

**Mechanisms of tissue compartmentalization
in human T cells**

Michelle Miron

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

©2019

Michelle Miron

All rights reserved

ABSTRACT

Mechanisms of tissue compartmentalization in human T cells

Michelle Miron

Mechanisms for human memory T cell differentiation and maintenance have predominantly been inferred from studies of peripheral blood, though the majority of T cells reside in lymphoid and non-lymphoid sites. Studies in mice have shown that memory T cells in non-lymphoid sites provide superior protection to pathogens compared to those in blood, defining a subset known as tissue-resident memory T cells (TRM), with emerging roles in lymphoid sites. There are many key unknown aspects of TRM biology in human tissues including if TRM have superior functional abilities, the mechanisms for maintenance of TRM in lymphoid and non-lymphoid sites, and the relatedness of tissue and blood localized T cell subsets.

Through a collaboration with the local organ procurement agency, we obtained samples from >15 tissue sites from healthy organ donors of all ages. We analyzed CD8⁺ T cells in diverse sites and found the majority of TRM cells in lymph nodes (LNs) display an increased proliferative capacity, increased expression of TCF-1, and decreased turnover compared to TRM and effector memory (TEM) cells in other sites including blood, bone marrow (BM), spleen and lung. Further, we identified that exposure to type 1 interferons results in increased downregulation of TCF-1 expression during cell divisions driven by T cell receptor (TCR) stimulation. We investigated the relatedness of CD4⁺ and CD8⁺ T cell subsets, including central memory (TCM), effector memory (TEM), TRM, and terminal effectors (TEMRA) by sequencing TCR rearrangements. From diversity analysis of TCR repertoires we found that effector and memory subsets are maintained in a hierarchy from most to least diverse (TCM > TEM and

TRM > TEMRA) that is largely conserved across tissues and CD4⁺ and CD8⁺ T cell lineages. Overlap analysis revealed the low and high relatedness of TCM and TEMRA cells respectively and this was highly conserved across tissues; in contrast, we found the relatedness of TEM and TRM was more dynamic across tissues. Together, these findings have implications for immune monitoring and modulation, highlighting that lymph nodes may function as reservoirs for long-lived memory T cells with high functional capacity; additionally, we identify cell extrinsic signals that regulate tissue-specific maintenance of T cell memory in lymph node sites.

TABLE OF CONTENTS

LIST OF FIGURES	iii
LIST OF TABLES	v
LIST OF ABBREVIATIONS	vi
ACKNOWLEDGEMENTS	vii
DEDICATION	ix
CHAPTER 1: INTRODUCTION	1
SECTION 1.1 OVERVIEW OF HUMAN T CELL RESPONSES	1
SECTION 1.2 DEVELOPMENT OF HUMAN T CELLS.....	8
SECTION 1.3 T CELL ACTIVATION AND DIFFERENTIATION	17
SECTION 1.4 MAINTENANCE AND FORMATION OF T CELL MEMORY.....	25
SECTION 1.5 TISSUE COMPARTMENTALIZATION OF MEMORY T CELLS	41
SECTION 1.6 THESIS OBJECTIVES	57
CHAPTER 2: MATERIALS AND METHODS	60
SECTION 2.1 HUMAN TISSUE ACQUISITION AND LYMPHOCYTE ISOLATION	60
SECTION 2.2 FLOW CYTOMETRY	65
SECTION 2.3 CYTOF	66
SECTION 2.4 WHOLE TRANSCRIPTOME PROFILING BY RNA SEQUENCING.....	70
SECTION 2.5 T CELL PROLIFERATION ASSAYS.....	73
SECTION 2.6 TCR SEQUENCING.....	75
SECTION 2.7 STATISTICAL TESTS	82
CHAPTER 3: HUMAN LYMPH NODES MAINTAIN TCF-1⁺ T CELLS WITH HIGH FUNCTIONAL POTENTIAL AND CLONAL DIVERSITY	83
SECTION 3.1 INTRODUCTION	85
SECTION 3.2 RESULTS	87
SECTION 3.3 DISCUSSION	128
CHAPTER 4: SUBSET-SPECIFIC COMPARTMENTALIZATION OF HUMAN T CELL RECEPTOR REPERTOIRES ACROSS BLOOD AND TISSUE SITES	133
SECTION 4.1 INTRODUCTION	135
SECTION 4.2 RESULTS	137
SECTION 4.3 DISCUSSION	165
CHAPTER 5: CONCLUSIONS	170

REFERENCES	178
APPENDICES	196
APPENDIX A. CMV-SPECIFIC T CELL RESPONSES IN TISSUES.	196
APPENDIX B. TRANSCRIPTIONAL PROFILING OF CD4 ⁺ AND CD8 ⁺ TRM IN BONE MARROW AND LYMPH NODES	202
APPENDIX C. SCRIPTS IN PYTHON AND R	222
APPENDIX D. ACCEPTED ABSTRACTS	249
APPENDIX E. ABSTRACTS OF CONTRIBUTING AUTHOR MANUSCRIPTS	253
APPENDIX F. CURRICULUM VITAE.....	262

LIST OF FIGURES

Figure 1-1. Overview of T cell development and responses.	4
Figure 1-2. Estimated percentage of the total number of T cells in different human organs.	6
Figure 1-3. Gene rearrangements forming the TCR.....	14
Figure 1-4. T cell activation and differentiation.	22
Figure 1-5. Tissue localization patterns and phenotypes human T cell subsets.	28
Figure 1-6. Expression of T-bet, Lef-1 and TCF-1 in human T cells.....	37
Figure 2-1. Human tissues acquired for research.....	63
Figure 2-2. Flowchart of TCR analysis.....	79
Figure 3-1. Expression of TCF-1 is maintained in LN CD8 ⁺ TEM cells.	88
Figure 3-2. Expression of TCF-1 in CD69 ⁻ and CD69 ⁺ CD8 ⁺ TEM cells in tissues.....	90
Figure 3-3. Sorting strategy for isolation of lymphocytes for RNA-seq and TCR-seq.	93
Figure 3-4. Human LN memory CD8 ⁺ T cells are phenotypically and transcriptionally distinct from peripheral blood, BM, and Spl T cells.	94
Figure 3-5. Relatedness of CD69 ⁺ and CD69 ⁻ CD8 ⁺ TEM from tissues and blood based on differentially expressed genes between CD8 ⁺ TEM in BM and LN.	102
Figure 3-6. Human LN memory CD8 ⁺ T cells exhibit similarity to murine CXCR5 ⁺ CD8 ⁺ T cells generated in response to chronic LCMV infection.	104
Figure 3-7. Expression of individual markers by total CD8 ⁺ T cells concatenated from spleen, lung, BM, and LN of three donors in tSNE plots by CyTOF.	108
Figure 3-8. Human LN memory CD8 ⁺ T cells exhibit a protein signature distinct from blood and other tissues.....	110
Figure 3-9. LN memory CD8 ⁺ T cells exhibit a differentiation profile with features of both naïve and memory T cells.....	112

Figure 3-10. Sorting strategy for isolation of CD8 ⁺ TEM cells for proliferation assays.....	115
Figure 3-11. Human LN memory CD8 ⁺ T cells exhibit enhanced proliferation and effector phenotype in response to TCR stimulation.....	116
Figure 3-12. Higher TCR diversity of CD8 ⁺ TEM cells in LN compared to BM.	121
Figure 3-13. IFN-I opposes maintenance of TCF-1 expression in proliferating CD8 ⁺ T cells. .	125
Figure 3-14. Summary of findings on human lymph node memory T cells.....	132
Figure 4-1. Overview of samples collected and experimental design for TCR sequencing of T cell populations from human lymphoid and non-lymphoid tissues.....	138
Figure 4-2. Gating strategy for T cell subsets isolated for TCR sequencing.....	139
Figure 4-3. T cell clone abundances attributed mainly to subset differences.....	145
Figure 4-4. Comparing DNA and RNA based methods for TCR sequencing.....	149
Figure 4-5 . Tissue distribution of T cell clones largely explained by cell subset and lineage identity.	153
Figure 4-6. TCR repertoires of CD4 ⁺ and CD8 ⁺ lineages are divergent and show similarity between replicate samples.....	154
Figure 4-7. Tissue-specific and subset-specific differences of CD4 ⁺ and CD8 ⁺ T cell subsets.	155
Figure 4-8. Top clones are widely distributed across tissues and compartmentalized in T cell subsets.	159
Figure 4-9. Substantial overlap of TEM and TRM in tissues with high intra-tissue and subset overlap.....	163
Figure 4-10. Overview of subset and tissue specific features of TCR repertoires.	167
Figure 5-1. Summary of findings on T cell compartmentalization in tissues.....	177

LIST OF TABLES

Table 1-1. Properties of human T cell subsets.....	32
Table 2-1. Antibody and reagents used for flow cytometry.	68
Table 2-2. Antibodies used for CyTOF	69
Table 2-3. PCR Primers used for TCR-V β amplification.....	77
Table 3-1. Pathways differentially regulated in CD8 ⁺ TEM cells in lymph nodes compared BM, spleen and blood.	96
Table 3-2. Total genes upregulated in LN CD8 ⁺ TEM.....	97
Table 3-3. Total genes downregulated in LN CD8 ⁺ TEM.	99
Table 3-4. Number of clones identified by TCR sequencing of CD8 ⁺ TEM in BM and LN. ...	120
Table 3-5. Potential upstream regulators of LN memory CD8 ⁺ T cells	123
Table 3-6. List of organ donors used in this study.....	126
Table 4-1. Age and characteristics of donors in this study.	140
Table 4-2. Number of T cell clones identified per sample by TCR sequencing.....	141

List of Abbreviations

TEM = Effector memory T cell

TCM = Central memory T cell

TEMRA = Terminally differentiated effector T cell

TRM = Tissue-resident memory T cell

APCs = Antigen presenting cells

TCF-1 = T cell factor 1

LN = Lymph node

BM = Bone Marrow

Spl= Spleen

Bld= Blood

CyTOF = Cytometry by time of flight

tSNE = t-Distributed Stochastic Neighbor Embedding

TCR = T cell antigen receptor

ACKNOWLEDGEMENTS

I had a great experience conducting my thesis work and for that I would first like to thank my advisor, Dr. Donna Farber. Throughout my graduate work I felt the lab environment, promoted by Donna and lab members, was extremely supportive and enabled me to accomplish what I did. I learned immensely from working with everyone in our lab and it made coming to work productive and fun. Additionally, Donna provided me with an opportunity to research cutting edge topics with a large degree of independence. Donna provided key guidance whenever asked, both when projects were going well and when projects were met with challenges. Her constructive feedback led me to grow as a scientist. I am extremely grateful for the experiences I had and the skills I have learned here. I am also grateful for the future opportunities I'll have as a result of my experiences here.

I would like to acknowledge the members of my thesis committee: Dr. Ran Reshef, Dr. Steven Reiner, Dr. Yufeng Shen, and Dr. Stephen Emerson for their scientific guidance and support throughout different stages of my graduate studies. In addition, the Department of Microbiology and Immunology as well as the Columbia Center for Translational Immunology were central for my training as a scientist, and I thank everyone for their support; including funding support by the T32 training grant led by Dr. David Fidock.

When I started in the lab, I worked with Dr. Claire Gordon. I would like to thank Claire; she taught me immunological techniques essential for my thesis work. She also supported my development as a scientist which helped me lead my own projects.

I would also like to thank my excellent collaborators. Dr. Adeeb Rahman provided essential help by generating the protein expression data by CyTOF. Additionally, the T cell receptor sequencing work was done in collaboration with members of the Luning Prak lab. In

particular, Dr. Wenzhao Meng provided help with designing and performing the sequencing experiments making this study possible. Dr. Nina Luning Prak provided key scientific guidance along the way. Thank you to Dr. Uri Hershberg, Aaron Rosenfeld and Brian Ji for help with the TCR analysis. It was a pleasure to work with everyone.

This would not have been possible without the help of my many lab members. I would like to especially thank Dr. Brahma Kumar, who provided key insights, help, strategies, and encouragement. I would also like to thank other members of the lab who have helped me in so many ways including: troubleshooting experiments, practicing talks for conferences and my qualifying exam, and enjoying good food, drinks and laughs. Thank you to Dr. Tomer Granot, Dr. Kyra Zens, Rebecca Guyer, Dr. Peter Szabo, Dr. Puspa Thapa, Dr. Pranay Dogra, Dr. Stu Weisberg, Dr. Tom Connors, Daniel Paik and Filip Cvetkovski. I look forward to keeping in touch!

I am also fortunate that I had access to the human tissue resources in our laboratory, through a collaboration with LiveOnNY, allowing me to make discoveries in human immunology. I want to thank the surgeons Dr. Dustin Carpenter and Dr. Takashi Senda, who worked hard to procure tissues at all hours and made this study possible.

Thank you to my previous mentors at the National Institutes of Health in Dr. Donald Court's lab. Thank you to Don for the opportunity to learn so much during my time in your lab. Thank you to Jim, Lynn and Nina for teaching me fundamental techniques and giving me advice that served to be important for my doctorate work.

Thank you to my significant other Ben Zax for his encouragement in all my endeavors. Last but foremost, I would like to thank my family for their support and important role in all of my successes.

DEDICATION

This work is dedicated to future scientists.

CHAPTER 1: Introduction

Adapted and expanded significantly from: **Miron M.**, Thome J.J.C., Gordon C.L., Farber D.L. (2017) Study of T Cell Immunosenescence in Various Tissue Compartments, *Springer*, and Szabo, P.A., **Miron M.**, Farber D.L. Staying put: Tissue resident memory T cells in mice and human. Under review at *Science Immunology*.

Section 1.1 Overview of human T cell responses

T cells are crucial responders in the immune system that coordinate adaptive immunity. Key functions that T cells perform include protection from pathogens, prevention of autoimmunity, attack of cancer cells, and helping generate antibody responses. T cell responses can also function in undesirable ways and can be major drivers of autoimmune and inflammatory diseases. In order to perform these diverse functions, there are different subsets of T cells which each carry out specific roles in an immune response.

T cells develop in the thymus, from precursor cells that migrate from the bone marrow. During development in the thymus, T cells acquire the ability to distinguish self from non-self molecules known as antigens. After development, naive T cells collectively have the ability to respond to millions of specific pathogens. These naïve T cells exit the thymus and seed lymphoid sites such as lymph nodes and spleen, where they recognize pathogen derived antigens. Upon recognition of antigen, naive T cells become activated and expand into pools of effector T cells [1, 2]. Effector cells disseminate to tissue sites of inflammation where they perform their diverse effector functions. Once the pathogen is cleared, most of the responding T cells die by apoptosis. A small portion of the initially activated pool of T cells differentiate into memory T cells that are maintained in the original sites of pathogen invasion, poised to respond where the pathogen may

likely re-appear. Upon antigen encounter, memory T cells produce a stronger response *in situ* resulting in more rapid clearance of the pathogen compared to primary responses. Additionally, smaller portions of memory T cells are maintained in lymphoid sites, with the ability to rapidly proliferate and disseminate if the pathogen re-appears. Memory T cells are long lived and can provide protection for years and up to decades. These important features of immunological memory make T cells a major target of vaccines and other immunotherapies.

Most of our current understanding regarding the development, differentiation, maintenance and function of T cells stems from *in vivo* studies of T cell responses to pathogens such as viruses and bacteria in mouse models. The phases of T cell development and response to an acute viral infection are graphically summarized in Figure 1-1. Mouse models are a powerful tool in elucidating T cell function and have been essential in understanding the molecular underpinnings of a T cell response, which has led to development of several effective immunotherapies for human disease including design of vaccines and immunotherapies for autoimmunity and cancer [3]. The use of mouse models confines the study of T cell response to a single pathogen or in a specific disease model, whereas humans are constantly exposed to a myriad of pathogens, chronic viruses, and cancer yet remain protected lifelong. It's therefore important to also study human immune system in order to translate findings in mouse models for their relevance in human health and disease.

Immunological memory was originally found to be maintained within the blood, both in serum through long-lived antibody responses and circulating memory T cells. More recent investigations have revealed the importance of tissue localized maintenance of immunological memory, especially T cell memory. Blood is the major source of material for human immune cell studies, but contains only 2-3% of the total T cell complement, with the majority in lymphoid sites (~76%), followed by skin and mucosal sites such as the lung and gut (~20%).[4, 5] The

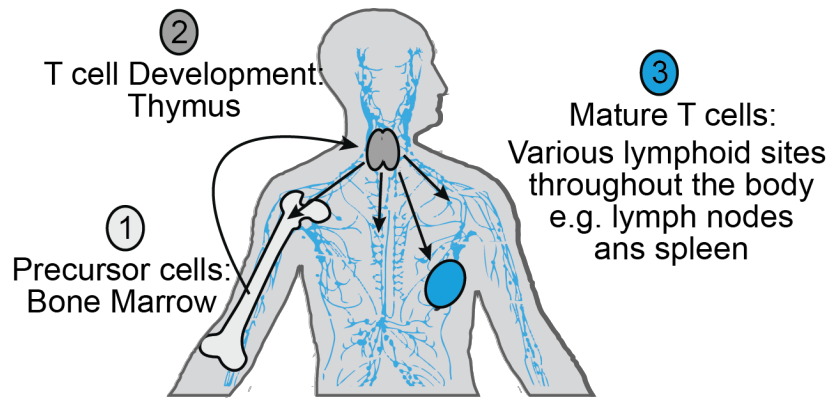
total number of T cells in the human body is approximately 330×10^9 cells. Estimates of the percentage of total T cells in different human organs is shown in Figure 1-2. How human T cells in tissue sites function and persist relative to subsets in blood is not well understood. Defining the nature of tissue localized T cell responses is important for monitoring and modulating immune responses. Investigating tissue T cell populations in humans typically involves obtaining tissue samples from either biopsy, surgical resections of diseased organs, or deceased organ donors whose individual tissues are harvested for life-saving transplantation. Our lab has extensively characterized the use of previously healthy organ donor tissue for the study of immune cells [6-12], demonstrating that this type of tissue resource effectively reveals snapshots of tissue immunity throughout all stages of life. We have made many discoveries on the human immune system using this resource including on T cell compartmentalization with age [13], response to chronic viruses [10] and as well as characterizing other immune cells [11, 14, 15]. However, there are still many open questions about how T cells are compartmentalized in human tissues and organ donor tissue serves as important resource for addressing these questions.

Figure 1-1. Overview of T cell development and responses.

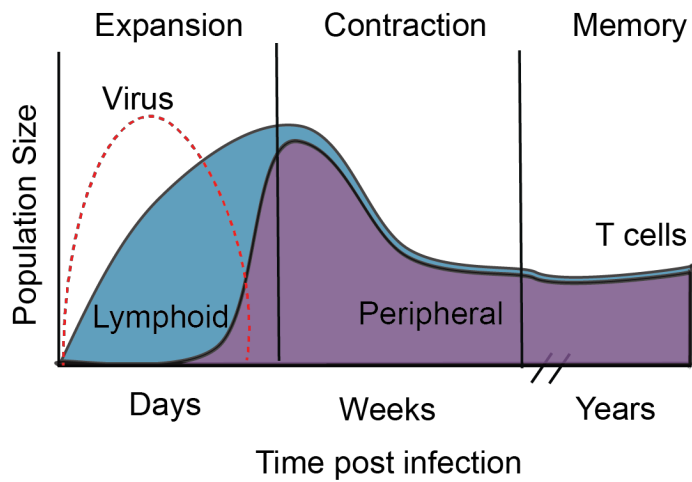
(A) T cells precursors originate in the bone marrow (white) and migrate to the thymus (blue) where they develop into mature T cells called naïve T cells. Naïve T cells are exported from the thymus and seed lymphoid sites throughout the body including lymph nodes and spleen (blue).

(B) Relative sizes and location of a T cell response to a viral infection and their respective kinetics. T cell population size is defined by number of total T cells over time. The peak of T cell expansion occurs after the peak of viral load (red), followed by contraction of the majority of T cells by cell death upon clearance of the pathogen. Most T cells die after resolution of inflammation; however, a portion of the initial responding T cells are maintained long-term as memory T cells. Localization of T cell populations during these phases are indicated by color, starting off in lymphoid tissues during activation and expansion (blue), and migration to diverse peripheral sites (purple) of inflammation, where effector function and maintenance of memory T cells occurs, with smaller portion maintained as memory in lymphoid sites. (C) The magnitude of a T cell response to secondary exposure (recall response) is larger and occurs more rapidly than the primary response. Additionally, while a primary response usually results in dispersion of effector cells throughout the body (purple), the recall response occurs more rapidly and with a greater magnitude than the primary response, due to tissue localized responding memory T cells *in situ* (yellow).

(A)



(B)



(C)

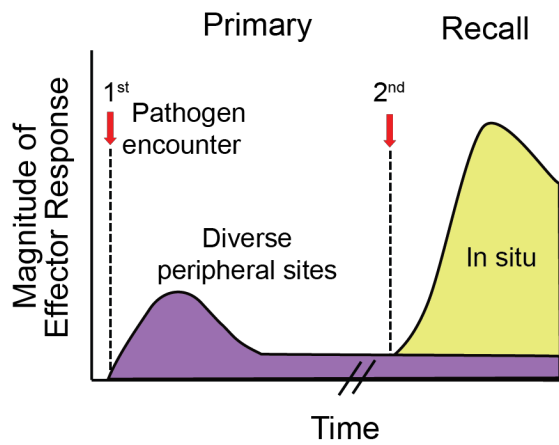
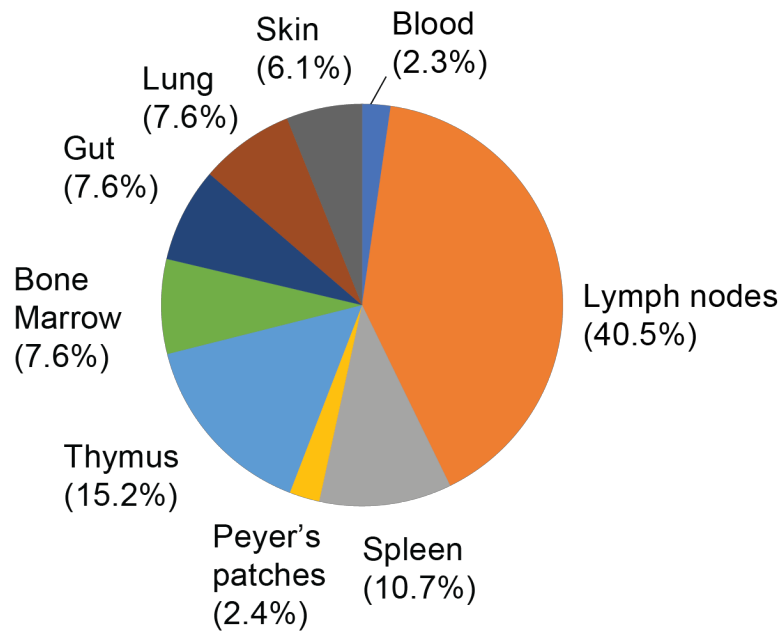


Figure 1-2. Estimated percentage of the total number of T cells in different human organs.

(A) The total number of T cells in the human body is approximately 330×10^9 cells. A pie chart of the estimated percentages of the total T cells in individual organs. It is estimated that the majority of T cells (76.4%) are in lymphoid tissues (lymph nodes, spleen, bone marrow, thymus, and Peyer's patches), followed by barrier sites (skin, lung and gut, totaling 21.3%), and the minority of T cells are found in blood (2.3%). These estimates are from three studies.[4-6]

(A)



Section 1.2 Development of Human T cells

Overview of T cell development

The primary sites for T cell development are the bone marrow and thymus; T cell precursors originate in the bone marrow and migrate to the thymus where mature T cells develop [16]. Two key processes occur in thymic development endowing T cell specificity and self-tolerance. First, rearrangement of T cell receptor (TCR) gene segments occurs to encode a mature $\alpha\beta$ TCR, with each T cell expressing a unique TCR; in humans this can comprise an estimated theoretical diversity of $\sim 5 \times 10^{11}$ distinct sequences [17] conferring different specificities [18, 19]. The second key process is for endowment of self-tolerance, in which T cells with strongly self-reactive TCRs are deleted by cell death, and only those T cells with TCR of optimal avidities to self-antigen emerge into the periphery as mature T cells [18]. The production of new T cells from the thymus is highest at birth and during infancy, and there is an established reduction in thymic function and volume beginning in puberty [20]. Upon generation of a sizable T cell repertoire early in life, this repertoire is maintained long term, mediates responses to new pathogens and provides life-long immunity [21].

Role of the thymus in T cell generation

The thymus is a unique organ in the body both in its exclusive role in generation and selection of new T cells, and its distinct developmental program for functional and structural degradation relatively early in life. The importance of the thymus in generation of immunity was first discovered in mouse models by surgical removal of the thymus (thymectomy) resulting in severe immune defects [22, 23]. Further, mice with genetic defects in the FOXP1 gene are not able to form a thymus due to defective development of the thymic epithelium [24]. This mutation also leads to abnormal hair growth (hairlessness) and therefore mice with this mutation are

referred to as *nude* mice[25]. Homozygous *nude* mice lack T cells and suffer from a lack of cell-mediated immunity. A rare case of the *nude* phenotype was also found in humans, mapped to the same gene and resulting in hairlessness and a complete lack of T cells leading to severe immunodeficiency[26]. In humans, a different genetic defect can lead the lack of a thymus and a rare syndrome known as DiGeorge syndrome. Individuals with DiGeorge syndrome can produce B cells but have a severe defect in the ability to form T cells. The genetic underpinning of DiGeorge syndrome has been mapped to a large deletion on chromosome 22 and leads to many complications and symptoms including congenital heart disease and immune deficiency[27]. Collectively these studies have shown the requirement of the thymus specifically for the production of T cells in mouse and human, and the importance of T cell development for cell-based immune responses.

A significant difference between T cell development in mice and humans is that in mice the thymus continues to develop for 3-4 weeks after birth, whereas in humans, the thymus is already fully developed at birth. As a result, mice are born lymphopenic, while humans are born with a full complement of T cells. Evidence for the timing of T cell development in humans derives from follow-up studies on infants thymectomized as a consequence of heart surgery to repair congenital abnormalities. Anatomically, the thymus is situated in front of the heart. In humans, neonatal thymectomy is performed during surgeries for congenital cardiac malformations in order to have optimal exposure of the surgical area. Longitudinal studies of patients thymectomized as neonates, show that patients can still develop naive T cells and do not experience more infections than healthy control children [28, 29]. However, long-term follow-up studies on adults who underwent thymectomy as neonates and were then in their 20s-30s did have significant decreases in T cell numbers and displayed increased declines in naïve T cell frequencies with age compared to control individuals [30, 31]. These studies show that while the

majority of T cells are produced early in life, new T cells can continue to be produced in the thymus during infancy and childhood [21]. Numerous studies have documented decreases in thymic function beginning in puberty and continuing into adulthood, with some changes seen immediately after birth [32]. It remains unknown whether later in life these individuals will be more prone to infections or cancer that are more common in the elderly due to already decreased immune function. In conclusion, the critical steps in T cell development occur before birth in humans. Additionally, once a T cell repertoire has been generated, this repertoire can be maintained for years in order to provide protection to new encounters throughout life.

T cell development in the thymus is a highly regulated process that involves interaction with MHC complexes and specialized epithelial cells within the thymus. Two different subsets of thymocytes are generated in the thymus: $CD4^+$ and $CD8^+$ T cells. T cell precursors that migrate from the bone marrow and develop into thymocytes are initially double negative (DN) for the co-receptors CD4 and CD8. The DN precursors give rise to two T cell lineages: the minority population of $\gamma\delta$ T cells which lack CD4 or CD8 when mature, and the majority population of $\alpha\beta$ T cells. The development of $\alpha\beta$ T cells is the focus of this work. Following rearrangement of the TCR chains, the prospective $\alpha\beta$ T cell populations proceed to into $CD4^+CD8^+$ double positive (DP) thymocytes. Most DP thymocytes die within the thymus, after becoming DP cells, but those cells whose receptors can interact with self-peptide:self-MHC molecular complexes lose expression of either CD4 or CD8 and increase the level of expression of the T-cell receptor, a process called positive selection[33]. Positive selection occurs through interaction with thymic epithelial cells. The opposite process called negative selection occurs through interaction with thymic antigen presenting cells (APCs) to eliminate T cells with too strong of self-reactivity by the TCR[34].

The cells that successfully pass both positive and negative selection events differentiate into single positive (SP) CD4⁺ or CD8⁺ naïve T cells through silencing transcription of one co-receptor locus. Lastly, mature SP T cells are exported from the thymus to peripheral lymphoid sites. Naïve T cells emerging from the thymus in humans express CD45RA, an isoform of CD45, and the lymph node homing receptor CCR7 [35]. This stepwise process of T cell development in the thymus is highly regulated and occurs within distinct micro-anatomical niches within the thymus, containing different stromal cell types essential for T cell education and formation [16]. CD4 and CD8 are important cell-surface molecules for identifying thymocyte subpopulations with distinct properties and functional abilities.

Formation of the T cell receptor

T cells sense the presence of antigens in their environment through cell surface T cell antigen receptor (TCR). Individual T cells express many copies of the same T cell receptor on the cell surface. Each developing T cell expresses a unique version of the TCR, that allows T cells to collectively have the ability to respond to millions of different antigens. If the human genome contained a unique gene to encode for every possible T cell receptor, this would add more genes than currently exist in the genome. Instead, the immune system has evolved a mechanism for generating highly variable proteins from a limited number of genes involving rearrangement of and joining of gene segments. The gene segments rearrange by somatic gene rearrangement to create variable regions of coding sequences. Gene rearrangements occur before the DP positive stage in T cell development as described above. These gene rearrangements allow the generation of an estimated theoretical diversity of 10^{20} different T cell receptors[36], each expressed by a single T cell. The mechanisms of gene rearrangements are common also to

B cells which produce antibodies. The generation of diverse T cell receptors is important for T cell ability to recognize diverse antigens and modulate their functions in protection.

The TCR is a transmembrane receptor that is made up of two protein chains: α and β . There are also other types of T cells that develop in the thymus, including $\gamma\delta$ T cells; the focus of this work is on majority T cells with $\alpha\beta$ TCRs. Each protein chain consists of two regions; a variable (V) region which binds to antigen, as well as an invariant constant region (C) which functions downstream of antigen recognition for intracellular signaling and subsequent effector responses[37]. The V-region is the region that undergoes gene rearrangement which accounts for generation of unique sequences to code for unique T cell receptor proteins. The variable region is coded by the following gene segments: variable (V), diversity (D), and junction (J) which generate a functional VDJ- β chain which is why this gene rearrangement is called VDJ-recombination. After rearrangement, a novel region in between the V and J gene segments is generated called the CDR3 β which is the most variable region of the T cell receptor, due to both junctional diversity (from combination of different V, D and J gene possibilities) as well as mutational diversity which arises due to errors prone to the gene rearrangement process of repaired breaks in the DNA. The V-region exon is spliced in order to join to the C region. The resulting mRNA is then translated to yield the β -chain of the T cell receptor. The α chain results from rearrangement of a different set of V and J gene segments. Once a productive β chain gene rearrangement has occurred, the α and β chains pair soon after they are synthesized to form the $\alpha:\beta$ TCR heterodimer that is expressed on the T cell surface [38]. Expression of a functional $\alpha:\beta$ TCR suppresses rearrangement of the second TCR allele through a process called allelic exclusion. The layout of the different gene segments within the V region is depicted in Figure 1-3, showing the germline encoded gene segments before and after somatic rearrangement as well as how splicing and translation form the T cell receptor.

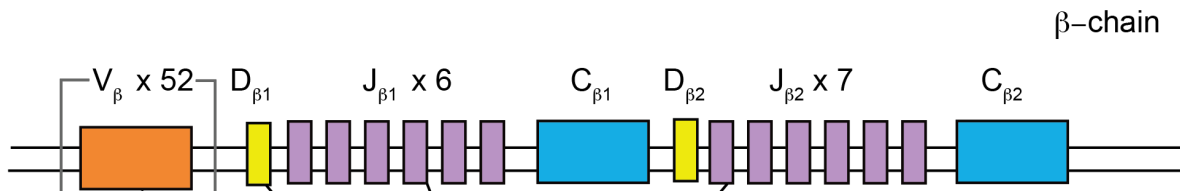
The $\alpha:\beta$ TCR alone is not sufficient to form a complete cell-surface antigen receptor. The TCR associates with invariant chains that carry out the signaling functions of the receptor to form a complete TCR complex[39-41]. The TCR complex consists of the $\alpha:\beta$ TCR heterodimer, four CD3 protein chains (CD3 γ , CD3 δ , and two CD3 ϵ chains), and the ζ -chain. The TCR complex is essential for T cell development, activation, differentiation and maintenance.

Figure 1-3. Gene rearrangements forming the TCR

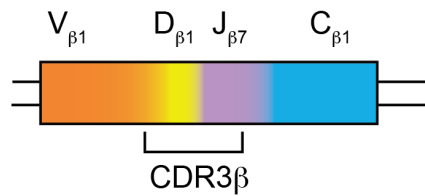
(A) The germline encoded locus for the human β -chain of the TCR before and after VDJ recombination and splicing. The locus coding for the β -chain of the TCR (TCR- β) located on chromosome 7 is comprised of 52 functional variable (V) gene segments distantly located from two segments each containing a diversity (D) gene and 6-7 joining (J) gene segments and single constant (C) gene. The TCR α -chain locus is not shown and has different numbers of gene segments with lower combinatorial diversity. After rearrangement, a novel region in between the V and J gene segments is generated called the CDR3 which is the most variable region of the T cell receptor, due to both combinatorial diversity (from combination of different V, D and J gene possibilities) as well as mutational diversity which arises due to errors prone to the gene rearrangement process of repaired breaks in the DNA. (B) The rearranged DNA is then transcribed, spliced and translated to form the β -chain of the TCR that binds most closely with antigen as shown in the diagram. Coloring of the β -chain corresponds to the rearranged locus that codes for that protein. Other parts of the TCR bind more with relatively conserved regions of MHC. The TCR α -chain (shown in dark grey), also important for binding antigen-MHC complexes, overall has lower diversity than the β -chains generated and structural data suggest it may be more important for binding to MHC.

(A)

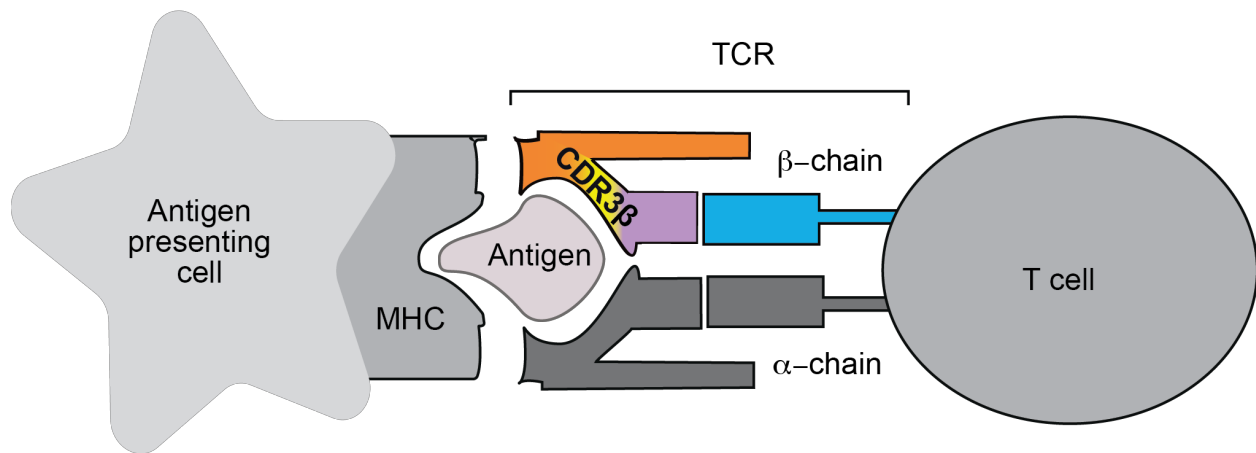
Before rearrangement:



After rearrangement:



(B)



Part A of this figure is adapted from Janeway's Immunobiology 8th Edition by Kenneth M. Murphy [38]

The mechanisms that coordinate VDJ-recombination include recombination signal sequences (RSSs) that flank the gene segments and are recognized by recombinase enzymes such

as RAG-1 and RAG-2. The roles of enzymes involved in the process of VDJ recombination have been elucidated through naturally occurring and induced mutations in mice. For instance, lack of RAG-1 and RAG-2 leads to faulty development of functional lymphocyte populations[42, 43]. Mutations in these proteins lead to severe immune deficiencies, part of a broader category of illnesses of the immune system known as severe combined immune deficiency (SCID), due to combine B and T cell defects [42]. The immune compromised state results from an inability to complete VDJ recombination; this further illustrates the absolute importance of these receptors in function of lymphocyte cell responses.

TCR gene rearrangement is central to the development of full naive T cell repertoire, conferring specificity to diverse antigens. Studies of human umbilical cord blood (UCB) show that Ag-specific T-cell precursors to common viruses and immunogenic antigens exist in the naive repertoire across a range between 1 and 352 cells per 100,000 CD8⁺ cells [44]. However, UCB is not necessarily representative of newborn T cell complement in terms of phenotype and transcriptional profile[45]. Therefore more studies are required to fully understand repertoires of T cells at birth by studying other tissues present and functioning after birth such as lymph nodes and thymus.

Age associated changes with T cell development

New T cells eventually cease to form with age. This is due to age-associated changes in the thymus which includes reduction of thymic volume, loss of thymic epithelial cells, increase in the perivascular space and predominance of adipose tissue [46]; however, the quantitative contribution of these changes to thymopoiesis has not been clear. A recent study of thymus tissue from surgical explants and organ donor tissues from infancy through 73 years of age found active thymus tissue from H&E staining at all ages, but large increases in adipose tissue and

fewer Hassel's corpuscles (a structural hallmark of the human thymus) in adult compared to infant thymus tissues [9]. Analysis of thymocyte populations in that study revealed that DP CD4⁺CD8⁺ thymocytes comprised 60-80% of total thymocytes (similar to DP frequencies in active human and mouse thymi) up until the fifth decade of life, after which the DP frequency was reduced to 5-15% thymocytes [13]. These findings show that active thymopoiesis does not exhibit a gradual decline but may cease abruptly at some discrete point in time after 40yrs of age. The cessation of development of new T cells, yet maintained ability to respond to new infections throughout life, is a testament for how long-lived T cell populations are, maintaining the ability to respond to new infections from T cell populations developed earlier in life.

Section 1.3 T cell activation and differentiation

T cell activation

T cell activation and differentiation occur in secondary lymphoid organs such as lymph nodes where naive T cells and APCs meet. This meeting occurs because the lymphatic system drains the fluid that leaks from our blood vessels and builds up in tissues, carrying that fluid to lymph nodes throughout the body. Traveling in lymph fluid are APCs that have picked up foreign material or antigens from distant tissues, storing that information to be delivered later to T cells. Upon arrival in lymph nodes, APCs deliver this information by direct contact with naive T cells, initiating the adaptive immune response. It is estimated that an adult human has over 500 lymph nodes throughout the body.[4] Additionally, other lymphoid organs can function as sites for T cell activation including bone marrow and spleen. [47, 48]

Naive T cells are first activated via contact with dendritic cells (DCs), a special class of APCs[49]. There are two important signals that DCs transfer to T cells; one signal is transferred to the TCR complex in the form of antigen complexed with MHC proteins (peptide-MHC

complexes), and the second signal is costimulation which is mediated through CD28 expression on T cells. Costimulation signal is important for naive T cell responses, and specifically binds to B7 receptor family members, for instance CD80 and CD86, on DCs. In naive T cells, CD28 costimulation enhances cell cycle entry, and stimulates IL-2 important for T cell survival, as well as induction of differentiation programs leading to helper T cell and cytotoxic effector responses T cells. Signals received by TCR and CD28 converge by intracellular T cell signaling elements that lead to downstream changes in the nucleus altering transcription[50]. During activation, T cells transiently upregulate expression of surface expression of CD69[51], a marker of early T cell activation which promotes retention of T cells through S1PR1 [52, 53]. This activation event, called T cell priming, initiates the T cell differentiation program.

CD4⁺ T cells recognize antigen by presentation on MHC Class II complexes [54-58], expressed on specialized subsets of antigen presenting cells including but not limited to dendritic cells and macrophages [59]. These antigen presenting cells can take up exogenous antigens and present them, therefore providing a mechanism for T cells to respond to extracellular pathogens like certain bacteria in addition to intracellular pathogens, like viruses. CD8⁺ T cells recognize antigen by presentation on MHC class I complexes which are expressed on all nucleated cells in the body.

Priming leads to cellular changes including entry into the cell cycle resulting in proliferation and differentiation into diverse types of effector cell populations which carry out pathogen clearance. Effector cells also gain expression of homing molecules such as CXCR3 and CCR5 that allow exit from priming sites and entry to sites of inflammation by trafficking through the blood. Additionally, in order to exit from lymphoid tissues, effector T cells downregulate chemokine receptor CCR7 and selectin molecule CD62L. Alternatively, for T cells that act to help antibody responses, expression of CXCR5 allows homing to B cell follicles[60, 61]. (For

review of integrins, chemokine receptors and their respective ligands important for T cell migration see here [62]).

CD4⁺ T cell function and differentiation

When activated, naive CD4⁺ T cells differentiate into several diverse subsets known as T helper (Th) cells including Th1, Th2, Th17, T follicular-helper (Tfh) and T regulatory (Treg) each with specific roles. Th1 cells produce INF- γ and are important for responses to intracellular pathogens and their effector responses include activating macrophages and inducing inflammation. Th2 cells produce IL-4, IL-5, IL-13 and respond to helminth infections. Tfh cells are specialized in helping B cell responses in germinal centers for production of antibodies. Th17 cells produce IL-17A, IL-17F and IL22 and control extracellular pathogens. Th cells have been associated with specific diseases including autoimmunity for Th1 and Th17 cells[63, 64], and allergic responses for Th2 cells [65-67]. Treg cells express IL-10 and prevent over-active immune responses and associated immunopathology[68-70]. There is also a subset of cytolytic T cells that produce granzyme B, perforin, and FASL to promote killing of cells. Additionally, there are emerging types of Th cells such as Th9 and Th22 cells [71, 72].

The intracellular signaling events induced by priming lead to changes in gene expression. These changes are induced by specific transcription factors that regulate the formation of distinct helper T cell subsets. Transcription factors act as “master regulators” that coordinate the expression of different transcriptional programs. Each CD4⁺ T cell subset has unique transcriptional regulator that is required for development: ROR γ for Th17 [73], FOXP3 for Treg [74, 75], GATA3 for Th2 [76, 77], Bcl-6 for Tfh [78], T-bet for Th1 [79, 80], and Eomes for cytolytic T cells (

Figure 1-4). While these transcription factors are important for lineage specification, in reality, the delineation of the subset may be more complicated with certain TFs have being important for multiple lineages [81, 82] as well as evidence of the possibility for plasticity within Th subsets depending on context of signals received by T cells for a recall response. [83, 84]

CD8⁺ T cell function and differentiation

Activated CD8⁺ T cells differentiate into CTLs that kill infected cells through granzymes and perforin and secrete cytokines such as IFN- γ and TNF- α . Upon activation, T cells rapidly proliferate into a large pool of cells which can then circulate though the body and home to the tissue where inflammation is occurring, and there perform killing of the targeted infected cells. Most of the effector cells die off after clearance of an infection and subsequently a small portion of those previously activated cells remain long lived and differentiate into memory T cells.

When CD8⁺ T cells differentiate into effector cells, they gain the ability to kill targeted cells. Functionally, this is achieved by production of several enzymes, including perforin and granzyme[85, 86]. This occurs through cell killing of targeting cells by apoptosis. This takes place when CTLs recognize antigen presented by MHC class I on the surface of an infected cell. Subsequently, T cells release granules containing cytotoxic molecules like perforin, which creates holes in the membrane of the target cells, and granzymes, which are enzymes that can induce apoptosis in the target cell after diffusing through pores created by perforin[86].

In order to regulate T cell responses, it is important for T cells to also be eliminated to decrease potential harmful effects of T cell responses including aberrant cell killing that can result in autoimmunity and tissue damage. Therefore, following proliferation and function of effector T cells, the majority of those cells die, with only 10-15% remaining as memory T cells. A major mechanism for cell death of T cells includes Fas pathways induced by TCR signaling

events[87]. Engagement of Fas (CD95) by Fas ligand (FasL) results in apoptotic cell death mediated by caspase activation. This pathway is important in regulating cell death within lymphoid compartment, and is also known as activation-induced cell death (AICD)[88, 89].

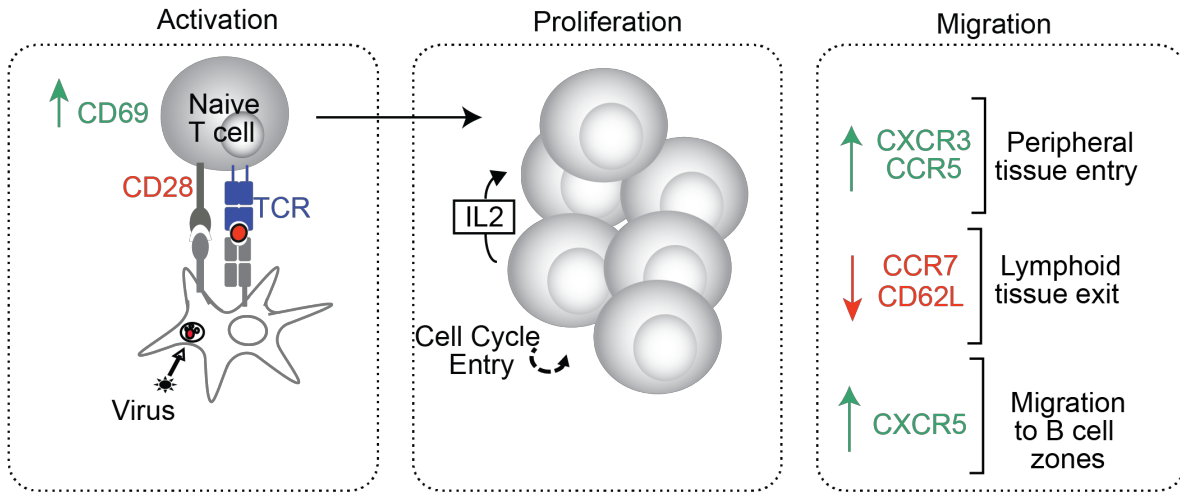
Transcription factors also play a key role in the differentiation of CD8⁺ T cells. Many of the same transcription factors important for CD4⁺ T cell differentiation are also important for CD8⁺ T cell differentiation. Transcription factors function to control the maintenance of naive T cells, and the generation of both effector T cells that provide shorter and longer term protection; also known as regulating formation of effector and memory T cell potential. Studies in mice have revealed the importance of Foxo transcription factors including Foxo1 in maintaining T cells in naive state. The expression of key factors in this process including CCR7 and L-selectin are regulated by Foxo1. Through conditional deletion of Foxo1, its importance in maintenance of naive T cell homeostasis was revealed through regulation of several genes critical for T cell trafficking and survival including transcription factor Klf2.[90] Further, deletion of Foxo1 led to severe defect in trafficking of naive T cells to lymph nodes upon cell transfer compared to wildtype cells as well as decreased expression of interleukin 7 receptor alpha chain. Together the maintenance of a naive T cell state seems to be regulated by Foxo1 by regulation of homing and survival signals. These studies also showed the importance of tissue localization in maintenance of T cell states. [90]

T-bet and Eomes, two T-box transcription factors, have crucial roles in formation and function of effector and memory T cells. They function with partially redundant roles to create CTLs by inducing expression of IFN- γ ., granzyme B, perforin, and CXCR3. [91] Eomes is

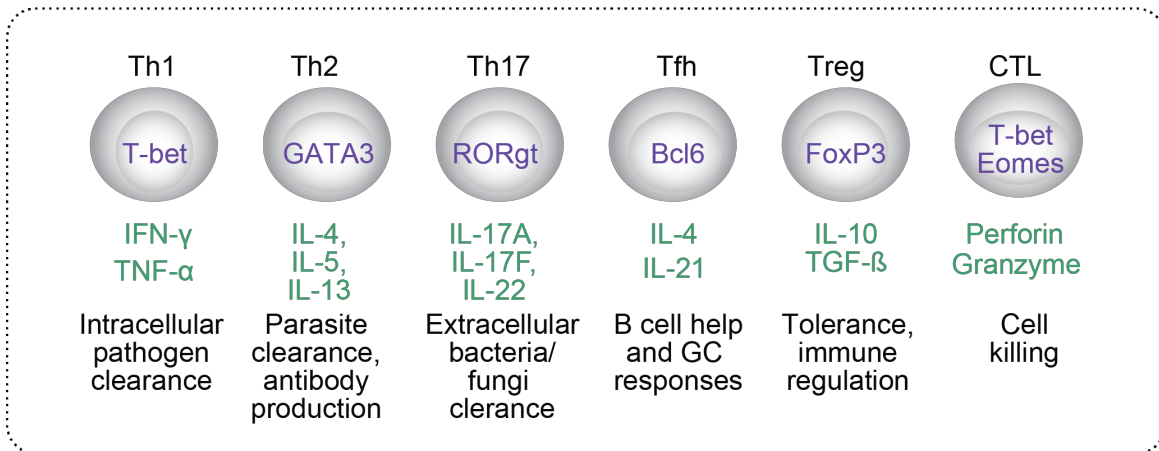
Figure 1-4. T cell activation and differentiation.

(A) Naive T cell activation is initiated by recognition of MHC/antigen complexes by their TCR (blue) and costimulation by CD28, resulting in transient upregulation of CD69. The differentiation program initiated by recognition of antigen results in cellular changes including cell cycle entry and production of IL-2 leading to cell proliferation, differential expression of chemokine receptors and selectin molecules that mediate migration to sites of inflammation, and the gain of effector functions including production of cytokines and cell killing molecules. Effector cells that gain expression of CXCR5 are able to enter to B cell follicles in lymphoid sites. Effector cells that gain expression of CXCR3 and CCR5 and lose expression of CCR7 and CD62L are able to exit from lymphoid sites and enter peripheral tissues. During differentiation, into individual T cell subsets, T cells acquire effector functions. Naive CD4⁺ T cells differentiate into several diverse subsets known as T helper (Th) cells including Th1, Th2, Th17, T follicular-helper (Tfh) and T regulatory (Treg) each with specific roles. The roles of helper T cells include: helping B cell and CD8⁺ T cell responses, performing cell killing, producing diverse types of effector cytokines, and regulating immune responses by preventing over-active responses. Naive CD8⁺ T cells differentiate into cytotoxic T lymphocytes (CTLs) that kill infected cells. The fate determination of T cells is thought to be regulated by the quality of signals that a T cell receives during stimulation which can depend on the context, strength and duration of peptide-MHC complex and costimulatory molecule binding and cytokine environment. [92]

1° Infection



Effector Function



expressed in activated CD8⁺ T cells. [93] T-bet and Eomes regulate cytotoxic and memory formation abilities of T cells in part by impacting interleukin 15 (IL-15) [94]. Additionally, these transcription factors can play non-redundant and additive roles in CD8⁺ T cell differentiation [95]. CD8⁺ T cells lacking T-bet and Eomes lose CTL identity and become IL-17 producing CD8⁺ T cells that cause a lethal inflammatory syndrome during LCMV infection.[96]

Evidence suggests that the same transcription factors also regulate T cell responses in humans. A study of human virus-specific T cells showed that T-bet and Eomes were markers for T cell capacity and ability for durable immune control. This study tracked virus specific T cells in response to chronic infection with cytomegalovirus (CMV) in lung transplant recipients who are mismatched from CMV⁺ donors and therefore at a greater risk for active CMV infection and higher mortality. This study found that the transcription factor expression profile of CMV-specific T cells was indicative of their *in situ* function in viral clearance. Specifically, transcription factor expression patterns of T-bet > Eomes differentiated lung transplant recipient controllers from viremic relapsers[97]. These findings were consistent with previous findings in mouse models that show Eomes expression was up-regulated in exhausted CD8⁺ T cells during chronic infection[98]. In summary, high levels of T-bet expression are correlated with better outcomes in an ongoing infection highlighting the importance of T-bet for CD8⁺ CTL effector functions.

In summary, T cell differentiation begins with the activation of naïve T cells in secondary lymphoid tissue, leading to T cell proliferation, differentiation to effector cells that acquire the ability to migrate to sites of inflammation and effector functions, graphically depicted in Figure 1-4.

Section 1.4 Maintenance and Formation of T cell memory

Human memory T cell subsets

Following clearance of pathogen, T cells undergo contraction where the majority of pathogen specific T cells die by apoptosis, and typically a small percentage survive and mature to become memory T cells. Due to an accumulation of antigen exposures over life, in adult humans, most T cells in the body are memory T cells [6, 7]. Memory T cells have two main functions: one is to self-renew, and the second is to mediate protective immunity upon re-exposure to pathogen. The ability to self-renew is mainly performed by central-memory (TCM) cells which also have the ability to create effector-memory (TEM) daughter cells, while the ability for immediate effector function is performed by TEM cells. Therefore, a TCM cell has been referred to as having “stem-like” properties, while TEM cells can be thought of as more differentiated [99].

The two most commonly used markers to distinguish human naive and memory T cells are CCR7 and CD45. CCR7 is a chemokine receptor that is required for entry into lymph nodes through high endothelial venules (HEV) [100-102]. Both naive and TCM cells express CCR7, indicative of their predominance in and migration patterns to lymphoid sites. TEM cells do not express CCR7 and are able to migrate to diverse peripheral tissue sites. [99] The lack of CCR7 expression on TEM cells is important for exit from lymphoid tissue and entrance into peripheral tissues and sites of inflammation [103, 104]. CD45 is a tyrosine phosphatase that regulates signaling through antigen receptors. Naive T cells express the long isoform of CD45 known as CD45RA, while memory T cells express the shorter CD45RO isoform [105, 106]. These markers also can be used to identify terminal effector T cells in humans, specifically CCR7 negative cells can re-express the CD45RA isoform and this population of cells are known as TEMRA cells. More recently an additional subset of memory T cells has been identified as a prominent T cell

population in mice and man: those that remain resident in diverse tissues termed tissue resident memory (TRM) cells, that often express CD69. The differentiation state of TRM in humans is unknown, and a topic of active investigation which will be address in future chapters within this dissertation.

In addition to the major delineations of memory versus naive T cells including markers CCR7 and CD45RA, there are many other phenotypic differences between naive and memory T cells[107-109]. Memory T cells express higher levels of IL-2R β -chain (CD122) a component of both the IL-2 and IL-15 receptors both thought to be important for maintenance of memory T cells. [110] Many of the markers upregulated on memory cells are also upregulated on effector cells, and therefore it is not always easy to distinguish an actively responding T cell to antigen (effector cell) from a memory T cells. In human blood TEMRA cells bear more features of an actively responding T cell. They express higher levels of perforin and granzyme and lower levels of CD27 and CD28, and are thought to arise from chronic exposure to antigen, aligning with the fact that CMV seropositive individuals, a chronic virus, tend to have higher frequencies of TEMRA cells predominantly in the CD8⁺ T cell compartment [10, 111, 112].

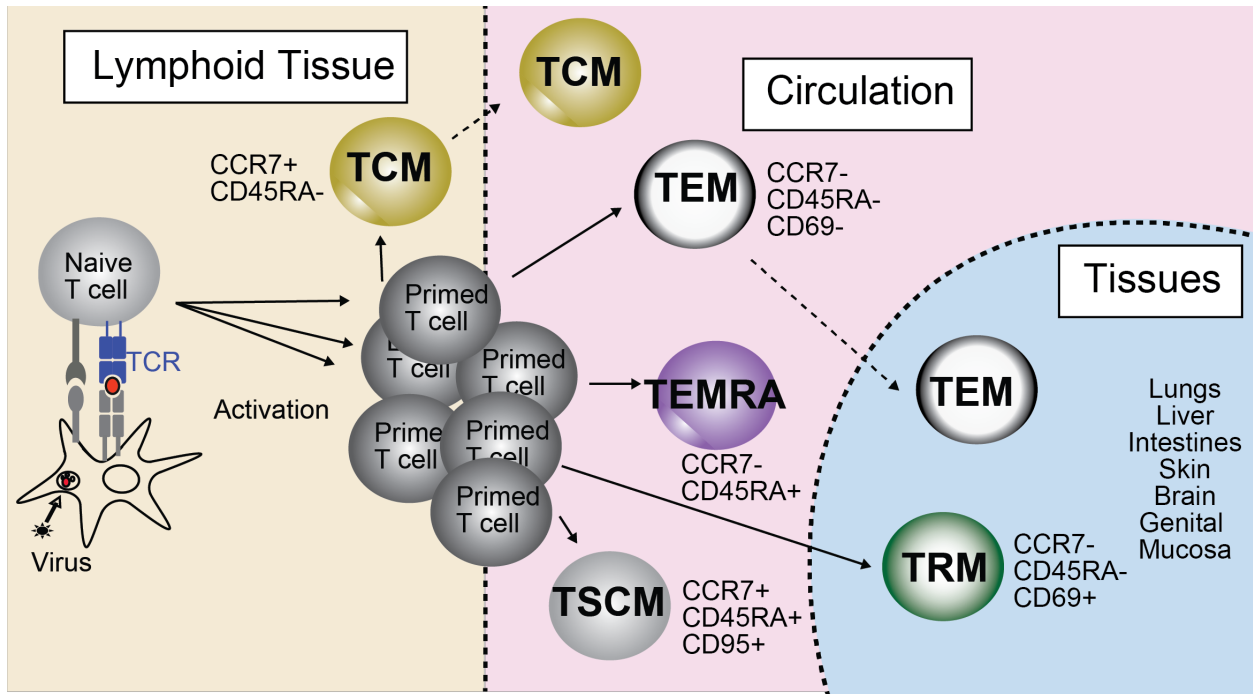
More recently, an additional subset of memory cells, called stem memory T cells (TSCM) were identified [113-115]. TSCM are phenotypically similar to naive T cells and express CCR7 and CD45RA however are still distinct from naïve T cells by their epigenetic marks indicating previous activation. To complicate matters in identification, they can be distinguished by other markers similar to memory T cells including CD95. TSCM cells have been found to be specific to both viral and self-tumor antigens. Additional markers expressed by TSCM similar to naive and TCM cells are higher levels CD27 and CD28. In contrast, they also express increased levels CXCR3 and LFA-1 both attributes of memory T cells. In humanized mouse models, these T cells displayed increased proliferative capacity, more efficient

reconstitution of immunodeficient hosts, and increased efficacy in anti-tumor responses. TSCM cells are a rare population of cells that can be detected within peripheral blood, and it remains unknown if they are present in other sites[116]. These subsets and their localization patterns as well as the markers used to identify them are outlined in Figure 1-5.

Figure 1-5. Tissue localization patterns and phenotypes human T cell subsets.

(A) Distinct subsets of T cells can be identified by their surface expression for the following markers: Naïve T cells, $CCR7^+CD45RA^-$; Central memory (TCM), $CCR7^+CD45RA^-$; Effector Memory (TEM), $CCR7^-CD45RA^-$; Tissue resident memory (TRM), $CCR7^-CD45RA^-CD69^+$, Stem cell-like memory (TSCM), $CCR7^+CD45RA^+CD95^+$ and Terminally differentiated effector cells (TEMRA) $CCR7^-CD45RA^+$. Each subset has a unique tissue localization pattern, with Naïve and TCM cells found in lymphoid tissue in addition to circulation, TEM cells found predominantly in circulation as well as peripheral tissues, TEMRA cells in circulation and tissues, TSCM cells in circulation, and finally TRM cells exclusively maintained in tissues.

(A)



Functional capacity of human memory T cells

The most important difference between memory and naive T cells is the ability of memory phenotype cells to respond more rapidly and robustly to antigen stimulation, known as a recall response. How does this occur? First, memory T cells are more sensitive to antigen stimulation and are less dependent on CD28-mediated co-stimulation. This could explain why naive T cells require dendritic cell antigen presentation, while memory T cells respond well to antigen presented on other APCs such as resting B-cells.[109, 117-119] Additionally, once activated memory T cells respond differently and produce different cytokines, specifically more IFN- γ and TNF- α compared to naive T cells[120].

In order to study human T cells in bulk, they can be stimulated polyclonally as their antigen specificity is not readily known. Upon polyclonal stimulation, naive T cells produce a lot of IL-2, central memory cells produce some and effector memory cells do not produce much IL-2[121]. However, for production of effector cytokines, memory cells produce these rapidly including IFN- γ and TNF- α in response to stimulation while naive T cells do not, and TCM and TEMRA cells produce some but not as high levels as effector memory T cells[121]. Studies in mice and humans have shown that multifunctional memory T cells producing IFN- γ , TNF- α and IL-2 define a correlate of vaccine mediated protection against *Leishmania major*, highlighting the importance of the quality of the response as well as the magnitude[122]. Additionally, a study of virus-specific CD4⁺ and CD8⁺ T cells in human peripheral blood found multi-functional IFN- γ ⁺TNF- α ⁺IL-2^{+/-} memory cells generated in response to chronic infection with CMV and are thought to be important for ongoing viral control[123, 124].

TEMRA cells produce cytokines, albeit at lower levels than TEM cells, have a dampened ability for proliferation, and express higher levels of cytotoxic molecules such as granzyme A/B and perforin compared to other T cell subsets[120, 124]. TEMRA also have a reduced capacity

for production of IL-2, an important cytokine for cell survival [125]. They are the most terminally differentiated T cell type, with expression of CD57, a molecule associated with replicative senescence and antigen-induced apoptotic death [126, 127]. The majority of TEMRA are CD8⁺ T cells, although rare populations of CD4⁺ TEMRA cells can also be present in certain individuals. The predominance of TEMRA cells increases with age [128] as well as infections with pathogens including CMV (CD4⁺ and CD8⁺ TEMRA), and dengue (CD4⁺ TEMRA) [10, 111, 112, 129, 130]. While TEMRA cells are most similar in phenotype to short lived effector cells (SLECs) in mice, it is unknown whether they represent the same population. In mouse models, SLECs can readily be defined and dissected by their short-term nature, their synchrony with timing of infection, and their specificity that can be precisely mapped using cell transfer and genetic models of manipulation. In contrast, in humans such experimental manipulation is not possible, and therefore understanding the lifespan of TEMRA cells and how they are generated *in vivo* is more difficult and remain open questions.

The different functional capacities, phenotypes and localization patterns of human T cell subsets are summarized in Table 1-1.

Table 1-1. Properties of human T cell subsets.

(++++)= Very High, (++++)= High, (++)= Medium, (+)= Low, N.D. = not detectable.

T cell subset	Identification	Cytokine production	Cytotoxic molecules	Proliferative capacity	Recall response	Localization
Naive	CCR7+ CD45RA+	IL-2 (++++)	N.D.	(++++)	(+)	Lymphoid tissues, blood
Central Memory (TCM)	CCR7+ CD45RA-	IL-2 (+++) TNF- α (+) IFN- γ (+)	Perforin (+) Granzyme (+)	(+++)	(+++)	Lymphoid tissues, blood
Effector Memory (TEM)	CCR7- CD45RA-	IL-2 (++) TNF- α (++++) IFN- γ (++++)	Perforin (+++) Granzyme (+++)	(++)	(+++)	Peripheral and lymphoid tissues, blood
Resident Memory (TRM)	CCR7- CD45RA- CD69+ CD103+	IL-2 (++) TNF- α (++++) IFN- γ (++++) IL-17 (++)	Perforin (++) Granzyme (++)	not known	not known	Peripheral tissues
Effector Memory RA (TEMRA)	CCR7- CD45RA+	IL-2 (+) TNF- α (+) IFN- γ (+)	Perforin (++++) Granzyme (+++++)	(+)	(+)	Blood, peripheral tissues
Stem-cell memory (TSCM)	CCR7+ CD45RA+ CD95+	IL-2 (++++) IFN- γ (+++)	Perforin (+) Granzyme (+)	(++++)	(++++)	Blood, other tissues not known

Molecular basis for memory T cell formation

The distinct features of memory versus naive and effector T cells are reflected in differences in gene expression [131]. Studies doing whole transcriptome profiling by RNA-sequencing, as well as transcriptome profiling by microarray, have found that the type of genes expressed in naive versus memory T cells are largely the same with over 95% overlap in the degree of similarity between naive and memory T cell transcriptomes, found by analysis of nine published studies (reviewed here [132]). However, the genes that are highly over expressed in either naive or memory T cells have very important roles in their respective functions for activation, effector function, and homeostasis. In memory T cells, upregulated genes include those involved in activation such as MHC class II genes (*HLA-DRA*, *HLA-DRB1*, *HLA-DPA1* and *HLA-DPB1*) as well as genes important for migration including chemokine receptor genes, such as *CCR5*, *CCR6*, *CXCR3* and *CXCR5*. Additionally, genes involved in intracellular signaling are also upregulated in memory T cells and these include mitogen-activated protein 3 kinase 5 (*MAP3K5*), dual specificity phosphatase 4 (*DUSP4*), regulator of G-protein signaling 1 (*RGS1*) and S100 calcium-binding protein A4 (*SI00A4*). Genes involved in effector functions that are upregulated in memory T cells include molecules important for cell killing such as granzyme A (*GZMA*) and *GZMK*. Finally genes important for homeostasis that are upregulated in memory T cells include expression of interleukin receptors *IL2RB* and *IL10RA*, among other cytokine receptor genes. Transcription factors regulate expression programs and these are also found to be differentially expressed in memory T cells. Notably, both *MAF*, thymocyte selection-associated high mobility group box (*TOX*) and *TBX21* (which encodes T-bet) are highly expressed in memory T cells compared to naive T cells[132].

Most of the highly expressed genes in memory T cells are shared between TCM and TEM subsets. However, there are several genes that are even more highly expressed in TEM

cells including *CCR2*, *LGALS1*, *LGALS3*, the MHC class II genes *HLA-DPB1*, *HLA-DQA1*, *HLA-DRA*, *HLA-DRB5* and *HLA-DRB6*, and integrin α M (*ITGAM*) for CD4⁺ T cells.[132] Genes more highly expressed in CD8⁺ TEM cells compared to TCM cells mostly relate to genes coding for cytotoxic molecules such as *GZMH*, as well as effector function genes such as *IFNG*, and genes encoding cell surface chemokine receptors including *CCR6* and *CCR9*. Together, these show that the highly expressed genes in TEM cells indicate their increased effector function and differential migration capacities.

This differential gene expression signature has been found to be driven by epigenetic changes including histone methylation patterns providing a chromatin basis for differential gene expression between naive and memory T cells[133]. Methylation patterns within the DNA occur predominantly at clusters of CpG dinucleotides (known as CpG islands). DNA methylation at CpG islands has been shown to regulate genes important for effector function and activation in memory T cells in numerous studies [134-138]. In a study examining naive, memory and effector T cells in mice, the *IFNG* locus was found to be differentially methylated according to subset. The *IFNG* locus was highly methylated in naive T cells, partially methylated in memory T cells, and unmethylated in effector cells, indicating that methylation contributes to lower production of IFN- γ observed in naive T cells. Interestingly, the *IFNG* gene was rapidly able to demethylated in memory but not naive T cells within 5 hours of antigen stimulation without requiring any cell division. [136]. Further, more recent studies conducting genome wide analysis of human naive, TEM and TEMRA cells show epigenetic networks regulate transcriptional programs. Interestingly, TEM and TEMRA cells shared similar epigenetic marks, and were both distinct from naive T cells. [139]

Role of transcription factors in memory T cell fate

Given the dual roles of memory T cells play, the ability for effector function and self-renewal, how is the formation of these diverse subsets regulated within T cells? Genetic ablation studies in mice revealed the importance of T-bet and Eomes in the differentiation of T cells important for their effector functions including IFN- γ , perforin, and granzyme B production [94, 95]. Studies in mice have shown that higher expression of T-bet promoted short lived effector cells, while lower expression of T-bet promoted memory T cell precursor formation[91]. Transcriptional activation of Wnt responsive genes by T-cell factor 1 (TCF-1) and Lymphoid enhancer-binding factor 1 (LEF1) is important for maintenance of memory T cells. [140-146] Interestingly, TCF-1 and LEF-1 also play essential roles in multiple stages of T cell development including earliest stages of T-cell development during thymopoiesis [147-149], therefore conditional knockout models were used in order to elucidate their important roles at later stages of differentiation including for Tfh commitment [150-153], and CD8⁺ T cell lineages [140, 154].

Genetic ablation studies have elucidated the cooperative roles of TCF-1 and LEF-1 in CD8⁺ T cell responses and shown that these factors are especially important for memory formation and maintenance. TCF-1 and LEF-1 are highly expressed in naive T cells, downregulated in effector cells, and then upregulated again in memory T cells [145]. TCF-1 is required for T cell development as shown in mouse models where lack Tcf-1 leads to partial block in T cell development, however some T cells are still able to develop and those that do develop still retain effector capacity[148]. Genetic ablation of Tcf-1 and Lef-1 is lethal, therefore to study their combined roles in T cells, Lef-1 was conditionally deleted using a Cre-lox system. This allowed Lef-1 to be intact for T cell development of naive T cells, and then deleted upon effector differentiation under the control of granzyme B promoter activity. Using this model, CD8⁺ T cells deficient for TCF-1 and LEF-1 lost the ability to form KLRG1^{lo}IL-7R α ^{hi} memory

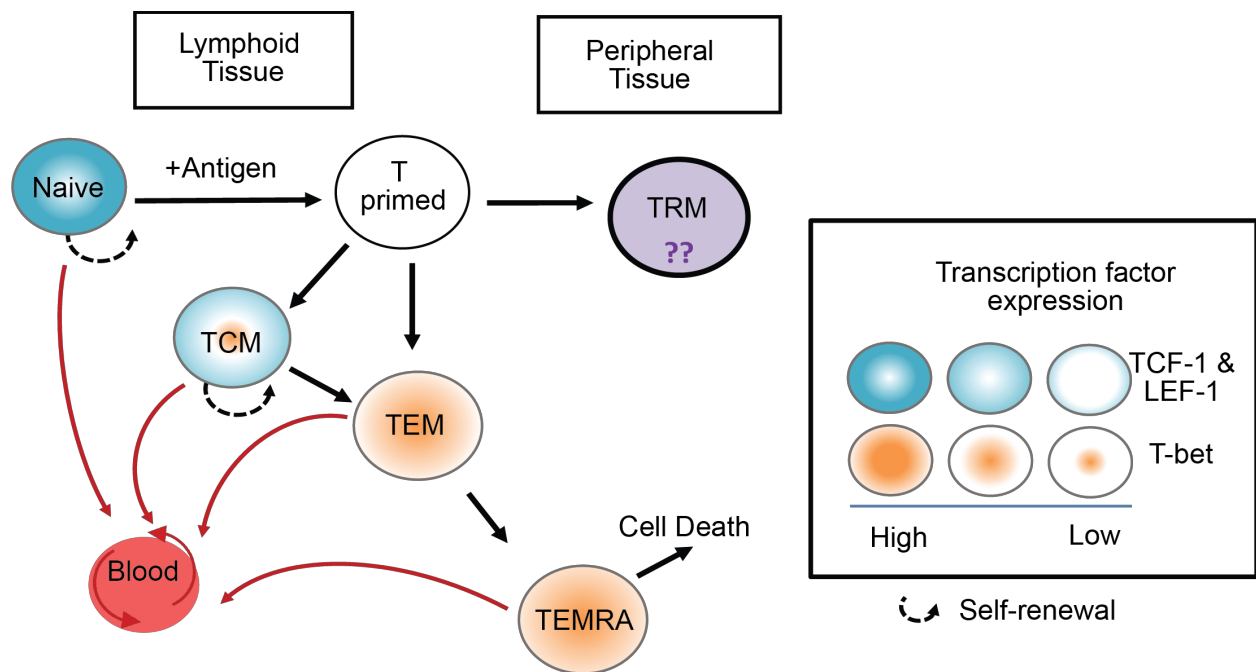
precursors, while maintaining the capacity to generate effector cells expressing IFN- γ , granzyme B, and perforin. Subsequently, these effector phenotype T cells were impaired in proliferation upon rechallenge [140, 155]. Conversely, transgenic mice with constitutively active Wnt signaling resulted in enhanced memory T cell formation and IL-2 production, resulting in enhanced immunity to secondary encounter to *Listeria monocytogenes*, a pathogenic bacterium. Taken together, these studies reveal the importance of Wnt signaling factors TCF-1 and LEF-1 in the generation and maintenance of CD8⁺ T cell memory.

In humans, less is understood about the mechanistic underpinning of memory T cell formation, but evidence suggests that the same transcription factors that regulate memory T cell formation in mice are also important in human T cells. Interestingly, T-bet expression is highest in TEMRA cells, also high in TEM cells, and lower in TCM cells, agreeing with mouse studies of the importance of controlled T-bet expression for maintenance of memory T cells [156]. Additionally, very recent studies have elegantly demonstrated that TCF-1 marks self-renewing cells in human blood. Naive T cells express TCF-1, and expression is maintained in central memory T cells and important for their self-renewal capacities [157]. The expression of LEF-1 follows similarly to TCF-1 with naive T cells having the highest expression, then TCM and TEM and TEMRA with the lowest expression.[158] The current knowledge of the expression of these T cell subsets is outlined in Figure 1-6 below, also indicating the gap in knowledge of expression of these key transcription factors in human TRM cells.

Figure 1-6. Expression of T-bet, Lef-1 and TCF-1 in human T cells

(A) Transcription factor expression with human T cell subsets is limited to those subsets which can be detected in peripheral blood. Naïve T cells express high levels of TCF-1 and LEF-1, TCM T cells express intermediate levels of TCF-1 as well as T-bet, TEM and TEMRA cells express high levels of T-bet and low levels of TCF-1. [156-158] For human TRM cells (purple) it remains unknown at what levels these transcription factors are expressed.

(A)



Models for mechanisms of T cell memory formation

The mechanism for how diverse memory T cell subsets develop from a single naïve T cell clone is debated and has been for years [159-162]. There are many different schools of thought of how this occurs [163], broadly speaking these can be broken down into two main models. In one model, a single naïve T cell upon priming, undergoes branching into distinct precursors, one that can form memory and other that will differentiate into effector populations. In this model, once a naïve T cell becomes an effector cell, it cannot go back to memory phenotype and will go off to proliferate and then die. Memory precursors in this model can also further differentiate into effector T cells indicating a single directionality of the differentiation process, known as the decreasing potential model. A second model is that all memory T cells first go through an effector phase, and then a portion of those effector cells re-express markers more similar to naïve T cells and become long lived memory, known as the de-differentiation model.

Experimental evidence to support both of these models exists. There are studies in mice and humans that support the de-differentiation model, including a study in humans tracking yellow fever virus (YFV) specific T cells. [164] In this study, participants were vaccinated with YFV vaccine (YFV-17D) and subsequently dosed with heavy water (D₂O) daily for two weeks after YFV-17D vaccination. Deuterium from D₂O is incorporated into DNA of cells that are actively dividing, and therefore quantification of deuterium levels in YFV-specific T cells derived from peripheral blood samples can reveal the rate of cell division during the time of D₂O exposure. Peripheral blood samples taken during the effector phase revealed the maximum levels of deuterium incorporated as virus-specific T cells underwent robust proliferation. From analysis of memory T cells from longitude peripheral blood samples taken 1-2 years later, they found high deuterium enrichment remains in the YFV-specific memory T

cells, indicating that this population originated from cells that underwent rapid proliferation during the effector stage of the immune response. Additionally, studies in mice provide supporting evidence that effector T cells de-differentiate to become memory T cells, by investigating changes in DNA methylation programming at naive and effector cell-associated genes in virus-specific CD8⁺ T cells during acute lymphocytic choriomeningitis virus infection in mice. The authors found evidence for elimination of *de novo* methylation profiles and re-expression of naive-associated genes after transfer of memory precursor (MP) cells, defined as KLRG1^{lo}CD127^{hi} cells[165]. However, there is an alternate explanation for the erasure of *de novo* methylation programs and re-expression of naive-associated genes observed. The analysis was of bulk MP cells which could comprise a mixture of more effector-like precursors as well as more memory-like precursors, and over time if the effector like precursors died and only the rare memory-like precursors remained, this could explain the apparent erasure of methylation at effector genes. Single cell analysis techniques can aid in elucidation of the origins of distinct memory T cell fates.

Evidence to support the decreasing potential linear model is that T cells carry epigenetic marks, and these differ between subsets with Naïve T cell being most similar to TSCM, then TCM, and lastly TEM and TEMRA. Further, TCM and TSCM cells have the ability [113] to convert to TEM cell phenotype, but not vice-versa, which is coupled to increased methylation of CCR7 and Tcf7 loci, indicating a directionality in T cell differentiation programs [166]. Additionally, there is evidence that TCR signaling strength can lead to generation of distinct memory and effector cells[167]. There is also substantial evidence that shows after several cell divisions of a naïve T cell due to activation with antigen, the resulting cell populations that bifurcate into effector cells are irreversibly effector and exhibit silencing of TCF-1 expression,

while the TCF-1⁺ cells maintain the ability to produce both TCF-1⁺ and TCF-1⁻ daughter cells.
[168-171]

Whether or not de-differentiation or the decreasing linear model is occurring, neither of these models explains how a single cell, whether naïve or effector, is able to produce two distinct daughter cells: effector and memory cells. Asymmetric cell division has been demonstrated to drive effector and memory T cell fates. During asymmetric cell division, cellular components including regulators of signaling and transcription are unequally partitioned resulting in two daughter cells with distinct fates. Asymmetric cell division is also broadly conserved in different phyla of life [172-175]. In T cells, unequal partitioning of PI3K signaling has been shown to mediate generation of TCF-1⁺ and TCF-1⁻ daughter cells with self-renewal and effector T cell fates respectively [168-171]. While the occurrence of asymmetric division has been shown, the mechanism underlying this occurrence- whether deterministic or stochastic is debated (reviewed in [176]).

Section 1.5 Tissue Compartmentalization of Memory T cells

Overview of Tissue Localized Memory T cells

Initially, immunological memory was thought to be maintained mainly by specialized circulating subsets of TEM and TCM cells in blood as described above to surveil lymphoid and peripheral tissues respectively. However, more recent investigations in mice have revealed that subsets of CD4⁺ and CD8⁺ T cells remain resident in tissues such as lung, skin, and gut long after infection resolution [177-183]. These studies established a new subset of cells designated tissue resident memory T cells (TRM). Functionally, TRM mediated immediate protection against diverse viral, bacterial and parasitic infections more effective than circulating memory T cell subsets [177-183]. Since these initial investigations, many additional studies continue to emerge finding TRM in almost every tissue examined including the salivary glands, brain, liver and lymphoid organs, with large populations in human tissues such as lung, skin, and liver [6, 8, 184-197]. TRM can also be generated in response to vaccination and cancer [182, 198-200], as well as playing pathogenic roles in human diseases[201, 202].

TRM provide localized protective immunity and immune surveillance in tissues. [203] TRM were established as resident and non-circulatory in mouse models by various methods including parabiosis assays [177, 189, 204], in-vivo antibody labeling to determine accessibility to circulation [179, 190, 205], and photo-conversion of T cells to track migration [206-208]. These studies showed that phenotypically the majority of TRM can be distinguished from circulating memory T cells by expression of CD69, [190, 205, 209] a surface marker which promotes retention of T cells through S1PR1 [52, 53] and the alpha E integrin CD103 for subsets of CD8⁺ T cells [209, 210]. TRM are mainly characterized in barrier tissues, but play emerging roles in lymphoid tissues as well [206, 208, 211]. Mouse models show how TRM development differs in infants, providing strategies and mechanism for better understanding of the molecular

underpinnings of TRM formation. Collectively, with a better understanding of TRM populations, we will have a better ability to target TRM generation with vaccination to promote long lived immunity.

Identification of TRM

TRM are generated in diverse sites in response to acute and chronic viruses, including influenza (lungs) [179], murine cytomegalovirus (MCMV) (salivary glands) [212], lymphocytic choriomeningitis virus (LCMV) (many sites) [213], and herpes simplex virus (HSV) (skin, vaginal mucosa) [198, 214]. An important experimental method in mouse models that has shown T cells remain resident in tissues is parabiosis, surgical joining the circulations of two mice, and tracking migration of T cells from one parabiont to the other. While blood T cells typically reach homeostasis between the partners within weeks, many T cells in tissues do not become equilibrated and remain in specific tissues of one parabiont. [177, 215-217] Another approach for identifying TRM in mice is the in-vivo labeling of circulating cells with fluorochrome-conjugated antibodies via intravascular injection, which can effectively identify T cells that are tissue resident at the time of injection and thus remain unlabeled. [179, 205, 218] Additionally, the transplantation of peripheral tissues into congenic or naive mouse recipients also permits the assessment of residency potential and persistence of putative donor TRM in the graft. [181, 213] The findings from these and other studies in mice have demonstrated that TRM are a distinct, non-circulating population of long-lived memory T cells.

Similar criteria using parabiosis and *in vivo* labeling to establish tissue residency cannot be applied to human T cells. However, certain clinical situations involving T cell deletional therapies, and sampling from organ and composite tissue transplantation have provided evidence for the persistence of tissue memory T cells maintained distinct from circulating counterparts.

Treatment of cutaneous T cell lymphoma (CTCL) patients with anti-CD52 depleting antibodies (alemtuzumab) eliminated circulating T cells from the blood, but spared a persisting resident population of CD4⁺ and CD8⁺ memory T cells in the skin[192]. Further, alemtuzumab is completely ineffective for a relatively benign skin-limited variant of CTCL called mycosis fungoides (MF). In MF patients, pathogenic T cells persist for decades in regionally defined inflammatory lesions in the skin. During treatment with steroids, the lesions can disappear but upon cessation of treatments, the lesions re-appear in exactly the same location in the skin, further indicating they are likely due to a resident T cell subset[219]. More severe variants of CTCL have been associated with circulating T cells and therefore require more systemic therapies. These studies show how TRM biology influences disease state and treatment considerations[201]. Transplantation of HLA-disparate tissues and organs containing endogenous TRM has created a natural experiment for assessing potential persistence of donor-derived TRM and development of tissue T cell populations from circulating recipient T cells. In intestinal transplants, donor-derived T cells were detected both in circulation and within the intestinal graft up to a year post-transplant with TRM phenotype cells in the intestinal graft [220, 221]. Similarly, epidermal CD8⁺T cells of donor origin in face transplant recipients were observed up to 2 years post-transplantation [222]. Together, these results indicate that human TRM persist in the tissue niche long-term, similar to mouse TRM.

While defining tissue resident in humans is more difficult due to limits in experimental methods, substantial evidence shows that these cells are predominant in many tissue sites. In both organ donor tissue and surgical resections, TRM-phenotype cells expressing CD69⁺/CD103 have been identified in virtually every tissue examined including lungs, liver, pancreas, lymphoid tissues, genital mucosa, the GI tract (stomach, jejunum, ileum, and colon), bone marrow (BM), and in brain obtained from autopsies [193, 195, 197, 223-226]. Transcriptional

profiling of CD69⁺ memory T cells from human lungs, spleen, liver and other sites has further revealed a conserved transcriptional profile distinct from blood memory T cells that exhibits key features with mouse TRM [8, 196, 227]. Importantly, the advantage of studying TRM in human tissues is the ability to directly associate TRM responses to protective immunity and specific disease states such as inflammatory and autoimmunity [201, 228].

In human tissues, the extent of CD69 and CD103 expression by CD4⁺ and CD8⁺ tissue memory T cells varies by tissue site and is highly conserved between diverse individuals [7, 8, 205]. The identification of TRM in both species has relied two key markers, CD69, originally defined as an early T cell activation marker, and expressed by a large proportion of tissue memory CD4⁺ and CD8⁺T cells in mice and humans, and CD103 (α E- subunit of the α E β 7-integrin that binds E-cadherin expressed on epithelial cells [229]), expressed by mouse and human CD8⁺memory T cells in mucosal and barrier tissue sites [6, 181, 187, 230]. Both of these molecules likely function in tissue retention or localization. CD69 serves as a signal for tissue retention by binding to and sequestering the sphingosine-1-phosphate receptor (S1PR1) which is required for tissue egress [52, 53, 231]; downregulation of S1PR1 is essential for TRM formation in mice [232]. While transient expression of CD69 on recently activated effector T cells similarly serves to retain them in the lymph nodes [52], CD69 expression by TRM is constitutive and not associated with expression of other activation markers like CD25, CD38 and HLA-DR [8]. Interestingly, genetic deletion of CD69 in mice results in a reduction (but not ablation) of CD8⁺TRM in the skin and lung [217, 233], and does not greatly affect CD4⁺TRM formation [190, 206]. Taken together, these findings demonstrate that CD69 is an important marker to identify TRM cells in humans; however, expression of CD69 is not itself sufficient for TRM formation.

CD103, the α E-subunit of the α E β 7-integrin that binds E-cadherin expressed on epithelial cells, is expressed by a subset of mouse and human CD8⁺memory T cells in mucosal and barrier tissue sites [6, 181, 187, 230]. CD103 expression is limited to certain subsets of TRM likely because E-cadherin: α E β 7-integrin interactions anchor CD103⁺ TRM in specific locations within the mucosal epithelium. Accordingly, the formation of CD8⁺ TRM in the skin of CD103⁻ deficient mice is reduced and CD103⁻ cells exhibit increased motility compared to wildtype controls [233, 234]. There are, however, populations of CD8⁺TRM outside of the epithelia in tissues, such as in secondary lymphoid organs and bone marrow [8, 15, 235-237], that do not express CD103 at steady state. Furthermore, the role of CD103 on CD4⁺ TRM is less clear, as the majority of CD4⁺ TRM in mice and humans do not express CD103, though smaller populations of CD103⁺CD4⁺ TRM persist in the lungs, intestines and skin [8, 238]. A parabiosis study in mice found that a proportion of CD69⁻CD4⁺ memory T cells egressing from the skin expressed CD103, suggesting that CD103 expression alone does not indicate tissue residency in memory T cells [239]. Therefore, while CD69 and CD103 expression do not alone determine tissue residency, stable expression of CD69⁺CD103^{+/-} remains the most effective way to phenotypically distinguish TRM from their circulating counterparts.

TRM also express other molecules on their surface that distinguish them from circulating T cells. For instance, mouse and human CD4⁺ and CD8⁺ TRM express the integrin CD49a in the lung and skin [8, 181, 191, 233, 240]. CD49a (integrin α 1) binds to CD29 (integrin β 1) to form VLA-1, an integrin specific to collagen [241] suggesting that CD49a may play an important role in adhesion of TRM near collagen-rich basement membrane of the epithelium. Accordingly, antibody blockade and genetic deletion of VLA-1 in mice results in impaired retention of CD8⁺ TRM in peripheral tissues [242]. CD49a expression may also delineate subsets of human TRM, particularly in skin as CD49a⁺CD8⁺TRM cells produced IFN- γ while CD49a⁻TRM cells

produced IL-17 [240]. CXCR6, a chemokine receptor which binds CXCL16, is another commonly identified marker of TRM in both human tissues [8] and in the mouse skin and liver, where it plays an essential role in the formation of CD8⁺TRM [234, 243]. We recently reported that most CD4⁺ and CD8⁺ TRM in human tissues commonly express the inhibitory receptors CD101 and PD-1 [8]. Relatively little is known about CD101 including its binding partner, though a mechanistic study in human T cells found that antibody stimulation of CD101 prevents CD3-induced T cell proliferation by interfering with IL-2 production [244].

Upon chronic stimulation, T cells can become exhausted and no longer responsive to antigen. T cell exhaustion has been well characterized in mouse models has a major role in failure to control chronic infection and neoplasia. Exhausted T cells have unique expression and cell phenotypes including high expression of PD-1, an inhibitory molecular on T cells [245, 246]. Further, the molecular underpinnings of PD-1 expression have been found to be negatively regulated by transcription factor T-bet, a master regulator of effector function[247]. However, expression of PD-1 alone is not enough to indicate T cell exhaustion, and further, T cells in tissues that are not exhausted also express PD-1, indicated its immunomodulatory roles beyond T cell exhaustion [248]. More specifically, PD-1 expression has been described in CD8⁺TRM generated through infection in mice [249, 250], and also in human multiple tissues at steady-state including in the lungs, spleen, tonsils and liver [8, 193, 196]. Interestingly, TRM expressing these inhibitory molecules mediate robust and rapid effector functions following TCR stimulation [8, 12, 193]. Additionally, PD-1 treated T cells can respond to immunotherapy and provide proliferative bursts to curtail viral infection. Additionally, human studies have shown efficiency in anti-PD1 therapies that reinvigorate T cells within tumors for clearance.[251] Constitutive expression of these inhibitory receptors on TRM may serve to fine-tune or self-

regulate TRM responses to optimize protective responses, while minimizing tissue pathology *in situ*.

Transcriptional regulation of tissue residency programs

The designation of TRM as a distinct subset with a unique transcriptional profile, raised the important question concerning the identity of the transcription factor(s) which drive TRM formation. Several transcription factors including Hobit, Blimp, Runx3 and Notch were found in mouse infection models to promote CD8⁺ TRM formation [196, 252-254]. These transcription factors (TFs) were identified based on upregulated expression of the TF or transcriptional targets in TRM compared to circulating memory T cells. Whether these TF function in TRM formation for CD4⁺T cells or for human T cells remains yet to be established. Hobit transcripts are upregulated in human T cells after activation [255] and are expressed at very low levels in resting human TRM cells, albeit with upregulated expression in lung CD69⁺ compared to CD69⁻ memory T cells [8, 196, 227]. Notch/RBPj is upregulated in human TRM and is part of the core gene signature [8, 196], although expression of Runx3 in human TRM has not yet been reported. Thus, the transcription factors driving human TRM formation remain unclear and whether mouse TF function similarly in human T cells has not yet been demonstrated. It is possible that TRM-driving TF are upregulated early in TRM formation, while in human tissue, TRM may be persisting for years or more and therefore TF expressed in human TRM may be reflecting a maintenance rather than development requirement. Ultimately, modulation of human TRM may need to target factors required for TRM maintenance or tissue retention such as TGF- β , produced within tissues shown to induce CD103 expression and promote tissue retention [210].

Mucosal Tissue TRM populations

In human skin, the majority of dermal TRM are CD4⁺CD69⁺CD103⁻ cells, with the epidermis containing mixed populations of CD103⁺CD4⁺ and CD8⁺TRM [238]. These TRM express high levels of the cutaneous lymphocyte antigen (CLA) and specific chemokine receptors such as CCR4 and a proportion upregulate the chemokine receptor CCR8 upon entry into the human skin niche [256, 257]. Stromal keratinocytes play key roles in the formation and maintenance of both CD4⁺ and CD8⁺ skin TRM through the activation of latent TGF- β that facilitates the upregulation of CD103 [238, 258]. Keratinocytes also produce IL-7 and IL-15 that is required for long-term maintenance and survival of memory T cells [259]. A recent study also found that skin TRM may alter their metabolism within specific sites; mouse dermal CD8⁺ TRM acquire a program of exogenous lipid uptake and increased oxidative metabolism to persist in the skin niche and mediate immunity [260]. Whether this metabolic reprogramming of TRM is a general mechanism for tissue adaptation or specific to lipid-rich epithelial tissues remains to be determined.

The lung is another mucosal site with abundant populations of CD4⁺ and CD8⁺ TRM identified in humans and mice [6, 179, 189, 191, 261]. Lung parenchyma and airway CD8⁺ TRM express either CD69⁺ CD103⁻ or CD69⁺ CD103⁺, while the majority of CD4⁺ TRM are CD69⁺ only. Similar to the skin, optimal formation of CD8⁺ TRM in the mouse lungs needs cross-priming DC within the local lymph node [262], but, also requires 41BB signaling and CD4⁺ T cell help via the production IFN- γ [263, 264]. Once in the lung niche, CD103⁺ dendritic cells facilitate CD103 upregulation and maintenance of CD8⁺ TRM through the production of TGF- β [265], and the long-term survival these CD103⁺ cells requires IL-15 [183]. In the lung, there is differential localization of CD4⁺ and CD8⁺ T cells that display unique clustering patterns within the tissue; CD8⁺ TRM localize within specific niches of tissue regeneration after lung injury

(termed repair-associated memory depots) that aid their formation and maintenance, while CD4⁺ TRM localize to the airways or around B cell follicles [217, 261].

Lymphoid TRM populations

Bone marrow is a prominent niche for populations of CD4⁺ and CD8⁺ memory T cells in both humans [266-268] and mice [188, 269, 270]. Human bone marrow contains significantly more TEM cells and fewer Naïve T cells compared to peripheral blood, suggesting increased maintenance of memory in the bone marrow [266-268]. Furthermore, the bone marrow CD4:CD8 ratio is about 1:2, in contrast with peripheral blood and lymph node sites which are closer to 2:1. [271] Interestingly, populations of bone marrow memory T cells also express canonical TRM marker CD69. In mice, it is a minority of memory (CD44^{hi}) T cells expressing CD69 (between 10-30%)[272], however in humans, a larger percentage of memory T cells in the bone marrow can express CD69 (between 40-60%), with higher expression on CD8⁺ T cell memory [266]. Expression of CD69 was initially thought of as indicating BM T cells are at an increased activation state compared to those in blood, due to simultaneous increased 4-1BB expression in CD4⁺ and CD8⁺ T cells and increased CD25 expression in CD8⁺ T cells [266]. However, follow up studies revealed that CD69⁺ T cells in human bone marrow are actually more quiescent rather than activated in phenotype, shown by decreased expression of Ki-67, a marker of active proliferation, as well as gene profiling showing a global resting gene expression signature of ex-vivo CD69⁺ compared to CD69⁻ CD4⁺ T cells from the bone marrow and T cells from peripheral blood [223, 273]. Genetic knockout studies in mice have shown that in CD69 knock-out animals, there is decreased accumulation of CD4⁺ T cells in the BM compared to WT counterparts, therefore suggesting that CD69 plays a role in retention of T cells in bone marrow niche [272].

T cells in the bone marrow also differ from blood T cells in that they express chemokine receptors such as CCR5 and CXCR4 which may aid in their retention within the bone marrow niche [268]. Despite the tissue-specific expression and possible mechanisms of maintenance within the bone marrow, evidence in mice suggests that bone marrow T cells are highly circulating and the markers of tissue residence may instead denote a temporary residence in bone marrow for homeostatic survival signals via IL-15 signaling [266, 270, 274]. The evidence for bone marrow memory being highly circulating is largely from experiments in mice by parabiosis showing that the ratio of host and donor T cells equilibrates to one after about two weeks [215], however whether this is indicative of T cell dynamics in humans remains unknown. Moreover, given mice have fewer putative CD69⁺ resident memory T cells than humans, perhaps there are increased levels of functional maintenance of T cells in human bone marrow compared to mice. Taken together, the bone marrow is a niche for memory T cell maintenance, however the extent to which any of the T cells are truly resident remains unknown as well the functional role of this niche for T cells in humans.

Recent studies in secondary lymphoid tissues of mice and humans demonstrate that CD4⁺ and CD8⁺ TRM also take up residence in LNs and spleen [7, 275-277]. In conventional inbred mouse models, the majority of virus-specific T cells within lymphoid tissue are circulating with <10% truly resident by parabiosis [189]. Compared to mice, a much greater proportion of CD4⁺ and CD8⁺ T cells in lymphoid tissues express phenotypes and transcriptional profiles of TRM cells [7, 8]. Interestingly, dirty mice possess a similarly increased number of TRM in their lymph nodes relative to conventional SPF-mice [278], suggesting that the greater levels of antigen exposure experienced by humans may relate to the size of the lymph node TRM pool. Much less is known about lymphoid TRM, possibly because these are a rare population within SPF mice, yet in humans they may play a big role. Future studies are required to better understand

populations of TRM within lymphoid tissues and is the focus of part of this dissertation work. TRM in humans specific for CMV and EBV have been found in tonsillar tissue [194]. Taken together, TRM in human lymphoid tissues are playing emerging roles, perhaps first overlooked due to their rarity in mouse models.

Maintenance of TRM populations over time

Once generated, the requirements and potential of TRM for long-term maintenance within the tissue are an active area of investigation in mouse models, and difficult to assess in humans. In the lung, IL-2 and IL-15 are required for CD4⁺ TRM formation [200], and given the known role of IL-15 as a homeostatic cytokine for peripheral memory T cell maintenance [279], it is possible that TRM use IL-15 for maintenance. Whether TRM can be maintained in tissues long-term is an area of debate: in the lung, CD8⁺ TRM generated from influenza virus infection were found to diminish over time due to apoptosis and lack of replenishment from circulating TEM cells [280, 281]. Lung CD4⁺ TRM generated from influenza did persist for months after infection and persistence did not require replenishment from circulation over the short-term [190], suggesting differential maintenance requirements for CD8⁺ and CD4⁺ TRM.

Although it is difficult to assess the persistence of human TRM in vivo and mechanisms for homeostatic turnover, there is evidence that TRM in certain human sites are long-lived. First, the overall proportion of TRM is set quite early in life, and this proportion is maintained at constant frequencies for decades of life—and into old age as assessed by extensive flow cytometry analysis of tissue memory T cells from key sites such as the lung and intestine from infancy through childhood and into the ninth decade of life [9, 190]. While this analysis does not assess antigen-specific populations, on the whole tissue level, there seem to be mechanisms for overall homeostatic maintenance of a stable population of TRM. In the human lung, influenza-

specific T cells can be readily detected [190, 282], although the timing of the initial infection that generated them cannot be determined. TRM in the lung and other human tissues exhibit lower frequencies of Ki67⁺ cells indicative of proliferating cells, compared circulating T cells in blood and other sites [190], suggesting lower rates of turnover of tissue compared to circulating T cells. In human skin, there is evidence, however, for long-term maintenance of antigen-specific TRM, based on studies of cutaneous T cell lymphoma (CTCL) [192], and skin lesions in psoriasis, an inflammatory skin condition. Lesions and flare-ups are known to occur occurs at the same location over years, and this is associated with the presence of clonal populations of skin TRM that produce IL-17 [283].

Aging is associated with decline in the ability to mount an immune response. This decline is associated with changes in various cell types of the immune system and the most defined changes are attributed to T lymphocytes which coordinate adaptive immunity. In general, there is a marked decline in T cell functionality associated with diminished responses to vaccines and infections, and immune dysregulation associated with increased autoimmunity in the elderly. As discussed earlier, T cell subsets are compartmentalized in tissues similar to adults with mucosal sites having predominant TEM populations, lymphoid tissue comprising both naïve and TEM cells, TCM cells found only among the CD4 subset in lymphoid sites and blood, and TEMRA cells found only among CD8 T cells in blood, spleen and lung, representing a blood-borne, circulatory subset [7]. Overall, this compartmentalization of memory T cells in mucosal sites, naïve T cells in lymphoid tissue, with blood containing all four subsets is maintained through many decades of adult life with a decrease in the frequency of naïve T cells being the major change with age [7]. These data indicate that the decline with age is not due to subset identity necessarily but maybe the quality of T cell responses with age. Further, effector memory CD8⁺ T cells isolated from older adults exhibited diminished proliferative responses and cytolytic

activity in response to *ex vivo* influenza A/H3N2 challenge [284]. By contrast, the phenotypic and functional characteristics of effector CD4⁺ T cells responding to influenza in older adults are preserved, suggesting a compensatory response to influenza infection when CD8⁺ T cells become compromised during immune aging [285]. The incidence of cancer increases rapidly with age and is hypothesized to be a result of the accumulation of somatic mutations over life. However, a new mathematical model, created by incorporating the decline in T cell production with age outperforms previous models based on somatic mutations.[286] Together, these findings show that age related decline in T cell function is a major risk factor for disease.

Lastly, in mouse models of influenza infection that generate TRM cells, it's been shown that TRM generation is significantly impaired in mice infected as infants [287]. The defect in ability to form TRM results in excess of effector cell at the expense of TRM generation. Further this increased effector generation is due to over-expression of T-bet, and T-bet heterozygote mice are better able to form TRM as infants more similar to adults [287, 288]. In conclusion, both infants and the elderly have impaired immune responses which may be due to inability to generate and maintain effective T cell responses.

TRM development by vaccination

The ultimate goal in translating the fundamental knowledge of TRM revealed in mouse models to humans is to develop novel vaccines for promoting protective immunity. Accordingly, preclinical mouse models of vaccination and infection have shown promising outcomes when targeting TRM responses. Intranasal administration of a live-attenuated IAV (LAIV) vaccine generated long term virus-specific CD4⁺ and CD8⁺ TRM in the lungs of mice, which mediated heterosubtypic protection independent of circulating memory T cells or neutralizing antibodies [182]. Importantly, intraperitoneal injection of inactivated virus or LAIV failed to generate TRM

responses and provide cross-strain protection [182], demonstrating that both route and vaccine formulation (i.e., live-attenuated virus) are key determinants for TRM formation. Vaccination with Bacille Calmette-Guérin (BCG) in mice similarly demonstrated that mucosal but not subcutaneous administration of BCG generated protective TRM in the airway [289]. Delivery of vaccine vectors to specific tissues has also proved successful in inducing protective TRM immune responses including those using IAV vectors expressing HIV antigens and HPV pseudovirus encoding antigens from HSV [290, 291] and RSV[292]. These results emphasize that the current immunization approaches administering vaccines peritoneally may be less effective in generating optimal protection compared to methods targeting sites where pathogens infect.

Another immunization approach to generate TRM, designated “prime and pull”, combines vaccination (prime) with local administration of chemokines or adjuvants to recruit TRM precursors to target tissues (pull). Subcutaneous immunization with an attenuated strain of HSV-2 coupled with topically applied chemokines to the vaginal mucosa generated virus-specific CD8⁺ TRM cells that protected mice from lethal HSV-2 challenge in the FRT [198, 217]. Variations of this approach combine prime and pull into a single inoculum, such as using antigen complexed to antibodies targeting tissue-specific DC populations [265]. A particularly successful strategy in mice used a hepatocyte-specific adenovirus expressing malaria antigens to target TRM formation in the liver and prevent liver-stage malaria, which has now progressed to phase I clinical trials in humans [293].

Together, these studies provide promising proof-of-principle results that protective TRM can be generated by vaccination. Whether TRM-based vaccines can be applied to humans to prevent acute and/or chronic infections, will likely be determined in the coming years. Therapeutically generated TRM must also conform to the homeostatic balance of tolerance and

effector responses within tissues to prevent immunopathology. The capacity of individual human tissues to generate and maintain TRM and the minimum threshold of TRM to provide protection are essential questions that will inform optimization of protective immunity in the next generation of vaccines.

Outstanding questions in field of tissue localized responses

Overall, studies in mouse and humans indicate that modulation of T cell responses in tissues can have great impact on many human pathologies including infections, therapies against cancer, vaccines and more. However, there are still key questions about TRM biology that remain not well understood. For instance, what is the developmental relationship between T cell subsets in blood and those within tissue sites? Additionally, much of the work on TRM has been derived from mouse models and specifically on TRM in barrier sites, and less is known about TRM in lymphoid tissues. In fact before the start of my thesis work, resident memory T cells in lymph nodes were not identified in humans, however now they are emerging as prominent populations. More specifically, given there are high levels of T cells expressing CD69, a marker of TRM, in lymphoid tissues in humans, are these cells similar to canonical TRM as defined in mouse models? TRM biology is further complicated by the fact that different tissue sites may contain T cells of different specificities. Additionally, humans are exposed to many pathogens and antigens at once, and how this shapes T cell response in tissue is unknown. How does exposure to myriad of antigens shape and influence T cell populations on the clonal level? Are properties of T cells on the clonal level more influenced by subset identity (TCM, TEM, TEMRA, and TRM) or tissue-specific features? What role does antigen play in generation of TRM populations, how different are TRM populations in different tissues, compared to those in blood? The use of T cell receptor sequencing methods and molecular profiling techniques as well

as access to human organ donor tissue has given us an unprecedented opportunity to investigate these questions.

Section 1.6 Thesis Objectives

Memory T cell differentiation, function, and maintenance occurs at diverse tissue sites. T cells travel via blood and lymph in order to enter into peripheral sites of inflammation including non-lymphoid barrier sites and several secondary lymphoid sites where T cells coordinate adaptive immunity. Lymphoid and non-lymphoid sites contain an estimated >95% of the total T cells in humans, yet most knowledge on human T cell differentiation and maintenance derives from extrapolation of studies on peripheral blood. Initially, two human memory T cell subsets, named TCM and TEM, were discovered in blood based on differential functional potential and homing to lymphoid and non-lymphoid sites respectively. More recently in mice, TRM cells, defined by ability to remain resident in non-lymphoid and lymphoid tissue sites have been found to play important roles in adaptive immunity. There are many key gaps in knowledge about human TRM biology. **Our goals are to understand how human T cells are localized in lymphoid sites and what regulates the maintenance of T cell subsets in lymphoid and non-lymphoid sites. This body of work investigates the hypothesis that there are tissue-specific mechanisms for tissue compartmentalization of human memory T cells.** With a better understanding of the healthy immune system in humans, we can better understand what goes wrong in disease.

Our first main objective is to investigate differentiation and mechanisms for maintenance of lymphoid memory T cells in BM, spleen and lymph nodes. Evidence from mouse models show that a small percentage of T cells within lymphoid tissue (<10%) establish residence, as shown by parabiosis studies. This is in contrast to T cells in non-lymphoid sites such as skin, intestine, and salivary glands of which nearly all are resident by the same parabiosis experiments [189]. In contrast, human data show that a large majority of memory T cells within lymphoid tissue express TRM marker CD69, indicating there may be increased amounts of TRM in human

lymphoid tissue compared to mouse models. *Therefore, our first objective is to examine memory T cells in lymph nodes, bone marrow, and spleen and investigate their properties and functions; and second to identify what, if any, tissue specific mechanisms are playing a role in their maintenance.*

Results from our first objective reveal a novel organ-specific, memory T cell subset that is maintained in LN but not in blood and other lymphoid or mucosal tissues. Interestingly, both LN TEM and TRM displayed a unique profile of protein and RNA expression that was distinct from those in other sites including expression of transcription factors Tcf-1 and Lef-1 known to be important for self-renewal and TFH differentiation [141, 143-145, 150-152, 154, 157, 170], and expression of chemokine receptors CXCR4 and CXCR5 that have roles in regulating lymphoid homing and migration to B-cell follicles [294]. The specific properties of LN TEM and TRM as more quiescent, having diverse receptor repertoires, and having high proliferative and differentiation capacities indicate that LN provide distinct niches for maintaining high capacity T cells poised for protective responses. Their maintenance throughout many decades of life, as shown here, provides immune reserves for eliciting protective responses and controlling chronic infections, and provide a mechanism for our previous findings of LN-restricted localization of CMV-specific T cells in certain individuals [10]. Finally, we identify type-1 IFN signaling as an opposing mechanism for TCF-1 maintenance in lymph node sites. Together these findings suggest that LN provide a specialized niche for maintenance of TEM and TRM over life.

Our second main objective is to investigate the relatedness of human effector and memory T cells in tissues. In mice, effector cells are widely disseminated during a primary infection, whilst short lived, and memory populations remain long-lived in specific tissue sites depending on localization of initial pathogen encounter and sites of inflammation. In humans, four major human CD4⁺ and CD8⁺ T cell subsets that have been identified including TEM,

TCM, TRM, and TEMRA cells, however how antigen has driven differentiation and compartmentalization of these subsets in human lymphoid and non-lymphoid tissues remains unknown. Sequencing of the TCR- β chain allows identification of T cell clones in order to investigate the relatedness of T cell subsets across tissues. *Our second objective is to investigate our hypothesis that tissue site has a major influence on the diversity of T cell clone repertoires and that a given T cell clone may be found widely disseminated in multiple sites in order to provide protection to many on-going systemic pathogens.*

Our results investigating our second objective reveal that the relatedness of T cells between sites, including blood, bone marrow, spleen, lymph nodes, and lung, is highly subset specific. From diversity analysis, we found a hierarchy of T cell diversity; it was highest among CD4⁺TCM, then CD4⁺ and CD8⁺ TEM and TRM, and lowest among CD8⁺TEMRA cells, a finding mostly conserved across tissues. We looked for tissue-specific features and found that CD4⁺ and CD8⁺ T cells in LN are more diverse than in BM. From our clonal overlap analysis, we found high degree of overlap between CD8⁺TEMRA cell populations across tissues; with few tissue-specific clones. Clonal overlap analysis across tissues revealed tissue-specific patterns and maintenance of TEM and TRM clones, including a high degree of similarity between CD4⁺TEM and CD4⁺TRM in lung, and a high degree of similarity between CD4⁺TEM in BM and CD4⁺TEM in Spl. Overall, our findings show that the relatedness of T cell clones is influenced mainly by subset and to a lesser degree, anatomical site. These results show how T cell responses are distributed in tissues and can serve as a baseline for future studies investigating T cell responses in individuals with pathologies such as infections, autoimmunity and cancer.

CHAPTER 2: Materials and Methods

Section 2.1 Human tissue acquisition and lymphocyte isolation

Acquisition of human samples.

Human tissues were obtained from deceased (brain dead) organ donors at the time of organ acquisition for life-saving clinical transplantation. Donor tissues were obtained through an approved protocol with LiveOnNY. Organ donors were free of chronic disease and cancer, negative for hepatitis B, hepatitis C and HIV. The tissues collected for research include: blood, bone marrow, lung, lymph nodes (lung-, mesenteric-, and iliac-draining), intestinal sites (ileum, jejunum, and colon), spleen, tonsils, and salivary glands. The tissues collected are depicted graphically in Figure 2-1. For isolation of blood from living volunteers, blood was drawn via venipuncture from consented volunteers, as approved by the Columbia University IRB.

Isolation of lymphocytes from human samples.

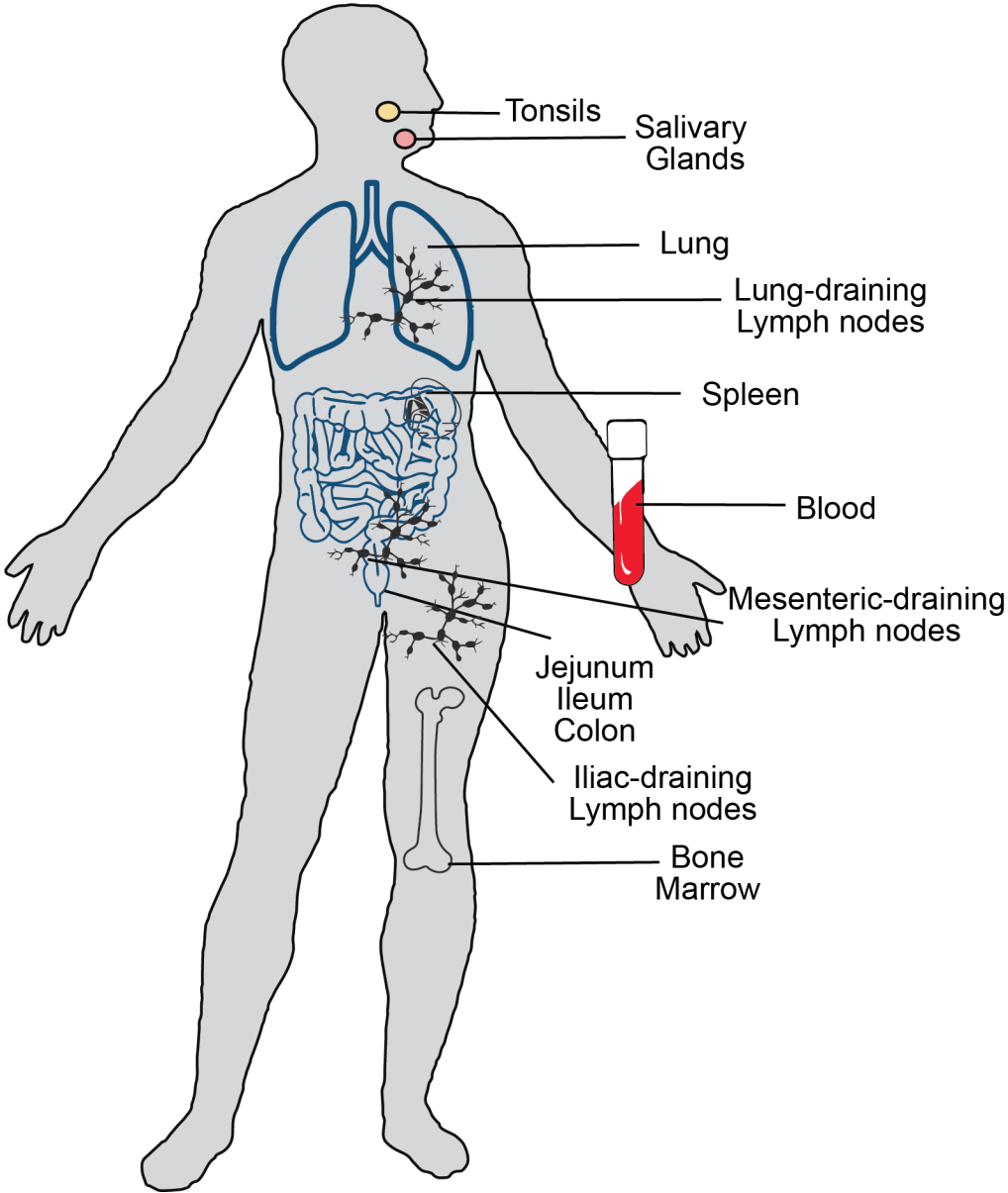
Tissue samples were maintained in cold saline and brought to the laboratory within 2-4h of procurement. Samples were rapidly processed using enzymatic and mechanical digestion to obtain lymphocytes with high viability. For spleen, lymph nodes, and mucosal tissues (including lung), tissues were minced and incubated at 37°C in enzymatic digestion media: RPMI (Thermo Fisher) containing 10% fetal bovine serum (FBS) (Thermo Fisher), L-glutamate (Thermo Fisher), sodium pyruvate (Thermo Fisher), nonessential amino acids (Thermo Fisher), penicillin-streptomycin (Thermo Fisher), collagenase D (1 mg/ml, Roche), trypsin inhibitor (1 mg/ml, Thermo Fisher) and DNase I (0.1 mg/ml, Roche). For the lung tissue, prior to incubation in digestion media, the tissue was inflated with digestion media using a 50cc syringe and 18g

precision needle. The incubation in digestion media times varied for each tissue; lymph nodes were incubated for 1 hour and spleen and lung were incubated for 1.5 hours. Digested tissue was further disrupted using the gentleMACS tissue dissociator (Miltenyi Biotech); the resulting suspension was passed through a tissue sieve (10–150 mesh size) and then pelleted through centrifugation. For only the spleen, red blood cells (RBC) were lysed using ACK lysis buffer (Corning Cellgro) by incubation for 5 minutes on ice. Following a wash with RPMI containing 2% FBS, dead cells and epithelial cells were removed via centrifugation through 30% Percoll (GE Healthcare Life Sciences). The settings for density separation using Percoll were 50g and breaks set at 0 (zero) and acceleration set at 3 for 15 minutes at room temperature. After centrifugation with Percoll, mononuclear cells that were below the top layer of epithelial cells were rescued and washed in RPMI media containing 2% FBS. To remove residual RBCs, RBCs were lysed using ACK lysis buffer by incubation for 5 minutes on ice. This was the second ACK lysis step for spleen, and is usually required for lung but not lymph nodes. Resulting cell suspensions were resuspended in RPMI containing 2% FBS. Finally, cells were passed through 100 μ m filter paper. Resulting cell suspensions were kept on ice until downstream use. Total processing time for lymphocyte isolation from each tissue is about 2.5 hours for spleen, 3.5 hours for lung, and 1 hour for lymph nodes.

For isolation of lymphocytes from blood and bone marrow, lymphocyte separation media (Cellgro) was used to perform density gradient centrifugation to isolate the mononuclear cells from the RBCs. The settings for the lymphocyte separation centrifugation were 400g and breaks set at 0 and acceleration at 0 for 20 minutes at room temperature. The mononuclear cell layer positioned between the plasma (top layer) and LSM (bottom layer) was extracted and washed with RMPI. If residual RBC contamination was present then cells were incubated in 5mL of AKC lysis buffer for 5 minutes at 4°C, followed by a wash and resuspension in RPMI

containing 2% FBS. Finally, cells were passed through 100 μm filter paper. Resulting cell suspensions were resuspended in RPMI containing 2% FBS and kept on ice until downstream use. Total processing time for mononuclear cell isolation from blood and bone marrow is about 1 hour depending on volume of blood and number of tubes used for processing.

Figure 2-1. Human tissues acquired for research.



Cryopreservation of lymphocytes

Single cell suspensions of lymphocytes were cryopreserved in Cryogenic Vials (Corning) in a solution of 10% dimethyl sulfoxide (DMSO) in FBS at a concentration of 10 million cells/mL. Vials containing the cell suspension were placed into Mr. Frosty Cryo Freezing Containers (Nalgene) and placed in the -80°C freezer. The “Mr. Frosty” containers are filled with a basin of ethyl alcohol, which allows for a cooling rate of 1°C/minute. After cooling to -80°C, vials containing frozen cell suspensions were transferred to liquid nitrogen freezer. The frozen suspensions can be stored for years and currently the Farber lab has cryopreserved lymphocyte suspensions from multiple sites from over 400 donors.

To thaw cryopreserved samples, vials are removed from liquid nitrogen and transferred to 37°C water bath until cell suspensions are fully thawed. Immediately upon thawing, pre-warmed RPMI containing 10% FBS added slowly according to a protocol shown to increase both T cell viability and functionality post-cryopreservation [295]. Initially 1 mL of media (to double the volume of cell suspension) was added with 0.1mg/mL DNase to minimize cell clumping of cell suspensions and allowed to sit for 1 minute. Subsequently, additional warmed RPMI containing 10% FBS is added one drop at a time until cell suspension solution is 10X in volume. Cells were spun down and resuspended in RPMI containing 10% FBS. Finally, cell suspensions were passed through 100 µm filter paper and cells are ready for downstream use.

Section 2.2 Flow Cytometry

Flow cytometry staining, acquisition, and analysis

For flow cytometric analysis, single cell suspensions were stained with fluorochrome conjugated antibodies for 30 minutes at 4°C in the dark. For the last 5 minutes of staining, DAPI was added for identification of dead cells (Biolegend). Staining and washing of cells was performed in FACS buffer (PBS/1% fetal bovine serum/0.1% sodium azide). Control samples included unstained, single fluorochrome–stained compensation beads (Ultra Comp eBeads, eBioscience) and fluorescence minus one (FMO) controls. Stained cells were acquired using a BD LSR II or BD Fortessa analytical flow cytometer in the CCTI flow cytometry core and analyzed using FlowJo (Treestar) and FCS Express (De Novo Software).

For intracellular staining, cells were stained with a fixable live-dead stain during the surface staining (Thermo Fisher). For detection of cytokine and transcription factor expression, cells were incubated in perm-fix buffer (eBioscience) for 1 hour, washed, resuspended in permeabilization buffer (eBioscience) with antibodies specific for intracellular markers for 1 hour at 20-25°C.

For isolation of subsets by fluorescent-activated cell sorting, lymphocyte suspensions were first enriched for T cells using a magnet based CD3 negative enrichment kit, MojoSort Human CD3 T cell Isolation Kit (Biolegend), stained for surface markers (as described above) in sort buffer (PBS/1% FBS), and sorted using an Influx high-speed cell sorter or Aria II cell sorter (both from BD Biosciences). For a list of antibodies and reagents used see Table 2-1

Section 2.3 CyTOF

CyTOF Sample Prep

Cryopreserved cell suspensions from each tissue were thawed and labeled with Rh103 intercalator as a viability marker. For each donor, cells from each tissue were first barcoded using a unique combinatorial barcode of CD45 antibodies conjugated with monoisotopic cisplatin and then pooled. The pooled tissue samples were then stained with a panel of antibodies against cell surface markers, washed, fixed and permeabilized (eBioscience Transcription Factor Staining Kit) and then stained with additional antibodies against intracellular targets. CyTOF antibodies were either purchased pre-conjugated from Fluidigm (formerly DVS Sciences) or purchased purified and conjugated in-house using MaxPar X8 Polymer Kits (Fluidigm) according to the manufacturer's instructions. For a complete list of antibodies used see Table 2-2. The samples were then washed and incubated in 0.125nM Ir intercalator (Fluidigm) diluted in PBS containing 2% formaldehyde, and stored at 4°C until acquisition.

Sample Acquisition and Analysis

Immediately prior to acquisition, samples were washed once with PBS, once with de-ionized water and then resuspended at a concentration of 1 million cells/ml in deionized water containing a 1/20 dilution of EQ 4 Element Beads (Fluidigm). The samples were acquired on a CyTOF2 (Fluidigm) equipped with a SuperSampler fluidics system (Victorian Airships) at an event rate of <500events/second. After acquisition, the data were normalized using bead-based normalization in the CyTOF software and uploaded to Cytobank for initial data processing.

For analysis, the data were gated to exclude residual normalization beads, debris, dead cells and doublets, and the derived from each tissue were deconvolved by Boolean gating on CD45

barcodes, leaving DNA⁺CD45⁺Rh103⁻ events for subsequent clustering and high dimensional analyses.

FCS express software (De Novo) was used for generating tSNE (t-Distributed Stochastic Neighbor Embedding) and PCA plots from CyTOF data[296, 297]. To generate tSNE plots, we used the Barnes-Hut implementation of the t-SNE algorithm called viSNE [298]. One tSNE analysis was generated for three individuals (D332, D333, D335) together. A separate tSNE analysis was generated for a fourth individual (D342) in which the panel lacked TCF-1 and included blood as an additional sampling site. Downstream statistical analysis was done using R programming language. Heatmaps of normalized marker expression were generated using R. We plotted the heatmap using the Z-score of average marker expression with samples clustered by unsupervised hierarchical clustering function *hclust* and visualized with *heatmap.2* in the *gplots* package. We performed multidimensional scaling (MDS) analysis using the *cmdscale* function in R on a distance matrix computed with *dist* function in R. MDS was plotted using *ggplot*.

Table 2-1. Antibody and reagents used for flow cytometry.

Reactivity/ Reagent	Target	Fluorochrome	Clone/target	Company
Human	CCR7	AF488	G043H7	Biologend
Human	CCR7	PE-Texas Red	G043H7	Biologend
Human	CD45RA	BV605	hi100	Biologend
Human	TCF1	PE	C63D9	Cell Signaling
Human	CD3	BV510	OKT3	Biologend
Human	CD4	PECy7	SK3	Biologend
Human	CD4	APCCy7	OKT4	Biologend
Human	CD4	BV650	SK3	Biologend
Human	CD8	APCCy7	SK1	BD Biosciences
Human	CD8	PE	HIT8a	Biologend
Human	CD8	BUV737	SK1	BD Biosciences
Human	CD8	PerCPCy5.5	RPA-T8	Tonbo
Human	CD45RO	PerCP-eFluor710	UCHL1	Ebiosciences
Human	CD69	BV421	fn50	Biologend
Human	CD69	APC	FN50	Biologend
Human	CD57	PE-Dazzle-964	HNK-1	Biologend
Human	Ki67	BV421	B-56	BD Biosciences
Human	Lef1	FITC	C12A5	Cell Signaling
Human	Peforin	PE-CF594	GG9	BD Biosciences
Human	Tbet	PE	4B10	Biologend
Human	Eomes	PE-Eflour610	WD1928	Ebiosciences
Human	CCR5	AF700	J418F1	Biologend
Human	CXCR4	APC	12G5	Biologend
Human	PD1	AF647	EH12.1	BD Biosciences
CMV-Multimer		APC	A2402, pp65 ₁₁₃₋₁₂₁	Proimmune
CMV-Multimer		APC	A0201, pp65 ₄₉₅₋₅₀₃	Proimmune
LIVE/DEAD™ Dead Cell Stain		Fixable Red	-	ThermoFisher
LIVE/DEAD™ Dead Cell Stain		Fixable Blue	-	ThermoFisher
Cell proliferation dye		efLuo450	-	eBiosciences

Table 2-2. Antibodies used for CyTOF

Target	Tag	Clone	Source	Source Cat#
CD57	113 In	HCD57	Biolegend	322302
CD28	141 Pr	CD28.2	Biolegend	302902
CD19	142 Nd	HIB19	Biolegend	302202
CD45RA	143 Nd	HI100	Biolegend	304102
CD103	144 Nd	Ber-Act8	Biolegend	350202
CD4	145 Nd	RPA-T4	Biolegend	300502
CD8a	146 Nd	RPA-T8	Biolegend	301002
Perforin	147 Sm	dG9	Biolegend	308102
CD16	148 Nd	3G8	Biolegend	302014
CD127	149 Sm	A019D5	Biolegend	351302
CD1c	150 Nd	L161	Biolegend	331502
CD123	151 Eu	6H6	Biolegend	306002
CD66b	152 Sm	G10F5	Biolegend	305102
PD-1	153 Eu	EH12.2H7	Biolegend	329912
ICOS	154 Sm	C398.4A	Biolegend	313502
CD27	155 Gd	O323	Biolegend	302802
CCR5	156 Gd	NP-6G4	Fluidigm	3156015A
Bcl6	158 Gd	IG191E/A8	Biolegend	648302
Tcf1	159 Tb	7F11A10	Biolegend	655202
CD14	160 Gd	M5E2	Biolegend	301810
CD56	161 Dy	B159	BD Biosciences	555513
CXCR5	163 Dy	J252D4	Biolegend	356902
CD69	164 Dy	FN50	Biolegend	310902
41BB	165 Ho	4B4-1	Biolegend	309802
CD25	166 Er	M-A251	Biolegend	356102
CCR7	167 Er	G043H7	Biolegend	353222
CD3	168 Er	UCHT1	Biolegend	300402
Tbet	169 Tm	4B10	Biolegend	644802
CD38	170 Er	HB-7	Biolegend	356602
CD95	171 Yb	DX2	Biolegend	305602
LAG3	172 Yb	11C3C65	Biolegend	369302
CXCR4	173 Yb	12G5	Fluidigm	3173001B
HLADR	174 Yb	L243	Biolegend	307602
TIGIT	175 Lu	MBSA43	eBioscience	16-9500-85
GranzymeB	176 Yb	GB11	Invitrogen	MA1-80734

Section 2.4 Whole transcriptome profiling by RNA sequencing

Isolation of RNA

T cells were purified by cell sorting and total RNA was isolated using an RNA/DNA kit (AllPrep DNA/RNA mini kit, QIAGEN) according to manufacturer's protocol. Briefly, sorted cells were pelleted, resuspend in RLT plus buffer and put through QIAshredder (Qiagen) columns to completely lyse cells. DNA and RNA were extracted using RNA/DNA columns by step-wise process outlined specifically in the manufacturer's manual. The final step in the RNA extraction protocol is to elute the purified RNA from the column. Next, we assessed RNA concentration and quality by RNA Integrity Number (RIN) using an Agilent 2100 Bioanalyzer (Agilent Technologies). For the majority of samples, >200ng of total RNA was submitted for sequencing with RIN>8.

RNA-sequencing pipeline

Library preparation and whole transcriptome profiling by RNA-sequencing was performed by the Columbia Genome Center. Library preparation was done using a poly-A pull-down to enrich mRNAs from total RNA samples followed by library preparation using Illumina TruSeq RNA prep kit. Libraries were sequenced using Illumina HiSeq2000 at Columbia Genome Center. Samples were multiplexed in each lane, which yields targeted number of paired-end 100bp reads for each sample, as a fraction of 180 million reads for the whole lane.

For base called, RTA (Illumina) was used and bcl2fastq (version 1.8.4) for converting BCL to fastq format, coupled with adapter trimming. Reads were mapped to a reference genome (Human: NCBI/build37.2) using Tophat (version 2.1.0) [299] with 4 mismatches (--read-mismatches = 4) and 10 maximum multiple hits (--max-multihits = 10). To map reads derived

from exon-exon junctions, Tophat infers novel exon-exon junctions ab initio, and combines them with junctions from known mRNA sequences (refgenes) as the reference annotation. The relative abundance (aka expression level) of genes and splice isoforms was estimated using cufflinks [300] (version 2.0.2) with default settings. Data are available at GEO accession # GSE106420.

Differential gene expression and pathway analysis.

Downstream statistical analysis was done using R programming language. We tested for differentially expressed genes under various conditions using DEseq2 [301], an R package based on a negative binomial distribution that models the number reads from RNA-seq experiments and tests for differential expression. We considered genes as significantly differentially expressed between two groups if $FDR \leq 0.05$ and the absolute value of \log_2 fold change > 1 . For downstream analysis and visualization, we first normalized gene counts with DeSeq2 and used the subsequent normalized counts (rlog normalized counts). For some analyses, the normalized counts were averaged between CD69⁻ and CD69⁺ samples in order to identify CD69 independent differences between transcriptomes of T cells derived from different tissue sites. We performed multidimensional scaling (MDS) analysis using the *cmdscale* function in R on a distance matrix computed with the *dist* function in R. Principle component analysis (PCA) was done using the *pca* function within DeSeq2 which uses the *prcomp* function. Both MDS and PCA were plotted using *ggplot*. For heatmap visualization of RNA-seq data Z-score of FPKM values were plotted using Microsoft Excel. Pathway analysis was conducted using Ingenuity Pathway Analysis software (IPA; QIAGEN). We used the Canonical Pathways function of IPA software to compare transcriptomes of memory T cells isolated from different sites (bone marrow, lymph nodes, spleen, lung, and blood).

Gene Set Enrichment Analysis (GSEA).

GSEA was implemented to compare human LN CD8⁺TEM cells to published datasets of mouse CXCR5-positive or CXCR5-negative T cells [302, 303] generated in response to chronic infection. We used two approaches. First, we obtained a gene list from microarray data [302]. For comparing to microarray data, we used a list of top upregulated genes (250 genes) and a list of the top downregulated genes (250 genes) ranked by fold change. The second approach used RNA-sequencing data. We derived a gene list from published RNA-sequencing DE analysis [303]. We split up the DE genes into two lists: a list of upregulated and a list of downregulated genes defined by *Deseq* analysis conducted in the paper and applied the following cutoffs: $FDR \leq 0.05$, $|\text{Foldchange}| \geq 1$). We used the gene lists from mouse populations as input into GSEA [304] and tested the null hypothesis that mouse genes have uniform distribution in ranks by the absolute value of log fold change between human TEM cells from lymph nodes and bone marrow on the x-axis. We reject the null hypothesis if the p-value is smaller than 0.05.

Section 2.5 T cell Proliferation Assays

Polyclonal T cell Proliferation Assay

T cells isolated from tissues were purified by cell sorting on an Influx cell sorter (BD Bioscience) in the CCTI flow cytometry core. TEM cells were sorted using the following antigens: CD3⁺CD8⁺CD4-CCR7-CD45RA⁻. After sorting, cells were stained with eFluor450 cell proliferation dye (eBioscience) according to manufacturer's instructions. Briefly, cells were washed twice with PBS to remove any serum, resuspended in Cell Proliferation Dye eFluor450 5uM final concentration in PBS and stained for 10 minutes in the dark at 37°C. Labeling was stopped by adding 4-5X volumes of cold RPMI-1640 containing 10% FBS and incubated for 5 minutes on ice. Cells were plated in a 48 well plate at 5X10⁵ cells/mL in complete media (RPMI-1640, 10% FBS, 1mM sodium pyruvate, 100 U/mL penicillin, 100ug/mL streptomycin, and 2mM L-glutamine) in the presence of ImmunoCult Human CD3/CD28 T cell Activator (STEMCELL technologies) at 25 uL/mL media. Cells were analyzed at day 3,4 and 5 after stimulation. Controls included unstimulated cells as well as cells unstained with eFluor450 in order to determine cell divisions. In some cases, 1000 Units of Recombinant Human IFN- α A (R&D Systems) and recombinant Human IFN- β (PeproTech) was added to culture medium.

CMV-specific T cell proliferation

Mononuclear cells (1-3X10⁶) isolated from blood or tissues as described above, were cultured in 96 well tissue culture plates at 5X10⁵ cells/mL Complete Media in presence or absence of 0.3 μ g/mL HCMV pp65 peptide mix (JPT Peptide Technology). IL-2 100U/mL was added on day 2 and cells were analyzed at day 8 or 9 after stimulation. Media was checked daily, and as needed was replenished with fresh media. HLA multimer reagents containing epitopes of

CMV (CMV-multimers) (see Table 2-1) were obtained from ProImmune. Staining with multimers was done according to the manufacturers protocols. Briefly, cells isolated from bone marrow, blood, and lymph nodes were washed with PBS and stained for 10mins at 37°C with tetramer reagent. Cells were washed and subsequently stained with other antibodies for downstream analysis as described in methods. Cells were analyzed by flow cytometry.

Section 2.6 TCR sequencing

DNA extraction

For most samples, sorted T cell subsets were pelleted and resuspended in cell lysis solution (Qiagen). DNA was isolated from cell lysate using the Genra Puregene kit (Qiagen). For the samples from 3 of 11 total donors, (D287, D280, and D229) DNA was extracted using an RNA/DNA kit (AllPrep DNA/RNA mini kit, QIAGEN) according to manufacturer's protocol. In the donor 383, the blood samples had fewer cells sequences, and therefore may not be quantitative, but the rest of the samples were equally sampled.

TCR V β amplification, library preparation and sequencing.

TCR sequencing was performed by the Human Immunology Core in collaboration with Dr. Nina Luning Prak and her laboratory. For TCR V β amplification, a cocktail of 23 V β families from framework region 2 (FR2) forward primers, and 13 J β region reverse primers, modified from the BIOMED2 primer series were used [305]. The PCR was performed with two mixes, all of which used the same 23 V β forward primers, but two different J β mixes. Primer sequences are provided in Table 2-3. The V β and J β primers mixes were used at 0.6 μ M in a reaction volume of 25 μ L using a Multiplex PCR kit (Qiagen, Cat. No. 158388). Amplification conditions for the PCR were: primary denaturation at 95°C for 10 minutes, cycling at 95°C 45s, Ta (57°C for J β mix 1, 61°C for J β mix 2) for 90s, extension at 72°C for 90s for 35 cycles, and a final extension step at 72°C for 10 minutes.

Amplicons were purified using the Agencourt AMPure XP beads system (Beckman Coulter, Inc) in a 1:1 ratio of beads to sample and eluted in 40 μ L of TE (0.1mM EDTA) buffer. Second-round PCRs to generate the sequencing libraries were carried out using 4 μ L of the first

round PCR product and 2.5 μ L each of NexteraXT Index Primers S5XX and N7XX, using the Qiagen Multiplex PCR kit in a reaction volume of 25 μ L. Amplification conditions for the PCR were primary denaturation at 95°C for 10 minutes, followed by cycling at 95°C 30s, 60°C 30s, extension at 72°C 45s for 8 cycles, and a final extension step at 72°C for 10 minutes. To confirm adequacy of amplification, aliquots of both the 1st and the 2nd round PCR products were run on agarose gels. Library quality was evaluated using Bioanalyzer 2100 (Agilent Technologies) and quantified by Qubit Fluorometric Quantitation (Thermo Fisher Scientific). A sharp single band from Bioanalyzer analysis indicated a good quality library and was used for sequencing. Readings from Qubit using the dsDNA HS (high sensitivity) assay kit (Cat. No. Q32851) were used to calculate the molarity of the library. Libraries were then loaded onto an Illumina MiSeq in the Human Immunology Core Facility at the University of Pennsylvania. 2x300 bp paired end kits were used for all experiments (Illumina MiSeq Reagent Kit v3, 600 cycle, Illumina Inc., Cat. No. MS-102-3003).

Table 2-3. PCR Primers used for TCR-V β amplification.

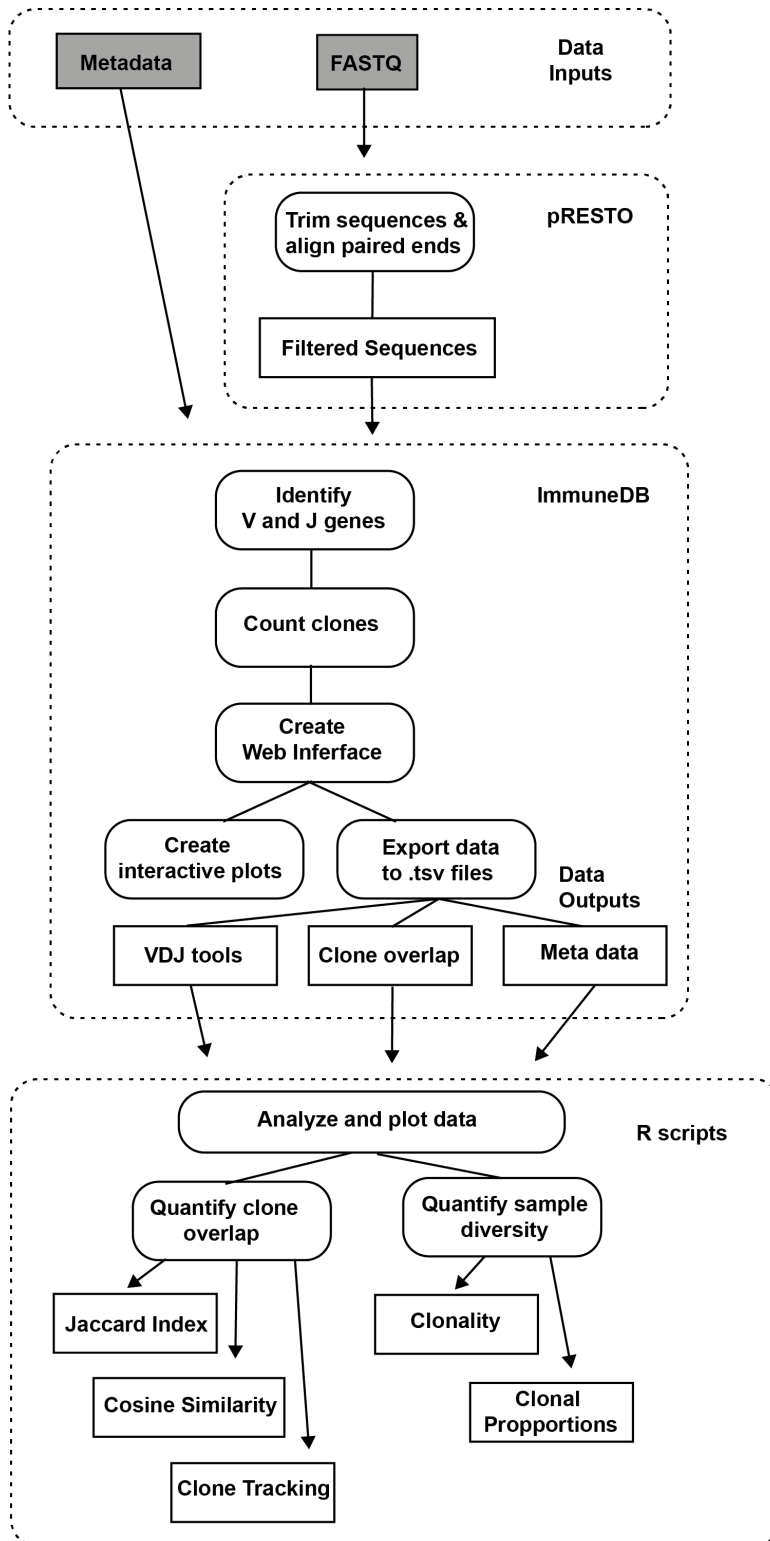
Primers to amplify TCR-V β rearrangements including 23 V β , 13 J β mix 1 and 2 (J β 1 and J β 2).

Name	Sequence
V β 2	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAACTATGTTTTGGTATCGTCA3'
V β 4	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCACGATGTTCTGGTACCGTCAGCA3'
V β 5/1	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAGTGTGTCCTGGTACCAACAG3'
V β 6a/11	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAACCCTTTATTGGTACCGACA3'
V β 6b/25	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATCCCTTTTTTGGTACCAACAG3'
V β 6c	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAACCCTTTATTGGTATCAACAG3'
V β 7	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCGCTATGTATTGGTACAAGCA3'
V β 8a	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTCCCGTTTTCTGGTACAGACAGAC3'
V β 9	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCGCTATGTATTGGTATAAACAG3'
V β 10	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTATGTTTACTGGTATCGTAAGAAGC3'
V β 11	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAAAATGTACTGGTATCAACAA3'
V β 12a/3/13a/15	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATACATGTACTGGTATCGACAAGAC3'
V β 13b	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGCCATGTACTGGTATAGACAAG3'
V β 13c/12b/14	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTATATGTCCTGGTATCGACAAGA3'
V β 16	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTAACTTTATTGGTATCGACGTGT3'
V β 17	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGCCATGTACTGGTACCGACA3'
V β 18	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCATGTTTACTGGTATCGGCAG3'
V β 19	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTATGTTTATTGGTATCAACAGAATCA3'
V β 20	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAACCTATACTGGTACCGACA3'
V β 21	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTACCCTTTACTGGTACCGGCAG3'
V β 22	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATACTTCTATTGGTACAGACAAATCT3'
V β 23/8b	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCACGGTCTACTGGTACCAGCA3'
V β 24	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCGTCATGTACTGGTACCAGCA3'
J β 1.1	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACTTACCTACAACCTGTGAATCTGGTG3'
J β 1.2	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGCTTACCTACAACGGTTAACCTGGTC3'
J β 1.3	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAAACCTTACCTACAACAGTGAGCCAACT T3'
J β 1.4	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGCGACATACCCAAGACAGAGAGCTGGG TTC3'
J β 1.5	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGANCTTACCTAGGATGGAGAGTCGAGTC3'
J β 1.6	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCATACCTGTCACAGTGAGCCTG3'
J β 2.1	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAACCTTCTTACCTAGCACGGTGA3'
J β 2.2	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNGACTTACCCAGTACGGTCAGCCT3'
J β 2.3	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTCCCGCTTACCGAGCACTGTCA3'
J β 2.4	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCAGCTTACCCAGCACTGAGA3'
J β 2.5	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNCGCGCACACCGAGCAC3'
J β 2.6	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCTCGCCAGCACGGTCAGCCT3'
J β 2.7	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTAAAACCTTACCTGTAACCGTGAGCCTG3'

TCR read counting and clone mapping

Raw reads were first processed using pRESTO [306] and filtered as described [307] with the script provided in Appendix A (page 222). Briefly, sequences are trimmed of poor-quality bases, paired reads are aligned into full length contiguous sequences, short sequences are then filtered out and bases with low quality score are replaced with an N, and any sequence containing more than 10 such bases is removed from further analysis. Filtered sequences are next processed by MiXCR (version 2.1) [308] or ImmuneDB [309, 310] for V and J gene identification and clone counting. The data from MiXCR and ImmuneDB were compared for initial data analysis and showed similar results; ImmuneDB was used for all analysis presented here. For the ImmuneDB pipeline, a non-default parameter was implemented during the clone counting step where we defined sequences with identical CDR3 amino acid sequences and the same V and J gene pairings to be collapsed into clones. In addition to clone assignment function of ImmuneDB, there is also a feature to create a web interface. We used the web interface for two purposes: 1) creating interactive plots for initial data visualization and exploration and 2) exporting the data to .tsv files in various formats for more in-depth downstream analysis scripted in R. The ImmuneDB webpage contains instructions for implementing the ImmuneDB pipeline (https://immunedb.readthedocs.io/en/latest/pipeline_example.html). To summarize, we used pRESTO, ImmuneDB, and R scripting for analysis of TCR sequences [see flowchart below in Figure 2-2].

Figure 2-2. Flowchart of TCR analysis.



Overview of TCR analysis in R

The ImmuneDB web interface (<http://clash.biomed.drexel.edu/databases/miron5/>) was used to export T cell receptor sequencing data in two formats for analysis. Downstream analysis included several methods of quantifying and visualizing T cell clone diversity of samples. Diversity analysis included clonal summary plots and calculating the clonality of each sample. We also did several analyses to quantify the extent of clonal overlap between samples. These overlap analyses including cosine similarity calculation, Jaccard index calculation, principle coordinate analysis of cosine similarity and clone tracking plots. Scripts for all of these analyses were written in R and can be found in the Appendix (page 222).

TCR diversity analysis

The clonal summary plots were generated using the *clonal.proportion* function from the tcR package in R. Clonality was measured for all TCR samples by normalizing the entropy values of each sample to the number of unique TCR sequences, resulting in a value ranging from 0 (most diverse) to 1 (least diverse). Given a clonotype x , the frequency of that clonotype, denoted as $p(x)$ was used to calculate Shannon entropy (H) of a TCR repertoire X by: $H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$. Normalized entropy (H_{norm}) is calculated in order to control for set size where L is the total number of unique clonotypes in the sample: $H_{\text{norm}}(X) = \frac{H(X)}{-\log_2 \frac{1}{L}}$. Next, clonality is calculated as follows: $\text{Clonality}(X) = 1 - H_{\text{norm}}(X)$. For clonality analysis, the same input T cell DNA was used for each sample from the same individual. True diversity is calculated using different orders as described [311].

Quantifying TCR clone overlap by Cosine Similarity and Jaccard Index

Cosine similarity was calculated using the cosine function in R. The mean of the four pairwise comparisons of two replicates per biological sample was reported in the heatmap. Hierarchical clustering was performed on the matrix of cosine similarity values for CD4⁺ and CD8⁺ T cells using the complete linkage, a method that find similar clusters, performed with the *pheatmap* function in R. Jaccard index (number of intersecting clonotypes over the union of clonotypes between two samples) was calculated using filtered data with abundance counts greater than 50% of the mean frequency to account for sequencing errors.

Principal coordinates analysis of T cell clonotype abundances

Raw read counts were first sampled down to a depth of 19,000 reads per sample, and any samples with total read counts below this cutoff were excluded. To control for the presence of any clonotypes that may have resulted from technical errors, a single sequence was required to be present in at least two samples to be included in the analysis. Pairwise cosine distances between samples were then calculated using the *pdist* function in MATLAB. The *cmdscale* function in MATLAB was used to perform principal coordinates analysis (PCoA).

Clone Tracking Analysis

The clone tracking plots were generated using VDJtools [312] software and specifically the Track Clonotypes command. The R script modified from the VDJtools software can be found in the Appendix (page 239).

Section 2.7 Statistical tests

Descriptive statistics (percent, mean, median, SEM) were calculated using Prism (Graphpad software) for flow cytometry data and R programming language for T cell receptor sequencing data. Significant differences in frequencies, ratios, gMFI, and density were assessed using a paired *t* test.

CHAPTER 3: Human lymph nodes maintain TCF-1⁺ T cells with high functional potential and clonal diversity

ABSTRACT:

Our understanding of T cell responses in human lymphoid sites and their relation to peripheral blood remains sparse. In this study, we used a unique human tissue resource to study human T cells in different anatomical compartments. We identify lymph nodes (LN) maintain memory CD8⁺ T cell with a distinct differentiation and functional profile compared with memory CD8⁺ T cells in blood, spleen, bone marrow, and lungs. LN memory CD8⁺ T cells express transcriptome and protein signatures of quiescence and self-renewal compared with corresponding populations in blood, spleen, bone marrow, and lung. Functionally, human LN memory T cells exhibit increased proliferation to TCR-mediated stimulation and maintain higher TCR clonal diversity compared with memory T cells from blood and other sites. These findings establish human LN as reservoirs for memory T cells with high capacities for expansion and diverse recognition and important targets for immunotherapies.

Chapter expanded from:

Miron, M., Kumar, B.V., Meng, W., Granot, T., Carpenter, D. J., Senda, T., Chen, D., Rosenfeld, A., Zhang, B., Lerner, H., Friedman, A., Hershberg, U., Shen, Y., Rahman, A., Luning-Prak, E., Farber, D.L. (2018) Human lymph nodes maintain TCF-1^{hi} memory T cells with high functional potential and clonal diversity throughout life. *Journal of Immunology*. 201(7), 2132-2140. doi: 10.4049/jimmunol.1800716.

Section 3.1 Introduction

T cells mediate adaptive immune responses and long-lived protective immunity, through their differentiation to effector and memory T cell populations, respectively. While the majority of effector T cells are short-lived *in vivo*, a subset of primed effector cells differentiate and persist as populations of long-lived memory T cells. In humans, T cell subset differentiation and memory maintenance have been extensively characterized from peripheral blood, revealing subsets of memory T cells that differ in phenotype and proliferative potential.[14, 313] However, the majority of T cells in the body are localized in tissue sites, and particularly in lymphoid tissues including bone marrow (BM), spleen and an estimated 700 lymph nodes[4], where T cell responses are initiated, regulated and maintained. How human T cells in lymphoid sites function and persist relative to subsets in blood is not well understood; defining their nature is important for monitoring and modulating immune responses, and translating findings from mouse models where T cells from spleen and lymph nodes are the predominant sites of investigation.

In human tissues, a significant fraction of memory T cells express markers including CD69 and CD103 which denote tissue-resident memory T cells (TRM).[7, 8, 15] In human lymphoid tissue, a lower proportion of memory T cells express CD69 and CD103 compared to mucosal sites.[8] It is not known whether lymphoid memory T cells persist with similar functional properties across tissues, or adopt compartment-specific attributes. Moreover, memory T cells in secondary lymphoid organs of laboratory mice generally do not exhibit TRM phenotypes[314], suggesting that lymphoid memory T cells in humans may exhibit features distinct from those in mouse tissues due to their longevity and/or increased exposure to diverse antigens encountered over human life.

T cell differentiation to effector and memory cell fates is regulated by key transcription factors.[170] T cell factor-1 (TCF-1) is essential for memory T cell formation and maintenance

in the periphery, through regulation of pathways for survival and quiescence[144, 315], while T-bet promotes effector over memory T cell generation.[91] In mice, differential expression of TCF-1 by memory T cells can influence their function and localization. In mouse models of chronic viral infection, several groups identified a TCF-1⁺ subset of memory CD8⁺T cells to be the responding population mediating viral clearance following anti-PD-1 therapy.[248, 303] This TCF-1⁺ memory subset was absent from peripheral blood, and only detected in lymphoid sites.[154, 248, 303, 316, 317] For human T cells, the role of specific transcription factors in the differentiation, maintenance and localization of effector and memory T cells has not been defined.

In this study, we used our unique tissue resource where we obtain blood, multiple lymphoid and mucosal tissues from individual organ donors of all ages [6, 7, 11] to investigate how T cells in lymphoid sites are transcriptionally and functionally related to those in blood and peripheral sites. By studying T cells across multiple tissues within and between individuals, we demonstrate here that memory CD8⁺T cells are maintained in human LN in an organ-specific manner throughout the human lifespan. Notably, LN memory CD8⁺T cells maintain expression of transcription factors, TCF-1 and LEF-1 associated with self-renewal and exhibit distinct transcriptional and protein expression signatures involved in T-follicular helper (Tfh) cell differentiation with downregulation of effector function and inflammatory signals compared to memory T cells in blood, other lymphoid (spleen, BM) and peripheral sites (e.g., lungs). LN memory CD8⁺T cells also exhibit higher proliferative capacity and increased T cell receptor clonal diversity compared to memory T cells in other sites. Together these findings establish human LN as reservoirs for maintenance of high potential memory T cells, and LN memory CD8⁺T cells as novel targets for immune modulation and adoptive immunotherapies.

Section 3.2 Results

Memory T cells exhibit increased TCF-1 expression in lymph nodes relative to other sites

Human T cells can be subdivided into major subsets based on CD45RA and CCR7 expression into naïve ($CD45RA^+CCR7^+$), central memory (TCM, $CD45RA^-CCR7^+$), effector-memory (TEM, $CD45RA^-CCR7^-$), and terminally differentiated effector cells (TEMRA, $CD45RA^+CCR7^-$).[7] For $CD8^+$ T cells, TEM phenotype cells are the predominant memory T cell subset in blood and diverse tissue sites, with low frequencies of TCM-phenotype cells[7], as shown here in BM, spleen, lung and lymph nodes (LN) (Figure 3-1A). We examined whether $CD8^+$ TEM cells exhibited tissue-specific variations in expression of TCF-1, a transcription factor associated with self-renewal.[145] In all donors, TCF-1 expression by $CD8^+$ TEM in LN was significantly higher than that found in $CD8^+$ TEM derived from blood, BM, spleen, and lung (Figure 3-1B). TCF-1 $^+$ $CD8^+$ TEM cells were observed in LN draining different anatomical sites (lungs, intestines, groin, Figure 3-1B) and across all ages (9-76 years, Figure 3-1C). Moreover, TCF-1 expression was comparably high within $CD69^+$ and $CD69^-$ subsets of LN memory T cells (Figure 3-2), delineating circulating and tissue-resident memory populations.[8] These results suggest that increased TCF-1 expression is an organ-specific (rather than subset-specific) feature of memory $CD8^+$ T cells within LN.

Figure 3-1. Expression of TCF-1 is maintained in LN CD8⁺ TEM cells.

(A) CD3⁺CD8⁺T cell subset composition showing frequencies of effector memory cells (TEM, CD45RA⁻CCR7⁻) in blood and tissues from a representative individual (D251). (B) TCF-1 expression is restricted to LN TEM cells. Left: TCF-1 expression by CD8⁺TEM cells in representative histograms from one donor (D334). Right: compiled frequencies from multiple donors (n=11). LLN=lung-draining, MLN= mesenteric-draining, ILN= iliac-draining lymph nodes; Spl= Spleen, Lng= Lung (C) LN-specific TCF-1 expression is maintained with age. Ratio of TCF-1 gMFI by CD8⁺TEM in tissues to that in blood. Error bars indicate SEM. * P<0.05, ** P<0.01, *** P<0.001, ****P<0.0001; NS=not significant by two-tailed t-test.

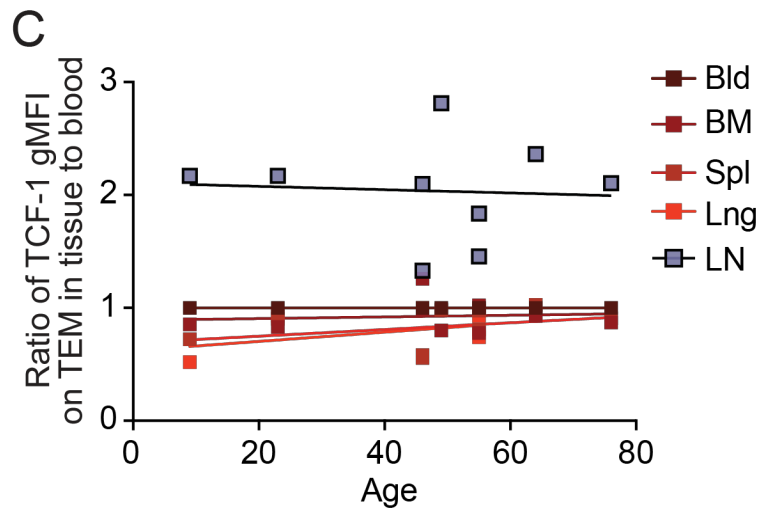
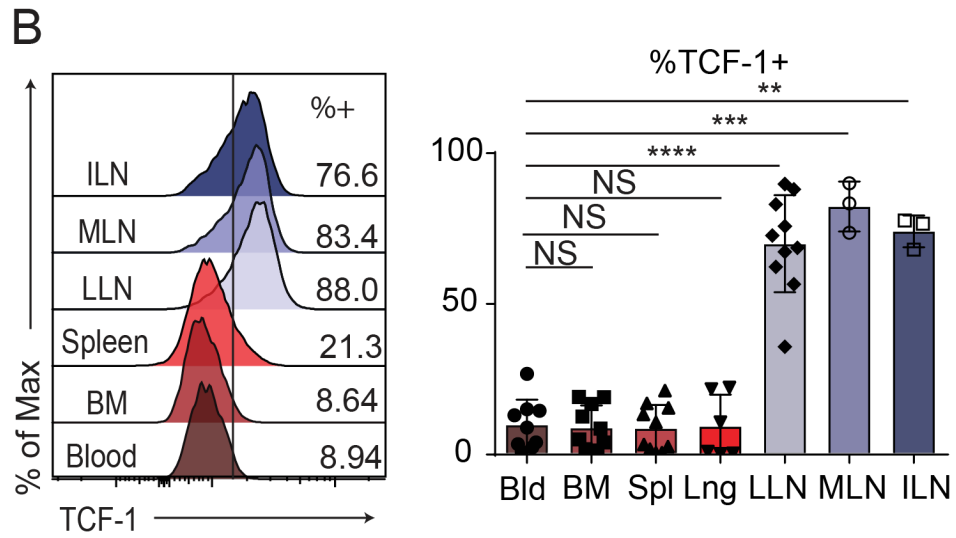
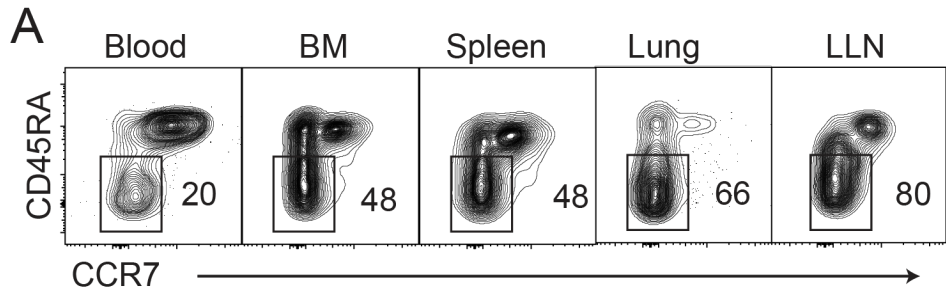


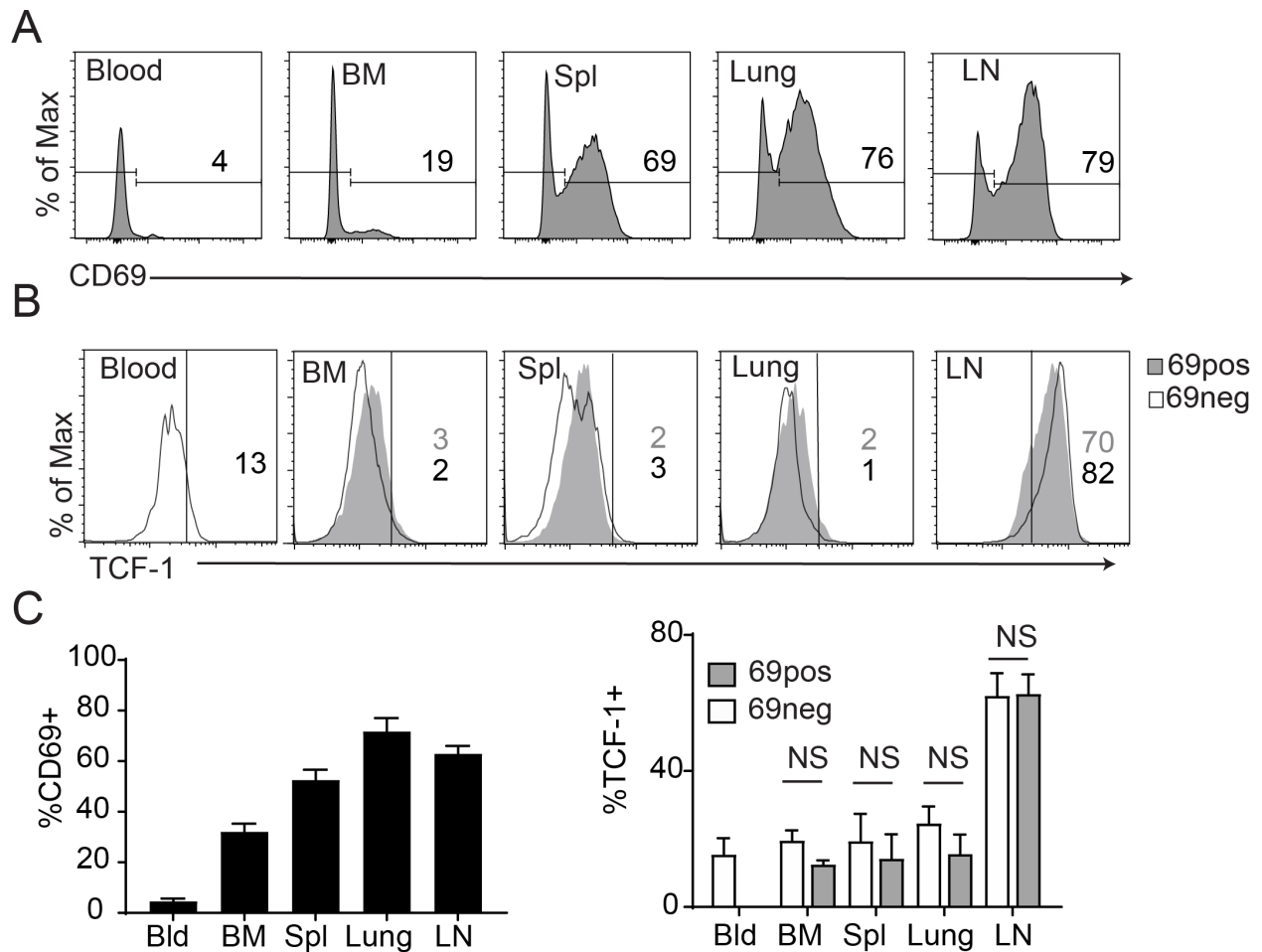
Figure 3-2. Expression of TCF-1 in CD69⁻ and CD69⁺ CD8⁺ TEM cells in tissues.

(A) Expression of CD69 by CD8⁺ TEM cells in indicated tissues from a representative donor (D259).

(B) Expression TCF-1 by CD69^{pos} and CD69^{neg} CD8⁺TEM cells in indicated tissues from a representative donor (D302).

(C) Left: compiled expression data from (A) (Bld n=19, BM n= 28, Spl n=24, Lung n=12, LN n=27).

Right: compiled expression data from (B) (Bld n=5, BM n= 8, Spl n=6, Lung n=4, LN n=8).



LN memory T cells exhibit a distinct phenotype and transcriptional profile

We used whole transcriptome profiling by RNAseq to investigate whether increased TCF-1 expression by LN memory CD8⁺T cells indicated a distinct transcriptional program compared to memory CD8⁺T cells in blood and other lymphoid sites. We initially analyzed differential gene expression between memory CD8⁺T cells in LN and BM of three individuals (see sorting strategy, Figure 3-3), identifying over 2,000 differentially expressed (DE) genes conserved between individuals, many of which included markers associated with T cell differentiation by pathway analysis shown in Table 3-1. In particular, LN memory CD8⁺T cells had increased expression of transcripts associated with self-renewal (LEF-1, TCF-7), follicular helper T (Tfh) cell differentiation [78, 152, 318] (BCL-6, CXCR5, CXCR4, CCR7), co-stimulation (CD28, ICOS), and reduced expression of effector transcripts (GZMA, PRF1) (Figure 3-4A). Genes involved in Wnt signaling (WNT10A, CD44, SOX13) and cell cycle control (CDKN1A, CDKN2C) were also differentially expressed in LN compared to BM memory CD8⁺T cells (Figure 3-4A). Analysis of protein expression confirmed increased CXCR4 and LEF-1 expression and decreased expression of Perforin and T-bet in human LN compared to BM CD8⁺memory T cells (Figure 3-4B).

We further compared the transcriptional profile of LN memory CD8⁺T cells to that in blood and other tissues including spleen and lungs, for which we previously obtained RNAseq profiles. [8] Based on the ~2,000 DE genes defined above, the transcriptional profile of LN memory CD8⁺T cells clustered together by principle component analysis (PCA), distinct from that of memory CD8⁺T cells in blood, BM, spleen and lung for all donors analyzed (Figure 3-4C). Through paired analysis of gene expression between LN and these other sites, we identified a gene signature comprising 330 upregulated and 340 downregulated genes in LN memory CD8⁺T cells compared to the corresponding subset in blood, spleen or BM (See gene

lists of genes upregulated in Table 3-2 and downregulated in Table 3-3). Notably, genes involved in T-follicular helper cell differentiation [319] were upregulated (IL21R, LEF-1, and ICOS), and genes involved in homing and cytotoxic function were downregulated (S1PR5, FCER1G, GRMH, and NKG7) in LN compared to blood, spleen and BM memory CD8⁺T cells (Figure 3-4D).

Figure 3-3. Sorting strategy for isolation of lymphocytes for RNA-seq and TCR-seq.

(A) Gating strategy shown in the following order (left to right): lymphocytes, singlets, memory T cells ($CD3^+CD45RO^+$), $CD8^+$ and $CD69^{+/-}$ cells from bone marrow (BM) of a representative individual for RNA-seq and TCR-seq.

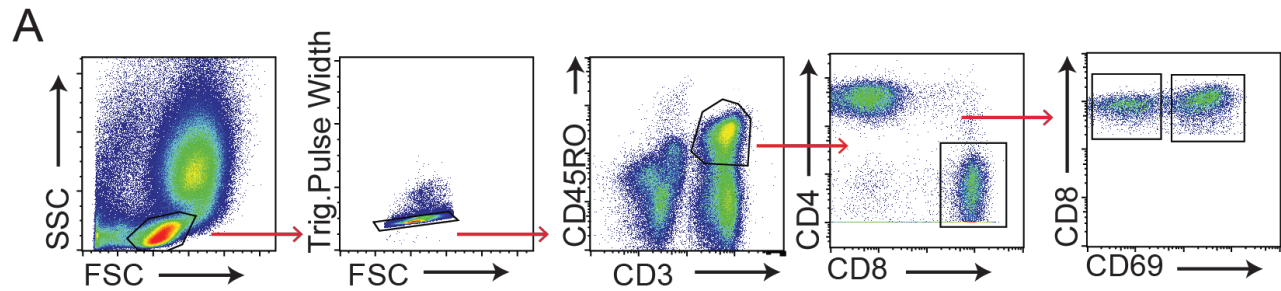


Figure 3-4. Human LN memory CD8⁺T cells are phenotypically and transcriptionally distinct from peripheral blood, BM, and Spl T cells.

(A) Heatmap of differentially expressed (DE) genes from whole transcriptome profiling of BM and LN (B and L respectively) CD8⁺TEM cells from three donors. (B) Protein expression of markers identified in (A) shown as histograms from one donor (top, from left to right: D259, D304, D227, D273 [see Table 3-6 for donor information]) and compiled: CXCR4, n=8; Perforin, n=5; Lef, n=7; T-bet, n=13 (bottom). (C) Principle component analysis (PCA) of transcriptional profiles of CD8⁺TEM cells from blood (Bld), bone marrow (BM), lung (Lng), spleen (Spl) and lymph node (LN) from nine individuals (1-9) based on the 2,521 DE genes between LN and BM memory CD8⁺T cells. (D) RNA expression of indicated genes among CD8⁺ TEM cells from blood and s tissue sites of nine individuals in (C). Error bars indicate SEM. * P<0.05, ** P<0.01, *** P<0.001, by two-tailed t-test.

Figure 3-4

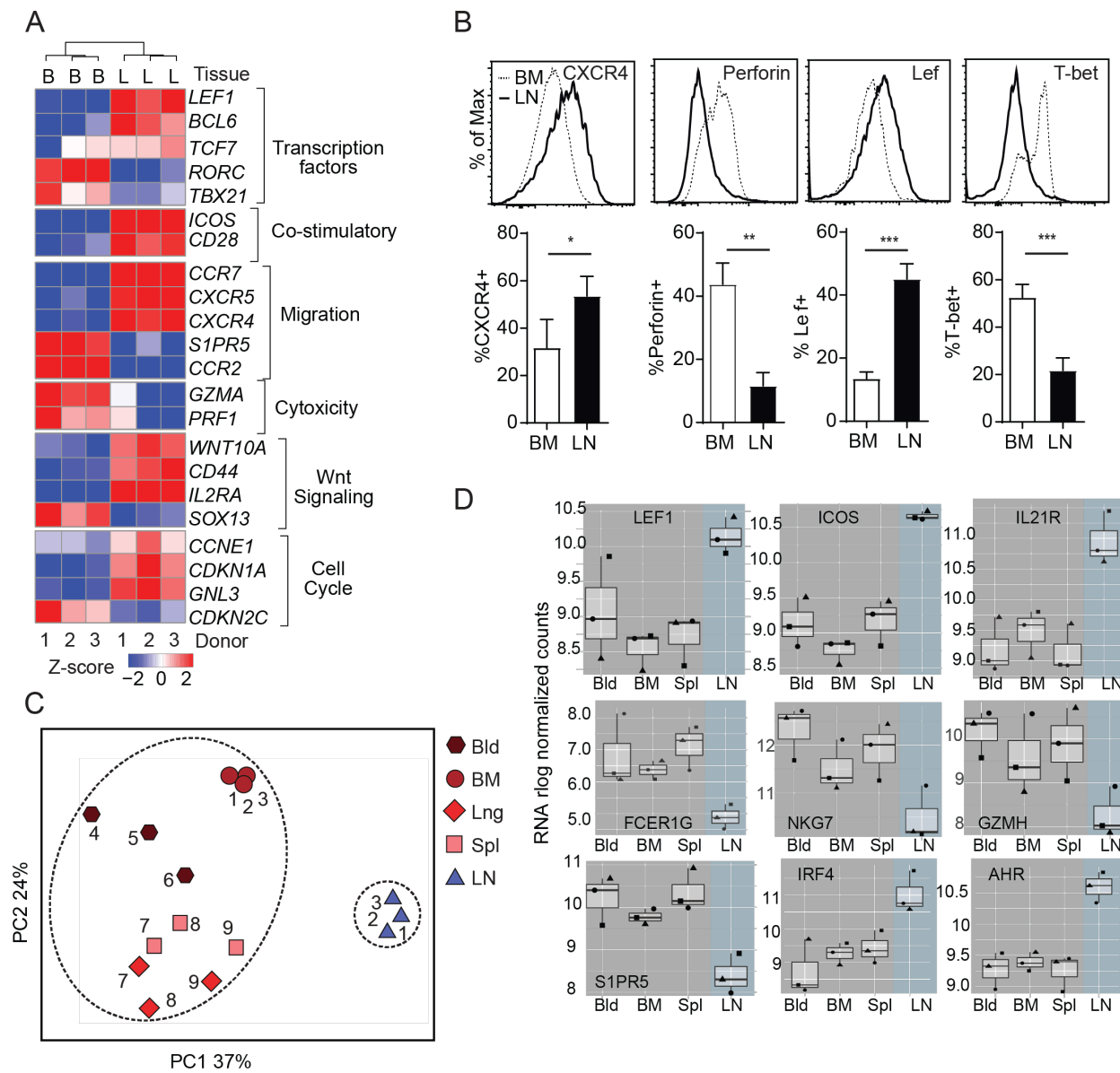


Table 3-1. Pathways differentially regulated in CD8⁺TEM cells in lymph nodes compared BM, spleen and blood.

Results were obtained from the Canonical Pathways function of Ingenuity Pathway Analysis software (QIAGEN) by comparison of memory T cells isolated from different tissues (LN vs. BM, LN vs. Blood, and LN vs. Spl).

Categories	Direction in LN	Ingenuity Canonical Pathways	z-score LN vs. BM	z-score LN vs. Spl	z-score LN vs. Bld
TFH or TH2	Up	iCOS-iCOSL Signaling in T Helper Cells	3.3	1	-0.343
TFH or TH2	Up	CD28 Signaling in T Helper Cells	3.153	1.698	-0.16
TFH or TH2	Up	Th2 Pathway	2.785	0.218	1.372
TFH or TH2	Up	CXCR4 Signaling	1.291	0.655	-0.48
Cell Cycle inhibition	Up	Cell Cycle: G1/S Checkpoint Regulation	0.447	1.265	0.577
Co-stimulatory	Up	4-1BB Signaling in T Lymphocytes	2.449	1	1.941
Stem cell & Development	Up	Role of NANOG [...] Stem Cell	2.646	1.265	0.688
Stem cell & Development	Up	Mouse Embryonic Stem Cell Pluripotency	2.496	1.807	0.667
Stem cell & Development	Up	Wnt/ β -catenin Signaling	1.043	1.414	2.359
T cell activation	Down	cAMP-mediated signaling	-0.667	-2.117	-2.219
T cell activation	Down	Phospholipase C Signaling	-0.784	-1.732	-1.987
TH1	Down	Cytotoxic T Lymphocyte	-0.447	n/a	n/a
Cell signaling	Down	Gas Signaling	-0.632	-0.905	-1.177
Cell signaling	Down	Gai Signaling	-0.632	-0.905	-1.177
Cell signaling	Down	Eicosanoid Signaling	-1.633	-1.265	-2.111
Inflammatory	Down	Inflammasome pathway	-0.632	-1.633	-0.632
Inflammatory	Down	Toll-like Receptor Signaling	-2.309	-2.53	-2.558
Inflammatory	Down	NF- κ B Signaling	-0.174	-0.962	-2.885

Table 3-2. Total genes upregulated in LN CD8⁺TEM.

Overlap of genes from differential expression analyses between CD8⁺TEM in different sites::

LN vs. BM, LN vs, Spl, and LN vs. Blood. Deseq2 with the following cut-offs: FDR <0.05 and fold change >2. Genes of interest in bold.

#	Gene	#	Gene	#	Gene	#	Gene
1	ABCG1	44	CCRN4L	85	EFCAB2	125	GNL3
2	ACBD7	45	CD55	86	EFNB2	126	GPCPD1
3	ACHE	46	CENPC1	87	EIF4A3	127	GPR132
4	ACRC	47	CH25H	88	EIF5	128	GRHL1
5	ADAM23	48	CHADL	89	ELL2	129	GTPBP1
6	ADAMTS17	49	CHD1	90	ELTD1	130	HAVCR1
7	AFAP1	50	CHD2	91	EMP1	131	HBP1
8	AHI1	51	CHRNA10	92	ENPP2	132	HECA
9	AHR	52	CHRNE	93	EPGN	133	HES4
10	AIM1	53	CILP2	94	EPHA1	134	HIC1
11	ALG13	54	CNKSR2	95	EPHA4	135	HNRNPU-AS1
12	AMD1	55	CNOT10	96	ETV3	136	HNRPLL
13	AOC2	56	COL18A1	97	EVC2	137	HSPA2
14	AP3M2	57	COL1A1	98	EZH2	138	HSPG2
15	APBA2	58	COL6A3	99	EZR	139	HTRA4
16	AREG	59	CRTAM	100	FAAH2	140	ICOS
17	AREGB	60	CSDA	101	FABP5	141	IFRD1
18	ARIH1	61	CSDAP1	102	FAM160A1	142	IL21R
19	ARL4A	62	CSRNP1	103	FAM177A1	143	IL6ST
20	ARL5B	63	CTH	104	FAM184A	144	INE1
21	ASAP1-IT1	64	CTLA4	105	FAM18B2	145	IRF4
22	ATP1B3	65	CTNNAL1	106	FAM190A	146	ITPKB
23	AURKAPS1	66	CTSL2	107	FAM53C	147	ITPKB-IT1
24	AXIN2	67	CXCR5	108	FAM83D	148	JMJD1C
25	BAAT	68	CXorf40A	109	FASN	149	JMJD6
26	BACH1	69	CXorf57	110	FER	150	JMY
27	BACH2	70	CYB5D1	111	FGD6	151	KBTBD8
28	BAIAP2	71	DCHS1	112	FGF18	152	KCNIP4-IT1
29	BRMS1L	72	DDX21	113	FGF9	153	KCNK5
30	BTBD19	73	DDX27	114	FLJ43663	154	KCNQ10T1
31	BTG2	74	DENND2C	115	FLT1	155	KDM6B
32	CACNA1I	75	DENND4A	116	FNDC4	156	KIAA0895
33	CASS4	76	DLC1	117	FRMD6	157	KIF13B
34	CCDC138	77	DLL1	118	FSCN1	158	KIF18A
35	CCDC141	78	DLL4	119	FZD6	159	KLF5
36	CCDC157	79	DNAAF2	120	GCNT4	160	KLHL15
37	CCDC168	80	DNAJB9	121	GEM	161	KLHL25
38	CCDC19	81	DSCAML1	122	GINS4	162	KLRK1
39	CCNE1	82	DUSP4	123	GK	163	KPNA2
40	CCNH	83	DYRK3	124	GLB1L3	164	LAMC3
41	CCNYL1	84	EDN3				
42	CCR4						
43	CCR7						

#	Gene
165	LEF1
166	LETM2
167	LIFR
168	LINC00239
169	LMO7
170	LRP5
171	LRP6
172	LRRC32
173	LRRC8B
174	MAFA
175	MALT1
176	MAN1C1
177	MAP1A
178	MARCKSL1
179	MASTL
180	MAT2A
181	MATN1
182	METTL12
183	METTL21CP1
184	MFSD2A
185	MMRN1
186	MORC4
187	MRPL1
188	NAB1
189	NAF1
190	NAMPT
191	NAV2
192	NBL1
193	NCAPH
194	NFKB1
195	NKX3-1
196	NLRP3
197	NLRP6
198	NR4A3
199	NRARP
200	NRIP3
201	NUP98
202	OLFML3
203	PBX4
204	PCDH17
205	PCSK5
206	PDE4B
207	PELI1

#	Gene
208	PER2
209	PGAP1
210	PHLDB1
211	PHLDB3
212	PIP5KL1
213	PLCD3
214	PLEKHH2
215	PLK3
216	PLS3
217	PNPLA8
218	POLR3E
219	POU3F1
220	PPIL4
221	PPP4R1L
222	PPP4R4
223	PRMT10
224	PTGER4
225	PVT1
226	PYGM
227	ProSAPiP1
228	RAB30
229	RAI2
230	RALGAPA1
231	RANBP2
232	RASGEF1B
233	RASL11A
234	RBBP8
235	REL
236	RGS13
237	RGS16
238	RIC3
239	RND1
240	RNF19A
241	RNU86
242	RPL21
243	RRN3P1
244	RRP12
245	SAMSN1
246	SC5DL
247	SCARNA4
248	SCARNA9L
249	SCML1
250	SDK2

#	Gene
251	SELP
252	SEMA3F
253	SEMA3G
254	SEMA4A
255	SERINC5
256	SERTAD1
257	SESN3
258	SFMBT1
259	SFPQ
260	SGK1
261	SH3RF3
262	SHANK3
263	SHF
264	SHISA2
265	SIK1
266	SIRT1
267	SKIL
268	SLC19A2
269	SLC25A19
270	SLC25A33
271	SLC26A5
272	SLC35A2
273	SLC38A2
274	SLC38A3
275	SLC7A5
276	SLC7A5P1
277	SLC7A5P2
278	SLCO4A1
279	SMCHD1
280	SNAI1
281	SNAPC1
282	SNHG7
283	SNX9
284	SPAG4
285	SPDYA
286	SPRY4
287	SRSF2
288	ST6GALNAC3
289	SYDE2
290	TAF4B
291	TAGAP
292	TAS2R31
293	TEK

#	Gene
294	TEX14
295	TFRC
296	TIMD4
297	TIPARP
298	TLE3
299	TMEM2
300	TMEM67
301	TMEM88
302	TNFRSF10D
303	TNFRSF13C
304	TNFRSF9
305	TP53BP2
306	TRIM39
307	TRPS1
308	TSC22D2
309	TSPY26P
310	TTC18
311	TUBA1A
312	TUBB2C
313	TUFT1
314	USP36
315	USP37
316	VWA1
317	VWDE
318	WDR47
319	WNT7A
320	ZBTB10
321	ZC3H12D
322	ZDBF2
323	ZDHHC11B
324	ZNF10
325	ZNF165
326	ZNF184
327	ZNF250
328	ZNF280C
329	ZNF331
330	ZNF571

Table 3-3. Total genes downregulated in LN CD8⁺ TEM.

Overlap of genes from differential expression analyses between CD8⁺TEM in different sites: LN vs. BM, LN vs, Spl, and LN vs. Blood. Deseq2 with the following cut-offs: FDR <0.05 and fold change >2. Genes of interest in bold.

#	Gene	#	Gene	#	Gene	#	Gene
1	AATK	44	CD14	85	CXCR2	127	GLI1
2	ABCC3	45	CD160	86	CXXC5	128	GLRX
3	ACPP	46	CD180	87	CYBB	129	GLT1D1
4	ADAMTS14	47	CD1D	88	CYP4F22	130	GPR114
5	AGTRAP	48	CD244	89	DAPK1	131	GPR141
6	AIF1	49	CD300C	90	DCTPP1	132	GPR56
7	AKR1C3	50	CD300E	91	DOK3	133	GPR68
8	ALAS2	51	CD300LB	92	DPY19L1P1	134	GPR97
9	ALDH2	52	CD300LF	93	EMR3	135	GZMH
10	ALDH3B1	53	CD33	94	EPB42	136	HBA1
11	ALOX5	54	CD79B	95	EPS8L1	137	HBA2
12	ALOX5AP	55	CDA	96	EPX	138	HBB
13	AMPH	56	CDK2AP1	97	FAM20A	139	HCAR2
14	ANPEP	57	CEACAM1	98	FAM49A	140	HCK
15	ANXA4	58	CEACAM3	99	FAM65C	141	HCN2
16	APOBEC3A	59	CEMP1	100	FAM70A	142	HDAC9
17	APOBEC3C	60	CFP	101	FBXO6	143	HHEX
18	APOL1	61	CHRN1	102	FCAR	144	HK3
19	AQP9	62	CHST15	103	FCER1G	145	HLA-DMB
20	ARG1	63	CLDND2	104	FCGR1A	146	HOMER3
21	ARL11	64	CLEC10A	105	FCGR2A	147	HOPX
22	ARMC7	65	CLEC12A	106	FCGR2C	148	HP
23	ASCL2	66	CLEC12B	107	FCGR3A	149	HSPA7
24	ASGR2	67	CLEC4D	108	FCGR3B	150	IFI30
25	ATP2C2	68	CLEC4E	109	FCN1	151	IGSF6
26	B3GALT2	69	CLEC7A	110	FCRL6	152	IL17RC
27	B3GNT8	70	CLIC3	111	FFAR2	153	IL17RE
28	BASP1	71	COL17A1	112	FGD2	154	IL18
29	BEST1	72	COLQ	113	FGFBP2	155	IL4I1
30	BMP8A	73	COPZ1	114	FGL2	156	INPP1
31	BTK	74	CPVL	115	FGR	157	IRAK3
32	C5AR1	75	CREB5	116	FHL3	158	IRF5
33	C5orf13	76	CSF2RB	117	FHOD3	159	ITGAM
34	C9orf100	77	CSF3R	120	FPR1	160	ITGAX
35	CACNA2D2	78	CSTA	121	FPR2	161	KCNE3
36	CALHM2	79	CTBP2	122	FTSJ2	162	KCNJ15
37	CARD16	80	CTIF	123	GAPT	163	KLRD1
38	CASP1	81	CTSH	124	GBP5	164	LAT2
39	CCDC28B	82	CTSW	125	GEMIN7	165	LILRA1
40	CCDC65	83	CX3CR1	126	GIMAP7	166	LILRA2
41	CCNJL	84	CXCR1				
42	CCR1						
43	CCR2						

#	Gene
167	LILRA3
168	LILRA5
169	LILRA6
170	LILRB1
171	LILRB2
172	LILRB3
173	LINGO3
174	LPPR3
175	LRRC25
176	LRRK2
177	LST1
178	LTBR
179	LTK
180	LY86
181	LYN
182	LYZ
183	MARCKS
184	MAST1
185	MCTP1
186	MEFV
187	METTL18
188	METTL21B
189	MGAM
190	MGST1
191	MLC1
192	MME
193	MMP25
194	MNDA
195	MOSC1
196	MPEG1
197	MPPED2
198	MS4A6A
199	NAPEPLD
200	NCAM1
201	NCF1
202	NCF1C
203	NCF2
204	NCR1
205	NCR3
206	NECAB1
207	NEO1
208	NFAM1
209	NFE2
210	NKG7
211	NLRC4

#	Gene
212	NLRP12
213	NMUR1
214	NR1D1
215	NTN4
216	NTSR1
217	OAZ3
218	OSCAR
219	OSTalpha
220	OXER1
221	P2RX1
222	P2RY13
223	PADI2
224	PADI4
225	PCTP
226	PDZD4
227	PF4
228	PFKFB4
229	PGLYRP1
230	PHOSPHO1
231	PLBD1
232	PLEK
233	PNKD
234	POP7
235	PPBP
236	PRAM1
237	PROCR
238	PROK2
239	PRSS30P
240	PRSS35
241	PTAFR
242	PTGDS
243	PTK6
244	PYGL
245	QPCT
246	RAB31
247	RAB32
248	RAB6B
249	RARG
250	RASGRP4
251	RASSF4
252	RBM47
253	RBP5
254	RGS18
255	RGS7BP
256	RNASE6

#	Gene
257	RNF135
258	RNF165
259	ROGDI
260	RORC
261	RTP4
262	S100A11
263	S100A12
264	S100A4
265	S100A8
266	S100P
267	S1PR5
268	SCRN1
269	SELRC1
270	SEPX1
271	SERINC2
272	SERINC4
273	SERPINA1
274	SH2D1B
275	SH3BGRL3
276	SIGLEC10
277	SIGLEC14
278	SIGLEC16
279	SIGLEC5
280	SIGLEC7
281	SIGLEC9
282	SIGLECP3
283	SIRPA
284	SIRPB2
285	SKAP2
286	SLAMF7
287	SLC11A1
288	SLC16A11
289	SLC1A7
290	SLC22A15
291	SLC24A4
292	SLC46A2
293	SLC4A10
294	SLC7A7
295	SLCO4C1
296	SLFN14
297	SOX13
298	SPI1
299	SPTB
300	SPTBN2
301	ST3GAL4

#	Gene
302	ST3GAL6
303	SYK
304	SYNGR1
305	TACO1
306	TARP
307	TBXAS1
308	TEF
309	TFEC
310	THAP8
311	TLR2
312	TLR7
313	TLR8
314	TM6SF1
315	TMEM150B
316	TMEM79
317	TNFAIP8L2
318	TNFSF13
319	TNFSF13B
320	TP53TG1
321	TREM1
322	TRIM21
323	TRIM58
324	TSPAN15
325	TSPO2
326	TTC38
327	TUBB1
328	TULP2
329	TXNDC3
330	TYROBP
331	VCAN
332	VEPH1
333	VNN1
334	VSTM1
335	ZBTB12
336	ZKSCAN4
337	ZNF137P
338	ZNF185
339	ZNF213
340	ZNF467

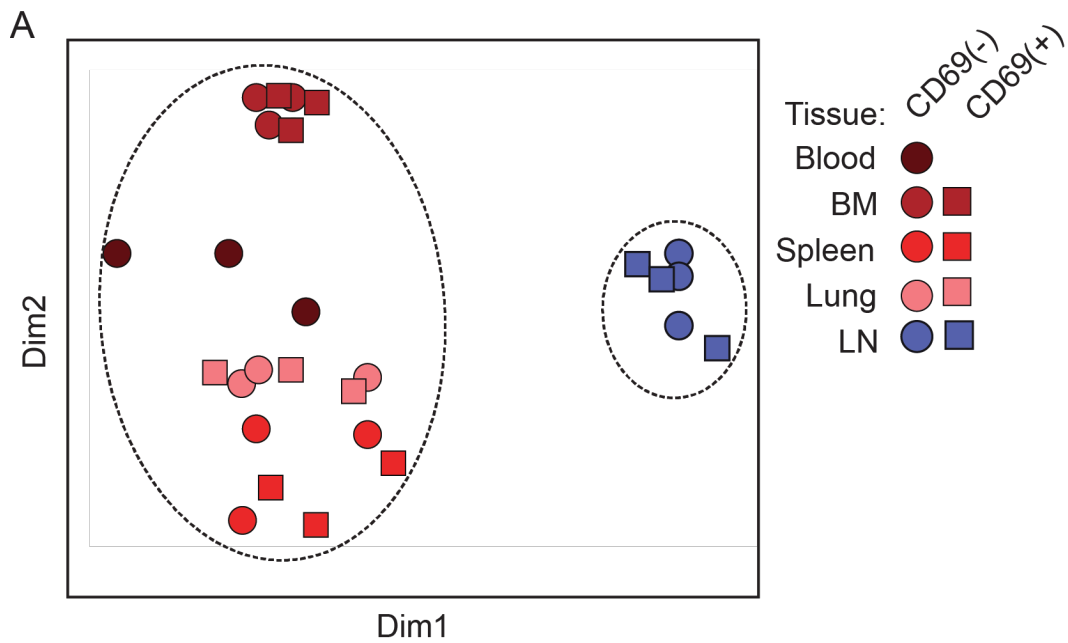
Human LN memory T cells are transcriptionally similar to murine TCF-1+ CD8+ T cells that respond to checkpoint blockade

We investigated whether the distinct transcriptional profile of LN memory CD8⁺T cells was due to enrichment of specific subset, such as TCM or TRM cells within the LN, or was organ specific. The LN-specific memory CD8⁺T cell profile identified in Figure 3-4 was not significantly similar to that previously defined for blood TCM CD8⁺TCM cells (18 shared genes of 785 TCM-associated genes).[320, 321] Moreover, the LN-specific gene signature identified in Figure 3-4 was conserved within both CD69⁺ (TRM-phenotype) and CD69⁻ (circulating) memory subsets (Figure 3-5). Therefore, the distinct transcriptional profile of human LN memory CD8⁺T cells is not attributed to an enrichment of TCM or TRM subsets.

Figure 3-5. Relatedness of CD69⁺ and CD69⁻ CD8⁺TEM from tissues and blood based on differentially expressed genes between CD8⁺TEM in BM and LN.

MDS plot of gene expression from CD8⁺T cell subsets from indicated tissues. Gene list derived DE (differential expression) of transcripts between LN and BM CD8⁺TEM cells (2,521 genes).

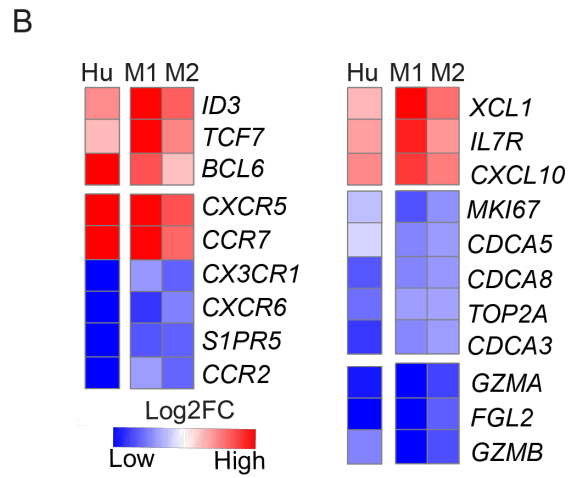
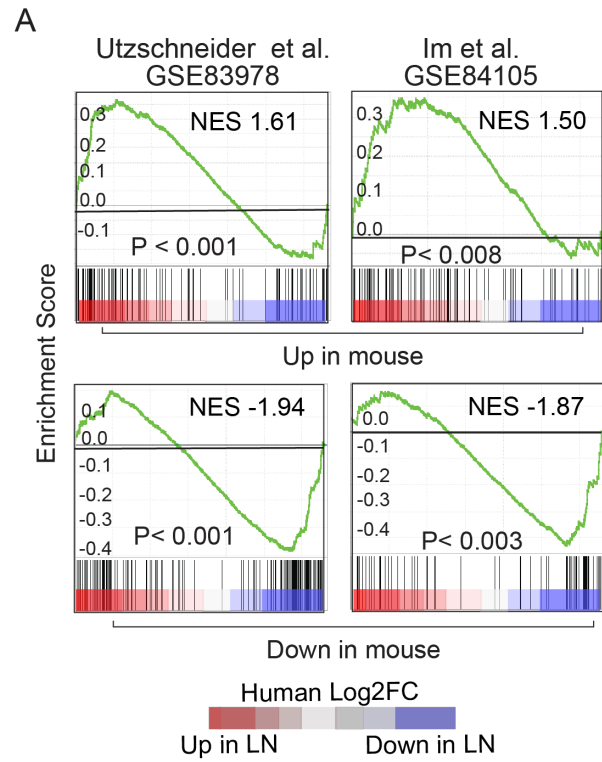
Bld= blood, BM= Bone marrow, LNG= Lung, SPL= Spleen, LLN= Lung draining lymph nodes.



We then considered whether the Tfh-like profile of human LN memory CD8⁺T cells could be similar to a recently defined subset of T cells in mice that was a PD-1⁺TCF-1⁺CXCR5⁺ memory CD8⁺T cell subset generated in response to chronic infection LCMV infection. This subset mediated viral clearance, and proliferated in response to anti-PD-1 checkpoint blockade.[154, 248, 303, 316, 317] By gene set enrichment analysis (GSEA)[304], we found significant homology in the complement of up- and downregulated genes between the human LN-specific gene signature and the gene signature of mouse TCF-1⁺ memory versus TCF-1- CD8⁺T cells from two independent studies [248, 303] (Figure 3-6A). Key genes shared between mouse and human subsets include those encoding transcriptional regulators (ID3, TCF7, BCL6), homing molecules (CXCR5, CCR7), cytokines, effector molecules (GZMB), and cell cycle regulators (CDCA3, CDC8) (Figure 3-6B). Human LN memory CD8⁺T cells therefore exhibit an organ-specific transcriptional profile similar to a T cell subset that maintains and restores immune responses to persistent pathogens in mice.

Figure 3-6. Human LN memory CD8⁺T cells exhibit similarity to murine CXCR5⁺CD8⁺ T cells generated in response to chronic LCMV infection.

(A) Comparison of mouse stem-like T cell transcriptome to human LN memory by gene set enrichment analysis. LN memory profile was generated from DE genes between LN and BM TEM cells. The list was ranked by log₂FC and shown on a color scale of red (high in LN), white (no change between BM and LN) to blue (low in LN). LN memory T cells are compared to ranked lists of up (top two plots) or down regulated (bottom two plots) genes between TCF-1⁺ versus TCF-1⁻ memory T cells generated in mouse models of chronic LCMV infection from two studies as indicated (Left, right) (12, 13). From the published studies we ranked the top 500 differentially expressed genes from microarray or the significantly differentially expressed genes from RNAseq (FDR ≤ 0.05, |Fold change| ≥ 1). (B) Heatmap showing relative expression (FC= Fold Change) of genes from human LN (Hu) versus BM CD8⁺TEM cells compared to mouse dataset 1 (M1 = GSE93978) and data set 2 (MS= GSE84105). NES= normalized enrichment score.



Defining an organ-specific protein signature of LN memory CD8⁺T cells by CyTOF

We used cytometry by time-of-flight (CyTOF) to define a protein expression signature of LN memory CD8⁺T cells relative to other sites based on the transcriptional profile identified above. We analyzed CD8⁺T cells from four tissue sites (LN, BM, spleen and lungs) of three individual donors using a 35 marker CyTOF panel including markers of T cell differentiation, function, proliferation as well as transcription factors (see Chapter 2 Methods, Table 2-2). Subset- and tissue-specific variations within and between donors were assessed using t-distributed stochastic neighbor embedding (t-SNE) analysis, a dimensionality reduction algorithm[297], on combined data for all tissues and donors. A total of 23 markers were used to generate the t-SNE plot with each the contribution of each individual marker and its expression density show in Figure 3-7. The concatenated analysis reveals that TCF-1 and CD28 segregate on distinct regions of the t-SNE plots relative to T-bet and Granzyme B expression (Figure 3-8A), which correspond to naïve (CD45RA⁺CCR7⁺) and TEMRA cells (CD45RA⁺CCR7⁻), respectively (Figure 3-8B). By contrast, CD8⁺TEM cells (Memory, CD45RA⁻CCR7⁻) cluster on broader regions of the tSNE plot with the higher density clusters on regions distinct from naive and TEMRA cells, indicating heterogeneity of memory CD8⁺T cells.

We dissected the heterogeneity of CD8⁺TEM cells by tissue site and individual. For both donors shown, CD8⁺TEM in BM and lung were most similar to each other and contained regions that overlapped with regions denoted by TEM and TEMRA cells and with splenic CD8⁺TEM cells (Figure 3-8B,C). In LN, CD8⁺TEM occupied a distinct region of the tSNE, showing overlap with CD8⁺TEM from spleen but not from BM and lung (Figure 3-8C). We further stratified memory CD8⁺T cells by CD69 expression to delineate putative tissue residency.[8] The tSNE profile of CD69⁺TEM cells was similar in the four tissue sites (LN, spleen, BM, lung) (Figure 3-8D), consistent with the finding that CD69⁺ tissue memory T cells exhibit a core signature of

TRM cells.[8] However, the tSNE profile of CD69-TEM, was distinct in LN and similar in spleen, lung and BM (Figure 3-8D), and the LN CD69-TEM cell profile was not observed in blood (Figure 3-8E). Interestingly the LN CD69-TEM cell profile was similar to LN CD69⁺TEM cells while in other sites, these two subsets were distinct (Figure 3-8D). Together, these findings indicate that LN memory CD8⁺T cells exhibit a protein expression signature that is organ specific, and not found in circulation or other tissues.

Figure 3-7. Expression of individual markers by total CD8⁺T cells concatenated from spleen, lung, BM, and LN of three donors in tSNE plots by CyTOF.

(A) Density plots show expression of each marker used in the t-SNE analysis. (B) Markers excluded from t-SNE analysis and used for gating of cell subsets. (C) Gating strategy used for T cell subset identification in Figure 3-8 and Figure 3-9.

Figure 3-7

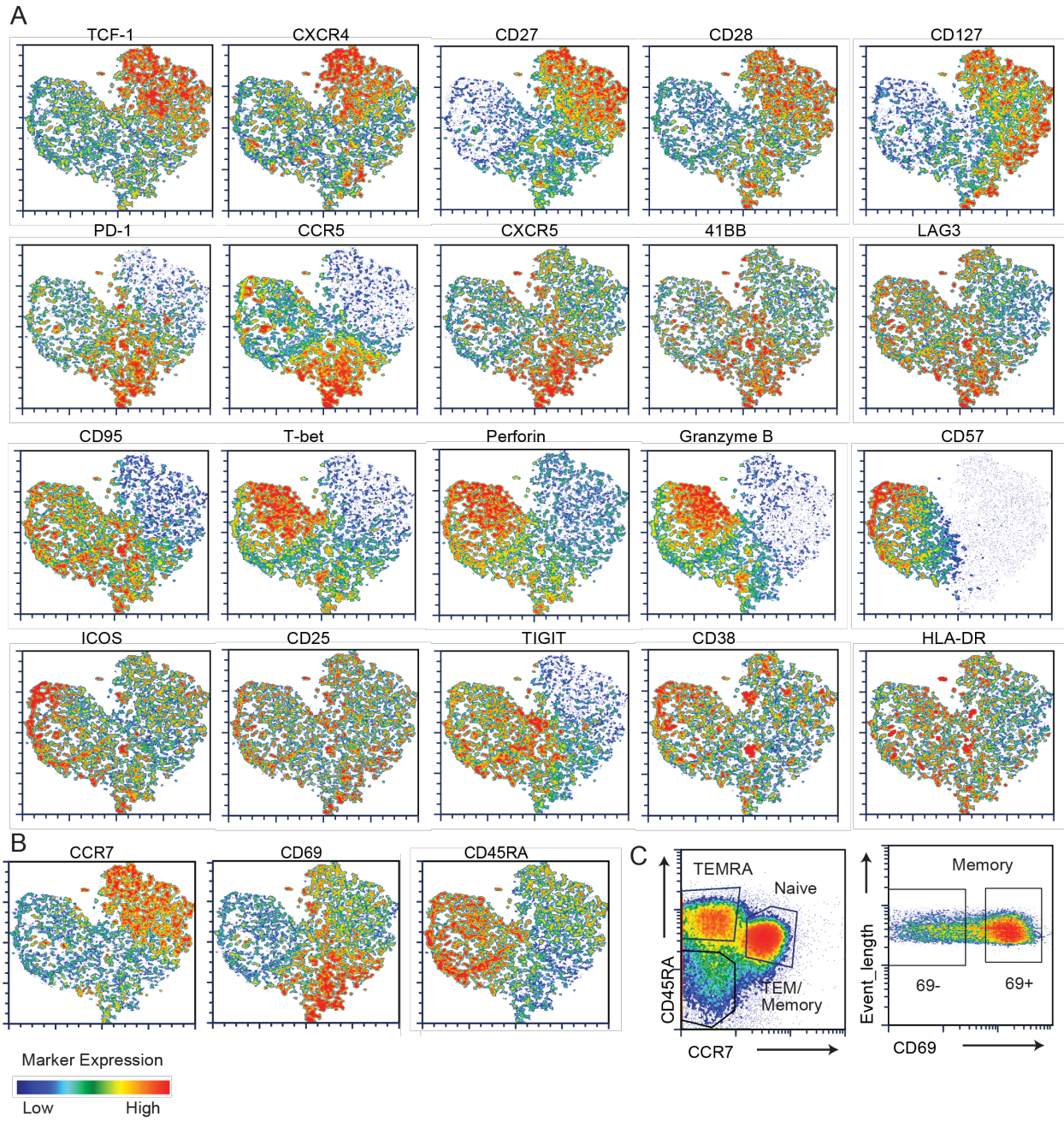


Figure 3-8. Human LN memory CD8⁺T cells exhibit a protein signature distinct from blood and other tissues.

T cells from LN, BM, spleen, lung +/- blood were stained with a 35-color panel (Table 2-2) and analyzed by CyTOF. (A) Expression of indicated markers on t-SNE plots of total CD8⁺T cells from BM, spleen, LN and lung of three individuals (D332, D333, and D335), based on 20 markers (CD57, CD28, Perforin, CD127, PD-1, ICOS, CD27, CCR5, TCF-1, CXCR5, 41BB, CD25, T-bet, CD38, CD95, LAG3, CXCR4, HLA-DR, TIGIT, and Granzyme-B; Fig. S2) (B) Density plots as in (A) gated on Naïve (CD45RA⁺CCR7⁺), Memory (CD45RA⁻CCR7⁻) and TEMRA (CD45RA⁺CCR7⁻) CD8⁺T cells. (C) Density plots as in (A) showing CD8⁺TEM (Donor 332, top row and Donor 335, bottom row). (D) Memory T cells from indicated tissues either CD69⁻ (top) or CD69⁺ (bottom). (E) Comparison of TEM CD69⁻ from blood and indicated tissues of one individual (D342) by t-SNE density plots based on markers in (A) minus TCF-1.

Figure 3-8

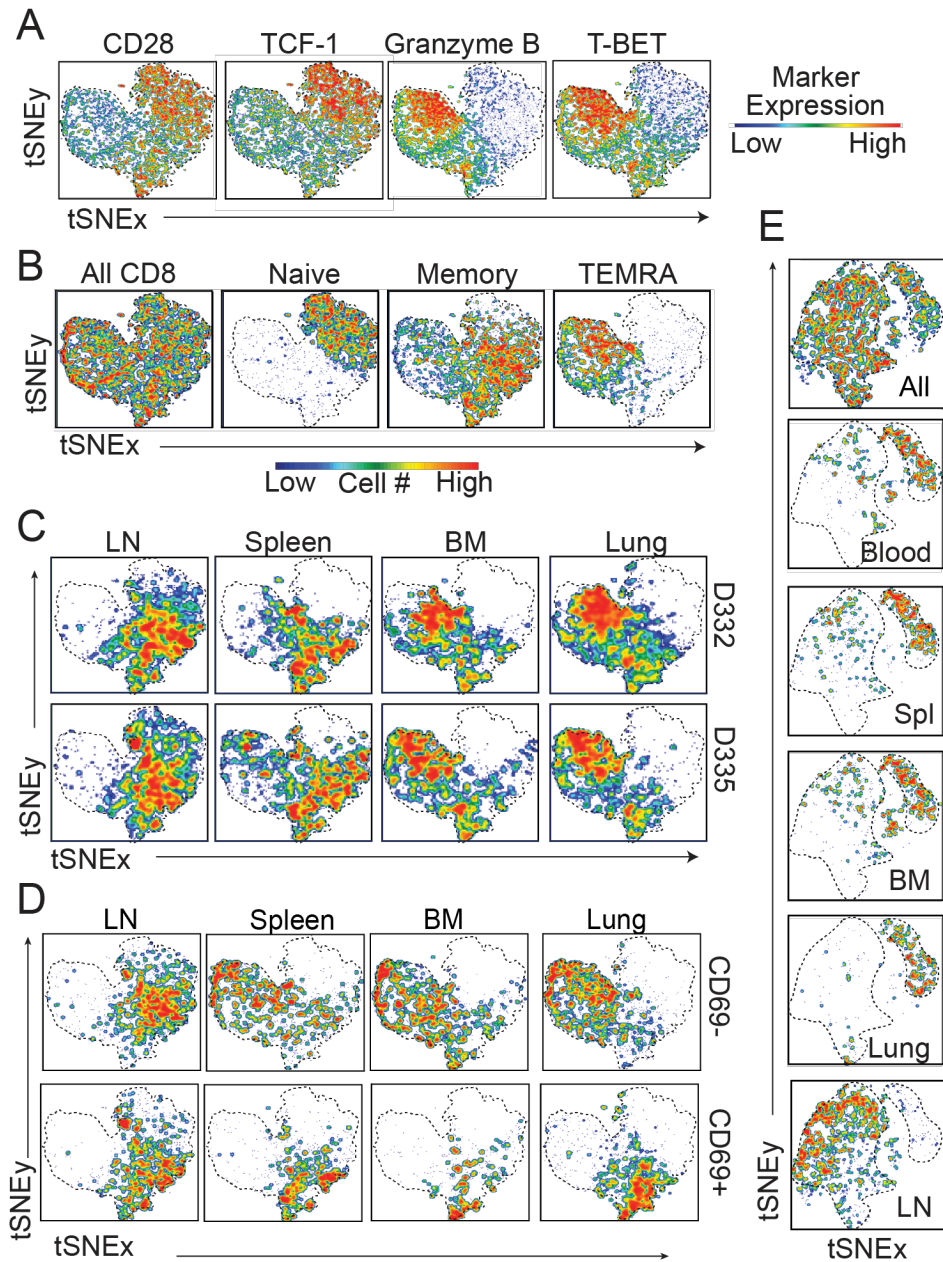
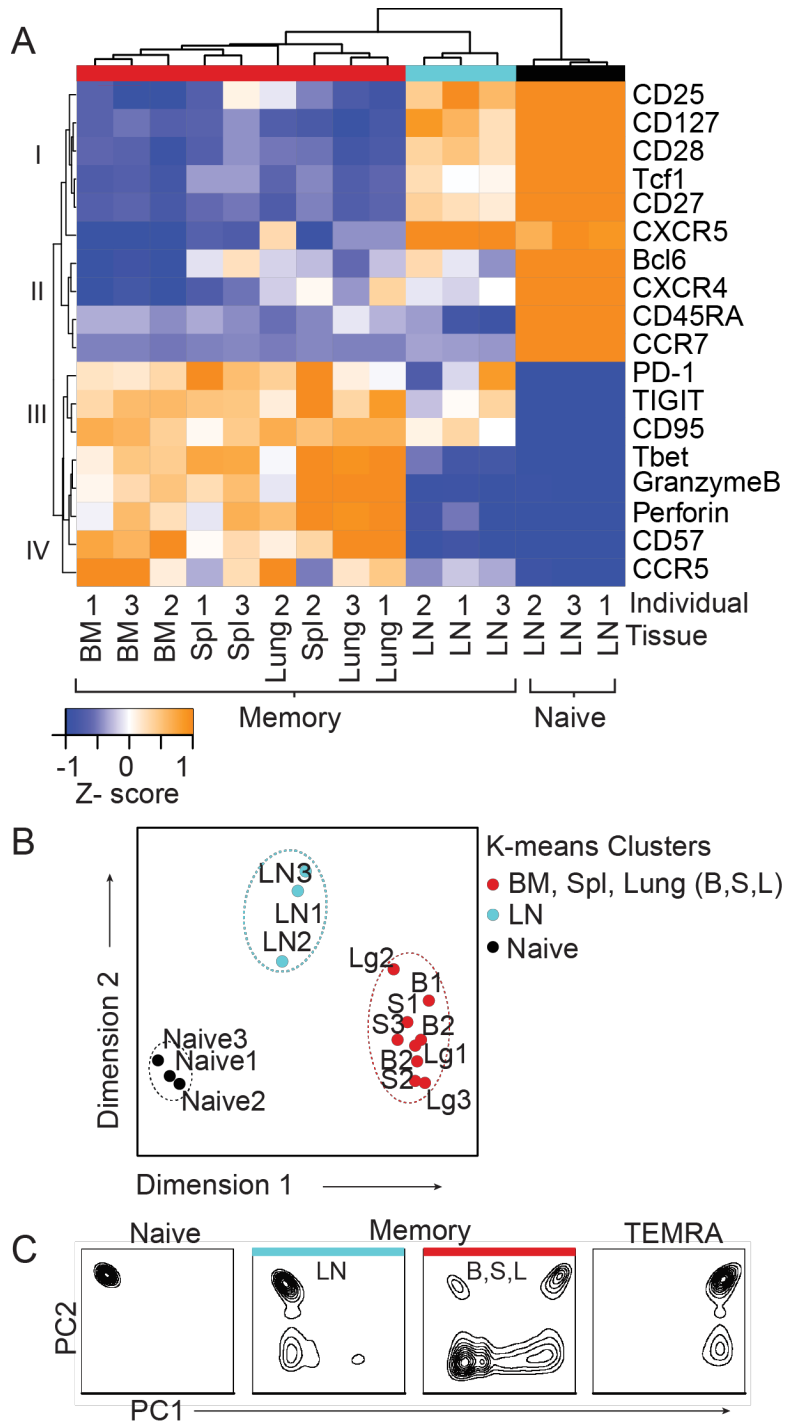


Figure 3-9. LN memory CD8⁺T cells exhibit a differentiation profile with features of both naïve and memory T cells.

(A) T cells from LN, BM, spleen, lung were stained with a 35-color panel (Table 2-2) and analyzed by CyTOF. Heatmap showing relative expression of indicated proteins by naïve and TEM CD69⁻ populations in tissues of three donors. (B) Multi-dimensional scaling (MDS) plot of T cell subsets by k means clustering. (C) PCA of indicated T cell subsets including total memory from LN (blue) and BM, spleen, and lung (red) and terminally differentiated effector cells (TEMRA) from all sites from donors as in (A). B= BM, S= Spl, L= Lung.

Figure 3-9



LN memory CD8⁺T cells exhibit differentiation signature with similarities to both naïve and memory T cells

To identify which markers were driving the LN-specific signature identified above, we used clustering analysis which revealed that LN memory CD8⁺T cells express a unique combination of markers—including those shared with naïve T cells (cluster I & IV), and others with memory CD8⁺T cells in blood and other sites (cluster II & III) (Figure 3-9A). Integrating the CyTOF data from multiple tissues and donors by MDS and K clustering analysis further shows that LN memory CD8⁺T cells have a protein expression signature that is distinct from naïve and memory T cells in other sites (Figure 3-9B). Principle component analysis (PCA) based on the protein expression data further reveals that LN TEM cells are more similar to naïve T cells, while tissue TEM are more similar to TEMRA cells representing the most differentiated subset (Figure 3-9C). These analyses show that LN memory CD8⁺T cells exhibit a less differentiated profile when compared to memory T cells in blood and other lymphoid and mucosal sites.

Figure 3-10. Sorting strategy for isolation of CD8⁺ TEM cells for proliferation assays.

(A) Gating strategy shown in the following order (left to right): lymphocytes, singlets, T cells (CD3⁺), CD8⁺, TEM (CCR7-CD45RA-) cells from lymph nodes.

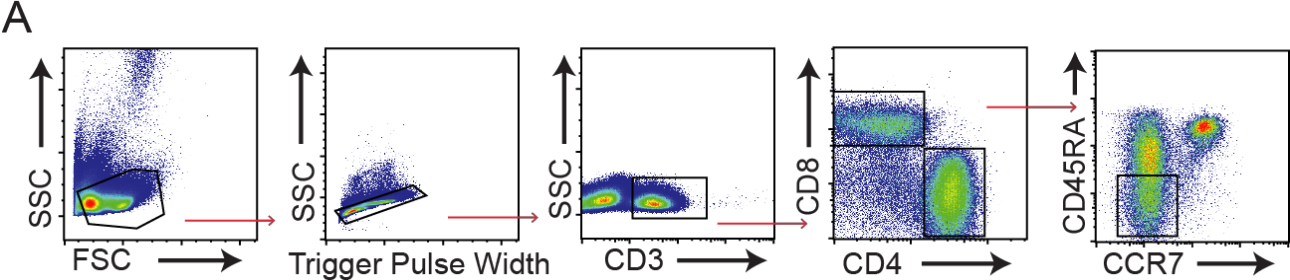
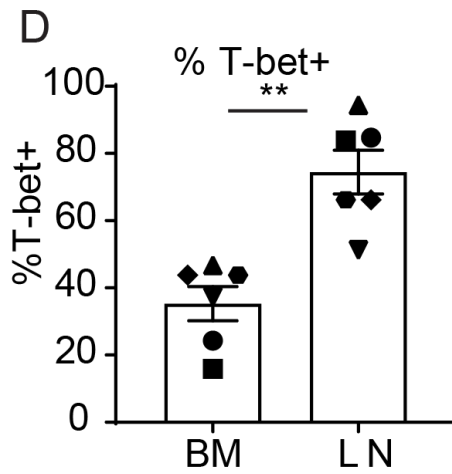
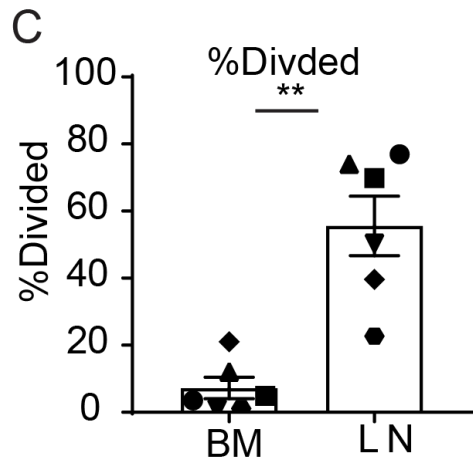
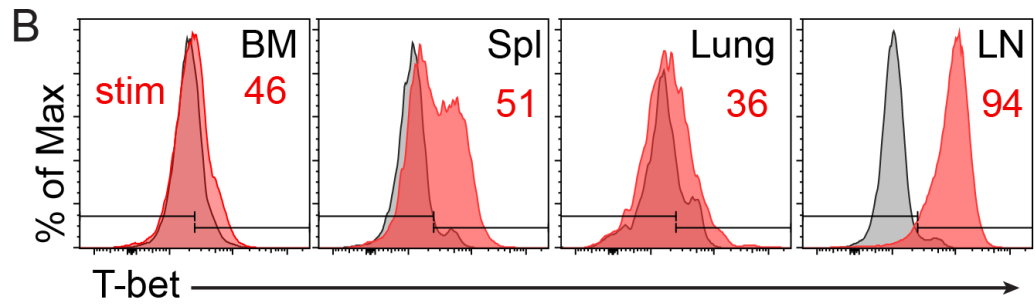
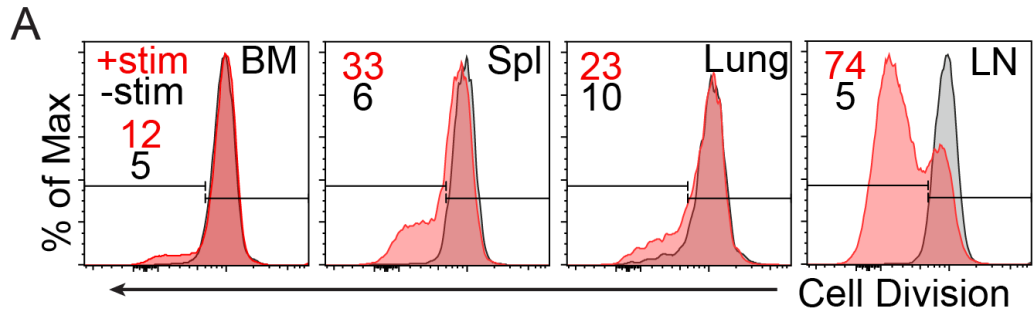


Figure 3-11. Human LN memory CD8⁺T cells exhibit enhanced proliferation and effector phenotype in response to TCR stimulation.

(A) Proliferation of CD8⁺TEM cells to anti-CD3/CD28 stimulation shown as %e-fluor450-negative (divided cells) (red) compared to unstimulated (black) from a representative individual (D377) (B) Expression of transcription factor T-bet shown as percent positive on stimulated T cells from indicated tissue sites from a representative individual (D377). (C) Data from A compiled from 6 individuals by percent divided (%eFluor450low) and (D) Data from B compiled from 6 individuals by percent positive for T-bet expression in stimulated cells. Error bars indicate SEM.* P<0.05, ** P<0.01, *** P<0.001 by two-tailed t-test.



Enhanced proliferative capacity of LN memory CD8⁺T cells

Functionally, TCF-1 expression is associated with a higher proliferative capacity to TCR-driven stimulation.[146] When stimulated with anti-CD3/anti-CD28 antibodies, LN memory CD8⁺T cells (see sorting strategy Figure 3-10) exhibited greater proliferation compared to counterparts in BM, spleen and lung(Figure 3-11A, C). After CD3/CD28 stimulation, LN CD8⁺TEM cell counts were between 98-106 fold higher than BM. T-bet expression, was also increased to a greater extent in LN CD8⁺TEM cells following stimulation compared to other sites both in the level of expression and the fold-change increase following activation (Figure 3-11B, D). Together, these results demonstrate that LN memory CD8⁺T cells have a higher capacity for proliferation and differentiation compared to memory CD8⁺T cells in other sites.

Increased TCR clonal diversity among LN compared to BM memory T cells

The results above indicate that LN maintain memory CD8⁺T cells exhibit a more quiescent state compared to other sites, including the BM which is a known reservoir for memory T cells of multiple specificities.[223] To assess whether there was differential maintenance of clonal populations between LN and BM memory T cells, we performed sequencing of the TCR CDR3 β variable (V) regions of memory CD8⁺T cells isolated from multiple donors. We found an increased number of different V region sequences from comparable numbers of LN compared to BM memory CD8⁺T cells for all donors (Table 3-4). Importantly, BM memory CD8⁺T cells consistently had higher clonality compared to LN memory CD8⁺T cells (Figure 3-12A). Conversely, an increase in TCR diversity was observed for LN compared to BM memory CD8⁺T cells for all donors even when controlling for clone size (Figure 3-12B). Together, these results show that memory CD8⁺T cells in LN represent a more

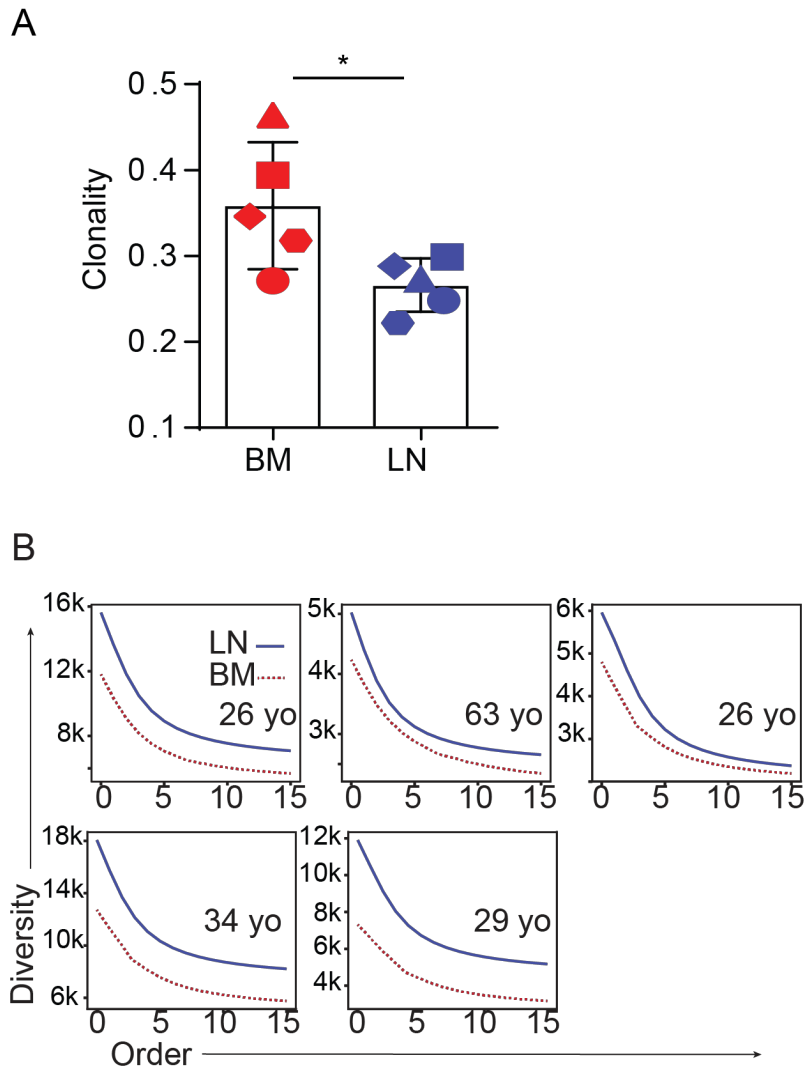
clonally diverse population than is maintained in blood and other tissue sites, consistent with our findings that LN memory CD8⁺T cells are maintained in a quiescent state.

Table 3-4. Number of clones identified by TCR sequencing of CD8+ TEM in BM and LN.

Donor	Tissue	Input DNA (ng)	# valid seq	# clones
D233	Bone Marrow	270.4	444267	5441
D233	Lymph node	270.4	522518	9195
D287	Bone Marrow	220	727259	8009
D287	Lymph node	220	552753	14248
D255	Bone Marrow	184	281123	2378
D255	Lymph node	184	420276	4605
D280	Bone Marrow	104	477615	4846
D280	Lymph node	104	504761	6694
D299	Bone Marrow	164	537904	2998
D299	Lymph node	164	772782	7705

Figure 3-12. Higher TCR diversity of CD8⁺TEM cells in LN compared to BM.

(A) Clonality of CD8⁺TEM TCR repertoires in BM and LN among 5 individuals. (B) Diversity of TCR sequences at different orders is shown for CD8⁺TEM cells from BM (dotted red) and LN (solid blue) among five individuals. At order zero, diversity is the total number of clones (richness). As order increases, the most abundant clones are given more weight.



Mechanisms for T cell maintenance in LN sites and the role of type I IFN signaling

To dissect mechanisms for the distinct gene expression profile of LN memory CD8⁺T cells, we analyzed pathways and upstream regulators. Pathway analysis showed conserved upregulated (Tfh, cell cycle, stem cell) and downregulated (T cell activation, signaling and inflammation) pathways in memory CD8⁺T cells from LN compared to BM, spleen and blood (Table 3-1). Potential upstream regulators of LN memory cells were associated with cell survival and growth factors (ie. FOXO1, NOTCH1, and EGFR), while molecules involving in interferon signaling and Type I IFN responses (ie. STAT1, IFNA2, IFNB1) were enriched in memory CD8⁺T cells derived from Blood, BM and spleen (Table 3-5). Therefore, LN memory CD8⁺T cells exhibit a distinct transcriptional program marked by quiescence and survival while memory CD8⁺T cells in BM, spleen, and blood exhibit cellular signatures associated with effector function and inflammatory signals.

Table 3-5. Potential upstream regulators of LN memory CD8⁺ T cells

Upstream Regulators ¹	Molecule Type	Direction in LN	Z-score
Ifnar	group	Inhibited	-3.27
IFN alpha/beta	group	Inhibited	-2.30
IFNA2	cytokine	Inhibited	-2.74
IFNB1	cytokine	Inhibited	-2.65
STAT1	transcription regulator	Inhibited	-2.67
STAT3	transcription regulator	Activated	1.48
FOXO1	transcription regulator	Activated	0.72
NOTCH1	transcription regulator	Activated	3.16
Vegf	group	Activated	3.47
HGF	growth factor	Activated	3.11
EGFR	kinase	Activated	3.48

¹Upstream regulators were identified as significantly enriched (p-value < 0.05). Z-score indicates predicted direction of regulator activity in memory CD8⁺ T cells from LN (inhibited or activated) compared to Blood, BM, and Spleen. Results were obtained from Ingenuity Pathway Analysis software (QIAGEN) by comparison of memory T cells isolated from different tissues (LN vs. BM, LN vs. Blood, and LN vs. Spleen).

IFN-I drives loss of TCF-1 expression in human memory CD8⁺ T cells

Following from the observation that IFN-I signaling was identified as a potential upstream regulator, specifically predicted as down regulated in LN memory T cells, we wondered how IFN-I regulates the proliferation of TCF-1⁺ CD8⁺ T cells in humans. We collected memory CD8⁺ T cells (CCR7⁻ CD45RA⁻ CD8⁺) from lymph nodes. The cells were labeled with Efluor450 and stimulated with anti-CD3 and anti-CD28 in the presence or absence of IFN-I as shown in Figure 3-13A. Interestingly, global proliferation was unaffected by IFN-I over the 5-day period (Figure 3-13B). However, there was a significant reduction in the number of proliferating cells that retained TCF-1 expression following IFN-I treatment (Figure 3-13C). These results show that IFN-I drives the loss of TCF-1 expression in proliferating CD8⁺ T cells.

Figure 3-13. IFN-I opposes maintenance of TCF-1 expression in proliferating CD8⁺ T cells.

(A) Experimental scheme for human lymph node CD8⁺ memory T cell treatment. (B) Percent of memory T cells from lymph nodes that divided in response to stimulation with CD3/CD28 antibodies treated with or without IFN- α/β . (C) Expression of TCF-1 with cell divisions of memory T cells from lymph nodes in response to stimulation with CD3/CD28 antibodies treated with or without IFN- α/β , measured by percent positive. Statistical comparison was performed using the ratio paired test. ** indicates $P < 0.01$ and not significant (NS) indicates $P > 0.05$.

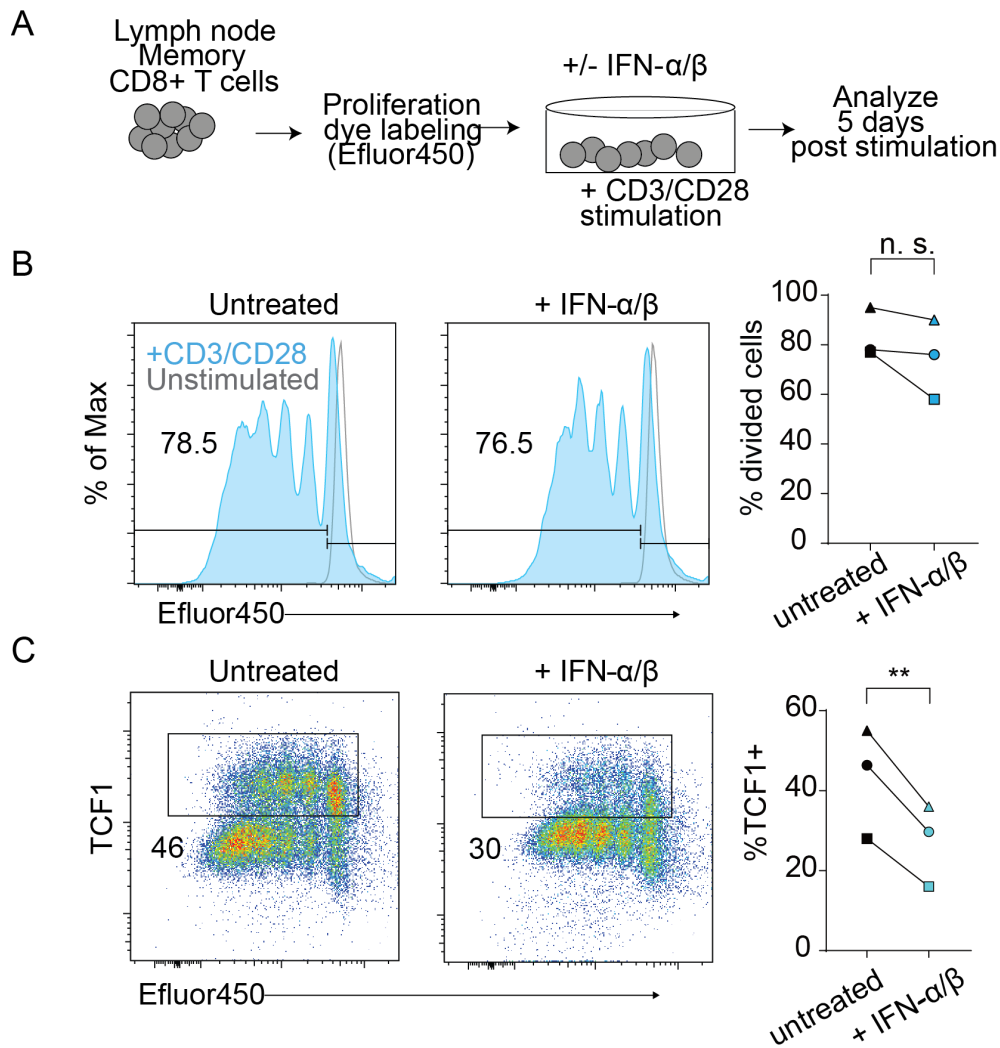


Table 3-6. List of organ donors used in this study.

Bld = Blood, Spl = Spleen, BM= Bone Marrow, LN= Lymph nodes.

Donor	Age	Sex	Tissue	Data
129	55	M	Bld, BM, Spl, LN, Lung	Flow
153	46	M	Bld, BM, Spl, LN, Lung	Flow
176	64	M	Bld, BM, Spl, LN, Lung	Flow
180	55	F	Bld, BM, Spl, LN, Lung	Flow
181	46	M	BM, LN, Spl	Flow
182	46	M	Bld, BM, Spl, LN, Lung	Flow
194	53	M	BM, LN	Flow
199	50	M	BM, LN	Flow
213	2	F	Bld, BM, Spl, LN, Lung	Flow
219	50	M	Bld, BM, Spl, LN, Lung	Flow
225	54	M	BM, LN	Flow
227	26	M	BM, LN	Flow
228	69	M	Bld, BM, Spl, LN, Lung	Flow
229	32	M	BM, LN	Flow, RNA-seq
230	52	F	BM, LN	Flow
232	53	M	Bld, BM, Spl, LN, Lung	Flow
233	26	F	Bld, BM, Spl, LN, Lung	Flow, TCR-seq
236	75	F	Bld, BM, Spl, LN, Lung	Flow
244	36	M	Bld, BM, Spl, LN, Lung	Flow
245	50	F	Bld, BM, Spl, LN, Lung	Flow
249	49	M	Bld, BM, Spl, LN, Lung	Flow
251	9	M	Bld, BM, Spl, LN, Lung	Flow
254	49	F	Bld, BM, Spl, LN, Lung	Flow
255	63	F	BM, LN	TCR-seq
259	46	M	Bld, BM, Spl, LN, Lung	Flow
262	73	M	Bld, BM, Spl, LN, Lung	Flow
273	67	F	Bld, BM, Spl, LN, Lung	Flow
275	31	M	Bld, BM, Spl, LN, Lung	Flow
277	21	M	BM, LN	Flow
280	26	M	BM, LN	RNA-seq, TCR-seq
287	34	M	Bld, BM, Spl, LN, Lung	Flow, RNA-seq, TCR-seq
288	32	M	Bld, BM, Spl, LN, Lung	Flow
289	58	M	Bld, BM, Spl, LN, Lung	Flow
291	26	F	Bld, BM, Spl, LN, Lung	Flow
293	53	F	BM, LN	Flow
297	59	F	Bld, BM, Spl, LN, Lung	Flow

299	20	M	Bld, BM, Spl, LN, Lung	Flow, TCR-seq
302	56	M	Bld, BM, Spl, LN, Lung	Flow
304	68	M	BM, LN	Flow
305	28	F	Bld, BM, Spl, LN, Lung	Flow
306	71	F	Bld, BM, Spl, LN, Lung	Flow
307	18	M	Bld, BM, Spl, LN, Lung	Flow
308	68	M	Bld, BM, Spl, LN, Lung	Flow
309	45	F	Bld, BM, Spl, LN, Lung	Flow
310	20	F	Bld, BM, Spl, LN, Lung	Flow
326	29	M	BM, LN	Flow
328	52	M	BM, LN	Flow
331	32	M	BM, LN	Flow
332	38	M	Bld, BM, Spl, LN, Lung	CYTOF
333	19	M	Bld, BM, Spl, LN, Lung	CYTOF
334	64	M	Bld, BM, Spl, LN, Lung	Flow
335	23	M	Bld, BM, Spl, LN, Lung	CYTOF, Flow
336	49	M	Bld, BM, Spl, LN, Lung	Flow
337	76	M	Bld, BM, Spl, LN, Lung	Flow
342	59	M	Bld, BM, Spl, LN, Lung	CYTOF
377	42	M	BM, Spl, LN, Lung	Flow

Section 3.3 Discussion

The majority of human T cells are found in lymphoid sites, although their function and relationship to blood T cells has been challenging to assess using conventional sampling. By examination of T cell subsets in human blood and tissue sites of individual organ donors, we demonstrate here that LN memory CD8⁺T cells exhibit organ-specific profiles not found in blood, other lymphoid (BM, spleen) or mucosal tissues. Specifically, LN memory T cells exhibit transcriptional and phenotypic signatures associated with quiescence, high proliferative capacity and enhanced TCR diversity compared to memory T cells in blood and other tissues which are enriched for effector function and inflammatory signatures. Our findings reveal that human LN serve as reservoirs for long-term maintenance of functional T cell responses and an important source for targeting in vaccines and immunotherapies.

We demonstrate here that memory T cells derived from LN maintain expression of the TCF-1 transcription factor associated with quiescence over decades of adult life. These results suggest differential priming of memory CD8⁺T cells that persist in LN compared to those that migrate and take up residence in other tissue sites. Recent studies in a mouse model of influenza infection showed differential TCF-1 expression by T cells primed in different LN sites: in the lung-draining LN there was rapid downregulation of TCF-1 expression, while TCF-1 expression was maintained by T cells primed to proliferate in a non-draining LN site, with reduced antigen and inflammatory signals.[168] We propose that memory T cells in LNs do not receive the full differentiation signals to promote LN egress and migration to the tissue site of infection, and therefore get retained and maintained with increased TCF-1 expression.

LN memory CD8⁺T cells exhibit a transcriptional and protein signature with homology to a mouse memory CD8⁺T cell subset generated during chronic virus infection which maintained functional capacity, proliferated and mediated viral clearance in response to anti-PD-1

therapy.[154, 248, 303, 316, 317] Whether human LN memory T cells provide similar roles in controlling chronic viral infection remains to be established; however, we previously showed that certain individuals with persistent cytomegalovirus (CMV) infection maintained CMV-specific CD8⁺T cells exclusively in LN, but not in circulation or other tissue sites[10].

Additionally, HIV-specific resident memory CD8⁺T cells were recently characterized in human LN.[275] These results suggest that human LNs can maintain reservoirs of anti-viral memory T cells. Moreover, the maintenance of higher TCR diversity among LN memory CD8⁺T cells suggest a unique role in maintaining functional T cell immunity over the lifespan with the potential to respond to different pathogens, of potential importance for targeting in vaccines.

More recent studies have found that in human lymph nodes of HIV-infected patients, CD8⁺ TRM appeared to be transcriptionally similar to bona fide mouse CD8⁺TRM, with phenotypic, functional and epigenetic signatures associated with residency [275]. Recent studies in secondary lymphoid tissues of mice and humans demonstrate that CD4⁺ and CD8⁺ TRM also take up residence in lymph nodes and spleen [7, 275-277]. However it remains unclear whether CD69⁺ T cells in lymph nodes are truly resident or maintained there temporarily with the ability to re-enter circulation. Importantly, CD69 expression by TCM cells does not necessarily confer tissue residence [211]. Additionally, CXCR4 is expressed on central memory T cells. Interestingly, in mouse model, CXCR4-deficient cells have impaired memory cell maintenance due to defective homeostatic proliferation, but are still able to re-expand and renew following antigen re-challenge. Perhaps CXCR4 plays a similar role in human T cell maintenance in lymph nodes, however this is yet to be tested.[322] Interestingly, in both human and mouse, LN TRM displayed increased proliferative potential; either in response to TCR stimulation in vitro or to checkpoint blockade immunotherapy respectively, denoting similar functional capacities. In humans, LN TRM populations also displayed high T cell receptor diversity compared to T cells

in BM, consistent with the possibility that these LN TRM may be reservoirs for protection against multiple antigens. If human LN TRM are similar to mouse, perhaps they also carry the ability to proliferate and re-enter circulation in response to checkpoint blockade therapy, an important line of research for future investigations. LN TRM may represent a niche for long term TRM maintenance and quiescence. Whether this is due mainly to sequestration of specific T cell populations from antigen within microanatomical niches of lymphoid tissue, or tissue specific signals for maintenance of these phenotypes remains unknown and an important area of investigation. These results are summarized by a graphical representation in Figure 3-14.

The mechanisms by which human LN maintain this distinct TCF-1⁺ memory subset are not known, although there is evidence that protection from IFN signaling may play a role. Transcriptionally, LN memory CD8⁺T cells exhibited downregulation of pathways involved type-1 interferon (IFN) signaling which were correspondingly upregulated in blood, BM, and spleen memory T cells. Furthermore, inhibiting type-1 IFN signaling during viral infection in mice, led to higher accumulation of TCF-1⁺ T cells in a T cell intrinsic manner.[154] Further, we found exogenous type-1 IFNs oppose the maintenance of TCF-1 expression in proliferating human T cells during TCR driven stimulation. Compared to mucosal and barrier sites which are constantly exposed to diverse microbial antigens, human LN are likely exposed to lower levels of inflammation and antigens, and thus provide protective niches for long-term maintenance of functional memory T cells.

Our results suggest that LN memory CD8⁺T cells are a potentially important source of functional T cells to target in immunotherapies for anti-tumor immunity. Given their homology to the mouse subset, LN memory T cells could analogously provide a source of highly proliferative T cells to respond to anti-PD-1 therapy. Memory T cells in LN draining sites of tumors could therefore be targeted for expansion of highly functional T cells to infiltrate the

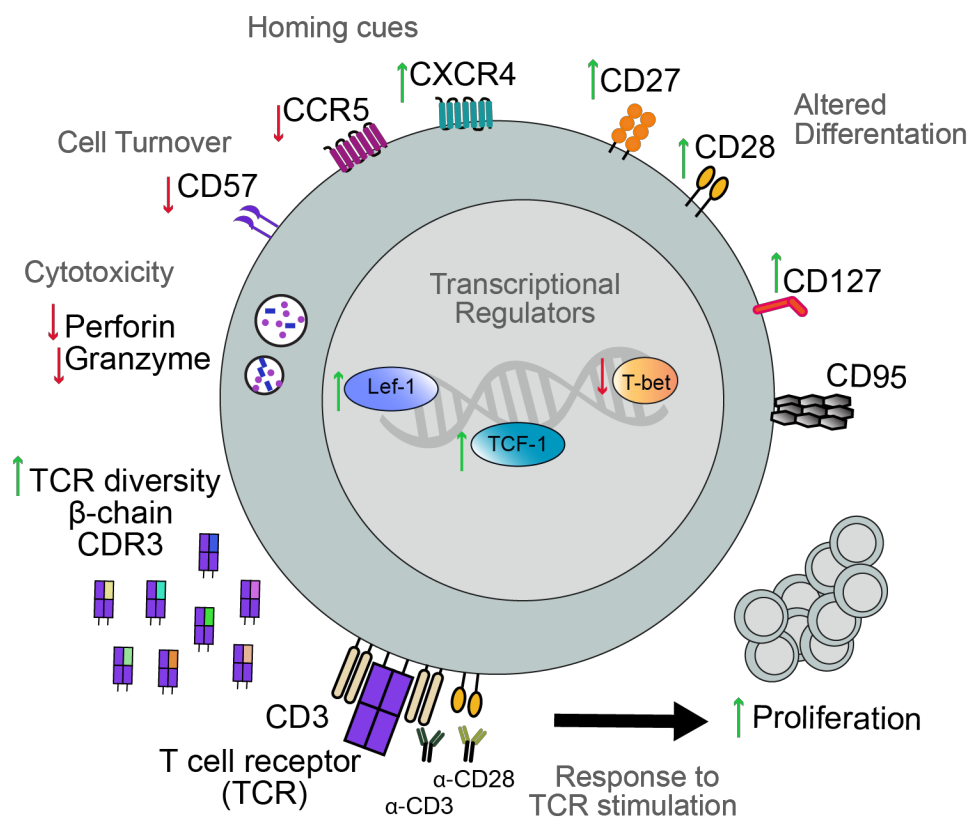
tumor site. Isolation of high potential memory subsets from blood has been a matter of considerable investigation in the field of adoptive cellular therapy, to optimize the persistence of expanded, differentiated cell populations[323], and to serve as a source of cells for transfection of targeted molecules for therapies in hematopoietic malignancies.[324] We propose that LN could be a more optimal source for memory T cells for adoptive immunotherapies compared to blood, based on their increased proliferative potential, and capacity to respond to diverse antigens.

In summary, our results provide evidence for long-term persistence of a distinct, organ-specific T cell subset in human LN. These findings demonstrate that tissue localization is a major determinant for human T cell differentiation, with LN providing long-term reservoirs for quiescent immunological memories important for targeting in vaccines and immunotherapies.

Figure 3-14. Summary of findings on human lymph node memory T cells.

We found lymph node memory T cells ($CD3^+CD8^+CCR7^-CD45RO^+CD45RA^-CD69^{+/-}$) differentially express several surface molecules and intracellular proteins relative to memory T cells in blood, BM, spleen and lung. Arrows indicate direction of differential expression in LN $CD8^+$ memory cells (green = upregulated, red = downregulated). Memory T cells in LN also displayed higher TCR diversity and proliferative response to TCR stimulation.

Human Lymph Node $CD8^+$ Memory T cells



CHAPTER 4: Subset-specific compartmentalization of human T cell receptor repertoires across blood and tissue sites

ABSTRACT:

T cell antigen receptor (TCR) activation leads to generation of diverse effector and memory subsets essential for immune protection. Here we investigated the relatedness of human CD4⁺ and CD8⁺ T cell subsets comprising central memory (TCM), effector memory (TEM), tissue-resident memory (TRM) and terminal effector cells (TEMRA) in diverse sites including blood, BM, Spl, LN and lungs by sequencing the β -chain of the TCR. We found that patterns of clonal expansion and distribution were predominantly subset-specific and to a lesser extent tissue-specific. Among subset-specific properties of T cells, we found that repertoires of CD4⁺TCM were the most diverse and CD8⁺TEMRA were the least diverse. We looked for tissue-specific features and found that repertoires of CD4⁺ and CD8⁺ T cells were more diverse in LN compared to BM, indicating tissue-specific reservoirs for T cells with previously low and high turnover respectively. We found substantial sharing of TEM and TRM cell clones across tissue sites; particularly abundant clones, suggesting systemic responses represent a significant portion of the T cell response in tissues. In summary, these results indicate the degree of clone sharing and diversity across sites is largely subset-specific, with implications for immune monitoring and modulation.

Section 4.1 Introduction

The ability of T lymphocytes to mediate protective immunity against pathogens depends on their specific recognition of pathogen-derived antigens, and their ability to migrate to diverse sites of pathogen encounter [179, 233]. Naïve T cells develop in the thymus and populate lymphoid tissues. Each naïve T cell expresses a unique T cell antigen receptor (TCR) comprised of α and β chains with variable CDR3 domains generated by gene rearrangement of *TCRa* and *TCRb* gene segments during T cell development. Naive T cells display a vast diversity of T cell receptor (TCR) proteins, conferring specificity to many possible antigen exposures with a theoretical diversity of $\sim 5 \times 10^{11}$ distinct β -chains [17]. When activated, antigen-specific naïve T cells clonally expand and differentiate into short-lived effector cells which acquire broad homing capacities for tissue-targeted migration. A fraction of these clonally expanded populations also develop into subsets of circulating and tissue-resident memory T cells that persist long-term and maintain protective responses. Mechanisms for the generation of circulating versus tissue-resident memory T cells and their lineage relationship remains unclear.

The pathways for differentiation of antigen-specific effector and memory T cells have been studied primarily in TCR-transgenic mouse models, where epitope-specific clones of T cells can be tracked. These studies have revealed that single clones of T cells can give rise to multiple lineages of effector and memory T cells [160, 162, 325], indicating that single T cell clones exhibit functional diversification during clonal expansion. However, it is not known whether certain clones or subsets are preferentially maintained or how tissue distribution is regulated on the clonal level. In humans, dissecting the clonal relatedness of T cell subsets has been challenging. Ex vivo studies have revealed that TCM can give rise to TEM phenotypes when activated [99]; however, assessing relatedness in vivo is limited by sampling constraints, with most studies on peripheral blood, representing only a small fraction ($\sim 3\%$) of total T cells

[4]. Assessing the relatedness of human T cell subsets in tissues and circulation has not been accomplished.

The primary subsets of CD4⁺ and CD8⁺ T cells in humans include central memory (TCM), effector memory (TEM), tissue-resident memory (TRM) and terminally differentiated effector cells (TEMRA).[99] TCM are distinguished from TEM based on expression of CCR7 and lymphoid homing capacities, while TRM are defined by expression of CD69 and CD103 which both function in tissue retention [7, 8, 121]. In humans, TRM cells are a non-circulating subset absent from peripheral blood, and have been identified in diverse sites such as bone marrow (BM), spleen (Spl), lymph nodes and lung and have distinct transcriptional properties compared to circulating T cells in blood and tissue to TRM [6, 8, 194, 196, 240, 275]. Although the role of TRM in human immune responses is not yet defined, they are implicated in a number of tissue-localized disease states, such as psoriasis in the skin and are associated with tumors specific to skin, liver, lung and breast tissues[326-329]. In humans little is known about how populations of previously activated T cells are distributed and maintained in blood and tissue sites throughout the body, important for optimal targeting in immunotherapies.

Next generation sequencing approaches have enabled the analysis of clonal populations within human blood and biopsy samples to high depth[330]. Here, we used CDR3 sequencing (along with sequencing V and J gene segments) in conjunction with validated computational approaches for identifying T cell clones to understand how clonally expanded effector and memory populations were maintained in human blood and tissue sites[309, 310]. We seek to understand the relatedness of human effector and memory T cell subsets on the basis of TCR sequence diversity and overlap. From diversity analysis we found that effector and memory subsets are maintained in a hierarchy of most diverse to least diverse (TCM > TEM and TRM > TEMRA) that is largely conserved across tissues and CD4⁺ and CD8⁺ T cell lineages. Overlap

analysis revealed the low and high relatedness of TCM and TEMRA cells respectively and this was highly conserved across tissues; in contrast, we found the relatedness of TEM and TRM was more dynamic across tissues. Taken together, these results reveal insights into human T cell differentiation in tissues and subset-specific patterns across sites.

Section 4.2 Results

Identifying T cell clones in lymphoid and non-lymphoid tissues

We isolated populations of previously primed T cell subsets comprising TCM, TEM, TEMRA and TRM cells from blood and tissue sites of 8 individual organ donors, along with corresponding populations in the blood of three living donors, resulting in 124 different biological samples (Figure 4-1A). From two individuals we obtained five sites (bone marrow (BM), spleen, lung and lymph nodes draining the lung (LN), +/- blood) and six cell subsets (gating strategy in Figure 4-2). The age and characteristics of individual donors are listed in Table 4-1. Total DNA was isolated from each purified cell subset and divided into two equal parts (biological replicates) from which separate PCR and library preparations were performed and sequenced (Figure 4-1B) from 239 samples in total (the amount of DNA per sample is listed in Table 4-2).

We used an analysis and storage tool designated ImmuneDB [309, 310], to identify V and J gene segments and clone counts from the sequencing data [309, 310]. We defined clonally related sequence as those which shared V and J gene segments and CDR3 amino acid sequence and we required that a specific sequence be detected at least twice in order to be designated a clone to reduce over estimation of clones due to sequencing errors.[331] We detected an average of 5,860 clones per CD4⁺ T cell sample and 3,438 clones per CD8⁺ T cell sample across all individuals (see Table 4-2 for individualized data).

Figure 4-1. Overview of samples collected and experimental design for TCR sequencing of T cell populations from human lymphoid and non-lymphoid tissues.

A) Origin and identity of T cell subsets purified from 11 individuals from blood and tissue sites including CD4⁺ TCM (central memory), CD4⁺ TEM (effector memory), CD4⁺ TRM (resident memory), and CD8⁺ TEMRA (terminally differentiated effector) and CD8⁺ TEM, CD8⁺ TRM cells. B) Schematic for sample preparation and TCR clone determination. DNA was extracted from purified memory T cells and subsequently separated into two equal parts (replicates) for sequencing. After PCR (with primers targeting the TCR V β -chain) and sequencing, the data were processed with ImmuneDB (13) to identify clones by unique V and J gene segment pairing and CDR3 amino acid sequence.

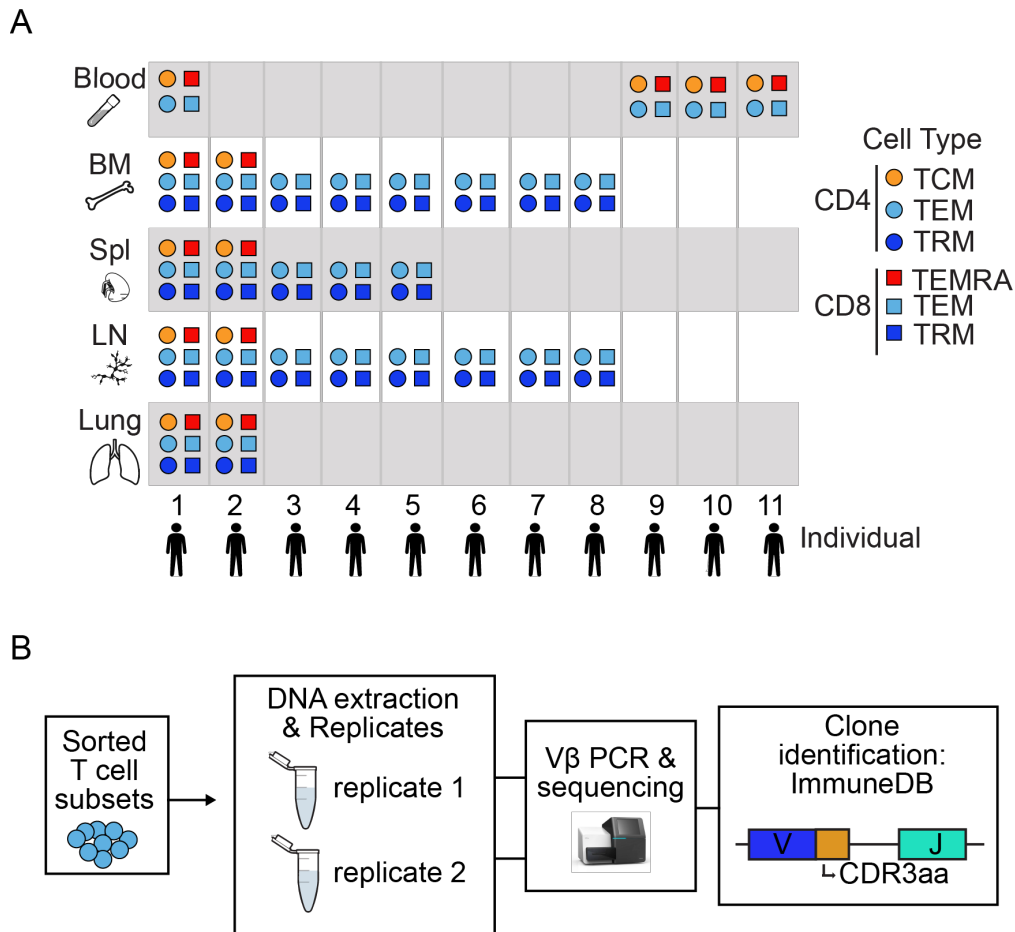


Figure 4-2. Gating strategy for T cell subsets isolated for TCR sequencing.

First, we identified lymphocytes by FSC-A and SSC-A, singlets by trigger cell width (not shown), and T cells by CD3⁺. Subsequently, we fractionated total T cells by CD4⁺CD8⁻ or CD8⁺CD4⁻ cells (negative gates not shown). Next, CD4⁺ T cells were gated on to identify TCM cells (CCR7⁺CD45RA⁻) and CD8⁺ T cells were gated on to identify TEMRA cells (CCR7⁺CD45RA⁺). Finally, both CD4⁺ and CD8⁺ T cells were further gated to identify TEM (CCR7⁻CD45RA⁻CD69⁻), and TRM (CCR7⁻CD45RA⁻CD69⁺) cells.

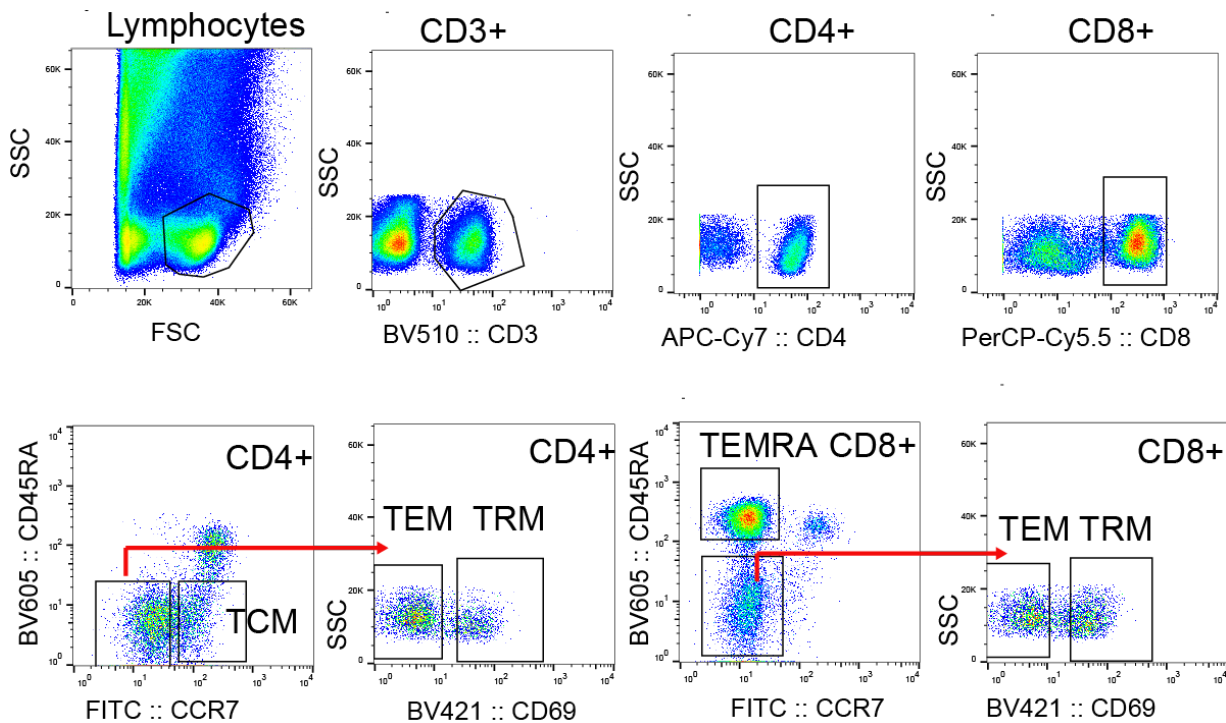


Table 4-1. Age and characteristics of donors in this study.

Information about individual organ donors used in this study including sex (M=male, F=female), age, cause of death (COD), and serology for chronic viruses CMV and EBV. N/A= Not applicable.

Individual	Donor	Age	Sex	Race	COD	CMV	EBV
1	383	39	M	White	Head trauma	0	1
2	324	56	M	White	Stroke	1	1
3	299	29	M	Hispanic	Anoxia	1	1
4	255	63	F	Hispanic	Stroke	1	1
5	233	26	F	White	Anoxia	0	1
6	287	34	M	White	Head trauma	1	1
7	280	26	M	White	Anoxia	0	1
8	229	32	M	White	Anoxia	0	1
9	LD1	55	F	White	N/A	-	-
10	LD2	32	F	White	N/A	-	-
11	LD3	29	M	White	N/A	-	-

Table 4-2. Number of T cell clones identified per sample by TCR sequencing.

The number of T cell clones identified by TCR sequencing for individual samples. Sample name indicates the donor number, anatomical site, lineage, subset, and replicate number of each sample respectively (example: D229_BM_CD4_TEM_1).

id	sample name	clones	DNA (ng)	id	sample name	clones	DNA (ng)
1	D229_BM_CD4_TEM_1	6413	23	51	D255_LN_CD4_TRM_2	8791	46
2	D229_BM_CD4_TEM_2	5458	23	52	D255_LN_CD8_TEM_1	2559	46
3	D229_BM_CD4_TRM_1	4417	11	53	D255_LN_CD8_TEM_2	3735	46
4	D229_BM_CD4_TRM_2	4609	11	54	D255_LN_CD8_TRM_1	1887	46
5	D229_BM_CD8_TEM_1	4603	23	55	D255_LN_CD8_TRM_2	2731	46
6	D229_BM_CD8_TEM_2	4732	23	56	D255_Spl_CD4_TEM_1	2221	46
7	D229_BM_CD8_TRM_1	5637	23	57	D255_Spl_CD4_TEM_2	3379	46
8	D229_BM_CD8_TRM_2	5102	23	58	D255_Spl_CD4_TRM_1	5144	46
9	D229_LN_CD4_TEM_1	6114	23	59	D255_Spl_CD4_TRM_2	12088	46
10	D229_LN_CD4_TEM_2	8482	23	60	D255_Spl_CD8_TEM_1	1791	46
11	D229_LN_CD4_TRM_1	5999	23	61	D255_Spl_CD8_TEM_2	2588	46
12	D229_LN_CD4_TRM_2	7489	23	62	D255_Spl_CD8_TRM_1	2163	46
13	D229_LN_CD8_TEM_1	5403	23	63	D255_Spl_CD8_TRM_2	3163	46
14	D229_LN_CD8_TEM_2	4809	23	64	D280_BM_CD4_TEM_1	6619	26
15	D229_LN_CD8_TRM_1	5285	23	65	D280_BM_CD4_TEM_2	7517	26
16	D229_LN_CD8_TRM_2	4938	23	66	D280_BM_CD4_TRM_1	1661	26
17	D233_BM_CD4_TEM_1	7054	67	67	D280_BM_CD4_TRM_2	1398	26
18	D233_BM_CD4_TEM_2	5162	67	68	D280_BM_CD8_TEM_1	1803	26
19	D233_BM_CD4_TRM_1	6543	67	69	D280_BM_CD8_TEM_2	2402	26
20	D233_BM_CD4_TRM_2	4373	67	70	D280_BM_CD8_TRM_1	2172	26
21	D233_BM_CD8_TEM_1	2972	67	71	D280_BM_CD8_TRM_2	2012	26
22	D233_BM_CD8_TEM_2	2798	67	72	D280_LN_CD4_TEM_1	2313	26
23	D233_BM_CD8_TRM_1	3455	67	73	D280_LN_CD4_TEM_2	1768	26
24	D233_BM_CD8_TRM_2	2489	67	74	D280_LN_CD4_TRM_1	3240	26
25	D233_LN_CD4_TEM_2	7570	67	75	D280_LN_CD4_TRM_2	3597	26
26	D233_LN_CD4_TRM_1	6064	67	76	D280_LN_CD8_TEM_1	2453	26
27	D233_LN_CD4_TRM_2	6968	67	77	D280_LN_CD8_TEM_2	2202	26
28	D233_LN_CD8_TEM_1	4093	67	78	D280_LN_CD8_TRM_1	1821	26
29	D233_LN_CD8_TEM_2	3161	67	79	D280_LN_CD8_TRM_2	2187	26
30	D233_LN_CD8_TRM_1	4556	67	80	D287_BM_CD4_TEM_1	10379	59
31	D233_LN_CD8_TRM_2	3411	67	81	D287_BM_CD4_TEM_2	11459	59
32	D233_Spl_CD4_TEM_1	4520	67	82	D287_BM_CD4_TRM_1	5917	50
33	D233_Spl_CD4_TEM_2	6931	67	83	D287_BM_CD4_TRM_2	6650	50
34	D233_Spl_CD4_TRM_1	6622	67	84	D287_BM_CD8_TEM_1	3724	69
35	D233_Spl_CD4_TRM_2	6521	67	85	D287_BM_CD8_TEM_2	4205	69
36	D233_Spl_CD8_TEM_1	3887	67	86	D287_BM_CD8_TRM_1	6653	35
37	D233_Spl_CD8_TEM_2	3018	67	87	D287_BM_CD8_TRM_2	6051	35
38	D233_Spl_CD8_TRM_1	2353	67	88	D287_LN_CD4_TEM_1	12303	63
39	D233_Spl_CD8_TRM_2	4359	67	89	D287_LN_CD4_TEM_2	11737	63
40	D255_BM_CD4_TEM_1	1816	46	90	D287_LN_CD4_TRM_1	11072	73
41	D255_BM_CD4_TEM_2	3494	46	91	D287_LN_CD4_TRM_2	12688	73
42	D255_BM_CD4_TRM_1	4051	46	92	D287_LN_CD8_TEM_1	7935	55
43	D255_BM_CD4_TRM_2	5623	46	93	D287_LN_CD8_TEM_2	8456	55
44	D255_BM_CD8_TEM_1	354	46	94	D287_LN_CD8_TRM_1	4454	14
45	D255_BM_CD8_TEM_2	2836	46	95	D287_LN_CD8_TRM_2	5407	14
46	D255_BM_CD8_TRM_1	2573	46	96	D299_BM_CD4_TEM_1	4608	41
47	D255_BM_CD8_TRM_2	3709	46	97	D299_BM_CD4_TEM_2	4005	41
48	D255_LN_CD4_TEM_1	4952	46	98	D299_BM_CD4_TRM_1	5815	41
49	D255_LN_CD4_TEM_2	6681	46	99	D299_BM_CD4_TRM_2	5627	41
50	D255_LN_CD4_TRM_1	6357	46	100	D299_BM_CD8_TEM_1	2924	41

id	sample name	clones	DNA (ng)	id	sample name	clones	DNA (ng)
101	D299_BM_CD8_TEM_2	1425	41	163	D324_Spl_CD8_TRM_2	635	15
102	D299_BM_CD8_TRM_1	3691	41	164	D383_Bld_CD4_TCM_1	5092	33
103	D299_BM_CD8_TRM_2	3631	41	165	D383_Bld_CD4_TCM_2	4889	33
104	D299_LN_CD4_TEM_1	4493	41	166	D383_Bld_CD4_TEM_1	3480	50
105	D299_LN_CD4_TEM_2	4443	41	167	D383_Bld_CD4_TEM_2	3707	50
106	D299_LN_CD4_TRM_1	4297	41	168	D383_Bld_CD8_TEM_1	4605	20
107	D299_LN_CD4_TRM_2	3846	41	169	D383_Bld_CD8_TEM_2	4865	20
108	D299_LN_CD8_TEM_1	4314	41	170	D383_Bld_CD8_TEMRA_1	910	2
109	D299_LN_CD8_TEM_2	4515	41	171	D383_Bld_CD8_TEMRA_2	661	2
110	D299_LN_CD8_TRM_1	5187	41	172	D383_BM_CD4_TCM_1	7022	50
111	D299_LN_CD8_TRM_2	5018	41	173	D383_BM_CD4_TCM_2	5628	50
112	D299_Spl_CD4_TEM_1	4561	41	174	D383_BM_CD4_TEM_1	6469	50
113	D299_Spl_CD4_TEM_2	4431	41	175	D383_BM_CD4_TEM_2	6224	50
114	D299_Spl_CD4_TRM_1	4921	41	176	D383_BM_CD4_TRM_1	4521	50
115	D299_Spl_CD4_TRM_2	4654	41	177	D383_BM_CD4_TRM_2	4948	50
116	D299_Spl_CD8_TEM_1	4008	41	178	D383_BM_CD8_TEM_1	1115	50
117	D299_Spl_CD8_TEM_2	3834	41	179	D383_BM_CD8_TEM_2	3475	50
118	D299_Spl_CD8_TRM_1	4270	41	180	D383_BM_CD8_TEMRA_1	2610	50
119	D299_Spl_CD8_TRM_2	4321	41	181	D383_BM_CD8_TEMRA_2	2192	50
120	D324_BM_CD4_TCM_1	4964	15	182	D383_BM_CD8_TRM_1	2640	50
121	D324_BM_CD4_TCM_2	4995	15	183	D383_BM_CD8_TRM_2	2532	50
122	D324_BM_CD4_TEM_1	5166	15	184	D383_LN_CD4_TCM_1	8538	50
123	D324_BM_CD4_TEM_2	4849	15	185	D383_LN_CD4_TCM_2	5032	50
124	D324_BM_CD4_TRM_1	4075	15	186	D383_LN_CD4_TEM_1	13606	50
125	D324_BM_CD8_TEM_1	831	12	187	D383_LN_CD4_TEM_2	11627	50
126	D324_BM_CD8_TEM_2	1926	12	188	D383_LN_CD4_TRM_1	11556	50
127	D324_BM_CD8_TEMRA_1	666	15	189	D383_LN_CD4_TRM_2	4359	50
128	D324_BM_CD8_TEMRA_2	658	15	190	D383_LN_CD8_TEM_1	6416	50
129	D324_BM_CD8_TRM_1	1544	11	191	D383_LN_CD8_TEM_2	6095	50
130	D324_BM_CD8_TRM_2	2466	11	192	D383_LN_CD8_TRM_1	6643	50
131	D324_LN_CD4_TCM_1	5421	15	193	D383_LN_CD8_TRM_2	3674	50
132	D324_LN_CD4_TCM_2	5120	15	194	D383_Lung_CD4_TCM_1	6870	50
133	D324_LN_CD4_TEM_1	5872	15	195	D383_Lung_CD4_TCM_2	7105	50
134	D324_LN_CD4_TEM_2	5505	15	196	D383_Lung_CD4_TEM_1	5334	50
135	D324_LN_CD4_TRM_1	5565	15	197	D383_Lung_CD4_TEM_2	5058	50
136	D324_LN_CD4_TRM_2	5166	15	198	D383_Lung_CD4_TRM_1	5963	50
137	D324_LN_CD8_TEM_1	3697	13	199	D383_Lung_CD4_TRM_2	3757	50
138	D324_LN_CD8_TEM_2	3274	13	200	D383_Lung_CD8_TEM_1	9858	50
139	D324_LN_CD8_TRM_1	2533	15	201	D383_Lung_CD8_TEM_2	9715	50
140	D324_LN_CD8_TRM_2	2668	15	202	D383_Lung_CD8_TEMRA_1	2557	50
141	D324_Lung_CD4_TCM_1	3578	15	203	D383_Lung_CD8_TEMRA_2	2137	33
142	D324_Lung_CD4_TCM_2	3553	15	204	D383_Lung_CD8_TRM_1	3787	50
143	D324_Lung_CD4_TEM_1	4550	15	205	D383_Lung_CD8_TRM_2	3845	50
144	D324_Lung_CD4_TEM_2	4144	15	206	D383_Spl_CD4_TCM_1	7641	50
145	D324_Lung_CD4_TRM_1	4575	15	207	D383_Spl_CD4_TCM_2	6868	50
146	D324_Lung_CD4_TRM_2	4439	15	208	D383_Spl_CD4_TEM_1	8150	50
147	D324_Lung_CD8_TEM_2	3898	15	209	D383_Spl_CD4_TEM_2	6816	50
148	D324_Lung_CD8_TEMRA_1	1027	15	210	D383_Spl_CD4_TRM_1	7139	50
149	D324_Lung_CD8_TEMRA_2	1314	15	211	D383_Spl_CD4_TRM_2	7094	50
150	D324_Lung_CD8_TRM_1	3577	15	212	D383_Spl_CD8_TEMRA_1	3364	50
151	D324_Lung_CD8_TRM_2	3158	15	213	D383_Spl_CD8_TEMRA_2	2372	50
152	D324_Spl_CD4_TCM_1	5965	15	214	D383_Spl_CD8_TMEM_1	5933	50
153	D324_Spl_CD4_TCM_2	5570	15	215	D383_Spl_CD8_TMEM_2	2116	50
154	D324_Spl_CD4_TEM_1	6169	15	216	LD1_Bld_CD4_TCM_1	10234	65
155	D324_Spl_CD4_TEM_2	5920	15	217	LD1_Bld_CD4_TCM_2	6699	65
156	D324_Spl_CD4_TRM_1	6161	15	218	LD1_Bld_CD4_TEM_1	5785	65
157	D324_Spl_CD4_TRM_2	6043	15	219	LD1_Bld_CD4_TEM_2	5184	65
158	D324_Spl_CD8_TEM_1	5072	15	220	LD1_Bld_CD8_TEM_1	2581	65
159	D324_Spl_CD8_TEM_2	4525	15	221	LD1_Bld_CD8_TEM_2	2607	65
160	D324_Spl_CD8_TEMRA_1	1365	15	222	LD1_Bld_CD8_TEMRA_1	1597	65
161	D324_Spl_CD8_TEMRA_2	1128	15	223	LD1_Bld_CD8_TEMRA_2	1666	65
162	D324_Spl_CD8_TRM_1	3190	15	224	LD2_Bld_CD4_TCM_1	9024	65

id	sample name	clones	DNA (ng)
225	LD2_Bld_CD4_TCM_2	8857	65
226	LD2_Bld_CD4_TEM_1	7488	65
227	LD2_Bld_CD4_TEM_2	4919	65
228	LD2_Bld_CD8_TEM_1	7295	65
229	LD2_Bld_CD8_TEM_2	6790	65
230	LD2_Bld_CD8_TEMRA_1	5929	65
231	LD2_Bld_CD8_TEMRA_2	3834	65
232	LD3_Bld_CD4_TCM_1	6689	65

id	sample name	clones	DNA (ng)
233	LD3_Bld_CD4_TCM_2	6654	65
234	LD3_Bld_CD4_TEM_1	9012	65
235	LD3_Bld_CD4_TEM_2	7883	65
236	LD3_Bld_CD8_TEM_1	2305	65
237	LD3_Bld_CD8_TEM_2	2507	65
238	LD3_Bld_CD8_TEMRA_1	1181	65
239	LD3_Bld_CD8_TEMRA_2	1106	65

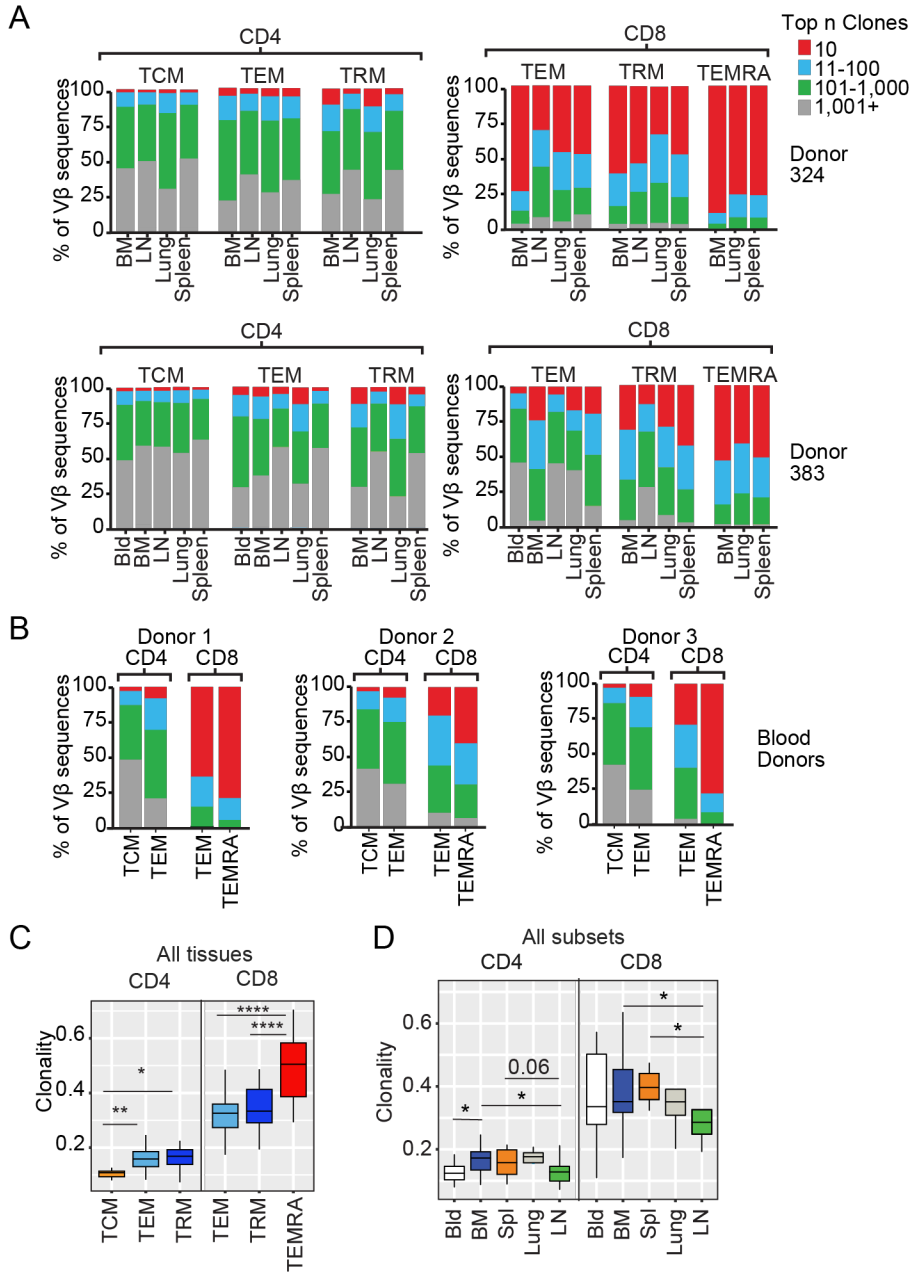
Evaluating diversity of T cell subsets from blood and tissues

We initially assessed the extent of TCR diversity and clonal expansion within each subset and tissue. To analyze the diversity of T cell clones within each sample we ranked all the clones from most expanded to least, and grouped clones based on their abundance into top 1-10, 11-100, 101-1,000, and 1,0001+ clones. We then calculated the percentage of the total sample repertoire that was represented within each clone ranked group. There were marked differences between CD4⁺ and CD8⁺ T cell subsets in terms of their overall clonal abundance and the proportion of total T cells that were represented by top clones. For CD4 T cells, the top 10 clones comprised 2-20% of total T cell repertoire; by contrast, for CD8 T cells, the top 10 clones comprised up to 80% of total T cell repertoire (Figure 4-3).

The proportion of the repertoire represented by the top 10 clones was consistent between subsets within CD4 and CD8 lineages. By different cell subsets, we found considerable differences between subsets that were conserved across tissues; the percentage of the top 10 clones for all subsets were as follows: ~5% for CD4⁺ TCM, ~25% for CD4⁺ TEM and TRM, between 30-75% for CD8⁺ TEM and TRM, and >75% for CD8⁺ TEMRA. The remaining three groups (top 11-100, 101-1,000, and 1,0001+ clones), followed similar patterns to the first groups (Figure 4-3). Remarkably, both individuals exhibited a similar hierarchy of clonal expansions from highest to lowest: TEMRA > TEM and TRM > TCM. We could not acquire blood for all subsets from organ donors, therefore we obtained blood from three living volunteers and found similar results; the percentage of the top clonal expansions accounted for ~5% of CD4⁺ TCM, ~15% of CD4⁺ TEM, between 25-70% of CD8⁺ TEM, and up to 80% of CD8⁺ TEMRA cell repertoires.

Figure 4-3. T cell clone abundances attributed mainly to subset differences.

A) Proportion of top n clones per sample for CD4⁺ (left) and CD8⁺ (right) memory T cells subsets (TCM, TEM, TRM and TEMRA) from two individual donors 324 (top) and 383 (bottom) and B) same from three live blood donors. C) Clonality across individuals by cell subset (all tissues combined) (CD4⁺: TCM, n=12; TEM, n=26; TRM, n=23; CD8⁺: TEM, n=25; TRM, n=22; TEMRA, n=10). D) or by tissue site (all subsets combined) for CD4⁺ and CD8⁺ T cells. (CD4⁺: Bld, n=8, BM, n=18; Spl, n=12; Lung, n=6; LN, n=17; CD8⁺: Bld, n=8; BM, n=18; Spl, n=10; Lung, n=5; LN, n=16). Students T-test. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.



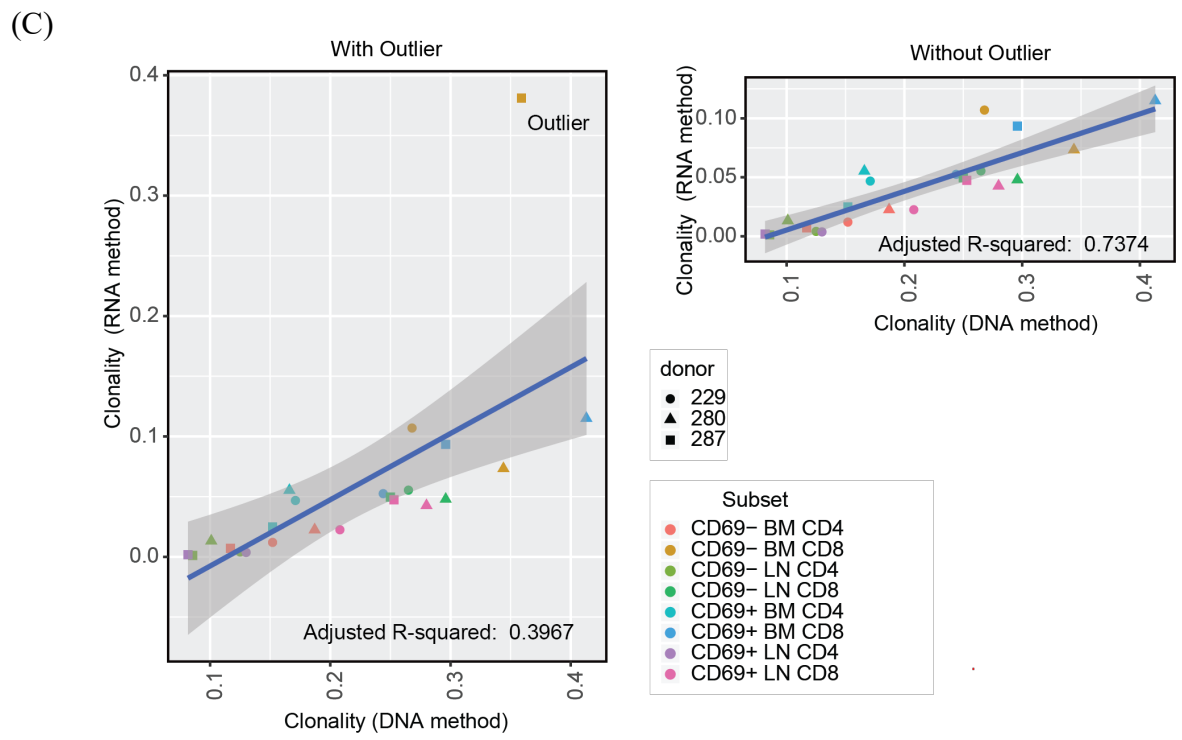
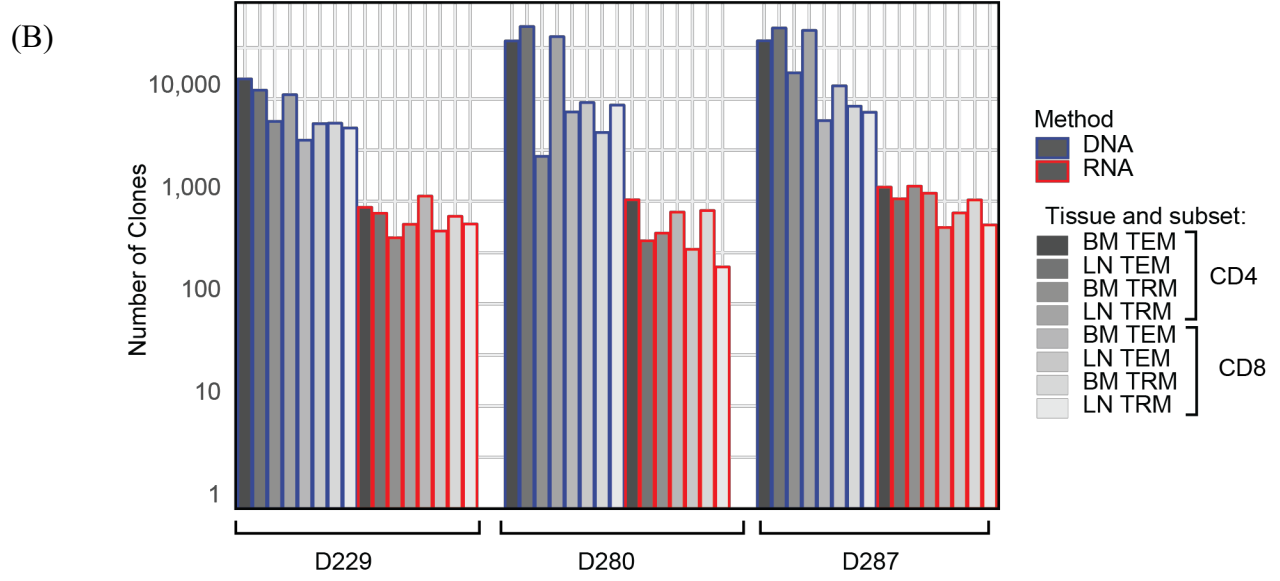
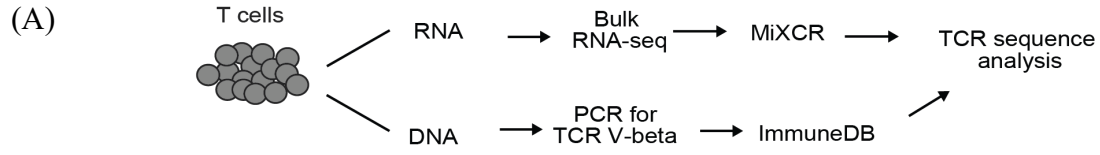
As another measure of TCR repertoire diversity, we calculated clonality, ranging from 0 (least clonally expanded; maximally diverse) to 1 (monoclonal, no diversity), from sequences derived from all 11 individuals (see methods). We compared clonality of different T cell subsets (from all tissues together) and we found clonality from highest to lowest was: CD8⁺ TEMRA > CD8⁺ TEM and TRM > CD4⁺ TEM and TRM > CD4⁺ TCM, consistent with the diversity analysis above. Next we looked for tissue specific differences, and we found that if we analyzed all subsets together, that clonality of T cells in LNs was lower than BM for both CD4⁺ and CD8⁺ T cells (Figure 4-3). In summary, the dominant factor determining clonality and diversity was subset, with CD8⁺TEMRA the most expanded/least diverse and CD4⁺ TCM, the least expanded/most diverse. Importantly, this subset clonality was independent of tissue. There were some tissue-specific effects with LN having the most diverse repertoires compared to BM.

Comparing DNA and RNA based methods for TCR sequencing

Increasingly, whole transcriptome profiling by RNASeq is being used to obtain TCR sequences. Software such as MiXCR[308], can be used to identify TCR sequences from T cell transcriptome profiles, without any targeted amplification of the TCR locus by PCR. One of the advantages of obtaining TCR sequence information from whole transcriptome profiling data is it allows simultaneous analysis of T cell functions coded by the transcriptome. However, the relative accuracy of TCR sequence determination from RNASeq compared to TCRseq has not been directly compared for human T cells. In order to compare these methods, we took advantage of a dataset we acquired using both DNA and RNA methods for TCR sequence identification from the same cell populations. We isolated total DNA and RNA from purified TEM and TRM cells from BM and LN for CD4⁺ and CD8⁺ lineages as described above from three donors (D299, D280, D287, see Table 4-1 for donor information) (Figure 4-4A). Whole transcriptome profiling was conducted from this RNA, as described in the methods section of Chapter 2. We found a substantial difference in the number of clones detected from DNA and RNA based TCR determination, with greater than 10 fold more clones detected using the DNA versus RNA-based method (Figure 4-4B). Interestingly, despite lower numbers of clones detected using the RNA method, there was a moderate linear relationship between clonality of samples from DNA and RNA methods (Adjusted R² = 0.40); removing an outlier resulted in an even stronger linear relationship (Adjusted R² = 0.74) (Figure 4-4C). In conclusion, with both methods we were able to analyze clonality and diversity of TCR sequences; however, the DNA method was superior due to higher coverage and the ability to detect far more unique T cell clones from the same samples.

Figure 4-4. Comparing DNA and RNA based methods for TCR sequencing.

(A) Schematic for the experimental procedure of DNA and RNA based methods for TCR sequencing and clone identification. Briefly, total RNA and DNA was isolated from purified T cell populations. RNA was proceeded for whole transcriptome sequencing by RNA-sequencing and MIXCR was used to identify T cell clones. DNA was used for PCR with specific primers targeted to the TCR V β -chain and ImmuneDB was used to identify clones. (B). Total number of unique clones identified by DNA (blue) or RNA (red) based methods for T cells, with tissue (BM or LN) and subset (TEM and TRM) identified by grey coloring for CD4⁺ and CD8⁺ T cells. (C) Clonality of samples from DNA versus RNA based methods with (left) or without (right) outlier (BM CD8⁺ TEM cells from D287) and the adjusted R² of the linear models.



Quantifying the overlap of T cell repertoires by cosine similarity

In order to determine how T cell clones are distributed across diverse sites within the same donor, we calculated the cosine similarity of clones for each subset between sites for the two donors (D324 and D383) that we had >4 sites analyzed. Cosine similarity, calculated by taking the cosine of the angle between two vectors of clone abundances, results in a value ranging from 0 (minimally overlapping) to 1 (complete overlap). [331] From this overlap analysis, we found that CD8⁺ T cells had overall higher cosine similarity compared with CD4⁺ T cells, indicating higher overlap. Importantly, there were substantial differences in the overlap between tissue T cell subsets of both CD4⁺ and CD8⁺ lineages (Figure 4-5). In order to analyze the subset differences, we performed hierarchical clustering on a matrix of pairwise cosine values which is shown as a heatmap in Figure 4-5. Hierarchical clustering revealed two main clusters for CD4⁺ T cells: TCM and TRM/TEM for D324, and two main clusters for CD8⁺ T cells: TEMRA and TEM/TRM for D324 and D383. For CD4⁺ T cells, TCM cells had low cosine similarity (close to 0) with all other samples, whereas CD4⁺ TEM and TRM cells displayed higher cosine similarity with non-TCM cells (up to 0.8) (Figure 4-5, left). For D324, CD8⁺ T cells, TEMRA cells had the highest cosine similarity with other TEMRA cells (close to 1), and very low cosine similarity with non-TEMRA cells (between 0-0.2); these samples clustered predominantly by subset (Figure 4-5, right). For donor 383, CD8⁺ TEMRA cells also had the highest degree of similarity with other TEMRA cells, compared to non-TEMRA cells, despite lower sampling of the blood sample (Figure 4-5, bottom right).

We next examined whether there was any tissue specific similarities between samples. For both individuals, we found that both CD4⁺ and CD8⁺ TEM and TRM from the lung had higher cosine similarity values with each other compared to non-lung samples (Figure 4-5, left). Additionally, for CD4⁺ T cells from both donors, Bld, Spl and BM TEM had higher cosine

similarity with each other compared to all non-TEM cells (Figure 4-5, left). Interestingly, cosine similarity values for CD8⁺ TEM and TRM cells were more variable donors, with D324 having overall higher levels of cosine similarity of TEM and TRM cells with other TEM and TRM cells across tissues (~0.7), whereas D383 displayed a larger range of similarity (between 0.2-0.8) (Figure 4-5, right).

To reveal the overall relatedness and clonal sharing of subsets and tissues visualized together, we conducted principle coordinate analyses (PCoA) of the cosine distances between samples. Analysis of total samples for each individual showed distinct clustering of CD4⁺ and CD8⁺ T cells, indicative of their divergent repertoires (Figure 4-6A). We then analyzed CD4⁺ and CD8⁺ T cells separately, based on replicate sampling, and found that replicates clustered near each other for both CD8⁺ and CD4⁺ T cells indicating replicate samples are similar to each other (Figure 4-6B). When examined in terms of subset-specific similarities, CD4⁺ T cells tended to cluster by subset, especially for CD4⁺TCM cells (Figure 4-7, left). When examined for tissue-specific similarities, CD4⁺ T cells, from the lung were clustered together (grey), while CD4⁺ T cells from the other tissue sites were intermingled together distinct from the lung (Figure 4-7, left). For CD8⁺ T cells subset similarities, TEMRA cells clustered closely together from different sites (Figure 4-7, right), while CD8⁺TEM and TRM subsets did not cluster by subset; they clustered more so by the tissue sites that these samples derived from, particularly for the lymph node samples which clustered together in both individuals (Figure 4-7, right). When analyzed with blood samples from donor 383, the blood TEM and bone marrow TEM and TRM samples clustered very closely together (Figure 4-7, right). These results agree with previous subset-specific clustering patterns from cosine similarity heatmaps and provide an overview of the relationships between cell populations, either by replicate sampling, subset or tissue site.

Figure 4-5 . Tissue distribution of T cell clones largely explained by cell subset and lineage identity.

A) Cosine similarity between pairwise cell populations labeled by Cell Type (TCM, TEM, and TRM) with indicated colors. Dendrogram created using complete linkage method. Each cell within the heatmap is the mean of the pairwise cosine similarity between two replicate samples per pairwise cell population comparison (four values).

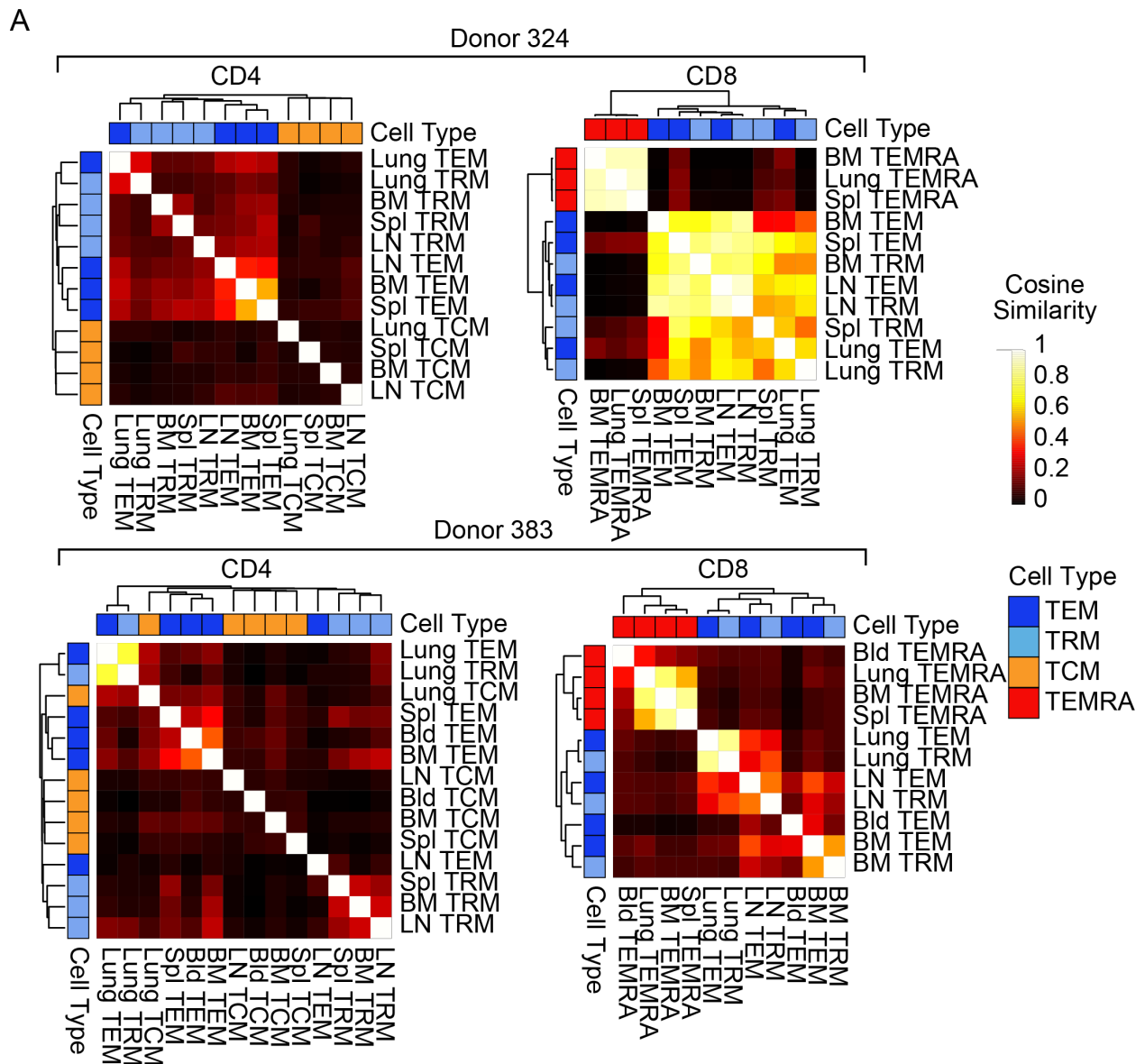


Figure 4-6. TCR repertoires of CD4⁺ and CD8⁺ lineages are divergent and show similarity between replicate samples.

(A) Principle coordinate analysis of cosine distances colored by lineage (CD4= blue, CD8 = pink) (B) or by replicate indicated by unique shape and color for D324 (left) and D383 (right).

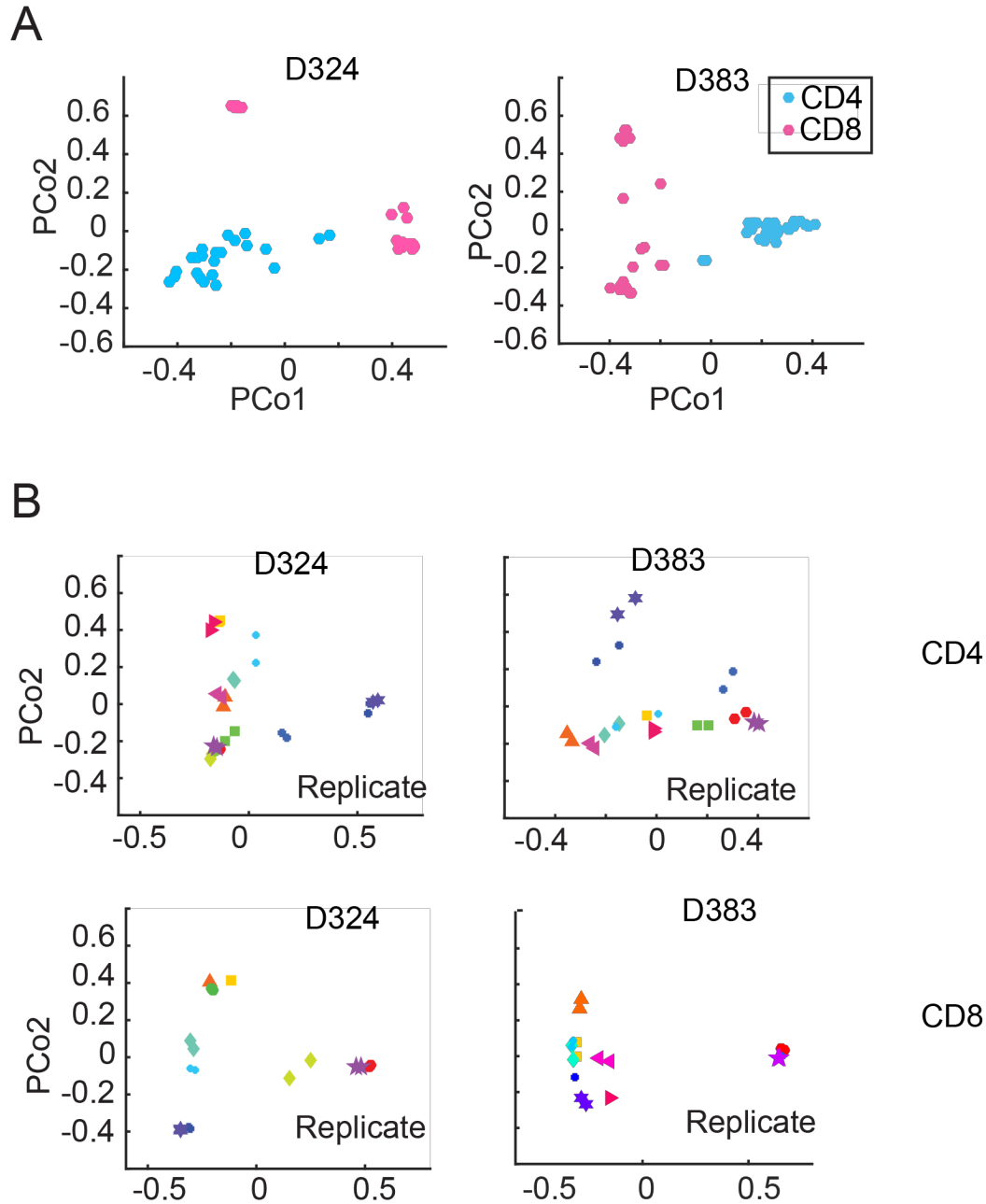
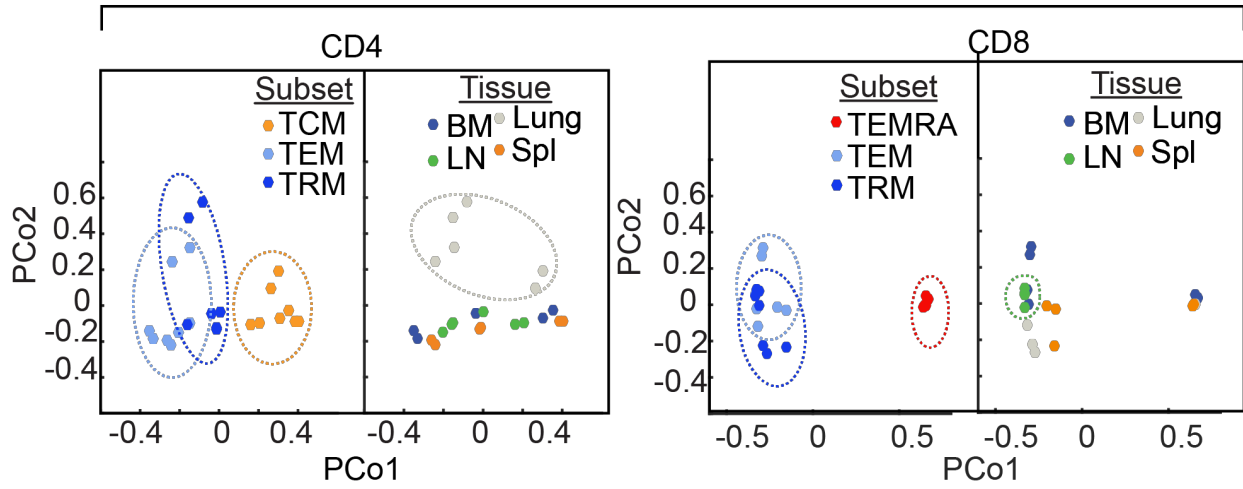


Figure 4-7. Tissue-specific and subset-specific differences of CD4⁺ and CD8⁺ T cell subsets.

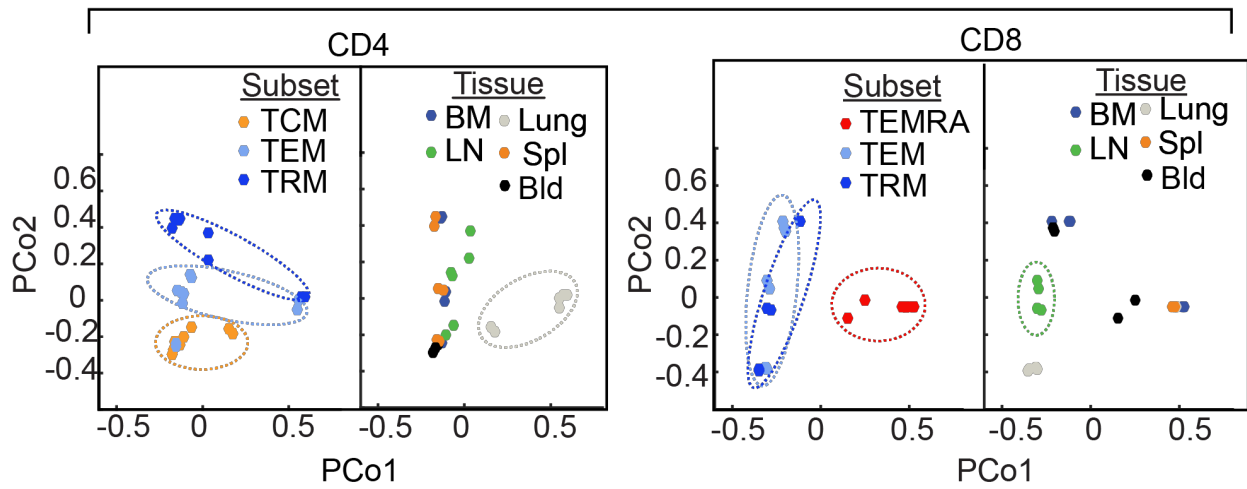
(A) Principle coordinate plots of cosine distances for CD4(left) and CD8 (right) samples from, D324 (top) and D383(bottom). Points are colored by subset (TCM = orange, TEM= light blue, TRM= dark blue) or Tissue (BM= dark blue, Lung= grey, LN = green, Spl= orange, Blood= black) as indicated with each plot.

A

Donor 324



Donor 383



Tracking the top clones and their abundances across tissues

The above analysis provided a quantitative assessment of the overlap of the entire complement of clones for each subset and tissue. In order to investigate how individual clones were distributed across subsets and tissues, we used a method of clone tracking which can be achieved with the most abundant clones. We examined the top 100 clones across all tissues within two individuals using clone tracking plots in which each clone is depicted as a colored segment of the plot, with the abundance of each clone proportional to the height of the segment; clones are stacked on top of one another (Figure 4-8). The color of each clone indicates the subset of the sample the clone was found in the highest frequency (TEM= light blue, TRM= dark blue, TCM= orange, TEMRA = red).

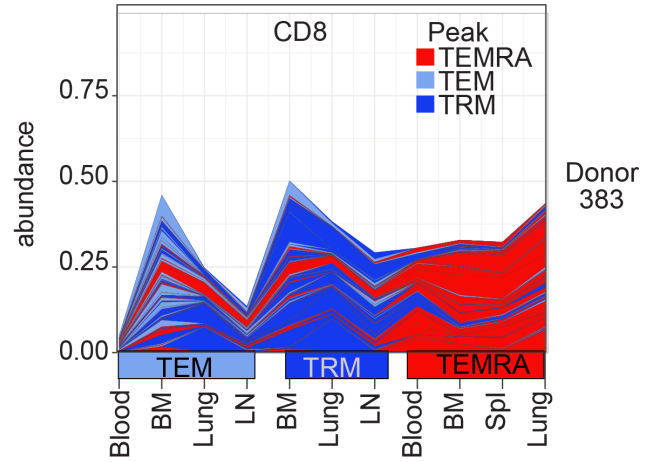
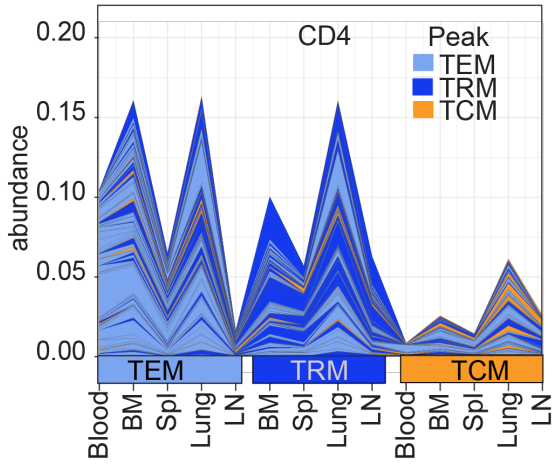
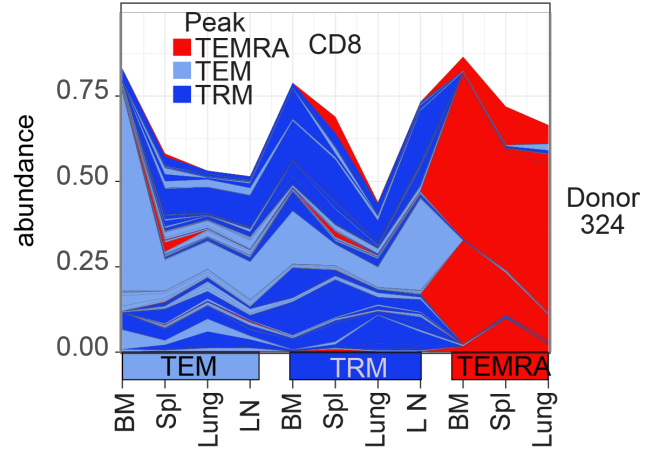
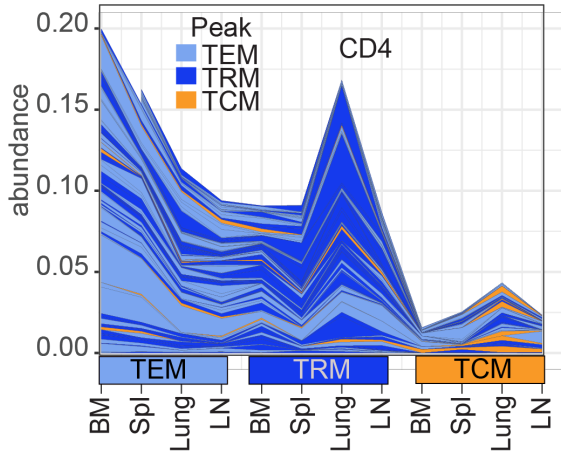
For CD4⁺ T cells, individual clones that were most abundant in TEM samples, tracked more frequently across TEM samples, but could also be detected in TRM and TCM samples across tissues (Figure 4-8, left). Likewise, individual clones that were most abundant in TRM samples, tracked more frequently across TRM samples, but could also be detected in TEM and TCM samples across tissues. This indicates some degree of increased sharing across subsets in different sites. Top clones in CD4⁺ TCM cells could also be detected in TEM and TRM samples across sites. For CD4⁺ T cells, we also observed some tissue-specific patterns, with lung and BM T cells which displayed highest proportion of these top clones, seen as higher peaks in the clone tracking plots. An interesting difference between the two individuals, was that donor 383 CD4⁺ LN TEM cells displayed very low overlap with other samples, looking more similar to the TCM samples across all tissues. CD8⁺ T cells displayed striking TEMRA specific sharing across sites; these top clones were either not found or found at very low frequency in other sites (Figure 4-8, right). Top clones in TEM and TRM samples displayed substantial sharing across subset and tissues. Similar to CD4⁺ T cells, the lung CD8⁺TEM T cells contained substantial TRM clones

shared with the lung TRM. Overall, this analysis shows the interconnectivity of the most abundant clones across all memory subsets analyzed, showing the dominant impact of T cell subset on clone identity properties and tissue sharing, as well as the impact of tissue site.

Figure 4-8. Top clones are widely distributed across tissues and compartmentalized in T cell subsets.

A) The top 100 most abundant clones by cumulative abundance across all tissue sites and their relative frequencies for each memory T cell population for donor 324 (top two plots) and donor 383 (bottom two plots) for CD4⁺ T cell populations (left) and CD8⁺ T cell populations (right). Each clone is colored by the memory T cell subset (TEM, TRM, TCM, or TEMRA) at which the clone was found at the highest frequency.

A



Similarity between TEM and TRM populations in lymphoid sites

In order to quantify similarity between samples by number of clones, rather than by clone abundance as calculated by cosine similarity, we analyzed the Jaccard index between samples. The Jaccard index was calculated by dividing the number of shared clones by the total number of clones (shared and non-shared) and multiplying that number by 100 to get a percentage. Using the Jaccard index, we examined the similarity of TEM and TRM subsets in lymphoid sites (BM, Spl, and LN) from eight individuals. The similarity between TEM and TRM populations within a tissue was between 5-15% for CD4⁺ T cells and 10-30% for CD8⁺ T cells and was higher overall than the overlap across tissues which was between 2-10% for CD4⁺ T cells and 8-20% for CD8⁺ T cells (Figure 4-9A,B). More detailed analysis of the intra-tissue overlap revealed that TEM populations overlap more with TEM populations (light blue boxplots), and likewise TRM populations overlap more with TRM populations (dark blue boxplots), compared to overlap between TEM and TRM populations within the same tissues, for CD4⁺ and CD8⁺ T cells (Figure 4-9A).

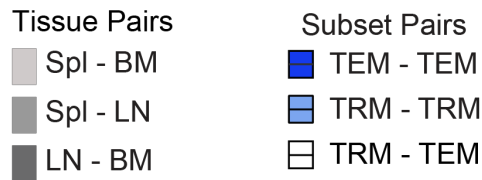
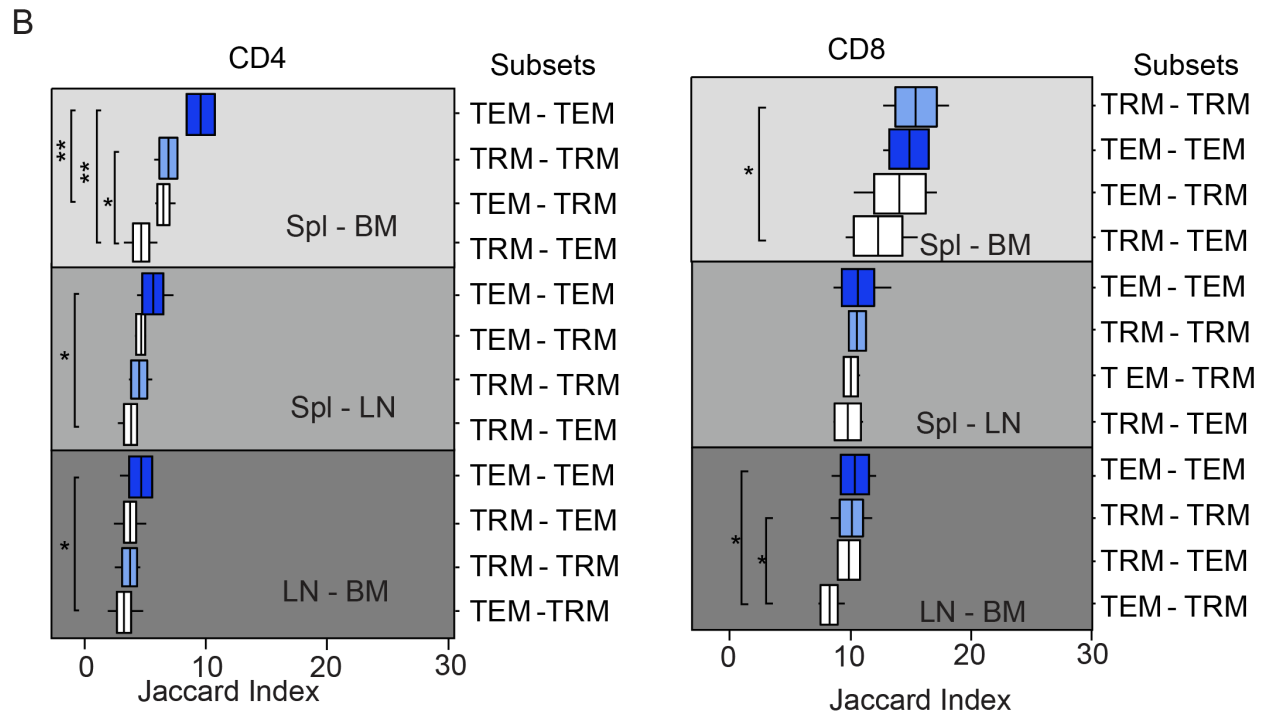
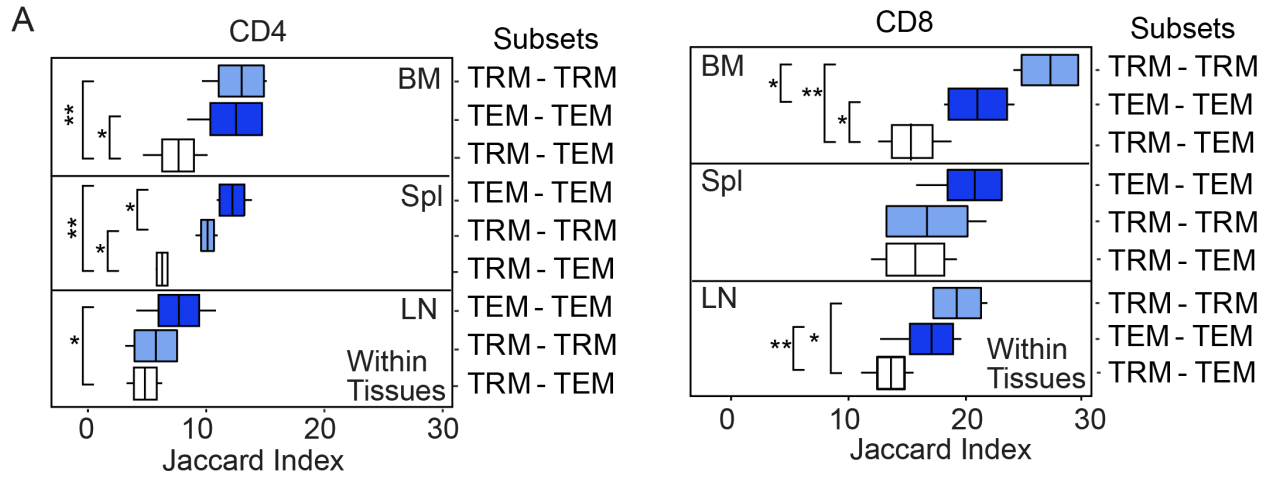
Next, we looked for subset-specific differences in the degree of overlap between different tissue sites. We found that CD4⁺ TEM cells overlapped more with CD4⁺ TEM cells in other tissues (dark blue boxplots, between 8-10%), compared to the degree of overlap with CD4⁺ TRM in other tissues (white boxplots, between 2-7%) in all lymphoid sites (Figure 4-9B, left). We also found that CD4⁺ TRM cells overlapped more with TRM cells in other tissues (light blue box plots, 5-8%) compared to with CD4⁺ TEM cells in other tissues (white boxplots, 2-7%) in all lymphoid sites, with the degree of overlap varying by tissue (Figure 4-9B, left). Interestingly, CD8⁺ T cells had fewer subset-specific differences than CD4⁺ T cells (Figure 4-9B, right).

Next, we examined tissue-specific differences in the degree of overlap between T cell populations across sites. We found that the overlap of CD4⁺ TEM and TRM between BM and

Spl (5-12%) and of CD8⁺ TEM and TRM between BM and Spl (15-20%) was overall higher than the overlap of CD4⁺ TEM and TRM between Spl and LN or BM and LN (2-5%) and the overlap between CD8⁺ TEM and TRM across BM and Spl (8-12%). Together, these findings indicate that there are subset and tissue specific difference in the degree of similarity of TEM and TRM subsets in lymphoid sites.

Figure 4-9. Substantial overlap of TEM and TRM in tissues with high intra-tissue and subset overlap.

A) T cell clones were counted in each sample and the jaccard index was calculated using the clones in samples 1 and 2 from the same tissues for CD4⁺ (left, from top: BM, n=7,8,8; Spl, n=4,5,5; LN, n=8) and CD8⁺ TEM and TRM cells (right, from top: BM, n=8, Spl, n=5,4,4, LN, n=8) B) and for samples from different tissues (top to bottom, CD4: BM-Spl, n=4,5,5,4; LN-Spl, n=5; LN-BM, n=8; CD8: Spl-BM, n=4,5,5,4, LN-Spl, n=5,4,5,4, LN-BM, n=8).. Students T-test. *P<0.05, **P<0.01. Cut-off for clones used in this analysis was 50% of the mean clone abundance.



Section 4.3 Discussion

Here we analyzed how human T cells were clonally distributed across blood and tissue sites and the role of subset and tissue in clonal maintenance. We report the discovery that human CD4⁺ and CD8⁺ memory T cell clones are widely disseminated to multiple lymphoid sites, lung, and blood, and terminal effectors cells further disseminated. Moreover, we found the degree of dissemination and maintenance of T cells clones was compartmentalized within specific T cell subsets. Several subset-specific features were conserved across tissue sites, including the high diversity among CD4⁺TCM clones, and low diversity among CD8⁺TEMRA. Further, we found tissue-specific differences in T cell clone dissemination, including a high degree of overlap between CD4⁺TEM and CD4⁺TRM clones within the lung, and a high degree of overlap between CD4⁺TEM from Spl and CD4⁺TEM from BM. Additionally, both CD8⁺TEM and CD8⁺TRM clones displayed a high degree of overlap across lymphoid sites. Together these results add new insights into T cell differentiation in tissues. We summarize these insights graphically in Figure 4-10, by indicating the subset-specific and tissue-specific features of T cell clone sharing and diversity across sites.

Cellular adaptive immunity is mediated by generation of diverse effector and memory T cell subsets that each have distinct functions and roles as elucidated by mouse models. A single naïve T cell with a given T cell receptor specificity can generate distinct memory subsets including both TCM and TRM cells in response to skin adjuvants. [159] Additionally, studies in mice tracking the response to vesicular stomatitis virus (VSV) show that T cell clones disseminate to multiple tissues regardless of their site of origin [184]. In humans, the distribution and fate of effector and memory T cell subsets that derive from clonal expansion of single naïve T cells is not known. Our data show that there is substantial overlap in T cell clones between different sites and subsets, but that the extent of overlap is influenced by the subset and lineage.

For example, CD8⁺ TEMRA cells are the most widely distributed across sites including BM, Spl blood, and lung. Additionally, recent studies on flu-specific responses in human tissues found a common CD8⁺ T cell clone in human lung, Spl and LN, consistent with our results that CD8⁺ T cell clones are widely disseminated in diverse sites. [332] While our results imply shared antigen specificities across tissues, we have not directly looked at antigen specificity in this study and this is an important area for further investigation.

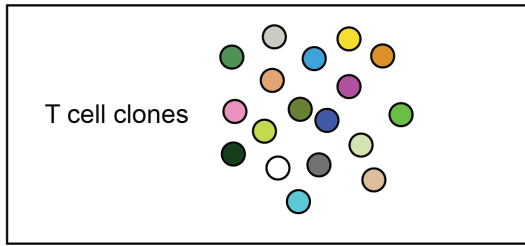
In general, it is thought that increased T cell receptor diversity is a valuable asset as this allows more potential to recognize pathogens for clearance. To support this, a study in humans examining T-cell receptor α - β diversity of pathogen-specific T cells to Cytomegalovirus (CMV), a chronic virus, found that a diverse T cell receptor repertoire is a biomarker of better outcomes to infection. Mechanistically this could be due to maintenance of T cells that are able to respond to different viruses. Alternatively, the fact that chronic antigen drives T cell clonal expansion and therefore decreases the diversity of a TCR repertoire may be due to viral replication and therefore be a result of infection rather than causative of increased protection. [333] Our results reveal that in LN sites CD4⁺ and CD8⁺ memory T cell repertoires display a higher degree of diversity compared to BM sites. We have previously shown evidence for this specifically in CD8⁺ T cells [334]. This suggests that LN may serve as a reservoir for diverse memory T cells that retain the capacity for recognition of different pathogens, to respond and populate tissue sites. Additionally, the high degree of clonal expansions observed in the BM suggest it is a niche for T cells that have previously undergone high levels of proliferation that has been suggested by mouse models [271, 335], and may be important for circulation, maintenance and dissemination of these effector clones throughout the body.

Our previous study of human CMV-specific T cell responses in tissues, indicates two patterns of responses: in one pattern individuals had abundant antigen specific responses in bone marrow that were highly disseminated across many sites (including blood, spleen and lungs), in

Figure 4-10. Overview of subset and tissue specific features of TCR repertoires.

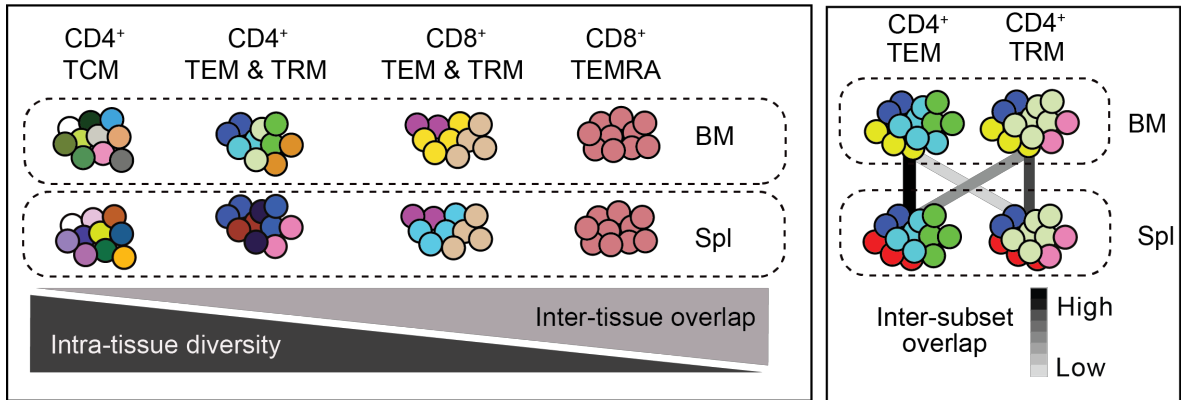
(A) Each T cell clone, identified by sequencing of the TCR V β -chain genomic locus, is represented by a circle filled with a unique color. (B) Summary of subset-specific features of T cell clones. The ranking of subsets in regard to diversity, from highest to lowest is CD4⁺ TCM, CD4⁺ TEM/TRM, CD8⁺ TEM/TRM, and CD8⁺ TEMRA cells and this is conserved across tissues. The ranking of subsets with regards to degree of overlap between tissues (inter-tissue) from lowest to highest is CD4⁺ TCM, CD4⁺ TEM/TRM, CD8⁺ TEM/TRM, and CD8⁺ TEMRA cells. These results are shown for BM and Spl as examples, with similar findings in blood, lung and LN sites (left box). The degree of overlap between CD4⁺ TEM and TRM cell populations in different tissues is indicated by the color of the connecting lines. Darker shades indicate higher degrees of overlap (right box). Results shown are representative of BM, Spl and LN sites. (C) Summary of tissue-specific features of T cell clones. High intra-tissue overlap between TEM and TRM cells in the lung and LN for CD4⁺ and CD8⁺ T cells respectively (left box). CD4⁺ and CD8⁺ T cells exhibit higher diversity in LN compared to BM (middle box). For TEM and TRM subsets, the degree of overlap between tissues (inter-tissue) is indicated by the color of the connecting arcs. Darker shades indicate higher degrees of overlap (right box).

(A)



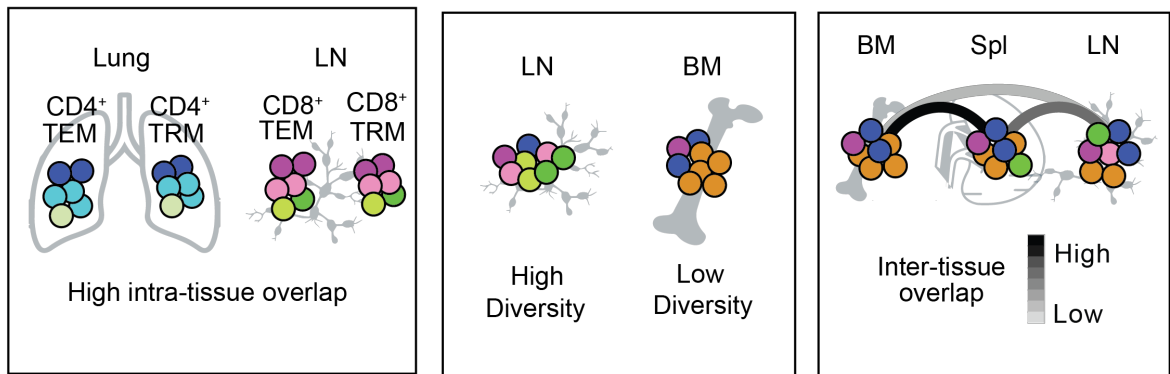
(B)

Subset-specific features



(C)

Tissue-specific features



the second pattern, individuals had antigen specific responses only in lymph nodes (including those draining the lung, mesentery, and iliac). [10] The results from our study showing highly expanded clones in BM compared to smaller clone sized in lymph nodes may be indicative of functioning in active responses versus maintenance of memory respectively.

Characterizations of the T cell receptor sequences in blood and disease tissues have revealed the potential of this information as a diagnostic biomarker, however the healthy immune response in tissues remains elusive and important for understanding protective immunity in humans [329, 336-338]. Our results indicate that sampling of T cells in peripheral blood by analyzing different subsets can be indicative of what is occurring in distant tissue sites. This provides support for the use of peripheral blood as an important resource for immune monitoring in human studies. Overall, our findings provide insight into memory T cell differentiation in tissues. Our results suggest that in terms of T cell lineage and differentiation, cell subset and tissue site are large determinants of T cell receptor repertoires. This data could be a useful resource for future modeling and estimates of T cell receptor diversity, which thus far have been based on blood sampling.

CHAPTER 5: Conclusions

Memory T cell generation, function and maintenance occur at diverse sites, yet most knowledge on human T cell immunity derives from study of peripheral blood samples. Initially, two subsets of T cells were identified in human peripheral blood based on homing capacity and function. Effector memory T cells (TEM) displayed increased abilities to migrate to non-lymphoid tissues and confer immediate effector function upon antigen encounter, while central memory T cells (TCM) displayed increased abilities to home to lymphoid sites and proliferate upon antigen encounter [121, 339]. Further, studies from mouse models have revealed a specialized subset of T cells that confer protective immunity remain resident in lymphoid and non-lymphoid sites termed tissue-resident memory (TRM). [177-183] Although the role of TRM in human immune responses is not yet defined, they are implicated in a number of tissue-localized disease states, such as psoriasis in the skin and are associated with tumors specific to skin, liver, lung and breast tissues[326-329]. In humans, how T cell differentiation is occurring to generate diverse subsets remains an active area of investigation. While mouse models cannot recapitulate the length and diversity of exposures to pathogens that takes place over many decades in humans, the extent to which this difference impacts the generalizability of findings on tissue immunity in mice is not known, and will be an important area of investigation in order to translate findings in mice to man.

A significant difference in memory T cell populations between mice and man includes that in humans, lymph nodes maintain a significant fraction (>50%) of CD8⁺TRM phenotype cells, marked by CD69 expression; however, in specific pathogen free (SPF) mice, CD8⁺TRM are infrequent. [211] Possible reasons for this discrepancy include differences in genetics or microbial exposure. Interestingly, in pet store mice which harbor more pathogens and

commensal microorganisms (non-SPF), CD8⁺CD69⁺ memory T cells are abundant and at similar levels to adult humans[278]. Further, mice co-housed with pet store mice also have similarly increased levels of CD8⁺CD69⁺ T cells [211]. Therefore, microbial exposure is likely to contribute to the apparent discrepancy in TRM phenotype cells in humans versus SPF mice. Further, previous studies in SPF mice have underestimated the importance of memory T cell populations, and in humans, lymph nodes are significant reservoirs of memory T cells. **Our hypothesis is that T cells are playing unique roles in LNs and we sought to define their functional capacity as well as mechanisms for maintenance.**

In studies comparing memory T cell populations in blood, LN, BM, spleen and lung, we identified substantial phenotypic and functional diversity of T cells in different anatomical sites. We found that CD8⁺TEM cells in blood, bone marrow, spleen and lung, displayed high effector function marked by expression of T-bet. In contrast, in lymph nodes, CD8⁺TEM cells displayed phenotypes of quiescence marked by expression of TCF-1. Lymph nodes also maintain a significant fraction (>50%) of CD8⁺TRM phenotype cells (not found in peripheral blood) and marked by CD69 expression. Interestingly, CD8⁺TRM cells also expressed higher levels of TCF-1 in LN compared to in BM, Spl, and lung. Additionally, we found that CD8⁺TEM and TRM cells displayed correspondingly higher expression of Lef-1 in LN compared to other sites. We found additional tissue-specific phenotypes of CD8⁺TEM cells in LN including lower expression of CD57 and Perforin and higher expression of CXCR5, CD28, and CD27. Together, we identified a protein signature that identified memory LN T cells as distinct from TEM and TRM in other sites and we designate this subset TLN. Upon TCR stimulation with CD3 and CD28, TLN had increased capacity for proliferation compared to total memory in BM, Spl, and lung. Additionally, upon stimulation, TLN highly upregulated T-bet indicating their capacity for differentiation into effector phenotype cells. We identified that exposure to IFN α/β during TCR

stimulation resulted in increased down regulation of TCF-1 expression in LN memory T cells. Therefore, we propose that perhaps lymph nodes provide a niche absent of type 1 interferon signaling, allowing maintenance of TCF-1⁺ T cells. Together, these results indicate maintenance of distinct subsets with high capacity for divisions or immediate effector functions.

What is the role of memory T cells in lymph nodes, and why are they maintained with distinct functional capacity? One hypothesis is that lymph node localized memory T cells are providing protective immunity to pathogens in lymph nodes, similar to roles in mice of influenza specific TRM in the lung [179]. Humans are exposed to several pathogens that selectively infect lymphoid cells including Epstein Barr Virus (EBV) and HIV which establish latent infection in B cells and CD4⁺ T cells respectively. Consistent with this hypothesis, one study found CD8⁺TRM phenotype cells specific for EBV in tonsils and spleen (up to 8% of total CD8⁺ T cells) [194]. Additionally, multiple studies have shown that increased numbers of cytotoxic CD8⁺T cells were localized in LNs of individuals infected with HIV, with T cell numbers inversely correlated with viral load[275, 316, 317]. HIV-antigen driven CD4⁺ T follicular helper cell clonal expansions were found in lymph nodes of HIV⁺ individuals [340], indicating a site-specific response, which may be due to either oligoclonal growth or cell killing.

Expression of TCF-1 and increased capacity for cell division, suggest an alternative hypothesis for the role memory T cells in lymph nodes; TLN may serve to regenerate effector cells and the self-renewing memory pool similar to central memory T cells [99, 341, 342]. Unlike canonical central memory T cells, TLN do not express high levels of CCR7, and instead the majority express TRM marker CD69, suggesting a CCR7-independent mechanisms for maintenance of memory T cells in lymph nodes. In mice, studies of Ag-specific responses to influenza challenge have shown that asymmetric division couples the generation of and

maintenance of TCF-1 expressing cells in non-draining lymph nodes with ability to self-renew and TCF-1 negative cells in the lung with ability for effector function.[168]

Successful immune responses generated against acute viruses such as influenza lead to rapid viral clearance and generation of memory T cells. In contrast, chronic viral infections and cancer cells evade immune mediated clearance and persist over long periods of time, presenting an ongoing challenge to the immune system. Several recent studies found lymphoid specific CXCR5⁺TCF-1⁺CD8⁺ T cell populations generated in response to chronic LCMV infection; they found these populations responded to PD-1 therapy, leading to robust proliferation of otherwise exhausted T cell responses.[154, 248, 303, 316, 317] The TLN cells we identified in humans displayed similar phenotype, function, and transcriptional profiles to these recently described subsets in mouse models. Perhaps human TLN cells are an analogous subset generated in response to chronic viral infection and are responsive to PD-1 therapy. A recent study in humans found that stem-like CXCR5⁺CD8⁺ T cells reside in tumors [343]. Evidence in mouse models further supports that successful response to cancer immunotherapy resulted in coordinated anti-tumor immunity across the organism.[344]

We found that T cell responses to chronic infection with CMV exhibit tissue-specific localization patterns. In certain individuals, CMV-specific T cell responses were compartmentalized in lymph nodes[10]. In contrast, other individuals had large populations of CMV-specific T cells disseminated in many sites including blood, BM, lung and spleen.[10] Perhaps LN reservoirs of memory T cells are poised to respond upon reactivation of latent viral reservoirs; whereas, disseminated responses are indicative of more recent viral reactivation and subsequent control. We have not been able to correlate this with what the virus is doing; we detected viral genomes in many sites, most often in the lung. [10] Further investigations of

relationship between T cell responses and virus localization will lead to better insights into human responses.

In our second study we investigated the relatedness of circulating and tissue T cells by clonal analysis of T cell subsets comprising TCM, TEM, TRM and TEMRA cells from Bld, BM, Spl, LN and Lung. From diversity analysis, we found that CD4⁺TCM cells were the most diverse, and CD8⁺TEMRA cells were the least diverse in all sites examined. CD4⁺TEM and TRM cells were more diverse than CD8⁺TEM and TRM cells. We looked for tissue specific differences and found that diversity of CD4⁺ and CD8⁺ T cells was greater in LN compared to BM. T cell repertoire diversity displayed an inverse relationship with the degree of clonal overlap between T cell subsets. CD4⁺TCM cells exhibited low overlap whereas CD8⁺TEMRA cells exhibited high overlap, particularly with TEMRA cells in different sites. In addition to these subset specific differences, we also found tissue-specific features of T clone sharing. We found CD4⁺TEM cells were most similar between BM and Spl, more so than with LN and Lung sites. We also found substantial overlap between TEM and TRM cells in the lung, suggestive of antigen driven compartmentalization in tissue sites. In conclusion, we found that the relatedness of T cell clones was largely subset-specific, and to a lesser degree tissue-specific.

Future studies tracking T cell responses over time with information on timing of pathogen encounter and control or clearance will provide more insights into the drivers of T cell differentiation. Additionally, investigations of the antigen specificity of T cell responses in human tissues would yield significant insights into history of antigen encounters drive T cell responses. In order to discover the Ag-specificity of a given T cell clone, T cells can be screened against large libraries of peptide-MHC complexes and reactive clones can be identified. This method was recently applied to questions in tumor immunology, where increased T cell infiltration in tumors has been associated with better prognosis.[345] In this study, authors

identified that only 10% of T cells derived from tumors were tumor-specific, with some tumors having even lower percentages, revealing that non-tumor antigens may drive a significant portion of T cell populations in tissues, or that they were unable to identify the right tumor antigen due to clonal deletion by immune systems or technical limitations. [345]

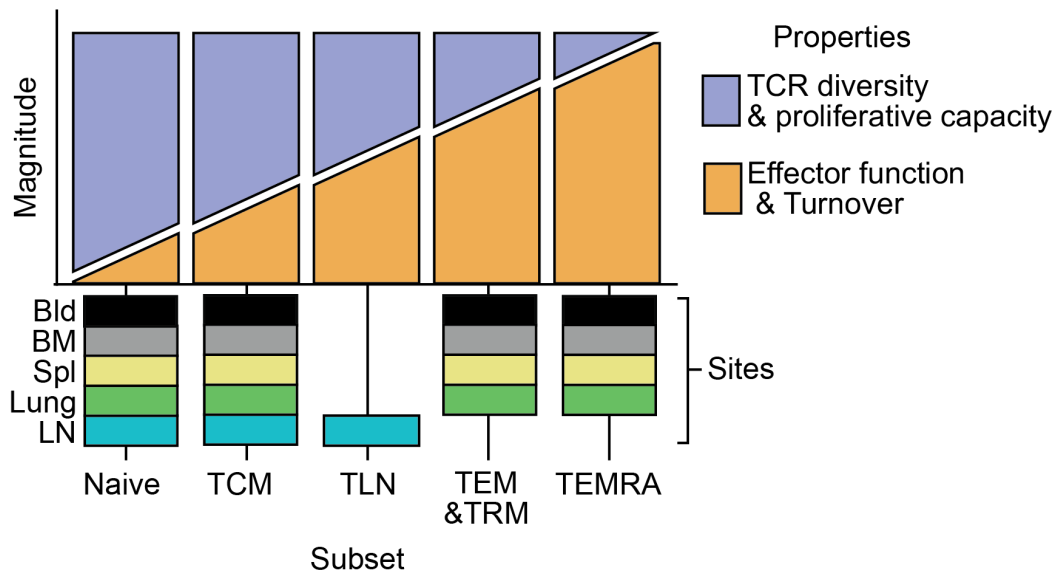
Our results reveal that T cell populations exhibit substantial diversity across tissue sites and subsets as depicted in Figure 5-1. Here we show how the magnitude of proliferative capacity (purple) and effector function (orange) gradually increase according to both localization and cell subset identity, with a novel subset we identify in LN sites as TLN. Additional properties that correspond with proliferative capacity and effector function are TCR diversity (purple) and turnover (orange) respectively. Naive T cells exhibit highest level of proliferative capacity and TCR diversity followed by TCM, TLN, TEM & TRM, and TEMRA cells in that order. Naive T cells also exhibit the lower levels of effector function and turnover, with increasing levels in TCM, TLN, TEM & TRM, and TEMRA cells in that order. Naive and TCM cells are found in all sites (Bld, BM, Spl, LN, and lung) however, TLN cells are maintained exclusively in LN sites, and distinct in phenotype and function from TEM and TRM cells in Bld, BM, Spl, and lung. Additionally, TEMRA are found in Bld, BM, Spl and lung. In summary, T cells exhibit substantial heterogeneity by subset and tissue site, leading to new insights into T cell differentiation and maintenance in humans including that lymph nodes are an important reservoir for distinct and quiescent subset memory T cells with increased T cell receptor diversity and proliferative capacity compared to subsets found in vascularized sites including blood, BM, spleen, and lung.

These results have major implications for lymph nodes as sites for maintenance of immune memory such as targeting for immunotherapies and vaccines. Further, lymph nodes may provide important reservoirs for maintenance of T cell clones with self-renewal abilities. In

contrast, our results suggest that BM may be a reservoir for highly circulating cells that have lower self-renewal abilities and more immediate effector functional capacity. Dissecting the role of systemic versus tissue specific responses as well as the interplay between the two in disease will be critical in translating findings in mouse models to humans.

Figure 5-1. Summary of findings on T cell compartmentalization in tissues.

T cell subsets: Naive (CCR7+CD45RA+), TCM (CCR7+CD45RA-), TLN (CCR7-CD45RA-CD69+/-), TEM (CCR7-CD45RA-CD69-), TRM (CCR7-CD45RA-CD69+), and TEMRA (CCR7-CD45RA+) shown according to magnitude of properties: 1) TCR diversity and proliferative capacity (purple) and 2) effector function and turnover (orange). Anatomical site which subset is found is indicated by colored bars below each column (Bld = black, BM = grey, Spl= yellow, Lung = green, LN= blue).



REFERENCES

1. Lefrancois, L. and D. Masopust, *T cell immunity in lymphoid and non-lymphoid tissues*. *Curr Opin Immunol*, 2002. **14**(4): p. 503-8.
2. Lefrancois, L., et al., *Migration of primary and memory CD8 T cells*. *Adv Exp Med Biol*, 2002. **512**: p. 141-6.
3. Rosenberg, S.A., *Decade in review-cancer immunotherapy: entering the mainstream of cancer treatment*. *Nat Rev Clin Oncol*, 2014. **11**(11): p. 630-2.
4. Ganusov, V.V. and R.J. De Boer, *Do most lymphocytes in humans really reside in the gut?* *Trends Immunol*, 2007. **28**(12): p. 514-8.
5. Clark, R.A., *Skin-resident T cells: the ups and downs of on site immunity*. *J Invest Dermatol*, 2010. **130**(2): p. 362-70.
6. Sathaliyawala, T., et al., *Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets*. *Immunity*, 2013. **38**(1): p. 187-97.
7. Thome, J.J., et al., *Spatial map of human T cell compartmentalization and maintenance over decades of life*. *Cell*, 2014. **159**(4): p. 814-28.
8. Kumar, B.V., et al., *Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites*. *Cell Reports*, 2017. **20**(12): p. 2921-2934.
9. Thome, J.J., et al., *Early-life compartmentalization of human T cell differentiation and regulatory function in mucosal and lymphoid tissues*. *Nat Med*, 2016. **22**(1): p. 72-7.
10. Gordon, C.L., et al., *Tissue reservoirs of antiviral T cell immunity in persistent human CMV infection*. *J Exp Med*, 2017. **214**(3): p. 651-667.
11. Carpenter, D.J., et al., *Human immunology studies using organ donors: Impact of clinical variations on immune parameters in tissues and circulation*. *Am J Transplant*, 2018. **18**(1): p. 74-88.
12. Kumar, B.V., et al., *Functional heterogeneity of human tissue-resident memory T cells based on dye efflux capacities*. *JCI Insight*, 2018. **3**(22).
13. Thome, J.J., et al., *Longterm maintenance of human naive T cells through in situ homeostasis in lymphoid tissue sites*. *Sci Immunol*, 2016. **1**(6): p. aah6506.
14. Farber, D.L., N.A. Yudanin, and N.P. Restifo, *Human memory T cells: generation, compartmentalization and homeostasis*. *Nat Rev Immunol*, 2014. **14**(1): p. 24-35.
15. Thome, J.J. and D.L. Farber, *Emerging concepts in tissue-resident T cells: lessons from humans*. *Trends Immunol*, 2015. **36**(7): p. 428-35.
16. Takahama, Y., *Journey through the thymus: stromal guides for T-cell development and selection*. *Nat Rev Immunol*, 2006. **6**(2): p. 127-35.
17. Robins, H.S., et al., *Overlap and effective size of the human CD8+ T cell receptor repertoire*. *Sci Transl Med*, 2010. **2**(47): p. 47ra64.
18. Qi, Q., et al., *Diversity and clonal selection in the human T-cell repertoire*. *Proc Natl Acad Sci U S A*, 2014. **111**(36): p. 13139-44.
19. Klein, L., et al., *Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see)*. *Nat Rev Immunol*, 2014. **14**(6): p. 377-91.
20. Aw, D. and D.B. Palmer, *The origin and implication of thymic involution*. *Aging Dis*, 2011. **2**(5): p. 437-43.

21. Haynes, B.F., et al., *Early events in human T cell ontogeny. Phenotypic characterization and immunohistologic localization of T cell precursors in early human fetal tissues.* J Exp Med, 1988. **168**(3): p. 1061-80.
22. Dalmaso, A.P., et al., *Studies on the Role of the Thymus in Immunobiology; Reconstitution of Immunologic Capacity in Mice Thymectomized at Birth.* J Exp Med, 1963. **118**: p. 1089-109.
23. Sjodin, K., et al., *Thymectomy in Newborn and Adult Mice.* Transplantation, 1963. **29**: p. 521-5.
24. Nowell, C.S., et al., *Foxn1 regulates lineage progression in cortical and medullary thymic epithelial cells but is dispensable for medullary sublineage divergence.* PLoS Genet, 2011. **7**(11): p. e1002348.
25. Nehls, M., et al., *New member of the winged-helix protein family disrupted in mouse and rat nude mutations.* Nature, 1994. **372**(6501): p. 103-7.
26. Frank, J., et al., *Exposing the human nude phenotype.* Nature, 1999. **398**(6727): p. 473-4.
27. de la Chapelle, A., et al., *A deletion in chromosome 22 can cause DiGeorge syndrome.* Hum Genet, 1981. **57**(3): p. 253-6.
28. Mancebo, E., et al., *Longitudinal analysis of immune function in the first 3 years of life in thymectomized neonates during cardiac surgery.* Clin Exp Immunol, 2008. **154**(3): p. 375-83.
29. Wells, W.J., et al., *Neonatal thymectomy: does it affect immune function?* J Thorac Cardiovasc Surg, 1998. **115**(5): p. 1041-6.
30. Prelog, M., et al., *Thymectomy in early childhood: significant alterations of the CD4(+)CD45RA(+)CD62L(+) T cell compartment in later life.* Clin Immunol, 2009. **130**(2): p. 123-32.
31. van den Broek, T., et al., *Neonatal thymectomy reveals differentiation and plasticity within human naive T cells.* J Clin Invest, 2016. **126**(3): p. 1126-36.
32. Steinmann, G.G., *Changes in the human thymus during aging.* Curr Top Pathol, 1986. **75**: p. 43-88.
33. Germain, R.N., *T-cell development and the CD4-CD8 lineage decision.* Nat Rev Immunol, 2002. **2**(5): p. 309-22.
34. Thien, C.B., et al., *Loss of c-Cbl RING finger function results in high-intensity TCR signaling and thymic deletion.* Embo J, 2005. **24**(21): p. 3807-19.
35. Forster, R., A.C. Davalos-Misnitz, and A. Rot, *CCR7 and its ligands: balancing immunity and tolerance.* Nat Rev Immunol, 2008. **8**(5): p. 362-71.
36. Zarnitsyna, V.I., et al., *Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire.* Front Immunol, 2013. **4**: p. 485.
37. Cao, Y., et al., *Pleiotropic defects in TCR signaling in a Vav-1-null Jurkat T-cell line.* Embo J, 2002. **21**(18): p. 4809-19.
38. Murphy, K., Travers, P., Walport, M., & Janeway, C., *Janeway's immunobiology.* 8th ed. 2012, New York: Garland Science.
39. Krishnan, S., et al., *The FcRgamma subunit and Syk kinase replace the CD3zeta-chain and ZAP-70 kinase in the TCR signaling complex of human effector CD4 T cells.* J Immunol, 2003. **170**(8): p. 4189-95.
40. Priatel, J.J., et al., *RasGRP1 transmits prodifferentiation TCR signaling that is crucial for CD4 T cell development.* J Immunol, 2006. **177**(3): p. 1470-80.
41. Carpino, N., et al., *Regulation of ZAP-70 activation and TCR signaling by two related proteins, Sts-1 and Sts-2.* Immunity, 2004. **20**(1): p. 37-46.

42. Shinkai, Y., et al., *RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement*. Cell, 1992. **68**(5): p. 855-67.
43. Mombaerts, P., et al., *RAG-1-deficient mice have no mature B and T lymphocytes*. Cell, 1992. **68**(5): p. 869-877.
44. Neller, M.A., et al., *Naive CD8 T-cell precursors display structured TCR repertoires and composite antigen-driven selection dynamics*. Immunol Cell Biol, 2015.
45. Olin, A., et al., *Stereotypic Immune System Development in Newborn Children*. Cell, 2018. **174**(5): p. 1277-1292 e14.
46. Taub, D.D. and D.L. Longo, *Insights into thymic aging and regeneration*. Immunol Rev, 2005. **205**: p. 72-93.
47. Tripp, R.A., et al., *Bone marrow can function as a lymphoid organ during a primary immune response under conditions of disrupted lymphocyte trafficking*. J Immunol, 1997. **158**(8): p. 3716-20.
48. Feuerer, M., et al., *Bone marrow as a priming site for T-cell responses to blood-borne antigen*. Nat Med, 2003. **9**(9): p. 1151-7.
49. Lanzavecchia, A. and F. Sallusto, *Regulation of T cell immunity by dendritic cells*. Cell, 2001. **106**(3): p. 263-6.
50. Michel, F. and O. Acuto, *CD28 costimulation: a source of Vav-1 for TCR signaling with the help of SLP-76?* Sci STKE, 2002. **2002**(144): p. PE35.
51. Ziegler, S.F., et al., *Molecular characterization of the early activation antigen CD69: a type II membrane glycoprotein related to a family of natural killer cell activation antigens*. Eur J Immunol, 1993. **23**(7): p. 1643-8.
52. Shiow, L.R., et al., *CD69 acts downstream of interferon-alpha/beta to inhibit S1P1 and lymphocyte egress from lymphoid organs*. Nature, 2006. **440**(7083): p. 540-4.
53. Mackay, L.K., et al., *Cutting edge: CD69 interference with sphingosine-1-phosphate receptor function regulates peripheral T cell retention*. J Immunol, 2015. **194**(5): p. 2059-63.
54. Doyle, C. and J.L. Strominger, *Interaction between CD4 and class II MHC molecules mediates cell adhesions*. Nature, 1987. **330**: p. 256.
55. Kaye, J., et al., *Selective development of CD4+ T cells in transgenic mice expressing a class II MHC-restricted antigen receptor*. Nature, 1989. **341**(6244): p. 746-9.
56. Murray, J.S., et al., *MHC control of CD4+ T cell subset activation*. J Exp Med, 1989. **170**(6): p. 2135-40.
57. Constant, S., et al., *Peptide and protein antigens require distinct antigen-presenting cell subsets for the priming of CD4+ T cells*. J Immunol, 1995. **154**(10): p. 4915-23.
58. Hong, S.C., et al., *The orientation of a T cell receptor to its MHC class II:peptide ligands*. J Immunol, 1997. **159**(9): p. 4395-402.
59. Ramirez, M.C. and L.J. Sigal, *Macrophages and dendritic cells use the cytosolic pathway to rapidly cross-present antigen from live, vaccinia-infected cells*. J Immunol, 2002. **169**(12): p. 6733-42.
60. Kohlmeier, J.E., et al., *CXCR3 directs antigen-specific effector CD4+ T cell migration to the lung during parainfluenza virus infection*. J Immunol, 2009. **183**(7): p. 4378-84.
61. Kobayashi, N., et al., *Functional and phenotypic analysis of human memory CD8+ T cells expressing CXCR3*. J Leukoc Biol, 2006. **80**(2): p. 320-9.
62. von Andrian, U.H. and C.R. Mackay, *T-cell function and migration. Two sides of the same coin*. N Engl J Med, 2000. **343**(14): p. 1020-34.

63. Kleinschek, M.A., et al., *IL-25 regulates Th17 function in autoimmune inflammation*. J Exp Med, 2007. **204**(1): p. 161-70.
64. Dardalhon, V., et al., *Role of Th1 and Th17 cells in organ-specific autoimmunity*. J Autoimmun, 2008. **31**(3): p. 252-6.
65. Zhang, D.H., et al., *Inhibition of allergic inflammation in a murine model of asthma by expression of a dominant-negative mutant of GATA-3*. Immunity, 1999. **11**(4): p. 473-82.
66. Piggott, D.A., et al., *MyD88-dependent induction of allergic Th2 responses to intranasal antigen*. J Clin Invest, 2005. **115**(2): p. 459-67.
67. Taylor, R.C., P. Richmond, and J.W. Upham, *Toll-like receptor 2 ligands inhibit TH2 responses to mite allergen*. J Allergy Clin Immunol, 2006. **117**(5): p. 1148-54.
68. Walsh, P.T., D.K. Taylor, and L.A. Turka, *Tregs and transplantation tolerance*. J Clin Invest, 2004. **114**(10): p. 1398-403.
69. Fehervari, Z. and S. Sakaguchi, *CD4+ Tregs and immune control*. J Clin Invest, 2004. **114**(9): p. 1209-17.
70. Mosmann, T.R. and R.L. Coffman, *TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties*. Annu Rev Immunol, 1989. **7**: p. 145-73.
71. Schlapbach, C., et al., *Human TH9 cells are skin-tropic and have autocrine and paracrine proinflammatory capacity*. Sci Transl Med, 2014. **6**(219): p. 219ra8.
72. Kaplan, M.H., M.M. Hufford, and M.R. Olson, *The development and in vivo function of T helper 9 cells*. Nat Rev Immunol, 2015. **15**(5): p. 295-307.
73. Ivanov, II, et al., *The orphan nuclear receptor ROR γ directs the differentiation program of proinflammatory IL-17+ T helper cells*. Cell, 2006. **126**(6): p. 1121-33.
74. Fontenot, J.D., M.A. Gavin, and A.Y. Rudensky, *Foxp3 programs the development and function of CD4+CD25+ regulatory T cells*. Nat Immunol, 2003. **4**(4): p. 330-6.
75. Hori, S., T. Nomura, and S. Sakaguchi, *Control of regulatory T cell development by the transcription factor Foxp3*. Science, 2003. **299**(5609): p. 1057-61.
76. Zhang, D.H., et al., *Transcription factor GATA-3 is differentially expressed in murine Th1 and Th2 cells and controls Th2-specific expression of the interleukin-5 gene*. J Biol Chem, 1997. **272**(34): p. 21597-603.
77. Zheng, W. and R.A. Flavell, *The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells*. Cell, 1997. **89**(4): p. 587-96.
78. Yu, D., et al., *The transcriptional repressor Bcl-6 directs T follicular helper cell lineage commitment*. Immunity, 2009. **31**(3): p. 457-68.
79. Szabo, S.J., et al., *A novel transcription factor, T-bet, directs Th1 lineage commitment*. Cell, 2000. **100**(6): p. 655-69.
80. Mullen, A.C., et al., *Role of T-bet in commitment of TH1 cells before IL-12-dependent selection*. Science, 2001. **292**(5523): p. 1907-10.
81. Kanno, Y., et al., *Transcriptional and epigenetic control of T helper cell specification: molecular mechanisms underlying commitment and plasticity*. Annu Rev Immunol, 2012. **30**: p. 707-31.
82. Nakayamada, S., et al., *Helper T cell diversity and plasticity*. Curr Opin Immunol, 2012. **24**(3): p. 297-302.
83. Ahmadzadeh, M. and D.L. Farber, *Functional plasticity of an antigen-specific memory CD4 T cell population*. Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11802-11807.
84. Krawczyk, C.M., H. Shen, and E.J. Pearce, *Functional plasticity in memory T helper cell responses*. J Immunol, 2007. **178**(7): p. 4080-8.

85. Kagi, D., et al., *The roles of perforin- and Fas-dependent cytotoxicity in protection against cytopathic and noncytopathic viruses*. Eur J Immunol, 1995. **25**(12): p. 3256-62.
86. Trapani, J.A. and M.J. Smyth, *Functional significance of the perforin/granzyme cell death pathway*. Nat Rev Immunol, 2002. **2**(10): p. 735-47.
87. Eischen, C.M., et al., *ZAP-70 tyrosine kinase is required for the up-regulation of Fas ligand in activation-induced T cell apoptosis*. J Immunol, 1997. **159**(3): p. 1135-9.
88. Gourley, T.S. and C.H. Chang, *Cutting edge: the class II transactivator prevents activation-induced cell death by inhibiting Fas ligand gene expression*. J Immunol, 2001. **166**(5): p. 2917-21.
89. Refaeli, Y., et al., *Interferon gamma is required for activation-induced death of T lymphocytes*. J Exp Med, 2002. **196**(7): p. 999-1005.
90. Kerdiles, Y.M., et al., *Foxo1 links homing and survival of naive T cells by regulating L-selectin, CCR7 and interleukin 7 receptor*. Nat Immunol, 2009. **10**(2): p. 176-84.
91. Joshi, N.S., et al., *Inflammation directs memory precursor and short-lived effector CD8(+) T cell fates via the graded expression of T-bet transcription factor*. Immunity, 2007. **27**(2): p. 281-95.
92. Lanzavecchia, A. and F. Sallusto, *Antigen decoding by T lymphocytes: from synapses to fate determination*. Nat Immunol, 2001. **2**(6): p. 487-92.
93. Sullivan, B.M., et al., *Antigen-driven effector CD8 T cell function regulated by T-bet*. Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15818-23.
94. Intlekofer, A.M., et al., *Effector and memory CD8+ T cell fate coupled by T-bet and eomesodermin*. Nat Immunol, 2005. **6**(12): p. 1236-44.
95. Pearce, E.L., et al., *Control of effector CD8+ T cell function by the transcription factor Eomesodermin*. Science, 2003. **302**(5647): p. 1041-3.
96. Intlekofer, A.M., et al., *Anomalous type 17 response to viral infection by CD8+ T cells lacking T-bet and eomesodermin*. Science, 2008. **321**(5887): p. 408-11.
97. Popescu, I., et al., *T-bet:Eomes balance, effector function, and proliferation of cytomegalovirus-specific CD8+ T cells during primary infection differentiates the capacity for durable immune control*. J Immunol, 2014. **193**(11): p. 5709-22.
98. Paley, M.A., et al., *Progenitor and terminal subsets of CD8+ T cells cooperate to contain chronic viral infection*. Science, 2012. **338**(6111): p. 1220-5.
99. Sallusto, F., et al., *Two subsets of memory T lymphocytes with distinct homing potentials and effector functions [see comments]*. Nature, 1999. **401**(6754): p. 708-12.
100. Forster, R., et al., *CCR7 coordinates the primary immune response by establishing functional microenvironments in secondary lymphoid organs*. Cell, 1999. **99**(1): p. 23-33.
101. Campbell, J.J., et al., *CCR7 expression and memory T cell diversity in humans*. J Immunol, 2001. **166**(2): p. 877-84.
102. Henning, G., et al., *CC chemokine receptor 7-dependent and -independent pathways for lymphocyte homing: modulation by FTY720*. J Exp Med, 2001. **194**(12): p. 1875-81.
103. Bromley, S.K., S.Y. Thomas, and A.D. Luster, *Chemokine receptor CCR7 guides T cell exit from peripheral tissues and entry into afferent lymphatics*. Nat Immunol, 2005. **6**(9): p. 895-901.
104. Unsoeld, H., et al., *Constitutive expression of CCR7 directs effector CD8 T cells into the splenic white pulp and impairs functional activity*. J Immunol, 2004. **173**(5): p. 3013-9.
105. Merckenschlager, M. and P.C. Beverley, *Evidence for differential expression of CD45 isoforms by precursors for memory-dependent and independent cytotoxic responses:*

- human CD8 memory CTLp selectively express CD45RO (UCHL1)*. Int Immunol, 1989. **1**(4): p. 450-9.
106. Wallace, D.L. and P.C.L. Beverley, *Phenotypic changes associated with activation of CD45RA⁺ and CD45RO⁺ T cells*. Immunol., 1990. **69**: p. 460-467.
 107. Rogers, P.R., C. Dubey, and S.L. Swain, *Qualitative changes accompany memory T cell generation: faster, more effective responses at lower doses of antigen*. J Immunol, 2000. **164**(5): p. 2338-46.
 108. Berard, M. and D.F. Tough, *Qualitative differences between naive and memory T cells*. Immunology, 2002. **106**(2): p. 127-38.
 109. Kedl, R.M. and M.F. Mescher, *Qualitative differences between naive and memory T cells make a major contribution to the more rapid and efficient memory CD8⁺ T cell response*. JI, 1998. **161**: p. 674-683.
 110. Zhang, X., et al., *Potent and selective stimulation of memory-phenotype CD8⁺ T cells in vivo by IL-15*. Immunity, 1998. **8**(5): p. 591-9.
 111. Khan, N., et al., *Cytomegalovirus seropositivity drives the CD8 T cell repertoire toward greater clonality in healthy elderly individuals*. J Immunol, 2002. **169**(4): p. 1984-92.
 112. Wills, M.R., et al., *Identification of naive or antigen-experienced human CD8(+) T cells by expression of costimulation and chemokine receptors: analysis of the human cytomegalovirus-specific CD8(+) T cell response*. J Immunol, 2002. **168**(11): p. 5455-64.
 113. Abdelsamed, H.A., et al., *Human memory CD8 T cell effector potential is epigenetically preserved during in vivo homeostasis*. J Exp Med, 2017. **214**(6): p. 1593-1606.
 114. Fuentes Marraco, S.A., et al., *Long-lasting stem cell-like memory CD8⁺ T cells with a naive-like profile upon yellow fever vaccination*. Sci Transl Med, 2015. **7**(282): p. 282ra48.
 115. Lugli, E., et al., *Superior T memory stem cell persistence supports long-lived T cell memory*. J Clin Invest, 2013. **123**(2): p. 594-9.
 116. Lugli, E., et al., *Identification, isolation and in vitro expansion of human and nonhuman primate T stem cell memory cells*. Nat Protoc, 2013. **8**(1): p. 33-42.
 117. Manickasingham, S.P., et al., *Qualitative and quantitative effects of CD28/B7-mediated costimulation on naive T cells in vitro*. J Immunol, 1998. **161**(8): p. 3827-35.
 118. Alam, S.M., et al., *Qualitative and quantitative differences in T cell receptor binding of agonist and antagonist ligands*. Immunity, 1999. **10**(2): p. 227-37.
 119. Croft, M., L.M. Bradley, and S.L. Swain, *Naive versus memory CD4 T cell response to antigen. Memory cells are less dependent on accessory cell costimulation and can respond to many antigen-presenting cell types including resting B cells*. J Immunol, 1994. **152**(6): p. 2675-85.
 120. von Fliedner, V., et al., *Production of tumor necrosis factor- α by naive or memory T lymphocytes activated via CD28*. Cell. Immunol., 1992. **139**: p. 198-207.
 121. Sallusto, F., et al., *Two subsets of memory T lymphocytes with distinct homing potentials and effector functions*. Nature, 1999. **401**(6754): p. 708-12.
 122. Darrah, P.A., et al., *Multifunctional TH1 cells define a correlate of vaccine-mediated protection against Leishmania major*. Nat Med, 2007. **13**(7): p. 843-50.
 123. Lachmann, R., et al., *Polyfunctional T cells accumulate in large human cytomegalovirus-specific T cell responses*. J Virol, 2012. **86**(2): p. 1001-9.
 124. Casazza, J.P., et al., *Acquisition of direct antiviral effector functions by CMV-specific CD4⁺ T lymphocytes with cellular maturation*. J Exp Med, 2006. **203**(13): p. 2865-77.

125. Henson, S.M., N.E. Riddell, and A.N. Akbar, *Properties of end-stage human T cells defined by CD45RA re-expression*. *Curr Opin Immunol*, 2012. **24**(4): p. 476-81.
126. Brenchley, J.M., et al., *Expression of CD57 defines replicative senescence and antigen-induced apoptotic death of CD8⁺ T cells*. *Blood*, 2003. **101**(7): p. 2711-20.
127. Beatty, G.L., et al., *Functional unresponsiveness and replicative senescence of myeloid leukemia antigen-specific CD8⁺ T cells after allogeneic stem cell transplantation*. *Clin Cancer Res*, 2009. **15**(15): p. 4944-53.
128. Koch, S., et al., *Human cytomegalovirus infection and T cell immunosenescence: a mini review*. *Mech Ageing Dev*, 2006. **127**(6): p. 538-43.
129. Tian, Y., et al., *Unique phenotypes and clonal expansions of human CD4 effector memory T cells re-expressing CD45RA*. *Nat Commun*, 2017. **8**(1): p. 1473.
130. Weiskopf, D., et al., *Dengue virus infection elicits highly polarized CX3CR1⁺ cytotoxic CD4⁺ T cells associated with protective immunity*. *Proc Natl Acad Sci U S A*, 2015. **112**(31): p. E4256-63.
131. Kaech, S.M., et al., *Molecular and functional profiling of memory CD8 T cell differentiation*. *Cell*, 2002. **111**(6): p. 837-51.
132. Weng, N.P., Y. Araki, and K. Subedi, *The molecular basis of the memory T cell response: differential gene expression and its epigenetic regulation*. *Nat Rev Immunol*, 2012. **12**(4): p. 306-15.
133. Araki, Y., et al., *Genome-wide analysis of histone methylation reveals chromatin state-based regulation of gene transcription and function of memory CD8⁺ T cells*. *Immunity*, 2009. **30**(6): p. 912-25.
134. Steinfeldt, S., et al., *Epigenetic modification of the human CCR6 gene is associated with stable CCR6 expression in T cells*. *Blood*, 2011. **117**(10): p. 2839-46.
135. Northrop, J.K., et al., *Epigenetic remodeling of the IL-2 and IFN-gamma loci in memory CD8 T cells is influenced by CD4 T cells*. *J Immunol*, 2006. **177**(2): p. 1062-9.
136. Kersh, E.N., et al., *Rapid demethylation of the IFN-gamma gene occurs in memory but not naive CD8 T cells*. *J Immunol*, 2006. **176**(7): p. 4083-93.
137. Syrbe, U., et al., *Differential regulation of P-selectin ligand expression in naive versus memory CD4⁺ T cells: evidence for epigenetic regulation of involved glycosyltransferase genes*. *Blood*, 2004. **104**(10): p. 3243-8.
138. Fitzpatrick, D.R., K.M. Shirley, and A. Kelso, *Cutting edge: stable epigenetic inheritance of regional IFN-gamma promoter demethylation in CD44^{high}CD8⁺ T lymphocytes*. *J Immunol*, 1999. **162**(9): p. 5053-7.
139. Rodriguez, R.M., et al., *Epigenetic Networks Regulate the Transcriptional Program in Memory and Terminally Differentiated CD8⁺ T Cells*. *J Immunol*, 2017. **198**(2): p. 937-949.
140. Zhou, X. and H.H. Xue, *Cutting edge: generation of memory precursors and functional memory CD8⁺ T cells depends on T cell factor-1 and lymphoid enhancer-binding factor-1*. *J Immunol*, 2012. **189**(6): p. 2722-6.
141. Sharma, A., et al., *T cell factor-1 and beta-catenin control the development of memory-like CD8 thymocytes*. *J Immunol*, 2012. **188**(8): p. 3859-68.
142. Xue, H.H. and D.M. Zhao, *Regulation of mature T cell responses by the Wnt signaling pathway*. *Ann N Y Acad Sci*, 2012. **1247**: p. 16-33.
143. Zhou, X., et al., *Differentiation and persistence of memory CD8(+) T cells depend on T cell factor 1*. *Immunity*, 2010. **33**(2): p. 229-40.

144. Jeannet, G., et al., *Essential role of the Wnt pathway effector Tcf-1 for the establishment of functional CD8 T cell memory*. Proc Natl Acad Sci U S A, 2010. **107**(21): p. 9777-82.
145. Zhao, D.M., et al., *Constitutive activation of Wnt signaling favors generation of memory CD8 T cells*. J Immunol, 2010. **184**(3): p. 1191-9.
146. Boudousquie, C., et al., *Differences in the transduction of canonical Wnt signals demarcate effector and memory CD8 T cells with distinct recall proliferation capacity*. J Immunol, 2014. **193**(6): p. 2784-91.
147. Germar, K., et al., *T-cell factor 1 is a gatekeeper for T-cell specification in response to Notch signaling*. Proc Natl Acad Sci U S A, 2011. **108**(50): p. 20060-5.
148. Verbeek, S., et al., *An HMG-box-containing T-cell factor required for thymocyte differentiation*. Nature, 1995. **374**(6517): p. 70-4.
149. Weber, B.N., et al., *A critical role for TCF-1 in T-lineage specification and differentiation*. Nature, 2011. **476**(7358): p. 63-8.
150. Wu, T., et al., *TCF1 Is Required for the T Follicular Helper Cell Response to Viral Infection*. Cell Rep, 2015. **12**(12): p. 2099-110.
151. Xu, L., et al., *The transcription factor TCF-1 initiates the differentiation of T(FH) cells during acute viral infection*. Nat Immunol, 2015. **16**(9): p. 991-9.
152. Choi, Y.S., et al., *LEF-1 and TCF-1 orchestrate T(FH) differentiation by regulating differentiation circuits upstream of the transcriptional repressor Bcl6*. Nat Immunol, 2015. **16**(9): p. 980-90.
153. Yu, Q., et al., *T cell factor 1 initiates the T helper type 2 fate by inducing the transcription factor GATA-3 and repressing interferon-gamma*. Nat Immunol, 2009. **10**(9): p. 992-9.
154. Wu, T., et al., *The TCF1-Bcl6 axis counteracts type I interferon to repress exhaustion and maintain T cell stemness*. Sci Immunol, 2016. **1**(6).
155. Schilham, M.W., et al., *Critical involvement of Tcf-1 in expansion of thymocytes*. J Immunol, 1998. **161**(8): p. 3984-91.
156. McLane, L.M., et al., *Differential localization of T-bet and Eomes in CD8 T cell memory populations*. J Immunol, 2013. **190**(7): p. 3207-15.
157. Kratchmarov, R., A.M. Magun, and S.L. Reiner, *TCF1 expression marks self-renewing human CD8(+) T cells*. Blood Adv, 2018. **2**(14): p. 1685-1690.
158. Willinger, T., et al., *Human naive CD8 T cells down-regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in vivo*. J Immunol, 2006. **176**(3): p. 1439-46.
159. Gaide, O., et al., *Common clonal origin of central and resident memory T cells following skin immunization*. Nat Med, 2015. **21**(6): p. 647-53.
160. Buchholz, V.R., et al., *Disparate individual fates compose robust CD8+ T cell immunity*. Science, 2013. **340**(6132): p. 630-5.
161. Gerlach, C., et al., *Heterogeneous differentiation patterns of individual CD8+ T cells*. Science, 2013. **340**(6132): p. 635-9.
162. Stemberger, C., et al., *A single naive CD8+ T cell precursor can develop into diverse effector and memory subsets*. Immunity, 2007. **27**(6): p. 985-97.
163. Restifo, N.P. and L. Gattinoni, *Lineage relationship of effector and memory T cells*. Curr Opin Immunol, 2013. **25**(5): p. 556-63.
164. Akondy, R.S., et al., *Origin and differentiation of human memory CD8 T cells after vaccination*. Nature, 2017. **552**(7685): p. 362-367.

165. Youngblood, B., et al., *Effector CD8 T cells dedifferentiate into long-lived memory cells*. Nature, 2017. **552**(7685): p. 404-409.
166. Durek, P., et al., *Epigenomic Profiling of Human CD4+ T Cells Supports a Linear Differentiation Model and Highlights Molecular Regulators of Memory Development*. Immunity, 2016. **45**(5): p. 1148-1161.
167. Farber, D.L., *Differential TCR signaling and the generation of memory T cells*. J Immunol, 1998. **160**(2): p. 535-9.
168. Nish, S.A., et al., *CD4+ T cell effector commitment coupled to self-renewal by asymmetric cell divisions*. J Exp Med, 2017. **214**(1): p. 39-47.
169. Adams, W.C., et al., *Anabolism-Associated Mitochondrial Stasis Driving Lymphocyte Differentiation over Self-Renewal*. Cell Rep, 2016. **17**(12): p. 3142-3152.
170. Lin, W.W., et al., *CD8(+) T Lymphocyte Self-Renewal during Effector Cell Determination*. Cell Rep, 2016. **17**(7): p. 1773-1782.
171. Lin, W.H., et al., *Asymmetric PI3K Signaling Driving Developmental and Regenerative Cell Fate Bifurcation*. Cell Rep, 2015. **13**(10): p. 2203-18.
172. Ciocca, M.L., et al., *Cutting edge: Asymmetric memory T cell division in response to rechallenge*. J Immunol, 2012. **188**(9): p. 4145-8.
173. Chang, J.T., et al., *Asymmetric proteasome segregation as a mechanism for unequal partitioning of the transcription factor T-bet during T lymphocyte division*. Immunity, 2011. **34**(4): p. 492-504.
174. Pollizzi, K.N., et al., *Asymmetric inheritance of mTORC1 kinase activity during division dictates CD8(+) T cell differentiation*. Nat Immunol, 2016. **17**(6): p. 704-11.
175. Verbist, K.C., et al., *Metabolic maintenance of cell asymmetry following division in activated T lymphocytes*. Nature, 2016. **532**(7599): p. 389-93.
176. Reiner, S.L. and W.C. Adams, *Lymphocyte fate specification as a deterministic but highly plastic process*. Nat Rev Immunol, 2014. **14**(10): p. 699-704.
177. Jiang, X., et al., *Skin infection generates non-migratory memory CD8+ TRM cells providing global skin immunity*. Nature, 2012. **483**(7388): p. 227-31.
178. Sakai, S., et al., *Cutting edge: control of Mycobacterium tuberculosis infection by a subset of lung parenchyma-homing CD4 T cells*. J Immunol, 2014. **192**(7): p. 2965-9.
179. Teijaro, J.R., et al., *Cutting edge: Tissue-retentive lung memory CD4 T cells mediate optimal protection to respiratory virus infection*. J Immunol, 2011. **187**(11): p. 5510-4.
180. Glennie, N.D., et al., *Skin-resident memory CD4+ T cells enhance protection against Leishmania major infection*. J Exp Med, 2015. **212**(9): p. 1405-14.
181. Gebhardt, T., et al., *Memory T cells in nonlymphoid tissue that provide enhanced local immunity during infection with herpes simplex virus*. Nat Immunol, 2009. **10**(5): p. 524-30.
182. Zens, K.D., J.-K. Chen, and D.L. Farber, *Vaccine-Generated Lung Tissue-Resident Memory T cells Provide Heterosubtypic Protection to Influenza Infection*. J. Clin. Invest. Insight, 2016. **1** (10): p. e85832.
183. Mackay, L.K., et al., *T-box Transcription Factors Combine with the Cytokines TGF-beta and IL-15 to Control Tissue-Resident Memory T Cell Fate*. Immunity, 2015. **43**(6): p. 1101-11.
184. Masopust, D., et al., *Activated primary and memory CD8 T cells migrate to nonlymphoid tissues regardless of site of activation or tissue of origin*. J Immunol, 2004. **172**(8): p. 4875-82.

185. Bingaman, A.W., et al., *Novel phenotypes and migratory properties distinguish memory CD4 T cell subsets in lymphoid and lung tissue*. Eur J Immunol, 2005. **35**: p. 3173-3186.
186. Masopust, D., et al., *Preferential localization of effector memory cells in nonlymphoid tissue*. Science, 2001. **291**(5512): p. 2413-7.
187. Wakim, L.M., A. Woodward-Davis, and M.J. Bevan, *Memory T cells persisting within the brain after local infection show functional adaptations to their tissue of residence*. Proc Natl Acad Sci U S A, 2010. **107**(42): p. 17872-9.
188. Tokoyoda, K., et al., *Professional memory CD4+ T lymphocytes preferentially reside and rest in the bone marrow*. Immunity, 2009. **30**(5): p. 721-30.
189. Steinert, E.M., et al., *Quantifying Memory CD8 T Cells Reveals Regionalization of Immunosurveillance*. Cell, 2015. **161**(4): p. 737-49.
190. Turner, D.L., et al., *Lung niches for the generation and maintenance of tissue-resident memory T cells*. Mucosal Immunol, 2014. **7**(3): p. 501-10.
191. Purwar, R., et al., *Resident memory T cells (T(RM)) are abundant in human lung: diversity, function, and antigen specificity*. PLoS One, 2011. **6**(1): p. e16245.
192. Clark, R.A., et al., *Skin effector memory T cells do not recirculate and provide immune protection in alemtuzumab-treated CTCL patients*. Sci Transl Med, 2012. **4**(117): p. 117ra7.
193. Pallett, L.J., et al., *IL-2high tissue-resident T cells in the human liver: Sentinels for hepatotropic infection*. J Exp Med, 2017. **214**(6): p. 1567-1580.
194. Woon, H.G., et al., *Compartmentalization of Total and Virus-Specific Tissue-Resident Memory CD8+ T Cells in Human Lymphoid Organs*. PLoS Pathog, 2016. **12**(8): p. e1005799.
195. Wong, M.T., et al., *A High-Dimensional Atlas of Human T Cell Diversity Reveals Tissue-Specific Trafficking and Cytokine Signatures*. Immunity, 2016. **45**(2): p. 442-56.
196. Hombrink, P., et al., *Programs for the persistence, vigilance and control of human CD8+ lung-resident memory T cells*. Nat Immunol, 2016. **17**(12): p. 1467-1478.
197. Booth, J.S., et al., *Characterization and functional properties of gastric tissue-resident memory T cells from children, adults, and the elderly*. Front Immunol, 2014. **5**: p. 294.
198. Shin, H. and A. Iwasaki, *A vaccine strategy that protects against genital herpes by establishing local memory T cells*. Nature, 2012. **491**(7424): p. 463-7.
199. Morawski, P.A., C.F. Qi, and S. Bolland, *Non-pathogenic tissue-resident CD8+ T cells uniquely accumulate in the brains of lupus-prone mice*. Sci Rep, 2017. **7**: p. 40838.
200. Hondowicz, B.D., et al., *Interleukin-2-Dependent Allergen-Specific Tissue-Resident Memory Cells Drive Asthma*. Immunity, 2016. **44**(1): p. 155-66.
201. Clark, R.A., *Resident memory T cells in human health and disease*. Sci Transl Med, 2015. **7**(269): p. 269rv1.
202. Park, C.O. and T.S. Kupper, *The emerging role of resident memory T cells in protective immunity and inflammatory disease*. Nat Med, 2015. **21**(7): p. 688-97.
203. Masopust, D. and L.J. Picker, *Hidden memories: frontline memory T cells and early pathogen interception*. J Immunol, 2012. **188**(12): p. 5811-7.
204. Gasteiger, G., et al., *Tissue residency of innate lymphoid cells in lymphoid and nonlymphoid organs*. Science, 2015. **350**(6263): p. 981-5.
205. Anderson, K.G., et al., *Intravascular staining for discrimination of vascular and tissue leukocytes*. Nat Protoc, 2014. **9**(1): p. 209-22.
206. Ugur, M., et al., *Resident CD4+ T cells accumulate in lymphoid organs after prolonged antigen exposure*. Nat Commun, 2014. **5**: p. 4821.

207. Tomura, M., et al., *Monitoring cellular movement in vivo with photoconvertible fluorescence protein "Kaede" transgenic mice*. Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10871-6.
208. Marriott, C.L., et al., *Retention of Ag-specific memory CD4(+) T cells in the draining lymph node indicates lymphoid tissue resident memory populations*. Eur J Immunol, 2017. **47**(5): p. 860-871.
209. Mueller, S.N. and L.K. Mackay, *Tissue-resident memory T cells: local specialists in immune defence*. Nat Rev Immunol, 2016. **16**(2): p. 79-89.
210. Bergsbaken, T. and M.J. Bevan, *Proinflammatory microenvironments within the intestine regulate the differentiation of tissue-resident CD8(+) T cells responding to infection*. Nat Immunol, 2015. **16**(4): p. 406-14.
211. Beura, L.K., et al., *T Cells in Nonlymphoid Tissues Give Rise to Lymph-Node-Resident Memory T Cells*. Immunity, 2018. **48**(2): p. 327-338.e5.
212. Thom, J.T. and A. Oxenius, *Tissue-resident memory T cells in cytomegalovirus infection*. Curr Opin Virol, 2016. **16**: p. 63-69.
213. Masopust, D., et al., *Dynamic T cell migration program provides resident memory within intestinal epithelium*. J Exp Med, 2010. **207**(3): p. 553-64.
214. Schenkel, J.M., et al., *Sensing and alarm function of resident memory CD8 T cells*. Nat Immunol, 2013. **14**(509-13).
215. Klonowski, K.D., et al., *Dynamics of blood-borne CD8 memory T cell migration in vivo*. Immunity, 2004. **20**(5): p. 551-62.
216. Teijaro, J.R., et al., *Tissue-Retentive lung memory CD4 T cells mediate optimal protection to respiratory virus infection*. J. Immunol., 2011. **187**(11): p. 5510-5514.
217. Takamura, S., et al., *Specific niches for lung-resident memory CD8+ T cells at the site of tissue regeneration enable CD69-independent maintenance*. J. Exp. Med., 2016. **213**(13): p. 3057-3073.
218. Anderson, K.G., et al., *Cutting edge: intravascular staining redefines lung CD8 T cell responses*. J Immunol, 2012. **189**(6): p. 2702-6.
219. Kim, Y.H., et al., *Long-term outcome of 525 patients with mycosis fungoides and Sezary syndrome: clinical prognostic factors and risk for disease progression*. Arch Dermatol, 2003. **139**(7): p. 857-66.
220. Zuber, J., et al., *Macrochimerism in Intestinal Transplantation: Association With Lower Rejection Rates and Multivisceral Transplants, Without GVHD*. Am J Transplant, 2015. **15**(10): p. 2691-703.
221. Weiner, J., et al., *Long-Term Persistence of Innate Lymphoid Cells in the Gut After Intestinal Transplantation*. Transplantation, 2016.
222. Lian, C.G., et al., *Biomarker evaluation of face transplant rejection: association of donor T cells with target cell injury*. Mod. Pathol., 2014. **27**(6): p. 788-799.
223. Okhrimenko, A., et al., *Human memory T cells from the bone marrow are resting and maintain long-lasting systemic memory*. Proc Natl Acad Sci U S A, 2014. **111**(25): p. 9229-34.
224. Swaims-Kohlmeier, A., et al., *Progesterone Levels Associate with a Novel Population of CCR5+CD38+ CD4 T Cells Resident in the Genital Mucosa with Lymphoid Trafficking Potential*. J Immunol, 2016. **197**(1): p. 368-76.
225. Radenkovic, M., et al., *Characterization of resident lymphocytes in human pancreatic islets*. Clin Exp Immunol, 2017. **187**(3): p. 418-427.

226. Smolders, J., et al., *Tissue-resident memory T cells populate the human brain*. Nat. Commun., 2018. **9**(1): p. 4593.
227. Oja, A.E., et al., *Trigger-happy resident memory CD4+ T cells inhabit the human lungs*. Mucosal Immunol., 2018. **11**(3): p. 654-667.
228. Muruganandah, V., et al., *A Systematic Review: The Role of Resident Memory T Cells in Infectious Diseases and Their Relevance for Vaccine Development*. Front. Immunol., 2018. **9**.
229. Cepek, K.L., et al., *Adhesion between epithelial cells and T lymphocytes mediated by E-cadherin and the alpha E beta 7 integrin*. Nature, 1994. **372**(6502): p. 190-3.
230. Gebhardt, T., et al., *Different patterns of peripheral migration by memory CD4+ and CD8+ T cells*. Nature, 2011. **477**(7363): p. 216-9.
231. Pham, T.H.M., et al., *SIP1 receptor signaling overrides retention mediated by G alpha i-coupled receptors to promote T cell egress*. Immunity, 2008. **28**(1): p. 122-133.
232. Skon, C.N., et al., *Transcriptional downregulation of *Slpr1* is required for the establishment of resident memory CD8+ T cells*. Nat Immunol, 2013. **14**(12): p. 1285-93.
233. Mackay, L.K., et al., *The developmental pathway for CD103(+)*CD8+* tissue-resident memory T cells of skin*. Nat Immunol, 2013. **14**(12): p. 1294-301.
234. Zaid, A., et al., *Chemokine Receptor-Dependent Control of Skin Tissue-Resident Memory T Cell Formation*. J. Immunol., 2017. **199**(7): p. 2451-2459.
235. Schenkel, J.M., et al., *Resident memory CD8 T cells trigger protective innate and adaptive immune responses*. Science, 2014.
236. Schenkel, J.M., et al., *T cell memory. Resident memory CD8 T cells trigger protective innate and adaptive immune responses*. Science, 2014. **346**(6205): p. 98-101.
237. Schenkel, J.M. and D. Masopust, *Tissue-Resident Memory T Cells*. Immunity, 2014. **41**(6): p. 886-897.
238. Watanabe, R., et al., *Human skin is protected by four functionally and phenotypically discrete populations of resident and recirculating memory T cells*. Sci Transl Med, 2015. **7**(279): p. 279ra39.
239. Collins, N., et al., *Skin CD4(+) *memory T cells exhibit combined cluster-mediated retention and equilibration with the circulation**. Nat Commun, 2016. **7**: p. 11514.
240. Cheuk, S., et al., *CD49a Expression Defines Tissue-Resident CD8+ T Cells Poised for Cytotoxic Function in Human Skin*. Immunity, 2017. **46**(2): p. 287-300.
241. Roberts, A.I., R.E. Brolin, and E.C. Ebert, *Integrin alpha1beta1 (VLA-1) mediates adhesion of activated intraepithelial lymphocytes to collagen*. Immunology, 1999. **97**(4): p. 679-685.
242. Ray, S.J., et al., *The collagen binding alpha1beta1 integrin VLA-1 regulates CD8 T cell-mediated immune protection against heterologous influenza infection*. Immunity, 2004. **20**(2): p. 167-79.
243. Tse, E. and Y.L. Kwong, *Epstein Barr virus-associated lymphoproliferative diseases: the virus as a therapeutic target*. Exp Mol Med, 2015. **47**: p. e136.
244. Soares, L.R., et al., *V7 (CD101) ligation inhibits TCR/CD3-induced IL-2 production by blocking Ca2+ flux and nuclear factor of activated T cell nuclear translocation*. J Immunol, 1998. **161**(1): p. 209-17.
245. Barber, D.L., et al., *Restoring function in exhausted CD8 T cells during chronic viral infection*. Nature, 2006. **439**(7077): p. 682-7.
246. Wherry, E.J., et al., *Molecular signature of CD8+ T cell exhaustion during chronic viral infection*. Immunity, 2007. **27**(4): p. 670-84.

247. Kao, C., et al., *Transcription factor T-bet represses expression of the inhibitory receptor PD-1 and sustains virus-specific CD8+ T cell responses during chronic infection*. Nat Immunol, 2011. **12**(7): p. 663-71.
248. Im, S.J., et al., *Defining CD8+ T cells that provide the proliferative burst after PD-1 therapy*. Nature, 2016. **537**(7620): p. 417-421.
249. Park, S.L., et al., *Local proliferation maintains a stable pool of tissue-resident memory T cells after antiviral recall responses*. Nat. Immunol., 2018. **19**(2): p. 183-191.
250. Shwetank, et al., *Maintenance of PD-1 on brain-resident memory CD8 T cells is antigen independent*. Immunol. Cell Biol., 2017. **95**(10): p. 953-959.
251. Ohaegbulam, K.C., et al., *Human cancer immunotherapy with antibodies to the PD-1 and PD-L1 pathway*. Trends Mol Med, 2015. **21**(1): p. 24-33.
252. Mackay, L.K. and A. Kallies, *Transcriptional Regulation of Tissue-Resident Lymphocytes*. Trends Immunol, 2017. **38**(2): p. 94-103.
253. Mackay, L.K., et al., *Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes*. Science, 2016. **352**(6284): p. 459-63.
254. Milner, J.J., et al., *Runx3 programs CD8+ T cell residency in non-lymphoid tissues and tumours*. Nature, 2017. **552**(7684): p. 253-257.
255. Vieira Braga, F.A., et al., *Blimp-1 homolog Hobit identifies effector-type lymphocytes in humans*. Eur J Immunol, 2015. **45**(10): p. 2945-58.
256. McCully, M.L., et al., *CCR8 Expression Defines Tissue-Resident Memory T Cells in Human Skin*. J. Immunol., 2018. **200**(5): p. 1639-1650.
257. Clark, R.A., et al., *The vast majority of CLA+ T cells are resident in normal skin*. J Immunol, 2006. **176**(7): p. 4431-9.
258. Mohammed, J., et al., *Stromal cells control the epithelial residence of DCs and memory T cells by regulated activation of TGF- β* . Nat. Immunol., 2016. **17**(4): p. 414-421.
259. Adachi, T., et al., *Hair follicle-derived IL-7 and IL-15 mediate skin-resident memory T cell homeostasis and lymphoma*. Nat. Med., 2015. **21**(11): p. 1272-1279.
260. Pan, Y., et al., *Survival of tissue-resident memory T cells requires exogenous lipid uptake and metabolism*. Nature, 2017. **543**(7644): p. 252-256.
261. Turner, D.L. and D.L. Farber, *Mucosal resident memory CD4 T cells in protection and immunopathology*. Front Immunol, 2014. **5**: p. 331.
262. Iborra, S., et al., *Optimal Generation of Tissue-Resident but Not Circulating Memory T Cells during Viral Infection Requires Crosspriming by DNGR-1+ Dendritic Cells*. Immunity, 2016. **45**(4): p. 847-860.
263. Zhou, A.C., et al., *Intrinsic 4-1BB signals are indispensable for the establishment of an influenza-specific tissue-resident memory CD8 T-cell population in the lung*. Mucosal Immunol., 2017. **10**(5): p. 1294-1309.
264. Laidlaw, B.J., et al., *CD4(+) T Cell Help Guides Formation of CD103(+) Lung-Resident Memory CD8(+) T Cells during Influenza Viral Infection*. Immunity, 2014. **41**(4): p. 633-45.
265. Wakim, L.M., et al., *Antibody-targeted vaccination to lung dendritic cells generates tissue-resident memory CD8 T cells that are highly protective against influenza virus infection*. Mucosal Immunol., 2015. **8**(5): p. 1060-1071.
266. Herndler-Brandstetter, D., et al., *Human bone marrow hosts polyfunctional memory CD4+ and CD8+ T cells with close contact to IL-15-producing cells*. J Immunol, 2011. **186**(12): p. 6965-71.

267. Palendira, U., et al., *Selective accumulation of virus-specific CD8⁺ T cells with unique homing phenotype within the human bone marrow*. 2008. **112**(8): p. 3293-3302.
268. Zhang, X., et al., *Human Bone Marrow: A Reservoir for “Enhanced Effector Memory” CD8⁺ T Cells with Potent Recall Function*. *The Journal of Immunology*, 2006. **177**: p. 6730-6737.
269. Becker, T.C., et al., *Bone marrow is a preferred site for homeostatic proliferation of memory CD8 T cells*. *J Immunol*, 2005. **174**(3): p. 1269-73.
270. Parretta, E., et al., *CD8 cell division maintaining cytotoxic memory occurs predominantly in the bone marrow*. *J Immunol*, 2005. **174**(12): p. 7654-64.
271. Di Rosa, F. and R. Pabst, *The bone marrow: a nest for migratory memory T cells*. *Trends Immunol*, 2005. **26**(7): p. 360-6.
272. Shinoda, K., et al., *Type II membrane protein CD69 regulates the formation of resting T-helper memory*. *Proc Natl Acad Sci U S A*, 2012. **109**(19): p. 7409-14.
273. Sercan Alp, O., et al., *Memory CD8(+) T cells colocalize with IL-7(+) stromal cells in bone marrow and rest in terms of proliferation and transcription*. *Eur J Immunol*, 2015. **45**(4): p. 975-87.
274. Schluns, K.S., K.D. Klonowski, and L. Lefrancois, *Transregulation of memory CD8 T-cell proliferation by IL-15R α + bone marrow-derived cells*. *Blood*, 2004. **103**(3): p. 988-994.
275. Buggert, M., et al., *Identification and characterization of HIV-specific resident memory CD8(+) T cells in human lymphoid tissue*. *Sci Immunol*, 2018. **3**(24).
276. Durand, A., et al., *Profiling the lymphoid-resident T cell pool reveals modulation by age and microbiota*. *Nat. Commun.*, 2018. **9**(1): p. 68.
277. Schenkel, J.M., K.A. Fraser, and D. Masopust, *Cutting edge: resident memory CD8 T cells occupy frontline niches in secondary lymphoid organs*. *J Immunol*, 2014. **192**(7): p. 2961-4.
278. Beura, L.K., et al., *Normalizing the environment recapitulates adult human immune traits in laboratory mice*. *Nature*, 2016. **532**(7600): p. 512-6.
279. Purton, J.F., et al., *Antiviral CD4+ memory T cells are IL-15 dependent*. *J Exp Med*, 2007. **204**(4): p. 951-61.
280. Wu, T., et al., *Lung-resident memory CD8 T cells (TRM) are indispensable for optimal cross-protection against pulmonary virus infection*. *J Leukoc Biol*, 2014. **95**(2): p. 215-24.
281. Slütter, B., et al., *Dynamics of influenza-induced lung-resident memory T cells underlie waning heterosubtypic immunity*. *Science Immunology*, 2017. **2**(7).
282. Piet, B., et al., *CD8(+) T cells with an intraepithelial phenotype upregulate cytotoxic function upon influenza infection in human lung*. *J Clin Invest*, 2011. **121**(6): p. 2254-63.
283. Matos, T.R., et al., *Clinically resolved psoriatic lesions contain psoriasis-specific IL-17-producing alphabeta T cell clones*. *J Clin Invest*, 2017. **127**(11): p. 4031-4041.
284. McElhaney, J.E., et al., *T cell responses are better correlates of vaccine protection in the elderly*. *J Immunol*, 2006. **176**(10): p. 6333-9.
285. Zhou, X. and J.E. McElhaney, *Age-related changes in memory and effector T cells responding to influenza A/H3N2 and pandemic A/H1N1 strains in humans*. *Vaccine*, 2011. **29**(11): p. 2169-77.
286. Palmer, S., et al., *Thymic involution and rising disease incidence with age*. *Proc Natl Acad Sci U S A*, 2018. **115**(8): p. 1883-1888.

287. Zens, K.D., et al., *Reduced generation of lung tissue-resident memory T cells during infancy*. J Exp Med, 2017. **214**(10): p. 2915-2932.
288. Mackay, L.K., et al., *T-box Transcription Factors Combine with the Cytokines TGF- β and IL-15 to Control Tissue-Resident Memory T Cell Fate*. Immunity, 2015. **43**(6): p. 1101-1111.
289. Perdomo, C., et al., *Mucosal BCG Vaccination Induces Protective Lung-Resident Memory T Cell Populations against Tuberculosis*. MBio, 2016. **7**(6).
290. Cuburu, N., et al., *Mouse model of cervicovaginal papillomavirus infection*. Methods Mol Biol, 2015. **1249**: p. 365-79.
291. Cuburu, N., et al., *Topical herpes simplex virus 2 (HSV-2) vaccination with human papillomavirus vectors expressing gB/gD ectodomains induces genital-tissue-resident memory CD8⁺ T cells and reduces genital disease and viral shedding after HSV-2 challenge*. J Virol, 2015. **89**(1): p. 83-96.
292. Cuburu, N., et al., *Intravaginal immunization with HPV vectors induces tissue-resident CD8⁺ T cell responses*. J Clin Invest, 2012. **122**(12): p. 4606-20.
293. Gola, A., et al., *Prime and target immunization protects against liver-stage malaria in mice*. Sci Transl Med, 2018. **10**(460).
294. Yopp, A.C., et al., *FTY720-enhanced T cell homing is dependent on CCR2, CCR5, CCR7, and CXCR4: evidence for distinct chemokine compartments*. J Immunol, 2004. **173**(2): p. 855-65.
295. Ramachandran, H., et al., *Optimal Thawing of Cryopreserved Peripheral Blood Mononuclear Cells for Use in High-Throughput Human Immune Monitoring Studies*. 2012. **1**(3): p. 313.
296. Van der Maaten, L., *Accelerating t-SNE using Tree-Based Algorithms*. Journal of Machine Learning Research, 2014. **15**: p. 3221-3245.
297. van der Maaten, L. and G. Hinton, *Visualizing data using t-SNE*. Mach. Learn. Res. , 2008. **9**: p. 2579-2605.
298. Amir el, A.D., et al., *viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia*. Nat Biotechnol, 2013. **31**(6): p. 545-52.
299. Trapnell, C., L. Pachter, and S.L. Salzberg, *TopHat: discovering splice junctions with RNA-Seq*. Bioinformatics, 2009. **25**(9): p. 1105-11.
300. Trapnell, C., et al., *Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation*. Nat Biotechnol, 2010. **28**(5): p. 511-5.
301. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biol, 2014. **15**(12): p. 550.
302. Im, S.J., et al., *Defining CD8(+) T cells that provide the proliferative burst after PD-1 therapy*. Nature, 2016. **537**(7620): p. 417-+.
303. Utschneider, D.T., et al., *T Cell Factor 1-Expressing Memory-like CD8(+) T Cells Sustain the Immune Response to Chronic Viral Infections*. Immunity, 2016. **45**(2): p. 415-27.
304. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles*. Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.
305. van Dongen, J.J., et al., *Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect*

- lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936.* Leukemia, 2003. **17**(12): p. 2257-317.
306. Vander Heiden, J.A., et al., *pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires.* Bioinformatics, 2014. **30**(13): p. 1930-2.
307. Meng, W., et al., *An atlas of B-cell clonal distribution in the human body.* Nat Biotechnol, 2017. **35**(9): p. 879-884.
308. Bolotin, D.A., et al., *MiXCR: software for comprehensive adaptive immunity profiling.* Nat Methods, 2015. **12**(5): p. 380-1.
309. Rosenfeld, A.M., et al., *ImmuneDB: a system for the analysis and exploration of high-throughput adaptive immune receptor sequencing data.* Bioinformatics, 2017. **33**(2): p. 292-293.
310. Rosenfeld, A.M., et al., *ImmuneDB, a Novel Tool for the Analysis, Storage, and Dissemination of Immune Repertoire Sequencing Data.* Front Immunol, 2018. **9**: p. 2107.
311. Schwartz, G.W. and U. Hershberg, *Conserved variation: identifying patterns of stability and variability in BCR and TCR V genes with different diversity and richness metrics.* Phys Biol, 2013. **10**(3): p. 035005.
312. Shugay, M., et al., *VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires.* PLoS Comput Biol, 2015. **11**(11): p. e1004503.
313. Sallusto, F., J. Geginat, and A. Lanzavecchia, *Central memory and effector memory T cell subsets: function, generation, and maintenance.* Annu Rev Immunol, 2004. **22**: p. 745-63.
314. Teijaro, J.R., et al., *Endothelial cells are central orchestrators of cytokine amplification during influenza virus infection.* Cell, 2011. **146**(6): p. 980-91.
315. Zhou, Q., et al., *Program death-1 signaling and regulatory T cells collaborate to resist the function of adoptively transferred cytotoxic T lymphocytes in advanced acute myeloid leukemia.* Blood, 2010. **116**(14): p. 2484-93.
316. He, R., et al., *Follicular CXCR5- expressing CD8(+) T cells curtail chronic viral infection.* Nature, 2016. **537**(7620): p. 412-428.
317. Leong, Y.A., et al., *CXCR5(+) follicular cytotoxic T cells control viral infection in B cell follicles.* Nat Immunol, 2016. **17**(10): p. 1187-96.
318. Chtanova, T., et al., *T follicular helper cells express a distinctive transcriptional profile, reflecting their role as non-Th1/Th2 effector cells that provide help for B cells.* J Immunol, 2004. **173**(1): p. 68-78.
319. Choi, Y.S., et al., *ICOS receptor instructs T follicular helper cell versus effector cell differentiation via induction of the transcriptional repressor Bcl6.* Immunity, 2011. **34**(6): p. 932-46.
320. Boddupalli, C.S., et al., *ABC transporters and NR4A1 identify a quiescent subset of tissue-resident memory T cells.* J Clin Invest, 2016. **126**(10): p. 3905-3916.
321. Tasaki, S., et al., *Multiomic disease signatures converge to cytotoxic CD8 T cells in primary Sjogren's syndrome.* Ann Rheum Dis, 2017. **76**(8): p. 1458-1466.
322. Chaix, J., et al., *Cutting edge: CXCR4 is critical for CD8+ memory T cell homeostatic self-renewal but not rechallenge self-renewal.* J Immunol, 2014. **193**(3): p. 1013-6.
323. Busch, D.H., et al., *Role of memory T cell subsets for adoptive immunotherapy.* Semin Immunol, 2016. **28**(1): p. 28-34.
324. June, C.H., et al., *CAR T cell immunotherapy for human cancer.* Science, 2018. **359**(6382): p. 1361-1365.

325. Plumlee, C.R., et al., *Environmental cues dictate the fate of individual CD8(+) T cells responding to infection*. *Immunity*, 2013. **39**(2): p. 347-56.
326. Boddupalli, C.S., et al., *Interlesional diversity of T cell receptors in melanoma with immune checkpoints enriched in tissue-resident memory T cells*. *JCI Insight*, 2016. **1**(21): p. e88955.
327. Zheng, C., et al., *Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing*. *Cell*, 2017. **169**(7): p. 1342-1356 e16.
328. Ganesan, A.P., et al., *Tissue-resident memory features are linked to the magnitude of cytotoxic T cell responses in human lung cancer*. *Nat Immunol*, 2017. **18**(8): p. 940-950.
329. Wang, T., et al., *The Different T-cell Receptor Repertoires in Breast Cancer Tumors, Draining Lymph Nodes, and Adjacent Tissues*. *Cancer Immunol Res*, 2017. **5**(2): p. 148-156.
330. Klarenbeek, P.L., et al., *Deep sequencing of antiviral T-cell responses to HCMV and EBV in humans reveals a stable repertoire that is maintained for many years*. *PLoS Pathog*, 2012. **8**(9): p. e1002889.
331. Rosenfeld, A.M., et al., *Computational Evaluation of B-Cell Clone Sizes in Bulk Populations*. *Front Immunol*, 2018. **9**: p. 1472.
332. Sant, S., et al., *Single-Cell Approach to Influenza-Specific CD8(+) T Cell Receptor Repertoires Across Different Age Groups, Tissues, and Following Influenza Virus Infection*. *Front Immunol*, 2018. **9**: p. 1453.
333. Wang, G.C., et al., *T cell receptor alphabeta diversity inversely correlates with pathogen-specific antibody levels in human cytomegalovirus infection*. *Sci Transl Med*, 2012. **4**(128): p. 128ra42.
334. Miron, M., et al., *Human Lymph Nodes Maintain TCF-1(hi) Memory T Cells with High Functional Potential and Clonal Diversity throughout Life*. *J Immunol*, 2018.
335. Di Rosa, F. and T. Gebhardt, *Bone Marrow T Cells and the Integrated Functions of Recirculating and Tissue-Resident Memory T Cells*. *Front Immunol*, 2016. **7**: p. 51.
336. Oakes, T., et al., *Quantitative Characterization of the T Cell Receptor Repertoire of Naive and Memory Subsets Using an Integrated Experimental and Computational Pipeline Which Is Robust, Economical, and Versatile*. *Front Immunol*, 2017. **8**: p. 1267.
337. Lin, K.R., et al., *T cell receptor repertoire profiling predicts the prognosis of HBV-associated hepatocellular carcinoma*. *Cancer Med*, 2018. **7**(8): p. 3755-3762.
338. Wong, Y.N.S., et al., *Urine-derived lymphocytes as a non-invasive measure of the bladder tumor immune microenvironment*. 2018.
339. Wherry, E.J., et al., *Lineage relationship and protective immunity of memory CD8 T cell subsets*. *Nat Immunol*, 2003. **4**(3): p. 225-34.
340. Wendel, B.S., et al., *The receptor repertoire and functional profile of follicular T cells in HIV-infected lymph nodes*. *Sci Immunol*, 2018. **3**(22).
341. Reinhardt, R.L., et al., *Visualizing the generation of memory CD4 T cells in the whole body*. *Nature*, 2001. **410**(6824): p. 101-5.
342. Zaph, C., et al., *Central memory T cells mediate long-term immunity to *Leishmania major* in the absence of persistent parasites*. *Nat Med*, 2004. **10**(10): p. 1104-10.
343. Brummelman, J., et al., *High-dimensional single cell analysis identifies stem-like cytotoxic CD8(+) T cells infiltrating human tumors*. *J Exp Med*, 2018. **215**(10): p. 2520-2535.
344. Spitzer, M.H., et al., *Systemic Immunity Is Required for Effective Cancer Immunotherapy*. *Cell*, 2017. **168**(3): p. 487-502 e15.

345. Scheper, W., et al., *Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers*. Nat Med, 2018.
346. Ljungman, P., M. Hakki, and M. Boeckh, *Cytomegalovirus in hematopoietic stem cell transplant recipients*. Infect Dis Clin North Am, 2010. **24**(2): p. 319-37.
347. Kotton, C.N., *CMV: Prevention, Diagnosis and Therapy*. Am J Transplant, 2013. **13** **Suppl 3**: p. 24-40; quiz 40.
348. Frantzeskaki, F.G., et al., *Cytomegalovirus reactivation in a general, nonimmunosuppressed intensive care unit population: incidence, risk factors, associations with organ dysfunction, and inflammatory biomarkers*. J Crit Care, 2015. **30**(2): p. 276-81.
349. Lichtner, M., et al., *Cytomegalovirus coinfection is associated with an increased risk of severe non-AIDS-defining events in a large cohort of HIV-infected patients*. J Infect Dis, 2015. **211**(2): p. 178-86.
350. Pera, A., et al., *CMV latent infection improves CD8+ T response to SEB due to expansion of polyfunctional CD57+ cells in young individuals*. PLoS One, 2014. **9**(2): p. e88538.
351. Arens, R., et al., *5(th) International Workshop on CMV and Immunosenescence - A shadow of cytomegalovirus infection on immunological memory*. Eur J Immunol, 2015. **45**(4): p. 954-7.
352. Furman, D., et al., *Cytomegalovirus infection enhances the immune response to influenza*. Sci Transl Med, 2015. **7**(281): p. 281ra43.

Appendices

Appendix A. CMV-specific T cell responses in tissues.

Only data resulting from experiments that I performed are presented in this section, and these data are taken from:

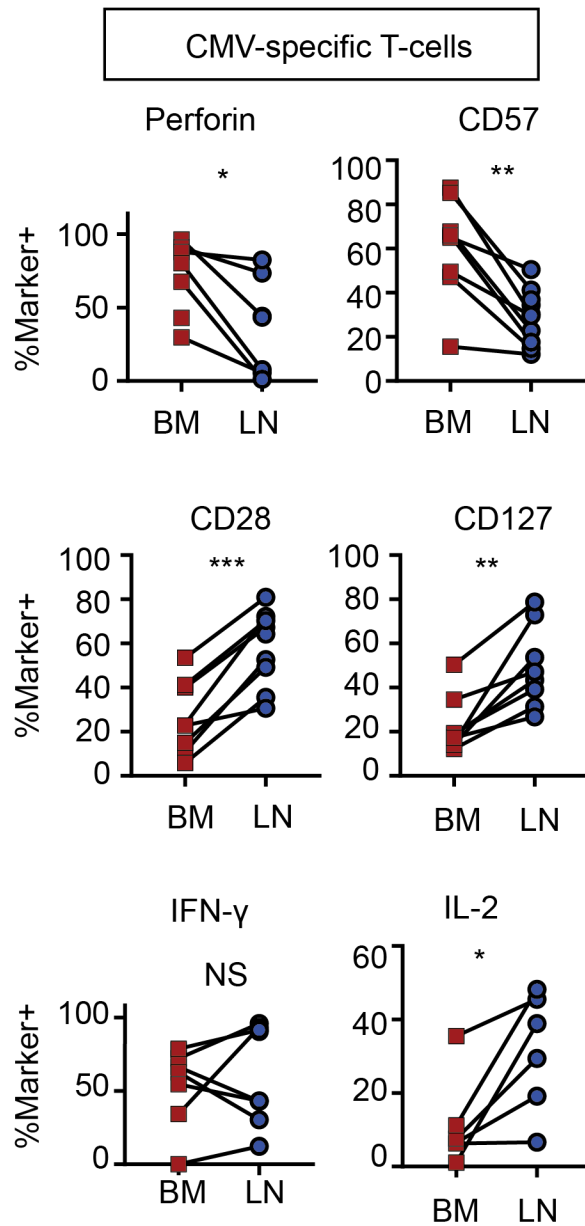
Gordon, C. L., **Miron, M.**, Thome, J. J., Matsuoka, N., Weiner, J., Rak, M. A., Igarashi, S., Granot, T., Lerner, H., Goodrum, F., Farber, D. L. (2017). Tissue reservoirs of antiviral T cell immunity in persistent human CMV infection. *J Exp Med*, 214(3), 651-667.
doi:10.1084/jem.20160758

T cell responses to viruses are initiated and maintained in tissue sites; however, knowledge of human antiviral T cells is largely derived from blood. Cytomegalovirus (CMV) persists in most humans, requires T cell immunity to control, yet tissue immune responses remain undefined. Here, we investigated human CMV-specific T cells, virus persistence and CMV-associated T cell homeostasis in blood, lymphoid, mucosal and secretory tissues of 44 CMV seropositive and 28 seronegative donors. CMV-specific T cells were maintained in distinct distribution patterns, highest in blood, bone marrow (BM), or lymph nodes (LN), with the frequency and function in blood distinct from tissues. CMV genomes were detected predominantly in lung and also in spleen, BM, blood and LN. High frequencies of activated CMV-specific T cells were found in blood and BM samples with low virus detection, whereas in lung, CMV-specific T cells were present along with detectable virus. In LNs, CMV-specific T cells exhibited quiescent phenotypes independent of virus. Overall, T cell differentiation was enhanced in sites of viral persistence with age. Together, our results suggest tissue T cell reservoirs for CMV control shaped by both viral and tissue-intrinsic factors, with global effects on homeostasis of tissue T cells over the lifespan.

Persistent CMV infection has important clinical relevance in the global population. While immune-mediated control of CMV in healthy individuals prevents disease and overt clinical symptoms, immune dysregulation due to immunosuppressive treatments in transplant and cancer patients, congenital immunodeficiencies, HIV/AIDS, and/or aging can result in CMV viremia, life threatening disease and even death [346-349]. When reactivated, CMV infects multiple tissues, causing pneumonitis, colitis, hepatitis and end organ failure [346, 347], for which anti-viral therapeutics are only partially effective. CMV persistence may also impact immunity in immunocompetent individuals, and has been associated with immunosenescence, and differential responses to infections or vaccinations [350-352]. The dynamic nature of viral persistence and anti-viral T cell responses indicates that both parameters need to be investigated to understand the mechanisms for immune-mediated escape and/or other effects on overall immune homeostasis.

We screened for the presence of CMV-specific CD8⁺ T cells in tissues from seropositive donors using up to five HLA-tetramer or -multimer reagents of the appropriate HLA containing peptide epitopes of immunodominant CMV proteins pp65, IE-1, and pp50 (CMV-multimers), compared to staining with negative control HLA-multimers. We used combinations of CMV multimers containing different CMV epitopes in order to maximize our ability to detect CMV-reactive T cells in each sites. We found that the phenotypes of CMV-specific T cells differed depending on the tissue site. CMV-specific T cells in LN displayed lower expression of CD57 and perforin and higher expression of CD28 and CD127 compared to CMV-specific T cells in BM (Figure A.A.1). Functionally, CMV-specific T cells in LN produced higher levels of IL-2 in response to PMA/INO stimulation, and comparable levels of IFN-gamma (Figure A.A.1).

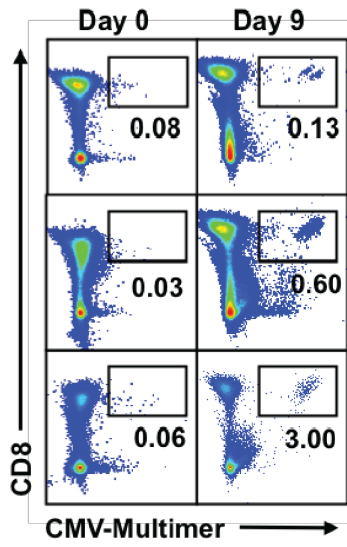
Figure A.A.1. Expression of various proteins in resting CD8⁺ CMV-specific T cells identified by MHC-peptide multimer staining of T cells from BM and LN and expression of cytokines in response to stimulation. Expression of Perforin, CD57 (top), CD28 and CD127 (middle). Expression of cytokines in response to PMA/Ionomycin stimulation (bottom). Paired t-test. NS $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Lines connect samples from the same donor.



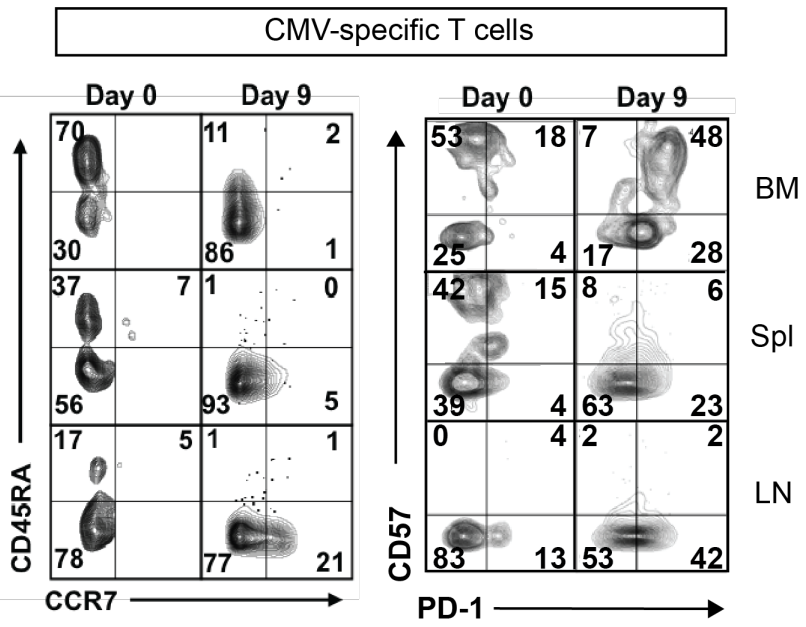
We analyzed the functional abilities of CMV-specific T cells in response to CMV-peptide, as described in the methods section, and found expansion of T cells was highest in LN, then Spl, and lowest in BM (Figure A.A.2). Further, we found differences in the phenotype of CMV-specific T cells in different sites; CMV-specific T cells in LN displayed TEM phenotype (CCR7-CD45RA-) and T cells in Spl and BM displayed TEM and TEMRA phenotypes (CCR7-CD45RA+). After expansion of CMV-specific T cells over the course of 9 days in culture, the phenotype of CMV-specific T cells changed; in the LN, CMV-specific T cells expressed CCR7, and in the BM and Spl CMV-specific T cells were mainly TEM phenotype. Additionally, CMV-specific T cells in the BM expressed CD57 before and after expansion, while CMV-specific T cells in the Spl expressed CD57 only before expansion. In the LN, T cells did not express CD57, but a portion expressed PD-1, and a higher proportion expressed PD-1 after expansion. Additionally, CMV-specific T cells displayed the highest PD-1 expression in the BM compared to Spl and LN after 9 days in culture (Figure A.A.2). We analyzed supernatant from Spl stimulated CMV-specific T cell cultures and detected increasing levels of IFN-gamma, and decreasing levels of IL-2 and TNF-alpha over the course of in-vitro stimulation (Figure A.A.2).

Figure A.A.2. Function and phenotype of BM, Spl, and LN CMV-specific T cells during antigen driven expansion. (A). Identification of CMV-specific T cells from D194 using the A24-QYD multimer. On day 0, total mononuclear cells from tissues were incubated with pp65 peptide and analyzed by flow cytometry on Day 9. See methods for detailed protocol. (B). Phenotype of CMV-specific T cell identified in (A) by expression of CCR7 and CD45RA (left) and expression of PD-1 and CD57 (right). (C). Supernatants from stimulated cells from spleen were analyzed on day 2,6, and 8 for detection of cytokines IL-2, TNF, and IFN-gamma.

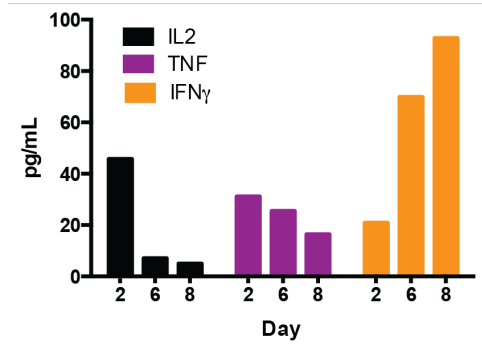
(A).



(B).



(C).



Appendix B. Transcriptional profiling of CD4⁺ and CD8⁺ TRM in bone marrow and lymph nodes

As discussed in Chapter 1, populations of TRM phenotype cells, marked by expression of CD69, are abundant in lymphoid sites such as BM and LN, however their properties relative to T cell subsets in circulation are not well characterized. We purified memory CD4⁺ and CD8⁺ T cells by FACS sorting and isolated total RNA for whole transcriptome profiling by RNA-sequencing (see Table A.B.1 below for description of samples). We used DESeq2 in order to identify genes differentially expressed between CD69⁺ and CD69⁻ memory CD4⁺ and CD8⁺ T cells in BM and LN (Figure A.B.1). For the complete list of genes differentially expressed see Table A.B.2 below.

In CD8⁺ T cells, we found about 200 genes differential expressed between CD69⁺ and CD69⁻ T cells in the BM; however, we found only 100 genes differentially expressed between CD69⁺ and CD69⁻ T cells in the LN indicating higher transcriptional similarity between CD69⁻ and CD69⁺ T cells in the LN compared to the BM (Figure A.B.3). When compared with the core TRM gene set identified in Kumar et al. 2017, CD69⁺CD8⁺ T cells were transcriptionally distinct from CD8⁺CD69⁻ memory T cells in BM and LN by MDS analysis, and CD69⁻ T cells were more similar to memory T cells from blood (Figure A.B.3). Included in the genes conserved with core TRM gene expression are downregulation of *KLF2*, *KLF3*, and *SIPRI* in CD69⁺ relative to CD69⁻ memory T cells in BM and LN (Figure A.B.3). In CD4⁺ T cells, ~350 genes were differentially expressed between CD69⁺ and CD69⁻ T cells in BM, and ~200 genes were differentially expressed in L; for both tissues more genes were differentially expressed in CD4⁺T cells than in CD8⁺ T cells (Figure A.B.4). The relationship between T cells from Bld,

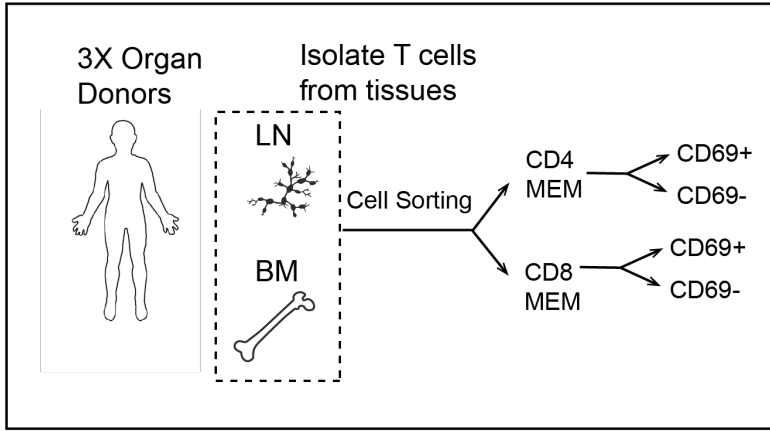
BM, Spl, LN and Lung was analyzed by MDS analysis using the list of genes differentially expressed between BM CD69- and CD69+ T cells. We found that CD69+ T cells from all sites clustered together, and CD69- T cells from all sites clustered together with blood (Figure A.B.4). Taken together, both CD4+ and CD8+, CD69+ and CD69- memory T cells display distinct transcriptional profiles in LN and BM. Memory T cells expressing CD69 were transcriptionally similar to TRM in spleen and lung, and memory T cells not expressing CD69 were transcriptionally similar to circulating memory T cells. We analyzed the abundance of T cells in BM biopsies by immunohistochemistry, and found 10-15% of total nucleated cells were T cells, by CD3 staining (Figure A.B.5). Further studies are needed to investigate whether CD69+ and CD69- T cells display differential localization patterns within the BM, or if CD69+ and CD69- T cells display differences in morphology or preferential interaction with specific cell types in the BM.

Figure A.B.1. Whole transcriptome profiling by RNA-seq of CD69⁺ and CD69⁻ memory CD8⁺ and CD4⁺ cells from BM and LN.

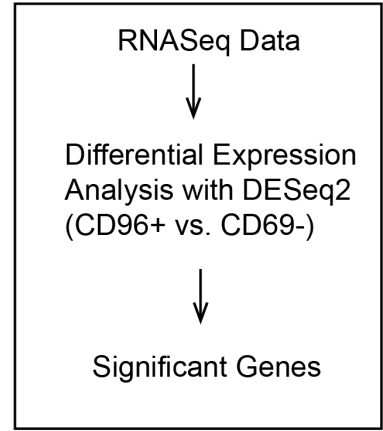
(A) Acquisition of purified T cells from BM and LN and subsequent analysis strategy using DESeq2. (B). Gating strategy (left to right): lymphocytes, singlets, memory T cells (CD3⁺CD45RO⁺), CD8⁺ or CD4⁺ and CD69^{+/-} from BM of a representative individual for RNASEq. (C) Stacked bar graph of the number of significantly up (grey) or (down) regulated genes by DESeq2 analysis (FDR<0.05 and |log₂FC| > 1). For differential expression analysis, CD69⁺ versus CD69⁻ memory T cells were compared to identify up and down regulated genes relative to CD69⁺ memory T cells. Results are for CD4⁺ and CD8⁺ T cells in LN and BM. Total genes, fold change, and significance are shown in Table A.B.2 below.

(A)

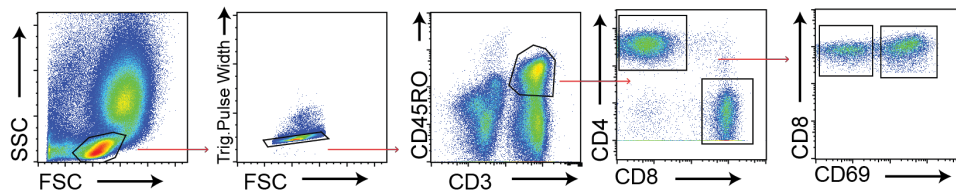
1. Data acquisition:



2. Analysis



(B)



(C)

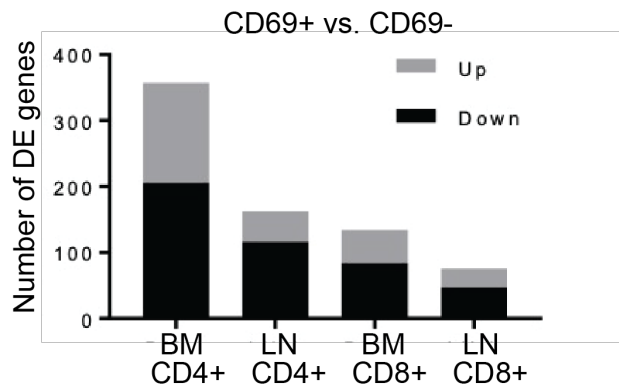


Table A.B.1. RNASeq sample information. Columns indicate tissue (BM or LN), CD69 expression (neg = negative, pos = positive), donor number, and lineage (CD4⁺ or CD8⁺) of samples.

Sample	Tissue	CD69	donor	CD4 or CD8
DM001	BM	neg	D280	CD4
DM002	BM	neg	D287	CD4
DM003	BM	neg	D229	CD4
DM004	LN	neg	D280	CD4
DM005	LN	neg	D229	CD4
DM006	LN	neg	D287	CD4
DM007	BM	neg	D280	CD8
DM008	BM	neg	D287	CD8
DM009	BM	neg	D229	CD8
DM010	LN	neg	D280	CD8
DM011	LN	neg	D229	CD8
DM012	LN	neg	D287	CD8
DM013	BM	pos	D280	CD4
DM014	BM	pos	D287	CD4
DM015	BM	pos	D229	CD4
DM016	LN	pos	D280	CD4
DM017	LN	pos	D229	CD4
DM018	LN	pos	D287	CD4
DM019	BM	pos	D280	CD8
DM020	BM	pos	D287	CD8
DM021	BM	pos	D229	CD8
DM022	LN	pos	D280	CD8
DM023	LN	pos	D229	CD8
DM024	LN	pos	D287	CD8

Table A.B.2. Differentially expressed genes in CD4⁺ and CD8⁺ TRM vs. TEM (CD69⁺ vs. CD69⁻) cells in BM and LN. Results from differential expression analysis from three donors for each group (samples specified in Table A.B.1 above).

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p <0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
GPR12	5.6	3.2	3.6	2.9	0.00	NA	NA	NA	6.7
GLIS1	5.6	1.8	0.8	2.8	0.00	NA	1.00	0.10	11.1
CDHR1	5.0	3.2	2.3	1.7	0.00	0.00	0.00	0.13	145.1
MEIS3	4.9	1.6	4.8	0.3	0.00	NA	NA	NA	4.5
LOC100129316	4.9	2.2	1.1	1.6	0.03	0.77	1.00	1.00	38.7
BCAM	4.5	0.0	2.0	1.4	0.04	1.00	NA	1.00	8.5
FLJ43390	4.5	0.0	4.0	2.3	0.02	NA	0.14	NA	4.6
PROKR2	4.1	1.5	4.6	1.9	NA	NA	0.00	NA	5.4
LOC100128909	4.0	1.4	1.1	1.3	0.00	1.00	0.28	0.85	53.6
PKDCC	3.9	3.5	2.3	2.1	0.00	0.10	0.26	NA	23.5
DAB2IP	3.7	0.3	1.6	2.3	0.00	1.00	1.00	0.02	199.8
IL2	3.7	3.1	2.2	2.2	0.00	0.00	0.49	0.03	34.5
LOC100507103	3.7	3.1	1.3	1.8	0.00	0.50	0.52	0.53	29.0
RAPGEF5	3.6	1.6	2.0	2.0	0.00	0.59	0.75	0.02	41.4
IL23R	3.1	1.5	1.1	2.4	0.00	0.61	1.00	0.14	159.8
KCNQ3	3.1	3.3	1.7	1.0	0.00	0.00	0.85	1.00	52.1
LOC285463	3.1	1.5	1.1	0.4	0.02	0.71	1.00	NA	7.9
CCDC165	3.0	1.0	1.7	1.1	0.00	1.00	0.72	1.00	12.6
SLC16A2	3.0	1.4	1.1	0.9	0.00	1.00	1.00	NA	8.3
AQP7P1	2.9	1.1	2.1	1.0	0.04	0.99	0.73	1.00	21.6
CD160	2.9	-1.0	3.2	2.3	0.01	NA	0.06	0.02	152.6
LOC100507390	2.9	0.2	0.8	4.1	0.03	NA	NA	NA	4.9
IL10	2.9	1.4	1.7	1.9	0.00	0.03	0.59	1.00	124.4
EPAS1	2.8	0.4	1.0	1.2	0.00	NA	0.89	1.00	213.7
GNB3	2.8	0.7	1.1	0.2	0.00	1.00	1.00	NA	11.3
XCL1	2.8	0.3	2.8	2.3	0.00	1.00	0.00	0.00	100.4
JAG2	2.7	1.0	2.5	2.0	0.00	0.73	0.00	0.03	57.2
WNK2	2.7	1.1	4.3	-0.6	NA	1.00	0.01	NA	6.0
GLT25D2	2.7	1.8	0.6	1.7	0.00	0.01	1.00	0.00	107.5
COL5A1	2.7	0.0	1.2	2.0	0.00	1.00	1.00	0.33	134.0
LOC440300	2.6	1.9	-0.3	0.3	0.01	0.22	1.00	1.00	17.7
FREM2	2.6	5.6	NA	3.2	NA	0.02	NA	NA	2.2
GLP1R	2.6	2.6	2.6	0.0	0.24	0.05	NA	NA	9.8
CRIM1	2.6	1.6	1.4	1.2	0.00	0.13	0.00	0.30	479.2
CCND1	2.6	1.9	2.6	1.5	0.01	0.02	0.00	0.29	42.5
IGSF11	2.5	3.6	1.4	2.0	0.03	0.00	NA	NA	16.0
WNT11	2.5	2.3	0.2	0.6	0.00	0.15	1.00	1.00	23.0
C2orf85	2.5	0.4	0.0	0.4	0.01	1.00	1.00	1.00	48.6
HTRA1	2.5	1.2	1.0	2.3	0.03	1.00	1.00	NA	11.7
MIR4772	2.5	2.2	0.7	1.6	0.01	0.56	1.00	NA	13.9
ADRB1	2.5	4.0	2.4	2.4	0.39	NA	0.03	0.20	13.8
FCRL6	2.5	0.4	-0.2	0.5	0.00	1.00	1.00	1.00	342.4
IL1RL1	2.4	1.2	1.0	1.4	0.00	0.43	1.00	NA	34.4
GFOD1	2.3	1.3	1.2	0.8	0.00	0.10	0.05	0.73	387.3
B3GAT1	2.3	1.1	-1.3	0.3	0.00	0.85	0.57	1.00	96.5
CXCR6	2.3	0.3	1.2	1.4	0.00	1.00	1.00	0.02	2023.4

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p <0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
ITGA1	2.3	2.3	2.4	1.7	0.00	0.00	0.00	0.09	308.0
SDK1	2.3	2.1	0.3	0.8	0.30	0.02	NA	NA	23.9
PDZD2	2.3	1.1	-1.3	1.6	0.01	1.00	1.00	NA	12.4
CCL4	2.3	1.7	1.1	1.2	0.00	0.01	0.25	1.00	1040.4
DTHD1	2.2	1.4	1.6	0.4	0.01	0.00	0.14	1.00	850.9
FER1L4	2.2	1.5	1.8	1.5	0.01	0.01	0.35	0.02	39.4
IL17RE	2.2	0.9	0.4	1.7	0.04	1.00	1.00	1.00	56.8
CPNE7	2.2	1.1	1.1	0.9	0.04	0.60	1.00	1.00	62.5
PTGDR	2.2	1.2	0.6	0.5	0.00	0.13	0.00	1.00	541.0
KCNA6	2.1	0.5	2.0	1.3	0.00	1.00	0.00	0.69	175.0
AATK	2.1	1.7	1.0	3.0	0.01	0.02	0.99	NA	55.8
TP53I1	2.1	1.4	0.6	0.7	0.00	0.39	1.00	1.00	236.6
RGS1	2.1	0.6	3.5	1.0	0.00	1.00	0.00	0.92	12373.5
LOC100130872	2.0	0.7	0.2	0.3	0.00	0.82	1.00	1.00	722.4
PLXDC1	2.0	1.0	1.5	0.7	0.00	0.00	0.00	0.89	415.4
MS4A1	2.0	0.8	1.8	0.2	0.00	0.85	0.00	1.00	128.6
LEPREL2	2.0	1.2	1.0	0.3	0.00	0.20	0.48	1.00	64.7
SHISA2	2.0	0.9	0.9	0.7	0.02	1.00	NA	1.00	71.1
LGR6	2.0	1.0	0.8	0.6	0.00	0.86	1.00	1.00	147.3
HIC1	2.0	0.7	1.8	1.4	0.01	1.00	0.00	1.00	366.0
TRPM2	2.0	0.4	1.2	1.6	0.00	1.00	0.24	0.43	198.0
LIF	1.9	2.0	2.3	0.8	0.25	0.02	0.08	1.00	29.2
P2RX5- TAX1BP3	1.9	1.2	0.7	0.9	0.01	0.42	1.00	1.00	34.0
CD101	1.9	0.0	1.4	1.1	0.00	1.00	0.09	0.64	118.8
ERRFI1	1.8	0.6	1.1	0.3	0.01	1.00	0.28	1.00	308.1
TGFBI	1.8	0.9	1.2	1.3	0.03	0.84	0.16	0.06	257.2
MYO1B	1.8	1.1	0.9	1.1	0.01	1.00	1.00	1.00	212.2
MCAM	1.8	0.3	1.0	0.7	0.02	1.00	1.00	1.00	254.1
MMRN1	1.8	1.4	0.2	0.1	0.03	0.22	1.00	1.00	118.6
ASB2	1.8	0.5	1.0	0.7	0.00	1.00	0.73	1.00	152.9
GNAO1	1.8	1.0	0.1	0.4	0.00	0.25	1.00	1.00	218.5
STARD9	1.7	0.6	1.3	1.0	0.00	0.47	0.00	0.01	679.6
ZCCHC12	1.7	2.2	-0.1	-0.3	0.50	0.01	NA	1.00	56.6
IL12RB2	1.7	0.1	0.8	1.3	0.00	1.00	1.00	0.00	241.9
IFNG	1.7	1.2	0.1	0.9	0.03	0.01	1.00	0.79	538.7
GFI1	1.7	1.3	0.7	1.1	0.00	0.00	0.00	0.00	775.4
SAMD4A	1.7	0.6	2.0	0.8	0.00	1.00	0.02	1.00	74.8
MIR3687	1.7	-0.7	1.2	0.3	0.00	1.00	0.02	1.00	6138.7
FAM110C	1.7	-0.9	1.5	-0.2	0.00	1.00	0.02	1.00	54.0
SPEG	1.7	1.0	1.3	1.4	0.05	0.48	1.00	0.24	110.1
CSF1	1.7	0.5	2.8	1.3	0.02	1.00	0.00	0.01	403.5
SOAT2	1.7	1.3	1.2	0.8	0.00	0.91	0.55	1.00	37.7
LOC100128342	1.6	0.4	0.6	1.5	0.03	1.00	1.00	1.00	93.6
IL18RAP	1.6	1.7	0.3	2.0	0.00	0.00	1.00	0.00	1114.8
CARD10	1.6	1.0	2.1	1.3	0.21	1.00	0.04	1.00	17.0
KCNK5	1.6	1.7	1.4	1.2	0.01	0.00	0.34	0.07	220.9
C8orf80	1.6	0.1	-0.1	0.4	0.00	1.00	1.00	1.00	185.1
F2R	1.6	0.8	0.3	0.5	0.00	0.78	1.00	1.00	840.2
YPEL1	1.6	0.4	0.2	0.5	0.01	1.00	1.00	1.00	422.6
FAM125B	1.6	1.1	0.2	0.7	0.00	0.23	1.00	1.00	585.4
ADAM23	1.6	2.4	0.2	0.2	0.04	0.31	1.00	1.00	483.3
GRID2IP	1.5	1.5	0.8	0.9	0.03	0.32	1.00	1.00	25.4
TOX2	1.5	0.9	2.7	0.4	0.34	0.41	0.00	1.00	56.8

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p <0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
AFF3	1.5	1.0	0.8	0.8	0.00	0.00	1.00	1.00	211.9
DUSP2	1.5	0.2	1.1	0.5	0.04	1.00	0.14	0.44	7510.6
ZNF618	1.5	0.3	0.5	0.9	0.04	1.00	1.00	1.00	36.7
IFNGR1	1.5	1.0	0.8	1.0	0.00	0.01	0.62	0.00	1653.8
SPRY1	1.5	1.6	1.0	1.5	0.03	0.01	0.65	0.00	271.5
SLC4A4	1.5	0.7	1.0	0.9	0.12	1.00	0.00	0.16	153.2
SLC41A2	1.5	1.7	0.8	1.1	0.05	0.02	0.65	0.45	61.8
DUSP6	1.5	1.3	1.0	1.2	0.01	0.00	0.00	0.27	863.5
PLXND1	1.5	1.0	0.2	1.1	0.06	0.02	1.00	0.00	1859.9
PTGFRN	1.5	0.3	1.5	0.6	0.00	1.00	1.00	1.00	33.9
SEMA7A	1.4	0.6	1.6	0.9	0.01	0.78	0.00	0.03	205.1
C9orf173	1.4	1.0	1.7	0.9	0.24	0.38	0.01	1.00	66.4
NRBP2	1.4	0.8	0.6	0.5	0.00	0.25	0.95	1.00	358.2
RYR1	1.4	0.8	1.3	0.7	0.00	0.94	0.10	1.00	98.6
SLC7A5	1.4	0.8	0.5	0.8	0.05	1.00	1.00	1.00	2519.9
LOC692247	1.4	0.5	1.0	0.9	0.00	1.00	0.06	0.53	105.2
COL6A1	1.4	0.9	0.4	0.7	0.00	0.22	1.00	1.00	98.8
PFKFB3	1.4	0.9	1.0	1.1	0.00	0.12	0.01	0.18	6098.7
RIMS3	1.4	0.4	1.2	0.8	0.02	1.00	0.16	1.00	130.9
PTGER4	1.4	0.3	0.8	0.7	0.00	0.37	0.00	0.03	7482.4
PHLDB1	1.4	0.4	1.5	0.4	0.03	1.00	0.01	1.00	143.9
SFMBT2	1.4	0.8	1.0	0.6	0.00	0.01	0.07	0.03	1433.9
ARHGAP18	1.3	0.6	1.3	0.8	0.00	0.85	0.00	0.96	398.9
NLRP6	1.3	1.2	0.6	0.7	0.00	0.00	0.07	0.08	566.4
ADAMTS17	1.3	0.3	2.2	1.0	0.00	0.94	0.01	0.00	336.7
UBE2Q2P1	1.3	0.7	0.3	0.5	0.01	0.57	1.00	1.00	121.9
TBX21	1.3	0.4	-0.3	0.5	0.00	1.00	1.00	0.29	1088.6
IGLL5	1.3	-0.5	1.1	2.9	0.92	1.00	1.00	0.01	71.2
GIPR	1.3	0.6	0.8	0.4	0.02	0.94	0.10	1.00	77.1
CD8A	1.3	0.1	0.0	0.3	0.05	1.00	1.00	1.00	4033.2
PDCD1	1.3	0.6	0.2	0.1	0.00	0.66	1.00	1.00	660.9
FLT4	1.3	1.3	0.2	1.0	0.27	0.01	1.00	0.45	528.3
VAV2	1.3	0.8	1.1	1.2	0.00	0.15	0.09	0.07	97.9
ID2	1.3	0.9	0.6	0.6	0.01	0.13	0.78	0.82	2959.4
CCR5	1.3	0.0	0.5	0.5	0.00	1.00	0.55	1.00	1351.9
BMF	1.3	0.2	1.0	0.5	0.00	1.00	0.02	1.00	195.3
CKLF	1.3	0.1	0.4	-0.2	0.02	1.00	1.00	1.00	33.5
PLXNB1	1.3	0.6	0.8	0.8	0.02	0.62	0.27	0.98	71.0
CD69	1.3	0.5	1.7	0.8	0.00	1.00	0.00	1.00	17091.7
STX1A	1.2	-0.3	0.2	0.4	0.00	1.00	1.00	1.00	75.0
KSR2	1.2	1.8	-1.6	-0.1	NA	0.02	NA	NA	11.5
SESN1	1.2	1.2	0.8	1.0	0.03	0.01	0.92	0.24	1722.0
COLQ	1.2	0.1	0.5	1.7	0.01	1.00	1.00	0.61	489.6
DUSP5	1.2	0.4	1.3	0.9	0.01	0.44	0.00	0.00	2630.0
EOMES	1.2	0.9	0.8	0.4	0.06	0.41	0.00	1.00	1113.7
SLC6A8	1.2	0.7	0.6	0.4	0.04	1.00	1.00	1.00	83.0
P2RX5	1.2	0.9	0.7	0.7	0.00	0.09	0.84	1.00	100.7
CDK5R1	1.2	1.1	1.2	0.6	0.01	1.00	0.04	1.00	651.8
NOXA1	1.2	0.2	0.9	0.3	0.01	1.00	0.30	1.00	93.2
CAPN5	1.2	0.2	0.2	0.3	0.03	1.00	1.00	1.00	40.3
SARDH	1.2	0.5	1.7	0.5	0.04	1.00	0.00	1.00	135.2
EGR3	1.2	0.3	1.2	0.4	0.03	1.00	0.00	1.00	223.3
NPTXR	1.2	1.1	0.8	0.6	0.00	0.10	0.88	1.00	193.3

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p <0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
PBX4	1.2	0.4	0.3	0.5	0.01	1.00	1.00	1.00	914.2
ZCCHC24	1.2	0.6	0.8	0.8	0.04	1.00	0.30	1.00	77.5
ATP10D	1.1	0.4	0.5	0.8	0.04	1.00	1.00	0.30	245.6
CH25H	1.1	2.2	0.1	0.6	0.24	0.00	1.00	1.00	453.9
IRGM	1.1	0.4	0.6	1.0	0.01	1.00	0.62	0.10	136.3
SGSM1	1.1	0.7	-0.4	0.5	0.03	0.53	1.00	1.00	81.7
LOC100506023	1.1	0.3	0.4	0.2	0.02	1.00	1.00	1.00	172.8
TBKBP1	1.1	0.7	0.1	0.5	0.02	0.15	1.00	1.00	310.5
TOX	1.1	0.8	0.7	0.4	0.03	0.15	0.17	1.00	481.3
ZNF80	1.1	0.4	1.8	-0.3	0.36	0.85	0.01	1.00	169.0
FAM18B2	1.1	0.5	0.3	0.1	0.02	0.57	1.00	1.00	350.9
HDAC9	1.1	0.5	1.8	1.3	0.16	1.00	0.01	0.47	162.1
AGR1	1.1	1.0	0.9	1.2	0.51	0.08	0.67	0.02	101.3
IL6ST	1.1	0.9	1.1	0.4	0.29	0.14	0.01	1.00	6786.1
SMOX	1.1	2.1	1.1	1.2	0.44	0.04	0.98	1.00	26.9
PDE4A	1.1	0.1	0.6	0.8	0.04	1.00	0.47	1.00	272.4
FGFRL1	1.1	0.5	0.8	0.6	0.03	1.00	0.01	1.00	329.5
PIK3AP1	1.1	0.1	0.8	1.5	0.15	1.00	0.53	0.00	878.0
MCC	1.0	1.0	0.6	0.6	0.00	0.07	1.00	1.00	460.8
C17orf107	1.0	0.5	1.5	0.7	0.02	1.00	0.01	1.00	1199.8
SNTB2	1.0	0.5	0.4	0.5	0.00	0.98	0.79	1.00	1341.2
FAM46C	1.0	-0.2	1.0	0.2	0.00	1.00	0.02	1.00	4817.6
GBGT1	1.0	1.6	1.0	0.7	0.06	0.00	0.76	1.00	70.1
SMPD3	1.0	0.4	2.3	1.4	0.43	1.00	0.00	1.00	54.2
PTPRN2	1.0	0.5	1.5	0.5	0.32	1.00	0.04	1.00	67.5
AOAH	1.0	1.1	0.5	0.3	0.50	0.04	1.00	1.00	1160.6
MIR155HG	1.0	1.5	1.0	0.6	0.34	0.02	0.73	1.00	89.0
MMP25	1.0	0.2	0.8	0.6	0.00	1.00	0.61	1.00	87.4
CRTAM	1.0	1.1	0.8	0.8	0.52	0.42	0.04	1.00	1039.1
MIR663	1.0	-0.6	0.6	0.1	0.04	1.00	0.12	1.00	9984.5
RAB15	1.0	0.3	1.6	0.6	0.38	1.00	0.00	1.00	71.5
CCL3	1.0	0.3	2.4	1.9	0.70	1.00	0.00	0.99	387.8
KLRB1	0.9	0.7	0.1	1.8	0.03	0.87	1.00	1.00	3315.9
LOC643733	0.9	1.0	0.6	0.5	0.15	0.01	1.00	1.00	139.0
PLCB1	0.9	1.1	0.2	0.8	0.01	0.01	1.00	0.71	379.9
CHN2	0.9	0.8	1.1	0.7	0.28	0.57	0.01	0.78	154.6
GJC2	0.9	0.3	2.0	0.8	0.67	1.00	0.02	1.00	84.6
SNTA1	0.8	0.5	0.5	1.2	0.31	0.73	0.89	0.01	138.0
SIPA1L2	0.7	1.2	-0.2	0.2	0.57	0.00	1.00	1.00	307.2
LOC100507209	0.7	1.4	0.2	-0.3	0.74	0.05	1.00	1.00	61.6
CCL20	0.6	2.0	1.3	3.3	0.88	0.00	1.00	0.01	60.3
PLCG2	0.6	0.0	1.3	0.3	0.43	1.00	0.00	1.00	393.5
PHLDA2	0.6	1.5	0.6	1.1	0.84	0.02	1.00	1.00	25.5
ITGAE	0.6	0.4	1.1	1.3	0.38	1.00	0.02	0.47	190.3
KLRD1	0.6	0.9	-0.6	1.1	0.93	1.00	1.00	0.00	842.5
MMP9	0.5	-0.5	0.2	2.5	0.97	1.00	1.00	0.01	307.2
FRMPD3	0.5	-1.2	-1.6	-1.4	0.95	NA	0.02	0.10	34.3
MAP7	0.5	1.7	0.2	-0.2	0.90	0.04	1.00	1.00	38.3
C5orf62	0.4	0.5	0.0	1.3	0.80	1.00	1.00	0.03	91.6
CACNA2D2	0.3	0.8	-1.4	0.1	0.89	0.22	0.00	1.00	185.1
FCGR3A	0.3	-1.7	-2.1	-0.9	0.98	0.84	0.02	1.00	325.8
ACVR2A	0.3	0.3	-1.5	-0.5	0.80	1.00	0.03	1.00	368.5
PLEKHH2	0.2	1.3	0.8	0.1	0.97	0.00	1.00	1.00	69.2

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p <0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
SYNE1	0.2	0.2	-1.1	-0.9	0.92	1.00	0.00	0.00	6891.7
PCDH1	0.1	-0.1	-3.9	0.9	NA	1.00	0.02	1.00	14.2
CHI3L1	0.0	-1.8	0.1	8.5	1.00	0.87	1.00	0.00	38.2
TGFBR3	0.0	-0.3	-1.2	-0.5	0.99	0.93	0.04	1.00	1224.4
MED12L	0.0	1.9	-0.3	0.5	1.00	0.01	1.00	1.00	41.7
ANK2	0.0	2.3	0.7	1.3	NA	0.01	NA	NA	12.4
LILRB1	-0.1	-1.4	-2.1	-0.1	0.99	1.00	0.02	1.00	79.8
SPRY2	-0.2	-1.1	2.2	1.7	0.99	NA	0.00	0.23	49.7
NCF2	-0.2	-2.8	0.0	-1.5	0.97	0.00	1.00	0.95	166.4
ABCA2	-0.2	-1.0	-0.3	-0.3	0.90	0.03	1.00	1.00	1792.9
ADAM8	-0.3	-1.2	-0.7	-0.2	0.88	0.00	0.42	1.00	1222.2
APOBEC3H	-0.3	-1.3	-1.3	-0.7	0.96	0.66	0.02	1.00	50.2
ITGAM	-0.3	-2.5	-1.5	-1.5	0.88	0.00	0.14	0.05	427.3
ATP10A	-0.3	-0.1	-1.2	0.0	0.53	1.00	0.01	1.00	602.5
TSPAN2	-0.3	-2.0	-1.6	-1.7	0.97	0.00	1.00	0.05	71.8
PYROXD2	-0.3	-0.1	-1.3	-0.7	0.85	1.00	0.00	1.00	90.8
HLA-DOA	-0.3	-1.5	0.3	-0.4	0.95	0.01	1.00	1.00	105.8
TFEB	-0.4	-0.6	-1.4	-0.6	0.75	0.97	0.00	1.00	156.2
SEMA4C	-0.4	-0.5	-1.0	-0.7	0.53	1.00	0.04	1.00	518.2
CEP78	-0.5	-0.4	-1.0	-0.6	0.21	0.67	0.00	0.14	347.1
CYB561	-0.5	-0.3	-1.1	-0.3	0.33	1.00	0.01	1.00	329.6
TMCC3	-0.6	-0.9	-2.8	-0.3	0.85	1.00	0.01	1.00	87.2
CLIC3	-0.6	-0.8	-1.2	-0.7	0.83	1.00	0.02	1.00	35.7
S100A4	-0.6	-1.1	-0.7	-0.4	0.32	0.01	0.84	1.00	1710.6
PHLDB2	-0.6	-0.1	-1.1	0.1	0.23	1.00	0.01	1.00	215.1
TNFAIP8L2	-0.6	-1.1	-0.8	-1.1	0.19	0.85	0.01	1.00	130.0
CASP10	-0.6	-1.0	-0.9	-0.7	0.35	0.02	0.01	0.49	367.4
ALOX5	-0.7	-2.5	-0.1	-1.4	0.46	0.00	1.00	0.59	147.2
GPR56	-0.7	-0.2	-2.6	-0.4	0.74	1.00	0.00	1.00	378.4
MYOIG	-0.7	-1.2	-0.7	-0.7	0.10	0.00	0.52	0.08	2302.2
MTSS1	-0.7	-1.0	-1.6	-0.9	0.59	0.01	0.03	0.09	259.0
CFP	-0.7	-1.7	0.0	-0.1	0.67	0.00	1.00	1.00	103.0
PLEKHG3	-0.7	-2.2	-2.0	-2.0	0.47	0.02	0.00	1.00	776.4
LRRC8C	-0.7	-0.5	-1.2	-0.8	0.19	0.62	0.02	0.01	2036.7
LGALS1	-0.8	-1.8	-1.3	-0.5	0.62	0.00	1.00	1.00	238.4
ANXA2P2	-0.8	-1.4	-0.4	-0.4	0.15	0.00	1.00	1.00	628.9
SSBP3	-0.8	0.1	-1.7	-0.5	0.11	1.00	0.02	1.00	121.8
TSHZ3	-0.8	-0.5	-2.5	-0.6	0.51	1.00	0.01	NA	32.7
HBEGF	-0.8	-1.6	0.2	-0.4	0.79	0.01	1.00	1.00	87.5
GRASP	-0.8	-1.1	-0.6	-0.4	0.01	0.62	1.00	1.00	656.3
ITGA5	-0.9	-0.6	-1.6	-0.6	0.00	0.68	0.00	0.82	888.2
CCDC65	-0.9	-1.6	-1.0	-0.8	0.01	0.02	0.09	1.00	127.1
KCNH2	-0.9	-1.5	0.5	-0.7	0.50	0.03	1.00	NA	32.2
DPYSL2	-0.9	-1.1	-1.0	-0.4	0.02	0.00	0.34	1.00	318.6
SAMD3	-0.9	-2.1	-0.7	-0.8	0.36	0.00	0.15	0.97	879.3
INF2	-0.9	-1.0	-1.0	-1.3	0.00	0.01	0.11	0.01	416.8
C14orf49	-0.9	-1.2	-0.6	-0.8	0.00	0.00	1.00	0.08	489.5
GABBR1	-0.9	-1.3	-0.4	-0.7	0.00	0.01	1.00	0.38	231.8
ANLN	-1.0	-2.1	-0.4	-1.9	0.59	0.01	1.00	0.32	72.8
CDCA7	-1.0	-1.6	-0.6	-1.6	0.27	0.09	1.00	0.02	69.9
HKDC1	-1.0	-2.4	-0.7	-1.2	0.07	0.01	0.72	1.00	101.2
HELLS	-1.0	-0.5	-0.2	-1.1	0.01	0.87	1.00	0.00	203.1
SPTBN1	-1.0	-0.8	-0.5	-0.6	0.00	0.00	0.84	0.49	4169.4

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p < 0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
CLSPN	-1.0	-1.2	-0.3	-1.4	0.18	0.02	1.00	0.09	84.6
S1PR5	-1.0	-0.1	-1.2	0.3	0.54	1.00	0.00	1.00	371.5
SLCO3A1	-1.0	-1.4	-1.3	-1.3	0.00	0.00	0.00	0.00	251.1
PRR11	-1.0	-1.8	-0.3	-1.1	0.10	0.00	1.00	0.63	76.7
NDC80	-1.0	-1.2	-0.1	-1.2	0.04	0.02	1.00	0.07	132.1
ABLIM1	-1.0	-0.5	-0.7	-0.5	0.00	0.62	0.32	1.00	2165.2
IL2RA	-1.0	-0.2	-0.2	1.2	0.02	1.00	1.00	0.01	725.7
SELPLG	-1.0	-1.3	-1.3	-1.0	0.00	0.00	0.00	0.00	2226.2
IGSF9B	-1.0	-1.2	0.2	-1.1	0.02	0.38	1.00	0.59	217.1
ITGB7	-1.1	-1.5	-0.8	-0.6	0.21	0.00	0.93	0.87	1224.5
STMN1	-1.1	-0.2	-0.4	-0.4	0.02	1.00	1.00	1.00	405.6
SYT11	-1.1	-0.7	-0.9	-0.7	0.00	0.41	0.32	0.08	482.1
AGPAT4-IT1	-1.1	-0.4	-2.8	-0.6	0.49	1.00	0.00	1.00	70.8
LINC00341	-1.1	-1.2	-1.0	-1.1	0.02	0.02	0.80	0.22	140.7
CCR4	-1.1	-0.3	0.0	-1.2	0.00	1.00	1.00	0.06	1618.0
NUSAP1	-1.1	-1.0	-0.4	-0.8	0.01	0.07	1.00	0.78	185.0
LOC728554	-1.1	-1.1	-0.4	-1.0	0.05	0.02	1.00	0.79	95.1
MYO6	-1.1	-0.3	-1.2	-0.2	0.21	1.00	0.02	1.00	142.8
BUB1	-1.1	-1.1	-0.3	-0.7	0.11	0.04	1.00	1.00	105.9
PITPNM2	-1.1	-1.0	0.2	-0.2	0.03	0.00	1.00	1.00	246.0
MAST4	-1.1	-0.4	-1.0	-0.7	0.00	1.00	0.19	1.00	486.8
CIT	-1.1	-0.7	0.1	-0.6	0.02	0.54	1.00	1.00	116.9
C11orf21	-1.1	-1.3	-0.9	-0.5	0.00	0.12	0.01	0.74	420.1
SCARNA16	-1.1	0.3	-0.5	0.8	0.04	1.00	1.00	1.00	52.8
CORO1B	-1.1	-1.1	-0.3	-0.4	0.00	0.16	1.00	1.00	136.2
SNX29	-1.1	-0.5	-0.5	-0.4	0.03	0.72	0.76	0.79	608.2
POLQ	-1.1	-1.4	-0.5	-1.6	0.17	0.05	1.00	0.07	67.2
CBLL1	-1.2	-0.3	-0.4	-0.2	0.00	1.00	0.91	1.00	484.0
CD52	-1.2	-1.1	-1.3	-1.0	0.00	0.01	0.00	0.11	2020.8
TIMP1	-1.2	-1.7	-1.2	-0.7	0.01	0.00	0.02	1.00	164.9
PDE3B	-1.2	-0.4	-1.3	-0.3	0.00	0.63	0.06	1.00	4253.6
TMEM169	-1.2	-1.3	-0.7	-1.8	0.04	0.64	1.00	NA	29.8
HIST1H3F	-1.2	-1.0	-0.2	-0.6	0.01	0.38	1.00	1.00	318.8
HIST1H2BH	-1.2	-0.7	-0.3	-0.4	0.00	0.70	1.00	1.00	482.5
SCARNA6	-1.2	0.5	-0.6	0.7	0.02	1.00	1.00	1.00	1668.0
TIAM1	-1.2	0.4	-1.5	-0.7	0.00	1.00	0.20	0.78	1110.3
AQP3	-1.2	-0.9	-0.1	-0.2	0.00	0.04	1.00	1.00	758.3
ARHGEF18	-1.2	-1.0	-0.7	-1.0	0.00	0.00	0.10	0.00	2538.2
MB21D2	-1.2	-0.7	-0.7	-1.3	0.00	0.69	1.00	1.00	84.0
BAIAP3	-1.2	-0.7	-1.0	-0.5	0.01	1.00	0.80	1.00	270.4
ADD3	-1.2	-0.8	-0.7	-0.9	0.00	0.00	0.01	0.00	3352.6
HIST1H4F	-1.2	-0.5	-0.6	-0.1	0.00	0.60	0.18	1.00	576.0
CENPE	-1.2	-1.1	-0.3	-1.0	0.00	0.00	1.00	0.35	181.5
BIRC3	-1.2	-0.6	-0.1	-0.2	0.00	0.17	1.00	1.00	6396.7
RASA3	-1.2	-0.8	-1.1	-0.8	0.00	0.06	0.00	0.00	2857.6
FAM129B	-1.3	-1.2	-0.9	-0.3	0.06	0.01	0.83	1.00	133.1
CXCR7	-1.3	-0.8	-1.5	-0.7	0.01	0.98	0.00	1.00	47.1
HPGD	-1.3	-1.1	0.0	-0.1	0.03	0.02	1.00	1.00	351.7
S1PR4	-1.3	-1.4	-1.0	-1.2	0.00	0.00	0.00	0.08	825.5
USP46	-1.3	-1.4	-0.7	-1.5	0.00	0.00	0.13	0.05	209.4
KIF11	-1.3	-1.5	-0.4	-1.4	0.02	0.00	1.00	0.01	164.3
TPX2	-1.3	-2.0	-0.6	-1.5	0.03	0.00	1.00	0.43	118.5
SNED1	-1.3	-0.9	-0.8	-1.0	0.01	1.00	1.00	0.78	271.3

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p <0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
NCAPG	-1.3	-1.3	-0.3	-1.3	0.05	0.01	1.00	0.28	78.1
RMRP	-1.3	0.1	-0.9	0.4	0.03	1.00	0.14	1.00	76059.3
KIF2C	-1.3	-1.7	-0.7	-1.5	0.03	0.01	1.00	0.89	61.3
LEF1	-1.3	-0.1	-2.0	-0.8	0.00	1.00	0.00	0.22	1874.7
NT5E	-1.3	-0.7	-0.6	-0.4	0.03	1.00	1.00	1.00	91.5
VCL	-1.3	-1.1	-1.5	-0.9	0.00	0.04	0.04	0.30	645.6
FLNA	-1.3	-1.2	-1.2	-0.8	0.00	0.07	0.06	0.85	21453.9
PLK1	-1.3	-1.5	-0.5	-1.0	0.01	0.19	1.00	1.00	52.0
CSGALNACT1	-1.3	-0.9	-0.9	-1.8	0.04	0.00	1.00	0.05	439.6
EEPDI	-1.3	-1.1	-1.1	-1.0	0.00	0.04	0.31	0.57	108.8
CENPF	-1.3	-0.9	-0.5	-1.5	0.03	0.05	1.00	0.01	320.3
ESPL1	-1.3	-1.2	-0.7	-1.5	0.02	0.55	1.00	0.74	64.2
KIF23	-1.3	-1.6	-0.7	-1.0	0.19	0.02	1.00	1.00	53.1
LOC100506609	-1.3	-0.8	-1.5	-1.0	0.20	1.00	0.03	1.00	38.4
ZNF365	-1.3	-0.5	-4.0	-1.8	0.34	0.91	0.02	0.46	62.8
C12orf75	-1.4	-1.8	-1.2	-1.0	0.00	0.00	0.08	0.09	186.9
JAM3	-1.4	-0.2	-0.7	-0.1	0.01	1.00	0.73	1.00	105.1
PGAP1	-1.4	0.2	-0.4	0.2	0.00	1.00	1.00	NA	1008.3
STK38	-1.4	-1.1	-1.2	-1.2	0.00	0.00	0.00	0.00	2196.2
EMP3	-1.4	-1.5	-1.1	-1.1	0.00	0.00	0.02	0.00	454.3
DLEC1	-1.4	-0.8	-1.0	-1.0	0.00	0.27	0.03	0.08	121.8
CCNA2	-1.4	-1.5	-0.8	-1.7	0.05	0.01	1.00	0.02	95.9
FMN1	-1.4	-0.9	-2.3	-1.8	0.16	1.00	0.00	0.17	445.1
FOXP3	-1.4	-1.1	0.0	0.0	0.00	0.63	NA	1.00	129.8
HIST1H3B	-1.4	-1.0	-0.5	-0.9	0.11	0.02	1.00	0.37	490.2
RRM2	-1.4	-2.2	-0.5	-2.1	0.14	0.00	1.00	0.11	114.3
CCR8	-1.4	-1.7	-3.2	-1.4	0.04	0.16	NA	1.00	61.2
GCNT4	-1.4	0.5	-0.4	-0.8	0.05	1.00	1.00	0.45	516.1
FAM63A	-1.4	-1.3	-0.8	-1.1	0.06	0.00	0.95	0.92	113.3
HJURP	-1.4	-1.7	-0.9	-1.9	0.04	0.18	1.00	0.91	39.3
FGFBP2	-1.4	-2.0	-4.9	-3.1	0.35	NA	0.00	0.02	156.5
NOSIP	-1.4	-1.1	-0.8	-0.9	0.00	0.00	0.06	0.03	613.7
KLF3	-1.4	-1.2	-1.4	-1.5	0.00	0.14	0.00	0.01	1367.5
MAL	-1.5	-0.2	-2.0	-0.7	0.00	1.00	0.13	0.98	347.5
KIFC1	-1.5	-1.3	-0.5	-0.6	0.00	0.21	1.00	1.00	69.8
TK1	-1.5	-1.4	-0.2	-1.5	0.52	0.05	1.00	1.00	39.4
CDCA8	-1.5	-1.3	-0.7	-1.5	0.04	0.49	1.00	0.98	40.5
CASC5	-1.5	-1.5	-0.2	-1.6	0.00	0.00	1.00	0.02	135.8
KIAA0101	-1.5	-3.1	-0.2	-2.0	0.26	0.00	1.00	0.33	34.5
MELK	-1.5	-2.0	-0.8	-1.5	0.03	0.01	1.00	1.00	29.6
ASPM	-1.5	-1.5	-0.6	-1.6	0.00	0.01	0.68	0.11	236.6
SH3BP5	-1.5	-2.1	-1.4	-1.5	0.00	0.00	0.00	0.00	627.6
SNORD89	-1.5	-0.1	-1.1	0.1	0.01	1.00	0.26	1.00	235.4
AGPAT4	-1.5	-0.5	-2.6	-0.7	0.00	1.00	0.00	1.00	350.1
CKAP2L	-1.5	-1.9	-1.0	-1.3	0.03	0.28	0.73	NA	25.4
FAM164A	-1.5	-1.0	-0.9	-1.1	0.05	0.23	1.00	1.00	54.6
TJP2	-1.6	-0.6	-1.4	-0.9	0.00	0.56	0.61	1.00	91.2
SNORD15B	-1.6	0.7	-0.9	0.9	0.02	1.00	1.00	1.00	139.3
PROCR	-1.6	-1.8	-2.0	-0.9	0.41	NA	0.00	NA	17.5
SNORD13	-1.6	0.5	-0.8	1.2	0.03	1.00	1.00	0.81	142.8
RBBP8	-1.6	-0.2	-0.4	0.5	0.02	1.00	1.00	1.00	128.6
KIF15	-1.6	-1.6	-0.6	-2.3	0.09	0.04	1.00	0.02	61.5
C1orf21	-1.6	-0.8	-2.5	-1.4	0.30	1.00	0.00	0.00	153.1

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p < 0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
ITGB1	-1.6	-1.2	-1.9	-1.1	0.00	0.00	0.01	0.00	3004.1
ODZ1	-1.6	-0.7	-2.2	-1.7	0.11	0.98	0.01	0.08	133.2
MYBL2	-1.6	-1.5	-0.4	-1.9	0.32	0.05	1.00	0.63	58.5
LOC100506029	-1.6	-0.8	-1.2	-1.4	0.02	0.76	0.30	0.35	34.4
C16orf45	-1.6	0.7	-2.3	-0.3	0.50	1.00	0.00	1.00	30.9
KIF14	-1.6	-1.7	-0.8	-1.9	0.08	0.05	1.00	0.09	75.2
STIL	-1.7	-0.6	-0.4	-0.7	0.03	0.82	1.00	1.00	65.3
TOP2A	-1.7	-1.5	-0.6	-1.8	0.02	0.00	1.00	0.00	350.0
LOC100506051	-1.7	-0.8	-1.2	-1.2	0.00	0.75	0.18	0.45	42.5
MSX2P1	-1.7	-2.2	-2.0	-1.8	0.00	0.18	0.00	1.00	35.2
KLF2	-1.7	-1.9	-1.1	-1.2	0.00	0.00	0.01	0.00	4440.4
NELL2	-1.7	-0.5	-0.7	-0.2	0.03	0.73	1.00	1.00	1084.4
SNORA48	-1.7	0.6	-0.7	1.2	0.01	0.72	1.00	0.47	442.0
TTC16	-1.7	-2.5	-1.0	-0.8	0.02	0.00	0.12	0.09	186.0
DLGAP5	-1.7	-2.4	-0.7	-1.8	0.13	0.00	1.00	0.46	66.8
ARHGEF10	-1.7	-0.6	-0.1	-0.1	0.00	0.54	1.00	1.00	58.8
PRSS23	-1.7	-0.2	-5.0	0.4	0.31	1.00	0.01	1.00	69.8
CDCA2	-1.7	-1.3	-0.6	-1.5	0.04	0.26	1.00	0.92	40.6
AURKB	-1.7	-1.3	-0.7	-1.9	0.00	0.36	1.00	1.00	31.6
HIST1H3C	-1.7	-0.9	-0.5	-0.8	0.00	0.49	1.00	0.89	204.6
NHSL2	-1.7	-2.4	-2.7	-1.6	0.00	0.00	0.00	0.08	116.3
CDCA5	-1.7	-1.6	-0.5	-1.6	0.02	0.22	1.00	0.57	38.9
LOC729041	-1.7	-2.4	-3.0	-1.1	0.03	0.05	0.00	1.00	26.4
DSEL	-1.7	-0.7	-2.0	-0.8	0.00	1.00	0.00	1.00	168.9
RNU4-2	-1.7	1.0	-1.3	1.1	0.00	1.00	0.57	1.00	807.8
ALS2CL	-1.8	-1.2	-0.3	-0.8	0.00	0.01	1.00	1.00	162.2
EEF1DP3	-1.8	-2.5	-2.8	-5.3	0.04	0.10	0.29	NA	15.8
FAM65B	-1.8	-1.6	-1.6	-1.5	0.00	0.00	0.00	0.00	4509.4
TRIM2	-1.8	-0.5	-1.7	-1.0	0.00	0.83	NA	1.00	59.7
BIRC5	-1.8	-2.4	-0.7	-2.5	0.03	0.00	1.00	0.16	48.9
CCNB2	-1.8	-1.9	-1.0	-2.0	0.06	0.03	1.00	0.38	45.5
SVIL	-1.8	-1.9	-1.7	-2.2	0.00	0.00	0.00	0.00	694.2
MFGE8	-1.8	-1.4	-0.8	-0.5	0.09	0.01	1.00	1.00	106.0
KLF7	-1.8	-0.8	-1.6	-1.2	0.00	0.22	0.09	0.42	138.4
D4S234E	-1.8	-1.0	-2.2	-0.7	0.08	0.37	0.00	1.00	190.0
GYLTL1B	-1.8	-2.4	-0.5	-1.6	0.05	0.01	1.00	0.51	23.1
KLF8	-1.8	-1.0	-1.5	-0.3	0.00	0.62	0.35	1.00	40.2
KIF19	-1.8	-2.3	-4.4	-1.9	0.12	0.82	0.00	NA	23.2
CX3CR1	-1.8	-1.5	-4.3	-2.9	0.08	0.38	0.00	0.02	269.7
GLB1L3	-1.8	-0.2	-3.5	-0.5	0.04	1.00	0.08	1.00	30.3
MKI67	-1.8	-2.5	-0.8	-2.5	0.00	0.00	1.00	0.01	486.2
HIST1H3G	-1.8	-2.4	-0.7	-2.2	0.01	0.00	1.00	0.02	189.7
CDC25B	-1.9	-1.4	-1.3	-1.0	0.00	0.00	0.00	0.00	1211.8
SNORD17	-1.9	0.8	-1.6	1.0	0.00	1.00	0.01	1.00	4213.8
LOC641518	-1.9	-0.6	-2.3	-1.6	0.03	1.00	0.36	0.52	30.2
SNORD94	-1.9	0.9	-1.3	1.5	0.02	1.00	0.56	1.00	86.3
RNU4-1	-1.9	0.8	-1.1	0.5	0.00	1.00	0.91	1.00	146.4
GPR146	-1.9	-2.0	-0.9	0.0	0.02	0.02	1.00	1.00	19.1
GPR15	-1.9	-1.0	-1.0	-0.7	0.01	1.00	1.00	1.00	331.1
ELOVL4	-1.9	-1.5	-1.0	-0.5	0.00	0.12	0.82	1.00	90.1
FAM111B	-1.9	-1.2	-1.0	-2.2	0.01	0.30	1.00	0.21	23.3
VSIG1	-1.9	-1.7	-1.3	-1.4	0.00	0.00	0.44	0.01	135.8
UBE2C	-1.9	-0.7	-0.5	-0.5	0.03	1.00	1.00	NA	33.9

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p <0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
PPP4R4	-1.9	-0.7	-0.8	-0.1	0.04	1.00	1.00	1.00	20.8
FOXM1	-2.0	-1.7	-0.3	-1.8	0.00	0.06	1.00	0.93	42.7
DNAI2	-2.0	-2.6	-3.2	-1.8	0.02	0.02	0.04	0.79	16.9
TTYH2	-2.0	-1.6	-2.1	-1.8	0.00	0.00	0.01	0.03	141.3
CRYBG3	-2.0	-1.6	-2.7	-1.7	0.00	0.00	0.00	0.00	273.8
SSTR3	-2.0	-0.4	-0.4	-0.5	0.01	1.00	NA	NA	20.0
PACSN1	-2.0	-0.5	-2.4	-1.7	0.02	1.00	0.06	0.01	48.5
SNORD10	-2.0	0.7	-1.3	1.2	0.00	1.00	0.56	0.82	495.7
MAP1A	-2.0	0.1	-0.2	0.2	0.04	1.00	1.00	1.00	97.7
VIPR1	-2.0	-1.2	-1.8	-1.3	0.01	0.53	1.00	0.73	292.9
SELL	-2.0	-0.7	-2.0	-1.8	0.18	0.76	0.00	0.00	2583.7
SNORD33	-2.1	-0.2	-1.4	0.1	0.04	1.00	0.64	1.00	39.2
SNORA52	-2.1	0.5	-1.2	0.8	0.02	1.00	0.80	1.00	50.9
NTN4	-2.1	-1.4	-1.4	-0.4	0.01	0.66	0.26	1.00	57.9
SBK1	-2.1	-1.8	-1.7	-1.4	0.00	0.00	0.00	0.00	140.4
SULT1B1	-2.1	-2.3	-0.6	-1.8	0.01	0.00	1.00	NA	29.6
MIR3609	-2.1	0.3	-0.6	1.1	0.00	1.00	1.00	1.00	1296.4
CR1	-2.1	-1.1	-1.9	-0.7	0.00	0.09	0.04	1.00	291.3
GAL3ST4	-2.1	-0.7	-0.3	-0.5	0.02	1.00	NA	NA	16.6
BCL7A	-2.2	-0.4	-1.8	-1.2	0.02	1.00	0.36	1.00	24.9
SNORA12	-2.2	0.6	-1.8	0.5	0.00	1.00	0.30	1.00	179.9
PKIA	-2.2	-1.0	-1.7	-1.3	0.00	0.00	0.01	0.02	144.2
RNU2-2	-2.2	1.2	-1.5	1.7	0.00	1.00	1.00	1.00	9485.0
PLXNA4	-2.2	-1.5	0.5	-1.2	0.01	0.50	1.00	1.00	34.1
NGFRAP1	-2.2	-1.0	-1.0	-1.1	0.00	0.05	1.00	0.58	74.0
GNLY	-2.2	-1.9	-3.8	-1.1	0.00	0.07	0.01	1.00	409.3
HPCAL4	-2.3	-1.3	-3.4	-1.7	0.00	0.03	0.11	0.01	141.2
SNORA54	-2.3	1.1	-1.0	2.3	0.00	1.00	1.00	1.00	703.6
EDA	-2.3	-0.7	-1.7	-0.9	0.00	1.00	0.64	1.00	30.7
LOC100652903	-2.3	-0.5	-2.0	-1.0	0.00	0.82	0.13	0.93	56.9
SNORA7B	-2.3	0.6	-1.6	1.3	0.01	1.00	0.44	1.00	35.6
SNORA73A	-2.3	0.3	-1.4	0.4	0.00	1.00	0.26	1.00	2320.3
RNU105A	-2.3	0.4	-1.5	0.5	0.00	1.00	0.08	1.00	2075.9
ARHGEF11	-2.4	-2.2	-1.1	-2.0	0.00	0.00	0.17	0.00	158.8
FHIT	-2.4	-1.0	0.0	-1.1	0.00	0.43	1.00	1.00	39.5
SNORA57	-2.4	0.7	-1.1	1.2	0.00	1.00	1.00	1.00	124.1
CCR7	-2.4	-0.3	-1.8	-0.6	0.00	1.00	0.03	1.00	2207.2
CRIP2	-2.4	-1.9	-0.1	-0.3	0.10	0.00	1.00	1.00	33.3
E2F2	-2.4	-3.1	-0.6	-3.5	0.00	0.00	1.00	0.00	120.0
SNORA49	-2.4	0.9	-1.3	1.4	0.00	1.00	0.61	1.00	142.7
SOX13	-2.5	-2.3	-2.7	-1.6	0.00	0.25	0.00	0.41	89.4
KRT72	-2.5	-2.8	-3.0	-2.0	0.00	0.00	0.00	0.14	36.7
TSHZ2	-2.5	-0.5	-1.1	-0.5	0.00	0.05	NA	1.00	290.0
SNORA63	-2.5	0.8	-1.5	1.1	0.00	1.00	1.00	1.00	616.9
CBR3	-2.5	-1.3	-1.7	-1.2	0.02	0.47	1.00	1.00	19.5
LOC100505551	-2.5	-1.2	-0.2	-1.0	0.04	0.66	1.00	1.00	13.7
SNORA22	-2.5	0.5	-0.7	1.0	0.04	1.00	1.00	1.00	17.0
C6orf105	-2.6	-0.2	-0.6	-0.3	0.02	1.00	1.00	1.00	118.8
AK5	-2.6	-2.2	-3.6	-2.6	0.06	0.02	0.30	0.24	48.6
LAMA2	-2.6	-1.1	-2.0	-1.1	0.01	0.98	0.02	1.00	39.4
SNORA34	-2.6	0.7	-1.5	1.0	0.02	1.00	0.77	1.00	65.7
KCTD15	-2.6	-2.8	-2.4	-1.6	0.00	0.00	0.10	1.00	46.5
EGLN3	-2.7	-1.2	-2.3	-1.9	0.00	0.66	0.83	0.72	17.2

Gene	Log2FC CD69(+) vs. CD69(-)				FDR Adjusted p-value, if p <0.05 (grey)				Avg. Base Mean
	BM CD4	LN CD4	BM CD8	LN CD8	BM CD4	LN CD4	BM CD8	LN CD8	
10-Sep	-2.7	-1.0	-1.1	-0.1	0.04	0.74	NA	1.00	16.6
PTGDS	-2.7	-3.7	-3.7	-2.4	0.00	0.00	0.00	0.06	61.2
RAP1GAP2	-2.7	-2.4	-3.7	-3.7	0.00	0.00	0.01	0.00	476.6
SPSB1	-2.7	0.3	-1.5	-0.6	0.00	1.00	0.48	1.00	66.7
TSPAN18	-2.7	-2.8	-2.1	-2.8	0.00	0.00	0.08	0.00	169.7
MCOLN3	-2.8	-3.0	-3.0	-3.2	0.02	0.12	NA	NA	8.1
KRT73	-2.8	-2.2	-2.1	-2.3	0.00	0.01	0.16	0.02	53.4
HAPLN3	-2.8	-1.6	-1.1	-1.3	0.00	0.00	0.22	0.00	250.9
F5	-2.8	-1.1	-0.5	-1.9	0.00	0.32	1.00	0.01	211.6
RGMB	-2.8	-1.4	-3.1	-1.9	0.00	0.00	0.01	0.00	88.5
FAM19A1	-2.8	-3.0	-2.1	-3.0	0.02	0.02	0.20	0.36	20.6
GP5	-2.9	-0.6	-1.4	0.0	0.00	1.00	0.60	1.00	39.8
EPHA4	-3.0	-2.1	-2.3	-1.9	0.00	0.00	0.00	0.00	688.7
GPA33	-3.0	-1.6	-1.7	-1.4	0.00	0.00	0.00	0.11	158.4
DST	-3.0	-1.0	0.2	0.1	0.00	0.01	1.00	1.00	112.0
ISM1	-3.1	-4.1	-2.3	-1.4	0.03	0.00	NA	NA	15.4
PI16	-3.1	-3.2	-2.8	-2.8	0.00	0.00	NA	0.05	38.8
FUT7	-3.2	-3.2	-1.4	-2.5	0.00	0.00	0.48	0.06	76.7
KALRN	-3.2	-2.0	-2.5	-2.2	0.00	0.00	0.58	1.00	27.2
NEFL	-3.3	-2.2	-2.5	-5.4	0.00	0.00	NA	NA	34.3
S1PR1	-3.3	-2.3	-3.1	-2.6	0.00	0.00	0.00	0.00	1315.8
WNT7A	-3.3	-2.5	-4.5	-3.4	0.00	0.00	0.17	0.00	46.1
SEMA3G	-3.3	-2.7	-0.4	-0.2	0.00	0.00	1.00	1.00	64.3
NPDC1	-3.4	-2.1	-2.8	-2.5	0.00	0.00	0.00	0.00	196.1
SGCD	-3.5	-2.3	-3.3	-1.1	NA	NA	0.01	1.00	12.0
CXCL9	-3.6	-2.8	2.3	6.5	0.05	0.10	1.00	0.00	22.1
DSC1	-3.7	-0.7	-0.7	-0.6	0.03	1.00	1.00	1.00	16.7
ST6GALNAC1	-3.8	-2.4	-0.1	-0.4	0.00	0.00	NA	NA	32.4
SEMA5A	-4.2	-1.4	-2.5	-2.1	0.01	0.43	NA	0.68	22.9
RASSF6	-5.1	-0.6	1.9	0.8	0.01	1.00	NA	NA	6.8
HSD11B1	-5.2	-2.8	1.8	-2.5	0.01	NA	NA	NA	3.3
SPP1	-5.2	0.0	3.4	3.4	0.00	1.00	1.00	1.00	34.0
CXCL10	-5.5	-2.8	1.5	7.9	0.00	0.05	1.00	0.03	25.1

Figure A.B.3. Human lymphoid tissue CD8⁺CD69⁺TEM cells display core features of tissue-resident memory T cells (TRM).

(A) Number of differentially expressed genes between transcriptomes of CD69⁻ and CD69⁺ CD8⁺TEM cells from BM and LN either up or down-regulated (P-adj. <0.05).

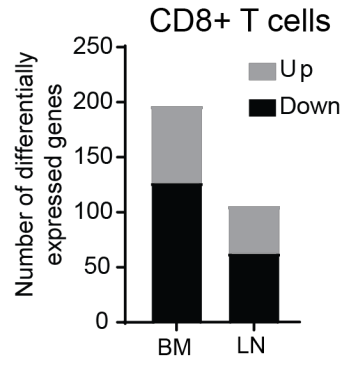
(B) Multidimensional scaling (MDS) analysis of gene expression of CD69⁻ and CD69⁺ CD8⁺TEM cells from indicated tissues of three individuals based on the core TRM gene

signature from Kumar et al [8]. Samples from LN and BM are from the same three individuals, lung and blood TEM cells are from an additional six individuals. (D) Normalized read counts of

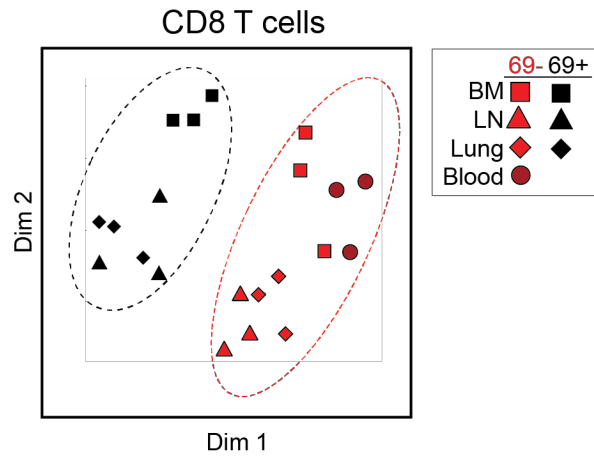
indicated genes from CD8⁺ TEM cells from BM and LN of three individuals. LN= lymph node.

LLN, MLN, and ILN = lung-, mesenteric-, and iliac- draining lymph nodes respectively. Error bars indicate SEM. * P<0.05, ** P<0.01 by two-tailed t-test.

(A)



(B)



(C)

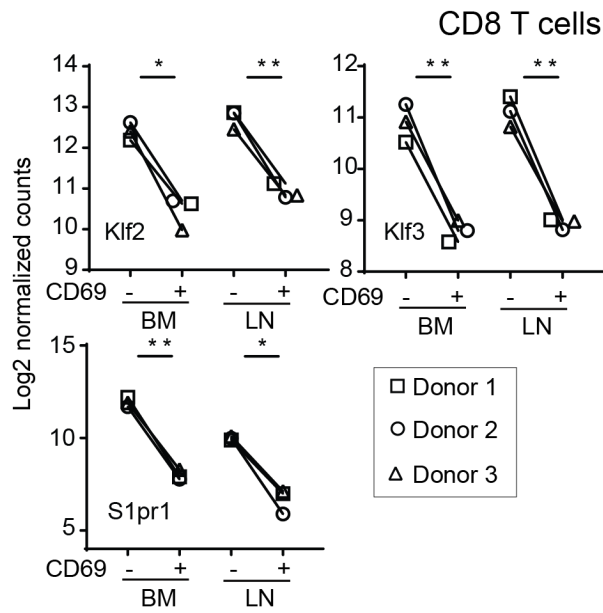


Figure A.B.4. Human lymphoid tissue CD4⁺CD69⁺ memory T cells display core features of TRM.

(A) Number of differentially expressed genes between transcriptomes of CD69⁺ CD4⁺TEM cells relative to CD69⁻CD4⁺TEM from BM and LN either up or down-regulated (P-adj. <0.05).

(B) Multidimensional scaling (MDS) analysis of gene expression of CD69⁻ and CD69⁺ CD4⁺TEM cells from indicated tissues of three individuals based on the differentially expressed genes between CD69⁺ and CD69⁻ CD4⁺TEM in BM (357 total genes). Samples from LN and BM are from the same three individuals, lung and blood TEM cells are from an additional six individuals from Kumar et al [8].

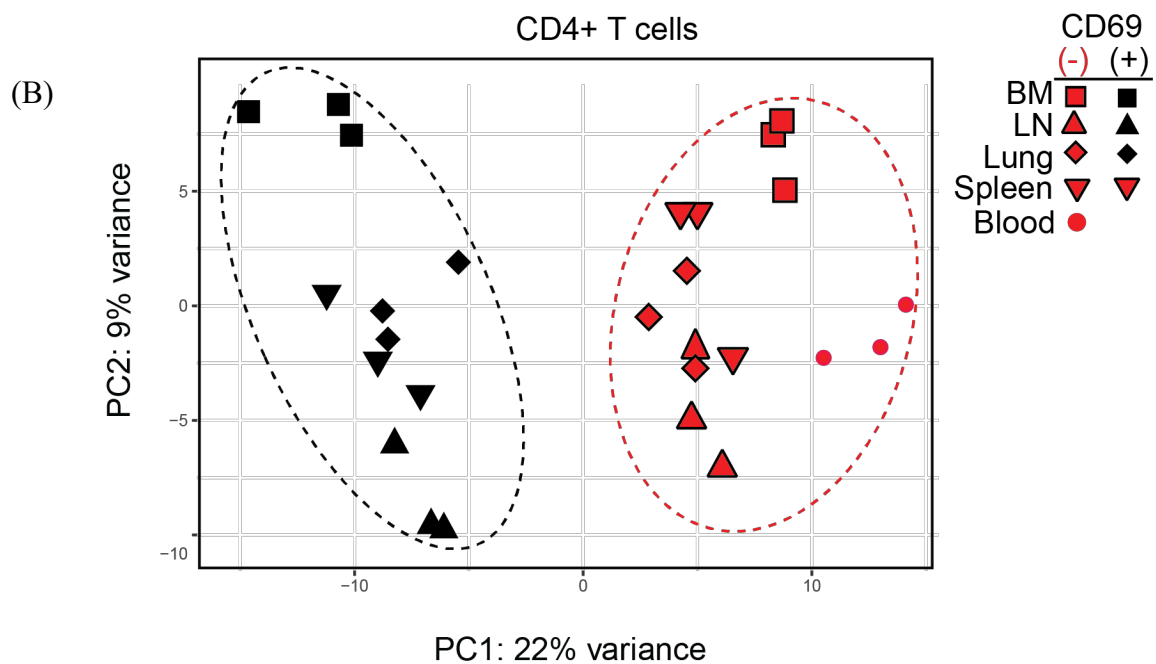
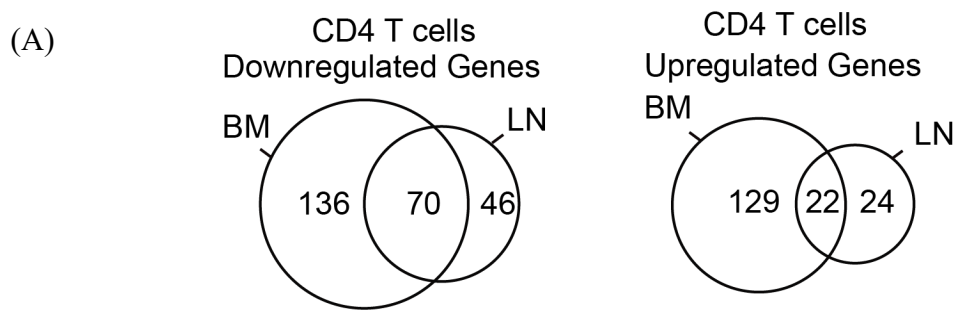
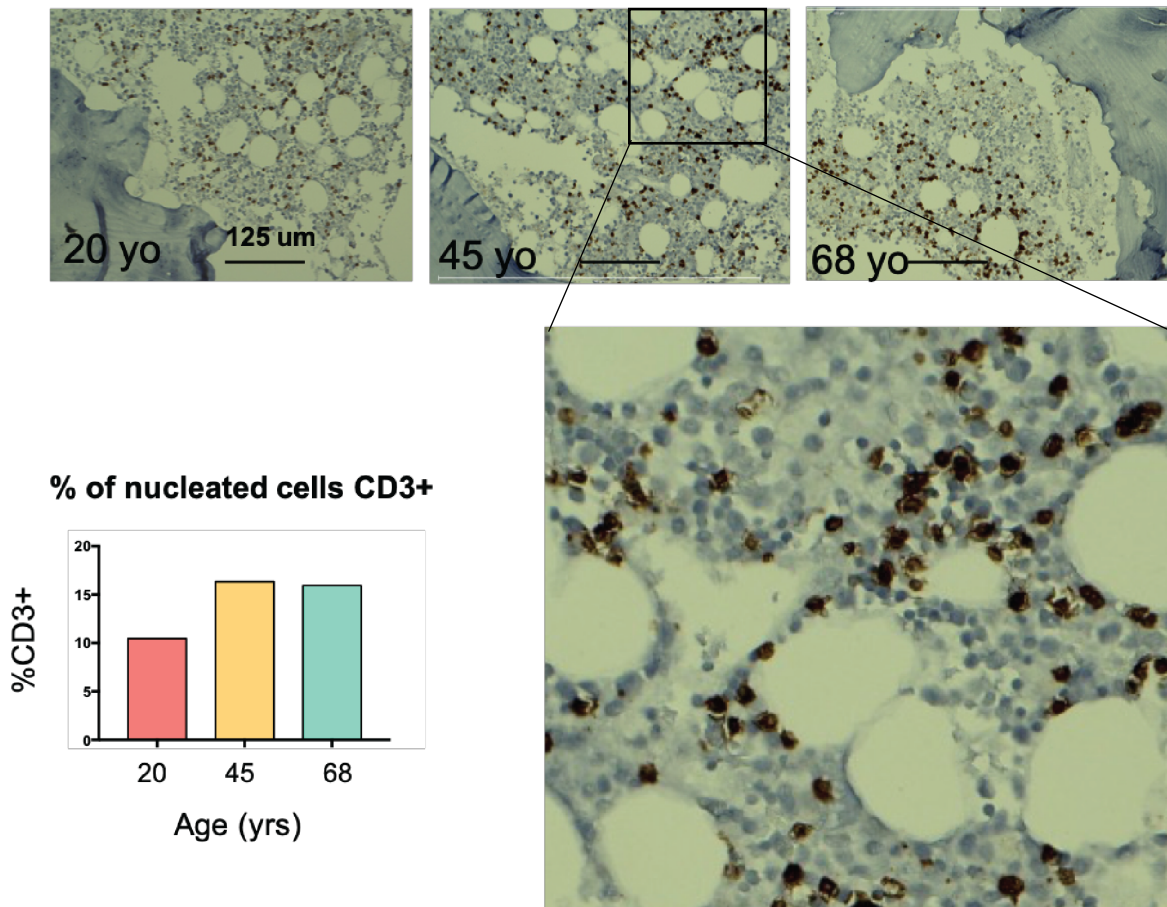


Figure A.B.5. Abundance of CD3⁺ T cells in bone marrow by immunohistochemistry.

Immunohistochemistry staining of a BM section using CD3. Paraffin embedded BM sections (5µM) were mounted onto slides for immunostaining with CD3 antibody. T cells are stained in brown (CD3⁺) and all other nucleated cells are stained in blue (nuclei). Bar graph shown the percentage of all nucleated cells in BM from three individuals of varying ages as indicated. Percentage was averaged from three equal sized sections per slide, as shown by the enlarged section below.



Appendix C. Scripts in Python and R

Section C-1: Cytometry by Time of Flight (CyTOF)

The code used for analyzing CyTOF data is located on my github page (<https://michellemiron.github.io/Human-T-cell-CyTOF/>). This page includes a reproducible analysis using *workflowr*, an R package.

Section C-2: T cell receptor sequencing

QC of raw sequencing reads

Before running the ImmuneDB pipeline, quality control of FASTQ files was performed using pRESTO. First, sequences were trimmed of poor-quality bases on the end farthest from the primer where base call confidence degrades. Using default parameters, sequences are then trimmed to the point where a window of 10 nucleotides has an average quality score of at least 20. If reads are paired, the next step is to align the R1 and R2 reads into full-length, contiguous sequences. Short sequences, those with less than 100 bases, are then removed from further analysis. Finally, any base with a quality score less than 20 is replaced with an N and any sequence containing more than 10 such bases is removed from further analysis. In the case of FASTA input which has no quality information, only paired-end assembly and short sequence removal are recommended. The script for running this process, from Dr. Wenzhao Meng, can be found below. After this process, the remaining filtered sequences are presumed to be of adequate quality for germline inference and clonal assignment using Immune DB.

```
1. """takes sequence files within a folder through the various modules from pRESTO to filter data prior to performing
2. VDJ alignment
3. """
4.
5. import os
6.
7. f = open('~/.list of files.txt', 'r')
8. files = f.readlines()
9. f.close()
10.
11. for line in files:
12.     line = line.strip()
13.     base_line = os.path.basename(line)
14.     directory = os.path.dirname(line)+'/'
15.     file_name = os.path.splitext(base_line)[0]
16.     ext = '.fastq'
17.     ext2 = '.fasta'
18.     extension1 = '_log.txt'
19.     extension2 = 'convert.fasta'
20.
21.     if line[-12:-10] == 'R1':
22.
23.         #Aligns sequences from Read 1 and Read 2
```

```

24. R1line = line
25. R2line = line[:11] + '2_001.fastq'
26. assemble_log_file_name = file_name + '_assemble'+ extension1
27.
28. os.system('python AssemblePairs.py align -1' + R1line + ' -2' + R2line + ' --rc tail --nproc 4 --coord illumina --
failed --log ' + directory + assemble_log_file_name)
29.
30. #qualtrim data based on a sliding window
31. assemble_file = line[:6] + '_assemble-pass'
32. qualtrim_log_file_name = file_name + '_trimqual' + extension1
33.
34. os.system('python FilterSeq.py trimqual -s ' + assemble_file + ext + ' -q 30 --nproc 4 --win 10 --failed --
log ' + directory + qualtrim_log_file_name)
35.
36. #trims sequences so that shorter sequences are removed
37. qual_file = assemble_file + '_trimqual-pass'
38. length_log_file_name = file_name + '_length' + extension1
39.
40. os.system('python FilterSeq.py length -s ' + qual_file + ext + ' -n 100 --nproc 4 --failed --
log ' + directory + length_log_file_name)
41.
42. #maskqual makes low scoring nucleotides become missing values (N)
43. len_file = qual_file + '_length-pass'
44. maskqual_log_file_name = qual_file + '_maskqual' + extension1
45.
46. os.system('python FilterSeq.py maskqual -s ' + len_file + ext + ' -q 30 --nproc 4 --failed')
47.
48. #eliminates sequences with too many missing nucleotides
49. maskqual_file = len_file + '_maskqual-pass'
50. missing_log_file_name = file_name + '_missing' + extension1
51.
52. os.system('python FilterSeq.py missing -s ' + maskqual_file + ext + ' -n 10 --nproc 4 --failed --
log ' + directory + missing_log_file_name)
53.

```

Calculating Clonality

The input of this script is the clone overlap data format exported from the ImmuneDB web interface. This script calculates the clonality for each sample and outputs the calculation for every sample in a single table as a .txt file.

```

1. ##### SET DATE AND LOAD PACKAGES -----
2. options(stringsAsFactors=FALSE)
3. date <- Sys.Date()
4. library(gsubfn)
5. library(lsa)
6. library(stringr)
7. library(dplyr)
8. library(data.table)
9. library(reshape)
10. library(pheatmap)
11. library(RColorBrewer)
12. library(colorspace)
13. library(ggplot2)
14.
15. dostats <- function(i) {
16. cloneoverlap <- read.csv("/Users/michellemiron/Desktop/TCR-paper-Oct-2018/data/2018-10-31-17-
26_clone_overlap.tsv",
17. sep="\t",
18. header=T)

```



```

19. ### remove the uniques column ###
20. clonoverlap2 <- cloneoverlap[,-c(3,5)]
21. ### filter data by donor and replicate ###
22.
23. clean<- clonoverlap2 %>%
24.   dplyr::filter(!str_detect(sample, "remove"), !str_detect(sample,"Fb"))
25. donors <- c("D299", "D287", "D280", "D233", "D383", "D324", "D229", "D255", "LD1", "LD2", "LD3")
26. donor <- donors[i]
27. donordf <- clean %>%
28.   dplyr::filter(str_detect(sample,donor))
29.
30. ##### cast the dataframe so each column is a sample and rows are unique clones #####
31. ddf<- dcast(donordf, clone_id ~ sample, value.var = "copies", fill=0)
32.
33. ##### THE FUNCTIONS ----
34. #These are the functions used to calculate clonality, obtained from Dr. Yufeng Shen's lab
35. normalize <- function(data) {
36.   nc = ncol(data)
37.   for (i in 1:nc) {
38.     data[,i] = data[,i] / sum(data[,i])
39.   }
40.   return(data)
41. }
42.
43. shannon.entropy <- function(p){
44.   if (min(p) < 0 || sum(p) <= 0)
45.     return(NA)
46.   p.norm <- p[p>0]/sum(p)
47.   -sum(log2(p.norm)*p.norm)
48. }
49.
50. Clonality <- function(p) {
51.   x = p[p>0] / sum(p)
52.   l = length(x)
53.   entropy = shannon.entropy(p)
54.   maxentropy = -log2(1/l)
55.   return(signif(1 - entropy / maxentropy, 3))
56. }
57.
58. calcSI<-function(vals){
59.   vals=vals[vals>0]
60.   fq=vals/sum(vals)
61.   si=sum(fq^2)
62.   return(si)
63. }
64.
65. calcr20 = function(X){
66.   X=sort(X,decreasing=T)
67.   X=X[X>0]
68.   CX=cumsum(X)
69.   num=length(which(CX/sum(X)<=0.2))
70.   den=length(X)
71.   return(num/den)
72. }
73. calcr50 = function(X){
74.   X=sort(X,decreasing=T)
75.   X=X[X>0]
76.   CX=cumsum(X)
77.   num=length(which(CX/sum(X)<=0.5))
78.   den=length(X)
79.   return(num/den)
80. }
81.

```

```

82. df <- ddf[, -1]
83.
84. # Apply function to all files in a given directory
85. entropy <- apply(df, 2, shannon.entropy)
86. clonality <- apply(df, 2, Clonality)
87. SI <- apply(df, 2, calcSI)
88. R20 <- apply(df, 2, calcr20)
89. R50 <- apply(df, 2, calcr50)
90. NumberUniqueClones <- apply(df, 2, function(c) sum(c != 0))
91. is.na(df) <- df == 0
92. Mean <- colMeans(df, na.rm = TRUE)
93. Mean50cutoffvalue <- Mean * 0.5
94.
95. statistics <- data.frame(entropy,
96.   clonality,
97.   SI,
98.   R20,
99.   R50,
100.  NumberUniqueClones,
101.  Mean,
102.  Mean50cutoffvalue)
103.
104. return(statistics)
105. }
106.
107. cstats <- data.frame()
108.
109. df1 <- dostats(1)
110. df2 <- dostats(2)
111. df3 <- dostats(3)
112. df4 <- dostats(4)
113. df5 <- dostats(5)
114. df6 <- dostats(6)
115. df7 <- dostats(7)
116. df8 <- dostats(8)
117. df9 <- dostats(9)
118. df10 <- dostats(10)
119. df11 <- dostats(11)
120.
121. T_df <- do.call("rbind", list(df1, df2, df3, df4, df5, df6, df7, df8, df9, df10, df11))
122.
123. write.table(T_df, file = paste("/Users/michellemiron/Desktop/TCR-paper-Oct-2018/data_outputs/Stats_allsamples",
124.   date,
125.   "txt",
126.   sep = ""),
127.   col.names = T,
128.   row.names = T,
129.   sep = "\t")

```

Plotting Clonality

This script inputs the file from the above script, and produces boxplots of clonality data.

```

1. ##### Load required packages #####
2. library(dplyr)
3. library(ggplot2)
4. library(ggbeeswarm)
5. ##### Load in the raw data, diversity metrics #####
6. rawdata <- read.table("/Users/michellemiron/Desktop/TCR-paper-Oct-2018/data_outputs/Stats_allsamples2018-11-01.txt", header = T)

```

```

7.
8. ##### convert rowname to a column to make it easier to filter ###
9. rawdata$sample <- rownames(rawdata)
10.
11. ##### plot of clonality vs. Number of unique clones of all samples #####
12. ggplot(rawdata) + geom_point(aes(clonality,NumberUniqueClones))
13.
14. ##### in order to color by donor and sample, add categorical data #####
15.
16. # load in metadata
17. metadata <- read.csv("/Users/michellemiron/Desktop/TCR-paper-Oct-2018/data/2018-11-01-00-
54_samples.tsv", sep = "\t", header = T )
18.
19. ##combine metadata and data into one df, dropping non matching rows in metadata ##
20.
21. metadata <- as_tibble(metadata)
22. metadata$name <- as.character(metadata$name)
23. rawdata <- as_tibble(rawdata)
24.
25. data a <- left_join(rawdata,metadata, by= c("sample" = "name"))
26.
27. ##filter data to remove water, and etc. ###
28. data_ac <- data_a %>%
29.   filter(subset == "CD4" | subset == "CD8" ) %>%
30.   filter(replicate != "3" & replicate != "4" & replicate != "3A"
31.     & replicate != "2" & replicate != "5" & replicate != "6"
32.     & replicate != "7") %>%
33.   filter(subset2 != "TMEM")
34.
35. data_ac$subset2 <- factor(data_ac$subset2, levels = c("TCM", "TEM", "TMEM",
36.   "TRM","TEMRA"))
37.
38. data_ac$tissue <- factor(data_ac$tissue, levels = c("Bld", "BM", "Spl",
39.   "Lung","LN"))
40.
41. ggplot(data_ac) + geom_point(aes(clonality,NumberUniqueClones, color=subject))
42. ggplot(data_ac) + geom_point(aes(entropy,clonality, color=subset2))
43.
44. ##comparing different subsets in each tissue ###
45. pdf("/Users/michellemiron/Desktop/TCR-paper-Oct-2018/Clonality/clonality-nov.pdf",
46.   width=6,
47.   height=4,
48.   paper='special',
49.   onefile=FALSE)
50.
51. ggplot(data_ac) +
52.   geom_boxplot(aes(subset2,clonality,
53.     color=interaction(subset2)),
54.     outlier.shape = NA) +
55.   facet_grid(subset~tissue, scales="free_x") +
56.   theme(axis.text.x = element_text(angle=90,hjust=1))
57. dev.off()
58.
59.
60. pdf("/Users/michellemiron/Desktop/TCR-paper-Oct-2018/Clonality/clonality-nov2.pdf",
61.   width=6,
62.   height=4,
63.   paper='special',
64.   onefile=FALSE)
65.
66. ggplot(data_ac) +
67.   geom_boxplot(aes(tissue,clonality,
68.     color=interaction(tissue)),

```

```

69.         outlier.shape = NA) +
70.     facet_grid(subset~subset2, scales="free_x") +
71.     theme(axis.text.x = element_text(angle=90,hjust=1))
72. dev.off()
73.
74. pdf("/Users/michellemiron/Desktop/TCR-paper-Oct-2018/Clonality/clonality-tissuenov.pdf",
75.     width=6,
76.     height=4,
77.     paper='special',
78.     onefile=FALSE)
79.
80. ggplot(data_ac) +
81.     geom_boxplot(aes(tissue,clonality,
82.                    color=interaction(tissue)),
83.              outlier.shape = NA) +
84.     facet_grid(~subset, scales="free_x") +
85.     theme(axis.text.x = element_text(angle=90,hjust=1))
86.
87.
88. dev.off()
89.
90. give.n <- function(x){
91.     return(c(y = mean(x), label = length(x)))
92. }
93.
94. ggplot(data_ac,aes(tissue,clonality)) +
95.     geom_boxplot(aes(color=interaction(tissue)),
96.              outlier.shape = NA) +
97.     stat_summary(fun.data = give.n, geom = "text") +
98.     facet_grid(~subset, scales="free_x") +
99.     theme(axis.text.x = element_text(angle=90,hjust=1))
100.
101.
102.
103. subsetc <- "CD4"
104. CDcdat <- data_ac %>% filter (subset == subsetc,tissue!="Bld")
105. pairwise.t.test(CDcdat$clonality,
106.                interaction(CDcdat$tissue),
107.                paired = FALSE,
108.                p.adjust="none",
109.                conf.level = 0.95)
110.
111.
112. ##Comparing subset ###
113. pdf("/Users/michellemiron/Desktop/TCR-paper-Oct-2018/Clonality/clonality-subsets-nov3.pdf",
114.     width=6,
115.     height=4,
116.     paper='special',
117.     onefile=FALSE)
118.
119. ggplot(data_ac) +
120.     geom_boxplot(aes(subset2,clonality,
121.                    color=interaction(subset2)),
122.              outlier.shape = NA) +
123.     facet_grid(~subset, scales="free_x")
124.
125. dev.off()
126.
127. ggplot(data_ac,aes(subset2,clonality)) +
128.     geom_boxplot(aes(color=interaction(subset2)),
129.              outlier.shape = NA) +
130.     stat_summary(fun.data = give.n, geom = "text") +
131.     facet_grid(~subset, scales="free_x")

```

Clonal Proportion Plots

The input to this script is the "vdjtools" data format exported from ImmuneDB web interface.

This script produces clonal proportion plots.

```

1. #Use tcR package for analysis of TCR sequencing data
2. library("tcR")
3.
4. donor<- c("D229", "D233", "D299", "D287", "D280", "D255", "D324", "D383", "LD1", "LD2", "LD3")
5.
6. i<-11
7. for (i in 1:length(donor)) {
8.   Location <-
9.     file.path("/Users/", "michellemiron", "Dropbox", "TCR data Brian and Michelle", "ImmuneDBdata", donor[i])
10.  List_of_files<- list.files(Location)
11.  listofdf <- list()
12.
13.  for (y in 1:length(List_of_files)){
14.    filepath <- paste(Location, "/", List_of_files[y], sep="")
15.    data <- parse.cloneset(filepath,
16.                          "cdr3nt",
17.                          "cdr3aa",
18.                          "count",
19.                          "count",
20.                          "v", "j", "d",
21.                          NA, NA, NA, NA, NA, NA, FALSE,
22.                          "\t")
23.    name <- List_of_files[y]
24.    tmp <- list(data)
25.    listofdf[[name]] <- tmp
26.  }
27.
28.
29.  plotname <- paste("clonal_proportion", "_", donor[i], sep="")
30.
31.  pdf(paste("~/Desktop/TCR-paper-Oct-2018/Clonal-Proportion/", plotname, ".pdf", sep=""),
32.      width=6,
33.      height=4,
34.      paper='special',
35.      onefile=FALSE)
36.
37.  print(vis.top.proportions(listofdf, c(10, 100, 1000, 10000, 20000), .col = "Read.count"))
38.  dev.off()
39. }

```

Calculating Cosine Similarity and Plotting Heatmaps

The input to this script is the clone overlap data format exported from ImmuneDB web interface.

This script produces heatmaps of calculated cosine similarity values.

```
1. ##### temporary donor IDS #####
2. d <- 3
3. j <- 2
4. cutoff = "no"
5. ##### Cluster method #####
6. cluster_method <- "complete"
7.
8. library(gsubfn)
9. library(lsa)
10. library(stringr)
11. library(dplyr)
12. library(data.table)
13. library(reshape)
14. library(pheatmap)
15. library(RColorBrewer)
16. library(colorspace)
17. library(ggplot2)
18.
19.
20. ##### Create list of all donors and lineages that will be plotted #####
21.
22. donors <- c("D287", "D233", "D280", "D299", "D229", "D255", "D383", "D324", "LD1", "LD2", "LD3")
23. lineages <- c("CD4", "CD8")
24.
25. ##### Load in the data #####
26. cloneoverlap <- read.csv("/Users/michellemiron/Desktop/TCR-paper-Oct-2018/data/2018-10-31-17-
26_clone_overlap.tsv",
27.   sep="\t",
28.   header=T)
29.
30. ##### filter the data for CD4 or CD8 T cells for reps 1 and 2 #####
31. donor <- donors[d]
32. lineage <- lineages[j]
33. clonoverlap2 <- cloneoverlap[,-3]
34.
35. ### filter data by donor and replicate ###
36. donordf <- clonoverlap2 %>% filter(str_detect(sample, donors[d]) &
37.   str_detect(sample, lineages[j]))
38. notmem <- donordf %>% filter(!str_detect(sample, "TMEM"))
39.
40. donor_remove <- notmem %>%
41.   filter(!str_detect(sample, "remove"), !str_detect(sample, "Fb"))
42. donordfRep12 <- donor_remove %>%
43.   filter(str_detect(sample, "_1") | str_detect(sample, "_2"))
44.
45. ##### cast the dataframe so each column is a sample and rows are unique clones #####
46. c_donordf <- dcast(donordf, clone_id ~ sample, value.var = "copies", fill=0)
47. c_donordfRep12 <- dcast(donordfRep12, clone_id ~ sample, value.var = "copies", fill=0)
48.
49. ##calculate mean freq for each sample
50. #filter df to have no zeros
51. nadf <- c_donordfRep12
52. nadf[nadf==0] <- NA
53. df.mean <- colMeans(nadf, na.rm=TRUE)
54. cut.off <- 0.5*df.mean
55. cut.off <- cut.off[-1]
```

```

56.
57. df.50mean <- c_donordfrep12[,-1]
58. head(df.50mean)
59. apply(df.50mean,2,min, na.rm=TRUE)
60. i <- 1
61.
62. for (i in 1:length(cut.off)) {
63.     df.50mean[,i][df.50mean[,i]<cut.off[i]] <- 0
64. }
65.
66.
67. ##### calculcate cosine #####
68. cosdf <- cosine(as.matrix(c_donordf[,-1]))
69. cosdf12reps <- cosine(as.matrix(c_donordfrep12[,-1]))
70.
71. cos.50meandf <- cosine(as.matrix(df.50mean[,-1]))
72.
73. ##### pick data to plot #####
74. if (cutoff == "yes") {
75.     df <- cos.50meandf } else {
76.     df <- cosdf12reps
77. }
78.
79. if (cutoff == "yes") {
80.     PlotName <- paste("Cosine",
81.         donors[d],
82.         cluster_method,
83.         lineages[j], "50MeanFreq")
84. } else {
85.     PlotName <- paste("Cosine",
86.         donors[d],
87.         cluster_method,
88.         lineages[j])
89. }
90.
91. ##make settings for heatmaps
92. breaksList = seq(0.0, max(df, na.rm=TRUE), by = 0.005)
93. customcolor = colorRampPalette(c("black", "red", "yellow", "white"))(length(breaksList))
94.
95. ### make cluster labels ###
96. conditions <- data.frame(
97.     Tissue = unlist(gsubfn::strapplyc(colnames(df), "Bld|BM|LLN|Lung|Spl|LN",simplify = TRUE)),
98.     CellType = unlist(gsubfn::strapplyc(colnames(df), "TCM|TEMRA|TEM|TRM|TMEM",simplify = TRUE)),
99.     Lineage = unlist(gsubfn::strapplyc(colnames(df), "CD4|CD8",simplify = TRUE)),
100.     Replicate = unlist(gsubfn::strapplyc(colnames(df), "_1|_2|_3|_4|_5|_6|_7",simplify = TRUE))
101. )
102.
103. conditions2 <- cbind(conditions,population=paste(conditions$Tissue,conditions$CellType,sep=""))
104. conditions2$Replicate <- gsub("_1","rep1",conditions2$Replicate)
105. conditions2$Replicate <- gsub("_2","rep2",conditions2$Replicate)
106. conditions2$Replicate <- as.factor(conditions2$Replicate)
107.
108.
109. rowlab <- gsub(paste(donors[d], "_",sep=""),"",rownames(df))
110. rowlab2 <- gsub(paste("CD4", "_",sep=""),"",rowlab)
111. rowlab3 <- gsub(paste("CD8", "_",sep=""),"",rowlab2)
112.
113. conditions2 <- conditions2[,c(1,2,5,4,3)]
114.
115. rownames(df) <- gsub(paste(donors[d], "_",sep=""),"",rownames(df))
116. colnames(df) <- gsub(paste(donors[d], "_",sep=""),"",colnames(df))
117. colnames(df) <- gsub(paste(lineages[j], "_",sep=""),"",colnames(df))
118. rownames(df) <- gsub(paste(lineages[j], "_",sep=""),"",rownames(df))

```

```

119. colnames(df) <- gsub(" ", "", colnames(df))
120. rownames(df) <- gsub("_", "", rownames(df))
121.
122. rownames(conditions2) <- colnames(df)
123.
124. palette <- brewer.pal(4, "Set2")
125. tisspal <- brewer.pal(11, "RdGy")
126.
127. ann_colors = list(
128.   Replicate = c(rep1 = "orange", rep2 = "blue"),
129.   Tissue = c(BM=tisspal[7],
130.     Spl=tisspal[8],
131.     Lung=tisspal[9],
132.     LN=tisspal[10],
133.     Bld=tisspal[11]
134.   ),
135.   CellType = c(TCM=palette[2],
136.     TEM =palette[1],
137.     TRM=palette[3],
138.     TEMRA=palette[4]
139.   )
140. )
141.
142. conditions3 <- conditions2[,c("Tissue", "Replicate",
143.   "CellType")]
144. pdf(paste("~/Desktop/CosineCalc/plots/cosineheatmap/", PlotName, ".pdf", sep=""),
145.   width=4,
146.   height=6,
147.   paper='special',
148.   onefile=FALSE)
149.
150. pheatmap(df,
151.   color = customcolor,
152.   cluster_rows=T,
153.   cluster_cols = T,
154.   border_color = NA,
155.   main=PlotName,
156.   breaks = breaksList,
157.   cellwidth = 5,
158.   cellheight = 5,
159.   fontsize_row = 5,
160.   fontsize_col = 5,
161.   legend = TRUE,
162.   annotation_legend = TRUE,
163.   annotation_row = conditions3,
164.   annotation_col = conditions3,
165.   annotation_colors = ann_colors,
166.   clustering_method = cluster_method,
167.   treeheight_row = 4,
168.   treeheight_col =4,
169.   fontsize=3,
170.   drop_levels= TRUE
171. )
172.
173. dev.off()
174. ##### take the mean of cosine between replicates #####
175.
176. ##how to calculate cosine between reps
177. ### keep X1 and X2 which are different from eachother
178. if (cutoff == "yes") {
179.   cosUnique <- melt(cos.50meandf) %>%
180.     filter(X1 != X2)
181. } else {

```



```

182. cosUnique <- melt(cosdf12reps) %>%
183.   filter(X1 != X2)
184. }
185.
186. ### keep X1 and X2 which are not reps of eachother
187. cos_ext <- cosUnique %>%
188.   filter(str_sub(
189.     cosUnique$X1,1,
190.     str_length(cosUnique$X1)-2)!=
191.     str_sub(cosUnique$X2,1,str_length(cosUnique$X2)-2))
192.
193.
194. cos_mean <- cos_ext
195.
196. cos_mean$X1 <- str_sub(as.character(cos_mean$X1),1,
197.   str_length(as.character(cos_mean$X1))-2)
198.
199. cos_mean$X2 <- str_sub(as.character(cos_mean$X2),1,
200.   str_length(as.character(cos_mean$X2))-2)
201.
202. cos_sumMean <- cos_mean %>%
203.   group_by(.dots=c("X1","X2")) %>%
204.   summarize(x=mean(value))
205.
206. cos_mean_matrix <- dcast(cos_sumMean,X1~X2, value.var= "x", fill=1)
207. rownames(cos_mean_matrix) <- cos_mean_matrix[,1]
208.
209. df <- cos_mean_matrix[,-1]
210.
211. ##set plot name
212. if (cutoff == "yes") {
213.   PlotName2 <- paste("Cosine",
214.     donor,
215.     "mean R1&2",
216.     cluster_method,
217.     lineages[j], "50MeanFreq")
218. } else {
219.   PlotName2 <- paste("Cosine",
220.     donor,
221.     "mean R1&2",
222.     cluster_method,
223.     lineages[j])
224. }
225.
226.
227. ##make settings for heatmaps
228. breaksList = seq(0.0, max(df, na.rm=TRUE), by = 0.005)
229. customcolor = colorRampPalette(c("black", "red", "yellow", "white"))(length(breaksList))
230.
231. ### make cluster labels ###
232. conditions <- data.frame(
233.   Tissue = unlist(gsubfn::strapplyc(colnames(df), "Bld|BM|LLN|Lung|Spl|LN",simplify = TRUE)),
234.   CellType = unlist(gsubfn::strapplyc(colnames(df), "TCM|TEMRA|TEM|TRM|TMEM",simplify = TRUE)),
235.   Lineage = unlist(gsubfn::strapplyc(colnames(df), "CD4|CD8",simplify = TRUE)))
236.
237.
238. rownames(conditions) <- colnames(df)
239. conditions2 <- as.data.frame(conditions[,c("Tissue", "CellType")])
240. rownames(conditions2) <- colnames(df)
241. rownames(df) <- gsub(paste(donors[d], " ", sep=""), "", rownames(df))
242. colnames(df) <- gsub(paste(donors[d], " ", sep=""), "", colnames(df))
243. colnames(df) <- gsub(paste(lineages[j], " ", sep=""), "", colnames(df))
244. rownames(df) <- gsub(paste(lineages[j], " ", sep=""), "", rownames(df))

```

```

245. colnames(df) <- gsub(" ", "", colnames(df))
246. rownames(df) <- gsub(" _", "", rownames(df))
247. rownames(conditions2) <- colnames(df)
248.
249. palette <- brewer.pal(4, "Set2")
250. tisspal <- brewer.pal(11, "RdGy")
251.
252. ann_colors = list(Tissue = c(BM=tisspal[7],
253.                               Spl=tisspal[8],
254.                               Lung=tisspal[9],
255.                               LN=tisspal[10],
256.                               Bld=tisspal[11]),
257.                  CellType = c(TCM=palette[2],
258.                                TEM =palette[1],
259.                                TRM=palette[3],
260.                                TEMRA=palette[4])
261.                )
262.
263.
264. pdf(paste("~/Desktop/CosineCalc/plots/cosineheatmap/mean/", PlotName2, ".pdf", sep=""),
265.     width=4,
266.     height=6,
267.     paper='special',
268.     onefile=FALSE)
269.
270. pheatmap(df,
271.          color = customcolor,
272.          cluster_rows=T,
273.          cluster_cols = T,
274.          main=PlotName2,
275.          breaks = breaksList,
276.          border_color = NA,
277.          cellwidth = 8,
278.          cellheight = 8,
279.          fontsize_row = 5,
280.          fontsize_col = 5,
281.          legend = TRUE,
282.          annotation_legend = TRUE,
283.          annotation_row = conditions2,
284.          annotation_col = conditions2,
285.          annotation_colors = ann_colors,
286.          clustering_method = cluster_method,
287.          treeheight_row = 5,
288.          treeheight_col = 5,
289.          fontsize=8,
290.          lables_row= gsub(paste(donor, " _", sep=""), "", rownames(df)),
291.          lables_col= gsub(paste(donor, " _", sep=""), "", colnames(df))
292.        )
293. dev.off()

```

Plotting Jaccard Index Boxplots

The input to this script are excel sheets containing matrices of Jaccard index calculations. This script produces boxplots of the Jaccard index values.

1. `##This code will input the 06082018 data from Wenzhao for Jaccard`

```

2.  ## file is Jaccard 50fc inFraction
3.
4.  library(readxl)
5.  library(dplyr)
6.  library(ggplot2)
7.  library(dplyr)
8.  library(ggsci)
9.  library(ggbeeswarm)
10. library(reshape2)
11.
12. ##### converting the excel sheet with multiple tabs into data frames in R
13. #####MAKING THE DATAFRAME #####
14. ##This is a help page with a function to read excel sheets and make data frames from the tabs:
15. #https://stackoverflow.com/questions/12945687/read-all-worksheets-in-an-excel-workbook-into-an-r-list-with-data-frames
16. read_excel_allsheets <- function(filename, tibble = FALSE) {
17.   # I prefer straight data.frames
18.   # but if you like tidyverse tibbles (the default with read_excel)
19.   # then just pass tibble = TRUE
20.   sheets <- readxl::excel_sheets(filename)
21.   x <- lapply(sheets, function(X) readxl::read_excel(filename, sheet = X))
22.   if(!tibble) x <- lapply(x, as.data.frame)
23.   names(x) <- sheets
24.   x
25. }
26.
27. mysheets <- read_excel_allsheets("~/Desktop/TCR-paper-Oct-2018/Jaccard Index/Jaccard_50fc_inFraction_Reps.xlsx")
28. # Now we have a list of dataframes, each data frame was a tab in the excel sheet
29.
30. str(mysheets[[1]])
31.
32. df <- melt(mysheets)
33. df <- df[,-4]
34. str(df)
35. #remove duplicated values
36. dfSingle <- df[!duplicated(df$value), ]
37. str(dfSingle)
38. colnames(dfSingle) <- c("Sample1", "Sample2", "value")
39. dfSingle$Sample2 <- as.character(dfSingle$Sample2)
40.
41. #rename a column in the dataframe
42.
43. celltype <- paste(gsubfn::strapplyc(dfSingle[,1], "CD4|CD8", simplify = TRUE),
44.   gsubfn::strapplyc(dfSingle[,1],
45.     "LN|Spl|BM|Lung|Blood", simplify = TRUE),
46.   gsubfn::strapplyc(dfSingle[,1],
47.     "TEM|TRM|TEMRA|TCM", simplify = TRUE),
48.   "-",
49.   gsubfn::strapplyc(dfSingle[,2], "CD4|CD8", simplify = TRUE),
50.   gsubfn::strapplyc(dfSingle[,2],
51.     "LN|Spl|BM|Lung|Blood", simplify = TRUE),
52.   gsubfn::strapplyc(dfSingle[,2], "TEM|TRM|TEMRA|TCM", simplify = TRUE))
53.
54.
55. celltype <- as.factor(celltype)
56. str(celltype)
57. donor <- gsubfn::strapplyc(dfSingle[,1], "D229|D233|D280|D287|D299|D383|D324|D255", simplify = TRUE)
58.
59. Lineage <- paste(gsubfn::strapplyc(dfSingle[,1], "CD4|CD8", simplify = TRUE),
60.   "-",
61.   gsubfn::strapplyc(dfSingle[,2], "CD4|CD8", simplify = TRUE))
62.

```

```

63. Tissue <- paste(gsubfn::strapplyc(dfSingle[,1],
64.           "LN|Spl|BM|Lung|Blood",simplify = TRUE),
65.           "-",
66.           gsubfn::strapplyc(dfSingle[,2],
67.           "LN|Spl|BM|Lung|Blood",simplify = TRUE))
68.
69. Subset <- paste(gsubfn::strapplyc(dfSingle[,1],"TEM|TRM|TEMRA|TCM",simplify = TRUE),
70.           "-",
71.           gsubfn::strapplyc(dfSingle[,2],"TEM|TRM|TEMRA|TCM",simplify = TRUE))
72.
73. Sample1SubsetTissue <- paste(gsubfn::strapplyc(dfSingle[,1],
74.           "TEM|TRM|TEMRA|TCM",simplify = TRUE),
75.           "-",
76.           gsubfn::strapplyc(dfSingle[,1],
77.           "LN|Spl|BM|Lung|Blood",simplify = TRUE)
78.           )
79.
80.
81. Sample2SubsetTissue <- paste(gsubfn::strapplyc(dfSingle[,2],
82.           "TEM|TRM|TEMRA|TCM",simplify = TRUE),
83.           "-",
84.           gsubfn::strapplyc(dfSingle[,2],
85.           "LN|Spl|BM|Lung|Blood",simplify = TRUE)
86.           )
87.
88. plotData <- dfSingle %>%
89.   cbind(celltype) %>%
90.   cbind(donor) %>%
91.   cbind(Lineage) %>%
92.   cbind(Tissue) %>%
93.   cbind(Subset) %>%
94.   cbind(Sample1SubsetTissue) %>%
95.   cbind(Sample2SubsetTissue)
96.
97.
98. library(summarytools)
99.
100. #order the data by mean
101. plotData$celltype <- with(plotData, reorder(celltype, value, mean))
102. CD4CD8 <- plotData %>% filter(Lineage == "CD4 - CD8" | Lineage == "CD8 - CD4")
103. mean(CD4CD8$value)
104. ##### PLOTTING ALL THE SPLEEN BM AND LN SAMPLES #####
105.
106. ##summarize the data to plot error bars toO!
107. #+++++
108. # Function to calculate the mean and the standard deviation
109. # for each group
110. #+++++
111. # data : a data frame
112. # varname : the name of a column containing the variable
113. #to be summarized
114. # groupnames : vector of column names to be used as
115. # grouping variables
116.
117.
118. plotCD4 <- plotData %>% filter(Lineage == "CD4 - CD4",
119.           Tissue != "Spl - Lung",
120.           Tissue != "BM - Lung",
121.           Tissue != "Blood - Lung",
122.           Tissue != "LN - Lung",
123.           Tissue != "Spl - Blood",
124.           Tissue != "BM - Blood",
125.           Tissue != "Blood - Lung",

```

```

126.         Tissue != "LN - Blood",
127.         Tissue != "Lung - Spl",
128.         Tissue != "Lung - BM",
129.         Tissue != "Lung - Blood",
130.         Tissue != "Lung - LN",
131.         Tissue != "Blood - Spl",
132.         Tissue != "Blood - BM",
133.         Tissue != "Lung - Blood",
134.         Tissue != "Blood - LN",
135.         Tissue != "Lung - Lung",
136.         Tissue != "Blood - Blood",
137.         Subset != "TCM - TCM",
138.         Subset != "TCM - TEM",
139.         Subset != "TCM - TRM",
140.         Subset != "TEM - TCM",
141.         Subset != "TRM - TCM"
142.     )
143.
144. plotCD8 <- plotData %>% filter(Lineage == "CD8 - CD8",
145.     Tissue != "Spl - Lung",
146.     Tissue != "BM - Lung",
147.     Tissue != "Blood - Lung",
148.     Tissue != "LN - Lung",
149.     Tissue != "Spl - Blood",
150.     Tissue != "BM - Blood",
151.     Tissue != "Blood - Lung",
152.     Tissue != "LN - Blood",
153.     Tissue != "Lung - Spl",
154.     Tissue != "Lung - BM",
155.     Tissue != "Lung - Blood",
156.     Tissue != "Lung - LN",
157.     Tissue != "Blood - Spl",
158.     Tissue != "Blood - BM",
159.     Tissue != "Lung - Blood",
160.     Tissue != "Blood - LN",
161.     Tissue != "Lung - Lung",
162.     Tissue != "Blood - Blood",
163.     Subset != "TEMRA - TEMRA",
164.     Subset != "TEMRA - TEM",
165.     Subset != "TEMRA - TRM",
166.     Subset != "TEM - TEMRA",
167.     Subset != "TRM - TEMRA"
168. )
169.
170. min.mean.sd.max <- function(x) {
171.     r <- c(min(x), mean(x) - sd(x), mean(x), mean(x) + sd(x), max(x))
172.     names(r) <- c("ymin", "lower", "middle", "upper", "ymax")
173.     r
174. }
175.
176. q1.mean.sd.q3 <- function(x) {
177.     r <- c(quantile(x, c(.25)), mean(x) - sd(x), mean(x), mean(x) + sd(x), quantile(x, c(.75)))
178.     names(r) <- c("ymin", "lower", "middle", "upper", "ymax")
179.     r
180. }
181.
182.
183. min.mean.se.max <- function(x) {
184.     sem <- function(X) {
185.         se <- sd(x)/sqrt(length(x))
186.         se
187.     }

```

```

188.   r <- c(min(x), mean(x) - sem(x), mean(x), mean(x) + sem(x), max(x))
189.   names(r) <- c("ymin", "lower", "middle", "upper", "ymax")
190.   r
191. }
192.
193. q1.mean.se.q3 <- function(x) {
194.   sem <- function(X) {
195.     se <- sd(x)/sqrt(length(x))
196.     se
197.   }
198.   r <- c(quantile(x, c(.25)), mean(x) - sem(x), mean(x), mean(x) + sem(x), quantile(x, c(.75)))
199.   names(r) <- c("ymin", "lower", "middle", "upper", "ymax")
200.   r
201. }
202.
203.
204. ####try plotting replicates, separate from non reps..separate from within tissue
205. #####-----Plot Reps CD4-----#####
206.
207. ##Reps:
208. RepsCD4 <- plotCD4 %>% filter((celltype == "CD4 BM TRM - CD4 BM TRM" |
209.                               celltype == "CD4 BM TEM - CD4 BM TEM" |
210.                               celltype == "CD4 LN TEM - CD4 LN TEM" |
211.                               celltype == "CD4 Spl TEM - CD4 Spl TEM" |
212.                               celltype == "CD4 LN TRM - CD4 LN TRM" |
213.                               celltype == "CD4 Spl TRM - CD4 Spl TRM"
214.                               )
215.                               )
216.
217. #####-----Plot Reps CD8-----#####
218.
219. RepsCD8 <- plotCD8 %>% filter((celltype == "CD8 BM TRM - CD8 BM TRM" |
220.                               celltype == "CD8 BM TEM - CD8 BM TEM" |
221.                               celltype == "CD8 LN TEM - CD8 LN TEM" |
222.                               celltype == "CD8 Spl TEM - CD8 Spl TEM" |
223.                               celltype == "CD8 LN TRM - CD8 LN TRM" |
224.                               celltype == "CD8 Spl TRM - CD8 Spl TRM"
225.                               )
226.                               )
227.
228. #####-----Plot within CD4 Tissue-----#####
229.
230.
231. ##within Tissue:
232. WithinCD4 <- plotCD4 %>% filter((celltype == "CD4 BM TRM - CD4 BM TEM" |
233.                               celltype == "CD4 BM TEM - CD4 BM TRM" |
234.                               celltype == "CD4 LN TEM - CD4 LN TRM" |
235.                               celltype == "CD4 Spl TEM - CD4 Spl TRM" |
236.                               celltype == "CD4 LN TRM - CD4 LN TEM" |
237.                               celltype == "CD4 Spl TRM - CD4 Spl TEM"
238.                               )
239.                               )
240.
241. #####-----Plot within CD8 Tissue-----#####
242.
243.
244. WithinCD8 <- plotCD8 %>% filter((celltype == "CD8 BM TRM - CD8 BM TEM" |
245.                               celltype == "CD8 BM TEM - CD8 BM TRM" |
246.                               celltype == "CD8 LN TEM - CD8 LN TRM" |
247.                               celltype == "CD8 Spl TEM - CD8 Spl TRM" |
248.                               celltype == "CD8 LN TRM - CD8 LN TEM" |
249.                               celltype == "CD8 Spl TRM - CD8 Spl TEM"
250.                               )

```

```

251.     )
252.
253. #####-----Plot between CD8 Tissue-----#####
254.
255.
256. BetweenCD8 <- plotCD8 %>% filter((celltype == "CD8 Spl TEM - CD8 BM TEM" |
257.     celltype == "CD8 Spl TRM - CD8 BM TRM" |
258.     celltype == "CD8 LN TEM - CD8 BM TEM" |
259.     celltype == "CD8 LN TRM - CD8 BM TRM" |
260.     celltype == "CD8 LN TEM - CD8 BM TEM" |
261.     celltype == "CD8 LN TEM - CD8 BM TRM" |
262.     celltype == "CD8 Spl TEM - CD8 LN TEM" |
263.     celltype == "CD8 Spl TRM - CD8 LN TRM" |
264.     celltype == "CD8 Spl TEM - CD8 BM TRM" |
265.     celltype == "CD8 Spl TEM - CD8 LN TRM" |
266.     celltype == "CD8 Spl TRM - CD8 LN TEM" |
267.     celltype == "CD8 Spl TRM - CD8 BM TEM"
268.     )
269.     )
270.
271.
272.
273. #####-----Plot between CD4 Tissue-----#####
274.
275. BetweenCD4 <- plotCD4 %>% filter((celltype == "CD4 Spl TEM - CD4 BM TEM" |
276.     celltype == "CD4 Spl TRM - CD4 BM TRM" |
277.     celltype == "CD4 LN TEM - CD4 BM TEM" |
278.     celltype == "CD4 LN TRM - CD4 BM TRM" |
279.     celltype == "CD4 LN TEM - CD4 BM TEM" |
280.     celltype == "CD4 LN TEM - CD4 BM TRM" |
281.     celltype == "CD4 Spl TEM - CD4 LN TEM" |
282.     celltype == "CD4 Spl TRM - CD4 LN TRM" |
283.     celltype == "CD4 Spl TEM - CD4 BM TRM" |
284.     celltype == "CD4 Spl TEM - CD4 LN TRM" |
285.     celltype == "CD4 Spl TRM - CD4 LN TEM" |
286.     celltype == "CD4 Spl TRM - CD4 BM TEM"
287.     )
288.     )
289.
290. #####-----Plotting the data-----#####
291. DataList <- list(RepsCD4,
292.     RepsCD8,
293.     WithinCD4,
294.     WithinCD8,
295.     BetweenCD8,
296.     BetweenCD4,
297.     plotCD4,
298.     plotCD8)
299.
300. for (i in 1:8) {
301.   data <- DataList[[i]]
302.
303.   give.n <- function(x){
304.     return(c(y = mean(x), label = length(x)))
305.   }
306.   pdf(paste("plot",i,".pdf",sep = ""))
307.   plot <- ggplot(aes(y = value, x = factor(celltype)), data = data)
308.   p <- plot +
309.     stat_summary(fun.data = q1.mean.se.q3, geom = "boxplot") +
310.     #geom_boxplot(width=0.5) +
311.     #geom_jitter(aes(shape= donor),
312.       position=position_jitter(width=.1), size=3) +
313.     ggtitle("Overlap between two samples line at mean,

```

```

314.         SEM, quantile 1 and 3") +
315.   xlab("Sample 1 and 2") +
316.   ylab("Overlap Index") +
317.   scale_shape_manual(values=c(0,1,5,6,15,16,17,18)) +
318.   theme(legend.key.size = unit(0.4,"cm"),
319.         axis.text.x = element_text(angle=90, hjust=1)) +
320.   geom_hline(yintercept = mean(CD4CD8$value)) +
321.   ylim(0,0.44) +
322.   coord_flip() +
323.   stat_summary(fun.data = give.n, geom = "text")
324. print(p)
325. dev.off()
326. }
327.
328. ggplot(aes(x=interaction(celltype), y=value )) +
329.   #colour=interaction(celltype)) +
330.   #geom_violin() +
331.   theme_minimal() +
332.   #geom_bar(position = "dodge", stat = "summary", fun.y = "mean") +
333.   geom_boxplot(outlier.shape = NA, fill= "grey") +
334.   #geom_quasirandom(aes(shape= donor),size=2) +
335.   scale_shape_manual(values=c(0,1,5,6,15,16,17,18)) +
336.   ggtitle("Jaccard CD8") +
337.   theme(legend.key.size = unit(0.35,"cm"),
338.         axis.text.x = element_text(angle=90, hjust=1)) +
339.   ylim(0,0.44) +
340.   coord_flip() +
341.   geom_hline(yintercept = mean(CD4CD8$value))

```

Clone Tracking Script #1

The input for this script is the file produced by Clone Tracking function from VDJtools

[312](test.tracking.strict.table.collapsed.txt). The output is a text file to be used for the Clone Tracking Script #2 below.

```

1.  ### take input file of test.tracking.strict.table.collapsed.txt
2.  #and fix the peak column to be the top frequency column
3.  print(getwd())
4.  setwd("../")
5.  print(getwd())
6.  library(data.table)
7.  dt <- read.table("VDJtools_output/test.tracking.strict.table.collapsed.txt",
8.                 header=T)
9.
10. meta <- read.table("data/metadata.txt", sep="\t",
11.                  header=T)
12.
13. freq.df <- dt[,c(as.character(meta$sample_id))]
14. freq.dt <- as.data.table(freq.df)
15. freq.dt[, MAX := colnames(.SD)[max.col(.SD, ties.method="first")]]
16. df.max <- freq.dt
17. df.peak.empty <- data.frame(maxsample = c(df.max$MAX), peak= c(NA))
18.
19. ##fill peak number based on key using match function
20. key <- data.frame(sample= meta$sample_id, number=

```



```

21.           c(0:(nrow(meta)-1))
22.         )
23.
24. peak.correct <- (match(df.peak.empty[,1],key[,1]))-1
25.
26. dt$peak <- peak.correct
27.
28.
29. file.loc <- "VDJtools_output/test.tracking.strict.table.collapsed2.txt"
30. write.table(dt,
31.           file= file.loc,
32.           sep="\t",
33.           quote= FALSE)
34.

```

Clone Tracking Script #2

The input for this script is the text file from the above script. This script is modified from the VDJtools software[312] and produced the clone tracking plots.

```

1. #currentRscript <- rstudioapi::getSourceEditorContext()$path
2. #setwd(gsub("/tracking_stackplot_custom.R*", "",currentRscript))
3.
4. args <- c("sample",
5.         "VDJtools_output/test.tracking.strict.table.collapsed2.txt",
6.         "plots/custom.stackplot.pdf",
7.         "data/metadata.txt")
8. .libPaths("/usr/local/Cellar/vdjtools/1.1.8/Rpackages/")
9. # data input
10. firstwd <-getwd()
11. setwd("../")
12. #args <- commandArgs(TRUE)
13.
14. require(ggplot2); require(reshape)
15. require(gridExtra); require(grid)
16. require(RColorBrewer); require(stringr)
17. require(dplyr)
18.
19. label <- args[1] #"time since HSCT, months"
20. file_in <- args[2] #"luc_table_collapsed.txt"
21. file_out <- args[3] #"out.pdf"
22. meta_in <- args[4] # "metadata.txt"
23.
24. #load metadata to make labels
25. meta <- read.table(meta_in, comment="", sep = "\t", header = TRUE)
26. custom.x.label <- as.character(meta$sample_id)
27.
28. # load time points, create some auxillary variables
29. x <- matrix(as.numeric(0:(nrow(meta)-1)))
30. n <- length(x)
31.
32. # load data
33. #df <- data.frame(read.delim(file_in))
34. df <- read.table(file_in, comment="", sep = "\t", header = TRUE)
35. xcols <- (ncol(df) - n + 1):ncol(df)
36. fcols <- 1:(ncol(df) - n)
37. xlbls <- colnames(df)[xcols]

```

```

38.
39. # convert abundance columns to numeric
40. df[, xcols] <- apply(df[, xcols], 2, as.numeric)
41.
42. # set up Non-overlapping and Not-shown
43. df$peak[nrow(df)-1] <- -1
44. df$peak[nrow(df)] <- -2
45.
46. # reshape data
47. df.m <- melt(df, id = fcols)
48.
49. # replace sample ids (factor) by time (numeric)
50. ind <- match(df.m$variable, xlbls)
51. df.m$variable <- x[ind]
52.
53. df.m$sign <- paste(df.m$cd3nt, df.m$v, df.m$d, df.m$j, sep=" ")
54. palette <- brewer.pal(4, "Set2")
55. pal <- colorRampPalette(c(palette[1],
56.                           palette[2],
57.                           palette[3],
58.                           palette[4])
59.                          )
60.
61. #draw
62.
63. if (grepl("\\.pdf$", file_out)){
64.   pdf(file_out)
65. } else if (grepl("\\.png$", file_out)) {
66.   png(file_out, width = 3.25,
67.        height = 3.25,
68.        units = "in",
69.        res = 1200,
70.        pointsize = 4)
71. } else {
72.   stop("Unknown plotting format")
73. }
74.
75. # make current pallet
76. ## make map of label to color
77.
78. df_map2 <- read.table("data/colorpalette.map.txt", header=T, sep="\t",
79.                      comment.char = "")
80.
81. #df_map <- data.frame(sample_map=sample_map, pal_map=pal_map)
82. # loop over custom.x.label
83. df_answer <- data.frame(sample_answer=custom.x.label, pal_answer=rep(NA,n))
84. df_answer$pal_answer <- df_map2$pal_map[match(df_answer$sample_answer, df_map2$sample_map)]
85. df_answer$peakNumber <- c(0:(n-1))
86.
87. #filter df_answer for rows that match the present factors in peak
88. peaks_to_filter_by <- as.integer(levels(as.factor(df.m$peak)))[-c(1,2)]
89. df_answer_c <- df_answer[df_answer$peakNumber %in% peaks_to_filter_by, ]
90.
91. # for each entry, look in the map to find the color
92. # add that color to the pallet
93.
94. # prepare plot
95. g<-ggplot() +
96.   geom_area(data = df.m,
97.            aes(variable, value, group = sign, fill = factor(peak)),
98.            colour = "gray25", position = 'stack', size = 0) +
99.   scale_fill_manual(
100.    name = "Peak position",

```

```

101. breaks = c(-2, -1, df_answer c$peakNumber),
102. labels = c("Non-overlapping", "Not-shown", as.character(df_answer_c$sample_answer)),
103. values = c("grey50", "grey70", as.character(df_answer_c$pal_answer))
104. ) +
105. scale_x_continuous(breaks = 0:(n-1),
106. labels = paste0(c(custom.x.label))) +
107. #scale_x_continuous(expand = c(0,0), limit = c(min(x), max(x)), breaks = x) +
108. scale_y_continuous(expand = c(0,0)) +
109. theme_bw() +
110. xlab(label) +
111. ylab("abundance") +
112. theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 0.5), panel.border = element_blank()) +
113. guides(size = F)
114.
115. # disable label cropping
116. gg_table <- ggplot_table(ggplot_build(g))
117. gg_table$layout$clip[gg_table$layout$name=="panel"] <- "off"
118.
119. grid.draw(gg_table)
120.
121. dev.off()
122.

```

Section C-3: RNA-seq analysis script

The input to this script are the text files from each sample containing a count matrix. This script performs differential expression analysis.

```

1. #Run Deseq pipeline ----
2. #HERE IS THE VIGNETTE: https://bioconductor.org/packages/3.7/bioc/vignettes/DESeq2/inst/doc/DESeq2.html
3.
4. #DEPENDENCIES ----
5. library("Rcpp")
6. library("colorspace")
7. library("DESeq2")
8.
9. # COUNT TABLE ALL SAMPLES ----
10. CountTable <-
11. read.table("/Volumes/Dom_CCTI/CCTI_USERS/Michelle Miron/BM and LN CD69 data/AllRawCounts.csv",
12. header=T,sep=";",
13. row.names=1)
14. Samples <- data.frame(row.names=colnames(CountTable), condition=as.factor(c(rep("LNCD4_RO+69-",3),
15. rep("LNCD4_RO+69+",3), rep("BMCD4_RO+69-",3),
16. rep("BMCD4_RO+69+",3),rep("LNCD8_RO+69-",3),
17. rep("LNCD8_RO+69+",3),rep("BMCD8_RO+69-",3),
18. rep("BMCD8_RO+69+",3)
19. )
20. )
21. )
22.
23. # COUNT TABLE FOR EACH SAMPLE PAIR ----
24. Cts_LNCD4 <- CountTable[1:25559,1:6]
25. Cts_BMCD4 <- CountTable[1:25559,7:12]
26. Cts_LNCD8 <- CountTable[1:25559,13:18]
27. Cts_BMCD8 <- CountTable[1:25559,19:24]

```

```

28.
29. Samp__LNCD4 <- data.frame(row.names=colnames(Cts_LNCD4),
30.   condition=as.factor(c(rep("69neg",3),
31.     rep("69pos",3)
32.   )
33. )
34. )
35.
36. Samp__BMCD4 <- data.frame(row.names=colnames(Cts_BMCD4),
37.   condition=as.factor(c(rep("69neg",3),
38.     rep("69pos",3)
39.   )
40. )
41. )
42.
43. Samp__LNCD8 <- data.frame(row.names=colnames(Cts_LNCD8),
44.   condition=as.factor(c(rep("69neg",3),
45.     rep("69pos",3)
46.   )
47. )
48. )
49.
50. Samp__BMCD8 <- data.frame(row.names=colnames(Cts_BMCD8),
51.   condition=as.factor(c(rep("69neg",3),
52.     rep("69pos",3)
53.   )
54. )
55. )
56.
57. # RUNNING DESEQ ----
58.
59. ddsLN4 <- DESeqDataSetFromMatrix(countData = Cts_LNCD4, colData=Samp__LNCD4, design=~condition)
60. ddsBM4 <- DESeqDataSetFromMatrix(countData = Cts_BMCD4, colData=Samp__BMCD4, design=~condition)
61. ddsLN8 <- DESeqDataSetFromMatrix(countData = Cts_LNCD8, colData=Samp__LNCD8, design=~condition)
62. ddsBM8 <- DESeqDataSetFromMatrix(countData = Cts_BMCD8, colData=Samp__BMCD8, design=~condition)
63.
64.
65. ddsLN4<- DESeq(ddsLN4)
66. ddsBM4<- DESeq(ddsBM4)
67. ddsLN8<- DESeq(ddsLN8)
68. ddsBM8<- DESeq(ddsBM8)
69.
70. resLN4 <- results(ddsLN4)
71. resBM4 <- results(ddsBM4)
72. resLN8 <- results(ddsLN8)
73. resBM8 <- results(ddsBM8)
74.
75. setwd("/Users/michellemiron/Dropbox/Dissertation/resources/")
76. write.csv(as.data.frame(resLN4),
77.   file="resLN4condition_treated_results.csv")
78. write.csv(as.data.frame(resBM4),
79.   file="resBM4condition_treated_results.csv")
80. write.csv(as.data.frame(resLN8),
81.   file="resLN8condition_treated_results.csv")
82. write.csv(as.data.frame(resBM8),
83.   file="resBM8condition_treated_results.csv")
84.
85.
86. browseVignettes("DESeq2")
87. #log fold change shrinkage for ranking and visualization
88.
89. resLN4LFC <- lfcShrink(ddsLN4, coef="condition_69pos_vs_69neg")
90. resBM4LFC <- lfcShrink(ddsBM4, coef="condition_69pos_vs_69neg")

```

```

91. resLN8LFC <- lfcShrink(ddsLN8, coef="condition 69pos vs 69neg")
92. resBM8LFC <- lfcShrink(ddsBM8, coef="condition_69pos_vs_69neg")
93.
94.
95.
96. rldLN4 <- rlog(ddsLN4, blind=FALSE)
97. ntdLN4 <- normTransform(ddsLN4)
98.
99. #MAKING VISUALIZATIONS -----
100.
101. plotMA(resLN4LFC, ylim=c(-2,2))
102. plotMA(resLN8LFC, ylim=c(-2,2))
103. plotMA(resBM4LFC, ylim=c(-2,2))
104. plotMA(resBM8LFC, ylim=c(-2,2))
105.
106.
107. #idx <- identify(resLN4$baseMean, resLN4$log2FoldChange)
108. #plotCounts(ddsLN4, gene=which.min(resLN4$padj), intgroup="condition")
109.
110. install.packages("pheatmap")
111. library("pheatmap")
112.
113. select <- order(rowMeans(counts(ddsLN4,normalized=TRUE)),
114.                 decreasing=TRUE)[1:100]
115.
116. ntdLN4 <- normTransform(ddsLN4)
117.
118. log2.norm.counts <- assay(ntdLN4)[select,]
119.
120. testnames <- c(colnames(log2.norm.counts))
121.
122. df <- as.data.frame(colData(ddsLN4)[,c("condition")])
123.
124. rownames(df) <- testnames
125.
126. pheatmap(log2.norm.counts, cluster_rows=FALSE, show_rownames=TRUE,
127.           cluster_cols=FALSE, annotation_col=df)
128.
129. plotPCA(rldLN4, intgroup=c("condition"))
130.
131.
132. #TESTING SOME HEATMAPS -----
133. #from this website: https://www.biostars.org/p/178748/
134. library(RColorBrewer)
135. breaksList = seq(-1, 1, by = 0.005)
136. plotUpDownSigGenes <- function(results, colNums, sampleNames, rld, title) {
137.
138.   # make the lists
139.   upgenes <- rownames(head(results[order(results$log2FoldChange),], n=200))
140.   downgenes <- rownames(head(results[order(-results$log2FoldChange),], n=200))
141.
142.   # this gives us the rows we want
143.   rows <- match(upgenes,row.names(rld))
144.   mat <- assay(rld)[rows,colNums]
145.   mat <- mat - rowMeans(mat)
146.
147.   # the labels are hard coded at the moment :(
148.   df <- as.data.frame(colData(rld)[,c("condition")])
149.   rownames(df) <- c(sampleNames)
150.   pheatmap(mat,
151.            color = colorRampPalette(rev(brewer.pal(n = 7, name = "RdYlBu")))(length(breaksList)),cluster_rows=FAL
SE,
152.            fontsize=3, annotation_col=df, main=paste(title, "top 200 DE genes"),breaks = breaksList)

```

```

153.
154. # this gives us the rows we want
155. rows <- match(downgenes, row.names(rld))
156. mat <- assay(rld)[rows,colNums]
157. mat <- mat - rowMeans(mat)
158.
159. df <- as.data.frame(colData(rld)[,c("condition")])
160. rownames(df) <- c(sampleNames)
161. pheatmap(mat,
162.           color = colorRampPalette(rev(brewer.pal(n = 7, name = "RdYlBu")))(length(breaksList)),cluster_rows=FAL
SE,
163.           fontsize=3, annotation_col=df, main=paste(title,"top 200 DE genes"),breaks = breaksList)
164. }
165. aCols <- c(1,2,3)
166. bCols <- c(4,5,6)
167.
168. ##make a combined heatmap up and down regulated genes
169. breaksList1 = seq(-3, -1.0, by = 0.5)
170. breaksList2 = seq(-1.01, 1, by = 0.05)
171. breaksList3 = seq(1.01, 3, by = 0.5)
172. breaksList <- c(breaksList1,breaksList2,breaksList3)
173. plotUpDownSigGenes <- function(results, colNums, sampleNames, rld, title) {
174.
175.   # make the lists
176.   upgenes <- rownames(head(results[order(results$log2FoldChange),], n=40))
177.   downgenes <- rownames(head(results[order(-results$log2FoldChange),], n=40))
178.
179.   # this gives us the rows we want
180.   rows <- match(upgenes,row.names(rld))
181.   mat <- assay(rld)[rows,colNums]
182.   mat <- mat - rowMeans(mat)
183.
184.   rows2 <- match(downgenes, row.names(rld))
185.   mat2 <- assay(rld)[rows2,colNums]
186.   mat2 <- mat2 - rowMeans(mat2)
187.
188.   # the labels are hard coded at the moment :(
189.   df <- as.data.frame(colData(rld)[,c("condition")])
190.   rownames(df) <- c(sampleNames)
191.
192.   pheatmap(rbind(mat,mat2),
193.            color = colorRampPalette(rev(brewer.pal(n = 11, name = "RdBu")))(length(breaksList)),
194.            cluster_rows=F,cluster_cols = F,
195.            fontsize=6, annotation_col=df, main=paste(title,"top 80 DE genes"),breaks = breaksList)
196. }
197.
198. #CD4 LN
199. contrastDEGenes <- subset(results(ddsLN4), padj < 0.05)
200. sampleNames <- c("DM004","DM005","DM006","DM016","DM017","DM018")
201. rldLN4 <- rlog(ddsLN4, blind=F)
202. plotUpDownSigGenes(contrastDEGenes,c(aCols, bCols),sampleNames,rldLN4,"Lymph node CD4 Memory")
203.
204. #CD4 BM
205. contrastDEGenes2 <- subset(results(ddsBM4), padj < 0.05)
206. sampleNames2 <- colnames(assay(ddsBM4))
207. rldBM4 <- rlog(ddsBM4, blind=F)
208. plotUpDownSigGenes(contrastDEGenes2,c(aCols, bCols),sampleNames2,rldBM4,"Bone Marrow CD4 Memory")
209.
210. #CD8 LN
211. contrastDEGenes3 <- subset(results(ddsLN8), padj < 0.05)
212. sampleNames3 <- colnames(assay(ddsLN8))
213. rldLN8 <- rlog(ddsLN8, blind=F)
214. plotUpDownSigGenes(contrastDEGenes3,c(aCols, bCols),sampleNames3,rldLN8,"LN CD8 Memory")

```

```

215.
216. #CD8 BM
217. contrastDEGenes4 <- subset(results(ddsBM8), padj < 0.05)
218. sampleNames4 <- colnames(assay(ddsBM8))
219. rldBM8 <- rlog(ddsBM8, blind=F)
220. plotUpDownSigGenes(contrastDEGenes4,c(aCols, bCols),sampleNames4,rldBM8,"Bone Marrow CD8 Memory")
221.
222. # OVERLAP BETWEEN BM AND LN ----
223.
224. numberDEgenes <- function(dds) {
225.   genelistPvalLNCD4 <- subset(results(dds, condition=c("69pos","69neg")), padj < 0.05)
226.   listUpLNCD4 <- rownames(subset(genelistPvalLNCD4,log2FoldChange>1))
227.   listDownLNCD4 <- rownames(subset(genelistPvalLNCD4,log2FoldChange<(-1)))
228.   print(c(length(listUpLNCD4),"up"))
229.   print(c(length(listDownLNCD4),"down"))
230. }
231.
232. # LN CD4 gene lists
233. genelistPvalLNCD4 <- subset(results(ddsLN4, condition=c("69pos","69neg")), padj < 0.05)
234. listUpLNCD4 <- rownames(subset(genelistPvalLNCD4,log2FoldChange>1))
235. listDownLNCD4 <- rownames(subset(genelistPvalLNCD4,log2FoldChange<(-1)))
236.
237.
238. # BM CD4 gene lists
239. genelistPvalBMCD4 <- subset(results(ddsBM4, condition=c("69pos","69neg")), padj < 0.05)
240. listUpBMCD4 <- rownames(subset(genelistPvalBMCD4,log2FoldChange>1))
241. listDownBMCD4 <- rownames(subset(genelistPvalBMCD4,log2FoldChange<(-1)))
242.
243.
244. ListLNBMC4UP <- intersect(listUpLNCD4,listUpBMCD4)
245. ListLNBMC4Down <- intersect(listDownLNCD4,listDownBMCD4)
246. ListLNBMC4UpDOWN <- intersect(listUpLNCD4,listDownBMCD4)
247. ListLNBMC4UpDOWN2 <- intersect(listDownLNCD4,listUpBMCD4)
248.
249. numberDEgenes(ddsBM4)
250. numberDEgenes(ddsLN4)
251. numberDEgenes(ddsBM8)
252. numberDEgenes(ddsLN8)
253.
254.
255. #making matrixes of gene lists
256. #Combined DEseq BM and LN
257.
258. Cts_CD4 <- CountTable[1:25559,1:12]
259. Samp_CD4 <- data.frame(row.names=colnames(Cts_CD4),
260.   condition=as.factor(c(rep("69neg",3),rep("69pos",3),
261.   rep("69neg",3),rep("69pos",3)
262.   )
263.   )
264.   )
265.
266.
267. ddsCD4 <- DESeqDataSetFromMatrix(countData = Cts_CD4, colData=Samp_CD4, design=~condition)
268. ddsCD4<- DESeq(ddsCD4)
269. rldCD4 <- rlog(ddsCD4, blind=F)
270.
271. breaksList1 = seq(-3, -1.0, by = 0.5)
272. breaksList2 = seq(-1.01, 1, by = 0.05)
273. breaksList3 = seq(1.01, 3, by = 0.5)
274. breaksList <- c(breaksList1,breaksList2,breaksList3)
275. plotOverlapSigGenes <- function(listUP1,listUP2,listDown1,listDown2,colNums, sampleNames, rld, title) {
276.
277.   # make the lists

```

```

278. ListUP <- intersect(listUP1,listUP2)
279. ListDown <- intersect(listDown1,listDown2)
280.
281. # this gives us the rows we want
282. rows <- match(ListUP,row.names(rld))
283. mat <- assay(rld)[rows,colNums]
284. mat <- mat - rowMeans(mat)
285.
286. rows2 <- match(ListDown, row.names(rld))
287. mat2 <- assay(rld)[rows2,colNums]
288. mat2 <- mat2 - rowMeans(mat2)
289.
290. # the labels are hard coded at the moment :(
291. df <- as.data.frame(colData(rld)[,c("condition")])
292. rownames(df) <- c(sampleNames)
293.
294. pheatmap(rbind(mat,mat2),
295.           color = colorRampPalette(rev(brewer.pal(n = 11, name = "RdBu")))(length(breaksList)),
296.           cluster_rows=F,cluster_cols = T,
297.           fontsize=6, annotation_col=df, main=paste(title,"Overlapping DE genes"),breaks = breaksList)
298. }
299.
300. plotOverlapSigGenes(listUpBMCD4,listUpLNCD4,listDownBMCD4,
301.                     listDownLNCD4, 1:12,colnames(ddsCD4),rldCD4,"LN and BM" )
302.
303. rowsLNCD4 <- match(listUpLNCD4,row.names(ddsLN4))
304. rldLN4 <- rlog(ddsLN4, blind=F)
305. matLNCD4 <- assay(ddsLN4)[rowsLNCD4,c(1:6)]
306. matLNCD4 <- mat - rowMeans(mat)
307.
308.
309.
310. ##MODIFY THIS TO GET OVERLAP BETWEEN TWO SAMPLES AND PLOT THEM BOTH
311. plotUpDownSigGenes2 <- function(results1, results2, colNums, sampleNames, rld, title) {
312.
313.   # make the lists
314.   upgenes <- rownames(results1[order(results1$log2FoldChange),])
315.   downgenes <- rownames(results1[order(-results1$log2FoldChange),])
316.
317.   print(length(upgenes))
318.   print(length(downgenes))
319.
320.   # make the lists from second data set
321.   upgenes2 <- rownames(results2[order(results2$log2FoldChange),])
322.   downgenes2 <- rownames(results2[order(-results2$log2FoldChange),])
323.
324.   print(length(upgenes2))
325.   print(length(downgenes2))
326.
327.   # this gives us the rows we want
328.   rows <- match(upgenes,row.names(rld))
329.   mat <- assay(rld)[rows,colNums]
330.   mat <- mat - rowMeans(mat)
331.
332.   rows2 <- match(downgenes, row.names(rld))
333.   mat2 <- assay(rld)[rows2,colNums]
334.   mat2 <- mat2 - rowMeans(mat2)
335.
336.   # the labels are hard coded at the moment :(
337.   df <- as.data.frame(colData(rld)[,c("condition")])
338.   rownames(df) <- c(sampleNames)
339.
340.   pheatmap(rbind(mat,mat2),

```



```
341.     color = colorRampPalette(rev(brewer.pal(n = 11, name = "RdBu")))(length(breaksList)),
342.     cluster_rows=F,cluster_cols = F,
343.     fontsize=6, annotation_col=df, main=paste(title,"top 80 DE genes"),breaks = breaksList)
344. }
345.
```

Appendix D. Accepted Abstracts

2017 Oral Presentation: **Tissue-reservoirs of anti-viral T cell immunity in persistent human CMV infection.**
Federation of Clinical Immunology Societies (FOCIS)
2017 Annual Meeting in Chicago, IL
FCE Fusion Award Recipient

Authors: **Michelle Miron**, Claire L. Gordon, Joseph J.C. Thome, Dustin Carpenter, Takashi Senda, Nobuhide Matsuoka, Joshua Weiner, Michael A. Rak, Suzu Igarashi, Tomer Granot, Harvey Lerner, Felicia Goodrum, and Donna L. Farber.

Abstract:

T cell responses to viruses are initiated and maintained in tissue sites; however, knowledge of human anti-viral T cells is largely derived from blood. Cytomegalovirus (CMV) persists in most humans, requires T cell immunity to control, yet tissue immune responses remain undefined. Here, we investigated human CMV-specific T cells, virus persistence and CMV-associated T cell homeostasis in blood, lymphoid, mucosal and secretory tissues of 44 CMV seropositive and 28 seronegative donors. CMV-specific T cells were maintained in distinct distribution patterns, highest in blood, bone marrow (BM), or lymph nodes (LN), with the frequency and function in blood distinct from tissues. CMV genomes were detected predominantly in lung and also in spleen, BM, blood and LN. High frequencies of activated CMV-specific T cells were found in blood and BM samples with low virus detection, while in lung, CMV-specific T cells were present along with detectable virus. In LNs, CMV-specific T cells exhibited quiescent phenotypes independent of virus. We observed differential responses of CMV-specific T cells in distinct tissues to stimulation by CMV antigen. Additionally, certain epitope-specific T cell populations in BM were tissue-resident in phenotype (CD69+) and localized to the BM, while others were TEM (CD69-) and circulatory. Polyclonal T cell differentiation was enhanced in sites of viral persistence with age. Together, our results suggest tissue T cell reservoirs for CMV control shaped by both viral and tissue-intrinsic factors, with global effects on homeostasis of tissue T cells over the lifespan.

2018

Oral Presentation: **Human lymph nodes maintain a distinct subset of TCF-1hi resident memory T cells throughout life.**

Keystone Symposia on Molecular and Cellular Biology
Translational Systems Immunology Meeting

Authors: **Michelle Miron**, Brahma Kumar, Wenzhao Meng, Tomer Granot, Dustin Carpenter, Takashi Senda, Adeeb Rahman, Eline T. Luning Prak, and Donna L. Farber.

Abstract:

Tissue resident memory T cells (TRM) predominate in barrier sites and mediate protective immunity, while their role in lymphoid tissues is undefined. Here we analyzed memory CD8+ T cells in different lymphoid compartments including bone marrow, spleen, and lymph nodes (LN) relative to lung within diverse individuals. We identify an organ-specific subset in human LN (TLN) not found in blood or other tissues, expressing high levels of TCF-1 and transcriptionally enriched for markers of quiescence, self-renewal and follicular-helper cells. High dimensional CyTOF analysis reveals TLN as intermediate in differentiation between naive and TRM cells, with circulating memory T cells the most differentiated. TLN exhibit higher TCR diversity, lower in vivo turnover, yet higher proliferative responses compared to memory cells in other lymphoid or mucosal sites. These findings establish human LN as reservoirs for diverse memory T cells poised for high expansion and TLN as important targets for in vivo immunotherapies.

2018

Oral Presentation: **Human lymph nodes are reservoirs for self-renewing memory T cells throughout life.**

Women in Science at Columbia (WISC)

2nd Annual Graduate Research Symposium

Authors: **Michelle Miron**, Brahma V. Kumar, Dustin J. Carpenter, Takashi Senda, Yufeng Shen⁴
Adeeb Rahman, and Donna L. Farber

Abstract:

In an immune response, T cells are activated and differentiate to effector and memory subsets that play distinct roles in adaptive immunity and immune homeostasis. Effector cells are short lived and secrete proinflammatory cytokines and cytotoxic mediators for pathogen destruction, while memory T cells are long-lived and can mediate rapid recall responses upon antigen re-encounter. Studies in mouse models have identified key transcription factors (TF) that determine effector versus memory T cell fate; T cell factor-1 (TCF-1) is essential for memory T cell formation and maintenance in the periphery and T-bet promotes effector over memory T cell differentiation. For human T cells, the role of specific TF in the differentiation and maintenance of effector and memory T cells remain unclear. In addition to T cell intrinsic factors, the tissue environment also influences T cell fate and memory maintenance. Human memory T cells are diversely distributed across multiple anatomic sites and comprise the predominant subset in most tissues for the majority of life. In this study, we investigated the tissue determinants of human T cell differentiation using our unique human tissue resource where we obtain blood, multiple lymphoid and mucosal tissues from individual organ donors of all ages through a longstanding collaboration with LiveOnNY, the organ procurement organization for the New York City metropolitan area. By studying T cells across tissues of individuals, we have identified that memory CD8⁺T cells maintained in human LN are organ-specific. Notably, LN memory CD8⁺T cells maintain high expression of transcription factor TCF-1 associated with cellular quiescence and self-renewal, exhibit low turnover and a higher proliferative capacity compared to memory T cells in other lymphoid (spleen, bone marrow) and peripheral sites (e.g., lungs). Together these findings establish human LN as important targets for promoting protection in vaccines and a source of high potential T cells for harnessing in immunotherapies.

2018 Oral Presentation: **Mapping Memory T-cell Clones in Human Tissues.**
Cold Spring Harbor Laboratories (CSHL)
Biennial Biological Data Science Meeting.

Authors: **Michelle Miron**, Wenzhao Meng, Eline T. Luning Prak, and Donna L. Farber.

Abstract:

New T cell responses result in the expansion of T cell clones which disseminate to lymphoid and barrier tissues where they are maintained as memory T cells and provide long-lived protective immunity. T cells that can be detected in peripheral blood provide a sampling of a T cell response but may not be representative of memory T cell responses in tissues that accumulate over lifetime of previous antigen encounters. We have established a collaboration and protocol with Live On NY, the organ donor network for the NY metropolitan area, which enables us to procure multiple healthy lymphoid and mucosal tissues from individual organ donors. Here we present a high-throughput T cell receptor (TCR) sequencing approach to characterize the human memory T cell response in diverse tissues. The large size and diversity of the human memory T cell response presents a great technical challenge; we therefore developed tools to quantify and statistically analyze T cell repertoire diversity and sharing between samples. Using both our unique resource and these newly developed tools, we present a map of memory T cell clones within the body. We sequenced 110 biological samples across eight individuals each with over 100,000 clones identified per sample. We mapped these clones to all four major memory T cell subsets (ie. TCM, central-; TEM, effector-; TEMRA, terminally differentiated effector; and TRM, tissue-resident-memory) and five tissue sites including blood, bone marrow, lymph nodes, spleen and lung. Our results reveal that all CD4⁺ T cell subsets are more clonally diverse than all CD8⁺ subsets independent of tissue site. In particular, the most clonally expanded CD8⁺ T cell subset are TEMRA cells, and the most diverse CD4⁺ T cell subset to be the TCM subset. We also find tissue-specific features of memory T cell responses, including an increased diversity of TEM and TRM cells within lymph nodes compared to other sites. We applied cosine similarity and Jaccard Index calculations to the TCR data to assess clonal overlap between subsets and tissues. Notably, TEM subsets are more connected to TEM subsets in other tissues than they are to TRM counterpart subsets in other tissues – and likewise for TRM subsets with other TRM subsets in tissues. This suggests a lineage connection between TRM subsets across tissues, and a lineage connection between TEM subsets across tissues. Overall, our findings provide a map of T-cell memory clone connectivity in tissues. These methods and results show what a healthy immune response looks like and can serve as a baseline for future studies investigating T cell responses in individuals with pathologies such as infections, autoimmunity and cancer.

Appendix E. Abstracts of contributing author manuscripts

Szabo, P.A., Levitin, H.M., **Miron, M.**, Snyder, M., Senda, T., Yuan, J., Cheng, Y.L., Bush, E.C., Dogra, P., Thapa, P., Farber, D.L., and Sims, P.A. (Submitted) Single cell transcriptomics defines activation states and relates human blood and tissue T cells.

Human T cells coordinate adaptive immunity by localization in diverse tissue sites, though blood T cells are the most readily studied. We investigated the role of tissue site in T cell responses by single-cell RNA-seq profiling of >50,000 resting and TCR-stimulated T cells isolated from human lungs (LG), lymph nodes (LN), bone marrow (BM) and blood. Using new factorization methods, we defined cellular states for resting and activated T cells conserved across tissues, including an IFN-response activation state in CD4⁺T cells and distinct effector states specific to CD8⁺T cells. Between sites, T cells from LG and LN were distinct; blood T cells were most similar to those in BM and contained trace numbers of cells bearing tissue resident profiles. Our results reveal signatures for the maintenance and activation of human tissue and blood-derived T cells important for interpreting and monitoring T cell immunity in health and diseases.

Fu, J., Zuber, J., Martinez, M, Shonts, B; Obradovic, A; Wang, H., Lau, S., Xia, A., Waffarn, E; Frangaj, K., Savage, T., Simson, M., Yang, S; Guo, X., **Miron, M.**, Senda, T., Rogers, K., Rahman, A., Ho, S., Shen, Y., Griesemer, A., Farber, D.L., Kato, T., Sykes, M. (2018) Human Intestinal Allografts Contain Functional Hematopoietic Stem and Progenitor Cells that Are Maintained by a Circulating Pool. *Cell Stem Cell*. doi: 10.1016/j.stem.2018.11.007

Human intestinal transplantation often results in long-term mixed chimerism of donor and recipient blood in transplant patients. We followed the phenotypes of chimeric peripheral blood cells in 21 patients receiving intestinal allografts over 5 years. Donor lymphocyte phenotypes suggested a contribution of hematopoietic stem and progenitor cells (HSPCs) from the graft. Surprisingly, we detected donor-derived HSPCs in intestinal mucosa, Peyer's patches, mesenteric lymph nodes, and liver. Human gut HSPCs are phenotypically similar to bone marrow HSPCs and have multilineage differentiation potential in vitro and in vivo. Analysis of circulating post-transplant donor T cells suggests that they undergo selection in recipient lymphoid organs to acquire immune tolerance. Our longitudinal study of human HSPCs carried in intestinal allografts demonstrates their turnover kinetics and gradual replacement of donor-derived HSPCs from a circulating pool. Thus, we have demonstrated the existence of functioning HSPCs in human intestines with implications for promoting tolerance in transplant recipients.

Senda, T., Dogra, P., Granot, T., Furuhashi, K., Snyder, M.E., Carpenter, D.J., Szabo, P.A., Thapa, P., **Miron, M.**, and Farber, D.L. (2018) Microanatomical dissection of human intestinal immunity reveals site-specific changes in gut-associated lymphoid tissues over life. *Mucosal Immunology*. doi: 10.1038/s41385-018-0110-8.

Defining adaptive immunity with the complex structures of the human gastrointestinal (GI) tract over life is essential for understanding immune responses to ingested antigens, commensal and pathogenic microorganisms, and dysfunctions in disease. We present here an analysis of lymphocyte localization and T cell subset composition across the human GI tract including mucosal sites (jejunum, ileum, colon), gut-associated lymphoid tissues (isolated lymphoid follicles (ILFs), Peyer's patches (PPs), appendix), and mesenteric lymph nodes (MLNs) from a total of 68 donors spanning eight decades of life. In pediatric donors, ILFs and PP containing naïve T cells and regulatory T cells (Tregs) are prevalent in the jejunum and ileum, respectively; these decline in frequency with age, contrasting stable frequencies of ILFs and T cell subsets in the colon. In the mucosa, tissue resident memory T cells develop during childhood, and persist in high frequencies into advanced ages, while T cell composition changes with age in GALT and MLN. These spatial and temporal features of human intestinal T cell immunity define signatures that can be used to train predictive machine learning algorithms. Our findings demonstrate an anatomic basis for age-associated alterations in immune responses, and establish a quantitative baseline for intestinal immunity to define disease pathologies.

Kumar, B.V., Kratchmarov, R., **Miron, M.**, Carpenter, D.J., Senda, T., Lerner, H., Friedman, A., Reiner, S.L., and Farber, D.L. (2018) Functional heterogeneity of human tissue-resident memory T cells based on dye efflux capacities. *JCI Insight*. doi: 10.1172/jci.insight.123568

Tissue-resident memory T cells (TRMs) accelerate pathogen clearance through rapid and enhanced functional responses in situ. TRMs are prevalent in diverse anatomic sites throughout the human lifespan, yet their phenotypic and functional diversity has not been fully described. Here, we identify subpopulations of human TRMs based on the ability to efflux fluorescent dyes [efflux(+) TRMs] located within mucosal and lymphoid sites with distinct transcriptional profiles, turnover, and functional capacities. Compared with efflux(-) TRMs, efflux(+) TRMs showed transcriptional and phenotypic features of quiescence including reduced turnover, decreased expression of exhaustion markers, and increased proliferative capacity and signaling in response to homeostatic cytokines. Moreover, upon activation, efflux(+) TRMs secreted lower levels of inflammatory cytokines such as IFN- γ and IL-2 and underwent reduced degranulation. Interestingly, analysis of TRM subsets following activation revealed that both efflux(+) and efflux(-) TRMs undergo extensive transcriptional changes following TCR ligation but retain core TRM transcriptional properties including retention markers, suggesting that TRMs carry out effector function in situ. Overall, our results suggest a model for tissue-resident immunity wherein heterogeneous subsets have differential capacities for longevity and effector function.

Carpenter, D.J., Matsuoka, N., Granot, T., Kumar, B.V., Senda, T., Thome, J.J.C., Gordon, C.L., **Miron, M.**, Weiner, J., Lerner, H., Friedman, A., Griesemer, A.D., Farber, D.L. (2018) Human immunology studies using organ donors: impact of clinical variations on immune parameters in tissues and circulation. *Am. J. Trans.*18:74-88. PMID:PMC5740015

Organ donors are sources of physiologically healthy organs and tissues for life-saving transplantation, and have been recently used for human immunology studies which are typically confined to the sampling of peripheral blood. Donors comprise a diverse population with different causes of death and clinical outcomes during hospitalization, and the effects of such variations on immune parameters in blood and tissues are not known. We present here a coordinate analysis of innate and adaptive immune components in blood, lymphoid (bone marrow, spleen, lymph nodes), and mucosal (lungs, intestines) sites from a population of brain-dead organ donors (2 months-93 years; n = 291) across eight clinical parameters. Overall, the blood of donors exhibited similar monocyte and lymphocyte content and low serum levels of pro-inflammatory cytokines as healthy controls; however, donor blood had increased neutrophils and serum levels of IL-8, IL-6, and MCP-1 which varied with cause of death. In tissues, the frequency and composition of monocytes, neutrophils, B lymphocytes and T cell subsets in lymphoid or mucosal sites did not vary with clinical state, and was similar in donors independent of the extent of clinical complications. Our results reveal that organ donors maintain tissue homeostasis, and are a valuable resource for fundamental studies in human immunology.

Kumar, B.V., Ma, W. **Miron, M.**, Granot, T., Guyer, R.S., Carpenter, D.J., Senda, T., Ho, S.-H., Lerner, H., Friedman, A.L., Shen, Y., and Farber, D.L. (2017) Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. *Cell Reports*. 20:2921-2934. PMID:PMC5646692.

Tissue-resident memory T cells (TRMs) in mice mediate optimal protective immunity to infection and vaccination, while in humans, the existence and properties of TRMs remain unclear. Here, we use a unique human tissue resource to determine whether human tissue memory T cells constitute a distinct subset in diverse mucosal and lymphoid tissues. We identify a core transcriptional profile within the CD69⁺ subset of memory CD4⁺ and CD8⁺ T cells in lung and spleen that is distinct from that of CD69⁻ TEM cells in tissues and circulation and defines human TRMs based on homology to the transcriptional profile of mouse CD8⁺ TRMs. Human TRMs in diverse sites exhibit increased expression of adhesion and inhibitory molecules, produce both pro-inflammatory and regulatory cytokines, and have reduced turnover compared with circulating TEM, suggesting unique adaptations for in situ immunity. Together, our results provide a unifying signature for human TRM and a blueprint for designing tissue-targeted immunotherapies.

Gordon, C.L., **Miron, M.**, Thome, J.J.C., Matsuoka, N., Weiner, J., Rak, M., Igarashi, S., Granot, T., Lerner, H., Goodrum, F. and Farber, D.L. (2017) Tissue-reservoirs of T cell immunity in persistent human CMV infection. *J. Exp. Med.* 214:651-667. PMID: PMC5339671

T cell responses to viruses are initiated and maintained in tissue sites; however, knowledge of human antiviral T cells is largely derived from blood. Cytomegalovirus (CMV) persists in most humans, requires T cell immunity to control, yet tissue immune responses remain undefined. Here, we investigated human CMV-specific T cells, virus persistence and CMV-associated T cell homeostasis in blood, lymphoid, mucosal and secretory tissues of 44 CMV seropositive and 28 seronegative donors. CMV-specific T cells were maintained in distinct distribution patterns, highest in blood, bone marrow (BM), or lymph nodes (LN), with the frequency and function in blood distinct from tissues. CMV genomes were detected predominantly in lung and also in spleen, BM, blood and LN. High frequencies of activated CMV-specific T cells were found in blood and BM samples with low virus detection, whereas in lung, CMV-specific T cells were present along with detectable virus. In LNs, CMV-specific T cells exhibited quiescent phenotypes independent of virus. Overall, T cell differentiation was enhanced in sites of viral persistence with age. Together, our results suggest tissue T cell reservoirs for CMV control shaped by both viral and tissue-intrinsic factors, with global effects on homeostasis of tissue T cells over the lifespan.

Granot, T., Senda, T., Carpenter, D., Matsuoka, N., Weiner, J., Gordon, C.L., **Miron, M.**, Kumar, B.V., Griesemer, A., Ho, S.-H., Lerner, H., Thome, J.J.C., Connors, T., Reizis, B., and Farber, D.L. (2017) Dendritic cells display subset and tissue-specific maturation dynamics over human life. *Immunity*, 46:504-515. PMID:PMC5415308

Maturation and migration to lymph nodes (LNs) constitutes a central paradigm in conventional dendritic cell (cDC) biology but remains poorly defined in humans. Using our organ donor tissue resource, we analyzed cDC subset distribution, maturation, and migration in mucosal tissues (lungs, intestines), associated lymph nodes (LNs), and other lymphoid sites from 78 individuals ranging from less than 1 year to 93 years of age. The distribution of cDC1 (CD141^{hi}CD13^{hi}) and cDC2 (Sirp- α ⁺CD1c⁺) subsets was a function of tissue site and was conserved between donors. We identified cDC2 as the major mature (HLA-DR^{hi}) subset in LNs with the highest frequency in lung-draining LNs. Mature cDC2 in mucosal-draining LNs expressed tissue-specific markers derived from the paired mucosal site, reflecting their tissue-migratory origin. These distribution and maturation patterns were largely maintained throughout life, with site-specific variations. Our findings provide evidence for localized DC tissue surveillance and reveal a lifelong division of labor between DC subsets, with cDC2 functioning as guardians of the mucosa.

Zens, K.D., Chen, J.-K., Guyer, R.S., Wu, F.L., Cvetkovski, F., **Miron, M.** and Farber, D.L. (2017) Reduced generation of lung-tissue resident memory T cells during infancy. *J. Exp. Med.* 214:2915- 2932. PMID:PMC5626403

Infants suffer disproportionately from respiratory infections and generate reduced vaccine responses compared with adults, although the underlying mechanisms remain unclear. In adult mice, lung-localized, tissue-resident memory T cells (TRMs) mediate optimal protection to respiratory pathogens, and we hypothesized that reduced protection in infancy could be due to impaired establishment of lung TRM. Using an infant mouse model, we demonstrate generation of lung-homing, virus-specific T effectors after influenza infection or live-attenuated vaccination, similar to adults. However, infection during infancy generated markedly fewer lung TRMs, and heterosubtypic protection was reduced compared with adults. Impaired TRM establishment was infant-T cell intrinsic, and infant effectors displayed distinct transcriptional profiles enriched for T-bet-regulated genes. Notably, mouse and human infant T cells exhibited increased T-bet expression after activation, and reduction of T-bet levels in infant mice enhanced lung TRM establishment. Our findings reveal that infant T cells are intrinsically programmed for short-term responses, and targeting key regulators could promote long-term, tissue-targeted protection at this critical life stage.

Appendix F. Curriculum Vitae

Education:

Ph.D.
(Anticipated 2019)

Columbia University, New York, NY
PhD in Microbiology, Immunology and Infection
Advisor: Donna L. Farber, PhD.
Thesis: Mechanisms of T cell compartmentalization in human tissues.

B.A.
(2013)

Colgate University, Hamilton, NY
Bachelor of Arts in Biology

Research:

July 2015– Present

Ph.D. Graduate Student

Laboratory of Donna L. Farber, PhD

Department of Microbiology & Immunology, Columbia University

Currently, I am studying mechanisms for tissue compartmentalization of human memory T cells. Memory T cells play an important role in defending against pathogens at peripheral sites; however, most of our knowledge about human memory T cells comes from studies of peripheral blood. We obtain human tissues through a collaboration with the organ procurement agency, and in my project I isolate memory T cells from different tissues and compare these cells to determine defining characteristics of human memory T cells in tissues. My data show that memory T cells in lymph nodes are a distinct memory T cell subset in humans with increased functional capacity and signs of decreased turnover in-vivo. In my second project, I developed a high-throughput sequencing technique to map T cell clones in different lymphoid and mucosal sites as well as blood. I discovered new insights into the relatedness of tissue resident and circulating T cell subsets.

June 2013- Jun2014
& Summer 2012

Post-Baccalaureate Fellow

Laboratory of Donald Court, PhD

I conducted research in the Molecular Control and Genetics Section in Dr. Court's lab, which has developed recombineering as a new way to engineer DNA in living cells. My research utilized recombineering in order to investigate interactions between *E. coli* and phage lambda. A large genomic region of lambda DNA was determined to be responsible for inhibiting several host functions in *E. coli*. Using genetic engineering, I mapped this phenotype to a

region containing a 29 amino acid open reading frame, identifying more precisely the genomic locus involved in a phage function.

Fall 2012

Undergraduate Researcher
Laboratory of Dr. Engda Hagos

I conducted research to assess the role of a transcription factor, KLF2, in DNA damage repair. I utilized mouse embryonic fibroblasts with genomic knock-out (KO) of the KLF2 gene, in order to compared wildtype to KO cell lines and investigate how these cells resonded to DNA damage. Using flow cytometric methods, I discovered alterations in cell cycle stages of WT and KO cells in response to DNA damage.

Publications:

Miron, M. Kumar, B.V., Meng, W., Granot, T., Carpenter, D. J., Senda, T., Chen, D., Rosenfeld, A., Zhang, B., Lerner, H., Friedman, A., Hershberg, U., Shen, Y., Rahman, A., Luning-Prak, E., Farber, D.L. (2018) Human lymph nodes maintain TCF-1hi memory T cells with high functional potential and clonal diversity throughout life. *Journal of Immunology*. 201(7), 2132-2140. doi: 10.4049/jimmunol.1800716.

Miron, M. Meng, W., Rosenfeld, A., Kumar, B.V., Ji, B., Carpenter, D. J., Senda, T., Hershberg, U., Luning-Prak, E., Farber, D.L. (In preparation) Mapping human memory T cell clones reveals clonality and compartmentalization of T cell receptor repertoires.

Miron, M. Thome, J. J., Gordon, C. L., Farber, D.L. (2017). Study of immunosenescence in various tissue compartments. In: Fulop T., Franceschi C., Hirokawa K., Pawlec G (eds) Handbook of Immunosenescence: Basic Understanding and Clinical Applications. Springer, New York.

Kumar, B.V., Kratchmarov, R., **Miron, M.**, Carpenter, D., Senda, T., Lerner, H., Friedman, A., Reiner, S.L., and Farber, D.L. (2018). Functional Heterogeneity of human tissue-resident memory T cell subsets based on dye efflux capacities. *JCI insight*. doi:10.1172/jci.insight.123568.

Takashi, S., Dogra, P., Granot, T., Furuhashi, K., Snyder, M., Carpenter, D., Szabo, P., Thapa, P., **Miron, M.**, Farber, D.L. (2018). Microanatomical dissection of human intestinal immunity reveals site-specific changes in gut-associated lymphoid tissues over life. *Mucosal Immunology*. doi: 10.1038/s41385-018-0110-8.

Fu, J., Zuber, J., Martinez, M, Shonts, B; Obradovic, A; Wang, H., Lau, S., Xia, A., Waffarn, E; Frangaj, K., Savage, T., Simson, M., Yang, S; Guo, X., **Miron, M.**, Senda, T., Rogers, K., Rahman, A., Ho, S., Shen, Y., Griesemer, A., Farber, D.L., Kato, T., Sykes, M. (2018) Human Intestinal Allografts Contain Functional Hematopoietic Stem and Progenitor Cells that Are Maintained by a Circulating Pool. *Cell Stem Cell*. doi: 10.1016/j.stem.2018.11.007.

Kumar, B.V., Ma, W., **Miron, M.**, Granot, T., Guyer, R.S., Carpenter, D.J., Senda, T., Sun, X., Ho, S.H., Lerner, H., *et al.* (2017). Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. *Cell Reports*, 20(12),2921-2934. doi: 10.1016/j.celrep.2017.08.078

Gordon, C. L., **Miron, M.**, Thome, J. J., Matsuoka, N., Weiner, J., Rak, M. A., Igarashi, S., Granot, T., Lerner, H., Goodrum, F., Farber, D. L. (2017). Tissue reservoirs of antiviral T cell immunity in persistent human CMV infection. *J Exp Med*, 214(3), 651-667. doi:10.1084/jem.20160758

Zens, K. D., Chen, J., Guyer, R. S., Wu, F. L., Cvetkovski, F., **Miron, M.**, Farber, D. L. (2017). Reduced generation of lung tissue-resident memory T cells during infancy. *J Exp Med*. 214:(10), 2915-2932. doi: 10.1084/jem.20170521

Carpenter, D. J., Granot, T., Matsuoka, N., Senda, T., Kumar, B.V., Thome, J.J.C., Gordon, C.L., **Miron, M.**, Weiner, J., Connors, T., Lerner, H., Friedman, A., Kato, T., Griesemer, A.D., Farber, D.L. (2017). Human immunology studies using organ donors: impact of clinical variations on immune parameters in tissues and circulation. *Am J Transplant*. 18(1), 74-88. doi:10.1111/ajt.14434

Granot, T., Senda, T., Carpenter, D. J., Matsuoka, N., Weiner, J., Gordon, C.L., **Miron, M.**, Kumar, B.V., Griesemer, A., Ho, S.H., Lerner, H., Thome, J.J., Connors, T., Reizis, B., Farber, D.L. (2017). Dendritic Cells Display Subset and Tissue-Specific Maturation Dynamics over Human Life. *Immunity*, 46(3), 504-515.

Bunin, A., Sisirak, V., Ghosh, H. S., Grajkowska, L. T., Hou, Z. E., **Miron, M.**, Yang, C., Ceribelli, M., Uetani, N., Chaperot, L., Plumas, J., Hendriks, W., Tremblay, M., Hacker, H., Staudt, L., Green, P., Bhagat, G., Reizis, B. (2015). Protein Tyrosine Phosphatase PTPRS Is an Inhibitory Receptor on Human and Murine Plasmacytoid Dendritic Cells. *Immunity*, 43(2), 277-288. doi:10.1016/j.immuni.2015.07.009

Oral and Poster Presentations:

2018 Oral Presentation: **Mapping memory T cell clones in human tissues.**
Biological Data Science Meeting by Cold Spring Harbor Laboratories
2018 Biennial Meeting, Cold Spring Harbor, NY
Abstract Award Recipient

2018 Oral Presentation: **Human lymph nodes maintain tissue-resident memory T cells with high proliferative capacity and clonal diversity.**
Columbia University Women in Science Symposia
Columbia University, New York, NY

2018 Oral and Poster Presentation: **Human lymph nodes maintain memory T cells with high proliferative capacity and clonal diversity.**
Translational Systems Immunology Meeting

Keystone Symposia, Snowbird, Utah

- 2017 Oral Presentation: **Tissue reservoirs of antiviral T cell immunity in persistent human CMV infection**
Federation of Clinical Immunology Societies (FOCIS)
2017 Annual Meeting in Chicago, IL
FCE Fusion Award Recipient
- 2017 Oral Presentation: **T cell responses to human CMV infection.**
Columbia Center of Translational Immunology (CCTI) Seminar Series
Columbia University, New York, NY

Other Relevant Experience:

2017-Present

Team Lead

Data Science Club, Columbia University Medical Center

I am a member and team leader of the Data Science Club, a student led organization on campus. Club members meet weekly in order to conduct code review sessions, discuss literature from recent research, as well as complete coursework. In addition to weekly meetings, I participated in monthly hackathons where individuals brought projects to work on independently or in groups. I also participated in two NIH hackathons. During one, my team worked on integrating NIH genotype and phenotype databases with other sequence archives in order to facilitate easier searching. In a second hackathon I was the team lead to create software for analysis of T cell receptor sequencing data, denoted PyClonal.

2017-2018

Science Alliance Leadership Training (SALT)

The New York Academy of Science, New York, NY

I participated in a week long in person leadership, communication skills, and team dynamics training program held at the New York Academy of Sciences. I was accepted to this program through a selective application process with 18 other participants from Universities across the US, mostly from the east-coast area. I also participated in 40 hours of follow-up training after the week long in-person workshops by webcasted courses and seminars.

2015-2018

Mentor

Girls Science Day, Columbia University

I led and planned experiments for three sessions of middle school aged girls (about 80 total students). We conducted color chromatography experiments, separating different dyes used in foods.