Deciphering the regulatory role of microRNAs during early developmental stages of B and T lymphocytes

A thesis
Submitted for the degree of
Doctor of Philosophy

By
Sameena Nikhat
(Reg. No. 10LAPH01)

Department of Animal Biology
School of Life Sciences
University of Hyderabad
Hyderabad - 500 046
India

October 2020

University of Hyderabad
(A Central University established in 1974 by an Act of Parliament)
Hyderabad - 500046, India

CERTIFICATE

This is to certify that **Ms. Sameena Nikhat (Reg. no. 10LAPH01)** has carried out the research work embodied in the present thesis under my supervision and guidance for a full period prescribed under the Ph.D. ordinance of this University. We recommend her thesis entitled "**Deciphering the regulatory role of microRNAs during early developmental stages of B and T lymphocytes**" for submission for the award of the degree of *Doctor of Philosophy* in Animal Biology.

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

16/12/2020
अध्यक्ष / HEAD
जंतु जैविकी विभाग
Department of Animal Biology

Dean,
School of Life Sciences

DEAN
School of Life Sciences
University of Hyderabad
Hyderabad - 500 046.

University of Hyderabad
(A Central University established in 1974 by an Act of Parliament)
Hyderabad - 500046, India

DECLARATION

I hereby declare that the work embodied in this thesis entitled "**Deciphering the regulatory role of microRNAs during early developmental stages of B and T lymphocytes**" has been carried out by me in the Department of Animal Biology, School of Life Sciences, University of Hyderabad, for the award of the degree *Doctorate in Philosophy*. This work has not been submitted to any other University for the award of any other degree or diploma.

Sameena

_____

Sameena Nikhat
(Reg. no: 10LAPH01)
PhD research scholar

16/10/2020

Prof. Jagan Pongubala
Ph.D. Supervisor

Prof. Jagan Pongubala
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.

I am thankful to my Doctoral Committee members Prof. S. Dayananda and Prof. Sharmistha Banerjee for their guidance, timely and helpful suggestions.

I am thankful to Prof. Anita Jagota, Head, Department of Animal Biology and the former Heads, Prof. Jagan Pongubala, Prof. B. Senthilkumaran and Prof. Manjula Sritharan for allowing me to use departmental facilities.

I am thankful to Prof. S. Dayananda, Dean, School of Life Sciences and the former Deans, Prof. P. Reddanna, Prof. K.V. Ramaiah and late Prof. Aparna Dutta Gupta for allowing me to use the school facilities.

I would like to thank Dr. Dasaradhi Palakodeti from InStem (The Institute of Stem Cell Science and Regenerative Medicine, Bengaluru) for his guidance, support and valuable advices during the course of my work, and Dr. Ullas Kolthur-Seetharam from TIFR (Tata Institute of Fundamental Research, Mumbai), for his help in procuring the miRNA-Sponges.

I also thank all the faculty members of School of Life Sciences for their valuable suggestions and timely help for allowing me to use their laboratory facilities, when required.

I would like to thank the members of the genomics facility at C-CAMP (Centre for Cellular And Molecular Platforms) for their help with the Next-generation sequencing (NGS) services. I would also like to thank Mrs. Rama Devi, NIAB (National Centre for Animal Biotechnology) for her help with Flow Cytometry.

I would like to thank my lab member, Ms. Anurupa Devi, for her efforts in bioinformatics analysis of my NGS data. I would like to thank all the current members of my lab: Ms. Anubhooti, Ms. Priyanka Kriti Narayan, Ms. Arpita Prusty, Mr. Ashok Kumar and Ms. Nidhi Singh for their help and support in the lab, and for sharing good food and happy times. I would also like to thank the previous members of my lab: Mr. Anil Kumar Moola, Mr. Meraj Hassan and Mr. Sambasiva Rao for their assistance and help during my early days at lab. I am very grateful to Ms. Manisha Mallick for encouraging me with my experiments, helping me with the lab office work and for managing my Pre-PhD seminar arrangements. I would also like to thank our lab alumnus, Dr. Ravi Boya. I would always be grateful to him for the immense help, assistance and cooperation that he had extended as we worked to set up our new lab. I also greatly appreciate his suggestions and help in carrying out my experiments during all these years.

I wish to thank all my friends and well-wishers from School of Life Sciences: Ms. Mitali Chaudhary for always being there for me whenever I needed her, Mr. Chava Suresh, Mr. Naveen Kumar, Mr. Ashok Kumar, Mr. Vinod, Ms. Deepa S, Ms. Amina Rafath and Mr. Tejaswi, for all their help, constant support and encouragement during my stay at School of Life Sciences. I would also like to thank Ms. Ghufrana Abdus Sami for being so kind and caring, for helping and encouraging me whenever I was low in spirits. I am grateful to Ms. Ghufrana, Mr. Fareed and Mr. Abdul Salam for being the first responders after my accident and helping me to reach hospital in-time. I am also grateful to Ms. Manisha Mallick for being so kind to change my bandages regularly during my recovery.

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

I sincerely thank UGC for funding my doctoral studies and DBT, DST, UGC – SAP and DBT – CREBB for providing financial support to lab and for the infrastructural facility of the School.

I am indebted to my school and college teachers for being sincere in carrying out their responsibility of teaching and mentoring; and their timeless efforts in instilling the love for science within me.

I am so gratified to have the empyrean, ineffable support and encouragement from my parents, Mr. M.A.Rasheed Khan and Mrs. Shakeela Naheed, all through my life. I am obliged to them for their unequivocal love, advices, prayers and their earnest efforts in teaching and inculcating virtues and good morals. I am so grateful to my grandfather, Mr. M. A. Khader Khan and my late grandma, Mrs. Hajera Begum, who were no less than parents to me in every possible way, for raising me up with love and diligence. I cherish and treasure the company and affection of my little sister, Dr. Humera Fatima, who has been by confidante and friend forever; and I appreciate her for being instrumental in increasing my awareness and interest in understanding human physiology and medicine.

I am extremely grateful to my husband, Dr. Abdul Razak, for being my companion for life; for providing unwavering support, advice and being ‘my rock’ whenever I fell weak in hope and spirits; and for his sheer patience during the long wait for all these years until I could bring this course to completion. I am privileged and blessed to have Mr. Ghulam Dastagir and Mrs. Beepasha Begum as my in-laws, who have been parental figures for me and have always prayed for my well-being and success.

Above and beyond all, I am ever grateful to Allah Almighty, for giving me the health, strength and courage to endure this colossal journey and for placing all the above people in my life to help and guide me in one way or other; and to help me learn invaluable lessons from my experiences to gear up for the challenges of life that lie ahead.

-Sameena Nikhat

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Ago2	Argonaute2
bp	Basepairs
cDNA	Complementary DNA
^{0}C	Degrees Celsius
DC	Dendritic Cell
d	Days
DTT	Dithiothreitol
dNTPs	Deoxyribonucleotide triphosphates
EBF1	Early B cell Factor 1
Flt3R	Fms-like Tyrosine Kinase Receptor
Flt3L	Fms-like tryrosine Kinase Ligand
FBS	Fetal Bovine Serum
ER	Estrogen Receptor
GMCSF	Granulocyte Macrophage Colony-Stimulating Factor
h	Hour
IL-7	Interleukin-7
IL-3	Interleukin-3
Kb	Kilobase
IP	Immunoprecipitation
MCSF	Macrophage Colony-Stimulating Factor
min	Minutes
miRNA	MicroRNAbp
mRNA	Messenger RNA
mL	Millilitre
mM	Millimolar
µM	Micromolar
µg	Microgram
nM	Nanomolar
NK	Natural Killer Cell
nt	Nucleotide
NEB	New England Biolabs
NP-40	Nonidet P-40
4-OHT	4-O-Hydroxy Tamoxifen
pH	Potential of hydrogen
PBS	Phosphate Buffered Saline
Pre-miRNA	Precursor microRNA
qRT-PCR	Quantitative Reverse transcription-Polymerase Chain Reaction
rpm	Rotations per minute
RIP	RNA Immunoprecipitation
SCF	Stem cell Factor
UTR	Untranslated region

LIST OF FIGURES

LIST OF TABLES

Page

Main Tables

Supplementary Tables

ABSTRACT

Commitment of multipotent hematopoietic progenitors towards a particular lineage involves activation of cell type-specific genes as well as silencing of genes that promote alternate cell-fates. The transcription factors (TFs) that drive multipotent progenitors towards B and T lineages have been extensively studied over the past two decades; however, the role played by microRNAs expressed during B and T cell-fate commitment, remains less explored. Despite the mutually exclusive gene expression programs regulating early B- and T-lymphocyte development, we show that Early-B and Early-T cells exhibit significantly correlated microRNA profiles. Interestingly, profiling of Argonaute2-associated mRNAs revealed that miRNA-targetomes of Early-B and Early-T cells are quite distinct, and predominated by transcripts that are closely associated with Natural Killer cell (NK), Dendritic Cell (DC) and Myeloid lineages. We demonstrate that combinatorial expression of multiple lymphocyte microRNAs - miR-186-5p, miR-128-3p and miR-330-5p in *Ebf1-/-* progenitors, significantly attenuates their myeloid differentiation capacity due to repression of myeloid-associated transcripts. Correspondingly, knockdown of these miRNAs during early B and T cell development resulted in de-repression of their myeloid-associated targets. Furthermore, B and T lineage-specific TFs potentially drive the expression of shared miRNAs in the respective cell-types. Collectively, using genome-wide and molecular analyses, this work demonstrates that microRNAs function as integrated components of gene regulatory networks that dictate lymphoid cell-fate commitment.

CHAPTER-1: INTRODUCTION

1.1. Introduction to Haematopoietic System

Haematopoiesis is defined as the process by which all the cellular components of blood and immune system are developed. It begins during embryonic development and continues throughout the adulthood to produce and maintain the cells of blood and the immune system. In mammals, the first primitive cells of blood develop in the yolk sac (called as 'extra-embryonic haematopoiesis' or 'Primitive haematopoeisis') which largely comprises of nucleated erythrocytes that serve to produce RBCs to assist oxygenation of tissues as the embryo undertakes rapid cell divisions (Orkin and Zon, 2008). However, this wave is temporary, as the erythroid progenitors do not have renewal capacity and are thus are not pluripotent. This is followed by a transient wave called 'Definitive haematopoiesis' which takes place in the blood islands and produces erythroid-myeloid progenitors (EMPs) as well as lymphoid progenitors, that seed the fetal liver in a transient manner (Bertrand et al., 2007; Böiers et al., 2013; McGrath et al., 2015).

However, the 'Definitive haematopoiesis', which eventually produces the permanent adult haematopoietic system, occurs later during the development, at particular time points in different species. It begins with emergence of the first identifiable Haematopoietic Stem cells (HSCs) in the evolutionarily conserved aorto-gonado-mesonephros (AGM) region (Ivanovs et al., 2011). Thereafter, haematopoiesis shifts to the fetal liver, then subsequently to the bone marrow, which is the location for HSCs in adults (Fig. 1.1).

Haematopoietic stem cells (HSCs) are the source of definitive haematopoiesis, i.e., they are responsible for continuous production and sustenance of various types of cells that constitute the blood and the immune system during the life of an organism. They hold an immense self-renewal capacity and are pluripotent i.e., capable of generating all the cell types that constitute the haematopoeitic system. The HSCs are thus programmed to allow efficient production of all the cellular components of blood that perform a variety of functions: from RBCs that serve to carry oxygen, megakaryocytes (which generate platelets) that regulate blood clotting; to the innate as well as the adaptive immune cells that protect against infections and cancers.

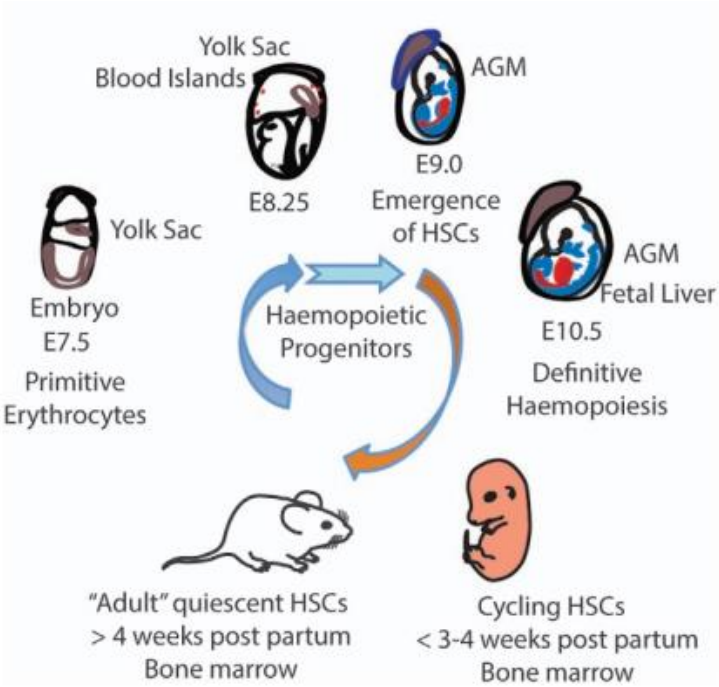


Figure 1.1. The journey from fetal to adult murine haematopoiesis. 'Primitive haematopoiesis' (represented by blue arrows) is followed by a wave of 'Definitive haematopoiesis' (represented by orange arrows), leading to emergence of HSCs, first in the aorto-gonado-mesonephros (AGM) region of embryo, followed by fetal liver and then bone marrow. HSCs drive a continuous supply of cells that eventually constitute the blood and immune system thorough out the adult lifespan (adapted from Ng et. al., 2017).

HSCs differ their self-renewal capacity; and they have been classified into various subsets depending upon the number of symmetric divisions that can be made during its lifetime. In humans and mice, HSCs that show repopulation capacity for greater than 16 weeks in a primary transplantation assay and at least once during the secondary transplantation are called Long-Term (LT-) HSCs. However, those HSCs that produce all differentiated cell types but show transient primary engraftment (also sometimes in secondary engraftment) are classified as Intermediate (IT-) HSCs or Short-Term (ST-) HSCs, depending upon the length and efficiency of the graft produced (Kent et al., 2009).The self-renewal capacity of HSCs appears to be dependent upon the time that is spent in quiescence as studies prove that HSCs which are most dormant show the most robust and longest repopulation capacity. Certain cell cycle factors which were shown to be correlated inversely with repopulation capacity include the frequency of cell division and the time a given HSC takes to exit quiescence. Interestingly, the mRNA and protein levels of CDK6 in quiescent HSCs is directly proportional to

dormancy, and may thus serve as a marker of the quiescence (Laurenti and Göttgens, 2018).

The niche of HSCs is a highly intricate ecosystem that promotes and sustains HSC function by facilitating the survival as well as long-term sustainment of the HSC pool. The mechanisms by which interactions within the niche shape the differentiation journeys and fate of individual HSC subsets, still remains unclear. Studies show that the distinct HSC subsets quite possibly prefer different niche environments, and HSCs are known to lose their self-renewal capacity quite rapidly upon isolation from their *in vivo* niches. Although most of HSCs in adults are located in the bone marrow, a small fraction are released into the blood, lungs and spleen under the regulation of circadian-clocks (Laurenti and Göttgens, 2018). The bone marrow comprises of a mixture of various types of cells, including mesenchymal cells, osteoblasts, osteoclasts, reticular cells, fat cells, blood cells, endothelial cells and other lesser-known cell types, thereby providing an environment suitable for sustenance and function of HSCs. These studies show that extrinsic signals from extracellular matrix-associated, membrane-bound soluble or ligands available within the niche play a crucial role for appropriate HSC behavior.

1.2. Lineage-restriction downstream of the HSPC continuum

In the classical model of haematopoiesis, long-term reconstituting HSCs (LT-HSCs) constitute the apex of the hierarchy and exhibit self-renewal as well as multi-lineage differentiation potential. LT-HSCs differentiate into short-term reconstituting HSCs (ST-HSCs) that exhibit a limited self-renewal potential but maintain multi-potency. ST-HSCs generate multipotent progenitors (MPPs), of which distinct subsets have been described. Multipotent progenitor cells express CD117 (c-Kit, which binds to a ligand called SCF or Stem Cell Factor) as well as Stem cell antigen (Sca)-1, but lack the expression of all mature cell markers (c-Kit$^+$Sca-1$^+$Lin$^-$). The MPPs can either directly differentiate into megakaryocyte-erythrocyte progenitors (MEPs) or into common-myeloid progenitor (CMPs). MEPs differentiate into megakaryocytes/platelets and erythrocytes while CMPs eventually give rise to granulocytes, macrophages, and dendritic cells via granulocyte–macrophage progenitors (GMPs) (Fig. 1.2).

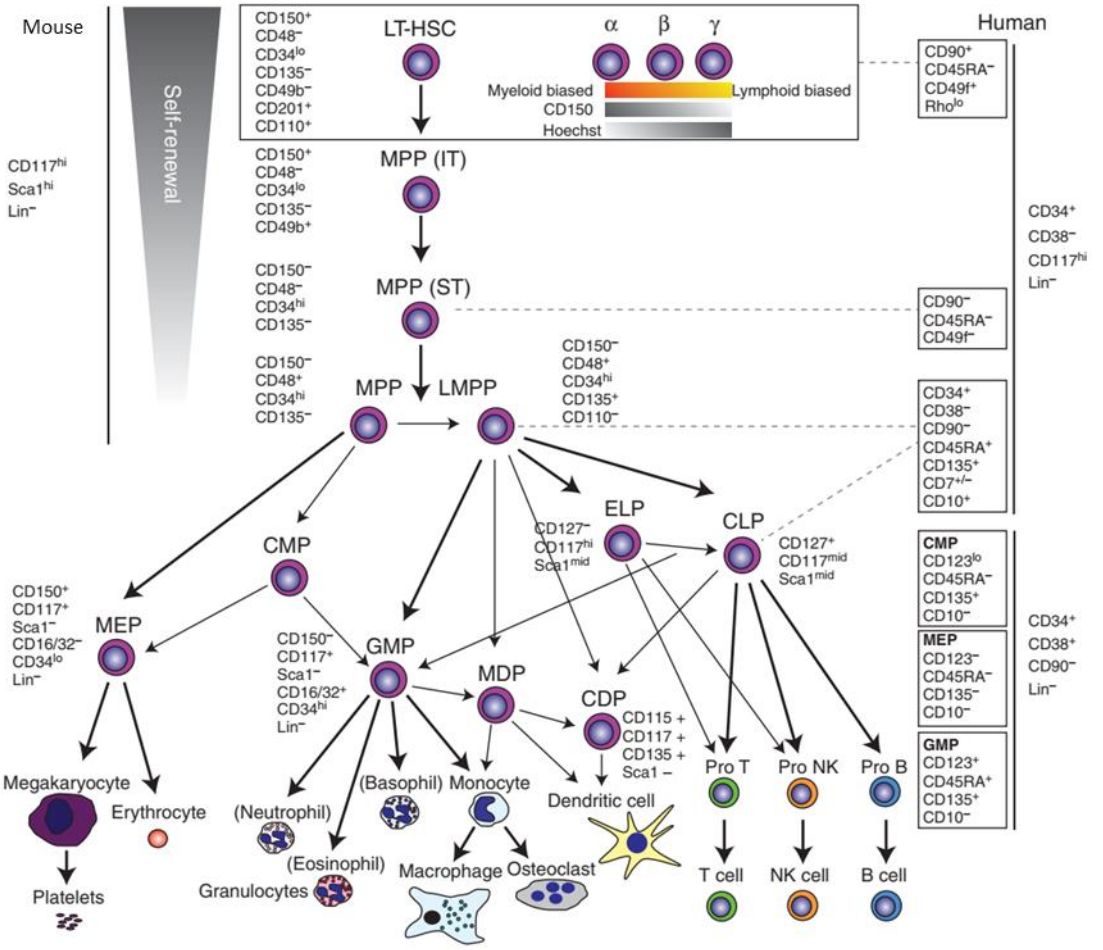


Figure 1.2. The classical haematopoietic hierarchy in the adult bone marrow. Murine adult haematopoiesis is represented here along with corresponding human haematopoietic population with their markers are indicated. In the classical model, long-term reconstituting HSCs (LT-HSCs)sit at the top of hierarchy and maintain self-renewal and multi-lineage differentiation potential. HSCs differentiate into all blood cell lineages via long described (bold arrows) and potentially also or alternatively more recently described differentiation routes (thin arrows). HSC, haematopoietic stem cell; MPP, multipotent progenitor; LT-, long-term repopulating; IT-, intermediate-term repopulating; ST-, short-term repopulating; LMPP, lymphoid-primed MPP; ELP, early lymphoid progenitor; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte–macrophage progenitor; MEP, megakaryocyte–erythrocyte progenitor; CDP, common dendritic progenitor; MDP, monocyte–dendritic cell progenitor; NK, natural killer cell. CD135, also known as FLK2 and FLT3; CD117 also known as IL-7R; Lin, lineage markers (which are a combination of markers found on mature blood cells but not HSCs or progenitors) (Adapted from Rieger and Schroeder, 2012).

On the other hand, MPPs can acquire CD135/ Flt3R to develop into lymphoid primed MPPs (LMPPs) that show lympho-myeloid potential. LMPPs are a subset of c-Kit$^+$Sca-1$^+$Lin$^-$CD34$^+$ cells that no more have self-renewal capacity or

megakaryocyte/erythrocyte differentiation potential, but possess the ability to generate all granulocyte/macrophage progenitors (GMPs) as well as lymphocytes. However, as LMPPs eventually acquire CD117/IL-7R, they become common lymphoid progenitors (CLPs) that can give rise into lymphocytes (B and T), Dendritic cells and Natural Killer cells, but lose myeloid developmental potential (Fig. 1.2).

The classical model is being revised further as new hierarchical progenitors are being discovered. A recent study that employs scRNA-seq (single cell RNA-Seq) in a human model suggests acquisition of cell fates is a continuous process, such that the uni-lineage restricted cells emerge directly from a continuum of undifferentiated, low-primed haematopoietic progenitors, without involving any major transitions via the multi- and bi-potent stages (Velten et al., 2017). A similar study in humans involving multi-lymphoid progenitors (MLPs), LMPPs or GMPs along with lympho-myeloid progenitor cells from cord blood, further suggests a model wherein myeloid and lymphoid lineage cells are generated from a continuum of progenitors, rather than only uni-lineage progenitors (Karamitros et al., 2018). Thus, the above mentioned studies change the perception of haematopoietic differentiation to be a continuous process rather than a discrete hierarchy; suggesting no apparent boundary between stem cells and progenitors.

1.3. Molecular control of Haematopoietic cell fate choice

Haematopoietic lineage differentiation is a complex process involving global changes in gene expression and must be appropriately regulated to allow generation of a given category and number of mature cells. Lineage specification and commitment involve induction as well as maintenance of lineage-affiliated genetic programs. This encompasses two aspects: the expression of lineage-specific genes as well as the repression of genes that are associated with other lineages, so as to establish the lineage identity. This is achieved by complex cellular networks that integrate lineage-specific transcription factors (TFs), cytokine signals along with several other epigenetic mechanisms (Lara-Astiaso et al., 2014; Javierre et al., 2016; Brown et al., 2018; Hu et al., 2018; Magilnick and

Boldin, 2018). Of these, the role played by cytokines and transcription factors in shaping the haematopoietic landscape, is described below.

1.3.1. Haematopoietic Cytokines

The cytokines, which are soluble protein factors secreted by progenitor or mature cells, provide an extrinsic mode of instruction for lineage choice of multipotent progenitors. Cytokines serve as significant regulators of haematopoiesis as they impact numerous aspects of haematopoietic cells, including proliferation, maturation, survival and function (Rieger et al., 2009). For example, TGF-β1 (at low levels) maintains the differentiation potential of HSCs or progenitors by negative regulation of cell cycling via activation of the cyclin-dependent kinase inhibitor, p57KIP2 (Batard et al., 2000). On the other hand, SCF (Stem Cell Factor), Flt-3L (Fms-like Tyrosine kinase-Ligand) and Tpo (Thyroperoxidase) along with IL-11 or IL-3, promote expansion of long-term repopulation of LSK cells, *in vitro*. However, cytokines do not merely allow survival and proliferation of the progenitors but act as lineage instructors. For example, granulocyte–macrophage progenitors (GMPs) when cultured in presence of G-CSF or M-CSF will generate either granulocytes or monocytes, respectively. Similarly, while IL-7 signaling is known to be crucial for early development of B-and T-lymphocytes, myeloid-biased progenitors were shown to exhibit poor response to IL-7 (Muller-Sieburg et al., 2004). The cell type-specific, pleiotropic and redundant regulation exhibited by cytokines perhaps explains our inability in precisely outlining their impact on cell fate choice. Moreover, even if they instruct specific lineages, other additional cell-intrinsic mechanisms that dictate lineage choice in absence of the cell-extrinsic signals from cytokines, have also been shown to important in regulating the differentiation process.

1.3.2. Transcription Factors in haematopoietic cell fate decisions

From CMPs to MEPs and GMPs

Unlike the cytokines, a transcription factor can independently instruct the lineage choice and are sometimes able to reprogram the committed cells. These lineage-specific factors act either independently, or in collaboration with each other or with lineage-specific cytokines to establish gene regulatory networks that bring about differentiation of multipotent progenitors to a given cell fate. Co-expression of PU.1

and GATA-1 initiates the commitment of HSCs/ MPPs towards CMPs (Common Myeloid progenitors). While these two factors are co-expressed in CMPs, their mutual exclusion is crucial for further differentiation of CMPs into either monocytic/granulocytic or megakaryocytic/erythroid progenitors (Fig. 1.3.1). Expression of GATA-1 is required absolutely for differentiation of CMPs towards MEPs. However, the GATA-1 interacts with two different factors i.e. FOG and C/EBPβ to produce either megakaryocytes or Eosinophil, respectively (Tsang et al., 1997; Querfurth et al., 2000).

On the other hand, PU.1 is required absolutely for differentiation of CMPs into GMPs, and its sustained expression in GMPs promotes the development of monocytes and granulocytes. High level of PU.1 along with its putative transcriptional partner, the Interferon Consensus Sequence Binding Protein (ICSBP or also called as IRF-8) favors differentiation towards monocytic development while inhibiting the development of neutrophils. However, low PU.1 levels along with an additional TF, C/EBPα, is required for differentiation of GMPs into neutrophils (Fig. 1.3.1).

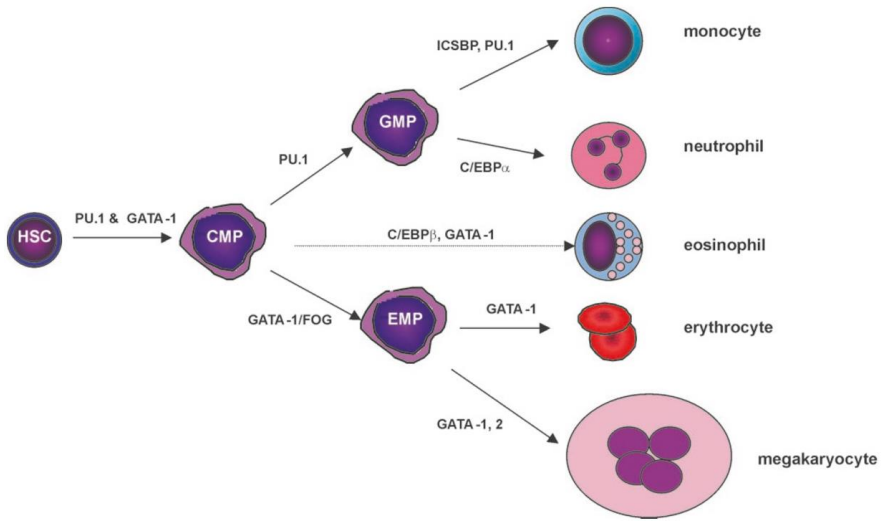


Figure 1.3.1. Transcriptional regulation of common myeloid progenitor (CMPs). CMPs differentiate into either common precursors for granulocytic and monocytic lineages (called as GMPs) or common precursors for both erythroid and megakaryocytic lineages (called as EMPs or MEPs). A separate pathway leading to Eosinophils is also indicated. Dual expression of PU.1 and GATA-1 leads HSCs to develop into CMPs, but then dominant expression of PU.1 is restricted to GMPs, while unopposed GATA-1 expression directs differentiation to EMPs (also called as MEPs). HSC, haematopoietic stem cell; CMP, common myeloid progenitor; GMP, granulocyte–macrophage progenitor; EMP, erythrocyte-megakaryocyte progenitor. (Adapted from Zhu and Emerson, 2002).

From LMPPs to CLPs

Lymphoid-primed multipotent progenitors (LMPPs) are the first lineage restriction point en-route from HSC to committed B- and T-lymphocytes. As discussed above, as described earlier, PU.1 is known to be crucial for differentiation of MPPs towards myeloid progenitors. However, it is also required for differentiation of MPPs to LMPPs, along with two important transcription factors called Ikaros (*Ikzf1*) and E2A, which belong to kruppel-like zinc-finger protein and Helix-loop-helix protein families, respectively. Due to its important role in both lymphoid and myeloid development, PU.1-deficient mice die within 24h of birth and lack mature macrophages, neutrophils, B, and T cells, while Red Blood Cells and Platelets are present (McKercher et. al., 1996). PU.1 plays a concentration-dependent role in fate decision between myeloid and lymphoid lineage. High levels of PU.1 promote macrophage development, whereas low expression levels support B-cell development (DeKoter and Singh, 2000). PU.1 mediates differentiation towards lymphoid lineage via upregulation of *Il7r* at CLP stage (reviewed in Boller and Grosschedl, 2014). However, as the LMPPs differentiate into lymphoid-biased, IL-7R^{+} cells, called CLPs, the levels of PU.1 decrease while E2A and Ikaros proteins continue to be expressed in CLPs until the early developmental stages of B and T lymphocytes (Miyazaki et al. 2014). Therefore, these three transcription factors set the stage for initiation of lymphocyte development from LMPPs (Fig. 1.3.2 and 1.4.1).

It is to be noted that the regulatory networks for a given lineage involve not only the sequential activation of various genes in a linear hierarchy, but also complex feed-forward and feedback loops as well as their cross-antagonism (Fig. 1.3.2). Induction of a lineage-instructive factors in the progenitor cells via extrinsic or intrinsic stimuli initiates the lineage-specific developmental program, resulting in activation of other downstream factors that drive them towards a given lineage along with concomitant repression of genes that are in-concordant with the differentiating cell-fate choice, thereby abolishing the mixed-lineage gene expression pattern.

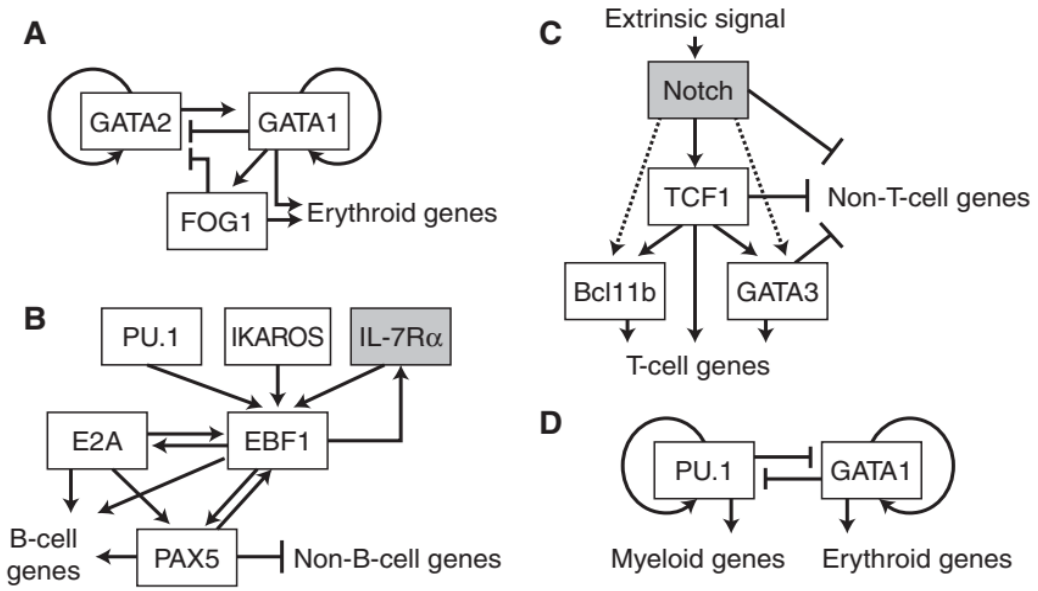


Figure 1.3.2. Network motifs for induction and maintenance of lineage commitment. Simplified examples of molecular mechanisms and networks of stable commitment induction and propagation in erythroid (A), B-lymphoid (B), T-lymphoid (C), and myeloid (D) cells. Direct or indirect activation (arrows) and repression (barred lines) of individual factor expression are indicated. Transcription factors are depicted in white, surface receptors in gray. Dashed lines represent indirect interactions. (Adapted from Rieger and Schroeder, 2012).

1.4. A Focus on early lymphoid development – Commitment of LMPPs towards B and T lineages

1.4.1. Transcriptional control of early B and T cell development

Downstream of the LMPPs, the CLPs (Common Lymphoid Progenitors) are widely considered to be the branch-point for development of B- and T-cells (Lai and Kondo, 2008; Miyazaki et al., 2014). The expression of Ly6D surface marker further divides the CLPs into All-lymphoid progenitors (ALPs) or B-cell biased lymphoid progenitors (BLPs). CLPs lacking Ly6D marker are called ALPs as they retain the potential for generating T cells, B cells, natural killer cells (NK) as well as dendritic cells (DC), (Fig. 1.4.1). However, the BLPs predominantly generate B cells and exhibit markedly low T and NK cell potential (Boller and Grosschedl, 2014). B-cell lymphopoiesis requires activation of a set of TFs including E2A, PU.1, IKAROS and FOXO1 that 'prime' the genomic cis-regulatory regions that are subsequently activated by the B lineage-specific TFs EBF1 and PAX-5. Since

IKAROS (encoded by *Ikzf1*) is involved in regulation of lymphopoiesis, a knockout of *Ikzf1* causes a developmental arrest at LMPP stage, as the cells fail to upregulate Flt3R and differentiate into CLPs. The E-protein TF, E2A, causes lymphocyte-specific priming by activating the lymphocyte-specific genes including *Rag1*, *Rag2*, *Dntt*, along with cytokine receptor, *Il7r*. Moreover, E2A is shown to collaborate along with FOXO1 to activate the key B-lineage determinant, EBF1. Besides E2A and FOXO1, the TF Runx1 was also found to be responsible for activation of EBF1 expression.

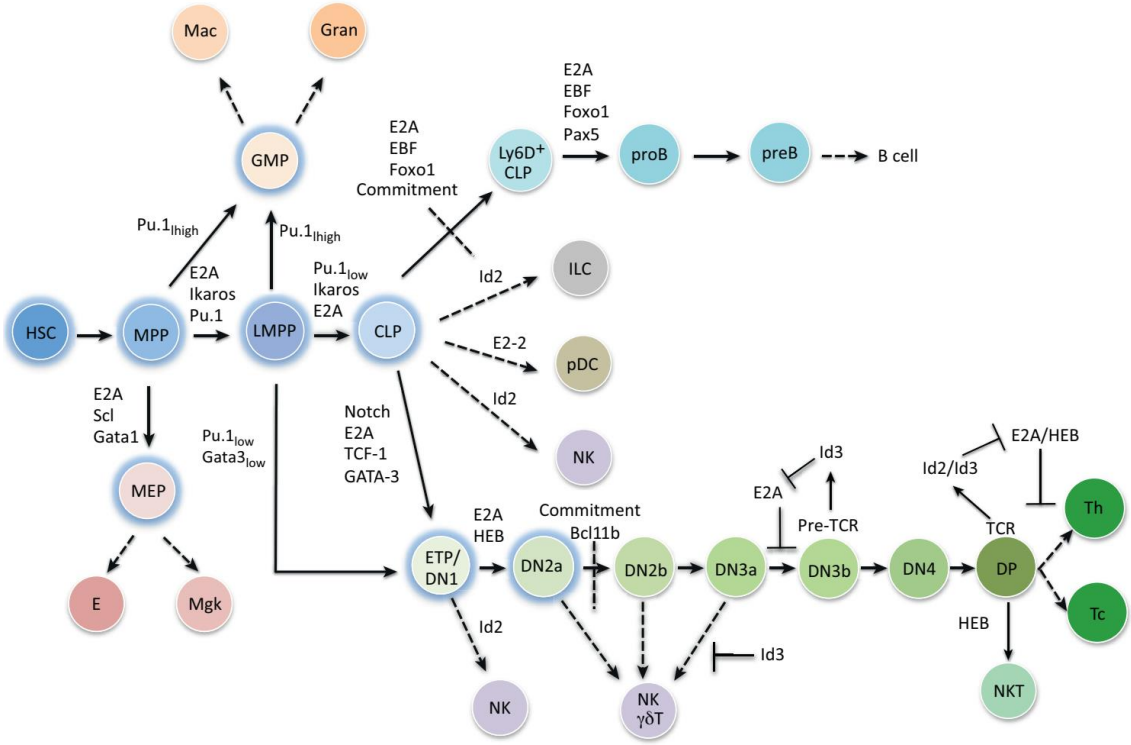


Figure 1.4.1. Transcriptional control of early B- and T-lymphocyte development. LMPPs, the earliest progenitors of B and T lymphocytes, possess both lymphoid and myeloid potential. The CLPs, are divided into ALP or BLP based on expression of Ly6D marker. ALPs give rise to B, T, NK and DC while BLPs are biased for differentiating into B cells. E2A, FOXO1, EBF1 and PAX5. ALP, All-Lymphoid Progenitor; BLP, B-cell biased lymphoid progenitor (Adapted from Miyazaki et al. 2014).

Factors regulating Early-B cell development

The transcription factor, EBF1, has been shown to be the primary cell fate determinant of B-lineage as *Ebf1*-knockout mice exhibit a complete arrest of B-cell development at a CLP-like stage characterized as c-Kit^{+}Sca-

1⁺Flt3⁺IL7R⁺CD43⁺B220⁺CD19⁻ (Lin and Grosschedl, 1995). These mutant cells fail to express key B-lineage genes, including *Cd79a, Cd79b, Vpreb3, Igll1* (that together constitute the Pre-B cell receptor, Pre-BCR) along with other important genes necessary for B cell commitment (i.e., *Pax5*) and are unable to undergo Ig heavy chain gene rearrangements (Zandi et al., 2008). The *Ebf1-/-* progenitors, despite being maintained under B-lymphoid conditions, exhibited developmental plasticity towards alternative lineages including myeloid, dendritic, NK and T cells, both *in vitro* and *in vivo* (Pongubala et al., 2008; Nechanitzky et al., 2013) (Fig. 1.4.2).

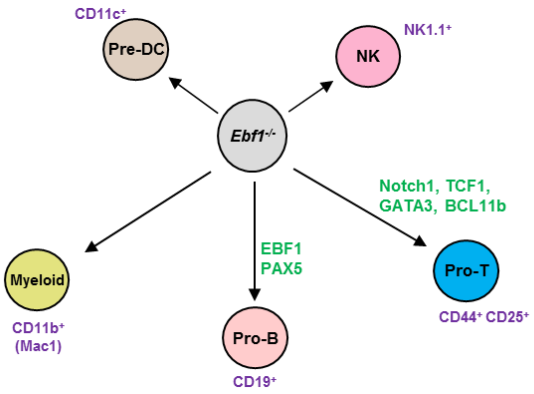


Figure 1.4.2. Multi-lineage developmental potential of *Ebf1-/-* progenitors. Loss of EBF1 arrests the early progenitors at a CLP-like stage. The arrested progenitors exhibit potential to differentiate into multiple lineages like T, Natural Killer (NK), Dendritic cell (DC) and Myeloid cells. The transcription factors that are responsible for differentiation of *Ebf1-/-* progenitors towards B and T lineages are indicated.

In contrast, enforced expression of EBF1 drives B-lineage differentiation at the expense of other haematopoietic lineages in murine HSCs and rescues the developmental arrest of PU.1-deficient MPPs (Medina et al., 2004). EBF1 antagonizes alternative cell fate choice of multipotent progenitors by direct repression of alternate lineage genes including *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). Moreover, the coordinated activity of E2A, EBF1 and FOXO1 activates the TF PAX5 in specified early-B cells, thereby generating a positive feed-back loop of EBF1 and PAX5 resulting in commitment of progenitors towards B-lineage at the CD19⁺ Pro-B cell stage. Therefore, EBF1 plays a non-redundant, pioneer role in specification and establishment of B-lineage identity.

Factors regulating Early-T cell development

On the other hand, upon entering thymus, the LMPPs predominantly differentiate towards T-lineage under the influence of Notch-DLL4 signaling, which drives the differentiation of progenitors towards T-lineage by activation of early T-lineage genes. Development of T-lymphocytes proceeds through the double-negative (DN) stages distinguished by the expression 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}) (Rothenberg et. al., 2008).

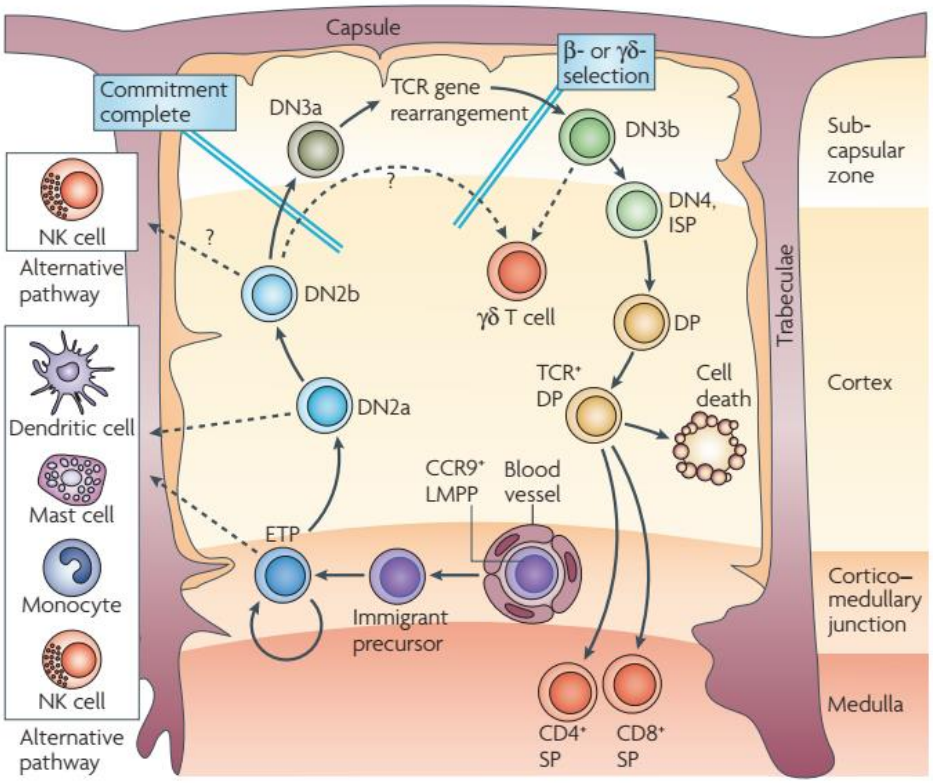


Figure 1.4.3. Early stages during development of T-lymphocytes. Cross-section through an adult thymic lobule showing the migration path of T-cell precursors during development. The early T-cell precursors (ETP) (also referred to as DN1 T cells) subsequently migrate, and differentiate from double negative (DN) to double positive (DP) and then to single positive (SP) stages, through the distinct microenvironments of the thymus. The markers expressed by ETPs, DN2, DN3 and DN4 T-cell subsets are defined as described in text. β-selection occurs during the accumulation of the DN3 T cells in the extreme outer portion of the thymus. Dotted arrows depict alternative developmental pathways that are still possible for ETPs and for different subsets of DN2 cells that are likely to correspond to DN2a and DN2b cells before they complete their commitment to the

T-cell-lineage. CCR9, CC-chemokine receptor 9; ISP, immature single positive; LMPP, lymphoid primed multipotent progenitor; NK, natural killer; TCR, T-cell receptor (Adaped from Rothenberg et al., 2008)

The expression of c-Kit is essential until the DN2 stage and decreases gradually through the DN3 stage, while IL-7Rα expression is upregulated at DN2 and DN3 stages until the β-selection is accomplished (Fig 1.4.3). The T-cell identity genes like *Cd3e*, *Cd3g*, *Il7ra* and *Zap70* that are expressed at low levels in ETPs, begin to be upregulated in the DN2 stage. Subsequently, during DN3 stage of development, the germline transcription of gene segments encoding the variable region of TCRβ is upregulated along with its V-(D)J rearrangement, thus establishing the T-cell-lineage identity. The β chain subsequently pairs up with the pre-Tα (surrogate α chain), and together with CD3 molecules produces a complex called pre-TCR, which leads to the survival, proliferation, arrest in further β chain loci rearrangement as well as allows further differentiation to the double positive (DP) stage by up-regulation of CD4 and CD8 receptors.

The thymic epithelium supports T-cell differentiation, proliferation and survival by providing a potent combination of cell surface ligands and growth factors. The most important environmental signal provided by thymic stromal cells are the delta-like ligand 1 (DLL1) and DLL4 ligands for the Notch cell-surface receptors, which trigger the development of ETPs towards T-lineage. Signaling via these ligands initiates and sustains the T cell developmental program throughout the pro-T-cell stages by activation of early-T genes like *Ptcra*, *Cd3e* and also important TF like TCF-1 (encoded by *Tcf7*), which in turn, is crucial for sustaining T-cell development by activating critical factors like GATA3 and BCL11B (Sambandam et al., 2005; Pai et al., 2003; Weber et al., 2011; Ikawa et al., 2010). Expression of the active Notch1 in the bone marrow induces thymus-independent differentiation of early lymphoid progenitors into T cells, at the expense of B lymphopoiesis (Pui et al., 1999). Conditional deletion of *Notch1* in HSCs results in arrest of T cell development at the earliest precursor stage (DN1), while simultaneously promoting B cell development in the thymus (Radtke et al., 1999; Wilson et al., 2001). While Notch1/ RBPJ and GATA3 are responsible for blocking of B-lineage potential in earliest thymic progenitors; BCL11B has been shown to play a role in

T-lineage commitment by repressing the development of Natural Killer (NK) cells (Li et al., 2010; Scripture-Adams et al., 2014).

1.4.2. Epigenetic control of early B and T cell development

Differentiation of MPPs into committed B or T cells is contingent upon various transcriptional as well as epigenetic changes. It is reported that multi-lineage transcriptional priming within the MPPs allows establishment of a permissive chromatin state in order to facilitate subsequent binding of lineage-specific transcription factors (Hu et al., 1997). Lymphoid-specific activation of chromatin regions including enhancers, has been found to be associated with expression of IKAROS, PU.1 and E2A. IKAROS was shown to function as an activator as well as repressor of transcription by recruiting chromatin-remodeling complexes, including SWI/SNF and Mi-2/NuRD (Kim et al., 1999). It also mediates chromatin accessibility at the IgH and Igκ locus and regulates the expression of *Rag* genes, thereby contributing to immunoglobulin gene rearrangements (Reynaud et al., 2008). Moreover, lymphoid cell priming also involves additional transcription factors like RUNX1, TAL1 and FLI1 that set the stage 'permissive' for the B-lineage program (reviewed in Boller and Grosschedl, 2014).

Once the permissive haematopoietic chromatin is established by the above factors, TF EBF1 initiates epigenetic changes at the loci of B-lineage genes for their subsequent activation. In particular, EBF1 initiates chromatin remodeling at the *Cd79a* promoter along with E2A and RUNX1 via DNA de-methylation, thereby increasing accessibility of *Cd79a* promoter for further regulation by other transcription factors like PAX5 (Hagman et al., 2005). EBF1 also induces D-J_H rearrangement of the IgH locus as well as Igλ locus. More recently, the C-terminal domain of EBF1 was found to show 'pioneering activity' by binding to naïve chromatin of B-cell genes and subsequently induce chromatin accessibility and active histone marks (H3K4me1 and H3K4me2) at the EBF1-occupied genomic regions (Boller et al., 2016).

On the other hand, in the early pre-commitment stages of T-cell development, PU.1 was shown to play a 'pioneering role' in the DN1 and DN2 stages by binding to thousands of open as well as closed genomic loci, leading to the opening of

chromatin at these sites (Ungerbäck et al., 2018). However, as the cells undergo transition from DN2 to DN3 stage which marks the T-cell fate commitment, PU.1 occupancy is lost at several of these loci, concomitant with the sharp decrease in PU.1 expression at DN3 stage. In contrast, the expression levels of BCL11B, a critical regulator of T-cell commitment, only begin to express at the DN2 stage. Comparison of chromatin interaction maps at BCL11B target genes before and after T cell commitment has taken place, indicates a remarkable increase in chromatin interactions during DN2-to-DN3 transition in TADs (Topologically Associated Domains) that exhibited increased BCL11B binding (Hu et al., 2018). Therefore, being a subunit of the ATP-dependent chromatin remodeling SWI/SNF complex, BCL11B may contribute to the maintenance of the three-dimensional nucleome of T cells resulting in local chromatin remodeling.

1.4.3. Non-coding RNAs in early B and T cell development

Although several studies have described the functions of lineage-specific transcription factors, cytokine signals and epigenetic modifications in regulating the development of B and T lymphocytes, the functional role of non-coding part of genome (i.e., Long non-coding RNAs and microRNAs) is still being investigated. While the long non-coding RNAs are now increasingly being shown to regulate the 3D architecture of genome during cell differentiation, studies carried out earlier suggest various roles for microRNAs in haematopoietic stem cell maintenance and differentiation (Van Schoonhoven et al., 2019; Chen et al., 2004; Monticelli et al., 2005; Petriv et al., 2010). Despite this, the precise set of miRNAs expressed during lymphocyte differentiation and their ability to act as an integral part of mechanisms that shape the cell fate choice of multipotent progenitors, has not yet been addressed comprehensively. In this perspective, we decided to pursue the role played by microRNAs during B and T cell-fate commitment.

1.5. MicroRNAs and gene regulation

MicroRNAs (miRNAs) are short ~22nt RNAs involved in the post-transcriptional regulation of genes across wide range of biological processes. By binding to sequences present primarily on target RNA's 3'UTR, they trigger translational

repression along with target destabilization and degradation, eventually resulting in decreased protein output (Bartel 2009).

1.5.1. Biogenesis of miRNAs

There are two pathways that describe miRNA biogenesis: the Canonical pathway and Non-canonical pathway.

The canonical Pathway

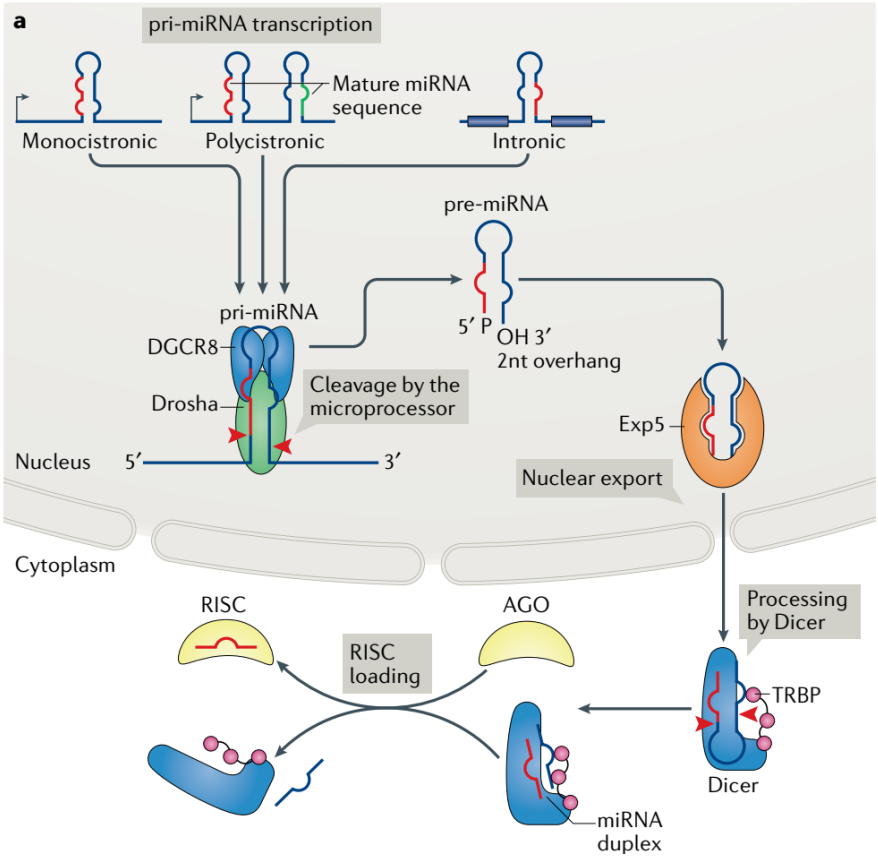


Figure 1.5.1. The canonical pathway of miRNA biogenesis. Animal microRNAs (miRNAs) are encoded as individual genes (monocistronic), as gene clusters (polycistronic) or in introns of host genes (intronic). RNA polymerase II (not shown) generates primary miRNA (pri-miRNA) transcripts containing hairpins and 5′ and 3′ flanking sequences. The microprocessor complex consisting of a DiGeorge critical region 8 (DGCR8) dimer and the enzyme Drosha, cleaves the pri-miRNA at the stem of the hairpin (red arrowheads) and liberates a precursor miRNA (pre-miRNA) with a characteristic 3′ hydroxyl group (OH), overhangs of 2 nucleotides (nt) and a 5′phosphate (P). Exportin-5 (Exp5) binds pre-miRNAs and facilitates their export into the cytoplasm, where Dicer cleaves them within the stem close to the terminal loop (red arrowheads) and generates a miRNA duplex intermediate. The Dicer, trans-activation-responsive RNA-binding protein (TRBP) and an Argonaute (AGO) protein assemble into the RNA-induced silencing

complex (RISC) loading complex, and one miRNA strand is transferred to the AGO protein, resulting in the formation of RISC (Adapted from Treiber et al., 2019).

In this pathway of miRNA biogenesis, RNA polymerase II transcribes primary transcripts that are capped and polyadenylated, called the 'Primary microRNAs'. These are generated from a single or multiple miRNA genes or from the introns of protein-coding genes. These primary transcripts they can be few hundred bases to several kilobases long. The functional portion of this primary transcript is the 20–25 region called 'mature miRNA' which is embedded in the stem region of its hairpin and needs to be excised out from the hairpin. This occurs in two major steps: In the first step that occurs inside nucleus, the 'primary transcript' is cleaved into 60-70nt single-hairpin structure called 'precursor miRNA', by an enzyme complex called microprocessor consisting of two proteins: Drosha AND DGCR8 (Fig. 1.5.1). In the second step, the precursor hairpins are transported by Exportin-5, from nucleus to cytoplasm where the proteins DICER and TRBP cleave out the loop region of hairpin to generate an miRNA duplex, consisting of a passenger strand (shown in blue) and the guide strand or 'mature miRNA' (in red) (Fig. 1.5.1). This miRNA duplex is then passed onto a member of Argonaute family of proteins (called AGO) whose slicing activity releases out the passenger strand while retaining the guide strand or mature miRNA (shown in red). This mature miRNA, along with the Ago protein and Ago-interacting partner called GW182, together constitute the core components of what is called as RISC, the RNA-induced silencing complex which binds to target mRNAs and regulates their fate.

The non-canonical pathway

Besides the classical canonical pathway, there exists the non-canonical pathways of biogenesis which involves generation of mature miRNAs from various types of RNA intermediates, e.g., spliced out introns that act as precursor miRNAs, from misfolded tRNAs when they lose their lariat structures, from snoRNA fragments or finally from random transcription events from RNA polymerase II (Fig. 1.5.2). All these types of intermediates are independent of cleavage by Drosha/ DGCR8 and are spontaneously exported to the cytoplasm for processing by Dicer. Since these are generated independent of Microprocessor complex and thus called 'Drosha-independent' miRNAs. However, mature miRNAs can arise in a 'Dicer-

independent' manner and such miRNAs are rather rare. An interesting example is that of miR-451, an erythrocyte specific miRNA that has a very short precursor and thus cannot be processed by Dicer; but is cleaved by Ago2 in its stem region followed by the activity of poly (A)-specific ribonuclease (PARN) that trims off the 3´end of precursor to generate the mature and functional miR-451 (Fig. 1.5.2). Therefore, the slicer activity of Ago2 required for erythrocyte maturation provides an example for existence of Dicer-independent miRNAs.

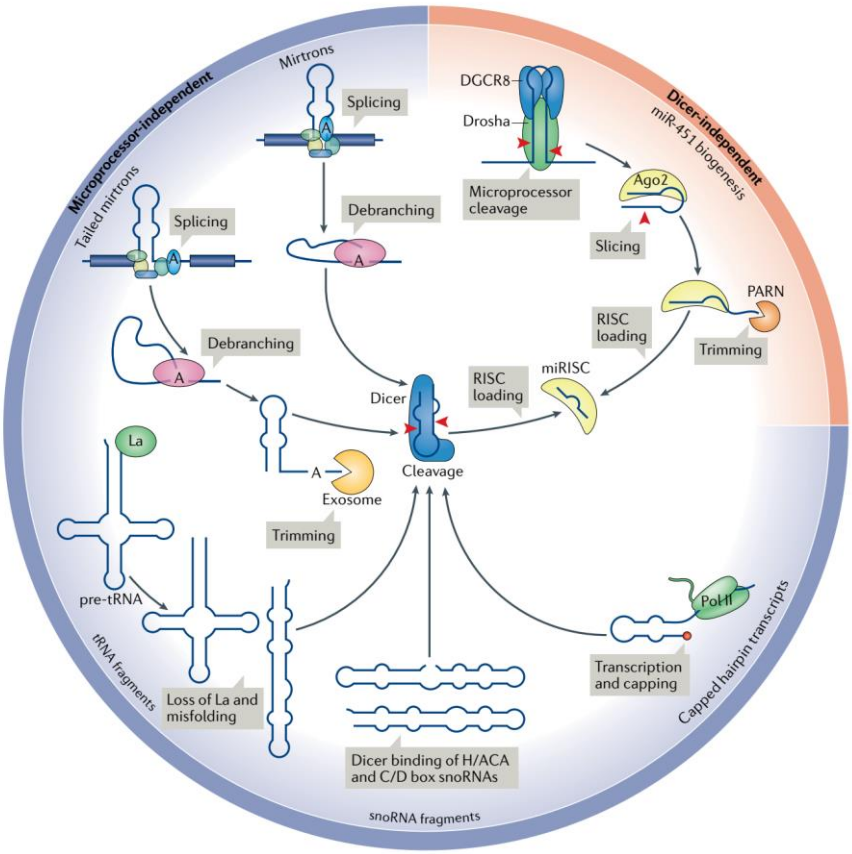


Figure 1.5.2. The non-canonical pathway of miRNA biogenesis. Microprocessor independent microRNAs include mirtrons and tailed mirtrons, which are generated by splicing and subsequent lariat debranching. Tailed mirtrons are further trimmed by the nuclear exosome (not shown). Depending on their sequence, some tRNAs can adopt alternative, hairpin-like structuresthat mimic precursor miRNAs (pre-miRNAs). In normal conditions, the RNA chaperone La prevents the formation of these structures and supports tRNA folding. Some small nucleolar RNAs (snoRNAs) of both major snoRNA families, the H/ACA and C/D box snoRNAs, can also serve as Dicer substrates. Finally, miRNAs can also be processed from hairpins directly generated by RNA polymerase II (Pol II) at specific transcription start sites. These pre-miRNAs are capped and exported through the exportin-1 pathway. The only known Dicer-independent miRNA biogenesis pathway involves the unusually short hairpin of miR-451, which is directly cleaved by Argonaute 2 (Ago2) and trimmed by poly (A)-specific ribonuclease (PARN). (Adapted from Treiber et al., 2019).

1.6. Mechanism of action of microRNAs

1.6.1. The miRNA seed sequence

Within the RISC complex, the 5′ end of the miRNA is held in a helical conformation by MID and PIWI domain of Ago2 protein, while the PAZ domain holds the 3′-terminal nucleotide (Fig. 1.6.1A).

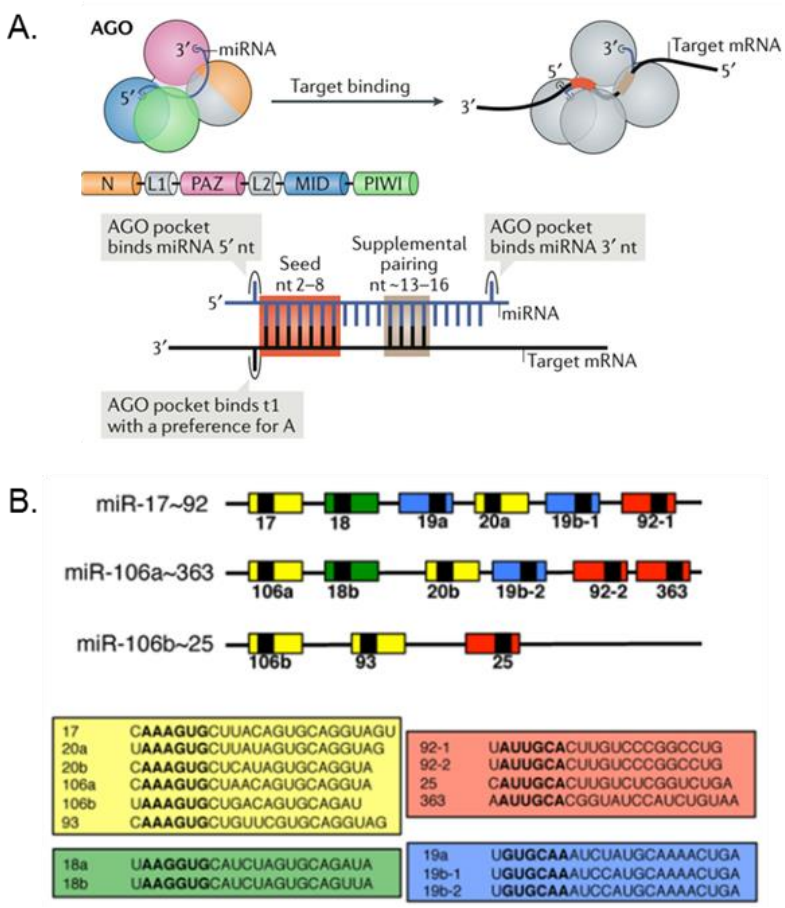


Figure 1.6.1. The importance of miRNA 'seed' sequence.
(A) Interactions of miRNA with the Ago2 domains within the RISC complex. miRNA regions identified as 'seed' and 'supplemental sites' are indicated.
(B) Representation of miRNAs that have the same seed sequence (indicated within coloured boxes) and are thus classified under the same 'family'.

But, how does miRNA base-pair with its target? The nucleotides 2-8 from the 5′ end of the miRNA constitute its 'seed', which is crucial for target mRNA recognition. Interactions at the seed regions involve nucleotides 2–8, 2–7 or 2–6. miRNAs that have the same 'seed' sequence are said to belong to same family. For example: miR-17, miR-20b, miR-20a, miR-93, miR-106a, miR-106b, all arise from different

genomic loci. However, they all are said to belong to the same family because they have the same nucleotides at 2-8th position (Fig. 1.6.1B). However, downstream of the seed, the nucleotides 13-16 can bind to target and enhance or supplement the interaction at the seed. These 'supplemental sites' can differ between the miRNA family members and can direct them to different target sites, thereby increasing the target repertoire (Fig. 1.6.1A).

1.6.2. Mechanism of miRNA-mediated gene silencing

It is known that miRNAs interact with the target mRNAs predominantly at their 3'UTRs. There are certain features of 3'UTRs that make them less or more susceptible to miRNA-mediated regulation. These include:

- Presence of sequences that are complementary to miRNA seed
- The Secondary structure – which may mask or expose the binding sites
- Presence of miRNA sites close to stop codon
- AU-rich regions (lower stability of secondary structures in such regions)
- Several miRNA sites within 50bp of each other – allows cooperative effect
- Interactions with RNA binding proteins (RBPs) – that remodel the target RNA to expose the miRNA binding sites.

So how do miRNAs search for their targets? As mentioned earlier; within the RISC complex, miRNA is held in a particular orientation by specific domains of Ago2, such that, that helix-7 of Ago2 masks the nt 6-8 of miRNA, thus allowing only nt 2-4 to be exposed outside (Fig. 1.6.2A). Subsequently, as the miRNA-RISC (miRISC) complex comes in contact with a potential target mRNA, it is these 2-4nt of the miRNA it that make the first brief contact with the target, which lasts for a few milliseconds. Upon favorable pairing of nucleotide 2-4, the helix-7 of Ago2 is docked away from the miRNA , thereby exposing the remaining portion of the seed to form a more stable interaction, which lasts as long as 300 seconds (Fig. 1.6.2B). However, if the initial pairing is not favourable, Ago2 releases itself off the site and either laterally diffuses along the mRNA to find another possible site or dissociates from it completely.

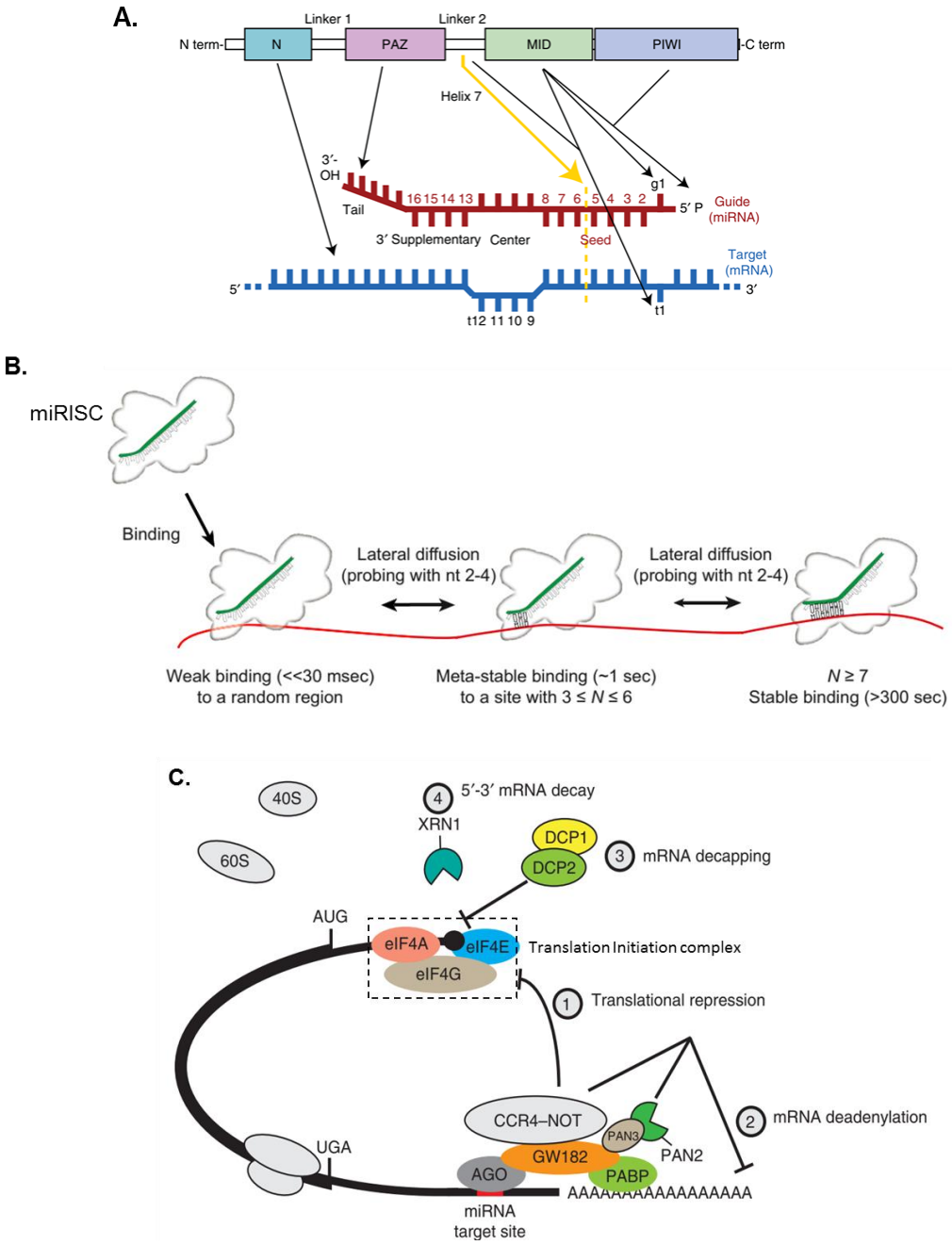


Figure 1.6.2. The mechanism and consequences of miRNA-target interaction.
(A) Schematic representation showing points of contact of Ago2 protein with miRNA and mRNA target.
(B) Illustration of proposed steps in target recognition by Ago2-miRNA. Ago2-miRNA binds the single stranded target RNA and diffuses along the RNA rapidly —sliding over 69nt in less than 100 millisec. Complementary sites of N≥3 cause the complex to pause and remain meta-stably bound (on the order of 1 sec). When site complementarity extends to the full length of the seed (nt 2–8) or longer, the complex remains bound stably for extended times (>>300 sec)
(C) Model of miRNA-mediated gene silencing. Step 1, repression of cap-dependent translation via the CCR4–NOT complex: miRNA-induced silencing complex (miRISC)

inhibits translation at the initiation step in a deadenylation-independent manner. Step 2, messenger RNA (mRNA) deadenylation: deadenylation is mediated by the CCR4-NOT complex and, to a lesser extent, the PAN2–PAN3 complex. Step 3, mRNA decapping: following deadenylation, the 5' cap moiety, denoted by a black circle, is removed by the DCP1–DCP2 decapping complex. Step 4, mRNA decay: once the mRNA 5'cap is hydrolyzed, the mRNA is degraded by the XRN1 5'-3' exonuclease. (Adapted from Duchaine and Fabian, 2018; Chandradoss et al., 2016).

However, once the RISC complex stably binds to target mRNA, the GW182 protein of this complex recruits several other proteins which, in turn, regulate target stability and protein expression in the following steps: The first step is repression of translation by dissociation of translation-initiation complex. The second step involves deadenylation of target mRNA by a deadenylase complex consisting of CCR4-NOT, and PAN2 and 3 polyA nucleases. In the third step, the 5'cap of mRNA is removed by DCP1 & 2 decapping complex. Once the 5'cap is hydrolyzed, the target is degraded by exonuclease XRN1 as the fourth step. Therefore, it is now widely accepted that miRNA-mediated gene silencing involves translational repression of target mRNA followed by its de-adenylation and subsequent decay (Fig. 1.6.2C).

1.7. Control of B and T lymphocyte development by microRNAs

The functional role of miRNAs is highly complex as a single miRNA can target multiple mRNAs and a given mRNA transcript can be targeted by more than one miRNA (Bartel 2009; Lewis et al., 2005). Given the ability of miRNAs to target a large number of genes simultaneously, they have been implicated in regulation of a number of genes that are critical for various physiological processes, including haematopoietic stem cell maintenance and differentiation (Chen et al., 2004; Monticelli et al., 2005; Petriv et al., 2010).

Studies involving lymphocyte-specific knockout of RNAi components have elucidated a significant role for microRNAs in lymphocyte survival, differentiation and function. Deletion of Dicer and DGCR8 during early stages of B cell development has shown to cause a developmental block at Pro-B cell stage, preventing their transition to Pre-B cells and a failure to upregulate Ig-H

expression (Koralov et al., 2008 and Brandl et al., 2016); whereas absence of Dicer in naïve B cells resulted in increased apoptosis and failure to produce antigen-specific, high affinity, class-switched antibodies (Xu et al. 2012, Daum et al., 2017). In thymocytes, depletion of Dicer at an early stage of T cell development decreased the survival of αβ T cells without significantly affecting the numbers of γδ T cells (Cobb et al. 2005). Moreover, deletion of Dicer at CD4+CD8+ DP stage severely decreases the maturation of CD8+ and CD4+ SP cells in the thymic periphery and causes defective T-helper cell proliferation and survival (Muljo et al., 2005).

Over the last decade, several studies have provided an overview of miRNA expression signatures in various developmental stages or subsets of lymphocytes (Landgraf et al., 2007; Neilson et al., 2007; Zhang et al., 2009; Kutchen et al., 2010; Petriv et al., 2010). However, the precise role played by microRNAs during lymphocyte differentiation and their ability to act as an integral part of mechanisms that determine the cell fate choice of multipotent progenitors, was yet to be addressed. In this direction, we aimed to identify the microRNAs expressed during early stages of B- and T-lymphocyte differentiation and validate their role in shaping the distinct gene expression programs of these lineages.

OBJECTIVES OF THE STUDY:

Previous studies have shown that multi-lineage transcriptional priming in progenitors allows promiscuous, low-level expression of genes associated with divergent cell-fates, thus allowing the progenitors to be 'primed' for differentiation towards various haematopoietic lineages. (Hu et al., 1997; Miyamoto et al., 2002). During progression to a specific lineage, the expression of lineage-inappropriate genes (i.e. genes that are fortuitously-expressed and are not required for development of that lineage) continues until the multipotent progenitors undergo commitment to the specified lineage. Thus, besides driving the expression of lineage-specific genes, the primary B and T cell-fate determinants (EBF1 or Notch1, respectively) are also responsible for repression of the inappropriately-expressed genes. To achieve this, EBF1 and Notch1 may exert a direct regulatory control not only at transcriptional level but also employ an additional layer of control, perhaps at the post-transcriptional level, to effectively regulate gene expression during the differentiation process. Considering this, we hypothesized that the primary lineage determinants potentially induce microRNAs (miRNAs) to repress the lineage-inappropriate genes and to fine-tune the random fluctuations in transcript abundance, perhaps providing robustness to the gene expression programs that govern B and T cell differentiation.

In this direction, we utilized the multi-lineage developmental potential of *Ebf1-/-* progenitors and differentiated them into 'Early-B' cells or 'Early-T' cells, *in vitro,* in order to obtain novel insights into the regulatory role played by miRNAs during early stages of B- and T-lymphocyte differentiation. Keeping this in view, we framed four objectives:

1. To investigate the microRNAs that are expressed during early developmental stages of B and T lymphocytes.
2. To identify genome-wide targets of microRNAs expressed during Early-B and Early-T cell development.
3. To test the functional role of Early-B and Early-T cell microRNAs during lymphocyte commitment.
4. To integrate miRNAs into the gene regulatory networks that dictate early developmental program of B and T lymphocytes.

CHAPTER-2: MATERIALS AND METHODS

2.1. Materials

The following list of Molecular biology reagents, kits and other consumables were used in this study.

2.1.1. Chemicals and Reagents

Table 2.1. List of Chemicals, enzymes and reagents used in this study.

Chemicals, Enzymes and Molecular biology reagents	Source
Ampicillin	Sigma-Aldrich
Agarose	SeaKem
Bacterial Agar	Himedia
Bovine Serum Albumin (Fraction V)	Roche
Chloroform	Sigma-Aldrich
Complete Protease Inhibitor Cocktail Tablets, EDTA-free	Roche
DNA polymerase I, Klenow Fragment	New England Biolabs
Deoxyadenosine triphosphate (dATP)	Invitrogen
Deoxycytidine triphosphate (dCTP)	Invitrogen
Deoxyguanosine triphosphate (dGTP)	Invitrogen
Deoxythymidine triphosphate (dTTP)	Invitrogen
Dithiothreitol	Promega
DH5α E.coli strain	Invitrogen
Ethanol, Molecular Biology grade	Hayman
EDTA	Sigma-Aldrich
Superscript III Reverse Transcriptase	Invitrogen
Hydrochloric acid	Sigma-Aldrich
Isopropanol	Sigma-Aldrich
Luria Bertani Broth	Himedia
Magnesium Chloride	Sigma
Nuclease-free water	Ambion
Nonidet P-40/IGEPAL	Sigma-Aldrich
Phusion Polymerase	New England Biolabs
Phosphatase, Calf Intestinal	New England Biolabs
Power Sybr Green Mastermix	Applied Biosystems
Protein-G Dynabeads	Invitrogen
Random Hexamers	Invitrogen
Restriction Enzymes	New England biolabs

RNAsin plus RNase inhibitor	Promega
Sodium Chloride	Sigma-Aldirch
Sodium Hydroxide pellets	Sigma-Aldrich
Taq polymerase	New England Biolabs
T4 Polynucleotide kinase	New England Biolabs
T4 DNA ligase	New England biolabs
Tris base	Sigma-Aldrich
Trizol/Trizol-LS	Invitrogen

2.1.2. Disposable material and Kits

Table 2.2. List of Disposable material and kits used in this study.

Disposables / Kits	**Source**
Cell strainer, 70µm	BD Biosciences
CellStar Cell culture plates (24-well)	Corning Inc.
CellStar Tissue culture dishes (60 mm)	Corning Inc.
CellStar Tissue culture flasks (T75, T25)	Corning Inc.
Cryovials, 1.5mL	Corning Inc.
DNA LoBind Tubes, 1.5 mL	Eppendorf
Falcon Tubes (15 ml, 50 ml)	BD Falcon, Tarsons, Nunc
Filter tips	Rainin, Mettler Toledo
Gloves	Mettler Toledo
mirVana miRNA isolation kit	Ambion
Nucleospin Gel Extraction and DNA purification Kit	Mackherey Nagel
Pipette tips	Tarsons
Plasmid purification Midiprep/Maxiprep kit	QIAgen
Reaction Tubes (1.5 ml, 2 ml)	Tarsons
Agilent high-sensitivity bioanalyzer kits	Agilent Technologies
BCA reagent kit	ThermoFisher
NEBNext Ultra II Directional RNA Library Prep kit	New England biolabs
NEBNext Multiplex Oligos for Illumina	New England biolabs
Trueseq small RNA library preparation kit	Illumina

2.1.3. Cell culture media and reagents

Table 2.3. List of cell culture media and other reagents used in this study.

Cell culture media and reagents	Source
4-O-Hydroxy Tamoxifen (4-OHT)	Sigma-Aldrich
DMEM	Gibco
Fetal Bovine Serum (FBS)	Gibco
Flt3L	R&D Systems
Fugen6	Promega
GM-CSF	R&D Systems
HEPES	Sigma-Aldrich
IL-3	R&D Systems
IL-7	R&D Systems
L-glutamine	Gibco
MCSF	R&D Systems
MEM alpha	Gibco
Opti-MEM	Gibco
Penicilllin Streptomycin Solution	Gibco
SCF	R&D Systems
β-Mercaptoethanol	Invitrogen
Polybrene	Sigma-Aldrich

2.1.4. Antibodies

Table 2.4. List of Antibodies used for Immunoprecipitation or Flow Cytometry.

Antibody (Clone)	Source
Anti-mouse Ago2 polyclonal antibody (RIP grade)	Abcam
Anti-mouse CD11b (M1/70)	BD Pharmingen
Anti-mouse CD44-PE (IM7)	BD Pharmingen
Anti-mouse CD25 (PC61)	BioLegend
Anti-human CD2 (RPA-2010)	BioLegend
Anti-mouse CD8a (53-6.7)	BD Pharmingen
Anti-mouse CD4 (RM4.5)	BD Pharmingen

2.1.5. Technical Equipment

Table 2.5. List of Important technical equipment used in this study.

Technical Equipment	Source
96-well plate reader	Pecan Instruments
Bacterial Incubator and Shaker	Thermo Scientitifc
Bioanalyzer	Agilent
Biosafety Cabinet	Thermo Scientitifc
Cell culture Incubator	Thermo Scientitifc
Refrigerated Centrifuge Sorval Legend Xtr	Thermo Scientitifc
FACSAriaIII	BD Biosciences
LSRFortessa (SORP)	BD Biosciences
GeneAmp PCR System 9700	Applied Biosystems
Nanodrop ND 1000	Thermo Fisher
StepOne Plus Real-time PCR	Applied Biosystems
Thermomixer	Eppendorf
Tube holder Magnetic stand	Invitrogen
Hiseq2000 sequencer	Illumina

2.1.6. Databases and Software

Table 2.6. List of databases or analytical software used in this study.

Database /Software	Source
Bowtie 1.2.2	Langmead et al., 2009
DESeq2 (v 1.24.0)	Love et. al., 2014
miRbase (V22)	http://www.mirbase.org/
multiMiR	https://bioconductor.org/packages/release/bioc/html/multiMiR.html
MyGeneset	Immunological Genome Project (ImmGen)
PANTHER	Mi et al., 2019
FIMO, MEME suite	Grant et al., 2011
FlowJo	Treestar Inc.
Sigma Plot V12.3	Sigma Plot Inc.
Statistical analysis	Microsoft Excel 2013, GraphPad Prism7

2.1.7. List of primers

Table 2.7: List of qRT-PCR primers used for quantifying mRNA transcripts.

Primer name	Sequence
q-Hprt-FP	ACCTCTCGAAGTGTTGGATA
q-Hprt-RP	CAACAACAAACTTGTCTGGA
q-Il10ra-FP	TCACGATCTCCAGCCTGAGC
q-Il10ra-RP	TCCAGTGGAGGATGTGCTGG
q-Itga8-FP	CATACCTGATGCGCGCACAG
q-Itga8-RP	TGAGGGCCAGGGACAGTAGT
q-Gsn-FP	CAGAAGCCATGCTGCAGGTG
q-Gsn-RP	TGCACCGTTGGAGACCTTGT
q-Atg4d-FP	CATCCTCAGGAAAGCTGTGGAGA
q-Atg4d-RP	CATCTGCCTTGTACACTGTGCAG
q-Nfam1-FP	TGTTCTGGTCTCAAATCCACCGT
q-Nfam1-RP	CCTTCATACAGGCGTAGACCTCC
q-Ctnnd1-FP	TGAAGACTGTAAGCCACGGCATA
q-Ctnnd1-RP	CCTCTCTGAGCTTACGTTCCGAA
q-Fgr-FP	GACTCCATACAGGACCTAGTGCG
q-Fgr-RP	CTAGAGTCTGGGGCTTAGTGGTG
q-Fam126a-FP	ATATGTGTCTGTGGCTATCCCCG
q-Fam126a-RP	TCCTGTTAGCATCTGCACCATGA
q-Tgfbr3-FP	AGAGAGAGGTTACCCTGCATCT
q-Tgfbr3-RP	TCTCTCTGTCTTCACATGCCAC
q-Zfhx4-FP	AAGGAATGTCCAGCGAATGAG
q-Zfhx4-RP	TCAGAGGGTCTCCTGCATCTT
q-Maf-FP	GGGACGCCTACAAGGAGAAAT
q-Maf-RP	CATGAAAAATTCGGGAGAGGA
q-Ptprg-FP	CCTTGACTACGCCACCCTGTA
q-Ptprg-RP	TGCTGCTCCGTGGTGAAGATA
q-Ctdspl-FP	GGCAGGTCATTCCCGTACCA
q-Ctdspl-RP	TGTGCACCAGGGTTTCGTCT
q-Egr1-FP	GCACCTGACCACAGAGTCCT
q-Egr1-RP	CAGGGAGAAGCGGCCAGTAT
q-tCd8a-FP	ATGCTCTTGGCTCTTCCAGA
q-tCd8a-RP	GCTTCTCGTCCCACGTTATC
q-Spi1-FP	TATCGCCACATGGAGCTGGA
q-Spi1-RP	ACATCCGGGGCATGTAGGAA

q-Mac1-FP	GATTGATGGCTCCGGTAGCAT
q-Mac1-RP	ACTCGTCCGAGTACTGCATCA
q-Csf1r-FP	CACGGCTCATGCTGATGAGAA
q-Csf1r-RP	ACAGGTCCTCCGTGAGTACA
q-Csf3r-FP	GCTTAGTGCCATGGGAAGACA
q-Csf3r-RP	TGTTGGTGGTGAGGTGCATGA
q-Cebpa-FP	AGTCGGTGGACAAGAACAGCA
q-Cebpa-RP	TCCAGCACCTTCTGTTGCGT
q-Vpreb3-FP	CCTGCCTCTGCTCCTGATAG
q-Vpreb3-RP	AGCTGAGATGAGCGTCTTGG
q-Igll1-FP	GTGCTCGCCCCATAGGCTT
q-Igll1-RP	TGGCCTTGTTGGCCTGAAGA
q-Mb1-FP	GAACCGCATCATCACAGCAGA
q-Mb1-RP	TAGTCATCTGGCATGTCCACC
q-CD79b-FP	GCAGAAATGTGACAGCGCCAA
q-CD79b-RP	GCCATCTTTCAGTGTGTTCCG
q-Foxo1-FP	GCTGGAGGAGAGTGAGGA
q-Foxo1-RP	CTGGAAGTCCCCGCACAG
q-Pax5-FP	GGACCGCGTGTTTGAGAGA
q-Pax5-RP	GCTCGTCAAGTTGGCTTTCAT
q-Gata3-FP	CTTATCAAGCCCAAGCGAAG
q-Gata3-RP	GTACAGCCCACAGGCATTG
q-Tcf7-FP	TCAATCTGCTCATGCCCTAC
q-Tcf7-RP	GTGGACTGCTGAAATGTTCG
q-Notch1-FP	AACAGTGCCGAATGTGAGTG
q-Notch1-RP	CTCCCGCAGAAAGTGGAAG
q-Bcl11b-FP	GCAATCCTTTCAACCTGCTG
q-Bcl11b -RP	GTGTGGGTCCAAGTGATGG
q-Ptcra-FP	GAATCTTCGACAGCCAGGAG
q-Ptcra-RP	TCGAGCAGAAGCAGTTTGAA
q-Il2ra-FP	GCCACGCTTGCTGATGTTG
q-Il2ra-RP	TCGGGTGGGTCATACAGACA
q-Cd3e-FP	CCAGTGCTGGGACATTGCT
q-Cd3e-RP	ACAAGGTGGCTGGCAGACAA

Table 2.8. List of primers used for making vector constructs.

Primer name	Sequence
SV40PolyA-FP	TTAAGATACATTGATGAGTTTGGACAAAC
SV40PolyA-RP	AACTTGTTTATTGCAGCTTATAATGGTT
tCD8a-FP	CCACCATGTCACCGTTGACCCG
tCD8a-RP	TTATGCCGCGATGAGAGTGATGATCAAGGACA
pPGK-FP	GGGTAGGGGAGGCGCTTT
pPGK-RP	CGAAAGGCCCGGAGATGAG
mER-FP	GAAATGAAATGGGTGCTTCAG
mER-RP	TCAGATCGTGTTGGGGAAG
PU.1-FP	CCACCATGTTACAGGCGTGCAAAATG
PU.1-RP	AGTGGGGCGGGAGGCGC
EBF1-FP	CCACCATGTTTGGGATCCAGGAAAG
EBF1-RP	GTGGATCAAGCATGGGAGGGACAATCATGCCAG
hCD2-FP	CCACCATGGGCTTTCCATG
hCD2-RP	TTAGGAAGTTGCTGGATTCTGAGG
tCD4-FP	CCACCATGTGCCGAGCCATCTCT
tCD4-RP	TCATGCCGCGCAGAGGATGCAGAGCC

Table 2.9. List of primers used for reporter assays.

Primer name	Sequence
Egr1 UTR-FP	ACATGACAGCGACCTTTTCTC
Egr1 UTR-RP	GATACACACCACATATCCCATG
Itga8 UTR-FP	GCCACTAACACTACAGAAACAATCCAC
Itga8 UTR-RP	TTTACTGTCTGTGAGTTTGGGTCATG
Hecw2 UTR-FP	ACAGTCTTTCAGTCAACAGTTCTTGAC
Hecw2 UTR-RP	CGCCTTTGGTACTTTCAGTGCAAAAT
asGFP-FP	GTGAGCAAGGGCGAGGAGCTG
asGFP-RP	CTTGTACAGCTCGTCCATGCCGAG

Table 2.10. List of primers used for quantification of microRNAs.

Primer name	Sequence
Universal reverse primer	GTGCAGGGTCCGAGGT
U6-FP	GCCGTTCGTGAAGCGTTC
miR-125a-5p-FP	AGGCGTCCCTGAGACCCTTTA
miR-125b-5p-FP	AGCCGATCCCTGAGACCCTA
miR-351-5p-FP	AGGTCCCTGAGGAGCCCTT
miR-128-3p-FP	GGGCCTCACAGTGAACCG
miR-330-5p-FP	AGGCTCTCTGGGCCTGTG
miR-186-5p-FP	AGCCGCCAAAGAATTCTCCT
miR-298-5p-FP	ACGGCAGAGGAGGGCTG
miR-223-3p-FP	GCGCGTGTCAGTTTGTCAA
RTP-U6 (Rnu6)	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAAAAATATG
RT-miR-125a-5p	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTCACAGGT
RTP-miR-125b-5p	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTCACAAGT
RTP-miR-351-5p	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCAGGCTCA
RTP-miR-298-5p	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGGAAGAA
RTP-miR-223-5p	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGGGGTAT
RTP-miR-186-5p	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAGCCCAAA
RTP-miR-128-3p	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAAAGAGAC
RTP-miR-330-5p	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGCCTAAGA

Table 2.11. List of primers used for cloning microRNAs in expression vectors.

Primer name	Sequence
Pre-miR-125a-FP	TCTTCTGGGGCTTAGGGTATCT
Pre-miR-125a-RP	AGCCTGTGATCAGAGATTGAGT
Pre-miR-125b-FP	ACCCAACTCTAATTCCCAAGCT
Pre-miR-125b-RP	ACCCTTGCTGATTAGAAAGAAAACC
Pre-miR-351-FP	ATTCTACATGGCAGTGGGGAC
Pre-miR-351-RP	AGACAAAGCAGCACGATTTCAC
Pre-miR-186-FP	GACTGTGGAGTCTTGGTGTGG
Pre-miR-186-RP	GAACACTGAAGCTGTGAAGTAGTCA
Pre-miR-128-FP	TGTCCTAAAATCCAGCGCCGG
Pre-miR-128-RP	TTCTCCCTGCAGTTTGGTCGG
Pre-miR-330-FP	TTTGGGGTGTTGCTGGGTAGG
Pre-miR-330-RP	GAGAAAGAGGGGTAGGGGAGAAG

2.1.8. Sponge sequences for miRNA knockdown

Table 2.12. MicroRNA-sponges used in this study.

Sponge sequences (7x repeats) for each of the miRNAs used in knockdown experiments. The sponge sequences for miR-125a-5p, miR-125b-5p and miR-351-5p are represented in Orange, Green and Blue, respectively. While the sponge sequences for miR-186-5p, miR-128-3p and miR-351-5p are represented in Red, Pink and Violet, respectively. Each repeat is separated by 4-nt spacer (ATGA, represented in black). The sponges were ordered from GenScript Inc., USA.

Sponge	Sequence
Sp(125a-5p.125b-5p.351.5p)	**TCACAGGTTAAACCACTCAGGGAATGATCACAGGTTAAAC CACTCAGGGAATGATCACAGGTTAAACCACTCAGGGAATG ATCACAGGTTAAACCACTCAGGGAATGATCACAGGTTAAA CCACTCAGGGAATGATCACAGGTTAAACCACTCAGGGAAT GATCACAAGTTACTACTCAGGGAATGATCACAAGTTACTA CTCAGGGAATGATCACAAGTTACTACTCAGGGAATGATCA CAAGTTACTACTCAGGGAATGATCACAAGTTACTACTCAG GGAATGATCACAAGTTACTACTCAGGGAATGATCACAAGT TACTACTCAGGGAATGACAGGCTCAAAGGAAACTCAGGG AATGACAGGCTCAAAGGAAACTCAGGGAATGACAGGCTC AAAGGAAACTCAGGGAATGACAGGCTCAAAGGAAACTCA GGGAATGACAGGCTCAAAGGAAACTCAGGGAATGACAGG CTCAAAGGAAACTCAGGGAATGACAGGCTCAAAGGAAAC TCAGGGA**
Sp-(186-5p.128-3p.330-5p)	**GCCTAAGACATCACCCAGAGAATGAGCCTAAGACATCACC CAGAGAATGAGCCTAAGACATCACCCAGAGAATGAGCCT AAGACATCACCCAGAGAATGAGCCTAAGACATCACCCAG AGAATGAGCCTAAGACATCACCCAGAGAATGAAAGAGA CCCCACACTGTGAATGAAAGAGACCCCACACTGTGAATG AAAAGAGACCCCACACTGTGAATGAAAGAGACCCCACA CTGTGAATGAAAAGAGACCCCACACTGTGAATGAAAGAG ACCCCACACTGTGAATGAAAGAGACCCCACACTGTGAAT GAAGCCCAAAAGCCTATTCTTTGATGAAGCCCAAAAGCCT ATTCTTTGATGAAGCCCAAAAGCCTATTCTTTGATGAAGCC CAAAAGCCTATTCTTTGATGAAGCCCAAAAGCCTATTCTTT GATGAAGCCCAAAAGCCTATTCTTTGATGAAGCCCAAAAG CCTATTCTTTG**

2.2. Methods

2.2.1. Vectors and constructs

For all cloning purposes, the sequences encoding cDNAs of genes or markers were amplified using Phusion polymerase, using primers that contained restriction sites (for sticky end cloning) or primers that were treated with T4-polynucleotide kinase (for blunt end cloning). The vectors were digested using appropriate restriction enzymes and blunted using DNA polymerase I, Klenow fragment (wherever necessary), followed by treatment with Calf Intestinal Phosphatase (CIP) for removal of 5'-Phosphates (wherever necessary). The amplified insert and vector were gel extracted or purified with Nucleospin Gel extraction and DNA purification columns and ligated using T4 DNA ligase. The ligation mixes were transformed into competent DH5α strain of bacteria. The colonies obtained were screened by 'Colony PCR' using Taq DNA polymerase. The clones were confirmed by restriction digestion and Sanger's sequencing. All constructs that were to be transfected, were produced in an ultrapure form on a large scale, using the Plasmid Maxiprep/Midiprep columns.

For making MigR1-EBF1.ER or MigR1-PU1.ER constructs, full length cDNA of mouse EBF1 or PU.1 were sub-cloned from vectors containing these sequences and fused with C-terminal domain of Estrogen Receptor (ER), into retroviral vector MigR1-IRES-hCD2 or MigR1-IRES-GFP, respectively.

For making Sensor-pcDH reporter vector; the tCd8a reporter cassette was first assembled in a separate vector by step-wise insertion of mouse PGK promoter sequence, full length mouse CD8a sequence (with truncated cytoplasmic tail) and SV40 polyA. The assembled tCD8a reporter cassette was subsequently inserted in an antisense orientation into the upstream region of pcDH-pEF1a-CopGFP lentiviral vector (Systems Biosciences), to obtain 'Empty Sensor-pcDH' reporter vector, which has two bidirectional cassettes, one for expression of selection marker (asEGFP) and the other cassette (containing tCD8a) for measuring reporter activity.

For making the 3'UTR Sensor-pcDH construct, the regions of 3'UTRs of desired transcripts were amplified from mouse genomic DNA and inserted immediately downstream of tCD8a in the reporter cassette. The Control Sensor-pcDH reporter vector was constructed by inserting the anti-sense sequence of EGFP immediately after tCD8a sequence in the reporter cassette.

For making miRNA-sponge constructs, the sponge sequences, Sp(125a.5p-125b.5p-351.5p) and Sp(186.5p-128.3p-330-5p), were synthesized at Genscript Inc. Sponge for each miRNA was designed to have 7x repeats of sequences antisense to the respective miRNA (each repeat has a 3-4nt central bulge and is separated by a 4-nt spacer). The above sponges were inserted immediately downstream of tCD8a sequence in Empty Sensor-pcDH reporter vector to generate miR-sponge constructs. Control Sensor-pcDH construct was used as 'Control Sponge'.

For making miRNA expression constructs, pcDH-pEF1a-CopGFP vector was first modified to replace its pEF1a-CopGFP with IRES-tCD4 (truncated cytoplasmic tail), to obtain pcDH-IRES-tCD4 vector (referred to as 'Empty pcDH'). For expression of individual miRNAs, its precursor sequences along with the ~90-110bp flanking region were PCR amplified and inserted into Empty pcDH vector. (Pre-miR-125a-125b-351)-pcDH and (Pre-miR-186-128-330)-pcDH vectors were made by step wise insertion of the precursor plus flanking sequences of the respective miRNAs next to each other.

2.2.2. Cell culture

OP9 or OP9-DL1 stromal cells were cultured in MEMalpha complete medium (supplemented with 10% (v/v) FBS, β-mercaptoethanol (50 µM), Penicillin-Streptomycin (10 U/mL) and 2mM L-gluatamine). When confluent, the cells were split up using Trypsin (0.25%) and plated in fresh complete medium. The cells were passaged every 2 days.

Plat-E and 293T cells were cultured in DMEM complete medium (supplemented with 10% (v/v) FBS, β-mercaptoethanol (50 µM), Penicillin-Streptomycin (10 U/mL) and 12.5mM HEPES). When confluent, the cells were split up using Trypsin (0.25%) and plated in fresh complete medium. The cells were passaged every 2 days.

Ebf1-/- progenitors were maintained in culture on OP9 stromal cells in the presence of OptiMEM complete medium (containing 4%(v/v) FBS, β-mercaptoethanol (50 µM), penicillin and streptomycin (10 U/ml), supplemented with cytokines SCF (10 ng/mL, Flt3L (10 ng/mL) and IL-7 (5 ng/mL). The cells were passaged every 2 days by tapping. The harvested cells were either sub-cultured in fresh complete medium or were used for assays.

2.2.3. Viral transductions

Plat-E and 293T cells were used for transfection of retroviral and lentiviral constructs, respectively. The retroviral or lentiviral constructs were transfected along with packaging vectors, containing the viral gag-pol and envelope proteins into Plat-E or 293T cells, respectively, as described previously (Pongubala et al., 2008). The viral supernatants were collected after 48h of transfection and subsequently used for spin infection of *Ebf1-/-* progenitors in presence of polybrene (at final concentration of 10ug/mL) at 2500rpm for 2.5h at 25°C in 24-well cell culture plates, using a swinging bucket rotor . The spin-infected cells were replenished with fresh medium supplemented with cytokines and plated on OP9/OP9-DL1 stroma. After 48h, transduced cells were FACS-sorted based on the selection marker(s) and maintained on OP9/OP9DL1 stroma under appropriate conditions or used for differentiation assays.

2.2.4. *In Vitro* differentiation assays

For differentiation of *Ebf1-/-* progenitors towards B-lineage, MigR1-EBF1.ER vector was expressed in *Ebf1-/-* progenitors via retroviral transduction. The transduced cells were FACS-sorted, plated on OP9 stroma and cultured in presence of cytokines SCF (10ng/mL), Flt3L (10ng/mL) and IL-7 (5ng/mL) and subsequently treated with 4-O-Hydroxy Tamoxifen (4.OHT, at 1uM final concentration) for 2 days, before performing any further experimental analyses.

For differentiation towards T-lineage, *Ebf1-/-* progenitors were plated on OP9-DL1 stroma and cultured in presence of cytokines SCF (10ng/mL), Flt3L (10ng/mL) and IL-7(5ng/mL) for 5 days, following which experimental analyses were performed. Cells were replenished with fresh medium supplemented with cytokines, every 2.5 days (Schmitt et al., 2002; Pongubala et al., 2008).

For differentiation towards myeloid-lineage, *Ebf1-/-* progenitors were plated on OP9 stroma and cultured for 3 days in presence of myeloid-promoting cytokine cocktail consisting SCF (10ng/mL), Flt3L (10ng/mL), GMCSF(10ng/mL), MSCF (10ng/mL) and IL-3 (5ng/mL), to obtain myeloid cells expressing low levels of Mac1 ($Mac1^{Lo}$). Alternatively, *Ebf1-/-* progenitors were transduced with MigR1-PU1.ER vector, FACS-sorted, plated on OP9 stroma and then cultured for a period of 2 days in presence of myeloid-promoting cytokines SCF (10ng/mL), Flt3L (10ng/mL) GMCSF (10ng/mL), MSCF (10ng/mL), IL-3 (5ng/mL) along with 4-OHT (at 1uM

final concentration), to obtain myeloid cells expressing high levels of Mac1 ($Mac1^{Hi}$).

2.2.5. Flow Cytometry Analysis

For analysis or sorting of cells, single-cell preparations were made in MACS Buffer (1X PBS pH=7.4, 0.5% BSA, 2mM EDTA). Wherever required, the cells were stained with one or more of the following antibodies - Anti-mouse CD25-BV421 (PC61), Anti-mouse CD44-PE (IM7), Anti-mouse CD11b/Mac1-PE (M1/70), Anti-human CD2 FITC (RPA-2010), Anti-mouse CD8a-PE (53-6.7), Anti-mouse CD4-PE (RM4.5). The stained cells were analyzed or sorted using BD LSRFortessa (SORP) and/or BD FACSAria III (BD Biosciences), respectively. Data was analyzed using FlowJO software (TreeStar Inc.).

2.2.6. Isolation of total and small RNAs

For quantification of mRNA transcripts, total RNA was used. Cells were suspended in Trizol and homogenized by pipetting. About 0.2mL of Chloroform was added per mL of Trizol, followed by vigorous shaking and incubation at room temperature for 5min. The samples were centrifuged at 4°C for 15-20 minutes to allow phase separation of aqueous and organic phases. The upper aqueous phase was transferred to fresh tube, leaving behind the interphase and lower organic layer. RNA was precipitated by addition of equal volume of Isopropanol and centrifuged at 4°C for 30 minutes to obtain a pellet, which was washed with 75% Ethanol, allowed to dry for few minutes and subsequently suspended in nuclease-free water.

For isolation of microRNAs, mirVana miRNA isolation kit was used, as per manufacturer's instructions. Briefly, the large RNA fraction was first precipitated and collected on column using buffer containing 25% Ethanol. Subsequently, the Ethanol content in the flow-through is increased to 55% to precipitate and obtain the small RNAs (<200nt) fraction, which contains microRNAs.

2.2.7. Quantification of mRNA transcripts

For quantification of mRNA transcripts, cDNA was synthesized from total RNAs of samples using 10mM dNTPs mix, Random Hexamers and cDNA synthesis kit containing Superscript III Reverse Transcriptase, as per manufacturer's instructions. qPCR was performed using Power Sybr green master mix using gene-specific primers. *Hprt* was used as endogenous normalization control in all cases.

2.2.8. Quantification of microRNAs

For quantifications of miRNAs, the small RNA fractions obtained using the mirVana kit were used. The desired miRNA present in the small RNA fraction was converted to cDNA using 10mM dNTPs mix, cDNA synthesis kit containing Superscript III Reverse Transcriptase, and miRNA-specific stem-loop RT primer (Chen et al., 2005). The 3'-end of stem-loop RT primer is complementary to the 3'-end of miRNA of interest, thereby allowing selective reverse transcription of that miRNA. The qPCR is then carried out with miRNA-specific forward primer and Universal reverse primer using Power Sybr green PCR master mix. Small RNA U6 was used as normalization control for all miRNA quantifications.

2.2.9. Preparation of Small RNA libraries

Small RNA libraries were prepared from *Ebf1-/-* progenitors, Early-B and Early-T cells, in replicates, using TruSeq Small RNA sample preparation kit v2 (Illumina), as per manufacturer's instructions. Briefly, cell pellets were suspended in Trizol reagent (Ambion) and total RNA was isolated, of which, 1ug was used for addition of 3' and 5' adapters that ligate specifically to small RNA molecules resulting from Drosha/Dicer cleavage. cDNA was prepared from adapter-ligated RNAs using RT primer and Superscript III reverse transcriptase. The cDNA was then PCR-amplified used indexed primers to obtain small RNA libraries. The library fragments corresponding to microRNAs were size selected and extracted, checked for quality using Agilent 2100 Bioanalyzer and sequenced on Illumina's Hiseq2000 platform.

2.2.10. Analysis pipeline for identifying differentially expressed miRNAs

A total of 60 million reads (7–13 million reads per sample) corresponding to *Ebf1-/-*, progenitors, Early-B cells and Early-T cells (in duplicates) with 35nt size length were obtained from these sequencing reactions. Most of the adaptor trimmed reads (65-85%) were within the 18-24 bp, with good quality range (Phred score >20) and were aligned to the mouse miRNA database, miRbase (V22) using bowtie1.2.2, allowing up to 2 mismatches (Langmead et al., 2009). About 75% of the reads were successfully mapped to the reference database. The total counts of all unique miRNAs for each samplewere aggregated to further quantify the miRNA expression using DESeq2 (v 1.24.0). For DESeq2, we used the DESeq function, which estimates size factors and dispersions and finally fits a model in order to perform differential expression tests using negative binomial generalized linear models. MicroRNAs showing Fold Change (Log_2) ≥ +1.5 or ≤-1.5 with p-val≤0.01, were considered to be differentially up- or down-regulated, respectively.

2.2.11. Ago2-RNA Immunoprecipitation and Sequencing (Ago2-RIPSeq)

Cell lysates were prepared for *Ebf1-/-* progenitors, Early-B and Early-T cells (all in replicates). For this, cells were washed once with 1X PBS, pH=7.4 and lysed in Cell Lysis buffer (50mM Tris-HCl (pH=7.4), 150mM NaCl, 5mM MgCl2, 1.0% NP-40), supplemented with 1mM DTT, 1X Protease Inhibitor cocktail and RNasin plus RNase inhibitor (at 200 U/mL final concentration) and stored at -80°C until further use. The frozen lysates were thawed on ice and centrifuged at 10,000xg at 4°C for 15 minutes to remove debris and then pre-cleared by incubating with Protein-G dynabeads (Invitrogen) at 4°C for 30 minutes, with rotation. For each RIP sample, about 4mg of cleared protein lysates were immunoprecipitated with anti-Ago2 antibody (Abcam) for 7-8hrs at 4°C, followed by pull down using Protein-G Dynabeads to capture the Ago2/RISC-complexes. The beads were washed 5 times with Cell lysis buffer and suspended in Trizol. RNA was isolated as per the manufacturer's instructions. For Input samples, cleared protein lysates were directly used for RNA isolation using Trizol. The Input and IP RNA samples were used for preparation of mRNA libraries using NEBNext Ultra II Directional RNA

Library Prep kit (NEB) and NEBNext Multiplex Oligos for Illumina, as per the manufacturer's instructions.

The raw reads obtained after sequencing were filtered for quality and aligned to mouse transcriptome (mm10) using Tophat(V2). Read quantification was performed using htseq-count with UCSC RefSeq gene annotation. DESeq2 was used to get the differentially Ago2-enriched genes in Early-B or Early-T cells compared to *Ebf1-/-* progenitors, by performing the likelihood ratio test to check whether the IP enrichment is different in condition B than in A. For this, we used the design term: '~ assay + condition + assay:condition' (where the interaction term 'assay:condition' represents the ratio of the ratios: (IP for B / Input for B) / (IP for A / Input for A)), to obtain statistically significant ($p \leq 0.05$) differentially Ago2-RIP enriched genes in Early-B or Early-T cells. Since transcripts targeted by miRNAs are usually found to undergo an overall decrease in their abundance (Guo et al., 2010), we accounted for any enrichment-bias (i.e. non-specific enrichment of a given transcript that may have resulted due to an increase in its expression levels during differentiation) using the filter, 'Input FC≤1'. Finally, transcripts that exhibited Ago2-RIP FC≥2 (p-val≤0.05) and Input FC≤1 were considered as 'miRNA targetome' for the respective lineages (286 and 781 transcripts in Early-B and Early-T cells, respectively).

2.2.12. RIP-qPCR analysis

For RIP-qPCR analysis, cell lysates were prepared as described in the previous section. The lysates from $20x10^6$ cells were used per IP. 5ug of either Anti-IgG or Anti-Ago2 antibody was incubated with the pre-cleared lysates for 7-8hrs at 4°C, followed by pull down using 50uL Protein-G Dynabeads to capture the Ago2/RISC-complexes. The beads were washed 5 times with Cell lysis buffer and suspended in Trizol for isolation of IP-RNA. Data is represented as Fold Enrichment of Ago2-IP vs IgG-IP.

2.2.13. Reporter Assays for measuring miRNA activity at 3'UTRs

For analysis of miRNA activity at 3'UTR of transcripts, Sensor-pcDH constructs containing the control (asEGFP) or 3'UTR sequences of *Hecw2*, *Egr1* or *Itga8* were transduced into *Ebf1-/-* progenitors either alone or along with MigR1-EBF1.ER-hCD2. The cells were sorted and differentiated into Early-T or Early-

B cells, respectively, as already described. For analysis of tCD8a expression by flow cytometry, the cells were harvested and stained with Anti-CD8a-PE. The cells were gated for CopGFP expression and then analyzed for measurement of tCd8a expression levels. The transcript levels of tCd8a were analyzed by qRT-PCR as described earlier.

2.2.14. Target prediction

The miRNAs upregulated in Early-B and Early-T cells were mapped with Ago2-enriched transcripts from the respective cell type, using R-package multiMiR. The user-defined cut off for target strength was set to 40%, in order to select only the top 40% predictions ranked by the primary score for each of the enlisted databases including MicroCosm, miRanda, miRDB, PicTar, TargetScan, DIANAmicroT, PITA, ElMMo, Tarbase, miRecords and miRTarbase. Subsequently, the miRNA-target interactions that were shown by at least two of the prediction tools were selected for further studies.

2.2.15. Comparative gene expression analysis using ImmGen

The Ago2-enriched transcripts in Early-B or Early-T cells, were submitted to 'MyGeneset' data browser (http://rstats.immgen.org/MyGeneSet_New/index.html) available at Immunological Genome Project (ImmGen), for comparative analysis of their expression levels across various immune cell sub-types. ImmGen's ULI RNA-Seq data option was used for the following populations that were selected for comparison: early stages of B cells (proB_CLP_BM, proB_FrA_BM and proB_FrBC_BM), early stages of T cells (preT_DN1_Th, preT_DN2a_Th, preT_DN2b_Th and preT_DN3_Th), NK cells (NKT_Sp, NK_27-11b+_BM, NK_27+11b-_BM and NK_27+11b+_BM), Dendritic cells (DC_4+_Sp**,** DC_8+_Sp and DC_pDC_Sp), Monocytes (Mo_6C+II-_Bl and Mo_6C-II-_Bl) and Granulocytes (GN_BM and GN_Sp).

2.2.16. Gene Ontology Analysis

Gene Ontology analysis was performed by submitting the gene list to the protein classification tool Protein ANalysis THrough Evolutionary Relationships (PANTHER, v14.0),

based on molecular function (Mi et al., 2019). Genes that play functionally related molecular roles were grouped together into broad categories and represented.

2.2.17. miRNA promoter analysis

H3K4me3 ChIP-Seq datasets from *Ebf1-/-* progenitors and Early-B cells (data unpublished) were analyzed to identify regions that gain H3K4me3 peaks during differentiation of *Ebf1-/-* progenitors into Early-B cells (unpublished data from our lab). The regions located 15kb upstream of miRNA precursor sequences were scanned for the presence of H3K4me3 marks and the peaks that were most proximal to the miRNA precursor sequences were considered to be the putative transcription start site (TSS) for the given miRNA. Subsequently, FIMO (Finding Individual Motif Occurrences, MEME suite) (Grant et al., 2011) was used to search for binding sites for B-lineage (EBF1 and PAX5) and T-lineage-specific (RBPJ, TCF1 and GATA3) transcription factors in the putative promoter region spanning 2kb upstream from the identified TSS.

2.2.18. Statistical analysis

Statistical analyses were performed using Microsoft Excel (2013) SigmaPlot (v12.3) and GraphPad Prism7. Data are expressed as means ± Standard Deviation (SD). Statistical significance was determined by unpaired two sample t-test. A value of $p \leq 0.05$ was considered statistically significant. Wherever indicated, statistical significance was assessed as *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns=not significant.

2.2.19. Data availability

All high-throughput sequencing data presented in this work were uploaded to the Gene Expression Omnibus (GEO) under one super-series with accession number GSE153544.

CHAPTER-3: RESULTS

3.1 *In vitro* differentiation of *Ebf1-/-* progenitors into Early-B and Early-T lymphocytes

Previous studies have demonstrated that restoration of transcription factor EBF1 in *Ebf1-/-* progenitors could drive their differentiation into committed $CD19^+$ pro-B cells (Pongubala et al., 2008). However, in order to achieve a more synchronized differentiation towards B-lineage and to capture the earliest changes that occur during this process, we expressed EBF1 in *Ebf1-/-* progenitors using an inducible retroviral construct encoding full length mouse EBF1 fused to C-terminal hormone binding domain of Estrogen Receptor (MigR1-EBF1.ER). The EBF1.ER-expressing *Ebf1-/-* cells were sorted and treated with 4-O-Hydroxy-Tamoxifen (4-OHT) for 2 days in presence of cytokines SCF, Flt3L and IL-7, to allow expression of EBF1 target genes (Fig. 3.1A). RT-PCR analysis showed robust expression of the well-known early B-lineage genes like *Vpreb3*, *Igll1*, *Mb-1*, *Cd79b*, *Foxo1* and *Pax5*, in EBF1.ER-*Ebf1-/-* cells treated with 4-OHT compared to untreated EBF1.ER-*Ebf1-/-* cells or *Ebf1-/-* progenitors (Fig. 3.1B). Moreover, the 4-OHT-treated EBF1.ER-*Ebf1-/-* cells showed repression of genes belonging to alternate lineages including *Tcf7*, *Gata3* and *Il2ra* (T lineage-specific) and *Cebpa* (Myeloid lineage-specific) which were reported to be downregulated upon EBF1 expression.

For differentiation towards T-lineage, the *Ebf1-/-* progenitors we cultured on OP-DL1 stroma in presence of cytokines SCF, Flt3L and IL-7, as described earlier (Schmitt et al., 2002). Flow Cytometry analysis revealed, by day-5 about 93% of *Ebf1-/-* cells cultured on OP9-DL1 were $CD44^+CD25^+$ (corresponding to DN2-stage of early thymocyte development) in comparison to *Ebf1-/-* cells cultured on OP9, that were mostly $CD44^+CD25^-$ (at DN1 stage) (Fig. 3.1C). Correspondingly, qRT-PCR analysis revealed a robust expression of early T-lineage genes including *Tcf7*, *Gata3*, *Il2ra*, *Ptcra, Bcl11b* and *Cd3e* in OP9DL1-cultured *Ebf1-/-* cells, indicating the advent of early-T lineage program (Fig. 3.1D). Therefore, by complementing *Ebf1-/-* progenitors with inducible EBF1 or culturing them under T-lymphoid conditions, we were able to efficiently generate the early B and T lymphocytes that are hereafter referred to as ‘Early-B’ and ‘Early-T’ cells, respectively.

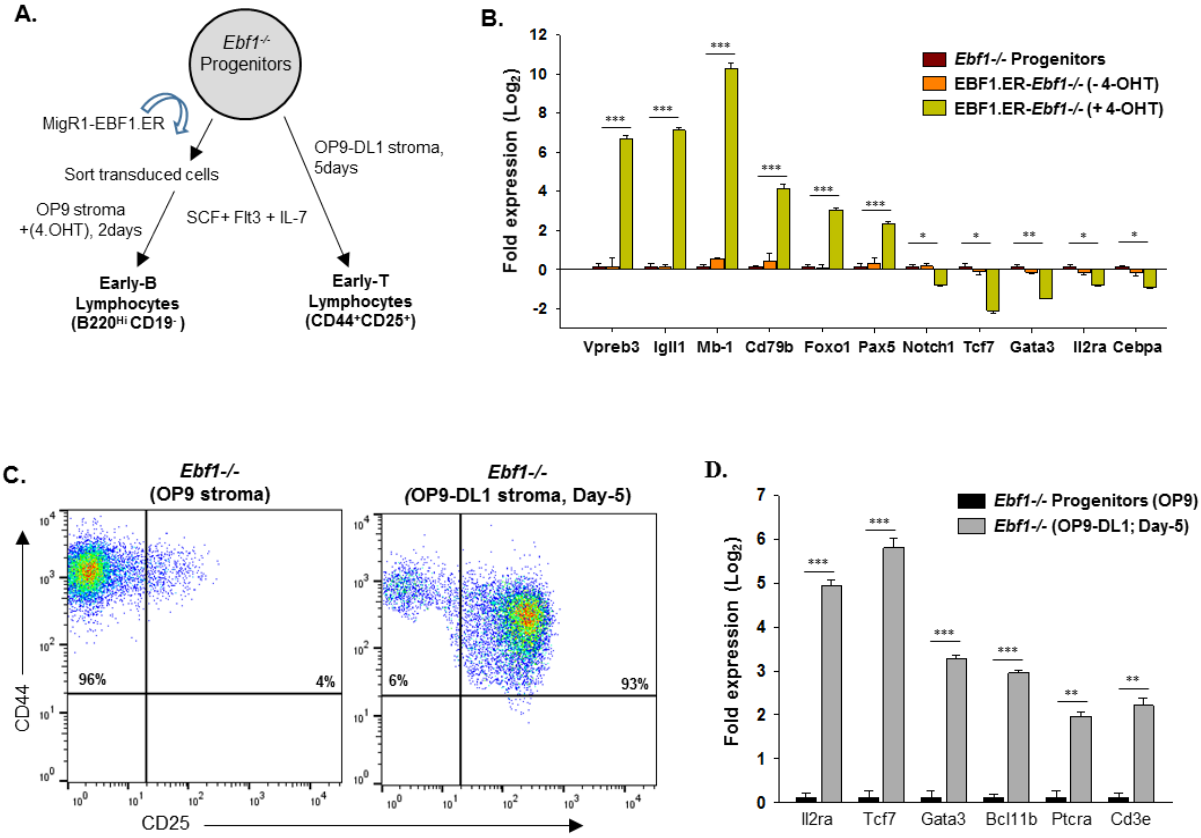


Figure 3.1. *In vitro* differentiation of *Ebf1-/-* progenitors into Early-B or Early-T cells.
(A) Experimental strategy for *in vitro* differentiation of *Ebf1-/-* progenitors into Early-B cells or Early-T cells. *Ebf1-/-* progenitors maintained under lymphoid-promoting conditions (OP9 stroma and cytokines – SCF, Flt3L and IL-7), were transduced with MigR1-EBF1.ER construct and cultured for 2 days on OP9 stroma in presence of 4-OHT, to obtain 'Early-B' cells. Alternatively, *Ebf1-/-* progenitors were transferred to OP9-DL1 stromal cells and cultured for 5 days to generate 'Early-T' cells.
(B) qRT-PCR analysis showing expression levels of early B-lineage genes (*Vpreb3, Igll1, Mb-1, Cd79b, Foxo1* and *Pax5*), early T-lineage genes (*Notch1, Tcf7, Gata3 and Il2r*) and myeloid gene (*Cebpa*) in *Ebf1-/-* progenitors, 4-OHT-untreated EBF1.ER-*Ebf1-/-* cells and 4-OHT-treated EBF1.ER-*Ebf1-/-* cells (at 1uM final concentration). Data is represented with *Ebf1-/-* progenitors as control.
(C) Flow cytometry analysis showing expression of early T-lineage markers, CD44 and CD25, on *Ebf1-/-* progenitors that were maintained on OP9 stroma or differentiated towards T-lineage by culturing on OP9-DL1 stroma for 5 days.
(D) qRT-PCR analysis showing expression levels of early T-lineage genes in *Ebf1-/-* progenitors maintained on OP9 stroma or cultured on OP9-DL1 stroma for 5 days. Data is represented with *Ebf1-/-* progenitors as control.
Data is shown as mean ± SD (*p ≤0.05; **p ≤0.01; ***p≤0.001).

3.2. MicroRNA signature of Early-B cells overlaps extensively with Early-T cells

As a first step to study the role of microRNAs in shaping the distinct gene expression programs initiated during B- and T-lineage commitment, we prepared small RNA libraries from total RNA of *Ebf1-/-* progenitors, Early-B and Early-T cells, in replicates. Subsequently, the library fragments corresponding to microRNAs were isolated and subjected to deep-sequencing (see Methods). After mapping the raw reads to mature miRNA sequences of mouse in miRBase (v22), we obtained about 7 million aligned reads per replicate of each of the three cell types (Fig. 3.2A). The aligned reads from each cell type were submitted to DeSeq2 to identify the differentially expressed miRNAs, each in Early-B and Early-T cells using *Ebf1-/-* progenitors as control. Finally, the microRNAs that showed Log_2-Fold Change, FC ≥+1.5 or FC≤-1.5 (p-value ≤ 0.01), were considered to be 'Upregulated' or 'Downregulated', respectively.

In Early-B cells, we identified a total of 150 miRNAs, of which, 76 were found to be upregulated while 74 were downregulated (Fig. 3.2B). The upregulated miRNAs include members of miR-181 family (miR-181a/b/c-5p), miR-125 family (miR-125a/b-5p and miR-351-5p), miR-221/222-3p, miR-199a/b-3p, miR-22-3p, miR-10a-5p, miR-142a-5p, miR-298-5p, miR-128-3p and miR-191-5p (Fig. 3.2C and Supplementary Table-1A). Whereas, members of miR-17 family (like miR-20a/b-5p, miR-106a/b-5p), miR-378 family (miR-378 a/b/c), miR-24-3p, miR-223-3p, miR-29b-3p and miR-146b-5p were found to be downregulated in Early-B cells (Figure 3.2C and Supplementary Table-1B).

In case of Early-T cells, a total of 199 miRNAs were found to be differentially expressed (Fig. 3.2B). Of these, 115 miRNAs were found to be upregulated, including miR-10b-5p, miR-22-3p and members of miR-125, miR-199, miR-181, miR-221/222-3p, miR-204/211, miR-212/132-3p, miR-486 and miR-99/100 families. On the contrary, 84 miRNAs were found to be downregulated in Early-T cells, including members of miR-17, miR-378, miR-467 and miR-669 families (Fig. 3.2D, Supplementary Table-2A and 2B).

Interestingly, upon careful comparison of miRNA expression profiles from both lineages, we identified that majority of differentially expressed miRNAs from Early-

B cells (~83%) exhibited a similar trend in expression pattern in Early-T cells (Fig. 3.2B and 3.2E). Consistently, these two cell types showed a significant correlation between their miRNA expression profiles, as indicated by the Pearson's correlation coefficient (r=0.85) (Fig. 3.2B and 3.2E). Therefore, we categorized the differentially expressed Early-B and Early-T cell miRNAs into two groups: 'common miRNAs' (Group-I) and 'unique miRNAs' (Group-II).

Within the Group-I, 66 miRNAs were found to be 'commonly-upregulated' in both the cell types and include miR-22-3p, miR-186-5p, miR-298-5p, miR-128-3p, miR-183-5p, miR-330-5p, miR-191-5p, miR-142a-5p as well as members from miR-181, miR-125, miR-199, miR-221/222, miR-130/301, miR-25/92/363/367, miR-326/330 and miR-200/429 families (Fig. 3.2B, 3.2E, 3.2H and Supplementary Table-3A). We would like to point out, of the 'commonly-upregulated' miRNAs ~77% (51 out of 66) were found to exhibit relatively higher fold-expression in Early-T cells than Early-B cells (Fig. 3.2J, Supplementary Table-3A). On the other hand, 58 miRNAs were found to be 'commonly-downregulated' in both cell types including miR-29b-3p, miR-142a-3p and members of miR-17, miR-378 and miR-669 families (Fig. 3.2B, 3.2E, 3.2H and Supplementary Table-3B).

Within the Group-II, 10 miRNAs were 'uniquely-upregulated' in Early-B cells compared to 49 in Early-T cells. Whereas, about 16 and 26 miRNAs were 'uniquely-downregulated' in Early-B and Early-T cells, respectively (Fig. 3.2B, Supplementary Table-4 and Table-5). However, we note that 13 miRNAs from Group-II displayed an 'inverse' expression pattern between Early-B and Early-T cells. For example, miR-127-3p, miR-210-3p, miR-30a-5p and miR-146b-5p were found to be upregulated in Early-T cells but were downregulated in Early-B cells. Conversely, miR-138-5p and members from miR-466 family were activated in Early-B cells but were downregulated in Early-T cells (Fig.3.2I and Supplementary Table-6). However, the fold change for majority of these inversely-expressed miRNAs was quite modest (within +/- 2-Fold). The expression levels of few of the miRNAs discussed above were validated using Stemloop qRT-PCR (Fig.3.2K).

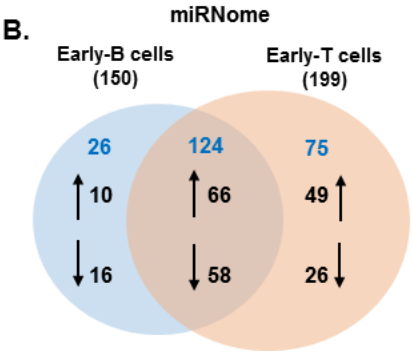
Interestingly, we note that many of the upregulated miRNAs that we discuss here, were shown to be implicated in proliferation and survival of HSCs (e.g. miR-125b) and development of B and T cells (e.g. miR-181 family) (Ooi et al., 2010; Chen et

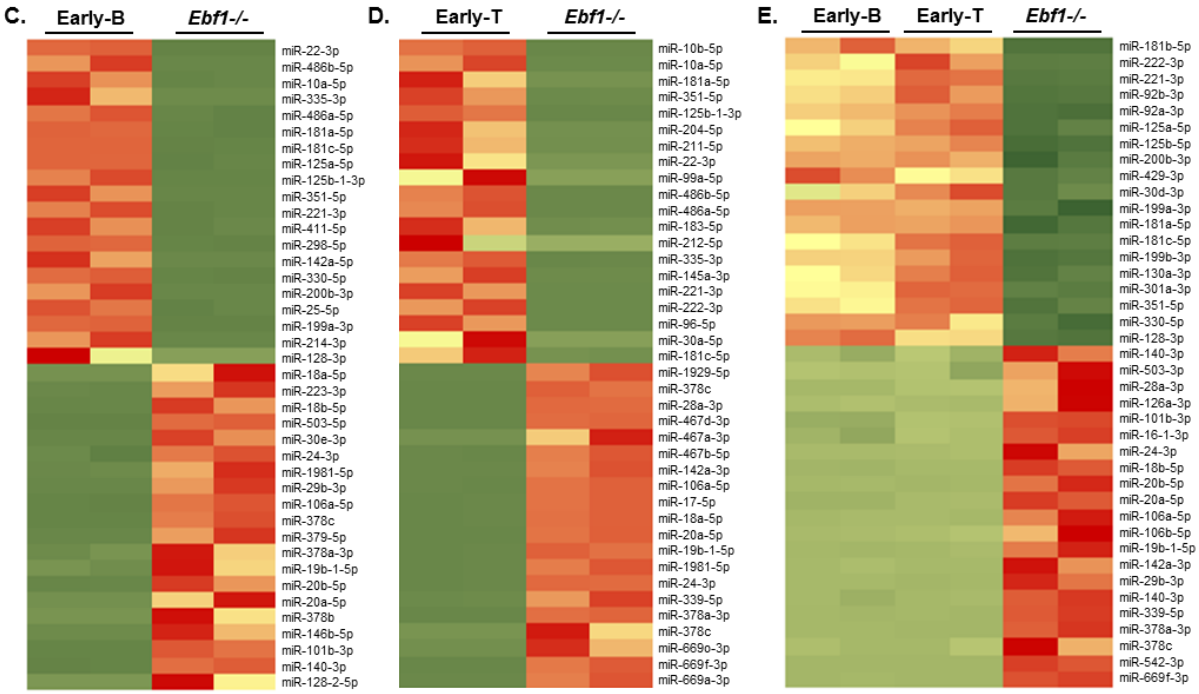
al., 2004; Li et al., 2007). On the other hand, miRNAs like miR-146a/b-5p, miR-223-3p and miR-23 cluster (miR-23a, miR-24 and miR-27a), which were previously reported to be expressed in myeloid cells, were found to be significantly downregulated in Early-B cells (Khalaj et al. 2014; Kurkewich et al., 2017; Rajasekhar et al., 2018). Similarly, miR-126 which is predominantly expressed in CLPs; and miR-30b which antagonizes Notch1, were both found to be downregulated during differentiation of *Ebf1-/-* progenitors into Early-B or Early-T cells (Petriv et al. 2010; Su et al., 2013). However, the miR-17~92 cluster which was reported to be downregulated during monocyte differentiation was found to be repressed even in Early-B and Early-T cells (Fontana et al., 2007).

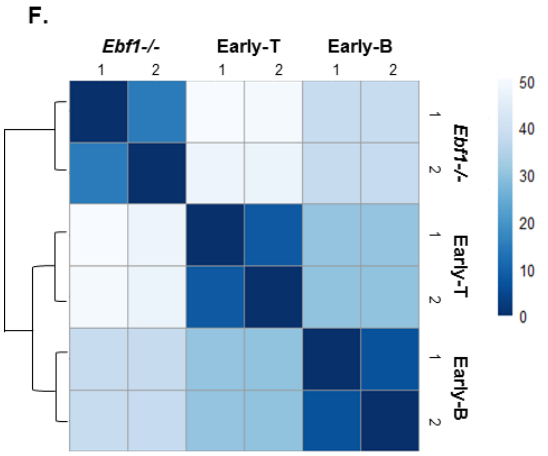
Collectively, our genome-wide miRNA analyses indicate that despite distinct identity and gene expression pattern, miRNA profiles of Early-B and Early-T cells are significantly correlated, with both cell types exhibiting a large fraction of commonly-regulated miRNAs and fewer unique miRNAs. Nonetheless, we note that the Early-T cells were found to exhibit a comparatively higher fold-expression for commonly-upregulated set of miRNAs besides the higher total number of upregulated miRNAs, compared to Early-B cells (115 vs 76, respectively) (Fig. 3.2B and 3.2J). With these observations, we decided to ascertain the functional role played by miRNAs expressed during differentiation of *Ebf1-/-* progenitors into Early-B or Early-T cells.

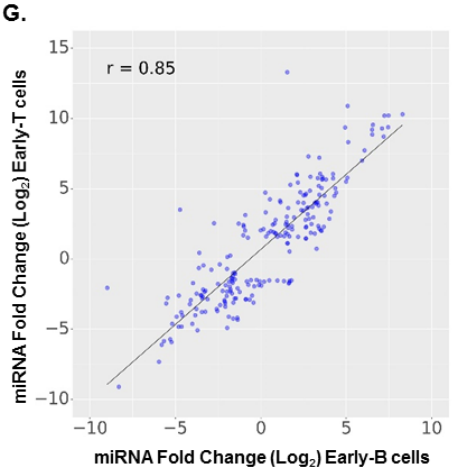
A.

SNO.	microRNA library	Reads aligned to miRbase (in Million)
1	*Ebf1-/-* Progenitors (Replicate 1)	7.26
2	*Ebf1-/-* Progenitors (Replicate 2)	8.07
3	Early-B cells (Replicate 1)	4.52
4	Early-B cells (Replicate 2)	5.89
5	Early-T cells (Replicate 1)	7.85
6	Early-T cells (Replicate 2)	7.72





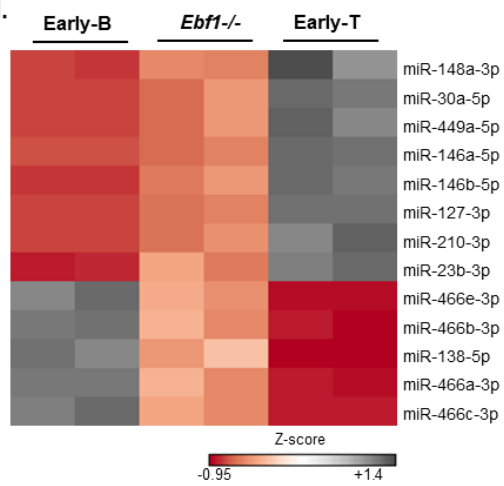




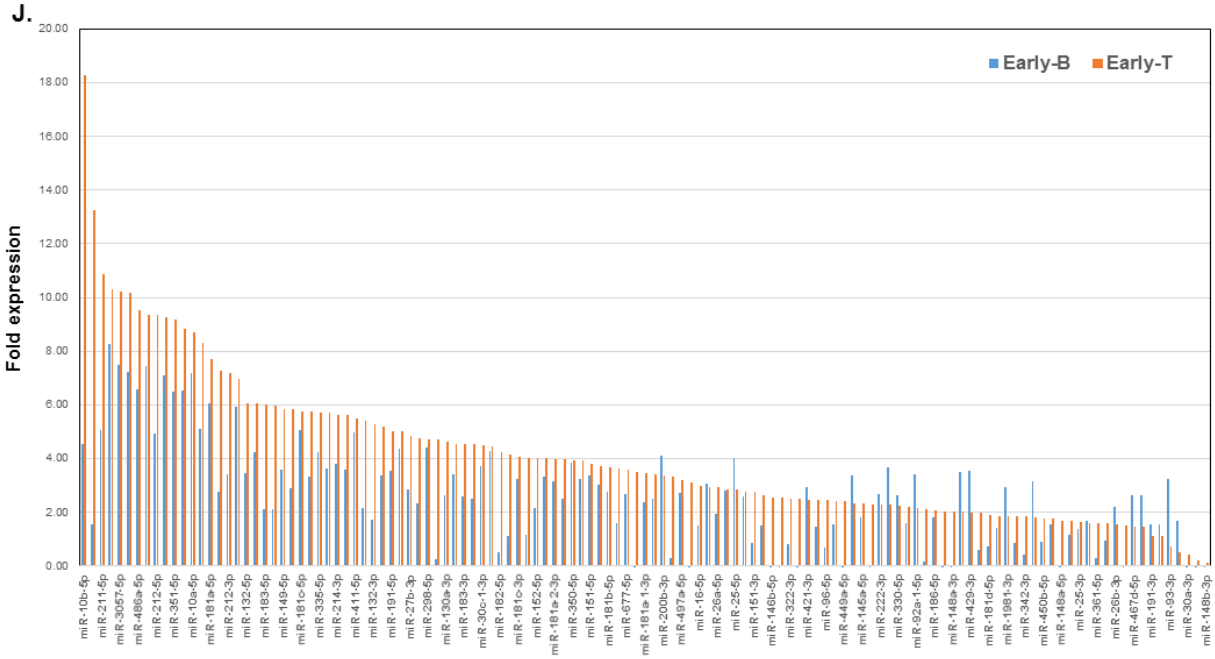
H.

miRNA family	Seed	Members expressed in Early-B and Early-T cells
COMMONLY-UPREGULATED		
miR-125ab-5p/351-5p	CCCUGAG	miR-125a-5p, miR-125b-5p, miR-351-5p
miR-181abcd/4262-5p	ACAUUCA	miR-181a-5p, miR-181b-5p, miR-181c-5p
miR-199ab-3p/3129-5p	CAGUAGU	miR-199a-3p, miR-199b-3p
miR-221/222/1928-3p	GCUACAU	miR-221-3p, miR-222-3p
miR-130ac/301ab-3p	AGUGCAA	miR-130a-3p, miR-301a-3p
miR-25/92/363/367-3p	AUUGCAC	miR-92a-3p, miR-92b-3p
miR-200bc/429-3p	AAUACUG	miR-200b-3p, miR-429-3p
miR-99/100-5p	ACCCGUA	miR-99a-5p, miR-99b-5p, miR-100-5p
COMMONLY-DOWNREGULATED		
miR-17/20ab/93/106ab-5p	AAAGUGC	miR-20a-5p, miR-20b-5p, miR-106a-5p, miR-106b-5p
miR-297abc-5p	UGUAUGU	miR-297a-5p, miR-297b-5p
miR-378a-3p/378bc	UGGACUU	miR-378a-3p, miR-378b, miR-378c
miR-18ab-5p/4735-3p	AAGGUGC	miR-18a-5p, miR-18b-5p

I.



J.



K.

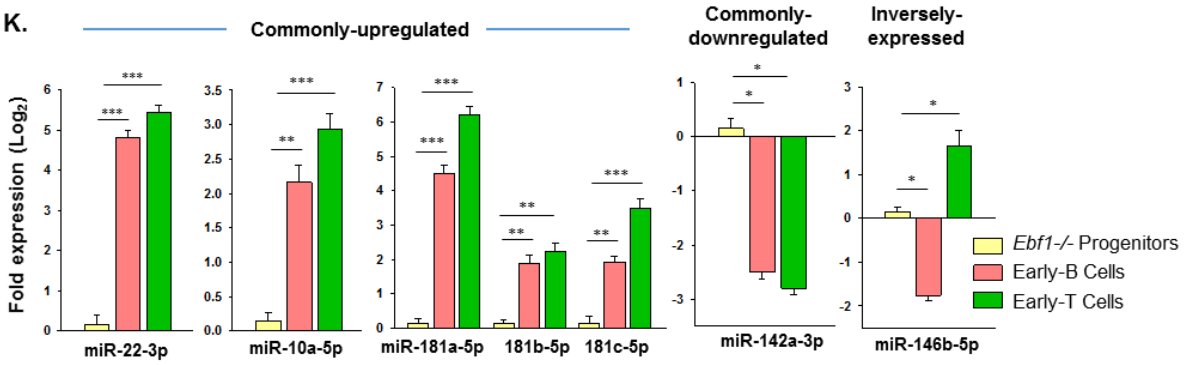


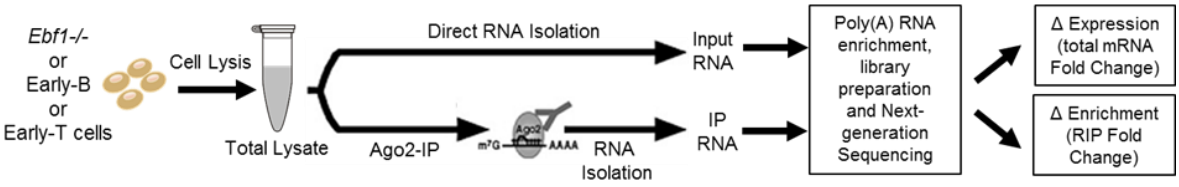
Figure 3.2. MicroRNA profiles of Early-B and Early-T cells.
(A) Number of aligned reads obtained after processing the sequenced reads from miRNA libraries of *Ebf1-/-* progenitors, Early-B and Early-T cells, in replicates.
(B) Venn diagram depicting the total number of differentially expressed miRNAs in Early-B and Early-T cells (146 and 199 miRNAs, respectively). The total miRNAs that are that are 'common' or 'unique' to each cell type are shown in blue while the upregulated or downregulated miRNAs are indicated by arrows.
(C) and (D) Heatmaps representing miRNAs that are differentially expressed by at least ±1.5-fold(Log_2) in Early-B or Early-T cells, respectively, in comparison to *Ebf1-/-* progenitors.
(E) Heatmap representing miRNAs that are commonly up- or downregulated in Early-B as well as Early-T cells, compared to *Ebf1-/-* progenitors.
(F) Heatmap showing sample distances between the miRNA profiles obtained from *Ebf1-/-* progenitors, Early-B and Early-T cells (in duplicates).
(G) Correlation plot showing fold expression (Log_2) of miRNAs in Early-B vs Early-T cells, indicating the Pearson's Correlation coefficient (r) between the miRNA profiles of these cell types.
(H) Table representing miRNA families and their corresponding members that are commonly up- or down-regulated in Early-B as well as Early-T cells.
(I) Heatmap showing miRNAs that exhibit inverse expression pattern in Early-B and Early-T cells.
(J) Comparison of fold-expression of upregulated miRNAs (identified from miRNA-seq data) in Early-B and Early-T cells.
(K) Experimental analysis of expression levels for few selected miRNAs that were differentially-expressed in Early-B and/or Early-T cells, using Stemloop qRT-PCR. Data is represented with *Ebf1-/-* progenitors as control. Data is shown as mean ± SD (*p ≤0.05; **p ≤0.01; ***p≤0.001).

3.3. Ago2-RIPSeq reveals microRNA targetome of Early-B and Early-T cells is largely distinct

Since miRNAs expressed in Early-B and Early-T cells were largely correlated, we speculated they perhaps regulate similar cellular processes or pathways in both the cell types. In order to investigate this, we sought to identify the genome-wide targets of miRNAs expressed in Early-B and Early-T cells. Using lysates of *Ebf1-/-* progenitors, Early-B and Early-T cells, we carried out Ago2-immunoprecipitation to capture and isolate the Ago2-associated mRNAs enriched within the cellular RISC complexes that were subsequently subjected to high-throughput sequencing (Ago2-RIPSeq). In parallel, we also carried out total mRNA-Seq (referred to as 'Input'), to quantify the expression levels of all mRNA transcripts present in each of these cell types (Fig. 3.3.1A and Methods). We obtained ~45 million aligned reads per 'Input' or 'Ago2-RIP' library (in duplicates) for each of these cell types (Fig. 3.3.1B). The Principal Component Analysis (PCA) revealed a significant distinction between Input and Ago2-RIP profiles of *Ebf1-/-* progenitors vs Early-B or Early-T cells (Fig. 3.3.1D).

Using DeSeq2, we identified the RISC-enriched, miRNA-targeted transcripts in three steps: Firstly, the Ago2-RIP data for all transcripts was normalized to the corresponding 'Input' data, for each of the three cell types: *Ebf1-/-* progenitors, Early-B and Early-T cells (Fig. 3.3.1C). Secondly, the normalized Ago2-RIP data from Early-B or Early-T cells was compared with that of *Ebf1-/-* progenitors to calculate the differential Ago2-RIP fold-enrichment of each transcript (represented as 'Ago2-RIP FC'). Thirdly, in order to account for the enrichment-bias that may have resulted owing to an increase in expression levels of a given Ago2-enriched transcript during differentiation, we implemented a cut-off for 'total mRNA-seq' or 'Input' as FC≤1, due to the fact that the miRNA-targeted genes were usually found to undergo an overall decrease in their mRNA transcript levels (Guo et al., 2010). From the above analyses we identified, about 286 transcripts in Early-B cells and 781 transcripts in Early-T cells, that exhibited Ago2-RIP FC≥2 ($p \leq 0.05$) and corresponding Input FC≤1, were considered to be significantly Ago2-enriched and thus referred to as 'miRNA targetome' of the respective cell type (Fig. 3.3.1C and Supplementary Table-7A and 7B).

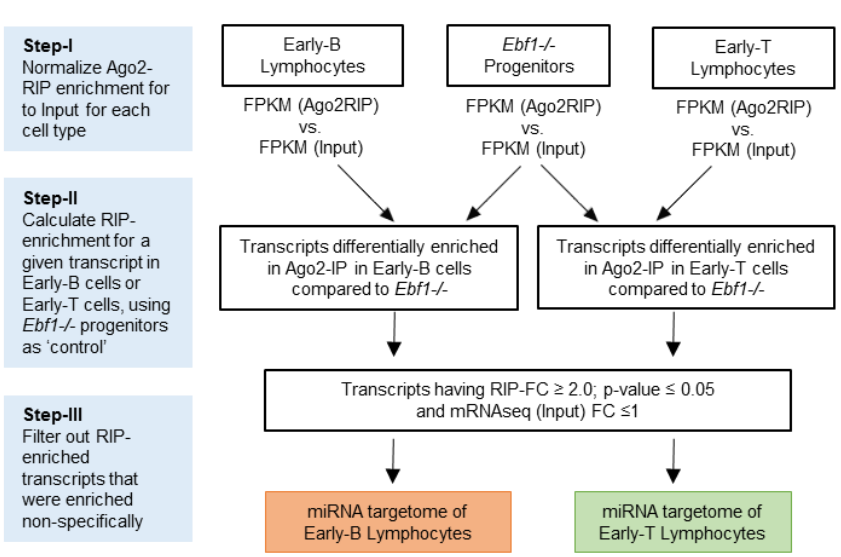
A.



B.

S.NO	Sample	Number of Raw reads per sample (Million)	% reads ≥ Q30	Mean Quality score	Number of Aligned reads per sample (Million)
1	Input_Ebf1-/-_Rep1	56.98	97.54	38.35	42.45
2	Input_Ebf1-/-_Rep2	54.46	97.59	38.36	50.54
3	RIP_Ebf1-/-_Rep1	58.34	97.06	38.21	43.46
4	RIP_Ebf1-/-_Rep2	52.2	97.54	38.35	37.53
5	Input_Early-B_Rep1	53.98	97.47	38.31	50.20
6	Input_Early-B_Rep2	53.7	97.64	38.38	49.99
7	RIP_Early-B_Rep1	84.76	97.3	38.29	43.57
8	RIP_Early-B_Rep2	106.28	97.37	38.3	51.33
9	Input_Early-T_Rep1	51.16	97.58	38.34	48.14
10	Input_Early-T_Rep2	56.62	97.55	38.33	53.39
11	RIP_Early-T_Rep1	49.44	97.00	38.19	34.36
12	RIP_Early-T_Rep2	51.42	97.42	38.32	37.33

C.



D.

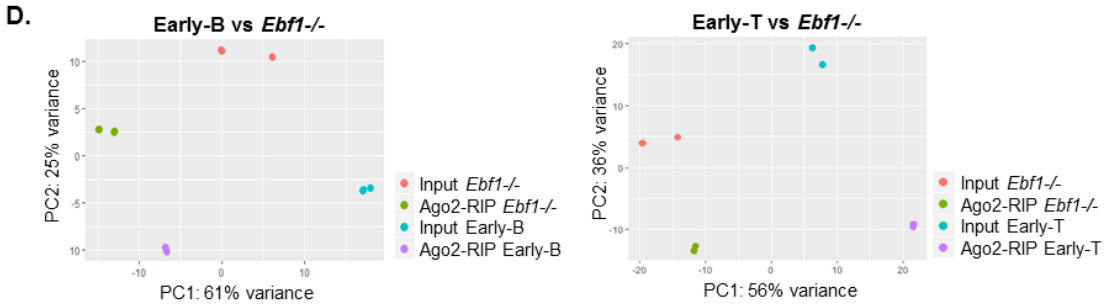


Figure 3.3.1. Ago2-RIPSeq data analysis pipeline for identifying microRNA targetome of Early-B and Early-T cells.
(A) A Schematic of experimental strategy for RNA-Immunoprecipitation and high-throughput sequencing of Ago2-associated mRNAs (Ago2-RIPseq).
(B) Summary of raw and aligned reads obtained from Input and RIP libraries of *Ebf1-/-* progenitors, Early-B and Early-T cells (in duplicates).
(C) Flowchart representing Ago2-RIP data analysis pipeline to identify miRNA targetomes of Early-B and Early-T cells.
(D) PCA plots showing distances between the 'Input' and 'Ago2-RIP' samples for *Ebf1-/-* progenitors vs Early-B cells or Early-T cells (left and right panels, respectively).

Given the significant correlation between miRNA profiles of Early-B and Early-T cells, we sought to determine whether their Ago2-RIP profiles were also correspondingly similar. Surprisingly, we found that only few (35 transcripts) were 'commonly Ago2-enriched' between Early-B or Early-T cells, while the remaining transcripts (251 in Early-B cells and 746 in Early-T cells) were found to be cell type dependent. (Fig. 3.3.2C). For instance, transcripts encoded by genes *Foxc2, Sarm1, Klf8, Il10ra, Itga8,* etc. were found to be Ago2-enriched in both Early-B and Early-T cells (Supplementary Table-7C). Whereas, transcripts encoded by genes like *Fhl2, Eomes, Tbx21, Hecw2, Bcl6*, etc. were enriched only in Early-B cells while *Ctdspl, Cd36, Egr1, Egr3, Trib1,Maf, etc.* were enriched only in Early-T cells (Fig. 3.3.2A and 3.3.2B). Notably, the miRNA targetome of Early-T cells is about three times larger compared to Early-B cells (781 vs 286 transcripts, respectively (Fig. 3.3.2C, Supplementary Table-7A and 7B). We suspect these differences exist due to the larger number (115 vs 76 miRNAs) and higher expression levels of upregulated miRNAs in Early-T cells than Early-B cells. Furthermore, the total mRNA or Input levels of Ago2-enriched transcripts displayed an overall decrease as *Ebf1-/-* progenitors differentiated into Early-B or Early-T cells (Fig. 3.3.2D), indicating that the functional effect of miRNAs on gene expression is closely reflected by destabilization and a subsequent decrease in target mRNA levels.

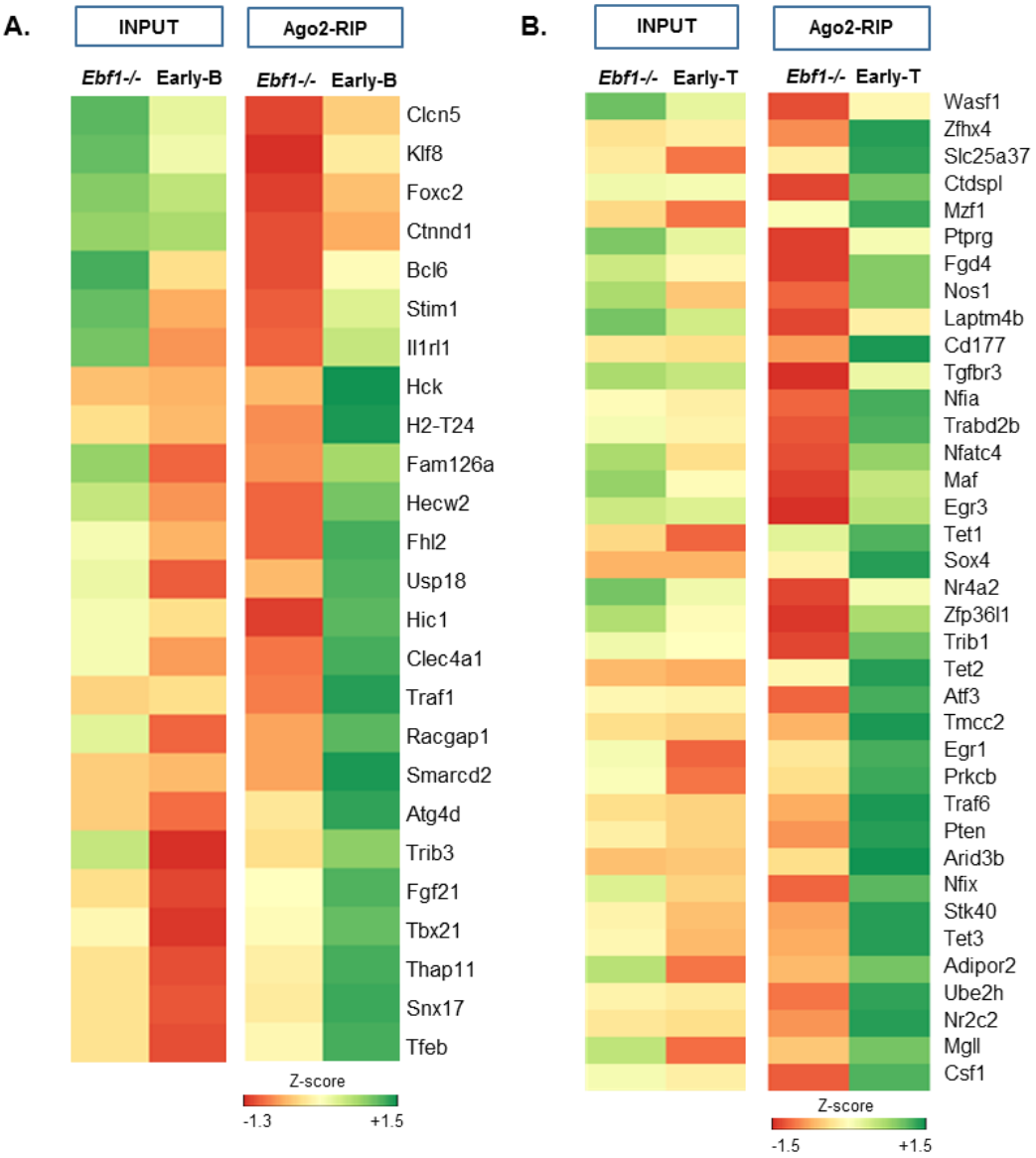
Next, we attempted to identify the miRNAs that potentially targeted each of the Ago2-enriched transcripts during Early-B or Early-T cell differentiation. To accomplish this, the upregulated miRNAs from Early-B or Early-T cells were mapped to Ago2-enriched transcripts of the corresponding lineage using an R-package called 'multiMiR', which provides miRNA-target pair interactions from multiple databases based on user-defined cut-offs (Ru et al., 2014). In order to

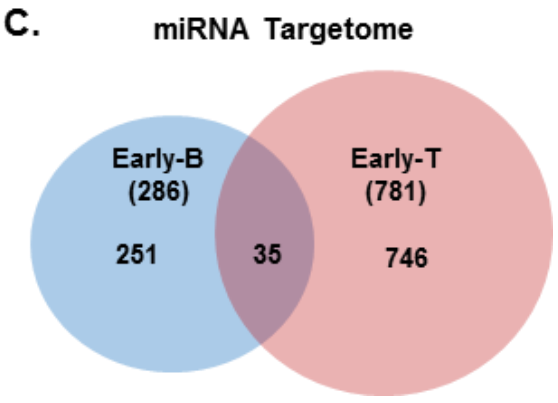
identify the miRNA-target pairs with high confidence, we selected only the top 40% of the targets predicted for all miRNAs by each database represented in multiMiR. Subsequently, only the miRNA-target pairs that were enlisted by at least two databases were selected for further analysis. Using these criteria, we were able to identify the miRNA-target pairs for about 80% of Ago2-enriched transcripts, in each cell type.

Subsequently, we sought to identify the transcripts that are regulated by the 'common' or 'unique' set of miRNAs in each cell type. We observed that a major fraction (~83.5%; 176 transcripts) of the Ago2-enriched transcripts of Early-B cells were mapped to commonly upregulated miRNAs, whereas a very small fraction of transcripts (~0.5%) were mapped to miRNAs that were unique to Early-B cells. Also, about 16% of the Early-B transcripts were found to be jointly targeted by both common as well as unique miRNAs (Fig. 3.3.2E, left panel). On the contrary, in Early-T cells, only few of the total Ago2-enriched transcripts were targeted specifically by either common or unique sets of miRNAs (~11.6% and 6.2%, respectively) while majority of the enriched transcripts (~82.2%) were found to be jointly targeted by common as well as unique miRNAs (Fig. 3.3.2E, right panel). This implies, nearly all the Ago2-enriched transcripts in both these lineages (99%; 210 transcripts in Early-B cells and 93.8%; 590 transcripts in Early-T cells) possess binding sites for commonly-upregulated miRNAs. Consistent with this observation, we noticed that the commonly-upregulated miRNAs like miR-186-5p, miR-214-3p, miR-27b-3p, miR-128-3p and members from families like miR-181, miR-125/351, miR-34b/449a, miR-132/212, miR-204/211 or miR-130a/301a, were shown to target a large number of transcripts, each in Early-B as well as Early-T cells (Fig. 3.3.2F).

Thus, it may be inferred, the upregulated miRNAs that are shared commonly between Early-B and Early-T cells perhaps target distinct sets of genes in these lineages. However, the differences in the distribution of common and unique miRNA binding sites within the Ago2-enriched transcripts of Early-B and Early-T cells (Fig. 3.3.2E) indicates that, besides the common miRNAs, the relatively larger number of unique miRNAs expressed in Early-T cells (than Early-B cells; 49 vs 10

miRNAs, respectively) accounts for the enlarged Early-T cell target repertoire compared to Early-B cells (781 vs 286, respectively).





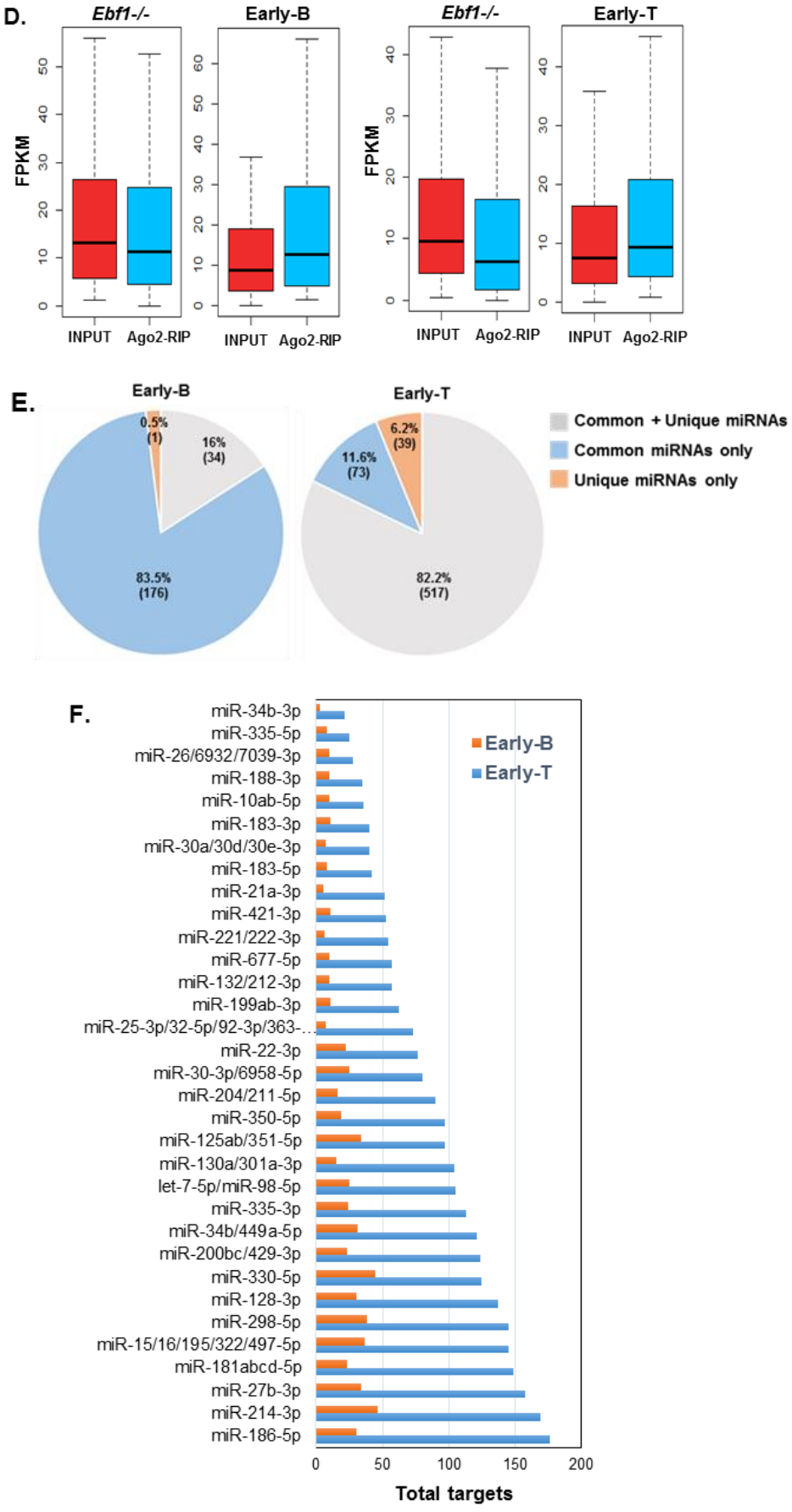

D.
Ebf1-/-
Early-B
Ebf1-/-
Early-T
FPKM
INPUT
Ago2-RIP
E.
Early-B
Early-T
0.5% (1)
16% (34)
83.5% (176)
6.2% (39)
11.6% (73)
82.2% (517)
Common + Unique miRNAs
Common miRNAs only
Unique miRNAs only
F.
miR-34b-3p
miR-335-5p
miR-26/6932/7039-3p
miR-188-3p
miR-10ab-5p
miR-183-3p
miR-30a/30d/30e-3p
miR-183-5p
miR-21a-3p
miR-421-3p
miR-221/222-3p
miR-677-5p
miR-132/212-3p
miR-199ab-3p
miR-25-3p/32-5p/92-3p/363-...
miR-22-3p
miR-30-3p/6958-5p
miR-204/211-5p
miR-350-5p
miR-125ab/351-5p
miR-130a/301a-3p
let-7-5p/miR-98-5p
miR-335-3p
miR-34b/449a-5p
miR-200bc/429-3p
miR-330-5p
miR-128-3p
miR-298-5p
miR-15/16/195/322/497-5p
miR-181abcd-5p
miR-27b-3p
miR-214-3p
miR-186-5p
Early-B
Early-T
0
50
100
150
200
Total targets

Figure 3.3.2. MicroRNA targetomes of Early-B and Early-T cells are distinct.
(A) and (B) Heatmaps representing relative enrichment levels of Ago2 RIP-enriched transcripts in Early-B and Early-T cells, respectively, in comparison to *Ebf1-/-* progenitors.
(C) Venn diagram showing miRNA targetome of Early-B and Early-T cells. Total number of Ago2-enriched transcripts are represented in brackets while the number of Ago2-enriched transcripts that are 'common' or 'unique' to each cell type is indicated within the circles.
(D) Box plots representing the 'Input' (total mRNASeq) levels for the set of transcripts enriched in Ago2-IP in Early-B cells and Early-T cells when compared to *Ebf1-/-* progenitors, (left and right panels, respectively).
(E) Pie charts indicating the binding sites of common, unique or both sets of miRNAs in Ago2-enriched transcripts from Early-B and Early-T cells.
(F) Total number of transcripts targeted by commonly-upregulated miRNA/miRNA families in Early-B or Early-T cells (shown by orange and blue bars, respectively), as predicted by multiMiR.

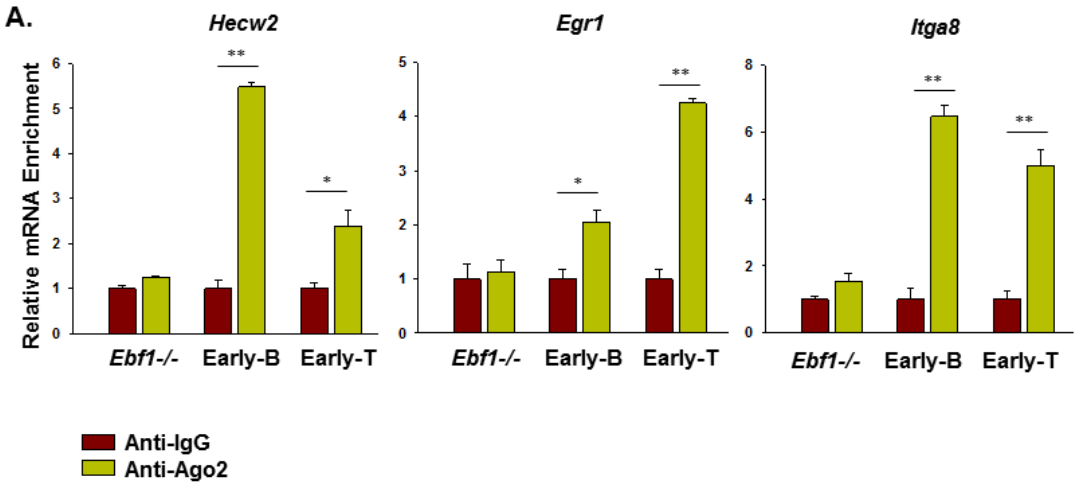
3.4. 3'UTR Reporter assays validate lineage-specific enrichment of RISC-bound transcripts

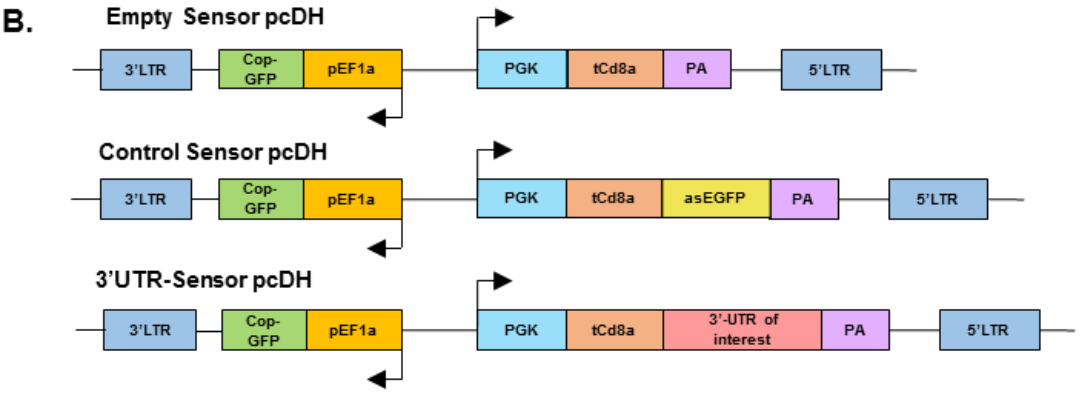
In order to validate if the distinctly Ago2-enriched transcripts in Early-B and Early-T cells were targeted by miRNAs in a truly lineage-specific manner, we performed reporter assays using 3'UTRs of transcripts of three selected Ago2-enriched transcripts - *Hecw2* (enriched only in Early-B cells), *Egr1* (enriched only in Early-T cells) and *Itga8* (enriched in both B and T cells) (Fig. 3.4A). The miRNAs that potentially bind to and regulate the 3'UTR regions of these transcripts are represented in Fig. 3.4C. To perform the reporter assays, we first constructed a lentiviral reporter vector named 'Empty Sensor-pcDH', which contains two bidirectional cassettes: one cassette expressing the CopGFP (Copepod GFP; driven by pEF1a promoter) as a selection marker and another cassette encoding the tCD8a (truncated CD8a; driven by PGK promoter) as a 'sensor or reporter' for measuring miRNA activity (Fig. 3.4B). The 3'UTR sequences of each of the above three Ago2-enriched transcripts or a control sequence (i.e., asEGFP, anti-sense sequence of Enhanced GFP) was inserted after the tCd8a reporter sequence within the 'Empty Sensor-pcDH' vector to generate the constructs, 'Control Sensor-pcDH' or '3'UTR-Sensor-pcDH', respectively (Fig. 3.4B).

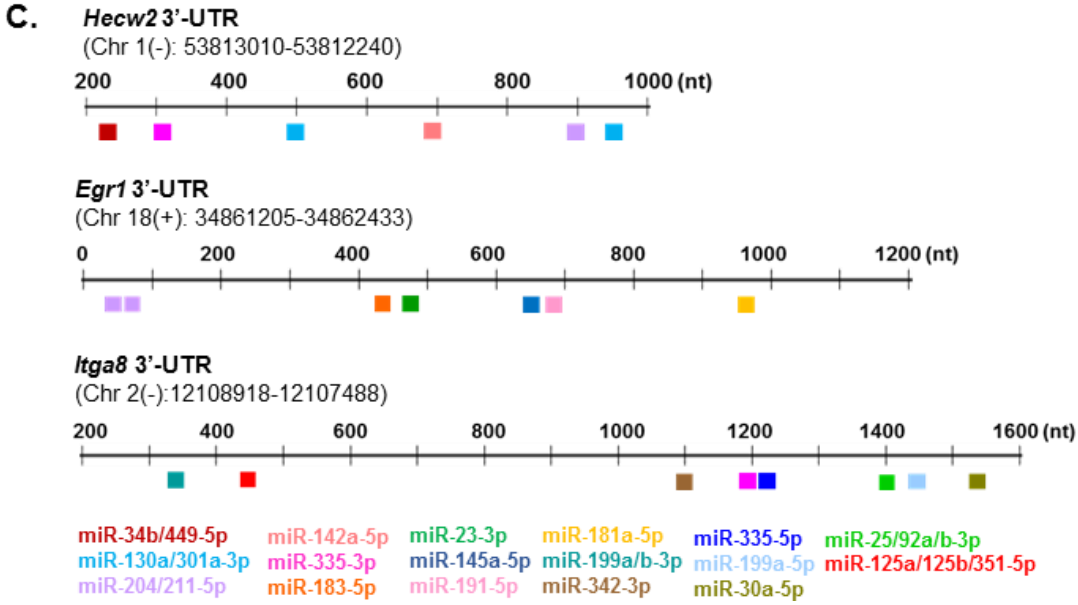
Subsequently, we transduced Control Sensor-pcDH vector or Sensor pcDH vector containing 3'UTR region of each of these transcripts into *Ebf1-/-* progenitors and allowed them to differentiate into Early-B or Early-T cells. Upon differentiation, cells were harvested and expression levels of tCD8a was analyzed by flow cytometry and qPCR. As shown in Fig. 3.4D, *Ebf1-/-* progenitors transduced with Control Sensor-pcDH vector exhibited high expression of tCD8a (tCD8a^{Hi}, 91.5%) which was found to be expressed at almost similar levels even upon differentiation into Early-B (tCD8a^{Hi}, 86%) or Early-T cells (tCD8a^{Hi}, 83%). However, when *Hecw2*.UTR-Sensor-pcDH construct was expressed, the levels of tCD8a decreased more dramatically when *Ebf1-/-* progenitors differentiated into Early-B cells (tCD8a^{Hi}, 7.2%) as compared to Early-T cells (tCD8a^{Hi}, 61%). On the contrary, when *Ebf1-/-* progenitors were transduced with *Egr1*.UTR-Sensor-pcDH, the expression of tCD8a decreased more significantly upon differentiation into Early-T cells (tCD8a^{Hi}, 19.4%) than Early-B cells (tCD8a^{Hi}, 66.8%). Interestingly, *Ebf1-/-* progenitors transduced with *Itga8*.UTR-Sensor-pcDH construct were found

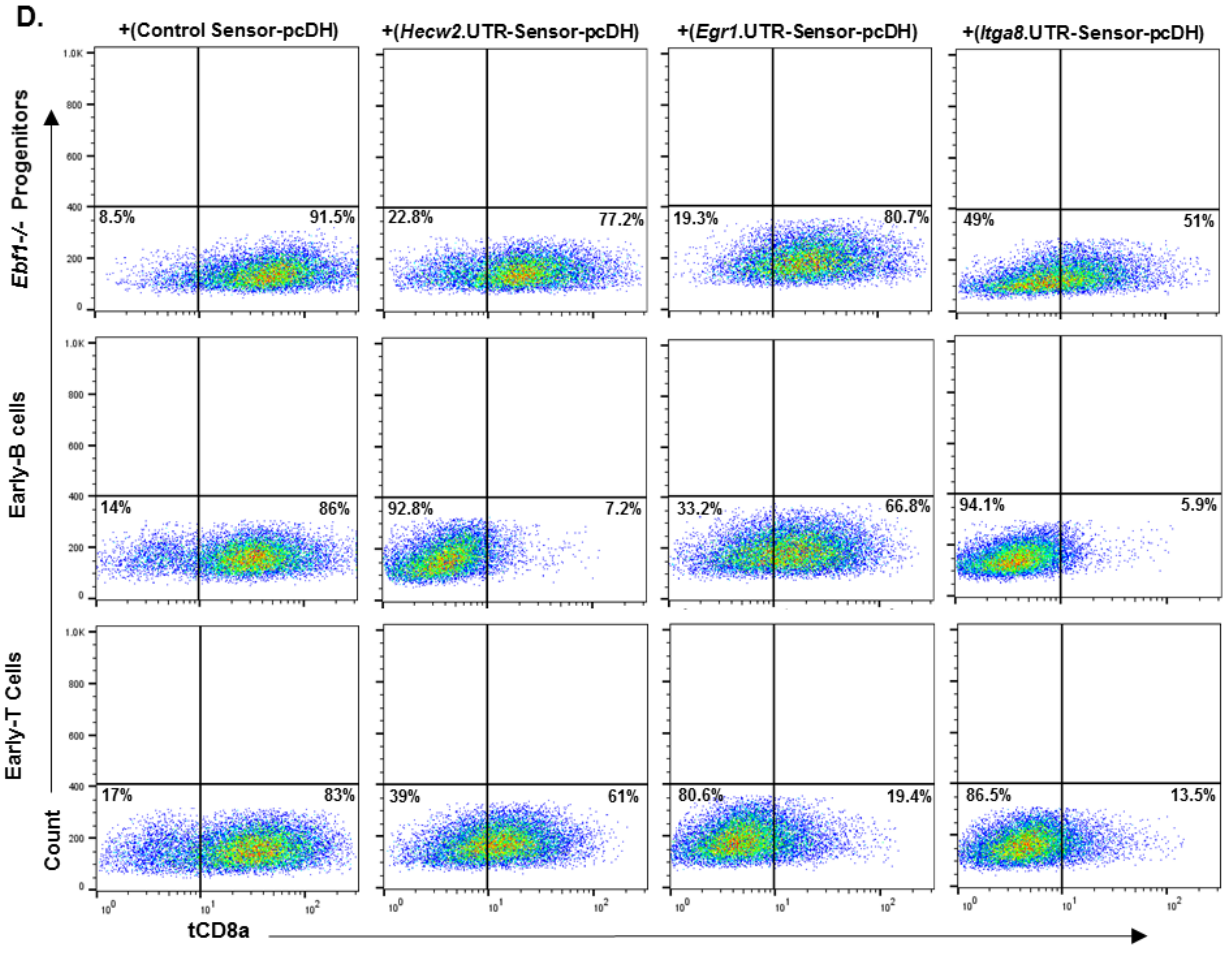
to undergo a decrease in tCD8a expression upon differentiation into Early-B as well as Early-T cells (tCD8a^{Hi}, 5.9% and 13.5%, respectively) (Fig. 3.4D). Consistent with the flow cytometry findings, qRT-PCR analysis also showed a corresponding decrease in the tCD8a mRNA levels (Fig. 3.4E).

Although most of the miRNAs which target *Hecw2* and *Egr1* transcripts were commonly-expressed in both Early-B and Early-T cells, we noticed a lineage-specific reduction in tCD8a reporter levels as a result of miRNA activity at their respective 3'UTRs. Thus, the 3'UTR reporter assays suggest that miRNA-mediated regulation is cell-type dependent and contingent not only upon the abundance of miRNAs and their targets, but may also involve other cell-intrinsic mechanisms.






C.
Hecw2 3'-UTR
(Chr 1(-): 53813010-53812240)
200 400 600 800 1000 (nt)
Egr1 3'-UTR
(Chr 18(+): 34861205-34862433)
0 200 400 600 800 1000 1200 (nt)
Itga8 3'-UTR
(Chr 2(-):12108918-12107488)
200 400 600 800 1000 1200 1400 1600 (nt)
miR-34b/449-5p
miR-130a/301a-3p
miR-204/211-5p
miR-142a-5p
miR-335-3p
miR-183-5p
miR-23-3p
miR-145a-5p
miR-191-5p
miR-181a-5p
miR-199a/b-3p
miR-342-3p
miR-335-5p
miR-199a-5p
miR-30a-5p
miR-25/92a/b-3p
miR-125a/125b/351-5p


D.
+(Control Sensor-pcDH)
+(Hecw2.UTR-Sensor-pcDH)
+(Egr1.UTR-Sensor-pcDH)
+(Itga8.UTR-Sensor-pcDH)
Ebf1-/- Progenitors
Early-B cells
Early-T Cells
Count
tCD8a
8.5% 91.5%
22.8% 77.2%
19.3% 80.7%
49% 51%
14% 86%
92.8% 7.2%
33.2% 66.8%
94.1% 5.9%
17% 83%
39% 61%
80.6% 19.4%
86.5% 13.5%

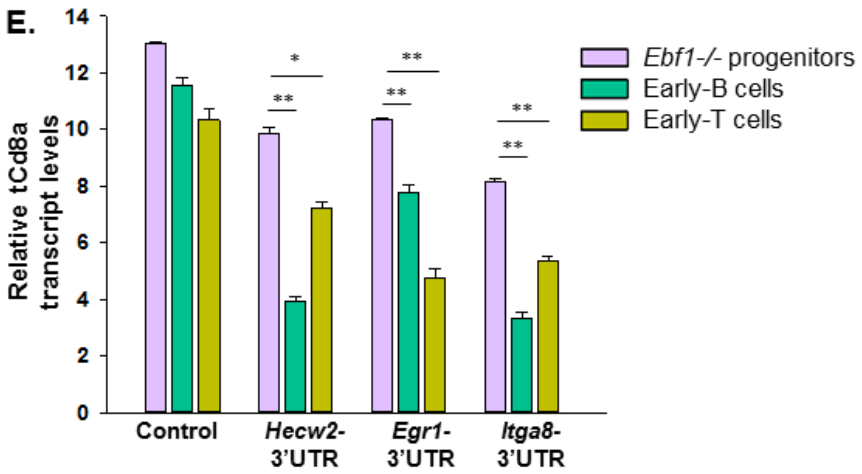


Figure 3.4. 3'UTR reporter assays showing miRNA-mediated target regulation is lineage-specific.

(A) Ago2 RIP-qPCR analysis to check lineage-specific enrichment of transcripts *Hecw2*, *Egr1* and *Itga8* in *Ebf1-/-* progenitors, Early-B and Early-T cells. For each transcript, the relative enrichment obtained after IP using anti-Ago2 or anti-IgG antibodies is represented.

(B) Design of the reporter vectors used for analysis of miRNA activity at 3'UTRs of Ago2-enriched transcripts. The vector 'Empty Sensor-pcDH' has CopGFP (as a marker for transfection) and tCd8a (as a reporter of miRNA activity), driven by two separate promoters (pEF1a and PGK, respectively) within two bidirectional cassettes (indicated by arrows). A control sequence, asEGFP (i.e. antisense sequence of EGFP) or 3'UTR region of a gene of interest was inserted after the tCd8a reporter to obtain Control Sensor-pcDH or 3'UTR-Sensor-pcDH, respectively. PA; SV 40 polyA tail.

(C) Binding sites predicted (by multiMiR) for miRNAs within the 3'UTR regions of selected Ago2-enriched transcripts - *Hecw2*, *Egr1* and *Itga8*, that were used in reporter assays to study lineage-specific miRNA-mediated repression.

(D) Flow cytometry analysis indicating surface levels of tCD8a reporter in *Ebf1-/-* progenitors, Early-B or Early-T cells, which expressed Control Sensor-pcDH or Sensor-pcDH vector containing 3'UTRs of *Hecw2*, *Egr1* or *Itga8*.

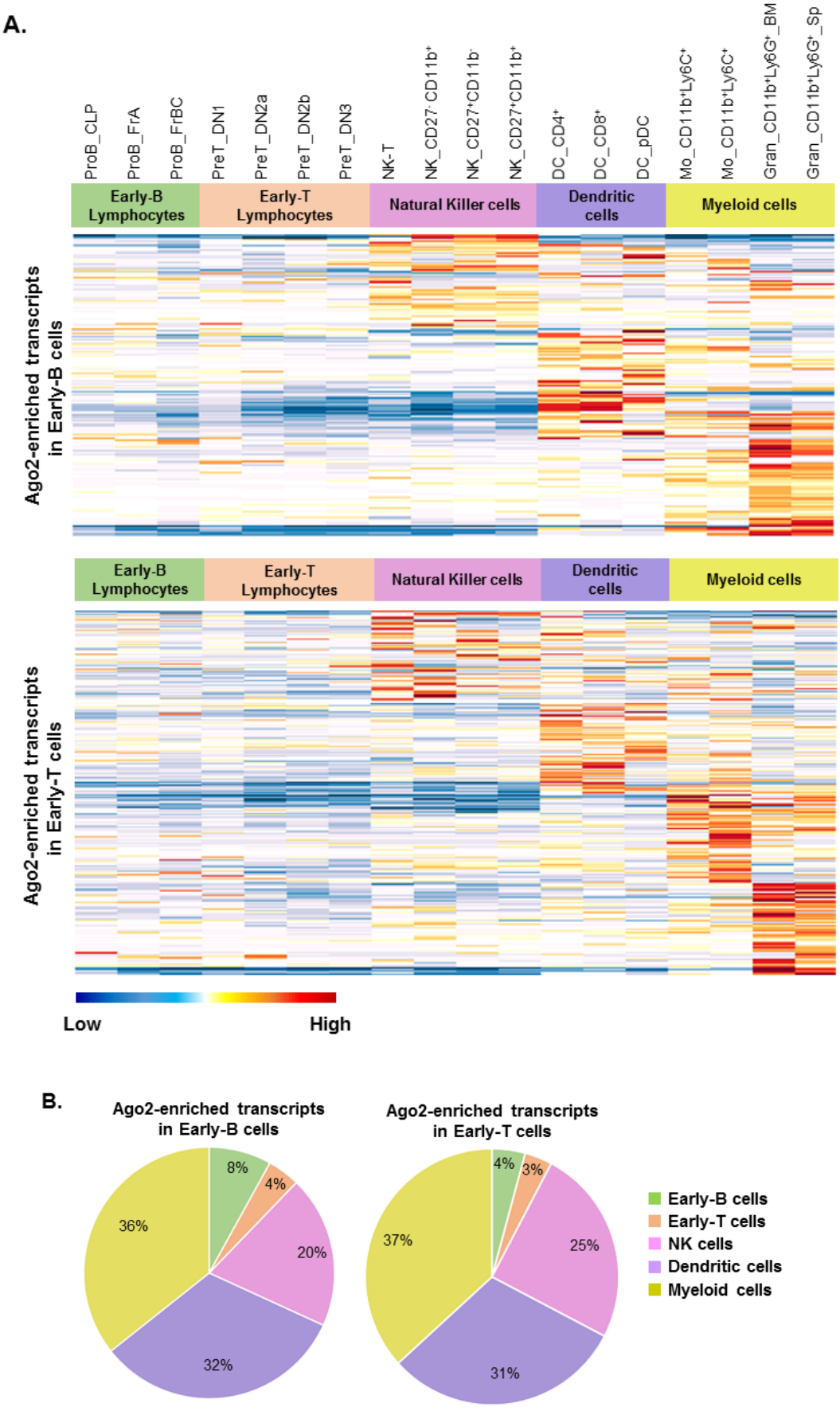
(E) Transcript levels of tCD8a reporter in *Ebf1-/-* progenitors, Early-B or Early-T cells that were transduced with Control Sensor-pcDH or Sensor-pcDH vector containing 3'UTRs of *Hecw2*, *Egr1* or *Itga8*. Data is shown as mean ± SD (*p ≤0.05; **p ≤0.01; ***p≤0.001).

3.5. miRNAs expressed in Early-B and Early-T cells target lineage-inappropriate genes associated with NK, DC and myeloid lineages

Since *Ebf1-/-* progenitors possess developmental plasticity towards other lineages besides their ability to readily differentiate into Early-B or Early-T cells; we speculated if the transcripts targeted by miRNAs in Early-B and Early-T cells could perhaps be associated with lineage-inappropriate genes. To explore this possibility, we attempted to ascertain the lineage identity of transcripts that were Ago2-enriched in Early-B or Early-T cells, by comparing their expression levels across various subsets of immune cell types or their developmental stages including Early-B lymphocytes, Early-T lymphocytes, Natural Killer cells (NK), Dendritic cells (DC), Monocytes and Neutrophils (Myeloid cells), using the RNAseq data available for these cell types from the Immunological Genome Project (ImmGen).

Strikingly, our analysis revealed that, the Ago2-enriched transcripts from Early-B or Early-T cells were expressed at comparatively higher levels in NK, DC and Myeloid lineages than B and T lymphocyte lineages (Fig. 3.5A). Of the total number of Ago2-enriched transcripts in Early-B cells, about 88% were mapped to alternate lineages i.e., 20% to Natural Killer cells (e.g., *Tbx21*, *Hic1*, *Nfatc2* and *Eomes*), 32% to Dendritic Cells (e.g., *Ciita*, *Mycl*, *H2-Aa* and *Prss30*) and 36% to Myeloid cells (e.g., *Rsad2*, *Fam126a*, *Ctnnd1* and *Fgr*), while the remaining 12% transcripts were mapped to early developmental stages of B and T lymphocytes (8% and 4%, respectively) (Fig. 3.5B, left panel and Supplementary Table-8A).

Similarly, in Early-T cells, 93% of the Ago2-enriched transcripts were mapped to the alternate lineages i.e., 25% to Natural Killer cells (e.g., *Il1rl1*, *Rapgef1*, *Zfp654* and *Kctd15*), 31% to Dendritic Cells (e.g., *Dkk3*, *Mast4*, *Nfat5* and *Klf8*) and 37% to Myeloid lineage (e.g., *Zfhx4*, *Klf13*, *Ctdspl* and *Tgfbr3*) (Fig. 3.5B, right panel and Supplementary Table-8B). Whereas, only a small percentage (7%) of these transcripts were mapped to early B and T lymphocytes (4% and 3%, respectively). It is noteworthy to mention that a handful of the Ago2-enriched transcripts from Early-B and Early-T cells were shown to be associated with more than one of the alternate lineages (e.g., *Shox2*, *Ptprg*, *Atf3* and *Trib1*).


A.
ProB_CLP
ProB_FrA
ProB_FrBC
PreT_DN1
PreT_DN2a
PreT_DN2b
PreT_DN3
NK-T
NK_CD27-CD11b+
NK_CD27+CD11b-
NK_CD27+CD11b+
DC_CD4+
DC_CD8+
DC_pDC
Mo_CD11b+Ly6C+
Mo_CD11b+Ly6C+
Gran_CD11b+Ly6G+_BM
Gran_CD11b+Ly6G+_Sp
Early-B Lymphocytes
Early-T Lymphocytes
Natural Killer cells
Dendritic cells
Myeloid cells
Ago2-enriched transcripts in Early-B cells
Ago2-enriched transcripts in Early-T cells
Low
High
B.
Ago2-enriched transcripts in Early-B cells
8%
4%
20%
32%
36%
Ago2-enriched transcripts in Early-T cells
4%
3%
25%
31%
37%
Early-B cells
Early-T cells
NK cells
Dendritic cells
Myeloid cells

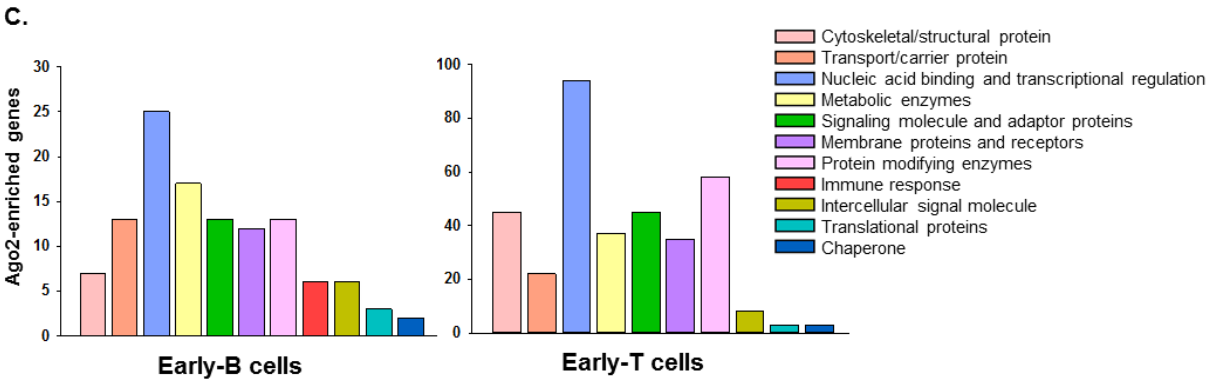


Figure 3.5. Lineage-association of Ago2-enriched transcripts identified using ImmGen.
(A) Heatmaps representing relative gene expression levels for Ago2-enriched transcripts from Early-B cells (top panel) or Early-T cells (bottom panel), analyzed in various subsets of early B and T lymphocytes, Natural Killer cells, Dendritic cells and Myeloid cells (monocytes and neutrophils), using the RNA-Seq data available at ImmGen.
(B) Pie chart showing percentages of Ago2-enriched transcripts from Early-B cells (top panel) or Early-T cells (bottom panel) that were mapped to various immune cell types using the RNA-Seq data available at ImmGen.
(C) Gene Ontology analysis of Ago2-enriched transcripts from Early-B and Early-T cells. The transcripts were classified into various categories based on their molecular function, using the annotation tool 'Panther'.

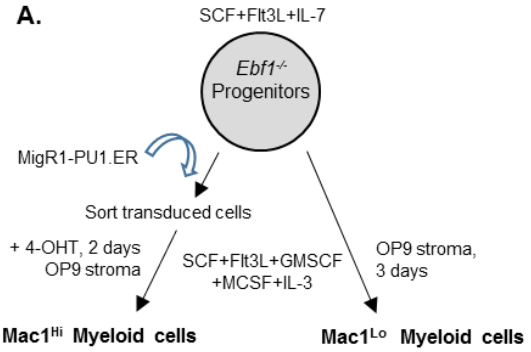
In order to assess the functional role of these transcripts, we performed Gene Ontology (GO) analysis and identified that majority of the Ago2-enriched transcripts in Early-B or Early-T cells encoded proteins associated with diverse biological functions such as signal transducers (e.g., *Ptprg* and *Map3k2*), surface receptors (e.g. *Tlr13*, *Tlr9*, and *Ciita*), metabolic regulators (e.g., *Ltf* and *Nos1*), cytoskeletal proteins (e.g., *Itga8*) as well as regulators of transcription (e.g., TFs encoded by *Tbx21*, *Eomes*, *Irf8*, *Tgfbr3*, *Nfatc2* or RNA-binding proteins encoded by *Cpeb3*) (Fig. 3.5C). Thus, despite the mutually exclusive gene regulatory networks that orchestrate B or T cell fate program, the presence of an overlapping set of miRNAs that target distinct sets of lineage-inappropriate genes, suggests that both B and T cells employ a common mechanism of post-transcriptional silencing to restrict the available alternate cell-fate options before they undergo commitment towards the respective lineages (Ikawa et al., 1999; Izon et al., 2001; Manz et al., 2001; Rumfelt et al., 2006; Wada et al., 2008; Rothenberg 2011; Zandi et al., 2012).

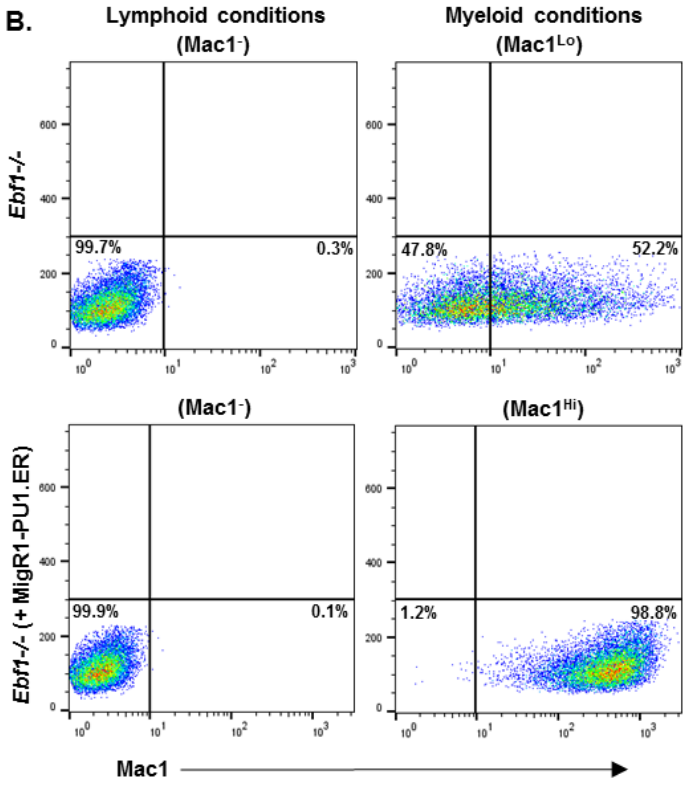
3.6. miRNA-targeted transcripts in Early-B or Early-T cells are upregulated during myeloid lineage differentiation

Of the Ago2-enriched transcripts in Early-B and Early-T cells, a high percentage (~37%) of them were predicted to be closely associated with myeloid lineage, as per the ImmGen analysis (Fig. 3.5B), raising the possibility that these genes may be necessary for the myeloid development. In order to verify the alternate lineage identity of Ago2-enriched transcripts from Early-B and Early-T cells, we analyzed their expression levels following the differentiation of *Ebf1-/-* progenitors into myeloid cells, *in vitro.* It has been previously shown that *Ebf1-/-* progenitors (which are maintained under lymphoid conditions) differentiate into a heterogeneous population of $Mac1^{+}$ ($CD11b^{+}$) myeloid cells, when transferred to myeloid-promoting conditions (Pongubala et al, 2008). However, to achieve a robust and synchronized differentiation of *Ebf1-/-* progenitors towards myeloid lineage, we expressed the transcription factor PU.1 fused to ligand binding domain of estrogen receptor using a retroviral construct, MigR1-PU1.ER (Fig. 3.6.1A). Interestingly, when PU1.ER-expressing *Ebf1-/-* progenitors (4-OHT untreated and maintained under lymphoid conditions) were transferred to myeloid-promoting conditions and treated with 4-OHT, they readily differentiated into myeloid cells expressing high levels of Mac1 (shown as $Mac1^{Hi}$, 98.8%), in comparison to myeloid cells generated by *Ebf1-/-* progenitors (in the absence of MigR1-PU1.ER) which exhibit heterogeneous Mac1 expression (shown as $Mac1^{Lo}$, 52.2%) (Fig. 3.6.1B). Correspondingly, the transcript levels of PU.1 and its well-known targets like *Csf1r* and *Cd11b* (Mac1) were found to be comparatively higher in $Mac1^{Hi}$ cells than compared to $Mac1^{lo}$ population, indicating a PU.1 dosage-dependent effect (Fig. 3.6.1C).

Subsequently, we used the $Mac1^{Lo}$ and $Mac1^{Hi}$ myeloid cells to analyze the expression levels of few randomly selected transcripts that were Ago2-enriched in Early-B or Early-T cells, and possibly associated with myeloid lineage (Supplementary Table-8A and 8B). These included transcripts that were Ago2-enriched in Early-B cells (e.g., *Fam126a, Fgr, Atg4d, Ctnnd1, Gsn and Nfam1)* (Fig. 3.6.2A) or in Early-T cells (e.g., *Maf, Ptprg, Ctdspl, Tgfbr3, Egr1* and *Zfhx4)* (Fig. 3.6.2B) or in both cell types (e.g., *Itga8* and *Il10ra)* (Fig. 3.6.2C). Interestingly, the expression levels of *Ctnnd1*, *Maf, Ptprg, Ctdspl, Tgfbr3, Gsn, Egr1, Zfhx4* and

Itga8 were upregulated in Mac1^{Lo} and more so in Mac1^{Hi} myeloid cells, compared to their levels in *Ebf1-/-* progenitors. However, we found only a modest induction of transcripts encoding the genes *Fam126a, Fgr, Atg4d, Nfam1* and *Il10ra,* in the Mac^{Lo} and Mac1^{Hi} myeloid cells. These findings confirm that the above mentioned Ago2-enriched transcripts are indeed associated with myeloid lineage; thus supporting the hypothesis that miRNAs expressed in Early-B and Early-T cells play a potential role in restricting the lineage-inappropriate genes during B and T cell-fate commitment.





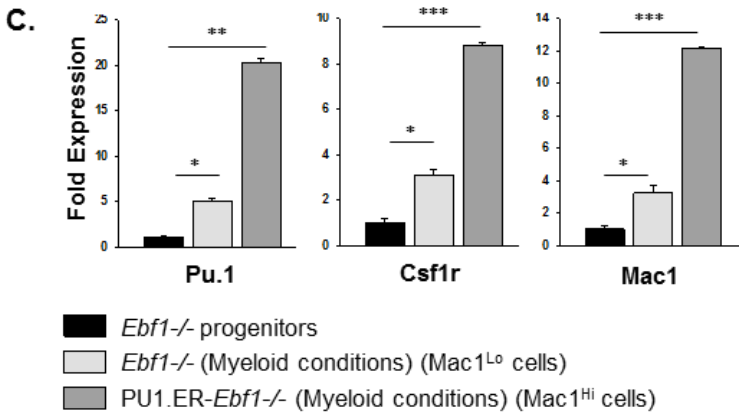


Figure 3.6.1. *In vitro* myeloid differentiation of *Ebf1-/-* progenitors.
(A) Experimental design and culture conditions used for *in vitro* differentiation of *Ebf1-/-* progenitors into myeloid ($Mac1^{Lo}$ or $Mac1^{Hi}$) cells. *Ebf1-/-* progenitors that were maintained under lymphoid-promoting cytokines (SCF+Flt3L+IL-7) were transferred to myeloid-promoting cytokines (SCF+Flt3L+GMSCF+MCSF+IL-3) and cultured for 3 days to obtain $Mac1^{Lo}$ myeloid cells. Alternatively, *Ebf1-/-* progenitors were first transduced with MigR1-PU1.ER construct and propagated under lymphoid-promoting cytokines in absence of 4-OHT, then subsequently transferred to myeloid-promoting cytokines and treated with 4-OHT (at 1uM final concentration) for 2 days, to generate $Mac1^{Hi}$ myeloid cells.
(B) Flow cytometry analysis showing expression levels of myeloid marker, Mac1 (or CD11b) on $Mac1^{Lo}$ or $Mac1^{Hi}$ myeloid cells generated by *Ebf1-/-* and PU1.ER-*Ebf1-/-* progenitors, respectively.
(C) qRT-PCR analysis for comparison of expression levels of known myeloid-associated genes, *Pu.1*, *Csf1r* and *Mac1* (*Cd11b*) in *Ebf1-/-* progenitors, $Mac1^{Lo}$ or $Mac1^{Hi}$ myeloid cells.

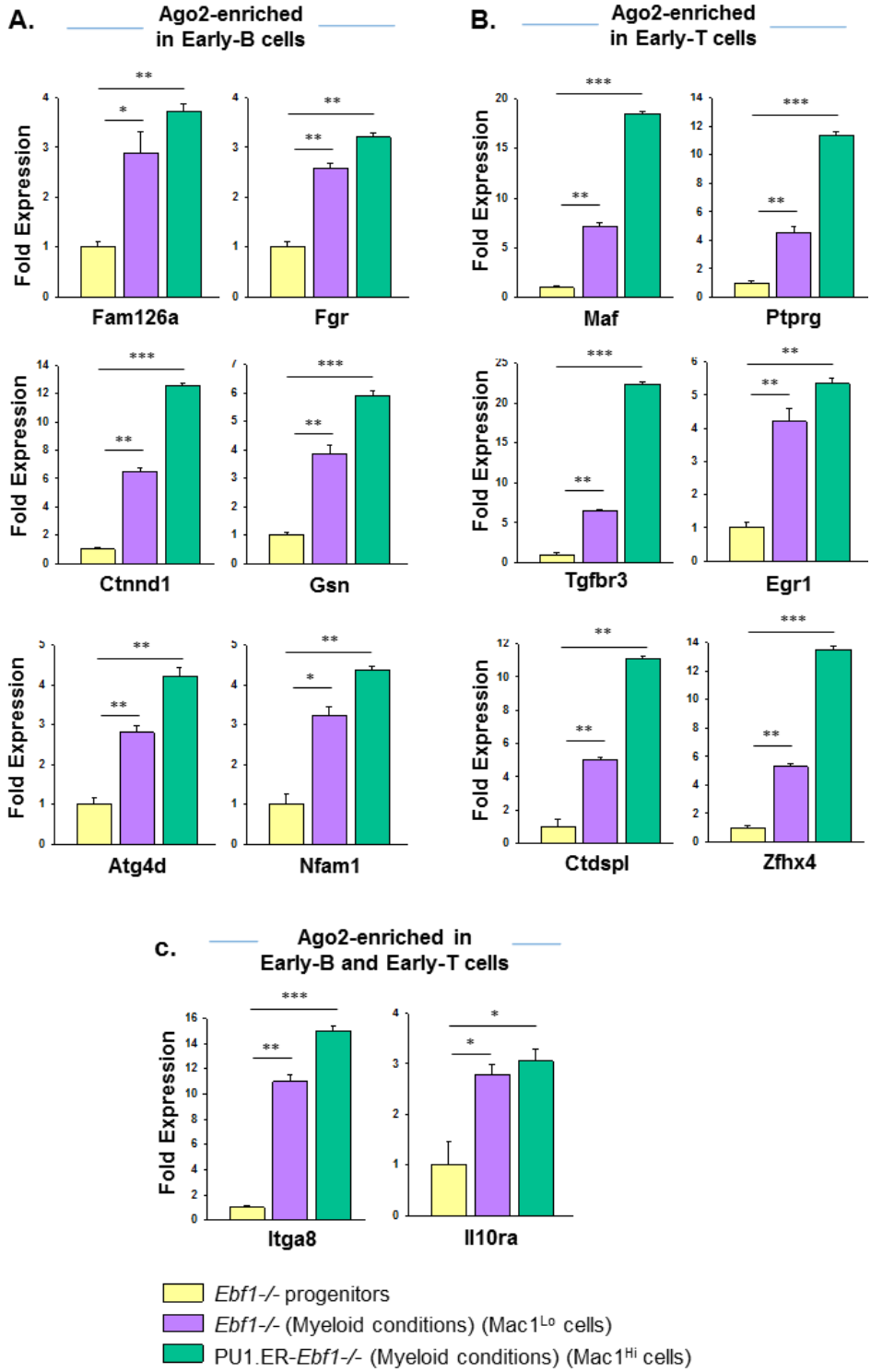


Figure 3.6.2. Expression levels of Ago2-enriched transcripts from Early-B or Early-T cells analyzed by qRT-PCR in myeloid cells.

Relative expression levels of transcripts that were that were suggested to be associated with myeloid-lineage (from ImmGen analysis) and were Ago2-enriched in Early-B cells (A), in Early-T cells (B), or in both these lineages (C), measured upon differentiation of *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors into myeloid cells ($Mac1^{Lo}$ or $MacI^{Hi}$, respectively). Data is shown as mean ± SD (*p ≤0.05; **p ≤0.01; ***p≤0.001).

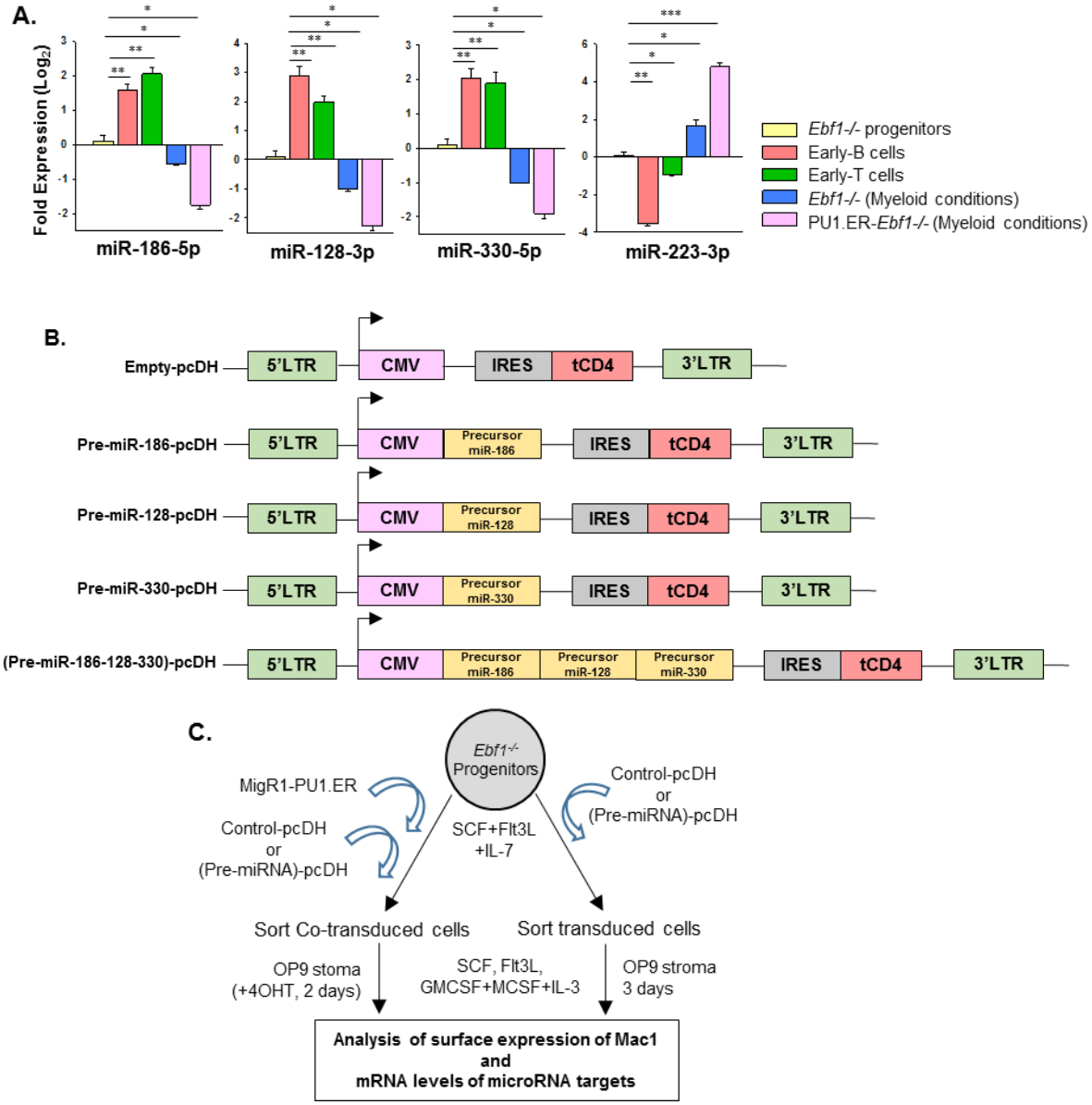
3.7. Lymphocyte miRNAs act combinatorially to suppress myeloid lineage option of multipotent progenitors

Given the finding that transcripts targeted by miRNAs in Early-B and Early-T cells are associated with alternate lineage, we attempted to interrogate the functional significance of miRNA-mediated repression of myeloid lineage genes. Hence, we tested the ability of lymphocyte miRNAs to attenuate the myeloid cell-fate option of the multipotent *Ebf1-/-* progenitors. In order to select the most potential miRNA candidates, we screened all the upregulated miRNAs from Early-B and Early-T cells on the basis of their abundance, total number of myeloid-associated targets and lymphocyte-specific expression pattern. Based on these criteria, we shortlisted three miRNAs: miR-186-5p, miR-128-3p and miR-330-5p which were commonly-upregulated in Early-B and Early-T cells but downregulated in myeloid cells (Mac1^{Lo} and Mac1^{Hi} populations), compared to *Ebf1-/-* progenitors (the myeloid-specific miRNA, miR-223-3p, was used as a positive control) (Fig. 3.7.1A). Subsequently, we chose to functionally test each of these three miRNAs for their ability to inhibit myeloid differentiation potential of *Ebf1-/-* progenitors.

To determine if forced expression of any of the selected miRNAs: miR-186-5p, miR-128-3p or miR-330-5p would impair the myeloid development, we expressed their individual precursor sequences (plus ~100nt flanking on its either sides), into *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors using the pcDH vector (Fig. 3.7.1B). Subsequently, the transduced *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors were allowed to differentiate towards myeloid lineage and their myeloid differentiation potential was determined by analyzing Mac1 expression (Fig. 3.7.1C). The levels of mature miRNAs (miR-186-5p, miR-128-3p or miR-330-5p) were found to be ~5-fold higher in cells transduced with the respective Pre-miRNA-pcDH vectors compared to cells transduced with Empty-pcDH (control vector), indicating effective processing of the precursor transcripts (Fig. 3.7.1D).

We found that expression of Pre-miR-186, Pre-miR-128 or Pre-miR-330 in *Ebf1-/-* progenitors resulted in a mild decrease in the generation of heterogeneous Mac1^{+} myeloid cells (46.5%, 46% or 46.2%, respectively) compared to *Ebf1-/-* progenitors that expressed Empty-pcDH (Mac1^{+}, 52.2%) (Fig. 3.7.2A, left panels). On the other

hand, PU1.ER-*Ebf1-/-* cells that were transduced with Pre-miR-186-pcDH exhibited a notable decrease in generation of Mac1^{Hi} myeloid cells (Mac1^{Hi}, 79.2%), leading to a corresponding increase in the accumulation of Mac1^{Lo} population (Mac1^{Lo}, 19.5%) in comparison to the Mac1^{Lo} population in PU1.ER-*Ebf1-/-* cells that expressed Empty-pcDH vector (Mac1^{Lo}, 7.1%) (Fig. 3.7.2A, right panels). Similarly, transduction of Pre-miR-128-pcDH and Pre-miR-330-pcDH in PU1.ER-*Ebf1-/-* cells also resulted in an increased accumulation of Mac1^{Lo} population (14.1% and 16.9%, respectively). Thus, expression of each of the above lymphocyte miRNAs individually, brings about a discernable, yet modest decrease in myeloid differentiation capacity of *Ebf1-/-* and PU1.ER-*Ebf1-/-* progenitors.



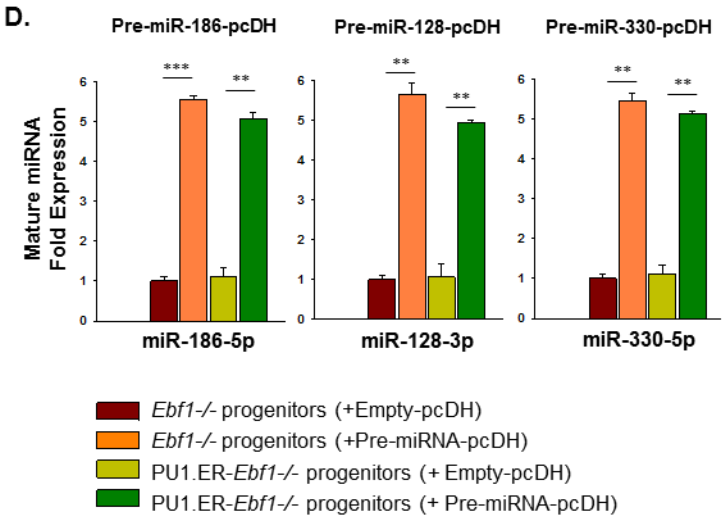
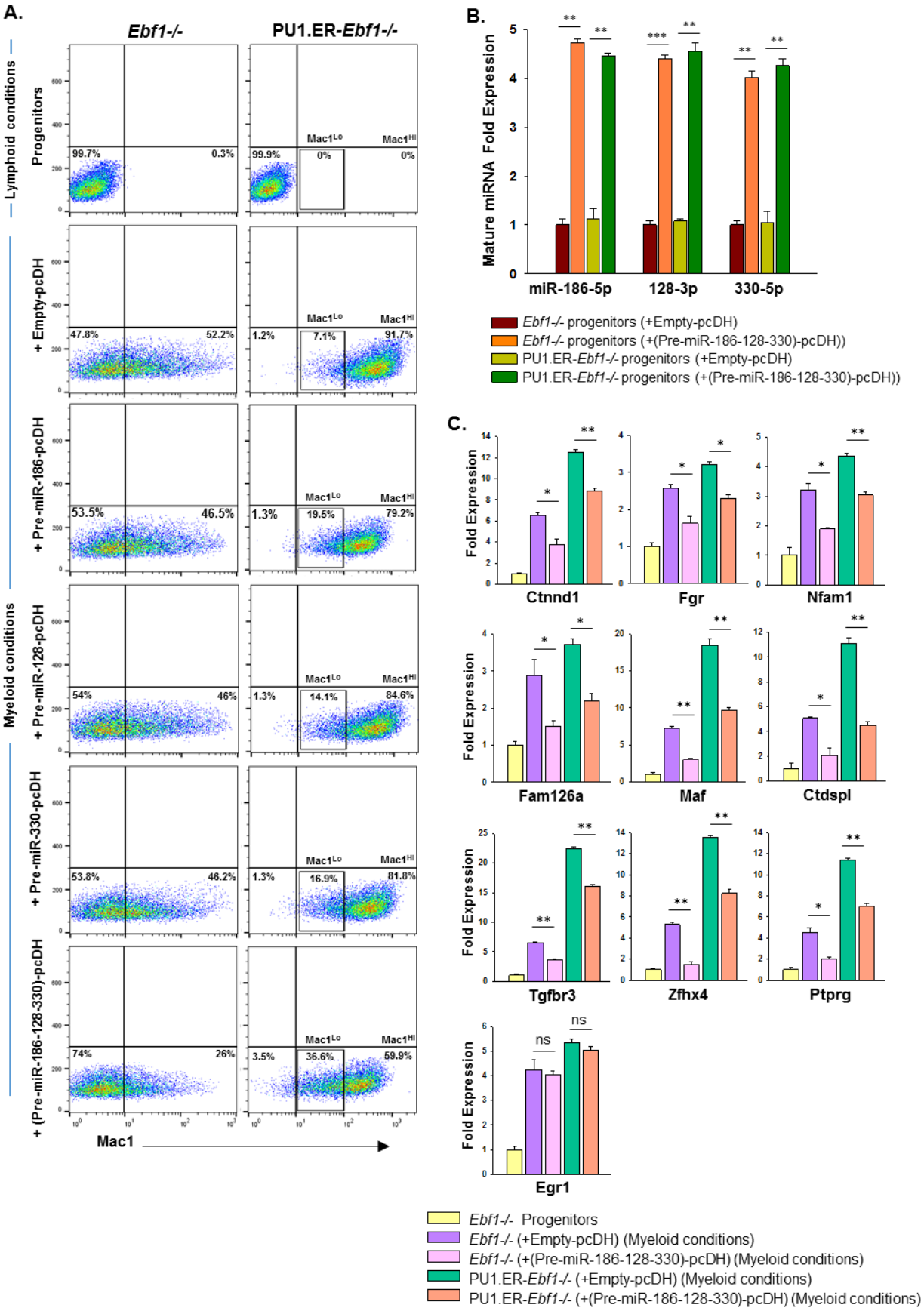


Figure 3.7.1. Experimental strategy for ectopic expression of lymphocyte miRNAs.
(A) Expression levels of miR-186-5p, miR-128-3p, miR-330-5p and miR-223-3p (myeloid-specific miRNA), measured using stemloop qRT-PCR in *Ebf1-/-* progenitors, Early-B cells, Early-T cells, Mac1^{Lo} and Mac1^{Hi} myeloid cells (generated from *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors, respectively, by culturing under myeloid-promoting conditions).
(B) Design of pcDH vector constructs for ectopic expression of miRNAs.
(C) Experimental strategy for studying the effect of ectopically expressed miRNAs on myeloid differentiation potential of *Ebf1-/-* progenitors.
(D) Levels of mature or effector miRNAs: miR-186-5p, miR-128-3p and miR-330-5p, in *Ebf1-/-* or PU1.ER-*Ebf1-/-* cells that were transduced with Empty-pcDH vector or pcDH vectors expressing precursor sequences of each of these miRNAs.

However, recent studies indicate that the efficiency as well as outcome of miRNA-mediated regulation of developmental processes is defined not only by the expression levels, binding strength or activity of individual miRNAs, but involves a combinatorial effect exhibited by multiple miRNAs that co-regulate functionally related transcripts (Pons-Espinal et al., 2017; Cursons et al., 2018; Bhaskaran et al., 2019). We may note that each of the three miRNAs: miR-186-5p, miR-128-3p and miR-330-5p, were predicted to target an average of 67 myeloid-associated transcripts, while in combination, they potentially target at least 148 myeloid-associated transcripts, together in Early-B and Early-T cells (Supplementary Table-9). Taking this into consideration, we analyzed the effect of simultaneous expression of these three miRNAs on myeloid differentiation capacity of *Ebf1-/-* and PU1.ER-*Ebf1-/-* progenitors. Strikingly, co-expression of Pre-miR-186, Pre-miR-128 and Pre-miR-330 in *Ebf1-/-* progenitors (via (Pre-miR-186-128-330)-pcDH construct resulted in a profound decrease in generation of heterogeneous

Mac1^{+} myeloid population (Mac1^{+}, 26%) (Fig. 3.7.2A, bottom-left panel) compared to the effect seen when each of these miRNAs were expressed individually (Mac1^{+}, ~46%). Similarly, PU1.ER-*Ebf1-/-* cells that co-expressed Pre-miR-186, Pre-miR-128 and Pre-miR-330 exhibited a substantial decrease in generation of Mac1^{Hi} populations (Mac1^{Hi}, 59.9%), leading to an increased accumulation of Mac1^{Lo} population (Mac1^{Lo}, 36.6%) (Fig. 3.7.2A, bottom-right panel) compared to PU1.ER-*Ebf1-/-* cells that expressed these miRNAs individually (Mac1^{Lo}, <20%). Although the corresponding effector miRNAs (miR-186-5p, miR-128-3p and miR-330-5p) were generated at similar levels in *Ebf1-/-* and PU1.ER-*Ebf1-/-* cells (Fig. 3.7.2B), we noted the myeloid differentiation capacity was abrogated to a lesser extent in PU1.ER-*Ebf1-/-* cells than *Ebf1-/-* cells (Mac1^{-} population, 3.5% vs 74%, respectively). This could perhaps be due to the limited ability of miRNAs to antagonize the robust myeloid developmental program induced in PU1.ER-*Ebf1-/-* cells, due to higher PU.1 levels.

Previously published studies suggest that many of the genes predicted to be targeted by miR-186-5p, miR-128-3p and miR-330-5p, were shown to play a role in development and survival of myeloid cells (Zheng et al., 2011; Wallner et al, 2014; Willman et al., 1991; Hegde et al., 1998; Hedge et al., 1999; Gemelli et al., 2008; Ohtsuka et al., 2004; Yasugi et. al., 2006; Peruta et al., 2010; Fieldler et al., 2013). So, we sought to determine if the observed attenuation in myeloid differentiation capacity is due to the repression of the important myeloid-associated genes that were predicted to be targeted by one or more of these miRNAs. As shown in Fig. 3.7.2C, the transcript levels of myeloid-associated genes like *Ctnnd1*, *Fgr*, *Nfam1*, *Fam126a*, *Maf*, *Tgfbr3*, *Ctdspl*, *Ptprg* and *Zfhx4* were found to be relatively lower in myeloid-differentiating *Ebf1-/-* and PU1.ER-*Ebf1-/-* cells that expressed (Pre-miR-186-128-330)-pcDH, compared to cells which expressed Empty-pcDH. Furthermore, transcripts that were targeted by two of these miRNAs (e.g., *Zfhx4*, *Maf*, *Ctdspl*, *Ptprg* and *Fam126a*) were repressed more profoundly than transcripts targeted by only one miRNA (e.g. *Ctnnd1*, *Fgr1*, *Nfam1* and *Tgfbr3*), indicating an additive effect (Fig. 3.7.2C). Importantly, we found no significant decrease in the transcript levels of the myeloid gene, *Egr1*, which is not a predicted target for any these miRNAs; thus indicating specificity of miRNA-mediated regulation.


A.
Ebf1-/-
PU1.ER-Ebf1-/-
Lymphoid conditions
Myeloid conditions
Progenitors
+ Empty-pcDH
+ Pre-miR-186-pcDH
+ Pre-miR-128-pcDH
+ Pre-miR-330-pcDH
+ (Pre-miR-186-128-330)-pcDH
Mac1Lo
Mac1Hi
99.7%
0.3%
99.9%
0%
0%
47.8%
52.2%
1.2%
7.1%
91.7%
53.5%
46.5%
1.3%
19.5%
79.2%
54%
46%
1.3%
14.1%
84.6%
53.8%
46.2%
1.3%
16.9%
81.8%
74%
26%
3.5%
36.6%
59.9%
Mac1
B.
Mature miRNA Fold Expression
miR-186-5p
128-3p
330-5p
Ebf1-/- progenitors (+Empty-pcDH)
Ebf1-/- progenitors (+(Pre-miR-186-128-330)-pcDH))
PU1.ER-Ebf1-/- progenitors (+Empty-pcDH)
PU1.ER-Ebf1-/- progenitors (+(Pre-miR-186-128-330)-pcDH))
C.
Fold Expression
Ctnnd1
Fgr
Nfam1
Fam126a
Maf
Ctdspl
Tgfbr3
Zfhx4
Ptprg
Egr1
ns
Ebf1-/- Progenitors
Ebf1-/- (+Empty-pcDH) (Myeloid conditions)
Ebf1-/- (+(Pre-miR-186-128-330)-pcDH) (Myeloid conditions)
PU1.ER-Ebf1-/- (+Empty-pcDH) (Myeloid conditions)
PU1.ER-Ebf1-/- (+(Pre-miR-186-128-330)-pcDH) (Myeloid conditions)

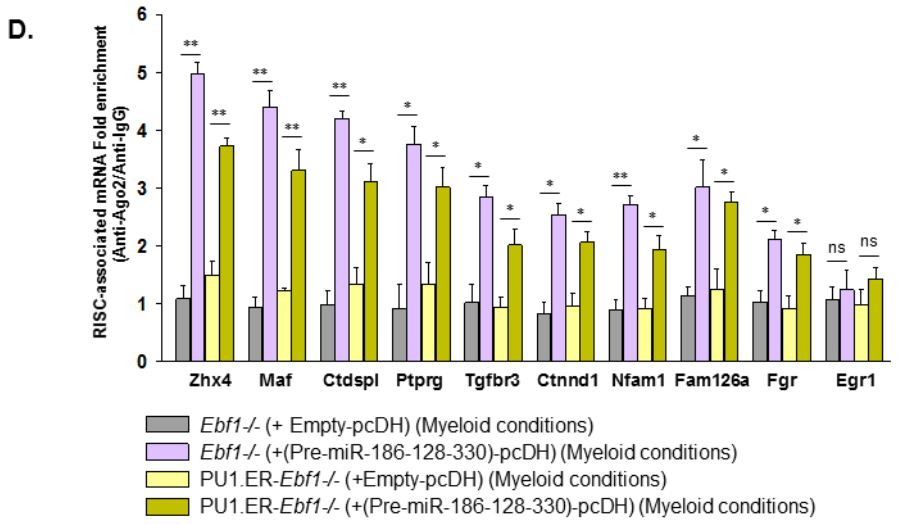


Figure 3.7.2. Effect of co-expression of multiple miRNAs with distinct seed sequences on myeloid differentiation potential of *Ebf1-/-* and PU1.ER-*Ebf1-/-* progenitors.

(A) Analysis of myeloid differentiation potential of *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors that were transduced with Empty-pcDH or pcDH vector expressing Pre-miR-186, Pre-miR-128 and Pre-miR-330, either individually or simultaneously. Transduced cells were cultured under myeloid-promoting conditions (see methods) and analyzed for Mac1 (CD11b) expression by flow cytometry.

(B) Levels of mature or effector miRNAs: miR-186-5p, miR-128-3p and miR-330-5p, in *Ebf1-/-* or PU1.ER-*Ebf1-/-* cells that were transduced with pcDH vectors co-expressing precursor sequences of each of these miRNAs.

(C) qRT-PCR analysis of transcript levels of myeloid–associated genes targeted by miRNAs – miR-186-5p, miR-128-3p or miR-330-5p, measured in *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors that were transduced with Empty-pcDH or (Pre-miR-186-128-330)-pcDH vector and subsequently differentiated towards myeloid lineage. Data is represented with *Ebf1-/-* progenitors as control sample.

(D) RIP-qPCR analysis for measurement of mRNA Fold enrichment (Ago2-IP vs IgG-IP) for myeloid-associated transcripts targeted by the above miRNAs in RISC-complexes of *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors that were transduced with Empty-pcDH or (Pre-miR-186-128-330)-pcDH vector and subsequently differentiated towards myeloid lineage.

(Data is shown as mean ± SD (*p ≤0.05; **p ≤0.01; ***p≤0.001; ns=not significant).

Consistent with these observations, RIP-qPCR analyses with Anti-Ago2 or Anti-IgG antibodies show an increased Ago2-enrichment of the above transcripts in RISC-complexes of myeloid-differentiating *Ebf1-/-* and PU1.ER-*Ebf1-/-* cells in presence of (Pre-miR-186-128-330)-pcDH, compared to Empty-pcDH (Fig. 3.7.2D). Moreover, the level of enrichment was found to positively correlate with the number of binding sites for these miRNAs in the 3'UTRs of transcripts i.e., genes like *Zfhx4*, *Maf*, *Ctdspl*, *Ptprg* and *Fam126a* that possess binding sites for

two of these miRNAs showed comparatively higher Ago2-enrichment than other transcripts having sites for only one miRNA (e.g. *Ctnnd1*, *Fgr1*, *Nfam1* and *Tgfbr3*). Taken together, these analyses provide an evidence that co-expression of lymphocyte-specific miRNAs: miR-186-5p, miR-128-3p and miR-330-5p, causes simultaneous repression of a large repertoire of myeloid-associated transcripts by recruiting them to cellular RISC complexes. Since many of these transcripts are important for myeloid differentiation and survival, it is quite likely that a cumulative decrease in their cellular levels resulted in attenuation of myeloid differentiation capacity of multipotent progenitors.

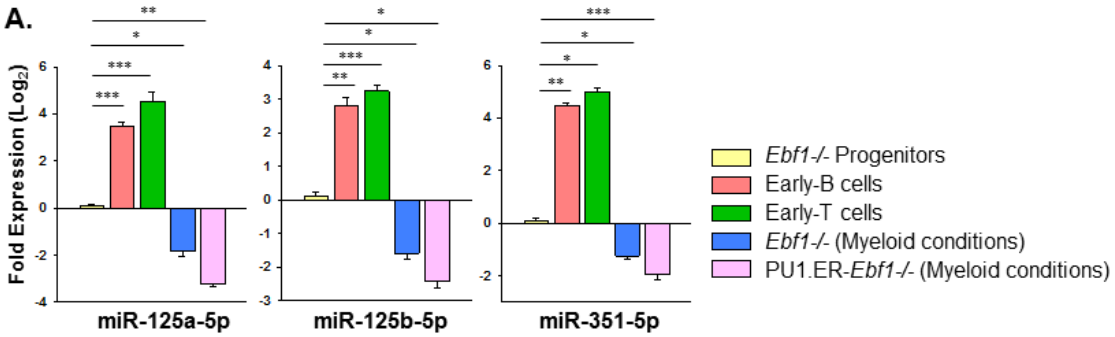
3.8. Co-expressed miRNAs having similar seed sequences are inefficient in attenuating myeloid differentiation potential

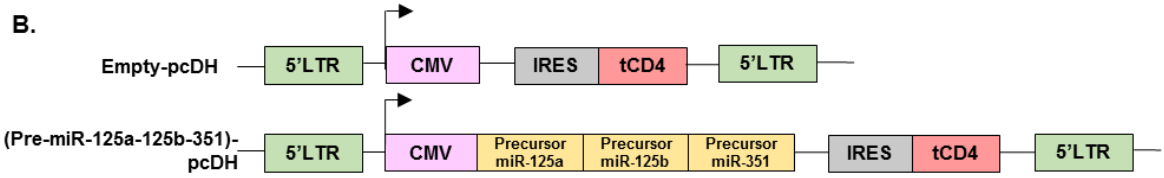
To further confirm if the observed decrease in myeloid differentiation potential in presence of miR-186-5p, miR-128-3p and miR-330-5p is a result of combinatorial and synergistic targeting of several myeloid-associated genes, but not merely due to miRNA overexpression; we tested the myeloid repression potential of a second set of miRNAs: miR-125a-5p, miR-125b-5p and miR-351-5p, which belong to miR-125 family. All the three members of this family were found to exhibit lymphocyte-specific expression pattern and shared overlapping set of targets due to similarity in their seed sequences (Fig. 3.8A). Collectively, they were predicted to target at least 62 myeloid-associated transcripts in Early-B and Early-T cells, which were also found to be important for myeloid differentiation and survival (Supplementary Table-10 and Ley et al., 2016; Jin et al., 2018; Ghosh et al., 2016; Hegde et al., 1998; Hedge et al., 1999; Gemelli et al., 2008; Zheng et al., 2011).

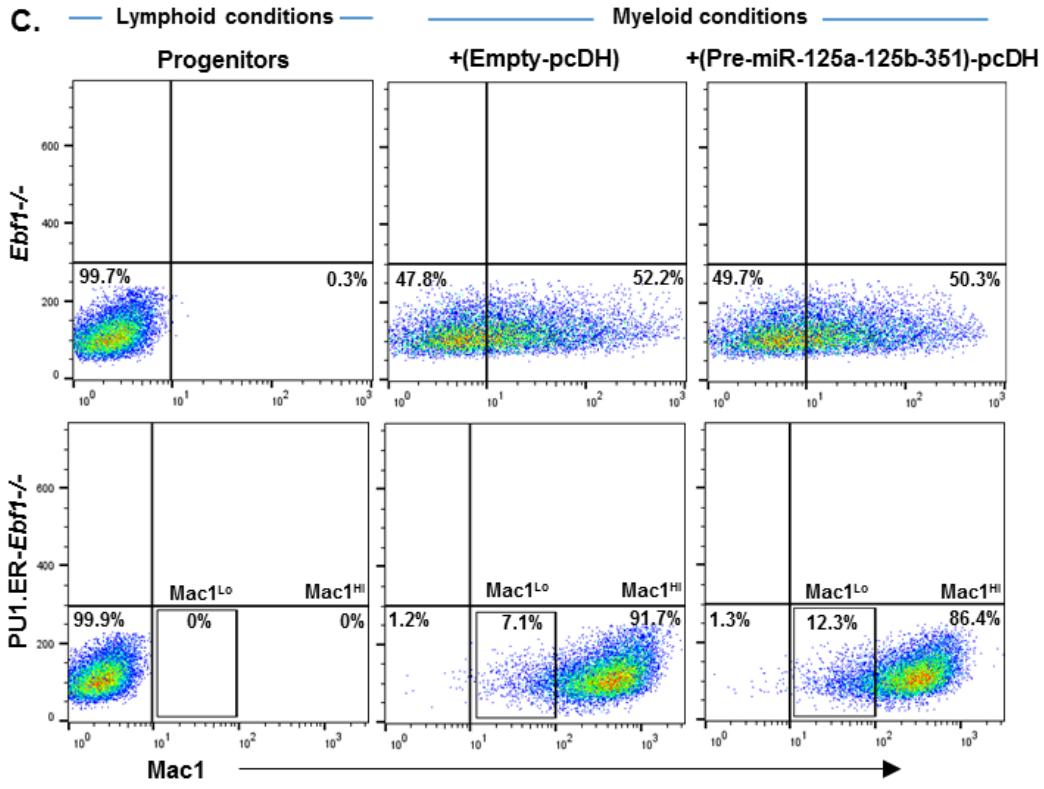
Forced co-expression of miR-125 family members in myeloid-differentiating *Ebf1-/-* or PU1.ER-*Ebf1-/-* cells via the construct (Pre-miR-125a-125b-351)-pcDH (Fig. 3.8B), resulted in a significant downregulation of their target genes (including *Gsn*, *Atg4d*, *Maf*, *Ctdspl*, *Itga8* and *Il10ra*) (Fig. 3.8D); but caused only a subtle decrease in myeloid differentiation potential of these cells (Fig. 3.8C). The inefficiency of myeloid lineage restriction by miR-125 family members does not seem to be due to a weak interaction with their targets, because the transcripts targeted by these miRNAs were found to be effectively enriched in the RISC-complexes of myeloid-differentiating *Ebf1-/-* or PU1.ER-*Ebf1-/-* cells that co-expressed the members of miR-125 family (Fig. 3.8E).

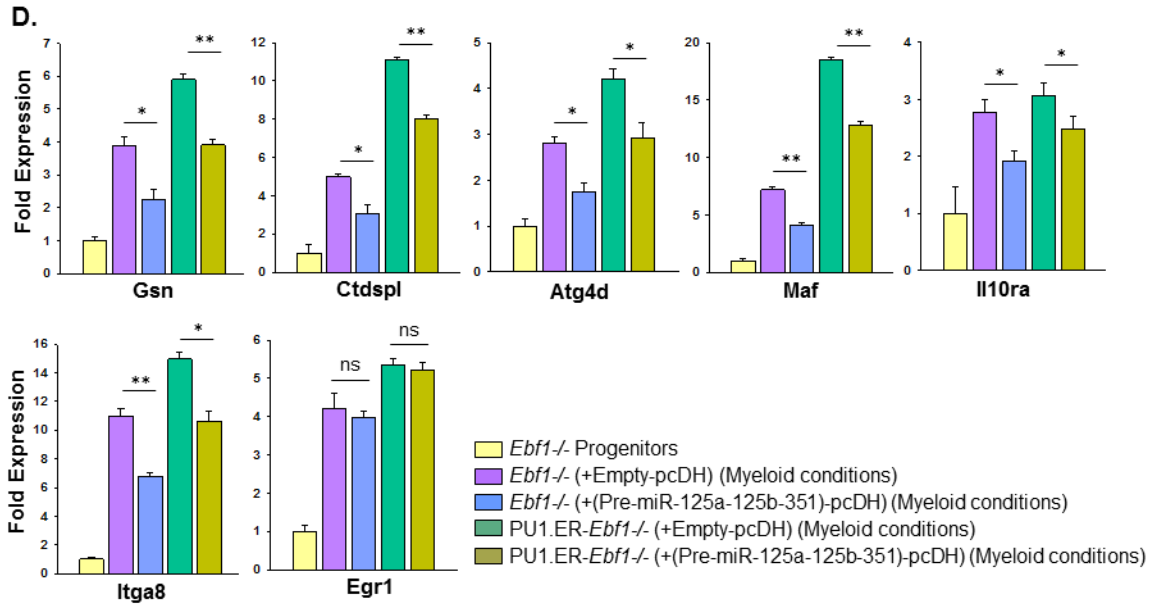
Moreover, the levels of effector miRNAs (i.e. miR-125a-5p, miR-125b-5p and miR-351-5p) in cells transduced with (Pre-miR-125a-125b-351)-pcDH were similar to that observed for miR-186-5p, miR-128-3p and miR-330-5p upon transduction of (Pre-miR-186-128-330)-pcDH; indicating the overall effector miRNA levels were comparable in both cases (Fig. 3.8F and 3.7.2B). These studies suggest, although co-expression of multiple lymphocyte miRNAs that share similar seed sequences could significantly lower the abundance of their respective myeloid-associated

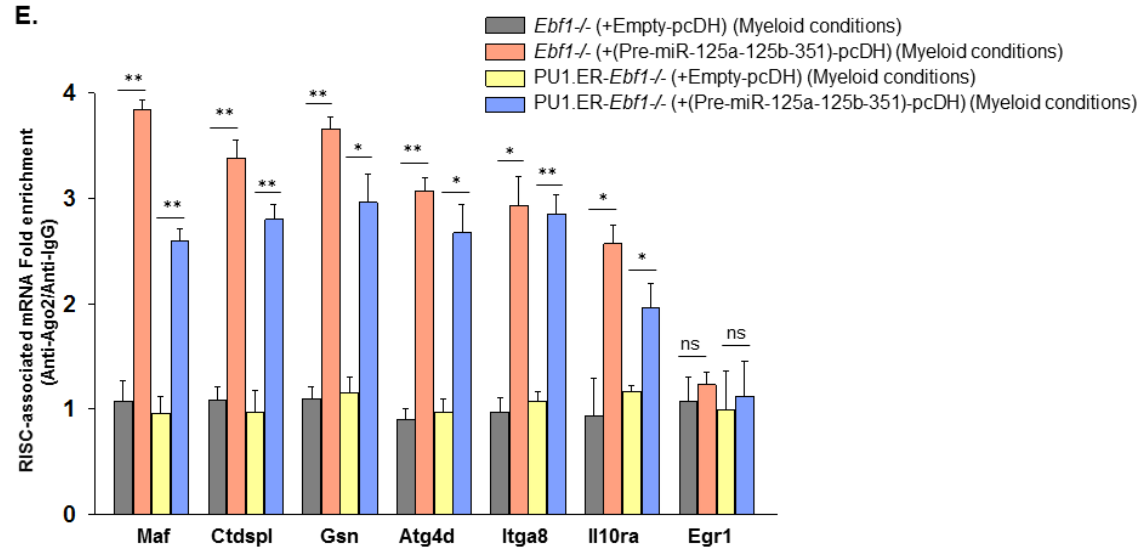
targets, it was ineffective in attenuating the myeloid differentiation capacity of *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors. Hence, effectual inhibition of a given lineage may require simultaneous repression of several lineage-associated genes, which is perhaps achieved by a concerted action of multiple miRNAs having distinct seed sequences.








D.
Fold Expression
Gsn
Ctdspl
Atg4d
Maf
Il10ra
Itga8
Egr1
ns
Ebf1-/- Progenitors
Ebf1-/- (+Empty-pcDH) (Myeloid conditions)
Ebf1-/- (+(Pre-miR-125a-125b-351)-pcDH) (Myeloid conditions)
PU1.ER-Ebf1-/- (+Empty-pcDH) (Myeloid conditions)
PU1.ER-Ebf1-/- (+(Pre-miR-125a-125b-351)-pcDH) (Myeloid conditions)


E.
Ebf1-/- (+Empty-pcDH) (Myeloid conditions)
Ebf1-/- (+(Pre-miR-125a-125b-351)-pcDH) (Myeloid conditions)
PU1.ER-Ebf1-/- (+Empty-pcDH) (Myeloid conditions)
PU1.ER-Ebf1-/- (+(Pre-miR-125a-125b-351)-pcDH) (Myeloid conditions)
RISC-associated mRNA Fold enrichment (Anti-Ago2/Anti-IgG)
Maf
Ctdspl
Gsn
Atg4d
Itga8
Il10ra
Egr1
ns

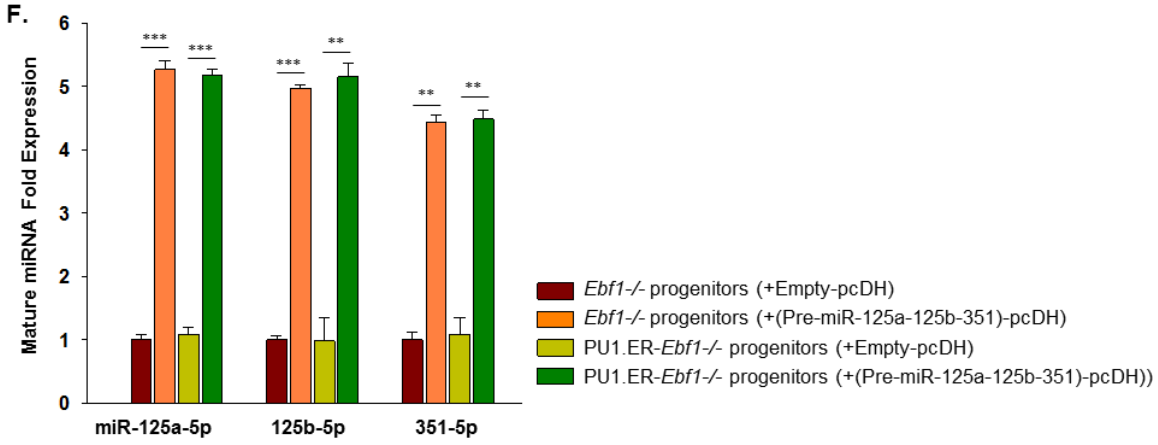

F.
Mature miRNA Fold Expression
miR-125a-5p
125b-5p
351-5p
Ebf1-/- progenitors (+Empty-pcDH)
Ebf1-/- progenitors (+(Pre-miR-125a-125b-351)-pcDH)
PU1.ER-Ebf1-/- progenitors (+Empty-pcDH)
PU1.ER-Ebf1-/- progenitors (+(Pre-miR-125a-125b-351)-pcDH))

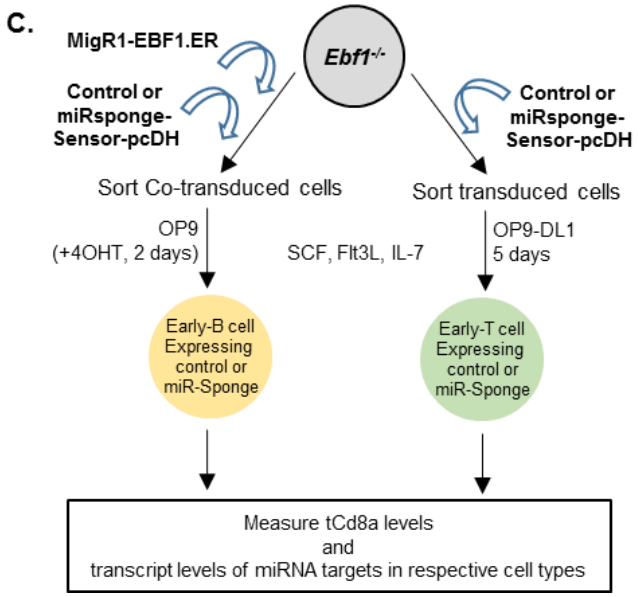
Figure 3.8. Effect of expression of miR-125-family on myeloid differentiation potential of *Ebf1-/-* and PU1.ER-*Ebf1-/-* progenitors.
(A) Expression levels of members of miR-125 family (miR-125a-5p, miR-125b-5p and miR-351-5p) in *Ebf1-/-* progenitors, Early-B cells, Early-T cells, $Mac1^{Lo}$ and $Mac1^{Hi}$ myeloid cells (generated from *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors, respectively, by culturing under myeloid-promoting conditions), quantified using stemloop qRT-PCR.
(B) Design of pcDH vector construct for ectopic expression of miR-125 family.
(C) Analysis of myeloid differentiation potential of *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors that were transduced with Empty-pcDH or pcDH vector co-expressing members of the miR-125 family. Transduced cells were cultured under myeloid-promoting conditions (see methods) and analyzed for Mac1 (CD11b) expression by flow cytometry.
(D) qRT-PCR analysis of transcript levels of myeloid–associated genes targeted by members of miR-125 family measured in *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors that were transduced with Empty-pcDH or (Pre-miR-125a-125b-351)-pcDH vector and subsequently differentiated towards myeloid lineage. Data is represented with *Ebf1-/-* progenitors as control sample.
(E) RIP-qPCR analysis for measurement of mRNA Fold enrichment (Ago2-IP vs IgG-IP) for myeloid-associated transcripts targeted by the above miRNAs in RISC-complexes of *Ebf1-/-* or PU1.ER-*Ebf1-/-* progenitors that were transduced with Empty-pcDH or (Pre-miR-125a-125b-351)-pcDH vector and subsequently differentiated towards myeloid lineage.
(F) Levels of mature or effector miRNAs - miR-125a-5p, miR-125b-5p and miR-351-5p, in *Ebf1-/-* or PU1.ER-*Ebf1-/-* cells that were transduced with Empty-pcDH or (Pre-miR-125a-125b-351)-pcDH vector.
(Data is shown as mean ± SD (*p ≤0.05; **p ≤0.01; ***p≤0.001; ns=not significant).

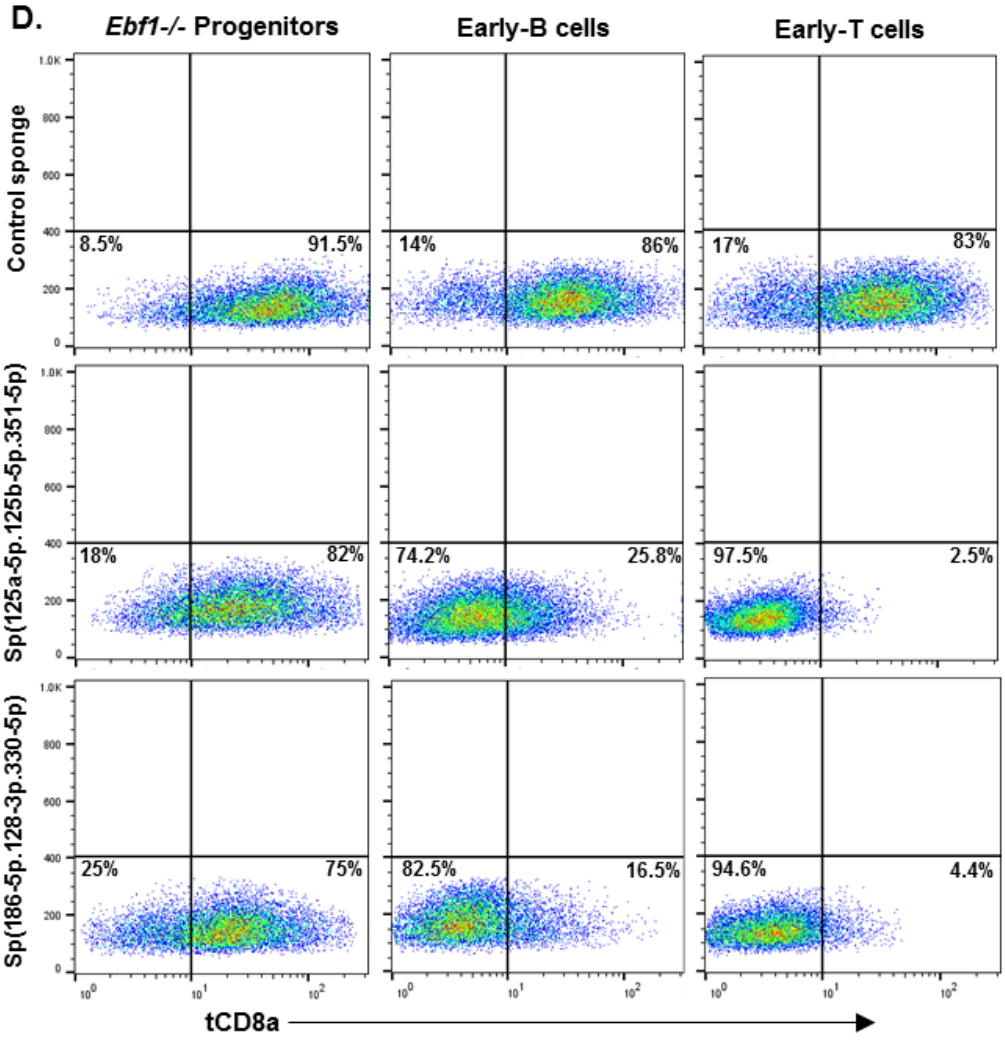
3.9. Knockdown of Ago2-enriched miRNAs rescues repression of myeloid-specific transcripts in Early-B and Early-T cells

The studies described above demonstrate that lymphocyte miRNAs can potentially suppress myeloid-associated genes and thus diminish the myeloid cell-fate option of multipotent progenitors. To rigorously test the importance of miRNAs expressed during lymphocyte differentiation, we examined if the Early-B and Early-T cell miRNAs could repress lineage-inappropriate genes independently of B or T lineage determinants (TFs). So, we carried out miRNA knockdown analysis in Early-B and Early-T cells, using miRNA sponges. We used two sets of sponges: one set for miR-186-5p, miR-128-3p and miR-330-5p and the other set for members of miR-125 family (Fig. 3.9.1A). The sponge sequences were inserted downstream of tCD8a reporter in Sensor-pcDH vector and the resulting sponge constructs, Sp(186-5p.128-3p.330-5p)-Sensor-pcDH and Sp(125a-5p.125b-5p.351-5p)-Sensor-pcDH, were transduced into *Ebf1-/-* progenitors that were allowed to differentiate into Early-B or Early-T cells (Fig. 3.9.1B and 3.9.1C). As shown in Fig. 3.9.1D, the surface levels of tCD8a in *Ebf1-/-* progenitors expressing Sp(125a-5p.125b-3p.351-5p) or Sp(186-5p.128-3p.330-5p) had decreased dramatically upon their differentiation into Early-B and Early-T cells, when compared to undifferentiated cells or those transduced with Control sponge (i.e., asGFP-Sensor-pcDH). Correspondingly, the transcript levels of tCD8a were also significantly decreased in Early-B and Early-T cells in presence of these sponges (Figure S7D), indicating effective sponge-mediated knockdown of these miRNAs in both the cell types (Fig. 3.9.1E).

Next, we checked if sequestration of these miRNAs causes de-repression of their respective myeloid-associated targets in Early-B or Early-T cells. Interestingly, the transcript levels of myeloid genes targeted by miRNAs – miR-186-5p, miR-128-3p, and miR-330-5p in Early-B cells (i.e. *Ctnnd1*, *Fam26a*, *Fgr* and *Nfam1*) or in Early-T cells (i.e., *Tgfbr3*, *Ctdspl*, *Zfhx4*, *Maf* and *Ptprg*) were found to be de-repressed in presence of the sponge Sp(186.5p-128.3p-330.5p) than the Control sponge (Fig. 3.9.2A). Similarly, the transcripts targeted by miR-125 family in Early-B cells (e.g., *Gsn* and *Atg4d*), or in Early-T cells (*Maf* and *Ctdspl)* or in both these cell

types (e.g., *Il10ra* and *Itga8*), were found to be comparatively higher in presence of Sp(125a-5p.125b-3p.351-5p) sponge compared to Control Sponge (Fig.3.9.2B). However, for most of these transcripts, the extent of de-repression was found to be modest, as they are perhaps also regulated by other miRNAs expressed in Early-B or Early-T cells. Although knockdown of miRNAs: miR-186-5p, miR-128-3p, miR-330-5p or miR-125 family could rescue the repression of their respective myeloid-associated targets, we observed no significant changes in the transcript levels of Early-B or Early-T genes (e.g., *Mb-1* and *Pax5* or *Tcf7* and *Gata3*, respectively) (Fig.3.9.2C and 3.9.2D). These analyses suggest, the contribution of Early-B and Early-T miRNAs towards repression of inappropriately-expressed genes is non-redundant and independent of the contribution by lineage-specific transcription factors. In addition, it is quite possible that miRNAs act together with the transcription factors to repress the lineage-inappropriate genes, in order to reinforce lineage identity during cell-fate commitment.


C.
MigR1-EBF1.ER
Control or miRsponge-Sensor-pcDH
Ebf1-/-
Control or miRsponge-Sensor-pcDH
Sort Co-transduced cells
Sort transduced cells
OP9 (+4OHT, 2 days)
SCF, Flt3L, IL-7
OP9-DL1 5 days
Early-B cell Expressing control or miR-Sponge
Early-T cell Expressing control or miR-Sponge
Measure tCd8a levels and transcript levels of miRNA targets in respective cell types


D.
Ebf1-/- Progenitors
Early-B cells
Early-T cells
Control sponge
Sp(125a-5p.125b-5p.351-5p)
Sp(186-5p.128-3p.330-5p)
8.5%
91.5%
14%
86%
17%
83%
18%
82%
74.2%
25.8%
97.5%
2.5%
25%
75%
82.5%
16.5%
94.6%
4.4%
tCD8a

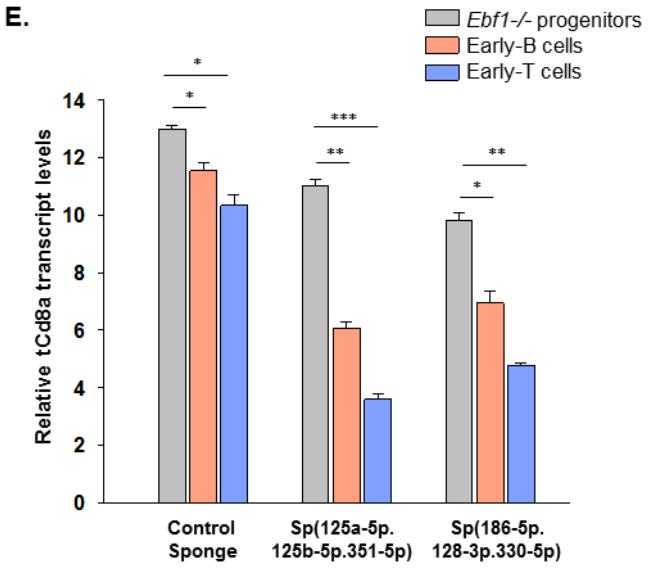


Figure 3.9.1. Design of sponges for knockdown of miRNAs in Early-B and Early-T cells.

(A) Sponge sequences (7x repeats) used for sequestering miRNAs in knockdown experiments. Each repeat sequence consists of regions that bind to 3' and 5' ends of mature miRNA (represented in green and red, respectively), and are separated by 3-4nt bulge (represented in black, unbolded). Each repeat is separated by 4-nt spacer (ATGA, represented in black, bolded).

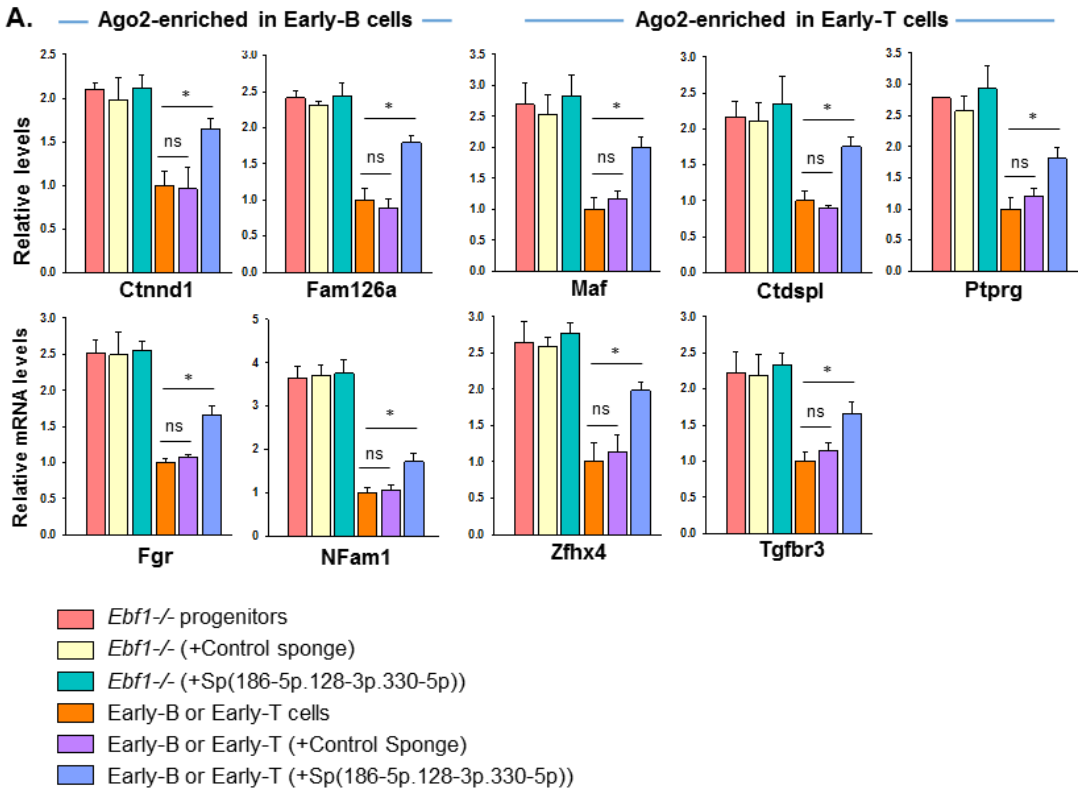
(B) Design of Control Sponge or miRNA-Sponge constructs that were made using Empty Sensor-pcDH vector.

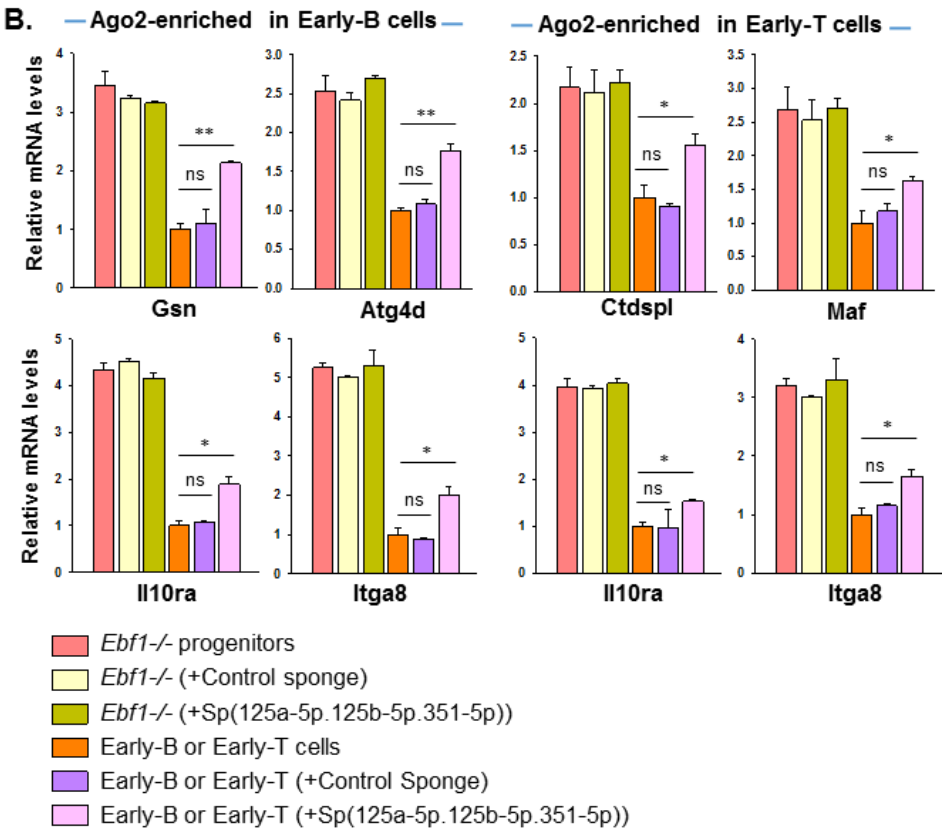
(C) Experimental strategy for studying the effect of knockdown of miRNAs expressed endogenously during differentiation of *Ebf1-/-* progenitors into Early-B and or Early-T cells.

(D) Flow cytometric measurement of tCD8a reporter levels in *Ebf1-/-* progenitors, Early-B cells and Early-T cells that expressed either the control sponge or Sp(125a.5p-125b.5p-351.5p) or Sp(186-5p.128-3p.330-5p).

(E) Transcript levels of tCD8a reporter upon expression of control or miRNA-sponges in *Ebf1-/-* progenitors, Early-B and Early-T cells.

(Data is shown as mean ± SD (*p ≤0.05; **p ≤0.01; ***p≤0.001; ns = not significant).


A.
Ago2-enriched in Early-B cells
Ago2-enriched in Early-T cells
Relative levels
Relative mRNA levels
Ctnnd1
Fam126a
Maf
Ctdspl
Ptprg
Fgr
NFam1
Zfhx4
Tgfbr3
ns
*
Ebf1-/- progenitors
Ebf1-/- (+Control sponge)
Ebf1-/- (+Sp(186-5p.128-3p.330-5p))
Early-B or Early-T cells
Early-B or Early-T (+Control Sponge)
Early-B or Early-T (+Sp(186-5p.128-3p.330-5p))


B.
Ago2-enriched in Early-B cells
Ago2-enriched in Early-T cells
Relative mRNA levels
Gsn
Atg4d
Ctdspl
Maf
Il10ra
Itga8
ns
*
**
Ebf1-/- progenitors
Ebf1-/- (+Control sponge)
Ebf1-/- (+Sp(125a-5p.125b-5p.351-5p))
Early-B or Early-T cells
Early-B or Early-T (+Control Sponge)
Early-B or Early-T (+Sp(125a-5p.125b-5p.351-5p))

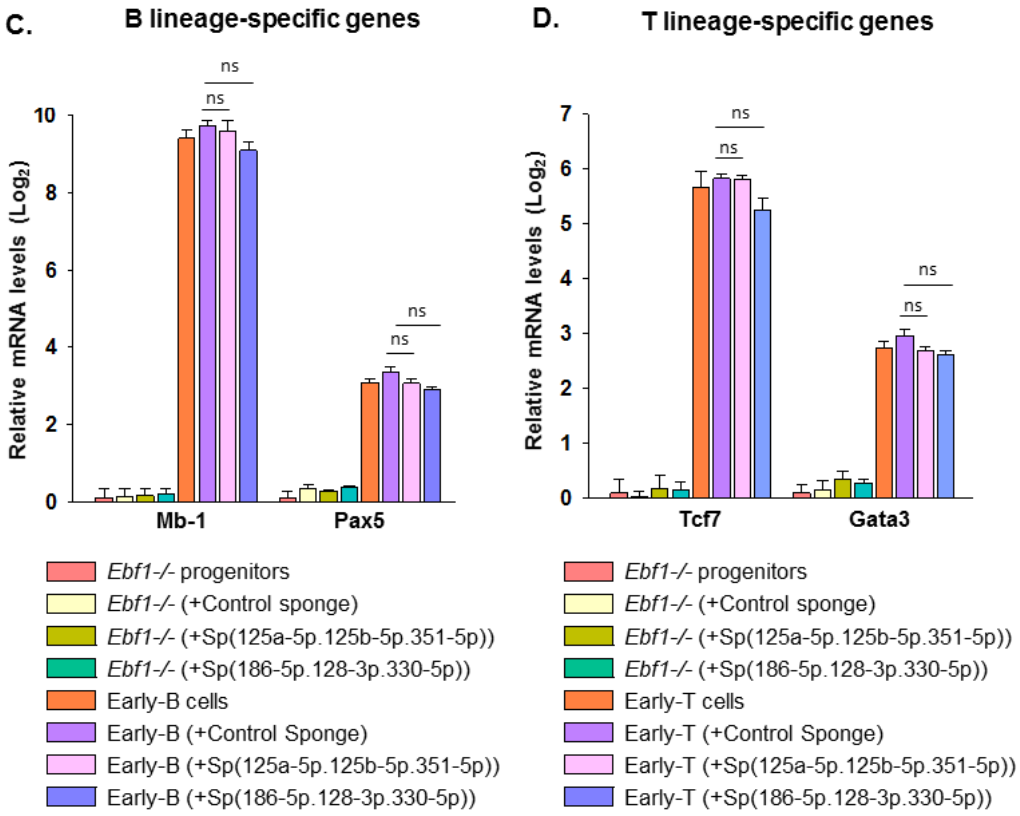


Figure 3.9.2. Effect of knockdown of lymphocyte miRNAs in Early-B and Early-T cells. (A) qRT-PCR analysis of transcript levels of myeloid-associated genes targeted by miRNAs - miR-186-5p, miR-128-3p or miR-330-5p, measured in *Ebf1-/-* progenitors and Early-B or Early-T cells that expressed Control sponge or Sp(186-5p.128-3p.330-5p). Data is represented with Early-B or Early-T cells as control sample.
(B) qRT-PCR analysis of transcript levels of myeloid-associated genes targeted by members of miR-125 family (miR-125a-5p, miR-125b-5p and miR-351-5p), measured in *Ebf1-/-* progenitors and Early-B or Early-T cells that expressed Control sponge or Sp(125a-5p.125b-5p.351-5p). Data is represented with Early-B or Early-T cells as control sample.
(C) Transcript levels of B lineage-specific genes upon differentiation of *Ebf1-/-* progenitors into Early-B cells, measured in presence of Control sponge, Sp(125a-5p.125b-5p.351-5p) or Sp(186-5p.128-3p.330-5p).
(D) Transcript levels of T lineage-specific genes upon differentiation of *Ebf1-/-* progenitors into Early-T cells, measured in presence of Control sponge, Sp(125a-5p.125b-5p.351-5p) or Sp(186-5p.128-3p.330-5p).
(Data is shown as mean ± SD (*p ≤0.05; **p ≤0.01; ***p≤0.001; ns = not significant).

3.10. Lineage-specific transcription factors potentially drive expression of conserved Early-B and Early-T miRNAs

Finally, to understand how Early-B and Early-T cells expressed similar repertoire of miRNAs in spite of having distinct gene expression programs, we decided to identify the TFs that possibly drive the expression of miRNAs in these cell types. However, identification of mechanisms that regulate miRNA expression has remained challenging primarily due to lack of definitive information about Transcription Start Sites (TSS) for locating the miRNA promoters as well as existence of non-canonical pathways for miRNA biogenesis. In line with previous attempts to characterize the miRNA promoter regions (Marson et al., 2008; Corcoran et al., 2009; Marsico et al., 2013), we used the H3K4me3 datasets (representing active promoters) from *Ebf1-/-* progenitors and Early-B cells (data unpublished) to identify the promoter regions of upregulated miRNAs (Santos-Rosa et al., 2002; Barski et al., 2007; Kouzarides, 2007).

In order to locate the TSS of the miRNA primary transcript, we scanned the genomic regions located upstream (-15kb) of the precursor miRNAs to identify the most proximal H3K4me3 marks that are enriched during differentiation. Subsequently, we used FIMO (Finding Individual Motif Occurrences, MEME suite) to search for binding sites for various B-lineage and T-lineage-specific transcription factors in a region spanning 2kb upstream of the putative TSS. Interestingly, as shown in Fig. 3.10, we identified the putative high-confidence binding sites for crucial B (EBF1 and PAX5) as well as T lineage-specific factors (TCF1 or RBPJ) within the promoter regions of miR-125 family members, miR-186 and miR-330. The promoter region of miR-128 possessed binding sites for GATA3 in addition to EBF1, PAX5, RBPJ, and TCF1. Thus, presence of binding sites for B as well as T lineage determinants within the promoter regions of the commonly-expressed miRNAs possibly suggests, these TFs are perhaps responsible for their conserved, lymphocyte-specific expression patterns in both, B and T lineages.

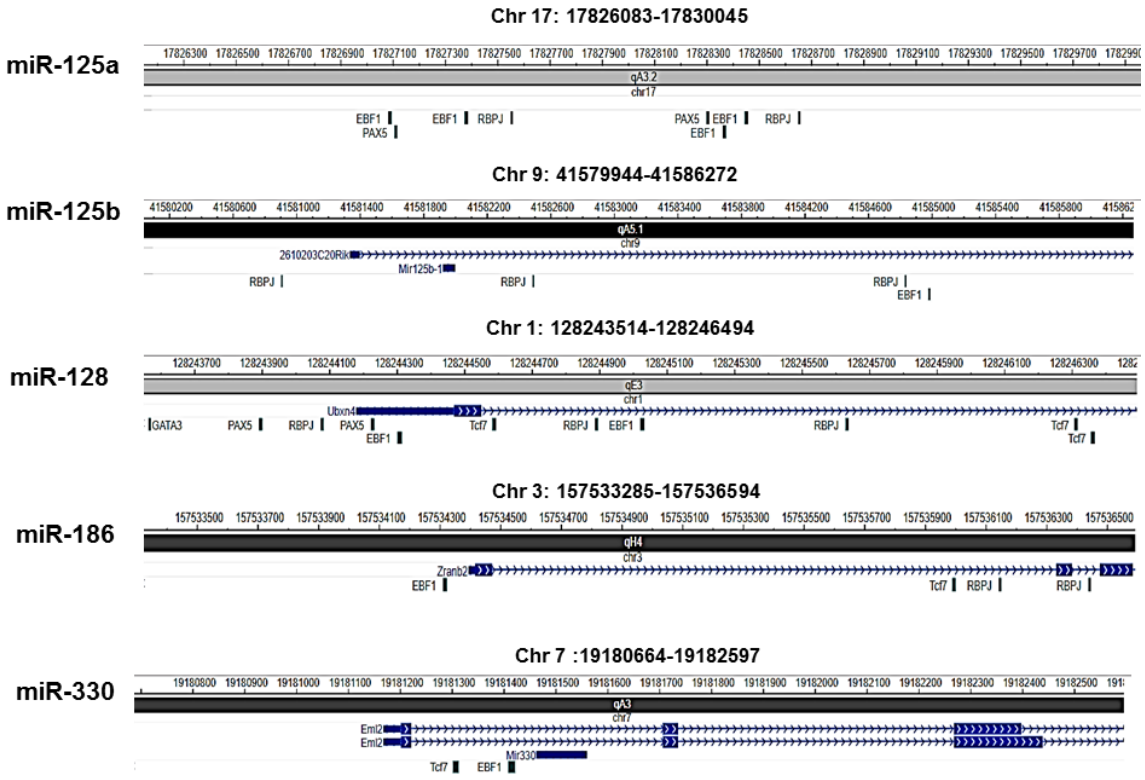


Figure 3.10. Lineage-specific transcription factors regulate expression of Early-B and Early-T miRNAs. Binding sites for B and T lineage-specific transcription factors (EBF1, PAX5 and RBPJ, TCF-1(*Tcf7*), GATA3 and BCL11B, respectively) in the promoter regions of commonly-expressed lymphocyte miRNAs.

CHAPTER-4: DISCUSSION

4.1. Early-B and Early-T cells share significantly correlated miRNA profiles but distinct Ago2-RIP profiles

The results presented here provide a comprehensive account of microRNAs that are associated with early development of B and T lymphocytes. Using the well-established model system of *Ebf1-/-* progenitors, we demonstrate that microRNAs exert an additional level of regulatory control during commitment of multipotent progenitors towards B or T cell-fates. Despite the fact that B- and T-lineage options are mutually exclusive, our studies show a significant correlation between microRNA profiles of Early-B and Early-T cells. One of the key findings of our study is, though majority of miRNAs that are activated in Early-B cells are also induced in Early-T cells, they were found to target distinct repertoire of transcripts in these two lineages, supporting the concept that miRNA targets are cell-type dependent.

The cell-type specific targeting by commonly-expressed miRNAs observed in Early-B and Early-T cells could be due to four possible reasons: Firstly, disparity in the transcriptome profiles along with the corresponding differences in transcript abundance between the two lineages, may generate variation in the availability, specificity, strength (site type), frequency and combination of miRNA binding sites within the 3'UTR regions of differentially-expressed transcripts; resulting in their differential Ago2-enrichment between the two lineages. We note that the although our miRNA-target maps may be largely restricted to the binding sites present in the 3'UTRs, the sites present in protein-coding sequences (CDS) may also be functionally relevant and possibly contribute towards target repression. Secondly, supplemental base-pairing or seed-distal interactions (involving 3'half of miRNA with target mRNA) have also been shown to increase the strength and specificity of targeting by compensating for seed-match imperfections (Broughton et al., 2016; Moore et al., 2015). Thus, despite similar seed sequences, each member of the miRNA families expressed in Early-B and Early-T cells may exhibit differential 3'-end base pairing with transcripts available in the respective cell type, thereby further contributing to diversity in target repertoires between the two lineages. Thirdly, the targeting specificity is also dictated by the expression levels of miRNAs or their relative abundance (Brancati and Großhans, 2018). So, it is quite likely that the relatively higher expression of commonly-expressed miRNAs in Early-T cells than Early-B cells allows targeting of transcripts that possess non-

canonical binding sites, in addition to transcripts that possess canonical binding sites for these miRNAs. Furthermore, our analyses show that besides the commonly-expressed miRNAs, the higher number of unique miRNAs expressed in Early-T cells than Early-B cells, contribute significantly towards the variability as well as increased targetome size of Early-T cells. Lastly, in addition to the above possibilities, the lineage-specific repression of *Hecw2* and *Egr1* transcripts (in Early-B and Early-T cells, respectively) as demonstrated by reporter assays, suggests the existence of additional cell-intrinsic mechanisms that support cell-type specific miRNA targeting, independent of miRNA or mRNA abundance. Such mechanisms may involve but are not limited to, differentially-expressed cellular RNA-Binding Proteins (RBPs) that bind to and alter the local structure of target mRNAs, subsequently influencing the accessibility or affinity of binding sites for one or more miRNAs, thus altering the targeting efficiency in a cell-type specific manner (Jacobsen et al., 2010; Hafezqorani et al., 2016; Doyle and Tenenbaum, 2014; Min et al., 2017).

We would like to point out, the progression dynamics of lineage commitment has been reported to be significantly different during differentiation of LMPPs towards B vs T lineages (Rothenberg, 2010). As a result, a possible implication of the cell-type specific miRNA-mediated targeting is, it perhaps allows differential regulation of genes which act as rate-limiting factors for commitment, and are thus required to be maintained at different thresholds in developing B and T cells.

4.2. Potential role for miRNAs in restricting lineage-inappropriate genes during B and T cell commitment

Interestingly, although the genes targeted by Early-B and Early-T cell miRNAs are distinct, a vast majority of them are associated with alternate cell-fates like NK, DC and myeloid which are required to be suppressed to enable lineage commitment. In this perspective, our study suggests that the developmental arrest seen upon absolute depletion of miRNAs in B and T cells (by abrogation of DGCR8 or Dicer) prior to lineage commitment (Koralov et al., 2008; Brandl et al., 2016; Cobb et al., 2005), could possibly be a consequence of accumulation of lineage-inappropriate genes, which in turn, prevents the transition of lineage-specified cells into a committed stage.

It is intriguing to note that, in both Early-B and Early-T cells, a large fraction of miRNA-targeted genes encode for proteins belonging to the signaling cascades like cell surface receptors and signal transducers. Since many of these molecules are involved in numerous cellular and metabolic processes, small changes in their threshold levels holds a potential to greatly impact the cellular output. Perhaps, this highlights the inconspicuous, yet significant role for miRNAs in maintaining optimal levels of such important modulators of cell development and function (Inui et al., 2010). Nonetheless, key TFs like *Tbx21* and *Eomes* for NK cells and *Irf8* for DCs were also found to be Ago2-enriched, indicating the versatility of target repertoire. Interestingly, despite the fact that PU.1 is expressed in both lymphoid and myeloid lineages (albeit at different dosage levels) (DeKoter and Singh, 2000); we find that PU.1 is not targeted by miRNAs in Early-B or Early-T cells. However, a large number of PU.1-dependent transcripts especially those specific to myeloid lineage, were found to be repressed by miRNAs in Early-B as well as Early-T cells. Such developmental similarities observed between these two divergent cell types could perhaps be partially explained by their common progenitor ancestry.

4.3. Combinatorial function of miRNAs regulates myeloid lineage restriction during lymphopoiesis

Another key observation we made is the functional significance of combinatorial action of miRNAs, which is often not captured in single miRNA studies. For instance, inhibition of myeloid lineage potential was more prominent when the miRNAs miR-186-5p, miR-128-3p and miR-330-5p, which possess distinct seed sequences, were co-expressed together than when they were expressed individually. More importantly, this effect seems to be a synergistic and cumulative outcome of simultaneous repression of several myeloid-associated genes and is not observed when multiple miRNAs having the same seed sequences were expressed together (e.g., miR-125 family). However, besides miR-186-5p, miR-128-3p, and miR-330-5p, it is quite possible that co-expression of other combinations of lymphocyte miRNAs may also generate such functional outcomes. Thus, our studies demonstrate how combinatorial targeting can impact cell-fate decisions, highlighting the fact that complex

processes like lineage differentiation may involve a multitude of individual miRNA-target interactions. However, we note that our study captures a static view of miRNAs and their target profiles, but such interactions are more likely to be highly dynamic *in vivo*, so as to allow tight buffering of protein levels and maintenance of mRNA thresholds, in order to achieve a precise control of differentiation process.

4.4. Lineage determinants drive expression of miRNAs during B and T cell development

Previous studies from our laboratory and others have led to the assembly of gene regulatory networks which govern B or T cell-fate choice of multipotent progenitors (Singh et al., 2005; Rothenberg, 2013; Boller and Grosschedl, 2014). These networks comprise of primary lineage determinants, their downstream factors and cell-extrinsic signal inputs; however, they do not integrate microRNAs, which perhaps play a potential role as critical components of these regulatory circuits. Given the functional significance of miRNAs demonstrated by this study, we propose a network for development of lymphocytes, which incorporates microRNAs as an additional tier of regulators, acting downstream of the primary determinants (TFs) for establishment of lineage identity (Fig. 4.1). The proposed network suggests three levels of regulatory control that drive the commitment of LMPPs towards B or T lineages: First, the TFs like E2A and PU.1, that are important for development of multiple lineages, set the stage for initiation of lymphocyte development within the LMPPs, in coordination with IL-7R-mediated signaling. Second, the primary lineage determinants (i.e. EBF1 or Notch1-RBPJ, activated in response to cell-intrinsic or cell-extrinsic signal inputs), initiate a self-sustaining circuit to activate and maintain the expression of additional lineage-promoting factors (e.g., PAX5 or TCF1, GATA3 and BCL11B, respectively), as well as induce the cell-type specific microRNAs. Thirdly, the lineage-specific factors act hand-in-hand with the microRNAs to antagonize the repertoire of inappropriately-expressed genes in order to block the available alternate cell-fate options, leading to establishment of lineage identity.

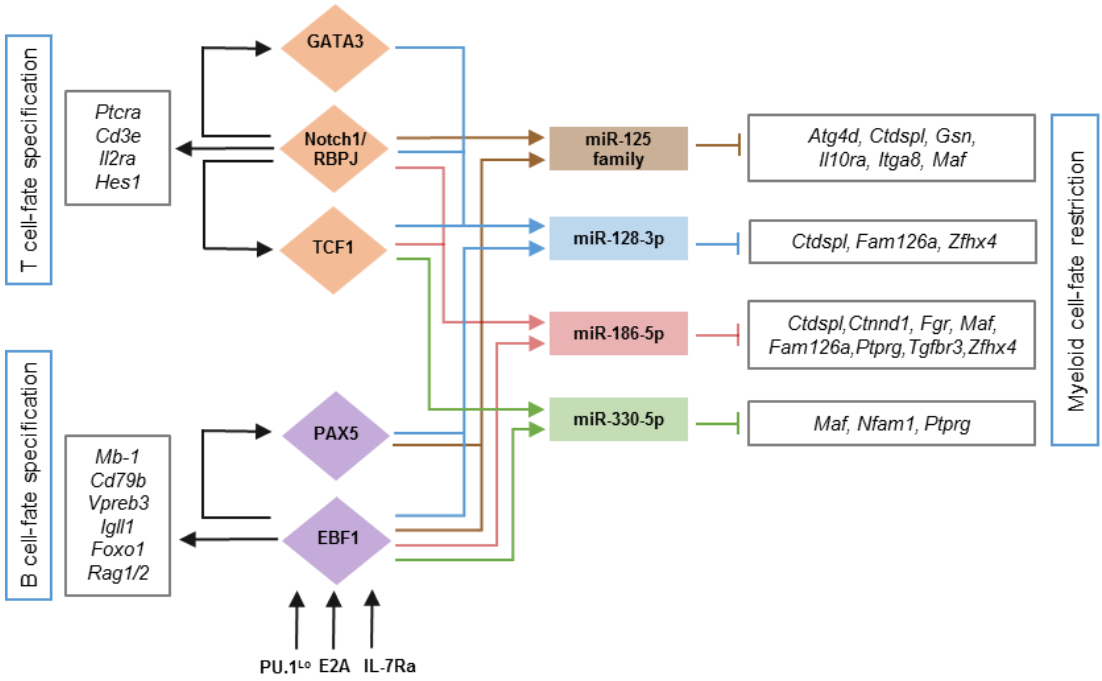


Figure 4.1. Schematic representation of the proposed gene regulatory network for B and T cell fate commitment, consisting of lineage-specific TFs that activate the cell type-specific genes as well as the miRNAs. The TFs and miRNAs act together in coordination with each other, to repress the inappropriately-expressed genes in order to establish the lineage identity.

CHAPTER-5: SUMMARY

During haematopoietic development, the multipotent progenitors were shown to promiscuously express genes associated with several lineages, which allows the progenitors to be 'primed' for differentiation towards more than one lineage. Commitment of multipotent haematopoietic progenitors towards a particular immune cell lineage involves activation of cell type-specific genes as well as silencing of genes that promote alternate cell-fates. So, the primary B- and T-lineage lineage determinants (EBF1 and Notch1-RBPJ, respectively) not only drive the expression of lineage-specific genes, but also have an obligate role to repress the inappropriately-expressed genes in order to undergo commitment towards the respective lineage. In order to effectively regulate the differentiation process, EBF1 and Notch1 may not only exert a direct regulatory control at transcriptional level, but quite likely also employ an additional layer of control perhaps at the post-transcriptional level. Based on this, we hypothesized that, the primary lineage determinants potentially induce microRNAs (miRNAs) to repress lineage-inappropriate genes and to fine-tune the random fluctuations in transcript abundance, thus providing robustness to the gene expression programs that govern B and T lymphocyte differentiation.

In the present study, we utilized the multi-lineage developmental potential of *Ebf1-/-* progenitors and differentiated them into Early-B or Early-T cells, *in vitro*, to obtain novel insights into the regulatory roles of microRNAs expressed during early developmental stages of B- and T-lymphocytes. Based on genome-wide expression analysis of miRNAs and a comprehensive analysis of their target genes, we have shown that miRNAs expressed during early stages of B and T cell development are significantly correlated, but target a distinct repertoire of genes in these lineages. Furthermore, we showed that the genes targeted by miRNAs in Early-B and Early-T cells function as TFs, signaling molecules surface receptors and metabolic regulators, and are associated predominantly with Natural Killer cell, Dendritic Cell and Myeloid lineages. Using gain-of-function experiments, we demonstrated that combinatorial expression of multiple lymphocyte microRNAs, miR-186-5p, miR-128-3p and miR-330-5p in *Ebf1-/-* progenitors, significantly attenuates their myeloid differentiation potential due to simultaneous repression of a large number of myeloid-associated transcripts. Loss-of-function experiments performed for depletion of these miRNAs in Early-B and Early-T cells resulted in

de-repression of their myeloid targets; suggesting the contribution of miRNAs towards repression of lineage-inappropriate genes is non-redundant and independent of the contribution of lineage-specific transcription factors.

In summary, this study provides functional insights on how miRNAs expressed during early development of B and T lymphocytes potentially target the inappropriately-expressed set of transcripts and exert a combinatorial and synergistic control to attenuate the alternate cell-fate choices of developing B and T cells, thus contributing towards the reinforcement of lineage identity. Collectively, our study establishes microRNAs as significant and non-redundant regulatory components of B and T lymphocyte developmental networks.

CHAPTER-6: SUPPLEMENTARY DATA

SUPPLEMENTARY TABLES

Supplementary Table-1A. List of microRNAs upregulated in Early-B cells.

miRNAs upregulated in Early-B cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)	miRNAs upregulated in Early-B cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)
miR-22-3p	8.29	miR-1981-3p	2.92
miR-486b-5p	7.23	miR-99b-5p	2.90
miR-10a-5p	7.19	miR-27b-3p	2.87
miR-335-3p	7.09	miR-34b-3p	2.82
miR-486a-5p	6.56	miR-99a-5p	2.78
miR-351-5p	6.51	miR-181b-5p	2.77
miR-181a-5p	6.07	miR-677-5p	2.69
miR-125a-5p	5.92	miR-222-3p	2.66
miR-125b-1-3p	5.11	miR-467d-5p	2.64
miR-211-5p	5.08	miR-130a-3p	2.64
miR-181c-5p	5.06	miR-130b-5p	2.63
miR-411-5p	4.99	miR-330-5p	2.63
miR-298-5p	4.41	miR-188-3p	2.61
miR-142a-5p	4.30	miR-92b-3p	2.50
miR-335-5p	4.23	miR-181a-1-3p	2.37
miR-200b-3p	4.10	miR-128-1-5p	2.31
miR-25-5p	4.01	miR-26b-3p	2.18
miR-350-5p	3.85	miR-21a-3p	2.17
miR-214-3p	3.81	miR-183-5p	2.10
miR-30c-1-3p	3.72	miR-26a-5p	1.94
miR-423-3p	3.68	miR-186-5p	1.83
miR-199b-3p	3.62	miR-466c-3p	1.79
miR-199a-3p	3.58	miR-132-3p	1.73
miR-191-5p	3.56	miR-484	1.68
miR-429-3p	3.56	miR-466a-3p	1.68
miR-128-3p	3.51	miR-221-3p	1.67
miR-671-3p	3.42	miR-466b-3p	1.66
miR-92a-1-5p	3.41	miR-466e-3p	1.63
miR-151-5p	3.38	miR-30d-3p	1.60
miR-1964-3p	3.36	miR-27b-5p	1.60
miR-100-5p	3.33	let-7i-5p	1.57
miR-34b-5p	3.33	miR-155-5p	1.56
miR-93-3p	3.26	miR-3068-3p	1.56
miR-181c-3p	3.25	miR-191-3p	1.56
miR-125b-5p	3.23	miR-16-5p	1.52
miR-92a-3p	3.16	miR-1843b-5p	1.50
miR-181a-2-3p	3.15		
miR-361-3p	3.07		
miR-301a-3p	3.01		
miR-421-3p	2.96		

Supplementary Table-1B. List of microRNAs downregulated in Early-B cells.

miRNAs downregulated in Early-B cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)	miRNAs downregulated in Early-B cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)
miR-128-2-5p	-8.99	miR-12191-3p	-2.55
miR-1929-5p	-8.31	miR-103-3p	-2.51
miR-378b	-5.96	miR-27a-5p	-2.41
miR-378a-3p	-5.80	miR-20a-5p	-2.38
miR-379-5p	-5.67	miR-345-3p	-2.24
miR-152-3p	-5.54	miR-18a-5p	-2.13
miR-24-3p	-5.46	miR-33-3p	-2.11
miR-378c	-5.30	miR-297a-5p	-2.09
miR-1936	-5.23	miR-186-3p	-2.08
miR-503-5p	-5.19	miR-365-3p	-2.06
miR-3068-5p	-4.93	miR-339-5p	-2.00
miR-1981-5p	-4.81	miR-126a-3p	-1.98
miR-542-3p	-4.80	miR-34a-5p	-1.98
miR-18b-5p	-4.66	miR-297c-5p	-1.97
miR-32-5p	-4.64	miR-148b-3p	-1.96
miR-16-1-3p	-4.60	miR-669f-3p	-1.88
miR-365-1-5p	-4.16	miR-181b-1-3p	-1.86
miR-3066-5p	-4.10	miR-350-3p	-1.82
miR-92a-2-5p	-4.01	miR-106b-5p	-1.79
miR-7a-5p	-3.97	miR-3535	-1.79
miR-142a-3p	-3.84	miR-33-5p	-1.77
miR-101b-3p	-3.77	miR-26a-2-3p	-1.71
miR-30e-3p	-3.76	miR-671-5p	-1.70
miR-19b-1-5p	-3.71	miR-219a-5p	-1.69
miR-223-3p	-3.63	miR-664-5p	-1.67
miR-30a-3p	-3.60	miR-16-2-3p	-1.65
miR-200c-3p	-3.50	miR-185-5p	-1.64
miR-28a-3p	-3.50	miR-669a-3-3p	-1.62
miR-3088-3p	-3.45	miR-700-5p	-1.62
miR-140-5p	-3.43	miR-29a-5p	-1.61
miR-140-3p	-3.40	mmu-let-7a-5p	-1.53
miR-3470b	-3.38	miR-669e-5p	-1.53
miR-30b-3p	-3.36	miR-23a-3p	-1.52
miR-24-1-5p	-3.30	miR-652-3p	-1.50
miR-29b-3p	-3.25		
miR-106a-5p	-3.24		
miR-20b-5p	-3.12		
miR-193b-5p	-2.90		
miR-146b-5p	-2.74		
miR-877-5p	-2.59		

Supplementary Table-2A. List of microRNAs upregulated in Early-T cells.

miRNAs upregulated in Early-T cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)	miRNAs upregulated in Early-T cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)	miRNAs upregulated in Early-T cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)
miR-10b-5p	18.28	miR-130a-3p	4.63	miR-1843a-5p	2.45
miR-204-5p	13.28	miR-671-3p	4.56	miR-96-5p	2.44
miR-211-5p	10.88	miR-183-3p	4.54	let-7i-5p	2.42
miR-22-3p	10.29	miR-1943-5p	4.52	miR-449a-5p	2.41
miR-3057-5p	10.20	miR-30c-1-3p	4.50	miR-1964-3p	2.34
miR-486b-5p	10.19	miR-142a-5p	4.48	miR-145a-5p	2.33
miR-486a-5p	9.54	miR-182-5p	4.24	miR-30a-5p	2.30
miR-1247-5p	9.37	miR-210-5p	4.14	miR-222-3p	2.29
miR-212-5p	9.35	miR-181c-3p	4.05	miR-423-3p	2.28
miR-335-3p	9.28	miR-199a-5p	4.03	miR-330-5p	2.24
miR-351-5p	9.19	miR-152-5p	4.02	miR-30d-3p	2.21
miR-145a-3p	8.85	miR-34b-5p	4.02	miR-92a-1-5p	2.16
miR-10a-5p	8.70	miR-181a-2-3p	3.98	miR-192-5p	2.12
miR-125b-1-3p	8.31	miR-92b-3p	3.98	miR-186-5p	2.09
miR-181a-5p	7.72	miR-350-5p	3.93	miR-141-3p	2.03
miR-99a-5p	7.29	miR-125b-5p	3.93	miR-148a-3p	2.02
miR-212-3p	7.20	miR-151-5p	3.80	miR-128-3p	2.02
miR-125a-5p	6.99	miR-301a-3p	3.70	miR-429-3p	2.00
miR-132-5p	6.06	miR-181b-5p	3.67	miR-26b-5p	1.97
miR-676-3p	6.04	miR-27b-5p	3.62	miR-181d-5p	1.91
miR-183-5p	6.02	miR-677-5p	3.60	miR-340-5p	1.88
miR-125b-2-3p	5.97	miR-127-3p	3.50	miR-1981-3p	1.86
miR-149-5p	5.84	miR-181a-1-3p	3.47	miR-195a-5p	1.84
miR-99b-5p	5.84	miR-31-5p	3.43	miR-342-3p	1.84
miR-181c-5p	5.77	miR-200b-3p	3.37	miR-92a-3p	1.80
miR-100-5p	5.75	miR-874-3p	3.31	miR-450b-5p	1.79
miR-335-5p	5.73	miR-497a-5p	3.18	miR-3068-3p	1.77
miR-199b-3p	5.70	miR-146a-5p	3.12	miR-148a-5p	1.68
miR-214-3p	5.65	miR-16-5p	3.00	miR-15a-3p	1.67
miR-199a-3p	5.65	miR-361-3p	2.96	miR-25-3p	1.62
miR-411-5p	5.50	miR-26a-5p	2.95	miR-221-3p	1.60
miR-21a-3p	5.43	miR-34b-3p	2.85	miR-361-5p	1.59
miR-132-3p	5.27	miR-25-5p	2.85	miR-126a-5p	1.58
miR-214-5p	5.19	miR-188-3p	2.78	miR-26b-3p	1.57
miR-191-5p	5.04	miR-151-3p	2.77	miR-23b-3p	1.51
miR-1933-3p	5.03	miR-1843b-5p	2.63		
miR-27b-3p	4.85	miR-146b-5p	2.54		
miR-128-1-5p	4.74	miR-322-3p	2.51		
miR-298-5p	4.72	miR-210-3p	2.51		
miR-143-3p	4.72	miR-421-3p	2.48		

Supplementary Table-2B. List of microRNAs downregulated in Early-T cells.

miRNAs downregulated in Early-T cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)	miRNAs downregulated in Early-T cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)	miRNAs downregulated in Early-T cells vs *Ebf1-/-* Progenitors	Fold Change (Log_2)
miR-1929-5p	-9.10	miR-140-5p	-2.82	miR-466a-3p	-1.56
miR-378b	-7.32	miR-669e-5p	-2.82	miR-466d-3p	-1.54
miR-378a-3p	-6.11	miR-17-3p	-2.81	miR-466c-3p	-1.53
miR-1936	-5.92	miR-24-3p	-2.77	miR-467b-5p	-1.53
miR-379-5p	-5.85	miR-140-3p	-2.69	miR-17-5p	-1.51
miR-378c	-5.72	miR-339-5p	-2.64	miR-669d-5p	-1.50
miR-19b-1-5p	-5.08	miR-466f-3p	-2.61		
miR-297c-5p	-4.91	miR-542-3p	-2.61		
miR-1981-5p	-4.83	miR-200c-3p	-2.60		
miR-32-5p	-4.82	miR-28a-3p	-2.57		
miR-503-5p	-4.62	miR-30e-3p	-2.47		
miR-92a-2-5p	-4.62	miR-219a-5p	-2.46		
miR-669a-3p	-4.33	miR-24-1-5p	-2.35		
miR-669a-3-3p	-4.25	miR-203-3p	-2.35		
miR-669o-3p	-4.19	miR-126a-3p	-2.33		
miR-669f-3p	-4.11	miR-106b-5p	-2.32		
miR-3068-5p	-4.11	miR-345-3p	-2.28		
miR-16-1-3p	-4.02	miR-106a-5p	-2.26		
miR-142a-3p	-4.01	miR-33-5p	-2.19		
miR-18b-5p	-3.79	miR-3470b	-2.18		
miR-3088-3p	-3.78	miR-20b-3p	-2.09		
miR-29b-3p	-3.76	miR-128-2-5p	-2.06		
miR-12191-3p	-3.71	miR-350-3p	-1.98		
miR-877-5p	-3.56	miR-30b-5p	-1.95		
miR-101b-3p	-3.56	miR-532-3p	-1.90		
miR-186-3p	-3.40	miR-365-3p	-1.87		
miR-20a-5p	-3.28	miR-872-3p	-1.86		
miR-365-1-5p	-3.22	miR-196b-5p	-1.84		
miR-297a-5p	-3.17	miR-103-3p	-1.78		
miR-152-3p	-3.17	miR-466e-3p	-1.74		
miR-29a-5p	-3.11	miR-7a-5p	-1.72		
miR-188-5p	-3.11	miR-16-2-3p	-1.70		
miR-33-3p	-3.04	miR-466b-3p	-1.69		
miR-18a-5p	-2.94	miR-181b-1-3p	-1.63		
miR-3066-5p	-2.93	miR-196b-3p	-1.62		
miR-466c-5p	-2.90	miR-466o-5p	-1.60		
miR-467a-3p	-2.88	miR-20a-3p	-1.60		
miR-20b-5p	-2.87	miR-138-5p	-1.58		
miR-467d-3p	-2.87	miR-671-5p	-1.58		

Supplementary Table-3A. List of microRNAs commonly-upregulated in Early-B and Early-T cells.

miRNAs commonly upregulated in Early-B and Early-T cells	Fold Change in Early-B cells (Log_2)	Fold Change in Early-T cells (Log2)	miRNAs commonly upregulated in Early-B and Early-T cells	Fold Change in Early-B cells (Log_2)	Fold Change in Early-T cells (Log2)
miR-211-5p	5.08	10.88	miR-677-5p	2.69	3.60
miR-22-3p	8.29	10.29	miR-181a-1-3p	2.37	3.47
miR-486b-5p	7.23	10.19	miR-200b-3p	4.10	3.37
miR-486a-5p	6.56	9.54	miR-16-5p	1.52	3.00
miR-335-3p	7.09	9.28	miR-361-3p	3.07	2.96
miR-351-5p	6.51	9.19	miR-26a-5p	1.94	2.95
miR-10a-5p	7.19	8.70	miR-34b-3p	2.82	2.85
miR-125b-1-3p	5.11	8.31	miR-25-5p	4.01	2.85
miR-181a-5p	6.07	7.72	miR-188-3p	2.61	2.78
miR-99a-5p	2.78	7.29	miR-1843b-5p	1.50	2.63
miR-125a-5p	5.92	6.99	miR-421-3p	2.96	2.48
miR-183-5p	2.10	6.02	let-7i-5p	1.57	2.42
miR-99b-5p	2.90	5.84	miR-1964-3p	3.36	2.34
miR-181c-5p	5.06	5.77	miR-222-3p	2.66	2.29
miR-100-5p	3.33	5.75	miR-423-3p	3.68	2.28
miR-335-5p	4.23	5.73	miR-330-5p	2.63	2.24
miR-199b-3p	3.62	5.70	miR-30d-3p	1.60	2.21
miR-214-3p	3.81	5.65	miR-92a-1-5p	3.41	2.16
miR-199a-3p	3.58	5.65	miR-186-5p	1.83	2.09
miR-411-5p	4.99	5.50	miR-128-3p	3.51	2.02
miR-21a-3p	2.17	5.43	miR-429-3p	3.56	2.00
miR-132-3p	1.73	5.27	miR-1981-3p	2.92	1.86
miR-191-5p	3.56	5.04	miR-92a-3p	3.16	1.80
miR-27b-3p	2.87	4.85	miR-3068-3p	1.56	1.77
miR-128-1-5p	2.31	4.74	miR-221-3p	1.67	1.60
miR-298-5p	4.41	4.72	miR-26b-3p	2.18	1.57
miR-130a-3p	2.64	4.63			
miR-671-3p	3.42	4.56			
miR-30c-1-3p	3.72	4.50			
miR-142a-5p	4.30	4.48			
miR-181c-3p	3.25	4.05			
miR-34b-5p	3.33	4.02			
miR-181a-2-3p	3.15	3.98			
miR-92b-3p	2.50	3.98			
miR-350-5p	3.85	3.93			
miR-125b-5p	3.23	3.93			
miR-151-5p	3.38	3.80			
miR-301a-3p	3.01	3.70			
miR-181b-5p	2.77	3.67			
miR-27b-5p	1.60	3.62			

Supplementary Table-3B. List of microRNAs commonly-downregulated in Early-B and Early-T cells.

miRNAs commonly downregulated in Early-B and Early-T cells	Fold Change in Early-B cells (Log_2)	Fold Change in Early-T cells (Log2)	miRNAs commonly downregulated in Early-B and Early-T cells	Fold Change in Early-B cells (Log_2)	Fold Change in Early-T cells (Log2)
miR-1929-5p	-8.31	-9.10	miR-33-3p	-2.11	-3.04
miR-378b	-5.96	-7.32	miR-18a-5p	-2.13	-2.94
miR-378a-3p	-5.80	-6.11	miR-3066-5p	-4.10	-2.93
miR-1936	-5.23	-5.92	miR-20b-5p	-3.12	-2.87
miR-379-5p	-5.67	-5.85	miR-140-5p	-3.43	-2.82
miR-378c	-5.30	-5.72	miR-669e-5p	-1.53	-2.82
miR-19b-1-5p	-3.71	-5.08	miR-24-3p	-5.46	-2.77
miR-297c-5p	-1.97	-4.91	miR-140-3p	-3.40	-2.69
miR-1981-5p	-4.81	-4.83	miR-339-5p	-2.00	-2.64
miR-32-5p	-4.64	-4.82	miR-542-3p	-4.80	-2.61
miR-503-5p	-5.19	-4.62	miR-200c-3p	-3.50	-2.60
miR-92a-2-5p	-4.01	-4.62	miR-28a-3p	-3.50	-2.57
miR-669a-3-3p	-1.62	-4.25	miR-30e-3p	-3.76	-2.47
miR-669f-3p	-1.88	-4.11	miR-219a-5p	-1.69	-2.46
miR-3068-5p	-4.93	-4.11	miR-24-1-5p	-3.30	-2.35
miR-16-1-3p	-4.60	-4.02	miR-126a-3p	-1.98	-2.33
miR-142a-3p	-3.84	-4.01	miR-106b-5p	-1.79	-2.32
miR-18b-5p	-4.66	-3.79	miR-345-3p	-2.24	-2.28
miR-3088-3p	-3.45	-3.78	miR-106a-5p	-3.24	-2.26
miR-29b-3p	-3.25	-3.76	miR-33-5p	-1.77	-2.19
miR-12191-3p	-2.55	-3.71	miR-3470b	-3.38	-2.18
miR-877-5p	-2.59	-3.56	miR-128-2-5p	-8.99	-2.06
miR-101b-3p	-3.77	-3.56	miR-350-3p	-1.82	-1.98
miR-186-3p	-2.08	-3.40	miR-365-3p	-2.06	-1.87
miR-20a-5p	-2.38	-3.28	miR-103-3p	-2.51	-1.78
miR-365-1-5p	-4.16	-3.22	miR-7a-5p	-3.97	-1.72
miR-297a-5p	-2.09	-3.17	miR-16-2-3p	-1.65	-1.70
miR-152-3p	-5.54	-3.17	miR-181b-1-3p	-1.86	-1.63
miR-29a-5p	-1.61	-3.11	miR-671-5p	-1.70	-1.58

Supplementary Table-4. List of microRNAs uniquely upregulated or downregulated in Early-B cells.

miRNAs uniquely upregulated in Early-B cells	Fold Change in Early-B cells (Log_2)
miR-93-3p	3.26
miR-467d-5p	2.64
miR-130b-5p	2.63
miR-484	1.68
miR-155-5p	1.56
miR-191-3p	1.56
miR-466e-3p	1.63
miR-466b-3p	1.66
miR-466a-3p	1.68
miR-466c-3p	1.79

miRNAs uniquely downregulated in Early-B cells	Fold Change in Early-B cells (Log_2)
miR-223-3p	-3.63
miR-30a-3p	-3.60
miR-30b-3p	-3.36
miR-193b-5p	-2.90
miR-146b-5p	-2.74
miR-27a-5p	-2.41
miR-34a-5p	-1.98
miR-148b-3p	-1.96
miR-3535	-1.79
miR-26a-2-3p	-1.71
miR-664-5p	-1.67
miR-185-5p	-1.64
miR-700-5p	-1.62
let-7a-5p	-1.53
miR-23a-3p	-1.52
miR-652-3p	-1.50

Supplementary Table-5. List of microRNAs uniquely upregulated or downregulated in Early-T cells.

miRNAs uniquely upregulated in Early-T cells	Fold Change in Early-T cells (Log_2)	miRNAs uniquely upregulated in Early-T cells	Fold Change in Early-T cells (Log_2)	miRNAs uniquely upregulated in Early-T cells	Fold Change in Early-T cells (Log_2)
miR-10b-5p	18.28	miR-210-3p	2.51	miR-192-5p	2.12
miR-204-5p	13.28	miR-1843a-5p	2.45	miR-141-3p	2.03
miR-3057-5p	10.20	miR-96-5p	2.44	miR-148a-3p	2.02
miR-1247-5p	9.37	miR-210-5p	4.14	miR-26b-5p	1.97
miR-212-5p	9.35	miR-199a-5p	4.03	miR-181d-5p	1.91
miR-145a-3p	8.85	miR-152-5p	4.02	miR-340-5p	1.88
miR-212-3p	7.20	miR-127-3p	3.50	miR-195a-5p	1.84
miR-132-5p	6.06	miR-31-5p	3.43	miR-342-3p	1.84
miR-676-3p	6.04	miR-874-3p	3.31	miR-450b-5p	1.79
miR-125b-2-3p	5.97	miR-497a-5p	3.18	miR-148a-5p	1.68
miR-149-5p	5.84	miR-146a-5p	3.12	miR-15a-3p	1.67
miR-214-5p	5.19	miR-151-3p	2.77	miR-25-3p	1.62
miR-1933-3p	5.03	miR-146b-5p	2.54	miR-361-5p	1.59
miR-143-3p	4.72	miR-322-3p	2.51	miR-126a-5p	1.58
miR-183-3p	4.54	miR-449a-5p	2.41	miR-23b-3p	1.51
miR-1943-5p	4.52	miR-145a-5p	2.33		
miR-182-5p	4.24	miR-30a-5p	2.30		

miRNAs uniquely down regulated in Early-T cells	Fold Change in Early-T cells (Log_2)	miRNAs uniquely down regulated in Early-T cells	Fold Change in Early-T cells (Log_2)
miR-669a-3p	-4.33	miR-196b-5p	-1.84
miR-669o-3p	-4.19	miR-466e-3p	-1.74
miR-188-5p	-3.11	miR-466b-3p	-1.69
miR-466c-5p	-2.90	miR-196b-3p	-1.62
miR-467a-3p	-2.88	miR-466o-5p	-1.60
miR-467d-3p	-2.87	miR-20a-3p	-1.60
miR-17-3p	-2.81	miR-138-5p	-1.58
miR-466f-3p	-2.61	miR-466a-3p	-1.56
miR-203-3p	-2.35	miR-466d-3p	-1.54
miR-20b-3p	-2.09	miR-466c-3p	-1.53
miR-30b-5p	-1.95	miR-467b-5p	-1.53
miR-532-3p	-1.90	miR-17-5p	-1.51
miR-872-3p	-1.86	miR-669d-5p	-1.50

Supplementary Table-6. List of microRNAs showing inverse pattern between Early-B and Early-T cells.

miRNAs showing inverse expression pattern between Early-B and Early-T cells	Fold Change in Early-B cells (Log_2)	Fold Change in Early-T cells (Log_2)
miR-148a-3p	-0.92	2.02
miR-30a-5p	-1.00	2.30
miR-449a-5p	-1.02	2.41
miR-146a-5p	-0.91	3.12
miR-146b-5p	-2.74	2.54
miR-127-3p	*-4.73*	*3.50*
miR-210-3p	*-1.30*	*2.51*
miR-23b-3p	-0.76	1.51
miR-466e-3p	1.63	-1.74
miR-466b-3p	1.66	-1.69
miR-138-5p	1.35	-1.58
miR-466a-3p	1.68	-1.56
miR-466c-3p	1.79	-1.53

Supplementary Table-7A. List of Ago2-enriched transcripts in Early-B cells.

Ago2-enriched transcripts in Early-B cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-B cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-B cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-B cells	Fold Enrich-ment
Sarm1	181.97	4930528A17Rik	14.13	Tbx21	3.79	8430408G22Rik	2.70
Clec4a1	178.08	4930538K18Rik	12.57	Tbxa2r	3.77	H2-Ab1	2.65
Hsf2bp	162.10	Oscar	12.45	Dgkh	3.76	Oasl2	2.64
Kcnb1	145.26	Tmtc1	11.93	Ssc4d	3.73	Crkl	2.63
Hecw2	101.58	Tex38	11.25	Sp7	3.61	F8a	2.62
Fhl2	88.59	Pcdhb22	10.77	Calml3	3.61	Rnase6	2.62
Hbb-y	86.16	Il1rl1	10.75	Npl	3.61	Trim32	2.61
Olfr433	82.42	Cyp4f39	10.57	Pilrb2	3.59	Dbnl	2.59
Pgbd5	78.18	Otub2	10.11	Slc4a8	3.57	Hck	2.57
Tenm4	68.57	Fam109b	8.76	Lcp2	3.57	Lrrc20	2.57
Slc34a1	61.20	Tmem221	8.45	Rdh12	3.49	Jmjd4	2.56
Doc2a	50.89	Vwa1	8.07	Tlr13	3.47	Bcdin3d	2.54
Egfem1	47.51	Klf8	7.27	Ltf	3.35	Zfp553	2.54
Lctl	44.52	Spon2	7.06	Tlr9	3.24	Tnfaip8l1	2.53
Cct6b	42.03	Gm21057	7.06	Scamp5	3.22	8030462N17Rik	2.52
Krba1	41.10	Gm19589	6.83	Scamp3	3.18	Il12rb1	2.50
Crb3	36.39	9430083A17Rik	6.81	Ipcef1	3.15	Aurka	2.50
Foxc2	36.08	Gpr155	6.79	Edc3	3.14	9530077C05Rik	2.50
L3hypdh	34.79	Apon	5.87	Eomes	3.12	Slc22a15	2.50
Nxpe5	33.65	Gm16023	5.77	Disc1	3.11	Il16	2.49
Angpt4	32.88	43346Rik	5.50	Mycl	3.09	Irf2bp1	2.48
Rsad2	28.70	Ciita	5.41	H2-Aa	3.07	Map3k2	2.48
Gm5801	25.39	Hic1	5.40	Prss30	3.06	Cd300a	2.47
Usp18	25.18	Tmem71	5.40	Map3k9	3.06	Rag1	2.47
2210416O15Rik	24.66	E330009J07Rik	5.35	Plcxd2	3.00	Irf8	2.47
Itga8	23.99	Cntd1	5.34	Sash3	3.00	Thap11	2.47
Crim1	23.92	Jrk	5.34	Cmpk2	2.95	Trim14	2.45
Wnt11	23.60	Cdh1	5.27	Adora3	2.95	Cbl	2.45
4732416N19Rik	23.40	St8sia1	5.25	Tmem51	2.94	Fam49a	2.45
Cited4	20.27	Kcnj1	4.86	Gpr162	2.91	Amica1	2.45
Vrtn	19.56	Neurl1a	4.70	Zfp862-ps	2.90	Wars	2.44
Ap1m2	19.35	Nfatc2	4.63	Lilra6	2.86	Mttp	2.44
Adgra3	18.56	Bcl6	4.30	Naif1	2.85	Prf1	2.44
Adora1	16.72	Clcn5	4.24	Frmd8	2.85	Ephb6	2.43
Gm11992	16.36	Slfn2	4.24	Sh3bp5l	2.85	Dffa	2.43
Apln	16.21	H2-T24	4.12	D11Wsu47e	2.83	Faap24	2.41
Exoc3l4	15.98	Arnt2	4.09	Cd74	2.74	Fbxl8	2.40
Prr18	15.94	Gpr68	3.96	Slc9a3r1	2.73	Lysmd1	2.40
Fgf21	14.39	1600010M07Rik	3.91	Xist	2.73	Chtf8	2.40
Ctnnd1	14.22	1810021B22Rik	3.87	H2-Eb1	2.72	Ttpal	2.39

Ago2-enriched transcripts in Early-B cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-B cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-B cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-B cells	Fold Enrich-ment
Ccr7	2.38	1600014C10Rik	2.18	Fam73b	2.08	Cdca5	2.00
Trib3	2.38	Ppfia4	2.17	Tbcc	2.08	Tcf19	2.00
Klhl11	2.36	Stim1	2.17	Pelo	2.08	Zfp41	2.00
D030056L22Rik	2.36	Ankrd46	2.17	Nxpe3	2.07	Glg1	2.00
Gpr183	2.36	Cyb5d2	2.17	Anxa6	2.07	Slc39a9	2.00
Plekhb2	2.35	Arhgap19	2.16	Hcls1	2.07	Thap4	2.00
Snx17	2.35	Zfp296	2.16	Bcl2l1	2.07		
Gsn	2.34	Itgam	2.16	Irf5	2.07		
Enam	2.34	Cnksr3	2.16	Tbc1d8	2.06		
Rtn4rl1	2.34	Bmf	2.16	Zfp775	2.06		
Rbm14	2.33	Necap2	2.16	Foxo3	2.05		
St3gal2	2.33	Tprn	2.15	Calhm2	2.05		
Sec22a	2.32	Etnk1	2.15	Tln1	2.05		
Unc119b	2.32	Grap	2.15	Kcnk5	2.05		
Cacnb3	2.31	Nipal3	2.15	Itpkc	2.05		
Asprv1	2.28	Dars2	2.15	Tmco4	2.05		
Irgm2	2.28	Ftsj2	2.15	Stx6	2.05		
Slc39a4	2.27	Ndrg1	2.14	Piwil2	2.05		
Ap5m1	2.27	Dph1	2.14	Zfp661	2.05		
Mlycd	2.27	Dlat	2.14	Kmo	2.04		
Sla	2.26	Wrnip1	2.14	Lig3	2.04		
Pilra	2.26	Trex1	2.14	Rad51d	2.04		
Rnf41	2.26	Pdxp	2.13	Cds1	2.04		
Slc9a3r2	2.26	Smad3	2.13	Rps6ka1	2.03		
1700021K19Rik	2.25	Gtf3c4	2.13	Nup62	2.03		
AI467606	2.25	Traf1	2.13	Ptger3	2.03		
Arhgap6	2.25	Sh2b2	2.12	Snai3	2.02		
Il10ra	2.25	Racgap1	2.12	Atg4d	2.02		
Narf	2.25	Pitpnc1	2.11	Cyb561a3	2.02		
Cnp	2.25	Smarcd2	2.10	Hs1bp3	2.02		
Tsc22d4	2.23	Nfam1	2.10	Fam126a	2.02		
E430018J23Rik	2.22	Mon1b	2.10	Nrros	2.02		
Ak4	2.21	Trmt12	2.10	Hn1	2.01		
2610035D17Rik	2.21	Rad54l	2.10	Fkbp4	2.01		
Tfeb	2.20	Trim45	2.10	Lrrc14	2.01		
Zdhhc24	2.20	Hmox2	2.09	Exoc8	2.01		
Aacs	2.20	Eme1	2.09	Bag2	2.01		
Adgre1	2.19	4930579G24Rik	2.09	Zswim1	2.01		
BC147527	2.18	Tmem229b	2.09	Alkbh5	2.01		
Fgr	2.18	Smap2	2.08	Rptor	2.00		

Supplementary Table-7B. List of Ago2-enriched transcripts in Early-T cells.

Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment
Zcchc14	479.34	Dysf	87.14	Akap3	50.30	Gm3230	31.76
Cd36	399.46	Bmper	87.01	Fhod3	50.30	2210416O15Rik	31.41
Fam13a	261.98	Hcar1	83.72	Rbm12b1	49.97	Trabd2b	30.97
Cidec	246.98	Myo1d	82.60	Fgd4	49.81	Gm9079	30.92
Ctdspl	236.90	Creb3l3	82.33	Rab40b	49.15	Pvrl1	30.88
Lgals12	233.57	Plekhg6	81.61	Cd177	48.99	Aqp7	30.74
Lctl	207.85	Lrg1	80.87	Chst2	48.42	Mir3113	30.42
Mast4	169.36	Car8	79.48	9330102E08Rik	48.39	Slc7a11	30.36
Mapk8ip2	168.90	Fbxo24	77.16	D130040H23Rik	47.30	Sparcl1	30.26
Cav1	152.77	Sec16b	76.80	Rbm20	43.95	Atp1a2	29.71
Tnnc2	151.39	Apba2	76.25	Nos1	43.60	Pappa2	29.64
Mir7087	149.02	Gas2l1	75.60	Laptm4b	43.53	Itgb6	29.32
Kif21a	142.94	Alms1-ps2	73.92	1700019B03Rik	43.39	Mgll	29.26
Bcar3	141.02	Meg3	73.75	Arhgap42	42.20	Plin4	29.21
Ppp2r2c	132.97	Cst6	72.92	Olfm2	41.40	Efna5	29.11
Hspb8	131.01	Timp4	71.43	Gpr161	40.98	Crb3	29.10
Zfhx4	129.06	Nckap5	70.14	Rusc2	40.97	Lrfn3	28.03
Sema6a	128.98	2610206C17Rik	69.75	Hopx	40.03	Otx1	27.93
Cts8	123.09	Ptprg	69.07	Fam213a	40.00	Gm2115	27.85
Col5a3	122.37	Cyp4v3	68.93	Zic1	39.49	Gas1	27.83
A530016L24Rik	119.83	Foxc2	67.32	Mroh2a	39.38	Rnf125	27.72
6330407A03Rik	119.28	Usp13	64.08	Hspa1b	39.30	Nfatc4	27.64
Adamts12	112.27	Cacna2d1	62.77	Tgfbr3	38.98	Loxl1	27.34
Aox1	111.87	Arhgap32	62.33	Csgalnact1	38.36	Prss35	27.03
Clstn3	109.39	Panx2	61.17	Ttll7	37.65	Tnxb	26.02
Cpeb3	106.10	Mir1839	60.73	Gm14827	37.58	A430078I02Rik	25.90
Rgl3	104.77	Gas2l2	60.42	Pcbp3	37.32	Lpl	25.13
Rgma	103.66	Arxes1	60.20	Ankrd33b	36.19	Ltbr	24.47
Tnfsf15	102.94	Pitx2	59.89	Cxx1c	35.95	Cav2	24.42
Epb4.1l3	102.48	Zfp423	59.77	Popdc2	35.70	Gm21319	24.39
Rnf144b	101.72	Apc2	58.78	Trim43a	35.55	Doc2a	24.31
Ptpdc1	99.87	Marveld2	57.04	Mir191	35.11	Slc34a1	24.24
Magi1	99.24	Arhgap44	55.72	Inha	34.81	Twist2	24.20
Acnat1	99.09	Vangl1	54.11	Pyroxd2	34.42	Col6a2	24.16
Col4a1	96.86	Ppl	53.40	Pygo1	33.85	Cldn15	23.70
Tmed6	93.05	Fam160a1	52.44	Clic5	33.61	Pparg	23.56
Zfp697	92.34	Draxin	52.30	Nid1	32.61	Aqp1	23.55
Pard3	89.69	Syt7	52.05	Nacad	32.33	Sdc2	23.42
Nfia	89.20	Gad1	52.05	Hivep3	32.18	Sardh	23.34

Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment
Slc16a3	23.30	Tmcc2	16.03	Myoz3	10.97	Matn2	7.52
Dgat2	23.03	Ehd2	16.01	Gdf15	10.93	Ecm2	7.50
Dkk3	22.80	Nr2f1	15.77	Flot2	10.90	Aldoc	7.49
4833427F10Rik	22.79	Slc1a2	15.63	Mir5107	10.87	Lin7b	7.47
Thbs2	22.49	Col4a2	15.20	Me1	10.81	Gal3st3	7.25
Lrrn3	22.44	Tbx18	15.07	Mmp2	10.39	Mgp	7.08
Mir8112	22.43	Arfgef3	14.96	Efcab9	10.12	Ggh	6.99
Adam23	22.40	Epb4.1l1	14.69	Lrrc14b	10.09	Hdgfrp3	6.95
G630055G22Rik	21.79	Hdac11	14.68	Nkx2-1	10.06	Csf1	6.57
9330188P03Rik	21.70	Sdr39u1	14.62	Gm19619	9.57	Itih2	6.53
Lingo3	21.66	Galm	14.42	Dcbld1	9.37	Abcg4	6.52
Olfr433	21.65	Wisp1	14.30	Cdcp2	9.33	Zfp941	6.45
Sytl4	21.60	Col6a3	14.29	Atf3	9.32	Cth	6.30
Sox2ot	21.00	Anks6	14.09	Jam2	9.31	9330159M07Rik	6.28
Btbd17	20.90	Kank1	14.02	Emp2	9.26	Rell2	6.20
Cygb	20.82	Masp2	13.78	Dapk1	9.20	Klf8	6.18
Scd1	20.64	1600023N17Rik	13.64	Arxes2	9.13	Cd300lg	6.11
Cspg5	20.52	Zfp248	13.61	Tnks1bp1	8.88	Gm16023	6.10
Cct6b	20.07	Nrap	13.56	Rnd1	8.82	Foxs1	6.08
Ank1	20.07	4933431E20Rik	13.34	Rasgef1b	8.75	Dmpk	5.97
Efemp1	19.90	A630020A06	13.34	Shox2	8.75	Ccdc78	5.97
Slc2a4	19.62	Tmem117	13.28	Zfp575	8.69	Rufy4	5.93
Ddr2	19.59	Gprc5b	13.19	Rin2	8.64	Itga8	5.90
Sarm1	19.14	Ltbp3	13.18	Sp7	8.62	Ablim2	5.73
Crx	19.04	Begain	12.85	Il1rl1	8.56	Tns2	5.72
Sult4a1	18.63	Trim52	12.69	Lrrn4cl	8.40	Eps8l2	5.70
4930452B06Rik	18.60	Gm10190	12.65	Akap5	8.36	St6galnac3	5.66
Fhad1	18.52	Onecut2	12.55	Adcy4	8.25	Trim2	5.65
Mical2	18.04	Tmem141	12.42	Serpinb9g	8.24	Acmsd	5.65
Cdkl4	17.49	Mir3061	12.42	Foxc1	8.15	Cmtm4	5.63
Dlg2	17.22	Amotl1	12.35	Mtss1l	8.14	Usp17la	5.63
Kctd15	17.21	Nacc2	12.25	Jag1	8.12	Cldn12	5.58
Gm10814	17.14	Mrc2	11.96	Grb14	8.10	Snx21	5.54
Maf	17.03	Wasf1	11.86	Nr2f2	8.09	Obscn	5.49
Pawr	16.83	Gm8615	11.16	4931406B18Rik	7.99	9330136K24Rik	5.49
Mmp14	16.81	9430083A17Rik	11.14	Fam213b	7.90	Mir3091	5.46
Apba1	16.64	Fut1	11.14	Hmgn3	7.75	Wfikkn1	5.44
Fn1	16.56	4932418E24Rik	11.01	Chst1	7.65	4631405J19Rik	5.38
Ppt2	16.41	Pclo	11.00	Gm2848	7.54	Cdc42bpb	5.29

Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment
Nipa1	5.28	A830082N09Rik	3.96	Amigo3	3.32	Bod1l	2.92
Frs3	5.20	Nr1h3	3.95	Rnf43	3.32	Scai	2.92
Rdh18-ps	5.15	Nav2	3.95	Neurl1b	3.29	Tnrc6b	2.92
Megf8	5.11	Fggy	3.91	Optn	3.27	Pth1r	2.90
Slfn8	5.04	Rtp3	3.91	Tmem59l	3.26	Zxda	2.89
Lamb1	4.89	Fam109b	3.86	Klf7	3.25	1700021K19Rik	2.89
Osgin1	4.85	Gm11127	3.85	Ankrd11	3.25	Armc2	2.88
Gucy1a2	4.84	Fndc3b	3.83	H2-T24	3.24	Sh3pxd2a	2.88
Glrp1	4.84	Slc27a2	3.79	Nr4a2	3.19	Hyal3	2.88
Slc6a8	4.82	Nfkbiz	3.78	Pram1	3.17	Mlph	2.87
Ank	4.79	Nr6a1	3.76	Rilp	3.16	Cass4	2.85
Egr3	4.79	Rspo1	3.70	Dfna5	3.15	Fktn	2.85
Smad6	4.77	Uprt	3.69	Kantr	3.14	Rel	2.85
Tanc2	4.72	Pou4f1	3.68	Mdc1	3.13	Pld2	2.84
Nupr1	4.67	Gm20362	3.68	Zcchc24	3.11	Erc1	2.84
Rps15a-ps4	4.53	Erp27	3.68	A630072M18Rik	3.11	Pde4dip	2.83
Ttc28	4.51	Mir142	3.61	Plek	3.08	Zfhx3	2.83
Tbc1d12	4.46	Pbx3	3.61	Cep170b	3.08	Cntnap1	2.82
P2rx3	4.42	Gm8234	3.60	Zfp36l1	3.07	Tmem81	2.82
Grik5	4.41	LOC106740	3.59	Apobr	3.06	Runx2	2.82
Mllt4	4.36	Golgb1	3.56	BC037704	3.06	Irs1	2.81
Slc6a9	4.34	Tnfaip2	3.53	Lpp	3.06	Josd1	2.81
Arhgef12	4.30	Car15	3.52	Stkld1	3.05	Paqr4	2.79
Rimbp3	4.28	Nes	3.49	Myo18b	3.05	Rps6ka2	2.79
Mmp15	4.27	Stk35	3.48	1700063D05Rik	3.05	Dock9	2.79
Hsd17b1	4.25	Syne1	3.48	Thsd1	3.03	Fam102a	2.77
Psd	4.22	Ubn2	3.46	Mylk	3.02	Sorbs1	2.77
Gpd1	4.20	Stau2	3.44	Akap12	3.00	Actn2	2.76
Gja4	4.18	Dnaic1	3.42	Zfr2	3.00	Traf6	2.76
Tanc1	4.17	Clmn	3.41	Gzmm	3.00	Traf3	2.75
Mfsd9	4.15	Npff	3.40	BC020402	3.00	Tle2	2.75
Egr1	4.13	Car5b	3.39	Cdk14	2.99	Braf	2.74
Zfp438	4.10	5031414D18Rik	3.39	Apbb1	2.98	6330549D23Rik	2.74
Mei4	4.09	A730090H04Rik	3.39	Alg6	2.98	Ss18l1	2.73
Espnl	4.08	Cep83os	3.39	Dtx4	2.95	Rrm2b	2.73
A530088E08Rik	4.08	Tmem8b	3.37	Kdm7a	2.95	Zdhhc15	2.72
Trib1	4.02	Gm7102	3.37	Fam49a	2.94	Usp31	2.72
Rnf223	4.02	Slc5a5	3.37	Rbm47	2.93	Nfat5	2.71
Gpt2	3.99	Prkcb	3.35	Plxna1	2.92	Card9	2.71

Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment
Ids	2.71	Rhbdd2	2.53	Bhlhe41	2.38	Nr2c2	2.26
Lhfpl2	2.70	Mthfr	2.53	Pcnx	2.38	Zfp871	2.26
Arhgap6	2.70	Rab11fip5	2.52	Marveld1	2.38	Mecp2	2.26
Iqsec1	2.70	Nr3c1	2.52	Arrb1	2.37	Cacnb1	2.26
Elovl7	2.69	Tceanc	2.52	Plekha1	2.36	Fam73b	2.26
Mzf1	2.69	1110002L01Rik	2.51	Bsn	2.36	Arhgap19	2.25
Slc30a4	2.68	Atf7	2.50	Eif3a	2.36	Zbtb34	2.25
Cpeb2	2.68	Snx27	2.49	Sh3rf1	2.35	Zfp275	2.25
Foxn3	2.67	Neat1	2.49	Slc7a5	2.35	Usp35	2.24
Sorcs2	2.67	Stap2	2.48	Pten	2.34	Cbfa2t3	2.24
Rassf3	2.67	Rc3h1	2.48	Sh2b2	2.34	Fam53c	2.23
Ppargc1b	2.67	Zfp937	2.48	4930480K23Rik	2.33	Ftx	2.23
Xist	2.66	Arhgef17	2.47	3110043O21Rik	2.32	Git1	2.23
Fam185a	2.66	Abhd1	2.47	Slc38a9	2.31	Atmin	2.23
Pvr	2.66	Zfp236	2.46	Adam22	2.31	Lix1l	2.23
Ccnyl1	2.65	Gfod1	2.46	Tbc1d2b	2.31	Smim13	2.22
Zdhhc21	2.64	Alkbh5	2.46	Zbtb4	2.31	Jag2	2.22
Zkscan8	2.64	Camk2a	2.46	Nfatc2	2.31	Tcof1	2.22
Tnfrsf10b	2.63	Dennd2c	2.45	Arl5a	2.31	Spata2	2.22
Hjurp	2.63	Rnf24	2.44	Gcnt2	2.30	Pank3	2.22
Mib2	2.62	Msantd1	2.44	Trib3	2.30	Prr14l	2.22
Cenpf	2.61	Fam43a	2.44	Gm29766	2.30	Igdcc4	2.22
Serac1	2.61	Zfp346	2.44	Setd1b	2.30	Epb4.1l4b	2.21
Purb	2.60	Golga4	2.44	Etnk1	2.30	Rps19-ps3	2.21
Hmox1	2.59	Slc29a4	2.43	Agps	2.29	Numa1	2.21
Crebl2	2.59	Tbcel	2.42	Arid3b	2.28	1700094D03Rik	2.21
Anxa4	2.58	Mmgt1	2.42	Syde2	2.28	Coa5	2.21
Ptges3l	2.57	Ube2h	2.41	Ulk2	2.28	Tbc1d2	2.20
Mn1	2.57	Ppm1k	2.40	Lmbrd2	2.28	A730017L22Rik	2.20
Ypel2	2.57	Cd69	2.40	Apof	2.28	Fgfrl1	2.20
Ubxn7	2.57	Exoc6b	2.39	Zbtb44	2.28	4833439L19Rik	2.20
Dbndd2	2.57	Mllt6	2.39	Nlrp1a	2.27	Tmem265	2.20
Ntsr1	2.57	Inf2	2.39	Srxn1	2.27	Gm16973	2.20
Caskin2	2.56	Fam120c	2.39	Fitm2	2.27	Pou2f2	2.20
Glg1	2.56	Trim65	2.39	Rhob	2.27	Ptprv	2.20
Kcnh2	2.55	Nol4l	2.39	Hmg20a	2.27	Rnf169	2.19
Gm17644	2.55	Klf13	2.39	Lin7c	2.27	Wipf2	2.19
Aqp9	2.54	Mdm4	2.39	Nfix	2.27	Nckipsd	2.19
Slc5a3	2.53	Zfp661	2.39	Pptc7	2.26	Creb1	2.19

Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment	Ago2-enriched transcripts in Early-T cells	Fold Enrich-ment
Il10ra	2.19	Tmem170b	2.14	Parvb	2.07	Per3	2.03
Dusp5	2.18	Fam20b	2.13	Pum2	2.07	Rab31	2.03
Lrrc58	2.18	Mapk1ip1	2.13	Blcap	2.07	Tspan5	2.03
Bcat1	2.18	Rb1	2.12	Phf19	2.07	Pigw	2.03
Zfp467	2.18	Fnip2	2.12	Stx12	2.07	Gtf2ird1	2.03
Mex3a	2.17	Slc20a2	2.12	Slc35e1	2.06	Gm7120	2.03
Slc17a8	2.17	Hivep1	2.12	Atxn7l3b	2.06	Ranbp10	2.03
Rnf150	2.17	AI597479	2.12	Egln1	2.06	Hsd3b7	2.03
Ogfrl1	2.17	Mafg	2.12	Ube3a	2.06	Zfp770	2.03
Dzip3	2.17	Safb2	2.12	Atf5	2.06	Dact3	2.03
Stk40	2.16	Ylpm1	2.11	Zdhhc24	2.06	Traf5	2.03
Irf2bp2	2.16	Socs7	2.11	Crkl	2.06	Rab11fip3	2.03
Dusp11	2.16	Apba3	2.11	Eif4ebp3	2.06	Fam219b	2.03
Tbc1d24	2.16	Ppapdc2	2.11	Tpm2	2.06	Ppp2r5b	2.02
Fam168a	2.16	Gltscr1	2.11	Jdp2	2.06	Epm2aip1	2.02
Crat	2.16	Ino80d	2.11	Stx16	2.05	Pign	2.02
Tet1	2.16	Dok4	2.11	Ano1	2.05	Taf3	2.02
Inpp4a	2.16	Zfp704	2.11	Wipf1	2.05	Jade1	2.02
Ctla2a	2.16	Mkl2	2.11	Foxp1	2.05	9230114K14Rik	2.02
Msantd2	2.16	Rapgef1	2.10	Rab43	2.05	Trim41	2.02
Slc39a3	2.16	Senp1	2.10	Bscl2	2.05	Speg	2.02
Cnksr3	2.15	Slc7a1	2.10	Tex2	2.05	Foxk1	2.02
Lrtm2	2.15	Gm9159	2.10	Pck2	2.05	Fam120a	2.02
Camkk2	2.15	Sox4	2.10	Chpf2	2.05	Fam53b	2.02
Rap1gds1	2.15	Tpst1	2.10	C2cd2l	2.05	Prkar2a	2.02
Slc25a37	2.15	Samd4b	2.09	6030419C18Rik	2.05	Rbms2	2.02
Vkorc1l1	2.15	Tgif2	2.09	BC030336	2.04	Mapk1ip1l	2.01
Sbk1	2.15	Adgrl1	2.09	Pigv	2.04	Tifab	2.01
Ric8b	2.15	Foxj2	2.08	Zfp235	2.04	Klhdc10	2.01
Hipk3	2.15	Whrn	2.08	Cebpg	2.04	A630001G21Rik	2.01
Zfp654	2.14	Sema6b	2.08	Pou2f1	2.04	Ap1s3	2.01
Tshz1	2.14	Bend4	2.08	Ubl3	2.04	Icmt	2.01
Tet3	2.14	Krcc1	2.08	Gatad2a	2.04	Notch2	2.00
Tmppe	2.14	Fzd7	2.08	Hirip3	2.04	Mtss1	2.00
Ube2w	2.14	6330416G13Rik	2.08	Ncoa3	2.04	Sh2b1	2.00
Reep6	2.14	Tbc1d8	2.07	Foxn2	2.03	Tnfsf8	2.00
Mthfd2	2.14	Pmaip1	2.07	Abl2	2.03	Filip1	2.00
Zdhhc1	2.14	B230219D22Rik	2.07	Tet2	2.03	Tmem245	2.00
Adipor2	2.14	Rhoq	2.07	AI661453	2.03	Arhgef6	2.00
						Socs5	2.00

Supplementary Table-7C. List of commonly-Ago2-enriched transcripts in Early-B and Early-T cells.

Ago2-enriched transcripts in Early-B and Early-T cells	Fold Enrichment in Early-B cells	Fold Enrichment in Early-T cells
Lctl	44.52	207.85
Foxc2	36.08	67.32
2210416O15Rik	24.66	31.41
Crb3	36.39	29.10
Doc2a	50.89	24.31
Slc34a1	61.20	24.24
Olfr433	82.42	21.65
Cct6b	42.03	20.07
Sarm1	181.97	19.14
9430083A17Rik	6.81	11.14
Sp7	3.61	8.62
Il1rl1	10.75	8.56
Klf8	7.27	6.18
Gm16023	5.77	6.10
Itga8	23.99	5.90
Fam109b	8.76	3.86
H2-T24	4.12	3.24
Fam49a	2.45	2.94
1700021K19Rik	2.25	2.89
Arhgap6	2.25	2.70
Xist	2.73	2.66
Glg1	2.00	2.56
Alkbh5	2.01	2.46
Zfp661	2.05	2.39
Sh2b2	2.12	2.34
Nfatc2	4.63	2.31
Trib3	2.38	2.30
Etnk1	2.15	2.30
Fam73b	2.08	2.26
Arhgap19	2.16	2.25
Il10ra	2.25	2.19
Cnksr3	2.16	2.15
Tbc1d8	2.06	2.07
Zdhhc24	2.20	2.06
Crkl	2.63	2.06

Supplementary Table-8A. Lineage association of Ago2-enriched transcripts from Early-B cells as shown by ImmGen.

Early-B cells	Natural Killer cells		Dendritic Cells			Myeloid cells		
Rad54l	Tenm4	Tmco4	Hecw2	Ppfia4	Il10ra	Clec4a1	Gpr155	Slc9a3r2
Aurka	Slc34a1	Fam126a	Egfem1	Zfp296	Tfeb	Rsad2	Kcnj1	AI467606
Trib3	Lctl	Foxc2	Nxpe5	Bmf	Rtn4rl1	Usp18	Bcl6	Narf
Rad54l	Crim1	Ap1m2	Usp18	Grap	Cnp	Ctnnd1	1600010M07Rik	Tsc22d4
Eme1	Cited4	Oscar	Wnt11	Traf1	Aacs	Oscar	Dgkh	Aacs
4930579G24Rik	Prr18	Enam	Vrtn	Nxpe3	BC147527	Cyp4f39	Pilrb2	Stim1
Fkbp4	Tex38	Unc119b	Pcdhb22	Irf5	Necap2	Neurl1a	Slc22a15	Ankrd46
Cdca5	Pcdhb22	Ap5m1	Cyp4f39	Tbc1d8	Etnk1	Slfn2	Cbl	Arhgap19
Tcf19	Spon2	Fgf21	Klf8	Kmo	Nipal3	H2-T24	Pilra	Itgam
Otub2	Hic1	Cyp4f39	Ciita	Ptger3	Sh2b2	Calml3	Tfeb	Ndrg1
Aurka	Nfatc2	Vwa1	Hic1	Nrros	Trim45	Tlr13	Fgr	Trmt12
9530077C05Rik	Gpr68	Calml3	St8sia1	Sarm1	Pelo	Tlr9	Atg4d	Hcls1
Trib3	Tbx21	Xist	Nfatc2	Olfr433	Cds1	Cmpk2	Hecw2	Tbc1d8
Arhgap19	Ipcef1	Nipal3	Gpr68	Itga8		Tmem51	Slc34a1	Piwil2
Dars2	Plcxd2	Kcnk5	1600010M07Rik	Exoc3l4		Lilra6	Itga8	Rps6ka1
Rad54l	Tnfaip8l1	Zswim1	1810021B22Rik	Prr18		Oasl2	Vrtn	Hs1bp3
Eme1	Il12rb1		Tmem229b	Ctnnd1		Hck	Gm11992	Thap4
4930579G24Rik	Ttpal		Cyb561a3	Neurl1a		Lrrc20	Apln	Gm19589
Fkbp4	Slc39a4		Npl	Bcl6		Cd300a	Il1rl1	Sp7
Rptor	Grap		Slc4a8	Adgre1		Thap11	Spon2	Irgm2
Cdca5	Nxpe3		Tlr9	Sp7		Fam49a	Gm21057	Foxo3
Tcf19	Alkbh5		Disc1	Adora3		Sla	Tmem71	Alkbh5
Zfp41	Fhl2		Mycl	Rnase6		Il10ra	E330009J07Rik	
Early-T cells	Doc2a		H2-Aa	Irf8		Adgre1	1810021B22Rik	
Fam109b	Cct6b		Prss30	Unc119b		Ssc4d	1600014C10Rik	
Ak4	Crb3		Tmem51	Asprv1		Ppfia4	Lcp2	
Nxpe3	Adora1		Gpr162	Slc9a3r2		Bmf	Rdh12	
Bag2	Tmtc1		Zfp862-ps	Anxa6		Necap2	Ltf	
Spon2	Il1rl1		Naif1	Tenm4		Nfam1	Scamp5	
Npl	Kcnj1		Cd74	Crb3		Irf5	Ipcef1	
Ephb6	Eomes		Xist	Angpt4		Calhm2	Map3k9	
Rad54l	F8a		H2-Eb1	Ap1m2		Fam126a	9530077C05Rik	
Eme1	Prf1		H2-Ab1	Oscar		Nrros	Il16	
Pelo	Rtn4rl1		Hck	Tex38		Sarm1	Trib3	
Kcnk5	Itgam		Cd300a	Tmem221		Egfem1	Gsn	
Fkbp4	Wrnip1		Fam49a	Gpr155		Gm5801	Enam	
	Smad3		Ccr7	Cdh1		Cited4	St3gal2	
	Pitpnc1		Gpr183	Clcn5		Fgf21	Asprv1	
	Anxa6		Cacnb3	Tbxa2r		Tex38	Slc39a4	
	Bcl2l1		Arhgap6	Calml3		Fam109b	Rnf41	

Supplementary Table-8B. Lineage association of Ago2-enriched transcripts from Early-T cells as shown by ImmGen.

Early-B cells	Early-T cells	Natural Killer cells				
Zfp423	Tmed6	Lctl	Rimbp3	Dok4	Serpinb9g	Fam53b
Aqp7	Arhgap32	Zfhx4	Hsd17b1	Adgrl1	Akap5	Rbms2
4930452B06Rik	Arhgap44	Alms1-ps2	Pou4f1	Egln1	Cth	Lgals12
9230114K14Rik	Aqp1	Cst6	Gm20362	Wipf1	Eps8l2	Cts8
4933431E20Rik	Sytl4	Ptprg	Dnaic1	BC030336	Cmtm4	A530016L24Rik
Gprc5b	Gprc5b	Arxes1	Clmn	Ranbp10	Ank	1110002L01Rik
Ltbp3	Me1	Zfp423	Myo18b	Rab11fip3	Arhgef12	4930480K23Rik
Zfp575	Dapk1	Rbm12b1	Timp4	Speg	Rnf223	1700063D05Rik
Chst1	Arxes2	Rab40b	Zfr2	Zcchc14	Pbx3	4933431E20Rik
Gal3st3	Matn2	Chst2	Dtx4	Cidec	Car15	1600023N17Rik
P2rx3	Smad6	Nos1	Plxna1	Cav1	Npff	Ppp2r5b
Mei4	Gm7102	Hopx	Armc2	Sema6a	Car5b	Rbm20
Espnl	Thsd1	Pygo1	Mlph	Car8	Rnf43	Lrfn3
Fggy	Hyal3	Atp1a2	Irs1	Sec16b	Tmem59l	Gm21319
Fam109b	6330549D23Rik	Gm2115	Fam102a	Usp13	Plek	Wasf1
Slc27a2	Kcnh2	Rnf125	Sorbs1	Apc2	Pth1r	Foxc1
Gm8234	Camk2a	Tnxb	Actn2	Marveld2	Cntnap1	Acmsd
Mylk	Fam120c	Slc34a1	Traf3	Ppl	Rrm2b	Mir3091
Dtx4	Syde2	Sardh	Usp31	Gad1	Zdhhc15	Lamb1
Tle2	Igdcc4	Thbs2	Ppargc1b	Fhod3	Elovl7	Gucy1a2
Sorcs2	Fgfrl1	Sox2ot	Xist	Olfm2	Dbndd2	Tanc1
Cenpf	Mex3a	Efemp1	Pvr	Zic1	Arhgap44	Gm7102
Mn1	Slc17a8	Sult4a1	Tnfrsf10b	Itgb6	Zfp937	A630072M18Rik
Trib3	Ctla2a	Mical2	Ptges3l	Plin4	Bhlhe41	Gzmm
Cacnb1	Reep6	Maf	Ntsr1	Crb3	Akap3	Paqr4
Ftx	Zdhhc1	Galm	Aqp9	Doc2a	Arl5a	Ctla2a
Jag2	Zfp704	Masp2	Slc29a4	Cct6b	Srxn1	Mkl2
Tet1	Sox4	Nrap	Ube2h	Ank1	Dusp5	Eif4ebp3
Zdhhc1	9230114K14Rik	Rgl3	Ppm1k	Fhad1	Dusp11	Gatad2a
Zfp704		Amotl1	Cd69	Kctd15	Lrtm2	Tnfsf8
Phf19		Emp2	Mllt6	Gm10814	Sbk1	Gas2l2
Tpm2		Mgp	Fam120c	Pawr	Fam20b	Loxl1
Pck2		Foxs1	Zbtb4	Col6a3	Socs7	Prss35
AI661453		Tns2	Nfatc2	Anks6	Rapgef1	Cspg5
Dlg2		St6galnac3	Hmg20a	Foxc2	Foxj2	Wisp1
		Wfikkn1	Zfp275	Ltbp3	Rhoq	Fggy
		Rdh18-ps	Smim13	Gm8615	Atf5	Sorcs2
		Slc6a8	Ptprv	Pclo	Ano1	AI661453
		Egr3	Rap1gds1	Myoz3	Pigv	Slc6a9
		Ttc28	Zfp654	Dcbld1	Per3	Tmppe
		Grik5	Tshz1	Jam2	Tspan5	Il1rl1

Dendritic Cells						
Mtss1	Fam213b	Stap2	Hcar1	Hyal3	Draxin	6030419C18Rik
Blcap	Aldoc	Abhd1	Apba2	Josd1	Akap3	2610206C17Rik
Tnfsf15	Itih2	Msantd1	Gas2l1	Paqr4	Arhgef6	D130040H23Rik
Magi1	Klf8	Ppm1k	Meg3	Pvr	Hivep3	9330136K24Rik
Bmper	Foxs1	Mllt6	Bscl2	Ccnyl1	Aqp7	9330102E08Rik
Alms1-ps2	Tns2	Plekha1	Marveld2	Tnfrsf10b	Efna5	4631405J19Rik
Chst2	Snx21	Adam22	Arhgap44	Hmox1	Crb3	3110043O21Rik
Nos1	Frs3	Nfatc2	Fam160a1	Ptges3l	Cav2	
Gpr161	Megf8	Apof	Syt7	Inf2	Sdc2	
Gm14827	Lamb1	Nlrp1a	Tex2	Nol4l	Cdkl4	
Popdc2	Ank	Usp35	Laptm4b	Marveld1	Dlg2	
Trim43a	Tbc1d12	Cbfa2t3	Arhgap42	Socs7	Apba1	
Nid1	Slc6a9	Lix1l	Zic1	Srxn1	Anks6	
Slc7a11	Psd	Il10ra	Ankrd33b	Cacnb1	Wasf1	
Sparcl1	Egr1	Rnf150	Trabd2b	Atmin	Gm8615	
Itgb6	Nr1h3	Mthfd2	Pappa2	Jag2	Matn2	
Nfatc4	Gm11127	Tmem170b	Otx1	Nckipsd	Obscn	
Lrrn3	Nfkbiz	Dok4	Prss35	Dusp5	Mir7087	
Adam23	Gm20362	Rapgef1	Dkk3	Fam168a	Rdh18-ps	
Cygb	Gm8234	Bend4	Thbs2	Vkorc1l1	Tanc2	
Fhad1	Gm7102	Tbc1d8	Olfr433	Fnip2	Gja4	
Mical2	Slc5a5	Pmaip1	Scd1	Slc20a2	Nav2	
Mmp14	Prkcb	Zdhhc24	Sarm1	Hivep1	Fndc3b	
Tmcc2	Stkld1	Eif4ebp3	Hdac11	Mkl2	Tmem8b	
Slc1a2	Cdk14	Stx16	Nacc2	Tpst1	H2-T24	
Tbx18	Fam49a	Rab43	Dapk1	Stx12	Zcchc24	
Masp2	Rbm47	Gm9159	Sp7	Tpm2	Erc1	
Jade1	Sh3pxd2a	Ncoa3	Adcy4	Jdp2	Runx2	
Trim52	Rel	Tet2	Mtss1l	Ano1	Rrm2b	
Tmem141	Pld2	Zfp770	Hmgn3	C2cd2l	Ypel2	
Mrc2	Pde4dip	Dact3	Rell2	Hsd3b7	Atf7	
Gdf15	Tle2	Trim41	Itga8	Rab11fip3	Sh3rf1	
Mir5107	Nfat5	Ap1s3	Eps8l2	Prkar2a	Sh2b2	
Me1	Ids	Tnfsf8	Whrn	Tifab	Ftx	
Gm19619	Arhgap6	Cd36	Cdc42bpb	Cav1	Irf2bp2	
Atf3	Slc30a4	Lgals12	Egr3	Bcar3	Inpp4a	
Shox2	Xist	Mast4	Arhgef12	Rgl3	Msantd2	
Rin2	Crebl2	Tnnc2	Trib1	Zfp697	Adipor2	
Foxc1	Anxa4	Rnf144b	Neurl1b	Pard3	9230114K14Rik	
Jag1	Ntsr1	Tmed6	Nr4a2	Arhgap32	A530016L24Rik	
Nr2f2	Gm17644	Dysf	Zfp36l1	Gas2l2	4933431E20Rik	

Myeloid cells							
Cidec	Masp2	Rhob	Panx2	Slc6a8	Zfp770	Rtp3	4930480K23Rik
Ctdspl	Trim52	Nfix	Ogfrl1	Tbc1d12	Fam219b	Erp27	6030419C18Rik
Cav1	Onecut2	Lix1l	Rbm20	Gpd1	Foxk1	Tnfaip2	4833427F10Rik
Kif21a	Tmem141	Pou2f2	Nos1	Trib1	Prkar2a	Syne1	4933431E20Rik
Ppp2r2c	Nacc2	Il10ra	Zic1	Nr1h3	Agps	Dnaic1	4931406B18Rik
Sema6a	Fut1	Zfp467	Mroh2a	Fam109b	Tnnc2	Clmn	9330102E08Rik
Aox1	Atf3	Crat	Tgfbr3	Gm11127	Mir7087	Optn	9330159M07Rik
Rgma	Emp2	Camkk2	Nacad	Nfkbiz	Bcar3	Klf7	A630001G21Rik
Col4a1	Rin2	Hipk3	Aqp7	Pbx3	Myo1d	H2-T24	Krcc1
Gas2l1	Fam213b	Tpst1	Pappa2	Nes	Lrg1	Pram1	Tbc1d8
Alms1-ps2	Chst1	Parvb	Itgb6	Stk35	Car8	Apobr	Blcap
Meg3	Lin7b	Stx12	Otx1	Myo18b	Sec16b	Kdm7a	Slc35e1
Cst6	Ggh	Bscl2	Gm2115	Cdk14	Vangl1	Tnrc6b	Atxn7l3b
Cyp4v3	Cd300lg	Chpf2	Twist2	Apbb1	Fam160a1	Zfhx3	Tspan5
Usp13	Mir3091	Pten	Col6a2	Rps6ka2	Fhod3	Josd1	Gtf2ird1
Gas2l2	Slfn8	Ubl3	Pparg	Usp31	Cd177	Braf	Speg
Ppl	Nupr1	Foxn2	Slc16a3	Foxn3	Csgalnact1	Rrm2b	Filip1
Syt7	Egr1	Rab31	Tbc1d24	Rassf3	Ankrd33b	Arhgap6	
Fgd4	Fndc3b	Hsd3b7	Mir8112	Crebl2	Slc7a11	Iqsec1	
Rusc2	Rspo1	Dact3	Btbd17	Dennd2c	Mgll	Elovl7	
Hopx	Slc5a5	Taf3	Scd1	Trim65	Slc34a1	Cpeb2	
Hspa1b	Rilp	Tifab	Ddr2	Gcnt2	Cldn15	Hjurp	
Pygo1	Zcchc24	Notch2	Sarm1	Lmbrd2	Dgat2	Mib2	
Sparcl1	Fam49a	Cd36	Sult4a1	Zbtb44	Cdkl4	Ubxn7	
Atp1a2	Mlph	Fam13a	Tmcc2	Srxn1	Gm10814	Aqp9	
Gas1	Ss18l1	Lgals12	Col4a2	Cbfa2t3	Galm	Arhgef17	
Loxl1	Card9	Zfhx4	Kank1	Atmin	Anks6	Rnf24	
Tnxb	Ids	Cts8	Inpp4a	Stk40	Flot2	Trib3	
Lpl	Xist	Adamts12	Mrc2	Fam168a	Mmp2	Reep6	
Ltbr	Serac1	Cpeb3	Wasf1	Tmppe	Gm19619	Tmem170b	
Cav2	Hmox1	Magi1	Efcab9	Adipor2	Rnd1	Mafg	
Gm21319	Anxa4	Acnat1	Nkx2-1	Rb1	Adcy4	Apba3	
Thbs2	Ypel2	Tmed6	Cdcp2	Gm9159	Nr2f2	Ino80d	
Crx	Dbndd2	Pard3	Shox2	Fzd7	Ric8b	Dok4	
Maf	Mthfr	Dysf	Jag1	Pmaip1	Dmpk	Foxj2	
Apba1	Rab11fip5	Creb3l3	Grb14	Rhoq	Rufy4	Sema6b	
Fn1	Fam43a	Plekhg6	Slc39a3	Zdhhc24	Itga8	Arid3b	
Nr2f1	Klf13	Fbxo24	Mgp	Jdp2	Cldn12	Arhgap19	
Tbx18	Pcnx	Timp4	Usp17la	Ano1	Nipa1	Zbtb34	
Wisp1	Arrb1	Ptprg	Cdc42bpb	Ncoa3	Osgin1	Tbc1d2	
Col6a3	Nlrp1a	Cacna2d1	Rdh18-ps	Tet2	Mfsd9	Ube2w	

Supplementary Table-9. Transcripts targeted by miRNAs miR-186-5p, miR-128-3p and miR-330-5p. Genes targeted by more than one miRNA are indicated in bold.

Myeloid-specifc targets of miR-186-5p in Early-B and Early-T cells		Myeloid-specifc targets of miR-128-3p in Early-B and Early-T cells		Myeloid-specifc targets of miR-330-5p in Early-B and Early-T cells	
9530077C05Rik	**Agps**	Anxa4	Atp1a2	**9530077C05Rik**	Adipor2
A630001G21Rik	**Ankrd33b**	**Arrb1**	Cav2	Agps	**Ankrd33b**
Alkbh5	Ano1	**Bmf**	**Chst1**	Al467606	Anks6
Ankrd46	**Apba1**	Cacna2d1	**Crebl2**	Arhgap19	**Apba1**
Bcar3	Atmin	**Cbfa2t3**	Cts8	**Arrb1**	Arid3b
Cbl	**Braf**	**Cpeb2**	Dennd2c	**Bmf**	**Cbfa2t3**
Cpeb3	Cdk14	**Cpeb3**	**Fam13a**	Dbndd2	Crat
Csgalnact1	**Chst1**	**Ctdspl**	Fhod3	Ddr2	**Crebl2**
Ctdspl	Clmn	**Fam126a**	**Fndc3b**	Foxk1	Emp2
Ctnnd1	**Cpeb2**	Fbxo24	Gpd1	**Gas2l1**	Fam168a
Fam126a	Cst6	**Fzd7**	Josd1	**Gcnt2**	Gpd1
Fgr	**Fam13a**	**Hipk3**	**Lpl**	Hs1bp3	**Ino80d**
Foxj2	**Fndc3b**	**Ino80d**	**Pparg**	Il10ra	Klf13
Foxn2	Foxn3	**Jag1**	Ube2w	Lrrc20	Myo18b
Fzd7	**Hipk3**	Kif21a	**Ubxn7**	**Maf**	Nfix
Gas2l1	Hjurp	Mafg	**Ypel2**	Map3k9	Pou2f2
Gcnt2	Hspa1b	**Magi1**	**Zbtb34**	Mthfr	**Ppp2r2c**
Ids	**Ino80d**	Mfsd9	**Zfhx4**	Myo1d	**Ptprg**
Lcp2	Itgb6	Neurl1a		Nacc2	**Rgma**
Maf	**Jag1**	Notch2		**Ncoa3**	Tbc1d12
Magi1	**Lpl**	Nr2f2		Necap2	**Tnrc6b**
Ncoa3	Nkx2-1	Ogfrl1		**Nes**	Tnxb
Nfkbiz	Optn	**Onecut2**		Nfam1	**Zbtb34**
Onecut2	Otx1	**Pcnx**		**Onecut2**	
Pcnx	Ppp2r2c	**Pten**		**Pcnx**	
Pten	Prkar2a	Rsad2		Piwil2	
Rnf24	**Ptprg**	**Sema6a**		Rab11fip5	
Scd1	Rab31	**Slc7a11**		Rbm20	
Sema6a	Rb1	Ss18l1		Rhoq	
Slc22a15	**Rgma**	Stk35		Rps6ka1	
Stk35	Slc35e1	Stk40		Sarm1	
Syt7	**Slc7a11**	**Tbc1d8**		Scamp5	
Tmem170b	Stx12	Tfeb		**Scd1**	
Tspan5	Syne1	**Tmem170b**		**Slc22a15**	
Usp17la	Tgfbr3	Tmem51		Slc9a3r2	
Vangl1	**Tnrc6b**	Ubl3		Slfn2	
Wasf1	**Ubxn7**	**Vangl1**		**Syt7**	
Wisp1	**Ypel2**	**Wisp1**		**Tbc1d8**	
Zbtb44	**Zcchc24**	**Zcchc24**		Usp31	
Zfhx3	**Zfhx4**	**Zfhx3**		**Zbtb44**	

Supplementary Table-10. Transcripts targeted by miR-125-family members: miR-125a-5p, miR-125b-5p and miR-351-5p.

Myeloid-specifc targets of miR-125 family in Early-B and Early-T cells	
Ankrd33b	Klf13
Apba1	Loxl1
Aqp7	Maf
Arhgef17	Map3k9
Arid3b	Mfsd9
Arrb1	Mthfr
Atg4d	Neurl1a
Atmin	Pygo1
Bmf	Rassf3
Braf	Rbm20
Clmn	Rgma
Cpeb3	Rhoq
Ctdspl	Rps6ka1
Fam168a	Rps6ka2
Flot2	Rusc2
Fndc3b	Sarm1
Foxk1	Slc35e1
Foxn2	Slc7a11
Fut1	Sp7
Gas2l1	Speg
Gcnt2	St3gal2
Ggh	Sult4a1
Gpd1	Syt7
Gsn	Tet2
Hck	Tifab
Hipk3	Tmem170b
Il10ra	Trim65
Il16	Ube2w
Ino80d	Vangl1
Itga8	Ypel2
Josd1	Zbtb34

CHAPTER-7: BIBLIOGRAPHY

Banerjee, A., Northrup, D., Boukarabila, H., Jacobsen, S.E.W., and Allman, D. (2013). Transcriptional Repression of Gata3 Is Essential for Early B-cell Commitment. Immunity *38*, 930–942.

Barski, A., Cuddapah, S., Cui, K., Roh, T., Schones, D.E., and Wang, Z. (2007). High-resolution profiling of histone methylations in the human genome. Cell *829*, 823–837.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233.

Batard, P., Monier, M.N., Fortunel, N., Ducos, K., SansilvestriMorel, P., Phan, T., Hatzfeld, A. and Hatzfeld, J.A. (2000). TGF-(beta)1 maintains hematopoietic immaturity by a reversible negative control of cell cycle and induces CD34 antigen up-modulation. J.Cell. Sci., *113*, 383 – 390.

Bertrand, J. Y., Kim, A. D., Violette, E. P., Stachura, D. L., Cisson, J. L. and Traver D. (2007). Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. Development *134*, 4147- 4156.

Bhaskaran, V., Nowicki, M.O., Idriss, M., Jimenez, M.A., Lugli, G., Hayes, J.L., Mahmoud, A.B., Zane, R.E., Passaro, C., Ligon, K.L., et al. (2019). The functional synergism of microRNA clustering provides therapeutically relevant epigenetic interference in glioblastoma. Nat. Commun. *10*, 1–13.

Böiers, C., Carrelha, J., Lutteropp, M., Luc, S., Green, J. C. A., Azzoni, E., Woll, P. S., Mead, A. J., Hultquist, A., Swiers, G. et al. (2013). Lympho-myeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. Cell Stem Cell *13*, 535-548.

Boller, S., and Grosschedl, R. (2014). The regulatory network of B-cell differentiation: A focused view of early B-cell factor 1 function. Immunol. Rev. *261*, 102–115.

Boller, S., Ramamoorthy, S., Akbas, D., Nechanitzky, R., Burger, L., Murr, R., Schübeler, D., and Grosschedl, R. (2016). Pioneering activity of the C-Terminal domain of EBF1 shapes the chromatin landscape for B cell programming. Immunity *44*, 527–541.

Brancati, G., and Großhans, H. (2018). An interplay of miRNA abundance and target site architecture determines miRNA activity and specificity. Nucleic Acids Res. *46*, 3259–3269.

Brandl, A., Daum, P., Brenner, S., Schulz, S.R., Yap, D.Y.H., Bösl, M.R., Wittmann, J., Schuh, W., and Jäck, H.M. (2016). The microprocessor component,

DGCR8, is essential for early B-cell development in mice. Eur. J. Immunol. *46*, 2710–2718.

Brown, G., and Ceredig, R. (2019). Modeling the hematopoietic landscape. Front. Cell Dev. Biol. *7*, 1–13.

Chandradoss, S.D., Schirle, N.T., Szczepaniak, M., Macrae, I.J., Chandradoss, S.D., Schirle, N.T., Szczepaniak, M., Macrae, I.J., and Joo, C. (2015). A Dynamic Search Process. Cell *162*, 96–107.

Chen, C., Ridzon, D.A., Broomer, A.J., Zhou, Z., Lee, D.H., Nguyen, J.T., Barbisin, M., Xu, N.L., Mahuvakar, V.R., Andersen, M.R., et al. (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res. *33*, 1–9.

Chen, C.Z., Li, L., Lodish, H.F., Bartel, D.P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. Science *303*, 83–86.

Cobb, B.S., Nesterova, T.B., Thompson, E., Hertweck, A., O'Connor, E., Godwin, J., Wilson, C.B., Brockdorff, N., Fisher, A.G., Smale, S.T., et al. (2005). T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer. J. Exp. Med. *201*, 1367–1373.

Corcoran, D.L., Pandit, K. V, Gordon, B., Bhattacharjee, A., and Kaminski, N. (2009). Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. Plos one *4(4)*, 1-10.

Cursons, J., Pillman, K.A., Scheer, K.G., Gregory, P.A., Foroutan, M., Hediyeh-Zadeh, S., Toubia, J., Crampin, E.J., Goodall, G.J., Bracken, C.P., et al. (2018). Combinatorial Targeting by MicroRNAs Co-ordinates Post-transcriptional Control of EMT. Cell Syst. *7*, 77-91.

Daum, P., Meinzinger, J., Pracht, K., Côrte-Real, J., Wittmann, J., and Jäck, H.-M. (2017). The microRNA processing subunit DGCR8 is required for immune response and germinal center formation. J. Immunol. *198*, 195.22 LP-195.22.

DeKoter, R.P., and Singh, H. (2000). Regulation of B lymphocyte and macrophage development by graded expression of PU.1. Science *288*, 1439–1442.

DeKoter, R.P., and Singh, H. (2000). Regulation of B lymphocyte and macrophage development by graded expression of PU.1. Science *288*, 1439–1442.

Doyle, F., and Tenenbaum, S.A. (2014). Trans-regulation of RNA-binding protein motifs by microRNA. Front. Genet. *5*, 1–10.

Duchaine, T.F., and Fabian, M.R. (2018). Mechanistic Insights into MicroRNA-Mediated Gene Silencing. Cold Spring Harb. Perspect. Biol. doi: 10.1101/cshperspect.a032771, 1–22.

Fiedler, K., Kokai, E., Bresch, E., and Brunner, C. (2013). MyD88 is involved in myeloid as well as lymphoid hematopoiesis independent of the presence of a pathogen. Am. J. Blood Res. *3(2)*, 124-140.

Fontana, L., Pelosi, E., Greco, P., Racanicchi, S., Testa, U., Liuzzi, F., Croce, C.M., Brunetti, E., Grignani, F., and Peschle, C. (2007). MicroRNAs 17-5p-20a-106a control monocytopoiesis through AML1 targeting and M-CSF receptor upregulation. Nat. Cell Biol. *9*, 775–787.

Gemelli, C., Orlandi, C., Zanocco, T., Martello, A., Vignudelli, T., Montanari, M., Parenti, S., Testa, A., Grande, A., Ferrari, S., et al. (2008). The Vitamin D 3 /Hox-A10 pathway supports MafB Function during the monocyte differentiation of human CD34+ hemopoietic progenitors. J Immunol. *181*, 5660-5672.

Ghosh, J., Chan, R.J., Kapur, R., Ghosh, J., Kobayashi, M., Ramdas, B., Chatterjee, A., and Ma, P. (2016). S6K1 regulates hematopoietic stem cell self-renewal and leukemia maintenance. J Clin Invest. *126(7)*, 2621-2625.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: Scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*, 835–840.

Hafezqorani, S., Lafzi, A., De Bruin, R.G., Van Zonneveld, A.J., Van Der Veer, E.P., Son, Y.A., and Kazan, H. (2016). Modeling the combined effect of RNA-binding proteins and microRNAs in post-transcriptional regulation. Nucleic Acids Res. *44*, 1–16.

Hagman, J., and Lukin, K. (2005). Early B-cell factor “pioneers” the way for B-cell development. Trends Immunol. *26*, 455–461.

Hegde, B.S.P., Zhao, J., Ashmun, R.A., and Shapiro, L.H. (1999). c-Maf induces monocytic differentiation and apoptosis in bipotent myeloid progenitors. Blood *94*, 1578–1589.

Hegde, S.P., Kumar, A., Kurschner, C., and Shapiro, L.H. (1998). c-Maf interacts with c-Myb to regulate transcription of an early myeloid gene during differentiation. Mol. Cell. Biol. *18*, 2729–2737.

Hu, G., Cui, K., Fang, D., Hirose, S., Wang, X., Wangsa, D., Jin, W., Ried, T., Liu, P., Zhu, J., et al. (2018). Transformation of accessible chromatin and 3D nucleome underlies lineage commitment of early T-cells. Immunity *48*, 227-242.e8.

Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., and Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. Genes Dev. *11*, 774–785.

Ikawa, T., Hirose, S., Masuda, K., Kakugawa, K., Satoh, R., Shibano-Satoh, A., Kominami, R., Katsura, Y., and Kawamoto, H. (2010). An essential developmental checkpoint for production of the T cell lineage. Science *329*, 93–96.

Ikawa, T., Kawamoto, H., Fujimoto, S. and Katsura, Y. (1999) Commitment of common T/Natural killer (NK) progenitors to unipotent T and NK progenitors in the murine fetal thymus revealed by a single progenitor assay. J Exp Med. 190(11), 1617-1626.

Inui, M., Martello, G., and Piccolo, S. (2010). MicroRNA control of signal transduction. Nat. Rev. Mol. Cell Biol. *11*, 252–263.

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 intra-embryonic aortagonad-mesonephros region. J. Exp. Med. *208*, 2417–2427.

Izon, D., Rudd, K., DeMuth, W., Pear, W.S., Clendenin, C., Lindsley, R.C., and Allman, D. (2001). A Common pathway for Dendritic Cell and Early B cell development. J. Immunol. *167*, 1387–1392.

Jacobsen, A., Wen, J., Marks, D.S., and Krogh, A. (2010). Signatures of RNA binding proteins globally coupled to effective microRNA target sites. Genome Res. *20*, 1010–1019.

Javierre, B.M., Sewitz, S., Cairns, J., Wingett, S.W., Várnai, C., Thiecke, M.J., Freire-Pritchett, P., Spivakov, M., Fraser, P., Burren, O.S., et al. (2016). Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. Cell *167*, 1369-1384.e19.

Jin, J., Britschgi, A., Schläfli, A.M., Humbert, M., Shan-krauer, D., Batliner, J., Federzoni, E.A., Ernst, M., Torbett, B.E., Yousefi, S., et al. (2018). Low Autophagy (ATG) gene expression is associated with an immature aml blast cell phenotype and can be restored during aml differentiation Therapy. Oxid. Med. Cell. Longev, 1-16, doi.org/10.1155/2018/1482795.

Karamitros, D., Stoilova, B., Aboukhalil, Z., Hamey, F., Reinisch, A., Samitsch, M., Quek, L., Otto, G., Repapi, E., Doondeea, J., et al. (2018). Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells article. Nat. Immunol. *19*, 85–97.

Kent, D.G., Copley, M.R., Benz, C., Wöhrer, S., Dykstra, B.J., Ma, E., Cheyne, J., Zhao, Y., Bowie, M.B., Zhao, Y., et al. (2009). Prospective isolation and molecular

characterization of hematopoietic stem cells with durable self-renewal potential. Blood *113*, 6342–6350.

Khalaj, M., Tavakkoli, M., Stranahan, A.W., and Park, C.Y. (2014). Pathogenic MicroRNA's in myeloid malignancies. Front. Genet. *5*, 1–18.

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

Koralov, S.B., Muljo, S.A., Galler, G.R., Krek, A., Chakraborty, T., Kanellopoulou, C., Jensen, K., Cobb, B.S., Merkenschlager, M., Rajewsky, N., et al. (2008). Dicer ablation affects antibody diversity and cell survival in the B lymphocyte lineage. Cell *132*, 860–874.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693–705.Broughton, J.P., Lovci, M.T., Huang, J.L., Yeo, G.W., and Pasquinelli, A.E. (2016). Pairing beyond the seed supports microRNA targeting specificity. Mol. Cell *64*, 320–333.

Kuchen, S., Resch, W., Yamane, A., Kuo, N., Li, Z., Chakraborty, T., Wei, L., Laurence, A., Yasuda, T., Peng, S., et al. (2010).

Kurkewich, J.L., Hansen, J., Klopfenstein, N., Zhang, H., Wood, C., Boucher, A., Hickman, J., Muench, D.E., Grimes, L., and Dahl, R. (2017). The miR-23a ~ 27a ~ 24-2 microRNA cluster buffers transcription and signaling pathways during hematopoiesis. PLoS Genet. *13*(7), 1–28.

Lai, A.Y., and Kondo, M. (2008). T and B lymphocyte differentiation from hematopoietic stem cell. Semin. Immunol. *20*, 207–212.

Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., et al. (2007). A Mammalian microRNA expression atlas based on small RNA library sequencing. Cell *129*, 1401–1414.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, 1-10.

Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretsky, I.,Jaitin, D.A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S., Friedman, N., and Amit, I. (2004). Chromatin state dynamics during blood formation. Science. *345*, 943-949.

Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. Nature *553*, 418–426.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15–20.

Ley, K., Pramod, A.K., Croft, M., Ravichandran K.S., Ting, J.P., (2016). How Mouse Macrophages Sense what is going on? Front. Immunol. *7*, 1–17.

Li, L., Leid, M., and Rothenberg, E. V (2010). An early T-cell lineage commitment Checkpoint dependent on the transcription factor Bcl11b. Science. *329*, 89-93.

Li, Q.J., Chau, J., Ebert, P.J.R., Sylvester, G., Min, H., Liu, G., Braich, R., Manoharan, M., Soutschek, J., Skare, P., et al. (2007). miR-181a Is an Intrinsic Modulator of T Cell Sensitivity and Selection. Cell *129*, 147–161.

Lin, H., Grosschedl, R. (1995). Failure of B-cell differentiation in mice lacking the transcription factor EBF. Nature *376*, 263–267.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol. *15*, 550.

Magilnick, N., and Boldin, M.P. (2018). Molecular Moirai: Long noncoding RNAs - mediators of HSC fate. Curr. Stem Cell Reports *4*, 158–165.

Manz, M.G., Traver, D., Miyamoto, T., Weissman, I.L., and Akashi, K. (2001). Dendritic cell potentials of early lymphoid and myeloid progenitors. Blood *97*, 3333–3341.

Marsico, A., Huska, M.R., Lasserre, J., Hu, H., Vucicevic, D., Musahl, A., Orom, U.A., and Vingron, M. (2013). PROmiRNA : a new miRNA promoter recognition method uncovers the complex regulation of intronic miRNAs. Genome Biology *4,* R84.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521–533.

McGrath, K. E., Frame, J. M., Fegan, K. H., Bowen, J. R., Conway, S. J., Catherman, S. C., Kingsley, P. D., Koniski, A. D. and Palis, J. (2015). Distinct sources of hematopoietic progenitors emerge before HSCs and provide functional blood cells in the mammalian embryo. Cell Rep. *11*, 1892-1904.

McKercher, S.R., Torbett, B.E., Anderson, K.L., Henkel, G.W., Vestal, D.J., Baribault, H., Klemsz, M., Feeney, A.J., Wu, G.E., Paige, C.J., et al. (1996). Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. EMBO J. *15*, 5647–5658.

Medina, K.L., Pongubala, J.M.R., Reddy, K.L., Lancki, D.W., DeKoter, R., Kieslinger, M., Grosschedl, R., and Singh, H. (2004). Assembling a gene regulatory network for specification of the B cell fate. Dev. Cell *7*, 607–617.

Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P.D. (2019). PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. *47*, D419–D426.

Min, K.W., Jo, M.H., Shin, S., Davila, S., Zealy, R.W., Kang, S.I., Lloyd, L.T., Hohng, S., and Yoon, J.H. (2017). AUF1 facilitates microRNA-mediated gene silencing. Nucleic Acids Res. *45*, 6064–6073.

Miyamoto, T., Iwasaki, H., Reizis, B., Ye, M., Graf, T., Weissman, I.L., and Akashi, K. (2002). Myeloid or Lymphoid Promiscuity as a Critical Step in Hematopoietic Lineage Commitment. *3*, 137–147.

Miyazaki, K., Miyazaki, M., and Murre, C. (2014). The establishment of B versus T cell identity. Trends Immunol. *35*, 205–210.

Monticelli, S., Ansel, K.M., Xiao, C., Socci, N.D., Krichevsky, A.M., Thai, T.H., Rajewsky, N., Marks, D.S., Sander, C., Rajewsky, K., et al. (2005). MicroRNA profiling of the murine hematopoietic system. Genome Biol. *6*.R71.

Moore, M.J., Scheel, T.K.H., Luna, J.M., Park, C.Y., Fak, J.J., Nishiuchi, E., Rice, C.M., and Darnell, R.B. (2015). miRNA-target chimeras reveal miRNA 3′-end pairing as a major determinant of Argonaute target specificity. Nat. Commun. *6, 1-16.*

Muljo, S.A., Mark Ansel, K., Kanellopoulou, C., Livingston, D.M., Rao, A., and Rajewsky, K. (2005). Aberrant T cell differentiation in the absence of Dicer. J. Exp. Med. *202*, 261–269.

Muller-Sieburg, C.E., Cho, R.H., Karlsson, L., Huang, J.F., and Sieburg, H.B. (2004). Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. Blood *103*, 4111–4118.

Nechanitzky, R., Akbas, D., Scherer, S., Györy, I., Hoyler, T., Ramamoorthy, S., Diefenbach, A., and Grosschedl, R. (2013). Transcription factor EBF1 is essential for the maintenance of B cell identity and prevention of alternative fates in committed cells. Nat. Immunol. *14*, 867–875.

Neilson, J.R., Zheng, G.X.Y., Burge, C.B., and Sharp, P.A. (2007). Dynamic regulation of miRNA expression in ordered stages of cellular development. Genes Dev. *21*, 578–589.

Ng, A.P., and Alexander, W.S. (2017). Haematopoietic stem cells: Past, present and future. Cell Death Discov. *3*, 2–5.

Ohtsuka, M., Arase, H., Takeuchi, A., Yamasaki, S., Shiina, R., Suenaga, T., and Saito, T. (2004). NFAM1, an immunoreceptor tyrosine-basedactivation motif-bearing molecule that regulates B cell development and signaling. Proc. Natl. Acad. Sci. U.S.A. *101*, 1–6.

Ooi, A.G.L., Sahoo, D., Adorno, M., Wang, Y., Weissman, I.L., and Park, C.Y. (2010). MicroRNA-125b expands hematopoietic stem cells and enriches for the lymphoid-balanced and lymphoid-biased subsets. Proc. Natl. Acad. Sci. U. S. A. *107*, 21505–21510.

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

Pai, S.Y., Truitt, M.L., Ting, C.N., Leiden, J.M., Glimcher, L.H., and Ho, I.C. (2003). Critical Roles for Transcription Factor GATA-3 in Thymocyte Development. Immunity *19*, 863–875.

Peruta, M. Della, Martinelli, G., Moratti, E., Pintani, D., Vezzalini, M., Mafficini, A., Grafone, T., Iacobucci, I., Soverini, S., Murineddu, M., et al. (2010). Protein Tyrosine Phosphatase Receptor Type γ is a functional tumor suppressor gene specifically downregulated in Chronic Myeloid Leukemia. Cancer Res. *70(21)*, 8896–8907.

Petriv, O.I., Kuchenbauer, F., Delaney, A.D., Lecault, V., White, A., Kent, D., Marmolejo, L., Heuser, M., Berg, T., Copley, M., et al. (2010). Comprehensive microRNA expression profiling of the hematopoietic hierarchy. Proc. Natl. Acad. Sci. U. S. A. *107*, 15443–15448.

Pongubala, J.M.R., Northrup, D.L., Lancki, D.W., Medina, K.L., Treiber, T., Bertolino, E., Thomas, M., Grosschedl, R., Allman, D., and Singh, H. (2008). Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. Nat. Immunol. *9*, 203–215.

Pons-Espinal, M., de Luca, E., Marzi, M.J., Beckervordersandforth, R., Armirotti, A., Nicassio, F., Fabel, K., Kempermann, G., and De Pietri Tonelli, D. (2017). Synergic Functions of miRNAs Determine Neuronal Fate of Adult Neural Stem Cells. Stem Cell Reports *8*, 1046–1061.

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., et al. (1999). Notch1 expression in early lymphopoiesis influences B versus T lineage determination. Immunity *11*, 299–308.

Querfurth, E., Schuster, M., Kulessa, H., Crispino, J. D., Döderlein, G., Orkin, S. H., Graf, T., & Nerlov, C. (2000). Antagonism between C/EBPbeta and FOG in eosinophil lineage commitment of multipotent hematopoietic progenitors. Genes dev. *14*(19), 2515–2525.

Radtke, F., Wilson, A., Stark, G., Bauer, M., Van Meerwijk, J., MacDonald, H.R., and Aguet, M. (1999). Deficient T cell fate specification in mice with an induced inactivation of Notch1. Immunity *10*, 547–558.

Rajasekhar, M., Schmitz, U., Flamant, S., Wong, J.J.L., Bailey, C.G., Ritchie, W., Holst, J., and Rasko, J.E.J. (2018). Identifying microRNA determinants of human myelopoiesis. Sci. Rep. *8*, 1–15.

Reynaud, D., Demarco, I.A., Reddy, K.L., Schjerven, H., Bertolino, E., Chen, Z., Smale, S.T., Winandy, S., and Singh, H. (2008). Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. Nat. Immunol. *9*, 927–936.

Rieger, M. A, Hoppe, P.S., Smejkal, B.M., Eitelhuber, A.C., and Schroeder, T. (2009). Haematopoietic cytokines instruct lineage Choice. Science. *325*, 217–218.

Rieger, M.A., and Schroeder, T. (2012). Hematopoiesis. Cold Spring Harb. Perspect. Biol. *4*,1–18.

Rothenberg, E. V (2010). Multilayered specification of the T-cell lineage fate. Immunol. Rev. *238*, 150–168.

Rothenberg, E. V (2011). T-Cell Lineage Commitment : Identity and renunciation. J Immunol. *186*, 6649-6655.

Rothenberg, E. V., Champhekar, A., Damle, S., Real, M.M. Del, Kueh, H.Y., Li, L., and Yui, M.A. (2013). Transcriptional establishment of cell-type identity: Dynamics and causal mechanisms of T-cell lineage commitment. Cold Spring Harb. Symp. Quant. Biol. *78*, 31–41.

Rothenberg, E. V., Moore, J.E., and Yui, M.A. (2008). Launching the T-cell-lineage developmental programme. Nat. Rev. Immunol. *8*, 9–21.

Rumfelt, L.L., Zhou, Y., Rowley, B.M., Shinton, S.A., and Hardy, R.R. (2006). Lineage specifi cation and plasticity in CD19− early B cell precursors. J. Exp. Med. *203*, 675–687.

Sambandam, A., Maillard, I., Zediak, V.P., Xu, L., Gerstein, R.M., Aster, J.C., Pear, W.S., and Bhandoola, A. (2005). Notch signaling controls the generation and differentiation of early T lineage progenitors. Nat. Immunol. *6*, 663–670.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C.T., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. Nat. letters *419*, 407–411.

Schmitt, T.M. & Zuniga-Pflucker, J.C. Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro. Immunity 17, 749–756 (2002).

Scripture-Adams, D.D., Damle, S.S., Li, L., Elihu, K.J., Qin, S., Arias, A.M., Butler, R.R., Champhekar, A., Zhang, J.A., and Rothenberg, E. V. (2014). GATA-3 dose-dependent checkpoints in Early T cell Commitment. J. Immunol. *193*, 3470–3491.

Singh, H., Medina, K.L., and Pongubala, J.M.R. (2005). Contingent gene regulatory networks and B cell fate specification. Proc. Natl. Acad. Sci. U. S. A. *102*, 4949–4953.

Su, X., Qian, C., Zhang, Q., Hou, J., Gu, Y., Han, Y., and Chen, Y. (2013). miRNomes of haematopoietic stem cells and dendritic cells identify miR-30b as a regulator of Notch1. Nat. Commun. *4*, 2903.

Treiber, T., Treiber, N., and Meister, G. (2019). Regulation of microRNA biogenesis
and its crosstalk with other cellular pathways Nat. Rev. Mol. Cell Biol. *20,* 5-20.

Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C., Weiss, M.J., Crossley, M., and Orkin, S.H. (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. Cell *90*, 109–119.

Ungerbäck, J., Hosokawa, H., Wang, X., Strid, T., Williams, B.A., Sigvardsson, M., and Rothenberg, E. V. (2018). Pioneering, chromatin remodeling, and epigenetic constraint in early T-cell gene regulation by SPI1 (PU.1). Genome Res*. 28*, 1508–1519.

Van Schoonhoven, A., Huylebroeck, D., Hendriks, R.W., and Stadhouders, R. (2019). 3D genome organization during lymphocyte development and activation. Brief. Funct. Genomics, 1–12.

Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D., et al. (2017). Human haematopoietic stem cell lineage commitment is a continuous process. Nat. Cell Biol. *19*, 271–281.

Wada, H., Masuda, K., Satoh, R., Kakugawa, K., Ikawa, T., Katsura, Y., and Kawamoto, H. (2008). Adult T-cell progenitors retain myeloid potential. Nat. Letters *452*, 768–773.

Wallner, S., Grandl, M., Konovalova, T., Sigru, A., Kopf, T., Peer, M., Orso, E., Liebisch, G., and Schmitz, G. (2014). Monocyte to macrophage differentiation goes along with modulation of the plasmalogen pattern through transcriptional regulation. PLoS ONE 9(4): e94102. doi:10.1371/journal.pone.0094102.

Weber, B.N., Chi, A.W.S., Chavez, A., Yashiro-Ohtani, Y., Yang, Q., Shestova, O., and Bhandoola, A. (2011). A critical role for TCF-1 in T-lineage specification and differentiation. Nature *476*, 63–69.

Willman, B.C.L., Stewart, C.C., Longacre, T.L., Head, D.R., Habbersett, R., Ziegler, S.F., and Perlmutter, R.M. (1991). Expression of the c-fgr and hck protein-tyrosine kinases in Acute Myeloid Leukemic blasts is associated with early commitment and differentiation events in the monocytic and granulocytic lineages Blood 77(4), 726–734.

Wilson, A., MacDonald, H.R., and Radtke, F. (2001). Notch 1-deficient common lymphoid precursors adopt a B cell fate in the thymus. J. Exp. Med. *194*, 1003–1012.

Xu, S., Shu, A., Pang, Y., Zhuo, Y., Huo, J., and Guo, K. (2015). The RNase III enzyme Dicer is essential for germinal center B cell formation during T cell-dependent immune response. Technology *119*, 767–777.

Yasugi, E., Horiuchi, A., Uemura, I., Okuma, E., Nakatsu, M., Saeki, K., Kamisaka, Y., Kagechika, H., Yasuda, K., and Yuo, A. (2006). Peroxisome proliferator-activated receptor γ ligands stimulate myeloid differentiation and lipogenensis in human leukemia NB4 cells. Dev. Growth Differ. *48*, 177–188.

Zandi, S., Åhsberg, J., Tsapogas, P., Stjernberg, J., Qian, H., and Sigvardsson, M. (2012). Single-cell analysis of early B-lymphocyte development suggests independent regulation of lineage specification and commitment in vivo. Proc. Natl. Acad. Sci. U. S. A. *109*, 15871–15876.

Zandi, S., Mansson, R., Tsapogas, P., Zetterblad, J., Bryder, D., and Sigvardsson, M. (2008). EBF1 is essential for B-lineage priming and establishment of a transcription factor network in common lymphoid progenitors. J. Immunol. *181*, 3364–3372.

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., et al. (2009). Patterns of microRNA expression characterize stages of human B-cell differentiation. Blood *113*, 4586–4594.

Zheng, Y., Zhang, H., Zhang, X., Feng, D., Luo, X., Zeng, C., Lin, K., Zhou, H., Qu, L., Zhang, P., et al. (2012). MiR-100 regulates cell differentiation and survival by

targeting RBSP3, a phosphatase-like tumor suppressor in acute myeloid leukemia. Oncogene *31*, 80–92.

Zhu, J., and Emerson, S.G. (2002). Hematopoietic cytokines, transcription factors and lineage commitment. Oncogene *21*, 3295–3313.

CHAPTER-8: APPENDIX

8.1. Publications

1. Boya R., Yadavalli A.D., **Nikhat S.**, Kurukuti S., Palakodeti D., Pongubala J.M.R. Developmentally regulated higher-order chromatin interactions orchestrate B cell fate commitment. Nucleic Acids Res. 2017 *45*(19), 11070-11087.

2. **Nikhat S.**, Yadavalli A.D., Prusty A., Palakodeti D., Pongubala J.M.R. A regulatory network of microRNAs confers lineage commitment during early developmental trajectories of B and T lymphocytes. (Manuscript submitted).

8.2. Plagiarism report

Attached on Next page

Deciphering the regulatory role of microRNAs during early developmental stages of B and T lymphocytes

ORIGINALITY REPORT

8%	4%	8%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Isabel Alcobia, Ana Sofia Quina, Hélia Neves, Nuno Clode, Leonor Parreira. "The spatial organization of centromeric heterochromatin during normal human lymphopoiesis: evidence for ontogenically determined spatial patterns", Experimental Cell Research, 2003 Publication	1%
2	"Encyclopedia of Signaling Molecules", Springer Science and Business Media LLC, 2018 Publication	1%
3	Boller, Sören, and Rudolf Grosschedl. "The regulatory network of B-cell differentiation: a focused view of early B-cell factor 1 function", Immunological Reviews, 2014. Publication	1%
4	www.nature.com Internet Source	<1%
5	M. A. Rieger, T. Schroeder. "Hematopoiesis", Cold Spring Harbor Perspectives in Biology,	<1%

2012
Publication

6	Sijia Shen, Yuxin Lin, Xuye Yuan, Li Shen, Jiajia Chen, Luonan Chen, Lei Qin, Bairong Shen. "Biomarker MicroRNAs for Diagnosis, Prognosis and Treatment of Hepatocellular Carcinoma: A Functional Survey and Comparison", Scientific Reports, 2016 Publication	<1%
7	Luca F. R. Gebert, Ian J. MacRae. "Regulation of microRNA function in animals", Nature Reviews Molecular Cell Biology, 2018 Publication	<1%
8	Ellen V. Rothenberg, Jonathan E. Moore, Mary A. Yui. "Launching the T-cell-lineage developmental programme", Nature Reviews Immunology, 2008 Publication	<1%
9	Hui Cheng, Zhaofeng Zheng, Tao Cheng. "New paradigms on hematopoietic stem cell differentiation", Protein & Cell, 2019 Publication	<1%
10	rsob.royalsocietypublishing.org Internet Source	<1%
11	support.bioconductor.org Internet Source	<1%

12	Submitted to uhasselt Student Paper	<1%
13	Alexandra Maslova, Ricardo N. Ramirez, Ke Ma, Hugo Schmutz et al. "Learning immune cell differentiation", Cold Spring Harbor Laboratory, 2019 Publication	<1%
14	diposit.ub.edu Internet Source	<1%
15	Andrejs Ivanovs, Stanislav Rybtsov, Elizabeth S. Ng, Edouard G. Stanley, Andrew G. Elefanty, Alexander Medvinsky. "Human haematopoietic stem cell development: from the embryo to the dish", Development, 2017 Publication	<1%
16	genesdev.cshlp.org Internet Source	<1%
17	repository.cshl.edu Internet Source	<1%
18	www.scribd.com Internet Source	<1%
19	www.mdpi.com Internet Source	<1%
20	Ivan Maillard, Terry Fang, Warren S. Pear. "REGULATION OF LYMPHOID	<1%

DEVELOPMENT, DIFFERENTIATION, AND FUNCTION BY THE NOTCH PATHWAY", Annual Review of Immunology, 2005
Publication

---

21 hss.ulb.uni-bonn.de <1%
Internet Source

---

22 Meinrad Busslinger. "Transcriptional Control of Early B Cell Development", Annual Review of Immunology, 2004 <1%
Publication

---

23 Xing, Yujiao, Jianhua Fu, Haiping Yang, Li Yao, Lin Qiao, Yanna Du, and Xindong Xue. "MicroRNA expression profiles and target prediction in neonatal Wistar rat lungs during the development of bronchopulmonary dysplasia", International Journal of Molecular Medicine, 2015. <1%
Publication

---

24 Claudia Ribeiro de Almeida, Rudi W. Hendriks, Ralph Stadhouders. "Dynamic Control of Long-Range Genomic Interactions at the Immunoglobulin κ Light-Chain Locus", Elsevier BV, 2015 <1%
Publication

---

25 Saïd Sif. "ATP-dependent nucleosome remodeling complexes: Enzymes tailored to deal with chromatin", Journal of Cellular <1%

	Biochemistry, 04/15/2004 Publication	
26	Ellen V. Rothenberg. "T Cell Lineage Commitment: Identity and Renunciation", The Journal of Immunology, 2011 Publication	<1%
27	publications.aston.ac.uk Internet Source	<1%
28	hdl.handle.net Internet Source	<1%
29	archiv.ub.uni-heidelberg.de Internet Source	<1%
30	bmcneurol.biomedcentral.com Internet Source	<1%
31	www.embopress.org Internet Source	<1%
32	documents.mx Internet Source	<1%
33	www.frontiersin.org Internet Source	<1%
34	bmcgenet.biomedcentral.com Internet Source	<1%
35	Song, Jinsoo, Dongkyun Kim, Churl-Hong Chun, and Eun-Jung Jin. "miR-370 and miR-373	<1%

regulate the pathogenesis of osteoarthritis by modulating one-carbon metabolism via SHMT-2 and MECP-2, respectively", Aging Cell, 2015.
Publication

36	Kyoko Ito, Massimo Bonora, Keisuke Ito. "Metabolism as master of hematopoietic stem cell fate", International Journal of Hematology, 2018 Publication	<1%
37	medcraveonline.com Internet Source	<1%
38	Hoffmann, R.. "A genomic view of lymphocyte development", Current Opinion in Immunology, 200306 Publication	<1%
39	Advances in Experimental Medicine and Biology, 2015. Publication	<1%
40	"Developmental Vascular Biology Workshop III", Angiogenesis, 2008 Publication	<1%
41	Wang, Minghua, Weiping Wang, Ping Zhang, Juanjuan Xiao, Jianguo Wang, and Chaoqun Huang. "Discrimination of the Expression of Paralogous microRNA Precursors That Share the Same Major Mature Form", PLoS ONE, 2014.	<1%

	Publication	
42	opus.bibliothek.uni-wuerzburg.de Internet Source	<1%
43	Casey G. Kowalik, Drew A. Palmer, Travis B. Sullivan, Patrick A. Teebagy et al. "Profiling microRNA from nephrectomy and biopsy specimens: predictors of progression and survival in clear cell renal cell carcinoma", BJU International, 2017 Publication	<1%
44	bmcplantbiol.biomedcentral.com Internet Source	<1%
45	prod--journal.elifesciences.org Internet Source	<1%
46	www.future-science.com Internet Source	<1%

Exclude quotes	On	Exclude matches	< 14 words
Exclude bibliography	On		