UNIVERSITY OF CALIFORNIA SAN DIEGO

Elucidating the Developmental Origins and Transcriptional Programming of $CD4^+$ Tissue-Resident Memory T cells in Anti-viral Immunity

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

in

Biology

by

Quynh Phuong Nguyen

Committee in charge:

Professor Ananda Goldrath, Chair
Professor John Chang
Professor Stephen M Hedrick
Professor Wendy Huang
Professor Susan Kaech

2021

The dissertation of Quynh Phuong Nguyen is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

"Công cha như núi Thái Sơn,
Nghĩa mẹ như nước trong nguồn chảy ra.
Một lòng thờ mẹ kính cha,
Cho tròn chữ hiếu mới là đạo con."

To my parents, Chinh M. Nguyen and Hue T. Do, who taught me the importance of hard work, humility, and integrity.

EPIGRAPH

"Like musicians, like mathematicians—like elite athletes—scientists peak early and dwindle fast. It isn't creativity that fades, but stamina: science is an endurance sport. To produce that single illuminating experiment, a thousand non-illuminating experiments have to be sent into the trash; it is battle between nature and nerve."

— Siddhartha Mukherjee, *The Gene: An Intimate History*

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

T_{RM}	Tissue-resident memory T cell
T_H1	Type 1 helper T cell
T_{FH}	Follicular helper T cell
LCMV-Arm	Lymphochoriomeningitis virus-Armstrong
SPL	Spleen
mLN	Mesenteric lymph node
IEL	Intraepithelial lymphocyte
LPL	Lamina propria lymphocyte
SLO	Secondary lymphoid organ
TF	Transcription factor

LIST OF FIGURES

ACKNOWLEDGEMENTS

Good science is never done in a bubble (despite what the pandemic may have forced us to think), and as with raising a child, nurturing a hypothesis into a project with publishable data takes a village. I owe the following people my endless gratitude:

To my PI, Ananda Goldrath: For giving me the chance to train as a scientist under your mentorship; for trusting me take the reins, try new things, and make mistakes; for your grace and patience while I tried to balance a PhD with all my other interests; for your honesty and tough love when you thought I was overtaxing myself; and for teaching me the value of collaboration in doing great science. It has been a privilege to be a member of the Goldrath family.

To members of the Goldrath lab, past and present: For teaching me, engaging with my random thoughts, feeding me, encouraging me, and pushing me to do my best science.

To my committee members – John Chang, Wendy Huang, Sue Kaech, and Steve Hedrick: For your advice and encouragement, even when I proposed things you knew wouldn't actually get done.

To all our collaborators, especially Shane Crotty, Matthew Pipkin, Jinyong Choi, and your lab members: For showing me that science is much more fun when you have others to commiserate your failures and celebrate your successes.

To Anthony Phan: For being a mentor and friend from the start, and for believing that I could in fact drive the bus.

To my closest friends and support system – Cong, Michelle, Evelyn, Andy, Ruby, Connie: For your friendship and for being my ride or dies; for listening to every rant, and afterwards, telling me to get back to work. You made sure I never lost sight of the bigger picture and my goals, even when I wanted to give up.

To my parents: Thank you for your sacrifices and hard work which has allowed me to pursue my dreams. I am here today because of you.

To my little brother, Kendrick: You are a much kinder and thoughtful person than I ever was at your age. Hold on tight to this no matter what life and the world throws at you.

To all animal care technicians, core facilities staff, administrative assistants, finance/grants administrators, custodial staff, and other essential workers at UCSD: You made sure I had all the necessary tools to do my job, and I may not know you all by name, but I will always be grateful for your support.

In loving memory of Arnaud Delpoux: When I think of flow cytometry, I think of you and your advice to me during late nights, "Just work hard and then have a beer. Life's good."

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

Chapters 2 and 3, in part, are adapted from a manuscript which has been submitted for publication. Nguyen QP, Deng TZ, O'Shea SM, Pipkin ME, Choi J, Crotty S, Goldrath AW. (2021). Blimp1, Id2, and Bcl6 balance effector and memory-associated programs to promote $CD4^+$ T_{RM} differentiation following viral infection. The dissertation author was the primary author of this paper.

VITA

2015 Bachelor of Science in Biology, Massachusetts Institute of Technology

2021 Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Nguyen QP, Deng TZ, O'Shea SM, Pipkin ME, Choi J, Crotty S, Goldrath AW. (2021). Blimp1, Id2, and Bcl6 balance effector and memory-associated programs to promote $CD4^+$ T_{RM} differentiation following viral infection. Manuscript submitted for review.

Shaw LA, Deng TZ, Omilusik KD, **Nguyen QP**, Goldrath AW. (2021). Id3 expression identifies mouse $CD4^+$ memory T_H1 cells. Manuscript submitted for review.

Milner JJ, Toma C, He Z, Kurd NS, **Nguyen QP**, McDonald B, Quezada L, Widjaja CE, Witherden DA, Crowl JT, Shaw LA, Yeo GW, Chang JT, Omilusik KD, Goldrath AW. (2020). Heterogenous Populations of Tissue-Resident CD8+ T Cells Are Generated in Response to Infection and Malignancy. *Immunity*, *52*(5), 808-824.e7. https://doi.org/10.1016/j.immuni.2020.04.007

Nguyen QP, Deng TZ, Witherden DA, Goldrath AW. (2019). Origins of $CD4^+$ circulating and tissue-resident memory T-cells. *Immunology*, *157*(1), 3–12. https://doi.org/10.1111/imm.13059

Rosowski EE, **Nguyen QP**, Camejo A, Spooner E, Saeij JPJ. (2014). Toxoplasma gondii inhibits gamma interferon (IFN-γ)-and IFN-β-induced host cell STAT1 transcriptional activity by increasing the association of STAT1 with DNA. *Infection and Immunity*, 8*2*(2), 706–719. https://doi.org/10.1128/IAI.01291-13

Melo MB, **Nguyen QP**, Cordeiro C, Hassan MA, Yang N, Mckell R, Rosowski EE, Julien L, Butty V, Dardé ML, Ajzenberg D, Fitzgerald K, Young LH, Saeij JPJ. (2013). Transcriptional Analysis of Murine Macrophages Infected with Different Toxoplasma Strains Identifies Novel Regulation of Host Signaling Pathways. *PLOS Pathogens*. https://doi.org/10.1371/journal.ppat.1003779

ABSTRACT OF THE DISSERTATION

Elucidating the Developmental Origins and Transcriptional Programming of CD4$^+$ Tissue-Resident Memory T cells in Anti-viral Immunity

by

Quynh Phuong Nguyen

Doctor of Philosophy in Biology

University of California San Diego, 2021

Professor Ananda Goldrath, Chair

CD4$^+$ T lymphocytes are a key element of adaptive immunity, acting to coordinate and enhance functions of innate cells, B cells, and CD8$^+$ T cells in response to diverse pathogens. Following clearance of the pathogen, a small proportion of effector CD4$^+$ T cells persists and

differentiates into long-lived memory cells, which enable a robust secondary response against reinfection and are pivotal in conferring lasting cellular immunity. While the majority of memory cells circulate between tissues and the secondary lymphoid organs (SLOs), tissue-resident memory T cells (T_{RM}) remain lodged in non-lymphoid barrier tissues, particularly at mucosal surfaces like the intestine and serve as sentinels at sites of potential re-exposure to pathogens.

In this dissertation, I aimed to address two overarching questions regarding the biology of virus-specific $CD4^+$ T_{RM} in the small intestine (SI) following acute lymphocytic choriomeningitis virus (LCMV) infection. First, I addressed the developmental origins of $CD4^+$ T_{RM} cells by examining how the resident population is related to circulating $CD4^+$ T helper subsets in SLOs. Second, I investigated potential transcriptional regulators in $CD4^+$ T_{RM} cells, specifically factors with known roles in driving effector versus memory T cell diffentiation. My work revealed that LCMV-specific $CD4^+$ T_{RM} at day 7 of infection shared a gene-expression program and chromatin profile with T_H1 cells and progressively acquired a mature T_{RM} program by a memory time point, supporting a developmental relationship between T_{RM} and T_H1 subsets. Furthermore, I demonstrated that T_{RM} cells expressed genes associated with both effector and memory T cell fates, including the transcriptional regulators Blimp1, Id2, and Bcl6 which were necessary for $CD4^+$ T_{RM} differentiation. T_H1-associated Blimp1 and Id2 were both required for early T_{RM} formation, while T_{FH}-associated Bcl6 initially inhibited T_{RM} differentiation but was critical for development of long-lived T_{RM} cells. These results identified new significance for transcription factors previously associated with circulating $CD4^+$ T cell populations and their roles in driving SI $CD4^+$ T_{RM} differentiation. This work may provide the basis to exploit the protective capacity of this essential memory T cell population and modulate their activity in the immune response.

Chapter 1 Origins of $CD4^+$ circulating and tissue-resident memory T-cells

1.1 Effective vaccines rely on development of memory cell populations

In late 2019, the first cases of COVID-19 caused by the novel respiratory coronavirus SARS-CoV-2 was reported in Wuhan, China. By early 2020, COVID-19 cases were on the rise, spreading across the globe and crippling the world's greatest healthcare systems; countries resorted to national lockdowns, mandated masking and social distancing, and restricted non-essential activities. Meanwhile, scientists raced to develop a vaccine for this unknown foe. In December 2020, the first vaccine was administered in the United States, and since then, millions across the globe have been vaccinated. While the vaccine is not a cure, as evidenced by a resurgence of COVID-19 cases due to new variants of the virus, vaccination has been the most effective method in reducing disease severity and mortality.

Vaccines have long been one of the most effective public health strategies for combating infectious diseases, from the first polio vaccine by Jonas Salk in 1955 to the annual flu vaccine for the common cold. They are essential training for our body's immune system, presenting innocuous or weakened pieces of the infectious pathogen to trigger a response and build up our defenses for future infections. This strategy relies on the biological process of "immunological memory" in which a population of pathogen-specific adaptive B and T cells remembers the offending agent and persists long-term in the body, waiting to re-encounter the infection and respond more rapidly and with a greater magnitude. Effective vaccines must ensure that the resulting immune response provides sufficient and appropriate signals for memory cell differentiation and maintenance.

1.2 $CD4^+$ T cells differentiate into distinct effector cell subsets following antigen encounter and develop into memory cells at resolution of infection.

When the body encounters a foreign pathogen such as a virus, the initial response is dominated by rapid innate cells and effector molecules which act on the order of hours to recognize canonical molecular patterns common among subsets of pathogens and neutralize the pathogen. Within the first few days, these innate cells pass along the message to the adaptive immune arm of the immune system, consisting of B and T cells which are specific for viral antigen and can launch a targeted second attack to eliminate the virus. B cells produce antibodies which bind the surface of the infectious agent to mark it for killing or to neutralize its interactions with other cells. T cells classically have been divided into two lineages, cytotoxic $CD8^+$ T cells, which kill infected cells, and helper $CD4^+$ T cells, which direct and enhance the functions of other immune cells. Effector $CD4^+$ T lymphocytes can differentiate into at least seven known functionally distinct T helper (T_H) subsets including T_H1, T_H2, T follicular helper (T_{FH}), and regulatory T cells (T_{REG}), each with unique effector functions within the circulation, SLOs, and infected tissues. Depending on the type of immunological threat, early host-pathogen interactions result in an infection milieu that directs naive $CD4^+$ T cells to acquire the specific helper functions for the appropriate immune response. Once infection has cleared, the majority of the adaptive cells die via apoptosis during a contraction phase, while a small proportion persists and differentiates into long-lived memory cells (Figure 1.1). This memory population enables a rapid and robust secondary response against recurring pathogens, and is pivotal in conferring lasting cellular immunity, particularly against pathogens where neutralizing antibodies alone are insufficient at providing long-term protection.

While significant advances have been made in understanding the generation and maintenance of memory $CD8^+$ T cells and B 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 continues to decline even after the antigen is fully cleared, resulting in fewer cells available for study, while the $CD8^+$ memory T cell population, if established, is typically stable[1–3]. Second, the existence of functionally distinct effector T_H cell subsets hinders our 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[4–9].

Memory T cells have been conventionally divided into central memory (T_{CM}) cells, which circulate between the blood and SLOs, and effector memory (T_{EM}) cells, which can migrate from the blood into non-lymphoid tissues[10,11]. Over the past decade, evidence of a novel subset of memory cells called tissue-resident memory T cells (T_{RM}) has emerged. T_{RM} are seeded in non-lymphoid tissues, particularly at barrier sites like the skin and intestinal mucosa[12–15]. T_{RM} have limited recirculation out of tissues and serve as sentinels at sites of potential reinfection, where they coordinate the initial response to pathogens and provide a substantial boost to tissue immunity via direct antigen recognition and recruitment of circulating immune cells[13,16]. The majority of studies on T_{RM} have focused on $CD8^+$ tissue-resident lymphocyte differentiation, survival, and function, while less is known about their $CD4^+$ counterparts, which also contribute to antiviral responses[17,18]. A better understanding of the precursors of $CD4^+$ T_{RM} and the molecular mechanisms mediating their differentiation will allow us to harness the protective capacity of this

memory population and modulate their activity in the context of infection or inflammatory diseases. In this chapter, I summarize recent studies addressing the identity of memory $CD4^+$ T cell populations and their precursors in both the periphery and non-lymphoid tissues.

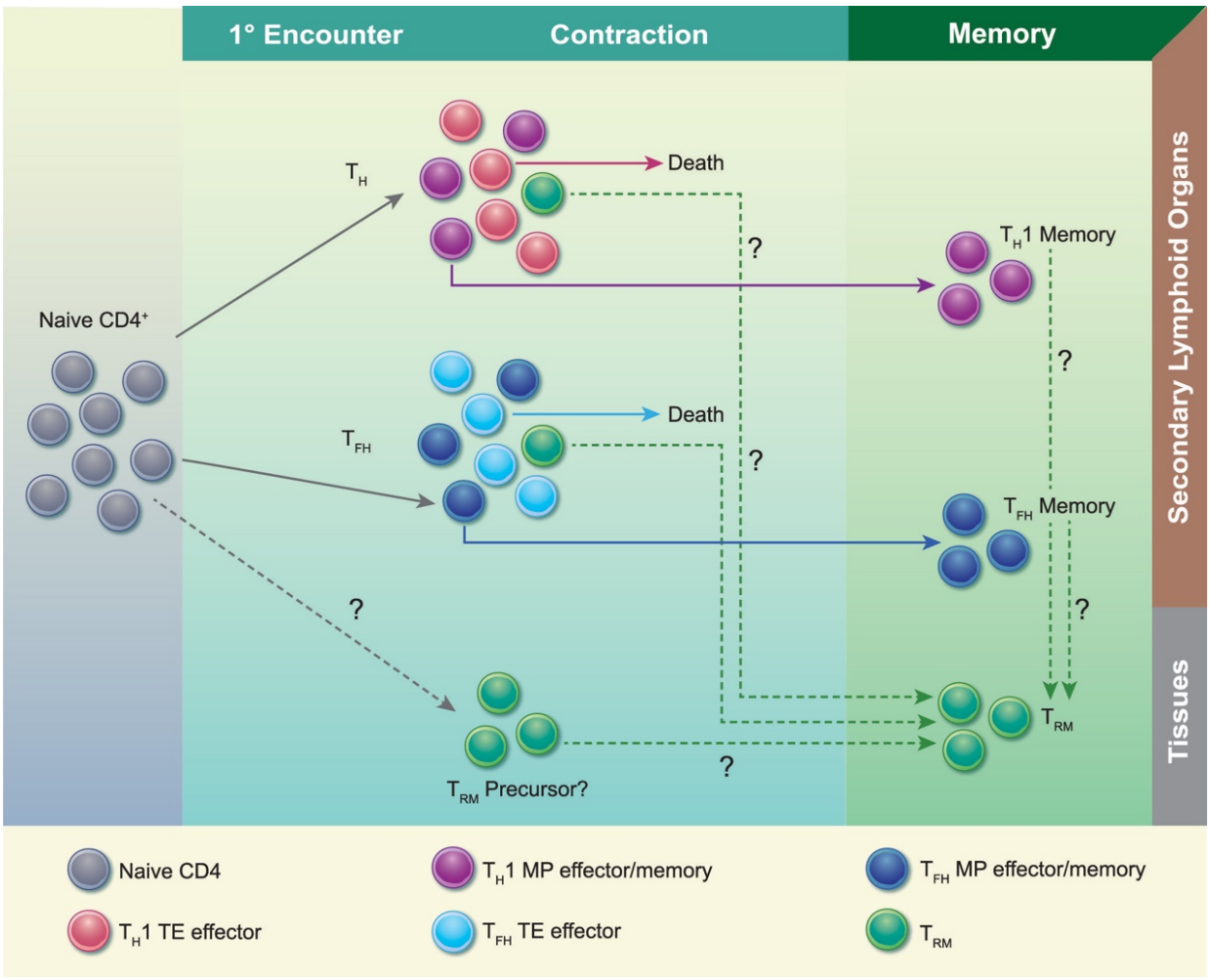


Figure 1.1 Effector and memory CD4$^+$ T cell differentiation. Upon antigen encounter, naive CD4$^+$ T-cells differentiate into effector subsets based on the type of infection. Within each effector CD4$^+$ subset, there potentially exist terminal effectors (TE) and memory-precursor (MP) effectors. The 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 $CD4^+$ 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[1], the models for memory $CD8^+$ T cell formation have served as a useful framework for investigation of memory $CD4^+$ T cells. 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 Killer Cell Lectin-like Receptor subfamily G member 1 (KLRG1) and interleukin-7 receptor-α (CD127)[19]. The $KLRG1^{hi}CD127^{lo}$ population, termed terminal effector cells (TE), is predominantly lost during the contraction phase, while the $KLRG1^{lo}CD127^{hi}$ subset contains memory-precursor cells (MP), which can differentiate into long-lived memory $CD8^+$ T cells[19]. $CD4^+$ T cells also express KLRG1 (ref.[20]) and CD127 (ref.[21]). However, the roles of these molecules in memory $CD4^+$ populations are not well established nor are there clear strategies for distinguishing shorter-lived effector cells 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 TCR transgenic T cells[22–25] and endogenous immune responses[23]. 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. addressed this issue by following the differentiation of individual $CD4^+$ T cells responding to infection[26]. Utilizing over 80 distinct TCR

clones that can specifically respond to *Listeria monocytogenes* (LM) infection, they demonstrated that all microbe-specific naive $CD4^+$ T cells have the potential to give rise to memory cells following acute infection[26]. 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 cell populations contain precursors of memory cells that retain their effector T_H characteristics (Figure 1.1). These data favor the idea that some $CD4^+$ memory T 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.

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 $CD4^+$ SMARTA (SM) T cells, which express an MHC Class II-restricted T cell antigen receptor (TCR) specific for LCMV glycoprotein (GP) amino acids 66–77[24,25,27,28]. 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[23,24,26]. 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, interferon gamma (IFNɣ), 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 (FACS) to isolate T_H1 and T_{FH} effector and memory cells based on known cell-surface receptors and studied their characteristics in the context of reinfection[22,25,26,27].

Marshall et al. found that within the primary effector populations from the spleen at day 8 of infection, two $CD4^+$ T cell subsets which 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[24]. 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 of infection, PSGL-1^{lo}Ly6C^{lo} effector cells showed high expression of known T_{FH} cell-surface receptors (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 memory T cell population at day 150 after infection, suggesting that MP of both T_H1 and T_{FH} phenotypes may persist long-term[24]. 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. that the PSGL-1^{hi}Ly6C^{lo} MP population actually contains both $CXCR5^-$ T_H1 and $CXCR5^+$ T_{FH} cells at comparable frequencies[25]. These results highlight the complexity and heterogeneity within $CD4^+$ memory T cells and the need for further studies.

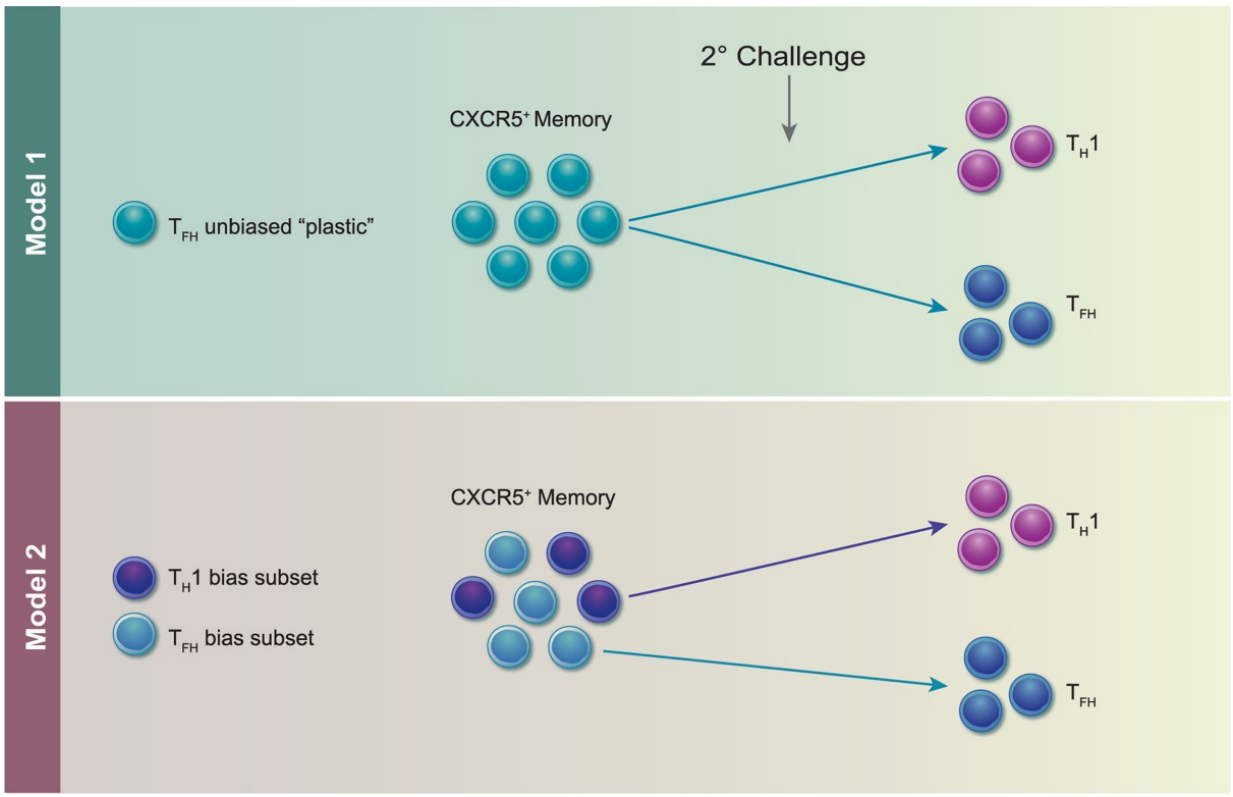


Figure 1.2 Two models for T_{FH} multi-potency. (1) T_{FH} memory cells retain cellular plasticity and can differentiate into T_H1 or T_{FH} secondary effectors based on signals present during secondary challenge; (2) T_{FH} memory cells are actually a heterogeneous population with subsets that are biased or primed towards a particular secondary effector lineage (T_H1 or T_{FH}).

To investigate the potential of T_{FH} memory cells for re-differentiation upon reinfection, Hale et al. utilized expression of CXCR5 and Ly6C to distinguish between T_H1 ($CXCR5^-Ly6C^{hi}$) and T_{FH} ($CXCR5^+Ly6C^{lo}$ & $CXCR5^+Ly6C^{int}$) memory populations following acute infection with LCMV-Armstrong[27], 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[29] as well as in viral influenza infection[28].

In a concurrent study, Pepper et al. addressed $CD4^+$ memory T cell differentiation using LM infection and the expression of CXCR5 and CC chemokine receptor 7 (CCR7), a chemokine receptor regulating trafficking to lymph nodes used as a marker in previous studies to identify central memory T cells (T_{CM}). During acute infection, antigen-specific effector T cells segregated into a $CXCR5^-$ population consistent with the T_H1 phenotype and a $CXCR5^+$ population[30]. 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 which likely include T_{FH} and GC T_{FH} [30]. Notably, Choi et al. found that precursors of T_{FH} or the $CXCR5^+$ populations show greater potential to develop into memory cells compared to T_H1 precursors and share gene expression signatures with memory $CD8^+$ T cells[25]. 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 T cell phenotypes.

Corroborative reports affirming the increased plasticity of T_{FH} memory relative to T_H1 memory upon re-challenge suggests 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 1.2). One possible explanation is that T_{FH} memory cells are inherently more plastic compared to 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 programed 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 down regulated 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. utilized single-cell RNA sequencing to investigate the heterogeneity of antigen-specific $CD4^+$ effector T cells in response to acute

LCMV infection[31]. Visualization of day 7 effector T cells using t-distributed stochastic neighbor embedding (t-SNE) 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.

T_H2 memory

T_H2 memory cells have been best characterized in the context of allergic inflammatory disorders[32], 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 to $CD8^+$ T cells[2], which is why early investigations into T_H2 memory relied on adoptive cell transfers of *in vitro* polarized T_H2 effectors[33]. This system involved activating $CD4^+$ T cells *in vitro* with antigen and antigen-presenting cells (APC) followed by culturing in T_H polarizing conditions[33] 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[34]. 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[34]. Utilizing a similar *in vitro* polarization system, Endo et al. 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 (AHR), and mucus hyper-production in a murine model of T_H2-driven allergic airway inflammation[35]. These studies provided early evidence of the potential existence of T_H2 memory populations, but data demonstrating direct *in vivo* generation was lacking until recently.

A study by Hondowicz et al. provided key insights into T_H2 memory studying the endogenous allergen-specific $CD4^+$ T cells induced in response to house dust mite (HDM) inoculation[36]. Using pMHC class II 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. The authors further explored characteristics of allergen-specific T_H2 resident in the lung, which will be discussed in the lung $CD4^+$ T_{RM} section below.

T_H17 memory

Though not as extensively characterized as other helper subsets, memory T_H17 cells have been documented in both humans and mice, primarily in the context of autoimmunity[37]. Early memory experiments using LM infection showed that T_H17 cells existed only transiently following intranasal infection[23]. However, it is worth noting that LM may not be an optimal infection for T_H17 studies as it is an intracellular pathogen[38] and most efficiently induces T_H1 cells. Muranski et al. 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[39]. In a recent study of dry eye disease (DED) Chen et al. 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[40]. Two cytokines associated with $CD4^+$ memory T lymphocytes, IL-7 and IL-15 (ref. [41]), 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 conjunctivae 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 T cells remains uncertain.

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. However, in the next section, we will review two biological requirements crucial for memory formation that appear to be conserved across all T_H subsets.

Memory differentiation cues: TCR signaling & IL-2

A comprehensive review of studies aimed at resolving the signals required for $CD4^+$ memory T cell formation[42] discussed the instructive signals both during the "early priming" phase of initial antigen recognition and activation as well as at "late-acting checkpoints" prior to contraction that play a role in the effector-to-memory transition. Much like the signals important for $CD8^+$ memory T cell generation, strengths of TCR and co-stimulatory signaling also have profound effects on memory T_H development[42,43]. Recent results from Snook et al. demonstrated that TCR signaling has a direct impact on T_H memory formation[44]. Utilizing a panel of TCRs specific for the same viral antigen, the authors showed substantial variability in TCR signal strength, expression of IL-2-receptor alpha (CD25), and activation of downstream TFs across the $CD4^+$ memory T cell population[44]. TCR clones with stronger TCR signaling appear to differentiate towards a more TE state and become largely depleted by memory time points, while clones with comparatively lower signaling were memory-like and able to persist after antigen clearance.

Interestingly, it seems that stronger TCR signaling was associated with higher expression of T_H1 cell-surface receptors, while weaker TCR signals correlated with higher expression of T_{FH} cell-surface receptors[44], suggesting that there may be a connection between lineage differentiation and memory potential for $CD4^+$ helper T cells.

Utilizing influenza A virus (IAV) as an infection model, Mckinstry et al. 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 cells[45]. 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 wildtype and IL-2 deficient donors showed similar cell numbers at the peak of infection and production of IFNɣ; 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. with 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[46], 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. identified a role for TF Oct1 and its coactivator OCA-B in poising the *Il2* locus for robust expression in memory $CD4^+$ T cells[47], unveiling an important mechanism by which memory $CD4^+$ T cells control IL-2 production. However, these studies regarding TCR signaling and IL-2 in $CD4^+$ memory T cells were completed without investigation of specific T_H lineages. Therefore, further investigation into the required transcriptional and epigenetic regulation for generation and maintenance of memory T_H subsets is

needed. Additionally, while targeting peripheral memory T cells in vaccination strategies can provide systemic protection, in some cases, a localized strategy in which tissue-resident memory T cells at barrier surfaces are activated as front-line defense against recurrent infections may be more effective, thus how these signals pertain to T_{RM} will be informative[48].

1.4 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. 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[49,50]. Thus, all circulating cells 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 labeled antibody to mark cells in the circulating system and distinguish them from cells in the tissues[51]. Any cells positive for the label are considered "circulating" while unlabeled cells are assumed to have limited access to circulation and are therefore "tissue-resident"[49]. To determine the protective functions of tissue-resident lymphocytes in secondary infection, immune mice were 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[52]. When these mice were re-challenged with the original pathogen, any immune response at the local site of infection would be mediated only by cells resident to that tissue[53].

Recent studies have highlighted a prominent population of long-lived $CD4^+$ T cells within many non-lymphoid tissues (NLTs) including the lungs[36,49,54–61], small intestine (SI)[12,18,62–64], skin[15,65–68], and female reproductive tract (FRT)[18,69,70] (Figure 1.3). A comprehensive assessment of $CD4^+$ T_{RM} populations in mucosal surfaces using parabiosis experiments revealed that the majority of $CD4^+$ T cells in the SI, FRT, salivary glands, kidneys, and liver were tissue-resident and expressed CD69 consistently[71]. Depletion of $CD4^+$ T_{RM} in the FRT resulted in decreased frequencies of effector $CD8^+$ T cells and dendritic cells in the tissue, highlighting the importance of $CD4^+$ T_{RM} in alarm and recruitment functions to facilitate rapid immune defense upon re-exposure to antigen[71,72]. To assess how $CD4^+$ and $CD8^+$ T_{RM} behave in the same tissues, cell-surface receptor expression, cytokine production, and gene expression profiles of LCMV-specific memory $CD4^+$ and $CD8^+$ T cells in the FRT and SI were analyzed following infection[71]. Both $CD4^+$ and $CD8^+$ T_{RM} expressed CD69, though Ly6C and CD103 expression was minimal for $CD4^+$ T_{RM} compared to $CD8^+$ T_{RM}. Additionally, the core $CD8^+$ T_{RM} signature was enriched in $CD4^+$ T_{RM} samples, suggesting shared transcriptional regulation for T_{RM}. However, this study and others also observed that while residency was the norm for $CD4^+$ T_{RM}, there was some equilibration between $CD4^+$ T_{RM} and circulating populations in most NLTs, compared to $CD8^+$ T_{RM}[71]. Further investigation is required to address the recirculation potential and developmental origins of $CD4^+$ T_{RM} and how tissue-specific cues can influence the development of this critical population

Lung $CD4^+$ T_{RM} cells

The lungs contain a population of $CD4^+$ T_{RM} that play a critical role in recruiting $CD8^+$ T cells and enhancing secondary immune responses against bacterial, viral, and worm infections[49,56,60,73,74]. In an influenza infection model, antigen-specific memory $CD4^+$ T cells migrated to the lungs and were retained in the tissue without recirculation, as demonstrated by

parabiosis experiments[49]. This subset of $CD4^+$ T cells showed a distinct phenotype from circulating populations, specifically high expression of CD69, a membrane bound type II C-lectin receptor and known marker of tissue retention[75,76]. Functionally, these lung-resident $CD4^+$ T cells provided protection from influenza virus when transferred to naive mice. Similar to its role in directing long-term memory fate as discussed above, IL-2 also supported the formation of antigen-specific lung $CD4^+$ T_{RM} with a transcriptional signature distinct from that of circulating $CD4^+$ T cell populations but similar to that of $CD8^+$ T_{RM}[56,77]. Interestingly, Strutt et al. also identified an IL-2-independent population of influenza-specific lung $CD4^+$ T_{RM} following infection, suggesting that IL-2 may not be the only cytokine regulating lung T_{RM} development and maintenance. In fact, IL-15 was shown to be essential for these IL-2-independent cells, acting as an "alarm" at local sites of infection to promote both $CD8^+$ T cell responses and induce long-lived $CD4^+$ T_{RM}[42,56]. Additionally, IL-17A producing $CD4^+$ T_{RM} cells in the lung following *Streptococcus pneumoniae* (pneumococcus) infection were shown to remodel the lung epithelial responses which increased recruitment of neutrophils into the tissue[78]. These results highlight the importance of lung $CD4^+$ T_{RM} in coordinating a network of cells within the microenvironment to ensure an effective response.

Studies of lung T_{RM} have also addressed the question of how circulating effector subsets contribute to the T_{RM} population. Using fluorescently labeled antibody injection, Hondowicz et al. showed that LCMV-specific $CD4^+$ T cells migrated to the lungs as a T-bet^{hi} T_H1 subset with CD69 expression similar to influenza experiments, suggesting a developmental relationship between T_{RM} and T_H1[54]. This establishment of LCMV-specific lung T_H1 T_{RM} cells required IL-2 signaling; however, for long-term maintenance and survival, additional interactions with B cells in tissues were necessary[54]. Alternatively, two recent studies by Swarnalekha *et al*. and Son *et al*. identified a subset of influenza-specific lung $CD4^+$ T_{RM} which were phenotypically similar to T_{FH} cells,

expressed the master regulator of TFH lineage Bcl6, and were critical for stimulating local $CD8^+$ T cell and B cell responses[79–81]. This "resident helper T cells" (T_{RH}) subset was distinct from the T_H1-like $CD4^+$ T_{RM} cells and depletion of the T_{RH} population resulted in a decrease of flu-specific antibody-secreting B cells. These findings raise further questions about the developmental origins of $CD4^+$ T_{RM}, specifically the cues determining differentiation of T_{FH}- versus T_H1-like subsets, and how this paradigm may apply in other tissues.

Several groups have shown that allergens in an asthma model also elicit $CD4^+$ T_{RM} responses in the lung. After exposure to HDM, allergen-specific $CD4^+$ T_{RM} in the lungs were shown to be resident by parabiosis experiments, expressed high levels of CD69, and required IL-2 signaling for their migration to the lungs[36]. Additionally, Bcl6, the fate-determining factor for T_{FH} cells, prevented HDM-specific $CD4^+$ T_{RM} from entering the lungs, and loss of Bcl6 actually increased the T_{RM} population in the tissue[36]. These results indicate an antagonistic relationship between T_{RM} and memory T_{FH} cells and highlight a role for TFs in directing the development of tissue-resident populations. In another study using the HDM model, lung $CD4^+$ T_{RM} were shown to be T_H2-like cells, though functionally and transcriptionally distinct from the circulating T_H2 subset, which has an important role in driving asthma[82]. Within the lungs, $CD4^+$ T_{RM} clustered around the airways at rest and rapidly reactivated upon secondary exposure within their clusters to produce IL-4, IL-5, and IL-17 and recruit dendritic cells[55.] Additional studies highlight a role for IL-7, along with IL-2, in the maintenance of allergen-specific $CD4^+$ T_{RM} cells in the lung parenchyma and airways. However, these IL-7-dependent $CD4^+$ T_{RM} did not express CD69 (ref. [57]), suggesting either that CD69 may not be a conclusive marker for $CD4^+$ T_{RM} or that $CD4^+$ T_{RM} are highly heterogeneous, and further studies must focus on elucidating the different subsets. Additionally, in both viral infection and allergen-induced asthma models, it remains unclear

whether the same cells can become resident in both the parenchyma and airways, and if similar survival signals sustain such subsets.

Skin CD4$^+$ T_{RM} cells

A second well-studied tissue for CD4$^+$ T_{RM} cells is the skin, particularly in a herpes simplex virus (HSV) infection model. Initial studies by Gebhardt et al. using intravital microscopy demonstrated that HSV-specific gDT-II CD4$^+$ T cells showed a migration pattern distinct from that of CD8$^+$ T cells in the skin following infection, homing to the dermis as opposed to the preference of CD8$^+$ T cells for the epidermis[15]. Additionally, dermal T_H cells were significantly more motile and had lower expression of CD103 compared to CD8$^+$ T cells. Thus, it appears that CD4$^+$ T_{RM} can migrate between the skin and circulation much more easily than CD8$^+$ T_{RM} and may not be a permanenntly resident population. In another study of skin HSV response, Collins et al. examined skin from naive mice and found resting CD4$^+$ T cells preferentially clustered around hair follicles. Upon infection, circulating CD4$^+$ T cells were rapidly recruited to the skin, specifically to the perifollicular regions; the majority of recruited cells were lost by day 30 but a small percentage survived in tissue long after initial antigen exposure[68]. These skin CD4$^+$ T_{RM}, as demonstrated by parabiosis experiments, also aggregated in APC-associated clusters, consistent with experiments done in FRT[70,83], as discussed below. Additional studies of skin CD4$^+$ T_{RM} following *Leishmania major* (*L. major*) infection showed evidence of a *L. major*-specific resident population in the dermis with increased expression of interferon (IFN) genes and chemokine signaling pathways[65]. These CD4$^+$ T_{RM} cells were able to control parasite growth following a secondary challenge, presumably through the aforementioned IFN response and recruitment of circulating immune cells. Through FTY-720 treatment experiments, resident CD4$^+$ T cells were found to be the main contributors to protection against *L. major*, recruiting inflammatory monocytes for production of

reactive oxygen species to kill the parasites[66]. These results highlight another function of $CD4^+$ T_{RM} in enhancing secondary responses and protection against recurring pathogens.

$CD4^+T_{RM}$ cells in other mucosal tissues

Through parabiosis experiments, $CD4^+$ T_{RM} have been shown to be active in the FRT in response to HSV-2 infection[70,84]. Intravaginal HSV-specific $CD69^+$ $CD4^+$ T_{RM} localized to memory lymphocyte clusters (MLCs) containing $CD8^+$ T cells, macrophages, and other APCs and were maintained in these structures by chemokines secreted by macrophages. Upon a secondary re-challenge with lethal HSV-2, resident T_H cells alone were sufficient for protection, largely mediated by $CD4^+$ T_{RM} production of IFNγ to inhibit viral replication. In this context, FRT-resident $CD4^+$ T cells were more directly involved in clearing the infection as opposed to their classic role in alarming and recruiting circulating populations.

The small intestine contains a potential subset of tissue-resident $CD4^+$ T cells, which can be found as both intraepithelial (IEL) and lamina propria lymphocytes (LPL), with a greater proportion of $CD4^+$ T cells in the LPL. The majority of studies of gut-resident $CD4^+$ T cell subsets have focused on endogenous polyclonal IEL which originate from the circulating population[63,85,86]. Upon entering the small intestine, these $CD4^+$ IELs downregulated the TF ThPOK responsible for thymic $CD4^+$ T cell fate and upregulated Runx3, the fate-determining factor for differentiating $CD8^+$ thymocytes. This switch in programming towards a more cytolytic function appeared to be important in the $CD4^+$ T cell response to endogenous gut microbiota[85]. Nevertheless, this study characterized polyclonal $CD4^+$ T cells with variable TCR specificity and affinity, and further investigations are necessary to determine whether similar transcriptional changes are seen in antigen-specific responses. In an LM infection, antigen-specific $CD4^+$ T cells preferentially migrated into the LPL (though a small population was seen in the IEL) to form a long-lived,

predominantly T_H1, antigen-specific memory population, expressed high levels of CD69, and unlike lung-resident $CD4^+$ T cells, was independent of IL-15 signaling[64.] Interestingly, this gut "T_{RM}" population had low Ly6C expression relative to other lymphocytes in circulation, similar to the MP population of circulating T_H1 cells discussed in the T_H1 and T_{FH} $CD4^+$ memory T cells section above. This suggests that circulating and resident memory cells may share a common precursor.

As discussed in the studies above, $CD4^+$ T_{RM} differ greatly between infection models and tissues, although some general principles can be drawn (Figure 1.3). As is the case with $CD8^+$ T_{RM}, the majority of $CD4^+$ T_{RM} express high levels of CD69, but CD103 expression varies within and between tissues[15,49,64,70]. Across multiple tissues – lung, skin, FRT – $CD4^+$ T_{RM} form clusters with other resident immune subsets including $CD8^+$ T cells, macrophages, and APCs[57,68,70]. These clusters position $CD4^+$ T_{RM} in close proximity to the cells they need to activate in case of a recurrent infection. This may be an optimized way for helper T_{RM} cells to perform their "sense-and-alarm" function[50,71], initiating the innate response to recruit circulating cells and activating $CD8^+$ T cells to fight off pathogen. Additionally, cytokines appear to be critical in the recruitment, formation, and maintenance of $CD4^+$ T_{RM} in many tissues[55–57,64]. However, which cytokines are important for which tissues remain unknown, although IL-2 may play a key role for all T_{RM} given its known effects in circulating memory populations[87]. It is also unclear which $CD4^+$ T_H effector subset, if any, gives rise to T_{RM}, or whether the tissue-resident precursor is distinct from the peripheral population completely, residing in the tissue until the right signals are detected to induce differentiation.

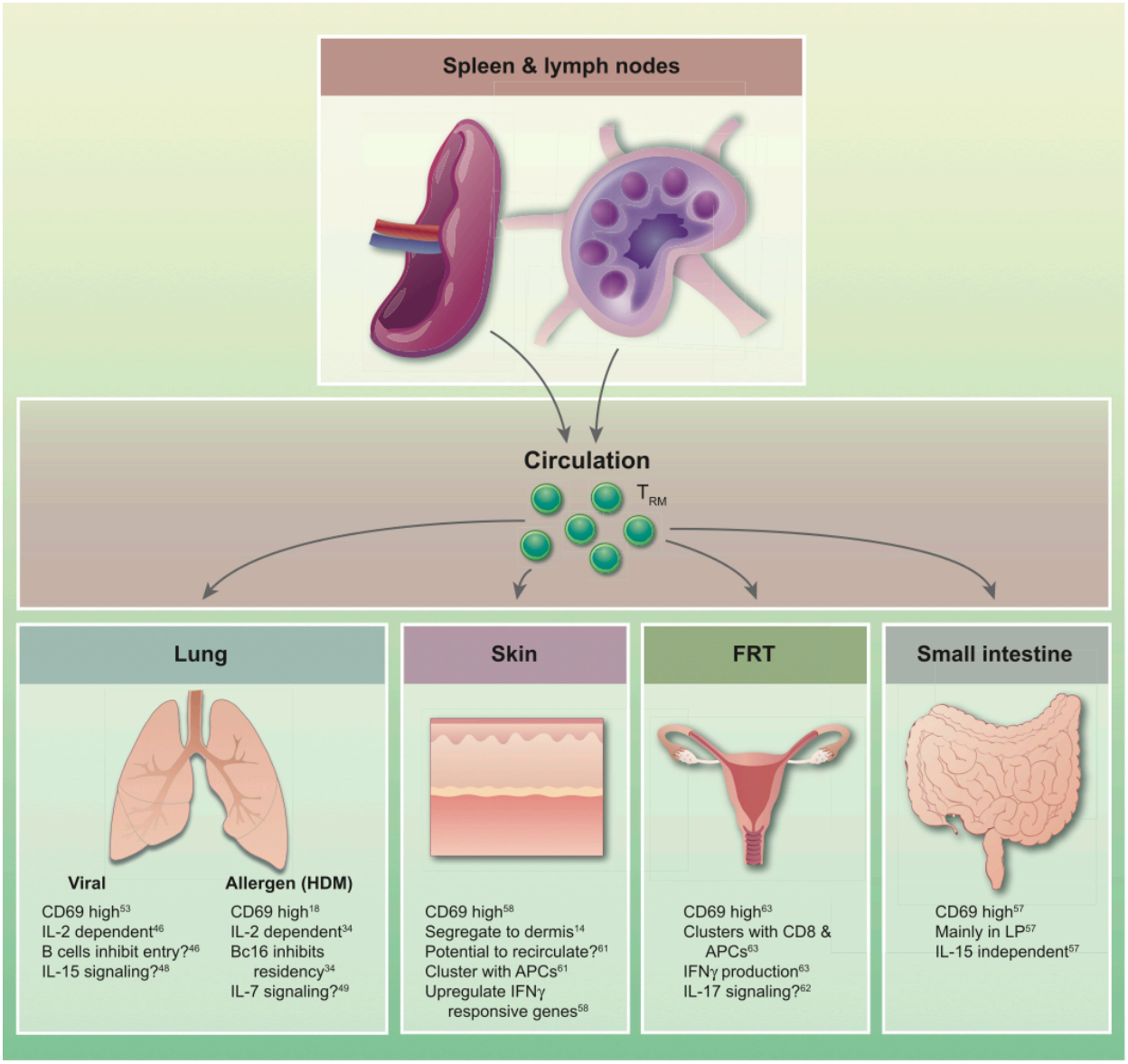


Figure 1.3 Following infection, CD4$^+$ tissue-resident memory cells (T_{RM}) are recruited to tissues from circulation and secondary lymphoid organs (SLOs). The majority of CD4$^+$ T_{RM} express high levels of CD69 and can form clusters with other resident immune subsets, including CD8$^+$ T_{RM}, macrophages and antigen-presenting cells (APCs). Cytokine signaling may also play a role in recruitment and retention, although specific cytokines are preferred by each tissue.

1.5 Conclusion

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 cell-surface receptors to differentiate between memory T_H1 and T_{FH} memory subsets[24,27], 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[25]. Likewise, in other infection systems which elicit T_H2 or T_H17 effector cells, we do not know how these effector subsets contribute to the final population of memory cells. Adding further to the complexity, $CD4^+$ T_{RM} populations are highly variable across tissues and, their establishment and retention are likely driven by tissue-specific cues similar to $CD8^+$ T_{RM} cells[88]. It is unclear whether the precursor for $CD4^+$ T_{RM} cells originates from the same memory-precursor population that yields circulating memory cells or if it is found even earlier, before the effector versus memory cell decision. Both non-T_{FH}, such as T_H1 or T_H2 cells, and T_{FH} subsets can contribute to the $CD4^+$ T_{RM} pool, but we do not know what cues drive their differentiation and if both populations exist in all tissues. Lastly, it is clear that cell-surface receptors do not accurately identify sub-populations in effector and memory cells, and further work requires examining TFs and transcriptional regulators which may direct the memory program.

In this dissertation, I aim to address these outstanding questions, focusing on the $CD4^+$ T_{RM} population, using bulk and single-cell epigenetic and transcriptional profiling of cells in the circulation and tissues over the course of an infection[77,89,90]. Chapter 2 addresses the relationship between $CD4^+$ T_{RM} cells and circulating effector subsets, specifically examining how T_H1 cells may contribute to the T_{RM} population. Chapter 3 focuses on the role of T_H1-fate associated factors Id2 and Blimp1 and T_{FH}-fate master regulator Bcl6 in intestinal $CD4^+$ T_{RM} differentiation to

elucidate the transcriptional program of this critical immune subset. This work will further inform our knowledge of transcriptional regulation involved in $CD4^+$ T_{RM} cell biology and identify potential targets for therapeutic purposes.

1.6 Acknowledgements

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

This review was supported by the American Heart Association Predoctoral Fellowship to QPN, and US National Institutes of Health R01 AI072117 and U19 AI109976 to AWG. The authors thank Drs. Kyla D. Omilusik and J. Ty Crowl for critical review of the manuscript.

Chapter 2 Anti-viral $CD4^+$ T_{RM} in the small intestine share a developmental relationship with effector T_H1 cells

2.1 Introduction

Prominent populations of long-lived $CD4^+$ T_{RM} cells reside in many tissues including lung[36,49,73,54–61], skin[16,65–68], female reproductive tract (FRT)[69,70], and small intestine (SI)[62–64]. T_{RM} are critical for recall responses by producing cytokines to create a pro-inflammatory environment and recruiting innate immune cells, $CD8^+$ T cells, and B cells to the specific tissue upon reinfection or antigen exposure[71]. While the phenotype, function, and regulation of $CD4^+$ T_{RM} differ between infection models and tissues, $CD4^+$ T_{RM} express high levels of CD69 and physically cluster and interact with other immune cell populations for optimized function and require tissue-specific cytokines for their recruitment and maintenance. Lung $CD4^+$ T_{RM} play a critical role in responses against bacterial, viral, and worm infections and in allergic asthma[36,54,60,79,80]. $CD4^+$ T_{RM} cells have also been detected in the skin in herpes simplex virus (HSV) and *Leishmania major* infections [65,66,68]. In the FRT, intravaginal HSV-specific $CD69^+$ $CD4^+$ T_{RM} localized in clusters containing $CD8^+$ T cells, macrophages, and other APCs, and produced interferon-gamma (IFNγ) to directly inhibit viral replication[70]. Furthermore, $CD4^+$ T cells can be found in either the intraepithelial compartment (IEL) or lamina propria (LPL) of the SI, with a greater proportion residing in the LPL[91].

It has been challenging to address the developmental origins of $CD4^+$ T_{RM} due to the heterogeneity of the effector T cell population and unresolved mechanisms of $CD4^+$ memory T cell differentiation[92]. Studies have shown that $CD4^+$ T_{RM} are developmentally linked to the T_H cell subsets with characteristics of the relevant helper effector cells, such as intestinal T_H1 cells in

Listeria infection, lung T_H2 cells in allergic asthma, and lung T_H17 cells in *Klebsiella pneumonia* immunization. This raises the question of whether there is a true T_{RM} precursor cell or whether T_{RM} cells are a subset of prolonged effector T cells[36,54,64,70,71,80,93]. However, $CD8^+$ T_{RM} cells, which also migrate to non-lymphoid tissues (NLT) to provide barrier protection against re-infection, have been compared to T_{FH} cells[94]. To facilitate trafficking to their respective tissues (GCs or NLT) and prevent egress, memory T_{FH} and $CD8^+$ T_{RM} cells both express the C-type lectin CD69[95,96] and downregulate transcription factor (TF) Kruppel-like factor 2 (Klf2)[97,98] and its downstream target, egress factor sphingosine-1-phosphate receptor 1 (S1pr1). Additionally, both T_{FH} and $CD8^+$ T_{RM} subsets express ICOS, a costimulatory molecule which is thought to induce downregulation of Klf2[99–102] and Bcl6, a transcriptional repressor that drives the T_{FH} lineage and promotes the $CD8^+$ T cell memory program[103–107]. Thus, it is not clear whether $CD4^+$ T_{RM} share a relationship with effector T_H subsets or if they are related to T_{FH} cells similar to their $CD8^+$ T_{RM} counterparts.

In the following chapter, we characterized $CD4^+$ T_{RM} in the SI following acute infection with Lymphocytic Choriomeningitis Virus-Armstrong (LCMV-Arm), focusing on key regulators of $CD4^+$ T_{RM} differentiation, tissue-residency, and homeostasis. $CD4^+$ T_{RM} exhibited an effector-like identity, with cell-surface molecule expression, cytokine production, and transcriptional enrichment of the T_H1 gene expression signature. Notably, bulk RNA sequencing (RNAseq) demonstrated that $CD4^+$ T_{RM} expressed genes of both the effector and memory T cell programs, and co-expression by the same cells was confirmed using single-cell RNA sequencing (scRNAseq). We also investigated the heterogeneity within the $CD4^+$ T_{RM} population at both effector and memory time points, including differences between the IEL and LPL subsets.

2.2 Results

Antiviral CD4$^+$ T_{RM} phenotypically resemble circulating T_H1 cells

To establish our experimental protocol and confirm the presence of CD4$^+$ T_{RM} in the SI, we first transferred TCR transgenic SMARTA CD4$^+$ T cells, which recognize an LCMV glycoprotein (GP) peptide presented by MHC-Class II, into congenically distinct host mice that were subsequently infected with LCMV-Arm. On day 20 of infection, we intravenously injected mice with anti-CD4 (RM4.5) antibody to mark CD4$^+$ T cells in the circulation before collecting tissues. We found that SMARTA CD4$^+$ T cells in the IEL and LPL were negative for the anti-CD4 IV label, confirming that these cells were "tissue-resident", while a high proportion of the cells in the vascularized tissues of spleen, liver, lung, and kidney were IV$^+$ (Figure 2.1A). Since all cells in the SI were IV$^-$, we did not continue with IV labeling in further studies of IEL and LPL T_{RM} populations.

The developmental origins of antiviral CD4$^+$ T_{RM}, specifically how T_{RM} cells are related to effector CD4$^+$ T_H subsets, remain largely unresolved. To address this question, we studied SMARTA CD4$^+$ T cells at days 7 and 21 of infection from the spleen, mesenteric lymph nodes (mLN), and both IEL and LPL compartments of the SI for characterization by flow cytometry. Following infection, circulating CD4$^+$ T cell effector subsets can be distinguished by expression of the cell-surface molecules signaling lymphocytic activation molecule (SLAM) and CXC-chemokine receptor type 5 (CXCR5), with T_H1 cells expressing high levels of SLAM while T_{FH} cells express CXCR5 to traffic to the germinal centers[108,109]. Consistent with this model, SMARTA CD4$^+$ T cells in the spleen and mLN on days 7 and 21 were either SLAM$^+$ T_H1 cells or CXCR5$^+$ T_{FH} cells, with a higher frequency of T_{FH} cells present on day 21 (Figure 2.1B,C). Conversely, in

the SI, the majority of $CD4^+$ T cells shared a phenotype with T_H1 cells at both day 7 and 21 after infection.

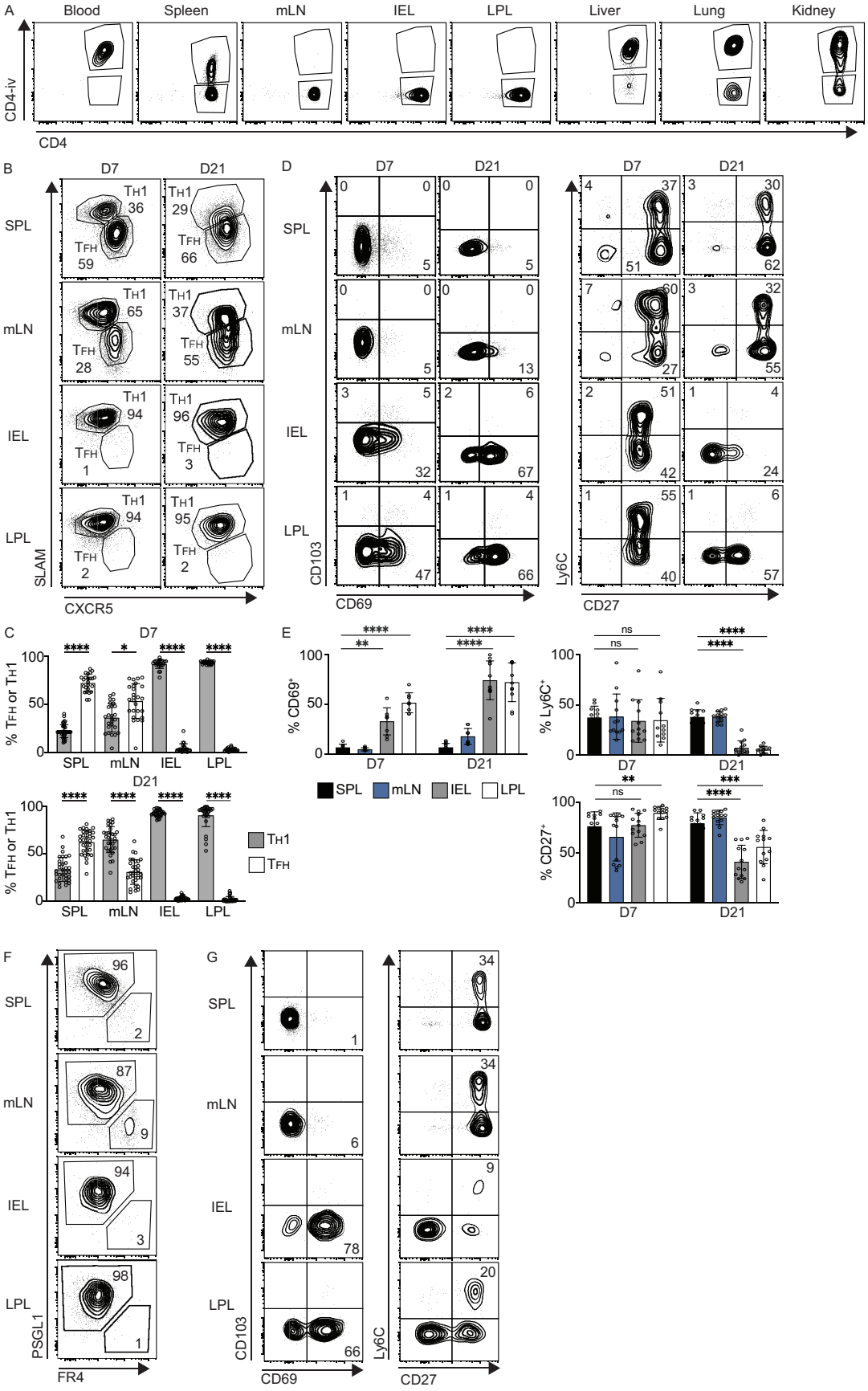
To distinguish SI $CD4^+$ T_{RM} from their circulating memory counterparts, we analyzed expression of molecules previously associated with tissue-residency and/or memory by SI $CD4^+$ T_{RM}. $CD8^+$ T_{RM} were distinguished from the circulating memory populations by their expression of the C-type lectin CD69 and the integrin CD103[76,97,110]. We confirmed that the majority of $CD4^+$ T_{RM} in both IEL and LPL expressed CD69 but not CD103 compared to $CD4^+$ T cells of the mLN and spleen (Figure 2.1D,E) consistent with previous studies[71].

CD27, a costimulatory marker expressed by naive T lymphocytes, has been shown to be critical in memory T cells development[111]. In humans, CD27 was used to distinguish between memory $CD4^+$ T cell subsets, with high CD27 expression associated with less antigen recall ability and greater plasticity[112,113]. At the peak of LCMV-Arm infection, circulating and SI SMARTA $CD4^+$ T cells were predominantly $CD27^+$ (Figure 2.1D,E). By day 21 of infection, the majority of spleen and mLN $CD4^+$ T cells remained $CD27^+$, while only half of IEL and LPL cells still expressed CD27. Lymphocyte antigen 6 complex (Ly6C) was also used in studies of circulating $CD4^+$ memory T cells to distinguish between terminally differentiated versus multi-potent subsets[24,27]. Specifically, $Ly6C^-$ memory cells were found to have increased proliferation upon rechallenge and ability to differentiate into either T_H1 or T_{FH} cells, compared to the $Ly6C^+$ population[24,27]. At the peak of infection, ~50% of $CD4^+$ T cells in all tissues expressed Ly6C (Figure 2.1D,E). However, by day 21, the majority of SI $CD4^+$ T_{RM} had downregulated Ly6C, while the frequencies in circulating populations remained stable. The low expression of both CD27 and Ly6C was consistent with other studies of intestinal $CD4^+$ T cells at memory time points[64,71].

A recent study identified two subsets of influenza-specific $CD4^+$ T cells in the lung, canonical T_{FH} memory cells and a new population of T resident helper (T_{RH}) cells with characteristics of T_{RM} and T_{FH} cells, which can be distinguished by their expression of P-selectin glycoprotein ligand-1 (PSGL1) and folate receptor 4 (FR4), respectively[79,80]. We assessed the expression of both molecules by SI $CD4^+$ T_{RM} populations on day 21 of infection and did not see equivalent subsets (Figure 2.1F), which suggested that infection model and tissue-specific cues likely influence the composition of the $CD4^+$ T_{RM} population.

To determine whether the SI $CD4^+$ T_{RM} phenotype differed between different antigens, we obtained an additional TCR transgenic mouse (NIP) which are $CD4^+$ T cells specific for the LCMV nucleoprotein (NP) peptide 311-325, the most abundantly produced viral protein during LCMV infection and the main target of the immunodominant antibody response[114]. Relative to GP-specific SMARTA cells, NP-specific NIP $CD4^+$ T cells more efficiently help NP-specific B cells and generate greater numbers of T_{FH} during infection[115]. We found that in response to LCMV-Arm infection, NIP and SMARTA $CD4^+$ T cells showed relatively similar numbers of intestinal $CD4^+$ T_{RM} at memory time points and shared $CD69^+$ expression (Figure 2.1G). There was also a population of $CD27^+Ly6C^+$ NIP cells in the IEL and LPL at day 21 compared to minimal Ly6C expression by SMARTA SI cells. The similarities between SMARTA and NIP T_{RM} responses provide us with another comparable model for studying $CD4^+$ memory T cell responses in the periphery and non-lymphoid tissues.

Figure 2.1 $CD4^+$ T_{RM} cells resemble circulating T_H1 effector cells during viral infection. (**A**) Representative flow plots showing staining for anti-CD4 i.v. label to distinguish circulating and resident SMARTA $CD4^+$ T cells from the spleen (SPL), mesenteric lymph nodes (mLN), intraepithelial lymphocytes (IEL) and lamina propria lymphocytes (LPL) of the small intestine (SI), liver, lung, and kidney. Mice were infected with LCMV-Arm for 20 days. (**B-C)** Expression of SLAM and CXCR5 by SMARTA $CD4^+$ T cells in the SPL, mLN, IEL, and LPL on days 7 and 21 of LCMV-Arm infection. Representative flow cytometry plots (B) and quantification (C) of T_H1 ($SLAM^+CXCR5^-$) and T_{FH} ($SLAM^-CXCR5^+$) SMARTA cells. (**D-E**) Expression of CD69, CD103, CD27 and Ly6C by SMARTA $CD4^+$ T cells in specified tissue on indicated day of infection. Representative flow cytometry plots (D) and quantification of frequencies of $CD69^+$, $CD27^+$, and $Ly6C^+$ cells from D. **(F)** FR4 and PSGL1 expression by specified tissue on day 21 of LCMV-Arm infection. Numbers in flow plots indicate percent of cells in corresponding gate. **(G)** Representative flow plots showing expression of CD69, CD103, CD27 and Ly6C by NIP $CD4^+$ T cells in specified tissue on day 21 of infection. Data are representative (B,D,F,G) or cumulative (C,E) of 3 experiments (B-E) with n=2-4 mice per experiment. Graphs show mean ± SD; *$p < 0.05$, ** $p<0.01$, *** $p<0.001$, ****$p< 0.0001$.


A
Blood
Spleen
mLN
IEL
LPL
Liver
Lung
Kidney
CD4-iv
CD4
B
D7
D21
SPL
mLN
IEL
LPL
TH1
TFH
SLAM
CXCR5
D
CD103
CD69
Ly6C
CD27
C
% TFH or TH1
E
% CD69+
% Ly6C+
% CD27+
SPL
mLN
IEL
LPL
F
PSGL1
FR4
G

Antiviral CD4$^+$ T cells with tissue migration potential are T_H1 cells and produce effector cytokines

C-C chemokine receptor type 9 (CCR9) and integrin alpha 4 (CD49d) are two gut-homing molecules that are upregulated by cells that traffic from the circulation to the SI at early stages of viral infection[116]. To investigate whether SI-homing cells were derived from T_H1 or T_{FH} cells at this early time point, we profiled CCR9 and CD49d expression by SMARTA cells in the spleen, mLN, and SI of mice infected with LCMV-Arm between days 6 to 8 of infection to capture the cells during peak effector response (Figure 2.2A,B). We observed an increase in CCR9$^+$CD49d$^+$ CD4$^+$ T cells in the IEL and LPL as cells migrated from the circulation between days 6 and 8 of infection. In contrast, only a small population of cells in the spleen expressed both CCR9 and CD49d during these three days. Interestingly, of the CCR9$^+$CD49d$^+$ cells, more than 80% of cells from the spleen and 90% of cells from the lymph nodes were SLAM$^+$ T_H1 cells (Figure 2.2C,D), indicating a developmental relationship between SI-specific T_{RM} and the T_H1 lineage.

To assess the functional capacity of anti-viral SI CD4$^+$ T_{RM} and splenic effector cell subsets, we isolated SMARTA cells from each tissue on day 7 or 21 of infection then re-stimulated them *ex vivo* with the LCMV-Arm GP66 peptide. On day 7 of infection, more CD4$^+$ T cells in the SI produced IFNγ and/or TNFα compared with the splenic population upon peptide re-stimulation (Figure 2.3A,B). However, at a memory time point, there were fewer cytokine-producing cells in the SI compared to circulation. SI CD4$^+$ T cells also expressed higher levels of Granzyme A and B at both effector and memory time points (Figure 2.3C,D). This was similar to the T_H1 cytokine profile and consistent with the effector nature of T_{RM} cells, in that they must be poised for rapid response upon reinfection at barrier surfaces.

It has been shown that following influenza A infection, a cytotoxic subset of CD4$^+$ T cells, called ThCTL, developed in the lung[117]. These cells mediated MHC-II-restricted cytotoxicity,

produced high levels of effector cytokines, and can be distinguished from other $CD4^+$ T cell subsets by expression of inhibitory molecule NKG2A/C/E[118]. To determine if there were parallels between ThCTL and T_{RM} in viral infection, we compared LCMV-specific $CD4^+$ T_{RM} in both the IEL and LPL on day 21 of infection. We found that $CD4^+$ T_{RM} in both compartments expressed minimal levels of NKG2A/C/E compared to circulating $CD4^+$ T cells, $CD8^+$ circulating and T_{RM} cells, or NK cells (Figure 2.3E). Despite their shared expression of IFNγ and Granzyme B, SI-resident $CD4^+$ T_{RM} cells did not express canonical cell-surface receptors of ThCTL, which further indicated that the T_{RM} program may be influenced by different infection and tissue-specific cues.

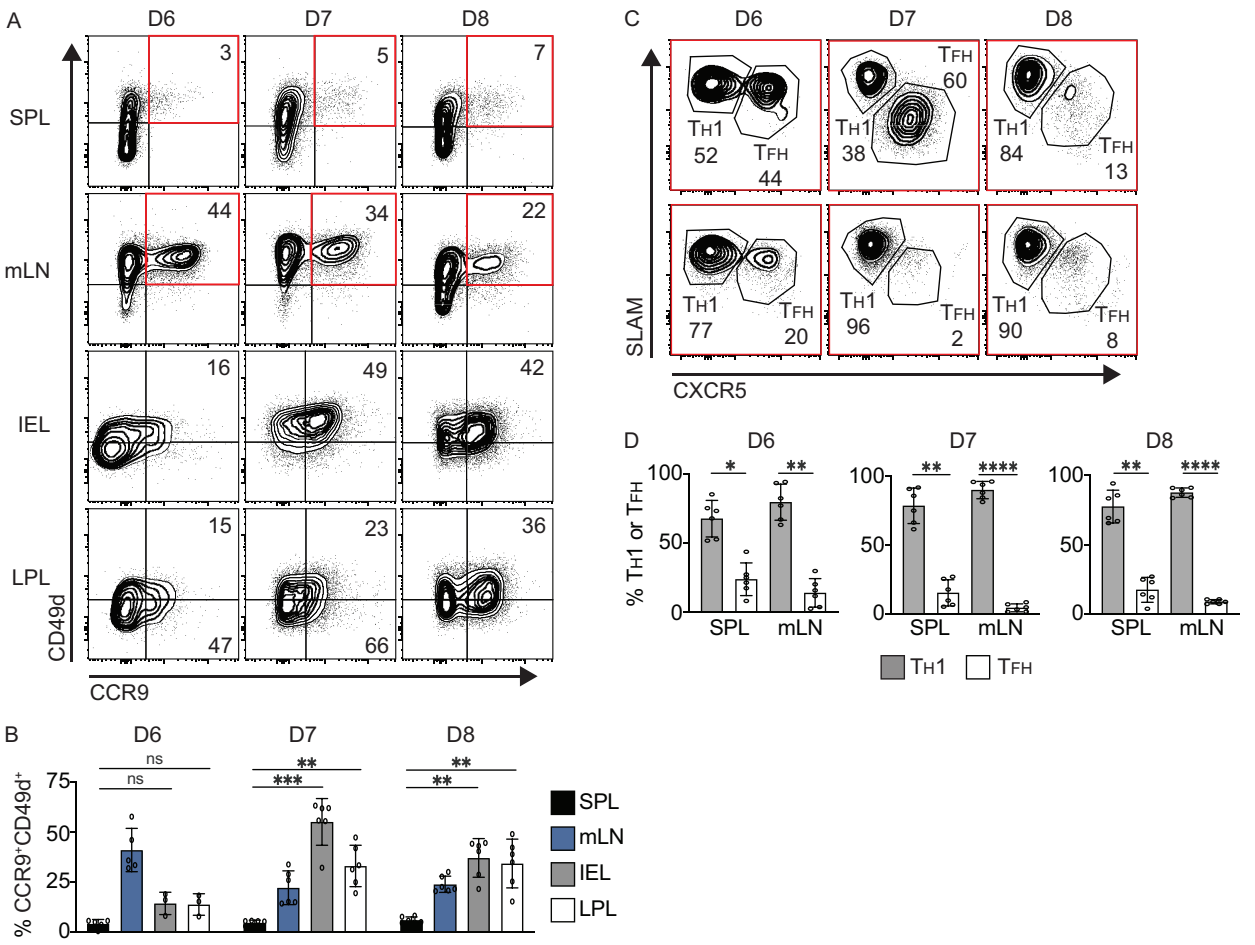
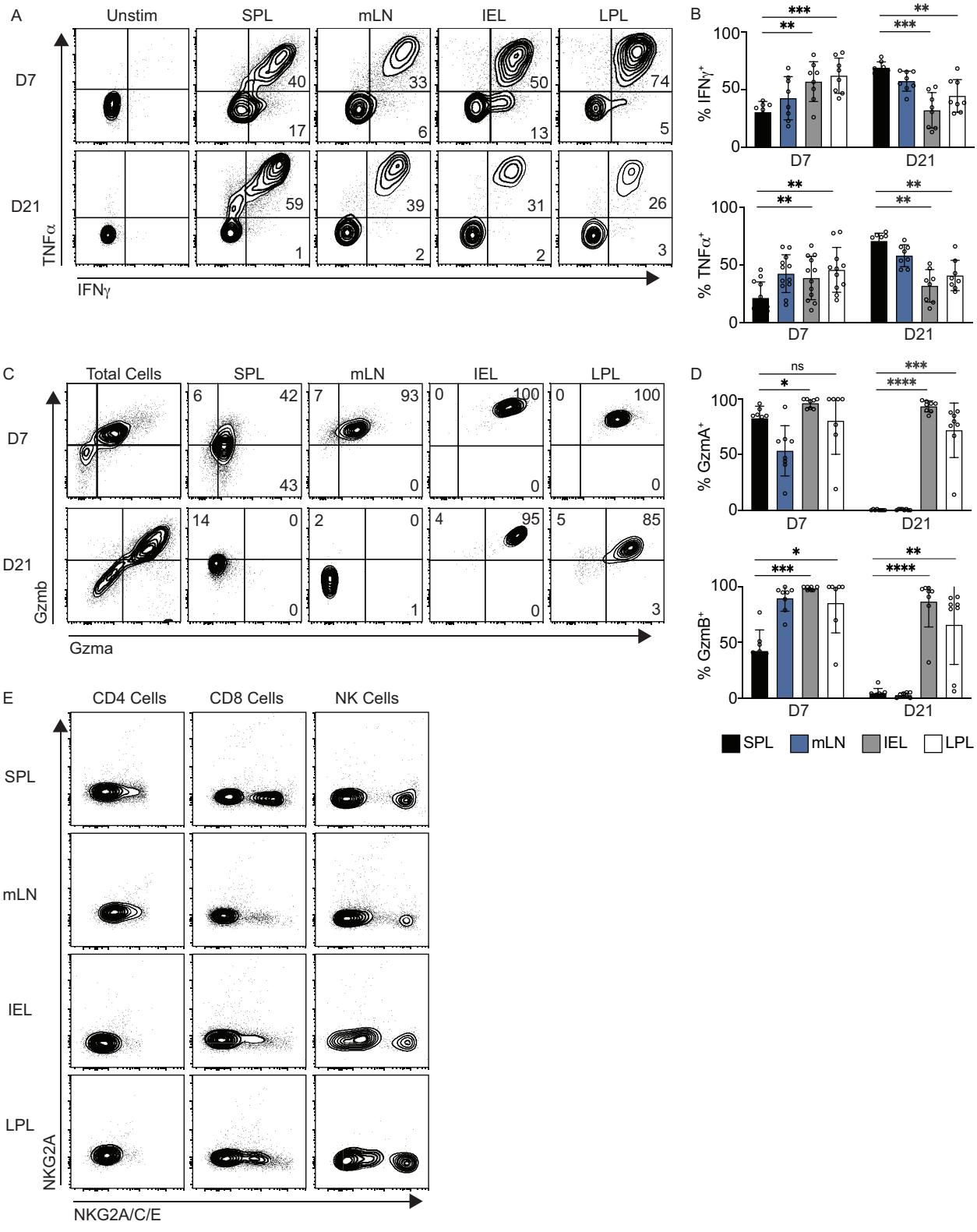


Figure 2.2 SI-homing cells from the spleen and mLNs are T_H1 effector cells. (A-B) CCR9 and CD49d expression by $CD4^+$ T cells in specified tissues on days 6-8 of LCMV-Arm infection. Representative flow cytometry plots (A) and quantification (B) of $CCR9^+CC49d^+$ SMARTA $CD4^+$ T cells. $CCR9^+CC49d^+$ quadrants are highlighted in red **(C-D)** Expression of SLAM and CXCR5 by $CCR9^+CC49d^+$ SMARTA $CD4^+$ T cells (red) in the spleen or mLN on days 6-8 of LCMV-Arm infection. Representative flow cytometry plots (C) and quantification (D) of T_H1 ($SLAM^+CXCR5^-$) and T_{FH} ($SLAM^-CXCR5^+$) SMARTA cells. Data are representative (A,C) or cumulative (B,D) of 2 experiments with n=2-4 mice per experiment. Graphs show mean ± SD; *$p < 0.05$, ** $p<0.01$, *** $p<0.001$, ****$p< 0.0001$.

Figure 2.3 $CD4^+$ T_{RM} in IEL and LPL produce T_H1-associated effector cytokines. (A,B) IFNγ and TNFα expression by $CD4^+$ T cells on indicated day of infection following *ex vivo* GP_{61} peptide stimulation in specified tissues. Representative flow cytometry plots (A) and quantification (B) of $IFN\gamma^+$ and $TNF\alpha^+$ SMARTA $CD4^+$ T cells. **(C,D)** Granzyme A (GzmA) and B (GzmB) expression by SMARTA $CD4^+$ T cells in specified tissues on indicated day of infection. Representative flow cytometry plots (C) and quantification (D) of $Gzma^+$ and $Gzmb^+$ SMARTA $CD4^+$ T cells. **(E)** NKG2A and NKG2A/C/E expression by specified tissue on day 21 of infection. Numbers in flow plots indicate percent of cells in corresponding gate. Data are representative (A,C,E) or cumulative (B,D) of 2 experiments with n=2-4 mice per experiment. Graphs show mean ± SD; *$p < 0.05$, ** $p<0.01$, *** $p<0.001$, ****$p< 0.0001$.


A
Unstim
SPL
mLN
IEL
LPL
D7
D21
TNFα
IFNγ
40
17
33
6
50
13
74
5
59
1
39
2
31
2
26
3
B
% IFNγ+
% TNFα+
D7
D21
C
Total Cells
Gzmb
Gzma
6
42
43
7
93
0
100
0
14
0
2
1
4
95
5
85
3
D
% GzmA+
% GzmB+
ns
E
CD4 Cells
CD8 Cells
NK Cells
SPL
mLN
IEL
LPL
NKG2A
NKG2A/C/E

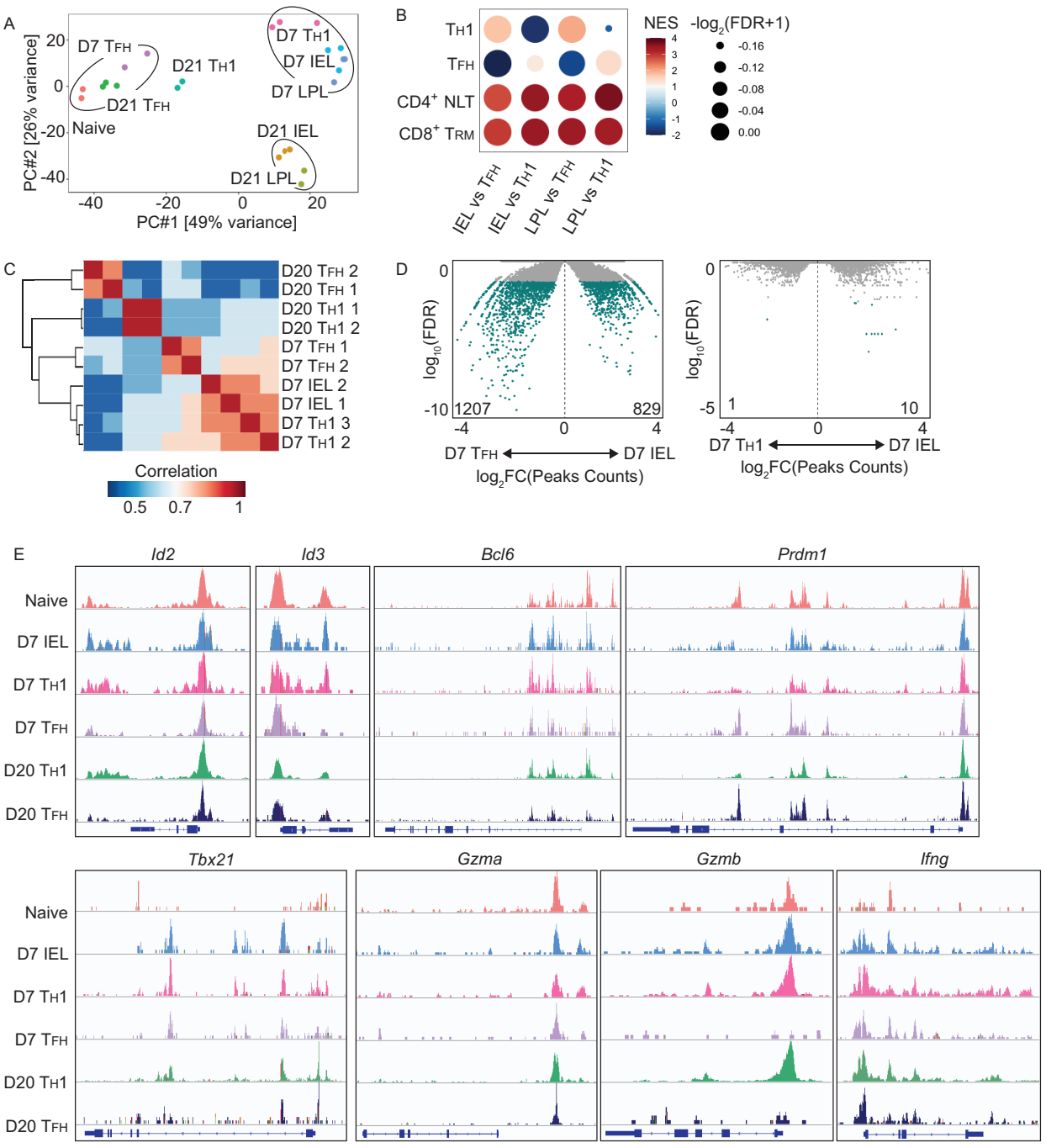
$CD4^+$ T_{RM} share a transcriptional and epigenetic profile with effector T_H1 cells

To investigate the transcriptional profile of SI $CD4^+$ T_{RM} cells and how they compare to effector T_H subsets, SMARTA $CD4^+$ T cells or $SLAM^+$ T_H1 cells and $CXCR5^+$ T_{FH} SMARTA cells were sorted from the SI or spleen, respectively, on day 7 or 21 of infection, and analyzed by bulk RNA sequencing. Principal component analysis (PCA) of gene expression revealed a clear separation of D21 IEL/LPL samples from D7 samples, T_{FH} samples, and naive cells (Figure 2.4A). Of note, day 7 IEL and LPL $CD4^+$ cells aggregated closely with day 7 T_H1 cells, consistent with the phenotypic similarities between $CD4^+$ T_{RM} and T_H1 cells we observed. We also observed some separation between D21 IEL and LPL samples, which may explain the heterogeneity of the overall SI population driven by tissue localization or functional requirements. Gene enrichment analysis (GSEA) was used to compare the transcriptional profiles of $CD4^+$ T cells from the SI and spleen on day 21 of infection (Figure 2.4B). Consistent with the T_H1-like qualities of $CD4^+$ T_{RM}, SI cells were highly enriched for the T_H1 signature[31] when compared to T_{FH} cells, and showed minimal enrichment of the T_{FH} signature[31]. Furthermore, $CD4^+$ T_{RM} were enriched for the published $CD8^+$ T_{RM}[119] and $CD4^+$ non-lymphoid tissue (NLT) signatures[71] when compared to both T_H1 or T_{FH} splenic subsets, regardless of the subset (Figure 2.4B).

Next, we performed the Assay for Transposase-Accessible Chromatin using sequencing (ATACseq) on $CD4^+$ T cells paired with bulk RNAseq timepoints to understand the epigenetic landscape of splenic and SI $CD4^+$ T cells and assess accessible regulatory regions in these populations. A correlation heatmap comparing differentially accessible regions among all samples revealed that D7 IEL cells were more closely associated with the T_H1 subset at both effector and memory time points (Figure 2.4C). We directly quantified peaks in differentially accessible regions (DARs) for D7 IEL, T_{FH}, or T_H1 cells, we found few peaks distinguishing IEL and T_H1

samples compared to 2,036 differentially accessible peaks between the IEL and T_{FH} subset (Figure 2.4D). When we compared differentially accessible peaks at specific gene loci among different timepoints and tissues, we observed similar profiles for D7 IEL and D7 T_H1 cells which were distinct from naive and D7 T_{FH} profiles (Figure 2.4E). Overall, these data revealed similarities between $CD4^+$ T_{RM} and effector T_H1 cells in surface-receptor phenotype, transcriptional profile, and chromatin landscape, suggesting that these two populations are related.

Figure 2.4 Transcriptional and epigenetic profile of CD4$^+$ T_{RM} in viral infection. (A) Principal component analysis (PCA) of bulk RNAseq of circulating and resident SMARTA CD4$^+$ T cells from spleen and SI, harvested on day 7 or 21 of LCMV-Arm infection. **(B)** Gene enrichment analysis (GSEA) of day 21 RNAseq data. **(C-E)** ATACseq of circulating and resident SMARTA CD4$^+$ T cells from spleen and SI on day 7 or spleen on day 21 of infection. **(C)** Pearson correlation for peaks in differentially accessible regions. **(D)** Volcano plots comparing peak counts between D7 IEL and T_H1 or T_{FH} subsets. Numbers in volcano plots indicate number of differentially accessible regions in IEL compared to either T_H1 or T_{FH} samples. (**E**) Genome browser tracks depict ATACseq chromatin accessibility across samples for indicated gene. Data are cumulative of 3 experiments (A,B) or 2 experiments (C-E) with n=4-5 mice per experiment for day 7 and n=12-15 mice per experiment for day 21.


A
PC#2 [26% variance]
PC#1 [49% variance]
D7 TFH
D21 TH1
D7 TH1
D7 IEL
D7 LPL
D21 TFH
Naive
D21 IEL
D21 LPL
B
TH1
TFH
CD4+ NLT
CD8+ TRM
NES
-log2(FDR+1)
IEL vs TFH
IEL vs TH1
LPL vs TFH
LPL vs TH1
C
D20 TFH 2
D20 TFH 1
D20 TH1 1
D20 TH1 2
D7 TFH 1
D7 TFH 2
D7 IEL 2
D7 IEL 1
D7 TH1 3
D7 TH1 2
Correlation
D
log10(FDR)
1207
829
D7 TFH
D7 IEL
log2FC(Peaks Counts)
1
10
D7 TH1
E
Id2
Id3
Bcl6
Prdm1
Naive
D7 IEL
D7 TH1
D7 TFH
D20 TH1
D20 TFH
Tbx21
Gzma
Gzmb
Ifng

Progressive acquisition of CD4$^+$ T_{RM} signature

Effector CD8$^+$ cells in the SI at day 7 of infection transcriptionally resembled CD8$^+$ T_{RM} cells at day 35 of infection, with "core" T_{RM} genes shared between the two time points[119]. However, SI CD4$^+$ cells at day 7 compared to SI CD4$^+$ cells at day 21 were transcriptionally distinct with D7 SI cells grouping very closely with D7 T_H1 cells (Figure 2.4A). Further analysis revealed that only 13-15% of the genes differentially expressed between mature D21 T_{RM} cells compared to splenic T_H1 or T_{FH} subsets were also differentially expressed between D7 T_{RM} cells compared to effector splenic cells (Figure 2.5A). We also observed changes in phenotype between the two time points, such as the loss of Ly6C expression by T_{RM} cells at a memory time point (Figure 2.1D).

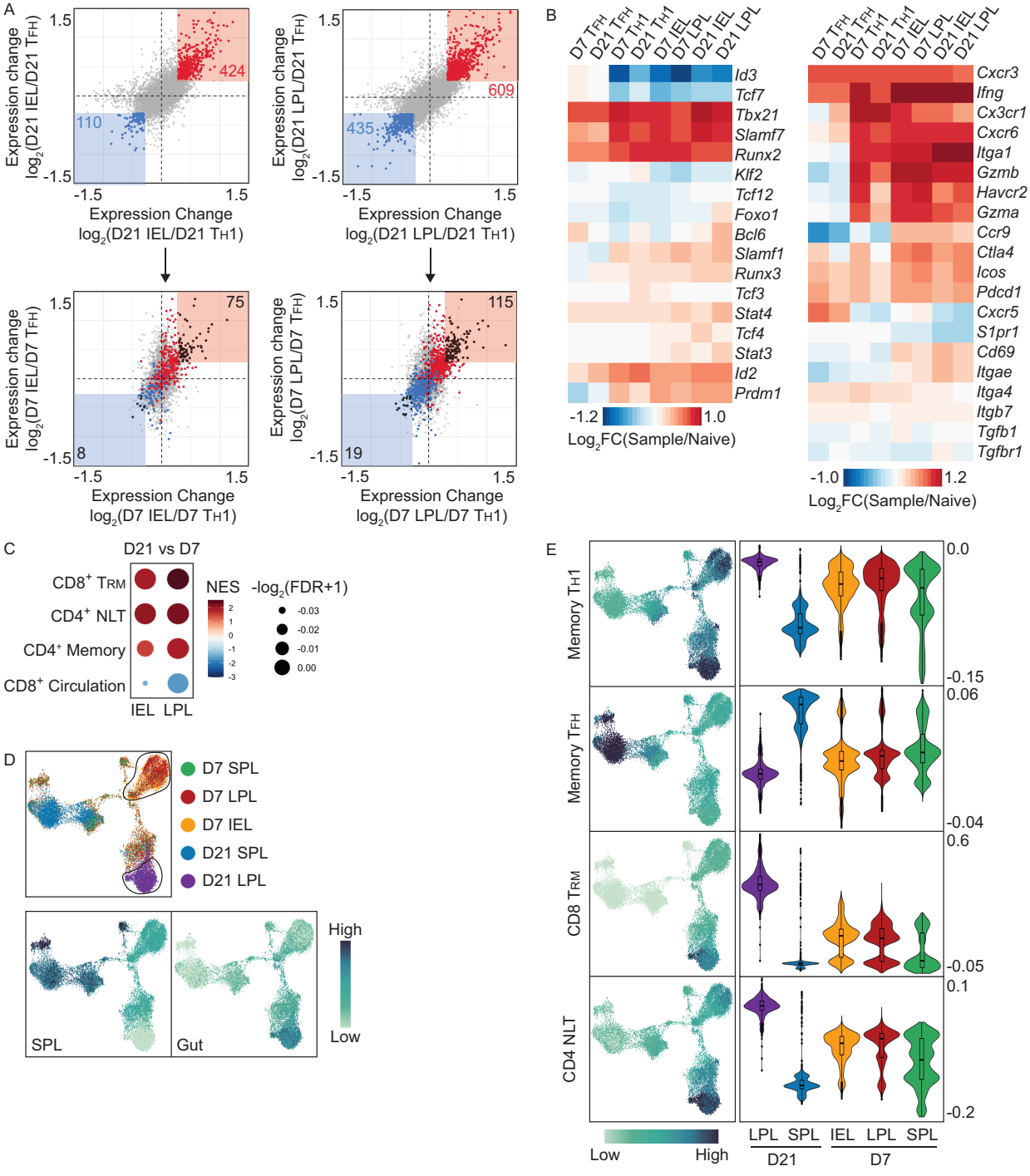
To further explore changes in T_{RM} between effector and memory time points, we examined the expression of specific genes encoding for cell-surface receptors, effector molecules, and TFs with known roles in effector and memory CD4$^+$ T cell development or the CD8$^+$ T_{RM} program by splenic and SI CD4$^+$ T cells on days 7 and 21 of infection (Figure 2.5B). When compared to naive CD4$^+$ T cells, IEL and LPL samples at both effector and memory timepoints were enriched for expression of genes associated with SI-homing and the T_{RM} program such as *Ccr9* and *Cd69*. Although there were low levels of CD103 protein on day 21 IEL and LPL cells (Figure 2.1D), we observed higher expression of *Itgae* gene in SI samples compared to splenic subsets. However, unlike with CD8$^+$ T_{RM}, there was no enrichment of *Tgfb* or *Tgfbr1*, which induces *Itgae* expression, suggesting that the regulation and function of CD103 protein were different between CD4$^+$ and CD8$^+$ T_{RM}. CD4$^+$ T_{RM} cells from IEL and LPL at day 21 expressed genes associated with both the effector/T_H1 program, such as *Gzma*, *Gzmb*, *Ifng*, *Tbx21*, *Prdm1*, *Id2*, *Stat4* and the memory/T_{FH} identity such as *Icos*, *Pdcd1*, *Stat3, Bcl6*. Additionally, D20 IEL, LPL, and T_H1 populations expressed both *Bcl6* and *Prdm1*, which was surprising given the defined reciprocal antagonism of

these two TFs in circulating T_H cell differentiation. GSEA of IEL and LPL cells on day 21 showed enrichment of the $CD8^+$ T_{RM}, $CD4^+$ NLT signature, and $CD4^+$ T cell memory signatures[31] compared to SI cells on day 7 of infection (Figure 2.4C). Together, these data indicate that SI $CD4^+$ cells at day 7 are still progressing towards the T_{RM} phenotype, and this effector population may require further regulation to become mature T_{RM}.

To explore differences in the transcriptional programs between effector and memory SI populations and to evaluate heterogeneity among cells within the populations, we utilized scRNAseq to compare gene expression of SMARTA $CD4^+$ T cells from the spleen, IEL, or LPL of LCMV-Arm infected mice on days 7 and 21 following infection. Due to the cell number requirements for high-quality scRNAseq results and the lower numbers of SMARTA cells in the IEL compartment at D21, we were only able to isolate and analyze D21 LPL sample. Unsupervised Uniform Manifold Approximation and Projection (UMAP) clustering separated samples by tissues, with D21 LPL cells (purple), and D21 SPL cells (blue) grouped separately from other samples and enriched for their respective gut or spleen signature (Figure 2.5D). While a subset of D7 IEL and LPL cells (yellow, red) clustered together, a high proportion of D7 SI cells were mixed with D7 SPL cells (green) and were diffused throughout the UMAP. This was consistent with our observations that SI cells at day 7 were not fully mature T_{RM}, and these effector cells did not yet express the gut signature. Next, we assigned a module score to each cell for the specified gene signatures generated from our bulk RNAseq data or published datasets (Figure 2.5E). The memory T_H1 signature (defined by fold-change >1.5 between D7 and D41 T_H1 cells) was enriched in both the D7 and D21 SI clusters compared to splenic cells, although the D7 spleen replicates did show variable expression of the T_H1 signature. The memory T_{FH} gene signature (defined by fold-change >1.5 between D7 and D41 T_{FH} cells) was only enriched in D21 SPL samples, highlighting the T_H1

versus T_{FH} division within $CD4^+$ splenocytes. The D21 LPL cluster showed highest enrichment of the $CD4^+$ NLT and $CD8^+$ T_{RM} signature, although the D7 SI samples were also moderately enriched for residency-related genes compared to splenic samples (Figure 2.4E). These data reveal differences between effector and memory SI $CD4^+$ T cells, with D7 SI cells sharing transcriptional similarities with D7 splenic cells instead of being distinct tissue-resident population.

Figure 2.5 Effector SI CD4$^+$ T cells in viral infection progress towards a mature T_{RM} program. (A) Top, comparison of gene expression of IEL (left) and LPL (right) T_{RM} cells relative to T_H1 and T_{FH} subsets on day 21 of LCMV infection. Red denotes genes increased in T_{RM} relative to T_H1 and T_{FH} cells; blue denotes genes increased in T_H1 and T_{FH} relative to T_{RM} cells. Bottom, comparison of differentially expressed genes in mature T_{RM} cells (from top panel) in cells from IEL or LPL on day 7 of infection. Black denotes genes from top panel (either blue or red) which are differentially expressed by day 7 SI cells compared to day 7 splenic subsets. **(B)** Heatmap showing gene expression of transcriptional regulators (left) or cell-surface receptors/cytokines (right). Values are calculated as a $\log_2$fold change between each sample and naive population. **(C)** Gene enrichment analysis (GSEA) of day 7 and 21 SI cells. **(D-E)** Single-cell RNAseq of circulating and resident SMARTA CD4$^+$ T cells from spleen and SI, harvested on day 7 or 21 post-infection. **(D)** UMAP dimensional reduction colored by tissues (top) and enrichment of indicated gene signatures (bottom). (**E)** UMAP reduction and violin plots showing enrichment of indicated gene signatures. Data are cumulative of 3 experiments (A-C) or 2 experiments (D-E) with n=4-5 mice per experiment for day 7 and n=12-15 mice per experiment for day 21.


A
Expression change log2(D21 IEL/D21 TFH)
Expression Change log2(D21 IEL/D21 TH1)
424
110
Expression change log2(D21 LPL/D21 TFH)
Expression Change log2(D21 LPL/D21 TH1)
609
435
Expression change log2(D7 IEL/D7 TFH)
Expression Change log2(D7 IEL/D7 TH1)
75
8
Expression change log2(D7 LPL/D7 TFH)
Expression Change log2(D7 LPL/D7 TH1)
115
19
B
D7 TFH
D21 TFH
D7 TH1
D21 TH1
D7 IEL
D7 LPL
D21 IEL
D21 LPL
Id3
Tcf7
Tbx21
Slamf7
Runx2
Klf2
Tcf12
Foxo1
Bcl6
Slamf1
Runx3
Tcf3
Stat4
Tcf4
Stat3
Id2
Prdm1
-1.2
1.0
Log2FC(Sample/Naive)
Cxcr3
Ifng
Cx3cr1
Cxcr6
Itga1
Gzmb
Havcr2
Gzma
Ccr9
Ctla4
Icos
Pdcd1
Cxcr5
S1pr1
Cd69
Itgae
Itga4
Itgb7
Tgfb1
Tgfbr1
-1.0
1.2
Log2FC(Sample/Naive)
C
D21 vs D7
CD8+ TRM
CD4+ NLT
CD4+ Memory
CD8+ Circulation
IEL
LPL
NES
-log2(FDR+1)
D
D7 SPL
D7 LPL
D7 IEL
D21 SPL
D21 LPL
SPL
Gut
High
Low
E
Memory TH1
Memory TFH
CD8 TRM
CD4 NLT
Low
High
LPL
SPL
IEL
LPL
SPL
D21
D7

CD4$^+$ T_{RM} are heterogeneous and express genes associated with both effector and memory fates

Single-cell RNAseq analysis combining datasets from both D7 and D21 time points may mask T_{RM} population heterogeneity within or between tissues. Therefore, we next performed a focused analysis of D7 (Figure 2.6) or D21 samples (Figure 2.7) to examine the heterogeneity within each timepoint and the differences between effector and memory phases.

Analysis of only spleen and SI cells at day 7 revealed 8 clusters, with D7 IEL and LPL cells mainly aggregated in clusters 0, 1, and 4 while splenic cells grouped in clusters 2, 3, 5, 6, and 7 (Figure 2.6A). All clusters had a mixture of cells from both spleen and SI, indicating that IEL and LPL cells at day 7 were not yet fully distinct from circulating cells which is consistent with our bulk RNAseq results (Figure 2.4A). The clusters were separated into T_H1 and T_{FH} subsets, with clusters 0, 1, 2, 4, 6, and 7 highly expressing effector T_H1 signature and the associated genes, Slamf1 and Id2, while clusters 3 and 5 were enriched for the effector T_{FH} signature and the associated genes, *Cxcr5* and *Id3* (Figure 2.6B). The CD4$^+$ NLT signature was also enriched in clusters 0, 1, and 4, with cluster 1 having the highest expression of tissue-associated genes. Cells in cluster 1 also showed the highest enrichment for genes upregulated in D20 SI cells compared to D7 cells, such as *Cd69* and *Itga1*. Meanwhile, cells in cluster 0 and 4 had higher expression of genes downregulated in D20 SI cells, suggesting that cluster 1 cells have more of the SI T_{RM} programming than remaining SI cells (Figure 2.6C). This supported our hypothesis that day 7 SI cells were not mature T_{RM} cells and might require specific tissue or memory cues for their differentiation to T_{RM}.

Reanalysis focusing on D21 spleen and LPL cells revealed 4 distinct clusters, with the LPL sample divided into clusters 0 and 1 and the spleen sample divided into clusters 2 and 3 (Figure

2.7A). The LPL clusters were enriched for the $CD4^+$ NLT[71], $CD8^+$ T_{RM}[119], and memory T_H1 gene signatures, with cluster 1 having the highest expression of the tissue-associated genes (Figure 2.7B). Cells in cluster 2 expressed the highest level of the memory T_{FH} signature, with moderate expression by cluster 3, compared with LPL cells. Among genes differentially expressed between the SI and splenic clusters, we were specifically interested in those closely associated with T cell memory and tissue-residency. We focused on transcriptional regulators involved in the effector versus memory programs, specifically *Bcl6* and *Stat3* for memory cells and *Id2* and *Prdm1* for T_H1 effector cells[11] (Figure 2.7C). All four genes showed elevated expression by the LPL compared to splenic clusters. The levels of expression of these factors varied between LPL clusters and even within cluster 1, indicating a level of heterogeneity within this population. We also examined expression of *Runx3* and *Klf2*, two TFs which regulate the generation of $CD8^+$ T_{RM}[97,119]. LPL cells were enriched for expression of *Runx3* and expressed low levels of *Klf2* compared to the spleen, with accompanying downregulation of its target *S1pr1* (Figure 2.7C). This suggests that similar mechanisms of tissue retention are shared by $CD4^+$ and $CD8^+$ T_{RM}. We also observed variation in the expression of *Ccr9, Cd69, and P2rx7*, which have been shown to be involved in migration to and maintenance of T cells in tissues, specifically the intestine[116,120] (Figure 2.7C).

It was unexpected for us to observe that both Blimp1 and Bcl6 were expressed by T_{RM} samples at levels higher than spleen populations in bulk RNAseq analysis. Therefore, we used scRNAseq to determine if the expression of these genes was from different subpopulations or if co-expression was in the same cells, and found overlapping expression. To determine whether any LPL cells co-expressed TFs associated with both effector and memory fates, we compared expression of either *Bcl6* and *Prdm1* or *Stat3* and *Stat4* in single cells (Figure 2.7D). Indeed, a small number of cells did co-express *Bcl6* and *Prdm1* mRNA, and the majority of cells in cluster

3 expressed both *Stat3* and *Stat4*, another pair of TFs with opposing roles in effector and memory T cell programs, revealing the intriguing possibility that $CD4^+$ T_{RM} do have a hybrid transcriptional program. We also transferred Blimp1-YFP SMARTA $CD4^+$ T cells to congenically distinct mice that were subsequently infected with LCMV-Arm, then sort-purified $Blimp1^+$ and $Blimp1^-$ SMARTA cells from the spleen and SI at day 14 of infection. Using qPCR, we observed a two to three-fold increase in *Bcl6* mRNA in $Blimp1^+$ and $Blimp1^-$ IEL and $Blimp1^+$ LPL cells compared to splenic populations, confirming the surprising co-expression of *Blimp1* and *Bcl6* (Figure 2.7E). Altogether, these data highlighted the heterogeneity within the SI $CD4^+$ T cell population starting at the effector phase, and implicated the TFs Blimp1, Id2, and Bcl6 as possible regulators of a dual effector-memory residency program.

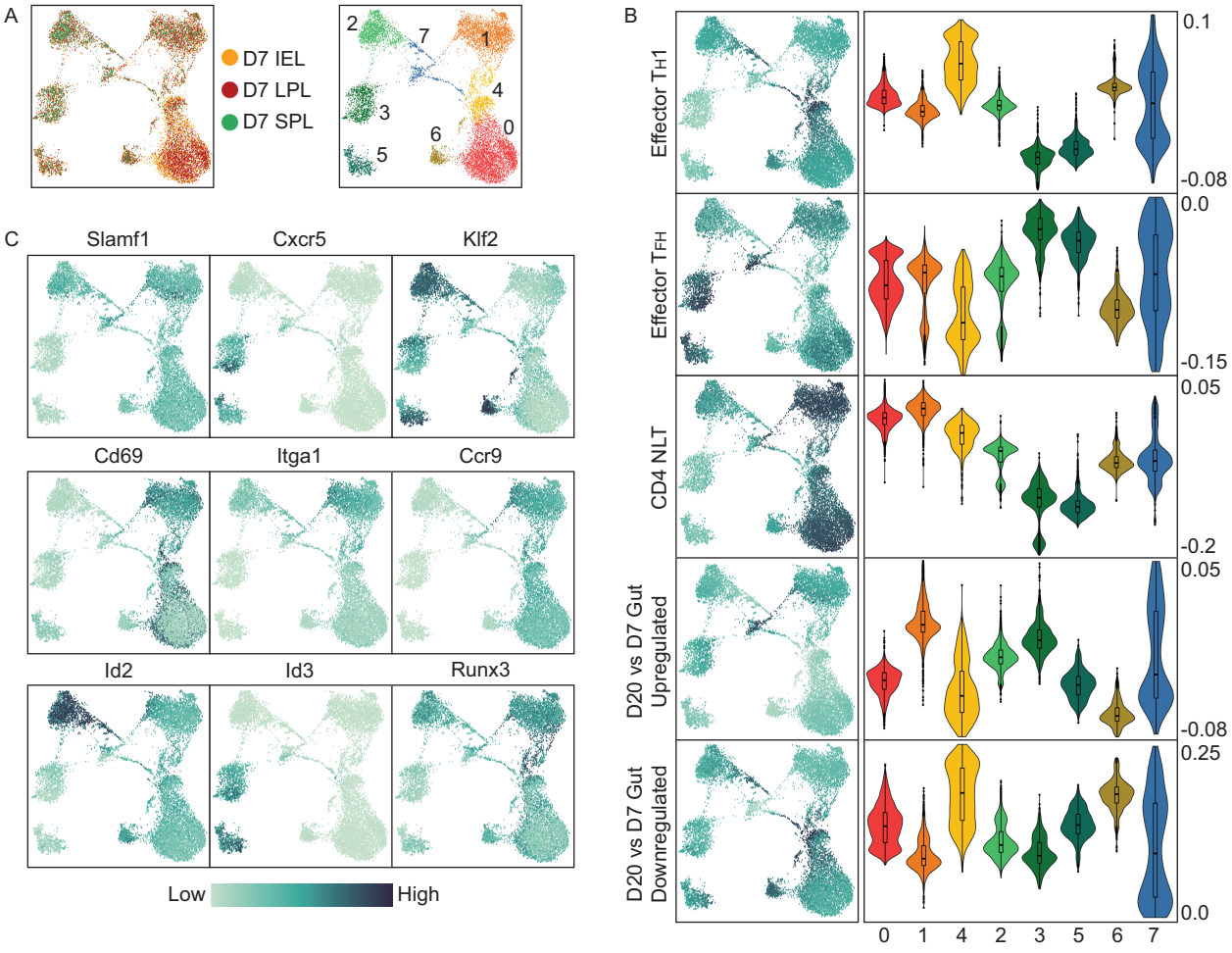


Figure 2.6 CD4$^+$ T_{RM} cells at day 7 of infection are heterogeneous and enriched for effector genes. (A) UMAP dimensional reduction of spleen and LPL CD4$^+$ T cells colored by tissue identity (left) or cluster (right). **(B)** UMAP reduction and violin plots showing enrichment of indicated gene signatures. **(C)** UMAP reduction showing relative expression of indicated transcriptional regulators or cell-surface receptors and cytokines in UMAP reduction (left) or violin plots by clusters (right).

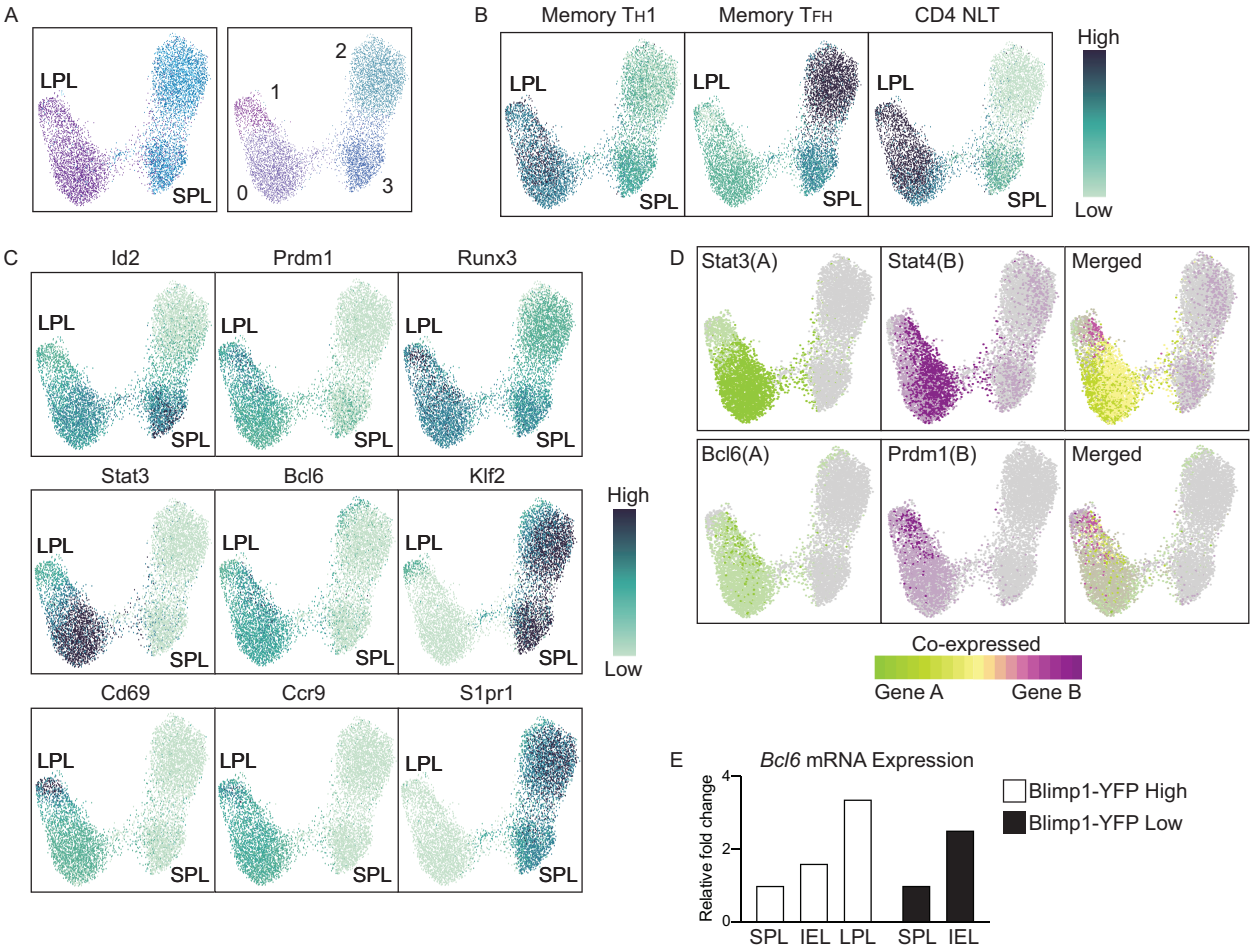


Figure 2.7 CD4$^+$ T_{RM} cells at day 21 exhibit heterogeneity and express genes associated with both effector and memory fates. (A) UMAP dimensional reduction of spleen and LPL CD4$^+$ T cells colored by tissue identity (left) or cluster (right). **(B)** UMAP reduction and violin plots showing enrichment of indicated gene signatures. **(C)** UMAP reduction showing relative expression of indicated transcriptional regulators or cell-surface receptors and cytokines in UMAP reduction (left) or violin plots by clusters (right). **(D)** UMAP dimensional reduction showing merged expression of indicated TF pairs. **(E)** Bar plot showing relative fold-change of *Bcl6* mRNA in indicated tissue compared to spleen in Blimp1$^+$ versus Blimp1$^-$ SMARTA cells from mice on day 15 of LCMV-Arm infection. Data are cumulative of 1-2 experiments with n=12-15 mice per experiment.

2.3 Discussion

The data presented in this chapter revealed a shared phenotype and transcriptional profile between $CD4^+$ T_{RM} and the T_H1 subset, which is considered more effector-like compared to T_{FH} cells. We found that SI-homing $CD4^+$ T cells from the spleen and the T_{RM} cells were predominantly $SLAM^+$ T_H1 cells, consistent with studies in other infection models and tissues where T_{RM} cells reflect the T_H effector program associated with the acute pathogen/antigen exposure[36,54,93]. At barrier surfaces, T_{RM} cells are among the first to encounter antigen and are required to rapidly coordinate a local recall response[71], which is consistent with enhanced effector traits compared to circulating memory populations to effectively respond to reinfection. Nevertheless, the relationship between T_H1 and T_{RM} is somewhat unexpected given that fate-committed T_{FH} cells are the main proportion of the circulating $CD4^+$ memory T cell population[25]. T_{FH} cells have a transcriptional program that has been associated with memory and long-term survival, and $CD4^+$ T_{RM} may require aspects of the T_{FH} program for their maintenance or share common progenitors with T_{FH} cells.

While $CD4^+$ T_{RM} populations were enriched for the core $CD8^+$ T_{RM} signature, there were clear differences between the two lineages. For example, while SI $CD8^+$ T_{RM} cells expressed both CD69 and CD103, $CD4^+$ T_{RM} cells had minimal levels of CD103. Additionally, while effector $CD8^+$ T cells in tissues are fully mature T_{RM}, effector SI $CD4^+$ T cells at day 7 are distinct from the memory SI cells and gradually gain the T_{RM} program as the immune response progresses. This slower development of the $CD4^+$ T_{RM} identity may be due to the varied T_H subsets in the effector stage and the resulting heterogenous memory population.

Work on circulating memory $CD4^+$ T cells identified distinct subsets with varying levels of CD27, where high expression is associated with memory and recall potential, and Ly6C, where

low expression is tied to increased longevity and plasticity. In a viral infection, we found that $CD4^+$ T_{RM} downregulated both CD27 and Ly6C compared to circulating populations, making the "memory" characteristics of these cells unresolved. Nevertheless, half of the population retained CD27 expression, especially within the LPL compartment, suggesting that LPL cells may be the more "memory-like" population compared to IEL cells which are directly at the barrier surface and constantly exposed to antigen from the microbiota. The variation in expression of cell-surface molecules within and between the IEL and LPL samples is also reflected in the transcriptional heterogeneity within our sequencing data.

In this chapter, we have demonstrated that $CD4^+$ T_{RM} in viral infection are a heterogeneous population which co-express genes important for both effector and memory T cell lineage differentiation. Future work will focus on identifying possible T_{RM} precursor cells within early T_H1 effector cells using adoptive transfers of T_H1 and T_{FH} cells and computational methods such as trajectory mapping. We also aim to identify key transcriptional regulators of $CD4^+$ T_{RM} differentiation, starting with known regulators in T_H1 versus T_{FH} differentiation and effector versus memory fate determination.

Chapter 2 was adapted from a manuscript which will be submitted for publication. **Nguyen QP**, Deng TZ, O'Shea SM, Pipkin ME, Choi J, Crotty S, Goldrath AW. (2021). *Blimp1, Id2, and Bcl6 balance effector and memory-associated programs to promote $CD4^+$ T_{RM} differentiation following viral infection.* The dissertation author was the primary author of this paper.

Chapter 3 Blimp1, Id2, and Bcl6 balance effector and memory-associated programs to promote $CD4^+$ T_{RM} differentiation following viral infection

3.1 Introduction

The expression and activity of transcriptional regulators critical in circulating $CD4^+$ T cells differentiation are not well understood in the T_{RM} population. Bcl6 and Blimp1 both have established roles in $CD4^+$ T cell fate determination, with Bcl6 promoting T_{FH} and memory $CD4^+$ populations, while Blimp1 directly inhibits T_{FH} development[30,103]. Bcl6-deficient $CD4^+$ T cells following influenza infection did not differentiate into lung T_{RH} cells at a memory time point, although loss of Bcl6 in effector $CD4^+$ T cells following house dust mite exposure resulted in more $CD4^+$ T cells in the lung, highlighting potential involvement of Bcl6[36,79,80]. Blimp1 has also been shown to be a central regulator of the $CD8^+$ tissue-resident program and may play a similar key role in $CD4^+$ resident populations[110]. Notably, Blimp and Bcl6 are both co-expressed by regulatory follicular helper cells (T_{FR}), a population that also displays hybrid functions of T_{reg} and T_{FH}[121,122].

Similarly, E protein TF and their inhibitors, Id proteins, also play critical regulatory functions in $CD4^+$ T cell differentiation[123,124]. Id2 is highly expressed by $CD4^+$ T_H1 cells and is essential for maintaining T_H1 identity, partly through inhibiting the T_{FH} transcriptional program[125]. Thus, Id2 is likely responsible for directing the generation of effector populations in $CD4^+$ T cells. Conversely, current work suggests that Id3, a known Id protein which drives T_{FH} differentiation and responsible for memory $CD8^+$ T cell differentiation[126], can mark memory-precursor $CD4^+$ T cells. In an acute LCMV infection, $Id3^{hi}$ $CD4^+$ T cells in the spleen survived following the contraction phase and exhibited multipotent potential upon rechallenge, indicating that Id3 marks

long-lived memory $CD4^+$ T cells in circulation (Shaw, under review). Despite these TFs being vital for the formation of circulating $CD4^+$ T cell subsets, it is unclear how these transcriptional regulators influence the T_{RM} population during the course of viral infection.

In this chapter, through studies of Blimp1, Id2, Id3, and Bcl6 function in antiviral $CD4^+$ T_{RM} populations, I demonstrated that Blimp1 and Id2 were both required for early T_{RM} formation, while Bcl6 initially suppressed T_{RM} formation but was required as the infection resolves for maintenance of long-lived T_{RM} cells.

3.2 Results

Blimp1 and Id2 are required for virus specific $CD4^+$ T_{RM} development

To better understand the relationship between $CD4^+$ T_{RM} and T_H1 population, we investigated the role of two transcriptional regulators of T_H1 development and effector function, Blimp1 and Id2[103,125]. Our bulk and scRNAseq data showed that D21 SI T_{RM} expressed both factors, suggesting that Blimp1 and Id2 may function in $CD4^+$ T_{RM} differentiation. We also utilized the PageRank algorithm as previously described to identify important TFs in our day 7 SI T_{RM} cells[89,119]. PageRank analysis of bulk RNAseq and ATACseq data highlighted *Prdm1* as a gene with a higher PageRank score and gene expression in SI samples relative to splenic cells at day 7 of infection. Additionally, PageRank also identified *Tcf12*, which encodes for the TF HEB and is a target of Id2 protein, as a gene with higher score and gene expression in the SI. These data suggest that Blimp1 and Id2, T_H1-associated factors, may be critical for early T_{RM} development (Figure 3.1A).

First, we examined the expression of Blimp1, Id2, and Id3 by $CD4^+$ T cells following acute viral infection. To assess Blimp1 expression in circulating and SI $CD4^+$ T cells following viral infection, we transferred Blimp1-YFP SMARTA $CD4^+$ T cells to congenically distinct mice that

were subsequently infected with LCMV-Arm. On day 7 and 21 of infection, SMARTA $CD4^+$ T cells were isolated from the spleen, mLN, and SI (Figure 3.1B,C). On day 7 of infection, the majority of SI $CD4^+$ T cells expressed Blimp1 compared to only 40-50% of $Blimp1^+$ $CD4^+$ T cells in the circulation. By day 21 of infection, the frequency of $Blimp1^+$ $CD4^+$ T cells in the SI dropped to ~60% but was still higher than in circulating $CD4^+$ T cell population. We also transferred SMARTA $CD4^+$ T cells expressing Id2-YFP and Id3-GFP reporter alleles to naive recipients that were subsequently infected with LCMV. We found that IEL and LPL $CD4^+$ T cells expressed only Id2 at effector and memory time points while the majority of circulating $CD4^+$ T cells expressed both Id2 and Id3 (Figure 3.1D,E). Recent work showed that Blimp1 and Id3 expression can identify functionally and transcriptionally distinct subsets in $CD8^+$ T_{RM} and tumor-infiltrating lymphocytes (TIL)[127]. We investigated whether similar subsets can be identified in SI $CD4^+$ T_{RM} using a mouse model expressing both Blimp1-YFP and Id3-GFP reporter alleles (Figure 3.1 F,G). On day 7 and 21 of infection, we observed distinct $Blimp1^+$ and $Id3^+$ populations in the spleen and mLN, while SI cells were mainly $Blimp1^+$. These results were consistent with our bulk RNAseq data in which the T_H1 subset and SI cells expressed *Id2* and *Prdm1* mRNA while only the T_{FH} populations expressed the *Id3* mRNA (Figure 2.5B).

To investigate the requirement for Id2 in SI $CD4^+$ T cells, we transferred a 50:50 mix of WT and $Id2^{f/f}$ CD4-Cre^+ (Id2-KO) SMARTA cells into congenically distinct hosts and infected the mice with LCMV-Arm. On days 7, 14, and 21 of infection, we isolated the transferred cells, and again, compared the frequency and numbers of WT and KO cells. On day 7, Id2-deficient $CD4^+$ T cells accumulated at a reduced frequency in the SI compared to WT cells, similar to that observed in the spleen and mLN (Figure 3.2A,B). By day 14 and 21 of infection, Id2-KO $CD4^+$ T cells were even further reduced in the SI compared to spleen and mLN. Consistent with the role

of Id2 in T_H1 differentiation, splenic Id2-KO cells did not express SLAM and were predominantly T_{FH} (Figure 3.2C). The reduction in $CD4^+$ T cells lacking Id2 was also reflected in the cell numbers at both effector and memory time points (Figure 3.2D).

To determine if Blimp1 contributes to regulating $CD4^+$ T_{RM} differentiation, we transferred $Blimp1^{f/f}CD4\text{-}Cre^+$ (Blimp1-KO) SMARTA T cells as a 50:50 mix with wild-type (WT) SMARTA T cells into congenically distinct hosts and infected the mice with LCMV-Arm. On days 7, 14, and 21 following infection, we isolated the transferred cells and compared the frequency and number of WT to Blimp1-KO cells. Loss of Blimp1 significantly impaired the accumulation of the $CD4^+$ T_{RM} population starting from the day 7 effector stage, with Blimp1-KO $CD4^+$ T cells making up ~10-15 % of the total SMARTA population isolated from the SI (Figure 3.2E,F). There was no evidence of delayed T_{RM} differentiation as the Blimp1-KO SMARTA $CD4^+$ T cells remained at a dramatically reduced proportion compared to WT cells at the memory time points. In spite of an absence of $SLAM^+$ T_H1 cells at day 7, there was a similar accumulation of cells in the spleen (Figure 3.2F). However, starting at day 14 of infection, there was an accompanying decrease in Blimp1-KO cells compared to WT cells in the spleen and mLN, although the deficit of Blimp1-KO $CD4^+$ T cells in circulation did not reach the levels seen in the SI. The difference between WT and Blimp1-deficient cells were also observed in absolute cell numbers, with significantly fewer Blimp1-KO cells compared to WT cells in the SI starting at day 7 of infection (Figure 3.2G). Blimp1 has also been shown to work in conjunction with the TF Hobit (encoded by *Zfp683)* to mediate T_{RM} development[110,128]; however, *Zfp683* was not detected in our bulk RNAseq data. All together, these data demonstrated a requirement for Blimp1 and Id2 in $CD4^+$ T_{RM} development and further strengthened a developmental relationship between T_{RM} and T_H1 cells.

Figure 3.1 $CD4^+$ T_{RM} express Blimp1 and Id2 during viral infection. (A)) Plot of PageRank scores and gene expression of SMARTA IEL and LPL samples over SMARTA spleen samples. Data are cumulative of 2 experiments with n=4-5 mice per experiment for day 7. Yellow points denote genes with an absolute value of greater than 1 for the log2 of fold-change between SI values and splenic values. **(B)** Representative histograms of Blimp1-YFP expression by SMARTA $CD4^+$ T cells from spleen, mLN, and SI on day 7 or 21 following LCMV-Arm infection. (**C**) Quantification of Blimp-YFP^+ cells of total SMARTA $CD4^+$ T cells in indicated tissues. **(D)** Representative flow plots of Id2-YFP and Id3-GFP expression by SMARTA $CD4^+$ T cells from SPL, mLN, and SI on day 7 or 21 of infection. **(E)** Quantification of the frequency of Id2-YFP^+ or ID2-YFP^+Id3-GFP^+ cells of total SMARTA $CD4^+$ T cells. **(F)** Representative flow plots of Blimp1-YFP and Id3-GFP expression by SMARTA $CD4^+$ T cells from SPL, mLN, and SI on day 7 or 21 of infection. Numbers in flow plots or histograms indicate percent of cells in corresponding gate. Data are representative (A,C,E) or cumulative (B,D) of 3 experiments with n=2-4 mice per experiment. Graphs show mean ± SD; *$p < 0.05$, ** $p<0.01$, *** $p<0.001$, ****$p< 0.0001$.

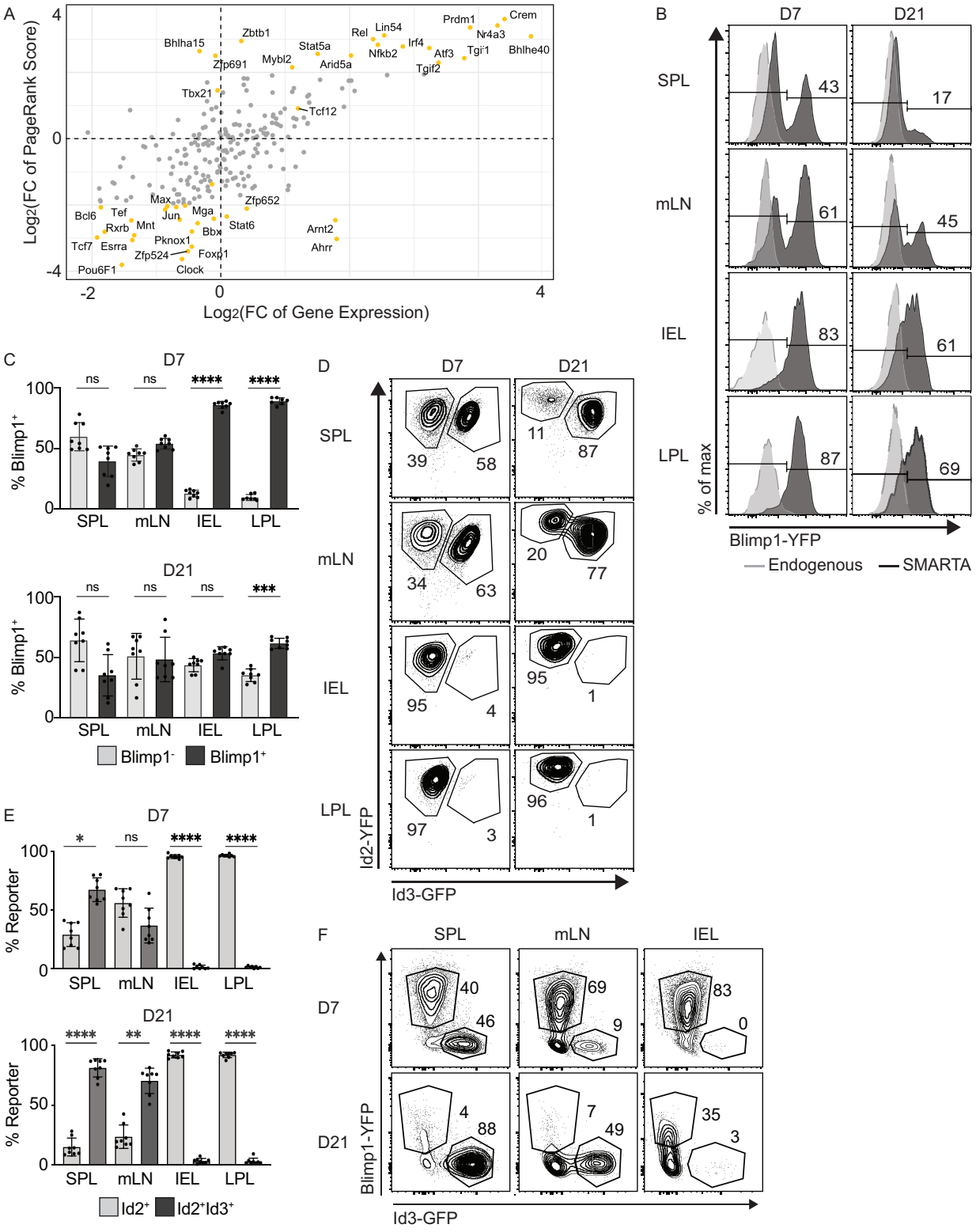
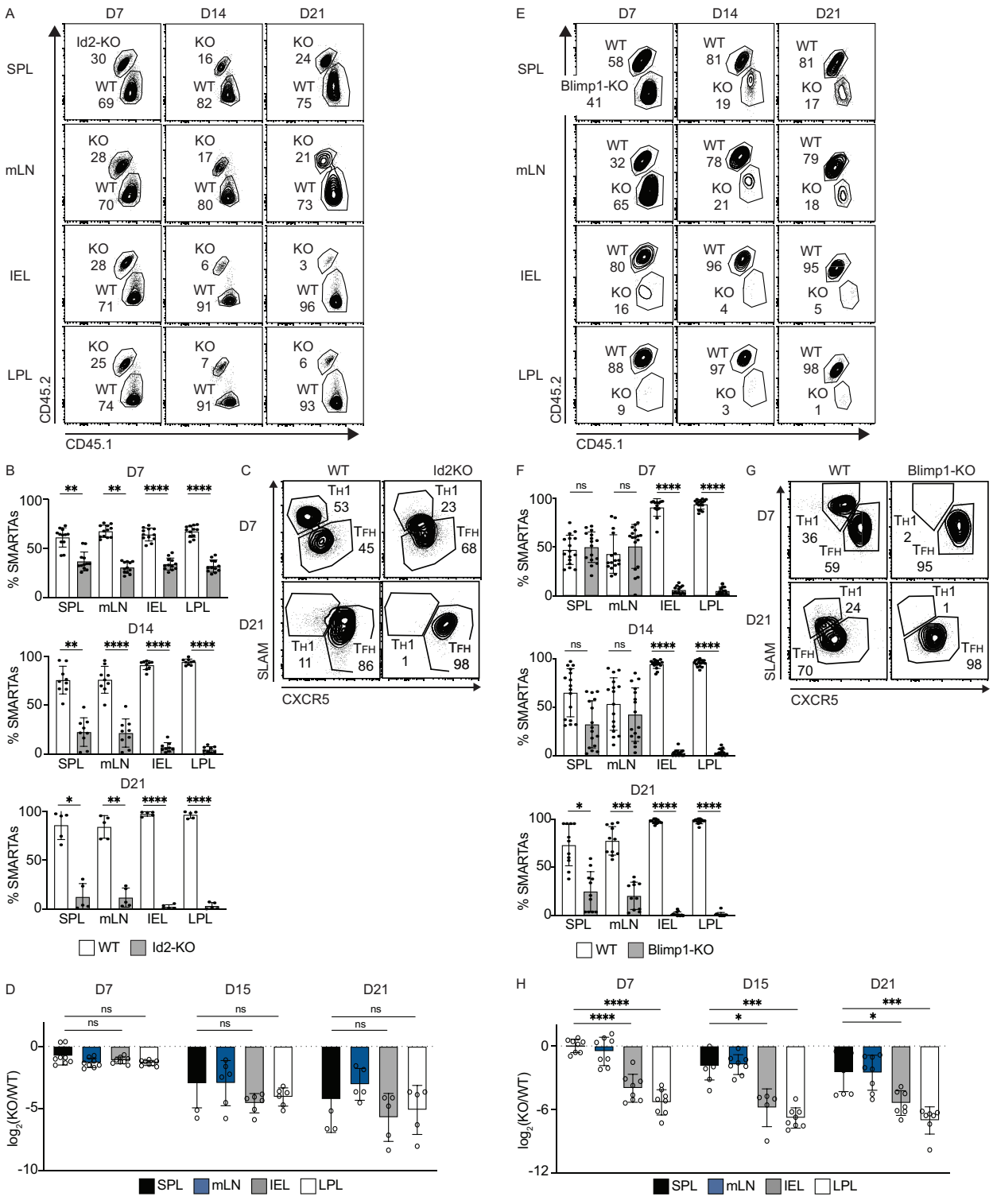

A
Log2(FC of PageRank Score)
Log2(FC of Gene Expression)
Prdm1
Crem
Nr4a3
Bhlhe40
Tgif1
Atf3
Irf4
Lin54
Rel
Nfkb2
Stat5a
Arid5a
Mybl2
Zbtb1
Bhlha15
Zfp691
Tbx21
Tcf12
Zfp652
Max
Mga
Jun
Tef
Bcl6
Rxrb
Mnt
Stat6
Bbx
Pknox1
Zfp524
Foxp1
Clock
Esrra
Tcf7
Pou6F1
Arnt2
Ahrr
B
D7
D21
SPL
mLN
IEL
LPL
43
17
61
45
83
61
87
69
% of max
Blimp1-YFP
Endogenous
SMARTA
C
D7
D21
% Blimp1+
ns
****
***
SPL
mLN
IEL
LPL
Blimp1-
Blimp1+
D
D7
D21
Id2-YFP
Id3-GFP
39
58
11
87
34
63
20
77
95
4
95
1
97
3
96
1
E
D7
D21
% Reporter
*
**
Id2+
Id2+Id3+
F
SPL
mLN
IEL
D7
D21
Blimp1-YFP
Id3-GFP
40
46
69
9
83
0
4
88
7
49
35
3

Figure 3.2 Loss of Id2 and Blimp1 impair CD4$^+$ T_{RM} differentiation. (A,E) Representative flow plots comparing the frequency of WT and Id2-KO **(A)** or Blimp1-KO **(E)** SMARTA CD4$^+$ T cells in indicated tissues at specific timepoints. **(B,F)** Quantification of the frequency of WT and Id2-KO **(B)** or Blimp1-KO **(F)** of total SMARTA CD4$^+$ T cell population. **(C,G)** SLAM and CXCR5 expression by WT and Id2-KO **(C)** or Blimp1-KO **(G)** SMARTA CD4$^+$ T cells from the SPL to distinguish T_H1 (SLAM$^+$CXCR5$^+$) and T_{FH} (SLAM$^-$CXCR5$^+$) subsets at indicated day of infection. **(D,H)** Cell numbers of WT and Id2-KO **(D)** or Blimp1-KO **(H)** SMARTA CD4$^+$ T cells presented as log_2fold-change of KO to WT cells on indicated day of infection. Data are representative (A,C,E,G) or cumulative (B,D,F,H) of 3 experiments with n=2-4 mice per experiment. Graphs show mean ± SD; *p < 0.05, ** p<0.01, *** p<0.001, ****p< 0.0001.


A
D7
D14
D21
SPL
mLN
IEL
LPL
Id2-KO
KO
WT
CD45.2
CD45.1
E
Blimp1-KO
B
% SMARTAs
SPL
mLN
IEL
LPL
WT
Id2-KO
C
WT
Id2KO
TH1
TFH
SLAM
CXCR5
F
Blimp1-KO
G
D
D15
log2(KO/WT)
H
ns
SPL
mLN
IEL
LPL

Bcl6 plays a dual role in the antiviral CD4$^+$ T_{RM} differentiation program.

The requirement for Blimp1 and Id2 in regulating the residency program suggests a relationship between the CD4$^+$ T_{RM} and T_H1 subsets. However, based on our sequencing data, we also observed expression of *Bcl6* by D21 SI CD4$^+$ T_{RM} (Figure 2.5B), which was unexpected given the antagonism of Blimp1 and Id2 by Bcl6. Thus, we hypothesized that Bcl6 may also be regulating the CD4$^+$ T_{RM} program in a manner that is distinct from its known function within circulating effector and T_{FH} populations.

To assess the role of Bcl6 in CD4$^+$ T_{RM} development, we transferred a 50:50 mix of WT and Bcl6$^{f/f}$CD4-Cre$^+$ (Bcl6-KO) SMARTA cells into congenically distinct recipients and infected with LCMV-Arm (Figure 6A). On days 7, 14, and 21 of infection, we isolated the transferred cells, and compared the frequency and numbers of WT and Bcl6-KO cells. In contrast to the Blimp1 and Id2 results, loss of Bcl6 resulted in a small but significant increase in SI CD4$^+$ T cell accumulation compared to the WT cells at day 7 of infection (Figures 3.3A,B). However, by day 14 of infection, the frequency of Bcl6-deficient cells was lower than that of WT cells in the SI compartment, and this reduction continued into the later memory time point. Loss of Bcl6 also significantly reduced the splenic and lymph node CD4$^+$ T cell populations starting at day 7 of infection, and this deficiency was amplified at memory time points (Figure 3.3A,B). The differences in frequencies of WT and Bcl6-KO CD4$^+$ T cells were mirrored in the absolute cell counts (Figure 3.3E). Consistent with the role of Bcl6 in T_{FH} differentiation, splenic Bcl6-KO cells on day 7 of infection did not express CXCR5 and were predominantly SLAM$^+$ (Figure 3.3C). However, by day 21 of infection, subsets of SLAM$^+$ and SLAM$^-$ were observed in Bcl6-KO CD4$^+$ T cell population from the spleen. Interestingly, Bcl6-deficient SI CD4$^+$ T cells displayed increased CD69 expression on day 7 and 21 and reduced CD27 expression on day 21 compared to WT cells, indicating that Bcl6

may inhibit the T_{RM} program at the peak of infection (Figure 3.3D). This was consistent with mRNA levels for day 7 in which IEL and LPL effector cells expressed lower *Bcl6* than splenic samples (Figure 2.5B). Thus, the impact of Bcl6 in the SI is more complex than previously thought, with different roles at effector versus memory time points and within the circulation versus non-lymphoid tissues.

To further assess the role of Bcl6 in early $CD4^+$ T_{RM} development, we performed bulk RNAseq of WT and Bcl6-KO SMARTA $CD4^+$ T cells at day 7 of infection. PCA of gene expression for these samples showed that SI Bcl6-KO $CD4^+$ T cells were a distinct population, while splenic Bcl6-KO $CD4^+$ T cells clustered closely with WT SI and T_H1 samples (Figure 3.4A). GSEA also revealed that loss of Bcl6 resulted in an enrichment of the $CD4^+$ NLT gene signatures by both splenic and SI populations and the $CD8^+$ T_{RM} signature by SI $CD4^+$ T cells (Figure 3.4B), which is further supported by module score analysis of scRNAseq data (Figure 3.4C). This indicated that Bcl6 may repress or antagonize T_{RM} characteristics at day 7 of infection. Interestingly, Bcl6-KO cells in the IEL and LPL at day 7 were enriched for genes upregulated in D21 SI samples compared to D7 SI samples, suggesting that Bcl6 might repress the progression of effector SI cells into mature T_{RM} (Figure 3.4B). Finally, we examined how the loss of Bcl6 influenced gene expression in the early $CD4^+$ T_{RM} program by comparing WT and Bcl6-KO cells in scRNAseq (Figures 3.4C,D). Loss of Bcl6 in IEL and LPL cells led to increased expression of direct gene targets of Bcl6 repression, *Id2*, *Prdm1*, and *Runx3*, which supported the role of Bcl6 in inhibiting T_{RM} cell differentiation during early infection[103,129]. There was decreased expression of T_{RM}-defining surface molecule *Cd69* by Bcl6-KO IEL cells and a very minimal increase in Bcl6-KO LPL cells; however, *Cd69* is not a known direct target of Bcl6 and may not be involved in this circuitry. Lastly, we observed no changes in *Klf2* expression nor any commensurate

decrease in *S1pr1* by the Bcl6-KO $CD4^+$ T cells in the SI, which is unexpected given that Bcl6 is known to inhibit expression of *Klf2* in circulating effector $CD4^+$ T cells[129]. Collectively, our data demonstrated that Bcl6 inhibited tissue-residency genes at an early phase of an immune response and was needed to maintain long-lived T_{RM} populations.

Figure 3.3 Loss of Bcl6 at day 7 enhances the T_{RM} differentiation program. (A) Representative flow plots showing the frequency of WT and Bcl6-KO SMARTA $CD4^+$ T cells in indicated tissues at specific time of infection. **(B)** Quantification of the frequency of WT and Bcl6-KO of total SMARTA $CD4^+$ T cell population. **(C)** SLAM and CXCR5 expression by WT and Bcl6-KO SMARTA $CD4^+$ T cells from the SPL to distinguish T_H1 ($SLAM^+CXCR5^+$) and T_{FH} ($SLAM^-CXCR5^+$) subsets at indicated day of infection. **(D)** Quantification of geometric mean fluorescence intensity (gMFI) of CD27 and CD69 by WT and Bcl6-KO SMARTA $CD4^+$ T cells in indicated tissues at specific time of infection. **(E)** Cell numbers presented as log_2fold-change of Bcl6-KO to WT cells at indicated day of infection. Data are representative (A,C,D) or cumulative (B,E) of 3 experiments with n=2-4 mice per experiment. Graphs show mean ± SD; *$p < 0.05$, ** $p<0.01$, *** $p<0.001$, ****$p<0.0001$.

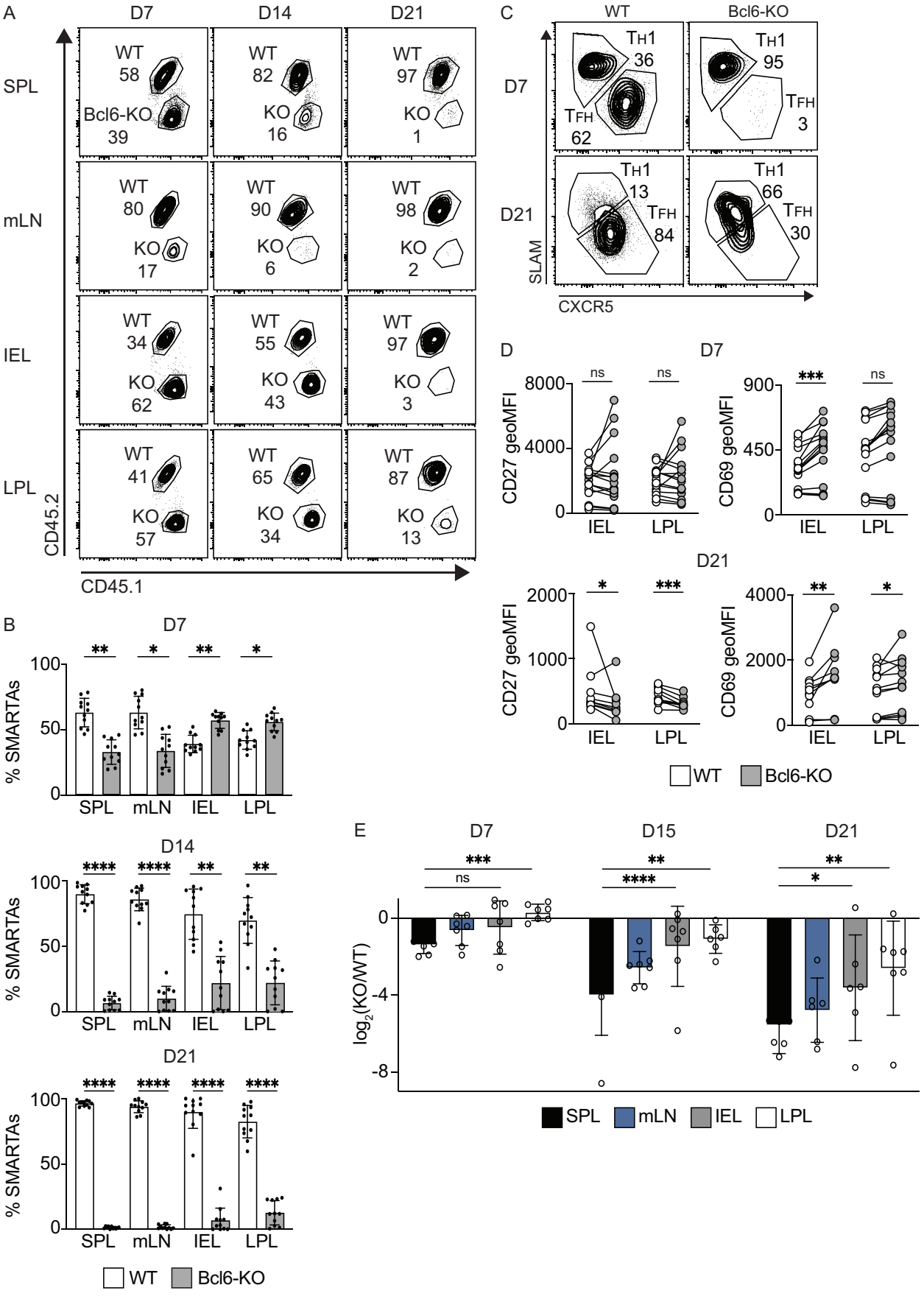
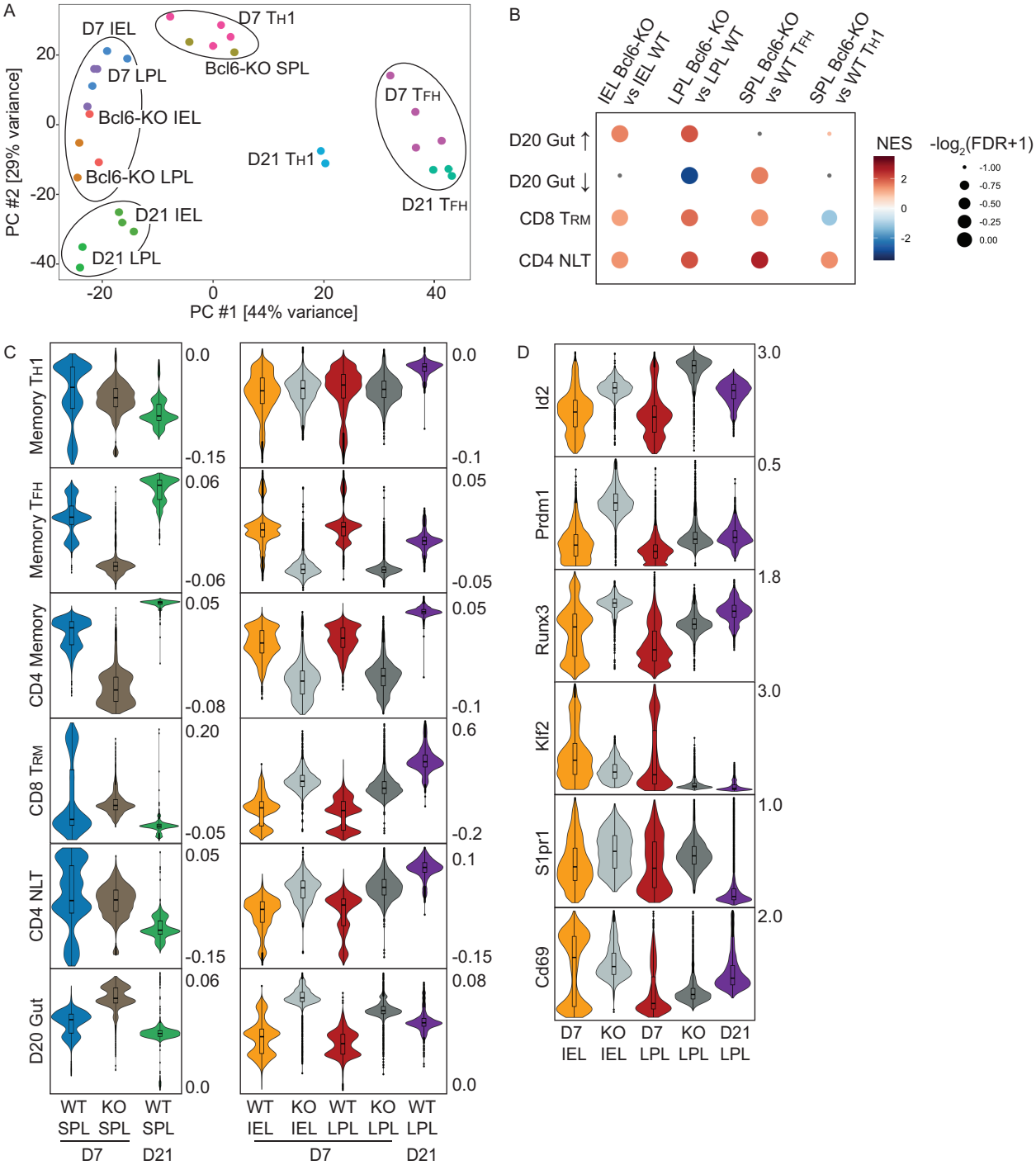

A
D7
D14
D21
SPL
mLN
IEL
LPL
WT
Bcl6-KO
KO
CD45.2
CD45.1
B
D7
D14
D21
% SMARTAs
SPL
mLN
IEL
LPL
WT
Bcl6-KO
C
WT
Bcl6-KO
D7
D21
TH1
TFH
SLAM
CXCR5
D
D7
D21
CD27 geoMFI
CD69 geoMFI
IEL
LPL
ns
WT
Bcl6-KO
E
D7
D15
D21
log2(KO/WT)
SPL
mLN
IEL
LPL

Figure 3.4 Loss of Bcl6 enhances the tissue-residency program in early CD4$^+$ T_{RM} differentiation. (A,B) Bulk RNAseq of WT and Bcl6-KO SMARTA CD4$^+$ T cells from SPL and SI, harvested at day 7 of LCMV-Arm infection. **(A)** Principal component analysis (PCA) of RNAseq data. **(B)** Gene enrichment analysis (GSEA) of Bcl6-KO compared to D7 WT RNAseq data. **(C,D)** Single-cell RNAseq of WT and Bcl6-KO SMARTA CD4$^+$ T cells from SI, harvested at day 7 and 21 of LCMV-Arm infection. **(C)** Violin plots showing module score enrichment of indicated gene signature in WT and Bcl6-KO SMARTA CD4$^+$ T cells from SPL and SI, harvested at day 7 and 21 of LCMV-Arm infection. **(D)** Violin plots showing expression of important transcriptional regulators or cell-surface molecules. Data are cumulative of 1-2 experiments with n=12-15 mice per experiment. Graphs show mean ± SD; *p < 0.05, ** p<0.01, *** p<0.001, ****p<0.0


A
D7 IEL
D7 TH1
Bcl6-KO SPL
D7 LPL
D7 TFH
Bcl6-KO IEL
D21 TH1
Bcl6-KO LPL
D21 TFH
D21 IEL
D21 LPL
PC #2 [29% variance]
PC #1 [44% variance]
B
IEL Bcl6-KO vs IEL WT
LPL Bcl6- KO vs LPL WT
SPL Bcl6-KO vs WT TFH
SPL Bcl6-KO vs WT TH1
D20 Gut ↑
D20 Gut ↓
CD8 TRM
CD4 NLT
NES
-log2(FDR+1)
C
Memory TH1
Memory TFH
CD4 Memory
CD8 TRM
CD4 NLT
D20 Gut
WT SPL
KO SPL
WT SPL
WT IEL
KO IEL
WT LPL
KO LPL
WT LPL
D7
D21
D
Id2
Prdm1
Runx3
Klf2
S1pr1
Cd69
D7 IEL
KO IEL
D7 LPL
KO LPL
D21 LPL

3.3 Discussion

Both our bulk and single-cell RNA sequencing analyses revealed the transcriptional heterogeneity of $CD4^+$ T_{RM}, with co-expression of genes important in both the effector and memory cell differentiation program. We thus investigated the function of three transcriptional regulators, Blimp1, Id2, and Bcl6, in driving the development of an SI $CD4^+$ T_{RM} population.

We showed that both Blimp1 and Id2, critical factors in T_H1 effector T cells, were required for the early seeding of $CD4^+$ T_{RM} in the SI, favoring T_H1 effector cells as the precursors of SI T_{RM} in LCMV-Arm infection. Given its established role as the T_H1 fate-defining TF and tissue residency, Blimp1 may be driving the T_H1 characteristics as well as the tissue-residency aspects of the $CD4^+$ T_{RM} program. This is consistent with previous studies on the functions of Blimp1 and its family member, Hobit, in regulating the transcriptional program of $CD8^+$ T_{RM} by inhibiting lymphocyte egress from the tissues and by repressing the development of circulating memory cells[110] and in driving T_{RM}-mediated colitis in human $CD4^+$ T_{RM} in the colon[128]. It is possible that Blimp1 and Hobit are also cooperating in SI $CD4^+$ T_{RM}, though Blimp1 seems to be the dominant TF between the two. We do not see evidence of a role for Hobit in $CD4^+$ T_{RM} as loss of Blimp1 depleted the T_{RM} population, and we did not observe significant Hobit expression. Id2-deficiency led to a significant decrease in $CD4^+$ T cells in both the SI and spleen on day 7 of infection. However, by day 14, when cells have transitioned into the memory phase, the loss of KO cells in the SI was significantly greater than the defect in spleen and mLN. This suggested that Blimp1 and Id2 are driving more than the T_H1-associated factors in T_{RM} cells, mediating expression of genes required for the migration into tissues or inhibiting genes that allow for egress.

In our study, loss of Bcl6 led to an initial relative increase in the SI $CD4^+$ T_{RM} population compared to WT SMARTA cells, suggesting that Bcl6 may act as an repressor of the early T_{RM}

developmental program. This result is consistent with the essential activity of Bcl6 in T_{FH} cells, repressing Blimp1 and Id2 to inhibit the T_H1 program[129]. In the absence of Bcl6, Blimp1 and Id2 are unrestrained and can direct more effector $CD4^+$ T cells towards the T_H1/T_{RM} fate. When we compared the transcriptional profiles of WT and Bcl6-KO cells at day 7, we observed an enrichment for T_{RM}-associated genes by KO cells from the SI and even within the splenic population. It is noteworthy that Bcl6 deletion did not increase the expression levels of *Klf2* or *S1pr1* by SI cells at day 7, even though Klf2 is a known target for Bcl6 inhibition. This may be due to higher observed expression of *Icos* in Bcl6-deficient IEL and LPL cells, which leads to greater inhibition of Foxo1 signaling by Akt, a downstream target of ICOS. It appears that at the effector stage, Bcl6 is actively inhibiting the ability of T_{RM} to become resident, more so than enhancing these cells' ability to stay in the circulation. Surprisingly, by a memory time point, Bcl6 was needed for the T_{RM} population, as *Bcl6* was expressed in day 21 SI cells and deletion of Bcl6 on day 14 and 21 decreased the frequency and number of T_{RM} cells. This may be because Bcl6 is required to repress Klf2 in order for T_{RM} cells to downregulate S1pr1 and prevent recirculation, which is supported by our bulk RNAseq analysis[97,129]. Bcl6-expressing T_{FH} cells have been shown to share a transcriptional profile with $CD8^+$ memory-precursor T cells and contribute to the $CD4^+$ memory T cell population, suggesting that Bcl6 promotes long-lived $CD4^+$ and $CD8^+$ T cells. Thus, in $CD4^+$ T_{RM} cells, Bcl6 may actively enhance the memory attributes of this population which otherwise is controlled by effector-associated factors including Blimp1 and Id2.

Here, we have highlighted the importance of Blimp1, Id2, and Bcl6 in the $CD4^+$ T_{RM} population in viral infection for regulating the differentiation of this critical immune subset. Blimp1 and Id2 were required at early stages of development to oppose the active inhibition of T_H1- and tissue-associated and genes by Bcl6 and to direct the T_{RM} precursors into a SI-oriented,

resident program. During the contraction phase, Bcl6 appeared to be promoting tissue residency, possibly repressing genes that would drive cells to leave the tissue and re-enter circulation. Further work is required to elucidate how $CD4^+$ T_{RM} are maintained in the tissues, particularly how Bcl6 may be regulating the T_{RM} program differently in effector and memory stages. A better understanding of the transcriptional regulation of $CD4^+$ T_{RM} in the intestine will enhance our knowledge of T_{RM} directed immune responses in the context of long-term memory and how to target these cells for therapeutic purposes.

Chapter 3 was adapted from a manuscript which will be submitted for publication. **Nguyen QP**, Deng TZ, O'Shea SM, Pipkin ME, Choi J, Crotty S, Goldrath AW. (2021). *Blimp1, Id2, and Bcl6 balance effector and memory-associated programs to promote $CD4^+$ T_{RM} differentiation following viral infection.* The dissertation author was the primary author of this paper.

Chapter 4 Perspectives

This dissertation aimed to elucidate the differentiation pathway and transcriptional programming of antiviral $CD4^+$ T_{RM} cells in the SI. T_{RM} cells, by nature, are required to be both rapid effector cells and long-lived memory cells, thus raising the question of how their transcriptional network balances these contrasting characteristics. In the context of viral infection, we found that $CD4^+$ T_{RM} were a heterogenous population with phenotypic, transcriptional, and epigenetic similarities to the T_H1 subset during the effector time point. As $CD4^+$ T_{RM} cells transitioned into a memory population, they developed a distinct transcriptional profile from splenic cells, specifically expressing both T_H1/effector (*Tbx21*, *Prdm1*, *Id2, Stat4*) and T_{FH}/memory-associated (*Icos*, *Pdcd1*, *Bcl6, Stat3)* genes in the same population, including co-expression of antagonistic factors, Blimp1 and Bcl6, in the same cells. We identified roles for Blimp1 and Id2 in driving early $CD4^+$ T_{RM} development consistent with their importance in T_H1 lineage differentiation and the T_H1 functional characteristics of $CD4^+$ T_{RM}. Bcl6 inhibited the early T_{RM} program, consistent with its role as a key TF in T_{FH} differentiation and known repressor of Blimp1 and Id2. However, by the later memory stage, loss of Bcl6 expression resulted in a significant decrease in frequency of $CD4^+$ T_{RM} in the SI, suggesting that Bcl6 is required for long-term persistence in tissues.

Multiple levels of regulation necessary for $CD4^+$ T_{RM} differentiation have been proposed, including elements which dictate tissue localization versus circulation, determine $CD4^+$ versus $CD8^+$ lineage commitment, and regulate stable residency versus recirculation potential[71]. Our data adds an additional level of regulation, one directing the effector versus memory characteristics of T_{RM} cells. We propose that $CD4^+$ T_{RM} required a gene expression signature with TFs associated with both effector and memory cell fates in order to balance the required rapid response of effector

lymphocytes against the persistence and plasticity inherent in memory populations. In acute viral infection, we find that $CD4^+$ T_{RM} cells likely originate from a T_H1-like progenitor, which accounts for many of the effector characteristics in their transcriptional program. Once T_{RM} migrate to the site of infection, tissue-intrinsic cues and antigen clearance can activate a second wave of transcriptional programming which enforces the memory qualities of this long-lived population.

The fine-tuning mechanisms required for such a "hybrid" cell-state remains unexplored, but our work on Blimp1, Id2, and Bcl6 provides a solid foundation for elucidating the interwoven transcriptional networks in T_{RM}. The repressor activity of Bcl6 in T_{RM}, especially in the memory phase, is particularly intriguing given that we do not yet know how Bcl6 is acting in circulating T_H1 and T_{FH} memory populations. *Bcl6* was expressed by D21 T_H1 in sequencing data, suggesting role for this gene in memory T_H1 cells. A potential strategy to address this question is using $Bcl6^{f/f}$ ER-Cre SMARTA mice for induced deletion of Bcl6 at later stages of infection such as day 7 with peak virus levels or day 14 when effector T cell contraction begins. These studies may reveal how Bcl6 contributes to the maintenance of the $CD4^+$ T_{RM} population and identify downstream targets of its repressor activity during the memory phase.

As the T_{RM} field moves forward, there are three outstanding questions which warrant further investigation. First, much of work on T_{RM} have examined these subsets in isolation, but in the body, $CD4^+$ T_{RM} are constantly interacting with $CD8^+$ T cells, B cells, innate immune cells, and non-immune cells in the tissue[130]. Future studies must consider the transcriptional regulation of the aggregate cell-cell interactions and how perturbations in one subset, such as T_{RM}, affect the remaining cell populations and the collective immune response. Visualization techniques which combine fluorescent imaging with proximity-based chemical tagging can elucidate how T_{RM} cells physically interact with their surroundings[131]. Additionally, technologies such as Single Cell

Optical Phenotyping and Expression (SCOPE-seq), which combines live single-cell imaging and scRNAseq, will enable real-time monitoring of live cells and allow for exploration of the inter- and intra-cellular transcriptional networks[132]. Second, as transcriptional regulators are identified for T_{RM} in mice, concurrent work must focus on exploring the role of similar transcriptional programs in human T_{RM} populations. Lastly, recent work has highlighted the ability of some T_{RM} to migrate out of the tissues and re-enter circulation and SLOs during the secondary response, an example of "outside-in" immunity[133]. However, it remains unclear which tissues or T_{RM} subsets have this recirculation potential and how this phenomenon is regulated in the overall tissue-residency program.

Tissue-resident T lymphocytes are becoming increasingly understood for their importance in immune defenses, and the potential to target T_{RM}, especially $CD4^+$ T cells, for tissue-specific vaccines and therapies is exciting. The majority of infectious pathogens invade through our mucosal barrier surfaces, including many which cause diseases for which we have no vaccines or treatment. We can address this need, especially in many developing countries, by targeting T_{RM} cells which are specific for these pathogens and localized to the likely sites of re-infection[134]. Additionally, in the context of cancers, T_{RM}-like cells have been identified as a component of tumor infiltrating lymphocytes within solid tumors and can be an effective target for cancer vaccines or immunotherapy[135,136]. Finally, $CD4^+$ T_{RM} cells have been implicated in the pathology of autoimmunity given the tissue-specific nature of many autoimmune diseases[137,138]. Enhancing T_{RM} formation, function, and maintenance can reduce pathogen spread in the earliest phases of infection and can be leveraged for vaccines which prevent infection from becoming established in the first place and therapies which protect against disease symptoms and improve clinical

outcomes. The work in this dissertation will provide a basis to better understand this essential memory T cell population and exploit their protective capacity in the immune response.

Appendix A

Material 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. Blimp1-YFP mice (stock #008828;The Jackson Laboratory), Id3-GFP mice [139], Id2-YFP mice [140], *Id2*$^{fl/fl}$ mice [141], CD4-Cre mice (stock#017336; The Jackson Laboratory), SMARTA mice (T cell receptor (TCR) transgenic for I-A^{b}-restricted LCMV glycoprotein amino acids 66–77 (Oxenius et al.), CD45.1, and CD45.1.2 congenic mice were bred in house. *Bcl6*$^{f/f}$CD4-Cre^{+}CD45.1^{+} SMARTA and *Prdm1*$^{f/f}$CD4-Cre^{+}CD45.1^{+} SMARTA mice were received from the Crotty lab at the La Jolla Institute. Recipient C57BL/6J (B6) mice were either bred at UCSD or received from The Jackson Laboratory.

T cell transfer and infection

Naive wild-type, Blimp1-YFP, Id2-YFP/Id3-GFP SMARTA CD4^{+} T cells congenically distinct for CD45 were adoptively transferred intravenously at 1x10^{5} cells. For cotransfers, *Prdm1*$^{f/f}$CD4-Cre^{+}, *Id2*$^{f/f}$CD4-Cre^{+}, *Bcl6*$^{f/f}$CD4-Cre^{+} and corresponding control SMARTA CD4^{+} T cells were mixed in a 1:1 ratio and adoptively transferred at 1x10^{5} total cells per recipient mouse. Mice were then infected intraperitoneally with 2x10^{5} plaque-forming units (PFU) of lymphocytic choriomeningitis virus-Armstrong (LCMV-Arm).

Preparation of single cell suspensions

Single-cell suspensions were prepared from spleen or lymph node by mechanical disruption. For small intestine (SI) preparations, Peyer's patches were excised, luminal contents were removed,

and tissue was cut longitudinally then into 1 cm pieces. The gut pieces were incubated while shaking in 10% HBSS/HEPES bicarbonate solution containing 15.4mg/100□L of dithioerythritol (EMD Millipore) at 37°C for 30 minutes to extract SI-IEL. For lamina propria lymphocyte isolation, gut pieces were further treated with 100U/ml type I collagenase (Worthington Biochemical) in RPMI-1640 containing 5% bovine growth serum, 2 mM $MgCl_2$ and 2 mM $CaCl_2$ at 37°C for 40 minutes. Cells were then filtered to remove any remaining tissue pieces. Lymphocytes from all tissue but skin, spleen and lymph node were purified on a 44%/67% Percoll density gradient.

Antibodies, flow cytometry, and cell sorting

The following antibodies were used for surface staining (all from eBioscience unless otherwise specified): CD4 (GK1.5), CD45.1 (A20), CD45.2 (104), B220 (RA3-6B2, 1:400), SLAM (TC15-12F12.2, BioLegennd), CD8 (53-6.7, BioLegend), CD69 (H1.2F3, Biolegend), CD103 (2E7), CD27 (LG.7F9), Ly6C (HK1.4, BioLegend), CD199/CCR9 (eBioCW-1.2), CD49d (R1-2), P2Rx7 (1F11, BioLegend). Cells were incubated for 30 minutes at 4°C in PBS supplemented with 2% bovine growth serum and 0.1% sodium azide, unless specified otherwise. CXCR5 staining was performed by incubating cells with purified anti-CXCR5 (SPRCL5, 1:50; Invitrogen), followed by PE-Cy7- or BV510-labeled streptavidin (1:1000, eBioscience) each for 30 minutes at 4°C. Intracellular staining was performed using the eBioscience FOXP3/Transcription Factor Staining Buffer Set and the following antibodies: Granzyme A (3G8.5, BioLegend), Granzyme B (GB12, Invitrogen), IFNγ (XMG1.2, BioLegend), IL-2 (JES6-5H4), and TNFα (MP6-XT22). For cytokine staining, $CD4^+$ T cells from the spleens, lymph nodes, and SI were incubated for 6 hours at 37°C in RPMI-1640 media containing 10% (v/v) bovine growth serum with 10nM GP_{66-81} peptide and Protein Transport Inhibitor (eBioscience). Stained cells were analyzed using LSRFortessa or

LSRFortessa X-20 cytometers (BD) and FlowJo software (TreeStar). All sorting was performed on the BD FACSAria instrument.

Bulk RNAseq library construction and sequencing

On day 7 or 21 of LCMV infection, $1x10^3$ T_H1 ($CD4^+SLAM^+CXCR5^-$) or T_{FH} ($CD4^+SLAM^+CXCR5^+$) SMARTA $CD4^+$ T cells from the spleen and $1x10^3$ total SMARTA $CD4^+$ T cells from the IEL or LPL were sorted into TCL buffer (Qiagen) with 1% 2-Mercaptoethanol. On day 7 of LCMV infection, $1x10^3$ *Bcl6*$^{f/f}$CD4-Cre^+ or WT SMARTA $CD4^+$ T cells from the spleen, IEL, or LPL were sorted similarly. RNAsequencing was performed on duplicate or triplicate samples as such: for day 7 LCMV sorts, 3-4 spleens and SI IEL or LPL were pooled per replicate; for day 21 LCMV sorts, 10-15 spleens and SI IEL or LPL were pooled per replicate; for Bcl6-KO experiments, 5-8 spleens and SI IEL or LPL were pooled per replicate.

For all samples, total RNA was captured and purified on RNAClean XP beads (Beckman Coulter). Polyadenylated mRNA was then selected using an anchored oligo(dT) primer (5′–AAGCAGTGGTATCAACGCAGAGTACT30VN-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.

Bulk RNAseq analysis

Normalized counts were analyzed using the Deseq2 R package to identify differentially expressed genes across all samples using a padj <0.05. For principal component analysis (PCA) plots, samples were graphed using the plotPCA command using only identified differentially expressed genes. Gene set enrichment analysis was performed on each cell subset using the GSEA software [143,144]. The $CD8^+$ T_{RM} signature is from Milner et al., 2017, the $CD4^+$ NLT Signature is from Beura et al., 2019, and the T_H1 Signature and T_{FH} Signature are from Ciucci et al.,2019. Normalized enrichment scores (NES) and false discovery rate (FDR) were visualized using the ggplot2 package in R. Heatmaps in Figure 3 were created using the seaborn python package, and genes for cell-surface molecules and transcription factors were cherry picked based on importance.

10x Genomics library preparation and sequencing

SMARTA $CD4^+$ T cells at day 7 or 21 of LCMV-Arm infection were sorted from the spleen or SI IEL and LPL and resuspended in PBS+0.04% (w/v) bovine serum albumin. Approximately 10,000 cells per sample were loaded into Single Cell A chips (10x Genomics) and partitioned into Gel Bead In-Emulsions (GEMs) in a Chromium Controller (10x Genomics). 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 RNAseq analysis

scRNAsequencing 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 (cite), 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.

Three cell-gene matrices were generated:

(1) Raw UMI matrix.

(2) UPM matrix. The raw UMI matrix was normalized to get UMIs per million reads (UPM), and was then log2 transformed. All downstream differential analysis was based on the UPM matrix.

(3) MAGIC matrix. UPM matrix was further permuted by MAGIC [145]. R package Rmagic 1.0.0 was used, and all options were kept as default. MAGIC aims to correct the drop-out effect of scRNAseq data; thus, we used MAGIC-corrected matrix for visualizing the gene expression pattern rather than using the UPM matrix. All gene expression overlaid on TSNE plots were based on the MAGIC matrix.

Overlay of gene signatures onto single cell data was done with AddModuleScore. Gene signatures were taken from published studies ($CD8^+$ T_{RM} signature from Milner et al., 2017, the $CD4^+$ NLT Signature from Beura et al., 2019) or calculated from bulk RNAseq data as follows: 1) SPL signature: genes upregulated in spleen samples compared to SI samples at day 7 and 20 by at least a fold change of 1.5

2) Gut signature: genes upregulated in SI samples compared to spleen samples at day 7 and 20 by at least a fold change of 1.5

3) Memory T_H1: genes upregulated in D40 T_H1 samples compared to D7 T_H1 samples by at least a fold change of 1.5

4) Memory T_{FH}: genes upregulated in D40 T_{FH} samples compared to D7 T_{FH} samples by at least a fold change of 1.5.

ATAC–seq library construction, sequencing, and analysis

Sorted cells (2.5-5.0 × 10^4) were resuspended in 25 μL of lysis buffer and spun down at 600g for 30 min at 4°C. The nuclear pellet was resuspended in 25 μl of Tn5 transposase reaction mixture (Nextera DNA Sample Prep Kit, Illumina) and incubated for 30 min at 37°C. Transposase-associated DNA was subsequently purified (Zymo DNA clean-up kit). For library amplification, DNA was amplified using indexing primer from Nextera kit and NEBNext High-Fidelity 2× PCR master mix. Then, the amplified DNA was size-selected to fragments less than 800 bp using NEB Ampure XP Beads. The library was sequenced using Hiseq 2500 for single- end 50-bp sequencing to yield at least 10 million reads. ENCODE pipeline [146] was used for mapping and peaks counting. Differential accessibility regions between samples were calculated using optimal peak files in DiffBind [147,148]. PageRank analysis was performed with Taiji software as described previously [89] as a separate bioinformatic algorithm based on differential gene expression and TF binding motifs that provides compatible results with differential ATAC–seq analysis, although the nature of the PageRank algorithm is more effective for identifying activator TFs rather than repressors. This algorithm works by generating a network of DAR and differentially expressed genes, and then predicting the transcription factors most likely to be driving these differences based on the TF motifs present in DARs. Each transcription factor is then assigned a score based on its weight in the network.

Statistical methods.

Statistical tests were performed using Prism (7.0/9.0) (Graphpad). Two-tailed paired or unpaired t-test was used for comparisons between groups. P values of <0.05 were considered significant.

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.

References

1. Foulds, K. E., Zenewicz, L. A., Shedlock, D. J., Jiang, J., Troy, A. E. & Shen, H. Cutting Edge: CD4 and CD8 T Cells Are Intrinsically Different in Their Proliferative Responses1. *J Immunol* **168**, 1528–1532 (2002).

2. Williams, M. A., Ravkov, E. V & Bevan, M. J. Rapid Culling of the CD4 + T Cell Repertoire in the Transition from Effector to Memory. *Immunity* **28**, 533–545 (2008).

3. Akondy, R. S., Monson, Nathan, D., Miller, J. D., Edupuganti, S., Teuwen, D., Wu, H., Quyyumi, F., Garg, S., Altman, J. D., Del Rio, C., Keyserling, H. L., Ploss, A., Rice, C. M., Orenstein, W. A., Mulligan, M. J. & Ahmed, R. The Yellow Fever Virus Vaccine Induces a Broad and Polyfunctional Human Memory CD8_ T Cell Response. *J. Immunol.* **183**, 7919–7930 (2009).

4. Zhou, L., Chong, M. M. W. & Littman, D. R. Plasticity of CD4 + T Cell Lineage Differentiation. *Immunity* **30**, 646–655 (2009).

5. Hirota, K., Turner, J.-E., Villa, M., Duarte, J. H., Demengeot, J., Steinmetz, O. M. & Stockinger, B. Plasticity of TH17 cells in Peyer's patches is responsible for the induction of T cell–dependent IgA responses. *Nat. Immunol.* **14**, 372–380 (2013).

6. Nakayamada, S., Takahashi, H., Kanno, Y. & O' Shea, J. J. Helper T cell diversity and plasticity. *Curr. Opin. Immunol.* **24**, 297–302 (2012).

7. Cannons, J. L., Lu, K. T. & Schwartzberg, P. L. T follicular helper cell diversity and plasticity. *Trends Immunol.* **34**, 200–207 (2013).

8. Yamane, H. & Paul, W. E. Memory CD4+ T Cells: fate determination, positive feedback and plasticity. *Cell. Mol. Life Sci.* **69**, 1577–1583 (2012).

9. MacLeod, M. K. L., Kappler, J. W. & Marrack, P. Memory CD4 T cells: generation, reactivation and re-assignment. *Immunology* **130**, 10–15 (2010).

10. Sallusto, F., Lenig, D., Förster, R., Lipp, M. & Lanzavecchia, A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nat. 1999 4026763* **402**, 34–38 (1999).

11. Chang, J. T., Wherry, E. J. & Goldrath, A. W. Molecular regulation of effector and memory T cell differentiation. *Nat. Immunol. Rev.* **15**, 1104–1115 (2014).

12. Mueller, S. N., Gebhardt, T., Carbone, F. R. & Heath, W. R. Memory T Cell Subsets, Migration Patterns, and Tissue Residence. *Annu. Rev. Immunol* **31**, 137–61 (2013).

13. Carbone, F. R., Mackay, L. K., Heath, W. R. & Gebhardt, T. Distinct resident and recirculating memory T cell subsets in non-lymphoid tissues. *Curr. Opin. Immunol.* **25**, 329–333 (2013).

14. Carbone, F. R. Tissue-Resident Memory T Cells and Fixed Immune Surveillance in Nonlymphoid Organs. *J. Immunol.* **195**, 17–22 (2015).

15. Gebhardt, T., Whitney, P. G., Zaid, A., Mackay, L. K., Brooks, A. G., Heath, W. R., Carbone, F. R. & Mueller, S. N. Different patterns of peripheral migration by memory CD4+ and CD8+ T cells. *Nature* **477**, 216–219 (2011).

16. Gebhardt, T., Mueller, S. N., Heath, W. R. & Carbone, F. R. Peripheral tissue surveillance and residency by memory T cells. *Trends Immunol.* **34**, 27–32 (2013).

17. Gebhardt, T., Palendira, U., Tscharke, D. C. & Bedoui, S. Tissue-resident memory T cells in tissue homeostasis, persistent infection, and cancer surveillance. *Immunol. Rev.* **283**, 54–76 (2018).

18. Turner, D. L. & Farber, D. L. Mucosal resident memory CD4 T cells in protection and immunopathology. *Frontiers in Immunology* 1–10 (2014) doi:10.3389/fimmu.2014.00331.

19. Kaech, S. M., Tan, J. T., Wherry, E. J., Konieczny, B. T., Surh, C. D. & Ahmed, R. Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nat. Immunol.* **4**, 1191–1198 (2003).

20. Beyersdorf, N., Ding, X., Tietze, J. K. & Hanke, T. Characterization of mouse CD4 T cell subsets defined by expression of KLRG1. *Eur. J. Immunol.* **37**, 3445–3454 (2007).

21. Kiazyk, S. A. K. & Fowke, K. R. Loss of CD127 expression links immune activation and CD4 + T cell loss in HIV infection. *Trends Microbiol.* **16**, 567–573 (2008).

22. Harrington, L. E., Janowski, K. M., Oliver, J. R., Zajac, A. J. & Weaver, C. T. Memory CD4 T cells emerge from effector T-cell progenitors. *Nature* **452**, 356–360 (2008).

23. Pepper, M., Linehan, J. L., Pagán, A. J., Zell, T., Dileepan, T., Cleary, P. P. & Jenkins, M. K. Different routes of bacterial infection induce long-lived TH1 memory cells and short-lived TH17 cells. *Nat. Immunol.* **11**, 83–89 (2010).

24. Marshall, H. D., Chandele, A., Jung, Y. W., Meng, H., Poholek, A. C., Parish, I. A., Rutishauser, R., Cui, W., Kleinstein, S. H., Craft, J. & Kaech, S. M. Differential Expression of Ly6C and T-bet Distinguish Effector and Memory Th1 CD4+ Cell Properties during Viral Infection. *Immunity* **35**, 633–646 (2011).

25. Choi, Y. S., Yang, J. A., Yusuf, I., Johnston, R. J., Greenbaum, J., Peters, B. & Crotty, S. Bcl6 Expressing Follicular Helper CD4 T cells are Fate Committed Early and Have the Capacity to Form Memory. *J. Immunol.* **190**, 4014–4026 (2013).

26. Tubo, N. J., Fife, B. T., Pagan, A. J., Kotov, D. I., Goldberg, M. F. & Jenkins, M. K. Most microbe-specific naïve CD4+ T cells produce memory cells during infection. *Sci. Immunol.* **351**, 511–514 (2016).

27. Hale, J. S., Youngblood, B., Latner, D. R., Ur, A., Mohammed, R., Ye, L., Akondy, R. S.,

Wu, T., Iyer, S. S., Ahmed, R., Mohammed, A. U. R., Ye, L., Akondy, R. S., Wu, T., Iyer, S. S. & Ahmed, R. 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**, 805–817 (2013).

28. Lüthje, K., Kallies, A., Shimohakamada, Y., Belz, G. T., Light, A., Tarlinton, D. M. & Nutt, S. L. The development and fate of follicular helper T cells defined by an IL-21 reporter mouse. *Nat. Immunol.* **13**, 491–498 (2012).

29. Liu, X., Chen, X., Zhong, B., Wang, A., Wang, X., Chu, F., Nurieva, R. I., Yan, X., Chen, P., Van Der Flier, L. G., Nakatsukasa, H., Neelapu, S. S., Chen, W., Clevers, H., Tian, Q., Qi, H., Wei, L. & Dong, C. Transcription factor achaete-scute homologue 2 initiates follicular T-helper-cell development. *Nature* **507**, 513–518 (2014).

30. Pepper, M., Pagán, A. J., Igyá, B. Z., Taylor, J. J., Jenkins, M. K., Igyártó, B. Z., Taylor, J. J. & Jenkins, M. K. Opposing Signals from the Bcl6 Transcription Factor and the Interleukin-2 Receptor Generate T Helper 1 Central and Effector Memory Cells. *Immunity* **35**, 583–595 (2011).

31. Ciucci, T., Vacchio, M. S., Gao, Y., Ardori, F. T., Candia, J., Mehta, M., Zhao, Y., Tran, B., Pepper, M., Tessarollo, L., McGavern, D. B. & Bosselut, R. The Emergence and Functional Fitness of Memory CD4+ T Cells Require the Transcription Factor Thpok. *Immunity* **50**, 91–105 (2019).

32. Endo, Y., Hirahara, K., Yagi, R., Tumes, D. J. & Nakayama, T. Pathogenic memory type Th2 cells in allergic inflammation. *Trends Immunol.* **35**, 69–78 (2014).

33. Nakayama, T. & Yamashita, M. Critical role of the Polycomb and Trithorax complexes in the maintenance of CD4 T cell memory. *Semin. Immunol.* **21**, 78–83 (2009).

34. Hegazy, A. N., Peine, M., Helmstetter, C., Panse, I., Fröhlich, A., Bergthaler, A., Flatz, L., Pinschewer, D. D., Radbruch, A. & Löhning, M. Interferons Direct Th2 Cell Reprogramming to Generate a Stable GATA-3+T-bet+ Cell Subset with Combined Th2 and Th1 Cell Functions. *Immunity* **32**, 116–128 (2010).

35. Endo, Y., Iwamura, C., Kuwahara, M., Suzuki, A., Sugaya, K., Tumes, D. J., Tokoyoda, K., Hosokawa, H., Yamashita, M. & Nakayama, T. Eomesodermin Controls Interleukin-5 Production in Memory T Helper 2 Cells through Inhibition of Activity of the Transcription Factor GATA3. *Immunity* **35**, 733–745 (2011).

36. Hondowicz, B. D., An, D., Schenkel, J. M., Altemeier, W. A., Masopust, D., Kim, K. S., Steach, H. R., Krishnamurty, A. T., Keitany, G. J., Garza, E. N., Fraser, K. A., Moon, J. J. & Pepper, M. Interleukin-2-Dependent Allergen-Specific Tissue-Resident Memory Cells Drive Asthma. *Immunity* **44**, 155–166 (2016).

37. Mcgeachy, M. J. Th17 memory cells: live long and proliferate. *J. Leukoc. Biol.* **94**, (2013).

38. Zúñiga, L. A., Jain, R., Haines, C. & Cua, D. J. Th17 cell development: from the cradle to

the grave. *Immunol. Rev.* **252**, 78–88 (2013).

39. Muranski, P. et al. Th17 Cells Are Long Lived and Retain a Stem Cell-like Molecular Signature. *Immunity* **35**, 972–985 (2011).

40. Chen, Y., Chauhan, S. K., Tan, X. & Dana, R. Interleukin-7 and -15 maintain pathogenic memory Th17 cells in autoimmunity. *J. Autoimmun.* **77**, 96–103 (2017).

41. Pepper, M. & Jenkins, M. K. Origins of CD4+ effector and central memory T cells. *Nat. Immunol.* **12**, 467–471 (2011).

42. Dhume, K. & McKinstry, K. K. Early programming and late-acting checkpoints governing the development of CD4 T cell memory. *Immunology* (2018) doi:10.1111/imm.12942.

43. Gasper, D. J., Tejera, M. M. & Suresh, M. CD4 T-Cell Memory Generation and Maintenance. *Crit. Rev. ™ Immunol.* **34**, 121–146 (2014).

44. Snook, J. P., Kim, C. & Williams, M. A. *Supplementary Materials for TCR signal strength controls the differentiation of CD4 + effector and memory T cells The PDF file includes*. *Sci. Immunol* vol. 3 http://immunology.sciencemag.org/ (2018).

45. Mckinstry, K. K., Strutt, T. M., Bautista, B., Zhang, W., Kuang, Y., Cooper, A. M. & Swain, S. L. Effector CD4 T-cell transition to memory requires late cognate interactions that induce autocrine IL-2. *Nat. Commun.* **5**, (2014).

46. DiToro, D., Winstead, C. J., Pham, D., Witte, S., Andargachew, R., Singer, J. R., Wilson, C. G., Zindl, C. L., Luther, R. J., Silberger, D. J., Weaver, B. T., Kolawole, E. M., Martinez, R. J., Turner, H., Hatton, R. D., Moon, J. J., Evavold, B. D. & Weaver, C. T. Differential IL-2 expression defines developmental fates of follicular versus nonfollicular helper T cells. *Science (80-.).* **361**, (2018).

47. Shakya, A., Goren, A., Shalek, A., German, C. N., Snook, J., Kuchroo, V. K., Yosef, N., Chan, C. R., Regev, A., Williams, M. A. & Tantin, D. Oct1 and OCA-B are selectively required for CD4 memory T cell function. *J. Exp. Med.* **212**, 2115–2131 (2015).

48. Wilk, M. M. & Mills, K. H. G. CD4 TRM Cells Following Infection and Immunization: Implications for More effective vaccine Design. *Front. Immunol.* **9**, 1–8 (2018).

49. Teijaro, J. R., Turner, D., Pham, Q., John Wherry, J. E., Lefrançois, L. & Farber, D. L. Tissue-Retentive Lung Memory CD4 T Cells Mediate Optimal Protection to Respiratory Virus Infection. *J. Immunol.* **187**, 5510–5514 (2011).

50. Schenkel, J. M., Fraser, K. A., Vezys, V. & Masopust, D. Sensing & alarm function of resident memory CD8+ T cells. *Nat. Immunol.* **14**, 509–513 (2013).

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

52. Hofmann, M., Brinkmann, V. & Zerwes, H.-G. FTY720 preferentially depletes naive T cells from peripheral and lymphoid organs. *Int. Immunopharmacol.* **6**, 1902–1910 (2006).

53. Schenkel, J. M. & Masopust, D. Tissue-Resident Memory T Cells. *Immunity* **41**, 886–897 (2014).

54. Hondowicz, B. D., Kim, K. S., Ruterbusch, M. J., Keitany, G. J. & Pepper, M. 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**, 80–86 (2018).

55. Turner, D. L., Goldklang, M., Cvetkovski, F., Paik, D., Trischler, J., Barahona, J., Cao, M., Dave, R., Tanna, N., D'Armiento, J. M. & Farber, D. L. Biased Generation and In Situ Activation of Lung Tissue-Resident Memory CD4 T Cells in the Pathogenesis of Allergic Asthma. *J. Immunol.* **21**, 0–0 (2018).

56. Strutt, T., Dhume, K., Finn, C., Hwang, J., Castonguay, C., Swain, S. & Mckinstry, K. IL-15 supports the generation of protective lung-resident memory CD4 T cells. *Mucosal Immunol.* (2017) doi:10.1038/mi.2017.101.

57. Yeon, S., Halim, L., Chandele, A., Perry, C. J., Hoon Kim, S., Kim, S.-U., Byun, Y., Hong Yuk, S., Kaech, S. M. & Woo Jung, Y. IL-7 plays a critical role for the homeostasis of allergen-specific memory CD4 T cells in the lung and airways. *Sci. Rep.* **7**, (2017).

58. Smith, N., Wasserman, G. A., Coleman, F. T., Hilliard, K., Yamamoto, K., Lipsitz, E., Malley, R., Dooms, H., Jones, M. R., Quinton, L. J. & Mizgerd, J. Regionally compartmentalized resident memory T cells mediate naturally acquired protection against pneumococcal pneumonia. *Mucosal Immunol.* **11**, (2017).

59. Wilk, M. M., Misiak, A., McManus, R. M., Allen, A. C., Lynch, M. A. & Mills, K. H. G. Lung CD4 Tissue-Resident Memory T Cells Mediate Adaptive Immunity Induced by Previous Infection of Mice with Bordetella pertussis. *J. Immunol.* **199**, 233–243 (2017).

60. Sakai, S., Kauffman, K. D., Schenkel, J. M., McBerry, C. C., Mayer-Barber, K. D., Masopust, D. & Barber, D. L. Control of Mycobacterium tuberculosis Infection by a Subset of Lung Parenchyma–Homing CD4 T Cells. *J. Immunol.* **192**, 2965–2969 (2014).

61. Chapman, T. J. & Topham, D. J. 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**, 3841–3849 (2010).

62. Steinfelder, S., Rausch, S., Michael, D., Kühl, A. A. & Hartmann, S. Intestinal helminth infection induces highly functional resident memory CD4 + T cells in mice. *Eur. J. Immunol* **47**, 353–363 (2017).

63. Faria, A. M. C., Reis, B. S. & Mucida, D. Tissue adaptation: Implications for gut immunity and tolerance. *JEM* 1211–1226 (2017).

64. Romagnoli, P., Fu, H., Qiu, Z., Khairallah, C., Pham, Q., Puddington, L., Khanna, K.,

Lefranc¸ois, L. & Sheridan, B. Differentiation of distinct long-lived memory CD4 T cells in intestinal tissues after oral Listeria monocytogenes infection. *Mucosal Immunol.* **10**, 520–530 (2017).

65. Glennie, N. D., Yeramilli, V. A., Beiting, D. P., Volk, S. W., Weaver, C. T. & Scott, P. Skin-resident memory CD4+ T cells enhance protection against Leishmania major infection. *J. Exp. Med.* **212**, 1405–1414 (2015).

66. Glennie, N. D., Volk, S. W. & Scott, P. Skin-resident CD4 + T cells protect against Leishmania major by recruiting and activating inflammatory monocytes. *PLOS Pathog.* (2017) doi:10.1371/journal.ppat.1006349.

67. Shin, H. & Iwasaki, A. Skin TRM mediates distributed border patrol. *Cell Res.* **22**, 1325–1327 (2012).

68. Collins, N., Jiang, X., Zaid, A., Macleod, B. L., Li, J., Park, C. O., Haque, A., Bedoui, S., Heath, W. R., Mueller, S. N., Kupper, T. S., Gebhardt, T. & Carbone, F. R. Skin CD4+ memory T cells exhibit combined cluster-mediated retention and equilibration with the circulation. *Nat. Commun.* **7**, (2016).

69. Bagri, P., Anipindi, V. C., Nguyen, P. V, Vitali, D., Stampfli, M. R. & Kaushic, C. 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**, (2017).

70. Iijima, N. & Iwasaki, A. A local macrophage chemokine network sustains protective tissue-resident memory CD4 T cells. *Science (80-.).* **346**, 93–98 (2014).

71. Beura, L. K., Fares-Frederickson, N. J., Steinert, E. M., Scott, M. C., Thompson, E. A., Fraser, K. A., Schenkel, J. M., Vezys, V. & Masopust, D. CD4 + resident memory T cells dominate immunosurveillance and orchestrate local recall responses. *J. Exp. Med.* (2019) doi:10.1084/jem.20181365.

72. Gray, J. I., Westerhof, L. M. & MacLeod, M. K. L. The roles of resident, central and effector memory CD4 T cells in protective immunity following infection or vaccination. *Immunology* 1–8 (2018) doi:10.1111/imm.12929.

73. Thawer, S., Horsnell, W., Darby, M., Hoving, J., Dewals, B., Cutler, A., Lang, D. & Brombacher, F. Lung-resident CD4+ T cells are sufficient for IL-4Rα-dependent recall immunity to Nippostrongylus brasiliensis infection. *Mucosal Immunol.* **7**, 239–248 (2014).

74. Pruner, K. B. & Pepper, M. Local memory CD4 T cell niches in respiratory viral infection. *J. Exp. Med.* **218**, (2021).

75. Gebhardt, T., Wakim, L. M., Eidsmo, L., Reading, P. C., Heath, W. R. & Carbone, F. R. Memory T cells in nonlymphoid tissue that provide enhanced local immunity during infection with herpes simplex virus. *Nat. Immunol.* **10**, 524–530 (2009).

76. Masopust, D., Vezys, V., John Wherry, E., Barber, D. L. & Ahmed, R. Cutting Edge: Gut Microenvironment Promotes Differentiation of a Unique Memory CD8 T Cell Population. *J Immunol* **176**, 2079–2083 (2006).

77. Milner, J. J. & Goldrath, A. W. Transcriptional programming of tissue-resident memory CD8 + T cells. *Curr. Opin. Immunol.* **51**, 162–169 (2018).

78. Shenoy, A. T., Wasserman, G. A., Arafa, E. I., Wooten, A. K., Smith, N. M. S., Martin, I. M. C., Jones, M. R., Quinton, L. J. & Mizgerd, J. P. Lung CD4+ resident memory T cells remodel epithelial responses to accelerate neutrophil recruitment during pneumonia. *Mucosal Immunol. 2019 132* **13**, 334–343 (2019).

79. Son, Y. M., Cheon, I. S., Wu, Y., Li, C., Wang, Z., Gao, X., Chen, Y., Takahashi, Y., Fu, Y.-X., Dent, A. L., Kaplan, M. H., Taylor, J. J., Cui, W. & Sun, J. *Tissue-resident CD4 + T helper cells assist the development of protective respiratory B and CD8 + T cell memory responses*. *Sci. Immunol* vol. 6 http://immunology.sciencemag.org/ (2021).

80. Swarnalekha, N., Schreiner, D., Litzler, L. C., Iftikhar, S., Kirchmeier, D., Künzli, M., Son, Y. M., Sun, J., Moreira, E. A. & King, C. G. *T resident helper cells promote humoral responses in the lung*. *Sci. Immunol* vol. 6 http://immunology.sciencemag.org/ (2021).

81. Schattgen, S. A. & Thomas, P. G. TRH cells, helpers making an impact in their local community. *Sci. Immunol.* **6**, 2886 (2021).

82. Rahimi, R. A., Nepal, K., Cetinbas, M., Sadreyev, R. I. & Luster, A. D. Distinct functions of tissue-resident and circulating memory Th2 cells in allergic airway disease. *J. Exp. Med.* **217**, (2020).

83. Iijima, N. & Iwasaki, A. Tissue instruction for migration and retention of TRM cells. *Trends in Immunology* (2015) doi:10.1016/j.it.2015.07.002.

84. Shin, H. & Iwasaki, A. A vaccine strategy that protects against genital herpes by establishing local memory T cells. *Nature* **491**, (2012).

85. Reis, B. S., Rogoz, A., Costa-Pinto, F. A., Taniuchi, I. & Mucida, D. Mutual expression of the transcription factors Runx3 and ThPOK regulates intestinal CD4+ T cell immunity. *Nat. Immunol.* **14**, 271–280 (2013).

86. Reis, B. S., Hoytema Van Konijnenburg, D. P., Grivennikov, S. I. & Mucida, D. Transcription Factor T-bet Regulates Intraepithelial Lymphocyte Functional Maturation. *Immunity* **41**, 244–256 (2014).

87. Ross, S. H. & Cantrell, D. A. Signaling and Function of Interleukin-2 in T Lymphocytes. *Annu. Rev. Immunol.* (2018) doi:10.1146/annurev-immunol-042617-053352.

88. Christo, S. N. et al. Discrete tissue microenvironments instruct diversity in resident memory T cell function and plasticity. *Nat. Immunol. 2021* 1–12 (2021) doi:10.1038/s41590-021-01004-1.

89. Yu, B., Zhang, K., Milner, J. J., Toma, C., Chen, R., Scott-Browne, J. P., Pereira, R. M., Crotty, S., Chang, J. T., Pipkin, M. E., Wang, W. & Goldrath, A. W. Epigenetic landscapes reveal transcription factors that regulate CD8+ T cell differentiation. *Nat. Immunol.* **18**, 573–582 (2017).

90. Kakaradov, B., Arsenio, J., Widjaja, C. E., He, Z., Aigner, S., Metz, P. J., Yu, B., Wehrens, E. J., Lopez, J., Kim, S. H., Zuniga, E. I., Goldrath, A. W., Chang, J. T. & Yeo, G. W. Early transcriptional and epigenetic regulation of CD8+ T cell differentiation revealed by single-cell RNA sequencing. *Nat. Immunol.* **18**, 422–432 (2017).

91. Van Wijk, F. & Cheroutre, H. Mucosal T cells in gut homeostasis and inflammation. *Expert Review of Clinical Immunology* vol. 6 559–566 (2010).

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

93. Amezcua Vesely, M. C. et al. Effector TH17 Cells Give Rise to Long-Lived TRM Cells that Are Essential for an Immediate Response against Bacterial Infection. *Cell* **178**, 1176-1188.e15 (2019).

94. Peng, C. & Jameson, S. C. The relationship between CD4+ follicular helper T cells and CD8+ resident memory T cells: Sisters or distant cousins? *International Immunology* vol. 32 583–587 (2020).

95. Mueller, S. N. & Mackay, L. K. Tissue-resident memory T cells: Local specialists in immune defence. *Nature Reviews Immunology* vol. 16 79–89 (2016).

96. Walsh, D. A., Borges da Silva, H., Beura, L. K., Peng, C., Hamilton, S. E., Masopust, D. & Jameson, S. C. The Functional Requirement for CD69 in Establishment of Resident Memory CD8 + T Cells Varies with Tissue Location . *J. Immunol.* **203**, 946–955 (2019).

97. Skon, C. N., Lee, J. Y., Anderson, K. G., Masopust, D., Hogquist, K. A. & Jameson, S. C. Transcriptional downregulation of S1pr1 is required for the establishment of resident memory CD8+ T cells. *Nat. Immunol.* **14**, 1285–1293 (2013).

98. Lee, J. Y., Skon, C. N., Lee, Y. J., Oh, S., Taylor, J. J., Malhotra, D., Jenkins, M. K., Rosenfeld, M. G., Hogquist, K. A. & Jameson, S. C. The Transcription Factor KLF2 Restrains CD4+ T Follicular Helper Cell Differentiation. *Immunity* **42**, 252–264 (2015).

99. Weber, J. P., Fuhrmann, F., Feist, R. K., Lahmann, A., Al Baz, M. S., Gentz, L. J., Vu Van, D., Mages, H. W., Haftmann, C., Riedel, R., Grün, J. R., Schuh, W., Kroczek, R. A., Radbruch, A., Mashreghi, M. F. & Hutloff, A. ICOS maintains the T follicular helper cell phenotype by down-regulating krüppel-like factor 2. *J. Exp. Med.* **212**, 217–233 (2015).

100. Stone, E. L., Pepper, M., Katayama, C. D., Kerdiles, Y. M., Lai, C. Y., Emslie, E., Lin, Y. C., Yang, E., Goldrath, A. W., Li, M. O., Cantrell, D. A. & Hedrick, S. M. ICOS coreceptor signaling inactivates the transcription factor FOXO1 to promote Tfh cell differentiation. *Immunity* **42**, 239–251 (2015).

101. Mackay, L. K., Rahimpour, A., Ma, J. Z., Collins, N., Stock, A. T., Hafon, M.-L., Vega-Ramos, J., Lauzurica, P., Mueller, S. N., Stefanovic, T., Tscharke, D. C., Heath, W. R., Inouye, M., Carbone, F. R. & Gebhardt, T. The developmental pathway for CD103+CD8+ tissue-resident memory T cells of skin. *Nat. Immunol.* **14**, (2013).

102. Kumar, B. V., Ma, W., Miron, M., Granot, T., Guyer, R. S., Carpenter, D. J., Senda, T., Sun, X., Ho, S. H., Lerner, H., Friedman, A. L., Shen, Y. & Farber, D. L. Human Tissue-Resident Memory T Cells Are Defined by Core Transcriptional and Functional Signatures in Lymphoid and Mucosal Sites. *Cell Rep.* **20**, 2921–2934 (2017).

103. Johnston, R. J., Poholek, A. C., DiToro, D., Yusuf, I., Eto, D., Barnett, B., Dent, A. L., Craft, J. & Crotty, S. Bcl6 and Blimp-1 Are Reciprocal and Antagonistic Regulators of T Follicular Helper Cell Differentiation. *Science (80-.).* **325**, (2009).

104. Crotty, S., Johnston, R. J. & Schoenberger, S. P. Effectors and memories: Bcl-6 and Blimp-1 in T and B lymphocyte differentiation. *Nature Immunology* vol. 11 114–120 (2010).

105. Yu, D., Rao, S., Tsai, L. M., Lee, S. K., He, Y., Sutcliffe, E. L., Srivastava, M., Linterman, M., Zheng, L., Simpson, N., Ellyard, J. I., Parish, I. A., Ma, C. S., Li, Q. J., Parish, C. R., Mackay, C. R. & Vinuesa, C. G. The Transcriptional Repressor Bcl-6 Directs T Follicular Helper Cell Lineage Commitment. *Immunity* **31**, 457–468 (2009).

106. Ichii, H., Sakamoto, A., Hatano, M., Okada, S., Toyama, H., Taki, S., Arima, M., Kuroda, Y. & Tokuhisa, T. Role for BcL-6 in the generation and maintenance of memory CD8+ T cells. *Nat. Immunol.* **3**, 558–563 (2002).

107. Ichii, H., Sakamoto, A., Kuroda, Y. & Tokuhisa, T. Bcl6 Acts as an Amplifier for the Generation and Proliferative Capacity of Central Memory CD8 + T Cells . *J. Immunol.* **173**, 883–891 (2004).

108. Castro, A. G., Hauser, T. M., Cocks, B. G., Abrams, J., Zurawski, S., Churakova, T., Zonin, F., Robinson, D., Tangye, S. G., Aversa, G., Nichols, K. E., De Vries, J. E., Lanier, L. L. & O'Garra, A. Molecular and functional characterization of mouse signaling lymphocytic activation molecule (SLAM): Differential expression and responsiveness in Th1 and Th2 cells. *J. Immunol.* **163**, 5860–5870 (1999).

109. Breitfeld, D., Ohl, L., Kremmer, E., Ellwart, J., Sallusto, F., Lipp, M. & Förster, R. Follicular B helper T cells express CXC chemokine receptor 5, localize to B cell follicles, and support immunoglobulin production. *J. Exp. Med.* **192**, 1545–1551 (2000).

110. Mackay, L., Kallies, A., Carbone, F. & Gisbergen, K. Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science (80-.).* **352**, 459–463 (2016).

111. Hendriks, J., Gravestein, L. A., Tesselaar, K., Van Lier, R. A. W., Schumacher, T. N. M. & Borst, J. CD27 is required for generation and long-term maintenance of T cell immunity. *Nat. Immunol.* **1**, 433–440 (2000).

112. Fritsch, R. D., Shen, X., Sims, G. P., Hathcock, K. S., Hodes, R. J. & Lipsky, P. E. Stepwise Differentiation of CD4 Memory T Cells Defined by Expression of CCR7 and CD27. *J. Immunol.* **175**, 6489–6497 (2005).

113. Schiött, Å., Lindstedt, M., Johansson-Lindbom, B., Roggen, E. & Borrebaeck, C. A. K. CD27- CD4+ memory T cells define a differentiated memory population at both the functional and transcriptional levels. *Immunology* **113**, 363–370 (2004).

114. West, B. R., Hastie, K. M. & Saphire, E. O. Structure of the LCMV nucleoprotein provides a template for understanding arenavirus replication and immunosuppression. *Acta Crystallogr.* **70**, 1764–9 (2014).

115. Nance, J. P., Bélanger, S., Johnston, R. J., Takemori, T. & Crotty, S. Cutting Edge: T Follicular Helper Cell Differentiation Is Defective in the Absence of Bcl6 BTB Repressor Domain Function. *DCSupplemental J. Immunol. by guest April* **194**, 5599–5603 (2015).

116. Masopust, D., Choo, D., Wherry, J., Newell, K. & Ahmed, R. Dynamic T cell migration program provides resident memory within intestinal epithelium. *J. Exp. Med* **207**, 553–564 (2010).

117. Marshall, N. B. & Swain, S. L. Cytotoxic CD4 T cells in antiviral immunity. *Journal of Biomedicine and Biotechnology* vol. 2011 (2011).

118. Marshall, N. B., Vong, A. M., Devarajan, P., Brauner, M. D., Kuang, Y., Nayar, R., Schutten, E. A., Castonguay, C. H., Berg, L. J., Nutt, S. L. & Swain, S. L. NKG2C/E Marks the Unique Cytotoxic CD4 T Cell Subset, ThCTL, Generated by Influenza Infection. *J. Immunol.* **198**, 1142–1155 (2017).

119. Milner, J. J., Toma, C., Yu, B., Zhang, K., Omilusik, K., Phan, A. T., Wang, D., Getzler, A. J., Nguyen, T., Crotty, S., Wang, W., Pipkin, M. E. & Goldrath, A. W. Runx3 programs CD8+ T cell residency in non-lymphoid tissues and tumours. *Nat. Publ. Gr.* **552**, (2017).

120. Stark, R., Wesselink, T. H., Behr, F. M., Kragten, N. A. M., Arens, R., Koch-Nolte, F., van Gisbergen, K. P. J. M., van Lier, R. A. W., M Kragten, N. A., Arens, R., Koch-Nolte, F., J M van Gisbergen, K. P. & W van Lier, R. A. TRM maintenance is regulated by tissue damage via P2RX7. *Sci. Immunol.* **3**, 1022 (2018).

121. Botta, D., Fuller, M. J., Marquez-Lago, T. T., Bachus, H., Bradley, J. E., Weinmann, A. S., Zajac, A. J., Randall, T. D., Lund1, F. E., León, B. & Ballesteros-Tato, A. Dynamic regulation of T follicular regulatory cell responses by interleukin 2 during influenza infection. *Nat. Immunol.* **18**, (2017).

122. Wing, J. B., Kitagawa, Y., Locci, M., Hume, H., Tay, C., Morita, T., Kidani, Y., Matsuda, K., Inoue, T., Kurosaki, T., Crotty, S., Coban, C., Ohkura, N., Sakaguchi, S., Graca, L. & Hori, S. A distinct subpopulation of CD25 T-follicular regulatory cells localizes in the germinal centers. doi:10.1073/pnas.1705551114.

123. Miyazaki, M., Miyazaki, K., Chen, S., Itoi, M., Miller, M., Lu, L. F., Varki, N., Chang, A. N., Broide, D. H. & Murre, C. Id2 and Id3 maintain the regulatory T cell pool to suppress inflammatory disease. *Nat. Immunol.* **15**, 767–776 (2014).

124. Lin, Y.-Y., Jones-Mason, M. E., Inoue, M., Lasorella, A., Iavarone, A., Li, Q.-J., Shinohara, M. L. & Zhuang, Y. Transcriptional Regulator Id2 Is Required for the CD4 T Cell Immune Response in the Development of Experimental Autoimmune Encephalomyelitis. *J. Immunol.* **189**, 1400–1405 (2012).

125. Shaw, L. A., Belanger, S., Omilusik, K. D., Cho, S., Scott-Browne, J. P., Nance, J. P., Goulding, J., Lasorella, A., Lu, L.-F., Crotty, S. & Goldrath, A. W. Id2 reinforces TH1 differentiation and inhibits E2A to repress TFH differentiation. *Nat. Immunol.* (2016) doi:10.1038/ni.3461.

126. Omilusik, K., Shaw, L. & Goldrath, A. Remembering one's ID/E-ntity: E/ID protein regulation of T cell memory. *Curr. Opin. Immunol.* (2013).

127. Milner, J. J., Toma, C., He, Z., Kurd, N. S., Nguyen, Q. P., McDonald, B., Quezada, L., Widjaja, C. E., Witherden, D. A., Crowl, J. T., Shaw, L. A., Yeo, G. W., Chang, J. T., Omilusik, K. D. & Goldrath, A. W. Heterogenous Populations of Tissue-Resident CD8+ T Cells Are Generated in Response to Infection and Malignancy. *Immunity* **52**, 808-824.e7 (2020).

128. Zundler, S., Becker, E., Spocinska, M., Slawik, M., Parga-Vidal, L., Stark, R., Wiendl, M., Atreya, R., Rath, T., Leppkes, M., Hildner, K., López-Posadas, R., Lukassen, S., Ekici, A. B., Neufert, C., Atreya, I., van Gisbergen, K. P. J. M. & Neurath, M. F. Hobit- and Blimp-1-driven CD4 + tissue-resident memory T cells control chronic intestinal inflammation. *Nat. Immunol.* **20**, 288–300 (2019).

129. Choi, J., Diao, H., Faliti, C. E., Truong, J., Rossi, M., Bélanger, S., Yu, B., Goldrath, A. W., Pipkin, M. E. & Crotty, S. Bcl-6 is the nexus transcription factor of T follicular helper cells via repressor-of-repressor circuits. *Nat. Immunol.* **21**, 777–789 (2020).

130. Paik, D. H. & Farber, D. L. Anti-viral protective capacity of tissue resident memory T cells. *Curr. Opin. Virol.* **46**, 20–26 (2021).

131. Bechtel, T. J., Reyes-Robles, T., Fadeyi, O. O. & Oslund, R. C. Strategies for monitoring cell–cell interactions. *Nat. Chem. Biol. 2021 176* **17**, 641–652 (2021).

132. Liu, Z., Yuan, J., Lasorella, A., Iavarone, A., Bruce, J. N., Canoll, P. & Sims, P. A. Integrating single-cell RNA-seq and imaging with SCOPE-seq2. *Sci. Reports 2020 101* **10**, 1–15 (2020).

133. Rosato, P. C., Wijeyesinghe, S., Stolley, J. M. & Masopust, D. Integrating resident memory into T cell differentiation models. *Curr. Opin. Immunol.* **63**, 35–42 (2020).

134. Lavelle, E. C. & Ward, R. W. Mucosal vaccines — fortifying the frontiers. *Nat. Rev. Immunol. 2021* 1–15 (2021) doi:10.1038/s41577-021-00583-2.

135. Dhodapkar, M. V. & Dhodapkar, K. M. Tissue-resident memory-like T cells in tumor immunity: Clinical implications. *Semin. Immunol.* **49**, 101415 (2020).

136. Okla, K., Farber, D. L. & Zou, W. Tissue-resident memory T cells in tumor immunity and immunotherapy. *J. Exp. Med.* **218**, (2021).

137. Sasson, S. C., Gordon, C. L., Christo, S. N., Klenerman, P. & Mackay, L. K. Local heroes or villains: tissue-resident memory T cells in human health and disease. *Cell. Mol. Immunol. 2020 172* **17**, 113–122 (2020).

138. Raphael, I., Joern, R. R. & Forsthuber, T. G. Memory CD4+ T Cells in Immunity and Autoimmune Diseases. *Cells* **9**, 531 (2020).

139. Miyazaki, M., Rivera, R. R., Miyazaki, K., Lin, Y. C., Agata, Y. & Murre, C. The opposing roles of the transcription factor E2A and its antagonist Id3 that orchestrate and enforce the naive fate of T cells. *Nat. Immunol.* **12**, 992–1001 (2011).

140. Yang, C. Y., Best, J. A., Knell, J., Yang, E., Sheridan, A. D., Jesionek, A. K., Li, H. S., Rivera, R. R., Lind, K. C., D'Cruz, L. M., Watowich, S. S., Murre, C. & Goldrath, A. W. The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8 + T cell subsets. *Nat. Immunol.* **12**, 1221–1229 (2011).

141. Niola, F., Zhao, X., Singh, D., Castano, A., Sullivan, R., Lauria, M., Nam, H. S., Zhuang, Y., Benezra, R., Di Bernardo, D., Iavarone, A. & Lasorella, A. Id proteins synchronize stemness and anchorage to the niche of neural stem cells. *Nat. Cell Biol.* **14**, 477–487 (2012).

142. Oxenius, A., Bachmann, M. F., Zinkernagel, R. M. & Hengartner, H. *Virus-specific MHC class II-restricted TCR-transgenic mice: effects on humoral and cellular immune responses after viral infection*. doi:10.1002/(SICI)1521-4141(199801)28:01.

143. Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. & Mesirov, J. P. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).

144. Mootha, V. K. et al. *PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes*. *NATURE GENETICS VOLUME* vol. 34 http://www.nature.com/naturegenetics (2003).

145. van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A. J., Burdziak, C., Moon, K. R., Chaffer, C. L., Pattabiraman, D., Bierie, B., Mazutis, L., Wolf, G., Krishnaswamy, S. & Pe'er, D. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**, 716-729.e27 (2018).

146. Dunham, I. et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

147. Stark, R. & Brown, G. *DiffBind: Differential binding analysis of ChIP-Seq peak data*.

148. Ross-Innes, C. S., Stark, R., Teschendorff, A. E., Holmes, K. A., Ali, H. R., Dunning, M. J., Brown, G. D., Gojis, O., Ellis, I. O., Green, A. R., Ali, S., Chin, S. F., Palmieri, C., Caldas, C. & Carroll, J. S. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).