

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

The identity and function of ROR γ t extrathymic Aire-expressing cells

Permalink

<https://escholarship.org/uc/item/3b6483tj>

Author

Wang, Jiaxi

Publication Date

2024

Peer reviewed|Thesis/dissertation

The identity and function of RORγt extrathymic Aire-expressing cells

by
Jiaxi Wang

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:



EDFBCD1914E943B...

Tiffany Scharschmidt

Chair

DocuSigned by:



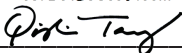
DocuSigned by: 14...

James Gardner



DocuSigned by: 406...

Mark Anderson



9BFF9AD92ACD4BA...

Qizhi Tang

Committee Members

This dissertation is dedicated to:

My parents Yanbin Wang and Huaxin Wang

My husband Tin Ngo

My emotional support Mona

Acknowledgements

I could not have completed my PhD without the support of many people before and during my graduate school. I would like to acknowledge the support and mentorship from my thesis advisor Jay Gardner and everyone from the Gardner lab. I joined Jay's lab as his first graduate student and have learned so much throughout my journey with his constant support both scientifically and mentally. Thank you to everyone from the Gardner lab who always made me feel welcomed and not alone. I have truly found my second family here in California and I appreciate everyone's support through thick and thin over the last five years. Thank you to Alex Gupta, Fang Xie, Hong Sun, Matt Arvedson, Anita Qualls, Nolan Horner, Joe Germino, Han Yin, Eva Gillis-Buck, Peter Zhong, Juan Du, and everyone else from the Bapat lab, Anderson lab, Waterfield lab, and many others from the Diabetes Center at UCSF. I would also like to thank my thesis committee members, Tiffany Scharschmidt, Mark Anderson, and Qizhi Tang, for all their insights and support on my thesis.

Thank you to my husband Tin Ngo who has always been there for me since college and all the way through the hardships of graduate school. I also could not have done it without the constant emotional support of my first pet and best cat Mona.

Chapter 2 Acknowledgements

We thank members of the Bluestone, Anderson, and Gardner laboratories for technical support and discussions including A. Chan, C. Miller, Y. Wang, J. Bridge, J. Phipps, H. Sun, and I. Proekt. We thank C. Brown, S. Rudensky, R. Kedmi, and D. Littman for discussions. We thank the members of the UCSF Flow Cytometry Core Facility for assistance with cell sorting. This work was supported by the ARCS Fellowship (J.W.), a Stanford Science Fellowship (C.A.L.), a Parker Scholar Ward (C.A.L.), NIH T32 FAVOR (J.L.B.), NIH P01AI118688 (M.S.A.), the ASTS Fellowship in Transplantation grant (J.M.G.), the UCSF Sandler Fellows PSSP grant (J.M.G.), NIH K08230188 and U01CA260852 (A.T.S.), a Cancer Research Institute Technology Impact Award (A.T.S.), and the Juvenile Diabetes Research Foundation and Lupus Research Alliance

(A.T.S.). Chapter 2 is reprinted as it appears in: Wang, J., Lareau, C.A., *et al.* Sci. Immunol. **6**, eabl5053 (2021). <https://www.science.org/doi/10.1126/sciimmunol.abl5053>

Chapter 3 Acknowledgements

We thank members of the Anderson and Gardner labs for their helpful discussions. Research in the Gardner laboratory is supported by the NIH (R01 AI145858), the Pew Biomedical Scholars Program, the Burroughs Wellcome Fund, the UCSF Physician-Scientist Scholars Program (PSSP), the Parker Institute for Cancer Immunotherapy, and the W.M. Keck Foundation. I.-H.S. was supported by NIH T32 training grant (5T32AI007334). J.W. is supported by a NIAID F31 predoctoral fellowship (1F31AI172348) and the ARCS Fellowship. A.Q. is supported by a MSTP NIH training grant (5T32GM141323-02). We thank Mark Anderson, Dan Littman, and Ranit Kedmi for their feedback and insight. We thank members of the UCSF Flow Cytometry Core Facility and Single Cell Genomics Core for their technical support. Chapter 3 is reprinted as a manuscript in preparation.

Contributions

Chapter 2 contributions

Jiaxi Wang^{1†}, Caleb A. Lareau^{2,3†}, Jhoanne L. Bautista⁴, Alexander R. Gupta^{1,4}, Katalin Sandor², Joe Germino¹, Yajie Yin², Matthew P. Arvedson¹, Gabriella C. Reeder¹, Nathan T. Cramer¹, Fang Xie^{1,4}, Vasilis Ntranos^{1,5}, Ansuman T. Satpathy^{2,3}, Mark S. Anderson^{1,6}, James M. Gardner^{1,4*}

¹Diabetes Center, University of California San Francisco, San Francisco, CA, USA.

²Department of Pathology, Stanford University, Stanford, CA, USA.

³Parker Institute for Cancer Immunotherapy, San Francisco, CA, USA.

⁴Department of Surgery, University of California San Francisco, San Francisco, CA, USA.

⁵Department of Epidemiology and Biostatistics, Computational Health Sciences Institute, University of California, San Francisco, San Francisco, CA, USA.

⁶Department of Medicine, University of California San Francisco, San Francisco, CA, USA.

*Corresponding author. Email: james.gardner@ucsf.edu

†These authors contributed equally to this work.

Author contributions: J.M.G. and M.S.A. conceived the study. J.W., J.L.B., C.A.L., A.R.G., and J.M.G. designed and performed experiments and analyzed data. C.A.L., J.G., J.M.G., V.N., and N.T.C. performed and supervised computational analysis. F.X. generated and cared for mice and performed experiments. J.M.G., J.W., C.A.L., and A.T.S. wrote and reviewed the manuscript. All authors read and approved the manuscript.

Chapter 3 contributions

Im-Hong Sun^{*1,2}, Jiaxi Wang^{*2}, Anita E. Qualls^{*2}, Han Yin², Nolan K. Horner^{1,2}, Matthew P. Arvedson², Joe Germino², Sheng Zhong^{1,2}, Vasilis Ntranos^{2,3}, Richard M. Locksley⁴⁻⁶, Roberto R. Ricardo-Gonzalez⁷, James M. Gardner^{†1,2}

¹Department of Surgery, University of California San Francisco, San Francisco, CA, USA.

²Diabetes Center, University of California San Francisco, San Francisco, CA, USA.

³Department of Epidemiology and Biostatistics, Computational Health Sciences Institute, University of California, San Francisco, San Francisco, CA, USA.

⁴Department of Medicine, University of California, San Francisco, San Francisco, CA, USA.

⁵Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA, USA.

⁶University of California, Howard Hughes Medical Institute, San Francisco, San Francisco, CA, USA.

⁷Department of Dermatology, University of California, San Francisco, San Francisco, CA, USA.

*These authors contributed equally to this work

[†]Corresponding author. Email: james.gardner@ucsf.edu

I.-H.S., J.W., A.Q., and J.M.G. designed experiments. I.-H.S., J.W., and A.Q. performed all experiments, analysed flow cytometry data, and wrote the manuscript. H.Y. analysed all computational data with oversight from J.G. and V.N. N.K.H., M.P.A., and S.Z. helped with experiments. R.M.L. and R.R.R-G. provided critical reagents and scientific feedback. J.M.G. provided scientific advice and expertise. All authors read and approved the manuscript.

The identity and function of ROR γ t extrathymic Aire-expressing cells

Jessica Wang

Abstract

Multicellular organisms require complex adaptive immune systems to maintain host defense from pathogens, but such systems must be precisely controlled to distinguish self from non-self and prevent autoimmunity. The autoimmune regulator (Aire) gene, a key transcriptional regulator expressed in medullary thymic epithelial cells (mTECs), has been shown to be crucial for central tolerance by inducing tissue specific antigen (TSA) expression in mTECs. Interestingly, we and others have also found extrathymic Aire-expressing cells (eTACs) in the secondary lymphoid organs such as the spleen and lymph nodes. eTACs are hematopoietic antigen-presenting cells (APCs) and consist of two similar cell types: CCR7⁺ Aire-expressing migratory dendritic cells (AmDCs) and an Aire-high population co-expressing Aire and retinoic acid receptor–related orphan receptor γ t (ROR γ t) that we termed Janus cells (JCs). Functionally, eTACs are capable of enforcing deletion and anergy on self-reactive T cells, and self-antigen expression in eTACs is sufficient to prevent autoimmunity. Single cell RNA and ATAC sequencing revealed the transcriptional and genomic symmetry between eTACs and mTECs. This potentially identifies a core program driven by Aire that may influence self-representation and tolerance across the spectrum of immune development. Outside of self-antigen tolerance, we also revealed the crucial role of ROR γ t⁺ eTACs in maintaining oral tolerance by inducing food antigen-specific regulatory T cells and suppressing follicular helper cells. In the absence of JCs, mice develop more severe delayed-typed hypersensitivity response against dietary allergens. Taken together, these findings reveal the identity of various eTAC populations and provides novel insights into the mechanisms of immune tolerance against self and innocuous dietary antigens. Understanding the biology of this population may have significance for a range of clinical applications from autoimmunity to food allergies to maternal-fetal tolerance.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
1.1 Overview	1
1.2 Immune regulation by Aire-expressing cells	2
1.3 Immune regulation by RORγt-lineage antigen presenting cells	3
CHAPTER 2: SINGLE-CELL MULTIOMICS DEFINES TOLEROGENIC EXTRATHYMIC AIRE-EXPRESSING POPULATIONS WITH UNIQUE HOMOLGY TO THYMIC EPITHELIUM	6
2.1 Abstract	6
2.2 Introductions	7
2.3 Results	8
2.3.1 eTACs consist of distinct populations of migDC-like cells	8
2.3.2 Single-cell chromatin accessibility supports the genomic identity of eTACs as myeloid populations with uniquely accessible chromatin	11
2.3.3 Focused multiomic analysis of eTAC populations identifies broad transcriptional up-regulation and antigen presentation in all eTACs with uniquely high TSA expression in JCs	12
2.3.4 Surface marker multiomics combined with functional flow cytometry and lineage tracing allows for identification and characterization of eTACs and their lineage relationships	13
2.3.5 eTACs are defined by transcriptional and regulatory homology to thymic medullary epithelium	15
2.5.6 RANK-RANKL signaling is required for Aire expression in eTACs	16
2.3.7 Pancreatic self-antigen expression in eTACs is sufficient to induce deletion of T cells escaping thymic selection and to prevent autoimmune diabetes	16

2.4 Discussion	18
2.5 Materials and Methods	37
CHAPTER 3: RORγt ETACS ARE MYELOID-DERIVED MEDIATORS OF ORAL	
TOLERANCE AND TREG INDUCTION	46
3.1 Abstract	46
3.2 Introduction	47
3.3 Results	48
3.3.1 ROR γ t-lineage APCs mediate oral tolerance to food antigen	48
3.3.2 Harmonizing the diversity of ROR γ t-lineage APC populations	49
3.3.3 ROR γ t eTACs are antigen presenting cells of myeloid lineage	50
3.3.4 ROR γ t eTACs mediate oral tolerance to food antigen	52
3.4 Discussion	54
3.5 Materials and Methods	70
CHAPTER 4: CLOSING AND FUTURE DIRECTIONS	77
REFERENCES	79

LIST OF FIGURES

Figure 2-1: Extrathymic Aire-expressing cells consist of distinct populations of migratory DC-like cells	24
Figure 2-2: Single cell RNA sequencing of Aire-expressing cells from mouse lymph nodes	26
Figure 2-3: Feature maps of genes and cell populations	27
Figure 2-4: Single-cell-chromatin accessibility supports the genomic identity of eTACs as myeloid populations with uniquely accessible chromatin	28
Figure 2-5: Single cell ATAC sequencing data	30
Figure 2-6: Focused multiomic analysis of eTAC populations identifies broad transcriptional upregulation and antigen presentation in all eTACs with uniquely high TSA expression in JCs	31
Figure 2-7: TSA in extrathymic Aire-expressing cells	32
Figure 2-8: Surface-marker multiomics combined with functional flow cytometry and lineage-tracing allows for identification and characterization of eTACs and their lineage relationships	33
Figure 2-9: eTACs are defined by transcriptional and genomic homology to thymic medullary epithelium	35
Figure 2-10: Merged thymic epithelial scRNA-seq data	37
Figure 2-11: RANK-RANK(L) signaling is required for Aire expression in eTACs	38
Figure 2-12: Pancreatic self-antigen expression in eTACs is sufficient to induce deletion of T cells escaping thymic selection, and to prevent autoimmune diabetes	39
Figure 3-1: ROR γ t-lineage APCs mediate oral tolerance to food antigen	62
Figure 3-2: Induction of oral antigen-specific T cells	63
Figure 3-3: Harmonizing the diversity of ROR γ t-lineage APC populations	64
Figure 3-4: ROR γ t-lineage APCs by scRNA-seq	66
Figure 3-5: ROR γ t eTACs are antigen presenting cells of myeloid lineage	68

Figure 3-6: RORyt eTACs are a myeloid antigen presenting cell population	70
Figure 3-7: RORyt eTACs mediate oral tolerance to food antigen	72
Figure 3-8: RORyt eTACs are essential for oral tolerance	74

Chapter 1 Introduction

1.1 Overview

Complex multicellular organisms require adaptive immune systems to maintain host defense from pathogens, but such systems must be precisely controlled to avoid aberrant inflammation and autoimmunity. Aire is crucial for thymic negative selection which deletes autoreactive T cells or converts them into regulatory T cells (Tregs) and a population of extrathymic Aire-expressing cells (eTACs) has also been described in the secondary lymphoid organs (SLOs). These cells have been defined as tolerogenic antigen presenting cells (APCs) and have also been shown to promote peripheral tolerance by inducing Tregs and functional inactivation/anergy in cognate T cells. Despite these compelling data, much of the basic biology and function of eTACs has remained elusive. By utilizing single-cell multiomics sequencing, high-dimensional single-cell flow cytometry, and various complex genetic mouse models, the aim of this thesis were as follows:

- 1) Define the identity of extrathymic Aire-expressing cells, described in chapter 2.
- 2) Determine the role of Aire⁺ ROR γ t-lineage antigen presenting cells in oral tolerance, described in chapter 3.

1.2 Immune Regulation by Aire-expressing Cells

The immune system's ability to distinguish pathogens from innocuous self and commensal antigens is essential for maintaining immune homeostasis and overall health. In vertebrates, one essential place for immune tolerance is achieved in the thymus, where developing T cells that react to self are eliminated or repurposed by specialized educator cells called medullary thymic epithelial cells (mTECs). Autoimmune Regulator (Aire) expression allows mTECs to express a wide range of tissue-specific antigens (TSAs) such as insulin¹⁻⁴. The essential role of Aire is demonstrated by mutations in the gene in humans, which cause the severe disease Autoimmune Polyglandular Syndrome Type 1^{5,6}.

While thymic self-education is crucial, complementary mechanisms also exist in the peripheral immune system to reinforce self-tolerance. Interestingly, a population of extrathymic Aire-expressing cells (eTACs) has been described in the secondary lymphoid organs (SLOs)⁷⁻¹⁰. These cells have been defined as tolerogenic antigen presenting cells (APCs) expressing high levels of MHC class II and low levels of co-stimulatory molecules CD80 and CD86. Furthermore, eTACs have been shown to promote peripheral tolerance by inducing regulatory T cells (Tregs) and functional inactivation/anergy in cognate T cells⁹. Despite these compelling functional data, the biology of these populations regarding their lineage relationship and migratory APC functions, the role of Aire within these populations, and the contribution of each population to normal immune homeostasis remains unclear.

Aire-expressing mTECs are responsible for autoreactive T cell deletion and Treg conversion in the thymus, but the function of eTACs in normal immune homeostasis is still unknown. Studies have suggested that Aire in eTACs may regulate expression of TSAs and promote a tolerogenic phenotype⁸. Loss of expression of certain TSAs like insulin in DCs leads to accelerated pancreatic autoimmunity¹¹, and overexpression of Aire in DCs can delay the onset of diabetes and experimental autoimmune encephalitis in mouse models^{12,13}. Others have shown that Aire may promote the expression of tolerogenic genes like IL-10 and Indoleamine 2,3-

dioxygenase (IDO) in DCs¹⁴ or suppress B cell responses by inhibiting the proliferation of T follicular helper cells. Additionally, we have demonstrated that deletion of eTACs as a population leads to a breakdown of maternal-fetal immune tolerance, and fetal rejection with an expansion of inflammatory T-follicular helper (Tfh) and T-helper 17 (Th17) cells and a contraction of regulatory T cells eTACs¹⁵. However, another report using an RORyt-lineage conditional deletion of Aire itself suggested that Aire expression in JCs is important for adequate host defense against the commensal *Candida albicans*¹⁶.

eTACs have been described as migratory dendritic cells (migDCs) and many reports suggest a crucial tolerogenic role of migDCs in peripheral tolerance. MigDCs are known to acquire self-antigen in the tissue and subsequently migrate to the LNs for peripheral T cell tolerance induction^{17–19}. However, the origin, identity, and stability of tolerogenic DCs is still largely up to debate in the field. Some have described tolerogenic DCs as naïve immature DCs, while others have shown that activated mature DCs can also induce T cell tolerance^{20–23}.

In conclusion, the function of eTACs as an antigen presenting cell is still largely up to debate. eTACs are similar to migDCs and can interact with cognate T cells, so they could play a unique role in the maintenance of immune homeostasis. Understanding the role of peripheral Aire and eTACs in this process will offer new insights into specialized DC subsets and mechanisms underlying peripheral immune tolerance.

1.3 Immune Regulation by RORyt-lineage Antigen Presenting Cells

Peripheral Tregs (pTregs) expressing RORyt are known to populate the gut where they induce tolerance to dietary antigens and commensal microbes^{24–26}. Originally, CD11c+ conventional DCs are thought to broadly induce pTregs in the gut²⁷, but several recent papers have attributed this role to distinct RORyt+ APCs instead. Different RORyt+ antigen presenting cells have been described and includes MHCII+ ILC3s and RORyt+ Aire+ eTACs that are similar to DCs. Despite all reports revealing RORyt+ APCs' indispensable role in shaping T cell responses to maintain commensal tolerance in the gut, they differ on their conclusion regarding

which specific RORyt⁺ APC is responsible^{28–32}. Since a subset of eTAC also express RORyt and MHCII, which we have termed Janus cells (JCs), further studies are needed to understand whether this Aire⁺ RORyt⁺ APC population is responsible for inducing immune tolerance to self and other innocuous foreign antigens.

Initial reports by Sonnenberg and colleagues have used RORyt-Cre x MHCII-flox conditional deletion to suggest that MHCII⁺ ILC3s are important for maintaining commensal tolerance through interactions with commensal-specific CD4⁺ T cells^{28,29}. In the absence of MHCII on LTI-like ILC3s, Th17 and Tfh in the gut are expanded, which suggest that MHCII⁺ ILC3s help suppress T cell responses. However, this was before the identification of Aire⁺ RORyt⁺ Janus cells as well as other RORyt⁺ DCs^{30,32}, so more detailed studies are needed to understand which population are affected in the RORyt-Cre x MHCII-flox mouse.

Three recent studies also utilized the RORyt-Cre x MHCII-flox mouse to understand key APC populations that are indispensable for the generation of microbiota-specific pTregs that express both FOXP3 and RORyt^{33–35}. Although there are differing conclusions on which specific RORyt⁺ population is responsible for pTreg generation, all three reports revealed that microbiota-specific RORyt⁺ Treg cells are significantly impaired in mice lacking MHC-II on RORyt⁺ APCs. Furthermore, the reports demonstrated the essential contribution of TGFβ-processing integrins (Itgb8 and Itgav) on RORyt⁺ APCs and a lack of a requirement for MHC-II on cDCs in this process. Kedmi *et al.* investigated pTreg induction against commensal *Helicobacter* species with the use of TCR-transgenic mice³³. CD11c-expressing cells were shown to be redundant in the generation of pTregs against *Helicobacter*. Instead, RORyt⁺ APCs are essential for this process as RORyt-Cre x MHCII-flox mice develop no pTregs against *Helicobacter*. They concluded that either MHCII⁺ LTI-like ILC3 or RORyt⁺ eTACs are responsible for commensal-specific pTreg generation. Akagbosu *et al.*³⁴ described a novel population of RORyt⁺ APCs and termed them Thetis cells which are CD11b⁺ ITGB8⁺ Rorc⁺ DC-like APCs that arise in early life with the development of pTregs in the mesenteric LN. With the conditional knockout of MHCII on RORyt⁺

lineage cells, mice exhibit with the loss of pTregs in the MLN and large intestine lamina propria with increased Th17 cells which results in colitis in adulthood. They concluded that antigen presentation by Thetis cells, but not ILC3s or cDCs, is crucial for this pTreg generation and the prevention of colitis. Interestingly, Thetis cells are transcriptionally similar to ROR γ t⁺ eTACs³⁶ and further studies are required to decipher the distinct role of Thetis cells and ROR γ t⁺ eTACs in this process. Lastly, Lyu *et al.* concluded that LTi-like ILC3s are the major ROR γ t⁺ APC subset that promotes the generation of microbiota-specific pTregs³⁵. With the use of Aire-Cre mouse, the study concluded that MHCII expression on Aire⁺ cells are dispensable for pTreg induction, while ILC3s are essential for the process. However, it was unclear whether the Aire-cre can efficiently target Janus cells. Furthermore, this study also found altered ROR γ t⁺ Tregs and MHCII⁺ ILC3s in the inflamed intestine of patients with inflammatory bowel disease, suggesting these populations might be important for the maintenance of normal intestinal homeostasis in humans.

Overall, additional tools and research are required to further decipher the contribution of each ROR γ t⁺ APC subset (LTi-like ILC3s, ROR γ t⁺ eTACs/Janus cells, and *Rorc*⁺ DC-like Thetis cells) to the induction of commensal-specific pTregs. While it is possible that all subsets could contribute in a spatial-temporal manner, this dissertation will focus on the role of ROR γ t⁺ eTACs/Janus cells in the induction of pTregs and the maintenance of immune homeostasis.

Chapter 2 Single-cell multiomics defines tolerogenic extrathymic Aire-expressing populations with unique homology to thymic epithelium

2.1 Abstract

The autoimmune regulator (Aire), a well-defined transcriptional regulator in the thymus, is also found in extrathymic Aire-expressing cells (eTACs) in the secondary lymphoid organs. eTACs are hematopoietic antigen-presenting cells and inducers of immune tolerance, but their precise identity has remained unclear. Here, we use single-cell multiomics, transgenic murine models, and functional approaches to define eTACs at the transcriptional, genomic, and proteomic level. We find that eTACs consist of two similar cell types: CCR7⁺ Aire-expressing migratory dendritic cells (AmDCs) and an Aire^{hi} population coexpressing Aire and retinoic acid receptor–related orphan receptor γ t (ROR γ t) that we term Janus cells (JCs). Both JCs and AmDCs have the highest transcriptional and genomic homology to CCR7⁺ migratory dendritic cells. eTACs, particularly JCs, have highly accessible chromatin and broad gene expression, including a range of tissue-specific antigens, as well as remarkable homology to medullary thymic epithelium and RANK-dependent Aire expression. Transgenic self-antigen expression by eTACs is sufficient to induce negative selection and prevent autoimmune diabetes. This transcriptional, genomic, and functional symmetry between eTACs (both JCs and AmDCs) and medullary thymic epithelium—the other principal Aire-expressing population and a key regulator of central tolerance—identifies a core program that may influence self-representation and tolerance across the spectrum of immune development.

2.2 Introduction

The autoimmune regulator (Aire) gene plays an essential role in the maintenance of central tolerance, and its function has been extensively defined in medullary thymic epithelial cells (mTECs), where, among other roles, it regulates the expression of a diverse range of otherwise tissue-specific self-antigens^{37,38}. Aire expression has also been observed in the secondary and tertiary lymphoid organs in both mice and humans but not in primary tissues^{8,14,39}, and self-antigen expression in such extrathymic Aire-expressing cells (eTACs) is sufficient to cause deletion or inactivation of cognate T cells^{8,9}. Recently, eTACs have also been demonstrated to be essential for the maintenance of normal maternal-fetal immune tolerance¹⁵. Thus, the precise identity and biology of these extra- thymic Aire-expressing populations is of substantial biological and clinical interest.

Whereas Aire expression has been described in various dendritic cell (DC) populations^{7–9,14,40–42}, the precise identity of eTACs has remained elusive. Recent reports have suggested that the principal Aire-expressing population in the secondary lymphoid organs may be a subset of group 3 innate lymphoid cells (ILC3s), based largely on the co-expression of Aire and retinoic acid receptor–related orphan receptor γ t (ROR γ t), encoded by the *Rorc* gene and known for its essential role in T helper cell type 17 (TH17) and ILC3 differentiation. A growing body of evidence has also shown that ROR γ t-expressing populations characterized as ILC3s can play important roles as antigen-presenting cells (APCs) to promote immune tolerance in a range of contexts, most notably in commensal tolerance in the gut^{28,29,31}. Furthermore, single-cell transcriptional mapping of murine DC lineages has also recently identified populations of ROR γ t-expressing DCs³⁰. Thus, the identification of an Aire- and ROR γ t coexpressing migratory DC (migDC)–like population may have implications for some of these divergent findings. Recent evidence suggests that migDCs may have important regulatory functions in the maintenance of immune tolerance^{18,43,44}. Initially characterized by their participation in the generation of memory immune responses to acquired antigens^{45,46}, migDCs have more recently been suggested to participate in

the regulation and repression of immune responses. Transcriptional characterization of migDCs by the Immunologic Genome Project (ImmGen) consortium suggested enrichment for an immunomodulatory and tolerogenic transcriptional program⁴⁷. Deletion of migDC populations, or interference with their migration, worsened autoimmunity in a range of mouse models^{18,43}. More recently, populations of tumor-associated migDCs (termed mregDCs) were described, which had both tolerogenic phenotype and function, and regulated immune responses to neoplasia⁴⁴. However, the signals and transcriptional regulatory circuits driving such tolerogenic phenotypes in migDCs remain obscure.

Here, we use high-throughput single-cell multiomics and a range of functional approaches to define eTACs as migDC-like populations and identify among them a new mixed-phenotype Aire- and ROR γ t expressing population, which we term Janus cells (JCs). We show that these populations have uniquely high degrees of broad chromatin accessibility and gene expression, including a range of tissue-specific antigens (TSAs) and immunomodulatory transcripts. Self-antigen expression in eTACs also leads to deletional tolerance among CD8⁺ T cells that escape thymic selection. Furthermore, we find that eTACs have remarkable transcriptional and genomic homology to medullary thymic epithelium, despite being hematopoietic in origin. This convergent transcriptional circuitry between thymic and peripheral Aire-expressing populations, despite their distinct stromal and hematopoietic origins, may provide unique insights into the fundamental mechanisms maintaining self-tolerance throughout immune development.

2.3 Results

2.3.1 eTACs consist of distinct populations of migDC-like cells

To define the identity and basic biology of extrathymic Aire- expressing populations in an unbiased fashion, we sought to characterize these cells using high-dimensional single-cell multiomics, including single-cell RNA sequencing (scRNA-seq) and ATAC (assay for transposase-accessible chromatin) with selected antigen profiling by sequencing (ASAP-seq). Because eTACs are present predominantly in the secondary lymphoid organs⁸, we enriched for

rare populations in the lymph nodes (LNs) by magnetic column depletion of T and B cells from the digested, pooled LNs of wild-type (WT) and Aire-driven IGRP-GFP (islet-specific glucose-6-phosphatase– related protein–green fluorescent protein) (Adig) mice, a previously characterized Aire reporter mouse strain⁸. Next, enriched cells were either directly subjected to parallel multiomic sequencing (scRNA-seq and ASAP-seq) from WT mice or flow-sorted for all live, GFP+ cells from Adig mice (Fig. 2-1A and Fig. 2-2A). From the scRNA-seq data (Fig. 2-2B and C), the unsorted WT sample (n = 6973 cells) allowed us to generate a map of rare and common secondary lymphoid populations, whereas the GFP-sorted sample (n = 2532 cells) from Adig mice allowed us to deeply profile eTAC populations and further map them into the immune landscape to define their identity. Using our combined data resource, we identified discrete populations of eTACs (GFP+ cells from the Adig LNs), which aligned closely with GFP and Aire transcript expression (Fig. 2-1B and D), confirming both sort purity and the fidelity of the GFP reporter. In the merged data, we were able to identify a broad range of rare lymphoid, myeloid, and stromal cell types and found that eTACs represented two distinct populations, both of which shared high transcriptional similarity to migDCs (Fig. 2-1C to G, and Fig. 2-2A and B). Most Aire expression was seen in migDC1/2s, and we referred to these populations as Aire-expressing migDCs (AmDCs), although these AmDCs did not cluster separately as a distinct group from Aire-negative migDCs (Fig. 2-1C). In contrast, the eTAC subset with the highest Aire expression was a discrete cluster, clearly distinct in RNA space after dimensional reduction, which expressed both high levels of Aire and *Rorc* (Fig. 2-1C, D, and G), a gene otherwise associated with innate and adaptive lymphocyte populations such as TH17 and ILC3s⁴⁸. These same eTAC populations were also all present in reference (non–Aire-enriched) samples alone, although at lower frequency (Fig. 2-1B and Fig. 2-2D). Despite its *Rorc* expression, this Aire-hi population otherwise appeared more closely related to the migDC lineage in t-distributed stochastic neighbor embedding (t-SNE) space and expressed comparable levels of *Zbtb46*, *Ccr7*, and major histocompatibility complex (MHC) class II (Fig. 2-1D). Because of the distinct phenotype and dual expression of Aire-

hi/Rorc+, we refer to the cell population as Janus cells, in reference to the two-faced classical deity.

To more precisely assign identity to extrathymic Aire-expressing populations in an unbiased fashion based on their entire transcriptome rather than individual marker genes, we next performed unsupervised two-dimensional hierarchical clustering and cosine similarity scoring of each pseudo bulk population against the entire ImmGen microarray database of 178 annotated reference immune populations (23) in RNA space (Fig. 2-1E). Our informatics approach verified clustering of eTAC populations (i.e., JCs with AmDCs) and high transcriptional similarity to reference ImmGen populations collectively identified as migDCs⁴⁹. Reciprocally, we then used select ImmGen reference populations to assign cosine similarity scores on a per-cell basis to each cell profiled by scRNA-seq, which again demonstrated that both JCs and AmDCs had high transcriptional similarity to migDCs (Fig. 2-1F and Fig. 2-3B). These unbiased, reciprocal approaches to cell identity assignment indicate that all eTAC populations appear to be most transcriptionally similar to migDCs.

Noting that migDCs have been characterized by a common canonical transcriptional signature⁴⁷, we verified that this signature was highly expressed in all Aire-expressing clusters (JCs and migDC1/2) (Fig. 2-1G). Despite a recent report that this Aire/Rorc-coexpressing population was ILC3-like³¹, we found that JCs clustered broadly with myeloid and not lymphoid lineages. JCs share less homology with defined populations of either NKp46+ or LTI ILC3s and do not express detectable levels of most canonical ILC3-family transcripts (e.g., Il17a/f, Il22, and Il23r)⁵⁰. However, some genes, including Ccr6 and Tmem176a/b, were shared between the populations. eTACs in general, and JCs in particular, shared some transcripts with epithelial/stromal lineages (Fig. 2-1G).

2.3.2 Single-cell chromatin accessibility supports the genomic identity of eTACs as myeloid populations with uniquely accessible chromatin

To further characterize eTACs at both the genomic and proteomic level, we subjected the same populations of cells purified from Aire reporter and WT mice described above (Fig. 2-1A) to combined single-cell ATAC (scATAC) and surface protein multiomics (ASAP-seq)⁵¹. In strong concordance with the scRNA-seq data, we again found that Aire accessibility was largely confined to migDCs and JCs (Fig. 2-4, A to C). To independently validate the cellular identity of these populations based on their chromatin profiles, we next performed unsupervised two-dimensional hierarchical clustering and cosine similarity scoring of each pseudo bulk population against the entire ImmGen ATAC-seq database that consisted of 89 populations⁵². Although the ImmGen ATAC database does not include migDC populations, JCs and migDCs again both coclustered with each other and showed the highest cosine similarity to DCs (Fig. 2-4D, red and blue arrows). Conversely, these eTAC populations showed low similarity to ILCs or other lymphoid lineages and were also similar to mTECs (Fig. 2-4D, green arrow). These identities were reciprocally validated using per-cell cosine similarity projections based on accessible chromatin (Fig. 2-5A).

Mapping the accessibility of the cis-regulatory region defining the Aire locus in these populations showed distinct peaks in JCs and migDCs in the conserved noncoding sequence 1 (CNS1) region, known to be essential for Aire expression (Fig. 2-4E)^{53,54}. Our analysis of this locus also revealed several additional peaks unique to Aire-expressing populations that had not been previously described (Fig. 2-4E and Fig. 2-5B). Global differential accessibility analyses of eTACs against all other Aire-negative populations showed significantly increased chromatin accessibility in Aire-expressing cells relative to all other LN populations, including at the Aire locus (Fig. 2-4, F and G). To support that this was not due to batch effects, we also evaluated the same chromatin accessibility between eTACs and other LN populations in the WT dataset alone and again observed significantly increased accessibility among eTACs (Fig. 2-5C). To confirm that this result was not due to skewing from a particular low-accessibility population, we also examined the per-

cell chromatin accessibility of each individual population and found the same trend—JCs had the highest chromatin accessibility of any cell type in the LN, followed closely by migDC1/2s (Fig. 2-4H). Overall, our analyses support a marked increase of chromatin accessibility in eTACs, a notable observation, given Aire's known role in chromatin regulation in the thymus^{55–57} and in governing the expression of a broad range of TSAs.

2.3.3 Focused multiomic analysis of eTAC populations identifies broad transcriptional up-regulation and antigen presentation in all eTACs with uniquely high TSA expression in JCs

To better define the biology and relationships of eTAC subsets, we next focused specifically on GFP-sorted populations. We again confirmed that eTACs consist principally of AmDC1/2s and JCs, as well as small numbers of cells from B cell and ILC lineages (Fig. 2-6A and Fig. 2-7A). Coincident with their increased chromatin accessibility, eTACs also had high levels of global gene expression compared with all other populations (Fig. 2-7B), with JCs having the highest level of global gene expression among eTAC clusters (Fig. 2-6B and Fig. 2-7B). Differentially up-regulated genes in JCs included a wide range of transcripts including both Aire and Rorc, as well as a distinct set of unique JC-specific genes (Fig. 2-6B). Among JC-specific genes, we identified a large number of transcripts previously classified as TSAs (Fig. 2-6C)⁵⁸. When we surveyed the global TSA expression, we found that JCs expressed the highest numbers of TSA genes, both when compared against AmDCs and more globally against all LN populations, presumably because of their overall increased levels of gene expression (Fig. 2-6D and Fig. 2-7, B and C) and greater chromatin accessibility. These TSAs were particularly enriched for neuronal-associated and germ cell/placental antigens (Fig. 2-7D), suggesting both potential disease models in which to interrogate the function of these populations as well as potential links to the established findings that loss of eTACs leads to a breakdown in maternal-fetal tolerance (7). JCs and AmDCs expressed a largely overlapping set of TSA genes, with JCs expressing a larger number of unique TSA transcripts (Fig. 2-7E).

Both JCs and AmDCs also expressed a diverse array of genes associated with antigen processing and presentation in both the class I and class II MHC pathways (Fig. 2-6E). In addition, eTACs also expressed a range of genes associated with immune tolerance and tolerogenic DC maturation^{48,59}, including *Socs2*, *Fas*, *Cd200*, *Ido1*, *Smad4*, and *Irf2bp2* (Fig. 2-6F). These findings suggest that these populations may be involved in both the expression and acquisition of self-antigens and support the existing evidence that eTACs may promote antigen-specific immune tolerance among interacting lymphocyte populations.

2.3.4 Surface marker multiomics combined with functional flow cytometry and lineage tracing allows for identification and characterization of eTACs and their lineage relationships

Although JC and AmDC subpopulations of eTACs appear transcriptionally similar, the lineage relationship between them was unclear. Because single-cell ASAP-seq allows for identification of a broad range of surface markers associated with each population, we next sought to define the surface proteins unique to AmDCs and JCs, which would allow us to distinguish and trace these populations. Differential analysis defined several surface proteins relatively more prevalent in either AmDCs (*CD11c*, *CD45*, and *MHCII*) or JCs (*CD200*; Fig. 2-8, A and B).

Furthermore, transcriptional data from the scRNA-seq analysis suggested that JCs were highly enriched for expression of a range of Wnt signaling pathway and progenitor-associated factors previously correlated with stem-like cell fates (Fig. 2-8C)⁶⁰⁻⁶⁴. Wnt signaling through E-cadherin and beta-catenin also has a role in maintaining tolerogenic function in dendritic and myeloid cells^{65,66}. To determine whether JCs' progenitor-like phenotype reflected their status as a precursor to AmDCs, we used both bioinformatic and functional approaches. RNA velocity⁶⁷ analysis showed no evidence of developmental vectors projecting from JCs to AmDCs (Fig. 2-8D). To verify this inference, we sought to characterize these relationships between respective eTAC populations by genetic lineage tracing and flow cytometry. First, using the Adig mouse, we verified that most eTACs consist of migDC1/2s in roughly equal proportions of the two populations

(Fig. 2-8E)—consistent with proportions found in our multiome data. Unexpectedly, assessed as a percentage of total migDCs, AmDCs represent more than 30% of the total migDC pool at baseline (Fig. 2-8F).

Next, to better identify JCs by flow cytometry and understand their lineage relationship to AmDCs, we used the previously described ROR γ t-Cre mouse⁶⁸. ROR γ t-Cre mice were crossed with Adig Aire reporter and Rosa26-LSL-TmR mice to generate ROR γ t expression and lineage tracer and Aire reporter (REALTAR) mice. The REALTAR mouse allowed us to identify cells actively transcribing Aire by GFP expression and cells expressing or having ever expressed ROR γ t by TmRed expression. Using this model, we could identify two populations: a larger, distinct population of GFP+/TmR⁻ cells and a smaller population of GFP+/TmR⁺ cells (Fig. 2-8G). Surface marker characteristics of these populations matched what we observed by scRNA-seq and ASAP-seq (Figs. 2-6B and 2-8, A and B)—that JCs (GFP+/TmR⁺) expressed markedly lower levels of CD11c and CCR7 than AmDCs (GFP+/TmR⁻; Fig. 2-8G). This was validated by staining for a number of other differential markers between AmDCs and JCs including MHCII, CD45, CCR6, and CD200 (Figs. 2-1G and 2-8, A and G), which are consistent with previously published results³¹. Thus, we confirmed that JCs have a unique CCR7 /CD11c surface marker phenotype distinct from AmDCs, despite expression of both genes at the transcript level in JCs. This lineage tracer system also validates that JCs express the ROR γ t isoform of Rorc. Furthermore, these data suggest that most AmDCs are not lineage-traced by ROR γ t and thus do not appear to arise from an ROR γ t-expressing JC precursor. Last, among MHCII^{hi} ROR γ t lineage-traced (TmRed⁺) cells, more than 50% of this population is Aire-expressing non-ILCs (GFP⁺ and CD90⁻; Fig. 2-8H). This suggests that perhaps JCs should be considered in the interpretation of previously described tolerogenic phenotypes attributed to MHCII-expressing ILC3s using ROR γ t-Cre systems^{28,29}.

2.3.5 eTACs are defined by transcriptional and regulatory homology to thymic medullary epithelium

Although the global transcriptional and functional profile of eTACs suggested a high similarity with migDCs, we noted that our cosine similarity scoring against reference ImmGen populations demonstrated significant transcriptional homology to another population—mTECs (Fig. 2-9, A and B). This was particularly pronounced among JCs. Indeed, the top differentially expressed genes defining JCs were highly enriched only in mTECs when assessed among all other annotated populations in the ImmGen RNA-seq database (Fig. 2-9C). To define the expression of these genes more precisely relative to Aire expression in the thymus, we aggregated an atlas of previously published scRNA-seq data on thymic epithelial cells and performed integrative analyses using scVI⁶⁹ to establish a dimensionally reduced map of thymic epithelial populations across the full spectrum of their development. The integrated data were annotated according to the population identities defined in the source data (Fig. 2-10)^{70–73}. In addition to being expressed in mTECs, many of these JC-specific genes appeared tightly associated with Aire expression during mTEC development (Fig. 2-9D). Most of these genes were not, themselves, Aire-regulated in the thymus [three of the top 50 differentially expressed (DE) genes, 5.5% overall]⁷⁴ but rather may represent Aire-associated transcriptional pathways with relevance for the biology and function of Aire in these populations. Differential transcription factor binding site accessibility analysis of the scATAC data demonstrated that the chromatin landscape of JCs was most highly enriched for transcription factor binding sites known for their roles driving Aire expression in mTECs, including RelB and nuclear factor kB (NF-kB), as well as less well-characterized human immunodeficiency virus type I enhancer binding protein (HIVEP) genes (Fig. 2-9E). Although some of this overlap might be explained by binding site sequence degeneracy, a number of these gene transcripts including the noncanonical NF-kB pathway genes Relb and Nfkb2, and Hivep1 and Hivep3, appeared to be highly expressed in JCs and/or migDCs (Fig. 2-9F). Surveying the ImmGen ATAC database showed that these same transcription factor binding

sites are also highly enriched in the open chromatin of medullary thymic epithelium (Fig. 2-9G, red arrow). Thus, our analyses indicate that JCs share a canonical transcriptional and accessible chromatin program with mTECs.

2.3.6 RANK-RANKL signaling is required for Aire expression in eTACs

Given the strong genomic and transcriptional evidence for RANK/NF- κ B signaling in eTACs and the well-established role of RANK- RANKL signaling in driving Aire expression in the thymus^{75,76}, we next asked whether these signals are also required for Aire expression in eTACs. First, we found that RANK (Tnfrsf11a) expression was highly up-regulated in eTACs. OPG (Tnfrsf11b), a soluble RANKL antagonist that often serves in a negative-feedback regulatory mechanism in cells receiving high levels of RANK signaling^{77,78}, was quite specifically expressed in JCs among peripheral immune populations (Fig. 2-11A). As previously demonstrated^{72,79,80}, inhibition of RANK signaling by antibody blockade of RANKL led to a significant loss of Aire expression among mTECs in the thymus (Fig. 2-11, B and D). The same was true in the periphery, where eTACs (all GFP+ cells) decreased in relative and absolute abundance after RANKL blockade (Fig. 2-11, C and D). JCs (GFP+ CCR7- CD11c-) almost entirely disappeared (or lost Aire expression), suggesting that, as in the thymus, Aire expression in JCs depends on RANK/NF- κ B signals. This result again demonstrates the symmetry between Aire expression in the thymus and periphery, provides valuable insight into signals required to induce Aire expression in hematopoietic populations, and suggests that studies using RANK blockade, for example, in the setting of tumor immunity^{80,81}, should also consider the impact on extrathymic populations.

2.3.7 Pancreatic self-antigen expression in eTACs is sufficient to induce deletion of T cells escaping thymic selection and to prevent autoimmune diabetes

Given these transcriptional and genomic similarities between mTECs and eTACs, despite their different origins, we asked whether self- antigen presentation by eTACs is sufficient to mediate tolerance for T cells that escape thymic negative selection. To assess this, we used a feature of the Adig mouse, which also expresses a pancreatic self-antigen, IGRP, under the

control of the Aire promoter⁸. To generate mice in which only eTACs and not mTECs expressed IGRP, we thymectomized WT and Adig nonobese diabetic (NOD) mice and simultaneously transplanted WT thymi under the kidney capsule and then sublethally irradiated and reconstituted them with bone marrow (BM) from the 8.3 mouse, a CD8⁺ T cell receptor (TCR)–transgenic specific for IGRP (Fig. 2-12A)⁸². Despite concern that peripheral antigen might traffic back to the transplanted thymus and mediate thymic negative selection, we saw no evidence of this, because fluorescence-activated cell sorting (FACS) analysis and tetramer staining of lymphocytes from these mice demonstrated no thymic negative selection and no loss of tetramer avidity (Fig. 2-12, B to D). By contrast, in the secondary lymphoid organs, we observed profound deletion of IGRP-specific CD8⁺ T cells and significantly decreased tetramer avidity of surviving CD8⁺ T cells (Fig. 2-12, B to D). Despite the large numbers of IGRP-specific T cells leaving the thymus in this permissive TCR transgenic system, eTAC-mediated deletion in the secondary lymphoid organs was sufficient to entirely prevent these mice from developing diabetes (Fig. 2-12E).

2.4 Discussion

A growing body of evidence has demonstrated that migDCs in the steady state may play a significant role in immune tolerance, with diverse examples in immune homeostasis, autoimmunity, and tumor immunity^{18,43,44}. Furthermore, although traditional paradigms suggested an inverse correlation between DC maturity and immunogenicity, some mature DCs, particularly migDCs, acquire a discrete tolerogenic profile⁵⁹. However, the transcriptional signals driving such tolerogenic maturation have remained unclear. In parallel, work defining the identity and biology of eTACs has suggested a potential migDC-like phenotype^{9,40}, but their precise identity and biology has remained unclear.

Here, we defined eTACs using high-dimensional multiomics as consisting of two distinct migDC-like populations, AmDCs and JCs, and show that self-antigen expression in eTACs is sufficient to cause deletion of autoreactive T cells escaping thymic selection and prevent autoimmune diabetes. We find a distinct transcriptional homology between eTACs and mTECs, and comparative analysis of the two populations identified a core transcriptional signature shared between the two populations. This suggests several promising future directions to define both the signals driving Aire expression and more broadly the transcriptional circuitry underlying self-education in the adaptive immune system. We show that eTACs share a common genomic accessibility signature with mTECs enriched in RANK/ NF- κ B signaling pathways, and as in the thymus, RANK/RANKL signaling is required for Aire expression in eTACs. Beyond these known signals, several other pathways appear common to both populations, including Wnt signaling and the *Hivep* genes. These parallel analyses also suggest novel candidate regulatory regions around the Aire locus, which may provide further insight into the regulation of Aire expression in both the thymus and periphery. Last, the unique genomic accessibility and high transcriptional activity of eTACs generally, and JCs specifically, raise numerous fascinating questions about their biology and homeostasis, and the role of TSA expression within these populations.

Although previous work had suggested that Aire/RORyt cells might be an ILC3-like population³¹, here we used unbiased whole-transcriptome and whole-chromatin single-cell profiling methods to strongly support the myeloid and DC-like phenotype of JCs. Their similarities to ILCs appear to be largely a product of RORyt-dependent gene expression, like *Tmem176a/b*⁵⁰. These Aire-hi/RORyt+ cells have previously been demonstrated to be capable of antigen processing and presentation³¹, and we show that these pathways are up-regulated in both JCs and AmDCs. Combined with their highly accessible chromatin and their broad transcriptional program including a range of TSAs, this suggests that the population may play a role in immune self-education. The representation of placental antigens among the TSAs expressed by JCs suggests a potential link to the recently reported requirement for eTACs in the maintenance of maternal-fetal tolerance in both syngeneic and allogeneic pregnancy¹⁵. Because of their lack of surface protein expression of either CD11c or CCR7, this population may have been overlooked in previous analyses of DC subsets. Furthermore, a number of previous studies have described tolerogenic phenotypes and attributed peripheral regulatory T cell induction to MHCII- expressing ILC3s using RORyt-Cre/MHCII-flox systems^{28,29}. The identification of RORyt-expressing JCs suggests that these cells should be considered as contributors to these phenotypes and deserves further study.

Numerous questions remain, including the contributions of peripheral Aire to normal immune homeostasis, the precise function of Aire within eTACs, and the mechanisms of eTAC-mediated tolerance induction. In addition, although we have demonstrated here that JCs are most likely not precursors to most AmDCs, JCs do share many phenotypic characteristics with migDCs and thus represent an intriguing discrete myeloid population whose developmental origin will be essential to further define. Similarly, the signals driving Aire expression in these populations, and in the thymus through shared programs, will require greater investigation to understand the physiologic significance of these regulatory networks.

Last, although eTACs have recently been demonstrated to be essential for normal immune homeostasis in maternal-fetal tolerance during pregnancy¹⁵, the precise mechanisms of that tolerance remain unclear. Although the expression of a range of pregnancy-associated and placental antigens in JCs offers a tantalizing clue, it remains to be demonstrated whether the observed phenotype—immune-mediated intrauterine growth restriction after ablation of eTACs—is a direct result of this antigen reservoir loss, of antigen acquisition by these migratory populations, or of some other mechanism.

The remarkable parallelism between thymic and peripheral Aire- expressing cells does, however, suggest a uniquely conserved central and peripheral core circuitry for the maintenance of immune tolerance, despite the divergent origin of these two cell types. Defining the biology of this core Aire-associated program may shed significant insight into basic Aire biology and its role in immune tolerance. Overall, work delineating the conserved transcriptional circuitry between these two tolerogenic populations could have broad relevance both for our understanding of normal immune homeostasis mechanisms and for diverse applications in autoimmunity, maternal-fetal tolerance, tumor immunity, and transplantation.

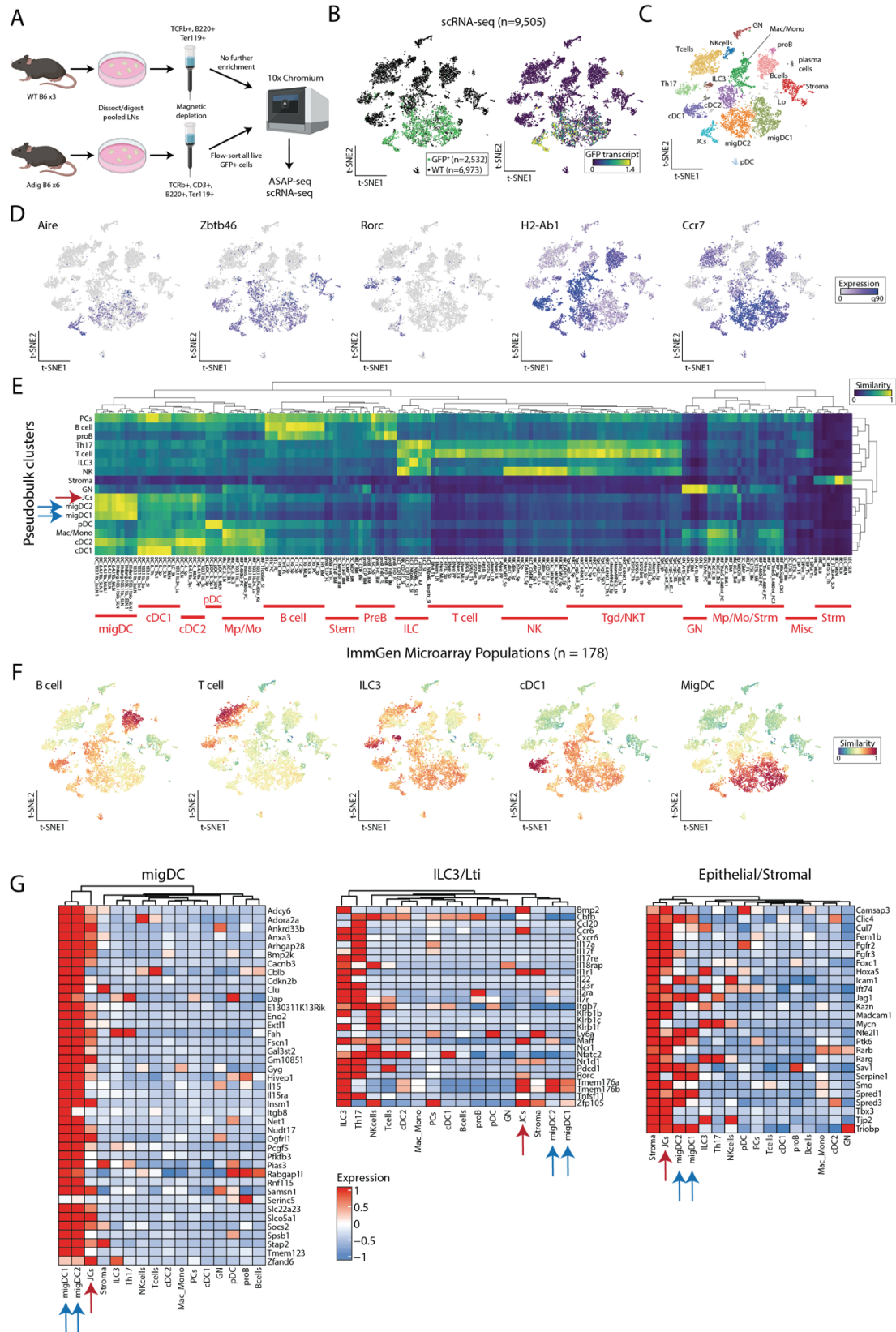


Figure 2-1

Figure 2-1: eTACs consist of distinct populations of migDC-like cells.

(A) Schematic of cell isolation and enrichment for scRNA-seq and ASAP-seq. (B) Reduced dimensionality representation of scRNA-seq data, indicating the GFP population and GFP transcript. (C) Annotated cell clusters of populations in scRNA-seq data. NK, natural killer. (D) Gene expression of selected transcripts used to define populations, including *Aire* and *Rorc*. (E) Hierarchical clustering of pseudobulk scRNA-seq clusters by all ImmGen microarray data using a scaled cosine similarity metric. Rows are min-max-normalized. ImmGen populations cluster identities annotated in red text as indicated. Aire-expressing populations are indicated with blue (migDC1/2) and red (JC) arrows. (F) Single-cell cosine similarity scores for scRNA-seq data using indicated ImmGen populations. (G) Gene expression heatmaps for selected gene sets, z score-normalized per row. Aire-expressing populations are indicated with blue (migDC1/2) and red (JC) arrows.

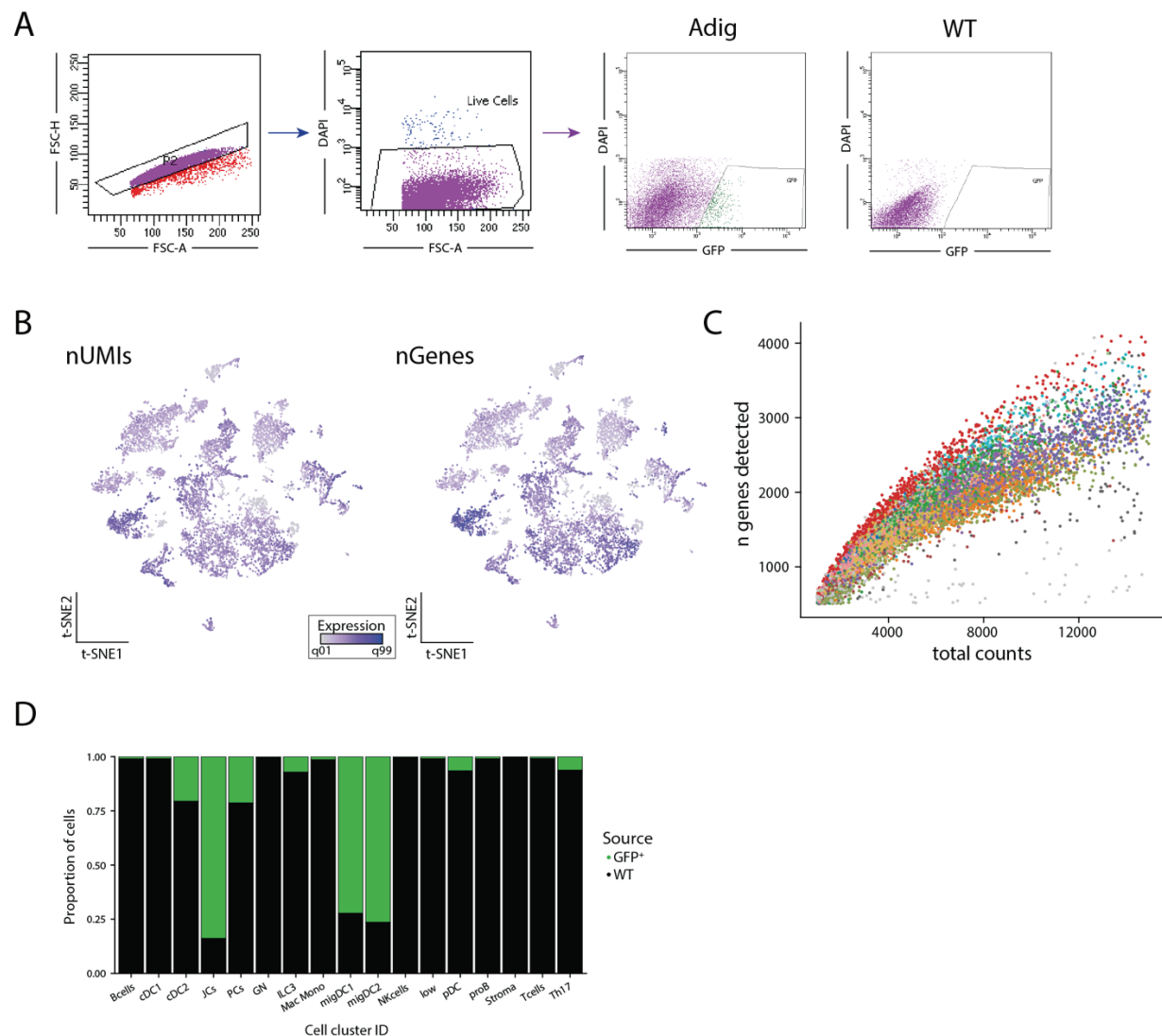


Figure 2-2: Single cell RNA sequencing of Aire-expressing cells from mouse lymph nodes. (A) Sorting strategy for single, live, GFP+ cells from Adig lymph nodes, with WT as GFP- negative control. (B) Summary of scRNA-seq quality control depicting number of unique molecular identifiers (UMIs; left) and genes (right) detected. (C) Scatter plot depicting genes and counts detected per cell. Colors represent cell clusters. (D) Proportion of cells per cluster from the GFP+ and wildtype (WT) mice lymph nodes.

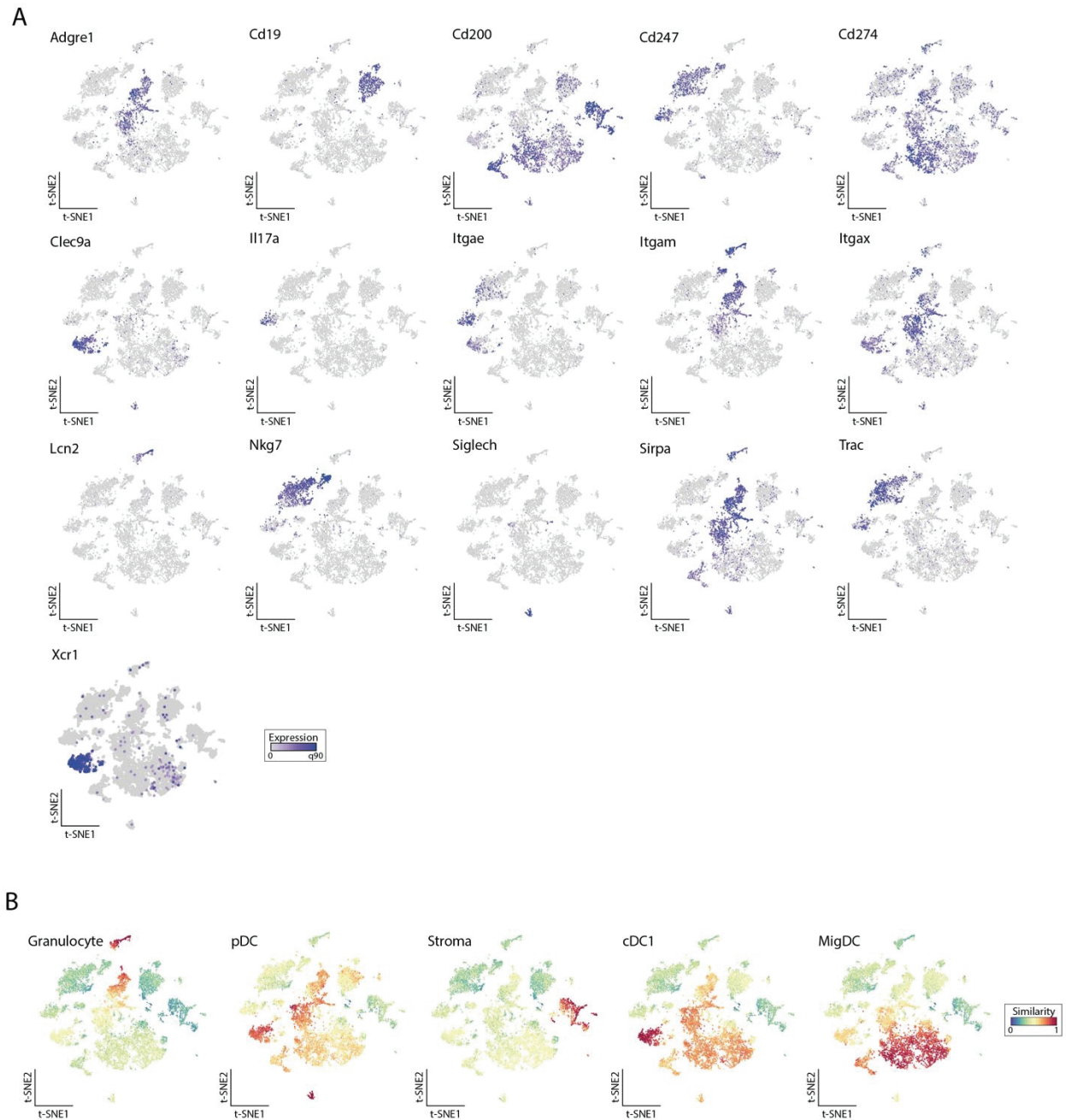


Figure 2-3: Feature maps of genes and cell populations.

(A) Feature maps of selected individual genes expressed in merged WT/GFP+ scRNA-seq tSNE space used to assign cluster identity. (B) ImmGen similarity scores supporting annotation of cell clusters from the scRNA-seq data with reference populations as indicated. cDC1 and MigDC are replicates from other ImmGen reference populations.

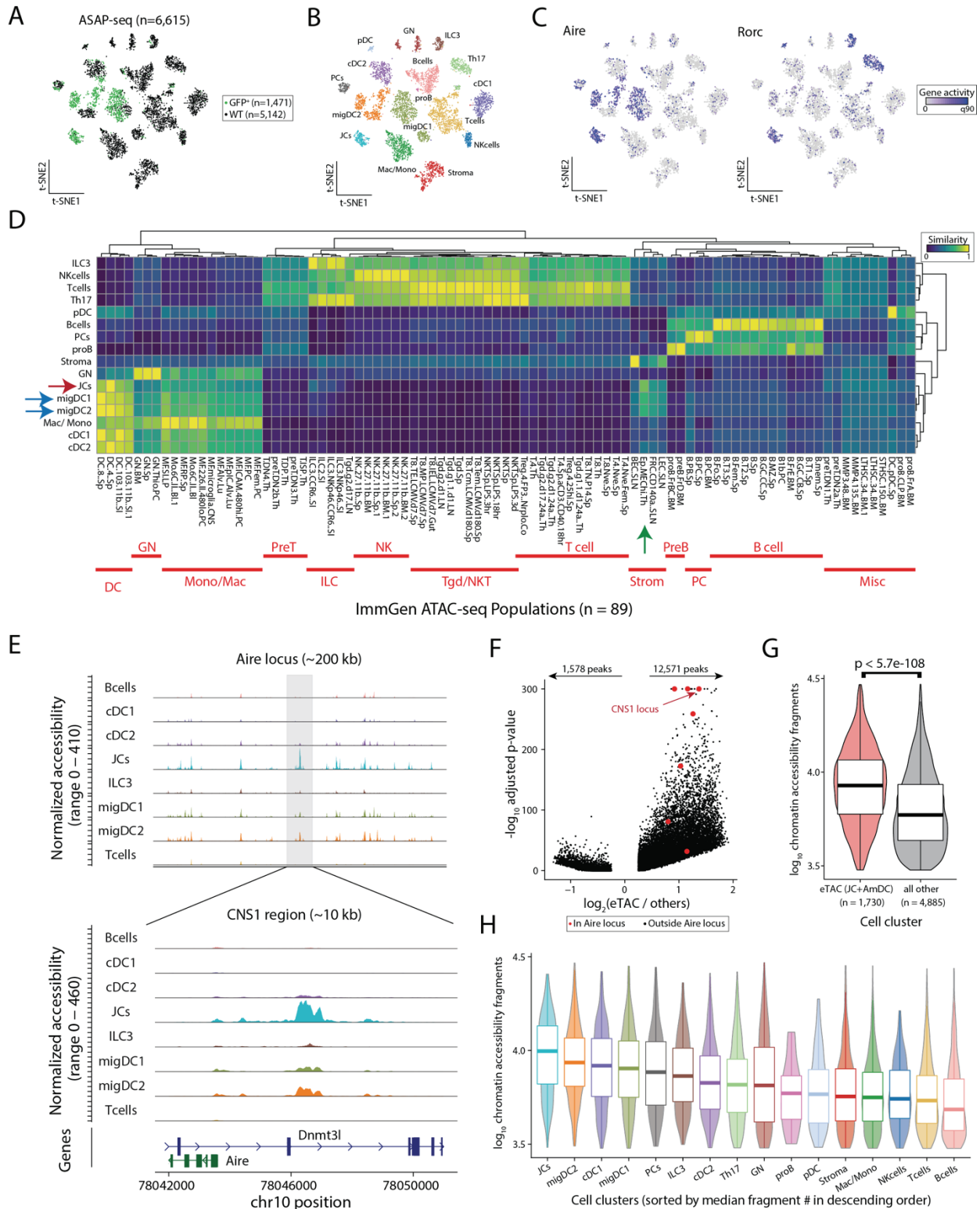


Figure 2-4

Figure 2-4: Single-cell-chromatin accessibility supports the genomic identity of eTACs as myeloid populations with uniquely accessible chromatin.

(A) Reduced dimensionality representation of ASAP-seq data, indicating the GFP and WT sample origins. (B) Annotated cell clusters of populations in ASAP-seq data. (C) Gene activity scores of *Aire* and *Rorc* across ASAP tSNE. (D) Hierarchical clustering of pseudobulk ASAP-seq clusters by all ImmGen ATAC-seq data using a scaled cosine similarity metric. Rows min-max normalized. Immgen populations cluster identities annotated in red as indicated. Aire-expressing populations are indicated with blue (migDC1/2) and red (JC) arrows, and mTECs from the ImmGen database with a green arrow. (E) Accessible chromatin landscape near the *Aire* locus, including the previously described CNS1 region (bottom). (F) Volcano plot of differential accessibility peaks, indicating the number of peaks with greater ($n=12,571$) or less ($n=1,578$) accessibility in Aire expressing populations. *Aire* locus peaks in large red dots; CNS1 locus as indicated. (G) Comparison of number of accessible chromatin fragments between eTACs and non-eTAC populations in the LN by box/violin overlay. Statistical test: Mann-Whitney test. (H) Per-cluster abundance of accessible chromatin fragments/cluster, box/violin overlay, noting highest accessibility in JCs. Boxplots: center line, median; box limits, first and third quartiles; whiskers, 1.5× interquartile range.

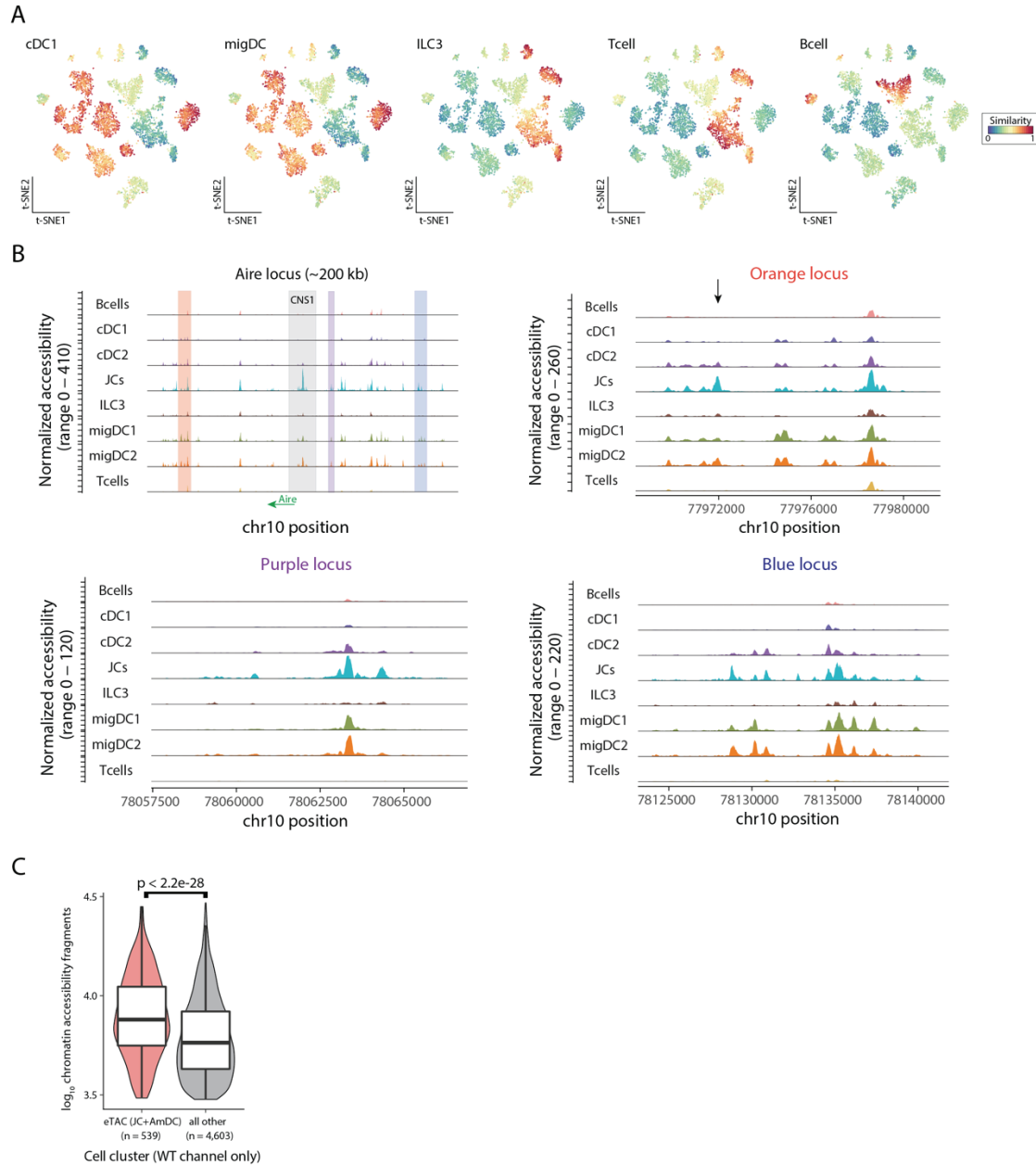


Figure 2-5: Single cell ATAC sequencing data.

(A) ImmGen similarity scores at single-cell resolution for ATAC-seq data for selected populations. (B) Accessible chromatin landscapes near the *Aire* locus. Individual accessible regions are indicated. (C) Comparison of number of accessible chromatin fragments among only the WT sample. Boxplots: center line, median; box limits, first and third quartiles; whiskers, 1.5 \times interquartile range. Statistical test: Mann-Whitney test.

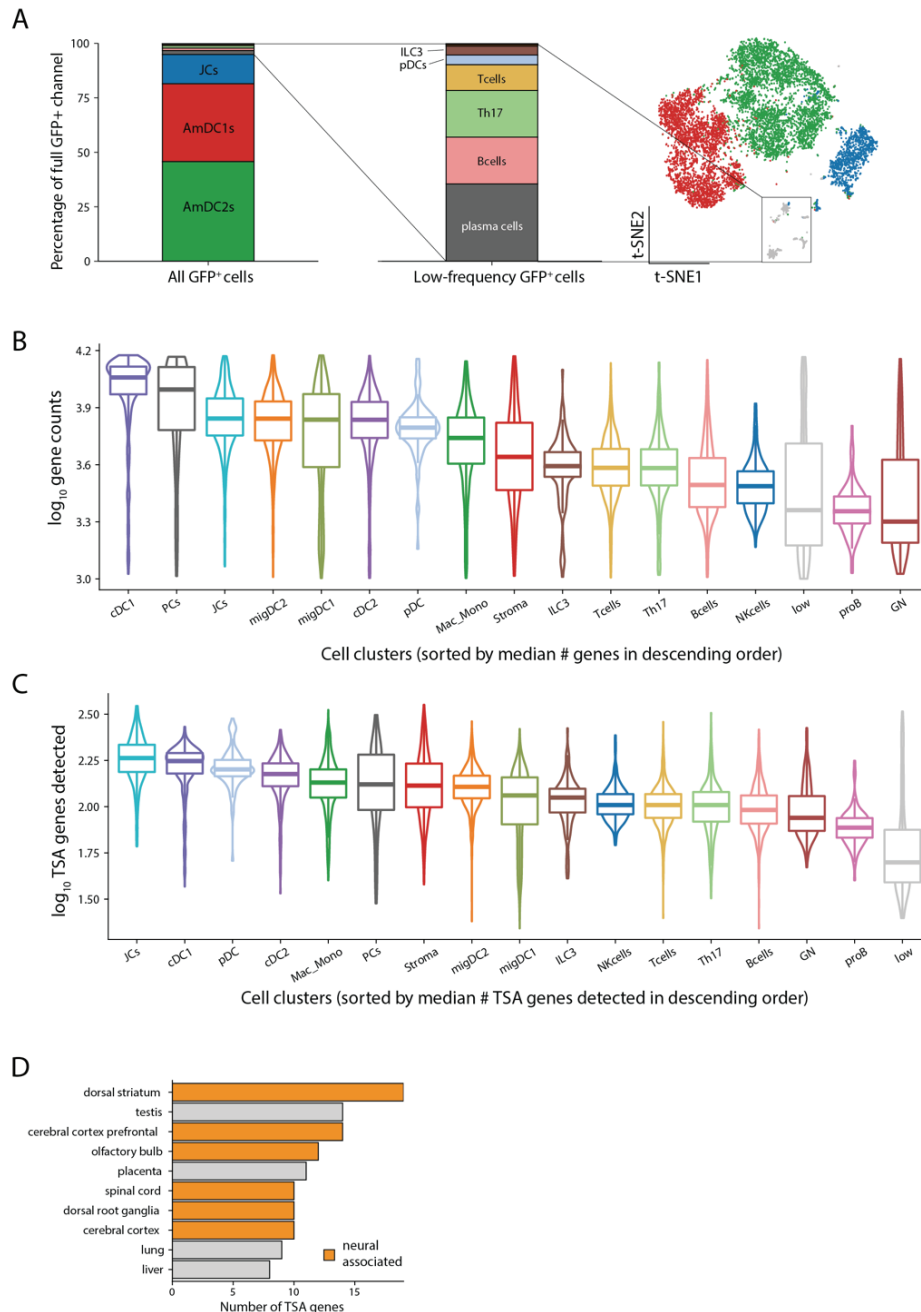


Figure 2-7: TSA in extrathymic Aire-expressing cells.

(A) Description of GFP+ populations, including low abundance populations (grey). (B) Number of genes detected and (C) number of TSAs per population. Boxplots: center line, median; box limits, first and third quartiles; whiskers, 1.5× interquartile range. (D) Count of tissue specific antigen (TSA) genes overexpressed in JCs by tissue as assigned by GeneAtlas. Neuronal tissues indicated in orange.

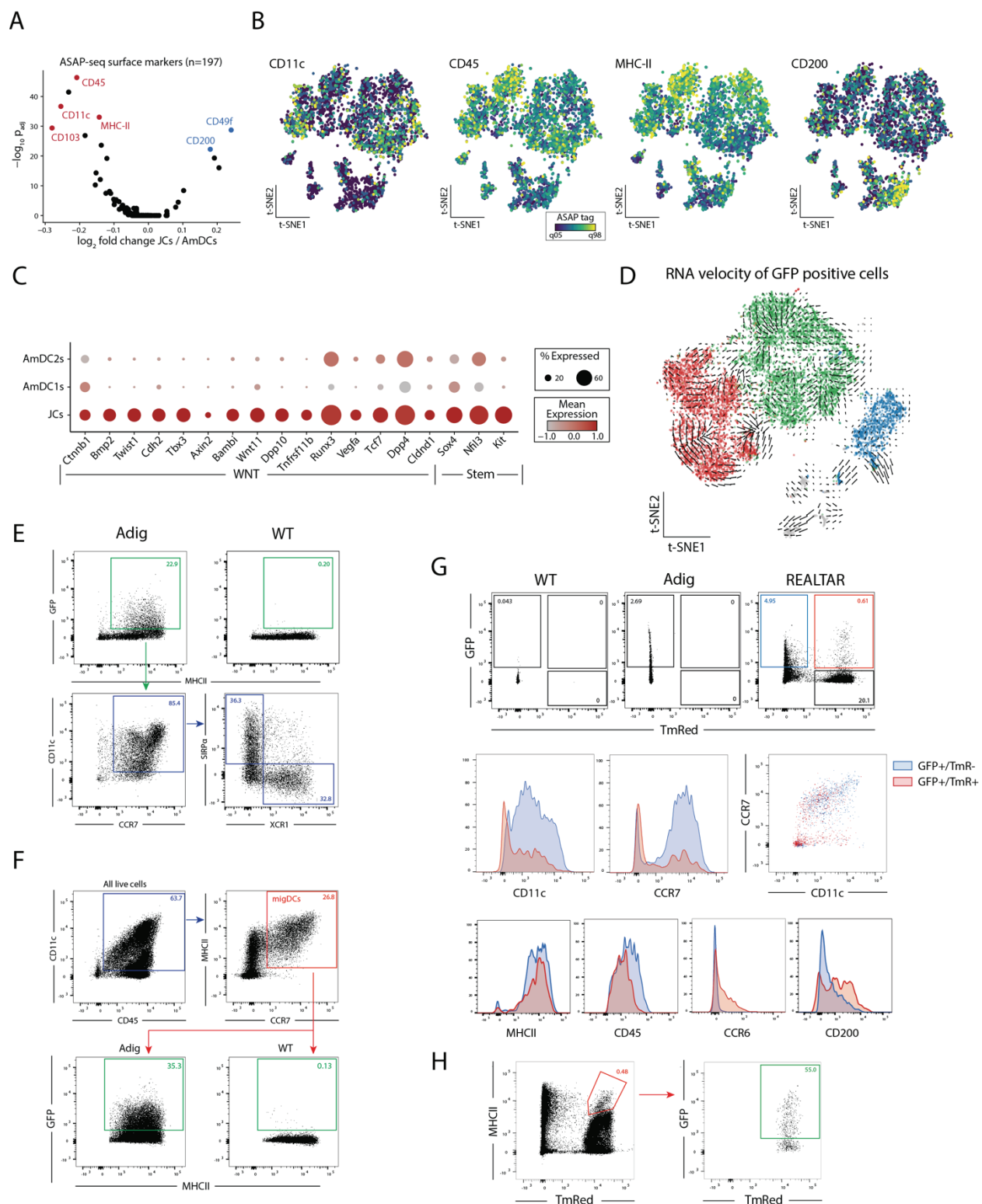


Figure 2-8

Figure 2-8: Surface-marker multiomics combined with functional flow cytometry and lineage-tracing allows for identification and characterization of eTACs and their lineage relationships. (A) Volcano plot of surface markers measured by ASAP-seq comparing JCs to AmDCs. (B) Per-cell visualization of selected surface markers indicated in panel A. (C) Dot plot summarizing expression of WNT and Stem-like genes over-expressed in JCs compared to Aire-expressing migratory DCs. (D) RNA velocity analysis of GFP⁺ populations. (E) Flow cytometry from WT and Adig lymph nodes, pre-gated on single, live, and dump (TCRbeta, CD19, SiglecF, F4/80, NK1.1, Ly6C)-negative cells. (F) Flow cytometry from WT and Adig lymph nodes, pre-gated on single and live cells. (G) Flow cytometry from WT, Adig, and REALTAR mice lymph nodes, pre-gated on single, live, and dump (TCRbeta, CD19, SiglecF, F4/80, NK1.1, Ly6C)-negative cells. (H) Flow cytometry from REALTAR mouse lymph nodes, pre-gated on single and live cells.

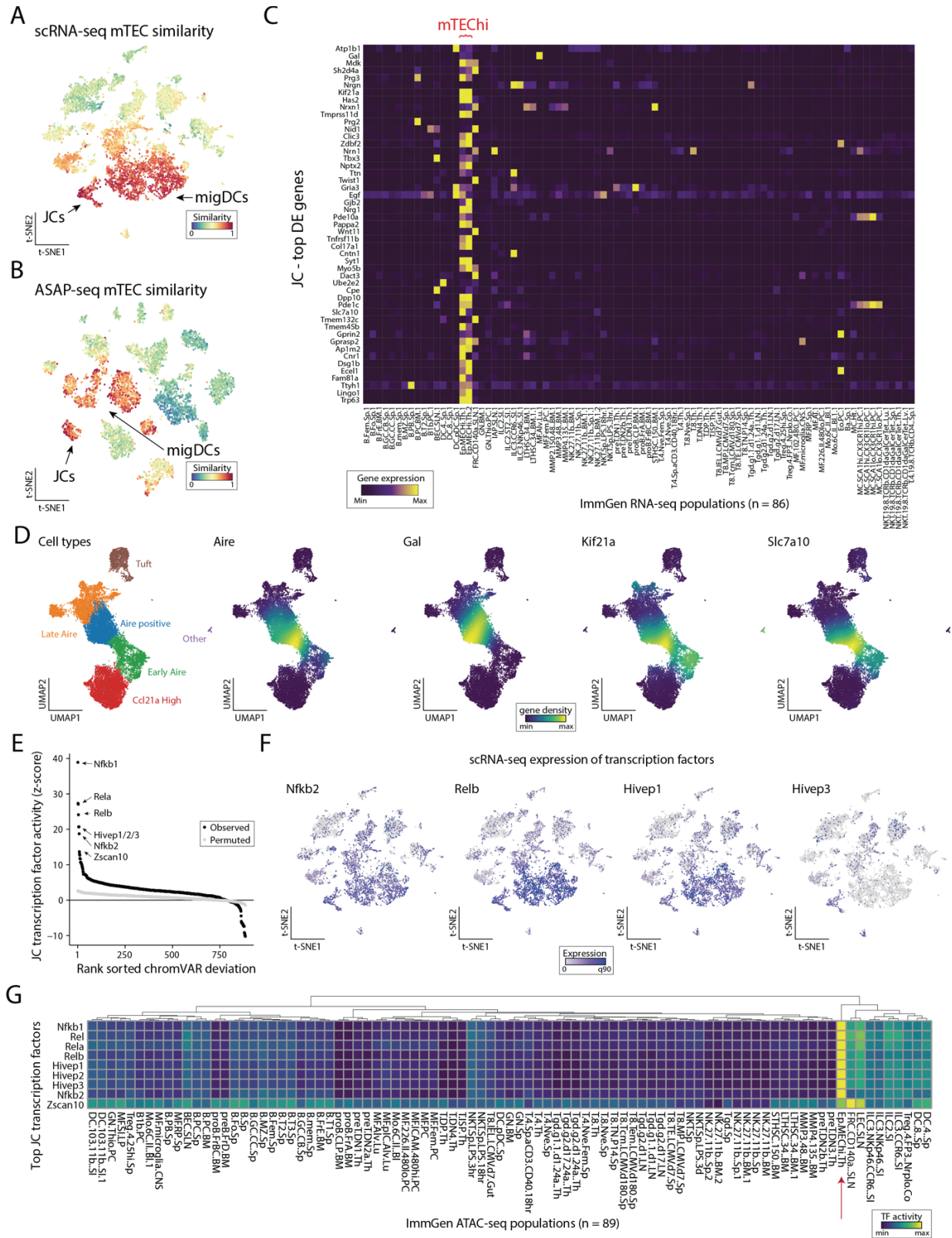


Figure 2-9

Figure 2-9: eTACs are defined by transcriptional and genomic homology to thymic medullary epithelium.

(A) ImmGen similarity scores for scRNA-seq and (B) ASAP-seq data for mTEC populations. JCs and migratory DCs clusters for each embedding are noted by arrow. (C) Heatmap of ImmGen bulk RNA-seq data for genes most overexpressed in JCs relative to all other lymph node populations. mTEChi populations indicated in red text on top. (D) UMAP of aggregated published scRNA data from thymic epithelium showing annotated TEC subsets (left) and gene-density expression for *Aire* and genes enriched in JCs (right). (E) Rank-ordering of transcription factor motifs enriched in JCs compared to other populations. (F) Expression of transcription factors from panel E in scRNA-seq data. (G) ImmGen bulk ATAC-seq chromVAR deviation scores for all 89 populations, highlighting the mTEC population with red arrow. Rows represent top transcription factors identified from JCs and are min-max normalized.

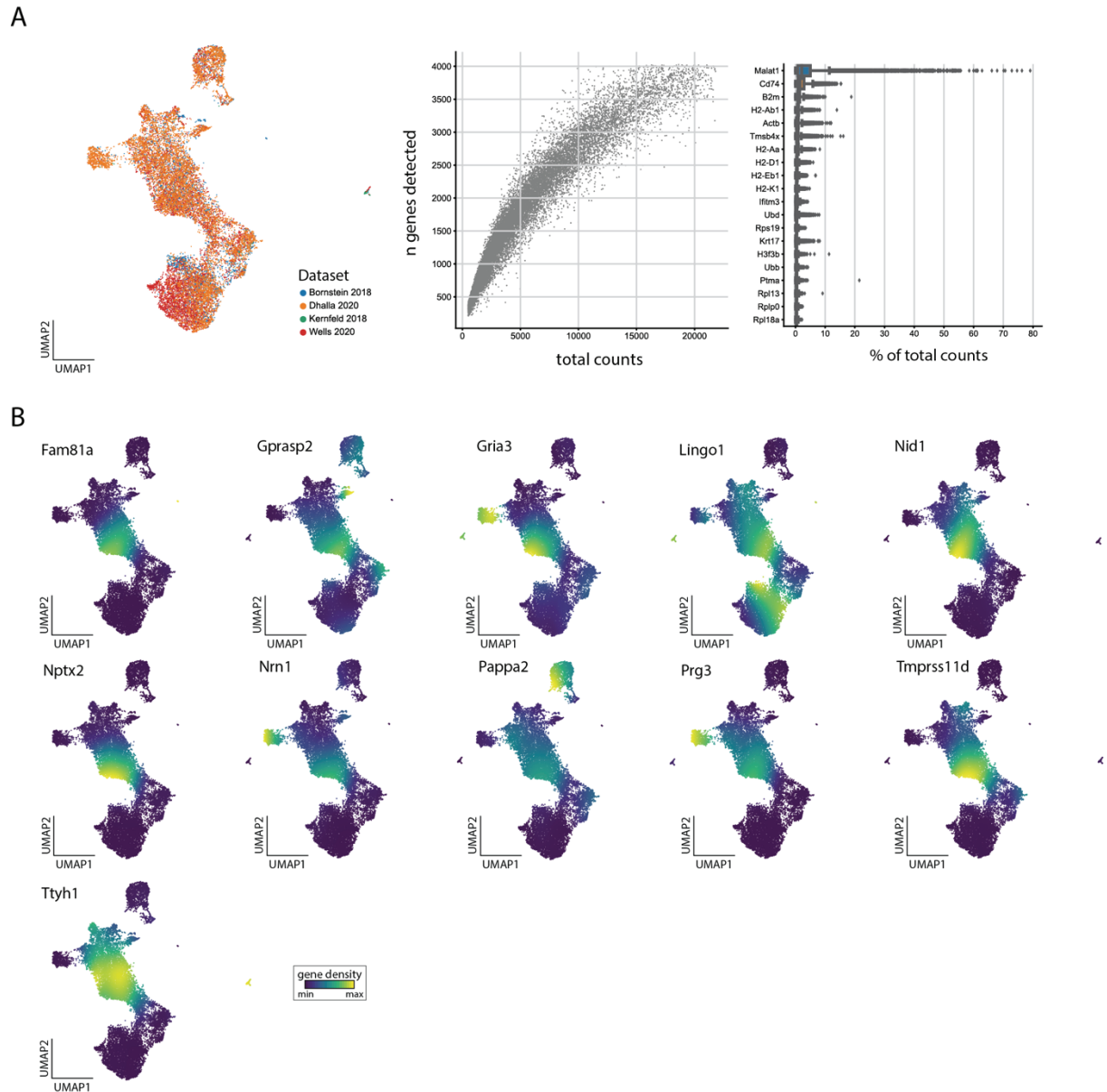


Figure 2-10: Merged thymic epithelial scRNA-seq data.

(A) Summary of quality-control for merged thymus scRNA-seq datasets. (B) Gene- expression densities of JC-specific genes in reduced dimension space from merged previously published thymic epithelial scRNAseq data.

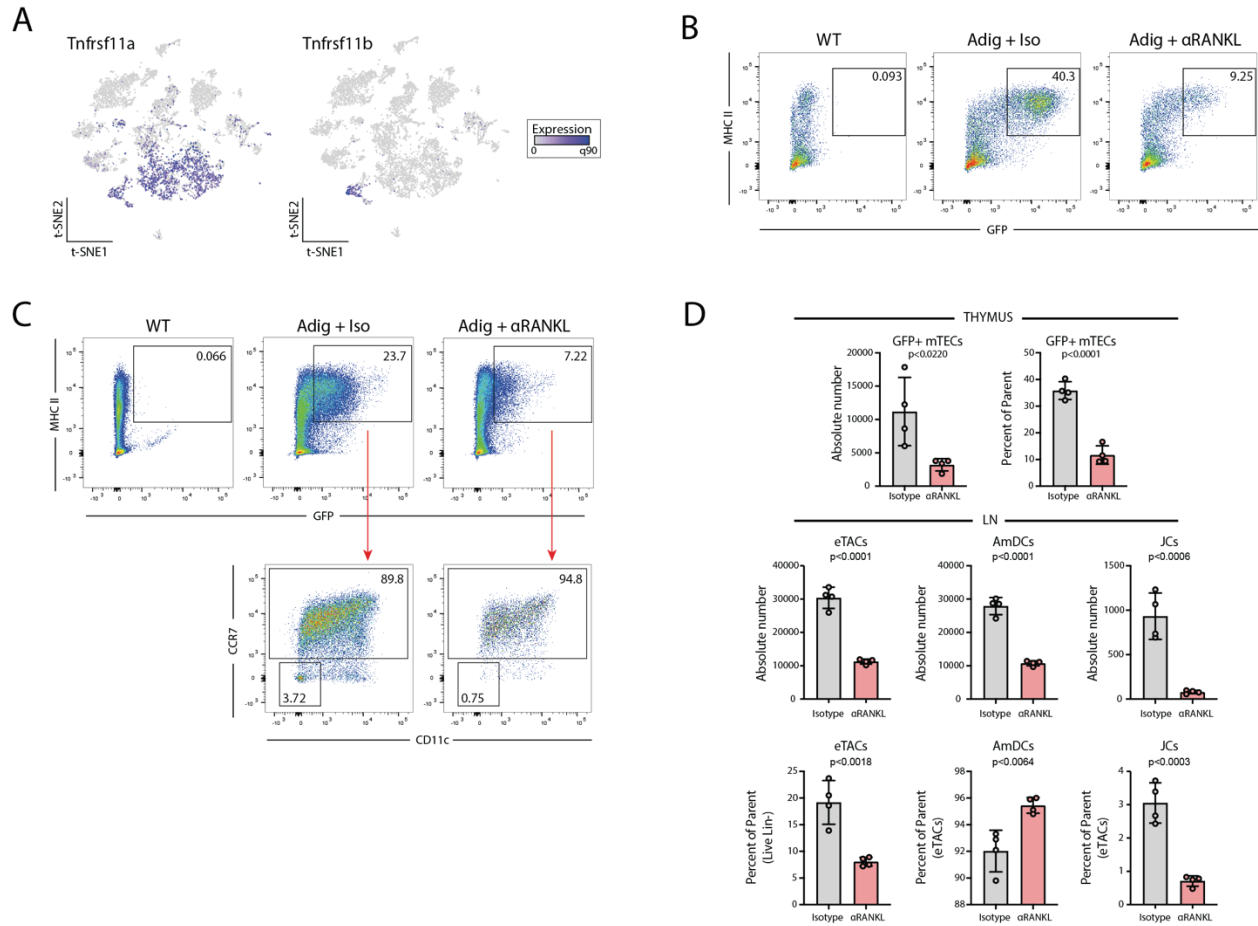


Figure 2-11: RANK-RANK(L) signaling is required for Aire expression in eTACs. (A) Gene expression of RANK (*Tnfrsf11a*) and OPG (*Tnfrsf11b*) from scRNA-seq data. (B) Flow cytometry of Percoll-enriched thymic epithelial cells from WT and Adig mice treated with isotype or α RANKL, pre-gated on single, live, CD11c⁻, CD45⁻, EpCAM⁺ cells. (C) Flow cytometry of lymph nodes from WT and Adig mice treated with isotype or α RANKL treatment, pre-gated on single, live, and dump (TCRbeta, CD19, SiglecF, F4/80, NK1.1, Ly6C)-negative cells. (D) Quantitation of percentage and absolute cell numbers from B and C, unpaired t-test.

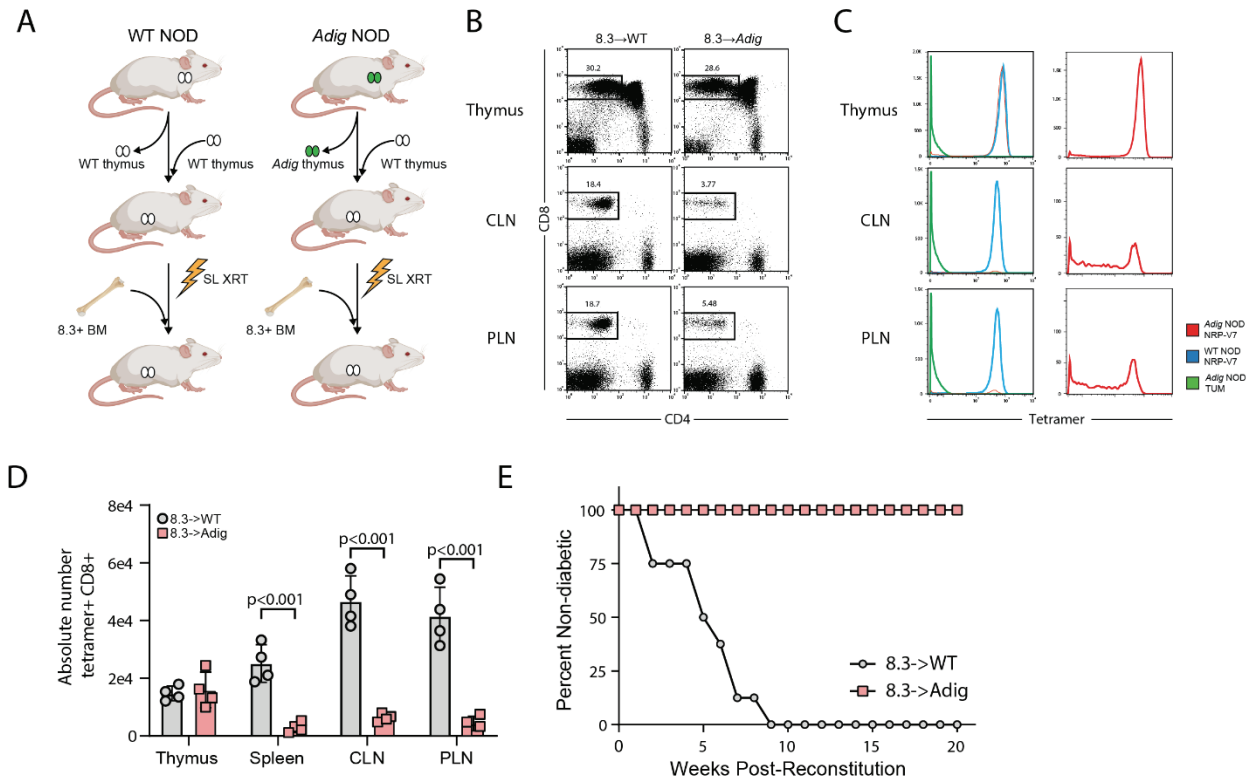


Figure 2-12: Pancreatic self-antigen expression in eTACs is sufficient to induce deletion of T cells escaping thymic selection, and to prevent autoimmune diabetes.

(A) Experimental design for thymic swap and BM chimerism in NOD and Adig NOD mice. SL XRT: sublethal irradiation. BM: bone marrow. (B) Flow cytometry showing lymphocyte populations from indicated tissues. CLN: cervical lymph node. PLN: pancreatic lymph node. (C) Flow cytometry tetramer staining of lymphocyte populations from indicated tissues; panels on right show Adig NOD NRP-V7 staining alone to show distribution of avidity in the periphery, given small number of surviving cells. (D) Quantitation of absolute cell numbers from B, unpaired t-test. (E) Kaplan-Meier curves, diabetes-free survival after bone-marrow reconstitution with 8.3+ BM.

2.5 Materials and Methods

Study design

The aim of this study was to define the identity and biology of eTACs using murine transgenic model systems in combination with advanced single-cell multiomics and integrative computational analyses. We used Adig Aire reporter mice to identify eTACs and sort them, along with a broad sample of immune populations from the mouse LN, and subjected these populations to parallel scRNA-seq and ASAP-seq. We performed flow cytometry using the Adig mouse and an additional REALTAR mouse to validate and extend these results and study lineage relationships between eTAC populations. We used publicly accessible bulk and single-cell RNA and ATAC sequencing data on mTECs, DCs, and other immune populations to define the migDC-like character of eTACs and the parallel biology between eTACs and mTECs. Furthermore, we used antibody blockade of the RANK signaling pathway to determine whether Aire expression in the periphery, like the thymus, depends on these pathways. Last, we used model antigen systems in combination with TCR transgenic BM chimeras and thymic “swap” experiments to determine whether self-antigen expression in eTACs was sufficient to prevent autoimmunity mediated by autoreactive T cells that escape thymic selection.

Mice

C57BL/6J (B6, H-2b), NOD, B6 RORYt-Cre, and B6 Rosa26-LSL-TmR mice were purchased from The Jackson Laboratory. B6 Adig and NOD Adig mice were generated by the laboratory and characterized as previously described⁸. NOD Thy-1.1 8.3 TCR transgenic mice, previously derived and characterized⁸², were bred and used as BM donors. All mice were housed in specific pathogen-free facilities at the University of California San Francisco, and all animal studies were approved by the Institutional Committee on Animal Use and Care at the University of California San Francisco.

Antibodies and tetramers

Fluorochrome- or biotin-tagged antibodies against the following mouse targets were purchased from BD Biosciences [CD11c (N418) and SiglecF (E50-2440)], Thermo Fisher Scientific [TCRb (H57-597), CD3e (145-2C11), B220 (RA3-6B2), Ter119 (TER-119), and MHCII (AF6-120.1)], and BioLegend [CD4 (GK1.5), CD8 (53-6.7), NK1.1 (PK136), Ly6C (HK1.4), CD90.2 (30-H12), CD19 (6D5), F4/80 (BM8), CCR7 (4B12), XCR1 (ZET), SIRPa (P84), TCRgd (GL3), CD45 (30- F11), CD200 (OX-90), CCR6 (29-2 L17), and EpCAM (G8.8)]. Iso- type control antibodies were purchased from the same source as target-specific antibodies. The Fixable Blue Dead Cell Stain Kit for ultraviolet excitation (Thermo Fisher Scientific) was used for viability staining for flow cytometry. DAPI (4',6-diamidino-2-phenylindole) was used for viability staining for cell sorting. MHC I tetramers (I-Ag7) were provided by the National Institutes of Health Tetramer Core Facility (Emory University, Atlanta, GA). I-Ag7 tetramers used for staining CD8⁺ T lymphocytes were the IGRP mimotope/MHC I tetramer NRP-V7-PE (KYNKANVFL) or irrelevant mock tetramer TUM-PE (KYQAVTTTL).

Thymectomy, thymic transplantation, and BM chimera generation

WT neonatal 2-day-old B6 pups were used as thymic donors; thymi were procured and cultured in 12-well transwell plates in 1.35 mM 2-deoxyguanosine (2-DG) in complete Dulbecco's modified Eagle's medium (DMEM)–10 at 37°C for 7 days. One day before transplant (on day 6), 2-DG medium was replaced with regular complete DMEM. At day 7 of thymic incubation, 6- to 8-week-old NOD and Adig NOD mice were anesthetized with ketamine/xylazine and im- mobilized, a midline submandibular incision was made, the salivary glands were reflected upward, and a partial sternotomy was made. The thymic lobes were suction-aspirated, confirming two lobes as- pirated per mouse, the salivary glands were reflected down to cover the partial sternotomy, and the skin was closed with staples. The mice were then rotated on their sides, and the left flank was shaved, after which a 1-cm incision was made and the kidney was exteriorized, the capsule was incised, and the donor thymi (both lobes) were transplanted in subcapsular fashion, after which

the capsular defect was cauterized and the kidney was replaced intra-abdominally. The flank incision was sutured closed with 6-0 Monocryl. Mice were recovered on covered warming pads with buprenorphine/flunixin analgesia and carefully monitored during recovery.

Two weeks after the procedure, experimental NOD thymic swap mice were sublethally irradiated with 900 rads (600/300), and donor BM was prepared from NOD 8.3 Thy-1.1 mice by complement depletion of CD4⁺ and CD8⁺ T cells (anti-CD4 clone GK1.5 and anti-CD8 clone YTS 164.5). Five million BM cells were transferred per recipient via tail vein injection. Mice were analyzed by flow cytometry at 8 weeks after reconstitution or followed for diabetes incidence by weekly urine dipstick or at more frequent intervals if they developed wasting or polyuria.

RANKL blockade

Anti-RANKL antibody (clone IK22/5, BioXCell) or isotype control antibody (clone 2A3, BioXCell) was administered at 100 ug per mouse in phosphate-buffered saline (PBS) every other day via intraperitoneal injections for a total of three injections.

Mouse LN processing and purification

Mouse secondary lymphoid organs (cervical, brachial, axillary, inguinal, and mesenteric LNs) were procured and pooled in digestion medium consisting of RPMI 1640 with 2% fetal bovine serum (FBS) (Sigma-Aldrich), with deoxyribonuclease (DNase) (100ug/ml; Roche) and Liberase (50ug/ml; Roche), minced and agitated at 37°C with gentle agitation for 30 min, and passed through a 70-um filter. Cells were resuspended for magnetic column enrichment (Miltenyi LD column depletion with streptavidin or anti-biotin microbeads and biotinylated antibodies against B220, Ter119, TCRb, ± CD3e). Cells were then either processed directly for 10× single-cell analysis (WT LN) or sorted by flow cytometry for all live, GFP⁺ cells, along with WT (nontransgenic) controls for mean fluorescence intensity (MFI) thresholding. Cells were sorted into PBS with 0.04% bovine serum albumin (BSA). Cell viability and counts were evaluated with Vi-CELL XR (Beckman Coulter), and samples with viability >85% were used for sequencing.

Thymus processing for flow cytometry

Mouse thymi were isolated and minced with razor blades, transferred into 15-ml Falcon tubes in total 10 ml of DMEM (Sigma-Aldrich) with 2% FBS, and vortexed briefly for 15 s. After settling for 5 min, the supernatant medium was removed and replaced with 4 ml of digestion medium consisting of DMEM with 2% FBS (Sigma-Aldrich), with DNase (100 µg/ml; Roche) and Liberase (50 µg/ml; Roche). Thymi were digested in 37°C water bath with mechanical aid by pipetting through a glass Pasteur pipette every 4 min. Every 8 min, tubes were spun briefly to pellet undigested fragments and the digested supernatant was transferred to 20 ml of 0.5% BSA (Sigma-Aldrich), 2 mM EDTA (Teknova), in PBS [magnetic-activated cell sorting (MACS) buffer]. This procedure was repeated for three rounds with 4 ml of fresh digestion medium each time until there were no remaining visible tissue fragments (total 24 min). The single-cell suspension was then pelleted and washed once in MACS buffer and passed through a 70-µm filter. To enrich for stromal cells, single-cell suspension of thymi was subjected to density gradient centrifugation with a three-layer Percoll gradient (GE Healthcare) with specific gravities of 1.115, 1.065, and 1.0. Stromal cells are isolated from the Percoll light fraction, between the 1.065 and 1.0 layers, and resuspended in PBS before counting and flow staining.

Single-cell RNA sequencing

scRNA-seq libraries were generated using the 10× Chromium Controller and the Chromium Single Cell 3' Library Construction Kit v3 according to the manufacturer's instructions. Briefly, the suspended cells were loaded on a chromium controller single-cell instrument to generate single-cell gel bead-in-emulsions (GEMs) followed by reverse transcription and sample indexing using a C1000 Touch Thermal cycler with 96-deep well reaction module (Bio-Rad). After breaking the GEMs, the barcoded complementary DNA (cDNA) was purified and amplified, followed by fragmenting, A tailing, and ligation with adaptors. Last, polymerase chain reaction (PCR) amplification was performed to enable sample indexing and enrichment of scRNA-seq libraries. The final libraries were quantified using a Qubit dsDNA HS Assay kit (Invitrogen) and a High

Sensitivity DNA chip run on a Bioanalyzer 2100 system (Agilent). All libraries were sequenced using NovaSeq S4 (Illumina) with paired-end reads (2 × 150 cycles), and per-sample sequencing libraries were demultiplexed using Cell Ranger v5.

ATAC with selected antigen profiling by sequencing

Cells were stained with a 198 Total-seq A–conjugated antibody panel for murine cells (BioLegend) as previously described⁸³ and outlined online (<https://cite-seq.com/asapseq/>). Briefly, after sorting, cells were fixed in 1% formaldehyde and processed as described for the mitochondrial single-cell assay for transposase-accessible chromatin (mtscATAC-seq) workflow⁵¹, with the modification that, during the barcoding reaction, 0.5 µl of 1 µM bridge oligo A (BOA for TSA) was added to the barcoding mix. For GEM incubation, the standard thermocycler conditions were used as described by 10x Genomics for scATAC-seq. Silane bead elution and solid phase reversible immobilization (SPRI) cleanup steps were modified as described to generate the indexed protein tag library. The final libraries were quantified using a Qubit dsDNA HS Assay kit (Invitrogen) and a High Sensitivity DNA chip run on a Bioanalyzer 2100 system (Agilent). All libraries were sequenced using NovaSeq S4 (Illumina) with paired-end reads (2 × 150 cycles), and per-sample sequencing libraries were demultiplexed using Cell Ranger ATAC v1.2.

scRNA-seq analyses

For each sequencing library, per-cell, per-gene counts were generated from raw sequencing reads (.fastq files) using Cell Ranger v5 count using the mm10 reference genome. Each library (GFP and WT) yielded high-quality data, including 11,697 and 10,953 cells for the GFP and WT libraries with a median number of genes detected per cell of 1831 and 1451, respectively. To perform integrative down- stream analyses, we used the Seurat analytical toolkit⁸⁴. First, we filtered putative cell doublets using DoubletFinder⁸⁵ and down- sampled the GFP to 3000 cells to allow for a more balanced representation of the LN populations in the reduced dimension space. Next, we filtered for cells meeting the following criteria: greater than 500 genes detected and 1000 unique molecular identifiers (UMIs), no more than 10% mitochondrial RNA UMIs, and no more

than 15,000 total UMIs detected per cell—thresholds based on inspection of the empirical density of cells—resulting in 2532 GFP+ cells and 6973 WT cells. After normalization and variable gene inference, we corrected the top 30 principal components for the sequencing batch using Harmony⁸⁶. The Harmony-corrected components were then used for identifying cell neighbors, clusters, and reduced dimensionality coordinates. Visualization of populations and gene abundances was performed using the default Seurat functionality. For GFP-positive analyses (Fig. 2-6), we used similar thresholds but did not down-sample the library to increase power to discriminate features between JCs and migDCs. RNA velocity analyses were performed using velocityto using the default parameters for this GFP+ population⁶⁷.

ASAP-seq analyses

For this multimodal capture, scATAC-seq data were aligned and processed using the Cell Ranger ATAC v1.2 pipeline with the mm10 reference genome. For the ASAP-seq antibody tag data, per-cell, per-antibody tag counts were enumerated via the kite/kallisto/bustools framework accounting for unique bridging events as previously described⁸³. Cells called by the Cell Ranger ATAC knee call were further filtered on the basis of abundance accessible chromatin (>3000 nuclear fragments) as well as accessible chromatin enrichment [>4 transcription start site (TSS) enrichment score] and minimal nonspecific antibody binding (<2000 total antibody tag counts). We further down-sampled the GFP-positive population to ensure a more balanced representation of all populations in the LN as described in the previous section. Next, gene activity score calculations [via Signac⁸⁷] were performed using the label transfer functionality originally described in Seurat v3⁸⁴ for our well-defined cluster annotations from the scRNA-seq data. We excluded cells with a <60% confidence for the maximum population assigned. Downstream analyses, including dimensionality reduction, transcription factor binding analysis, and accessible chromatin visualization, were all performed using functions from the Signac package⁸⁷, including those that wrap chromVAR⁸⁸.

Cell type annotation/ImmGen similarity scores

To assign identities to the resulting cell subpopulations/clusters in an unbiased way, we developed an approach to query the similarity of single-cell RNA-seq profiles against all 178 publicly available reference microarray profiles in the ImmGen database⁴⁹; all ImmGen primary data were downloaded from haemosphere (www.haemosphere.org/datasets/show). The similarity measure that we used is based on the cosine distance between appropriately normalized microarray/ gene expression profiles—considering only genes ($n = 2000$) that were identified as highly variable in scRNA-seq. For each reference subpopulation, the corresponding ImmGen microarray profile was normalized to unit sum. The single-cell gene expression profiles were first scaled to 1×10^4 total counts and logarithmized before being normalized to unit sum. Each expression dataset was then subsequently scaled. For each single-cell profile, the resulting cosine similarities to all ImmGen profiles were further minimum-maximum (min-max) scaled in a range between 0 and 1 to obtain the final ImmGen similarity scores. We applied this approach for scoring both individual cells (t-SNE visualization) and pseudobulk cluster profiles (clustered heatmap visualization). Final cell type annotation was carried out in a semi-supervised way, cross-validating the resulting ImmGen scores with subpopulation-specific gene expression of known markers and shown in Fig. 2-3. An analogous workflow was devised for the ATAC scoring of populations where the top ~9000 variable peaks were used as a feature set before performing similarity estimates between the pseudobulk and scATAC-seq data generated here and the bulk ImmGen ATAC-seq database.

TSA assignment, annotation, and expression analysis

To examine TSAs, we downloaded the corresponding list from Sansom et al.⁵⁸, consisting of 6610 genes, 202 of which were significantly up-regulated in JCs compared with other Aire-expressing populations. To map these TSAs to their corresponding tissue, we downloaded the GeneAtlas MOE430 dataset from BioGPS and excluded cell lines from the data resource, resulting in 89

tissues for further examination. The distribution of tissues with JC-specific TSAs represented the tissue with the maximum expression per gene from this resource.

Merging and analysis of previously published mTEC scRNA-seq datasets

Raw transcript data from four previously published thymic scRNA-seq datasets were collected and merged, filtering on genes that were detected across all four datasets⁷⁰⁻⁷³ and on cells isolated from WT mice. Data likely to correspond to doublets were excluded by filtering out cells with the total reads and number of genes in the top 1%. Further quality control was performed by excluding cells containing fewer than three genes. Batch correction and normalization of raw read counts were performed with scVI⁶⁹ using the default parameters for model training and batch keys corresponding to each sequencing run within the four datasets. An scVI model trained on the top 5000 highly variable genes was used to generate a uniform manifold approximation and projection (UMAP) for the merged data, which was then subset on the 18,349 cells in mTEC clusters based on cell type labels assigned in the published analysis of the individual datasets. A new UMAP was generated for this subset of the filtered data, and an scVI model trained using all 16,786 genes in the dataset was used to normalize gene expression for visualization.

Statistical analyses and visualization

All data points are shown in the graphs as mean value indicated and individual data points as shown. For all boxplots: center line, median; box limits, first and third quartiles; whiskers, 1.5× interquartile range. All statistical analyses were performed using GraphPad Prism 8 and R v4.0.3. Statistical significance was calculated using one-way analysis of variance (ANOVA) with Tukey multiple comparison, the un-paired two-sided t test, or the nonparametric Mann-Whitney U test. Diabetes-free survival was analyzed using the log-rank test. *P < 0.05, **P < 0.005, ***P < 0.0005, and ****P < 0.00005 were considered significant. Experimental graphics were generated with BioRender (<https://biorender.com/>).

Data and materials availability:

The sequencing data are available through the Gene Expression Omnibus under accession GSE176282. All other data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials.

Chapter 3 RORyt eTACs are myeloid-derived mediators of oral tolerance and Treg induction

3.1 Abstract

A critical feature of immune tolerance is the ability to distinguish pathogens from innocuous signals such as dietary antigens, but the precise mechanisms of oral tolerance induction and the key cellular actors have remained unclear. Classical dendritic cells (cDCs) have been implicated in the generation of regulatory T cells (Tregs) and T-cell anergy in response to food antigens⁸⁹, but here we demonstrate that RORyt-lineage antigen-presenting cells (APCs), specifically RORyt-lineage APCs expressing the Autoimmune Regulator (Aire) gene – referred to as RORyt eTACs – are a critical population required for oral tolerance induction. Using lineage tracing combined with next-generation gel bead-in emulsion (GEM) single-cell sequencing, we map the identity of these populations at unprecedented depth and reconcile the identity of RORyt-lineage APC populations across all known single-cell transcriptomic datasets to establish consensus identities. We show that RORyt eTACs are of a unique myeloid-derived lineage and are both necessary and sufficient for oral tolerance induction. Upon selective depletion of RORyt eTACs, mice fail to develop adequate dietary antigen-specific Tregs and display severe delayed-type hypersensitivity (DTH) responses in recall assays. Antigen presentation by lymphoid RORyt-lineage APCs (ILC3 and lymphoid tissue inducer cells) is not required for oral tolerance induction, as functional RORyt eTACs alone rescue diet-specific Treg generation. These findings establish RORyt eTACs as critical myeloid-derived mediators of oral tolerance induction and suggest novel cellular targets to modulate allergic sensitisation and immune tolerance.

3.2 Introduction

Food allergies are increasingly prevalent worldwide, but effective treatments and prevention strategies are lacking^{90–92}. While it is known that food allergy arises from inappropriate immune reactivity to orally digested dietary antigens, there is a limited understanding of oral tolerance mechanisms, and the antigen-presenting cell (APC) populations involved. Conventional dendritic cells (cDCs) have been suggested as key players in tolerance against both commensal and dietary antigens via induction of antigen-specific regulatory T cells (Tregs) in the mesenteric lymph node (mLN)^{93–98}. However, recent reports suggest that RORyt⁺ APCs, not cDCs, induce tolerance to intestinal commensal bacteria^{33–35}. We thus sought to determine if RORyt-lineage APCs are also involved in the induction of oral tolerance and to better define their identity, lineage, and taxonomy.

We found that RORyt-lineage APCs induce oral antigen-specific Treg differentiation while suppressing T follicular helper cells (Tfh). We and others have previously described a unique APC population referred to by several names including Janus cells (JCs)³², Thetis cells (TCs)³⁴, Aire⁺ ILC3-like cells³¹, and RORyt⁺ eTACs⁹⁹ that express Aire and high levels of major histocompatibility complex class II (MHCII) and are capable of inducing cognate CD4 and CD8 T cell tolerance^{8,9}. Recently, the term “RORyt⁺ eTACs” was suggested as a consensus nomenclature^{99,100} and will be utilised here, with the understanding that this remains an evolving field. Here, we show that RORyt eTACs, innate lymphoid cells type 3 (ILC3s) / lymphoid tissue inducer cells (LTis), and minor populations of dendritic cells (DCs) are the three principal RORyt-lineage APC populations in mouse secondary lymphoid organs (SLOs) and are consistent across both age and location. While ILC3s and LTis are of lymphoid lineage, mixed bone marrow chimaera (BMC) studies reveal that RORyt eTACs are of a myeloid lineage. Oral tolerance is intact in mice with functional RORyt eTACs but not ILC3s/LTis. In contrast, the ablation of RORyt eTACs leads to impaired dietary antigen-specific Tregs and exaggerated delayed-type

hypersensitivity (DTH) reactions to food antigens. These findings highlight the critical role of ROR γ t eTACs in maintaining immune tolerance at mucosal barriers.

3.3 Results

3.3.1 ROR γ t-lineage APCs mediate oral tolerance to food antigen

Recent reports have demonstrated the essential role of antigen presentation by ROR γ t-lineage APCs in the induction of Tregs against commensal bacteria in the gut using *ROR γ t^{Cre} × MHCII^{fl/fl}* (MHCII Δ ROR γ t) mice^{28,29,33–35}. Like commensal tolerance, the immune system must also recognise and tolerate food antigens to prevent allergy, and we sought to understand if ROR γ t-lineage APCs are also crucial for inducing oral tolerance to dietary antigens. To monitor T-cell responses against dietary antigens we adoptively transferred naïve OT-II CD4⁺ T cells – specific for chicken ovalbumin (OVA) – into wildtype or MHCII Δ ROR γ t mice (Fig. 3-1A). After oral administration of OVA, our model dietary antigen, we found diminished OT-II ROR γ t⁻ and ROR γ t⁺ Tregs in the mLN of MHCII Δ ROR γ t mice compared to WT control mice (Fig. 3-1B-C, Fig. 3-2A). The induction of anergic T cell clones have also been implicated in oral tolerance^{101–103}, and MHCII Δ ROR γ t mice exhibited decreased percentages of FR4⁺CD73⁺ anergic OT-II T cells (Fig. 3-2A). Consistent with previous reports, percentages of endogenous ROR γ t⁺ Tregs were decreased while T helper 17 (Th17) cells were increased in MHCII Δ ROR γ t mice, but ROR γ t⁻ Tregs were not different from WT control mice (Fig. 3-2B). After 7 days of OVA feeding via oral gavage, MHCII Δ ROR γ t mice had similar reductions in the percentage of OVA-specific Tregs and elevated Tfh in both the mLN and Peyer's patches (PP) (Fig. 3-1D, Fig. 3-2C). To measure the long-term impact of ROR γ t-lineage APC loss on oral tolerance, we next utilised an established mouse model of systemic DTH^{95,104}. As MHCII Δ ROR γ t mice develop gut dysbiosis in adulthood due to the absence of these APC populations in early life^{28,29} – which could impair oral tolerance simply as a secondary effect of general gut inflammation – we first generated BMC mice with either WT or MHCII Δ ROR γ t bone marrow (BM). After 8 weeks of reconstitution, BMC mice were given oral OVA by gavage on days 0 and 1 and immunised with OVA in complete Freund's adjuvant (CFA) on day 9, followed by subcutaneous

injection of OVA in the ear 14 and 24 days after immunisation (Fig. 3-1E). Tolerance was assessed by measurement of ear swelling after subcutaneous injection of OVA into the ear. MHCII^{ΔRORyt} mice fed with OVA had markedly increased ear swelling compared to controls (Fig. 3-1F), demonstrating that RORyt-lineage APCs are crucial for establishing tolerance against oral antigens via induction of dietary antigen-specific Tregs.

3.3.2 Harmonising the diversity of RORyt-lineage APC populations

To precisely define the landscape of RORyt-lineage APC populations, we performed next-generation gel bead-in emulsion (GEM) single-cell RNA sequencing (scRNA-seq) analysis on all RORyt-lineage MHCII⁺ APCs in mouse SLOs. We obtained mLN, spleen and other LNs (cervical, inguinal, axillary, brachial, and para-aortic) from adult RORyt-lineage mice (*RORyt*^{Cre}×*Rosa26-LSL-tdTomato*) and FACS-sorted tdTomato⁺ MHCII⁺ cells for subsequent scRNA-seq. We found that RORyt-lineage APCs consist of three main populations: ILC3s/LTis, RORyt eTACs (R-eTACs; previously referred to as Janus Cells, Thetis Cells, and ILC-like Aire⁺ Cells)¹⁴, and minor populations of DCs (Fig. 3-3A, Fig. 3-4A-B). All three populations showed tdTomato (tdTmt) reporter transcript (Fig. 3-4A). R-eTACs were distributed across all SLOs and not unique to the mesentery, contrary to previous report³⁴ (Fig. 3-4C-D). R-eTACs were previously described as expressing both *Aire* and *Rorc*, and our current scRNA-seq data revealed differential expression of *Aire* and *Rorc* among 3 distinct R-eTAC subsets. R-eTAC1 and 3 have higher *Aire* expression, while R-eTAC2 have the highest *Rorc* expression (Fig. 3-3B). All R-eTACs subsets express MHCII, but otherwise have quite distinct transcriptional signatures (Fig. 3-3C, Fig. 3-4A and E). Expression of *Itgav* and *Itgb8* has been shown to be critical for RORyt⁺ APCs to induce commensal-specific Tregs^{33,34}, and here we found *Itgav* and *Itgb8* expression in both R-eTACs and DC subsets (Fig. 3-4F). We also carried out parallel scRNA-seq of RORyt-lineage MHCII⁺ populations in SLOs of early-life (2–3-week-old) mice and found high concordance between these populations and their adult counterparts (Fig. 3-3D, Fig. 3-4G). Finally, because of the recent confusion around these populations, including varied nomenclature and sorting strategies, we

reconciled our scRNA-seq data with all previously published single-cell data on ROR γ t⁺ APCs^{32–35}. We found high transcriptional homology across all datasets, with R-eTAC1-3 corresponding to eTACs I/II and TC I-III, with less concordance with TC IV, which were not found in our data or others (Fig. 3-3E). Utilising single-cell variational inference (scVI), we merged and clustered all known ROR γ t APC datasets into one UMAP which demonstrated the consistency of these identities (Fig. 3-3F, Fig. 3-4H). In summary, we have identified ILC3s/LTis, ROR γ t eTACs, and rare DCs as the principal populations of ROR γ t-lineage APCs in mouse SLOs, encompassing the landscape of putative populations responsible for the induction of oral tolerance.

3.3.3 ROR γ t eTACs are antigen presenting cells of myeloid lineage

To identify R-eTACs subsets by flow cytometry, we utilised the previously described ROR γ t Expression And Lineage Tracer and Aire Reporter (REALTAR) mouse (ROR γ t^{Cre}×Adig *Aire*-reporter×Rosa26-LSL-tdTmt)³². This allowed us to trace cells actively expressing *Aire* by GFP expression and all ROR γ t-lineage cells by tdTmt. Additionally, we can precisely identify R-eTACs in non-reporter mice with intracellular Aire staining by flow cytometry (Fig. 3-5A). We found >90% of Aire⁺ cells by flow cytometry are tracked by tdTmt in REALTAR mice, indicating that they are R-eTACs (Fig. 3-5B). By gating on GFP⁺ tdTmt⁺ from REALTAR mice, we found that R-eTACs are enriched in early life between 2-3 weeks of age, decline after weaning, and persist into late adulthood (Fig. 3-6A), concordant with prior findings^{34,105}. R-eTACs are present in all SLOs in mice, including all lymph nodes and spleen, as seen in both our scRNA-seq (Fig. 3-5C) and flow cytometry data (Fig. 3-5D). By cell number, R-eTACs are slightly more abundant within mesenteric and cervical LNs, supporting their potential tolerogenic roles at mucosal sites (Fig. 3-6B). Furthermore, we can differentiate R-eTAC1 and R-eTAC2/3 subsets by flow cytometry, concordant with our scRNA-seq results (Fig. 3-5E); R-eTAC1 can be identified as CCR7⁺-EpCAM⁺ while R-eTAC2/3 are EpCAM⁺ (Fig. 3-5F).

Next, we sought to determine the antigen processing and presentation ability of R-eTACs by co-culturing with naïve OT-II in the presence of OVA protein. R-eTACs were sorted from

REALTAR SLOs as GFP⁺ tdTmt⁺, along with control DC populations. In the presence of whole OVA protein after 72 hours of co-culture, R-eTACs were able to process and present OVA and induce OT-II proliferation to a similar extent as control GFP⁻ DCs (Fig. 3-5G, Fig. 3-6C). Interestingly, R-eTACs also induced markedly higher frequency of Foxp3⁺ OT-II Tregs compared to GFP⁻ DCs, again supporting their tolerogenic capabilities (Fig. 3-5G).

We have previously shown that ROR γ t eTACs are transcriptionally most similar to migratory DCs³², but their hematopoietic lineage remains unknown. Extrathymic ROR γ t⁺ Aire⁺ cells have been previously described as a population of ILC3-like cells³¹, with suggested potential interconversion between ILCs and ROR γ t eTACs¹⁰⁶. To determine whether R-eTACs share a lymphoid lineage with ILCs, we next utilised *Il7r*^{-/-} mice, which lack common lymphoid progenitors and fail to generate both lymphocytes and ILCs¹⁰⁷. We generated mixed BMC mice with WT and *Il7r*^{-/-} BM at a 1:1 ratio and quantified both lymphoid and myeloid cell subsets in the lymph nodes after 8 weeks of reconstitution. As expected, *Il7r*^{-/-} BM failed to give rise to ILCs but retained comparable levels of DCs (Fig. 3-5H). Surprisingly, *Il7r*^{-/-} BM was able to give rise to normal numbers of R-eTACs, suggesting that this population is not of lymphoid lineage or does not pass through a common lymphoid progenitor stage, as previously suggested¹⁰⁸ (Fig. 3-5H). By gating Aire⁺ R-eTAC1 and eTAC2/3 based on EpCAM and CCR7 staining, we further confirmed that *Il7r*^{-/-} BM gave rise to all R-eTAC populations (Fig. 3-6D-E). We further confirmed these findings in a parallel mixed BMC model using *Rag2*^{-/-}*Il2rg*^{-/-} mice, which also lack adaptive lymphocytes and ILCs¹⁰⁹. Here again, ILCs were not found among the populations derived from the *Rag2*^{-/-}*Il2rg*^{-/-} BM, while R-eTACs developed normally and showed no difference between WT or *Rag2*^{-/-}*Il2rg*^{-/-} BM (Fig. 3-6F).

To further map the presumptive myeloid lineage of R-eTACs, we used both *Csf1r*^{Cre}×*Rosa26*-LSL-tdTmt¹¹⁰ and *Cx3cr1*^{CreER}×*Rosa26*-LSL-tdTmt¹⁰⁰ lineage-tracing systems and found that in both models R-eTACs were labelled at similar percentages to DCs, while lymphocytes were not traced (Fig. 3-5I, Fig. 3-6G). Given the transcriptional homology between

R-eTACs and migratory DCs^{32,34,111,112}, we next sought to address whether they share common progenitors. Recombinant Flt3L (rFlt3L) administration, which has been shown to induce expansion of DC progenitors^{113,114}, resulted in a 30-fold increase in the absolute numbers of DCs but no change in R-eTACs (Fig. 3-6H). We also utilised the *Clec9a*^{CreER}×*Rosa26*-LSL-tdTmt mice¹¹⁵ lineage-tracing system for tracking DC progenitors and determined that less than 5% of R-eTACs were labelled (Fig. 3-6I), consistent with prior results^{34,116}. Thus, while R-eTACs and their subsets can be identified by lineage tracing and intracellular Aire staining, and are functional, tolerogenic APCs of a myeloid lineage, they are distinct from lymphocytes, ILCs, and cDCs.

3.3.4 ROR γ t eTACs mediate oral tolerance to food antigen

To determine if ROR γ t eTACs or ILC3/LTis are responsible for diet-specific Treg induction, we next generated mixed BMC mice with MHCII Δ ROR γ t and *Il7r*^{-/-}. Since *Il7r*^{-/-} BM is deficient in all adaptive and innate lymphoid cells but has intact R-eTACs (Fig. 3-5H), mice with MHCII Δ ROR γ t and *Il7r*^{-/-} mixed BM lack all ILC3/LTis as APCs but have functional antigen presentation from R-eTACs (Fig. 3-7A). By comparing MHCII Δ ROR γ t+*Il7r*^{-/-} mixed BMC with WT or MHCII Δ ROR γ t BMC, we can thus determine whether antigen presentation by R-eTACs is sufficient to rescue the loss of oral antigen-specific Tregs in MHCII Δ ROR γ t mice (Fig. 3-7A). Indeed, antigen presentation by ILC3s/LTis is dispensable for oral tolerance induction as functional R-eTACs alone in MHCII Δ ROR γ t+*Il7r*^{-/-} mixed BMC mice restored food-specific Tregs and lowered the percentage of Tfh to similar levels as mice with WT BM (Fig. 3-7B). While this approach does not exclude a potential role for ROR γ t-lineage DCs, recent parallel evidence argues strongly against a functional role for these populations¹¹⁷.

Next, to determine if Aire-expressing R-eTACs are responsible for inducing food-specific Tregs, we utilised our AireDTR mice, which induce ablation of all Aire-expressing cells, including R-eTACs, upon intraperitoneal administration of diphtheria toxin (DT) (Fig. 3-7C). Dosage frequency was determined based on pulse-chase experiments which show a reconstitution half-life for R-eTACs of approximately 7 days (Fig. 3-8B). To determine if R-eTACs are necessary for

oral tolerance, we transferred naïve OT-II CD4⁺ T cells into DT-treated AireDTR⁻ or AireDTR⁺ mice followed by 2 oral gavages of OVA (Fig. 3-7D). AireDTR⁺ mice harvested after 2 days of oral OVA had significantly decreased percentages of OVA-specific RORγt⁻ and RORγt⁺ OT-II Tregs and anergic OT-II with increased percentage of Bcl6⁺ Tfh in the mLN (Fig. 3-7E, Fig. 3-8C). Aire-DTR⁺ mice harvested after 7 days of oral gavage also had diminished percentage of Foxp3⁺ OT-II and elevated Tfh in the mLN and Peyer's Patches (Fig. 3-7F, Fig. 3-8D-E).

To determine whether Aire-expressing RORγt eTACs are required to rescue the loss of dietary antigen-specific Treg observed in MHCII^{ΔRORγt} mice, we generated mixed BMC with WT, MHCII^{ΔRORγt}, and AireDTR mice and performed similar experiments as Fig 3-7D. Validation of BMC showed the loss of MHCII on Aire-expressing cells from MHCII^{ΔRORγt} and MHCII^{ΔRORγt} + AireDTR mice (Fig. 3-8F). After the transfer of OT-II followed by OVA gavage in these mice, we observed marked reductions in both RORγt⁻ and RORγt⁺ Tregs from MHCII^{ΔRORγt} + AireDTR mice compared to WT BMC, with minor rescue compared to MHCII^{ΔRORγt} BMC mice (Fig. 3-7G). Of note, although extrathymic Aire expression in RORγt-lineage APCs has been shown to be important in the induction of Th17 cells in response to *Candida albicans*¹⁰⁵, Aire itself was dispensable for oral tolerance induction, as crossing Aire^{fl/fl} mice with pan-hematopoietic Vav1^{iCre} (Aire^{ΔVav1}) or RORγt^{iCre} (Aire^{ΔRORγt}) showed no defect in OT-II Treg generation (Fig. 3-8G-H).

Finally, to assess the requirement for RORγt eTACs in preventing recall allergic responses, we again utilised the DTH model, in which BMC mice were reconstituted with either AireDTR⁻ or AireDTR⁺ BM cells and challenged with subcutaneous OVA ear injection (Fig. 3-8I). AireDTR⁺ mice in which eTACs had been ablated had significantly more ear swelling than AireDTR⁻ mice after second subcutaneous OVA injection (Fig. 3-7H), highlighting a critical role for R-eTACs in memory recall responses to oral tolerance.

Overall, deletion of peripheral Aire-expressing cells results in impaired T-cell tolerance against dietary antigens in the gut, while antigen presentation by R-eTACs alone was able to rescue food-specific Tregs in MHCII^{ΔRORγt} BM mice. These data support RORγt eTACs as a

unique antigen-presenting population playing a critical role in the induction and maintenance of oral tolerance.

3.4 Discussion

The immune system's ability to distinguish pathogens from innocuous self and commensal antigens is essential for maintaining gut homeostasis and overall health. Recent reports have established that ROR γ ⁺ APCs induce tolerance to commensals via peripheral Treg differentiation, but their involvement in tolerance to food antigens remains unclear. Here, we aimed to determine which ROR γ -lineage APC populations are required to induce food-specific Tregs. We found that ROR γ -lineage APCs in mouse SLOs consist of three main subsets: Aire-expressing R-eTACs, ILC3s/LTis, and rare DCs. In the absence of R-eTACs, mice have attenuated induction of food-specific Tregs and thus do not establish intestinal tolerance to food antigens.

Interestingly, MHCII ^{Δ ROR γ} mice had more severe ear inflammation after subcutaneous OVA injections than Aire-DTR⁺ mice in the DTH model, and MHCII ^{Δ ROR γ} + Aire-DTR mixed BMC had minor rescue of oral antigen-specific Tregs compared to MHCII ^{Δ ROR γ} BMC mice. This could be evidence that other ROR γ -lineage APCs also play a role in the long-term maintenance of oral tolerance. In our scRNA-seq analysis, R-eTAC1 and R-eTAC3 have high levels of Aire expression while R-eTAC2 have low Aire expression—thus, R-eTAC2 could be partially spared from DT treatment ablation in AireDTR⁺ mice. Further investigation is required to decipher which R-eTAC subpopulation is responsible for oral tolerance using targeted, subset-specific deletion mice. As MHCII ^{Δ ROR γ} + *Il7r* ^{$-/-$} mixed BMC mice – with only functional R-eTACs but not ILC3s/LTis – rescued food-specific Treg differentiation, this suggests that antigen presentation by ILC3/LTis are dispensable for inducing oral tolerance in the gut. The minor Treg rescue in MHCII ^{Δ ROR γ} + Aire-DTR mixed BMC compared to MHCII ^{Δ ROR γ} BMC mice also suggests that other ROR γ -lineage APCs could help induce oral antigen-specific Tregs, such as the minor populations of DCs found in our scRNA-seq data. However, findings from other groups utilising the MHCII-ON ^{Δ Clec9a} mouse suggest that conventional DCs alone are not capable of inducing food antigen-specific Tregs³⁸. It

is also important to note that transient deletion of R-eTACs in adult AireDTR mice could affect oral tolerance differently than MHCII^{ΔROR γ t} mice, which constitutively lack antigen-presenting function in all ROR γ t-lineage APCs since conception. Studies suggest the peri-weaning period (~3 weeks) is a critical window of opportunity for the establishment and lifelong maintenance of oral tolerance via Treg differentiation¹¹⁸. Therefore, future studies are required to determine how the consequences of R-eTAC deletion during the weaning period may differ from adulthood.

In conclusion, we highlight ROR γ t eTACs as myeloid-derived mediators of gut homeostasis, modulating dietary antigen-specific T cell responses toward tolerance instead of inflammation. Therapeutically, understanding the biology of tolerogenic ROR γ t-lineage APC populations is of significant interest. Indeed, oral tolerance has been implicated as an alternative prevention or treatment for several autoimmune diseases, including experimental autoimmune encephalomyelitis and type I diabetes^{119,120}. Defining and ultimately shaping the biology of ROR γ t eTACs could thus provide a critical pathway to modulate systemic immunity across a range of disease states.

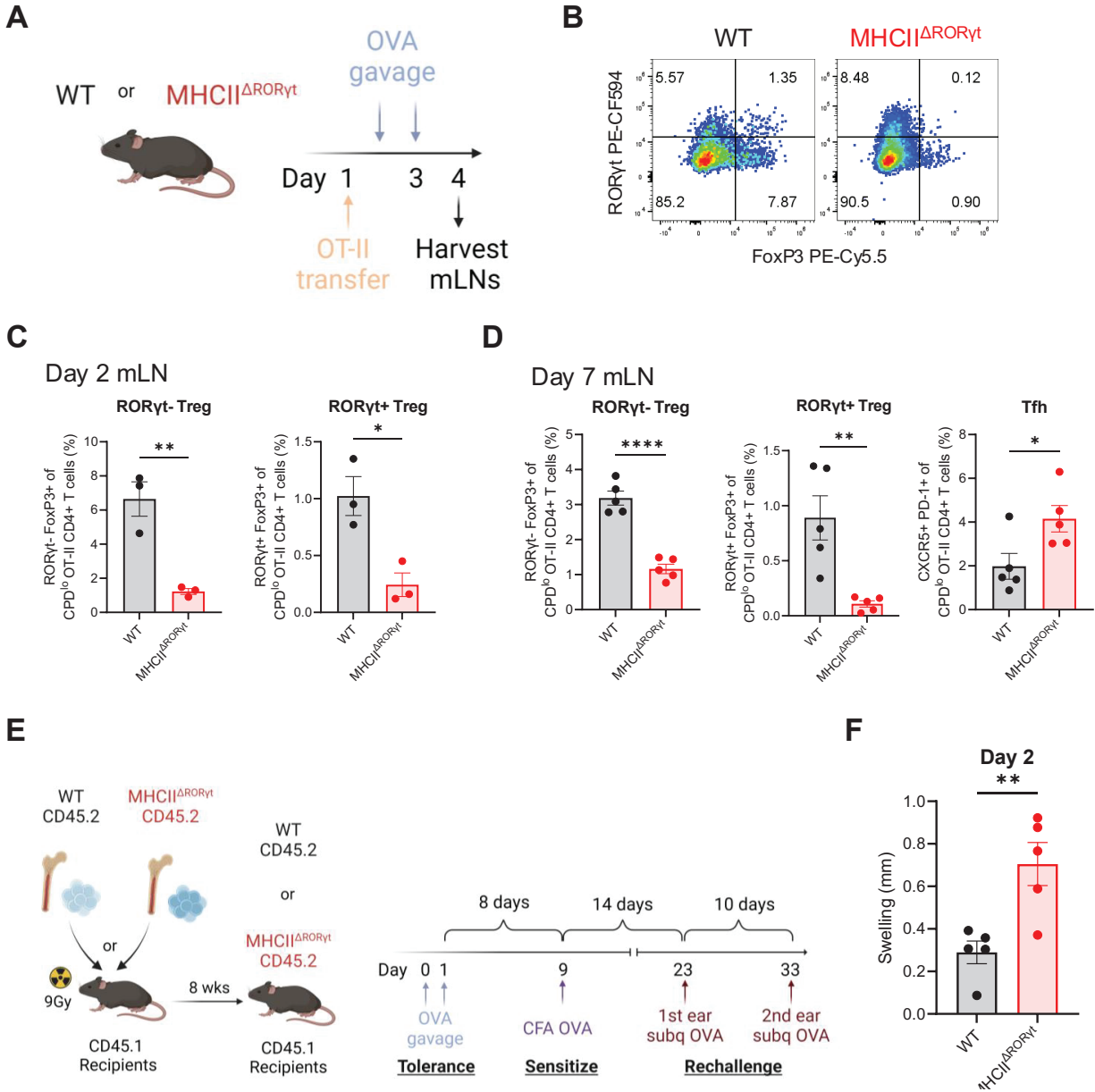


Figure 3-1: RORγt-lineage APCs mediate oral tolerance to food antigen.

(A) Experimental design of OT-II adoptive transfer and OVA oral gavage. (B) Flow cytometry and (C) quantification of CPD^{lo} OT-II in $MHCII^{\Delta ROR\gamma t}$ and WT mLN after 2 days of oral OVA ($n = 3$ in each group). (D) Quantification of CPD^{lo} OT-II in $MHCII^{\Delta ROR\gamma t}$ and WT mLN after 7 days of oral OVA ($n = 5$ in each group). (E) Experimental design of BMC generation and DTH model with OVA. (F) Ear swelling measurement at 48hr after the second subcutaneous injection of OVA in the ear ($n = 5$). Data in C, D, and F is representative of 2 independent experiments. Error bars: mean \pm s.e.m.; statistics were calculated by unpaired two-sided t -test; * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$, **** $P < 0.001$.

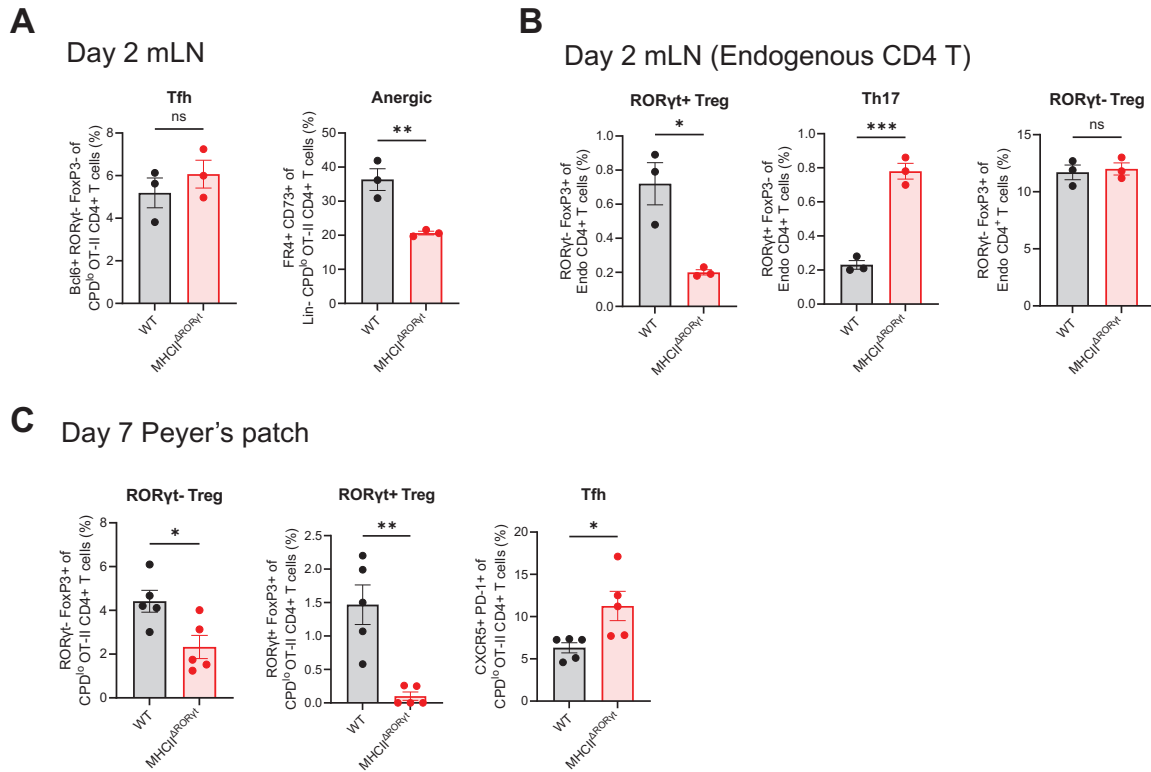


Figure 3-2: Induction of oral antigen-specific T cells.

(A) Quantification of Th17, Bcl6+, and FR4+ CD73+ OT-II in mLN after 2 days of OVA gavage ($n = 3$ from each group). (B) Quantification of RORγt+ Treg and Th17 from endogenous CD4 T cell subsets in mLN after 2 days of OVA gavage ($n = 3$ from each group). (C) Quantification of Peyer's patch RORγt- Treg, and RORγt+ Treg, and Tfh 7 days after OVA gavage ($n = 5$ from each group). Data is representative of 2 independent experiments. Error bars: mean \pm s.e.m.; statistics were calculated by unpaired two-sided t -test; * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$.



Figure 3-3

Figure 3-3: Harmonizing the diversity of RORyt-lineage APC populations.

(A) Reduced dimensionality representation of scRNA-seq data with annotated cell clusters in adult mice. (B) Gene expression of *Aire* and *Rorc*. (C) Differential gene expression heatmap of R-eTAC subsets. (D) CellTypist projection of our 2-week-old scRNA-seq data onto our adult scRNA-seq data. (E) CellTypist projection of Akagbosu *et al.* and Lyu *et al.* datasets onto our adult scRNA-seq data. (F) UMAP of the integrated adult and early life RORyt-lineage APC data with the Lyu *et al.*, Akagbosu *et al.*, Wang *et al.*, and Kedmi *et al.* datasets using scVI (left). Gene expression of *Aire* and *Rorc* (right).

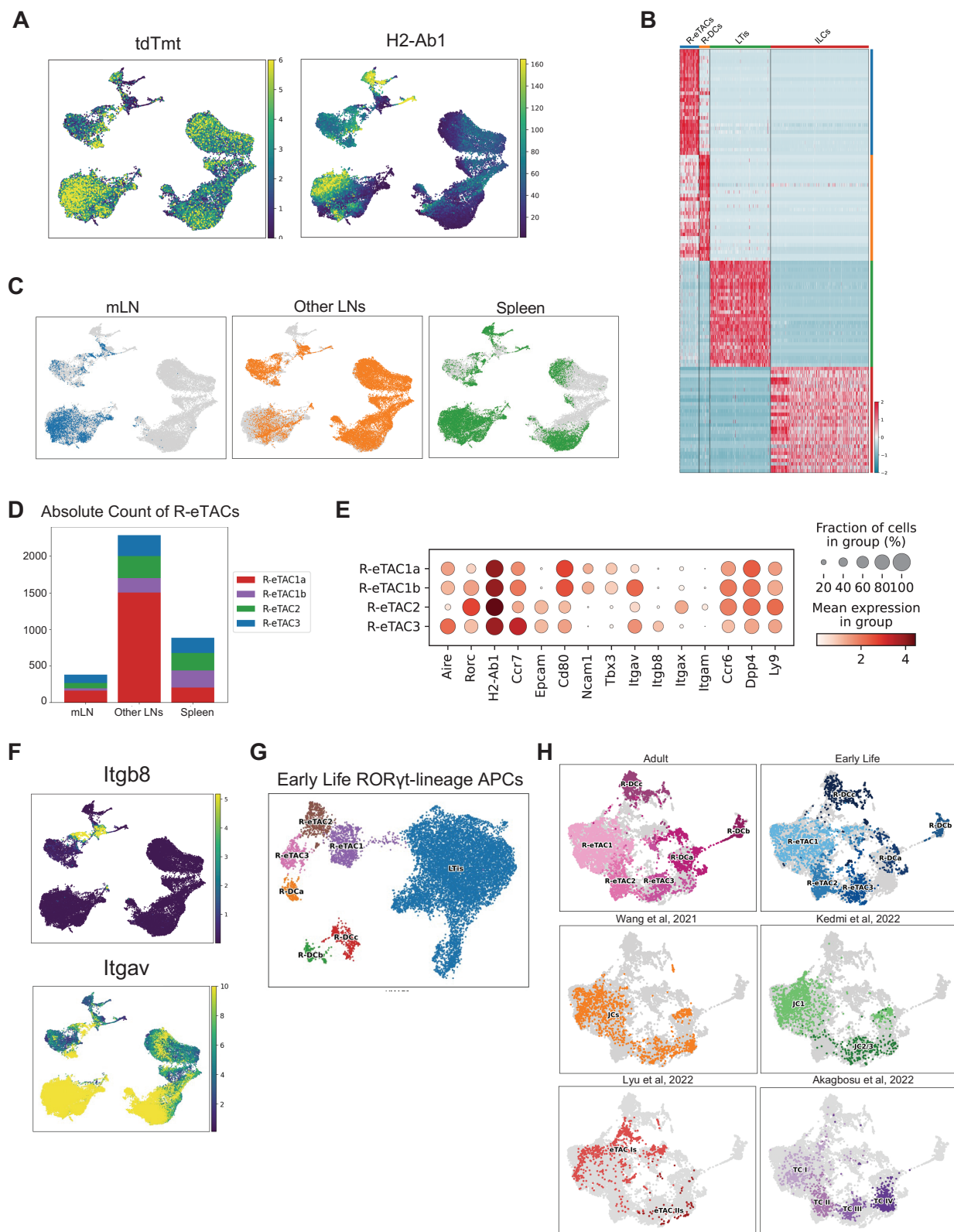


Figure 3-4

Figure 3-4: RORyt-lineage APCs by scRNA-seq.

(A) Expression of *tdTmt* and *H2-Ab1* in scRNA-seq clusters. (B) Heatmap with differential gene expression of cell clusters from scRNA-seq data. (C) Distribution of R-eTAC subset clusters found from each sample origin. (D) Number of cells among R-eTAC subset clusters found in each sample origin. (E) Select gene expression in R-eTAC subsets. (F) Expression of *Itgb8* and *Itgav* in our scRNA-seq dataset. (G) Reduced dimensionality representation of scRNA-seq data with annotated cell clusters in 2-week-old mice. (H) Integrated UMAP representation indicating R-eTAC equivalent cell clusters from each dataset.

Figure 3-5: RORyt eTACs are antigen presenting cells of myeloid lineage.

(A) Flow cytometry of intracellular Aire staining between WT and germline Aire-KO mice. (B) Flow cytometry from REALTAR mouse lymph nodes pre-gated on single, live, Lin⁻ cells. (C-D) Distribution of R-eTAC subsets found in different SLOs by (C) scRNA-seq and by (D) flow cytometry from REALTAR mice (right, $n = 6$). (E) Gene expression of select markers from scRNA-seq. (F) Flow cytometry plot of Aire⁺ cells showing gating of R-eTAC1 and R-eTAC2/3 subsets based on CCR7 and EpCAM expression. (G) Flow cytometry plots (left) and quantification (right) of GFP⁺ tdTmt⁺ R-eTACs versus DCs co-cultured with naïve OT-II cells after 72hr. (H) Quantification of ILCs, DCs, and intracellular Aire⁺ cells from lymph nodes of WT+*Il7r*^{-/-} mixed BMC mice ($n = 5$ from each group). (I) Quantification of tdTmt⁺ cells in *Cx3cr1*^{CreER}×*Rosa26-LSL-tdTmt* mouse lymph nodes ($n = 3$). Data in A, B, F and H are representative of 3 independent experiments. Data in D is representative of 2 independent experiments. Data in G and I is pooled from 3 independent experiments. Error bars: mean ± s.e.m.; statistics were calculated by one-way ANOVA with Tukey's multiple comparisons test (D); * $P < 0.05$, ** $P < 0.005$, **** $P < 0.0001$.

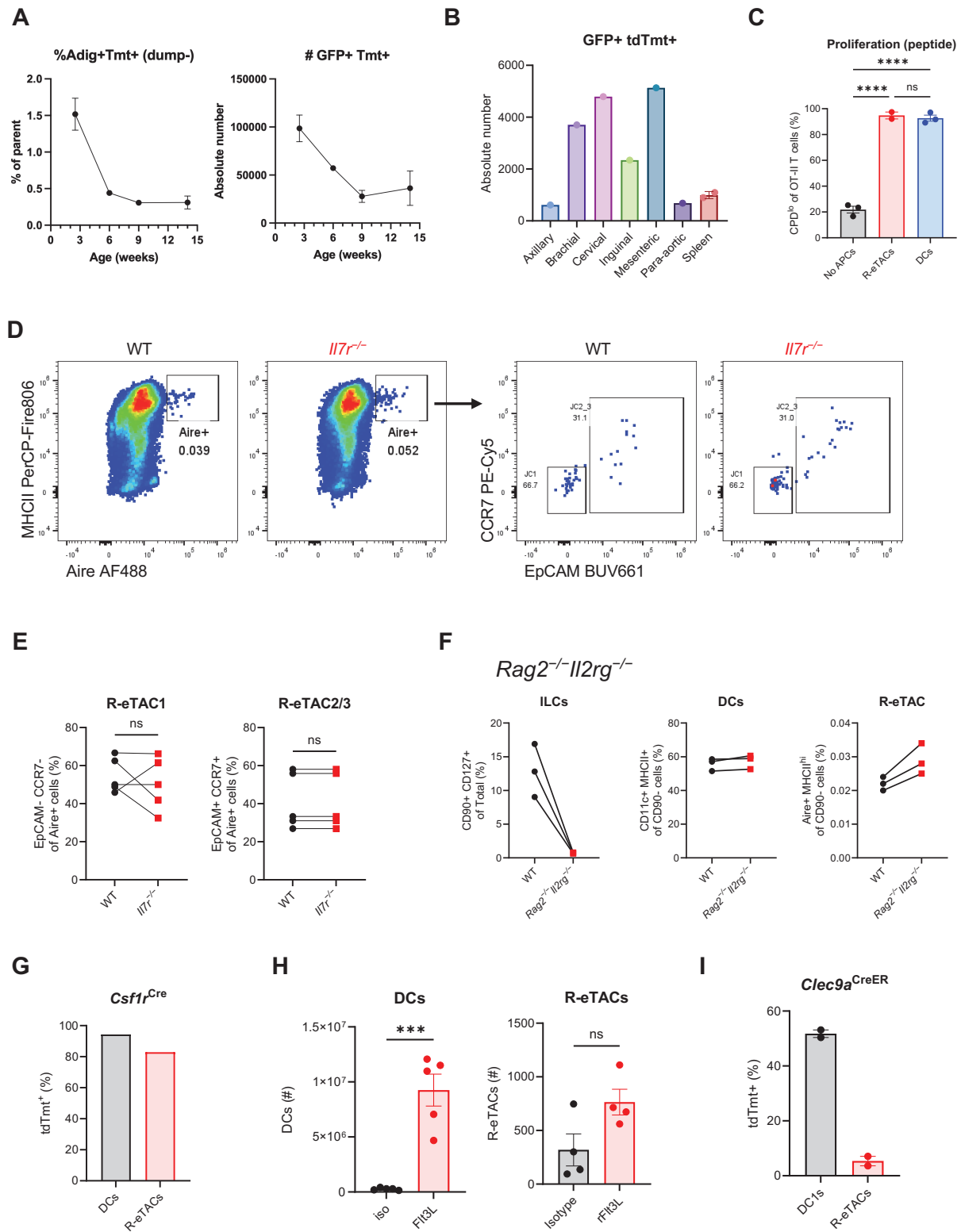


Figure 3-6

Figure 3-6: RORYt eTACs are a distinct myeloid population.

(A) Flow cytometry quantification of GFP⁺ Tmt⁺ cells from REALTAR mice lymph nodes at various ages, gated on single, live, Lin⁻ cells ($n = 3$). (B) Flow cytometry quantification of GFP⁺ Tmt⁺ cells from various lymph node locations of REALTAR mice, gated on single, live, Lin⁻ cells ($n = 6$). (C) Quantification of proliferation of OT-II cells co-cultured with GFP⁺ tdTmt⁺ R-eTACs versus DCs in the presence of OVA-II peptide after 72hr in Fig 3-5G. (D) Flow cytometry and (E) quantification of Aire⁺ cells and R-eTAC subsets from *Il7r^{-/-}* or WT BM-derived portion ($n = 5$). (F) Quantification of ILCs, DCs, and intracellular Aire⁺ cells from lymph nodes of mixed BMC mice ($n = 3$). (G) Flow cytometry quantification of Aire⁺ eTACs and DCs from *Csf1r^{Cre}* × Rosa26-LSL-tdTmt mice lymph nodes ($n = 1$). (H) Quantification of DCs and Aire⁺ cells from lymph nodes of mice treated with isotype (Iso) or recombinant Flt3L (rFlt3L) ($n = 5$ from each group). (I) Flow cytometry quantification of Aire⁺ eTACs and DCs from *Clec9a^{CreER}* × Rosa26-LSL-tdTmt mice lymph nodes ($n = 2$). Data in B is representative of 2 independent experiments. Data in C is pooled from 3 independent experiments. Data in D and E is representative of 3 independent experiments. Data in F and H is representative of 2 independent experiments. Data in A, G and I is from 1 experiment. Error bars: mean \pm s.e.m.; statistics were calculated by one-way ANOVA with Tukey's multiple comparisons test (C) or unpaired two-sided *t*-test (H); *** $P < 0.0005$, **** $P < 0.0001$.

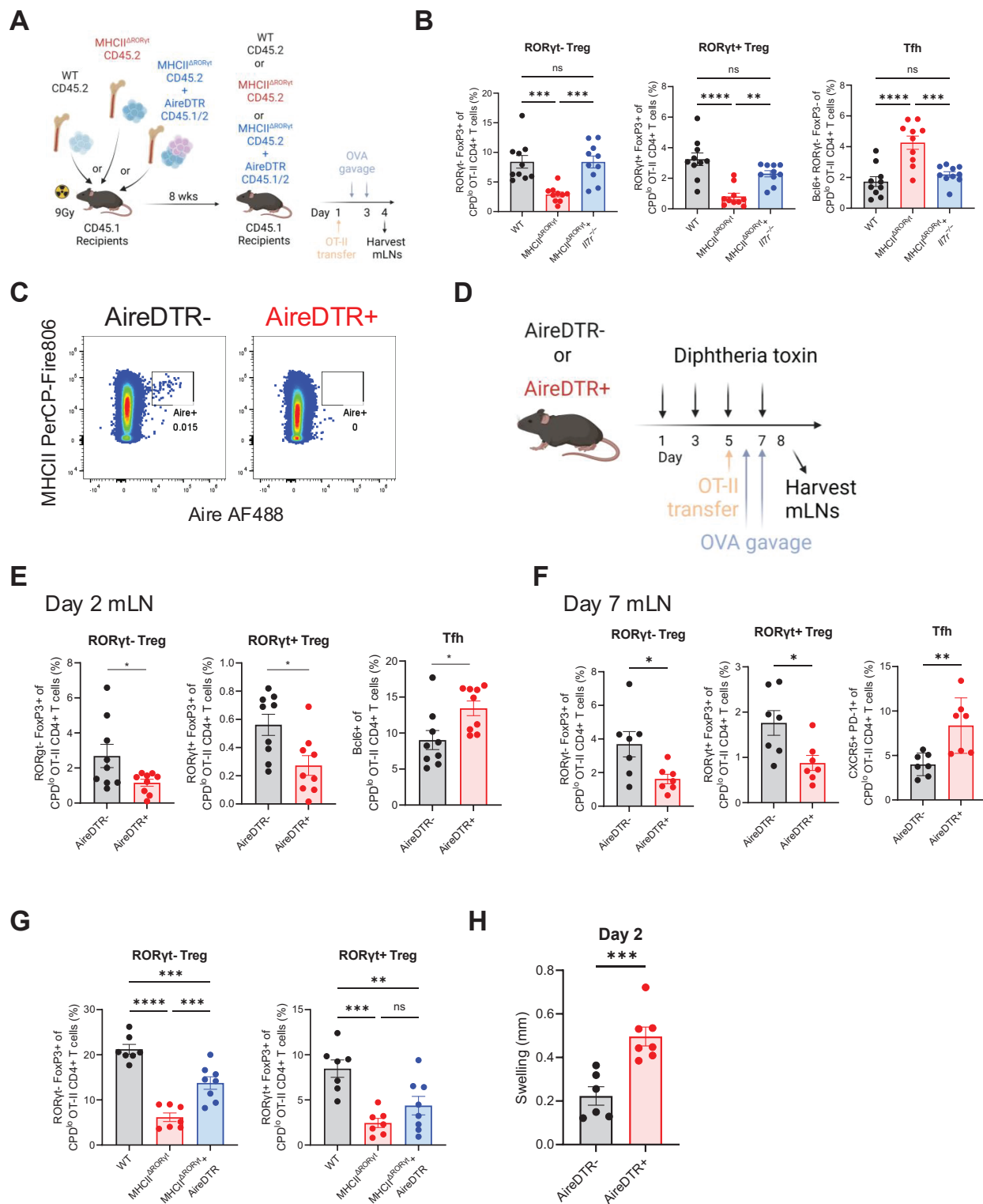


Figure 3-7

Figure 3-7: ROR γ t eTACs mediate oral tolerance to food antigen.

(A) Experimental design of mixed BMC generation with WT, MHCII Δ ROR γ t, and *I17r^{-/-}* mice and OVA oral gavage. (B) Flow cytometry quantification of CPD^{lo} OT-II in BMC mice mLN after 2 days of oral OVA ($n = 10$ from each group). (C) Flow cytometry of cells pre-gated on single, live, Lin⁻ from AireDTR⁻ and AireDTR⁺ mouse lymph nodes. (D) Treatment regimen for DT, OVA gavage and OT-II transfer. (E-F) Quantification of OT-II T cells from AireDTR⁻ and AireDTR⁺ mLN by flow cytometry after (E) 2 days ($n = 9$ from each group) and (F) 7 days ($n = 7$ from each group) of OVA gavage. (G) Quantification of OT-II T cells from mixed BMC generated from WT, MHCII Δ ROR γ t, and AireDTR mLN after 2 days of OVA gavage ($n = 7$ for WT and MHCII Δ ROR γ t, $n = 8$ for MHCII Δ ROR γ t+AireDTR). (H) Ear swelling measurement 2 days after the second subcutaneous injection of OVA in the ear. Data in B, C, and E is pooled from 2 independent experiments. Data in F and H is representative of 2 independent experiments. Data in G is from 1 experiment. Error bars: mean \pm s.e.m.; statistics were calculated by one-way ANOVA with Tukey's multiple comparisons test (B and G) and unpaired two-sided *t*-test (E, F, and H); * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$, **** $P < 0.0001$.

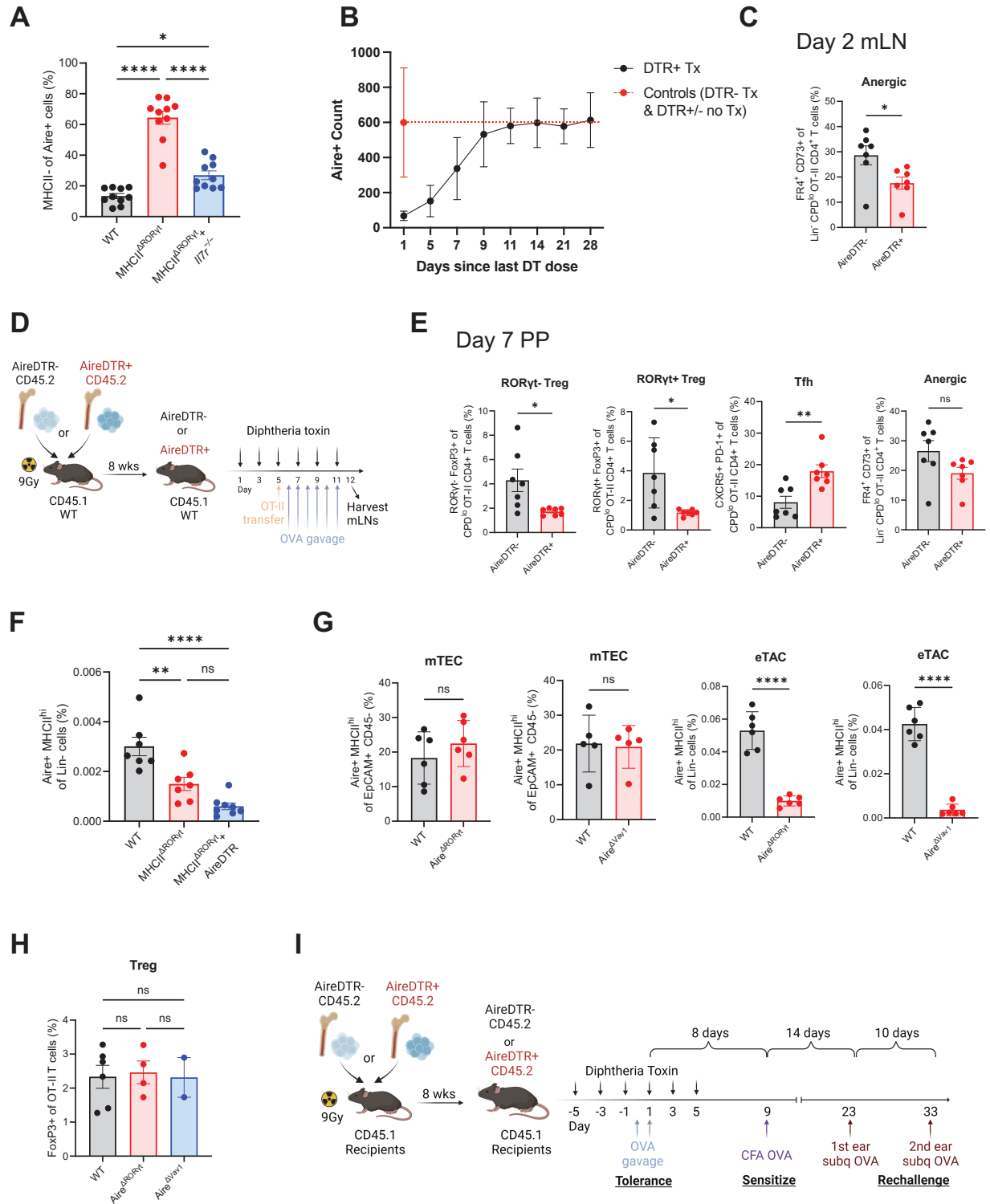


Figure 3-8

Figure 3-8: RORyt eTACs are essential for oral tolerance.

(A) Validation of MHCII expression between WT, MHCII^{ΔRORyt}, and MHCII^{ΔRORyt} + *Il7r*^{-/-} mixed BMC experiment ($n = 10$ from each group). (B) Quantification of the repopulation kinetics of Aire⁺ cells post DT treatment ($n = 10$ for controls (red), $n = 2-4$ for each experimental group timepoint (black)). (C) Quantification of FR4⁺ CD73⁺ OT-II CD4⁺ T cells frequency from mLN after 2 days of OVA gavage ($n = 7$ from each group). (D) Extended treatment regimen for DT, OVA gavage and OT-II transfer. (E) Quantification of RORyt⁻ Treg, and RORyt⁺ Treg, Tfh, and anergic CD4 T cells 7 days after OVA gavage in PP ($n = 7$ from each group). (F) Validation of MHCII expression between WT, MHCII^{ΔRORyt}, and MHCII^{ΔRORyt} + AireDTR mixed BMC experiment ($n = 7$ from WT and MHCII^{ΔRORyt}, $n = 8$ for MHCII^{ΔRORyt} + AireDTR). (G) Quantification of Aire⁺ cells from the thymus (top) and SLOs (bottom) in Aire^{ΔRORyt}, Aire^{ΔVav1}, or WT mice ($n = 6$ from each group). (H) Flow cytometry quantification of OT-II T cells from Aire^{ΔRORyt}, Aire^{ΔVav1}, or WT mouse mLN after 2 days of OVA gavage ($n = 6$ for WT, $n = 4$ for Aire^{ΔRORyt}, $n = 2$ for Aire^{ΔVav1}). (I) Experimental design of BMC generation with WT and AireDTR, and DTH model with OVA. Data in A, C, and F is pooled from 2 independent experiments. Data in B is pooled from 4 independent experiments. Data in E is representative of 2 independent experiments. Data in G and H is representative of 3 independent experiments. Error bars: mean \pm s.e.m.; statistics were calculated by one-way ANOVA with Tukey's multiple comparisons test (A, F, and H) and unpaired two-sided *t*-test (C, E, and G); * $P < 0.05$, ** $P < 0.005$, **** $P < 0.0001$.

3.5 Materials and Methods

Mice

Six- to eight-week-old male or female mice were used for all the experiments in this study. C57BL/6J (000664), Rosa26-LSL-dtTmt (007914), Aire-flox (031409), MHCII-flox (037709), *Roryt*^{Cre} (022791), *Vav1*^{iCre} (008610), *Cx3cr1*^{CreER} (021160), and *Clec9a*^{CreER} (037804) were purchased from The Jackson Laboratory. Adig, AireDTR, and germline Aire-KO mice were generated and provided by the laboratory of Dr. Mark S. Anderson and were characterized as previously described^{8,37,79}. *Il7r*^{-/-} and *Rag2*^{-/-}*Il2rg*^{-/-} mice were generously provided by Dr. Roberto R. Ricardo-Gonzalez (UCSF). *Csf1r*^{Cre} mice were generously provided by Dr. Matthew Spitzer (UCSF). Lineage tracing mice were generating by crossing lineage-specific Cre with Rosa26-LSL-dtTmt. All mice were housed in specific pathogen-free facilities at the University of California San Francisco, and all animal studies were approved by the Institutional Committee on Animal Use and Care at the University of California San Francisco.

Mouse lymph node processing and purification

For single-cell RNA sequencing, mouse spleen, mesenteric, and other lymph nodes (cervical, brachial, axillary, inguinal, and para-aortic) were procured and pooled in digestion medium consisting of RPMI 1640 with 2% foetal bovine serum (FBS) (Sigma-Aldrich), with deoxyribonuclease (DNase) (100 µg/mL; Roche) and LiberaseTM (96 µg/mL; Roche), minced and agitated at 37°C with gentle agitation for 30 min, and passed through a 70-micron filter. Cells were then sorted for live, tdTmt+ MHCII+ cells and processed for 10× single-cell analysis.

For eTAC characterization by flow cytometry, digested cells from lymph nodes were resuspended for magnetic column enrichment (Miltenyi LS column depletion with anti-biotin microbeads and biotinylated antibodies against B220, Ter119, TCRβ, ± CD3). Cell viability and counts were evaluated with Vi-CELL XR (Beckman Coulter). For OTI/II characterization, mesenteric lymph nodes were procured, and cells were passed through a 70-micron cell strainer to achieve single cell suspension for subsequent flow cytometry staining and analysis.

Antibodies and reagents

Fluorochrome- or biotin-conjugated antibodies against the following mouse targets were purchased from BD Biosciences: CD3e (145-2C11), B220 (RA3-6B2), Ly6G (1A8), EpCAM(G8.8), CD80 (16-10A1), TCR β (H57-5970), MHCII (AF6-120.1), CD8a (53-6.7), Bcl6 (K112-91), CD56 (809220), ROR γ t (Q31-378); ThermoFisher Scientific: Aire (5H12), SiglecF (1RNM44N), TCR β (H57-597), CD3e (145-2C11), B220 (RA3-6B2), Ter119 (TER-119), CD19 (1D3), Thy1.2 (30-H12), FoxP3 (FJK-16s); and BioLegend: CD4 (GK1.5), CD8 (53-6.7), NK1.1 (PK136), Ly6C (HK1.4), CD90.2 (30-H12), CD19 (6D5), F4/80 (BM8), CCR7 (4B12), XCR1 (ZET), SIRP α (P84), TCR β (GL3), CD45 (30- F11), CD200 (OX-90), CCR6 (29-2 L17), CD73 (TY/11.8), FR4 (12A5), CD44 (IM7), CD95 (SA367H8), CD45.2 (104), CD45.1 (A20), CXCR5 (L138D7), PD1 (29F.1A12). Fc block (2.4G2) was purchased from BioXCell. Isotype control antibodies were purchased from the same source as target-specific antibodies. The Zombie NIR Fixable Viability Kit from Biolegend was used for viability staining for flow cytometry. DAPI (4',6-diamidino-2-phenylindole) was used for viability staining for cell sorting.

Flow cytometry staining

CCR7 antibody staining was done at 37°C for 20 min prior to live/dead staining (10 min, room temperature). Then, surface antibodies and Fc block were incubated with cells for 30 min at 4°C. For intracellular staining, cells were fixed and permeabilized with eBioscience Foxp3/Transcription Factor Fixation/Permeabilization solution, then washed with 1 \times Permeabilization Buffer before FoxP3 antibody incubation. For intracellular Aire staining, cells were then fixed in 4% PFA for 15 min at RT, then incubated with eBioscience Foxp3/Transcription Factor Fixation/Permeabilization solution for 30 min at 4°C. Cells were washed and resuspended in 1 \times Permeabilization Buffer overnight at 4°C. The next day, intracellular antibody staining was incubated with cells in 1 \times Permeabilization Buffer for 1 hour at RT. Finally, cells were washed three times in 1 \times Permeabilization Buffer and resuspended in the staining buffer for flow cytometry.

Flow cytometry data were acquired on Cytex Aurora and analysed using FlowJo (TreeStar v10.10.0). For eTACs gating, Lin⁻ gate includes (SiglecF⁻ Ly6G⁻ NK1.1⁻ TCRb⁻ B220⁻ F4/80⁻).

Bone marrow chimaera

CD45.1 BM recipients were irradiated with two doses of 450 rads (X-RAD 320, Precision X-Ray) 4 hours apart before receiving $5-10 \times 10^6$ of donor bone marrow cells by retro-orbital injection. Mice were allowed 6-8 weeks for reconstitution before experimental use.

Adoptive T cell transfer

Single cell suspension of lymphocytes was obtained from mashing lymph nodes (inguinal, axillary, brachial) and spleen from TCR transgenic mice through a 70-micron cell strainer. Naive CD4 T cells were isolated from OTII TCR transgenic mice using the Naive CD4 T Cell Isolation Kit (Miltényi Biotec, 130-104-453). After magnetic column isolation, naive T cells were stained with eBioscience Cell Proliferation Dye eFluor 450 (ThermoFisher, 65-0842-90) at 10 μ M for 10 minutes at 37°C, then quenched with RPMI with 2% FBS. Labelled T cells were resuspended in HBSS without Ca²⁺ and Mg²⁺ before transferring into recipient mice by retro-orbital injections at $0.5-1 \times 10^6$ cells in 200 μ L HBSS per mouse.

Oral OVA administration

OVA (grade III, Sigma-Aldrich, A5378) was administered through oral gavage using plastic gavage needles. For short-term oral tolerance experiments, OVA was gavaged 1 and 2 days after naive OTI/II adoptive transfer, at 50 mg in 200 μ L PBS per mouse each day, and mice were sacrificed on day 3. For long-term oral OVA challenge, 50 mg in 200 μ L PBS OVA was administered on days 1 and 2 after naive OTI/II transfer, and subsequently 25 mg in 100 μ L PBS OVA on days 3-6 until experiment endpoint on day 7.

CFA-OVA immunization and subcutaneous ear challenge

For CFA-OVA immunization, CFA (BD, 263810) and OVA in PBS were emulsified prior to injection. 8 days after oral OVA administration, each mouse received 100 μ L CFA and 300 μ g endotoxin-free OVA (InvivoGen, vac-pova-100) dissolved in 100 μ L PBS through subcutaneous injection

between the shoulder blades. Then, 14 and 24 days after CFA-OVA immunization, mice were challenged with 30 μ L of OVA (2.5 mg/mL endotoxin free, dissolved in PBS) subcutaneously on one ear and PBS control on the other ear. Ear thickness was measured daily from the pre-challenge until the experimental endpoint with a digital micrometre (Mitutoyo) and swelling was calculated as follows: [(OVA post-challenge) - (OVA pre-challenge)] - [(PBS post-challenge) - (PBS pre-challenge)]

Diphtheria Toxin administration

Mice were given Diphtheria Toxin (Enzo, BML-G135-0001) at 0.75 μ g in 100 μ L PBS by intraperitoneal injection on days -5, -3, -1, 1, 3, 5 days while OVA was administered orally on days 1 and 2.

Tamoxifen treatment

For induction of gene recombination in *Cx3cr1*^{CreER} mice, 30 mg/mL tamoxifen (Sigma-Aldrich, T5648) was dissolved in corn oil and administered 100 μ L via oral gavage 3 times a week for 4 weeks for complete labelling. Mice were harvested the day after the last tamoxifen treatment. For induction of gene recombination in *Clec9a*^{CreER} mice, mice were put on tamoxifen-containing chow (Inotiv, TD.130860) for a week before analysis.

Recombinant Flt3L treatment

Wildtype C57BL/6J mice were treated with 10 μ g of either isotype (BioXCell, BE0096) or recombinant Flt3L (BioXCell, BE0098) for 7 days. Mice were sacrificed the next day, and lymph nodes were harvested for subsequent flow analysis.

Single-cell RNA sequencing

scRNA-seq libraries were generated from pooled spleen, mesenteric lymph node, and other lymph node samples. Early life samples were processed using 10x Genomics Chromium 3' sequencing platform according to manufacturer protocols. The adult samples were generated using the 10x Genomics GEM-X 3' sequencing platform according to manufacturer protocols. The resulting libraries were sequenced on the Illumina NovaSeq 6000.

scRNA-seq analyses

Early Life Sample: Raw sequencing reads were aligned to the mm10 reference genome using the standard Cell Ranger pipeline¹²¹. Cells that were not transcript positive for tdTomato and MHCII were excluded from downstream analysis. Cells were filtered out if they met any of the following criteria – greater than 5% mitochondrial reads, greater than 5,500 unique genes, and greater than 35,000 total reads.

Adult Sample: Raw sequencing reads were aligned to the mm10 reference genome using the standard Cell Ranger pipeline¹²¹. Cells were filtered out if they met any of the following criteria – greater than 5% mitochondrial reads, greater than 8,000 unique genes, and greater than 60,000 total reads.

Analysis: The early life and adult samples were separately analysed in python using Scanpy package and scVI¹²² integration methodology. Briefly, each tissue sample (spleen, mLN and other LNs) was restricted to the top jointly highly variable genes and combined. In order to impute gene expression in our combined dataset, a scVI model was trained on the combined data using tissue as the batch identifier, ribosomal gene percentage as a continuous covariate and the negative binomial distribution of gene likelihood. Principle components and a UMAP embedding were calculated from the scVI latent space. Clusters were identified using the Leiden algorithm. Cell types were identified based on expression of known marker genes. Three unique populations of putative R-eTACs were identified and termed R-eTAC1-3. Three unique populations of lineage-traced DCs were also identified and termed R-DCa, R-DCb and R-DCc. R-DC, ILC and LTi subclusters were group together for ease of visualization. Non ILC, LTi, R-DC or R-eTAC clusters are not shown. Batched corrected differentially expressed genes were called using scVI with a false discovery rate of 0.05. For heatmaps, units are z-scored gene-level expression and for feature plots, the units are scVI normalized expression.

RORyt Lineage Landscape Integration

The Lyu *et al.* (GSE184175), Kedmi *et al.* (GSE200147) and Akagbosu *et al.* (GSE174405) datasets were downloaded from the Gene Expression Omnibus. The Wang *et al.* (GSE176282) data was taken from our previous publications. For ease of integration, the integration was restricted to R-eTACs, R-DCs and equivalent populations. Integration was performed in python using methods detailed above, except the scVI model was trained using data set source as a batch identifier and both MT gene percentage and ribosomal gene percentage were supplied as continuous covariates. Dimensionality reduction, clustering and cell type identification were performed in the same fashion as above. To validate our integration, CellTypist¹²³ was used to project cell type labels from Lyu *et al.* and Akagbosu *et al.* and our early life dataset to our adult dataset.

APC and OTII Co-culture

Myeloid cells were obtained from digested lymph nodes and spleen and isolated using magnetic column depletion from REALTAR mice (see digestion and purification method above). Lymphocyte-depleted samples were then FACS sorted on single, single, live, CD19- CD90.2-, then GFP+ tdTmt+ (R-eTACs), as well as GFP- CD11c+ (DCs). 3500-6500 APCs were then co-cultured with CPD-labelled naïve OT-II T cells at a 1:10 ratio in the presence of either 100 µg/mL whole OVA protein or OVA³²³⁻³³⁹ peptide (Anaspec, AS-27024) for 72 hours in a U-bottom 96 well plate. Culture media was made of RPMI with 10% FBS, 1× Pen/Strep/Glutamine, NEAA, HEPES, and β-mercaptoethanol.

Statistical analyses and visualization

All data points are shown in the graphs as mean value indicated and individual data points as shown. All statistical analyses were performed using GraphPad Prism (v10.2.3). Statistical significance was calculated using one-way analysis of variance (ANOVA) with Tukey multiple comparison or the unpaired two-sided *t* test. **P* < 0.05, ***P* < 0.005, ****P* < 0.0005, and *****P* <

0.0001 were considered significant. Experimental graphics were generated with BioRender (<https://biorender.com/>).

Chapter 4 Closing and Future Directions

We have used single cell multiomics to identify and define the population of Janus cells that express Aire and ROR γ t. Aire-expressing cells in the thymus has been thoroughly studied for their role in T cell negative selection and is a crucial part of immune tolerance to prevent autoimmunity. Janus cells are found outside of the thymus with high transcriptional and genomic homology to Aire-expressing stromal cells in the thymus. This could suggest these divergent cell populations are using a similar transcriptional program to promote immune tolerance throughout the body and throughout life. This not only advances our knowledge on cell types carrying out immune tolerance against self and other innocuous non-self-antigens, but also provides insights into potential transcriptional programs behind the development of tolerogenic antigen presenting cells.

Despite elucidating the identity of JCs and understanding their transcriptional program, the specific cell-intrinsic role of Aire in eTACs is yet to be define. The cell-intrinsic functions of Aire in eTACs have been controversial largely due to the lack of tools to specifically target peripheral Aire without affecting thymic Aire expression. To address this issue, we have generated and validated our Vav1 Aire Knockout (VAKO) mouse (Vav1-Cre x Aire^{fl/fl}) to selectively delete Aire in all eTACs and ROR γ t Aire knockout (RAKO) mouse (ROR γ t-Cre Aire^{fl/fl}) to specifically delete Aire in ROR γ t⁺ eTACs. Although Aire is known as a transcriptional regulator of TSA expression in mTECs, some recent data has suggested its role in chromatin modification may be paradoxically suppressive⁵⁷. Using these conditional Aire KO mice, further scRNA-seq and scATAC-seq studies would need to be performed on eTACs to understand the gene expression and chromatin accessibility regulated by Aire in eTACs. To determine the functional roles of extrathymic Aire in the induction of T-cell tolerance and the maintenance of normal immune homeostasis, additional functional studies are needed to compare between RAKO and VAKO mice with control counterparts. Another interesting next step would also be to understand the role

of ROR γ t in the development of ROR γ t⁺ eTACs and whether ROR γ t promotes a tolerogenic transcriptional program in APCs.

Our studies have defined three distinct subsets of ROR γ t⁺ eTACs, but the lineage relationship between these subsets is yet to be defined. Because we still see residual Aire expression in eTACs from RAKO mice, it suggests that Aire expression precedes ROR γ t expression. RNA velocity on scRNA-seq data could provide insight into lineage relationship as well as barcoding studies which would inform lineage relationships in an unbiased manner.

Since our studies suggest that ROR γ t⁺ eTACs are responsible for inducing oral antigen specific Tregs, further studies would be required to understand which subset is important or if all three subsets are indispensable in inducing oral tolerance. More mouse genetic tools are required to delete specific subsets and determine their roles separately.

Taken together, we are deciphering a key population of antigen presenting cells that are potentially responsible for T cell tolerance and the prevention of aberrant inflammation. Understanding their transcriptional, genomic, and functional biology will help elucidate mechanisms of immune tolerance as well as potential disease mechanisms. These studies could help provide insights into the development of treatments for various autoimmune and inflammatory diseases as well as cancer and maternal-fetal tolerance.

References

1. Klein, L. Recollections of the discovery of promiscuous antigen expression in mTECs. *Nat. Immunol.* **21**, 1303–1305 (2020).
2. Smith, K. M., Olson, D. C., Hirose, R. & Hanahan, D. Pancreatic gene expression in rare cells of thymic medulla: Evidence for functional contribution to T cell tolerance. *Int. Immunol.* **9**, 1355–1365 (1997).
3. Derbinski, J., Schulte, A., Kyewski, B. & Klein, L. Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. *Nat. Immunol.* **2**, 1032–1039 (2001).
4. Hanahan, D. Peripheral-antigen-expressing cells in thymic medulla: Factors in self-tolerance acid autoimmunity. *Curr. Opin. Immunol.* **10**, 656–662 (1998).
5. Nagamine, K. *et al.* Positional cloning of the APECED gene. *Nat. Genet.* (1997) doi:10.1038/ng1297-393.
6. Aaltonen, J. *et al.* An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. *Nat. Genet.* **17**, 399–403 (1997).
7. Kogawa, K. *et al.* Expression of AIRE gene in peripheral monocyte/dendritic cell lineage. *Immunol. Lett.* **80**, 195–198 (2002).
8. Gardner, J. M. *et al.* Deletional Tolerance Mediated by Extrathymic Aire-Expressing Cells. *Science* (80-.). **321**, 843–847 (2008).
9. Gardner, J. M. *et al.* Extrathymic aire-expressing cells are a distinct bone marrow-derived population that induce functional inactivation of CD4+ T cells. *Immunity* (2013) doi:10.1016/j.immuni.2013.08.005.
10. Lee, J. W. *et al.* Peripheral antigen display by lymph node stroma promotes T cell tolerance to intestinal self. *Nat. Immunol.* **8**, 181–190 (2007).
11. Grupillo, M. *et al.* Essential roles of insulin expression in Aire + tolerogenic dendritic cells in maintaining peripheral self-tolerance of islet β -cells. *Cell. Immunol.* (2012)

doi:10.1016/j.cellimm.2011.12.010.

12. Li, D. *et al.* Transplantation of Aire-overexpressing bone marrow-derived dendritic cells delays the onset of type 1 diabetes. *Int. Immunopharmacol.* **49**, 13–20 (2017).
13. Ko, H. J. *et al.* Transplantation of autoimmune regulator-encoding bone marrow cells delays the onset of experimental autoimmune encephalomyelitis. *Eur. J. Immunol.* **40**, 3499–3509 (2010).
14. Poliani, P. L. *et al.* Human peripheral lymphoid tissues contain autoimmune regulator-expressing dendritic cells. *Am. J. Pathol.* **176**, 1104–1112 (2010).
15. Gillis-Buck, E. *et al.* Extrathymic Aire -expressing cells support maternal-fetal tolerance. *Sci. Immunol.* **6**, eabf1968 (2021).
16. Dobeš, J. *et al.* Extrathymic expression of Aire controls the induction of effective TH17 cell-mediated immune response to *Candida albicans*. *Nat. Immunol.* (2022)
doi:10.1038/s41590-022-01247-6.
17. Scheinecker, C., McHugh, R., Shevach, E. M. & Germain, R. N. Constitutive presentation of a natural tissue autoantigen exclusively by dendritic cells in the draining lymph node. *J. Exp. Med.* **196**, 1079–1090 (2002).
18. Idoyaga, J. *et al.* Specialized role of migratory dendritic cells in peripheral tolerance induction. *J. Clin. Invest.* **123**, 844–854 (2013).
19. Baratin, M. *et al.* Homeostatic NF- κ B Signaling in Steady-State Migratory Dendritic Cells Regulates Immune Homeostasis and Tolerance. *Immunity* **42**, 627–639 (2015).
20. Vander Lugt, B. *et al.* Transcriptional determinants of tolerogenic and immunogenic states during dendritic cell maturation. *J. Cell Biol.* **216**, 779–792 (2017).
21. Lutz, M. B. & Schuler, G. Immature, semi-mature and fully mature dendritic cells: Which signals induce tolerance or immunity? *Trends Immunol.* **23**, 445–449 (2002).
22. Devi, K. S. P. & Anandasabapathy, N. The origin of DCs and capacity for immunologic tolerance in central and peripheral tissues. *Semin. Immunopathol.* **39**, 137–152 (2017).

23. Hasegawa, H. & Matsumoto, T. Mechanisms of tolerance induction by dendritic cells in vivo. *Front. Immunol.* **9**, (2018).
24. Lochner, M. *et al.* In vivo equilibrium of proinflammatory IL-17+ and regulatory IL-10+ Foxp3+ RORyt+ T cells. *J. Exp. Med.* **205**, 1381–1393 (2008).
25. Ohnmacht, C. *et al.* The microbiota regulates type 2 immunity through RORyt+ T cells. *Science (80-.)*. **349**, 989–993 (2015).
26. Sefik, E. *et al.* Individual intestinal symbionts induce a distinct population of RORyt+ regulatory T cells. *Science (80-.)*. **349**, 993–997 (2015).
27. Coombes, J. L. *et al.* A functionally specialized population of mucosal CD103+ DCs induces Foxp3+ regulatory T cells via a TGF- β -and retinoic acid-dependent mechanism. *J. Exp. Med.* **204**, 1757–1764 (2007).
28. Hepworth, M. R. *et al.* Innate lymphoid cells regulate CD4 + T-cell responses to intestinal commensal bacteria. *Nature* **498**, 113–117 (2013).
29. Hepworth, M. R. *et al.* Group 3 innate lymphoid cells mediate intestinal selection of commensal bacteria-specific CD4+ T cells. *Science (80-.)*. **348**, 1031–1035 (2015).
30. Brown, C. C. *et al.* Transcriptional Basis of Mouse and Human Dendritic Cell Heterogeneity. *Cell* **179**, 846-863.e24 (2019).
31. Yamano, T. *et al.* Aire-expressing ILC3-like cells in the lymph node display potent APC features. *J. Exp. Med.* **216**, (2019).
32. Wang, J. *et al.* Single-cell multiomics defines tolerogenic extrathymic Aire-expressing populations with unique homology to thymic epithelium. *Sci. Immunol.* **6**, (2021).
33. Kedmi, R. *et al.* A RORyt+ cell instructs gut microbiota-specific Treg cell differentiation. *Nature* **610**, 737–743 (2022).
34. Akagbosu, B. *et al.* Novel antigen-presenting cell imparts Treg-dependent tolerance to gut microbiota. *Nature* **610**, 752–760 (2022).
35. Lyu, M. *et al.* ILC3s select microbiota-specific regulatory T cells to establish tolerance in

- the gut. *Nature* (2022) doi:10.1038/s41586-022-05141-x.
36. Park, T., Leslie, C., Rudensky, A. Y. & Brown, C. C. Reconciling the spectrum of ROR γ ⁺ antigen-presenting cells. *bioRxiv* 2023.11.01.565227 (2023).
 37. Anderson, M. S. *et al.* Projection of an immunological self shadow within the thymus by the aire protein. *Science* (80-.). (2002) doi:10.1126/science.1075958.
 38. Anderson, M. S. *et al.* The cellular mechanism of Aire control of T cell tolerance. *Immunity* **23**, 227–239 (2005).
 39. Ramsey, C. *et al.* Increased antigen presenting cell-mediated T cell activation in mice and patients without the autoimmune regulator. *Eur. J. Immunol.* **36**, 305–317 (2006).
 40. Fergusson, J. R. *et al.* Maturing human CD127⁺ CCR7⁺ PDL1⁺ dendritic cells express AIRE in the absence of tissue restricted antigens. *Front. Immunol.* **10**, (2019).
 41. Lindh, E. *et al.* AIRE regulates T-cell-independent B-cell responses through BAFF. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18466–18471 (2008).
 42. Lindmark, E. *et al.* AIRE expressing marginal zone dendritic cells balances adaptive immunity and T-follicular helper cell recruitment. *J. Autoimmun.* (2013) doi:10.1016/j.jaut.2012.11.004.
 43. Anandasabapathy, N. *et al.* Classical Flt3L-dependent dendritic cells control immunity to protein vaccine. *J. Exp. Med.* **211**, 1875–1891 (2014).
 44. Maier, B. *et al.* A conserved dendritic-cell regulatory program limits antitumour immunity. *Nature* **580**, 257–262 (2020).
 45. Itano, A. A. *et al.* Distinct dendritic cell populations sequentially present antigen to CD4 T cells and stimulate different aspects of cell-mediated immunity. *Immunity* **19**, 47–57 (2003).
 46. Allenspach, E. J., Lemos, M. P., Porrett, P. M., Turka, L. A. & Laufer, T. M. Migratory and Lymphoid-Resident Dendritic Cells Cooperate to Efficiently Prime Naive CD4 T cells. *Immunity* **29**, 795–806 (2008).

47. Miller, J. C. *et al.* Deciphering the transcriptional network of the dendritic cell lineage. *Nat. Immunol.* **13**, 888–899 (2012).
48. Eberl, G. ROR γ t, a multitask nuclear receptor at mucosal surfaces. *Mucosal Immunol.* **10**, 27–34 (2017).
49. Heng, T. S. P. *et al.* The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
50. Pokrovskii, M. *et al.* Characterization of Transcriptional Regulatory Networks that Promote and Restrict Identities and Functions of Intestinal Innate Lymphoid Cells. *Immunity* **51**, 185-197.e6 (2019).
51. Lareau, C. A. *et al.* Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nat. Biotechnol.* **39**, 451–461 (2021).
52. Yoshida, H. *et al.* The cis-Regulatory Atlas of the Mouse Immune System. *Cell* **176**, 897-912.e20 (2019).
53. Haljasorg, U. *et al.* A highly conserved NF- κ B-responsive enhancer is critical for thymic expression of Aire in mice. *Eur. J. Immunol.* **45**, 3246–3256 (2015).
54. LaFlam, T. N. *et al.* Identification of a novel cis-regulatory element essential for immune tolerance. *J. Exp. Med.* **212**, 1993–2002 (2015).
55. Guha, M. *et al.* DNA breaks and chromatin structural changes enhance the transcription of autoimmune regulator target genes. *J. Biol. Chem.* **292**, 6542–6554 (2017).
56. Koh, A. S. *et al.* Aire employs a histone-binding module to mediate immunological tolerance, linking chromatin regulation with organ-specific autoimmunity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15878–15883 (2008).
57. Koh, A. S. *et al.* Rapid chromatin repression by Aire provides precise control of immune tolerance article. *Nat. Immunol.* **19**, 162–172 (2018).
58. Sansom, S. N. *et al.* Population and single-cell genomics reveal the Aire dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic

- epithelia. *Genome Res.* **24**, 1918–1931 (2014).
59. Ardouin, L. *et al.* Broad and Largely Concordant Molecular Changes Characterize Tolerogenic and Immunogenic Dendritic Cell Maturation in Thymus and Periphery. *Immunity* **45**, 305–318 (2016).
 60. Wang, B., Zhao, L., Fish, M., Logan, C. Y. & Nusse, R. Self-renewing diploid Axin2 + cells fuel homeostatic renewal of the liver. *Nature* **524**, 180–185 (2015).
 61. Nusse, R. & Clevers, H. Wnt/ β -Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985–999 (2017).
 62. Zhang, H., Ye, M., Welner, R. S. & Tenen, D. G. Sox4 Is Required for the Formation and Maintenance of Multipotent Progenitors. *Blood* **124**, 1577–1577 (2014).
 63. Kashiwada, M., Pham, N.-L. L., Pewe, L. L., Harty, J. T. & Rothman, P. B. NFIL3/E4BP4 is a key transcription factor for CD8 α + dendritic cell development. *Blood* **117**, 6193–6197 (2011).
 64. Ogawa, M. *et al.* Expression and function of c-kit in hemopoietic progenitor cells. *J. Exp. Med.* **174**, 63–71 (1991).
 65. Jiang, A. *et al.* Disruption of E-Cadherin-Mediated Adhesion Induces a Functionally Distinct Pathway of Dendritic Cell Maturation. *Immunity* **27**, 610–624 (2007).
 66. Manicassamy, S. *et al.* Activation of β -catenin in dendritic cells regulates immunity versus tolerance in the intestine. *Science* (80-.). **329**, 849–853 (2010).
 67. La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
 68. Ebertl, G. & Litman, D. R. Thymic origin of intestinal $\alpha\beta$ T cells revealed by fate mapping of ROR γ t+ cells. *Science* (80-.). **305**, 248–251 (2004).
 69. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).
 70. Bornstein, C. *et al.* Single-cell mapping of the thymic stroma identifies IL-25-producing tuft epithelial cells. *Nature* **559**, 622–626 (2018).

71. Kernfeld, E. M. *et al.* A Single-Cell Transcriptomic Atlas of Thymus Organogenesis Resolves Cell Types and Developmental Maturation. *Immunity* **48**, 1258-1270.e6 (2018).
72. Wells, K. L. *et al.* Combined transient ablation and single-cell RNA-sequencing reveals the development of medullary thymic epithelial cells. *Elife* **9**, 1–27 (2020).
73. Dhalla, F. *et al.* Biologically indeterminate yet ordered promiscuous gene expression in single medullary thymic epithelial cells. *EMBO J.* **39**, 1–18 (2020).
74. Tomofuji, Y. *et al.* Chd4 choreographs self-antigen expression for central immune tolerance. *Nat. Immunol.* **21**, 892–901 (2020).
75. Rossi, S. W. *et al.* RANK signals from CD4+3- inducer cells regulate development of Aire-expressing epithelial cells in the thymic medulla. *J. Exp. Med.* **204**, 1267–1272 (2007).
76. Hikosaka, Y. *et al.* The Cytokine RANKL Produced by Positively Selected Thymocytes Fosters Medullary Thymic Epithelial Cells that Express Autoimmune Regulator. *Immunity* **29**, 438–450 (2008).
77. Simonet, W. S. *et al.* Osteoprotegerin: A novel secreted protein involved in the regulation of bone density. *Cell* **89**, 309–319 (1997).
78. Udagawa, N. *et al.* Osteoprotegerin produced by osteoblasts is an important regulator in osteoclast development and function. *Endocrinology* **141**, 3478–3484 (2000).
79. Metzger, T. C. *et al.* Lineage Tracing and Cell Ablation Identify a Post-Aire-Expressing Thymic Epithelial Cell Population. *Cell Rep.* **5**, 166–179 (2013).
80. Khan, I. S. *et al.* Enhancement of an anti-tumor immune response by transient blockade of central T cell tolerance. *J. Exp. Med.* **211**, 761–768 (2014).
81. Bakhru, P. *et al.* Combination central tolerance and peripheral checkpoint blockade unleashes antimelanoma immunity. *JCI Insight* **2**, (2017).
82. Krishnamurthy, B. *et al.* Autoimmunity to Both Proinsulin and IGRP Is Required for Diabetes in Nonobese Diabetic 8.3 TCR Transgenic Mice. *J. Immunol.* **180**, 4458–4464

- (2008).
83. Mimitou, E. P. *et al.* Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nature Biotechnology* (Springer US, 2021). doi:10.1038/s41587-021-00927-2.
 84. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
 85. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* **8**, 329-337.e4 (2019).
 86. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
 87. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat. Methods* **18**, 1333–1341 (2021).
 88. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. ChromVAR: Inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat. Methods* **14**, 975–978 (2017).
 89. Pabst, O. & Mowat, A. M. Oral tolerance to food protein. *Mucosal Immunol.* **5**, 232–239 (2012).
 90. Dunlop, J. H. & Keet, C. A. Epidemiology of Food Allergy. *Immunol. Allergy Clin. North Am.* **38**, 13–25 (2018).
 91. Garkaby, J. *et al.* The Sesame-Peanut Conundrum in Israel: Reevaluation of Food Allergy Prevalence in Young Children. *J. Allergy Clin. Immunol. Pract.* **9**, 200–205 (2021).
 92. Vale, S. L. *et al.* A systematic review of infant feeding food allergy prevention guidelines – can we AGREE? *World Allergy Organ. J.* **14**, 100550 (2021).
 93. Worbs, T. *et al.* Oral tolerance originates in the intestinal immune system and relies on antigen carriage by dendritic cells. *J. Exp. Med.* **203**, 519–527 (2006).

94. Mazzini, E., Massimiliano, L., Penna, G. & Rescigno, M. Oral tolerance can be established via gap junction transfer of fed antigens from CX3CR1⁺ macrophages to CD103⁺ dendritic cells. *Immunity* **40**, 248–261 (2014).
95. Esterházy, D. *et al.* Classical dendritic cells are required for dietary antigen-mediated induction of peripheral T reg cells and tolerance. *Nat. Immunol.* **17**, 545–555 (2016).
96. Coombes, J. L. *et al.* A functionally specialized population of mucosal CD103⁺ DCs induces Foxp3⁺ regulatory T cells via a TGF- β – and retinoic acid–dependent mechanism. *J. Exp. Med.* **204**, 1757–1764 (2007).
97. Loschko, J. *et al.* Absence of MHC class II on cDCs results in microbial-dependent intestinal inflammation. *J. Exp. Med.* **213**, 517–534 (2016).
98. Esterházy, D. *et al.* Compartmentalized gut lymph node drainage dictates adaptive immune responses. *Nat.* 2019 5697754 **569**, 126–130 (2019).
99. Abramson, J., Dobeš, J., Lyu, M. & Sonnenberg, G. F. The emerging family of ROR γ t⁺ antigen-presenting cells. *Nat. Rev. Immunol.* 2023 241 **24**, 64–77 (2023).
100. Feng, J. *et al.* Clonal lineage tracing reveals shared origin of conventional and plasmacytoid dendritic cells. *Immunity* **55**, 405–422.e11 (2022).
101. Hong, S. W. *et al.* Immune tolerance of food is mediated by layers of CD4⁺ T cell dysfunction. *Nat.* 2022 6077920 **607**, 762–768 (2022).
102. Friedman, A. & Weiner, H. L. Induction of anergy or active suppression following oral tolerance is determined by antigen dosage. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6688–6692 (1994).
103. Chen, Y. *et al.* Peripheral deletion of antigen-reactive T cells in oral tolerance. *Nat.* 1995 3766536 **376**, 177–180 (1995).
104. Hadis, U. *et al.* Intestinal Tolerance Requires Gut Homing and Expansion of FoxP3⁺ Regulatory T Cells in the Lamina Propria. *Immunity* **34**, 237–246 (2011).
105. Dobeš, J. *et al.* Extrathymic expression of Aire controls the induction of effective TH17

- cell-mediated immune response to *Candida albicans*. *Nat. Immunol.* 2022 237 **23**, 1098–1108 (2022).
106. Lyu, M. *et al.* ILC3s select for ROR γ t⁺ Tregs and establish tolerance to intestinal microbiota. *bioRxiv* (2022) doi:<https://doi.org/10.1101/2022.04.25.489463>.
 107. Peschon, J. J. *et al.* Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *J. Exp. Med.* **180**, 1955–1960 (1994).
 108. Paucar Iza, Y. A. & Brown, C. C. Early life imprinting of intestinal immune tolerance and tissue homeostasis. *Immunol. Rev.* **323**, 303–315 (2024).
 109. Cao, X. *et al.* Defective lymphoid development in mice lacking expression of the common cytokine receptor gamma chain. *Immunity* **2**, 223–238 (1995).
 110. Loschko, J. *et al.* Inducible targeting of cDCs and their subsets in vivo. *J. Immunol. Methods* **434**, 32–38 (2016).
 111. Poliani, P. L. *et al.* Human peripheral lymphoid tissues contain autoimmune regulator-expressing dendritic cells. *Am. J. Pathol.* **176**, 1104–1112 (2010).
 112. Antonova, A. U. *et al.* A distinct human cell type expressing MHCII and ROR γ t with dual characteristics of dendritic cells and type 3 innate lymphoid cells. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2318710120 (2023).
 113. Waskow, C. *et al.* The receptor tyrosine kinase Flt3 is required for dendritic cell development in peripheral lymphoid tissues. *Nat. Immunol.* 2008 96 **9**, 676–683 (2008).
 114. Maraskovsky, E. *et al.* Dramatic increase in the number of functionally mature dendritic cells in Flt3 ligand-treated mice: Multiple dendritic cell subpopulations identified. *J. Exp. Med.* **184**, 1953–1962 (1996).
 115. Schraml, B. U. *et al.* XGenetic tracing via DNCR-1 expression history defines dendritic cells as a hematopoietic lineage. *Cell* **154**, 843–858 (2013).
 116. Narasimhan, H. *et al.* Cross-species analyses reveal ROR γ t-expressing dendritic cells are a lineage of antigen presenting cells conserved across tissues. *bioRxiv*

2024.05.06.592772 (2024) doi:10.1101/2024.05.06.592772.

117. Rudnitsky, A., Oh, H., Talmor, J. & Kedmi, R. Coordinated network of T cells and antigen presenting cells regulate tolerance to food. *bioRxiv* (2024)
doi:<https://doi.org/10.1101/2024.07.11.603064>.
118. Cheifetz, T. R. & Knoop, K. A. The right educational environment: Oral tolerance in early life. *Immunol. Rev.* 1–18 (2024) doi:10.1111/imr.13366.
119. Weiner, H. L. Oral tolerance for the treatment of autoimmune diseases. *Annu. Rev. Med.* **48**, 341–351 (1997).
120. Pinheiro-Rosa, N. *et al.* Oral tolerance as antigen-specific immunotherapy. *Immunother. Adv.* **1**, 1–21 (2021).
121. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 2017 81 **8**, 1–12 (2017).
122. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* 2018 1512 **15**, 1053–1058 (2018).
123. Xu, C. *et al.* Automatic cell-type harmonization and integration across Human Cell Atlas datasets. *Cell* **186**, 5876–5891.e20 (2023).

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

Jiaxi Wang

7CFD96D66B594B2...

Author Signature

7/31/2024

Date