UNIVERSITY OF CALIFORNIA SAN DIEGO

Investigating transcriptional regulators of memory helper T cells

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Biology

by

Tianda Deng

Committee in charge:

Professor Stephen Hedrick, Chair
Professor Jack Bui
Professor Ananda Goldrath
Professor Enfu Hui
Professor Susan Kaech
Professor Li-fan Lu

2021

The Dissertation of Tianda Deng is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

I dedicate this dissertation to loving my mother, whose sacrifice made this achievement a possibility.

EPIGRAPH

Science knows no country, because knowledge belongs to humanity,
and is the torch which illuminates the world.

Louis Pasteur

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge my mentors Stephen Hedrick and Ananda Goldrath. I have learned so much through their guidance and nurturing, both in- and outside of the laboratory. Without their support, this could not have been possible. I want to thank Ananda for the countless delicious Regents pizzas throughout the years! A huge thanks to the members of the Goldrath and Hedrick labs, both past and present. I want to thank Kyla especially for all her guidance during my PhD. This was a fruitful yet challenging journey and I definitely could not have done it without my fellow lab members. I also want to acknowledge all my committee members for their guidance and input throughout this journey. It has been an absolute pleasure learning from all of you.

Chapter 1, in part, is a reprint of the material as it appears in Immunology. Nguyen, Q. P., Deng, T. Z., Witherden, D. A., and Goldrath, A. W. *Origins of CD4$^+$ circulating and tissue-resident memory t-cells.* Immunology 157, (2019), 3–12. The dissertation author was a primary author of this paper.

Chapter 2, in full, has been submitted for publication of the material as it may appear in PNAS, 2021. Shaw, L. A., Deng, T. Z., Omilusik, K. D., Nguyen, Q.P., and Goldrath, A. W. *Id3 expression identifies mouse CD4$^+$ memory T_H1 cells.* The dissertation author was a primary investigator and a first author of this paper.

Chapter 3, in part, is currently being prepared for submission for publication of the material. Deng, T. Z., Goldrath A. W. The dissertation author was a primary investigator and author of this material.

VITA

2011-2015	B.S. in Biomechanical Engineering, Stanford University
2015–2016	Staff Research Associate, Stanford University.
2017–2020	Teaching Assistant, Division of Biological Sciences University of California, San Diego
2016–2021	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

Loh K. M, Chen A., Koh P. W., **Deng T.Z.**, et al. *Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types.* Cell 166, (2016), 451-467.

Nguyen, Q. P., **Deng, T. Z.**, Witherden, D. A., and Goldrath, A. W. *Origins of CD4$^+$ circulating and tissue-resident memory t-cells.* Immunology 157, (2019), 3–12.

ABSTRACT OF THE DISSERTATION

Investigating transcriptional regulators of memory helper T cells

by

Tianda Deng

Doctor of Philosophy in Biology

University of California San Diego, 2021

Professor Stephen Hedrick, Chair

Upon infection, naive T cells proliferate and differentiate into highly specialized effector cells to combat the invading pathogen. Naive $CD4^+$ T cells have the potential to differentiate into multiple functionally distinct T helper (T_H) subsets based on the type of infection. Differing pathogens elicit distinct infection milieux that help direct the differentiation of naive $CD4^+$ T cells to ensure that each class of pathogen is countered with the appropriate immune response. The majority of effector T cells will die as the infection wanes, while a small proportion of cells will survive to established a long-lived memory population. This memory population is essential for improved antibody responses and also enables a rapid and robust secondary response against recurring pathogens, thus conferring lasting cellular immunity.

However due to the functional breadth of $CD4^+$ T cell lineages, the identification of a conserved memory $CD4^+$ T cell precursor and memory population across T_H lineages has proved challenging. Lack of such knowledge impedes the ability to investigate conserved mechanisms of memory $CD4^+$ T cell formation and regulation. To better understand the biology of $CD4^+$ memory T cells, we sought to identify a conserved marker of memory $CD4^+$ T cells across different T_H subsets. Utilizing fluorescent reporter mice, we found that expression of Id3, an inhibitor of E protein transcription factors, identified a population of cells within both the $CD4^+$ T_{FH} and T_H1 helper lineages that exhibited memory potential in response to secondary infection. Notably, a subset of T_H1 memory cells expressing Id3 exhibited enhanced expansion upon response to pathogen, generating both T_H1 and T_{FH} secondary effector cell populations, and displayed enrichment of key molecules associated with memory potential when compared to $Id3^{lo}$ T_H1 cells. Relative to $Id3^{lo}$ T_H1 memory cells, $Id3^{hi}$ T_H1 cells exhibited a transcriptomic profile more akin to that of memory T lymphocytes. Thus, we found that Id3 expression serves as an important marker of multipotent memory $CD4^+$ T cells.

To investigate novel regulators of $CD4^+$ memory T cells, we took a computational approach by using Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-seq) and bulk RNA sequencing (RNA-seq) of effector and memory $CD4^+$ T cell populations. We leveraged the PageRank algorithm to first predict putative regulators based on changes in transcriptomic expression as well as chromatin accessibility between effector and memory $CD4^+$ T cells. Validation of predicted targets utilized the electroporation of CRISPR/Cas9 RNP complex to achieve loss-of-function disruptions of target genes in $CD4^+$ T cell prior to LCMV-Armstrong infection. Although initial testing of predicted targets Srebf2 and Rorb did not reveal significant effects in $CD4^+$ memory T cell formation, the optimization of the CRISPR/Cas9 RNP system has provided an efficient and reliable method for gene-disruption in T cells that undoubtedly expands our ability to investigate T cell biology. Future experiments utilizing this workflow have the potential to identify conserved regulators of $CD4^+$ precursor and memory T cell populations across T_H lineages, shedding light on possible mechanism for $CD4^+$ T cell

memory formation.

Chapter 1

Introduction

The invention of vaccines is undoubtedly one of the greatest achievements in human history. However despite this incredible innovation, infectious diseases still remain one of the leading causes of death globally. In 2019 the world was ambushed by the SARS-Cov2 virus, ultimately leading to the COVID-19 pandemic. Aside from mandatory facial barriers, social distancing policies were implemented forcing countries all over the world into a lock-down state. The world ground to a halt while financial and social consequences continued to accrue. Despite setbacks, humanity was able to manufacture a safe vaccine against the virus at an unprecedented pace of just one year. Our ability to rapidly respond to an emergency public health crisis relies heavily on our scientific understanding of the immune system, the combined product of decades of dedicated research. The COVID-19 pandemic perfectly exemplifies the tangible benefits of the scientific endeavor and more importantly highlights the necessity of our unrelenting pursuit for understanding the unknown.

1.1 The T cell lineage

Our body's immune system is a complex network of cells and proteins that are constantly operating in unison to combat against invading harmful pathogens - bacteria, fungi, viruses or other disease-causing agents. The first line of defense for our immune system is simply an outer covering, such as skin which prevents pathogens from entering the body. However sealing off the

entire body surface is impossible due to gas exchange, nutrition and reproduction requirements. The body's entrance and exits are guarded by secretions that trap or kill microbes, while epithelial linings of the airway, digestive track and other exchange surfaces provide additional barriers against infection. However, once a pathogen gains entry into the body, the strategy for its elimination changes vastly.

After breaching barrier defenses, the pathogen is no longer an outsider, therefore the immune system must be able to detect foreign particles and cells within the body. Such an achievement is contingent upon the body's ability to distinguish self from non-self. This is accomplished via molecular recognition between the binding of immune cell surface receptors and molecules from foreign pathogens. This specific binding of immune receptors to foreign particles is the principal event in identifying non-self molecules from self. There are two main types of molecular recognition which are central to two different types of immune defenses: innate immunity, which is found in all animals, and adaptive immunity, which is found only in vertebrate species.

Innate immunity, which encompasses the body's barrier defenses, utilizes a small set of surface protein receptors that recognize molecules or structures that are absent from the body but are commonly expressed by groups of viruses, bacteria or other pathogens. This recognition event triggers the innate immune responses that eliminate a very broad range of pathogens. Immune cells directly involved in the innate response include macrophages, natural killer (NK) cells, dendritic cells, neutrophils, eosinophils, and basophils. Adaptive immunity utilizes a much broader range of receptors, and each of which recognizes a specific feature found only on a certain part of a certain molecule in a certain pathogen. This allows adaptive immune cells to achieve incredible recognition specificity against invading pathogens. However the adaptive immune response is typically activated after the innate immune response and develops more slowly. Adaptive immune cells include B and T lymphocytes as well as natural killer T cells. The T cell lineage can be further subcategorized into helper and killer T cells based on the surface expression of CD4 or CD8 respectively.

B cells are a lineage of lymphocytes that are responsible for the production of antibodies in humoral immunity. Progenitor B cells arise in the bone marrow, similar to T cell progenitors. However B cells mature in the bone marrow while T cell progenitors emigrate into thymus for maturation. B cells utilize B cell receptors (BCRs) on their cell surface to specifically bind a wide range of antigens. Upon activation, B cells proliferate and can differentiate into antibody-secreting cells or memory B cells that release antibodies which can elicit a wide range of immune function based on antibody type. Natural killer T (NKT) cells are a specialized hybrid between innate and adaptive immune cells. They are an innate-like T cell subset that typically express an invariant T cell receptor α-chain which recognizes lipids presented on the surface protein CD1d. Upon activation, NKT cells can directly eliminate target cells or influence innate and adaptive immune cells to indirectly kill infected cells. These NKT cells differentiate in the thymus into distinct subsets that are highly analogous to subsets in $CD4^+$ helper T cells, affording them a wide range of possible immune functions.

1.1.1 $CD4^+$ Helper T cells (The Helpers)

Our immune system is continuously challenged by divergent threats ranging from viruses to bacteria to fungi, which require different strategies to limit and eliminate. Therefore, diversity amongst invading pathogens necessitates flexibility within the immune response. Within the adaptive branch of the immune system, $CD4^+$ T cells can best exemplify this immune adaptability. In response to invading pathogens, naive $CD4^+$ T cells proliferate and have the potential to differentiate into at least seven functionally distinct T helper (T_H) subsets with unique effector functions [95]. These T_H subsets play pivotal roles in adaptive immunity but can also influence the innate immune response [132]. Depending on the type of immunological threat, early host-pathogen interactions will result in an infection milieu that directs differentiation of naive $CD4^+$ T cells to acquire the helper functions to ensure that each class of pathogen is countered with the appropriate immune response [155].

Naive $CD4^+$ T cells originate as common lymphoid progenitors from the bone marrow,

which then undergo development in the thymus to finally execute their immune functions in peripheral tissues as well as secondary lymphoid organs (SLOs) such as the spleen, lymph nodes and Peyer's patches [77]. The thymic selection process ensures that $CD4^+$ T cells can efficiently recognize foreign antigens without eliciting an autoimmune response against self-antigens. To become an activated T cell, naive $CD4^+$ T cells are first activated in lymphatic tissues by professional antigen presenting cells (APCs). APCs will present pathogenic antigens on major histocompatibility complex (MHC) class II molecules, while also providing a costimulatory signal to fully activated the naive T_H cell. This activation event will drive the T_H cell to differentiate into distinct T effector (T_{eff}) populations based on the nature of the infection. $CD4^+$ T_{eff} cells mediate defensive mechanisms against invading pathogens through the production of effector cytokines, such as interferon (IFN) and interleukins (ILs), that serve a myriad of functions including: activating innate (macrophages) and adaptive ($CD8^+$ T cells) immune cells to kill infected cells as well as helping B cells undergo class-switching and somatic hypermutation to produce high-affinity antibodies for mediating humoral immune responses [3, 155, 116].

Based on the expression of signature T_H cytokines as well as lineage-specifying master transcriptional factors (TFs), circulatory $CD4^+$ T_H cells can be broadly categorized into five major lineages: T_H1 (IFN-γ and T-bet) T_H2 (IL-4 and GATA3), T_H17 (IL-17 and RORγt), T_{FH} (T Follicular Helper)(IL-21 and Bcl6), and T_{reg} (T regulatory)(IL-10 and Foxp3) (Figure 1.1) [37]. Other T_H subsets have also been proposed including T_H3 [20], T_H9 [30, 139], T_H22 [36], and T_R1 [48]. However the underlying biology behind these subsets are not as well established and thus will not be discussed further in this dissertation. Among $CD4^+$ Helper T cell subsets, the T_H1 lineage mediates type 1 responses to protect against intracellular pathogens such as viruses, bacteria, and protozoa via the activation of type 1 macrophages as well as other immune cells [155]. The T_H2 lineage "coincidentally" is responsible for eliciting type 2 responses against infections such as helminth parasites. T_H2 cells do so by the activating type 2 macrophages as well as recruiting basophils, eosinophils, and mast cells to the infection site [49]. The "rebellious" T_H17 lineage induces type 3 responses against extracellular bacteria and fungi via induction of

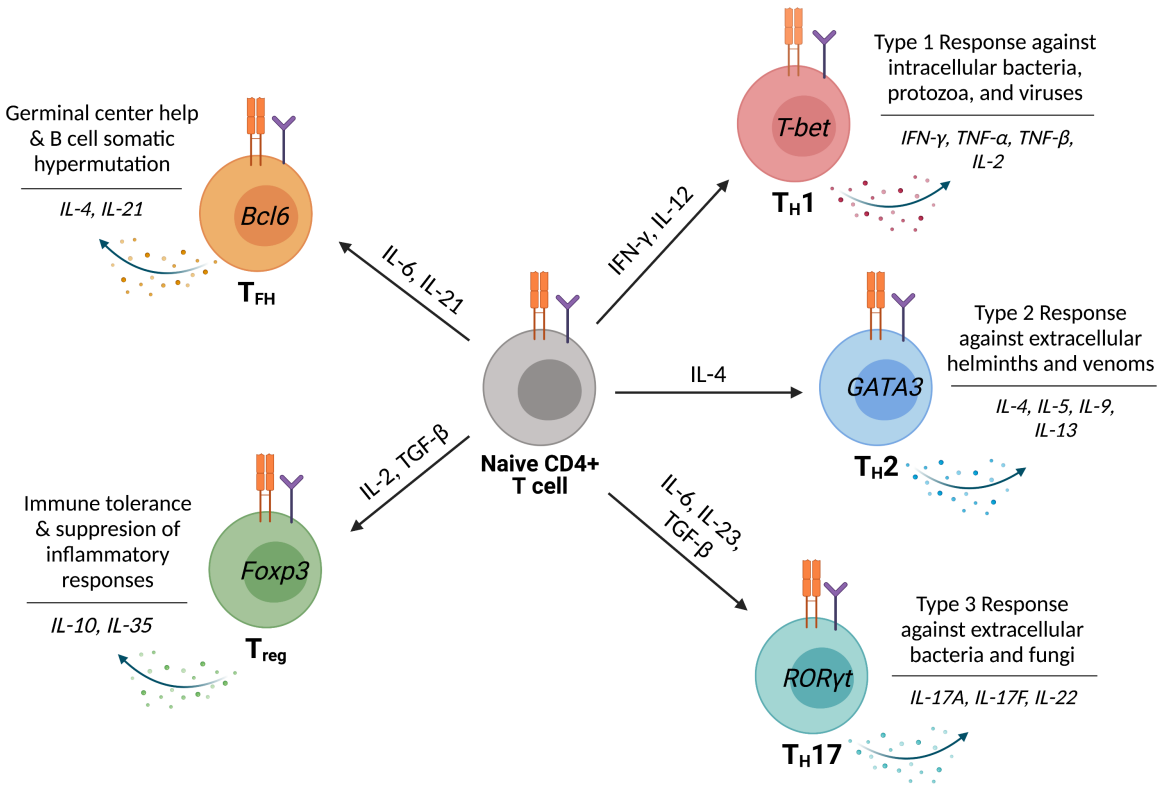


Figure 1.1. CD4$^+$ helper T cell subsets. Differentiation of naive CD4$^+$ T cell into distinct T_H subsets is controlled by key transcription factors and cytokines.

antimicrobial peptide production by epithelial barrier tissues as well as neutrophil recruitment [129]. The T_{FH} lineage are vital to supporting B cells in antibody responses. T_{FH} cells promote germinal center formation, antibody affinity maturation, and immunoglobulin class switching [28]. Most licensed human vaccines work on the basis of long-term protective antibody responses, therefore T_{FH} cells are essential mediators of vaccine-based protective immunity. Lastly the T_{reg} lineage is critical for ensuring immune tolerance and preventing autoimmune illnesses [33].

1.1.2 CD8$^+$ Cytotoxic T cells (The Killers)

The other subset of T lymphocytes is the CD8$^+$ T cell, which is required for eradicating intracellular pathogens and plays a crucial role in the clearance of infections as well as tumors [17, 89]. During an infection, naive CD8$^+$ T cells encounter pathogenic peptides presented in

the contest of MHC class I molecules presented by APCs in the secondary lymphoid tissues. T cell receptor (TCR) binding along with co-stimulation will activate the $CD8^+$ T cell to rapidly proliferate and differentiate into a potent cytotoxic T cell (T_C), which then migrates to the sites of infection [67, 63]. These T_C cells then produce effector cytokines (i.e. IFNγ and TNFα) and can release cytolytic molecules (i.e. perforin and granzymes) or utilize FAS ligands binding to eliminate infected cells and clear the infection. This expanded population of effector $CD8^+$ T cells is a heterogenous population composed of terminal effector (TE) T cells (identified by high cell surface expression of killer cell lectin-like receptor G1 (KLRG1) and low levels of Interleukin-7 receptor (IL-7Rα, CD127)) and memory precursor (MP) T cells (identified by $KLRG1^{lo}CD127^{hi}$) [67].

1.2 The importance of memory

Once the $CD4^+$ and $CD8^+$ T_{eff} cells eliminate the pathogen, the majority of the effector cells die via apoptosis during contraction phase, while a small portion persists and differentiates into long-lived memory cells. This memory population enables a rapid and robust secondary response against recurring pathogens and, thus, is pivotal in conferring lasting cellular immunity, particularly against pathogens where neutralizing antibodies alone are insufficient at providing long-term protection [107]. Memory T cells are able to respond faster and more efficiently against a recurring pathogen due to several characteristics. Firstly, memory T cells can have a lower threshold of TCR activation compared to naive T cells, and also are less dependent on co-stimulation [86]. A lower activation requirement allows the memory T cells to proliferate and generate robust T_{eff} faster compared to a primary T cell response [98, 21]. It is worth noting that much like the effector response of a primary infection, the secondary response by T_{eff} cells is also dependent on the context of their infection based on cytokine signals in the microenvironment [10]. Secondly, the sheer number of memory T cells are much higher than those of naive T cells during a primary infection [12, 38]. The increased frequency of pathogen recognizing T cells

raises the likelihood of encounter with a potentially recurring pathogens, thus enabling a more rapid response [85, 84]. Lastly, unlike naive T lymphocytes, which only circulate between SLOs and blood, memory T lymphocytes can survey the lymphatic, blood and peripheral tissues (i.e. gut, lungs, or skin) [145, 100].

Memory T cell subsets

Memory T cells within the circulation have been conventionally divided into central memory (T_{CM}) cells, which circulate between the blood and SLO, and effector memory (T_{EM}) cells, which can migrate from the blood into non-lymphoid tissues [95, 86, 119]. T_{EM} are defined by low expression of L-selectin (CD62L), a cell adhesion molecule and C-C chemokine receptor type 7 (CCR7), a cell-surface receptor. T_{EM} cells have access to non-lymphoid sites and the ability to produce effector cytokines within hours following TCR stimulation. T_{CM} cells are characterized by high levels of CD62L and CCR7, and the ability to recirculate through lymph nodes, secrete IL-2 upon reactivation and undergo significant proliferation to generate secondary effector T cells [113]. Recently a new subset of non-circulating and tissue-homing memory T cells was identified and these tissue resident memory T cells were termed T_{RM}. T_{RM} cells migrate to specific peripheral tissues like the skin, liver, guts, and lungs to take up permanent residency. These T_{RM} cells are strategically positioned at many barrier sites such as mucosal linings and the skin, areas of the body where pathogens are most likely to invade, to increase the likelihood of pathogen detection and elimination [107]. Memory T cell subsets utilizes these unique characteristics that aid them in the long-term protection from harmful recurring pathogens.

1.2.1 Memory differentiation cues

TCR signaling

T cell receptor signaling strength is a crucial factor in memory T cell development. This TCR binding strength is collectively determined by the affinity of the receptor to the antigen

and MHC molecule complex, the density of the antigen complex presented by the APCs, and the duration of the binding event [8, 27]. Based on varying binding strengths, different T cells will receive differing degrees of biochemical pathway activation, resulting in changes in their transcriptomic landscape to either favor or deter memory T cell formation [76, 29]. These transcriptomic changes involve several key TFs including Bcl-6, Blimp-1, Eomes, and T-bet [29] as well as the upstream regulators nuclear factor of activated T-cells (NFAT) and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) [113]. Alternatively, transcriptional changes can also be achieved directly via the binding of cytokines to cytokine receptors like IL-2 and IL-2 receptor alpha chain (CD25), which can synergize with TCR signaling to induce memory T cell development [127, 114]. For memory $CD8^+$ T cells, TCR signaling strength seems to be inversely correlated with memory cell development [29].

Much like the signals important for $CD8^+$ memory T cell generation, strengths of TCR and co-stimulatory signalling also have profound effects on memory T_H development [31, 43]. Recent results from Snook et al. [127] demonstrated that TCR signalling has a direct impact on T_H memory formation. Utilizing a panel of TCRs specific for the same viral antigen, the authors showed substantial variability in TCR signal strength, expression of CD25 and activation of downstream TFs across the $CD4^+$ memory T-cell population [127]. TCR clones with stronger TCR signalling appear to differentiate towards a more TE-like state and become largely depleted by memory time points, while clones with comparatively lower signalling were memory-like and able to persist after antigen clearance. Interestingly, it seems that stronger TCR signalling was associated with higher expression of T_H1 markers, while weaker TCR signals correlated with higher expression of T_{FH} markers [127], suggesting that there may be a connection between lineage differentiation and memory potential for $CD4^+$ helper T cells.

Interleukin-2

Utilizing influenza A virus (IAV) as an infection model, McKinstry et al. [94] showed that IL-2 is crucial at a late checkpoint for effector helper T cells to survive the contraction

phase, allowing for the transition into resting memory T cells. To circumvent defects in initial T-cell priming caused by IL-2 deficiency, the authors first activated $CD4^+$ T-cells *in vitro* with exogenous IL-2 and then transferred these cells into naive mice for infection. Following IAV challenge, both *in vitro* primed wild type and IL-2-deficient donors showed similar cell numbers at the peak of infection and production of IFNγ; however, the IL-2-deficient population quickly declined and was undetectable by day 28 of infection. Exogenous administration of IL-2 during days 5–7 of infection successfully restored memory cell numbers for IL-2-deficient $CD4^+$ T cells, demonstrating the importance of IL-2 for $CD4^+$ memory T cell generation in this context. Furthermore, a recent study by DiToro et al. [32] with Listeria monocytogenes (LM) infection showed that as early as 20 hours after antigen exposure *in vivo*, IL-2 production in $CD4^+$ T_H effector cells strongly correlated with T_H fate differentiation during infection, again supporting a link between lineage specification and memory formation. To further highlight the importance of IL-2 in T_H memory, Shakya et al. [122] identified a role for TF Oct1 and its co-activator OCA-B in poising the Il2 locus for robust expression in memory $CD4^+$ T cells, unveiling an important mechanism by which memory $CD4^+$ T cells control IL-2 production. However, these studies regarding TCR signalling and IL-2 in $CD4^+$ memory T cells were done without investigation of specific T_H lineages. Therefore, further investigations into the required transcriptional and epigenetic regulation for generation and maintenance of memory T_H subsets are still needed.

1.2.2 Transcriptional regulation of memory T cells

As alluded to in the previous section, transcription factors serves as central arbiters for the cell-fate decisions between subsets of both effector as well as memory T cells [17]. These TFs can drastically influence gene expression changes as well as alter the epigenetic landscape of the cell's genome [137, 141, 151]. Due to its pertinence in human health, the transcriptional circuitry underlying CD8 memory formation has been under fervent investigation and has yielded many fruitful insights into $CD8^+$ T-cell differentiation. An overarching theme has emerged from these studies where cell fate decisions, between terminal-effector and memory, seem to be directed by

an equipoise between pairs of opposing transcription factors [117, 124, 69, 58, 15, 59, 64, 148]. As an example, increased expression of the transcription factor T-bet foster the differentiation of effector $CD8^+$ T cells towards the terminally differentiated $KLRG^{hi}$ fate [61, 65]. The expression level of T-bet positively correlates with the amount of inflammatory signals like IL-12 [65]. This correlation brilliantly provides the immune system the ability to balance the potential severity of the infection to the amount of effector T cells needed to combat the infection [4, 5].

Genetic loss of T-bet results in the loss of the $KLRG^{hi}$ subset of $CD8^+$ T_{eff} cells but minor effects for the $CD127^{hi}$ subset [65]. The TF T-bet is highly homologous to another transcription factor Eomesodermin (Eomes), however Eomes seems to foster the development of memory $CD8^+$ T cells rather than the $KLRG^{hi}$ $CD8^+$ effector T cells [61]. Genetic deletion of Eomes leads to a modest effect on the effector T cell pool but results in the attrition of memory T cells [7] as well as the inability to form self-renewing T_{CM} cells [110]. Taken together, this pair of two related transcription factors directs the differentiation of $CD8^+$ T cell towards the terminal effector fate versus the memory population. It is worth noting here that the amount of T-bet and Eomes is important as a gradient effect clearly exists, where the increase of T-bet fosters the increase in terminal effector differentiation [65].

Compared to memory $CD8^+$ T cells, much less is known about the transcriptional regulators of $CD4^+$ memory T cells. One investigation has highlighted the importance of transcription factor Thpok, the master regulator of $CD4^+$ T cell maturation in the thymus, in the generation of $CD4^+$ memory T cells [24]. The authors report that deletion of Thpok induces an exhaustion-like program in $CD4^+$ T cells which results in the loss off protective $CD4^+$ memory responses [24].

Transcriptional inhibitors: the Id proteins

A similar parallel can also be observed in the expression of transcriptional inhibitor proteins: Inhibitors of DNA-binding, Id2 and Id3. The Id family of proteins are a group of transcriptional inhibitors that specifically bind to E protein transcription factor family members

[91, 9]. The E protein TF family contains transcription factors in the basic helix-loop-helix (bHLH) family which control numerous aspects of T cell biology including controlling the TCR rearrangement, progression, survival, and proliferation of T cell progenitors [103]. E proteins can interact as homo- or hetero-dimers via the HLH domains and bind DNA at E-box-consensus sequences to either activate or suppress transcriptional activation [70]. However, the ability of E proteins to bind DNA is inhibited by a family of similar proteins called the Id proteins, which share the HLH domain while lacking a DNA-binding domain. This means that Id proteins will form heterodimers with E proteins via the shared HLH domain but since they lack the DNA-binding domain, this E and Id protein heterodimer will be unable to bind DNA and thus inhibiting E protein function [9, 91]. The aforementioned Id2 and Id3 are Id protein family members shown to be relevant in regulation T cell differentiation [91, 103]

Genetic deletion of Id3 in $CD8^+$ T cells results in the ablation of long-lived $CD8^+$ memory T cells. Genetic knock-out of Id2 in $CD8^+$ T cells leads to a defect in the generation of terminal effector T cells [148]. Much like T-bet and Eomes, Id2 and Id3 are related and share many common characteristics (i.e. suppressing E protein function) yet there is also a clear distinction in their roles in $CD8^+$ T cell differentiation. Using fluorescent protein knock-in reporter mice (Id2-YFP and Id3-GFP), our lab identified $CD8^+$ T_{eff} cells with high expression of Id3-GFP and intermediate levels of Id2-YFP (Id3-GFP^{hi}Id2-YFP^{int}) preferentially differentiated into the $KLRG^{lo}CD127^{hi}$ memory precursor subset which survive longer and respond better to secondary challenge compared to the Id3-GFP^{lo}Id2-YFP^{hi} [148]. Interestingly, the Id3-GFP^{hi} effector $CD8^+$ T cells were more similar in transcriptional profile to long-lived memory cells, even before the surface expression of known $CD8^+$ memory markers, hence making them memory precursors.

While not much is known about the role of Id proteins in memory $CD4^+$ T cells, their functions in effector $CD4^+$ T cells have been extensively studied [123]. An report by Shaw et al. shows that T_H1 effector cells have high expression of Id2 while T_{FH} effector cells express Id3 at a high level [123]. Id2 deficiency in $CD4^+$ T cells results in an increase of T_{FH} effector cells while a reduction in T_H1 effector cells, supporting a regulatory role for Id2 in T_H1 effector cell

differentiation. Further investigations are need to understand the roles of Id proteins in memory $CD4^+$ T cell differentiation.

1.3 Memory Helper T cells

While significant advances have been made in understanding the generation and maintenance of memory $CD8^+$ T cells, the molecular mechanisms underlying the generation of memory $CD4^+$ T cells remain relatively elusive. Two major obstacles have contributed to this knowledge deficit. First, $CD4^+$ T cells are inherently less proliferative and the $CD4^+$ memory T cell population appears to decline following antigen clearance, while the $CD8^+$ memory T cell population, if established, is typically stable [40, 143, 1] and thus, fewer cells are available for study. Second, the existence of functionally distinct effector T_H cell subsets hinders the ability to characterize a common $CD4^+$ memory T cell precursor. Further, T_H effector and memory T cells also exhibit significant plasticity and can interconvert between lineages, both *in vivo* and *in vitro*, adding an additional layer of complexity to identifying memory precursor cells in $CD4^+$ T memory studies [152, 54, 105, 16, 147, 86].

In spite of the challenges in studying memory $CD4^+$ T cells, efforts in recent years focusing on different aspects of memory development have begun to elucidate a more comprehensive picture for the generation and maintenance of memory $CD4^+$ T cells. In this section, we look at recent studies addressing the identity of memory $CD4^+$ T cell populations and their precursors in both the periphery and non-lymphoid tissues.

$CD4^+$ T cell memory in secondary lymphoid organs

Despite clear differences between memory $CD4^+$ and $CD8^+$ T cell populations, including the range of effector cell heterogeneity [40], the models for memory $CD8^+$ T cell formation have served as a useful framework for the investigation of memory $CD4^+$ T cells. As mentioned previously, during the primary response of antigen-specific cytotoxic $CD8^+$ T cells (CTL), two effector $CD8^+$ T cell populations can be identified based on surface expression of KLRG1 and

CD127 [66]. The KLRG1^{hi}CD127^{lo} terminal effector population is predominantly lost during the contraction phase, while the KLRG1^{lo}CD127^{hi} subset contains memory precursor cells, which can differentiate into long-lived memory CD8$^+$ T cells [66]. CD4$^+$ T cells also express KLRG1 [11] and CD127 [71]. However, the roles of these molecules in memory CD4$^+$ populations are not well established nor are clear strategies for distinguishing shorter-lived effector and precursors of memory T_H populations.

Evidence for long-lived CD4$^+$ memory T cells capable of responding to pathogen re-challenge has been documented in studies of adoptive transfer of T-cell receptor transgenic T cells [52, 112, 88, 23] and endogenous immune responses [112]. However, the diversity of functional T_H phenotypes has made identification of distinct CD4$^+$ TE and MP effector populations challenging. Additionally, it is unclear whether all CD4$^+$ T_H effector T cells possess the same potential to differentiate into long-lived memory cells. A separate MP may exist for each subset or there may be a unique effector subset with an inherent memory program that can give rise to memory populations with the potential to generate T_H subsets with all or some effector functions (T_H1, T_H2, T_H17, T_{FH}, T_{reg}) in a secondary infection. An elegant study by Tubo et al. [134] addressed this issue by following the differentiation of individual CD4$^+$ T cells responding to infection. Utilizing over 80 distinct TCR clones that can specifically respond to Listeria monocytogenes infection, they demonstrated that all microbe-specific naive CD4$^+$ T cells have the potential to give rise to memory cells following acute infection [134]. Different individual naive CD4$^+$ T cells generated antigen-specific effector populations with varying frequencies of T_H1 and T_{FH} effector cells. Notably, the relative frequencies of these subsets were preserved into the memory phase, suggesting that both T_H1 and T_{FH} effector populations contain precursors of memory cells that retain their effector T_H characteristics (Figure 1.2). These data favour the idea that some CD4$^+$ memory cells are relatively lineage-committed; however, a range of expansion potential and plasticity among progeny was also observed, suggesting that not all CD4$^+$ memory precursor cells may be equivalent.

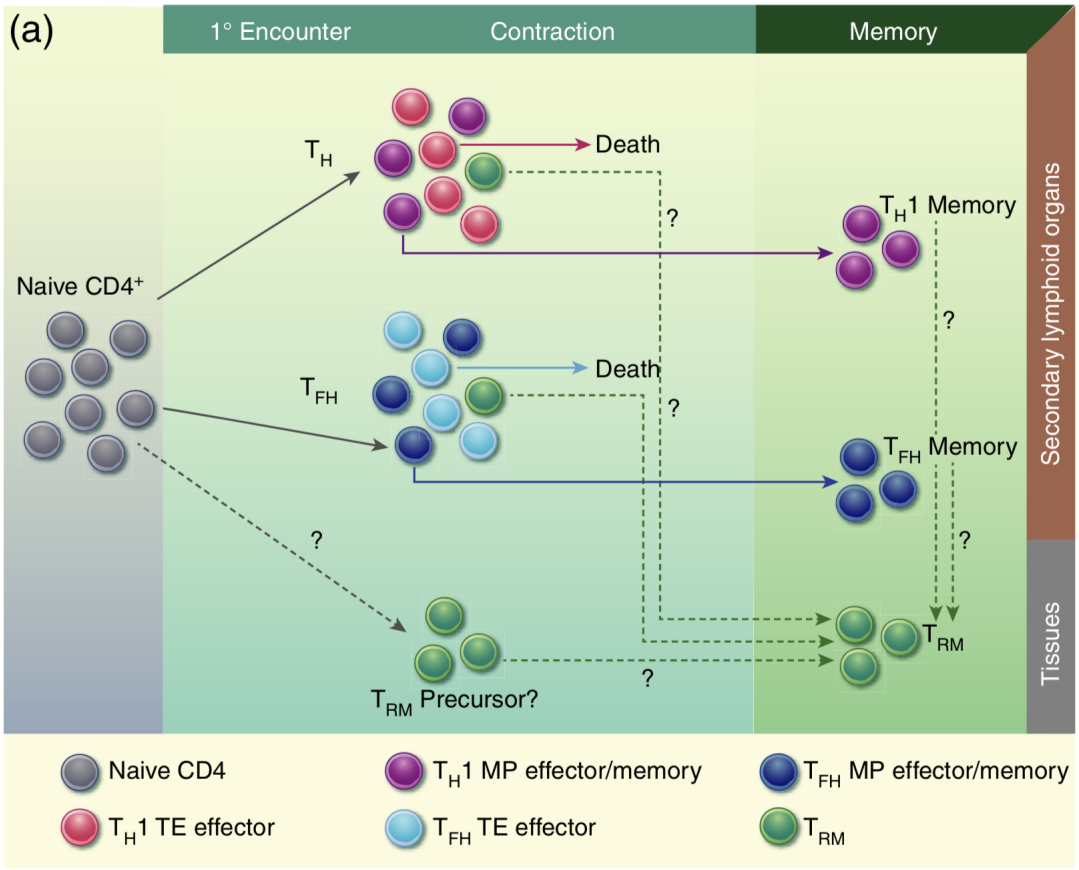


Figure 1.2. A model of CD4$^+$ memory T cell formation. Upon antigen encounter, naive CD4$^+$ T cells differentiate into effector subsets based on the type of infection. Within each effector CD4$^+$ subset, there potentially exist terminal effector (TE) and memory precursor (MP) effector cells. The majority of TEs die during the contraction, while MPs can survive and transition into resting memory cells. CD4$^+$ tissue-resident memory cells (T_{RM}) may differentiate from: (1) the naive subset; (2) MP cells within the effector population; or (3) committed memory cells.

1.3.1 T_H1 and T_{FH} CD4$^+$ memory T cells

In efforts to address these questions, several groups have used lymphocytic choriomeningitis virus (LCMV) to characterize the response of adoptively transferred SMARTA (SM) cells, which have transgenic expression of an MHC Class II-restricted T-cell antigen receptor (TCR) specific for LCMV glycoprotein amino acids 66–77 [88, 23, 51, 83]. Meanwhile, other investigators have studied the endogenous polyclonal response by utilizing the peptide-loaded major histocompatibility complex class II (pMHCII) tetramer-based approach to identify antigen-specific CD4$^+$ T-cells [52, 112, 23]. During acute infection with LCMV-Armstrong,

antigen-specific $CD4^+$ T cells differentiate into two main helper subtypes in the spleen and lymph nodes: T_H1 and T_{FH}. T_H1 cells express the transcriptional regulator T-bet and are known for secreting their signature effector molecule IFNγ, while T_{FH} cells express Bcl6 and their hallmark surface molecule C-X-C chemokine receptor type 5 (CXCR5), which allows for homing to germinal centers to support B cell responses. To explore the origins of T_H1 and T_{FH} memory cells, investigators utilized fluorescence-activated cell sorting to isolate T_H1 and T_{FH} effector and memory cells based on known markers, and studied their characteristics in the context of reinfection [71, 88, 23, 134].

Marshall et al. [88] found that within the primary effector populations from the spleen at day 8, two $CD4^+$ T cell subsets that resembled the $CD8^+$ TE and MP T cells were observed. The TE like population was marked by high expression of both P-selectin glycoprotein ligand-1 (PSGL-1) and lymphocyte antigen 6 complex (Ly6C), while the MP-like effector cells were PSGL-1^{hi}Ly6C^{lo}. In contrast to the PSGL-1^{hi}Ly6C^{hi} cells, the PSGL-1^{hi}Ly6C^{lo} MP-like population exhibited greater longevity in uninfected hosts, increased proliferation following antigen re-challenge, and similar gene-expression profiles with day 60 PSGL-1^{hi} memory $CD4^+$ T cells [88]. These results led the authors to propose that differential expression of Ly6C can distinguish TE from MP cells within the T_H1 subset. At day 8, PSGL-1^{lo}Ly6C^{lo} effector cells showed high expression of known T_{FH} markers (ICOS, CXCR5, PD-1). This PSGL-1^{lo}Ly6C^{lo} subset was found along with PSGL-1^{hi}Ly6C^{hi} and PSGL-1^{hi}Ly6C^{lo} T_H1 cells within the memory cells at day 150 after infection, suggesting that MP of both T_H1 and T_{FH} phenotypes may persist long term [88]. Interestingly, while the PSGL-1^{hi}Ly6C^{lo} MP population was thought to be primarily T_H1 cells, it was later shown by Choi et al. [23] that the PSGL-1^{hi}Ly6C^{lo} MP population actually contains both $CXCR5^-$ T_H1 and $CXCR5^+$ T_{FH} cells at comparable frequencies. These results highlight the complexity and heterogeneity within $CD4^+$ memory T cells and the need for further studies to fully understand the nature of the $CD4^+$ memory T cell pool.

To investigate the potential of T_{FH} memory cells for re-differentiation upon reinfection, Hale et al. [51] utilized expression of CXCR5 and Ly6C to distinguish between T_H1

($CXCR5^{-}Ly6C^{hi}$) and T_{FH} ($CXCR5^{+}Ly6C^{lo}$ and $CXCR5^{+}Ly6C^{int}$) memory populations following acute infection with LCMV-Armstrong, then transferred each of the three subsets into naive hosts for reinfection. T_H1 memory cells mostly maintained high Ly6C expression with few effector cells gaining CXCR5 expression, while T_{FH} memory cells were able to give rise to both $CXCR5^{-}Ly6C^{hi}$ T_H1 cells and $CXCR5^{+}Ly6C^{lo/int}$ T_{FH} cells. This multi-potency of T_{FH} memory cells during re-challenge has also been observed in acute bacterial infection with LM [81] as well as in viral influenza infection [83].

In a concurrent study, Pepper et al. addressed $CD4^{+}$ memory T cell differentiation using LM infection and the expression of CXCR5 and CCR7, a marker used in previous studies to identify T_{CM}. During acute infection, antigen-specific effector cells segregated into a $CXCR5^{-}$ population favouring the T_H1 phenotype and a $CXCR5^{+}$ population [113]. A fraction of the $CXCR5^{-}$ T_H1 population, which the authors termed T_H1 effector memory cells, survived to a memory time point and, upon re-challenge, produced T_H1 effector cells. The $CXCR5^{+}$ effector population included cells with high expression of the lineage-defining factor Bcl6, were localized to follicles and were termed T_{FH}, while cells with lower Bcl6 levels showed co-expression of CCR7 and were termed T_{CM}. It is worth noting that the T_{FH} subset resembled what some studies term germinal center T_{FH} cells (GC T_{FH}); GC T_{FH} can lose expression of Bcl6 after infection, suggesting that, depending on the time point, the $CXCR5^{+}$ population can include cells that did not enter the GC as well as those that were transiently in the GC. While both T_{FH} and T_{CM} in this study expressed CXCR5, T_{CM} were not seen in the follicles and, upon re-challenge, produced both T_H1 effector cells and $CXCR5^{+}$ cells that likely include T_{FH} and GC T_{FH} [113]. Notably, Choi et al. [23] found that precursors of T_{FH} or the $CXCR5^{+}$ populations show greater potential to develop into memory cells compared with T_H1 precursors and share gene-expression signatures with memory $CD8^{+}$ T cells. These results suggest that both T_H1 and T_{FH} effector T cells can give rise to memory cells, and $CXCR5^{+}$ T_{FH}-derived memory cells have greater plasticity in generating secondary effector phenotypes.

Corroborative reports affirming the increased plasticity of T_{FH} memory relative to T_H1

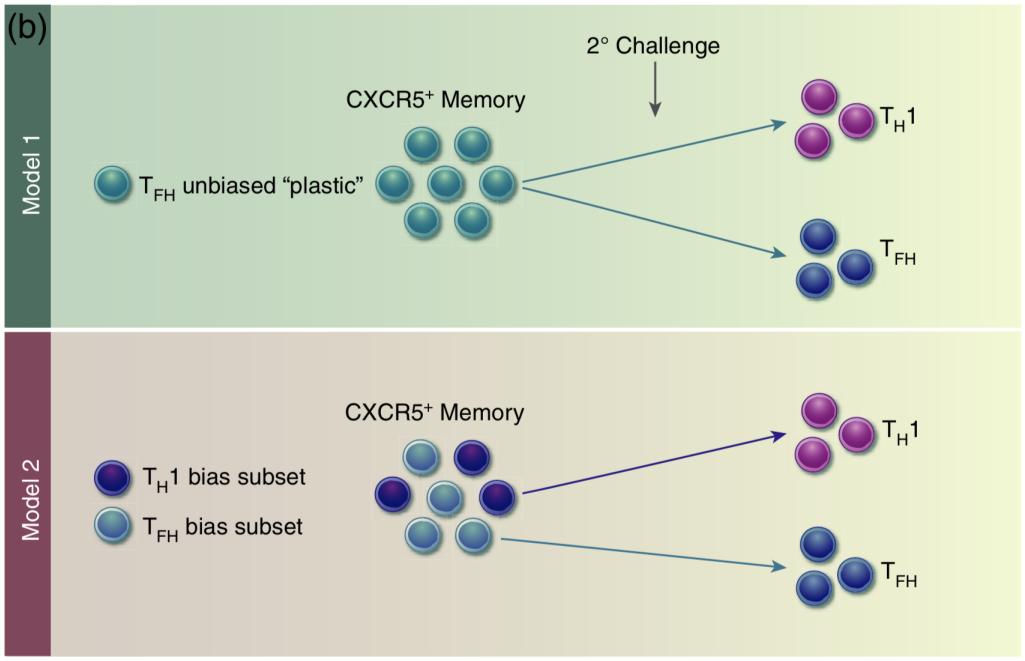


Figure 1.3. Two models for T follicular helper cell (T) multi-potency: (1) T_{FH} memory cells retain cellular plasticity and can differentiate into T_H1 or T_{FH} secondary effector cells based on signals present during secondary challenge; (2) T_{FH} memory cells are actually a heterogeneous population with subsets that are biased or primed towards a particular secondary effector lineage (T_H1 or T_{FH}).

memory upon re-challenge suggest that T_{FH} memory populations may retain a greater cellular 'stem-ness' and are capable of providing a more comprehensive and robust secondary response during re-infection. Two possible models can explain the multi-potency demonstrated by $CXCR5^+$ memory cells (Figure 1.3). One possible explanation is that T_{FH} memory cells are inherently more plastic compared with other T_H memory cells and, therefore, retain the ability to differentiate into alternative helper lineages upon reinfection. A second possibility is that the $CXCR5^+$ memory population actually contains distinct subsets that are programmed or biased towards a specific T_H lineage upon secondary challenge. In this case, $CXCR5^+CCR7^+$ could distinguish memory cells with the greater potential for re-expansion, while $CXCR5^+CCR7^-$ cells may be long-lived T_{FH}/GC T_{FH} cells that have downregulated Bcl6 and PD-1, and are more similar to long-lived effector subsets. Based on the data currently available, neither hypothesis can be eliminated and further characterization of T_{FH} memory cells, perhaps using single-cell

approaches, is needed to determine whether the multi-potency of T_{FH} memory is the result of cellular plasticity or population heterogeneity, or both.

In line with this idea, a recent study by Ciucci et al. [24] utilized single-cell RNA sequencing to investigate the heterogeneity of antigen-specific $CD4^+$ effector T cells in response to acute LCMV infection. Visualization with t-distributed stochastic neighbour embedding (t-SNE) of day 7 effector T cells yielded multiple transcriptionally distinct clusters showcasing the heterogeneity exhibited by T_H1 and T_{FH} effector cells. At 30 days post-infection, single cell analysis also showed multiple distinct transcriptional clusters with shared T_{FH} features, supporting the idea that memory $CXCR5^+$ T_{FH} multi-potency may be the result of population heterogeneity.

1.3.2 T_H2 $CD4^+$ memory T cells

T_H2 memory cells have been best characterized in the context of allergic inflammatory disorders [34], though some studies have highlighted this population's role in defense against parasitic worm infection. As mentioned previously, antigen-experienced $CD4^+$ T_H cells contract more rapidly after pathogen clearance compared with $CD8^+$ T cells [143], which is why early investigations into T_H2 memory relied on adoptive cell transfers of *in vitro* polarized T_H2 effector cells [104]. This system involved activating CD4 T-cells *in vitro* with antigen and antigen-presenting cells followed by culturing in T_H polarizing conditions [104] and subsequent adoptive transfer. Interestingly, *in vitro* generated T_H1 and T_H2 cells retained their expression of lineage-defining transcription factors (TFs), T-bet and GATA3, respectively, for months after transfer into naive hosts [53]. However, upon viral infection with LCMV, *in vitro*-derived T_H2 memory cells were able to adapt a T_H1 phenotype and persist as a 'hybrid' memory cell with combined T_H1 and T_H2 characteristics [53]. Utilizing a similar *in vitro* polarization system, Endo et al. [35] identified an interleukin-5 (IL-5)-producing subset of T_H2 memory cells in the spleen that is primarily responsible for asthmatic symptoms, such as eosinophilic infiltration into the airway, airway hyper-responsiveness and mucus hyper-production in a murine model of

T_H2-driven allergic airway inflammation. These studies provided early evidence of the potential existence of T_H2 memory populations, but data demonstrating direct *in vivo* generation were lacking until recently.

A study by Hondowicz et al. [56] provided key insights into T_H2 memory studying the endogenous allergen-specific $CD4^+$ T cells induced in response to house dust mite (HDM) inoculation. Using pMHCII tetramers to follow antigen-specific $CD4^+$ T cells, the authors showed an expansion of allergen-specific $CD4^+$ T_H2 cells in SLOs and the lung following intranasal HDM administration. Notably, this allergen system induces both antigen-specific T_H2 and T_{FH} cells, analogous to the T_H1 and T_{FH} response against LCMV-Armstrong. The allergen-specific memory pool in the SLOs consisted of $CXCR5^+$ and $CXCR5^-$ cells that also expressed $CCR7^+$, consistent with the earlier observations that memory T cells retain characteristics of T_H effector phenotypes.

1.3.3 T_H17 $CD4^+$ memory T cells

Though not as extensively characterized as other helper subsets, memory T17 cells have been documented in both humans and mice, primarily in the context of autoimmunity [93]. Early memory experiments using LM infection showed that T_H17 cells existed only transiently following intranasal infection [112]. However, it is worth noting that LM may not be an optimal infection for T_H17 studies as it is an intracellular pathogen [156] and most efficiently induces T_H1 cells. Muranski et al. [102] reported on long-lived memory T_H17 cells but, similar to early T_H2 studies, these cells required *in vitro* polarization prior to transfer into host mice. In a recent study of dry eye disease, Chen et al. [19] utilized a pre-clinical murine model of autoimmune ocular disease, where mice were subjected to 14 days of environmental desiccating stress followed by rest in normal conditions for 14 days, and found disease-specific pathogenic memory T_H17 cells in both the inflamed site and draining lymph nodes. Two cytokines associated with $CD4^+$ memory T lymphocytes, IL-7 and IL-15 [111] were shown to be crucial in the maintenance of these pathogenic T_H17 cells. Neutralization of these cytokines with topical application of

anti-IL-7 or anti-IL- 15 antibody decreased the number of T_H17 cells in both the conjunctiva and lymph nodes, offering a potential therapy for autoimmune disorders. One crucial caveat to note is that these 'memory' T_H17 cells were studied under the chronic inflammatory environment of autoimmunity, perhaps under prolonged or recurrent exposure to antigen; therefore, this population's identity as true resting memory cells remains uncertain.

1.3.4 Tissue-resident CD4$^+$ memory T cells

Much like circulating CD4$^+$ memory T cells, studies of tissue-resident lymphocytes have predominantly focused on CD8$^+$ T_{RM} due to the heterogeneity of CD4$^+$ memory T cells and the existing gaps in knowledge regarding mechanisms governing memory CD4$^+$ T cell formation. Nevertheless, recent studies have highlighted a prominent population of long-lived CD4$^+$ T cells within many non-lymphoid tissues, including the lungs [56, 57, 136, 130, 149, 126, 142, 118, 133, 18], small intestine (SI) [101, 135, 128, 39, 115], skin [44, 46, 45, 125, 25], and female reproductive tract (FRT) [135, 6, 60]. Similar to their CD8$^+$ T cell counterparts, CD4$^+$ T_{RM} have been shown to facilitate rapid immune defense upon re-exposure to antigen and can supplant innate immunity in recognizing and responding to recurrent infections [47]. However, much remains to be explored about the phenotype, function and maintenance of CD4$^+$ T_{RM} during infections. Additionally, differences exist between CD4$^+$ and CD8$^+$ T_{RM} in tissue localization, surface marker expression and cytokine cues driving T_{RM} formation; these outstanding questions in the field need further clarification to better define the identity and differentiation of CD4$^+$ T_{RM}.

Classically, tissue-resident memory T lymphocytes have been defined using parabiosis experiments in which a naive mouse and an immune mouse, previously exposed to antigen, are surgically joined to create a shared circulatory system [133, 120]. Thus, all circulating cells will normalize between both partners while the non-circulating tissue-resident cells remain lodged in the tissues of the immune mouse. Alternative methods have been developed and validated to assess whether cells remain in tissues, including intravenous injection of a fluorescently

labelled antibody to mark cells in the circulating system [2]. Any cells positive for the label are considered 'circulating', while unlabelled cells are assumed to have limited access to circulation and are therefore 'tissue-resident' [133]. To determine the protective functions of tissue-resident lymphocytes in secondary infection, immune mice are treated with FTY-720, an agonist of sphingosine-1-phosphate receptor 1 (S1PR1), which causes decreased surface expression of S1PR1 and therefore prevents egress of circulating memory cells from lymph nodes [55]. When these mice are re-challenged with the original pathogen, any immune response at the local site of infection will be mediated only by cells resident to that tissue [121].

1.4 Discussion

The diversity and plasticity of effector $CD4^+$ T cells create a heterogeneous memory pool, making the study of helper T cell memory differentiation complex. While there are some promising markers to differentiate between memory T_H1 and T_{FH} memory subsets [88, 51], it is still unclear whether both helper memory populations originate from their respective effector cells or whether the 'stem-like' properties of T_{FH} cells make them the primary precursor [23]. Likewise, in other infection systems that elicit T_H2 or T_H17 effector cells, we do not know how these effector populations contribute to the final pool of memory cells. It is clear that current surface marker knowledge does not accurately identify subpopulations in effector and memory pools, and further work requires examining TFs and regulators that may direct the memory program. One possible approach to parse the heterogeneity of $CD4^+$ memory T cells is through bulk and single-cell epigenetic and transcriptional profiling of cells in the circulation and tissues over the course of an infection to identify whether an early memory precursor exists or whether memory potential is programmed as effector cells contract and die [96, 150, 68]. Paired with adoptive transfers of putative subsets, it will be possible to identify key factors at steps in the effector-to-memory transition, and in the formation and survival of the tissue-resident subset. Uncovering the origin and identity of resting memory or MP cells within a particular helper

T cell lineage will lay the foundation for future molecular studies into how each memory T_H lineage is uniquely regulated.

In chapter 2, we investigate the role of transcriptional inhibitor protein Id3, which was previously shown to be important for $CD8^+$ memory T cells, in $CD4^+$ memory T cells. We examine whether Id3 can serve as a conserved marker of $CD4^+$ memory T cells across differing T_H lineages and further inspect the transcriptomic signatures associated with Id3 expression in $CD4^+$ memory T cells. In chapter 3, we highlight an unbiased approach for investigating putative transcriptional regulators of $CD4^+$ memory T cells. We will use a novel bioinformatics analysis to generate a ranked list of putative transcriptional regulators important for memory T_H1 and T_{FH} cells. Then we will showcase a CRISPR/Cas9 ribonucleoprotein transfection workflow for an efficient genome-editing system in $CD4^+$ T cells which will be utilized for *in vivo* validation of predicted targets.

Chapter 1, in part, is a reprint of the material as it appears in Immunology. Nguyen, Q. P., Deng, T. Z., Witherden, D. A., and Goldrath, A. W. *Origins of CD4$^+$ circulating and tissue-resident memory t-cells.* Immunology 157, 1 (2019), 3–12. The dissertation author was a primary author of this paper.

Chapter 2

Id3 expression identifies mouse CD4$^+$ memory T_H1 cells

2.1 Introduction

Generation of T cell memory is crucial in conferring vaccine-induced immunity, particularly against pathogens where neutralizing antibodies alone are insufficient at providing long-term protection. Antigen-specific CD4$^+$ T cells expand upon pathogen recognition and, depending on the infection milieu, differentiate into distinct effector cell types including T_H1, T_H2, T_H17, T_{FH}, and T_{reg} cells. Following the resolution of infection, a residual population of CD4$^+$ memory T cells remain within the circulation or in tissues that persist long-term and provide protection from reinfection [107]. The memory CD4$^+$ T cell population within the circulation has conventionally been divided into two subsets: effector-memory T cells (T_{EM}) and central-memory T cells (T_{CM})[113, 119]. T_{EM} are defined by low expression of CD62L and CCR7, with access to non-lymphoid sites and the ability to produce effector cytokines within hours following TCR stimulation. T_{CM} cells are characterized by high levels of CD62L and CCR7, and the ability to recirculate through lymph nodes, secrete IL-2 upon reactivation and undergo significant proliferation to generate secondary effector CD4$^+$ T cells [113].

Considerable efforts have been made to classify CD4$^+$ T cell memory precursor and memory T cell populations based on expression of cell-surface receptors, transcription factors and effector molecules such as cytokines. The fact that expression of many of these molecules occurs

along a continuum rather than being polarized between subsets, compounded by the existence of lineage plasticity amongst the $CD4^+$ T cell subsets during primary and secondary responses, have added substantial complexity to this effort [86, 105, 147, 16, 74]. Studies by several groups have attempted to relate unique phenotypic markers found on effector $CD4^+$ T cells with their intrinsic potential to form long-lived memory cells [113, 23, **?**, 88]. Two prominent models have emerged: one positing that the MP and memory population are heterogeneous whereby each T_H subset contains a portion of cells that is long-lived with expansion potential [88, 51], or alternatively, that the T_{CM} or T_{FH} subset serves as a unique source of memory $CD4^+$ T cells and a proportion of these cells are able to survive following the contraction phase to seed the memory T cell compartment [23, 51].

The enriched multipotency of $CXCR5^+$ T_{FH} memory cells (compared to $CXCR5^-$) has been described following Lymphocytic choriomeningitis virus (LCMV)-Armstrong [51, 78, 24, 83, 81]. Utilizing the T_{EM} and T_{CM} paradigm for characterization, Pepper et al. found that $CD4^+$ T_{EM} ($CXCR5^-CCR7^-$) cells primarily gave rise to $CXCR5^-$ (T_H1) secondary effector cells, while T_{CM} ($CXCR5^+CCR7^+$) cells gave rise to both $CXCR5^+$ (T_{FH}) and $CXCR5^-$ (T_H1) secondary effector cells in response to Listeria monocytogenes infection [113]. Additional phenotypic subsetting of memory $CD4^+$ T cells revealed a PSGL-1^{hi}Ly6C^{lo} MP subset following acute LCMV infection that was shown to exhibit greater longevity and increased proliferation following antigen re-challenge compared to the PSGL-1^{hi}Ly6C^{hi} subset [88]. While the PSGL-1^{hi}Ly6C^{lo} MP population was originally presumed to be primarily T_H1 cells, it was later suggested to also contain T_{FH} cells at comparable frequencies [23]. Collectively, the evidence suggests that T_H1 memory cells can persist and form secondary effector cells of only the T_H1 lineage, while T_{FH} memory cells exhibit greater multipotency in the context of pathogen re-challenge. Additionally, formation of $CD4^+$ T_{CM} phenotype cells was recently shown to require Thpok, which is also necessary for T_{FH} formation via suppression of T_H1 associated transcription factors Blimp-1 and Runx3 [24]). Thus, it remains a question if the pluripotent memory $CD4^+$ T cell subset is necessarily contained within the T_{FH} $CXCR5^+CCR7^+$ population in all infection and

inflammation contexts [116].

Despite clear differences between memory $CD4^+$ and $CD8^+$ T cell populations [40], the model of $CD8^+$ T cell memory formation can serve as a valuable guiding framework for memory $CD4^+$ T cell investigations. Our lab and others have demonstrated the role of E and Id proteins in the differentiation of both short-lived effector and MP populations of $CD8^+$ T cells [148, 15, 64, 92]. Notably, Id3 expression identified $CD8^+$ T cells with memory potential at effector timepoints [148], which raises the possibility of an analogous role for Id3 in memory $CD4^+$ T cells. E/Id proteins cooperate to regulate transcriptional programs necessary for T_H cell specification in naive, infection and autoimmune settings [23, 81, 123, 42, 99, 90, 41, 80], however, their role in differentiation and persistence of memory $CD4^+$ T cells has not been studied as extensively. Here, we find that a population of $Id3^{hi}T_H1$ memory cells emerges following acute LCMV infection which exhibits enhanced expansion potential and increased expression of memory-associated molecules such as CD127, Bcl2 and Tcf1 when compared to $Id3^{lo}T_H1$ cells at memory timepoints. While the majority of T_H1 memory $CD4^+$ T cells appear limited in their ability to form both T_H1 and T_{FH} secondary effector cells, the $Id3^{hi}T_H1$ memory $CD4^+$ T cells present as a small durable subset with enhanced multipotent recall potential.

2.2 Results

2.2.1 Helper $CD4^+$ T cells share transcriptomic characteristics with cytotoxic $CD8^+$ T cells

To assess the possibility of common memory T cell differentiation programs between $CD4^+$ and $CD8^+$ T cells, we compared global gene expression of effector and memory $CD4^+$ SMARTA T cells (recognizing LCMV gp66-77 presented by MHC Class II I-Ab) with changes in gene-expression in $CD8^+$ T cells responding to LCMV-Armstrong infection. Strikingly, the majority of genes upregulated by T_H1 and T_{FH} subsets at day 7 and 41 following LCMV infection compared to naive SMARTA $CD4^+$ T cells were those found within the effector or memory $CD8^+$

T cell gene signatures, respectively (Figure 2.1a,c). Thus, despite biological differences among $CD4^+$ and $CD8^+$ T cells and T_H1 versus T_{FH} populations, the two lineages shared unexpected similarities in gene expression at both effector and memory time points (Figure 2.1b,d). Further, geneset enrichment analysis (GSEA) indicated that the T central memory precursor (T_{cmp}) signature recently defined by Ciucci et al. [24] is enriched in both the T_H1 and T_{FH} effector populations, suggesting that both of these lineages may harbor T cell memory potential (Figure 2.1e).

2.2.2 Id3-GFP-expressing memory $CD4^+$ T cells expand and give rise to T_H1 and T_{FH} secondary effector cell populations

Given the evident similarities in transcriptional signatures we observed between $CD4^+$ and $CD8^+$ T cell populations, we hypothesized that, akin to $CD8^+$ MP T cells, Id3 may serve as a marker of memory potential within effector $CD4^+$ T cell populations. We assessed kinetics of Id3 expression by adoptive transfer of $CD4^+$ T cells using $Id3^{GFP/+}$ reporter SMARTA TCR transgenic T cells [123]. $CD4^+$ T cells from these mice were transferred into congenically distinct hosts, which were infected 1 day later with LCMV. Consistent with high Id3 expression by some T cells, prior to infection, more than 95% of $CD4^+$ T cells expressed Id3-GFP (Figure 2.2a,b). Following infection, the proportion of effector $CD4^+$ T cells with low Id3 expression signficiantly increased, but as the infection was cleared, upwards of 90% of the remaining memory cells expressed Id3-GFP (Figure 2.2a,b), with a greater absolute number of Id3-expressing cells surviving the contraction phase and persisting to memory timepoints (Figure 2.2c)

Consistent with our previous studies, at effector timepoints Id3-GFP^{hi} cells were almost exclusively T_{FH} ($CXCR5^+SLAM^{lo}$ or $CXCR5^+PD\text{-}1^{lo}$) and GC T_{FH} cells ($CXCR5^+PD\text{-}1^+$), while the vast majority of Id3-GFP^{lo} cells displayed a T_H1 phenotype ($SLAM^+CXCR5^-$ or $CXCR5^-PD\text{-}1^-$) (data not shown) [123]. Since Id3 positive cells encompass the majority of the memory $CD4^+$ T cell population, we evaluated whether Id3-GFP^{hi} T cells had any advantages over the Id3-GFP^{lo} memory cells in the context of re-infection. Id3-GFP^{lo} or Id3-GFP^{hi} SMARTA

memory CD4$^+$ T cells (28-32 days following primary infection) were transferred into a new cohort of B6 hosts, which were then infected 1 day later with LCMV (Figure 2.2d). Following LCMV rechallenge, we found that both Id3-GFP^{lo} and Id3-GFP^{hi} SMARTA CD4$^+$ T cell populations were able to generate secondary effector T_H cells. However, the phenotype and abundance of the expanded progeny were strikingly dissimilar (Figure 2.2e-k). We recovered 3.4-fold more secondary effector T cells derived from Id3-GFP^{hi} memory T cells than from Id3-GFP^{lo} memory cells (Figure 2.2e), indicating that Id3-GFP^{hi} memory cells have significantly greater expansion potential. Secondary effector T cells generated from the transfer of Id3-GFP^{lo} memory cells also maintained low expression of Id3-GFP, whereas Id3-GFP^{hi} memory cells generated a mixed population of Id3-GFP^{lo} and Id3-GFP^{hi} secondary effector T cells (Figure 2.2f,g). The majority of secondary effector cells derived from Id3- GFP^{lo} memory cells were SLAM$^+$CXCR5$^-$ T_H1 cells. Conversely, the Id3-GFP^{hi} memory T cells repopulated the CD4$^+$ T cell compartment with both T_H1 (SLAM$^+$CXCR5$^-$) cells and T_{FH} (SLAM^{lo}CXCR5$^+$) secondary effector T cells (Figure 2.2h,i). Further, the Id3-GFP^{hi} cells also generated a higher frequency of PD-1$^+$CXCR5$^+$ GC T_{FH} cells when compared to the Id3-GFP^{lo} cells (Figure 2.2j,k). These data suggest that Id3-GFP-expressing memory T cells have enhanced expansion and multipotent recall potential, capable of differentiating into both T_H1 and T_{FH} cells upon rechallenge.

2.2.3 Id3-GFP^{hi} T_H1 memory cells accumulate at memory time points

Both T_H1 and T_{FH} memory T cells persist following LCMV infection (Figure 2.3a), but, we observe a decrease in the frequency of T_H1 (CXCR5$^-$) cells over time. Within this waning population however, we observed the emergence of an Id3-GFP-expressing memory T_H1 population, where 15% of T_H1 memory cells expressed Id3-GFP by day 41 following infection (Figure 2.3a,b). T_{CM} CD4$^+$ T cells are known to possess enhanced differentiation potential, and are traditionally marked by CCR7 expression. To assess how this small population of Id3-GFP^{hi} T_H1 memory cells might factor into the broader paradigm of T_{CM} and T_{EM} CD4$^+$ populations, we analyzed the expression of CCR7 by this Id3-GFP-expressing T_H1 memory

population [113, 119, 88]. While a significant portion of Id3-GFP^{hi} T_{FH} memory cells exhibited CCR7 expression (Figure 2.3c,d), Id3-GFP^{hi} T_H1 memory cells did not gain expression of CCR7 (Figure 2.3c,d), suggesting that they do not fit the canonical central memory T cell criteria.

2.2.4 Id3-GFP^{hi} T_H1 memory cells give rise to T_H1 and T_{FH} cells in a secondary response

As we previously found that the Id3-GFP^{hi} memory T cell population exhibited greater multipotent potential during secondary challenge, we next evaluated whether this was the case specifically within the T_H1 memory lineage. Id3-GFP^{hi} and Id3-GFP^{lo} T_H1 primary memory cells were sorted and transferred into a new cohort of B6 hosts, which were infected 1 day later with LCMV (Figure 2.4a,b). Following re-infection, we found that Id3-GFP^{lo} T_H1 donors primarily generated secondary effector cells with low expression of Id3-GFP, whereas Id3-GFP^{hi} T_H1 donors were able to generate 3.3-fold more secondary effector cells with a mixed population of both Id3-GFP^{lo} and Id3-GFP^{hi} cells (Figure 2.4d-i). Secondary effector T cells from Id3-GFP^{lo} T_H1 donors were predominantly T_H1 cells, while the secondary effector T cells from Id3-GFP^{hi} T_H1 donors were composed of both T_H1 (SLAM^{+}CXCR5^{-}) cells and T_{FH} (SLAM^{lo}CXCR5^{+}) cells (Figure 2.4f,g). Further, the Id3-GFP^{hi} T_H1 cells also generated a higher frequency of PD-1^{+}CXCR5^{+} GC T_{FH} cells when compared to the Id3-GFP^{lo} T_H1 cells (Figure 2.4h,i). Collectively, the data suggest that Id3-GFP^{hi} T_H1 memory cells have greater multipotent recall potential compared to Id3-GFP^{lo} T_H1 memory cells, despite the absence of CCR7 expression and canonical T_{CM} phenotype [119].

We further characterized the Id3-GFP^{hi} and Id3-GFP^{lo} T_H1 memory populations for expression of key molecules associated with long-lived memory T cells. Notably, Id3-GFP^{hi} T_H1 memory cells expressed significantly more IL-7-receptor (CD127), suggesting a greater responsiveness to IL-7 that would promote memory T cell survival and homeostasis [138, 75, 109]. Correspondingly, the Id3-GFP^{hi} T_H1 cells expressed increased levels of the anti-apoptotic molecule BCL2, further supporting the notion that Id3-GFP^{hi} T_H1 cells have an

increased capacity for survival compared to their Id3-GFP^{lo} counterparts. Finally, TCF1, a transcription factor important for memory CD8^{+} T cell formation and function [154, 153] and T_{FH} development [82, 22, 146], was expressed at higher levels in Id3-GFP^{hi} T_H1 memory cells compared to the Id3-GFP^{lo} T_H1 memory population (Figure 2.4j-m). These data suggest that expression of Id3-GFP imbues a population of T_H1 memory cells with enhanced memory T cell characteristics including greater survival, expansion and multipotent differentiation potential.

2.2.5 Id3 expression defines a transcriptionally distinct T_H1 memory population

To examine the transcriptional differences between the Id3-GFP^{lo} and Id3-GFP^{hi} T_H1 memory cells, we performed RNA-seq on these sorted populations (30 of LCMV infection). Id3-GFP^{hi} T_H1 memory cells were enriched for transcripts encoding key memory genes including Bcl2 and Tcf1 compared to the Id3-GFP^{lo} T_H1 memory cells (Figure 2.5a). Relative to their Id3-GFP^{hi} counterparts, the Id3-GFP^{lo} T_H1 memory cells were enriched for effector molecule transcripts (Prdm1,Gzma, Gzmb and Gzmk) suggesting a more "effector-like" transcriptional profile (Fig 4a), whereas expression of Id2 and Bcl6 were equivalent, confirming their T_H1 identity. When directly comparing Id3-GFP^{hi} T_H1 memory cells to the Id3-GFP^{hi} T_{FH} memory cells, the Id3-GFP^{hi} T_H1 memory cells expressed T_H1-associated transcripts compared to the Id3-GFP^{hi} T_{FH} memory cells, emphasizing that these were distinct memory populations (Figure 2.62a). Notably, the T_{FH} memory population was also enriched for memory-associated genes (Figure 2.6a), suggesting that CD4^{+} T cell memory subsets may persist along a cell-state continuum as previously described [74]. Using gene set enrichment analysis (GSEA), we found that when compared to the Id3-GFP^{lo} T_H1 cells, the Id3-GFP^{hi} memory cells were significantly enriched for the T_{CM} memory signature (defined by upregulated transcripts compared to T_{EM}cells) as well as the T_{cmp} signature [24] (Figure 2.5b).

As we observed marked phenotypic and functional differences between Id3-GFP^{hi} and Id3-GFP^{lo} T_H1 memory cells, we investigated the heterogeneity within the CD4^{+} effector and

memory T cell populations by single-cell RNA sequencing (scRNA-Seq) of SMARTA $CD4^+$ T cells at days 7, 21, and 41 of LCMV infection (Figure 2.5). Unsupervised clustering with visualization via UMAP (Uniform Manifold Approximation and Projection) revealed distinct clusters of cells that correlated with T_{FH} and T_H1 subsets based on expression of canonical lineage markers at each time point (Figure 2.5c,e). Memory T cell-associated genes (including Tcf7 and Il7r) were enriched in the day 21 and day 41 samples, while key markers of the T_{FH} (Cxcr5) and T_H1 (Tbx21, Cxcr6, Slamf1) lineages exhibited mutually-exclusive enrichment in the clusters (Figure 2.5e). To expand this analysis, we next defined the gene-expression signature of T_H1 and T_{FH} memory cells by performing bulk RNA sequencing on sorted T_H1 and T_{FH} memory SMARTA populations on day 21 of LCMV infection. We found that 2325 (1049 T_{FH} + 1276 T_H1) genes were differentially expressed with a fold change of ≥ 2 between these two populations (Figure 2.6b). We overlayed the T_{FH} and T_H1 memory signatures onto the single-cell projections to identify individual cells enriched for expression of the T_H1 or T_{FH} memory cell transcriptome (Figure 2.5e). The T_H1 signature-enriched cells showed expression of Id3 corresponding to the small population of Id3-GFP^{hi} T_H1 memory cells identified *in vivo* that also expressed memory-associated genes including Tcf7, Bcl2 and Il7r (Figure 2.6e). Notably, cells on day 7 of the response also showed moderate enrichment for our T_H1 and T_{FH} signatures, consistent with the presence of a $Id3^{hi}$ $CD4^+$ memory-precursor population as previously described in endogenous $CD4^+$ T cells [24].

To understand how the T_{CM} vs T_{EM}dichotomy broadly applies to memory $CD4^+$ T cells, we overlayed T_{CM} and T_{EM}signatures from memory $CD8^+$ T cell subsets onto the memory $CD4^+$ T cell single-cell RNAseq data (Figure 2.6f). Enrichment of the T_{EM}signature correlated with day 7 effector cell samples as well as "T_H1-signature"-enriched clusters, while the T_{CM} signature exhibited greater overlap with memory time point cells (days 21 and 44) as well as "T_{FH}-signature"- enriched clusters. This analysis suggests that without prior subsetting, T_{FH} memory cells overall exhibit a more T_{CM}-like phenotype than T_H1 memory cells. The T_{cmp} signature [24] also shows enrichment in memory T_H1 cells (cluster 2), supporting our observation

of that these cells are bona fide T_H1 memory cells (Figure 2.6g).

To focus on the hetereogenity of memory $CD4^+$ T cells specifically, we performed unsupervised clustering of memory time-point cells (days 21 and 44), which revealed 3 major clusters (Figure 2.5f). When we overlayed T_H1 and T_{FH} memory signatures onto these data, we found the T_{FH} signature enriched in one distinct cluster while the T_H1 signature spanned the remaining two clusters (Figure 2.5g). To test whether the two T_H1-enriched clusters represented the Id3-GFP^{hi} and Id3-GFP^{lo} T_H1 memory cells, we overlayed T_H1 Id3-GFP^{hi} and Id3-GFP^{lo} memory signatures (defined by fold change >1.75) generated from the bulk RNA-sequencing (Figure 2.5a) onto the single- cell analyses. Indeed, the T_H1 Id3-GFP^{hi} and T_H1 Id3-GFP^{lo} memory signatures exhibited mutually-exclusive polarization within T_H1-enriched clusters, with the T_H1 Id3-GFP^{hi}-signature highlighting both T_H1 and T_{FH} memory clusters (Figure 2.5h). Taken together, the single-cell RNAseq data definitively show that the Id3-GFP^{hi} T_H1 population as a subset of long-lived memory $CD4^+$ T cells that is transcriptionally distinct from T_{FH} memory cells and Id3-GFP^{lo} T_H1 cells.

2.3 Discussion

Given the pivotal role $CD4^+$ memory T cells play in mediating long-term cellular and humoral immunity, key parameters of their identity and differentiation are relevant in vaccine development. However, the existence of functionally distinct $CD4^+$ helper T cell subsets complicate the identification of $CD4^+$ memory T cells, and the current literature remains conflicted in whether there is a single $CD4^+$ memory T cell precursor that gives rise to secondary T_H subsets or there are memory precursor cells within each helper lineage. Collectively, we found that much like $CD8^+$ T cells, the transcriptional regulator Id3 defines a transcriptionally distinct population of $CD4^+$ T cells with enhanced memory potential. Importantly, within the T_H1 compartment, we identify for the first time, a novel subset of Id3-GFP^{hi} cells that accumulate in frequency over time following acute viral infection. Compared to Id3-GFP^{lo} T_H1 memory cells, the Id3-GFP^{hi}

T_H1 memory cells exhibited greater multipotency and proliferative potential upon secondary challenge. Furthermore, the Id3-GFP^{hi} T_H1 cells showed increased expression of molecules critical for T cell memory formation and survival when compared to Id3-GFP^{lo} memory T_H1 cells. Single-cell RNA-sequencing revealed that the the Id3-GFP^{hi} T_H1 memory cells formed a distinct population that retained genes associated with T_H1 polarization while also upregulating memory-associated molecules enriched in the T_{CM} and T_{FH} memory compartments. Our data uniquely define a subset of CXCR5$^-$ T_H1 memory cells with stem-like properties, a characteristic previously associated primarily with CXCR5$^+$ memory CD4$^+$ T cells.

Although our investigation identified Id3 as an important marker of CD4$^+$ memory T cells, the cellular origin of Id3^{hi} T_H1 memory cells remains unclear. One can imagine a Id3^{lo} T_H1 effector gaining Id3 expression past the infection peak and becoming a long-lived memory cell. However a scenario whereby a Id3^{hi} T_{FH} effector T cell potentially converting into a T_H1 phenotype during effector to memory T-cell conversion is also feasible. To probe whether the Id3^{hi} T_H1 memory population arise from T_H1 or T_{FH} effector T cells, experimental models with T-bet or Bcl6 deletion can be used to see how defects in T_H1 or T_{FH} effector cell generation impacts the formation of Id3^{hi} T_H1 memory T cells. Further investigation is also needed to determine possible differentiation signals for Id3^{hi} T_H1 memory cells. Examination into localization patterns of Id3$^+$ T_H1 memory cells, particularly within the secondary lymphoid organs, can yield potentially hints to the mechanisms underlying their generation. Compared to Id3^{lo} T_H1 memory cells, perhaps Id3^{hi} T_H1 memory T cells preferentially locate closer to the T/B cell border within SLOs, thus potentially receiving more cytokine signaling from T_{FH} or B cells, leading towards a Id3^{hi} fate. Additional investigation is needed to uncover the cellular origin and mechanism underlying the generation of Id3^{hi} T_H1 memory cells, however identifying universal markers of memory such as Id3 across T cell populations will undoubtedly aid in deconvoluting the complexities associated with defining CD4$^+$ T cell memory.

Chapter 2, in full, has been submitted for publication of the material as it may appear in PNAS, 2021. Shaw, L. A., Deng, T. Z., Omilusik, K. D., Nguyen, Q.P., and Goldrath, A. W.

Id3 expression identifies mouse CD4$^+$ memory T_H1 cells. The dissertation author was a primary investigator and a first author of this paper.

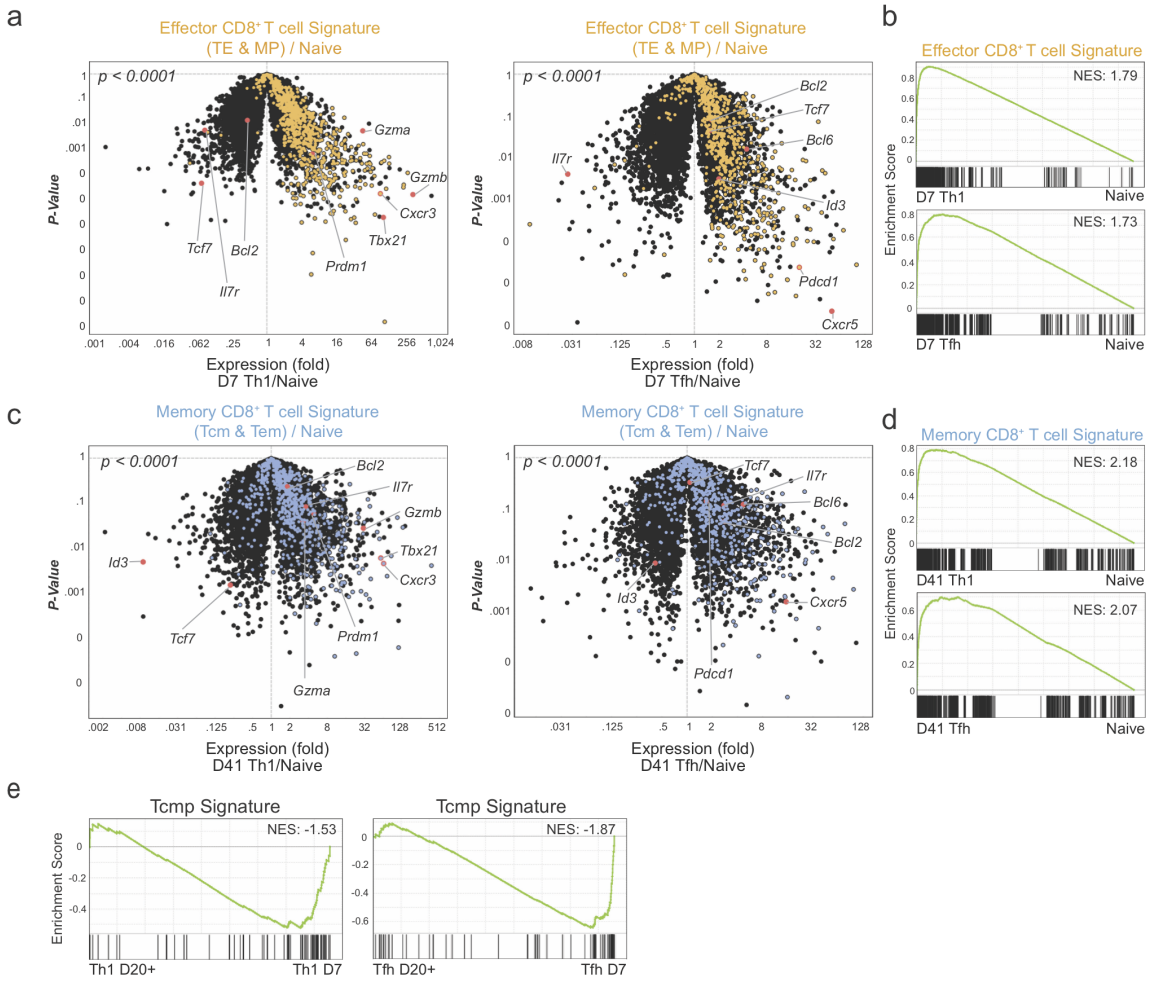
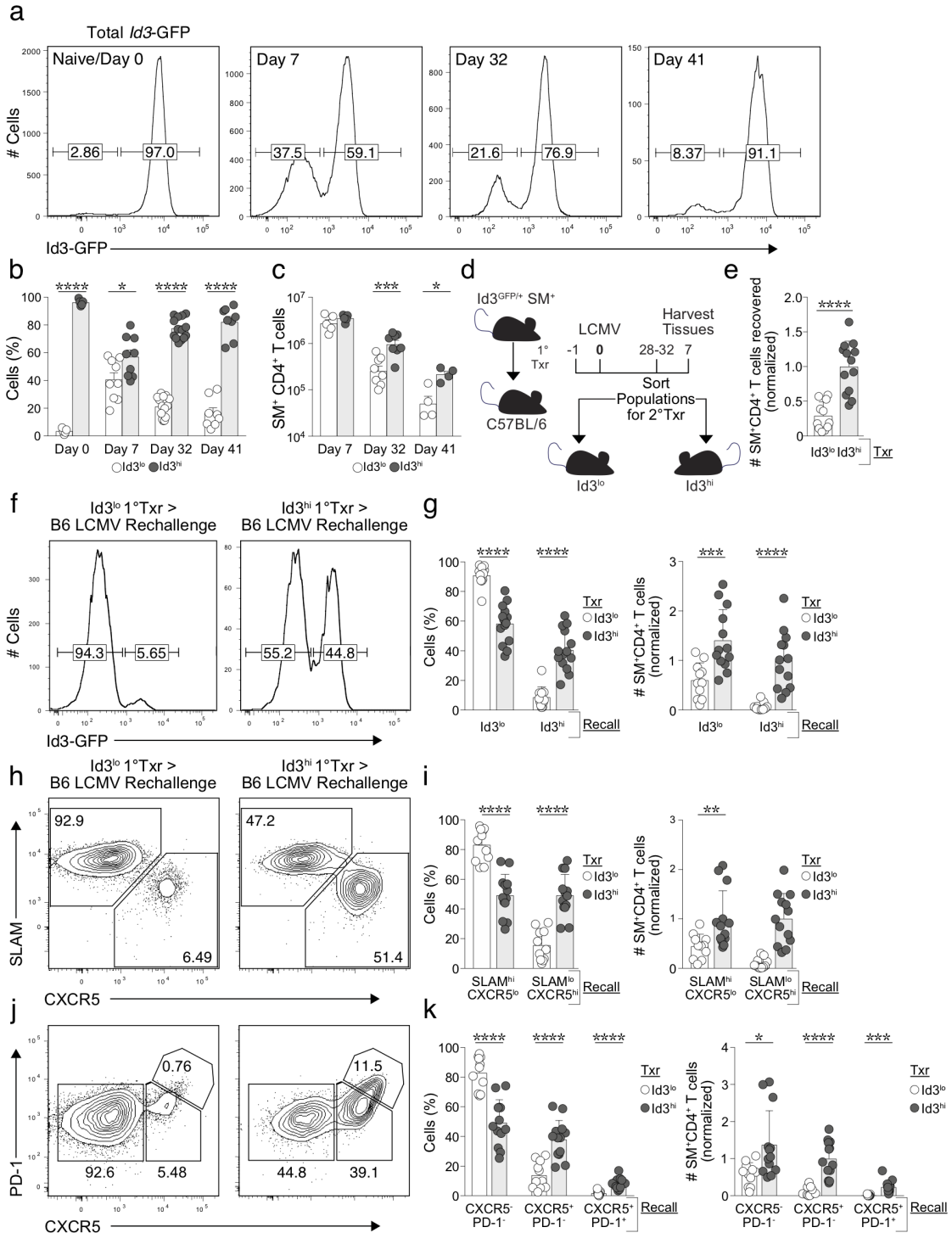


Figure 2.1. Shared transcriptional features between CD4$^+$ and CD8$^+$ memory T cells. (a) Volcano plots showing averaged mRNA expression (3-4 independent replicates) of T_H1 (left) or T_{FH} (right) effector SMARTA CD4$^+$ T cells vs averaged mRNA expression (2 independent replicates) of naive SMARTA CD4$^+$ T cells. The effector CD8$^+$ T cell gene-expression signature defined by transcripts enriched in both terminal effector (TE) and memory precursor (MP) CD8$^+$ T cells over naive CD8$^+$ T cells (FC $\geq$ 2) from previous published data (Milner et al., PNAS, 2020) [97] is overlaid (highlighted in orange). (b) GSEA of effector CD8$^+$ T cell gene-signature in effector vs naive CD4$^+$ T cells. (c) Volcano plots showing averaged mRNA expression (2 independent replicates) of T_H1 (left) or T_{FH} (right) memory SMARTA CD4$^+$ T cells vs averaged mRNA expression (2 independent replicates) of naive SMARTA CD4$^+$ T cells. The memory CD8$^+$ T cell gene-expression signature defined by transcripts enrichment in both T_{EM} and T_{CM} populations over naive CD8$^+$ T cells (FC $\geq$ 2) from previous published data (Milner et al., PNAS, 2020) is overlaid (highlighted in blue). (d) GSEA of memory CD8$^+$ T cell gene-signature in memory vs naive CD4$^+$ T cells. (e) GSEA of T_{cmp} signature in memory vs naive CD4$^+$ T cells.

Figure 2.2. Id3 expression defines CD4$^+$ T cells with increased memory potential. (a) Flow cytometric analysis of donor Id3$^{GFP/+}$ SMARTA CD4$^+$ T cells from C57BL/6 host mice over the course of an LCMV infection. (b,c) Frequency of Id3 expressing cells among SMARTA CD4$^+$ T cells (b) or total SMARTA CD4$^+$ T cells on indicated days of infection (c). (d-k) Id3-GFP^{lo} or Id3-GFP^{hi} memory SMARTA CD4$^+$ T cells were sorted on days 28-32, transferred to naive C57BL/6 hosts that were then infected with LCMV to be analyzed 7/8 days later. (d) Experimental schematic for isolation of memory T cells based on expression of Id3. (e) Total SMARTA CD4$^+$ T cells recovered from host mice at day 7/8 of secondary LCMV infection. (f) Analysis of Id3-GFP expression at day 7/8 of infection in donor cells from host mice that received transfers of either Id3-GFP^{lo} (left) or Id3-GFP^{hi} (right) memory SMARTA CD4$^+$ T cells. Numbers on histogram peaks indicate percent of cells within indicated gates. (g) Frequency among SMARTA CD4$^+$ T cells (left) and total SMARTA CD4$^+$ T cells (right) generated from indicated transferred populations in (f). (h) Analysis of the percent SLAM$^+$CXCR5$^-$ (T_H1) cells or SLAM^{lo}CXCR5$^+$ (T_{FH}) cells generated from indicated memory populations following secondary infection. (i) Frequency among SMARTA CD4$^+$ T cells (left) and total SMARTA CD4$^+$ T cells (right) from indicated populations in (h) are shown. (j) Percent CXCR5$^-$PD-1$^-$ (T_H1) cells, CXCR5$^+$PD-1$^-$ (T_{FH}) cells or CXCR5$^+$PD-1$^+$ (GC T_{FH}) cells formed from indicated memory populations following secondary infection. (k) Frequency among SMARTA CD4$^+$ T cells (left) and total SMARTA CD4$^+$ T cells (right) from indicated populations in (j). *P $<$0.05, **P $<$0.01, ***P$<$0.001 and ****P $<$0.0001 (two-tailed unpaired Student s t test). Data are representative of 3 experiments (a,c), each with n= 3-10 mice per group (mean $\pm$ s.e.m.) or pooled from three (b, e-k) independent experiments with n= 3-10 mice per group (mean $\pm$ s.e.m.)


a
Total Id3-GFP
Naive/Day 0
Day 7
Day 32
Day 41
Cells
2.86
97.0
37.5
59.1
21.6
76.9
8.37
91.1
Id3-GFP
b
Cells (%)
Day 0
Day 7
Day 32
Day 41
Id3lo Id3hi
c
SM+ CD4+ T cells
d
Id3GFP/+ SM+
1° Txr
C57BL/6
LCMV
Harvest Tissues
-1 0 28-32 7
Sort Populations for 2°Txr
Id3lo
Id3hi
e
SM+CD4+ T cells recovered (normalized)
Txr
f
Id3lo 1°Txr > B6 LCMV Rechallenge
Id3hi 1°Txr > B6 LCMV Rechallenge
94.3
5.65
55.2
44.8
g
Recall
h
92.9
6.49
47.2
51.4
SLAM
CXCR5
i
SLAMhi CXCR5lo
SLAMlo CXCR5hi
j
0.76
92.6
5.48
11.5
44.8
39.1
PD-1
k
CXCR5- PD-1-
CXCR5+ PD-1-
CXCR5+ PD-1+

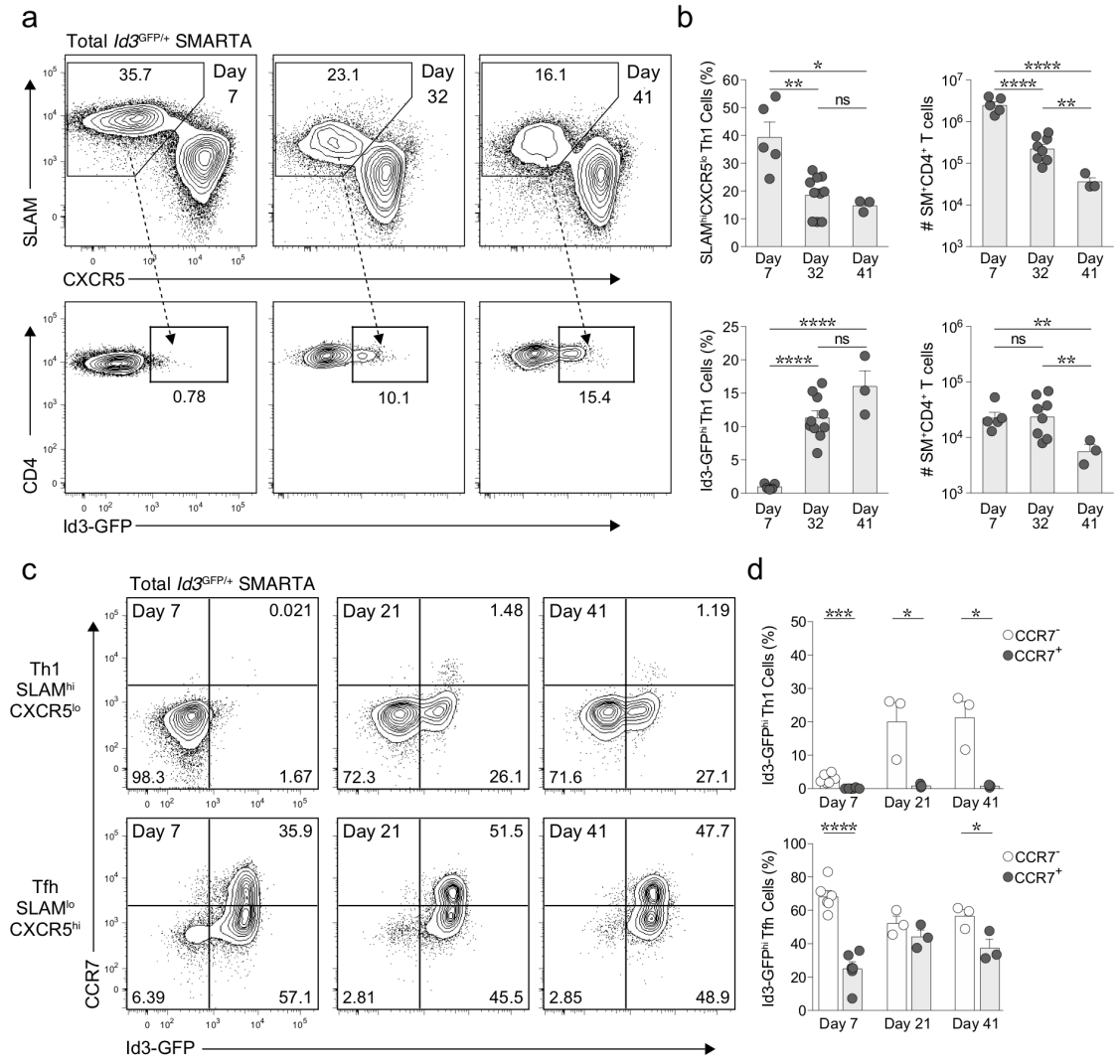


Figure 2.3. The T_H1 population contain Id3-expressing cells. (a) Flow cytometric analysis of donor $Id3^{GFP/+}$ SMARTA $CD4^+$ T cells from C57BL/6 host mice at indicated day of LCMV infection (7, 32 and 41). Numbers in outlined area indicate percent $SLAM^+CXCR5^-$ (T_H1) cells (top) and the expression of Id3-GFP within the T_H1 compartment (bottom). (b) Frequency of $SLAM^+CXCR5^-$ (T_H1) cells (left top) and total SMARTA $CD4^+$ T cells (right top); frequency of Id3-GFP-expressing cells among $SLAM^+CXCR5^-$ (T_H1) cells (left bottom) and total SMARTA $CD4^+$ T cells (right bottom). (c) Analysis of CCR7 expression on $Id3^{GFP/+}$ SMARTA $CD4^+$ T cells at days 7, 32, and 41 of LCMV infection. (d) Frequency of $Id3^+CCR7^-$ and $Id3^+CCR7^+$ cells among T_H1 and T_{FH} cells. *P <0.05, **P <0.01, ***P <0.001 and ****P <0.0001 (two-tailed unpaired Student s t test). Data are representative of two experiments each with n= 3-8 mice per group (mean $\pm$ s.e.m.).

Figure 2.4. Id3-GFP^{hi} T_H1 memory cells exhibit increased accumulation and multipotency upon rechallenge. (a-i) C57BL/6 host mice received a transfer of either Id3-GFP^{lo} or Id3-GFP^{hi} SLAM^{+}CXCR5^{-} T_H1 memory SMARTA CD4^{+} T cells, and were infected with LCMV for 7/8 days before analysis. (a,b) Experimental schematic for sorting SLAM^{+}CXCR5^{-} T_H1 cells based on expression of Id3. (c) Total SMARTA CD4^{+} T cells recovered from Id3lo T_H1 or Id3hi T_H1 secondary transfer on day 7/8 of reinfection. (d) Expression of Id3 in indicated memory populations on day 7/8 of reinfection. (e) Frequency among SMARTA CD4^{+} T cells (left) and total SMARTA CD4^{+} T cells (right) generated from indicated transferred populations in (d). (f) Percent of SLAM^{+}CXCR5^{-} (T_H1) cells or SLAM^{lo}CXCR5^{+} (T_{FH}) cells generated from indicated memory populations on day 7/8 of reinfection. (g) Frequency among SMARTA CD4^{+} T cells (left) and total SMARTA CD4^{+} T cells (right) from indicated populations in (f). (h) Percent CXCR5^{-}PD-1^{-} (T_H1) cells, CXCR5^{+}PD-1^{-} (T_{FH}) cells or CXCR5^{+}PD-1^{+} (GC T_{FH}) cells formed from indicated memory populations at day 7/8 of reinfection. (i) Frequency among SMARTA CD4^{+} T cells (left) and total SMARTA CD4^{+} T cells (right) from indicated populations in (h). (j-m) Analysis of memory Id3GFP/+ SMARTA CD4^{+} T cells at day 30 of LCMV infection. (j) Schematic of experimental timeline. (k) Frequency of SLAM^{+}CXCR5^{-} (T_H1) cells or SLAM^{lo}CXCR5^{+} (T_{FH}) cells among total donor cells (right) and Id3 expression in indicated donor subset (right). (l) Histograms show expression of indicated protein on Id3-GFP^{lo} (black) or Id3-GFP^{hi} (grey shaded) T_H1 memory populations. (m) Quantification of median fluorescence intensity for indicated protein. *P $<$0.05, **P $<$0.01, ***P$<$0.001 and ****P $<$0.0001 (two-tailed unpaired Student s t test). Data are normalized and pooled from two independent experiments with n= 3-10 mice per group (mean $\pm$ s.e.m.).

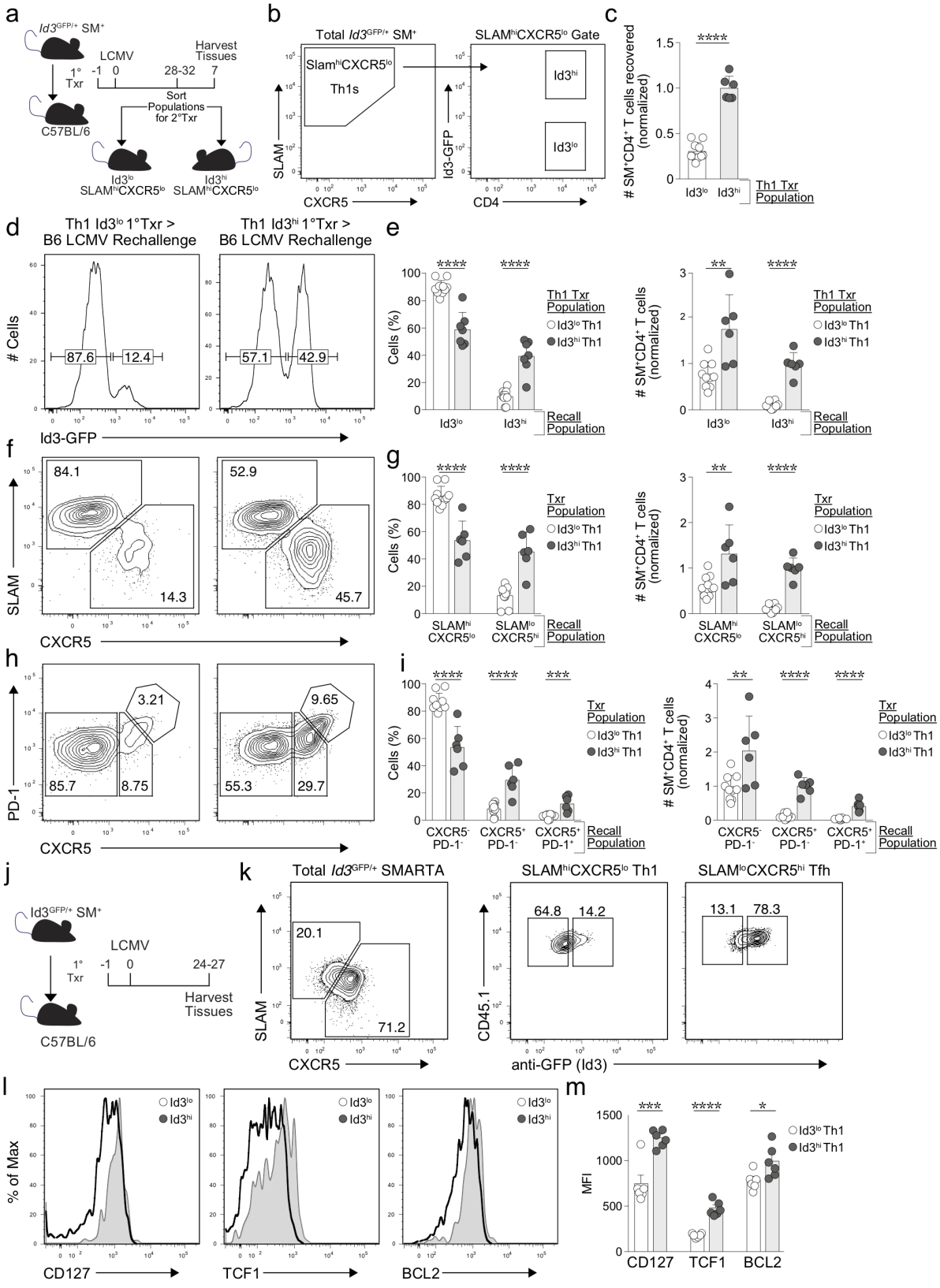

a
Id3GFP/+ SM+
LCMV
Harvest Tissues
1° Txr
-1 0
28-32 7
Sort Populations for 2°Txr
C57BL/6
Id3lo SLAMhiCXCR5lo
Id3hi SLAMhiCXCR5lo
b
Total Id3GFP/+ SM+
SLAMhiCXCR5lo Gate
SlamhiCXCR5lo
Th1s
SLAM
CXCR5
Id3-GFP
CD4
Id3hi
Id3lo
c
SM+CD4+ T cells recovered (normalized)
****
Id3lo Id3hi
Th1 Txr Population
d
Th1 Id3lo 1°Txr > B6 LCMV Rechallenge
Th1 Id3hi 1°Txr > B6 LCMV Rechallenge
Cells
87.6
12.4
57.1
42.9
Id3-GFP
e
Cells (%)
SM+CD4+ T cells (normalized)
Th1 Txr Population
Id3lo Th1
Id3hi Th1
Recall Population
f
84.1
14.3
52.9
45.7
SLAM
CXCR5
g
Txr Population
SLAMhi CXCR5lo
SLAMlo CXCR5hi
h
3.21
85.7
8.75
9.65
55.3
29.7
PD-1
i
CXCR5- PD-1-
CXCR5+ PD-1-
CXCR5+ PD-1+
j
Id3GFP/+ SM+
LCMV
1° Txr
-1 0
24-27
Harvest Tissues
C57BL/6
k
Total Id3GFP/+ SMARTA
SLAMhiCXCR5lo Th1
SLAMloCXCR5hi Tfh
20.1
71.2
64.8
14.2
13.1
78.3
CD45.1
anti-GFP (Id3)
l
% of Max
CD127
TCF1
BCL2
m
MFI
CD127
TCF1
BCL2

Figure 2.5. Id3-GFP^{hi} cells are a transcriptionally distinct T_H1 memory cell subset. (a) Averaged (3 independent replicates) mRNA expression by expression plot of T_H1 Id3-GFP^{hi} vs Id3-GFP^{lo} cells from bulk RNA-sequencing. Highlighted genes (grey) indicate fold change ≥ 1.75 (b) GSEA of T_{CM} [97] and T_{cmp} [24] signature in memory CD4$^+$ Id3-GFP^{hi} T_H1 vs Id3-GFP^{lo} T_H1 cells. (c-g) SMARTA CD4$^+$ T cells were adoptively transferred into congenically distinct hosts 1 day before infection with LCMV. Splenocytes were harvested and SMARTA CD4$^+$ T cells were sorted at 7, 21, and 41 days of infection and subsequently processed for scRNA-seq with the 10x Genomics platform. (c) uMAP plot of samples colored by sample ID (d) uMAP plot of relative enrichment of memory T_H1 (left) and T_{FH} (right) gene-signatures generated from bulk-RNA sequencing of sorted T_{FH} and T_H1 memory cells (e) Relative expression of indicated genes including known T_H1- and T_{FH}-associated genes. (f) uMAP plot of samples colored by sample ID. (g) Relative enrichment of memory T_H1 (left) and T_{FH} (right) gene-expression signatures generated from bulk-RNA sequencing of sorted T_{FH} and T_H1 memory cells. (h) Relative enrichment of T_H1 Id3-GFP^{lo} (left) and Id3-GFP^{hi} (right) memory gene-expression signatures generated from bulk-RNA sequencing.

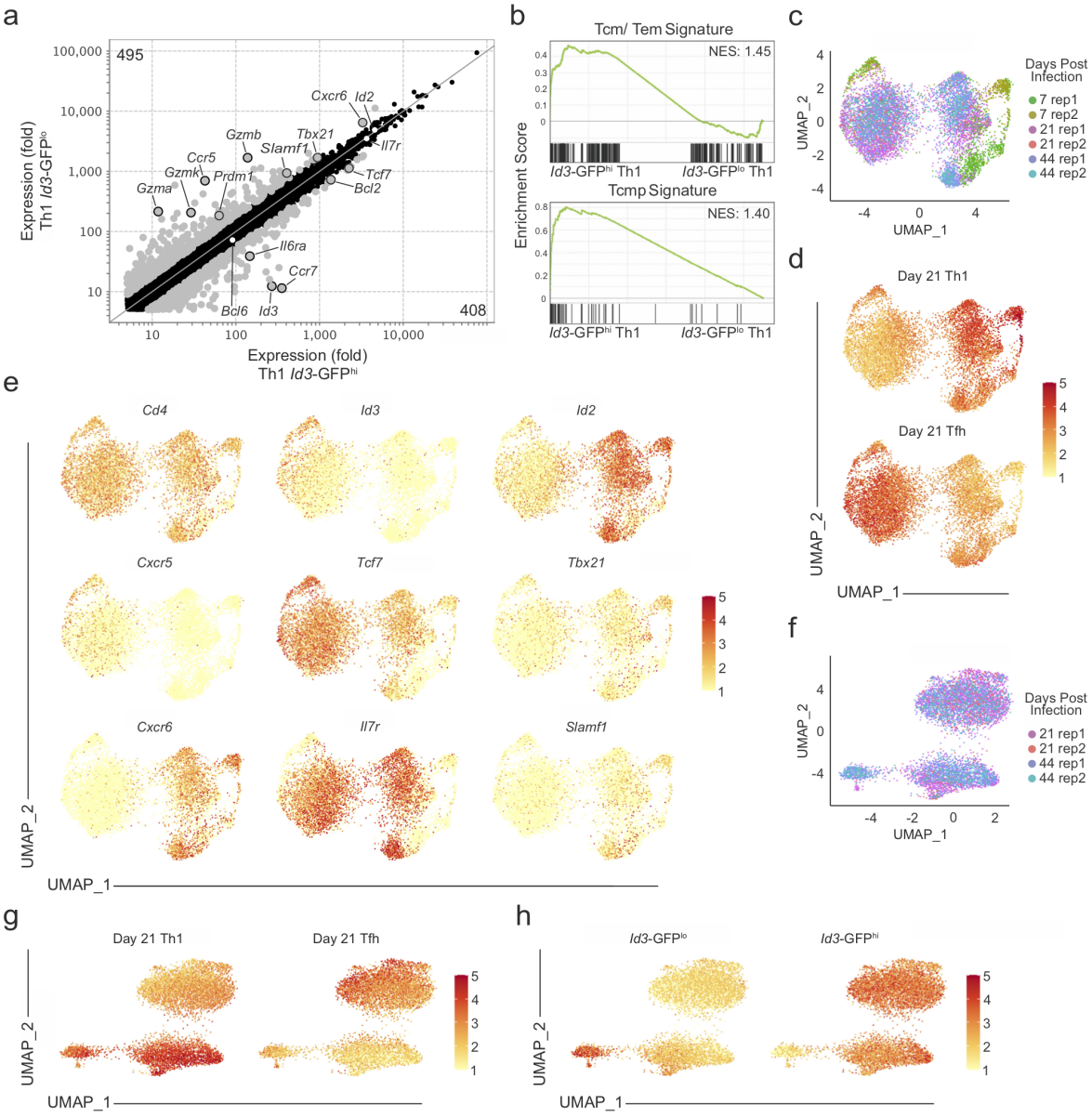
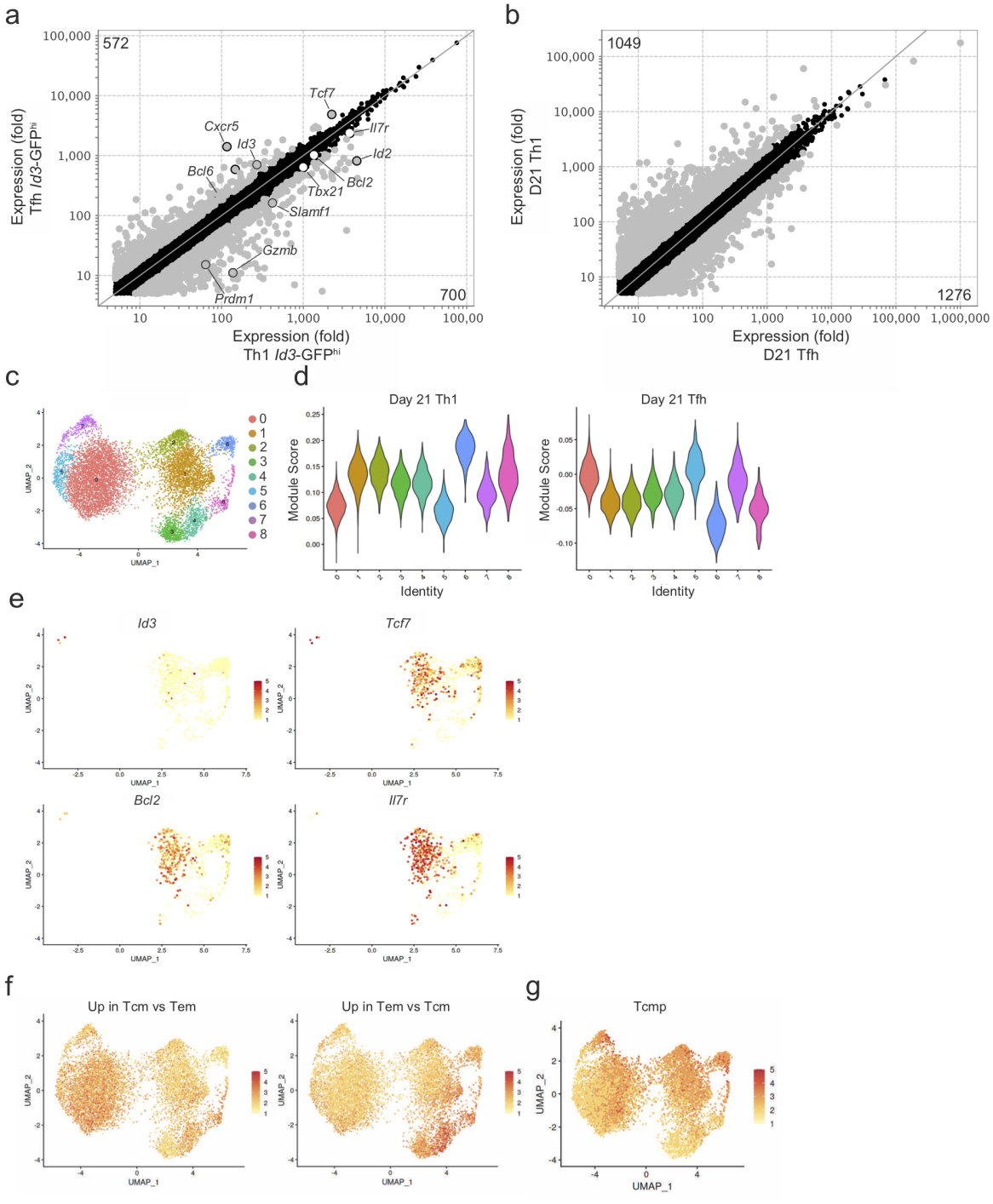

a
495
408
Cxcr6
Id2
Gzmb
Tbx21
Slamf1
Il7r
Ccr5
Gzmk
Prdm1
Tcf7
Gzma
Bcl2
Il6ra
Ccr7
Bcl6
Id3
Expression (fold)
Th1 Id3-GFPlo
Expression (fold)
Th1 Id3-GFPhi
b
Tcm/ Tem Signature
NES: 1.45
Tcmp Signature
NES: 1.40
Enrichment Score
Id3-GFPhi Th1
Id3-GFPlo Th1
c
Days Post Infection
7 rep1
7 rep2
21 rep1
21 rep2
44 rep1
44 rep2
UMAP_1
UMAP_2
d
Day 21 Th1
Day 21 Tfh
e
Cd4
Id3
Id2
Cxcr5
Tcf7
Tbx21
Cxcr6
Il7r
Slamf1
f
g
h
Id3-GFPlo
Id3-GFPhi

Figure 2.6. Single-cell RNA-seq and bulk-RNA sequencing of memory SMARTA CD4$^+$ T cells. (a) Averaged (3 independent replicates) mRNA expression by expression plot of >day 30 memory T_{FH} Id3-GFP^{hi} vs T_H1 Id3-GFP^{hi}. Highlighted genes (grey) indicate fold change $\geq$1.75. (b) Averaged (4 independent replicates) mRNA expression by expression plot of day 21 T_{FH} vs T_H1 memory cells. Highlighted genes (grey) indicate those of the T_H1 (left) or T_{FH} (right) memory gene- signatures defined by FC $\geq$ 2. (c) uMAP plot of samples colored by cluster. (d) Violin plots of relative enrichment of memory T_H1 (left) and T_{FH} (right) gene-signatures generated from bulk- RNA sequencing of sorted T_{FH} and T_H1 memory cells. (e) Relative expression of Id3 and indicated memory-associated genes on uMAP plots containing cells enriched for the T_H1 memory gene- signature. For filtering, violin plots were used to set thresholds to select memory cells with an enriched T_H1 cell gene-signature (>0.10) as well as a diminished T_{FH} cell gene-signature (<-0.05). (f) Relative enrichment of T_{CM} (left) and T_{EM}(right) (Milner et al., PNAS, 2020) [97] or (g) T_{cmp} (Ciucci et al., Immunity, 2019) [24] gene-signatures on uMAP plots generated from scRNA-seq of sorted SMARTA CD4$^+$ T cells over the course of an LCMV infection (day 7, 21, and 41).


a
572
700
Tcf7
Il7r
Cxcr5
Id3
Id2
Bcl6
Bcl2
Tbx21
Slamf1
Gzmb
Prdm1
Expression (fold)
Tfh Id3-GFPhi
Expression (fold)
Th1 Id3-GFPhi
b
1049
1276
Expression (fold)
D21 Th1
Expression (fold)
D21 Tfh
c
d
Day 21 Th1
Day 21 Tfh
Module Score
Identity
e
Id3
Tcf7
Bcl2
Il7r
UMAP_1
UMAP_2
f
Up in Tcm vs Tem
Up in Tem vs Tcm
g
Tcmp

Chapter 3

Unbiased approach for investigating putative transcriptional regulators of memory CD4$^+$ T cells

3.1 Introduction

The previous chapter identified a marker of CD4$^+$ memory T cells by borrowing knowledge from CD8$^+$ memory T cells. In this chapter, we elucidated an alternate approach towards investigating transcriptional regulators of memory CD4$^+$ T cells. We first explored uniquely accessible regions of the genome for effector and memory T_H1 as well as T_{FH} cells using ATAC- (Assay for Transposase-Accessible Chromatin-) sequencing. Then utilizing a novel bioinformatics analysis that combines bulk RNA-sequencing and ATAC-sequencing, we are able to predict putative transcription factors that may serve a role in the biology of memory T_H cells. To validate these putative regulators, we employed CRISPR/Cas9 technology to knock out potential genes of interest in SMARTA T cells prior to adoptive cell transfer and infection. A similar approach was utilized by Yu et al. [150] to successfully identify two TFs important in the differentiation of CD8$^+$ TE and MP effector T cells. By leveraging sequencing technology and bioinformatics approaches, we are equipped to further expand our understanding of CD4$^+$ memory T cell populations.

3.2 Effector and Memory T_H1 and T_{FH} cells possess unique epigenetic signatures

To investigate chromatin accessibility of T_H1 and T_{FH} cell as they differentiated to effector and memory populations, we adoptively transferred SMARTA T cells into C57BL/6J recipients followed by LCMV-Armstrong infection (Figure 3.1). At effector and memory timepoints, we sorted out T_H1 and T_{FH} populations for ATAC-sequencing as well as bulk RNA-sequencing. Utilizing the ATAC-seq results, we identified accessible regions of the genome

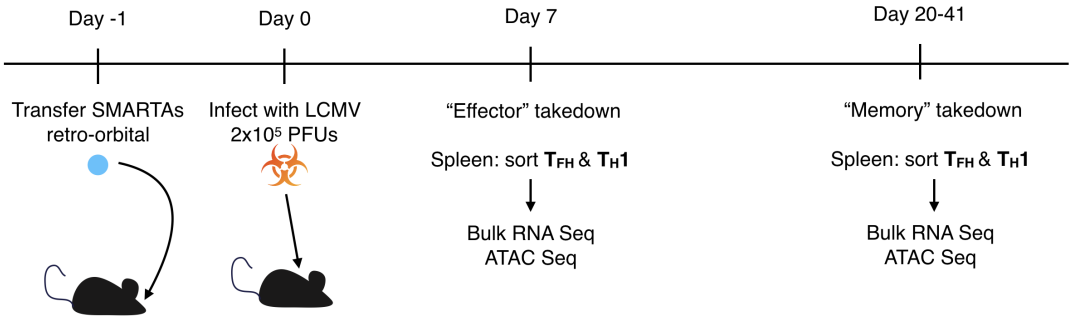


Figure 3.1. Experimental schematics for RNA- and ATAC-sequencing. Data from effector and memory SMARTA T cells will be used in the PageRank workflow to generate putative regulators of $CD4^+$ memory T cells.

using the ENCODE ATAC-seq pipeline [26]. The pipeline generated a list of all accessible peaks from the samples as well as quantify the number of reads within each peak. Replicate samples for each subset showed strong similarity based on Spearman correlation (Figure 3.2a). Interestingly, principal component analysis (PCA) of ATAC data revealed that while D20 T_{FH} cells exhibited a chromatin state more similar to D7 T_{FH} effector cells, D20 T_H1 cells actually correlated more closely with D40 T_H1 memory cells (Figure 3.2b). This pattern of D20 T_{FH} cells being more similar to effector T_{FH} cells, while D20 T_H1 cells being more similar to memory T_H1 cells was similarly replicated by PCA of RNA-seq data of the same subsets (Figure 3.2c), perhaps suggesting that the T_{FH} lineage take a longer time course to fully differentiate into a long-lived memory cell-state, perhaps due to the prolonged presence of germinal centers after viral clearance. Subsequently we identified 11,940 shared differentially accessible peaks between

D7, D20, and D40 cells for the T_H1 and T_{FH} lineages. These peaks were then annotated and subsequently filtered based on differential expression patterns in RNA-Seq before being clustered based on expression profile similarities using k-means clustering (Figure 3.2d). Clustering results of filtered differentially accessible peaks reaffirmed the previous finding that D20 T_{FH} cells more closely resemble T_H cells from the effector timepoint, while the D20 T_H1 cells look more similar to D40 memory T_H cells (Figure 3.2d). Peaks in differentially accessible regions (DAR) are enriched with intronic and distal intergenic regions compared to all peaks (Figure 3.2e), suggesting differentially accessible non-coding regions may regulate these populations. Collectively the data suggests $CD4^+$ T_H effector and memory cells have unique chromatin signatures that is shared by both the T_H1 and T_{FH} subsets.

3.3 PageRank analysis generates putative TFs for $CD4^+$ memory T cells

Despite being critical transcription factors for two opposing $CD8^+$ effector lineages, regulators of $CD8^+$ TE effector cells (T-bet, Id2, Irf4, Batf, Zeb2) and MP effector cells (Tcf7, Eomes, Id3, Bcl6) do not always exhibit differential mRNA expression between the two subsets [154, 65, 66, 17]. This issue makes the identification of key TFs for differing subsets via bulk RNA sequencing alone extremely challenging and inefficient. To circumvent this issue, the PageRank analysis cleverly enlists the power of Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-seq) to globally probe open chromatin regions of the genome [14, 144]. As previously noted, transcription factors can directly influence transcriptional activity via DNA binding, but can also recruit co-factors to alter the epigenetic landscape of the T cell's genome [137, 141, 151]. This epigenetic shift then allows previously inaccessible regions of the genome to become available for transcription and regulation. Thus the PageRank analysis first utilizes ATAC-seq data to scan for TF binding motifs within the open regions of chromatin to infer all potential TF binding sites and create TF networks consisting of all potential gene targets

of any known TF [106, 150]. In the next step, the workflow assigns a weight to every gene within each constructed TF network based on the level of differential expression in RNA-seq data across different cell types or conditions. Lastly, the workflow applies the PageRank algorithm onto the TF regulatory network to adjust the weight of each gene based on two criteria: 1) the number of regulated genes and 2) the weight or importance of regulated genes (Figure 3.3). In essence, the algorithm ensures that TFs that regulate more genes or regulate more important genes would receive a higher weight or GeneRank score.

3.3.1 Putative regulators of T_H1 and T_{FH} memory

To investigate novel putative transcriptional regulators of $CD4^+$ memory T cells, we first transferred SMARTA T cells into host recipients and then infected them with LCMV-Armstrong (Figure 3.1). At both effector and memory time points, we sorted out T_H1 and T_{FH} helper T cells for bulk RNA sequencing as well as ATAC sequencing (Figure 3.1). With RNA-seq and ATAC-seq data from effector and memory T_H cells from both the T_H1 and T_{FH}, we employed the PageRank analysis workflow to predict putative regulators of T_H1 and T_{FH} memory T cells. To identify potential targets, we assessed the fold change of GeneRank scores (weight generated by PageRank analysis) between memory and effector T_H cells within both the T_H1 and T_{FH} lineage. We selected a significant cutoff of fold change >2 and identified 88 putative regulators (highlighted in purple) of T_{FH} memory cells (Figure 3.4) and 64 putative regulators (highlighted in purple) of T_H1 memory cells (Figure 3.5). The top predicted targets, i.e. genes that had the highest positive fold change in GeneRank score between the memory timepoint and the effector timepoint, for the T_H1 include: Zfp957, Rhox9, Hoxc9, Rorb, and Srebf2 (Figure 3.5). Top putative regulators for T_{FH} memory are: Rorb, Hes1, Lhx5, Srebf2, and Pitx3 (Figure 3.4). The full list of putative memory regulators can be found in Figure 3.6. After generating putative regulators of memory $CD4^+$ T cells, we aimed to functionally validate these TFs in *in vivo* experiments through an loss-of-function, gene-knockout approach.

3.4 Functional validation of predicted TFs via Cas9-gRNA ribonucleoproteins

Since the advent of CRISPR (clustered, regularly interspaced, short palindromic repeats)/Cas9 (CRISPR-associated protein 9) technology, it has become the prominent tool used for genome editing. The Cas9 protein forms a ribonucleoprotein (RNP) complex with a functional guide RNA (gRNA), which contains a CRISPR RNA (crRNA) portion as well as a tracer RNA (tracrRNA) component. The crRNA contains sequences that target a specific region of the genome while the tracrRNA anchors the crRNA to the Cas9 protein. The CRISPR-Cas9 complex then searches for genomic targets by binding to DNA and unwinding the double helix to test for potential matches against the specific guide RNA [62]. Once the target sequence is found within the genome, the RNP complex facilitates a DNA break at the specific location to be repaired by cell endogenous mechanisms. The prominent repair mechanism in somatic cells is called mutagenic nonhomologous end-joining (NHEJ) which often creates insertions or deletion (InDels) at the DNA break site, leading to loss-of-function mutations in the target gene. Early adaptations of CRISPR/Cas9 editing in T cells primarily utilized viral delivery of Cas9 and gRNA [140, 79] plasmids or transfection via electroporation of gRNA/Cas9 constructs [87, 131]. These methods resulted either in low targeting efficiencies or cell toxicity due to DNA electroporation [108].

More recently, an approach utilizing electroporation of preformed Cas9/gRNA RNP exhibited positive results for rapid and efficient target-gene knock outs in primary T cells [108]. While the previous methods required transcription and translation of various CRISPR/Cas9 components prior to gene-editing, the RNP system introduces a preformed CRISPR/Cas9 complex directly into the cell via electroporation (Figure 3.7), allowing gene-editing activities to commence immediately. The RNP complexes are quickly degraded and removed within the cell, making the editing activity short-lived, consequently reducing the chances of off-target effects.

To validate putative $CD4^+$ memory T cell regulators, we adapted the RNP electroporation

protocol by Oh et al. [108] to our SMARTA T cell system (Figure 3.8). To initiate Cas9-mediated gene knockout, we first enriched and activated naive SMARTA T cells *in vitro* with anti-CD3 and anti-CD28 stimulation. After 48 hours, we begin the generation of CRISPR/Cas9 RNP complexes with crRNA against the pan T cell marker Thy1.2 (CD90.2) as well as putative memory regulators. We utilized the Neon Invitrogen system to electroporate our activated SMARTA T cells with preformed RNPs. One day post electroporation, we measured the efficiency using flow cytometry by detecting fluorescence emitted by the special crRNA we use in our RNP complexes, which contains the fluorescent label ATTO550. We then transfer the RNP-transfected SMARTA T cells into B6 hosts followed by LCMV-Armstrong infection. To investigate the effects of the CRISPR-mediated knockout on SMARTA T cells post infection, we performed immuno-phenotyping on splenocytes at both an effector and memory timepoint (Figure 3.8). An expected result from knocking out an important memory regulator would likely entail minimal consequences for effector cell generation accompanied by a significant loss of memory cell frequency.

3.4.1 Functional validation of putative regulators: *Srebf2* and *Rorb*

For functional validation of the PageRank generated putative regulators, we prioritized investigating the transcription factors Srebf2 and Rorb due to their status as top predicted memory regulator in both the T_H1 and T_{FH} lineage. While not much is known about the role of Rorb in T cells, Srebf2 has been shown to have important roles in regulating T cell lipid homeostasis and effector T cell generation [72]. To test the effects of knocking out *Srebf2* and *Rorb* in $CD4^+$ T cells, we first activated SMARTA T cells *in vitro* followed by RNP electroporation 48 hours after. As noted previously, we detect RNP electroporation efficiency via the detection of ATTO550, a fluorescent component of the RNP complex. However, since the RNP will be degraded over time, we will lose the ability to longitudinally track our RNP-tranduced SMARTA cells throughout an LCMV-Armstrong infection. To overcome this obstacle, we utilized crRNA against Thy1.2 (CD90.2) as our control baseline condition.

Thy1.2 or CD90.2 is a commonly expressed pan-immune surface marker that has no

known impact on T cell fitness [73]. More specifically, Thy1 is a Glycosylphosphatidylinositol (GPI) anchored surface protein that is highly expressed on neurons, thymocytes, and mature T cells [50]. Embryonic deletion of Thy1 results in no defects in thymocyte numbers but a reduction in mature $CD8^+$ T cells [50]. Though no physiological ligand or potential receptor has been identified in murine, *in vitro* cross-linking of Thy1 with monoclonal antibodies in the context of strong CD28 co-stimulatory signals can partially substitute for TCR signaling, a phenomenon also seen with other GPI-anchored proteins [50]. Thy1-null peripheral T cells show a modest decrease in proliferative responses to immobilized anti-CD3 mAb *in vitro* but responses to Concanavalin A or PMA plus Ca^{2+} ionophore are normal with no effects on cytokine production [50]. Therefore targeting Thy1.2 will not impact the SMARTA cell's ability to respond to LCMV infection but will allow us to detect RNP-transfected cells via the depletion of cell surface CD90.2. Depletion of Thy1.2 provides a new method of tracking of RNP-transfected cells longitudinally throughout the course of an infection *in vivo*. One day post RNP electroporation, we measure the transfection efficiency and found that $>$94% of SMARTA cells are RNP-positive (Figure 3.9a). It is worth noting that at 24 hours post RNP electroporation, we already detected reduced surface expression of Thy1.2 (Figure 3.9b), evidence of CRISPR/Cas RNP mediated gene-knockout. To confirm CRISPR-mediated disruption of *Srebf2* and *Rorb*, we extracted genomic DNA post RNP electroporation and performed sanger sequencing around the targeted region. To estimate spectrum and frequency of Crispr-mediated InDels , we utilized the Tracking of Indels by DEcomposition (TIDE) tool [13]. Though the percentage of RNP^+ ($ATTO550^+$) cells were consistently $>$94% across all samples, the actual InDel frequencies around the target site were variable (Figure 3.10). While the *Srebf2* targeting guides were able to generate InDels around the intended disruption site, with Srebf2-AC having the highest InDel efficiency at 47.6% (Figure 3.10a-d), the two *Rorb* targeting guides, AA and AB, seemed unable to induce disruptions at the target site (Figure 3.10e-h). It is worth noting here that despite $>$96% of cells displayed successful RNP transfection with *Rorb* targeting guides (Figure 3.9), actual genetic disruption was minimal. One possible explanation for this is that the target genomic region is perhaps

inaccessible in naive CD4$^+$ T cells, and thus successful RNP transfection would be unable to mediate genomic editing.

Following the assessment of RNP efficiency, we transferred the transfected SMARTA T cells into host mice followed by LCMV-Armstrong infection. Although *Srebf* and *Rorb* are predicted regulators of memory CD4$^+$ T cells, an evaluation of the effector time point is necessary as a compromised effector phase will undoubtedly impact the formation of memory populations. We detected no significant difference in effector T cell frequency aside from a slight increase in the Srebf2-AC treated group compared to the control (Figure 3.11a,b). However an analysis of absolute numbers of SMARTA effector T cells no longer reflects this difference (Figure 3.11b). To assess potential unwanted effects on the differentiation of CD4$^+$ effector T cells, we first gate on our CRISPR-transfected SMARTA T cells based on low expression of CD90.2 (Figure 3.11c). It is worth noting that we observed very consistent Thy1.2 depletion efficiency across all test conditions despite the two Rorb targeting guides being ineffective at inducing target gene disruption (Figure 3.11d). After gating on CD90$^-$ SMARTA T cells, we observed no significant changes in the frequency of T_H1, T_{FH}, or GC-T_{FH} effector CD4$^+$ T cells (Figure 3.11e-h). Collectively the data would suggest that CRISPR targeting of *Srebf2* and *Rorb* does not affect CD4$^+$ effector T cell generation or differentiation. However it is important to remember that the *Rorb* targeting conditions were unsuccessful in mediating InDels at the intended site, and therefore no definitive conclusions can be made regarding the role of *Rorb*. At the memory timepoint, we did not observe any significant differences in memory SMARTA T cell frequency or total memory SMARTA T cells between test conditions and the control (Figure 3.12a,b). However we were able to recover ample numbers of RNP-transfected memory SMARTA T cells for *ex vivo* analysis, which was previously unfeasible when using short hairpin RNA knockdown approaches. After gating on CD90$^-$ memory T cells (Figure 3.12c,d), we found no significant changes in the frequency of T_H1 and T_{FH} memory T cells (Figure 3.12e,f). Analysis of CRISPR-edited memory CD4$^+$ T cells would suggest that *Srebf2* does not affect CD4$^+$ memory T cell generation or differentiation, however it is also possible that

the deletion efficiency was not sufficient to significantly disrupt gene function and therefore no phenotypic changes were observed. Target deletion of *Rorb* appeared unsuccessful and therefore no definitive conclusions can be drawn regarding its role in $CD4^+$ memory T cells.

3.5 Discussion

While chapter 2 demonstrated how memory $CD8^+$ T cell knowledge can be leveraged for memory $CD4^+$ T cell investigations, here we reported a possible strategy for an unbiased approach on investigating novel putative transcriptional regulators of memory $CD4^+$ T cells. We first compared transcriptomic landscape and chromatin accessibility between effector and memory T_H1 and T_{FH} cells. Interestingly, we found that while D20 T_{FH} cells exhibited a transcriptomic as well as chromatin state that was more similar to D7 T_{FH} effector cells, D20 T_H1 cells actually correlated more closely with D40 T_H1 memory cells. RNA-seq data and ATAC-seq data collectively suggest that perhaps T_{FH} memory cells take a longer trajectory to reach their long-lived memory cell-state. This could be explained by the prolonged presence of germinal centers post viral clearance that serve as a source for T_{FH} effector inducing signals. We then took conserved differentially accessible regions between D7, D20, and D40 T_H1 and T_{FH} cells and filtered the DARs based on differential mRNA expression from RNA-seq. The resulting peaks, or DARs that can presumably induce transcriptional changes, were clustered based on k-means clustering and again we found that D20 T_{FH} cells more closely resemble T_{FH} cells from the effector timepoint, while D20 T_H1 cells looked more similar to D40 memory T_H1 cells.

To predict putative transcriptional regulators of $CD4^+$ memory T cells, we utilized the PageRank bioinformatics analysis which uses ATAC-seq data to scan for TF binding motifs to construct a TF to gene network. Subsequently the workflow utilizes RNA-seq data to assign a ”weight” to each gene based on differential expression. Lastly the program then re-evaluates the network and assigns TFs that either regular more genes or ”heavier” genes a higher GeneRank score. To generate putative $CD4^+$ memory T cell regulators, we calculated the fold change in

GeneRank score between memory and effector T_H1 and T_{FH} cells. The top predicted TFs have the highest fold change in GeneRank score of memory over effector populations. *Srebf2* and *Rorb* were top predicted targets for both T_H1 and T_{FH} memory lineages and therefore were selected for *in vivo* validation. Here we demonstrated an optimized gene-editing system in $CD4^+$ T cells using electroporation of CRISPR/Cas9 ribonucleoproteins. The CRISPR-RNP system leverages the genetic-editing capability of a preformed CRISPR/Cas9 ribonucleoprotein, which when transfected into the cell, can mediate immediate gene deletion and will be degraded shortly after to prevent off-target effects. The RNP complex contains a fluorescent component which can be detected via flow cytometry in order to measure transfection efficiency shortly after electroporation. We demonstrated that knockout of Thy1 can be perform in conjunction with other target genes with no defects on $CD4^+$ T cell proliferation or memory formation. This is valuable in providing the ability to longitudinally track CRISPR-edited T cells through an infection *in vivo*. The two guides targeting *Rorb* were unable to induce InDels at the target site so no definitive conclusions can be drawn about its role in $CD4^+$ memory T cells. CRISPR-mediated editing of *Srebf2* did not significantly impact effector or memory $CD4^+$ T cell generation and differentiation. However this could be due to incomplete nullification of target gene function, highlighting the important of guide design when utilizing this system.

Despite negative results on CRISPR-targeting of *Srebf2* and *Rorb*, this optimized CRISPR-RNP system can serve as a valuable resource for an efficient and simple method for mediating gene-disruption in T cells that inevitably expands our capacity to investigate T lymphocytes. This system can be used to effectively recover genomically-altered T cells at a memory timepoint post viral infection, which was previously not possible with the shRNA knockdown system. It is important to note that while RNP transfection efficiency, Thy1 knockout efficiency and actual target gene InDel frequency should all be theoretically correlated, this is not necessarily the case in practice. The *Rorb* targeting guides showed highly efficient ($>$96%) RNP transfection efficiency as well as excellent Thy1.2 knockout efficiency but very little InDel frequency. This is possibly due to poor guide design or more likely chromatin inaccessibility of the target gene.

One possible method to circumvent this issue is to perform ATAC-sequencing on naive SMARTA T cells to confirm the chromatin accessibility of intended targets in the future. Overall the optimization of this gene-editing system shows great promise in expanding the scope of our investigative capabilities. This system is also easily adaptable towards studies on other aspects of T cell biology and hopefully can yield fruitful discoveries in the future.

Chapter 3, in part, is currently being prepared for submission for publication of the material. Deng, T. Z., Goldrath A. W. The dissertation author was a primary investigator and author of this material.

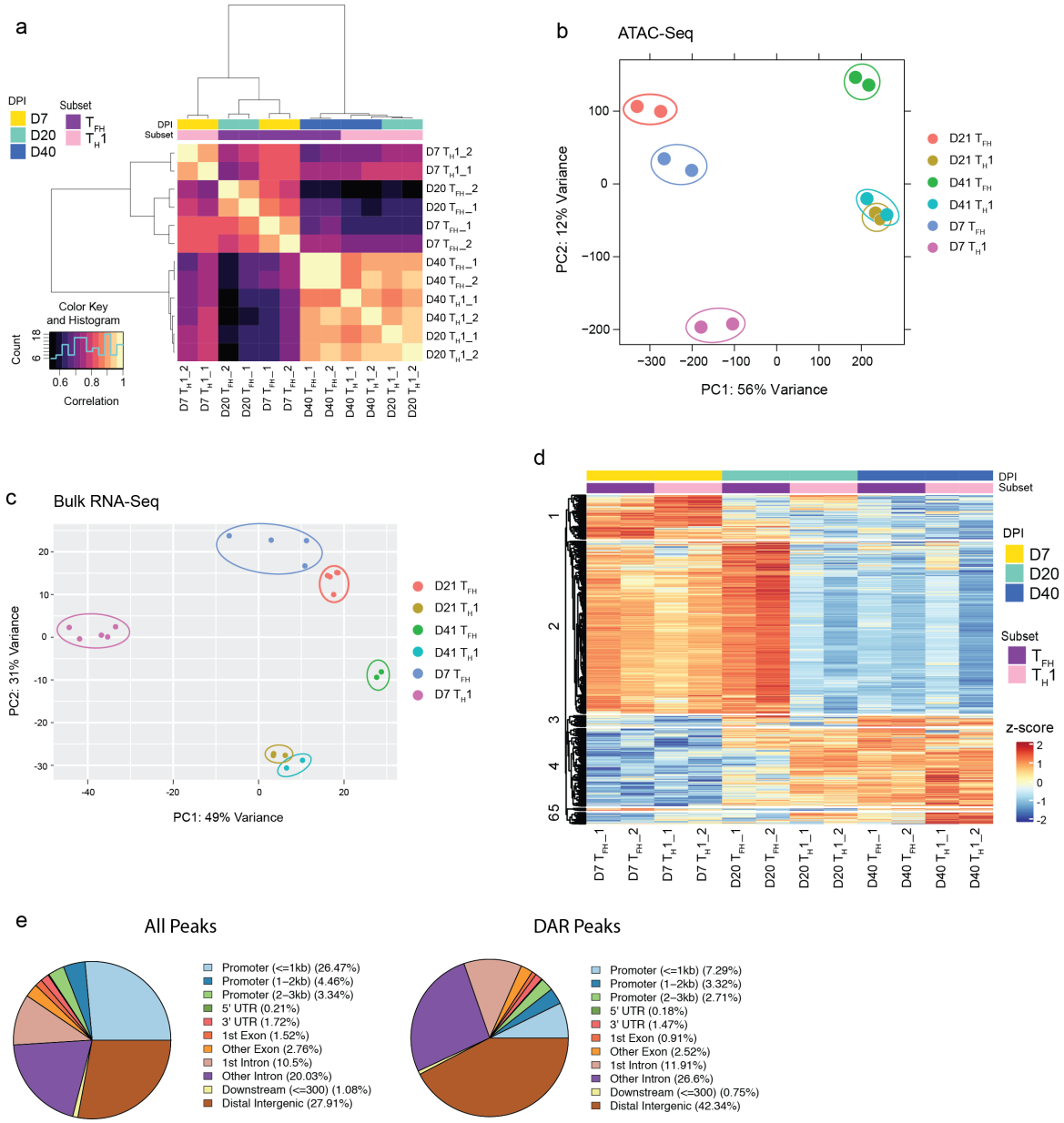


Figure 3.2. Epigenetic landscape of effector and memory CD4$^+$ helper T cells. (a) Pearson correlation for all peaks across all samples. (b) Principal component analysis of all ATAC-seq samples. (c) Principal component analysis of all RNA-seq samples. (d) Heatmap of differentially accessible regions (DAR) clustered based on k-means = 6. (e) Annotation of genomic region type for all peaks (left) and DAR peaks (right).

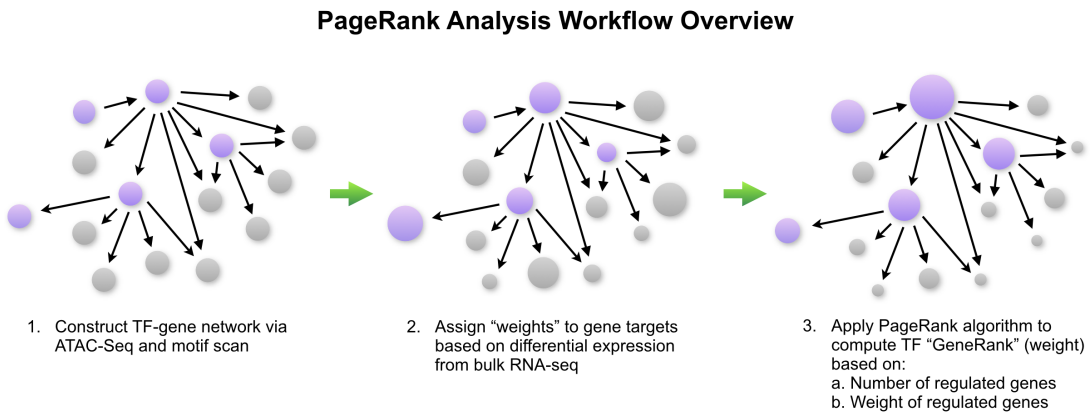


Figure 3.3. PageRank Analysis Workflow Overview.

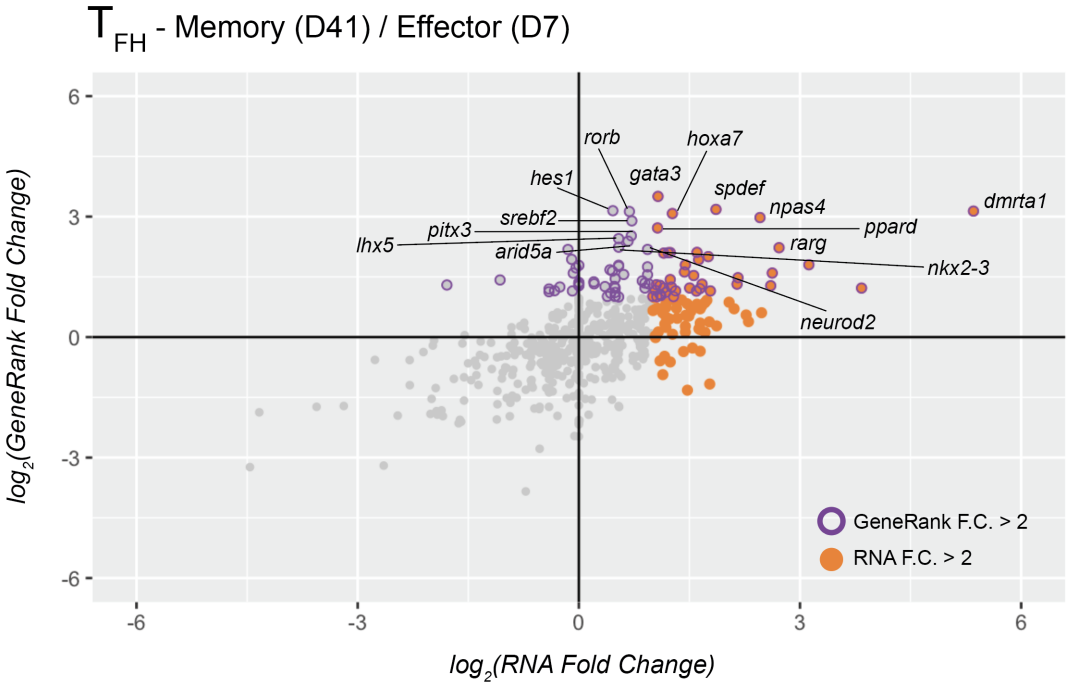


Figure 3.4. Putative regulators of memory T_{FH}. Top 15 putative regulators (highest GeneRank F.C.) are annotated.

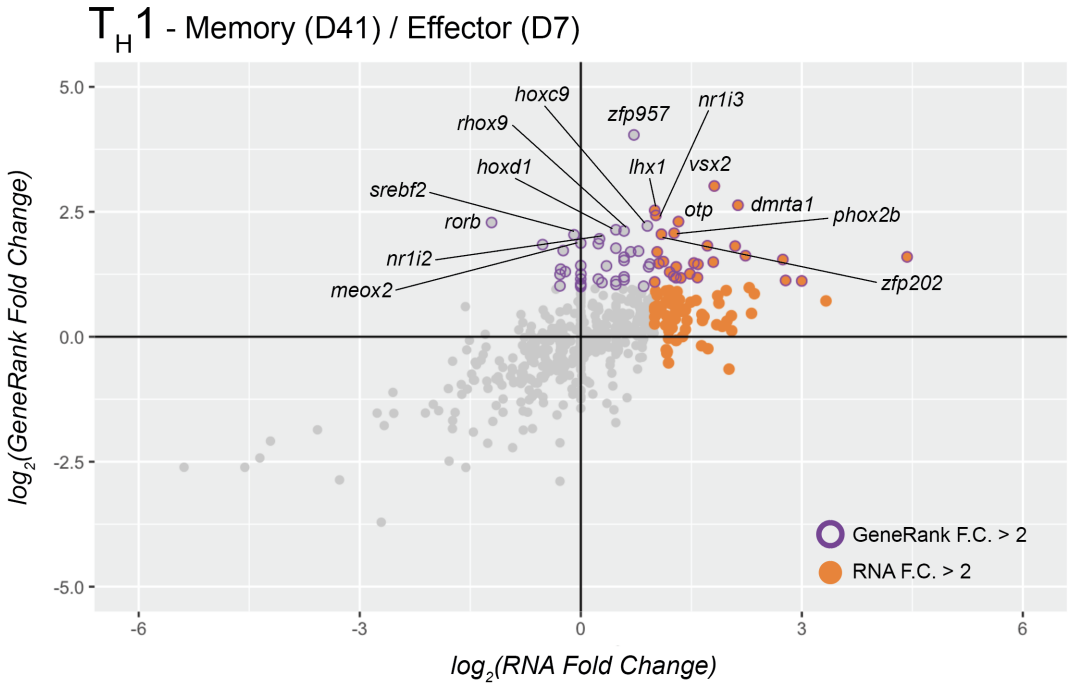


Figure 3.5. Putative regulators of memory T_H1. Top 15 putative regulators (highest GeneRank F.C.) are annotated.

Putative Memory T_{FH} regulators

Genes	GeneRank F.C.	RNA F.C.
gata3	3.50	1.07
spdef	3.18	1.86
hes1	3.15	0.46
dmrta1	3.14	5.35
rorb	3.13	0.69
hoxa7	3.08	1.27
npas4	2.98	2.46
srebf2	2.89	0.72
ppard	2.72	1.07
pitx3	2.52	0.71
lhx5	2.45	0.54
arid5a	2.39	0.67
nkx2-3	2.24	0.54
rarg	2.23	2.72
neurod2	2.18	0.93
ppara	2.18	-0.15
dbx1	2.11	1.60
zfp957	2.10	1.24
vax1	2.10	1.21
olig3	2.09	1.15
dmrtc2	2.00	1.76
glis1	1.94	-0.10
foxd2	1.92	1.62
nr3c2	1.80	3.12
hoxb13	1.80	1.44
otx2	1.79	0.54
nkx2-6	1.79	0.00
hoxc4	1.76	0.54
emx1	1.75	0.93
gata5	1.72	-0.04
zbtb4	1.68	0.42
obox1	1.64	0.46
prdm1	1.62	1.43
rfx2	1.60	2.62
alx4	1.59	-0.08
hmg20b	1.56	0.61
evx2	1.55	0.93
six4	1.53	1.56
sox6	1.48	2.16
lhx3	1.45	0.49
cdx4	1.43	1.24
mafg	1.42	-1.07
neurog1	1.40	0.86
hoxd13	1.38	0.20

Genes	GeneRank F.C.	RNA F.C.
dlx6	1.36	0.00
nkx3-2	1.35	0.00
hoxd9	1.34	0.90
foxo4	1.33	0.21
scrt1	1.32	0.00
sebox	1.32	0.95
nr6a1	1.32	2.15
klf9	1.32	1.67
dlx1	1.31	0.00
gbx1	1.30	1.05
pax5	1.30	-1.79
tgif2	1.29	1.10
foxa1	1.28	2.60
nkx2-9	1.28	0.00
hnf4a	1.26	0.48
usf2	1.26	0.35
hbp1	1.25	-0.26
tbr1	1.25	1.25
hoxa9	1.22	1.15
tbx19	1.22	1.50
pou4f3	1.22	0.90
gsc2	1.22	0.49
rara	1.22	3.83
pou3f1	1.21	1.64
hnf4g	1.20	-0.40
mbd1	1.16	1.04
kdm2b	1.15	-0.33
prrx1	1.15	1.31
zic1	1.15	1.78
pitx2	1.15	-0.09
myt1l	1.15	1.59
nkx2-4	1.13	-0.40
pou1f1	1.11	1.18
msx1	1.11	0.49
hoxb5	1.09	0.43
zfp263	1.06	1.10
pou3f3	1.04	1.11
crx	1.02	0.49
nkx2-1	1.02	0.49
sox8	1.02	1.00
nr2c1	1.01	0.39
zfp105	1.01	1.28
hsf1	1.00	1.06
six1	1.00	0.54

Putative Memory T_H1 regulators

Genes	GeneRank F.C.	RNA F.C.
zfp957	4.04	0.72
vsx2	3.02	1.81
dmrta1	2.63	2.13
lhx1	2.53	1.00
nr1i3	2.43	1.02
otp	2.30	1.33
rorb	2.29	-1.21
hoxc9	2.22	0.90
hoxd1	2.14	0.48
rhox9	2.12	0.59
phox2b	2.07	1.27
zfp202	2.05	1.09
srebf2	2.04	-0.09
nr1i2	1.96	0.25
meox2	1.87	0.00
hoxa1	1.86	0.24
mnx1	1.84	-0.52
msx1	1.82	1.72
prdm11	1.81	2.09
insm1	1.77	0.48
barhl2	1.72	-0.24
dmrt2	1.71	0.78
lbx1	1.70	0.68
dlx3	1.70	1.04
zfp523	1.62	2.24
foxp3	1.60	4.43
foxl1	1.59	0.59
tef	1.54	2.74
hoxd4	1.53	0.59
dbx2	1.50	1.12
bhlha15	1.49	1.80
six2	1.48	1.54
zfp740	1.47	1.06
lbx2	1.46	0.94
zic2	1.45	1.59
barhl1	1.42	0.00
hmga2	1.42	0.35
scrt1	1.40	0.92
nr2f6	1.40	1.29
pax3	1.35	-0.27
myt1l	1.30	-0.21
rarg	1.29	1.21
mzf1	1.26	1.48
hsf2	1.25	0.00

Genes	GeneRank F.C.	RNA F.C.
barx2	1.25	-0.28
zscan20	1.21	1.26
foxb1	1.20	0.59
tcf7	1.19	1.58
zbtb7b	1.18	1.29
arid5a	1.18	1.36
hoxb5	1.15	0.00
gbx2	1.15	0.24
tlx2	1.14	0.59
hic1	1.13	2.78
bcl6	1.12	3.00
barx1	1.11	0.48
foxa1	1.10	1.00
nr1h2	1.08	0.29
irx3	1.06	0.00
hoxa6	1.04	0.48
hoxd13	1.03	0.00
msx2	1.02	-0.28
hsf1	1.01	0.85
dmrt1	1.00	0.00

Figure 3.6. Complete list of putative regulators for memory CD4$^+$ T cells. All values displayed were $Log_2()$ transformed. Transcription factors are sorted based on highest fold change in GeneRank from memory over effector T cells. Fold change in RNA expression also shown.

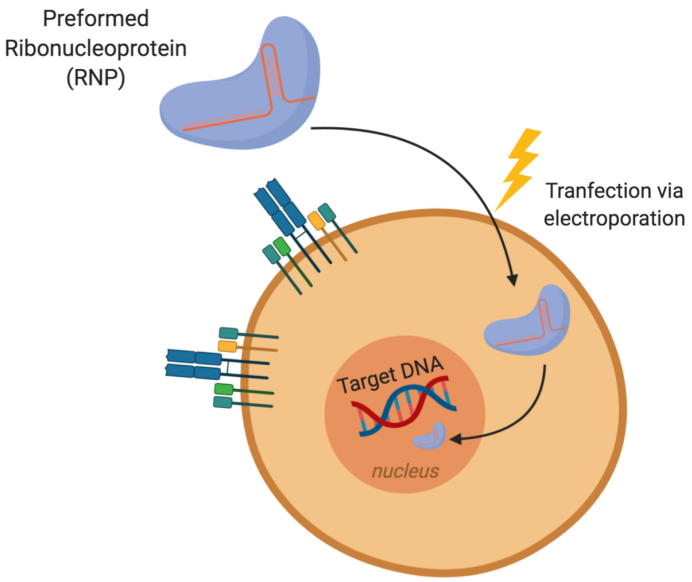


Figure 3.7. Electroporation of CRISPR/Cas9 ribonucleoprotein complex. When preformed RNP is introduced via electroporation into T cells, the enzyme complex rapidly starts cutting targeted genomic DNA.

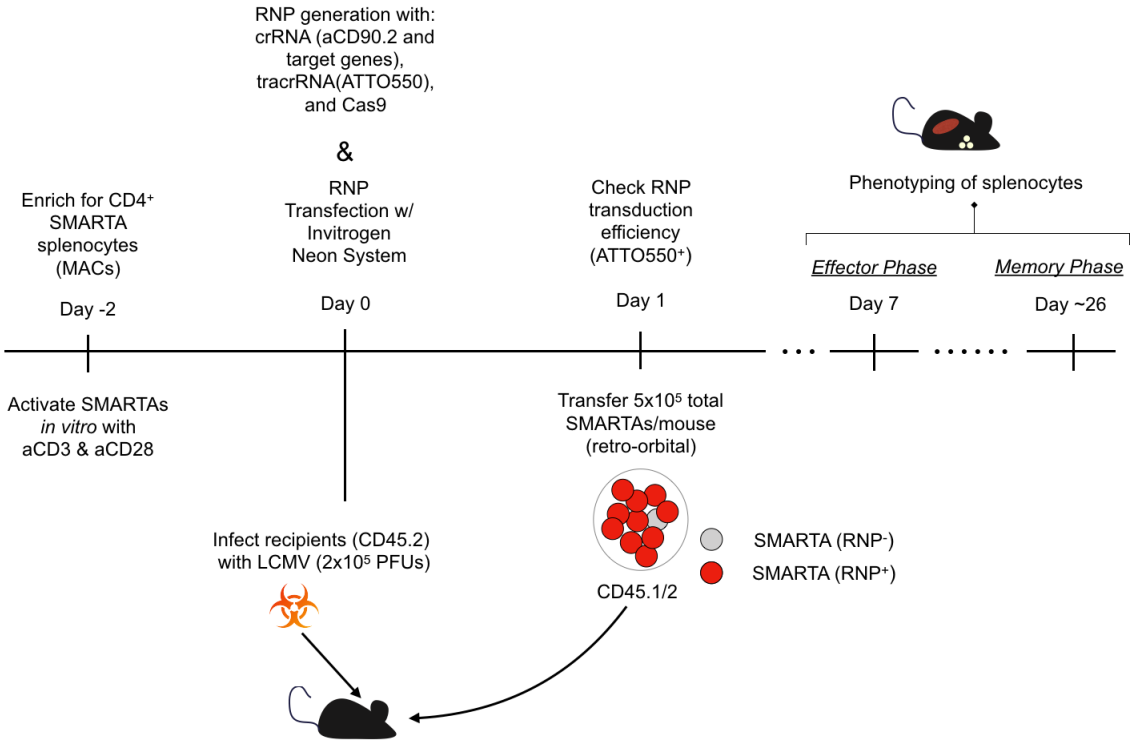


Figure 3.8. Experimental schematics for validating putative regulators of memory CD4+ T cells.

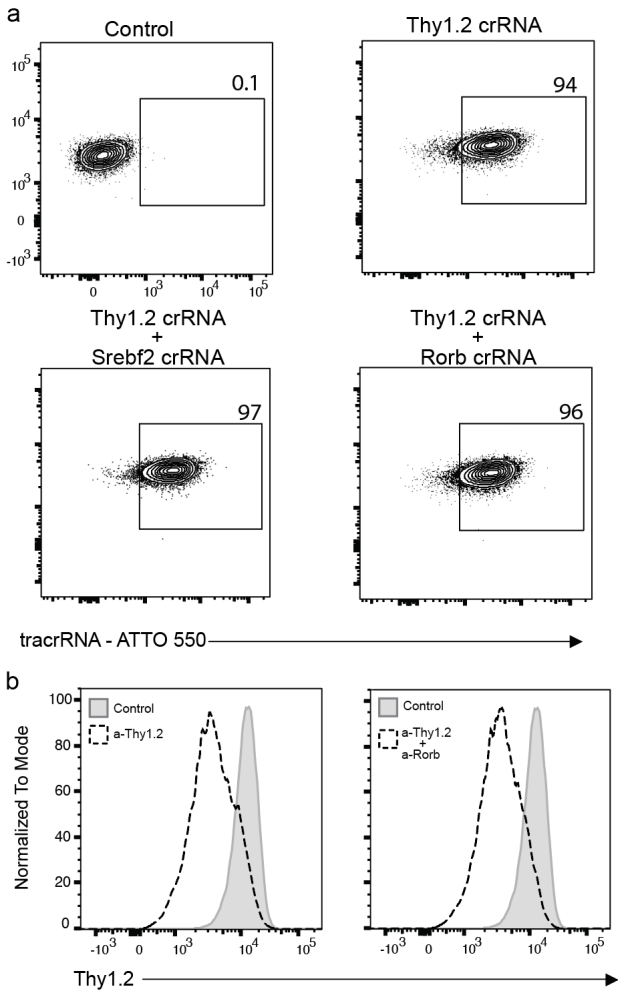


Figure 3.9. CRISPR/Cas RNP Electroporation Efficiency. (a) Representative flow plots showing fluorescent expression level of ATTO550 24 hours post RNP electroporation. (b) Representative flow histogram depicting surface expression of Thy1.2 24 hours post RNP electroporation.

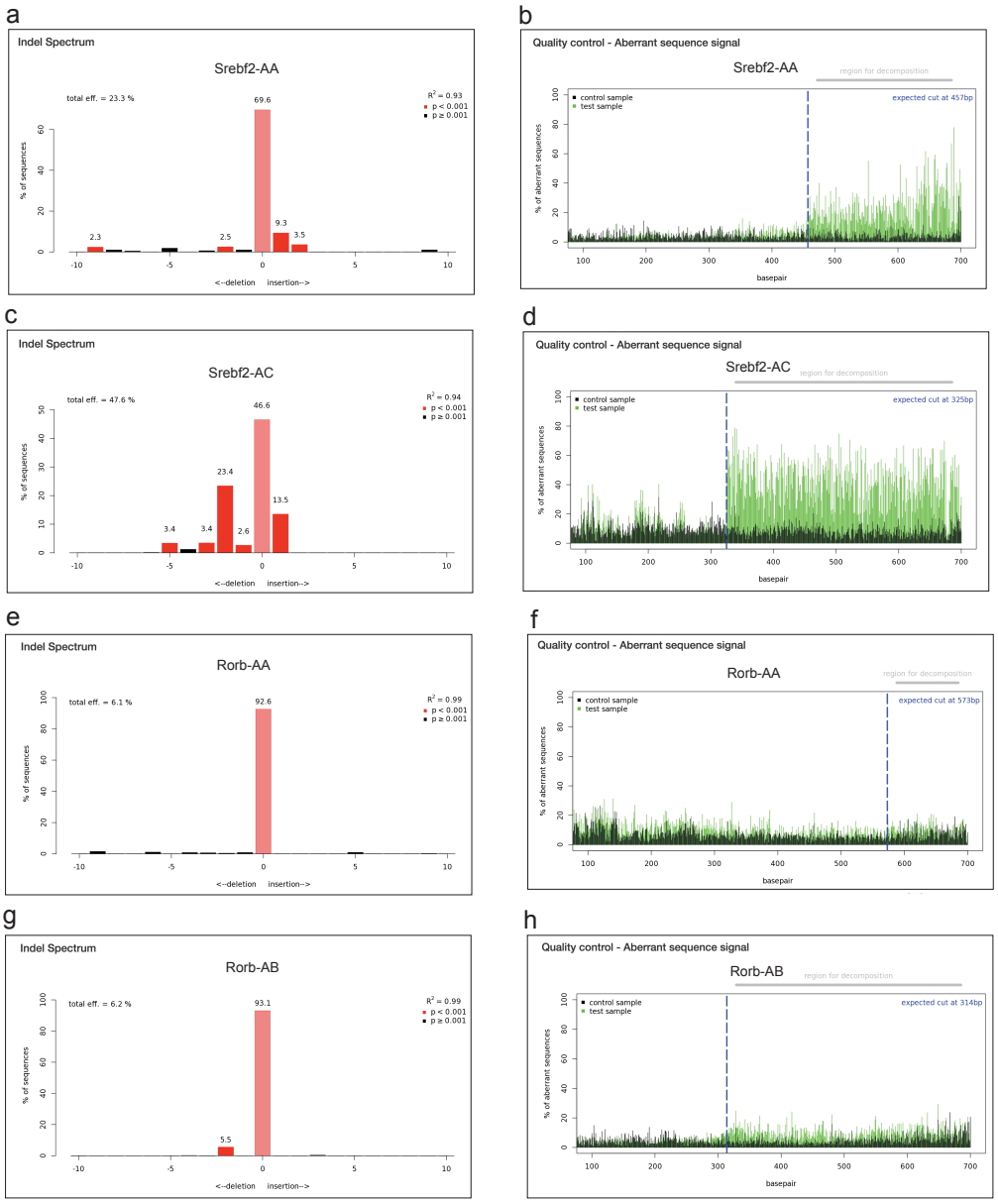


Figure 3.10. TIDE validation of Cripsr-mediated InDels. (a,c,e,g) InDel spectrum plot depicting frequency of small InDels created around the target side; Bars to the left of 0 represent the number of nucleotide deletions while bars to the right of 0 represent the number of nucleotide insertions, frequency of such edits are represented at the top. (b,d,f,h) Aberrant sequence signal plot representing the variability of nucleotide sequence between control (black) and test (green). Increased test (green) signals after the expected cut site is representative of accurate disruption at the intended site.

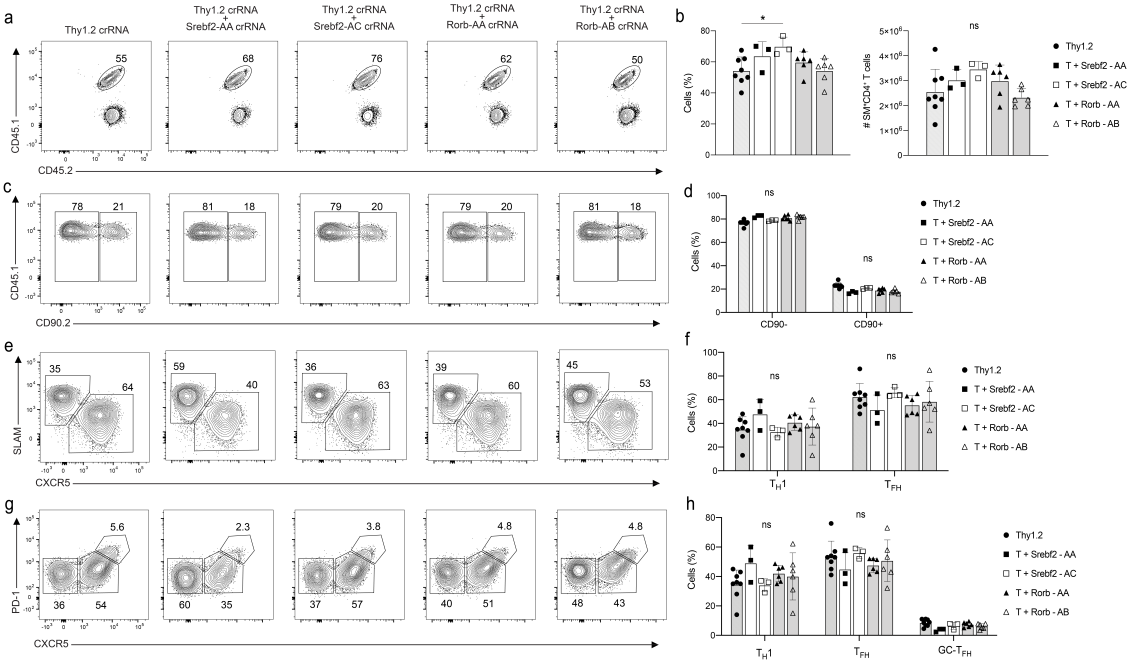


Figure 3.11. Effector timepoint analysis of CRISPR-edited T cells. (a) Flow cytometric analysis on frequency of donor SMARTA T cells at the effector timepoint. (b) Frequency of SMARTA T cells among $CD4^+$ T cells (left) and total number of SMARTA T cells recovered from the spleen (right). (c) Representative flow plots on CD90.2 expression. (d) Frequency of $CD90^+$ and $CD90^-$ SMARTA T cells. (e) Representative flow plots for $SLAM^+CXCR5^-$ (T_H1) and $SLAM^-CXCR5^+$ (T_{FH}). (f) Frequency of T_H1 and T_{FH} effector T cells. (g) Representative flow plots for $CXCR5^-PD\text{-}1^-$ (T_H1), $CXCR5^+PD\text{-}1^-$ (T_{FH}) and $CXCR5^+PD\text{-}1^+$ (GC-T_{FH}). (h) Frequency of T_H1, T_{FH} and GC-T_{FH} effector T cells. *$P<0.05$, (two-tailed unpaired Student's T test).

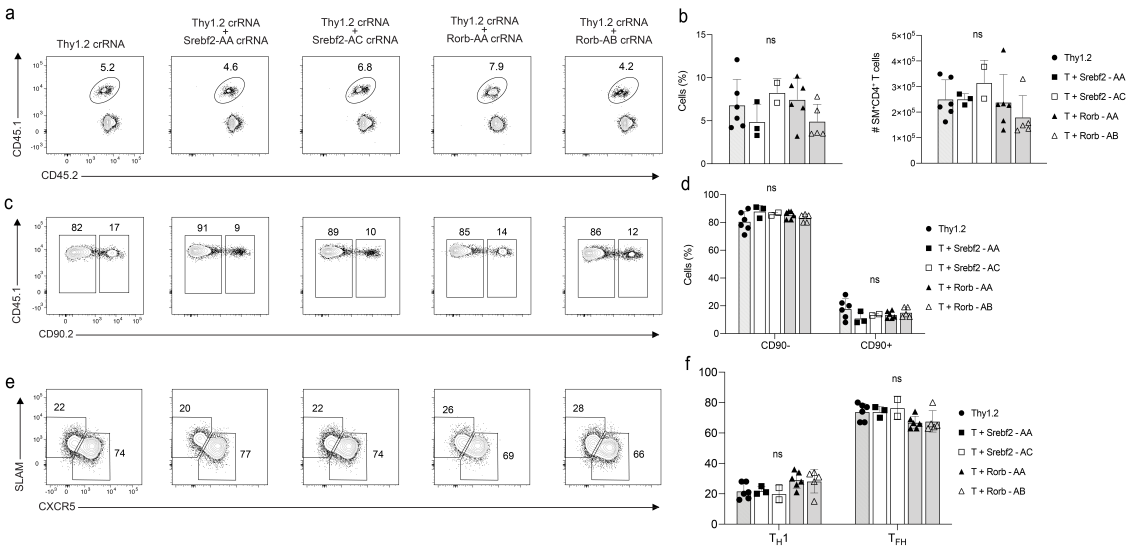


Figure 3.12. Memory timepoint analysis of CRISPR-edited T cells. (a) Flow cytometric analysis on frequency of donor SMARTA T cells at the memory timepoint. (b) Frequency of SMARTA T cells among $CD4^+$ T cells (left) and total number of SMARTA T cells recovered from the spleen (right). (c) Representative flow plots on CD90.2 expression. (d) Frequency of $CD90^+$ and $CD90^-$ SMARTA T cells. (e) Representative flow plots for $SLAM^+CXCR5^-$ (T_H1) and $SLAM^-CXCR5^+$ (T_{FH}). (f) Frequency of T_H1 and T_{FH} memory T cells.

Chapter 4

Conclusion

A typical antigen-driven T-cell response against an acute infection can be broadly broken down into two main phases: a short-lived effector phase where large numbers of effector T cells will have direct functional response against the infection followed by a contraction phase where majority of the effector cells will die leaving behind only a small portion of long-lived memory T cells that persists up to decades post antigen clearance. This memory T cell population is especially important against re-infection with the same pathogen because it ensures both a rapid and robust secondary response against the recurring pathogen, and thus is pivotal in conferring long-term immunity against harmful pathogens. T lymphocytes are further subcategorized into two main cell types based on expression of cell-surface molecules CD4 and CD8. The primary function of $CD8^+$ cytotoxic T cells is to eradicate infected cells while $CD4^+$ helper T cells can differentiate into multiple functionally distinct helper cell types depending on the type of infection and the cytokines present during their activation [95]. To ensure that each class of pathogen is countered with the appropriate immune response, early host-pathogen interactions will result in an infection environment that directs differentiation of naive $CD4^+$ T cells to acquire specific helper functions. These helper cell types have been described to each have their own unique transcriptional regulators as well as signature effector mediating cytokines.

Memory T cells have been under fervent investigation in the past decades due to their indispensable role in mediating vaccine-based immunity. While significant advances have been

made in understanding the biology of $CD8^+$ memory T cells, mechanisms underlying $CD4^+$ memory T cell remain relatively uncharted compared to the $CD8^+$ lineage. Two main factors have contributed to this knowledge deficit; firstly, $CD4^+$ T cells are inherently less proliferative and $CD4^+$ memory T cell population appears to decline following antigen clearance, while the $CD8^+$ memory T cell population is typically stable [40, 143, 1]. Secondly, the functional breadth of the $CD4^+$ T cells combined with documented instances of lineage interconversion between helper T cell lineages have added substantial complexity in regards to memory studies.

Current literature remains conflicted in whether there is a single $CD4^+$ memory T cell precursor that gives rise to secondary T_H subsets or if there are memory precursor cells within each helper lineage. In chapter 2, we reported that expression of Id3, an inhibitor of E protein transcription factors, definitively identifies a memory-precursor population within both the $CD4^+$ T_{FH} and T_H1 helper lineages. Id3 expressing T_H memory cells exhibited greater potential in response to secondary infection than their $Id3^{lo}$ counterparts. Notably, a small subset of T_H1 memory cells expressing Id3 exhibited enhanced expansion upon response to pathogen, gives rise to both T_H1 and T_{FH} secondary effector cell populations, and is enriched for key molecules associated with memory potential when compared to $Id3^{lo}$ T_H1 cells. Thus, Id3 serves as an important conserved marker of multipotency for $CD4^+$ memory T cells. Identification of such markers allow for further investigation into possible mechanisms underlying the cell-fate decisions between an $Id3^{hi}$ versus an $Id3^{lo}$ memory phenotype, perhaps shedding light on broader principals of $CD4^+$ memory T cell formation.

While chapter 2 identified a universal marker of $CD4^+$ memory T cells by leveraging knowledge from $CD8^+$ memory T cells, chapter 3 highlighted a possible strategy for an unbiased approach towards investigating putative regulators of $CD4^+$ memory T cells. To generate potential transcriptional regulators for memory helper T cells, we employed the PageRank bioinformatics analysis which utilizes RNA expression as well as chromatin accessibility data to predict potentially important transcription factors for memory $CD4^+$ T cells. Top predicted regulators were taken for functional *in vivo* validation utilizing an optimized CRISPR/Cas9

ribonucleoprotein electroporation protocol. Despite negative results on putative targets, the RNP protocol provides a reliable and highly efficient method for achieving loss-of-function knockout deletions in $CD4^+$ T cells. This system can be easily adapted toward other T cell investigations as well which further empowers our ability to explore different aspects of T cell biology. We also found that by deleting surface marker Thy1 in conjunction with potential targets, we were able to successfully identify and recover genomically-edited cells throughout an *in vivo* infection, which was not feasible with the shRNA system. Perhaps future exploration into putative regulators of $CD4^+$ T cells could employ the CRISPR RNP protocol on Id3-GFP reporter SMARTA T cells to probe possible effects on Id3 expression. Even further, perhaps the PageRank analysis should be performed specifically on sorted $Id3^{hi}$ and $Id3^{lo}$ memory T_H populations to investigate possible conserved regulators of Id3 expression in $CD4^+$ memory T cells. With concurrent advancements in RNA-seq, ATAC-seq and CRISPR genome editing technologies, we are empowered with more effective and efficient ways to explore T cell biology. By identifying a potent marker of $CD4^+$ memory, we can leverage efficient gene-editing systems to further map out the molecular networks underlying memory $CD4^+$ T cell formation in the hopes of improving future vaccine design.

Appendix A

Materials and Methods

Mice

All mice were housed under specific pathogen-free conditions in an American Association of Laboratory Animal Care–approved facility at the University of California, San Diego (UCSD), and all procedures were approved by the UCSD Institutional Animal Care and Use Committee. Id3- GFP mice (37), SMARTA mice (38) (with transgenic expression of an I-Ab-restricted TCR specific for LCMV glycoprotein amino acids 66–77) and recipient C57BL/6J mice were either bred at UCSD or received from The Jackson Laboratory.

T cell transfer and infection

Naive CD45.1+ or CD45.1.2+ SMARTA CD4+ T cells (25,000 cells/mouse) were adoptively transferred into congenically distinct wildtype C57BL/6J recipients 1 day before infection with 2 x 105 plaque-forming units (PFU) of LCMV-Armstrong, injected intraperitoneally.

Cell Preparation and Flow cytometry

Single-cell suspensions of spleen were prepared by standard mechanical disruption. Surface staining for flow cytometry was performed with monoclonal antibodies against CD4 (RM4-5, 1:400), CD45.1 (A20, 1:400), CD45.2 (104, 1:400), B220 (RA3-6B2, 1:400), PD-1 (J43, 1:400), SLAM (TC15-12F12.2, 1:400), CD4 (GK1.5, 1:400), CD127 (A7R34, 1:400) and CCR7 (4B12, 1:200). Staining was done for 30 min at 4 °C in PBS supplemented with 0.5% bovine serum albumin and 0.1% sodium azide, unless specified otherwise. CCR7 staining was

completed prior to other surface markers at 37°C for 45 min. CXCR5 staining was performed using purified anti- CXCR5 (SPRCL5, 1:50; Invitrogen) for 30 min, followed by PE-Cy7- or BV510-labeled streptavidin (1:1000, eBioscience) at 4°C. Intracellular staining was performed with monoclonal antibodies to Bcl2 (clone 3F11; BD Pharmingen, 1:20), TCF1 (clone C63D9; Cell Signaling, 1:200), and polyclonal antibodies against GFP (cat. A21331; Invitrogen) using the Foxp3 ICS kit according to manufacturer's instructions (eBioscience). Stained cells were analyzed using LSRII, LSRFortessa or LSRFortessa X-20 (BD) and FlowJo software (TreeStar). All sorting was completed on a FACSAria (BD Biosciences).

Bulk RNA-seq Library construction and Sequencing

Sorted cell lysates (5l) were used for Smart-seq2 library construction, prepared as previously described (39, 40) with slight modifications. Briefly, total RNA was captured and purified on RNAClean XP beads (Beckman Coulter). Polyadenylated mRNA was then selected using an anchored oligo(dT) primer (5–AAGCAGTGGTAT CAACGCAGAGTACTVN-3) and converted to cDNA via reverse transcription. First strand cDNA was subjected to limited PCR amplification followed by Tn5 transposon based fragmentation using the Nextera XT DNA Library Preparation Kit (Illumina). Samples were then PCR amplified for 12 cycles using barcoded primers such that each sample carries a specific combination of eight base Illumina P5 and P7 barcodes and pooled together prior to sequencing. Smart-seq paired-end sequencing was performed on an Illumina NextSeq500 (two full NextSeq runs per batch of 96 samples, for 10M raw reads/sample on average) using 2 x 38bp reads with no further trimming.

10x Genomics library preparation and sequencing

Sorted cells were washed and resuspended in phosphate-buffered saline and 0.04% (w/v) bovine serum albumin per the manufacturer's guidelines. Single-cell libraries were prepared according to the protocol for 10x Genomics for Single Cell V(D)J and 5 Gene Expression. About 10,000 sorted SMARTA cells were loaded and partitioned into Gel Bead In-Emulsions. scRNA libraries were sequenced on a HiSeq4000 (Illumina).

Single-Cell RNA-Seq Analysis

scRNA-seq analysis was performed using cellranger software and Seurat version 3.5.1 in R Studio. Cellranger was used with default parameters. Seurat Analysis of 10x counter matrices was done by following these steps: low-quality cells, identified by percent mitochondria ¡ 10, nFeatures RNA ¡ 200 or ¿ 3,000, were removed, counts were normalized with FastMNN, dimensionality reduction and cluster identification were done with uMAP (dims = 1:30), FindNeighbors (dims = 1:30), and FindClusters (resolution = 0.6). FindAllMarkers function with default default parameters and min.pct = 0.25 and logfc.threshold = 0.25. Overlay of gene signatures onto single cell data was done with AddModuleScore. Statistical methods. Statistical tests were performed using Prism (7.0/9.0) (Graphpad). Significance was determined by unpaired Student's t-test with 95% confidence interval.

Assay for transposase-accessible chromatin sequencing and analysis

Spleens were isolated and pooled from 3–5 mice per group. Then, 2-5×10^4 $CXCR5^{hi}SLAM^{lo}$ T_{FH} and $CXCR5^{lo}SLAM^{hi}$ T_H1 SMARTA cells were sorted using a FACSAria. Cells were pelleted and resuspended in 25 microliter lysis buffer, and pelleted again. The nuclear pellet was resuspended into 25 microliter transposition reaction mixture containing Tn5 transposase from a Nextera DNA Sample Prep Kit (Illumina) and incubated at 37°C for 30 min. Then, the transposase-associated DNA was purified using a MinElute Purification kit (Qiagen). To amplify the library, the DNA was amplified for twelve cycles using a KAPA Real-Time Library amplification kit (KAPA Biosystems) with Nextera indexing primers. The total amplified DNA was purified using AmPureXP beads. The quantity and size of amplified DNA was examined by TapeStation to confirm that independent samples exhibited similar fragment distributions. The libraries were sequenced using a HiSeq 4000 with paired-end sequencing (Illumina). Replicates were generated from two independent experiments. Analysis of ATAC data was done using the ENCODE ATAC-seq pipeline [26].

CRISPR–Cas9-mediated gene deletion of murine $CD4^+$ T cells

High-ranked guide sequences with the highest on-target and off-target scores were selected by CHOP-CHOP or IDT. crRNA and ATTO-550-conjugated trans-activating CRISPR

RNA (tracrRNA) were purchased from Integrated DNA Technologies. Purified Streptococcus pyogenes Cas9-NLS protein was purchased from QB3 Macrolab of University of California, Berkeley. crRNA and tracrRNA were duplexed by heating at 95°C for 5 min. RNP complexes were generated by mixing crRNA–tracrRNA duplexes (240 pmol) and Cas9-NLS protein (80 pmol) for 10 min at 24–26°C. Isolated $CD4^+$ T cells were stimulated in 6-well plates pre-coated with anti-CD3 and anti-CD28 for 2 days. The cells were then transfected with an RNP mixture by electroporation using Invitrogen NEON transfection system (1600 Volts, 10ms Width, 3 Pulses). The transfected cells were cultured in R10 + 50 M 2-ME + 10 ng ml1 human IL-2 without TCR stimulation for 1 d, followed by *in vivo* cell transfer. crRNA sequences used in the study were as follows: crCD90.2 (5-CGTGTGCTCGGGTATCCCAA-3), crSrebf2-AA (5-GCACCTGCGGGGACGT CACG-3), crSrebf2-AC (5-CTTCAGCGTGGTCAACACAA-3), crRorb-AA (5-CAG GTCAATGACGTGCCCGT-3), rRorb-AB (5-TACGGAGTCATCACGTGT GA-3).

Study Approval

All animal studies were approved by the Institutional Animal Care and Use Committees of the University of California, San Diego (UCSD) and performed in accordance with UC guidelines.

Bibliography

[1] R. S. Akondy, N. D. Monson, J. D. Miller, S. Edupuganti, D. Teuwen, H. Wu, F. Quyyumi, S. Garg, J. D. Altman, C. Del Rio, H. L. Keyserling, A. Ploss, C. M. Rice, W. A. Orenstein, M. J. Mulligan, and R. Ahmed. The yellow fever virus vaccine induces a broad and polyfunctional human memory cd8+ t cell response. *J Immunol*, 183(12):7919–30, 2009.

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

[3] F. Annunziato, C. Romagnani, and S. Romagnani. The 3 major types of innate and adaptive cell-mediated effector immunity. *J Allergy Clin Immunol*, 135(3):626–35, 2015.

[4] V. P. Badovinac, K. A. Messingham, A. Jabbari, J. S. Haring, and J. T. Harty. Accelerated cd8+ t-cell memory and prime-boost response after dendritic-cell vaccination. *Nat Med*, 11(7):748–56, 2005.

[5] V. P. Badovinac, B. B. Porter, and J. T. Harty. Cd8+ t cell contraction is controlled by early inflammation. *Nat Immunol*, 5(8):809–17, 2004.

[6] P. Bagri, V. C. Anipindi, P. V. Nguyen, D. Vitali, M. R. Stampfli, and C. Kaushic. Novel role for interleukin-17 in enhancing type 1 helper t cell immunity in the female genital tract following mucosal herpes simplex virus 2 vaccination. *J Virol*, 91(23), 2017.

[7] A. Banerjee, S. M. Gordon, A. M. Intlekofer, M. A. Paley, E. C. Mooney, T. Lindsten, E. J. Wherry, and S. L. Reiner. Cutting edge: The transcription factor eomesodermin enables cd8+ t cells to compete for the memory cell niche. *J Immunol*, 185(9):4988–92, 2010.

[8] C. K. Baumgartner, H. Yagita, and L. P. Malherbe. A tcr affinity threshold regulates memory cd4 t cell differentiation following vaccination. *J Immunol*, 189(5):2309–17, 2012.

[9] R. Benezra, R. L. Davis, A. Lassar, S. Tapscott, M. Thayer, D. Lockshon, and H. Weintraub. Id: a negative regulator of helix-loop-helix dna binding proteins. control of terminal myogenic differentiation. *Ann N Y Acad Sci*, 599:1–11, 1990.

[10] M. Berard and D. F. Tough. Qualitative differences between naive and memory t cells. *Immunology*, 106(2):127–38, 2002.

[11] N. Beyersdorf, X. Ding, J. K. Tietze, and T. Hanke. Characterization of mouse cd4 t cell subsets defined by expression of klrg1. *Eur J Immunol*, 37(12):3445–54, 2007.

[12] J. N. Blattman, R. Antia, D. J. Sourdive, X. Wang, S. M. Kaech, K. Murali-Krishna, J. D. Altman, and R. Ahmed. Estimating the precursor frequency of naive antigen-specific cd8 t cells. *J Exp Med*, 195(5):657–64, 2002.

[13] E. K. Brinkman, T. Chen, M. Amendola, and B. van Steensel. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res*, 42(22):e168, 2014.

[14] J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, and W. J. Greenleaf. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, dna-binding proteins and nucleosome position. *Nat Methods*, 10(12):1213–8, 2013.

[15] M. A. Cannarile, N. A. Lind, R. Rivera, A. D. Sheridan, K. A. Camfield, B. B. Wu, K. P. Cheung, Z. Ding, and A. W. Goldrath. Transcriptional regulator id2 mediates cd8+ t cell immunity. *Nat Immunol*, 7(12):1317–25, 2006.

[16] J. L. Cannons, K. T. Lu, and P. L. Schwartzberg. T follicular helper cell diversity and plasticity. *Trends Immunol*, 34(5):200–7, 2013.

[17] J. T. Chang, E. J. Wherry, and A. W. Goldrath. Molecular regulation of effector and memory t cell differentiation. *Nat Immunol*, 15(12):1104–15, 2014.

[18] T. J. Chapman and D. J. Topham. Identification of a unique population of tissue-memory cd4+ t cells in the airways after influenza infection that is dependent on the integrin vla-1. *J Immunol*, 184(7):3841–9, 2010.

[19] Y. Chen, S. K. Chauhan, X. Tan, and R. Dana. Interleukin-7 and -15 maintain pathogenic memory th17 cells in autoimmunity. *J Autoimmun*, 77:96–103, 2017.

[20] Y. Chen, V. K. Kuchroo, J. Inobe, D. A. Hafler, and H. L. Weiner. Regulatory t cell clones induced by oral tolerance: suppression of autoimmune encephalomyelitis. *Science*, 265(5176):1237–40, 1994.

[21] B. K. Cho, C. Wang, S. Sugawa, H. N. Eisen, and J. Chen. Functional differences between memory and naive cd8 t cells. *Proc Natl Acad Sci U S A*, 96(6):2976–81, 1999.

[22] Y. S. Choi, J. A. Gullicksrud, S. Xing, Z. Zeng, Q. Shan, F. Li, P. E. Love, W. Peng, H. H. Xue, and S. Crotty. Lef-1 and tcf-1 orchestrate t(fh) differentiation by regulating differentiation circuits upstream of the transcriptional repressor bcl6. *Nat Immunol*, 16(9):980–90, 2015.

[23] Y. S. Choi, J. A. Yang, I. Yusuf, R. J. Johnston, J. Greenbaum, B. Peters, and S. Crotty. Bcl6 expressing follicular helper cd4 t cells are fate committed early and have the capacity to form memory. *J Immunol*, 190(8):4014–26, 2013.

[24] T. Ciucci, M. S. Vacchio, Y. Gao, F. Tomassoni Ardori, J. Candia, M. Mehta, Y. Zhao, B. Tran, M. Pepper, L. Tessarollo, D. B. McGavern, and R. Bosselut. The emergence and functional fitness of memory cd4(+) t cells require the transcription factor thpok. *Immunity*, 50(1):91–105 e4, 2019.

[25] N. Collins, X. Jiang, A. Zaid, B. L. Macleod, J. Li, C. O. Park, A. Haque, S. Bedoui, W. R. Heath, S. N. Mueller, T. S. Kupper, T. Gebhardt, and F. R. Carbone. Skin cd4(+) memory t cells exhibit combined cluster-mediated retention and equilibration with the circulation. *Nat Commun*, 7:11514, 2016.

[26] Encode Project Consortium. An integrated encyclopedia of dna elements in the human genome. *Nature*, 489(7414):57–74, 2012.

[27] E. Corse, R. A. Gottschalk, and J. P. Allison. Strength of tcr-peptide/mhc interactions and in vivo t cell responses. *J Immunol*, 186(9):5039–45, 2011.

[28] S. Crotty. T follicular helper cell biology: A decade of discovery and diseases. *Immunity*, 50(5):1132–1148, 2019.

[29] M. A. Daniels and E. Teixeiro. Tcr signaling in t cell memory. *Front Immunol*, 6:617, 2015.

[30] V. Dardalhon, A. Awasthi, H. Kwon, G. Galileos, W. Gao, R. A. Sobel, M. Mitsdoerffer, T. B. Strom, W. Elyaman, I. C. Ho, S. Khoury, M. Oukka, and V. K. Kuchroo. Il-4 inhibits tgf-beta-induced foxp3+ t cells and, together with tgf-beta, generates il-9+ il-10+ foxp3(-) effector t cells. *Nat Immunol*, 9(12):1347–55, 2008.

[31] K. Dhume and K. K. McKinstry. Early programming and late-acting checkpoints governing the development of cd4 t-cell memory. *Immunology*, 155(1):53–62, 2018.

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

[33] M. Dominguez-Villar and D. A. Hafler. Regulatory t cells in autoimmune disease. *Nat Immunol*, 19(7):665–673, 2018.

[34] Y. Endo, K. Hirahara, R. Yagi, D. J. Tumes, and T. Nakayama. Pathogenic memory type th2 cells in allergic inflammation. *Trends Immunol*, 35(2):69–78, 2014.

[35] Y. Endo, C. Iwamura, M. Kuwahara, A. Suzuki, K. Sugaya, D. J. Tumes, K. Tokoyoda, H. Hosokawa, M. Yamashita, and T. Nakayama. Eomesodermin controls interleukin-5 production in memory t helper 2 cells through inhibition of activity of the transcription factor gata3. *Immunity*, 35(5):733–45, 2011.

[36] S. Eyerich, K. Eyerich, D. Pennino, T. Carbone, F. Nasorri, S. Pallotta, F. Cianfarani, T. Odorisio, C. Traidl-Hoffmann, H. Behrendt, S. R. Durham, C. B. Schmidt-Weber, and A. Cavani. Th22 cells represent a distinct human t cell subset involved in epidermal immunity and remodeling. *J Clin Invest*, 119(12):3573–85, 2009.

[37] D. Fang and J. Zhu. Dynamic balance between master transcription factors determines the fates and functions of cd4 t cell and innate lymphoid cell subsets. *J Exp Med*, 214(7):1861–1876, 2017.

[38] D. L. Farber, N. A. Yudanin, and N. P. Restifo. Human memory t cells: generation, compartmentalization and homeostasis. *Nat Rev Immunol*, 14(1):24–35, 2014.

[39] A. M. C. Faria, B. S. Reis, and D. Mucida. Tissue adaptation: Implications for gut immunity and tolerance. *J Exp Med*, 214(5):1211–1226, 2017.

[40] K. E. Foulds, L. A. Zenewicz, D. J. Shedlock, J. Jiang, A. E. Troy, and H. Shen. Cutting edge: Cd4 and cd8 t cells are intrinsically different in their proliferative responses. *J Immunol*, 168(4):1528–32, 2002.

[41] Jr. Frias, A. B., E. J. Hyzny, H. M. Buechel, L. Y. Beppu, B. Xie, M. J. Jurczak, and L. M. D'Cruz. The transcriptional regulator id2 is critical for adipose-resident regulatory t cell differentiation, survival, and function. *J Immunol*, 203(3):658–664, 2019.

[42] P. Gao, X. Han, Q. Zhang, Z. Yang, I. J. Fuss, T. G. Myers, P. J. Gardina, F. Zhang, and W. Strober. Dynamic changes in e-protein activity regulate t reg cell development. *J Exp Med*, 211(13):2651–68, 2014.

[43] D. J. Gasper, M. M. Tejera, and M. Suresh. Cd4 t-cell memory generation and maintenance. *Crit Rev Immunol*, 34(2):121–46, 2014.

[44] T. Gebhardt, P. G. Whitney, A. Zaid, L. K. Mackay, A. G. Brooks, W. R. Heath, F. R. Carbone, and S. N. Mueller. Different patterns of peripheral migration by memory cd4+ and cd8+ t cells. *Nature*, 477(7363):216–9, 2011.

[45] N. D. Glennie, S. W. Volk, and P. Scott. Skin-resident cd4+ t cells protect against leishmania major by recruiting and activating inflammatory monocytes. *PLoS Pathog*, 13(4):e1006349, 2017.

[46] N. D. Glennie, V. A. Yeramilli, D. P. Beiting, S. W. Volk, C. T. Weaver, and P. Scott. Skin-resident memory cd4+ t cells enhance protection against leishmania major infection. *J Exp Med*, 212(9):1405–14, 2015.

[47] J. I. Gray, L. M. Westerhof, and M. K. L. MacLeod. The roles of resident, central and effector memory cd4 t-cells in protective immunity following infection or vaccination. *Immunology*, 2018.

[48] H. Groux, A. O'Garra, M. Bigler, M. Rouleau, S. Antonenko, J. E. de Vries, and M. G. Roncarolo. A cd4+ t-cell subset inhibits antigen-specific t-cell responses and prevents colitis. *Nature*, 389(6652):737–42, 1997.

[49] R. K. Gurram and J. Zhu. Orchestration between ilc2s and th2 cells in shaping type 2 immune responses. *Cell Mol Immunol*, 16(3):225–235, 2019.

[50] S. M. Haeryfar and D. W. Hoskin. Thy-1: more than a mouse pan-t cell marker. *J Immunol*, 173(6):3581–8, 2004.

[51] J. S. Hale, B. Youngblood, D. R. Latner, A. U. Mohammed, L. Ye, R. S. Akondy, T. Wu, S. S. Iyer, and R. Ahmed. Distinct memory cd4+ t cells with commitment to t follicular helper- and t helper 1-cell lineages are generated after acute viral infection. *Immunity*, 38(4):805–17, 2013.

[52] L. E. Harrington, K. M. Janowski, J. R. Oliver, A. J. Zajac, and C. T. Weaver. Memory cd4 t cells emerge from effector t-cell progenitors. *Nature*, 452(7185):356–60, 2008.

[53] A. N. Hegazy, M. Peine, C. Helmstetter, I. Panse, A. Frohlich, A. Bergthaler, L. Flatz, D. D. Pinschewer, A. Radbruch, and M. Lohning. Interferons direct th2 cell reprogramming to generate a stable gata-3(+)t-bet(+) cell subset with combined th2 and th1 cell functions. *Immunity*, 32(1):116–28, 2010.

[54] K. Hirota, J. E. Turner, M. Villa, J. H. Duarte, J. Demengeot, O. M. Steinmetz, and B. Stockinger. Plasticity of th17 cells in peyer's patches is responsible for the induction of t cell-dependent iga responses. *Nat Immunol*, 14(4):372–9, 2013.

[55] M. Hofmann, V. Brinkmann, and H. G. Zerwes. Fty720 preferentially depletes naive t cells from peripheral and lymphoid organs. *Int Immunopharmacol*, 6(13-14):1902–10, 2006.

[56] B. D. Hondowicz, D. An, J. M. Schenkel, K. S. Kim, H. R. Steach, A. T. Krishnamurty, G. J. Keitany, E. N. Garza, K. A. Fraser, J. J. Moon, W. A. Altemeier, D. Masopust, and M. Pepper. Interleukin-2-dependent allergen-specific tissue-resident memory cells drive asthma. *Immunity*, 44(1):155–166, 2016.

[57] B. D. Hondowicz, K. S. Kim, M. J. Ruterbusch, G. J. Keitany, and M. Pepper. Il-2 is required for the generation of viral-specific cd4(+) th1 tissue-resident memory cells and b cells are essential for maintenance in the lung. *Eur J Immunol*, 48(1):80–86, 2018.

[58] H. Ichii, A. Sakamoto, M. Arima, M. Hatano, Y. Kuroda, and T. Tokuhisa. Bcl6 is essential for the generation of long-term memory cd4+ t cells. *Int Immunol*, 19(4):427–33, 2007.

[59] H. Ichii, A. Sakamoto, Y. Kuroda, and T. Tokuhisa. Bcl6 acts as an amplifier for the generation and proliferative capacity of central memory cd8+ t cells. *J Immunol*, 173(2):883–91, 2004.

[60] N. Iijima and A. Iwasaki. T cell memory. a local macrophage chemokine network sustains protective tissue-resident memory cd4 t cells. *Science*, 346(6205):93–8, 2014.

[61] A. M. Intlekofer, N. Takemoto, E. J. Wherry, S. A. Longworth, J. T. Northrup, V. R. Palanivel, A. C. Mullen, C. R. Gasink, S. M. Kaech, J. D. Miller, L. Gapin, K. Ryan, A. P. Russ, T. Lindsten, J. S. Orange, A. W. Goldrath, R. Ahmed, and S. L. Reiner. Effector and memory cd8+ t cell fate coupled by t-bet and eomesodermin. *Nat Immunol*, 6(12):1236–44, 2005.

[62] I. E. Ivanov, A. V. Wright, J. C. Cofsky, K. D. P. Aris, J. A. Doudna, and Z. Bryant. Cas9 interrogates dna in discrete steps modulated by mismatches and supercoiling. *Proc Natl Acad Sci U S A*, 117(11):5853–5860, 2020.

[63] S. C. Jameson and D. Masopust. Diversity in t cell memory: an embarrassment of riches. *Immunity*, 31(6):859–71, 2009.

[64] Y. Ji, Z. Pos, M. Rao, C. A. Klebanoff, Z. Yu, M. Sukumar, R. N. Reger, D. C. Palmer, Z. A. Borman, P. Muranski, E. Wang, D. S. Schrump, F. M. Marincola, N. P. Restifo, and L. Gattinoni. Repression of the dna-binding inhibitor id3 by blimp-1 limits the formation of memory cd8+ t cells. *Nat Immunol*, 12(12):1230–7, 2011.

[65] N. S. Joshi, W. Cui, A. Chandele, H. K. Lee, D. R. Urso, J. Hagman, L. Gapin, and S. M. Kaech. Inflammation directs memory precursor and short-lived effector cd8(+) t cell fates via the graded expression of t-bet transcription factor. *Immunity*, 27(2):281–95, 2007.

[66] S. M. Kaech, J. T. Tan, E. J. Wherry, B. T. Konieczny, C. D. Surh, and R. Ahmed. Selective expression of the interleukin 7 receptor identifies effector cd8 t cells that give rise to long-lived memory cells. *Nat Immunol*, 4(12):1191–8, 2003.

[67] S. M. Kaech and E. J. Wherry. Heterogeneity and cell-fate decisions in effector and memory cd8+ t cell differentiation during viral infection. *Immunity*, 27(3):393–405, 2007.

[68] B. Kakaradov, J. Arsenio, C. E. Widjaja, Z. He, S. Aigner, P. J. Metz, B. Yu, E. J. Wehrens, J. Lopez, S. H. Kim, E. I. Zuniga, A. W. Goldrath, J. T. Chang, and G. W. Yeo. Early transcriptional and epigenetic regulation of cd8(+) t cell differentiation revealed by single-cell rna sequencing. *Nat Immunol*, 18(4):422–432, 2017.

[69] A. Kallies, A. Xin, G. T. Belz, and S. L. Nutt. Blimp-1 transcription factor is required for the differentiation of effector cd8(+) t cells and memory responses. *Immunity*, 31(2):283–95, 2009.

[70] B. L. Kee. E and id proteins branch out. *Nat Rev Immunol*, 9(3):175–84, 2009.

[71] S. A. Kiazyk and K. R. Fowke. Loss of cd127 expression links immune activation and cd4(+) t cell loss in hiv infection. *Trends Microbiol*, 16(12):567–73, 2008.

[72] Y. Kidani, H. Elsaesser, M. B. Hock, L. Vergnes, K. J. Williams, J. P. Argus, B. N. Marbois, E. Komisopoulou, E. B. Wilson, T. F. Osborne, T. G. Graeber, K. Reue, D. G. Brooks, and S. J. Bensinger. Sterol regulatory element-binding proteins are essential for the metabolic programming of effector t cells and adaptive immunity. *Nat Immunol*, 14(5):489–99, 2013.

[73] N. Killeen. T-cell regulation: Thy-1 - hiding in full view. *Curr Biol*, 7(12):R774–7, 1997.

[74] E. Kiner, E. Willie, B. Vijaykumar, K. Chowdhary, H. Schmutz, J. Chandler, A. Schnell, P. I. Thakore, G. LeGros, S. Mostafavi, D. Mathis, C. Benoist, and Consortium Immunological Genome Project. Gut cd4(+) t cell phenotypes are a continuum molded by microbes, not by th archetypes. *Nat Immunol*, 22(2):216–228, 2021.

[75] R. M. Kondrack, J. Harbertson, J. T. Tan, M. E. McBreen, C. D. Surh, and L. M. Bradley. Interleukin 7 regulates the survival and generation of memory cd4 cells. *J Exp Med*, 198(12):1797–806, 2003.

[76] M. S. Kuhns and M. M. Davis. Tcr signaling emerges from the sum of many parts. *Front Immunol*, 3:159, 2012.

[77] B. V. Kumar, T. J. Connors, and D. L. Farber. Human t cell development, localization, and function throughout life. *Immunity*, 48(2):202–213, 2018.

[78] M. Kunzli, D. Schreiner, T. C. Pereboom, N. Swarnalekha, L. C. Litzler, J. Lotscher, Y. I. Ertuna, J. Roux, F. Geier, R. P. Jakob, T. Maier, C. Hess, J. J. Taylor, and C. G. King. Long-lived t follicular helper cells retain plasticity and help sustain humoral immunity. *Sci Immunol*, 5(45), 2020.

[79] C. Li, X. Guan, T. Du, W. Jin, B. Wu, Y. Liu, P. Wang, B. Hu, G. E. Griffin, R. J. Shattock, and Q. Hu. Inhibition of hiv-1 infection of primary cd4+ t-cells by gene editing of ccr5 using adenovirus-delivered crispr/cas9. *J Gen Virol*, 96(8):2381–2393, 2015.

[80] Y. Y. Lin, M. E. Jones-Mason, M. Inoue, A. Lasorella, A. Iavarone, Q. J. Li, M. L. Shinohara, and Y. Zhuang. Transcriptional regulator id2 is required for the cd4 t cell immune response in the development of experimental autoimmune encephalomyelitis. *J Immunol*, 189(3):1400–5, 2012.

[81] X. Liu, X. Chen, B. Zhong, A. Wang, X. Wang, F. Chu, R. I. Nurieva, X. Yan, P. Chen, L. G. van der Flier, H. Nakatsukasa, S. S. Neelapu, W. Chen, H. Clevers, Q. Tian, H. Qi, L. Wei, and C. Dong. Transcription factor achaete-scute homologue 2 initiates follicular t-helper-cell development. *Nature*, 507(7493):513–8, 2014.

[82] T. Lonnberg, V. Svensson, K. R. James, D. Fernandez-Ruiz, I. Sebina, R. Montandon, M. S. Soon, L. G. Fogg, A. S. Nair, U. Liligeto, M. J. Stubbington, L. H. Ly, F. O. Bagger, M. Zwiessele, N. D. Lawrence, F. Souza-Fonseca-Guimaraes, P. T. Bunn, C. R. Engwerda, W. R. Heath, O. Billker, O. Stegle, A. Haque, and S. A. Teichmann. Single-cell rna-seq and computational analysis using temporal mixture modelling resolves th1/tfh fate bifurcation in malaria. *Sci Immunol*, 2(9), 2017.

[83] K. Luthje, A. Kallies, Y. Shimohakamada, G. T. Belz, A. Light, D. M. Tarlinton, and S. L. Nutt. The development and fate of follicular helper t cells defined by an il-21 reporter mouse. *Nat Immunol*, 13(5):491–8, 2012.

[84] C. R. Mackay. Dual personality of memory t cells. *Nature*, 401(6754):659–60, 1999.

[85] M. K. MacLeod, E. T. Clambey, J. W. Kappler, and P. Marrack. Cd4 memory t cells: what are they and what can they do? *Semin Immunol*, 21(2):53–61, 2009.

[86] M. K. MacLeod, J. W. Kappler, and P. Marrack. Memory cd4 t cells: generation, reactivation and re-assignment. *Immunology*, 130(1):10–5, 2010.

[87] P. K. Mandal, L. M. Ferreira, R. Collins, T. B. Meissner, C. L. Boutwell, M. Friesen, V. Vrbanac, B. S. Garrison, A. Stortchevoi, D. Bryder, K. Musunuru, H. Brand, A. M. Tager, T. M. Allen, M. E. Talkowski, D. J. Rossi, and C. A. Cowan. Efficient ablation of genes in human hematopoietic stem and effector cells using crispr/cas9. *Cell Stem Cell*, 15(5):643–52, 2014.

[88] H. D. Marshall, A. Chandele, Y. W. Jung, H. Meng, A. C. Poholek, I. A. Parish, R. Rutishauser, W. Cui, S. H. Kleinstein, J. Craft, and S. M. Kaech. Differential expression of ly6c and t-bet distinguish effector and memory th1 cd4(+) cell properties during viral infection. *Immunity*, 35(4):633–46, 2011.

[89] M. D. Martin and V. P. Badovinac. Defining memory cd8 t cell. *Front Immunol*, 9:2692, 2018.

[90] T. Maruyama, J. Li, J. P. Vaque, J. E. Konkel, W. Wang, B. Zhang, P. Zhang, B. F. Zamarron, D. Yu, Y. Wu, Y. Zhuang, J. S. Gutkind, and W. Chen. Control of the differentiation of regulatory t cells and t(h)17 cells by the dna-binding inhibitor id3. *Nat Immunol*, 12(1):86–95, 2011.

[91] M. E. Massari and C. Murre. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol*, 20(2):429–40, 2000.

[92] F. Masson, M. Minnich, M. Olshansky, I. Bilic, A. M. Mount, A. Kallies, T. P. Speed, M. Busslinger, S. L. Nutt, and G. T. Belz. Id2-mediated inhibition of e2a represses memory cd8+ t cell differentiation. *J Immunol*, 190(9):4585–94, 2013.

[93] M. J. McGeachy. Th17 memory cells: live long and proliferate. *J Leukoc Biol*, 94(5):921–6, 2013.

[94] K. K. McKinstry, T. M. Strutt, B. Bautista, W. Zhang, Y. Kuang, A. M. Cooper, and S. L. Swain. Effector cd4 t-cell transition to memory requires late cognate interactions that induce autocrine il-2. *Nat Commun*, 5:5377, 2014.

[95] K. K. McKinstry, T. M. Strutt, and S. L. Swain. The potential of cd4 t-cell memory. *Immunology*, 130(1):1–9, 2010.

[96] J. J. Milner and A. W. Goldrath. Transcriptional programming of tissue-resident memory cd8(+) t cells. *Curr Opin Immunol*, 51:162–169, 2018.

[97] J. J. Milner, H. Nguyen, K. Omilusik, M. Reina-Campos, M. Tsai, C. Toma, A. Delpoux, B. S. Boland, S. M. Hedrick, J. T. Chang, and A. W. Goldrath. Delineation of a molecularly distinct terminally differentiated memory cd8 t cell population. *Proc Natl Acad Sci U S A*, 117(41):25667–25678, 2020.

[98] T. Mishima, S. Toda, Y. Ando, T. Matsunaga, and M. Inobe. Rapid proliferation of activated lymph node cd4(+) t cells is achieved by greatly curtailing the duration of gap phases in cell cycle progression. *Cell Mol Biol Lett*, 19(4):638–48, 2014.

[99] M. Miyazaki, K. Miyazaki, S. Chen, M. Itoi, M. Miller, L. F. Lu, N. Varki, A. N. Chang, D. H. Broide, and C. Murre. Id2 and id3 maintain the regulatory t cell pool to suppress inflammatory disease. *Nat Immunol*, 15(8):767–76, 2014.

[100] J. R. Mora and U. H. von Andrian. T-cell homing specificity and plasticity: new concepts and future challenges. *Trends Immunol*, 27(5):235–43, 2006.

[101] S. N. Mueller, T. Gebhardt, F. R. Carbone, and W. R. Heath. Memory t cell subsets, migration patterns, and tissue residence. *Annu Rev Immunol*, 31:137–61, 2013.

[102] P. Muranski, Z. A. Borman, S. P. Kerkar, C. A. Klebanoff, Y. Ji, L. Sanchez-Perez, M. Sukumar, R. N. Reger, Z. Yu, S. J. Kern, R. Roychoudhuri, G. A. Ferreyra, W. Shen, S. K. Durum, L. Feigenbaum, D. C. Palmer, P. A. Antony, C. C. Chan, A. Laurence, R. L. Danner, L. Gattinoni, and N. P. Restifo. Th17 cells are long lived and retain a stem cell-like molecular signature. *Immunity*, 35(6):972–85, 2011.

[103] C. Murre. Helix-loop-helix proteins and lymphocyte development. *Nat Immunol*, 6(11):1079–86, 2005.

[104] T. Nakayama and M. Yamashita. Critical role of the polycomb and trithorax complexes in the maintenance of cd4 t cell memory. *Semin Immunol*, 21(2):78–83, 2009.

[105] S. Nakayamada, H. Takahashi, Y. Kanno, and J. J. O'Shea. Helper t cell diversity and plasticity. *Curr Opin Immunol*, 24(3):297–302, 2012.

[106] S. Neph, A. B. Stergachis, A. Reynolds, R. Sandstrom, E. Borenstein, and J. A. Stamatoyannopoulos. Circuitry and dynamics of human transcription factor regulatory networks. *Cell*, 150(6):1274–86, 2012.

[107] Q. P. Nguyen, T. Z. Deng, D. A. Witherden, and A. W. Goldrath. Origins of cd4(+) circulating and tissue-resident memory t-cells. *Immunology*, 157(1):3–12, 2019.

[108] S. A. Oh, A. Seki, and S. Rutz. Ribonucleoprotein transfection for crispr/cas9-mediated gene knockout in primary t cells. *Curr Protoc Immunol*, 124(1):e69, 2019.

[109] L. C. Osborne, S. Dhanji, J. W. Snow, J. J. Priatel, M. C. Ma, M. J. Miners, H. S. Teh, M. A. Goldsmith, and N. Abraham. Impaired cd8 t cell memory and cd4 t cell primary responses in il-7r alpha mutant mice. *J Exp Med*, 204(3):619–31, 2007.

[110] M. A. Paley, S. M. Gordon, E. K. Bikoff, E. J. Robertson, E. J. Wherry, and S. L. Reiner. Technical advance: Fluorescent reporter reveals insights into eomesodermin biology in cytotoxic lymphocytes. *J Leukoc Biol*, 93(2):307–15, 2013.

[111] M. Pepper and M. K. Jenkins. Origins of cd4(+) effector and central memory t cells. *Nat Immunol*, 12(6):467–71, 2011.

[112] M. Pepper, J. L. Linehan, A. J. Pagan, T. Zell, T. Dileepan, P. P. Cleary, and M. K. Jenkins. Different routes of bacterial infection induce long-lived th1 memory cells and short-lived th17 cells. *Nat Immunol*, 11(1):83–9, 2010.

[113] M. Pepper, A. J. Pagan, B. Z. Igyarto, J. J. Taylor, and M. K. Jenkins. Opposing signals from the bcl6 transcription factor and the interleukin-2 receptor generate t helper 1 central and effector memory cells. *Immunity*, 35(4):583–95, 2011.

[114] M. J. Richer, J. C. Nolz, and J. T. Harty. Pathogen-specific inflammatory milieux tune the antigen sensitivity of cd8(+) t cells by enhancing t cell receptor signaling. *Immunity*, 38(1):140–52, 2013.

[115] P. A. Romagnoli, H. H. Fu, Z. Qiu, C. Khairallah, Q. M. Pham, L. Puddington, K. M. Khanna, L. Lefrancois, and B. S. Sheridan. Differentiation of distinct long-lived memory cd4 t cells in intestinal tissues after oral listeria monocytogenes infection. *Mucosal Immunol*, 10(2):520–530, 2017.

[116] M. Ruterbusch, K. B. Pruner, L. Shehata, and M. Pepper. In vivo cd4(+) t cell differentiation and function: Revisiting the th1/th2 paradigm. *Annu Rev Immunol*, 38:705–725, 2020.

[117] R. L. Rutishauser, G. A. Martins, S. Kalachikov, A. Chandele, I. A. Parish, E. Meffre, J. Jacob, K. Calame, and S. M. Kaech. Transcriptional repressor blimp-1 promotes cd8(+) t cell terminal differentiation and represses the acquisition of central memory t cell properties. *Immunity*, 31(2):296–308, 2009.

[118] S. Sakai, K. D. Kauffman, J. M. Schenkel, C. C. McBerry, K. D. Mayer-Barber, D. Masopust, and D. L. Barber. Cutting edge: control of mycobacterium tuberculosis infection by a subset of lung parenchyma-homing cd4 t cells. *J Immunol*, 192(7):2965–9, 2014.

[119] F. Sallusto, D. Lenig, R. Forster, M. Lipp, and A. Lanzavecchia. Two subsets of memory t lymphocytes with distinct homing potentials and effector functions. *Nature*, 401(6754):708–12, 1999.

[120] J. M. Schenkel, K. A. Fraser, V. Vezys, and D. Masopust. Sensing and alarm function of resident memory cd8(+) t cells. *Nat Immunol*, 14(5):509–13, 2013.

[121] J. M. Schenkel and D. Masopust. Tissue-resident memory t cells. *Immunity*, 41(6):886–97, 2014.

[122] A. Shakya, A. Goren, A. Shalek, C. N. German, J. Snook, V. K. Kuchroo, N. Yosef, R. C. Chan, A. Regev, M. A. Williams, and D. Tantin. Oct1 and oca-b are selectively required for cd4 memory t cell function. *J Exp Med*, 212(12):2115–31, 2015.

[123] L. A. Shaw, S. Belanger, K. D. Omilusik, S. Cho, J. P. Scott-Browne, J. P. Nance, J. Goulding, A. Lasorella, L. F. Lu, S. Crotty, and A. W. Goldrath. Id2 reinforces th1 differentiation and inhibits e2a to repress tfh differentiation. *Nat Immunol*, 17(7):834–43, 2016.

[124] H. Shin, S. D. Blackburn, A. M. Intlekofer, C. Kao, J. M. Angelosanto, S. L. Reiner, and E. J. Wherry. A role for the transcriptional repressor blimp-1 in cd8(+) t cell exhaustion during chronic viral infection. *Immunity*, 31(2):309–20, 2009.

[125] H. Shin and A. Iwasaki. Skin t(rm) mediates distributed border patrol. *Cell Res*, 22(9):1325–7, 2012.

[126] N. M. Smith, G. A. Wasserman, F. T. Coleman, K. L. Hilliard, K. Yamamoto, E. Lipsitz, R. Malley, H. Dooms, M. R. Jones, L. J. Quinton, and J. P. Mizgerd. Regionally compartmentalized resident memory t cells mediate naturally acquired protection against pneumococcal pneumonia. *Mucosal Immunol*, 11(1):220–235, 2018.

[127] J. P. Snook, C. Kim, and M. A. Williams. Tcr signal strength controls the differentiation of cd4(+) effector and memory t cells. *Sci Immunol*, 3(25), 2018.

[128] S. Steinfelder, S. Rausch, D. Michael, A. A. Kuhl, and S. Hartmann. Intestinal helminth infection induces highly functional resident memory cd4(+) t cells in mice. *Eur J Immunol*, 47(2):353–363, 2017.

[129] B. Stockinger and S. Omenetti. The dichotomous nature of t helper 17 cells. *Nat Rev Immunol*, 17(9):535–544, 2017.

[130] T. M. Strutt, K. Dhume, C. M. Finn, J. H. Hwang, C. Castonguay, S. L. Swain, and K. K. McKinstry. Il-15 supports the generation of protective lung-resident memory cd4 t cells. *Mucosal Immunol*, 11(3):668–680, 2018.

[131] S. Su, B. Hu, J. Shao, B. Shen, J. Du, Y. Du, J. Zhou, L. Yu, L. Zhang, F. Chen, H. Sha, L. Cheng, F. Meng, Z. Zou, X. Huang, and B. Liu. Crispr-cas9 mediated efficient pd-1 disruption on human primary t cells from cancer patients. *Sci Rep*, 6:20070, 2016.

[132] S. L. Swain, K. K. McKinstry, and T. M. Strutt. Expanding roles for cd4(+) t cells in immunity to viruses. *Nat Rev Immunol*, 12(2):136–48, 2012.

[133] J. R. Teijaro, D. Turner, Q. Pham, E. J. Wherry, L. Lefrancois, and D. L. Farber. Cutting edge: Tissue-retentive lung memory cd4 t cells mediate optimal protection to respiratory virus infection. *J Immunol*, 187(11):5510–4, 2011.

[134] N. J. Tubo, B. T. Fife, A. J. Pagan, D. I. Kotov, M. F. Goldberg, and M. K. Jenkins. Most microbe-specific naive cd4(+) t cells produce memory cells during infection. *Science*, 351(6272):511–4, 2016.

[135] D. L. Turner and D. L. Farber. Mucosal resident memory cd4 t cells in protection and immunopathology. *Front Immunol*, 5:331, 2014.

[136] D. L. Turner, M. Goldklang, F. Cvetkovski, D. Paik, J. Trischler, J. Barahona, M. Cao, R. Dave, N. Tanna, J. M. D'Armiento, and D. L. Farber. Biased generation and in situ activation of lung tissue-resident memory cd4 t cells in the pathogenesis of allergic asthma. *J Immunol*, 200(5):1561–1569, 2018.

[137] G. Vahedi, H. Takahashi, S. Nakayamada, H. W. Sun, V. Sartorelli, Y. Kanno, and J. J. O'Shea. Stats shape the active enhancer landscape of t cell populations. *Cell*, 151(5):981–93, 2012.

[138] E. M. van Leeuwen, J. Sprent, and C. D. Surh. Generation and maintenance of memory cd4(+) t cells. *Curr Opin Immunol*, 21(2):167–72, 2009.

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

[140] W. Wang, C. Ye, J. Liu, D. Zhang, J. T. Kimata, and P. Zhou. Ccr5 gene disruption via lentiviral vectors expressing cas9 and single guided rna renders cells resistant to hiv-1 infection. *PLoS One*, 9(12):e115987, 2014.

[141] G. Wei, L. Wei, J. Zhu, C. Zang, J. Hu-Li, Z. Yao, K. Cui, Y. Kanno, T. Y. Roh, W. T. Watford, D. E. Schones, W. Peng, H. W. Sun, W. E. Paul, J. J. O'Shea, and K. Zhao. Global mapping of h3k4me3 and h3k27me3 reveals specificity and plasticity in lineage fate determination of differentiating cd4+ t cells. *Immunity*, 30(1):155–67, 2009.

[142] M. M. Wilk, A. Misiak, R. M. McManus, A. C. Allen, M. A. Lynch, and K. H. G. Mills. Lung cd4 tissue-resident memory t cells mediate adaptive immunity induced by previous infection of mice with bordetella pertussis. *J Immunol*, 199(1):233–243, 2017.

[143] M. A. Williams, E. V. Ravkov, and M. J. Bevan. Rapid culling of the cd4+ t cell repertoire in the transition from effector to memory. *Immunity*, 28(4):533–45, 2008.

[144] D. R. Winter, S. Jung, and I. Amit. Making the case for chromatin profiling: a new tool to investigate the immune-regulatory landscape. *Nat Rev Immunol*, 15(9):585–94, 2015.

[145] D. L. Woodland and J. E. Kohlmeier. Migration, maintenance and recall of memory t cells in peripheral tissues. *Nat Rev Immunol*, 9(3):153–61, 2009.

[146] L. Xu, Y. Cao, Z. Xie, Q. Huang, Q. Bai, X. Yang, R. He, Y. Hao, H. Wang, T. Zhao, Z. Fan, A. Qin, J. Ye, X. Zhou, L. Ye, and Y. Wu. The transcription factor tcf-1 initiates the differentiation of t(fh) cells during acute viral infection. *Nat Immunol*, 16(9):991–9, 2015.

[147] H. Yamane and W. E. Paul. Memory cd4+ t cells: fate determination, positive feedback and plasticity. *Cell Mol Life Sci*, 69(10):1577–83, 2012.

[148] C. Y. Yang, J. A. Best, J. Knell, E. Yang, A. D. Sheridan, A. K. Jesionek, H. S. Li, R. R. Rivera, K. C. Lind, L. M. D'Cruz, S. S. Watowich, C. Murre, and A. W. Goldrath. The transcriptional regulators id2 and id3 control the formation of distinct memory cd8+ t cell subsets. *Nat Immunol*, 12(12):1221–9, 2011.

[149] S. M. Yeon, L. Halim, A. Chandele, C. J. Perry, S. H. Kim, S. U. Kim, Y. Byun, S. H. Yuk, S. M. Kaech, and Y. W. Jung. Il-7 plays a critical role for the homeostasis of allergen-specific memory cd4 t cells in the lung and airways. *Sci Rep*, 7(1):11155, 2017.

[150] B. Yu, K. Zhang, J. J. Milner, C. Toma, R. Chen, J. P. Scott-Browne, R. M. Pereira, S. Crotty, J. T. Chang, M. E. Pipkin, W. Wang, and A. W. Goldrath. Epigenetic landscapes reveal transcription factors that regulate cd8(+) t cell differentiation. *Nat Immunol*, 18(5):573–582, 2017.

[151] V. P. Zediak, J. B. Johnnidis, E. J. Wherry, and S. L. Berger. Cutting edge: persistently open chromatin at effector gene loci in resting memory cd8+ t cells independent of transcriptional status. *J Immunol*, 186(5):2705–9, 2011.

[152] L. Zhou, M. M. Chong, and D. R. Littman. Plasticity of cd4+ t cell lineage differentiation. *Immunity*, 30(5):646–55, 2009.

[153] X. Zhou 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*, 189(6):2722–6, 2012.

[154] X. Zhou, S. Yu, D. M. Zhao, J. T. Harty, V. P. Badovinac, and H. H. Xue. Differentiation and persistence of memory cd8(+) t cells depend on t cell factor 1. *Immunity*, 33(2):229–40, 2010.

[155] J. Zhu, H. Yamane, and W. E. Paul. Differentiation of effector cd4 t cell populations (*). *Annu Rev Immunol*, 28:445–89, 2010.

[156] L. A. Zuniga, R. Jain, C. Haines, and D. J. Cua. Th17 cell development: from the cradle to the grave. *Immunol Rev*, 252(1):78–88, 2013.