UNIVERSITY OF CALIFORNIA SAN DIEGO

Multi-layered epigenetic control of T cell fate decisions

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

in

Biology

by

Bingfei Yu

Committee in charge:

Professor Ananda Goldrath, Chair
Professor John Chang
Professor Stephen Hedrick
Professor Cornelis Murre
Professor Wei Wang

2018

The dissertation of Bingfei Yu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California San Diego

2018

DEDICATION

To my parents who have been giving me countless love, trust and support to make me who I am.

EPIGRAPH

Stay hungary. Stay foolish.

— Steve Jobs

quoted from the back cover of the 1974 edition of the Whole Earth Catalog

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge my mentor Ananda, the best mentor I would ever hope for. She is such an inspiring scientist and wonderful mentor. I would thank her for giving me a chance, an international girl with no background of immunology, who never touched a mouse, for taking time to understand authentic jokes to grow as an independent researcher with confidence, enthusiasm and the sense of humor for both science and life. I would thank my committee members: Professors Kees Murre, Stephen Hedrick, John Chang and Wei Wang for giving me so many valuable suggestions to keep me on the right track. I would also thank all Goldrath lab members past and present. They are not only my colleagues but also make me feel like I am surrounded by family members with laughter and love. I learned and enjoyed so much from exciting scientific discussions to funny pop culture education with them. I would also thank all my friends for being supportive, engaging and caring. I will remember every card game and Karaoke night. They made my graduate school life so much fun and memorable.

Chapter 2, is a reprint of the material as it appears in Nature Immunology 2017. Yu B, Zhang K, Milner JJ, Toma C, Chen R, Scott-Browne JP, Pereira RM, Crotty S, Chang JT, Pipkin ME, Wang W, Goldrath AW. Epigenetic landscapes reveal transcription factors that regulate CD8+ T cell differentiation. *Nature Immunology.* 2017 18(5):573-582

The dissertation author was a primary investigator and the first author of this paper.

Chapter 3, in part, is a subset of the material as it appears in Nature 2017. Milner JJ, Toma C*, Yu B*, Zhang K, Omilusik KD, Phan A, Wang DP, Getzler A, Crotty S, Wang W, Pipkin ME, Goldrath AW. Runx3 programs CD8+ T cell residency in non-lymphoid tissues and tumors. *Nature*. 2017 552(7684)

The dissertation author was a primary investigator and the co-second author of this paper.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Yu B, Goldrath AW.

The dissertation author was a primary investigator and the first author of this material.

VITA

2012-2018	Ph.D. in Biology, University of California San Diego
2009-2012	M.S in Molecular Biology and Biochemistry, Xiamen University
2005-2009	B.S in Biology, Xiamen University

PUBLICATIONS

Omilusik KD, Nadjsombati M, Shaw L, **Yu B**, Milner J, Goldrath AW. Sustained regulation of E-protein transcription factors by Id2 enforces terminal differentiation of effector CD8+ T cells. *Journal of Experimental Medicine*. 2018 215(3):773-783

Milner JJ, Toma C*, **Yu B***, Zhang K, Omilusik KD, Phan A, Wang DP, Getzler A, Crotty S, Wang W, Pipkin ME, Goldrath AW. Runx3 programs CD8+ T cell residency in non-lymphoid tissues and tumors. *Nature*. 2017 552(7684) * equally contributed

Yu B, Zhang K, Milner JJ, Toma C, Chen R, Scott-Browne JP, Pereira RM, Crotty S, Chang JT, Pipkin ME, Wang W, Goldrath AW. Epigenetic landscapes reveal transcription factors that regulate CD8+ T cell differentiation. *Nature Immunology*. 2017 18(5):573-582

Kakaradov B, Arsenio J, Widjaja CE, He Z, Aigner S, Metz PJ, **Yu B**, Wehrens EJ, Lopez J, Kim SH, Zuniga EI, Goldrath AW, Chang JT, Yeo GW. Early transcriptional and epigenetic regulation of CD8+ T cell differentiation revealed by single-cell RNA sequencing. *Nature Immunology*. 2017 18(4):422-432

Omilusik KD, Best JA, **Yu B**, Goossens S, Weidemann A, Nguyen JV, Seuntjens E, Stryjewska A, Zweier C, Roychoudhuri R, Gattinoni L, Bird LM, Higashi Y, Kondoh H, Huylebroeck D, Haigh J, AW Goldrath. Transcriptional repressor ZEB2 promotes terminal differentiation of CD8+ effector and memory T cell populations during infection. *Journal of Experimental Medicine*. 2015. 212(12): 2027-39

Ai N, Hu X, Ding F, **Yu B**, Wang H, Lu X, Zhang K, Li Y, Han A, Lin W, Liu R, Chen R. Signal-induced Brd4 release from chromatin is essential for its role transition from chromatin targeting to transcriptional regulation. *Nucleic Acids Res*. 2011. 39, 95929604

* indicates co-authorship

ABSTRACT OF THE DISSERTATION

Multi-layered epigenetic control of T cell fate decisions

by

Bingfei Yu

Doctor of Philosophy in Biology

University of California San Diego, 2018

Professor Ananda Goldrath, Chair

$CD8^+$ T cells are a central component of the adaptive immune system. Upon infection, a naive $CD8^+$ T cell will differentiate into a heterogeneous population of effector T cells composed of terminal-effector and memory-precursor $CD8^+$ T cells. Terminal-effector T cells rapidly decay after pathogens are eradicated while memory-precursor T cells survive during the contraction phase to become memory T cells, providing long-term protection from reinfection. Similar to the heterogeneity of effector T cells, memory T cells can also be divided into central-memory, effector-memory and tissue-resident memory subsets based on

trafficking, location, proliferation potential and cytotoxic function. These memory subsets collaborate together to enhance the pathogen clearance and vaccine efficacy. The differentiation of a naive $CD8^+$ T cell into a specific effector or memory subset is influenced by cell-extrinsic environmental signals and cell-intrinsic factors including transcription factors (TFs), epigenetic modification and chromatin organization. Considerable advances have been made to identify key TFs that regulate these T cell fate decisions. However, the TF-mediated transcriptional network responsible for specific subset differentiation and how epigenetic modifications and chromatin configuration modulate $CD8^+$ T cell fate determination is still largely unknown. To address these questions, we deciphered epigenetic landscapes in effector and memory $CD8^+$ T cells in response to bacterial infection by characterizing the genome-wide histone modification, chromatin accessibility and transcriptional program. Integrative analysis of epigenomics data showed that subset-specific enhancers established by key TFs foreshadow the specific lineage differentiation. To better identify cruicial TFs from multilayered epigenetic landscapes, we developed a webpage ranking-based algorithm (PageRank) to rank the importance of TFs from transcriptional network and identified a novel function of two TFs: YY1 and Nr3c1 to regulate terminal-effector and memory-precursor subset differentiation, respectively. By leveraging the PageRank analysis and chromatin accessibility data, we developed a computational screen to predict key TFs for tissue-resident memory T cell differentiation. Combining this approach with shRNA functional screen, we identified the role of Runx3 in programming tissue-residency signatures in non-lymphoid tissues and tumors. Finally, we discovered a novel role for the genome organizer CTCF in $CD8^+$ T cell fate decisions illustrating the impact of

chromatin organization on effector and memory T cell differentiation. Taken together, we uncovered a multi-layered regulation of chromatin state, accessibility and organization, that influences T cell fate decisions by fine-tuning transcriptional circuits. We further constructed a computational framework that integrates these high-dimensional data, facilitating identification of key transcriptional regulators and providing valuable biological insights.

Chapter 1

Introduction

1.1 The differentiation path of CD8$^+$ T cells in combating infectious and malignant diseases

The battle between the immune system and invading pathogens has lasted and evolved for many eons. Despite the development of vaccines, infectious diseases still remain the leading cause of death worldwide. The growing impact of HIV/AIDS epidemics, the emergence of multidrug-resistant pathogens, and the threat of lethal pathogens like Ebola and Zika viruses have all urged the development of better vaccines[1]. T cell-based vaccination has been a major focus for the prevention of infectious diseases such as HIV, Tuberculosis (TB) and malaria due to the inefficiency of the neutralizing antibody for protection[2]. CD8$^+$ T cells are a critical component in host protection from infectious and malignant diseases given the unique ability of effector CD8$^+$ T cells to kill pathogen-infected or malignant cells and memory CD8$^+$ T cells to protect against re-infection or tumor

growth. Thus, a thorough understanding of how $CD8^+$ effector and memory T cells are generated and maintained will greatly inform the development of vaccines and immunotherapy.

In response to pathogen infection, naive $CD8^+$ T cells undergo clonal expansion and differentiation establishing a heterogeneous population of pathogen-specific effector $CD8^+$ T cells. These effector $CD8^+$ T cells can produce effector molecules and cytokines such as IFNγ, TNFα and granzyme B to clear the invading intracellular pathogens. While the majority of these $CD8^+$ T cells die through apoptosis during the contraction phase of infection, a small fraction persists as memory cells, providing lasting protection from re-infection. This memory T cell population can "remember" the antigen and elicit a more rapid and robust immune response when re-encountering the same pathogen, which is the molecular basis of vaccine development[2]. Recent studies demonstrate that the commitment to the shorter-lived effector versus longer-lived memory $CD8^+$ T cell fates occurs early after infection, and the differential expression of cell-surface markers, such as killer lectin-like receptor (KLRG1) and interleukin-7 receptor (IL-7R) may be used to distinguish two effector $CD8^+$ T cell subsets with distinct memory potential: shorter-lived terminally-differentiated effector (TE, $KLRG1^{hi}IL\text{-}7R^{lo}$) and longer-lived memory precursor effector (MP, $KLRG1^{lo}IL\text{-}7R^{hi}$) $CD8^+$ T cells[3, 4]. Although both TE and MP $CD8^+$ T cell subsets exhibit similar effector functions, only the MP subset can survive long-term to form the heterogeneous memory $CD8^+$ T cell population[3]. Effective immune response for host protection from repetitive infection largely relies on the diversity and heterogeneity of memory T cells. Memory T cells can be typically segregated into three populations: cen-

tral memory (Tcm), effector memory (Tem) and tissue-resident memory (Trm) subsets[5, 6]. Tcm and Tem are circulating memory cells that ensure durable immunosurveillance. Tcm cells primarily locate in secondary lymphoid organs and Tem cells traffic in blood and recirculate between non-lymphoid and lymphoid organs. Tcm cells express high level of lymph node homing receptors CD62L and CCR7 and display a higher proliferation rate while Tem cells lack CD62L and CCR7 expression and display a higher cytotoxic ability[7]. The recently identified memory subset, Trm, primarily resides in non-lymphoid tissues including barrier (skin, lung and gut) and non-barrier (brain, liver and kidney) tissues[8]. Unlike Tem, which will recirculate after homing to non-lymphoid tissues, Trm cells permanently reside in non-lymphoid tissues without recirculation, providing the front-line protection against re-infection[8, 9]. A coordination of all three memory subsets is a key feature of protective immune response. When re-encountering pathogens, Trm cells are the first responders that can rapidly trigger the inflammatory response and produce alarm signals to attract circulating memory cells into the infected tissues[10]. Therefore, a better understanding of how specific memory subsets are formed and cooperate together to fight against infection will greatly impact the efficacy of vaccination.

1.2 Transcriptional regulation of $CD8^+$ T cell fate decisions

Upon infection, naive $CD8^+$ T cells undergo activation and proliferation to become a heterogeneous population of effector $CD8^+$ T cells that include TE

and MP subsets. The fate of a naive $CD8^+$ T cell is influenced by cell-extrinsic factors including precursor frequency, TCR strength and exposure to inflammatory cytokines[4, 11]. For instance, increased antigen and inflammatory cytokines such as IL-12 promote the formation of the TE subset[3, 4]. As well, cell-intrinsic factors including TFs, epigenetic landscapes and chromatin organization shape the fate of the differentiating T cells[4, 11]. Specifically, a number of TFs have been reported to promote the cell fate decision: T-bet, Zeb2, Prdm1 and Id2 are essential for the TE subset differentiation while Tcf7, Foxo1, Bcl6, Eomes and Id3 play critical roles in the MP subset development[12, 3, 4, 11]. Importantly, it is the extrinsic cues that are integrated by the differentiating T cells that ultimately regulate the cell-intrinsic TFs that direct the effector and memory subset specification. For example, IL-2 and IL-12 signaling induce Blimp1 and T-bet to promote TE formation and effector function[13]. TFs work hierarchically or in concert to direct to specify cell subset fates. For instance, Zeb2, which is recently reported to repress IL-7R expression in effector $CD8^+$ T cells, is a downstream target of T-bet[14, 15]. T-bet collaborates with Zeb2 to promote the TE subset differentiation while repressing the transcriptional program favoring the central memory $CD8^+$ T cell differentiation[14, 15]. Interestingly, many TFs are similarly expressed in both TE and MP subsets although they have been shown to regulate specific subset differentiation. This suggests other regulatory mechanisms in addition to TF expression contribute to cell-type specific differentiation. With the accumulating evidence of the transcriptional regulation of TE versus MP cell fate decision, it still remains unclear how the TF-mediated transcriptional network participates in effector cell subset specification and it is likely that a number of TFs controlling

this cell fate determination has yet to be elucidated.

As the heterogeneity of the memory pool is increasingly appreciated, accumulating studies have focused on the factors influencing the cell fate decision to become Tcm, Tem or Trm. A thorough understanding of the molecular mechanism underlying the memory subset specification is important for specific vaccine development: targeting Tcm may be beneficial for traditional vaccines against infecious diseases while targeting Trm may be beneficial for preventing infections at barrier sites. Extensive evidence revealed that TFs that regulate the TE versus MP subset differentiation are also involved in Tem versus Tcm cell fate decisions. For example, the loss of T-bet, Blimp1 or Id2 impairs the differentiation of Tem while Tcf7, Foxo1 and Id3 are essential for Tcm formation and maintenance[11]. Interestingly, emerging evidence shows that the formation and differentiation of Trm relies on a hybrid TFs mediated transcriptional program, requiring both the effector and memory associated TFs. T-bet and Eomes have both been shown to repress Trm differentiation, although T-bet and Eomes are important for differentiation of effector/TE and memory/MP populations, respectively[11, 9]. In contrast to the repressive role of T-bet, Blimp1 can promote the differentiation of Trm cells, although T-bet and Blimp1 are both effector-favored TFs[9, 13]. Taken together, there is still much to be learned about the transcriptional pathways and molecular drivers modulating the Trm differentiation, maintenance and function.

1.3 Epigenetic modulation of CD8$^+$ T cell differentiation

The organization of chromatin is revealed as a hierarchy of compaction and regulation levels: (1) The genomic DNA is wrapped in 200bp segments around the histone octamers to form nucleosomes and the interaction between DNA and histones is regulated by DNA methylation and histone modification; (2) The chromatin fibers containing nucleosomes are densely packaged into heterochromatin and euchromatin with distinct chromatin accessibility; (3) The topologically-associated domains (TADs) and lamina-associated domains (LADs) form specific chromatin territories in the nucleus[16]. Chromatin serves as a platform for storing the complex DNA information in a high-order organization. It also contributes to balancing the access of transcriptional machinery to key regulatory elements via DNA methylation and histone modification. These heritable changes to gene function without alterations in DNA sequence during cell division are referred to as epigenetic regulation.

DNA methylation primarily occurs at cytosine residues within CG dinucleotide (CpG)-dense regions. When CpG methylation occurs at promoter regions, it is always associated with transcriptional repression[17]. Upon LCMV infection, naive-associated genes acquire *de novo* DNA methylation in effector T cells. Subsequently, during the transition from effector to memory T cells, naive-associated genes are demethylated in memory T cells. This suggests that the epigenetic repression of naive-associated genes in effector T cells can be reversed in long-lived memory T cells while some key effector genes remain demethylated[18]. Impor-

tantly, the deficiency of *de novo* methyltransferase Dnmt3a in early effector T cells lead to decreased methylation of naive-associated genes, ultimately accelerating the memory T cell differentiation[18]. These studies provide novel insights into how epigenetic programs are modified at the DNA methylation level to modulate effector and memory $CD8^+$ T cell differentiation.

Histone modification primarily occurs at N-terminal histone tails including methylation, acetylation, phosphorylation and ubiquitination[19]. These post-translational modification of histones are critical epigenetic programs regulating the DNA accessibility to transcription factors and chromatin remodelers. It has been shown that certain combinations of histone modifications can mark key regulatory regions including enhancers and promoters with either active or repressive status. For example, H3K4 monomethylation (H3K4me1) is associated with enhancers whereas H3K4 trimethylation (H3K4me3) ususally marks promoters. H3K27 acetylation (H3K27ac) is associated with active transcription while H3K27 trimethylation (H3K27me3) is associated with transcriptional repression. Active enhancers are often associated with the combination of H3K4me1 and H3K27ac, positively correlating with active gene expression[20, 21, 22]. The presence of H3K4me3 and H3K27me3 at promoter regions represents a bivalent chromatin signature, which is often associated with transcriptionally poised genes[23, 24]. In naive T cells, the promoter regions of effector-associated genes display this bivalent signature. During effector $CD8^+$ T cell differentiation, the repressive mark H3K27me3 is lost but the level of H3K4me3 is maintained, concomitant with increased expression of effector-associated genes[23, 24]. Similarly, the memory-associated genes are also bivalently modifed at promoter regions in effector T

cells and resolve to a H3K4me3 only state when differentiating into memory T cells[25, 26]. Notably, conditional deletion of Ezh2, a histone methyltransferase responsible for H3K27me3, results in increased expression of memory-associated genes including *Tcf7* and *Foxo1* and decreased expansion of effector T cells, suggesting that epigenetic repression of memory-associated genes is critical to maintain effector T cell identity[25, 26].

In addition to gradual changes of histone modifications at promoter regions, the enhancer landscape marked by specific histone modifications often exhibit dynamic changes in a cell-type-specific manner. Accumulating evidence suggests that enhancers play a key role in fine-tuning gene expression via long-distance communication with promoter elements, providing better cellular specificity compared with promoters[27, 28, 29, 30, 31]. It is now well accepted that distinct enhancer landscapes in different cell types are bound and established by lineage-specific TFs to modulate cell-type-specific transcriptional programs. Therefore, global mapping of enhancer and promoter landscapes in $CD8^+$ T cell subsets upon infection is essential to understand how epigenetic programs are involved in $CD8^+$ T cell differentiation, and to inform the key TFs that establish specific enhancer and promoter repertoires responsible for specific T cell subset differentiation.

1.4 Outline

In this dissertation, I present the results in three chapters, aiming to address the following questions:

How can key TFs for $CD8^+$ T cell fate decisions be identified by deciphering

epigenetic landscapes?

In Chapter 2, we characterized the epigenetic landscape, chromatin accessibility and transcriptional program of terminal-effector, memory-precursor and memory $CD8^+$ T cells in response to bacterial infection *in vivo*. Integration of these data demonstrated that expression and binding of TFs established subset-specific enhancers impacting $CD8^+$ T cell differentiation. We leveraged our genome-wide analyses to develop a PageRank-based ranking system to more accurately predict novel TFs influencing effector and memory cells. Highlighting the power of this new strategy, we showed TFs YY1 and Nr3c1, both constitutively expressed during $CD8^+$ T cell differentiation, regulated the formation of terminal-effector and memory-precursor cell fates, respectively.

What are the molecular drivers for Trm and tumor-infiltrating T cell differentiation?

In Chapter 3, we showed that Trm precursors represent a unique $CD8^+$ T cell subset that is distinct from the precursors of circulating memory populations at the levels of gene expression and chromatin accessibility. We developed a computational screen integrating ATAC-seq, RNA-seq and PageRank analysis to infer key TFs for Trm specification and identified a novel role for the TF Runx3 in promoting Trm differentiation and homeostasis. We further showed that tumor-infiltrating T cells shared a core tissue-residency gene signature with Trm, and Runx3 can induce this core tissue-residency gene-expression program to promote the tissue-residency of $CD8^+$ T cells in both non-lymphoid tissues and tumors.

What is the role of the genome organizer, CTCF, in $CD8^+$ T cell differentiation?

In Chapter 4, we revealed that the ubiquitously expressed architectural protein CTCF is essential for terminal-effector, effector-like memory and secondary effector $CD8^+$ T cell differentiation. Interestingly, the loss of CTCF promoted the accumulation and differentiation of IEL Trm, especially $CD103^+$ Trm in response to LCMV infection. We further performed CTCF ChIP-seq and showed that CTCF can bind and regulate gene expression of the key TF, T-bet, which promotes terminal-effector $CD8^+$ T cell differentiation while repressing Trm formation. Finally, analysis of RNA-seq data from patients carrying *de novo* mutation of CTCF showed that the haploinsufficiency of CTCF impacted the TE gene signature in peripheral blood lymphocytes, suggesting a critical role of CTCF in TE subset differentiation in humans.

These data combined represent a hierarchy of transcriptional regulation of $CD8^+$ T cell differentiation at chromatin accessibility, chromatin state and chromatin organization levels.

1.5 References

[1] Victor Appay, Daniel C. Douek, and David A. Price. CD8+T cell efficacy in vaccination and disease. *Nature Medicine*, 14(6):623–628, 2008. ISSN 10788956. doi: 10.1038/nm.f.1774.

[2] R Ahmed and D Gray. Immunological memory and protective immunity: understanding their relation. *Science*, 272(5258):54–60, 1996. ISSN 0036-8075. doi: 10.1126/science.272.5258.54.

[3] Nikhil S. Joshi, Weiguo Cui, Anmol Chandele, Heung Kyu Lee, David R. Urso, James Hagman, Laurent Gapin, and Susan M. Kaech. Inflammation Directs Memory Precursor and Short-Lived Effector CD8+ T Cell Fates via the Graded Expression of T-bet Transcription Factor. *Immunity*, 27(2):281–295, 2007. ISSN 10747613. doi: 10.1016/j.immuni.2007.07.010.

[4] Susan M Kaech, Joyce T Tan, E John Wherry, Bogumila T Konieczny, Charles D Surh, and Rafi Ahmed. Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nature immunology*, 4(12):1191–8, 2003. ISSN 1529-2908. doi: 10.1038/ni1009.

[5] E. John Wherry, Volker Teichgräber, Todd C. Becker, David Masopust, Susan M. Kaech, Rustom Antia, Ulrich H. von Andrian, and Rafi Ahmed. Lineage relationship and protective immunity of memory CD8T cell subsets. *Nature Immunology*, 4(3):225–234, 2003. ISSN 15292908. doi: 10.1038/ni889.

[6] Stephen C. Jameson and David Masopust. Understanding Subset Diversity in T Cell Memory. *Immunity*, 48(2):214–226, 2018. ISSN 10747613. doi: 10.1016/j.immuni.2018.02.010.

[7] W M Flanagan, B Corthesy, R J Bram, and G R Crabtree. Two subsets of memory T lymphocytes with distinct homing potential and effector functions. 6(October):6270–6273, 1979.

[8] J. M. Schenkel, K. A. Fraser, L. K. Beura, K. E. Pauken, V. Vezys, and D. Masopust. Resident memory CD8 T cells trigger protective innate and adaptive immune responses. *Science*, 346(6205):98–101, 2014. ISSN 0036-

8075. doi: 10.1126/science.1254536.

[9] T Cd, J Justin Milner, and Ananda W Goldrath. Transcriptional programming of tissue-resident memory CD8+ T cells. *Current Opinion in Immunology*, 51:162–169, 2018. ISSN 09527915. doi: S0952791517300985.

[10] T Cd, Jason M Schenkel, Kathryn A Fraser, Vaiva Vezys, and David Masopust. Sensing and alarm function of resident memory. *Nature immunology*, 14(5):509–513, 2013. doi: 10.1038/ni.2568.

[11] John T Chang, E John Wherry, and Ananda W Goldrath. Molecular regulation of effector and memory T cell differentiation. *Nature reviews. Immunology*, 15(12), 2014. doi: 10.1038/ni.3031.1104.

[12] Xinyuan Zhou, Shuyang Yu, Dong-Mei Zhao, John T Harty, Vladimir P Badovinac, and Hai-Hui Xue. Differentiation and Persistence of Memory CD8(+) T Cells Depend on T Cell Factor 1. *Immunity*, 33(2):229–240, 2010. ISSN 1097-4180. doi: 10.1016/j.immuni.2010.08.002.

[13] Annie Xin, Frederick Masson, Yang Liao, Simon Preston, Tianxia Guan, Renee Gloury, Moshe Olshansky, Jian-Xin Lin, Peng Li, Terence P Speed, Gordon K Smyth, Matthias Ernst, Warren J Leonard, Marc Pellegrini, Susan M Kaech, Stephen L Nutt, Wei Shi, Gabrielle T Belz, and Axel Kallies. A molecular threshold for effector CD8+ T cell differentiation controlled by transcription factors Blimp-1 and T-bet. *Nature Immunology*, 17(4), 2016. ISSN 1529-2908. doi: 10.1038/ni.3410.

[14] Kyla D Omilusik, J Adam Best, Bingfei Yu, Steven Goossens, Alexander Weidemann, Jessica V Nguyen, Eve Seuntjens, Agata Stryjewska, Christiane Zweier, Rahul Roychoudhuri, Luca Gattinoni, Lynne M Bird, Yujiro Higashi, Hisato Kondoh, Danny Huylebroeck, Jody Haigh, and Ananda W Goldrath. Transcriptional repressor ZEB2 promotes terminal differentiation of CD8+ effector and memory T cell populations during infection. *The Journal of experimental medicine*, pages jem.20150194–, 2015. ISSN 1540-9538. doi: 10.1084/jem.20150194.

[15] Claudia X Dominguez, Robert A Amezquita, Tianxia Guan, Heather D Marshall, Nikhil S Joshi, Steven H Kleinstein, and Susan M Kaech. The tran-

scription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. *The Journal of experimental medicine*, 212(12):2041–2056, 2015. ISSN 1540-9538. doi: 10.1084/jem.20150186.

[16] Adriana Gonzalez-sandoval and Susan M Gasser. On TADs and LADs : Spatial Control Over Gene Expression. *Trends in Genetics*, 32(8):485–495, 2016. ISSN 0168-9525. doi: 10.1016/j.tig.2016.05.004.

[17] Peter A. Jones. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, 13(7):484–492, 2012. ISSN 14710056. doi: 10.1038/nrg3230.

[18] Ben Youngblood, J. Scott Hale, Haydn T. Kissick, Eunseon Ahn, Xiaojin Xu, Andreas Wieland, Koichi Araki, Erin E. West, Hazem E. Ghoneim, Yiping Fan, Pranay Dogra, Carl W. Davis, Bogumila T. Konieczny, Rustom Antia, Xiaodong Cheng, and Rafi Ahmed. Effector CD8 T cells dedifferentiate into long-lived memory cells. *Nature*, 552(7685):404–409, 2017. ISSN 14764687. doi: 10.1038/nature25144.

[19] Tony Kouzarides. Chromatin Modifications and Their Function. *Cell*, 128(4): 693–705, 2007. ISSN 00928674. doi: 10.1016/j.cell.2007.02.005.

[20] Yin Shen, Feng Yue, David F. McCleary, Zhen Ye, Lee Edsall, Samantha Kuan, Ulrich Wagner, Jesse Dixon, Leonard Lee, Victor V. Lobanenkov, and Bing Ren. A map of the cis-regulatory sequences in the mouse genome. *Nature*, 488(7409):116–120, 2012. ISSN 0028-0836. doi: 10.1038/nature11243.

[21] Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics*, 13(9):613–626, 2012. ISSN 1471-0056. doi: 10.1038/nrg3207.

[22] Eliezer Calo and Joanna Wysocka. Modification of Enhancer Chromatin: What, How, and Why?, 2013. ISSN 10972765.

[23] Yasuto Araki, Zhibin Wang, Chongzhi Zang, William H. Wood, Dustin Schones, Kairong Cui, Tae Young Roh, Brad Lhotsky, Robert P. Wersto,

Weiqun Peng, Kevin G. Becker, Keji Zhao, and Nan ping Weng. Genome-wide Analysis of Histone Methylation Reveals Chromatin State-Based Regulation of Gene Transcription and Function of Memory CD8+ T Cells. *Immunity*, 30 (6):912–925, 2009. ISSN 10747613. doi: 10.1016/j.immuni.2009.05.006.

[24] Brendan E. Russ, Moshe Olshanksy, Heather S. Smallwood, Jasmine Li, Alice E. Denton, Julia E. Prier, Angus T. Stock, Hayley A. Croom, Jolie G. Cullen, Michelle L T Nguyen, Stephanie Rowe, MatthewR R. Olson, David B. Finkelstein, Anne Kelso, Paul G. Thomas, Terry P. Speed, Sudha Rao, and Stephen J. Turner. Distinct epigenetic signatures delineate transcriptional programs during virus-specific CD8+ T cell differentiation. *Immunity*, 41(5): 853–865, 2014. ISSN 10974180. doi: 10.1016/j.immuni.2014.11.001.

[25] Boyko Kakaradov, Janilyn Arsenio, Christella E. Widjaja, Zhaoren He, Stefan Aigner, Patrick J. Metz, Bingfei Yu, Ellen J. Wehrens, Justine Lopez, Stephanie H. Kim, Elina I. Zuniga, Ananda W. Goldrath, John T. Chang, and Gene W. Yeo. Early transcriptional and epigenetic regulation of CD8 + T cell differentiation revealed by single-cell RNA sequencing. *Nature Immunology*, 18(4):422–432, 2017. ISSN 15292916. doi: 10.1038/ni.3688.

[26] Simon M. Gray, Robert A. Amezquita, Tianxia Guan, Steven H. Kleinstein, and Susan M. Kaech. Polycomb Repressive Complex 2-Mediated Chromatin Repression Guides Effector CD8+T Cell Terminal Differentiation and Loss of Multipotency. *Immunity*, 46(4):596–608, 2017. ISSN 10974180. doi: 10.1016/j.immuni.2017.03.012.

[27] Nathaniel D Heintzman, Gary C Hon, R David Hawkins, Pouya Kheradpour, Alexander Stark, Lindsey F Harp, Zhen Ye, Leonard K Lee, Rhona K Stuart, Christina W Ching, Keith a Ching, Jessica E Antosiewicz-Bourget, Hui Liu, Xinmin Zhang, Roland D Green, Victor V Lobanenkov, Ron Stewart, James a Thomson, Gregory E Crawford, Manolis Kellis, and Bing Ren. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, 459(7243):108–112, 2009. ISSN 0028-0836. doi: 10.1038/nature07829.

[28] D. Lara-Astiaso, A. Weiner, E. Lorenzo-Vivas, I. Zaretsky, D. A. Jaitin, E. David, H. Keren-Shaul, A. Mildner, D. Winter, S. Jung, N. Friedman, and I. Amit. Chromatin state dynamics during blood formation. *Science*, 345 (6199):943–9, 2014. ISSN 0036-8075. doi: 10.1126/science.1256271.

[29] R. David Hawkins, Antti Larjo, Subhash K. Tripathi, Ulrich Wagner, Ying Luu, Tapio Lnnberg, Sunil K. Raghav, Leonard K. Lee, Riikka Lund, Bing Ren, Harri Lhdesmki, and Riitta Lahesmaa. Global Chromatin State Analysis Reveals Lineage-Specific Enhancers during the Initiation of Human T helper 1 and T helper 2 Cell Polarization. *Immunity*, 38(6):1271–1284, 2013. ISSN 10747613. doi: 10.1016/j.immuni.2013.05.011.

[30] Golnaz Vahedi, Hayato Takahashi, Shingo Nakayamada, Hong Wei Sun, Vittorio Sartorelli, Yuka Kanno, and John J. OShea. STATs shape the active enhancer landscape of T cell populations. *Cell*, 151(5):981–993, 2012. ISSN 00928674. doi: 10.1016/j.cell.2012.09.044.

[31] Michael Bulger and Mark Groudine. Functional and mechanistic diversity of distal transcription enhancers, 2011. ISSN 00928674.

Chapter 2

Epigenetic landscapes reveal transcription factors regulating effector and memory $CD8^+$ T cell differentiation

2.1 Introduction

A number of TFs have been identified as critical regulators of the TE $CD8^+$ T cell fate including T-bet, Blimp1, Id2, Irf4, Batf, and Zeb2. Conversely, Tcf7, Eomes, Id3, E proteins, Bcl6, and Foxo1 are known to be important in the formation of MP and memory populations[1, 2, 3, 4]. Notably, not all of these factors exhibit differential expression between TE and MP, suggesting that additional mechanisms contribute to their activity in promoting cell fates. Further, how these

TF function within a coherent regulatory network is still unknown, and it is likely that additional TFs relevant in $CD8^+$ T cell differentiation remain unidentified.

We reasoned that integrated analysis of TF expression, binding and the expression of their gene targets would provide additional insights to identify previously unappreciated TF involved in $CD8^+$ effector and memory T cell differentiation. Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-seq) has recently been used to globally probe open chromatin to map regions of TF binding with high genomic resolution requiring minimal cellular input[5, 6]. By scanning all TF binding motifs on accessible chromatin regions, it is possible to infer the binding of hundreds of TF, as well as identify potential gene targets of all known TFs simultaneously, generating a global view of TF binding that has previously been technically impossible to achieve[7]. This provides a feasible strategy in the context of *in vivo* cellular immune responses where cell numbers are limiting. ATAC-seq proves especially powerful when combined with the characterization of the relevant regulatory elements such as: enhancers associated with histone acetyltransferase p300 and histone modifications such as monomethylation of histone H3 lysine 4 (H3K4me1); acetylation of histone H3 lysine 27 (H3K27ac) associated with active transcription; trimethylation of histone H3 lysine 4 (H3K4me3) associated with active promoters; and, trimethylation of histone H3 lysine 27 (H3K27me3) associated with gene repression[8, 9, 10]. Recent studies utilizing data generated by ATAC-seq in the context of chromatin modifications have facilitated the successful prediction of tissue-specific TFs and enhancers active in shaping the identity of macrophages in different tissues and of lineage determining factors in hematopoiesis[11, 12]. In the case of naive $CD8^+$

T cells, co-deposition of H3K4me3 and H3K27me3 at promoter regions is a signature of genes important for cellular differentiation, suggesting an epigenetic mechanism underlying $CD8^+$ T cell differentiation[13, 14]. However, these studies focused exclusively on promoter regions. Accumulating evidence suggests that enhancers play a key role in fine-tuning gene expression via long-distance communication with promoter elements, providing better cellular specificity compared with promoters[15, 11, 16, 17, 18], and enhancer landscapes important for effector and memory $CD8^+$ T cell differentiation remain largely unknown.

Here, we characterized the epigenetic landscapes of naive, terminal-effector, memory-precursor, and memory $CD8^+$ T cells generated in response to bacterial infection to identify both enhancer and promoter regions important for $CD8^+$ T cell effector and memory cell fates. Using ATAC-seq to identify accessible regulatory regions, we predicted TF candidates and further constructed a transcriptional regulatory network for each subset. To facilitate the identification of key TF regulators, we applied PageRank algorithm to rank the importance of TFs in each regulatory network. We identified TFs known to be central to $CD8^+$ T cell differentiation as well as numerous TFs not previously associated with $CD8^+$ effector versus memory T cell-fate specification. Among the novel TFs, we experimentally validated that Yin And Yang-1 (YY1) which acts as both transcriptional activator and repressor in numerous contexts, promotes a TE phenotype, and, Nuclear Receptor Subfamily 3 Group C Member 1 (Nr3c1) which binds to glucocorticoid response elements, supports MP differentiation. Taken together, our results yield a comprehensive catalog of the regulatory elements of $CD8^+$ T cells, revealing unexpected regulators controlling $CD8^+$ T cell fates. Furthermore, our computational

framework can be applied generally to any cell or tissue type to decipher regulatory networks and identify biologically significant TFs.

2.2 Results

2.2.1 TE and MP subsets reflect transcriptional differences between effector and memory CD8$^+$ T cells.

We previously characterized the transcriptome of the total antigen-specific population of CD8$^+$ T cells over the course of infection, identifying signatures of gene expression at different phases of the immune response[19].However, the effector population is characterized by extensive phenotypic and functional heterogeneity, including terminal-effector (TE) and memory-precursor (MP) CD8$^+$ T cell subsets[2]. Using microarray analysis to delineate transcriptional differences between these two effector subsets, we revealed 321 genes expressed $>$ 1.5 fold higher in the TE compared to the MP subset, and 653 genes expressed $>$ 1.5-fold higher in the MP compared to TE subset (Figure 2.8a). We further compared this with our previous gene expression data for total effector and memory CD8$^+$ T cells and found that 79% of genes upregulated in the TE subset are enriched in total effector CD8$^+$ T cells and 84% of genes upregulated in the MP subset are enriched in memory CD8$^+$ T cells (Figure 2.8b). This result indicates the unique transcriptional identities of effector and memory CD8$^+$ T cells can be captured on day 8 of infection in the TE and MP subsets, consistent with the idea that the MP subset is a precursor of memory CD8$^+$ T cells[2, 3]. Interestingly, the differences in mRNA expression of the majority of known TF that control the differentiation of

TE versus MP are subtle (Figure 2.8c), suggesting that expression differences alone do not account for the differential dependence on TF in distinct subsets. Likewise, the protein abundance of certain polarizing TF are not dramatically different between the TE and MP subsets when measured by flow cytometry (Figure 2.8d). These results indicated that in addition to TF expression, additional regulatory mechanisms, such as the control of TF binding, may be involved in controlling differentiation of these two subsets and subsequent formation of long-lived memory cells.

2.2.2 $CD8^+$ T cell subsets exhibit distinct enhancer repertoires

Spatial and temporal regulation of gene expression requires the specific binding of TF at regulatory elements, which is affected by chromatin state and accessibility. We combined ChIP-seq of histone modifications and ATAC-seq to provide insight into chromatin state and accessibility, respectively, allowing prediction of TF binding at specific regulatory elements. To characterize potential enhancer and promoter elements, we performed ChIP-seq with antibodies that specifically detect histone modifications H3K4me1 (enhancer mark), H3K4me3 (promoter mark), H3K27ac (active mark) and H3K27me3 (repressive mark) (Figure 2.1a)[20]. We used OT-I TCR transgenic $CD8^+$ T cells that specifically recognize ovalbumin (OVA); $2x10^4$ naive cells were transferred to congenically distinct host mice, followed by infection with Listeria monocytogenes engineered to express recombinant OVA (Lm-OVA)[19]. Using differential expression of CD45 alleles to identify donor cells, we sorted $KLRG1^{hi}IL\text{-}7R^{lo}$ TE $CD8^+$ T cells and $KLRG1^{lo}$IL-

7R^{hi} MP CD8$^+$ T cells on day 8 of infection or KLRG1^{lo}IL-7R^{hi} memory CD8$^+$ T cells $>$ 60 days after infection. Additionally, CD44^{lo} naive OT-I CD8$^+$ T cells were used as a reference. Indicated ChIP-seq and ATAC-seq were performed for two ex vivo replicates of sort-purified populations. Samples were sequenced to approximately 20 million reads per sample and sequences were aligned to mouse genome mm10 and unique, high quality reads were retained for subsequent analyses.

Previous studies showed that bivalent chromatin domains, comprising H3K4 me3 and H3K27me3 modifications, exist in the promoter regions of effector genes in naive CD8 T cells, and that H3K27me3 occupancy at these promoters was reduced upon differentiation into effector CD8$^+$ T cells[13, 14]. Consistent with this, we observed a similar pattern in the change of bivalent modification during differentiation. For instance, *Tbx21*, a gene encoding an essential TF promoting effector CD8$^+$ T cell differentiation, was co-occupied by both H3K4me3 and H3K27me3 in naive CD8$^+$ T cells. This gene locus lost the repressive mark H3K27me3 and maintained H3K4me3 upon differentiation (Figure 2.1b). Interestingly, we found that naive T cell-enriched genes were repressed in effector CD8$^+$ T cells, concomitant with increased H3K27me3 occupancy at promoter regions. For example, the promoter of *Tcf7*, a TF important for naive and memory CD8$^+$ T cells, was associated with H3K27me3 only in TE and MP effector subsets (Figure 2.1b). Upon further analysis, we found that the percentage of genes with occupancy of H3K27me3 at promoter regions was higher during differentiation to effector CD8$^+$ T cell subsets compared to memory CD8$^+$ T cells, suggesting that epigenetic repression of genes enriched in naive CD8$^+$ T cells may be essential for terminal differentiation of effector CD8$^+$ T cells (Figure 2.1c).

We examined the distal regulatory regions over the course of the $CD8^+$ T cell immune response, observing dynamic changes of histone modifications. Focusing on several well-characterized genes in $CD8^+$ effector and memory T cells, we found both gains and losses of enhancers and repressive H3K27me3 marks. For example, *Gzma*, a key effector gene highly expressed in TE, was associated with increased levels of H3K4me1 and H3K27ac upon differentiation from naive $CD8^+$ T cells to the TE subset (Figure 2.1d). Conversely, *Il7r* exhibited higher levels of H3K4me1 and H3K27ac in MP and memory $CD8^+$ T cells compared to the TE subset, consistent with its role promoting long-term survival for memory $CD8^+$ T cells (Figure 2.1d). Alternatively, *Id2*, a well-known transcriptional regulator of $CD8^+$ T cell differentiation, exhibited high levels of H3K4me1 in all $CD8^+$ T cells, but was associated with the increased level of H3K27ac, and decreased level of H3K27me3 only in activated $CD8^+$ T cells (Figure 2.1d)[21]. In contrast, *Id3* gene, a critical transcriptional regulator enforcing a naive state of T cells, was associated with an increased level of H3K27me3, and a concomitant decrease of gene expression upon differentiation (Figure 2.1d)[22]. Thus, as expected, combinatorial epigenetic marks set the stage for gene expression.

Collectively, we characterized the dynamics of chromatin states consisting of alterations of enhancer mark and active/repressive marks during $CD8^+$ T cell differentiation. To systematically identify putative enhancers from chromatin states in different $CD8^+$ T cell subsets, we applied a computational method called RFECS (Random Forest based Enhancer Identification from Chromatin States)[23]. RFECS is a machine-learning classifier algorithm that uses 3 histone marks - H3K4me1, H3K4me3 and H3K27ac - as features to predict enhancers. We

identified 27236, 26561, 23302, 21883 enhancers in naive, TE, MP and memory $CD8^+$ T cells respectively, which when merged comprised a non-redundant set of 52331 putative enhancers that may regulate gene expression of effector and memory $CD8^+$ T cells. Evaluation of the enhancer landscape as naive $CD8^+$ T cells differentiate revealed that TE gain a great number of newly formed enhancers compared to MP and memory CD8 T cells while all populations lose a similar number of enhancers enriched in naive CD8 T cells (Figure 2.2a). To better understand the dynamic usage of enhancers at different stages of $CD8^+$ T cell differentiation, we performed k-means clustering on the set of 52331 enhancers according to their H3K4me1 signal level. The enhancers were separated into five distinct clusters (Figure 2.1b). Cluster V represented common enhancers shared by all $CD8^+$ T cell subsets, while the other 4 clusters showed dynamics of enhancer establishment. In cluster V, H3K4me1 marks were maintained at equivalent levels across the $CD8^+$ T cell subsets and genes associated with this cluster were highly expressed in $CD8^+$ T cells such as *Cd8a* and *Lck*. Notably, *Tcf7* and *Prdm1* were also found in this cluster despite differential expression during differentiation, perhaps indicating other marks such as H3K27ac and H3K27me3 influence their expression. As such, the *Tcf7* locus was only associated with repressive mark H3K27me3 in effector $CD8^+$ T cell subsets, illustrating the relevance of gene repression by H3K27me3 and exemplifying how complete characterization of chromatin states provide a deeper understanding of subset-specific gene regulation. In cluster I and II, H3K4me1 marks were increased during differentiation from naive to effector and memory $CD8^+$ T cells, but were more enriched in the TE subset compared to MP and memory $CD8^+$ T cells. Genes associated with enhancers in these clusters were

associated with TE differentiation, such as *Klrg1* and *Tbx21* (Figure 2.2b)[2]. In cluster III, H3K4me1 marks were enriched in all differentiated compared to naive $CD8^+$ T cells and genes associated with enhancers in this cluster associated with $CD8^+$ T cell activation such as *Prf1*. In contrast, in cluster IV, H3K4me1 marks were lost during differentiation from naive $CD8^+$ T cell to the TE subset, and were more enriched in MP and memory $CD8^+$ T cells compared to the TE subset. Enhancers of canonical regulators of memory potential and homeostasis of $CD8^+$ T cells were found in cluster IV, such as *Il7r* and *Cxcr4* (Figure 2.2b)[3, 24].

Differential establishment of enhancers may be important for regulating subset-specific gene expression. To test this hypothesis, we assigned enhancers to the nearest genes and compared gene expression during $CD8^+$ T cell differentiation. Cluster I, II and III enhancers were associated with genes upregulated in activated $CD8^+$ T cells and cluster IV enhancers were associated with genes enriched in naive $CD8^+$ T cells (Figure 2.9a). Notably, the expression of genes associated with cluster I and II enhancers were more enriched in the TE subset compared to the MP subset, while those associated with cluster IV enhancers were more enriched in the MP subset (Figure 2.2b). Further, we observed genes encoding receptors for specific signaling pathway such as *Il12rb2* and *Il18rap* were associated with cluster I and II enhancers whereas *Tgfbr1* and *Tgfbr2* were associated with cluster IV enhancers. To gain insight into the biological function and molecular pathways associated with these enhancer clusters, we performed Gene Ontology (GO) analysis using the GREAT tool (Genomic Regions Enrichment of Annotations Tool)[25] and found that the IL-12 signaling pathway was enriched in cluster I and II, consistent with the role of IL-12 in promoting the TE subset

differentiation[2]. In addition, TGF-β and EGF signaling pathways were enriched in cluster IV, suggesting that these signaling pathways may favor the naive and/or memory T cell state, consistent with data showing TGF-β signaling is required for memory T cell differentiation (Figure 2.2c)[26]. We further observed that genes associated with increased number of enhancers correlated with higher expressed compared with those associated with a single enhancer (Figure 2.2d). Enhancer dosage has not previously been evaluated during the context of $CD8^+$ T cell differentiation, and here we demonstrated these analyses reveal novel insight into regulation of $CD8^+$ T cell gene expression.

2.2.3 TF expression and putative binding contribute to establishment of subset-specific regulatory elements

We reasoned that TF binding motifs would be enriched in accessible regulatory regions, which could be used to discover TF important for $CD8^+$ T cell differentiation. To better predict TF important for $CD8^+$ T cell subsets, we performed ATAC-seq to probe open chromatin and overlapped the ATAC-peaks with enhancer and promoter regions to uncover accessible regulatory regions. We identified subset-specific open enhancers and promoters and then scanned 761 unique known TF binding motifs at the center of the ATAC-peaks of these regulatory regions (Figure 2.3a). For instance, we discovered that the T-bet binding motif was enriched in a TE-specific accessible enhancer near the *Zeb2* gene locus, which was exclusively expressed in TE subset, supporting previous findings that T-bet is essential in regulating TE subset differentiation and directly regulates Zeb2 in this subset (Figure 2.3b)[27, 28]. Using randomly-selected open chromatin regions

as background, we computed the p-value for TF enrichment or depletion using a binomial test to identify candidate TF. Consistent with previous reports, our approach predicted putative binding of known TF such as T-bet, Batf, Srebf2, Eomes, Tcf7, Lef1 and E2A at promotes and enhancers [29, 30, 31, 32, 33, 1]. T-bet, Batf and Srebf2 binding motifs were depleted in naive and enriched in all differentiated $CD8^+$ T cells, consistent with their crucial roles in $CD8^+$ T cell activation and effector function (Figure 2.3c). Notably, the T-bet motif was highly enriched in the TE subset compared to MP, consistent with its role in TE subset differentiation. Tcf7, Lef1 and E2A binding motifs were depleted in TE and enriched in naive, MP and memory $CD8^+$ T cells, consistent with well-characterized roles in maintaining/promoting memory potential of $CD8^+$ T cells (Figure 2.3c)[1, 33]. Further, the binding motifs of AP-1 TF family members like Jun, JunB, JunD and Fosl2 were highly enriched in all activated $CD8^+$ T subsets, consistent with TCR signaling inducing AP-1 TF binding to regulatory elements, promoting $CD8^+$ T cell activation in all subsets (Figure 2.10)[34]. Enrichment of some TF binding motifs (Tcf7 and T-bet) was highly correlated with gene expression and function [1, 2];in contrast, the enrichment of other TF binding motifs was consistent with their putative roles yet were not differentially expressed. For example, the Srebf2 binding motif is depleted in naive $CD8^+$ T cells and highly enriched in effector T cell subsets, which is consistent with the role of Srebf2 in maintaining effector T cell growth and activation, although the expression of Srebf2 is similar between naive and effector $CD8^+$ T cells[30]. Likewise, E2A binding motif is depleted in TE but highly enriched in MP and memory CD8 T cells, consistent with the function of E2A in promoting MP and memory CD8 T cell differentiation[33]. Thus, while

expression of E2A does not change during CD8 T cell differentiation, differential binding to distinct regulatory elements provides a plausible mechanism for differential E protein activity in TE versus MP and memory fates (Figure 2.3d). These data indicated that subset-specific enhancers and promoters might be established by TF where differential expression and binding both contributed, and that putative TF binding, in addition to expression, must be considered when predicting key TF. As such, previous studies primarily relying on mRNA expression data to predict regulators of memory formation have likely missed TF that impact TE versus MP or memory fates.

2.2.4 Constructing a TF regulatory network during CD8$^+$ T cell differentiation

Alterations in TF binding allows the same TF without differential expression to produce distinct transcriptional outputs through regulation of distinct gene targets in various contexts. To elucidate TF-mediated regulatory mechanisms underlying CD8$^+$ T cell differentiation, we sought to construct a TF regulatory network in different CD8$^+$ T cell subsets. Previous studies have applied gene co-expression correlation to construct regulatory networks that connect the TF with potential gene targets if their expression patterns are highly correlated[35, 36]. However, such analysis does not take into account direct TF-binding. In addition, subtle variations in expression of key TF can lead to dramatic changes in expression of downstream gene targets and thus it is challenging to identify these TF based on gene expression data alone. We combined TF binding motifs, chromatin states and chromatin accessibility information to predict and link available TF

binding sites to their potential gene targets (Figure 2.4a). Using the predicted regulation, we reconstructed TF regulatory networks and uncovered critical regulatory circuits responsible for $CD8^+$ T cell differentiation. Here, we focused on T-bet-regulated circuits in TE and MP subsets, given the essential role of T-bet in supporting $CD8^+$ T cell effector function and TE subset differentiation despite the fact that T-bet is expressed at relatively high levels in both TE and MP subsets (Figure 2.8)[37]. In connection, T-bet regulated gene network in different subsets is not well described.

Globally a substantial number of putative targets regulated by T-bet were identified in both TE and MP subsets, indicative of similar T-bet expression in both populations (Figure 2.4b). We compared predicted T-bet regulated genes between the TE and MP subsets and found that 61.4% of candidate genes were shared in both subsets, such as *Ifnγ* and *Cxcr3*, well-established T-bet regulated targets important for $CD8^+$ T cell effector function (Figure 2.4c)[38, 39]. Interestingly, the subset-specific T-bet regulatory circuits predicted that T-bet uniquely controls expression of *Zeb2*, *Gzma* and *Klrb1c* in TE and *Bcl2*, *Crtam* (an activation-induced surface receptor important for retention of lymphocytes in lymph node) and *Pou6f1* (a transcription factor highly expressed in memory $CD8^+$ T cells) in MP. To validate our analyses and test the role of T-bet in regulating these predicted genes in a subset-specific manner, we co-transferred T-bet-WT and T-bet-KO OT-I $CD8^+$ T cells into hosts that were then infected with Lm-OVA. Given the loss of the TE subset in T-bet KO mice, we sort-purified either total donor $CD8^+$ T cells or the MP subset from T-bet WT and KO mice and compared the mRNA expression of candidate gene targets by qPCR. If the expression of gene

targets are primarily affected in total donor CD8$^+$ T cells rather than the MP subset, this is indicative of a TE-specific T-bet mediated gene expression. In the absence of T-bet, there was a 260-fold decrease of Zeb2 expression in total donor CD8$^+$ T cells compared to a 5-fold decrease in the MP subset, consistent with T-bet regulating Zeb2 expression in TE rather than MP subset as shown previously (Figure 2.4d-e)[27, 28]. To avoid the bias of loss of TE subset in T-bet KO mice, we compared the gene expression in TE subset of T-bet WT and HET (where a sizable TE population remains) and confirmed a loss of expression of *Zeb2*, *Gzma* and *Klrb1c* in the TE subset (Figure 2.4f-g). Interestingly, the loss of T-bet (HET and KO) impacted the expression of *Bcl2*, *Crtam* and *Pou6f1* mRNA in the MP subset, suggesting that T-bet regulates *Bcl2*, *Crtam* and *Pou6f1* in a MP-specific manner (Figure 2.4f-g). As the role for T-bet in MP cells has not been previously explored, we compared the frequency of WT versus T-bet KO CD8$^+$ T cells during infection. The absence of T-bet resulted in a defect in accumulation of MP cells consistent with T-bet regulating memory differentiation (Figure 2.4h). Thus, we concluded that T-bet positively regulates different genes in distinct subsets, such as *Zeb2*, *Gzma* in TE or *Bcl2*, *Crtam* and *Pou6f1* in MP subsets. Taken together, we have provided a comprehensive view of TF regulatory networks during CD8$^+$ T cell differentiation, allowing prediction of potential gene targets unique to different CD8$^+$ T cell subsets for similarly expressed TFs.

2.2.5 Identification of novel TF from PageRank-based TF ranking

While it is well-established that constitutively expressed TF can exert cell-type-specific functions through regulating specific gene expression, incorporating this knowledge to identify key TF has proven challenging, because the downstream targets of TF are largely unknown. To overcome this limitation, we took advantage of the TF regulatory network generated in Figure 4 which yields downstream putative gene targets for all available TF and we further applied advanced graph algorithms to identify biologically significant TF. We used the personalized PageRank algorithm to assess the importance of each TF in the TF regulatory network. The personalized PageRank algorithm measures global influence of each node in a network, used by Google and many other companies to order search engine results[40]. In an internet network, nodes are web pages and edges are links between websites. The PageRank algorithm was designed to find out how likely a specific web page is visited if web surfers who start on a random page sampled from a given distribution have α probability of choosing a random link from the page they are currently visiting and $1-\alpha$ probability of jumping to a random page chosen from the entire set of web pages. When applying the personalized PageRank algorithm to a TF regulatory network, the rank of a given TF is determined by two factors: the number of regulatees and the importance of regulatees that are regulated by the query TF. In other words, TF that regulate more genes or regulate more important genes would receive higher ranks. The importance of genes were assessed by relative expression of genes across different cell types/conditions

generated from microarray data (Figure 2.5a).

Using PageRank analysis, we predicted 100 key TF important for $CD8^+$ T cell differentiation. Prior studies have used motif enrichment analysis to predict TF candidates[11, 12]. Therefore, we compared our PageRank analysis with motif enrichment analysis to investigate how many known TF reported previously as essential regulators of $CD8^+$ T cell differentiation can be recovered from predicted TF pools. We found that approximately half of the predicted TF were identified by both analyses, and 25% of these shared TF regulators are shown to be key TFs based on previous studies (Figure 2.11a-b). PageRank analysis revealed additional known TF regulators compared with motif enrichment analysis (22% in PageRank-specific compared to 5% in motif enrichment-specific) (Figure 2.5b). For example, PageRank analysis revealed Stat3 scores higher in memory $CD8^+$ T cells compared to the TE subset (Figure 2.5c). This was consistent with the role of Stat3 in promoting the formation of mature and self-renewing memory $CD8^+$ T cells[41]. Thus, these data highlighted the robustness of PageRank analysis in the identification of key TF, suggesting that unknown TF predicted by PageRank analysis might be critical for $CD8^+$ T cell differentiation.

2.2.6 PageRank analysis accurately predicts essential roles for YY1 and Nr3c1 in TE and MP subset differentiation

To highlight the power of PageRank analysis, we focused on YY1 and Nr3c1, which were predicted by PageRank analysis (Figure 2.5) but not by the standard

motif analysis (Figure 2.4). Although the expression of YY1 and Nr3c1 did not change during $CD8^+$ T cell differentiation (Figure 2.11c), PageRank analysis determined that YY1 scored highly in the TE subset while Nr3c1 scored highly in the MP subset (Figure 2.5c). YY1 is a transcription factor shown to act as both an activator and a repressor of gene expression[42]. Prior studies have shown that YY1 is important in immune cell development such as the maturation of thymocytes and differentiation of B and regulatory T cells[43, 44, 45]. On the other hand, Nr3c1, encodes glucocorticoid receptor, which can function as a TF by translocating into nucleus to regulate gene expression after binding to glucocorticoid hormone in the cytosol. Nr3c1 plays a critical role in development, metabolism and the immune response (i.e. thymocyte development and anti-inflammatory response)[46, 47, 48]. However, the role of YY1 and Nr3c1 in $CD8^+$effector or memory T cell differentiation in response to infection are unknown.

Based on the PageRank analysis, we hypothesized abrogated expression of YY1 and Nr3c1 would affect the formation of TE or MP subsets, respectively. To test if YY1 is essential for TE $CD8^+$ T cell differentiation, we co-transfered $CD8^+$ T cells infected with retrovirus encoding shRNA targeting YY1 or control shRNA (targeting CD19 which is not expressed in $CD8^+$ T cells) into congenically distinct recipient mice followed by Lm-OVA infection and followed effector T cell differentiation (Figure 2.6a). RT-qPCR confirmed knockdown of YY1 levels resulting in a 54% reduction in mRNA levels (Figure 2.6b). Flow cytometric analysis of $CD8^+$ T cell subsets at day 7 of infection showed a significant reduction in both the frequency and number of the TE $CD8^+$ T cells in comparison to control shRNA transduced donor cells (Figure 2.6c-d). In addition, the expression

of MP-associated markers such as CD27 and CXCR3 and memory-important TF, Tcf7, were significantly increased by donor cells after knockdown of YY1 (Figure 2.6e)[1]. Furthermore, we examined cytokine production by flow cytometric analysis following *in vitro* restimulation with OVA peptide of shYY1 and control shRNA transduced cells. Both IFNγ expression and the number of IFNγ producing cells was reduced in the absence of YY1 while the number of TNFα producing cells was not changed (Figure 2.6f). Together, these data confirmed that YY1 is important for differentiation of TE subset $CD8^+$ T cells.

Similarly, we determined how lowering Nr3c1 expression impacted the differentiation of $CD8^+$ T cell MP subset differentiation. RT-qPCR showed that Nr3c1 mRNA knockdown resulted in 86% reduction compared to control shRNA (Figure 2.7a). Over the course of infection, we observed that both frequency and number of MP were significantly decreased when NR3c1 was knocked down compared to control group (Figure 2.7b-c). Consistent with a loss of CD127-expressing cells, levels of MP-associated genes CD27 and CXCR3 and memory-associated TF Tcf7 were significantly reduced after knockdown of Nr3c1 compared to control-transduced cells (Figure 2.7d), supporting a functional role for Nr3c1 in MP subset differentiation. To further test if the defect in MP subset differentiation persists during the contraction phase, we monitored the percentage of MP subset from day 8 to day 30 of infection and observed a dramatic decrease of the percentage of MP subset on day 30 after the loss of Nr3c1 (Figure 2.7e-f). Nr3c1 has been shown to interact with multiple co-factors such as Nuclear receptor co-repressor 1 (Ncor1) to modulate hormone-response gene expression[49]. In connection, we performed Ncor1 shRNA knockdown to determine if the loss of Nr3c1 cofactor, Ncor1,

affected MP subset differentiation. The percentage of MP subset was decreased after knockdown of Ncor1, consistent with the phenotype of Nr3c1 knockdown (Figure 2.12a). To further confirm the role of glucocorticoid receptor Nr3c1 in MP subset differentiation, we treated mice with synthetic glucocorticoids agonist Dexamethasone (Dex) for 7 days and observed that the percentage of MP subset was significantly increased after Dex treatment (Figure 2.12b). Collectively, these data demonstrated that glucocorticoid receptor Nr3c1 is important in promoting MP subset differentiation.

2.3 Discussion

The specificity and function of immune cells are controlled by TF that sense environmental cues to regulate specific gene expression. Efficient transcriptional regulation requires the interplay between TF and chromatin remodelers to control TF binding with high fidelity[50]. Key information is encoded in regulatory elements that contain specific TF binding sequences and are associated with specific histone modifications that influence accessibility, structure and location of that element[15]. Therefore, to identify the TF-mediated regulatory circuits critical for $CD8^+$ T cell differentiation, we systematically characterized the epigenome of $CD8^+$ T cell subsets during the response to pathogen infection. Our global map of regulatory elements revealed a dynamic pattern of enhancer establishment that foreshadows specific gene expression-programs.

Our combined TF gene-expression profile and TF motif enrichment analysis demonstrate that gene expression alone may not fully explain cell-fate determina-

tion, supporting the idea that both potential TF binding and gene expression should be considered together to facilitate the identification of important TF. Numerous crucial TF in the context of $CD8^+$ T cell differentiation have been identified based on differential gene expression as well as construction of TF-mediated gene regulatory networks based on gene-gene co-expression correlation[19, 35, 36]. However, alterations of TF binding without changes in TF expression also result in differential expression of downstream gene targets, making it clear that identification of key TF based only on gene-expression analysis provides only a partial understanding of the relevant TF networks. Differential TF binding can be achieved in several ways including: 1) Differential chromatin state and accessibility (compacted nucleosome occupancy could prevent TF access to binding sites even if the TF is substantially expressed), 2) TF translocation to the nucleus (location or activation changes of a TF without changes in TF levels), 3) Availability of co-factor or inhibitor (differential TF binding regulated by other proteins). The approach we describe represents an important advance in generating a complete view of the regulatory networks that establish $CD8^+$ effector and memory T cell fates by integrating data describing mRNA expression and chromatin states and accessibility, providing new insights into the subset-specific targets of known factors such as T-bet as well as identifying many factors that have not previously been implicated in $CD8^+$ effector and memory T cell differentiation.

Our network analysis of subset-specific T-bet regulatory circuits emphasizes the potential for differential activity in distinct cell fates as well as the importance of gene targets potentially regulated by TF. As a well-established TF important in TE subset differentiation, the function of T-bet in MP subset has been over-

looked and remains unknown. Here, we first revealed the novel function of T-bet in maintaining MP subset survival, potentially through regulating anti-apoptotic protein Bcl2. Additionally, since Crtam is important in lymphocyte migration and adhesion[51], the MP-specific regulation of Crtam by T-bet might suggest a role for T-bet in modulating the migration of MP subset.

To prioritize these data, it is imperative to develop new methods that rank the potential importance of TF based on the quantity and quality of gene targets regulated by that TF. Here, we constructed a TF regulatory network and applied the Personalized PageRank algorithm to measure the importance of TF by taking into account of both TF binding and gene expression information. This new strategy is robust, as it allows us to recover TF known to play key roles in $CD8^+$ T cell differentiation. For example, PageRank analysis revealed Foxo3 scored highly in $CD8^+$ memory T cells, consistent with a cell-intrinsic regulatory role of Foxo3 in $CD8^+$ memory T cells[52]. Furthermore, PageRank analysis revealed that Gfi1 scored highly in $CD8^+$ memory T cells (as it regulates more memory-relevant genes). This is consistent with that Gfi1 repress IL-7r expression to maintain early IL-$7r^{lo}$ effector $CD8^+$ T cells formation[53]. As a repressor, it is likely that Gfi1 repress genes important for $CD8^+$ memory T cell differentiation.

In addition to known TF, we identified a range of novel TF predicted to modulate $CD8^+$ T cell differentiation. Using shRNA knockdowns, we validated two of the predicted TF, YY1 and Nr3c1, demonstrating their essential roles in TE and MP subset differentiation, respectively. YY1 is a ubiquitously expressed TF and Nr3c1 can translocate to nucleus upon binding of glucocorticoids in cytosol. Thus, it is difficult to identify these TF when using traditional approach like differential

gene expression analysis. YY1 has been shown to modulate chromatin remodeling and mediate long-range chromatin interaction of Th2 cytokine loci in Th2 cells[54]. Thus how YY1 regulates TE subset differentiation and if YY1 controls chromatin interaction in the TE subset remains to be determined. Glucocorticoid receptor (Nr3c1) has been shown to regulate immune response such as inducing thymocytes apoptosis and inhibiting inflammation. Here, we first showed that Nr3c1 regulates MP subset differentiation, consistent with a role of glucocorticoids in promoting IL-7r expression[47]. As expected, treatment of synthetic glucorcorticoid agonist Dex increased the proportion of MP subset during differentiation. Thus, using our framework, we can both identify key TF and predict critical microenvironmental signals involved in regulating $CD8^+$ T cell differentiation and function.

Despite the successful experimental validation of TF predicted by our computational framework, additional factors could be integrated to refine our results. The current TF binding motif library is not complete, therefore limiting the breadth of our analysis. Global investigation of TF binding motifs using new approaches, such as protein binding microarray, may be beneficial to broaden the database of known TF binding motifs[55]. Moreover, TF function with binding partners or cofactors to regulate specific gene expression. Thus, TF-TF co-binding analysis could be incorporated to improve our network construction, and pooled screens using targeted libraries of shRNAmirs could be used to interrogate experimentally the functionality of predicted TF networks and their interaction[56, 50, 57]. Furthermore, the assignment of enhancers to the nearest genes is a limited heuristic, and being able to better associate long-range enhancers to gene targets would enhance the power of our approach considerably.Recent stud-

ies have shown the interaction of enhancer and promoters is confined in topologically associated domains[58, 59]. Future exploration of chromatin organization using high-resolution Hi-C or ChIA-PET of enhancer marks will facilitate the assignment of enhancers to their targets[59, 60].

2.4 Acknowledgement

Chapter 2, is a reprint of the material as it appears in Nature Immunology 2017. Yu B, Zhang K, Milner JJ, Toma C, Chen R, Scott-Browne JP, Pereira RM, Crotty S, Chang JT, Pipkin ME, Wang W, Goldrath AW. Epigenetic landscapes reveal transcription factors that regulate CD8+ T cell differentiation. *Nature Immunology.* 2017 18(5):573-582

The dissertation author was a primary investigator and the first author of this paper.

2.5 Methods

2.5.1 Mice, cell transfer, infection.

All mice were maintained in a specific pathogen-free condition according to the instructions of Institutional Animal Care and Use Committee (IACUC) of the University of California, San Diego (UCSD). OT-I TCR transgenic (specific for $OVA_{257\text{-}x}$)-MHC H2-K^b), *Tbx21* knockout, CD45.1 congenic, and C57BL/6J mice were either bred at UCSD or received from The Jackson Laboratory. We

transferred 2×10^4 OT-I TCR transgenic $CD8^+$ T cells into congenically distinct mice by i.v. injection and then infected i.v. with 5×10^3 cfu Listeria monocytogenes expressing OVA (Lm-OVA) one day later. For T-bet KO experiment, we co-transfered 1×10^4 T-bet WT OT-I and T-bet KO OT-I $CD8^+$ T cells into host mice and then i.v. infected with 5×10^3 cfu Lm-OVA.

2.5.2 Antibodies and flow cytometry.

KLRG1 (2F1), CD127 (A7R34), CD8 (53-6.7), CD45.1 (A20-1.7), CD45.2 (104), CXCR3 (CXCR3-173), CD27 (LG-7F9), T-bet (4B10), Bcl6 (K112-91),were purchased from eBioscience. Foxo1(C29H4), Tcf7(C63D9), IFNγ(XMG1.2), TNFα (MP6-XT22) were from cell signaling technology. Antibodies for ChIP-seq were used as follows: H3K4me3 (Ab8580), H3K4me1 (Ab8895) and H3K27ac (Ab4729) were from Abcam. H3K27me3 (07-449) was from Millipore. For intracellular staining of cytokines, splenocytes were in vitro restimulated with 1μg/mL OVA peptide (SIINFEKL) with Protein Transport Inhibitor (eBioscience) for 4h and then fixed and permeabilized using BD cytofix/cytoperm kit (BD Biosciences). Foxp3-transcription factor staining buffer kit (eBioscience) were used for intracellular staining of transcription factors. For intracellular staining of shRNA transduced cells containing Ametrine-reporter, cells were fixed using freshly made 2% formaldehyde for 45 min on ice and then permeabilized. All flow cytometry data were acquired by BD LSRFortessa X-20 and all sorts were performed on BD FACS Aria.

2.5.3 shRNA knockdown by retroviral transduction.

The detailed protocol was described previously[57]. PLAT-E cells were transfected with shRNAmir using TransIT-LT1 Reagent (Mirus). Retrovirus-containing supernatant was harvested after 48 hours and mixed with 2-mercaptoeth anol and polybrene (Millipore) for transduction. Purified naive OT-I $CD8^+$ T cells were *in vitro* activated by anti-CD3 and anti-CD28 for at least 18 hours and then spun at 2,000 rpm with retrovirus for 1 hour at 37C. After 4 hr incubation, the retrovirus-containing medium was replaced by T cell medium. Transduction efficiency were measured by flow cytometric analysis of ametrine-reporter after 24 hr and $1x10^4$ shRNA transduced cells were transferred into host mice followed by Lm-OVA infection.

2.5.4 RT-PCR and qPCR.

For RT-PCR, RNA was extracted using Trizol (Life Technologies) followed by precipitation of isopropanol. CDNA was synthesized using Superscript II kit (Life Technologis) following the manufacturers instruction. For qPCR, cDNA was quantitatively amplified using Stratagene Brilliant II Syber Green master mix (Agilent Technologies). The abundance of transcripts were normalized to housekeeping gene *Hprt*. The following primers were used: Zeb2 forward: CATGAACCCATTTAGTGCCA, Zeb2 reverse: AGCAAGTCTCCCTGAAATCC, Bcl2 forward: ACTTCGCAGAGATGTCCAGTCA, Bcl2 reverse: TGGCAAAGCGTCCCCTC, Gzma forward:TGCTGCCCACTGTAACGTG, Gzma reverse:GGTAGGTGAAGGATAGCCACAT, Klrb1c forward: GACACAGCAAGTATCTACCT, Klrb 1c reverse: TACTAAGACTCGCACTAAGAC, Pou6f1 forward: GTCAGATCCT-

CACGAATGCTC, Pou6f1 reverse: GAGTCACGGCTTGGACCTG, Crtam forward: CCTTTTCATCATCGTTCAGCTCT, Crtam reverse: GGAGCCTGGCTGCTATTCTC, YY1 forward: CATGTGGTCCTCGGATGAAA, YY1 reverse: GGGAGTTTCTTGCCTGTCATA, Nr3c1 forward: CCGGGTCCCCAGGTAAAGA, Nr3c1 reverse: TGTCCGGTAAAATAAGAGGCTTG, Hprt forward: GGCCAGACTTTGTTGGATTT, Hprt reverse: CAACTTGCGCTCATCTTAGG

2.5.5 Microarray analysis.

The protocol was described previously[19]. KLRG1^{hi}IL-7R^{lo} TE and KLRG1^{lo}IL-7R^{hi} MP CD8^{+} T cells (2x10^{4}) were sorted into TRIzol on day 8 of Lm-OVA infection. RNA was amplified and labeled with biotin followed by hybridized to Affymetrix Mouse Gene ST 1.0 microarrays (Affymetrix). Microarray analysis was performed using GenePattern Multiplot Studio module. All data was generated in collaboration with the Immgen project (www.immgen.org) and passed ImmGen quality control pipeline.

2.5.6 Chromatin Immunoprecipitaion (ChIP), ChIP-seq library construction and sequence alignment.

Cells were fixed in 1% formaldehyde for 10 min and then quenched with 125mM glycine for 5 min. Cells were lysed for 5 min on ice and sonicated to generate 200-500 bp fragments using Bioruptor sonicator (Diagenode). Sonicated DNA was used as input control. 30μL magnetic-dynabeads were washed with blocking buffer twice and then mixed with 5μg antibody in 500μL blocking buffer

and rotated at 4C. The sonicated lysates were first diluted to a final 0.1% SDS concentration. The diluted lysates were added to antibody-conjugated Dynabeads incubated at 4C. Beads were washed by Wash Buffer I, II and III for 5 min and then washed twice by TE buffer for 5 min. The beads were resuspended in 200 μL Elution Buffer and reverse-crosslinked at 65C overnight and then treated with RNAse for 30 min at 37C and Proteinase K at 55C for 1h. DNA was purified by Zymo DNA Clean & Concentrator kit (Zymo Research). The purified DNA was end-repaired using End-it End-repair kit (Epicentre) and then added an "A" base to the 3 end of DNA fragments using Klenow (NEB). Then DNA was ligated with adaptors using quick DNA ligase (NEB) at 25C for 15 min followed by size selection of 200-400 bp using AMPure SPRI beads (Beckman Coulter). The adaptor ligated DNA was amplified using NEBNext High-Fidelity 2X PCR master mix (NEB). To prevent PCR overamplification, 1μL DNA was first quantitatively amplified using Syber Green I master mix to determine the best amplification cycle. Then the amplified library was size-selected as 200-400bp using SPRI beads and quantified by Qubit dsDNA HS assay kit (ThermoFisher). Finally the library was sequenced using Hiseq 2500 for single-end 50bp sequencing to get around 20 million reads for each sample. We used BWA to map raw reads to the Mus musculus genome (mm10) with following parameters: "-q 5 -l 32 -k 2"[61]. Reads with low quality (MAPQ < 30) were filtered out. If multiple reads were mapped to the exact same location, only one read was kept.

2.5.7 ATAC-seq and peak calling.

The detailed protocol was described previously[5]. Cells were sorted (2.5×10^4) into 1mL FACS buffer and spun down 1500rpm for 20 min at 4C. The cell pellet was resuspended in 25μL lysis buffer and then spun down 2500rpm for 30 min at 4C. The nuclei pellet was resuspended into 25μL transposition reaction mixture containing Tn5 transposase from Nextera DNA Sample Prep Kit (Illumina) and incubated at 37C for 30 min. Then the transposased DNA was purified using Zymo DNA clean-up kit. To amplify the library, the DNA was first amplified for 5 cycles using indexing primer from Nextera kit and NEBNext High-Fidelity 2X PCR master mix. To reduce the PCR amplification bias, 5μL of amplified DNA after the first 5 cycles was used to do qPCR of 20 cycles to decide the number of cycles for the second round of PCR. Usually the maximum cycle of the second found of PCR is 5 cycles. Then the total amplified DNA was size selected to fragments less than 800 bp using SPRI beads. Quantification of the ATAC-seq library was based on KAPA library quantification kit (KAPAbiosystems). The size of the pooled library was examined by TapeStation. Finally, the library was sequenced using Hiseq 2500 for single-end 50bp sequencing to get at least 10 million reads. To obtain confident peaks, we performed each ATAC-seq experiment at least twice and used the Irreproducibility Discovery Rate (IDR) framework to identify the reproducible peaks. In particular, we called peaks for each individual replicate as well as the pooled data from the two replicates using MACS2 with a relaxed threshold (p-value 0.01)[62]. These 3 sets of peaks were input for IDR analysis using a threshold of 0.05 to identify the confident set of peaks.

2.5.8 Enhancer prediction.

Enhancers were identified by RFECS algorithm using 3 histone marks (H3K 4me1, H3K4me3 and H3K27ac)[23]. The RFECS model was trained on the active and distal P300 ChIP-seq peaks (at least 2 kb away from any TSS), which were taken as the representatives of enhancers (the positive set). For the non-enhancer class (the negative set), we chose promoters overlapping with DNase I hypersensitivity (DHS) peaks, and random 100 bp bins that are distal (2 kb away) to any P300 site and TSS. The trained model was used to scan the whole genome except the 2000 bp upstream of TSS and 500 bp downstream of TSS and classify each 100 bp bin as an enhancer or non-enhancer based on the histone modification pattern. To further reduce false positive rate, we filtered the predicted enhancers using a false discover rate (FDR) of 1

2.5.9 Predicting putative TF binding sites.

To identify putative binding sites of TFs, we first collected 761 unique motifs from 3 TF motif database, JASPAR, UniPROBE and Jolma et al. study[63, 64, 65]. We then searched for motifs binding sites in 150bp regions centered around ATAC-seq peak summits, using the algorithm described in Grant's paper with of P-value cutoff of 1e-5[66].

2.5.10 Motif enrichment analysis at open chromatin regions.

To compute the enrichment of a TF motif over cell-type-specific open chromatin regions, we first identified the number of regions that contain at least one motif binding site, denoted by m. Let N be the number of all regions, then $\frac{m}{N}$ is considered as the enrichment score of the query motif. To construct the null model for P-value calculation, we randomly selected 10,000 regions from all open chromatin sites and computed the fraction of those regions, denoted by p, containing at least one occurrence of the motif. The P-value for enrichment or depletion is then computed using the binomial test with p as the population proportion of null hypothesis.

2.5.11 Constructing TF regulatory networks.

We selected active promoters as the 5-kb-regions around TSS (4 kb upstream and 1 kb downstream) that are marked by H3K4me3 peaks. Enhancers were predicted using the RFECS based on enhancer-associated histone modification signatures. Enhancers were linked to the nearest genes. We connected a TF to a gene if the TF had any predicted binding site in the gene's promoter or linked enhancers. We assembled all the regulatory interactions between TFs and genes into a genetic network.

2.5.12 Personalized PageRank.

PageRank is the stationary distribution of a random walk which, at each step, with a certain probability α jumps to a random node, and with probability $1-\alpha$ follows a randomly chosen outgoing edge from the current node. Personalized PageRank is an extension of PageRank in which all the jumps are made according to a pre-defined probability distribution[67]. To give a formal definition, let $\mathcal{G} = (V, E)$ denote a directed graph, where V is a set of nodes and E contains a directed edge $\langle u, v\rangle$ if and only if node u links to node v. Let A be the transition matrix. We define $A_{ij} = \frac{1}{O(j)}$ if node j links to node i, and $A_{ij} = 0$ otherwise, where $O(j)$ is the out-degree of node j. Given a seed vector s, the personalized PageRank vector v is calculated by

$$v = (1 - \alpha)Av + \alpha s$$

In a TF regulatory network, we set the weight of each gene to e^{z_i}, where z_i is the z-score of expression levels of gene i under different conditions or in different cell states. The weights of genes are then normalized and used as the seed vector for computing personalized PageRank.

2.6 Figures

Figure 2.1: **Epigenetic landscape of CD8$^+$ T cells in response to bacterial infection. a,** Schematic view of experimental design for characterization of global epigenetic landscape and gene expression of naive(N), terminal effector(TE), memory precursor(MP), and memory(M) CD8$^+$ T cell subsets using ChIP-seq, ATAC-seq and microarray analyses. **b,** Representative genes displaying bivalent modification of H3K4me3 and H3K27me3 at promoter regions during CD8+ T cell differentiation. **c,** Comparison of the percentage of genes with increased H3K27me3 at promoter regions of genes with decreased expression upon differentiation. **d,** Representative genes displaying dynamic change of enhancer H3K4me1 and active H3K27ac marks (left). Representative genes displaying a unique pattern of active H3K27ac and repressive H3K27me3 marks (right). Bar graphs showing the gene expression generated from microarray analysis. Data in (b,c) are representative of two independent experiments (n=10) and data in (d) are representative of three independent experiments (n=3).

Figure 1

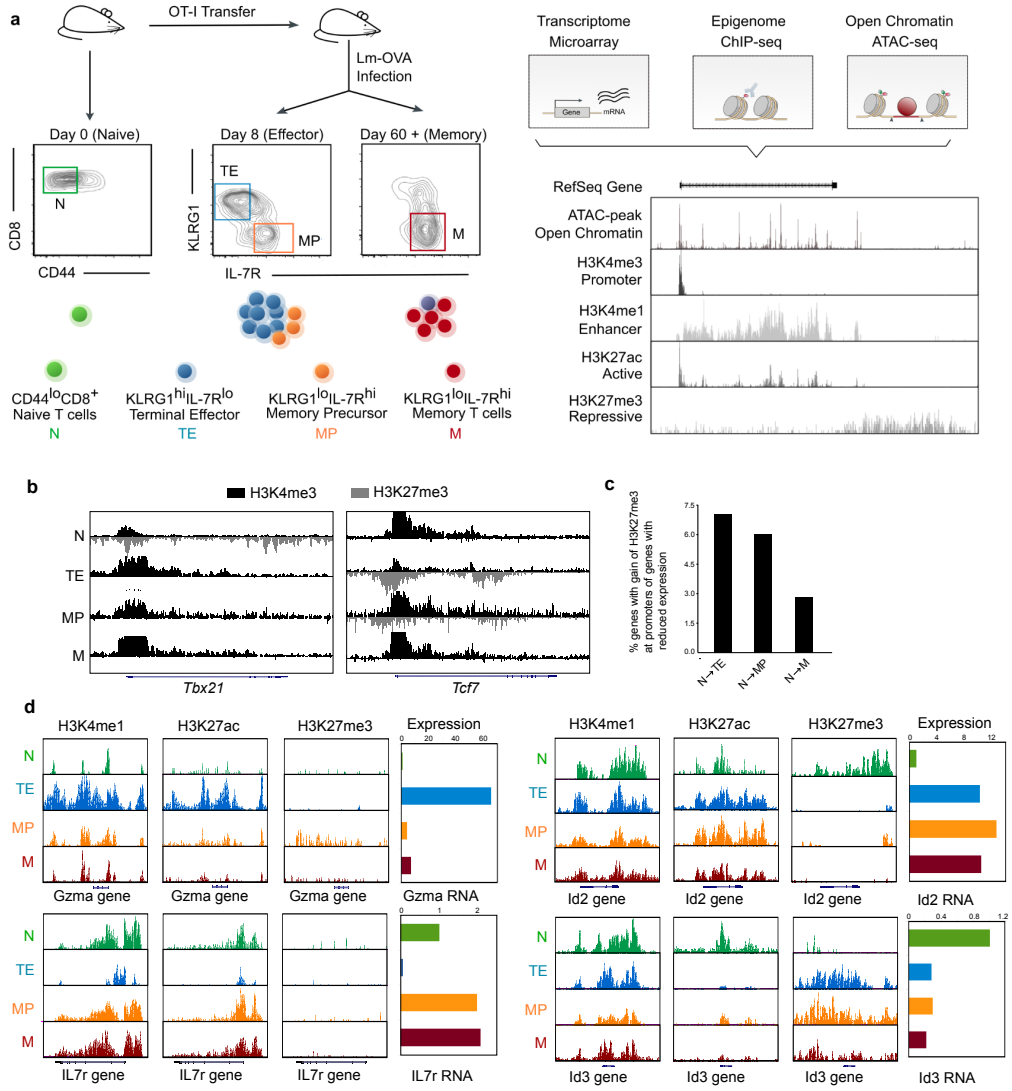

a
OT-I Transfer
Lm-OVA Infection
Day 0 (Naive)
Day 8 (Effector)
Day 60 + (Memory)
CD8
CD44
KLRG1
IL-7R
N
TE
MP
M
CD44loCD8+ Naive T cells
KLRG1hiIL-7Rlo Terminal Effector
KLRG1loIL-7Rhi Memory Precursor
KLRG1loIL-7Rhi Memory T cells
Transcriptome Microarray
Epigenome ChIP-seq
Open Chromatin ATAC-seq
RefSeq Gene
ATAC-peak Open Chromatin
H3K4me3 Promoter
H3K4me1 Enhancer
H3K27ac Active
H3K27me3 Repressive
b
H3K4me3
H3K27me3
Tbx21
Tcf7
c
% genes with gain of H3K27me3 at promoters of genes with reduced expression
N→TE
N→MP
N→M
d
H3K4me1
H3K27ac
H3K27me3
Expression
Gzma gene
Gzma RNA
Id2 gene
Id2 RNA
IL7r gene
IL7r RNA
Id3 gene
Id3 RNA

Figure 2.2: **Dynamic use of enhancers is associated with differentially expressed genes during CD8$^+$ T cell differentiation. a,** Bar plot showing the number of enhancers gained (black) and lost (gray) during differentiation from naive CD8$^+$ T cells to TE, MP and memory CD8$^+$ T cells. **b,** Heatmap of k-means clustering (k=5) of total 52331 enhancers across CD8$^+$ T cell subsets. **c,** Gene Ontology analysis of clusters in (a) using a binomial test with top 2 pathways shown (cut off as binomial P value < 0.001). **d,** Box plots showing fold change of mRNA expression of genes with the indicated number of enhancers in clusters I-III (left) and cluster IV (right) during differentiation of naive CD8$^+$ T cells to effector CD8$^+$ T cells. Data in (a-c) are representative of two independent experiments (n=10). Data in (d) are representative of three independent experiments (n=3). P value was calculated by unpaired two tailed Student's t-test: *: $p<0.0001$

Figure 2

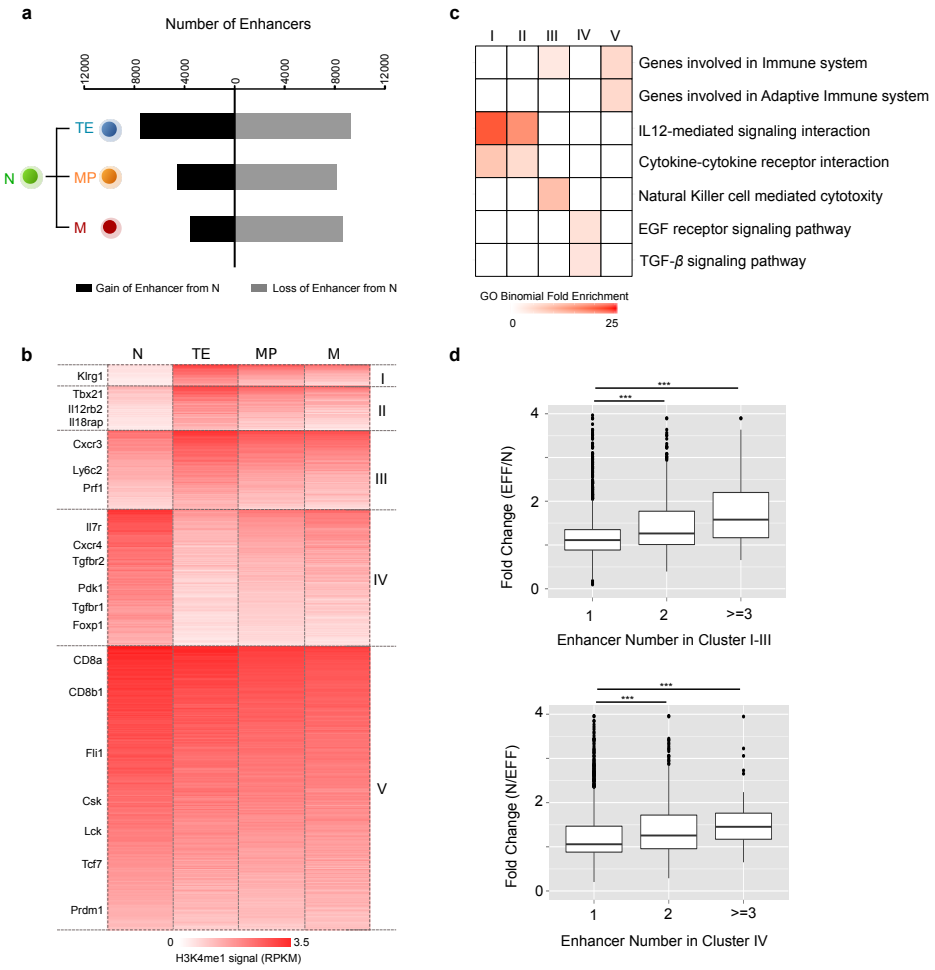

a
Number of Enhancers
12000
8000
4000
0
4000
8000
12000
TE
N
MP
M
Gain of Enhancer from N
Loss of Enhancer from N
c
I
II
III
IV
V
Genes involved in Immune system
Genes involved in Adaptive Immune system
IL12-mediated signaling interaction
Cytokine-cytokine receptor interaction
Natural Killer cell mediated cytotoxity
EGF receptor signaling pathway
TGF-β signaling pathway
GO Binomial Fold Enrichment
0
25
b
N
TE
MP
M
Klrg1
Tbx21
Il12rb2
Il18rap
Cxcr3
Ly6c2
Prf1
Il7r
Cxcr4
Tgfbr2
Pdk1
Tgfbr1
Foxp1
CD8a
CD8b1
Fli1
Csk
Lck
Tcf7
Prdm1
I
II
III
IV
V
0
3.5
H3K4me1 signal (RPKM)
d
Fold Change (EFF/N)
Enhancer Number in Cluster I-III
Fold Change (N/EFF)
Enhancer Number in Cluster IV
1
2
>=3
***

Figure 2.3: **Accessible regulatory regions allow prediction of TF regulators.** **a,** Schematic view of identification of candidate transcription factors from ATAC-seq and ChIP-seq. **b,** Representative gene displaying subset-specific accessible enhancer containing known TF motif. **c,** Heatmap showing the P-value of TF motif enrichment at subset-specific enhancers or promoters calculated by binomial test using randomly-picked open chromatin regions as background. Motif enrichment or depletion are indicated as red or blue, respectively. Known TFs that are key to effector or/and memory $CD8^+$ T cell differentiation are highlighted in red. **d,** Bar graphs showing mRNA expression of indicated TFs generated by microarray. Data in (b,c) are representative of two independent experiments (n=10). Data in (d) are representative of three independent experiments (n=3).

Figure 3

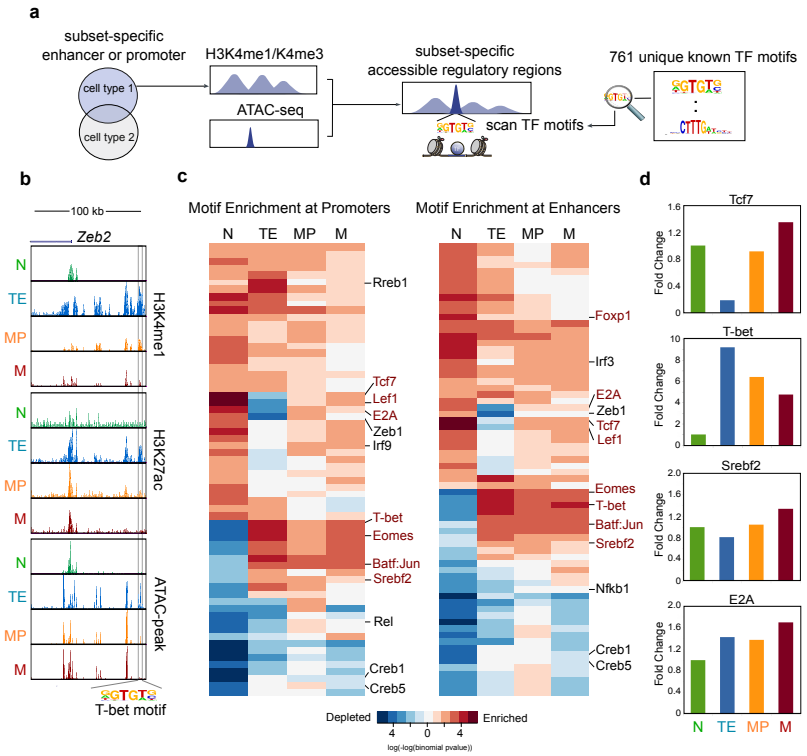

a
subset-specific enhancer or promoter
cell type 1
cell type 2
H3K4me1/K4me3
ATAC-seq
subset-specific accessible regulatory regions
scan TF motifs
761 unique known TF motifs
b
100 kb
Zeb2
N
TE
MP
M
H3K4me1
H3K27ac
ATAC-peak
T-bet motif
c
Motif Enrichment at Promoters
Motif Enrichment at Enhancers
N TE MP M
Rreb1
Tcf7
Lef1
E2A
Zeb1
Irf9
T-bet
Eomes
Batf:Jun
Srebf2
Rel
Creb1
Creb5
Foxp1
Irf3
Nfkb1
Depleted
Enriched
log(-log(binomial pvalue))
d
Tcf7
T-bet
Srebf2
E2A
Fold Change
N TE MP M

Figure 2.4: **Network analysis reveals subset-specific T-bet regulatory circuits. a,** Schematic view of network construction from ChIP-seq and ATAC-seq. **b,** Global regulatory network in the TE and MP subsets. T-bet regulated genes are highlighted in red; T-bet is labeled in blue. **c,** Comparison of T-bet regulated genes between the TE and MP subsets. **d-e,** $Tbx21^{+/+}$ and $Tbx21^{-/-}$ OT-I cells were co-transferred into recipient mice followed by Lm-OVA infection. At day 9 of infection, total $CD8^{+}$ T cells or the MP subset of $Tbx21^{+/+}$ and $Tbx21^{-/-}$ populations were sorted to and mRNA expression levels of subset-specific T-bet regulated gene targets determined by qPCR (d). The dashed line indicates that $Tbx21^{+/+}/Tbx21^{-/-}$ ratio=1. **f-g,** $Tbx21^{+/+}$ OT-I and $Tbx21^{+/-}$ OT-I cells were co-transferred into recipient mice followed by Lm-OVA infection. At day 8 of infection, TE and MP subset of $Tbx21^{+/+}$ and $Tbx21^{+/-}$ were sorted to measure RNA expression of subset-specific T-bet regulated gene targets (f). The dashed line indicates that $Tbx21^{+/+}/Tbx21^{+/-}$ ratio=1. **h,** Kinetic analysis of the percentage of MP of $Tbx21^{+/+}$ and $Tbx21^{-/-}$ cells during Lm-OVA infection. Data in (e-h) are representative of two independent experiments (n=3 in (d-g), n=4 in (h)). P value was calculated by paired two tailed Student's t-test: n.s. *: $p<0.05$; **: $p<0.01$

Figure 4

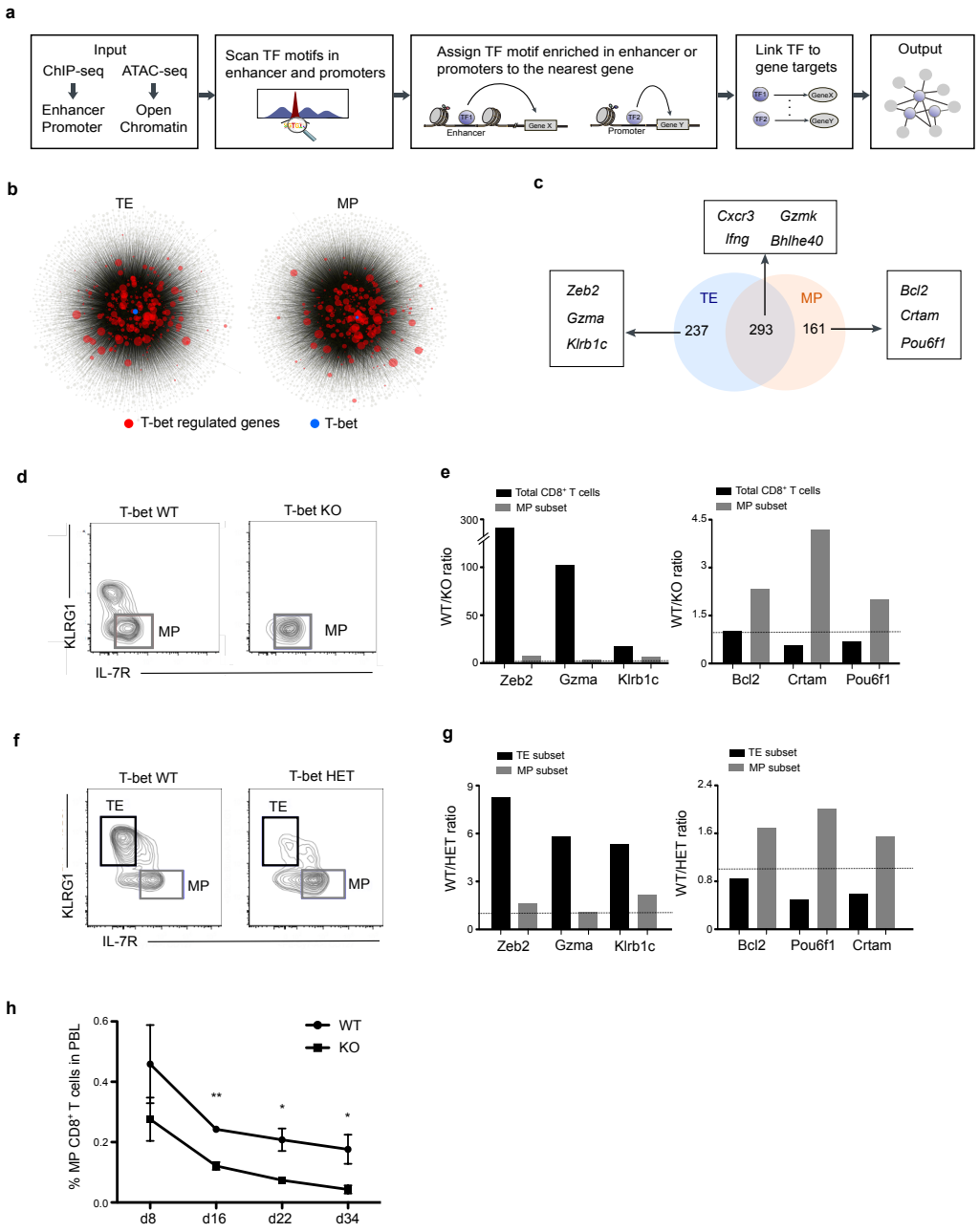

a
Input
ChIP-seq
ATAC-seq
Enhancer
Promoter
Open
Chromatin
Scan TF motifs in enhancer and promoters
Assign TF motif enriched in enhancer or promoters to the nearest gene
Enhancer
Promoter
Link TF to gene targets
Output
b
TE
MP
T-bet regulated genes
T-bet
c
Cxcr3
Gzmk
Ifng
Bhlhe40
Zeb2
Gzma
Klrb1c
TE
MP
237
293
161
Bcl2
Crtam
Pou6f1
d
T-bet WT
T-bet KO
KLRG1
IL-7R
MP
e
Total CD8+ T cells
MP subset
WT/KO ratio
Zeb2
Gzma
Klrb1c
Bcl2
Crtam
Pou6f1
f
T-bet WT
T-bet HET
TE
MP
KLRG1
IL-7R
g
TE subset
MP subset
WT/HET ratio
Zeb2
Gzma
Klrb1c
Bcl2
Pou6f1
Crtam
h
WT
KO
% MP CD8+ T cells in PBL
d8
d16
d22
d34

Figure 2.5: **PageRank-based TF ranking highlights key TF candidates.** **a,** Schematic view of PageRank-based TF motif ranking. The size of circles in the network represents the importance of gene targets which are assessed by relative expression across different cell types generated by microarray data. **b,** Comparison of PageRank analysis with motif enrichment analysis in Figure 3. Bar graph showing the percentage of known TFs reported previously recovered from predicted TF candidates for each analysis. **c,** Heatmap of PageRank fold enrichment of TFs across $CD8^+$ T cell subsets. Data in (c) is based on network generated from reproducible peaks of two independent ATAC-seq and ChIP-seq experiments (n=10).

Figure 5

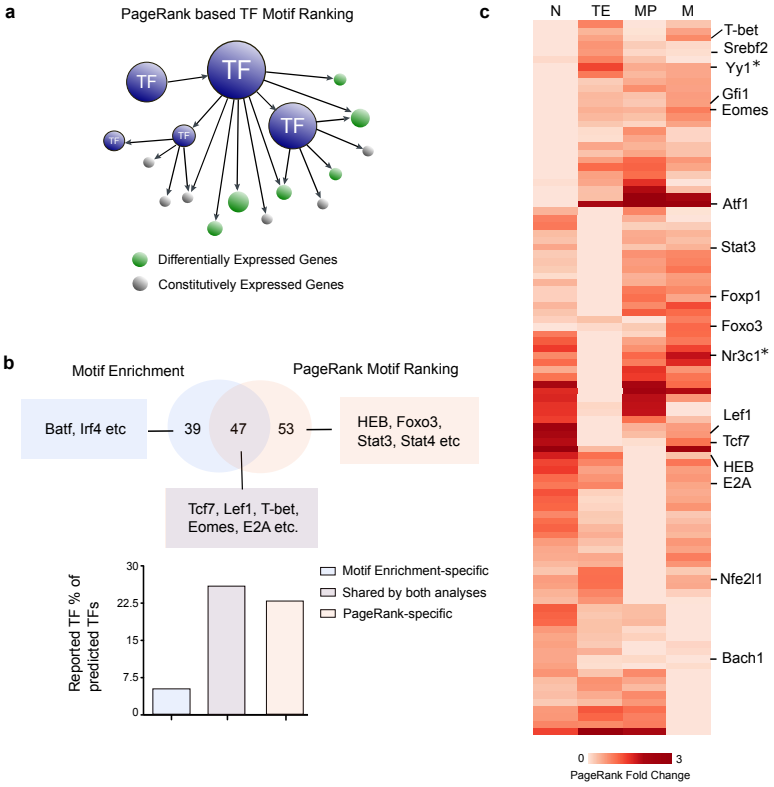

a
PageRank based TF Motif Ranking
TF
TF
TF
TF
TF
Differentially Expressed Genes
Constitutively Expressed Genes
b
Motif Enrichment
PageRank Motif Ranking
Batf, Irf4 etc
39
47
53
HEB, Foxo3, Stat3, Stat4 etc
Tcf7, Lef1, T-bet, Eomes, E2A etc.
Reported TF % of predicted TFs
30
22.5
15
7.5
0
Motif Enrichment-specific
Shared by both analyses
PageRank-specific
c
N
TE
MP
M
T-bet
Srebf2
Yy1*
Gfi1
Eomes
Atf1
Stat3
Foxp1
Foxo3
Nr3c1*
Lef1
Tcf7
HEB
E2A
Nfe2l1
Bach1
0
3
PageRank Fold Change

Figure 2.6: **YY1 is a transcriptional regulator of terminal-effector CD8$^+$ T cells differentiation. a,** Schematic view of experimental design. OT-I CD8$^+$ T cells were in vitro activated and transduced with retrovirus containing control shRNA shCD19 or shYY1 for 24 hours, co-transfered into recipient mice followed by i.v. infection with Lm-OVA. Splenocytes were isolated and analyzed on day 7 of infection. **b,** mRNA expression of YY1 quantified by RT-qPCR after YY1 shRNA knockdown in CD8$^+$ T cells after 72 hours activation in vitro. **c,** Flow cytometric analysis of KLRG1 and CD127 expression for cells transudeced with shCD19 and shYY1. **d,** The percentage and the number of terminal-effector and memory-precursor CD8$^+$ T cells after knockdown of YY1. **e,** Histogram (left) and MFI expression (right) of CD27, CXCR3 and Tcf7 after knockdown of YY1. **f,** Flow cytometric analysis of the frequency and number of IFNγ producing cells and IFNγ MFI and the number of TNFα producing cells using intracellular cytokine staining of isolated splenocytes restimulated by OVA peptide for 4 hours. Data are representative of two (b,e,f) or three (c,d) experiments with three mice per group. For comparision of two groups, two tailed student't t-test was performed. *: $p<0.05$; **: $p<0.001$; ***: $p<0.0001$

Figure 6

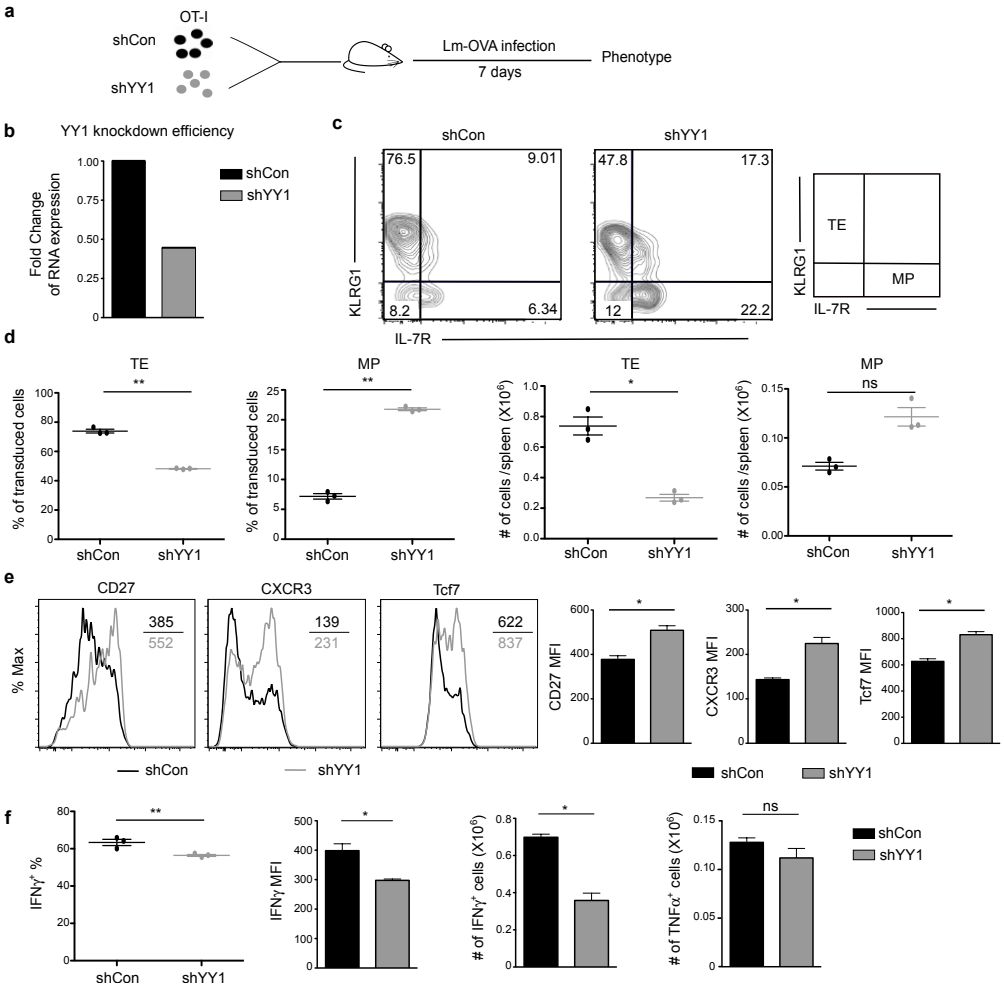

a
OT-I
shCon
shYY1
Lm-OVA infection
7 days
Phenotype
b
YY1 knockdown efficiency
Fold Change of RNA expression
shCon
shYY1
c
shCon
shYY1
76.5
9.01
8.2
6.34
47.8
17.3
12
22.2
KLRG1
IL-7R
TE
MP
d
TE
MP
% of transduced cells
# of cells /spleen (X10^6)
ns
e
CD27
CXCR3
Tcf7
% Max
385
552
139
231
622
837
CD27 MFI
CXCR3 MFI
Tcf7 MFI
f
IFNγ+ %
IFNγ MFI
# of IFNγ+ cells (X10^6)
# of TNFα+ cells (X10^6)

Figure 2.7: **Glucocoticoid receptor Nr3c1 is essential for the formation and maturation of CD8$^+$ memory T cells. a,** mRNA expression of Nr3c1 quantified by RT-qPCR after YY1 shRNA knockdown in CD8$^+$ T cells after 72 hours activation in vitro. **b-d,** OT-I CD8$^+$ T cells were in vitro activated and transduced with retrovirus containing control shRNA shCD19 or shNr3c1 for 24 hours, co-transfered into recipient mice followed by i.v. infection with Lm-OVA. Splenocytes were isolated and analyzed on day 7 of infection. Flow cytometric analysis of KLRG1 and CD127 expression shown in (b) and histogram and MFI of CD27, CXCR3 and Tcf7 after knockdown of Nr3c1 shown in (d). The frequency and number of TE and MP CD8$^+$ T cells after knockdown of Nr3c1 shown in (c). **e,** Kinetic analysis of MP subset frequency after knockdown of Nr3c1. **f,** Flow cytometric analysis of KLRG1 and IL-7R expression of shRNA transduced cells in the spleen on day 30 of Lm-OVA infection. Data are representative of two independent experiments (n=3 in (a-d), n=5 in (e)). For comparison of two groups, two tailed paired Student's t-test was performed. n.s. *: $p<0.05$; **: $p<0.01$; ***: $p<0.001$

Figure 7

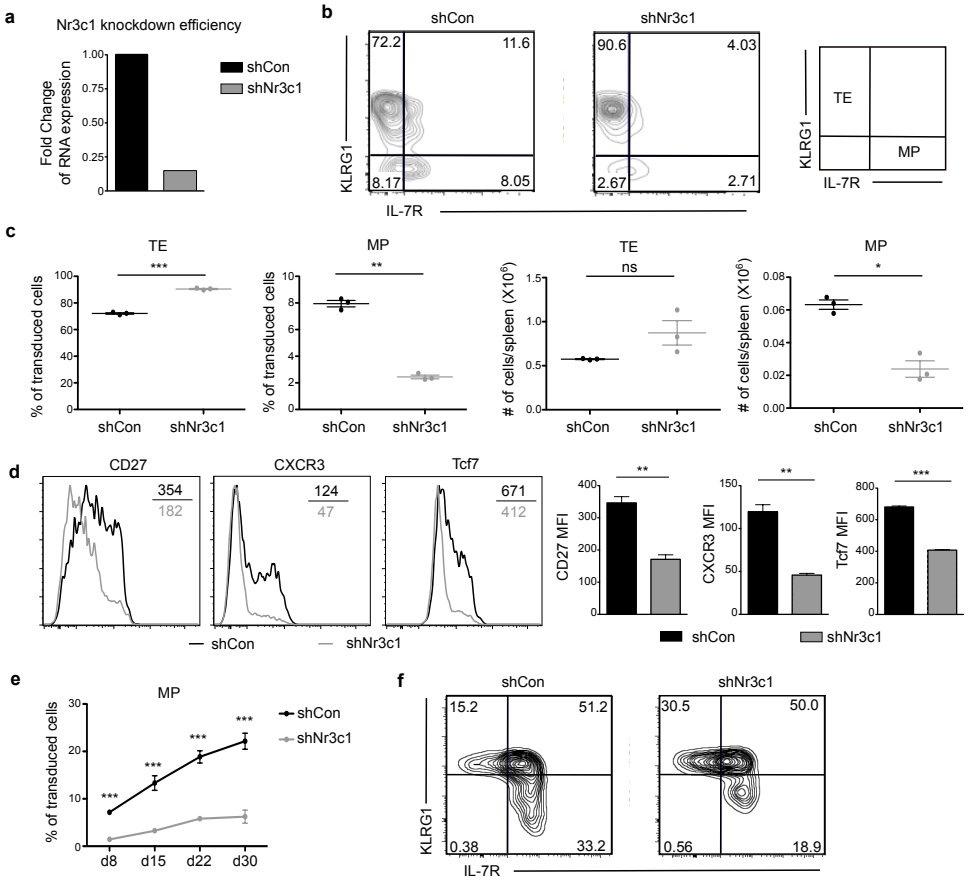

a
Nr3c1 knockdown efficiency
Fold Change of RNA expression
shCon
shNr3c1
b
shCon
shNr3c1
72.2
11.6
8.17
8.05
90.6
4.03
2.67
2.71
KLRG1
IL-7R
TE
MP
c
TE
MP
% of transduced cells
# of cells/spleen (X10^6)
***
**
ns
*
shCon
shNr3c1
d
CD27
CXCR3
Tcf7
354
182
124
47
671
412
CD27 MFI
CXCR3 MFI
Tcf7 MFI
e
MP
% of transduced cells
d8
d15
d22
d30
f
15.2
51.2
0.38
33.2
30.5
50.0
0.56
18.9

Figure 2.8: **Transcriptional program of TE and MP CD8$^+$ T cell subsets.** **a,** Comparison of gene expression of TE and MP CD8$^+$ T cell subsets by microarray. Genes that are 1.5-fold upregulated in TE or MP CD8$^+$ T cell subsets are highlighted as blue or orange, respectively. Transcripts did not differ significantly in expression following correction with FDR therefore a fold-change cut-off of 1.5-fold was used for comparisons. **b,** Volcano plots of the comparison of total effector and memory CD8$^+$ T cells highlighting TE- or MP-enriched genes. Numbers in bottom corners indicate the number of highlighted genes in that region. **c,** The ratio of gene expression of known TFs in TE versus MP subset from microarray. **d,** Histograms of protein abundance of key TFs in TE versus MP subset.

Supplementary Figure 1

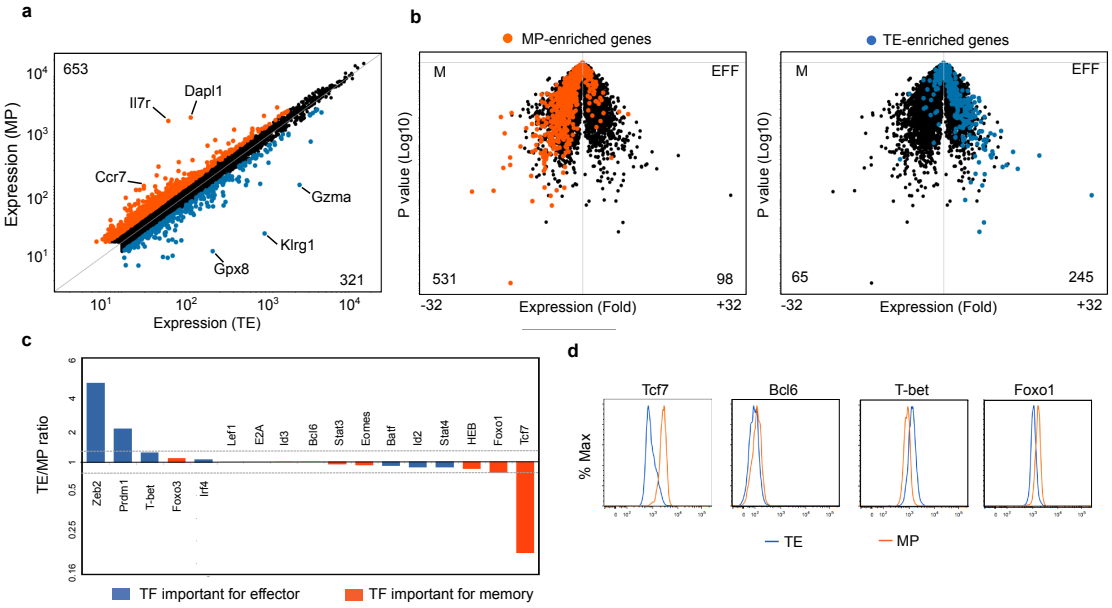


Figure 2.9: **Dynamic enhancer establishment is associated with gene expression during CD8$^+$ T cell differentiation.** **a,** Violin plots showing the expression of genes associated with different enhancer clusters generated from Figure 2a in naive (N), total effector (EFF) and memory (M) CD8+ T cells generated from microarray data in Best *et al.* study[19]. **b,** Volcano plots of the comparison of TE and MP CD8$^+$ T cells showing expression of enhancer cluster associated genes. Data are representative of three independent experiments with three mice per group (median value). The statistical analysis was performed by a nonparametric Wilcoxon rank-sum test. n.s. * : p value <0.0001.

Supplementary Figure 2

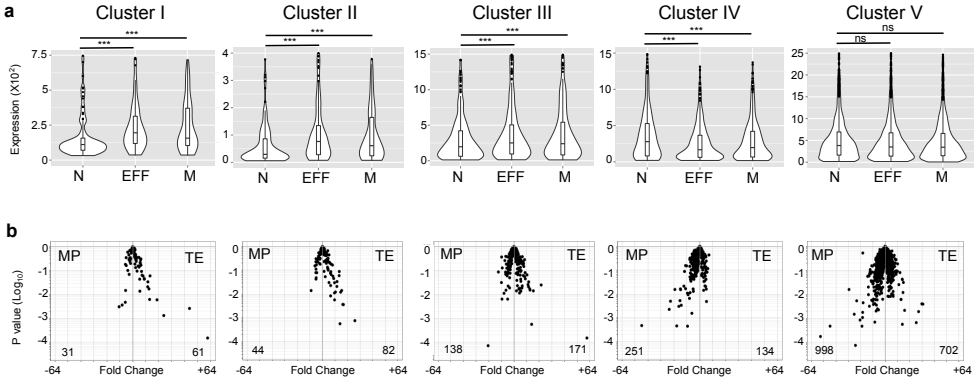


Figure 2.10: **A full list of TF motifs enriched in subset-specific regulatory elements. a,** Venn diagram showing the overlap of enhancers between $CD8^+$ T cell subsets. **b,** Heatmap showing the p-value of transcription factor motif enrichment at subset-specific promoters (left) or enhancers (right) calculated by binomial test using randomly-picked open chromatin regions as background. Motif enrichment or depletion are indicated as red or blue, respectively.

Supplementary Figure 3

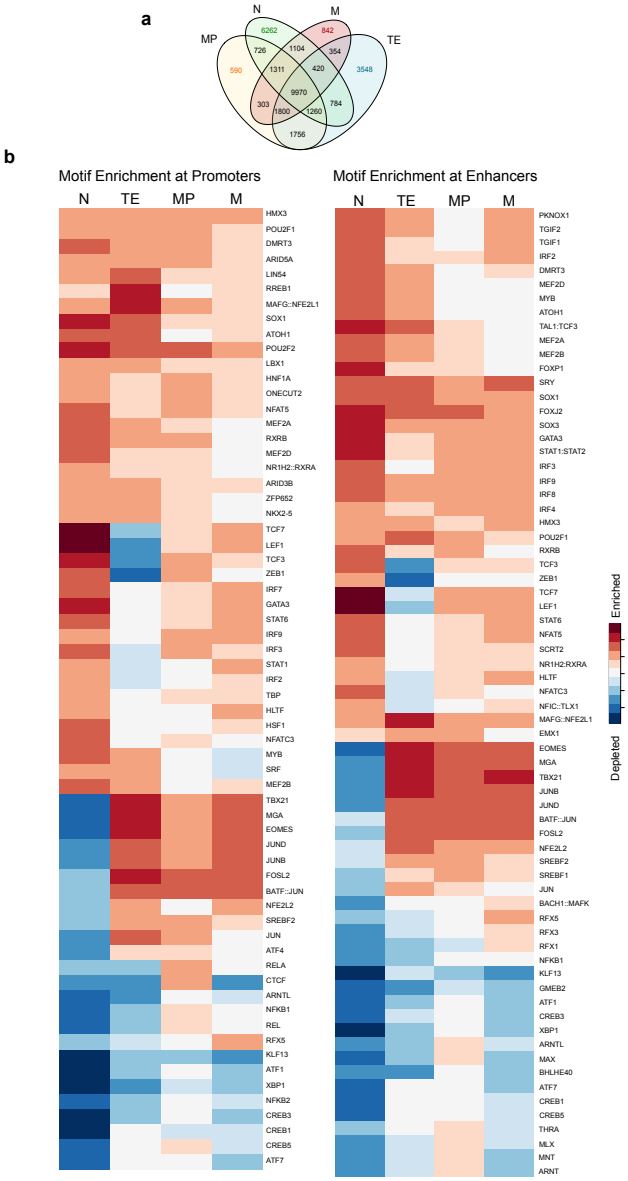


Figure 2.11: **A full list of TFs identified by PageRank. a,** heatmap showing PageRank fold enrichment of TFs across $CD8^+$ T cell subsets. **b,** A list of TFs revealed by PageRank analysis and motif enrichment in Figure 3. Known TFs important for $CD8^+$ T cell differentiation are highlighted in red. **c,** Bar graphs showing the fold change of YY1 and Nr3c1 gene expression generated from microarray. Data in (e) are the mean value of gene expression from three independent experiments with pooled spleens from three mice.

Supplementary Figure 4

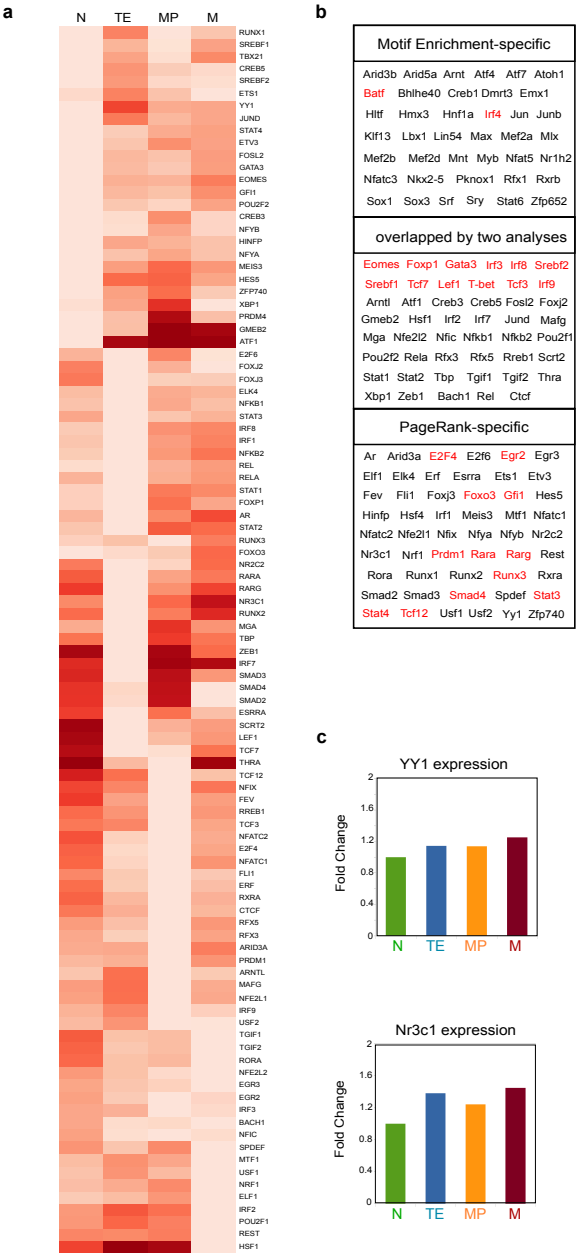
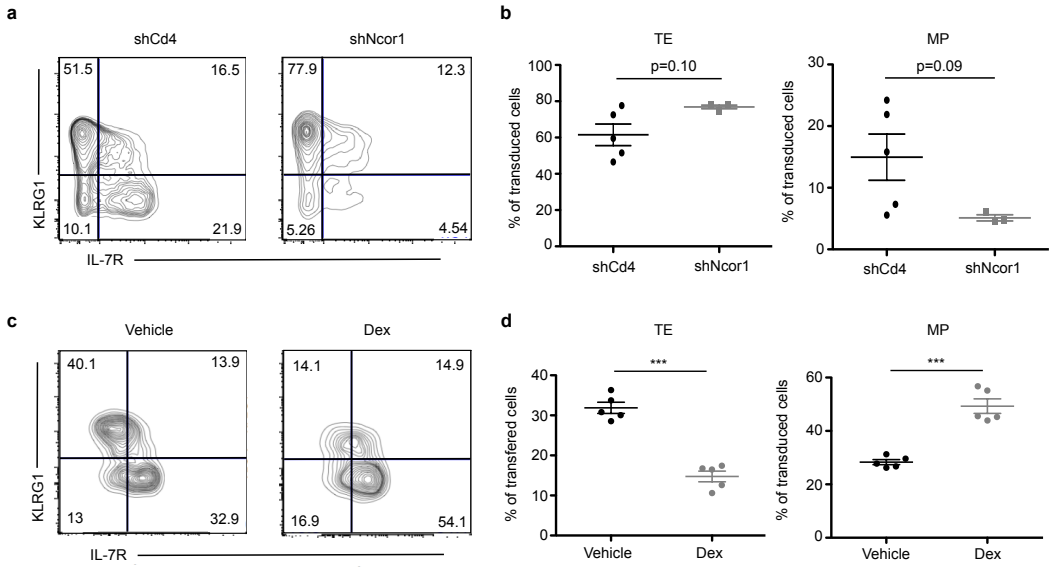


Figure 2.12: **Ablation of Nr3c1 cofactor Ncor1 and treatment with dexamethasone affect MP CD8$^+$ T cell differentiation.** **a,** Flow cytometric analysis of KLRG1 and IL-7R expression for cells transduced with shCd4 and shNcor1 in PBL on day 8 of infection. **b,** The percentage of TE and MP CD8$^+$ T cells gated on transduced cells on day 8 of infection after knockdown of Ncor1. **c,** Flow cytometric analysis of KLRG1 and IL-7R expression for donor cells in mice treated with either vehicle or dexamethasone for 7 days. **d,** The percentage of TE and MP CD8$^+$ T cells gated on donor cells on day 8 of Lm-OVA infection after drug treatment. Data are representative from two independent experiments with 5 mice per group. The statistical analysis was performed by two-tailed paired t-test in (b) and two-tailed unpaired t-test in (d). *: $p<0.001$

Supplementary Figure 6



2.7 References

[1] Xinyuan Zhou, Shuyang Yu, Dong-Mei Zhao, John T Harty, Vladimir P Badovinac, and Hai-Hui Xue. Differentiation and Persistence of Memory CD8(+) T Cells Depend on T Cell Factor 1. *Immunity*, 33(2):229–240, 2010. ISSN 1097-4180. doi: 10.1016/j.immuni.2010.08.002.

[2] Nikhil S. Joshi, Weiguo Cui, Anmol Chandele, Heung Kyu Lee, David R. Urso, James Hagman, Laurent Gapin, and Susan M. Kaech. Inflammation Directs Memory Precursor and Short-Lived Effector CD8+ T Cell Fates via the Graded Expression of T-bet Transcription Factor. *Immunity*, 27(2):281–295, 2007. ISSN 10747613. doi: 10.1016/j.immuni.2007.07.010.

[3] Susan M Kaech, Joyce T Tan, E John Wherry, Bogumila T Konieczny, Charles D Surh, and Rafi Ahmed. Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nature immunology*, 4(12):1191–8, 2003. ISSN 1529-2908. doi: 10.1038/ni1009.

[4] John T Chang, E John Wherry, and Ananda W Goldrath. Molecular regulation of effector and memory T cell differentiation. *Nature reviews. Immunology*, 15(12), 2014. doi: 10.1038/ni.3031.1104.

[5] Jason D Buenrostro, Paul G Giresi, Lisa C Zaba, Howard Y Chang, and William J Greenleaf. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods*, 10(12):1213–8, 2013. ISSN 1548-7105. doi: 10.1038/nmeth.2688.

[6] Deborah R Winter, Steffen Jung, and Ido Amit. Making the case for chromatin profiling: a new tool to investigate the immune-regulatory landscape. *Nature reviews. Immunology*, 15(9):585–94, 2015. ISSN 1474-1741. doi: 10.1038/nri3884.

[7] Shane Neph, Andrew B. Stergachis, Alex Reynolds, Richard Sandstrom, Elhanan Borenstein, and John A. Stamatoyannopoulos. Circuitry and dynamics of human transcription factor regulatory networks. *Cell*, 150(6):1274–1286, 2012. ISSN 00928674. doi: 10.1016/j.cell.2012.04.040.

[8] Yin Shen, Feng Yue, David F. McCleary, Zhen Ye, Lee Edsall, Samantha Kuan, Ulrich Wagner, Jesse Dixon, Leonard Lee, Victor V. Lobanenkov, and Bing Ren. A map of the cis-regulatory sequences in the mouse genome. *Nature*, 488(7409):116–120, 2012. ISSN 0028-0836. doi: 10.1038/nature11243.

[9] Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics*, 13(9):613–626, 2012. ISSN 1471-0056. doi: 10.1038/nrg3207.

[10] Eliezer Calo and Joanna Wysocka. Modification of Enhancer Chromatin: What, How, and Why?, 2013. ISSN 10972765.

[11] D. Lara-Astiaso, A. Weiner, E. Lorenzo-Vivas, I. Zaretsky, D. A. Jaitin, E. David, H. Keren-Shaul, A. Mildner, D. Winter, S. Jung, N. Friedman, and I. Amit. Chromatin state dynamics during blood formation. *Science*, 345 (6199):943–9, 2014. ISSN 0036-8075. doi: 10.1126/science.1256271.

[12] Yonit Lavin, Deborah Winter, Ronnie Blecher-Gonen, Eyal David, Hadas Keren-Shaul, Miriam Merad, Steffen Jung, and Ido Amit. Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell*, 159(6):1312–1326, 2014. ISSN 10974172. doi: 10.1016/j.cell.2014.11.018.

[13] Yasuto Araki, Zhibin Wang, Chongzhi Zang, William H. Wood, Dustin Schones, Kairong Cui, Tae Young Roh, Brad Lhotsky, Robert P. Wersto, Weiqun Peng, Kevin G. Becker, Keji Zhao, and Nan ping Weng. Genome-wide Analysis of Histone Methylation Reveals Chromatin State-Based Regulation of Gene Transcription and Function of Memory CD8+ T Cells. *Immunity*, 30 (6):912–925, 2009. ISSN 10747613. doi: 10.1016/j.immuni.2009.05.006.

[14] Brendan E. Russ, Moshe Olshanksy, Heather S. Smallwood, Jasmine Li, Alice E. Denton, Julia E. Prier, Angus T. Stock, Hayley A. Croom, Jolie G. Cullen, Michelle L T Nguyen, Stephanie Rowe, MatthewR R. Olson, David B. Finkelstein, Anne Kelso, Paul G. Thomas, Terry P. Speed, Sudha Rao, and Stephen J. Turner. Distinct epigenetic signatures delineate transcriptional programs during virus-specific CD8+ T cell differentiation. *Immunity*, 41(5): 853–865, 2014. ISSN 10974180. doi: 10.1016/j.immuni.2014.11.001.

[15] Nathaniel D Heintzman, Gary C Hon, R David Hawkins, Pouya Kheradpour,

Alexander Stark, Lindsey F Harp, Zhen Ye, Leonard K Lee, Rhona K Stuart, Christina W Ching, Keith a Ching, Jessica E Antosiewicz-Bourget, Hui Liu, Xinmin Zhang, Roland D Green, Victor V Lobanenkov, Ron Stewart, James a Thomson, Gregory E Crawford, Manolis Kellis, and Bing Ren. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, 459(7243):108–112, 2009. ISSN 0028-0836. doi: 10.1038/nature07829.

[16] R. David Hawkins, Antti Larjo, Subhash K. Tripathi, Ulrich Wagner, Ying Luu, Tapio Lnnberg, Sunil K. Raghav, Leonard K. Lee, Riikka Lund, Bing Ren, Harri Lhdesmki, and Riitta Lahesmaa. Global Chromatin State Analysis Reveals Lineage-Specific Enhancers during the Initiation of Human T helper 1 and T helper 2 Cell Polarization. *Immunity*, 38(6):1271–1284, 2013. ISSN 10747613. doi: 10.1016/j.immuni.2013.05.011.

[17] Golnaz Vahedi, Hayato Takahashi, Shingo Nakayamada, Hong Wei Sun, Vittorio Sartorelli, Yuka Kanno, and John J. OShea. STATs shape the active enhancer landscape of T cell populations. *Cell*, 151(5):981–993, 2012. ISSN 00928674. doi: 10.1016/j.cell.2012.09.044.

[18] Michael Bulger and Mark Groudine. Functional and mechanistic diversity of distal transcription enhancers, 2011. ISSN 00928674.

[19] J Adam Best. Transcriptional insights into the CD8+ T cell response to infection and memory T cell formation. *Nature immunology*, 29(4):997–1003, 2013. ISSN 15378276. doi: 10.1016/j.biotechadv.2011.08.021.Secreted.

[20] Menno P Creyghton, Albert W Cheng, G Grant Welstead, Tristan Kooistra, Bryce W Carey, Eveline J Steine, Jacob Hanna, Michael A Lodato, Garrett M Frampton, Phillip A Sharp, Laurie A Boyer, Richard A Young, and Rudolf Jaenisch. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America*, 107(50):21931–21936, 2010. ISSN 1091-6490. doi: 10.1073/pnas.1016071107.

[21] Cliff Y Yang, J Adam Best, Jamie Knell, Edward Yang, Alison D Sheridan, Adam K Jesionek, Haiyan S Li, Richard R Rivera, Kristin Camfield Lind, Louise M D'Cruz, Stephanie S Watowich, Cornelis Murre, and Ananda W Goldrath. The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8+ T cell subsets. *Nature immunology*, 12(12):1221–9,

2011. ISSN 1529-2916. doi: 10.1038/ni.2158.

[22] Masaki Miyazaki, Richard R. Rivera, Kazuko Miyazaki, Yin C. Lin, Yasutoshi Agata, and Cornelis Murre. The opposing roles of E2A and Id3 that orchestrate and enforce the naïve T cell fate. *Nature Immunology*, 12(10):992–1001, 2012. ISSN 1878-5832. doi: 10.1016/j.micinf.2011.07.011.Innate.

[23] Nisha Rajagopal, Wei Xie, Yan Li, Uli Wagner, Wei Wang, John Stamatoyannopoulos, Jason Ernst, Manolis Kellis, and Bing Ren. RFECS: A Random-Forest Based Algorithm for Enhancer Identification from Chromatin State. *PLoS Computational Biology*, 9(3), 2013. ISSN 1553734X. doi: 10.1371/journal.pcbi.1002968.

[24] Julie Chaix, Simone A Nish, Wen-Hsuan W Lin, Nyanza J Rothman, Lei Ding, E John Wherry, and Steven L Reiner. Cutting edge: CXCR4 is critical for CD8+ memory T cell homeostatic self-renewal but not rechallenge self-renewal. *Journal of immunology (Baltimore, Md. : 1950)*, 193(3):1013–6, 2014. ISSN 1550-6606. doi: 10.4049/jimmunol.1400488.

[25] Cory Y McLean, Dave Bristor, Michael Hiller, Shoa L Clarke, Bruce T Schaar, Craig B Lowe, Aaron M Wenger, and Gill Bejerano. GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology*, 28(5): 495–501, 2010. ISSN 1546-1696. doi: 10.1038/nbt.1630.

[26] Chaoyu Ma and Nu Zhang. Transforming growth factor- β signaling is constantly shaping memory T-cell population. *Pnas*, 2015(35):11013–11017, 2015. ISSN 10916490. doi: 10.1073/pnas.1510119112.

[27] Kyla D Omilusik, J Adam Best, Bingfei Yu, Steven Goossens, Alexander Weidemann, Jessica V Nguyen, Eve Seuntjens, Agata Stryjewska, Christiane Zweier, Rahul Roychoudhuri, Luca Gattinoni, Lynne M Bird, Yujiro Higashi, Hisato Kondoh, Danny Huylebroeck, Jody Haigh, and Ananda W Goldrath. Transcriptional repressor ZEB2 promotes terminal differentiation of CD8+ effector and memory T cell populations during infection. *The Journal of experimental medicine*, pages jem.20150194–, 2015. ISSN 1540-9538. doi: 10.1084/jem.20150194.

[28] Claudia X Dominguez, Robert A Amezquita, Tianxia Guan, Heather D Mar-

shall, Nikhil S Joshi, Steven H Kleinstein, and Susan M Kaech. The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. *The Journal of experimental medicine*, 212(12):2041–2056, 2015. ISSN 1540-9538. doi: 10.1084/jem.20150186.

[29] Andrew M Intlekofer, Naofumi Takemoto, E John Wherry, Sarah a Longworth, John T Northrup, Vikram R Palanivel, Alan C Mullen, Christopher R Gasink, Susan M Kaech, Joseph D Miller, Laurent Gapin, Kenneth Ryan, Andreas P Russ, Tullia Lindsten, Jordan S Orange, Ananda W Goldrath, Rafi Ahmed, and Steven L Reiner. Effector and memory CD8+ T cell fate coupled by T-bet and eomesodermin. *Nature Immunology*, 6(12):1236–1244, 2005. ISSN 1529-2908. doi: 10.1038/ni1268.

[30] Yoko Kidani, Heidi Elsaesser, M Benjamin Hock, Laurent Vergnes, Kevin J Williams, Joseph P Argus, Beth N Marbois, Evangelia Komisopoulou, Elizabeth B Wilson, Timothy F Osborne, Thomas G Graeber, Karen Reue, David G Brooks, and Steven J Bensinger. Sterol regulatory element-binding proteins are essential for the metabolic programming of effector T cells and adaptive immunity. *Nature immunology*, 14(5):489–99, 2013. ISSN 1529-2916. doi: 10.1038/ni.2570.

[31] Makoto Kurachi, R Anthony Barnitz, Nir Yosef, Pamela M Odorizzi, Michael A DiIorio, Madeleine E Lemieux, Kathleen Yates, Jernej Godec, Martin G Klatt, Aviv Regev, E John Wherry, and W Nicholas Haining. The transcription factor BATF operates as an essential differentiation checkpoint in early effector CD8+ T cells. *Nature immunology*, 15(4):373–83, 2014. ISSN 1529-2916. doi: 10.1038/ni.2834.

[32] Bryan R Fonslow, Benjamin D Stein, Kristofor J Webb, Tao Xu, Jeong Choi, Sung Kyu, and John R Yates Iii. Generation of memory precursors and functional memory CD8+ T cells depends on TCF-1 and LEF-1. 10(1):54–56, 2013. ISSN 1751-7370. doi: 10.1038/nmeth.2250.Digestion.

[33] Louise M. D'Cruz, Kristin Camfield Lind, Bei Bei Wu, Jessica K. Fujimoto, and Ananda W. Goldrath. Loss of E protein transcription factors E2A and HEB delays memory-precursor formation during the CD8 + T-cell immune response. *European Journal of Immunology*, 42(8):2031–2041, 2012. ISSN 00142980. doi: 10.1002/eji.201242497.

[34] M Rincón and R a Flavell. AP-1 transcriptional activity requires both T-cell receptor-mediated and co-stimulatory signals in primary T lymphocytes. *The EMBO journal*, 13(18):4370–4381, 1994. ISSN 02614189.

[35] Xunshan Ding, Jamie Boney-montoya, Bryn M Owen, Angie L Bookout, Colbert Coate, David J Mangelsdorf, and Steven A Kliewer. Network Analysis Reveals Centrally Connected Genes and Pathways Involved in CD8+ T Cell Exhaustion versus Memory. 16(3):387–393, 2013. ISSN 1878-5832. doi: 10.1016/j.cmet.2012.08.002.

[36] Guangan Hu and Jianzhu Chen. A genome-wide regulatory network identifies key transcription factors for memory CD8 T-cell development. *Nature communications*, 4:2830, 2013. ISSN 2041-1723. doi: 10.1038/ncomms3830.

[37] Nikhil S Joshi, Weiguo Cui, Claudia X Dominguez, Jonathan H Chen, Timothy W Hand, and Susan M Kaech. Increased numbers of preexisting memory CD8 T cells and decreased T-bet expression can restrain terminal differentiation of secondary effector and memory CD8 T cells. *The Journal of Immunology*, 187(8):4068–4076, 2011. ISSN 1550-6606. doi: 10.4049/jimmunol.1002145.

[38] Susanne J Szabo, Brandon M Sullivan, Claudia Stemmann, Abhay R Satoskar, Barry P Sleckman, and Laurie H Glimcher. Distinct Effects of T-bet in TH1 Lineage Commitment and IFN-γ Production in CD4 and CD8 T Cells. *Science*, 295(5553):338–342, 2002. ISSN 1095-9203. doi: 10.1126/science.1065543.

[39] Graham M. Lord, Ravi M. Rao, Hyeryun Choe, Brandon M. Sullivan, Andrew H. Lichtman, F. William Luscinskas, and Laurie H. Glimcher. T-bet is required for optimal proinflammatory CD4+ T-cell trafficking. *Blood*, 106 (10):3432–3439, 2005. ISSN 00064971. doi: 10.1182/blood-2005-04-1393.

[40] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank Citation Ranking: Bringing Order to the Web. *World Wide Web Internet And Web Information Systems*, 54(1999-66):1–17, 1998. ISSN 1752-0509. doi: 10.1.1.31.1768.

[41] Weiguo Cui, Ying Liu, Jason S. Weinstein, Joseph Craft, and Susan M.

Kaech. An interleukin-21- interleukin-10-STAT3 pathway is critical for functional maturation of memory CD8 + T cells. *Immunity*, 35(5):792–805, 2011. ISSN 10747613. doi: 10.1016/j.immuni.2011.09.017.

[42] K M Galvin and Y Shi. Multiple mechanisms of transcriptional repression by YY1. *Molecular and cellular biology*, 17(7):3723–3732, 1997. ISSN 0270-7306.

[43] Soo Seok Hwang, Sung Woong Jang, Min Kyung Kim, Lark Kyun Kim, Bong-Sung Kim, Hyeong Su Kim, Kiwan Kim, Wonyong Lee, Richard A. Flavell, and Gap Ryol Lee. YY1 inhibits differentiation and function of regulatory T cells by blocking Foxp3 expression and activity. *Nature Communications*, 7: 10789, 2016. ISSN 2041-1723. doi: 10.1038/ncomms10789.

[44] Guangchao Sui, El Bachir Affar, Yujiang Shi, Chrystelle Brignone, Nathan R. Wall, Peng Yin, Mary Donohoe, Margaret P. Luke, Dominica Calvo, Steven R. Grossman, and Yang Shi. Yin Yang 1 is a negative regulator of p53. *Cell*, 117 (7):859–872, 2004. ISSN 00928674. doi: 10.1016/j.cell.2004.06.004.

[45] Huifei Liu, Marc Schmidt-Supprian, Yujiang Shi, Elias Hobeika, Natasha Barteneva, Hassan Jumaa, Roberta Pelanda, Michael Reth, Jane Skok, Klaus Rajewsky, and Yang Shi. Yin Yang 1 is a critical regulator of B-cell development. *Genes & development*, 21:1179–1189, 2007. ISSN 0890-9369. doi: 10.1101/gad.1529307.

[46] M. J. Herold, K. G. McPherson, and H. M. Reichardt. Glucocorticoids in T cell apoptosis and function, 2006. ISSN 1420682X.

[47] Denis Franchimont, Jérôme Galon, Melanie S Vacchio, Samuel Fan, Roberta Visconti, David M Frucht, Vincent Geenen, George P Chrousos, Jonathan D Ashwell, and John J O'Shea. Positive effects of glucocorticoids on T cell function by up-regulation of IL-7 receptor alpha. *Journal of immunology (Baltimore, Md. : 1950)*, 168:2212–2218, 2002. ISSN 0022-1767. doi: 10.4049/jimmunol.168.5.2212.

[48] Kathleen A. Smoak and John A. Cidlowski. Mechanisms of glucocorticoid receptor signaling during inflammation, 2004. ISSN 00476374.

[49] Qi Wang, John A Blackford, Liang-nian Song, Ying Huang, Sehyung Cho, and S Stoney Simons. Equilibrium interactions of corepressors and coactivators with agonist and antagonist complexes of glucocorticoid receptors. *Molecular endocrinology (Baltimore, Md.)*, 18(6):1376–1395, 2004. ISSN 0888-8809. doi: 10.1210/me.2003-0421.

[50] Alexander M Tsankov, Hongcang Gu, Veronika Akopian, Michael J Ziller, Julie Donaghey, Ido Amit, Andreas Gnirke, and Alexander Meissner. Transcription factor binding dynamics during human ES cell differentiation. *Nature*, 518(7539):344–9, 2015. ISSN 1476-4687. doi: 10.1038/nature14233.

[51] Arata Takeuchi, Yasushi Itoh, Akiko Takumi, Chitose Ishihara, Noriko Arase, Tadashi Yokosuka, Haruhiko Koseki, Sho Yamasaki, Yoshimi Takai, Jun Miyoshi, Kazumasa Ogasawara, and Takashi Saito. CRTAM confers late-stage activation of CD8+ T cells to regulate retention within lymph node. *Journal of immunology (Baltimore, Md. : 1950)*, 183(7):4220–4228, 2009. ISSN 1550-6606. doi: 10.4049/jimmunol.0901248.

[52] Jeremy A. Sullivan, Eui Ho Kim, Erin H. Plisch, Stanford L. Peng, and M. Suresh. FOXO3 regulates CD8 T cell memory by T cell-intrinsic mechanisms. *PLoS Pathogens*, 8(2), 2012. ISSN 15537366. doi: 10.1371/journal.ppat.1002533.

[53] Anmol Chandele, Nikhil S Joshi, Jinfang Zhu, William E Paul, Warren J Leonard, and Susan M Kaech. Formation of IL-7Ralphahigh and IL-7Ralphalow CD8 T cells during infection is regulated by the opposing functions of GABPalpha and Gfi-1. *Journal of immunology (Baltimore, Md : 1950)*, 180(8):5309–5319, 2008. ISSN 0022-1767. doi: 180/8/5309[pii].

[54] Soo Seok Hwang, Young Uk Kim, Sumin Lee, Sung Woong Jang, Min Kyung Kim, Byung Hee Koh, Wonyong Lee, Joomyeong Kim, Abdallah Souabni, Meinrad Busslinger, and Gap Ryol Lee. Transcription factor YY1 is essential for regulation of the Th2 cytokine locus and for Th2 cell differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(1):276–81, 2013. ISSN 1091-6490. doi: 10.1073/pnas.1214682110.

[55] Gwenael Badis, Michael F Berger, Anthony A Philippakis, Andrew R Gehrke, Savina A Jaeger, Esther T Chan, Anastasia Vedenko, Xiaoyu Chen, Hanna Kuznetsov, Chi-fong Wang, Daniel E Newburger, Quaid Morris, Timothy R

Hughes, and L Bulyk. Diversity and Complexity in DNA Recognition by Transcription Factors. 324(5935):1720–1723, 2010. doi: 10.1126/science.1162327. Diversity.

[56] Michael J Ziller, Reuven Edri, Yakey Yaffe, Julie Donaghey, Ramona Pop, William Mallard, Robbyn Issner, Casey A Gifford, Alon Goren, Jeff Xing, Hongcang Gu, Davide Cachiarelli, Alexander Tsankov, Chuck Epstein, R John, Tarjei S Mikkelsen, Oliver Kohlbacher, Andreas Gnirke, E Bradley, Yechiel Elkabetz, and Alexander Meissner. Dissecting neural differentiation regulatory networks through epigenetic footprinting. 518(7539):355–359, 2015. ISSN 1476-4687. doi: 10.1038/nature13990.Dissecting.

[57] Runqiang Chen, Simon Bélanger, Megan A. Frederick, Bin Li, Robert J. Johnston, Nengming Xiao, Yun Cai Liu, Sonia Sharma, Bjoern Peters, Anjana Rao, Shane Crotty, and Matthew E. Pipkin. In vivo RNA interference screens identify regulators of antiviral CD4+ and CD8+ T cell differentiation. *Immunity*, 41(2):325–338, 2014. ISSN 10974180. doi: 10.1016/j.immuni.2014.08.002.

[58] Jesse R Dixon, Inkyung Jung, Siddarth Selvaraj, Yin Shen, Jessica E Antosiewicz-Bourget, Ah Young Lee, Zhen Ye, Audrey Kim, Nisha Rajagopal, Wei Xie, Yarui Diao, Jing Liang, Huimin Zhao, Victor V Lobanenkov, Joseph R Ecker, James A Thomson, and Bing Ren. Chromatin architecture reorganization during stem cell differentiation. *Nature*, 518(7539):331–336, 2015. ISSN 0028-0836. doi: 10.1038/nature14222.

[59] Jesse R. Dixon, Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398): 376–380, 2012. ISSN 0028-0836. doi: 10.1038/nature11082.

[60] M J Fullwood, M H Liu, Y F Pan, J Liu, H Xu, Y B Mohamed, Y L Orlov, S Velkov, A Ho, P H Mei, E G Chew, P Y Huang, W J Welboren, Y Han, H S Ooi, P N Ariyaratne, V B Vega, Y Luo, P Y Tan, P Y Choy, K D Wansa, B Zhao, K S Lim, S C Leow, J S Yow, R Joseph, H Li, K V Desai, J S Thomsen, Y K Lee, R K Karuturi, T Herve, G Bourque, H G Stunnenberg, X Ruan, V Cacheux-Rataboul, W K Sung, E T Liu, C L Wei, E Cheung, and Y Ruan. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature*, 462(7269):58–64, 2009. ISSN 1476-4687. doi: 10.1038/nature08497.

[61] Heng Li and Richard Durbin. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14):1754–1760, 2009. ISSN 13674803. doi: 10.1093/bioinformatics/btp324.

[62] Yong Zhang, Tao Liu, Clifford A Meyer, Jérôme Eeckhoute, David S Johnson, Bradley E Bernstein, Chad Nussbaum, Richard M Myers, Myles Brown, Wei Li, and X Shirley Liu. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*, 9(9):R137, 2008. ISSN 1465-6906. doi: 10.1186/gb-2008-9-9-r137.

[63] Anthony Mathelier, Oriol Fornes, David J Arenillas, Chih Yu Chen, Grégoire Denay, Jessica Lee, Wenqiang Shi, Casper Shyr, Ge Tan, Rebecca Worsley-Hunt, Allen W Zhang, François Parcy, Boris Lenhard, Albin Sandelin, and Wyeth W Wasserman. JASPAR 2016: A major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*, 44(D1):D110—-D115, 2016. ISSN 13624962. doi: 10.1093/nar/gkv1176.

[64] Daniel E. Newburger and Martha L. Bulyk. UniPROBE: An online database of protein binding microarray data on protein-DNA interactions. *Nucleic Acids Research*, 37(SUPPL. 1):77–82, 2009. ISSN 03051048. doi: 10.1093/nar/gkn660.

[65] Arttu Jolma, Jian Yan, Thomas Whitington, Jarkko Toivonen, Kazuhiro R. Nitta, Pasi Rastas, Ekaterina Morgunova, Martin Enge, Mikko Taipale, Gonghong Wei, Kimmo Palin, Juan M. Vaquerizas, Renaud Vincentelli, Nicholas M. Luscombe, Timothy R. Hughes, Patrick Lemaire, Esko Ukkonen, Teemu Kivioja, and Jussi Taipale. DNA-binding specificities of human transcription factors. *Cell*, 152(1-2):327–339, 2013. ISSN 00928674. doi: 10.1016/j.cell.2012.12.009.

[66] Charles E. Grant, Timothy L. Bailey, and William Stafford Noble. FIMO: Scanning for occurrences of a given motif. *Bioinformatics*, 27(7):1017–1018, 2011. ISSN 13674803. doi: 10.1093/bioinformatics/btr064.

[67] Glen Jeh and Jennifer Widom. Scaling personalized web search. In *Proceedings of the twelfth international conference on World Wide Web - WWW '03*, volume ofthe12th, page 271, 2003. ISBN 1581136803. doi: 10.1145/775152.775191.

Chapter 3

Runx3 programs CD8$^+$ T cell residency in non-lymphoid tissues and tumors

3.1 Introduction

During an immune response to pathogen infection, a naive CD8$^+$ T lymphocyte can give rise to terminal effector cells that eradicate invaded pathogens and functionally distinct subsets of memory cells that provide long-term protection. Memory T cells can be segregated into distinct subsets based on localization, function and phenotype: central memory cells (Tcm) found in the blood and lymphoid tissues or effector memory cells (Tem). The appreciation of a third subset of memory CD8$^+$ T cells that strictly resides within tissues and provides essential sentinel protection at body surfaces, referred to as tissue-resident memory cells (Trm), has

driven the revision of our understanding of memory cell differentiation[1]. Trm cells do not recirculate after seeding within non-lymphoid tissues, and during effector phase, Trm cells differentiate from KLRG1^{lo} populations[2]. Recent studies have revealed a list of TF regulators that can promote or repress Trm cell differentiation[3]. For example, Blimp1 and its homolog Hobit can synergistically regulate TRM differentiation by instructing a universal transcriptional program of tissue residency and tissue egress[4]. In contrast, T-box family members T-bet and Eomes have been shown to repress Trm cell differentiation and suppression of T-bet is essential to maintain optimal TGF-β signaling that favors Trm differentiation[5]. Although accumulating studies highlighted the critical role of Trm in immunological memory, little is known about the transcriptional pathways and the molecular drivers that regulating Trm formation, function and maintenance.

In infectious settings, Trm display the following features: (1) successful infiltration and long-term residency in non-lymphoid tissues; (2) high cytotoxicity to kill pathogen-infected cells; (3) production of cytokines and chemokines to recruit other innate and adaptive immune cells. These features allow Trm to become a critical component in non-infectious settings including tumor malignancy and autoimmunity. Indeed, the tumor microenvironment is very similar to non-lymphoid tissues in terms of low oxygen tension, specific cytokine environment and nutrient availability. Recent work have demonstrated that retention molecule CD103, a marker of "bona fide" Trm, is also expressed on tumor-infiltrating T cells (TILs) in particularly the cancers of epithelial origin[6]. More importantly, CD8^{+} TILs that express CD103 also display a Trm-like transcriptional program and is predictive of a better survival outcome in human lung cancer patients[7]. Thus, a deep

understanding of how TILs employ the Trm-like gene expression program and how to specifically target TFs that control Trm program in TILs would greatly enhance the efficacy of immunotherapy especially the adoptive T cell therapy.

3.2 Results

3.2.1 Trm precursors exhibit a unique transcriptional program and chromatin state.

A number of studies have highlighted the distinct gene-expression profiles exhibited by circulating memory $CD8^+$ T cells and Trm[8]; however, the early transcriptional identity of differentiating Trm and the signals controlling their fate are not well understood. We utilized an established infection model in which transgenic $CD8^+$ T cells responsive to lymphocytic choriomeningitis virus (LCMV) $GP_{33\text{-}41}$ presented by H-2Db (P14) were transferred into recipient mice one day prior to infection with LCMV. In this acute infection model, P14 cells located in non-lymphoid tissues on day 7 of infection began to upregulate retention molecules characteristic of Trm[9], including canonical Trm markers CD103 and CD69. Gene-expression analysis revealed that 90-96% of the genes upregulated in mature Trm cells (day 35 of infection) were elevated in Trm precursors relative to splenic effector cells (Figure 3.1a). Furthermore, analysis of genes differentially expressed between splenic and non-lymphoid populations on day 7 of infection revealed two distinct gene programs segregating circulating (PBL, Spleen, Tcm and Tem) from non-lymphoid (kidney and IEL) $CD8^+$ T cells responding to infection, independent of infection time point (Figure 3.1b). Previous experiments suggest lymph node

(LN) or splenic KLRG1^{lo}CD127^{hi} memory-precursor (MP) cells preferentially give rise to long-lived memory populations compared to shorter-lived KLRG1^{hi}CD127^{lo} terminal-effector (TE) cells[10, 11]. Unexpectedly, day 7 Trm precursors within the small intestine intraepithelial lymphocyte (IEL) compartment were transcriptionally distinct from splenic MP cells (Figure 3.1c). Thus, the Trm-precursor populations in non-lymphoid tissues are transcriptionally distinct from splenic and blood effector cells as well as memory-precursor cells on day 7 of infection, and the majority of the Trm transcriptional program is established at this time point, prior to CD8$^+$ T cell contraction.

As chromatin accessibility is a key determinant of cell identity and fate, we profiled kidney, IEL, and splenic effector P14 populations as well as splenic MP and TE subsets using ATAC-seq on day 7 of infection. Principal component analysis highlighted that, despite day 7 being an effector time point, the global chromatin landscape dramatically differs between effector CD8$^+$ T cells located in the spleen, including MP cells, relative to non-lymphoid tissues (Figure 3.1d). Uniquely accessible sites were identified in IEL P14 cells for genes characteristic of mature Trm such as Cd69 and Nr4a1, whereas genes that promote T cell recirculation such as Klf2 and S1pr1 exhibited loss of accessible sites, correlating with gene expression (Figure 3.1e). Thus, the unique chromatin state of differentiating Trm is consistent with the striking transcriptional differences observed (Figure 3.1a-c) and foreshadows the distinct fates of antigen-specific cells in the spleen relative to non-lymphoid tissues. Taken together, these transcription and chromatin profiling studies establish precursors of Trm cells in non-lymphoid sites as a unique and distinct CD8$^+$ T cell subset relative to effectors in the lymphoid

compartment, including KLRG1^{lo}CD127^{hi} memory precursors.

3.2.2 Computational and functional *in vivo* RNAi screen identify transcriptional regulators of Trm differentiation.

Specification of CD8$^+$ T cell fates during infection is dependent on the integrated activity of transcription factors (TFs)[12]. Although a number of TFs that promote effector (T-bet, Blimp1, Id2, Zeb2) or memory differentiation (Bcl6, Id3, Foxo1, Tcf1) have been identified[12], there are relatively few TFs currently known to control Trm formation[13]. To facilitate identification of regulators of Trm differentiation, we utilized a novel combined screening approach, consisting of: a computational strategy integrating ATAC-seq data, transcriptional profiling and personalized PageRank analysis to predict regulatory TFs, and a functional *in vivo* RNAi screen targeting putative Trm regulators identified through our computational approach (Figure 3.2a). We have recently used the integrated analysis of accessible TF binding motifs (assessed through ATAC-seq) and the regulation of their target gene expression to yield insight into TFs with regulatory functions in differentiation of splenic MP, TE, and long-lived memory CD8$^+$ T cells[14]. Leveraging this approach and the personalized PageRank analysis, we predicted a number of TFs with established regulatory roles in controlling Trm versus circulating CD8$^+$ T cell differentiation (Blimp1, Nr4a1, Bach2, Eomes, T-bet[5, 4, 15, 16] and many with no previously described role in Trm (Figure 3.2b). We evaluated both barrier (IEL) and non-barrier (kidney) Trm sites to reveal TFs important to

Trm differentiation independent of the tissue. To further establish functional roles for predicted regulators of Trm formation identified through PageRank analysis, we utilized an RNAi screening strategy allowing testing of hundreds of individual shRNA constructs in parallel for activity in promoting or repressing Trm differentiation *in vivo*[17] (Figure 3.2c). Several TFs with established roles in regulating Trm were identified (e.g. Blimp1, Klf2 and T-bet[4, 5]) as well as TFs with previously unknown functions in controlling $CD8^+$ Trm formation such as Nr4a3 and Runx3 (Figure 3.2b). Additionally, a key strength of this computational screen is that influential roles of differentially expressed TFs as well as TFs with homogenous expression can be anticipated Figure 3.2d). Thus, this dual computational-functional screening strategy identified a number of TFs that had not been previously recognized as regulators of Trm fate.

3.2.3 Runx3 is essential for Trm differentiation.

Runx3 has well-established functions in controlling $CD8^+$ T cell lineage specification during thymocyte development[18] and cytotoxic effector function of mature $CD8^+$ T cells[19]. Runx3 is thought to be important in $CD4^+$ T cell localization within the intestinal epithelium[20], yet how Runx3 may impact CD8+ Trm differentiation and homeostasis remains unexplored. Our computational analysis ranked Runx3 high among TFs predicted to regulate key Trm genes (Figure 3.2b) despite relatively uniform Runx3 expression in circulating and resident CD8+ T cell subsets (Figure 3.3a), and the RNAi screen indicated that Runx3 shRNAs impaired Trm differentiation (Figure 3.2c). We validated Runx3 as a regulator of Trm differentiation in a 1:1 mixed transfer of P14 cells transduced with control (Cd19

shRNA) or Runx3 shRNA-encoding retroviruses into C57BL/6 mice that were subsequently infected with LCMV, allowing assessment of a cell-intrinsic role for Runx3 as both populations differentiate in the same environment in vivo. Runx3 shRNA-mediated knockdown suppressed mRNA expression of Runx3 (Figure 3.3b) and reduced the proportion of transduced cells in the IEL relative to the spleen, consistent with the screening results (Figure 3.3c-d). In addition, Runx3 shRNA-mediated knockdown impaired expression of surface molecules CD69 and CD103, which are characteristically expressed by mature Trm[9] (Figure 3.3e). These data demonstrate that Runx3 is critical for Trm differentiation.

3.2.4 Runx3 regulates the core Trm transcriptional program to promote CD8$^+$ T cell tissue-residency.

We next assessed how ectopic expression of Runx3 impacted Trm differentiation. Congenically distinct P14 cells transduced with a control retrovirus (GFP-RV) or Runx3 cDNA-encoding retrovirus (Runx3-RV) were mixed 1:1 and transferred into recipient mice that were subsequently infected with LCMV. Overexpression of Runx3 accelerated IEL P14 CD69$^+$CD103$^+$ Trm differentiation on day 8 of infection, but did not impact migration to the small intestine, as the ratio of transferred cells was similar between the spleen, mLN and IEL populations at this time point (Figure 3.4a). Evidence of enhanced Trm differentiation was further confirmed by the greater abundance of IEL Trm on day 13 of infection as well as increased CD103 expression, consistent with a reported role for Runx3 in regulating CD103 expression[21, 20] (Figure 3.4b).

Given that manipulation of Runx3 impacted Trm formation in a range of

tissue microenvironments, we constructed a core Trm transcriptional signature by computational integration of $CD8^+$ Trm gene expression data sets from the IEL, kidney, lung[8], skin[8] and brain[22], to evaluate the hypothesis that Runx3 is a universal regulator of $CD8^+$ Trm differentiation. This analysis identified a core Trm gene-expression signature consisting of 157 genes expressed at higher levels in all five Trm populations relative to their circulating splenic counterparts (Fig. 3e). Further, 114 genes were downregulated in all Trm populations relative to splenic memory cells, yielding a core gene-expression signature of circulating T cell populations. Next, gene expression by $Runx3^{fl/fl}$, WT, or Runx3-RV transduced $CD8^+$ T cells was evaluated by RNA-seq analysis, and notably we found the majority of the core Trm signature genes were upregulated in Runx3-overexpressing cells and downregulated in Runx3-deficient cells. Conversely, the core signature of circulating $CD8^+$ T cell memory cells appeared to be predominantly upregulated in Runx3-deficient cells and downregulated in Runx3-overexpressing cells (Figure 3.4c). These findings were statistically validated through gene set enrichment analyses (Figure 3.4c). Therefore, Runx3 promoted expression of Trm signature genes and repressed genes characteristic of circulating cells.

3.2.5 $CD8^+$ TIL share transcriptional similarity with Trm and require Runx3 for tumor residency.

It has been noted that $CD8^+$ tumor infiltrating lymphocytes (TIL) can exhibit characteristics of Trm, and a positive disease outcome has been correlated with TIL that display characteristic features of Trm, such as CD103 expression[23, 24, 7]. However, the relationship between Trm and TIL is unclear and understud-

ied. As Runx3 is central to Trm differentiation and homeostasis, we assessed the transcriptional similarities of Trm and TIL and evaluated a role for Runx3 in controlling TIL accumulation within tumors. TILs isolated from mouse melanoma[24] or mammary tumors[24] shared 70% of the core Trm gene-expression program relative to splenic $CD8^+$ T cells (Figure 3.5a), and this relationship was further highlighted through principal component analysis (Figure 3.5b). Given that the core Trm and re-circulating gene-expression programs are dynamically regulated by Runx3 (Fig. 3i-j), we next tested if Runx3 was required by TIL residency. Utilizing an adoptive therapy model, Runx3-knockdown or Runx3-overexpressing P14 cells were mixed with control (Con shRNA or GFP-RV) P14 cells at a 1:1 ratio and transferred into mice with palpable melanoma tumors expressing GP33-41. Strikingly, Runx3-deficiency impaired TIL accumulation while overexpression enhanced TIL abundance ((Figure 3.4c,d). In line with this observation, RNA-seq analysis of TILs from GFP-RV and Runx3-RV group and splenic counterparts showed that overexpression of Runx3 promted the expression of the core tissue-residency gene signature (Figure 3.5e). In clinical settings, TIL frequency strongly correlates with positive prognoses in cancer patients[25]. In connection, Runx3-deficient P14 cells were impaired in their ability to control tumor growth, resulting in greater mortality (Figure 3.5f). Conversely, Runx3-overexpressing P14 cells delayed tumor growth and prolonged survival (Figure 3.5g). Furthermore, single-cell RNAseq analysis of mouse and human melanoma samples indicated that $CD44^+CD8^+$ T cells expressing Runx3 exhibited enrichment of the Trm signature relative to $CD44^+CD8^+$ TIL with low Runx3 expression levels (Figure 3.5h). These data indicate that both human and murine TIL share transcriptional similarities with Trm

and CD8$^+$ T cell residency within the tumor microenvironment is likely dependent on Runx3.

3.3 Conclusion

Here we show that Runx3 is a central regulator of Trm differentiation and describe an unappreciated role for Runx3 in regulating TIL accumulation. These findings provide a new strategy that may be used to augment adoptive cell therapies, including CAR therapy where limited efficacy in solid tumor settings has been reported[26]. Manipulation of TFs promoting Trm differentiation may yield more effective TIL through supplementing transferred CD8$^+$ T cells with a gene expression program that better supports features important to both Trm and TIL such as *in situ* survival and repression of tissue egress programs, ultimately fostering accumulation of protective T cells in tissues.

3.4 Acknowledgement

Chapter 3, in part, is a subset of the material as it appears in Nature 2017. Milner JJ, Toma C*, Yu B*, Zhang K, Omilusik KD, Phan A, Wang DP, Getzler A, Crotty S, Wang W, Pipkin ME, Goldrath AW. Runx3 programs CD8+ T cell residency in non-lymphoid tissues and tumors. *Nature*. 2017 552(7684)

The dissertation author was a primary investigator and the co-second author of this paper.

3.5 Methods

3.5.1 Mice

Mice were maintained in specific-pathogen-free conditions in accordance with the Institutional Animal Care and Use Committees (IACUC) of the University of California, San Diego (UCSD) and The Scripps Research Institute, Jupiter, FL (TSRI-FL). All mice were of a C57BL6/J background and bred at UCSD and TSRI-FL or purchased from the Jackson Laboratory, including: wild-type or P14 mice with distinct expression of the congenic molecules CD45.1, CD45.1.2, CD45.2, Thy1.1, Thy1.1.2, and Thy1.2 as well as control Thy1.2^{+} Runx3$^{+/+}$ Ert2-Cre YFP+ P14 mice and Runx3 inducible deletion Thy1.1^{+} Runx3$^{fl/fl}$ Ert2-Cre YFP+ P14 mice.

3.5.2 Naive T cell transfer, infection and treatment

Naive P14 CD8^{+} T cells were transferred intravenously (i.v.) into congenically distinct sex matched recipient mice, or female P14 cells were transferred into male mice. For all microarray, RNA-seq or ATAC-seq experiments, 1×10^5 P14 cells were transferred. For co-transfer experiments, naive Thy1.2^{+} Runx3$^{+/+}$ Ert2-Cre YFP+ P14 cells and naive Thy1.1^{+} Runx3$^{fl/fl}$ Ert2-Cre YFP+ P14 cells were mixed 1:1 and a total of 3×10^4 P14 cells were transferred into Thy1.2^{+} recipient mice. Recipient mice were subsequently infected i.p. with 2×10^5 PFU of the Armstrong strain of lymphocytic choriomeningitis virus (LCMV) one day after cell transfer. To distinguish vascular associated CD8^{+} T cells in non-lymphoid tissues, 3μg of CD8a (53-6.7) conjugated to APC eFlour780 was injected i.v. into mice

four minutes prior to sacrifice and organ excision. CD8a^{neg} cells were considered to be localized within non-lymphoid tissues.

3.5.3 Preparation of cell suspensions.

Isolation of CD8$^+$ T cells was performed similarly as described. For isolation of CD8$^+$ T cells from the small intestine intraepithelial lymphocyte (IEL) compartment, Peyers patches were removed and the intestine was cut longitudinally and subsequently cut laterally into 0.5-1cm^2 pieces that were then incubated with 0.154mg/mL dithioerythritol (DTE) in 10% HBSS/HEPES bicarbonate for 30min at 37C while stirring. Kidneys, salivary glands and lungs were cut into pieces and digested for 30min with 100 U/mL type I collagenase (Worthington) in RPMI 1640, 5% FBS, 2mM MgCl2, 2mM CaCl2 at 37C while shaking. Skin was processed similarly as described in which a 2cm^2 area of the right flank was excised, pre-digested for 30min at 37C and then enzymatically digested with 0.7 mg/mL collagenase D. After enzymatic incubations (skin, lungs, kidneys, salivary glands), tissues were further dissociated over a 70m nylon cell strainer (Falcon). For isolation of lymphocytes, single-cell suspensions were then separated using a 44/67% Percoll density gradient. Spleens and lymph nodes were processed with the frosted ends of microscope slides. Red blood cells were lysed with ACK buffer (140 mM NH4Cl and 17 mM Tris-base, pH 7.4).

3.5.4 Antibodies and flow cytometry.

The following antibodies were obtained from eBioscience: CD8a (53-6.7), CD8b (eBio H35-17.2), CD62L (MEL-14), CD127 (A7R34), KLRG1 (2F1), CD103

(2E7), CD69 (H1.2F3), CD45.1 (A20-1.7), CD45.2 (104), Thy1.1 (OX-7, HIS51), Thy1.2 (53-2.1), CCR9 (Ebio CW-1.2), CXCR3 (CXCR3-173), CD49d (R1-2), and T-bet (4B10). For intracellular staining of T-bet while preserving ametrine or YFP reporter expression in transduced or Cre-YFP+ populations, cells were fixed and permeabilized through a 15min incubation with BD cytofix/cytoperm (BD Biosciences). Intracellular staining was subsequently performed using the Perm Buffer of the Foxp3-transcription factor staining buffer kit (eBioscience). For flow cytometry, all events were acquired on a BD LSRFortessa X-20 or a BD LSRFortessa. Cell sorting was performed on BD FACSAria or BD FACSAria Fusion instruments.

3.5.5 RNAi screening approach.

We have described this screening approach in detail previously[17]. The targeted shRNA library was generated based on key genes identified from the computational screening approach as well as genes with known roles in regulating Trm from literature. The library was produced by cloning shERWOOD-designed shRNAmir sequences, after PCR of synthetic 97mer oligos, into our pLMPd-Amt vector. Purified DNA from sequence-verified clones was used to package retroviral particles in PLAT-E cells. For transfections, PLAT-E cells were seeded in the middle 60 wells of a 96-well flat bottom plate at a density of $4 - 6 \times 10^4$ cells/well) one day prior to transfection. Next, each well was individually transfected with 0.2μg of DNA from each pLMPd-Amt clone and 0.2μg of pCL-Eco using TransIT-LT1 (Mirus). Retroviral supernatant was harvested 36, 48, and 60h after transfection, and RV sup from each well was used to individually transduce

in vitro activated P14 cells in 96-well round bottom plates. For CD8+ T cell activation in vitro, naive CD8+ T cells from spleen and lymph nodes were negatively enriched and 2×10^5 P14 cells were plated in the middle 60 wells of 96-well round bottom plates pre-coated with 100μg/mL goat anti-hampster IgG (H+L, Thermoscientific) and 1g/mL anti-CD3 (145-2C11) and 1g/mL anti-CD28 (37.51) (both from eBioscience). Culture media was removed 18h after activation, and replaced with retroviral supernatant supplemented with 50μM BME and 8g/mL polybrene (Millipore) followed by spinfection (60min. centrifugation at 2000 rpm, 37C). Two hours after the spinfection, the P14 cells were washed 3 times with cold PBS and 90% of each well of cells (individually transduced with distinct retroviral constructs) was harvested, pooled and 5×10^5 pooled P14 cells were transferred into recipient mice which were then infected 1h later with 1.5×10^5 PFU of LCMV clone 13 ip 1h later, resulting in an acute infection. The remaining cells in vitro were cultured for an additional 24h and either pooled for input sequencing (6×10^5 P14 cells) or were used to test transduction efficiency of each construct using flow cytometry to detect the percentage of ametrine+ cells in each well. Twelve days after infection, spleens and small intestines were harvested from 15-18 mice and splenocytes and IEL P14 cells were processed as described above. Prior to sorting, all IEL or splenic samples were pooled. CD62L+ P14 cells (Tcm) from the spleen as well as P14 cells from the IEL were sorted ($3.5 - 6 \times 10^5$ cells total). Genomic DNA was then harvested from sorted cells using the FlexiGene kit (Qiagen). The integrated proviral passenger strand shRNAmir sequences in each cell subset were amplified from 20-100ng total genomic DNA per reaction, with 23-28 cycles of PCR using Ion Proton-compatible barcoded primers that anneal to the common

5 mir30 and shRNA loop sequences. 2-3 replicate reactions were performed for each genomic DNA sample and the replicates were pooled after amplification. The pooled reactions were purified using AMPure XP beads, the amplicons in each sample were quantified using a Bioanalyzer, and then pooled in a 1:1 molar ratio for sequencing.In each replicate of the screen, a minimum of 2.5 million reads per sample were generated and retained, after filtering low-quality reads. Reads assigned to each barcode were aligned to a reference database of all shRNAmirs in the library using BLAST and a custom script to count the top alignment of each read and summarize the number of reads aligned to each shRNAmir. For analysis of shRNA representation in Tcm relative to IEL Trm, the total number of reads in each of the samples was normalized, and the number of reads for each shRNA was scaled proportionally. Subsequently, the normalized number of reads in the IEL Trm cells for a given shRNA was divided by the normalized number of reads for the same shRNA in the Tcm sample and then log2 transformed. The mean and standard deviation of the ratios of each of the 25 negative control shRNA constructs (targeting Cd19, Cd4, Cd14, Ms4a1, Cd22, Hes1, Klf12, Mafb, Plagl1, Pou2af1 and Smarca1) were used to calculate the Z-score for each shRNA construct. The screen was repeated three times and the Z-score of each construct from each individual screen was averaged and plotted (Fig. 2c). Some shRNAs were only included in two independent screens. Eighty-four percent (21/25) of all negative control shRNA constructs had an average Z-score between 0.9 and -0.9.

3.5.6 T cell transduction, cell transfer, and infection for individual analysis of retroviral constructs.

Activation, transfections and transductions were carried out as described for the RNAi screening approach except in some experiments 2×10^6 P14 cells were activated per well in 6-well plates. Congenically distinct P14 cells transduced with the Runx3.2 shRNA or Cd19.1 shRNA (control) retroviruses were mixed 1:1 within 24h of transduction and a total of $1 - 5 \times 10^5$ P14 cells were transferred i.v. into recipient mice. One hour after adoptive transfer, recipient mice were infected i.p. or intratracheally (i.t.) with 2×10^5 PFU LCMV armstrong or intradermally (i.d.) with 2×10^4 PFU clone 13. In similar experiments, P14 cells were transduced with MigR1-based retroviruses that were empty (GFP-RV) or that contained Runx3 cDNA (Runx3-RV), mixed 1:1 and transferred to recipient mice for subsequent infections.

3.5.7 Adoptive therapy tumor model.

For adoptive therapy experiments, 5×10^5 B16-GP33 cells were transplanted subcutaneously into the right flank of wild-type mice. After tumors became palpable, 7-8 days post-transplant, in vitro expanded P14 cells were transferred i.v. For comparison of TIL accumulation in a mixed transfer setting, naive P14 cells were activated, transduced and expanded with 100U/mL of IL-2 for 2-3 days; cells transduced with control constructs (Cd19.1 shRNA or GFP-RV) or experimental constructs (Runx3.2 shRNA or Runx3-RV) were mixed 1:1 and 0.5-1x106 P14 cells were transferred i.v. For efficacy studies, transduced cells were expanded for 5-6

days; transduced cells were then sorted (or not sorted with a Runx3-RV and GFP-RV transduction efficiency >83%), and $1-2.5\times10^6$ cells were transferred i.v. into mice with established B16-GP33 tumors. Tumors were monitored daily and mice with ulcerated tumors or tumors exceeding 1500 mm^3 were euthanized.

3.5.8 qPCR, Microarray, RNA-seq and ATAC-seq analysis.

For validation of the Runx3-RV overexpression construct and Runx3.2 shRNA construct, enriched $CD8^+$ T cells were activated, transduced and expanded for 4-6 days in 100U/mL IL-2. Cells were sorted on ametrine (Runx3 shRNA or Con shRNA) or GFP (Runx3-RV or GFP-RV) directly into TRIzol (Life Technologies) and RNA was extracted per manufacturers specifications. Next, cDNA was synthesized using Superscript II (Life Technologies) and qPCR was performed using the Stratagene Brilliant II Syber Green master mix (Agilent Technologies). Runx3 expression levels were normalized to the housekeeping gene Hprt. We have previously validated the Tbx21.3 shRNA. The following primers were used for qPCR: Runx3 forward, 5-CAGGTTCAACGACCTTCGATT-3, and Runx3 reverse, 5-GTGGTAGGTAGCCACTTGGG-3; Hprt forward, 5-GGCCAGACTTTGTTGGATTT-3, and Hprt reverse: 5-CAACTTGCGCTCATCTTAGG-3. On day 7 of infection, tissues from 2-3 mice were pooled and $2-3\times10^4$ P14 cells from the IEL, kidney, spleen, or blood were sorted into TRIzol. On day 35 of infection, tissues from 5-10 mice were pooled and $1-2\times10^4$ CD62L+ Tcm, CD62L- Tem, kidney Trm and IEL Trm P14 cells were sorted into TRIzol. As described previously, RNA was amplified and labeled with biotin and hybridized to Affymetrix Mouse

Gene ST 1.0 micrroarrays (Affymetrix). Analyses were performed using GenePattern Multiplot Studio. Differentially expressed genes in IEL Trm compared to Tcm and Tem as well as kidney Trm compared to Tcm and Tem were identified with a fold change (FC) >1.5 and an expression value (EV) >120 (Fig. 1a). Genes with >1.5 FC and >120 EV between day 7 spleen, day 7 IEL, day 7 kidney samples were identified (2206 probes) and evaluated in day 7 and day 35 subsets, which were ordered with Pearson correlation using the HierarchicalClustering module of GenePattern (Fig. 1b); data was row centered, row normalized and visualized with the HierarchicalClusteringViewer module within GenePattern.

The core Trm and circulating signatures were generated by integrating differential expression (>1.5 FC) data comparing Trm from the following tissues to circulating splenic memory cells: D35 IEL (LCMV), D35 kidney (LCMV), D30 skin $CD103^{+}CD8^{+}$ (herpes simplex virus), D30 lung $CD103^{+}CD8^{+}$ (influenza virus), and D20 $CD103^{+}$ brain (vesicular stomatitis virus); overlapping genes upregulated in all Trm populations comprised the Trm core signature (157 genes) and genes downregulated in all populations comprised the circulating signature (114 genes). The TIL microarray datasets were generated previously[24]. For RNA-seq analysis of D7 IEL, D7 MP and D7 TE as well as naive P14 cells, spleens or IEL samples from 2-3 mice were pooled and 5×10^3 cells were sorted at day 7 of LCMV Arm infection. For library preparation, isolation of polyA+ RNA was performed as detailed online (www.immgen.org/Protocols/11cells.pdf). For RNA-seq analyses of Runx3-manipulated cells, CD8+ T cells from naive $Runx3^{+/+}$ YFP+ (WT) and $Runx3^{fl/fl}$ YFP+ ($Runx3^{fl/fl}$) mice were enriched by negative isolation and transduced (as detailed above) with a Cre cDNA expressing retrovirus (Cre-RV). Runx3-

overexpressing cells were generated similarly by transducing Runx3$^{+/+}$ YFP+ CD8 T cells with a Runx3-cDNA expressing retrovirus (Runx3-RV). Forty-eight hours after TCR activation, the CD8^{+} T cells were resuspended and re-cultured in fresh media supplemented with 100U/mL rhIL-2; twenty-four hours later, YFP+ (WT or Runx3$^{fl/fl}$) or GFP+ (Runx3-RV) were FACS-purified and then recultured in 100U/mL IL-2. The cells were expanded until day 6 by reculturing at 5×10^5 cells/mL every 24h in fresh 100U/mL IL-2 media. On day 6 post-activation, cells were harvested and total RNA was extracted in TRIzol. Purified RNA was depleted of ribosomal RNA and strand-specific paired-end libraries were prepared and sequenced using an Illumina Nextseq 500. Samples were generated from two biological replicates, and approximately 20 million paired-reads were generated per sample. Reads were mapped using Tophat and aligned reads in transcripts were counted with HTseq. Gene-set-enrichment analysis was performed by using the GSEA module in GenePattern, and the normalized enrichment scores and false-discovery rate q values were determined by using the permutation test.

ATAC-seq anlalysis was performed as described in detail previously[27]. Sorted cells (2.5×10^4) were resuspended in 25μL of lysis buffer and spun down 600g for 30min at 4C. The nuclear pellet was rususpended in 25μL of Tn5 transposase reaction mixture (Nextera DNA Sample Prep Kit, Illumina) and incubated for 30min at 37C. Transposase-associated DNA was subsequently purified (Zymo DNA clean-up kit). For library amplification, DNA was amplified for five cycles using indexing primer from Nextera kit and NEBNext High-Fidelity 2X PCR master mix. Then the amplified DNA was size selected to fragments less than 800 bp using SPRI beads. Quantification of the ATAC-seq library was based on KAPA

library quantification kit (KAPAbiosystems). The size of the pooled library was examined by TapeStation. The library was sequenced using Hiseq 2500 for single-end 50-bp sequencing to yield at least 10 million reads. We used bowtie to map raw reads to the Mus musculus genome (mm10) with following parameters: –best -m 1. We called peaks for each individual replicate as well as the pooled data from the two replicates using MACS2 with a relaxed threshold (P-value 0.01). Each ATAC-seq experiment was performed twice and we used the Irreproducibility Discovery Rate (IDR) framework to identify reproducible peaks. For the Runx3 ChIP-seq analysis, the fastq files were downloaded from GSE50131 and mapped to mm10 mouse genome using bowtie. Runx3 peaks were called by HOMER using findPeaks command with parameters: -style factor and visualized using the UCSC genome browser.

For the single cell RNA-seq analysis of human melanoma TIL, the preprocessed single cell TIL gene expression data was downloaded from GEO database GSE72056. Activated CD8 TILs (CD8a expression >5 and CD44 expression>2) in melanoma #75 and melanoma #79 were used and classified into Runx3hi TILs which express high level of Runx3 (Runx3 expression>3) and Runx3lo TILs without expression of Runx3 (Runx3 expression$\bar{0}$). GSEA was performed to look at the enrichment of core Trm gene signature in Runx3hi TILs relative to Runx3lo TILs.

3.5.9 Computational screen.

TF regulatory networks and PageRank analysis was performed similarly as described[14] except that gene expression and ATAC-seq data from D7 IEL, D7

kidney and D7 spleen samples were used. To identify putative binding sites of TFs, we first collected 761 unique motifs from 3 TF motif database, JASPAR, UniPROBE and Jolma et al. study[28, 29, 30]. We then searched for motifs binding sites in 150bp regions centered around ATAC-seq peak summits, using the algorithm described in Grant's paper with of P-value cutoff of 1e-5[31]. Then we connected a TF to a gene if the TF had any predicted binding motif in the ATAC-seq peak of the nearest gene. We assembled all the interactions between TFs and genes into a regulatory network. To identify important TF regulators for Trm differentiation, we performed personalized PageRank analysis in the TF regulatory network constructed above using the pipeline described previously[14]. The importance of a TF is based on the quantity and quality of its regulated gene targets. A TF would receive a higher PageRank score if it regulates more important genes where the importance is evaluated by the differential expression from the microarray or RNA-seq analysis.

3.6 Figures

Figure 3.1: **Trm precursors exhibit a distinct transcriptional program and chromatin state from splenic memory precursors. a,** Comparison of gene-expression of Trm isolated from the small intestine IEL (left) and kidney parenchyma (right) relative to splenic memory subsets, Tcm and Tem, on day 35 of LCMV infection; red, genes elevated in Trm relative to Tcm and Tem; blue, genes elevated in Tcm and Tem relative to Trm (top). Comparison of differentially expressed genes in mature Trm (from top panel) in cells from the spleen, IEL or kidney on day 7 of infection (bottom). **b,** Differentially expressed genes between splenic, IEL and kidney populations on day 7 of infection were compared among effector and memory CD8+ T cell subsets. Populations are ordered by hierarchical clustering with Pearson correlation. **c,** Principal component analysis of differentially expressed genes among TE, MP and IEL CD8+ T cell subsets on day 7 of infection and naive P14 cells. **d,** Principal-component analysis of differential global chromatin accessibility for CD8+ T cell populations on day 7 of infection identified by ATAC-seq analysis. **e,** ATAC-seq analysis of the S1pr1, Klf2, Cd69 and Nr4a1 loci for IEL and splenic CD8+ T cells on day 7 of infection (left) and corresponding gene expression (right). For graphs in a,b, each timepoint represents an individual experiment consisting of 2-3 biological replicates where n=2-10 mice were pooled for each replicate; for c,d,e, 2 biological replicates where 2-5 mice were pooled for each replicate

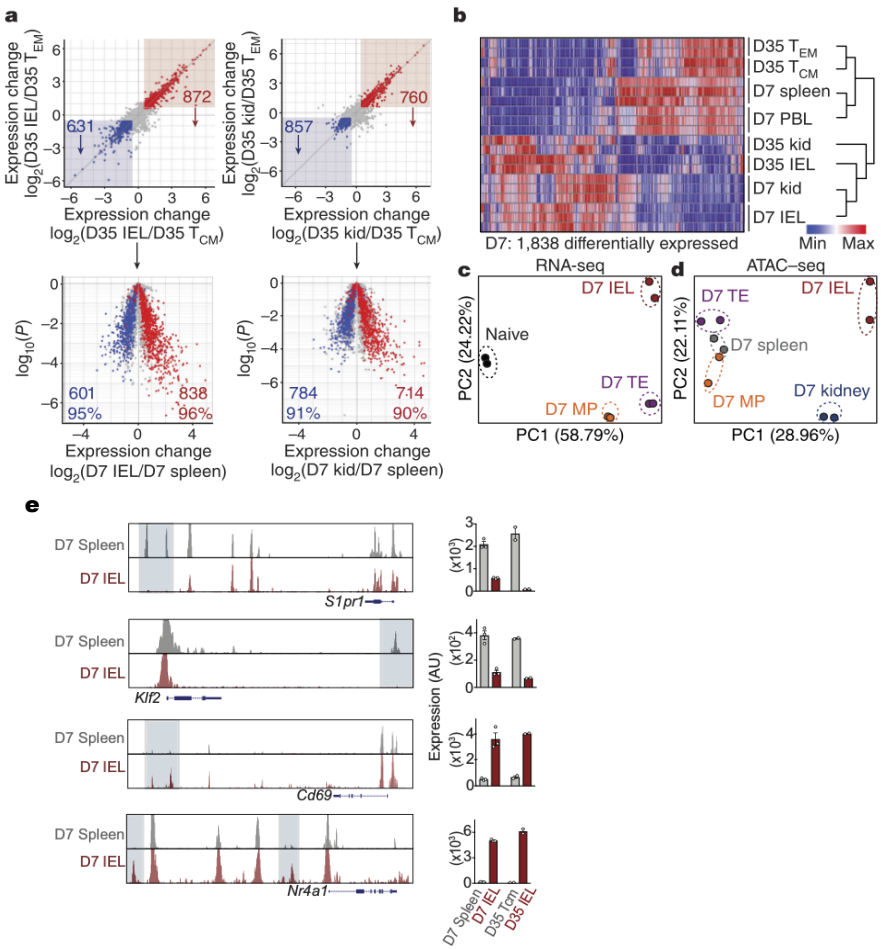

a
Expression change log2(D35 IEL/D35 TEM)
Expression change log2(D35 IEL/D35 TCM)
631
872
Expression change log2(D35 kid/D35 TEM)
Expression change log2(D35 kid/D35 TCM)
857
760
log10(P)
601
95%
838
96%
Expression change log2(D7 IEL/D7 spleen)
784
91%
714
90%
Expression change log2(D7 kid/D7 spleen)
b
D35 TEM
D35 TCM
D7 spleen
D7 PBL
D35 kid
D35 IEL
D7 kid
D7 IEL
D7: 1,838 differentially expressed
Min
Max
c
RNA-seq
D7 IEL
Naive
D7 TE
D7 MP
PC2 (24.22%)
PC1 (58.79%)
d
ATAC–seq
D7 IEL
D7 TE
D7 spleen
D7 MP
D7 kidney
PC2 (22.11%)
PC1 (28.96%)
e
D7 Spleen
D7 IEL
S1pr1
Klf2
Cd69
Nr4a1
Expression (AU)
(x10³)
(x10²)
D7 Spleen
D7 IEL
D35 Tcm
D35 IEL

Figure 3.2: **Computational and functional RNAi screens identify transcriptional regulators of Trm differentiation.** **a,** Combinatorial screening approach consisting of an integrated computational analysis and an *in vivo* RNAi functional screen. **b,** Predicted regulators of Trm differentiation identified through the computational screen; genes with known function in regulating Trm formation are in bold font. TFs with a PageRank score of at least 2 fold change across spleen, IEL and kidney CD8 T cells are shown in the heatmap. **c,** Relative enrichment of shRNAs in IEL Trm relative to splenic Tcm from the functional RNAi screen, reported as the average Z-score from three independent screens. **d,** Personalized PageRank score and gene-expression for predicted regulators of IEL Trm differentiation. For c, each of the three independent screens was performed by pooling genomic DNA from P14 cells isolated from tissues of 15-18 mice.

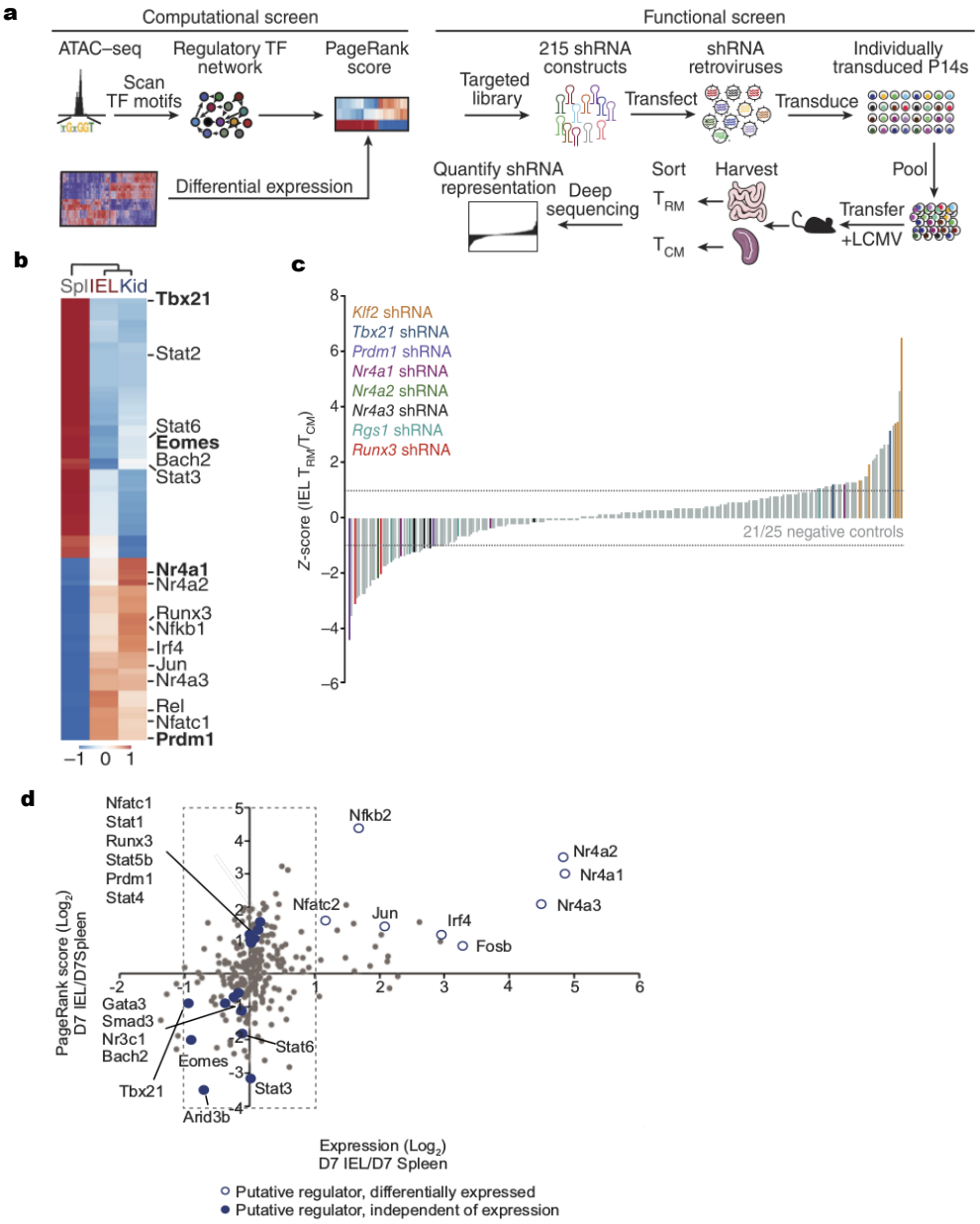

a
Computational screen
ATAC–seq
Scan TF motifs
Regulatory TF network
PageRank score
Differential expression
Functional screen
Targeted library
215 shRNA constructs
Transfect
shRNA retroviruses
Transduce
Individually transduced P14s
Pool
Transfer +LCMV
Harvest
Sort
T_{RM}
T_{CM}
Deep sequencing
Quantify shRNA representation
b
Spl IEL Kid
Tbx21
Stat2
Stat6
Eomes
Bach2
Stat3
Nr4a1
Nr4a2
Runx3
Nfkb1
Irf4
Jun
Nr4a3
Rel
Nfatc1
Prdm1
−1 0 1
c
Klf2 shRNA
Tbx21 shRNA
Prdm1 shRNA
Nr4a1 shRNA
Nr4a2 shRNA
Nr4a3 shRNA
Rgs1 shRNA
Runx3 shRNA
Z-score (IEL T_{RM}/T_{CM})
21/25 negative controls
d
PageRank score (Log_2) D7 IEL/D7Spleen
Expression (Log_2) D7 IEL/D7 Spleen
Nfatc1
Stat1
Runx3
Stat5b
Prdm1
Stat4
Nfkb2
Nr4a2
Nr4a1
Nr4a3
Nfatc2
Jun
Irf4
Fosb
Gata3
Smad3
Nr3c1
Bach2
Eomes
Tbx21
Stat6
Stat3
Arid3b
Putative regulator, differentially expressed
Putative regulator, independent of expression

Figure 3.3: **Runx3 is essential for Trm differentiation. a,** Runx3 gene expression from microarray analyses. Graphs indicate mean of 2-3 replicates (n=2-10 mice per replicate) per cell type. Symbols represent an individual replicate. **b,** Runx3 shRNA knockdown efficiency using RT-qPCR data to show Runx3 mRNA expression of *in vitro* cultured cells transduced with Con shRNA or Runx3 shRNA retroviruses. **c,** Schematic of experimental design. Congenically distinct P14 cells were transduced with Runx3 shRNA or control shRNA retroviruses, mixed at a 1:1 ratio, and transferred to recipient mice that were subsequently infected with LCMV. **d,** Representative flow cytometry plots (left) or quantification (right) of the proportion of transduced P14 cells in indicated tissues on day 12 of infection, normalized to splenic cells. **e,** Representative flow cytometry plots (left) and quantification (right) of the frequency of $CD69^+$ and $CD103^+$ cells of Con shRNA or Runx3 shRNA cells. Graphs indicate mean and representative of two independent experiments with n=2-10 mice (a) or n=6-8 mice per group (d,e), *$P<0.05$, **$P<0.005$, ***$P<0.0005$. Symbols represent an individual mouse (d,e).

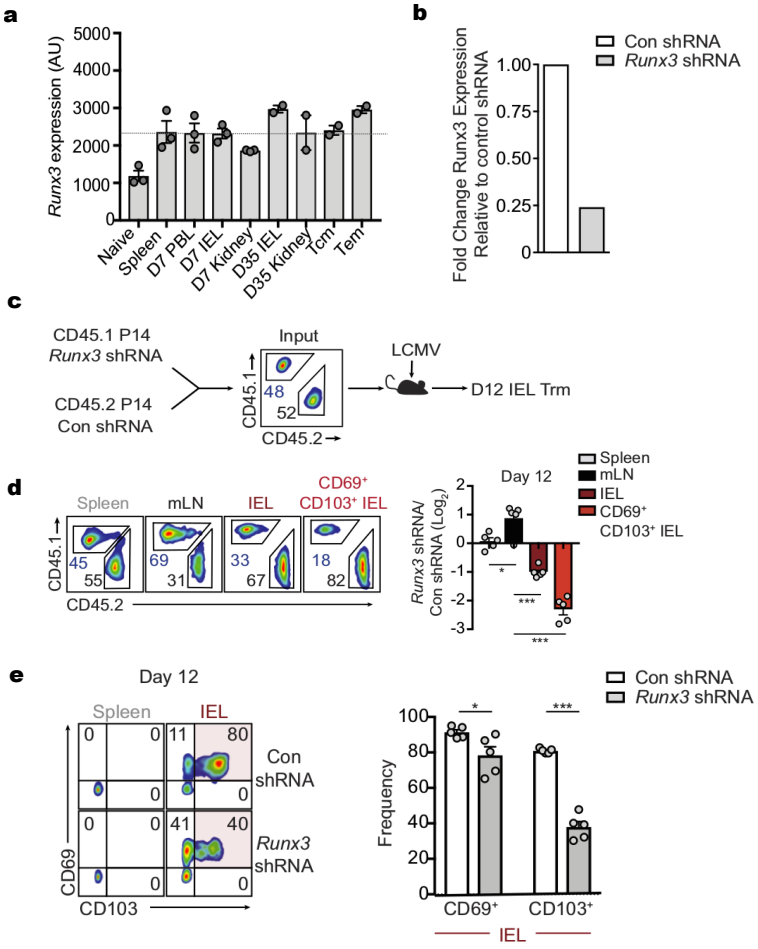

a
Runx3 expression (AU)
5000
4000
3000
2000
1000
0
Naive
Spleen
D7 PBL
D7 IEL
D7 Kidney
D35 IEL
D35 Kidney
Tcm
Tem
b
Fold Change Runx3 Expression Relative to control shRNA
1.00
0.75
0.50
0.25
0
Con shRNA
Runx3 shRNA
c
CD45.1 P14
Runx3 shRNA
CD45.2 P14
Con shRNA
Input
CD45.1
48
52
CD45.2
LCMV
D12 IEL Trm
d
Spleen
mLN
IEL
CD69+ CD103+ IEL
CD45.1
45
55
69
31
33
67
18
82
CD45.2
Day 12
Runx3 shRNA/ Con shRNA (Log2)
2
1
0
-1
-2
-3
*
***
***
Spleen
mLN
IEL
CD69+ CD103+ IEL
e
Day 12
Spleen
IEL
0
0
11
80
0
0
0
0
0
0
41
40
0
0
Con shRNA
Runx3 shRNA
CD69
CD103
Frequency
100
80
60
40
20
0
*
***
CD69+
CD103+
IEL
Con shRNA
Runx3 shRNA

Figure 3.4: **Runx3 regulates the core Trm transcriptional program to promote CD8$^+$ T cell tissue-residency. a,** Congenically distinct P14 cells were transduced with Runx3-RV (CD45.1+ cells) or GFP-RV (CD45.1.2+ cells), mixed at a 1:1 ratio, and transferred to recipient mice subsequently infected with LCMV. Representative flow cytometry plots (left) and quantification (right) of the ratio of transduced cells evaluated on days 8 and 12 or 13 of infection. **b,** Representative flow cytometry plots and quantification of the frequency of CD69$^+$ and CD103$^+$ cells. **c,** The core Trm transcriptional signature was generated by integration of all genes elevated in brain, lung, skin, IEL and kidney Trm relative to splenic memory counterpart cells, whereas the core circulating signature consists of genes elevated in splenic memory cells relative to the Trm populations (top panel). Relative expression of the core circulating and Trm genes was evaluated by RNAseq analysis of Runx3-RV, Runx3fl/fl or WT CD8+ T cells (middle panel). Gene set enrichment analysis of the circulating and Trm expression signatures of Runx3-RV, Runx3fl/fl or WT CD8+ T cells (bottom panel).

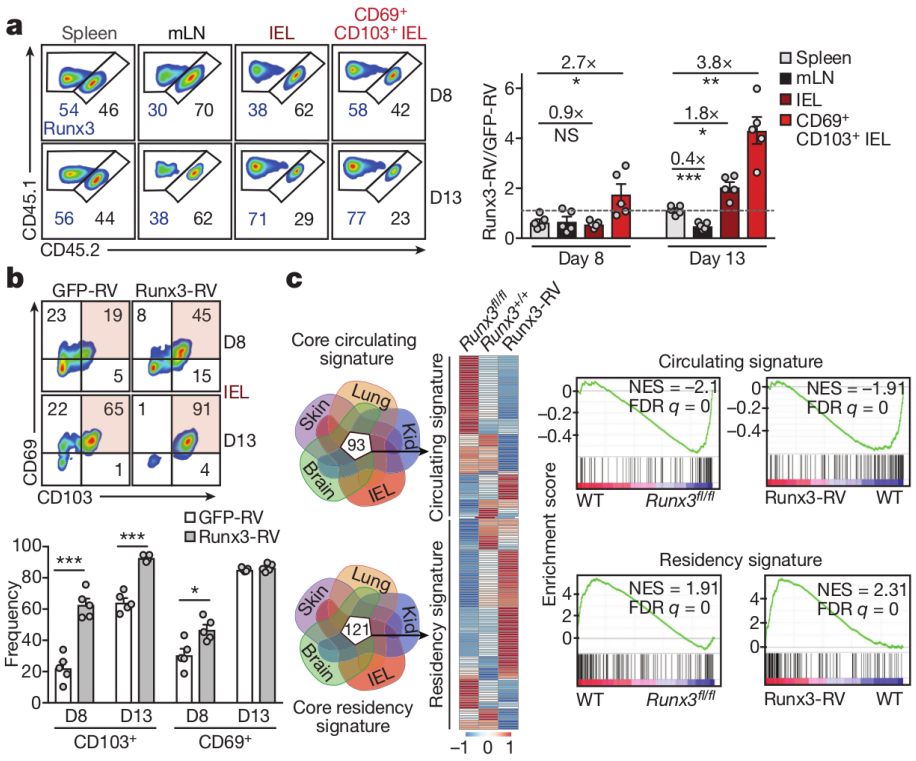
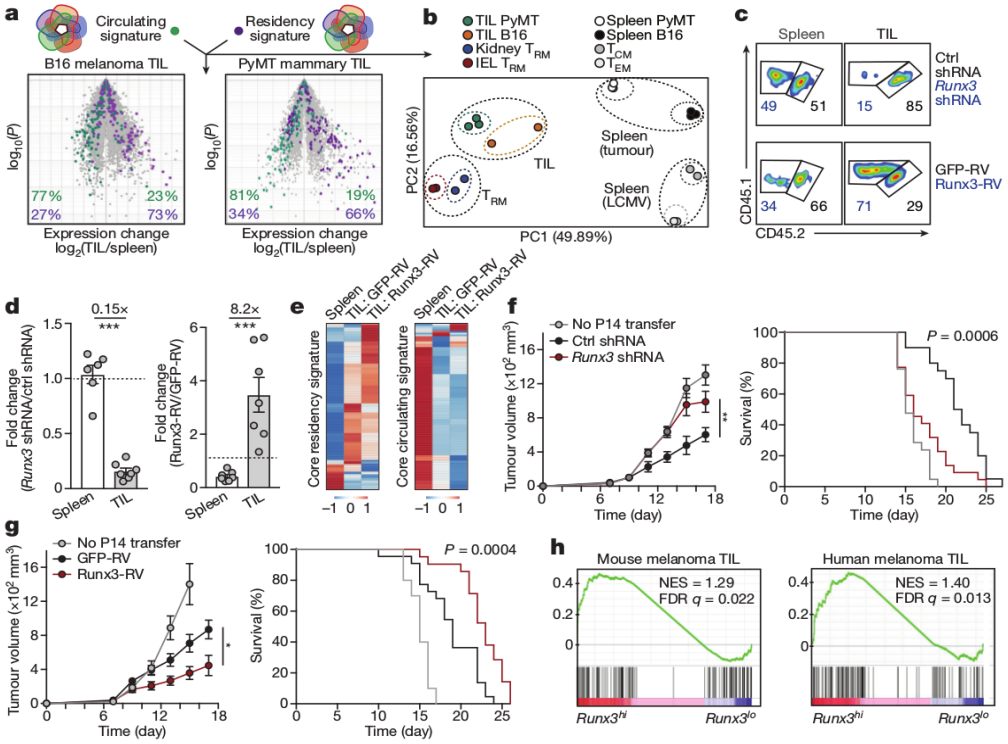

a
Spleen
mLN
IEL
CD69+ CD103+ IEL
Runx3
54 46
30 70
38 62
58 42
D8
56 44
38 62
71 29
77 23
D13
CD45.1
CD45.2
Runx3-RV/GFP-RV
2.7×
*
0.9×
NS
3.8×
**
1.8×
*
0.4×
***
Day 8
Day 13
Spleen
mLN
IEL
CD69+ CD103+ IEL
b
GFP-RV
Runx3-RV
23 19 5
8 45 15
D8
IEL
22 65 1
1 91 4
D13
CD69
CD103
GFP-RV
Runx3-RV
Frequency
***
***
*
D8
D13
D8
D13
CD103+
CD69+
c
Core circulating signature
Lung
Skin
Kid
Brain
IEL
93
Core residency signature
121
Runx3fl/fl
Runx3+/+
Runx3-RV
Circulating signature
Residency signature
−1 0 1
Circulating signature
NES = −2.1
FDR q = 0
NES = −1.91
FDR q = 0
Enrichment score
WT
Runx3fl/fl
Runx3-RV
WT
Residency signature
NES = 1.91
FDR q = 0
NES = 2.31
FDR q = 0
WT
Runx3fl/fl
Runx3-RV
WT

Figure 3.5: **CD8$^+$ TIL share transcriptional similarity with Trm and require Runx3 for tumor residency. a,** Comparison of the core Trm signature and circulating signature (from Fig. 4c) in B16 melanoma CD8$^+$ TIL or PyMT mammary tumor CD8$^+$ TIL relative to corresponding splenic CD8$^+$ T cells. **b,** Principal-component analysis of gene expression of core Trm and circulation gene sets for TIL, Trm or circulating CD8$^+$ T cell subsets. **c-d,** Congenically distinct P14 cells were transduced with retroviruses encoding Runx3 shRNAmir or Runx3-RV (CD45.1$^+$ cells) and control shRNAmir or GFP-RV (CD45.1.2$^+$ cells), mixed at a 1:1 ratio and transferred into mice with established B16-GP 3341 melanoma tumours. Flow plots (c) and graphs (d) indicate ratio of transduced cells. **e,** Relative expression of the core tissue-residency and core circulating gene sets in GFP-RV splenocytes, GFP-RV TILs, and Runx3-RV TILs following the same approach as in (c). **f,** Tumour growth and survival after adoptive transfer of the Con shRNA and Runx3 shRNA transduced cells. A log-rank (MantelCox) test was used to compare survival rates. **g,** Tumour growth and survival after adoptive transfer of the GFP-RV and Runx3-RV transduced cells. A log-rank (MantelCox) test was used to compare survival rates. **h,** Gene set enrichment analysis of the core Trm and core circulating gene signatures in Runx3^{hi} versus Runx3^{lo} from the single cell RNAseq analysis of mouse (left) and human (right) melanoma TILs. Graphs indicate mean of n=3-7 mice per group (c,d) from one representative experiment of 2-3 independent experiments or data pooled from 3 independent experiments consisting of n=10-21 mice per group (f, g), *P<0.05, **P<0.005, ***P<0.0005. Symbols represent an individual mouse (c,d).


a
Circulating signature
Residency signature
B16 melanoma TIL
PyMT mammary TIL
log10(P)
77%
23%
27%
73%
81%
19%
34%
66%
Expression change log2(TIL/spleen)
b
TIL PyMT
TIL B16
Kidney TRM
IEL TRM
Spleen PyMT
Spleen B16
TCM
TEM
PC2 (16.56%)
PC1 (49.89%)
TIL
TRM
Spleen (tumour)
Spleen (LCMV)
c
Spleen
TIL
Ctrl shRNA
Runx3 shRNA
GFP-RV
Runx3-RV
49
51
15
85
34
66
71
29
CD45.1
CD45.2
d
0.15×
***
8.2×
***
Fold change (Runx3 shRNA/ctrl shRNA)
Fold change (Runx3-RV/GFP-RV)
Spleen
TIL
e
Spleen
TIL: GFP-RV
TIL: Runx3-RV
Core residency signature
Core circulating signature
f
No P14 transfer
Ctrl shRNA
Runx3 shRNA
Tumour volume (×10² mm³)
Time (day)
**
P = 0.0006
Survival (%)
g
No P14 transfer
GFP-RV
Runx3-RV
Tumour volume (×10² mm³)
Time (day)
*
P = 0.0004
Survival (%)
h
Mouse melanoma TIL
NES = 1.29
FDR q = 0.022
Human melanoma TIL
NES = 1.40
FDR q = 0.013
Runx3hi
Runx3lo

3.7 References

[1] J. M. Schenkel, K. A. Fraser, L. K. Beura, K. E. Pauken, V. Vezys, and D. Masopust. Resident memory CD8 T cells trigger protective innate and adaptive immune responses. *Science*, 346(6205):98–101, 2014. ISSN 0036-8075. doi: 10.1126/science.1254536.

[2] Brian S. Sheridan, Quynh Mai Pham, Young Tae Lee, Linda S. Cauley, Lynn Puddington, and Leo Lefrançois. Oral infection drives a distinct population of intestinal resident memory cd8+ t cells with enhanced protective function. *Immunity*, 40(5):747–757, 2014. ISSN 10974180. doi: 10.1016/j.immuni.2014.03.007.

[3] T Cd, J Justin Milner, and Ananda W Goldrath. Transcriptional programming of tissue-resident memory CD8+ T cells. *Current Opinion in Immunology*, 51:162–169, 2018. ISSN 09527915. doi: S0952791517300985.

[4] Laura K. Mackay, Martina Minnich, Natasja A.M. Kragten, Yang Liao, Benjamin Nota, Cyril Seillet, Ali Zaid, Kevin Man, Simon Preston, David Freestone, Asolina Braun, Erica Wynne-Jones, Felix M. Behr, Regina Stark, Daniel G. Pellicci, Dale I. Godfrey, Gabrielle T. Belz, Marc Pellegrini, Thomas Gebhardt, Meinrad Busslinger, Wei Shi, Francis R. Carbone, René A.W. Van Lier, Axel Kallies, and Klaas P.J.M. Van Gisbergen. Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science*, 352(6284):459–463, 2016. ISSN 10959203. doi: 10.1126/science.aad2035.

[5] Laura K. Mackay, Erica Wynne-Jones, David Freestone, Daniel G. Pellicci, Lisa A. Mielke, Dane M. Newman, Asolina Braun, Frederick Masson, Axel Kallies, Gabrielle T. Belz, and Francis R. Carbone. T-box Transcription Factors Combine with the Cytokines TGF-β and IL-15 to Control Tissue-Resident Memory T Cell Fate. *Immunity*, 43(6):1101–1111, 2015. ISSN 10974180. doi: 10.1016/j.immuni.2015.11.008.

[6] Anusha-preethi Ganesan, James Clarke, Oliver Wood, Eva M Garrido-martin, Serena J Chee, Toby Mellows, Daniela Samaniego-castruita, Divya Singh, Grégory Seumois, Aiman Alzetani, Edwin Woo, Peter S Friedmann, Emma V King, Gareth J Thomas, Tilman Sanchez-elsner, Pandurangan Vijayanand, and Christian H Ottensmeier. Tissue-resident memory features are linked to the magnitude of cytotoxic T cell responses in human lung cancer. 18(8),

2017. doi: 10.1038/ni.3775.

[7] Fayçal Djenidi, Julien Adam, Aïcha Goubar, Aurélie Durgeau, Guillaume Meurice, Vincent de Montpréville, Pierre Validire, Benjamin Besse, and Fathia Mami-Chouaib. CD8 ⁺ CD103 ⁺ TumorInfiltrating Lymphocytes Are Tumor-Specific Tissue-Resident Memory T Cells and a Prognostic Factor for Survival in Lung Cancer Patients. *The Journal of Immunology*, 194(7):3475–3486, 2015. ISSN 0022-1767. doi: 10.4049/jimmunol.1402711.

[8] Laura K Mackay, Azad Rahimpour, Joel Z Ma, Nicholas Collins, Angus T Stock, Ming-Li Hafon, Javier Vega-Ramos, Pilar Lauzurica, Scott N Mueller, Tijana Stefanovic, David C Tscharke, William R Heath, Michael Inouye, Francis R Carbone, and Thomas Gebhardt. The developmental pathway for CD103+CD8+ tissue-resident memory T cells of skin. *Nature Immunology*, 14(12):1294–1301, 2013. ISSN 1529-2908. doi: 10.1038/ni.2744.

[9] D. Masopust, V. Vezys, E. J. Wherry, D. L. Barber, and R. Ahmed. Gut Microenvironment Promotes Differentiation of a Unique Memory CD8 T Cell Population. *The Journal of Immunology*, 176(4):2079–2083, 2006. ISSN 0022-1767. doi: 10.4049/jimmunol.176.4.2079.

[10] Nikhil S. Joshi, Weiguo Cui, Anmol Chandele, Heung Kyu Lee, David R. Urso, James Hagman, Laurent Gapin, and Susan M. Kaech. Inflammation Directs Memory Precursor and Short-Lived Effector CD8+ T Cell Fates via the Graded Expression of T-bet Transcription Factor. *Immunity*, 27(2):281–295, 2007. ISSN 10747613. doi: 10.1016/j.immuni.2007.07.010.

[11] Susan M Kaech, Joyce T Tan, E John Wherry, Bogumila T Konieczny, Charles D Surh, and Rafi Ahmed. Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nature immunology*, 4(12):1191–8, 2003. ISSN 1529-2908. doi: 10.1038/ni1009.

[12] John T Chang, E John Wherry, and Ananda W Goldrath. Molecular regulation of effector and memory T cell differentiation. *Nature reviews. Immunology*, 15(12), 2014. doi: 10.1038/ni.3031.1104.

[13] Laura K. Mackay and Axel Kallies. Transcriptional Regulation of Tissue-Resident Lymphocytes. *Trends in Immunology*, 38(2):94–103, 2017. ISSN 14714981. doi: 10.1016/j.it.2016.11.004.

[14] Bingfei Yu, Kai Zhang, J. Justin Milner, Clara Toma, Runqiang Chen, James P. Scott-Browne, Renata M. Pereira, Shane Crotty, John T. Chang, Matthew E. Pipkin, Wei Wang, and Ananda W. Goldrath. Epigenetic landscapes reveal transcription factors that regulate CD8 + T cell differentiation. *Nature Immunology*, 18(5):573–582, 2017. ISSN 15292916. doi: 10.1038/ni.3706.

[15] Chandra Sekhar Boddupalli, Shiny Nair, Simon M. Gray, Heba N. Nowyhed, Rakesh Verma, Joanna A. Gibson, Clara Abraham, Deepak Narayan, Juan Vasquez, Catherine C. Hedrick, Richard A. Flavell, Kavita M. Dhodapkar, Susan M. Kaech, and Madhav V. Dhodapkar. ABC transporters and NR4A1 identify a quiescent subset of tissue-resident memory T cells. *Journal of Clinical Investigation*, 126(10):3905–3916, 2016. ISSN 15588238. doi: 10.1172/JCI85329.

[16] Rahul Roychoudhuri, David Clever, Peng Li, Yoshiyuki Wakabayashi, Kylie M. Quinn, Christopher A. Klebanoff, Yun Ji, Madhusudhanan Sukumar, Robert L. Eil, Zhiya Yu, Rosanne Spolski, Douglas C. Palmer, Jenny H. Pan, Shashank J. Patel, Derek C. Macallan, Giulia Fabozzi, Han Yu Shih, Yuka Kanno, Akihiko Muto, Jun Zhu, Luca Gattinoni, John J. O'Shea, Klaus Okkenhaug, Kazuhiko Igarashi, Warren J. Leonard, and Nicholas P. Restifo. BACH2 regulates CD8 + T cell differentiation by controlling access of AP-1 factors to enhancers. *Nature Immunology*, 17(7):851–860, 2016. ISSN 15292916. doi: 10.1038/ni.3441.

[17] Runqiang Chen, Simon Bélanger, Megan A. Frederick, Bin Li, Robert J. Johnston, Nengming Xiao, Yun Cai Liu, Sonia Sharma, Bjoern Peters, Anjana Rao, Shane Crotty, and Matthew E. Pipkin. In vivo RNA interference screens identify regulators of antiviral CD4+ and CD8+ T cell differentiation. *Immunity*, 41(2):325–338, 2014. ISSN 10974180. doi: 10.1016/j.immuni.2014.08.002.

[18] Ruka Setoguchi, Masashi Tachibana, Yoshinori Naoe, Sawako Muroi, Kaori Akiyama, Chieko Tezuka, Tsukasa Okuda, and Ichiro Taniuchi. Repression of the transcription factor Th-POK by Runx complexes in cytotoxic T cell development. *Science*, 319(5864):822–825, 2008. ISSN 00368075. doi: 10.

1126/science.1151844.

[19] Fernando Cruz-Guilloty, Matthew E. Pipkin, Ivana M. Djuretic, Ditsa Levanon, Joseph Lotem, Mathias G. Lichtenheld, Yoram Groner, and Anjana Rao. Runx3 and T-box proteins cooperate to establish the transcriptional program of effector CTLs. *The Journal of Experimental Medicine*, 206(1): 51–59, 2009. ISSN 0022-1007. doi: 10.1084/jem.20081242.

[20] B. Grueter, M. Petter, T. Egawa, K. Laule-Kilian, C. J. Aldrian, A. Wuerch, Y. Ludwig, H. Fukuyama, H. Wardemann, R. Waldschuetz, T. Moroy, I. Taniuchi, V. Steimle, D. R. Littman, and M. Ehlers. Runx3 Regulates Integrin E/CD103 and CD4 Expression during Development of CD4-/CD8+ T Cells. *The Journal of Immunology*, 175(3):1694–1705, 2005. ISSN 0022-1767. doi: 10.4049/jimmunol.175.3.1694.

[21] Bernardo Sgarbi Reis, Aneta Rogoz, Frederico Azevedo Costa-Pinto, Ichiro Taniuchi, and Daniel Mucida. Mutual expression of the transcription factors Runx3 and ThPOK regulates intestinal CD4 + T cell immunity. *Nature Immunology*, 14(3):271–280, 2013. ISSN 15292916. doi: 10.1038/ni2518.

[22] L. M. Wakim, A. Woodward-Davis, R. Liu, Y. Hu, J. Villadangos, G. Smyth, and M. J. Bevan. The Molecular Signature of Tissue Resident Memory CD8 T Cells Isolated from the Brain. *The Journal of Immunology*, 189(7):3462–3471, 2012. ISSN 0022-1767. doi: 10.4049/jimmunol.1201305.

[23] Chandra Sekhar Boddupalli, Noffar Bar, Krishna Kadaveru, Michael Krauthammer, Natopol Pornputtapong, Zifeng Mai, Stephan Ariyan, Deepak Narayan, Harriet Kluger, Yanhong Deng, Rakesh Verma, Rituparna Das, Antonella Bacchiocchi, Ruth Halaban, Mario Sznol, Madhav V. Dhodapkar, and Kavita M. Dhodapkar. Interlesional diversity of T cell receptors in melanoma with immune checkpoints enriched in tissue-resident memory T cells. *JCI Insight*, 1(21), 2016. ISSN 2379-3708. doi: 10.1172/jci.insight.88955.

[24] A. L. Doedens, M. P. Rubinstein, E. T. Gross, J. A. Best, D. H. Craig, M. K. Baker, D. J. Cole, J. D. Bui, and A. W. Goldrath. Molecular Programming of Tumor-Infiltrating CD8+ T Cells and IL15 Resistance. *Cancer Immunology Research*, 4(9):799–811, 2016. ISSN 2326-6066. doi: 10.1158/2326-6066.CIR-15-0178.

[25] M. J.M. Gooden, G. H. De Bock, N. Leffers, T. Daemen, and H. W. Nijman. The prognostic influence of tumour-infiltrating lymphocytes in cancer: A systematic review with meta-analysis. *British Journal of Cancer*, 105(1):93–103, 2011. ISSN 00070920. doi: 10.1038/bjc.2011.189.

[26] Carmen S M Yong, Valerie Dardalhon, Christel Devaud, Naomi Taylor, Phillip K Darcy, and Michael H Kershaw. CAR T-cell therapy of solid tumors. *Immunology and Cell Biology*, 95(4):356–363, 2017. ISSN 0818-9641. doi: 10.1038/icb.2016.128.

[27] Jason D Buenrostro, Paul G Giresi, Lisa C Zaba, Howard Y Chang, and William J Greenleaf. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods*, 10(12):1213–8, 2013. ISSN 1548-7105. doi: 10.1038/nmeth.2688.

[28] Anthony Mathelier, Oriol Fornes, David J Arenillas, Chih Yu Chen, Grégoire Denay, Jessica Lee, Wenqiang Shi, Casper Shyr, Ge Tan, Rebecca Worsley-Hunt, Allen W Zhang, François Parcy, Boris Lenhard, Albin Sandelin, and Wyeth W Wasserman. JASPAR 2016: A major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*, 44(D1):D110—-D115, 2016. ISSN 13624962. doi: 10.1093/nar/gkv1176.

[29] Daniel E. Newburger and Martha L. Bulyk. UniPROBE: An online database of protein binding microarray data on protein-DNA interactions. *Nucleic Acids Research*, 37(SUPPL. 1):77–82, 2009. ISSN 03051048. doi: 10.1093/nar/gkn660.

[30] Arttu Jolma, Jian Yan, Thomas Whitington, Jarkko Toivonen, Kazuhiro R. Nitta, Pasi Rastas, Ekaterina Morgunova, Martin Enge, Mikko Taipale, Gonghong Wei, Kimmo Palin, Juan M. Vaquerizas, Renaud Vincentelli, Nicholas M. Luscombe, Timothy R. Hughes, Patrick Lemaire, Esko Ukkonen, Teemu Kivioja, and Jussi Taipale. DNA-binding specificities of human transcription factors. *Cell*, 152(1-2):327–339, 2013. ISSN 00928674. doi: 10.1016/j.cell.2012.12.009.

[31] Charles E. Grant, Timothy L. Bailey, and William Stafford Noble. FIMO: Scanning for occurrences of a given motif. *Bioinformatics*, 27(7):1017–1018,

2011. ISSN 13674803. doi: 10.1093/bioinformatics/btr064.

Chapter 4

An essential role of chromatin architectural protein CTCF in regulating effector and memory $CD8^+$ T cell differentiation

4.1 Introduction

TFs bind to distal and proximal regulatory elements including enhancers and promoters to modulate specific gene expression programs responsible for cell differentiation. Spatial control of long-range interactions between enhancers and promoters plays a critical role in regulating specific gene expression profiles[1, 2]. These long-range interactions are precisely regulated by a number of factors including CTCF, YY1, cohesin, condensin and Brg1[3, 4].

CTCF is a ubiquitous zinc finger protein essential for the establishment of three-dimensional chromatin organization[3]. CTCF can facilitate or insulate long-distance interactions between enhancers and promoters to activate or repress gene expression[5, 6]. In this way, CTCF is important for the regulation of hematopoietic cell lineage specification and function. For instance, in early thymocytes development, CTCF deficiency specifically impairs differentiation and proliferation of double-positive T cells due to increased expression of cyclin-CDK inhibitors p21 and p27[7]. CTCF also regulates distinct cytokine production in $CD4^+$ T helper (Th) cells in different polarization conditions[8, 9]. For example, under Th2 polarization condition, the deletion of CTCF abrogates the expression of IL-4, IL-5 and IL-13[9]. In contrast, disruption of CTCF binding site in the IL-21 locus promotes IL-21 expression induced by the proinflammatory cytokine IL-6, suggesting that CTCF functions as an enhancer-blocking insulator to repress IL-21 expression[8]. CTCF-dependent long-distance chromatin interactions are vital for TCR rearrangement and immunoglobulin recombination in T and B cells respectively[10], while in macrophages, CTCF is involved in fine-tuning the production of cytokines IL-10 and TNFα[11].

Similar to CTCF, another ubiquitously expressed TF YY1 has also been reported to regulate long-distance chromatin interactions and subsequently influences immune cell differentiation[4, 12, 13, 14]. Our lab recently discovered an essential function of YY1 in $CD8^+$ TE subset differentiation in response to pathogen infection[15]. However, the role of CTCF in effector and memory $CD8^+$ T cell differentiation remains largely unknown.

4.2 Results

4.2.1 CTCF is essential for terminal differentiation of effector CD8$^+$ T cells.

To investigate the role of CTCF in CD8$^+$ T cell differentiation in response to pathogen infection, we utilized a retroviral shRNA knockdown system to ablate the expression of CTCF in CD8$^+$ T cells. Using two different shRNAs targeting CTCF, we achieved 50% (shCTCF#1) and 80% (shCTCF#2) knockdown efficiency at the mRNA and protein levels (Figure 4.1a). Given the critical role of CTCF in TCR rearrangement, we specifically knocked down CTCF in TCR transgenic CD8$^+$ T cells to bypass TCR rearrangement defects from a lack of CTCF[10]. We transduced congenically distinct OT-I CD8$^+$ T cells with retrovirus expressing shRNA targeting CTCF (shCTCF) or CD19 as a negative control, and then co-transferred the cells mixed at a 1:1 ratio into recipient mice followed by Lm-OVA infection (Figure 4.1b). Seven days after infection, T cells with 80% knockdown of CTCF (shCTCF#2) failed to accumulate as the same level as T cells transduced with shCtrl. In contrast, T cells with 50% knockdown of CTCF (shCTCF#1) exhibited similar expansion as the control population (Figure 4.1c). The accumulation defect in the shCTCF#2 group is consistent with findings that full deletion of CTCF abrogated the proliferation of thymocytes[7].

To further test the function of CTCF in effector CD8$^+$ T cell differentiation, we characterized KLRG1 and CD127 expression after knockdown of CTCF. Both the frequency and number of KLRG1^{hi}CD127^{lo} populations (defined as the TE subset) were dramatically decreased in both shCTCF#1 and #2 groups compared

to shCtrl group (Figure 4.1d,e), suggesting that CTCF is crucial for terminal differentiation of effector $CD8^+$ T cells. Additionally, the frequency of $KLRG1^{lo}CD127^{hi}$ populations (defined as the MP subset) were significantly increased after the loss of CTCF (Figure 4.1d,e). Given that cytokine production is a hallmark function of effector $CD8^+$ T cells, we compared IFNγ, TNFα and IL-2 production after knockdown of CTCF. We found a comparable frequency of IFNγ and TNFα producing T cells between shCtrl and shCTCF groups (Figure 4.1f), however, the proportion of IL-2 producing cells was significantly higher after CTCF knockdown, consistent with an enhanced frequency of MP $CD8^+$ T cells that produce more IL-2 than TE $CD8^+$ T cells (Figure 4.1f)[16]. Taken together, our results indicate that shRNA-mediated knockdown of CTCF impairs terminal effector $CD8^+$ T cell differentiation.

4.2.2 The loss of CTCF impairs the differentiation of effector like memory subset and secondary effector $CD8^+$ T cells.

We examined the phenotype of $CD8^+$ T cells over the course of Lm-OVA infection, and observed a significant decrease of $KLRG1^{hi}CD127^{lo}$ population after CTCF knockdown at both effector and memory phase (Figure 4.2a,b). This $KLRG1^{hi}CD127^{lo}$ population that persists at the memory phase is considered as a subset of memory cells displaying effector-like features, suggesting that CTCF apprears to be essential for the differentiation of the long-lived $KLRG1^{hi}CD127^{lo}$ effector-like memory subset. It has been recently shown that $CD27^{lo}CD43^{lo}$ popula-

tion resembles a effector-like memory subset and shows optimal protection against Listeria[17]. To further test if CTCF is important for effector-like memory T cell differentiation, we characterized CD27 and CD43 expression at day 30 and found that the $CD27^{lo}CD43^{lo}$ population was indeed dramatically decreased after the loss of CTCF (Figure 4.2c). We observed an increased $CD27^{hi}$ population and higher expression of CXCR3 after knockdown of CTCF, both indicating a higher recall ability of memory T cells in shCTCF group compared to shCtrl group (Figure 4.2c,d). The $CD27^{hi}$ memory T cell pool is composed of central-memory (Tcm) and effector memory (Tem), based on the expression of CD62L[18, 19]. We observed that CTCF knockdown has no effect on CD62L expression, suggesting that the loss of CTCF may not affect Tcm and Tem subset differentiation (Figure 4.2e).

Although there is a higher percentage of $KLRG1^{lo}CD127^{hi}$ and $CD27^{hi}$ $CD8^{+}$ T cells after CTCF knockdown, both suggesting a better recall ability, it is necessary to examine if a more robust recall response is formed in shCTCF-transduced memory T cells when encountering the secondary antigen. To test this, we co-transferred shCtrl and shCTCF OT-I cells into recipient mice and initially infected with Lm-OVA. After 30 days of primary infection, we re-infected the mice with VSV-OVA(Figure 4.2f). Although there were more $KLRG1^{lo}CD127^{hi}$ memory T cells in the shCTCF group before re-infection, we observed an impaired expansion of secondary effector T cells deficient for CTCF, especially secondary $KLRG1^{hi}CD127^{lo}$ TE population (Figure 4.2g,h), suggesting that CTCF is critical for the ability of memory $CD8^{+}$ T cells to differentiate into secondary effector $CD8^{+}$ T cells.

4.2.3 CTCF suppresses Trm differentiation in response to LCMV infection.

Tissue-resident memory (Trm) $CD8^+$ T cells reside in non-lymphoid tissues, and are critical to provide the first-line protection at barrier surfaces such as skin, lung and gut[20, 21]. To investigate if the deficiency of CTCF impacts the differentiation of Trm, we co-transferred shCtrl and shCTCF P14 $CD8^+$ T cells into host mice followed by acute LCMV-armstrong infection (Figure 4.3a). Similar to bacterial infection, the frequency and number of the $KLRG1^{hi}CD127^{lo}$ subset were remarkably reduced after knockdown of CTCF in response to LCMV infection (Figure 4.3b,c). Furthermore, the frequency of $KLRG1^{lo}CD127^{hi}$ population, which has been shown to preferentially give rise to Trm cells, was significantly increased in the absence of CTCF (Figure 4.3b,c). Upon characterization of the Trm cells in the small intestine, we observed a significant increase of IEL Trm cells but not the memory cells in the spleen after knockdown of CTCF (Figure 4.3d), suggesting that CTCF might suppress Trm differentiation. Given that the activation marker CD69 and the retention molecule CD103 mark bona fide Trm in the small intestine[22], we further examined the expression of CD69 and CD103 and observed a robust increase of both the frequency and the number of $CD69^+CD103^+$ IEL Trm cells in the absence of CTCF (Figure 4.3e,f). Taken together, these results demonstrate that the loss of CTCF promotes the formation of Trm $CD8^+$ T cells.

4.2.4 CTCF controls gene expression of key TFs for effector and memory CD8$^+$ T cell differentiation.

The differentiation of distinct effector and memory CD8$^+$ T cell subsets is coordinated by a network of key TFs that regulate specific transcriptional programs. To better investigate how CTCF promotes terminal effector and effector-like memory T cell differentiation while suppressing gut-resident memory CD8$^+$ T cell differentiation, we profiled gene expression of key TFs for CD8$^+$ T cell subsets differentiation after knockdown of CTCF. To avoid the bias introduced by a decrease of KLRG1^{hi} population in the absence of CTCF, we co-transferred shCtrl and shCTCF transduced cells and sorted KLRG1^{hi} and KLRG1^{lo} subsets from each group at day 8 of Lm-OVA infection to assess key TFs gene expression (Figure 4.4a). Upon CTCF knockdown, we observed a decrease of T-bet and Eomes expression and an increase of Prdm1, Hobit and Tcf7 expression in both subsets(Figure 4.4b). This observation is consistent with increased Trm formation after CTCF knockdown since T-bet and Eomes have been shown to repress CD103$^+$ Trm development while Prdm1 and Hobit are essential for Trm differentiation[23]. The repression of Prdm1 by CTCF has been reported in germinal center B cells, indicating that the CTCF-Prdm1 regulatory axis is conserved in both B and T cells[13]. However, the increase of Prdm1 in the absence of CTCF cannot fully explain the loss of the TE subset since Prdm1 promotes differentiation of this subset[24]. This observation suggests that additional key TFs for the TE subset are dramatically impacted after knockdown of CTCF. Indeed, extensive studies have shown that T-bet is critical for terminal effector and effector-like memory

$CD8^+$ T cell differentiation[25, 16]. Given the downregulation of T-bet in the absence of CTCF and the corresponding phenotype, we postulated that T-bet is a direct target of CTCF (Figure 4.4c). To test this, we performed ChIP-seq of CTCF in naive and TE $CD8^+$ T cells to detect CTCF binding at the T-bet locus. We observed an increase of CTCF binding at active enhancer regions of the T-bet gene in the TE subset, correlating with a higher deposition of the active enhancer mark H3K27ac in TE compared to naive $CD8^+$ T cells (Figure 4.4d). This suggests that CTCF might activate T-bet expression by directly binding to active enhancers. In sum, these results demonstrate a role for CTCF in regulating the expression of key TFs responsible for promoting the TE subset and suppressing Trm $CD8^+$ T cell differentiation.

4.2.5 Heterozygotic mutation of CTCF in patients impacts the TE gene signature in peripheral blood lymphocytes.

To investigate if the defect in the TE subset differentiation in shCTCF groups (leading to 50% knockdown) can be translated into a human phenotype, we analyzed published data from patients with *de novo* mutations of CTCF, who showed intellectual disability, microcephaly, and growth retardation[26]. RNA-seq of peripheral blood lymphocytes from three patients with heterozygous mutations of CTCF and eight healthy control individuals were performed previously[26]. To detect if the TE subset gene signature was impacted in patients with haploinsufficiency of CTCF, we re-analyzed the RNAseq data specifically focusing on

enrichment of TE and MP gene signatures. We observed that 62.5% of the TE genes were downregulated in patients while the TE but not the MP gene signature was significantly enriched in healthy control individuals compared to patients, suggesting that the TE subset might be affected in patients with CTCF mutation (Figure 4.5a,b). Similarly, we found a significant enrichment of the Tem gene signature in the healthy group compared to patients, partially due to a remarkable overlap between TE and Tem gene signatures (Figure 4.5c,d). To further examine the influence of haploinsufficiency of CTCF on the effector subsets' transcriptional programs, we categorized the TE- or MP-associated genes into key TFs, cytokines, chemokines and other markers and compared their expression between healthy individuals and patients. We observed a downregulation of TE-associated TFs such as Blimp1, Zeb2 and T-bet and an upregulation of MP-associated TFs including Tcf7 and Id3 in patients compared to healthy controls (Figure 4.5e). Additionally, we found that expression of IL-2 was dramatically increased in patients, in line with the mouse phenotype of a higher production of IL-2 after knockdown of CTCF (Figure 4.5e). Taken together, these analyses suggest that the TE subset transcriptional program may be impacted in humans with haploinsufficiency of CTCF, however, differences in subset compositions may also impact this and we are working to obtain new samples with age-matched controls.

4.3 Discussion

Here, we discovered a novel function for the genome organizer CTCF in promoting the TE subset and suppressing Trm $CD8^+$ T cell differentiation. Upon

depletion of CTCF using two different CTCF shRNAs with graded knockdown efficiency, we observed a dose-dependent role for CTCF in expansion and proliferation of effector $CD8^+$ T cells. This is consistent with the findings that CTCF participates in the proliferation and cell-cycle progression of $\alpha\beta$ T cells in the thymus in a dose-dependent manner[7]. Therefore, the CTCF shRNA with intermediate knockdown efficiency allowed us to interrogate the role of CTCF in effector and memory T cell differentiation due to the mild impact on cell proliferation. We further observed impaired differentiation of terminal-effector, effector-like memory and secondary effector $CD8^+$ T cells after knockdown of CTCF. These subsets likely require CTCF to establish a high-ordered three-dimentional chromatin organization and a fine-tuned spatial control of gene expression to maintain their "effector-like" transcriptional program. Intriguingly, the loss of CTCF promotes the differentiation of Trm especially $CD103^+$ Trm, possibly due to upregulation of TFs that are critical for Trm differentiation including Blimp1 and Hobit and downregulation of TFs that inhibit Trm differentiation including T-bet. The increased binding of CTCF at the T-bet enhancer region and increased three-dimentional interactions at the T-bet locus (data not shown) in the TE subset compared to naive T cells suggest that CTCF may facilitate enhancer-promoter interactions of the T-bet gene to promote TE-specific expression of T-bet. Given the essential role of T-bet in promoting the TE subset while suppressing Trm differentiation, the lower level of T-bet upon depletion of CTCF partially explains the observed phenotypes including the defect in the TE and effector-like memory cells and the increased Trm population. It is very likely that CTCF can activate or repress expression of key genes for specific subset differentiation through facilitating or block-

ing the enhancer-promoter interactions in addition to its central role in establishing high-order chromatin structures including TADs. A global characterization of the transcriptional program and long-range interactions in T cells after knockdown of CTCF would provide a snapshot of the CTCF-dependent enhancer-promoter interactome that is critical for T cell differentiation.

4.4 Acknowledgement

Chapter 4, in part, is currently being prepared for submission for publication of the material. Yu B, Goldrath AW.

The dissertation author was a primary investigator and the first author of this material.

4.5 Methods

4.5.1 Mice

All mice were on a C57BL6/J background and maintained/bred in specific-pathogen-free conditions in accordance with the Institutional Animal Care and Use Committees (IACUC) of the University of California, San Diego (UCSD) or purchased from the Jackson Laboratory, including wild type mice, TCR transgenic P14 and OT-I mice.

4.5.2 T cell transfer and infection

For shRNA knockdown experiments, total 2×10^5 OT-I or P14 cells were transferred into recipient mice by intravenous injection. For ChIP-seq experiment, 1×10^4 OT-I cells were transferred into host mice intravenously. One day after T cell transfer, mice were infected intravenously by 5×10^3 Lm-OVA or intraperitoneally by 2×10^5 PFU LCMV-Armstrong. For secondary infection, mice were rechallenged by intravenous injection of 1×10^6 VSV-OVA.

4.5.3 Tissue processing and cell preparation.

Spleens and lymph nodes were processed to get single-cell solution. Red blood cells were lysed with ACK buffer (140 mM NH4Cl and 17 mM Tris-base, pH 7.4). For isolation of lymphocytes from small intestine IEL compartment, Peyers patches were removed and the intestine was cut longitudinally and subsequently cut laterally into 0.5-1 cm^2 pieces that were then incubated with 0.154 mg/ml dithioerythritol (DTE) in 10% HBSS/HEPES bicarbonate for 30 min at 37C while stirring. Then single-cell suspensions were separated using a 44/67% Percoll density gradient to isolate lymphocytes.

4.5.4 Antibodies and flow cytometry.

The following antibodies were obtained from eBioscience: CD8a (53-6.7), CD8b (eBio H35-17.2), CD62L (MEL-14), CD27 (LG-7F9), CD43(1B11), CD127 (A7R34), KLRG1 (2F1), CD103 (2E7), CD69 (H1.2F3), CD45.1 (A20-1.7), CD45.2 (104), CXCR3 (CXCR3-173), IFNγ (XMG1.2), TNF (MP6-XT22), IL2 (JES6-

SH4), and T-bet (4B10). For intracellular cytokine staining, splenocytes were re-stimulated with XX OVA peptide for 4 hours with the presence of 1X Protein Transport Inhibitor Cocktail (ebioscience). To better preserve the ametrine reporter signal in transduced populations, samples were fixed and permeabilized using the Cytofix/Cytoperm Fixation/Permeabilization kit (BD). For flow cytometry, all events were acquired on a BD LSRFortessa X-20 or a BD LSRFortessa and analyzed using FlowJo software. Cell sorting was performed on BD FACSAria or BD FACSAria Fusion instruments.

4.5.5 shRNA knockdown.

The CTCF shRNA sequence were shown as follows. CTCF#1: CCAGATGAAGACTGAAGTCAT; CTCF#2:GCAGAGCATTCAGAACAGTGA. For tran sfections, PLAT-E cells were seeded in the middle 60 wells of a 96-well flat-bottom plate at a density of $4 - 6 \times 10^4$ cells per well one day before transfection. Next, each well was individually transfected with 0.2 μg of DNA from each pLMPd-Amt clone and 0.2 μg of pCL-Eco using TransIT-LT1 (Mirus) in Opti-MEM medium. The medium was replaced by T cell medium after 16h and the retroviral supernatant were collected 36, 48 and 60h after transfection. For $CD8^+$ T cell activation, naive $CD8^+$ T cells from spleens and lymph nodes were negatively enriched using MACS columns and 2×10^5 OT-I or P14 cells were plated in the middle 60 wells of 96-well round-bottom plates pre-coated with 100 μg/ml goat anti-hamster IgG (H+L, Thermo Fisher Scientific) and 1 μg/ml anti-CD3 (145-2C11) and 1 μg/ml anti-CD28 (37.51) (eBioscience). Culture medium was replaced after 18h of activation with retroviral supernatant mixed with 50 μM BME and 8 μg/ml polybrene

(Millipore) followed by spin-infection (60 min centrifugation at 2000 rpm, 37 C). The plate was incubated at 37 C for 2h after spin-infection and then the retroviral supernatant were removed and replaced by T cell medium for 24h. Cogenically-distinct OT-I or P14 cells that were transduced with shCD19 or shCTCF, were mixed at 1:1 ratio after 24h transduction and then transferred into host mice intravenously followed by Lm-OVA or LCMV-Arm infection.

4.5.6 RT-qPCR.

$0.2 - 1 \times 10^6$ cells were sorted directly into Trizol and RNA was extracted by chloroform and isopropanol precipitation. CDNA was synthesized using Superscript II (Life Technologies) following manufacturer's instructions and quantitative PCR (qPCR) was performed using the Stratagene Brilliant II Syber Green master mix (Agilent Technologies). The qPCR primers were used as follows: CTCF-forward: 5'-TGACACAGTCATAGCCCGAAAA-3', CTCF-reverse: 5'-TGCCGT GATCAATATAGGAATGC-3', Hprt-foward: 5-GGCCAGACTTTGTTGGATTT-3, Hprt-reverse: 5-CAACTTGCGCTCATCTTAGG-3.

4.5.7 Western Blotting.

Naive $CD8^+$ T cells isolated from spleens and lymph nodes were in vitro activated by 1 μg/ml anti-CD3 and 1 μg/ml anti-CD28 for 24h and then transduced with shCD19 or shCTCF and then cultured for 48h with 100 U/ml IL-2. 2×10^6 Ametrine+ cells were sorted, lysed and resolved by SDS-PAGE. CTCF (07-729, Millipore) and β-actin (Santa Cruz Biotechnology) were detected by immunoblotting.

4.5.8 ChIP-seq and computational analysis.

5×10^6 KLRG1^{hi} CD8$^+$ T cells were sorted from spleens and lymph nodes from mice infected for 8 d with Lm-OVA, fixed in 1% formaldehyde for 10 min, and subsequently quenched with 0.125 M glycine. Cells were lysed and sonicated to generate 250-500 bp fragments using Bioruptor. 30μl magnetic-dynabeads were mixed with 5μg CTCF antibody (07-729, Millipore) in 500μl blocking buffer and rotated for at least 4h and then mixed with diluted lysate and rotated overnight at 4C. Beads were washed, eluted and reverse-crosslinked at 65C overnight and then treated with RNAse for 30 min at 37C and Proteinase K at 55C for 1h. DNA was purified by Zymo DNA Clean & Concentrator kit (Zymo Research). The ChIPed DNA was end-repaired using End-it End-repair kit(Epicentre) and then added an A base to the 3 end of DNA fragments using Klenow (NEB). Then DNA was ligated with adaptors using quick DNA ligase (NEB) at 25C for 15 min followed by size selection of 200-400 bp using AMPure SPRI beads(Beckman Coulter). The adaptor ligated DNA was amplified using NEBNext High-Fidelity 2X PCR master mix(NEB). Then the amplified library was sized selected as 200-400bp using SPRI beads and quantified by Qubit dsDNA HS assay kit (ThermoFisher). Finally the library was sequenced using Hiseq 2500 for single-end 50bp sequencing to get around 10 million reads for each sample. ChIP-seq sequence reads were aligned to mm10 using bowtie2 and analyzed to generate bedgraph files for UCSC genome browser visualization using Homer. For human PBL RNA-seq analysis, the normalized gene expression data was downloaded from GEOXX and the volcano plots were generated by Genepattern multiplot studio module. GSEA was performed by using the GSEA module in GenePattern, and the normalized enrichment scores

and false-discovery rate q values were determined by using the permutation test.

4.6 Figures

Figure 4.1: **CTCF is essential for terminal differentiation of effector CD8$^+$ T cells. a,** Naive CD8$^+$ T cells were first activated in vitro by anti-CD3 and anti-CD28 for 24h and then transduced with retrovirus expressing shRNA targeting CD19 (shCtrl) or two different shRNAs targeting CTCF (shCTCF) and further cultured for 48h. The CTCF shRNA knockdown efficiency was measured at both CTCF mRNA level using quantitative RT-qPCR (left, results were presented as fold change relative to shCtrl) and CTCF protein level using Western Blot (right, β-actin was used as loading control). **b,** Schematic of experimental design. Cogenically distinct OT-I CD8$^+$ T cells were transduced with retrovirus expressing shCtrl or shCTCF, mixed at a 1:1 ratio, then transferred to host mice followed by Lm-OVA infection for 7 days. **c,** Representative FACS plots showing the ratio of transduced cells between shCtrl and shCTCF(left: shCTCF#1; right: shCTCF#2) isolated from spleens before co-transfer and 7 days after transfer and Lm-OVA infection. Numbers in plots represent the percentage of cells. **d,** Representative FACS plots showing the expression of KLRG1 and CD127 of transduced cells isolated from spleen on day 7 after infection. **e,** Quantification of the frequency (left) and the absolute number (right) of KLRG1$^{\text{hi}}$CD127$^{\text{lo}}$ and KLRG1$^{\text{lo}}$CD127$^{\text{hi}}$ population from shRNA-transduced cells represented in (d). **f,** Representative FACS plots showing the expression of intracellular cytokines IFNγ and TNFα (top left) or IL2 (bottom left) in transduced cells isolated from spleens on day 7 Lm-OVA infection and restimulated by OVA peptide for 4h. Quantification of the frequency of IFNγ^+TNFα^+ (top right) and IFNγ^+IL2$^+$ (bottom right) populations from shRNA-transduced cells. Data shown are representative of two independent experiment; n=3-5 mice per group. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$ (two-tailed paired Students t test).

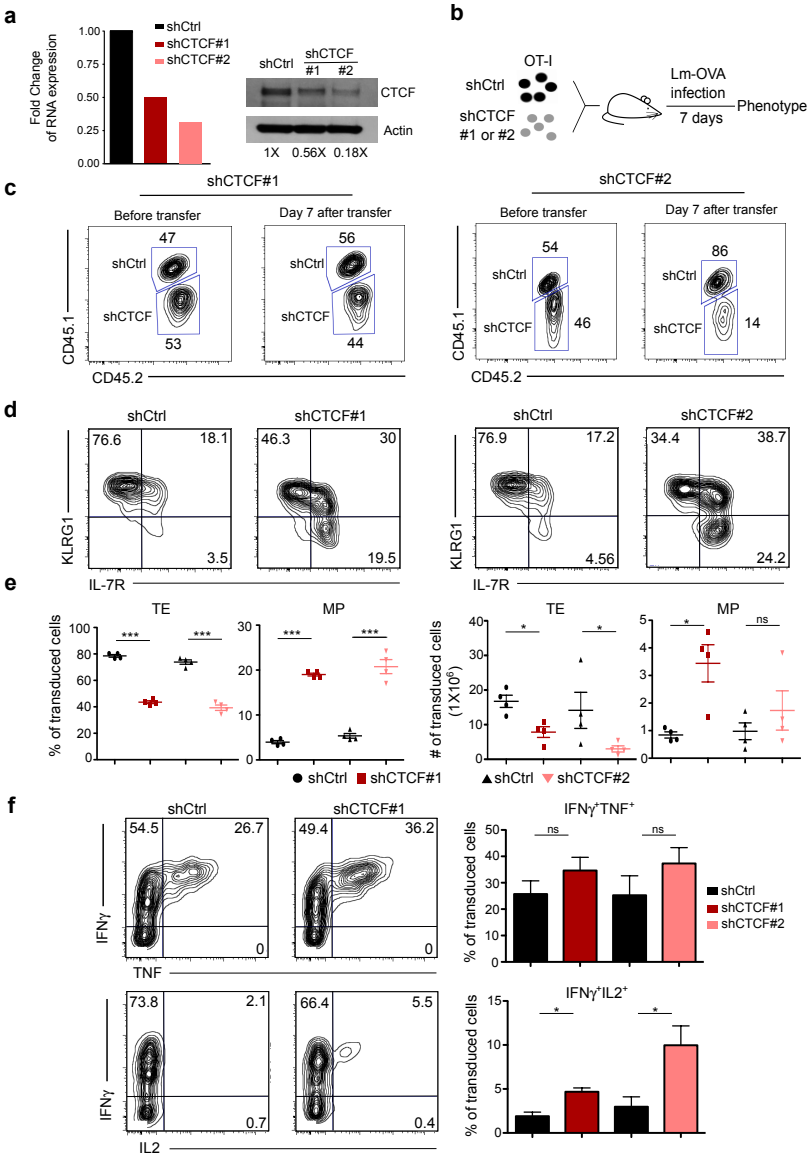

a
Fold Change of RNA expression
shCtrl
shCTCF#1
shCTCF#2
shCtrl
shCTCF #1 #2
CTCF
Actin
1X 0.56X 0.18X
b
OT-I
shCtrl
shCTCF #1 or #2
Lm-OVA infection
7 days
Phenotype
c
shCTCF#1
shCTCF#2
Before transfer
Day 7 after transfer
CD45.1
CD45.2
d
shCtrl
shCTCF#1
shCTCF#2
KLRG1
IL-7R
e
TE
MP
% of transduced cells
# of transduced cells (1X10⁶)
shCtrl
shCTCF#1
shCTCF#2
f
IFNγ
TNF
IL2
IFNγ⁺TNF⁺
IFNγ⁺IL2⁺

Figure 4.2: **The loss of CTCF impacts specific memory subset and secondary effector T cell differentiation. a,** shCtrl and shCTCF#1 transduced OT-I $CD8^+$ T cells were co-transferred into host mice followed by Lm-OVA infection. The frequency of $KLRG1^{hi}CD127^{lo}$ (left) and $KLRG1^{lo}CD127^{hi}$ (right) population from shRNA-transduced cells in PBL on indicated days after infection is shown. **b-c,** Representative FACS plot showing the expression of KLRG1 and CD127 (b), CD43 and CD27 (c) of transduced cells isolated from spleens on day 30 after infection. **d,** Quantification of the frequency of $KLRG1^{hi}CD127^{lo}$ and $KLRG1^{lo}CD127^{hi}$ population in transduced cells in spleen on day 30 after infection. **e,** Representative FACS plot showing the expression of CD44 and CD62L (left) of transduced cells same as in (B) and the frequency of $CD44^{hi}CD62L^{hi}$ in transduced cells (right). **f,** Schematic of recall experiment design. shCtrl and shCTCF#1 transduced OT-I $CD8^+$ T cells were co-transferred into host mice followed by Lm-OVA infection. 30 days after the primary infection, recipient mice were re-infected with VSV-OVA. **g,** The frequency of transduced cells in PBL on day 30 of Lm-OVA infection before re-infection. **h,** Kinetics of the number of total transduced cells (left), $KLRG1^{hi}$ (middle) and $KLRG1^{lo}$ (right) population in PBL per million before and after VSV-OVA reinfection on indicated days. Numbers in plots represent the percentage of cells. Data shown are representative of two independent experiment; n=3-5 mice per group. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$ (two-tailed paired Students t test).

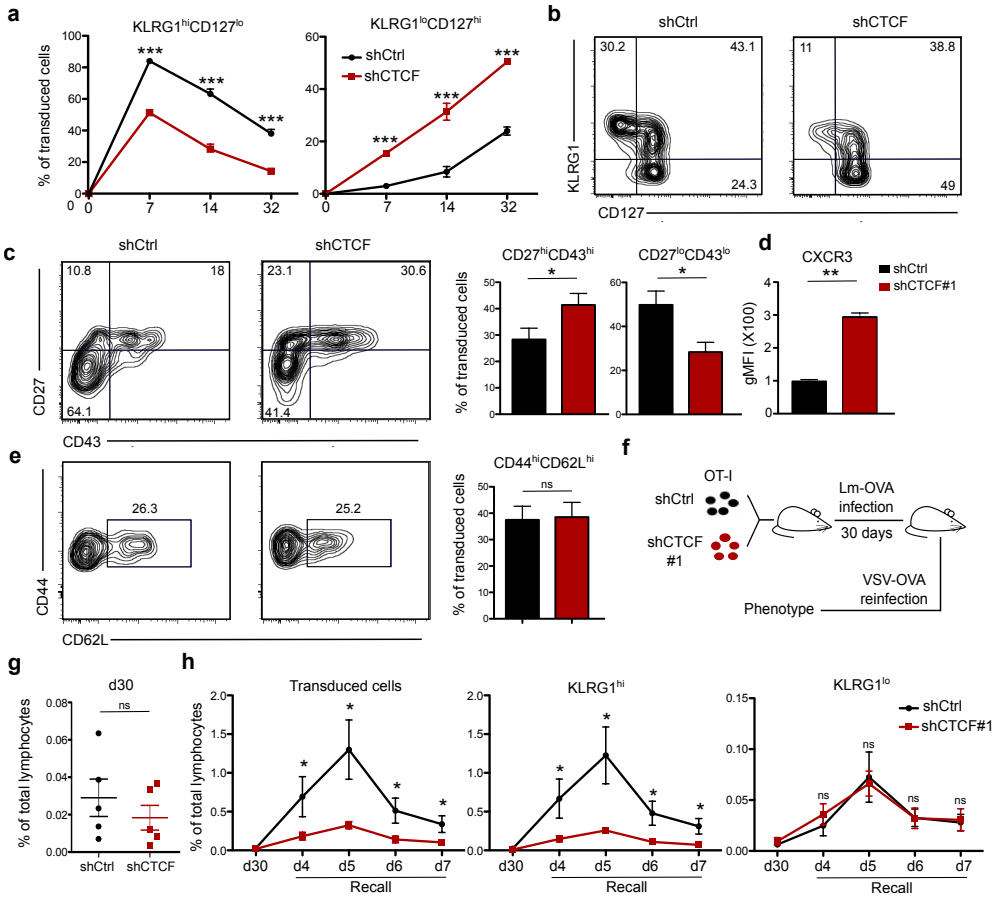

a
KLRG1hiCD127lo
KLRG1loCD127hi
% of transduced cells
shCtrl
shCTCF
b
shCtrl
shCTCF
KLRG1
CD127
30.2
43.1
24.3
11
38.8
49
c
shCtrl
shCTCF
CD27
CD43
10.8
18
64.1
23.1
30.6
41.4
CD27hiCD43hi
CD27loCD43lo
% of transduced cells
d
CXCR3
gMFI (X100)
shCtrl
shCTCF#1
e
CD44
CD62L
26.3
25.2
CD44hiCD62Lhi
ns
f
OT-I
shCtrl
shCTCF #1
Lm-OVA infection
30 days
VSV-OVA reinfection
Phenotype
g
d30
% of total lymphocytes
ns
shCtrl
shCTCF
h
Transduced cells
KLRG1hi
KLRG1lo
% of total lymphocytes
d30
d4
d5
d6
d7
Recall
shCtrl
shCTCF#1
ns

Figure 4.3: **CTCF suppresses Trm differentiation in response to LCMV infection. a,** Schematic of experimental design. Congenically distinct P14 $CD8^+$ T cells were transduced with retrovirus expressing shCtrl or shCTCF#1, mixed at 1:1 ratio, and then transferred to host mice followed by LCMV-Armstrong infection. Lymphocytes isolated from spleen and small intestine were analyzed on day 15 after infection. **b,** Representative FACS plot showing expression of KLRG1 and CD127 of transduced cells isolated from spleen. **c-d,** Quantification of the frequency (c) and the absolute number (d) of $KLRG1^{hi}CD127^{lo}$ (left) and $KLRG1^{lo}CD127^{hi}$ (right) population in transduced cells isolated from spleen on day 15 after infection. **e,** Quantification of the frequency of shCtrl and shCTCF transduced cells in small intestine (IEL) and spleen (Sp) as indicated in (a). **f,** Representative FACS plot showing expression of CD69 and CD103 of transduced cells from IEL. **g,** Quantification of the frequency (left) and the absolute number (right) of $CD69^+CD103^+$ populations of transduced cells from IEL on day 15 after infection. Numbers in plots represent the percentage of cells. Data shown are representative (b,c, d and f) and cumulative (e and g) of three independent experiment; n=3-5 mice per group. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$ (two-tailed paired Students t test).

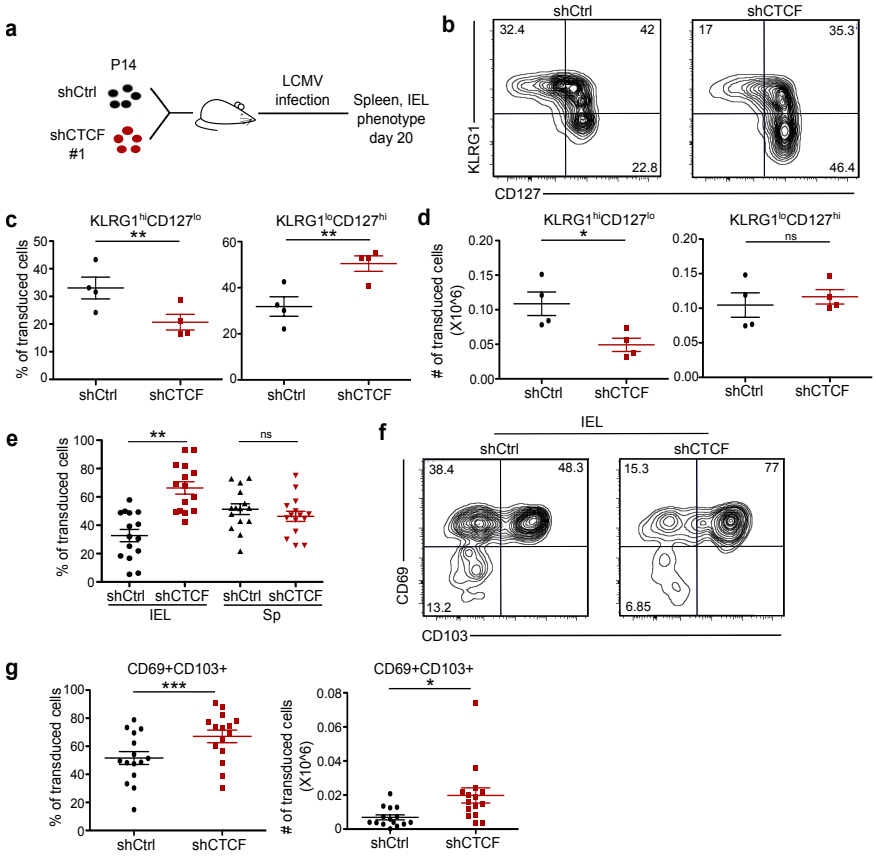

a
P14
shCtrl
shCTCF #1
LCMV infection
Spleen, IEL phenotype day 20
b
shCtrl
shCTCF
32.4
42
22.8
17
35.3
46.4
KLRG1
CD127
c
KLRG1hiCD127lo
KLRG1loCD127hi
% of transduced cells
shCtrl
shCTCF
d
KLRG1hiCD127lo
KLRG1loCD127hi
# of transduced cells (X10^6)
ns
e
% of transduced cells
IEL
Sp
f
IEL
shCtrl
shCTCF
38.4
48.3
13.2
15.3
77
6.85
CD69
CD103
g
CD69+CD103+
% of transduced cells
# of transduced cells (X10^6)

Figure 4.4: **CTCF regulates gene expression of key TFs for effector and memory CD8$^+$ T cell differentiation.** **a,** Schematic of experimental design. Congenically distinct OT-I CD8$^+$ T cells were transduced with retrovirus expressing shCtrl or shCTCF#1, mixed at a 1:1 ratio, and then transferred to host mice followed by Lm-OVA infection. KLRG1^{hi} and KLRG1^{lo} subsets from shCtrl and shCTCF transduced populations isolated from spleen were sorted in Trizol on day 8 after infection followed by RNA extraction and RT-qPCR. **b,** A heatmap generated from RT-qPCR data showing the fold change of gene expression of key TFs in shCtrl versus shCTCF cells. **c,** A proposed model for CTCF-T-bet regulatory axis responsible for specific effector and memory CD8$^+$ T cell subsets differentiation. **d,** ChIP-seq analysis of CTCF and H3K27ac at T-bet loci in naive (TN) and terminal effector (TE) CD8$^+$ T cells. The differential CTCF binding peaks are highlighted in grey.

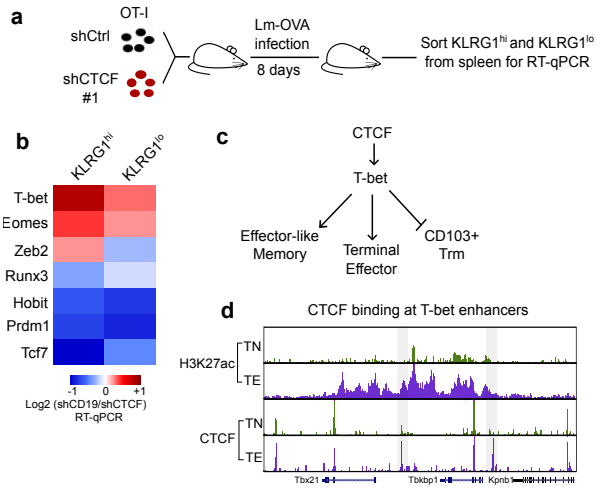
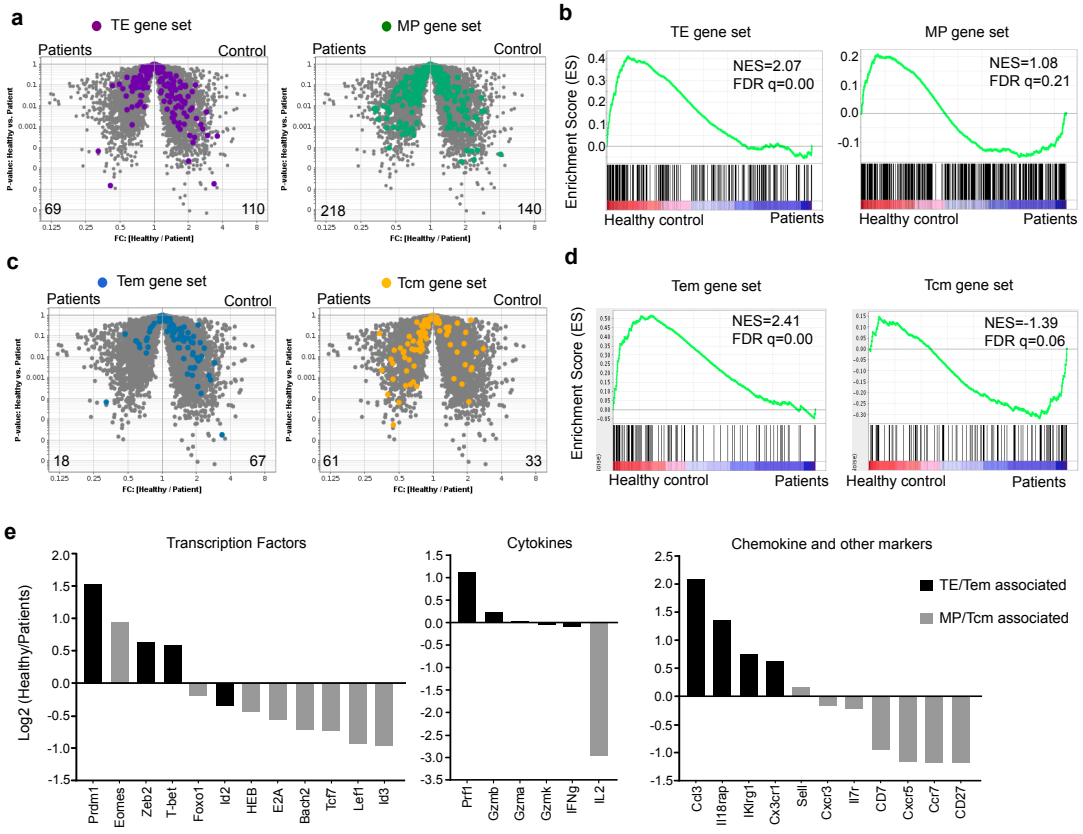

a
OT-I
shCtrl
shCTCF
#1
Lm-OVA
infection
8 days
Sort KLRG1hi and KLRG1lo
from spleen for RT-qPCR
b
KLRG1hi
KLRG1lo
T-bet
Eomes
Zeb2
Runx3
Hobit
Prdm1
Tcf7
-1
0
+1
Log2 (shCD19/shCTCF)
RT-qPCR
c
CTCF
T-bet
Effector-like
Memory
Terminal
Effector
CD103+
Trm
d
CTCF binding at T-bet enhancers
H3K27ac
TN
TE
CTCF
TN
TE
Tbx21
Tbkbp1
Kpnb1

Figure 4.5: **Heterozygotic mutation of CTCF in patients impacts the TE gene signature in peripheral blood lymphocytes. a,** Volcano plot comparing gene expression between healthy control individuals and patients harboring *de novo* mutation of CTCF overlaid with genes that are >1.5 fold upregulated in TE compared to MP (TE gene set, left) or in MP compared to TE (MP gene set, right). **b,** Gene Set Enrichment Analysis (GSEA) of RNA-seq from lymphocytes between healthy control individuals and patients with TE-specific (left) or MP-specific (right) gene signature. NES: normalized enrichment score; FDR q: false discovery rate q value. **c,** Volcano plot comparing gene expression between healthy control individuals and patients overlaid with genes that are >1.5 fold upregulated in Tem compared to Tcm (Tem gene set, left) or in Tcm compared to Tem (Tcm gene set, right). **d,** GSEA of RNA-seq from lymphocytes between healthy control individuals and patients with Tem- (left) or Tcm- (right) associated gene signature. **e,** Relative expression of key transcription factors (left), cytokines (middle), and chemokines (right) between healthy control and patients. Bars colored in black are genes associated with TE and Tem differentiation and bars colored in grey are genes associated with MP and Tcm differentiation.


a
TE gene set
MP gene set
Patients
Control
69
110
218
140
b
TE gene set
MP gene set
Enrichment Score (ES)
NES=2.07
FDR q=0.00
NES=1.08
FDR q=0.21
Healthy control
Patients
c
Tem gene set
Tcm gene set
18
67
61
33
d
Tem gene set
Tcm gene set
NES=2.41
FDR q=0.00
NES=-1.39
FDR q=0.06
e
Transcription Factors
Cytokines
Chemokine and other markers
Log2 (Healthy/Patients)
TE/Tem associated
MP/Tcm associated
Prdm1
Eomes
Zeb2
T-bet
Foxo1
Id2
HEB
E2A
Bach2
Tcf7
Lef1
Id3
Prf1
Gzmb
Gzma
Gzmk
IFNg
IL2
Ccl3
Il18rap
Klrg1
Cx3cr1
Sell
Cxcr3
Il7r
CD7
Cxcr5
Ccr7
CD27

4.7 References

[1] Adriana Gonzalez-sandoval and Susan M Gasser. On TADs and LADs : Spatial Control Over Gene Expression. *Trends in Genetics*, 32(8):485–495, 2016. ISSN 0168-9525. doi: 10.1016/j.tig.2016.05.004.

[2] Jesse R. Dixon, Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398): 376–380, 2012. ISSN 0028-0836. doi: 10.1038/nature11082.

[3] Chin-Tong Ong and Victor G Corces. CTCF: an architectural protein bridging genome topology and function. *Nature reviews. Genetics*, 15(4):234–46, 2014. ISSN 1471-0064. doi: 10.1038/nrg3663.

[4] Michael L. Atchison. Function of YY1 in long-distance DNA interactions, 2014. ISSN 16643224.

[5] Matthias Merkenschlager and Duncan T Odom. CTCF and Cohesin : Linking Gene Regulatory Elements with Their Targets. *Cell*, 152(6):1285–1297, 2013. ISSN 0092-8674. doi: 10.1016/j.cell.2013.02.029.

[6] Rodolfo Ghirlando and Gary Felsenfeld. CTCF : making the right connections. pages 881–891. doi: 10.1101/gad.277863.116.This.

[7] Helen Heath, Claudia Ribeiro De, Frank Sleutels, Suzanne Van De Nobelen, Kam-wing Ling, Joost Gribnau, Rainer Renkawitz, Frank Grosveld, W Hendriks, and Niels Galjart. CTCF regulates cell cycle progression of ab T cells in the thymus. (May):2839–2850, 2008. doi: 10.1038/emboj.2008.214.

[8] Joo-hong Park, Yeeun Choi, Min-ji Song, Keunhee Park, Jong-joo Lee, and Hyoung-pyo Kim. Dynamic Long-Range Chromatin Interaction Controls Expression of IL-21 in CD4 + T Cells. 2017. doi: 10.4049/jimmunol.1500636.

[9] Claudia Ribeiro de Almeida, Helen Heath, Sanja Krpic, Gemma M Dingjan, Jan Piet van Hamburg, Ingrid Bergen, Suzanne van de Nobelen, Frank Sleutels, Frank Grosveld, Niels Galjart, and Rudi W Hendriks. Critical role for the transcription regulator CCCTC-binding factor in the control of Th2 cy-

tokine expression. *Journal of immunology (Baltimore, Md. : 1950)*, 182(2): 999–1010, 2009. ISSN 1550-6606. doi: 10.4049/jimmunol.182.2.999.

[10] Han-yu Shih, Jiyoti Verma-gaur, Ali Torkamani, Ann J Feeney, Niels Galjart, and Michael S Krangel. Tcra gene recombination is supported by a Tcra enhancer- and CTCF-dependent chromatin hub. (D), 2012. doi: 10.1073/pnas.1214131109.

[11] Tatjana Nikolic, Dowty Movita, Margaretha E H Lambers, Claudia Ribeiro De Almeida, Paula Biesta, Kim Kreefft, Marjolein J W De Bruijn, Ingrid Bergen, Niels Galjart, and Andre Boonstra. The DNA-binding factor Ctcf critically controls gene expression in macrophages. *Cellular and Molecular Immunology*, 11(1):58–70, 2014. ISSN 1672-7681. doi: 10.1038/cmi.2013.41.

[12] Abraham S Weintraub, Charles H Li, Alicia V Zamudio, James E Bradner, Nathanael S Gray, Richard A Young, Abraham S Weintraub, Charles H Li, Alicia V Zamudio, Alla A Sigova, Nancy M Hannett, and Daniel S Day. Article YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell*, 171(7): 1573–1579.e28, 2017. ISSN 0092-8674. doi: 10.1016/j.cell.2017.11.008.

[13] Jose M Ligos, Niels Galjart, Ester Marina-za, Arantxa Pe, and Almudena R Ramiro. CTCF orchestrates the germinal centre transcriptional program and prevents premature plasma cell differentiation . (May), 2017. doi: 10.1038/ncomms16067.

[14] Xuan Pan, Madhusudhan Papasani, Yi Hao, Marco Calamito, Fang Wei, William J Quinn Iii, Arindam Basu, Junwen Wang, Suchita Hodawadekar, Kristina Zaprazna, Huifei Liu, Yang Shi, David Allman, Michael Cancro, and Michael L Atchison. YY1 controls Igκ repertoire and B-cell development, and localizes with condensin on the Igκ locus. *The EMBO journal*, 32(8):1168–82, 2013. ISSN 1460-2075. doi: 10.1038/emboj.2013.66.

[15] Bingfei Yu, Kai Zhang, J. Justin Milner, Clara Toma, Runqiang Chen, James P. Scott-Browne, Renata M. Pereira, Shane Crotty, John T. Chang, Matthew E. Pipkin, Wei Wang, and Ananda W. Goldrath. Epigenetic landscapes reveal transcription factors that regulate CD8 + T cell differentiation. *Nature Immunology*, 18(5):573–582, 2017. ISSN 15292916. doi: 10.1038/ni.3706.

[16] Nikhil S. Joshi, Weiguo Cui, Anmol Chandele, Heung Kyu Lee, David R. Urso, James Hagman, Laurent Gapin, and Susan M. Kaech. Inflammation Directs Memory Precursor and Short-Lived Effector CD8+ T Cell Fates via the Graded Expression of T-bet Transcription Factor. *Immunity*, 27(2):281–295, 2007. ISSN 10747613. doi: 10.1016/j.immuni.2007.07.010.

[17] Janelle A. Olson, Cameron McDonald-Hyman, Stephen C. Jameson, and Sara E. Hamilton. Effector-like CD8+ T Cells in the Memory Population Mediate Potent Protective Immunity. *Immunity*, 38(6):1250–1260, 2013. ISSN 10747613. doi: 10.1016/j.immuni.2013.05.009.

[18] E. John Wherry, Volker Teichgräber, Todd C. Becker, David Masopust, Susan M. Kaech, Rustom Antia, Ulrich H. von Andrian, and Rafi Ahmed. Lineage relationship and protective immunity of memory CD8T cell subsets. *Nature Immunology*, 4(3):225–234, 2003. ISSN 15292908. doi: 10.1038/ni889.

[19] W M Flanagan, B Corthesy, R J Bram, and G R Crabtree. Two subsets of memory T lymphocytes with distinct homing potential and effector functions. 6(October):6270–6273, 1979.

[20] J. M. Schenkel, K. A. Fraser, L. K. Beura, K. E. Pauken, V. Vezys, and D. Masopust. Resident memory CD8 T cells trigger protective innate and adaptive immune responses. *Science*, 346(6205):98–101, 2014. ISSN 0036-8075. doi: 10.1126/science.1254536.

[21] T Cd, J Justin Milner, and Ananda W Goldrath. Transcriptional programming of tissue-resident memory CD8+ T cells. *Current Opinion in Immunology*, 51:162–169, 2018. ISSN 09527915. doi: S0952791517300985.

[22] D. Masopust, V. Vezys, E. J. Wherry, D. L. Barber, and R. Ahmed. Gut Microenvironment Promotes Differentiation of a Unique Memory CD8 T Cell Population. *The Journal of Immunology*, 176(4):2079–2083, 2006. ISSN 0022-1767. doi: 10.4049/jimmunol.176.4.2079.

[23] Laura K. Mackay, Martina Minnich, Natasja A.M. Kragten, Yang Liao, Benjamin Nota, Cyril Seillet, Ali Zaid, Kevin Man, Simon Preston, David Freestone, Asolina Braun, Erica Wynne-Jones, Felix M. Behr, Regina Stark, Daniel G. Pellicci, Dale I. Godfrey, Gabrielle T. Belz, Marc Pellegrini, Thomas

Gebhardt, Meinrad Busslinger, Wei Shi, Francis R. Carbone, René A.W. Van Lier, Axel Kallies, and Klaas P.J.M. Van Gisbergen. Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science*, 352(6284):459–463, 2016. ISSN 10959203. doi: 10.1126/science.aad2035.

[24] Rachel L. Rutishauser, Gislâine A. Martins, Sergey Kalachikov, Anmol Chandele, Ian A. Parish, Eric Meffre, Joshy Jacob, Kathryn Calame, and Susan M. Kaech. Transcriptional Repressor Blimp-1 Promotes CD8+ T Cell Terminal Differentiation and Represses the Acquisition of Central Memory T Cell Properties. *Immunity*, 31(2):296–308, 2009. ISSN 10747613. doi: 10.1016/j.immuni.2009.05.014.

[25] Nikhil S Joshi, Weiguo Cui, Claudia X Dominguez, Jonathan H Chen, Timothy W Hand, and Susan M Kaech. Increased numbers of preexisting memory CD8 T cells and decreased T-bet expression can restrain terminal differentiation of secondary effector and memory CD8 T cells. *The Journal of Immunology*, 187(8):4068–4076, 2011. ISSN 1550-6606. doi: 10.4049/jimmunol.1002145.

[26] Anne Gregor, Martin Oti, Evelyn N. Kouwenhoven, Juliane Hoyer, Heinrich Sticht, Arif B. Ekici, Susanne Kjaergaard, Anita Rauch, Hendrik G. Stunnenberg, Steffen Uebe, Georgia Vasileiou, André Reis, Huiqing Zhou, and Christiane Zweier. De novo mutations in the genome organizer CTCF cause intellectual disability. *American Journal of Human Genetics*, 93(1):124–131, 2013. ISSN 00029297. doi: 10.1016/j.ajhg.2013.05.007.

Chapter 5

Conclusion

In response to pathogen infection, the $CD8^+$ T cell commitment to a specific effector or memory subset is sophisticatedly mediated by a series of cell-intrinsic TF regulators that can sense cell-extrinsic factors including antigen load, TCR strength and cytokine environment. Efficient transcriptional regulation requires cooperation between TF regulators and chromatin remodelers to instruct TF binding to specific regulatory elements encoded in the chromatin landscape in a spatial-temporal manner. Considerable effort has been invested to identify key TF regulators responsible for T cell fate determination; however, how the chromatin landscape/state impacts the TF behavior in distinct subset differentiation remains poorly explored. To fill this gap, we performed ChIP-seq for histone marks H3K4me1, H3K4me3, H3K27ac and H3K27me3 to characterize the chromatin state for effector and memory $CD8^+$ T cells. Ultimately, we generated a catalog of regulatory elements containing enhancers and promoters for T cell differentiation (Chapter 2). Combining the chromatin state and chromatin accessibility generated from ATAC-seq data, we showed that subset-specific enhancers are established by

key TFs to regulate unique transcriptional programs for T cell subset differentiation (Chapter 2). Proper and stable long-range interactions between enhancers and promoters are critical to maintain the expression of key genes responsible for specific subset differentiation. Indeed, disruption of enhancer-promoter looping by knocking down the genome organizer CTCF dramatically impaired terminal effector, effector-like memory and secondary effector $CD8^+$ T cell differentiation (Chapter 4). Interestingly, a mild pertubation of CTCF promoted $CD103^+$ Trm differentiation, possibly due to a decreased expression of T-bet which suppresses Trm formation (Chapter 4). It is still unclear how CTCF, a constitutively expressed chromatin architecture protein functions in a subset-specific manner. Future studies examining how interactions between CTCF and lineage-specific TFs or chromatin remodelers impacts gene expression and genome organization will provide insight into the role of CTCF in cell-type-specific regulation.

Naive T cells responding to infection can establish and effector T cell population with a range of differentiation states. The spectra of cell states are generated when the TF-gene regulatory circuits integrate varied inputs. These inputs include *trans* inputs such as expression levels of TFs and *cis* regulatory modules including enhancer and promoter elements that are bound by TFs and chromatin remodelers to regulate gene expression[1]. Regulatory circuits can be identified using a *trans* genomic approach based on the correlation of gene expression between TFs and putative targets[2, 3, 4]. However, this co-expression pattern does not mean a causal relationship or a direct binding of a TF to its target. In addition, this approach may miss the TF regulators that are not differentially expressed in distinct subsets but regulate different gene targets owning to the impact of chromatin

state, chromatin accessibility, chromatin organization and availability of cofactors. On the other hand, the *cis* genomic approach using TF motifs enriched in *cis* regulatory elements of genes to identify TF-gene regulatory circuits also has its own limitations. For example, it is difficult to assign the distal enhancers to the gene promoters given the varying long distances between enhancers and their targets. A binding of TFs on the regulatory elements of genes does not necessarily give rise to functional expression outputs. To resolve the limitations of these genomic approaches, it is essential to exploit more than one strategy and integrate multiple datasets including global expression profiles, chromatin state/accessibility and chromatin organization information. A comprehensive integration of these data types requires a more systemic and sophisticated computational method to decipher the complex molecular network, extract the useful transcriptional circuits and prioritize the biologically meaningful TF regulators with increased magnitude and complexity. In this dissertation, we developed a computational framework combining gene expression profiles, chromatin state and chromatin accessibility datasets to identify key molecular drivers for specific T cell subset differentiation (Chapter 2). We first determined accessible *cis* regulatory elements from ChIP-seq of histone modifications and ATAC-seq, and then scanned for enriched TF motifs and assigned them to gene targets to construct a TF-gene network. Differential gene expression across different states and conditions can be assigned to gene targets with distinct weights in the network. Finally, we applied the PageRank agolrithm, a Google-based webpage ranking agolrithm to prioritize TF regulators based on the quantity and quality of gene targets potentially regulated by this TF (Chapter 2). Future incorporation of chromatin organization data like HiC or HiChIP

of enhancer marks will provide a more accurate linking of TFs to their gene targets. In addition to assigning weights to nodes (gene targets), we can further assign distinct weights to the edges (the links between TF and gene targets) based on TF-gene co-expression pattern, TF-binding affinity and TF-bound enhancer activity. To validate the computational discoveries, we used retroviral-mediated shRNA knockdown to uncover novel roles of two TFs YY1 and Nr3c1 in regulating TE and MP subsets, respectively (Chapter 2). We further expanded this shRNA knockdown strategy to a large-scale genetic perturbation using a pooled shRNA screen combined with a PageRank-based computational screen which narrows down the predicted TF candidates. This double screen enabled us to probe molecular drivers in a highly efficient and unbiased way. In particular, we identified Runx3, a transcription factor with ubiquitous expression during differentiation, to have an essential role in the regulation of the tissue-residency gene expression program in non-lymphoid tissues and tumors (Chapter 3). Importantly, when we manipulated this regulatory circuit by overexpressing Runx3 in $CD8^+$ T cells, we observed a higher infiltration and residency of T cells in tumors that ultimately prevented tumor growth (Chapter 3). Taken together, our PageRank-based computational framework can be applied to any cell differentiation or tissue development to decode the transcriptional network and identify important TF regulators. With the advance of single-cell RNA-seq, single-cell ATAC-seq and CRISPR genome editing technologies, we can further refine the molecular network at the single cell level and precisely manipulate regulatory circuits to control specific cell type differentiation and function.

5.1 References

[1] I Amit, A Regev, and N Hacohen. Strategies to discover regulatory circuits of the mammalian immune system. *Nature Reviews Immunology*, 11(12):873–880, 2011. ISSN 1474-1733. doi: 10.1038/nri3109.

[2] J Adam Best. Transcriptional insights into the CD8+ T cell response to infection and memory T cell formation. *Nature immunology*, 29(4):997–1003, 2013. ISSN 15378276. doi: 10.1016/j.biotechadv.2011.08.021.Secreted.

[3] Xunshan Ding, Jamie Boney-montoya, Bryn M Owen, Angie L Bookout, Colbert Coate, David J Mangelsdorf, and Steven A Kliewer. Network Analysis Reveals Centrally Connected Genes and Pathways Involved in CD8+ T Cell Exhaustion versus Memory. 16(3):387–393, 2013. ISSN 1878-5832. doi: 10.1016/j.cmet.2012.08.002.

[4] Guangan Hu and Jianzhu Chen. A genome-wide regulatory network identifies key transcription factors for memory CD8 T-cell development. *Nature communications*, 4:2830, 2013. ISSN 2041-1723. doi: 10.1038/ncomms3830.