

Regulation and Derailment of an Innate-like T Cell Thymic Developmental Pathway

by

Sumedha Roy

Department of Immunology
Duke University

Date: _____

Approved:

Yuan Zhuang, Supervisor

Michael S. Krangel, Committee Chair

Qi-Jing Li

Maria Ciofani

Bernard Mathey-Prevot

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Immunology in the Graduate School
of Duke University

2018

ABSTRACT

Regulation and Derailment of an Innate-like T Cell Thymic Developmental Pathway

by

Sumedha Roy

Department of Immunology
Duke University

Date: _____

Approved:

Yuan Zhuang, Supervisor

Michael S. Krangel, Committee Chair

Qi-Jing Li

Maria Ciofani

Bernard Mathey-Prevot

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Immunology in the Graduate School of
Duke University

2018

Copyright by
Sumedha Roy
2018

Abstract

Invariant Natural Killer T (iNKT) and $\gamma\delta$ NKT cells are well-characterized innate-like counterparts of $\alpha\beta$ and $\gamma\delta$ T cells respectively that express semi-invariant T cell receptors (TCRs) and are capable of mounting rapid immune responses. Although many key regulatory molecules have been shown to play important roles in the development of these cells, the mechanism of their lineage specification and acquisition of effector functions remain to be fully addressed.

Id proteins, or inhibitor of DNA binding and differentiation, act as antagonists of transcription factors known as E proteins. Id proteins are known to promote the differentiation of conventional T cells, and suppress the expansion of innate-like T cells. We have previously found that expansion of iNKT and another subset of innate-like T cells leads to rapid lymphoma development in Id2/Id3-deficient mice. The goal of this dissertation is to elucidate the mechanisms by which Id proteins differentially regulate the lineage choice between the concurrently developing innate and conventional lineages in early stages of T cell development, as well as the mechanisms driving the malignant transformation of these expanding innate-like T cells.

Firstly, I tested the hypothesis whether uninhibited E2A activity in the absence of Id proteins transcriptionally promotes the development of iNKT cells. Indeed, I found E2A-mediated upregulation of critical genes, and biased rearrangement at the DP stage promotes iNKT cell lineage development in Id-deficient mice. The observed expansion

of the iNKT cells in these mice is not abrogated by blocking pre-TCR signaling, which is required for conventional $\alpha\beta$ T cell development. Finally, E2A is found to be a key transcriptional regulator of both iNKT and $\gamma\delta$ NKT lineages, which appear to have shared lineage history. Therefore, my study revealed a previously unappreciated role of E2A in the regulation of lineage choice between conventional and innate-like T cell fate as early as the pre-TCR checkpoint.

Second, I explored the origins and pathways that drive innate-like lymphomas in Id2/Id3-deficient mice. I found that CD1dTet⁺ innate-like T cells develop independent of CD1d-mediated selection, and start expansion in neonatal mice. The transcriptional program in expanding neonatal iNKT cells is significantly modified, including upregulation of the cytokine-cytokine receptor interaction pathway which can promote their expansion and migration, ultimately leading to their malignant transformation. I also discovered shared dysregulation of the NF- κ B pathway and genes with reported driver mutations between our iNKT-derived lymphomas and human iNKT tumors. My study demonstrates that Id2 plays a tumor suppressive role in collaboration with Id3 in developing T cells in mice. Contrary to the perception of Id proteins as potential therapeutic targets in some cancer models, these results also highlight the possibility of aggravated tumorigenesis upon non-targeted suppression of Id2 and Id3.

Dedication

To all my family and friends who were my beacon of inspiration and support in this journey.

Contents

Abstract	iv
Dedication	vi
List of Tables	xi
List of Figures	xii
1. Introduction	1
1.1 Conventional T cell development in the thymus	2
1.1.1 Commitment to the T cell lineage	4
1.1.2 Generating a diverse repertoire	5
1.1.2.1 TCR β , TCR γ/δ rearrangement at the DN stage	6
1.1.2.2 TCR α rearrangement at the DP stage	8
1.1.3 Selection checkpoints for appropriate T cell development	9
1.1.3.1 Pre-TCR selection at the DN stage	9
1.1.3.2 TCR-mediated selection at DP stage	10
1.2 Innate-like T cells	11
1.2.1 Innate-like $\alpha\beta$ invariant natural killer T cells	11
1.2.2 Innate-like $\gamma\delta$ natural killer T cells	13
1.2.3 Innate-like T cell lineage commitment	14
1.2.4 Transcriptional Regulation of iNKT and $\gamma\delta$ NKT lineages	18
1.2.5 Regulation of iNKT subsets	19
1.3 E and Id proteins: Gatekeepers of lymphocyte development	20
1.3.1 Regulation of conventional T cell development	22

1.3.2 Regulation of innate-like T cell development	25
1.3.2.1 iNKT TCR expression and selection into the iNKT lineage.....	28
1.3.2.2 Initiation of the iNKT developmental program	30
1.3.2.3 iNKT population expansion	32
1.3.2.4 iNKT differentiation and maturation.....	34
1.4 Id proteins and cancer.....	38
1.4.1 Tumor suppressive roles of E2A in T cell lymphomas.....	44
1.4.2 Paradoxical role of Id proteins in T cell lymphomas	48
2. Materials and methods.....	51
2.1 Mice	51
2.2 Cell sorting and RNA extraction.....	51
2.3 ChIP-Seq analysis	51
2.4 RNA-Seq analysis.....	52
2.5 TCR α repertoire sequencing.....	53
2.6 Flow cytometry	54
2.7 Anti-CD3 injection.....	55
2.8 Histopathological analysis	55
2.9 Adoptive transfer of lymphoma cells.....	55
2.10 PCR.....	56
2.11 Microarray	57
2.12 Bioinformatics analysis.....	57
2.12.1 ChIP-Seq.....	57

2.12.2 RNA-Seq.....	57
2.12.3 TCR Repertoire analysis.....	58
2.12.4 Innate gene signature and network analysis.....	58
2.12.5 Gene Set Enrichment Analysis (GSEA).....	59
2.12.6 Correlation analysis	60
2.12.7 Microarray data analysis.....	60
2.13 Data plotting, visualization and statistics.....	61
2.14 Data availability.....	61
3. Regulation of innate-like T cell fate choice by E and Id proteins.....	62
3.1 Introduction.....	62
3.2 Results	63
3.2.1 Absence of Id proteins allows E2A to induces genes involved in iNKT cell development and function.....	63
3.2.2 E2A supports iNKT cell fate at the DP stage.....	67
3.2.3 Biased V α 14-J α 18 rearrangement in pre-selection cells	69
3.2.4 Pre-TCR signaling-independent expansion of iNKT population	75
3.2.5 Initiation of an innate-like transcriptional program	80
3.2.6 Transcriptional similarity between iNKT cells and $\gamma\delta$ NKT cells.....	82
3.2.7 E2A orchestrates an innate-like gene network	86
3.3 Conclusions and discussion.....	90
4. Malignant transformation of innate-like T cells	93
4.1 Introduction.....	93
4.2 Results	96

4.2.1 Neonatal expansion of CD1dTet ⁺ innate-like T cells	96
4.2.2 CD1dTet ⁺ expand independent of CD1d-mediated selection.....	96
4.2.3 Malignant innate-like T cells invade healthy tissues	97
4.2.4 Transcriptional program in neonatal expanding iNKT cells.....	102
4.2.5 Acquisition of multiple tumorigenic programs.....	108
4.3 Conclusions and discussion.....	114
5. Conclusions.....	118
6. Future Directions	120
6.1 Identifying downstream E2A targets critical for the large innate-like T cell populations in Id-deficient mice	120
6.2 Does Id-deficiency have an impact on other innate-like T cells?.....	120
6.3 Delineating the distinct roles and contribution of E2A in promoting the iNKT lineage prior to and upon TCR selection	121
6.4 How/do E proteins regulate the outcomes and timing of innate-like TCR α rearrangements?	122
6.5 Identifying associated mutations upon loss of Id2 and Id3	123
References	125
Biography.....	148

List of Tables

Table 1: Role of E and Id proteins in iNKT cell development.	38
Table 2: List of genotyping primers.....	56
Table 3: Cluster definitions according to Self-Organizing Map (SOM) analysis	107
Table 4: List of genes shared in key pathways with other innate-like tumors.....	112

List of Figures

Figure 1: Current paradigm of T cell development in the thymus	4
Figure 2: T cell Receptor (TCR) loci	9
Figure 3: Conflicting models of iNKT cell development.....	16
Figure 4: Transcriptional regulation by E proteins and their inhibitors, Id proteins	22
Figure 5: Key roles of E2A in T cell development	25
Figure 6: Role of E and Id proteins in iNKT cell development	27
Figure 7: Context-dependent roles of E and Id proteins in development and cancer	43
Figure 8: Upregulation of transcriptional programs that support iNKT cell development and function in the absence of Id proteins	65
Figure 9: Uninhibited E2A activity directly regulates genes that support the iNKT lineage in DP and iNKT cells.....	66
Figure 10: E2A regulates and collaborates with downstream transcription factors to promote the iNKT lineage fate choice at the DP stage	68
Figure 11: Id2/Id3 deficiency increases V α 14-J α 18 rearrangement in pre-selection DP pool of CD1d-deficient mice.....	71
Figure 12: Biased J α repertoire in pre-selection DP cells eligible for iNKT selection.....	72
Figure 13: E2A binds to E α and Trv11 promoters.....	74
Figure 14: Impaired IL-7R expression and proliferation in L-DKO DN4 cells.....	75
Figure 15: A block in pre-TCR signaling does not arrest enhanced iNKT cell and $\gamma\delta$ NKT populations in the absence of Id proteins.....	77
Figure 16: Development of innate-like DP cells in L-DKO pT $\alpha^{-/-}$ mice	79
Figure 17: Adoption of innate-like transcriptional program by DP cells in the absence of pre-TCR signaling and Id proteins	81
Figure 18: Transcriptional similarity between iNKT and $\gamma\delta$ NKT cells.....	84

Figure 19: Lineage competition between iNKT and $\gamma\delta$ NKT cells is regulated by Id proteins.....	85
Figure 20: E2A-driven transcriptional network.....	88
Figure 21: E2A drives a gene network that promotes iNKT and $\gamma\delta$ NKT fate in the absence of Id proteins.....	89
Figure 22: Proposed developmental model for divergence of innate-like and conventional T cell lineages, as regulated by E and Id proteins	92
Figure 23: Summary of phenotypes observed in Id-deficient mice	95
Figure 24: Neonatal development of Tet ⁻ cells in LDKO mice	98
Figure 25: CD1dTet ⁻ cells develop independent of CD1d.....	99
Figure 26: CD1dTet ⁻ TCR β ⁺ T cells give rise to lymphomas in TKO mice.....	100
Figure 27: Adoptive transfer of lymphoma cells gives rise to tumors in WT hosts	101
Figure 28: Microarray analysis of premalignant iNKT and tumor iNKT cells.....	103
Figure 29: Gene expression patterns in premalignant neonatal iNKT cells	106
Figure 30: Pathway and gene expression analysis in iNKT tumors	110
Figure 31: Distinct pathways drive expansion and tumorigenesis in L-DKO mice	111
Figure 32: Resemblance to human NK/T tumor	114

1. Introduction

Our body is defended against damage from a plethora of harmful agents and microorganisms by our complex immune system. Effector cells react to, manage and eliminate invading pathogens through diverse defense mechanisms. Broadly, the immune system is classified into innate and adaptive branches, each serving its unique function in protection. Cells of the innate immune system serve as front-line defenders, responding rapidly to antigens and doing damage control until cells of the adaptive immune system are able to mount a delayed, but highly specific immune response against the antigen. The adaptive immune system comprises of T and B lymphocytes, which develop in the thymus and bone marrow respectively. Starting with broad specificities for different types of antigens, these cells undergo selection and specific amplification upon exposure to a particular pathogen, leading to a targeted, highly effective immune response mediated by antibody production, direct killing, and activation of other effector cells through cytokines. There are several mechanisms and checkpoints enforced at different stages to ensure that immune cells only recognize and respond appropriately to foreign antigen, and not to self. In certain situations, a failure in these mechanisms can give rise to aberrant cells that cause autoimmunity.

A diverse and abundant population of conventional CD4⁺ and CD8⁺ $\alpha\beta$ T cells arising out of the thymus is necessary for a specific and effective immune response to antigens. A smaller but significant population of unconventional T cells concomitantly

develops in the thymus, with innate-like capabilities of mounting a rapid and potent immune response [1]. Innate-like T cells play a very important role in activating several immune cell types through the production of a broad range of Th1 (IFN- γ), Th2 (IL-4) and Th17 (IL-17) cytokines. These cells can play protective roles in infectious diseases, certain tumors and autoimmune diseases, but aggravating roles in allergy, asthma etc. Thus, it is critical to elucidate key transcription factors regulating the development of these innate-like cells for better insight into the roles they play in diseases, as well as for their therapeutic applications. In this dissertation, I will primarily focus on innate-like $\alpha\beta$ iNKT and $\gamma\delta$ NKT cells.

In the following section, I discuss conventional and innate-like T cell development in the thymus, how critical checkpoints along this development are regulated, particularly by transcription factors E proteins and their antagonists Id proteins. I also describe the complex relationship between E and Id proteins, and tumorigenesis.

1.1 Conventional T cell development in the thymus

Conventional $\alpha\beta$ T cell development and commitment in the thymus proceeds in a sequential manner, starting from the early thymic progenitor (ETP) to CD4⁻ CD8⁻ double negative (DN) cells (DN2, DN3, DN4) and CD4⁺ CD8⁺ double positive (DP) cells, which ultimately differentiate into CD4⁺ or CD8⁺ single positive (SP) cells (Figure 1). Each of these cell stages is unique in its signature transcription factor, surface markers,

extent of proliferation and T cell receptor (TCR) rearrangement status. Progression through these stages allows commitment to the T cell fate, followed by establishment of T cell identity defined by a TCR, and finally maturation and differentiation into distinct subtypes.

Starting from the same lymphoid progenitor as $\alpha\beta$ T cells, $\gamma\delta$ and iNKT cells diverge at distinct points in this developmental pathway upon expression of the appropriate TCR and positive selection. Cells with productive V(D)J rearrangement of the TCR γ and TCR δ at the DN2 or DN3a stage are directed towards the $\gamma\delta$ T cell lineage, whereas cells rearranging the β locus at the DN3a stage are driven towards the $\alpha\beta$ fate. In order to ensure that only cells with functional β rearrangements continue further, the β -selection checkpoint is enforced at the DN3b stage where the TCR β chain pairs with a pre-TCR alpha (pre-T α) chain and is able to generate an appropriate TCR signal only if the β rearrangement is functional. These cells proceed beyond the β -selection checkpoint to the DP stage, where rearrangement of the TCR α locus takes place, followed by positive selection to test the affinity of the fully rearranged TCR. DP cells must recognize self-antigens presented on MHC molecules on thymic epithelial cortical cells with only moderate affinity to receive survival signals. The type of Major Histocompatibility Complex (MHC) recognized by the TCR also determines whether the DP cells become conventional CD4⁺ or CD8⁺ SP cells.

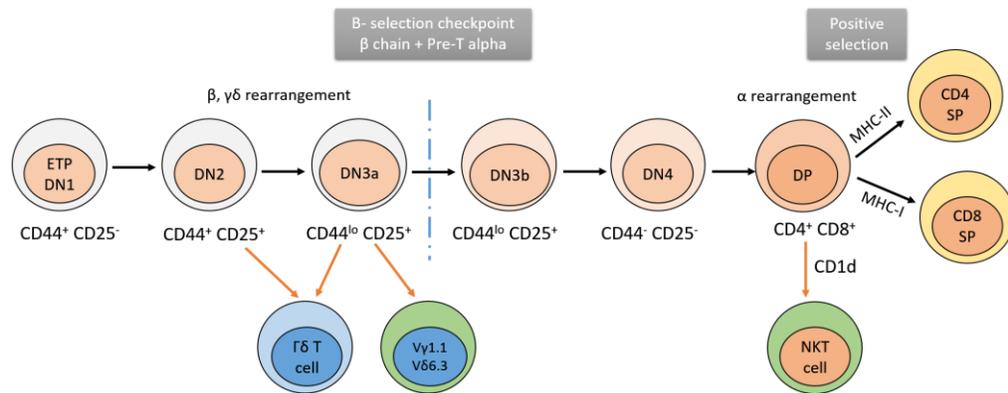


Figure 1: Current paradigm of T cell development in the thymus

T cell development in the thymus entails several sequential steps, starting from Early Thymic Progenitor (ETP) to Double Negative (DN1 – DN4) stages to Double Positive (DP) and finally Single Positive (SP) cells. These stages are distinguishable by surface expression of markers such as CD44, CD25, CD4 and CD8, and represent stages of T cell lineage commitment, TCR rearrangement and selection. Cells must pass through developmental checkpoints, such as the pre-TCR (or TCR β) and positive TCR selection checkpoints, which ensure development of cells with surface expression of an appropriate, functional TCR. Majority of cells adopt the conventional $\alpha\beta$ T cell fate by expressing a TCR with productive α and β rearrangements, and upon selection by Major Histocompatibility Class (MHC) molecules, become either CD4 or CD8 SP cells. A smaller percentage of cells with a productive $\gamma\delta$ TCR diverge at the DN2/3 stage to become $\gamma\delta$ T cells. Only a minor population of innate-like T cells arises in the thymus. It is considered that $\gamma\delta$ T cells that happen to express a V γ 1.1V δ 6.3 TCR become $\gamma\delta$ NKT cells, and $\alpha\beta$ T cells expressing a V α 14J α 18 TCR adopt iNKT fate upon TCR-mediated selection, alongside their conventional T cell counterparts.

1.1.1 Commitment to the T cell lineage

Hematopoietic Stem Cells (HSCs) from the bone marrow migrate to the thymus to undergo T cell lineage commitment and differentiation. Until these cells reach the DN2 stage, they still possess the potential to adopt other lineage fates. Irreversible commitment to the T cell lineage, and thereby loss of potential for other lineages, occurs

during the transition from DN2 to DN3 stage. This crucial process is regulated by expression of lineage-specific transcription factors that enforce the T cell lineage identity, while suppressing alternative lineage potentials, as well as cues from the thymic micro-environment.

The intricate regulation of timing and expression of transcription factors is critical for T cell commitment. For instance, PU.1 is important for early T cell progenitors, but must be downregulated by the DN3 stage to suppress the potential for myeloid lineages [2-4]. Factors, such as BCL11B, IKAROS, GATA3 and E proteins also play fundamental roles in this process [5, 6]. Additionally, Notch signaling activated by interaction with delta-like ligands from the environment is known to mediate T cell commitment and specification [7-9].

1.1.2 Generating a diverse repertoire

One of the most important features of our immune system is the ability to respond to a very wide range of harmful pathogens. This necessitates lymphocytes, including T cells, to be able to express cell surface receptors with wide range of specificities. T cells therefore undergo an elegant and elaborate process named V(D) J recombination, which allows random, but somewhat orchestrated combinations of variable (V), diversity (D) and joining (J) gene segments. In this way, a limited number of gene segments can give rise to a vast and diverse repertoire of T cell receptors. TCRs are not only important for antigen recognition, but also for MHC-mediated selection,

which a critical checkpoint for T cell development [10]. The TCR also defines the identity of T cells, such that $\gamma\delta$ T cells must successfully rearrange $\text{TCR}\gamma$ and $\text{TCR}\delta$ at the DN stage, whereas $\alpha\beta$ T cells successfully rearrange $\text{TCR}\beta$ at the DN stage followed by $\text{TCR}\alpha$ at the DP stage (Figure 2). The δ gene segments are located within the $\text{TCR}\alpha$ locus, which is why it is commonly referred to as the $\text{TCR}\alpha/\delta$ locus [11].

Recombination is mediated by recombination activation genes 1 and 2 (RAG1 and RAG2), through binding and formation of complexes between two random but preferred recombination signal sequences (RSS) flanking unique TCR gene segments, followed by cleavage of the RSSs and joining of corresponding TCR gene segments via non-homologous end joining (NHEJ). Further diversity is added through the loss of nucleotides or addition of non-template nucleotides during the joining process. On the flip side, creating DNA breaks during the recombination process poses the risk of undesirable mutations leading to tumorigenesis. Therefore, recombination is tightly regulated with respect to cell and stage-specific, cell cycle-dependent RAG expression, as well as RSS accessibility and chromatin remodeling [11-13].

1.1.2.1 $\text{TCR}\beta$, $\text{TCR}\gamma/\delta$ rearrangement at the DN stage

At the DN2/3 stage of T cell development, cells undergo $\text{TCR}\beta$, $\text{TCR}\gamma$ and $\text{TCR}\delta$ rearrangement. $\text{TCR}\beta$ recombination involves joining a $\text{D}\beta$ to $\text{J}\beta$ gene segment followed by combining that to a $\text{V}\beta$ segment [14, 15]. Upon successful rearrangement of

TCR β on one allele and pairing with pre-T α , downstream signaling downregulates Rag expression, preventing further β rearrangement on the other allele [16, 17].

Rearrangement for TCR γ and TCR δ is similar in many respects, but evidently different in terms of the number and location (for TCR δ) of gene segments (Figure 2). Given the fewer gene segments available for γ and δ chains, a lot of the repertoire diversity relies upon junction variability. $\gamma\delta$ T cells are also unique in terms of their development, as they arise in “waves” during different phases in life, and migrate to different tissues specific for their effector function. For instance, the first subset of $\gamma\delta$ T cells arises at embryonic day 13 (E13) of a mouse, with rearrangement of V γ 3. This is rapidly followed by the next waves of V γ 4 at E14 and V γ 2 and V γ 5 at E16 [18, 19]. It is therefore important to also consider the predominant $\gamma\delta$ TCR usage among these different developmental waves of $\gamma\delta$ T cells, and the possible regulatory mechanisms that tightly control developmental timing and rearrangement outcomes in these cells [20]. The importance of the $\gamma\delta$ TCR in determining the fate choice has also been debated. While some believe that expression of lineage-specific factors such as Sox13 precedes and determines the rearrangement of a $\gamma\delta$ TCR, many believe in the instructive or signal strength model, which proposes that a stronger TCR signaling event triggered by a productive $\gamma\delta$ rearrangement is imperative for commitment to the lineage [21, 22].

1.1.2.2 TCR α rearrangement at the DP stage

Unlike $\gamma\delta$ T cells that complete rearrangement and divert towards effector differentiation at the DN stage, cells with a productive TCR β chain need to proceed towards the DP stage to undergo TCR α rearrangement. Successful rearrangement of a TCR α chain is necessary for the expression of a functional TCR with both TCR α and TCR β chains. The presence of TCR δ gene segments within the TCR α locus, and their rearrangement at the DN stage, makes TCR α rearrangement at the DP stage more complex than TCR β rearrangement (Figure 2). While the E δ enhancer regulates TCR δ rearrangement, the E α enhancer facilitates the rearrangement of V α to J α segments [23]. Interestingly, the positioning of interspersed TCR δ gene segments within the TCR α locus doesn't seem to be simply fortuitous, but rather plays a role in the diversity of TCR α rearrangement outcomes as well [24]. Multiple attempts are possible towards achieving a successful TCR α rearrangement [25, 26]. However, there is a preferential order of proximal to distal rearrangements for usage of TCR α gene segments. Accordingly, achieving a distal 5' V α or 3' J α would require multiple rounds of rearrangement and a longer lifespan at the DP stage [14].

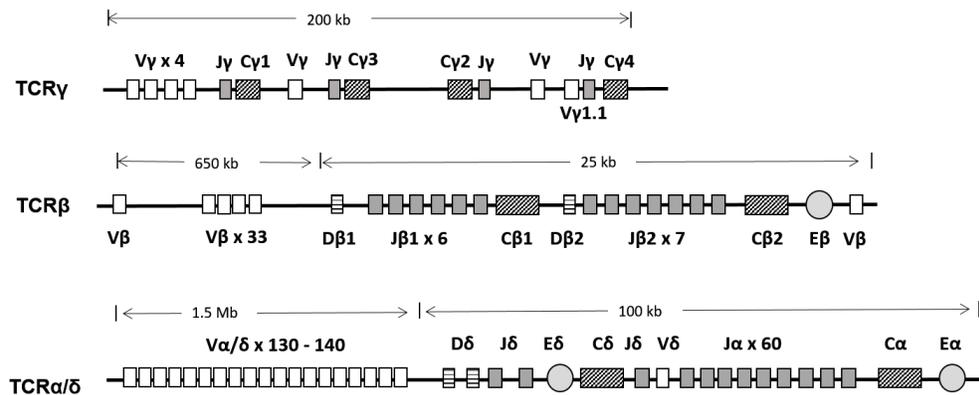


Figure 2: T cell Receptor (TCR) loci

Schematic representation of the TCR β , TCR γ TCR α/δ loci in mice with the variable (V), diversity (D), joint (J) and constant (C) gene segments, and enhancers (E) labeled. Number of gene segments of each type, and distance between parts of the loci are also indicated.

1.1.3 Selection checkpoints for appropriate T cell development

1.1.3.1 Pre-TCR selection at the DN stage

Cells that have rearranged TCR β at the DN3 stage must progress towards the DP stage to seal their $\alpha\beta$ T cell lineage fate choice. However, there is a halt in cell cycle and progression enforced by the β -selection checkpoint [27]. A functional pre-TCR signal requires the pairing of a successfully rearranged TCR β chain with the surrogate α chain, better known as pre-T α . This pre-TCR-mediated release of cell cycle and development arrest at the DN stage is distinct from the role of TCR signaling at the DP stage [28]. The pre-TCR signal activates a downstream pathway involving signaling molecules and transcription factors that are necessary for cell cycle progression, survival and differentiation towards the next stage [29]. This includes activation of c-Myc, NF- κ B and the Ras/Raf/ERK pathway that upregulates Id3 [29-32].

1.1.3.2 TCR-mediated selection at DP stage

The halt in cell cycle and progression is resumed upon productive pre-TCR signaling, leading cells into the DN4 stage, where they continue proliferating until they reach the DP stage via the Immature Single Positive (ISP) stage. The process of TCR α rearrangement is started at the DP stage with the goal of achieving a productive rearrangement, which will allow them to overcome the second checkpoint to finally become CD4 or CD8 SP effector cells. This selection checkpoint requires the assembly of an $\alpha\beta$ TCR on the cell surface, which is then tested for its ability to recognize antigens presented by MHC molecules.

Unlike the pre-TCR signaling checkpoint however, more caution is warranted for the TCR selection checkpoint as it qualifies T cells to differentiate, mature and migrate to perform effector functions in the periphery. Therefore, the selection process must positively select for weak recognition of MHC-presented self-peptides, but negatively select against strong affinity for self-peptides. Cells that do not pass these selection criteria die from neglect, i.e. lack of a survival signal from MHC recognition, or apoptosis, i.e. an activation-induced cell death signal due to self-reactivity. TCR recognition of MHC Class I and Class II molecules drive differentiation towards the CD8 SP and CD4 SP fate respectively. An optimum TCR signal allows survival of the positively selected cell through upregulation of Bcl-2 family members [33, 34]. It also facilitates the progression of these cells from the DP to SP stage through upregulation of

CCR7, allowing migration to the medulla, as well as factors like Id3, Foxo1, IL-7R α [27, 35]. It is evident that optimum signal strength is critical for the TCR selection checkpoint. The threshold of TCR signaling and downstream pathways that determine differential outcomes remains an active area of research [36, 37].

1.2 Innate-like T cells

Innate-like T cells have garnered increasing interest as their memory phenotype can be harnessed in the context of allergies, infections and tumors. Innate-like T cell populations include TCR $\alpha\beta$ ⁺ natural killer T (iNKT) cells, TCR $\gamma\delta$ ⁺ iNKT cells ($\gamma\delta$ NKT), innate-like CD8⁺ T cells, CD8 $\alpha\alpha$ intraepithelial lymphocytes (IELs), and mucosal-associated invariant T (MAIT) cells.

1.2.1 Innate-like $\alpha\beta$ invariant natural killer T cells

A distinct population of $\alpha\beta$ T cells possessing NK (Natural Killer) cell markers and the innate-like ability to mount a potent immune response within hours of exposure to antigens, is referred to as NKT (Natural Killer T) cells. A unique feature that distinguishes these cells from most conventional $\alpha\beta$ T cells is the ability to recognize microbial and self-lipids presented on the non-canonical CD1d molecule, which is MHC (Major Histocompatibility Complex) Class-I like in structure. It has been found that NKT cells can be activated directly by antigen recognition, or indirectly by APCs (Antigen presenting cells) [38, 39], to produce a wide range of cytokines. Further, unlike the diverse TCR (T cell receptor) repertoire represented by conventional $\alpha\beta$ T cells, most of

these cells express a semi-invariant TCR. The most well characterized subset of these are the type I NKT cells, or invariant Natural Killer T (iNKT) cells that express an invariant V α 14-J α 18 TCR α chain paired with primarily V β 8.2, V β 7 or V β 2 chain in mice, or an invariant V α 24-J α 18 V β 11 TCR in humans [40, 41]. This semi-invariant TCR allows these cells to recognize α -GalCer, a marine sponge derived α -galactosylceramide, among other closely related lipids, which is also utilized for their tetramer-based identification across mice, humans and non-human primates [42]. Type II NKT cells have more diverse TCR pairings and recognize other CD1d-presented lipids, but will not be focused upon in this dissertation.

Although iNKT cells constitute only a small fraction of T cells in the thymus and periphery of mice and humans, their invariant TCR and recognized ligands have been evolutionarily highly conserved across species, indicating a critical role in the immune system. iNKT cells are known to play a contextual role in diseases, as they are found to be protective in infectious diseases, tumors and certain autoimmune diseases but harmful in asthma and allergy [43-45]. Their ability to cross-activate dendritic cells (DCs) and other immune effectors through cytokines and chemokines has also garnered a lot of attention to their potential as vaccine adjuvants [46].

It is currently accepted that these cells arise from conventional $\alpha\beta$ T cell progenitors and follow the same developmental program until the DP (CD4⁺CD8⁺ double positive) stage, where the stochastic expression of the semi-invariant V α 14-J α 18

TCR (henceforth referred to as iNKT TCR) them to undergo TCR selection mediated by a CD1d molecule expressed on other conventional DP thymocytes, bifurcating them from conventional SP (CD4⁺ or CD8⁺ single positive) fate [47-49]. This developmental pathway is known to be regulated at different stages by several transcription factors including PLZF (Promyelocytic leukaemia zinc finger protein) [50] and EGR2 (Early growth response 2) [51], but the exact developmental regulatory programs in iNKT cells are far from fully elucidated.

iNKT cells are either CD4SP or DN cells that proceed through 4 broad stages of development - commitment (stage 0, CD24^{hi}), expansion (stage 1, CD24^{lo} CD44^{lo}), differentiation (stage 2, CD44^{hi}) and maturation (CD44⁺ NK1.1⁺) [52]. Previously, it was thought that stage 1 iNKTs secrete only IL-4 on stimulation, stage 2 cells produce both IL-4 and IFN- γ and stage 3 cells produce mostly IFN- γ , but it has now been shown that 3 terminally differentiated iNKT subsets (iNKT1, iNKT2, iNKT17) similar to the conventional T helper cell subsets exist within these linearly ordered stages of iNKT development, each with distinct cytokine profiles [53-56].

1.2.2 Innate-like $\gamma\delta$ natural killer T cells

$\gamma\delta$ NKT cells are yet another population of innate-like $\gamma\delta$ T cells that express a restricted V γ 1.1V δ 6.3 TCR, along with NK cell surface markers [57]. Among the waves of $\gamma\delta$ T cell development, $\gamma\delta$ NKT cells expand primarily during the neonatal window, perhaps being derived from late fetal precursors [58]. Both iNKT and $\gamma\delta$ NKT cells are

characterized by high levels of expression of the innate-like transcription factor, promyelocytic zinc finger (PLZF), and readily produce effector cytokines like IL-4 [59, 60]. Similarities with iNKT cells in terms of effector functions and dependence on PLZF expression has inspired the name $\gamma\delta$ NKT cells for this subset of $\gamma\delta$ T cells [57, 58, 60, 61].

Similar to their conventional counterparts, a correlation between TCR signaling strength and $\gamma\delta$ NKT lineage choice has been proposed. This is primarily due to the observed expansion of $\gamma\delta$ NKT cells in mice with deficiency of downstream TCR signaling molecules, Id3 or ITK, as well as the overexpression of DOK1, that plays a negative role in TCR signaling [61-64]. Besides the preference for lower TCR signal strength, $\gamma\delta$ NKT cells also depend on SLAM/SAP mediated signaling, similar to iNKT cells, which is further described in the next section [65].

1.2.3 Innate-like T cell lineage commitment

Ever since their discovery, the lineage history of iNKT cells has been a topic of debate in the field, with researchers supporting either the “pre-commitment” or the “TCR-instructive” model of iNKT cell development (Figure 3) [66]. The first model suggests that iNKT lineage fate is preprogrammed, i.e. an intrinsic iNKT-specific program exists, or is turned on, in precursors even prior to CD1d-mediated selection, which separates them from conventional T cells. These pre-committed iNKT cells then undergo further lineage specification upon signaling through the iNKT TCR and CD1d-mediated selection [67]. In support of this theory, $V\alpha 14J\alpha 18$ rearrangements have been

detected prior to development of conventional $\alpha\beta$ T cells during fetal life indicating a possibly distinct iNKT developmental pathway [68]. Also, iNKT “precursor” DN4 cells with V α 14J α 18 transcripts (but no surface TCR expression) have been detected in CD1d-deficient mice, and shown to give rise to iNKT cells when cultured with normal thymocytes^[69]. This hints towards the existence of early precursors with an iNKT-specific program, and possibly biased towards iNKT TCR rearrangement prior to selection. Recent data from our lab also hints towards the role of E proteins in promoting iNKT TCR rearrangement before the DP stage [70]. In addition, it has been proposed that the distinct transcriptional machinery required for development of conventional T and iNKT cells supports the idea of distinct precursors [71, 72]. In line with this, it has been found that several mutations selectively impair iNKT development without any impact on conventional T cell development [71, 73]. It has also been suggested that innate-like $\alpha\beta$ and $\gamma\delta$ T cells follow a different developmental program from conventional T cells, such that the TCR expression of the opposite isotype is not selected against in these cells^[74]. The complete block in conventional T cell development in TCR α ^{-/-} animals with a V α 14 transgene again indicates that the iNKT TCR may primarily be expressed only in iNKT precursor cells [75]. Thus, data in support of the pre-commitment theory is isolated but substantial.

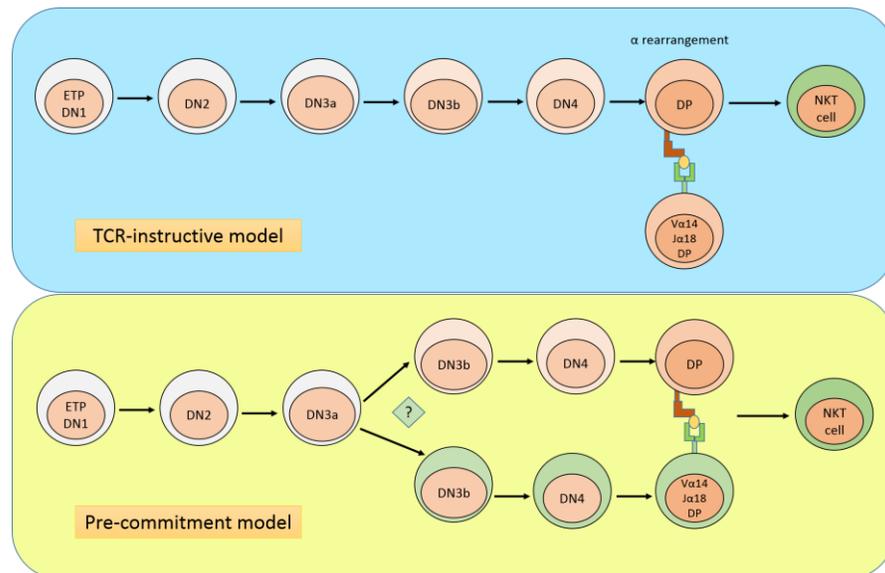


Figure 3: Conflicting models of iNKT cell development

The TCR-instructive model is shown in the upper panel, where lineage fate choice is determined solely upon receiving the appropriate iNKT TCR-mediated signal. The lower panel depicts the pre-commitment model where precursor cells already turn on some lineage specific programs that support the iNKT TCR rearrangement, which finally allows selection into the lineage.

The TCR-instructive model postulates that DP cells which have stochastically rearranged a $V\alpha 14$ - $J\alpha 18$ TCR, undergo CD1d-mediated selection and subsequent SLAM/SAP (Signaling lymphocytic activation molecule/ SLAM associated protein) signaling^[76], that drives their iNKT lineage commitment. In favor of this model, there is strong evidence to suggest that the iNKT TCR is generated from the same random rearrangement process as conventional T cells, rather than being uniquely regulated as suggested by the pre-commitment model. For instance, the TCR repertoire in iNKT cells reflects the same stochastic patterns as conventional $TCR\alpha$ rearrangement [40, 77]. Furthermore, it has been found that the iNKT-specific $V\alpha 14$ - $J\alpha 18$ rearrangement is also

subject to the same rearrangement bias of pairing with proximal V α and J α segments [78]. This would imply that only secondary rearrangements can give rise to the distally rearranged V α 14-J α 18 TCR. Along these lines, mice deficient in ROR γ t (RAR-related orphan receptor gamma t), an important survival factor for DP cells that prolongs the window of secondary rearrangements [26], lack iNKT cells due to impaired V α 14-J α 18 rearrangement [47, 48]. iNKT development can also be rescued by expression of a Bcl-xL (B-cell lymphoma – extra-large) transgene, which increases DP cell survival and hence promotes distal rearrangements [48]. iNKT cell development has also been shown to consist of a DP intermediate [49], and to proceed through the DP stage using fate mapping techniques [48]. It has been proposed that similar to $\gamma\delta$ T cells, strong TCR signaling is critical for selection into the iNKT lineage [79], which was also implied in a reporter mouse model [80]. A role of strong TCR signaling in PLZF expression, a key transcription factor in iNKT development, has also been shown [51, 60]. These observations have garnered great support in favor of the TCR-instructive model. However, it is important to note that both models heavily stress and argue upon the timing and regulation of V α 14-J α 18 rearrangement. While it is clear that CD1d-mediated TCR engagement provides the decisive signal for iNKT lineage commitment and specification, it still needs to be determined whether all or just a fraction of DP cells are competent to become iNKT cells. Re-evaluating published data in this context may allow

us to uncover the transcription program that distinguishes iNKT TCR from other TCR signals during iNKT commitment.

1.2.4 Transcriptional Regulation of iNKT and $\gamma\delta$ NKT lineages

Selection into any lineage is followed by turning on the appropriate transcriptional program, marking commitment into the lineage. This program must ensure the right level of expression of transcription factors needed for development and maturation, and is usually distinctly recognizable by the expression of a “master regulator”. After an intensive screening in search of a unique iNKT-specific transcription factor, PLZF was identified as a key regulator of iNKT development [50, 59]. Later it was found that iNKT cells shared PLZF expression with other innate-like cells as well [60]. PLZF is found to be upregulated soon after TCR signaling in stage 0 and stage 1 iNKTs. Its expression is critical for iNKT development and function, as is demonstrated by the decrease in iNKT numbers, disturbed cytokine production and iNKT localization in lymph nodes in PLZF^{-/-} mice [50].

While the transcriptional programs that drive conventional CD4⁺ and CD8⁺ T cell specification and development have been well characterized, little is known about the innate-specific transcriptional programs upstream of PLZF that are responsible for the divergence of innate-like T cells from conventional T cells [81].

1.2.5 Regulation of iNKT subsets

Similar to the T helper subsets, three iNKT cell subsets iNKT1, iNKT2 and iNKT17, have recently been characterized [53]. The definition of conventional stage 2 CD44^{hi} NKTs has accordingly been revised to comprise of a heterogeneous population of terminally differentiated iNKT2 and iNKT17 cells, and iNKT1 progenitor cells [54]. These subsets later migrate to the periphery or stay in the thymus to mature, and have varying transcriptional programs to direct production of distinct cytokines in the periphery. While iNKT1 cells are defined by their IFN- γ production with high T-Bet (or T-box transcription factor, TBX21) and low PLZF expression, NKT2s are IL-4 and IL-13 producing cells with high GATA3 (GATA Binding Protein 3) and PLZF and moderate IL-17R β expression, and iNKT17 cells are identified as IL-17 producing ROR γ t⁺ cells with low expression of PLZF and T-BET [53]. Similar to skewing of T helper responses in different mouse strains, the distribution of iNKT subsets also varies widely between mouse strains [54]. Due to the prevalence of iNKT1 cells in C57BL/6J mice [82, 83], this subset has been extensively studied and is shown to mature in the thymus. While some consider mature iNKT1 cells to be same as the conventional mature stage 3 NKTs [84] (NK1.1⁺ CD44⁺), others believe that stage 3 can consist of all 3 mature subsets [85]. All three subsets localize to different organs in the periphery - iNKT1 cells are found in the liver and spleen, iNKT2 cells in the lung and iNKT17 cells are found mostly in lymph nodes [53].

The discovery of these terminally differentiated subsets of iNKT cells has altered our understanding of the development of this innate-like lineage. Although their key transcription factors and cytokine profiles have been well characterized, their differentiation from either a common or distinct iNKT precursor(s) is not yet known. It has been proposed that a separate checkpoint may exist for directing iNKT cells towards iNKT1, iNKT2 or iNKT17 fate, which is regulated by either transcription factors or TCR signal strength [86]. Since all three subsets reside in distinct locations in the periphery, it is hypothesized that the cytokine milieu in the microenvironment of the iNKT subset may also play a major role in their homeostasis [85]. Further, the plasticity of the subsets makes it harder to uncouple the intrinsic transcriptional requirements from external stimuli. TCR signals and signal strength are unlikely to be the only external stimuli affecting the transcriptional program, and the role of cytokines, chemokines and growth factors in this setting are still waiting to be explored.

1.3 E and Id proteins: Gatekeepers of lymphocyte development

A family of Class I bHLH (basic Helix Loop Helix) proteins known as E proteins can regulate transcription by binding to E-box (CANNTG) domains, and thereby play key roles in regulating both B and T cell development [87]. These transcription factors possess activation domain(s) (AD), a basic DNA binding domain and a helix-loop-helix (HLH) domain (Figure 4). E protein family members (E2A, HEB and E2-2) homo- or hetero- dimerize with each other using the HLH domain in order to bind DNA. Unlike

the canonical E proteins that have two activation domains, there are alternative HEB and E2-2 forms with only one AD, accompanied by a unique N-terminal coding exon known as the Alt domain. Based on the tissue-specific expression of different E proteins and their alternate forms, E proteins can mediate regulation of diverse downstream programs [88, 89]. This DNA binding and regulation can be prevented by the formation of heterodimers with members of another HLH (Helix Loop Helix) family known as Id (Inhibitor of DNA-binding or differentiation) proteins (Figure 4) [87, 90, 91].

Id proteins, including Id1, Id2, Id3 and Id4, are members of Class V HLH proteins that can dimerize with other HLH protein classes, but unlike their binding partners, lack an additional basic DNA binding domain. They possess structural and conformational flexibility that is necessary for them to interact with different binding partners, although E protein binding is reported to be the strongest *in vivo* [92, 93]. All four Id protein family members share highly conserved HLH domains, but have differences in the N and C terminals, which can allow them to bind to diverse partner protein classes, albeit with varying binding preferences [93-96]. Id2 and Id3 are the key Id family members that are known to inhibit E protein activity in lymphocytes [89]. The interaction between the HLH Id proteins and bHLH E proteins precludes the DNA binding ability of the latter, making Id proteins potent inhibitors of E protein function. Interestingly, the antagonistic interaction between E and Id proteins allows them to serve as gatekeepers or natural “brakes” in lymphocyte development, such that there is

halt in progression into the next stage of development or maturation, until appropriate signals that regulate their expression tip the E/Id protein balance to allow cells to progress, accompanied by the expression of appropriate target genes for the next stage.

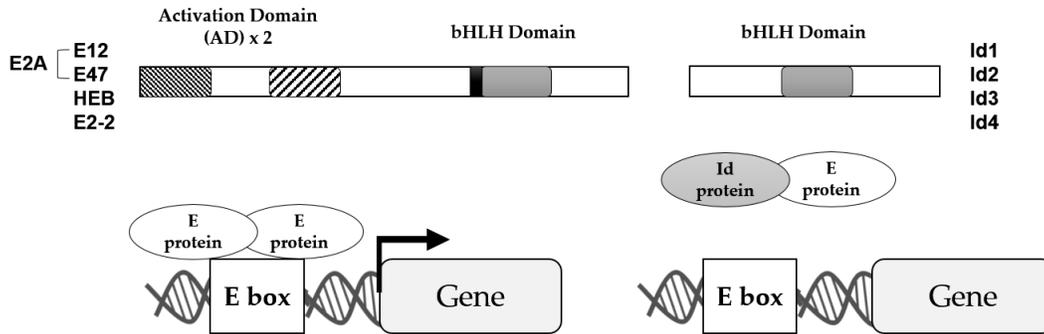


Figure 4: Transcriptional regulation by E proteins and their inhibitors, Id proteins

Top panel depicts structural domains in E and Id proteins including activation domains (AD) and basic-helix-loop-helix (bHLH) domains. Bottom panel shows the interaction between E proteins, allowing them to regulate genes with E-box domains in their regulatory regions. In the presence of Id proteins, this dimerization and associated DNA-binding of E proteins is lost, such that E proteins' regulatory function is inhibited.

1.3.1 Regulation of conventional T cell development

E proteins play important roles in driving T cell development even before precursors enter the thymus, by supporting HSCs and lymphoid fate choice. As ETPs enter the thymus, E proteins initiate lymphoid priming by upregulating key T cell factors like IL-7R, Notch1 and CCR9 [88, 97-100]. E proteins continue to play crucial roles in early stages of T cell commitment, lineage specification and differentiation upon appropriate T cell receptor (TCR) expression (Figure 5) [87, 101-104]. Their activity must

be down regulated upon pre-TCR signaling to allow DN to DP transition, and subsequently must be further repressed by Id proteins upon TCR signaling to allow DP to SP transition [105-107]. In line with this, there is a significant developmental block at the early Double Negative (DN) stage in the absence of E2A [108, 109]. E2A has been found to regulate several T cell-related genes in a heterodimer form with HEB, as a mutant HEB that is unable to bind DNA, results in a partial block at the DN stage and impaired recombination [110, 111]. E protein family member, E2-2, has also been proposed to play redundant roles in T cell development [112].

E2A is found to activate Notch signaling to further induce T cell-specific genes like *Tcf1*, *Gata3* and *Bcl11b* [113]. E2A blocks developmental progression at pre-TCR and TCR selection checkpoints such that only a productive signal can activate downstream events that upregulate Id3 expression to ultimately overcome the E2A-mediated block [103]. A proliferative burst after pre-TCR selection is associated with the induction of cyclin D3 (*Ccnd3*) by E protein family member E47 [114, 115]. Several other target genes that help E47 in regulating cell survival, cell cycle and maturation have been elucidated, and involve the induction of JAK/STAT pathway and SOC signaling [116]. Given the important function of E proteins, the regulation of Id proteins becomes very critical in early T cell development, especially considering the redundancy between Id protein family members and potential impaired development and tumorigenic outcomes [117, 118]. Some aspects of the role of E proteins in T cell development are still under

examination. This includes the relationship between E2A and Notch as well as the role of E2A in V(D)J recombination and allelic exclusion [119-123].

E and Id proteins continue regulating key developmental steps beyond the pre-TCR checkpoint (Figure 5). E proteins, particularly HEB, are known to impact DP cell survival through ROR γ t expression [124, 125]. This regulation of DP lifespan can impact TCR α repertoire, as discussed in section 1.1.2. The transition from DP to SP stage, as regulated by the TCR-selection checkpoint is also tightly controlled by E proteins. This checkpoint is impacted by the dosage of E2A and HEB, such that E protein deficiency leads to increased selection independent of the TCR [126-128]. Beyond TCR selection, the lineage choice between CD4 SP and CD8 SP cells is also subject to E protein dosage. While high E protein activity (or Id protein deficiency) preferentially supports CD4SP fate, E protein deficiency supports CD8 SP fate choice [127, 129]. These outcomes are controlled by E proteins through regulation of lineage-specific transcription factors, as well as other important chemokine and cytokine signaling receptors. Therefore, the dosage of E and Id proteins can dramatically impact all stages of conventional $\alpha\beta$ T cell development and maturation.

The $\gamma\delta$ TCR selection checkpoint is also regulated in a similar manner by E and Id proteins. E2A and HEB are deemed important for $\gamma\delta$ T cell development, and regulation of $\gamma\delta$ TCR rearrangement since deficiency in E proteins leads to reduced $\gamma\delta$ T cell numbers and skewed TCR usage [61, 130-132].

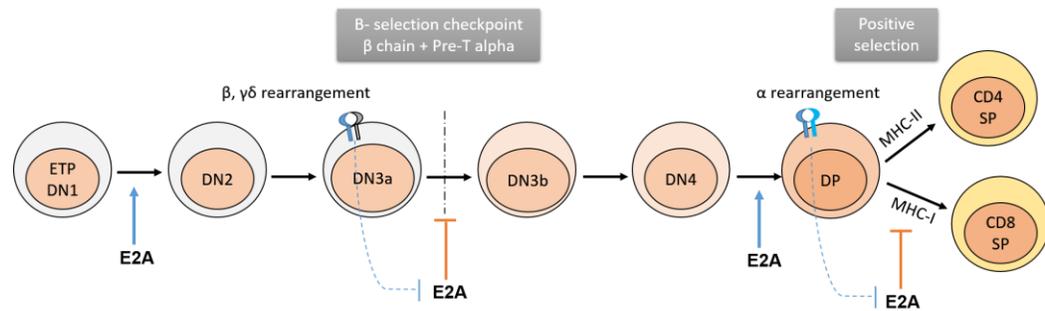


Figure 5: Key roles of E2A in T cell development

Schematic depicting the regulation of conventional T cell development by E2A. Developmental checkpoints or steps supported by E2A are represented by thick blue arrows, and those inhibited by E2A are shown by orange flat head arrows. Signaling downstream of pre-TCR and TCR signaling is depicted by dotted flat head arrows that block E2A activity (through Id3 upregulation, not shown).

1.3.2 Regulation of innate-like T cell development

Majority of this section is adapted from the publication "Orchestration of invariant natural killer T cell development by E and Id proteins", Sumedha Roy, and Yuan Zhuang, Critical Reviews in Immunology. 2015;35(1):33-48

E and Id proteins have been previously demonstrated to play roles in iNKT development [113, 133, 134] such as the critical role of HEB in iNKT TCR rearrangement [125] and Id2 in hepatic iNKT homeostasis [135]. In contrast to the developmental block in conventional T cell development, large populations of iNKT, $\gamma\delta$ NKT and innate variant T_{FH} cells have been observed in the same Id3- and Id2/Id3-deficient animals, indicating a negative role for Id proteins in regulating innate-like T cell development [61, 70, 136-139]. However, the mechanism that drives the development and expansion of these innate-like T cell populations in Id-deficient mice is still elusive. Recently, more

substantial evidence has been published to indicate various roles of these proteins in iNKT lineage commitment, development and function (summarized in Figure 6 and Table 1) [70, 84-86, 139].

In the summary figure I have also marked several “steps” in iNKT development, which may be regulated by transcription factors, particularly E proteins. These steps include iNKT TCR rearrangement, CD1d-mediated selection and subsequent transitions between stages. For example, EGR2 is found to be important for PLZF expression [51], while NF- κ B [140] (nuclear factor kappa-light-chain-enhancer of activated B cells) and MYC [141] are important for the proliferative burst at stage 1. Although the roles of E or Id proteins in iNKT development have so far been limited to that of HEB [125] and Id2, [135] several key roles have been revealed in more recent papers which are discussed in subsequent sections.

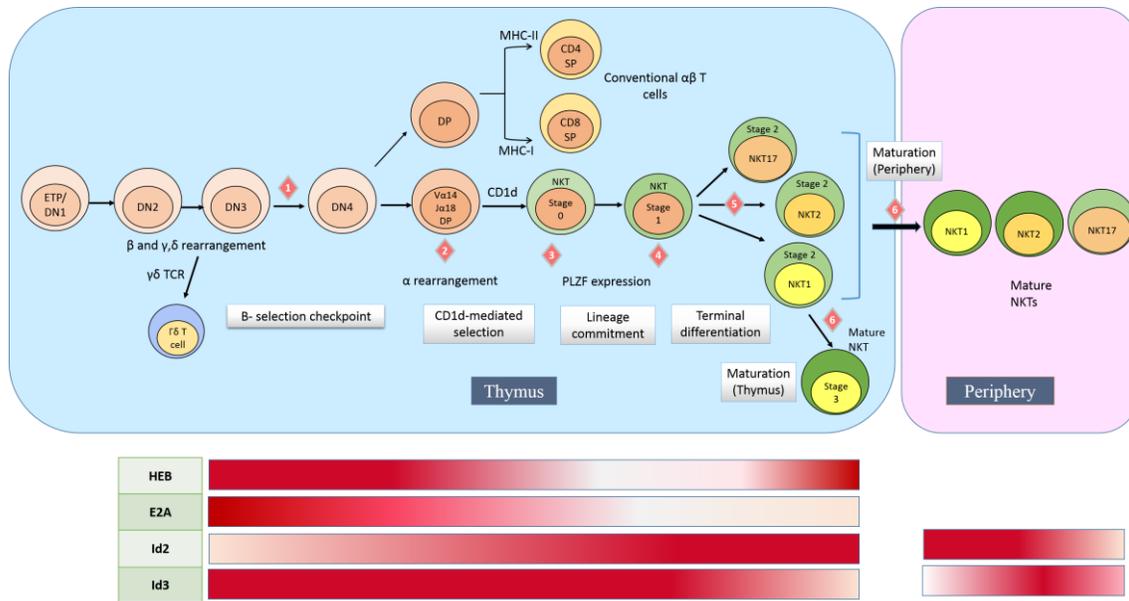


Figure 6: Role of E and Id proteins in iNKT cell development

iNKT cell development is considered to occur concurrently with conventional $\alpha\beta$ T cell development up to the DP stage. There are several important “steps” (depicted as numbered diamonds) that are critical for their appropriate lineage specification and functional maturation, and are regulated by E and Id proteins to influence their development. A functional TCR β rearrangement allows DN3 cells to proceed beyond pre-TCR selection (step 1) and reach the DP stage where α rearrangement takes place. DP cells that then stochastically express a Va14Ja18TCR (step 2) typically diverge from conventional $\alpha\beta$ T cells upon CD1d-mediated selection and undergo iNKT lineage commitment marked by PLZF expression at stage 0 and 1 (step 3). This is followed by an expansion phase at stage 1 (step 4) and terminal differentiation into iNKT1, iNKT2 or iNKT17 stage 2 subsets (step 5). These subsets with varying cytokine profiles, then migrate to different parts of the periphery to mature (step 6). The iNKT cells that mature within the thymus were previously characterized as stage 3 iNKTs (step 6), but have now been identified to be iNKT1 cells. E and Id protein expression during iNKT cell development (beyond DN3 stage) is graphically represented below the corresponding stages (dark red represents highest expression, white represents little to no expression).

1.3.2.1 *iNKT TCR expression and selection into the iNKT lineage*

After pre-TCR selection at the DN3 stage, the expression of a hallmark V α 14-J α 18 TCR in DP cells and selection by CD1d-presented antigen on other DP cells, allows commitment to the iNKT lineage (step 2, figure 6). This process of “selection” may be randomly governed by TCR expression, or may be pre-determined by an iNKT precursor, depending on which model of development is being considered. In this section, I will primarily discuss the role of E proteins in influencing iNKT TCR rearrangement, but it is worthy of mentioning that consistently high levels of E proteins can impair the pre-TCR checkpoint and limit selection into the $\alpha\beta$ lineage, which would also take a toll on selection into the iNKT lineage [70].

It has been demonstrated in the past that Id2 deficiency allows normal iNKT development, but impairs iNKT survival and localization to the liver [135]. However, the first concrete evidence of E proteins playing a role in iNKT development was the lack of the E protein HEB leading to a severe reduction in iNKT cell number at all stages in the thymus and periphery [125]. This phenotype could be explained by the associated reduction of ROR γ t expression, a direct target of E proteins [127, 142], and the downregulation in ROR γ t target Bcl-x $_L$, an important survival factor for DP cells [124]. As mentioned before, a similar phenotype is also observed in ROR γ t^{-/-} mice due to reduced secondary TCR rearrangements, particularly V α 14-J α 18 [48]. Another interesting observation made by this group was that even an iNKT TCR transgene could

not fully rescue development past stage 1, suggesting that HEB may play other roles in influencing iNKT differentiation and maturation. It is important to note that E2A and HEB have different expression levels and often regulate T cell development as heterodimer partners, and hence most likely play different roles during iNKT development [125].

Kee's group has demonstrated that Id3 deficiency leads to increased selection into the iNKT pathway partially because it allows continued expression of E protein targets ROR γ t, RAG1 and RAG2 [84]. Thus, even after positive selection, a fraction of Id3^{-/-} DP thymocytes continue rearranging their TCR, resulting in a larger fraction of cells with the distal V α 14-J α 18 rearrangement. We have also found that combined deletion of both Id2 and Id3 at the DN3 stage using Lck (lymphocyte-specific protein tyrosine kinase) Cre leads to a dramatic expansion of both CD4⁺ and DN iNKT cells [70]. Using high-throughput sequencing of the TCR repertoire of pre-selection CD69⁻ DP cells from these mice, we observed that there was random pairing of J α chains in cells with the common V α 8 rearrangement but a positively biased pairing of the J α 18 chain in cells with V α 14 rearrangements, as compared to WT mice. This suggests that the lack of Id proteins or very high E protein levels can perhaps promote a skewed TCR repertoire in favor of selection into the iNKT lineage [70].

Thus, E proteins are critical for iNKT development in order to allow more secondary rearrangements at DP stage, and Id proteins limit selection by preventing the

same. However, it still remains to be determined if E proteins are also capable of biasing the lineage decision by specifically regulating iNKT TCR expression, and if so, how.

1.3.2.2 Initiation of the iNKT developmental program

Expression of PLZF is considered pivotal for the iNKT cell developmental program, and hence our need for understanding its regulation in these cells is critical (step 3, Figure 6). Strong TCR signaling, as is observed in case of iNKT selection, can immediately lead to EGR2 upregulation followed by PLZF expression [51]. The presence of an EGR2 binding site in the PLZF promoter region and the failure to upregulate PLZF in EGR1/2 double deficient mice indicates that TCR signaling mediated EGR2 upregulation leads to PLZF expression in stage 0 iNKT cells. On the other hand, another group's work has demonstrated that E proteins bind to two E-box sites in the PLZF promoter region to directly regulate PLZF expression [85]. Sustained E protein expression can also cause an increase in PLZF expression [86]. In contrast however, Id2/Id3 double deficient mice have been found to have a 2-fold reduction in PLZF expression, as well as a naïve iNKT phenotype similar to PLZF^{-/-} mice [84]. Further, the authors found an increase in EGR2 levels in these iNKT cells, implying that Id2 and Id3 are important for PLZF expression in an EGR2-independent manner.

There are several possibilities that can explain these contrasting observations. It is possible that Egr2 positively regulates either E or Id proteins. The fact that TCR signaling promotes Egr2, which lies upstream of Id3 in $\gamma\delta$ T cells [136], suggests that

Egr2 may downregulate E proteins in these cells. E proteins and Egr2 may also coordinate to promote PLZF, which would explain why loss of either E proteins or Egr2 leads to a similar phenotype. Furthermore, Egr2 deficiency has also been demonstrated to reduce Id2 expression in iNKT cells [51]. This could imply that Egr2, Id and E proteins can also regulate PLZF expression at different stages during iNKT development. For instance, initially Egr2 and E proteins can promote PLZF together, until Egr2-mediated Id2 expression suppresses E protein to prevent further PLZF upregulation. Alternatively, Id2 and Id3 may positively regulate PLZF expression in the early stages of iNKT development. This would be supported by the observation that ET-2 mice with sustained E protein levels throughout iNKT development, showed low PLZF in stage 1 but high PLZF in stages 2 and 3 [86].

While the exact mechanism of PLZF regulation in iNKT development is unclear, it is certain that E and Id proteins are critical for appropriate expression of PLZF, and hence, are important for iNKT-specific program initiation. It is important to note that PLZF is also considered important for the development of innate-like $\gamma\delta$ NKT cells [60], and so E and Id proteins can promote both iNKTs and these $\gamma\delta$ T cells. Indeed, our data has shown that reducing E protein levels to 50% lower (Id2^{f/f} Id3^{f/f} E2A^{+/-} HEB^{+/-} LckCre⁺) than Id2^{f/f} Id3^{f/f} LckCre⁺ mice, led to a drastic change in phenotype from predominantly iNKT cells to mostly $\gamma\delta$ NKT cells, suggesting that E protein levels can influence lineage

choices between these two closely linked cell types [70]. The downstream E protein targets that assist in this lineage choice are yet to be determined.

1.3.2.3 *iNKT population expansion*

In comparison to conventional T cells that undergo antigen-driven clonal expansion in the periphery [143], stage 1 iNKT cells start undergoing a major population expansion in the thymus immediately following positive selection and iNKT lineage commitment (step 4, figure 6) [144]. Several transcription factors, including NF- κ B [140, 145] and c-Myc [141, 146], have been shown to be critical for proliferation of stage 1 and/or stage 2 iNKT cells. It should be noted that although PLZF expression is required for proper development of stage 1 and stage 2 NKTs, it doesn't seem to influence the proliferation capacity of these cells [59]. I have demonstrated that high E protein levels achieved by deletion of both Id2 and Id3 leads to a 3 fold increase in proliferation (and hence number) of stage 1 CD4⁺ iNKTs and DN TCR β^{lo} type II NKTs [70]. As mentioned in the previous section, this increase in iNKT cells is dependent on E protein levels, as deletion of one copy each of E2A and HEB in Id2-Id3 deficient mice leads to an expansion of $\gamma\delta$ NKT cells instead [70]. A similar increase in total NKTs has also been reported upon deletion of 3 out of 4 Id alleles [70, 85]. Further, a mouse model with E2A and HEB deletion at the DP stage along with a V α 14 TCR transgene (to study roles of E proteins in an iNKT selection-independent context) showed a 50% reduction in iNKT numbers and a block beyond stage 0 partly due to defective proliferation in stage 1 cells

[85]. It was also noted that there was an expansion of Id3-deficient iNKT cells at stages 2 and 3 in Id3 deficient/wild type mixed bone marrow chimeras, suggesting that Id3 limits iNKT expansion at not just stage 1, but at all stages of iNKT development. This increase in iNKT numbers was explained by increased proliferation in Id3-deficient cells, particularly at stage 2. Loss of Id2 however led to no change in iNKT proliferation [85].

The iNKT expansion phase allows the lineage to make up for the limited numbers of CD1d-selected cells initially and generate high copy numbers, which are important for quick response to antigens. It is interesting to note that both E protein and c-Myc play roles in iNKT proliferation [85]. However, it is unclear whether E proteins upregulate c-Myc or if they act in a concerted manner as observed in Burkitt's lymphoma [85, 147]. Further, it has been suggested that E proteins may upregulate Id3 during this phase to keep a check on the proliferation [85]. Given that Id3 limits iNKT expansion, it is yet to be determined if its expression needs to be reduced upon selection to increase iNKT numbers [85], and if so, how it is suppressed post TCR signaling. It has been suggested that the induction of the ITK/RAS/ERK (interleukin-2-inducible T-cell kinase/ Ras/ Extracellular-signal-regulated kinases) pathway, SLAM/SAP signaling or β -catenin may play a role in this [84]. Since Id2 can negatively regulate Id3 expression [85, 148], it is also possible that Id2 upregulation through cytokine signaling [148] suppresses Id3 expression to allow proliferation. Therefore, high E protein levels promote iNKT

population expansion during stages 1 and 2 of iNKT development, while Id3 limits iNKT cell numbers.

1.3.2.4 iNKT differentiation and maturation

The emergence of iNKT subsets has not only changed the paradigm of a linear iNKT developmental program, but has also enabled study of subset-specific roles of NKTs in disease contexts and expanded the potential applications of these cells as adjuvants to prime other cell types. E and Id proteins can therefore be imagined to play distinct roles in influencing the linear part of iNKT development until stage 1, which includes lineage selection, commitment and expansion. Beyond stage 2, E and Id proteins play subset specific roles as differentiation into distinct iNKT subsets requires unique transcriptional programs (Step 5 and 6, Figure 6). The more recently discovered regulatory iNKT (or NKT10) cells are not yet well characterized, and hence are not discussed here [149, 150].

For iNKT1 development, both Id2 and Id3 have been demonstrated to be important as there is a significant reduction in the frequency of Id3-deficient T-bet⁺ iNKT1 cells in Id3^{+/+}/Id3^{-/-} mixed bone marrow chimeras, as well as reduced frequency and numbers of Id2-deficient iNKT1 cells in Id2^{+/+}/Id2^{-/-} mixed bone marrow chimeras [85]. This suggests that wild-type level of E protein activity, with both Id2 and Id3 present, supports iNKT1 development, as observed in B6 mice. Further, the expression of T-bet, the critical transcription factor for iNKT1 development as well as IFN- γ

production and IL-2R β expression for stage 3 maturation [53, 151, 152], has also been deemed a critical checkpoint in iNKT cell development [84]. Stage 2 iNKTs in ET-2 mice with sustained high E protein levels throughout iNKT cell development failed to upregulate T-bet, causing a block in stage 3 maturation with impaired IFN- γ production [86]. These data suggest that E proteins can directly or indirectly suppress T-bet expression to influence iNKT maturation and effector functions. Id2 or Id3 deficiency in the BM chimeras described above also result in increased PLZF but impaired T-bet expression, explaining the decrease in the iNKT1 population [85]. Id3^{-/-} mice also have reduced numbers of stage 3 cells and IFN- γ producing iNKTs, along with significantly aberrant gene expression from as early as stages 1 and 2 [84]. Based on the expression profile of Id3 (high in stages 1, 2 and iNKT2 cells, but low in stage 3) and its requirement for the development of stage 3 iNKT1 cells, Id3 has been proposed to be important at earlier stages for initiation of the stage 3 maturation program. Deficiency of both Id2 and Id3 also leads to fewer IFN- γ only producing (i.e. iNKT1) cells [84], but it is still unclear whether Id2 and Id3 play distinct roles in the development of the subset, or the observations are merely artifacts of functional compensation by Id2 and Id3. Moreover, one of the early pieces of evidence suggesting that Id or E proteins play a role in iNKT development was the role of Id2 in maintaining the homeostasis of hepatic NKTs [135], which are now known to be mostly iNKT1 cells. These cells are drastically reduced in Id2-deficient mice due to greater apoptosis caused by a reduction in CXCR6 (chemokine

(C-X-C motif) receptor 6), Bcl-2 (B-cell lymphoma 2), Bcl-xL and the inability to counter the pro-apoptotic protein BIM, which is also an E protein target.

The role of Id3 in iNKT2 development is ambiguous as Id3 deficiency leads to a larger population of PLZF^{hi} iNKT2 cells, but these cells also have the highest Id3 expression [85]. It is possible that the high level of Id3 is essential for selection into the iNKT2 subset, but limits the numbers after selection by inhibiting E proteins that would otherwise drive increased differentiation into this subset. In support of this idea, ET-2 mice with sustained E protein expression have greater numbers of iNKT2 cells [86]. Along similar lines, iNKT17 cells are also promoted by high E protein levels, as reported by increase in this subset in ET-2 mice [86]. It is interesting to note that this subset has low to medium expression of both Id2 and Id3, as opposed to a dominant expression of one Id protein, as seen in the other subsets [85]. Also, this increase in iNKT2 and iNKT17 cells is accompanied by a reduction in the iNKT1 population, indicating that higher than normal E protein levels promote iNKT2 and iNKT17 subset differentiation at the expense of the otherwise predominant iNKT1 lineage [86]. Moreover, retroviral overexpression of T-bet in immature NKTs from these mice can rescue IFN- γ production and some other features of iNKT1 cells, hinting at the plasticity of these subsets. All the above data suggest that E and Id protein levels can differentially influence the maturation and development of iNKT subsets. High E protein levels promote iNKT2 and iNKT17 development, whereas Id proteins are necessary for iNKT1 development.

The implications of these findings can be better understood by bearing in mind their resemblance to T helper subsets. For instance, E proteins are capable of directly regulating GATA3 expression [153], and Id2-deficiency in mice causes a Th2-dominant phenotype [154]. Id3 and E proteins have also been demonstrated to be critical for Th17 development by promoting ROR γ t expression and blocking IL-4 production [155]. It is interesting to note that besides influencing iNKT differentiation, high E protein levels were also shown to impair thymic exit, causing a reduction in peripheral iNKT cells [86]. Overall, these results indicate that E and Id proteins can regulate key transcription factors to bias iNKT subset differentiation and their effector function.

Table 1: Role of E and Id proteins in iNKT cell development.

E and Id proteins regulate the iNKT lineage at several key developmental steps which are marked in figure 6, and summarized here.

Step No.	Developmental step	E or Id protein role
1	Pre-TCR / β selection : DN3b stage	<ul style="list-style-type: none"> High E proteins suppress selection [70]
2	iNKT TCR rearrangement/ iNKT selection : DP stage	<ul style="list-style-type: none"> HEB promotes secondary rearrangement through increased survival [125] Id3 prevents secondary rearrangement by turning off <i>Rag</i> [84] Id2 and Id3 prevent early $V\alpha 14$-$J\alpha 18$ rearrangement [70]
3	iNKT program initiation: iNKT Stage 0	<ul style="list-style-type: none"> E proteins bind to PLZF directly and promote expression [85] Id protein deficiency lead to lower PLZF [84]
4	iNKT proliferation : mainly Stage 1	<ul style="list-style-type: none"> Id2 and Id3 limit iNKT proliferation [70] Id3 limits iNKT numbers through all development stages [85]
5	iNKT subset differentiation : Stage 2	<ul style="list-style-type: none"> iNKT1 development needs Id2 and Id3; iNKT2 cells expand on Id3 deficiency [85] iNKT2 and iNKT17 are supported by high E protein (or low Id proteins) [86]
6	iNKT Stage 3 maturation	<ul style="list-style-type: none"> Id3 needed at Stages 1 and 2 for initiating appropriate Stage 3 program [84] E proteins prevent T-bet expression which is needed for maturation [85]

1.4 Id proteins and cancer

Majority of this chapter is adapted from the publication "Paradoxical role of Id proteins in regulating tumorigenic potential of lymphoid cells", Sumedha Roy, and Yuan Zhuang, Frontiers of Medicine. 2018 (Under Review for special issue on "Cancer and Immunology")

The relationship between Id proteins and cancer is complex, necessitating a thorough understanding of the context of study. Since their discovery almost three decades ago, Id proteins have predominantly been associated with the maintenance of a state of anaplasia or dedifferentiation in cells [92, 93, 156-158]. They are known to facilitate this through inhibition of functions of various classes of ubiquitous and tissue-specific Helix-loop-Helix (HLH) transcription factors, including E proteins, ETS, PAX, MYOD and RB proteins. The role of Id proteins in preventing cell differentiation and promoting stem cell behavior has begged the question regarding their potential in giving rise to cancer stem cells, and therefore as oncogenes [92, 93, 158, 159]. In support of their oncogenic functions, Id proteins have been reported to have high expression in multiple tumor types, and to contribute to dedifferentiation, angiogenesis and metastasis of tumors. The upregulation of Id proteins can be mediated by upstream oncogenic events that promote their overexpression, increase their stability, or abolish their suppression by tumor suppressors. However, gain-of-function alterations in Id genes in human tumors, that can conclusively verify the oncogenic nature of Id proteins, are a key piece to the puzzle that is still missing. Nonetheless, efforts spearheading the inhibition of Id proteins as potential therapeutic targets are underway [92, 93, 96][160-164]. The notion of Id proteins acting as oncogenes, however, has also now been challenged by recent data demonstrating their tumor suppressive roles in some cancers, most prominently in Burkitt's lymphoma [165-167]. Id3^{-/-} mice have also been reported

to develop $\gamma\delta$ Hepatosplenic T-cell lymphoma (HSTCL) as a consequence of $\gamma\delta$ NKT cell population expansion [168]. The overall role of Id proteins in normal development and tumorigenic pathways has been reviewed extensively [92, 93, 156-159].

Considering the different potential interaction partners and cell-specific functions of Id proteins, it is hard to uniformly and exhaustively predict the role of Id proteins in distinct tumor types. Blood cancers, including lymphomas and leukemias, are estimated to constitute up to 8% of new cancer cases diagnosed this year (Lymphoma and Leukemia Society statistics). Phenotypic similarities between lymphocyte-derived tumors that develop in Id-overexpression and E protein-null mutant mice suggest that Id proteins often mediate their functions in these tumors through inhibition of E protein activity [158]. In this section, I have summarized the underlying pathways that drive E and Id-mediated tumorigenesis or tumor suppression in T cells.

Overall E and Id proteins act as gatekeepers or “brakes” to ensure smooth regulation of cell proliferation and differentiation during development. The appropriate expression of, as well as the interaction between, E and Id proteins tightly control developmental outcomes, and therefore, their dysregulation causes the brakes to “fail”, often resulting in cells being predisposed to malignant transformation. It is crucial to understand the various factors that can determine physiological versus tumorigenic outcomes driven by E and Id proteins (Figure 7).

While I do not wish to enumerate all observations that have led to predominant beliefs about the oncogenic roles of Id proteins, and corresponding tumor suppressor roles of E proteins, it is worth mentioning a few of them to provide a historical perspective. The observation of induction of high Id expression in cells that are undifferentiated and proliferating supported the idea that Id proteins could contribute to dedifferentiation [158, 169]. Even though just overexpression of Id proteins was not enough for immortalization of cell lines, it supported a pluripotent, stem cell-like fate in the presence of leukemia inhibitor factor (LIF) [93, 170]. It was proposed that Id2 and Id3 may act together to control cyclin E (*Ccne1*) and cyclin A (*Ccna2*) [156]. Id proteins were also found to suppress the regulators of cell senescence, such as *Rb* and cyclin dependent kinase inhibitors *p16INK4a*, *p21* and *p27*, which can be activated by Ets and E protein family members [30, 171]. In stark contrast, E2A was found to activate *p21*, and to play a role in apoptosis and growth suppression in NIH3T3 cells [172, 173]. A forced MyoD-E47 dimer that can no longer bind to Id proteins can overcome the effects of Id1 overexpression to promote differentiation in target cells [156, 174]. Id protein half-life is also extended by binding to E2A [175]. Therefore, the dynamics between E and Id proteins is crucial in maintaining a balance between cell proliferation, survival and differentiation either through their direct regulation of relevant downstream targets, or through indirect inhibition of binding to other HLH interaction partners, such as RB and ETS proteins.

An oncogene is typically defined as a gene, which when mutated, confer cells with properties that facilitate adoption of a tumorigenic program [93]. Id proteins can upregulate anti-apoptotic and pro-survival factors that overcome programmed cell death, promoting tumorigenesis. In many cases, Id proteins are often dysregulated by upstream oncogenic events, such as Myc, Ras and Notch signaling, and also inhibited by tumor suppressor genes. The activation of Id3 by oncogenic Ras signaling leads to E protein inhibition, tipping the balance. This inhibition via Ras supports the role of E proteins as tumor suppressors [30, 104]. However, overexpression of Id proteins is not enough to cause tumorigenesis in mice unless combined with anti-apoptotic genes [170, 173, 176, 177]. It is also dangerous to draw conclusions about the oncogenic roles of E or Id proteins based on their expression levels or physiological activity, since these may be impaired or hijacked, irrespective of expression patterns. For instance, Id3 is overexpressed several fold in Burkitt lymphoma samples [147, 165, 166], but its function of mediating E protein inhibition is blocked through mutations in Id3 or Tcf3, revealing the tumor suppressive function, instead of the oncogenic role of Id3. Therefore, RNA expression data alone without validation of protein functions may not be sufficient to determine whether Id3 plays a tumor suppressing or promoting role.

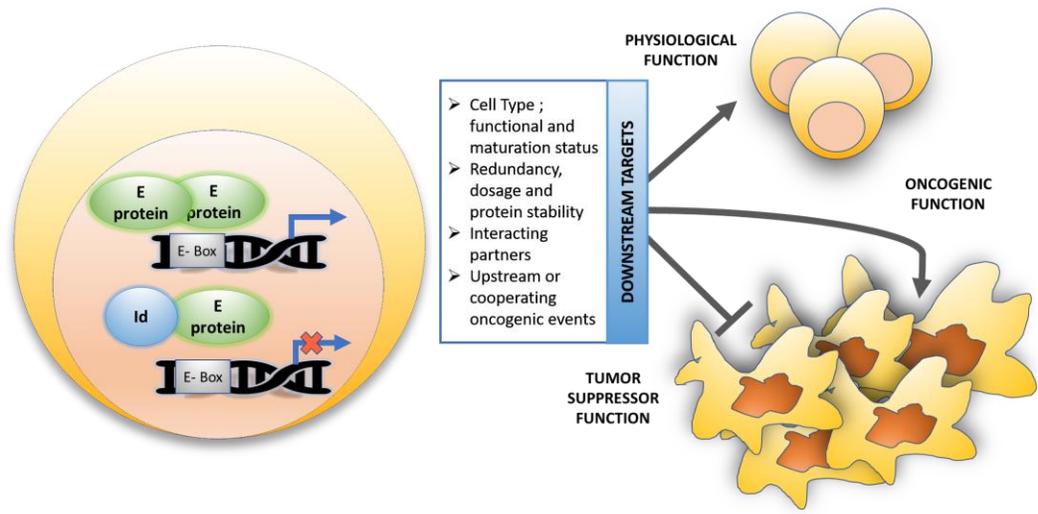


Figure 7: Context-dependent roles of E and Id proteins in development and cancer

The interaction between E and Id proteins is represented in a cell on the right, and the various factors that determine the regulated or dysregulated downstream targets are listed, which ultimately determine whether E and Id proteins are able to perform their intended physiological function, or whether they play oncogenic or tumor suppressive roles.

Various factors can determine the downstream targets of E and Id proteins, and therefore their function in physiological development, tumor suppression or tumorigenesis (Figure 1). This includes important considerations such as the target cell type, dosage of E proteins, or transient upregulation of Id proteins, and functional redundancy between family members. Given the complexity that determines the role of Id proteins in cancers, inhibition of Id proteins may be clinically beneficial in certain tumor settings, but not in others.

1.4.1 Tumor suppressive roles of E2A in T cell lymphomas

Malignant T cells can give rise to different kinds of lymphomas and leukemia in humans. A recurrent gene deletion is found in E2A in 70% of patients with Sezary syndrome, a type of T cell lymphoma, which is the first report demonstrating inactivating alterations in the E2A gene in human lymphomas [178]. This gene deletion results in upregulation of the cell cycle regulator *Cdk6*, which is a known E47 target [116] and can execute the anti-apoptotic and proliferative effects of NOTCH in these tumors. Interestingly, all chromosomal losses of E2A detected in these tumors were heterozygous. The authors propose that one copy of E2A is necessary for the survival of these cells, and therefore homozygous mutations are selected against. The remaining copy of WT E2A may also be post-translationally degraded by NOTCH1 [178].

Anaplastic Large cell lymphoma (ALCL) is a type of Peripheral T Cell Lymphoma (PTCL), where gene translocations are commonly observed. It was reported that the translocation that gives rise to an NPM-ALK fusion protein, also leads to abnormal upregulation of genes that are proximal to the breakpoint, including *Id2*. This *Id2* overexpression impairs E2A-mediated regulation of T cell specific genes like *Cd3*, *Lck*, *Fyn*, *Tcf7*, *Tbx21*, and *Gata3*, explaining the loss of T cell phenotype observed in ALCL cases. This supports tumor suppressor activity of E2A that is sensitive to Id protein inhibition [179].

T-ALL, or T-cell acute lymphoblastic leukemia, is a classic example for the role played by E and Id proteins in the tumorigenic process. Activation of the *Tal1* (also known as SCL) or the *Tal2* genes is often a major causal factor for T-ALL, and is achieved by translocation and other mechanisms in over 60% of T-ALL patients [180-182]. TAL1 is a bHLH transcription factor that forms an obligate heterodimer with E2A and HEB. It is typically required for HSC generation, and is not expressed in the thymus. Human T-ALLs show no sign of changes in E2A expression, but it has long been hypothesized that TAL1 inactivates E2A or usurps its tumor suppressor activity to give rise to T-ALL. Similarly, *Lyl1* is another gene that is activated in T-ALL and can bind to E2A, but results in regulation of different target genes [108]. TAL1 and LYL1 are involved in the maintenance of short-term and long-term HSCs respectively, where they each function with E proteins in a complex. This may be reflective of gene targets that drive T-ALL [89]. The mechanisms of T-ALL tumorigenesis have been explored by using mouse models and human lymphoma cell lines. Aberrant TAL1 overexpression in human tumor as well as mouse models gives rise to a developmental block at the DP stage [182]. TAL1 is found to mediate lymphoma development through a regulatory complex that comprises of E2A, HEB, LMO1/2, GATA3, RUNX1 and TAL1, where the latter three transcription factors are also part of a positive autoregulatory loop. Knockdown of these components inhibits cell growth and induces apoptosis [182]. Haplo-insufficiency of either E2A or HEB leads to a block in development at the DN

stage, and accelerated T-ALL development in mice with *Tal1* overexpression [180]. However, the dosage of E proteins is critical for lymphomagenesis as both E2A and HEB are required for the TAL1 complex. Co-repression of typical E2A targets was also observed, partly mediated by the mSin3A/HDAC1 complex.

A study done in the Jurkat T-ALL cell line, where *Tal1* is overexpressed and complexed with E2A, found low E2A transcriptional activity for an E-box reporter gene. Further, cells transduced with E-T/2 (E2A-Tal1) to restore E protein activity, underwent growth arrest and apoptosis, supporting an inhibition model where Tal1 inhibits the downstream tumor suppressive genes of E2A, such as p21, which could not be induced in Jurkat T cells [181]. While some normal E2A targets were still conserved in T-ALL tumor cell lines, such as *Ptcra*, *Notch3*, *Rag1/2* and *Gfi1*, TAL1 was found to promote some genes that are otherwise suppressed by E2A and HEB, and vice versa. One of those genes, TRIB2, is normally repressed by E proteins, but upregulated in the context of TAL1 complex, perhaps through TAL1 and NOTCH1 activity. TRIB2 has been reported as an oncogene in AML, and has now been shown to be important for T-ALL survival through downstream activation of XIAP, inhibitor of apoptosis [182, 183]. TRIB2 can also control the expression of TAL1 partners GATA3 and RUNX1, and destabilizes E2A through proteasomes [183]. These observations support the tumor suppressor role of E2A in T-ALL.

Independent of the tumor suppressor roles of E2A observed in T-ALL, it was observed that E2A-deficient mice exhibit an early block in T cell development and develop lymphomas derived from immature DP or SP thymocytes [108, 109, 178]. Tumors, but not premalignant cells, overexpress c-Myc, which can possibly be explained by the gain of an extra copy of chromosome 15. But it is unclear if E2A and c-Myc work together, or if c-Myc is the sole culprit in this lymphoma model [108]. These thymocytes aren't hyperproliferating or resistant to apoptosis. Overexpression of E2A in cell lines derived from lymphomas in E2A-deficient mice led to programmed cell death in the cells rather than growth arrest [114]. On the other hand, homozygous knockout of Id1 in these E2A-null mice improved survival but they still failed to rescue the tumor phenotype. This can be explained by a dominant tumor suppressor role of E2A, rather than being a direct Id-mediated effect [109]. The function of Notch acting together with, or in opposition to E2A in T cell development is still open to debate [104]. Their role in E2A-deficient or T cell lymphomas is equally debatable [184-188]. Notch has been proposed to represent a second hit in T-ALL as Notch1 and Notch3 are found to be aberrantly activated by translocations and somatic mutations in a large fraction of T-ALL cases, where the tumor depends on Notch signaling for survival [187]. Others have reported that Notch activity is not essential for tumorigenesis in T-cell derived tumors with impaired E protein activity [185].

1.4.2 Paradoxical role of Id proteins in T cell lymphomas

Since E proteins were suggested to play tumor suppressive roles in T cell-derived human and murine cancer models, it was natural to assume the oncogenic roles of Id proteins in these contexts, except in T-ALL where TAL1 overexpression modifies E2A targets through Id-independent mechanisms. The overexpression of Id2 in ALCL, PTCL and multiple other types of T cell lymphomas, and its induction by Myc, supported this hypothesis [189]. In mouse models expressing an Id2 transgene, there is a developmental block at early steps of T cell development, accompanied by hyperproliferating T cell lymphomas in most mice [117]. These tumors have high level of *Myc* expression upon malignant transformation, but not in premalignant stage, therefore c-Myc may be important for the tumor, but is not driven by Id2 overexpression. Not all mice with the Id2 transgene develop tumor, suggesting that it is an important but not sufficient event for causing tumorigenesis. It is likely that overexpression of *Id2* (or E protein deficiency) blocks further differentiation in proliferating cells, predisposing them to secondary mutations, that represent either early transformation events after which the T cells can independently rearrange their TCR, or divergent transformation events, that give rise to polyclonal tumors [117]. A similar phenotype is observed in Id1 transgenic mice, such that there is a severe block at T cell progenitor step with massive apoptosis triggered in most T cells, ultimately leading to lymphoma development. c-Myc may play a role in this phenotype since it can function both as an oncogene, and

inducer of apoptosis. Induction of p21 by E2A may partly be able to block cell cycle progression in these cells [118]. These studies support the oncogenic roles of Id proteins in T cell-derived lymphoma development.

In direct contrast, there is also strong evidence in favor of tumor suppressor roles played by Id proteins in a subset of lymphomas derived from innate-like T cells. Knocking out Id3 in mice gives rise to $\gamma\delta$ T cell lymphomas, resembling hepatosplenic $\gamma\delta$ T cell lymphomas (HSTCL) in humans. There is no dysregulation of genes that are commonly found in $\alpha\beta$ T cell lymphomas, such as those described in the previous sections, including *Lyl1*, *Tal1* and *p21*. However, an increase in Myc expression was observed among some samples [168]. We have recently found that Id2 and Id3 also play tumor suppressor roles in invariant natural killer T (iNKT-) and innate-like tumors in mice [190]. These lymphomas develop much more rapidly than the $\gamma\delta$ T cell lymphomas in Id3-deficient mice, and display signs of chromosomal instability. As described in section 4.2.4, the dichotomy observed in the gene expression programs for premalignant innate-like T cells in these mice is interesting – there is both downregulation and upregulation of tumor suppressor, anti-proliferative and cell cycle arrest genes, demonstrating a somewhat “balanced” state. However, after a presumed second hit, lymphoma cells have enrichment for genes often dysregulated in cancers, cytokine-cytokine interaction genes and NF- κ B signaling, similar to pathways reported in human NK/T tumors. Similarly, deficiency of Id2 and/or Id3 in mice also causes an expansion of

innate variant T_{FH}-like cells, and $\alpha\beta$ T cell lymphomas in these Id2/Id3-deficient mice that display an increase in Myc expression and reduction in the tumor suppressor Cdkn2a [103, 137]. There is evidence to suggest tumor suppressor functions of *Id4* in T cell tumors. Exogenous expression of Id4 in lymphoid tumor cell lines induces caspase-dependent apoptosis. Additionally, methylation is found at the Id4 promoter locus in multiple tumor models, including chronic lymphocytic leukemia (CLL), and a mouse model of T/NK acute lymphoblastic leukemia upon malignant transformation, that correlates with reduced Id4 expression and lower patient survival [191-194]. Therefore, Id proteins can play both oncogenic and tumor suppressive roles in lymphomagenesis of T cells.

2. Materials and methods

2.1 Mice

Id2^{fl/fl}Id3^{fl/fl}LckCre⁺ (L-DKO), Id3^{-/-} and Id3^{-/-} TCRδ^{-/-} mice were generated as previously described [70, 195]. CD1d^{-/-} mice were purchased from Jackson Laboratory (Strain 008881), and bred with L-DKO mice to generate L-DKO CD1d^{-/-} mice. L-DKO pTα^{-/-} mice were generated by breeding L-DKO mice with pTα^{-/-} mice [196], which were a generous gift from David L. Weist (Fox Chase Cancer Center, Philadelphia). All mice were bred in a specific pathogen-free facility of Duke University Division of Laboratory Animal Resources, and all procedures were performed according to protocols approved by the Institutional Animal Care and Use Committee.

2.2 Cell sorting and RNA extraction

All cells were sorted in FACS buffer using a MoFlo XDP cell sorter. Total mRNA from sorted cells was extracted using an RNAqueous Kit (Life Technology) according to manufacturer's protocol.

2.3 ChIP-Seq analysis

26 × 10⁶ iNKT and 30 × 10⁶ DP cells were sorted and pooled from multiple L-DKO mice for the E2A ChIP-Seq analysis. 2 × 10⁶ γδ T cells were also sorted from Id3^{-/-} mice for H3K4me2 analysis. iNKT (CD1dTet⁺ TCRβ⁺) and DP (CD1dTet⁻ CD4⁺ CD8⁺) cells were sorted from 3-5 weeks old L-DKO mice. For H3K4me2 ChIP-Seq, cells were fixed with 1% formaldehyde, whereas for E2A ChIP-Seq, cells were additionally fixed with 1.5

mM EGS (ethylene glycol-bis(succinic acid N-hydroxysuccinimide ester)). Crosslinked cells were lysed, nuclei were extracted, and sonicated using Bioruptor Plus (Diagenode) and immunoprecipitated with E2A (V-18, Santa Cruz Biotechnology, Lot G0814) or H3K4me2 (Millipore, 07-030, Lot 2430486) antibody. After elution and reverse crosslinking, RNA and proteins were digested, followed by DNA purification using a ChIP DNA Clean and Concentrator kit (Zymoresearch). Libraries were prepared with the NEBNext primer set, which included subjecting ChIP DNA to end repair, A-tailing, adapter ligation and PCR amplification. Samples were cleaned and size selected by 8% PAGE or AMPure beads (Agencourt). Sequencing was done on HiSeq4000 platform (Illumina).

2.4 RNA-Seq analysis

iNKT (CD1dTet⁺ TCRβ⁺), γδNKT (TCRγδ⁺ CD3⁺) and DP (TCRγδ⁻ CD4⁺ CD8⁺) cells were sorted from 4-5 weeks old WT (total 2 x 10⁵ γδ T, 2 x 10⁵ iNKT cells, 2 x 10⁶ DP cells), Id3^{-/-} (total 15 x 10⁶ γδ T cells) or L-DKO (total 15 x 10⁶ iNKT cells, 15 x 10⁶ DP cells) mice. After RNA extraction using the RNAqueous kit, quality was assessed using the Agilent Bioanalyzer RNA Pico chip. Ribosomal RNA was depleted using the RiboErase method from Kapa Biosystems. In short, 1ug of total RNA was hybridized with 1ug of hybridization oligos tiling the 18s, 28s, 5.8s and mitochondrial rRNA sequences. Each sample was then RNaseH treated to degrade complementary rRNA sequence. The product was cleaned and purified using 2.2X AMPure beads (Agencourt). The cleaned

product was DNase treated to degrade the DNA oligo mix. The remaining rRNA depleted samples were then purified using 2.2X AMPure XP beads. The Kapa Stranded RNA-Seq Kit was used to generate stranded Illumina sequencing libraries (Kapa Biosystems). RNA from was fragmented at 94°C for 6 minutes. Briefly, RNA was hybridized to random primers, followed by first-strand cDNA synthesis, second-strand cDNA synthesis with marking, A-tailing, ligation of Illumina paired-end adapters with 8 bp barcodes, and 9 cycles of PCR amplification. Reactions were purified with Agencourt AMPure XP beads where necessary. Libraries were multiplexed in equimolar amounts, and sequenced as paired-end 50-bp reads using a HiSeq2500 platform (Illumina).

A second, independent round of RNA-Seq was done with DP (TCR $\gamma\delta$ ⁻ CD4⁺ CD8⁺) cells sorted from 5 weeks old WT (9×10^5 cells), pT α KO (4.5×10^5 cells), L-DKO (9×10^5 cells) and L-DKO pT α KO (9×10^5 cells) mice. 150 bp paired-end sequencing was done on the HiSeq2500 platform (Illumina).

2.5 TCR α repertoire sequencing

Pre-selection DP (CD4⁺ CD8⁺ CD69⁻) cells were sorted from 3-4 weeks old CD1d^{-/-} or TKO mice. RNA was extracted from sorted cells, and reverse transcribed into cDNA by murine leukemia virus reverse transcriptase (Life Technology). Sequences specific for V α 8⁺ and V α 14⁺ cells were isolated and amplified using nested PCR with V α -specific and C α primers, followed by indexed V α primers. Barcoded sequences were finally

amplified with common adapter-specific primers, gel purified, and sequenced using Ion Torrent technology (Applied Biosystems).

Primers used for V α 8- and V α 14-specific amplification (20 cycles) were as follows: V α 8/ C α (forward) 5' TCCACGCCACTCTCCATAAG 3', (reverse) 5' ACAGCAGGTTCTGGGTTCTG 3'; V α 14/ C α (forward) 5' GCAACTCTGGATAAAGATGC 3', (reverse) 5' GTCGGTGAACAGGCAGAGG 3'. For barcoding individual samples, PCR product from the previous step was diluted 1:2 times and amplified for 25 cycles, using the designed barcoded primer (forward) 5' CCATCTCATCCCTGCGTGTCTCCGACTCAG|“Barcode X”| Barcode Adapter| V α 8 or V α 14 3', and reverse adapter attached with C α primer (reverse) 5' CCTCTCTATGGGCAGTCGGTGATACAGCAGGTTCTGGGTTCTG 3'. Finally, the successfully barcoded samples were amplified for 25 cycles using primers corresponding to adapters (forward) 5' CCATCTCATCCCTGCGTGTC 3', (reverse) 5' CCTCTCTATGGGCAGTCGGTGAT 3'.

2.6 Flow cytometry

Surface marker antibodies were used according to manufacturer's protocol (Biolegend). Intracellular staining with PLZF antibody (eBioscience) was done using the Foxp3 staining buffer kit (eBioscience). CD1d tetramers were received from the Tetramer Facility of the National Institutes of Health. Stained samples were run on a FACSCanto II machine (BD Biosciences) and data was further analyzed with FlowJo software (Tree

Star). Bar graphs were drawn using GraphPad Prism (GraphPad Software). Two-tailed student's t-test was used for statistics, with p values less than 0.05 considered significant.

2.7 Anti-CD3 injection

4 week old pT $\alpha^{-/-}$ and L-DKO pT $\alpha^{-/-}$ were intraperitoneally injected with a 200 μ L sterile mix containing 150 μ L of 1mg/mL LEAF purified α -CD3 (Biolegend) and 50 μ L of PBS. Cre- and pT $\alpha^{-/-}$ control littermates were simultaneously injected with 200 μ L sterile PBS. Mice were sacrificed 10 days after injection, thymi were harvested and analyzed by flow cytometry.

2.8 Histopathological analysis

Tissue sections were removed immediately after sacrificing the mice, and fixed in 10% phosphate-buffered saline (PBS)-formalin. Embedding, sectioning and staining (hematoxylin and eosin, and Masson's Trichrome) were done by the Pathology service core at Duke University.

2.9 Adoptive transfer of lymphoma cells

Enlarged thymi from L-DKO donor mice were minced in PBS with 5% BCS, filtered, lysed for red blood cells using BD Pharm Lyse lysing buffer (BD Bioscience) and washed with PBS. 5 - 7 weeks old WT mice were sublethally irradiated with 300 rads, and injected with 5×10^6 tumor cells 24 hours after irradiation. Recipient mice were

sacrificed 4 - 10 weeks after transfer, and tissues were collected for FACS analysis and H&E staining.

2.10 PCR

Genotyping of mice was done as described previously [34]. Total RNA was extracted using an RNAqueous Kit (Life Technology) according to manufacturer's protocol, and reverse transcribed into cDNA by murine leukemia virus reverse transcriptase (Life Technology). Primers are listed in Table 2.

Table 2: List of genotyping primers

Gene	Forward Primer	Reverse Primer
Id2 ^{fl/fl}	5' TGTGCATAATTAATCGCATCA 3'	5' TTGGGAAGTCACATTTGTAGTG 3'
Id3 ^{fl/fl}	5' GCTCTGAGGTCATAAATCCC 3'	5' CCATTTGGTTCTATGTATGCCCGTG 3'
LckCre	5' GCAGGAAGTGGGTAAGTACTAGACTAAC 3'	5' TCTCCCACCGTCAGTACGTGAGATATC 3'
Id3 WT	5'-GTT TTG AAC ATA GGT CTG CC-3'	5'-CAC CGG GCT CAG CGC CTT CAT-3'
Id3 ^{-/-}	5'-GTT TTG AAC ATA GGT CTG CC-3'	5'-TCG CAG CGC ATC GCC TTC TA-3'
TCR δ WT	5'-TGAGGGAGGAAGGAAAGGAGAGAC-3'	5'-TCG CAG CGC ATC GCC TTC TA-3'
TCR δ ^{-/-}	5'-TGAGGGAGGAAGGAAAGGAGAGAC-3'	5' - AAAGTGGAAAGAGGGAGCGGTG-3'
CD1d WT	5' AGGGCTGTGTAGAAGTCTGGCGCTA 3'	5' GCAGGGAGCGGAAGGTGTAATT 3'
CD1d ^{-/-}	5' AGGGCCAGCTCATTCCTCCACT 3'	5' GCAGGGAGCGGAAGGTGTAATT 3'
pT α WT	5'-CAC ACT GCT AGA TGG AA-3'	5'-GCG TTT GCC ACA GAG TAA AG-3'
pT α ^{-/-}	5'-CAC ACT GCT AGA TGG AA-3'	5'-CCA CAG TCG ATG AAT CCA GAA A-3'

2.11 Microarray

Pre-malignant iNKT cells were sorted as TCR β +CD1dTet+ cells from L-DKO mice at 20 days of age. Lymphoma cells were sorted from tissues of 18-37 week old mice as T cells that are CD1dTet+ or CD1dTet-. Total RNA was extracted as described for RT-PCR. mRNA expression profiling was done by the Duke Microarray Core Facility using GeneChip Mouse Genome 430A 2.0 arrays (Affymetrix).

2.12 Bioinformatics analysis

2.12.1 ChIP-Seq

ChIP-Seq sequenced reads were aligned to the mm9 genome using Bowtie [197] software (version 1.1.2, parameters: `--chunkmbs 128 --mm -m1 --best --strata -p4 -S -q`). For alignment to the TCR locus, up to 3 repeats were allowed using `-k 3` option along with default options. Peaks were called using MACS [198] (version 1.4.2, default parameters). Peaks were then annotated by the NGS: Peak Annotation tool on Nebula [199]. Bed and wiggle files were generated by MACS for visualization using the Integrative Genomic Viewer [200]. De-novo motif analysis was done using the findmotifs.pl program with HOMER [201] (v4.7.2, 50 or 200 bp within each peak).

2.12.2 RNA-Seq

RNA-Seq sequencing reads were first trimmed using Trimmomatic [202]. Read alignment was done using Tophat and expression quantification was done using Cufflinks [203]. Log₂ transformed FPKM (fragments per kilobase exon-model per

million reads mapped) were used for downstream analyses. Further filtering of low quality genes, Principal Component Analysis, statistical analysis and visualizations were done using R [204]. Pathway analysis was done using the Molecular Signatures Database (MSigDB) v5.2 [205].

2.12.3 TCR Repertoire analysis

All TCR repertoire analysis was done using IMGT HighV-QUEST and its statistical tool with default parameters [206].

2.12.4 Innate gene signature and network analysis

Raw microarray expression data was requested and downloaded from Immgen for selected subsets: preT_DN3A_Th (DN3a), preT_DN3B_Th (DN3b), T_DN4_Th (DN4), T_DP_Th (DP), T_4SP69+_Th (post-selection CD4SP), iNKT_44-NK1_1-_Th (stage 0 and 1 iNKT cells), Tgd_Th (total thymic $\gamma\delta$ T cells), Tgd_vg1+vd6+24ahi_Th (immature V γ 1.1V δ 6.3 cells). Average gene expression among DN3a, DN3b, DN4, DP and CD4SP cells was assumed to be the reference conventional $\alpha\beta$ T cell population. Total thymic $\gamma\delta$ T cells were considered as reference for conventional $\gamma\delta$ T cell population. Fold change in expression for iNKT and $\gamma\delta$ NKT cells was calculated with respect to the reference conventional $\alpha\beta$ and $\gamma\delta$ T cell populations respectively. Genes that had more than 1.5-fold upregulation or 0.6-fold downregulation among both iNKT and $\gamma\delta$ NKT cells were considered to represent the “innate-like gene signature”. These moderately relaxed fold change parameters allowed us to ensure that maximal numbers

of appropriate genes were captured in this analysis. 189 genes were therefore identified from these specific expression patterns among WT iNKT and $\gamma\delta$ NKT cells. Additionally, I also included 7 other genes - *Tcf3* (E2A), *Id2*, *Id3*, *Lef1*, *Sox13*, *Blk* and *Sox4* – which have been reported to play important roles in iNKT and $\gamma\delta$ NKT lineage development, but did not have expression patterns that fit our criteria, i.e. being significantly upregulated or downregulated in both cell types as compared to reference populations. The total 197 genes constituted our innate-like gene signature, derived from Immgen and literature.

111 of the 197 signature genes were found to be dysregulated in *Id3*^{-/-} $\gamma\delta$ and/or in L-DKO iNKT cells. Other known interactions between these 111 genes were retrieved from GeneMania [207]. 83 of the 111 genes were also identified as E2A targets, which had E2A binding to the enhancer, promoter, intragenic, intergenic or downstream regions of these genes, as annotated by Nebula. These interactions, ChIP-Seq targets and gene expression patterns of the 111 genes were represented as a network using Cytoscape3.4.0 [208].

2.12.5 Gene Set Enrichment Analysis (GSEA)

The GSEA [205, 209] desktop application (v2.0) was used to analyze the log₂FPKM expression patterns in L-DKO and L-DKO pT α DP samples. 9245 genes that were unchanged in L-DKO DP samples as compared to WT DP samples were included in this analysis. Enrichment in L-DKO pT α DP samples over L-DKO samples was determined using weighted, log₂(ratio of classes) parameters and 1000 permutations.

The innate-gene signature included 196 genes derived from Immgen, and listed in Supplementary Table 2. The iNKT development and maturation gene set [152] (Msigdb gene set M18517) and inflammatory responses gene sets (Msigdb gene set M5932) were downloaded from Msigdb and used as is.

2.12.6 Correlation analysis

To determine correlation with *Zbtb16* expression, Pearson and Spearman correlation coefficients were determined for all genes across six samples, including replicates of WT DP, pTαKO DP, L-DKO pTαKO DP and L-DKO DP, as derived from RNA-Seq analysis. Genes with both coefficients greater than or equal to 0.7 were considered to be positively correlated, and those with both coefficients less than or equal to -0.7 were considered to be negatively correlated with *Zbtb16* expression. Scatter plots were generated using a custom R script.

2.12.7 Microarray data analysis

Microarray data for pre-malignant and lymphoma cells was normalized using RMA, and differential analysis was done using the limma package available through Bioconductor [210]. Publicly available, normalized Immgen data for WT cells was requested and downloaded from <http://rstats.immgen.org/DataRequest/> [211]. The two normalized datasets were combined according to the Empirical Bayes method using the web tool ArrayMining (www.arraymining.net) [212].

2.13 Data plotting, visualization and statistics

Gene expression (fold change) heatmaps were generated using Gene-E (<http://www.broadinstitute.org/cancer/software/GENE-E/>). Self-organizing Maps (SOM) were generated by the Partek Genomics Suite made available by the Duke Center for Genome and Computational Biology. Principal Component Analysis (PCA) was performed and plotted using the in-built R functions, `prcomp` and `plot3d`, in the open-source RStudio software. Gene overlaps in the form of Venn diagram was drawn using the `eulerAPE` software [213]. Pathway analysis was done using the HOMER software [201]. Survival curves and bar graphs were drawn using GraphPad Prism (GraphPad Software). Two-tailed student's t-test was used for statistical analyses, with p values less than 0.05 considered significant.

2.14 Data availability

Complete datasets can be accessed from NCBI GEO: E2A ChIP-Seq and RNA-Seq (GSE8984), microarray (GSE83761).

3. Regulation of innate-like T cell fate choice by E and Id proteins

The majority of this chapter is adapted from the publication “Id proteins suppress E2A-driven iNKT cell development prior to TCR selection”, Sumedha Roy, Amanda J. Moore, Cassandra Love, Anupama Reddy, Deepthi Rajagopalan, Sandeep Dave, Leping Li, Cornelis Murre, and Yuan Zhuang, Frontiers in Immunology. 2018 9:42. doi: 10.3389/fimmu.2018.00042

3.1 Introduction

Given the reciprocal nature of Id proteins in supporting conventional T cells and suppressing innate-like T cells, and the function of Id3 in promoting conventional T cell development downstream of the TCR signaling [214], I decided to test the hypothesis that Id proteins control innate-like T cell development through a somewhat distinct mechanism from conventional T cells. Since Id proteins have been shown to significantly modulate E protein activity during early stages of T cell development [87], I predicted that Id-mediated suppression of these innate-like T cells is not only limited to cell expansion after selection and commitment, but can also influence their lineage choice at earlier stages of development.

I tested this hypothesis in multiple Id-deficient models, and found biased V α 14-J α 18 rearrangements and E2A-driven regulation of genes that promote the iNKT lineage in DP cells of Id-deficient mice. Further, a block in pre-TCR signaling hinders

conventional $\alpha\beta$ T cell development but fails to eliminate the expanded innate-like iNKT and $\gamma\delta$ NKT cells in Id-deficient mice. This study reveals a distinct regulatory event that separates iNKT cell lineage from the conventional $\alpha\beta$ T cell lineage prior to the TCR signal. Additionally, I define an E2A-mediated transcription network that supports innate-like iNKT and $\gamma\delta$ NKT lineages.

3.2 Results

3.2.1 Absence of Id proteins allows E2A to induces genes involved in iNKT cell development and function

Our laboratory and others have shown that the loss of function of Id3 or Id2/Id3 results in a significant increase in numbers of iNKT cells [70, 84-86, 139]. I hypothesized that uninhibited E2A activity in the absence of Id proteins may induce genes important for the iNKT developmental program. Therefore, I sought to identify specific downstream gene targets that drive the expansion of iNKT cells in Id2/Id3-deficient (Id2^{fl/fl}Id3^{fl/fl}LckCre⁺, LckCre-mediated double knockout or L-DKO) mice by performing RNA-Seq and E2A ChIP-Seq analysis in L-DKO DP and L-DKO iNKT cells, as representative populations prior to, and after CD1d-mediated selection (Figure 8a). Comparing the transcription profile of L-DKO iNKT cells to wild type (WT) iNKT cells, I found 552 genes to be upregulated by more than 2-fold in L-DKO iNKT cells with respect to WT iNKT cells (Figure 8b). Pathway analysis confirmed significant upregulation of genes related to iNKT differentiation and effector function (Figure 8c,

8d). Genes essential for iNKT development and function, such as *Tcf7*, *Sox4* and *Gzma*, were highly upregulated in iNKT cells deficient in Id proteins [215, 216]. A subset of genes upregulated in L-DKO iNKT cells were also upregulated in L-DKO DP cells compared to WT DP cells (Figure 8d). *Zbtb16*, which is highly expressed in wild type iNKT cells, was found to be prematurely activated in L-DKO DP cells [59].

ChIP-Seq analysis of L-DKO DP and iNKT cells further verified strong E2A binding to the promoter and/or enhancer regions of the highly upregulated genes (Figure 8a, 9), indicating a direct role for E2A in initiating and/or maintaining the transcription of these target genes. Overall, these findings suggested E2A-mediated promotion of iNKT cell development in the absence of Id proteins.

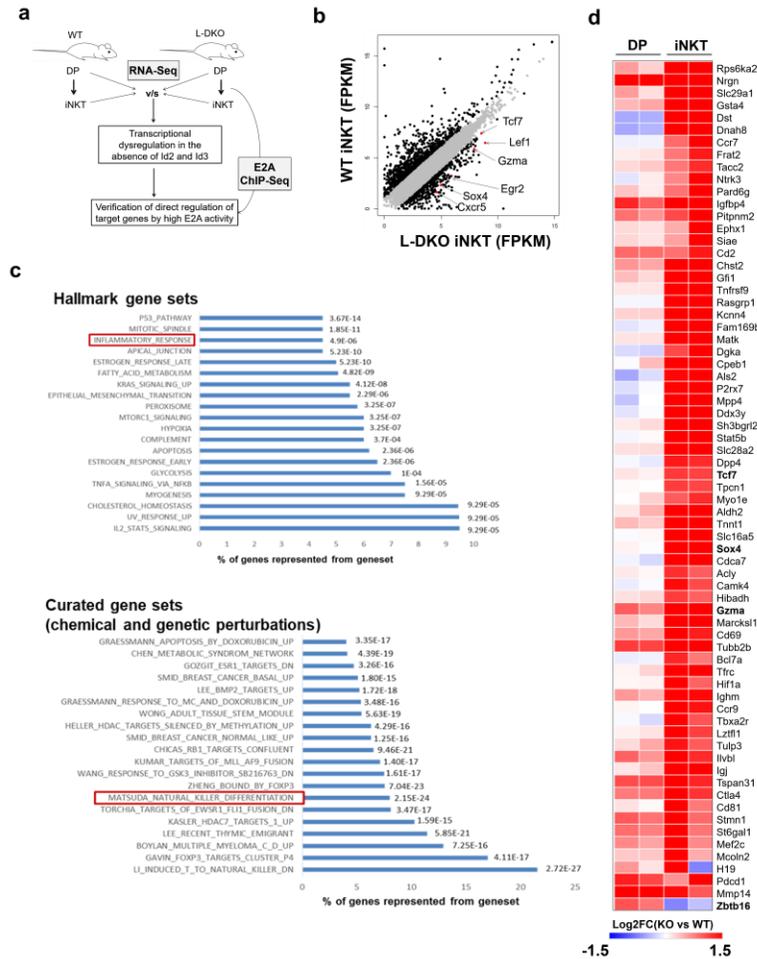


Figure 8: Upregulation of transcriptional programs that support iNKT cell development and function in the absence of Id proteins

(a) Schematic showing RNA-Seq analysis depicted in (b-d) and subsequent ChIP-Seq analysis. (b) Gene expression (FPKM) in WT and L-DKO iNKT cells with key genes highlighted (red). Genes with greater than 2-fold difference in L-DKO or WT iNKT cells are marked in black, and rest in gray. (c) Pathway analysis of genes that are more than 2-fold upregulated in L-DKO iNKT compared to WT iNKT cells. Top pathways from hallmark and curated gene sets are displayed with the percentage of genes that were shared with the respective pathways, along with the p values. Select gene sets representing inflammatory response (Msigdb ID: M5913) and iNKT differentiation (Msigdb ID: M18517) programs are highlighted in red. (d) Gene expression patterns in L-DKO DP and L-DKO iNKT cells (fold change over corresponding wild type cells) for genes that are more than two-fold upregulated in L-DKO iNKT cells compared to WT iNKT cells, and are part of the highlighted pathways.

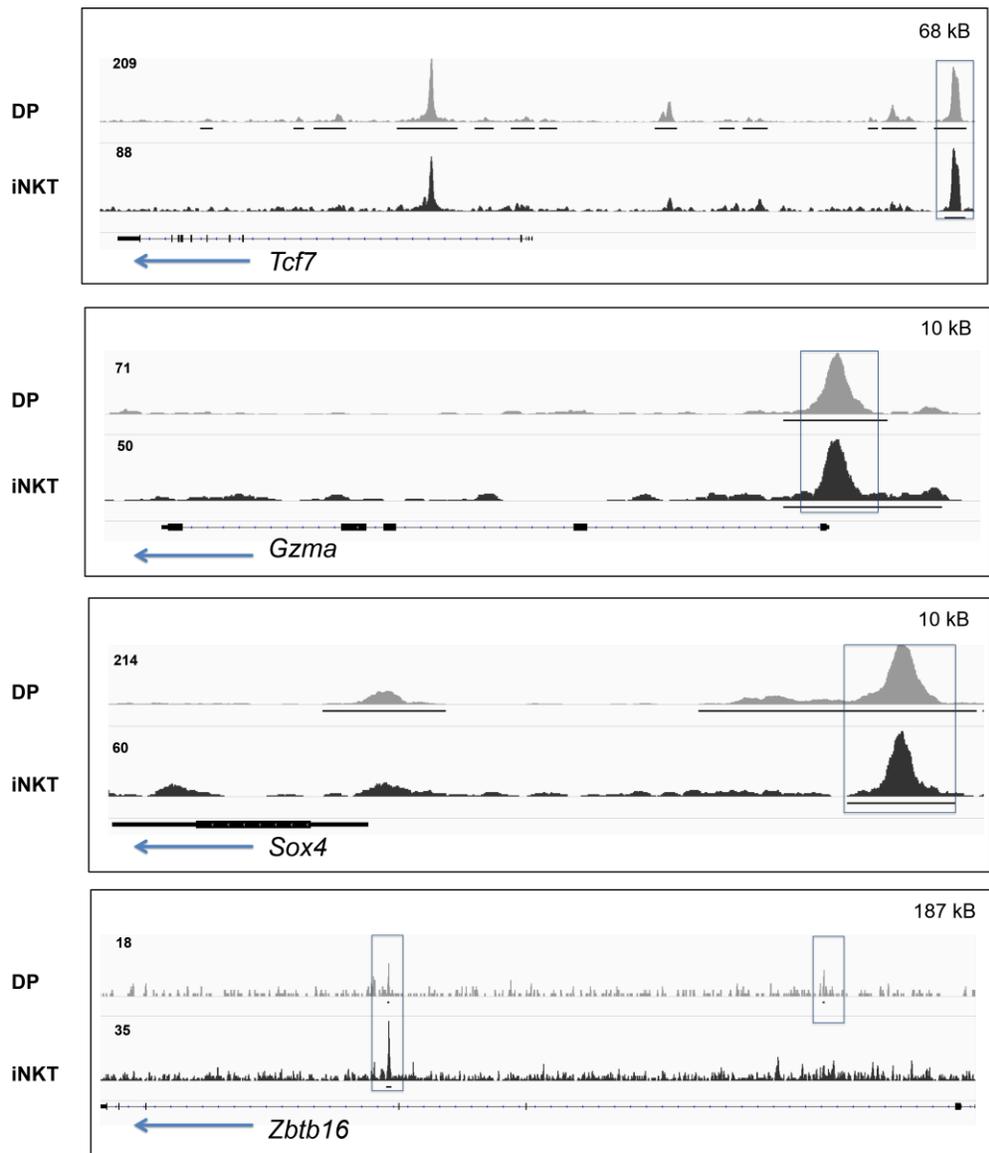


Figure 9: Uninhibited E2A activity directly regulates genes that support the iNKT lineage in DP and iNKT cells

E2A ChIP-Seq peaks in L-DKO DP and L-DKO iNKT samples. Solid black lines underneath the tracks indicate significant (p value less than 10^{-5}) peaks called by MACS. Important E2A peaks are highlighted by boxes within each panel. Length of the genome in each panel is indicated on the top right. Numbers at the top of each track indicate the maximum peak height.

3.2.2 E2A supports iNKT cell fate at the DP stage

I noted that Id-deficient DP cells seemed to upregulate a handful of genes related to the iNKT lineage, and many of the E2A targets in L-DKO iNKT cells were also occupied by E2A in DP cells (Figure 8d, 9, 10a). The overall number of peaks in L-DKO DP cells was also much greater as compared to iNKT cells (Figure 10a). It is possible that the shared downstream targets are important for driving and sustaining iNKT fate during and after TCR selection. The choice of conventional CD4⁺ and CD8⁺ T cell fate upon TCR selection is determined by the lineage-specific transcription factors ThPOK and RUNX3 respectively [217]. Along similar lines, I decided to examine involvement of transcription factors that might cooperate with E2A in differentially promoting iNKT lineage fate choice in DP cells.

I used *de novo* motif analysis to predict transcription factors that can bind to regulatory regions of identified ChIP-Seq gene targets. Besides the expected binding by E2A, this analysis demonstrated enrichment for RUNX1, TCF7, LEF1, GATA3 and ROR γ t motifs in our peaks, deeming them as potential partners of E2A in L-DKO DP and iNKT cells (Figure 10b). Furthermore, I found E2A peaks at the genes encoding these transcription factors, indicating that E2A may directly regulate and subsequently collaborate with these factors to modulate gene expression (Figure 10c). These transcription factors have been well documented to play critical roles in iNKT cell development [48, 216, 218]. RUNX1 has also been recently reported to regulate PLZF

kinetics and expression through the *Zbtb16* enhancer region where E2A was bound (Figure 9) [81]. These data suggest that E2A may promote iNKT lineage fate choice in DP cells by activating and collaborating with downstream transcription factors.

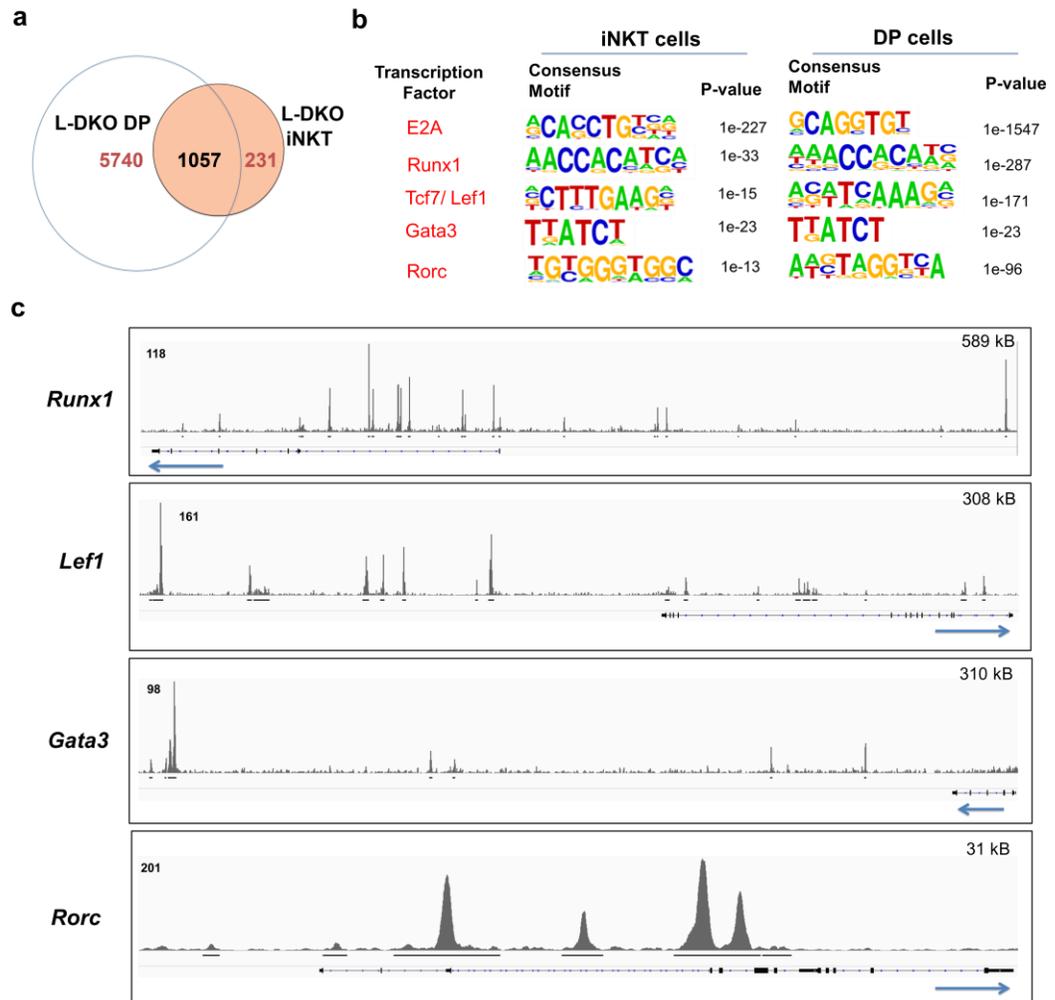


Figure 10: E2A regulates and collaborates with downstream transcription factors to promote the iNKT lineage fate choice at the DP stage

(a) Number of genes with E2A binding in L-DKO DP and/or L-DKO iNKT cells. (b) *De novo* motif analysis in L-DKO DP and iNKT cells, with predicted consensus motifs within E2A peaks, corresponding transcription factors, and p values. (c) E2A peaks in L-DKO DP cells at loci for motifs identified in (b).

3.2.3 Biased V α 14-J α 18 rearrangement in pre-selection cells

The previous observations suggested regulation of iNKT lineage fate choice at the DP stage. ROR γ t, which regulates survival of DP cells and, consequently, distal iNKT TCR α (V α 14-J α 18) rearrangement, was also predicted to be a co-factor of E2A in L-DKO DP cells (Figure 10b) [125]. Id gene deletion has been reported to lead to prolonged Rag1 and Rag2 expression, and one possible outcome of the persistent expression is a higher frequency of secondary, distal TCR α rearrangements, including the iNKT-specific rearrangement [84]. However, I wanted to determine if Id gene deletion also has a specific impact on V α 14-J α 18 rearrangement in DP cells that have not yet undergone TCR selection. I evaluated TCR α usage in L-DKO CD1d^{-/-} mice that lack iNKT cells due to *Cd1d1* and *Cd1d2* deficiency, thereby allowing us to study the impact on TCR α rearrangement independent of CD1d-mediated selection of iNKT cells [219].

I sorted pre-selection DP (CD4⁺CD8⁺CD69⁻) cells from L-DKO CD1d^{-/-} and CD1d^{-/-} control mice, and sequenced V α 8⁺ and V α 14⁺ populations to compare the J α diversity among these cells, as indicators of total pre-selection DP cells and potential iNKT precursors respectively. I found no difference in the breadth of the J α repertoire among V α 8⁺ cells in L-DKO CD1d^{-/-} or CD1d^{-/-} control mice (Figure 11). There was also no evidence to suggest increased distal J α rearrangements in L-DKO CD1d^{-/-} mice as compared to CD1d^{-/-} mice (Figure 12a). This implied that the absence of Id proteins did not promote an overall increase in distal J α rearrangements. However, I did find a

preferential increase in the frequency of J α 18 rearrangements among V α 14⁺ cells in L-DKO CD1d^{-/-} mice as compared to CD1d^{-/-} control mice (Figure 11, 12b). This increase was found in both productive and non-productive rearrangements, which verified that this outcome was not due to TCR selection (Figure 12c). Overall, this indicated that the loss of function of Id proteins causes a specific, CD1d-independent increase in the frequency of pre-selection DP cells that are eligible for selection into the iNKT lineage. The expression of a V α 14-J α 18 transgene can partially rescue defects in iNKT cell development and lead to an increase in iNKT cells [125, 220]. Therefore, the increased bias towards iNKT-specific rearrangement in pre-selection DP cells is likely to contribute to the increased iNKT population in Id-deficient mice in synergy with additional E2A-mediated transcriptional programs.

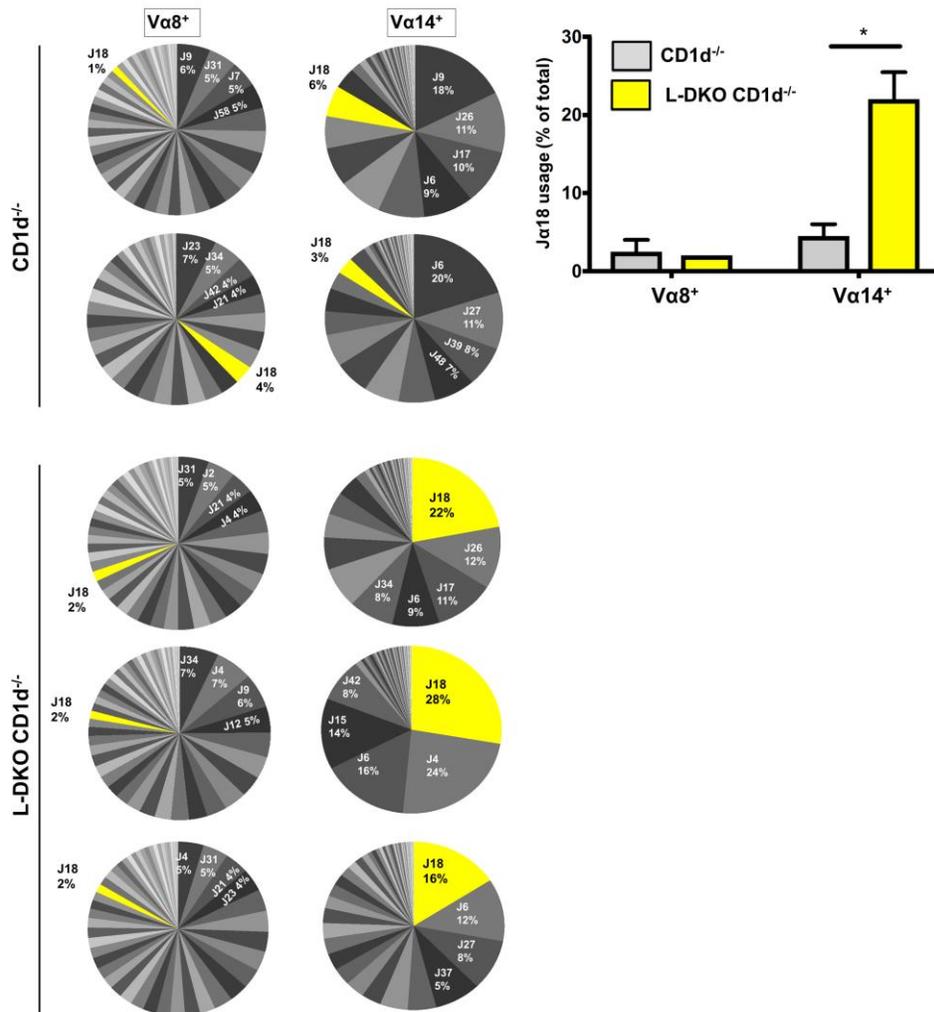


Figure 11: Id2/Id3 deficiency increases Vα14-Jα18 rearrangement in pre-selection DP pool of CD1d-deficient mice

Jα repertoire for Vα8⁺ (left panel) and Vα14⁺ (center panel) CD4⁺ CD8⁺ CD69⁺ cells in CD1d^{-/-} (n = 2) and L-DKO CD1d^{-/-} (n = 3) mice. Each pair of Vα8⁺ and Vα14⁺ pie charts represents the repertoire in a single mouse. Percentages represented by top 4 Jα chains, as well as Jα18 are indicated for each mouse. Jα chains are labeled according to new HUGO Gene Nomenclature Committee (TRAJ is abbreviated as J). Average number of total reads for CD1d^{-/-} control mice was 37,936 for Vα8⁺ cells and 27,619 for Vα14⁺ cells. For TKO mice, it was 40,925 for Vα8⁺ cells and 28,758 for Vα14⁺ cells. Bar graph (right) represents summary of Jα18 usage in CD1d^{-/-} and L-DKO CD1d^{-/-} mice. Error bars represent SEM. Statistical significance is represented by p values (* < 0.05).

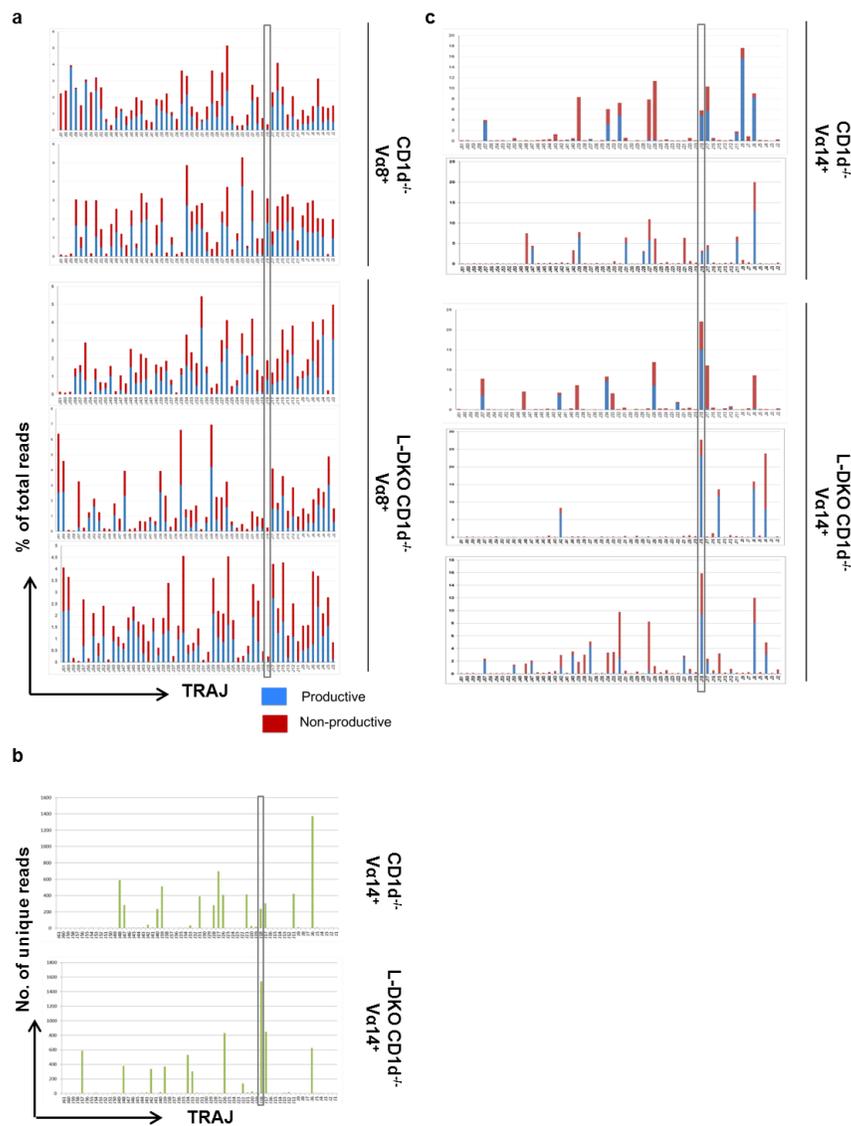


Figure 12: Biased $J\alpha$ repertoire in pre-selection DP cells eligible for iNKT selection

(a) Productive (blue) and non-productive (red) rearrangements among $V\alpha 8^+$ pre-selection DP cells in $CD1d^{-/-}$ (n=2) and TKO (n=3) mice respectively. (b) Representative $J\alpha$ repertoire among $V\alpha 14^+$ cells in $CD1d^{-/-}$ and TKO mice, as analyzed using IMGT HighV-QUEST statistical analysis tool. The counts for each TRAJ genes represents only single copies which are distinct in length and/or sequence. (c) Productive (blue) and non-productive (red) rearrangements among $V\alpha 14^+$ pre-selection DP cells in $CD1d^{-/-}$ (n=2) and TKO (n=3) mice respectively.

I delved deeper into possible mechanisms that can drive this bias towards V α 14-J α 18 rearrangements. Re-analysis of our ChIP-Seq data to allow for repeats in the TCR locus revealed strong E2A binding to both the promoter of Trv11 loci (representing V α 14) as well as the TCR α enhancer (E α enhancer), which regulates the TCR α locus accessibility and rearrangement (Figure 13) [221]. RNA-Seq analysis of DP cells from Id-deficient mice also demonstrated increased V α 14 transcription (Figure 13). This suggests that E2A might be capable of directly regulating TCR α rearrangement outcomes, in a manner that favors V α 14-J α 18 rearrangements.

It is known that the lack of IL-7 signaling can block proliferation and allow premature TCR α rearrangement [222]. Upon closer examination, I found a significant reduction in IL-7R expression among Id2/Id3-deficient DN4 cells, but not in DN3 cells (Figure 14a). Ki-67 analysis also hinted towards the reduced proliferation in L-DKO DN4 cells (Figure 14b). This data implies that in the absence of Id proteins E2A impacts the timing and outcome of TCR α rearrangement, which can contribute to the iNKT TCR rearrangement bias prior to TCR selection at the DP stage.

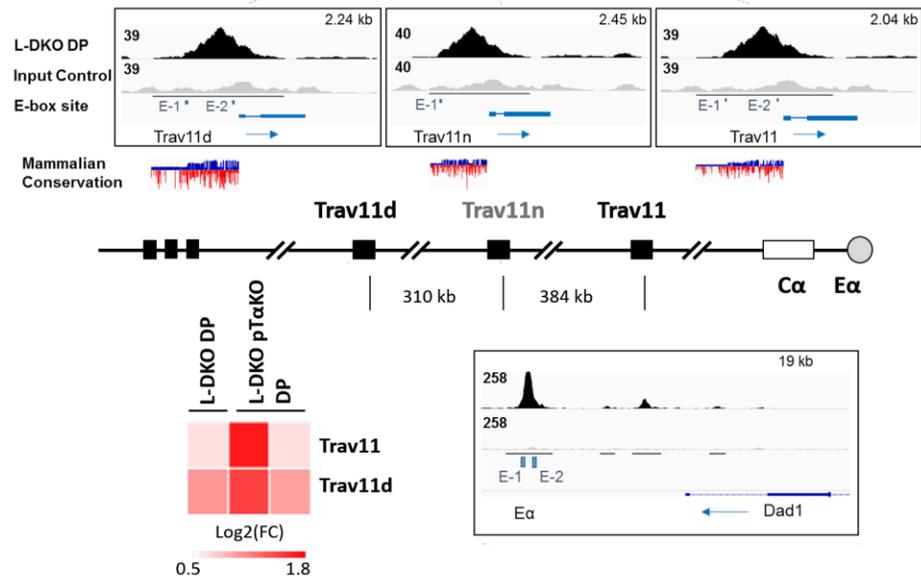


Figure 13: E2A binds to Ealpha and Trav11 promoters

E2A ChIP-Seq peaks in L-DKO DP cells at the Trav11 triplicate gene segments (Trav11, 11d and 11n) and E α enhancer region, along with canonical E-box sites within peaks marked as E-1 and E-2. RNA-Seq expression (log₂ fold change with respect to WT DP cells) of Trav11 and Trav11d transcripts in Id-deficient cells is also shown (lower panel, left). A simplified schematic representation of the TCR α/δ locus is drawn in the middle panel with zoomed in views of the peaks at depicted loci. Solid black lines underneath each track indicate significant (p value less than 10^{-5}) peaks called by MACS. Length of the genome depicted in each panel is indicated on the top right. Numbers at the top of each track indicate the maximum peak height.

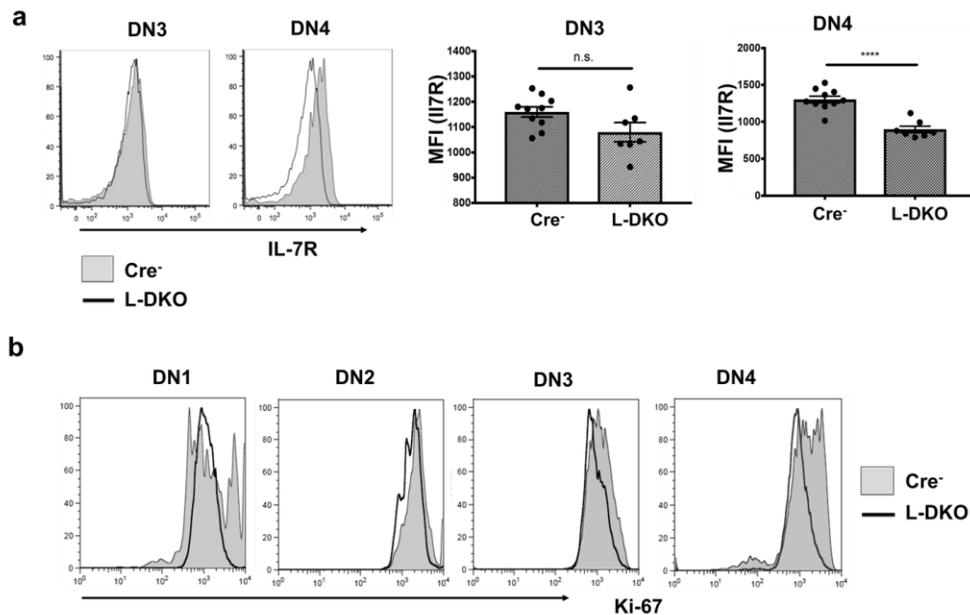


Figure 14: Impaired IL-7R expression and proliferation in L-DKO DN4 cells

(a) Representative surface expression (left) and Median Fluorescence Intensity (MFI, right) of IL-7R on DN3 and DN4 cells from 3-4 weeks old Cre^{-/-} and L-DKO mice. (b) Representative Ki-67 surface expression in DN1-4 cells from 3 week old Cre^{-/-} and L-DKO mice.

3.2.4 Pre-TCR signaling-independent expansion of iNKT population

E2A-mediated regulation of iNKT-relevant downstream targets in DP cells, and an early bias in iNKT TCR α rearrangement among pre-selection cells in our Id-deficient mouse models further prompted me to consider the possibility that Id proteins may suppress lineage specification for iNKT cells early in T cell development. Therefore, I decided to investigate if these populations are impacted by the deficiency of Id proteins at the pre-TCR checkpoint (at the DN3 stage), which regulates conventional $\alpha\beta$ T cell development. Mice deficient in pre-T α have restricted T cell development, with the

majority of cells blocked at the DN stage [196]. $pT\alpha^{-/-}$ mice are also known to completely lack iNKT cells [223]. I generated L-DKO $pT\alpha^{-/-}$ mice to examine how blocking pre-TCR signaling impacted the expanded iNKT population in L-DKO mice.

Despite the complete absence of iNKT cells in $pT\alpha^{-/-}$ mice, I surprisingly found a robust iNKT population in L-DKO $pT\alpha^{-/-}$ mice (Figure 15a-c). These iNKT cells also expressed high levels of PLZF (Figure 15d). It is known that $pT\alpha^{-/-}$ mice have an increase in $\gamma\delta$ T cells, and L-DKO $pT\alpha^{-/-}$ mice showed a similar increase in the $\gamma\delta$ population compared to WT mice (Figure 15e) [196]. However, the $\gamma\delta$ T cells in L-DKO $pT\alpha^{-/-}$ mice were predominantly $V\gamma 1.1^+V\delta 6.3^+$ and uniformly upregulated PLZF, reflecting a specific increase in innate-like $\gamma\delta$ NKT cells in these mice (Figure 15f-i).

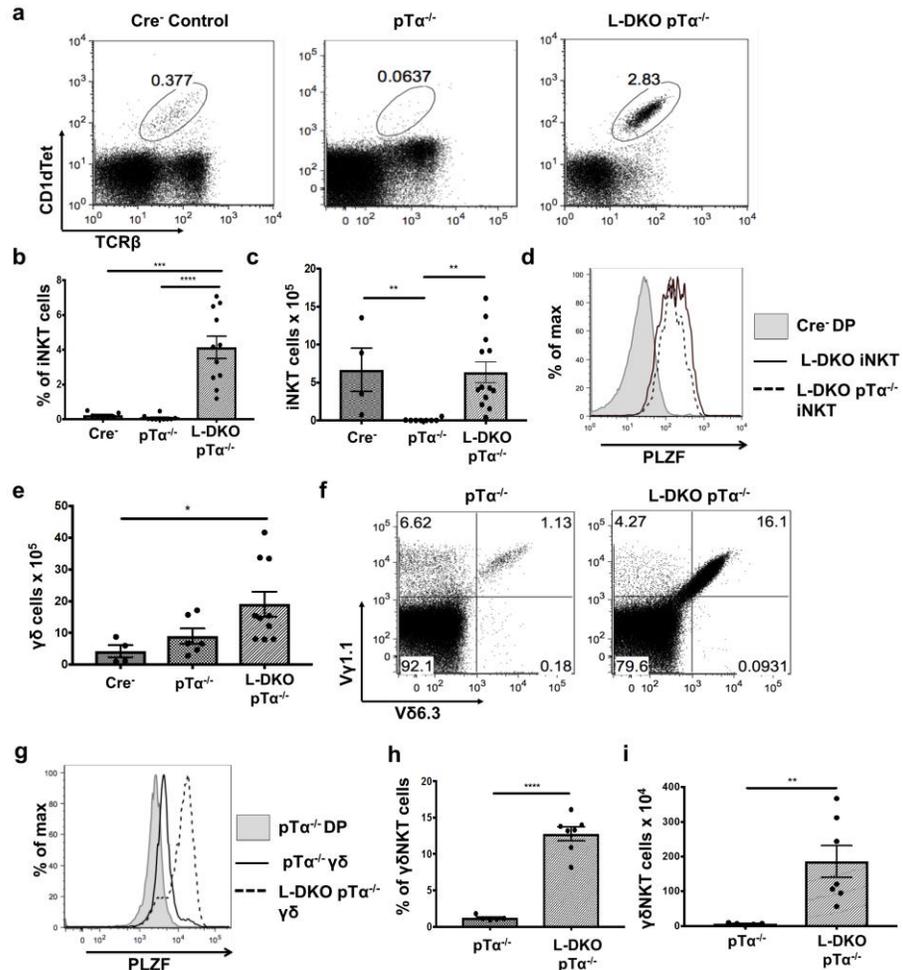


Figure 15: A block in pre-TCR signaling does not arrest enhanced iNKT cell and $\gamma\delta$ NKT populations in the absence of Id proteins

(a) Representative iNKT cells in 3-5 week old Cre⁻ (n=4), pTα^{-/-} (n=10) and L-DKO pTα^{-/-} (n=11) mice. (b) Percentage and (c) numbers of iNKT cells in 3-5 week old mice. (d) Representative PLZF expression levels in iNKT cells from 3-5 week old Cre⁻ (n=3), L-DKO (n=3) and L-DKO pTα^{-/-} (n=4) mice. (e) Total number of $\gamma\delta$ T cells in 3-5 week old Cre⁻ (n=4), pTα^{-/-} (n=6), L-DKO pTα^{-/-} (n=10) mice. (f) Representative $\gamma\delta$ NKT (V γ 1.1+V δ 6.3⁺) populations in 3-5 week old pTα^{-/-} (n=5), L-DKO pTα^{-/-} (n=7) mice. (g) Representative PLZF expression levels in $\gamma\delta$ T cells from 3-5 week old Cre⁻ (n=3), L-DKO (n=3) and L-DKO pTα^{-/-} (n=4) mice. (h) Percentages and (i) numbers of $\gamma\delta$ NKT cells in mice shown in (f). Error bars represent SEM. Statistical significance is represented by p values (* < 0.05, ** < 0.005, *** < 0.0005, n.s > 0.05).

As expected, L-DKO pT α ^{-/-} mice still had a profound block in conventional $\alpha\beta$ T cell development due to the lack of pre-TCR signaling (Figure 16a, b). Interestingly, despite the pre-TCR block in L-DKO pT α ^{-/-} mice, the deletion of Id proteins seemed to partially rescue the development of DP cells (Figure 16a, c). This was particularly unexpected since E2A normally blocks progression from the DN3 to the DP stage in the absence of a productive pre-TCR signal (Figure 5). Upon careful investigation, I found that these DP cells upregulate PLZF (Figure 16d). Our gating strategy excluded iNKT cells recognizing the CD1d tetramer and $\gamma\delta$ T cells expressing TCR $\gamma\delta$ to ensure that these PLZF^{hi} DP cells are not an artifact of aberrant upregulation of CD4 and CD8 by iNKT or $\gamma\delta$ NKT cells. Total thymocytes from L-DKO pT α ^{-/-} mice also displayed a prevalent innate-like phenotype, as indicated by their PLZF expression pattern (Figure 16e). Thus, blocking conventional $\alpha\beta$ T cells with pT α deficiency revealed a pre-TCR independent pathway that drives iNKT and innate-like lineage development in Id2/Id3-deficient mice. To further verify that the lack of a pre-TCR signal was selectively blocking conventional T cell development, I used anti-CD3 injection to potentially rescue these cells by mimicking the pre-TCR signal, as demonstrated in Rag-deficient mice. [224] Indeed, a partial rescue was observed in DP development, as indicated by the loss of PLZF expression on rescued DP cells (Figure 16f).

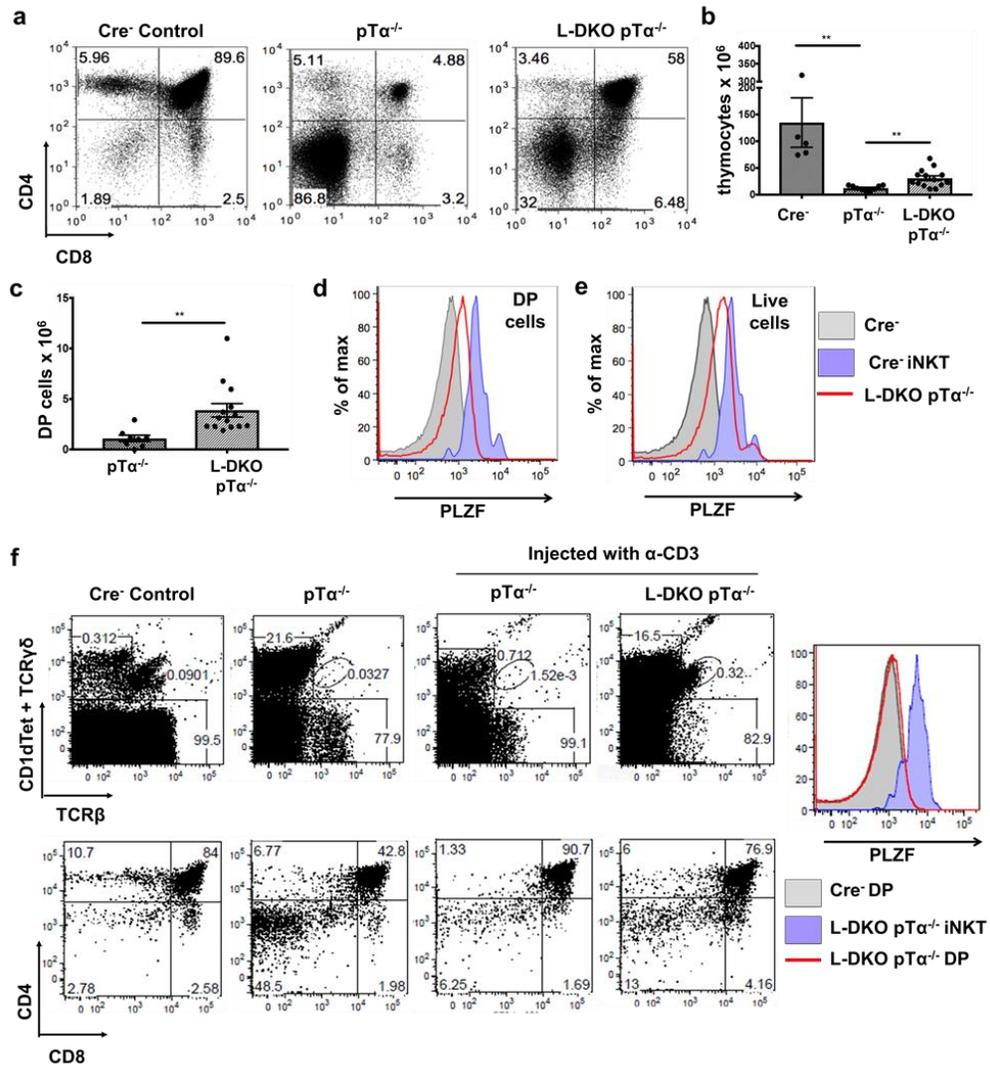


Figure 16: Development of innate-like DP cells in L-DKO pTα^{-/-} mice

(a) Representative thymocytes in 3-5 week old Cre⁻, pTα^{-/-} and L-DKO pTα^{-/-} mice as shown by CD4 and CD8 staining. (b) Total number of thymocytes in 3-5 week old mice. (c) Number of DP cells in mice shown in (a). 5 to 13 mice included for each genotype shown in (a-c). (d, e) Representative PLZF expression levels in (d) DP cells (CD1d^{Tet} TCRγδ⁺) and (e) total live cells from 3-5 week old Cre⁻ and L-DKO pTα^{-/-} mice. (f) Thymocytes in pTα^{-/-} and L-DKO pTα^{-/-} mice upon anti-CD3 injection in 4 weeks old, analyzed 10 days after injection. Uninjected Cre⁻ and pTα^{-/-} littermates are included as controls. PLZF expression in DP and iNKT cells from injected L-DKO pTα^{-/-} mouse is depicted on the right, with DP cells from uninjected Cre⁻ mouse as control.

3.2.5 Initiation of an innate-like transcriptional program

My previous data suggested that pre-TCR signaling and Id protein activity is necessary to enforce conventional T cell fate, such that the absence of both gave rise to predominantly innate-like T cell populations in the thymus. I further verified the innate-like phenotype of PLZF^{hi} DP (CD1d^{Tet} TCR $\gamma\delta$ ⁻ CD4⁺ CD8⁺) cells in L-DKO pT α ^{-/-} mice by RNA-Seq analysis (Figure 17a). Gene Set Enrichment Analysis (GSEA) verified innate-like genes associated with iNKT cell development and inflammatory responses, including *Zbtb16*, *Gzma* and *Il2rb*, to be enriched in L-DKO pT α ^{-/-} DP cells compared to both L-DKO DP and WT DP cells (Figure 17a, b) [152, 225]. Since genes with similar expression patterns can be expected to function together and/or be involved in similar biological processes, I examined expression patterns for genes that correlated positively or negatively with *Zbtb16* across all samples, including WT, pT α ^{-/-}, L-DKO and L-DKO pT α ^{-/-} DP cells, to discern innate-like genes. I found that most genes positively correlated with *Zbtb16* were specifically upregulated in L-DKO pT α ^{-/-} DP cells, whereas most genes negatively correlated with *Zbtb16* were downregulated in these cells (Figure 17c). RNA-Seq analysis of the PLZF^{hi} L-DKO pT α ^{-/-} DP cells also revealed these cells had undergone TCR α rearrangement with a fairly broad V-J usage (data not shown). These analyses demonstrate the early initiation and adoption of an innate-like transcriptional program specifically in DP cells that arise in the absence of pre-TCR signaling and Id function in the thymus.

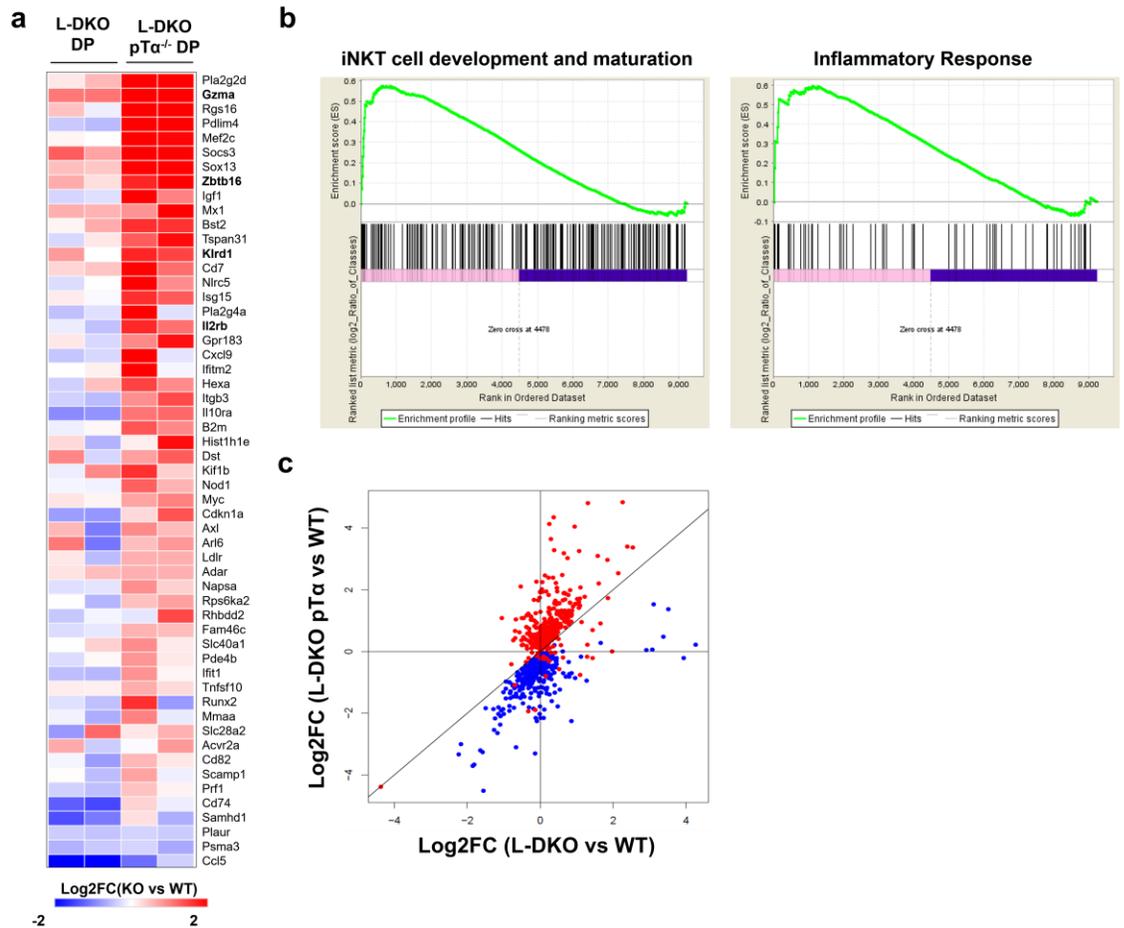


Figure 17: Adoption of innate-like transcriptional program by DP cells in the absence of pre-TCR signaling and Id proteins

(a) Gene expression patterns in L-DKO DP and L-DKO pTα^{-/-} DP cells for genes that are more than two-fold upregulated in the latter compared to WT DP cells, and are part of the iNKT transcriptional program (Msigdb ID: M18517) or inflammatory responses (Msigdb ID: M5913). (b) Enrichment of gene sets in L-DKO pTα^{-/-} DP cells over L-DKO DP cells pertaining to iNKT development and maturation (left), and inflammatory responses (right) using Gene Set Enrichment Analysis (GSEA). Enriched genes from these gene sets are shown in (a). (c) Gene expression patterns for genes that are positively (red) or negatively (blue) correlated with Zbtb16 expression.

3.2.6 Transcriptional similarity between iNKT cells and $\gamma\delta$ NKT cells

In the previous section, our study of L-DKO pT $\alpha^{-/-}$ mice revealed an expansion of both innate-like $\gamma\delta$ and iNKT cells. Recent publications have also reported the sharing of transcriptional and effector programs between iNKT cells and $\gamma\delta$ NKT cells, even though these two lineages are considered to independently diverge at the DP and DN3 stages respectively [226, 227]. Since Id3-deficient (Id3 $^{-/-}$) mice support the expansion of innate-like $\gamma\delta$ NKT cells, I compared WT and Id3 $^{-/-}$ $\gamma\delta$ NKT cells to WT and L-DKO iNKT and DP cells by RNA-Seq [61, 138]. Interestingly, both Principal Component Analysis (PCA) and clustering analysis highlighted that iNKT cells are transcriptionally more similar to $\gamma\delta$ NKT cells than to conventional $\alpha\beta$ DP cells, regardless of whether the comparisons were of WT populations or mutant populations (Figure 18a, b). While there were differences in the gene expression patterns of L-DKO iNKT and Id3 $^{-/-}$ $\gamma\delta$ NKT cells, around 25% of dysregulated genes were similarly affected in both populations (Figure 18c, d). The genes with similar expression patterns, including *Egr2*, *Slamf6*, *Rorc* and *Ifngr1*, were largely specific to the innate-like populations, and had distinct expression patterns in L-DKO DP cells (Figure 18d). Given the transcriptional and functional similarities between iNKT and $\gamma\delta$ NKT cells, and their expansion in Id-deficient mice, I wanted to further examine if Id proteins regulate the lineage competition between these two innate-like populations by eliminating $\gamma\delta$ lineage development and expansion.

I found a significant increase in the iNKT population in $Id3^{-/-}TCR\delta^{-/-}$ mice as compared to $Id3^{-/-}$ mice (Figure 19a-c). A modest but significant increase was observed in iNKT-committed stage 0 cells, but not in proliferating stage 1 cells (Figure 19d, e). In one study, iNKT cells and $\gamma\delta$ NKT cells were shown to compete for a thymic niche, based on the reduction in iNKT cells upon expansion of $\gamma\delta$ NKT cells [228]. In contrast, another study has reported that a reduction in iNKT cells does not lead to a corresponding increase in $\gamma\delta$ NKT cells [229]. In order to address this issue, I decided to separate ongoing T cell development from homeostatic expansion associated with void space, and examined pre-weaning pups that had not yet undergone full expansion and stabilization of the thymic architecture. A large increase in the iNKT population was observed again in pre-weaning age $Id3^{-/-}TCR\delta^{-/-}$ mice (Figure 19f, g). Age matched $TCR\delta^{-/-}$ mice, which lack total $\gamma\delta$ T cells but are wild-type for $Id3$, did not exhibit a corresponding increase in iNKT cells. These data suggest lineage competition between $\gamma\delta$ NKT and iNKT lineages in Id -deficient mice. It is also possible that despite an apparent disparate development, these cells may share the same evolutionary history.

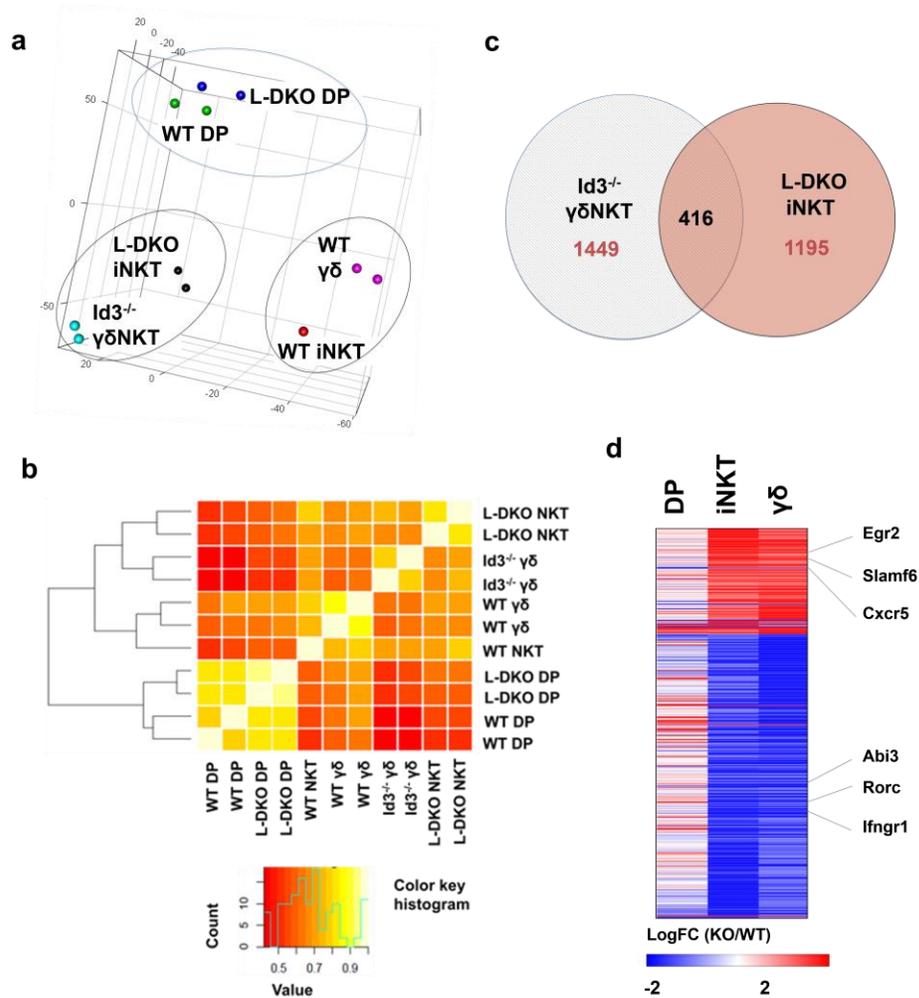


Figure 18: Transcriptional similarity between iNKT and $\gamma\delta$ NKT cells

(a) Principal Component Analysis of RNA-Seq expression data, with samples grouped according to similarity. (b) Hierarchical clustering analysis based on RNA-Seq expression patterns. (c) Genes found to be differentially expressed (fold change > 2) in L-DKO iNKT and $Id3^{-/-}$ $\gamma\delta$ samples, as compared to their WT counterparts. Numbers indicate unique or shared gene dysregulation between samples. (d) Expression patterns of 416 differentially expressed genes identified in (c), in L-DKO DP, L-DKO iNKT and $Id3^{-/-}$ $\gamma\delta$ T cells compared to WT DP, WT iNKT and WT $\gamma\delta$ T cells respectively.

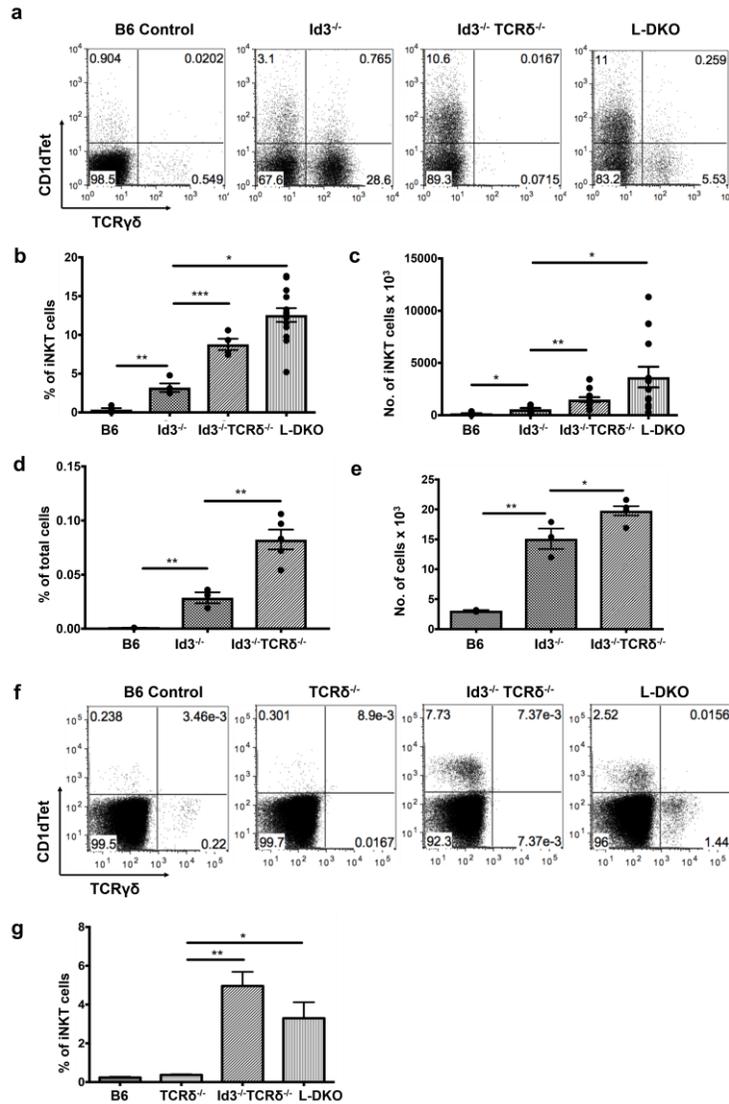


Figure 19: Lineage competition between iNKT and $\gamma\delta$ NKT cells is regulated by Id proteins

(a) Representative distribution of iNKT (CD1dTet⁺) versus $\gamma\delta$ T cells in 20 day old B6 (n=4), Id3^{-/-} (n=4), Id3^{-/-}TCRδ^{-/-} (n=4) and L-DKO (n=4) mice. (b) Percentage and (c) numbers of iNKT cells in 3-5 weeks old B6 (n=5), Id3^{-/-} (n=7), Id3^{-/-}TCRδ^{-/-} (n=13) and L-DKO (n=14) mice. (d) Percentage and (e) numbers of stage 0 iNKT cells in 2-5 weeks old B6 (n=3), Id3^{-/-} (n=3) and Id3^{-/-}TCRδ^{-/-} (n=5) mice. (f) Representative distribution of iNKT (CD1dTet⁺) versus $\gamma\delta$ T cells in 13 day old B6 (n=3), TCRδ^{-/-} (n=3), Id3^{-/-} TCRδ^{-/-} (n=5) and L-DKO (n=4) mice. (g) Percentage of iNKT cells in 13 day old mice shown in (f).

3.2.7 E2A orchestrates an innate-like gene network

So far, I observed expansion of iNKT, $\gamma\delta$ NKT and innate-like DP cells in Id-deficient mice, and our previous results implicated E2A in control of transcription programs that drive iNKT cell development. I wanted to further explore the role of E2A in orchestrating innate-like T cell development. Compared to conventional T cells, however, our understanding of transcriptional programs in innate-like T cells is still in its nascent stages. I therefore decided to explore the innate-like transcription program by compiling a reference innate-like gene set from publicly available Immgen data [211].

I hypothesized that genes that are upregulated or downregulated significantly in *both* iNKT and $\gamma\delta$ NKT cells over other T cell populations in the thymus, would be unique to these innate-like lineages and most likely be important for their development and/or function. My motivation was to delineate genes that are representative of a broad innate-like program, and important for both iNKT and $\gamma\delta$ NKT cell development, and specifically relevant for stage 0/1 iNKT cells and CD24^{hi} immature $\gamma\delta$ NKT cells that expand most dramatically in Id-deficient mice [70, 130]. Therefore, I compared the gene expression in these WT innate-like T cells against multiple WT conventional T cell populations, including DN3a, DN3b, DN4, DP, post-selection CD4SP and thymic $\gamma\delta$ T cells. This analysis resulted in 189 reference genes which are significantly overexpressed or repressed in both WT iNKT and $\gamma\delta$ NKT populations. I added 7 other genes to this list that were culled from literature to be important for the development of these

populations, but are not significantly overexpressed or repressed in these comparisons (Supplementary Table 1). It is important to note that these strict criteria would filter out genes important for only one of the populations, or for latter stages of maturation and function of either lineage.

When this reference innate-like gene set was compared to the genes identified in our RNA-Seq analysis, I found more than 50% (111 genes) to be dysregulated by at least two-fold in either one or both cell populations in Id-deficient mice (Figure 20, 21, Supplementary Table 2). Importantly, E2A directly bound to many (83 of 111 genes) of these differentially expressed reference genes (Supplementary Table 2). In order to further delineate the role of E proteins in regulating the developmental programs of iNKT and $\gamma\delta$ NKT cells through these downstream mediators, I divided the 111 genes into 3 groups based on their expression profiles in L-DKO iNKT and Id3^{-/-} $\gamma\delta$ NKT cells (Figure 20). Groups 1 and 2 included “biased” genes that were upregulated or downregulated by a significantly larger fold change in one mutant population as compared to the other, i.e. either in L-DKO iNKT or Id3^{-/-} $\gamma\delta$ NKT cells, compared to their WT counterparts. On the other hand, group 3 included “common” genes that were significantly and similarly upregulated or downregulated in both innate-like populations (Figure 20, 21). By combining known interactions between these genes with our RNA-Seq and CHIP-Seq data, I created a network map with the three groups of genes demarcated (Figure 21). The distribution of E2A targets across all three groups

strongly supported the role of E2A in orchestrating innate-like T cell developmental programs. Our previous observation of diminished iNKT and $\gamma\delta$ NKT populations in $Id2^{f/f}Id3^{f/f}E2A^{f/f}HEB^{f/f}Lck^{Cre+}$ (or Q-KO) mice that lack E protein activity further supports the pivotal role of E proteins in the development of these cells [70].

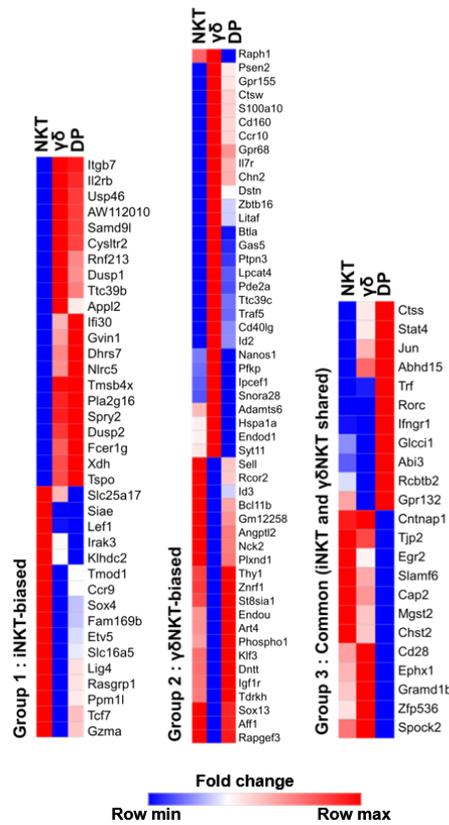


Figure 20: E2A-driven transcriptional network

Expression patterns (fold change in mutant cells compared to WT) of genes in groups identified in Figure 7, based on more than 2-fold dysregulation in one or more cell types, and part of the innate-gene reference gene set derived from WT Immgen data. Numerical fold change values provided in Supplementary Table 2.

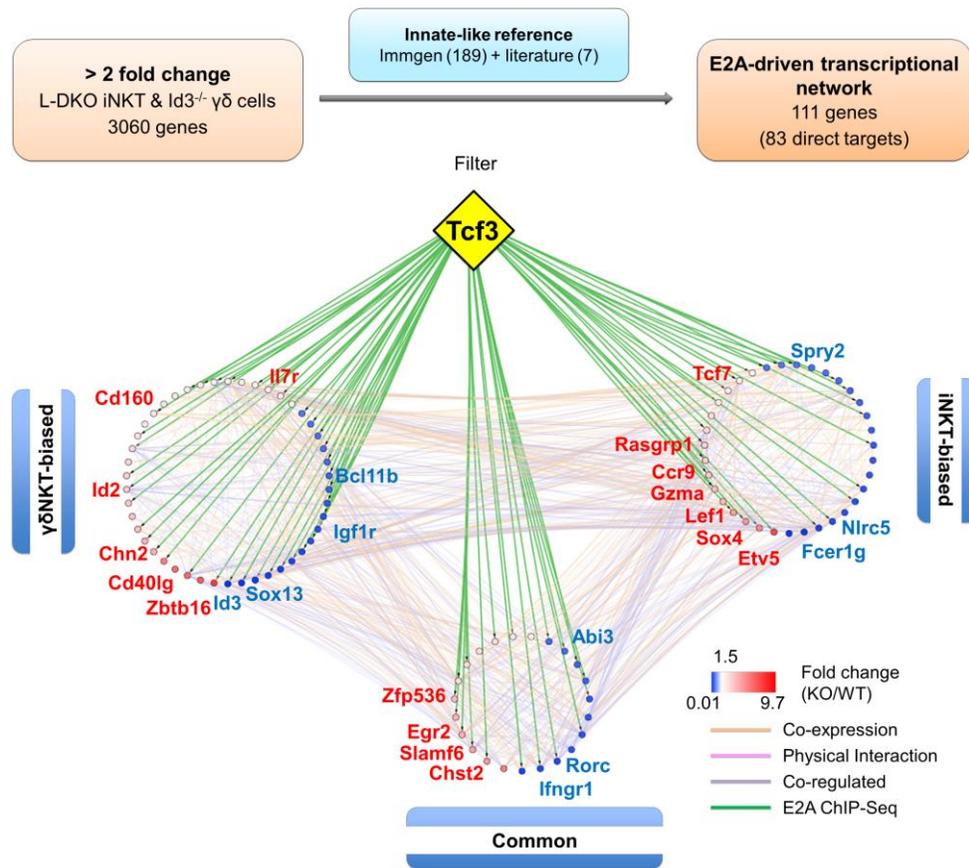


Figure 21: E2A drives a gene network that promotes iNKT and $\gamma\delta$ NKT fate in the absence of Id proteins

Schematic (top panel) representing the analysis pipeline used to derive the E2A-driven innate-like T cell transcriptional network. Network analysis (lower panel) depicting RNA-Seq expression and E2A ChIP-Seq data for genes part of the innate-like reference gene set. Genes in the “ $\gamma\delta$ NKT-” or “iNKT-biased” groups are colored according to their expression in Id3^{-/-} $\gamma\delta$ T cells and L-DKO iNKT cells respectively. Genes in the “common” group are represented by their average expression in Id3^{-/-} $\gamma\delta$ T cells and L-DKO iNKT cells. ChIP-Seq binding of E2A (encoded by Tcf3 gene) to gene targets is represented by green lines. Other interactions between gene targets, classified by GeneMania as co-expression, physical interaction or co-regulation, are represented by orange, pink and purple lines respectively.

3.3 Conclusions and discussion

iNKT cells are typically described as a lineage developing in parallel to conventional $\alpha\beta$ T cells, bifurcating after TCR-mediated selection at the DP stage. In order to investigate the mechanisms by which Id proteins suppress iNKT development, I studied lineage outcomes and transcriptional programs in Id-deficient mouse models that have a marked expansion of iNKT cells. This study shows that iNKT lineage development can be distinguished from the conventional $\alpha\beta$ T lineage as early as the pre-TCR checkpoint, albeit in the context of a unique genetic background. The divergence of iNKT and innate-like T cells from conventional T cells prior to TCR selection has also been proposed in other mouse models with physiological levels of E protein activity. A recent study has used fate-mapping and restriction of Rag2 expression to demonstrate an alternate developmental pathway for a subset of the iNKT population, such that these cells do not arise from the conventional DP stage [230]. In line with this, I uncovered a pre-TCR independent pathway for the development of iNKT cells using L-DKO $pT\alpha^{-/-}$ mice. It is likely that the depletion of Id proteins unleashes the “early”, pre-TCR-independent developmental program for iNKT and other innate-like T cells, which otherwise occurs at much lower frequencies on a wild-type genetic background. Consequently, these innate-like $\alpha\beta$ T cells can also give rise to lymphomas in Id2/Id3-deficient mice, as described in the next chapter [137, 190]. Cumulatively, our findings support a layered [215], rather than a parallel developmental

structure that coordinates the distinct fates of iNKT and conventional $\alpha\beta$ T cells during T cell development in the thymus (Figure 22).

The loss of iNKT cells in L-DKO CD1d^{-/-} mice emphasizes the critical role of the selection step in iNKT cell development [190]. However, TCR α repertoire sequencing of pre-selection DP cells from these mice demonstrated an increased frequency of V α 14-J α 18 rearrangements, suggesting that the lack of Id proteins can promote iNKT-specific rearrangements prior to, and independent of their selection. It remains to be determined if E2A can regulate the timing or outcomes of TCR α rearrangement to favor expression of the iNKT TCR. The combined genome-wide binding and transcriptional data revealed E2A-mediated transcription programs that support the development of $\gamma\delta$ NKT and iNKT lineages, providing a direct explanation to the several previous reports of expansion of iNKT and innate-like populations in the absence of Id proteins. This analysis identified E2A as an upstream regulator of genes critical for iNKT and $\gamma\delta$ NKT lineage differentiation, including *Zbtb16*, *Slamf6* and *Egr2* [51, 231, 232]. Genes that are associated with iNKT1 and iNKT17 cytokine profiles, such as *Ifngr1* and *Rorc*, were found to be significantly downregulated in both Id3^{-/-} $\gamma\delta$ NKT and L-DKO iNKT cells, supporting the involvement of E2A in preferentially driving iNKT2 and $\gamma\delta$ NKT lineage development in the absence of Id proteins. My data also serves as a framework and repository to add new genes as they are increasingly identified by our group and others to be important for innate-like T cell development and effector functions.

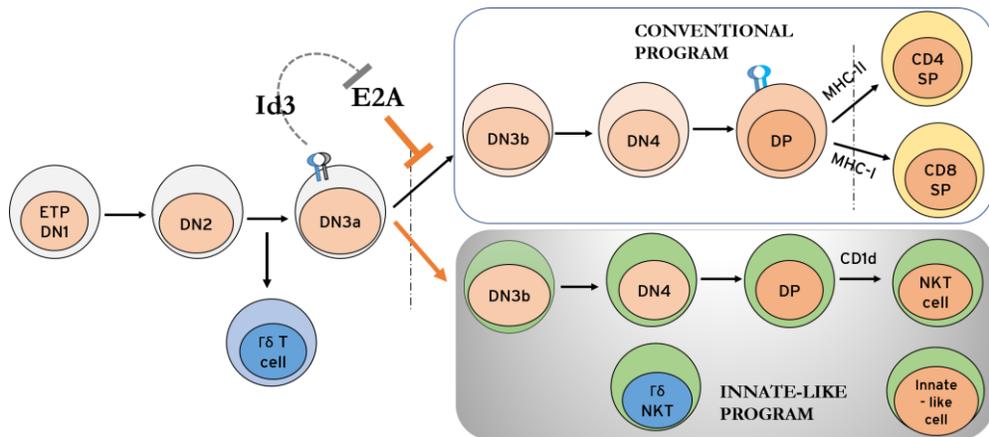


Figure 22: Proposed developmental model for divergence of innate-like and conventional T cell lineages, as regulated by E and Id proteins

A revised model of T cell development such that pre-TCR signaling and downstream Id3 suppress E2A activity at the pre-TCR checkpoint in order to predominantly drive developing T cells towards the conventional T cell program while suppressing the innate-like T cell program. Normal E and Id protein activity drives development along black arrows. Orange arrows show outcomes when E2A activity is high, i.e. in the absence of Id proteins. The pre-TCR and TCR selection checkpoints are labelled with vertical dotted lines. Precursor and mature T cells that turn on the conventional program have orange outer circles, whereas cells with an innate-like program are depicted in green outer circles.

4. Malignant transformation of innate-like T cells

The majority of this chapter is adapted from the publication “Id2 collaborates with Id3 to suppress iNKT and innate-like tumors”, Jia Li, Sumedha Roy, Young-Mi Kim, Shibo Li, Baojun Zhang, Cassandra Love, Anupama Reddy, Deepthi Rajagopalan, Sandeep Dave, Anna Mae Diehl, and Yuan Zhuang, Journal of Immunology. 2017 Apr 15;198(8):3136-3148.

4.1 Introduction

A significant portion of cancer research is dedicated to the identification of underlying factors that contribute to the hallmarks of tumorigenesis [233-235] such as dysregulated proliferation and self-renewal [236, 237]. Compared to Id3 and Id4, there is only limited evidence in favor of a tumor suppressive role played by Id2 [238]. Surprisingly, as we were studying iNKT cell development in Id2/Id3-deficient L-DKO mice, we found that these mice also rapidly develop tumors and start dying between 3-11 months of age [190]. L-DKO mice showed evident signs of splenomegaly and hepatomegaly, and developed lymphoma in several organs, including thymus, lymph nodes, bone marrow, spleen, liver and gut. Histopathological H&E (Hematoxylin and eosin) staining revealed infiltration of lymphoma cells in these organs and the disruption of normal tissue structures in these tumor mice. A large fraction of the lymphomas were derived from iNKT cells, that start expansion in neonatal mice as reported by us earlier and described in the previous chapter [70]. The identity of the cells was also verified by high PLZF expression and the typical V α 14-J α 18 rearrangement.

These cells were indeed tumorigenic as Rag2^{-/-} recipients died within 7-12 weeks of transfer of L-DKO lymphoma cells [190]. This observation was unexpected, as iNKT cells are primarily potent anti-tumor effector cells, and only in rare cases, they can undergo malignant transformation and give rise to tumors. Human NK/T tumors were initially described as PTCL cases which displayed both typical T cell and NK cell surface markers [239,240]. While the role of IL-15 in driving NK and iNKT cell activation has been reported in both development and tumor settings, much remains to be understood in terms of the mechanisms of tumorigenesis [241-245].

We also found that even though iNKT cells expand rapidly in neonatal L-DKO mice, only 36% of L-DKO tumors were derived from CD1dTet⁺TCRβ⁺ iNKT tumors, while the rest mostly developed CD1dTet⁻TCRβ⁺ tumors. Interestingly, the CD1dTet⁻TCRβ⁺ cells were also found to be PLZF⁺, distinguishing them from conventional αβ T cells. The Murre lab has also described the expansion of innate-like T_{FH} cells and lymphoma development in Id2^{fl/fl}Id3^{fl/fl}IL7RCre mice [137]. Other groups have also reported the expansion of PLZF⁺CD1dTet⁻TCRβ⁺ populations in the absence of ID proteins, the different stage of Id deletion and lack of comprehensive surface markers make these populations difficult to compare [65, 103]. Nonetheless, the PLZF⁺CD1dTet⁻TCRβ⁺ cells in our L-DKO mice had low surface expression of type II NKT-associated markers and a broader and more evenly spread TCRα chains, with no clear preference for Vα14 or Vα3, unlike iNKT and type II NKT cells. A unique feature of innate-like

lymphocytes is their ability to produce IL4 at steady state. Q-PCR analysis revealed that the CD4⁺CD1dTet⁺ population had a higher IL-4 expression at steady state than iNKT cells and conventional CD4 T cells, which indicated a possible regulatory role of this subset [190].

In this section, I further identify the expanding CD1dTet⁺ precursor populations and delineate dysregulated pathways that account for lymphoma development from expanding iNKT cells. Given that iNKT lymphomas are extremely rare and highly lethal in humans [246], this study provides a much needed animal model for understanding the genetic basis of similar types of tumors in humans.

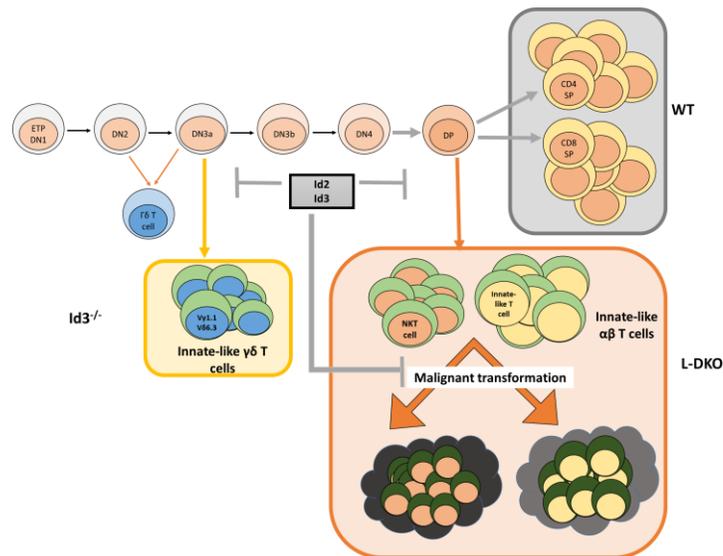


Figure 23: Summary of phenotypes observed in Id-deficient mice

WT thymic development is depicted by black and gray arrows, such that Id2 and Id3 suppress innate-like T cell fates and lead to predominant development of conventional T cells. Expansion of $\gamma\delta$ NKT cells is found in Id3^{-/-} mice, and iNKT and innate-like $\alpha\beta$ T cells expand and give rise to innate-like lymphomas in L-DKO mice.

4.2 Results

4.2.1 Neonatal expansion of CD1dTet⁻ innate-like T cells

Similar to the neonatal expansion of iNKT cells that ultimately give rise to iNKT tumors in L-DKO mice, I wanted to explore the possible neonatal expansion of these CD1dTet⁻ cells in L-DKO mice. After gating out $\gamma\delta$ T cells, which aberrantly upregulate CD4 and CD8 in L-DKO mice, I found that the CD1dTet⁻TCR β ⁺ population was indeed expanded in 20-day old L-DKO mice (Figure 24). This population had markedly upregulated PLZF expression, as observed in CD1dTet⁻ L-DKO tumor cells. Interestingly, the PLZF level in CD4SP CD1dTet⁻ cells was even higher than that in iNKT cells. These observations are indicative of neonatal expansion of the innate-like CD1dTet⁻ population.

4.2.2 CD1dTet⁻ expand independent of CD1d-mediated selection

Based on our prior examination of surface markers and TCR α repertoire, the expanding CD1dTet⁻ cells did not seem to belong to traditional iNKT or iNKT type II categories [190]. It is known that all types of $\alpha\beta$ iNKT cells depend on CD1d-mediated selection for their development, [247] and therefore I decided to examine if these innate-like T cells could develop in the absence of CD1d-mediated selection.

In the previous chapter I reported that L-DKO CD1d^{-/-} (henceforth referred to as triple knockout, or TKO) mice have a bias towards the iNKT TCR prior to selection. These mice lack CD1dTet⁺ iNKT cells due to their dependence on CD1d-mediated

selection (Figure 25a, second column). However, I found that the CD1dtet⁻TCRβ⁺ population was still existent in the absence of *CD1d* (Figure 25a). PLZF expression also verified that these CD1dtet⁻TCRβ⁺ cells in TKO mice were innate-like (Figure 25b). Cumulatively, these data verified that these CD1dTet⁻ cells are innate-like, and a novel type of CD1d-independent iNKT cells, similar to γδ iNKT cells.

4.2.3 Malignant innate-like T cells invade healthy tissues

I found that the innate-like CD1dTet⁻ population not only expanded, but also gave rise to tumors in TKO mice (Figure 26a). Similar to L-DKO mice, lymphomas were observed in several organs, including thymus, lymph nodes, bone marrow, spleen, liver and gut. Histopathological H&E (Hematoxylin and eosin) staining of the spleen, liver, lung and kidney of TKO tumor mice revealed infiltration of lymphoma cells in these organs, and the disruption of normal tissue structures (Figure 26b). The infiltrating cells in the various peripheral organs were also verified to be CD1dTet⁻, with visibly lower TCRβ expression than conventional T cells, as has been described by us for innate-like T cells in L-DKO mice (Figure 26c) [70].

Adoptive transfer of CD1dTet⁻ lymphoma cells from L-DKO mice into wild type (WT) hosts also gave rise to secondary lymphomas within 10 weeks (Figure 27a, b), indicating that these tumors have acquired the ability to evade immune surveillance. These secondary lymphomas match the original CD1dTet⁻ phenotype of the donor innate-like lymphoma cells, and sustained high PLZF expression levels (Figure 27c).

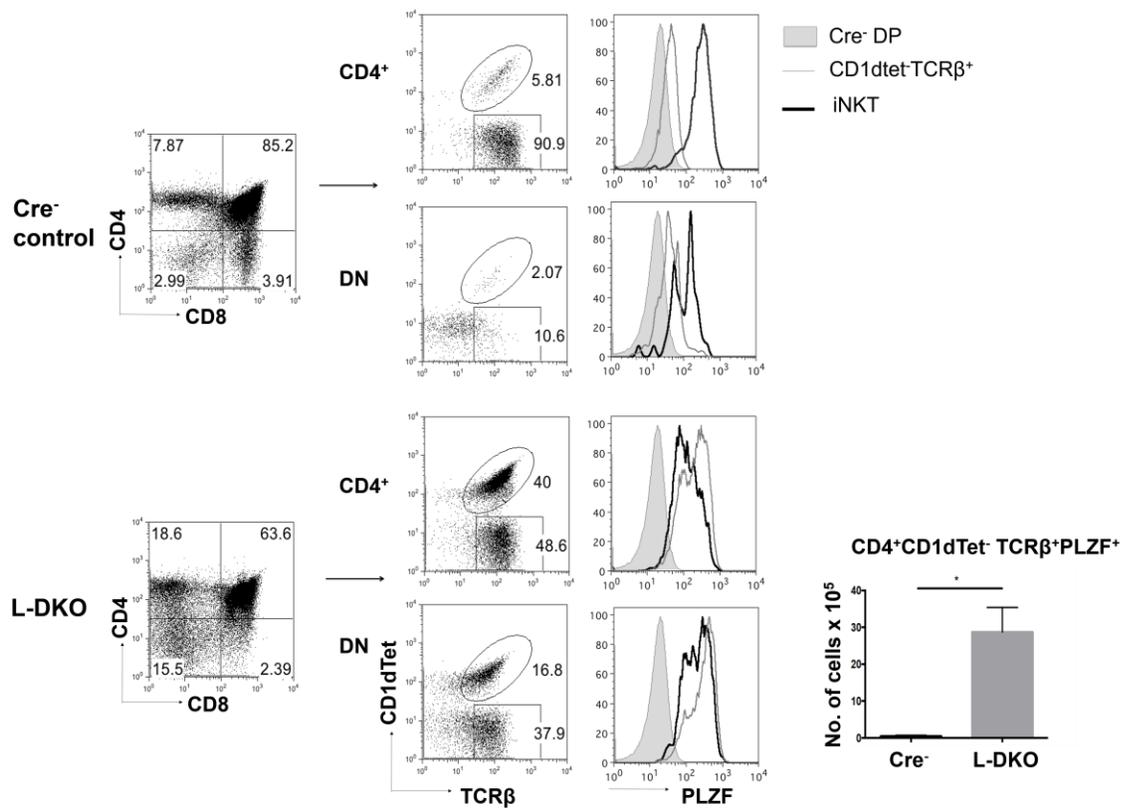


Figure 24: Neonatal development of Tet⁻ cells in LDKO mice

Representative flow cytometry analysis of thymocytes (TCR $\gamma\delta$ ⁺ cells gated out) from 20 day old L-DKO and Cre⁻ control mice. Cells were stained for CD4 and CD8 to separate the CD4⁺ and DN populations, which were further analyzed with TCR β and CD1dTet markers. Intracellular PLZF staining is shown for the corresponding CD1dTet⁺ and CD1dTet⁻ (iNKT) cells from the CD4 SP and DN fractions, along with Cre⁻ DP cells as controls. Absolute numbers of CD4⁺CD1dTet⁻TCR β ⁺PLZF⁺ cells are shown for 20 day old L-DKO and Cre⁻ mice (n = 3 for both Cre⁻ and L-DKO mice).

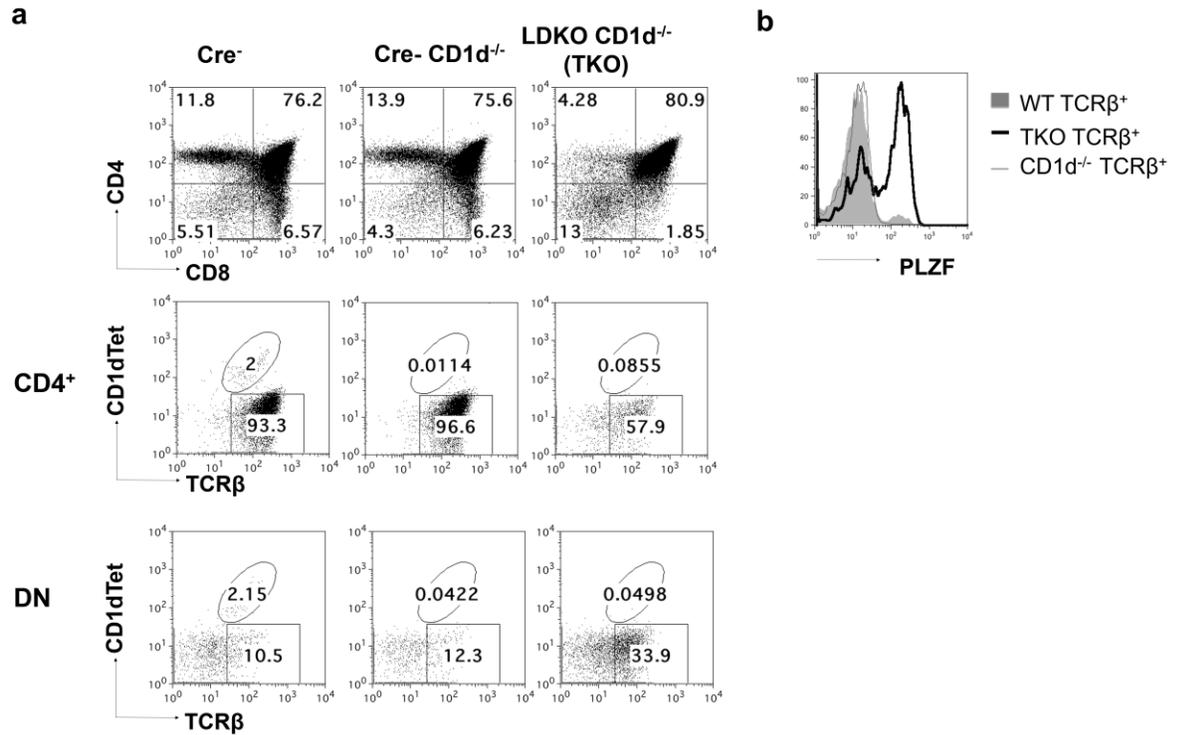


Figure 25: CD1dTet cells develop independent of CD1d

(a) Representative staining of thymocytes from 20 day old Cre⁻ control, Cre⁻ CD1d^{-/-} control or L-DKO CD1d^{-/-} (TKO) mice using CD4 and CD8 markers (top panel). CD4⁺ and DN gated cells were further analyzed for CD1dTet and TCRβ expression. (b) Representative intracellular PLZF staining for TCRβ⁺ cells from 20 day old WT (Cre⁻), Cre⁻ CD1d^{-/-} and TKO mice (n = 3).

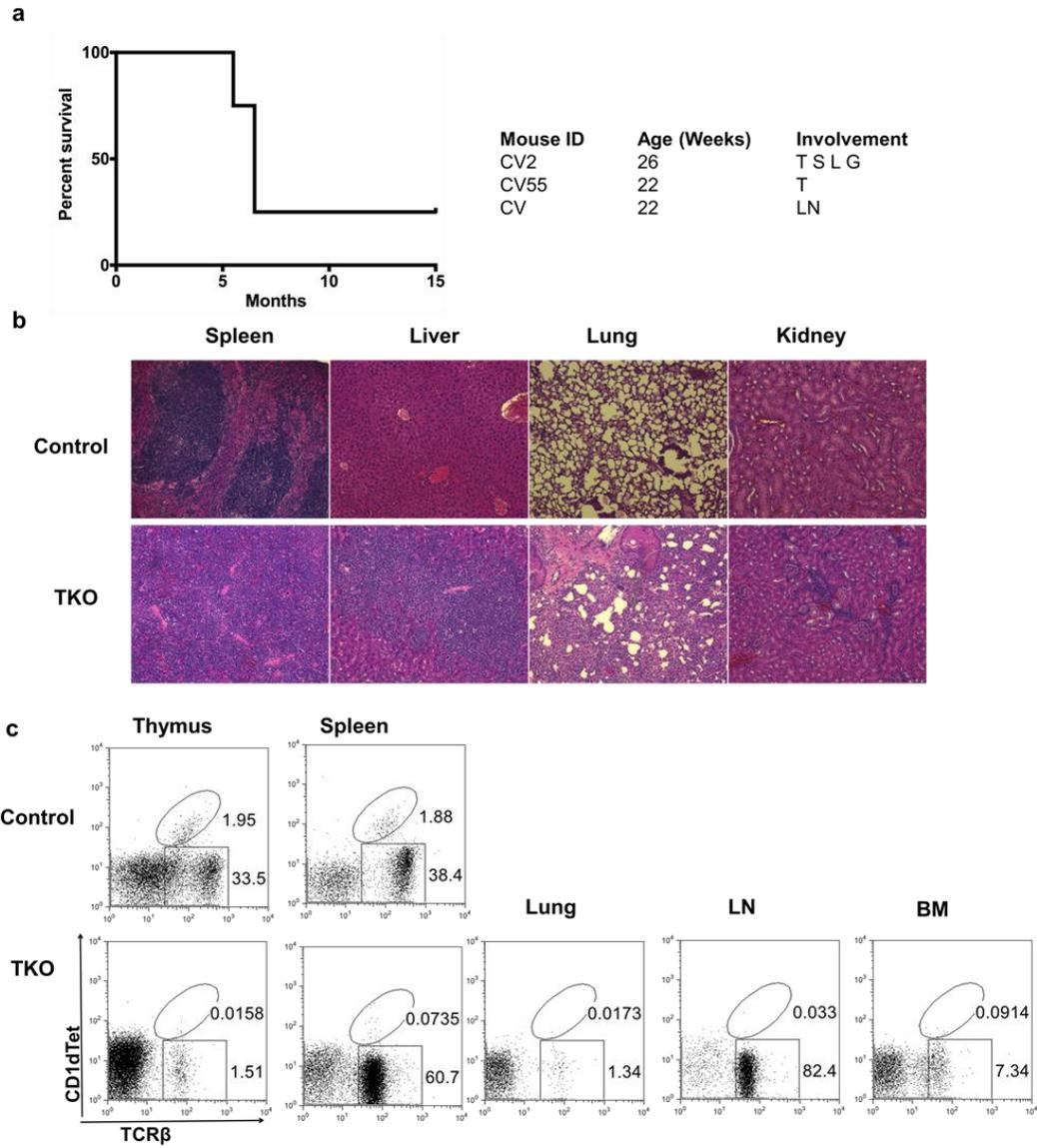


Figure 26: CD1dTet⁺TCRβ⁺ T cells give rise to lymphomas in TKO mice

(a) Survival curve for TKO mice (n = 4) along with details of mice that developed tumor. (b) Representative pictures of H&E staining of spleen, lung, liver and kidney tissue sections from Cre⁻ control and TKO mice (100X). (c) Representative distribution of TCRβ⁺CD1dTet⁺ and TCRβ⁺CD1dTet⁻ cells in thymus, spleen, lung, lymph node (LN) and bone marrow (BM) of Cre⁻ mice and TKO mice with tumors.

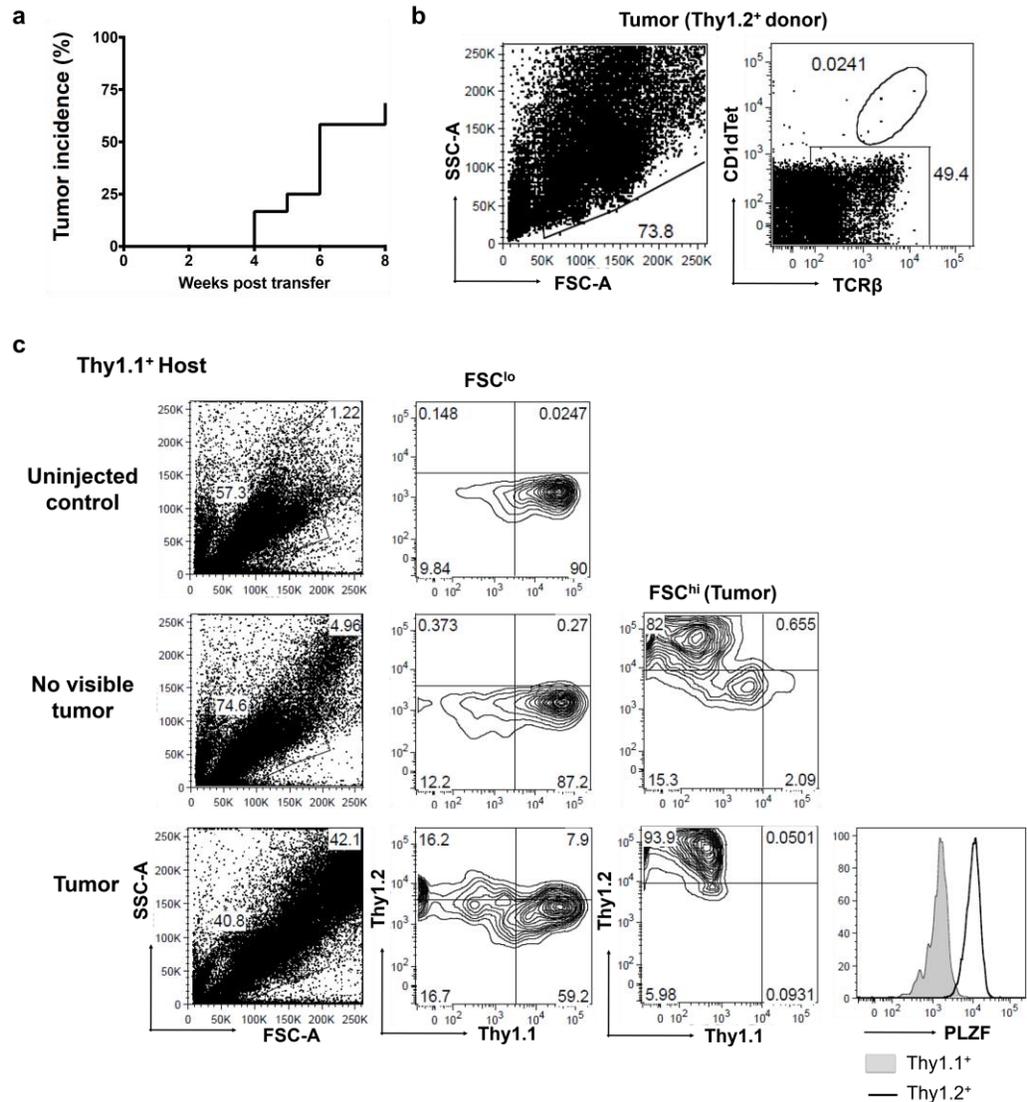


Figure 27: Adoptive transfer of lymphoma cells gives rise to tumors in WT hosts

(a) Tumor incidence (left) in WT hosts ($n = 10$) after receiving 5×10^6 lymphoma cells from L-DKO donor mice with CD1dTet tumor ($n = 2$). Representative picture of host with tumor in liver, spleen and LN (marked with arrows). (b, c) Representative flow cytometry analysis of (b) CD1dTet⁻ lymphoma donor cells from spleen, and (c) cells recovered from spleen of Thy1.1⁺ WT hosts, 10 weeks after receiving Thy1.2⁺ lymphoma donor cells, or uninjected control. Retention of the innate-like phenotype of lymphoma donor cells was verified by upregulated PLZF expression in FSC^{hi} TCRβ⁺Thy1.2⁺ cells (large lymphoma cells) as compared to FSC^{lo}TCRβ⁺Thy1.1⁺ cells.

4.2.4 Transcriptional program in neonatal expanding iNKT cells

After examining the CD1dTet⁻ cells that give rise to tumors in L-DKO mice, I next sought to identify the dysregulated molecular mechanisms responsible for tumor initiation and development. I did a microarray analysis to compare pre-malignant iNKT cells from 20 day old L-DKO mice, and lymphoma cells (either CD1dTet⁻ or CD1dTet⁺ iNKT in origin) from L-DKO mice with well-developed tumors (Figure 28a). Hierarchical clustering analysis (data not shown) and unsupervised Self-Organizing Maps demonstrated a clear distinction between lymphoma and pre-malignant iNKT cells, and more variability between the tumor samples (Figure 28b). In order to allow better and direct comparisons of our L-DKO pre-malignant and tumor cells with control WT iNKT and WT DP cells, I combined our microarray data with publicly available ImmGen data [211]. Principal Component Analysis (PCA) showed clear segregation between the WT iNKT, L-DKO neonatal and tumor samples (Figure 28c). Since the CD1dTet⁻ tumor sample was significantly distant from the CD1dTet⁺ tumors, it was treated as a unique outlier, and analyzed later using RNA-Seq (Figure 28c, Table 4). I then focused our attention to the iNKT tumor samples, and identified genes with more than 2-fold expression in L-DKO iNKT, iNKT tumors or WT iNKT compared to WT DP cells. Of these genes, the iNKT tumor cells had a largely unique gene profile with only a small fraction retained in common with WT iNKT and premalignant L-DKO iNKT cells (Figure 28d).

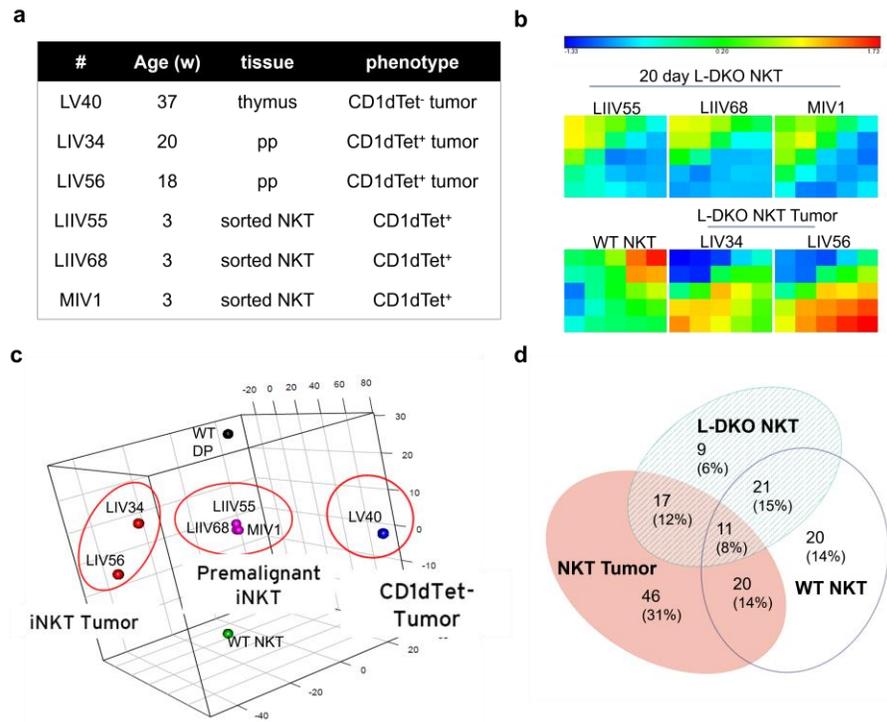


Figure 28: Microarray analysis of premalignant iNKT and tumor iNKT cells

(a) Age of mice, tissue origin of cells and phenotype, as used for mouse genome microarray analysis. Neonatal iNKT cells from 20 day old L-DKO mice represent pre-malignant state whereas tumor cells (CD1dTet⁺ and CD1dTet⁻) were obtained from 18-37 week old mice. (b) Self-organized map (SOM) showing gene expression in clusters of genes for tumor or iNKT cells from mice listed above, or from WT control mice (combined data). Colors represent low (blue) to high (red) log₂ fold change in gene expression with respect to WT DP cells. (c) Principal Component Analysis (PCA) for L-DKO iNKT, iNKT tumor and CD1dTet⁺ tumor samples (described in (a)), combined with WT iNKT and WT DP cells from Immgen. (d) Venn diagram represents the number and percentage of iNKT-specific genes ($p < 0.05$ and absolute fold change greater than two in WT iNKT with respect to WT DP) that are unique or shared between WT iNKT, neonatal L-DKO iNKT and iNKT tumor cells.

Based on the expression patterns of gene clusters in the samples, I identified genes that were differentially expressed in either WT or L-DKO iNKT cells (Figure 28b, Table 3). Interestingly, a significant fraction of genes were downregulated in L-DKO iNKT cells as compared to WT iNKT cells, which included anti-proliferative and pro-apoptotic genes *Paar* [248] and *Lgals1* [249], and tumor suppressor genes *Cebpb* [250] and *Irf5* [251] (Figure 29a). Genes implicated in cell cycle progression and metastasis, such as *Vangl2* (Wnt pathway), *Cdk1* (p53 pathway), *Ccr7* and *Igfbp4* were significantly upregulated in L-DKO iNKT cells. *Dgka*, which has been reported to be important for iNKT development [252] as well as to promote tumorigenesis [253], was also found to be more than 2-fold upregulated in L-DKO iNKT cells. These expression patterns supported the tumorigenic potential of these cells.

On the other hand, these premalignant cells also demonstrated the upregulation of several cell cycle arrest, tumor suppressor and anti-proliferative genes such as *Rprm* [254], *Ptpn14* [255] and *Btg2* [256] (Figure 29a). Other genes that commonly contribute to tumor development, or are overexpressed in tumors, such as *Pkd2* [257], *Mmp2* [258], *Adm* [259] and *Vcam1* [260], had reduced expression in these cells. Genes involved in cytokine-cytokine receptor interaction, many of which have been implicated in facilitating tumor metastasis, were also downregulated (Figure 29b) [261]. This data hints towards the existence of a tumor suppression program in these cells that prevents tumorigenesis at this stage.

Rag2 was found to be upregulated by more than 3 fold in L-DKO iNKT cells, which would allow prolonged TCR α rearrangement to increase chances of the distal V α 14J α 18 rearrangement to promote iNKT cell development (Figure 29a) [125]. We have previously described a block in iNKT development beyond stage 1 in L-DKO mice, which allows these cells to constantly proliferate without undergoing maturation [70]. I found downregulation of *Relb* in L-DKO iNKT cells, which has been described to be critical for developmental progression of iNKT cells to stage 2 and 3 (Figure 29a) [262]. Overall, the gene expression and pathway analysis revealed that *Id2/Id3* deletion initiates a modified transcriptional program in L-DKO iNKT cells that supports their prolific expansion while maintaining a pre-malignant state.

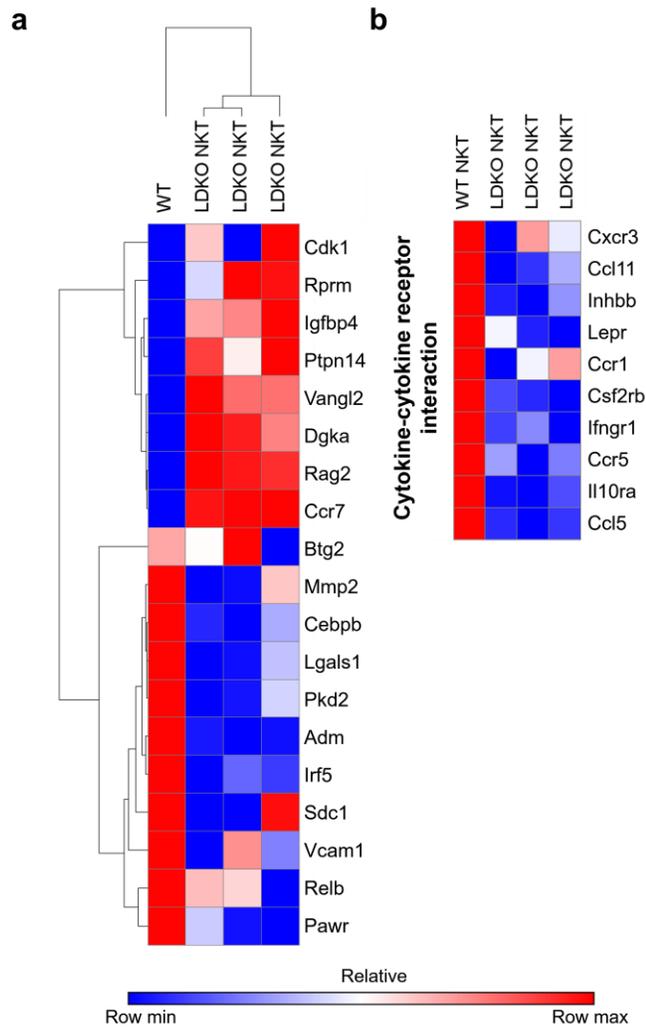


Figure 29: Gene expression patterns in premalignant neonatal iNKT cells

(a) Heat map showing hierarchical clustering and relative log fold change of gene expression in WT iNKT and neonatal L-DKO iNKT cells with respect to WT DP cells. Genes were selected based on expression patterns of SOM clusters (listed in Table 3). (b) Heat map showing the relative log fold change for genes involved in cytokine-cytokine receptor interaction. Colors represent the lowest (blue) to highest (red) fold change of a particular gene among the different samples.

Table 3: Cluster definitions according to Self-Organizing Map (SOM) analysis

Cluster ID	Row	Column	Gene Names
1	1	1	Ifit1, Sh3bgrl2, Secisbp2l, Ifih1, Cmpk2, Irgm2, Irf7
2	1	2	Scn11a, Rps4y2, Hivep3, Ddx24, AY036118, Tprgl
3	1	3	F2rl1, Btg2, Coro2a, Atg14, Chd7, Rasa1, Osbp15
4	1	4	Arl4d, Itgae, Trip4, Pcgf2, Rgs3
5	1	5	Marveld2, Ahnak, Smpdl3a, Fgd6
6	2	1	Phxr4, Brdt, Arl4c, Ndr3, Dgka, Ifi27l2a, Vangl2
7	2	2	Parg, H2-T24, Cox6b2, Ccs, Hmbox1, Ldlrap1, Emilin1, Fam189b, Ndr4
9	2	4	Cxcr3, Rhob, Cd244
10	2	5	Adam19, Hip1, Ctsd
11	3	1	Ccr7, Igfbp4, Rag2
13	3	3	Tmem176a, Tnc, Pld4, Lgals1, Wbp5, Colec12, Ggta1, Cdh5, Vcam1, Mmp2, Rab34, Sdc1, Rhoj
14	3	4	Pygl, Cst7, Slc39a4, Ifngr1, Syng2
15	3	5	Ifitm3, Ms4a6c, Ccl5, Swap70, Il10ra, Ccr5, Naip2
16	4	1	Figl1, Cdk1, Emp3
17	4	2	Tbc1d24, Sepp1, Fhl1
18	4	3	Cpe, Adm, Cfp, Slc15a3, Cd83, Cd86, Ms4a1, Cd79b, Fabp5, Ccl11
19	4	4	Frmd4b, Grn, Mt2, Ctsz, Hgfac
20	4	5	Icosl, Naip5, Irf5, Txndc5
21	5	1	Prss23, Cd81, Pdzd3, Slc27a2
22	5	2	Myh3, Fscn1, Mertk, Cd74, Lepr, Lpcat1, H2-DMA
23	5	3	Ms4a7, Spon1, Papss2, Bgn, Ccdc88a, Relb, Gns, Procr, Mospd2, Ankrd33b, Nfkbie, Pawr
24	5	4	Csf2rb, Lpl, 2610528A11Rik, Ryr1, Cebpdl, Ly6d, Myadm, Ncf1, Serpina10, Cp, Slc4a8
25	5	5	Chi3l3, Mrc1, Gja1, Rsph1, Pkd2, Lrrk1, Steap4, Emp1, Rgs18, Uap1, Ccr1, Prcp

4.2.5 Acquisition of multiple tumorigenic programs

Since the tumors in L-DKO mice share lineage identity with iNKT or CD1dTet cells that have undergone persistent expansion starting at neonatal age, it is likely that the acquisition of secondary and tertiary oncogenic mutations ultimately leads to tumor development. The clear transcriptional distinction observed between the tumor samples could indicate the existence of multiple, varying aberrant pathways that may be reflective of the history of accumulated mutations in the parent cells (Figure 28c). This hypothesis is also supported by their oligoclonal expansion patterns reported by us previously [190].

I next did pathway analysis to identify the pathways over-represented in the tumor samples. This analysis revealed many aberrant pathways specific to the tumor samples, particularly corresponding to cytokine-cytokine receptor interaction, NF- κ B signaling and transcriptional misregulation in cancer (Figure 30a). Interestingly, the cytokine-cytokine receptor pathway, which was downregulated in premalignant iNKT cells, was found to be upregulated upon tumorigenesis. This included genes such as *Csf1r* that aid in proliferation and can act as proto-oncogenes [263], as well as chemokines and their ligands (*Cxcl12*, *Ccl11*, *Ccl9*, *Ccr1*, *Ccr5*) which have been shown to contribute to metastasis (Figure 30b) [264, 265]. *Icam1* and *Cxcl12*, which have been reported to regulate iNKT homing to the liver and bone marrow (BM) [266, 267], were also upregulated in the tumor cells, and therefore could potentially promote

accumulation of iNKT cells and tumor formation [268, 269] in these organs. *Lta* and *Ltb* have been shown to play roles in iNKT thymic emigration [270], and were significantly upregulated in iNKT tumors. Furthermore, there were several interesting genes involved in the transcriptional misregulation of cancer, such as *Runx2*, *Mmp3*, *Mmp9*, *Egln3* and *Vegfa* [271-274]. I also found upregulation of the *Id2* transcript in iNKT tumors as compared to premalignant iNKT cells (Figure 30b). This could represent overexpression of *Id2* exon 3, which is the only remaining exon in *Id2^{fl}* mice [275].

I then focused my attention to identifying interesting genes that were unique to L-DKO iNKT or iNKT tumor samples (Figure 30c). Genes supporting cell proliferation, such as *Cd74* [276] and *Adm* [259], were downregulated in premalignant iNKT cells but highly upregulated in iNKT tumors. The tumor suppressor, *Pcgf2* [277], was also significantly downregulated in iNKT tumor cells. Interestingly, *Ccr7* [278] and *Egr2* [51, 140], which have both been shown to play critical roles in iNKT development, were specifically upregulated in premalignant iNKT cells. Therefore, gene expression profiling revealed several known oncogenic pathways that may contribute to the development of iNKT tumors. This analysis also indicated that there were distinct pathways involved in the pre-malignant expansion of iNKT cells and the ultimate transformation of iNKT cells leading to uncontrolled tumor growth (Figure 31). Since the microarray data showed divergence of the CD1dTet⁺ tumor sample, I also did RNA-Seq analysis of two CD1dTet⁺ tumor samples, which showed shared and distinct pathways

with the iNKT tumor samples (data not shown). Comparison of these identified genes and pathways, however, revealed only a small overlap with other innate-like lymphomas described by other groups (Table 4) [137, 279].

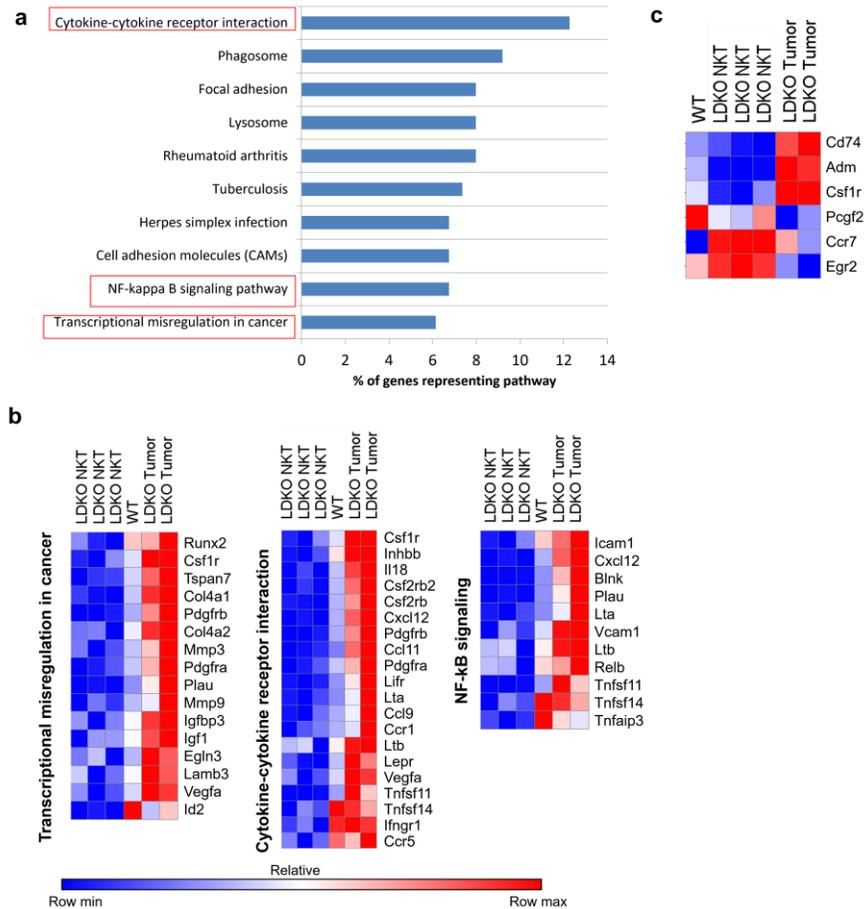


Figure 30: Pathway and gene expression analysis in iNKT tumors

(a) Pathways overrepresented by genes with greater than 2 fold gene expression in iNKT tumor samples according to gene sets annotated by Kyoto Encyclopedia of Genes and Genomes (KEGG) [280]. Percentages represent fraction of genes from each pathway that are overexpressed in the samples. (b) Heat maps showing relative log fold change of gene expression with respect to WT DP cells for genes from pathways identified in the above analysis. (c) Heat map showing select genes with significant differential expression among premalignant and tumor samples.

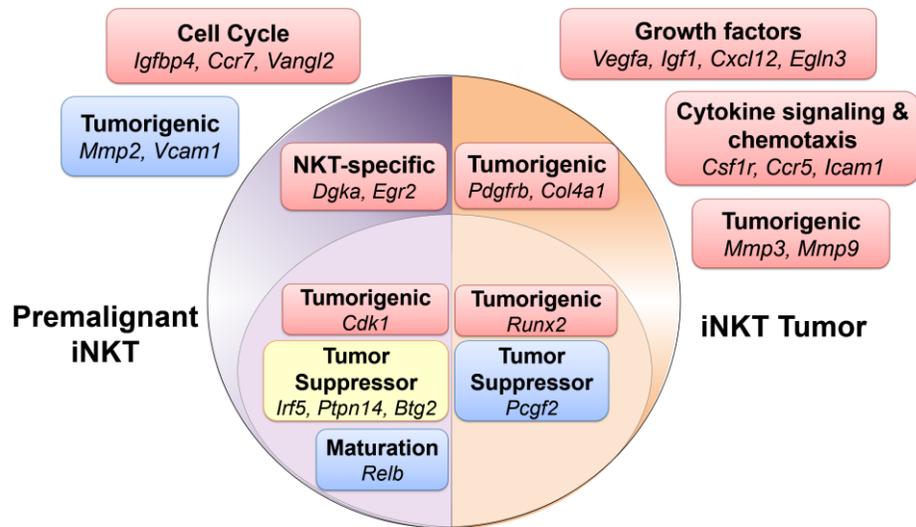


Figure 31: Distinct pathways drive expansion and tumorigenesis in L-DKO mice

Graphic depicting a few key overrepresented pathways in premalignant iNKT cells from 20 day old L-DKO mice (purple), or in CD1dTet⁺ iNKT lymphoma cells from older L-DKO mice (orange). Downregulation (blue), upregulation (red) or partial upregulation and downregulation (yellow) of selected genes from the pathways is also shown.

Table 4: List of genes shared in key pathways with other innate-like tumors

Lymphoma type L-DKO (Source)	Innate-like T _{FH} tumor [137] (Miyazaki et al)		CD1d-restricted tumor [279] (Bachy et al)	
	Cytokine-cytokine receptor signaling (downregulated)	JAK-STAT signaling (downregulated)	Positive regulation of gene expression (upregulated)	Cell cycle (upregulated)
iNKT (Microarray)	Il7r Il10ra Il12rb2 Lifr Pdgfrb Il18 Cxcl12 Ccr8 Tnfrsf1b	Il7r Il10ra Il12rb2 Lifr Stat4	Zxdc Srpk2 Akt2 Phip Tfdp2 Ppm1a Sirt2 Prkab1 Brd7 Cdk6	Srpk2 akt2 phip tfdp2 ppm1a sirt2 prkab1 brd7 cdk6
	Cell cycle (Upregulated)	Pathways in Cancer (Upregulated)		
CD1dTet (RNASeq)	Ywhaq Ywhaz Ywhab Ywhae Ccne2 Cdk1 Smc1a Mdm2 Pcna Mcm6 Mcm7 Myc Chek1 Ccnh	Mapk1 Cdc42 Myc Hsp90aa1 Hsp90ab1 Chuk Hsp90b1 Jak1 Hif1a Tceb1 Tceb2 Tpr Stat3		

Additionally, to determine similarity with human iNKT tumors, I compared our L-DKO tumor data to that from a recent publication characterizing key driver mutations and pathways in patients with iNKT lymphomas [281]. I found many of the mutated genes in human patients from their study [281] to also be dysregulated in the L-DKO tumor model. The shared genes had differential expression patterns in L-DKO iNKT tumor samples and neonatal iNKT cells as compared to WT iNKT cells (Figure 32). The publication also implicated the upregulation of the NF- κ B pathway in driving tumorigenesis in a subset of patients with poor prognosis. [281] Several genes of the NF- κ B pathway were also found to be uniquely upregulated or downregulated in the L-DKO iNKT tumors, but not in the premalignant, neonatal iNKT cells (Figure 30b). This data suggests that this is a potential mouse model to investigate mechanisms of iNKT and innate-like tumors in humans.

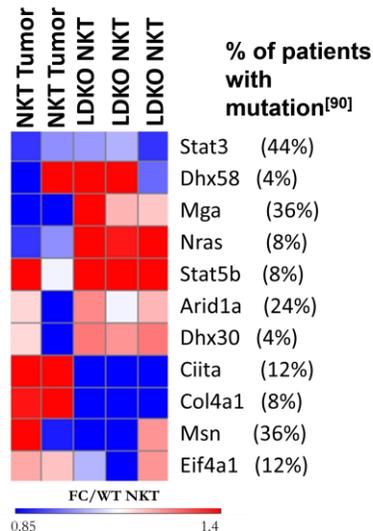


Figure 32: Resemblance to human NK/T tumor

Fold change in gene expression with respect to WT iNKT cells, of genes that are implicated in human iNKT tumors [281], and also dysregulated in L-DKO iNKT tumors. Percentages in top panel indicate the percent of patients with iNKT lymphomas (total n = 25) that have mutations in the listed genes, as characterized by Jiang et al [281].

4.3 Conclusions and discussion

A previous study of *Id3*^{-/-} mice has revealed a tumor suppressor role of ID3 in the development of HSTCL-like tumors [168]. We have previously reported iNKT and innate-like T cell tumors upon deletion of both *Id2* and *Id3* [190]. It is interesting to note the difference in kinetics of lymphoma development and progression in *Id3*^{-/-} mice and L-DKO mice. *Id3*^{-/-} mice often develop autoimmune diseases [195], but tumor development is much more infrequent and delayed, such that these mice live for at least one year. On the other hand, L-DKO mice start dying of tumor by 3 months of age. This rapid development of $\alpha\beta$ T cell lymphomas in *Id2*^{f/f}*Id3*^{f/f}IL7RCre⁺ mice [137], and iNKT

and CD1dTet tumors in L-DKO mice argue in favor of Id2 playing novel compensatory and non-redundant roles in the regulation and suppression of tumorigenesis of developing T cells in the murine thymus along with Id3.

In line with previous publications that have demonstrated the role of Id proteins in suppressing the development of innate-like $\gamma\delta$ and iNKT cell development [62, 139], as well as the initiation of an innate-like transcription program described in the previous chapter, we observe expansion of another population of innate-like T cells, namely the CD1dTet⁺ T cells, in L-DKO mice. Interestingly, this population bears resemblance to the CD4⁺PLZF⁺ cells reported to expand in mice deficient in *Itk*, *Id3* or with early deletion of both Id2 and Id3 [65, 137, 282]. Additional characterization of surface markers and gene expression programs in these cells would be important for comparing these innate-like T cells, and for understanding the cellular origin of iNKT and innate-like T cell tumors in humans [239, 245, 246, 283].

To delve into the mechanism(s) of tumor formation in L-DKO mice, I performed meta-analysis by combining our L-DKO microarray data with WT data. This allowed us to perform direct comparisons between our mutant cells and WT iNKT cells, which were originally not included in our microarray analysis. However, this approach limits the analysis to only the common gene probes in the two microarray datasets, which can lead to the omission of some potentially interesting genes in this model. It is also important to note that the expanded iNKT cells in L-DKO mice are a heterogeneous population

consisting primarily of stage 1 and stage 2 iNKT cells [70]. Due to the lack of availability of an appropriate control population, I used mature iNKT cells from B6 mice as reference. Despite this distinction between the populations, it is reasonable to make this comparison as a reflection of changes in transcriptional programs in iNKT cells lacking Id2 and Id3 that lead to lymphoma development. With the microarray datasets combined, I was able to observe the deviation in the genetic program in L-DKO iNKT cells as compared to WT iNKT cells. I found that several cell cycle genes were upregulated, while pro-apoptotic genes were downregulated in neonatal L-DKO iNKT cells. This modified program allowed their dramatic expansion, but also kept their tumorigenic potential in check. I inferred that these expanding neonatal innate-like cells are stochastically driven towards tumorigenesis via different pathways, giving rise to heterogeneous tumors in these mice.

Among the L-DKO tumors, I observed dysregulation of genes in pathways such as transcriptional misregulation in cancer and cytokine-cytokine receptor interactions, as well as others that are commonly overexpressed in various cancer types. While it is difficult to determine with certainty which genes contributed to, versus those that were upregulated as a result of lymphoma development, I verified the sharing of key genes and pathways in L-DKO tumors with human iNKT tumors. Therefore, these mice serve as an appropriate mouse model to study iNKT and innate-like tumors. Furthermore, the striking resemblance between all premalignant iNKT samples and the divergence of

tumor samples leads us to the enticing prospect of treating tumors by identifying and targeting early tumorigenic pathways. Such a study of tumor initiation and gradual progression is only possible in a mouse model, and would be useful in determining common genes that lead to malignant transformation.

It is indeed an interesting proposition that *Id2* and *Id3*, through their suppression of innate-like cell fate, prevent unchecked expansion of iNKT and CD1dTet⁺ cells under WT conditions. Therefore, upon deletion of *Id2* and *Id3*, the rapid proliferation and expansion of these cells makes them prone to accrual of additional mutations leading to tumorigenesis by various mechanisms.

5. Conclusions

Although innate-like T cells normally represent only a small fraction of the thymic population, their indispensable roles in mounting rapid immune responses in different contexts warrants a holistic understanding of the regulation of their concurrent development with conventional T cells in the thymus. Transcriptional regulation by E proteins, and their antagonism by Id proteins, acts as gatekeepers to ensure appropriate T cell development at key checkpoints. However, derailment of these regulatory mechanisms in lymphocytes can drive lymphoma development.

Overall, Id and E protein dosage can influence lineage choice and differentiation in iNKT cells. For instance, the small iNKT population in WT mice suggests that low E protein activity after the pre-TCR checkpoint in WT mice is not sufficient to promote selection into the iNKT lineage. However it seems that deletion of Id2 and Id3 can enhance the E protein levels sufficiently to allow them to specifically promote the iNKT lineage fate choice [70]. In this dissertation, I characterized E2A-driven transcription programs that promote innate-like T cell development prior to TCR selection and independent of pre-TCR signaling, which are otherwise suppressed by Id proteins. I also delineated the pathways that drive iNKT expansion, versus those that can drive innate-like lymphomagenesis in Id-deficient mice.

Not surprisingly, phylogenetic analysis of innate-like T cells and their associated transcription factors indicates that these cells emerged much earlier than conventional T

cells in the course of evolution [215, 284]. Hence, I propose that innate-like lineage specification precedes conventional $\alpha\beta$ T cells in the thymus, and that evolutionary pressures necessitated Id-mediated suppression to ensure the predominance of conventional $\alpha\beta$ T cells. This data also suggests that Id proteins are potent suppressors of iNKT cell fate at the pre-TCR checkpoint.

6. Future Directions

Based on the data presented in this dissertation, there are other compelling questions that need to be addressed in the future.

6.1 Identifying downstream E2A targets critical for the large innate-like T cell populations in Id-deficient mice

This is the first report of E2A driving key genes for innate-like T cell development and function. The network analysis also serves as a rich database and framework for further exploration of novel downstream targets that drive or suppress these innate-like populations in Id-deficient and WT mice respectively. It would also be interesting to search for potential lineage-specific transcription factor(s) that cooperate with E2A in definitively bifurcating DP cells towards the iNKT lineage. Some preliminary candidates for studying these questions include Etv5, Ets1 and Runx1, based on their expression patterns, enrichment of binding motifs within E2A peaks, and supportive data in literature [48, 81, 285-288].

6.2 Does Id-deficiency have an impact on other innate-like T cells?

My study in Id-deficient study was primarily restricted to examination of iNKT and $\gamma\delta$ NKT cells. However, expansion of innate-like T_{FH} cells has also been reported in Id2/Id3 deficient mice [137]. Based on the observation that E2A can drive an innate-like transcriptional program in the absence of Id proteins, it would be interesting to study the impact of high E protein activity on the development of other innate-like T cells like

MAIT cells and IELs [134]. The proposed innate-like reference gene set defined here would also aid in the identification of lineage-specific transcription factors for different innate-like T cell populations.

6.3 Delineating the distinct roles and contribution of E2A in promoting the iNKT lineage prior to and upon TCR selection

Our group and others have observed large innate-like populations in various Id-deficient mouse models. Based on the summary in table 1, and results presented in chapter 3, it is also clear that the role of E proteins in driving the iNKT lineage is multifaceted. While on one hand an innate-like transcriptional program and TCR rearrangement bias in DN and DP cells increases the number of cells that can potentially undergo CD1d-mediated selection, these cells do undergo significant expansion after selection and commitment. Therefore, it becomes hard to tease out the extent to which these roles contribute to the iNKT population size. Not to mention, mouse models with most genetic perturbations of E and Id proteins will also have effects on conventional T cell development. For instance, we have previously reported impaired conventional T cell, particularly CD8 SP, development in L-DKO mice [129]. Ideally, mosaic mouse models with mutations in a smaller fraction of cells, with most cells in WT condition, would provide the most unbiased, physiological study of expansion of these cells. One limitation of such models can be deletion efficiency, and elimination of the small number of cells by immune surveillance, making it difficult to observe lymphoma development. Comparison of iNKT phenotypes between mice with Id protein deletion

at different stages like LckCre and E8^{III}-Cre [289] can also partly resolve the issue of examining impact at different checkpoints. Nonetheless, development of better mouse models and markers other than binding to the CD1d tetramer is needed for easy identification and possible tracking of iNKT precursors and proliferating cells. There is also potential for exploring the role of HEB in influencing iNKT development beyond cell survival and secondary rearrangements, such as through the dysregulation of metabolic genes and cytokines like IL17R β and IL12R β , which can also be expected to play important roles in this lineage [125].

6.4 How/do E proteins regulate the outcomes and timing of innate-like TCR α rearrangements?

My data suggests that high E2A activity can specifically bind to both E α enhancer and Trav11 promoter regions, which can possibly lead to the biased rearrangement favoring iNKT cell fate. I also found lower IL-7R and proliferation rates among Id-deficient DN4 cells, which can potentially activate premature TCR α rearrangement. However, the mechanism(s) by which E proteins promote the iNKT TCR rearrangement remains unclear.

One approach to studying the impact of E2A-mediated regulation on iNKT TCR rearrangement and lineage choice is to mutating the E-box sites within the E2A binding peaks for Trav11 and E α enhancer (Figure 13). Any impact on TCR rearrangement outcomes or iNKT population size would indicate a role for E2A in this process. Another approach would be to examine the possibility of early TCR α rearrangement in L-DKO

DN4 cells. By sorting these cells and doing RT-PCR, I have found increased V α 14-J α 18 transcripts in these cells over WT control DN4 cells. Better primer design to distinguish between germline transcription and rearranged product would be important for drawing a conclusion. Further, designing a technique a mouse model for detection and tracking of cells with rearranged V α 14-J α 18 transcripts would not only allow us to determine timing of rearrangement, but also serve as surrogate marker for precursors cells that are poised for CD1d selection. I think another interesting aspect to focus on in the future would be to study TCR α repertoire in innate-like LDKO pT α ^{-/-} DP cells to expand our appreciation of TCR α chains in other innate-like T cells.

6.5 Identifying associated mutations upon loss of Id2 and Id3

Id proteins and E proteins are powerful regulators of cell cycle and senescence, allowing them to interchangeably play tumor suppressive and oncogenic roles. In almost all cases, simple dysregulation of E and Id proteins is not sufficient to cause tumors, and a secondary mutation hit is often required. It is possible that the dysregulation of E and Id proteins causes a growth arrest at a stage where most cells are cycling, leading to accumulation of mutations that predispose these cells to lymphomageneses. This predisposition to accumulating mutations to undergo malignant transformation, referred to as cellular pliancy, however, differs between different cell types and developmental stages. Profiling of associated mutations, such as Id3 and c-Myc mutations reported in Burkitt's lymphoma, can aid in our understanding

of major pathways that drive NK/T and CD1dTet⁺ tumorigenesis in Id-deficient mice. Based on physiological expression patterns of identified genes, we can also infer the stage of development for the second mutational hit. For therapeutic applications, it would also be crucial to identify targets that can block innate-like T cell tumor development and metastasis. For instance, the efficacy of CDK4/6 inhibitors can be tested in this system.

Overall, answering these questions will improve our understanding of the basic biology of innate-like T cell development that would help us design better strategies for controlled applications of these cells, as well as improve therapies for NK/T tumors.

References

1. Godfrey, D.I., et al., *The burgeoning family of unconventional T cells*. Nat Immunol, 2015. **16**(11): p. 1114-23.
2. Dakic, A., et al., *PU.1 regulates the commitment of adult hematopoietic progenitors and restricts granulopoiesis*. J Exp Med, 2005. **201**(9): p. 1487-502.
3. Iwasaki, H., et al., *Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation*. Blood, 2005. **106**(5): p. 1590-600.
4. Dionne, C.J., et al., *Subversion of T lineage commitment by PU.1 in a clonal cell line system*. Dev Biol, 2005. **280**(2): p. 448-66.
5. Rothenberg, E.V., *T cell lineage commitment: identity and renunciation*. J Immunol, 2011. **186**(12): p. 6649-55.
6. David-Fung, E.S., et al., *Transcription factor expression dynamics of early T-lymphocyte specification and commitment*. Dev Biol, 2009. **325**(2): p. 444-67.
7. Lefort, N., et al., *Short exposure to Notch ligand Delta-4 is sufficient to induce T-cell differentiation program and to increase the T cell potential of primary human CD34+ cells*. Exp Hematol, 2006. **34**(12): p. 1720-9.
8. Rothenberg, E.V. and D.D. Scripture-Adams, *Competition and collaboration: GATA-3, PU.1, and Notch signaling in early T-cell fate determination*. Semin Immunol, 2008. **20**(4): p. 236-46.
9. Schmitt, T.M. and J.C. Zuniga-Pflucker, *Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro*. Immunity, 2002. **17**(6): p. 749-56.
10. Dudley, D.D., et al., *Mechanism and control of V(D)J recombination versus class switch recombination: similarities and differences*. Adv Immunol, 2005. **86**: p. 43-112.
11. Schatz, D.G. and Y. Ji, *Recombination centres and the orchestration of V(D)J recombination*. Nat Rev Immunol, 2011. **11**(4): p. 251-63.
12. Liu, Y., L. Zhang, and S. Desiderio, *Temporal and spatial regulation of V(D)J recombination: interactions of extrinsic factors with the RAG complex*. Adv Exp Med Biol, 2009. **650**: p. 157-65.

13. Schatz, D.G. and P.C. Swanson, *V(D)J recombination: mechanisms of initiation*. *Annu Rev Genet*, 2011. **45**: p. 167-202.
14. Krangel, M.S., *Mechanics of T cell receptor gene rearrangement*. *Curr Opin Immunol*, 2009. **21**(2): p. 133-9.
15. Wang, X., et al., *Regulation of Tcrb recombination ordering by c-Fos-dependent RAG deposition*. *Nat Immunol*, 2008. **9**(7): p. 794-801.
16. Jackson, A.M. and M.S. Krangel, *Turning T-cell receptor beta recombination on and off: more questions than answers*. *Immunol Rev*, 2006. **209**: p. 129-41.
17. Khor, B. and B.P. Sleckman, *Allelic exclusion at the TCRbeta locus*. *Curr Opin Immunol*, 2002. **14**(2): p. 230-4.
18. Prinz, I., B. Silva-Santos, and D.J. Pennington, *Functional development of gammadelta T cells*. *Eur J Immunol*, 2013. **43**(8): p. 1988-94.
19. Xiong, N. and D.H. Raulet, *Development and selection of gammadelta T cells*. *Immunol Rev*, 2007. **215**: p. 15-31.
20. Carding, S.R. and P.J. Egan, *Gammadelta T cells: functional plasticity and heterogeneity*. *Nat Rev Immunol*, 2002. **2**(5): p. 336-45.
21. Melichar, H.J., et al., *Regulation of gammadelta versus alphabeta T lymphocyte differentiation by the transcription factor SOX13*. *Science*, 2007. **315**(5809): p. 230-3.
22. Hayes, S.M., E.W. Shores, and P.E. Love, *An architectural perspective on signaling by the pre-, alphabeta and gammadelta T cell receptors*. *Immunol Rev*, 2003. **191**: p. 28-37.
23. Hernandez-Munain, C., B.P. Sleckman, and M.S. Krangel, *A developmental switch from TCR delta enhancer to TCR alpha enhancer function during thymocyte maturation*. *Immunity*, 1999. **10**(6): p. 723-33.
24. Carico, Z.M., et al., *Tcrd Rearrangement Redirects a Processive Tcra Recombination Program to Expand the Tcra Repertoire*. *Cell Rep*, 2017. **19**(10): p. 2157-2173.
25. Hawwari, A. and M.S. Krangel, *Role for rearranged variable gene segments in directing secondary T cell receptor alpha recombination*. *Proc Natl Acad Sci U S A*, 2007. **104**(3): p. 903-7.

26. Guo, J., et al., *Regulation of the TCR α repertoire by the survival window of CD4(+)CD8(+) thymocytes*. Nat Immunol, 2002. **3**(5): p. 469-76.
27. Carpenter, A.C. and R. Bosselut, *Decision checkpoints in the thymus*. Nat Immunol, 2010. **11**(8): p. 666-73.
28. Borowski, C., et al., *Pre-TCR α and TCR α are not interchangeable partners of TCR β during T lymphocyte development*. J Exp Med, 2004. **199**(5): p. 607-15.
29. Yamasaki, S. and T. Saito, *Molecular basis for pre-TCR-mediated autonomous signaling*. Trends Immunol, 2007. **28**(1): p. 39-43.
30. Engel, I. and C. Murre, *E2A proteins enforce a proliferation checkpoint in developing thymocytes*. EMBO J, 2004. **23**(1): p. 202-11.
31. Dose, M., et al., *c-Myc mediates pre-TCR-induced proliferation but not developmental progression*. Blood, 2006. **108**(8): p. 2669-77.
32. Voll, R.E., et al., *NF-kappa B activation by the pre-T cell receptor serves as a selective survival signal in T lymphocyte development*. Immunity, 2000. **13**(5): p. 677-89.
33. Campbell, K.J., et al., *Elevated Mcl-1 inhibits thymocyte apoptosis and alters thymic selection*. Cell Death Differ, 2012. **19**(12): p. 1962-71.
34. Sohn, S.J., J. Thompson, and A. Winoto, *Apoptosis during negative selection of autoreactive thymocytes*. Curr Opin Immunol, 2007. **19**(5): p. 510-5.
35. Gascoigne, N.R. and E. Palmer, *Signaling in thymic selection*. Curr Opin Immunol, 2011. **23**(2): p. 207-12.
36. Daniels, M.A., et al., *Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling*. Nature, 2006. **444**(7120): p. 724-9.
37. Kane, L.P. and S.M. Hedrick, *A role for calcium influx in setting the threshold for CD4+CD8+ thymocyte negative selection*. J Immunol, 1996. **156**(12): p. 4594-601.
38. Holzapfel, K.L., et al., *Antigen-dependent versus -independent activation of invariant NKT cells during infection*. J Immunol, 2014. **192**(12): p. 5490-8.
39. Brigl, M., et al., *Innate and cytokine-driven signals, rather than microbial antigens, dominate in natural killer T cell activation during microbial infection*. J Exp Med, 2011. **208**(6): p. 1163-77.

40. Lantz, O. and A. Bendelac, *An invariant T cell receptor alpha chain is used by a unique subset of major histocompatibility complex class I-specific CD4+ and CD4-8- T cells in mice and humans*. J Exp Med, 1994. **180**(3): p. 1097-106.
41. Exley, M., et al., *Requirements for CD1d recognition by human invariant Valpha24+ CD4-CD8- T cells*. J Exp Med, 1997. **186**(1): p. 109-20.
42. Brossay, L., et al., *CD1d-mediated recognition of an alpha-galactosylceramide by natural killer T cells is highly conserved through mammalian evolution*. J Exp Med, 1998. **188**(8): p. 1521-8.
43. Bendelac, A., P.B. Savage, and L. Teyton, *The biology of NKT cells*. Annu Rev Immunol, 2007. **25**: p. 297-336.
44. Matsuda, J.L., et al., *CD1d-restricted iNKT cells, the 'Swiss-Army knife' of the immune system*. Curr Opin Immunol, 2008. **20**(3): p. 358-68.
45. Albacker, L.A., et al., *Invariant natural killer T cells recognize a fungal glycosphingolipid that can induce airway hyperreactivity*. Nat Med, 2013. **19**(10): p. 1297-304.
46. Cerundolo, V., et al., *Harnessing invariant NKT cells in vaccination strategies*. Nat Rev Immunol, 2009. **9**(1): p. 28-38.
47. Bezbradica, J.S., et al., *Commitment toward the natural T (iNKT) cell lineage occurs at the CD4+8+ stage of thymic ontogeny*. Proc Natl Acad Sci U S A, 2005. **102**(14): p. 5114-9.
48. Egawa, T., et al., *Genetic evidence supporting selection of the Valpha14i NKT cell lineage from double-positive thymocyte precursors*. Immunity, 2005. **22**(6): p. 705-16.
49. Gapin, L., et al., *NKT cells derive from double-positive thymocytes that are positively selected by CD1d*. Nat Immunol, 2001. **2**(10): p. 971-8.
50. Savage, A.K., et al., *The transcription factor PLZF directs the effector program of the NKT cell lineage*. Immunity, 2008. **29**(3): p. 391-403.
51. Seiler, M.P., et al., *Elevated and sustained expression of the transcription factors Egr1 and Egr2 controls NKT lineage differentiation in response to TCR signaling*. Nat Immunol, 2012. **13**(3): p. 264-71.

52. Godfrey, D.I. and S.P. Berzins, *Control points in NKT-cell development*. Nat Rev Immunol, 2007. **7**(7): p. 505-18.
53. Constantinides, M.G. and A. Bendelac, *Transcriptional regulation of the NKT cell lineage*. Curr Opin Immunol, 2013. **25**(2): p. 161-7.
54. Lee, Y.J., et al., *Steady-state production of IL-4 modulates immunity in mouse strains and is determined by lineage diversity of iNKT cells*. Nat Immunol, 2013. **14**(11): p. 1146-54.
55. Watarai, H., et al., *Development and function of invariant natural killer T cells producing T(h)2- and T(h)17-cytokines*. PLoS Biol, 2012. **10**(2): p. e1001255.
56. Coquet, J.M., et al., *Diverse cytokine production by NKT cell subsets and identification of an IL-17-producing CD4-NK1.1- NKT cell population*. Proc Natl Acad Sci U S A, 2008. **105**(32): p. 11287-92.
57. Azuara, V., et al., *A novel subset of adult gamma delta thymocytes that secretes a distinct pattern of cytokines and expresses a very restricted T cell receptor repertoire*. Eur J Immunol, 1997. **27**(2): p. 544-53.
58. Grigoriadou, K., L. Boucontet, and P. Pereira, *Most IL-4-producing gamma delta thymocytes of adult mice originate from fetal precursors*. J Immunol, 2003. **171**(5): p. 2413-20.
59. Kovalovsky, D., et al., *The BTB-zinc finger transcriptional regulator PLZF controls the development of invariant natural killer T cell effector functions*. Nat Immunol, 2008. **9**(9): p. 1055-64.
60. Kreslavsky, T., et al., *TCR-inducible PLZF transcription factor required for innate phenotype of a subset of gammadelta T cells with restricted TCR diversity*. Proc Natl Acad Sci U S A, 2009. **106**(30): p. 12453-8.
61. Verykokakis, M., et al., *Inhibitor of DNA binding 3 limits development of murine slam-associated adaptor protein-dependent "innate" gammadelta T cells*. PLoS One, 2010. **5**(2): p. e9303.
62. Alonzo, E.S. and D.B. Sant'Angelo, *Development of PLZF-expressing innate T cells*. Curr Opin Immunol, 2011. **23**(2): p. 220-7.
63. Yin, C.C., et al., *The Tec kinase ITK regulates thymic expansion, emigration, and maturation of gammadelta NKT cells*. J Immunol, 2013. **190**(6): p. 2659-69.

64. Besin, G., et al., *Dok-1 overexpression promotes development of gammadelta natural killer T cells*. Eur J Immunol, 2012. **42**(9): p. 2491-504.
65. Verykokakis, M., et al., *SAP protein-dependent natural killer T-like cells regulate the development of CD8(+) T cells with innate lymphocyte characteristics*. Immunity, 2010. **33**(2): p. 203-15.
66. Kronenberg, M. and L. Gapin, *The unconventional lifestyle of NKT cells*. Nat Rev Immunol, 2002. **2**(8): p. 557-68.
67. MacDonald, H.R., *Development and selection of NKT cells*. Curr Opin Immunol, 2002. **14**(2): p. 250-4.
68. Makino, Y., et al., *Development of Valpha4+ NK T cells in the early stages of embryogenesis*. Proc Natl Acad Sci U S A, 1996. **93**(13): p. 6516-20.
69. Dashtsoodol, N., et al., *Identification of CD4(-)CD8(-) double-negative natural killer T cell precursors in the thymus*. PLoS One, 2008. **3**(11): p. e3688.
70. Li, J., et al., *Combined deletion of id2 and id3 genes reveals multiple roles for e proteins in invariant NKT cell development and expansion*. J Immunol, 2013. **191**(10): p. 5052-64.
71. Pear, W.S., L. Tu, and P.L. Stein, *Lineage choices in the developing thymus: choosing the T and NKT pathways*. Curr Opin Immunol, 2004. **16**(2): p. 167-73.
72. Matsuda, J.L. and L. Gapin, *Developmental program of mouse Valpha14i NKT cells*. Curr Opin Immunol, 2005. **17**(2): p. 122-30.
73. Eberl, G., B. Lowin-Kropf, and H.R. MacDonald, *Cutting edge: NKT cell development is selectively impaired in Fyn- deficient mice*. J Immunol, 1999. **163**(8): p. 4091-4.
74. Kisielow, J., et al., *Evidence for the divergence of innate and adaptive T-cell precursors before commitment to the alphabeta and gammadelta lineages*. Blood, 2011. **118**(25): p. 6591-600.
75. Taniguchi, M., et al., *Essential requirement of an invariant V alpha 14 T cell antigen receptor expression in the development of natural killer T cells*. Proc Natl Acad Sci U S A, 1996. **93**(20): p. 11025-8.

76. Griewank, K., et al., *Homotypic interactions mediated by Slamf1 and Slamf6 receptors control NKT cell lineage development*. *Immunity*, 2007. **27**(5): p. 751-62.
77. Shimamura, M., et al., *Lack of directed V alpha 14-J alpha 281 rearrangements in NK1+ T cells*. *Eur J Immunol*, 1997. **27**(6): p. 1576-9.
78. Hager, E., et al., *Multiple constraints at the level of TCRalpha rearrangement impact Valpha14i NKT cell development*. *J Immunol*, 2007. **179**(4): p. 2228-34.
79. Zhou, D., et al., *Lysosomal glycosphingolipid recognition by NKT cells*. *Science*, 2004. **306**(5702): p. 1786-9.
80. Moran, A.E., et al., *T cell receptor signal strength in Treg and iNKT cell development demonstrated by a novel fluorescent reporter mouse*. *J Exp Med*, 2011. **208**(6): p. 1279-89.
81. Mao, A.P., et al., *A shared Runx1-bound Zbtb16 enhancer directs innate and innate-like lymphoid lineage development*. *Nat Commun*, 2017. **8**(1): p. 863.
82. McNab, F.W., et al., *The influence of CD1d in postselection NKT cell maturation and homeostasis*. *J Immunol*, 2005. **175**(6): p. 3762-8.
83. Benlagha, K., et al., *A thymic precursor to the NK T cell lineage*. *Science*, 2002. **296**(5567): p. 553-5.
84. Verykokakis, M., et al., *Essential functions for ID proteins at multiple checkpoints in invariant NKT cell development*. *J Immunol*, 2013. **191**(12): p. 5973-83.
85. D'Cruz, L.M., et al., *E and Id proteins influence invariant NKT cell sublineage differentiation and proliferation*. *J Immunol*, 2014. **192**(5): p. 2227-36.
86. Hu, T., et al., *Increased level of E protein activity during invariant NKT development promotes differentiation of invariant NKT2 and invariant NKT17 subsets*. *J Immunol*, 2013. **191**(10): p. 5065-73.
87. Engel, I. and C. Murre, *The function of E- and Id proteins in lymphocyte development*. *Nat Rev Immunol*, 2001. **1**(3): p. 193-9.
88. de Pooter, R.F. and B.L. Kee, *E proteins and the regulation of early lymphocyte development*. *Immunol Rev*, 2010. **238**(1): p. 93-109.
89. Kee, B.L., *E and ID proteins branch out*. *Nat Rev Immunol*, 2009. **9**(3): p. 175-84.

90. Loveys, D.A., M.B. Streiff, and G.J. Kato, *E2A basic-helix-loop-helix transcription factors are negatively regulated by serum growth factors and by the Id3 protein*. *Nucleic Acids Res*, 1996. **24**(14): p. 2813-20.
91. Massari, M.E. and C. Murre, *Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms*. *Mol Cell Biol*, 2000. **20**(2): p. 429-40.
92. Lasorella, A., R. Benezra, and A. Iavarone, *The ID proteins: master regulators of cancer stem cells and tumour aggressiveness*. *Nat Rev Cancer*, 2014. **14**(2): p. 77-91.
93. Perk, J., A. Iavarone, and R. Benezra, *Id family of helix-loop-helix proteins in cancer*. *Nat Rev Cancer*, 2005. **5**(8): p. 603-14.
94. Benezra, R., et al., *The protein Id: a negative regulator of helix-loop-helix DNA binding proteins*. *Cell*, 1990. **61**(1): p. 49-59.
95. Langlands, K., et al., *Differential interactions of Id proteins with basic-helix-loop-helix transcription factors*. *J Biol Chem*, 1997. **272**(32): p. 19785-93.
96. Roschger, C. and C. Cabrele, *The Id-protein family in developmental and cancer-associated pathways*. *Cell Commun Signal*, 2017. **15**(1): p. 7.
97. Dias, S., et al., *E2A proteins promote development of lymphoid-primed multipotent progenitors*. *Immunity*, 2008. **29**(2): p. 217-27.
98. Ng, S.Y., et al., *Genome-wide lineage-specific transcriptional networks underscore Ikaros-dependent lymphoid priming in hematopoietic stem cells*. *Immunity*, 2009. **30**(4): p. 493-507.
99. Semerad, C.L., et al., *E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors*. *Proc Natl Acad Sci U S A*, 2009. **106**(6): p. 1930-5.
100. Spooner, C.J., et al., *A recurrent network involving the transcription factors PU.1 and Gfi1 orchestrates innate and adaptive immune cell fates*. *Immunity*, 2009. **31**(4): p. 576-86.
101. Engel, I., et al., *Early thymocyte development is regulated by modulation of E2A protein activity*. *J Exp Med*, 2001. **194**(6): p. 733-45.
102. Murre, C., *Role of helix-loop-helix proteins in lymphocyte development*. *Cold Spring Harb Symp Quant Biol*, 1999. **64**: p. 39-44.

103. Miyazaki, M., et al., *The opposing roles of the transcription factor E2A and its antagonist Id3 that orchestrate and enforce the naïve fate of T cells*. Nat Immunol, 2011. **12**(10): p. 992-1001.
104. Murre, C., *Helix-loop-helix proteins and lymphocyte development*. Nat Immunol, 2005. **6**(11): p. 1079-86.
105. Wojciechowski, J., et al., *E2A and HEB are required to block thymocyte proliferation prior to pre-TCR expression*. J Immunol, 2007. **178**(9): p. 5717-26.
106. Jones, M.E. and Y. Zhuang, *Stage-specific functions of E-proteins at the beta-selection and T-cell receptor checkpoints during thymocyte development*. Immunol Res, 2011. **49**(1-3): p. 202-15.
107. Bain, G., et al., *Regulation of the helix-loop-helix proteins, E2A and Id3, by the Ras-ERK MAPK cascade*. Nat Immunol, 2001. **2**(2): p. 165-71.
108. Bain, G., et al., *E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas*. Mol Cell Biol, 1997. **17**(8): p. 4782-91.
109. Yan, W., et al., *High incidence of T-cell tumors in E2A-null mice and E2A/Id1 double-knockout mice*. Mol Cell Biol, 1997. **17**(12): p. 7317-27.
110. Barndt, R., M.F. Dai, and Y. Zhuang, *A novel role for HEB downstream or parallel to the pre-TCR signaling pathway during alpha beta thymopoiesis*. J Immunol, 1999. **163**(6): p. 3331-43.
111. Barndt, R.J., M. Dai, and Y. Zhuang, *Functions of E2A-HEB heterodimers in T-cell development revealed by a dominant negative mutation of HEB*. Mol Cell Biol, 2000. **20**(18): p. 6677-85.
112. Wikstrom, I., et al., *A role for E2-2 at the DN3 stage of early thymopoiesis*. Mol Immunol, 2008. **45**(11): p. 3302-11.
113. Miyazaki, K., M. Miyazaki, and C. Murre, *The establishment of B versus T cell identity*. Trends Immunol, 2014. **35**(5): p. 205-10.
114. Engel, I. and C. Murre, *Ectopic expression of E47 or E12 promotes the death of E2A-deficient lymphomas*. Proc Natl Acad Sci U S A, 1999. **96**(3): p. 996-1001.
115. Sicinska, E., et al., *Requirement for cyclin D3 in lymphocyte development and T cell leukemias*. Cancer Cell, 2003. **4**(6): p. 451-61.

116. Schwartz, R., et al., *Gene expression patterns define novel roles for E47 in cell cycle progression, cytokine-mediated signaling, and T lineage development*. Proc Natl Acad Sci U S A, 2006. **103**(26): p. 9976-81.
117. Morrow, M.A., et al., *Overexpression of the Helix-Loop-Helix protein Id2 blocks T cell development at multiple stages*. Mol Immunol, 1999. **36**(8): p. 491-503.
118. Kim, D., X.C. Peng, and X.H. Sun, *Massive apoptosis of thymocytes in T-cell-deficient Id1 transgenic mice*. Mol Cell Biol, 1999. **19**(12): p. 8240-53.
119. Ikawa, T., et al., *E proteins and Notch signaling cooperate to promote T cell lineage specification and commitment*. J Exp Med, 2006. **203**(5): p. 1329-42.
120. Yashiro-Ohtani, Y., et al., *Pre-TCR signaling inactivates Notch1 transcription by antagonizing E2A*. Genes Dev, 2009. **23**(14): p. 1665-76.
121. Tremblay, M., et al., *Regulation of pT alpha gene expression by a dosage of E2A, HEB, and SCL*. J Biol Chem, 2003. **278**(15): p. 12680-7.
122. Schlissel, M., A. Voronova, and D. Baltimore, *Helix-loop-helix transcription factor E47 activates germ-line immunoglobulin heavy-chain gene transcription and rearrangement in a pre-T-cell line*. Genes Dev, 1991. **5**(8): p. 1367-76.
123. Agata, Y., et al., *Regulation of T cell receptor beta gene rearrangements and allelic exclusion by the helix-loop-helix protein, E47*. Immunity, 2007. **27**(6): p. 871-84.
124. Sun, Z., et al., *Requirement for RORgamma in thymocyte survival and lymphoid organ development*. Science, 2000. **288**(5475): p. 2369-73.
125. D'Cruz, L.M., et al., *An essential role for the transcription factor HEB in thymocyte survival, Tcr α rearrangement and the development of natural killer T cells*. Nat Immunol, 2010. **11**(3): p. 240-9.
126. Rivera, R.R., et al., *Thymocyte selection is regulated by the helix-loop-helix inhibitor protein, Id3*. Immunity, 2000. **12**(1): p. 17-26.
127. Jones, M.E. and Y. Zhuang, *Acquisition of a functional T cell receptor during T lymphocyte development is enforced by HEB and E2A transcription factors*. Immunity, 2007. **27**(6): p. 860-70.
128. Bain, G., et al., *Thymocyte maturation is regulated by the activity of the helix-loop-helix protein, E47*. J Exp Med, 1999. **190**(11): p. 1605-16.

129. Jones-Mason, M.E., et al., *E protein transcription factors are required for the development of CD4(+) lineage T cells*. *Immunity*, 2012. **36**(3): p. 348-61.
130. Ueda-Hayakawa, I., J. Mahlios, and Y. Zhuang, *Id3 restricts the developmental potential of gamma delta lineage during thymopoiesis*. *J Immunol*, 2009. **182**(9): p. 5306-16.
131. Jones, M.E. and Y. Zhuang, *Regulation of V(D)J recombination by E-protein transcription factors*. *Adv Exp Med Biol*, 2009. **650**: p. 148-56.
132. Bain, G., et al., *Positive and negative regulation of V(D)J recombination by the E2A proteins*. *J Exp Med*, 1999. **189**(2): p. 289-300.
133. Rankin, L. and G.T. Belz, *Diverse roles of inhibitor of differentiation 2 in adaptive immunity*. *Clin Dev Immunol*, 2011. **2011**: p. 281569.
134. Verykokakis, M., E.C. Zook, and B.L. Kee, *ID'ing innate and innate-like lymphoid cells*. *Immunol Rev*, 2014. **261**(1): p. 177-97.
135. Monticelli, L.A., et al., *Transcriptional regulator Id2 controls survival of hepatic NKT cells*. *Proc Natl Acad Sci U S A*, 2009. **106**(46): p. 19461-6.
136. Alonzo, E.S., et al., *Development of promyelocytic zinc finger and ThPOK-expressing innate gamma delta T cells is controlled by strength of TCR signaling and Id3*. *J Immunol*, 2010. **184**(3): p. 1268-79.
137. Miyazaki, M., et al., *The E-Id protein axis modulates the activities of the PI3K-AKT-mTORC1-Hif1a and c-myc/p19Arf pathways to suppress innate variant TFH cell development, thymocyte expansion, and lymphomagenesis*. *Genes Dev*, 2015. **29**(4): p. 409-25.
138. Zhang, B., et al., *Id3 and Id2 act as a dual safety mechanism in regulating the development and population size of innate-like gammadelta T cells*. *J Immunol*, 2014. **192**(3): p. 1055-63.
139. Roy, S. and Y. Zhuang, *Orchestration of invariant natural killer T cell development by E and Id proteins*. *Crit Rev Immunol*, 2015. **35**(1): p. 33-48.
140. Lazarevic, V., et al., *The gene encoding early growth response 2, a target of the transcription factor NFAT, is required for the development and maturation of natural killer T cells*. *Nat Immunol*, 2009. **10**(3): p. 306-13.

141. Dose, M., et al., *Intrathymic proliferation wave essential for Valpha14+ natural killer T cell development depends on c-Myc*. Proc Natl Acad Sci U S A, 2009. **106**(21): p. 8641-6.
142. Xi, H., et al., *Interplay between RORgammat, Egr3, and E proteins controls proliferation in response to pre-TCR signals*. Immunity, 2006. **24**(6): p. 813-26.
143. van Stipdonk, M.J., E.E. Lemmens, and S.P. Schoenberger, *Naive CTLs require a single brief period of antigenic stimulation for clonal expansion and differentiation*. Nat Immunol, 2001. **2**(5): p. 423-9.
144. Benlagha, K. and A. Bendelac, *CD1d-restricted mouse V alpha 14 and human V alpha 24 T cells: lymphocytes of innate immunity*. Semin Immunol, 2000. **12**(6): p. 537-42.
145. Godfrey, D.I., S. Stankovic, and A.G. Baxter, *Raising the NKT cell family*. Nat Immunol, 2010. **11**(3): p. 197-206.
146. Mycko, M.P., et al., *Selective requirement for c-Myc at an early stage of V(alpha)14i NKT cell development*. J Immunol, 2009. **182**(8): p. 4641-8.
147. Schmitz, R., et al., *Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics*. Nature, 2012. **490**(7418): p. 116-20.
148. Yang, C.Y., et al., *The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8+ T cell subsets*. Nat Immunol, 2011. **12**(12): p. 1221-9.
149. Sag, D., et al., *IL-10-producing NKT10 cells are a distinct regulatory invariant NKT cell subset*. J Clin Invest, 2014. **124**(9): p. 3725-40.
150. Lynch, L., et al., *Regulatory iNKT cells lack expression of the transcription factor PLZF and control the homeostasis of Treg cells and macrophages in adipose tissue*. Nat Immunol, 2015. **16**(1): p. 85-95.
151. Townsend, M.J., et al., *T-bet regulates the terminal maturation and homeostasis of NK and Valpha14i NKT cells*. Immunity, 2004. **20**(4): p. 477-94.
152. Matsuda, J.L., et al., *T-bet concomitantly controls migration, survival, and effector functions during the development of Valpha14i NKT cells*. Blood, 2006. **107**(7): p. 2797-805.
153. Xu, W. and B.L. Kee, *Growth factor independent 1B (Gfi1b) is an E2A target gene that modulates Gata3 in T-cell lymphomas*. Blood, 2007. **109**(10): p. 4406-14.

154. Kusunoki, T., et al., *TH2 dominance and defective development of a CD8+ dendritic cell subset in Id2-deficient mice*. J Allergy Clin Immunol, 2003. **111**(1): p. 136-42.
155. Zhang, F., et al., *Transcription of RORgammat in developing Th17 cells is regulated by E-proteins*. Mucosal Immunol, 2014. **7**(3): p. 521-32.
156. Norton, J.D., et al., *Id helix-loop-helix proteins in cell growth and differentiation*. Trends Cell Biol, 1998. **8**(2): p. 58-65.
157. Lasorella, A., T. Uo, and A. Iavarone, *Id proteins at the cross-road of development and cancer*. Oncogene, 2001. **20**(58): p. 8326-33.
158. Sikder, H.A., et al., *Id proteins in cell growth and tumorigenesis*. Cancer Cell, 2003. **3**(6): p. 525-30.
159. Hasskarl, J. and K. Munger, *Id proteins--tumor markers or oncogenes?* Cancer Biol Ther, 2002. **1**(2): p. 91-6.
160. Lee, S.H., et al., *The Id3/E47 axis mediates cell-cycle control in human pancreatic ducts and adenocarcinoma*. Mol Cancer Res, 2011. **9**(6): p. 782-90.
161. Tsuchiya, T., et al., *Targeting Id1 and Id3 inhibits peritoneal metastasis of gastric cancer*. Cancer Sci, 2005. **96**(11): p. 784-90.
162. Sharma, P., D. Patel, and J. Chaudhary, *Id1 and Id3 expression is associated with increasing grade of prostate cancer: Id3 preferentially regulates CDKN1B*. Cancer Med, 2012. **1**(2): p. 187-97.
163. Kamalian, L., et al., *Inhibition of tumourigenicity of small cell lung cancer cells by suppressing Id3 expression*. Int J Oncol, 2010. **37**(3): p. 595-603.
164. Gray, M.J., et al., *Therapeutic targeting of Id2 reduces growth of human colorectal carcinoma in the murine liver*. Oncogene, 2008. **27**(57): p. 7192-200.
165. Love, C., et al., *The genetic landscape of mutations in Burkitt lymphoma*. Nat Genet, 2012. **44**(12): p. 1321-5.
166. Richter, J., et al., *Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing*. Nat Genet, 2012. **44**(12): p. 1316-20.
167. Schmitz, R., et al., *Oncogenic mechanisms in Burkitt lymphoma*. Cold Spring Harb Perspect Med, 2014. **4**(2).

168. Li, J., et al., *Mutation of inhibitory helix-loop-helix protein Id3 causes gammadelta T-cell lymphoma in mice*. *Blood*, 2010. **116**(25): p. 5615-21.
169. Israel, M.A., et al., *Id gene expression as a key mediator of tumor cell biology*. *Cancer Res*, 1999. **59**(7 Suppl): p. 1726s-1730s.
170. Norton, J.D. and G.T. Atherton, *Coupling of cell growth control and apoptosis functions of Id proteins*. *Mol Cell Biol*, 1998. **18**(4): p. 2371-81.
171. Wong, Y.C., X. Wang, and M.T. Ling, *Id-1 expression and cell survival*. *Apoptosis*, 2004. **9**(3): p. 279-89.
172. Slattery, C., M.P. Ryan, and T. McMorrow, *E2A proteins: regulators of cell phenotype in normal physiology and disease*. *Int J Biochem Cell Biol*, 2008. **40**(8): p. 1431-6.
173. Tang, J., et al., *The helix-loop-helix protein id-1 delays onset of replicative senescence in human endothelial cells*. *Lab Invest*, 2002. **82**(8): p. 1073-9.
174. Neuhold, L.A. and B. Wold, *HLH forced dimers: tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id*. *Cell*, 1993. **74**(6): p. 1033-42.
175. Deed, R.W., S. Armitage, and J.D. Norton, *Nuclear localization and regulation of Id protein through an E protein-mediated chaperone mechanism*. *J Biol Chem*, 1996. **271**(39): p. 23603-6.
176. Alani, R.M., et al., *Immortalization of primary human keratinocytes by the helix-loop-helix protein, Id-1*. *Proc Natl Acad Sci U S A*, 1999. **96**(17): p. 9637-41.
177. Nickoloff, B.J., et al., *Id-1 delays senescence but does not immortalize keratinocytes*. *J Biol Chem*, 2000. **275**(36): p. 27501-4.
178. Steininger, A., et al., *Genomic loss of the putative tumor suppressor gene E2A in human lymphoma*. *J Exp Med*, 2011. **208**(8): p. 1585-93.
179. Mathas, S., et al., *Gene deregulation and spatial genome reorganization near breakpoints prior to formation of translocations in anaplastic large cell lymphoma*. *Proc Natl Acad Sci U S A*, 2009. **106**(14): p. 5831-6.
180. O'Neil, J., et al., *TAL1/SCL induces leukemia by inhibiting the transcriptional activity of E47/HEB*. *Cancer Cell*, 2004. **5**(6): p. 587-96.

181. Park, S.T., G.P. Nolan, and X.H. Sun, *Growth inhibition and apoptosis due to restoration of E2A activity in T cell acute lymphoblastic leukemia cells*. J Exp Med, 1999. **189**(3): p. 501-8.
182. Sanda, T., et al., *Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia*. Cancer Cell, 2012. **22**(2): p. 209-21.
183. Tan, S.H., et al., *TRIB2 reinforces the oncogenic transcriptional program controlled by the TAL1 complex in T-cell acute lymphoblastic leukemia*. Leukemia, 2016. **30**(4): p. 959-62.
184. Spaulding, C., et al., *Notch1 co-opts lymphoid enhancer factor 1 for survival of murine T-cell lymphomas*. Blood, 2007. **110**(7): p. 2650-8.
185. Wang, H.C., et al., *Enhanced Notch activation is advantageous but not essential for T cell lymphomagenesis in Id1 transgenic mice*. PLoS One, 2012. **7**(2): p. e32944.
186. Grabher, C., H. von Boehmer, and A.T. Look, *Notch 1 activation in the molecular pathogenesis of T-cell acute lymphoblastic leukaemia*. Nat Rev Cancer, 2006. **6**(5): p. 347-59.
187. Reschly, E.J., et al., *Notch1 promotes survival of E2A-deficient T cell lymphomas through pre-T cell receptor-dependent and -independent mechanisms*. Blood, 2006. **107**(10): p. 4115-21.
188. Talora, C., et al., *Pre-TCR-triggered ERK signalling-dependent downregulation of E2A activity in Notch3-induced T-cell lymphoma*. EMBO Rep, 2003. **4**(11): p. 1067-72.
189. Cotta, C.V., et al., *The helix-loop-helix protein Id2 is expressed differentially and induced by myc in T-cell lymphomas*. Cancer, 2008. **112**(3): p. 552-61.
190. Li, J., et al., *Id2 Collaborates with Id3 To Suppress Invariant NKT and Innate-like Tumors*. J Immunol, 2017. **198**(8): p. 3136-3148.
191. Yu, L., et al., *Global assessment of promoter methylation in a mouse model of cancer identifies ID4 as a putative tumor-suppressor gene in human leukemia*. Nat Genet, 2005. **37**(3): p. 265-74.
192. Chen, S.S., et al., *Silencing of the inhibitor of DNA binding protein 4 (ID4) contributes to the pathogenesis of mouse and human CLL*. Blood, 2011. **117**(3): p. 862-71.

193. Cen, J., et al., *Association between lymphoma prognosis and aberrant methylation of ID4 and ZO-1 in bone marrow and paraffin-embedded lymphoma tissues of treatment-naive patients*. *Oncol Rep*, 2013. **30**(1): p. 455-61.
194. Hagiwara, K., et al., *Frequent DNA methylation but not mutation of the ID4 gene in malignant lymphoma*. *J Clin Exp Hematop*, 2007. **47**(1): p. 15-8.
195. Belle, I., et al., *Aberrant production of IL-13 by T cells promotes exocrinopathy in Id3 knockout mice*. *Cytokine*, 2014. **69**(2): p. 226-33.
196. Fehling, H.J., et al., *Crucial role of the pre-T-cell receptor alpha gene in development of alpha beta but not gamma delta T cells*. *Nature*, 1995. **375**(6534): p. 795-8.
197. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome*. *Genome Biol*, 2009. **10**(3): p. R25.
198. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS)*. *Genome Biol*, 2008. **9**(9): p. R137.
199. Boeva, V., et al., *Nebula--a web-server for advanced ChIP-seq data analysis*. *Bioinformatics*, 2012. **28**(19): p. 2517-9.
200. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration*. *Brief Bioinform*, 2013. **14**(2): p. 178-92.
201. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities*. *Mol Cell*, 2010. **38**(4): p. 576-89.
202. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data*. *Bioinformatics*, 2014. **30**(15): p. 2114-20.
203. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks*. *Nat Protoc*, 2012. **7**(3): p. 562-78.
204. Team, R.C., *R: A language and environment for statistical computing*. 2013.
205. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles*. *Proc Natl Acad Sci U S A*, 2005. **102**(43): p. 15545-50.

206. Li, S., et al., *IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling*. *Nat Commun*, 2013. **4**: p. 2333.
207. Warde-Farley, D., et al., *The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function*. *Nucleic Acids Res*, 2010. **38**(Web Server issue): p. W214-20.
208. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks*. *Genome Res*, 2003. **13**(11): p. 2498-504.
209. Mootha, V.K., et al., *PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes*. *Nat Genet*, 2003. **34**(3): p. 267-73.
210. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies*. *Nucleic Acids Res*, 2015. **43**(7): p. e47.
211. Heng, T.S., M.W. Painter, and C. Immunological Genome Project, *The Immunological Genome Project: networks of gene expression in immune cells*. *Nat Immunol*, 2008. **9**(10): p. 1091-4.
212. Glaab, E., J.M. Garibaldi, and N. Krasnogor, *ArrayMining: a modular web-application for microarray analysis combining ensemble and consensus methods with cross-study normalization*. *BMC Bioinformatics*, 2009. **10**: p. 358.
213. Micallef, L. and P. Rodgers, *eulerAPE: drawing area-proportional 3-Venn diagrams using ellipses*. *PLoS One*, 2014. **9**(7): p. e101717.
214. Lauritsen, J.P., et al., *Marked induction of the helix-loop-helix protein Id3 promotes the gammadelta T cell fate and renders their functional maturation Notch independent*. *Immunity*, 2009. **31**(4): p. 565-75.
215. Kang, J. and N. Malhotra, *Transcription factor networks directing the development, function, and evolution of innate lymphoid effectors*. *Annu Rev Immunol*, 2015. **33**: p. 505-38.
216. Berga-Bolanos, R., et al., *Cell-autonomous requirement for TCF1 and LEF1 in the development of Natural Killer T cells*. *Mol Immunol*, 2015. **68**(2 Pt B): p. 484-9.
217. Egawa, T., *Regulation of CD4 and CD8 coreceptor expression and CD4 versus CD8 lineage decisions*. *Adv Immunol*, 2015. **125**: p. 1-40.

218. Carr, T., et al., *The transcription factor lymphoid enhancer factor 1 controls invariant natural killer T cell expansion and Th2-type effector differentiation*. J Exp Med, 2015. **212**(5): p. 793-807.
219. Sonoda, K.H., et al., *CD1-reactive natural killer T cells are required for development of systemic tolerance through an immune-privileged site*. J Exp Med, 1999. **190**(9): p. 1215-26.
220. Bendelac, A., R.D. Hunziker, and O. Lantz, *Increased interleukin 4 and immunoglobulin E production in transgenic mice overexpressing NK1 T cells*. J Exp Med, 1996. **184**(4): p. 1285-93.
221. Winoto, A. and D. Baltimore, *A novel, inducible and T cell-specific enhancer located at the 3' end of the T cell receptor alpha locus*. EMBO J, 1989. **8**(3): p. 729-33.
222. Boudil, A., et al., *IL-7 coordinates proliferation, differentiation and Tcr α recombination during thymocyte beta-selection*. Nat Immunol, 2015. **16**(4): p. 397-405.
223. Eberl, G., et al., *Absolute requirement for the pre-T cell receptor alpha chain during NK1.1+ TCR α cell development*. Eur J Immunol, 1999. **29**(6): p. 1966-71.
224. Shinkai, Y. and F.W. Alt, *CD3 epsilon-mediated signals rescue the development of CD4+CD8+ thymocytes in RAG-2 $^{-/-}$ mice in the absence of TCR beta chain expression*. Int Immunol, 1994. **6**(7): p. 995-1001.
225. Liberzon, A., et al., *The Molecular Signatures Database (MSigDB) hallmark gene set collection*. Cell Syst, 2015. **1**(6): p. 417-425.
226. Lee, Y.J., et al., *Lineage-Specific Effector Signatures of Invariant NKT Cells Are Shared amongst $\gamma\delta$ T, Innate Lymphoid, and Th Cells*. J Immunol, 2016. **197**(4): p. 1460-70.
227. Bezman, N.A., et al., *Molecular definition of the identity and activation of natural killer cells*. Nat Immunol, 2012. **13**(10): p. 1000-9.
228. Pereira, P. and L. Boucontet, *Innate NKT $\gamma\delta$ and NKT $\alpha\beta$ cells exert similar functions and compete for a thymic niche*. Eur J Immunol, 2012. **42**(5): p. 1272-81.
229. Sandrock, I., et al., *MicroRNA-181a/b-1 Is Not Required for Innate $\gamma\delta$ NKT Effector Cell Development*. PLoS One, 2015. **10**(12): p. e0145010.

230. Dashtsoodol, N., et al., *Alternative pathway for the development of V α 14⁺ NKT cells directly from CD4⁻CD8⁻ thymocytes that bypasses the CD4⁺CD8⁺ stage*. Nat Immunol, 2017.
231. Dutta, M., et al., *A role for Ly108 in the induction of promyelocytic zinc finger transcription factor in developing thymocytes*. J Immunol, 2013. **190**(5): p. 2121-8.
232. Gapin, L., *Development of invariant natural killer T cells*. Curr Opin Immunol, 2016. **39**: p. 68-74.
233. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. Cell, 2011. **144**(5): p. 646-74.
234. Hainaut, P. and A. Plymoth, *Targeting the hallmarks of cancer: towards a rational approach to next-generation cancer therapy*. Curr Opin Oncol, 2013. **25**(1): p. 50-1.
235. Ellenbroek, S.I. and J. van Rheenen, *Imaging hallmarks of cancer in living mice*. Nat Rev Cancer, 2014. **14**(6): p. 406-18.
236. Evan, G.I. and K.H. Vousden, *Proliferation, cell cycle and apoptosis in cancer*. Nature, 2001. **411**(6835): p. 342-8.
237. DeBerardinis, R.J., et al., *The biology of cancer: metabolic reprogramming fuels cell growth and proliferation*. Cell Metab, 2008. **7**(1): p. 11-20.
238. Russell, R.G., et al., *Id2 drives differentiation and suppresses tumor formation in the intestinal epithelium*. Cancer Res, 2004. **64**(20): p. 7220-5.
239. Ohshima, K., et al., *Classification of cell lineage and anatomical site, and prognosis of extranodal T-cell lymphoma -- natural killer cell, cytotoxic T lymphocyte, and non-NK/CTL types*. Virchows Arch, 2002. **440**(4): p. 425-35.
240. Matsuo, Y., et al., *Megakaryoblastic leukemia cell line MOLM-16 derived from minimally differentiated acute leukemia with myeloid/NK precursor phenotype*. Leuk Res, 2003. **27**(2): p. 165-71.
241. Carson, W.E., et al., *Interleukin (IL) 15 is a novel cytokine that activates human natural killer cells via components of the IL-2 receptor*. J Exp Med, 1994. **180**(4): p. 1395-403.
242. Caligiuri, M.A., *Human natural killer cells*. Blood, 2008. **112**(3): p. 461-9.

243. Gordy, L.E., et al., *IL-15 regulates homeostasis and terminal maturation of NKT cells*. J Immunol, 2011. **187**(12): p. 6335-45.
244. Matsuda, J.L., et al., *Homeostasis of V alpha 14i NKT cells*. Nat Immunol, 2002. **3**(10): p. 966-74.
245. Yu, J., et al., *NKp46 identifies an NKT cell subset susceptible to leukemic transformation in mouse and human*. J Clin Invest, 2011. **121**(4): p. 1456-70.
246. Matsuo, Y. and H.G. Drexler, *Immunoprofiling of cell lines derived from natural killer-cell and natural killer-like T-cell leukemia-lymphoma*. Leuk Res, 2003. **27**(10): p. 935-45.
247. Godfrey, D.I., et al., *NKT cells: what's in a name?* Nat Rev Immunol, 2004. **4**(3): p. 231-7.
248. El-Guendy, N. and V.M. Rangnekar, *Apoptosis by Par-4 in cancer and neurodegenerative diseases*. Exp Cell Res, 2003. **283**(1): p. 51-66.
249. Perillo, N.L., et al., *Apoptosis of T cells mediated by galectin-1*. Nature, 1995. **378**(6558): p. 736-9.
250. Marigo, I., et al., *Tumor-induced tolerance and immune suppression depend on the C/EBPbeta transcription factor*. Immunity, 2010. **32**(6): p. 790-802.
251. Yanai, H., et al., *Role of IFN regulatory factor 5 transcription factor in antiviral immunity and tumor suppression*. Proc Natl Acad Sci U S A, 2007. **104**(9): p. 3402-7.
252. Shen, S., et al., *Tight regulation of diacylglycerol-mediated signaling is critical for proper invariant NKT cell development*. J Immunol, 2011. **187**(5): p. 2122-9.
253. Dominguez, C.L., et al., *Diacylglycerol kinase alpha is a critical signaling node and novel therapeutic target in glioblastoma and other cancers*. Cancer Discov, 2013. **3**(7): p. 782-97.
254. Ohki, R., et al., *Reprimo, a new candidate mediator of the p53-mediated cell cycle arrest at the G2 phase*. J Biol Chem, 2000. **275**(30): p. 22627-30.
255. Liu, X., et al., *PTPN14 interacts with and negatively regulates the oncogenic function of YAP*. Oncogene, 2013. **32**(10): p. 1266-73.
256. Winkler, G.S., *The mammalian anti-proliferative BTG/Tob protein family*. J Cell Physiol, 2010. **222**(1): p. 66-72.

257. Zou, Z., et al., *PKD2 and PKD3 promote prostate cancer cell invasion by modulating NF-kappaB- and HDAC1-mediated expression and activation of uPA*. J Cell Sci, 2012. **125**(Pt 20): p. 4800-11.
258. Pellikainen, J.M., et al., *Expression of matrix metalloproteinase (MMP)-2 and MMP-9 in breast cancer with a special reference to activator protein-2, HER2, and prognosis*. Clin Cancer Res, 2004. **10**(22): p. 7621-8.
259. Oehler, M.K., et al., *Adrenomedullin inhibits hypoxic cell death by upregulation of Bcl-2 in endometrial cancer cells: a possible promotion mechanism for tumour growth*. Oncogene, 2001. **20**(23): p. 2937-45.
260. Banks, R.E., et al., *Circulating intercellular adhesion molecule-1 (ICAM-1), E-selectin and vascular cell adhesion molecule-1 (VCAM-1) in human malignancies*. Br J Cancer, 1993. **68**(1): p. 122-4.
261. Johnston, B., et al., *Differential chemokine responses and homing patterns of murine TCR alpha beta NKT cell subsets*. J Immunol, 2003. **171**(6): p. 2960-9.
262. Elewaut, D., et al., *NIK-dependent RelB activation defines a unique signaling pathway for the development of V alpha 14i NKT cells*. J Exp Med, 2003. **197**(12): p. 1623-33.
263. Lamprecht, B., et al., *Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma*. Nat Med, 2010. **16**(5): p. 571-9, 1p following 579.
264. Koizumi, K., et al., *Chemokine receptors in cancer metastasis and cancer cell-derived chemokines in host immune response*. Cancer Sci, 2007. **98**(11): p. 1652-8.
265. Zlotnik, A., A.M. Burkhardt, and B. Homey, *Homeostatic chemokine receptors and organ-specific metastasis*. Nat Rev Immunol, 2011. **11**(9): p. 597-606.
266. Franitza, S., et al., *Differential usage of VLA-4 and CXCR4 by CD3+CD56+ NKT cells and CD56+CD16+ NK cells regulates their interaction with endothelial cells*. Eur J Immunol, 2004. **34**(5): p. 1333-41.
267. Emoto, M., et al., *Critical role of leukocyte function-associated antigen-1 in liver accumulation of CD4+NKT cells*. J Immunol, 1999. **162**(9): p. 5094-8.
268. Teicher, B.A. and S.P. Fricker, *CXCL12 (SDF-1)/CXCR4 pathway in cancer*. Clin Cancer Res, 2010. **16**(11): p. 2927-31.

269. Imai, H., et al., *Clinicopathological and therapeutic significance of CXCL12 expression in lung cancer*. *Int J Immunopathol Pharmacol*, 2010. **23**(1): p. 153-64.
270. Franki, A.S., et al., *A unique lymphotoxin {alpha}beta-dependent pathway regulates thymic emigration of V{alpha}14 invariant natural killer T cells*. *Proc Natl Acad Sci U S A*, 2006. **103**(24): p. 9160-5.
271. Blyth, K., et al., *Runx2 and MYC collaborate in lymphoma development by suppressing apoptotic and growth arrest pathways in vivo*. *Cancer Res*, 2006. **66**(4): p. 2195-201.
272. Egeblad, M. and Z. Werb, *New functions for the matrix metalloproteinases in cancer progression*. *Nat Rev Cancer*, 2002. **2**(3): p. 161-74.
273. Sternlicht, M.D., et al., *The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis*. *Cell*, 1999. **98**(2): p. 137-46.
274. Bouchard, F., et al., *EGR-1 activation by EGF inhibits MMP-9 expression and lymphoma growth*. *Blood*, 2010. **116**(5): p. 759-66.
275. Niola, F., et al., *Id proteins synchronize stemness and anchorage to the niche of neural stem cells*. *Nat Cell Biol*, 2012. **14**(5): p. 477-87.
276. Starlets, D., et al., *Cell-surface CD74 initiates a signaling cascade leading to cell proliferation and survival*. *Blood*, 2006. **107**(12): p. 4807-16.
277. Wang, W., et al., *The novel tumor-suppressor Mel-18 in prostate cancer: its functional polymorphism, expression and clinical significance*. *Int J Cancer*, 2009. **125**(12): p. 2836-43.
278. Cowan, J.E., et al., *Differential requirement for CCR4 and CCR7 during the development of innate and adaptive alphabetaT cells in the adult thymus*. *J Immunol*, 2014. **193**(3): p. 1204-12.
279. Bachy, E., et al., *CD1d-restricted peripheral T cell lymphoma in mice and humans*. *J Exp Med*, 2016. **213**(5): p. 841-57.
280. Ogata, H., et al., *KEGG: Kyoto Encyclopedia of Genes and Genomes*. *Nucleic Acids Res*, 1999. **27**(1): p. 29-34.
281. Jiang, L., et al., *Exome sequencing identifies somatic mutations of DDX3X in natural killer/T-cell lymphoma*. *Nat Genet*, 2015. **47**(9): p. 1061-6.

282. Prince, A.L., et al., *Innate PLZF+CD4+ alphabeta T cells develop and expand in the absence of Itk*. J Immunol, 2014. **193**(2): p. 673-87.
283. McGregor, S., et al., *PLZF staining identifies peripheral T-cell lymphomas derived from innate-like T-cells with TRAV1-2-TRAJ33 TCR-alpha rearrangement*. Blood, 2014. **123**(17): p. 2742-3.
284. Edholm, E.S., M. Banach, and J. Robert, *Evolution of innate-like T cells and their selection by MHC class I-like molecules*. Immunogenetics, 2016. **68**(8): p. 525-36.
285. Jojic, V., *Identification of transcriptional regulators in the mouse immune system*. 2013. **14**(6): p. 633-43.
286. Eyquem, S., et al., *The Ets-1 transcription factor is required for complete pre-T cell receptor function and allelic exclusion at the T cell receptor beta locus*. Proc Natl Acad Sci U S A, 2004. **101**(44): p. 15712-7.
287. Walunas, T.L., et al., *Cutting edge: the Ets1 transcription factor is required for the development of NK T cells in mice*. J Immunol, 2000. **164**(6): p. 2857-60.
288. Balmelle, N., et al., *Developmental activation of the TCR alpha enhancer requires functional collaboration among proteins bound inside and outside the core enhancer*. J Immunol, 2004. **173**(8): p. 5054-63.
289. Park, J.H., et al., *Signaling by intrathymic cytokines, not T cell antigen receptors, specifies CD8 lineage choice and promotes the differentiation of cytotoxic-lineage T cells*. Nat Immunol, 2010. **11**(3): p. 257-64.

Biography

Sumedha Roy was born on September 18th, 1989 in Calcutta, India. She completed her Bachelors in Technology (B. Tech.) in 2011 with a major in Biochemical Engineering, and her Masters in Technology (M. Tech.) in 2012 with a focus on Bioinformatics, from Indian Institute of Technology (IIT), Delhi, India. She started her graduate studies in Immunology at Duke University as a Chancellor's scholar in August 2012, and successfully defended her Doctor of Philosophy (Ph. D.) thesis in March 2018.

Publications

Li J*, **Roy S***, Kim Y-M, Zhang B, Love C, Reddy A, Rajagopalan D, Dave S, Li S, Dave S, Diehl A M, and Zhuang Y. Id2 Collaborates with Id3 To Suppress Invariant iNKT and Innate-like Tumors, *J. Immunol.*, 2017, 198(8):3136–48. *- Co-first author

Roy S, Moore A, Love C, Reddy A, Rajagopalan D, Dave S, Li L, Murre C, and Zhuang Y. Id proteins suppress E2A-driven iNKT cell development prior to TCR selection, *Front. Immunol.*, 2018, doi: 10.3389/fimmu.2018.00042

Roy S and Zhuang Y, Orchestration of Invariant Natural Killer T cell development by E proteins, *Crit. Rev. Immunol.*, 2015, 35(1):3348.

Roy S and Zhuang Y, Paradoxical role of Id proteins in regulating tumorigenic potential of lymphoid cells, *Front. Med.*, 2018 (Submitted).

Roy S, Dutta S, Khanna K, Singla S, Sundar D., Prediction of DNA-binding specificity in zinc finger proteins, *J. Bioscience*, 2012, 37(3):483491