Mechanisms of naive T cell heterogeneity

Dakota Rogers

Department of Physiology
McGill University

Tiohtià:ke, unceded land of the Kanien'kehá:ka Nation

Montreal, Quebec, Canada

December 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

LAND ACKNOWLEDGEMENT

This work was done in Tiohtià:ke/Montreal on the traditional territory of the Kanien'kehá:ka, a place which has long served as a site of meeting and exchange amongst many First Nations including the Kanien'kehá:ka of the Haudenosaunee Confederacy, Huron/Wendat, Abenaki, and Anishinaabeg. We honour, recognize, respect, and thank these nations and the diverse Indigenous peoples whose presence marks this territory as the traditional stewards of these lands and waters.

TABLE OF CONTENTS

ABSTRACT

The T cell receptor (TCR) is essential for T cells to recognize foreign antigens from the diverse world of pathogens. While the TCR evolved to selectively recognize peptide presented by major histocompatibility complexes (MHC) on the cell surface, it must also be tested for being able to discriminate self from non-self. The education of T cells occurs during their maturation in the thymus where they interact with self-peptide MHC (self-pMHC). Importantly, while overtly self-reactive T cells are deleted, T cells must also obtain sub-threshold self-pMHC signals to be selected into the mature repertoire at all. This establishes a self-reactivity spectrum among naive T cells. Once naive T cells enter the periphery, they continue to obtain sub-threshold survival signals from self-ligand interactions. However, whether T cells across the self-reactivity spectrum are wired differently at the transcriptional level, whether this impacts their function post-activation, to what extent the self-pMHC signals received in the thymus or periphery contribute to T cell heterogeneity, and why they are necessary for naive T cell survival, remained unknown.

Here, we observed that naive T cells occupying opposing ends of the self-reactivity spectrum possessed distinct chromatin landscapes and transcriptional networks. The heterogeneity of naive T cells prior to their encounter with cognate antigen imparted biases in early signaling post-activation, regulatory T cell differentiation, and impacted follicular helper T (T_{FH}) cell versus non-T_{FH} effector lineage choice. Moreover, by depriving naive $CD4^+$ T cells of tonic self-interactions we identified gene expression and chromatin accessibility differences that did not rely on continuous peripheral self-ligand interactions and are thus likely thymically imprinted during development. Interestingly, one of the highest differentially expressed genes between naive $CD4^+$ T cells with low versus high self-reactivity was terminal deoxynucleotidyl transferase (TdT), in line with a prior hypothesis that longer TCRs may be overrepresented in T cells with lower self-reactivity. Strikingly, we observed that during chronic infection, low self-reactive T cells predominated in the immune response, and that in $TdT^{-/-}$ mice chronic viral control was delayed while acute viral infection was cleared normally. Finally, we investigated why continuous self-pMHC interactions are essential for naive T cell maintenance and survival. We identified gene circuits involved in autophagy, lipid metabolism, response to starvation, and catabolic metabolism, as being upregulated upon self-ligand deprivation. Interestingly, we found that naive T cells underwent transient self-deprivation during viral infection, resulting in impaired function following *in vitro* stimulation.

Together, our work sheds new light on how the strength of self-reactivity of naive T cells regulates effector responses during infection, impacts T cell heterogeneity pre- and post-activation, and maintains T cell function. This has important implications for understanding which gene circuits may be amenable to therapeutics to modify T cell function, while others are epigenetically imprinted. Overall, we highlight the importance of the co-evolutionary relationship between a T cell and its pMHC ligands within an individual.

RÉSUMÉ

Le récepteur des lymphocytes T (TCR) est essentiel pour reconnaître les antigènes venant du monde très diversifié des agents pathogènes. Alors que le TCR a évolué pour reconnaître sélectivement les pathogènes sous la forme de peptides présentés par des complexes majeurs d'histocompatibilité (pCMH), il doit aussi pouvoir pour discriminer le soi du non-soi. L'éducation des cellules T se produit pendant leur maturation dans le thymus par leurs interactions avec l'ensemble des pCMH du soi. Bien que les cellules T les plus auto-réactives soient supprimées, les autres doivent tout de même obtenir un signal suffisant des pCMH afin d'être sélectionnées dans le répertoire mature, ce qui établit un spectre de réactivité au soi parmi les cellules T naïves. Une fois en périphérie, elles continuent d'obtenir ces signaux qui leur permettent de survivre. Cependant, on ne sait pas si les cellules T sont transcriptionnellement identiques les unes aux autres au sein du spectre d'auto-réactivité et si cela a un impact sur leur fonction après leur activation. On ne sait pas non plus dans quelle mesure les signaux reçus par les pCMH du soi dans le thymus ou dans la périphérie contribuent à l'hétérogénéité des cellules T, ni pourquoi ces signaux sont nécessaires à la survie des lymphocytes T naïfs.

Nous avons constaté que les lymphocytes T naïfs situés aux extrémités du spectre d'auto-réactivité possèdent des paysages chromatiniens et des réseaux transcriptionnels distincts. Cette hétérogénéité pré-activation des lymphocytes T naïfs entraîne des biais de signalisation précoce lors de leur activation, ce qui résulte en des disparités dans leur potentiel de différenciation en lymphocytes T régulateurs (T_{reg}) ou T auxiliaires folliculaires (T_{FH}). En privant les cellules T $CD4^+$ naïves d'interactions toniques, nous avons identifié les différences d'expression génétiques et d'accessibilité à la chromatine ne reposant pas sur les interactions continues avec les peptides du soi périphériques et donc probablement encodées dans le thymus au cours du développement. De façon intéressante, l'un des gènes les plus différentiellement exprimé parmi les lymphocytes T $CD4^+$ naïfs peu auto-réactifs est la désoxynucléotidyl transférase terminale (TdT), en ligne avec une hypothèse selon laquelle les TCR les plus longs sont surreprésentés parmi les cellules T peu auto-réactives. De manière surprenante, nous avons observé que lors d'une infection chronique, les lymphocytes T peu auto-réactifs prédominent dans la réponse immunitaire, et que la perte de TdT induit un retard dans résolution cette infection, un effet absent lors de la résolution d'une infection virale aiguë. Enfin, nous avons étudié pourquoi les interactions toniques continues avec les pCMH du soi sont essentielles dans le maintien et la survie des cellules T naïves. Nous avons identifié

l’induction de circuits génétiques impliqués dans l'autophagie, le métabolisme lipidique, et la réponse au manque de nutriments lorsque les signaux obtenus par les pCMH du soi ont été supprimés. De façon intéressante, nous avons constaté que les cellules T naïves sont transitoirement privées des signaux venant du soi lors d’une infection virale, ce qui inhibe leur fonction après leur stimulation *in vitro*.

Ensemble, nos travaux apportent une nouvelle vision sur le rôle de l'auto-réactivité des lymphocytes T naïfs dans les réponses effectrices aux infections, leur hétérogénéité avant et après leur activation, et le maintien de leurs fonctions. Nos découvertes ont des implications importantes pour comprendre les circuits génétiques et épigénétiques susceptibles d'être ciblés thérapeutiquement afin d'altérer la fonction des cellules T. Somme toute, nous montrons l'importance de la relation co-évolutive entre une cellule T et le pCMH de l’individu.

ACKNOWLEDGMENTS

Ten years ago, if you were to ask this small-town rural Nova Scotian kid where he would be, 'finishing a PhD' is not an answer I would have given you. When I was figuring out what to do with my life after undergrad, I knew exactly what I was not going to do. I was not going to do a graduate degree, I was not going to go to McGill, and I was not going move to Quebec. I don't think I quite understand the meaning of 'not'.

To Judith, words cannot express how thankful I am for your guidance over the years. I remember when I was (not) looking for a graduate supervisor, my then-supervisor had warned me that starting in a lab with a new PI was a risk. I am so glad I took that risk. My growth as a scientist and as a person over the years is in large part thanks to you. It has been an absolute pleasure to be part of the first generation of students in the lab and to watch the lab grow into the amazing collection of weirdos that we are today!

To my lab mates: Caitlin, Connie, Jérémy, Aanya, Angela, Maryl, and Dhanesh as well as our honorary lab members Sam and Vincent, they say that drinking is an essential part of grad school. Boy, were they right! But they also say that drinking with good company is one of life's greatest pleasures. Thank you for being the best company.

To my friends, old and new, you are who kept me sane throughout all these years. Special shout-out to Corbin who has been one of the constants throughout my degree, thank you for all that you have done.

To Julie Haughn, it is said that a great teacher can change your life. You were mine. Thank you for lighting my passion to learn.

To my cats, Gluten-Free and Kat von D, your companionship over the years has saved me more than once. Buckle up girls, I hope you're ready for the next journey in our lives.

To my parents, thank you for everything you have done for me! I am who I am today because of you. I am where I am today because of you. I will never forget that.

As a first-generation student, academia can be extremely intimidating and imposter syndrome can run deep (some days deeper than others). While the past ~~7 years~~ have been some of my most challenging, I do not regret a single decision. I may be the first person in my family to attend university and the first to complete a PhD, but I certainly hope I will not be the last.

This thesis is dedicated to my nephews, Liam and Kamden, and my niece, Braxlyn. I hope someday, when your time comes, you get to pursue your passions.

RESEARCH PUBLICATIONS

This is a manuscript-based thesis. The work enclosed has or will be published as detailed below. Unpublished chapters may change based on experiments in progress and revisions for publication.

Chapters 1 and 5

Part of these chapters are adapted from a review article on the self-immunopeptidome and its impact on T cell development and homeostasis.

This S.*, **Rogers D.***, Mallet Gauthier È., Mandl J.N.[#], Melichar H.J.[#] (2023). What's self got to do with it: Sources of heterogeneity among naive T cells. Seminars in Immunology. 65: 101702.

* Equal contribution

[#] Equal contribution

Chapter 2

Rogers D., Sood A., Wang H., van Beek J.P., Rademaker T.J., Artusa P., Schneider C., Shen C., Wong D. C, Lebel M., Condotta S.A., Richer M.J., Martins A.J., Tsang J.S., Barreiro L., Francois P., Langlais D., Melichar H.J., Textor J., Mandl J.N. (2021). Pre-existing chromatin accessibility and gene expression differences among naive $CD4^+$ T cells influence effector potential. Cell Reports. 37(9): 110064.

Chapter 3

The data presented in this chapter are included in two unpublished manuscripts or as additional unpublished data. A pre-print of both manuscripts can be found on BioRxiv.

Jamaleddine S.*, **Rogers D.***, Perreault G., Mandl J.N.[#], Khadra A[#]. (2022). Chronic infection control relies on T cells with lower foreign antigen binding strength generated by N-nucleotide diversity. Biorxiv. https://doi.org/10.1101/2022.06.26.497644. *Under Review.*

Textor J., Buytenhuijs F., **Rogers D.**, Mallet Gauthier É., Sultan S., Wortel I.M.N., Kalies K., Fähnrich A., Pagel R., Melichar H.J., Westermann J., Mandl J.N. (2022). Machine learning analysis of the T cell receptor repertoire identifies sequence features that predict self-reactivity. BioRxiv. https://doi.org/10.1101/2022.11.23.517563. *Under Review.*

* Equal contribution

[#] Equal contribution

Chapter 4

Rogers D., Mandl J.N. Naive $CD4^+$ T cell changes at the transcriptional level when deprived of self may have consequences during infection. *Manuscript in preparation*.

Other research contributions

I have also contributed to the following publications that are not included in this thesis:

Schneider C., Shen C., Gopal A.A., Douglas T., Forestell B., Kauffman K.D., **Rogers D.**, Artusa P., Zhang Q., Jing H., Freeman A.F., Barber D.L., King I.L., Saleh M., Wiseman P.W., Su H.C., Mandl J.N. (2020). Migration-induced cell shattering due to DOCK8 deficiency causes a type 2-biased helper T cell response. *Nature Immunology* 21(12): 1528-1539.

Sood A., Lebel M., Fournier M., **Rogers D.**, Mandl J.N., Melichar H.J. (2019). Differential interferon-gamma production potential among naïve $CD4^+$ T cells exists prior to antigen encounter. *Immunology & Cell Biology* 97(10): 931-940.

CONTRIBUTION OF AUTHORS

The work in this thesis was completed under the supervision of Dr. Judith N. Mandl.

Chapter 2

The project idea behind the work in Chapter 2 was conceived by Dr. Mandl. Dr. Mandl and I designed the research. I performed most of the experiments with help from Patricio Artusa, Caitlin Schneider, Connie Shen, Dylan Wong, Aanya Bhagrath, and Dr. Marie-Éve Lebel. The RNA-seq dataset was generated by Dr. Mandl, Dr. Andrew Martins and Dr. John Tsang. The ATAC-seq dataset was generated by myself and Dr. Mandl. I performed the analysis of the RNA-seq and ATAC-seq datasets with help from HanChen Wang, Dr. Aditi Sood, Dr. Luis Barreiro, and Dr. David Langlais. The single-cell RNA-seq was performed and analyzed by Drs. Jasper van Beek and Johannes Textor. Data were analyzed by myself, Dr. Mandl, Dr. Thomas Rademaker, Dr. Paul François, and Dr. Johannes Textor. Critical reagents or intellectual input were provided by Dr. Stephanie Condotta, Dr. Martin Richer, and Heather Melichar. The manuscript was written by myself and Dr. Mandl.

Chapter 3

The project idea for Chapter 3 was conceived by Dr. Mandl. Dr. Mandl and I designed the research. I performed the experiments with help from Hassan Jamaleddine, Geneviève Perreault, and Dhanesh Patel. The ATAC-seq dataset was generated by me and Dr. Mandl. I performed the analysis of the ATAC-seq dataset. Cell sorts were performed by either me or Dr. Mandl. Bone marrow chimeras were made by me. Data were analyzed by me and Dr. Mandl. The manuscript was written by me with feedback from Dr. Mandl.

Chapter 4

The project idea for Chapter 4 was conceived by myself and Dr. Mandl. The RNA-seq data was generated as described in Chapter 2. I performed all experiments and wrote the manuscript with feedback from Dr. Mandl.

CONTRIBUTION TO ORIGINAL KNOWLEDGE

The work presented here contributes substantial original findings to our understanding of pre-existing naive T cell heterogeneity, effector lineage decisions, and the impact of self-pMHC signals on naive T cells.

Chapter 2

Prior to the publication of this study, it was becoming increasingly appreciated that differences in self-reactivity impart biases in naive $CD4^+$ T cell effector choices post-activation. We demonstrated that even prior to activation the naive $CD4^+$ T cell population possesses remarkably diverse chromatin landscapes and transcriptional networks that were maintained following activation and predicted effector lineage differentiation. Additionally, we showed that a large portion of the pre-existing heterogeneity does not rely on continuous self-ligand interactions and is likely set in the thymus during development.

Chapter 3

Naive $CD4^+$ T cells with high self-pMHC reactivity dominate the T cell response during acute infection. However, the role that low self-reactive naive $CD4^+$ T cells play during infection has not been studied. We found that naive $CD4^+$ T cells with low-pMHC reactivity dominate the T cell response during chronic infection. We further show that T cells that lack TdT expression produce TCRs that may have, on average, higher self-pMHC reactivity and take longer to control chronic viral infection.

Chapter 4

Naive T cells require continuous self-pMHC interactions in the periphery for survival and maintenance. Yet, naive T cells can survive for long periods upon deprivation of self-pMHC signals. How naive T cells can survive for so long without self-pMHC was unknown. We show that deprivation of self-signals upregulates transcriptional networks that support survival like autophagy, catabolic metabolism, lipid metabolism and response to starvation in naive $CD4^+$ T cells. We also demonstrate that pathogen non-responding naive T cells experience self-deprivation during infection that is dependent on type-I interferon signaling and impairs response to subsequent activation.

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF ABBREVIATIONS

AIRE – Autoimmune regulator

APC – Antigen presenting cell

Arm - Armstrong

ATAC – Assay for transposase-accessible chromatin

BCR – B cell receptor

CD – Cluster of differentiation

CDR – Complementarity-determining region

CFU – Colony forming units

Cl13 – Clone 13

cTEC – Cortical thymic epithelial cell

DAG - Diacylglycerol

DAR – Differentially accessible region

DC – Dendritic cell

DEG – Differentially expressed gene

DN – Double negative

DP – Double positive

FRC – Fibroblast reticular cells

GO – Gene ontology

GSEA – Gene set enrichment analysis

IFN – Interferon

IISD – Infection-induced self-deprivation

IL - Interleukin

IP3 – Inositol triphosphate

ITAM – Immunoreceptor tyrosine-based activated motifs

KO – Knockout

LAT – Linker for activation of T cells

LCMV – Lymphocytic choriomeningitis virus

LIP – Lymphopenia-induced proliferation

MAPK – Mitogen activated protein kinase

MCMV – Murine cytomegalovirus

MFI – Mean fluorescence intensity

MHC – Major histocompatibility complex

mTEC – Medullary thymic epithelial cell

NK – Natural killer

PCA – Principal component analysis

PFU – Plaque forming units

PIP2 – phosphatidylinositol biphosphatePLCγ1 – Phospholipase Cγ1

pMHC – Peptide MHC

PRR – Pattern recognition receptors

RBC – Red blood cell

RFI – Relative fluorescence intensity

RTE – Recent thymic emigrant

Self-pMHC – Self-peptide MHC

SHP – SH2 domain-containing tyrosine phosphatase

SLO – Secondary lymphoid organ

SP – Single positive

SPR – Surface plasmon resonance

TCR – T cell receptor

TdT – Terminal deoxynucleotidyl transferase

Tg – Transgenic

T_{FH} – T follicular helper

T_H – T helper

TR – Transcriptional regulator

TRA – Tissue-restricted antigen

Treg – Regulatory T cell

TSS – Transcription start site

TSSP – Thymus-specific serine protease

VLR – Variable lymphocyte receptors

WT – Wildtype

CHAPTER 1: INTRODUCTION

1.1 Scope

Since 2001 *Nature* has issued milestone collections to highlight the most influential developments in science. Such collections have profiled the fields of diabetes, cancer, vaccines, and microbiota, among others. On December 6, 2022 (just in time for this thesis), *Nature* published its newest milestone: T cells. Some of the discoveries highlighted in the collection that are also discussed in this thesis include the thymus, T cell subsets, the T cell receptor (TCR), and distinguishing self from non-self, all impressive contributions in their own right. In 10 years, I hope we will see naive T cell heterogeneity receiving the attention that I think it deserves and be included in an upgraded version of the T cell milestone collection.

It has long been appreciated that $CD4^+$ T cells are able to differentiate into a variety of effector cell types upon recognition of foreign antigens. However, historically it has generally been thought that prior to encounter with foreign antigen the naive $CD4^+$ T cell population is homogeneous, despite massive diversity in the TCRs expressed by T cells. Prior to my time in the lab, we were only just beginning to understand that differences in self-ligand signal strength establish a heterogeneous naive $CD4^+$ T cell population, even prior to activation [1, 2]. For my work, we then asked: do self-signals pre-program the naive T cell population to possess biases in response to foreign antigen? Throughout my years in this lab, a body of literature from our group (presented in this thesis) and others have now shown that indeed self-signals pre-wire the naive T cell population. This thesis also addresses several other questions that we developed along my PhD journey. The rest of Chapter 1 is a literature review that explores three major areas of research essential to the work in this thesis: what is a T cell (Chapter 1.2), how do T cells 'see' antigens (Chapter 1.3), and why do T cells need to 'see' self (Chapter 1.4).

Chapter 2 presents the bulk of what we did to understand how differences in strength of self-reactivity impact naive $CD4^+$ T cell predispositions for effector cell responses post-activation. We revealed that, even prior to foreign antigen encounter, naive $CD4^+$ T cells with low and high self-reactivity have distinct pre-existing chromatin landscapes and gene expression networks that predict biases in effector cell function and differentiation. Importantly, we also demonstrate that the transcriptional heterogeneity is only partially maintained by peripheral self-peptide MHC

interactions and that thymic development may imprint some of the chromatin and transcriptional networks identified.

Most studies that have looked at naive T cell heterogeneity have focused on cells with higher self-reactivity. Yet, despite half the T cell population consisting of low self-reactive T cells, little work has asked how the low-reactive T cells contribute to the repertoire [3]. In Chapter 3, we show that low self-reactive cells dominate the T cell response in chronic infection and aid in pathogen control.

The work in Chapter 4 asks how T cells cope with the deprivation of self-ligand interactions at steady-state. While self-interactions are required for T cell survival and maintenance [4], surprisingly, naive T cells can survive for up to a month without these signals. We show that upon deprivation of self-signals naive $CD4^+$ T cells upregulate several transcriptional networks suggesting an active survival response. Chapter 4 also presents a novel finding that during infection non-responding naive T cells undergo self-signal deprivation leading to impaired responses upon subsequent activation. We also show that infection-induced self-deprivation relies on type-I interferon (IFN) signaling.

The thesis concludes with Chapter 5, which discusses the main findings from Chapters 2-4 with a focus on outstanding and new questions we have developed from the work presented here.

1.2 T is for Thymus

T cells are some of the most remarkable cells in the body. So special they have an entire organ dedicated to their development, the thymus, from which the 'T' in their name is derived. However, for several centuries, the thymus remained an enigmatic organ with no clear function. The first recorded hypothesis of thymic function was in the 14^{th} century by Joannes Galenus, a philosopher, in his commentary of the *Iliad* where he proposed the thymus was located between the collar bones to protect the trachea from injury [5, 6]. As early as 1777 William Hewson observed that the thymus was filled with 'particles' that resembled what was found in the blood and lymph [7], which we now know today are leukocytes. The thymus was the last organ to have its function identified. It was not until 1961 that Jacques Miller published his seminal paper describing the thymus as being essential for the development of 'immunologically competent cells' [8].

(i) Development – where the porridge is juuust right

The development of immunologically competent cells in the thymus, later renamed T cells, is a well-timed highly choreographed process dependent on interactions with specialized cells in the thymic cortex and medulla. Common lymphoid progenitors first arise in the bone marrow and migrate to the thymus via the blood. Once in the thymus, the progenitor cells lose their potential to develop into B cells and natural killer (NK) cells [9, 10]. Progenitor cells become double negative (DN, $CD4^-$ and $CD8^-$) T cell committed precursors which are further subclassified into four sequential stages of development based on the expression of CD44 and CD25: DN1 ($CD44^+CD25^-$), DN2 ($CD44^+CD25^+$), DN3 ($CD44^-CD25^+$), and DN4 ($CD44^-CD25^-$) [11].

During this stage of development, DN thymocytes can diverge to produce either γδ or αβ TCR expressing cells [12]. The TCR acts as the 'eyes' in developing thymocytes and mature T cells that allow them to 'see' the world via peptides presented in the context of major histocompatibility complexes (pMHC) (discussed further in Section 1.2). For αβ thymocytes, at the DN3 stage, the invariant pre-TCRα is first expressed [13]. The pre-TCRα pairs with a TCRβ-chain produced from an arrangement of Vβ, Dβ, and Jβ gene segments that have undergone somatic recombination dependent on RAG1 and RAG2 expression to form the pre-TCR [14, 15]. The expression of a functional TCRβ is required for the cell to complete β-selection [16].

Thymocytes that survive β-selection subsequently undergo several rounds of division and transition into the largest population of developing cells in the thymus, the double positive (DP) stage, where they upregulate CD4 and CD8 coreceptors and rearrange the TCRα locus (composed of Vα and Jα gene segments) [17]. The fate of DP thymocytes is determined by four processes – death by neglect, negative selection, positive selection, and lineage commitment [18]. It is TCR interactions with self-peptide MHC (self-pMHC) that provide the filtering criteria for selecting TCRs. The majority (~90%) of TCRs in the DP population interact weakly, if at all, with self-pMHC and die by neglect [19]. Low to moderate strength interactions with self-pMHC provide the necessary signals to survive (positive selection) and promote differentiation to either the CD8 or CD4 single positive (SP) stage when the TCR interacts with peptide presented by either MHCI or MHCII respectively (lineage commitment) [20]. A small fraction of thymocytes bear TCRs that interact with self-pMHC too strongly, which triggers apoptotic cell death, or clonal deletion, and are pruned from the repertoire [21-23]. Alternatively, in some instances where self-pMHC signals would normally induce deletion, thymocytes can develop into NK-T cells, intestinal intraepithelial

lymphocytes, and regulatory T cells (Tregs) [24]. By the time thymocytes complete development only a small fraction (~5%) of all possible TCRs pass the ruthless selection process and while there is a strict threshold between positive and negative selection that is permissible to survival, the 'goldilocks' zone, the breadth of self-pMHC interactions that support positive selection covers a broad range [20, 25-27]. If a $CD4^+$ SP or $CD8^+$ SP thymocyte has passed the gauntlet of development, it will leave the thymus within a few days and enter the bloodstream where it will circulate as a mature naive T cell.

(ii) Periphery – location, location, location

The fundamental role of a T cell is to search for and destroy anything non-self. Like in the thymus, T cells continue to use pMHC interactions to test their environment. To cover the most ground, naive T cells lead a nomadic life and are continuously circulating throughout the body via the blood, secondary lymphoid organs (SLOs), and the lymph [28, 29]. Once in the SLO, it might be expected that brute force migration toward a pathogen would be the most efficient way to eliminate a threat [30]. However, such behaviour is not observed in migrating naive T cells because cognate antigens presented by APCs do not provide signal cues until TCR engagement. Moreover, given the rarity of antigen-specific naive T cells, only a few cells per 10^5-10^6 total naive T cells [31-33], the best migratory strategy would be to provide all cells with equal opportunity to scan an APC [34]. Indeed, in the T cell zone of the lymph nodes, naive T cells undergo integrin-independent diffusive (Brownian-type) random walking [35-37]. While naive T cell migration lacks a collective strategy, movement is not entirely random. Naive cells utilize an interconnected network of stromal fibroblast reticular cells (FRCs) to aid in 'guided random walks' [38]. The FRCs produce the chemoattractants CCL19 and CCL21 which T cells use to migrate through the SLOs at speeds of ~11μm/min [37, 39-41]. However, during this migration, there are fluctuations in slow and fast movements that are partially linked to integrins [39]. For example, LFA-1 interactions with ICAM1 on dendritic cells (DCs) promote high-speed motility bursts in naive T cells [42]. Additionally, CCR7, the receptor for CCL19 and CCL21, promotes retention in the lymph node (LN) and is important for maintaining naive T cell speed but not directionality [43]. A naive T cell will typically spend ~4-24 hours within a given LN, with naive $CD4^+$ T cells on average spending less time in the LN than naive $CD8^+$ T cells [44, 45]. During their time in the LN, a naive $CD4^+$ T cell will scan ~200 DCs whereas a naive $CD8^+$ T cell scans ~300 DCs [44].

If a naive T cell does not find a foreign antigen it recognizes (cognate antigen), it relies on sphingosine-1-phosphate receptor 1 expression to promote egress into the efferent lymph and the migratory cycle begins anew [46].

(iii) Wonder T cell powers … Activate!

At steady-state, naive T cells are maintained in a quiescent state, which is an active process characterized by low metabolic, transcriptional, and translational activity [47]. Once a TCR engages its cognate antigen the T cell will exit quiescence and undergo activation. Activation is a complex signalling cascade that utilizes co-receptors, enzymes, and adaptor proteins to promote transcription factor activity and rapid changes within the T cells. For full activation a naive T cell requires three signals: 1) TCR signaling, 2) co-stimulation, and 3) cytokines.

Signal 1 – the TCR handshake

The first signal that kickstarts T cell activation is when the TCR engages its cognate antigen presented by MHC, triggering a downstream signaling cascade (**Figure 1**). Upon stabilization of the TCR:pMHC interaction with the co-receptors CD4 and CD8, the protein kinase, LCK, is brought to the immunoreceptor tyrosine-based activated motifs (ITAMs) on the TCR complex [48]. Phosphorylation of the TCR ITAMs provides docking sites for the cytoplasmic kinase Zap-70 [49]. Zap-70 will subsequently phosphorylate tyrosines on the linker for activation of T cells (LAT), which LCK is unable to phosphorylate [50]. This stage is often considered the rate-limiting step during cognate antigen stimulation (Discussed further in Chapter 1.3.iv). LAT serves as a hub for several signaling molecules that have wide-ranging downstream effects. For example, the phosphorylation of phospholipase C cleaves phosphatidylinositol biphosphate (PIP2) into diacylglycerol (DAG) and inositol triphosphate (IP3) [51]. IP3 then activates the release of internal calcium stores from the endoplasmic reticulum as well as entry of calcium from the extracellular environment [52]. The increased concentration of cytosolic calcium activates the phosphatase calcineurin which dephosphorylates the transcription factor NFAT allowing its translocation into the nucleus to bind target genes [53, 54]. Additionally, the increased calcium levels coupled with DAG and PIP2 activate protein kinase C and promote NF-κB signaling [54].

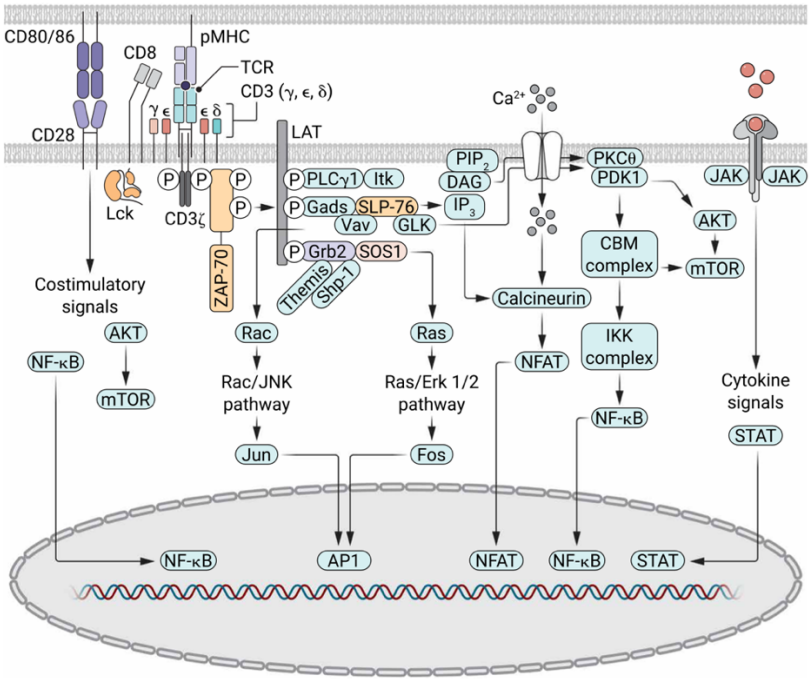


Figure 1. TCR signaling. Adapted from [55].

Signal 2 – co-receptors co-llaborate

The second signal provides necessary help for sustained T cell activation. CD28 was the first of the co-stimulatory receptors to be described in T cells [56]. Binding to the B7-1/B7-2 ligands on APCs, CD28 induces phosphoinositol-3-kinase activation which promotes ERK and mitogen activated protein kinase (MAPK) signaling to lower the threshold for activation and enhances T cell proliferation and survival [57]. Since the discovery of CD28, several other co-stimulatory molecules have been identified, such as CD44, LFA-1, 4-1BB, OX40, and ICOS [58]. When T cells do not get these second signals they become anergic or die [59].

To counter the effects of co-stimulation during TCR signaling, T cells also use co-inhibitory receptors. The first co-inhibitor described was CTLA-4 which limits co-stimulation by competing for the same ligands as CD28 [60]. In addition to CLTA-4, other co-inhibitory molecules (PD-1, LAG3, and TIGIT) play a key role in dampening the TCR signal [58]. Most inhibitory markers regulate the TCR signals by utilizing phosphatases like SH2 domain-containing tyrosine phosphatase 1 (SHP-1), SHP-2, and serine/threonine protein phosphatase 2A to dephosphorylate components of the TCR signaling cascade [58]. The interplay of the co-stimulatory and co-inhibitory molecules work together to promote T cell activation but also shut down the TCR signaling cascade when necessary.

Signal 3 – cytokines, Kool-Aid for T cells

The third signal is provided by cytokines which are crucial proteins in cell-to-cell communication and are secreted by a wide range of cells. While fibroblasts and epithelial cells can produce cytokines, cells of the innate immune system provide the primary source of polarizing cytokines that instruct lineage decisions [61]. Following the recognition of pathogen associate molecular patterns, the pattern recognition receptors (PRRs) expressed on innate cells tailor cytokine secretion based on the pathogen encountered (bacterial, viral, parasitic, and fungal) allowing the fine-tuned response of effector T cells that are best suited to control the invading pathogen.

(iv) The different flavours of T cells

The cytokines that a naive T cell is exposed to steer effector cell differentiation. The first effector subsets identified were the T helper 1 (T_H1) and T_H2 [62] which have now expanded to include the T_H9, T_H17, T_{FH}, and Treg lineages (**Figure 2A**) [63]. Effector lineages have historically been classified by a lineage-defining transcription factor and the cytokines produced. However, importantly, there is plasticity in effector lineage choices where differentiation into one lineage does not preclude the cell from adopting the function of another [64]. Additionally, effector cells can express the transcription factors of multiple lineages [65].

T_H1: T_H1 cells typically develop in response to intracellular pathogens (type I responses). Differentiation is initiated by interleukin (IL)-12 binding with the IL-12 receptor (IL-12R) which signals downstream via STAT4 to upregulate the transcription factor T-bet leading to IFN-γ, TNF-α, and IL-2 production [66, 67]. Specifically, IFN-γ can then act in a positive feedback loop and promote further T_H1 differentiation via STAT1 signaling [68]. These cytokines work together to recruit macrophages to the site of infection and enhance their ability to eliminate pathogens [69]. They also provide support for the destruction of cancer and infected cells by inducing apoptosis and enhancing $CD8^+$ T cell and NK cell cytotoxicity [65].

T_H2: Type II responses are elicited by allergens and multicellular pathogens, like helminths, which drive IL-4 production that signals through the IL-4R upregulating GATA3 and committing the T cell to the T_H2 lineage [70, 71]. T_H2 cells then produce the cytokines IL-4, IL-5, and IL-13 [72]. Like IFN-γ for T_H1 cells, IL-4 produced by T_H2 cells can sustain prolonged T_H2 responses. Because parasites are large multicellular organisms, the cytokines produced by T_H2 cells do not work by killing the pathogen but rather by assisting in the expulsion of the worms [69]. T_H2 cytokines will promote the activation and degranulation of eosinophils and mast cells, smooth muscle contraction, goblet cell hyperplasia and mucus secretion, as well as immunoglobulin E (IgE) class switching in B cells [73].

T_H9: A relatively new class of effector cells, T_H9 cells were first described in 2008 when it was found that T_H2 cells could be reprogrammed into cells primarily expressing IL-9 by stimulating them with TGF-β and IL-4 [74]. Commitment to the T_H9 lineage requires the transcription factors IRF4 and PU-1 leading to the expression of IL-9 and IL-10 [75, 76]. T_H9 cells, like T_H2 cells, play an important role in tissue inflammation and anti-parasitic responses and have recently been implicated as a key contributor to the anti-tumor response [77].

T_H17: T_H17 cells are involved in type III responses and are critical for the elimination of extracellular single-cell pathogens and fungi. Commitment to the T_H17 lineage is guided by TGF-β and IL-6 signaling through STAT3 to induce RORγt expression [78]. The signature cytokines produced by T_H17 cells include IL-17 (IL-17A and IL-17F), IL-21, IL-22, and TNF-α [79]. IL-17 signaling through the IL-17R, which is expressed on many cell types including stromal cells like fibroblasts and epithelial cells, leads to the production of β-defensin and C-type lectins that kill bacteria directly [80]. In addition, the cytokines secreted by T_H17 cells aid in the recruitment of neutrophils and macrophages to the site of infection [81].

Treg: While the above-mentioned T_H cell lineages function to promote immune responses against pathogens, Tregs dampen the immune response. First described in 1995, $CD4^+$ T cells expressing the IL-2Rα chain (CD25) were shown to prevent spontaneous autoimmune diseases [82]. In 2003, mutations in the transcription factor Foxp3 were responsible for the development of severe multi-

organ autoimmune diseases [83]. Independently in 2003, it was discovered that these $CD25^+$ $CD4^+$ T cells were dependent on Foxp3 expression, later termed Tregs [84].

Tregs can develop either in the thymus or in the periphery. Tregs formed in the thymus require access to IL-2 and are derived from CD4-committed thymocytes that typically receive stronger self-signals during development, referred to as natural Tregs [84]. In the periphery, naive $CD4^+$ T cells can be instructed to become Tregs in the presence of IL-2 and TGF-β, induced Tregs [84]. Tregs utilize several mechanisms to inhibit effector T cell function such as IL-2 sequestration, co-inhibitory signaling, and immunosuppressive cytokine production. Tregs express increased levels of the high-affinity IL-2Rα and act as a sink for IL-2 by mopping up available cytokine preventing use by other effector T cells [85]. Additionally, Tregs express high levels of inhibitory receptors like CTLA-4, PD-1, LAG3, and TIGIT [86]. Another mechanism that Tregs use to dampen T cell responses is by producing the cytokines TGF-β and IL-10 which inhibit T cell proliferation and reduce antigen presentation and pro-inflammatory cytokine production by APC [87].

T_{FH}: T_{FH} cells are rather complicated cells, their differentiation is not defined by a single cytokine but instead, fate commitment relies on multistage multifactorial signals. Upon priming by a DC in the SLO T cell zone, if a naive $CD4^+$ T cell upregulates the chemokine receptor CXCR5 it will migrate to the border of the B cell follicle [88]. This early T_{FH} differentiation phase is regulated by TCR signal strength, IL-6, ICOS, and IL-2. On average, naive T cells that have longer dwell times with pMHC, thus presumably receiving stronger TCR signals, preferentially become T_{FH} cells [89]. IL-6 signalling transiently upregulates the lineage-defining transcription factor BCL6 which is necessary for CXCR5 expression [90]. Signalling by ICOS synergizes with IL-6 to reinforce BCL6 expression for early T_{FH} differentiation and migration to the B cell follicle [91]. In addition, IL-2 signals are a potent inhibitor of T_{FH} commitment early during T cell priming [92, 93].

The next stage in T_{FH} commitment occurs when T cells interact with antigen-specific B cells in the follicles of the B cell zone. As the early T_{FH} cell upregulates CXCR5 it decreases the expression of CCR7 allowing the cell to leave to T cell zone. Once in the follicle, B cells become the primary APC population for the early T_{FH} cells and provide ICOS-ICOS ligand signals that promote proper positioning [94]. The final stage of T_{FH} lineage commitment involves the germinal

center (GC). The GC is a structure within the B cell follicles consisting of GC T_{FH} cells, GC B cells, follicular dendritic cells, macrophages, and stromal cells [95]. Within the GC, T_{FH} cells produce IL-21 and provide help for B cells undergoing affinity maturation (reviewed in [96]).

Until the discovery of T_{FH} cells, it was thought that cytokine signaling fixed lineage fate decisions of effector cells. However, neither IL-6 nor IL-21 are able to induce sustained expression of BCL6 and stably polarize T cells to the T_{FH} lineage and, unlike other effector T cells, *in vitro* differentiation (which uses cytokines to polarize activated $CD4^+$ T cells) of T_{FH} cells has been unsuccessful to date. Additionally, T_{FH} cells arise in diverse immune environments and can produce several context-dependent cytokines like IFN-γ, IL-4, and IL-17, thus blurring whether T_{FH} cells are discrete effector lineages or not [97]. Recently, a new paradigm in the $CD4^+$ T cell fate decision has been proposed whereby early in activation a bifurcation event occurs when a T cell 'decides' to become a *bona-fide* T_{FH} cell or undergo effector differentiation (**Figure 2B**). This bifurcation event can thus produce any of the classical effector cells but also a T_{FH} counterpart ($T_{FH}1$, $T_{FH}2$, $T_{FH}17$, and T follicular regulatory [T_{FR}] cells) [98].

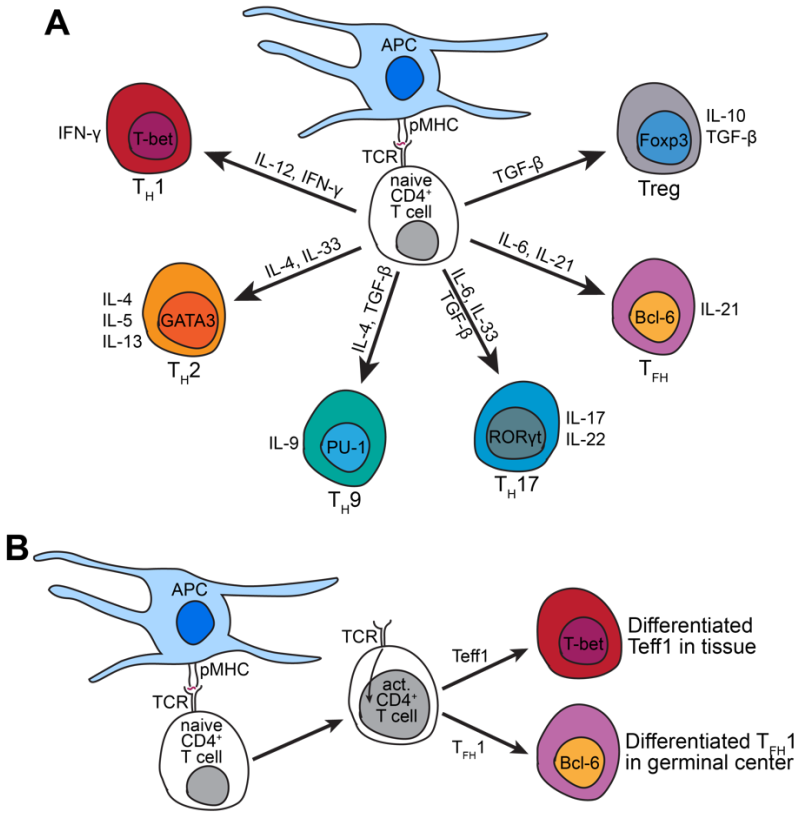


Figure 2. $CD4^+$ T effector lineage choices. (**A**) The different effector lineages with polarizing cytokines, lineage-defining transcription factors, and cytokine production. (**B**) The new paradigm of early activated $CD4^+$ T cell bifurcation in fate decision with a focus on type I polarization.

(v) One versus many – single-cell and population-based diversity

Flexibility and diversification are key features of an effective T cell response. While the number of precursor frequencies ranges from a couple dozen to several hundred epitope-specific naive T cells per mouse [31], during systemic infection nearly all of the naive T cells specific to the invading pathogen are recruited into the response [99]. Thus, diversity within a T cell response must arise from several naive T cells recognizing the same epitope. A number of non-mutually exclusive ways to generate this T cell diversity have been proposed for individual cells as well as by population averaging [100]. At the individual cell level, the two hypotheses proposed are 'one cell, one fate' and 'one cell, multiple fates'. From the population level, diverse responses can be generated in a fractal (the diversity proportion in single T cells scales for the entire epitope-specific population) or composite manner (each T cell within an epitope-specific population possesses a distinct pattern of diversification). Several studies have shown that T cell responses support the 'one cell, multiple fates' hypothesis and that population-level T cell responses are built in a composite fashion [89, 101-103].

(vi) Conclusion

T cells have the extraordinary capacity to become several flavours of effector cells and to tailor their responses to the type of pathogen being encountered. But they also face what seems like an insurmountable obstacle, they must find their cognate antigens. I imagine this search is like finding a needle in a haystack, except the haystack is the size of a mountain, it is dark, and you only know you found the needle once it pokes you. To top it off, it is likely that the needle you are searching for is on a different mountain. The next section discusses how T cells learn to recognize the diverse world of ligands and differentiate self from non-self.

1.3 Seeing the world through the perspective of a T cell

The adaptive immune systems of both jawless fish and jawed vertebrates have evolved receptors, the variable lymphocyte receptors (VLRs) A and C and the TCR respectively, that recognize foreign antigens from the immensely diverse world of pathogens [104]. However, while it is likely that these receptors evolved as a mechanism to selectively recognize pathogens, a key feature of antigen recognition via the TCR is that T cells must be able to discriminate self from non-self and that self-recognition does not overtly activate naive T cells. In this section, we will

explore how TCR diversity is generated, how T cells differentiate self from foreign, and tools that are used to measure TCR signal strength.

(i) Be fruitful and diversify – TCR generation

As thymocytes develop they rearrange the α- and β-chains of the TCR. Germline TCR sequences establish an evolutionary baseline for TCR diversity encoded by somatic recombination of V, J, and D (β-chain only) gene segments [105]. In mice, the α-chain locus constitutes ~100 V and 60 J segments and 35 V, 12 J, and 2 D segments in the β-chain [106, 107]; for humans, ~70 V and 61 J segments of the α-chain and 52 V, 13 J, and 2 D segments of the β-chain [108]. As the V(D)J gene segments are arranged the DNA polymerase terminal deoxynucleotidyl transferase (TdT) adds ~2-7 additional non-templated (N)-nucleotides at the junction sites of the V(D)J segments accounting for >90% of TCR diversity [109-111]. However, the addition of these N-nucleotides is not entirely random where TdT biases the insertion of dGMP and dCMP over dAMP or dTMP [110, 112]. Estimations have placed the number of possible TCRs in the range of 10^{15}-10^{20} [113-115], a staggering number akin to the estimated number of stars in the observable universe (10^{24}). Yet, not all TCR gene recombination events are useful and stringent selection processes weed out T cells bearing non-functional and auto-reactive antigen receptors [116]. Still, the naive T cell population likely consists of hundreds of millions of different TCRs representing individual T cell clonotypes and is a major source of heterogeneity in this population [113, 117]. Thus, the repertoire of an individual cannot possibly possess all TCRs and a subset of potential sequences must be sufficiently large enough to respond to all foreign challenges. Studies have suggested that the likelihood of a particular TCR being represented in the mature T cell repertoire is dependent on the probability of that TCR being encoded by multiple sequences (ie. sequences with multiple codons that translate the same amino acid are more likely to be present than sequences that encode amino acids with few redundant codons) [118]. However, it remains unclear if there are biases and patterns in the choice of TCRs and how a T cell repertoire is selected based on the TCRs used.

(ii) Features of the TCR

Different structural components of the TCR interact with peptide and the MHC molecule itself. The complementarity-determining region 1 (CDR1) and CDR2 of the TCR are encoded by

germ-line sequences whereas the CDR3 is composed of the V(D)J gene segments along with the TdT-mediated N-nucleotides insertions that form the antigen-binding site of the TCR. Upon binding of ligands, the residues of the CDR1 and CDR2 interact with the MHC, whereas the CDR3 interacts with the peptide [113, 119]. Together, the CDRs of the TCR adopt an overall shape that is complementary to pMHC. Because the CDR1 and CDR2 are conserved they only adopt a few secondary structures which limit structural diversity among the TCR thus establishing consistent MHC contacts across the T cell repertoire [120, 121]. Conversely, there is little complementarity in the shape of the TCR with pMHC ligands [122]. Additionally, amino acid usage has been shown to aid in MHC ligand specificity. Hydrophobic amino acid residues at positions 6 and 7 within TCRα and TCRβ chain CDR3s increase the frequency and strength of TCR interactions with pMHC likely by excluding water molecules from the binding site [55]. Together, these structural features of the TCR inform sequence-specific effects of the TCR on ligand binding.

The physical features of the TCR promote stable interactions with pMHCs. However, whether TCR engagement of pMHCs results in catch bonds (force prolongs the TCR-pMHC interaction) versus slip bonds (force shortens the TCR-pMHC interaction) has remained a topic of debate [123-125]. Despite this, it is clear there are distinct differences in the binding kinetics of the TCR with long-lived foreign-pMHCs versus the short half-life of self-pMHCs [26, 126-128]. It is the integration of these TCR-pMHC interactions and the downstream cellular signaling networks that are required for T cells to surpass the digital (on/off) activation threshold.

(iii) Setting self-pMHC signal strength

Once a T cell rearranges an αβTCR it is tested by interacting with self-peptide presented by thymic APCs. It is generally thought that self-reactivity is determined at the DP stage during positive selection when the mature αβTCR is first expressed. Increasing evidence suggests however that, in fact, self-peptides during β-selection may play a previously underappreciated role in shaping T cell self-reactivity. Historically, the pre-TCR has been thought to not be necessary for DN thymocytes to pass through β-selection [129-131]. More recently, however, biochemical studies have demonstrated that the pre-TCR interacts with pMHC similarly to the αβTCR but in a distinct conformation and, likely, with broader specificity [132-135]. Binding of the pre-TCR to pMHC at this stage initiates the TCR signaling cascade and mutations that inhibit pre-TCR interactions with pMHC partially inhibit progression to the DP stage of development [135]. At the

β-selection stage, the absence of self-pMHC decreases cell proliferation, TCR repertoire diversity, and promotes abnormal differentiation [134, 136]. Together these studies suggest that both self-ligand-dependent and -independent β-selection may establish an initial distribution in self-reactivity amongst developing thymocytes before a mature TCR is even expressed (**Figure 3**).

At the DP stage of development, thymocytes undergo positive selection by receiving self-signals from cortical thymic epithelial cells (cTECs) that support commitment to the $CD4^+$ or $CD8^+$ SP stage. The strict threshold in strength of self-reactivity sets the bounds for positive and negative selection and only ~5% of all possible TCRs survive this strict selection criteria [20, 25-27]. Notably, the efficiency with which an individual thymocyte completes positive selection may impact the overall self-reactivity of the TCR repertoire. Thymocytes that bear TCRs with higher affinity for self-pMHC complete positive selection and differentiate into SP thymocytes more rapidly [137]. Moreover, in the $CD4^+$ T cell lineage, there is a non-normal skewed distribution of CD5, a marker of self-reactivity (discussed in section 1.3.v), suggesting that thymocytes expressing TCRs at the higher end of the permissible self-reactivity spectrum preferentially pass the positive selection bottleneck [1]. Overall, the various developmental stages that thymocytes receive self-pMHC signals likely play a role in collectively establishing the breadth of self-reactivity among the mature T cell repertoire (**Figure 3**).

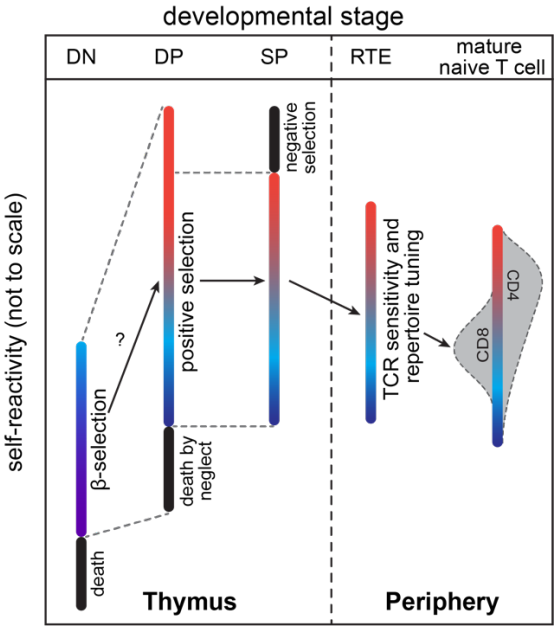


Figure 3. The self-immunopeptidome in establishing naive T cell heterogeneity. Double negative (DN) thymocytes rearrange the β chain of the T cell receptor (TCR) pairing it with the pre-TCRα. These DN cells undergo β-selection which may establish an initial setpoint for T cell self-reactivity. Upon successful rearrangement of the α chain, double

positive (DP) thymocytes undergo positive selection whereby self-peptides are presented by cortical thymic epithelial cells (cTEC). Cells that do not receive self-signals at this stage die by neglect. DP thymocytes then commit to either $CD4^+$ or $CD8^+$ single positive (SP) thymocyte lineages and are negatively selected on self-pMHC presented by medullary thymic epithelial cells (mTEC), dendritic cells (DC), and B cells. The thymocytes that respond too strongly to self are removed from the T cell repertoire. As a population, SP thymocytes that survive negative selection represent a distribution of self-reactivities, which in $CD4^+$ SP T cells, is skewed towards cells with higher self-reactivity. Upon egress into the periphery, recent thymic emigrants (RTE) undergo retuning of TCR signal sensitivity and are incorporated into the mature naive T cell pool. Adapted from [4].

The breadth of self-reactivity may also be a function of the TCR or an outcome of the mechanisms generating TCR diversity. In particular, the addition of non-templated nucleotides by TdT at the V(D)J junctions during TCR recombination may broaden the self-reactivity of the TCR repertoire [138]. Both in naive $CD4^+$ and $CD8^+$ T cells, expression of the TdT gene, *Dntt*, has been shown to be substantially greater (~10 fold) in low-affinity T cells [139-142]. Indeed, it has recently been shown that there is a correlation between greater expression level of *Dntt* during TCR rearrangement and more added n-nucleotides that, on average, produce TCRs of lower self-reactivity [143]. Moreover, thymocytes from TdT-deficient mice more efficiently undergo positive selection suggesting the TCR repertoire may be enriched for relatively higher affinity clones that are more cross-reactive [144, 145].

(iv) Discriminating self from non-self

Discriminating between self and foreign is a formidable challenge given that peptides from pathogens closely resemble those encoded by mammalian genomes [146]. To borrow a language-learning parallel that has been proposed for their training in the thymus, T cells learn to 'read' the language of peptide presented in the context of MHC and are able to distinguish between the peptides (words) originating from self versus foreign proteins (two different languages) [146]. Ultimately, how T cells are mostly able to successfully make this distinction is still not fully understood, but likely requires a broad diversity in TCRs for possible pMHCs [147].

The kinetic proofreading model provides the best explanation for how T cells can discriminate self from foreign to receive tonic signals or undergo activation respectively. Multi-step reversible biochemical reactions downstream of the TCR accumulate to create a time delay between the pMHC interaction and the induction of T cell activation, or not [148] (**Figure 3**, reviewed in detail here [149]). Given that such biochemical steps are reversible, the TCR-pMHC

interactions must either be of sufficient affinity and density, long enough half-life, or are stable enough to activate the T cell. While both self- and foreign-pMHC signals induce phosphorylation of the ITAMs on CD3 or TCRζ, self-ligand interactions with the TCR are weaker with shorter half-lives preventing naive T cells from reaching the irreversible activation threshold within the signaling cascade [50, 150, 151]. Whereas, foreign-pMHC signals are of sufficient length and strength to overcome the necessary steps to kickstart T cell activation. Recent work has shown that while low-affinity ligands phosphorylate the CD3 chains and induce Zap70 association similarly to high-affinity ligands, low-affinity ligands cannot lead to Zap70 phosphorylation resulting in impaired LAT and CD6 signalosomes (the multi-protein signaling complexes assembled by LAT and CD6) [152]. Importantly, while the physical location of the TCR intracellular signaling machinery plays an important role in self versus foreign discrimination [153]; it remains unknown which step in the kinetic proofreading model acts as the switch for *bona-fide* T cell activation.

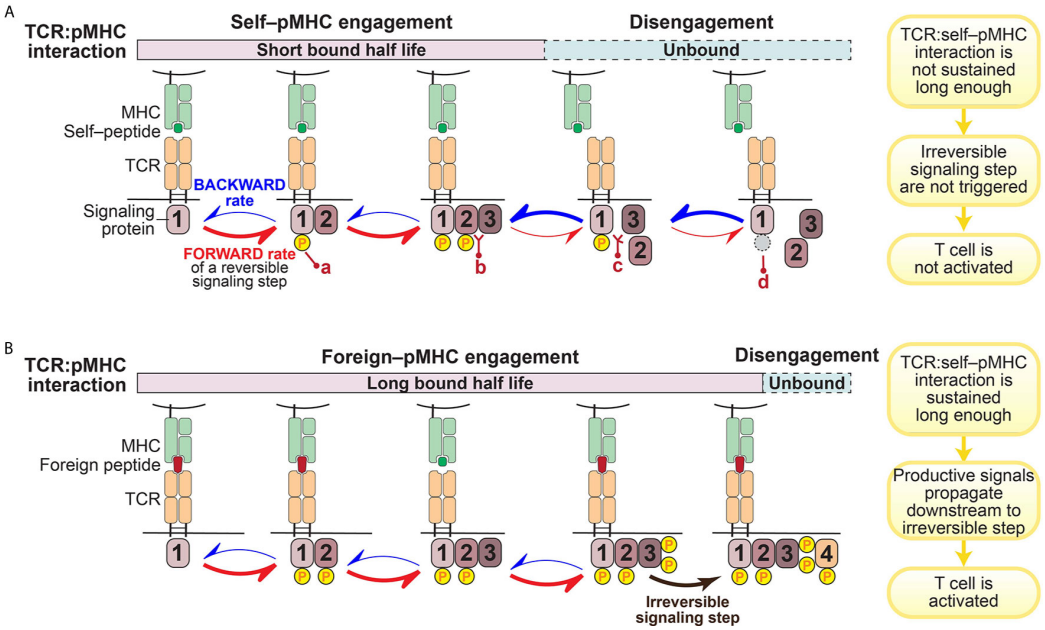


Figure 4. Kinetic proofreading model for discrimination of self versus foreign ligands. Propagation of TCR signals is triggered by a series of reversible biochemical reactions that ultimately lead to T cell activation. These biochemical reactions include phosphorylation (a), protein-protein interactions (b), protein-protein dissociation (c), and dephosphorylation (d). (**A**) TCR signals from self-pMHC are short-lived weak interactions that do not propagate down the signaling cascade to induce irreversible activation before the TCR disengages with self-pMHC. (**B**) When TCR interactions with foreign-pMHC are sufficiently long they accumulate phosphorylation across the reversible steps and overcome the irreversible T cell activation threshold. Adapted from [149].

The LAT signalosome plays a pivotal role in maintaining naive T cell ability to differentiate self and foreign antigen. During tonic self-ligand interactions, the LAT signalosome includes phospholipase Cγ1 (PLCγ1)-DAG which subsequently phosphorylates the transcriptional repressor HDAC7 leading to cytoplasmic accumulation [154]. Loss of LAT will allow HDAC7 to accumulate in the nucleus repressing Nur77 and IRF4 thus blunting T cell proliferation and differentiation upon encounter with cognate antigen [154]. In fact, the phosphorylation of LAT at tyrosine 132 appears to be an essential rate-limiting step for ligand discrimination of self versus foreign [155, 156].

Following the formation of the proximal TCR signalosomes, downstream transcription factors fully activate naive T cells characterized by extensive chromatin remodeling, transcription, metabolic shifts, proliferation, and effector cell commitment. Most notably following LAT activation, self- versus foreign-ligand discrimination is controlled by competing ERK (positive) and SHP1 (negative) feedback loops [157, 158]. Weak pMHC binding, such as recognition of self-ligands, recruits the phosphatase SHP-1 to the proximal TCR signaling machinery and desensitizes the TCR by inactivating LCK; following strong foreign ligand binding, phosphorylation of ERK via LCK prevents SHP-1 recruitment and promotes sustained signaling leading to gene activation [158]. However, given the complexity of the TCR signaling cascade, it is likely that self vs foreign discrimination is the result of the cumulative contribution of many interconnected signaling networks which have not yet been fully elucidated.

(v) Measuring TCR signal strength

We classify ligands as either of self or foreign origin where self-pMHC interactions are generally weaker than foreign. However, within both sources of pMHC, there is heterogeneity in binding strength that can vary by several orders of magnitude [159, 160]. A great deal of interest has gone into developing tools to measure pMHC binding strength and assess T cell heterogeneity.

Surface plasmon resonance: One of the first techniques developed to measure TCR interactions with pMHC was surface plasmon resonance (SPR) in 1994 [161, 162]. TCR is flowed over pMHC bound to a sensor surface. As TCRs bind the pMHC, mass accumulates on the sensor surface thus increasing the signal which can be used to obtain binding kinetics and affinities. However, this

technique is not easy to scale to a polyclonal population and requires prior knowledge of the epitope to be loaded onto the MHC.

Two-dimensional binding assays: Shortly after SPR was developed, the micropipette adhesion frequency assay (later renamed the two-dimensional [2D] binding assay) was introduced [163]. For this assay, pMHC is coated onto the membrane of a red blood cell (RBC) through a biotin-streptavidin system. The RBC and a T cell are immobilized on a micropipette and brought into contact with each other to establish TCR:pMHC interaction. The kinetics and affinity are determined based on the frequency of contacts within a set timeframe and how much the RBC membrane deforms upon retraction. Like SRP, this technique also suffers from the same requirement of *a priori* knowledge of the pMHC specificity for the TCR being tested.

Tetramers: Tetramers, first described in 1996 [164], revolutionized the way antigen-specific $CD4^+$ and $CD8^+$ T cells were studied. Briefly, recombinant MHCI or MHCII monomers loaded with a peptide of interest are biotinylated and tetramerization is achieved by adding a streptavidin molecule conjugated with a fluorophore [165]. Tetramers can measure the binding strength for a particular antigen-specific T cell by assessing the dissociation rate of the tetramer off the cell, although low-affinity T cells tend to be missed by tetramers [128, 166]. The main disadvantage of using tetramers is that you can only assess one clonotype at a time and, like SPR and 2D binding assays, prior knowledge of the epitope specificity is required.

TCR transgenic mice: Developed in 1988 [167], TCR transgenic mice have now become a foundational tool for studying antigen-specific T cell responses. Unlike tetramers which assess single T cell clones within a polyclonal population, TCR transgenic mice only have one clonotype that makes up its entire T cell pool. A repertoire composed of a single T cell clone simplifies the interpretation of the T cell response to antigen but can make it difficult to draw parallels with the polyclonal T cell population.

Surrogate markers: The sub-threshold signals obtained from self-pMHC at steady-state are challenging to measure directly, particularly as only very few self-peptides have been identified so far for known TCRs [126, 168-170]. Thus, the characterization of functional heterogeneity

among naive T cells has relied on markers whose expression levels provide read-outs of self-reactivity. One of these is CD5, a surface glycoprotein expressed on all T cells that modulates TCR signaling [142, 171-175]. Two others are the orphan nuclear receptor member *Nr4a1* encoding Nur77, for which a GFP reporter can identify sub-threshold TCR signal strength [176-178], and Ly6C, a GPI-linked receptor that is upregulated on a subset of naive T cells shortly upon egress from the thymus [179]. While both CD5 and Nur77 are broadly distributed on $CD4^+$ and $CD8^+$ T cells and directly correlate with self-reactivity, Ly6C has a bimodal distribution on naive $CD4^+$ T cells whereby $CD4^+$ T cells with relatively low self-reactivity are $Ly6C^+$ and $CD4^+$ T cells with higher levels of self-reactivity are $Ly6C^-$ [179]. In some studies, the combination of all three markers is used to subset naive $CD4^+$ T cells [180, 181], while other studies have focused on comparing $CD5^{lo}$ to $CD5^{hi}$ naive T cells [2, 3, 140-142, 182-184] or $Ly6C^+$ with $Ly6C^-$ naive $CD4^+$ T cells [139, 179]. Of note, recent findings suggest that CD6 might be a more stable marker of self-reactivity than CD5 with differences that are more obviously maintained following activation [140]. However, CD6 has not been used to date to assess functional heterogeneity among naive T cells.

(vii) Conclusion

T cells face strong selective pressure in the thymus, it is more likely than not for a thymocyte to die than survive development. For those lucky few thymocyte soldiers that do survive, they have been trained by interactions with self-ligands that establish their rank within a polyclonal population as defined by the strength of self-reactivity. Once a T cell leaves basic training (the thymus) and is deployed to the periphery it needs to reliably be able to identify civilians (self ligands) from enemies (foreign ligands). *In this thesis, we are interested in parsing naive T cell heterogeneity by self-reactivity and asking whether there are predispositions for T cells in response to activation and lineage choices. This is the topic of interest for Chapters 2 and 3.*

1.4 Γνῶθι σεαυτόν [Greek, Oracle of Delphi]: "Know thyself"

The idea that the body can discriminate self from foreign has been traced back as far as ancient Egypt and India where skin from a person's cheek or forehead was used as skin grafts in early rhinoplasty and otoplasty procedures [185, 186]. However, why skin grafts from the same

person survived (sometimes) and did not if the skin was from a different individual remained unknown. It was not until the 1950s that immunological tolerance was proposed where the introduction of foreign antigens in fetal development prevented the 'antibody-forming system' (T and B cells were not identified yet) from rejecting grafted tissue in adulthood [187]. In the 1980s, T cells were identified as the linchpin in establishing tolerance to self by clonally deleting T cells with TCRs that recognize a selecting self-ligand [167, 188]. However, we now know that not all self is 'bad' self and, in fact, self-pMHC signals are essential for all T cell development and in the periphery naive T cells will receive subthreshold self-signals for the rest of their lives. In this section, we discuss what is self, what self-signals do for naive T cells, and how self-signals maintain a functionally heterogeneous population.

(i) Self-immunopeptidome – all the things you are

The self-immunopeptidome is the full set of peptides that are derived from proteins encoded by an individual's genes and that can be processed and presented by APCs on their MHC molecules. It is likely that the set of peptides that constitute the self-immunopeptidome is highly individualized due to genetic polymorphisms, variations in protein expression levels, splice variants, and binding differences of distinct MHC alleles. Estimates have suggested that the self-immunopeptidome consists of 10^4-10^6 peptides by calculating non-overlapping peptides from the proteome that can be bound by MHC [146, 189-191]. Of note, the availability of better MHC I peptide binding prediction software, and the greater ease with which mass-spectrometry studies have been able to elute and sequence MHC I presented peptides have meant that while our understanding of the self-peptidome is far from complete, it is less well defined for MHC II- than MHC I-bound peptides. Adding complexity, our understanding of what 'self' is has evolved over recent years with the emerging view that 'self' may extend beyond the genome-encoded proteome to include peptides from the microbiome or endogenous viral elements, as well as cryptic peptides generated from apparently untranslated genomic regions or non-contiguous protein segments [191-193]. Moreover, with age, changes in protein translation, apoptotic cell uptake, and/or antigen presentation both in the involuting thymus and in the periphery by DCs may also lead to shifts in the composition of the self-immunopeptidome [194-197]. Importantly, the self-immunopeptidome is also a function of the organ, the presenting cell type, and whether it expresses only MHC I or also MHC II, as well as the peptide processing machinery available.

(ii) Self-immunopeptidomes in the thymus and periphery

Self-proteins are degraded either in the cytoplasm by proteasomes or in endosomes by proteases where peptide fragments are then loaded onto MHCI or MHCII to be presented to the developing thymocytes. These different proteolytic processes presumably provide unique activity that alters the repertoire of degraded self-peptides generated. The cTECs, which positively select thymocytes, uniquely express the β5t-proteasome subunit as well as the cathepsin-L endopeptidase and thymus-specific serine protease (TSSP) [198-201]. The β5t-proteasome degrades proteins for peptide loading onto MHCI and cathepsin-L and TSSP for MHCII [198-204]. Conversely, the medullary thymic epithelial cells (mTEC), which negatively select thymocytes, do not express the β5t-proteasome subunit or cathepsin-L, and instead predominantly rely on the constitutive proteasome and immunoproteasome to generate their MHC-bound peptide repertoires [205]. Additionally, mTECs broaden their self-immunopeptidome by using the transcription factor autoimmune regulator (AIRE) to promiscuously express a 'shadow self' in the thymus that includes tissue-restricted antigens (TRA) [206-208]. More recently, mTECs have been shown to differentiate into cells that mimic peripheral cell types including microfold cells and tuft cells, among others [209-211]. Microbiome-derived peptides can also be shuttled to the thymus by migratory DCs further expanding the self that is used to train T cells [212-215]. However, it remains unknown whether the self-immunopeptidome that is presented to thymocytes is a random sample of all possible self-peptides generated or whether there are features of presented self that facilitate self- versus non-self-discrimination and, ultimately, T cell functionality, as has been suggested [146].

In the periphery, many hematopoietic cells such as macrophages and DCs express both the constitutive proteasome subunit and immunoproteasome, and the immunoproteasome can be upregulated in hematopoietic and non-hematopoietic cells in the presence of pro-inflammatory cytokines to impact peptide processing as well as gene expression [216]. Thus, the medullary self-immunopeptidome is likely more reflective of the peripheral peptide repertoire than that found in the thymic cortex. However, how much the thymic and peripheral self-immunopeptides overlap remains an open question. Importantly, restricting the self-immunopeptidome leads to severely curtailed TCR diversity emphasizing the critical crosstalk between self and T cells in determining the TCR repertoire composition [138, 217].

(iii) Maintaining T cell homeostasis and repertoire diversity

Once T cells leave their thymic training ground and become part of the mature naive T cell pool, their key function is to survey for foreign-pMHC. In doing so, they transit between SLOs via the blood and move within each SLO, scanning pMHC on DCs and stromal cells [29, 218]. Given that an individual naive T cell's usefulness hinges on this continued surveillance activity, it is perhaps not surprising that their scanning function is intimately tied to naive T cell survival, providing an opportunity to remove T cells from the pool that have ceased to make productive contacts with self-pMHC. The reliance on tonic TCR signals for naive T cell survival has been shown by inducible ablation of TCR or TCR-proximal signaling molecules, as well as adoptive transfers into MHC I or II-deficient mice [219-223]. For reasons that are still unclear, loss of tonic self-pMHC signals results in only a gradual decline in naive T cell numbers. Self-deprived naive $CD8^+$ T cells have a half-life of 16-19 days, and naive $CD4^+$ T cells have a half-life of 30-45 days [221, 223]. In contrast, the loss of IL-7 signaling, another key survival signal for naive T cells, leads to a much more rapid decay [224]. The ability of T cells to survive without self-signals is perhaps even more striking when compared to B cells which die within 3-6 days of tonic signal deprivation [225]. Interestingly, it is also not understood how tonic sub-threshold TCR signals provide naive T cells with survival cues. The absence of IL-7 signals leads to reduced cell size, lower levels of the glucose transporter GLUT1, reduced mitochondrial function, and decreased ATP levels [226], but whether lack of access to self-pMHC results in a similar 'starvation' phenotype in naive T cells has not been examined.

While the continued relationship with self-pMHC in the periphery is a key means by which the naive T cell surveillance function is enforced, self-pMHC contacts and competition for these contacts within the naive T cell pool are also thought to play a role in maintaining TCR repertoire diversity. Precursor frequencies of T cells with the same foreign-pMHC specificity within a polyclonal naive repertoire range from a few dozen to several hundred antigen-specific cells [32, 227, 228], but whether similar frequencies of T cells recognize the same self-pMHC and how foreign pMHC specificity relates to the particular self-peptide(s) able to provide tonic survival signals remains unknown. It is likely that at least a degree of overlap in self and foreign recognition between T cells plays a role in maintaining distinct precursor frequencies given that TCR cross-reactivity is confined to a narrow set of related peptides [147, 229]. However, examples have also

been described where T cells with the same foreign-pMHC specificity are not selected on the same self-pMHC, and, conversely, where T cells with distinct foreign pMHC specificity do recognize the same self-ligand [126, 168, 230]. Evidence for the idea that naive T cells compete for tonic self-ligands comes from transferring TCR transgenic $CD4^+$ T cells into lymphoreplete recipients and showing that monoclonal naive T cells have a longer half-life when added in small compared to large numbers to an established polyclonal repertoire [228]. Indeed, when naive T cell numbers drop, increased availability of IL-7 and reduced competition for self-pMHC pushes the normally quiescent naive T cells into division [231]. Notably, naive T cells with distinct self-pMHC reactivities differ in their ability to contribute to the reconstitution of a T cell repertoire during lymphopenia-induced proliferation (LIP), likely as a function of the intensity of competition for access to particular self-ligands [232-234]. Thus, weakly self-reactive T cells undergo slow LIP and maintain a naive phenotype whereas high self-reactive T cells rapidly proliferate under lymphopenic conditions and adopt memory-like phenotype [232, 233, 235, 236]. Interestingly, CD5 levels increase as competition for self-peptide is reduced in lymphoreplete environments though the functional impact of this has not been tested [232].

(iv) Plasticity, tuning and readiness

While there is evidence that functional biases among naive T cells are at least partly imprinted during thymic selection and subsequently maintained in the periphery, some level of plasticity exists in mature naive T cells that is influenced by the quality of continued interactions with the self-immunopeptidome in SLOs. Evaluation of the similarities and differences in the function of recent thymic emigrants (RTE) and their mature naive T cell counterparts provides valuable information regarding which characteristics might be set in the thymus versus maintained or induced in the periphery. RTEs retain a heightened sensitivity to TCR engagement with self-pMHC, display elevated levels of CD5 and Nur77 as compared to more mature naive T cells, and have an enhanced ability to reconstitute a lymphopenic environment [177, 237, 238]. In some respects, the functional characteristics of RTEs are similar to those at the high end of the self-reactivity spectrum of the mature naive T cell population in that RTEs preferentially differentiate into Tregs [239]. However, other functional biases associated with higher self-reactivity are not replicated in RTEs, including biases in $CD4^+$ T helper differentiation and the increased contribution to acute immune responses; these may be thymically imprinted characteristics that

are independent of the strength with which the T cells interact with the self-immunopeptidome in the periphery. In line with this, CD5 expression itself, while modulated in response to access to self-ligands in the periphery, remains distinct between different TCR transgenic populations with different CD5 levels in transitioning from thymic SP T cells to mature naive $CD4^+$ T cells, and sorted $CD5^{hi}$ naive CD4 T cells do not become $CD5^{lo}$ even when deprived of self-ligand access [1, 44, 140, 240].

Interestingly, the pathways that have been implicated in imposing naive T cell quiescence, as defined by their low metabolic activity and G0 cell cycle state, do not seem to be a function of self-pMHC reactivity, relying on gatekeepers such as Tsc1 and VISTA expression, although their relationship with self-reactivity has perhaps not been fully explored [47, 93, 140, 241]. Despite being actively quiescent, naive T cells nonetheless remain in a state of readiness for activation, rapidly turning over a subset of proteins, maintaining a reservoir of glycolytic enzymes, and preserving a large number of idling ribosomes, all of which facilitate the rapid changes T cells undergo during activation [242]. Whether these 'readiness' features of naive T cells are modulated by self-pMHC interactions has not been investigated, but these studies emphasize that there are features of naive T cell function that rely on plastic, ongoing processes. Indeed, transient variations in protein levels of SHP1 and ERK can impact the diversity in T cell responsiveness to antigen in $CD8^+$ T cells, explaining some, but perhaps not all, of the intra-clonal variability observed within TCR transgenic populations [243]. Similarly, the transcriptional wiring of naive $CD4^+$ T cells was recently shown to have both imprinted (presumably through epigenetic modifications) as well as continually adjusted components, and at least some of the plasticity in gene expression arises as a result of continued engagement with self-pMHC [140] (as described in Chapter 2), although specific molecular mediators (such as type I interferon and asymmetric cell division) may also play a role [241, 244-247]. Moreover, it has been shown that the removal of tonic self-pMHC signals leads to the rapid dephosphorylation of the TCRζ chain and ZAP70, and results in stunted TCR signaling, proliferation, and IL-2 production upon foreign pMHC recognition [248]. In addition, in the $CD8^+$ T cell lineage, disruption of interactions with self-peptide leads to increases in coreceptor expression and, ultimately, an increase in sensitivity to low-affinity ligands [249]. Whether and which functional traits underlying naive T cell heterogeneity can be ascribed to being fixed versus plastic is therefore an important outstanding question.

(v) A heterogeneous naive T cell population dependent on self

There is a long-standing assumption that naive $CD4^+$ and $CD8^+$ T cells are largely homogeneous populations despite the extraordinary diversity of their TCRs. While, on the population level, naive $CD4^+$ T cells express higher CD5 and Nur77 compared to naive $CD8^+$ T cells suggesting that naive $CD4^+$ T cells may receive stronger self-pMHC interactions. The space within the self-reactivity spectrum of a polyclonal population that an individual naive T cells occupies may not entirely be dependent on the affinity and/or avidity of the TCR interaction with self-pMHC. It has been suggested that the tissue distribution of self-ligands may also impact the intensity of self-recognition by naive T cells. For example, naive T cells that are higher in self-reactivity may be selected on 'public' peptides that would be readily accessible in the periphery; whereas, some naive T cells of lower self-reactivity may be selected by 'private' peptides that have limited expression in the periphery [250]. Additionally, evidence has emerged to suggest that variability in basal levels of self-reactivity among individual cells within the T cell population differentially regulates homeostasis and function.

Gene expression differences and functional biases among individual T cells are reflective of the spectrum of self-reactivity across the population (**Figure 5**). Of note, the naive $CD8^+$ T cell population has a substantially narrower distribution of cell surface CD5 as compared to naive $CD4^+$ T cells possibly a result of 'coreceptor tuning'; unlike the CD4 coreceptor, CD8 levels on individual T cells are adjusted in the periphery to balance the self-reactivity of the TCR [249, 251]. As a result, the underlying heterogeneity in naive $CD4^+$ T cells and the role of self-pMHC reactivity in driving this variation is better defined. There are increases in TCR signaling molecules and negative regulators of T cell activity at the gene and protein level that correlate with heightened basal interactions with self-peptide, the balance of which impacts responsiveness to antigen challenge and, perhaps, limits autoimmune pathology at the precarious threshold between T cell readiness and overt autoreactivity [1, 2, 139, 140, 180, 181]. On average, naive T cells with relatively higher levels of self-reactivity are poised to respond to antigen challenge and preferentially contribute to the immune response in the context of acute infection, though those T cells at the highest end of the spectrum may be refractory to activation [1, 2, 140, 141, 180, 181]. The increased rigor with which $CD4^+$ T cells that receive comparatively stronger signals from basal interactions with self-peptide respond to antigen challenge also limits the response as they are more sensitive to activation-induced cell death [2]. This may explain, in part, why their $CD5^{lo}$

counterparts play a larger role in a chronic infection setting [252, 253], as described in Chapter 3. Interestingly, pre-existing transcriptional networks present within the naive $CD4^+$ T cell population are predictive of some, but not all, cell fate biases [139, 140]. $CD4^+$ T cells with low self-reactivity preferentially produce IFNγ [3, 142, 179], while cells at the high end of the spectrum are prone to differentiate into regulatory T cells (Treg) and T helper 17 (T_H17) cells (**Figure 5**) [139, 140, 179, 183]. T cells that experience stronger tonic TCR signaling express a follicular helper (T_{FH}) cell gene signature [140, 181], and these cells are skewed toward the T_{FH} fate in a lymphoreplete environment [140, 254]; however, there may be exceptions to this general rule [255]. In a broader context, differences in thymic selection that lead to changes in the basal self-reactivity of the naive T cell population at both ends of the life spectrum may play a role in the unique T cell response to antigen challenge in neonates and the elderly population [256, 257].

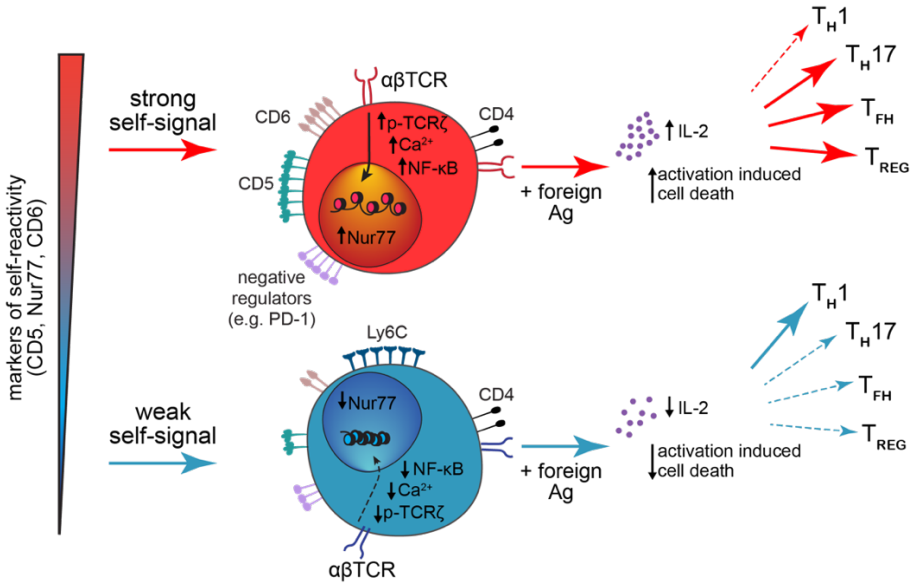


Figure 5. Naive $CD4^+$ T cell heterogeneity. Distinct steady-state transcriptional and chromatin-accessibility states have been described for naive T cells at the high compared to the low end of self-reactivity, as identified by surrogate markers for the strength of self-reactivity, including CD5, CD6, and Nur77. Among naive $CD4^+$ T cells, differences have been described in basal phosphorylation of the TCR z chain, in levels of cytosolic calcium, and in NFkB signalling among others, as well as in the expression levels of negative regulators of TCR signaling. Importantly, these differences impact responses following activation with foreign antigen, including IL-2 production, susceptibility to activation-induced cell death, and the relative contributions to distinct effector lineages. Adapted from [4].

(vi) Conclusion

T cells have a life-long relationship with self-ligands that begins in the thymus and continues in the periphery. The self-signals naive T cells receive in the periphery play several supporting roles: they induce cell division in lymphopenic conditions, promote retention in the SLOs, can act as coagonists, increase T cell sensitivity to foreign antigen by polarizing the TCR on the cell membrane, and provide survival signals (**Figure 6**) [138]. In fact, the importance of self-pMHC in providing survival cues for naive T cells has been appreciated for over 25 years. Yet, we know surprisingly little about the pathways T cells use when receiving self-signals to maintain survival and sensitivity. Conversely, we know even less about the amazing ability of naive T cells (especially $CD4^+$ T cells) to survive for well over a month when self-signals are limiting. *In Chapter 4 we tackle these questions and ask if there are other physiological settings that might impact access to self-ligands.*

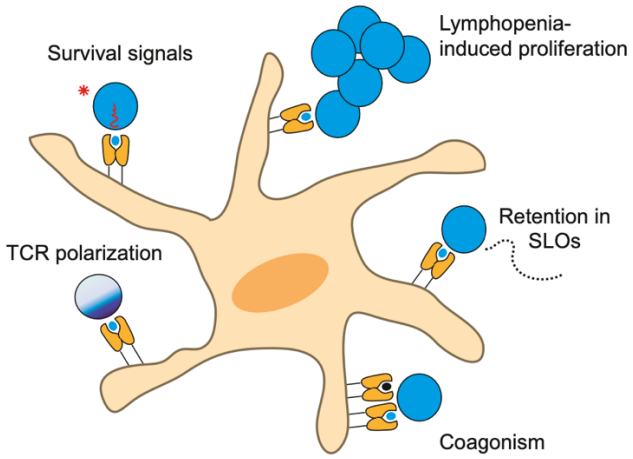
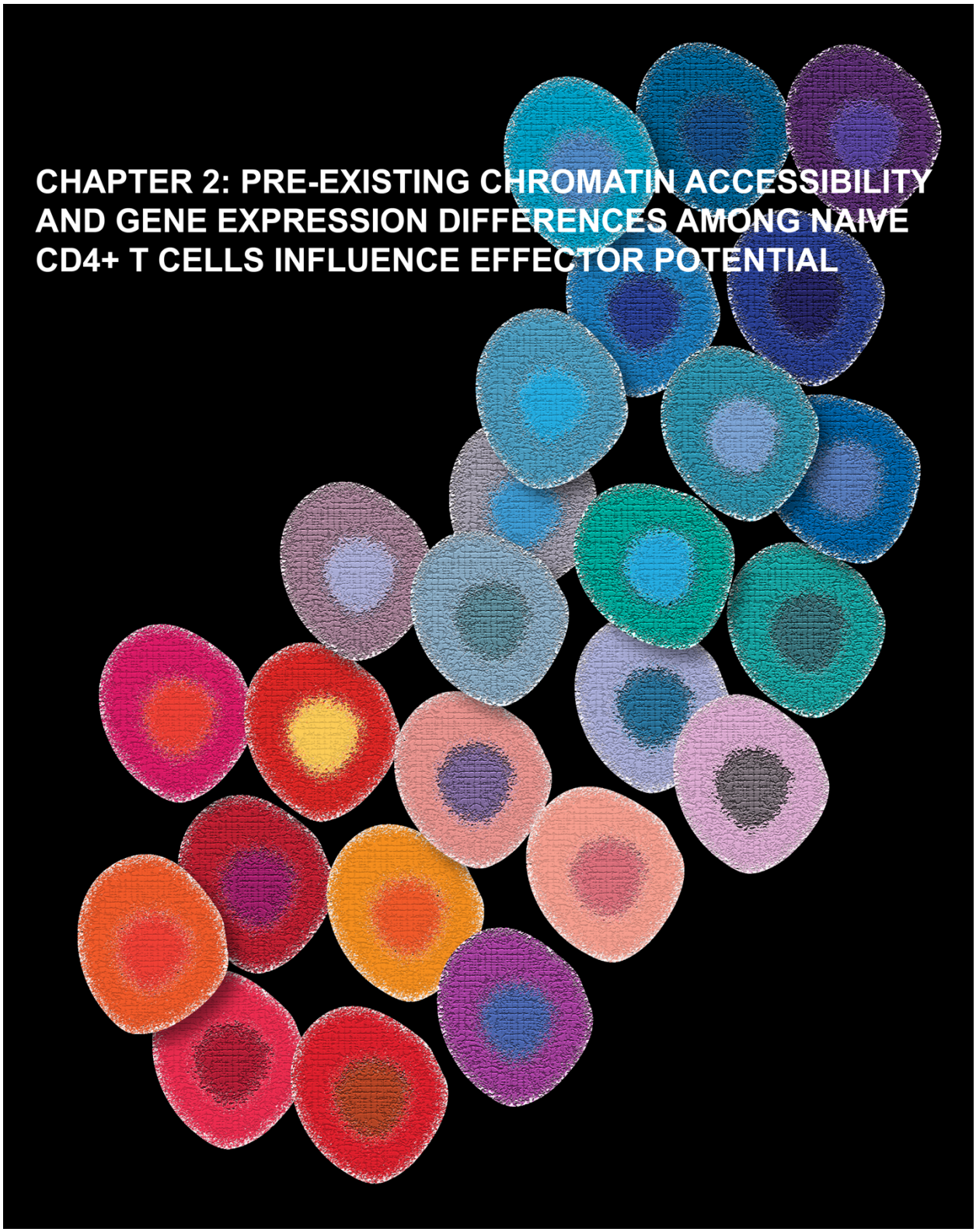


Figure 6. Role of self-signals on peripheral naive T cells. Adapted from [138].

PREFACE TO CHAPTER 2

Once $CD4^+$ T cells encounter their cognate antigen they can differentiate into several effector cell populations. It has historically been thought that prior to foreign antigen encounter, naive $CD4^+$ T cells are a homogeneous population despite the fact that there are differences in the strength that T cells can recognize ligands. The goal of this study was to determine whether differences in strength of self-reactivity pre-wire the naive $CD4^+$ T cell population and bias response to activation and effector lineage decisions. In addition, we ask to what extent transcriptional heterogeneity is dependent on tonic self-ligand interactions.

CHAPTER 2: PRE-EXISTING CHROMATIN ACCESSIBILITY AND GENE EXPRESSION DIFFERENCES AMONG NAIVE CD4+ T CELLS INFLUENCE EFFECTOR POTENTIAL



On the cover:

Naive $CD4^+$ T cells are heterogeneous with respect to their strength of reactivity to self (coded from red to purple). In this chapter, Rogers et al. show that this self-reactivity spectrum is established by pre-wired transcriptional and chromatin states that are partially imprinted during thymic development. Artwork by Judith N. Mandl and Dakota Rogers.

This image was featured as the cover of *Cell Reports* Volume 37 Issue 9 (2021).

Pre-existing chromatin accessibility and gene expression differences among naïve CD4$^+$ T cells influence effector potential

Dakota Rogers[1,2], Aditi Sood[3,4], HanChen Wang[1,2], Jasper J. P. van Beek[6], Thomas J. Rademaker[7], Patricio Artusa[1,2], Caitlin Schneider[2,8], Connie Shen[2,8], Dylan C. Wong[2,8], Aanya Bhagrath[1,2], Marie-Ève Lebel[3], Stephanie A. Condotta[8,†], Martin J. Richer[8,†], Andrew J. Martins[9], John S. Tsang[9], Luis Barreiro[10], Paul Francois[7], David Langlais[2,11,12], Heather J. Melichar[3,5], Johannes Textor[6], and Judith N. Mandl[1,2,8,13*]

[1]Department of Physiology, McGill University, Montreal, Canada.
[2]McGill University Research Centre for Complex Traits, Montreal, Canada.
[3]Immunology-Oncology Unit, Maisonneuve-Rosemont Hospital Research Center, Montreal, Canada.
[4]Department of Microbiology, Immunology, and Infectious Disease, Université de Montréal, Montréal, Canada.
[5]Department of Medicine, Université de Montréal, Montréal, Canada.
[6]Department of Tumor Immunology, Radboud Institute for Molecular Life Sciences, Nijmegen, The Netherlands.
[7]Department of Physics, McGill University, Montreal, Canada.
[8]Department of Microbiology and Immunology, McGill University, Montreal, Canada.
[9]Systems Genomics and Bioinformatics Unit, Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, USA.
[10]Department of Genetic Medicine, University of Chicago, Chicago, USA.
[11]Department of Human Genetics, McGill University, Montreal, Canada.
[12]McGill University Genome Centre, Montreal, QC, Canada.
[13]Lead contact
[†]Current address: Department of Microbiology & Immunology, Indiana University School of Medicine, Indianapolis, USA.

*To whom correspondence should be addressed:
judith.mandl@mcgill.ca
Bellini Building, room 364
3649 Promenade Sir William Osler
Montreal, QC, Canada H3G 0B1

Cell Reports. 2021 Nov 30; 37(9): 110064. Doi: 10.1016/j.celrep.2021.110064.

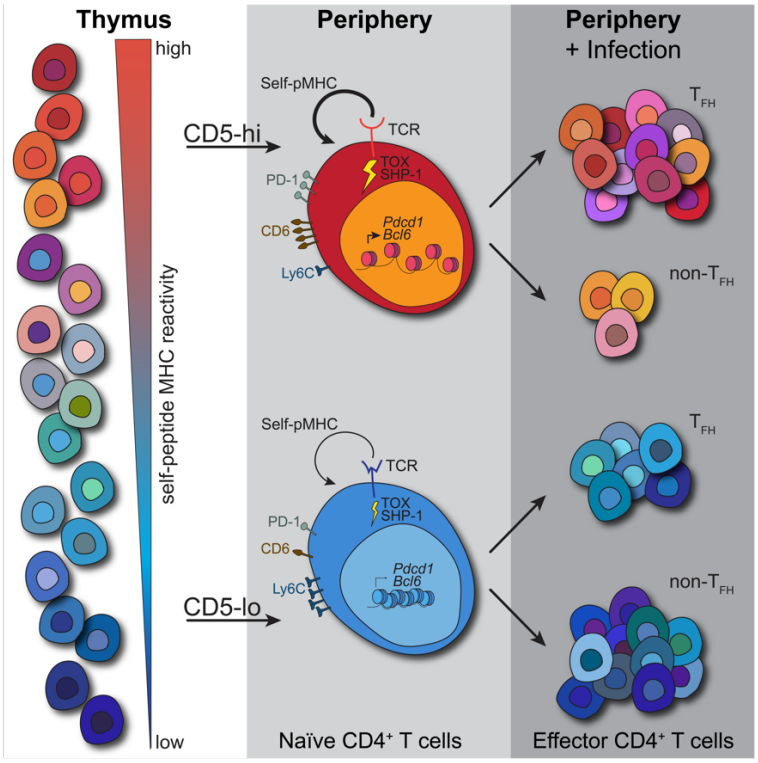
2.1 Summary

CD4$^+$ T cells have a remarkable potential to differentiate into diverse effector lineages following activation. Here, we probe the heterogeneity present among naïve CD4$^+$ T cells before encountering their cognate antigen to ask whether their effector potential is modulated by pre-existing transcriptional and chromatin landscape differences. Single-cell RNA sequencing shows that key drivers of variability are genes involved in T cell receptor (TCR) signaling. Using CD5 expression as a read-out of the strength of tonic TCR interactions with self-peptide MHC, and sorting on the ends of this self-reactivity spectrum, we find that pre-existing transcriptional differences among naïve CD4$^+$ T cells impact follicular helper cell (T_{FH}) versus non-T_{FH} effector lineage choice. Moreover, our data implicate TCR signal strength during thymic development in establishing differences in naïve CD4$^+$ T cell chromatin landscapes that ultimately shape their effector potential.

Keywords

Cell heterogeneity, CD4$^+$ T cells, CD5, T cell differentiation, follicular helper T cells, chromatin landscapes, T cell response, thymus, T cell development

2.2 Graphical Abstract



2.3 Introduction

Heterogeneity is a fundamental property of cellular systems [258, 259]. Even clonally-derived cell populations exhibit variations in gene expression which impact cell fate decisions [260, 261]. Single-cell studies of immune cells have begun to reveal the diversity present in populations thought to be homogeneous and emphasized its role in the immune response [262, 263]. Such heterogeneity is perhaps nowhere as intimately tied to cellular function as it is in adaptive immune cells. T cell populations comprise a breadth of T cell receptors (TCRs), with individual cells expressing unique TCRs generated by somatic recombination [105]. $CD4^+$ T cells, in particular, possess a remarkable capacity for diversification into distinct effector lineages following activation that are defined by the cytokines they make and immune cells they act on. Indeed, $CD4^+$ T cell effector fate is critical in tailoring an immune response to the specific pathogen encountered [264, 265].

$CD4^+$ T cell differentiation relies on dynamic metabolic and transciptional changes [266, 267]. An early $CD4^+$ T cell fate decision is between effector subsets (T helper 1 (T_H1), T_H2, T_H9, and T_H17) and follicular helper T cells (T_{FH}). T_{FH} cells are essential to germinal center (GC) formation and generation of high-affinity plasma and memory B cells. T_{FH} development requires TCR engagement, expression of the lineage-defining transcription factor Bcl6 with inhibition of Blimp1, and upregulation of surface proteins such as PD-1, CXCR5, and ICOS, enabling interaction with B cells [97, 268]. Remarkably, a single $CD4^+$ T cell clone can expand into both T_{FH} and non-T_{FH} cells [89, 269]. $CD4^+$ T cell clones are not equal, however. The propensity of a single $CD4^+$ T cell to differentiate into one helper subset over another varies between cells expressing distinct TCRs [89, 103, 270]. Whether this clonal variablity in cell fate is determined entirely by TCR signal strength upon antigen encounter, or whether naïve T cells are pre-wired with specific effector biases remains incompletely understood.

It is increasingly appreciated that naïve $CD4^+$ T cells already differ prior to antigen stimulation. Despite their quiescence, naïve $CD4^+$ T cells remain ready to rapidly respond to antigen [47, 242, 248]. This is mediated at least in part by sub-threshold interactions with self-peptide presented on MHC (self-pMHC) that provide tonic survival signals [138]. Several markers provide read-outs of self-pMHC signal strength. Ly6C distinguishes naïve $CD4^+$ T cells with high ($Ly6C^–$) from low ($Ly6C^+$) self-pMHC reactivity [139, 179]. The expression levels of two other markers, CD5, a surface protein and negative regulator of TCR signaling, and the orphan nuclear

receptor Nur77 (*Nr4a1*), positively correlate with sub-threshold TCR signal strength [172, 177]. Measures of CD5 expression revealed that self-reactivity of naïve T cells spans a wide spectrum, whereby the upper and lower bounds are likely set in the thymus by positive and negative selection [1, 141]. Even within a monoclonal TCR transgenic (Tg) population, cells may vary with regard to self-pMHC reactivity [180]. Importantly, not only does self-pMHC reactivity impact competition between T cells for homeostatic signals [232], it also influences their responses to antigen. $CD5^{hi}$ naïve $CD4^+$ T cells express higher basal levels of NFκB, phosphorylated TCR-ζ and ERK, make more IL-2 post-activation, preferentially differentiate into regulatory T cells (Tregs), predominate in acute infections, and contribute disproportionately to the memory T cell pool [1, 2, 141, 182, 183]. Similarly, $Ly6C^-$ cells preferentially differentiate into Tregs and T_H17 cells [179]. In contrast, $CD5^{lo}$ naïve $CD4^+$ T cells produce more IFN-γ upon activation [3]. However, it remains unclear whether T cells are pre-programmed entirely by tonic peripheral TCR signals or whether early self-ligand encounters in the thymus play a role.

Here, we adopted an unbiased systems approach, combining single-cell (sc)RNA sequencing (RNA-seq) with bulk RNA-seq, and assay for transposase-accessible chromatin using sequencing (ATAC-seq), to comprehensively investigate the drivers of heterogeneity among naïve $CD4^+$ T cells. We report that individual cell-level biases in the expression of modulators of TCR signal strength, and in T_{FH} versus non-T_{FH} effector lineage choice, are driven, at least in part, by pre-existing transcriptional and epigenetic (chromatin accessibility) differences between cells with low ($CD5^{lo}$) versus high ($CD5^{hi}$) self-pMHC reactivity. Unexpectedly, our data reveal that many of the differences in gene expression among naïve $CD4^+$ T cells do not require continuous signaling through the TCR via self-pMHC interactions but are likely a result of variable signal strengths obtained during development.

2.4 Results

TCR signaling induced gene expression differences are drivers of variability among individual naïve $CD4^+$ T cells

To define transcriptional differences among naïve $CD4^+$ T cells with single cell resolution, we performed scRNA-seq on 1152 cells sorted from the spleen, of which 697 passed quality control (**Table S1**). For each individual T cell, we measured CD5 protein expression during sorting given prior studies implicating CD5 as a key read-out of diversity among naïve T cells [1, 2, 141].

We verified that *Cd5* transcript counts correlated with CD5 protein (**Figure S1A)**, consistent with prior evidence for CD5 transcriptional regulation [271, 272]. In total we detected 14,040 genes; an average of 1,389 genes per cell. Among the top 5% most variable genes, after accounting for technical variation, non-detection drop-outs, and removing mitochondrial genes (**Table S2**) [273], we detected genes important for trafficking between blood and lymphoid organs including *Cd69*, *S1pr1*, *Sell*, *Klf2*, *Itag4*, *Tln1*, and *Foxo1*; genes involved in TCR signaling, such as *Ptpn6*, *Folr4*, *Il7r*, *Cd4*, *Jun*, *Il2rg*, *Thy1*, *Lck*, *Klf6*, *Bcl2*, *Nr4a1*, and also *Cd5*; as well as chromatin modifiers, including *Dnmt1*, *Hdac4*, *Sirt1*, and *Smc4* (**Figure 1A**). Gene ontology (GO) enrichment analyses of the most variable genes showed an enrichment of GO terms associated with TCR signaling, including αβ T cell activation and T cell selection (**Figure 1B**). UMAP analysis of the 2000 most variable genes did not reveal any detectable sub-clustering of naïve cells by CD5 expression (**Figure S1B)**. However, principal component analyses (PCA) based on expression of 55 genes in the GO terms involved in T cell activation from Figure 1B **(Table S3)** showed that the transcriptional state of $CD5^{lo}$ naïve $CD4^+$ T cells differed from that of $CD5^{hi}$ naïve $CD4^+$ T cells, albeit with considerable overlap between the two populations (**Figure 1C**). Overall, our data highlighted the diversity between individual naïve $CD4^+$ T cells that was detectable at the transcript-level and identified expression differences in genes involved in TCR signaling as key drivers of cellular variability.

CD5 expression reveals the existence of distinct gene-expression profiles and chromatin landscapes in naïve $CD4^+$ T cells

To characterize the transcriptional heterogeneity of naïve $CD4^+$ T cells in greater depth, we performed bulk RNA-seq, focusing on the top and bottom 15% of the self-reactivity spectrum defined by CD5 expression. We used $FoxP3^{GFP+}$ reporter mice to exclude Tregs – including them as an 'outgroup' in our initial analyses (**Figure 2A, 2B**, **S2A,** and **S2B**) as Tregs express high levels of CD5 [274]. We confirmed that sorted naïve T cells did not contain Tregs (**Figure S2C and S2D**). Given the detection of chromatin modifiers in our scRNA-seq (**Figure 1A**), we also performed ATAC-seq to investigate the open-chromatin landscape of $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells. PCA highlighted that at the transcriptional level, $CD5^{lo}$ and $CD5^{hi}$ cells clustered into distinct populations (**Figure 2C**). The $CD5^{lo}$ and $CD5^{hi}$ cells were distinguishable in PC2 and separated from Tregs in PC1 (**Figure 2C**). The $CD5^{lo}$ and $CD5^{hi}$ segregation was similarly

reflected by their chromatin accessibility profiles (**Figure 2D**). Overall, $CD5^{lo}$ and $CD5^{hi}$ cells had comparable numbers and genomic annotation of accessible peaks (**Figure S2E** and **S2F**). Importantly, we identified a total of 1,006 differentially expressed genes (DEGs) at FDR ≤ 0.01 between $CD5^{lo}$ and $CD5^{hi}$ cells, of which 90% were found among detected differentially accessible regions (DARs) (**Figure 2E**). Indeed, there was significantly greater correspondence between gene expression and chromatin accessibility among DEGs than for a random gene set (**Figure 2F** and **S2G**). Plotting the fold changes (FC) between $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells from the bulk RNA-seq versus the ATAC-seq emphasized the dataset concordance (**Figure S2H**). Among the DEGs and DARs identified was *Cd5* itself, which had a ~6.7 fold greater transcript expression in $CD5^{hi}$ cells and was less accessible in $CD5^{lo}$ cells **(Figure 2G** and **2H**). Consistent with Nur77 expression reflecting tonic signal strength, *Nr4a1* was among the DEGs (**Figure 2G**) and, as previously described [177], *Nr4a1* expression was greater in Tregs than in $CD5^{hi}$ $CD4^+$ T cells. Like *Cd5*, the *Nr4a1* locus was more open in $CD5^{hi}$ cells (**Figure 2H**). Together, our findings revealed considerable differences both at the transcriptional and the chromatin level between $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells.

Differences in expression of transcriptional regulators and chromatin modifiers may contribute to functional differences among $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells

We next examined the DEGs identified between $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cell populations. More DEGs were upregulated in $CD5^{hi}$ cells and more transcripts were expressed at a FC ≥2 compared to $CD5^{lo}$ cells (**Figure 3A** and **Table S4**). Importantly, we found strong concordance between the bulk and scRNA-seq datasets with regard to specific genes. For instance, *Cd5*, *Folr4*, *Cd6*, *Nr4a1, Tox*, *Ptpn6,* and *Tcf25* were more highly expressed in $CD5^{hi}$ cells, while *Ly6c1* and *Dntt* were higher among $CD5^{lo}$ cells (**Figure 3B**). Indeed, as was described in CD5-sorted naïve $CD8^+$ T cells [141] and in human naïve $CD4^+$ T cells [142], one of the top DEGs was *Dntt* (encoding the DNA polymerase terminal deoxynucleotidyl transferase, TdT, which inserts non-templated nucleotides during V(D)J TCR rearrangement), at a 14x greater abundance in $CD5^{lo}$ cells. Given the greater *Dntt* expression in $CD5^{lo}$ cells in both $CD4^+$ and $CD8^+$ T cells, we asked whether other DEGs were shared by naïve $CD8^+$ T cells sorted on CD5. Interestingly, in line with the narrower CD5 distribution in $CD8^+$ T cells [1], far fewer DEGs with a FC cut-off ≥2 were identified in $CD8^+$ T cells than in our $CD4^+$ dataset and there was little overlap (**Figure S3A**).

Among overlapping DEGs, all but 2 of the 24 DEGs showed expression concordance in $CD4^+$ and $CD8^+$ T cells (**Figure S3B**). Unlike $CD8^+$ T cells, $CD5^{hi}$ naïve $CD4^+$ T cells were smaller and expressed less CD44 compared to $CD5^{lo}$ cells (**Figure S3C**). Overall, there were few parallels between naïve $CD4^+$ and $CD8^+$ T cells sorted on CD5 expression with regard to the specific DEGs.

To investigate patterns in DEGs increased among $CD5^{lo}$ or $CD5^{hi}$ naïve $CD4^+$ T cells, we performed GO analyses. In $CD5^{lo}$ cells we found an enrichment for tumor-mediated immunity and regulation of IFN-γ responses (**Figure 3C**), consistent with work showing that $CD5^{lo}$ $CD4^+$ T cells produce more IFNγ than $CD5^{hi}$ cells upon activation [3]. In contrast, $CD5^{hi}$ cells were enriched for gene networks involved in leukocyte activation, regulation of signaling, and cell migration (**Figure 3C**). In line with this, gene set enrichment analysis (GSEA) indicated that genes associated with $CD4^+$ T cell activation and effector responses [275, 276] were overrepresented in $CD5^{hi}$ cells (**Figure S3D**). Of note, genes involved in the active maintenance of a quiescent state among naïve T cells [47, 241, 277, 278], were detected in both $CD5^{lo}$ and $CD5^{hi}$ cells but not significantly different (**Table S4** and **S5**). Moreover, no cytokines were among the DEGs, except *Il16* which is constitutively expressed in naïve $CD4^+$ T cells [279] and was slightly increased in $CD5^{lo}$ cells (**Table S4** and **S5**). Indeed, the chromatin loci for effector cytokines such as *Ifng*, *Il5*, *Il17a*, and *Il10* were closed; only the transcription start site (TSS) for *Il2* and a few peaks surrounding the TSS for *Il21* had marginally greater accessibility in $CD5^{hi}$ cells (**Figure S3E**). These data indicated that although $CD5^{hi}$ cells were enriched for gene networks associated with activation, both gene expression and accessible chromatin regions were largely consistent with an equally quiescent and non-differentiated cell state among naïve $CD4^+$ T cells with different self-reactivities.

Quiescence exit occurs when naïve T cells are activated by antigen, after which they undergo chromatin remodeling, transcriptional changes, and ultimately effector differentiation. To understand whether naïve $CD4^+$ T cells were poised to respond differently to activation as a function of our identified transcriptome and chromatin accessibility differences, we first investigated whether they differed in expression of transcriptional regulators (TRs). We used a predefined list of 1680 known or putative TRs [280] and detected 31 TRs upregulated in $CD5^{hi}$ cells involved in T cell proliferation or survival (*Atf6*, *Myb*, and *Bcl3*), and T cell activation or differentiation (*Egr1*, *Egr2*, *Egr3*, *Nfatc3*, *Ikfz3*, *Ikzf4*, *Tox*, *Tox2*, *Nr4a1*, *Nr4a3*, *Klf9*, *Lef1*, *Bcl6*, *Eomes*, *Irf6*, *Id2*, and *Id3*) (**Figure 3D**). We also detected 21 TRs that mediate chromatin modifications such as acetylation (*Hdac5*, *Hdac9*, and *Etv6*) and methylation (*Dmnt3a*, *Klf10*,

Kdm2b, and *Gfi1*) (**Figure 3D**), as well as other TRs (**Figure S3F**). Twenty TRs were enriched in $CD5^{lo}$ cells including repressors (*Hdac7, Nr1d1, Prdm1, Rara,* and *Trps1*) (**Figure 3D** and **S3F**).

We next asked if there were TR binding motifs enriched among chromatin peaks that were unique to either $CD5^{lo}$ or $CD5^{hi}$ cells. Indeed, $CD5^{hi}$ cells were enriched in binding motifs for TR networks downstream of TCR activation, including AP-1 and JNK transcription factors, FOS, FOSL2, AP-1, JUNB, as well as NUR77, NFAT, and NFkB (**Figure 3E** and **S3G**). Conversely, $CD5^{lo}$ cells were enriched in binding motifs for IRF-4 and PRDM1 (BLIMP-1), both of which promote non-T_{FH} effector differentiation [281, 282]. Notably, while not represented in TR binding motif analysis, TOX and TOX2 were among identified upregulated TRs in $CD5^{hi}$ cells, which function downstream of the TCR through NFAT signaling [283-285] and promote T_{FH} differentiation by increasing Bcl6 expression through enhanced chromatin accessibility of TCF-1 and LEF-1 bound regions of the *Bcl6* locus [286]. Thus, we asked whether TOX, LEF-1, and TCF-1 expression was greater at the protein level in $CD5^{hi}$ cells. While the detected differences were small (1.3-1.4 fold) they were robust across mice (**Figure 3F**), were not observed in unstained controls (**Figure S3H**), and corresponded to increased chromatin accessibility in the loci for *Tox*, *Tox2*, *Lef1*, and *Tcf7* in $CD5^{hi}$ cells (**Figure S3I**).

Together, our data suggested the possibility that a network of TRs and unique chromatin accessibility profiles results in differences in cell states among naïve $CD5^{lo}$ and $CD5^{hi}$ $CD4^{+}$ T cells, impacting their function and/or differentiation upon activation, particularly with regard to the early T_{FH} vs. non-T_{FH} bifurcation.

Pre-existing expression differences in regulators of TCR signaling among naïve $CD4^{+}$ T cells are maintained after activation

To investigate the cell signaling and lymphocyte activation signatures in $CD5^{hi}$ $CD4^{+}$ T cells, we curated a list of DEGs involved in regulating T cell activation. $CD5^{hi}$ cells had increased expression of genes involved in co-stimulation, such as *Icos*, *Rankl*, *Itgb2*, and *Gitr* (**Figure 4A**). More predominantly, $CD5^{hi}$ cells had a higher expression of genes involved in the negative regulation of TCR signaling or T cell activation, including *Cd6*, *Nt5e* (CD73), *Ptpn6* (SHP-1), *Ctla4*, *Pdcd1*, *Btla, IL10ra, P2rx7, Nrp1, Nrp2, Cd200,* and *Adora2a* (**Figure 4A**). Given the role of negative regulators in T cell exhaustion during chronic antigen stimulation, and perhaps indicative of a greater frequency and/or strength of self-pMHC signals obtained by $CD5^{hi}$ T cells,

GSEA identified an enrichment in exhaustion-associated genes among $CD5^{hi}$ cells (**Figure S4A**). Corroborating our RNA-seq, we observed expression differences between $CD5^{lo}$ and $CD5^{hi}$ cells at the protein level for a subset of the regulators of TCR signaling including GITR, LFA-1, CD6, FolR4, PD-1, and CD73 (**Figure 4B** and **4C**). Of note, FCs in CD6, Ly6C, FolR4, and CD73 between $CD5^{lo}$ and $CD5^{hi}$ cells were detectable in neonates and remained stable with age (**Figure S4B** and **S4C**). Interestingly, signaling through LFA-1 was recently shown to promote Bcl6 expression and be required for T_{FH} differentiation [287], and some of the other genes modulating the TCR signal were also shown to play a role in T_{FH} differentiation, including *Tox*, *Tox2*, *FolR4*, and *Icos*, all of which are expressed at greater levels in $CD5^{hi}$ cells (**Figure 4A**). In addition, we confirmed greater protein expression in $CD5^{hi}$ cells of SHP-1, a negative regulator of TCR signaling that modulates T cell antigen sensitivity [158, 243], associates with CD5, CD6 [174, 288], and other negative regulators of TCR signaling such as PD-1, BTLA, and CTLA-4 [289] (**Figure 4B** and **4C**). Further, $CD5^{lo}$ cells expressed higher levels of Ly6C. The non-significant DEGs CD4, CD98 (LAT1), and CD45 were used as controls to establish a cut-off for biologically significant protein FCs (**Figure 4B** and **4C**). Thus, at least for the subset of tested DEGs, transcriptional differences among naïve $CD4^+$ cells correlated with differences in protein expression.

We next asked whether the differential expression of TCR signal regulators was maintained following TCR stimulation. We found that sorted $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells stimulated with αCD3/CD28, which bypasses individual TCR affinity for their agonist peptides, retained expression differences 24 hours after stimulation, albeit for some proteins at reduced FCs (**Figure 4D** and **S4D**). In unsorted $CD4^+$ T cells, SHP-1 and CD6 expression increased with greater CD5 levels, and this relationship was maintained following activation (**Figure 4E**). In contrast, Ly6C had a bimodal distribution, with most cells becoming Ly6C-negative after activation (**Figure 4D and 4E**). In addition, the transcription factor TOX, which promotes expression of exhaustion-associated genes [284, 285], followed a similar expression pattern as SHP-1 across the full spectrum of CD5, but its expression range was reduced post-activation (**Figure S4E**). Mutual information analysis confirmed that CD5 was a strong predictor of both CD6 and SHP-1 expression pre- and post-activation, whereas TCRβ only became predictive post-activation for SHP-1 and was a poor predictor of CD6 (**Figure S4F** and **S4G**).

Cd5 and *Cd6* are gene homologs located on the same chromosome and have functional similarities [290, 291]. Our data indicated a tight correlation between CD5 and CD6 levels, with CD5 expression being a robust predictor of CD6 expression both pre- and post-activation (**Figure 4E** and **S4G**). Interestingly, the FC difference in CD6 expression on sorted $CD5^{lo}$ and $CD5^{hi}$ $CD4^{+}$ T cells remained slightly greater post-activation than CD5, suggesting that CD6 might more reliably read out self-pMHC reactivity after activation (**Figure 4F**). The maintenance of differences among naïve $CD4^{+}$ T cells in the expression of regulators of TCR signaling even following a strong TCR stimulus raised the possibility that these might be regulated through epigenetic modifications rather than modulated entirely by TCR signal strength. Indeed, akin to the differences in chromatin accessibility of the *Cd5* locus (**Figure 2H)**, the *Cd6* and *Ptpn6* loci were more accessible in $CD5^{hi}$ cells (**Figure 4G**). In summary, we corroborated expression differences between $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^{+}$ T cells at the protein-level for genes important in tuning TCR signal strength and showed that these differences were maintained even after strong TCR stimulation.

$CD5^{hi}$ cells have a greater propensity to develop into T_{FH} cells than $CD5^{lo}$ naïve $CD4^{+}$ T cells

We next investigated whether the observed transcriptional and chromatin accessibility differences alter T cell differentiation. GSEA of our RNA-seq indicated that $CD5^{hi}$ cells were enriched for genes expressed in natural Tregs (**Figure S5A**). To determine whether $CD5^{hi}$ cells were more likely to differentiate into Tregs *in vivo* than $CD5^{lo}$ cells, we adoptively transferred Treg-depleted $CD4^{+}$ T cells into $TCRb^{-/-}$ mice in a model of inflammatory bowel disease [292, 293]. Consistent with prior *in vitro* Treg polarization experiments [183] and studies of $Ly6C^{-}$ naïve $CD4^{+}$ T cells [179], $TCRb^{-/-}$ mice given $CD5^{hi}$ cells subsequently had an increased survival time likely due to increased frequency of Tregs, compared to mice that received $CD5^{lo}$ cells (**Figure S5B** and **S5C**).

Interestingly, GSEA also showed that $CD5^{hi}$ cells were enriched for T_{FH} gene signatures (**Figure 5A**). Across all replicates, $CD5^{hi}$ cells expressed higher levels of *Pdcd1*, *Cxcr5*, and *Bcl6*, while $CD5^{lo}$ cells expressed higher levels of the T_{FH} repressor *Prdm1* (Blimp-1) (**Figure 5B**). Moreover, the *Pdcd1*, *Cxcr5*, and *Bcl6* loci were more accessible in $CD5^{hi}$ cells (**Figure 5C**). These data suggested the possibility of a pre-existing disposition among $CD5^{hi}$ cells to become T_{FH} cells relative to their $CD5^{lo}$ counterparts. In support of this hypothesis, $CD5^{hi}$ naïve $CD4^{+}$ T cells were

shown to produce more IL-2 than $CD5^{lo}$ cells post-stimulation [2], and data has highlighted the importance of early IL-2 production in T_{FH} lineage choice, with IL-2 producers becoming T_{FH} cells and paracrine IL-2 signaling reinforcing non-T_{FH} lineage commitment of $CD4^+$ T cells obtaining weaker TCR signals [92, 254].

To directly assess whether $CD5^{hi}$ naïve $CD4^+$ T cells gave rise to a greater proportion of T_{FH} cells *in vivo* than $CD5^{lo}$ cells, we first sorted 15% $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells, adoptively transferred ~$1x10^7$ cells of each into separate recipients that were infected with LCMV. An estimated 8 LCMV-GP66 specific $CD4^+$ T cells are present per 10^6 naïve $CD4^+$ T cells [31, 294], and so it was not surprising that the proportion of activated cells post infection (day 8) was highly stochastic among the transferred cells. However, as in our *in vitro* assays, the CD5 expression difference between transferred $CD5^{lo}$ and $CD5^{hi}$ $CD4^+$ T cells was maintained following activation (**Figure 5D** and **5E**). As an alternative approach, we asked whether we would be better able to assess differences in T_{FH} differentiation potential by transferring sorted populations into $TCR\beta^{-/-}$ mice. Given the lack of competitor T cells in these recipients, we first investigated whether this would impact T_{FH} lineage choice in $CD4^+$ T cells with a fixed TCR. We performed adoptive transfers of LCMV-specific TCR Tg SMARTA $CD4^+$ T cells into both wild-type (WT) and $TCR\beta^{-/-}$ recipients, infected them 1 day later with LCMV and then assessed the response 8 days post infection (**Figure S5D**). Overall, the clonal expansion of SMARTA TCR Tg cells was greater in the $TCR\beta^{-/-}$ mice (**Figure S5E**). Interestingly, we found that when SMARTA TCR Tg cells were transferred into $TCR\beta^{-/-}$ recipients, CD5 surface expression was increased compared to cells transferred into WT mice (**Figure S5F**), suggesting that SMARTA Tg cells were receiving stronger TCR signals upon antigen encounter in the absence of other competitor T cells, as has been previously shown in lymphopenia-induced expansion [232]. Importantly, the percent of T_{FH} cells among transferred SMARTA TCR Tg cells was halved in the infected $TCR\beta^{-/-}$ compared to WT recipients (**Figure S5G**). Given the role of IL-2 in T_{FH} differentiation and the modulation of IL-2 production by TCR signal strength [254], we postulated that the decreased T_{FH} differentiation of SMARTA TCR Tg cells in $TCR\beta^{-/-}$ mice was a result of greater IL-2 production mediated by greater TCR signaling. Indeed, we observed that SMARTA TCR Tg cells produced more IL-2 and had greater CD25 (IL-2Rα) surface expression in the infected $TCR\beta^{-/-}$ mice (**Figure S5H** and **S5I**). Due to the impact of the lack of other T cells in the $TCR\beta^{-/-}$ mice on T_{FH} frequency

post-infection we concluded that this experimental approach would not accurately identify T_{FH} potential differences between $CD5^{lo}$ and $CD5^{hi}$ polyclonal naïve $CD4^+$ T cells.

Instead, based on the observation that CD5 expression remained detectably different after transfer, we quantified T_{FH} differentiation among activated endogenous $CD5^{lo}$ and $CD5^{hi}$ $CD4^+$ T cells post-LCMV infection. We found ~2 fold increased T_{FH} differentiation within the 15% $CD5^{hi}$ activated $CD4^+$ T cells compared to $CD5^{lo}$ cells (**Figure 5F**, **5G**, and **S5J**). Moreover, among $CD5^{hi}$ cells, T_{FH} ($PD\text{-}1^{hi}$ $CXCR5^+$) and $PD\text{-}1^{hi}$ $CXCR5^-$ cells expressed greater levels of PD-1 than their $CD5^{lo}$ counterparts (**Figure S5K**). T_{FH} cells were also overrepresented among $CD6^{hi}$ $CD4^+$ T cells and had increased PD-1 expression over their $CD6^{lo}$ counterparts (**Figure 5F-G** and **S5J-K**). In line with an increased T_{FH} population from $CD5^{hi}$ cells, these T_{FH} cells expressed higher TOX, LEF-1, and TCF-1 (**Figure S5L**). Thus, our data suggested that T_{FH} and non-T_{FH} cell fate decision was altered by pre-existing differences present in naïve $CD4^+$ T cells prior to foreign antigen encounter.

Removing naïve $CD4^+$ T cells from continuous self-pMHC interactions reveals gene expression differences that are independent of post-thymic self-ligand recognition

Continuous tonic self-pMHC signals obtained by naïve T cells in the periphery facilitate antigen recognition by maintaining partial TCRζ-chain phosphorylation, and through polarization of the TCR and its signaling components [248]. Interrupting signals from self-pMHC interactions for only 30 minutes leads to a loss of antigen sensitivity [248]. It is unknown whether described differences between naïve $CD4^+$ T cells of low and high self-reactivity are similarly dependent on subthreshold tonic self signals. An alternative hypothesis is that naïve T cells are pre-wired by TCR signals obtained during selection in the thymus, and that these differences are then epigenetically imprinted. Given our findings that some modulators of TCR signal strength remained distinct post-activation, we examined which DEGs between $CD5^{lo}$ and $CD5^{hi}$ cells were dependent on continuous self-pMHC interactions, and which might be a result of thymic imprinting.

To investigate this, we sorted 15% $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells and cultured them *ex vivo* in the absence of self-pMHC for 22 hours with IL-7 and performed bulk RNA-seq. After resting, sorted cells segregated into clusters distinct from their freshly isolated counterparts along PC1, while differences between $CD5^{lo}$ and $CD5^{hi}$ cells were preserved in PC2 (**Figure 6A**). As

described, levels of CD5 protein and transcript rapidly decreased upon resting (**Figure 6B**) [44, 173]. Interestingly, however, CD5 mRNA and protein expression remained distinct in rested $CD5^{lo}$ compared to $CD5^{hi}$ cells (**Figure 6C**). Therefore, while CD5 expression levels are maintained by peripheral self-interactions, our data suggested that the retention of CD5 expression differences between $CD5^{lo}$ and $CD5^{hi}$ cells were independent of tonic self-signals.

To examine whether other genes in our identified $CD5^{lo}$ versus $CD5^{hi}$ DEGs followed a similar pattern to *Cd5*, we compared expression levels in fresh and rested cells and designed classification criteria to subset genes into 2 groups: DEG-ND, genes where expression differences between CD5-sorted populations were not dependent on tonic self-pMHC interactions; DEG-D, genes where differences were lost upon resting (i.e. expression differences were dependent on tonic self-pMHC interactions) (**Figure 6D**). A gene was classified as being DEG-ND if it was a DEG (FDR≤0.01) in the fresh comparison and remained a DEG (FDR≤0.01) after resting. Conversely, to be classified as DEG-D, the gene was a DEG (FDR≤0.01) in the fresh comparison but became non-significant (FDR≥0.3) post resting. With these criteria, 513 DEGs were DEG-ND, and 212 DEG-D, while 281 DEGs did not fall into either group (**Figure 6D, Table S4** and **S5**). Plotting the $CD5^{lo}$ versus $CD5^{hi}$ FC of the fresh versus the rested confirmed that DEG-ND were unaffected by the absence of self-pMHC, while DEG-D converged to zero after resting (**Figure 6E**). Thus, transcriptional heterogeneity among the naïve $CD4^{+}$ T cell population was explained both by differences which depended on continuous self-pMHC interactions and which were independent of tonic self-pMHC signals. Interestingly, most of the negative regulators of TCR signaling and *Dntt* were DEG-ND, consistent with their expression level being set during thymic development (**Figure 6E**). In addition, more than half of the T_{FH}-associated genes were DEG-ND, including *Icos*, *Itgb2*, *P2rx7*, and *Adora2a*, whereas 20% of T_{FH} associated genes were DEG-D, including *Tox2* and *Dusp6*. Of note, the TRs *Jun, Egr1, Egr3, Ikzf3*, and *Nfatc3* were classified as DEG-D, while *Stat1, Eomes, Ikzf2, Id2, Sox4,* and *Nr4a1* were DEG-ND, indicating that even at the level of regulation of TRs some expression differences rely on continuous self-ligand interactions while others do not (**Figure 6E**). To verify whether the maintenance of differences among transcripts following resting was also observed at the protein level, we chose a subset involved in TCR signalling (CD6, CD73, FolR4, Ly6C, PD-1, and SHP-1) from the DEG-ND group and measured expression in sorted $CD5^{lo}$ versus $CD5^{hi}$ cells rested for 5 days to give protein levels time to turn over. We observed no loss in cell viability during this period (**Figure S6A**),

and, while protein levels did change upon resting, the differences between $CD5^{lo}$ and $CD5^{hi}$ cells were retained (**Figure 6F, Figure S6B**).

Our classification of genes into two distinct groups based on their reliance on tonic self-pMHC interactions suggested that at least a subset of the described transcriptionally-wired heterogeneity among naïve $CD4^+$ T cells may reflect epigenetic differences established in the thymus. To test this hypothesis, we probed our ATAC-seq dataset to ask whether there were detectable differences in chromatin states between the genes in the two groups. We predicted that loci of DEG-ND genes would show a greater difference in accessible regions between $CD5^{lo}$ and $CD5^{hi}$ cells than would the loci of DEG-D genes. Indeed, on average, the fold difference in chromatin accessibility was greater in the DEG-ND compared to DEG-D group (**Figure 6G** and **6H)**. In line with this, gene expression difference between $CD5^{lo}$ and $CD5^{hi}$ cells were greater in the DEG-ND set than the DEG-D set (**Figure 6I**). Collectively these data suggested that there were two sources of heterogeneity among naïve $CD4^+$ T cells which impacted their responsiveness to foreign antigen and their effector differentiation: differences that required continuous self-pMHC interactions in the periphery, and differences that did not require self-pMHC interactions that may be imprinted in the thymus.

2.5 Discussion

$CD4^+$ T cells play a critical role in orchestrating an immune response, with early fate decisions between T_{FH} and non-T_{FH} effector subset lineage decisions thought to be primarily determined by TCR engagement with cognate pMHC during priming [97]. Here we showed that there are transcriptional and open-chromatin differences between $CD4^+$ T cells that are present prior to their activation, maintained after activation in the short-term, and impact TCR signal strength and early lineage choice upon antigen encounter. At the single cell level, our data highlight that heterogeneity among naïve $CD4^+$ T cells is mediated by many interacting genes. Importantly, while we show that CD5 expression alone does not establish discrete clusters among naïve $CD4^+$ T cells, other work has defined subpopulations which change with infectious challenge [241, 295]. Here, we identified expression in modulators of TCR signaling, chromatin modifiers, and steady-state T cell trafficking genes as key drivers of between-cell variability. Our data suggest that CD5 expression is a better predictor of cellular behaviour at the population-level than at the single cell level. Population averages of sorted polyclonal naïve $CD4^+$ T cells are therefore not always

mirrored when studying the behaviour of a few specific TCR clonotypes, explaining some of the contradictory results with regard to T_{FH}-lineage differentiation biases described based on specific pairs of TCR Tg clones [2, 255].

Our findings expand on previous work implicating self-pMHC reactivity in the effector potential of $CD4^+$ T cells [3, 179, 183]. Our work implicates self-reactivity in early T_{FH} cell lineage bifurcation. Previous work showed that T cells obtaining stronger TCR signals during activation become IL-2 producing cells and signal in *trans* to IL-2-non-producing cells to reinforce non-T_{FH} effector differentiation [92, 254]. We found that increasing the strength of self-pMHC signals obtained by SMARTA TCR Tg $CD4^+$ T cells transferred into T cell deficient mice led to a decrease in T_{FH} differentiation due to enhanced IL-2 production. Thus, removing competitor T cells, by modulating both IL-2 and TCR signal strength, led to the opposite outcome with regard to T_{FH} differentiation predicted by TCR signal strength alone – in line with recent observations that $Nur77^{lo}$ $CD4^+$ T cells adoptively transferred into $TCRa^{-/-}$ gave rise to a greater T_{FH} frequency than transferred $Nur77^{hi}$ cells [255]. A recent study has corroborated the use of CD5 as a marker for the self-ligand reactivity in human T cells [142], but it will be important to investigate whether the relationships described between self-reactivity and effector potential hold.

While we implicate thymically-imprinted epigenetic differences in the transcriptional heterogeneity among naïve $CD4^+$ T cells, we did not address here whether such signals are impacted at different times during development. It is increasingly appreciated that T cell development is a layered process, with neonatally-derived T cells being distinct from adult-derived T cells with regard to responsiveness to stimulation and effector potential [296]. Moreover, in both mice and humans, CD5 expression on naïve $CD4^+$ T cells is lower in adults than in neonates [1, 256]. While we showed that differences between $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells remained stable between neonates and older mice, it will be important to determine if T cell developmental origins (fetal versus adult) are a contributing factor in their diversity.

Whether differences among naïve $CD4^+$ T cells can ultimately be related back to features of their specific TCRs remains an open question. It is intriguing that in $CD5^{lo}$ naïve $CD4^+$ T cells one of the top DEGs was *Dntt* (encoding TdT), as was also observed in mouse naïve $CD8^+$ T cells and human naïve $CD4^+$ T cells [141, 142]. TdT is responsible for adding n-nucleotides during V(D)J recombination and thus diversifying the TCR repertoire [297, 298]. It has been proposed that $CD5^{hi}$ T cells have a greater proportion of germline TCRs (lacking n-nucleotides) and were

evolutionarily optimized to strongly bind to pMHC [138]. It is possible that differences in *Dntt* expression play a role in dictating TCR sequence length and self-pMHC reactivity. Ultimately, patterns in TCR sequences may exist that enable some prediction of self-pMHC reactivity and, therefore, the differentiation potential of individual T cell clones.

Together, our data shed light on which pre-existing transcriptome-level differences among naïve $CD4^+$ T cells are accessible to interventions targeting self-pMHC peripheral interactions, compared to others that would require modulation at the chromatin level, which may aid in optimizing protocols for enhancing desirable T cell functions in clinical settings such as in adoptive cell therapies [299, 300].

Limitations of the study

Our data imply that naïve $CD4^+$ T cell heterogeneity is partly thymically imprinted and retained independent of interactions with self-pMHC in the periphery. Thus, our work reconciles prior studies that have described specific heterogeneous traits among naïve $CD4^+$ T cells, not all of which were dependent on tonic TCR signals [1, 2, 182, 248]. One limitation of our work is that not all DEGs identified could be classified into DEG-ND or DEG-D groups and, given the lack of commercially available antibodies for all genes of interest, only some of our classified genes were validated at the protein level. Further, future work will be needed to more clearly define the link between expression differences in chromatin modifiers, epigenetic changes in individual T cells, and the interactions with self-pMHC made in the thymus.

2.6 Acknowledgements

We would like to thank G. Perreault, P. D'Arcy, and the animal facility staff at McGill for their excellent care of our animal colony. We are grateful to R. Germain (NIH), T. Baldwin (U Alberta), N. Singh (U Maryland), N. Vrisekoop (Utrecht University), and Mandl lab members for critical feedback and comments on the manuscript. We thank C. Stegen and J. Leconte at the Cell Vision Core Facility and the core facility at the Institut de Recherches Cliniques de Montréal. We thank S. Lesage (U de Montréal) for sharing RAG2^{GFP} mice. Computational analyses were enabled in part by support provided by Calcul Québec and Compute Canada. D.R. was funded by a Frederick Banting and Charles Best Canada Graduate Doctoral Award (CIHR) and a Tomlinson Doctoral Fellowship (McGill). A.S. was funded by a Cole Foundation Postdoctoral Fellowship. J.J.P.vB was funded by a NOW Vici grant (016.140.655). D.L. and H.J.M. are junior 1 and junior 2 scholars of the FRQS, respectively. J.N.M. is a Canada Research Chair for Immune Cell Dynamics. This research was supported by a NSERC Discovery Grant (2016-03808) and McGill start-up fund to J.N.M., and a Cancer Research Society (22344) and CIHR (PJT-168862) grant jointly to J.N.M and H.J.M.

Author Contributions

J.N.M. conceived and supervised the project and research, with input from H.J.M. and J.T. D.R. designed and performed most experiments with input from C.Sc., C.Sh., D.W., P.A. and A.B. Bulk RNA-seq was performed by J.N.M, with help from A.J.M. and J.S.T., and D.R. analyzed the data. D.R., A.S. and M.L. generated the ATAC-seq dataset with input from H.J.M., and data was analysed by D.R. and H.W. The scRNA-seq data was generated and analysed by J.vB. and J.T. with input from D.R. Mutual information was determined by T.R. and P.F. Critical reagents or intellectual input were provided by S.A.C., M.J.R., L.B., P.F., and D.L. The manuscript was written by J.N.M. and D.R., with critical input from H.J.M and J.T., and feedback from all authors.

Declaration of Interests

The authors declare no competing interests.

Inclusion and Diversity Statement

We worked to ensure sex balance in the selection of non-human subjects. One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science. One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community. While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

2.7 Figures

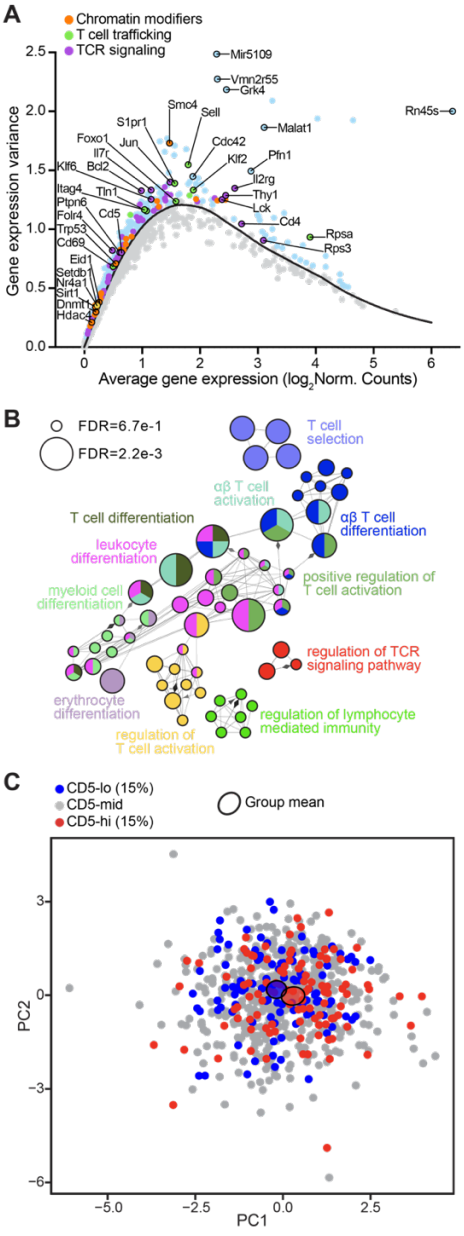


Figure 1. Cellular variability among naïve CD4^{++} T cells at the single cell level is driven by T cell receptor signaling gene expression. (**A**) Plot of top 5% most variable genes (blue) across individual naïve CD4^{++} T cells (sorted on TCRβ^{+}, CD4^{++}, CD8^{+-}, CD25^{-}, CD44^{lo}, CD62L^{hi}). Select genes involved in TCR signaling (purple), T cell lymphoid organ trafficking (green), and chromatin modification (orange) are labelled. (**B**) Immune system process GO enrichment analyses of top 5% variable genes from A. Circles correspond to unique GO ontology groups with related groups coded in same color. Circle size reflects enrichment significance (FDR cut-offs

shown). (**C**) PCA projection of scRNA-seq profiles based on genes that were in the top 5% most variable genes and also in GO terms involved in T cell activation from B. The surface protein expression of CD5 is overlaid with the 15% ^{CD5lo} (blue) and 15% ^{CD5hi} (red) naïve $CD4^{++}$ T cells. Shaded colored circles represent group means. Each data point represents a single cell. See also Figure S1.

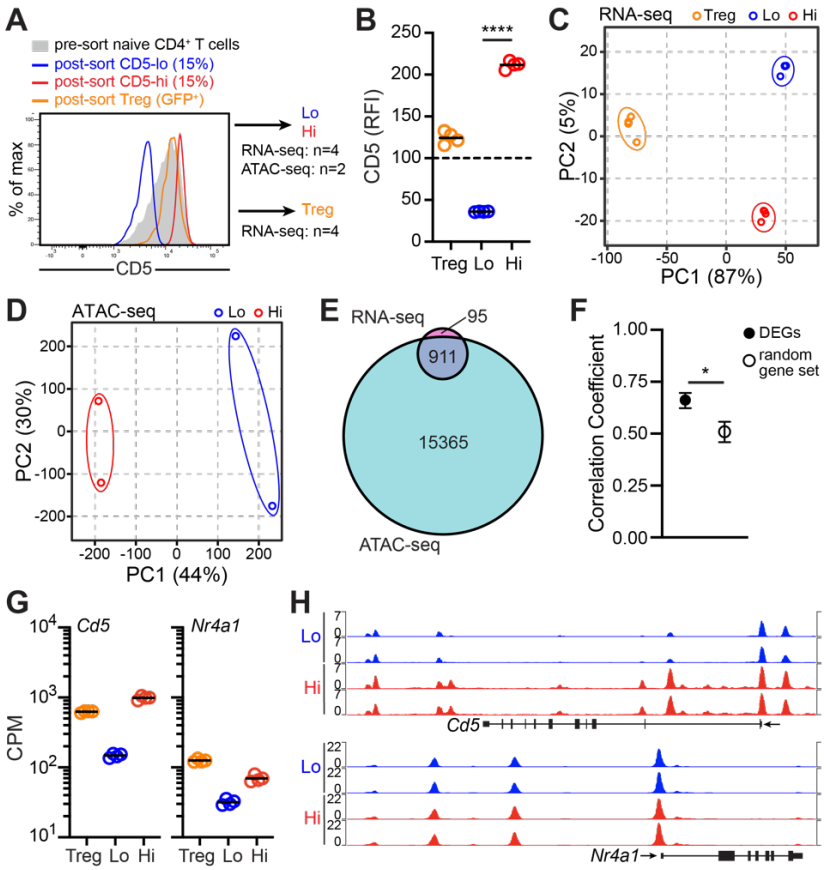


Figure 2. Sorting on CD5 expression reveals that naïve CD4^{++} T cells encompass distinct gene-expression profiles and chromatin landscapes. (**A**) Sorted ^{CD5lo}, ^{CD5hi} naïve CD4^{++} T cells, and Tregs for RNA- and ATAC-seq. (**B**) Relative fluorescence intensity (RFI) of surface CD5 expression on sorted populations from A, relative to pre-sorted naïve CD4^{++} T cells. Lines represent group means (n=4). (**C,D**) PCA of populations in A, for RNA-seq (n=4) (C), and ATAC-seq (n=2) (D). Percents in parentheses indicate the variation described by each PC. I Venn diagram illustrating overlap in number of DEGs (FDR≤0.01) identified by RNA-seq and DARs identified by ATAC-seq in ^{CD5lo} versus ^{CD5hi} cells. (**F**) Correlation between ^{CD5hi} versus ^{CD5lo} FC differences in identified DEGs and corresponding DARs as compared to the correlation with a random gene set. (**G**) *Cd5* and *Nr4a1* mRNA expression (RNA-seq). All group comparisons are significant at FDR<0.01. Lines represent group means (n=4). (**H**) ATAC-seq signal profiles across *Cd5* and *Nr4a1* gene loci from 2 biological replicates of sorted samples.

Statistics: Significant DEGs had FDR<0.01 (E, G), correlation coefficient with 95% confidence intervals (F). See also Figure S2.

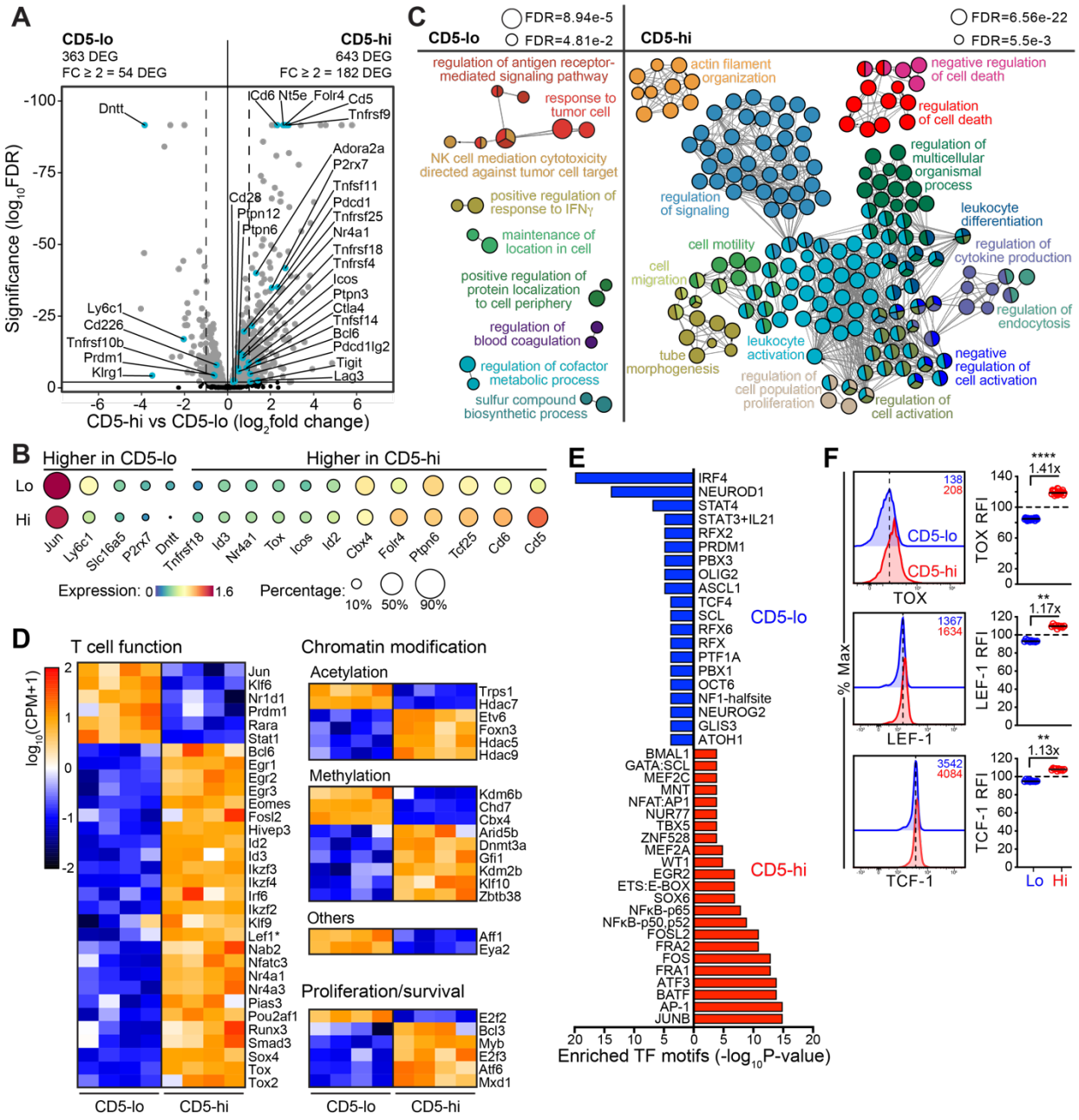


Figure 3. Transcriptional diversity among naïve CD4^{++} T cells suggest functional differences through transcriptional regulators activity and chromatin modifiers. (**A**) Volcano plot of DEGs identified by RNA-seq comparison of ^{CD5lo} and ^{CD5hi} naïve CD4^{++} T cells. Positive FC values indicate increased expression in ^{CD5hi} naïve CD4^{++} T cells. Significant DEGs (FDR<0.01) are indicated in grey and a subset of DEGs are labeled (blue). Dotted lines are drawn at FC = 2. (**B**) scRNA-seq gene expression level (colour) and frequency of cells with detectable expression (circle size) of select DEGs. (**C**) GO enrichment analysis for genes upregulated in ^{CD5lo} (left) or ^{CD5hi}

(right) cells. Circles correspond to unique GO ontology groups; related groups coded in same color. Circle size reflects enrichment significance (FDR cut-offs shown). (**D**) Heatmap of all differentially expressed TRs between ^{CD5lo} and ^{CD5hi} naïve $CD4^{++}$ T cells (n=4), grouped by function. I Significant (p-value $\leq 10^{-4}$) unique enriched TF motifs in DARs from ATAC-seq for ^{CD5lo} (blue) and ^{CD5hi} (red) cells. (**F**) Protein expression TOX, LEF-1, and TCF-1 in ^{CD5lo} and ^{CD5hi} cells (RFI is relative to total naïve $CD4^{++}$ T cells). Representative flow cytometry histograms are shown (with mean fluorescence int$CD5^{lo}$y (MFI) in top right). Data are summarized from 2-5 independent experiments. Dotted lines in histograms denote ^{CD5lo} modes; data points in graphs represent individual mice (n=10-20); lines denote group means and average fold differences are indicated.

Statistics: All TRs had FDR<0.01 except *Lef1* (FDR=0.012) I, Wilcoxon matched-pairs signed rank test I. $^{**}p<0.01$, $^{****}p<0.0001$. See also Figure S3.

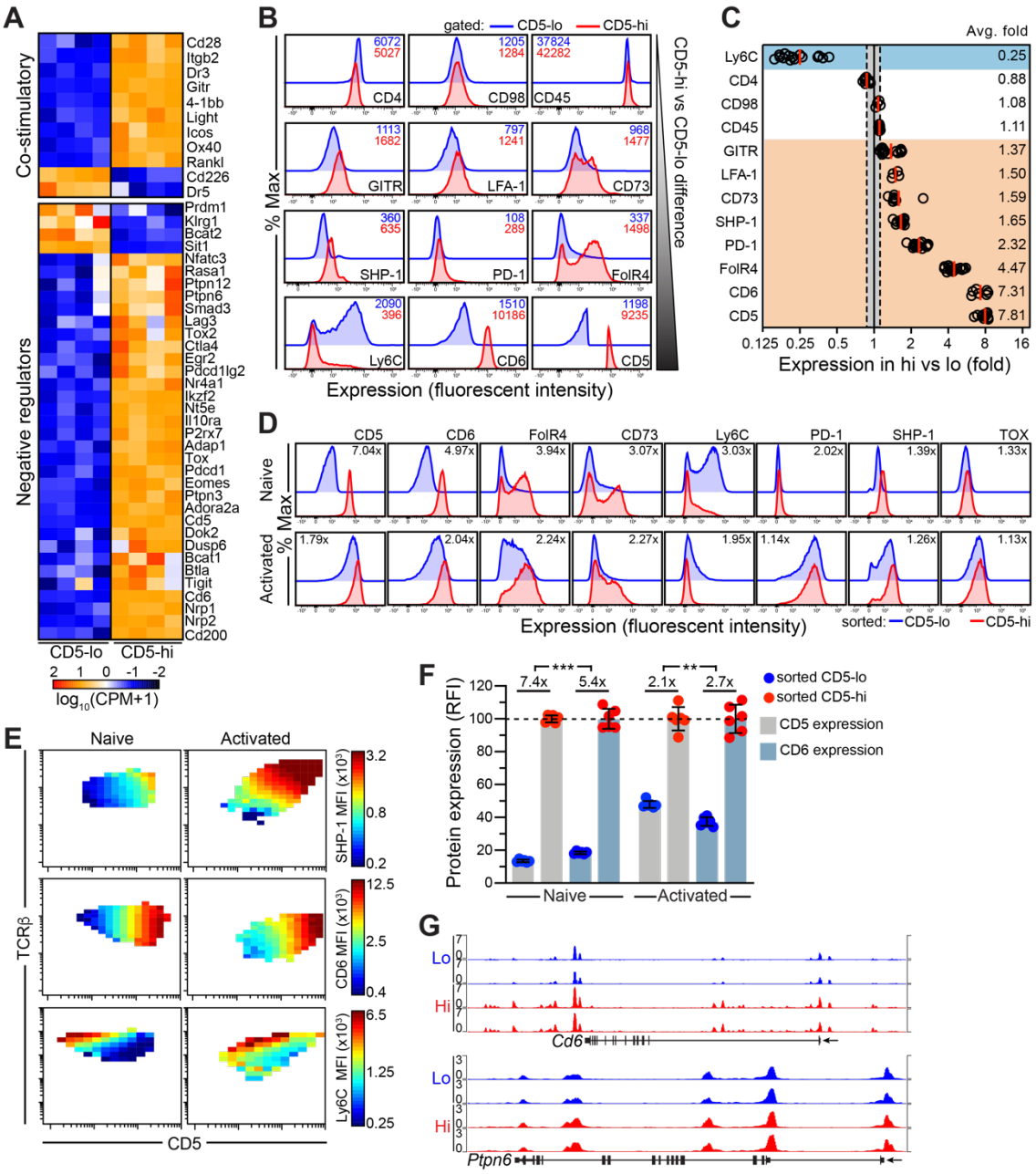


Figure 4. Pre-existing transcriptional and protein differences among naïve CD4^{++} T cells are maintained post-activation. (**A**) Heatmap of curated list of DEGs involved in positive and negative regulation of T cell activation and TCR signaling (n=4). (**B**) Representative histograms of protein expression (measured by flow cytometry) of select DEGs identified from bulk RNA-seq analyses comparing ^{CD5lo} and ^{CD5hi} naïve CD4^{++} T cells. Numbers in top right of histograms indicate MFI in ^{CD5lo} (blue) and ^{CD5hi} (red) cells. (**C**) Summary of fold expression differences of proteins measured in B that were significantly different between ^{CD5hi} and ^{CD5hi} groups from 2-5

independent experiments. Non-significant DEGs ($CD4^+$, CD98, and CD45) were used as controls to establish a cut-off for biological significance. Greater expression in ^{CD5lo} (blue shading) and greater expression in ^{CD5hi} (orange shading). Each data point is from an individual mouse (n=4-15); red lines denote group means. (**D**) Representative histograms of protein expression measured by flow cytometry in sorted 15% ^{CD5lo} and ^{CD5hi} naïve $CD4^{++}$ T cells pre- and 24 hours post-activation with anti-CD3/CD28. Numbers in top of histograms indicate fold difference between ^{CD5lo} (blue) and ^{CD5hi} (red) populations. I Single-cell flow cytometry analysis for SHP-1, CD6, and Ly6C expression among naïve $CD4^{++}$ T cells pre- and 24 hours post-activation. Color scale represents MFI of proteins of interest in bins of at least 10 cells across the distribution of CD5 (x-axis) and TCRβ (y-axis). (**F**) Naïve $CD4^{++}$ T cells were sorted into ^{CD5lo} and ^{CD5hi}, and CD5 and CD6 protein expression measured pre- and post-activation (anti-CD3/CD28). RFI normalized to CD5 or CD6 expression in the ^{CD5hi} population. Data are summarized from 2 independent experiments; each data point represents 4-5 pooled mice (n=6); error bars represent mean ± S.D. (**G**) ATAC-seq signal profiles across *Ptpn6* and *Cd6* gene loci shown from 2 biological replicates of sorted samples.

Statistics: All genes had FDR<0.01 (A), paired t test (F). $^{**}p<0.01$, $^{***}p<0.001$. See also Figure S4.

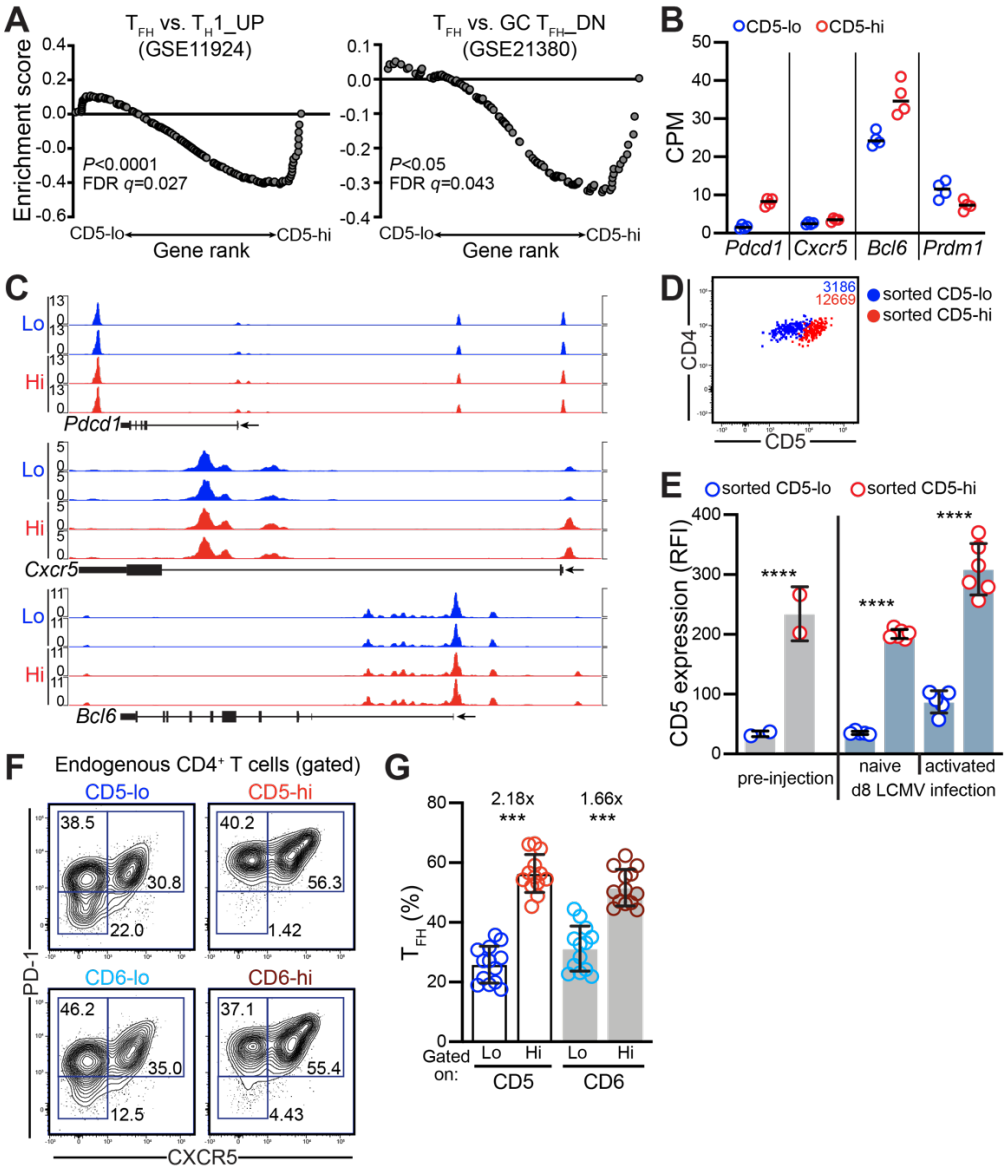


Figure 5. ^{CD5hi} naïve CD4^{++} T cells are enriched for T_{FH}-associated genes and have a greater T_{FH} differentiation potential upon infection than ^{CD5lo} cells. (**A**) GSEA of T_{FH} signatures enriched in ^{CD5lo} and ^{CD5hi} naïve CD4^{++} T cells. (**B**) *Pdcd1*, *Cxcr5*, *Bcl6*, and *Prdm1* mRNA expression from bulk RNA-seq for sorted ^{CD5lo} and ^{CD5hi} naïve CD4^{++} T cell populations. Lines indicate group means (n=4). (**C**) ATAC-seq signal profiles across *Pdcd1*, *Cxcr5*, and *Bcl6* gene loci from 2 biological replicates of sorted ^{CD5lo} and ^{CD5hi} naïve CD4^{++} T cell samples. (**D, E**) Naïve 15% ^{CD5lo} and ^{CD5hi} CD4^{++} T cells were sorted and adoptively transferred into recipients and infected 1 day later with LCMV. Representative flow cytometry plot of activated (CD44^{hi}) transferred cells with numbers in top right indicating CD5 MFI of sorted ^{CD5lo} (blue) and ^{CD5hi} (red)

cells (D), and summary of CD5 RFI (relative to endogenous total naïve $CD4^{++}$ T cells) of sorted ^{CD5lo} and ^{CD5hi} naïve or activated $CD4^{++}$ T cells I. Data are summarized from 2 independent experiments (D, E); each data point is from an individual recipient mouse (n=2-6); error bars represent mean ± S.D. (**F,G**) Activated ($CD44^{hi}$) $CD4^{++}$ T cells isolated day 8 post LCMV infection were gated on the top and bottom 15% CD5- or CD6-expressing cells and the percent T_{FH} determined. Representative flow cytometry plots are shown, with numbers indicating percent cells within each gate (F), data are summarized from 2 independent experiments; each data point is from an individual mouse (n=13); error bars represent mean ± S.D. (G).

Statistics: All genes had FDR<0.01 except *Cxcr5* (FDR=0.044) (B), one-way ANOVA with Tukey's multiple comparison I, Wilcoxon matched-pairs signed rank test (G). ***p<0.001, ****p<0.0001. See also Figure S5.

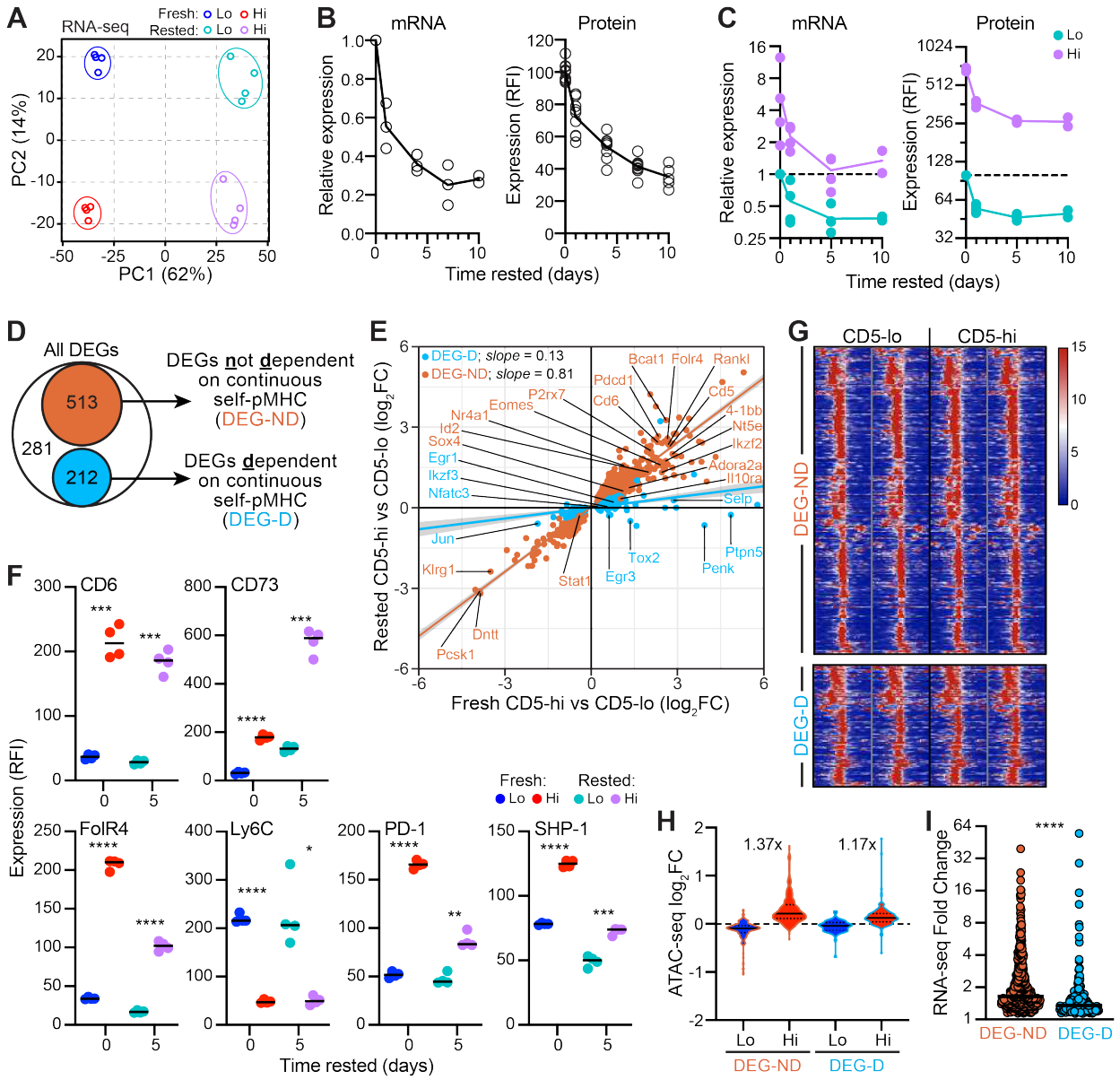


Figure 6. Withdrawal of naïve CD4^{++} T cells from self-pMHC identifies transcriptional and chromatin differences between ^{CD5lo} and ^{CD5hi} cells that do not rely on continuous self-pMHC interactions. (**A**) PCA of RNA-seq data from fresh and rested ^{CD5lo} and ^{CD5hi} naïve CD4^{++} T cells (n=4). (**B**) CD5 mRNA (normalized to *Gapdh*) and protein expression, relative to day 0, measured for naïve CD4^{++} T cells rested in culture in the presence of IL-7. Data are summarized from 2 independent experiments; each data point is from an individual mouse (n=3-9). (**C**) CD5 mRNA (normalized to *Tbp*) and protein expression determined after resting as in B of sorted 15% ^{CD5lo} and ^{CD5hi} naïve CD4^{++} T cells. Data from 1 experiment; each data point represented 9-10 pooled mice. Day 0, 1 and 5 (n=4) and day 10 (n=2). There is a significant effect between ^{CD5lo} and ^{CD5hi}

in both mRNA (p<0.05) and protein (p<0.0001) groups post-resting. (**D**) Venn diagram dividing DEGs identified by RNA-seq into two groups: DEG-ND (orange) were differentially expressed (FDR≤0.01) in the fresh ^{CD5hi} vs. ^{CD5lo} comparison and remained differentially expressed (FDR≤0.01) in the rested ^{CD5hi} vs. ^{CD5lo} comparison; DEG-D (blue) were differentially expressed (FDR≤0.01) in the fresh ^{CD5hi} vs. ^{CD5lo} comparison but were not differentially expressed (FDR≥0.3) in the rested comparison. I FC expression of DEG-ND and DEG-D identified in D, in both the fresh and rested RNA-seq. Lines indicate best fits, slopes are shown. (**F**) Naïve $CD4^{++}$ T cells were sorted into ^{CD5lo} and ^{CD5hi} and rested in culture in the presence of IL-7. RFI (normalized to total naïve $CD4^{++}$ T cells) from genes identified from D. Data from 1 experiment; each data point represents 3 mice pooled (n=4); lines denote group means. (**G**) ATAC-seq chromatin accessibility heatmap for open chromatin regions among DEG-ND and DEG-D. (**H**) Summary of data in G with fold differences between ^{CD5hi} and ^{CD5hi} open chromatin peaks indicated on the graph. (**I**) Gene expression FCs from RNA-seq between ^{CD5lo} and ^{CD5hi} cells in DEG-ND and DEG-D groups identified in D. Lines denote group means.

Statistics: Two-way ANOVA with Sidak's multiple comparison I, Paired t test (F), Mann-Whitney test (I). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns = non-significant. See also Figure S6.

2.8 Star Methods

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse TCRβ (H57-597)	BioLegend	Cat #109243;; RRID: AB_2629564
Anti-mouse CD4 (RM4.5)	BioLegend	Cat #100552; RRID: AB_2563053
Anti-mouse CD8a (53-6.7)	BioLegend	Cat #100742; RRID: AB_2563056
Anti-mouse CD5 (53-7.3)	BioLegend	Cat #100627; RRID: AB_2563930
Anti-mouse Foxp3 (FJK-16 s)	Life Technologies	Cat #11-5773-82; RRID: AB_465243
Anti-mouse CD44 (IM7)	Life Technologies	Cat #45-0441-82; RRID: AB_925746
Anti-mouse CD62L (MEL-14)	BioLegend	Cat #104440; RRID: AB_2629685
Anti-mouse CD25 (PC61.5)	Life Technologies	Cat #48-0251-82; RRID: AB_10671550
Anti-mouse CD45 (30F11)	Life Technologies	Cat #17-0451-83; RRID: AB_469393
Anti-mouse CD98 (RL388)	Life Technologies	Cat #12-0981-81; RRID: AB_465792
Anti-mouse GITR (DTA-1)	Life Technologies	Cat #25-5874-80; RRID: AB_10548516
Anti-mouse LFA-1 (H155-78)	BioLegend	Cat #141007; RRID: AB_10694861
Anti-mouse CD73 (TY/11.8)	BioLegend	Cat #127210; RRID: AB_11218786
Anti-mouse PD-1 (29F.1A12)	BioLegend	Cat #135216; RRID: AB_10689635
Anti-mouse FolR4 (eBio12A5)	BioLegend	Cat #125009; RRID: AB_1134201
Anti-mouse Ly6C (HK1.4)	BioLegend	Cat #128012; RRID: AB_1659241
Anti-mouse CD6 (OX-129)	BioLegend	Cat #146404; RRID: AB_2562753
Anti-mouse CXCR5 (SPRCL5)	Life Technologies	Cat #13-7185-82; RRID: AB_2572800
Anti-mouse CD45.1 (A20)	Life Technologies	Cat #12-0453-82; RRID: AB_465675
Anti-mouse CD45.2 (104)	BioLegend	Cat #109806; RRID: AB_313443
Anti-mouse TOX (TXRX10)	Life Technologies	Cat #12-6502-82; RRID: AB_10855034
Anti-mouse CD69 (H1.2F3)	Life Technologies	Cat #25-0691-82; RRID: AB_469637
Anti-mouse B220 (RA3-6B2)	BioLegend	Cat #103241; RRID: AB_11204069
Anti-mouse F4/80 (BM8)	Life Technologies	Cat #48-4801-82; RRID: AB_1548747
Anti-mouse Ly6G (1A8)	BioLegend	Cat #127641; RRID: AB_2565881
Anti-mouse CD11b (M1/70)	BioLegend	Cat #101259; RRID: AB_2566568
Anti-mouse CD11c (N418)	Life Technologies	Cat #48-0114-82; RRID: AB_1548654
Anti-mouse NK1.1 (PK136)	BioLegend	Cat #108722; RRID: AB_2132712
Anti-mouse CD19 (eBio1D3)	Life Technologies	Cat #48-0193-82; RRID: AB_2734905
Anti-mouse IL-2 (JES6-5H4)	Life Technologies	Cat #12-7021-82; RRID: AB_466150
LEF-1 (C12A5) Rabbit mAb	Cell Signaling Technology	Cat #2230S; RRID: AB_823558
TCF-1 (C46C7) Rabbit mAb	Cell Signaling Technology	Cat #2206S; RRID: AB_2199300
SHP-1 (C14H6) Rabbit mAb	Cell Signaling Technology	Cat #3759S; RRID: AB_2173694
Goat anti-Rabbit IgG (H+L) Cross-Absorbed Secondary Antibody, Alexa Fluor 488	Life Technologies	Cat #A-11008; RRID: AB_143165
Donkey anti-Rabbit IgG (H+L) Highly Cross-Absorbed Secondary Antibody, Alexa Fluor 647	Life Technologies	Cat #A-31573; RRID: AB_2536183
Ultra-LEAF Purified anti-mouse CD3e	BioLegend	Cat #100340; RRID: AB_11149115
Ultra-LEAF Purified anti-mouse CD28	BioLegend	Cat #102116; RRID: AB_11147170
TruStain FcX (anti-mouse CD16/32)	BioLegend	Cat #101320; RRID: AB_1574975
Bacterial and virus strains		
LCMV Armstrong	Dr. John Harty Lab	University of Iowa, Iowa City, USA

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
RPMI 1640	Wisent	Cat #350-030-CL
FBS	Wisent	Cat #080-450
L-glutamine (200mM)	Wisent	Cat #609-065-EL
Penicillin/streptomycin (5000IU)	Wisent	Cat #450-201-EL
2-Mercaptoethanol (50mM)	Life Technologies	Cat #21985023
Non-essential amino acids (100X)	Wisent	Cat #321-011-EL
Sodium pyruvate (100mM)	Wisent	Cat #600-110-EL
HEPES (1M)	Wisent	Cat #330-050-EL
0.5M EDTA (pH 8.0)	Life Technologies	Cat #AM9262
IL-7	BioLegend	Cat #577804
Zombie UV Fixable Viability Kit	BioLegend	Cat #423108
Fixable Viability Dye eFluor 780	Life Technologies	Cat #65-0865-18
Fixable Viability Dye eFluor 506	Life Technologies	Cat #65-0866-18
PBS (1X)	Wisent	Cat #311-010-CL
Permeabilization Buffer (10X)	Life Technologies	Cat #00-8333-56
TRIzol	ThermoFisher Scientific	Cat #15596026
Brefeldin A solution (1000X)	BioLegend	Cat #420601
Monensin solution (1000X)	Life Technologies	Cat #00-4505-51
ACK lysis buffer	Life Technologies	Cat #A1049201
FoxP3/Transcription Factor Fixation/ Permeabilization Concentrate and Diluent	Life Technologies	Cat #00-5521-00
Trypan Blue Solution, 0.4%	ThermoFisher Scientific	Cat# 15250061
Critical commercial assays		
Transcription factor phospho buffer set	BD Biosciences	Cat #563239
EasySep Mouse CD4 T cell Isolation Kit	Stemcell	Cat #19852
EasySep Mouse T Cell Isolation Kit	Stemcell	Cat #19851
Quick-RNA miniprep kit	Zymo Research	Cat #R1054
Purelink RNA Mini Kit	Life Technologies	Cat #12183018A
High-Capacity cDNA reverse transcription kit	Life Technologies	Cat #4368814
Deposited data		
GEO SuperSeries (bulk ATAC-seq, bulk RNA-seq, and scRNA-seq)	This paper	GEO: GSE185677
Bulk ATAC-seq	This paper	GEO: GSE185674
Bulk RNA-seq	This paper	GEO: GSE185675
scRNA-seq	This paper	GEO: GSE185676
scRNA-seq source code	Zenodo, this paper	https://doi.org/10.5281/zenodo.5570794
Experimental models: Organisms/strains		
C57BL/6 mice	Jackson Laboratories	Cat #000664; RRID: IMSR_JAX:000664
B6 CD45.1 mice	Jackson Laboratories	Cat #002014; RRID: IMSR_JAX:002014
$Foxp3^{GFP}$ transgenic mice	(Oukka, 2007)	N/A
SMARTA TCR transgenic mice	(Oxenius et al., 1998)	N/A
$TCR\beta^{-/-}$ mice	(Mombaerts et al., 1992)	N/A
$RAG2^{GFP}$ transgenic mice	(Yu et al., 1999)	N/A
Oligonucleotides		
Cd5 (Mm00432417_m1) FAM	Life Technologies	Cat #4331182
Foxp3 (Mm00475162_m1) FAM	Life Technologies	Cat #4331182
Gapdh (Mm99999915_g1) FAM	Life Technologies	Cat #4351370

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Tbp (Mm01277042_m1) VIC	Life Technologies	Cat #4448489
Software and algorithms		
GraphPad Prism	GraphPad	https://www.graphpad.com; RRID: SCR_002798
FlowJo	BD Biosciences	https://www.flowjo.com; RRID:SCR_008520
Integrative Genomics Viewer 2.9.4	(Thorvaldsdóttir et al., 2013)	https://software.broadinstitute.org/software/igv/; RRID: SCR_011793
MeV	(Howe et al., 2011)	https://mev.tm4.org/#/about; RRID: SCR_001915
GSEA	(Subramanian et al., 2005)	https://www.gsea-msigdb.org/gsea/index.jsp; RRID: SCR_003199
Cytoscape 3.8.1	(Shannon et al., 2003)	https://cytoscape.org; RRID: SCR_003032
ClueGO 2.5.4	(Bindea et al., 2009)	http://www.ici.upmc.fr/cluego/; RRID: SCR_005748
ScatterSlice	(Cotari et al., 2013)	N/A
HOMER	(Heinz et al., 2010)	http://homer.ucsd.edu/homer; RRID: SCR_010881
CASAVA 1.8.2	Illumina	https://www.illumina.com; RRID: SCR_001802
TopHat 2.0.11	(Kim et al., 2013)	http://ccb.jhu.edu/software/tophat/index.shtml; RRID: SCR_013035
featureCounts 1.4.5 in SubRead package	(Liao et al., 2014)	N/A
EdgeR	(Robinson et al., 2010)	http://bioconductor.org/packages/release/bioc/html/edgeR.html; RRID: SCR_012802
BWA	(Li and Durbin, 2010)	http://bio-bwa.sourceforge.net/; RRID: SCR_010910
scater	(McCarthy et al., 2017)	https://bioconductor.org/packages/release/bioc/html/scater.html; RRID: SCR_015954
scran	(Lun et al., 2016)	https://bioconductor.org/packages/release/bioc/html/scran.html; RRID: SCR_016944
uwot	(Melville, 2019)	N/A
umap	(McInnes et al., 2020)	https://github.com/lmcinnes/umap; RRID: SCR_018217
FastQC 0.10.1	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/; RRID: SCR_014583
Trimmomatic 0.33	(Bolger et al., 2014)	http://www.usadellab.org/cms/index.php?page=trimmomatic; RRID: SCR_011848
Bowtie 1.0.0	(Langmead et al., 2009)	http://bowtie-bio.sourceforge.net/index.shtml; RRID: SCR_005476
MACS1.4.1	(Zhang et al., 2008)	N/A
PeakAnalyzer	(Salmon-Divon et al., 2010)	http://www.bioinformatics.org/peakanalyzer/wiki/;; RRID: SCR_001194
pheatmap	(Raivo, 2018)	https://www.rdocumentation.org/packages/pheatmap/versions/0.2/topics/pheatmap; RRID: SCR_016418
ggplot2	(Wickham, 2016)	https://cran.r-project.org/web/packages/ggplot2/index.html; RRID: SCR_014601
Other		
UltraComp eBeads compensation beads	Life Technologies	Cat #01-2222-42

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Judith Mandl (judith.mandl@mcgill.ca).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Additional information required to reanalyze the data reported in this paper are available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

C57BL/6 mice, CD45.1^{+}, Foxp3^{GFP+} Tg [301], SMARTA TCR Tg [302], TCRβ$^{-/-}$ [303], and MHCII$^{-/-}$ mice [304] were purchased from Jackson Laboratories (Bar Harbor, ME) and bred in-house. RAG2^{GFP} Tg mice [305] were obtained from Sylvie Lesage (Université de Montréal). All mice were on a B6 background, both male and female mice were used, and were 6-12 weeks of age. Animal housing, care and research were in accordance with the Guide for the Care and Use of Laboratory Animals and all procedures performed were approved by the McGill University, Maisonneuve-Rosemont Hospital Research Center, Radboud University, or NIAID Animal Care Committee.

Infections

LCMV Armstrong was propagated from a strain provided by Dr. J. Harty (University of Iowa). Mice were infected with 2x10^{5} plaque forming units (PFU) by intra-peritoneal injection [306]. Cellular responses were assessed 8 days post-infection.

METHOD DETAILS

Lymphocyte isolation, resting, activation, and restimulation

Spleen and peripheral lymph nodes (inguinal, axillary, brachial, superficial cervical, and mesenteric) were passed through a 70μm filter with 1% RPMI (1% FBS, L-glutamine, and pen/strep). ACK lysis buffer (Life Technologies) was added for 3 minutes and samples were refiltered and resuspended in 1% RPMI. Cell counts were determined by diluting a single-cell suspension 1:10 in Trypan Blue (ThermoFisher Scientific) and manually counting live single cells (Trypan Blue-negative) on a hemocytometer. For experiments where naïve $CD4^+$ T cells were rested in culture, cells (either total or sorted, as specified) were kept in complete RPMI (10% FBS, 1% L-glutamine, 1% pen/strep, 1% HEPES buffer, 1% Sodium Pyruvate, 1% non-essential Amino Acids, and 0.1% 2-mercapto-ethanol 1000X solution) supplemented with IL-7 (10 ng/mL, Biolegend). To activate T cells, sorted cells or total splenocytes were cultured in complete RPMI in 96-well plates coated with α-CD3 and α-CD28 (Invitrogen; both at 3μg/mL). Restimulation of splenocytes for cytokine production was performed with α-CD3 and α-CD28 (Invitrogen; both at 2μg/mL) in 96-well plates with brefeldin A and monensin (Invitrogen, both diluted 500X) for 5 hours at 37°C.

Flow cytometry

Samples were incubated in Fixable Viability Dye (AF780 or eF506, Life Technologies; Zombie UV, BioLegend) diluted in PBS for 20 minutes at 4°C. Extracellular antibodies were diluted in FACS buffer (2% FBS and 5mM EDTA in PBS) with Fc Block (Life Technologies) and incubated for 30 minutes at 4°C. Samples requiring intracellular staining were subsequently incubated in FoxP3 Transcription Factor Fixation/Permeabilization Concentrate and Diluent (Life Technologies) for 30 minutes at 4°C. Intracellular antibodies were diluted in permeabilization wash buffer and incubated for 30-60 minutes at 4°C. Directly conjugated antibodies used were as follows: TCRβ (H57-597), CD4 (RM4.5), CD8a (53-6.7), CD5 (53-7.3), Foxp3 (FJK-16s), CD44 (IM7), CD62L (MEL-14), CD25 (PC61.5), CD45 (30F11), CD98 (RL388), GITR (DTA-1), LFA-1 (H155-78), CD73 (TY/11.8), PD-1 (29F.1A12), FolR4 (eBio12A5), Ly6C (HK1.4), CD6 (OX-129), CXCR5 (SPRCL5), CD45.1 (A20), CD45.2 (104), TOX (TXRX10), CD69 (H1.2F3), B220 (RA3-6B2), F4/80 (BM8), Ly6G (1A8), CD11b (M1/70), CD11c (N418), NK1.1 (PK136), CD19 (eBio1D3), and IL-2 (JES6-5H4). Primary unconjugated antibodies used were: LEF-1 (C12A5), TCF-1 (C46C7), SHP-1 (C14H6). Secondary conjugated antibodies used were either Goat anti-Rabbit IgG (H+L) Alexa Fluor 488 or Donkey anti-Rabbit IgG (H+L) Alexa Fluor 647. For

samples assessed for SHP-1 expression, cells were fixed with 1X TFP Fix/Perm Buffer for 50 minutes at 4°C, then incubated in Perm Buffer III (BD Biosciences) for 20 minutes on ice. Fc Block, surface, and intracellular antibodies were diluted in 1X TFP Perm/Wash Buffer and incubated for 50 minutes at 4°C, and secondary antibody diluted in 1X TFP Perm/Wash Buffer was added for an additional 50 minutes at 4°C. For all flow cytometry experiments, cells were acquired using an LSRFortessa (BD Bioscience) and analyzed with FlowJo software (BD Bioscience).

Cell sorts

Lymphocytes from B6 or CD45.1$^+$ congenic mice were isolated in single cell suspension as described. Samples for bulk RNA-seq, ATAC-seq, *in vivo* or *in vitro* assays were pooled from spleen and lymph nodes (inguinal, axillary, brachial, mesenteric, and cervical) from 2-10 mice. Samples for scRNA-seq were from a spleen from a single mouse. Total isolated cells or cells magnetically enriched for CD4 or total T cells (Stemcell EasySep mouse total T cell or CD4$^+$ T cell enrichment kits) were then incubated in fixable viability dye, and subsequently stained with surface antibodies for 1 hour at 4°C. Naïve CD4$^+$ T cell were sorted on CD5 expression (top and bottom 15%) for bulk analyses; single naïve CD4$^+$ T cells were sorted into 384-well plates for subsequent scRNA-seq. Naïve CD4$^+$ T cells were sorted on singlets, live, dump-negative (RNA-seq and ATAC-seq), TCRβ$^+$ (bulk- and scRNA-seq), CD4$^+$, CD8$^-$, CD25$^-$ (scRNA-seq and *in vivo* and in *vitro* assays) or Foxp3^{GFP-} (RNA-seq), CD44^{lo}, CD62L^{hi}, and 15-25% CD5^{lo} and CD5^{hi} (RNA-seq, ATAC-seq, and *in vivo* and *in vitro* assays). Dump channel included B220, CD11b, CD11c, F4/80, Ly6G, NK1.1, and CD69 for RNA-seq; the ATAC-seq dump channel also included CD19 and CD25 (ATAC-seq). Sorts were performed on either a FACS Aria Fusion, Aria III, or Aria II SORP (BD Bioscience). All cell populations were sorted to >90% purity for bulk populations.

Adoptive cell transfers

For all adoptive cell transfer experiments, donors and recipients were sex-matched.

TCRβ$^{-/-}$ adoptive transfer: 25% CD5^{lo} and CD5^{hi} cells were sorted from CD45.2$^+$ mice and 1x10^6 cells were adoptively transferred by i.v. injection into CD45.1$^+$ TCRβ$^{-/-}$ recipients and mice were followed for 91 days.

LCMV infection: 15-18 CD45.1$^+$ or CD45.2$^+$ mice were used as donors to obtain a total of 12-20x10^6 cells from each of 15% CD5^{lo} and 15% CD5^{hi} cells sorted as detailed above. 6-10x10^6 sorted donor cells were adoptively transferred by i.v. injection into CD45.2$^+$ or CD45.1$^+$ recipients (n=2 per group), respectively. One day post-transfer, mice were infected with LCMV as described. Cells were isolated from the spleens and peripheral lymph nodes of recipient mice 8 days post-infection.

SMARTA transgenic T cell adoptive transfer: 1x10^4 CD45.2$^+$ SMARTA CD4$^+$ T cells were adoptively transferred by i.v. injection into CD45.1$^+$ recipients. One day post-transfer, mice were infected with LCMV as described. Cells were isolated form the spleens of recipient mice 8 days post-infection.

Bulk RNA sequencing

1x10^6 cells from four independent samples, each with cells pooled from 2 mice, were sorted as described and CD5^{lo} and CD5^{hi} naïve CD4$^+$ T cells were either directly added to 500μL TRIzol (ThermoFisher Scientific) or rested in complete RPMI supplemented with IL-7 for 22 hrs first. RNA was purified using RNA miniprep kit (Zymo Research) according to manufacturers' recommendations. 500ng of purified RNA was used to prepare RNA-seq libraries using TruSeq mRNA library preparation kit v2 (Illumina). Libraries were sequenced on an Illumina HiSeq 2000 using v3 chemistry and 50 cycle paired end reads. Illumina bcl files were converted to FASTQ using CASAVA1.8.2 and mapped to the UCSC mm9 mus musculus genome annotation using Tophat 2.0.11 [307]. Reads overlapping exons were counted using featureCounts version 1.4.5 from the SubRead package [308], with a minimum read mapping quality score of 10. Normalized read counts differential gene expression analysis was performed with EdgeR [309]. In order for a gene to be included in the matrix a minimum CPM value of 5 in at least 3 of the 4 replicates was required. The p-values were corrected using the Benjamini-Hochberg method and an FDR threshold of <0.01 was considered significant. No FC threshold was set unless stated otherwise.

Single cell RNA sequencing

Each well within a 384-well plate contained CEL-Seq2 primers covered by mineral oil. Primers consisted of a 24bp polyT stretch, a 6bp random molecular barcode (UMI), a cell-specific barcode, the 5' Illumina TruSeq small RNA kit adaptor and a T7 promoter. After sorting, the plates were

frozen at -80°C until further use. Single cell RNA-seq library preparation and sequencing was performed by Single Cell Discoveries (Utrecht, Netherlands) [310]. Libraries were prepared following the SORT-seq protocol [311], which consists of an automated and improved version of the CEL-Seq2 protocol [312]. Briefly, cells were first lysed for 5 minutes at 65°C, and reverse transcription and second strand mixes were dispensed by the Nanodrop II liquid handling platform (GC Biotech). Single cell double stranded cDNAs were pooled together and *in vitro* transcribed for linear amplification. Illumina sequencing libraries were prepared using the TruSeq small RNA primers (Illumina) and sequenced paired-end at 75 bp read length the Illumina NextSeq. Paired-end reads from Illumina sequencing were aligned to the mouse transcriptome genome by BWA [313]. Read 1 contained the barcode information and was used for assigning reads to correct cells and libraries, while read 2 was mapped to gene models. Reads that mapped equally well to multiple locations were discarded. Read counts were first corrected for UMI barcode by removing duplicate reads that had identical combinations of library, cell-specific, and molecular barcodes and were mapped to the same gene. For each cell barcode the number of UMIs for every transcript was counted, and transcript counts were then adjusted to the expected number of molecules based on counts, 256 possible UMI's and poissonian counting statistics [314]. A unique feature of this protocol is the combination of both flow cytometry staining and RNA sequencing; this allowed for the simultaneous tracking of select protein expression and gene expression on single cells.

Single-cell RNA-seq data analysis

Raw read counts were first subjected to quality control. We identified two blocks of wells with fewer than 500 non-spike-in reads. To exclude these and similar low-content wells, we applied a UMAP clustering on all wells (including the spike-in reads) and excluded the cluster of cells that was mainly composed by the low-read wells. After quality control, 697 wells out of 1152 were kept in the analysis. R packages 'scater' [315] and 'scran' [273] were used for further processing, the spike-in reads were removed and expression values were normalized to library size and normalized log expression values and gene variance were determined as described previously described [273]. Mitochondrial genes were excluded before modeling gene variance. The processed data was then plotted or subjected to UMAP clustering [316], performed with the R package 'uwot' [317] using the cosine distance and a neighbourhood size of 30. For defining CD5

low, mid, and high cells the CD5 mean fluorescent intensity was logged by the flow cytometer when we sorted the cells into individual wells.

ATACseq library preparation, sequencing, and visualization

Two independent biological replicates of $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells were sorted as described, counted and $1x10^5$ nuclei pelleted. ATAC-seq libraries were prepared from the fresh nuclei pellets using Illumina Tagment DNA TDE1 Kit by the Institut de recherches cliniques de Montréal (Montreal, Canada). Briefly, paired-end 42bp sequencing reads were generated by Illumina sequencing (using a NovaSeq6000). The quality of the sequenced reads was checked using FastQC tool v0.10.1 (Babraham Bioinformatics), and low-quality bases removed using Trimmomatic v.0.33 [318]. The trimmed reads were mapped to the mouse UCSC mm9 genome using Bowtie 1.0.0 [319], in paired-end mode with --best parameter. Peak calling was performed using MACS1.4.1 [320] with p-values $<10^{-7}$. Subpeaks were identified using PeakAnalyzer [321], with parameters: *valley*=0.5 and *cutoff*=5 counts per million (cpm). Normalized sequenced read density profiles (bigwig) were generated using makeUCSCfile from Homer package [322], normalizing the total number of reads in each sample to 10^6, and visualized on Integrative Genomics Viewer (IGV) [323]. When visualizing ATAC-seq signal profiles for individual genes, group scaling was performed. Peaks identified in the biological replicates were pooled using mergePeaks from Homer package, merging peak summits within 50bp to each other. Read densities around the peak summits were retrieved using annotatePeaks from Homer package and quantiles normalized for FC comparison between $CD5^{lo}$ and $CD5^{hi}$ replicates [324]. Transcription factor binding motif enrichment analysis was performed using Homer package on unique peaks found only in $CD5^{lo}$ or $CD5^{hi}$ replicates with a *P*-value $<10^{-4}$. Hierarchical clustering of the peaks near the DEG-ND and DEG-D gene sets were performed using Pearson correlation with complete linkage method.

RNA extraction and quantitative real-time PCR

To assess *Cd5* and *Foxp3* expression in total and sorted $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells or Tregs, lymphocytes from B6 or $CD45.1^+$ congenic mice were isolated in single cell suspension as described. Lymphocytes were then magnetically enriched for total T cells or $CD4^+$ T cells (Stemcell EasySep mouse total T cell or $CD4^+$ T cell enrichment kits) and sorted as described.

RNA was extracted using Purelink™ RNA Mini Kit (Life Technologies) and cDNA converted using High-Capacity cDNA Reverse Transcription Kit (Life Technologies) according to manufacturers' recommendations. RT-qPCR analysis was performed with TaqMan™ Gene Expression Master Mix (Life Technologies) and TaqMan™ Gene Expression Assay (Life Technologies; FAM, *Cd5*, Mm00432417_m1; FAM, *Foxp3*, Mm00475162_m1). Housekeeping genes *Gapdh* (Life Technologies, FAM, Mm99999915_g1) or *Tbp* (Life Technologies, VIC, Mm01277042_m1) were used.

QUATIFICATION AND STATISTICAL ANALYSIS

Heatmaps

For RNA-seq these were created by either showing individual replicates or average expression within replicates. Log_{10}(CPM+1) were visualized using the pheatmap package in R-Project on a color scale of black-blue-white-orange-red [325]. For ATAC-seq, heatmaps were created using annotatePeaks from Homer package, taking read densities ±2.5kb with bin size of 50bp for the highest peak summit near each gene TSS. Images were generated using MeV tool with blue-white-red scale [326].

Geneset enrichment analysis (GSEA)

GSEAs were performed as previously described [327] using gene sets defined by the Molecular Signatures Database [328] or otherwise described.

Gene ontology pathway analysis

Enrichment of GO terms in naïve $CD4^+$ T cells was performed using ClueGO (version 2.5.4) [329] on Cytoscape (version 3.8.1) [330]. The following parameters were used when running ClueGO on the top 5% most variable genes from the scRNA-seq: Min GO Level =4, Max GO Level =6, Minimum Number of Genes associated to GO term =6, and Minimum Percentage of Genes associated to GO term =6. The following parameters were used when running ClueGO on bulk RNA-seq DEGs: Min GO Level =3; Max GO Level =4. For $CD5^{lo}$ cells: Minimum Number of Genes associated to GO term =3; Minimum Percentage of Genes associated to GO term =5. For $CD5^{hi}$ cells: Minimum Number of Genes associated to GO term =20; Minimum Percentage of Genes associated to GO term =10. Enrichment p-values were based on a hypergeometric test and

Benjamini-Hochberg method used for multiple testing correction. For bulk RNA-seq only pathways with $P \leq 0.05$ were considered significant.

PCA

PCA plots were built using filtered $\log_2$CPM (RNA-seq) or $\log_2$-transformed read densities around peak summits (ATAC-seq) using ggplot2 package in R-Project [331].

Mutual Information

Mutual Information (MI) is a robust, non-parametric measure of the statistical relationship between observables with distinct advantages over simple correlation measures [332]. MI is computed as:

$$MI = \sum_{x,y} p(x,y) \log_2 \left(\frac{p(x,y)}{p(x)p(y)} \right)$$

(Joint) probability distributions are obtained by binning the data into 96 geometrically spaced bins over the full Mean Fluorescent Intensity (MFI) range (10^0-10^6) for TCRβ, CD5, CD6, and SHP-1.

ScatterSlice analysis

Scale values corresponding to single $CD4^+$ T cells with expression of TCRβ, Ly6C, CD5, CD6, TOX, and SHP-1 from flow cytometry data were identified and exported as csv files for analysis in the R-Project ScatterSlice [333]. Cells were divided into defined bins (15x15 matrix with a minimum of 15 cells per bin) and within each bin, the average MFI of SHP-1, Ly6C, TOX, or CD6 was projected in false-color onto a plot of TCRβ versus CD5 expression.

Statistical analyses

Group comparisons were performed using Prism V9 (GraphPad). Unless specified, data are presented as mean ± standard deviations (SD) with each data point representing an individual mouse. The cut-off for significance considered was $p \leq 0.05$ for all analyses unless otherwise stated. Information about specific statistical tests used for each experiment are listed in the figure legends.

2.9 Supplemental Information

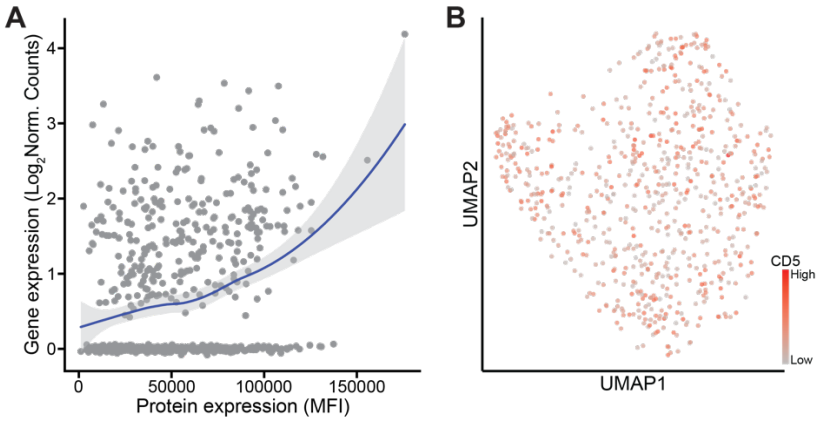


Figure S1. CD5 expression among single naïve CD4$^+$ T cells. Related to Figure 1. (**A**) Comparison of *Cd5* gene expression from scRNA-seq (y-axis) and CD5 protein expression from corresponding cells (x-axis) among naïve CD4$^+$ T cells. Each data point represents a single cell. The line shows the polynomial regression line across all data. (**B**) UMAP projection of the 2000 most variable genes from scRNA-seq profiles of 697 cells. CD5 protein expression level is overlayed as a grey to red intensity scale.

Statistics: Pearson correlation coefficient = 0.09, 95% CI 0.02, 0.17, *p*-value = 0.01 (A).

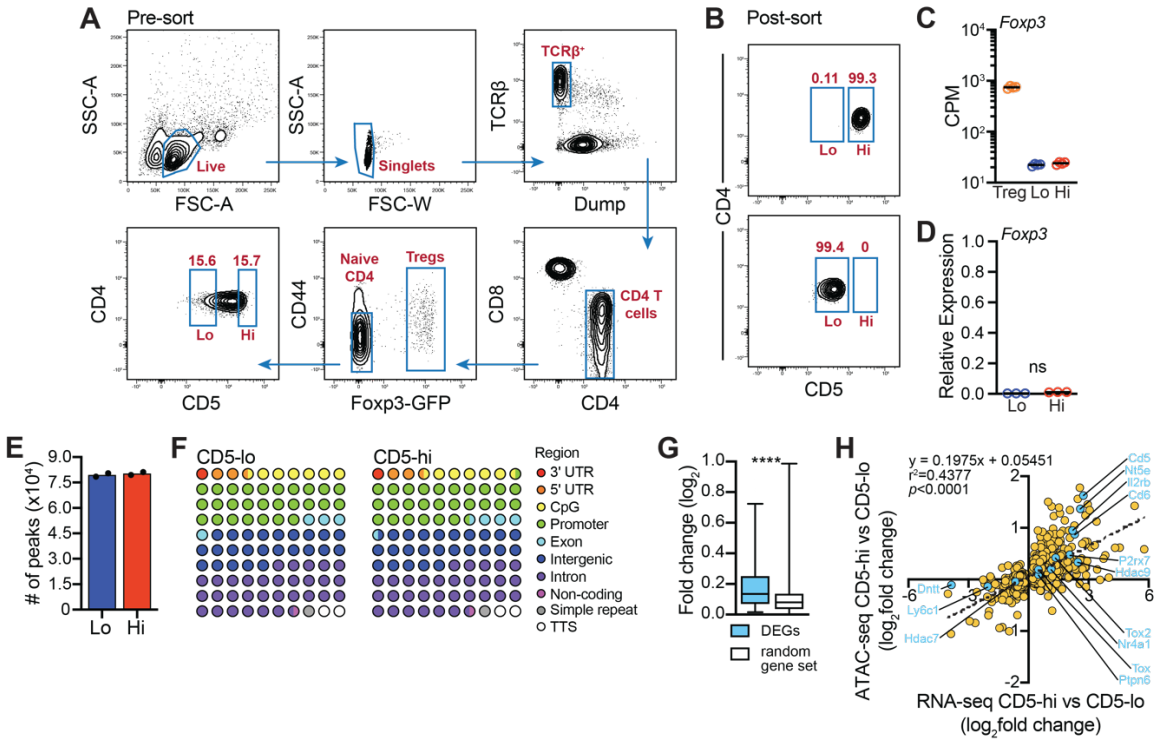


Figure S2. Sort strategy and differences in chromatin accessible regions between CD5^{lo} and CD5^{hi} naïve CD4^{+} T cells. Related to Figure 2. (**A**) Flow cytometry plots illustrating the sort strategy used for RNA- and ATAC-seq. Dump channel included B220, CD11b, CD11c, F4/80, Ly6G, NK1.1, and CD69. (**B**) Representative post-sort purity for RNA- and ATAC-seq; numbers above gates indicate percent of cells in gate. (**C,D**) *Foxp3* mRNA expression in sorted CD5^{lo} and CD5^{hi} naïve CD4^{+} T cells from RNA-seq data (n=4) (C) and verified by RT-qPCR in independently sorted CD5^{lo} and CD5^{hi} cells normalized to the housekeeping gene *Gapdh* and relative to Treg population (n=3) (D). Sorted Tregs (TCRβ^{+}, CD4^{+}, CD8^{-}, CD25^{+}) were used as positive control. Lines represent group means. (**E,F**) Global analysis of DARs in sorted CD5^{lo} and CD5^{hi} naïve CD4^{+} T cells from ATAC-seq. Total number of peaks (n=2) (E) and proportion of peaks annotated to genomic regions (F). (**G**) Absolute fold expression of DARs from identified DEGs between CD5^{hi} versus CD5^{lo} as compared DARs corresponding to a random set of genes. Box plot shows the quartiles and the 5th and 95th percentiles as whiskers. (**H**) FC expression of the 911 significant DEGs between CD5^{hi} versus CD5^{lo} indentified in **Fig. 2E** from bulk RNA-seq dataset with corresponding DARs from ATAC-seq dataset. Dotted line denotes line of best fit from linear regression (equation shown).

Statistics: Mann-Whitney test (D and G). ****p<0.0001, ns = non-significant.

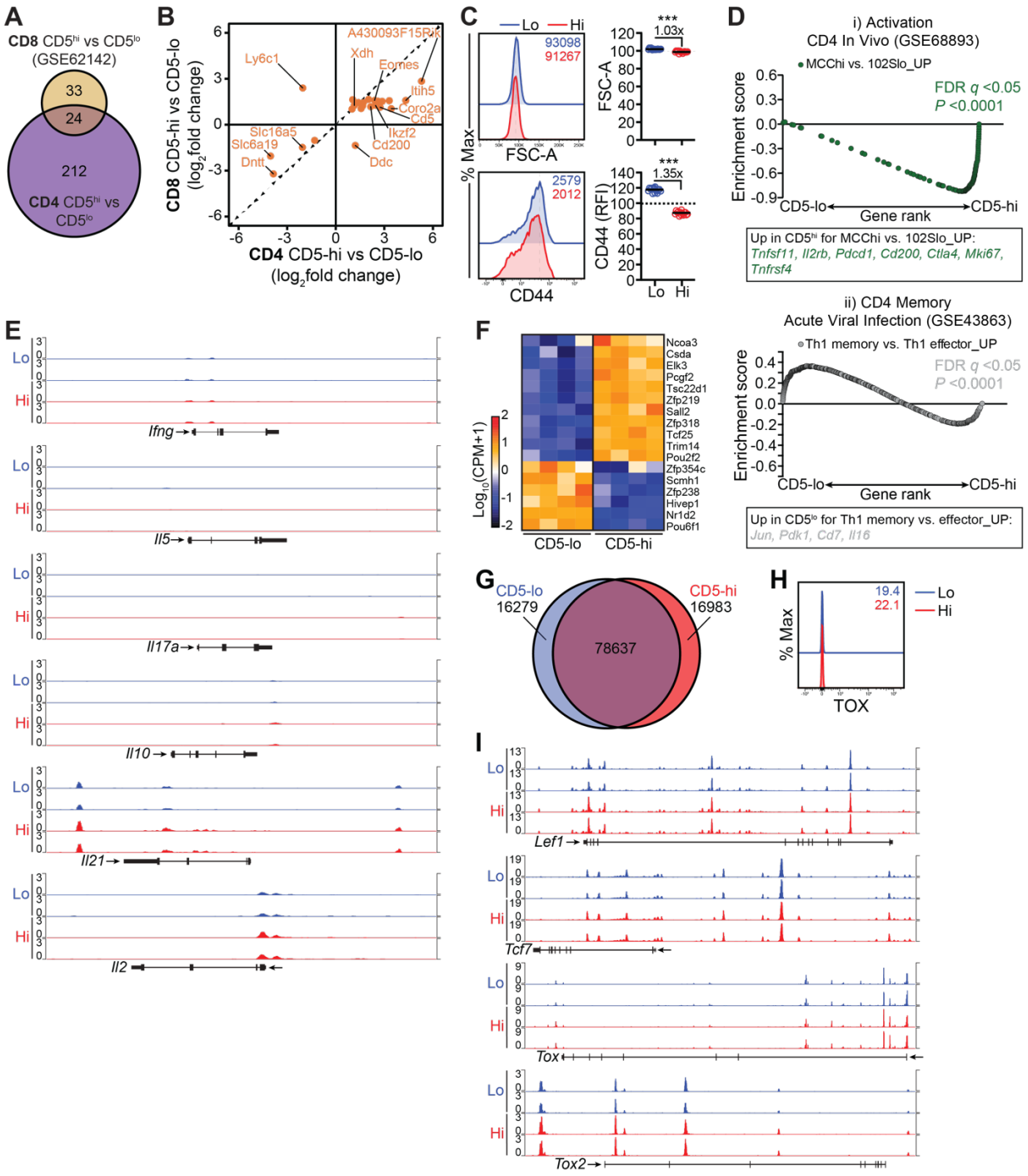


Figure S3. Comparison of chromatin accessible regions and transcriptomes of CD5^{hi} versus CD5^{lo} CD4^{+} T cells. Related to Figure 3. (**A**) Venn diagram of DEGs identified in **Fig. 2** with FC≥2, compared to published DEGs determined for naïve CD8^{+} T cells sorted into CD5^{hi} and CD5^{lo} (GSE62142). (**B**) Fold change expression differences between CD5^{hi} and CD5^{lo} of the 24 DEGs identified in the venn diagram overlap in A. (**C**) FSC-A and CD44 protein expression levels

of $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells (CD44 RFI is relative to total naïve $CD4^+$ T cells). Representative flow cytometry histograms are shown and data summarized from 3 independent experiments. Dotted lines in histograms denote $CD5^{lo}$ modes; data point in graphs represent individual mice (n=14); lines denote group means. (**D**) GSEA showing enrichment of gene signatures in (i) $CD4^+$ T cell activation *in vivo*, GSE68893, and (ii) $CD4^+$ T cell memory during acute LCMV infection, GSE43863, relative to the $CD5^{lo}$ versus $CD5^{hi}$ naïve $CD4^+$ T cell comparison. (**E**) ATAC-seq signal profiles of T cell effector cytokines from 2 biological replicates of sorted $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cell samples. (**F**) Heatmap of differentially expressed transcription factors (TF) between $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells not classified in **Fig. 3D**. (**G**) Venn diagram of unique DARs from ATAC-seq for $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells used for TF binding motif enrichment analysis in **Fig. 3E**. (**H**) Representative flow cytometry histogram of TOX fluorescence minus one (FMO) control in gated $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells. (**I**) ATAC-seq signal profiles of *Lef1*, *Tcf7*, *Tox*, and *Tox2* gene loci from 2 biological replicates of sorted $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cell samples.

Statistics: Wilcoxon matched-pairs signed rank test (C). ***$p<0.001$.

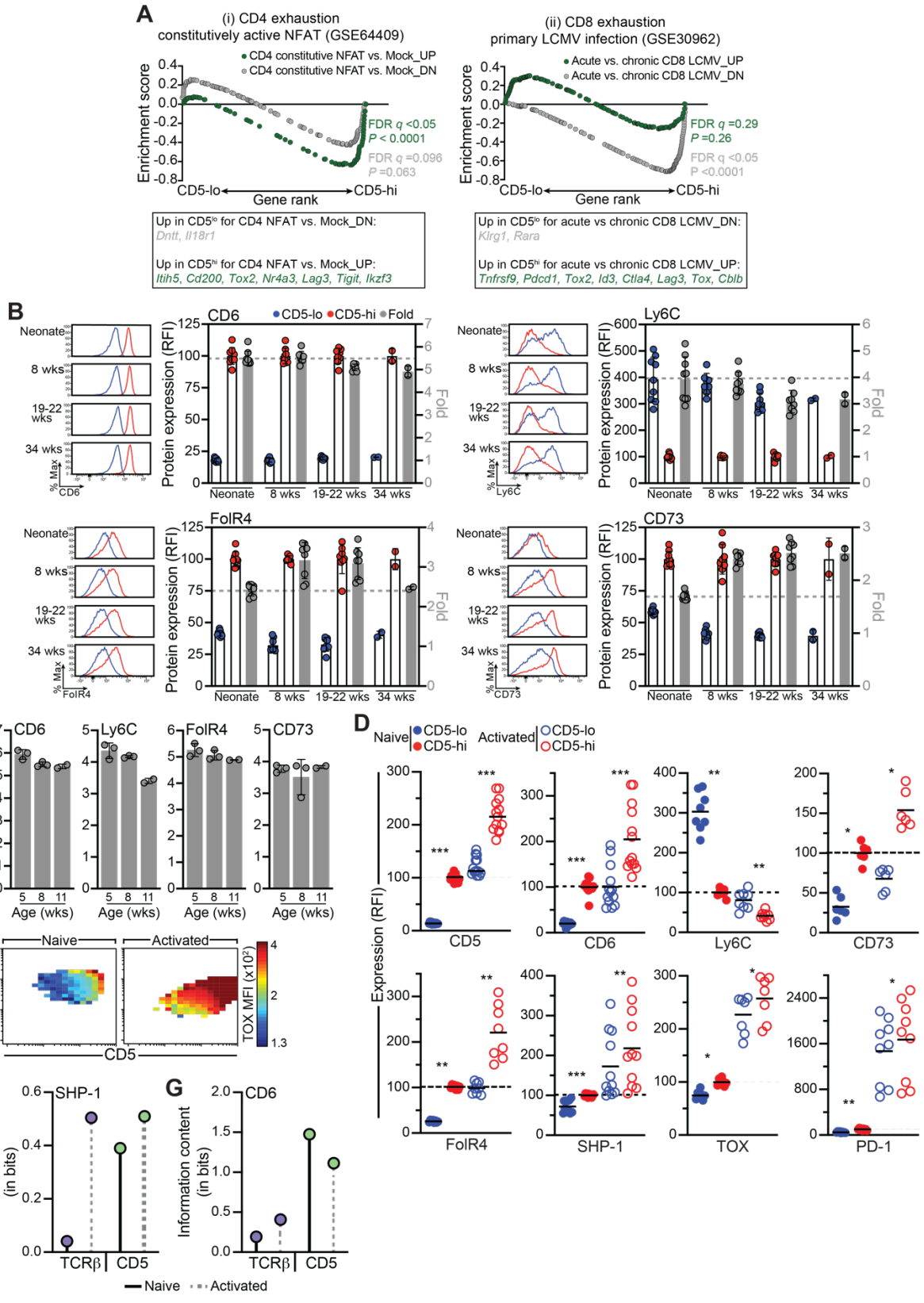

A
(i) CD4 exhaustion
constitutively active NFAT (GSE64409)
CD4 constitutive NFAT vs. Mock_UP
CD4 constitutive NFAT vs. Mock_DN
Enrichment score
FDR q <0.05
P < 0.0001
FDR q =0.096
P =0.063
CD5-lo
Gene rank
CD5-hi
Up in CD5lo for CD4 NFAT vs. Mock_DN:
Dntt, Il18r1
Up in CD5hi for CD4 NFAT vs. Mock_UP:
Itih5, Cd200, Tox2, Nr4a3, Lag3, Tigit, Ikzf3
(ii) CD8 exhaustion
primary LCMV infection (GSE30962)
Acute vs. chronic CD8 LCMV_UP
Acute vs. chronic CD8 LCMV_DN
FDR q =0.29
P =0.26
FDR q <0.05
P <0.0001
Up in CD5lo for acute vs chronic CD8 LCMV_DN:
Klrg1, Rara
Up in CD5hi for acute vs chronic CD8 LCMV_UP:
Tnfrsf9, Pdcd1, Tox2, Id3, Ctla4, Lag3, Tox, Cblb
B
CD6
CD5-lo CD5-hi Fold
Ly6C
FolR4
CD73
Neonate
8 wks
19-22 wks
34 wks
% Max
Protein expression (RFI)
Fold
C
CD6
Ly6C
FolR4
CD73
Fold
Age (wks)
D
Naive CD5-lo CD5-hi
Activated CD5-lo CD5-hi
Expression (RFI)
CD5
CD6
Ly6C
CD73
FolR4
SHP-1
TOX
PD-1
E
Naive
Activated
TCRβ
CD5
TOX MFI (x10²)
F
SHP-1
Information content (in bits)
TCRβ
CD5
G
CD6
Naive
Activated

Figure S4. CD5 expression on naive CD4$^+$ T cells identifies protein differences that are stable with age and that predict divergence of functional differences post-activation. Related to Figure 4. (**A**) GSEA of gene signatures in mutant/constitutively-active form of NFAT1 overexpressing CD8$^+$ T cells (GSE64409) and T cell exhaustion during chronic viral infection (GSE30962) in CD5^{lo} or CD5^{hi} naïve CD4$^+$ T cells. (**B**) Protein expression of key markers (CD6, Ly6C, FolR4, and CD73) were measure at 8 days (neonate), 19-22 weeks, and 34 weeks old B6 mice in CD5^{lo} versus CD5^{hi} naïve CD4$^+$ T cells. Representative histograms, protein expression (black bars, relative to CD5^{hi} cells at each age group), and fold enrichment (grey bars) are shown. Data summarized from 2 independent experiments; data points are from individual mice (n=2-9); error bars represent standard deviation; grey hashed line represents the average fold enrichment of protein expression in neonates. (**C**) Fold enrichment of protein expression of markers as in (B) except mice were RAG2^{GFP} and cells were gated on non-recent thymic emigrants (GFP$^-$). Data summarized from 1 experiment; data points are from individual mice (n=2-3); error bars are standard deviations. (**D**) Protein expression (relative to sorted CD5^{hi} naïve CD4$^+$ T cells) measured by flow cytometry in sorted 15% CD5^{lo} and CD5^{hi} naïve CD4$^+$ T cells pre- and 24 hours post-activation with anti-CD3/CD28. Activated cells were gated on CD44^{hi}CD25$^+$ or CD44^{hi}CD62L$^-$. Data summarized from 2-4 independent experiments; individual data points are from 2-5 pooled mice (n=4-13); lines denote group means. (**E**) 3-dimensional single-cell flow cytometry analysis for TOX expression among naïve CD4$^+$ T cells pre- and 24 hours post-activation. Color scale represents MFI of TOX in bins of at least 10 cells across the full protein expression spectrum of CD5 (x-axis) and TCRβ (y-axis). (**F,G**) Mutual information analysis to determine the ability of CD5 or TCRβ to predict SHP-1 (F) or CD6 (G) expression in naïve CD4$^+$ T cells pre- and 24 hours post-activation.

Statistics: Wilcoxon matched-pairs signed rank test (D). *p<0.05, **p<0.01, ***p<0.001.

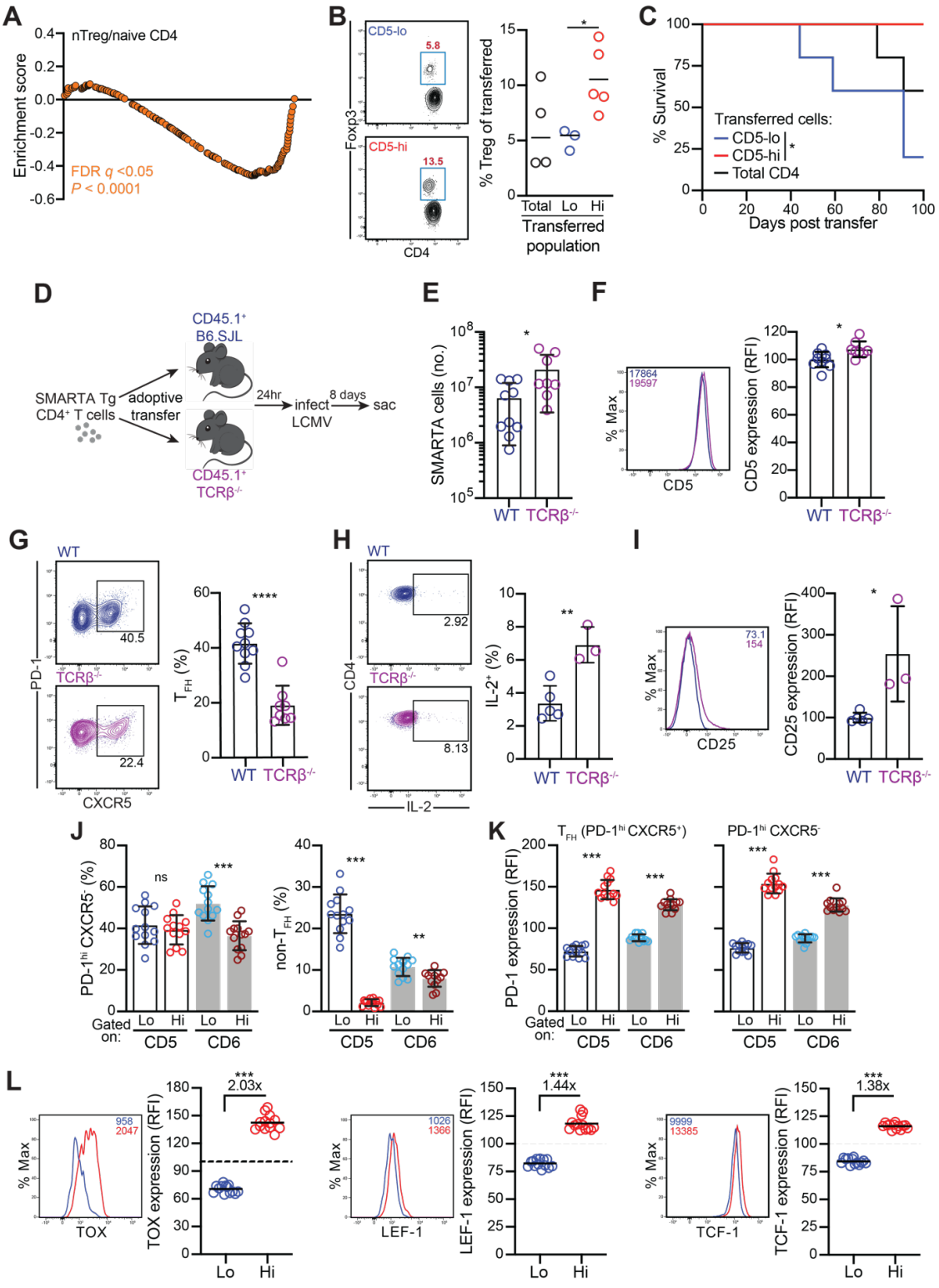


Figure S5. Differences in CD5 and CD6 expression predicts biased effector differentiation in CD4$^+$ T cells. Related to Figure 5. (**A**) GSEA of natural (nTreg) gene signature (GSE14308) in CD5^{lo} or CD5^{hi} naïve CD4$^+$ T cells. (**B**) Treg conversion of total and sorted 25% CD5^{lo} and CD5^{hi}

naïve $CD4^+$ T cells 91 days post-adoptive transfer into $TCR\beta^{-/-}$ recipients. Representative flow cytometry plots (left, numbers are percent of cells within indicated gate) and summary plot (right, lines denote group means, data points are individual mice (n=3-5)) of percent Tregs among transferred cells. (**C**) Survival of mice adoptively transferred with total and sorted 25% $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells. All data in B and C are summarized from 1 experiment. (**D-I**) Congenically labelled ($CD45.2^+$) SMARTA transgenic $CD4^+$ T cells were transferred into WT or $TCR\beta^{-/-}$ mice that were infected 1 day later with LCMV, as shown in schematic of experimental setup (D). Total number of activated ($CD44^{hi}$) transferred SMARTA cells on day 8 post-infection (E). CD5 protein expression (relative to SMARTA cells transferred into WT mice) of transferred SMARTA cells (F) and percent T_{FH} of transferred SMARTA cells 8 days post-infection (G). Percent IL-2^+ (H) and CD25 protein expression (I) (relative to SMARTA cells transferred into WT mice) of transferred SMARTA cells 8 days post-infection. All data in D-I are summarized from 1-2 independent experiments; each data point in summary graphs is from an individual mouse (n=3-10). Numbers in representative flow cytometry plots are percent of cells within indicated gate (G, H) or mean fluorescent intensity (F, I). (**J ,K**) Activated ($CD44^{hi}$) $CD4^+$ T cells isolated on day 8 post-infection were gated on the top and bottom 15% CD5- or CD6-expressing cells and the percent PD-$1^{hi}CXCR5^-$ and non-T_{FH} (PD-$1^{lo}CXCR5^-$) determined (J) and PD-1 RFI (relative to total T_{FH} cells) of gated cells from **Fig. 5F** (K).All data in J,K are summarized from 2 independent experiments; each data point is from an individual mouse (n=13), error bars represent mean ± S.D. (**L**) TOX, LEF-1, and TCF-1 protein expression (relative to total T_{FH} cells) in T_{FH} cells from gated $CD5^{lo}$ and $CD5^{hi}$ $CD4^+$ T cells. Numbers in representative flow cytometry plots are mean fluorescent intensity. Data is summarized from 2 independent experiments; each data point is from an individual mouse (n=13).

Statistics: Mann-Whitney test (B), Mantel-Cox test (C), unpaired t test (E-I), and Wilcoxon matched-pairs signed rank test (J-L). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, ns= non-significant.

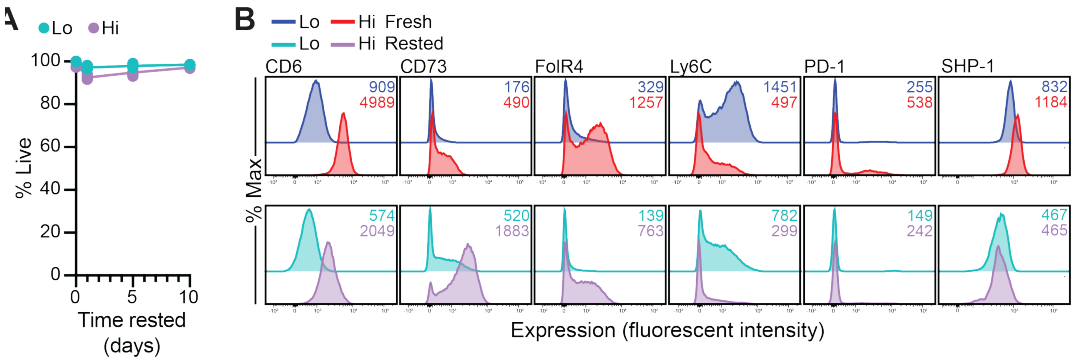


Figure S6. Withdrawal of naïve CD4$^+$ T cells from self-pMHC identifies proteins that do not rely on continuous self-pMHC interactions. Related to Figure 6. (**A**) Naïve CD4$^+$ T cells were sorted into 15% CD5^{lo} and CD5^{hi} populations, rested in dissociated culture in the presence of IL-7 for up to 5 days, and cell viability determined. Data is summarized from 1 experiment; each data point is from an individual mouse (n=4 for day 0, 1, and 5; n=2 for day 10). (**B**) Representative flow cytometry plots of proteins from sorted CD5^{lo} and CD5^{hi} naïve CD4$^+$ T cells rested in dissociated culture in the presence of IL-7 for 5 days. Numbers in top right of histograms indicate mean fluorescence intensity in fresh or rested CD5^{lo} and CD5^{hi} populations.

Statistics: Two-way ANOVA with Sidak's multiple comparisons showed a significant effect (p<0.01) between CD5^{lo} and CD5^{hi} naive CD4$^+$ T cells at Day 1, but not at Day 0, 5, or 10 (A).

Table S1. Counts matrix for transcripts in scRNA-seq analysis performed on individual naïve $CD4^+$ T cells. CD5 surface protein expression (cd5.level) was measured by flow cytometry for each cell. Only cells that passed quality control are included. Related to Figure 1.

Table S2. List of top 5% most variable genes among naïve $CD4^+$ T cells from single cell RNA-seq. Log2 normalized counts, gene expression variance, and gene classification are indicated. Related to Figure 1.

Table S3. List of non-redundant genes enriched in the defined GO terms associated with T cell activation used for PCA analysis of naïve $CD4^+$ T cells from single cell RNA-seq. Related to Figure 1.

Table S4. List of DEGs identified in comparing $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells (FDR < 0.01) using bulk RNA-seq. Log2 fold changes, absolute fold changes, log2 CPMs and p-values are indicated. Gene group classifications defined in Figure 6 are indicated in column G. Related to Figures 3 and 6.

Table S5. Counts matrix for all detected transcripts in bulk RNA-seq in sorted fresh and rested $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells. Related to Figures 3 and 6.

PREFACE TO CHAPTER 3

In Chapter 2 we found that the naive $CD4^+$ T cell population possesses remarkably diverse transcriptional and chromatin heterogeneity that is based on strength of self-reactivity. We show that $CD5^{hi}$ (high self-reactive) naive $CD4^+$ T cells are more likely to become Tregs and T_{FH} cells. We next sought to understand what low self-reactive T cells contribute to the T cell response during infection. We examined the role of $CD5^{lo}$ (low self-reactive) $CD4^+$ T cells during chronic infection. We also investigated the impact of acute and chronic pathogen control when a T cell population has on average fewer low pMHC-reactive T cells.

CHAPTER 3: T CELLS WITH LOW ANTIGEN BINDING STRENGTH GENERATED BY N-NUCLEOTIDE DIVERSITY AID IN CHRONIC PATHOGEN CONTROL

On the cover:

T cells express one of a possible 10^{20} T cell receptors (depicted as individual dots) where this diversity establishes a T cell population that spans a spectrum of strength in reactivity to peptide (encoded from red for high-reactivity to blue for low-reactivity). In this chapter, Rogers et al. show that when the T cell population is missing low-reactive receptors (blue dots) chronic pathogen control is delayed. Image concept inspired by the work of Yayoi Kusama. Artwork by Dakota Rogers.

T cells with low antigen binding strength generated by N-nucleotide diversity aid in chronic pathogen control

Dakota Rogers[1,2], Hassan Jamaleddine[1], Geneviève Perreault[2], Dhanesh Patel[1,2], and Judith N. Mandl[1,2,3]*

[1]Department of Physiology, McGill University, Montreal, Quebec, Canada.
[2]McGill University Research Centre on Complex Traits, Montreal, Quebec, Canada.
[3]Department of Microbiology and Immunology, McGill University, Montreal, Quebec, Canada.

*Corresponding author: Judith Mandl
Email: judith.mandl@mcgill.ca

Unpublished. Incorporated into two manuscripts:

Jamaleddine S.*, Rogers D.*, Perreault G., Mandl J.N., Khadra A. (2022). Chronic infection control relies on T cells with lower foreign antigen binding strength generated by N-nucleotide diversity. Biorxiv. https://doi.org/10.1101/2022.06.26.497644. *Under Review.*

* Equal contribution

Textor J., Buytenhuijs F., Rogers D., Mallet Gauthier É., Sultan S., Wortel I.M.N., Kalies K., Fähnrich A., Pagel R., Melichar H.J., Westermann J., Mandl J.N. (2022). Machine learning analysis of the T cell receptor repertoire identifies sequence features that predict self-reactivity. BioRxiv. https://doi.org/10.1101/2022.11.23.517563.

3.1 Summary

The breadth of pathogens a T cell can respond to is determined by the T cell receptors (TCRs) present within a person's repertoire. While N-nucleotide insertions at V(D)J gene segment junctions by terminal deoxynucleotidyl transferase (TdT) during recombination of the αβTCR contribute more than 90% of TCR sequence diversity, the importance of these additions has remained elusive. Here, we probe whether TCRs with higher diversity, via TdT, differentially contribute to acute or chronic pathogen control. Using CD5 expression as a surrogate for TCR signal strength to self-peptide MHC (pMHC), which correlates with foreign pMHC reactivity, we found that $CD5^{lo}$ (low self-reactive) $CD4^{+}$ T cells dominated in the responding T cell population during chronic infection. Supporting the idea that N-nucleotide additions skew TCRs, on average, to have lower pMHC-reactivity, pre-selection TdT-KO thymocytes had increased levels of CD5. Moreover, when TdT-KO was restricted to the T cell compartment, chronic viral control was delayed, while acute viral infections were unaffected. Together these data suggested that TCR diversity imparted by N-nucleotide additions preferentially produced TCRs of lower reactivity to pMHC and that these T cells support chronic pathogen control.

Keywords

Acute and chronic infections; Lymphocytic choriomeningitis virus; *Cryptococcus neoformans*; T cell receptor affinity; DNA nucleotidylexotransferase, terminal deoxynucleotidyl transferase; CD5**.**

3.2 Introduction

T cells have the remarkably difficult task of recognizing and responding to any possible pathogen. T cells use a specialized receptor called the T cell receptor (TCR) to recognize the pathogen-derived peptides through presentation on major histocompatibility complexes (pMHC) [334, 335]. To make a T cell population that is diverse enough to recognize all potential foreign peptides each T cell clone expresses a unique αβTCR that is generated by somatic recombination of V and J gene segments for the α-chain and V, D, and J gene segments for the β-chain [105]. Further increasing TCR diversity, the DNA polymerase, terminal deoxynucleotidyl transferase (TdT), adds non-templated (N) nucleotides into the junctions of the V(D)J segments [105, 297, 336, 337]. N-nucleotides account for ~90-95% of the TCR diversity [109, 338, 339]. Together, TdT in combination with somatic recombination of the V(D)J gene segments generates one of a possible 10^{15}-10^{20} TCRs [113-115]. Many of the TCR recombination events produce receptors that are either non-functional or react to pMHC too strongly, which are removed from the T cell population during development in the thymus [116]. Still, the mature naive T cell repertoire consists of millions of TCRs representing individual clonotypes [113, 117].

Upon encounter with a pathogen, many antigen-specific naive T cells within the T cell repertoire are mobilized to contribute to the response. Each T cell clonotype within the anti-pathogen response recognizes the same or different pMHCs and the strength with which the TCRs bind their ligands can differ by several orders of magnitude [166, 340]. The heterogeneity among responding T cells in their TCR binding strength to pMHC, here referred to as pMHC reactivity, correlates with their self-pMHC reactivity and differences in T cell function. For $CD4^+$ T cells, differences in pMHC reactivity impact effector lineage choices and memory cell differentiation [1, 4, 341, 342]; while among $CD8^+$ T cells pMHC reactivity correlates with lytic activity, proliferation, and memory cell formation [340]. However, to what extent the distribution of pMHC-reactivity for pathogen-responding T cells affects how quickly a pathogen is cleared remains incompletely understood.

Currently, there are few techniques that permit an in-depth analysis of the responding T cell population and how pMHC-reactivity evolves over the course of the infection. First, two-dimensional binding assays measure the force required to dissociate the TCR interaction from its cognate pMHC as a T cell is physically pulled from the antigen presenting cell (APC) [343]. Second, surface plasmon resonance measures a change in mass on a sensor-surface with

immobilized pMHC as TCR molecules are flown over the pMHC to bind [344]. A third common method uses pMHC tetramers to identify antigen-specific T cells, although pMHC tetramers primarily identify T cell clonotypes with higher pMHC-reactivity and may miss low-reactive T cells [166]. However, all three techniques assess only one TCR at a time and require prior knowledge of the epitope that is recognized by the T cell clonotype of interest. Thus, by focusing on only a subset of known epitope-specific T cells, the contribution of the remaining pathogen-specific T cell population is missed.

While TdT is the primary source of TCR diversity that establishes a polyclonal T cell population, the benefit of added N-nucleotide from TdT has remained unclear. TdT-knockout (KO) mice, while possessing severely diminished TCR diversity, do not appear to be more susceptible to infection nor are their T cell responses impaired during acute infection [336]. Despite normal control of acute pathogens in TdT-KO mice, TdT is highly conserved across all jawed vertebrates [345, 346], suggesting an evolutionarily important role of TdT which has yet to be fully elucidated. One hypothesis proposes that by adding N-nucleotides, TdT generates non-germline TCRs that expand the scope of peptides a T cell can recognize outside the germline-encoded repertoire but may consequently have lower pMHC-reactivity because of the introduced bases [138]. Consistent with this idea, the neonatal T cell repertoire is first produced in the absence of TdT and has increased pMHC reactivity [296]. In addition, TdT-KO T cells undergo positive selection in the thymus more efficiently, suggesting they may have, on average, TCRs that have higher pMHC reactivity [144]. Moreover, during influenza A virus infection, $CD8^+$ T cells specific to the HA_{518} epitope from TdT-KO mice were ~10 time more sensitive to antigen stimulation compared to the same epitope-specific $CD8^+$ T cells from wildtype (WT) mice, further suggesting that TdT-independent TCRs may have higher pMHC reactivity [347]. However, in settings where there is chronic antigen stimulation, such as in long-term infections and cancers, T cells with high pMHC reactivity are more likely to undergo exhaustion, characterized by epigenetic remodelling leading to dysfunction, compared to T cells with lower pMHC reactivity [306, 348, 349]. As such, a possible benefit for TdT is to allow for better control of chronic infection when germline TCRs that, on average, have higher pMHC reactivity may otherwise become exhausted.

Here we investigated whether differences in the strength of pMHC reactivity within the polyclonal T cell population impact the clearance of acute and chronic pathogens. We show that naive CD4$^+$ T cells with low self-reactivity, and thus lower foreign reactivity, preferentially contributed to the responding T cell pool during chronic infection. We also demonstrated that mice lacking T cells with TdT expression, thus having a T cell population that is presumably shifted towards higher-reactive T cells, have delayed chronic pathogen control. Taken together, our data suggest that one evolutionary benefit for TdT may be to produce T cells with lower pMHC reactivity and in doing so provide better control of pathogens during chronic infection.

3.3 Results

Strength of self-reactivity inversely correlates with *Dntt* expression in CD4$^+$ T cells

Previous work from our group and others had unexpectedly identified *Dntt* as one of the most enriched genes in low self-reactive T cells [3, 139-141]. *Dntt* encodes for the non-templated DNA polymerase TdT which has been proposed to generate TCRs during development that, on average, have lower self-pMHC reactivity [138]. To investigate the relationship between naive CD4$^+$ T cell self-reactivity and *Dntt* expression, we first sought to understand whether the difference in *Dntt* expression was due to residual mRNA expression in recent thymic emigrants (RTE) where *Dntt* might still be expressed. We used CD5 protein expression as a surrogate to measure the strength of self-reactivity for naive T cells, as has been suggested by prior studies [1, 2, 140, 141, 172]. Using a RAG2-GFP reporter mouse, we found that GFP$^+$ RTEs were enriched for CD5^{hi} cells rather than CD5^{lo} cells (**Figure S1A**). This is consistent with RTEs possessing increased sensitivity to pMHC interactions [238]. Regardless, we excluded GFP$^+$ RTEs when sorting naive CD4$^+$ T cells with low, mid, and high CD5 expression to assess *Dntt* mRNA expression by qPCR (**Figures S1B and S1C**). In addition, we sorted double negative (DN), double positive (DP), and CD4$^+$ single positive (SP) thymocytes to assess *Dntt* gene expression during development (**Figures S1D and S1E**). Consistent with previously published data, we found that *Dntt* expression was inversely correlated with CD5 protein in naive CD4$^+$ T cells (**Figures 1A and 1B**). Of note, even though CD5^{lo} naive CD4$^+$ T cells express higher *Dntt* than CD5^{hi} cells, their *Dntt* expression was ~40-fold lower than in the developing DN and DP thymocytes (**Figure 1A**). In addition to greater *Dntt* gene expression in peripheral naive CD4$^+$ T cells, CD5^{lo} cells had greater chromatin accessibility for the *Dntt* locus (**Figure 1C**). Notably, the labelled intergenic

peak synergizes with the *Dntt* promoter to activate transcription specifically in T cells, but not B cells [350]. To gain a better understanding of the expression dynamics of TdT (encoded by *Dntt*) we further assessed protein expression at the different stages of thymocyte development (**Figure S1F**). As thymocytes progressed through development, TdT expression was first detected at the DN2 and DN3 stages, which is when the TCRβ chain is rearranged [351]. There was a transient decrease in TdT expression in the DN4 stage then re-expression in the DP stage where the TCRα chain is rearranged, albeit not to the same level as the DN2-DN3 stages. TdT protein was no longer detected in $CD4^+$ SP thymocytes. Together, these data led us to hypothesize that greater *Dntt* expression during thymic development may produce TCRs with more N-nucleotide insertions that increase the diversity of the TCR repertoire but may also, on average, have lower self-pMHC reactivity.

To determine whether TdT-dependent TCRs are indeed biased for lower self-reactivity, we assessed CD5 expression throughout thymic development in TdT-KO mice. In theory, if TdT is used to make low self-reactive TCRs, the TdT-KO mice should possess a T cell repertoire that is enriched for higher self-reactive T cells, thus expressing more CD5. Pre-selection DP TdT-KO thymocytes indeed expressed higher levels of CD5 compared to WT thymocytes (**Figure 1D**). However, following selection in the DP stage as well as the $CD4^+$ and $CD8^+$ SP stages, CD5 expression was no longer different between WT and TdT-KO thymocytes. These data suggest that following selection CD5 expression may be renormalized by different bounds of positive and negative selection within the TdT-KO mice. Indeed these data are supported by published work that shows TdT-KO T cells are more efficiently positively selected, implying that they have, on average, higher reactivity to pMHC [144].

T cells with low pMHC reactivity dominate during chronic infection

During acute infection, antigen-responding $CD4^+$ T cells are dominated by high self-reactive ($CD5^{hi}$) $CD4^+$ T cells [1]. However, it remains unclear what role the low self-reactive $CD4^+$ T cells play during infection. We asked if low self-reactive T cells, which may on average have more N-nucleotide insertions, are important for chronic pathogen control, as has been previously hypothesized [138]. To test this hypothesis, we sorted naive $CD4^+$ T cells on the bottom and top 20% CD5 expressing cells, denoted as $CD5^{lo}$ and $CD5^{hi}$ respectively. The two sorted naive T cell populations (identified by congenic markers, $CD45.1^+$ or $Thy1.1^+$) were mixed at a 1:1 ratio

and adoptively transferred into CD45.2$^+$ Thy1.2$^+$ recipient mice that were infected 3 days prior with *Cryptococcus neoformans* (**Figures 2A and 2B, S2A and S2B**). Notably, exposure to the fungal pathogen, *C. neoformans*, establishes a persistent pulmonary infection [352]. Following 20 days post-transfer (23 days post-infection) the *C. neoformans* responding (activated) CD4$^+$ T cell population in the lung was dominated by CD5^{lo} CD4$^+$ T cells, outnumbering CD5^{hi} CD4$^+$ T cells 4-fold (**Figures 2C and 2D**). Of note, within the infected mice, CD5 expression increased slightly upon activation; however, a CD5^{lo} CD4$^+$ T cell remained CD5^{lo}, and a CD5^{hi} CD4$^+$ T cell remained CD5^{hi} as has been previously shown during LCMV-Arm infection (**Figures S2C and S2D**) [140]. Our data is consistent with prior data from both antigen-specific CD4$^+$ and CD8$^+$ T cells showing that T cells with lower TCR reactivity predominate during chronic infection [253, 353].

Absence of low pMHC reactive T cells generated by N-nucleotide diversity delays chronic pathogen control

We next sought to address whether a T cell repertoire devoid of TdT expression, which presumably disproportionately generates T cells with lower pMHC-reactivity, was important for pathogen control. TdT-KO mice are not impaired in their ability to control acute infection [336]; however, it is not known whether these mice can control chronic pathogens. We infected total TdT-KO mice with the chronic pathogen LCMV-Cl13 and assessed viral loads as the infection progressed. Across early (day 6), middle (day 25), and late (day 45) stages of infection viral titers were not impacted in TdT-KO mice, relative to WT mice (**Figure 3A**). Given that B cells also utilize TdT to diversify their B cell receptors (BCR) [354], it is possible that changes in the B cell repertoire in total TdT-KO mice impact viral replication. Thus, we restricted TdT deficiency to the T cell compartment alone, with normal TdT expression in B cells. To do this, we generated mixed bone marrow chimeras where TCRβ-KO bone marrow (no T cells, WT B cells) was mixed 1:1 with either WT bone marrow (WT T cells, WT B cells) leading to reconstitution with WT T and B cells, or with JH and TdT double KO bone marrow (TdT KO T cells, no B cells) developing TdT-KO T cells and WT B cells (**Figure 3B**). Prior to bone marrow cell transfer, we verified the cells were mixed at 1:1 ratio from each of the donors (CD45.1$^+$ or CD45.2$^+$) (**Figure S3A**). Consistent with prior data from total TdT-KO mice [336], there were no differences in viral titers during acute LCMV-Arm infection between the WT and TdT-KO T cell chimera groups in the serum at day 2 or in the spleen at day 6 (**Figure 3C**). However, during the chronic phase of LCMV-

Cl13 infection at days 26-28 and days 40-45 TdT-KO T cell chimeras had higher viral loads at 3.05-fold and 3.15-fold respectively (**Figure 3D**). Since some endogenous hematopoietic stem cells are retained following irradiation in bone marrow chimera recipients, which may skew the T cell response to virus, we repeated the chimera setup using TCRβ-KO mice as recipients (**Figure S3B**). Like the chimeras made with CD45.1$^+$ WT mice as recipients, TCRβ-KO CD45.1$^+$ recipients also had viral titers 3.74-fold greater than WT chimeras (**Figure S3C**). Intriguingly, antigen-responding TdT-KO T cells expressed higher levels of PD-1 expression during the chronic phase of LCMV-Cl13 infection (days 38-50) (**Figure S3D**). These data together suggest that T cells with TCRs that are not dependent on N-nucleotide insertions from TdT are, on average, higher affinity and may be more prone to exhaustion. Ultimately our findings highlighted the importance of low-affinity TCRs in the control of chronic pathogens.

3.4 Discussion

As T cells rearrange their TCRs during development TdT incorporates random nucleotides at the V(D)J gene junctions that are responsible for an explosion in TCR diversity outside the germline-encoded TCRs [297, 339, 355]. Yet, despite TdT being highly evolutionarily conserved [337], the impact these non-germline TCRs have on T cell responses has remained incompletely understood. Our data showed that T cells with lower self-reactivity, as determined by low CD5 protein levels, express more *Dntt*, the gene encoding TdT, and may consequently have longer TCRs. These low self-reactive CD4$^+$ T cells subsequently predominated in the pathogen-responding T cell pool, consistent with a previous hypothesis [138]. When fewer low-reactive T cells are present in the T cell repertoire, as we predicted in TdT-KO mice, there was a delay in chronic, but not acute, pathogen control. Importantly, our data corroborated previous experiments showing that TdT-KO mice were not more susceptible to acute pathogens such as Vesicular Stomatitis, Sendai, Influenza A, and LCMV-WE [336, 347]. Together, our work shows a previously undescribed benefit for TCR repertoire diversification by TdT in the control of chronic infection.

In settings where there is chronic antigen stimulation, high-affinity T cells are more prone to undergo permanent epigenetic remodelling, referred to as epigenetic scarring, which results in irreversible T cell exhaustion [306, 349, 356]. We found that TdT-KO T cells have increased expression of PD-1 during the chronic phase of infection. While PD-1 alone is not sufficient to

show that TdT-KO T cells are more prone to exhaustion, it does indicate that TdT-KO T cells receive greater TCR signals [357]. It will be interesting to evaluate the dynamic epigenetic changes that occur in TdT-KO T cells in chronic infection settings, where there is continuous antigen stimulation, to determine whether these cells indeed have an increased propensity to become exhausted.

The impact of a TdT-deficient repertoire in pathogen control is not only relevant for understanding the evolutionary conservation of TdT in jawed vertebrates but is also relevant to neonatal immunity, given that the T cell repertoire in the early stages of life is generated in the absence of TdT. In mice, TdT is first detected at 3-5 days after birth, while in humans TdT expression is detected at ~20 weeks of pregnancy [358, 359]. As a consequence, the neonatal TCR repertoire lacks N-nucleotide additions, TCR sequences are shorter, are more likely to be found across multiple individuals (public), and are more cross-reactive [145, 360, 361]. Indeed, the neonatal T cell repertoire has been shown to have higher self-reactivity [296], which is consistent with TdT-KO T cells. With these parallels in mind, it will be important to consider to what extent the neonatal versus the adult T cell repertoire differs in their responses to pathogens.

TdT-KO mice have been an enigma since first being characterized [336]. We found that peripheral TdT-KO naive $CD4^+$ or $CD8^+$ T cells do not express higher levels of CD5, contrary to what would be expected if cells had high pMHC-reactivity as we predict with the TdT-KO T cell repertoire. However, prior to thymic selection, TdT-KO thymocytes did express higher CD5, suggesting that while the T cell repertoire may be enriched for TCRs of high pMHC-reactivity, thymic selection re-normalizes the bounds of CD5 expression to be comparable to WT levels. Given these caveats, we additionally turned to computational modelling to characterize the T cell responses during acute and chronic infections and to modify a theoretical T cell repertoire and track pMHC-reactivity distributions over time during infection [252]. In line with our data, the model predicted a decrease in pMHC-reactivity in the pathogen-responding T cell population during chronic infection *in silico* [252]. Additionally, when a theoretical TCR repertoire was shifted to be higher in pMHC-reactivity the model predicted a prolonged chronic infection [252]. These experimental and computational data show complementary findings for how TdT impacts the T cell repertoire in response to pathogens.

Using bulk TCRβ-sequencing and machine learning to asked whether there were differences in TCR usage for low ($CD5^{lo}$) and high ($CD5^{hi}$) self-reactive cells, our group recently

found that $CD5^{lo}$ naive $CD4^+$ T cells had longer TCRs, on average, compared to $CD5^{hi}$ cells [143]. Intriguingly, sequences from $CD5^{lo}$ TCRs were slightly enriched in pre-selection DP thymocytes, suggesting that low-reactive TCRs may take longer to receive the necessary signals to be positively selected; conversely, $CD5^{hi}$ TCR sequences were enriched in the $CD4^+$ SP thymocytes [143]. Indeed, experimental data has shown that weakly self-reactive thymocytes take longer to undergo positive selection [137]. Additionally, this new study found that $CD5^{hi}$ TCR sequences were preferentially identified in acute LCMV infection and in the memory response [143], consistent with previously published data [1]. Lastly, $CD5^{lo}$ TCR sequences were overrepresented in chronic LCMV infection [143]. These data provide the first direct evidence that low self-reactive T cells expressing higher *Dntt* indeed have longer TCR sequences, at least for the β-chain.

In summary, our work provides new insights into the differential contribution of T cells with low- and high-reactivity to pMHC during acute and chronic immune responses. We show that by adding N-nucleotides to the TCR, TdT plays an essential role in setting the strength of pMHC reactivity. These data are important given recent findings in both mice and humans, for both $CD4^+$ and $CD8^+$ T cells, that differences in TCR signal strength play a critical role in effector cell outcomes [140-142]. Overall, our findings suggest that low pMHC-reactive T cells may be best suited for immune settings where chronic antigen exposure occurs such as other chronic infections, cancers, and autoimmunity.

3.5 Acknowledgements

We would like to thank the animal facility staff at McGill University for their excellent care of our animal colony, C. Stegen and J. Leconte at the Cell Vision Core Facility for cell sorting, F. Buytenhuijs and J. Textor at Radboud University for key intellectual input and for conducting separate analyses not included in this manuscript, and P. Artusa for initial experiments at an earlier stage of the project that were also not included. We would like to thank A. Khadra for key intellectual input. D.R. was supported by a Frederick Banting and Charles Best Canada Graduate Doctoral Award (CIHR) and a Tomlinson Doctoral Fellowship (McGill). J.N.M. is a Canada Research Chair for Immune Cell Dynamics. This work was supported by NSERC grant #2016-03808 to J.N.M.

Author Contributions

J.N.M. conceived the study, supervised, led the study implementation, and acquired financial support for this project. D.R. and J.N.M. designed and performed experiments, with key experimental input from H.J., G.P., and D.P. D.R. analysed the experiments, prepared the figures, and wrote the manuscript with input from J.N.M.

Declaration of Interests

The authors declare no competing interests.

3.6 Figures

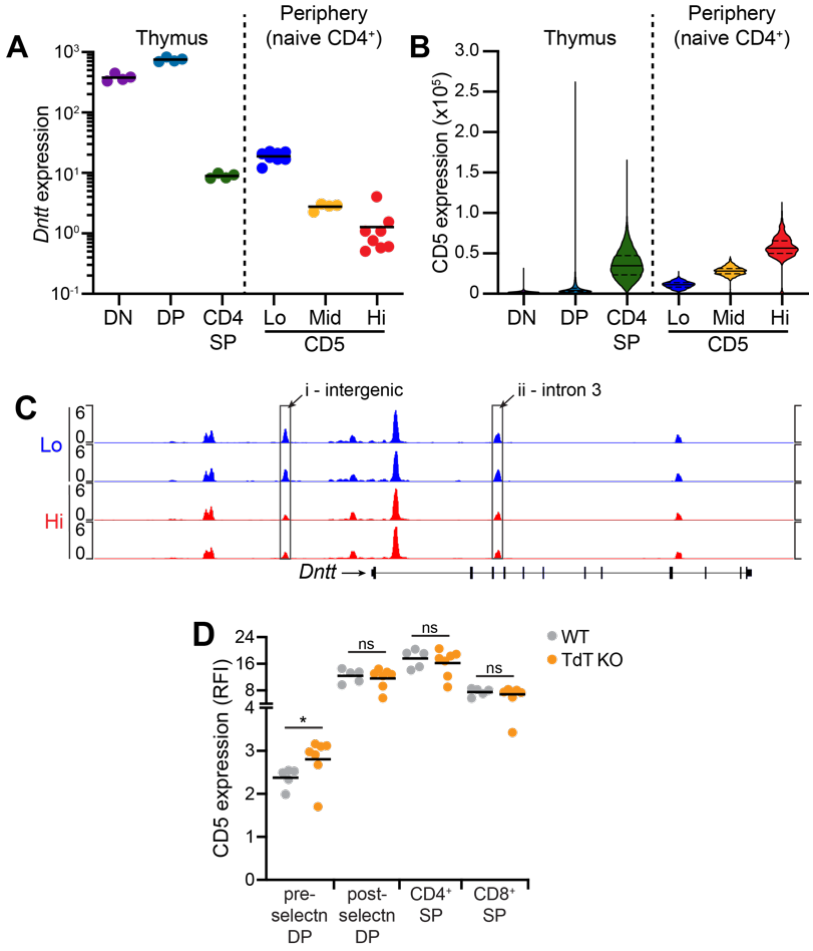


Figure 1. *Dntt* expression inversely correlates with CD5 expression in naive CD4$^+$ T cells. (**A**) *Dntt* mRNA expression during thymic development and in peripheral naive CD4$^+$ T cells. Thymocytes were sorted for double negative (DN; CD4$^-$CD8$^-$TCRβ$^-$), double positive (DP; CD4$^+$CD8$^+$TCRβ$^-$), and single positive (SP; CD4$^+$TCRβ$^+$) populations. Naive CD4$^+$ T cells were sorted by 20% lowest, middle, and highest CD5 expressing populations (CD5^{lo}, CD5^{mid}, and CD5^{hi} respectively). n=4-8 mice. (**B**) CD5 protein expression (mean fluorescent intensity [MFI]) throughout development and in peripheral naive CD4$^+$ T cells. DN, DP, and CD4$^+$ SP thymocytes were defined by gating within the total thymocyte population. Peripheral naive CD4$^+$ T cells were sorted into CD5^{lo}, CD5^{mid}, CD5^{hi} populations. (**C**) ATAC-seq signal profile of *Dntt* from 2 biological replicates of sorted CD5^{lo} and CD5^{hi} naive CD4$^+$ T cell samples. (**D**) CD5 protein expression levels in pre-selection (CD3$^-$CD69$^-$) and post-selection (CD3$^+$CD69$^+$) DP (CD4$^+$CD8$^+$) thymocytes as well as CD4$^+$ and CD8$^+$ SP thymocytes in wild-type (WT) and TdT-KO mice. CD5

relative fluorescent intensity (RFI) is determined relative to the DN ($CD4^-CD8^-$) thymocyte population. n=5-7 mice per group.

Statisitics: Mann-Whitney test (D). ns = not-significant, *$p<0.05$.

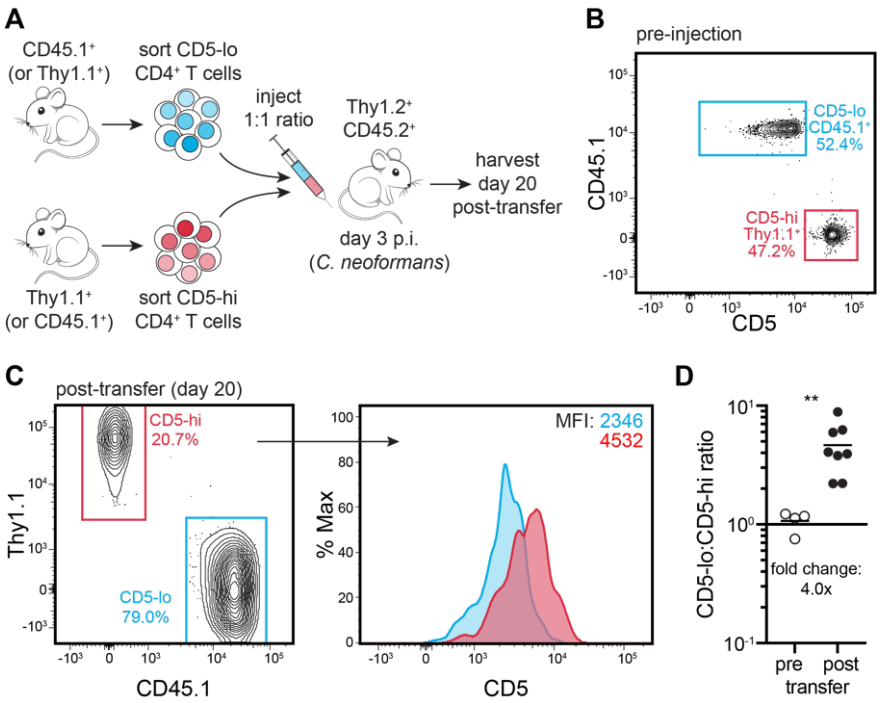


Figure 2. CD5^{lo} CD4^{+} T cells predominate during chronic infection. (**A**) Schematic of experimental approach for adoptive cell transfer. Sorted CD5^{lo} and CD5^{hi} naive CD4^{+} T cells (either Thy1.1^{+} or CD45.1^{+}) were mixed at a 1:1 ratio and adoptively transferred into congenic CD45.2^{+} Thy1.2^{+} recipient mice. Recipient mice were infected with *C. neoformans* 3 days prior to adoptive transfer. (**B**) Representative flow cytometry plot of sorted CD5^{lo} CD45.1^{+} and CD5^{hi} Thy1.1^{+} co-transferred naive T cell populations prior to adoptive transfer. (**C**) Representative flow cytometry plot (left) of activated (CD44^{hi}) CD5^{lo} CD45.1^{+} and CD5^{hi} Thy1.1^{+} 20 days post-transfer (23 days post-infection) with corresponding CD5 protein expression levels shown as a histogram (right) and MFI indicated in blue (CD5^{lo}) and red (CD5^{hi}) text. (**D**) Ratio of co-transferred activated CD5^{lo} to CD5^{hi} CD4^{+} T cells pre-infection and 20 days post-transfer. n=8 recipient mice.

Statistics: Wilcoxon rank sum test (D). **p<0.01.

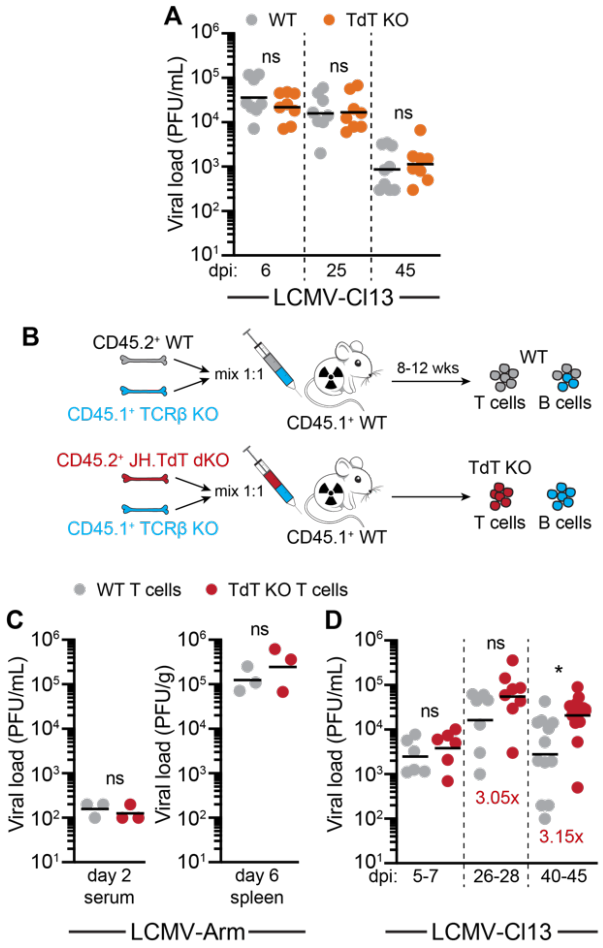


Figure 3. Loss of TdT expression in T cells impairs chronic pathogen control. (**A**) LCMV-Cl13 viral loads in the serum of WT or TdT-KO mice in early (6 days), mid (25 days), and late (45 days) stages of viral replication. n=8-9 mice. (**B**) Schematic for the generation of mixed bone marrow chimeric mice possessing WT B cells and either WT T cells (irradiated mice reconstituted with 1:1 ratio of bone marrow from WT mice and TCRβ KO mice) or TdT KO T cells (irradiated mice reconstituted with 1:1 ratio of bone marrow from JH and TdT double KO mice and TCRβ KO mice). (**C**) LCMV-Arm viral loads in serum at day 2 (left) and spleen at day 6 (right) post-infection in WT or TdT KO T cell chimeras. n=3 mice. (**D**) LCMV-Cl13 viral loads in serum at day 5-7, 26-28, and 40-45 days post-infection of chimeras with WT or TdT KO T cells. Fold changes indicate viral loads in TdT KO T cell mice relative to WT T cell mice. n=6-12 mice.

Statisitics: Mann-Whitney test (A), unpaired t-test (C), and Kruskal-Wallis test (D). ns = not significant, *$p<0.05$.

3.7 Materials and Methods

Mice

C57BL/6, congenic CD45.1^{+}, congenic Thy1.1^{+}, RAG2-GFP [305], and TCRβ$^{-/-}$ mice [303] were purchased from Jackson Laboratories (Bar Harbor, ME). The TdT$^{-/-}$ mice were shared by Dr. A. Feeney (The Scripps Research Institute) [338] and the JH$^{-/-}$ mice were shared by Dr. J. Fritz (McGill) [362]. All mice were on a C57BL/6 background, bred in-house and experiments were performed at 6-12 weeks of age with both males and females. Animal housing, care, and research were in accordance with the Guide for the Care and Use of Laboratory Animals and all procedures performed were approved by the McGill University Animal Care Committee.

Pathogens stocks and infections

LCMV. LCMV-Arm and -Cl13 strains were propagated from stocks provided by Dr. M. Richer (University of Indiana) on BHK-21 or L929 cells (ATCC). Briefly, virus was added at MOI 0.01, incubated for 90 minutes in serum-free media at 37ºC in 5% CO_2, then topped up with complete media for incubation for another 48 hours before harvesting the supernatant. BHK-21 cells were cultured in EMEM supplemented with 0.1% penicillin/streptomycin, 1% L-glutamine, 1% non-essential amino acids, 1% sodium pyruvate, and 10% FBS and maintained at 37ºC in 5% CO_2. L929 cells were cultured in RPMI supplemented with 10% FBS, 1% L-glutamine, and 1% penicillin/streptomycin. Mice were infected with $2x10^5$ plaque forming units (PFU) of LCMV-Arm by intra-peritoneal injection or $2x10^6$ PFU by intravenous injection for LCMV-Cl13 as previously described [306, 363]. Mice were bled by either tail artery or cardiac puncture into sterile eppendorf tubes kept on ice, blood was spun down at 12,000 rpm for 10 minutes and serum was aliquoted and frozen for viral titer determination. Spleens were collected in 1% RPMI and weighed. Spleens were placed in Lysing Matrix D tubes (MP Biomedicals) and homogenized with a MagNA Lyser (Roche) at 6000 rpm for 40 seconds. Spleen homogenate was then spun down at 12,000 rpm for 10 minutes at 4ºC and supernatant was transferred to a separate sterile tube and re-spun at 12,000 rpm for 10 minutes at 4ºC then aliquoted and frozen for viral titer determination. Viral titers (stocks used, mouse serum, and tissue samples) were determined by plaque assay with Vero cells [364]. Briefly, Vero cell monolayers were infected with 100 µL of serially diluted serum (1 in 10 dilutions from 10^{-1} to 10^{-7}) and incubated for 90 minutes at 37ºC in 5% CO_2. Infected cells were then overlaid with 1% agarose (Wisent) and incubated for 3 days at 37ºC in 5% CO_2. A

second agarose overlay supplemented with 1% neutral red was then added and cells incubated for 24 hours at 37°C in 5% CO_2, after which plaques were counted.

C. neoformans. The H99 strain was provided by K. Kwon-Chung (NIH). Frozen stocks (-80°C) were prepared in 15% glycerol from fresh cultures from a YPD agar plate. Three days before infection, *C. neoformans* was scraped from the frozen stock and streaked onto a YPD agar plate. One day prior to infection, a single colony was inoculated and incubated for 12-16 hours at 30°C with continuous agitation in YPD broth. Immediately before infection, C. neoformans was resuspended in cold PBS. Mice were then anesthetized with isoflurane and infected by intrapharyngeal aspiration with $5x10^3$ colony forming units (CFU) in 20 µL of PBS. Mice were sacrificed and tissue collected 23 days after infection. *C. neoformans* CFUs were determined as previously described [352].

Lymphocyte and thymocyte isolation

For the *C. neoformans* infections, prior to harvest, an intravascular stain using 2.5 µg anti-CD45 (30F11) was performed as previously described [365]. Infected lungs were harvested in cold PBS and minced with scissors. Lung was then digested at 37°C with agitation for 30 minutes in digestion buffer (1 mg/mL collagenase D, 50 U/mL DNase I, 1 mg/mL hyaluronidase, 1% L-glutamine, 1% pen/strep in RPMI). Tissue was then passed through a 100-µm filter with PBS supplemented with 1% FBS and resuspended in 10mL of 37% Percoll in RPMI. Samples were centrifuged at 3000 rpm for 20 minutes at 22°C. ACK lysis buffer (Life Technologies) was added for 3 minutes, samples were washed with PBS, refiltered, and resuspended in complete RPMI (10% FBS, 1% L-glutamine, 1% HEPES buffer, 1% pen/strep, 1% sodium pyruvate, 1% non-essential amino acids, 0.1% 2-mercapto-ethanol 1000X solution). Dilution of single-cell suspensions at 1:10 in Trypan Blue and manual counting of live cells (Trypan Blue-negative) on a hemacytometer was used to determine total cell counts.

Spleen, thymi, and peripheral lymph nodes (inguinal, axillary, brachial, and mesenteric) were collected and passed through a 70-µm filter with 1% RPMI (1% penicillin/streptomycin, 1% L-glutamine, and 1% FBS). ACK lysis buffer (Life Technologies) was added for 3 minutes, samples were washed with PBS, refiltered and resuspended in 1% RPMI. For experiments with thymocytes, thymi were harvested and passed through a 70µm filter with 1% RPMI (1% FBS, 1% L-glutamine, and 1% pen/strep). Dilution of single-cell suspensions at 1:10 in Trypan Blue and

manual counting of live cells (Trypan Blue-negative) on a hemacytometer was used to determine total cell counts.

Bone marrow chimeras

Bone marrow was collected from the femurs and tibias of donor mice (either $JH^{-/-}$, $TdT^{-/-}$, $TCR\beta^{-/-}$, or B6 WT) by flushing the marrow from the bones with cold 1% RPMI. Bone marrow cells were then passed through a 70-µm filter with 1% RPMI and red blood cells were lysed with ACK lysis buffer (Life Technologies) and cell counts were determined as above. Recipient mice (either B6 $CD45.1^{+}$ or $TCR\beta^{-/-}$ $CD45.1^{+}$) were irradiated twice at 550 rads 3 hours apart and reconstituted with a 1:1 mix of $2.5x10^{6}$ cells per genotype that were injected i.v. within 5 hours of the first irradiation. To establish the WT chimera (WT T cells and B cells) B6 and $TCR\beta^{-/-}$ bone marrow cells were mixed at equal proportions; to make the T cell restricted $TdT^{-/-}$ chimeras (WT B cells) bone marrow cells from $JH^{-/-}$ $TdT^{-/-}$ mice were mixed 1:1 with $TCR\beta^{-/-}$ bone marrow cells. Recipient mice were given neomycin water (2g/L) 2 days prior to bone marrow transfer and kept on the antibiotic water for 2 weeks following transfer. Mice were used 8-12 weeks post-irradiation and bone marrow reconstitution.

Flow cytometry

Samples were incubated in Fixable Viability Dye (AF780, Life Technologies) diluted in PBS for 20 minutes at 4°C. Extracellular antibodies were diluted in FACS buffer (2% FBS and 5mM EDTA in PBS) with Fc Block (Life Technologies) and incubated for 30 minutes at 4°C. For intracellular staining, samples were then incubated in FoxP3 Transcription Factor Fixation/Permeabilization Concentrate and Diluent (Life Technologies) for 30 minutes at 4°C. Intracellular antibodies were diluted in Permeabilization Wash Buffer (Life Technologies) and samples were incubated for 30-60 minutes at 4°C. Directly conjugated antibodies used were as follows: TCRb (H57-597), CD4 (RM4.5), CD8a (53-6.7), CD5 (53-7.3), Foxp3 (FJK-16 s), CD44 (IM7), CD62L (MEL-14), CD25 (PC61.5), CD45.1 (A20), CD45.2 (104), PD-1 (29F.1A12), B220 (RA3-6B2), Ly6C (HK1.4), B220 (RA3-6B2), CD11b (M1/70), CD11c (N418), F4/80 (T45-2342), NK1.1 (PK126), and TdT (19-3). For all flow cytometry experiments, cells were acquired using an LSRFortessa (BD Bioscience) and analyzed with FlowJo software (BD Bioscience).

Cell sorts

Cell sorts were performed as previously described and detailed briefly below.

Sorts for quantitative RT-PCR (qRT-PCR): Thymocytes and lymphocytes from C57BL/6 or C57Bl/6.SJL (CD45.1$^+$) congenic mice were isolated in single-cell suspension as described. Total isolated thymocytes were directly stained for sorting. Total lymphocytes were magnetically enriched for CD4$^+$ T cells (Stemcell EasySep mouse CD4$^+$ T cell enrichment kit). Cells were then incubated in Fixable Viability Dye and subsequently stained with surface antibodies for 1 hour at 4°C. Sorts were performed on either a FACS Aria Fusion, Aria III, or Aria II SORP (BD Bioscience). All cell populations were sorted to >90% purity for bulk populations. Thymocytes were sorted on dump negative and TCRβ^{lo}, CD8$^-$, and CD4$^-$ for DN; TCRβ^{lo}, CD8$^+$, CD4$^+$ for DP; and TCRβ^{hi}, CD8$^-$, CD4$^+$ for CD4$^+$ SP. Peripheral naive CD4$^+$ were sorted on dump negative, TCRβ$^+$, CD4$^+$, CD8$^-$, RAG2$^{GFP-,}$ CD44$^-$, and 20% CD5^{lo}, CD5^{mid}, or CD5^{hi}. Dump channel included CD25 (periphery only), B220, CD11b, CD11c, F4/80, Ly6C, and NK1.1.

Sorts for adoptive cell transfers: Lymphocytes from Thy1.1$^+$ or CD45.1$^+$ congenic mice were isolated in single-cell suspension as described. Spleens and lymph nodes (inguinal, axillary, brachial, mesenteric, and cervical) were pooled from 15 mice for each congenic marker. Cells were then magnetically enriched for total CD4$^+$ T cells (Stemcell EasySep CD4$^+$ T cell Enrichment kit or Miltenyi Biotec CD4$^+$ T cell Isolation Kit). Enriched CD4$^+$ T cells were stained with surface antibodies for 1 hour at 4^{o}C. Naive CD4$^+$ T cells were sorted on singlets, CD4$^+$, CD8$^-$, CD62L^{hi}, CD44^{lo}, and 20% CD5^{lo} or CD5^{hi}. Sorts were performed on a FACS Aria III (BD Bioscience). All cell populations were sorted to >90% purity.

RNA extraction and quantitative real-time RT-PCR

RNA from sorted DN, DP, and SP thymocytes and 20% CD5^{lo}, CD5^{mid}, CD5^{hi} naive CD4$^+$ T cells was extracted using RNAqueousTM-Micro Total RNA Isolation Kit (Life Technologies) and cDNA converted using High-Capacity cDNA Reverse Transcription Kit (Life Technologies). qPCR analysis was performed with TaqManTM Gene Expression Master Mix (Life Technologies) and TaqManTM Gene Expression Assay (FAM, Dntt, Mm00493500_m1, Life Technologies). Average Ct values across technical duplicates were determined for *Dntt* and fold change was calculated as Log2-transformed 2ΔΔ*Ct* values relative to the expression of the housekeeping gene *Gapdh*.

Adoptive cell transfers

C. neoformans infection. All donors and recipients were sex matched. 15 $CD45.1^+$ or $Thy1.1^+$ mice were used as donors to obtain a total of 8-14x10^6 cells for each of 20% $CD5^{lo}$ and 20% $CD5^{hi}$ cells sorted as described above. Sorted $CD5^{lo}$ and $CD5^{hi}$ were then mixed in a 1:1 ratio. 4-7x10^6 of each sorted population was adoptively transferred into $CD45.2^+$ $Thy1.2^+$ recipients that were infected with 5x10^3 CFU of *C. neoformans* 3 days prior to transfer. Cells were isolated from the lungs of recipient mice 20 days post-transfer.

ATAC-seq library preparation, sequencing, and visualization

Two independent biological replicates of $CD5^{lo}$ and $CD5^{hi}$ naïve $CD4^+$ T cells were sorted, counted, and 1x10^5 nuclei pelleted. ATAC-seq analysis was done as previously described [140]. ATAC-seq signal profiles were visualized on Integrative Genomics Viewer (IGV) [323].

Statistical analyses of experimental data

Group comparisons were performed using Prism V9 (GraphPad). The cut-off for significance considered was $p<0.05$. Information about implemented statistical tests and sample sizes for individual experiments is provided in the figure legends.

3.8 Supplemental Information

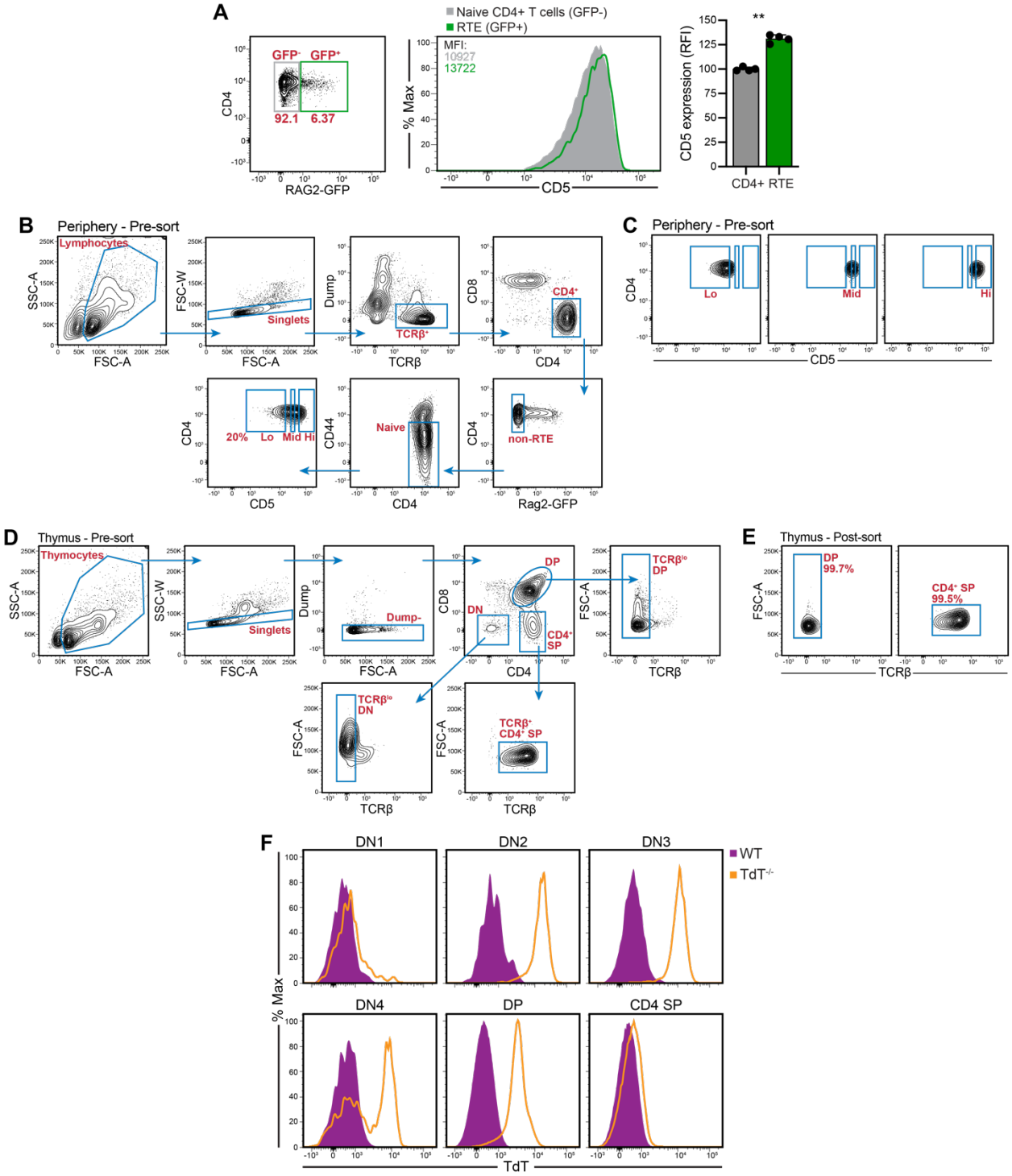


Figure S1. Greater *Dntt* expression in CD5^{lo} naive CD4^{+} T cells is not explained by recent thymic emigrants. (**A**) Representative flow cytometry plot (left) and histogram (middle) and summary data (right) of CD5 protein expression on recent thymic emigrants (RTE) compared to non-RTE naive CD4^{+} T cells. Relative fluorescence intensity (RFI) is relative to non-RTE naive

$CD4^+$ T cells RTEs were identified as GFP^+ naive $CD4^+$ T cells using a RAG2-GFP reporter. n=4 mice. (**B-E**) Gating strategy for pre- (B) and post- (C) sort peripheral naive $CD4^+$ T cells, as well as pre- (D) and post- (E) sort thymocytes used for qPCR and CD5 protein expression. (**F)** Representative flow cytometry plots of TdT protein expression in DN1 ($CD4^-CD8^-CD44^+CD25^-$), DN2 ($CD4^-CD8^-CD44^+CD25^+$), DN3 ($CD4^-CD8^-CD44^-CD25^+$), DN4 ($CD4^-CD8^-CD44^-CD25^-$), DP ($CD4^+CD8^+$), and $CD4^+$ SP thymocytes in WT and TdT-KO mice.

Statistics: Paired t-test (A). **p<0.01.

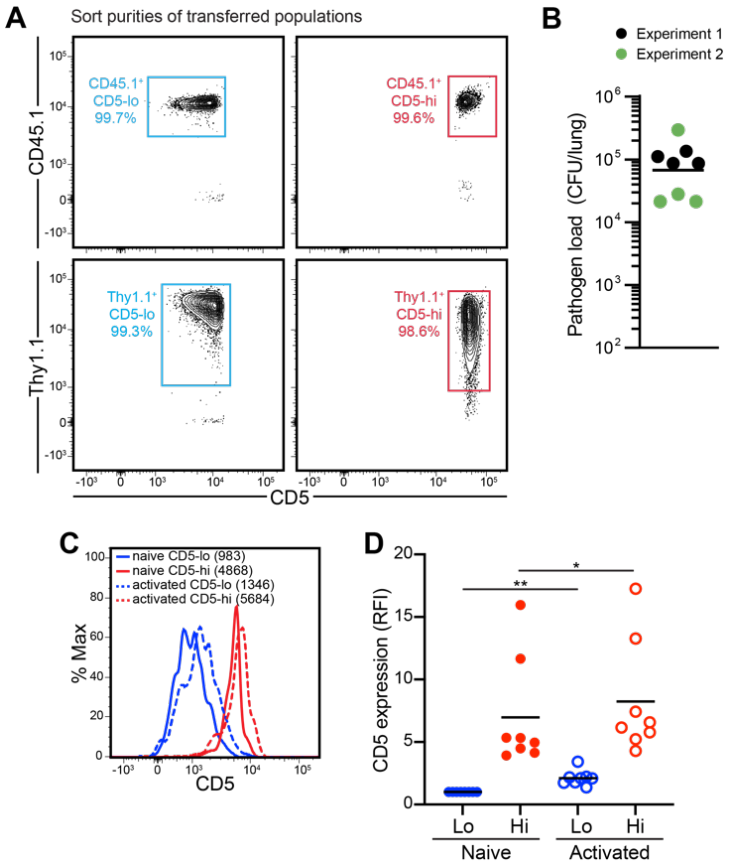


Figure S2. CD5 expression is maintained post-activation in sorted CD5^{lo} and CD5^{hi} CD4^{+} T cells. (**A**) Flow cytometry plots showing percent purities of sorted CD45.1^{+} or Thy1.1^{+} CD5^{lo} and CD5^{hi} naive (CD62L^{+}CD44^{lo}) CD4^{+} T cells prior to mixing and co-transfer into *C. neoformans* infected mice. (**B**) *C. neoformans* fungal burdens in the lungs of infected mice. Data is from 2 independent experiments. n=8 mice. (**C**) Representative histograms of naive (CD44^{lo}) and activated (CD44^{hi}) sorted CD5^{lo} and CD5^{hi} CD4^{+} T cells that were adoptively co-transferred. The numbers in parentheses are the MFIs. (**D**) CD5 expression (RFI is relative to naive CD5^{lo} cells) of transferred naive and activated sorted CD5^{lo} and CD5^{hi} CD4^{+} T cells. Data is from 2 independent experiments. n=8 mice.

Statistics: Wilcoxon matched-pairs signed rank test (D). *p<0.05, **p<0.01.

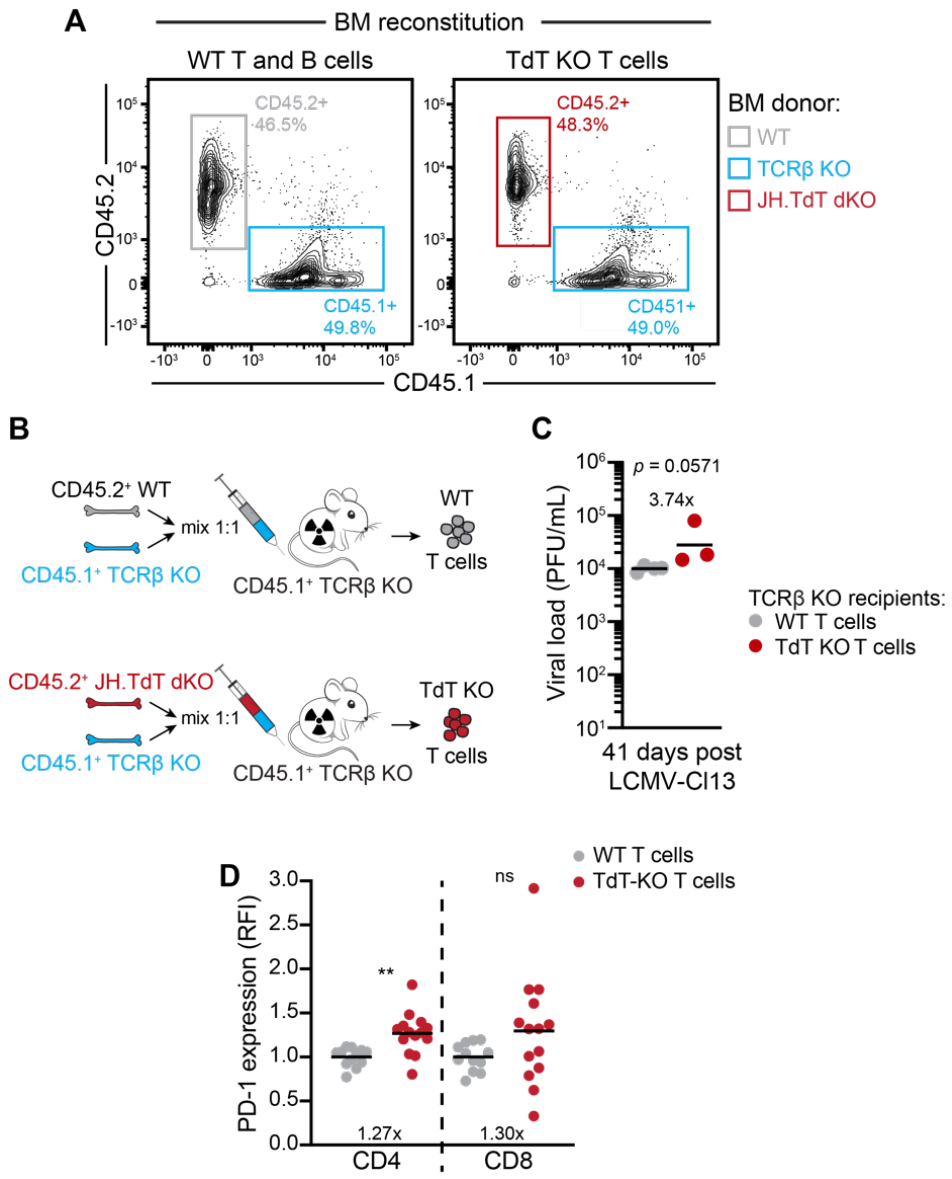


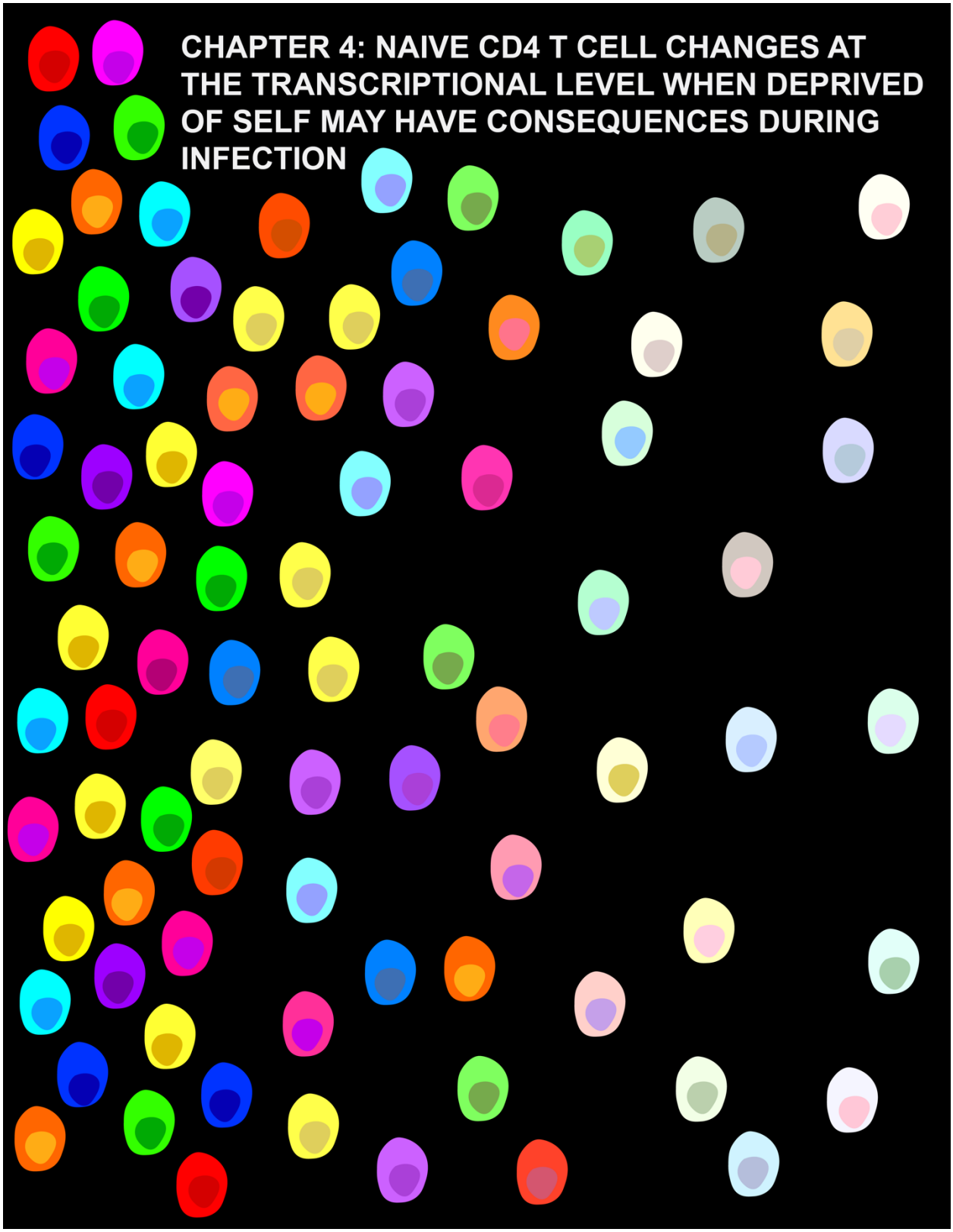
Figure S3. Activated TdT-KO T cells have increased PD-1 expression during chronic infection. (**A**) Representative flow cytometry plots showing the percent of mixed bone marrow cells from each set of donor mice. Donor bone marrow were CD45.1^{+} cells from TCRβ-KO mice and CD45.2^{+} cells from either WT or JH x TdT double KO mice. (**B**) Modified schematic from Fig. 3B for mixed bone marrow chimera generation using irradiated TCRβ-KO mice as bone marrow recipients. (**C**) Serum LCMV-Cl13 viral titers 41 days post-infection of TCRβ-KO recipient mice with either WT or TdT-KO T cells. (**D**) PD-1 expression on activated (CD44^{hi})

$CD4^+$ and $CD8^+$ T cells from WT or TdT-KO T cell chimeras at 38-50 days post-infection with LCMV-Cl13. N=12-14 mice.

Statistics: two-tailed Wilcoxon rank sum test on geometric means (C), Mann-Whitney test (D). ns = not significant, **p<0.01.

PREFACE TO CHAPTER 4

In Chapters 2 and 3, we showed that naive $CD4^+$ T cells with low and high self-reactivity possess biases in function post-activation. Common among all naive T cells, they continuously receive self-signals as they migrate throughout the body. When self-signals are removed naive T cells can survive for several weeks; however, it is not known why. In Chapter 4, we delve into what transcriptional changes occur when naive $CD4^+$ T cells are deprived of self-signals. Additionally, we investigate physiological settings where self-deprivation may occur.


CHAPTER 4: NAIVE CD4 T CELL CHANGES AT THE TRANSCRIPTIONAL LEVEL WHEN DEPRIVED OF SELF MAY HAVE CONSEQUENCES DURING INFECTION

On the cover:

Strength of reactivity to self-ligands establishes a heterogeneous naive $CD4^+$ T cells population (coded by the rainbow of colors). In this chapter, Rogers et al. begin to unravel why continuous tonic self-signals are essential for naive T cell maintenance. When naive T cells are deprived of self-ligands their response to activation is impaired (as depicted by a decrease in color vibancy from left to right) and ultimately begin to die (as depicted by a decease in number of cells from left to right). Artwork by Dakota Rogers.

Naive CD4$^+$ T cell changes at the transcriptional level when deprived of self may have consequences during infection

Dakota Rogers[1,2] and Judith N. Mandl[1,2,3]*

[1]Department of Physiology, McGill University, Montreal, Quebec, Canada.
[2]McGill University Research Centre on Complex Traits, Montreal, Quebec, Canada.
[3]Department of Microbiology and Immunology, McGill University, Montreal, Quebec, Canada.

*Corresponding author: Judith Mandl
Email: judith.mandl@mcgill.ca

Unpublished

4.1 Summary

T cells use their T cell receptors (TCRs) to discriminate self from foreign by recognizing peptides presented in the context of major histocompatibility complexes (pMHC). In the periphery, naive T cells rely on continuous self-pMHC interactions to maintain their readiness in response to pathogens. When naive T cells are deprived of self-pMHC interactions they gradually die, yet why naive T cells can still survive for several weeks without these signals remains incompletely understood. Here, we probe the protein and transcriptional changes that occur when naive T cells are deprived of self-pMHC interactions. Self-deprivation of naive $CD4^+$ T cells at steady-state leads to dynamic changes in protein expression and upregulation of transcriptional networks like autophagy, lipid metabolism, response to starvation, and catabolic metabolism that implicate adaptations that may promote cell survival. We found that pathogen-non-specific naive $CD4^+$ and $CD8^+$ T cells undergo transient infection-induced self-deprivation (IISD), which impairs their response to subsequent activation. Additionally, we observed that IISD is dependent on type-I interferon signaling. Together, our work provides insights into how naive $CD4^+$ T cells cope with self-deprivation and show that self-deprivation can occur during infection.

Keywords

Self-pMHC, self-ligands, self-deprivation, steady-state, infection-induced self-deprivation, naive T cells, infection, type-I interferons

4.2 Introduction

The fundamental role of T cells is to differentiate self from foreign. To do this, T cells use the T cell receptor (TCR) to recognize peptides presented in the context of major histocompatibility complexes (MHC) on antigen presenting cells (APCs) [366]. T cells have a life-long relationship with self-ligands, which begins in the thymus, where self-peptide MHC (self-pMHC) interactions are necessary for T cell development [116]. At homeostasis, referred to here as steady-state, mature naive T cells continue to require self-pMHC interactions in the periphery. Naive T cells will eventually die if they are deprived of self-pMHC interactions. In fact, self-deprived naive $CD4^+$ T cells have a half-life of 30-45 days, while naive $CD8^+$ T cells have a half-life of 16-19 days [221, 223]. In contrast, the loss of IL-7 signals, another critical survival signal for naive T cells, leads to a more rapid loss of naive T cells [224]. However, the reason why naive T cells gradually decline during self-deprivation, instead of quickly dying, is still not clear.

Removal of self-pMHC signals also impairs naive T cell responses to cognate antigen. Removal of self-pMHC leads to the rapid dephosphorylation of the TCRζ chain and ZAP70 [248], critical TCR signalling molecules. The dephosphorylation of the TCR cascade components desensitizes the T cell and stunts TCR signaling, proliferation, and IL-2 production upon foreign pMHC recognition [248]. Additionally, the disruption of self-signals in $CD8^+$ T cells leads to the upregulation of the $CD8^+$ coreceptor that increases the sensitivity of cells to low-affinity ligands [249]. As such, self-ligands provide signals that are essential for maintaining naive T cell readiness for encounters with foreign antigens.

Upon infection, pathogens are first detected by innate immune and non-immune cells through pattern recognition receptors (PRRs), which recognize microbial products and trigger the induction of cytokines and chemokines that shape the adaptive immune response [367]. Some of the first cytokines to be produced are the type-I interferons (IFN), IFNα and IFNβ [368]. Historically, type-I IFNs are essential for viral control; however, it has been shown that parasites, bacteria, and fungi can also induce type-I IFN signaling [368]. Within a few hours of infection, there is an initial wave of IFNβ followed by IFNα that typically lasts for ~48 hours [369, 370]. Type-I IFN signaling through the IFNαβ receptor (INFAR) can have broad-ranging effects on cells; notably, signaling can impact peptide presentation by APCs and can directly signal on T cells. In APCs, IFNα/β stimulates the generation of the immunoproteasome in both immune and non-immune cells [216, 371], which may lead to changes in peptide presentation on MHC

molecules. For example, following PRR recognition, dendritic cells (DCs) preferentially present foreign peptides on MHCs over self peptides [372, 373]. For $CD4^+$ T cells, type-I IFN signals can impact lineage differentiation, survival, and expansion; and in CD8 T cells can promote proliferation [374, 375]. In addition, type-I IFN signaling has recently been implicated in naive T cell homeostasis [241, 244]; however, we are only recently beginning to unravel the role these signals have on naive T cells.

Here we investigated why naive $CD4^+$ T cells can survive for well over a month without self-pMHC signals and explore the physiological settings where naive T cells can experience self-deprivation. We show that self-ligand deprivation leads to the upregulation of transcriptional networks that may be implicated in promoting naive $CD4^+$ T cell survival. In addition, we show that non-pathogen-specific naive $CD4^+$ and $CD8^+$ T cells experience transient self-ligand deprivation during infection that impairs their subsequent response to activation. We also show that infection-induced self-deprivation (IISD) is dependent on type-I interferon signaling. Taken together, we shed light on the potential pathways that naive T cells may utilize to survive in settings where self-ligands may be limiting, such as during infections.

4.3 Results

Naive $CD4^+$ T cells induce cell survival transcriptional networks upon in vitro self-deprivation

In our previous work, we showed that self-pMHC interactions maintain some, but not all, of the transcriptional and protein heterogeneity within the naive $CD4^+$ T cell population [140]. We thus sought to assess how the removal of self-pMHC interactions, also referred to as self-deprivation, affected protein expression dynamics in the total naive $CD4^+$ T cell population. To deprive naive $CD4^+$ T cells of self-ligand interactions, we cultured them *in vitro* with media supplemented with IL-7 (**Figure 1A**), a critical survival signal for T cells [376]. Culturing naive T cells in this way removes the APCs that present self-pMHC to the naive T cells. Over 5 days of self-deprivation there were dynamic changes in protein expression on naive $CD4^+$ T cells, at least for the markers we assessed (**Figure 1B**). CD5, a surrogate marker for the strength of self-signals [172], CD6, and folate receptor 4 (FolR4) expression progressively decreased over time. The decrease in CD5 expression has been previously shown for naive $CD4^+$ T cells that have been deprived of self-pMHC interactions [2, 140]. For Ly6C, where expression inversely correlates with

access to self-signals on $CD4^+$ T cells [179], there was a transient decrease in protein levels that then progressively increased over time, consistent with a previous study where naive $CD4^+$ T cells were deprived of self-signals *in vivo* [179]. The protein CD73, an ectonucleotidase that converts extracellular AMP to adenosine [377], saw a rapid rise in protein expression that was maintained for the full 5 days. CD98, a neutral amino acid transporter [378], saw an initial decrease in expression which normalized back to baseline levels by day 5. Lastly, CD4, a protein that has been shown to be unaffected by self-deprivation [2], remained constant across the 5 days. Together, the surface expression changes of these markers defined a phenotype on naive $CD4^+$ T cells that we used to investigate self-deprivation on naive T cells. We next sought to ensure that the changes in protein expression were indeed due to self-deprivation and not other environmental cues like physical interactions and chemokine signaling in the secondary lymphoid organs. Naive $CD4^+$ T cells were adoptively transferred into $MHCII^{-/-}$ mice and the self-deprivation protein phenotype was assessed 5 days post-transfer (**Figure 1C**). Consistent with the *in vitro* data, CD5, CD6, and FolR4 expression decreased while Ly6C and CD73 expression increased following 5 days of *in vivo* self-deprivation (**Figure 1D**).

We next sought to investigate whether self-deprivation led to changes in the transcriptomes of naive $CD4^+$ T cells. To address this, we utilized a previously published naive $CD4^+$ T cell dataset [140]. Briefly, naive $CD4^+$ T cells with low ($CD5^{lo}$) or high ($CD5^{hi}$) self-pMHC reactivity were sorted and either fresh cells were taken for RNA sequencing, or naive cells were cultured as described above to deprive cells of self-pMHC interactions for 22 hours then sequenced. To define a common T cell signature, we took the overlap in differentially expressed genes (DEGs) for the fresh $CD5^{hi}$ versus self-deprived $CD5^{hi}$ and fresh $CD5^{lo}$ and self-deprived $CD5^{lo}$ comparisons, thus genes found across the whole T cell population. The T cell signature consisted of 4150 DEGs (**Figure S1A**). We further classified the common T cell signature as genes with increased expression in fresh (positive fold-change [FC] values, 2085 genes) and increased gene expression during self-deprivation (negative FC values, 2063 genes) (**Figure S1B**). Gene ontology (GO) analysis showed that pathways enriched in fresh cells included activation, proliferation, and response to pathogens suggesting maintenance of T cell sensitivity to activation (**Figure S1C**). The GO terms found in the fresh cells are consistent with previous data showing that continuous self-pMHC signals promote naive T cell sensitivity to cognate antigen encounter [248]. Interestingly, when naive T cells are deprived of self-signals in culture, they rapidly (within 22

hours) upregulate transcriptional networks for autophagy, catabolic metabolism, lipid metabolism, and response to starvation (**Figure 1E**). Together, these data suggest that naive $CD4^+$ T cells may adapt to self-pMHC signal deprivation by upregulating pathways that promote cell survival.

Infection induces a transient state of self-deprivation on naive T cells

We next asked if there were other physiological settings where self-deprivation may occur. This may be particularly important during infection where APCs upregulate different proteomes [372, 373], which may reflect in changes in the peptides presented to T cells where foreign peptides are biased for presentation over self-peptides. Additionally, several pathogens can promote the downregulation of MHC molecules as a way of hiding from the immune system [379-381], but consequently may also impact naive T cell access to self-peptides. We thus asked whether infection would lead to self-deprivation in the non-responding (non-pathogen specific) naive, hereafter referred to simply as naive, $CD4^+$ and $CD8^+$ T cell pools. Indeed, 8 days post-infection with lymphocytic choriomeningitis virus-Armstrong (LCMV-Arm), the naive $CD4^+$ and $CD8^+$ T cells had decreased expression of CD5, CD6, and FolR4, as well as increased Ly6C (**Figures 2A, 2B, and S2A**). These protein expression changes are consistent with naive $CD4^+$ T cells deprived of self at steady-state. Additionally, the IISD phenotype was found in multiple model pathogens including influenza A (PR8 strain), Murine cytomegalovirus (MCMV), and LCMV-Clone 13 (Cl13); however, it was not observed when mice were infected with the fungal pathogen *Cryptococcus neoformans* (**Figure 2C**). Importantly, IISD was not dependent on the dose of the pathogen as mice infected with the normal dose ($2x10^5$ PFU) and high dose ($2x10^6$ PFU) of LCMV-Arm experienced the same degree of decreased CD5 expression (**Figure S2B**). Additionally, when infected with the acute pathogen LCMV-Arm CD5 expression decreased by 8 days post-infection but rebounded to steady-state levels by day 12 (**Figure 2D**). Conversely, mice infected with the chronic pathogen, LCMV-Cl13, had decreased levels of CD5 47 days post-infection, albeit to a lesser degree than 8 days post-infection with LCMV-Cl13 (**Figures 2C and 2E**). This suggested that naive T cells experience IISD so long as the pathogen has not been cleared and foreign antigen remains present in the system.

Removal of self-signals at steady state impairs naive T cell responses to cognate antigen [248]. We next asked if this was also true for non-responding naive $CD4^+$ T cells that are experiencing IISD. To address this, we enriched naive $CD4^+$ T cells from LCMV-Arm infected

mice 8 days post-infection and subsequently activated them with α-CD3 and α-CD28 for 24 hours (**Figure 2F**). Both CD5 and SHP-1 expression were impaired compared to uninfected $CD4^+$ T cells (**Figure 2G and 2H**), consistent with the impaired response of $CD4^+$ T cells during steady-state self-deprivation. In summary, infection leads to transient depivation of self-ligand signals that impairs the pathogen-non-specific naive T cell pool's ability to respond to subsequent activation.

Infection-induced self-deprivation relies on type-I interferon signaling

As part of the early response to pathogens, type-I IFNs are expressed by almost all cells in the body in response to stimulation of the PRRs by microbial products. The type-I IFNs promote a shift in the peptide processing machinery potentially impacting the peptides available for naive T cells [368, 371]. We thus explored whether the IISD, which was most prominent during viral infection, was due to IFNα/β signaling. $IFNAR1^{-/-}$ mice, which cannot receive IFNα or IFNβ signals, were infected with LCMV-Arm and IISD was assessed 8 days post-infection. The changes in protein expression during IISD for decreased CD5 and FolR4 and increased Ly6C on naive $CD4^+$ and $CD8^+$ T cells in WT mice were not recapitulated in the $IFNAR^{-/-}$ mice (Figure 3). These data identify that IISD is dependent on type-I interferon signals.

4.4 Discussion

The importance of self-pMHC interactions in maintaining the naive T cell population has been understood for over 25 years [4]. Yet, we know surprisingly little about how naive T cells can still survive for several weeks when self-MHC signals are removed. Here, we show that self-ligand interactions provide signals that upregulate gene networks maintaining naive $CD4^+$ T cell readiness for activation, consistent with previously published data [248]. Conversely, the deprivation of self-signals leads to the upregulation of transcriptional networks that implicate pathways that may promote cell survival such as autophagy, lipid metabolism, starvation response, and catabolic metabolism in naive $CD4^+$ T cells. Importantly, we found that infections induce a state of self-deprivation in non-responding naive T cells that impair T cell response to subsequent activation, a finding that is strikingly similar to steady-state self-deprivation. We also uncover the surprising requirement of type-I IFN signals to induce IISD in the naive T cells. Together, our work implicates several cell survival pathways that naive $CD4^+$ T cells may utilize to promote survival in times when self-pMHC signals are limiting such as in the context of infection.

To begin to uncover to what extent IISD mimics steady-state self-deprivation, we will perform RNA-sequencing on self-deprived naive $CD4^+$ and $CD8^+$ T cells at both steady-state and during infection. If IISD upregulates the same survival transcriptional networks as steady-state self-deprivation such as autophagy or response to starvation we can then target these pathways and assess their impact on the naive T cell pool *in vivo*. Using autophagy as an example, if the naive T cells are prevented from recycling cellular components via autophagy during self-deprivation it might be expected that the naive T cells would die faster. Thus, parsing the pathways that naive T cells use to prolong their survival during self-deprivation will help us to understand why T cells can survive without self-pMHC, at least short-term, even though self-signals are essential for maintaining naive T cell readiness for pathogen.

It will be important to determine which signals induce IISD. In infection, we show that type-I IFN signals are required for naive T cells to experience self-deprivation. However, whether these type-I IFN signals induce self-deprivation by changing the peptides presented by APCs or IFNα/β are signaling directly onto the T cells remains to be seen. Additionally, increasing evidence has shown that type-I IFN signals at steady-state may also be important for maintaining naive T cell homeostasis [241, 244], but more work is required to fully understand the impact of these signals and whether they are also required for steady-state self-deprivation.

Together, our data provide novel insights into the physiological settings where self-ligand deprivation may occur and the consequences on T cell function. Ultimately, our work may have implications in elucidating how infections change the naive T cell repertoire and affect T cell responses to subsequent infection, which may be particularly important when considering co-infections with multiple pathogens [382].

4.5 Acknowledgements

We would like to thank G. Perreault and the animal facility staff at McGill University for their excellent care of our animal colony. We thank C. Stegen and J. Leconte at the Cell Vision Core Facility. D.R. was supported by a Frederick Banting and Charles Best Canada Graduate Doctoral Award (CIHR) and a Tomlinson Doctoral Fellowship (McGill). J.N.M. is a Canada Research Chair for Immune Cell Dynamics. This work was supported by a CIHR grant awarded to J.N.M.

Author Contributions

J.N.M. and D.R. conceived the study. J.N.M. supervised, led the study implementation, and acquired financial support for this project. D.R. designed and performed experiments, with key experimental input from J.N.M. D.R. analyzed experiments, prepared the figures, and wrote the manuscript with input from J.N.M.

4.6 Figures

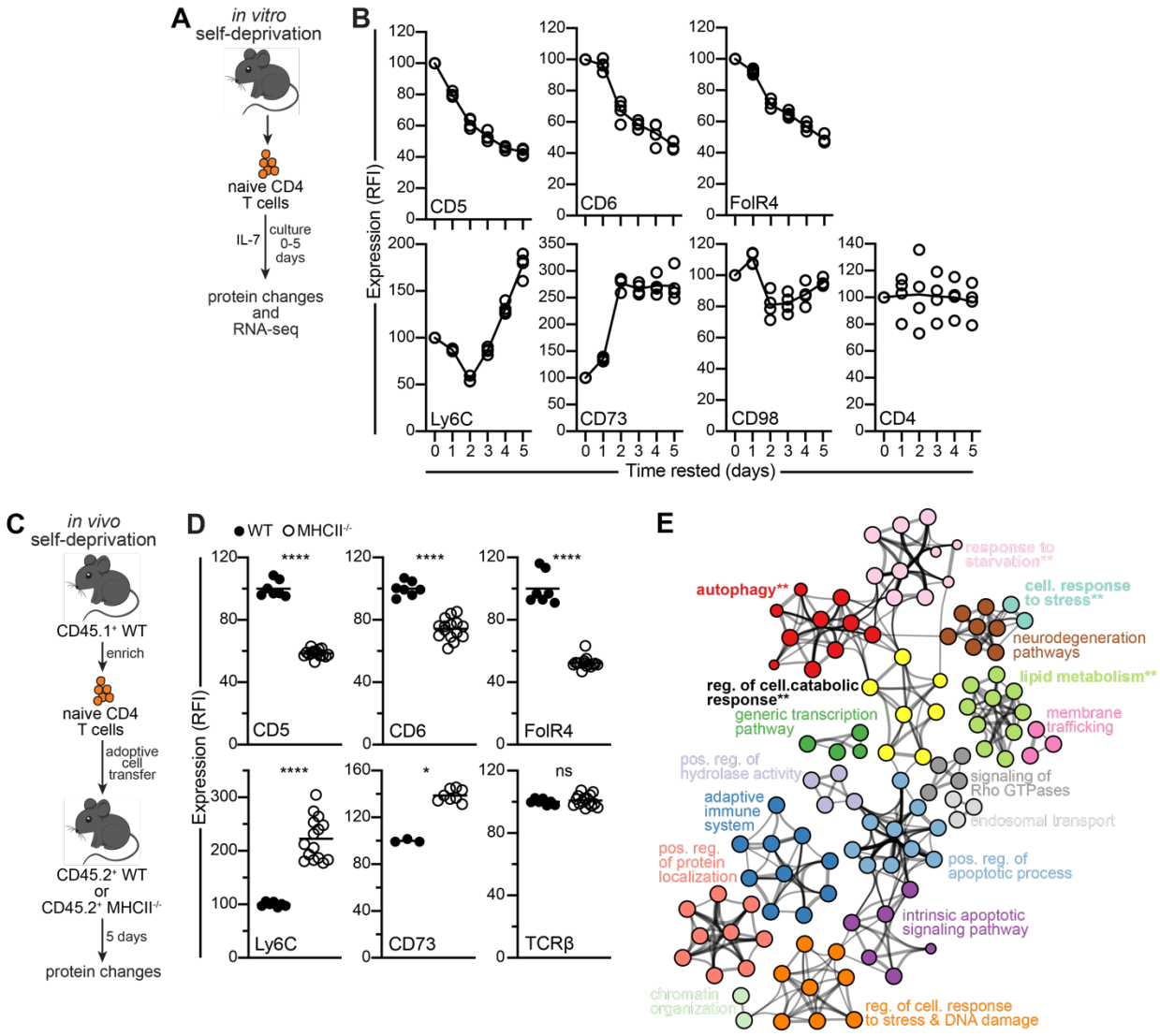


Figure 1. *In vitro* self-deprivation results in protein and gene expression changes in naive CD4$^+$ T cells and implicates adaptations for cell survival. (**A**) Schematic for *in vitro* self-deprivation of naive CD4$^+$ T cells. (**B**) Protein expression of naive CD4$^+$ T cells that were deprived of self-ligand interactions in culture in the presence of IL-7. Relative fluorescence intensity (RFI) is indicated for each protein relative to day 0. n=4 mice. (**C**) Schematic of *in vivo* self-deprivation for adoptively transferred naive CD4$^+$ T cells. (**D**) Protein expression of congenic CD45.1$^+$ naive (CD44^{lo}CD62L^{hi}) CD4$^+$ T cells adoptively transferred into CD45.2$^+$ WT or MHCII$^{-/-}$ mice. Expression (RFI) is relative to the average protein mean fluorescence intensity of cells transferred into WT mice. Data are from 1-2 independent experiments; n=3-15. (**E**) Gene ontology (GO)

enrichment analysis for genes upregulated in self-deprived naive $CD4^+$ T cells. Gene pathways of particular interest are bolded with a '**' following the GO term. Circles correspond to unique GO groups; related groups are coded in the same colour; edges link similar GO terms.

Statistics: Mann-Whitney test (D). ns = not-significant, *p<0.05, ****p<0.001.

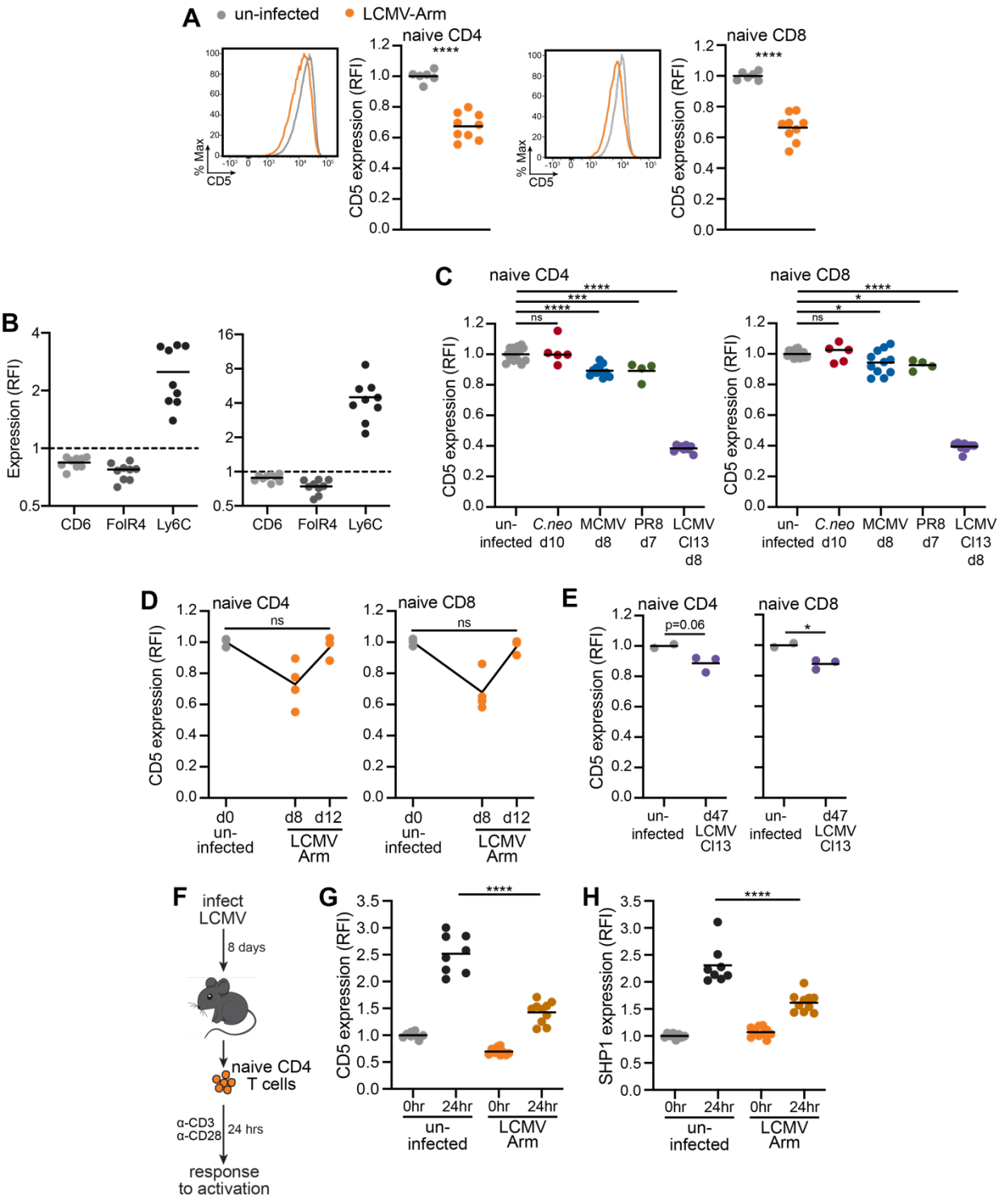


Figure 2. Infection leads to transient self-deprivation impairing non-responding naive T cell sensitivity to activation. (**A**) Representative flow cytometry histograms and summary data of CD5 expression on non-responding naive ($CD44^{lo}CD62L^{hi}$) T cells at day 8 post-infection with LCMV-Arm. RFI is relative to the average CD5 expression in mice not infected with the virus. Data are

from two independent experiments; n=6-9 mice. (**B**) CD6, Ly6C, and FolR4 protein expression on non-responding naive ($CD44^{lo}CD62L^{hi}$) T cells at day 8 post-infection with LCMV-Arm. RFI is relative to the average expression of the respective protein in mice not infected with the virus. The hashed line represents the baseline level in uninfected mice that were used to calculate RFI. Data are from two independent experiments; n=6-9 mice. (**C**) Summary data of CD5 expression on non-responding naive ($CD44^{lo}CD62L^{hi}$) T cells from several infection systems. CD5 expression was assessed on day 7, day 8, and day 10 for IAV PR8, MCMV and LCMV-Cl13, and *C. neoformans* respectively. RFI is relative to the average CD5 expression in mice not infected with the virus. Data are from 1-2 independent experiments; n=4-15 mice. (**D**) Time course of CD5 expression on naive T cells in the blood at day 8 and 12 post-infection with LCMV-Arm. RFI is relative to the average CD5 expression in mice not infected with virus (day 0). n=3-4 mice. (**E**) Summary data of CD5 expression on non-responding naive ($CD44^{lo}CD62L^{hi}$) T cells in LCMV-Cl13 47 days post-infection. RFI is relative to the average CD5 expression in mice not infected with the virus. (**F**) Schematic of naive $CD4^+$ T cell isolation from LCMV-Arm infected mice for *in vitro* activation with α-CD3 and α-CD28. (**G-H**) CD5 (G) and SHP1 (H) expression of naive $CD4^+$ T cells enriched from either mice not infected or infected with LCMV-Arm then subsequently activated for 24 hours. Protein expression is relative to uninfected mice at 0 hours. Data are from two independent experiments; n=8-10.

Statistics: unpaired t-test (A, D, E), one-way ANOVA (C, G, H). *$p<0.05$, **$p<0.01$, ****$p<0.001$.

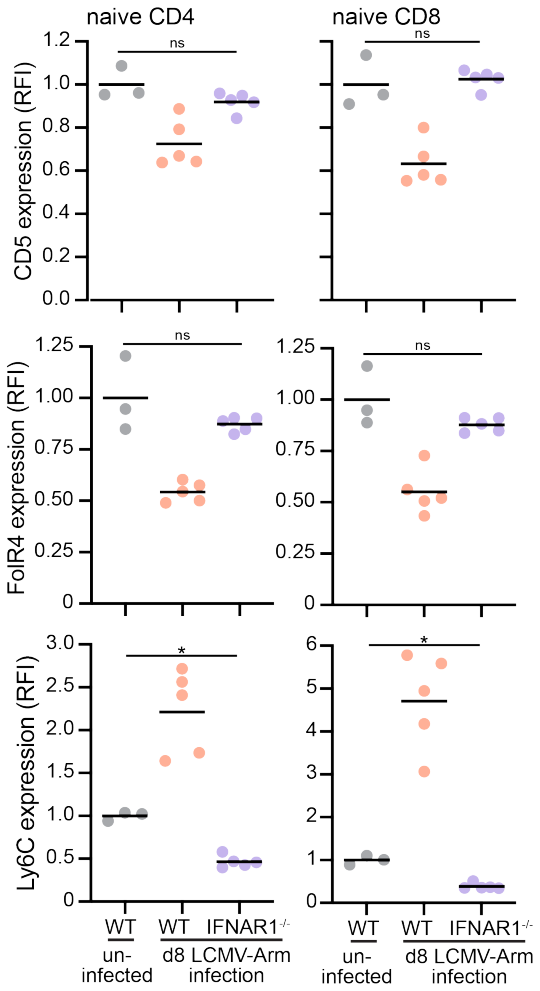


Figure 3. Infection-induced self-deprivation is type-I interferon signaling dependent. CD5, Ly6C, and FolR4 protein expression on naive $CD4^+$ and $CD8^+$ T cells from uninfected WT mice and at day 8 post-infection in WT and $INFAR1^{-/-}$ mice. Protein RFI is relative to the expression of uninfected WT mice. n=3-5.

Statistics: Mann-Whitney test. ns = not-significant, *p<0.05.

4.7 Materials and Methods

Mice

C57BL/6, congenic CD45.1^{+}, MHCII$^{-/-}$ [304] mice were purchased from Jackson Laboratories (Bar Harbor, ME). The IFNAR1$^{-/-}$ mice were shared by Dr. J Rauch (McGill University Health Centre) [383]. All mice were on a C57BL/6 background, bred in-house and experiments were performed at 6-12 weeks of age with both males and females. Animal housing, care, and research were in accordance with the Guide for the Care and Use of Laboratory Animals and all procedures performed were approved by the McGill University Animal Care Committee.

Pathogens

LCMV. LCMV-Arm and -Cl13 strains were propagated from stocks provided by Dr. M. Richer (University of Indiana) as previously described [252]. Mice were infected with 2x10^5 plaque forming units (PFU) of LCMV-Arm by intraperitoneal injection or 2x10^6 PFU by intravenous injection for LCMV-Cl13 as previously described [306, 363]. Tissue was harvested from mice at either day 8 or 47 post-infection.

Cryptococcus neoformans. The *C. neoformans* H99 strain was provided by Dr. K. Kwon-Chung (NIH). Stocks used for infection were prepared as previously described [352]. Mice were anesthetized and infected by intrapharyngeal aspiration with 5x10^3 colony forming units (CFU) in 20 µL of PBS. Mice were sacrificed and tissue was collected 10 days after infection. *C. neoformans* CFUs were determined as previously described [352].

IAV. Mice were anesthetized and infected intranasally with 2x10^3 PFU with the PR/8 strain of IAV in PBS. Mice were euthanized and tissue was harvested 7 days post-infection.

MCMV. The MCMV-Δm157 stain was provided by Dr. J. Sun (Memorial Sloan Kettering Cancer Center). Mice were infected with 2x10^6 PFU intravenously and tissue was harvested 8 days post-infection.

Leukocyte isolation

Spleen was collected and crushed through a 70-µm filter with 1% RPMI (1% penicillin/streptomycin, 1% L-glutamine, and 1% FBS). Red blood cell lysis was done using ACK lysis buffer (Life Technologies) for 3 minutes. Cells were then washed with PBS, refiltered, and resuspended in 1% RPMI. For lymphocyte isolate from the blood, whole blood

was centrifuged at 2000rpm for 20 minutes using a Histopaque-1077 (Sigma-Aldrich) density gradient and lymphocytes were collected at the plasma/Histopaque-1077 interface. Single-cell suspensions were dilution 1:10 in Trypan Blue and manual counting of live cells (Trypan Blue-negative) on a hemacytometer was used to determine total cell counts. To purify total T cells or $CD4^+$ T cells Stemcell EasySep mouse total T cell or $CD4^+$ T cell enrichment kits were used. To enrich naive $CD4^+$ T cells Stemcell EasySep mouse $CD4^+$ T cell enrichment kits were supplemented with CD44-bio antibody (IM7).

For experiments where naive $CD4^+$ T cells were deprived of self-pMHC interactions, cells (either total or enriched cells, as specified) were cultured in 96-well plates in complete RPMI (10% FBS, 1% L-glutamine, 1% penicillin/streptomycin, 1% HEPES buffer, 1% Sodium Pyruvate, 1% non-essential amino acids, and 0.1% 2-mercapto-ethanol 1000X solution) supplemented with 10 ng/mL of IL-7 (Biolegend). To activate T cells, cells were cultured with complete RPMI in 96-well plates coated with α-CD3 and α-CD28 (Biolegend; both at 3 mg/mL).

Flow cytometry

Cells were incubated with Fixable Viability Dye (AF780, Life Technologies) diluted in PBS for 20 minutes at 4°C. Extracellular antibodies were diluted in FACS buffer (2% FBS and 5mM EDTA in PBS) containing Fc Block (Life Technologies) and were incubated for 30 minutes at 4°C. If intracellular staining was required, cells were subsequently incubated in Foxp3 Transcription Factor Fixation/Permeabilization Concentrate and Diluent (Life Technologies) for 30 minutes at 4°C. Intracellular antibodies were then diluted in permeabilization wash buffer and incubated for 30-60 minutes at 4°C. Directly conjugated antibodies used were as follows: TCRb (H57-597), CD4 (RM4.5), CD8a (53-6.7), CD5 (53-7.3), Foxp3 (FJK-16 s), CD44 (IM7), CD62L (MEL-14), CD98 (RL388), CD73 (TY/11.8), FolR4 (eBio12A5), Ly6C (HK1.4), CD6 (OX- 129), CD45.1 (A20), and CD45.2 (104). Primary unconjugated antibodies were: SHP-1 (C14H6). The secondary conjugated antibody used was Donkey anti-Rabbit IgG (H+L) Alexa Fluor 647. To assess SHP-1 expression, cells were fixed with 1X TFP Fix/Perm Buffer for 50 minutes at 4°C, washed, then incubated in Perm Buffer III (BD Biosciences) for 20 minutes on ice. Fc block and all surface and intracellular antibodies were diluted together in permeabilization wash buffer and incubated for 50 minutes at 4°C then washed. The secondary

antibody was also diluted in 1X TFP Perm/Wash Buffer and stained for an additional 50 minutes at 4ºC. For all flow cytometry experiments, cells were acquired using an LSRFortessa (BD Biosciences) and analyzed with FlowJo software (BD Biosciences).

Adoptive cell transfers

All donors and recipients were sex matched. 13 CD45.1^{+} mice were used to isolate CD4^{+} T cells as described above. A total of 6-8x10^{6} CD4^{+} T cells were adoptively transferred into either CD45.2^{+} WT or CD45.2^{+} MHCII$^{-/-}$ mice. Transferred cells were isolated from the spleen and peripheral lymph nodes of recipient mice 5 days post-transfer.

RNA-seq

Cell sorts, RNA-seq preparation, library prep, analysis, and data visualizing were done as previously described [140]. Briefly, 1x10^{6} CD5^{lo} and CD5^{hi} naive CD4^{+} T cells were sorted and were either directly added to 500uL TRIzol (ThermoFisher Scientific) or were first rested in complete RPMI supplemented with IL-7 for 22 hours. FASTQ files were mapped to the UCSC mm9 *Mus musculus* genome and normalized read counts DEG analysis was performed with EdgeR. An FDR threshold of <0.01 was considered significant and No FC threshold was imposed on DEGs. The overlap from the DEGs in the fresh CD5^{hi} versus self-deprived CD5^{hi} and fresh CD5^{lo} versus self-deprived CD5^{lo} group comparisons were used to define the 'common T cell signature'. The data reported in this paper can be found in the Gene Expression Omnibus database under accession number GSE185675.

Statistical analyses

Group comparisons were performed using Prism V9 (GraphPad). Unless specified, data are presented as mean with each data point representing an individual mouse. The cut-off for significance was $p<0.05$ for all analyses unless otherwise stated. Information about specific statistical tests used for each experiment is listed in the figure legends.

4.8 Supplemental information

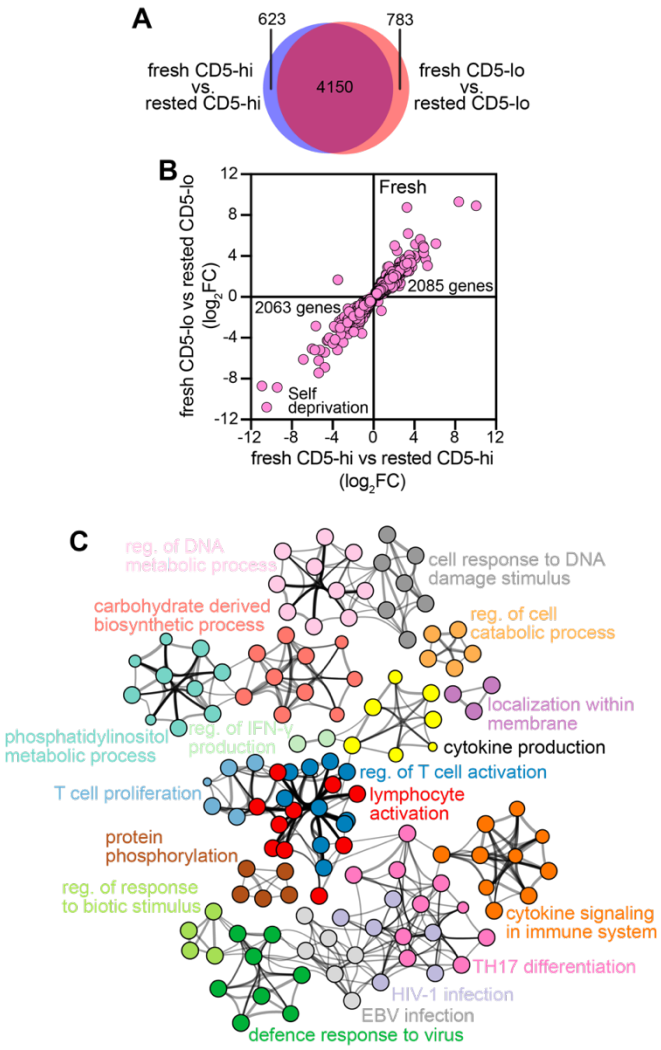


Figure S1. Self-pMHC interactions maintain transcriptional networks that support naive CD4$^+$ T cell readiness. (**A**) Venn diagram of differentially expressed genes (DEGs) (FDR cut-off <0.01) for fresh CD5^{hi} versus self-deprived CD5^{hi} and fresh CD5^{lo} versus self-deprived CD5^{lo} used to define the T cell signature common to all T cells. (**B**) Expression (log_2Fold Change) of the DEGs in the T cell signature. Genes enriched in fresh naive CD4$^+$ T cells are found in the upper-right quadrant, and gene enrichment upon self-deprivation of naive CD4$^+$ T cells is located in the lower-left quadrant. (**C**) Gene ontology enrichment analysis of the fresh naive CD4$^+$ T cell gene signature. Circles correspond to unique GO groups; related groups are coded in the same colour; edges link similar GO terms.

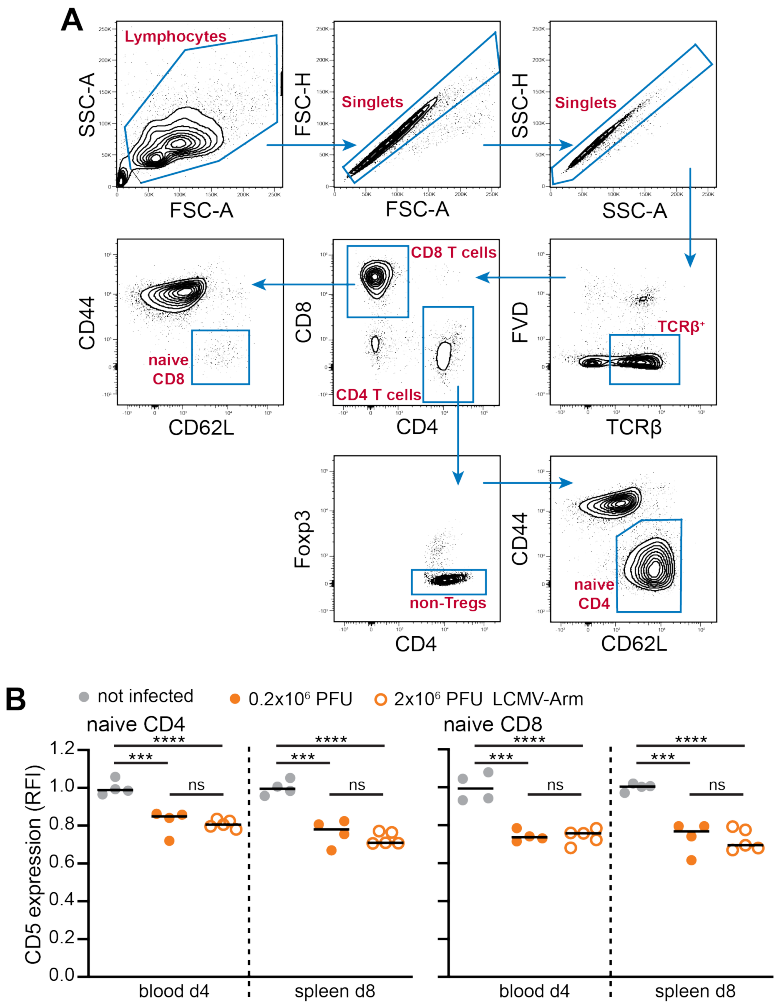


Figure S2. Infection-induced self-deprivation is not dependent on the dose of pathogen. (**A**) Flow cytometry plots illustrate the gating strategy used to assess non-responding naive $CD4^+$ and $CD8^+$ T cells during infection. (**B**) Summary data of CD5 expression on non-responding naive ($CD44^{lo}CD62L^{hi}$) T cells in the blood at day 4 and spleen day 8 post-infection with LCMV-Arm. RFI is relative to the average CD5 expression in mice un-infected with the virus. n=4-5 mice.

Statistics: One-way ANOVA (B). ns = not-significant, ***p<0.001, ****p<0.0001

CHAPTER 5: DISCUSSION

5.1 Overview

Our understanding of the factors that influence the fate of an individual T cell when called upon to combat infection is evolving. While T cell receptor (TCR) affinity/avidity for foreign ligand and the cytokine milieu in which a T cell is activated are major players in the fate of a T cell upon antigen challenge, pre-existing functional biases among individual naive T cells may act to ensure functional diversity in the responding cells. In Chapter 2, we identify that the strength at which naive $CD4^+$ T cells recognize self-peptides presented by major histocompatibility complexes (self-pMHCs), also called self-reactivity, imparts diverse chromatin and transcriptional networks. We also show that the transcriptional diversity is established by two sources of heterogeneity, one that relies on continuous self-pMHC signals and the second where differences are imprinted in the thymus during development, presumably by epigenetic changes. In Chapter 3, we found that low self-reactive naive $CD4^+$ T cells preferentially contributed to the T cell response during chronic infection and that chronic pathogen control is delayed when fewer low self-reactive T cells are present within the T cell population. In Chapter 4, we begin to probe the consequences of self-pMHC deprivation. We observed that naive $CD4^+$ T cells may utilize cell survival pathways to adapt to self-deprivation. Additionally, non-responding naive T cells undergo transient self-deprivation during infection, which is dependent on type-I interferon (IFN) signaling. How these findings fit into the larger context of naive T cell heterogeneity and maintenance and what their implications might be are discussed below.

5.2 Nature vs nurture – naive T cell heterogeneity and thymic imprinting

During the DN progenitor stage of T cell development, the newly generated TCRβ chain is paired with the invariant pre-TCRα chain, forming the pre-TCR. However, beyond being a gatekeeper to ensure functional TCRβ rearrangement [16], β-selection may contribute to shaping T cell self-reactivity, thus impacting the mature T cell compartment in ways that have previously not been appreciated. The functional consequences of pre-TCR engagement with self-pMHC on the TCR repertoire and subsequent selection processes are still emerging. The strength of pre-TCR signaling at the DN T cell stage can shape the mature T cell repertoire such that decreasing pre-TCR signals limits TCR diversity [384]. It has been reported that CD5 is upregulated after the β-

selection checkpoint [172, 385], but whether the strength of signaling and CD5 levels during β-selection are dependent on individual TCRβ clones (and their interactions with self-pMHC) is not yet known. It is tempting to speculate that CD5 levels correlate with the strength of pre-TCR signaling for individual TCRβ chains and, given that CD5 modulates TCR signaling, that CD5 might participate in setting unique positive selection thresholds for each TCRβ. Thus, β-selection may ultimately contribute to establishing the self-reactivity of a T cell even before the mature TCR is expressed.

A series of encounters with positively selecting self-peptides is required for differentiation of double positive (DP) cells into mature single positive (SP) thymocytes to reach a 'signal threshold' that may be specific for each TCR. Thymocytes undergoing positive selection are characterized by serial, transient interactions with self-peptide that occur over time [386, 387]. The magnitude and frequency of these individual interactions is likely influenced by the interactions with the self-peptide(s) providing the survival signals (**Figure 1**). Given a gradual increase in basal intracellular calcium levels in thymocytes during this period [388], it was hypothesized that thymocytes remember individual TCR signals and accumulate signaling intermediates until they reach a 'signal threshold'. Evidence of this is further supported by the finding that a dose-dependent inhibition of TCR signals during positive selection of a monoclonal thymocyte population leads to a decrease in the frequency of cells that pass the positive selection test; however, the mature T cells that develop have accumulated similar amounts of signal [387]. This might suggest that 'signal thresholds' are set for individual TCRs prior to positive selection, and possibly as early as the β-selection stage (**Figure 1**). Differences in the amount of accumulated calcium/signal may lead to distinct transcriptional programs and epigenetic imprinting that ultimately impacts the functional bias of the T cell in the periphery.

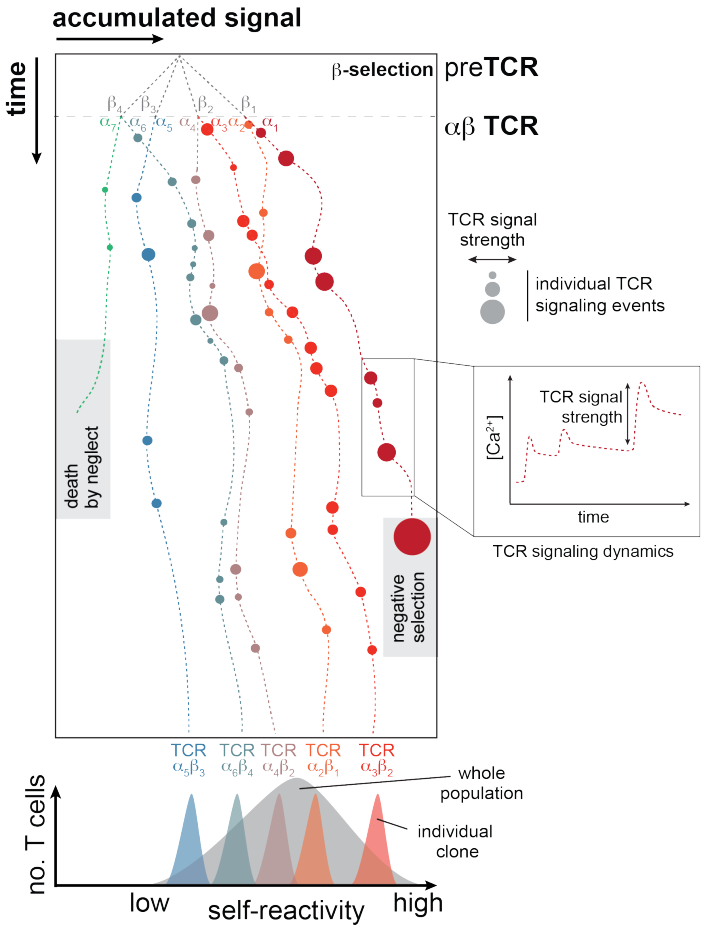


Figure 1. Model predicting how cumulative self-signals establish the self-reactivity of individual naive T cells. Following TCRβ rearrangement (hypothetical $\beta_{1\text{-}4}$), β-selection may establish initial differences in self-reactivity. Subsequent rearrangement of the TCRα chain (hypothetical $\alpha_{1\text{-}7}$) and interactions with self-pMHC (denoted as circles) induce Ca^{2+} fluxes with baseline intracellular Ca^{2+} levels increasing over time. Inset depicts the Ca^{2+} signal received at each self-pMHC interaction during selection dependent on the strength of interaction. If a thymocyte fails to receive self-signals, it dies by neglect (green clone) or if a thymocyte interacts with self too strongly, it will be negatively selected from the repertoire (dark red clone). Those clones that complete development then occupy distinct niches of self-reactivity that together establish a polyclonal repertoire. Adapted from [4].

TCR repertoire selection in the thymus appears to play an important role in establishing functional biases in naive T cells across the self-reactivity spectrum that may impact how they ultimately respond to pathogens. Functional biases that correlate with self-reactivity are already evident in mature thymocytes [2], and it has been suggested that self-pMHC affinity during positive selection influences the responsiveness of T cells at least at early time points after thymic egress [389]. Despite changes in TCR sensitivity and the self-immunopeptidome from the thymus to the periphery, the relative self-reactivity of individual TCRs is generally maintained in both

compartments as read out by surrogate markers of TCR signal strength [1]. In Chapter 2, we showed that when the ability of naive T cells to engage self-pMHC is removed, many of the functionally relevant gene expression differences between T cells with relatively low versus high affinity for self are retained [140]. These characteristics are reminiscent of the Treg lineage: while continued interactions with self-peptide in the periphery are essential for Treg function, interactions with self-pMHC are not required for maintaining the expression of the majority of genes contributing to the Treg gene signature which were epigenetically imprinted in the thymus during their development [390-392]. Ultimately, however, which functional biases are thymically imprinted and which require ongoing interactions with the self-immunopeptidome in the periphery, is still unknown.

Identifying the sources of heterogeneity in the naive T cell population, which functional attributes of T cells can be modulated through post-thymic interventions versus those that are fixed during thymic development, and what the characteristics of the TCR are that are associated with the extent of self-reactivity, could significantly contribute to improving the efficacy of T cell therapies – developing or isolating populations with desirable functions in terms of cytokine production or persistence. Thus, populations of T cells with relevant traits might be identified and isolated based on their self-reactivity or engineered with antigen-specific TCRs known to provide relatively lower or higher basal signals.

5.3 Limp handshakes do not always reflect a weak character – low-reactive naive T cells generated by N-nucleotide diversity

As developing thymocytes rearrange their TCR in the thymus, the combination of V(D)J gene segments and non-templated (N)-nucleotide insertions from terminal deoxynucleotidyl transferase (TdT) produce an estimated 10^{20} possible TCRs [113-115]. Without N-nucleotide additions, TCR diversity is reduced by over 90% [297, 338, 339]. Even though TdT is conserved in jawed vertebrates [337], its evolutionary benefit has remained elusive. The T cell repertoire in TdT-KO mice is not impaired during infection [336]; however, the pathogens that have been tested all establish acute infections. Several non-mutually exclusive hypotheses may explain the evolutionary advantage of why introducing somatic TCRs (containing N-nucleotides) would benefit immune responses instead of a repertoire made entirely of germline sequences [138]. First, while non-germline TCRs might have a higher probability of not binding pMHC as well,

occasionally TdT may produce a TCR that binds pMHC better that provides essential help to a T cell response [393]. Second, the TdT-dependent TCRs may be important for controlling pathogens that easily mutate and escape the T cell response [394, 395]. Third, having somatic TCRs that may have lower pMHC binding strength present within the repertoire might benefit the T cell response during chronic infection where T cells that are highly responsive to pMHC are more likely to become exhausted and dysfunctional [396]. In Chapter 3 we focus on the third hypothesis, and, for the first time, we provide evidence that TdT produces TCRs that have, on average, lower pMHC reactivity and that this has consequences for chronic infection control. It should be noted that the results outlined in Chapter 3 do not rule out other potential advantages of TCR repertoire diversification by TdT in different contexts, but rather suggest the added benefit that TdT-dependent TCRs possess biases in pMHC binding strength that benefit hosts challenged specifically with chronic infection.

TCRs that are produced independently of TdT, the germline TCRs, have been proposed to be of higher reactivity to pMHC [138]. The neonatal TCR repertoire is initially generated without the use of TdT and does not have N-nucleotide insertions; consequently, these TCRs have higher self-reactivity [296, 361]. Additionally, TdT-KO T cells are more efficiently positively selected indicating that TdT-independent TCRs have, on average, higher reactivity to pMHC [144]. We found that TdT-KO pathogen-responding (activated) $CD4^+$ and $CD8^+$ T cells express higher levels of PD-1 during chronic infection suggesting that TdT-KO T cells receive stronger TCR signals. Ultimately, prolonged TCR signals such as chronic antigen stimulation during infection can drive T cell exhaustion [397]. It is increasingly appreciated that assessing exhaustion by cell surface marker expression alone (like PD-1) is insufficient. Exhaustion is now defined by complex epigenetic phenotypes that position T cells on a spectrum of dysfunction [356, 398, 399]. In the future, it will be important to investigate the epigenetic changes of TdT-KO T cells in the chronic setting. It is possible that TdT-independent TCRs, while on average being of higher pMHC-reactivity, may also be more prone to exhaustion.

Our work provides new insight into how T cells with TdT-dependent TCRs, and presumably lower pMHC reactivity, contribute to the immune response during chronic infection. Chronic infections provide long-term sources of foreign-antigen and inflammation that induce exhaustion and dysfunction within the responding and memory T cell populations. However, chronic inflammation can arise in sterile or non-sterile conditions. Such examples include

microbiota dysbiosis, exposure to xenobiotics, physical inactivity, diet, obesity, social isolation, psychological stress, and disrupted sleep [400]. It will be intriguing to investigate how TdT-dependent TCRs contribute to the immune responses in instances of sterile and other non-sterile chronic inflammation, aside from chronic infection.

Loss of TdT expression within the T cell compartment delayed the control of chronic pathogen. When we infected full TdT-KO mice (possessing both TdT-KO T and B cells) there was no impact on pathogen control. Much like T cells, B cells also use TdT to diversify the B cell receptor (BCR) [354]. In Chapter 3, we did not consider the effect of TdT deficiency in the BCR repertoire and have not addressed whether this may similarly impact acute or chronic infection dynamics. However, given that BCRs are not constrained by having to recognize antigen in the context of MHC, the distance of the BCR from the germline would presumably not have the same implication with regard to reducing antigen binding strength. Investigating this question will certainly be interesting in the future.

Our data show that one possible explanation for the evolutionary origin of TdT is to produce TCRs with, on average, lower-pMHC reactivity that aid in chronic pathogen control. Of note, no people with a loss-of-function in TdT have been identified to date. Whether this is because TdT loss-of-function is lethal or that these people live relatively normal lives is not known. However, aberrant expression and mutations in TdT are used as key biomarkers in the diagnosis of several leukemias [401, 402].

5.4 When the going gets tough, T cells endure – impact of self-deprivation on the naive T cell population

The peripatetic nature of naive T cells ensures they cover the largest area possible when scanning for foreign antigens [29]. As naive T cells migrate between the secondary lymphoid organs they are continuously encountering self-pMHC on antigen presenting cells (APCs) that provide them with tonic self-signals [366]. Despite self-signals being important for maintaining naive T cell numbers and sensitivity to foreign antigen [221, 223, 248], little is known about what happens when naive T cells are deprived of self-pMHC interactions. Notably, when self-signals are deprived, naive T cells are surprisingly resilient to attrition and can survive for several weeks. In contrast, the absence of IL-7 signals leads to a much more rapid decay in T cell numbers [224]. It is not understood how tonic self-signals provide naive T cells with survival cues. In the absence

of IL-7 signals, naive T cells have reduced cell size, less expression of the glucose transporter GLUT1, reduced mitochondrial function, and lower ATP levels [376], but whether self-deprivation induces a similar 'starvation' phenotype has not been studied. Here, we show that naive $CD4^+$ T cells may cope with self-deprivation by upregulating transcriptional networks that promote cell survival pathways like autophagy and response to starvation. It will be important to investigate if naive $CD4^+$ T cells use these survival pathways to prevent cell death when self-signals are limiting. Additionally, whether naive $CD8^+$ T cells upregulate similar survival pathways upon self-deprivation as naive $CD4^+$ T cells have not been studied.

In Chapter 2, we show that many of the transcriptional differences present between low and high self-reactive naive $CD4^+$ T cells are not dependent on continuous self-pMHC interactions in the periphery and may be imprinted in the thymus during development. We and others have also shown that the strength of self-pMHC reactivity imparts biases in lineage choices [3, 139, 140, 183, 255]. It is unclear if some of the pre-disposition in fate decisions are also thymically imprinted. It will be interesting to see whether self-deprivation impacts $CD4^+$ T cell effector cell differentiation.

We also wondered if there were physiological settings where self-deprivation may occur. We reasoned that infection, where APCs preferentially present foreign-peptides on MHCs over self-peptides [372, 373], may inadvertently deprive pathogen-non-specific naive T cells of the tonic self-signals they require for maintenance and sensitivity to cognate antigens. Indeed, we found that infection leads to self-deprivation in non-responding naive T cells and that this infection-induced self-deprivation (IISD) is dependent on type-I IFN signals. However, whether type-I IFN signaling establishes IISD by upregulating the immunoproteosome, thus changing the peptide repertoire presented by APCs, or directly acts on naive T cells remains unknown. Additionally, at steady-state, IFNβ is constitutively expressed at low levels [403], which may provide an additional tonic signal to naive T cells. Indeed, there is increasing evidence that suggests type-I IFN signals may play a role in naive T cell maintenance [241, 244]. Particularly, naive $CD8^+$ T cells with high self-pMHC reactivity are more responsive to tonic type-I IFN signals and these IFN signals are crucial for maintaining the $Ly6C^+$ high-self-reactive population [244]. For naive $CD4^+$ T cells, single-cell RNA-sequencing has identified a type-I IFN gene signature within the population [241]; but, the impact of tonic type-I IFN signaling within the naive $CD4^+$ T cell population has yet to be described.

5.5 Concluding remarks

T cells can differentiate into several effector cell populations in response to foreign threats. Historically, it has been thought that prior to pathogen encounter, naive T cells represent a homogenous population. It is only within the past 10 years that we now understand that the naive T cell pool is remarkably diverse. The work presented in this thesis found a previously unappreciated level of chromatin and transcriptional heterogeneity within the naive $CD4^+$ T cell population that predisposes T cells to certain effector lineage choices based on strength of self-reactivity. Additionally, we show that a portion of the naive $CD4^+$ T cell transcriptional heterogeneity may be thymically imprinted during development. We also show that low self-reactive T cells are important for chronic pathogen control. Furthermore, we implicate pathways that naive $CD4^+$ T cells may use to promote survival upon removal of self-pMHC signals at steady-state and perhaps during infection. More work is needed to determine exactly how naive T cells adapt to self-signal deprivation and the signals that regulate access to self-ligands.

REFERENCES

[1] J.N. Mandl, J.P. Monteiro, N. Vrisekoop, R.N. Germain, T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens, Immunity 38(2) (2013) 263-274.

[2] S.P. Persaud, C.R. Parker, W.L. Lo, K.S. Weber, P.M. Allen, Intrinsic CD4+ T cell sensitivity and response to a pathogen are set and sustained by avidity for thymic and peripheral complexes of self peptide and MHC, Nat Immunol 15(3) (2014) 266-74.

[3] A. Sood, M.E. Lebel, M. Fournier, D. Rogers, J.N. Mandl, H.J. Melichar, Differential interferon-gamma production potential among naive CD4(+) T cells exists prior to antigen encounter, Immunol Cell Biol 97(10) (2019) 931-940.

[4] S. This, D. Rogers, È. Mallet Gauthier, J.N. Mandl, H.J. Melichar, What's self got to do with it: Sources of heterogeneity among naive T cells, Semin Immunol 65 (2022) 101702.

[5] T.W. Allen, Homeri Ilias, E typographeo Clarendoniano1931.

[6] K. Laios, The thymus gland in ancient Greek medicine, Hormones 17(2) (2018) 285-286.

[7] W. Hewson, G. Gulliver, The works of William Hewson, F.R.S, Printed for the Sydenham Society, London, 1846.

[8] J.F. Miller, Immunological function of the thymus, Lancet 2(7205) (1961) 748-9.

[9] J.C. Pui, D. Allman, L. Xu, S. DeRocco, F.G. Karnell, S. Bakkour, J.Y. Lee, T. Kadesch, R.R. Hardy, J.C. Aster, W.S. Pear, Notch1 expression in early lymphopoiesis influences B versus T lineage determination, Immunity 11(3) (1999) 299-308.

[10] A.M. Michie, J.R. Carlyle, T.M. Schmitt, B. Ljutic, S.K. Cho, Q. Fong, J.C. Zúñiga-Pflücker, Clonal characterization of a bipotent T cell and NK cell progenitor in the mouse fetal thymus, J Immunol 164(4) (2000) 1730-3.

[11] D.I. Godfrey, J. Kennedy, T. Suda, A. Zlotnik, A developmental pathway involving four phenotypically and functionally distinct subsets of CD3-CD4-CD8- triple-negative adult mouse thymocytes defined by CD44 and CD25 expression, J Immunol 150(10) (1993) 4244-52.

[12] H.R. MacDonald, F. Radtke, A. Wilson, T cell fate specification and alphabeta/gammadelta lineage commitment, Curr Opin Immunol 13(2) (2001) 219-24.

[13] H. von Boehmer, H.J. Fehling, Structure and function of the pre-T cell receptor, Annu Rev Immunol 15 (1997) 433-52.

[14] P. Mombaerts, J. Iacomini, R.S. Johnson, K. Herrup, S. Tonegawa, V.E. Papaioannou, RAG-1-deficient mice have no mature B and T lymphocytes, Cell 68(5) (1992) 869-77.

[15] Y. Shinkai, S. Koyasu, K. Nakayama, K.M. Murphy, D.Y. Loh, E.L. Reinherz, F.W. Alt, Restoration of T cell development in RAG-2-deficient mice by functional TCR transgenes, Science 259(5096) (1993) 822-5.

[16] A. Dutta, B. Zhao, P.E. Love, New insights into TCR β-selection, Trends Immunol 42(8) (2021) 735-750.

[17] T.K. Starr, S.C. Jameson, K.A. Hogquist, Positive and negative selection of T cells, Annu Rev Immunol 21 (2003) 139-76.

[18] E. Robey, B.J. Fowlkes, Selective events in T cell development, Annu Rev Immunol 12 (1994) 675-705.

[19] R.N. Germain, T-cell development and the CD4-CD8 lineage decision, Nat Rev Immunol 2(5) (2002) 309-22.

[20] R.G. Scollay, E.C. Butcher, I.L. Weissman, Thymus cell migration. Quantitative aspects of cellular traffic from the thymus to the periphery in mice, Eur J Immunol 10(3) (1980) 210-8.

[21] G.L. Stritesky, S.C. Jameson, K.A. Hogquist, Selection of self-reactive T cells in the thymus, Annu Rev Immunol 30 (2012) 95-114.

[22] E. Palmer, Negative selection--clearing out the bad apples from the T-cell repertoire, Nat Rev Immunol 3(5) (2003) 383-91.

[23] L. Klein, M. Hinterberger, G. Wirnsberger, B. Kyewski, Antigen presentation in the thymus for positive selection and central tolerance induction, Nat Rev Immunol 9(12) (2009) 833-44.

[24] T.A. Baldwin, K.A. Hogquist, S.C. Jameson, The fourth way? Harnessing aggressive tendencies in the thymus, J Immunol 173(11) (2004) 6515-20.

[25] B.D. McDonald, J.J. Bunker, S.A. Erickson, M. Oh-Hora, A. Bendelac, Crossreactive αβ T Cell Receptors Are the Predominant Targets of Thymocyte Negative Selection, Immunity 43(5) (2015) 859-69.

[26] M.A. Daniels, E. Teixeiro, J. Gill, B. Hausmann, D. Roubaty, K. Holmberg, G. Werlen, G.A. Holländer, N.R. Gascoigne, E. Palmer, Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling, Nature 444(7120) (2006) 724-9.

[27] K. Shortman, D. Vremec, M. Egerton, The kinetics of T cell antigen receptor expression by subgroups of CD4+8+ thymocytes: delineation of CD4+8+3(2+) thymocytes as post-selection intermediates leading to mature T cells, J Exp Med 173(2) (1991) 323-32.

[28] J.L. Gowans, E.J. Knight, THE ROUTE OF RE-CIRCULATION OF LYMPHOCYTES IN THE RAT, Proc R Soc Lond B Biol Sci 159 (1964) 257-82.

[29] U.H. von Andrian, T.R. Mempel, Homing and cellular traffic in lymph nodes, Nat Rev Immunol 3(11) (2003) 867-78.

[30] A. James, M.J. Plank, R. Brown, Optimizing the encounter rate in biological interactions: Ballistic versus Lévy versus Brownian strategies, Phys Rev E Stat Nonlin Soft Matter Phys 78(5 Pt 1) (2008) 051128.

[31] M.K. Jenkins, J.J. Moon, The role of naive T cell precursor frequency and recruitment in dictating immune response magnitude, J Immunol 188(9) (2012) 4135-40.

[32] J.J. Moon, H.H. Chu, M. Pepper, S.J. McSorley, S.C. Jameson, R.M. Kedl, M.K. Jenkins, Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude, Immunity 27(2) (2007) 203-13.

[33] J.N. Blattman, R. Antia, D.J. Sourdive, X. Wang, S.M. Kaech, K. Murali-Krishna, J.D. Altman, R. Ahmed, Estimating the precursor frequency of naive antigen-specific CD8 T cells, J Exp Med 195(5) (2002) 657-64.

[34] M.F. Krummel, F. Bartumeus, A. Gérard, T cell migration, search strategies and mechanisms, Nat Rev Immunol 16(3) (2016) 193-201.

[35] C. Beauchemin, N.M. Dixit, A.S. Perelson, Characterizing T cell movement within lymph nodes in the absence of antigen, J Immunol 178(9) (2007) 5505-12.

[36] J. Textor, A. Peixoto, S.E. Henrickson, M. Sinn, U.H. von Andrian, J. Westermann, Defining the quantitative limits of intravital two-photon lymphocyte tracking, Proc Natl Acad Sci U S A 108(30) (2011) 12401-6.

[37] E. Woolf, I. Grigorova, A. Sagiv, V. Grabovsky, S.W. Feigelson, Z. Shulman, T. Hartmann, M. Sixt, J.G. Cyster, R. Alon, Lymph node chemokines promote sustained T lymphocyte motility without triggering stable integrin adhesiveness in the absence of shear forces, Nat Immunol 8(10) (2007) 1076-85.

[38] T.R. Mempel, T. Junt, U.H. von Andrian, Rulers over randomness: stroma cells guide lymphocyte migration in lymph nodes, Immunity 25(6) (2006) 867-9.

[39] M.J. Miller, S.H. Wei, I. Parker, M.D. Cahalan, Two-photon imaging of lymphocyte motility and antigen response in intact lymph node, Science 296(5574) (2002) 1869-73.

[40] M. Bajénoff, J.G. Egen, L.Y. Koo, J.P. Laugier, F. Brau, N. Glaichenhaus, R.N. Germain, Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes, Immunity 25(6) (2006) 989-1001.

[41] S.A. Luther, H.L. Tang, P.L. Hyman, A.G. Farr, J.G. Cyster, Coexpression of the chemokines ELC and SLC by T zone stromal cells and deletion of the ELC gene in the plt/plt mouse, Proc Natl Acad Sci U S A 97(23) (2000) 12694-9.

[42] T. Katakai, K. Habiro, T. Kinashi, Dendritic cells regulate high-speed interstitial T cell migration in the lymph node via LFA-1/ICAM-1, J Immunol 191(3) (2013) 1188-99.

[43] T. Worbs, T.R. Mempel, J. Bölter, U.H. von Andrian, R. Förster, CCR7 ligands stimulate the intranodal motility of T lymphocytes in vivo, J Exp Med 204(3) (2007) 489-95.

[44] J.N. Mandl, R. Liou, F. Klauschen, N. Vrisekoop, J.P. Monteiro, A.J. Yates, A.Y. Huang, R.N. Germain, Quantification of lymph node transit times reveals differences in antigen surveillance strategies of naive CD4+ and CD8+ T cells, Proc Natl Acad Sci U S A 109(44) (2012) 18036-41.

[45] S.R. Schwab, J.G. Cyster, Finding a way out: lymphocyte egress from lymphoid organs, Nat Immunol 8(12) (2007) 1295-301.

[46] T.H. Pham, T. Okada, M. Matloubian, C.G. Lo, J.G. Cyster, S1P1 receptor signaling overrides retention mediated by G alpha i-coupled receptors to promote T cell egress, Immunity 28(1) (2008) 122-33.

[47] N.M. Chapman, M.R. Boothby, H. Chi, Metabolic coordination of T cell quiescence and activation, Nat Rev Immunol 20(1) (2020) 55-70.

[48] D.H. Chu, C.T. Morita, A. Weiss, The Syk family of protein tyrosine kinases in T-cell activation and development, Immunol Rev 165 (1998) 167-80.

[49] D. Qian, M.N. Mollenauer, A. Weiss, Dominant-negative zeta-associated protein 70 inhibits T cell antigen receptor signaling, J Exp Med 183(2) (1996) 611-20.

[50] B.B. Au-Yeung, N.H. Shah, L. Shen, A. Weiss, ZAP-70 in Signaling, Biology, and Disease, Annu Rev Immunol 36 (2018) 127-156.

[51] W. Zhang, J. Sloan-Lancaster, J. Kitchen, R.P. Trible, L.E. Samelson, LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor to cellular activation, Cell 92(1) (1998) 83-92.

[52] J.B. Imboden, J.D. Stobo, Transmembrane signalling by the T cell antigen receptor. Perturbation of the T3-antigen receptor complex generates inositol phosphates and releases calcium ions from intracellular stores, J Exp Med 161(3) (1985) 446-56.

[53] A.M. Cameron, J.P. Steiner, A.J. Roskams, S.M. Ali, G.V. Ronnett, S.H. Snyder, Calcineurin associated with the inositol 1,4,5-trisphosphate receptor-FKBP12 complex modulates Ca2+ flux, Cell 83(3) (1995) 463-72.

[54] M. Oh-Hora, M. Yamashita, P.G. Hogan, S. Sharma, E. Lamperti, W. Chung, M. Prakriya, S. Feske, A. Rao, Dual functions for the endoplasmic reticulum calcium sensors STIM1 and STIM2 in T cell activation and tolerance, Nat Immunol 9(4) (2008) 432-43.

[55] E.S. Huseby, E. Teixeiro, The perception and response of T cells to a changing environment are based on the law of initial value, Sci Signal 15(736) (2022) eabj9842.

[56] P.S. Linsley, E.A. Clark, J.A. Ledbetter, T-cell antigen CD28 mediates adhesion with B cells by interacting with activation antigen B7/BB-1, Proc Natl Acad Sci U S A 87(13) (1990) 5031-5.

[57] D. Cantrell, T cell antigen receptor signal transduction pathways, Annu Rev Immunol 14 (1996) 259-74.

[58] L. Chen, D.B. Flies, Molecular mechanisms of T cell co-stimulation and co-inhibition, Nature Reviews Immunology 13(4) (2013) 227-242.

[59] M.K. Jenkins, The ups and downs of T cell costimulation, Immunity 1(6) (1994) 443-6.

[60] M. Azuma, D. Ito, H. Yagita, K. Okumura, J.H. Phillips, L.L. Lanier, C. Somoza, B70 antigen is a second ligand for CTLA-4 and CD28, Nature 366(6450) (1993) 76-9.

[61] C. Dong, Cytokine Regulation and Function in T Cells, Annual Review of Immunology 39(1) (2021) 51-76.

[62] M.J. Butcher, J. Zhu, Recent advances in understanding the Th1/Th2 effector choice, Fac Rev 10 (2021) 30.

[63] S. Jiang, C. Dong, A complex issue on CD4(+) T-cell subsets, Immunol Rev 252(1) (2013) 5-11.

[64] J.J. O'Shea, W.E. Paul, Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells, Science 327(5969) (2010) 1098-102.

[65] D. DiToro, R. Basu, Emerging Complexity in CD4(+)T Lineage Programming and Its Implications in Colorectal Cancer, Front Immunol 12 (2021) 694833.

[66] R. Nishikomori, T. Usui, C.Y. Wu, A. Morinobu, J.J. O'Shea, W. Strober, Activated STAT4 has an essential role in Th1 differentiation and proliferation that is independent of its role in the maintenance of IL-12R beta 2 chain expression and signaling, J Immunol 169(8) (2002) 4388-98.

[67] S.J. Szabo, S.T. Kim, G.L. Costa, X. Zhang, C.G. Fathman, L.H. Glimcher, A novel transcription factor, T-bet, directs Th1 lineage commitment, Cell 100(6) (2000) 655-69.

[68] M. Afkarian, J.R. Sedy, J. Yang, N.G. Jacobson, N. Cereb, S.Y. Yang, T.L. Murphy, K.M. Murphy, T-bet is a STAT1-induced regulator of IL-12R expression in naïve CD4+ T cells, Nat Immunol 3(6) (2002) 549-57.

[69] K.M. Murphy, C. Weaver, Janeway's Immunobiology: Ninth International Student Edition, Garland Science2016.

[70] W. Zheng, R.A. Flavell, The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells, Cell 89(4) (1997) 587-96.

[71] W.E. Paul, Interleukin 4/B cell stimulatory factor 1: one lymphokine, many functions, Faseb j 1(6) (1987) 456-61.

[72] J. Zhu, B. Min, J. Hu-Li, C.J. Watson, A. Grinberg, Q. Wang, N. Killeen, J.F. Urban, Jr., L. Guo, W.E. Paul, Conditional deletion of Gata3 shows its essential function in T(H)1-T(H)2 responses, Nat Immunol 5(11) (2004) 1157-65.

[73] C.M. Lloyd, R.J. Snelgrove, Type 2 immunity: Expanding our view, Sci Immunol 3(25) (2018).

[74] M. Veldhoen, C. Uyttenhove, J. van Snick, H. Helmby, A. Westendorf, J. Buer, B. Martin, C. Wilhelm, B. Stockinger, Transforming growth factor-beta 'reprograms' the differentiation of T helper 2 cells and promotes an interleukin 9-producing subset, Nat Immunol 9(12) (2008) 1341-6.

[75] H.C. Chang, S. Sehra, R. Goswami, W. Yao, Q. Yu, G.L. Stritesky, R. Jabeen, C. McKinley, A.N. Ahyi, L. Han, E.T. Nguyen, M.J. Robertson, N.B. Perumal, R.S. Tepper, S.L. Nutt, M.H. Kaplan, The transcription factor PU.1 is required for the development of IL-9-producing T cells and allergic inflammation, Nat Immunol 11(6) (2010) 527-34.

[76] V. Staudt, E. Bothur, M. Klein, K. Lingnau, S. Reuter, N. Grebe, B. Gerlitzki, M. Hoffmann, A. Ulges, C. Taube, N. Dehzad, M. Becker, M. Stassen, A. Steinborn, M. Lohoff, H. Schild, E. Schmitt, T. Bopp, Interferon-regulatory factor 4 is essential for the developmental program of T helper 9 cells, Immunity 33(2) (2010) 192-202.

[77] J. Lee, B. Lozano-Ruiz, F.M. Yang, D.D. Fan, L. Shen, J.M. González-Navajas, The Multifaceted Role of Th1, Th9, and Th17 Cells in Immune Checkpoint Inhibition Therapy, Front Immunol 12 (2021) 625667.

[78] Ivanov, II, B.S. McKenzie, L. Zhou, C.E. Tadokoro, A. Lepelley, J.J. Lafaille, D.J. Cua, D.R. Littman, The orphan nuclear receptor RORgammat directs the differentiation program of proinflammatory IL-17+ T helper cells, Cell 126(6) (2006) 1121-33.

[79] T. Korn, E. Bettelli, M. Oukka, V.K. Kuchroo, IL-17 and Th17 Cells, Annual Review of Immunology 27(1) (2009) 485-517.

[80] S.C. Liang, X.Y. Tan, D.P. Luxenberg, R. Karim, K. Dunussi-Joannopoulos, M. Collins, L.A. Fouser, Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides, J Exp Med 203(10) (2006) 2271-9.

[81] C.T. Weaver, R.D. Hatton, Interplay between the TH17 and TReg cell lineages: a (co-)evolutionary perspective, Nat Rev Immunol 9(12) (2009) 883-9.

[82] S. Sakaguchi, N. Sakaguchi, M. Asano, M. Itoh, M. Toda, Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases, J Immunol 155(3) (1995) 1151-64.

[83] E. Gambineri, T.R. Torgerson, H.D. Ochs, Immune dysregulation, polyendocrinopathy, enteropathy, and X-linked inheritance (IPEX), a syndrome of systemic autoimmunity caused by mutations of FOXP3, a critical regulator of T-cell homeostasis, Curr Opin Rheumatol 15(4) (2003) 430-5.

[84] S.Z. Josefowicz, L.F. Lu, A.Y. Rudensky, Regulatory T cells: mechanisms of differentiation and function, Annu Rev Immunol 30 (2012) 531-64.

[85] P. Pandiyan, L. Zheng, S. Ishihara, J. Reed, M.J. Lenardo, CD4+CD25+Foxp3+ regulatory T cells induce cytokine deprivation-mediated apoptosis of effector CD4+ T cells, Nat Immunol 8(12) (2007) 1353-62.

[86] L.E. Lucca, M. Dominguez-Villar, Modulation of regulatory T cell function and stability by co-inhibitory receptors, Nature Reviews Immunology 20(11) (2020) 680-693.

[87] N. Askenasy, A. Kaminitz, S. Yarkoni, Mechanisms of T regulatory cell function, Autoimmun Rev 7(5) (2008) 370-5.

[88] R. Goenka, L.G. Barnett, J.S. Silver, P.J. O'Neill, C.A. Hunter, M.P. Cancro, T.M. Laufer, Cutting edge: dendritic cell-restricted antigen presentation initiates the follicular helper T cell program but cannot complete ultimate effector differentiation, J Immunol 187(3) (2011) 1091-5.

[89] N.J. Tubo, A.J. Pagán, J.J. Taylor, R.W. Nelson, J.L. Linehan, J.M. Ertelt, E.S. Huseby, S.S. Way, M.K. Jenkins, Single naive CD4+ T cells from a diverse repertoire produce different effector cell types during infection, Cell 153(4) (2013) 785-96.

[90] R.I. Nurieva, Y. Chung, G.J. Martinez, X.O. Yang, S. Tanaka, T.D. Matskevitch, Y.H. Wang, C. Dong, Bcl6 mediates the development of T follicular helper cells, Science 325(5943) (2009) 1001-5.

[91] Y.S. Choi, R. Kageyama, D. Eto, T.C. Escobar, R.J. Johnston, L. Monticelli, C. Lao, S. Crotty, ICOS receptor instructs T follicular helper cell versus effector cell differentiation via induction of the transcriptional repressor Bcl6, Immunity 34(6) (2011) 932-46.

[92] A. Ballesteros-Tato, B. León, B.A. Graf, A. Moquin, P.S. Adams, F.E. Lund, T.D. Randall, Interleukin-2 inhibits germinal center formation by limiting T follicular helper cell differentiation, Immunity 36(5) (2012) 847-56.

[93] R.J. Johnston, Y.S. Choi, J.A. Diamond, J.A. Yang, S. Crotty, STAT5 is a potent negative regulator of TFH cell differentiation, J Exp Med 209(2) (2012) 243-50.

[94] H. Xu, X. Li, D. Liu, J. Li, X. Zhang, X. Chen, S. Hou, L. Peng, C. Xu, W. Liu, L. Zhang, H. Qi, Follicular T-helper cell recruitment governed by bystander B cells and ICOS-driven motility, Nature 496(7446) (2013) 523-7.

[95] S. Crotty, T follicular helper cell differentiation, function, and roles in disease, Immunity 41(4) (2014) 529-42.

[96] N.S. De Silva, U. Klein, Dynamics of B cells in germinal centres, Nature Reviews Immunology 15(3) (2015) 137-148.

[97] M. Ruterbusch, K.B. Pruner, L. Shehata, M. Pepper, In Vivo CD4(+) T Cell Differentiation and Function: Revisiting the Th1/Th2 Paradigm, Annu Rev Immunol 38 (2020) 705-725.

[98] X. Ma, S. Nakayamada, Multi-Source Pathways of T Follicular Helper Cell Differentiation, Front Immunol 12 (2021) 621105.

[99] J.W. van Heijst, C. Gerlach, E. Swart, D. Sie, C. Nunes-Alves, R.M. Kerkhoven, R. Arens, M. Correia-Neves, K. Schepers, T.N. Schumacher, Recruitment of antigen-specific CD8+ T cells in response to infection is markedly efficient, Science 325(5945) (2009) 1265-9.

[100] V.R. Buchholz, T.N. Schumacher, D.H. Busch, T Cell Fate at the Single-Cell Level, Annu Rev Immunol 34 (2016) 65-92.

[101] C. Stemberger, K.M. Huster, M. Koffler, F. Anderl, M. Schiemann, H. Wagner, D.H. Busch, A single naive CD8+ T cell precursor can develop into diverse effector and memory subsets, Immunity 27(6) (2007) 985-97.

[102] C. Gerlach, J.W. van Heijst, E. Swart, D. Sie, N. Armstrong, R.M. Kerkhoven, D. Zehn, M.J. Bevan, K. Schepers, T.N. Schumacher, One naive T cell, multiple fates in CD8+ T cell differentiation, J Exp Med 207(6) (2010) 1235-46.

[103] Y.L. Cho, M. Flossdorf, L. Kretschmer, T. Höfer, D.H. Busch, V.R. Buchholz, TCR Signal Quality Modulates Fate Decisions of Single CD4(+) T Cells in a Probabilistic Manner, Cell Rep 20(4) (2017) 806-818.

[104] T. Boehm, M. Hirano, S.J. Holland, S. Das, M. Schorpp, M.D. Cooper, Evolution of Alternative Adaptive Immune Systems in Vertebrates, Annu Rev Immunol 36 (2018) 19-42.

[105] D.G. Schatz, Y. Ji, Recombination centres and the orchestration of V(D)J recombination, Nat Rev Immunol 11(4) (2011) 251-63.

[106] N. Bosc, M.P. Lefranc, The mouse (Mus musculus) T cell receptor beta variable (TRBV), diversity (TRBD) and joining (TRBJ) genes, Exp Clin Immunogenet 17(4) (2000) 216-28.

[107] M. Migalska, A. Sebastian, J. Radwan, Profiling of the TCRβ repertoire in non-model species using high-throughput sequencing, Sci Rep 8(1) (2018) 11613.

[108] T. Dupic, Q. Marcou, A.M. Walczak, T. Mora, Genesis of the αβ T-cell receptor, PLoS Comput Biol 15(3) (2019) e1006874.

[109] E.A. Motea, A.J. Berdis, Terminal deoxynucleotidyl transferase: the story of a misguided DNA polymerase, Biochim Biophys Acta 1804(5) (2010) 1151-66.

[110] H.S. Robins, S.K. Srivastava, P.V. Campregher, C.J. Turtle, J. Andriesen, S.R. Riddell, C.S. Carlson, E.H. Warren, Overlap and effective size of the human CD8+ T cell receptor repertoire, Sci Transl Med 2(47) (2010) 47ra64.

[111] S. Mickelsen, C. Snyder, K. Trujillo, M. Bogue, D.B. Roth, K. Meek, Modulation of terminal deoxynucleotidyltransferase activity by the DNA-dependent protein kinase, J Immunol 163(2) (1999) 834-43.

[112] M. Basu, M.V. Hegde, M.J. Modak, Synthesis of compositionally unique DNA by terminal deoxynucleotidyl transferase, Biochem Biophys Res Commun 111(3) (1983) 1105-12.

[113] M.M. Davis, P.J. Bjorkman, T-cell antigen receptor genes and T-cell recognition, Nature 334(6181) (1988) 395-402.

[114] V.I. Zarnitsyna, B.D. Evavold, L.N. Schoettle, J.N. Blattman, R. Antia, Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire, Front Immunol 4 (2013) 485.

[115] A. Murugan, T. Mora, A.M. Walczak, C.G. Callan, Jr., Statistical inference of the generation probability of T-cell receptors from sequence repertoires, Proc Natl Acad Sci U S A 109(40) (2012) 16161-6.

[116] L. Klein, B. Kyewski, P.M. Allen, K.A. Hogquist, Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see), Nat Rev Immunol 14(6) (2014) 377-91.

[117] G. Lythe, R.E. Callard, R.L. Hoare, C. Molina-París, How many TCR clonotypes does a body maintain?, J Theor Biol 389 (2016) 214-24.

[118] V. Venturi, D.A. Price, D.C. Douek, M.P. Davenport, The molecular basis for public T-cell responses?, Nat Rev Immunol 8(3) (2008) 231-8.

[119] K.C. Garcia, M. Degano, R.L. Stanfield, A. Brunmark, M.R. Jackson, P.A. Peterson, L. Teyton, I.A. Wilson, An alphabeta T cell receptor structure at 2.5 A and its orientation in the TCR-MHC complex, Science 274(5285) (1996) 209-19.

[120] D. Housset, B. Malissen, What do TCR-pMHC crystal structures teach us about MHC restriction and alloreactivity?, Trends Immunol 24(8) (2003) 429-37.

[121] B. Al-Lazikani, A.M. Lesk, C. Chothia, Canonical structures for the hypervariable regions of T cell alphabeta receptors, J Mol Biol 295(4) (2000) 979-95.

[122] X. Ysern, H. Li, R.A. Mariuzza, Imperfect interfaces, Nat Struct Biol 5(6) (1998) 412-4.

[123] B. Liu, W. Chen, B.D. Evavold, C. Zhu, Accumulation of dynamic catch bonds between TCR and agonist peptide-MHC triggers T cell signaling, Cell 157(2) (2014) 357-368.

[124] B. Liu, E.M. Kolawole, B.D. Evavold, Mechanobiology of T Cell Activation: To Catch a Bond, Annu Rev Cell Dev Biol 37 (2021) 65-87.

[125] L. Limozin, M. Bridge, P. Bongrand, O. Dushek, P.A. van der Merwe, P. Robert, TCR-pMHC kinetics under force in a cell-free system show no intrinsic catch bond, but a minimal encounter duration before binding, Proc Natl Acad Sci U S A 116(34) (2019) 16943-16948.

[126] W.L. Lo, N.J. Felix, J.J. Walters, H. Rohrs, M.L. Gross, P.M. Allen, An endogenous peptide positively selects and augments the activation and survival of peripheral CD4+ T cells, Nat Immunol 10(11) (2009) 1155-61.

[127] J. Juang, P.J. Ebert, D. Feng, K.C. Garcia, M. Krogsgaard, M.M. Davis, Peptide-MHC heterodimers show that thymic positive selection requires a more restricted set of self-peptides than negative selection, J Exp Med 207(6) (2010) 1223-34.

[128] J.D. Stone, A.S. Chervin, D.M. Kranz, T-cell receptor binding affinities and kinetics: impact on T-cell activity and specificity, Immunology 126(2) (2009) 165-76.

[129] M.J. Grusby, H. Auchincloss, Jr., R. Lee, R.S. Johnson, J.P. Spencer, M. Zijlstra, R. Jaenisch, V.E. Papaioannou, L.H. Glimcher, Mice lacking major histocompatibility complex class I and class II molecules, Proc Natl Acad Sci U S A 90(9) (1993) 3913-7.

[130] B.A. Irving, F.W. Alt, N. Killeen, Thymocyte development in the absence of pre-T cell receptor extracellular immunoglobulin domains, Science 280(5365) (1998) 905-8.

[131] A.L. Crump, M.J. Grusby, L.H. Glimcher, H. Cantor, Thymocyte development in major histocompatibility complex-deficient mice: evidence for stochastic commitment to the CD4 and CD8 lineages, Proc Natl Acad Sci U S A 90(22) (1993) 10739-43.

[132] D.K. Das, R.J. Mallis, J.S. Duke-Cohan, R.E. Hussey, P.W. Tetteh, M. Hilton, G. Wagner, M.J. Lang, E.L. Reinherz, Pre-T Cell Receptors (Pre-TCRs) Leverage Vβ Complementarity Determining Regions (CDRs) and Hydrophobic Patch in Mechanosensing Thymic Self-ligands, J Biol Chem 291(49) (2016) 25292-25305.

[133] R.J. Mallis, H. Arthanari, M.J. Lang, E.L. Reinherz, G. Wagner, NMR-directed design of pre-TCRβ and pMHC molecules implies a distinct geometry for pre-TCR relative to αβTCR recognition of pMHC, J Biol Chem 293(3) (2018) 754-766.

[134] X. Li, R. Mizsei, K. Tan, R.J. Mallis, J.S. Duke-Cohan, A. Akitsu, P.W. Tetteh, A. Dubey, W. Hwang, G. Wagner, M.J. Lang, H. Arthanari, J.H. Wang, E.L. Reinherz, Pre-T cell receptors topologically sample self-ligands during thymocyte β-selection, Science 371(6525) (2021) 181-185.

[135] R.J. Mallis, K. Bai, H. Arthanari, R.E. Hussey, M. Handley, Z. Li, L. Chingozha, J.S. Duke-Cohan, H. Lu, J.H. Wang, C. Zhu, G. Wagner, E.L. Reinherz, Pre-TCR ligand binding impacts thymocyte development before αβTCR expression, Proc Natl Acad Sci U S A 112(27) (2015) 8373-8.

[136] J.S. Duke-Cohan, A. Akitsu, R.J. Mallis, C.M. Messier, P.H. Lizotte, J.C. Aster, W. Hwang, M.J. Lang, E.L. Reinherz, Pre-T cell receptor Self-MHC Sampling Restricts Thymocyte Dedifferentiation, Nature (2022).

[137] L.K. Lutes, Z. Steier, L.L. McIntyre, S. Pandey, J. Kaminski, A.R. Hoover, S. Ariotti, A. Streets, N. Yosef, E.A. Robey, T cell self-reactivity during thymic development dictates the timing of positive selection, bioRxiv (2021) 2021.01.18.427079.

[138] N. Vrisekoop, J.P. Monteiro, J.N. Mandl, R.N. Germain, Revisiting thymic positive selection and the mature T cell repertoire for antigen, Immunity 41(2) (2014) 181-90.

[139] V. Guichard, N. Bonilla, A. Durand, A. Audemard-Verger, T. Guilbert, B. Martin, B. Lucas, C. Auffray, Calcium-mediated shaping of naive CD4 T-cell phenotype and function, Elife 6 (2017).

[140] D. Rogers, A. Sood, H. Wang, J.J.P. van Beek, T.J. Rademaker, P. Artusa, C. Schneider, C. Shen, D.C. Wong, A. Bhagrath, M. Lebel, S.A. Condotta, M.J. Richer, A.J. Martins, J.S. Tsang, L.B. Barreiro, P. François, D. Langlais, H.J. Melichar, J. Textor, J.N. Mandl, Pre-existing chromatin accessibility and gene expression differences among naive CD4(+) T cells influence effector potential, Cell Rep 37(9) (2021) 110064.

[141] R.B. Fulton, S.E. Hamilton, Y. Xing, J.A. Best, A.W. Goldrath, K.A. Hogquist, S.C. Jameson, The TCR's sensitivity to self peptide-MHC dictates the ability of naive CD8(+) T cells to respond to foreign antigens, Nat Immunol 16(1) (2015) 107-17.

[142] A. Sood, M.E. Lebel, M. Dong, M. Fournier, S.J. Vobecky, E. Haddad, J.S. Delisle, J.N. Mandl, N. Vrisekoop, H.J. Melichar, CD5 levels define functionally heterogeneous populations of naive human CD4(+) T cells, Eur J Immunol (2021).

[143] J. Textor, F. Buytenhuijs, D. Rogers, È. Mallet Gauthier, S. Sultan, I.M.N. Wortel, K. Kalies, A. Fähnrich, R. Pagel, H.J. Melichar, J. Westermann, J.N. Mandl, Machine learning analysis of the T cell receptor repertoire identifies sequence features that predict self-reactivity, bioRxiv (2022) 2022.11.23.517563.

[144] S. Gilfillan, C. Waltzinger, C. Benoist, D. Mathis, More efficient positive selection of thymocytes in mice lacking terminal deoxynucleotidyl transferase, Int Immunol 6(11) (1994) 1681-6.

[145] M.A. Gavin, M.J. Bevan, Increased peptide promiscuity provides a rationale for the lack of N regions in the neonatal T cell repertoire, Immunity 3(6) (1995) 793-800.

[146] I.M.N. Wortel, C. Keşmir, R.J. de Boer, J.N. Mandl, J. Textor, Is T Cell Negative Selection a Learning Algorithm?, Cells 9(3) (2020).

[147] M.E. Birnbaum, J.L. Mendoza, D.K. Sethi, S. Dong, J. Glanville, J. Dobbins, E. Ozkan, M.M. Davis, K.W. Wucherpfennig, K.C. Garcia, Deconstructing the peptide-MHC specificity of T cell recognition, Cell 157(5) (2014) 1073-87.

[148] T.W. McKeithan, Kinetic proofreading in T-cell receptor signal transduction, Proc Natl Acad Sci U S A 92(11) (1995) 5042-6.

[149] W.L. Lo, A. Weiss, Adapting T Cell Receptor Ligand Discrimination Capability via LAT, Front Immunol 12 (2021) 673196.

[150] A.K. Chakraborty, A. Weiss, Insights into the initiation of TCR signaling, Nat Immunol 15(9) (2014) 798-807.

[151] N.R. Gascoigne, J. Casas, J. Brzostek, V. Rybakin, Initiation of TCR phosphorylation and signal transduction, Front Immunol 2 (2011) 72.

[152] G. Voisinne, M. Locard-Paulet, C. Froment, E. Maturin, M.G. Menoita, L. Girard, V. Mellado, O. Burlet-Schiltz, B. Malissen, A. Gonzalez de Peredo, R. Roncagalli, Kinetic proofreading through the multi-step activation of the ZAP70 kinase underlies early T cell ligand discrimination, Nat Immunol 23(9) (2022) 1355-1364.

[153] R.S. Ganti, W.L. Lo, D.B. McAffee, J.T. Groves, A. Weiss, A.K. Chakraborty, How the T cell signaling network processes information to discriminate between self and agonist ligands, Proc Natl Acad Sci U S A 117(42) (2020) 26020-26030.

[154] D.R. Myers, T. Lau, E. Markegard, H.W. Lim, H. Kasler, M. Zhu, A. Barczak, J.P. Huizar, J. Zikherman, D.J. Erle, W. Zhang, E. Verdin, J.P. Roose, Tonic LAT-HDAC7 Signals Sustain Nur77 and Irf4 Expression to Tune Naive CD4 T Cells, Cell Rep 19(8) (2017) 1558-1571.

[155] W.L. Lo, N.H. Shah, S.A. Rubin, W. Zhang, V. Horkova, I.R. Fallahee, O. Stepanek, L.I. Zon, J. Kuriyan, A. Weiss, Slow phosphorylation of a tyrosine residue in LAT optimizes T cell ligand discrimination, Nat Immunol 20(11) (2019) 1481-1493.

[156] W.-L. Lo, M. Kuhlmann, G. Rizzuto, H.A. Ekiz, E.M. Kolawole, M.P. Revelo, R. Andargachew, Z. Li, Y.-L. Tsai, A. Marson, B.D. Evavold, D. Zehn, A. Weiss, A single-amino acid substitution in the adaptor LAT accelerates TCR proofreading kinetics and alters T-cell selection, maintenance and function, Nature Immunology (2023).

[157] G. Altan-Bonnet, R.N. Germain, Modeling T cell antigen discrimination based on feedback control of digital ERK responses, PLoS Biol 3(11) (2005) e356.

[158] I. Stefanová, B. Hemmer, M. Vergelli, R. Martin, W.E. Biddison, R.N. Germain, TCR ligand discrimination is enforced by competing ERK positive and SHP-1 negative feedback pathways, Nat Immunol 4(3) (2003) 248-54.

[159] M. Hebeisen, S.G. Oberle, D. Presotto, D.E. Speiser, D. Zehn, N. Rufer, Molecular insights for optimizing T cell receptor specificity against cancer, Front Immunol 4 (2013) 154.

[160] G. Dolton, E. Zervoudi, C. Rius, A. Wall, H.L. Thomas, A. Fuller, L. Yeo, M. Legut, S. Wheeler, M. Attaf, D.M. Chudakov, E. Choy, M. Peakman, A.K. Sewell, Optimized Peptide-MHC Multimer Protocols for Detection and Isolation of Autoimmune T-Cells, Front Immunol 9 (2018) 1378.

[161] K. Matsui, J.J. Boniface, P. Steffner, P.A. Reay, M.M. Davis, Kinetics of T-cell receptor binding to peptide/I-Ek complexes: correlation of the dissociation rate with T-cell responsiveness, Proc Natl Acad Sci U S A 91(26) (1994) 12862-6.

[162] M. Corr, A.E. Slanetz, L.F. Boyd, M.T. Jelonek, S. Khilko, B.K. al-Ramadi, Y.S. Kim, S.E. Maher, A.L. Bothwell, D.H. Margulies, T cell receptor-MHC class I peptide interactions: affinity, kinetics, and specificity, Science 265(5174) (1994) 946-9.

[163] S.E. Chesla, P. Selvaraj, C. Zhu, Measuring two-dimensional receptor-ligand binding kinetics by micropipette, Biophys J 75(3) (1998) 1553-72.

[164] J.D. Altman, P.A. Moss, P.J. Goulder, D.H. Barouch, M.G. McHeyzer-Williams, J.I. Bell, A.J. McMichael, M.M. Davis, Phenotypic analysis of antigen-specific T lymphocytes, Science 274(5284) (1996) 94-6.

[165] S. Sims, C. Willberg, P. Klenerman, MHC–peptide tetramers for the analysis of antigen-specific T cells, Expert Review of Vaccines 9(7) (2010) 765-774.

[166] R. Andargachew, R.J. Martinez, E.M. Kolawole, B.D. Evavold, CD4 T Cell Affinity Diversity Is Equally Maintained during Acute and Chronic Infection, J Immunol 201(1) (2018) 19-30.

[167] P. Kisielow, H. Blüthmann, U.D. Staerz, M. Steinmetz, H. von Boehmer, Tolerance in T-cell-receptor transgenic mice involves deletion of nonmature CD4+8+ thymocytes, Nature 333(6175) (1988) 742-6.

[168] P.J. Ebert, S. Jiang, J. Xie, Q.J. Li, M.M. Davis, An endogenous positively selecting peptide enhances mature T cell responses and becomes an autoantigen in the absence of microRNA miR-181a, Nat Immunol 10(11) (2009) 1162-9.

[169] K.A. Hogquist, A.J. Tomlinson, W.C. Kieper, M.A. McGargill, M.C. Hart, S. Naylor, S.C. Jameson, Identification of a naturally occurring ligand for thymic positive selection, Immunity 6(4) (1997) 389-99.

[170] F.R. Santori, W.C. Kieper, S.M. Brown, Y. Lu, T.A. Neubert, K.L. Johnson, S. Naylor, S. Vukmanović, K.A. Hogquist, S.C. Jameson, Rare, structurally homologous self-peptides promote thymocyte positive selection, Immunity 17(2) (2002) 131-42.

[171] J.J. Perez-Villar, G.S. Whitney, M.A. Bowen, D.H. Hewgill, A.A. Aruffo, S.B. Kanner, CD5 negatively regulates the T-cell antigen receptor signal transduction pathway: involvement of SH2-containing phosphotyrosine phosphatase SHP-1, Mol Cell Biol 19(4) (1999) 2903-12.

[172] H.S. Azzam, A. Grinberg, K. Lui, H. Shen, E.W. Shores, P.E. Love, CD5 expression is developmentally regulated by T cell receptor (TCR) signals and TCR avidity, J Exp Med 188(12) (1998) 2301-11.

[173] K. Smith, B. Seddon, M.A. Purbhoo, R. Zamoyska, A.G. Fisher, M. Merkenschlager, Sensory adaptation in naive peripheral CD4 T cells, J Exp Med 194(9) (2001) 1253-61.

[174] G. Blaize, H. Daniels-Treffandier, M. Aloulou, N. Rouquié, C. Yang, M. Marcellin, M. Gador, M. Benamar, M. Ducatez, K.D. Song, O. Burlet-Schiltz, A. Saoudi, P.E. Love, N. Fazilleau, A. Gonzalez de Peredo, R. Lesourne, CD5 signalosome coordinates antagonist TCR signals to control the generation of Treg cells induced by foreign antigens, Proc Natl Acad Sci U S A 117(23) (2020) 12969-12979.

[175] E. Burgueño-Bucio, C.A. Mier-Aguilar, G. Soldevila, The multiple faces of CD5, J Leukoc Biol 105(5) (2019) 891-904.

[176] L. Odagiu, J. May, S. Boulet, T.A. Baldwin, N. Labrecque, Role of the Orphan Nuclear Receptor NR4A Family in T-Cell Biology, Front Endocrinol (Lausanne) 11 (2020) 624122.

[177] A.E. Moran, K.L. Holzapfel, Y. Xing, N.R. Cunningham, J.S. Maltzman, J. Punt, K.A. Hogquist, T cell receptor signal strength in Treg and iNKT cell development demonstrated by a novel fluorescent reporter mouse, J Exp Med 208(6) (2011) 1279-89.

[178] J. Zikherman, R. Parameswaran, A. Weiss, Endogenous antigen tunes the responsiveness of naive B cells but not T cells, Nature 489(7414) (2012) 160-4.

[179] B. Martin, C. Auffray, A. Delpoux, A. Pommier, A. Durand, C. Charvet, P. Yakonowsky, H. de Boysson, N. Bonilla, A. Audemard, T. Sparwasser, B.L. Salomon, B. Malissen, B. Lucas, Highly self-reactive naive CD4 T cells are prone to differentiate into regulatory T cells, Nat Commun 4 (2013) 2209.

[180] W.M. Zinzow-Kramer, A. Weiss, B.B. Au-Yeung, Adaptation by naïve CD4(+) T cells to self-antigen-dependent TCR signaling induces functional heterogeneity and tolerance, Proc Natl Acad Sci U S A 116(30) (2019) 15160-15169.

[181] W.M. Zinzow-Kramer, E.M. Kolawole, J. Eggert, B.D. Evavold, C.D. Scharer, B.B. Au-Yeung, Strong Basal/Tonic TCR Signals Are Associated with Negative Regulation of Naive CD4⁺ T Cells, Immunohorizons 6(9) (2022) 671-683.

[182] C.A. Matson, S. Choi, F. Livak, B. Zhao, A. Mitra, P.E. Love, N.J. Singh, CD5 dynamically calibrates basal NF-κB signaling in T cells during thymic development and peripheral activation, Proc Natl Acad Sci U S A 117(25) (2020) 14342-14353.

[183] J.G. Henderson, A. Opejin, A. Jones, C. Gross, D. Hawiger, CD5 instructs extrathymic regulatory T cell development in response to self and tolerizing antigens, Immunity 42(3) (2015) 471-83.

[184] M.J. Palmer, V.S. Mahajan, J. Chen, D.J. Irvine, D.A. Lauffenburger, Signaling thresholds govern heterogeneity in IL-7-receptor-mediated responses of naïve CD8(+) T cells, Immunol Cell Biol 89(5) (2011) 581-94.

[185] G.C. Ang, History of skin transplantation, Clin Dermatol 23(4) (2005) 320-4.

[186] M. Kohlhauser, H. Luze, S.P. Nischwitz, L.P. Kamolz, Historical Evolution of Skin Grafting—A Journey through Time, Medicina 57(4) (2021) 348.

[187] R.E. Billingham, L. Brent, P.B. Medawar, Actively acquired tolerance of foreign cells, Nature 172(4379) (1953) 603-6.

[188] J.W. Kappler, T. Wade, J. White, E. Kushnir, M. Blackman, J. Bill, N. Roehm, P. Marrack, A T cell receptor V beta segment that imparts reactivity to a class II major histocompatibility complex product, Cell 49(2) (1987) 263-71.

[189] V. Detours, R. Mehr, A.S. Perelson, Deriving quantitative constraints on T cell selection from data on the mature T cell repertoire, J Immunol 164(1) (2000) 121-8.

[190] V. Müller, S. Bonhoeffer, Quantitative constraints on the scope of negative selection, Trends Immunol 24(3) (2003) 132-5.

[191] D.P. Granados, C.M. Laumont, P. Thibault, C. Perreault, The nature of self for T cells-a systems-level perspective, Curr Opin Immunol 34 (2015) 1-8.

[192] S.R. Starck, V. Jiang, M. Pavon-Eternod, S. Prasad, B. McCarthy, T. Pan, N. Shastri, Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I, Science 336(6089) (2012) 1719-23.

[193] N. Vigneron, B.J. Van den Eynde, Proteasome subtypes and the processing of tumor antigens: increasing antigenic diversity, Curr Opin Immunol 24(1) (2012) 84-91.

[194] C. Wong, D.R. Goldstein, Impact of aging on antigen presentation cell function of dendritic cells, Curr Opin Immunol 25(4) (2013) 535-41.

[195] Z. Liang, X. Dong, Z. Zhang, Q. Zhang, Y. Zhao, Age-related thymic involution: Mechanisms and functional impact, Aging Cell 21(8) (2022) e13671.

[196] J. Srinivasan, J.N. Lancaster, N. Singarapu, L.P. Hale, L.I.R. Ehrlich, E.R. Richie, Age-Related Changes in Thymic Central Tolerance, Front Immunol 12 (2021) 676236.

[197] Y. Gonskikh, N. Polacek, Alterations of the translation apparatus during aging and stress response, Mech Ageing Dev 168 (2017) 30-36.

[198] C. Viret, C. Lamare, M. Guiraud, N. Fazilleau, A. Bour, B. Malissen, A. Carrier, S. Guerder, Thymus-specific serine protease contributes to the diversification of the functional endogenous CD4 T cell receptor repertoire, J Exp Med 208(1) (2011) 3-11.

[199] T. Nakagawa, W. Roth, P. Wong, A. Nelson, A. Farr, J. Deussing, J.A. Villadangos, H. Ploegh, C. Peters, A.Y. Rudensky, Cathepsin L: critical role in Ii degradation and CD4 T cell selection in the thymus, Science 280(5362) (1998) 450-3.

[200] S. Murata, K. Sasaki, T. Kishimoto, S. Niwa, H. Hayashi, Y. Takahama, K. Tanaka, Regulation of CD8+ T cell development by thymus-specific proteasomes, Science 316(5829) (2007) 1349-53.

[201] K. Honey, T. Nakagawa, C. Peters, A. Rudensky, Cathepsin L regulates CD4+ T cell selection independently of its effect on invariant chain: a role in the generation of positively selecting peptide ligands, J Exp Med 195(10) (2002) 1349-58.

[202] T. Nitta, S. Murata, K. Sasaki, H. Fujii, A.M. Ripen, N. Ishimaru, S. Koyasu, K. Tanaka, Y. Takahama, Thymoproteasome shapes immunocompetent repertoire of CD8+ T cells, Immunity 32(1) (2010) 29-40.

[203] Y. Xing, S.C. Jameson, K.A. Hogquist, Thymoproteasome subunit-β5T generates peptide-MHC complexes specialized for positive selection, Proc Natl Acad Sci U S A 110(17) (2013) 6979-84.

[204] C.S. Hsieh, P. deRoos, K. Honey, C. Beers, A.Y. Rudensky, A role for cathepsin L and cathepsin S in peptide generation for MHC class II presentation, J Immunol 168(6) (2002) 2618-25.

[205] M. Frantzeskakis, Y. Takahama, I. Ohigashi, The Role of Proteasomes in the Thymus, Front Immunol 12 (2021) 646209.

[206] J. Derbinski, J. Gäbler, B. Brors, S. Tierling, S. Jonnakuty, M. Hergenhahn, L. Peltonen, J. Walter, B. Kyewski, Promiscuous gene expression in thymic epithelial cells is regulated at multiple levels, J Exp Med 202(1) (2005) 33-45.

[207] D. Mathis, C. Benoist, Aire, Annu Rev Immunol 27 (2009) 287-312.

[208] M.S. Anderson, E.S. Venanzi, L. Klein, Z. Chen, S.P. Berzins, S.J. Turley, H. von Boehmer, R. Bronson, A. Dierich, C. Benoist, D. Mathis, Projection of an immunological self shadow within the thymus by the aire protein, Science 298(5597) (2002) 1395-401.

[209] D.A. Michelson, K. Hase, T. Kaisho, C. Benoist, D. Mathis, Thymic epithelial cells co-opt lineage-defining transcription factors to eliminate autoreactive T cells, Cell 185(14) (2022) 2542-2558.e18.

[210] C.N. Miller, I. Proekt, J. von Moltke, K.L. Wells, A.R. Rajpurkar, H. Wang, K. Rattay, I.S. Khan, T.C. Metzger, J.L. Pollack, A.C. Fries, W.W. Lwin, E.J. Wigton, A.V. Parent, B. Kyewski, D.J. Erle, K.A. Hogquist, L.M. Steinmetz, R.M. Locksley, M.S. Anderson, Thymic tuft cells promote an IL-4-enriched medulla and shape thymocyte development, Nature 559(7715) (2018) 627-631.

[211] C. Bornstein, S. Nevo, A. Giladi, N. Kadouri, M. Pouzolles, F. Gerbe, E. David, A. Machado, A. Chuprin, B. Tóth, O. Goldberg, S. Itzkovitz, N. Taylor, P. Jay, V.S. Zimmermann, J. Abramson, I. Amit, Single-cell mapping of the thymic stroma identifies IL-25-producing tuft epithelial cells, Nature 559(7715) (2018) 622-626.

[212] M. Ennamorati, C. Vasudevan, K. Clerkin, S. Halvorsen, S. Verma, S. Ibrahim, S. Prosper, C. Porter, V. Yeliseyev, M. Kim, J. Gardecki, S. Sassi, G. Tearney, B.J. Cherayil, L. Bry, B. Seed, N. Jain, Intestinal microbes influence development of thymic lymphocytes in early life, Proc Natl Acad Sci U S A 117(5) (2020) 2570-2578.

[213] D.F. Zegarra-Ruiz, D.V. Kim, K. Norwood, M. Kim, W.H. Wu, F.B. Saldana-Morales, A.A. Hill, S. Majumdar, S. Orozco, R. Bell, J.L. Round, R.S. Longman, T. Egawa, M.L. Bettini, G.E. Diehl, Thymic development of gut-microbiota-specific T cells, Nature 594(7863) (2021) 413-417.

[214] R. Bonasio, M.L. Scimone, P. Schaerli, N. Grabie, A.H. Lichtman, U.H. von Andrian, Clonal deletion of thymocytes by circulating dendritic cells homing to the thymus, Nat Immunol 7(10) (2006) 1092-100.

[215] J. Li, J. Park, D. Foss, I. Goldschneider, Thymus-homing peripheral dendritic cells constitute two of the three major subsets of dendritic cells in the steady-state thymus, J Exp Med 206(3) (2009) 607-22.

[216] D. de Verteuil, T.L. Muratore-Schroeder, D.P. Granados, M.H. Fortier, M.P. Hardy, A. Bramoullé, E. Caron, K. Vincent, S. Mader, S. Lemieux, P. Thibault, C. Perreault, Deletion of immunoproteasome subunits imprints on the transcriptome and has a broad impact on peptides presented by major histocompatibility complex I molecules, Mol Cell Proteomics 9(9) (2010) 2034-47.

[217] B. Lucas, R.N. Germain, T-cell repertoire: political correctness in the immune system, Curr Biol 6(7) (1996) 783-7.

[218] J.N. Mandl, P. Torabi-Parizi, R.N. Germain, Visualization and dynamic analysis of host-pathogen interactions, Curr Opin Immunol 29 (2014) 8-15.

[219] C. Tanchot, F.A. Lemonnier, B. Pérarnau, A.A. Freitas, B. Rocha, Differential requirements for survival and proliferation of CD8 naïve or memory T cells, Science 276(5321) (1997) 2057-62.

[220] S. Takeda, H.R. Rodewald, H. Arakawa, H. Bluethmann, T. Shimizu, MHC class II molecules are not required for survival of newly generated CD4+ T cells, but affect their long-term life span, Immunity 5(3) (1996) 217-28.

[221] N. Labrecque, L.S. Whitfield, R. Obst, C. Waltzinger, C. Benoist, D. Mathis, How much TCR does a T cell need?, Immunity 15(1) (2001) 71-82.

[222] B. Seddon, R. Zamoyska, TCR signals mediated by Src family kinases are essential for the survival of naive T cells, J Immunol 169(6) (2002) 2997-3005.

[223] B. Polic, D. Kunkel, A. Scheffold, K. Rajewsky, How alpha beta T cells deal with induced TCR alpha ablation, Proc Natl Acad Sci U S A 98(15) (2001) 8744-9.

[224] S.R. Jacobs, R.D. Michalek, J.C. Rathmell, IL-7 is essential for homeostatic control of T cell metabolism in vivo, J Immunol 184(7) (2010) 3461-9.

[225] K.P. Lam, R. Kühn, K. Rajewsky, In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death, Cell 90(6) (1997) 1073-83.

[226] J.C. Rathmell, M.G. Vander Heiden, M.H. Harris, K.A. Frauwirth, C.B. Thompson, In the absence of extrinsic signals, nutrient utilization by lymphocytes is insufficient to maintain either cell size or viability, Mol Cell 6(3) (2000) 683-92.

[227] J.J. Obar, K.M. Khanna, L. Lefrançois, Endogenous naive CD8+ T cell precursor frequency regulates primary and memory responses to infection, Immunity 28(6) (2008) 859-69.

[228] J. Hataye, J.J. Moon, A. Khoruts, C. Reilly, M.K. Jenkins, Naive and memory CD4+ T cell survival controlled by clonal abundance, Science 312(5770) (2006) 114-6.

[229] J.N. Mandl, R.N. Germain, Focusing in on T cell cross-reactivity, Cell 157(5) (2014) 1006-8.

[230] N.J. Singh, J.K. Bando, R.H. Schwartz, Subsets of nonclonal neighboring CD4+ T cells specifically regulate the frequency of individual antigen-reactive T cells, Immunity 37(4) (2012) 735-46.

[231] S.C. Jameson, T cell homeostasis: keeping useful T cells alive and live T cells useful, Semin Immunol 17(3) (2005) 231-7.

[232] N. Vrisekoop, P. Artusa, J.P. Monteiro, J.N. Mandl, Weakly self-reactive T-cell clones can homeostatically expand when present at low numbers, Eur J Immunol 47(1) (2017) 68-73.

[233] W.C. Kieper, J.T. Burghardt, C.D. Surh, A role for TCR affinity in regulating naive T cell homeostasis, J Immunol 172(1) (2004) 40-4.

[234] C.T. Moses, K.M. Thorstenson, S.C. Jameson, A. Khoruts, Competition for self ligands restrains homeostatic proliferation of naive CD4 T cells, Proc Natl Acad Sci U S A 100(3) (2003) 1185-90.

[235] T. Hogan, A. Shuvaev, D. Commenges, A. Yates, R. Callard, R. Thiebaut, B. Seddon, Clonally diverse T cell homeostasis is maintained by a common program of cell-cycle control, J Immunol 190(8) (2013) 3985-93.

[236] A. Yates, M. Saini, A. Mathiot, B. Seddon, Mathematical modeling reveals the biological program regulating lymphopenia-induced proliferation, J Immunol 180(3) (2008) 1414-22.

[237] E.G. Houston, Jr., L.E. Higdon, P.J. Fink, Recent thymic emigrants are preferentially incorporated only into the depleted T-cell pool, Proc Natl Acad Sci U S A 108(13) (2011) 5366-71.

[238] A.M. Berkley, P.J. Fink, Cutting edge: CD8+ recent thymic emigrants exhibit increased responses to low-affinity ligands and improved access to peripheral sites of inflammation, J Immunol 193(7) (2014) 3262-6.

[239] R.S. Paiva, A.C. Lino, M.L. Bergman, I. Caramalho, A.E. Sousa, S. Zelenay, J. Demengeot, Recent thymic emigrants are the preferential precursors of regulatory T cells differentiated in the periphery, Proc Natl Acad Sci U S A 110(16) (2013) 6494-9.

[240] L.K. Lutes, Z. Steier, L.L. McIntyre, S. Pandey, J. Kaminski, A.R. Hoover, S. Ariotti, A. Streets, N. Yosef, E.A. Robey, T cell self-reactivity during thymic development dictates the timing of positive selection, Elife 10 (2021).

[241] M.A. ElTanbouly, Y. Zhao, E. Nowak, J. Li, E. Schaafsma, I. Le Mercier, S. Ceeraz, J.L. Lines, C. Peng, C. Carriere, X. Huang, M. Day, B. Koehn, S.W. Lee, M. Silva Morales, K.A. Hogquist, S.C. Jameson, D. Mueller, J. Rothstein, B.R. Blazar, C. Cheng, R.J. Noelle, VISTA is a checkpoint regulator for naive T cell quiescence and peripheral tolerance, Science 367(6475) (2020).

[242] T. Wolf, W. Jin, G. Zoppi, I.A. Vogel, M. Akhmedov, C.K.E. Bleck, T. Beltraminelli, J.C. Rieckmann, N.J. Ramirez, M. Benevento, S. Notarbartolo, D. Bumann, F. Meissner, B. Grimbacher, M. Mann, A. Lanzavecchia, F. Sallusto, I. Kwee, R. Geiger, Dynamics in protein translation sustaining T cell preparedness, Nat Immunol 21(8) (2020) 927-937.

[243] O. Feinerman, J. Veiga, J.R. Dorfman, R.N. Germain, G. Altan-Bonnet, Variability and robustness in T cell activation from regulated heterogeneity in protein levels, Science 321(5892) (2008) 1081-4.

[244] Y.J. Ju, S.W. Lee, Y.C. Kye, G.W. Lee, H.O. Kim, C.H. Yun, J.H. Cho, Self-reactivity controls functional diversity of naive CD8(+) T cells by co-opting tonic type I interferon, Nat Commun 12(1) (2021) 6059.

[245] J.T. Chang, M.L. Ciocca, I. Kinjyo, V.R. Palanivel, C.E. McClurkin, C.S. Dejong, E.C. Mooney, J.S. Kim, N.C. Steinel, J. Oliaro, C.C. Yin, B.I. Florea, H.S. Overkleeft, L.J. Berg, S.M. Russell, G.A. Koretzky, M.S. Jordan, S.L. Reiner, Asymmetric proteasome segregation as a mechanism for unequal partitioning of the transcription factor T-bet during T lymphocyte division, Immunity 34(4) (2011) 492-504.

[246] K.N. Pollizzi, I.H. Sun, C.H. Patel, Y.C. Lo, M.H. Oh, A.T. Waickman, A.J. Tam, R.L. Blosser, J. Wen, G.M. Delgoffe, J.D. Powell, Asymmetric inheritance of mTORC1 kinase activity during division dictates CD8(+) T cell differentiation, Nat Immunol 17(6) (2016) 704-11.

[247] K.C. Verbist, C.S. Guy, S. Milasta, S. Liedmann, M.M. Kamiński, R. Wang, D.R. Green, Metabolic maintenance of cell asymmetry following division in activated T lymphocytes, Nature 532(7599) (2016) 389-93.

[248] I. Stefanová, J.R. Dorfman, R.N. Germain, Self-recognition promotes the foreign antigen sensitivity of naive T lymphocytes, Nature 420(6914) (2002) 429-34.

[249] K. Takada, S.C. Jameson, Self-class I MHC molecules support survival of naive CD8 T cells, but depress their functional sensitivity through regulation of CD8 expression levels, J Exp Med 206(10) (2009) 2253-69.

[250] K.A. Hogquist, S.C. Jameson, The self-obsession of T cells: how TCR signaling thresholds affect fate 'decisions' and effector function, Nat Immunol 15(9) (2014) 815-23.

[251] J.H. Park, S. Adoro, P.J. Lucas, S.D. Sarafova, A.S. Alag, L.L. Doan, B. Erman, X. Liu, W. Ellmeier, R. Bosselut, L. Feigenbaum, A. Singer, 'Coreceptor tuning': cytokine signals transcriptionally tailor CD8 coreceptor expression to the self-specificity of the TCR, Nat Immunol 8(10) (2007) 1049-59.

[252] H. Jamaleddine, D. Rogers, G. Perreault, J.N. Mandl, A. Khadra, Chronic infection control relies on T cells with lower foreign antigen binding strength generated by N-nucleotide diversity, bioRxiv (2022) 2022.06.26.497644.

[253] A. Tsitsiklis, D.J. Bangs, L.K. Lutes, S.W. Chan, K.M. Geiger, A.J. Modzelewski, L. Labarta-Bajo, Y. Wang, E.I. Zuniga, S. Dai, E.A. Robey, An Unusual MHC Molecule Generates Protective CD8+ T Cell Responses to Chronic Infection, Front Immunol 11 (2020) 1464.

[254] D. DiToro, C.J. Winstead, D. Pham, S. Witte, R. Andargachew, J.R. Singer, C.G. Wilson, C.L. Zindl, R.J. Luther, D.J. Silberger, B.T. Weaver, E.M. Kolawole, R.J. Martinez, H. Turner, R.D. Hatton, J.J. Moon, S.S. Way, B.D. Evavold, C.T. Weaver, Differential IL-2 expression defines developmental fates of follicular versus nonfollicular helper T cells, Science 361(6407) (2018).

[255] J.M. Bartleson, A.A. Viehmann Milam, D.L. Donermeyer, S. Horvath, Y. Xia, T. Egawa, P.M. Allen, Strength of tonic T cell receptor signaling instructs T follicular helper cell-fate decisions, Nat Immunol 21(11) (2020) 1384-1396.

[256] M. Dong, P. Artusa, S.A. Kelly, M. Fournier, T.A. Baldwin, J.N. Mandl, H.J. Melichar, Alterations in the Thymic Selection Threshold Skew the Self-Reactivity of the TCR Repertoire in Neonates, J Immunol 199(3) (2017) 965-973.

[257] N.R. Deshpande, H.L. Parrish, M.S. Kuhns, Self-recognition drives the preferential accumulation of promiscuous CD4(+) T-cells in aged mice, Elife 4 (2015) e05949.

[258] S.J. Altschuler, L.F. Wu, Cellular heterogeneity: do differences make a difference?, Cell 141(4) (2010) 559-63.

[259] A. Mayer, T. Mora, O. Rivoire, A.M. Walczak, Diversity of immune strategies explained by adaptation to pathogen statistics, Proc Natl Acad Sci U S A 113(31) (2016) 8630-5.

[260] H.H. Chang, M. Hemberg, M. Barahona, D.E. Ingber, S. Huang, Transcriptome-wide noise controls lineage choice in mammalian progenitor cells, Nature 453(7194) (2008) 544-7.

[261] B. Carter, K. Zhao, The epigenetic basis of cellular heterogeneity, Nat Rev Genet (2020).

[262] A.C. Villani, R. Satija, G. Reynolds, S. Sarkizova, K. Shekhar, J. Fletcher, M. Griesbeck, A. Butler, S. Zheng, S. Lazo, L. Jardine, D. Dixon, E. Stephenson, E. Nilsson, I. Grundberg, D. McDonald, A. Filby, W. Li, P.L. De Jager, O. Rozenblatt-Rosen, A.A. Lane, M. Haniffa, A. Regev, N. Hacohen, Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors, Science 356(6335) (2017).

[263] X. Xie, Q. Shi, P. Wu, X. Zhang, H. Kambara, J. Su, H. Yu, S.Y. Park, R. Guo, Q. Ren, S. Zhang, Y. Xu, L.E. Silberstein, T. Cheng, F. Ma, C. Li, H.R. Luo, Single-cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection, Nat Immunol 21(9) (2020) 1119-1133.

[264] L. Zhou, M.M. Chong, D.R. Littman, Plasticity of CD4+ T cell lineage differentiation, Immunity 30(5) (2009) 646-55.

[265] J. Zhu, H. Yamane, W.E. Paul, Differentiation of effector CD4 T cell populations, Annu Rev Immunol 28 (2010) 445-89.

[266] L. Almeida, M. Lochner, L. Berod, T. Sparwasser, Metabolic pathways in T cell activation and lineage differentiation, Semin Immunol 28(5) (2016) 514-524.

[267] R.M. Rodriguez, C. Lopez-Larrea, B. Suarez-Alvarez, Epigenetic dynamics during CD4(+) T cells lineage commitment, Int J Biochem Cell Biol 67 (2015) 75-85.

[268] S. Crotty, T Follicular Helper Cell Biology: A Decade of Discovery and Diseases, Immunity 50(5) (2019) 1132-1148.

[269] S. Becattini, D. Latorre, F. Mele, M. Foglierini, C. De Gregorio, A. Cassotta, B. Fernandez, S. Kelderman, T.N. Schumacher, D. Corti, A. Lanzavecchia, F. Sallusto, T cell immunity. Functional heterogeneity of human memory CD4^{+} T cell clones primed by pathogens or vaccines, Science 347(6220) (2015) 400-6.

[270] A. Khatun, M.Y. Kasmani, R. Zander, D.M. Schauder, J.P. Snook, J. Shen, X. Wu, R. Burns, Y.G. Chen, C.W. Lin, M.A. Williams, W. Cui, Single-cell lineage mapping of a diverse virus-specific naive CD4 T cell repertoire, J Exp Med 218(3) (2021).

[271] M. Arman, J. Calvo, M.E. Trojanowska, P.N. Cockerill, M. Santana, M. Lopez-Cabrera, J. Vives, F. Lozano, Transcriptional regulation of human CD5: important role of Ets transcription factors in CD5 expression in T cells, J Immunol 172(12) (2004) 7519-29.

[272] J.W. Tung, S.S. Kunnavatana, L.A. Herzenberg, L.A. Herzenberg, The regulation of CD5 expression in murine T cells, BMC Mol Biol 2 (2001) 5.

[273] A.T. Lun, D.J. McCarthy, J.C. Marioni, A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor, F1000Res 5 (2016) 2122.

[274] D. Ordoñez-Rueda, F. Lozano, A. Sarukhan, C. Raman, E.A. Garcia-Zepeda, G. Soldevila, Increased numbers of thymic and peripheral CD4+ CD25+Foxp3+ cells in the absence of CD5 signaling, Eur J Immunol 39(8) (2009) 2233-47.

[275] R.A. Gottschalk, M.M. Hathorn, H. Beuneu, E. Corse, M.L. Dustin, G. Altan-Bonnet, J.P. Allison, Distinct influences of peptide-MHC quality and quantity on in vivo T-cell responses, Proc Natl Acad Sci U S A 109(3) (2012) 881-6.

[276] J.S. Hale, B. Youngblood, D.R. Latner, A.U. Mohammed, L. Ye, R.S. Akondy, T. Wu, S.S. Iyer, R. Ahmed, Distinct memory CD4+ T cells with commitment to T follicular helper- and T helper 1-cell lineages are generated after acute viral infection, Immunity 38(4) (2013) 805-17.

[277] S.E. Hamilton, S.C. Jameson, CD8 T cell quiescence revisited, Trends Immunol 33(5) (2012) 224-30.

[278] I. Yusuf, D.A. Fruman, Regulation of quiescence in lymphocytes, Trends Immunol 24(7) (2003) 380-6.

[279] F. Ren, X. Zhan, G. Martens, J. Lee, D. Center, S.K. Hanson, H. Kornfeld, Pro-IL-16 regulation in activated murine CD4+ lymphocytes, J Immunol 174(5) (2005) 2738-45.

[280] M. Mingueneau, T. Kreslavsky, D. Gray, T. Heng, R. Cruse, J. Ericson, S. Bendall, M.H. Spitzer, G.P. Nolan, K. Kobayashi, H. von Boehmer, D. Mathis, C. Benoist, A.J. Best, J. Knell, A. Goldrath, V. Joic, D. Koller, T. Shay, A. Regev, N. Cohen, P. Brennan, M. Brenner, F. Kim, T. Nageswara Rao, A. Wagers, T. Heng, J. Ericson, K. Rothamel, A. Ortiz-Lopez, D. Mathis, C. Benoist, N.A. Bezman, J.C. Sun, G. Min-Oo, C.C. Kim, L.L. Lanier, J. Miller, B. Brown, M. Merad, E.L. Gautier, C. Jakubzick, G.J. Randolph, P. Monach, D.A. Blair, M.L. Dustin, S.A. Shinton, R.R. Hardy, D. Laidlaw, J. Collins, R. Gazit, D.J. Rossi, N. Malhotra, K. Sylvia, J. Kang, T. Kreslavsky, A. Fletcher, K. Elpek, A. Bellemare-Pelletier, D. Malhotra, S. Turley, The transcriptional landscape of αβ T cell differentiation, Nat Immunol 14(6) (2013) 619-32.

[281] V. Krishnamoorthy, S. Kannanganat, M. Maienschein-Cline, S.L. Cook, J. Chen, N. Bahroos, E. Sievert, E. Corse, A. Chong, R. Sciammas, The IRF4 Gene Regulatory Module Functions as a Read-Write Integrator to Dynamically Coordinate T Helper Cell Fate, Immunity 47(3) (2017) 481-497 e7.

[282] R.J. Johnston, A.C. Poholek, D. DiToro, I. Yusuf, D. Eto, B. Barnett, A.L. Dent, J. Craft, S. Crotty, Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation, Science 325(5943) (2009) 1006-10.

[283] O. Khan, J.R. Giles, S. McDonald, S. Manne, S.F. Ngiow, K.P. Patel, M.T. Werner, A.C. Huang, K.A. Alexander, J.E. Wu, J. Attanasio, P. Yan, S.M. George, B. Bengsch, R.P. Staupe, G. Donahue, W. Xu, R.K. Amaravadi, X. Xu, G.C. Karakousis, T.C. Mitchell, L.M. Schuchter, J. Kaye, S.L. Berger, E.J. Wherry, TOX transcriptionally and epigenetically programs CD8(+) T cell exhaustion, Nature 571(7764) (2019) 211-218.

[284] A.C. Scott, F. Dündar, P. Zumbo, S.S. Chandran, C.A. Klebanoff, M. Shakiba, P. Trivedi, L. Menocal, H. Appleby, S. Camara, D. Zamarin, T. Walther, A. Snyder, M.R. Femia, E.A. Comen, H.Y. Wen, M.D. Hellmann, N. Anandasabapathy, Y. Liu, N.K. Altorki, P. Lauer, O. Levy, M.S. Glickman, J. Kaye, D. Betel, M. Philip, A. Schietinger, TOX is a critical regulator of tumour-specific T cell differentiation, Nature 571(7764) (2019) 270-274.

[285] H. Seo, J. Chen, E. González-Avalos, D. Samaniego-Castruita, A. Das, Y.H. Wang, I.F. López-Moyado, R.O. Georges, W. Zhang, A. Onodera, C.J. Wu, L.F. Lu, P.G. Hogan, A. Bhandoola, A. Rao, TOX and TOX2 transcription factors cooperate with NR4A transcription factors to impose CD8(+) T cell exhaustion, Proc Natl Acad Sci U S A 116(25) (2019) 12410-12415.

[286] W. Xu, X. Zhao, X. Wang, H. Feng, M. Gou, W. Jin, X. Wang, X. Liu, C. Dong, The Transcription Factor Tox2 Drives T Follicular Helper Cell Development via Regulating Chromatin Accessibility, Immunity 51(5) (2019) 826-839.e5.

[287] A.P. Meli, G. Fontés, D.T. Avery, S.A. Leddon, M. Tam, M. Elliot, A. Ballesteros-Tato, J. Miller, M.M. Stevenson, D.J. Fowell, S.G. Tangye, I.L. King, The Integrin LFA-1 Controls T Follicular Helper Cell Generation and Maintenance, Immunity 45(4) (2016) 831-846.

[288] C.M. Gonçalves, S.N. Henriques, R.F. Santos, A.M. Carmo, CD6, a Rheostat-Type Signalosome That Tunes T Cell Activation, Front Immunol 9 (2018) 2994.

[289] U. Lorenz, SHP-1 and SHP-2 in T cells: two phosphatases functioning at many levels, Immunol Rev 228(1) (2009) 342-59.

[290] O. Lecomte, J.B. Bock, B.W. Birren, D. Vollrath, J.R. Parnes, Molecular linkage of the mouse CD5 and CD6 genes, Immunogenetics 44(5) (1996) 385-90.

[291] O. Padilla, J. Calvo, J.M. Vilà, M. Arman, I. Gimferrer, L. Places, M.T. Arias, M.A. Pujana, J. Vives, F. Lozano, Genomic organization of the human CD5 gene, Immunogenetics 51(12) (2000) 993-1001.

[292] B. Martin, A. Banz, B. Bienvenu, C. Cordier, N. Dautigny, C. Bécourt, B. Lucas, Suppression of CD4+ T lymphocyte effector functions by CD4+CD25+ cells in vivo, J Immunol 172(6) (2004) 3391-8.

[293] F. Powrie, M.W. Leach, S. Mauze, L.B. Caddle, R.L. Coffman, Phenotypically distinct subsets of CD4+ T cells induce or protect from chronic intestinal inflammation in C. B-17 scid mice, Int Immunol 5(11) (1993) 1461-71.

[294] R.W. Nelson, D. Beisang, N.J. Tubo, T. Dileepan, D.L. Wiesner, K. Nielsen, M. Wuthrich, B.S. Klein, D.I. Kotov, J.A. Spanier, B.T. Fife, J.J. Moon, M.K. Jenkins, T cell receptor cross-reactivity between similar foreign and self peptides influences naive cell population size and autoimmunity, Immunity 42(1) (2015) 95-107.

[295] A.P. Meli, Y. Wang, D.A. de Kouchkovsky, Y. Kong, M.K. Basu, S. Ghosh, C.V. Rothlin, IL-4-induced hysteresis in naïve T cell activation, bioRxiv (2020) 2020.08.31.275842.

[296] B.D. Rudd, Neonatal T Cells: A Reinterpretation, Annu Rev Immunol 38 (2020) 229-247.

[297] J.P. Cabaniols, N. Fazilleau, A. Casrouge, P. Kourilsky, J.M. Kanellopoulos, Most alpha/beta T cell receptor diversity is due to terminal deoxynucleotidyl transferase, J Exp Med 194(9) (2001) 1385-90.

[298] C.L. Benedict, S. Gilfillan, T.H. Thai, J.F. Kearney, Terminal deoxynucleotidyl transferase and repertoire development, Immunol Rev 175 (2000) 150-7.

[299] J. Borst, T. Ahrends, N. Bąbała, C.J.M. Melief, W. Kastenmüller, CD4(+) T cell help in cancer immunology and immunotherapy, Nat Rev Immunol 18(10) (2018) 635-647.

[300] E. Alspach, D.M. Lussier, A.P. Miceli, I. Kizhvatov, M. DuPage, A.M. Luoma, W. Meng, C.F. Lichti, E. Esaulova, A.N. Vomund, D. Runci, J.P. Ward, M.M. Gubin, R.F.V. Medrano, C.D. Arthur, J.M. White, K.C.F. Sheehan, A. Chen, K.W. Wucherpfennig, T. Jacks, E.R. Unanue, M.N. Artyomov, R.D. Schreiber, MHC-II neoantigens shape tumour immunity and response to immunotherapy, Nature 574(7780) (2019) 696-701.

[301] M. Oukka, Interplay between pathogenic Th17 and regulatory T cells, Ann Rheum Dis 66 Suppl 3(Suppl 3) (2007) iii87-90.

[302] A. Oxenius, M.F. Bachmann, R.M. Zinkernagel, H. Hengartner, Virus-specific MHC-class II-restricted TCR-transgenic mice: effects on humoral and cellular immune responses after viral infection, Eur J Immunol 28(1) (1998) 390-400.

[303] P. Mombaerts, A.R. Clarke, M.A. Rudnicki, J. Iacomini, S. Itohara, J.J. Lafaille, L. Wang, Y. Ichikawa, R. Jaenisch, M.L. Hooper, et al., Mutations in T-cell antigen receptor genes alpha and beta block thymocyte development at different stages, Nature 360(6401) (1992) 225-31.

[304] L. Madsen, N. Labrecque, J. Engberg, A. Dierich, A. Svejgaard, C. Benoist, D. Mathis, L. Fugger, Mice lacking all conventional MHC class II genes, Proc Natl Acad Sci U S A 96(18) (1999) 10338-43.

[305] W. Yu, H. Nagaoka, M. Jankovic, Z. Misulovin, H. Suh, A. Rolink, F. Melchers, E. Meffre, M.C. Nussenzweig, Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization, Nature 400(6745) (1999) 682-7.

[306] E.J. Wherry, J.N. Blattman, K. Murali-Krishna, R. van der Most, R. Ahmed, Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment, J Virol 77(8) (2003) 4911-27.

[307] D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S.L. Salzberg, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions, Genome Biol 14(4) (2013) R36.

[308] Y. Liao, G.K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features, Bioinformatics 30(7) (2014) 923-30.

[309] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics 26(1) (2010) 139-40.

[310] B. Artegiani, A. Lyubimova, M. Muraro, J.H. van Es, A. van Oudenaarden, H. Clevers, A Single-Cell RNA Sequencing Study Reveals Cellular and Molecular Dynamics of the Hippocampal Neurogenic Niche, Cell Rep 21(11) (2017) 3271-3284.

[311] M.J. Muraro, G. Dharmadhikari, D. Grün, N. Groen, T. Dielen, E. Jansen, L. van Gurp, M.A. Engelse, F. Carlotti, E.J. de Koning, A. van Oudenaarden, A Single-Cell Transcriptome Atlas of the Human Pancreas, Cell Syst 3(4) (2016) 385-394.e3.

[312] T. Hashimshony, N. Senderovich, G. Avital, A. Klochendler, Y. de Leeuw, L. Anavy, D. Gennert, S. Li, K.J. Livak, O. Rozenblatt-Rosen, Y. Dor, A. Regev, I. Yanai, CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq, Genome Biol 17 (2016) 77.

[313] H. Li, R. Durbin, Fast and accurate long-read alignment with Burrows-Wheeler transform, Bioinformatics 26(5) (2010) 589-95.

[314] D. Grün, L. Kester, A. van Oudenaarden, Validation of noise models for single-cell transcriptomics, Nat Methods 11(6) (2014) 637-40.

[315] D.J. McCarthy, K.R. Campbell, A.T. Lun, Q.F. Wills, Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R, Bioinformatics 33(8) (2017) 1179-1186.

[316] L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, arXiv:1802.03426, 2020.

[317] J. Melville, uwot: The Uniform Manifold Approximation and Projection (UMAP) Method for Dimensionality Reduction. R package version 0.1.3., (2019).

[318] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30(15) (2014) 2114-20.

[319] B. Langmead, C. Trapnell, M. Pop, S.L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, Genome Biol 10(3) (2009) R25.

[320] Y. Zhang, T. Liu, C.A. Meyer, J. Eeckhoute, D.S. Johnson, B.E. Bernstein, C. Nusbaum, R.M. Myers, M. Brown, W. Li, X.S. Liu, Model-based analysis of ChIP-Seq (MACS), Genome Biol 9(9) (2008) R137.

[321] M. Salmon-Divon, H. Dvinge, K. Tammoja, P. Bertone, PeakAnalyzer: genome-wide annotation of chromatin binding and modification loci, BMC Bioinformatics 11 (2010) 415.

[322] S. Heinz, C. Benner, N. Spann, E. Bertolino, Y.C. Lin, P. Laslo, J.X. Cheng, C. Murre, H. Singh, C.K. Glass, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities, Mol Cell 38(4) (2010) 576-89.

[323] H. Thorvaldsdóttir, J.T. Robinson, J.P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration, Brief Bioinform 14(2) (2013) 178-92.

[324] B. Bolstad, preprocessCore: A collection of pre-processing functions, R package version 1.50.0 (2020).

[325] K. Raivo, pheatmap: Pretty Heatmaps. R package version 1.0.12, 2018.

[326] E.A. Howe, R. Sinha, D. Schlauch, J. Quackenbush, RNA-Seq analysis in MeV, Bioinformatics 27(22) (2011) 3209-10.

[327] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc Natl Acad Sci U S A 102(43) (2005) 15545-50.

[328] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J.P. Mesirov, P. Tamayo, The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell Syst 1(6) (2015) 417-425.

[329] G. Bindea, B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W.H. Fridman, F. Pagès, Z. Trajanoski, J. Galon, ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks, Bioinformatics 25(8) (2009) 1091-3.

[330] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res 13(11) (2003) 2498-504.

[331] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, (2016).

[332] T.E. Chan, M.P.H. Stumpf, A.C. Babtie, Gene Regulatory Network Inference from Single-Cell Data Using Multivariate Information Measures, Cell Syst 5(3) (2017) 251-267.e3.

[333] J.W. Cotari, G. Voisinne, O.E. Dar, V. Karabacak, G. Altan-Bonnet, Cell-to-cell variability analysis dissects the plasticity of signaling of common γ chain cytokines in T cells, Sci Signal 6(266) (2013) ra17.

[334] M.D. Cooper, M.N. Alder, The evolution of adaptive immune systems, Cell 124(4) (2006) 815-22.

[335] M.B. Yassai, Y.N. Naumov, E.N. Naumova, J. Gorski, A clonotype nomenclature for T cell receptors, Immunogenetics 61(7) (2009) 493-502.

[336] S. Gilfillan, M. Bachmann, S. Trembleau, L. Adorini, U. Kalinke, R. Zinkernagel, C. Benoist, D. Mathis, Efficient immune responses in mice lacking N-region diversity, Eur J Immunol 25(11) (1995) 3115-22.

[337] G.W. Litman, J.P. Rast, S.D. Fugmann, The origins of vertebrate adaptive immunity, Nat Rev Immunol 10(8) (2010) 543-53.

[338] S. Gilfillan, A. Dierich, M. Lemeur, C. Benoist, D. Mathis, Mice lacking TdT: mature animals with an immature lymphocyte repertoire, Science 261(5125) (1993) 1175-8.

[339] T. Komori, A. Okada, V. Stewart, F.W. Alt, Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes, Science 261(5125) (1993) 1171-5.

[340] E.M. Kolawole, T.J. Lamb, B.D. Evavold, Relationship of 2D Affinity to T Cell Functional Outcomes, Int J Mol Sci 21(21) (2020).

[341] J.P. Snook, C. Kim, M.A. Williams, TCR signal strength controls the differentiation of CD4(+) effector and memory T cells, Sci Immunol 3(25) (2018).

[342] N. van Panhuys, TCR Signal Strength Alters T-DC Activation and Interaction Times and Directs the Outcome of Differentiation, Front Immunol 7 (2016) 6.

[343] J. Huang, V.I. Zarnitsyna, B. Liu, L.J. Edwards, N. Jiang, B.D. Evavold, C. Zhu, The kinetics of two-dimensional TCR and pMHC interactions determine T-cell responsiveness, Nature 464(7290) (2010) 932-6.

[344] S.D. Gunasinghe, N.G. Peres, J. Goyette, K. Gaus, Biomechanics of T Cell Dysfunctions in Chronic Diseases, Front Immunol 12 (2021) 600829.

[345] A. Lee, E. Hsu, Isolation and characterization of the Xenopus terminal deoxynucleotidyl transferase, J Immunol 152(9) (1994) 4500-7.

[346] J.D. Hansen, Characterization of rainbow trout terminal deoxynucleotidyl transferase structure and expression. TdT and RAG1 co-expression define the trout primary lymphoid tissues, Immunogenetics 46(5) (1997) 367-75.

[347] S.M. Haeryfar, H.D. Hickman, K.R. Irvine, D.C. Tscharke, J.R. Bennink, J.W. Yewdell, Terminal deoxynucleotidyl transferase establishes and broadens antiviral CD8+ T cell immunodominance hierarchies, J Immunol 181(1) (2008) 649-59.

[348] M.A. Alexander-Miller, G.R. Leggatt, A. Sarin, J.A. Berzofsky, Role of antigen, CD8, and cytotoxic T lymphocyte (CTL) avidity in high dose antigen induction of apoptosis of effector CTL, J Exp Med 184(2) (1996) 485-92.

[349] M. Shakiba, P. Zumbo, G. Espinosa-Carrasco, L. Menocal, F. Dündar, S.E. Carson, E.M. Bruno, F.J. Sanchez-Rivera, S.W. Lowe, S. Camara, R.P. Koche, V.P. Reuter, N.D. Socci, B. Whitlock, F. Tamzalit, M. Huse, M.D. Hellmann, D.K. Wells, N.A. Defranoux, D. Betel, M. Philip, A. Schietinger, TCR signal strength defines distinct mechanisms of T cell dysfunction and cancer evasion, J Exp Med 219(2) (2022).

[350] M. Cherrier, M.F. D'Andon, F. Rougeon, N. Doyen, Identification of a new cis-regulatory element of the terminal deoxynucleotidyl transferase gene in the 5' region of the murine locus, Mol Immunol 45(4) (2008) 1009-17.

[351] M.S. Krangel, Mechanics of T cell receptor gene rearrangement, Curr Opin Immunol 21(2) (2009) 133-9.

[352] C. Schneider, C. Shen, A.A. Gopal, T. Douglas, B. Forestell, K.D. Kauffman, D. Rogers, P. Artusa, Q. Zhang, H. Jing, A.F. Freeman, D.L. Barber, I.L. King, M. Saleh, P.W. Wiseman, H.C. Su, J.N. Mandl, Migration-induced cell shattering due to DOCK8 deficiency causes a type 2-biased helper T cell response, Nat Immunol (2020).

[353] K. Schober, F. Voit, S. Grassmann, T.R. Müller, J. Eggert, S. Jarosch, B. Weißbrich, P. Hoffmann, L. Borkner, E. Nio, L. Fanchi, C.R. Clouser, A. Radhakrishnan, L. Mihatsch, P. Lückemeier, J. Leube, G. Dössinger, L. Klein, M. Neuenhahn, J.D. Oduro, L. Cicin-Sain, V.R. Buchholz, D.H. Busch, Reverse TCR repertoire evolution toward dominant low-affinity clones during chronic CMV infection, Nat Immunol 21(4) (2020) 434-441.

[354] K.J. Jackson, M.J. Kidd, Y. Wang, A.M. Collins, The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor, Front Immunol 4 (2013) 263.

[355] S. Gilfillan, C. Benoist, D. Mathis, Mice lacking terminal deoxynucleotidyl transferase: adult mice with a fetal antigen receptor repertoire, Immunol Rev 148 (1995) 201-19.

[356] M.S. Abdel-Hakeem, S. Manne, J.C. Beltra, E. Stelekati, Z. Chen, K. Nzingha, M.A. Ali, J.L. Johnson, J.R. Giles, D. Mathew, A.R. Greenplate, G. Vahedi, E.J. Wherry, Epigenetic scarring of exhausted T cells hinders memory differentiation upon eliminating chronic antigenic stimulation, Nat Immunol 22(8) (2021) 1008-1019.

[357] J.M. Jubel, Z.R. Barbati, C. Burger, D.C. Wirtz, F.A. Schildberg, The Role of PD-1 in Acute and Chronic Infection, Front Immunol 11 (2020) 487.

[358] A. Bonati, P. Zanelli, M. Savi, T.M. Neri, TCR-beta chain gene rearrangement and expression in human T-cell development and in leukemia, Leukemia 8(6) (1994) 918-23.

[359] M. Bogue, S. Gilfillan, C. Benoist, D. Mathis, Regulation of N-region diversity in antigen receptors through thymocyte differentiation and thymus ontogeny, Proc Natl Acad Sci U S A 89(22) (1992) 11011-5.

[360] M. Yassai, K. Ammon, J. Goverman, P. Marrack, Y. Naumov, J. Gorski, A molecular marker for thymocyte-positive selection: selection of CD4 single-positive thymocytes with shorter TCRB CDR3 during T cell development, J Immunol 168(8) (2002) 3801-7.

[361] M. Yassai, J. Gorski, Thymocyte maturation: selection for in-frame TCR alpha-chain rearrangement is followed by selection for shorter TCR beta-chain complementarity-determining region 3, J Immunol 165(7) (2000) 3706-12.

[362] H. Gu, Y.R. Zou, K. Rajewsky, Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting, Cell 73(6) (1993) 1155-64.

[363] M.J. Richer, J.C. Nolz, J.T. Harty, Pathogen-specific inflammatory milieux tune the antigen sensitivity of CD8(+) T cells by enhancing T cell receptor signaling, Immunity 38(1) (2013) 140-52.

[364] R. Ahmed, A. Salmi, L.D. Butler, J.M. Chiller, M.B. Oldstone, Selection of genetic variants of lymphocytic choriomeningitis virus in spleens of persistently infected mice. Role in suppression of cytotoxic T lymphocyte response and viral persistence, J Exp Med 160(2) (1984) 521-40.

[365] K.G. Anderson, K. Mayer-Barber, H. Sung, L. Beura, B.R. James, J.J. Taylor, L. Qunaj, T.S. Griffith, V. Vezys, D.L. Barber, D. Masopust, Intravascular staining for discrimination of vascular and tissue leukocytes, Nat Protoc 9(1) (2014) 209-22.

[366] C.D. Surh, J. Sprent, Homeostasis of naive and memory T cells, Immunity 29(6) (2008) 848-62.

[367] H. Kumar, T. Kawai, S. Akira, Pathogen recognition by the innate immune system, Int Rev Immunol 30(1) (2011) 16-34.

[368] F. McNab, K. Mayer-Barber, A. Sher, A. Wack, A. O'Garra, Type I interferons in infectious disease, Nat Rev Immunol 15(2) (2015) 87-103.

[369] J.R. Teijaro, C. Ng, A.M. Lee, B.M. Sullivan, K.C. Sheehan, M. Welch, R.D. Schreiber, J.C. de la Torre, M.B. Oldstone, Persistent LCMV infection is controlled by blockade of type I interferon signaling, Science 340(6129) (2013) 207-11.

[370] A.J. Lee, A.A. Ashkar, The Dual Nature of Type I and Type II Interferons, Front Immunol 9 (2018) 2061.

[371] E.C. Shin, U. Seifert, T. Kato, C.M. Rice, S.M. Feinstone, P.M. Kloetzel, B. Rehermann, Virus-induced type I IFN stimulates generation of immunoproteasomes at the site of infection, J Clin Invest 116(11) (2006) 3006-14.

[372] J.M. Blander, R. Medzhitov, Toll-dependent selection of microbial antigens for presentation by dendritic cells, Nature 440(7085) (2006) 808-12.

[373] P. Nair-Gupta, A. Baccarini, N. Tung, F. Seyffer, O. Florey, Y. Huang, M. Banerjee, M. Overholtzer, P.A. Roche, R. Tampé, B.D. Brown, D. Amsen, S.W. Whiteheart, J.M. Blander, TLR signals induce phagosomal MHC-I delivery from the endosomal recycling compartment to allow cross-presentation, Cell 158(3) (2014) 506-21.

[374] M. Kuka, M. De Giovanni, M. Iannacone, The role of type I interferons in CD4(+) T cell differentiation, Immunol Lett 215 (2019) 19-23.

[375] R.M. Welsh, K. Bahl, H.D. Marshall, S.L. Urban, Type 1 interferons and antiviral CD8 T-cell responses, PLoS Pathog 8(1) (2012) e1002352.

[376] J.C. Rathmell, E.A. Farkash, W. Gao, C.B. Thompson, IL-7 enhances the survival and maintains the size of naive T cells, J Immunol 167(12) (2001) 6869-76.

[377] A.L. Horenstein, A. Chillemi, G. Zaccarello, S. Bruzzone, V. Quarona, A. Zito, S. Serra, F. Malavasi, A CD38/CD203a/CD73 ectoenzymatic pathway independent of CD39 drives a novel adenosinergic loop in human T lymphocytes, Oncoimmunology 2(9) (2013) e26246.

[378] J.M. Cantor, M.H. Ginsberg, CD98 at the crossroads of adaptive immunity and cancer, J Cell Sci 125(Pt 6) (2012) 1373-82.

[379] G. Holloway, F.E. Fleming, B.S. Coulson, MHC class I expression in intestinal cells is reduced by rotavirus infection and increased in bystander cells lacking rotavirus antigen, Scientific Reports 8(1) (2018) 67.

[380] C.V. Harding, W.H. Boom, Regulation of antigen presentation by Mycobacterium tuberculosis: a role for Toll-like receptors, Nature Reviews Microbiology 8(4) (2010) 296-307.

[381] M. Koutsakos, H.E.G. McWilliam, T.E. Aktepe, S. Fritzlar, P.T. Illing, N.A. Mifsud, A.W. Purcell, S. Rockman, P.C. Reading, J.P. Vivian, J. Rossjohn, A.G. Brooks, J.M. Mackenzie, J.D. Mintern, J.A. Villadangos, T.H.O. Nguyen, K. Kedzierska, Downregulation of MHC Class I Expression by Influenza A and B Viruses, Front Immunol 10 (2019) 1158.

[382] P. Devi, A. Khan, P. Chattopadhyay, P. Mehta, S. Sahni, S. Sharma, R. Pandey, Co-infections as Modulators of Disease Outcome: Minor Players or Major Players?, Front Microbiol 12 (2021) 664386.

[383] J.R. Prigge, T.R. Hoyt, E. Dobrinen, M.R. Capecchi, E.E. Schmidt, N. Meissner, Type I IFNs Act upon Hematopoietic Progenitors To Protect and Maintain Hematopoiesis during Pneumocystis Lung Infection in Mice, J Immunol 195(11) (2015) 5347-57.

[384] E.R. Bovolenta, E.M. García-Cuesta, L. Horndler, J. Ponomarenko, W.W. Schamel, M. Mellado, M. Castro, D. Abia, H.M. van Santen, A set point in the selection of the αβTCR T cell repertoire imposed by pre-TCR signaling strength, Proc Natl Acad Sci U S A 119(22) (2022) e2201907119.

[385] A.S. Chann, M. Charnley, L.M. Newton, A. Newbold, F. Wiede, T. Tiganis, P.O. Humbert, R.W. Johnstone, S.M. Russell, Stepwise progression of β-selection during T cell development involves histone deacetylation, Life Sci Alliance 6(1) (2023).

[386] H.J. Melichar, J.O. Ross, P. Herzmark, K.A. Hogquist, E.A. Robey, Distinct temporal patterns of T cell receptor signaling during positive versus negative selection in situ, Sci Signal 6(297) (2013) ra92.

[387] B.B. Au-Yeung, H.J. Melichar, J.O. Ross, D.A. Cheng, J. Zikherman, K.M. Shokat, E.A. Robey, A. Weiss, Quantitative and temporal requirements revealed for Zap70 catalytic activity during T cell development, Nat Immunol 15(7) (2014) 687-94.

[388] J.O. Ross, H.J. Melichar, B.B. Au-Yeung, P. Herzmark, A. Weiss, E.A. Robey, Distinct phases in the positive selection of CD8+ T cells distinguished by intrathymic migration and T-cell receptor signaling patterns, Proc Natl Acad Sci U S A 111(25) (2014) E2550-8.

[389] U.S. Khanom, I. Ohigashi, S. Fujimori, K. Kondo, K. Takada, Y. Takahama, TCR Affinity for In Vivo Peptide-Induced Thymic Positive Selection Fine-Tunes TCR Responsiveness of Peripheral CD8(+) T Cells, J Immunol 203(4) (2019) 881-887.

[390] S. Herppich, A. Toker, B. Pietzsch, Y. Kitagawa, N. Ohkura, T. Miyao, S. Floess, S. Hori, S. Sakaguchi, J. Huehn, Dynamic Imprinting of the Treg Cell-Specific Epigenetic Signature in Developing Thymic Regulatory T Cells, Front Immunol 10 (2019) 2382.

[391] A.G. Levine, A. Arvey, W. Jin, A.Y. Rudensky, Continuous requirement for the TCR in regulatory T cell function, Nat Immunol 15(11) (2014) 1070-8.

[392] J.C. Vahl, C. Drees, K. Heger, S. Heink, J.C. Fischer, J. Nedjic, N. Ohkura, H. Morikawa, H. Poeck, S. Schallenberg, D. Rieß, M.Y. Hein, T. Buch, B. Polic, A. Schönle, R. Zeiser, A. Schmitt-Gräff, K. Kretschmer, L. Klein, T. Korn, S. Sakaguchi, M. Schmidt-Supprian, Continuous T cell receptor signals maintain a functional regulatory T cell pool, Immunity 41(5) (2014) 722-36.

[393] I. Messaoudi, J.A. Guevara Patiño, R. Dyall, J. LeMaoult, J. Nikolich-Zugich, Direct link between mhc polymorphism, T cell avidity, and diversity in immune defense, Science 298(5599) (2002) 1797-800.

[394] D. Meyer-Olson, N.H. Shoukry, K.W. Brady, H. Kim, D.P. Olson, K. Hartman, A.K. Shintani, C.M. Walker, S.A. Kalams, Limited T cell receptor diversity of HCV-specific T cell responses is associated with CTL escape, J Exp Med 200(3) (2004) 307-19.

[395] K.P. van Gisbergen, P.L. Klarenbeek, N.A. Kragten, P.P. Unger, M.B. Nieuwenhuis, F.M. Wensveen, A. ten Brinke, P.P. Tak, E. Eldering, M.A. Nolte, R.A. van Lier, The costimulatory molecule CD27 maintains clonally diverse CD8(+) T cell responses of low antigen affinity to protect against viral variants, Immunity 35(1) (2011) 97-108.

[396] C.U. Blank, W.N. Haining, W. Held, P.G. Hogan, A. Kallies, E. Lugli, R.C. Lynn, M. Philip, A. Rao, N.P. Restifo, A. Schietinger, T.N. Schumacher, P.L. Schwartzberg, A.H. Sharpe, D.E. Speiser, E.J. Wherry, B.A. Youngblood, D. Zehn, Defining 'T cell exhaustion', Nature Reviews Immunology 19(11) (2019) 665-674.

[397] C.M. Bucks, J.A. Norton, A.C. Boesteanu, Y.M. Mueller, P.D. Katsikis, Chronic antigen stimulation alone is sufficient to drive CD8+ T cell exhaustion, J Immunol 182(11) (2009) 6697-708.

[398] J.C. Beltra, S. Manne, M.S. Abdel-Hakeem, M. Kurachi, J.R. Giles, Z. Chen, V. Casella, S.F. Ngiow, O. Khan, Y.J. Huang, P. Yan, K. Nzingha, W. Xu, R.K. Amaravadi, X. Xu, G.C. Karakousis, T.C. Mitchell, L.M. Schuchter, A.C. Huang, E.J. Wherry, Developmental Relationships of Four Exhausted CD8(+) T Cell Subsets Reveals Underlying Transcriptional and Epigenetic Landscape Control Mechanisms, Immunity 52(5) (2020) 825-841.e8.

[399] M. Philip, L. Fairchild, L. Sun, E.L. Horste, S. Camara, M. Shakiba, A.C. Scott, A. Viale, P. Lauer, T. Merghoub, M.D. Hellmann, J.D. Wolchok, C.S. Leslie, A. Schietinger, Chromatin states define tumour-specific T cell dysfunction and reprogramming, Nature 545(7655) (2017) 452-456.

[400] D. Furman, J. Campisi, E. Verdin, P. Carrera-Bastos, S. Targ, C. Franceschi, L. Ferrucci, D.W. Gilroy, A. Fasano, G.W. Miller, A.H. Miller, A. Mantovani, C.M. Weyand, N. Barzilai, J.J. Goronzy, T.A. Rando, R.B. Effros, A. Lucia, N. Kleinstreuer, G.M. Slavich, Chronic inflammation in the etiology of disease across the life span, Nature Medicine 25(12) (2019) 1822-1832.

[401] M.S. Coleman, J.J. Hutton, P. De Simone, F.J. Bollum, Terminal deoxyribonucleotidyl transferase in human leukemia, Proc Natl Acad Sci U S A 71(11) (1974) 4404-8.

[402] F. Campagnari, E. Bombardieri, F. de Braud, L. Baldini, A.T. Maiolo, Terminal deoxynucleotidyl transferase, TdT, as a marker for leukemia and lymphoma cells, Int J Biol Markers 2(1) (1987) 31-42.

[403] D.J. Gough, N.L. Messina, C.J. Clarke, R.W. Johnstone, D.E. Levy, Constitutive type I interferon modulates homeostatic balance through tonic signaling, Immunity 36(2) (2012) 166-74.