

Probing long-range genomic interactions and gene expression dynamics that instruct T cell fate

Thesis submitted for the award of the degree of

DOCTOR OF PHILOSOPHY

By

Arpita Prusty

16LAPH01



**Department of Animal Biology
School of Life Sciences
University of Hyderabad
Hyderabad-500046, India**

April 2023



UNIVERSITY OF HYDERABAD

SCHOOL OF LIFE SCIENCES

CERTIFICATE

This is to certify that the thesis entitled “**Probing long-range genomic interactions and gene expression dynamics that instruct T cell fate**” submitted by **Ms. Arpita Prusty**, bearing registration number **16LAPH01**, in partial fulfilment of the requirements for awarding a Doctor of Philosophy in the School of Life Sciences, is a bona fide work carried out by her under my supervision and guidance.

This thesis does not have any plagiarism and has not been previously submitted, in part or in full, to this or any other university or institution for the award of any degree or diploma.

The results of this thesis have been presented in the following conferences:

1. Meeting on the Molecular Mechanisms that Underpin gamma/delta T cell Development
New York, 12-13 November 2021.
2. 4D Nucleome Annual Meeting, Washington (DC), 3-6 December 2019.

Further, the student has passed the following courses towards fulfilment of the coursework requirement for the Ph.D.

<u>Course code</u>	<u>Name</u>	<u>Credits</u>	<u>Pass/Fail</u>
AB-801	Analytical Techniques	4	Pass
AB-802	Research Ethics, Data Analysis and Biostatistics	3	Pass
AB-803	Lab work & Seminar	5	Pass

Prof. Jagan Pongubala

Ph.D. Supervisor
Prof. Jagan Pongubala
Department of Animal Biology
School of Life Sciences
University of Hyderabad
Gachibowli, Hyderabad-500 046.

Head

Department of Animal Biology
अध्यक्ष / HEAD
जंतु जैविकी विभाग
Department of Animal Biology

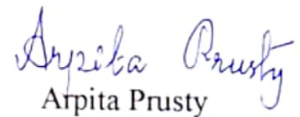
Dean

School of Life Sciences
संकाय अध्यक्ष / Dean
जीव विज्ञान संकाय / School of Life Sciences
हैदराबाद विश्वविद्यालय / University of Hyderabad
हैदराबाद / Hyderabad-500 046

DECLARATION

I hereby declare that the results of the study presented in the thesis entitled **“Probing long-range genomic interactions and gene expression dynamics that instruct T cell fate”** have been carried out by me under the supervision of Prof. Jagan Pongubala at the Department of Animal Biology, School of Life Sciences. The work presented in this thesis is a bona fide research and has not been submitted for any degree or diploma in any other University or Institute.

Date : 10-04-2023


Arpita Prusty

(16LAPH01)



Prof. Jagan Pongubala

Prof. Jagan Pongubala
(Supervisor)
Department of Animal Biology
School of Life Sciences
University of Hyderabad
Gachibowli, Hyderabad-500 046.

ACKNOWLEDGEMENTS

I would like to express a deep sense of gratitude to my supervisor Prof. Jagan Pongubala, for his guidance, encouragement and support, throughout my PhD. I am very grateful to him for his insightful advices, positive and encouraging outlook and for training me to grow and emerge as a thoughtful researcher. I am grateful to him for giving me the opportunity to work in his lab; for all the help he has provided to me during these years and for his supervision to enable me to complete my PhD successfully.

A word of special gratitude to Prof. Kees Murre from UCSD for his keen interest, insightful knowledge and for his unwavering encouragement, motivation, and eternal willingness to help me overcome all the hurdles during my scientific adventure at UCSD and made my stay everyday an enjoyable day.

I sincerely thank Dr. David Wiest, Dr. Juan Carlos Zúñiga-Pflücker and Dr. Erin Adams for their guidance, support, valuable advice and provided resources during my work.

I convey sincere thanks to my doctoral committee members Prof. B. Senthilkumaran and Dr. Nooruddin Khan for their encouragement and feedbacks.

I am thankful to Prof. Srinivasulu Kurukuti, Head, Department of Animal Biology, and the former Heads, Prof. Anita Jagota, Prof. Jagan Pongubala and Prof. B Senthilkumaran for their support and in allowing me to use departmental facilities. I extend thanks to Prof. N. Siva kumar, Dean, School of Life Sciences and the former Deans, to Prof. S. Dayananda, Prof. P. Reddanna and Prof. K.V. Ramaiah for permitting me to use all necessary School resources for my research work. I thank faculty members of School of Life Sciences for their timely and useful suggestions.

I would like to thank my former lab members, Dr. Ravi Boya, Dr. Sameena Nikhat and Dr. Anurupa Devi for scientific discussions and helpful suggestions. I would like to thank all the current members of my lab: Ms. Anubhooti, Ms. Priyanka Kriti Narayan, Mr. Ashok Kumar and Ms. Nidhi Singh, Ms. Shravya Kurella and Ms. Apoorva Soni for their help and support in the lab and making a pleasant working atmosphere.

Special thanks to Dr. Indumathi Patta, Dr. Mauricio, Dr. Sergio, Dr. Joy, Dr. Fernando, Dr. Yina, Dr. Hanbin, Megan and the entire research team of Prof. Kees Murre, who constantly encouraged me to chase my research goals and helped me stay on track.

I also remain indebted to Dr. Zachary Warburg, Research Scholar, UCSD for all his help and suggestions in our joint project. Objective-1 of my thesis is done in collaboration with Zach. I'm thankful to him for sharing a part of his thesis work with me.

I would like to thank all the previous and current non-teaching staff of Dept. of Animal Biology: Mrs. Vijaya Lakshmi, Mr. Nikhil, Mr. Jagan, Mr. Gopi, Mr. Rangaswamy and Mr. Sreenu for their help in office and administrative work.

I sincerely thank CSIR-UGC, DBT-IUSSTF and NIH for funding my doctoral studies and DST and DBT for providing financial assistance to the lab and the infrastructural facility to the school. I would like to thank core facilities at UCSD for their help.

I would like to thank all my dear friends Aurobinda, Amit, Debakanta, Deepak, Nimi, Poorna, Ravish, Raghav, Sachin, View for being with me through good and bad times and for endorsing me at all times in pursuit of my goals. I would also like to thank all the Research Scholars in School of Life Sciences for their cooperation and help.

My family members have been my pillars of strength with their persistent faith, patience, and enormous encouragement all through my life. I am obliged to my father Dr. Kishore Kumar Prusty, my mother Mrs. Sasmita Prusty for their unequivocal love, advice, prayers and their earnest efforts in teaching and inculcating virtues and good morals. I cherish and treasure the company and affection of my little sister Dr. Aveepsa Prusty who has been a friend forever. I am indebted to my husband, Dr. Ranay Mohan Yadav, for being my life partner. I owe him an enormous debt of gratitude for being my constant companion and source of encouragement and strength. I shall ever remain indebted to each one of them for their immense support in this entire journey.

Above all, I am thankful to the God Almighty for every opportunity, for every learning experience and for showing me the right perspective in all situations.

-Arpita Prusty

Dedicated

to

my parents, sister, and the Almighty God

TABLE OF CONTENTS

List of Abbreviations	I-II
List of Figures	III-IV
List of Tables	V
Abstract	VI
Chapter-1: Introduction	01-21
1.1 General Introduction to hematopoietic system	
1.2 Classical hematopoietic hierarchy in the adult bone marrow	
1.3 Early B cell development	
1.4 Developmental progression of early T-lineage cells	
1.5 Transcriptional control of T cell development	
1.6 E and Id proteins in T cell development	
1.7 $\alpha\beta$ and $\gamma\delta$ T cell fate choice	
1.8 Regulation of genome 3D structure in immune cell development	
1.9 Non-coding RNAs in early T cell development	
1.9.1 Role of lncRNA-ThymoD in T cell development	
1.10 Dynamics of transcription	
1.11 What triggers a transcription burst?	
1.11.1 Proximal Regulation	
1.11.2 Distal Regulation	
1.12 Visualization of transcription dynamics	
1.12.1 Fixed-cell techniques for RNA detection	
1.12.2 Live cell RNA probes	
1.13 Overview of MS2-MCP system	
Objectives of the study	22-23

2.1. Materials 25-34

- 2.1.1 List of chemicals, enzymes and reagents
- 2.1.2 List of kits and disposables
- 2.1.3 List of cell culture media and reagents used
- 2.1.4 List of antibodies used
- 2.1.5 List of important technical equipment used
- 2.1.6 Experimental models: cell lines/ organisms/strains used
- 2.1.7 Recombinant DNAs used
- 2.1.8 List of analytical software used
- 2.1.9 CRISPR injection mix used for generation of CR4 KO mice
- 2.1.10 CRISPR injection mix used for generation of Id3-MS2 knock-in mice
- 2.1.11 Guide RNAs used for CTCF binding site deletion
- 2.1.12 List of primers used for genotyping CTCF binding site deletions
- 2.1.13 Guide RNAs used to generate CR4 knockout mice
- 2.1.14 List of qRT-PCR primers used for quantifying mRNA transcripts
- 2.1.15 List of primers used for making vector constructs
- 2.1.16 List of primers used for genotyping Id3-MS2 transgenic cell lines and mice

2.2 Methods: Objective-I 34-38

- 2.2.1 Designing and cloning gRNA constructs to target CTCF binding sites
- 2.2.2 Cell line cultures
- 2.2.3 Deleting CR4 CTCF binding regions in the Scid.adh cell line
- 2.2.4 FACS sorting
- 2.2.5 Genotyping of CR4 and B4 KO cell colonies
- 2.2.6 Real-time PCR
- 2.2.7 Generation of CR4 KO transgenic mice
- 2.2.8 Genotyping and Sequencing of CR4 knockout Mice
- 2.2.9 Co-culture of bone marrow progenitors and stromal cells
- 2.2.10 Cut&Run
- 2.2.11 Analysis for Cut&Run

2.2.12 RNA-seq	
2.2.13 Analysis for RNA-seq	
2.2.13 Flow cytometry	
2.2.14 Statistical analysis	
2.3 Methods: Objective-II	39-42
2.3.1 Vectors and constructs cloning	
2.3.2 Designing and cloning gRNA constructs to insert MS2 loops into Id3 3'UTR	
2.3.3 Generating Id3-MS2 Pro-B cell line	
2.3.4 Id3 MS2 clone screening strategy	
2.3.5 Generation of Id3-MS2 mice	
2.3.6 Generation of Rag1 ^{-/-} : Id3-MS2 mice	
2.3.7 Mouse Pro-B cell culture	
2.3.8 Mouse T cell culture	
2.3.9 Flow cytometry	
2.3.10 Viral transduction	
2.3.11 Generation of MCP-GFP expression constructs for signal to noise ratio optimization	
2.3.12 Anti CD3ε/ ERK inhibitor treatment	
2.3.13 Generation of TCR plasmids	
2.3.14 Real-time PCR	
2.3.15 MHC tetramer treatment	
2.3.16 Live cell imaging	
2.3.17 Image analysis	
Chapter-3: Results	43-67
Objective-I	43-50
3.1.1 Single CTCF binding sites specifically influence Bcl11b expression over vast genomic distances	
3.1.2 Generation of the CR4 deletion mutant mouse line	
3.1.3 Disruption of CR4 greatly abolishes its CTCF binding	
3.1.4 CR4 CTCF binding is essential for T-lineage gene expression program	

3.1.5	Early T cell developmental program is mildly affected in CR4 knockout mice	
3.1.6	CR4KO bone marrow progenitors are impaired to differentiate into early T cells in vitro	
Objective-II		51-67
3.2.1	Generation of Id3-MS2 transgenic pro-B cell line	
3.2.2	Generation of Id3-MS2 transgenic mice	
3.2.3	Establishment of In-vitro differentiation of T cell progenitors	
3.2.4	Optimization of signal-to-noise ratio for MCP-GFP mediated fluorescence	
3.2.5	Id3 bursting frequency is rapidly activated by pre-TCR signaling	
3.2.6	The ERK-MAPK signaling pathway regulates Id3 bursting	
3.2.7	TCR $\gamma\delta$ -mediated signaling induces higher Id3 bursting frequencies as compared to signaling induced by the TCR β chain	
3.2.8	Id3 expression levels correlate with T cell antigen receptor affinity	
Chapter-4:	Discussion	68-72
Chapter-5:	Summary	73
Chapter-6:	Bibliography	74-86
Chapter-7:	Appendix	87
8.1.	Publications	
8.2.	Plagiarism report	

LIST OF ABBREVIATIONS

HSC	Haematopoietic Stem cells
MPP	Multipotent Progenitors
LMPP	Lymphoid Primed Multipotent Progenitors
CLP	Common Lymphoid Progenitors
ChIP	Chromatin Immuno-Precipitation
TADs	Topologically associated domains
3D	Three Dimensional
FISH	Fluorescence In Situ Hybridization
TCR	T cell receptor
RAG	Recombination Activating Gene
CD19	Cluster of Differentiation 19
Mb	Megabase
mg	Milligram
min	minutes
ml	millilitre
PBS	Phosphate Buffered Saline
rpm	Revolutions per minute
RNA pol II	RNA polymerase II
RNase A	Ribonuclease A
SCF	Stem Cell Factor
ug	Microgram
uM	Micromolar
ul	Microlitre
°C	Degree Celsius
BSA	Bovine Serum Albumin
CTCF	CCCTC-binding factor
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
IL-7	Interleukin-7
Flt3L	FMS-like tyrosine kinase 3 ligand
GFP	Green Fluorescent Protein
EDTA	Ethylenediaminetetraacetic acid

dATP	Deoxyadenosine triphosphate
dTTP	Deoxythymidine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
cDNA	Complementary DNA
FBS	Fetal Bovine Serum
hr	Hour
mRNA	Messenger RNA
nM	Nanometer
NK	Natural Killer Cell
UTR	Untranslated Region
qRT-PCR	Quantitative Reverse Transcription- Polymerase Chain Reaction
MHC	Major Histocompatibility Complex
ThymoD	Thymocyte Differentiation Factor
MCP	MS2-coat protein
Id3	Inhibitor of differentiation 3
ERK	Extracellular signal-regulated kinase
KO	Knock-out
gRNA	Guide RNA

LIST OF FIGURES

- Figure 1.1 The journey from fetal to adult murine haematopoiesis
- Figure 1.2 The classical haematopoietic hierarchy in the adult bone marrow
- Figure 1.3 Early B-cell development in the bone marrow
- Figure 1.4 Developmental progression of early T-lineage cells
- Figure 1.5 Transcriptional control of T cell development and gene regulatory network.
- Figure 1.6 E and Id proteins in T cell development
- Figure 1.7 $\alpha\beta$ and $\gamma\delta$ T cell fate choice
- Figure 1.8 Regulation of genome 3D structure in immune cell development
- Figure 1.9 Non-coding RNAs in early T cell development
- Figure 1.10 Dynamics of transcription
- Figure 1.12 Visualization of transcription dynamics: Methods of fluorescently labelling RNA in fixed or live samples
- Figure 1.13 Overview of MS2-MCP system
- Figure 3.1.1 Single CTCF binding sites specifically influence Bcl11b expression over vast genomic distances
- Figure 3.1.2 Generation of the CR4 deletion mutant mouse line
- Figure 3.1.3 Disruption of CR4 greatly abolishes its CTCF binding
- Figure 3.1.4 CR4 CTCF binding is essential for T-lineage gene expression program
- Figure 3.1.5 Early T cell developmental program is mildly affected in CR4 knockout mice
- Figure 3.1.6 CR4KO bone marrow progenitors are impaired to differentiate into early T cells in vitro
- Figure 3.2.1 Generation of Id3-MS2 transgenic pro-B cell line
- Figure 3.2.2 Generation of Id3-MS2 transgenic mice
- Figure 3.2.3 Establishment of in-vitro differentiation of T cell progenitors

Figure 3.2.4 Optimization of signal-to-noise ratio for MCP-GFP mediated fluorescence

Figure 3.2.5 Id3 bursting frequency is rapidly activated by pre-TCR signaling

Figure 3.2.6 The ERK-MAPK signaling pathway regulates Id3 bursting

Figure 3.2.7 TCR $\gamma\delta$ -mediated signaling induces higher Id3 bursting frequencies as compared to signaling induced by the TCR β chain

Figure 3.2.8 Id3 expression levels correlate with T cell antigen receptor affinity

LIST OF TABLES

- 2.1.1 List of chemicals, enzymes and reagents
- 2.1.2 List of kits and disposables
- 2.1.3 List of cell culture media and reagents used
- 2.1.4 List of antibodies used
- 2.1.5 List of important technical equipment used
- 2.1.6 Experimental models: cell lines/ organisms/strains used
- 2.1.7 Recombinant DNAs used
- 2.1.8 List of analytical software used
- 2.1.9 CRISPR injection mix used for generation of CR4 KO mice
- 2.1.10 CRISPR injection mix used for generation of Id3-MS2 knock-in mice
- 2.1.11 Guide RNAs used for CTCF binding site deletion
- 2.1.12 List of primers used for genotyping CTCF binding site deletions
- 2.1.13 Guide RNAs used to generate CR4 knockout mice
- 2.1.14 List of qRT-PCR primers used for quantifying mRNA transcripts
- 2.1.15 List of primers used for making vector constructs
- 2.1.16 List of primers used for genotyping Id3-MS2 transgenic cell lines and mice

ABSTRACT

Cell growth, differentiation, and response to external stimulation all necessitate precise and coordinated control of gene expression. It is becoming clear that gene expression programs rely on cis-regulatory interactions mediated through the binding of combinatorial transcription factors. Here, we have focused on studying molecular mechanisms that underlie Bcl11b promoter-enhancer interactions and the transcription dynamics of Id3 that are essential for the development of diverse T cell lineages, by establishing genetically engineered animal models. We have pursued these, using high-throughput molecular approaches and single cell live imaging technology. These studies have revealed a novel CTCF binding region necessary for Bcl11b activation. Deletion of this CTCF region impeded Bcl11b expression, indicating a change in chromatin conformation that is incompatible with its expression. Consistent with these observations, multipotent progenitors (MPPs) isolated from CTCF-mutant animals failed to differentiate into T cells, as evidenced by a block at the DN2 cell stage. Furthermore, genome-wide expression analysis has revealed that mutation of the CTCF binding region within the Bcl11b locus perturbed the expression pattern of a large number of genes, indicating a key role for CTCF binding within the Bcl11b locus. In parallel, single cell live imaging studies have shown that Id3 transcription is found to be dynamically switching between on and off states. Furthermore, the threshold of Id3 transcription is positively correlated with the externally regulated signals that drive diverse T cell lineages ($\gamma\delta$ and $\alpha\beta$ TCR signaling). It is also remarkable that upon antigenic stimulation double negative T cells exhibit increased Id3 burst transcription as determined by high intensity bursting activity. This was antagonized by inhibiting ERK signaling. The transcription variations could be explained by changes in the proximity interaction between cis-regulatory elements and/or binding of transcription factors, suggesting that they constitute a transcriptional regulatory event that controls nascent mRNA output. Collectively, these studies emphasize the importance of CTCF binding for chromatin conformation, which controls the enhancer-promoter communication required for Bcl11b expression, and the transcription dynamics of Id3 transcription, both of which are crucial for the development of diverse T lineages.

Chapter 1

INTRODUCTION

1.1. General introduction to the hematopoietic system

Hematopoiesis is the process through which all cellular components of the blood and immune system are formed. It starts with embryonic development and continues into adulthood to develop and maintain blood cells. In mammals, the first primitive blood cells originate in the yolk sac (known as 'primitive hematopoiesis'), which primarily includes nucleated erythrocytes that produce RBCs to help oxygenate tissues while the embryo undergoes multiple cell divisions (Orkin and Zon, 2008). These erythroid progenitors lack self-renewal capacity; hence, this is a transient wave. This is followed by a brief wave known as "definitive hematopoiesis," which begins in the blood islands and generates lymphoid and erythroid-myeloid progenitors (EMPs), that temporarily seed the fetal liver (Bertrand et al., 2007; Böiers et al., 2013; McGrath et al., 2015).

The stable adult hematopoietic system eventually develops during "definitive hematopoiesis," which takes place later in the development and at specific times in different animals. The first detectable hemopoietic stem cells (HSCs) develop in the evolutionarily conserved aorto-gonado-mesonephros (AGM) area (Ivanovs et al., 2011). Following that, hematopoiesis switches to the fetal liver, then to the bone marrow, where HSCs are found in adulthood (**Fig 1.1**). Hematopoietic stem cells (HSCs) are the source of definitive hematopoiesis. They help in the continuous generation and maintenance of many types of blood and immune cells throughout an organism's life. They have a remarkable ability for self-renewal and are pluripotent, which means they can generate all blood cell types. Thus, HSCs are programmed to enable efficient production of blood cell components that have differential functions, particularly innate and adaptive immune cells, which safeguard against cancer and infectious diseases.

HSCs differ in their ability to self-renew, and they have been divided into distinct subsets according to the number of symmetric divisions which can be achieved during their lifetime. Long-Term HSCs (LT-HSCs) are the HSCs that show regeneration capacity for more than 16 weeks in a primary transplantation experiment and at least once during subsequent transplantation in humans and mice. HSCs that produce all differentiated cell types but exhibit transient primary engraftment (also known as secondary engraftment) are defined as intermediate (IT-) HSCs or Short-Term (ST-) HSCs, based on the efficiency and length of the graft formed (Kent et al., 2009). The HSC niche is a complex ecosystem that supports and preserves HSC function by facilitating the long-term survival of the HSC pool. The mechanisms that influence the fate of specific HSC

subsets inside the niche are currently unclear. Previous studies revealed that diverse HSC subsets may prefer different niche habitats, and HSCs are known to lose their self-renewal capacity when isolated from their *in vivo* niches. Extrinsic cues from extracellular matrix-associated and membrane-bound ligands present within the niche play a critical role in HSC function (Laurenti and Göttgens, 2018).

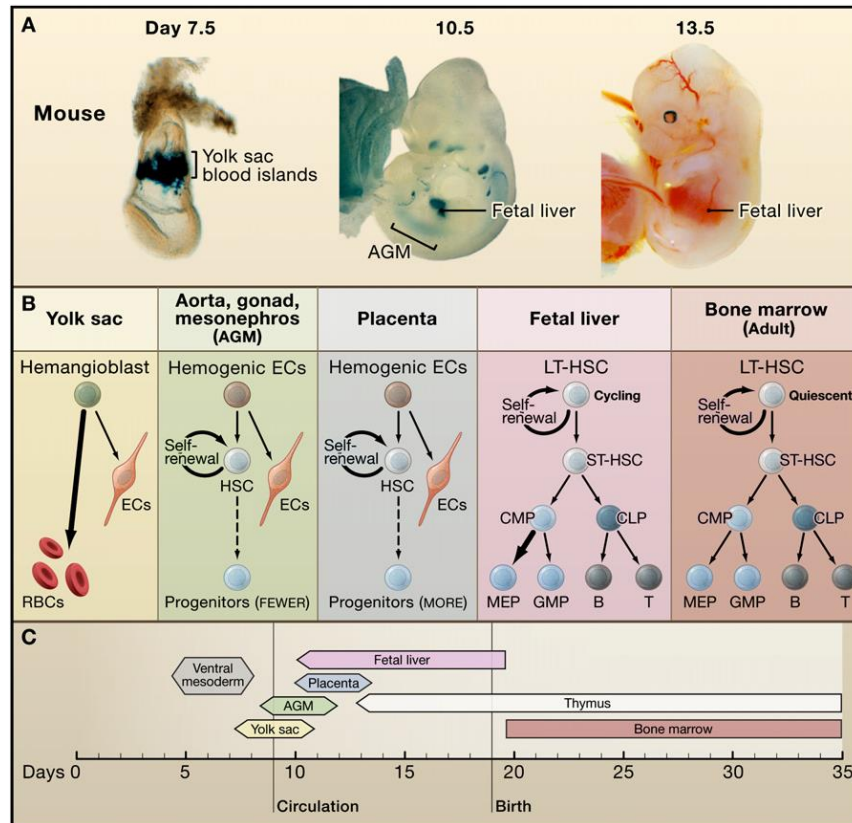


Figure 1.1. The journey from fetal to adult murine hematopoiesis. A wave of "Definitive haematopoiesis" that results in the development of HSCs occurs after "Primitive hematopoiesis," first in the aorto-gonado-mesonephros (AGM) area of the embryo, then in the fetal liver, and finally in the bone marrow. Throughout the adult lifespan, HSCs maintain a steady flow of cells that eventually make up the immune and blood systems (adapted from Orkin and Zon, 2008).

1.2 Classical hematopoietic hierarchy in the adult bone marrow

Self-renewing HSCs are found within the lineage-negative $c\text{-kit}^{\text{hi}}\text{Sca-1}^+$ population of cells in the bone marrow (LSK) (Spangrude et al., 1989). HSC loses its ability to self-renew while differentiating into a multipotent progenitor, MPP (a transit multiplying cell), which can perform restricted cycles of cell division and differentiate into a number of successive lineage-restricted progenitors. Elevated levels of the Flt3 receptor are associated with a decrease in self-renewal potential (Adolfsson et al., 2001). Gfi-1 and Bmi-1 regulatory proteins have been demonstrated to

be required for HSC self-renewal, whereas C/EBP and c-Myc appear to promote differentiation (Hock et al., 2004; Park et al., 2003; Wilson et al., 2004; Zhang et al., 2004; Laslo et al., 2008). These findings suggest that the decision of an HSC to differentiate may involve the transitory activation of C/EBP and c-Myc, which may in turn oppose the expression of Gfi-1 and Bmi-1. Initial studies have suggested that MPPs produce two primary lineage restricted intermediates: a common myeloid progenitor (CMP that produces megakaryocytic, erythroid, granulocytic, and macrophage progeny) and a common lymphoid progenitor (CLP that gives rise to B and T lymphoid cells) (Akashi et al., 2000). Subsequently, it has been shown that Flt3⁻ MPPs develop predominantly along the erythroid/megakaryocyte lineage, whereas Flt3⁺ MPPs have substantially decreased megakaryocyte and erythrocyte potential and give rise mostly to lymphoid (B and T) and myeloid lineages (macrophages and granulocytes) (Adolfsson et al., 2005). These studies have resulted in the idea that MPPs make a binary decision to become a megakaryocyte/erythroid (MEP) progenitor or a lymphoid/myeloid multipotential progenitor (LMPP) (Fig 1.2). This finding implies that the innate (myeloid) and adaptive (lymphoid) lineages share a common origin and presumably share one or more regulatory elements, such as the transcription factor PU.1 (Singh et al., 1999). MEPs develop into megakaryocytes/platelets, and erythrocytes, whereas CMPs result in macrophages, dendritic cells, and granulocytes, via granulocyte-macrophage progenitors (GMPs) (Laslo et al., 2008).

The CLPs (Common Lymphoid Progenitors) are usually regarded as the branch-point for the formation of B- and T-cells downstream of the LMPPs (Lai and Kondo, 2008; Miyazaki et al., 2014). The surface expression of the Ly6D marker distinguishes CLPs as all-lymphoid progenitors (ALPs) or B-cell biased lymphoid progenitors (BLPs). ALPs are CLPs that lack the Ly6D expression but retain the ability to generate B, T, NK, and dendritic cells (DC). Nevertheless, BLPs mainly develop B cells and have very limited T and NK cell potential (Boller and Grosschedl, 2014).

1.3 Early B cell development

Transcription factors can independently drive lineage selection, and at times, reprogram committed cells. These lineage determinants induce a distinct set of factors that act in concert with one another or with lineage-specific cytokines to form gene regulatory networks that cause multipotent progenitors to differentiate into a particular cell fate. B-cell lymphopoiesis necessitates the activation of a number of transcription factors (TFs) including PU.1, E2A, IKAROS, and

FOXO1 that 'seed' the genomic cis-regulatory areas that are then induced by the B lineage-specific TFs, EBF1 and PAX-5 (**Fig 1.3a**). EBF1 has been proven to be the principal B- cell fate determinant, since Ebf1-knockout mice exhibit a complete developmental block of B-cells at a CLP-like stage known as c-Kit+Sca-1+Flt3+IL7R+CD43+B220+CD19- (Lin and Grosschedl, 1995). These mutant cells lack expression of critical B-lineage genes such as Vpreb3, Igll1 Cd79a, Cd79b (which together form the pre-B cell receptor) and other essential genes required for B cell commitment such as Pax5, as well as the ability to undergo Ig heavy chain gene rearrangements (Zandi et al., 2008).

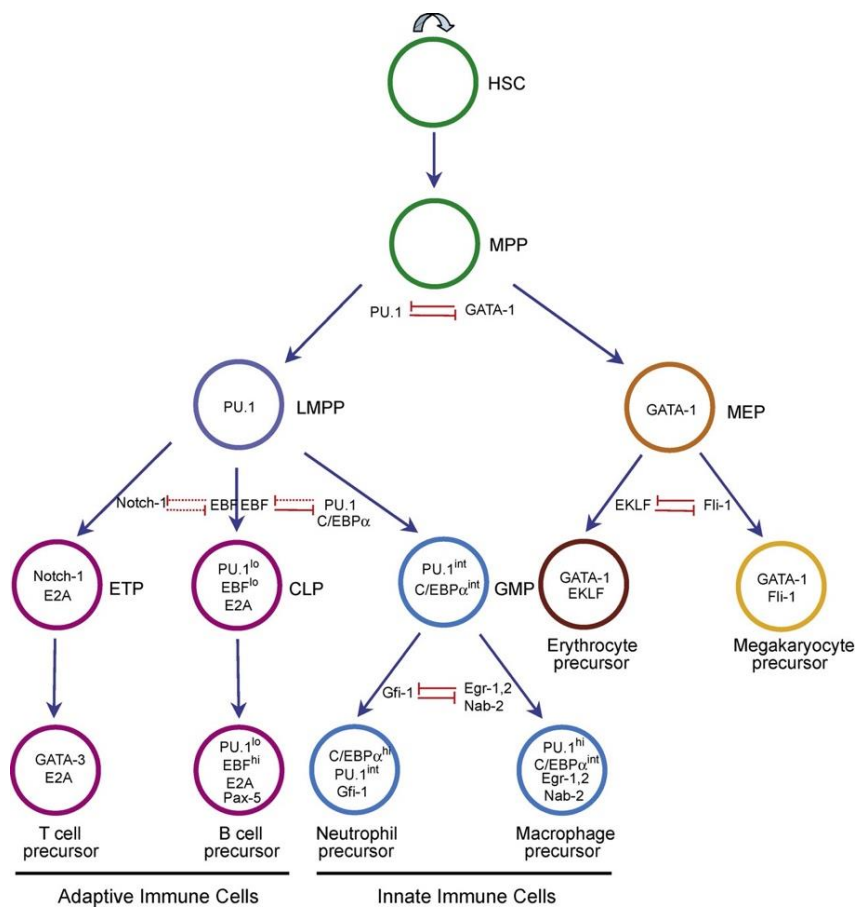


Figure 1.2. The classical hematopoietic hierarchy in the adult bone marrow The scheme describes hematopoiesis that involves a lymphoid-primed multipotent progenitor (LMPP) from which all innate (myeloid) and adaptive (lymphoid) lineages of the immune system are generated. Cross-antagonism between key transcription factors that function to regulate binary cell fate choices is noted at the appropriate bifurcation points in the developmental scheme. Transcription factors that are important for the generation of particular intermediates are noted within colored circles representing such cells. HSC (hematopoietic stem cell), MPP (Multipotential progenitor), LMPP (Lymphoid-primed multipotential progenitor), MEP (Megakaryocyte-Erythrocyte progenitor), ETP (Early thymic progenitor), CLP (Common lymphoid progenitor), GMP (Granulocyte-Macrophage progenitor) (adapted from Laslo et al., 2008).

Loss-of-function and gain-of-function analyses indicate Ebf1's critical role in B-cell fate determination. Also, Ebf1^{-/-} cells have the capacity to develop into a variety of additional lineages, including T cells, myeloid cells, and natural killer (NK) cells (Pongubala et al., 2008). Overexpression of Ebf1 but not Pax5 promotes B cell development from MPPs at the expense of myeloid lineage by suppressing myeloid determinants such as Cebpa, Sfpi1, and Id2. Similarly, inducing Ebf1 expression in Pax5^{-/-} pro-B cells reduces their T-cell potential by suppressing GATA3 expression. As a result, Ebf1 limits the alternate lineage selection of multipotent progenitors independent of Pax5 (Pongubala et al., 2008).

Furthermore, induced EBF1 expression promotes B-lineage fate choice at the cost of other hematopoietic branches in murine HSCs and overcomes the developmental block of MPPs lacking PU.1 (Liu et al., 2003). Alternate cell fate choice of MPPs is inhibited by EBF1 expression through direct suppression of alternate lineage genes such as Tcf7 and Gata3 (T-cells), Id2 and Id3 (NK cells), and Cebpa (myeloid cells) (Banerjee et al., 2013; Pongubala et al., 2008; Nechanitzky et al., 2013; Nikhat et al., 2021). Furthermore, the synchronized activity of EBF1, E2A, and FOXO1 stimulates PAX5 in specific early-B cells, leading to a positive feed-back loop of EBF1 and PAX5 that leads the progenitors to commit B-lineage at the CD19⁺ pro-B cell stage. As a result, EBF1 serves as a non-redundant pioneer in the specification and development of B-lineage identification. These findings show that the differentiation of multipotent progenitors into committed B-cells is facilitated by a complex and hierarchical gene regulation network comprised of numerous signalling molecules and lineage-specific transcription factors. (**Fig 1.3b**).

1.4 Developmental progression of early T-lineage cells

LMPPs, on the other hand, preferentially develop towards T-lineage upon homing to the thymus under the impact of Notch-DLL4 signalling. Notch1 promotes progenitor differentiation along the T-lineage by expressing early T-lineage genes. T-lymphocytes development progresses through a number of intermediate stages, including double-negative (DN) stages, which are characterized by the presence of surface markers CD44, CD25, and c-Kit. These populations can be identified as ETPs or DN1 stage (c-Kit^{Hi}CD44^{Hi}CD25⁻), DN2 stage (c-Kit^{Hi}CD44^{Hi}CD25^{Hi}), DN3 stage (c-Kit^{Lo}CD44^{Lo}CD25^{Med}) and DN4 stage (c-Kit^{Lo}CD44^{Lo}CD25^{Lo}) (**Figure 1.4**). (Rothenberg et al., 2008; Naito et al. 2011; Thompson and Zuniga-Pflucker 2011; Yui and Rothenberg 2014; Hosokawa et al., 2018).

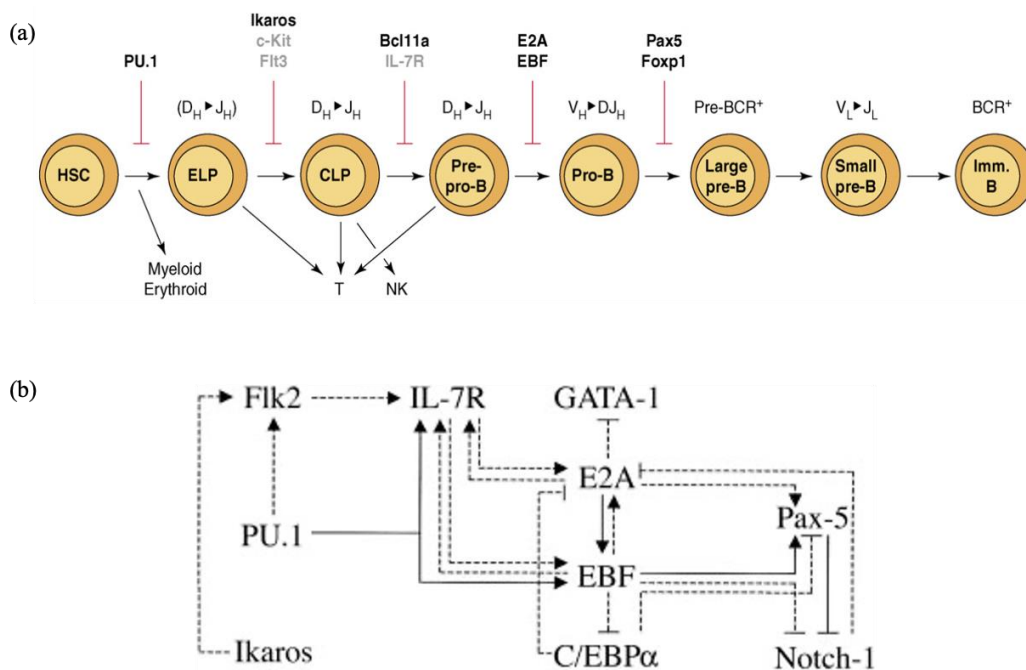


Figure 1.3 Early B cell development (a) Stages of B-cell development. B-cell development occurs in both the bone marrow. In the bone marrow, development progresses through the pro-B-cell, pre-B-cell and immature-B-cell stages. During this differentiation, rearrangements at the immunoglobulin locus result in the generation and surface expression of the pre-B-cell receptor (pre-BCR, which is comprised of an Ig heavy chain and surrogate light chains (VpreB or V5) and finally a mature BCR (comprised of rearranged heavy- and light-chain genes) that is capable of binding antigen ([adapted from Fuxa et al., 2007](#)). (b) A self-sustaining regulatory network established in a B cell precursor. It is suggested that the establishment of this network depends on transient signaling and inputs from the cytokine receptor Flk2 and the transcription factors PU.1 and Ikaros, respectively. Positive feedback loops involving the cytokine receptor IL-7R and the transcription factors EBF, E2A, and Pax-5 may generate a self-sustaining circuit. The network architecture also features cross-antagonism with alternate cell fate-determining transcription factors such as GATA-1, C/EBPα, and Notch-1. Stimulatory and inhibitory inputs are indicated as solid or dashed lines, depending on the strength of the experimental evidence ([adapted from Singh et al., 2005](#)).

c-Kit expression is required until the DN2 stage and then declines progressively through the DN3 stage, whereas IL-7R expression is increased at the DN2 and DN3 stages until β selection is achieved. T-cell identifying genes such as Cd3 γ , Il7R α , Cd3 ϵ , and Zap70, which are present in low amounts in ETPs, continue to be elevated in the DN2 stage. The DN3 stage is associated with the germline transcription of different genes encoding TCR variable regions, coupled with V-(D)J rearrangement that defines T cell lineage commitment. The β chain then pairs with the pre-T α (surrogate chain), and in combination with CD3 molecules, forms a complex known as the pre-TCR, which leads to sustenance, proliferation, and arrest in further β chain loci rearrangement, as

well as differentiation to the double positive (DP) stage via up-regulation of CD4 and CD8. (Rothenberg 2005; Petrie and Zuniga-Pflucker 2007; Love and Bhandoola 2011).

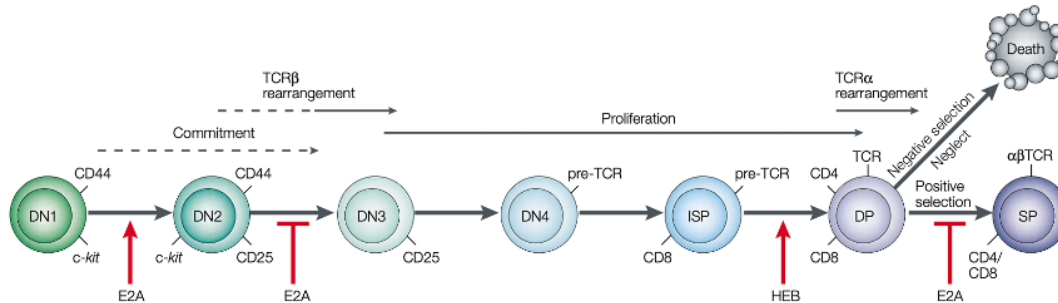


Figure 1.4 Developmental progression of early T-lineage cells. A schematic representation of the developmental process of T cells, illustrating the various cell surface markers that are expressed during the various stages of T cell maturation in the mouse. The thymus is the organ responsible for the generation of T cells, which progress through the developmental phases of double negative (DN), double positive (DP), and single positive in a sequential order (SP). DN cells further differentiated into the phases of DN1, DN2, DN3, and DN4. During the DN2 and DN3 stages, the TCR rearrangement process takes place. During the DN3 stage, molecules of TCR are expressed; this causes -selection, which in turn leads to the development of DP. DP go through a process called TCR rearrangement and then continue to develop into CD4, CD8SP, etc (adapted from Engel and Murre, 2001).

The presence of cell surface ligands and growth factors in the thymic epithelium promotes the T cell fate choice of multipotent progenitors. Particularly, the delta-like ligand 1 (DLL1) and DLL4 ligands, which are expressed at high levels by thymic stromal cells, have been shown to drive the growth of ETPs towards the T-lineage. Signaling through these ligands leads to the activation of the Notch intra-cellular domain that translocate into the nucleus and induces the T cell developmental programme by activating early-T genes such as *Ptcra* and *Cd3e*, as well as important TFs such as TCF-1 (encoded by *Tcf7*), which is essential for preserving T-cell development by activating key factors such as GATA3 and BCL11B (Sambandam et al., 2005; Pai et al., 2003; Weber et al., 2011; Ikawa et al., 2010). Gain-of-function studies have demonstrated that NOTCH1 activation in the bone marrow promotes thymus-independent differentiation of early lymphoid progenitors into T cells at the expense of B lymphopoiesis (Pui et al., 1999). Consistent with these studies, loss-of-function studies have shown that conditional ablation of the Notch1 gene in HSCs blocks the developmental progression of the T cells at the DN1 stage while increasing B cell differentiation in the thymus (Radtke et al., 1999; Wilson et al., 2001). NOTCH1/RBPJ and GATA3 inhibit B-lineage T-lineage commitment by

suppressing the formation of natural killer (NK) cells (Li et al., 2010; Scripture-Adams et al., 2014).

1.5 Transcriptional control of T cell development

T cell development is tightly controlled by multiple transcriptional regulators. First, the E-proteins are turned on, which activates a network of genes implicated in the Notch signaling pathway (Bain and Murre, 1998; Ikawa et al., 2006; Miyazaki et al., 2017). Notch signaling then activates Bcl11b, GATA-3, and TCF1 expression (Yui and Rothenberg, 2014). To help cells progress from the DN2a to DN2b stage, Bcl11b is activated at this step. When combined with E2A, Bcl11b further activates the lineage defining gene network. The alternate cell fate determining gene is also repressed by Bcl11b and E2A (Liu et al., 2010; Ikawa et al., 2010; Li et al., 2010a; Longabaugh et al., 2017). It is well established that ETP require Gata3 in order to mature into DN2 thymocytes. E2A and HEB, both members of the basic helix-loop-helix (bHLH) transcription factor family, are also crucial in the first stages of T-cell maturation (Dias et al., 2008). Also crucial for early thymocyte development is the RUNT-related transcription factor (Runx) complex, which consists of a Runx protein and an obligatory non-DNA binding partner, core-binding factor-beta (Cbf-beta) (Talebian et al., 2007). High levels of IL-7 were found to be necessary for the in vitro maintenance of early-DN2 status. When IL-7 levels were lowered, DN3 cells began to differentiate into DP thymocytes (Ikawa et al., 2010).

Previous studies (Ikawa et al., 2010; Li et al., 2010) have found that a transcription factor, Bcl11b, is essential for complete commitment to the T-lineage, providing molecular evidence for an essential branching point at the DN2 stage. Detailed analyses of Bcl11b germ line knockout mice revealed that thymocyte development was blocked at the c-Kit⁺ CD25⁺ DN2 stage in the absence of Bcl11b (Ikawa et al., 2010; Li et al., 2010). Conditional removal of Bcl11b from T-lineage-committed cells such as DN3 or even DP thymocytes results in the emergence of NK-like cells, presumably via de-differentiation or trans-differentiation (Li et al., 2010). These results demonstrated that Bcl11b is essential for T-lineage commitment, which happens during early-to-late DN2 transition. Interestingly, at the DN2 stage, an enhancer within the intergenic locus control region of Bcl11b induces the expression of Bcl11b. Transcription factors such as Notch, GATA-3, TCF1, and RUNX1 were found to bind to the Major Peak enhancer (Guo et al., 2008; Weber et al., 2011; Garcia-Ojeda et al., 2013; Li et al., 2013). Furthermore, a recent study has shown that the chromatin state transitioning from inactive to active is the rate limiting step of full activation

of Bcl11b (Kueh et al., 2016). It was also found that E2A is necessary to preserve the integrity of the β -checkpoint, which occurs during the DN2 to DN3 transition (Engel et al., 2001; Engel et al., 2004). Consistent with these findings, the expression of a dominant-negative HEB protein disrupts the function of E2A and HEB, which leads to dysregulation of V(DJ) rearrangement and, as a result, an arrest in embryonic progression at the DN3 stage (Barndt et al., 2000). When pre-TCR signaling occurs, there is an increase in the amount of an inhibitor of E2A called inhibitor of DNA binding 3 (Id3), which in turn down-regulates the activity of E2A (Figure 1.5 a & b) (Engel et al., 2001; Naito et al., 2011).

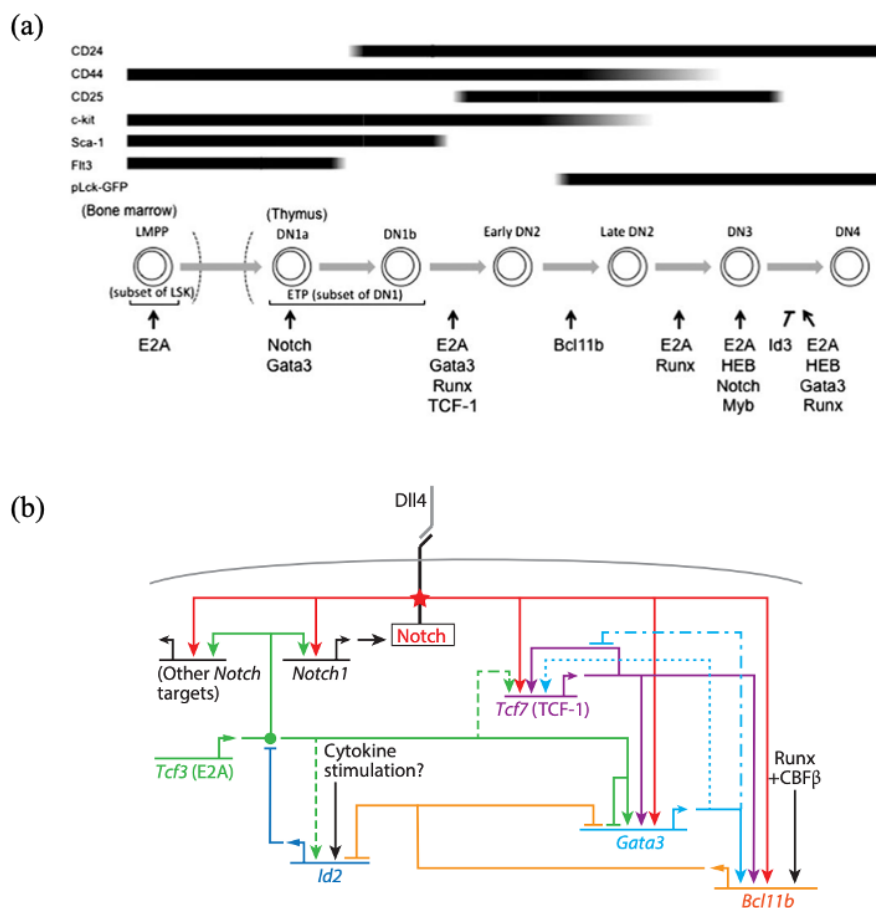


Figure 1.5 Transcriptional control of T cell development and gene regulatory network. (a) In the middle of the diagram are depicted the stages of early T-cell development. The expression of the markers that define each step is displayed at the top of the chart. At the bottom is a diagram that illustrates the stages at which each transcription factor is active. The areas denoted by black arrows are those in which transcription factors are essential, whereas the areas denoted by the black bar are those in which Id3 inhibits transcription (adapted from Naito et al., 2011). (b) Schematic representation of T cell fate gene regulatory network. The horizontal lines denote genes coding for regulatory factors; bent arrows denote transcription and translation of gene leading to product (itself a regulatory factor). The star indicates that the Notch transmembrane protein is activated by binding to its ligand Delta-like 4 (Dll4). Although many genes require Notch input for activation, Gata3, Tcf7, and Bcl11b do not require its continuation for maintenance

of expression. Solid lines denote relationships with strong molecular and functional support within the context of T-lineage specification, dashed lines indicate relationships seen in related cells, dotted lines indicate interactions suggested by DNA binding, and the long dash-dot line denotes an inferred effect of GATA-3 on Tcf7 (adapted from Rothenberg, 2014).

1.6 E and Id proteins in T cell development

During the first stages of the production of T cells, the E2A gene is also very significant. The activation of E2A expression in thymocytes that have not yet committed to a specific lineage is a prerequisite for the successful launch of the T-cell developmental program. It is also important to note that the deficiencies in E2A null DN thymocytes appear to influence something other than the production of T lymphocytes (Engel and Murre 2001). There are four mammalian E-proteins that have been discovered: E12, E47, E2-2, and HEB. Each of these mammalian E-proteins can bind to E-box elements in either a homodimeric or heterodimeric configuration, depending on the type of bHLH protein that is present in the tissue. Two of these proteins, E12 and E47, are generated from the E2A transcript following alternative splicing. The only difference between E12 and E47 is the exon that codes for the bHLH domain. Both proteins have identical amino-terminal transactivation domains (Murre et al., 1989). Even though HEB and E2-2 proteins are similar to the E2A proteins, they are encoded by distinct genes (Hu et al., 1992; Henthorn et al., 1990; Bain and Murre, 1998).

The proteins encoded by the Id gene operate as antagonists to the DNA-binding activities of the E-proteins (Benezra et al. 1990). There have been four different Id proteins discovered in mammalian genomes, and they have been given the names Id1–4. Id proteins have an HLH domain, but they don't have a basic area. When they heterodimerize, they stop the DNA-binding activities of bHLH proteins since they don't have a basic region. Studies show that a decrease in E-protein activity overcomes the need for Id expression to promote embryonic progression, indicating that the primary function of Id proteins is to target E proteins (**Fig 1.6a**) (Yan et al. 1997; Rivera et al. 2000; Boos et al. 2007; Miyazaki et al. 2011, 2017; Zook et al. 2018). Id proteins are involved in a diverse array of developmental pathways that are involved in both health and disease (Lasorella et al. 2014). They perform most of their functions by influencing the progression of the cell cycle, developmental progression, and tumor suppression (Lyden et al. 1999; Yokota et al. 1999; Lasorella et al. 2014; Miyazaki et al. 2015; Murre, 2019).

It has been shown that the E-Id protein axis controls the expression of a large set of genes involved in immune cell development, depending on the type of cell and the stage of development (**Fig 1.6b**). Genes encoding transcription factors, signaling pathway components, antigen receptors, chemokine receptors, DNA repair factors, enzymes involved in somatic recombination, and many other functions are among them. Expression of Id2 and Id3 is dynamically and intricately regulated in developing T cells. Pre-TCR and TCR signaling via the ERK-MAPK-EGR1 pathway activates Id3 expression during thymocyte selection ([Bain et al. 2001](#)). Blocking the activation of the kinases MEK1 and MEK2 of the mitogen-activated protein (MAP) kinase family, which are involved in signal transduction via the RAS extracellular-signal-related kinase (ERK)-MAP kinase pathway, reduces E-box binding and prevents the induction of Id3 ([Alessi et al., 1995](#); [Dudley et al., 1995](#)). It has been established that this signaling pathway is crucial for positive selection ([Pages et al., 1999](#), [Alberola-Ila et al., 1995](#)). More evidence suggests that activation of the MAP kinase-responsive transcription factor Egr1 is responsible for inducing Id3 transcription (early growth-response factor 1) ([Bain et al., 2001](#)). Overall, these findings support a concept in which TCR ligation mediates the suppression of E-protein activity by inducing Id3 via the activation of the RAS-ERK-MAP kinase cascade (**Fig 1.6c**) ([Engel and Murre 2001](#)).

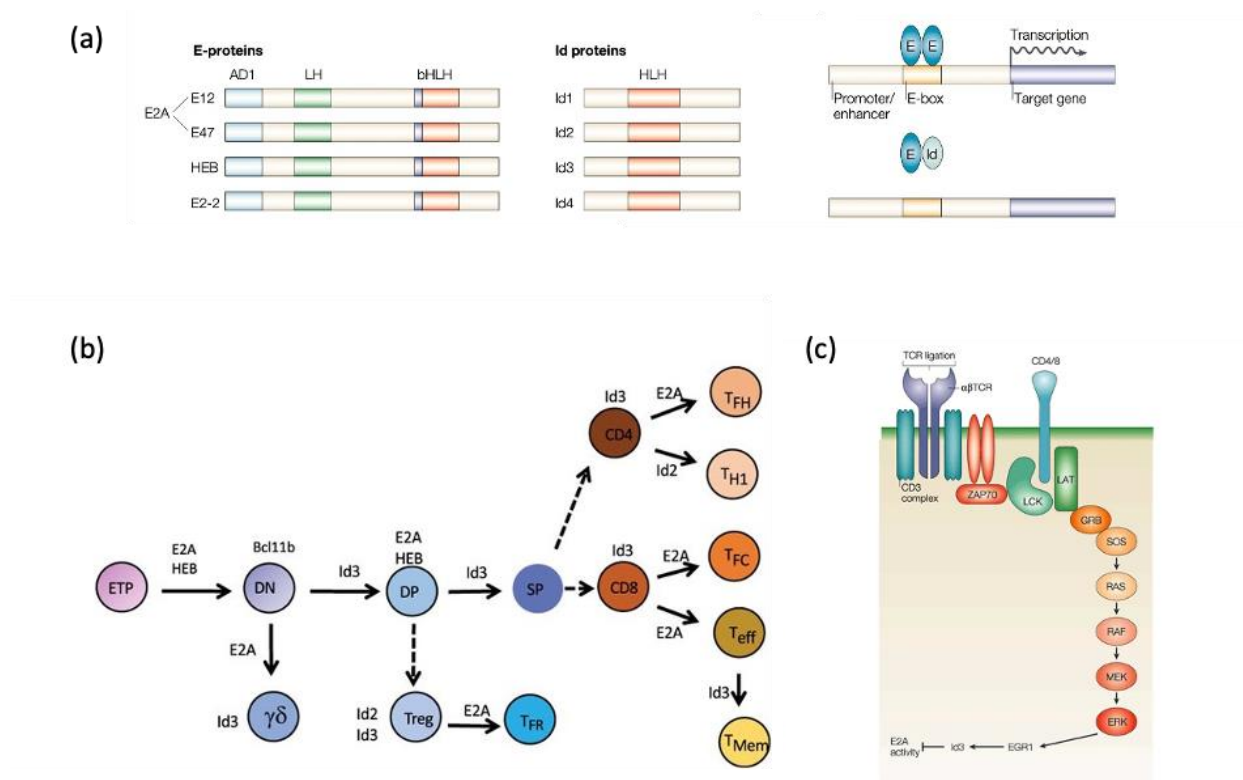


Figure 1.6 E and Id proteins in T cell development (a) Schematic representation of the E and Id proteins, emphasizing the basic and helix-loop-helix (HLH) DNA-binding and dimerization domains as well as the AD1 and LH (AD2) activation domains. The binding of E-protein dimers to E-box sites in promoters and enhancers to induce transcription of target genes is depicted. Because Id proteins are deficient in the basic

domain necessary for DNA binding, Id-E-protein dimers are unable to bind promoters and initiate transcription ([adapted from Engel and Murre, 2001](#)) (b) Immune cell diversity and HLH proteins. Evidence points to HLH proteins playing a function in hematopoiesis and immune cell formation. Transitions between the primary and secondary lymphoid compartments are represented by dashed lines, indicating the presence of several intermediates during development ([adapted from Murre, 2019](#)). (c) Reduced E2A activity due to TCR signaling inhibition. Illustration of how the T-cell receptor (TCR)-mediated stimulation of the RAS-ERK-MAP kinase pathway suppresses E2A activity by increasing EGR1 and, ultimately, Id3 expression ([adapted from Engel and Murre, 2001](#)).

1.7 $\alpha\beta$ and $\gamma\delta$ T cell fate choice

There are two major categories of T cells that can be distinguished from one another: $\alpha\beta$ and $\gamma\delta$ T cells. It is generally agreed upon that after progenitors have committed to the T-cell lineage, the first decision they make regarding their lineage is whether or not to become a $\alpha\beta$ or $\gamma\delta$ T cell. Historically, $\alpha\beta$ and $\gamma\delta$ lineages were distinguished from one another based on the kind of T-cell receptor (TCR). CD4⁻CD8⁻ thymocytes, also known as "double negative" (DN) cells, are responsible for the rearrangement of three of the four TCR loci. These loci are TCR β , TCR γ , and TCR δ . At this point, the proliferation of the cells has ceased, and in order for them to restart the cell cycle, the expression of TCR is necessary. The TCR chain is expressed as a complex with the pre-TCR (pT α) chain, which is encoded in the germline, if an in-frame TCR β rearrangement is successful in each cell. When this complex, known as pre-TCR, is expressed, a cascade of events occurs, including a proliferation surge, the overexpression of the CD4 and CD8 coreceptors, the silencing of TCR γ , and the commencement of TCR α rearrangement (which results in the excision of the TCR δ locus). Thymocytes that are CD4⁺ and CD8⁺ (double positive (DP) because of a functional rearrangement of TCR α express TCR on their cell surfaces and can go on to specialize in either the CD4⁺ (helper) or CD8⁺ (killer) lineages. It is often held that reaching the DP level is indicative of a dedicated lineage.

These progenitors' express surface TCR and are responsible for productive rearrangement of TCR γ and TCR δ loci. These cells also experience a proliferation burst, but in wildtype (WT) mice, the vast majority of them bypass the DP stage and exit to the periphery with a CD4⁻CD8⁻ (or, less frequently, a CD4⁻CD8⁺ or a CD4⁺CD8⁻) phenotype. The $\alpha\beta$ and $\gamma\delta$ lineages are currently classified on the basis of passage through the DP stage (lineage) or absence of $\gamma\delta$ lineage, as a common molecular program provides a better basis for lineage determination than the expression of a single receptor (TCR). Targeted activation of resting thymocytes occurs through either β - or $\gamma\delta$ -selection, depending on whether or not they express pre-TCR or $\gamma\delta$ -TCR ([Kreslavsky et al., 2010](#)). TCR signals, depending on their intensity, trigger Id3 expression. E

proteins are essential regulators of T cell growth, and Id3 blocks their activity. Transduction of potent TCR signals by the $\gamma\delta$ TCR results in dramatic upregulation of Id3, downregulation of E protein activity, and the selection of the $\gamma\delta$ T cell destiny. $\alpha\beta$ T cell fate commitment is achieved in contrast by weak, temporary signals induced by the pre-TCR complex, which induce less Id3 and sustain more E protein function (Fahl et al., 2016).

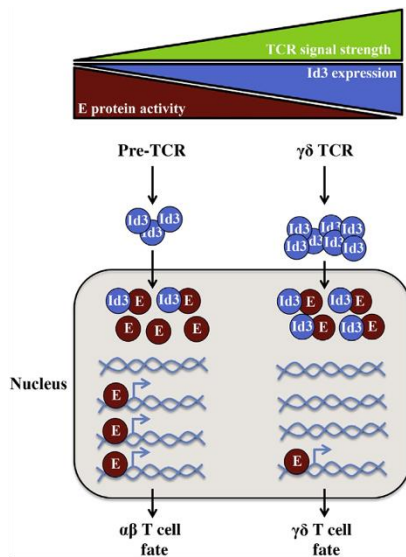


Figure 1.7 $\alpha\beta$ and $\gamma\delta$ T cell fate choice Gradual E protein activity determines T lineage commitment. Id3 expression is proportional to TCR signal intensity. Id3 inhibits T cell growth regulators E box-binding family DNA binding proteins. E protein activity decreases with TCR signal strength. Strong TCR signals from the TCR transduce substantial increases in Id3, significant reduction of E protein function, and $\gamma\delta$ T cell destiny. The pre-TCR complex generates weak, transient signals that suppress Id3 and sustain E protein activity, resulting in $\alpha\beta$ T cell destiny (adapted from Fahl et al., 2016).

1.8 Regulation of genome 3D structure in immune cell development

In a metazoan cell, there are nearly 20,000 genes that span across a genome of 6469.66 Mb. It is estimated that the genome of a single nucleus measures a total length of about 3 meters, while the average diameter is about 6 μ M. Majority of the genes always start with a genetic regulatory element known as a promoter containing the TATA box, whose activity is regulated by cis-regulatory elements such as enhancers and insulators in a cell- and developmental-stage specific manner. Thus, the genome is organized non-randomly to ensure that the regulatory molecules, including transcriptional factors, mediators, polymerases, etc. can find their target efficiently in the crowded nucleus. With the advance of technologies such as the chromosome capture conformation assay (3C) and its related high-throughput approaches, 4C, 5C, HiC and HiC-ChIP, we have revealed the genome architecture at multiple scales. First, chromosomes have their own chromosome territories that rarely intermingle, except for nucleoli (Boya et al., 2017; Sexton and Cavalli, 2015; Maharana et al., 2016). Second, each of the chromosomes is partitioned into a highly self-interacting domain known as topologically associated domains (TADs), which can be either euchromatic (A) or heterochromatic (B) compartments (Lieberman-Aiden et al., 2009; Boya et al., 2017). It is well established that the euchromatic genome is positioned in the nuclear interior, while

the heterochromatic compartments tend to be associated with the nuclear lamina (Peric-Hupkes et al., 2010; Kind et al., 2015). Both the A and B compartments exhibit strong enrichment of intra-domain interaction and few inter-domain interactions. The purpose of dividing the chromosome into compartments is thought to be minimizing unwanted interaction between genes and regulatory elements not in the same domain. During developmental progression, gene elements often reposition from the lamina to the nuclear interior and vice versa to modulate gene expression (Lin et al., 2012; Isoda et al., 2017).

Enhancers were found to regulate their target genes from a distant genomic location independent of their orientation. Enhancers acquire their activity through the combinatorial binding of various transcription factors. It has been proposed that enhancers interact with their cognate target gene promoters by looping intervening DNA, which thereby promotes transcription. The enhancer-promoter interaction can form large loops known as ‘regulatory loops’ on the chromatin and bring distant enhancers into proximity with the promoter (Fig 1.8). Enhancer promoter interaction can be regulated by many different factors, including CCCTC-binding factor (CTCF), cohesin, and some non-coding RNAs. While NIPBL and MAU2 promote the formation of the cohesin complex on the DNA, Wapl and its cofactors unload the cohesin. Recruitment of cohesin results in extrusion of chromatin until a pair of convergent CTCF binding sites is reached. (Pongubala and Murre, 2021).

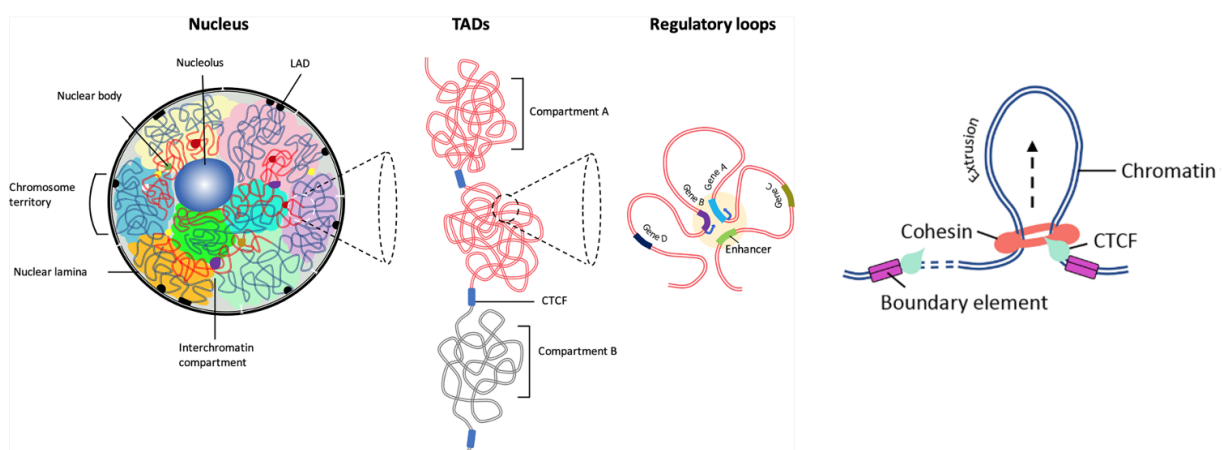


Figure 1.8 Regulation of genome 3D structure in immune cell development. The multi-scale organization that ultimately results in a hierarchical genome architecture, with chromosome territory formation occurring at the nuclear scale, formation of compartments occurring at the chromosomal scale, TADs arising at the sub-chromosomal scale, and regulatory loops occurring at the sub-megabase scale. (adapted from Pongubala and Murre, 2021).

1.9 Non-coding RNAs in early T cell development

Even though a number of studies have shown how lineage-specific transcriptional regulators, cytokine signals, and epigenetic modifications control the development of B and T lymphocytes, the significance of the non-coding part of the genome (i.e., non-coding RNAs) remains unclear. Non-coding RNAs include microRNAs (miRNAs), small RNAs (sRNAs), piwi-interacting RNAs (piRNAs), siRNAs, circular RNAs (circRNAs), and long non-coding RNAs (lncRNAs) ([Homback and Kretz, 2016](#)). These RNAs have been known to play a significant role in various physiological processes in the cell. LncRNAs have emerged as important regulatory molecules with complex gene regulatory functions ([Liu et al., 2018](#)). LncRNAs were found to be transcribed by RNA polymerase II, localized primarily in the nucleus, unstable, and expressed in a cell type specific manner ([Hube and Francastel, 2018](#)). LncRNAs regulate gene expression in a variety of ways at epigenetic, chromatin remodelling, transcriptional, and translational levels. For instance, long non-coding RNA X inactive specific transcript (lncRNA Xist) regulates X chromosome inactivation leading to the silencing of one of the X-chromosomes during female cell development ([Starmer and Magnuson 2009](#); [Wang and Chang 2011](#); [Lee et al., 2013](#); [Yang Z et al., 2018](#)). Some lncRNAs bind to proteins and recruit them to specific sites, leading to the regulation of the expression of target genes. This is essentially mediated through the recruitment of chromatin modifying enzymes that alter the chromatin state ([Rinn et al., 2007](#); [Loewe et al 2010](#)). LncRNAs can also control transcription by sequestering transcription factors, miRNAs, and chromatin modifiers ([Wang and Chang 2011](#)) away from their respective target sites. Other lncRNA loci regulate gene expression in cis by having a transcriptional enhancer-like function for genes that are located on the same chromosome.

Recent studies indicate that the expression of a nearby protein-coding gene can be controlled by the transcription of long noncoding RNA (lncRNA) as well as by the splicing of transcripts ([Engreitz et al., 2016](#)). Even though the level of lncRNA expression within a certain cell type may be able to predict its biological activity ([Liu et al., 2016](#); [Arun et al., 2016](#)), lncRNA loci are still capable of having enhancer-like action even when there is no transcription taking place ([Groff et al., 2016](#)). In addition, lncRNA loci that are known to act in cis can have distinct biological functions in trans ([Paralkar et al., 2016](#)).

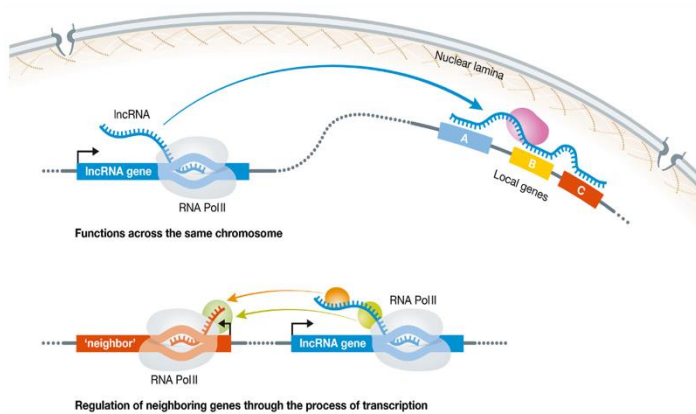


Figure 1.9 Non-coding RNAs in early T cell development: mechanisms of lncRNA activity that are mediated by cis-acting elements. The chromosome from which the lncRNA transcript is transcribed is the same chromosome that the transcript functions along. Through the mechanism of their own transcription, lncRNAs are also able to function in the cis orientation (adapted from Liu et al., 2018)

1.9.1 Role of lncRNA-ThymoD in T cell development

According to findings from previous research, a lncRNA called ThymoD is necessary for the transmission of precursors along the T-cell lineage. During the T-cell progenitor stage, transcription of ThymoD causes the Bcl11b enhancer to move from the nuclear lamina, which is a repressive compartment, to the nuclear interior, which is a permissive compartment. This leads to enhanced gene expression. The compartment switch helps to bring the distant enhancer closer to the appropriate promoter so that they can function together more effectively. These regions then stay within the same topologically associated domain (TAD) throughout the early stages of T cell development. As a result, they can interact with one another at a high frequency in a manner that is lineage-specific, which in turn enables the expression of the T-cell gene expression program. Following the formation of the loop domain, the expression of Bcl11b can then promote a gene program for the T-lineage while simultaneously blocking the expression of alternate cell lineage programs (Isoda et al., 2017). Even though these studies show that ThymoD transcription is necessary for the expression of Bcl11b, which is an essential checkpoint for the early determination of the T lineage, they did not completely elaborate the framework by which loop extrusion takes place in these progenitors.

1.10 Dynamics of transcription

Even though transcription is a very dynamic process, we do not yet have a complete understanding of the function of the dynamic features of the transcription apparatus or the implications they have. New technologies during the past two decades have shed light on the fact that transcription is an inherently discontinuous process. This process is characterized by transitory bursts of transcriptional activity, interrupted by periods of stillness. There are two distinct forms of transcriptional dynamics that are displayed by the promoter while transcription is being initiated.

The first type is one in which the amount of nascent RNA that is created at the promoter oscillates around a positive value that is constant. RNA Pol II starts the transcription process at a pace that is stable on average. The promoter is active and non-bursting. The second type of transcription process is one in which many polymerases transcribe the gene at once during distinct periods of promoter activity. This period of activity may be continuous in time, or it may be punctuated by stretches of inactivity during which no polymerases begin transcription. The term "bursts" or "pulses" in transcription describes these periods of intense activity. Expression of developmentally regulated genes typically occurs in such bursts, although housekeeping genes' bursts during development are much weaker because of large reductions in their amplitude (Ferraro et al., 2016).

Frequency, intensity, and duration of bursts are the defining features of a bursting promoter. Burst amplitude (i.e., the quantity of transcripts being created during the burst), burst duration (i.e., the time frame when the promoter is active), and burst frequency (i.e., the duration between two sequential bursts) all contribute to the overall rate of transcript synthesis in a cell.

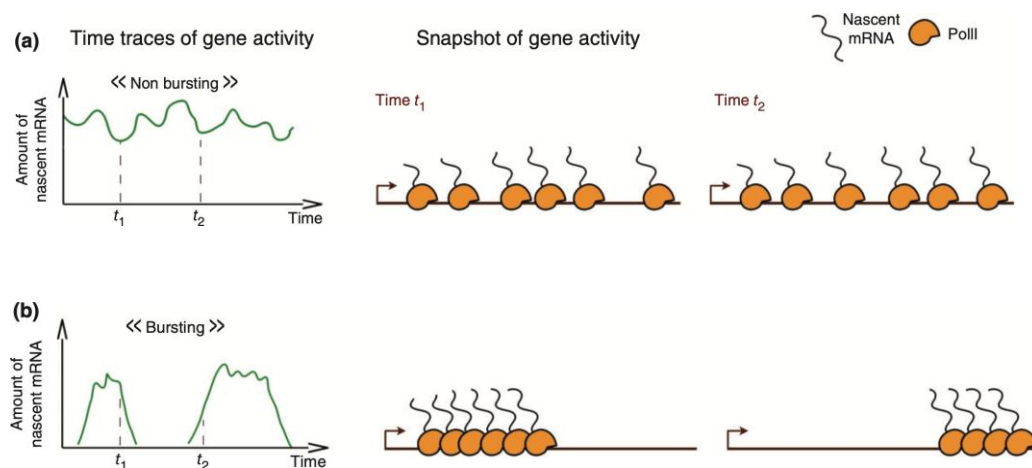


Figure 1.10 Dynamics of transcription: Transcription dynamics of a promoter: (a) non-bursting and (b) bursting (adapted from Ferraro et al., 2016).

1.11 What triggers a transcription burst?

1.11.1 Proximal Regulation

The assembly of the pre-initiation complex (PIC) of RNA polymerase II (Pol II) is the rate-limiting step in transcriptional activation. After the binding of sequence-specific factors to the promoter regions, the nascent transcript starts to grow. It is well known that transcription factors bind to their sites based on not only consensus DNA sequences, but also flanking nucleotides and adjacent

binding factors. Also, expression levels of transcription factors can either increase or decrease binding affinity. Therefore, it is highly likely there exists a positive correlation between burst frequency and binding affinity if transcription factors are more frequently bound at a given promoter. Longer stays by transcription factors might improve PIC formation and the subsequent transcription cycle (Rodriguez and Larson, 2020).

1.11.2 Distal Regulation

Enhancers can be found mega-bases away from their target genes but are essential for the transcriptional activation of genes during development, circadian rhythms, and nuclear receptor activity (Banerji et al., 1981; Catarino et al., 2018; Fang et al., 2014; Hah et al., 2011; Lettice et al., 2003). A higher concentration of co-factors and Pol II may be present in the vicinity of a target gene if an enhancer is present there. This is because enhancers typically contain several transcription factors binding sites that contribute to the recruitment of these factors. Enhancing transcription factor levels locally may improve promoter sampling rate and hence the likelihood of triggering a transcriptional burst (Dufourt et al., 2018; Pennacchio et al., 2013). The second model proposes that mediators, transcription factors, or chromatin re-modelers can be deposited on the target gene promoter through transitory chromosomal looping between enhancers and promoters. There is some evidence that this temporary looping can lead to the creation of a permanent complex that attracts the machinery required to activate target genes (Thanos et al., 1995). There is a gap in our understanding of the relationship between the stability of this enhancer-promoter complex and bursting dynamics. If stability is required for transcriptional initiation, however, persistent stability may result in re-initiation events, which would increase the burst size (Rodriguez and Larson, 2020).

1.12 Visualization of transcription dynamics

The visualization of mRNA in its natural environment has been accomplished using a variety of optical approaches. The study of the subcellular localization of messenger RNA (mRNA) in both fixed (Sun et al., 2011) and live cells (Lukinavicius et al., 2014) has provided new insights on the kinetics of transcription (Lang et al., 2014; Snijder et al., 2012), mRNA export (Kaiser et al., 2014), and translation (Lukinavicius et al., 2014; Zhao et al., 2015).

1.12.1 Fixed-cell techniques for RNA detection

A single-molecule RNA FISH, also known as smFISH, is a type of fluorescence in situ hybridization that makes it possible to count individual mRNA molecules in cells or nuclei that have been fixed ([Femino et al., 1998](#)). During this process, cells are permeabilized, fluorescently labelled DNA probes are hybridized to specific mRNA molecules. This produces strong fluorescent dots both at active sites of transcription and in other parts of the cell ([Raj et al., 2008](#)). Calculating the probability distribution of the number of molecules that are found in each cell or nucleus is one way to analyze these data. This method holds great potential for the development of sophisticated systems biology models to anticipate global cellular responses and to elucidate the dynamics of cellular heterogeneity. Even if it is insightful, this method does not provide the temporal resolution that is required to investigate the expression program ([Specht et al., 2017](#)).

1.12.2 Live cell RNA probes

Live-cell imaging was developed for in-depth research on transcription kinetics. Live-cell RNA imaging probes come in a few different varieties, and most of them offer spatial resolution within the cell. Fluorescence is activated in the RNA aptamer-fluorogen system by the binding of a cell-permeable small molecule GFP like DFHBI to a tiny aptamer (named Spinach/Broccoli) that has been genetically integrated as a fusion of the RNA of interest. Aptamer-fluorogen systems are adequate for imaging highly expressed non-coding RNAs (tRNAs, rRNAs, etc.), but they are not effective for imaging messenger RNA (mRNA) because of their low brightness ([Rodriguez and Larson, 2020](#); [Specht et al., 2017](#)).

In the CRISPR/Cas9 system, an FP is joined to a genetically modified form of Cas9 that stays in the nucleus, unless it is kept in the cytosol by binding to an endogenous RNA of interest using a PAMmer that is sequence-specific. In this case, it is moved into the cytosol (protospacer adjacent motif-presenting oligomer). If a complementary DNA oligonucleotide is placed at the appropriate target location, Cas9 nuclease is able to recognise ssRNA even though its primary function is to bind DNA ([O'Connell et al., 2014](#)). This method was extended to RNA imaging in living cells via a GFP-Cas9 fusion, and it was discovered to produce no disturbance to RNA levels or localization ([Nelles et al., 2016](#)).

Many different RNA-binding proteins that have developed over time are used to direct fluorescent proteins to specific RNAs. The MCP (MS2 coat protein) and PCP (PP7 coat protein) orthogonal systems were developed from bacteriophage proteins that bind specific RNA stem-

loops and are widely employed (Buxbaum et al., 2014). To improve the fluorescence contrast, several aptamers are usually added to a single transcript. In the MS2 system, the 3' end of the target RNA is modified with repeating short RNA aptamers. Two FPs are joined to an aptamer-binding protein in each aptamer. Which generates a strong signal at the transcriptional promoter (Specht et al., 2017).

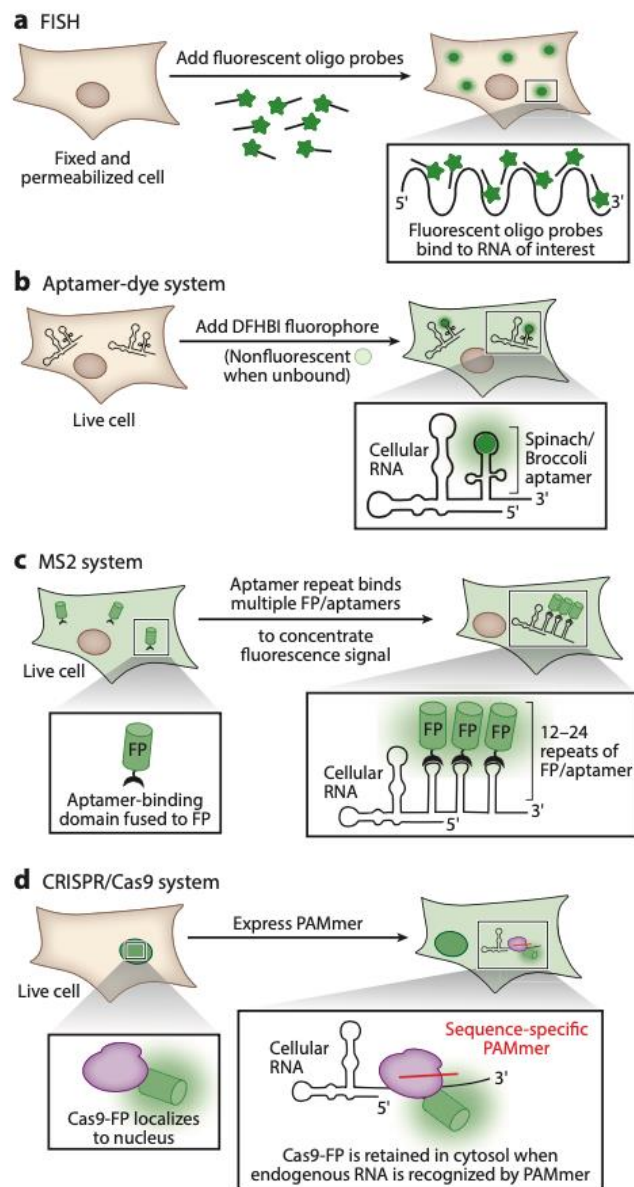


Figure 1.12 Visualization of transcription dynamics Methods of fluorescently labelling RNA in fixed or live samples (a) Sequence-specific fluorescent oligo probes bind the RNA of interest in fixed permeabilized cells, hence labelling endogenous RNA. (b) The target RNA is fused genetically with a tiny aptamer (named Spinach/Broccoli), and fluorescence is activated following binding of the cell-permeable small chemical DFHBI. (c) Cas9 has been engineered so that it stays in the nucleus unless it binds to an endogenous RNA of interest via a sequence specific PAMmer, in which instance it is kept in the cytoplasm. Fluorescent protein; PAMmer; protospacer adjacent motif-presenting oligomer. (d) A series of short aptamers are inserted at the 3' end of the target RNA. Each aptamer recognizes a specific aptamer-binding protein that has been fused to one or more FPs (adapted from Specht et al., 2017).

1.13 Overview of the MS2-MCP system

The MS2 system makes it possible to take real-time measurements of dynamic processes in living cells. Coat proteins that are genetically expressed and fluorescently tagged (MCP-GFP, green circles) attach to repeats (8 to 128) of a stem-loop sequence that has been introduced into the gene sequence that is of interest. The MS2-MCP system is comprised of two different parts. The MS2 sequence is an RNA aptamer that was derived from bacteriophages. It is incorporated into the 3'UTR of the mRNA sequence as an array of 24 repeats that form 24 stem loops (Larson et al., 2013). The second component of the system is a homodimer of the MCP (MS2 Capsid Protein), and each MCP is fused to GFP. In cells that express both components, the selective binding of the MCP to the MS2 provisionally decorates each mRNA with 48 FPs. This occurs only in the cells that express both components. The measurement of transcription in real time is made possible thanks to the highly focused fluorescent signal (Tutucci et al., 2018).

If a gene is expressed constitutively or in bursts, the frequency (f), amplitude (A), duration (t_{on}), and time between transcriptions (t_{off}) can all be evaluated with this technique. This allows one to identify whether a gene is expressed continuously or intermittently (Bentovim et al., 2017). The capability of making these inferences about kinetic parameters opens up the possibility of offering a deeper understanding of the spatial and temporal controls of bursting at single-cell resolution (Bowles et al., 2022).

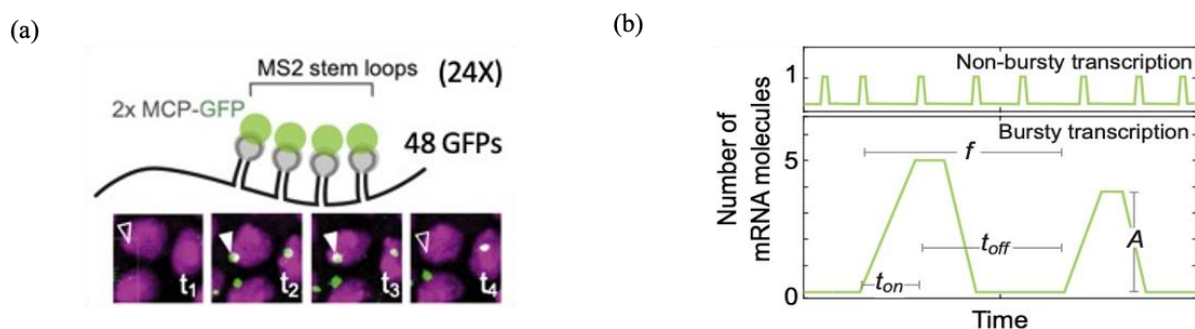


Figure 1.13 Overview of MS2-MCP system (a) MS2-MCP system to measure transcription in living organisms (Bentovim et al., 2017). This image shows the same field of view across time, exhibiting nuclei with fluorescent MS2 signal at active sites of transcription (solid arrowheads); the sites before and after transcription takes place are also indicated (open arrowhead). (b) Determining whether a gene is produced constitutively (i.e., non-bursty) or in bursts requires characterization of transcription kinetics, including the frequency (f), amplitude (A), length (t_{on}), and time between transcription (t_{off}) (below) (adapted from Bentovim et al., 2017).

OBJECTIVES OF THE STUDY

A hallmark of immune cell development is the precise timing of gene expression in different types of cells. Inappropriate gene expression has been linked to a wide range of diseases in mammals, including cancer and neurodegenerative disorders like Alzheimer's disease ([Wapinski et al., 2011](#)). Long non-coding RNAs (lncRNAs) add an important layer of control when it comes to limiting important genetic pathways to cell states and lineages. Recently, it was found that ThymoD, a lncRNA, is essential for T-cell development and lineage maintenance ([Isoda et al., 2017](#)). Bcl11b, a gene involved in subsets of T-cell acute lymphoblastic leukemia (T-ALL), is activated by ThymoD, which acts to promote T-cell commitment ([Gutierrez et al., 2011](#)). At the T-cell progenitor stage, transcription of ThymoD shifts the Bcl11b enhancer from the repressive nuclear lamina to the permissive nuclear interior. Bcl11b expression is activated when relevant promoter elements interact with the Bcl11b enhancer. Loop extrusion, mediated by cohesin and CTCF, is responsible for facilitating this interaction ([Isoda et al., 2017](#)).

Although previous research has suggested that CTCF occupancy is necessary for Bcl11b expression, experimental evidence for the precise CTCF binding region is lacking. Multiple areas across the Bcl11b locus are enriched for CTCF binding when ThymoD is transcribed. These regions have been designated as CR4, CR5, B1, B2, B3, and B4. To facilitate connections between the Bcl11b super-enhancer and promoter regions, we predicted that these CTCF binding sites would serve as anchors in regulatory loops. Our primary hypothesis is that CTCF and cohesin, are required for Bcl11b expression and T-cell development by facilitating cis-regulatory interactions inside a CTCF-dependent loop domain. The specific goals of this research are as follows:

1. To identify the functional CTCF binding sites necessary for Bcl11b expression in in vitro studies
2. To generate CTCF mutant animals using CRISPR/Cas9 technology
3. To validate CTCF occupancy across the Bcl11b locus using mutant mice
4. To determine the functional role of CTCF binding sites during T cell fate choice in vivo and in vitro

The onset of T-lineage specification is marked by the induction of T-lineage genes, and repression of key inappropriate genes. The temporal induction of gene expression requires precise regulation of transcription ([Cramer, 2019](#)). Transcription itself is a dynamic process. However, the

majority of studies focused on transcriptional regulation, initiation, and elongation along a gene stem from molecular and biochemical studies. However, these approaches are not suited to acquiring insight into the real-time dynamics of transcription. Some of these restrictions have been lifted thanks to the use of cutting-edge cell biology techniques, primarily focused on in vivo imaging, to examine transcription in both its native environment and in real-time ([Misteli, 2001](#); [Darzacq et al., 2009](#)). These recent advances have allowed us to monitor transcription in live immune cells ([Hager et al., 2009](#)).

As was previously established, the TCR is responsible for transducing TCR signals that lead to substantial upregulation of Id3, potent suppression of E protein function, and the selection of T cell destiny. In contrast, the pre-TCR complex generates weak and temporary signals that promote low levels of Id3 expression while protecting overall E protein function, ultimately leading to the selection of the T cell fate. Id3 expression is induced by TCR signals in a manner proportional to the intensity of the TCR signal ([Fahl et al., 2016](#)). Thus, we aimed to determine whether and how different distinct transcriptional bursting signatures are associated with establishing T cell fate. Do different modes of TCR signaling elevate Id3 burst frequencies, or do they modulate distinct ON and OFF times? To address these questions, we implemented the following strategy:

1. Development of Id3-MS2 transgenic animal model
2. Visualization of Id3 transcription during $\gamma\delta$ vs $\alpha\beta$ T cell fate choice
3. Determine the dynamics of Id3 transcription in response to antigenic stimulation.

Chapter 2

MATERIALS AND METHODS

2.1 Materials

2.1.1 List of chemicals, enzymes and reagents

Chemicals, Enzymes and Molecular biology reagents	Source
Ampicillin	Sigma-Aldrich
Agarose	Biopioneer
Bacterial Agar	Fisher Scientific
Bovine Serum Albumin (Fraction V)	Sigma-Aldrich
Chloroform	Fisher
Complete Protease Inhibitor Cocktail Tablets, EDTA-free	Roche
Calcium Chloride	Sigma-Aldrich
Deoxyadenosine trisphosphate (dATP)	Promega
Deoxythymidine trisphosphate (dTTP)	Promega
Deoxycytidine trisphosphate (dCTP)	Promega
Deoxyguanosine trisphosphate (dGTP)	Promega
DH5 α E.coli strain	Invitrogen
DH10 β E.coli strain	Invitrogen
Stbl3 E.coli strain	Invitrogen
Ethanol, Molecular Biology grade	Millipore Sigma
EDTA	Sigma-Aldrich
Hydrochloric Acid	Sigma-Aldrich
HEPES	Sigma-Aldrich
Tris Base	Sigma-Aldrich
Superscript III Reverse transcriptase	Invitrogen
Random Hexamers	Invitrogen
Sybr green master mix	Bimake
Restriction enzymes	New England Biolabs
Isopropanol	Sigma-Aldrich
2-Butanol	Sigma-Aldrich
Luria Bertani Broth	Fisher Scientific
Nuclease-free water	Corning

Chemicals, Enzymes and Molecular biology reagents	Source
Phusion HSII polymerase	Invitrogen
Phire green HSII polymerase	Invitrogen
Phosphatase, Calf Intestinal	New England Biolabs
T4 DNA Ligase	New England Biolabs
Sodium Chloride	Sigma-Aldrich
Sodium Hydroxide Pellets	Sigma-Aldrich
Poly-L-Lysine solution	Sigma-Aldrich
Anti-biotin microbeads	Miltenyi Biotec
PE-microbeads	Miltenyi Biotec
AmpureXP Microbeads	Beckman Coulter
Sodium Acetate	Sigma-Aldrich
Engen Spy Cas9 NLS	New England Biolabs

2.1.2 List of kits and disposables

Kits/Disposables	Source
Gloves	Kimberley Clark
Pipette Tips	Rainin, Olympus
Reaction Tubes	Eppendorf
DNA LoBind Tubes, 1.5 ml	Eppendorf
Falcon Tubes (15 ml, 50 ml)	Eppendorf
Cryogenic vials	VWR
35mm Glass bottom dish	Mat Tek corporation
Cell strainer, 70uM	Fisher Scientific
Cell culture dishes	Corning
Cell culture plates	Corning
Genejet Plasmid Miniprep Kit	Thermo Fisher
RNeasy Mini Plus Kit	Qiagen
Hyperprep Library Kit Illumina	Kapa Biosystems
Zymoclean gel DNA recovery Kit	Zymo Research
DNeasy Blood and Tissue Kit	Qiagen
Easysep CD25 positive selection Kit	Stem Cell

Lineage cell depletion kit	Miltenyi Biotec
Neon transfection kit	Invitrogen
Sterile Petri dish	Fisher Scientific
Neon transfection Kit	Invitrogen
CD19 Microbeads	Miltenyi Biotec
NEBuilder HiFi DNA Assembly Master Mix	New England Biolabs
dsDNA HS qubit kit	Life Technologies
Phire tissue direct PCR Kit	Invitrogen

2.1.3 List of cell culture media and reagents used

Cell culture media and reagents	Source
DMEM	Gibco
Opti-MEM	Gibco
IMDM	Gibco
α -MEM	Gibco
Fetal Bovine Serum (FBS)	Corning
Penicillin Strpetomycin Glutamax Solution	Gibco
β -Mercaptoethanol	Invitrogen
Polybrene	Sigma-Aldrich
Retronectin	Takara Bio
Retro-X-concentrator	Takara Bio
SCF	Peprtech
Flt3-L	Peprtech
IL-7	Peprtech
ERK signaling Inhibitor (PD98059)	Cell Signaling Technology
DPBS	Gibco

2.1.4 List of antibodies used

Antibody (Clone)	Source
Anti-mouse CD3e (145-2C11)	Invitrogen
Anti-mouse CD4 (GK1.5; RM4-5)	Biolegend
Anti-mouse CD8a (53-6.7)	Biolegend
Anti-mouse CD11b (M1/70)	Invitrogen
Anti-mouse CD19 (MB19-1)	Invitrogen
Anti-mouse B220 (RA3-6B2)	Invitrogen
Anti-mouse Gr1 (RB6-8C5)	Biolegend
Anti-mouse NK1.1(PK136)	Biolegend
Anti-mouse CD25 (PC61)	Biolegend
Anti-mouse CD44 (IM7)	Invitrogen
Anti-mouse CD45.2 (104)	Biolegend
Anti-mouse CD117 (ack45)	Biolegend
Anti-mouse TCR β (H57-597)	Biolegend
Anti-mouse TCR $\gamma\delta$ (eBioGL3)	e-Biosciences
Anti-mouse Ter119 (TER119)	Biolegend
Anti-human CD25 (PC61.5)	e-Biosciences
Annexin-V	Becton Dickinson
Anti-mouse CD3 (17A2)	Biolegend
Anti-mouse CD3e (145-2C11)	Becton Dickinson
Anti-CTCF antibody (2899)	Cell Signaling
Anti-Rabbit IgG (C15410206)	Diagenode
H3K27me3 (9733T)	Cell signaling technologies

2.1.5 List of important technical equipment used

Technical Equipment	Source
Thermomixer	Eppendorf
CFX96 C1000 Touch Thermocycler	Bio-Rad
Nanodrop ND 1000	Thermo Fisher
PCR system	Applied Biosystems
FACSAriaIII	BD Biosciences
LSRFortessa	BD Biosciences
Cell culture incubator	Thermo Scientific

Biosafety cabinet	Thermo Scientific
Bacterial incubator and shaker	Thermo Scientific
Magnetic stand tube holder	Invitrogen
880 Airy scan LSM microscope	Zeiss
Neon electroporation system	Invitrogen
Refrigerated sorval centrifuge	Thermo Scientific
NovaSeq S4 instrument (PE100)	Illumina

2.1.6 Experimental models: cell lines/ organisms/strains used

Experimental model	Source
Human: 293T	ATCC
Mouse: OP9-DL1	Schmitt and Zuniga-Pflucker, 2002
445.3 cells	Barajas-Mora et al., 2019
Mouse: SCID.adh cells	Carleton et al., 1999
C57BL/6 Mice	Jackson
CR4 KO Mice	This study
Id3-MS2 Mice	This study
Id3-MS2: Rag ^{-/-} Mice	This study

2.1.7 Recombinant DNAs used

Recombinant DNA	Source
pX330-mCherry	Addgene #98750
pX458	Addgene #48138
pET259-pUC57-24XMS2V6	Addgene #104391
MSCV-Puro-CMV-GFP	Addgene #68485
LMP-PGK-TetR EGFP	Lucas et al., 2014
pUbc-NLS-ha-stdMCP-stdGFP	Addgene #98916
pMIY-KN6- $\gamma\delta$	Coffey et al., 2014
pMIG-V β 1	Ciofani et al., 2004
MS2-GFP	Addgene #61764
pCL-Eco	Addgene #12371

2.1.8 List of analytical softwares used

Software	Source
Bowtie 1.2.2	Langmead et al., 2009
DESeq2 (v 1.24.0)	Love et. al., 2014
FlowJo	Treestar Inc.
Statistical analysis	Microsoft Excel 2013, GraphPad Prism7
Zenblck	Zeiss
Fiji	Open source
Metascape	Zhou et al., 2009
SAMtools	Li et al., 2009
Pluto data analysis	https://pluto.bio
Feature counts	Liao et al., 2014
Homer	http://homer.ucsd.edu/homer

2.1.9 CRISPR injection-mix used for generation of CR4 KO mice

Reagent	Initial conc. (in uM)	Final conc. (in uM)	Volume used (in ul)
Cas9 protein	20	1.2	3
Tracr/crRNA1	50	0.6	6
Tracr/crRNA2	50	0.6	6
IDTE	NA	NA	35

2.1.10 CRISPR injection-mix used for generation of Id3-MS2 knock-in mice

Reagent	Initial conc. (in uM)	Final conc. (in uM)	Volume used (in ul)
Cas9 protein	20	1.2	3
Tracr/crRNA	50	0.6	6
Donor fragment	500ng	-	2.5
IDTE	NA	NA	35

2.1.11 Guide RNAs used for CTCF binding site deletion

Site	Target sequences (crRNA)
CR4 CTCF-1	GTACCCTAGAGAAGTCCAGG
CR4 CTCF-2	TGGCCTTAGCACTTCTTGAT
CR5 CTCF-1	GGTGGGGACACCATGCGGCT
CR5 CTCF-2	ACCCATCCCCCATACTGTT

Bcl11b1-1	TAAACCCCTCCTTACCTGGC
Bcl11b1-2	CGTAAGTAGACTGTTTTACA
Bcl11b2-1	TGGCGGAGTCCGGGTCTCC
Bcl11b2-2	CCGGTAGCGCAAAAGAGGGT
Bcl11b3-1	CGCGCGCCTTCCCACCGTCG
Bcl11b3-2	TTGCTCGACCCGCCTGACTT
Bcl11b4-1	TTTCAATGGACTTTACAATG
Bcl11b4-2	GATGACCTCTAGAGTTCAGG

2.1.12 List of primers used for genotyping to determine the deletion mutants of CTCF binding region

Primer ID	Primer Sequence
CR4 CTCF Geno F2	TGGCCTTAGCACTTCTTGAT
CR4 CTCF Geno R	CCATTGACTCTGGCCCATGT
CR5 CTCF Geno F	CTCTTCCACCGAGGGACTTC
CR5 CTCF Geno R	CCCGACTCCAAGTCAGAAGG
Bcl11b1 Geno F	TCAGGTTGTCTTGGGTCCAC
Bcl11b1 Geno R	CCCAGTGAAGACAACCACGA
Bcl11b2 Geno F	CTTGTCTCGTACCCCTCA
Bcl11b2 Geno R	AAGTTTCTACAAAATGCTCCCCTCTC
Bcl11b3 Geno F	ATGGGAAGGAAGGCCAGAGA
Bcl11b3 Geno R	TAGGCCAAACGCAGGTACAG
Bcl11b4 Geno F	GGCCTGAGTCAGATCAAGTCT
Bcl11b4 Geno R	CAGGTCAGGGTTTGGAGGAC

2.1.13 Guide RNAs used to generate CR4 knockout mice

gRNA	Sequence
CR4 CTCF crRNA1	GTACCCTAGAGAAGTCCAGG
CR4 CTCF crRNA2	TGGCCTTAGCACTTCTTGAT

2.1.14 List of qRT-PCR primers used for quantifying desired mRNA transcripts

Primer Name	Sequence
ThymoD-qPCR-F	GGGCAGACGAAACTGACTGT
ThymoD-qPCR-R	AAGCCCTGCCTTGACTGTAA
Bcl11b-qPCR-F	ATGTTCTCTCAGGCTTCTCC
Bcl11b-qPCR-R	TGGGACACAGTCAAGTTACC
Id3-qPCR-F1	CGACCGAGGAGCCTCTTAG
Id3-qPCR-R1	GCAGGATTTCCACCTGGCTA
Arp-qPCR-F	CGACCTGGAAGTCCAACTAC
Arp-qPCR-R	ATCTGCTGCATCTGCTTG
Id3-qPCR-F2	GACTCTGGGACCCTCTCTC
Id3-qPCR-R2	ACCCAAGTTCAGTCCTTCTC

2.1.15 List of primers used for making vector constructs necessary for generating Id3-MS2 knock-in system

Primer Name	Sequence
Id33'Uex_5'HAEcoRI_F	TAAGCAGAATTCGATCCAGGTGCGAGAGGG
Id33'Uex_5'HAKpnI_R	TAAGCAGGTACCCCATTTCTCGGAAAAGCCAGT
Id33'Uex_3'HASmaI_F	TAAGCACCCGGGGGTGTCGAGAGGGTGTGGGG
Id3 3'Uex_3'HASaII_R	TAAGCAGTCGACTGAGTTTCCCAGCAAGCCTA
pMSCV_fwd_1	ATTGTACAAATAAACGCGTAGAATTCGATAAAA TAAAAG
pMSCV_rev_1	AATAATCAATGTCTGGCGCCTAGAGAAGGAGTG
CMV_fwd_1	TCTCTAGGCGCCGACATTGATTATTGACTAGTTA TTAATAGTAATC
CMV_rev_1	GTGGCGCGGCCGAGCTCTGCTTATATAGACCTC
WTStdMCPStdGFP_fwd_1	TATAAGCAGAGCTCGGCCGCGCCACCATGGG
WTStdMCPStdGFP_rev_1	AATTCTACGCGTTTATTTGTACAATTCATCCATAC CATGGGTAATACCAGC
pMSCV_rev_2	CCTCCCCTACCCGGCGCCTAGAGAAGGAGTG
PGK_fwd_2	TTCTCTAGGCGCCGGGTAGGGGAGGCGCTTTTC
PGK_rev_2	GGTGGCGCGGCCGCGAAAGGCCCGGAGATGAG
WT-StdMCPStdGFP_fwd_2	CCGGGCCTTTCGCGGCCGCGCCACCATGGG
pMSCV_rev_3	GTAATAAACCCGGGCGCCTAGAGAAGGAGTG
UbC_fwd_3	TTCTCTAGGCGCCCGGGTTTATTACAGGGACAG

UbC_rev_3	GGTGGCGCGGCCCGCCAAGTGACGATCACAGC
WT-StdMCPStdGFP_fwd_3	ATCGTCACTTGGCGGCCGCGCCACCATGGG
pMSCV_rev_4	AATAATCAATGTCGGCGCCTAGAGAAGGAGTG
CMV_fwd_4	TCTCTAGGCGCCGACATTGATTATTGACTAGTTA TTAATAGTAATC
CMV_rev_4	GAGGCGCGGCCGAGCTCTGCTTATATAGACCTC
Mut-StdMCPStdGFP_fwd_4	TATAAGCAGAGCTCGGCCGCGCCTCCATGGG
pMSCV_rev_5	CCTCCCCTACCCGGCGCCTAGAGAAGGAGTG
PGK_fwd_5	TTCTCTAGGCGCCGGGTAGGGGAGGCGCTTTTC
PGK_rev_5	GGAGGCGCGGCCGCGAAAGGCCCGGAGATGAG
Mut-StdMCPStdGFP_fwd_5	CCGGGCCTTTCGCGGCCGCGCCTCCATGGG
pMSCV_rev_6	GTAATAAACCCGGGCGCCTAGAGAAGGAGTG
UbC_fwd_6	TTCTCTAGGCGCCCGGGTTTATTACAGGGACAG
UbC_rev_6	GGAGGCGCGGCCCGCCAAGTGACGATCACAGC
Mut-StdMCPStdGFP_fwd_6	ATCGTCACTTGGCGGCCGCGCCTCCATGGG
Std GFP_Mid_Rev_All	CGGTGGTGCAGATGAACTTC
Gag_Mid_Fwd	CCTACATCGTGACCTGGGAA
CMV_Promoter_Fwd	GGGATTTCCAAGTCTCCACC
pPGK_midFwd	TCCTTCGCTTTCTGGGCTCA
pUbC_mid_Fwd	ATTCTGCGGAGGGATCTCCG
Mut-Koz-Fwd-all	GCCGCGCCTCCATGGGCCCA
Mut-Koz-Rev-all	TGGGCCCATGGAGGCGCGGC

2.1.16 List of primers used for genotyping Id3-MS2 transgenic cell lines and mice

Primer Name	Sequence
Id33'UTRExon5'HA_seqF	CTCAGAGCTGTGGGTTCGAA
Id33'UTRint5'HA_seqF	TGCATTCTTAGACACGCTG
Id3_3'HA_Fwd_cpcr	GCTCTGGGATCCATGATTGG
Id3_3'HA_Rev_cpcr	GCAGGGTTCCTTTCTCCAAG
24x MS2 v6 Fwd	TAGATCTTCCGTGTGAGGGT
24x MS2 v6 Rev	CCAATCATGGATCCCAGAGC
24x MS2 v6 mid Fwd	TGGGATCTTCCGTGTGAG
24x MS2 v6 mid Rev	CACACGGAAGATCCCAGAGC
Id3_3U_Ex5HA_flank_FP1	CTTGGCGGTCTGTTTTGAAT
Id3_3U_Ex5HA_flank_FP2	CATCTCCCGATCCAGGTG

Id3_3UEx3HA_flnk_RP1	TCTCCCATTGGTTCATCAGA
Id3_3UEx3HA_flnk_RP2	CCAAACATTAAGGGAAGTTGAGT
Id3_3UEx3HA_GenoRev	TGATTACAGAAAGTCACCTTCCT
RAG1-L	CAGTACCAAGCTTCTTGCC
RAG1-R	ATCTTGCGCGGGACACTTG
PGK-P	CAAAGCTGCTATTGGCCGC

2.2 METHODS: OBJECTIVE-I

2.2.1 Designing and cloning gRNA constructs to target CR4 and B4 CTCF binding regions

The guide RNAs targeting the CTCF binding regions were designed using online tools: ChopChop (<https://chopchop.cbu.uib.no/>) or through IDT's custom guide RNA design (https://sg.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM). Two gRNAs were selected, flanking each CTCF enrichment region (Table 2.1.11) that lost enrichment when ThymoD expression was blocked (Isoda et al., 2017). The sgRNAs were cloned into the BbsI site of PX-458 vector (<https://www.addgene.org/48138/>). Thus, this expression vector contains Cas9 along with the gRNA scaffold and GFP (Ran et al., 2013). This plasmid was later used in conjunction with the Neon electroporation system to generate deletions in Scid.adh cells.

2.2.2 Cell line cultures

Scid.adh cells and derivative cell lines were grown in IMDM supplemented with 10% FBS (Corning) and 1x PSG at 37°C and 5% CO₂. These cells were frozen in 90% FBS and 10% DMSO in cryogenic tubes and stored at -80°C for overnight before being transferred over to liquid nitrogen storage. The cell lines were stabilized with a Bcl2 transgene to improve viability when conducting serial electroporation for paired knockouts.

2.2.3 Deleting CR4 CTCF binding regions in Scid.adh cell line

Plasmid DNA from selected clones were isolated using the Thermo Scientific GeneJet Endo-free Maxiprep Kit (K0861). DNA constructs were mixed with gRNA plasmids and transferred into Scid.adh cell via electroporated using the Invitrogen Neon Transfection System with the following settings: 1300 V, 40 m/s, 1 pulse. Following electroporation, cells were recovered and maintained in growth media without PSG. After 48 hours, single cells were sorted based on the GFP expression into 96-well plates containing 200 µL of growth media and allowed to expand. Cells derived from a single mutant cell were expanded and genomic DNA was extracted, and PCR

reactions were carried out using the Phire Tissue Direct PCR kit (Thermo Fisher F170L) to determine the CTCF deletions.

2.2.4 FACS sorting

For sorting, cell cultures were washed once with PBS and resuspended in IMDM supplemented with 10% FCS and 1x PSG. Cells were strained through 70 μ M filters prior to sorting single cells into 96-well round bottom plates using a 70-micron nozzle in a BD FACS Aria II. To reduce cell condensation the outer wells were filled with PBS while the inner 60 wells contained growth media. Cells were gated first by SSC-A vs FSC-A (to eliminate debris), then by FSC-H vs FSC-A (to discriminate for singlets), and finally by Fitc-A vs FSC-A (to distinguish cells with high GFP fluorescence).

2.2.5 Genotyping of CR4 and B4 KO cells

The cells were genotyped and sequenced to confirm the deletion of respective CTCF binding regions before conducting further analysis. PCR reactions were carried out using the genomic DNA as a template by employing Phire PCR kit (according to manufacturer's protocol). Genotyping was conducted with specific primers to detect the deletion of CTCF region. In parallel distant (>800 nucleotides from the cut site) primers were also used to control for larger deletions away from the genotyping primer annealing sites. Multiple clones containing different deletions for CR4 and B4 region were considered for all analyses. Sequential deletions were generated to produce double knockout cell lines. Primer sequences that were used to genotype each of the colonies are shown in Table 2.1.12.

2.2.6 Real-time PCR

For qPCR, RNA was extracted from at least 10 million cells using the QIAGEN RNA-Easy Plus Mini Kit (74134). The cells were lysed using the lysis buffer supplemented with 2-Mercaptoethanol and frozen at -80°C prior to RNA isolation. RNAs were isolated according to the manufacturer protocol. cDNAs were generated reverse transcription using 0.5 - 1 ug of RNA using the Qiagen QuantiTect Reverse Transcription Kit (205311). Real time PCR was performed to monitor the expression of ThymoD, Bcl11b, and Vrk1 using the primers that were described previously ([Isoda et al., 2017](#)). Real time PCR was conducted using the SYBR Green PCR master mix in the Bio-Rad CFX96 C1000 Touch Thermocycler. The data was analyzed using Microsoft Excel and Prism.

2.2.7 Generation of CR4 KO transgenic mice

CR4 guide designs were optimized in cell lines to improve efficiency and reduce the frequency of off-target mutations. These crRNAs were resuspended with nuclease free IDTE at pH 8.0. For generation of transgenic mice, CRISPR mix containing Cas9 protein, crRNA and tracrRNA was prepared (Table 2.1.13). For this, first, tracrRNA and crRNA were combined and incubated at 95°C for 5 minutes. After cooling to room temperature for 5-10 minutes, this mix was diluted 1:10. Then 0.6 μ m of tracrRNA/crRNA was mixed with 1.2 μ m Cas9 protein in total volume of 50 μ l and incubated for 5 minutes at room temperature. Following this incubation, it was centrifuged at 10,000 rpm for 1 minute and the supernatant was transferred to a new tube containing 40-50 μ l of final mixture prior to being given to the UC San Diego transgenic core for blastocyte injection into a C57BL/6n background.

2.2.8 Genotyping and Sequencing of CR4 Knockout Mice

After receiving transgenic pups from the transgenic core, we isolated genomic DNA from ear clips using the Phire Tissue Direct PCR Master Mix kit from Thermo Fisher. Genotyping was conducted with the primers described in the table 2.1.12. More than half of the mice received back from both injections were characterized by heterozygous or homozygous CBS deletions. Sequencing was conducted by IDT using the forward and reverse genotyping primers. The genomic region surrounding each mutation is shown below, with the genotyping primers and gRNAs in red and the deleted sequence crossed out. Upon confirmation of the knockout by sequencing, these mice were backcrossed to wildtype C57BL/6 mice. The backcrossed mice were sequenced again to confirm that the correct mutation was preserved after breeding.

2.2.9 Co-culture of bone marrow progenitors and stromal cells

OP9-DL1 stromal cells were obtained from Dr. Zúñiga-Pflücker ([Schmitt and Zuniga-Pflucker, 2002](#)) and cultured in MEM α complete medium supplemented with 20% (v/v) FBS, β -mercaptoethanol (1X) Penicillin- Streptomycin-Glutamine (1X). The cells were split upon confluency using Trypsin (0.25%) and cultured in fresh complete medium. The cells were passaged every 2 days to avoid over-confluency. Multipotent progenitors were isolated from bone marrow of 6–8-week-old male mice. During this process, progenitors from the bone marrow cell suspension were enriched by depleting mature lineage cells using biotinylated antibodies against CD3 ϵ , CD19, B220, NK1.1, CD11b, CD11c, Gr1 and Ter119 by negative selection using MACS LS magnetic columns. Progenitors were co-cultured in the presence of OP9-DLL1 using alpha-MEM in the presence of IL-7 and Flt3L by varying their concentrations: 10 ng/ml each from day0

– day7, 5 ng/ml each from day7- day10, and 1 ng/ml each from day10 – day18 in OP9 medium (α -MEM, 20% FBS, 1X Penicillin-Streptomycin-Glutamine and β -mercaptoethanol). All in vitro cultures were maintained at 37°C in the presence of 5% CO₂ environment.

2.2.10 Cut&Run strategy

For ex vivo analysis, DN2 cells obtained from OP9-DLL1 co-cultures on day 7 were stained using the DN panel described under “Flow Cytometry” and sorted to obtain 100,000-150,000 cells. For in vivo analysis, thymocytes were MACS depleted of CD4⁺CD8⁺ double positives as described previously (Zhou et al., 2019). Following depletion, the purity of the DN population was checked by cytometry and the cells were directly used in analysis. Cut&Run analysis was carried out as described previously using pA-MNase (provided by Dr. Ananda Goldrath’s lab) (Skene et al., 2017). Spike-in DNA was provided by Dr. Henikoff, Fred Hutchinson Cancer Center. DNA fragments were enriched using anti-CTCF (Cell Signaling 2899), Anti-Rabbit IgG (Diagenode C15410206) or H3K27me3 (Cell signaling technologies 9733T) and purified by phenol-chloroform extraction. Libraries were prepared using the Kapa Hyper Prep Kit (KR0961) and Kapa Unique-Dual Indexed Adapters (08861919702) according to the manufacturer protocol. The libraries were pooled and purified using AMPure XP beads (A63882) and subject to deep sequencing using NovaSeq S4 instrument (PE100).

2.2.11 Analysis for Cut&Run

Raw DNA sequences were mapped to the mm10 reference genome using Bowtie2 with the following parameters: “very-sensitive-local”, “no-mixed”, “no-unal”, “no-discordant”, “phred33”, and specifying 10 minimum and 700 maximum fragment lengths. Quality control was conducted with FastQC. The bam files were processed by sorting using SAM tools and removing duplicates with Picard-tools. The sorted bam files with duplicates were removed and then processed with Homer: Tag directories were created with makeTagDirectory using -tbp 1, and then UCSC compatible bedGraph files were generated with makeUCSCfile.

2.2.12 RNA-Seq

DN2 cells were generated from bone marrow progenitors following differentiation under T cell conditions for 6-8 days on OP9-DLL1. RNA was isolated from DN2 sorted populations using RNeasy miniplus kit (Qiagen) and purified by phenol-chloroform extraction followed by ethanol precipitation. Ribo-depleted RNAs were obtained, and libraries were prepared for paired-end sequencing by the UC San Diego Institute for Genomic Medicine. Sample were sequenced using

NovaSeq S4 instrument (PE100) with minimum depth of 40 million reads per sample.

2.2.13 Analysis for RNA-seq

Raw sequence reads were trimmed with cutadapt (v1.16 38), aligned using STAR (v 2.5.2a 39), and then exonic read counts were quantified with feature Counts using subread package (v1.6.2 40). Differential gene expression analysis was performed utilizing the DESeq2 R package and then a volcano plot was generated using the Pluto data analysis software (<https://pluto.bio>). Gene ontology analysis was conducted with Metascape. Venn diagrams were generated using an online tool (<https://bioinformatics.psb.ugent.be/webtools/Venn/>).

2.2.13 Flow cytometry

For phenotypic analysis of CR4 KO animals, thymocyte cells were isolated from 5–7-week-old mice and T cell populations were characterized by flow cytometry analysis. In parallel, ex vivo experiments, bone marrow progenitors were obtained from the CR4 KO and cultured in the presence of OP9 stromal cells expressing notch ligands. Cells were harvested at various time points and T cell populations were characterized by staining with the following antibodies obtained from BioLegend and e-Biosciences at a 1:200 dilution unless otherwise noted: Live/Dead near-IR (Invitrogen), c-Kit (PE), CD25 (PE-Cy7), CD44 (BV711), CD4 (PerCP-5.5), CD8a (APC), Thy1.2 (PE), B220 (BV711), NK1.1 (APC), TCR β (PE-Cy7), $\gamma\delta$ TCR (BV421), CD3 (PE), CD19 (FitC), B220 (FitC), CD11b (FitC), CD11c (FitC), GR1 (FitC), NK1.1 (FitC), TER119 (FitC). Cells were analyzed using BD LSR Fortessa X-20 instrument, whereas cells were sorted using BD FACS Aria II Fusion instrument. Compensation was calculated using BD CompBeads (552845). Data was analyzed using FlowJo software.

2.2.14 Statistical analysis

Data sets were analyzed using appropriate statistical tools such as Microsoft Excel (2013) SigmaPlot (v12.3) and GraphPad Prism7. Data are expressed as means $\pm$ Standard Deviation (SD). Statistical significance was determined by unpaired two sample t-test wherein a value of $p \leq 0.05$ was considered statistically significant. The statistical significance was assessed as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ns=not significant.

2.3 METHODS: OBJECTIVE-II

2.3.1 Vectors and constructs cloning

All PCR amplifications were carried out by PCR using Phusion polymerase in the presence of target specific primers carrying restriction sites (for sticky end cloning). The vectors were linearized using appropriate restriction enzymes for cloning the inserts. The amplified inserts and linearized vectors were purified with gel extraction kits as per the manufacturers protocol (Zymo Research). Purified fragments were ligated using T4 DNA ligase and transformed into competent DH10 beta strain of bacteria. The colonies obtained were screened by colony PCR using Phire tissue direct PCR master mix (Thermo Scientific). The clones were confirmed by restriction digestion as well as Sanger's sequencing. Large scale DNA samples were prepared using Plasmid Maxiprep columns (Qiagen) or cesium chloride density-gradient centrifugation method and resuspended in nuclease-free water.

2.3.2 Designing and cloning gRNA constructs to insert MS2 loops into Id3 3'UTR

The guide RNAs targeting the CTCF binding regions were designed using online tools: ChopChop (<https://chopchop.cbu.uib.no/>) or through IDT's custom guide RNA design (https://sg.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM). Based on score, gRNAs having highest on target activity and lowest off target activity were chosen for cloning into pX330-mCherry. To construct a repair template, 1000 bps homology arms flanking the gRNA target site were amplified from genomic DNA with primers using appropriate restriction sites. These homology arms were cloned into pET259-pUC57-24XMS2V6 (Addgene plasmid # 104391).

2.3.3 Generating Id3-MS2 Pro-B cell line

445.3 cells were cultured in RPMI 10% FBS, PSG at 37°C in 5% CO₂. MS2 knock-in was conducted using a Neon transfection system (Invitrogen). During this, 2.5×10^6 of 445.3 cells were harvested in a 15 mL tube and washed with PBS. Cells were suspended in 100 µl of buffer R. sgRNA and repair template were added in cell suspension. Cells were transfected and then transferred to 6 well plates and fed with 2.5 mL of RPMI-1640. On day two, single cells were sorted into 96 well plates based on mCherry signal and cultured for 7-10 days. Genotyping was performed for the cells isolated from each well to detect the presence of MS2 insertion.

2.3.4 Id3 MS2 clone screening strategy

Two sets of primers, which are internal to the endogenous Id3 locus and located outside of the homology arms were used to identify MS2 insertions. Generation of a single PCR product of 2317 bps indicates that MS2 is integrated in both alleles (homozygous). On the other hand, amplification of 573 bp product indicates wild type or absence of MS2 insertions. Appearance of 573 and 2317 bp PCR products indicate that MS2 is integrated in only one allele.

2.3.5 Generation of Id3-MS2 mice

A sgRNA target site targeting Id3 3'UTR was designed using online tools. The target specificity of sgRNA and repair template were verified by transfection of sgRNAs into a pro-B cell line. For generation of transgenic mice, a mixture of sgRNA, Cas9 protein and repair template were injected in mouse zygotes using standard procedures at UCSD transgenic core facility. All mice were bred in specific pathogen-free conditions in accordance with the Institutional Animal Care and Use Committee (IACUC) of the University of California, San Diego.

2.3.6 Generation of Rag1^{-/-}: Id3-MS2 mice

To avoid the heterogeneity of T cell populations from Id3 MS2 transgenic mice, we crossed Id3-MS2 mice in a Rag1^{-/-} background. Id3-MS2; Rag1^{-/-} homozygous mice were generated by backcrossing crossing the heterozygotes.

2.3.7 Mouse Pro-B cell culture

CD19⁺ B cells were purified from bone marrow suspensions by positive selection and cultured for 5 days in RPMI medium containing 10% fetal bovine serum, 1X Penicillin-Streptomycin-Glutamine, 1X β -mercaptoethanol and 10 ng/ml of both IL-7 and SCF.

2.3.8 Flow cytometry

For flow cytometry analysis, single cell suspensions of OP9 DL1 cultured bone marrow cells were prepared and analyzed as follows. Cells were stained with CD11b (M1/70), CD19 (MB19-1), B220 (RA3-6B2), Gr1 (RB6-8C5), NK1.1(PK136) and Ter119 (TER119) conjugated with Fitc to remove the lineage positive cells. T cell subsets were characterized based on the staining with CD25 (PC61)-PECy7, CD44 (IM7)-BV711, CD45.2 (104)-BV510, CD117 (ack45)-PE, TCR β

(H57-597)-PECy7, TCR- $\gamma\delta$ (eBioGL3)BV-421, CD3e (145-2C11)-BV-650, CD4 (GK1.5; RM4-5)-APC, CD8a (53-6.7)-BV-786 and huCD25 (PC61.5)-PE antibodies (Becton Dickinson and e-Biosciences). Data were collected on a LSRII (BD Biosciences) and analyzed with FlowJo software (TreeStar).

2.3.9 Viral transduction

HEK293T cells were used to generate retroviral supernatants. HEK293T cells were cultured in DMEM complete medium (supplemented with 10% (v/v) FBS, β -mercaptoethanol (50 μ M), Penicillin-Streptomycin (10 U/mL) and 12.5mM HEPES). To generate retroviral supernatant pMSCV-MCP GFP vectors were initially transfected into HEK293T packaging cells using CaPO4 transfection method. Following overnight, cells were fed with fresh media. Viral supernatants were collected at 48 hr post-transfection and concentrated using Retro-X-concentrator. For retroviral transductions, non-tissue culture plates were coated with 50 μ g/ml RetroNectin (Takara bio) at 4°C overnight. After removal of excess RetroNectin, Pro-B cells or bone marrow derived T cells were transduced with viral supernatant by spin infection in the presence of polybrene. The cells were replenished with fresh culture medium, supplemented with cytokines and plated on OP9-DL1 stroma. After 48 hours of spin infection, GFP expression was measured, and live cell imaging was performed.

2.3.10 Generation of MCP-GFP expression constructs for signal to noise ratio optimization

To optimize signal to noise ratio of MCP-GFP mediated fluorescence, we constructed various vectors following two distinct strategies: (1) use of divergent promoters CMV, PGK or UbC to drive the expression of MCP-EGFP fusion protein, (2) addition of single nucleotide mutations in the kozak sequence to control the expression levels for each of the promoters. MSCV-Puro-CMV-GFP (Addgene# 68485) vector was used as a template for amplification CMV promoter. Whereas PGK promoter was amplified from LMP-PGK-TetR EGFP plasmid and UbC promoter was amplified from pUbC-NLS-ha-stdMCP-stdGFP (Addgene# 98916). Kozak sequence was mutated by overlap extension PCR. The fragments of divergent promoters along with WT and mutant kozak sequences were cloned into pMSCV retroviral backbone by Gibson assembly.

2.3.11 Anti CD3 ϵ / ERK inhibitor treatment

Anti-CD3 ϵ (clone 145-2C11) were added to OP9-DL1 cocultures in suspension at the concentrations of 15 μ g/ml. For live cell imaging experiments, bone marrow derived Id3-MS2-

RAG-1^{-/-} DN2/3 cells were harvested on day10 of coculture with OP9-DL1. Equal volumes of DMSO or ERK inhibitor (serially diluted in DMSO) were added to the co-culture medium. Cells were preincubated with 30μM of PD98059 for 2 h before the addition of 15μg/ml anti- CD3ε mAb. Cells were treated with Anti- CD3ε just before conducting live cell imaging.

2.3.12 Generation of TCR plasmids

pMIY-KN6-γδ (Coffey et al., 2014) and pMIG-Vβ1 (Ciofani et al., 2004) plasmids were obtained from Dr. David L. Wiest (Fox Chase Cancer Center) and Dr. Zúñiga-Pflücker (University of Toronto), respectively. γδTCR and Vβ1TCR sequences were PCR amplified and cloned into a pMSCV retroviral construct carrying human CD25, a selection marker.

2.3.13 MHC tetramer treatment

Soluble MHC tetramers were obtained from Prof. Erin Adams (University of Chicago). Cells transduced with retrovirus encoding γδTCR were treated with soluble MHC tetramers having various affinities towards γδTCR (Adams et al., 2008; Crowley et al., 2000). The stock concentration of MHC tetramers are 1μg/ml. These tetramers were added to the cells with a dilution of 1:100 prior to live cell imaging. Cells were treated

2.3.14 Live cell imaging

Images were captured using Zenblack software with a cooled CCD camera attached to a fluorescence microscope (Zeiss, 880 Airyscan laser scanning confocal microscope) using 37°C incubation and 5% CO₂. Imaging was performed using 488 nm excitation, zoom of 1.7 at 932x932, pixel size of 0.09 microns for all data. Z stacks were acquired spanning 5-10 microns at 0.4micron intervals. Imaging of Z stacks occurred every 300 seconds for 10-12 hours. The maximum intensity projections were performed and used for image analysis.

2.3.15 Image analysis

Raw images were processed using airyscan processing by Zenblack software. The maximum intensity projections were obtained. Bursting frequencies were calculated by using Image J software. Total number of cells and the number of cells that show bursting at different time intervals (1-5 hours) were calculated. The bursting frequencies were plotted by using Microsoft excel and Prism.

Chapter 3

RESULTS

CHAPTER III: RESULTS (Objective-I)

3.1.1 Binding of CTCF influence Bcl11b expression over vast genomic distances

ThymoD transcripts have been shown to orchestrate chromatin compartmentalization and promote folding by CTCF and cohesin-dependent looping, facilitating interaction between the Bcl11b promoter and enhancer. In the absence of non-coding transcription, there is a loss of cis-regulatory interactions and a subsequent reduction in Bcl11b expression, resulting in blockage of T cell development at DN2 stage. CTCF ChIP-seq analysis revealed a loss of CTCF enrichment at five distinct regions within the Bcl11b locus (**Fig 3.1.1 a and b**). However, the precise CTCF binding sites (CBS) that stabilize the Bcl11b promoter-enhancer interactions remain to be determined. To probe the CTCF control of genome organization and their effect on Bcl11b expression, we implemented CRISPR/Cas9 technology to disrupt these CBS located within the Bcl11b locus using Scid.adh cells, a cell line that represents the DN3 stage of T cell development (**Fig 3.1.1c**). We employed CrisprCas9 to design a highly specific sgRNA to target the CBS and constructed retroviral vectors for delivery of CRISPR/Cas9 and donor templates carrying fluorescent markers, mCherry and GFP, respectively. To target the CBS within the Bcl11b loci, Scid.adh cells were electroporated with vectors carrying CRISPR/Cas9 and Id3 homology arms under optimal conditions. Three days after electroporation, single transduced cells were FACS sorted based on fluorescent marker, expanded, and subjected to genotyping to identify the homozygous deletion mutant clones (**Fig 3.1.1d**). Based on PCR analysis and sequencing of the targeted loci, we have selected homozygous deletion clones of CR4, and CR5 as well as B2, B3, B4, and B5 Scid.adh cells. To test, if these deletions could influence Bcl11b expression, we carried out qPCR analysis. Interestingly, Bcl11b expression was significantly reduced in the CR4 mutant, while a moderate reduction was observed in CR5, B2, B3, B4, and B5 clones. We also tested whether deletion of any of these CTCF sites had caused a change in the expression of Vrk1, located adjacent to the Bcl11b locus. However, the expression levels of Vrk1 remained unchanged in all CTCF deletion mutant clones (**Fig 3.1.1e-f**). These results together suggested that CBSs are important for the formation of single loop domains that facilitate proximity promoter-enhancer interactions of Bcl11b that drive its expression. Thus, it is reasonable to speculate that CTCF promote formation of the cis-regulatory interactome, and acts as a chromatin organizer and establishes T cell identity.

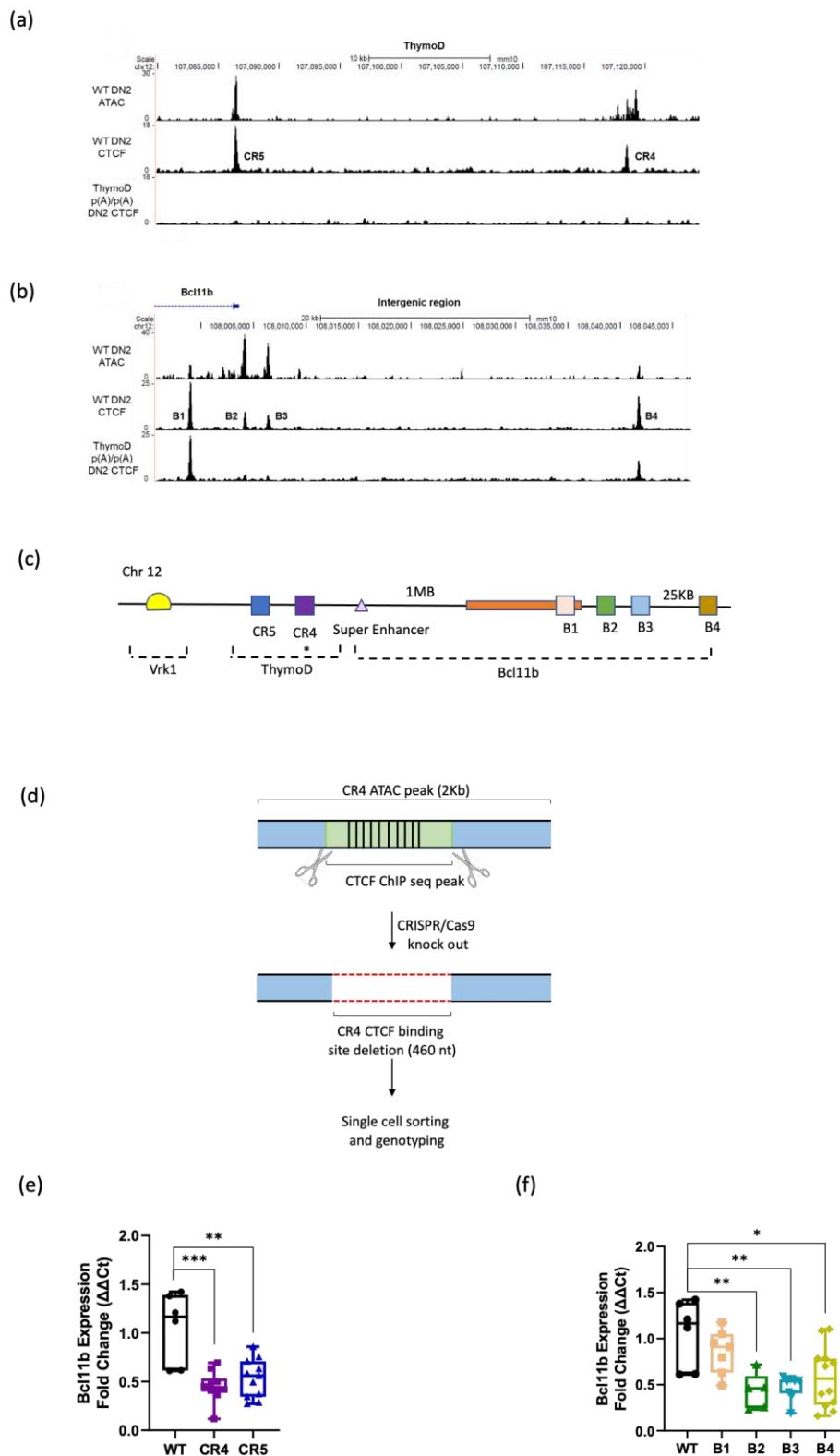


Figure 3.1.1 Single CTCF binding sites specifically influence *Bcl11b* expression over vast genomic distances. (a) Schematic representation of CTCF binding sites across *Bcl11b* locus. (b) ChIP- Seq analysis indicating CR4 and CR5 CTCF binding peaks in WT and ThymoD p(A)/p(A) mice. (c) ChIP seq peaks representing enrichment of CTCF binding at B1, B2, B3, B4 sites. (d) qPCR analysis for *Bcl11b* expression in CR4 and CR5 mutant Scid.adh cell lines. (e) qPCR analysis for *Bcl11b* expression in B1, B2, B3 and B4 mutant Scid.adh cell lines. Data is represented with WT Scid.adh cells as control. Data is shown as mean $\pm$ SD (* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$).

3.1.2 Generation of the CR4 deletion mutant mouse line

Next, we sought to determine the functional significance of CR4 CBS, as this site had the highest impact on *Bcl11b* expression by generating CR4 knockout (CR4KO) animals. For this, we designed specific crRNAs to use in conjunction with tracrRNA and the Cas9 protein. All guides were designed as described in materials and methods, and care was taken to avoid any sequence with known binding by other transcription factors or markers for promoters or enhancers while encapsulating all predicted CTCF sites within the flanked guide RNAs. Most of the donor vectors were designed to carry homology arms less than 1 kb to facilitate vector construction, and PCR screening. After testing these guides in Scid.adh cells, the crRNA, tracrRNA, and Cas9 protein were injected into fertilized zygotes and implanted into pseudo-pregnant females to generate CR4KO lines at the UC San Diego Transgenic Core Facility following previously approved animal protocol (**Fig 3.1.2**). Founder mice were genotyped using PCR and heterozygous mice were selected for initial testing and backcrossing. After testing the first litter for differences in immune compartments, we selected the lines with the smallest deletions containing all predicted CTCF sites to move forward for all experiments. The genotyping and sequencing of target loci confirmed the complete deletion of the CR4 CTCF region.

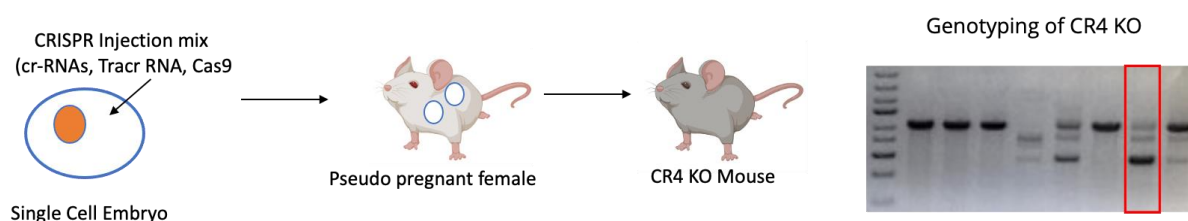


Figure 3.1.2 Strategy for the generation of CR4 deletion mutant mouse line: crRNA, tracrRNA and Cas9 protein were injected into fertilized zygotes and implanted into pseudo-pregnant females to generate CR4KO lines. Founder mice were genotyped using PCR and heterozygous mice were selected for initial testing and backcrossing.

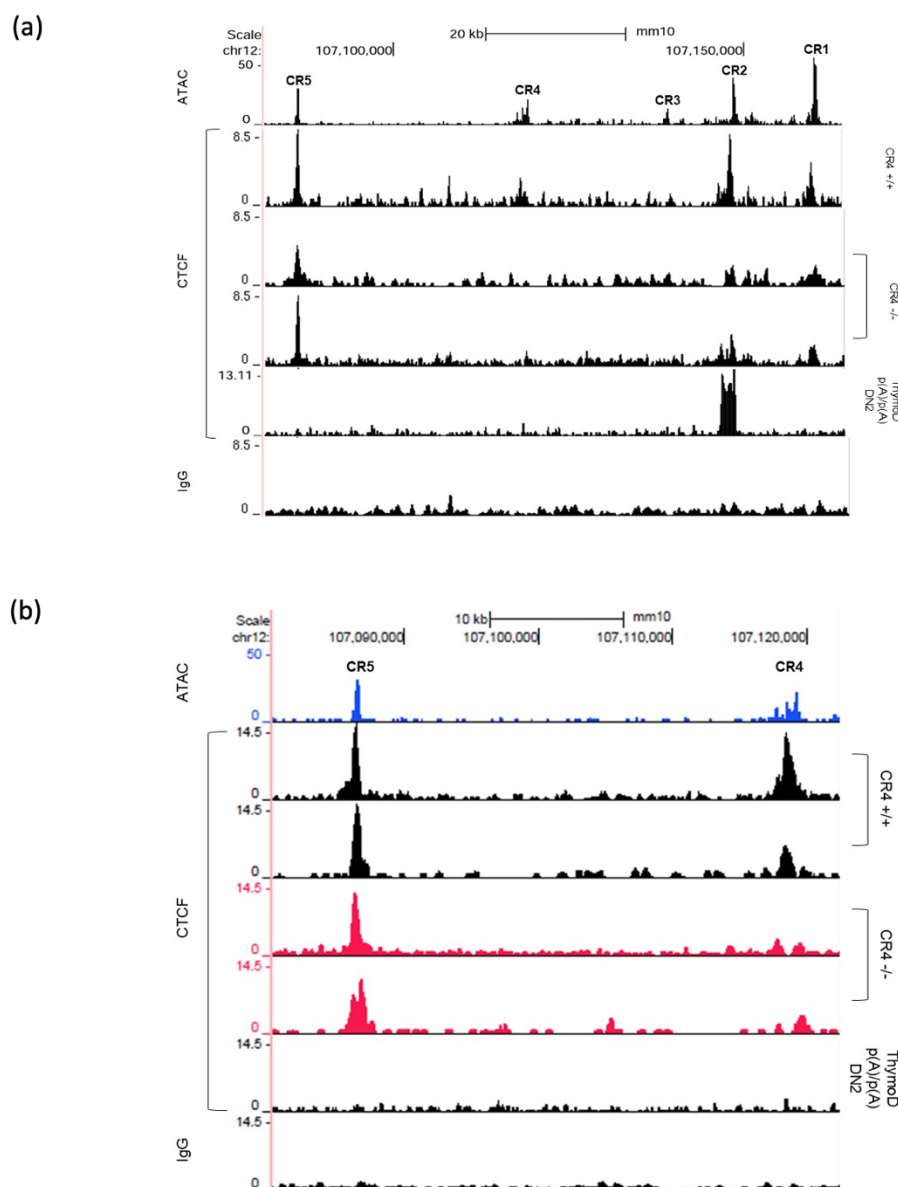
3.1.3 Disruption of CR4 greatly abolishes its CTCF binding

Next, to determine the effect of deletion of the CR4 region on CTCF binding, we have implemented CTCF Cut & Run (Cleavage Under Targets and Release Using Nuclease) analysis ([Schmid et al 2004](#), [Skene and Henikoff, 2017](#)), which allows genome-scale profiling of chromatin associated proteins using a limited number of cells. During these experiments, we first sorted bone marrow progenitors and cultured on the OP9-DLL1 system as previously described ([Schmitt and](#)

[Zuniga-Pflucker, 2002](#)). On day 8, the DN2 population ($CD25^+CD44^+$) was FACS purified. Then, nuclei were isolated and incubated with anti-CTCF antibody along with recombinant Protein A-micrococcal nuclease, thus targeting MNase to antibody bound proteins in the presence of calcium. Nuclease activity was chelated, and cleaved DNA fragments were purified. Libraries were generated by ligating with adapter sequences (Illumina) and subjected to high-throughput sequencing. The raw sequences were filtered and CTCF enrichment was determined. Surprisingly, the number of aligned reads as well as the intensity of CTCF peaks were not satisfactory (**Figure 3.1.3a**). This may be due to the low number of cells generated under in vitro conditions. To overcome this, we resorted to purifying the DN population from the CR4KO thymocytes following positive lineage depletion of double positive ($CD4^+CD8^+$) or single positive ($CD4^+$ or $CD8^+$) T cells. Then, the purified double negative (DN) populations were used for CTCF CUT&RUN analysis. These studies have revealed that genome wide enrichment for CTCF binding was similar between wildtype and CR4KO DN populations. Most of the regions enriched for CTCF were positioned in intergenic regions. However, we observed a substantial loss of CTCF enrichment spanning the CR4 region in CR4KO cells when compared to wild type cells (**Figure 3.1.3b**). We note that there were no significant changes in the CTCF enrichment in the CR5 region in both wild type and CR4KO cells. These studies together indicate that CTCF association at CR4 is T cell specific and may potentially involve the formation of the Bcl11b loop domain.

3.1.4 CR4 CTCF binding is essential for T-lineage gene expression program

We then checked for the changes in Bcl11b gene expression pattern associated with the disruption of the CR4 region using real-time PCR measurements. During these experiments, thymocyte progenitors (DN2; $CD25^+CD45^+$) were FACS sorted from wildtype or CR4KO mice. Interestingly, Bcl11b expression was significantly lower in CR4KO progenitors compared to wild type cells, though not as severely as in ThymoD p(A)p(A) mutant thymic progenitors. On the other hand, similar expression levels of Vrk1, which is located adjacent to Bcl11b, were observed in wild type and CR4KO cells (**Fig.3.1.4a**). These results demonstrated that the CR4 CTCF binding site, located 1Mb away from the Bcl11b TSS, is important for efficient Bcl11b expression. Given the role of the CR4 binding region in Bcl11b expression, we posited that the CR4 binding region could play at least in part an important role in genome organization and developmental gene expression programs that were observed in the ThymoD p(A)p(A) mice. To address this, we conducted RNA-Seq using DN2 cells isolated from CR4KO and wild-type mice. Consistent with the expression analysis, Bcl11b levels found to be lower in DN2 cells of CR4 KO mice. Interestingly, gene-ontology analysis revealed downregulation of a number of genes involved in



Bone marrow progenitors were differentiated on OP9-DLL1 stromal cells for 6-8 days and then FACS sorted to isolate DN2 cells from CR4 +/+ and -/- populations. These cells were then used for Cut&Run analysis. The tracks from this study show CTCF occupancy across a genomic region containing the first five CR ATAC-seq peaks, with the negative control (Rabbit IgG) included at the bottom. Other tracks included show WT DN2 ATAC-seq and ThymoD p(A)/p(A) CTCF ChIP-seq are from a previous publication (b) Double negative 1-4 T cells were isolated from CR4 +/+

and +/- thymic populations by MACS sorting and used for Cut&Run analysis. The tracks from this study show CTCF occupancy across a genomic region containing CR4 and CR5 ATAC-seq peaks, with the negative control (Rabbit IgG) included at the bottom. Other tracks included show WT DN2 ATAC-seq and ThymoD p(A)/p(A) CTCF ChIP-Seq are from a previous publication.

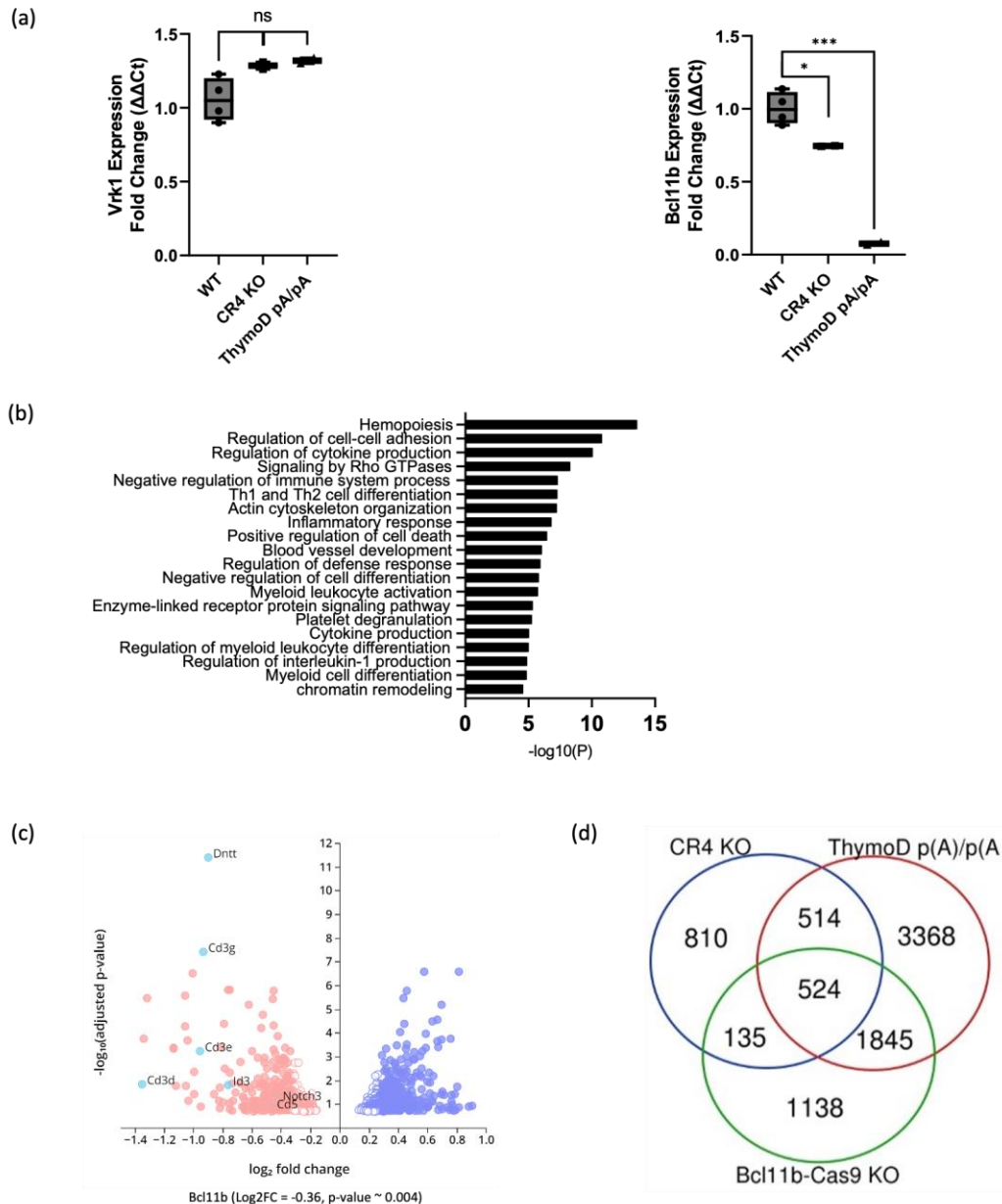


Figure 3.1.4 CR4 CTCF binding instigates a T-lineage gene expression profile while suppressing alternative immune gene programs. Bone marrow progenitors were isolated from CR4 knockout (KO) and wildtype (WT) mice and cultured on OP9 stromal cells for 8 days. qPCR and RNA-seq were conducted on these DN cells to determine the role of CR4 on early thymocyte development and expression. (a) Depicted here is the fold change in *Vrk1* (left) and *Bcl11b* (right) expression in ex vivo-cultured DN2 cells isolated from CR4 $-/-$, CR4 $+/+$ and ThymoD p(A)/p(A) mice. qPCR analysis was conducted to determine

the impact of these deletions on downstream expression. Gene expression was first normalized to the housekeeping gene ARP, and then to wildtype DN2 expression. The graph and paired sample t-tests were prepared with PRISM. (b) Gene ontology analysis was derived from RNA-seq analysis of sorted DN2 cells from four wildtype and four CR4 knockout pooled littermates in two separate experiments. Deseq2 was used for comparison and calculations of p-value and fold change. Transcripts that were differentially regulated with a p-value < 0.05 and fold change less than 0.8 or greater than 1.2 were analyzed with Metascape and the graph was prepared using Prism. (c) Volcano plot depicting differential gene expression analysis between CR4 knockout and wildtype pooled littermates. Two pseudogenes that were extreme data points were excluded from this figure. Figure generated using Pluto (<https://pluto.bio>). (d) Previously published RNA-seq data from ThymoD p(A)/p(A) and +/+ ex vivo-derived and Bcl11b -/- and +/+ in vitro-derived DN3 progenitors were analyzed using Deseq2.^{3,20} Venn diagrams were generated to look at the intersection of differentially regulated genes from ThymoD, Bcl11b and CR4 mutants using an online platform (<https://bioinformatics.psb.ugent.be/webtools/Venn/>).

3.1.5 Early T cell developmental program is mildly affected in CR4 knockout mice

We then wanted to know if CR4KO mice would have immune compartment defects like ThymoD mutant mice. Phenotypic analysis of thymocytes indicated that CR4KO mice displayed a milder impact on development than a full ThymoD knockout (**Fig 3.1.5a**). Particularly, the percentage of DN2 population in CR4KO mice was significantly higher when compared to wild type cells (**Fig 3.1.5b**). We propose that loss of CTCF binding at the CR4 region may have altered the genes that are critical for the developmental transition from DN2 to DN3. Consistently, the DN3 population is lower in CR4KO mice as compared to wild type. Nevertheless, there was no change at the DN4 cell stage in either of the mice (**Figure 3.1.5c**). These findings suggest that CR4 is important for early thymic T cell development.

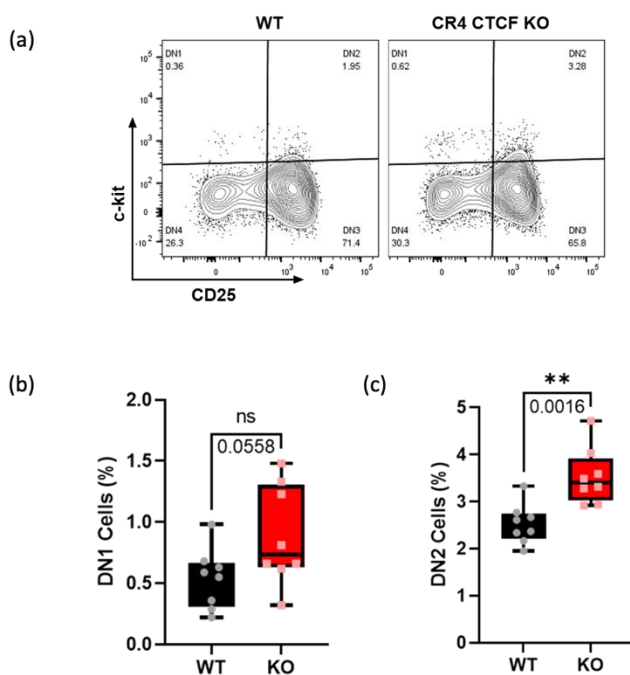


Figure 3.1.5 Mild differences in early T cell progenitors in CR4 knockout mice (a) Representative cytometry plots from littermate mice produced using FlowJo. (b) and (c) A summary of the cytometry analysis from the DN1 and DN2 populations respectively in wildtype (WT, n=8) and CR4 knockout (KO, n=8) mice. P-values are displayed for DN1 and DN2 populations below the asterisks. Littermates were used in three separate experiments and then pooled for this analysis. Cells were gated first on FSC/SSC, singlets, live-dead, CD4/8 double negative and lineage negative populations. The graph and paired sample t-tests were prepared with PRISM (p < 0.05).

3.1.6 CR4KO bone marrow progenitors are impaired to differentiate into early T cells

To rigorously test the T cell developmental capacity of CR4KO multipotent progenitors, we isolated bone marrow (BM) progenitors ($\text{Lin}^- \text{cKit}^+$) from CR4KO mutant mice and co-cultured them in the presence of OP9-DL1 or OP9-DL4 for 18-22 days under lymphoid cytokines (IL-7, Flt3L). Cells were harvested at various time points, and T cell developmental progression was monitored (**Fig 3.1.6a**). These studies have revealed that the developmental arrest of the CR4KO cells at the DN2 stage was more pronounced when compared to wild type progenitors. This could be because the progenitors are relatively homogeneous and differentiated under similar conditions in a synchronous manner. Moreover, dramatic differences were observed in all compartments that define early T cell development (**Fig 3.1.6 b-c**) like previous studies involving ThymoD p(A)/p(A) mice ([Isoda et al., 2017](#)). These findings were reproducible and were repeated in three separate experiments, using littermates derived from three separate litters. These results indicate that CTCF binding in the Bcl11b intergenic region promotes a T cell gene program to support early T cell development.

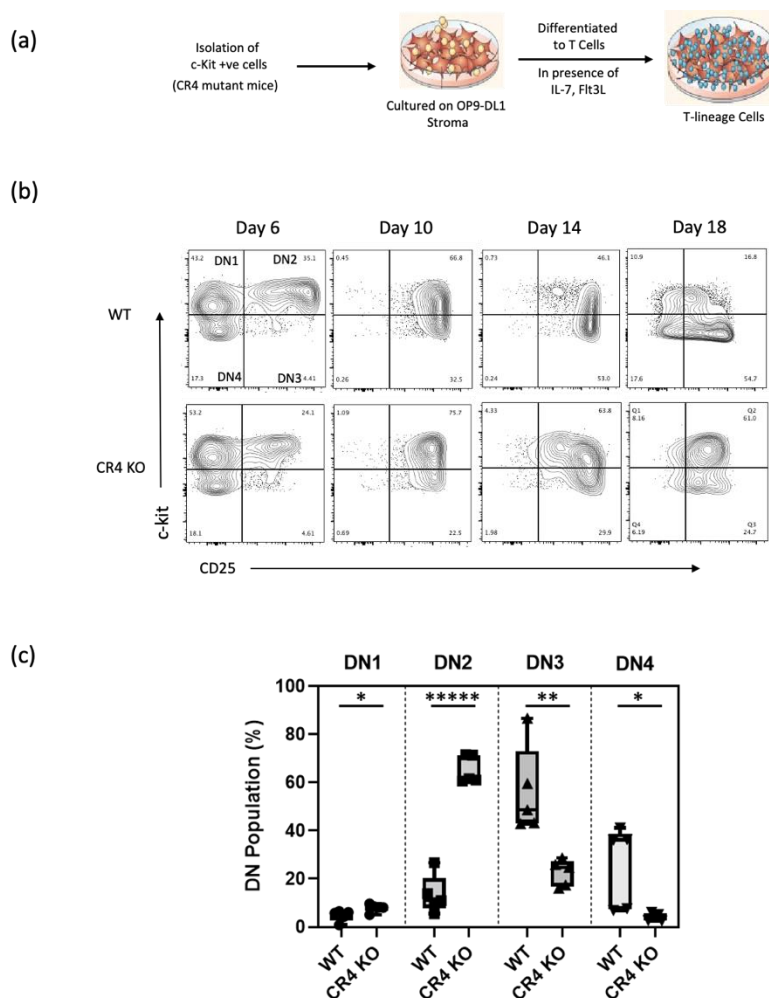


Figure 3.1.6 Dramatic differences in early T cell progenitor differentiation in CR4 CBS knockout mice ex-vivo. Bone marrow progenitors were isolated from CR4 knockout (KO) and wildtype (WT) mice and cultured on OP9 stromal cells for 18 days. Cytometry was conducted on these DN cells to determine the role of CR4 on early thymocyte development and expression. (a) Experimental strategy for in vitro differentiation of multipotent progenitors into early-T cells. (b) Representative cytometry plots from littermate mice over an 18-day time course after bone marrow isolation. (c) A comparison of DN1-4 cells derived from bone marrow progenitors cultured for 18 days. Littermates were used in all comparisons in two independent experiments. Cells were gated based on FSC/SSC, singlets, live-dead, lineage and then c-kit and cd25. The graph and paired sample t-tests were prepared with PRISM ($p < 0.05$).

CHAPTER III: RESULTS- OBJECTIVE-II

3.2.1 Generation of Id3-MS2 transgenic pro-B cell line

To visualize the transcriptional dynamics of Id3, we employed the MS2 system ([Beach et al 1999](#); [Bertrand et al 1998](#)). We have integrated 24 MS2-binding sites into the endogenous locus of the Id3 UTR region using CRISPR/Cas9 (clustered regularly interspaced short palindromic repeat-CRISPR associated protein 9) technology in pro-B cells 445.3, which involves the generation of multiple plasmid constructs. First, we have generated a donor vector, 24xMS2V6, carrying 24 repeats of MS2 binding sites, each consisting of a 19-nucleotide RNA stem-loop, flanked by 1 kb of upstream (5') and downstream (3') homology arms of the Id3 UTR region. Second, a suitable gRNA was designed, using the CHOPCHOP online tool to target the Id3-3'UTR region, and cloned into the pSpCas9(BB) plasmid, which carries Cas9-mCherry. For MS2 integration into the endogenous Id3 locus, pro-B cells were co-transfected with gRNA/Cas9 and 24xMS2V6 constructs by electroporation. After 48 hrs, cells were harvested and sorted at the single-cell level based on mCherry fluorescence (**Figure 3.2.1a**). Stable clonal cells carrying Id3-MS2 were confirmed by genotyping using the PCR primers flanking the 5'- and 3'-homology arms of the Id3 as well as the MS2 internal primers. The PCR-amplified products were also cloned and sequenced for confirmation of MS2 insertion into the Id3 locus (**Figure 3.2.1b**).

To detect the Id3 transcription, we generated a retroviral construct encoding MS2 Coat Protein-GFP (MCP-GFP) and transduced it Id3-MS2 knock-in pro-B cells. After 48 hrs, the transduced cells were harvested, plated onto optical dishes, and maintained under optimal conditions on the microscope stage. Images were captured using Zenblack software with a cooled CCD camera attached to a fluorescence microscope (Zeiss, 880 Airyscan laser scan microscope). A three-dimensional data set with Z-stacks spanning 10 μ at 0.4 μ intervals, which comprises 25 Z planes, was acquired using a 488-excitation laser. These studies revealed that the integration of MS2 loops was successful, and we were able to capture the Id3 transcription as a concentrated GFP puncta signal in Id3-MS2 knock-in pro-B cells. In the control cells, without Id3-MS2 integration, a background GFP signal could be seen that was not scored as puncta. Accumulation of Id3 nascent transcripts and spontaneous binding of MS2 facilitated the formation of multi-molecular GFP puncta. Because GFP puncta were observed only in Id3-MS2 knock-in cells and could not be detected in control cells, indicates that the GFP puncta are directly associated with real-time transcription events of the Id3 gene (**Figure 3.2.1c**). These data show that MS2 is being

targeted to the Id3 locus, and the MS2coat protein is able to bind to the MS2 stem-loop sites as revealed using confocal imaging.

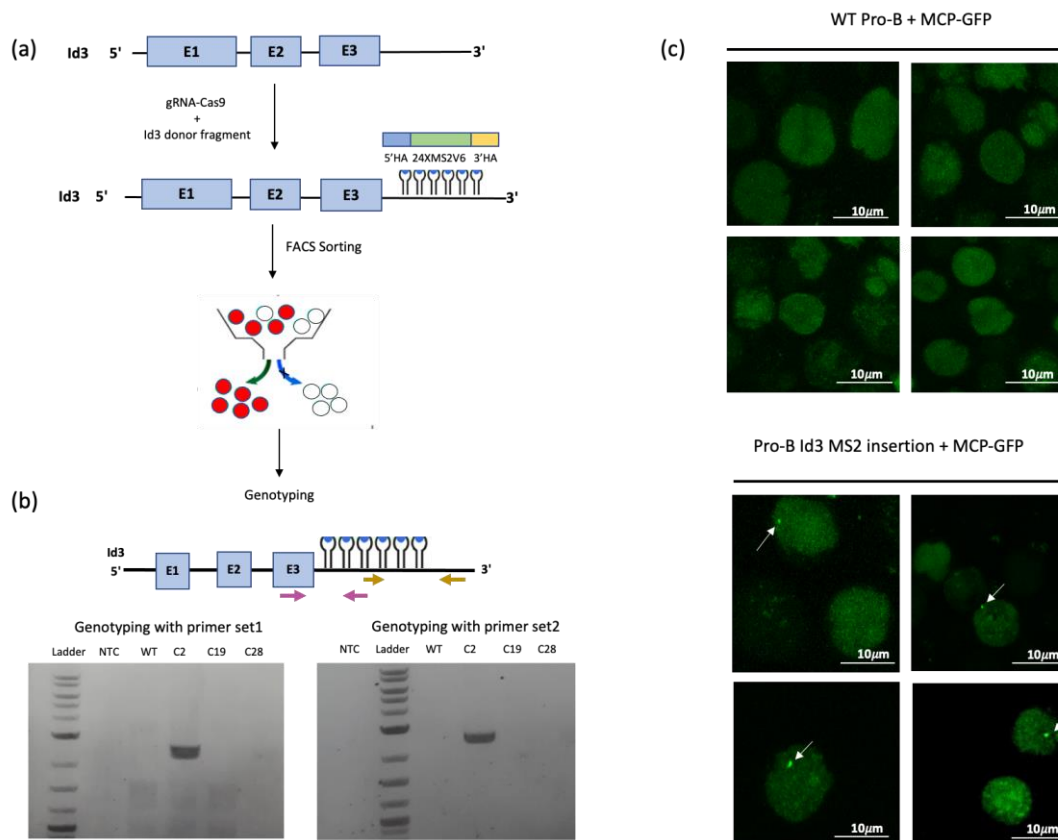


Figure 3.2.1 Generation of the Id3-MS2 pro-B Cell Line. (a) Experimental strategy for generating an Id3-MS2-pro-B cell line by using CRISPR/Cas9 engineering. (b) Genotyping of cell colonies obtained from single cell sorting by PCR from genomic DNA to confirm MS2 insertion into Id3 locus is shown. (c) Snapshots of live pro-B cells indicating Id3 transcription in control as well as MS2 knock-in cell lines after transduction with MCP-GFP retrovirus are displayed.

3.2.2 Generation of Id3-MS2 transgenic mice

As the Id3-MS2 system is enabled to detect Id3 transcription in live pro-B cells, we aimed to generate mouse MS2-Id3 transgenic lines to monitor Id3 transcription under various physiological conditions. Using the same gRNAs that were employed for establishing MS2-knockin Id3 pro-B cell lines, we generated MS2-Id3 transgenic lines. During this, CRISPR injection mixes containing crRNA, tracrRNA, and Cas9 protein were injected into zygotes and implanted into pseudo-pregnant females. The pups derived from these zygotes were genotyped for the Id3-MS2 insertion using PCR (**Fig 3.2.2a**). These transgenic lines were generated at the UCSD transgenic core facility.

To identify homozygous mice, primers external to MS2 loop sequences were used for amplification. Based on the primer design, the parental cells should give rise to a 573 bp DNA fragment, whereas homozygous knock-in alleles should generate a 2.5 kb amplified fragment. The appearance of amplification products of 573 bp and 2.5 kb fragments suggests heterozygosity (**Fig 3.2.2b**).

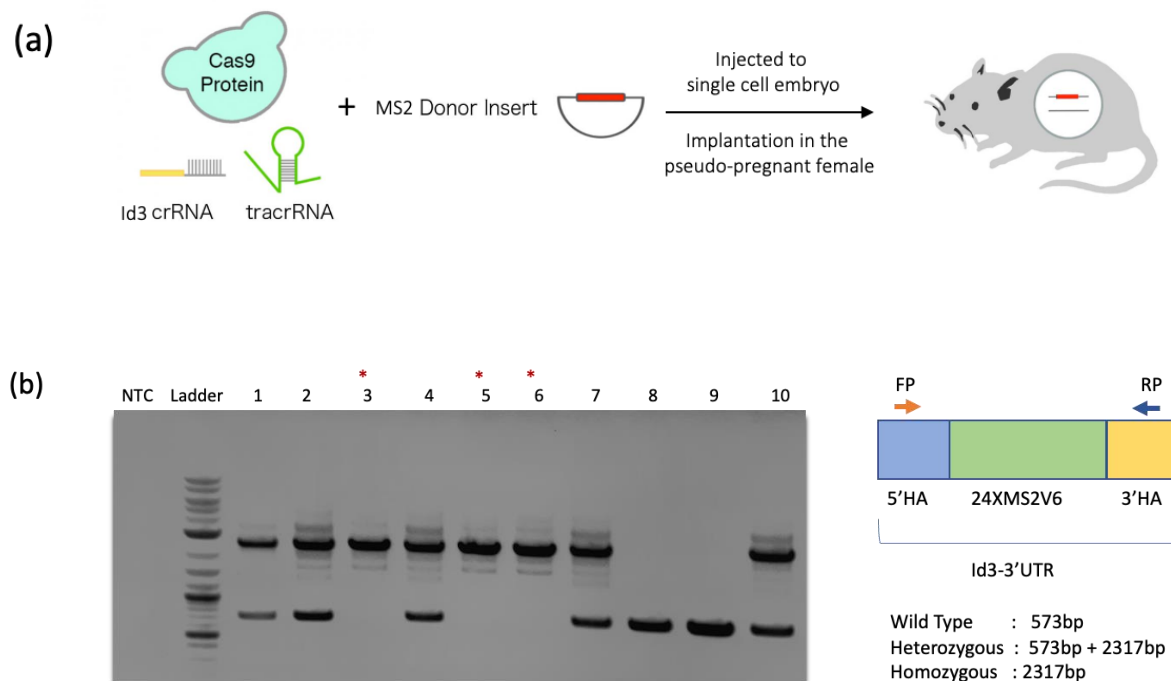


Figure 3.2.2 Generation of Id3-MS2 transgenic mice. (a) Strategy for generating Id3-MS2 mice. (b) Genotyping of pups obtained after CRISPR injection using primers external to the site of MS2 insertion.

3.2.3 Establishment of In-vitro differentiation of T cell progenitors

Differentiation of hematopoietic stem cells (HSCs) from the bone marrow into T-lymphocytes can be achieved in vitro with the support of OP9-DL1 cells, a bone-marrow-derived stromal cell line that ectopically expresses the Notch ligand, Delta-like 1 (Dll1). This approach provides a simple, versatile, and efficient culture system that allows for the commitment, differentiation, and proliferation of T-lineage cells from different sources of stem cells (Holmes et al., 2009).

Multipotent progenitors were purified from wild type mice bone marrow (**Fig3.2.3a**). and allowed to differentiate on the OP9-DLL1 stroma in the presence of IL-7 and FLT3L for a period of 20 days. Cells were fed with fresh media every three days. The cells were harvested at various

time points, and the development of T cells was monitored by staining with c-Kit and CD25. Flow cytometry analysis revealed, using these culture conditions, that MPPs gave rise to DN1 through DN4 T cell progenitors (**Fig3.2.3b**). The in vitro differentiated cells were next stained for TCR $\gamma\delta$ and TCR β . $\gamma\delta$ T cells developed swiftly as early as day 5 but declined in later days of differentiation. On the other hand, the $\alpha\beta$ population showed slower expansion in the initial days of differentiation. The percentage of the population continued to increase by day 20, indicating that almost 20% of cells expressed the TCR β chain (**Fig3.2.3c**). Thus, by culturing the progenitors under T- lymphoid conditions, we were able to efficiently generate early T lymphocytes.

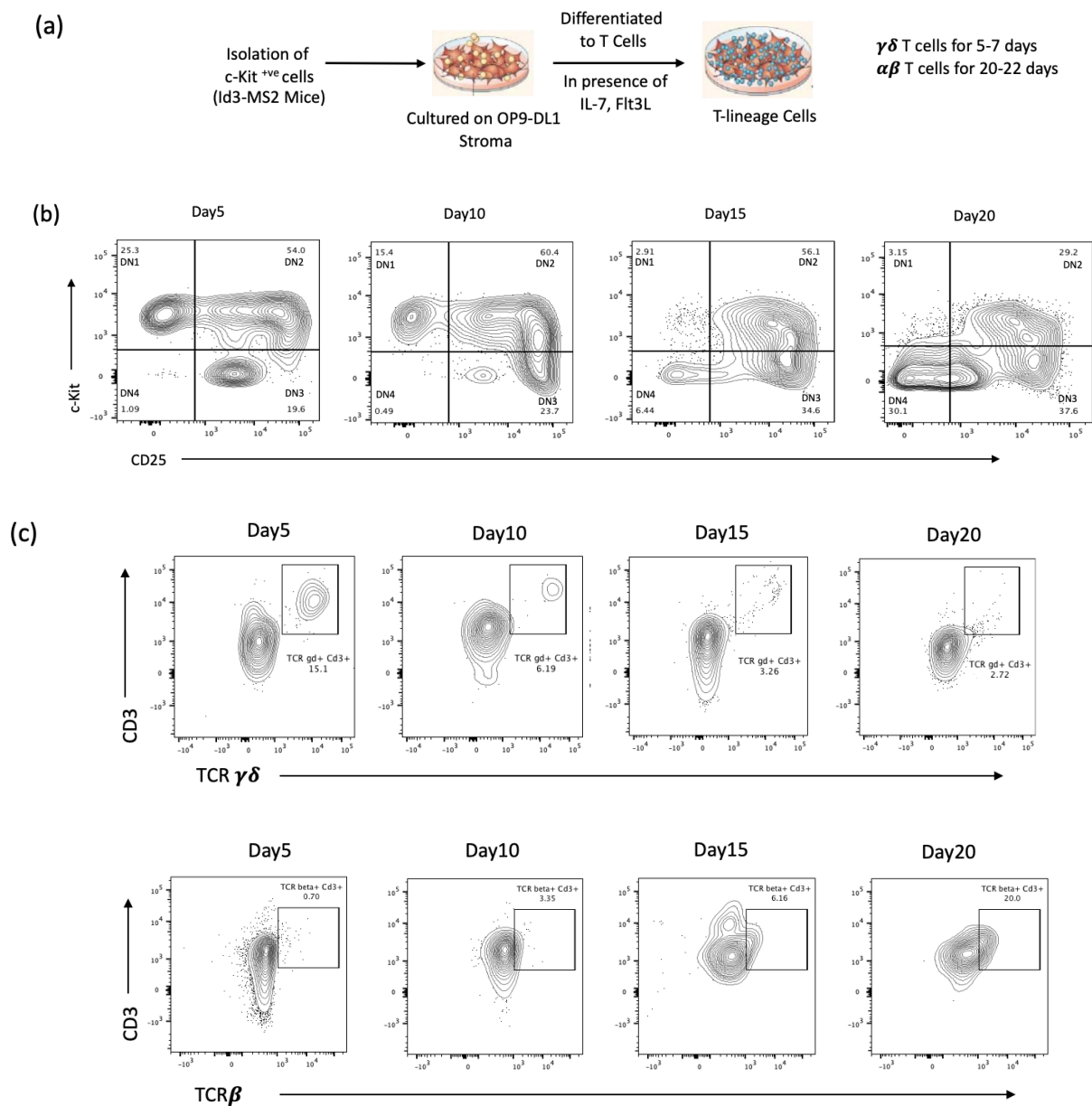


Figure 3.2.3 In-vitro differentiation of T cell progenitors (a) Experimental strategy for in vitro differentiation of multipotent progenitors into T cell progenitors. MPPs maintained under T lymphoid-promoting conditions (OP9 DLL1 stroma and cytokines– Flt3L and IL-7) and cultured for 20 days to

differentiate into T cells. (b) Flow cytometry analysis showing expression of early T-lineage markers, c-Kit and CD25, on MPPs differentiated into T-lineage cells by culturing on OP9-DL1 stroma for 20 days. (c) Flow cytometry analysis showing expression of TCR $\gamma\delta$ and TCR β in differentiating T cells. Upper panel shows TCR $\gamma\delta$ expression and lower panel shows TCR β expression in differentiated cells.

3.2.4 Optimization of signal-to-noise ratio for MCP-GFP mediated fluorescence

Previous studies utilized UbC-std MCP-std GFP fusion proteins to visualize transcription in mammalian cell lines (Wu et al., 2015). Our initial attempts to detect Id3 transcription mediated through MCP-GFP fluorescence in mouse primary pro-B cells were specific with distinct puncta and showed a high background of GFP fluorescence. To overcome this, we attempted to optimize the signal-to-noise ratio of MCP-GFP mediated fluorescence when bound to the Id3 transcript. We constructed various vectors following two distinct strategies: (1) use of divergent promoters CMV, PGK, or UbC to drive the expression of MCP-EGFP fusion protein, (2) addition of single nucleotide mutations close to the translation initiation site (AUG) to control the expression levels for each of the promoters (**Fig 3.2.4a**). These constructs were transfected and generated retroviral supernatants using HEK 293T packaging cells. Then, mouse pro-B cells were transduced with the viral supernatants obtained from wild type or mutant vectors, and MCP-GFP signal intensities were examined using flow cytometry following two days post-transduction. These studies revealed that the CMV and PGK promoters in conjunction with WT Kozak sequences showed high levels of GFP expression, whereas the UbC promoter with WT Kozak sequences showed moderate expression levels. The PGK and UbC promoters carrying mutated Kozak sequences also showed low levels of GFP expression. On the other hand, CMV promoter carrying mutated Kozak sequences displayed optimal levels of MCP-GFP when compared with the other constructs (**Fig 3.2.4b**). Furthermore, we carried out Id3 transcription live imaging using virus supernatants obtained from these vectors to vigorously test the optimal signal-to-noise ratios of MCP-GFP fluorescence. Consistent with the flow-cytometry analysis, fluorescence image analysis indicated that the CMV vector carrying a mutant Kozak sequence turned out to be the best vector for an optimal signal to noise ratio (**Fig 3.2.4c**). Based on these studies, we employed the CMV vector carrying the mutant Kozak sequence (CMK-MCP-GFP) for all subsequent live cell imaging studies.

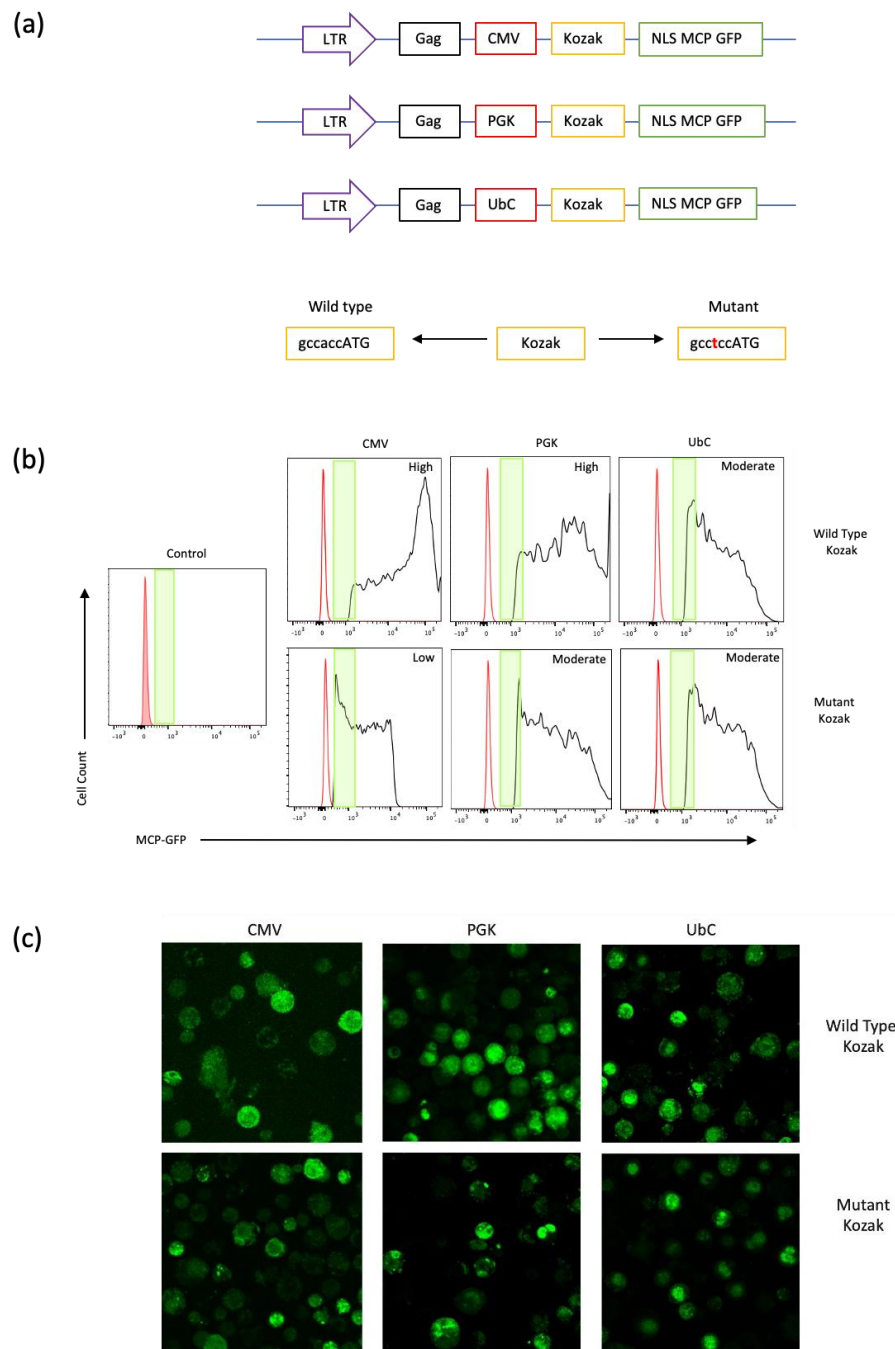


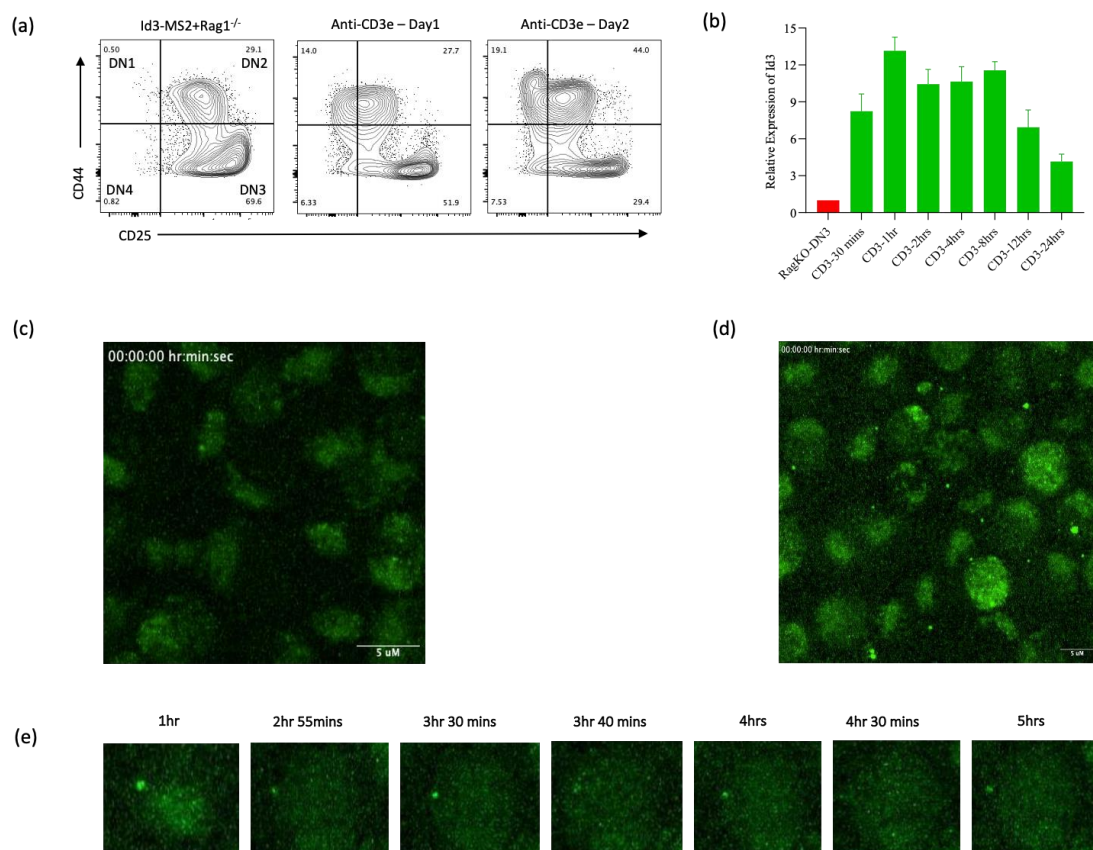
Figure 3.2.4 Optimization of signal-to-noise ratio for MCP-GFP mediated fluorescence (a) Construction of MCP-GFP-expressing retroviral vectors. The NLS-MCP-GFP coding sequence was cloned into three distinct retroviral vectors containing CMV, PGK and UbC promoters. Each construct was made using either a wild-type or mutant Kozak sequence. **(b)** Mouse primary pro-B cells were infected with NLS-MCP-GFP-expressing retrovirus, and expression levels were monitored using flow cytometry. Green bars show approximate expression levels leading to high signal to noise ratios of MCP-GFP mediated fluorescence. **(c)** Mouse primary pro-B cells infected with NLS-MCP-GFP expressing retroviruses and Id3 transcription was visualized using fluorescence microscopy. The dots in each cell show active sites of Id3 transcription.

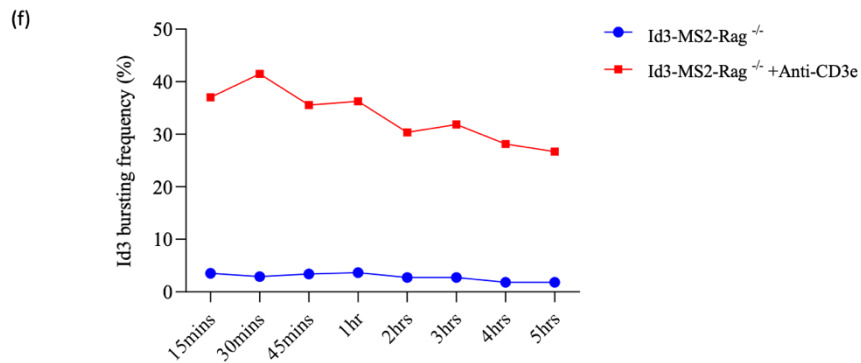
3.2.5 Id3 bursting frequency is rapidly activated by pre-TCR signaling

Previous studies have demonstrated that Id3 expression levels were found to be a rate-limiting factor during the $\gamma\delta$ vs $\alpha\beta$ T cell-fate choice of multipotent progenitors and that its expression levels positively correlated with the extrinsic signals, particularly T-cell antigen receptor signaling (Zarin et al., 2014). Two distinct lineages of T cells, $\alpha\beta$ and $\gamma\delta$ populations, employ distinct TCR complexes. One lineage employs TCR proteins termed $\alpha\beta$ while the other uses the TCR $\gamma\delta$ pair of proteins. Since the developing thymocyte progenitors are heterogeneous and respond differently to TCR stimuli, we have attempted to carry out measurements of Id3 expressions during the development of $\alpha\beta$ and $\gamma\delta$ T cells in a controlled cell type-specific manner. To accomplish this, we crossed Id3-MS2 mice with Rag-1 (Recombinase activating gene-1) knockout mice to generate the Id3-MS2; Rag1^{-/-} compound transgenic line. Mutation of Rag-1 or -2 (Recombination activating genes-1 and -2) results in T cell development blockage at early stages of development, and the mutant progenitors fail to undergo rearrangement of these antigen receptor proteins. Restoration of expression of $\alpha\beta$ and $\gamma\delta$ proteins would allow the mutant progenitors to differentiate towards $\alpha\beta$ and $\gamma\delta$ early T lineage cells, respectively. These early T lineage cells can be progressively differentiated into the DP stage by anti-CD3 stimulation when cultured in the presence of OP9-DL1 stromal cells (Ciofani et al., 2004). Therefore, it is anticipated that multipotent progenitors isolated from Id3-MS2;Rag1^{-/-} transgenic mice would allow monitoring Id3 expression in response to alpha-beta or gamma-delta antigen TCR in a conditional and cell-type specific manner.

To test this, initially, DN3 thymocytes were isolated from Id3-MS2;Rag1^{-/-} transgenic mice and allowed to differentiate on OP9-DLL1 stroma in presence of lymphoid cytokines (IL-7 and FLT3L) and in the absence or presence of anti-CD3 antibodies. Cells were analyzed by staining with cell surface receptors, CD44 and CD25, that represent early T lineage markers. In parallel, Id3 expression levels were analyzed following anti-CD3e treatment at various time points. FACS analyses revealed that in vitro engagement of TCR complex allowed DN3 cells to readily differentiate into the DN4 compartment (**Fig 3.2.5a**). Interestingly, induction of Id3 expression was observed as early as 30 min post anti-CD3e stimulation, and continued to sustain high levels for 8 hrs., and then downregulated progressively (**Fig 3.2.5b**) reflecting a positive correlation between TCR signaling and Id3 expression.

To validate these findings, we performed live cell imaging on DN3 cells derived from Rag^{-/-}-Id3-MS2 mice following the expression of MCP-GFP, before and after stimulation with anti-CD3e. No Id3 puncta were observed in control cells before anti-CD3e treatment (**Fig 3.2.5c**) whereas in vitro stimulation of cells with anti-CD3e antibodies resulted in the formation of distinct fluorescent puncta that were associated with Id3 nascent transcripts (**Fig 3.2.5d**). The live cell images were captured for several hours. We found that Id3 expression followed a burst pattern. Snapshots of single cells at various time points indicated that transcription showed variability in signal intensity and underwent a distinct ON and OFF pattern (**Fig 3.2.5e**). We captured images for 150-200 cells and calculated bursting frequencies at various time points. We found that cell responses to the anti-CD3e stimuli were instinctive and immediate. This data indicates that pre-TCR mediated signaling rapidly induces a distinct pattern of Id3 transcriptional bursting (**Fig 3.2.5f**).





3.2.5 Id3 burst frequency is rapidly activated by pre-TCR signaling

(a) Flow cytometry plots showing developmental progression of DN3 cells to DN4 upon stimulation with anti CD3e. Left plot indicates developmental arrest of differentiated T-cells at the DN3 cell stage derived from

Rag^{-/-}-Id3-MS2 bone marrow. Middle and right plots show differentiation of DN3 cells to DN4 cell stage post anti-CD3e stimulation at day 1 and day 2, respectively. (b) Real-time PCR measurements of Id3 mRNA expression levels at different time points starting from 30 minutes and 24 hours after anti-CD3e stimulation. (c) Snapshot of live cell imaging representing Rag^{-/-}-Id3-MS2 control cells. No Id3 signal was observed. (d) Snapshot of live cell imaging representing Rag^{-/-}-Id3-MS2 cells stimulated with anti-CD3e. A large fraction of cells started bursting within 15 minutes of stimulation. (e) Snapshot of a single cell at various time points indicating that transcription shows a variability in the intensity of the signal and undergoes an ON and OFF pattern. (f) Calculation of bursting frequencies at various time points spanning 15 minutes-5 hours post-stimulation. Data were collected for 110 cells for Rag^{-/-} control cells and 135 cells stimulated with anti-CD3e. The frequencies were calculated in percentages.

3.2.6 The ERK-MAPK signaling pathway regulates Id3 bursting

Previous studies have shown that the defect in thymocyte maturation observed in mice that lack Id3 is similar to the defect observed in mice that lack MEK1 or Ras. Moreover, in the presence of a MAP kinase inhibitor, induction of Id3 is impaired, indicating a potential role of the ERK pathway in regulation of Id3 gene expression. These studies indicate that MAPK plays an important role during T cell positive selection (Pages et al., 1999; Alberola-Ila et al., 1995). Furthermore, it has been shown that Id3 transcription is induced by Egr1, a downstream target of the MAP kinase pathway (Bain et al., 2001). These studies supported a model in which TCR ligation promotes the inhibition of E-protein activity by activation of the RAS–ERK–Id3 cascade (Engel et al., 2001).

To determine whether and how ERK signaling modulates the Id3 transcription pattern, we isolated DN3 cells from Id3-MS2 transgenic mice and allowed them to differentiate on OP9-DLL1 stroma under T cell culture conditions. Cells were stimulated by anti-CD3e antibodies either in the presence or absence of the ERK inhibitor PD98059. Flow cytometry analysis indicated that the ability of DN3 cells to differentiate into the DN4 stage is impaired in the presence of an ERK inhibitor as compared to the absence of an inhibitor (**Fig 3.2.6a**). Consistently, we found that interference with ERK signaling substantially decreased Id3 mRNA abundance, as evidenced by the real time PCR measurements. Id3 mRNA expression was less abundant upon stimulation with

anti-CD3e antibodies in the presence of an ERK inhibitor when compared to the cells treated in the absence of ERK inhibitor (**Fig 3.2.6b**).

As expected, live cell imaging indicated that Id3 expression levels were greatly diminished in the presence of an ERK inhibitor. **Fig 3.2.6c** shows Id3 transcription in cells upon stimulation with anti-CD3e, and **Fig 3.2.6d** shows cells stimulated with anti-CD3e in the presence of an ERK inhibitor. We detected a few Id3 puncta when cells were treated with PD98059 prior to anti CD3e stimulation. Together, the image analysis data show $Rag^{-/-}$ do not show Id3 expression, whereas upon anti-CD3e stimulation highest induction of Id3 bursting was observed whereas with ERK signaling inhibition the bursting frequency of Id3 transcription was significantly reduced (**Fig 3.2.6e and Fig 3.2.6f**).

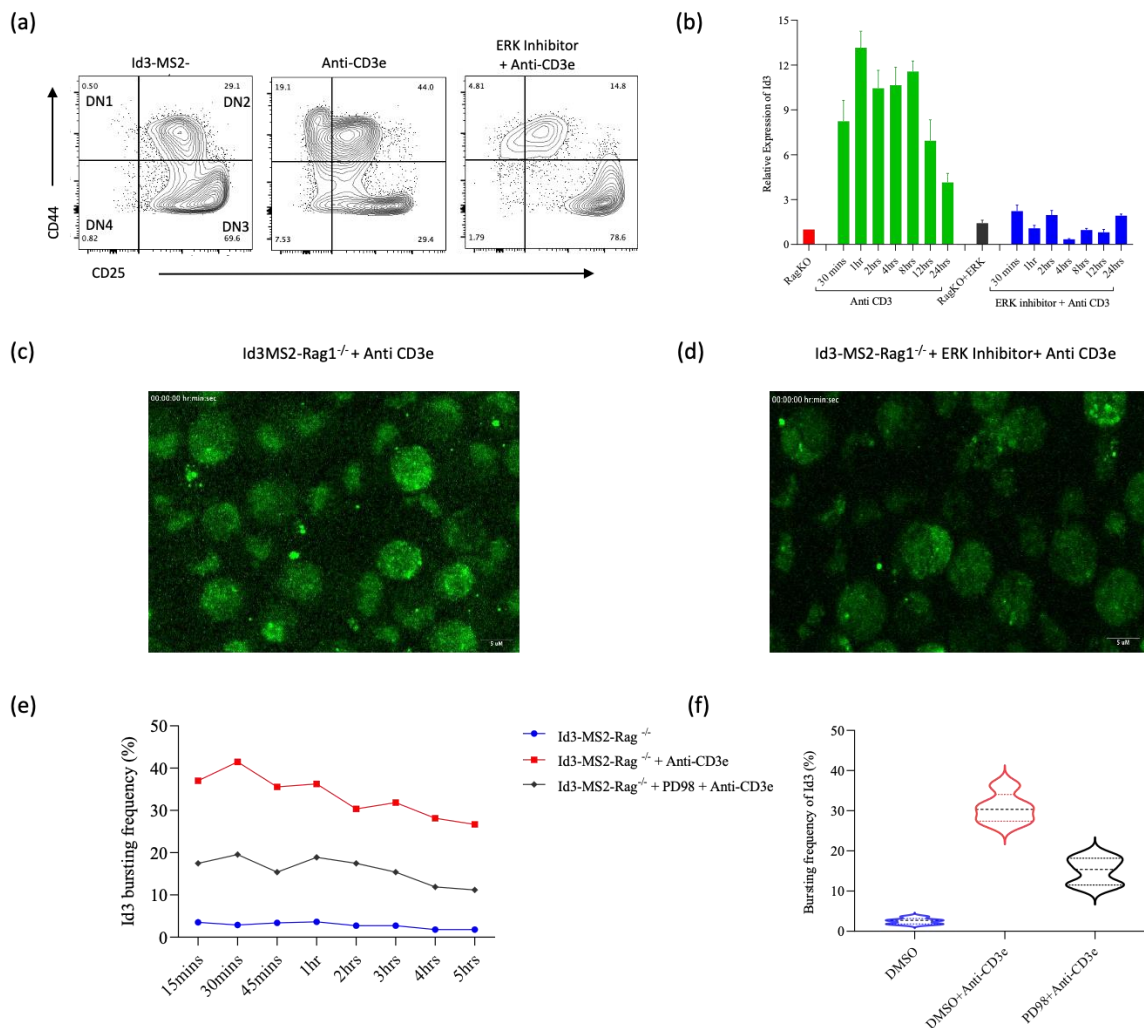


Figure 3.2.6 The ERK-MAPK signaling pathway regulates Id3 bursting. (a) Flow cytometry plots showing developmental progression of DN3 to DN4 cells. Left plot indicates developmental arrest at DN3 for cells derived from $Rag^{-/-}$ -Id3-MS2 mice bone marrow. Middle plot shows transition from DN3 to DN4

cell stage post anti-CD3e stimulation at day 2. Right plot indicates slow progression of DN3 to DN4 cells in presence of ERK inhibitor PD98059 2 days post-treatment. (b) qPCR measurements of Id3 transcript levels at different time points starting from 30 minutes -24 hours after anti-CD3e stimulation in the absence or presence of ERK inhibitor PD98059. Treatment with ERK inhibitor greatly reduced Id3 expression (c) Snapshot of a live cell movie representing Rag^{-/-}-Id3-MS2 cells stimulated with anti CD3e in the absence of ERK inhibitor. (d) Snapshot of a live cell movie representing Rag^{-/-} cells stimulated with anti CD3e in the presence of ERK inhibitor. Comparatively fewer puncta were observed. (e) Calculation of bursting frequencies at various time points at 15 minutes- 5 hours-time-points for Rag^{-/-} control cells, cells stimulated with anti CD3e in the presence or absence of ERK inhibitor. The frequencies are presented as percentages. (f) Calculation of cumulative bursting frequencies. Data collected from 110 cells for Rag^{-/-} control cells, 135 cells for cells stimulated with anti CD3e in absence of ERK inhibitor and 143 cells for cells stimulated with anti-CD3e in presence of ERK inhibitor. The frequencies are shown as percentages.

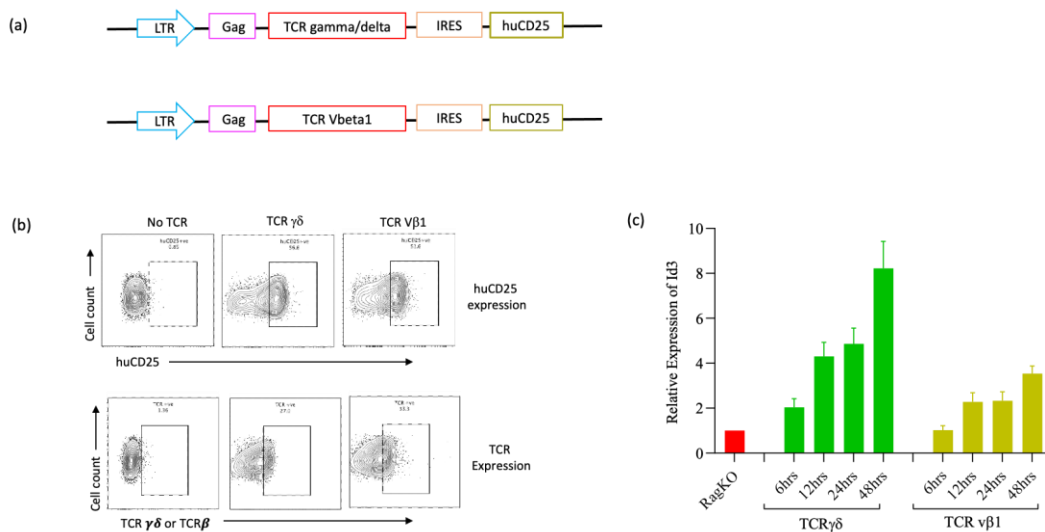
3.2.7 TCR $\gamma\delta$ -mediated signaling induces higher Id3 bursting frequencies as compared to signaling induced by the TCR β chain

The early thymic progenitors, ETPs, bipotent precursors that differentiate either into $\alpha\beta$ - or $\gamma\delta$ -lineage T cells. The diverse T cell fate found to be under the control of pre-TCR and $\gamma\delta$ -TCR chains (Zarin et al., 2014). It is proposed that the productive rearrangement and the signal strength promote the T cell fate choice of the ETPs (Hayes et al., 2005; Haks et al., 2005). The apparent connection between lineage choice and the TCR expressed by the cell can be severed by manipulations of TCR signal strength. Previous studies demonstrated that stimulating stronger signals via expression of ERK/MAPK-induced Id3 appears to promote the $\gamma\delta$ -lineage fate in developing DN3 cells, suggesting a critical role for Id3 in mediating $\alpha\beta$ - versus $\gamma\delta$ -lineage decisions at this developmental checkpoint. Id3 is required for strong TCR signals to both promote adoption of the $\gamma\delta$ - fate and oppose the $\alpha\beta$ -fate outcome (Lauritsen et al., 2009).

To track the transcription dynamics of Id3 in real-time during the course of $\gamma\delta$ versus $\alpha\beta$ cell differentiation, DN3 cells derived from Id3-MS2;Rag1^{-/-} mice were subjected to over-expression of TCR $\gamma\delta$ and TCR β genes. To achieve this, retroviral plasmids were constructed carrying either TCR $\gamma\delta$ or TCR $\nu\beta$ 1 genes and huCD25, allowing rapid purification of TCR positive cells (Fig 3.2.7a). Specifically, MPPs from Id3-MS2;Rag1^{-/-} mice were cultured under T cell differentiation conditions as described earlier, then transduced with retrovirus encoding either TCR $\gamma\delta$ or TCR β polypeptide chains. Expression of TCR $\gamma\delta$ or TCR β was confirmed by flow cytometry analysis using antibodies against TCR $\gamma\delta$, TCR β and huCD25. Flow cytometry analysis revealed the expression of respective TCR along with the huCD25, indicating successful transduction (Fig 3.2.7b). To check Id3 mRNA abundance in cells transduced with TCR

retroviruses, cells were collected for RT-PCR measurements at different time intervals. We found higher levels of Id3 expression in $\gamma\delta$ T cells as compared to TCR β expressing cells. However, Id3 mRNA abundance increased gradually for both cell types with increasing time periods after infection (**Fig 3.2.7c**).

Next, time lapse live cell imaging was carried out to determine whether the different T cell receptors induce distinct and varying Id3 bursting patterns. A higher number of puncta signals associated with Id3 transcription were observed in cells expressing TCR $\gamma\delta$ as compared to TCR β (**Fig 3.2.7d and e**). The snapshot of a single cell detected Id3 transcript levels at various time points, indicating different magnitudes of signaling. Id3 expression was not only higher in cells expressing $\gamma\delta$ TCR but also many cells showed bi-allelic bursting as compared to TCR β expressing cells. Moreover, $\gamma\delta$ T cells displayed high levels of Id3 expression, as indicated by the GFP abundance, while cells transduced with TCR β expressing retroviruses exhibited only moderate levels of expression. These studies indicated that $\gamma\delta$ T cell signals positively correlate with Id3 transcriptional bursting signatures when compared to signals emanating from the TCR β chain (**Fig 3.2.7f**). We analyzed 150-200 cells, and bursting frequencies were calculated at different time intervals for both conditions. The measurements indicated a fluctuation of bursting frequencies at different time points while imaging (**Fig 3.2.7 g and h**). Cumulative bursting frequencies indicated that TCR $\gamma\delta$ mediated signaling induced high Id3 bursting frequencies as compared to TCR β expression (**Fig 3.2.7i**).



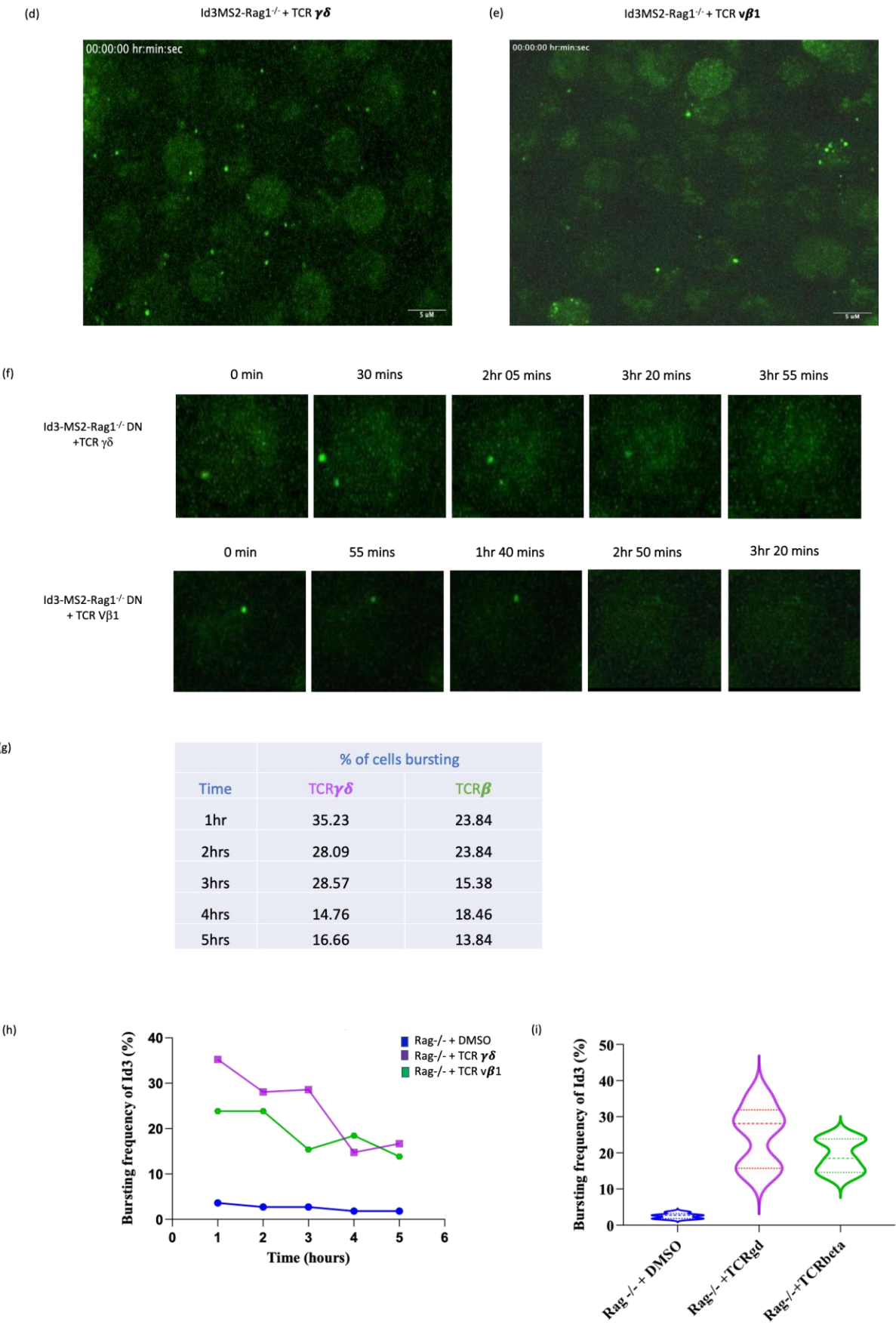


Figure 3.2.7 TCR $\gamma\delta$ induces higher Id3 bursting frequencies as compared to TCR β (a) Construction of TCR retroviral plasmids containing huCD25 marker. (b) Flow cytometry plots showing expression of

TCR 48 hours post-transduction. Upper panel shows expression of huCD25 and lower panel shows $\gamma\delta$ and β TCR expression in 4DN3 cells. (c) Id3 mRNA expression measured by RT-PCR at different hours post TCR transduction. $\gamma\delta$ TCR induced high Id3 mRNA levels as compared to β TCR. (d) Snapshot of a field of view depicting cells transduced with $\gamma\delta$ TCR. (e) Snapshot of a field of view showing cells transduced with β TCR. (f) Snapshot of a single cell from each TCR at various time points indicating different magnitudes of Id3 signal intensity and variability. (g) Id3 bursting frequencies calculated in percentages for Rag^{-/-}-Id3-MS2 DN3 cells transduced with TCR $\gamma\delta$ vs TCR β (h) Id3 bursting frequencies plotted as a function of time for Rag^{-/-}-control cells, cells transduced with TCR $\gamma\delta$ and cells transduced with TCR β . Control cells do not show Id3 bursting over the time. TCR β expressing cells showed less Id3 bursting than $\gamma\delta$ T cells. (i) Calculation of cumulative bursting frequencies. Data collected from 110 cells for Rag^{-/-}-control cells, 130 cells for cells transduced with TCR β and 200 cells for cells transduced with TCR $\gamma\delta$. The frequencies are calculated in percentages.

3.2.8 Id3 expression levels correlate with T cell antigen receptor affinity

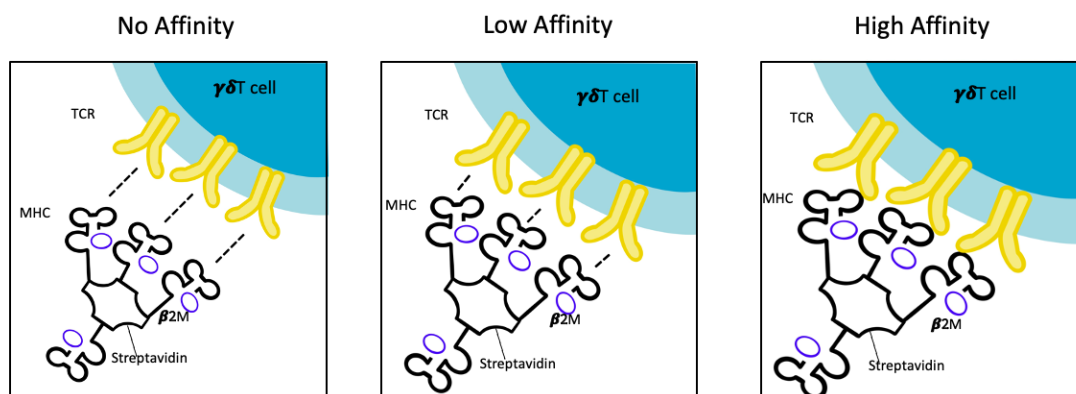
It is found that the $\gamma\delta$ T cells are a minority population in the circulating blood and lymph, typically composing only 1–5% of circulating lymphocytes. However, $\gamma\delta$ T cells can constitute nearly 50% of epithelial lymphocytes (Carding et al., 2002) and play a number of physiological roles, such as wound healing (Jameson et al., 2002), pathogenic clearance (Mombaerts et al., 1993, Morita et al., 2007, Mishra et al., 2010), and tumor surveillance (Morita et al., 2007, Mishra et al., 2010, Tanaka et al., 1995). The molecular mechanisms by which $\gamma\delta$ T cells recognize cellular distress through their TCR and carry out their effector functions remain largely unknown. Previous studies have shown that $\gamma\delta$ and $\alpha\beta$ T cells contribute differently to host immune defence. Mice deficient in $\gamma\delta$ T cells generally exhibit more profound defects in the regulation of immune function than in the clearance of intracellular pathogens (Crowley et al., 2000). The signal strength hypothesis asserts that the quantity of the signal downstream of the TCR can direct lineage choice (Zarin et al., 2015). TCR signaling activates the ERK-Egr3 pathway, which results in the upregulation of Id3 in direct proportion to the strength of the signal (Lauritsen et al., 2009). The KN6 TCR was initially cloned from IFN γ producing $\gamma\delta$ T-cells that recognize the non-classical MHC1b molecules T10 and T22 (Bonneville et al., 1989 ; Felix et al., 1990). As both strong (T22) and weak (T10) ligands are known for the KN6 $\gamma\delta$ TCR, it is a very useful model for understanding the impact of TCR signal strength on $\gamma\delta$ T-cell effector function (Bonneville et al., 1989).

Previous studies demonstrated that T cells respond differently depending on the strength of the interaction between the TCR and MHC ligand (Fig 3.2.8a). To evaluate the differential levels of Id3 expression in response to variations in the affinity of antigen stimulation for T cells, we used 3 different kinds of well-known ligands with varying affinity to the KN6- $\gamma\delta$ TCR, including the no affinity ligand T10d, low affinity T10b ligand, and the T22 tetramer with high

affinity. For these studies, isolated multipotent progenitors from Id3-MS2; Rag1^{-/-} mice bone marrow were differentiated on OP9-DLL1 stroma under T cell culture conditions. Next, DN3 cells were transduced with TCR $\gamma\delta$ and TCR β retroviruses. The transduced cells were treated with soluble tetramers of T10d, T10b, and T22 ligands. RT PCR analysis showed, $\gamma\delta$ T cells treated with T10d had no effect on Id3 mRNA. T10b showed a slight increase in Id3 transcripts, while T22 showed a significant increase in Id3 mRNA levels (**Fig 3.2.8b**). On the other hand, T cells expressing the TCR β chain that were stimulated with these ligands showed no change in Id3 expression (**Fig 3.2.8c**).

Next, live cell imaging was captured under these conditions. Some snapshots of live cells are shown in **Fig 3.2.8(d)**. The control $\gamma\delta$ T cells and $\gamma\delta$ T cells treated with T10d showed similar Id3 bursting frequencies. The T10b haplotype induced a slight increase in Id3 expression. T22 tetramers stimulated more Id3 bursting, showing a greater number of puncta. Live cell imaging for TCR β cells upon stimulation with tetramers showed no response to the stimulation. TCR β cells did not respond to the stimulation with these tetramers. Data was collected from 120-150 cells for each condition, and Id3 bursting frequencies were calculated in percentages. The image analysis shows high bursting frequencies induced by high affinity T22 tetramer in $\gamma\delta$ TCR expressing cells. Low bursting frequencies are induced by the low affinity T10b tetramer in $\gamma\delta$ TCR expressing cells. There was no change in the bursting frequencies upon stimulation with the T10d haplotype (**Fig 3.2.8e**). TCR β expressing cells were non-reactive to different ligands (**Fig 3.2.8f**).

(a)



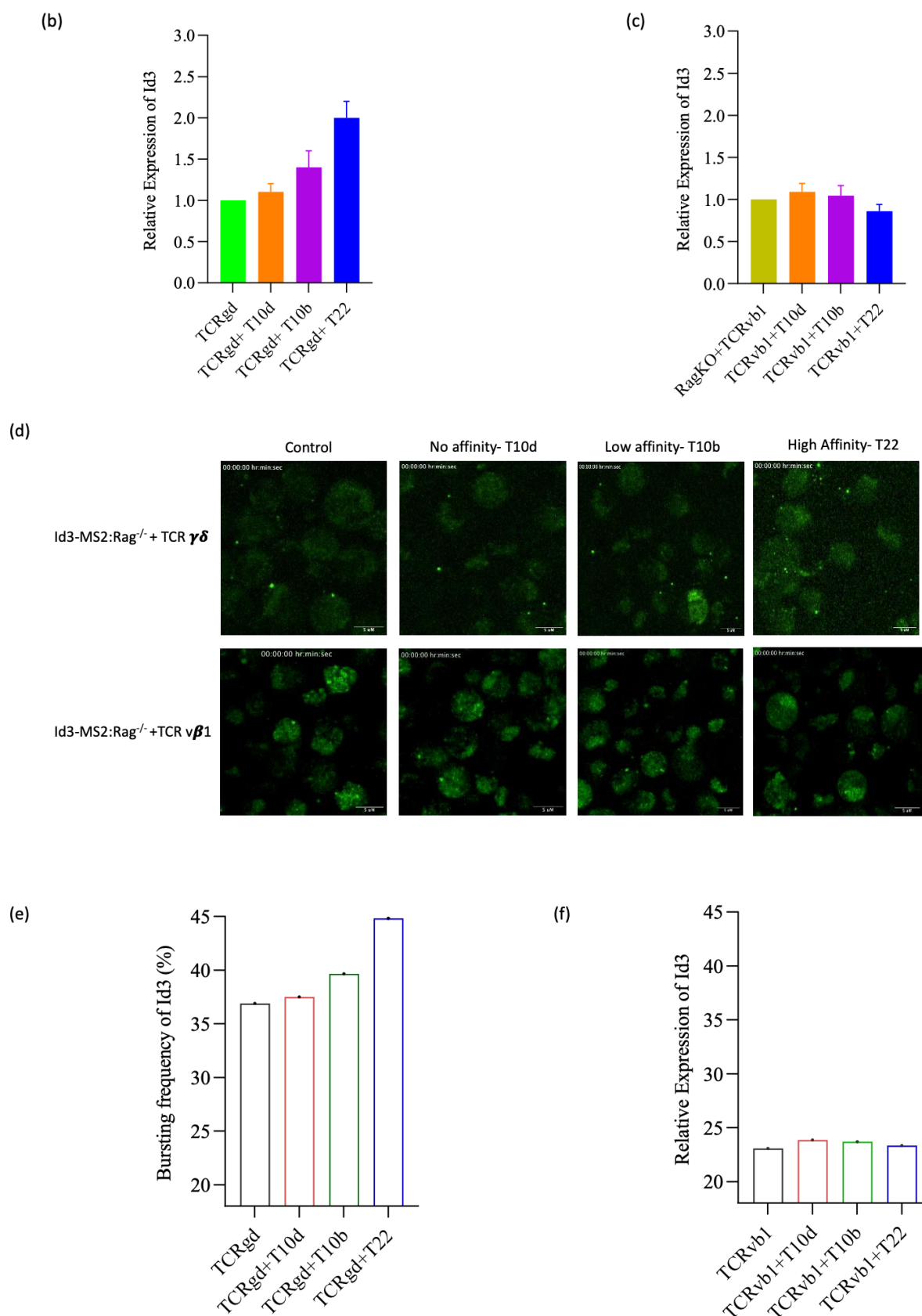


Figure 3.2.8 Id3 expression levels correlate with T cell antigen receptor affinity (a) TCR-MHC interaction in response to variations in the affinity of antigen affinity. T10d had no effect on Id3 mRNA levels, T10b showed a small increase in mRNA expression whereas T22 tetramer showed high levels of Id3 mRNA abundance. (c) Cells expressing the TCR β showed no impact on Id3 mRNA abundance. (d)

Upper panel shows snapshots of a live cell imaging movies performed on $\gamma\delta$ T cells upon stimulation with T10d, T10b and T22 tetramers. Lower panel represents snapshots of a live cell imaging movie performed on cells transduced with TCR β upon stimulation with T10d, T10b and T22 tetramers. (e) Id3 bursting frequencies in cells expressing the $\gamma\delta$ TCR in response to T10d, T10b and T22 tetramer stimuli. (f) Id3 bursting frequencies in DN3 cells infected with TCR β upon stimulation with T10d, T10b and T22 tetramers.

Chapter 4

DISCUSSION

DISCUSSION: OBJECTIVE - I

In an earlier study, it was demonstrated how non-coding transcription directs chromatin folding to allow early progenitors to undergo T cell differentiation ([Isoda et al., 2017](#)). Here, we illustrate how CTCF binding sites coordinate the upregulation of Bcl11b in DN2 cells. We report here that multiple CTCF sites are tethered together to promote Bcl11b expression. We believe that CTCF binds to these sites to allow and strengthen the looping of the Bcl11b enhancer and promoter together by preventing further elongation of this genomic region by cohesin after it reaches these bound CTCF sites. These sites act redundantly to facilitate a stable upregulation of Bcl11b. While Bcl11b expression is dramatically reduced in these CBS deletions, it is not completely absent. This indicates that while certain tethered domains are optimal, one of the other sites can compensate to allow limited loop stability. It is also possible that some of this redundancy is afforded by the activity of other CBSs that were not disrupted, including the sites that were excluded from this study. Altogether, these data demonstrate that ThymoD transcription regulates Bcl11b expression by facilitating paired CTCF occupancy and activating Bcl11b expression. These studies inform a mechanism by which ThymoD transcription creates an epigenetic landscape permitting CTCF to bind to CBSs nearby the Bcl11b enhancer and promoter. Cohesin then extrudes DNA until reaching these enriched CBSs, bringing the Bcl11b enhancer and promoter into close proximity to instruct early T-cell developmental gene programs. We expect that the mechanism informed by this study will be reflected in other loci that share similarities found throughout the genome.

Our data supports a mechanism by which ThymoD transcription promotes the binding of CTCF to multiple binding sites to allow the juxtaposition of the Bcl11b enhancer and promoter by cohesin in late DN2 cells. When CTCF binding was ablated at the CR4 site in mice, there was an increase in DN2 cells *in vivo*, and a more dramatic change to all double negative cells. In our analyses of the CR4 knockout mice, we further observed a specific loss of CTCF binding at the locus responsible for this phenotype. We further observed that loss of CTCF binding at the CR4 binding site led to the downregulation of a multitude of genes essential for early T cell development. While we have focused our studies on the CR4, we anticipate that similar phenotypes may be observed at the other four sites identified. An earlier publication identified a differentially methylated region that overlaps the B2 and B3 CBS and has a significant influence on T-cell specific expression ([Li et al., 2013](#)). We chose not to focus on these sites as they were too close to

the promoter for Bcl11b. As mentioned earlier, single CTCF site deletions generally resulted in only a 50% reduction in Bcl11b expression. We further anticipate a more dramatic effect by double CBS knockouts in mice, closer to the phenotype observed in ThymoD mutant *in vivo*.

While these findings elucidate a larger mechanism for the orchestration of T cell development by ThymoD, they do not differentiate between the roles of the non-coding transcription and the transcript in this process. Future researchers may investigate this question further by conducting a knockdown of the mature transcript by treating cells with siRNA, shRNA, or ASOs. The ThymoD transcript could facilitate CTCF binding through three mechanisms: stabilizing CTCF-DNA interaction by enhancing the binding affinity of CTCF, facilitating CTCF binding by promoting phase separation due to the accumulation of the transcript and/or RNA polymerase II; or doing both. Recent studies have clarified how the carboxy-terminal domain of RNA Polymerase II is sufficient to form phase-separated droplets (Boehning et al., 2018). Further evidence was presented that RNA can increase the affinity of CTCF binding to DNA through its interaction with RNA binding regions (Hansen et al., 2019). It is unclear whether the RNA stabilizes the CTCF-DNA interaction through phase separation or through an unknown mechanism, as a recent genome-wide study found minimal effect of transcription inhibition of CTCF binding, while depletion of CTCF resulted in a widescale loss of transcriptional condensates (Lee et al., 2022). However, this study was conducted in a cancer cell line and did not focus on lncRNAs, many of which have been shown to directly interact with CTCF, with thousands more interactions predicted (Guo et al., 2020; Kuang et al., 2020).

Some of the most enriched gene subsets revealed by our RNA-seq analysis include the genes involved in brain development and tumorigenesis. These results suggest a possible role for CTCF binding within Bcl11b in these processes, with some of these differentially expressed genes also presenting themselves in my analysis of the previously published ThymoD p(A)/p(A) RNA-seq. As mentioned earlier, mutations in Bcl11b are found in a large subset of patients diagnosed with T-ALL. Similarly, ThymoD mutants had high incidences of tumors and increased morbidity (Isoda et al., 2017). These tumors were often associated with antisense ThymoD transcription, allowing them to express Bcl11b transcription. Less is known about the effects of intergenic mutations on Bcl11b expression and the resulting shifts in immune populations. Mutations in Bcl11b have also been linked to a range of neurodevelopmental disorders, but the role of ThymoD

and these CTCF sites remains unexplored (Lessel et al., 2018). We hope that these findings may aid in the identification of intergenic mutations that could lead to T-ALL, neurodevelopmental, and related diseases in patients. In fact, recent studies have discovered mutations in CTCF sites that led to aberrant limb development (Ushiki et al., 2021). These studies indicate that further clarifying the roles of CTCF sites genome-wide in developmental contexts would be instrumental in understanding the establishment and severity of disease outcomes. Together, our findings provide a basis for the requirement of multiple CTCF sites for developing appropriately balanced immune repertoires from early T cell progenitors downstream of ThymoD transcription. We suggest that this process is further specified by gene expression patterns in these progenitors. The role of CTCF binding in T cell developmental processes supports the hypothesis that T cell development is orchestrated by ThymoD-mediated loop extrusion at a restricted developmental stage.

OBJECTIVE-II

Previous studies have demonstrated that the developmental progression of T cells is closely associated with the gene activity of key transcription factors, including Notch1, Ikaros, GATA3, E2A, Bcl11b, and Id3 that are involved in the assembly of stage-specific gene regulatory networks (Rothenberg, 2014; Murre, 2019). Transcriptional control of these factors is often initiated and sustained by unique extra-cellular signals that forge diverse T cell identities and promote T cell activation. Particularly, the antagonistic E2A-Id axis orchestrates the developmental choice between gammadelta ($\gamma\delta$) and alphabeta (ab) T cells as well as CD4 and CD8 single-positive T cells. While the role for E2A and Id3 in promoting distinct T cell fate choices is well established, insights on their transcriptional characteristics in real-time are yet to be understood. To address this, we successfully implemented the MS2 system (Buxbaum et al., 2014; Park et al 2014), which allows for single-cell transcription measurements, and show how these experiments broadened our mechanistic understanding of Id3 transcriptional regulation. Although transcription kinetics can be assessed in several ways, we focus here on dynamic measurements in live single cells for several reasons. First, to visualize the transcription dynamics at a single cell level in real-time. Second, to assess the transcription kinetics, particularly during physiological processes such as cell differentiation and stimulation. Third, to quantitatively measure mRNA production over time, which provides a distinct opportunity to connect the external signals and expression behaviour. Fourth, the time resolution of single cell studies far surpasses the currently available transcript

measurements. We were successfully able to establish mouse transgenic lines carrying the MS2 system, which allows to detect and track specific mRNA transcripts in single cells through live-cell imaging, to monitor Id3 expression in real-time at a single cell level. As a first attempt, we visualized Id3 bursting frequencies and transcription rate during T cell developmental choice of multipotent progenitors and during stimulation of $\gamma\delta$ T cells with MHC like tetramers.

The visualization of Id3 transcription over time yielded information on the relation between the transcription and its variance during differentiation and stimulation. These studies revealed that in differentiating multipotent progenitors, Id3 transcription is a stochastic process that varies from cell to cell. This may be due to short-lived dynamic interactions between promoter and enhancer as well as binding of transcription factors with their target motifs in the Id3 locus or pausing of RNA polymerase. This variability can be explained by the assumption that transcription occurs in bursts, with periods of high activity followed by periods of inactivity. Additionally, the genome position of Id3 and transcription factors', which bind at the cis-regulatory elements, impinges on the rate of transcription (Blake et al., 2006; Hornung et al., 2012). These findings support a recent model that suggests that a promoter switches between two states, on and off, that transcription occurs only in the on state (burst size), and that the rate of switching to the on state (burst frequency) varies from cell to cell. Consistent with this hypothesis, mammalian genes were found to be dominated by burst transcription (Bahar et al, 2015; Singer et al., 2014). However, the transcription patterns are relatively gene-specific, with varying characteristics of burst size, burst frequency, and burst duration. These characteristics are regulated by promoter strength, genomic position, as well as interactions with their target cis-regulatory elements that bind varying transcription factors. Thus, cells undergo major changes in gene expression patterns during the course of a developmental program and potentially employ the binary mode of regulation to reduce the noise of developmentally regulated genes.

These studies show that the amount of Id3 transcription is much higher in differentiating T cells, especially $\gamma\delta$ T cells, than in $\alpha\beta$ T cells. These observations are consistent with the findings that $\gamma\delta$ TCR exerts high thresholds of signal, which may be important for the selective activation of lineage specific gene expression programs as compared to pre-TCR (TCRb with invariant pre-Ta). Similarly, in $\gamma\delta$ T cells, Id3 expression readily increases to relatively high levels upon antigen stimulation with T10b and T22 tetramers, and a concomitant increase in Id3 bursting frequencies is observed. The increased spike in transcription in response to TCR stimulation indicates that Id3

is an inducible gene. The observed increase in transcription may be due to an increased rate and frequency of bursts under the influence of externally regulated signals. This is further supported by the findings that high affinity tetramers but not the low or no affinity tetramers failed to activate Id3 transcription. These findings raise the question of the functional significance of high levels of Id3 expression in response to $\gamma\delta$ TCR or an antigen response. Previous studies indicate that Id3 modulates E2A activity by inhibiting its DNA binding activity. It is highly likely that high levels of Id3 are important for opposing E2A function, which is necessary for the $\gamma\delta$ T cell fate option and for the antigenic response.

Over the years, a lot of progress has been made in understanding the regulatory components of immune cell development, including transcription factors and epigenetic factors (Smale, 2014). However, insights into the temporal transcription dynamics of the key factors that underlie immune cell development remain to be studied. By putting together the quantitative measurements of transcription over a time component, the dynamic distribution of transcripts in a single cell would show the relationship between the rate of transcription and its variation. Different physical models of gene expression can be developed, considering the time and frequency of gene expression. But it is possible that the dynamics seen in the transcription may not match the steady-state levels of Id3 protein in each cell. Because mRNA has been shown to be transcribed more than once (Cai et al., 2008; Yu et al., 2006). Therefore, it is possible that the amount of protein distribution at a single cell level is a cumulative outcome of transcriptional and translational events. Combining the relative contributions of heterogeneity in transcriptional dynamics and posttranslational processes remains to be explored (Csardi et al., 2015). Collectively, the studies proposed here would reveal for the first time the real-time quantitative transcription measurements of Id3, which performs distinct roles in T lymphoid development and function.

Chapter 5

SUMMARY

SUMMARY

Multiple lines of evidence suggest that regulatory factors such as promoters, enhancers and insulators exhibit long-range intra- and inter-chromosomal interactions that are often closely linked with modulation of distinct transcription programs. Consequently, dynamic alterations in higher-order DNA loops have a significant role in tissue-specific patterns of gene expression by bringing divergent gene segments into close proximity. However, the generality of enhancer-promoter looping remains to be understood. Furthermore, recent studies indicate that transcription is a stochastic process, with the number of mRNA molecules varying greatly in genetically homogeneous populations. Cell-to-cell variation was explained by assuming that transcription occurs in bursts of high activity followed by periods of inactivity. These fluctuations result in persistent cell individuality, thereby rendering a clonal population heterogeneous. By using genetically engineered animal models, we investigated the mechanisms underlying promoter-enhancer communication of *Bcl11b*, a transcription factor essential for $\alpha\beta$ T cell development, as well as the transcription characteristics of *Id3*, a key factor necessary for $\gamma\delta$ T cell fate choice of multipotent progenitors. The key findings of our studies are:

Objective I: CTCF dependent cis-regulatory interactions are important for *Bcl11b* expression and T cell development

- Mutation of CR4 greatly reduces CTCF binding and expression of *Bcl11b*
- CR4 mutation impairs T cell development in-vitro and in-vivo
- CR4 CTCF mediated promoter-enhancer interactions are important for activation of *Bcl11b* expression and development of T cells

Objective II: *Id3* transcriptional dynamics are directly related to TCR signal threshold

- Pre-TCR mediated signaling swiftly activates the *Id3* bursting frequency.
- Inhibition of ERK signaling interferes with *Id3* bursting frequencies.
- TCR $\gamma\delta$ induces higher *Id3* bursting frequencies as compared to TCR β .
- *Id3* bursting frequencies in $\gamma\delta$ T cells positively correlate with affinity for self-reactive ligands.

Chapter 6
BIBLIOGRAPHY

- Adams, E. J., Strop, P., Shin, S., Chien, Y. H., & Garcia, K. C. (2008). An autonomous CDR3delta is sufficient for recognition of the nonclassical MHC class I molecules T10 and T22 by gammadelta T cells. *Nature immunology*, 9(7), 777–784. <https://doi.org/10.1038/ni.1620>.
- Adolfsson, J., Borge, O. J., Bryder, D., Theilgaard-Mönch, K., Astrand-Grundström, I., Sitnicka, E., Sasaki, Y., & Jacobsen, S. E. (2001). Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity*, 15(4), 659–669. [https://doi.org/10.1016/s1074-7613\(01\)00220-5](https://doi.org/10.1016/s1074-7613(01)00220-5)
- Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O. J., Thoren, L. A., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M., & Jacobsen, S. E. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell*, 121(2), 295–306. <https://doi.org/10.1016/j.cell.2005.02.013>.
- Akashi, K., Traver, D., Miyamoto, T., & Weissman, I. L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature*, 404(6774), 193–197. <https://doi.org/10.1038/35004599>.
- Alberola-Ila, J., Forbush, K. A., Seger, R., Krebs, E. G., & Perlmutter, R. M. (1995). Selective requirement for MAP kinase activation in thymocyte differentiation. *Nature*, 373(6515), 620–623. <https://doi.org/10.1038/373620a0>.
- Alessi, D. R., Cuenda, A., Cohen, P., Dudley, D. T., & Saltiel, A. R. (1995). PD 098059 is a specific inhibitor of the activation of mitogen-activated protein kinase kinase in vitro and in vivo. *The Journal of biological chemistry*, 270(46), 27489–27494. <https://doi.org/10.1074/jbc.270.46.27489>.
- Arun, G., Diermeier, S., Akerman, M., Chang, K. C., Wilkinson, J. E., Hearn, S., Kim, Y., MacLeod, A. R., Krainer, A. R., Norton, L., Brogi, E., Egeblad, M., & Spector, D. L. (2016). Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. *Genes & development*, 30(1), 34–51. <https://doi.org/10.1101/gad.270959.115>.
- Aubrey, M., Warburg, Z. J., & Murre, C. (2022). Helix-Loop-Helix Proteins in Adaptive Immune Development. *Frontiers in immunology*, 13, 881656. <https://doi.org/10.3389/fimmu.2022.881656>.
- Bahar Halpern, K., Tanami, S., Landen, S., Chapal, M., Szlak, L., Hutzler, A., Nizhberg, A., & Itzkovitz, S. (2015). Bursty gene expression in the intact mammalian liver. *Molecular cell*, 58(1), 147–156. <https://doi.org/10.1016/j.molcel.2015.01.027>
- Bain, G., & Murre, C. (1998). The role of E-proteins in B- and T-lymphocyte development. *Seminars in immunology*, 10(2), 143–153. <https://doi.org/10.1006/smim.1998.0116>.
- Bain, G., & Murre, C. (1998). The role of E-proteins in B- and T-lymphocyte development. *Seminars in immunology*, 10(2), 143–153. <https://doi.org/10.1006/smim.1998.0116>
- Bain, G., Cravatt, C. B., Loomans, C., Alberola-Ila, J., Hedrick, S. M., & Murre, C. (2001). Regulation of the helix-loop-helix proteins, E2A and Id3, by the Ras-ERK MAPK cascade. *Nature immunology*, 2(2), 165–171. <https://doi.org/10.1038/84273>.
- Banerjee, A., Northrup, D., Boukarabila, H., Jacobsen, S. E., & Allman, D. (2013). Transcriptional repression of Gata3 is essential for early B cell commitment. *Immunity*, 38(5), 930–942. <https://doi.org/10.1016/j.immuni.2013.01.014>
- Banerji, J., Rusconi, S., & Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell*, 27(2 Pt 1), 299–308. [https://doi.org/10.1016/0092-8674\(81\)90413-x](https://doi.org/10.1016/0092-8674(81)90413-x).
- Barndt, R. J., Dai, M., & Zhuang, Y. (2000). Functions of E2A-HEB heterodimers in T-cell development revealed by a dominant negative mutation of HEB. *Molecular and cellular biology*, 20(18), 6677–6685. <https://doi.org/10.1128/MCB.20.18.6677-6685.2000>.

- Beach, D. L., Salmon, E. D., & Bloom, K. (1999). Localization and anchoring of mRNA in budding yeast. *Current biology: CB*, 9(11), 569–578. [https://doi.org/10.1016/s0960-9822\(99\)80260-7](https://doi.org/10.1016/s0960-9822(99)80260-7).
- Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., & Weintraub, H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell*, 61(1), 49–59. [https://doi.org/10.1016/0092-8674\(90\)90214-y](https://doi.org/10.1016/0092-8674(90)90214-y).
- Bentovim, L., Harden, T. T., & DePace, A. H. (2017). Transcriptional precision and accuracy in development: from measurements to models and mechanisms. *Development* (Cambridge, England), 144(21), 3855–3866. <https://doi.org/10.1242/dev.146563>.
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S. M., Singer, R. H., & Long, R. M. (1998). Localization of ASH1 mRNA particles in living yeast. *Molecular cell*, 2(4), 437–445. [https://doi.org/10.1016/s1097-2765\(00\)80143-4](https://doi.org/10.1016/s1097-2765(00)80143-4).
- Bertrand, J. Y., Kim, A. D., Violette, E. P., Stachura, D. L., Cisson, J. L., & Traver, D. (2007). Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. *Development* (Cambridge, England), 134(23), 4147–4156. <https://doi.org/10.1242/dev.012385>.
- Blake, W. J., Balázsi, G., Kohanski, M. A., Isaacs, F. J., Murphy, K. F., Kuang, Y., Cantor, C. R., Walt, D. R., & Collins, J. J. (2006). Phenotypic consequences of promoter-mediated transcriptional noise. *Molecular cell*, 24(6), 853–865. <https://doi.org/10.1016/j.molcel.2006.11.003>
- Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A. S., Yu, T., Marie-Nelly, H., McSwiggen, D. T., Kokic, G., Dailey, G. M., Cramer, P., Darzacq, X., & Zweckstetter, M. (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nature structural & molecular biology*, 25(9), 833–840. <https://doi.org/10.1038/s41594-018-0112-y>.
- Böiers, C., Carrelha, J., Lutteropp, M., Luc, S., Green, J. C., Azzoni, E., Woll, P. S., Mead, A. J., Hultquist, A., Swiers, G., Perdiguero, E. G., Macaulay, I. C., Melchiori, L., Luis, T. C., Kharazi, S., Bouriez-Jones, T., Deng, Q., Pontén, A., Atkinson, D., Jensen, C. T., ... Jacobsen, S. E. (2013). Lymphomyeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. *Cell stem cell*, 13(5), 535–548. <https://doi.org/10.1016/j.stem.2013.08.012>
- Boller, S., & Grosschedl, R. (2014). The regulatory network of B-cell differentiation: a focused view of early B-cell factor 1 function. *Immunological reviews*, 261(1), 102–115. <https://doi.org/10.1111/imr.12206>
- Bonneville, M., Ito, K., Krecko, E. G., Itoharu, S., Kappes, D., Ishida, I., Kanagawa, O., Janeway, C. A., Murphy, D. B., & Tonegawa, S. (1989). Recognition of a self-major histocompatibility complex TL region product by gamma delta T-cell receptors. *Proceedings of the National Academy of Sciences of the United States of America*, 86(15), 5928–5932. <https://doi.org/10.1073/pnas.86.15.5928>.
- Boos, M. D., Yokota, Y., Eberl, G., & Kee, B. L. (2007). Mature natural killer cell and lymphoid tissue-inducing cell development requires Id2-mediated suppression of E protein activity. *The Journal of experimental medicine*, 204(5), 1119–1130. <https://doi.org/10.1084/jem.20061959>.
- Bowles, J. R., Hoppe, C., Ashe, H. L., & Rattray, M. (2022). Scalable inference of transcriptional kinetic parameters from MS2 time series data. *Bioinformatics* (Oxford, England), 38(4), 1030–1036. <https://doi.org/10.1093/bioinformatics/btab765>.
- Boya, R., Yadavalli, A. D., Nikhat, S., Kurukuti, S., Palakodeti, D., & Pongubala, J. M. R. (2017). Developmentally regulated higher-order chromatin interactions orchestrate B cell fate commitment. *Nucleic acids research*, 45(19), 11070–11087. <https://doi.org/10.1093/nar/gkx722>

- Buxbaum, A. R., Wu, B., & Singer, R. H. (2014). Single β -actin mRNA detection in neurons reveals a mechanism for regulating its translatability. *Science* (New York, N.Y.), 343(6169), 419–422. <https://doi.org/10.1126/science.1242939>.
- Cai, L., Dalal, C. K., & Elowitz, M. B. (2008). Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature*, 455(7212), 485–490. <https://doi.org/10.1038/nature07292>
- Cantrell D. A. (2002). T-cell antigen receptor signal transduction. *Immunology*, 105(4), 369–374. <https://doi.org/10.1046/j.1365-2567.2002.01391.x>
- Carding, S. R., & Egan, P. J. (2002). Gammadelta T cells: functional plasticity and heterogeneity. *Nature reviews. Immunology*, 2(5), 336–345. <https://doi.org/10.1038/nri797>.
- Carleton, M., Ruetsch, N. R., Berger, M. A., Rhodes, M., Kaptik, S., & Wiest, D. L. (1999). Signals transduced by CD3epsilon, but not by surface pre-TCR complexes, are able to induce maturation of an early thymic lymphoma in vitro. *Journal of immunology Baltimore, Md. : 1950*, 163(5), 2576–2585.
- Catarino, R. R., & Stark, A. (2018). Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. *Genes & development*, 32(3-4), 202–223. <https://doi.org/10.1101/gad.310367.117>.
- Chen, B., Mao, B., Huang, S., Zhou, Y., Tsuji, K. & Ma, F. (2014). Human Embryonic Stem Cell-Derived Primitive and Definitive Hematopoiesis. Book> Pluripotent Stem Cell Biology-Advances in Mechanisms, Methods and Models, DOI: 10.5772/58628.
- Ciofani, M., Schmitt, T. M., Ciofani, A., Michie, A. M., Cuburu, N., Aublin, A., Maryanski, J. L., & Zúñiga-Pflücker, J. C. (2004). Obligatory role for cooperative signaling by pre-TCR and Notch during thymocyte differentiation. *Journal of immunology (Baltimore, Md.: 1950)*, 172(9), 5230–5239. <https://doi.org/10.4049/jimmunol.172.9.5230>.
- Coffey, F., Lee, S. Y., Buus, T. B., Lauritsen, J. P., Wong, G. W., Joachims, M. L., Thompson, L. F., Zúñiga-Pflücker, J. C., Kappes, D. J., & Wiest, D. L. (2014). The TCR ligand-inducible expression of CD73 marks $\gamma\delta$ lineage commitment and a metastable intermediate in effector specification. *The Journal of experimental medicine*, 211(2), 329–343. <https://doi.org/10.1084/jem.20131540>.
- Crowley, M. P., Fahrner, A. M., Baumgarth, N., Hampl, J., Gutgemann, I., Teyton, L., & Chien, Y. (2000). A population of murine gammadelta T cells that recognize an inducible MHC class Ib molecule. *Science* (New York, N.Y.), 287(5451), 314–316. <https://doi.org/10.1126/science.287.5451.314>.
- Csárdi, G., Franks, A., Choi, D. S., Airoidi, E. M., & Drummond, D. A. (2015). Accounting for experimental noise reveals that mRNA levels, amplified by post-transcriptional processes, largely determine steady-state protein levels in yeast. *PLoS genetics*, 11(5), e1005206. <https://doi.org/10.1371/journal.pgen.1005206>
- Dias, S., Månsson, R., Gurbuxani, S., Sigvardsson, M., & Kee, B. L. (2008). E2A proteins promote development of lymphoid-primed multipotent progenitors. *Immunity*, 29(2), 217–227. <https://doi.org/10.1016/j.immuni.2008.05.015>.
- Dimitrova, N., Zamudio, J. R., Jong, R. M., Soukup, D., Resnick, R., Sarma, K., et al. (2014). LincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint. *Mol. Cell*. 54, 777–790. doi: 10.1016/j.molcel. 2014.04.025.
- Dudley, D. T., Pang, L., Decker, S. J., Bridges, A. J., & Saltiel, A. R. (1995). A synthetic inhibitor of the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences of the United States of America*, 92(17), 7686–7689. <https://doi.org/10.1073/pnas.92.17.7686>.
- Dufourt, J., Trullo, A., Hunter, J., Fernandez, C., Lazaro, J., Dejean, M., Morales, L., Nait-Amer, S., Schulz, K. N., Harrison, M. M., Favard, C., Radulescu, O., & Lagha, M. (2018). Temporal control of gene expression by the pioneer factor Zelda through transient

- Engel, I., & Murre, C. (2001). The function of E- and Id proteins in lymphocyte development. *Nature reviews. Immunology*, 1(3), 193–199. <https://doi.org/10.1038/35105060>.
- Engel, I., & Murre, C. (2004). E2A proteins enforce a proliferation checkpoint in developing thymocytes. *The EMBO journal*, 23(1), 202–211. <https://doi.org/10.1038/sj.emboj.7600017>.
- Engreitz, J. M., Haines, J. E., Perez, E. M., Munson, G., Chen, J., Kane, M., McDonel, P. E., Guttman, M., & Lander, E. S. (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature*, 539(7629), 452–455. <https://doi.org/10.1038/nature20149>.
- Fahl, S.P., Coffey, F., MacCormack, D., Wiest, D. L. (2016). Control of early T cell development by Notch and T cell receptor signals, *Encyclopedia of Immunobiology*, 1, 234–241.
- Fang, B., Everett, L. J., Jager, J., Briggs, E., Armour, S. M., Feng, D., Roy, A., Gerhart-Hines, Z., Sun, Z., & Lazar, M. A. (2014). Circadian enhancers coordinate multiple phases of rhythmic gene transcription in vivo. *Cell*, 159(5), 1140–1152. <https://doi.org/10.1016/j.cell.2014.10.022>.
- Femino, A. M., Fay, F. S., Fogarty, K., & Singer, R. H. (1998). Visualization of single RNA transcripts in situ. *Science (New York, N.Y.)*, 280(5363), 585–590. <https://doi.org/10.1126/science.280.5363.585>.
- Ferraro, T., Lucas, T., Clémot, M., De Las Heras Chanes, J., Desponds, J., Coppey, M., Walczak, A. M., & Dostatni, N. (2016). New methods to image transcription in living fly embryos: the insights so far, and the prospects. *Wiley interdisciplinary reviews. Developmental biology*, 5(3), 296–310. <https://doi.org/10.1002/wdev.221>.
- Fuxa, M., & Skok, J. A. (2007). Transcriptional regulation in early B cell development. *Current opinion in immunology*, 19(2), 129–136. <https://doi.org/10.1016/j.coi.2007.02.002>.
- García-Ojeda, M. E., Klein Wolterink, R. G., Lemaître, F., Richard-Le Goff, O., Hasan, M., Hendriks, R. W., Cumano, A., & Di Santo, J. P. (2013). GATA-3 promotes T-cell specification by repressing B-cell potential in pro-T cells in mice. *Blood*, 121(10), 1749–1759. <https://doi.org/10.1182/blood-2012-06-440065>.
- Gorentla, B. K., & Zhong, X. P. (2012). T cell Receptor Signal Transduction in T lymphocytes. *Journal of clinical & cellular immunology*, 2012(Suppl 12), 5. <https://doi.org/10.4172/2155-9899.S12-005>
- Groff, A. F., Sanchez-Gomez, D. B., Soruco, M. M. L., Gerhardinger, C., Barutcu, A. R., Li, E., Elcavage, L., Plana, O., Sanchez, L. V., Lee, J. C., Sauvageau, M., & Rinn, J. L. (2016). In Vivo Characterization of Linc-p21 Reveals Functional cis-Regulatory DNA Elements. *Cell reports*, 16(8), 2178–2186. <https://doi.org/10.1016/j.celrep.2016.07.050>.
- Guo, C. J., Xu, G., & Chen, L. L. (2020). Mechanisms of Long Noncoding RNA Nuclear Retention. *Trends in biochemical sciences*, 45(11), 947–960. <https://doi.org/10.1016/j.tibs.2020.07.001>
- Guo, R., Wan, C. K., Carpenter, J. H., Mousallem, T., Boustany, R. M., Kuan, C. T., Burks, A. W., & Zhong, X. P. (2008). Synergistic control of T cell development and tumor suppression by diacylglycerol kinase alpha and zeta. *Proceedings of the National Academy of Sciences of the United States of America*, 105(33), 11909–11914. <https://doi.org/10.1073/pnas.0711856105>.
- Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M. C., Hung, T., Argani, P., Rinn, J. L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R. B., van de Vijver, M. J., Sukumar, S., & Chang, H. Y. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, 464(7291), 1071–1076. <https://doi.org/10.1038/nature08975>.

- Hah, N., Danko, C. G., Core, L., Waterfall, J. J., Siepel, A., Lis, J. T., & Kraus, W. L. (2011). A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell*, 145(4), 622–634. <https://doi.org/10.1016/j.cell.2011.03.042>.
- Haks, M. C., Lefebvre, J. M., Lauritsen, J. P., Carleton, M., Rhodes, M., Miyazaki, T., Kappes, D. J., & Wiest, D. L. (2005). Attenuation of gammadeltaTCR signaling efficiently diverts thymocytes to the alphabeta lineage. *Immunity*, 22(5), 595–606. <https://doi.org/10.1016/j.immuni.2005.04.003>.
- Hansen, A. S., Hsieh, T. S., Cattoglio, C., Pustova, I., Saldaña-Meyer, R., Reinberg, D., Darzacq, X., & Tjian, R. (2019). Distinct Classes of Chromatin Loops Revealed by Deletion of an RNA-Binding Region in CTCF. *Molecular cell*, 76(3), 395–411.e13. <https://doi.org/10.1016/j.molcel.2019.07.039>Henthorn, P.,
- Hayes, S. M., Li, L., & Love, P. E. (2005). TCR signal strength influences alphabeta/gammadelta lineage fate. *Immunity*, 22(5), 583–593. <https://doi.org/10.1016/j.immuni.2005.03.014>.
- Hock, H., Meade, E., Medeiros, S., Schindler, J. W., Valk, P. J., Fujiwara, Y., & Orkin, S. H. (2004). Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. *Genes & development*, 18(19), 2336–2341. <https://doi.org/10.1101/gad.1239604>
- Holmes, R., & Zúñiga-Pflücker, J. C. (2009). The OP9-DL1 system: generation of T-lymphocytes from embryonic or hematopoietic stem cells in vitro. *Cold Spring Harbor protocols*, 2009(2), pdb.prot5156. <https://doi.org/10.1101/pdb.prot5156>.
- Hornung, G., Bar-Ziv, R., Rosin, D., Tokuriki, N., Tawfik, D. S., Oren, M., & Barkai, N. (2012). Noise-mean relationship in mutated promoters. *Genome research*, 22(12), 2409–2417. <https://doi.org/10.1101/gr.139378.112>
- Hosokawa, H., Ungerback, J., Wang, X., Matsumoto, M., Nakayama, K. I., Cohen, S. M., Tanaka, T., & Rothenberg, E. V. (2018). Transcription Factor PU.1 Represses and Activates Gene Expression in Early T Cells by Redirecting Partner Transcription Factor Binding. *Immunity*, 48(6), 1119–1134.e7. <https://doi.org/10.1016/j.immuni.2018.04.024>.
- Hu, J. S., Olson, E. N., & Kingston, R. E. (1992). HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. *Molecular and cellular biology*, 12(3), 1031–1042. <https://doi.org/10.1128/mcb.12.3.1031-1042.1992>.
- Hubé, F., & Francastel, C. (2018). Coding and Non-coding RNAs, the Frontier Has Never Been So Blurred. *Frontiers in genetics*, 9, 140. <https://doi.org/10.3389/fgene.2018.00140>
- Hwang, J. R., Byeon, Y., Kim, D., & Park, S. G. (2020). Recent insights of T cell receptor-mediated signaling pathways for T cell activation and development. *Experimental & molecular medicine*, 52(5), 750–761. <https://doi.org/10.1038/s12276-020-0435-8>
- Ikawa, T., Hirose, S., Masuda, K., Kakugawa, K., Satoh, R., Shibano-Satoh, A., Kominami, R., Katsura, Y., & Kawamoto, H. (2010). An essential developmental checkpoint for production of the T cell lineage. *Science (New York, N.Y.)*, 329(5987), 93–96. <https://doi.org/10.1126/science.1188995>.
- Ikawa, T., Kawamoto, H., Goldrath, A. W., & Murre, C. (2006). E proteins and Notch signaling cooperate to promote T cell lineage specification and commitment. *The Journal of experimental medicine*, 203(5), 1329–1342. <https://doi.org/10.1084/jem.20060268>.
- interactions in hubs. *Nature communications*, 9(1), 5194. <https://doi.org/10.1038/s41467-018-07613-z>.
- Isoda, T., Moore, A. J., He, Z., Chandra, V., Aida, M., Denholtz, M., Piet van Hamburg, J., Fisch, K. M., Chang, A. N., Fahl, S. P., Wiest, D. L., & Murre, C. (2017). Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. *Cell*, 171(1), 103–119.e18. <https://doi.org/10.1016/j.cell.2017.09.001>.
- Ivanovs, A., Rybtsov, S., Welch, L., Anderson, R. A., Turner, M. L., & Medvinsky, A. (2011). Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-

- gonad-mesonephros region. *The Journal of experimental medicine*, 208(12), 2417–2427. <https://doi.org/10.1084/jem.20111688>
- Jameson, J., Ugarte, K., Chen, N., Yachi, P., Fuchs, E., Boismenu, R., & Havran, W. L. (2002). A role for skin gammadelta T cells in wound repair. *Science (New York, N.Y.)*, 296(5568), 747–749. <https://doi.org/10.1126/science.1069639>.
- Kaiser, P. D., Maier, J., Traenkle, B., Emele, F., & Rothbauer, U. (2014). Recent progress in generating intracellular functional antibody fragments to target and trace cellular components in living cells. *Biochimica et biophysica acta*, 1844(11), 1933–1942. <https://doi.org/10.1016/j.bbapap.2014.04.019>.
- Kent, D. G., Copley, M. R., Benz, C., Wöhrer, S., Dykstra, B. J., Ma, E., Cheyne, J., Zhao, Y., Bowie, M. B., Zhao, Y., Gasparetto, M., Delaney, A., Smith, C., Marra, M., & Eaves, C. J. (2009). Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. *Blood*, 113(25), 6342–6350. <https://doi.org/10.1182/blood-2008-12-192054>
- Kiledjian, M., & Kadesch, T. (1990). Two distinct transcription factors that bind the immunoglobulin enhancer microE5/kappa 2 motif. *Science (New York, N.Y.)*, 247(4941), 467–470. <https://doi.org/10.1126/science.2105528>.
- Kind, J., Pagie, L., de Vries, S. S., Nahidiazar, L., Dey, S. S., Bienko, M., Zhan, Y., Lajoie, B., de Graaf, C. A., Amendola, M., Fudenberg, G., Imakaev, M., Mirny, L. A., Jalink, K., Dekker, J., van Oudenaarden, A., & van Steensel, B. (2015). Genome-wide maps of nuclear lamina interactions in single human cells. *Cell*, 163(1), 134–147. <https://doi.org/10.1016/j.cell.2015.08.040>.
- Kreslavsky, T., Gleimer, M., Garbe, A. I., & von Boehmer, H. (2010). $\alpha\beta$ versus $\gamma\delta$ fate choice: counting the T-cell lineages at the branch point. *Immunological reviews*, 238(1), 169–181. <https://doi.org/10.1111/j.1600-065X.2010.00947.x>.
- Kuang, S., & Wang, L. (2020). Identification and analysis of consensus RNA motifs binding to the genome regulator CTCF. *NAR genomics and bioinformatics*, 2(2), lqaa031. <https://doi.org/10.1093/nargab/lqaa031>.
- Kueh, H. Y., Yui, M. A., Ng, K. K., Pease, S. S., Zhang, J. A., Damle, S. S., Freedman, G., Siu, S., Bernstein, I. D., Elowitz, M. B., & Rothenberg, E. V. (2016). Asynchronous combinatorial action of four regulatory factors activates Bcl11b for T cell commitment. *Nature immunology*, 17(8), 956–965. <https://doi.org/10.1038/ni.3514>.
- Lai, A. Y., & Kondo, M. (2008). T and B lymphocyte differentiation from hematopoietic stem cell. *Seminars in immunology*, 20(4), 207–212. <https://doi.org/10.1016/j.smim.2008.05.002>.
- Lang, K., & Chin, J. W. (2014). Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chemical reviews*, 114(9), 4764–4806. <https://doi.org/10.1021/cr400355w>.
- Larson, M. H., Gilbert, L. A., Wang, X., Lim, W. A., Weissman, J. S., & Qi, L. S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature protocols*, 8(11), 2180–2196. <https://doi.org/10.1038/nprot.2013.132>.
- Laslo, P., Pongubala, J. M., Lancki, D. W., & Singh, H. (2008). Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system. *Seminars in immunology*, 20(4), 228–235. <https://doi.org/10.1016/j.smim.2008.08.003>
- Lasorella, A., Benenzra, R., & Iavarone, A. (2014). The ID proteins: master regulators of cancer stem cells and tumour aggressiveness. *Nature reviews. Cancer*, 14(2), 77–91. <https://doi.org/10.1038/nrc3638>.
- Laurenti, E., & Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature*, 553(7689), 418–426. <https://doi.org/10.1038/nature25022>
- Lauritsen, J. P., Wong, G. W., Lee, S. Y., Lefebvre, J. M., Ciofani, M., Rhodes, M., Kappes, D. J., Zúñiga-Pflücker, J. C., & Wiest, D. L. (2009). Marked induction of the helix-loop-helix

- protein Id3 promotes the gammadelta T cell fate and renders their functional maturation Notch independent. *Immunity*, 31(4), 565–575. <https://doi.org/10.1016/j.immuni.2009.07.010>.
- Lee S., Kopp F., Chang T. C., Sataluri A., Chen B. B., Sivakumar S., et al. (2016). Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. *Cell* 164 69–80. [10.1016/j.cell.2015.12.017](https://doi.org/10.1016/j.cell.2015.12.017).
- Lee, J. T., & Bartolomei, M. S. (2013). X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell*, 152(6), 1308–1323. <https://doi.org/10.1016/j.cell.2013.02.016>
- Lee, R., Kang, M. K., Kim, Y. J., Yang, B., Shim, H., Kim, S., Kim, K., Yang, C. M., Min, B. G., Jung, W. J., Lee, E. C., Joo, J. S., Park, G., Cho, W. K., & Kim, H. P. (2022). CTCF-mediated chromatin looping provides a topological framework for the formation of phase-separated transcriptional condensates. *Nucleic acids research*, 50(1), 207–226. <https://doi.org/10.1093/nar/gkab1242>
- Lessel, D., Gehbauer, C., Bramswig, N. C., Schluth-Bolard, C., Venkataramanappa, S., van Gassen, K. L. I., Hempel, M., Haack, T. B., Baresic, A., Genetti, C. A., Funari, M. F. A., Lessel, I., Kuhlmann, L., Simon, R., Liu, P., Denecke, J., Kuechler, A., de Kruijff, I., Shoukier, M., Lek, M., Kubisch, C. (2018). BCL11B mutations in patients affected by a neurodevelopmental disorder with reduced type 2 innate lymphoid cells. *Brain : a journal of neurology*, 141(8), 2299–2311. <https://doi.org/10.1093/brain/awy173>
- Lettice, L. A., Heaney, S. J., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., Goode, D., Elgar, G., Hill, R. E., & de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Human molecular genetics*, 12(14), 1725–1735. <https://doi.org/10.1093/hmg/ddg180>.
- Li, L., Leid, M., & Rothenberg, E. V. (2010). An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. *Science (New York, N.Y.)*, 329(5987), 89–93. <https://doi.org/10.1126/science.1188989>.
- Li, L., Zhang, J. A., Dose, M., Kueh, H. Y., Mosadeghi, R., Gounari, F., & Rothenberg, E. V. (2013). A far downstream enhancer for murine Bcl11b controls its T-cell specific expression. *Blood*, 122(6), 902–911. <https://doi.org/10.1182/blood-2012-08-447839>
- Li, S., Feng, X., Li, T., Zhang, S., Zuo, Z., Lin, P., Konoplev, S., Bueso-Ramos, C. E., Vega, F., Medeiros, L. J., & Yin, C. C. (2013). Extranodal NK/T-cell lymphoma, nasal type: a report of 73 cases at MD Anderson Cancer Center. *The American journal of surgical pathology*, 37(1), 14–23. <https://doi.org/10.1097/PAS.0b013e31826731b5>.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (New York, N.Y.)*, 326(5950), 289–293. <https://doi.org/10.1126/science.1181369>.
- Lin, Y. C., Benner, C., Mansson, R., Heinz, S., Miyazaki, K., Miyazaki, M., Chandra, V., Bossen, C., Glass, C. K., & Murre, C. (2012). Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nature immunology*, 13(12), 1196–1204. <https://doi.org/10.1038/ni.2432>.
- Liu, H., Chi, A. W., Arnett, K. L., Chiang, M. Y., Xu, L., Shestova, O., Wang, H., Li, Y. M., Bhandoola, A., Aster, J. C., Blacklow, S. C., & Pear, W. S. (2010). Notch dimerization is required for leukemogenesis and T-cell development. *Genes & development*, 24(21), 2395–2407. <https://doi.org/10.1101/gad.1975210>.
- Liu, N., Ohnishi, N., Ni, L., Akira, S., & Bacon, K. B. (2003). CpG directly induces T-bet expression and inhibits IgG1 and IgE switching in B cells. *Nature immunology*, 4(7), 687–693. <https://doi.org/10.1038/ni941>

- Liu, S. J., & Lim, D. A. (2018). Modulating the expression of long non-coding RNAs for functional studies. *EMBO reports*, 19(12), e46955. <https://doi.org/10.15252/embr.201846955>.
- Liu, S. J., Nowakowski, T. J., Pollen, A. A., Lui, J. H., Horlbeck, M. A., Attenello, F. J., He, D., Weissman, J. S., Kriegstein, A. R., Diaz, A. A., & Lim, D. A. (2016). Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome biology*, 17, 67. <https://doi.org/10.1186/s13059-016-0932-1>.
- Loewer, S., Cabili, M. N., Guttman, M., Loh, Y. H., Thomas, K., Park, I. H., Garber, M., Curran, M., Onder, T., Agarwal, S., Manos, P. D., Datta, S., Lander, E. S., Schlaeger, T. M., Daley, G. Q., & Rinn, J. L. (2010). Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nature genetics*, 42(12), 1113–1117. <https://doi.org/10.1038/ng.710>
- Longabaugh, W. J. R., Zeng, W., Zhang, J. A., Hosokawa, H., Jansen, C. S., Li, L., Romero-Wolf, M., Liu, P., Kueh, H. Y., Mortazavi, A., & Rothenberg, E. V. (2017). Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network. *Proceedings of the National Academy of Sciences of the United States of America*, 114(23), 5800–5807. <https://doi.org/10.1073/pnas.1610617114>.
- Love, P. E., & Bhandoola, A. (2011). Signal integration and crosstalk during thymocyte migration and emigration. *Nature reviews. Immunology*, 11(7), 469–477. <https://doi.org/10.1038/nri2989>.
- Lukinavičius, G., Reymond, L., D'Este, E., Masharina, A., Göttfert, F., Ta, H., Güther, A., Fournier, M., Rizzo, S., Waldmann, H., Blaukopf, C., Sommer, C., Gerlich, D. W., Arndt, H. D., Hell, S. W., & Johnsson, K. (2014). Fluorogenic probes for live-cell imaging of the cytoskeleton. *Nature methods*, 11(7), 731–733. <https://doi.org/10.1038/nmeth.2972>.
- Lyden, D., Young, A. Z., Zagzag, D., Yan, W., Gerald, W., O'Reilly, R., Bader, B. L., Hynes, R. O., Zhuang, Y., Manova, K., & Benezra, R. (1999). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts. *Nature*, 401(6754), 670–677. <https://doi.org/10.1038/44334>.
- Marshall, J. S., Warrington, R., Watson, W., & Kim, H. L. (2018). An introduction to immunology and immunopathology. *Allergy, asthma, and clinical immunology : official journal of the Canadian Society of Allergy and Clinical Immunology*, 14(Suppl 2), 49. <https://doi.org/10.1186/s13223-018-0278-1>
- McGrath, K. E., Frame, J. M., Fegan, K. H., Bowen, J. R., Conway, S. J., Catherman, S. C., Kingsley, P. D., Koniski, A. D., & Palis, J. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. *Cell reports*, 11(12), 1892–1904. <https://doi.org/10.1016/j.celrep.2015.05.036>
- McGrath, K. E., Frame, J. M., Fromm, G. J., Koniski, A. D., Kingsley, P. D., Little, J., Bulger, M., & Palis, J. (2011). A transient definitive erythroid lineage with unique regulation of the β -globin locus in the mammalian embryo. *Blood*, 117(17), 4600–4608. <https://doi.org/10.1182/blood-2010-12-325357>.
- Mishra, R., Chen, A. T., Welsh, R. M., & Szomolanyi-Tsuda, E. (2010). NK cells and gammadelta T cells mediate resistance to polyomavirus-induced tumors. *PLoS pathogens*, 6(5), e1000924. <https://doi.org/10.1371/journal.ppat.1000924>.
- Miyazaki, K., Miyazaki, M., & Murre, C. (2014). The establishment of B versus T cell identity. *Trends in immunology*, 35(5), 205–210. <https://doi.org/10.1016/j.it.2014.02.009>
- Miyazaki, M., Miyazaki, K., Chen, K., Jin, Y., Turner, J., Moore, A. J., Saito, R., Yoshida, K., Ogawa, S., Rodewald, H. R., Lin, Y. C., Kawamoto, H., & Murre, C. (2017). The E-Id Protein Axis Specifies Adaptive Lymphoid Cell Identity and Suppresses Thymic Innate Lymphoid Cell Development. *Immunity*, 46(5), 818–834.e4. <https://doi.org/10.1016/j.immuni.2017.04.022>.
- Miyazaki, M., Miyazaki, K., Chen, S., Chandra, V., Wagatsuma, K., Agata, Y., Rodewald, H. R., Saito, R., Chang, A. N., Varki, N., Kawamoto, H., & Murre, C. (2015). The E-Id protein

- axis modulates the activities of the PI3K-AKT-mTORC1-Hif1a and c-myc/p19Arf pathways to suppress innate variant TFH cell development, thymocyte expansion, and lymphomagenesis. *Genes & development*, 29(4), 409–425. <https://doi.org/10.1101/gad.255331.114>.
- Miyazaki, M., Rivera, R. R., Miyazaki, K., Lin, Y. C., Agata, Y., & Murre, C. (2011). The opposing roles of the transcription factor E2A and its antagonist Id3 that orchestrate and enforce the naive fate of T cells. *Nature immunology*, 12(10), 992–1001. <https://doi.org/10.1038/ni.2086>.
- Morita, C. T., Jin, C., Sarikonda, G., & Wang, H. (2007). Nonpeptide antigens, presentation mechanisms, and immunological memory of human Vgamma2Vdelta2 T cells: discriminating friend from foe through the recognition of prenyl pyrophosphate antigens. *Immunological reviews*, 215, 59–76. <https://doi.org/10.1111/j.1600-065X.2006.00479.x>.
- Munk, M. E., Elser, C., & Kaufmann, S. H. (1996). Human gamma/delta T-cell response to *Listeria monocytogenes* protein components in vitro. *Immunology*, 87(2), 230–235. <https://doi.org/10.1046/j.1365-2567.1996.470549.x>.
- Murre C. (2019). Helix-loop-helix proteins and the advent of cellular diversity: 30 years of discovery. *Genes & development*, 33(1-2), 6–25. <https://doi.org/10.1101/gad.320663.118>.
- Murre, C., McCaw, P. S., & Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell*, 56(5), 777–783. [https://doi.org/10.1016/0092-8674\(89\)90682-x](https://doi.org/10.1016/0092-8674(89)90682-x).
- Naito, T., Tanaka, H., Naoe, Y., & Taniuchi, I. (2011). Transcriptional control of T-cell development. *International immunology*, 23(11), 661–668. <https://doi.org/10.1093/intimm/dxr078>.
- Nechanitzky, R., Akbas, D., Scherer, S., Györy, I., Hoyler, T., Ramamoorthy, S., Diefenbach, A., & Grosschedl, R. (2013). Transcription factor EBF1 is essential for the maintenance of B cell identity and prevention of alternative fates in committed cells. *Nature immunology*, 14(8), 867–875. <https://doi.org/10.1038/ni.2641>.
- Nelles, D. A., Fang, M. Y., O'Connell, M. R., Xu, J. L., Markmiller, S. J., Doudna, J. A., & Yeo, G. W. (2016). Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell*, 165(2), 488–496. <https://doi.org/10.1016/j.cell.2016.02.054>.
- Nikhat, S., Yadavalli, A. D., Prusty, A., Narayan, P. K., Palakodeti, D., Murre, C., & Pongubala, J. M. R. (2021). A regulatory network of microRNAs confers lineage commitment during early developmental trajectories of B and T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 118(46), e2104297118. <https://doi.org/10.1073/pnas.2104297118>.
- Orkin, S. H., & Zon, L. I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. *Cell*, 132(4), 631–644. <https://doi.org/10.1016/j.cell.2008.01.025>.
- Pagès, G., Guérin, S., Grall, D., Bonino, F., Smith, A., Anjuere, F., Auberger, P., & Pouyssegur, J. (1999). Defective thymocyte maturation in p44 MAP kinase (Erk 1) knockout mice. *Science (New York, N.Y.)*, 286(5443), 1374–1377. <https://doi.org/10.1126/science.286.5443.1374>.
- Pai, S. Y., Truitt, M. L., Ting, C. N., Leiden, J. M., Glimcher, L. H., & Ho, I. C. (2003). Critical roles for transcription factor GATA-3 in thymocyte development. *Immunity*, 19(6), 863–875. [https://doi.org/10.1016/s1074-7613\(03\)00328-5](https://doi.org/10.1016/s1074-7613(03)00328-5).
- Paralkar, V. R., Taborda, C. C., Huang, P., Yao, Y., Kossenkova, A. V., Prasad, R., Luan, J., Davies, J. O., Hughes, J. R., Hardison, R. C., Blobel, G. A., & Weiss, M. J. (2016). Unlinking an lncRNA from Its Associated cis Element. *Molecular cell*, 62(1), 104–110. <https://doi.org/10.1016/j.molcel.2016.02.029>.
- Park, H. Y., Lim, H., Yoon, Y. J., Follenzi, A., Nwokafor, C., Lopez-Jones, M., Meng, X., & Singer, R. H. (2014). Visualization of dynamics of single endogenous mRNA labeled in

- live mouse. *Science* (New York, N.Y.), 343(6169), 422–424. <https://doi.org/10.1126/science.1239200>
- Park, I. K., Qian, D., Kiel, M., Becker, M. W., Pihalja, M., Weissman, I. L., Morrison, S. J., & Clarke, M. F. (2003). Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature*, 423(6937), 302–305. <https://doi.org/10.1038/nature01587>
- Parker, M. E., & Ciofani, M. (2020). Regulation of $\gamma\delta$ T Cell Effector Diversification in the Thymus. *Frontiers in immunology*, 11, 42. <https://doi.org/10.3389/fimmu.2020.00042>.
- Pennacchio, L. A., Bickmore, W., Dean, A., Nobrega, M. A., & Bejerano, G. (2013). Enhancers: five essential questions. *Nature reviews. Genetics*, 14(4), 288–295. <https://doi.org/10.1038/nrg3458>.
- Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S. W., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R. M., van Lohuizen, M., Reinders, M., Wessels, L., & van Steensel, B. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Molecular cell*, 38(4), 603–613. <https://doi.org/10.1016/j.molcel.2010.03.016>.
- Petrie, H. T., & Zúñiga-Pflücker, J. C. (2007). Zoned out: functional mapping of stromal signaling microenvironments in the thymus. *Annual review of immunology*, 25, 649–679. <https://doi.org/10.1146/annurev.immunol.23.021704.115715>.
- Pongubala, J. M. R., & Murre, C. (2021). Spatial Organization of Chromatin: Transcriptional Control of Adaptive Immune Cell Development. *Frontiers in immunology*, 12, 633825. <https://doi.org/10.3389/fimmu.2021.633825>.
- Pongubala, J. M., Northrup, D. L., Lancki, D. W., Medina, K. L., Treiber, T., Bertolino, E., Thomas, M., Grosschedl, R., Allman, D., & Singh, H. (2008). Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. *Nature immunology*, 9(2), 203–215. <https://doi.org/10.1038/ni1555>
- Pui, J. C., Allman, D., Xu, L., DeRocco, S., Karnell, F. G., Bakkour, S., Lee, J. Y., Kadesch, T., Hardy, R. R., Aster, J. C., & Pear, W. S. (1999). Notch1 expression in early lymphopoiesis influences B versus T lineage determination. *Immunity*, 11(3), 299–308. [https://doi.org/10.1016/s1074-7613\(00\)80105-3](https://doi.org/10.1016/s1074-7613(00)80105-3)
- Quang, C. T., Zaniboni, B., & Ghysdael, J. (2017). A TCR-switchable cell death pathway in T-ALL. *Oncoscience*, 4(3-4), 17–18. <https://doi.org/10.18632/oncoscience.342>
- Radtke, F., Wilson, A., Stark, G., Bauer, M., van Meerwijk, J., MacDonald, H. R., & Aguet, M. (1999). Deficient T cell fate specification in mice with an induced inactivation of Notch1. *Immunity*, 10(5), 547–558. [https://doi.org/10.1016/s1074-7613\(00\)80054-0](https://doi.org/10.1016/s1074-7613(00)80054-0)
- Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A., & Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nature methods*, 5(10), 877–879. <https://doi.org/10.1038/nmeth.1253>.
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature protocols*, 8(11), 2281–2308. <https://doi.org/10.1038/nprot.2013.143>.
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., & Chang, H. Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, 129(7), 1311–1323. <https://doi.org/10.1016/j.cell.2007.05.022>
- Rivera, R., & Murre, C. (2001). The regulation and function of the Id proteins in lymphocyte development. *Oncogene*, 20(58), 8308–8316. <https://doi.org/10.1038/sj.onc.1205091>.
- Rodriguez, J., & Larson, D. R. (2020). Transcription in Living Cells: Molecular Mechanisms of Bursting. *Annual review of biochemistry*, 89, 189–212. <https://doi.org/10.1146/annurev-biochem-011520-105250>.

- Rothenberg E. V. (2014). Transcriptional control of early T and B cell developmental choices. *Annual review of immunology*, 32, 283–321. <https://doi.org/10.1146/annurev-immunol-032712-100024>.
- Rothenberg, E. V., & Taghon, T. (2005). Molecular genetics of T cell development. *Annual review of immunology*, 23, 601–649. <https://doi.org/10.1146/annurev.immunol.23.021704.115737>.
- Rothenberg, E. V., Moore, J. E., & Yui, M. A. (2008). Launching the T-cell-lineage developmental programme. *Nature reviews. Immunology*, 8(1), 9–21. <https://doi.org/10.1038/nri2232>.
- Sallusto, F., Geginat, J., & Lanzavecchia, A. (2004). Central memory and effector memory T cell subsets: function, generation, and maintenance. *Annual review of immunology*, 22, 745–763. <https://doi.org/10.1146/annurev.immunol.22.012703.104702>
- Sambandam, A., Maillard, I., Zediak, V. P., Xu, L., Gerstein, R. M., Aster, J. C., Pear, W. S., & Bhandoola, A. (2005). Notch signaling controls the generation and differentiation of early T lineage progenitors. *Nature immunology*, 6(7), 663–670. <https://doi.org/10.1038/ni1216>
- Samelson L. E. (2002). Signal transduction mediated by the T cell antigen receptor: the role of adapter proteins. *Annual review of immunology*, 20, 371–394. <https://doi.org/10.1146/annurev.immunol.20.092601.111357>
- Schmid, M., Durussel, T., & Laemmli, U. K. (2004). ChIC and ChEC; genomic mapping of chromatin proteins. *Molecular cell*, 16(1), 147–157. <https://doi.org/10.1016/j.molcel.2004.09.007>.
- Schmitt, T. M., & Zúñiga-Pflücker, J. C. (2002). Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro. *Immunity*, 17(6), 749–756. [https://doi.org/10.1016/s1074-7613\(02\)00474-0](https://doi.org/10.1016/s1074-7613(02)00474-0).
- Scripture-Adams, D. D., Damle, S. S., Li, L., Elihu, K. J., Qin, S., Arias, A. M., Butler, R. R., 3rd, Champhekar, A., Zhang, J. A., & Rothenberg, E. V. (2014). GATA-3 dose-dependent checkpoints in early T cell commitment. *Journal of immunology (Baltimore, Md. : 1950)*, 193(7), 3470–3491. <https://doi.org/10.4049/jimmunol.1301663>
- Singh, H., DeKoter, R. P., & Walsh, J. C. (1999). PU.1, a shared transcriptional regulator of lymphoid and myeloid cell fates. *Cold Spring Harbor symposia on quantitative biology*, 64, 13–20. <https://doi.org/10.1101/sqb.1999.64.13>
- Singh, H., Medina, K. L., & Pongubala, J. M. (2005). Contingent gene regulatory networks and B cell fate specification. *Proceedings of the National Academy of Sciences of the United States of America*, 102(14), 4949–4953. <https://doi.org/10.1073/pnas.0500480102>
- Skene, P. J., & Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife*, 6, e21856. <https://doi.org/10.7554/eLife.21856>.
- Smale S. T. (2014). Transcriptional regulation in the immune system: a status report. *Trends in immunology*, 35(5), 190–194. <https://doi.org/10.1016/j.it.2014.03.003>
- Singer, Z. S., Yong, J., Tischler, J., Hackett, J. A., Altinok, A., Surani, M. A., Cai, L., & Elowitz, M. B. (2014). Dynamic heterogeneity and DNA methylation in embryonic stem cells. *Molecular cell*, 55(2), 319–331. <https://doi.org/10.1016/j.molcel.2014.06.029>
- Snijder, B., Sacher, R., Rämö, P., Liberali, P., Mench, K., Wolfrum, N., Burleigh, L., Scott, C. C., Verheije, M. H., Mercer, J., Moese, S., Heger, T., Theusner, K., Jurgeit, A., Lamparter, D., Balistreri, G., Schelhaas, M., De Haan, C. A., Marjomäki, V., Hyypiä, T., ... Pelkmans, L. (2012). Single-cell analysis of population context advances RNAi screening at multiple levels. *Molecular systems biology*, 8, 579. <https://doi.org/10.1038/msb.2012.9>.
- Spangrude G. J. (1989). Enrichment of murine haemopoietic stem cells: diverging roads. *Immunology today*, 10(10), 344–350. [https://doi.org/10.1016/0167-5699\(89\)90192-8](https://doi.org/10.1016/0167-5699(89)90192-8).
- Specht, E. A., Braselmann, E., & Palmer, A. E. (2017). A Critical and Comparative Review of Fluorescent Tools for Live-Cell Imaging. *Annual review of physiology*, 79, 93–117. <https://doi.org/10.1146/annurev-physiol-022516-034055>.

- Starmer, J., & Magnuson, T. (2009). A new model for random X chromosome inactivation. *Development* (Cambridge, England), 136(1), 1–10. <https://doi.org/10.1242/dev.025908>
- Sun, M., Schwalb, B., Schulz, D., Pirkel, N., Etzold, S., Larivière, L., Maier, K. C., Seizl, M., Tresch, A., & Cramer, P. (2012). Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome research*, 22(7), 1350–1359. <https://doi.org/10.1101/gr.130161.111>.
- Taleblian, L., Li, Z., Guo, Y., Gaudet, J., Speck, M. E., Sugiyama, D., Kaur, P., Pear, W. S., Maillard, I., & Speck, N. A. (2007). T-lymphoid, megakaryocyte, and granulocyte development are sensitive to decreases in CBFbeta dosage. *Blood*, 109(1), 11–21. <https://doi.org/10.1182/blood-2006-05-021188>.
- Tanaka, Y., Morita, C. T., Tanaka, Y., Nieves, E., Brenner, M. B., & Bloom, B. R. (1995). Natural and synthetic non-peptide antigens recognized by human gamma delta T cells. *Nature*, 375(6527), 155–158. <https://doi.org/10.1038/375155a0>.
- Thanos, D., & Maniatis, T. (1995). Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell*, 83(7), 1091–1100. [https://doi.org/10.1016/0092-8674\(95\)90136-1](https://doi.org/10.1016/0092-8674(95)90136-1).
- Thompson, P. K., & Zúñiga-Pflücker, J. C. (2011). On becoming a T cell, a convergence of factors kick it up a Notch along the way. *Seminars in immunology*, 23(5), 350–359. <https://doi.org/10.1016/j.smim.2011.08.007>.
- Tutucci, E., Vera, M., Biswas, J., Garcia, J., Parker, R., & Singer, R. H. (2018). An improved MS2 system for accurate reporting of the mRNA life cycle. *Nature methods*, 15(1), 81–89. <https://doi.org/10.1038/nmeth.4502>.
- Ushiki, A., Zhang, Y., Xiong, C., Zhao, J., Georgakopoulos-Soares, I., Kane, L., Jamieson, K., Bamshad, M. J., Nickerson, D. A., University of Washington Center for Mendelian Genomics, Shen, Y., Lettice, L. A., Silveira-Lucas, E. L., Petit, F., & Ahituv, N. (2021). Deletion of CTCF sites in the SHH locus alters enhancer-promoter interactions and leads to acheiropodia. *Nature communications*, 12(1), 2282. <https://doi.org/10.1038/s41467-021-22470-z>
- Wang, K. C., & Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAs. *Molecular cell*, 43(6), 904–914. <https://doi.org/10.1016/j.molcel.2011.08.018>
- Weber, B. N., Chi, A. W., Chavez, A., Yashiro-Ohtani, Y., Yang, Q., Shestova, O., & Bhandoola, A. (2011). A critical role for TCF-1 in T-lineage specification and differentiation. *Nature*, 476(7358), 63–68. <https://doi.org/10.1038/nature10279>.
- Wilson, A., MacDonald, H. R., & Radtke, F. (2001). Notch 1-deficient common lymphoid precursors adopt a B cell fate in the thymus. *The Journal of experimental medicine*, 194(7), 1003–1012. <https://doi.org/10.1084/jem.194.7.1003>
- Wilson, A., Murphy, M. J., Oskarsson, T., Kaloulis, K., Bettess, M. D., Oser, G. M., Pasche, A. C., Knabenhans, C., Macdonald, H. R., & Trumpp, A. (2004). c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. *Genes & development*, 18(22), 2747–2763. <https://doi.org/10.1101/gad.313104>
- Wu, B., Buxbaum, A. R., Katz, Z. B., Yoon, Y. J., & Singer, R. H. (2015). Quantifying Protein-mRNA Interactions in Single Live Cells. *Cell*, 162(1), 211–220. <https://doi.org/10.1016/j.cell.2015.05.054>.
- Yan, W., Young, A. Z., Soares, V. C., Kelley, R., Benezra, R., & Zhuang, Y. (1997). High incidence of T-cell tumors in E2A-null mice and E2A/Id1 double-knockout mice. *Molecular and cellular biology*, 17(12), 7317–7327. <https://doi.org/10.1128/MCB.17.12.7317>.
- Yang, Z., Jiang, X., Jiang, X., & Zhao, H. (2018). X-inactive-specific transcript: A long noncoding RNA with complex roles in human cancers. *Gene*, 679, 28–35. <https://doi.org/10.1016/j.gene.2018.08.071>

- Yokota, Y., Mansouri, A., Mori, S., Sugawara, S., Adachi, S., Nishikawa, S., & Gruss, P. (1999). Development of peripheral lymphoid organs and natural killer cells depends on the helix-loop-helix inhibitor Id2. *Nature*, 397(6721), 702–706. <https://doi.org/10.1038/17812>.
- Yu, J., Xiao, J., Ren, X., Lao, K., & Xie, X. S. (2006). Probing gene expression in live cells, one protein molecule at a time. *Science (New York, N.Y.)*, 311(5767), 1600–1603. <https://doi.org/10.1126/science.1119623>
- Yui MA, Rothenberg EV. Developmental gene networks: a triathlon on the course to T cell identity. *Nat Rev Immunol*. 2014 Aug;14(8):529-45. doi: 10.1038/nri3702. PMID: 25060579; PMCID: PMC4153685.
- Zandi, S., Mansson, R., Tsapogas, P., Zetterblad, J., Bryder, D., & Sigvardsson, M. (2008). EBF1 is essential for B-lineage priming and establishment of a transcription factor network in common lymphoid progenitors. *Journal of immunology (Baltimore, Md. : 1950)*, 181(5), 3364–3372. <https://doi.org/10.4049/jimmunol.181.5.3364>
- Zarin, P., Chen, E. L., In, T. S., Anderson, M. K., & Zúñiga-Pflücker, J. C. (2015). Gamma delta T-cell differentiation and effector function programming, TCR signal strength, when and how much?. *Cellular immunology*, 296(1), 70–75. <https://doi.org/10.1016/j.cellimm.2015.03.007>.
- Zarin, P., Wong, G. W., Mohtashami, M., Wiest, D. L., & Zúñiga-Pflücker, J. C. (2014). Enforcement of $\gamma\delta$ -lineage commitment by the pre-T-cell receptor in precursors with weak $\gamma\delta$ -TCR signals. *Proceedings of the National Academy of Sciences of the United States of America*, 111(15), 5658–5663. <https://doi.org/10.1073/pnas.1312872111>.
- Zhang, J., Jima, D. D., Jacobs, C., Fischer, R., Gottwein, E., Huang, G., Lugar, P. L., Lagoo, A. S., Rizzieri, D. A., Friedman, D. R., Weinberg, J. B., Lipsky, P. E., & Dave, S. S. (2009). Patterns of microRNA expression characterize stages of human B-cell differentiation. *Blood*, 113(19), 4586–4594. <https://doi.org/10.1182/blood-2008-09-178186>
- Zhao, Y., & Yang, Y. (2015). Profiling metabolic states with genetically encoded fluorescent biosensors for NADH. *Current opinion in biotechnology*, 31, 86–92. <https://doi.org/10.1016/j.copbio.2014.08.007>.
- Zhou, W., Yui, M. A., Williams, B. A., Yun, J., Wold, B. J., Cai, L., & Rothenberg, E. V. (2019). Single-Cell Analysis Reveals Regulatory Gene Expression Dynamics Leading to Lineage Commitment in Early T Cell Development. *Cell systems*, 9(4), 321–337.e9. <https://doi.org/10.1016/j.cels.2019.09.008>.
- Zook, E. C., Li, Z. Y., Xu, Y., de Pooter, R. F., Verykokakis, M., Beaulieu, A., Lasorella, A., Maienschein-Cline, M., Sun, J. C., Sigvardsson, M., & Kee, B. L. (2018). Transcription factor ID2 prevents E proteins from enforcing a naïve T lymphocyte gene program during NK cell development. *Science immunology*, 3(22), eaao2139. <https://doi.org/10.1126/sciimmunol.aao2139>.
- Zúñiga-Pflücker J. C. (2004). T-cell development made simple. *Nature reviews. Immunology*, 4(1), 67–72. <https://doi.org/10.1038/nri1257>

Chapter 7

APPENDIX

8.1. PUBLICATION

Nikhat, S., Yadavalli, A. D., Prusty, A., Narayan, P. K., Palakodeti, D., Murre, C., & Pongubala, J. M. R. (2021). A regulatory network of microRNAs confers lineage commitment during early developmental trajectories of B and T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 118(46), e2104297118. <https://doi.org/10.1073/pnas.2104297118>

8.2. PLAGIARISM REPORT

Attached on Next page.

Probing long-range genomic interactions and gene expression dynamics that instruct T cell fate

by Arpita Prusty



Librarian
Indira Gandhi Memorial Library
UNIVERSITY OF HYDERABAD
Central University P.O.
HYDERABAD-500 046.

Submission date: 10-Apr-2023 02:41PM (UTC+0530)

Submission ID: 2060425581

File name: Arpita_Prusty.pdf (648.2K)

Word count: 17734

Character count: 96922

Probing long-range genomic interactions and gene expression dynamics that instruct T cell fate

ORIGINALITY REPORT

21%

SIMILARITY INDEX

11%

INTERNET SOURCES

15%

PUBLICATIONS

6%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to University of Hyderabad,
Hyderabad

Student Paper

5%

I M R N Y

2

escholarship.org

Internet Source

2%

I M R N Y

3

Takeshi Isoda, Amanda J. Moore, Zhaoren He,
Vivek Chandra et al. "Non-coding
Transcription Instructs Chromatin Folding and
Compartmentalization to Dictate Enhancer-
Promoter Communication and T Cell Fate",
Cell, 2017

Publication

1%

I M R N Y

4

www.annualreviews.org

Internet Source

1%

5

genesdev.cshlp.org

Internet Source

1%

6

academic.oup.com

Internet Source

1%

7

www.ncbi.nlm.nih.gov

The source items
from 1-6 covers publications
from my lab, universal
nomenclature of cell phenotypes,
& me from I M R N Y
Dr. Jagan Pondubata
Department of Applied
School of Life Sciences
University of Hyderabad
Gachibowli, Hyderabad-500 046.

8

www.researchgate.net

Internet Source

1 %

9

Taras Kreslavsky. "αβ versus γδ fate choice: counting the T-cell lineages at the branch point : αβ versus γδ fate choice", Immunological Reviews, 11/2010

Publication

1 %

10

Laslo, P.. "Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system", Seminars in Immunology, 200808

Publication

1 %

11

Lital Bentovim, Timothy T. Harden, Angela H. DePace. "Transcriptional precision and accuracy in development: from measurements to models and mechanisms", Development, 2017

Publication

<1 %

12

www.nature.com

Internet Source

<1 %

13

Anshika Chowdhary, Venkata Satagopam, Reinhard Schneider. "Long Non-coding RNAs: Mechanisms, Experimental, and Computational Approaches in Identification, Characterization, and Their Biomarker

<1 %

Potential in Cancer", Frontiers in Genetics, 2021

Publication

14

Elizabeth A. Specht, Esther Braselmann, Amy E. Palmer. "A Critical and Comparative Review of Fluorescent Tools for Live-Cell Imaging", Annual Review of Physiology, 2017

Publication

<1 %

15

authors.library.caltech.edu

Internet Source

<1 %

16

www.jbc.org

Internet Source

<1 %

17

www.pnas.org

Internet Source

<1 %

18

hdl.handle.net

Internet Source

<1 %

19

"Poster Sessions", European Journal of Immunology, 09/2009

Publication

<1 %

20

dokumen.pub

Internet Source

<1 %

21

Bain, G.. "The role of E-proteins in B- and T-lymphocyte development", Seminars in Immunology, 199804

Publication

<1 %

22 Shawn P. Fahl, Francis Coffey, Dermot MacCormack, David L. Wiest. "Control of Early T Cell Development by Notch and T Cell Receptor Signals", Elsevier BV, 2016

Publication

<1 %

23 Joseph Rodriguez, Gang Ren, Christopher R. Day, Keji Zhao, Carson C. Chow, Daniel R. Larson. "Intrinsic Dynamics of a Human Gene Reveal the Basis of Expression Heterogeneity", Cell, 2019

Publication

<1 %

24 www.science.gov

Internet Source

<1 %

25 www.science.org

Internet Source

<1 %

26 Fiona M. Docherty, Kent A. Riemondy, Roberto Castro-Gutierrez, JaeAnn M. Dwulet et al. "ENTPD3 Marks Mature Stem Cell-Derived β -Cells Formed by Self-Aggregation In Vitro", Diabetes, 2021

Publication

<1 %

27 estudogeral.sib.uc.pt

Internet Source

<1 %

28 Tineke L. Lenstra, Joseph Rodriguez, Huimin Chen, Daniel R. Larson. "Transcription Dynamics in Living Cells", Annual Review of Biophysics, 2016

<1 %

29

biotechnologyforbiofuels.biomedcentral.com

Internet Source

<1 %

30

escholarship.umassmed.edu

Internet Source

<1 %

31

jdc.jefferson.edu

Internet Source

<1 %

32

Isaac Engel. "THE FUNCTION OF E- AND ID PROTEINS IN LYMPHOCYTE DEVELOPMENT", Nature Reviews Immunology, 12/2001

Publication

<1 %

33

"Epigenetic Regulation of Lymphocyte Development", Springer Science and Business Media LLC, 2012

Publication

<1 %

34

Catherine V. Laiosa, Matthias Stadtfeld, Thomas Graf. "DETERMINANTS OF LYMPHOID-MYELOID LINEAGE DIVERSIFICATION", Annual Review of Immunology, 2006

Publication

<1 %

35

Takaya Abe, Ken-ichi Inoue, Yasuhide Furuta, Hiroshi Kiyonari. "Pronuclear Microinjection during S-Phase Increases the Efficiency of CRISPR-Cas9-Assisted Knockin of Large DNA Donors in Mouse Zygotes", Cell Reports, 2020

Publication

<1 %

- | | | |
|----|---|------|
| 36 | Maria Ciofani, Juan Carlos Zúñiga-Pflücker.
"The Thymus as an Inductive Site for T
Lymphopoiesis", Annual Review of Cell and
Developmental Biology, 2007
Publication | <1 % |
| 37 | www.frontiersin.org
Internet Source | <1 % |
| 38 | www.jimmunol.org
Internet Source | <1 % |
| 39 | topics.sciencedirect.com
Internet Source | <1 % |
| 40 | www.stanfordchildrens.org
Internet Source | <1 % |
| 41 | Hiroyuki Hosokawa, Ellen V. Rothenberg.
"Cytokines, Transcription Factors, and the
Initiation of T-Cell Development", Cold Spring
Harbor Perspectives in Biology, 2018
Publication | <1 % |
| 42 | April M. Fischer, Carol D. Katayama, Giles
Pagès, Jacques Pouyssegur, Stephen M.
Hedrick. "The Role of Erk1 and Erk2 in
Multiple Stages of T Cell Development",
Immunity, 2005
Publication | <1 % |
| 43 | Ellen V. Rothenberg. "Programming for T-
lymphocyte fates: modularity and
mechanisms", Genes & Development, 2019 | <1 % |

44

Shawn P. Fahl, Alejandra V. Contreras, Anjali Verma, Xiang Qiu et al. "The E protein-TCF1 axis controls $\gamma\delta$ Tcell development and effector fate", Cell Reports, 2021

Publication

<1 %

45

David C. Klein, Sarah J. Hainer. "Genomic methods in profiling DNA accessibility and factor localization", Chromosome Research, 2019

Publication

<1 %

46

Jonathan R. Bowles, Caroline Hoppe, Hilary L. Ashe, Magnus Rattray. "Scalable inference of transcriptional kinetic parameters from MS2 time series data", Cold Spring Harbor Laboratory, 2020

Publication

<1 %

47

Reinhard Hohlfeld. "Does inflammation stimulate remyelination?", Journal of Neurology, 2007

Publication

<1 %

48

Submitted to The University of Manchester

Student Paper

<1 %

49

Masaki Miyazaki, Kazuko Miyazaki, Shuwen Chen, Manami Itoi et al. "Id2 and Id3 maintain the regulatory T cell pool to suppress

<1 %

inflammatory disease", Nature Immunology,
2014

Publication

50

Spooner, C.J.. "A Recurrent Network Involving
the Transcription Factors PU.1 and Gfi1
Orchestrates Innate and Adaptive Immune
Cell Fates", Immunity, 20091016

Publication

<1 %

51

dspace.mit.edu

Internet Source

<1 %

52

repub.eur.nl

Internet Source

<1 %

53

www.biorxiv.org

Internet Source

<1 %

54

www.i-scholar.in

Internet Source

<1 %

55

Banerjee, Anupam, Daniel Northrup, Hanane
Boukarabila, StenErikW. Jacobsen, and David
Allman. "Transcriptional Repression of Gata3
Is Essential for Early B Cell Commitment",
Immunity, 2013.

Publication

<1 %

56

David-Fung, E.S.. "Transcription factor
expression dynamics of early T-lymphocyte
specification and commitment",
Developmental Biology, 20090115

Publication

<1 %

57	Florian Wiede, Jarrod A. Dudakov, Kun-Hui Lu, Garron T. Dodd et al. "PTPN2 regulates T cell lineage commitment and $\alpha\beta$ versus $\gamma\delta$ specification", The Journal of Experimental Medicine, 2017 Publication	<1 %
58	Jichuan Zhang, Jingyi Fei, Benjamin J. Leslie, Kyu Young Han, Thomas E. Kuhlman, Taekjip Ha. "Tandem Spinach Array for mRNA Imaging in Living Bacterial Cells", Scientific Reports, 2015 Publication	<1 %
59	Edouard Bertrand, Pascal Chartrand, Matthias Schaefer, Shailesh M. Shenoy, Robert H. Singer, Roy M. Long. "Localization of ASH1 mRNA Particles in Living Yeast", Molecular Cell, 1998 Publication	<1 %
60	bioone.org Internet Source	<1 %
61	theses.gla.ac.uk Internet Source	<1 %
62	www.cell.com Internet Source	<1 %
63	www.prolekare.cz Internet Source	<1 %

Exclude quotes On

Exclude matches

< 14 words

Exclude bibliography On