibility, or mechanistically influence the enucleation process. Thus, the epigenetic logic that links chromatin remodeling to terminal erythroid maturation remains unresolved. In particular, the timing and mechanisms by which epigenetic complexes modulate GATA1 binding and nuclear condensation remain unknown and there is no epigenetic-focused approach that has been applied in physiologically normal erythroid cells to identify specific epigenetic regulators of terminal erythropoiesis. Therefore, how chromatin regulators coordinate transcription factor access, cytoskeletal remodeling, and nuclear extrusion stages remains a major unresolved question that current study directly addresses.

1.3. EP300 and histone acetyltransferases

1.3.1 *Histone Acetylation and Acetyltransferases*

Histone acetylation is a key epigenetic modification that plays a central role in regulating the chromatin structure and gene expression. This process involves the addition of acetyl groups to lysine residues located on the N-terminal tails of histone proteins, primarily histones H3 and H4 (Verdone, Caserta, & Di Mauro, 2005). Acetylation neutralizes the positive charge of lysine residues, weakening the electrostatic interaction between histones and DNA. As a result, chromatin adopts a more relaxed or open conformation, which facilitates the recruitment of transcriptional machinery and promotes gene activation.

The acetylation of histones is dynamically regulated by two opposing enzyme families: histone acetyltransferases (HATs) and histone deacetylases (HDACs). HATs catalyze the addition of acetyl groups from acetyl-CoA tail onto histone substrates, whereas HDACs remove these groups, leading to chromatin compaction and transcriptional repression. Although originally characterized for their roles in histone modification, many HATs also acetylate non-histone proteins, thereby influencing diverse biological processes such as DNA repair, signaling pathways, and cell cycle regulation. For example, EP300 and CBP acetylate the tumor suppressor p53, thereby enhancing its stability and transcriptional activation during DNA damage responses (Ito et al., 2001). Similarly, these HATs acetylate the NF- κ B subunit RelA, modulating its nuclear retention and transcriptional activity in inflammatory signaling pathways (Calao, Burny, Quivy, Dekoninck, & Van Lint, 2008). Other examples include the acetylation of E2F1, which promotes its role in cell-cycle control and apoptosis, and the modification of HIF-1 α , which regulates its stability and activity under hypoxic conditions (Fang et al., 2025; Geng et al., 2012; Manickavinayaham et al., 2019). These examples highlight that histone acetyltransferases function as broad regulators of cellular physiology through acetylation of both histone and non-histone substrates.

1.3.2 Role of EP300 in Hematopoietic Stem Cell Biology

EP300 is an essential transcriptional coactivator and histone acetyltransferase with profound regulatory functions in hematopoietic stem cell (HSC) biology. Through its capacity to acetylate histones and numerous transcription factors, EP300 integrates external signals with transcriptional programs that govern HSC maintenance, quiescence, self-renewal, and lineage differentiation.

Several studies have demonstrated that EP300 is required for proper HSC homeostasis. Loss-of-function models reveal that EP300 deficiency impairs long-term repopulating capacity, disrupts quiescence, and skews the balance between self-renewal and differentiation (S. Wang et al., 2021). These effects are largely attributed to the role of EP300 in activating gene networks that maintain stemness while repressing inappropriate differentiation pathways. EP300 also participates in the regulation of cell-cycle genes, metabolic pathways, and stress-response programs that are critical for HSC survival in the bone marrow niche (C. Wang, Liao, Wang, Hu, & Xu, 2022).

In addition to its canonical HAT activity, EP300 facilitates protein–protein interactions through its multiple functional domains, enabling the formation of transcriptional complexes with key hematopoietic regulators such as GATA2, RUNX1, and TAL1/SCL (Heshusius et al., 2022; Z. Wang et al., 2021; Zheng et al., 2014). By acetylating these transcription factors, EP300 modulates their stability, DNA-binding affinity, and transcriptional activity, thereby influencing lineage specification during hematopoiesis.

The importance of EP300 in hematopoiesis extends to stress response and leukemogenesis. Under conditions such as inflammation or oxidative stress, EP300 helps to orchestrate adaptive transcriptional responses that allow HSCs to maintain functionality (Krosel et al., 2023). Conversely, mutations or dysregulation of EP300 have been associated with hematologic malignancies, including acute myeloid leukemia (AML) and myelodysplastic syndromes, where altered acetylation patterns contribute to aberrant gene expression and impaired differentiation (Blobel, 2000; Nicosia et al., 2023).

Overall, EP300 acts as a central epigenetic regulator that links environmental cues to chromatin dynamics, ensuring proper maintenance and functional integrity of the hematopoietic stem cell compartment.

1.4. SWI/SNF (BAF) Complex

1.4.1 Functions of BAF complex and subunits

The compaction of DNA into chromatin limits the ability of TFs to reach their specific binding sites. ATP-dependent chromatin-remodeling complexes help to promote TF binding by sliding, reorganizing, or removing nucleosomes, thereby creating accessible chromatin at promoters, enhancers, and insulators (Clapier, Iwasa, Cairns, & Peterson, 2017; Koehler, Merchant, & Schatz, 1999; Panier & Boulton, 2014). In mammals, these remodelers are classified into four main families according to their ATPase subunits: SWI/SNF (also referred to as BRG1/BRM-associated factor, BAF), imitation switch (ISWI), chromodomain helicase DNA-binding (CHD), and INO80 (Atkins et al., 2013; Indio do Brasil et al., 2013; Nina et al., 2013). Although different remodelers may occupy the same genomic regions, each family also displays selectivity for particular loci and influences the binding of specific TFs.

The mammalian SWI/SNF (mSWI/SNF) complex is a large, multi-subunit assembly encoded by 29 genes. It forms three distinct subcomplexes—canonical BAF (cBAF), polybromo-associated BAF (PBAF), and noncanonical BAF (ncBAF)—all of which incorporate one of two mutually exclusive ATPase subunits, BRG1 (SMARCA4) or BRM (SMARCA2) (He et al., 2020; Kadoch & Crabtree, 2015; **Figure 3**). Although they share the core subunits and rely on the same ATPase engine, they differ in their unique subunits, chromatin preferences, and biological functions. cBAF contains defining components such as ARID1A or ARID1B and SMARCB1, and it predominantly functions at enhancer regions (Mashtalir et al., 2020; X. Wang et al., 2017). It plays essential roles in developmental gene regulation, enhancer activation, and the control of lineage-specific transcription. PBAF, which incorporates ARID2, PBRM1, and BRD7, tends to occupy promoter regions enriched in histone acetylation (Chandrasekaran & Thompson, 2007; Charlop-Powers, Zeng, Zhang, & Zhou, 2010; Nakayama et al., 2017).

This complex contributes to initiation of the transcription and often exhibits tumor-suppressive activity. In contrast, ncBAF includes BRD9 and GLTSCR1/GLTSCR1L and lacks several structural subunits found in cBAF and PBAF, including SMARCB1. ncBAF is generally enriched at promoter-proximal regions and supports housekeeping gene expression and proliferative or oncogenic programs (Fiterman & Raz, 2019; Meng et al., 2019; Schick et al., 2019). These functional distinctions arise from the combinatorial assembly of their subunits, which directs each complex to distinct genomic elements.

All three complexes contain one of two mutually exclusive ATPase subunits, SMARCA4 (BRG1) or SMARCA2 (BRM). These ATPases provide the mechanical force required to remodel chromatin by hydrolyzing ATP and repositioning or ejecting nucleosomes. SMARCA4 is more widely expressed, has stronger catalytic activity, and is frequently mutated in cancer (Bai et al., 2013; Hodges, Kirkland, & Crabtree, 2016; Reisman, Sciarrotta, Wang, Funkhouser, & Weissman, 2003). SMARCA4 loss leads to severe defects in enhancer regulation and is associated with aggressive tumor phenotypes (Mardinian, Adashek, Botta, Kato, & Kurzrock, 2021). SMARCA2, although being a paralogue, is often epigenetically silenced rather than mutated, and can compensate for SMARCA4 loss; this paralogous relationship underlies synthetic lethal interactions in SMARCA4-deficient cancers (Glaros et al., 2007). The identity of the ATPase subunit influences the efficiency of chromatin remodeling, the transcriptional programs activated, and the specific roles of each SWI/SNF subcomplex. Because nucleosome movement is the essential biochemical activity of these complexes, the ATPase subunit is central to their function.

SMARCB1 is a core component of cBAF and PBAF but is absent in ncBAF. It acts as both a structural stabilizer and a key biological regulator (Nakayama et al., 2017). Structurally, SMARCB1 helps anchor the ATPase module to the nucleosome and is required for proper assembly and integrity of cBAF and PBAF (He et al., 2020). Loss of SMARCB1 destabilizes these complexes and shifts the cellular balance toward ncBAF, which relies heavily on BRD9 for structural cohesion (Langer, Ward, & Archer, 2019; Sima et al., 2019). Biologically, SMARCB1 serves as an important tumor

suppressor. Its loss impairs typical enhancer activity, disrupts lineage-specific transcription, and allows super-enhancers to dominate the gene expression programs (Langer et al., 2019). This reorganization supports oncogenic transcriptional states and contributes to the development of malignancies such as malignant rhabdoid tumors and synovial sarcoma (Clark et al., 1994). SMARCB1 also plays a key role in opposing polycomb-mediated repression, and its absence leads to increased PRC2 activity and aberrant gene silencing (Nakayama et al., 2017). Overall, SMARCB1 is essential for both the stability and the proper gene-regulatory functions of SWI/SNF complexes.

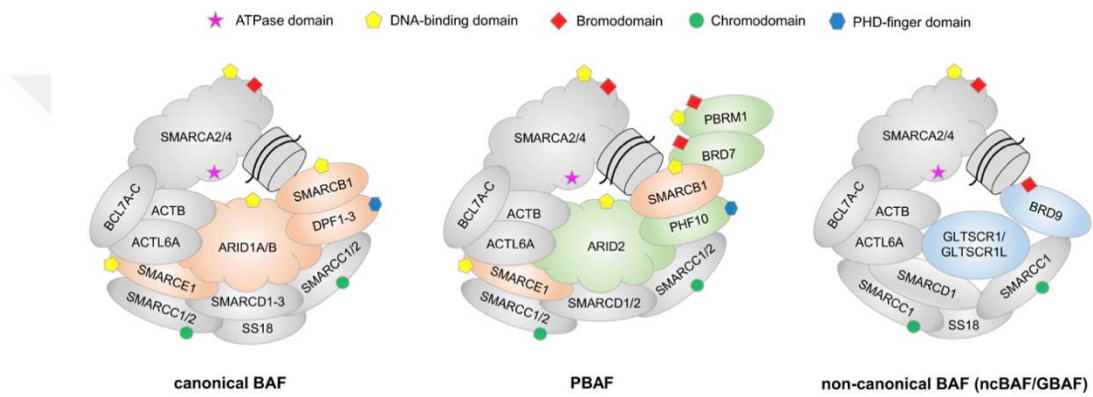


Figure 3 – Members of three different BAF complexes

SWI/SNF subunits are shown in grey, BAF-specific components absent from ncBAF are highlighted in orange, PBAF-specific elements in green, and ncBAF-restricted subunits in blue (Adopted from Wanior et al., 2021).

1.4.2 Roles of BAF complexes in HSC differentiation and erythropoiesis

SWI/SNF (BAF) chromatin-remodeling complexes regulate gene expression in hematopoiesis by physically repositioning the nucleosomes to create the accessible chromatin architecture required for TF binding. As ATP-dependent remodelers, SWI/SNF complexes generate nucleosome-depleted regions at promoters and enhancers, enabling key hematopoietic TFs to engage their regulatory motifs and activate lineage-specific transcriptional programs. HSCs, SWI/SNF activity preserves an open enhancer landscape that supports quiescence, metabolic stability, and genomic integrity (Wanior, Kramer, Knapp, & Joerger, 2021). BRG1-containing BAF complexes counteract nucleosome encroachment at stemness-associated enhancers and maintain the regulatory flexibility essential for long-term self-renewal. Consistent with this, disruption of core BAF components—including SMARCA4/BRG1 or ARID1A—causes enhancer closure,

loss of HSC quiescence, diminished long-term repopulating capacity, and broad defects in myeloid and lymphoid differentiation (Helming et al., 2014).

In AML, pharmacological inhibition of SWI/SNF ATPases (BRM014) reduces the chromatin accessibility at PU.1-bound enhancers, decreases PU.1 occupancy, and redirects PU.1 from enhancers to promoter-proximal sites (Chambers et al., 2023). Genetic knockdown of SMARCA4/SMARCA2 or BRM014 treatment similarly suppresses MYC expression and promotes leukemic differentiation (Rago et al., 2022; Shi et al., 2013). Subunit-specific roles within BAF complexes are also very important: loss of ARID1A leads to widespread enhancer decommissioning and profound hematopoietic failure, whereas ARID1B deficiency causes a milder phenotype but selectively impairs stress-induced myelopoiesis. Combined ARID1A/ARID1B deletion abolishes BAF remodeling activity, resulting in global nucleosome compaction, loss of enhancer–promoter communication, and rapid bone marrow failure (Madan et al., 2023). Collectively, these findings highlight that SWI/SNF complexes regulate hematopoiesis at multiple levels, and deficiencies in individual subunits disrupt lineage fidelity, predispose to differentiation bias, and contribute to leukemogenesis.

Additionally, recent work illustrates how BAF subcomplexes distinctively shape erythroid gene regulation. A previous study (Bagchi et al., 2025) demonstrated that the BRG1-containing ncBAF complex selectively activates HBG (fetal globin) and HBD (minor adult β -like globin) in adult-type erythroblasts. Using CRISPR perturbations, they identified BRD9 and BAF60A as ncBAF-specific components required for this regulatory activity. Loss of these subunits reduced HBG and HBD transcription while accelerating terminal erythroid maturation, suggesting that ncBAF establishes a chromatin environment permissive for fetal and minor β -like globin expression. This work reveals an additional layer of lineage-specific BAF function and underscores how the modular composition of SWI/SNF complexes enables precise control of hematopoietic gene expression programs.

Erythropoiesis is a tightly regulated differentiation process that progresses through well-defined cellular stages and culminates in terminal maturation and enucleation. While

lineage-determining transcription factors are essential for establishing and maintaining erythroid identity, the contribution of epigenetic regulators to stage-specific transcriptional control during erythroid maturation remains incompletely understood.

The hypothesis of this thesis is that epigenetic regulation during human erythropoiesis regulates erythrocyte maturation in a cell stage-dependent context, rather than through a uniform regulatory mechanism across differentiation. Moreover, this work proposes that erythroid lineage differentiation is driven by distinct chromatin regulatory factors acting during early differentiation and during terminal maturation and enucleation.

To further address this hypothesis, we employed a epigenome focus CRISPR/Cas9-based functional screening strategy to systematically interrogate stage-specific epigenetic dependencies during human erythroid differentiation. To understand epigenetic dependencies associated with terminal erythroid differentiation, we established experimental strategies to separate nucleated erythroid cells from their expelled nuclei, referred to as pyrenocytes. This separation allowed cell state-resolved analyses, enabling the direct comparison of genetic and epigenetic requirements between nucleated erythroid precursors and enucleated erythroid cells.

Chapter 2

MATERIALS and METHODS**2.1. Cell culture experiments and erythroid differentiation***2.1.1 BEL-A cell culture conditions*

BEL-A (Bristol Erythroid Line-A) is an adult human erythroid cell line, immortalized predominantly at the proerythroblast stage of erythroid differentiation using inducible HPV16 E6 and E7 (Kurita et al., 2013). The erythroid cells were differentiated from bone marrow CD34⁺ cells. BEL-A cell line was created by Professor Jan Frayne, Professor David Anstee and Dr Kongtana Trakarnsanga at Bristol University (Kurita et al., 2013). Expansion Medium: StemSpan SFEM containing 50 ng/ml stem cell factor (SCF), 3 U/ml erythropoietin (EPO), 1 µM dexamethasone (DEXA) and 1 µg/ml doxycycline (dox). BEL-A cells were maintained in a tissue culture incubator at 37 °C with 5% CO₂ in the expansion medium at a density of $1-1.5 \times 10^5$ cells/ml. Cell culture media was changed every day or every other day for maintenance.

2.1.2 BEL-A cell differentiation

Erythroid Differentiation Culture:

Primary medium: Iscove's medium containing 3% (v/v) human AB serum, 2% (v/v) fetal calf serum, 10 µg/ml insulin, 3 U/ml heparin, 3 U/ml EPO, 200 µg/ml transferrin, 10 ng/ml SCF and 1 ng/ml IL-3 and 1 U/ml penicillin/streptomycin.

Tertiary medium: Iscove's medium containing 3% (v/v) human AB serum, 2% (v/v) fetal calf serum, 10 µg/ml insulin, 3 U/ml heparin, 3 U/ml EPO, 500 µg/ml transferrin and 1 U/ml penicillin/streptomycin.

To induce the erythroid differentiation, BEL-A cells were transferred to the primary medium containing 1 µg/ml doxycycline and maintain for 4 days (Day 0: 2×10^5 cells/ml (Primary + dox), Day 2: 3.5×10^5 cells/ml (Primary + dox), Day 4: 5×10^5 cells/ml (Primary - dox) and Day 6: 5×10^5 cells/ml (Primary - dox). Cells are counted every other day and maintained at the correct cell density. On day 8, cells are transferred to the tertiary medium. Count the cells every other day and perform total medium changes to obtain a final concentration at 1×10^6 cells/ml. K562 was used to optimize lentiviral transduction efficiency. Cells were resuspended in pre-warmed (37°C) RPMI (Invitrogen Ltd) containing 10% (v/v) fetal bovine serum (Invitrogen). Cells were incubated in a tissue culture incubator at 37°C with 5% CO₂ and counted using and maintained at 1×10^6 cells/mL by total media changes every two to three days.

2.1.3 CD34⁺ hematopoietic stem cell culture conditions

Human CD34⁺ HSPCs, isolated from the mobilized peripheral blood of healthy adult donors, were obtained from the Weill Cornell Medicine Blood Bank (New York City, USA). Cells were thawed and cultured in StemSpan SFEM II medium (Stemcell Technologies, 09655), supplemented with StemSpan™ CD34⁺ Expansion Supplement (10X) (Stemcell Technologies, 02691). After 2 days of culture, the cells were transferred to a three-phase erythroid culture medium to initiate the erythroid differentiation.

2.1.4 CD34⁺ hematopoietic stem cell differentiation

Erythroid differentiation of primary HSPCs was performed in a three-phase culture system. Cells were maintained at a concentration of 200,000–1,000,000 cells/mL, and the medium was changed every 2 days to maintain the cell concentration within this range.

Phase I (day 0-8): IMDM (Sigma, FG0465) supplemented with 3% human AB serum (Sigma-Aldrich, H4522), 2% fetal bovine serum (Thermo, 10437028), 1% penicillin/streptomycin (Life Technologies, 15140-122), 10 µg/mL insulin (Sigma-Aldrich, 9278), 3 IU/mL heparin (Sigma-Aldrich, H3149), 200 µg/mL holo-transferrin (Sigma-Aldrich, T0665), 10 ng/mL stem cell factor (SCF) (Peprotech, 300-07), 3 IU/mL erythropoietin (EPO) (Stemcell Tech, 78007) and 5 ng/mL IL-3 (Peprotech, 200-03).

Phase II (days 8-13): IL-3 was removed from the medium.

Phase III (days 14-21): IL-3 and SCF were removed, and the holo-transferrin concentration was increased to 1 mg/ml.

2.2. Chemicals and antibodies

The chemicals and antibodies used in the study are listed in **Table-1**.

Table 1 – Chemicals and antibodies used in the study

Chemical Name	Catalog number
Puromycin	Sigma Aldrich, P8833
Hygromycin	GoldBio, H-270-1
Blasticidin	Thermo Fisher Scientific, R210001
Ampicilin	GoldBio, A-301-5
GATA1 antibody	Abcam, ab11852
Cas9 antibody	Cell signaling, 14697
EP300 antibody	Active motif, 61401
SMARCB1 antibody	Proteintech, 20654-1-AP

2.3. gRNA sequences and gRNA cloning

The sgRNA sequences were generated using Benchling (<https://www.benchling.com>) and are provided in **Table 2**. Complementary sgRNA oligos were first phosphorylated

with T4 PNK (New England Biolabs, M0201S) and subsequently annealed. In parallel, the pLentiGuide-Puro, pLentiGuide-Puro-GFP, pLentiGuide-Puro-mCherry was digested with BsmBI, treated with Antarctic Phosphatase (NEB, M0289) to remove residual phosphates, and purified using the NucleoSpin Gel and PCR Cleanup kit (Macherey-Nagel, 740609). Ligation of the phosphorylated oligos into the linearized backbone was carried out using T4 DNA ligase (NEB, M0202S) with 200 ng of vector and a 7:1 insert–vector molar ratio, incubating at 25°C for 2 hours followed by an overnight incubation at 16°C. Bacterial transformation, plasmid isolation, and lentiviral particle generation.

Table 2 – sgRNA sequences for pLentiGuide-puro and pLentiGuide-GFP backbone

Oligonucleotide name	Sequence
SMARCB1-g1-F	caccGCAGATCGAGTCCTACCCCA
SMARCB1-g1-R	aaacTGGGGTAGGACTCGATCTGC
EP300-g1-F	caccgCTGTAATAAGTGGCATCACG
EP300-g1-R	aaacCGTGATGCCACTTATTACAG
EP300-g2-F	caccGTGGCACGAAGATATTACTC
EP300-g2-R	aaacGAGTAATATCTTCGTGCCAC
DPF2-g1-F	caccgAGAGGCGAGCATTGTAATTG
DPF2-g1-R	aaacCAATTACAATGCTCGCCTCT
DPF2-g2-F	caccGAAGATACTCCCAAGCGTCG
DPF2-g2-R	aaacCGACGCTTGGGAGTATCTTC
DPF3-g1-F	caccgACAGCCGTGAGTTGTAACTC
DPF3-g1-R	aaacGAGTTACAACCTCACGGCTGT
DPF3-g2-F	caccgAGCGCTACAAGAACCGACCG
DPF3-g2-R	aaacCGGTGCGTTCTTGTCGCGCT
GFP-g1-F	caccGGGCGAGGAGCTGTTACCG
GFP-g1-R	aaacCGGTGAACAGCTCCTCGCCC
GFP-g2-F	caccGCTGCACGCCGTAGGTCAGGG
GFP-g2-R	aaacCCCTGACCTACGGCGTGCAGC

2.4. HSC Electroporation

Primary human CD34⁺ hematopoietic stem and progenitor cells were genome edited using CRISPR–Cas9 ribonucleoprotein (RNP) electroporation. sgRNAs were synthesized with terminal chemical modification (2'-O-methyl-3'-phosphorothioate at three terminal nucleotides of 5' and 3' ends) to enhance stability and were resuspended to 100 μ M immediately prior to use. sgRNA sequences are listed on **Table-3**. Cas9 sgRNA RNP complexes were assembled in P3 Primary Cell Nucleofactor Solution and

incubated at room temperature for 20 minutes to allow efficient complex formation. CD34⁺ cells were washed with PBS, counted, and resuspended in P3 solution, using approximately 1×10^5 cells per electroporation reaction. For each condition, 17.5 μ L of the cell suspension was combined with 7.5 μ L of the pre-assembled RNP mix and transferred into a 20 μ L Nucleocuvette. Electroporation was performed on a Lonza 4D-Nucleofector X Unit using program EO-100, which is optimized for primary HSCs. Immediately after pulsing, cells were transferred into pre-warmed StemSpan SFEM II medium supplemented with CD34⁺ Expansion Supplement and 100 nM UM171 and cultured for 12-24 hours to allow recovery and editing. After 24 hours GFP⁺ cells were sorted with FACS Aria (Becton Dickinson). Genomic DNA was collected after 5 days of post-nucleofection to assess editing efficiency using targeted amplicon sequencing.

Table 3 – sgRNA sequences for electroporation of HSC cells

Gene	Guide Sequence	Target Region
SMARCB1 g1	GGCUUUUAAACAGGGUCACAC	GGCTTTTAACAGGGTCACAC
SMARCB1 g2	UGUUUGUCUGUUACAGAUCA	TGTTTGTCTGTTACAGATCA
EP300 g1	AGGAAAUAUGGCUUGGACGC	AGGAAATATGGCTTGGACGC
EP300 g2	CCGGCGUAGGAAAUAUGGCU	CCGGCGTAGGAAATATGGCT
Non targeting gRNA	GCGAGGUACUUGAAGUACGU	GCGAGGT ACTTGAAGTACGT

2.5. Lentivirus generation and BEL-A cell spininfection

To produce the lentiviral particles, 293T cells were seeded to 10-cm petri dishes. For transfection, 2500 ng of the transfer vector, 2250 ng of the packaging vector (psPAX2), and 250 ng of the envelope vector (pCMV-VSV-G) were mixed with 15 μ L PEI (Polysciences, 23966) diluted in DMEM. The cells were then transfected with the prepared mixture. The viral supernatant was collected at 48- and 72-hours post-transfection and concentrated 100-fold using 50% (w/v) PEG 8000 (Sigma, P2139).

For the spin-infection, 2×10^5 BEL-A cells were seeded and infected with the lentiviral particles by centrifugation. The virus was supplemented with 8 μ g/ml polybrene and 10 mM HEPES buffer. Spin-infections were carried out at 900xg for 2 hours at room temperature in 12-well plates. To select for transduced cells, 8 μ g/ml blasticidin or 1

µg/ml puromycin was added 48-hours post-infection and maintained for 7 days to generate stable cell lines.

The transfer plasmids used in the study and generated cell lines are listed in **Table-4**.

Table 4 – Lentiviral plasmids used in the study

Plasmid Name	Plasmid Source	Cell line name
Cas9-blast	Addgene; 73310	BEL-A-Cas9
pCMV-VSV-G	Addgene; 14888	Lenti virus production
psPAX2	Addgene; 12260	Lenti virus production
pDECKO_mCherry-sg-AAVS1	Addgene; 140109	BEL-A-Cas9 mCherry-sgAAVS1
pDECKO_mCherry-sg-NT	Addgene; 141212	BEL-A-Cas9 mCherry-NT
Lenti-H2B-GFP	Addgene; 51004	BEL-A-Cas9-H2B-GFP
Lentiguide-GFP	Addgene; 185473	BEL-A-Cas9-sgRNA-GFP
RRL-GFP	Gift from Onder Lab	BEL-A-Cas9-RRL-GFP

2.6. Epigenome targeted CRISPR screen

2.6.1 Production of EPIKOL gRNA library

Epigenome targeted CRISPR library (EPIKOL) was a gift from Dr. Tugba Bagci Onder. Plasmid pools coding gRNAs were first electroporated to the electrocompetent *E. coli* (Lucigen, 60242) cells and propagated on 35-cm BioAssay dishes (Thermo, 240835) containing LB Agar for 16 hours at 30°C. Then, colonies were collected by washing the plates with cold LB Broth solution. Plasmids were isolated using a commercial Midiprep kit (Macharey Negel, 740410.50). sgRNA production coverage was calculated by counting colonies on 1:100 dilution plates.

2.6.2 Plasmid library sequencing

After isolating the plasmid pooled libraries, first, the quality of library production was determined by next-generation sequencing. Plasmids were PCR amplified by using a primer couple (details of primer sequences are listed at Section 2.7.5) that generates Illumina chemistry compatible sequencing libraries and sequenced on NovaSeq 6000 (Illumina). gRNA sequence distributions were then compared with the source library using MageCK algorithm (**Figure 4**).

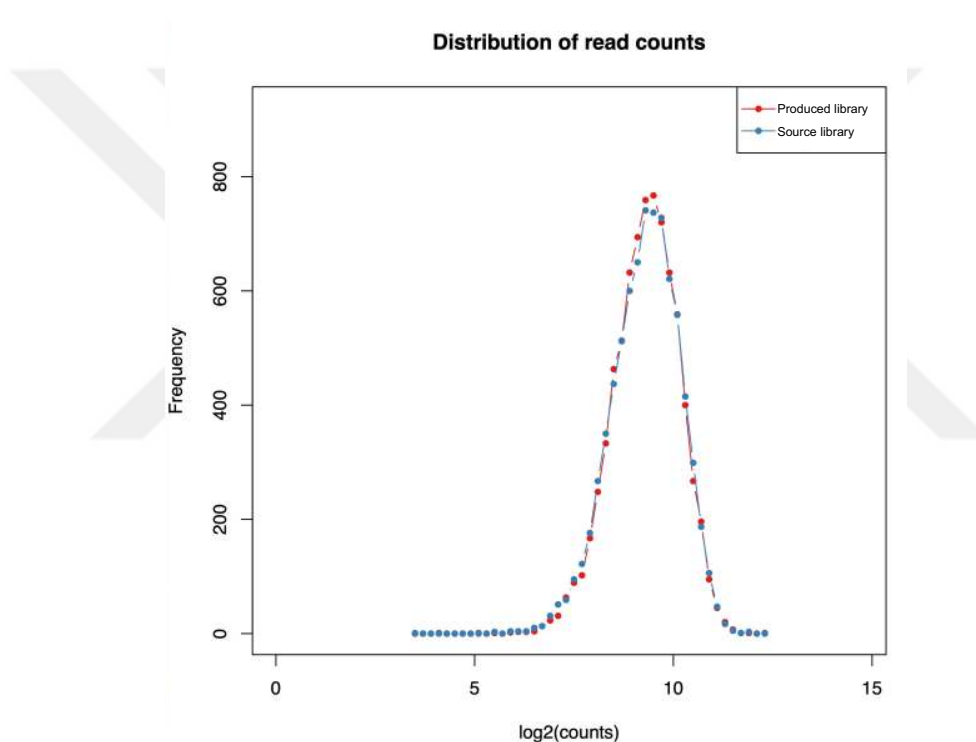


Figure 4 – Distribution of gRNA counts of produced library and source library

2.6.3 MOI calculation and library transduction of BEL-A-Cas9 cells

After the lentiviral particles were produced in HEK293-T cells as described in Section 2.3, lentiviral transducing units (TU) were calculated using 100.000 cells and various amounts of pooled lentivirus. After 24 hours of puromycin selection, multiplicity of infection (MOI) was calculated by the ratio of surviving cells, and TU was calculated with the following formula:

$$TU/mL = \frac{(Number\ of\ Transduced\ Cells \times MOI)}{Volume\ of\ Lentiviral\ Stock\ in\ mL}$$

After the TU determination of the lentivirus stock, BEL-A-Cas9 cells were transduced with MOI:0.6 for CRISPR screening experiments.

2.6.4 Sample collection and extraction of genomic DNA

EPIKOL library-transduced BEL-A-Cas9 H2B-GFP cells were collected from 4 different samples as follows: (i) Initial library transduced cells (Time0), (ii) Undifferentiated erythroblasts (+dox), (iii) Differentiated nucleated cells (orthochromatic erythroblast enriched population), and (iv) enucleated pyrenocytes. Differentiation was induced with removal of the doxycycline and following the BEL-A differentiation protocol. During the cell culture, a library coverage of 300X was maintained by counting the cells during cell passaging. Differentiated GFP+ cells were separated by their size on a flow cytometer and collected by cell sorting. Experiments were conducted in three separate biological replicates.

Genomic DNA from the samples were extracted by using a commercial kit (Macharey Negel, 740952), following the protocol. Briefly, frozen cell pellets were lyzed, and processed, and the DNA concentrations were measured by a Nanodrop spectrometer (Thermo).

2.6.5 Library preparation and NGS sequencing

CRISPR screen libraries for NGS sequencing were prepared by a two-step PCR reaction, described before by Yedier-Bayram et al. (Yedier-Bayram et al., 2022). Briefly, in the first reaction, the lentiviral region was amplified by using external primers. In the second reaction, Illumina sequencing regions as i5 and i7 indexes, and Illumina adaptors were added to the region that contains the sgRNA. All samples were amplified with same forward (P5) primer mix, but with different reverse primers,

containing sample specific 8 bp indexes. The primer sequences were provided in **Table 5**.

The reactions were performed as follows: 10 ng template DNA were mixed with 0.5 µl Phusion High-Fidelity DNA Polymerase (NEB, M0530), 10 µl 5x GC Buffer, 1 µl dNTP mix (10 mM each) (Thermo, R0192), 2 µl Forward Stagger Mix (10 µM) and 2 µl Reverse Primer (10 µM) specific to each vector backbone and Nuclease-free water up to 50 µl. Thermal cycler conditions were as follows: denaturation for 30 seconds (s) at 98 C, followed by 10 s at 98 C, 15 s at 62 C, 20 s at 72 C for 16 cycles, final extension for 5 minutes at 72 C. PCR amplicons were gel extracted using NucleoSpin Gel and PCR clean-up (Macherey-Nagel, 740609) kit according to manufacturer's instructions and quantified using Nanodrop. Next generation sequencing was performed at Gen Era (Turkey) by using NovaSeq 6000 sequencing system (Illumina) with at least 10 million reads/plasmid library.

Table 5 – Primer sequences used in CRISPR screen gRNA NGS sequencing libraries

Name	Flowcell adapter		Illumina sequencing	Stagger region	Vector binding sequence
library_p5_0	AATGATACGGCGACCAACCGAGATCT		ACACTCTTTCCCTACACGACGCTCTTCCGATCT		TTGTGAAAGGACGAAACACCG
library_p5_1	AATGATACGGCGACCAACCGAGATCT		ACACTCTTTCCCTACACGACGCTCTTCCGATCT	C	TTGTGAAAGGACGAAACACCG
library_p5_2	AATGATACGGCGACCAACCGAGATCT		ACACTCTTTCCCTACACGACGCTCTTCCGATCT	GC	TTGTGAAAGGACGAAACACCG
library_p5_3	AATGATACGGCGACCAACCGAGATCT		ACACTCTTTCCCTACACGACGCTCTTCCGATCT	AGC	TTGTGAAAGGACGAAACACCG
library_p5_4	AATGATACGGCGACCAACCGAGATCT		ACACTCTTTCCCTACACGACGCTCTTCCGATCT	CAAC	TTGTGAAAGGACGAAACACCG
library_p5_5	AATGATACGGCGACCAACCGAGATCT		ACACTCTTTCCCTACACGACGCTCTTCCGATCT	TGCACC	TTGTGAAAGGACGAAACACCG
library_p5_6	AATGATACGGCGACCAACCGAGATCT		ACACTCTTTCCCTACACGACGCTCTTCCGATCT	ACGCAAC	TTGTGAAAGGACGAAACACCG
library_p5_7	AATGATACGGCGACCAACCGAGATCT		ACACTCTTTCCCTACACGACGCTCTTCCGATCT	GAAGACCC	TTGTGAAAGGACGAAACACCG
Reverse P7 primers					
Name	Flowcell adapter	Index	Illumina sequencing		Vector binding sequence
library_rev_1	CAAGCAGAAGACGGCATACGAGAT	CGGTTCAA	GTGACTGGAGTTTCAGACGTGTGCTCTTCCGATCT		TCTACTATTCTTCCCTGCACTGT
library_rev_2	CAAGCAGAAGACGGCATACGAGAT	GCTGGATT	GTGACTGGAGTTTCAGACGTGTGCTCTTCCGATCT		TCTACTATTCTTCCCTGCACTGT
library_rev_3	CAAGCAGAAGACGGCATACGAGAT	TAACTCGG	GTGACTGGAGTTTCAGACGTGTGCTCTTCCGATCT		TCTACTATTCTTCCCTGCACTGT
library_rev_4	CAAGCAGAAGACGGCATACGAGAT	TAACAGTT	GTGACTGGAGTTTCAGACGTGTGCTCTTCCGATCT		TCTACTATTCTTCCCTGCACTGT
library_rev_5	CAAGCAGAAGACGGCATACGAGAT	ATACTCAA	GTGACTGGAGTTTCAGACGTGTGCTCTTCCGATCT		TCTACTATTCTTCCCTGCACTGT
library_rev_6	CAAGCAGAAGACGGCATACGAGAT	GCTGAGAA	GTGACTGGAGTTTCAGACGTGTGCTCTTCCGATCT		TCTACTATTCTTCCCTGCACTGT
library_rev_7	CAAGCAGAAGACGGCATACGAGAT	ATTGGAGG	GTGACTGGAGTTTCAGACGTGTGCTCTTCCGATCT		TCTACTATTCTTCCCTGCACTGT
library_rev_8	CAAGCAGAAGACGGCATACGAGAT	TAGTCTAA	GTGACTGGAGTTTCAGACGTGTGCTCTTCCGATCT		TCTACTATTCTTCCCTGCACTGT
library_rev_9	CAAGCAGAAGACGGCATACGAGAT	CGGTGACC	GTGACTGGAGTTTCAGACGTGTGCTCTTCCGATCT		TCTACTATTCTTCCCTGCACTGT

Sequences are provided as 5' -> 3' direction.

2.6.6 CRISPR screen results analysis with MageCK

Sequencing results were downloaded as “.fastq.gz” files after demultiplexing, and analysis was run with MageCK algorithm in a Python (v. 3.7) environment (W. Li et al., 2014). Briefly, “-count” command was used for combining the gRNA specific counts, and “-test” command was used for running the pairwise comparisons and

generating the log Fold Change ($\log_2FC$) and p values. Total reads are provided in **Table 6**. Generated files were combined and visualized in R (v. 4.5.1).

Table 6 – Total reads of sequencing results

File	Label	Total Read Count
time0_biol1_r1.fastq.gz	time0_b1	23M
time0_biol2_r1.fastq.gz	time0_b2	10M
time0_biol3_r1.fastq.gz	time0_b3	27M
dox_biol1_r1.fastq.gz	dox_b1	12M
dox_biol2_r1.fastq.gz	dox_b2	17M
dox_biol3_r1.fastq.gz	dox_b3	11M
ortho_biol1_r1.fastq.gz	ortho_b1	17M
ortho_biol2_r1.fastq.gz	ortho_b2	8.9M
ortho_biol3_r1.fastq.gz	ortho_b3	4.5M
pyre_biol1_r1.fastq.gz	pyre_b1	10M
pyre_biol2_r1.fastq.gz	pyre_b2	7M
pyre_biol3_r1.fastq.gz	pyre_b3	3M

2.7. Flow cytometry experiments

For assessment of erythroid differentiation, 1×10^5 live cells were harvested, washed once with PBS, and resuspended in 100 μ L PBS. 1ul of each antibody was added, followed by incubation at room temperature for 30 minutes in the dark. After staining, cells were washed in PBS and resuspended in PBS containing 1% FBS for acquisition using a flow cytometer.

To determine reticulocyte levels within the erythroid cultures, 200 μ L of cell suspension was collected and stained with Hoechst 33342 (Sigma-Aldrich) at a final concentration of 5 μ g/mL to label nucleated cells. Samples were incubated at 37°C for 30 minutes in the dark and subsequently analyzed by flow cytometry. Staining reagents used in flow cytometry experiments are provided in **Table 7**.

Table 7 – Reagents used in flow cytometry experiments

Antibody and Chemical Name	Catalog number
CD71-FITC	Biolegend, 334104
CD235a-APC	Biolegend, 11076

Alpha 4 integrin (CD49d-APC)	Biolegend, 304308
CD233 (Band3) - FITC	American Research Products, NC1433740
Hoechst 33342	Invitrogen, 33342
DRAQ5	Biolegend, 424101
SYTO16	Thermo Fisher, S7578
IgG Control	Biolegend, 410711
Anti-Mouse-APC	Invitrogen, A-11005
Anti-Rb-APC	Invitrogen, A-11012

2.8. May-Grunwald Giemsa Staining

Enucleation was assessed by Cytospin preparation followed by MGG staining. Slides were fixed in methanol for 10 minutes, stained with May-Grunwald solution (Sigma, MG1L) for 10 minutes, and subsequently stained with Giemsa solution (Sigma, GS500) for 5 minutes. Stained slides were rinsed, air-dried, and examined under a light microscope.

2.9. RNA sequencing (RNAseq)

For each biological replicate, 5×10^6 cells were seeded in expansion medium. After 2 days, the culture was split into two conditions: two-thirds of the cells were transferred into erythroid differentiation medium, while one-third remained in expansion medium. Following 6 days of differentiation and expansion, cells were harvested for RNA isolation. Total RNA was extracted using a commercial kit (Macherey-Nagel, 740955), following the manufacturer's instructions. Cells were lysed in the provided lysis buffer, homogenized, and loaded onto a silica-based RNA purification columns. An on-column DNase I digestion step was included to remove genomic DNA contamination. Purified RNA was washed thoroughly and eluted in RNase-free water.

2.10. Single Cell RNA sequencing (sc-RNAseq)

Single-cell RNA sequencing was performed using the 10x Genomics Chromium Single Cell 3' protocol, following the manufacturer's instructions. CD34⁺ hematopoietic stem

cells were collected at differentiation days 0, 5, 8, 18, and 24. Single cell libraries were prepared according to manufacturer's instructions and sequenced with NovaSeq 6000 (Illumina). Reads were aligned to human genome (Hg38) with CellRanger software. Output files were analyzed with Seurat (v. 5.3.1). Normalization was made with "NormalizeData" function in Seurat, which use logarithmic normalization and scaling (scale factor = 10000). Data integration was performed with Harmony in Seurat. Pseudobulk expression analysis was performed with DElegate (v. 1.2.1). All analyses were conducted in R (v. 4.5.1).

2.11. Sanger sequencing of CRISPR edits in HSCs

For CRISPR/Cas9 knockout validation, genomic DNA was isolated from KO and control cells using a standard genomic Macherey-Nagel DNA extraction kit according to the manufacturer's instructions. A genomic region of approximately 300 bp surrounding the CRISPR target site was amplified by PCR using locus-specific primers (**Table 8**). PCR reactions were performed under optimized conditions to obtain a specific product, which was purified by agarose gel electrophoresis. Indel regions were quantified using the TIDE analysis.

Table 8 – Primer sequences for Sanger sequencing

SMARCB1-sgRNA-F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACCTTGATGTGCTGCATCCACT
SMARCB1-sgRNA-R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATCGTTGCCATCCAGAATCTCTT
SMARCB1-sgRNA-F	TTTGAACCCGCTGACGTTTG
SMARCB1-sgRNA-R	AAAGACGGGCTTCTGCTGAG
EP300-sgRNA-F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACATTTTCTTTCTCTATGTGTTCA
EP300-sgRNA-R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGATTGTTGCAGATTCATACATG
EP300-sgRNA-F	CAGACACTGTGTTGCTACGG
EP300-sgRNA-R	GTATGTGTGCTGGCCCATCC
EP300-sgRNA-F	CTCCCTGGGTCCTATGCCAA
EP300-sgRNA-R	GGTGGGAAAGGTTGAGAAGGAA

2.12. Bulk Docking & Deamination – Sequencing (D&D-Seq) analysis

This protocol is adopted from Chi et al. (Chi et al., 2025). D&D protocol is composed of 4 main steps: cell fixation and permeabilization, antibody binding, D&D binding and activation, bulk ATAC-Seq (**Figure 5**). Cells were processed using a standardized centrifugation scheme, in which all spins prior to fixation were carried out at $300 \times g$ for 5 min and all spins after fixation at $600 \times g$ for 5 min using a swing-bucket rotor. We collected between 2×10^5 for each condition. To fix the cells, we added methanol-free 16% formaldehyde (Thermo Fisher Scientific, PI28906) to a final concentration of 0.1% and incubated the suspension for 5 min at room temperature. Fixation was quenched by adding glycine to a final concentration of 125 mM, followed by one wash in 1 ml PBS.

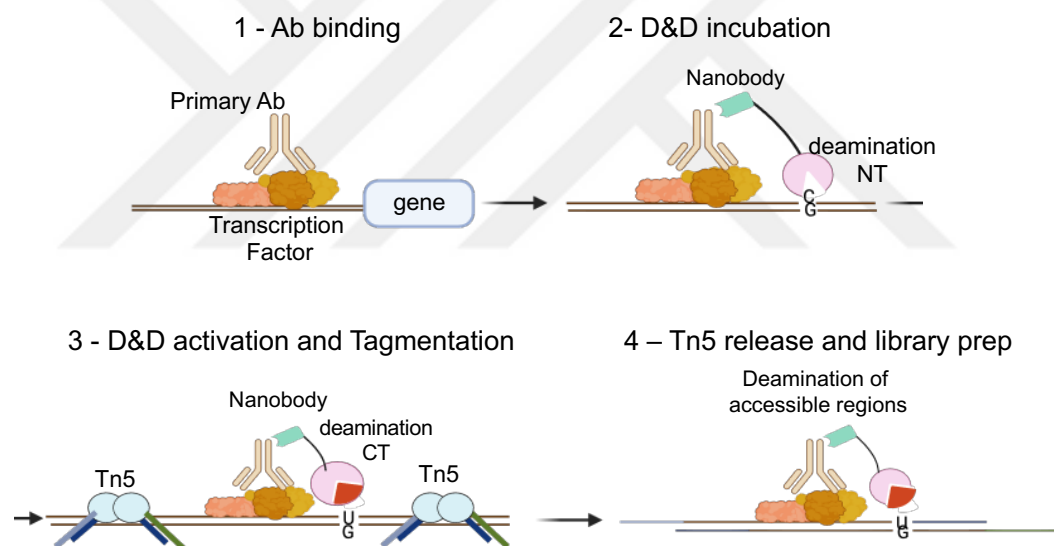


Figure 5 – D&D-Seq workflow
Created in BioRender. ID: WU2961DSTB

For permeabilization, fixed cells were incubated on ice in 200 μ l permeabilization buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 3 mM MgCl₂, 0.1% NP-40, 0.1% Tween-20, 1% BSA, and protease inhibitors). Permeabilization was performed for 7 min for cell lines and 5 min for primary cells. Cells were then washed in 200 μ l ice-cold wash buffer (20 mM HEPES pH 7.6, 150 mM NaCl, 0.5 mM spermidine, 1% BSA, and protease inhibitors).

Following cell counting, we transferred 2×10^5 cells into PCR tubes and resuspended them in 100 μ l antibody buffer (20 mM HEPES pH 7.6, 150 mM NaCl, 2 mM EDTA, 0.5 mM spermidine, 1% BSA, and protease inhibitors). Primary antibodies were added at a 1:100 dilution (1 μ l per reaction). Cells were incubated overnight at 4°C with gentle rotation to allow efficient antibody binding.

After overnight incubation, D&D binding and activation was followed. Cells were first washed with 200 μ l wash buffer and then with an equal volume of D&D binding buffer. The pellet was resuspended in 50 μ l D&D activation buffer containing 100 μ M DddA_NT, and samples were rotated at room temperature for 1 h to allow binding. After this step, cells were washed again in 200 μ l D&D binding buffer, resuspended in 50 μ l fresh D&D activation buffer, and incubated for an additional hour at 37 °C to induce deamination.

After deamination reaction is completed Tn5 adaptors were obtained from IDT and prepared according to standard procedures (**Table 9**). In brief, 100 μ M adaptors were annealed in TE buffer to generate double-stranded mosaic-end (MEDS) oligos by heating to 95 °C for 5 min followed by controlled cooling (0.2 °C per second) to 12 °C. Equal volumes of MEDS-A and MEDS-B were combined, and 6.25 μ l of annealed ME-A and ME-B was incubated with 10 μ l of in-house Tn5 enzyme at room temperature for 30mins to assemble Tn5. Once D&D activation was completed, cells were washed once with 200 μ l 1 $\times$ Tris-TD buffer and resuspended in 50 μ l of the same buffer containing 2 μ l loaded Tn5 enzyme. The optimal amount of Tn5 was found by titration test and titration test was repeated in every new batch of Tn5. Then, reactions were incubated at 37 °C for 1 h for tagmentation. To lyse the cells, we added 1 μ l 10% SDS, 2 μ l proteinase K (NEB, P8107S), and 3 μ l 0.5 M EDTA, followed by incubation at 55 °C for 1 h. DNA was then purified using a column-based protocol (Zymo, D5205) according to the manufacturer's instructions. For library amplification, we used an uracil-tolerant DNA polymerase to be able to amplify deamination footprint (NEBNext Q5U Master Mix, NEB M0597S) together with Nextera-compatible barcoding primers. To improve the initial gap-filling step, 1 μ l of Bst 3.0 polymerase (NEB, M0374) was added to the reaction. PCR conditions were: 72 °C for 5 min; 98 °C for 30 s; 12 cycles

of 98 °C for 10 s, 55 °C for 30 s, and 72 °C for 30 s; and a final extension at 72 °C for 5 min.

Table 9 – Tn5 tagmentation oligonucleotide sequences

Mosaic end Adapter A	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
Mosaic end Adapter B	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG
Mosaic end reverse	PO4-CTGTCTCTTATACACATCT

2.13. Single cell Combined D&D-Seq and Multiome analysis

Following D&D activation, cells were transferred into 1× nuclei buffer and quantified using trypan blue exclusion on a Countess II FL Automated Cell Counter. All subsequent steps were performed according to the single-cell Multiome protocol provided by 10x Genomics (CG000209), with the modification described below. During the GEM generation and barcoding reaction (step 2.1), 2 ul of uracil-tolerant DNA polymerase (NEBNext Q5U Master Mix, NEB M0597S) was added to the barcoding reaction mixture to facilitate the first few rounds of PCR.

The sequencing libraries were run on a NovaSeq6000 using dual-indexed, paired-end 150-bp configurations. Index and read lengths were as follows: i5, 8 bp (16 bp for single-cell sequencing); i7, 8 bp; read1 150 bp; and read2 150 bp.

2.14. Computational analysis of D&D-Seq

To identify and quantify the D&D-induced genome edits, we analyzed scMultiome-seq and bulk ATAC-seq datasets using an in-house pipeline. First, scMultiome-seq BAM files were generated by CellRanger, and separated by cell type using sinto, while bulk ATAC-seq reads aligned with BWA-MEM2 were processed using the same pipeline. In each BAM, PCR duplicates were removed with picard MarkDuplicates function, and only high-quality reads were retained.

Variant calling was performed using bcftools mpileup, and only variants supported by multiple reads were considered as candidate edits. Germline SNPs were removed using

genomeAD annotations and variant allele frequency thresholds, after which we isolated C→T and G→A transitions as the characteristic signatures of D&D activity.

ATAC peaks were then analyzed with MACS2, filtered against ENCODE blacklist regions, and annotated for transcription factor motifs using MEME-SEA with the HOCOMOCO motif library. Peaks containing the relevant motif (or overlapping ChIP-seq binding tracks) were defined as target peaks, whereas motif-lacking peaks were treated as background. Both peak groups were resized around the motif or peak summit and intersected with called edits to measure on- versus off-target mutation signals.

The final output of these analyses contained the VCF files that lists D&D edits, BED files containing motif-centered peak coordinates, and the BAM files composed only of edited reads. These edited-read BAMs were converted to bigWig format and visualized in IGV to evaluate binding sides and distribution of D&D edits at regulatory sites.

Analyses were conducted with Seurat (v. 5.3.1), Signac (v. 1.15.0), and Monocle3 (v. 1.4.26) in R (v. 4.5.1). Single cell D&D edit counts were analyzed with an in-house pipeline. Python and R scripts for D&D-seq data analysis are available on GitHub <https://github.com/sangho1130/DnD/>.

For differential GATA1-binding-associated D&D edit analysis, D&D edits located within ± 3000 bp of annotated transcription start sites (TSS) were assigned to genes using R packages GenomicRanges (v. 1.61.5), AnnotationDbi (v. 1.71.1) and org.Hs.eg.db (v. 3.21.0). Edit counts mapped to the same gene were summed, and differential analysis was performed using a linear modeling framework implemented in limma (v. 3.65.4).

2.14. Statistical Analyses

Statistical analyses were performed using GraphPad Prism (version 10), Microsoft Excel (2024), and R (v. 4.5.1). For comparisons between two groups, two-tailed Student's t-tests were used where appropriate. For experiments involving multiple

groups or variables, one-way or two-way analysis of variance (ANOVA) was applied, followed by appropriate post hoc tests as indicated in the figure legends.

Differential GATA1 binding analysis for D&D-Seq experiments was conducted using the limma framework. Bulk RNA-sequencing differential expression analyses were performed using DESeq2. For single-cell RNA-sequencing data, cells were aggregated into pseudobulk profiles by cell sources, and differential expression analysis was performed using DESeq2. Unless otherwise stated, p values < 0.05 were considered statistically significant.



Chapter 3

RESULTS**3.1. Validation of erythroid differentiation conditions required for CRISPR screening**

To evaluate the epigenetic dependencies of erythroid differentiation, we first established and validated a suitable cell model for a CRISPR-Cas9 screening platform. We utilized the human erythroid progenitor BEL-A cell line, engineering it to stably express both the Cas9 nuclease and a nuclear-localized GFP (H2B-GFP). The H2B-GFP signal was intended to serve as a marker to identify and isolate nucleated cells during flow cytometry, thereby enabling the efficient enrichment of differentiated erythroid cells.

To optimize the experimental model system, BEL-A cells were first induced to differentiate, and cell morphologies were evaluated using live-cell microscopy and Giemsa staining. After 4 and 7 days of differentiation, the populations predominantly consisted of basophilic and polychromatic erythroblasts, whereas orthochromatic erythroblasts appeared after day 10. By day 12, the cultures contained mainly reticulocytes and pyrenocytes, indicating completion of the terminal erythropoiesis program (**Figure 6-A**). To separate the orthochromatic erythroblasts and enucleated pyrenocytes in mixed differentiated cell populations, cells were sorted by size using flow cytometry (orthochromatic erythroblasts: DAPI⁺/large; pyrenocytes: DAPI⁺/small). Subsequent staining of the sorted fractions confirmed high purity and efficient separation (**Figure 6-B**). In addition, GPA staining of the sorted populations revealed a reduction in GPA levels in both erythrocytes and pyrenocytes as expected (**Figure 6-C**).

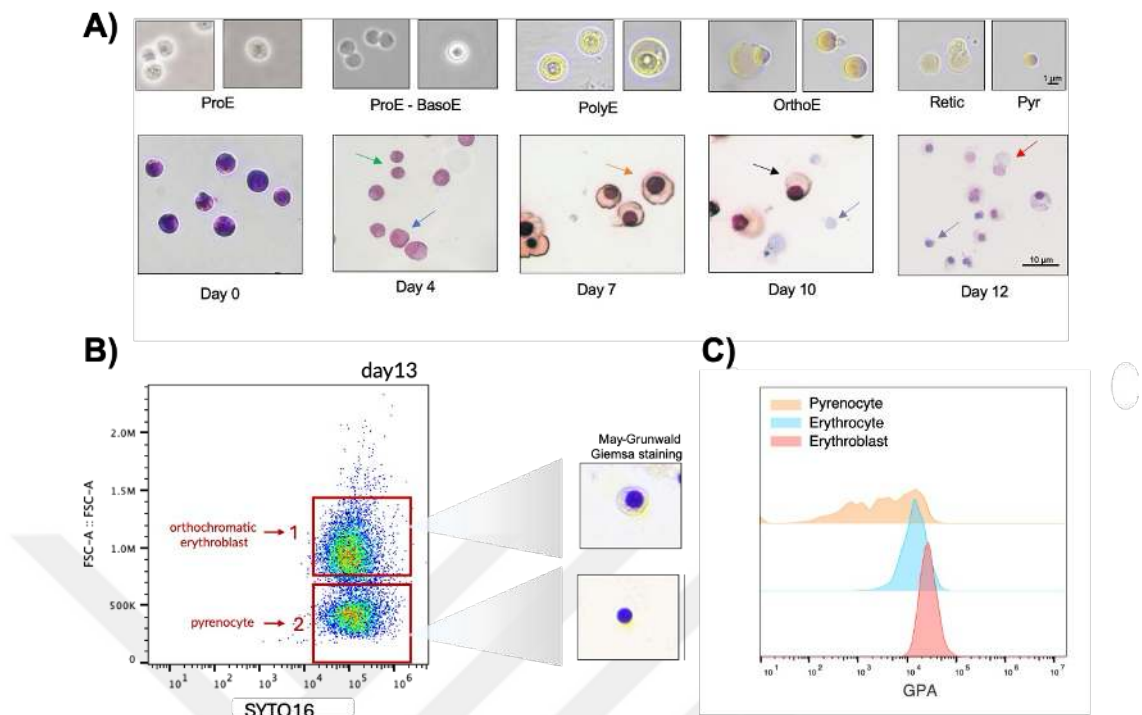


Figure 6 – Characterization of different differentiation stages of BEL-A cells

A) Morphologies of different cell populations in differentiating BEL-A cells. B) Separation of pyrenocytes and orthochromatic erythroblasts after differentiation. C) Glycophorin A staining intensity of different cell populations.

Next, stable Cas9-expressing BEL-A cells were generated for use in CRISPR screen experiments. Cas9 expression was confirmed by Western blotting (**Figure 7-A**). To ensure that Cas9 expression did not adversely impact erythropoiesis, we evaluated differentiation outcomes and enucleation efficiency and observed no significant differences in cell-type distributions following differentiation (**Figure 7-B**).

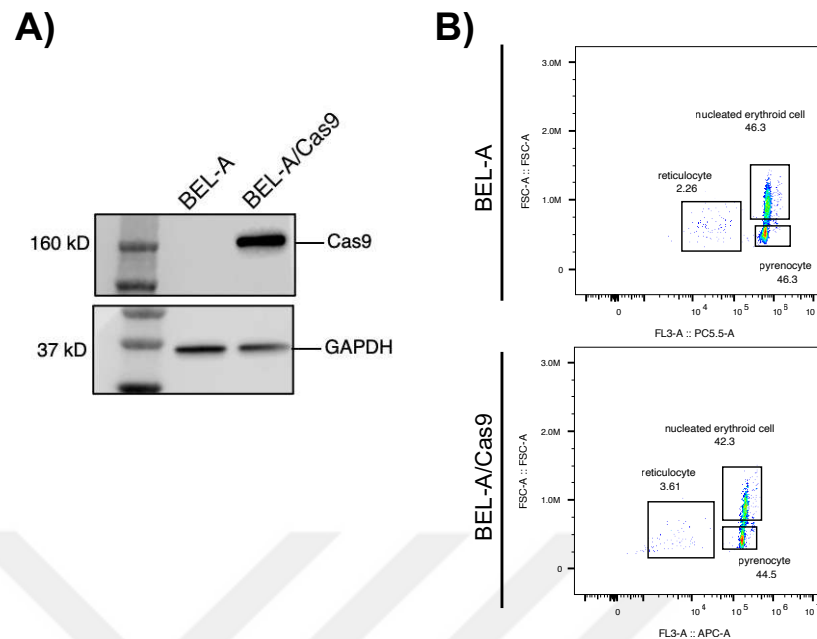


Figure 7 – Cas9 expression does not affect differentiation and enucleation of BEL-A cells

A) Cas9 expression in BEL-A cells. B) Cas9 expression does not alter differentiation and enucleation efficiency of BEL-A cells.

Next, Cas9 activity in the engineered BEL-A-Cas9 cells was evaluated. This was achieved by transducing the cells with a GFP expression cassette alongside two guide RNA sequences designed to target the GFP sequence. Analysis by flow cytometry demonstrated a significant reduction in GFP-positive cells, confirming highly efficient Cas9-mediated editing (**Figure 8**). Specifically, we observed a total decrease in GFP fluorescence after 13 days, compared to control cells transduced with a non-targeting gRNA, thereby validating the functionality of the Cas9-expression system.

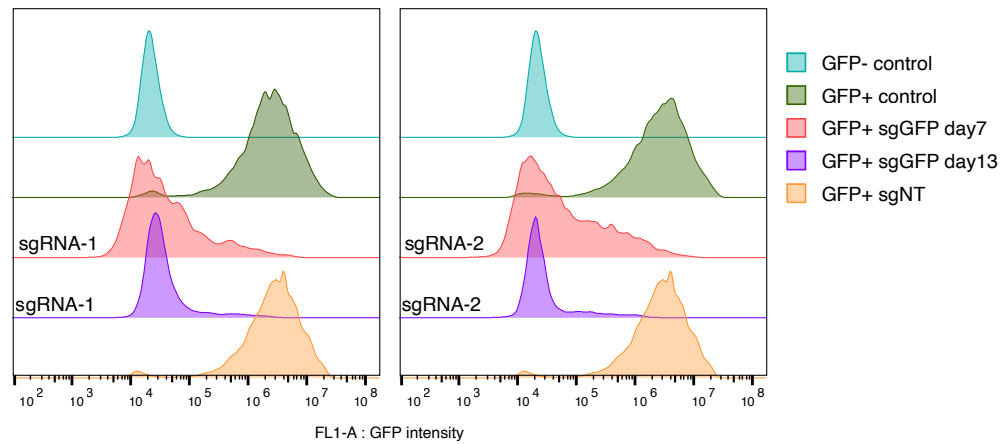


Figure 8 – Cas9 activity assay reveals the effectiveness of Cas9 in BEL-A cells

While SYTO16 staining confirmed successful differentiation and enucleation, its application in a large-scale CRISPR screen was deemed impractical due to the prohibitively long durations required for cell sorting. To overcome this limitation, we employed the H2B-GFP marker as an alternative. To confirm that H2B-GFP expression did not alter the intrinsic differentiation capacity of the cells, we performed comparative differentiation assays. These experiments confirmed that H2B-GFP-expressing BEL-A cells underwent terminal erythroid differentiation, including enucleation, with kinetics and efficiency comparable to the parental BEL-A cell line (**Figure 9**).

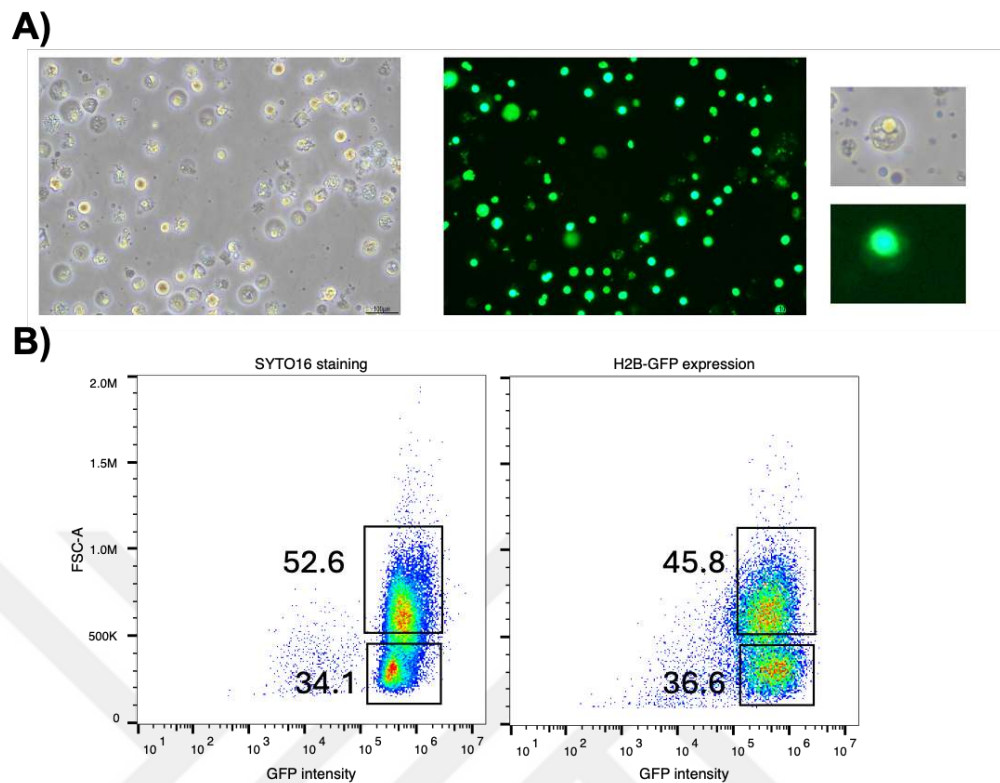


Figure 9 – H2B-GFP expression is maintained in pyrenocytes

A) Live cell imaging of H2B-GFP expression in BEL-A cells. B) H2B-GFP expression allows separation of different cell populations with flow cytometry and does not alter enucleation efficiency.

3.2. CRISPR screen identifies the epigenetic dependencies of erythroid differentiation

The epigenome targeted CRISPR screen in BEL-A cells was done by separating and sorting final populations by FACS. Briefly, BEL-A-Cas9 H2B-GFP cells were transduced with EPIKOL library, selected with puromycin, and expanded for 3 days in cell growth media. Then, cells were separated into two conditions; one condition was maintained in growth media, while differentiation was induced in the second condition. After 13 days of differentiation, nucleated differentiated cells (orthochromatic erythroblasts) and enucleated pyrenocytes were sorted into separate tubes. Time-matched undifferentiated population was also collected to evaluate the growth associated epigenetic dependencies (**Figure 10**).

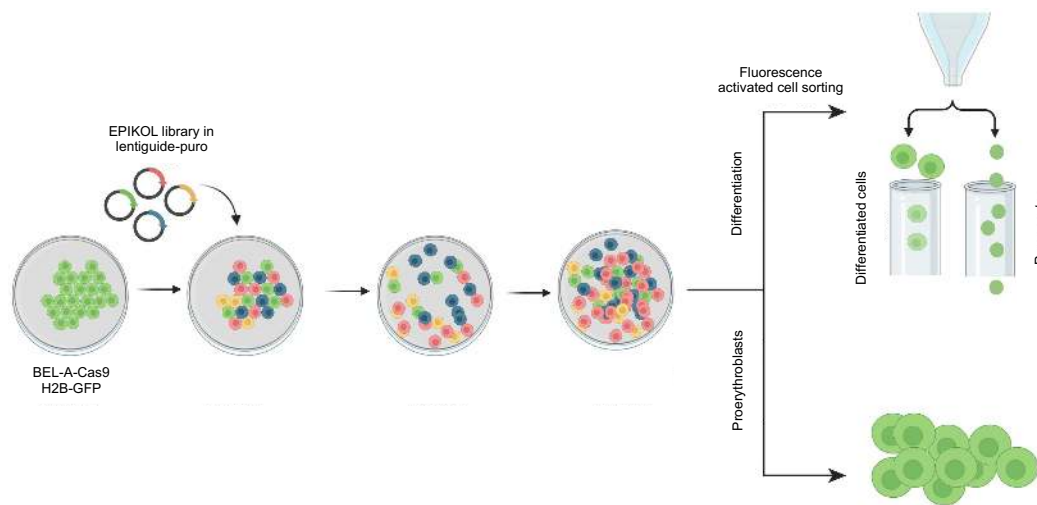


Figure 10 – Experimental design of the epigenome targeted CRISPR screen
Created in BioRender. ID: ZQ2961KH3M

CRISPR screen experiments were conducted in three different biological replicates, and differentiated cells were separated by FACS (**Figure 11-A**). Then, libraries were PCR amplified (**Figure 11-B**) and sequenced. Sequences were evaluated with pairwise comparisons by MageCK algorithm.

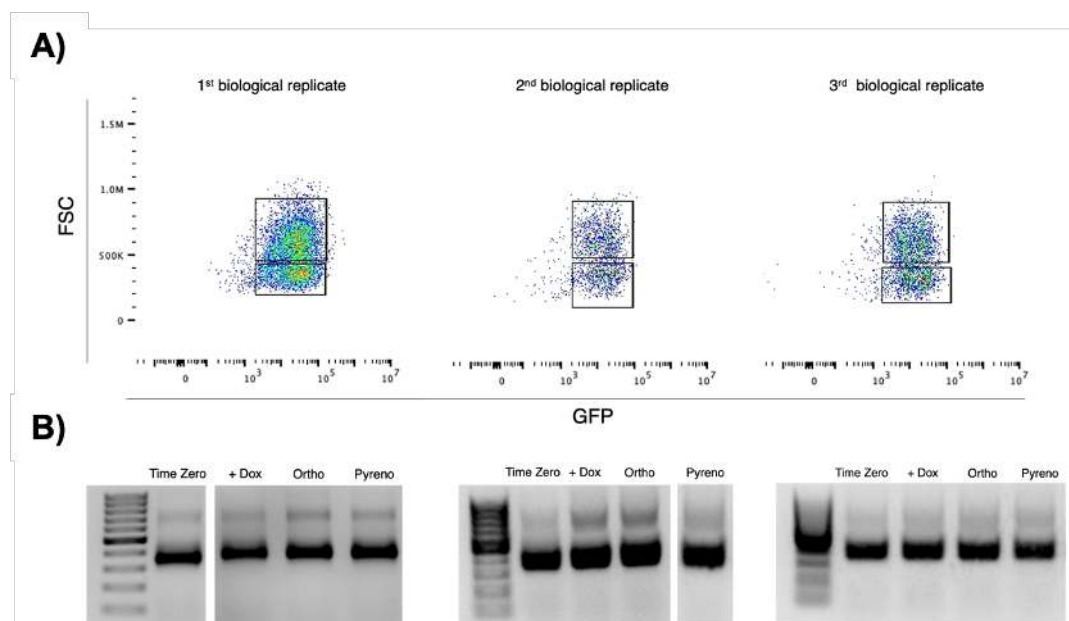


Figure 11 – CRISPR screen experiments in three biological repeats

A) Sorted populations in CRISPR screen experiments after 13 days of differentiation. B) PCR amplification results of NGS libraries in 3 biological repeats.

Initially, we compared the dependencies of cell growth in SFEM media, and differentiated cells, by comparing both populations to Initial library transduced cells. Comparative plots revealed 3 different groups of gene depletions; (i) Genes only depleted in growth comparison, showing that SFEM-growth dependencies, (ii) Genes depleted in both comparisons, showing that genes required for survival in both conditions, and (iii) Genes only depleted in differentiation comparison, showing the differentiation associated dependencies (**Figure 12**). PRPF31, TFPT, EP400, DPF2 and ARID2 were among the top growth depleted genes. SF3B1, RUVBL1, and SS18L2 were depleted both in growth and differentiation conditions. PPP4R2, SFPQ, KDM3B, and EP300 were only depleted in differentiation conditions, with no depletion effect in growth comparison (**Figure 12-A**). Ranking the genes with differential LFC values revealed that WAC and EP300 were top differentiation associated hits (**Figure 12-B**).

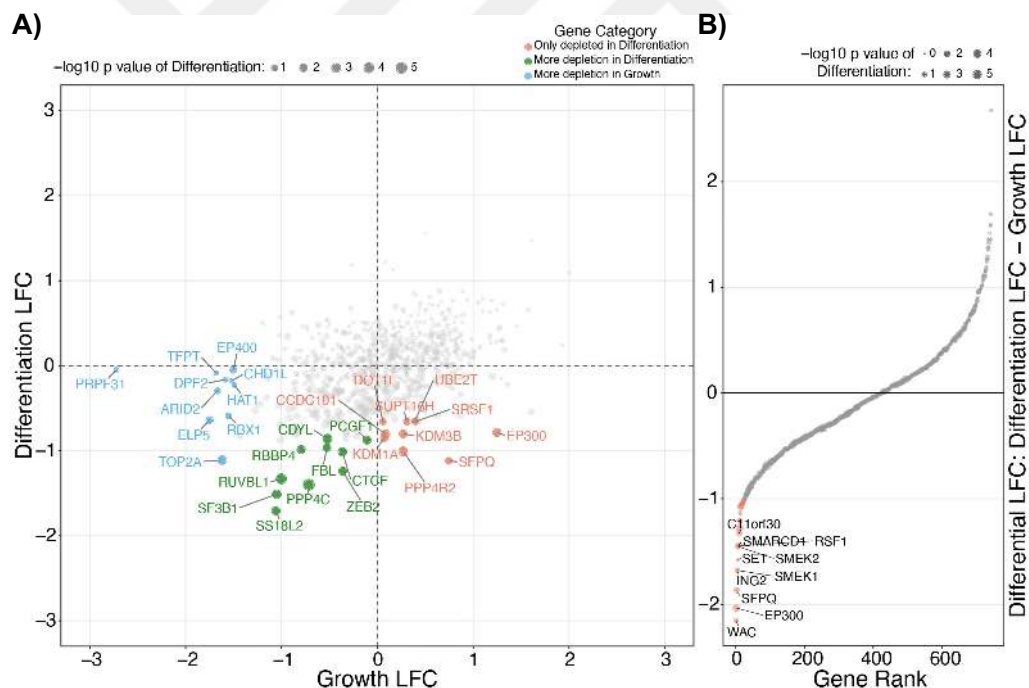


Figure 12 – Growth and differentiation associated comparisons

A) Comparison of growth and differentiation associated hits. B) Ranked plot of top differentiation associated hits.

3.3. CRISPR screen identifies stage specific epigenetic dependencies of erythropoiesis

After determining the general differentiation and growth effects, we focused on individual populations and pairwise comparisons. 4 different comparisons showed different requirements of distinct stages of erythroid differentiation (**Figure 13**):

- 1) **Ortho vs Time 0**: This comparison revealed gene dependencies of differentiation into Orthochromatic erythroblasts.
- 2) **Dox vs Time 0**: This comparison showed the required genes for BEL-A cell survival in an undifferentiated state.
- 3) **Pyre vs Time 0**: This comparison showed genes depleted in pyrenocyte population compared to initial time point.
- 4) **Pyre vs Ortho**: This comparison showed genes depleted in pyrenocytes compared to orthochromatic erythroblasts; revealing the late stage (enucleation) dependencies of erythroid differentiation.

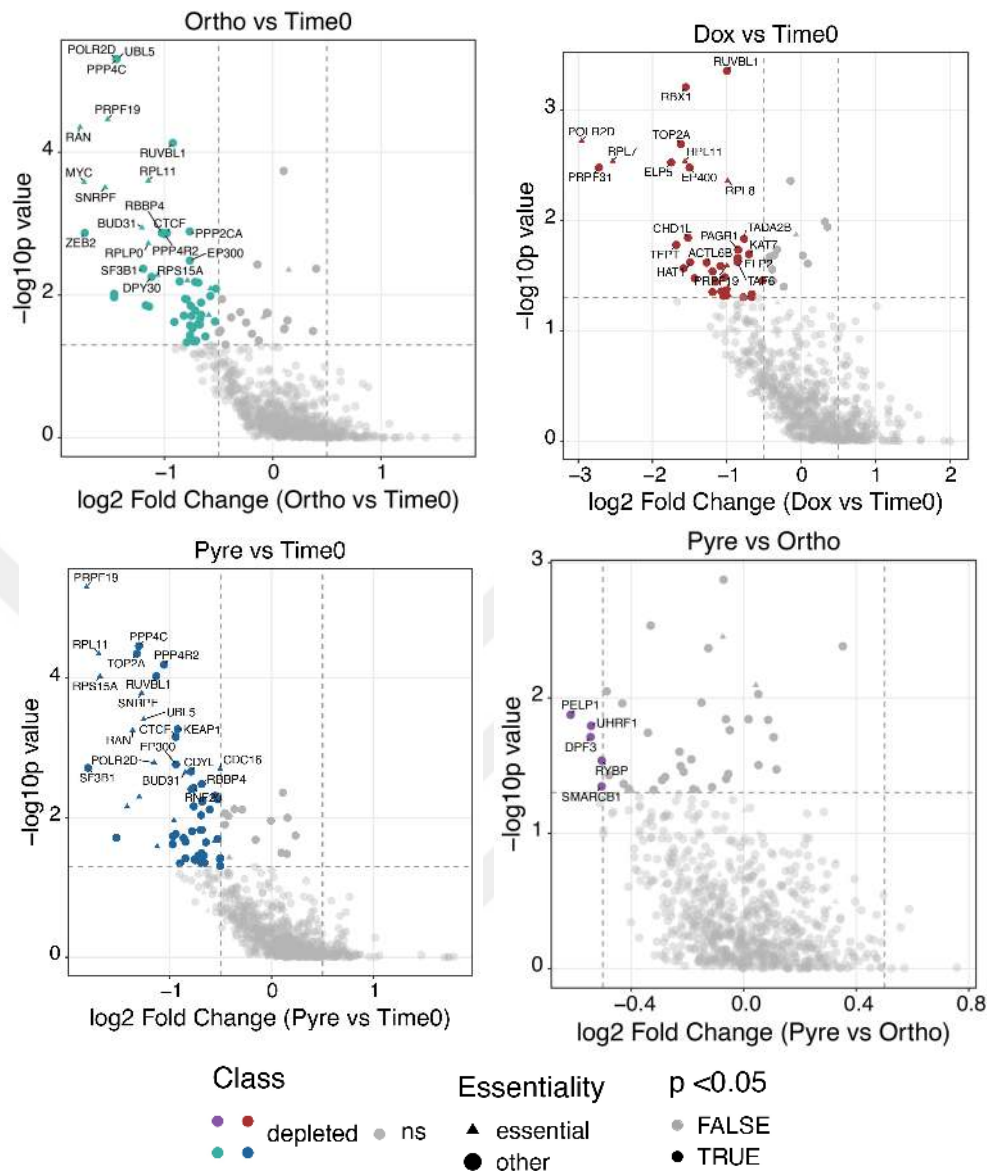


Figure 13 – Volcano plots of stage specific comparisons

Excluding the essential genes in the CRISPR library, the top significantly depleted hits in orthochromatic erythroblasts and pyrenocytes were PPP4C, PPP4R2, KEAP1 and EP300 (**Figure 13**). The most depleted genes in undifferentiated BEL-A cells involved EP400, RUVBL1, PRPF31, and CHD1L, marking the survival associated dependencies (**Figure 13**). Intriguingly, when depleted genes in pyrenocytes were compared with depletions in orthochromatic erythroblasts, PELP1, UHRF1 and SMARCB1 were top genes (**Figure 13**). Pairwise comparisons also marked EP300 as a unique differentiation dependency (**Figure 14-A** and **Figure 14-B**), while

SMARCB1, PELP1, UHRF1 and ZMYM3 were top hits in pyrenocytes, compared to orthochromatic erythroblasts (**Figure 14-C**).

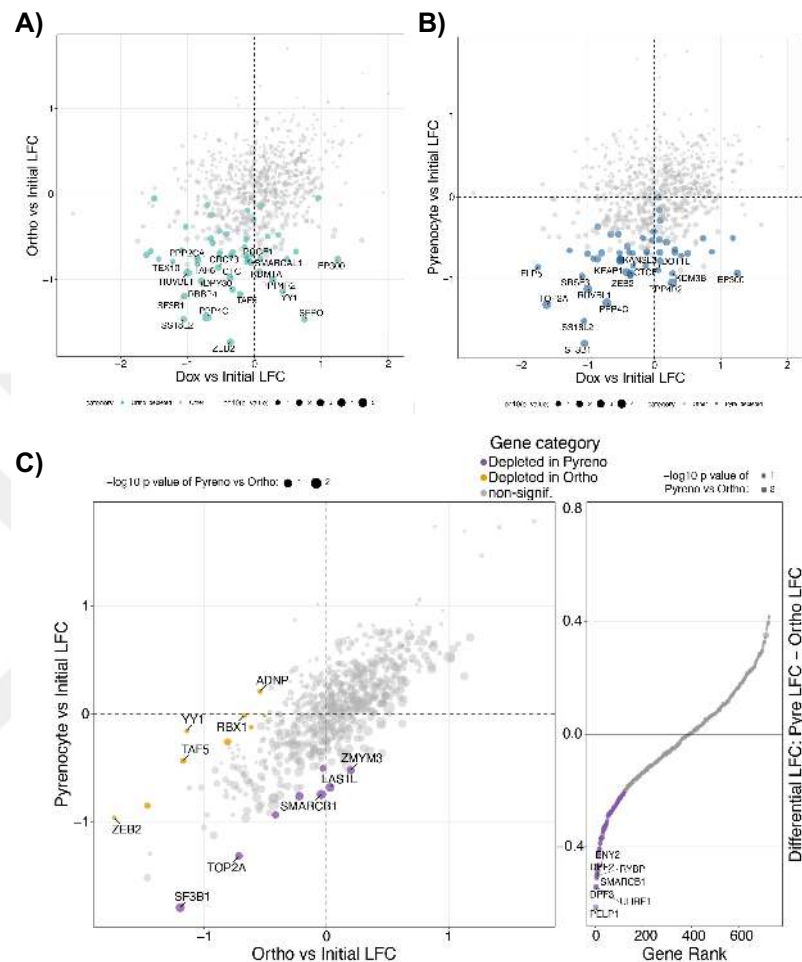


Figure 14 – Pairwise comparisons of depletions in different cell types

A) Comparison of CRISPR screen results in dox and ortho comparisons. B) Comparison of CRISPR screen results in dox and pyrenocyte comparisons. C) Left: Comparison of ortho and pyrenocyte associated hits. Right: Ranked plot of top pyrenocytes associated hits.

Additionally, combined results of the top 20 hits of all comparisons revealed different depletion patterns of several genes. As examples, even though UHRF1, SETD8 and RCOR1 were among the top 20 hits in pyrenocyte vs orthochromatic erythroblasts comparison, they were also depleted in undifferentiated cells vs initial time point comparison (**Figure 15**). This approach revealed that EP300, CTCF, and PPP4C as

true differentiation dependencies, and SMARCB1 and L3MBTL4 as true enucleation (late stage) dependencies (**Figure 15**).



Figure 15 – Combined results of top20 hits of all comparisons

Protein-protein interaction (PPI) analyses of all significant CRISPR screen hits revealed several pathway level enrichments. Subclusters of MCL clustering enrichment revealed pathways as SWI/SNF complex, chromatin modifying enzymes, protein phosphatases, and transcriptional repressors (**Figure 16**). GO

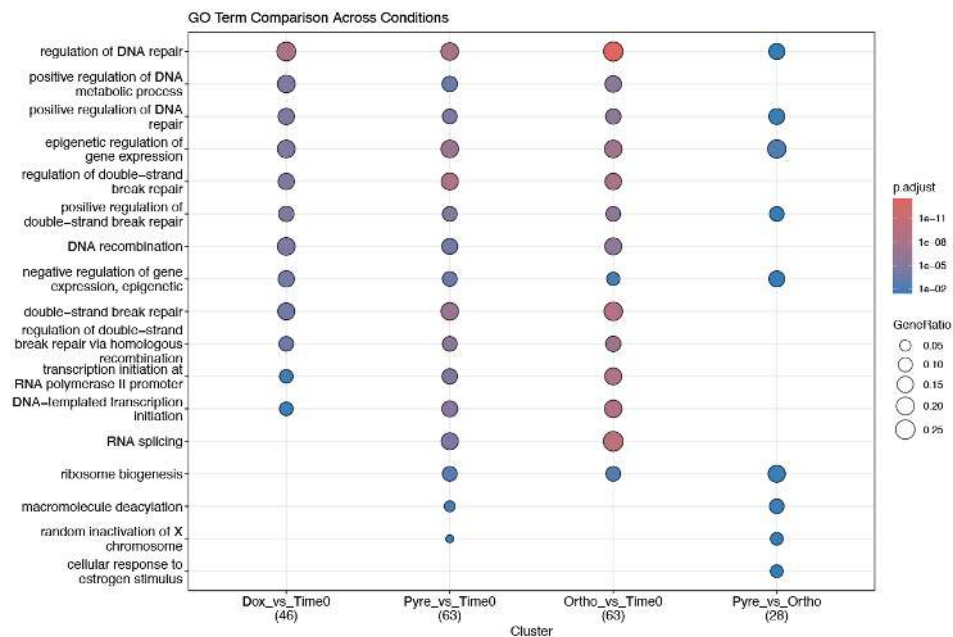


Figure 17 – GO enrichment analysis of CRISPR screen hits

3.4. Dual-color competition experiments validate CRISPR screen results

To dissect the temporal requirements of EP300 and SMARCB1 during erythroid maturation, we first performed time-resolved dual-color competition assays (**Figure 18**). In mixed cultures, sgSMARCB1-GFP cells showed a significant reduction at late differentiation time points (day 10), whereas sgEP300-GFP cells were depleted much earlier (day 4) (**Figure 19**). These observations indicate that EP300 is essential for the survival of early differentiating BEL-A cells, while SMARCB1 function is required predominantly during late maturation.

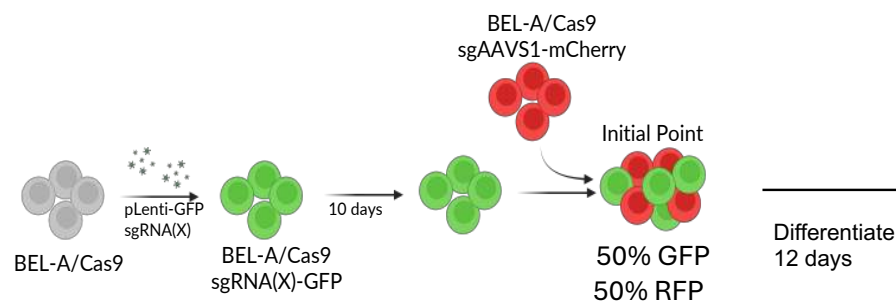


Figure 18 – Overview of competition experiments.

Created in BioRender. ID: AS2961I7Z0

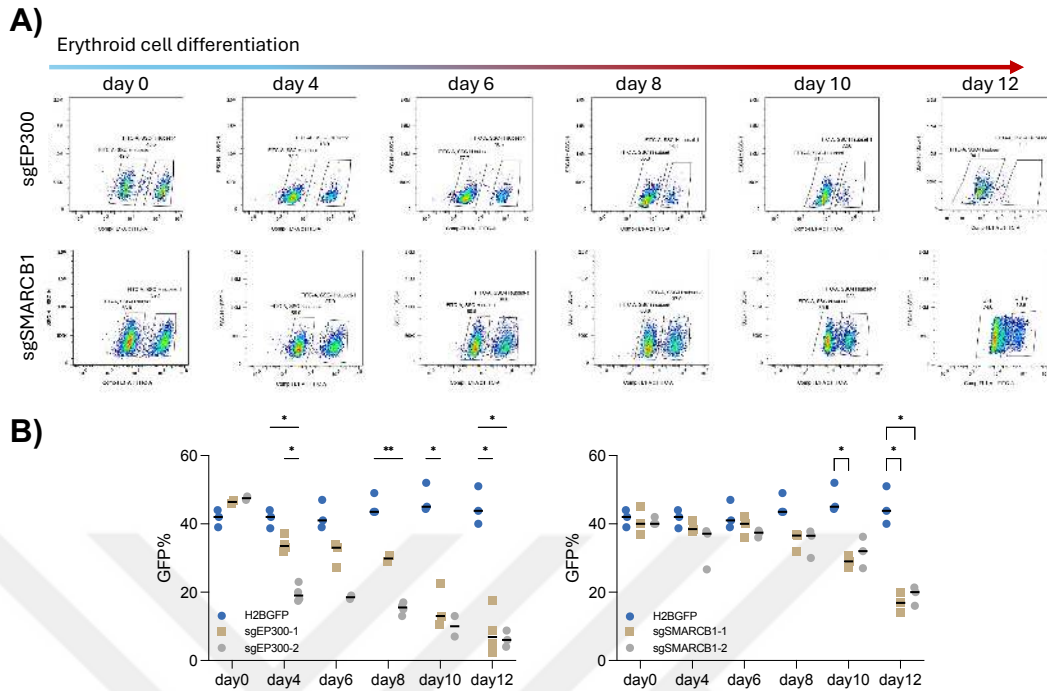


Figure 19 – EP300 is required for survival in early differentiation stages

A) Flow cytometry plots showing erythroid differentiation over time (Day 0 to Day 12, left to right, arrow indicated). Analysis based on FITC-A and SSC-A parameters to assess changes in sgRNA-GFP⁺ (%) cells during differentiation (n=3). B) Quantification of GFP⁺ (%) cells assessed by flow cytometry in EP300-KO and SMARCB1-KO erythroid cells.

Following these temporal analyses, we quantified enucleation phenotypes in mixed mCherry-sgAAVS1 and sgRNA-eGFP populations after 12 days of differentiation. Four cell types as nucleated GFP⁻, nucleated GFP⁺, GFP⁻ pyrenocytes, and GFP⁺ pyrenocytes were compared, and the yielding ratios were determined as follows: 30% / 35% / 20.5% / 18.4% (**Figure 20-A**).

Using this baseline, we then evaluated percentage of cell types of targeted sgRNAs (EP300, SMARCB1, DPF2, DPF3) against mCherry-sgAAVS1 controls. The results showed: (i) sgEP300 cells displayed reduced proportions of both nucleated cells and pyrenocytes (47.5% / 6.95% / 43.9% / 1.63%), (ii) sgSMARCB1 did not alter nucleated cell frequency but markedly decreased pyrenocyte production (43.0% / 29.7% / 25.1% / 2.24%), (iii) sgDPF2, similar to sgEP300, reduced both nucleated and enucleated populations (59.4% / 12.9% / 22.9% / 4.85%), and (iv) sgDPF3

selectively reduced pyrenocytes (48.0% / 25.0% / 27.9% / 1.15%) (Figure 20-B & 20-C).

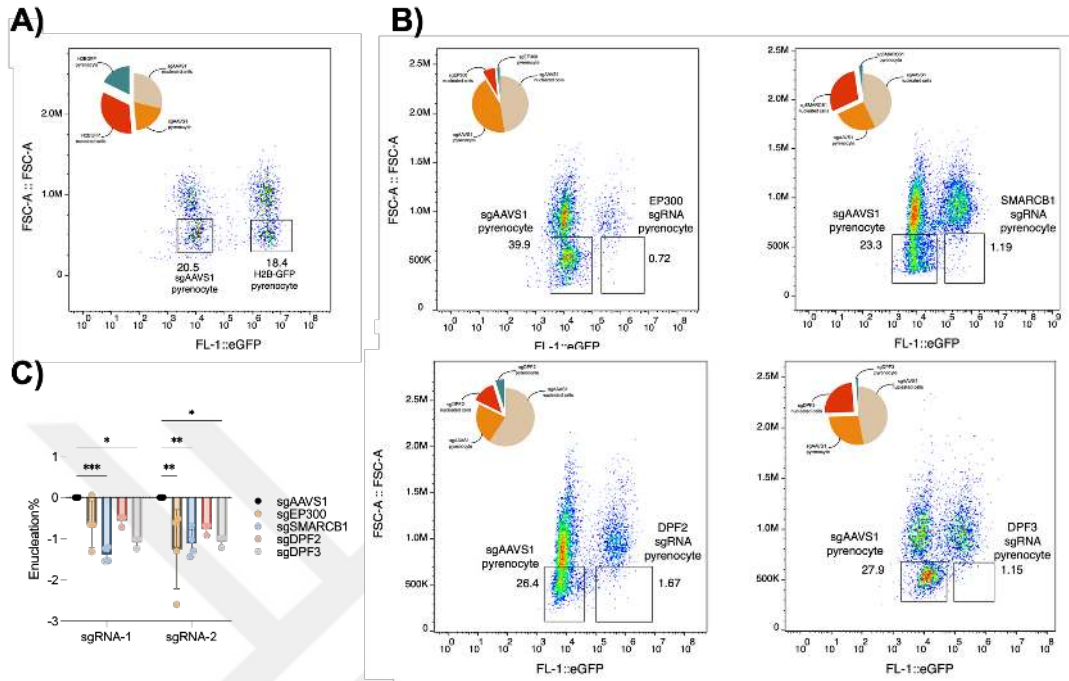


Figure 20 – Dual color competition experiments for enucleation

A) Competition experiment with sgAAVS1-mCherry- and H2B-GFP-expressing BEL-A cells. B) The effects on pyrenocyte formation across conditions by flow cytometry. Robust pyrenocyte frequencies in control sgAAVS1 cells (23.3–39.9%); pronounced reduction upon EP300 or SMARCB1 targeting (0.72% and 1.19%). Moderate effects upon DPF2 and DPF3 perturbation (1.67% and 1.15%). C) Quantification of pyrenocyte frequencies in competition experiments across indicated sgRNA conditions. Mean $\pm$ SEM from six measurements derived from three independent biological replicates. Two-way ANOVA with Šídák multiple comparisons test; multiplicity-adjusted P values relative to sgAAVS1 control $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***), $P < 0.0001$ (****).

3.5. EP300 is required for differentiation of BEL-A cells

Since the CRISPR screen and dual-color competition assays suggested distinct, time-dependent dependencies for EP300 and SMARCB1, we next examined how targeting these genes affects the dynamics of erythroid differentiation. We monitored cells at day 2, day 5, and day 7 after induction of differentiation and assessed progression using the surface markers CD49d (downregulated during maturation) and CD233

(upregulated during maturation). As expected, sgNT control cells showed a steady decrease in CD49d and increase in CD233, reflecting normal erythroid differentiation (**Figure 21**).

In sgSMARCB1 cells, the CD49d/CD233 trajectory closely resembled the control, indicating that SMARCB1 loss does not disrupt the early stages of differentiation. By contrast, sgEP300 cells displayed a clear stall in differentiation at day 5, failing to downregulate CD49d or upregulate CD233 to the same extent as controls (**Figure 21**).

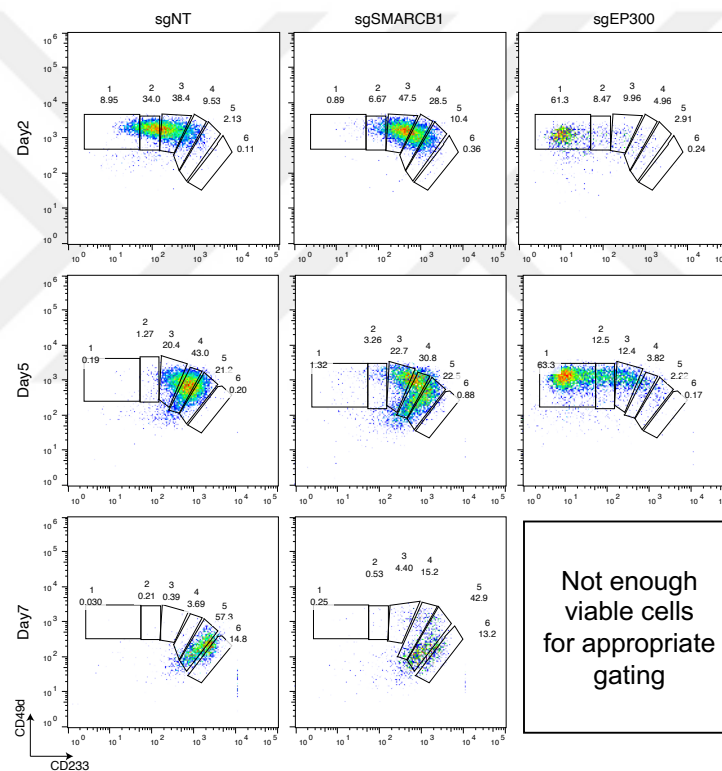


Figure 21 – EP300 knock-out inhibits differentiation of BEL-A cells.

Additionally, EP300 staining confirmed a strong reduction in endogenous EP300 levels in sgEP300 cells (**Figure 22-A**). Consistent with impaired erythroid maturation, differentiated EP300-KO cell pellets did not display the characteristic red coloration observed in BEL-A WT cells (**Figure 22-B**), indicating reduced hemoglobin production.

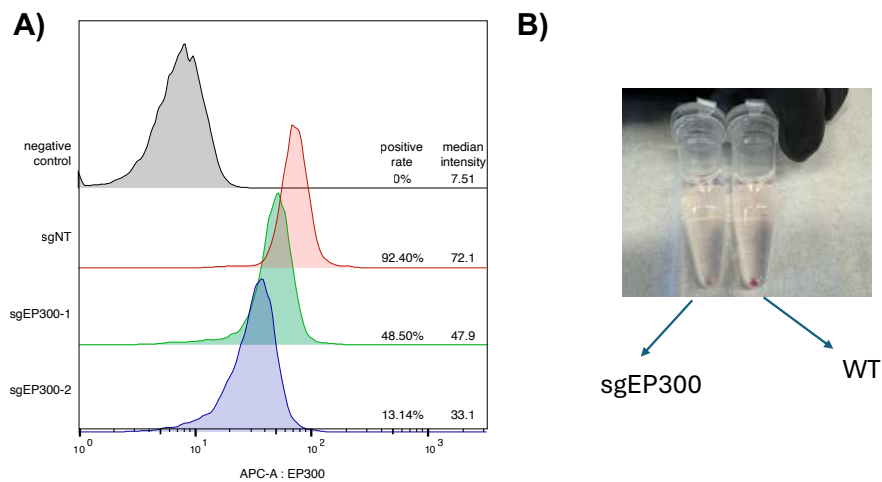


Figure 22 – EP300 knock-out inhibits hemoglobin production in differentiating BEL-A cells

These findings reinforce the conclusion that EP300 is required at early differentiation stages, whereas SMARCB1 functions later, consistent with our screening results.

3.6. cBAF/pBAF complex members represent a late-stage erythroid dependency

Since we identified SMARCB1 as a late-stage enucleation dependency—and observed the depletion of several other SWI/SNF (BAF) complex components across different comparisons—we next examined how individual SWI/SNF subunits behaved in orthochromatic erythroblasts versus pyrenocytes, focusing on potential complex-specific dependencies.

We first organized the SWI/SNF subunits according to their known complex affiliations:

Common core members: *SMARCA4*, *SMARCA2*, *SMARCC1*, *SMARCC2*, *BCL7A*, *BCL7B*, *BCL7C*, *ACTB*, *ACTL6A*, *SS18*, *SMARCD1*, *SMARCD2*, *SMARCD3*, *SS18L1*, *SS18L2*

pBAF-specific members: *ARID2*, *PHF10*, *BRD7*, *PBRM1*

ncBAF-specific members: *GLTSCR1*, *GLTSCR1L*, *BRD9*

cBAF/pBAF shared members: *SMARCE1*, *SMARCB1*, *DPF1*, *DPF2*, *DPF3*

cBAF-specific members: *ARID1A*, *ARID1B*

We then compared the depletion of these subunits in CRISPR dependency data from orthochromatic erythroblasts and pyrenocytes. As expected, most *core/common* SWI/SNF members showed similar depletion levels in both contexts (**Figure 23-A**), consistent with their broad roles in chromatin remodeling. However, when we ranked genes by the difference in log2 fold-change (Pyrenocyte – Ortho), a striking pattern emerged: cBAF/pBAF shared subunits were disproportionately more depleted in pyrenocytes (**Figure 23-B**). This trend suggests that late-stage erythroid differentiation and/or enucleation imposes an increased requirement for the cBAF/pBAF shared proteins, beyond a general SWI/SNF dependency, or a target specific dependency.

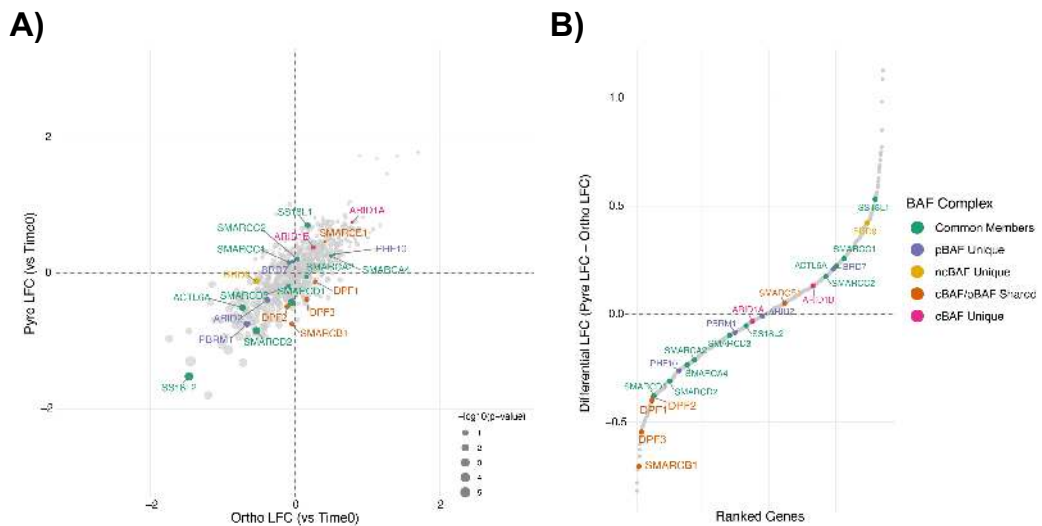


Figure 23 – Results of SWI/SNF complex member genes in CRISPR screen

A) Distribution of SWI/SNF complex members in Ortho and Pyro comparisons.

B) Ranked plot showing SWI/SNF members in top depleted Pyro hits.

3.7. SMARCB1 is required for enucleation stage of erythropoiesis

To determine how SMARCB1 loss affects the terminal stages of erythropoiesis, we evaluated enucleation efficiency of BEL-A cells. First, the efficiency of sgSMARCB1 was checked with flow cytometry, and a 50% decrease of endogenous SMARCB1 staining was observed (**Figure 24-A**). In sgNT control cells, differentiation proceeded normally, culminating in a high proportion of enucleated pyrenocytes (42%), indicative

of successful terminal maturation (**Figure 24-B**). In contrast, sgSMARCB1-edited cells showed a significant reduction in enucleation (16% - 19%), with a substantial proportion of cells arrested at the orthochromatic erythroblast stage and failing to extrude their nuclei (**Figure 24-C**).

Together with the temporal dependencies observed in our CRISPR-based assays, these findings demonstrate that SMARCB1 is required for the late-stage enucleation program, functioning specifically at the final transition from orthochromatic erythroblast to enucleated pyrenocyte.

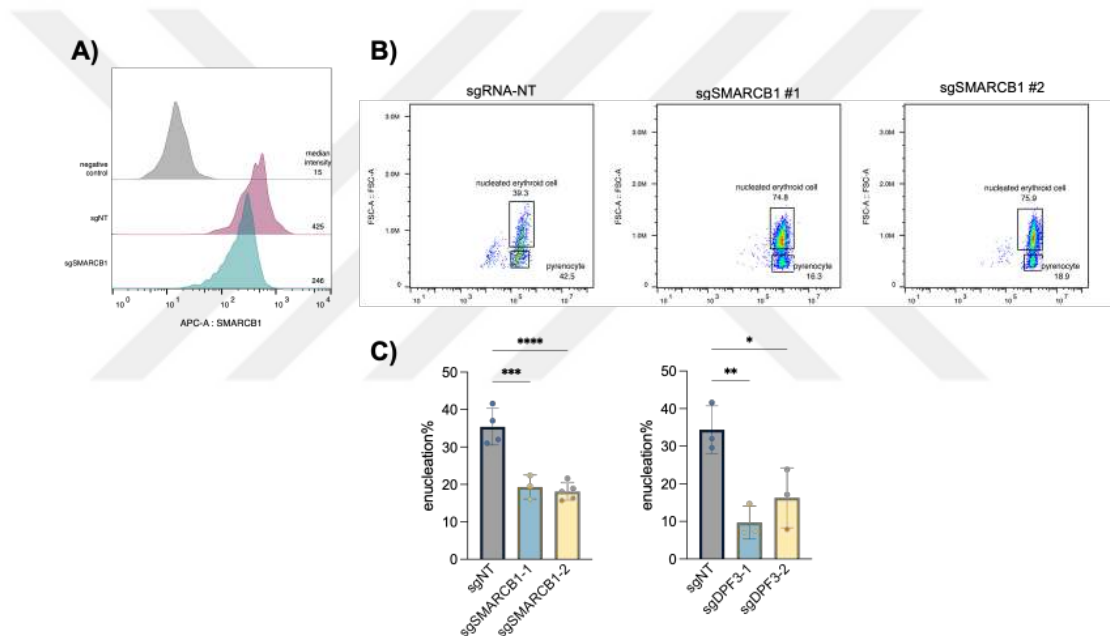


Figure 24 – SMARCB1 knock-out inhibits enucleation

A) Endogenous SMARCB1 staining in SMARCB1-KO cells. >50% reduction in SMARCB1 expression. B) Pyrenocyte and nucleated erythroid cell percentages in sgNT and sgSMARCB1 BEL-A cells. C) Quantification of enucleation. Reduced pyrenocyte formation upon SMARCB1 and DPF3 knockout (~50% vs control). One-way ANOVA with Sidak multiple comparisons test; adjusted *P* values shown. Mean $\pm$ SD, 3 independent experiments. *P* < 0.05(*), *P* < 0.01(**), *P* < 0.001(***), *P* < 0.0001(****).

3.8. Targeted experiments validate CRISPR screen results in primary HSCs

To evaluate our CRISPR screen findings in primary hematopoietic cells, we first characterized the erythroid differentiation trajectory of CD34⁺ hematopoietic stem and progenitor cells (HSPCs). Cells were differentiated for 13 days, and progression was monitored using surface expression of GYPA (Glycophorin A/CD235a) and CD71 (Transferrin receptor). As expected, GYPA expression continuously increased until day 11, whereas CD71 expression rose until day 8 and then declined through day 13, consistent with the transition from proliferative early erythroblasts to late, more mature stages (**Figure 25-A & 25-B**). By day 13, the fold increase of CD235a⁺ cells relative to CD235a⁻ cells reached approximately 10-fold, indicating robust maturation toward late erythroid states (**Figure 25-C**).

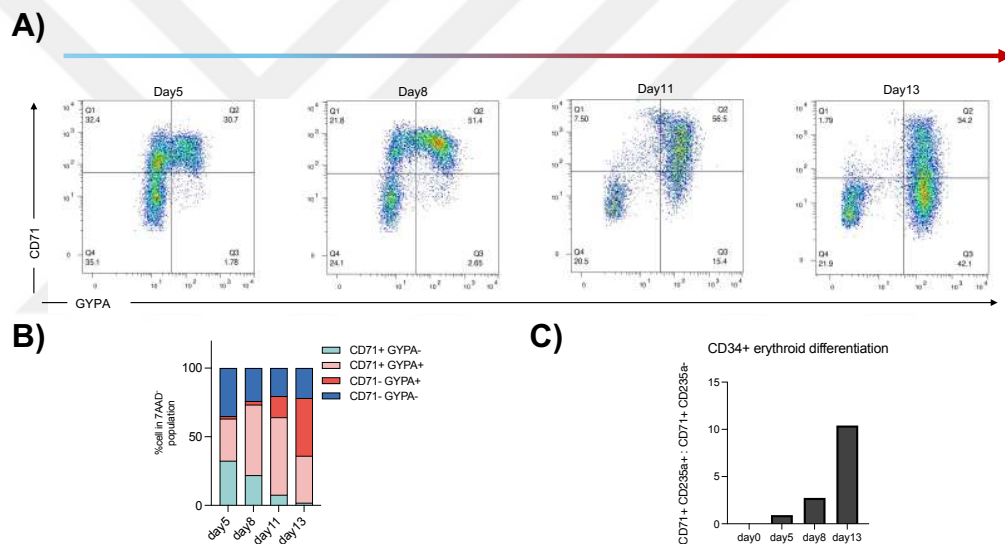


Figure 25 – Erythroid differentiation of primary CD34⁺ HSCs

A) GYPA and CD71 staining in erythroid differentiation-induced HSCs. Dynamic CD71 expression with initial increase followed by decline during maturation; progressive increase in GYPA expression. B) Rate of erythroid differentiation. C) Quantification of CD71⁺ CD235a⁺ and CD71⁺ CD235a⁻ erythroid populations during differentiation.

Next, we investigated how SMARCB1 knockout influences the differentiation and enucleation of primary erythroid cells. CD34⁺-derived erythroid progenitors were electroporated with either control or SMARCB1-targeting sgRNA RNPs using a GFP-tagged Cas9 protein, and GFP⁺ cells were sorted 8 hours later (**Figure 26**).

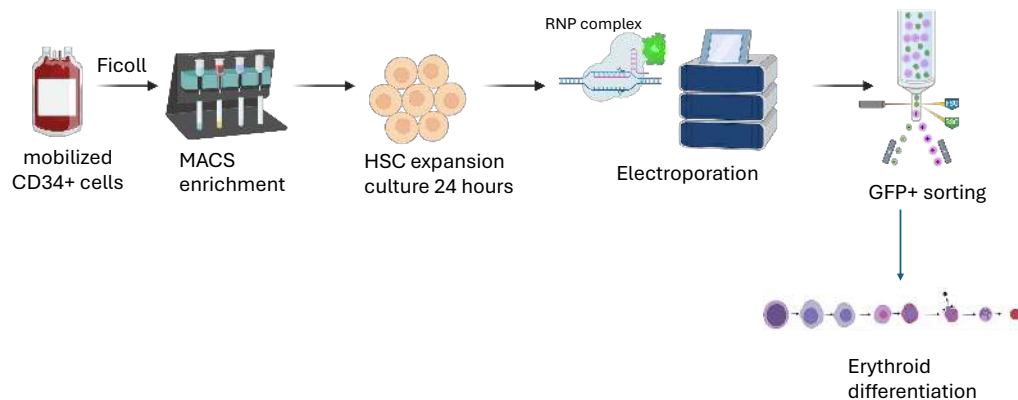


Figure 26 – Experimental flow of HSC experiments

Electroporation efficiency was first assessed by flow cytometry, revealing that >98% of cells were GFP⁺ (**Figure 27-A**). Following sorting, cells were induced to undergo erythroid differentiation. To confirm the sgRNA activity, we examined the targeted loci by Sanger sequencing and detected sequence alterations downstream of the expected Cas9 cut sites (**Figure 27-B**). Quantification of editing efficiency using TIDE analysis (Methods) showed that the polyclonal populations contained 44% and 51% DNA edits (**Figure 28**), indicating robust and efficient SMARCB1 disruption.

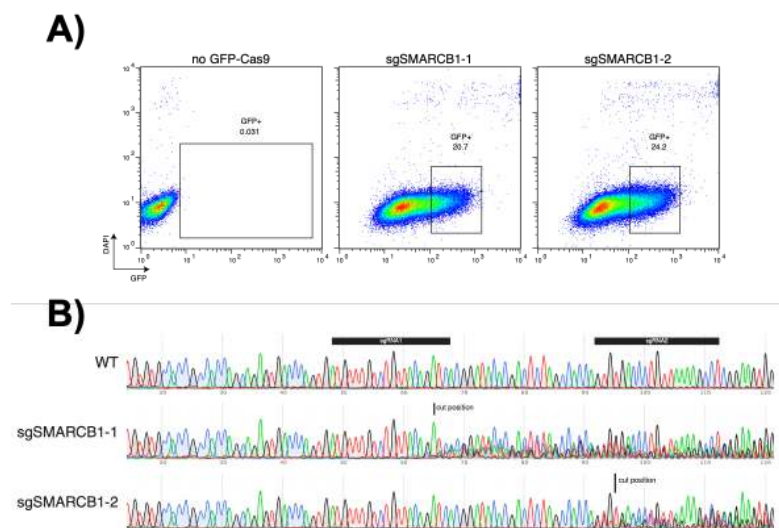


Figure 27 – Electroporation of HSCs with RNP complexes and Cas9-GFP

A) Electroporation efficiency of RNP complexes and sorted cell populations. B) Sanger sequencing results of regions targeted by SMARCB1 gRNAs.

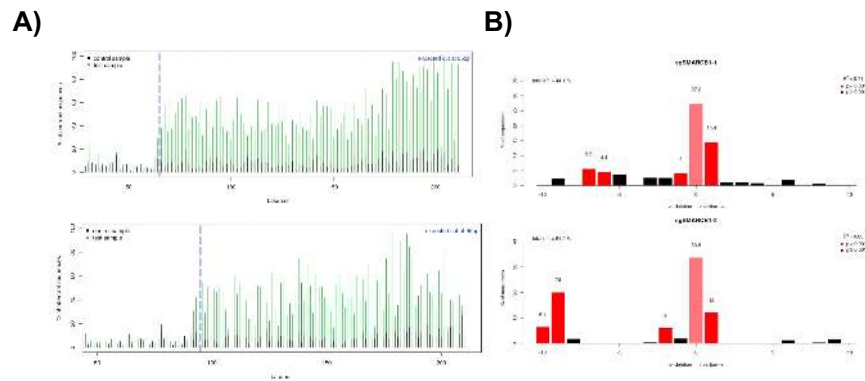


Figure 28 – TIDE analysis results of sgSMARCB1 induced DNA edits

A) TIDE analysis of aberrant DNA sequences with SMARCB1 gRNAs. B) Deletion/insertion events in SMARCB1 gene with gRNAs.

First, we evaluated how sgSMARCB1 and sgEP300 alter erythroid differentiation of HSCs. CD233 and CD49d staining revealed that while SMARCB1 KO did not affect the percentages of mature erythroid cells and proerythroblasts after 7 days of differentiation, EP300 knockout cells accumulated at the proerythroblast stage (**Figure 29**).

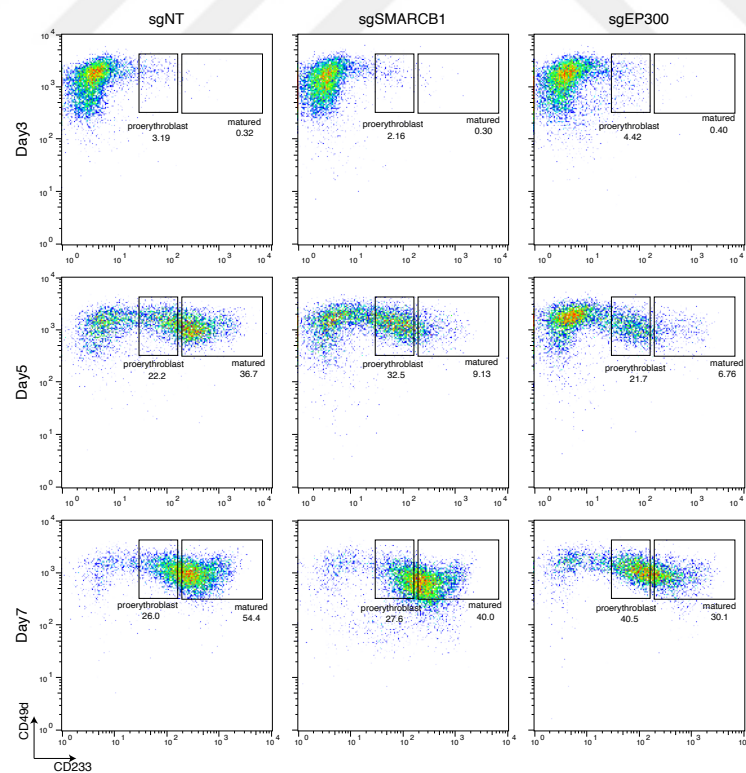


Figure 29 – sgEP300 reduces the differentiation of HSCs

Flow cytometry analysis of erythroid differentiation in day3, 6 and 7 using CD233 and CD49d surface markers. EP300 perturbation results in delayed erythroid maturation, whereas SMARCB1 perturbation shows differentiation profiles comparable to control cells, consistent with earlier observations.

Then, we assessed how SMARCB1 knockout affects erythrocyte production during primary HSPC differentiation. In control sgNT-transfected cells, 50.9% of the population differentiated into erythrocytes after 14 days. In contrast, sgSMARCB1-edited cells generated only 25% and 31% erythrocytes in two independent replicates (**Figure 30**). These findings demonstrate a substantial impairment in erythrocyte output upon SMARCB1 loss, consistent with its essentiality in late-stage erythroid maturation.

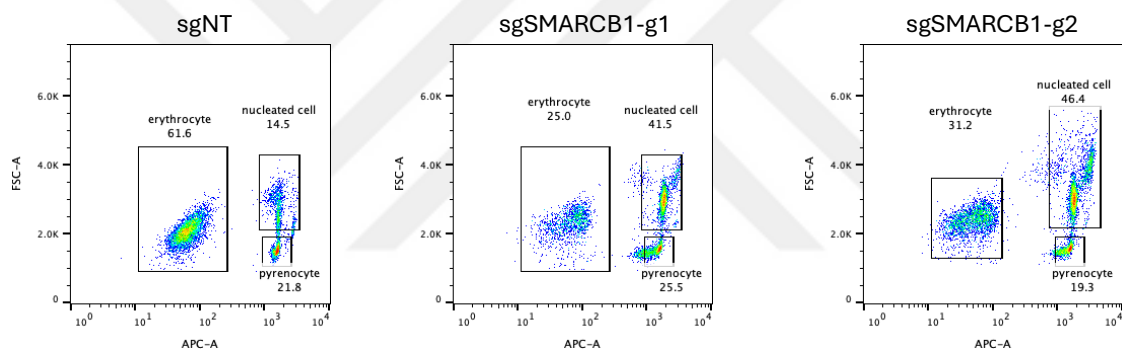


Figure 30 – Erythrocyte production of sgNT and sgSMARCB1 transfected differentiated HSCs

Finally, histological assessment clarified how EP300 and SMARCB1 regulate erythroid differentiation in primary HSCs. After 14 days of differentiation, approximately 60% of wild-type (WT) cells were classified as mature erythrocytes, indicating efficient completion of the differentiation program (**Figure 31-A, 31-D**). In SMARCB1 knockout cultures, we observed a modest accumulation of basophilic, polychromatic, and orthochromatic erythroblasts, accompanied by a significant reduction in mature erythrocytes (**Figure 31-B, 31-D**), consistent with a block in late-stage maturation. In contrast, EP300 knockout cells showed a pronounced accumulation at the basophilic erythroblast stage, indicating an earlier differentiation defect (**Figure 31-C, 31-D**).

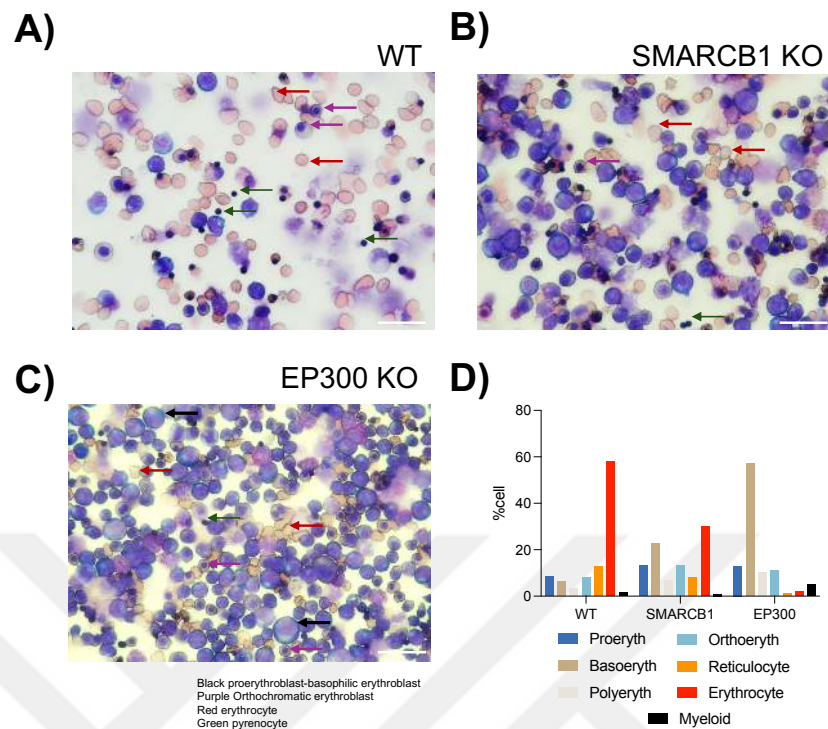


Figure 31 – Histological assessment of SMARCB1 and EP300 knock-out on HSC differentiation

- A) Histological evaluation of WT HSC cells after 14 days differentiation.
 B) Histological evaluation of SMARCB1 KO HSC cells after 14 days differentiation.
 C) Histological evaluation of EP300 KO HSC cells after 14 days differentiation.
 D) Quantification of results. Scale bars in (A), (B) and (C): 20 microns.

3.9. scRNA-seq reveals stage specific expression of CRISPR screen hits

To better evaluate the stage specific roles of CRISPR screen hits, a single cell RNA-sequencing experiment was conducted with using primary human CD34⁺ hematopoietic stem cells. Cells were differentiated for 0, 8, 16 days, and all samples were collected and combined. Single cell RNA-seq libraries were prepared, sequenced and analyzed, described in detail in Methods section.

When results were analyzed with UMAP dimensionality reduction approach, we observed that mRNA expression information of 13,807 cells passed the quality filters, but the counts of captured late-stage orthochromatic erythroblasts were lower than expected (**Figure 32-A**). Then, we combined the acquired results with a previously

published public dataset by Huang et al. (GSE150774) (Huang et al., 2020) (**Figure 32-A**). In the combined dataset, we were able to capture a total of 43,663 cells, and an increased total number of erythroid differentiated cells as polychromatic erythroblasts and orthochromatic erythroblasts (**Figure 32-B**). Cell cycle phase associated gene expression patterns also revealed an accumulation of G1 cells in orthochromatic erythroblast population, confirming the cell annotations (**Figure 32-C**).

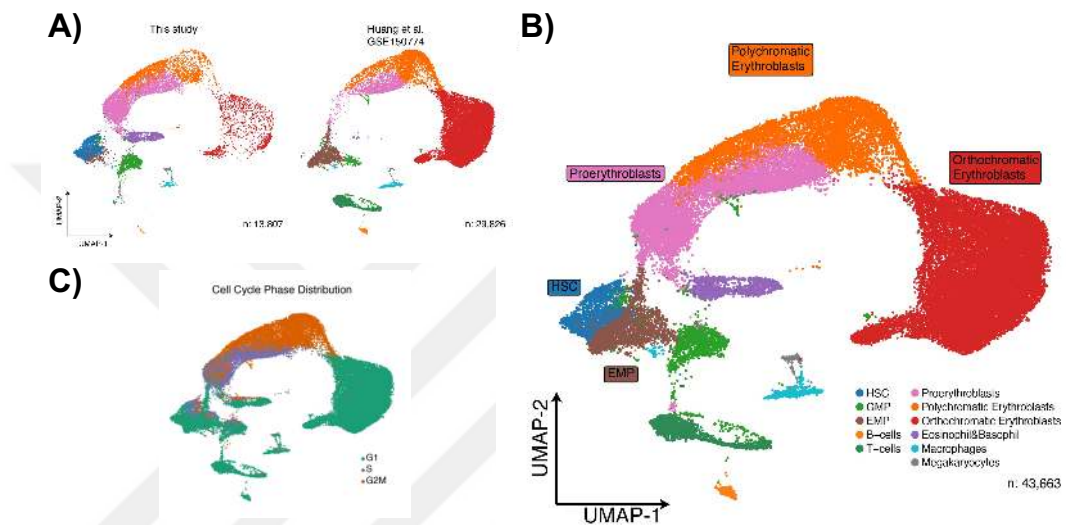


Figure 32 – Integrated scRNA-seq analysis of erythroid differentiation induced HSCs
A) Comparison of cell distribution between studies. B) UMAP visualization of different clusters. C) Cell cycle phase distribution of cells in different clusters.

Then, we analyzed the expression patterns of top cluster specific expressed genes. CD34 was mostly expressed in non-differentiated hematopoietic stem cells, and G2 phase associated Aurora kinase A (AURKA) and NUSAP1 was mostly expressed in polychromatic erythroblasts, reflecting the mid-stage erythroid dividing cells (**Figure 33**). XPO7 and ARL4A was among the top expressed genes in orthochromatic erythroblasts, showing the late-stage expression patterns of erythroid differentiation (**Figure 33**).

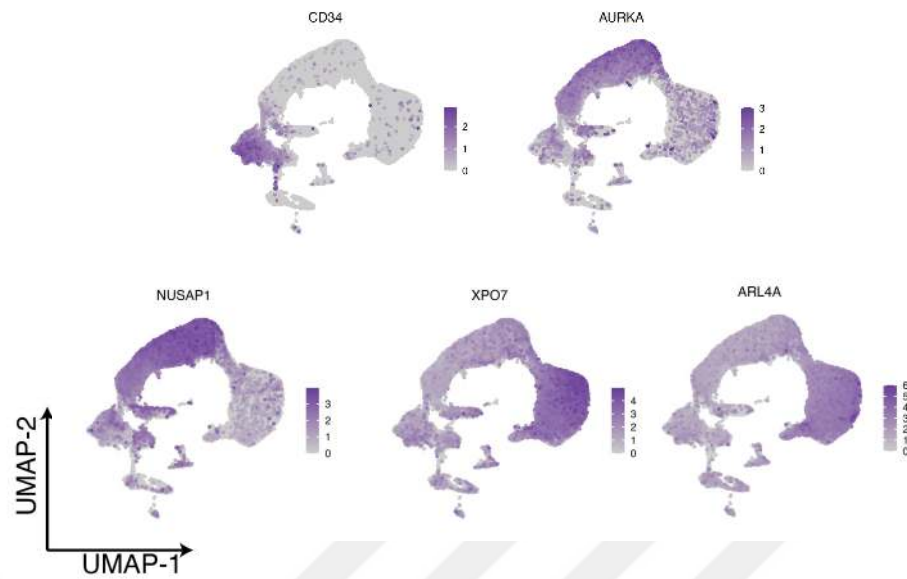


Figure 33 – Cell stage specific markers of erythroid cells

Analysis of different CRISPR screen hits revealed different expression patterns. UHRF1 expression was limited to proerythroblasts and polychromatic erythroblasts, with no detected expression in orthochromatic erythroblasts. EP300 and SS18L2 were widely expressed in many cell types, including orthochromatic erythroblasts. SETD8 was mostly expressed in polychromatic erythroblasts and with a low percentage of expressing orthochromatic erythroblasts. SMARCB1 was mostly expressed in erythroid progenitor cells, proerythroblasts, and polychromatic erythroblasts, with a low expression in orthochromatic erythroblasts. Finally, DPF3 was mostly expressed in proerythroblasts and polychromatic erythroblasts (**Figure 34**). Other than the erythroid lineage, some of these epigenetic regulators, as examples, EP300 and SMARCB1, were also expressed in myeloid cells as megakaryocytes (**Figure 35**). We also checked how expression of key erythroid transcription factors was changed in different erythroid cell stages. While GATA1, NFE2, TAL1, and KLF1 were highly expressed across all stages of erythroid differentiation, GATA2 expression was predominantly detected in granulocytes (**Figure 36**). LDB1 showed its highest expression in proerythroblasts and progressively decreased as cells advanced through differentiation (**Figure 36**).

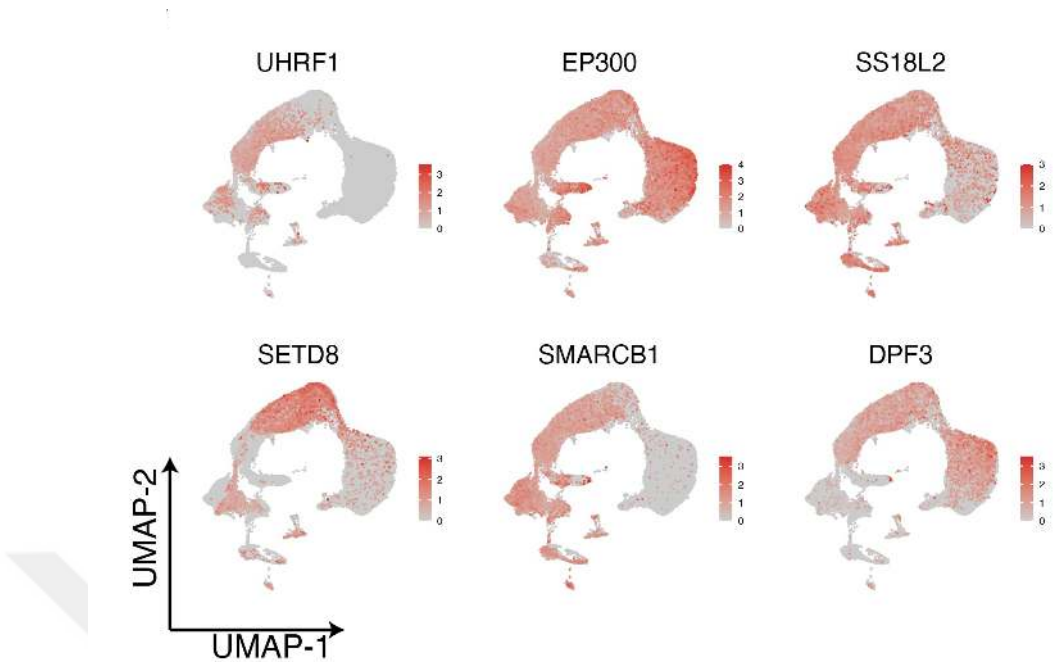


Figure 34 – Stage specific expression patterns of CRISPR screen hits in different cell clusters

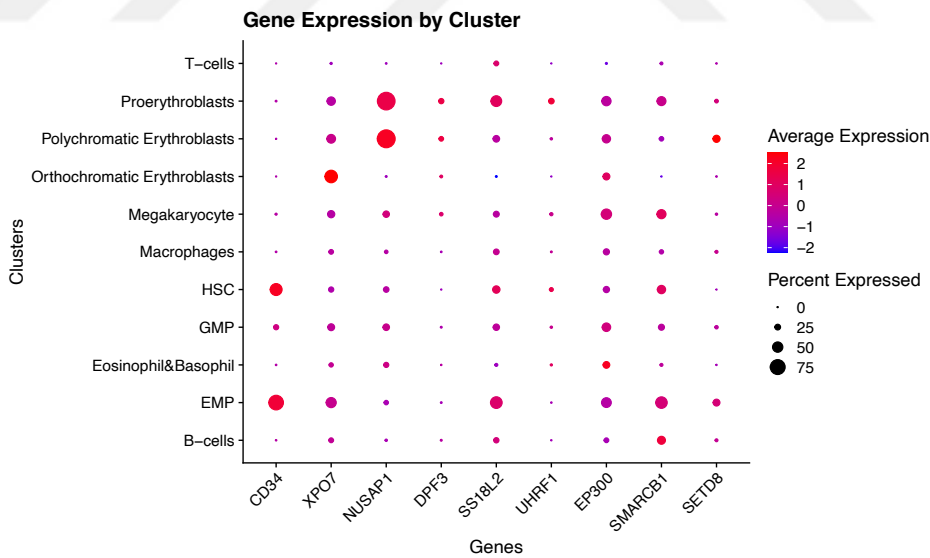


Figure 35 – Expression of CRISPR screen hits in different cell clusters

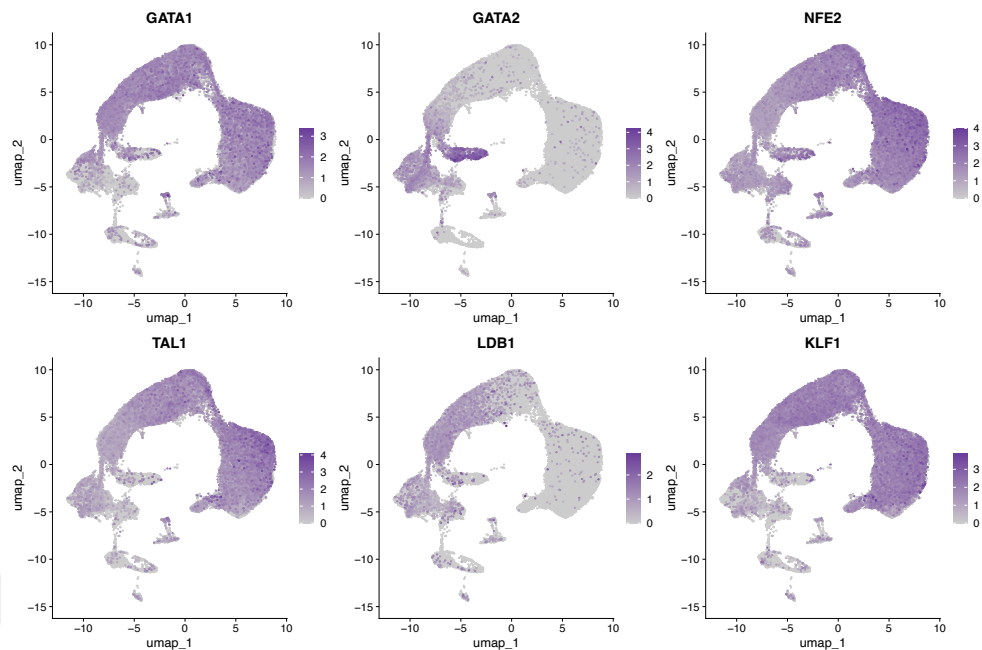


Figure 36 – Expression of key erythroid transcription factors in different stages

When results were analyzed by a pseudo-bulking approach, AURKA, CDCA8, HMMR, CKAP2, and NEK2 was identified as top differential expressed genes in polychromatic erythroblasts, and CLIC2, ARL4A, HBP1, FOXO3, and TBCEL was identified as top differential expressed genes in orthochromatic erythroblasts, compared to proerythroblasts (**Figure 37-A**). Gene ontology (GO) analyses revealed the enrichment of cell division associated terms as mitotic nuclear division and chromosome segregation among the upregulated genes in polychromatic erythroblasts, and terms as autophagy, ubiquitin-dependent protein catabolic process and vesicle organization in orthochromatic erythroblasts (**Figure 37-B**).

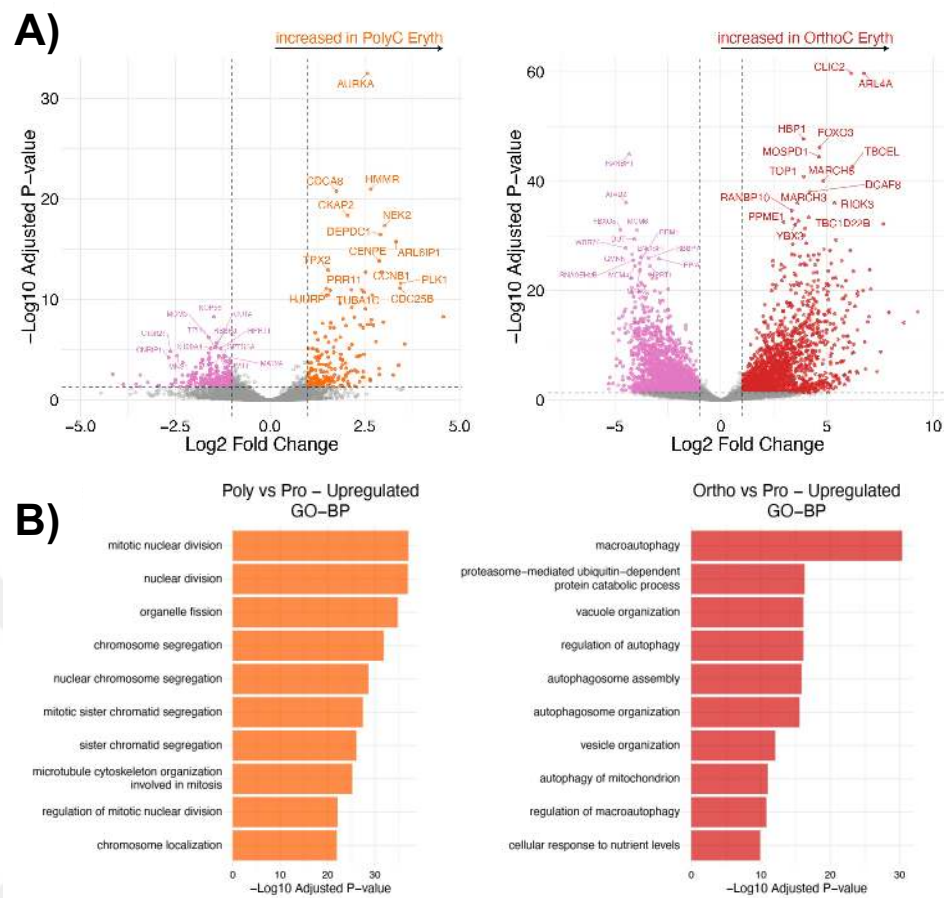


Figure 37 – Gene expression changes in erythroid differentiation stages

A) Upregulated genes in different erythroid differentiation stages. B) GO-BP enrichment analysis of upregulated genes

3.10. Six days of BEL-A differentiation reflects a transcriptomic signature of orthochromatic erythroblasts

To characterize transcriptomic changes associated with BEL-A cell differentiation, we performed bulk RNA sequencing on sgAAVS1-expressing BEL-A-Cas9 cells after 6 days of in vitro differentiation and compared them to undifferentiated controls. As expected, genes involved in erythroid maturation and heme biosynthesis, including HBB, HBA1, HBA2, HBM, HBD, FTL, and ALAS2 were significantly upregulated (Figure 38).

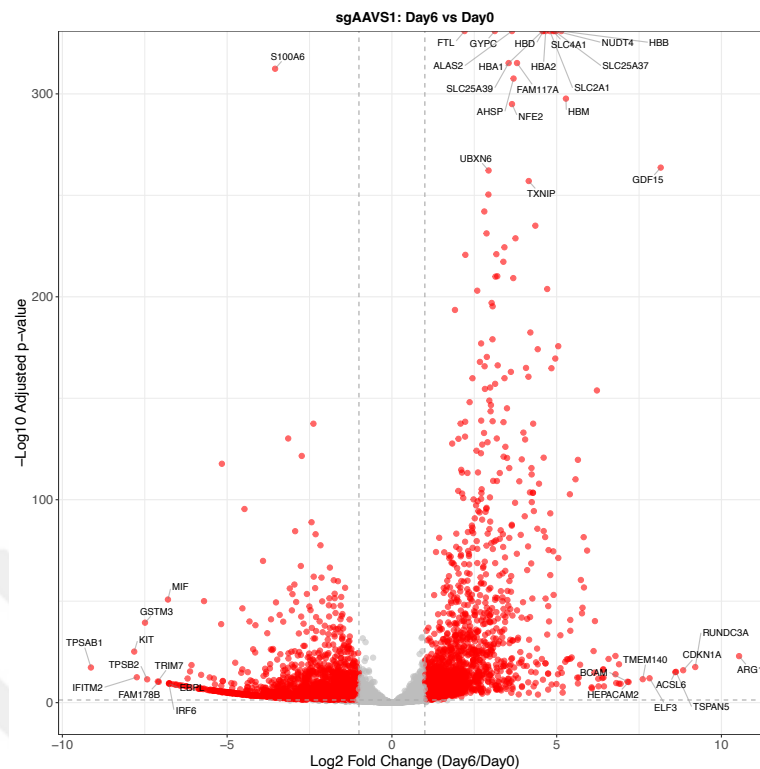


Figure 38 – Differentiation induced gene expression changes in BEL-A cells

To place these changes in the context of our single-cell RNA-seq–defined differentiation signatures, we next examined the expression of polychromatic- and orthochromatic erythroblast–associated genes. As anticipated, polychromatic erythroblast proliferation–associated genes (AURKA, CDCA8, HMMR, NEK2, NUSAP1, CKAP2, and CENPE) was increased slightly following differentiation, indicating that the cells had progressed beyond this proliferative stage. In contrast, orthochromatic erythroblast–specific genes, including ARL4A, RIOK3, XPO7, CPEB4, FOXO3, and CLIC2, were significantly upregulated more (**Figure 39**).

Together, these results demonstrate that 6 days of differentiation is sufficient to drive BEL-A cells to an orthochromatic erythroblast–like transcriptional state.

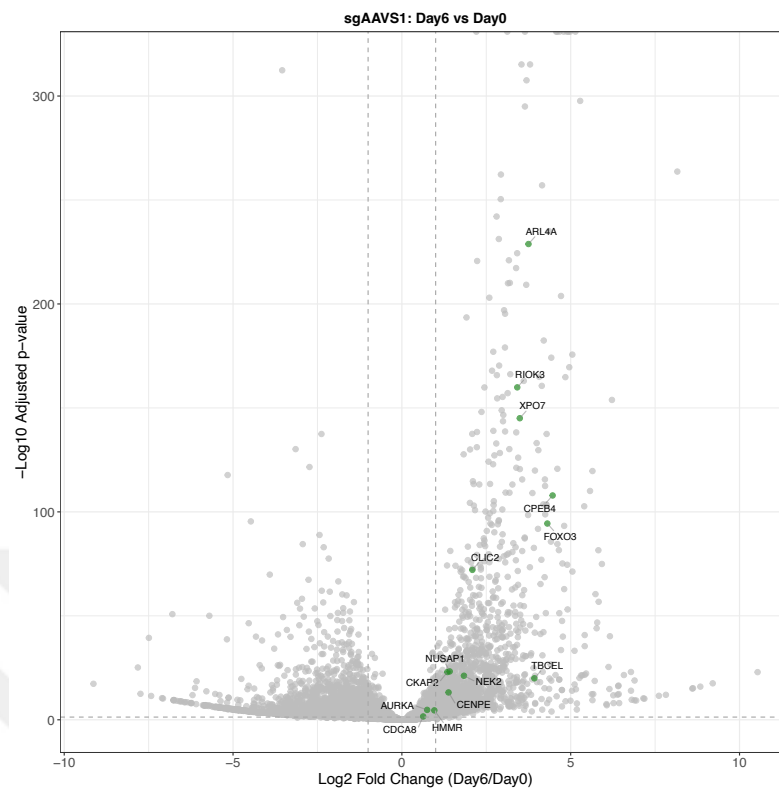


Figure 39 – Expression of polychromatic and orthochromatic erythroblast markers from scRNAseq in differentiated BEL-A cells

3.11. EP300 knock-out results with changes in differentiation associated gene expressions

Since EP300 emerged as a key differentiation dependency in our CRISPR screen, we next sought to evaluate how loss of EP300 influences the transcriptome during erythroid maturation. To this end, we performed RNA sequencing on sgAAVS1 (control) and sgEP300 BEL-A-Cas9 cells under both undifferentiated and 6-day differentiated conditions. This design allowed us to assess (i) the transcriptional impact of EP300 knockout at baseline and (ii) how EP300 loss alters the differentiation-associated gene expression program.

First, we examined how EP300 knockout affected the transcriptome of undifferentiated BEL-A cells. EP300 loss resulted in modest downregulation of several genes, including *LXN*, *GSTM3*, *CD36*, *HBB*, and *EPB41L3*, alongside modest upregulation of genes such as *CAI*, *JUN*, *EGR1*, *ACTG1*, and *LGALS1* (**Figure 40**). These relatively subtle changes

indicate that EP300 knockout has limited impact on the baseline transcriptome prior to the onset of erythroid differentiation.



Figure 40 – Effect of EP300 knock-out on the transcriptome of undifferentiated BEL-A cells

We next compared the polychromatic and orthochromatic erythroblast gene-expression signatures defined from our single-cell dataset to determine how EP300 loss affects stage-specific transcriptional programs (**Figure 41**). In control sgAAVS1 cells, all polychromatic-associated genes were robustly upregulated upon differentiation, consistent with normal progression through intermediate erythroid stages. In contrast, sgEP300 cells failed to induce these polychromatic signatures (**Figure 41**). A similar trend was observed for orthochromatic erythroblast-specific genes, which increased as expected in sgAAVS1 cells upon differentiation but were markedly attenuated in sgEP300 cells (**Figure 41**).

Together, these data demonstrate that EP300 knockout disrupts both polychromatic and orthochromatic transcriptional programs, confirming a strong inhibitory effect on erythroid differentiation.

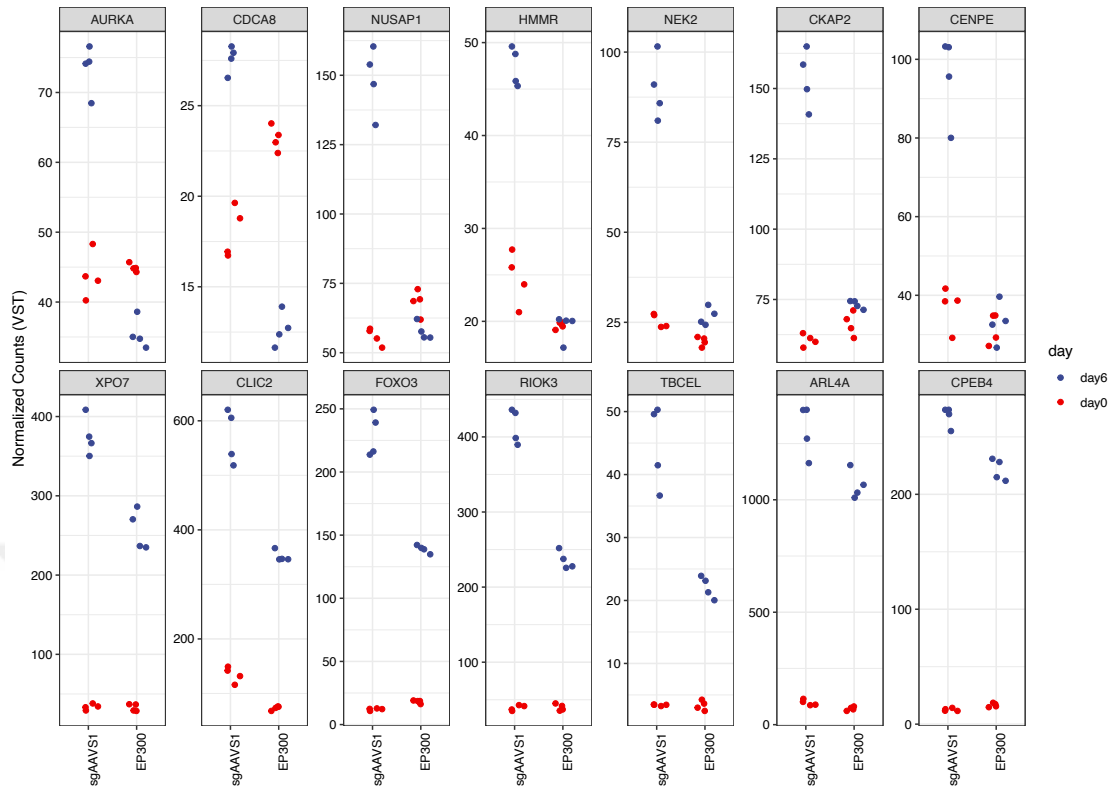


Figure 41 – sgEP300 associated gene expression changes in differentiating BEL-A cells

Top panel: Polychromatic erythroblast-associated genes.

Bottom panel: Orthochromatic erythroblast-specific genes

3.12. D&D-Seq infers GATA1 binding events in HSC cells

Given that GATA1 is the central transcription factor governing erythroid lineage commitment and differentiation, we next evaluated how our CRISPR screen hits might influence GATA1 binding and activity. To this end, we applied D&D-Seq, a single-cell multiomic method approach combining RNA-seq and ATAC-seq, to profile GATA1 binding footprints across differentiating hematopoietic cells. Cell identities were first assigned using lineage-specific RNA markers (**Figure 42-A**). We then assessed GATA1 footprints in erythroid versus myeloid populations. Analysis of 200 randomly sampled accessible peaks revealed strong GATA1 motif enrichment in erythroid cells, whereas myeloid cells displayed only background-level footprinting, consistent with lineage-specific GATA1 function (**Figure 42-B**).

Mapping the total D&D edits—representing GATA1 binding events—showed high edit counts in proerythroblasts and early erythroid cells (polychromatic erythroblasts) (**Figure 42-C & 42-D**). Correspondingly, GATA1 expression was predominantly detected in proerythroblasts, and GATA1 motif accessibility extended from proerythroblasts into early erythroid stages (**Figure 42-E**). These data confirm that our multiome dataset robustly captures the expected GATA1 activity gradient across early erythroid differentiation.

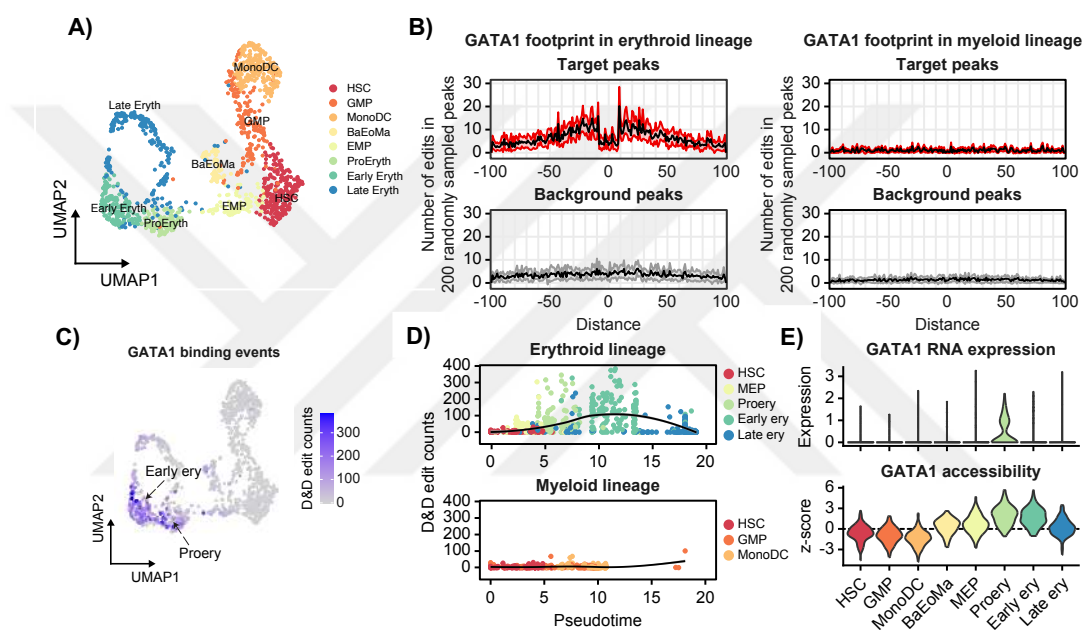
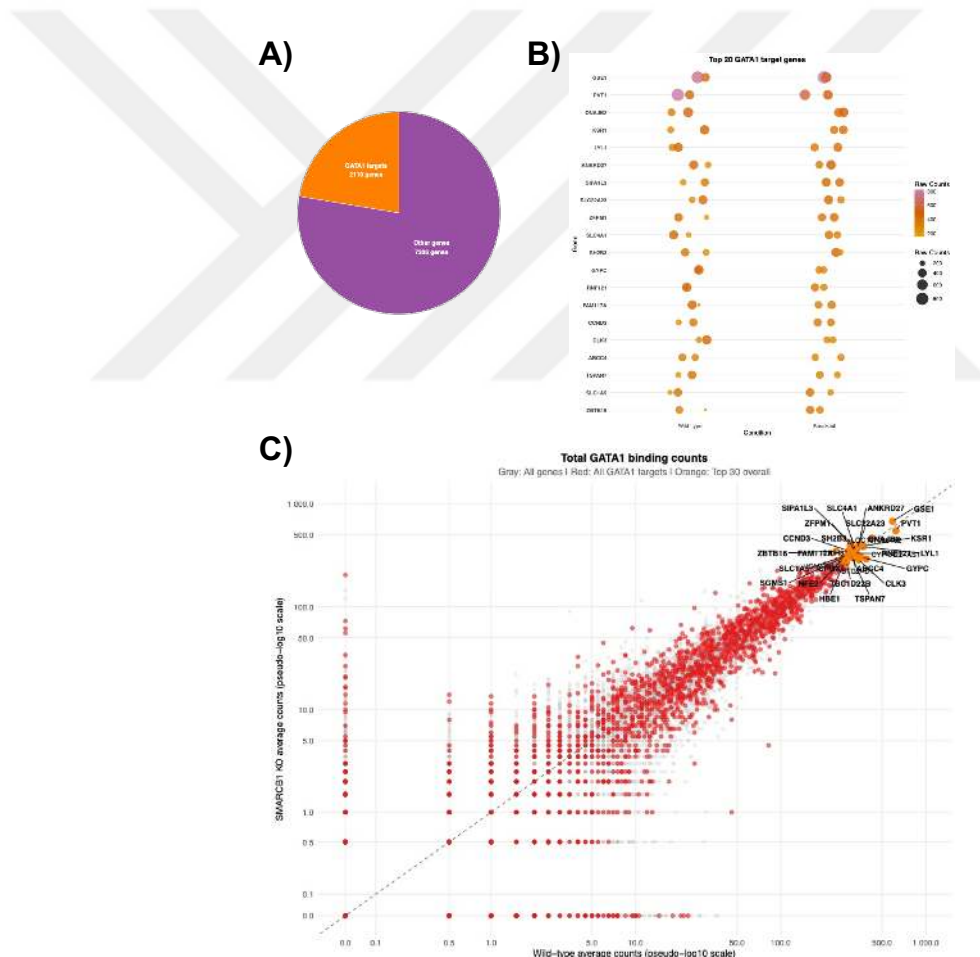


Figure 42 – D&D-Seq identifies GATA1 binding in differentiating HSCs

3.13. SMARCB1 knock-out affects GATA1 binding in BEL-A cells

To evaluate how SMARCB1 is associated with GATA1 binding, the bulk D&D method was applied. Binding sites were deaminated using a GATA1-specific antibody, as in the single-cell workflow, and deaminated regions were subsequently filtered and analyzed. First, identified regions were assessed for known GATA1 target loci, and bulk D&D analysis successfully detected binding events across 2110 gene regions (**Figure 43-A**). Among these, the top GATA1-enriched targets in both conditions included GSE1, PVT1, and DNAJB2 (**Figure 43-B**). Overall, GATA1 binding profiles between BEL-A WT and

Additionally, we evaluated whether SMARCB1 regulates endogenous GATA1 levels and observed no differences between WT and SMARCB1-KO cells (**Figure 44-A**). Consistently, analysis of ATAC reads in the D&D-Seq dataset revealed no global changes in chromatin accessibility (**Figure 44-B**). Moreover, GATA1 occupancy at its own promoter remained unchanged in SMARCB1-KO cells (**Figure 44-C**), suggesting that SMARCB1 does not regulate GATA1 expression or global accessibility in BEL-A cells.



A) Distribution of previously known GATA1 targets in combined D&D results.
B) Enrichment of top 20 GATA1 binding in WT and SMARCB1 KO BEL-A cells.
C) Comparison of GATA1 binding events in WT and SMARCB1 KO BEL-A cells.

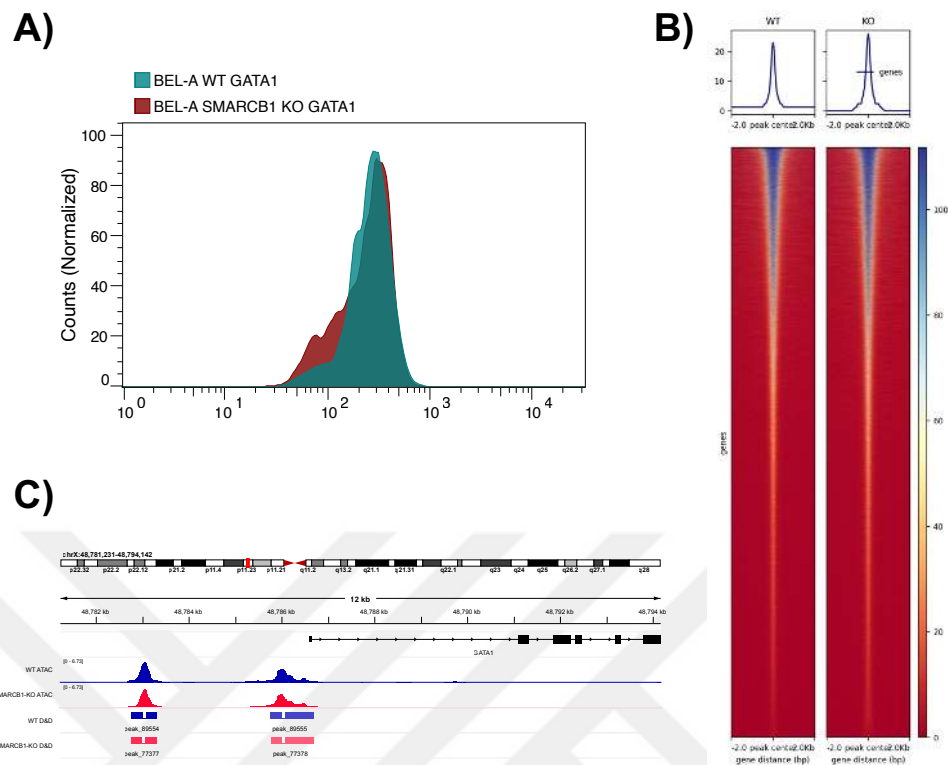


Figure 44 – SMARCB1-KO does not alter the GATA1 expression and global gene accessibility profile of the BEL-A cells

A) GATA1 levels in WT and SMARCB1 KO BEL-A cells. B) Global chromatin accessibility profile of WT and KO cells. C) GATA1 binding to its promoter is not regulated by SMARCB1.

Differential binding analysis of BEL-A SMARCB1 KO cells compared to BEL-A WT cells revealed several altered GATA1-binding events (**Figure 45-A**). SMARCB1 KO was associated with increased GATA1 binding at regions surrounding genes such as *DRAM1*, *FURIN*, and *BLVRB*, and decreased binding at loci including *ACAD9*, *CLDN10*, *PLD5*, *SNCA*, *UGCG*, and *CD226* (**Figure 45-B**).

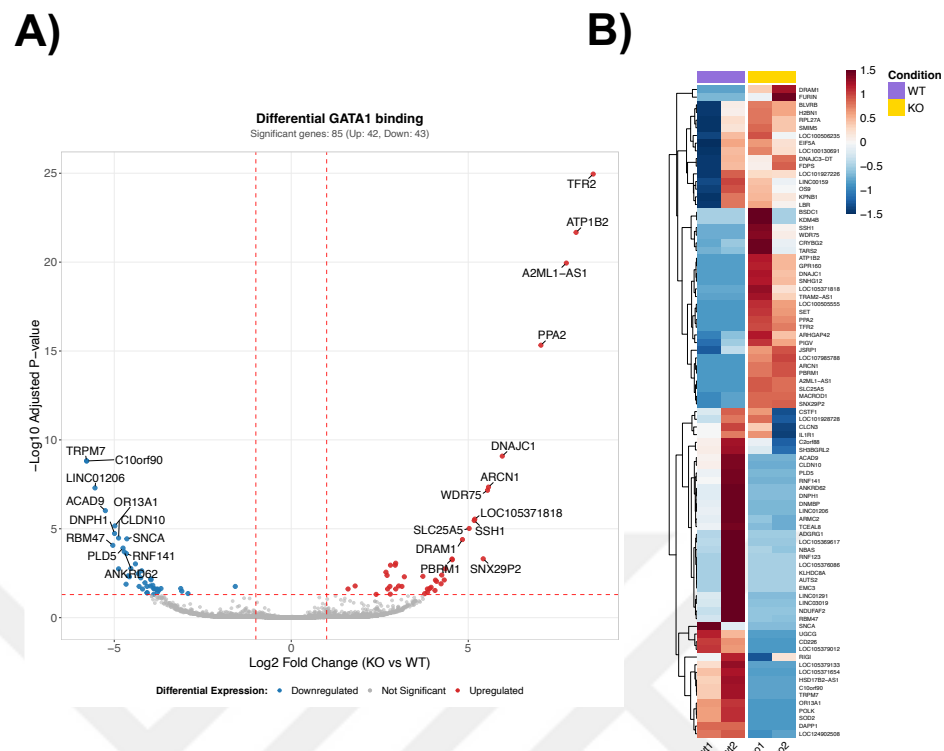


Figure 45 – SMARCB1 KO associated changes of GATA1 binding

A) Volcano plot of GATA1 binding differences in SMARCB1 KO. B) Heatmap visualization of SMARCB1 associated GATA1 binding changes

We then filtered these genes for previously annotated GATA1 binding sites to refine the results, yielding a smaller subset of high-confidence changes. The top three genes with decreased GATA1 binding were identified as *CLDN10*, *SNCA*, and *CD226*. Conversely, *TFR2*, *PPA2*, *DNAJC*, and *WDR75* showed increased GATA1 binding related D&D edits (**Figure 46-A** and **Figure 46-B**). Enrichment analysis of the D&D-Seq results revealed that downregulated GATA1-bound targets were enriched for pathways related to reactive oxygen species regulation and cytokine production, whereas upregulated targets were associated with sterol biosynthesis pathways (**Figure 47-A**). Among the genes contributing to these enriched categories, *SOD2*, *CLCN3*, and *SNCA* were linked to terms involving the regulation of reactive oxygen species (**Figure 47-B**).

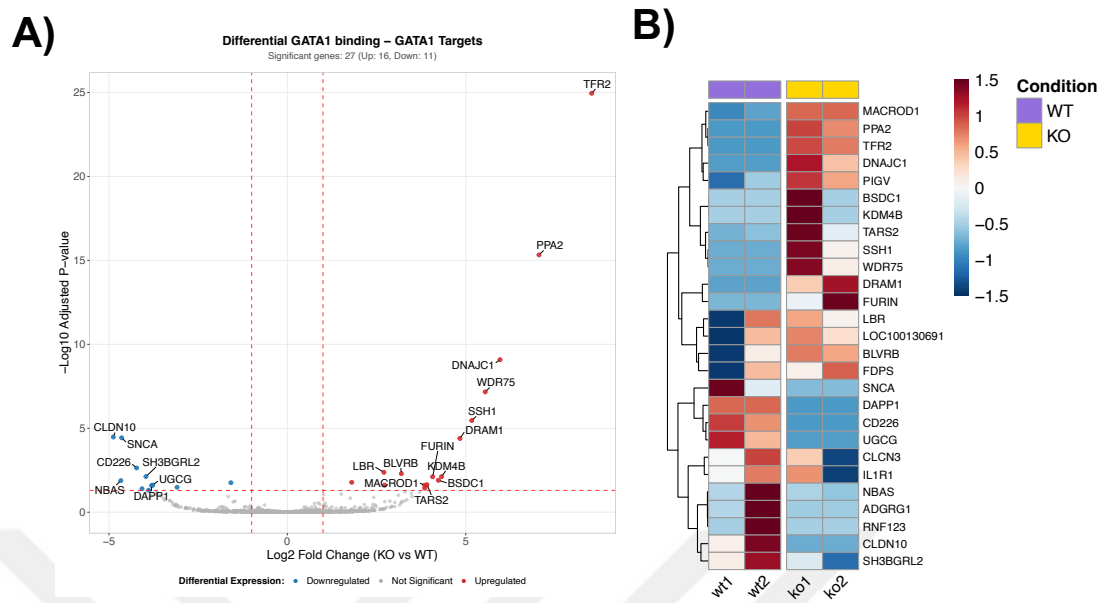


Figure 46 – SMARCB1 associated GATA1 binding changes of known GATA1 target genes

A) Volcano plot of GATA1 binding differences in SMARCB1 KO B) Heatmap visualization of SMARCB1 associated GATA1 binding changes

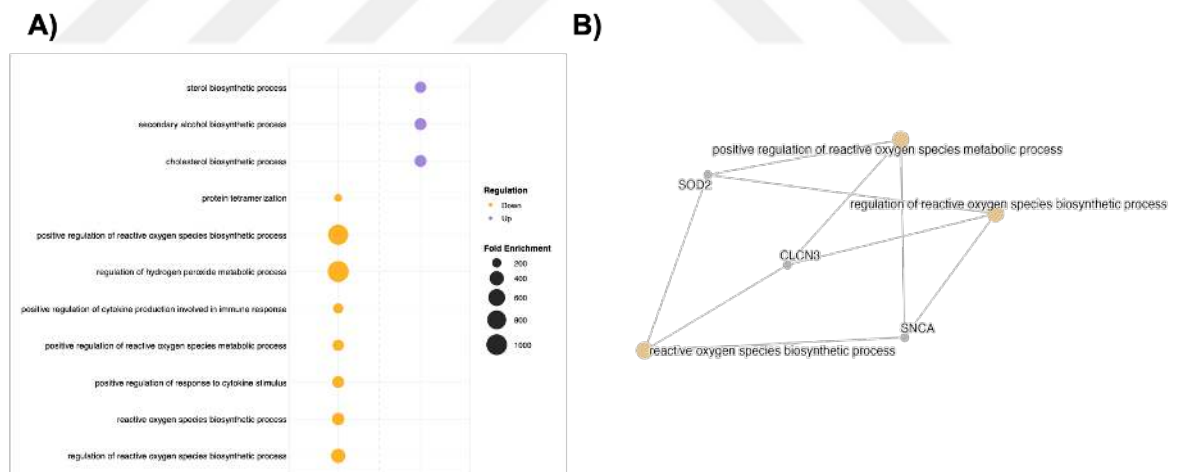


Figure 47 – Enrichment analysis of SMARCB1-KO associated D&D-Seq hits

A) GO-CC analysis of SMARCB1-KO D&D-Seq hits. B) Network analysis of ROS related genes in enrichment analysis

To determine how the expression of these genes is regulated during erythroid differentiation, we next examined their transcriptional patterns in the single-cell RNA-seq dataset. Among the top genes exhibiting increased or decreased GATA1 binding in the SMARCB1 KO D&D-Seq analysis, *TFR2* and *SNCA* were identified as highly

expressed during erythroid differentiation, consistent with their potential regulation by GATA1. In contrast, *MACROD1*, *CD226*, and *UGCG* showed low expression across erythroid populations, whereas *PPA2*, *DNAJC1*, and *DAPP1* displayed expression in myeloid-differentiated cells (**Figure 48**). Violin plot analyses further demonstrated that *SNCA* expression increases progressively with erythroid maturation, reaching its highest levels in late-stage orthochromatic erythroblasts (**Figure 49**).

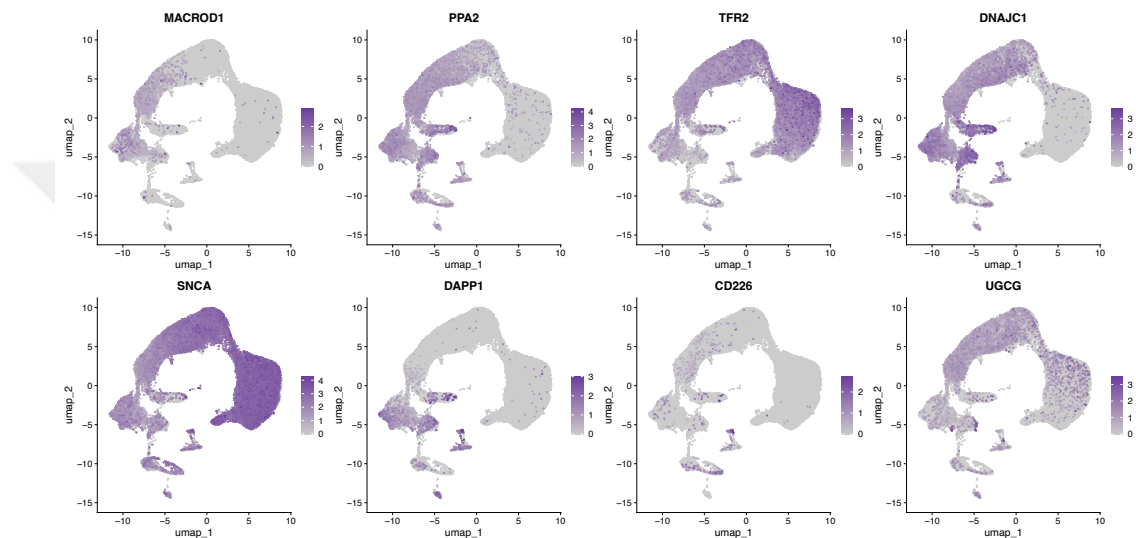


Figure 48 – Expression patterns of genes that have SMARCB1 associated GATA1 binding changes in different single cell RNA-seq clusters

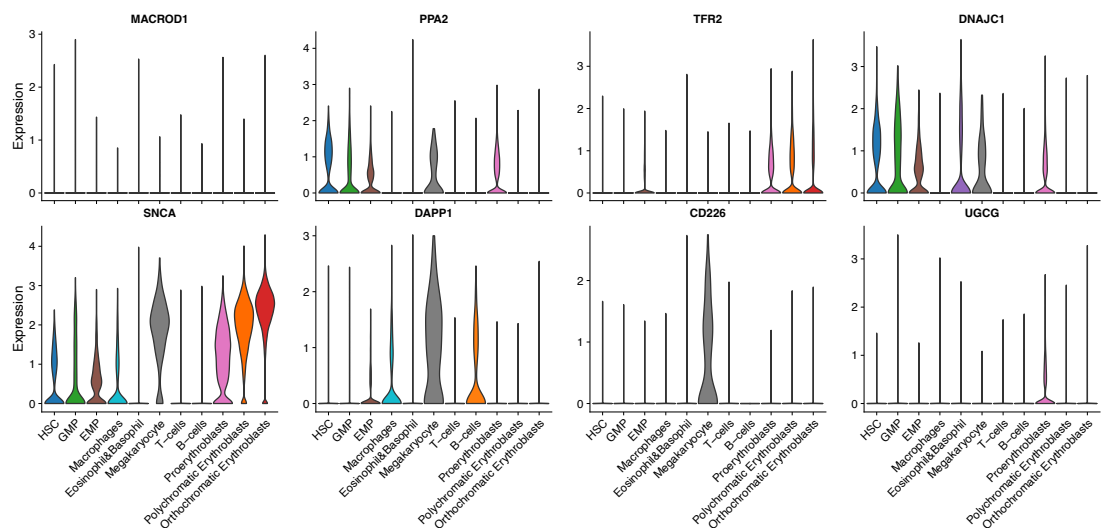


Figure 49 – SNCA expression increases with erythroid differentiation

Next, we evaluated the regulation of these genes using the scMultiome dataset. Comparison of erythroid and myeloid GATA1-binding motifs within ATAC fragments revealed that many SMARCB1-KO-associated GATA1 motif sites were preferentially enriched in erythroid cells (**Figure 50**). Furthermore, although total ATAC fragment counts were comparable across cell clusters, GATA1 D&D-Seq edits were specifically enriched in proerythroblasts and polychromatic erythroblasts. This pattern closely mirrored the differentiation-associated increase in SNCA expression, further supporting its erythroid-specific regulation (**Figure 51**).

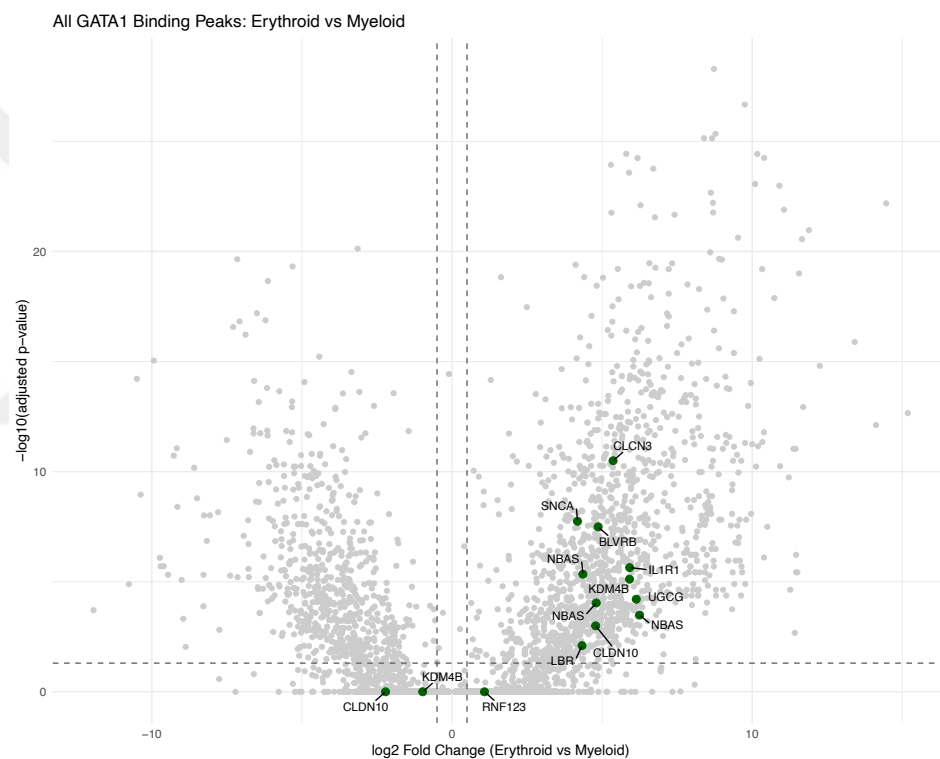


Figure 50 – SMARCB1 associated GATA1 binding events are enriched in erythroid differentiated cells

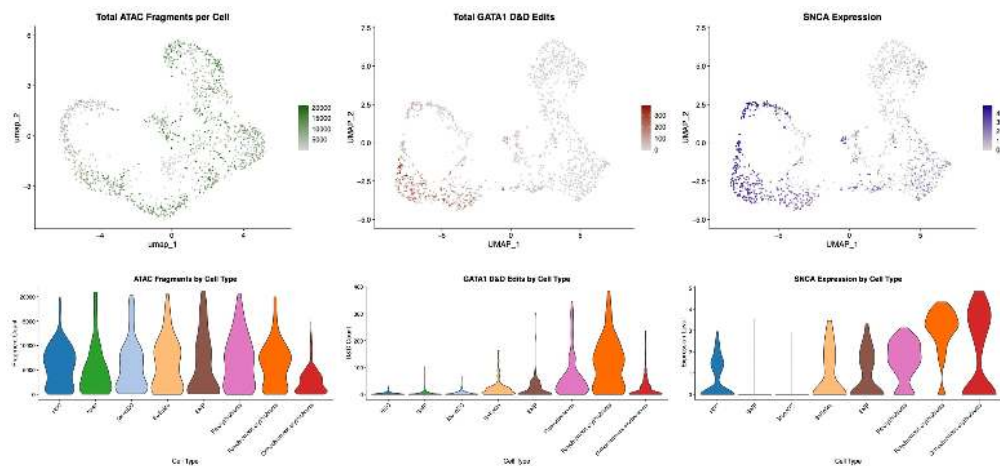


Figure 51 – GATA1 D&D edits and SNCA expression are increased with erythroid differentiation

Overall, these results indicate that SMARCB1 KO reduces GATA1 occupancy at the SNCA locus, while SNCA expression normally increases during late erythroid differentiation, particularly in orthochromatic erythroblasts. When combined with the observed defect in enucleation in SMARCB1-deficient cells, this suggests that SMARCB1 may be required to correctly activate late-stage erythroid programs—including SNCA expression—necessary for proper terminal maturation. Thus, SMARCB1 loss likely disrupts the coordination between chromatin remodeling, GATA1 binding, and gene expression required for successful enucleation.

Chapter 4

DISCUSSION

Erythropoiesis is a highly controlled differentiation process in which HSPCs progressively differentiate through morphologically and transcriptionally distinct stages to ultimately form enucleated erythrocytes from pluripotent cells. In terminal erythropoiesis, cells undergo dramatic chromatin condensation, cytoskeletal reorganization, membrane remodeling, and nuclear extrusion, however the epigenetic regulators that enable these unique processes remain incompletely defined. In this thesis, our primary goal was to identify and functionally characterize the epigenetic dependencies that control the enucleated erythroid cells. To achieve this, we applied CRISPR gRNA screen to identify epigenetic dependencies of enucleated cells by isolating and profiling pyrenocytes, the expelled nuclei of erythroblasts. This approach enabled us to uncover stage-specific epigenetic programs, including early requirements for the histone acetyltransferase EP300 and late-stage dependencies on specific SWI/SNF (BAF) chromatin-remodeling complexes. Furthermore, we uncover a previously unrecognized link between GATA1, SWI/SNF activity, and the regulation of late-stage genes such as SNCA, LBR, CD226, SH3B, and GRL2 suggesting a novel epigenetic programming for terminal erythroid maturation. Together, these findings illuminate how epigenetic regulators orchestrate the structural and functional transitions required for erythroid cells to successfully mature into erythrocytes.

In this study, we implemented a tetracycline-inducible BEL-A erythroid differentiation system to synchronize the onset of maturation and enable high-resolution interrogation of chromatin factor requirements during terminal erythropoiesis. BEL-A cells provide a homogeneous proerythroblast population with preserved GATA1/KLF1-driven regulatory programs and physiologic chromatin condensation kinetics while retaining enucleation capacity (Daniels et al., 2020). This controlled in vitro platform ensures

uniform sgRNA editing at a defined developmental stage, supports the high coverage required for chromatin-focused CRISPR libraries (Bock et al., 2022), and produces quantitative readouts that are not achievable in more heterogeneous primary systems.

Although several previous CRISPR screen studies were conducted in primary HSPCs, their differentiation trajectories exhibit substantial developmental asynchrony, lineage variability, and intrinsically low sgRNA representation (Waarts et al., 2024). These features undermine statistical power and render primary systems as not always compatible with the scale and depth required for whole genome perturbation studies, despite their high scale occurrence. In contrast, synchronized BEL-A maturation overcome these limitations by providing a uniform, reproducible, and high-throughput platform ideally suited to identify chromatin regulators acting specifically during late erythroid maturation.

In vivo CRISPR screening offers unmatched physiological context but is fundamentally restricted by pronounced clonal bottlenecks (Keys & Knouse, 2022). Following transplantation, only a small subset of transduced HSPCs engraft, resulting in stochastic expansion of a limited number of clones and distortion of sgRNA abundances independent of gene function. Clone-intrinsic fitness differences, variable lentiviral transduction efficiencies, and niche-dependent selection pressures further exacerbate this bias (Y. Wang, Xie, Armendariz, & Hon, 2022), confounding the identification of true epigenetic dependencies during terminal erythropoiesis.

Single-cell CRISPR modalities provide transcriptional resolution, yet low sgRNA capture rates and limited gRNA count markedly reduce their sensitivity (Martin-Rufino et al., 2025; Mimitou et al., 2019), particularly for rare late-stage erythroid states characterized by chromatin condensation and low RNA content while new single cell CRISPR/Cas9 libraries and CRISPR screening with accessibility profiling technologies has been developed (Roh et al., 2025; Uijttewaal et al., 2025). These constraints hinder the ability of single cell approaches to resolve regulators that acts exclusively at the final stages of erythroid maturation. Together, these limitations highlight the unique suitability of the

synchronized BEL-A system for uncovering chromatin regulators that governs terminal erythropoiesis in a biological window that remains inaccessible to current *in vivo* and single-cell CRISPR screening platforms. For these reasons, we employed the EPIKOL epigenome-targeted CRISPR library, which targets ~800 chromatin regulators with ten sgRNAs per gene and incorporates stringent internal controls. The ability of BEL-A cells to progress through late-stage erythroid maturation, including hemoglobin expression, allows this platform to quantitatively assess chromatin-dependent mechanisms driving terminal erythropoiesis and enucleation.

Across the screen, distinct depletion patterns emerged that reflect the fundamentally different biological demands of proliferating proerythroblasts versus erythroid cell differentiation. In the +dox condition, cells largely retain a proerythroblast-like identity characterized by open chromatin architecture and a high proliferative output. These pathways are heavily utilized and cannot be compensated once disrupted, leading to strong depletion of +dox-specific hits like EP400, HAT1, CHD1L, and ELP5. That means, the loss of many differentiation-specific regulators is well tolerated because the pathways they control such as differentiation-driven transcriptional activation, histone modification turnover, chromatin compaction, and nuclear quality-control are not yet engaged yet. However, once erythropoiesis progresses, cells encounter a set of unique structural and transcriptional challenges that sharply increase dependence on these pathways. Terminal differentiation requires large-scale chromatin reorganization, including the consolidation of heterochromatic domains, rewiring of lineage-defining transcriptome, extensive histone modification resetting, and the progressive reduction of nuclear volume. Genes depleted selectively during differentiation therefore represent factors whose functions become critical only once these maturation-specific processes are activated. For example, as; EP300, CTCF, SFPQ, DOT1L, KDM3B, DPF3, UHRF1, SMEK2 are only depleted during differentiation. This chromatin architecture is important to interpret the chromatin regulator dependencies during differentiation. For example, CTCF is the one of the mainly illustrated genes. According to results, CTCF is strongly depleted among differentiated erythroid cells. This finding is not surprising, as CTCF is a master architectural protein that establishes topologically associating domain (TAD) boundaries, insulates enhancers from inappropriate promoters, and organizes long-range chromatin

loops. In erythropoiesis, CTCF and cohesin (e.g., SMC3) maintain the 3D genome architecture required for large-scale chromatin condensation and proper enhancer rewiring during terminal maturation (D. Li, Wu, Zhou, Huang, & Lee, 2023). Terminal erythropoiesis demands an extreme degree of chromatin compaction and TAD reorganization that CTCF is a central regulator. Therefore, depletion during erythroid differentiation likely reflects a failure to maintain the higher-order genome structure necessary for nuclear condensation and extrusion, explaining why CTCF loss disproportionately affects the pyrenocyte stage while earlier differentiation remains permissive.

In contrast, genes depleted in the growth state were enriched for universally required processes as DNA replication, transcriptional elongation, mRNA splicing, ribosome synthesis, and basal chromatin maintenance, whose loss compromises the survival of highly proliferative progenitors regardless of erythroid identity. Genes depleted in both growth and differentiation represent global essential regulators required across all developmental stages. Together, these patterns illustrate that differentiation-specific vulnerabilities emerge precisely because differentiating erythroblasts must execute complex chromatin, transcriptional, and structural transitions in proerythroblasts have not yet initiated. For example, in this study, the most specifically depleted gene in erythroid differentiation group is EP300 and we identify that EP300 is as a chromatin regulator required for human erythropoiesis. EP300 loss resulted in a significant reduction in erythroid differentiation capacity, accompanied by impaired enucleation efficiency and persistence of progenitor-like populations, suggesting that EP300 is required for the transition through late-stage erythropoietic checkpoints (**Figure 21, 23 & 31**). EP300 loss caused a strong impairment in erythroid maturation, marked by reduced enucleation and the accumulation of progenitor-like cells, indicating to inhibit at late stages of differentiation. A previous chromatin profiling study showed that EP300 depletion resulted with a global reduction in H3K27ac at erythroid enhancers, especially those bound by GATA1 and KLF1, highlighting EP300-dependent acetylation as a key prerequisite for transcription factor recruitment and activation of terminal erythroid programs (Zheng et al., 2014). Together, our findings show that EP300 functions as a key co-activator that opens and activates chromatin regions needed for erythroid gene

expression. When EP300 is lost, erythroid genes essential for human erythropoiesis become dysregulated, and terminal differentiation cannot proceed efficiently. Unlike EP300, depletion of other histone acetyltransferases (HATs) is not identified as a significant dependency for erythroid differentiation. HATs have a diverse family of chromatin regulators that catalyze lysine acetylation on histone tails, generally promoting the transcriptional activation by increasing chromatin accessibility (Marei, 2025). Although multiple HAT families are well characterized and known to catalyze lysine acetylation such as GNAT (GCN5/PCAF), MYST (KAT6A/KAT7/TIP60), and the KAT3 family (EP300/CBP); their genomic targets and biological functions are highly context-specific (Chen, Koutelou, & Dent, 2022). Most HATs have a role on mostly promoter-proximal acetylation or participate in general transcriptional and replication-associated processes, roles that are not unique to during erythroid maturation. In contrast, erythroid differentiation depends heavily on enhancer-driven transcription (Elnitski, Miller, & Hardison, 1997; Tagoh et al., 2004; Xu et al., 2012), and among all HATs, EP300 is the predominant enhancer-activating acetyltransferase. EP300 is preferentially recruited by lineage-defining transcription factors such as GATA1, TAL1, and KLF1 and activation of erythroid-specific regulatory enhancers (Su et al., 2013). Because other HATs do not target these enhancers with comparable specificity or efficiency, their loss is either redundant or non-limiting in our system, whereas EP300 loss produces a strong differentiation inhibition.

In contrast to EP300, which functions early in differentiation, several SWI/SNF chromatin-remodeling subunits displayed selective importance at the final stages of erythropoiesis, particularly during enucleation. The most striking of these was SMARCB1, a core subunit of the cBAF and pBAF complexes. SMARCB1 depletion caused minimal impairment in early differentiation, CD49d and CD233 dynamics in SMARCB1 knock-out cells resembled those of control cells yet produced a dramatic reduction in pyrenocyte formation in BEL-A cells, and erythrocyte formation in primary HSCs (**Figure 24, 30 & 31**). This phenotype pointed in primary HSPCs, where SMARCB1 knockout reduced erythrocyte yield and promoted accumulation of basophilic, polychromatic, and orthochromatic erythroblasts. These findings indicate that SMARCB1 is not required for lineage commitment or early maturation but is crucial for

orchestrating the final transition from orthochromatic erythroblast to enucleated erythrocyte.

By deconstructing the CRISPR screen results according to SWI/SNF complex specificity, we uncovered a new picture of stage-specific dependency. Core SWI/SNF members (e.g., SMARCA4, SMARCC1/2) showed relatively uniform depletion patterns across stages, consistent with their broad roles in chromatin remodeling (**Figure 23**). However, cBAF/pBAF-shared members—including SMARCB1, SMARCE1, DPF1/2/3—were disproportionately depleted in pyrenocytes relative to orthochromatic erythroblasts. This bias was not observed for ncBAF-specific components, suggesting that cBAF/pBAF complexes are uniquely required for the late-stage maturation and enucleation, whereas ncBAF complexes play minimal roles in this context. This brings us a new question that is regarding the mechanistic basis of the stage-specific behavior exhibited by different SWI/SNF subcomplexes.

Early erythroid gene activation relies on EP300 histone acetylation, which facilitates transcriptional induction without requiring large-scale chromatin architectural changes. In contrast, enucleation requires profound restructuring: chromatin condensation, nuclear polarization, coordination with the cytoskeleton, and transcriptional silencing. These processes align more closely with the mechanistic capabilities of BAF complexes, which reposition or evict nucleosomes, restructure chromatin loops, and regulate accessibility at promoters and late-acting enhancers (**Figure 52**).

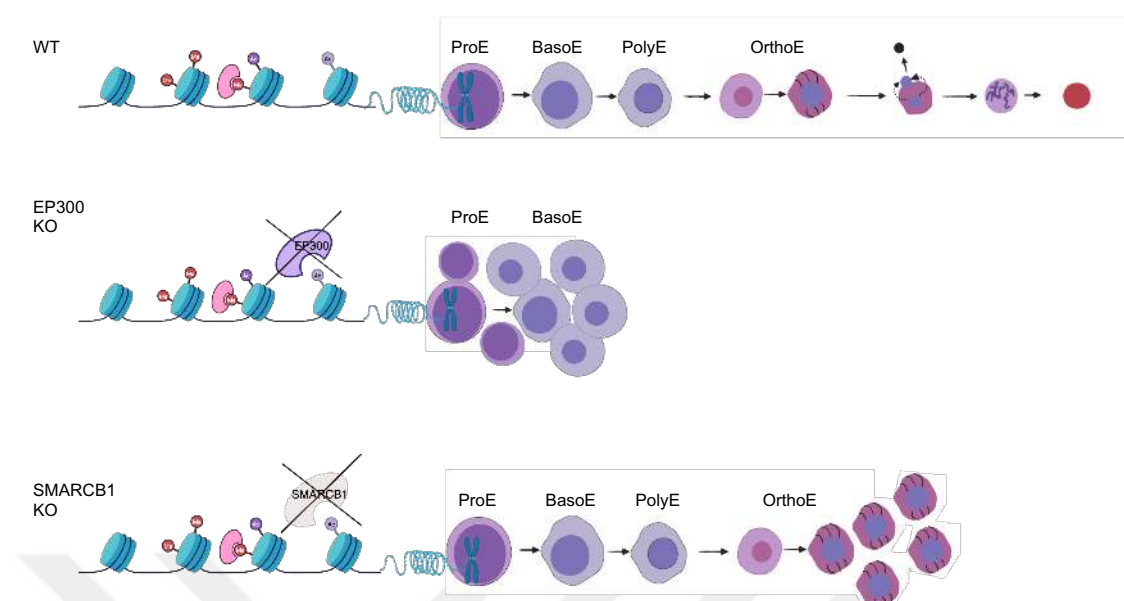


Figure 52 – Regulation of erythropoiesis by EP300 and SMARCB1
created in BioRender: licensed under CC BY 4.0.

To understand better the epigenetic mechanism of SMARCB1-dependent differentiation, we analyzed GATA1 binding using bulk D&D-Seq. While global GATA1 occupancy remained highly close between wild-type and SMARCB1 knockout cells, a subset of loci displayed significant SMARCB1-dependent gains or losses of GATA1 binding. These changes were not random because among the genes with reduced GATA1 occupancy was *SNCA*, a gene whose expression increases progressively during erythroid maturation and is highest in orthochromatic erythroblasts (Renella, Schlehe, Selkoe, Williams, & LaVoie, 2014; Scherzer et al., 2008). Genes that lose GATA1 binding in the absence of SMARCB1—including *CLDN10*, *UGCG*, *CD226*, and several lipid- and membrane-associated regulators—are enriched for pathways involved in membrane organization, vesicle dynamics, metabolic remodeling, and cytoskeletal coordination.

SMARCB1 appears to be required for maintaining or stabilizing GATA1 binding at specific late-acting enhancers, including *SNCA*. As SMARCB1 knockout cells approach the orthochromatic stage, they fail to induce *SNCA* and similar genes to appropriate levels. Because *SNCA* and related pathways have roles in vesicle trafficking, membrane dynamics, and cytoskeletal organization, their insufficient activation may underlie the

block in nuclear extrusion observed in SMARCB1-deficient cells. This shows that altered GATA1 binding by SMARCB1 in early erythroid cells may influence the later engagement of erythroid and nucleus-movement-associated genes such as SNCA, potentially contributing to their insufficient activation during terminal maturation. In this thesis, we found that the established role of GATA1 in erythropoiesis is expanded by its additional contribution to the regulation of enucleation during terminal erythroid differentiation (**Figure 53**).

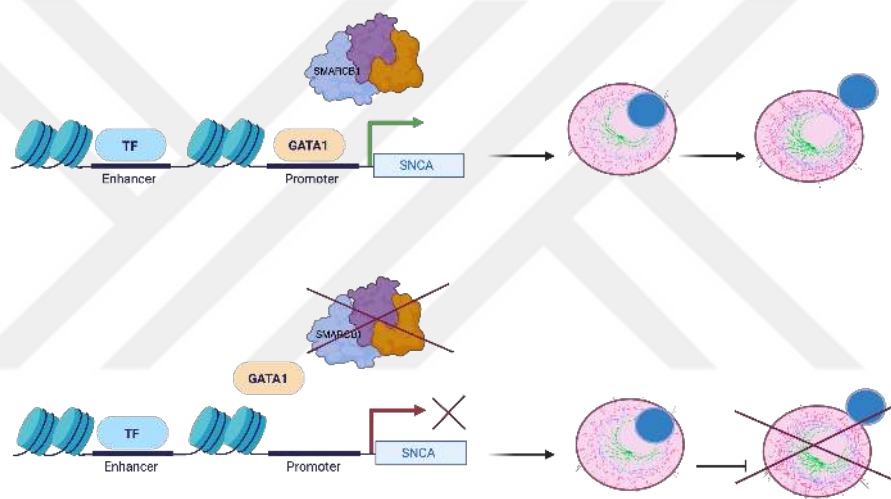


Figure 53 – Regulation of enucleation by SMARCB1
Created with Biorender.com ID: CS2960VYRA

Chapter 5

CONCLUSION

In this thesis, we systematically dissected the epigenetic mechanisms that govern human erythroid differentiation, with a particular emphasis on the final transition from orthochromatic erythroblast to enucleated erythrocyte. By integrating a synchronized BEL-A erythroid differentiation model with a pyrenocyte-focused CRISPR screening strategy, we established a uniquely powerful framework for identifying stage-specific chromatin dependencies in erythropoiesis.

Our results reveal that erythropoiesis is organized through a sequential and modular epigenetic architecture. Early differentiation relies heavily on enhancer activation mediated by the histone acetyltransferase EP300, which primes lineage-defining regulatory elements and enables the induction of erythroid transcriptional programs. Loss of EP300 disrupts acetylation, compromises erythroid gene expression, and blocks cell differentiation, demonstrating that EP300 functions as a central, non-redundant co-activator in human erythropoiesis.

In contrast, the final steps of maturation, including chromatin condensation, membrane remodeling, and nuclear extrusion, depend on the activity of specific SWI/SNF chromatin-remodeling complexes, particularly the cBAF/pBAF branches defined by SMARCB1 and DPF family members. SMARCB1 emerged as the most selective dependency of the enucleation stage, with its loss leaving early differentiation largely intact but profoundly reducing pyrenocyte formation. SMARCB1 alters GATA1 binding at a discrete subset of late-acting control genes involved in membrane organization and cytoskeletal coordination (SNCA, processes indispensable for nuclear extrusion. These findings delineate a division of labor among epigenetic regulators in which EP300 primes early

erythroid, while SMARCB1-dependent SWI/SNF remodeling ensures their functional execution during terminal maturation.

Beyond identifying EP300 and SMARCB1 as key regulators, this work provides broader conceptual insight into erythropoiesis by demonstrating that differentiation-stage-specific chromatin requirements cannot be inferred from early progenitor systems alone.

Our findings reveal epigenetic mechanism in which cBAF-mediated remodeling stabilizes GATA1 binding at late-acting enhancers, thereby enabling the structural programs required for enucleation. Moreover, by demonstrating that GATA1 mis-binding contributes directly to defective nuclear extrusion, this work assigns a new, mechanistically grounded role to GATA1 in orchestrating the physical displacement of the nucleus during terminal erythropoiesis.

Together, this thesis defines a multi-layered epigenetic model for erythroid maturation, uncovers new regulatory principles governing cell differentiation and nuclear extrusion with establish experimental approaches that will enable deeper mechanistic studies of chromatin architecture in human blood cell development. By resolving previously inaccessible stages of terminal erythropoiesis, these findings lay the groundwork for future efforts to understand congenital anemias, clonal hematopoiesis, and other disorders in which dysregulated chromatin remodeling disrupts erythrocyte maturation.

BIBLIOGRAPHY

- Arlet, J. B., Dussiot, M., Moura, I. C., Hermine, O., & Courtois, G. (2016). Novel players in beta-thalassemia dyserythropoiesis and new therapeutic strategies. *Curr Opin Hematol*, 23(3), 181-188. doi:10.1097/MOH.0000000000000231
- Atkins, B. Z., Hashmi, Z. A., Ganapathi, A. M., Harrison, J. K., Hughes, G. C., Rogers, J. G., & Milano, C. A. (2013). Surgical correction of aortic valve insufficiency after left ventricular assist device implantation. *J Thorac Cardiovasc Surg*, 146(5), 1247-1252. doi:10.1016/j.jtcvs.2013.05.019
- Bagchi, A., Billakanti, S., Shehu, V., Farrell, K. M., Cotteta, S. A., Godfrey, B. N., . . . Khandros, E. (2025). Highly discriminative globin gene activation by the non-canonical BAF chromatin remodeling complex. *Blood*. doi:10.1182/blood.2025029429
- Bai, J., Mei, P., Zhang, C., Chen, F., Li, C., Pan, Z., . . . Zheng, J. (2013). BRG1 is a prognostic marker and potential therapeutic target in human breast cancer. *PLoS One*, 8(3), e59772. doi:10.1371/journal.pone.0059772
- Bauer, D. E., Kamran, S. C., & Orkin, S. H. (2012). Reawakening fetal hemoglobin: prospects for new therapies for the beta-globin disorders. *Blood*, 120(15), 2945-2953. doi:10.1182/blood-2012-06-292078
- Blobel, G. A. (2000). CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood*, 95(3), 745-755. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/10648382>
- Bock, C., Datlinger, P., Chardon, F., Coelho, M. A., Dong, M. B., Lawson, K. A., . . . Zhuang, X. (2022). High-content CRISPR screening. *Nat Rev Methods Primers*, 2(1). doi:10.1038/s43586-022-00098-7
- Bond, D. R., Lee, H. J., & Enjeti, A. K. (2020). Unravelling the Epigenome of Myelodysplastic Syndrome: Diagnosis, Prognosis, and Response to Therapy. *Cancers (Basel)*, 12(11). doi:10.3390/cancers12113128
- Boyault, C., Sadoul, K., Pabion, M., & Khochbin, S. (2007). HDAC6, at the crossroads between cytoskeleton and cell signaling by acetylation and ubiquitination. *Oncogene*, 26(37), 5468-5476. doi:10.1038/sj.onc.1210614
- Calao, M., Burny, A., Quivy, V., Dekoninck, A., & Van Lint, C. (2008). A pervasive role of histone acetyltransferases and deacetylases in an NF-kappaB-signaling code. *Trends Biochem Sci*, 33(7), 339-349. doi:10.1016/j.tibs.2008.04.015
- Capellera-Garcia, S., Pulecio, J., Dhulipala, K., Siva, K., Rayon-Estrada, V., Singbrant, S., . . . Flygare, J. (2016). Defining the Minimal Factors Required for Erythropoiesis through Direct Lineage Conversion. *Cell Rep*, 15(11), 2550-2562. doi:10.1016/j.celrep.2016.05.027

- Caulier, A. L., & Sankaran, V. G. (2022). Molecular and cellular mechanisms that regulate human erythropoiesis. *Blood*, 139(16), 2450-2459. doi:10.1182/blood.2021011044
- Centis, F., Tabellini, L., Lucarelli, G., Buffi, O., Tonucci, P., Persini, B., . . . Schrier, S. L. (2000). The importance of erythroid expansion in determining the extent of apoptosis in erythroid precursors in patients with beta-thalassemia major. *Blood*, 96(10), 3624-3629. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/11071663>
- Chambers, C., Cermakova, K., Chan, Y. S., Kurtz, K., Wohlan, K., Lewis, A. H., . . . Hodges, H. C. (2023). SWI/SNF Blockade Disrupts PU.1-Directed Enhancer Programs in Normal Hematopoietic Cells and Acute Myeloid Leukemia. *Cancer Res*, 83(7), 983-996. doi:10.1158/0008-5472.CAN-22-2129
- Chandrasekaran, R., & Thompson, M. (2007). Polybromo-1-bromodomains bind histone H3 at specific acetyl-lysine positions. *Biochem Biophys Res Commun*, 355(3), 661-666. doi:10.1016/j.bbrc.2007.01.193
- Chaparro, C. M., & Suchdev, P. S. (2019). Anemia epidemiology, pathophysiology, and etiology in low- and middle-income countries. *Ann N Y Acad Sci*, 1450(1), 15-31. doi:10.1111/nyas.14092
- Charlop-Powers, Z., Zeng, L., Zhang, Q., & Zhou, M. M. (2010). Structural insights into selective histone H3 recognition by the human Polybromo bromodomain 2. *Cell Res*, 20(5), 529-538. doi:10.1038/cr.2010.43
- Chen, Y. C., Koutelou, E., & Dent, S. Y. R. (2022). Now open: Evolving insights to the roles of lysine acetylation in chromatin organization and function. *Mol Cell*, 82(4), 716-727. doi:10.1016/j.molcel.2021.12.004
- Chi, W. Y., Yoon, S. H., Mekerishvili, L., Ganesan, S., Potenski, C., Izzo, F., . . . Raimondi, I. (2025). Single-cell mapping of regulatory DNA:Protein interactions. *bioRxiv*. doi:10.1101/2024.12.31.630903
- Clapier, C. R., Iwasa, J., Cairns, B. R., & Peterson, C. L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol*, 18(7), 407-422. doi:10.1038/nrm.2017.26
- Clark, J., Rocques, P. J., Crew, A. J., Gill, S., Shipley, J., Chan, A. M., . . . Cooper, C. S. (1994). Identification of novel genes, SYT and SSX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. *Nat Genet*, 7(4), 502-508. doi:10.1038/ng0894-502
- Daniels, D. E., Downes, D. J., Ferrer-Vicens, I., Ferguson, D. C. J., Singleton, B. K., Wilson, M. C., . . . Frayne, J. (2020). Comparing the two leading erythroid lines BEL-A and HUDEP-2. *Haematologica*, 105(8), e389-e394. doi:10.3324/haematol.2019.229211
- Ding, J., Zheng, Y., Zhu, F., Wang, M., Fang, L., Li, H., . . . Chai, D. (2023). Adenovirus-assembled DC vaccine induces dual-targeting CTLs for tumor antigen and adenovirus to eradicate tumors. *Int Immunopharmacol*, 123, 110722. doi:10.1016/j.intimp.2023.110722
- El Hoss, S., Cochet, S., Godard, A., Yan, H., Dussiot, M., Frati, G., . . . El Nemer, W. (2021). Fetal hemoglobin rescues ineffective erythropoiesis in sickle cell disease. *Haematologica*, 106(10), 2707-2719. doi:10.3324/haematol.2020.265462

- Elnitski, L., Miller, W., & Hardison, R. (1997). Conserved E boxes function as part of the enhancer in hypersensitive site 2 of the beta-globin locus control region. Role of basic helix-loop-helix proteins. *J Biol Chem*, 272(1), 369-378. doi:10.1074/jbc.272.1.369
- Esrick, E. B., McConkey, M., Lin, K., Frisbee, A., & Ebert, B. L. (2015). Inactivation of HDAC1 or HDAC2 induces gamma globin expression without altering cell cycle or proliferation. *Am J Hematol*, 90(7), 624-628. doi:10.1002/ajh.24019
- Fang, Z., Gong, C., Hu, Y., Cui, T., Lin, M., Lin, S., & Ye, M. (2025). Acetylation of E2F1 at K125 facilitates cell apoptosis under serum stress. *Transl Oncol*, 52, 102259. doi:10.1016/j.tranon.2024.102259
- Feng, Y., Yang, Y., Ortega, M. M., Copeland, J. N., Zhang, M., Jacob, J. B., . . . Fields, P. E. (2010). Early mammalian erythropoiesis requires the Dot1L methyltransferase. *Blood*, 116(22), 4483-4491. doi:10.1182/blood-2010-03-276501
- Fiterman, O., & Raz, S. (2019). Cognitive, neural and endocrine functioning during late pregnancy: An Event-Related Potentials study. *Horm Behav*, 116, 104575. doi:10.1016/j.yhbeh.2019.104575
- Flores-Figueroa, E., Gutierrez-Espindola, G., Guerrero-Rivera, S., Pizzuto-Chavez, J., & Mayani, H. (1999). Hematopoietic progenitor cells from patients with myelodysplastic syndromes: in vitro colony growth and long-term proliferation. *Leuk Res*, 23(4), 385-394. doi:10.1016/s0145-2126(98)00176-3
- Gaspersic, J., Kristan, A., Kunej, T., Zupan, I. P., & Debeljak, N. (2021). Erythrocytosis: genes and pathways involved in disease development. *Blood Transfus*, 19(6), 518-532. doi:10.2450/2020.0197-20
- Geng, H., Liu, Q., Xue, C., David, L. L., Beer, T. M., Thomas, G. V., . . . Qian, D. Z. (2012). HIF1alpha protein stability is increased by acetylation at lysine 709. *J Biol Chem*, 287(42), 35496-35505. doi:10.1074/jbc.M112.400697
- Gifford, S. C., Derganc, J., Shevkoplyas, S. S., Yoshida, T., & Bitensky, M. W. (2006). A detailed study of time-dependent changes in human red blood cells: from reticulocyte maturation to erythrocyte senescence. *Br J Haematol*, 135(3), 395-404. doi:10.1111/j.1365-2141.2006.06279.x
- Glaros, S., Cirrincione, G. M., Muchardt, C., Kleer, C. G., Michael, C. W., & Reisman, D. (2007). The reversible epigenetic silencing of BRM: implications for clinical targeted therapy. *Oncogene*, 26(49), 7058-7066. doi:10.1038/sj.onc.1210514
- Hasserjian, R. P. (2019). Myelodysplastic Syndrome Updated. *Pathobiology*, 86(1), 7-13. doi:10.1159/000489702
- Hattangadi, S. M., Wong, P., Zhang, L., Flygare, J., & Lodish, H. F. (2011). From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. *Blood*, 118(24), 6258-6268. doi:10.1182/blood-2011-07-356006
- He, S., Wu, Z., Tian, Y., Yu, Z., Yu, J., Wang, X., . . . Xu, Y. (2020). Structure of nucleosome-bound human BAF complex. *Science*, 367(6480), 875-881. doi:10.1126/science.aaz9761
- Helming, K. C., Wang, X., Wilson, B. G., Vazquez, F., Haswell, J. R., Manchester, H. E., . . . Roberts, C. W. (2014). ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat Med*, 20(3), 251-254. doi:10.1038/nm.3480

- Heshusius, S., Grech, L., Gillemans, N., Brouwer, R. W. W., den Dekker, X. T., van, I. W. F. J., . . . Philipsen, S. (2022). Epigenomic analysis of KLF1 haploinsufficiency in primary human erythroblasts. *Sci Rep*, 12(1), 336. doi:10.1038/s41598-021-04126-6
- Hodges, C., Kirkland, J. G., & Crabtree, G. R. (2016). The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer. *Cold Spring Harb Perspect Med*, 6(8). doi:10.1101/cshperspect.a026930
- Huang, P., Zhao, Y., Zhong, J., Zhang, X., Liu, Q., Qiu, X., . . . Xu, X. (2020). Putative regulators for the continuum of erythroid differentiation revealed by single-cell transcriptome of human BM and UCB cells. *Proc Natl Acad Sci U S A*, 117(23), 12868-12876. doi:10.1073/pnas.1915085117
- Indio do Brasil, C. K., Avezum, A., Jr., Uint, L., Del Monaco, M. I., Barros, V. M., Campos, S. Y., & Sousa, A. M. (2013). Cardiovascular prevention in coronary heart disease patients: guidelines implementation in clinical practice. *Rev Bras Cir Cardiovasc*, 28(2), 238-247. doi:10.5935/1678-9741.20130034
- Ito, A., Lai, C. H., Zhao, X., Saito, S., Hamilton, M. H., Appella, E., & Yao, T. P. (2001). p300/CBP-mediated p53 acetylation is commonly induced by p53-activating agents and inhibited by MDM2. *EMBO J*, 20(6), 1331-1340. doi:10.1093/emboj/20.6.1331
- Ji, P., Yeh, V., Ramirez, T., Murata-Hori, M., & Lodish, H. F. (2010). Histone deacetylase 2 is required for chromatin condensation and subsequent enucleation of cultured mouse fetal erythroblasts. *Haematologica*, 95(12), 2013-2021. doi:10.3324/haematol.2010.029827
- Kadoch, C., & Crabtree, G. R. (2015). Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci Adv*, 1(5), e1500447. doi:10.1126/sciadv.1500447
- Keys, H. R., & Knouse, K. A. (2022). Genome-scale CRISPR screening in a single mouse liver. *Cell Genom*, 2(12). doi:10.1016/j.xgen.2022.100217
- Koehler, C. M., Merchant, S., & Schatz, G. (1999). How membrane proteins travel across the mitochondrial intermembrane space. *Trends Biochem Sci*, 24(11), 428-432. doi:10.1016/s0968-0004(99)01462-0
- Krosel, M., Gabathuler, M., Moser, L., Maciukiewicz, M., Zullig, T., Seifritz, T., . . . Klein, K. (2023). The histone acetyl transferases CBP and p300 regulate stress response pathways in synovial fibroblasts at transcriptional and functional levels. *Sci Rep*, 13(1), 17112. doi:10.1038/s41598-023-44412-z
- Kuppers, R. (2019). CD81 as target for B cell lymphomas. *J Exp Med*, 216(7), 1469-1470. doi:10.1084/jem.20190733
- Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroyama, T., Miyoshi, H., . . . Nakamura, Y. (2013). Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS One*, 8(3), e59890. doi:10.1371/journal.pone.0059890
- Langer, L. F., Ward, J. M., & Archer, T. K. (2019). Tumor suppressor SMARCB1 suppresses super-enhancers to govern hESC lineage determination. *Elife*, 8. doi:10.7554/eLife.45672
- Lara-Astiaso, D., Goni-Salaverri, A., Mendieta-Esteban, J., Narayan, N., Del Valle, C., Gross, T., . . . Huntly, B. J. P. (2023). In vivo screening characterizes chromatin

- factor functions during normal and malignant hematopoiesis. *Nat Genet*, 55(9), 1542-1554. doi:10.1038/s41588-023-01471-2
- Lee, J. C., Gimm, J. A., Lo, A. J., Koury, M. J., Krauss, S. W., Mohandas, N., & Chasis, J. A. (2004). Mechanism of protein sorting during erythroblast enucleation: role of cytoskeletal connectivity. *Blood*, 103(5), 1912-1919. doi:10.1182/blood-2003-03-0928
- Lewis, J. L., Marley, S. B., Blackett, N. M., Szydlo, R., Goldman, J. M., & Gordon, M. Y. (1998). Interleukin 3 (IL-3), but not stem cell factor (SCF) increases self-renewal by human erythroid burst-forming units (BFU-E) in vitro. *Cytokine*, 10(1), 49-54. doi:10.1006/cyto.1997.0256
- Li, D., Wu, F., Zhou, S., Huang, X. J., & Lee, H. Y. (2023). Heterochromatin rewiring and domain disruption-mediated chromatin compaction during erythropoiesis. *Nat Struct Mol Biol*, 30(4), 463-474. doi:10.1038/s41594-023-00939-3
- Li, M., Zhang, H., Wu, X., Yu, M., Yang, Q., Sun, L., . . . Chen, L. (2025). IDH1 regulates human erythropoiesis by eliciting chromatin state reprogramming. *Elife*, 13. doi:10.7554/eLife.100406
- Li, W., Xu, H., Xiao, T., Cong, L., Love, M. I., Zhang, F., . . . Liu, X. S. (2014). MAGECK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol*, 15(12), 554. doi:10.1186/s13059-014-0554-4
- Li, X., Mei, Y., Yan, B., Vitriol, E., Huang, S., Ji, P., & Qiu, Y. (2017). Histone deacetylase 6 regulates cytokinesis and erythrocyte enucleation through deacetylation of formin protein mDia2. *Haematologica*, 102(6), 984-994. doi:10.3324/haematol.2016.161513
- Madan, V., Shyamsunder, P., Dakle, P., Woon, T. W., Han, L., Cao, Z., . . . Koeffler, H. P. (2023). Dissecting the role of SWI/SNF component ARID1B in steady-state hematopoiesis. *Blood Adv*, 7(21), 6553-6566. doi:10.1182/bloodadvances.2023009946
- Malik, J., Lillis, J. A., Couch, T., Getman, M., & Steiner, L. A. (2017). The Methyltransferase Setd8 Is Essential for Erythroblast Survival and Maturation. *Cell Rep*, 21(9), 2376-2383. doi:10.1016/j.celrep.2017.11.011
- Manickavinayaham, S., Velez-Cruz, R., Biswas, A. K., Bedford, E., Klein, B. J., Kutateladze, T. G., . . . Johnson, D. G. (2019). E2F1 acetylation directs p300/CBP-mediated histone acetylation at DNA double-strand breaks to facilitate repair. *Nat Commun*, 10(1), 4951. doi:10.1038/s41467-019-12861-8
- Mantel, C. R., O'Leary, H. A., Chitteti, B. R., Huang, X., Cooper, S., Hangoc, G., . . . Broxmeyer, H. E. (2015). Enhancing Hematopoietic Stem Cell Transplantation Efficacy by Mitigating Oxygen Shock. *Cell*, 161(7), 1553-1565. doi:10.1016/j.cell.2015.04.054
- Mardinian, K., Adashek, J. J., Botta, G. P., Kato, S., & Kurzrock, R. (2021). SMARCA4: Implications of an Altered Chromatin-Remodeling Gene for Cancer Development and Therapy. *Mol Cancer Ther*, 20(12), 2341-2351. doi:10.1158/1535-7163.MCT-21-0433
- Marei, H. E. (2025). Epigenetic regulators in cancer therapy and progression. *NPJ Precis Oncol*, 9(1), 206. doi:10.1038/s41698-025-01003-7
- Martin-Rufino, J. D., Caulier, A., Lee, S., Castano, N., King, E., Joubran, S., . . . Sankaran, V. G. (2025). Transcription factor networks disproportionately enrich

- for heritability of blood cell phenotypes. *Science*, 388(6742), 52-59.
doi:10.1126/science.ads7951
- Mashtalir, N., Suzuki, H., Farrell, D. P., Sankar, A., Luo, J., Filipovski, M., . . . Kadoch, C. (2020). A Structural Model of the Endogenous Human BAF Complex Informs Disease Mechanisms. *Cell*, 183(3), 802-817 e824.
doi:10.1016/j.cell.2020.09.051
- Mei, Y., Liu, Y., & Ji, P. (2021). Understanding terminal erythropoiesis: An update on chromatin condensation, enucleation, and reticulocyte maturation. *Blood Rev*, 46, 100740. doi:10.1016/j.blre.2020.100740
- Meng, X. Y., Yau, L. F., Huang, H., Chan, W. H., Luo, P., Chen, L., . . . Wang, J. R. (2019). Improved approach for comprehensive profiling of gangliosides and sulfatides in rat brain tissues by using UHPLC-Q-TOF-MS. *Chem Phys Lipids*, 225, 104813. doi:10.1016/j.chemphyslip.2019.104813
- Mimitou, E. P., Cheng, A., Montalbano, A., Hao, S., Stoeckius, M., Legut, M., . . . Smibert, P. (2019). Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat Methods*, 16(5), 409-412.
doi:10.1038/s41592-019-0392-0
- Myers, G., Friedman, A., Yu, L., Pourmandi, N., Kerpet, C., Ito, M. A., . . . Khoriaty, R. (2025). A genome-wide screen identifies genes required for erythroid differentiation. *Nat Commun*, 16(1), 3488. doi:10.1038/s41467-025-58739-w
- Nakayama, R. T., Pulice, J. L., Valencia, A. M., McBride, M. J., McKenzie, Z. M., Gillespie, M. A., . . . Kadoch, C. (2017). SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters. *Nat Genet*, 49(11), 1613-1623. doi:10.1038/ng.3958
- Nicosia, L., Spencer, G. J., Brooks, N., Amaral, F. M. R., Basma, N. J., Chadwick, J. A., . . . Somervaille, T. C. P. (2023). Therapeutic targeting of EP300/CBP by bromodomain inhibition in hematologic malignancies. *Cancer Cell*, 41(12), 2136-2153 e2113. doi:10.1016/j.ccell.2023.11.001
- Nina, V. J., Matias, M. M., Brito, D. J., Figueiredo Neto, J. A., Coutinho, L. B., Rodrigues, R. F., . . . Gaspar, S. F. (2013). Acute kidney injury after coronary artery bypass grafting: assessment using RIFLE and AKIN criteria. *Rev Bras Cir Cardiovasc*, 28(2), 231-237. doi:10.5935/1678-9741.20130033
- Nocka, K., Majumder, S., Chabot, B., Ray, P., Cervone, M., Bernstein, A., & Besmer, P. (1989). Expression of c-kit gene products in known cellular targets of W mutations in normal and W mutant mice--evidence for an impaired c-kit kinase in mutant mice. *Genes Dev*, 3(6), 816-826. doi:10.1101/gad.3.6.816
- Palis, J. (2014). Primitive and definitive erythropoiesis in mammals. *Front Physiol*, 5, 3. doi:10.3389/fphys.2014.00003
- Panier, S., & Boulton, S. J. (2014). Double-strand break repair: 53BP1 comes into focus. *Nat Rev Mol Cell Biol*, 15(1), 7-18. doi:10.1038/nrm3719
- Pauling, L., Itano, H. A., & et al. (1949). Sick cell anemia a molecular disease. *Science*, 110(2865), 543-548. doi:10.1126/science.110.2865.543
- Pretini, V., Koenen, M. H., Kaestner, L., Fens, M., Schiffelers, R. M., Bartels, M., & Van Wijk, R. (2019). Red Blood Cells: Chasing Interactions. *Front Physiol*, 10, 945. doi:10.3389/fphys.2019.00945
- Priam, P., Krasteva, V., Rousseau, P., D'Angelo, G., Gaboury, L., Sauvageau, G., & Lessard, J. A. (2017). SMARCD2 subunit of SWI/SNF chromatin-remodeling

- complexes mediates granulopoiesis through a CEBPvarepsilon dependent mechanism. *Nat Genet*, 49(5), 753-764. doi:10.1038/ng.3812
- Rago, F., Rodrigues, L. U., Bonney, M., Sprouffs, K., Kurth, E., Elliott, G., . . . Jagani, Z. (2022). Exquisite Sensitivity to Dual BRG1/BRM ATPase Inhibitors Reveals Broad SWI/SNF Dependencies in Acute Myeloid Leukemia. *Mol Cancer Res*, 20(3), 361-372. doi:10.1158/1541-7786.MCR-21-0390
- Reisman, D. N., Sciarrotta, J., Wang, W., Funkhouser, W. K., & Weissman, B. E. (2003). Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis. *Cancer Res*, 63(3), 560-566. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/12566296>
- Renella, R., Schlehe, J. S., Selkoe, D. J., Williams, D. A., & LaVoie, M. J. (2014). Genetic deletion of the GATA1-regulated protein alpha-synuclein reduces oxidative stress and nitric oxide synthase levels in mature erythrocytes. *Am J Hematol*, 89(10), 974-977. doi:10.1002/ajh.23796
- Roh, H., Shen, S. P., Hu, Y., Kwok, H. S., Siegenfeld, A. P., Lee, C., . . . Liao, B. B. (2025). Coupling CRISPR scanning with targeted chromatin accessibility profiling using a double-stranded DNA deaminase. *Nat Methods*, 22(10), 2083-2093. doi:10.1038/s41592-025-02811-2
- Scherzer, C. R., Grass, J. A., Liao, Z., Pepivani, I., Zheng, B., Eklund, A. C., . . . Schlossmacher, M. G. (2008). GATA transcription factors directly regulate the Parkinson's disease-linked gene alpha-synuclein. *Proc Natl Acad Sci U S A*, 105(31), 10907-10912. doi:10.1073/pnas.0802437105
- Schick, S., Rendeiro, A. F., Runggatscher, K., Ringler, A., Boidol, B., Hinkel, M., . . . Kubicek, S. (2019). Systematic characterization of BAF mutations provides insights into intracomplex synthetic lethality in human cancers. *Nat Genet*, 51(9), 1399-1410. doi:10.1038/s41588-019-0477-9
- Schroit, A. J., Madsen, J. W., & Tanaka, Y. (1985). In vivo recognition and clearance of red blood cells containing phosphatidylserine in their plasma membranes. *J Biol Chem*, 260(8), 5131-5138. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/3988747>
- Shearstone, J. R., Golonzhka, O., Chonkar, A., Tamang, D., van Duzer, J. H., Jones, S. S., & Jarpe, M. B. (2016). Chemical Inhibition of Histone Deacetylases 1 and 2 Induces Fetal Hemoglobin through Activation of GATA2. *PLoS One*, 11(4), e0153767. doi:10.1371/journal.pone.0153767
- Shi, J., Whyte, W. A., Zepeda-Mendoza, C. J., Milazzo, J. P., Shen, C., Roe, J. S., . . . Vakoc, C. R. (2013). Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes Dev*, 27(24), 2648-2662. doi:10.1101/gad.232710.113
- Sima, X., He, J., Peng, J., Xu, Y., Zhang, F., & Deng, L. (2019). The genetic alteration spectrum of the SWI/SNF complex: The oncogenic roles of BRD9 and ACTL6A. *PLoS One*, 14(9), e0222305. doi:10.1371/journal.pone.0222305
- Sperling, A. S., Gibson, C. J., & Ebert, B. L. (2017). The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. *Nat Rev Cancer*, 17(1), 5-19. doi:10.1038/nrc.2016.112
- Strang, K. H., Golde, T. E., & Giasson, B. I. (2019). MAPT mutations, tauopathy, and mechanisms of neurodegeneration. *Lab Invest*, 99(7), 912-928. doi:10.1038/s41374-019-0197-x

- Su, M. Y., Steiner, L. A., Bogardus, H., Mishra, T., Schulz, V. P., Hardison, R. C., & Gallagher, P. G. (2013). Identification of biologically relevant enhancers in human erythroid cells. *J Biol Chem*, 288(12), 8433-8444. doi:10.1074/jbc.M112.413260
- Subudhi, P., Subudhi, B. N. R., Patro, S., & Sitaram, S. (2022). Commentary: Endoexpression technique of nucleus delivery in manual small-incision cataract surgery. *Indian J Ophthalmol*, 70(11), 4064-4065. doi:10.4103/ijo.IJO_2237_22
- Sui, X., Krantz, S. B., & Zhao, Z. J. (2000). Stem cell factor and erythropoietin inhibit apoptosis of human erythroid progenitor cells through different signalling pathways. *Br J Haematol*, 110(1), 63-70. doi:10.1046/j.1365-2141.2000.02145.x
- Tagoh, H., Melnik, S., Lefevre, P., Chong, S., Riggs, A. D., & Bonifer, C. (2004). Dynamic reorganization of chromatin structure and selective DNA demethylation prior to stable enhancer complex formation during differentiation of primary hematopoietic cells in vitro. *Blood*, 103(8), 2950-2955. doi:10.1182/blood-2003-09-3323
- Uijttewaal, E. C. H., Lee, J., Sell, A. C., Botay, N., Vainorius, G., Novatchkova, M., . . . Elling, U. (2025). CRISPR-StAR enables high-resolution genetic screening in complex in vivo models. *Nat Biotechnol*, 43(11), 1848-1860. doi:10.1038/s41587-024-02512-9
- Verdone, L., Caserta, M., & Di Mauro, E. (2005). Role of histone acetylation in the control of gene expression. *Biochem Cell Biol*, 83(3), 344-353. doi:10.1139/o05-041
- Waarts, M. R., Mowla, S., Boileau, M., Martinez Benitez, A. R., Sango, J., Bagish, M., . . . Levine, R. L. (2024). CRISPR Dependency Screens in Primary Hematopoietic Stem Cells Identify KDM3B as a Genotype-specific Vulnerability in IDH2- and TET2-mutant Cells. *Cancer Discov*, 14(10), 1860-1878. doi:10.1158/2159-8290.CD-23-1092
- Wang, C., Liao, S., Wang, Y., Hu, X., & Xu, J. (2022). Computational Identification of Guillain-Barre Syndrome-Related Genes by an mRNA Gene Expression Profile and a Protein-Protein Interaction Network. *Front Mol Neurosci*, 15, 850209. doi:10.3389/fnmol.2022.850209
- Wang, S., Xiang, T., Yu, L., Wen, J., Liu, F., Yang, D., . . . Hu, L. (2021). Novel Molecular Subtypes and Related Score Based on Histone Acetylation Modification in Renal Clear Cell Carcinoma. *Front Cell Dev Biol*, 9, 668810. doi:10.3389/fcell.2021.668810
- Wang, X., Lee, R. S., Alver, B. H., Haswell, J. R., Wang, S., Mieczkowski, J., . . . Roberts, C. W. (2017). SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat Genet*, 49(2), 289-295. doi:10.1038/ng.3746
- Wang, Y., Li, W., Schulz, V. P., Zhao, H., Qu, X., Qi, Q., . . . An, X. (2021). Impairment of human terminal erythroid differentiation by histone deacetylase 5 deficiency. *Blood*, 138(17), 1615-1627. doi:10.1182/blood.2020007401
- Wang, Y., Xie, S., Armendariz, D., & Hon, G. C. (2022). Computational identification of clonal cells in single-cell CRISPR screens. *BMC Genomics*, 23(1), 135. doi:10.1186/s12864-022-08359-1
- Wang, Z., Wang, P., Li, Y., Peng, H., Zhu, Y., Mohandas, N., & Liu, J. (2021). Interplay between cofactors and transcription factors in hematopoiesis and hematological

- malignancies. *Signal Transduct Target Ther*, 6(1), 24. doi:10.1038/s41392-020-00422-1
- Wanior, M., Kramer, A., Knapp, S., & Joerger, A. C. (2021). Exploiting vulnerabilities of SWI/SNF chromatin remodelling complexes for cancer therapy. *Oncogene*, 40(21), 3637-3654. doi:10.1038/s41388-021-01781-x
- Xu, J., Shao, Z., Glass, K., Bauer, D. E., Pinello, L., Van Handel, B., . . . Orkin, S. H. (2012). Combinatorial assembly of developmental stage-specific enhancers controls gene expression programs during human erythropoiesis. *Dev Cell*, 23(4), 796-811. doi:10.1016/j.devcel.2012.09.003
- Yedier-Bayram, O., Gokbayrak, B., Kayabolen, A., Aksu, A. C., Cavga, A. D., Cingoz, A., . . . Bagci-Onder, T. (2022). EPIKOL, a chromatin-focused CRISPR/Cas9-based screening platform, to identify cancer-specific epigenetic vulnerabilities. *Cell Death Dis*, 13(8), 710. doi:10.1038/s41419-022-05146-4
- Zhao, B., Yang, J., & Ji, P. (2016). Chromatin condensation during terminal erythropoiesis. *Nucleus*, 7(5), 425-429. doi:10.1080/19491034.2016.1226717
- Zheng, W. W., Dong, X. M., Yin, R. H., Xu, F. F., Ning, H. M., Zhang, M. J., . . . Yang, X. M. (2014). EDAG positively regulates erythroid differentiation and modifies GATA1 acetylation through recruiting p300. *Stem Cells*, 32(8), 2278-2289. doi:10.1002/stem.1723