2020

# Molecular Mechanisms and Antigen Receptor Requirements for Lymphocyte Adaptation to Intestinal Tissues

Mariya B. London


# MOLECULAR MECHANISMS AND ANTIGEN RECEPTOR REQUIREMENTS FOR LYMPHOCYTE ADAPTATION TO INTESTINAL TISSUES

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfillment of the Requirements for
the degree of Doctor of Philosophy

by
Mariya B. London
June 2020

# Molecular mechanisms and antigen receptor requirements for lymphocyte adaptation to intestinal tissues

Mariya B. London, Ph.D.
The Rockefeller University 2020

The intestine plays a crucial role in food digestion, nutrient absorption, water retention, and waste excretion. It contains the most populous immune cell reservoir in the body and is continuously exposed to a large and diverse number of diet- and microbiota- derived antigens. The highly stimulating luminal environment is separated from the core of the body, the lamina propria (LP), by just a single layer of epithelial cells. The intestinal immune system is thus tasked with being able to tolerate innocuous stimuli while mounting an effective response against potential pathogens in a controlled manner. To ensure appropriate balance between tolerance and resistance, T cells undergo tissue adaptation upon migrating from the gut-draining mesenteric lymph nodes (mLN) to the intestinal lamina propria and epithelium (IE). We sought to elucidate the transcriptional mechanisms and T cell receptor (TCR) signaling requirements of $CD4^+$ T cell plasticity and adaptation in the intestinal tissues.

Within the intestine, peripherally induced $Foxp3^+$ regulatory T cells (iTregs), which are instrumental in limiting inflammatory responses to non-self antigen, are located primarily in the lamina propria. However, $CD8\alpha\alpha$-expressing intraepithelial $CD4^+$ T cells (CD4-IELs), which also exhibit anti-inflammatory properties and depend on similar environmental cues, reside in the epithelium. Using intravital microscopy, we find distinct cell dynamics of intestinal Tregs and CD4-IELs. We addressed the molecular imprinting of the gut epithelium on T cells by integrating mouse genetics with single-cell RNA-sequencing analyses. Transcriptionally, $CD4^+$ T cells from mLN, LP and IE segregate based on the intestinal layer they occupy; trajectory analysis suggests a stepwise loss of CD4-programming and acquisition of an intraepithelial profile as $CD4^+$ T cells adapt to the epithelium and convert to CD4-IELs.

We found that upon migration to the epithelium, Tregs can lose Foxp3 expression and convert to CD4-IELs in a microbiota-dependent fashion, an effect in part attributed to loss of the CD4 lineage-defining transcription factor ThPOK. Treg fate-mapping coupled with RNA- and ATAC-sequencing revealed that the Treg program shuts down before an intraepithelial program becomes fully accessible at the epithelium. Ablation of *Thpok* results in premature acquisition of an IEL profile by mLN Tregs, partially recapitulating epithelium imprinting. Furthermore, we demonstrate that iTregs and CD4-IELs perform complementary roles in the regulation of intestinal inflammation in response to dietary antigen.

To uncover the specific role of the T cell receptor in the process of CD4-IEL development, we combined *in vivo* fate-mapping and gene ablation models with single cell TCR-

sequencing. Single-cell TCR repertoire and transcriptomic analysis of intraepithelial $CD4^+$ T cells revealed different extents of clonal expansion and TCR overlap between cell states; fully differentiated CD4-IELs from regulatory or conventional $CD4^+$ T cells were the least diverse. Conditional deletion of TCR on differentiating $CD4^+$ T cells or of MHC-II on intestinal epithelial cells prevented CD4-IEL differentiation. However, TCR ablation on developed CD4-IELs did not affect their accumulation.

Overall, our results reveal an inter- and intra-tissue specialization of anti-inflammatory $CD4^+$ T cells shaped by discrete niches of the intestine. We uncovered the stepwise molecular mechanisms and TCR-signaling requirements for T cells to adapt to the intestinal epithelium. We found that the coordinated replacement of the circulating lymphocyte program with site–specific transcriptional and chromatin changes is necessary for tissue imprinting. Furthermore, our results indicate that local recognition of possibly a limited set of antigens is an essential signal for the differentiation and adaptation of T cells to the epithelium. Taken together, the work presented in this thesis demonstrates that a combination of genetic, TCR, and environmental triggers is crucial in driving T cell plasticity and adaptation to the tissues within the intestine.

*I dedicate this thesis to my family.*
*Thank you for the unconditional love + support, and for pushing me to excel.*

# Acknowledgements

Throughout my time at Rockefeller, many people have positively influenced my education, scientific pursuits, well-being and overall experience. To all of them, I am extremely grateful. First of all, I'd like to thank my mentor, **Daniel Mucida**. Your commitment to science, mentorship and improvement (both to self- and non-self) are very much appreciated. Thank you for your guidance, patience, general open-door policy, and always taking advantage of teachable moments. I'd also like to extend my gratitude to **Angelina Bilate**, it's been a sincere pleasure to work together. I've learned a lot from you on many different fronts: science, life and even Português. Thank you for always being ready to share your friendship and mentorship! A big thank you to **Tiago B RdC**, your hard work, scientific analysis and puns are much appreciated. Muito obrigado por seu amor e apoio, por estar ao meu lado e por ser meu Xuxu!

The friendly and collaborative lab environment has been a great place to grow, so thank you to all of the members of the **Mucida Lab**, both past and present! **Ciça**, thank you for your much-appreciated friendship, positivity and always being there for me, from fruit-time to gym-time! **Ainsley**, your friendship, gut-prep and editing help, brain-boat commiserating, and readiness with a hug, a wine or a travel plan as needed are much appreciated! **Fanny**, danke schön for our walks, and always having an ear, a hug, and a tea for me! **Bernardo**, your positivity, patience and wisdom (regarding science and life) have made lab a great place to be, thank you for being the sunshine in the lab! **Roham**, it's always great chatting about life and IELs with you, thank you for your advice regarding both! **Juliana**, thank you for your positivity and inclusiveness, for looking out for me and my mice. **Carla**, thanks for always being there for a de-stress chat, a wine or a crêpe! **Paul**, your wisdom, commiseration, colorful remarks + markers have made sharing many hours across the bench a pleasure (and thanks for making my side of the bench look clean)! **Aneta**, your dedication to keeping the lab organized and running is much appreciated, as are your amazing baked and knitted goods! **Virginia**, thank you for your friendship, mentorship, and chocolate-drawer! **David**, your overall positivity in and out of the lab has been a great contributor to my happiness. **Daria**, thank you for teaching me how to do a gut prep, and your your patience and mentorship from the very beginning! **Tomo**, it was very nice to both work and tour Japan with you! A big thanks to **Tomasz** and **Greg** for your grad-school advice and calming remarks, and **Aubrey** and **Begum** for always having a big smile. I'd like to also extend my appreciation to the **Victora Lab** and the **Nussenzweig Lab**, for being great neighbors and collaborators. Thank you for always being ready to share scientific input, reagents, and the lunch table.

The Rockefeller community as a whole has provided a lot of support throughout my time here. I'd like to thank my committee members, **Luciano Marraffini** and **Sohail Tavazoie**, your scientific input and encouragement throughout the years have been very helpful! Thank you to **Ivaylo Ivanov** for agreeing to serve as my external committee member! The **dean's office** has made all of the non-scientific aspects of graduate school easy, for which I'm grateful! Especially to **Sid** and **Emily** for your support and to **Cris** for your quick

responsiveness and always being there to help! **Uwe**, you were my 1st mentor at Rockefeller, and I'll always appreciate you setting the stage for me on a positive note! Thank you to the **sequencing facility** staff (especially **Connie** + **Christine**), **Sara** and the rest of the **animal facility** staff, and the **flow cytometry** core (**Kristie**, **Kal**, **Klara**, and **Neena**), and the rest of the staff at Rockefeller for supporting my research!

My friends outside of the lab have really contributed to making my graduate school journey enjoyable! Thank you to **Erin** for your constant support and always checking in on me (and to **Melody** for providing walks and cuddles), to **Audrey** for keeping me sane and in shape, my local Jhop friends (**Mara, Coco, Alyssa, Dina, Ali**) for your positivity, friendship and support, **Kunal** for the puns, cubes, and laughs, **Ed** for all the tea-time, and **Elissa** for always being there for me.

Last, but definitely not least, I'd like to thank my family. Thank you for raising me, feeding me, teaching me, pushing me to excel, believing in me, always supporting and loving me, and making home a great place to be! Thank you to my **grandparents** for instilling a love for science, my **mom** for always supporting me, **Vika** for teaching me to both tie my shoes and use a pipette, **Igor** for keeping us sane and positive, and **Bryan** for the upbeat demeanor and the interesting conversations!

# Disclaimer

This thesis contains portions of manuscripts written together with Dr. Daniel Mucida and Dr. Angelina M. Bilate. The work presented was performed in collaboration with others, and is either published or submitted for peer-reviewed journals as follows:

**Chapter 2:** Published in *Science*[1], 2016
Contributing authors: Tomohisa Sujino[α], Mariya London[α], David P. Hoytema van Konijnenburg[α,β], Tomiko Rendon[α], Thorsten Buch[γ], Hernandez M. Silva[δ], Juan J. Lafaille[δ], Bernardo S. Reis[α] and Daniel Mucida[α]

**Chapter 3**: Submitted to *Nature Immunology,* 2020
Contributing authors: Mariya London[α♦], Angelina M. Bilate[α♦], Tiago B. R. Castro[α♦], Daniel Mucida[α]

**Chapter 4:** Submitted to *Immunity*, 2019
Contributing authors: Angelina M. Bilate[α♦], Mariya London[α♦], Tiago B. R. Castro[α], Luka Mesin[ε], Suppawat Kongthong[α], Audrey Harnagel[α], Gabriel D. Victora[ε], Daniel Mucida[α]

♦Equal contribution

[α]Laboratory of Mucosal Immunology, The Rockefeller University, New York, NY 10065, USA.
[β]Laboratory of Translational Immunology, University Medical Center Utrecht, the Netherlands
[γ]Institute of Experimental Immunology, University of Zurich, Zurich, Switzerland.
[δ]Skirball Institute, New York University School of Medicine, New York, NY 10016, USA.
[ε]Laboratory of Lymphocyte Dynamics, The Rockefeller University, New York, NY 10065, USA.

# Table of Contents

# List of Figures

# List of Abbreviations

| | |
|---|---|
| ABX | Antibiotics |
| APC | Antigen presenting cells |
| ATAC | Assay for transposase-accessible chromatin |
| CDR3 | Complementarity determining region 3 |
| ChIP | Chromatin immunoprecipitation |
| CNS | Conserved noncoding sequences |
| D50 | Diversity 50 |
| DACR | Differentially accessible chromatin regions |
| DC | Dendritic cell |
| DEG | Differentially expressed genes |
| EAE | Experimental autoimmune encephalomyelitis |
| GF | Germ-free |
| GO | Gene ontology |
| GSEA | Gene set enrichment analysis |
| i.p. | Intraperitoneally |
| IE | Intestinal epithelium |
| IEC | Intestinal epithelial cell |
| IEL | Intraepithelial lymphocyte |
| iTreg | Peripherally induced regulatory T cell |
| IVM | Intravital multi-photon microscopy |
| LN | Lymph node |
| LP | Lamina propria |
| LRT | Likelihood ratio test |
| MHC | Major histocompatibility complex |
| mLN | Mesenteric lymph nodes |
| nTreg | Thymically derived natural regulatory T cell |
| OT-II | Ovalbumin specific T cell receptor transgenic system |
| OVA | Ovalbumin |
| PCA | Principal component analysis |
| RA | Retinoic acid |
| scRNA-seq | Single-cell RNA-sequencing |
| SLO | Secondary lymphoid organ |
| SPF | Specific pathogen-free |
| Tconv | Conventional T cell |
| TCR | T cell receptor |
| TF | Transcription factor |
| TGF$\beta$ | Transforming growth factor $\beta$ |
| TL | Thymic leukemia antigen |
| $T_{RM}$ | Tissue resident memory |
| TSS | Transcriptional start site |
| UMAP | Uniform manifold approximation and projection |
| WT | Wild type |

# Chapter 1: Introduction

## 1.1 Lymphocytes and the immune response

The adaptive immune system must be able to mount an effective, yet controlled response against potential pathogens while tolerating harmless stimuli. Lymphocytes carry antigen receptors that allow for the recognition of virtually any antigen, defined as a substance that can induce an immune response, they may encounter during a lifetime of an individual. They should be able to distinguish and tolerate self-antigens, yet actively respond to harmful or non-self antigens without damaging the host. T lymphocytes are derived from hematopoietic stem cells in the bone marrow and develop in the thymus, where they undergo a series of selection steps. Although some T cells express the $\gamma\delta$ T cell receptor (TCR), this work focuses on $\alpha\beta$TCR-bearing T cells.

Approximately 1-2 % of thymocytes are selected to exit the thymus as mature T cells[2, 3], in part on the basis of the affinity and avidity of their T cell receptor complex for the major histocompatibility (MHC)-self-peptide complex[4]. The strength with which the TCR complex engages with the peptide-MHC complex influences thymocyte fate[5]. Interactions that are too weak or too strong preclude survival, while those that fall within the appropriate range are selected to mature. Thymocytes with a permissibly stronger TCR signal strength further develop into certain T cell subtypes, such as natural regulatory T cells (nTregs), natural killer T cells and natural intraepithelial lymphocytes[6].

The theoretical diversity of the TCR repertoire is $10^{15}$-$10^{20}$ in humans[7] and $10^{14}$ in mice[8, 9] , values estimated by the combinations of germline variable (V), diversity (D), and joining (J) segments, along with random nucleotide insertions and deletions at junction sites and the pairing of the two TCR chains, $\alpha$ and $\beta$[7]. The complementarity-determining region 3 (CDR3) is encoded by the V(D)J regions of each chain, has the highest variability, is in direct contact with the peptide loaded on the MHC complex, and is what determines the T cell's ability to recognize antigen[10, 11]. Although in general quite diverse, the TCR repertoire can be further shaped by thymic selection, cell subtype and the extrathymic environment.

In addition to TCR signaling, developing thymocytes transition through several stages marked by the absence of surface co-receptor molecules CD8 and CD4, then their dual expression, and finally the presence of either CD8 or CD4[12, 13]. The successful interaction of thymocytes with MHC class I or II determines their $\alpha\beta$TCR T cell lineage, CD8 or CD4, respectively.[14] Cells of the former express the CD8$\alpha\beta$ heterodimer on their surface and the CD8-lineage defining transcription factor (TF) *Runx3*[15, 16, 17]. On the other hand, cells of the latter express the CD4-lineage defining TF *Zbtb7b* (referred to as *Thpok*)[18, 19]. Expression of Runx3 and ThPOK is mutually exclusive and this dichotomy is established in the thymus during T cell

selection and development[20]. Partially in response to regulation of its distal enhancer, *Runx3* is upregulated in CD8-single positive thymocytes. Runx3 not only silences *Cd4* but also *Thpok* by binding to its silencer region[21], precluding its expression, and driving the upregulation of CD8-related genes, many of which are considered to be cytotoxic. *Thpok* is progressively upregulated during MHC-II induced selection as thymocytes down-regulate CD8[22]. ThPOK binds to its own silencer region in $CD4^+$ T cells, allowing for its continued expression[23], along with the regulation of two enhancers. In turn, CD8 genes are repressed and the CD4 lineage is established and maintained throughout subsequent peripheral $CD4^+$ T cell activation and differentiation[20, 24, 25].

## 1.2 Peripheral T cell differentiation and plasticity

Upon successful maturation in the thymus, known as a primary lymphoid organ, conventional mature, yet naïve T lymphocytes exit to the peripheral, or secondary lymphoid organs (SLO)[26]. Although a subset of T cells proceeds directly to tissues, naïve CD8 and CD4 T cells in the SLO can be activated and undergo further differentiation upon encounter with and recognition of peptide antigen presented on MHC-I or II complexes, respectively[26]. While all nucleated cells express MHC-I, only a subset of cells express MHC-II. Professional antigen presenting cells (APCs) specialize in internalizing and processing antigen and displaying their processed peptides on the surface of MHC-class II molecules to peripheral naïve $CD4^+$ T cells[27]. APCs, particularly dendritic cells (DCs), migrate after capturing antigen in various tissues to their respective draining lymph nodes (LN). Upon recognition of cognate peptide-MHC-II complexes in secondary lymphoid organs, and influenced by specific environments, $CD4^+$ T cells can develop into one of several canonical subtypes[28, 29], including $T_H1$, $T_H2$, $T_H17$, $T_{FH}$, and induced regulatory T cells (iTregs). In addition to activation, primed T cells also receive initial migration cues from DCs in the draining LNs, resulting in their migration back to the location of stimuli. For instance, $CX_3CR_1$-expressing[30] $CD103^+$ DCs can migrate to the intestine and capture luminal contents before returning to the gut draining mesenteric lymph nodes (mLN) where they activate naïve $CD4^+$ T cells[31, 32, 33, 34, 35], resulting in their differentiation and migration to intestinal tissues, thus completing the gut-LN axis.

T cells retain a level of plasticity that allows them to adapt to different stimuli and tissue environments. In this context, plasticity is defined as the ability of a single T cell to undertake characteristics canonically ascribed to multiple distinct T cell subsets, either simultaneously or sequentially throughout its life cycle[36]. Many parallels can be extended between the plasticity of $CD4^+$ T cells and that of stem cells, to which this term was initially ascribed. Two important features of stem cells are multipotency and asymmetric division, marked by their capability to both generate daughter cells that retain the initial undifferentiated state and developmental potential as well as the capability to generate a variety of more

differentiated daughter cells[37]. Simultaneous multipotency and self-renewal has been demonstrated in central memory $CD8^+$ T cells[38, 39] . Stem cell regulatory features have also been noted in subsets of $CD4^+$ T cells. microRNAs that have been found to tune key aspects of stem cells[40] have also been linked to the development of $T_H1$ and $T_H17$ subsets under inflammatory conditions[41]. However, it is not known if these stem cell characteristics, simultaneous multipotency and self-renewal, can be extended to conventional non-memory $CD4^+$ T cells. Furthermore, one key distinction is that unlike stem cells which further develop into distinct, more differentiated cell subtypes, T cells can either progress towards a more terminally differentiated state or obtain aspects from multiple cell types simultaneously. For example, in some inflammatory contexts, differentiated $T_H17$ cells can produce the $T_H1$-associated cytokine IFN$\gamma$ along with IL-17A, in turn undertaking a more pathogenic phenotype[42]. Additionally, regulatory T cells can co-express transcription factors associated with other helper subtypes, such as $T_H17$ cells, upon specific bacterial stimuli[43, 44]. Overall, the plasticity of T cells, which display some level of stem cell features, is important for their adaptation to distinct environments.

## 1.3 Regulatory T cells in the periphery

While natural Tregs are thymically-derived with high affinity for self-antigen, induced Tregs upregulate the Treg-defining transcription factor Foxp3, along with a cohort of associated regulatory genes, in the periphery[45]. Although they share a core set of Treg genes, natural and induced Tregs can be distinguished by certain markers such as *Helios* and *Neuropilin-1*[46, 47]. Distinct gene expression has been observed in regulatory T cells in different tissues, including the small and large intestines[48]. In addition, the relative frequencies of the two types of Tregs, natural and induced, differ depending on location. While induced Tregs are enriched in intestinal tissues, natural Tregs are more frequent in intestinal draining lymph nodes. Together, these observations suggest possible specialization or a division of labor of regulatory cells within different tissues.

The main function of Tregs is to actively suppress immune response, thus maintaining tolerance and homeostasis. Loss of function mutations in *Foxp3*, such as *scurfy* in mice, lead to severe autoimmunity[49]. Similar *FOXP3* mutations in humans result in the immunodysregulation polyendocrinopathy enteropathy, X-linked syndrome (IPEX)[50]. Tregs are crucial for oral tolerance[51], an active process in which the immune system response to antigens to which it has had previous oral exposure is highly controlled to prevent damage. This is particularly important in the intestine due to the high local food-derived antigen-burden. In mouse models, genetic ablation of Tregs results in inflammatory mucosal disease such as colitis, further highlighting the role of Tregs in the intestine[52].

Although the functional importance of Tregs is well studied, their mechanisms of action remain an active area of research. Nevertheless, Tregs have been described to suppress effector T cells in multiple ways. For instance, Tregs express inhibitory molecules like CTLA-4[53], which binds to CD80/CD86 on APCs and blocks other T cells from receiving the co-stimulatory CD28 signaling. Similarly, LAG-3 on Treg surfaces binds to MHC-II on APCs with high affinity, precluding their ability to activate T cells[54]. Tregs can also lead to the increase of intracellular cyclic AMP in effector T cells, leading to their suppression[55]. Furthermore, Tregs produce anti-inflammatory cytokines, such as IL-35 and IL-10 to mitigate inflammation[56, 57]. IL-10 ablation from Tregs results in a pro-inflammatory effect primarily in the skin and mucosal tissues[58], suggesting that Treg function can be tissue-dependent.

In addition to Tregs modulating the environment, the environment in turn influences Treg differentiation. For example, the presence of certain commensal bacteria in the intestine, such as *Helicobacter hepaticus* and *Clostridium* species has been linked to Treg differentiation[59, 60]. Furthermore, single cell TCR sequencing has revealed that colonic Tregs use TCRs that are both distinct from other $CD4^+$ T cells in the colon and from Tregs in other organs such as skin[61]. In fact, this study found that colonic-derived bacterial antigens are able to stimulate colonic Treg TCRs *in vitro*, suggesting that the local antigen milieu shapes the Treg TCR repertoire. A sufficiently diverse TCR repertoire is suggested to be important to be able to respond to a vast array of antigen and maintain homeostasis. In the presence of a diverse microbiota, "limited mice" harboring T cells which only utilize a fixed TCR$\alpha$ and TCR$\beta$ chain develop spontaneous colitis, induced by an increased $IL\text{-}17^+$ $IFN\gamma^+$ producing $T_H17$ population[62]. This study suggested that if TCR diversity is limited, Tregs are unable to induce tolerance to commensal microbiota stimuli in part by failure to suppress local DC activation and priming of pro-inflammatory T cells. Another model with a fixed TCRV$\beta$ chain revealed that while the TCR repertoire of a colonic $Foxp3^+ROR\gamma\tau^+$ Treg population is mostly unique when compared to other colonic T cell subsets, it was in part shared with local $ROR\gamma\tau^+$ $T_H17$ cells[63]. This unique sharing suggests that Tregs with a similar TCR repertoire and co-expression of effector TF's with corresponding effector T cell subtypes functions to specifically suppress those cells.

## 1.4 Intraepithelial lymphocytes

T cell adaptation to tissues within the intestine, which plays a crucial role in food digestion, nutrient absorption, water retention, and waste excretion is important for its homeostasis. Two adjacent tissues, the intestinal epithelium (IE) and underlying lamina propria (LP) contain more immune cells than any other organ. Although only a basement membrane separates the single layer of intestinal epithelial cells (IECs) and the IE environment from the LP, the two compartments are populated by very different immune cell populations. While a heterogenous population of

conventional CD4+ and CD8+ T cells as well as the majority of intestinal APCs reside in the lamina propria, the predominant lymphocyte populations within the intestinal epithelium consist of CD8αα-expressing intraepithelial lymphocytes (IELs)[64]. The four types of IELs are thought to be terminally differentiated and have an activated yet resting phenotype[65]. IEL subsets seed the intestine at different times and in response to different stimuli, but all establish permanent residency. The two types of natural IELs, αβTCR CD8αα+ and γδTCR CD8αα+ (both CD4− CD8β−), acquire gut homing molecules and the IEL profile directly in the thymus and populate the intestine shortly after birth in a largely microbiota-independent way. On the other hand, the two peripherally induced αβTCR IEL types; CD8αβ+CD8αα+ (CD8-IELs) and CD4+CD8αα+ (CD4-IELs) develop *in situ* over time, in a microbiota-dependent manner[65]. Of note, CD4-IELs are derived from conventional CD4+ T cells, which display a unique level of lineage-defying T cell plasticity within the epithelium. Upon migrating to the IE, CD4-IEL precursors downregulate ThPOK and upregulate the surface CD8αα homodimer and a cohort of CD8 and cytotoxic genes including *Runx3*[66, 67].

In addition to cytotoxic CD8-related genes such as granzymes, IELs express effector cytokines, activating and inhibitory natural killer cell receptors, and anti-inflammatory receptors[65]. IELs also express tissue resident markers such as CD69 and CD103[65, 68]. This profile, in combination with the longevity of IELs resembles that of tissue resident memory ($T_{RM}$) cells. Despite their similarities, IELs have not been definitively classified as $T_{RM}$ cells.

Much of $T_{RM}$ knowledge stems from *bona fide* CD8+ T cells, with only slight parallels drawn to CD4+ T cells. $T_{RM}$ cells are antigen-experienced T cells which reside in tissue, do not recirculate, and do not require the presence of antigen for maintenance[69, 70, 71, 72]. They are thought to derive from a stimulus, such as an infection, and function as an antigen-specific first responder cell poised to rapidly and effectively alarm the neighborhood upon re-infection. $T_{RM}$ cells are effector-like, terminally differentiated, and retain proliferative capacity upon re-stimulation[71]. These cells are highly adapted to their tissue niche, displaying environment-dependent distinct gene expression profiles. $T_{RM}$ cells in different tissues detect and respond to recurring pathogens with different speeds and morphologies. For instance, CD8+ $T_{RM}$ cells in the skin epidermis adopt a dendritic cell-like morphology to be able to scan more host cells for pathogens[73, 74] and CD8+ $T_{RM}$ cells increase their speed upon intestinal infection[75]. Similarly, γδTCR-IELs change the movement and speed of their homeostatic surveillance behavior upon *Salmonella* infection, suggesting a $T_{RM}$-like behavior for these cells in the gut[76]. Although a core $T_{RM}$ signature has been ascribed to CD8+ $T_{RM}$ cells, such as *Hobit* and *Blimp1*[77] being their master TF regulators, they differ based on their tissue environments[78]. For instance, some $T_{RM}$ cells require CD103 expression for maintenance, but not those found in the gut[79]. It is postulated that CD103 expression by CD4-IELs physically retains them in the intestinal epithelium (by

binding to E-cadherin on IECs) but is not required for a $T_{RM}$ state.[71] Likewise, while the CD69 expression by IELs resembles $T_{RM}$, they still make Granzyme B (GzmB), suggesting that they are recently stimulated, and not memory cells. CD69 can be transiently expressed through TCR stimulation or inflammatory cytokines (such as IFN$\gamma$ and TNF$\alpha$,) to prevent tissue egress and allow for proper programing within the tissue, but it lasts for a few days[80]. However, IELs constitutively express CD69, suggesting they are always sensing stimuli. Re-activated $T_{RM}$ cells produce IFN$\gamma$, TNF$\alpha$, and IL-2, enabling them to recruit leukocytes and trigger interferon-stimulated genes in neighboring cells to resist re-infection in the local area. The IFN$\gamma$ and GzmB secretion by CD4-IELs, in combination with their IE-residency, further suggests that they may be $T_{RM}$ cells. However, as it is still unknown what CD4-IELs respond to, one cannot truly classify them as $T_{RM}$ cells as their response to re-occurring antigen cannot yet be measured. Furthermore, most studies regarding $CD4^+$ $T_{RM}$ cells have been limited to polyclonal cells without accessing antigen-specificity. While such specificity of CD4-IELs is unknown, they display a limited TCR repertoire, particularly in older animals, suggesting they may have memory to specific antigens.

Despite a potential for recognition of certain antigens, IELs are thought to have a dampened perception of TCR signaling due to their expression of the CD8$\alpha\alpha$ homodimer[65]. When it is paired with CD8$\beta$, the CD8$\alpha$ co-receptor functions to enhance intracellular TCR activation signaling events by associating with Src kinase $p56^{lck}$ and LAT to allow for proper phosphorylation, and thus a cascade of downstream signaling[81]. However, when CD8$\alpha$ is paired in a homodimer, as is the case on IELs, it sequesters this downstream signaling due to associating with a separate lipid raft from that of the TCR complex, therefore dampening TCR signaling[82]. Instead, CD8$\alpha\alpha$ interacts with thymic leukemia antigen (TL)[83, 84], a non-classical MHC-I molecule that is expressed on IECs as well as in the thymus. It is thought that TL does not present antigen as it lacks the helical domains that define the antigen-binding grove for MHC-I molecules. It is hypothesized that the TCR-dampening role of CD8$\alpha\alpha$ on IELs functions to either attenuate ongoing stimulation or increase the minimal signal strength required for re-stimulation of antigen-experienced IELs in the antigen-rich environment of the intestine[65, 85].

The function of IELs as a whole is poorly characterized. $\gamma\delta$TCR IELs have been shown to function in surveillance of the IE at steady state, defense against enteric pathogens, and repair of the epithelial barrier[76, 86, 87]. Their absence has been linked to increased susceptibility to infections by *Toxoplasma gondii*, *Salmonella Typhimurium*, and *Listeria monocytogenes*[76, 87, 88]. Similarly, CD8-IELs have been described to protect against LCMV, rotavirus, and *Toxoplasma gondii* infection[89]. Both CD8- and CD4-IELs have been implicated in contributing to pro-inflammatory damage in celiac disease[90], and inflammatory bowel disease[91]. These detrimental roles are in part mediated by NKG2D-mediated targeting of IECs by CD8-IELs, leading to villous atrophy, as well as the production of pro-inflammatory cytokines

by CD4-IELs[91, 92]. However, CD4-IELs have also been implicated in anti-inflammatory roles[93, 94] such as mediating immune responses to dietary antigens[1]. Overall, IELs may sense their environment and accordingly adapt in both their expression programs and functional responses.

## 1.5 Immune crosstalk in the intestinal environment

The intestine is considered to be in a chronically, yet physiological, inflamed state, in part due to the high abundance of microbial and dietary antigens[95]. In addition to food, the human intestine harbors $10^{13}$ commensal bacteria[96], paralleling the number of host cells that make up an average human. The intestinal immune system is thus tasked with tolerating the harmless stimuli, in order to reap their benefits, while still being able to mount an effective response against potential pathogens in a controlled manner[91].

Within the intestinal environments, the interplay of immune cells, intestinal epithelial cells, food and the microbiota is crucial for T cell adaptation and a healthy homeostasis. Both direct and indirect interactions of the commensal microbiota and food metabolites with IECs impact their differentiation and secretion of factors which further influence the local environment. For instance, IEC production of antimicrobial peptides[97, 98], mucous and tight junction proteins, all of which can be modulated by the microbiota, in turn restructure the intestinal bacterial communities[99, 100, 101] and contribute to barrier protection[102]. Some bacteria can be directly sensed by pattern recognition receptors (such as toll-like[103] and nucleotide-binding oligomerization domain-like receptors[104]) on IECs, while others, such as *Segmented filamentous bacteria* (SFB) can directly attach to the epithelium, leading to changes in their gene expression profiles[105, 106, 107]. Moreover, certain members of the microbiota, such as *Bifidobacteria,* specialize in fermenting dietary fibers, a process that results in the release of short chain fatty acids[108], which are then sensed by IECs. In turn, the proliferative[109] and secretory capacities of IECs are modulated, which can further influence the differentiation of local immune cells, displaying the interplay of diet, microbiota, IECs and lymphocytes.

Pathogens also contribute to the intestinal environment crosstalk. For example, helminths and *Salmonella Typhimurium* drive the development of Tuft cells[110] and microfold cells[111] (M cells), respectively, which then facilitate an immune response against these pathogens. In response to helminths, Tuft cells act in an auto-regulatory loop by inducing a type 2 immunity response, which leads to further Tuft cell development until pathogen clearance[112]. M cells contribute to the connection between the microbiota and immune system by transporting bacterial antigens from the lumen to the lymphoid follicle, where APCs can sense them via toll like receptors, uptake and subsequently present antigen to T cells[113].

Secretion of factors such as transforming growth factor β (TGFβ), cytokines and the expression of gut homing ligands by IECs, can all be promoted by the bacterial and diet components and impact the local T cell population[114, 115, 116, 117, 118]. For instance, while TGFβ directs $CD4^+$ T cells to differentiate towards inflammatory $T_H17$ cells in the presence of IL-6, the additional presence of diet-derived metabolite retinoic acid (RA) inhibits the added contribution of the latter, and instead re-directs differentiation towards induced Tregs[119]. Moreover, TGFβ and RA can induce the upregulation of the gut homing integrin $\alpha_E\beta_7$, which binds to E-cadherin on IECs and allows for T cell homing and retention in the gut[120]. RA can also promote CCR9 expression, which binds to IEC-expressed CCL25, although this effect is more important to homing to the small, rather than the large intestine[121]. Overall, the crosstalk between multiple components within the intestine is crucial for pathogen clearance and also for the maintenance of homeostasis.

## 1.8 Overall goal of the thesis

The local immune system not only contributes to, but also adapts to the chronically inflamed state of the gut. It must be able to tolerate harmless stimuli while maintaining the ability to respond to pathogens in a controlled manner. As an imbalance in this dual function of the intestinal immune system may lead to disease, it is important to understand the underlying functional mechanisms.

The overall goal of this thesis work is to understand the mechanisms of intestinal T cell plasticity and tissue adaptation. We sought to investigate inter-and intra-environmental differences in $CD4^+$ T cells in order to understand how they adapt to different intestinal tissues. We focused on the environment-driven destabilization of $CD4^+$ T cells and their subsequent development into intraepithelial lymphocytes. To this end, we aimed to identify the transcriptional and chromatin changes that $CD4^+$ T cells undergo during CD4-IEL development and elucidate the functional consequences of this plasticity. Furthermore, we coupled our studies of T cell plasticity with that of T cell receptor stimulation and repertoire to uncover the functional role of the TCR in tissue adaptation. Overall, this work provides mechanistic insight into $CD4^+$ T cell adaptation to intestinal tissues in a microbial-derived antigen dependent manner.

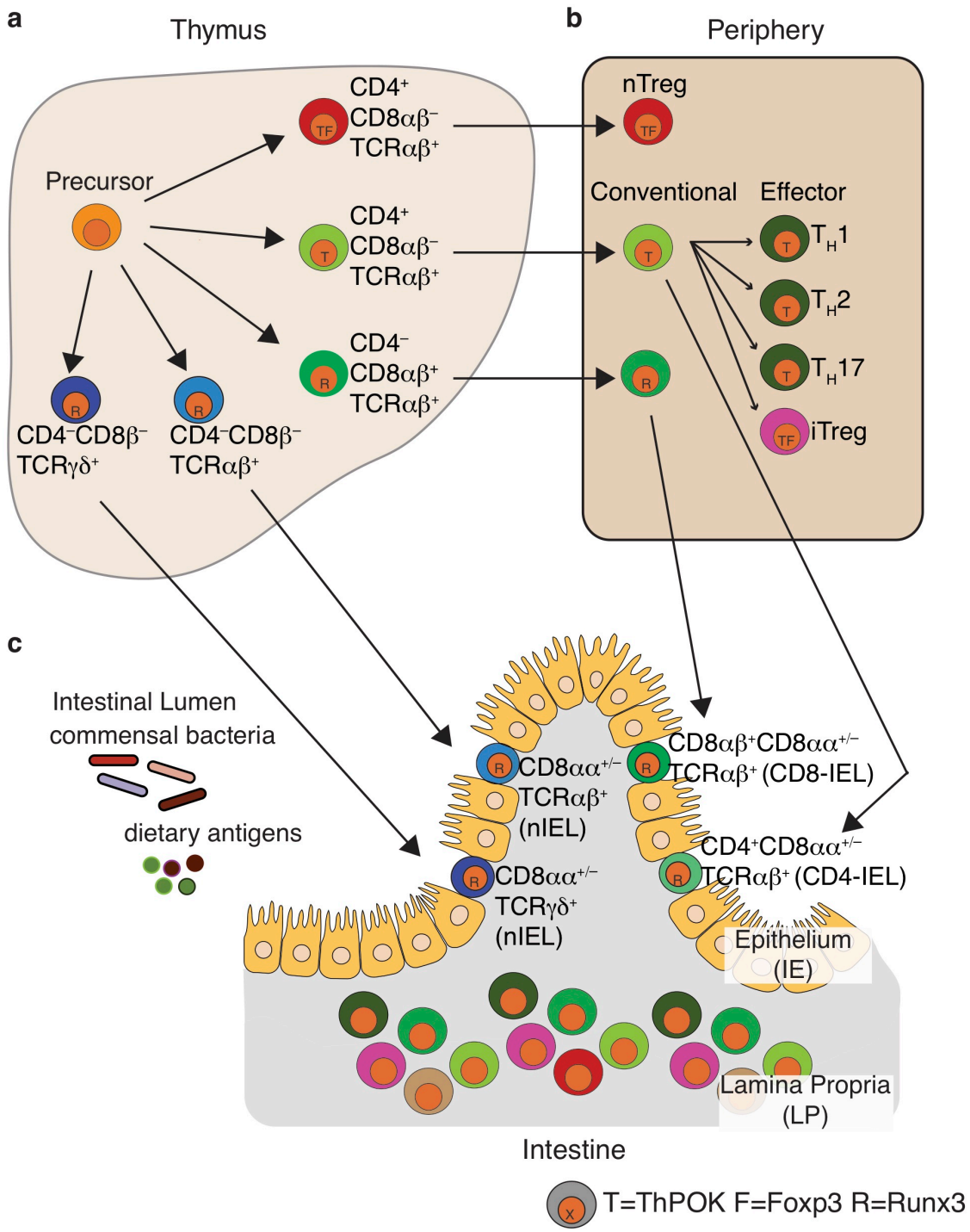



**Figure 1.1 Schematic representation of T lymphocytes.** (**a-c**) Cells include conventional CD4 or CD8 (green), natural or induced regulatory T cells (nTreg; red, iTreg; pink), natural intraepithelial lymphocytes (TCRαβ+ nIEL, TCRγδ+nIEL; blue) and induced intraepithelial lymphocytes (CD8-IEL, CD4-IEL; green). Letter inside cell represents expressed transcription factor (T; ThPOK, F; Foxp3, R; Runx3). Cells within the thymus (**a**), peripheral lymph node (**b**), and intestine (**c**), including lamina propria (LP), intestinal epithelium (IE), and lumen.

# Chapter 2: Tissue adaptation of regulatory and intraepithelial CD4+ T cells controls gut inflammation

## 2.1 Reciprocal localization between Tregs and CD4-IELs in intestinal tissues correlates with ThPOK levels

The gut mucosa is exposed to large amounts of both harmless and potentially pathogenic stimuli on a daily basis, hence diverse immune regulatory mechanisms must operate to avoid inflammatory diseases [122]. Peripherally induced Foxp3-expressing regulatory T cells (iTregs) mediate suppression of a variety of immune cells and actively prevent inflammatory bowel diseases and food allergies [45, 51, 123, 124, 125, 126]. Similar to iTregs, Foxp3$^{-}$CD8$\alpha\alpha^{+}$CD4$^{+}$ intraepithelial lymphocytes (CD4-IELs) depend on retinoic acid (RA) and transforming growth factor (TGF)-$\beta$ signaling for their development and also have anti-inflammatory properties [51, 66, 67, 127, 128, 129, 130]. However, while CD4-IELs accumulate at the intestinal epithelium (IE), very few total Tregs (including iTregs and thymically-derived natural Tregs) can be found at this site (**Figure 2.1a, b**). We asked whether and how the intestinal environment segregates iTregs and CD4-IELs, the transcriptional control involved in this regulation and how this balance affects gut inflammation.

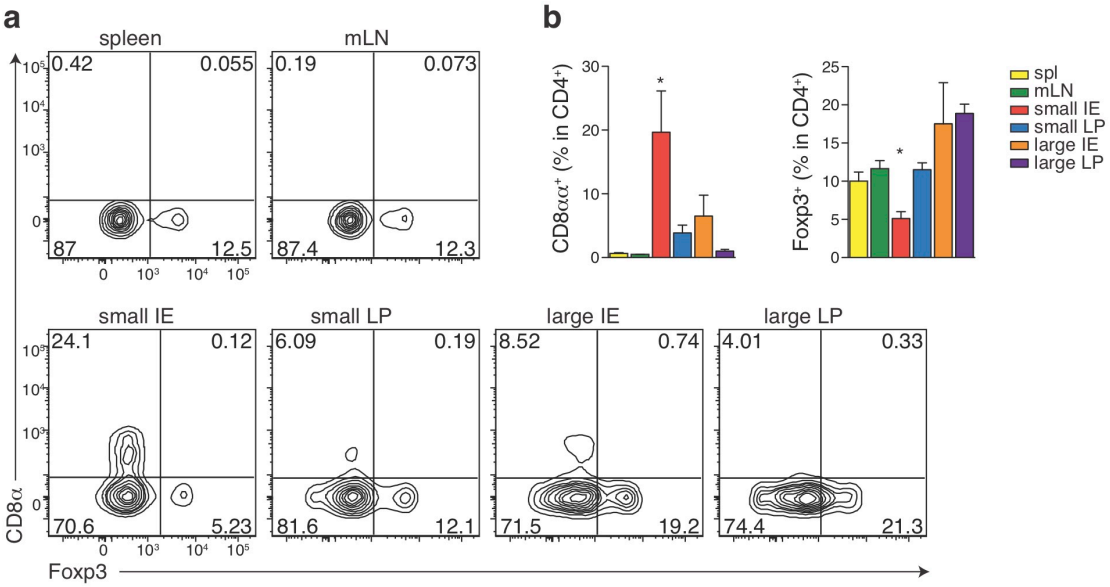



**Figure 2.1 Few Tregs reside in the intestinal epithelium.** (**a-b**) Flow cytometry analysis of surface CD8$\alpha$ and intracellular Foxp3 expression among TCR$\beta^{+}$CD4$^{+}$CD8$\beta^{-}$ cells from from spleen (spl), mesenteric lymph node (mLN), small intestine epithelium (IE) and lamina propria (LP) or large intestine IE and LP in WT mice. (**a**) Representative flow cytometry contour plots. (**b**) Frequency of CD8$\alpha^{+}$ cells (left) or Foxp3$^{+}$ cells (right). Data are expressed as mean ± SD of individual mice (n=6).

We used intravital multi-photon microscopy (IVM) to investigate whether Tregs are actively excluded from the gut epithelium. For tracking *in vivo* Treg dynamics, we utilized tamoxifen-inducible *Foxp3*$^{CreER\text{-}eGFP}$:*Rosa26*$^{lsl\text{-}tdTomato}$ (i*Foxp3*$^{Tomato}$) mice [131], which allow Treg fate mapping and the distinction between cells that currently express (Tomato$^+$Foxp3$^+$) and cells that once expressed Foxp3 (Tomato$^+$Foxp3$^-$). We compared Treg movement patterns in these mice shortly (24 hours) after tamoxifen administration, allowing for the visualization of *bona fide* Tregs (Tomato$^+$) and thymically-derived γδIEL cell movement patterns, which in the gut, reside exclusively in the epithelium [132]. We found that more than 80% of TCRγδ$^{GFP}$ cells preferentially remained in the epithelium, while 68% and 14% of Tomato$^+$ cells were considered lamina propria and IE residents, respectively, throughout the duration of image recordings (**Figure 2.2a, b**). Among the remaining (~18%) migrating Tomato$^+$ cells, cell tracking indicated that cells moved into the epithelial layer from the lamina propria more frequently than vice-versa (**Figure 2.2b, c**). Because *ex vivo* analysis of the intestinal epithelium revealed a low frequency of Foxp3$^+$ Tregs (**Figure 2.1a**), these IVM results suggested either preferential cell-death of Tregs or their conversion into another T cell subtype.

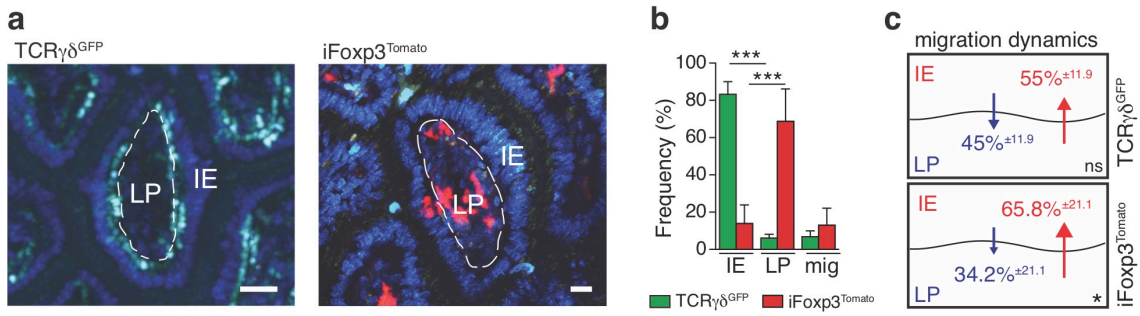



**Figure 2.2 Tregs and IELs exhibit reciprocal localization in the intestinal tissues.** (**a-c**) Intravital microscopy (IVM) analysis of ileal villi of TCRγδ$^{GFP}$ or i*Foxp3*$^{Tomato}$ mice, 24 hours after tamoxifen administration. Mice were injected with Hoechst prior to imaging to visualize all nuclei (blue). Scale bar=10μm. (**a**) Time-stacked image of TCRγδ$^{GFP}$ (green channel; left) mice and i*Foxp3*$^{Tomato}$ (red channel; right) mice. Images are representative of 20-22 villi from at least 3 independent experiments. (**b-c**) Frequency of intraepithelial (IE), lamina propria (LP) or migratory TCRγδ$^{GFP}$ and iFoxp3$^{Tomato}$ cells. (**c**) Percentages within migrating cells. Cells were tracked using *Imaris* software (Bitplane UK). Data are expressed as mean ± SD from independent movies. ***$P$<0.001 (Student's $t$ test).

While all Tregs express the CD4-lineage transcription factor *Zbtb7b* (encoding ThPOK), CD8αα$^+$CD4$^+$ and more than 50% of Foxp3$^-$CD8α$^-$CD4$^+$ cells in the small intestine epithelium lack ThPOK expression (**Figure 2.3a**). We thus asked whether down-modulation of ThPOK influences CD4$^+$ T cell dynamics. We performed IVM using an ovalbumin (OVA)-specific T cell receptor (TCR) transgenic system (OT-II), in which oral OVA exposure leads to incremental ThPOK loss by intestinal CD4$^+$ T cells [67, 133]. To allow for the visualization of T cells upon ThPOK loss, we crossed

OT-II (*Rag1*$^{-/-}$) with *Thpok*$^{GFP}$ knock-in reporter and ubiquitous mRFP1 mice. Sorted naïve (RFP$^{+}$GFP$^{+}$) CD4$^{+}$ T cells from OT-II (RFP *Thpok*$^{GFP}$) mice were transferred to *Rag1*$^{-/-}$ mice kept on an OVA-containing diet for one week before IVM analysis. Computational tracking revealed that roughly 60% of the transferred cells that downregulated ThPOK (RFP$^{+}$GFP$^{-}$), and only 20% of ThPOK$^{high}$ cells (RFP$^{+}$GFP$^{+}$) remained in the epithelial layer (**Figure 2.3b, c**). Migrating ThPOK$^{high}$ cells showed similar movement patterns to Treg cells, with preferential displacement from the lamina propria into the epithelial layer, suggesting that part of these cells convert into ThPOK$^{low}$ cells, or die, in that compartment (**Figure 2.3d**). These observations indicate that loss of ThPOK corresponds to an IEL-like behavior in CD4$^{+}$ T cells. Additionally, the discrepancy between the capacity of Tregs to visit the intestinal epithelium and their low frequency suggests that this environment may favour Treg plasticity.

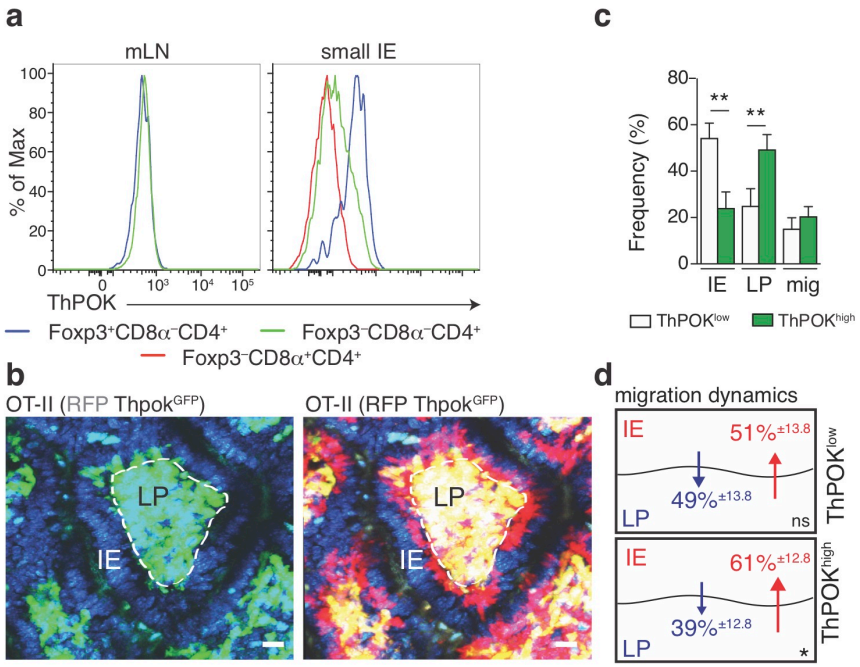



**Figure 2.3 ThPOK levels correlate with reciprocal Treg and CD4-IEL localization and migration dynamics in the intestine.** (**a**) Representative flow cytometry histograms of intracellular ThPOK expression by Foxp3$^{+}$CD8αβ$^{-}$, Foxp3$^{-}$CD8αβ$^{-}$, Foxp3$^{-}$CD8α$^{+}$CD8β$^{-}$ (TCRβ$^{+}$CD4$^{+}$) cells from WT mice. (**b-d**) Sorted naïve CD4$^{+}$ T cells from OT-II (RFP *Thpok*$^{GFP}$) mice were transferred to *Rag1*$^{-/-}$ mice and recipient mice were fed an OVA-containing diet for 7 days before ileal intravital microscopy analysis. Mice were injected with Hoechst prior to imaging to visualize all nuclei (blue). Scale bar=10μm. (**b**) Time-stacked image of GFP$^{+}$ (green channel and blue channel overlay; left) and GFP$^{+}$ (yellow) and GFP$^{-}$ (red) cells (green, red and blue channel overlay; right). Time-stacked images are representative of at least 50 villi from 4 independent experiments. (**c**) Quantification of tracked GFP$^{+}$ and GFP$^{-}$ cell dynamics from 4 different movies (total 6-paired villi) in 2 independent experiments. (**d**) Percentages within migrating cells. Cells were tracked using *Imaris* software (Bitplane UK). Data are expressed as mean ± SD from independent movies., $^{**}P<0.01$ (Student's *t* test).

## 2.2 CD4-IEL levels in the intestinal epithelium are microbiota-dependent

The abundance of CD4-IELs varies with age as well as diet and depends on the indigenous microbiota[134 135 136]. We found that the reciprocity between Tregs and CD4-IELs within the intestinal epithelium maintained with differences in age and microbiota. The frequency of CD4-IELs among total CD4$^+$ T cells in aged mice (up to 75 weeks old) was much higher than in adult (8-16 weeks old) mice, while the relative Treg frequency was lower (**Figure 2.4a**). Furthermore, adult animals housed in different facilities (RU or MSKCC), harboring distinct microbiotas, also displayed reciprocal frequencies of CD4-IELs and Tregs, further suggesting that although the microbial composition influences CD4-IELs, the reciprocity between Tregs and CD4-IEL frequencies is largely maintained (**Figure 2.4b**). Upon antibiotic treatment (vancomycin and metronidazole), animals from both facilities displayed a decline in CD4-IEL frequencies, although only MSKCC -derived mice showed a reciprocal increase in Tregs upon antibiotic treatment (**Figure 2.4b**). This data highlights the microbiota contribution to development and/or maintenance of CD4-IELs.

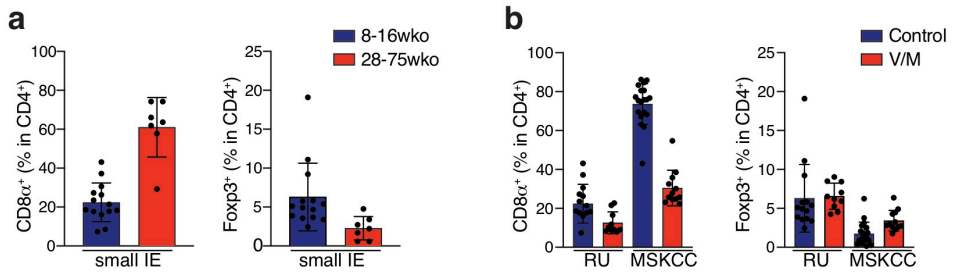



**Figure 2.4 The microbiome influences CD4-IEL frequencies.** (**a-b**) Frequencies of CD4-IELs (CD8$\alpha^+$CD4$^+$CD8$\beta^-$) and Tregs (Foxp3+CD8$\alpha^-$) among CD4$^+$ T cells in the small intestine epithelium (IE). (**a**) CD4-IEL (left) and Treg (right) frequencies of adult mice (8-16 weeks old; blue) and older mice (28-75 weeks old; red) housed in the RU facility. (**b**) CD4-IEL (left) and Treg (right) frequencies of adult mice born and housed until adult age in either RU or MSKCC animal facilities. Mice were either maintained on control drinking water (blue) or were given vancomycin and metronidazole antibiotics (V/M; red) for 4 weeks prior to analysis. (n=6-14).

Commensal bacteria has been shown to play a major role in the induction of large intestine lamina propria iTregs [123, 124, 125, 126, 137]. In contrast, we observed an increased frequency of iTregs (Neuropilin-1$^-$ Foxp3$^+$) in the small intestinal epithelium isolated from germ-free (GF) mice when compared to specific pathogen-free (SPF) controls (**Figure 2.5a**). The total Treg number in the epithelial compartment was comparable between GF and SPF mice. In agreement with the Treg-IEL reciprocity observed above, GF mice showed an almost ten-fold

reduction in the number of intraepithelial CD4+ T cells (**Figure 2.5a**). Moreover, consistent with a reciprocal ThPOK or Foxp3 expression and CD4-IEL differentiation, we found an increased frequency of ThPOK$^{high}$ CD4+ T cells and significantly reduced CD4-IELs in GF mice (**Figure 2.5b**).

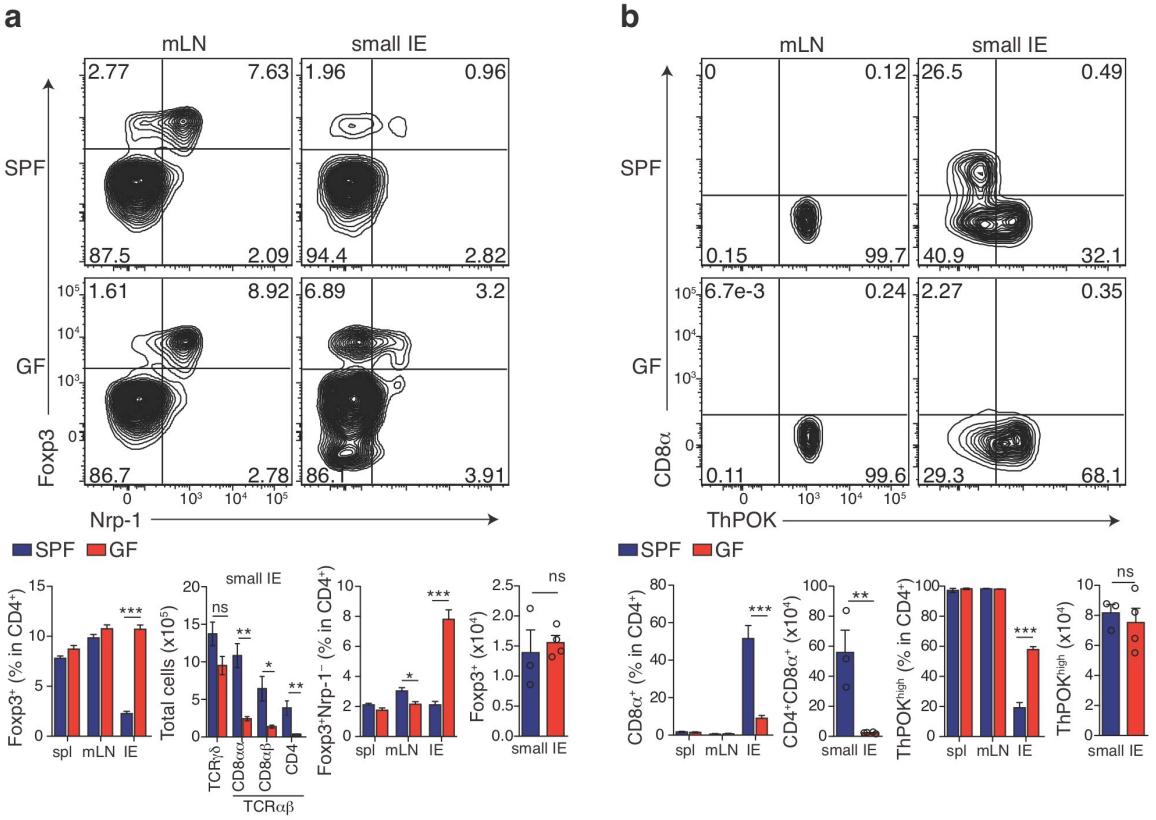



**Figure 2.5 Microbiota-dependent CD4-IEL levels in the intestinal epithelium.** (**a-b**) Flow cytometry analysis of TCRβ+CD4+CD8β− lymphocytes from spleen (spl), mesenteric lymph nodes (mLN) and small intestine epithelium (IE) of wild-type C57BL/6 mice maintained under specific pathogen-free (SPF) or germ-free (GF) conditions. (**a**) Surface neuropilin-1 (Nrp-1) and intracellular Foxp3 expression (top). Bar graphs represent frequency and total number of Foxp3+ cells (left) or Foxp3+Nrp-1− cells (iTregs) (right). Total cell number for T cell populations isolated from the sIE is also shown (middle). (**b**) Surface CD8α and intracellular ThPOK expression (top). Bar graphs represent frequency and total number of CD8α+ (left) or ThPOK$^{high}$ (right). Data expressed as mean ± SD of individual mice (n=6). **P*<0.05, ***P*<0.01, ****P*<0.001 (Student's *t* test).

## 2.3 Lineage-defining transcription factors play a key role in CD4-IEL development

We interrogated whether the reciprocal properties of iTregs and CD4-IELs were influenced by lineage-defining transcription factors. We first determined the role of ThPOK by crossing *Thpok*$^{fl/fl20}$ with *Cd4*$^{CreER}$ mice[138] for *Thpok* depletion upon

tamoxifen treatment[133]. *In vivo* administration of tamoxifen to i*Cd4*(Δ*Thpok*) mice [133] one week before analysis resulted in a 32% to 48% reduction in the frequency of Foxp3$^+$ Tregs in lymphoid and intestinal tissues, respectively, whereas an increase in CD4-IELs was observed only in intestinal tissues (**Figure 2.6a,b**).

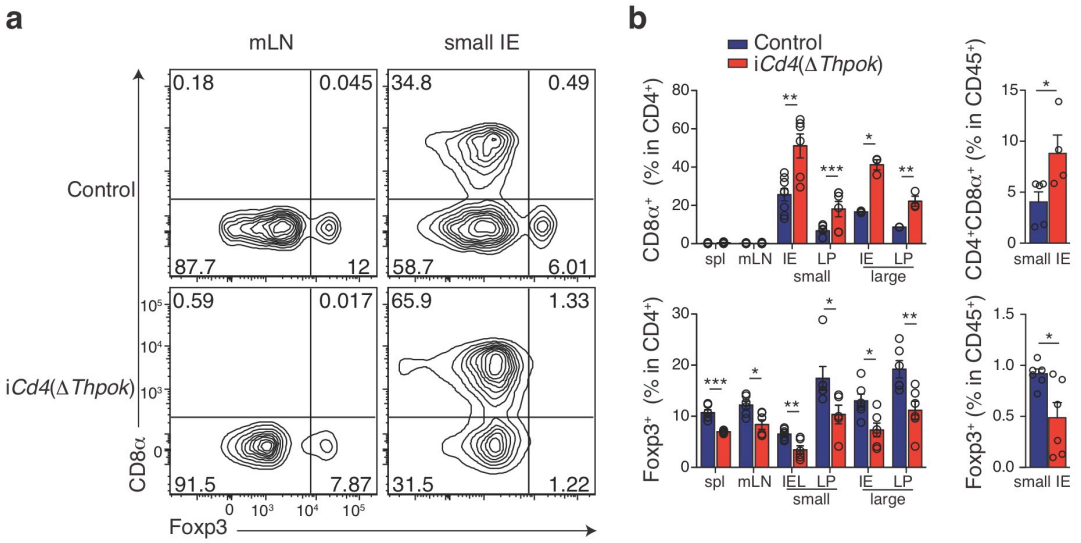



**Figure 2.6 ThPOK expression by intestinal epithelial CD4$^+$ T cells plays a key role in the reciprocal development of Tregs and CD4-IELs.** (**a-b**) Flow cytometry analysis of lymphocytes from spleen (spl), mesenteric lymph nodes (mLN), small and large intestine epithelium (IE) and lamina propria (LP) of inducible conditional *Thpok*-deficient mice: *Cd4*$^{CreER}$:*Thpok*$^{fl/fl}$ (i*Cd4*(Δ*Thpok*)) and (Cre$^-$) *Thpok*$^{f/f}$ littermate controls 7 days post-tamoxifen administration. (**a**) Representative contour plot for surface CD8α and intracellular Foxp3 among TCRβ$^+$CD4$^+$CD8β$^-$ cells. (**b**) Frequency of CD8α$^+$ (upper) and Foxp3$^+$ (lower) among TCRβ$^+$CD4$^+$CD8β$^-$ or among total CD45$^+$ cells in the indicated tissues. Data expressed as means ± SD of individual mice (n=4-5). **P*<0.05, ***P*<0.01, ****P*<0.001 (Student's *t* test).

We sought to target the CD8- lineage defining transcription factors, which have also been described to play a positive role in CD4-IEL development. We generated mice with conditional deletion of *Runx3* or *Tbx21* (encoding T-bet), which mediate downregulation of ThPOK in developing CD4-IELs [67, 127]. Analysis of the epithelial compartment of *Cd4*(Δ*Runx3*) mice revealed a reduction in the frequency and number of CD4-IELs and, conversely, enrichment in Tregs (**Figure 2.7a, b**). Next, we analyzed mice in which *Tbx21* was excised early (driven by *Ox40*$^{Cre}$ [139]), or late (driven by E8I$^{Cre}$ [140]) during CD4-IEL differentiation [133]. *Ox40*(Δ*Tbx21*) mice also showed reduced CD4-IEL and roughly a two-fold increase in Treg in the epithelium, while E8I(Δ*Tbx21*) mice showed reduced CD4-IELs, but normal numbers of Tregs in the epithelium when compared to Cre$^-$ control mice (**Figure 2.7a, b**). Collectively, this data provide a possible mechanism for the reduced

number of Tregs in the gut epithelium, where collaboration between Runx3 and T-bet results in downmodulation of ThPOK in CD4+ T cells [67, 127].

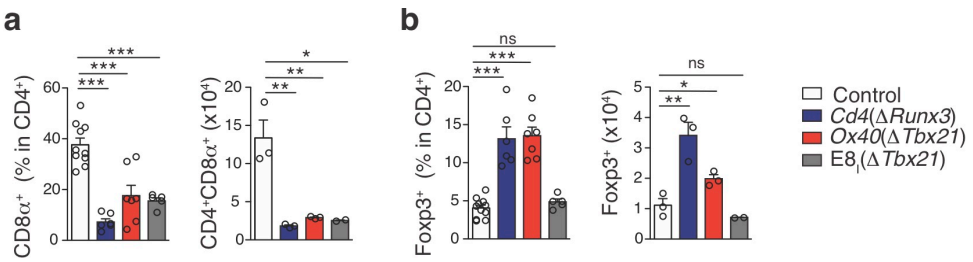



**Figure 2.7 Runx3 and T-bet expression by intestinal epithelial CD4+ T cells play a key role in the reciprocal development of Tregs and CD4-IELs.** (**a-b**) Frequency (left) and total number(right) of CD8α+ (**a**) and Foxp3+ (**b**) among TCRβ+CD4+CD8β− cells from small intestinal epithelium of *Ox40*(Δ*Tbx21*), *Cd4*(Δ*Runx3*), E8I(Δ*Tbx21*), and wild-type (control) mice. Data expressed as mean ± SD of individual mice (n=3-9). **P*<0.05, ***P*<0.01, ****P*<0.001 (One-way ANOVA with Tukey post-test).

## 2.4 Microbiota contributes to Treg plasticity in the intestinal epithelium

To directly address Treg plasticity in the gut tissue, we performed Treg fate mapping using naïve adult *Foxp3*$^{Cre\text{-}YFP}$:*Rosa26*$^{lsl\text{-}DsRed}$ (*Foxp3*$^{DsRed}$) mice [141]. Analysis of peripheral lymphoid tissues isolated from *Foxp3*$^{DsRed}$ mice revealed an almost complete concurrence between Foxp3 reporter (YFP) and DsRed expression, confirming previously described stability of the Treg lineage [131, 142]. However, over 40% of DsRed+ CD4+ T cells in the small intestine epithelium did not express Foxp3 (YFP−), indicating that a considerable number of Tregs lost Foxp3 expression (**Figure 2.8a**). Because previous studies have demonstrated that the majority of "ex-Foxp3" cells in steady state were derived from uncommitted precursors that transiently upregulated Foxp3 [142], we also performed fate mapping after pulse-labeling i*Foxp3*$^{Tomato}$ mice with tamoxifen [131], a strategy more likely to target *bona fide* Tregs [143]. Nevertheless, while stable Foxp3 expression was again observed in several peripheral tissues examined, over 50% of Tomato+ CD4+ T cells that accumulated in the small intestine and almost 10% in the large intestine epithelium isolated from i*Foxp3*$^{Tomato}$ mice no longer expressed Foxp3 five weeks post tamoxifen administration (**Figure 2.8b,c**). The contribution of Tomato+ cells to the CD8αα+ and CD8αβ+ CD4-IEL pools was roughly 10% and 25%, respectively (**Figure 2.8d**). Consistent with a ThPOK-dependent process, ex-Tregs that underwent IEL differentiation showed low levels of ThPOK (**Figure 2.8e**). These results indicate that a substantial proportion of intestinal Tregs physiologically convert to CD4-IELs.

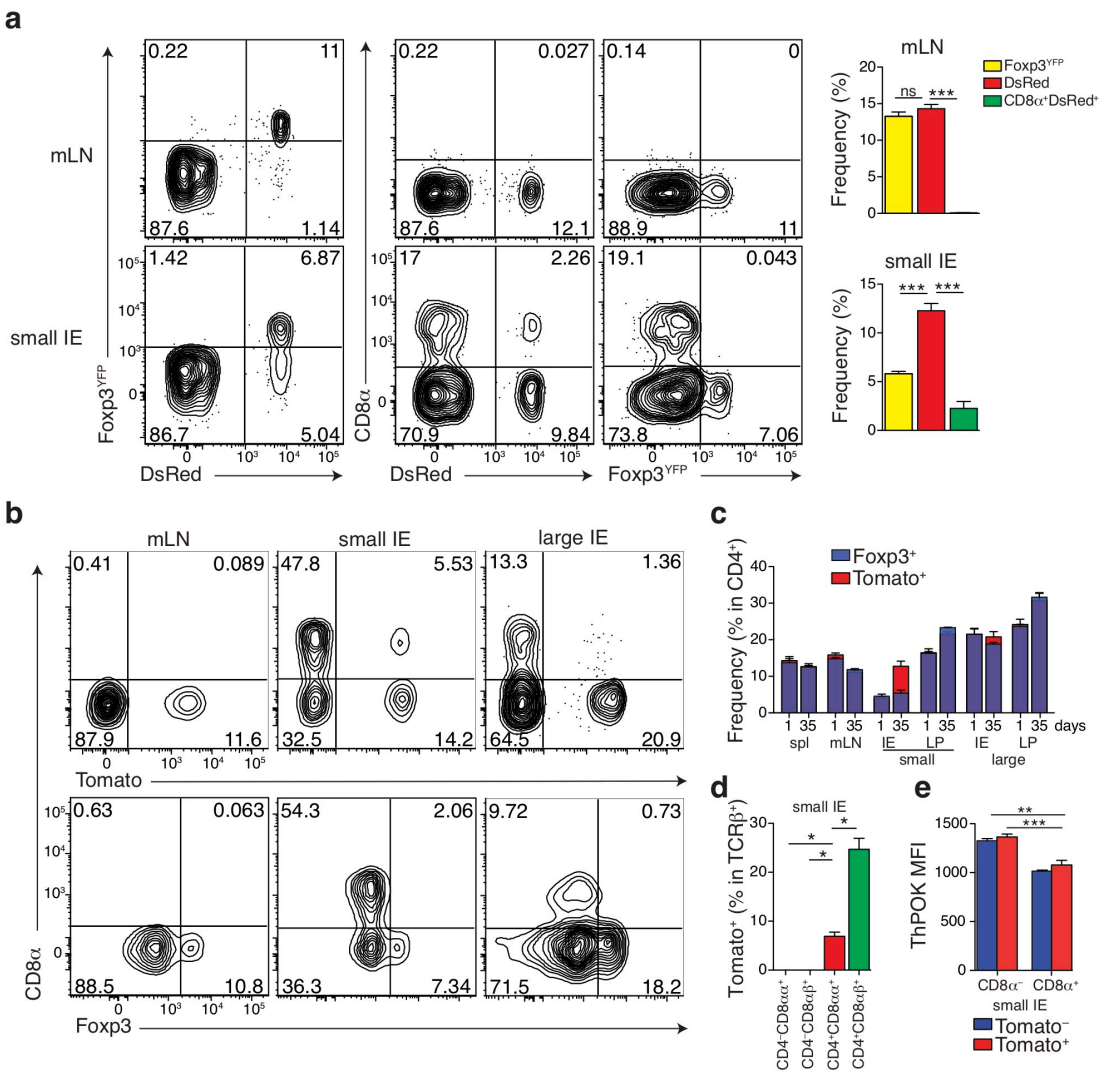



**Figure 2.8 A subset of Tregs converts to CD4-IELs in the intestinal epithelium.** (**a**) Flow cytometry analysis of mesenteric lymph nodes (mLN) and small intestine epithelium (IE) of *Foxp3*$^{\text{Cre-YFP}}$:*Rosa26*$^{\text{lsl-DsRed}}$ (*Foxp3*$^{\text{DsRed}}$) mice, which constitutively labels cells that currently or once expressed Foxp3. Representative plot for YFP, CD8α and DsRed among TCRβ$^{+}$CD4$^{+}$CD8β$^{-}$ cells. Bar graphs represent frequency of indicated cell types. (**b-e**) Flow cytometry analysis of spl, mLN, small or large IE and lamina propria (LP) of *Foxp3*$^{\text{CreER-eGFP}}$:*Rosa26*$^{\text{lsl-tdTomato}}$ (i*Foxp3*$^{\text{Tomato}}$) mice 1 day (shown in c) and 35 days post-tamoxifen administration (**b**) Surface CD8α and Tomato expression or intracellular Foxp3 among TCRβ$^{+}$CD4$^{+}$ cells. (**c**) Frequency of Foxp3$^{+}$ (blue) and Tomato$^{+}$ (red) among TCRβ$^{+}$ CD4$^{+}$ cells in the indicated tissues. (**d**) Frequency of Tomato$^{+}$ cells among TCRβ$^{+}$ CD4$^{-}$CD8α$^{+}$β$^{-}$, CD4$^{-}$CD8αβ$^{+}$, CD4$^{+}$CD8α$^{+}$β$^{-}$, and CD4$^{+}$CD8αβ$^{+}$ sIELs. (**e**) MFI for intracellular ThPOK expression by Tomato$^{-}$ (blue) or Tomato$^{+}$ (red) cells among TCRβ$^{+}$CD4$^{+}$ CD8α$^{-}$ or CD8α$^{+}$ sIELs. Data expressed as mean ± SD of individual mice (n=5) **P*<0.05, ***P*<0.01, ****P*<0.001 (Student's *t* test).

We next asked if the microbiota influences Treg plasticity similarly to CD4-IEL development and/or maintenance. To address this question, we treated i*Foxp3*$^{Tomato}$ mice with broad-spectrum antibiotics for five weeks, immediately after tamoxifen exposure. We observed that microbiota depletion prevented Foxp3 loss within the Tomato$^+$ CD4$^+$ T cell population, resulting in an accumulation of Tregs in the epithelial compartment (**Figure 2.9a**). The direct contribution of microbial metabolites versus microbial or dietary antigens to the differentiation of CD4-IELs or iTregs occupying the small intestinal epithelium [144] remains to be fully determined. Nevertheless, provision of a TCR ligand can overcome the strict microbiota requirement for CD4-IEL differentiation, as demonstrated by the relatively normal CD4-IEL differentiation in antibiotic-treated OT-II *Thpok*$^{GFP}$ *Rag1*$^{-/-}$ mice exposed to oral OVA (**Figure 2.9b**). Along with the above data, these observations substantiate a microbial-induced plasticity of Treg cells in the epithelium, corroborating intra-tissue specialization and conversion of gut CD4$^+$ T cells.

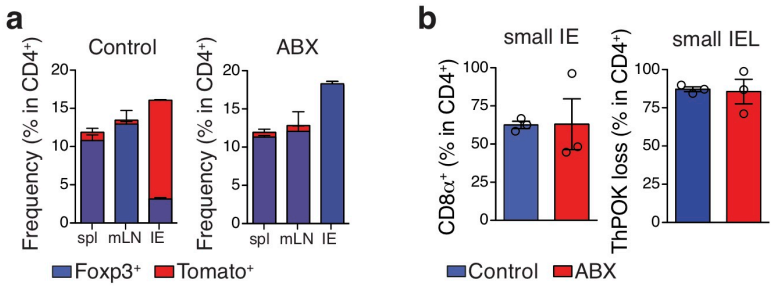



**Figure 2.9 Microbiota-dependent plasticity of Tregs in the intestinal epithelium.** (**a**) Frequency of Foxp3$^+$ (blue) and Tomato$^+$ (red) among TCR$\beta^+$ CD4$^+$ cells from spl, mLN and small intestine IEL of *Foxp3*$^{CreER\text{-}eGFP}$:*Rosa26*$^{lsl\text{-}tdTomato}$ (i*Foxp3*$^{Tomato}$) mice treated with tamoxifen and maintained with a mix of broad range antibiotics (ABX) containing vancomycin, ampicillin, neomycin, metronidazole and sucralose-base sweetener or sucralose (Control) for 5 weeks. (**b**) OT-II-*Rag1*$^{-/-}$ animals were treated with a mix of broad range ABX or sucralose (Control) for 4 weeks. After treatment, animals were fed 1% OVA diet for 21 days prior to flow cytometry analysis for CD8$\alpha$ (left) and ThPOK (right) expression. Data expressed as means ± SD of individual mice (n=3). *$P$<0.05, **$P$<0.01, ***$P$<0.001 (Student's *t* test).

## 2.5 ThPOK downmodulation plays a key role in Treg plasticity and conversion to CD4-IELs

We next asked whether disrupting ThPOK and Runx3 could directly affect the Treg to IEL progression. To simultaneously fate-map Tregs and abrogate their ThPOK and/or Runx3 expression, we crossed i*Foxp3*$^{Tomato}$ mice with ThPOK$^{fl/fl}$ and/or

*Runx3*$^{fl/fl}$ mice[20, 145] and administered tamoxifen to the i*Foxp3*$^{Tomato}$(ΔThPOK, ΔRunx3, and ΔThPOK ΔRunx3 vs WT) mice over 10 weeks. Removal of ThPOK promoted the destabilization of Tregs and their acquisition of the CD4-IEL phenotype, as determined by CD8α expression by Treg–fate-mapped cells, suggesting that the forced deletion of ThPOK leads to premature differentiation of IELs from Tregs (**Figure 2.10a, b**). Concomitant *Thpok* and *Runx3* ablation in part abrogated the effect of *Thpok* deletion, confirming a positive role of Runx3 in CD4-IEL conversion[67]. *Runx3* ablation in Tregs alone did not play a significant role in Treg stability or CD8α-acquisition, implying that it acts primarily in the absence of ThPOK at this stage (**Figure 2.10a, b**). In contrast to cells in the epithelium, all mLN Tregs remained stable, highlighting the finding that Treg plasticity is tissue–dependent (**Figure 2.10a, b**). Furthermore, we did not observe any overt signs of disease, suggesting that the remaining stable Tregs, together with the augmented CD4-IELs, were sufficient to maintain intestinal homeostasis.

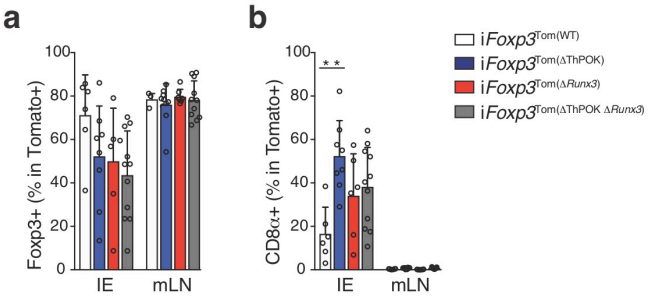



**Figure 2.10. ThPOK downmodulation leads to CD8α induction in intraepithelial Tregs.** (**a**-**b**) Flow cytometry analysis of CD45$^+$TCRβ$^+$CD4$^+$CD8β$^{low}$ Tomato$^+$ cells in the small intestinal epithelium (IE) and mesenteric lymph nodes (mLN) of *Zbtb7b*$^{fl/+}$ x *Runx3*$^{fl/+}$ x *Rosa26*$^{lsltdTomato}$ x *Foxp3*$^{CreER}$(i*Foxp3*), i*Foxp3* x *Runx3*$^{fl/+}$ (i*Foxp3*$^{(\Delta ThPOK)}$), i*Foxp3*x *Zbtb7b*$^{fl/+}$x *Runx3*$^{fl/fl}$ (i*Foxp3*$^{(\Delta Runx3)}$), i*Foxp3* x *Zbtb7b*$^{fl/fl}$ x *Runx3*$^{fl/fl}$ (i*Foxp3*$^{(\Delta ThPOK\ \Delta Runx3)}$) mice after 10 weeks of tamoxifen treatment. (**a**) Frequency of total Foxp3$^+$ cells or (**b**) total CD8α$^+$ cells. Data are expressed as mean +/- SEM of individual mice (n=5-11 per genotype). **p<0.01, [one-way ANOVA and Bonferonni test].

As previously noted, one key difference between mLN and IE Tregs is that while the former mostly consists of natural Tregs (nTregs), the latter is predominantly composed of less stable peripherally induced Tregs (iTregs)[45, 146] (**Figure 2.11a**). Thus, we next examined if potential effects by ThPOK and Runx3 on Treg plasticity in our *in* vivo results were masked by differences in nTreg vs iTreg stability. Moreover, to dissect the functional roles of ThPOK and Runx3 in these two types of Tregs, we used the classical model of T cell-transfer colitis in *Rag1*-deficient mice[147, 148, 149]. Mice were co-transferred with naïve CD45.1 CD4$^+$ T cells along with peripheral CD45.2 Tom$^+$ nTregs or iTregs (neuropilin-1 high or low, respectively) sorted from iFoxp3$^{Tomato}$ (ΔThPOK, ΔRunx3, ΔThPOK ΔRunx3 vs

WT) mice (**Figure 2.11b**). While iTregs lost more Foxp3 expression than nTregs as expected, both nTregs and iTregs were stable in the mLN but less so in the IE compartment (**Figure 2.11c, d**). We again found that ThPOK abrogation led to an increased frequency of ex-Treg-IELs but without a significant loss of Foxp3-expressing cells in comparison to the mice transferred with WT iTregs; a result likely due to an increase in Foxp3-CD8$\alpha$ double positive cells (**Figure 2.11e**). Moreover, concomitant deletion of *Runx3* and *Thpok* did not result in increased accumulation of CD8$\alpha$-expressing cells. In the presence of ThPOK, Runx3 appeared to be stabilizing to iTregs, reinforcing previously proposed roles for this TF in Tregs[150, 151]. Furthermore, we did not observe significant weight loss or intestinal inflammation as shown by low fecal lipocalin-2 levels (**Figure 2.11f, g)**. Thus, both remaining Tregs and/or Tregs that differentiated into CD4-IELs were able to confer protection against colitis. Overall, while deletion of ThPOK was not sufficient to destabilize nTregs as shown by Foxp3 expression, in iTregs it led to the development of CD4-IELs in the IE (**Figure 2.11d, e**). Overall, we conclude that ThPOK down-regulation in both conventional CD4$^+$ T cells and iTregs leads to the development of CD4-IELs. However, conventional CD4$^+$ T cells (Tconvs) and iTregs may have different requirements for the CD8-lineage defining TF *Runx3* in CD4-IEL development.

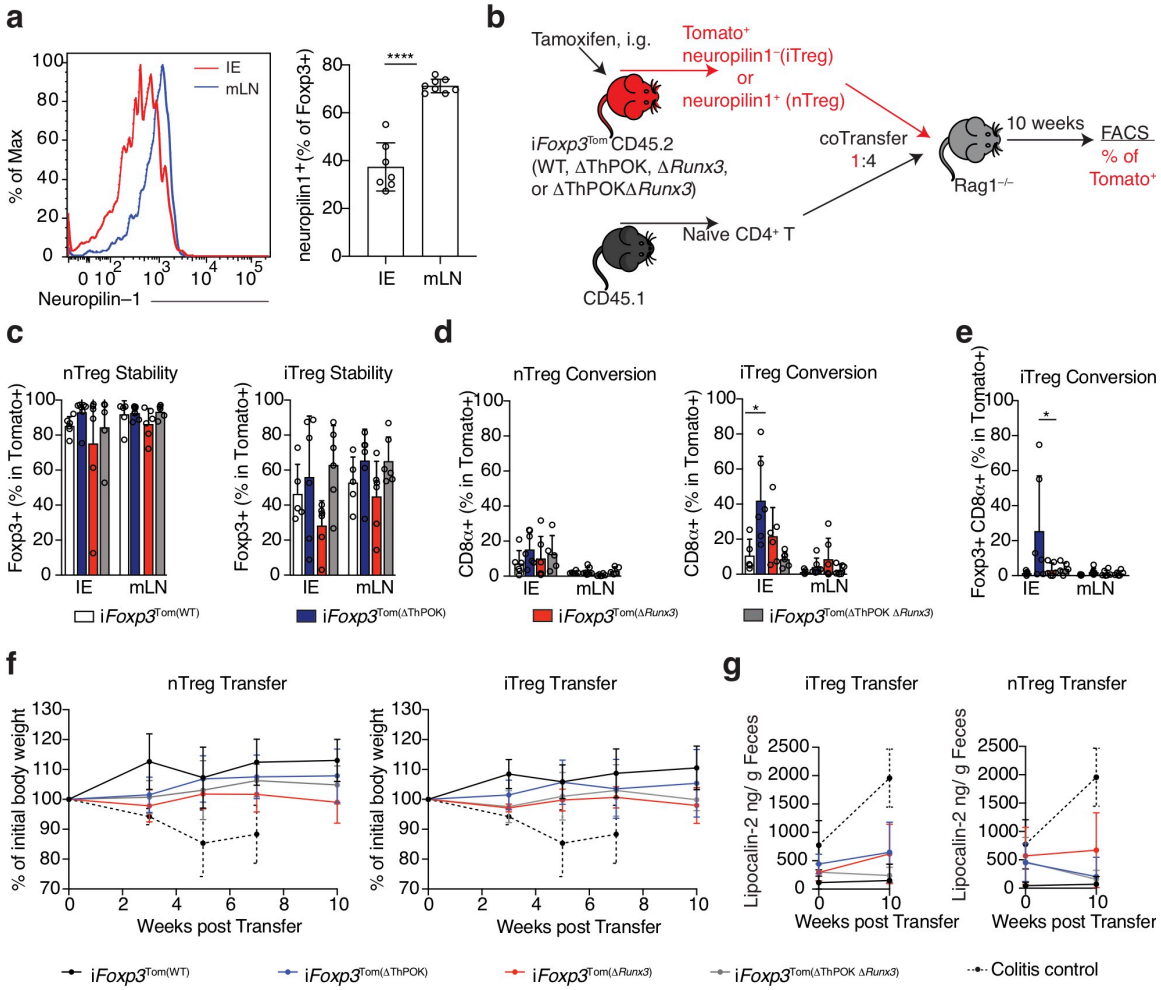
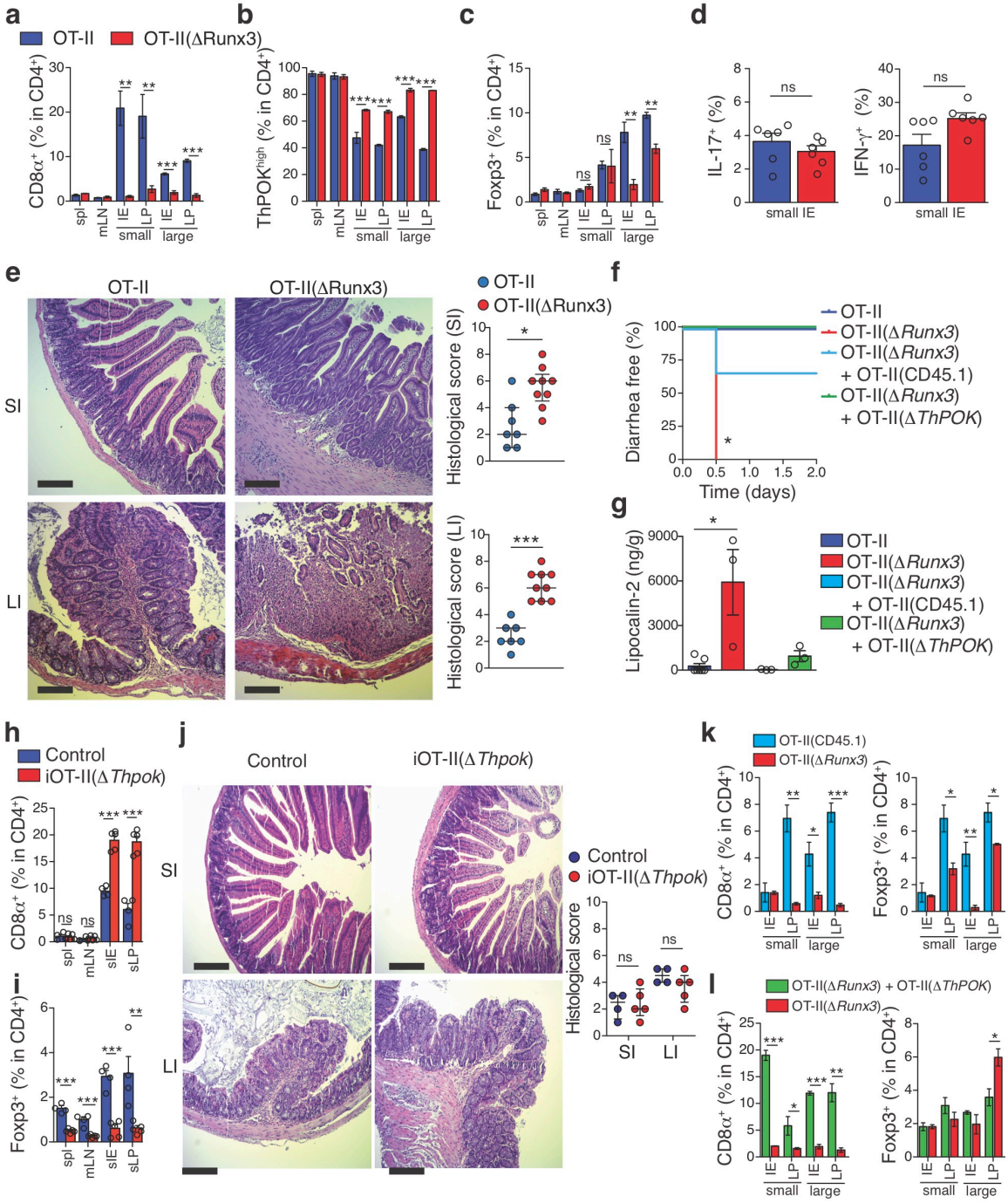



**Figure 2.11. Impact of ThPOK and Runx3 on Treg stability and function.** (**a**) Neuropilin-1 expression among Foxp3+ Tregs in the intraepithelial compartment (IE) and mesenteric lymph nodes (mLN) of WT mice. Histogram (left) and frequency (right). (**b-g**) nTreg (neuropilin-1−) or iTregs (neuropilin-1+) were sorted from spleens and mLNs of CD45.2 *Zbtb7b*$^{fl/+}$ x *Runx3*$^{fl/+}$ x *Rosa26*$^{lsltdTomato}$ x *Foxp3*$^{CreER}$ (i*Foxp3*), i*Foxp3*x *Runx3*$^{fl/+}$ (i*Foxp3*$^{(\Delta ThPOK)}$), i*Foxp3* x *Zbtb7b*$^{fl/+}$ x *Runx3*$^{fl/fl}$ (i*Foxp3*$^{(\Delta Runx3)}$), i*Foxp3*x *Zbtb7b*$^{fl/fl}$ x *Runx3*$^{fl/fl}$ (i*Foxp3*$^{(\Delta ThPOK\,\Delta Runx3)}$) mice after tamoxifen administration and co-transferred with CD45.1 naïve CD4+ T cells to *Rag1*$^{-/-}$ hosts. CD45.2+TCRβ+CD4+CD8β-Tomato+ lymphocytes from the IE and mLN were analyzed 10 weeks after transfer. (**b**) Experimental layout. (**c- d**) Frequencies of total intracellular Foxp3+ (**c**) and total surface CD8α+(**d**) expressing cells after nTreg (left) or iTreg (right) transfer. (**e**) Frequency of Foxp3+CD8α+ cells after iTreg transfer. (**f-g**) Body weight of *Rag1*$^{-/-}$ recipients after nTreg (left) or iTreg (right) transfers (**f**) and levels of fecal lipocalin-*2* before and after per cell type transferred (**g**). Dashed lines represent colitis control transfer of naïve CD45.1+ CD4+ T cells only. Data are expressed as mean +/- SEM of individual mice (n=5-6 per genotype). *p<0.05, **p<0.01, ***p<0.001 [one-way ANOVA and Bonferonni test].

## 2.6 iTregs and CD4-IELs have complementary roles in regulating local inflammatory responses toward dietary antigens

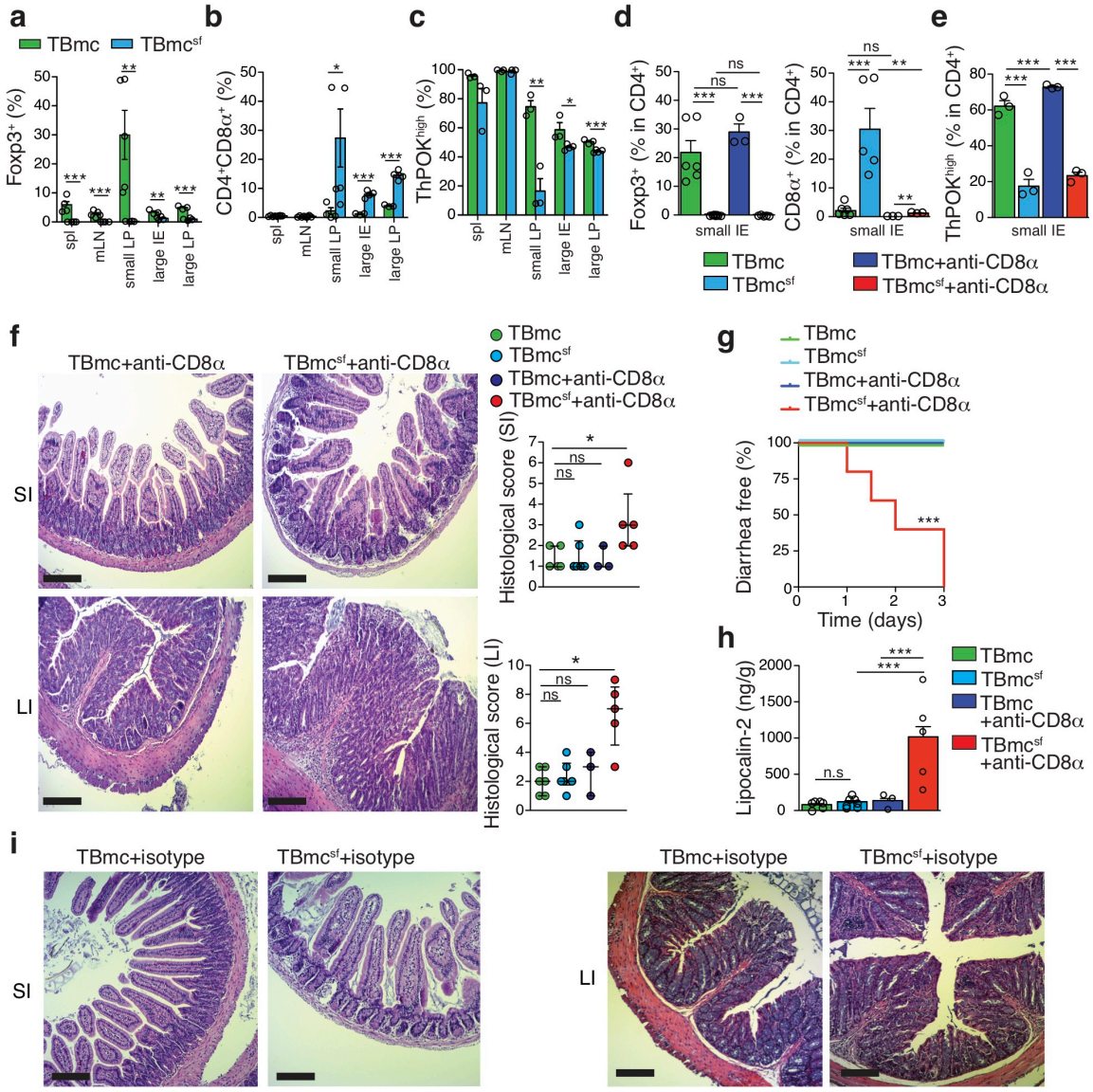
Oral exposure to TCR ligands results in both iTreg and CD4-IEL differentiation in a TGF-β-dependent manner [51, 66, 67, 128, 130, 152]. We asked whether the intra-tissue adaptation of iTregs and CD4-IELs influences the outcome of T cell responses to dietary antigens by employing a transcription factor-based targeting of these lineages in OVA-specific TCR transgenic mice on *Rag1*$^{-/-}$ background [133]. Conditionally targeting *Runx3* in the OT-II model (OT-II(Δ*Runx3*)) prevented ThPOK loss and CD4-IEL differentiation, and also impacted iTreg differentiation in the large intestine while no differences in cytokine production were found when compared to control OT-II mice (**Figure 2.12a-d**). Whereas OVA-challenged control OT-II mice showed few or no signs of intestinal inflammation, OT-II(Δ*Runx3*) mice readily developed diarrhea and severe pathology, as confirmed by faecal lipocalin-2 levels (**Figure 2.12e-g**). We concluded that prevention of ThPOK loss and CD4-IEL differentiation resulted in a local inflammatory response towards dietary antigens, although the reduction in iTregs could also contribute to the exaggerated inflammatory phenotype observed in the large intestine of OT-II(Δ*Runx3*) mice. In contrast, conditionally targeting *Thpok* by administering tamoxifen to OVA-fed iOT-II(Δ*Thpok*) mice [133] resulted in a severe impairment of iTreg development in all tissues examined with a concomitant increase in CD4-IELs in the intestine, while no inflammatory phenotype was observed (**Figure 2.12h-j**). We then tested whether the OT-II(Δ*Runx3*) disease phenotype could be rescued, or prevented by wild-type or *Thpok*-deficient OT-II cells [133]. We found that transferred wild-type CD45.1 OT-II cells, which could differentiate to both iTreg and CD4-IELs, as well as CD4$^+$ T cells from tamoxifen-treated iOT-II(Δ*Thpok*) mice, which could only differentiate to CD4-IELs, rescued the diarrhea and inflammatory phenotype observed in OT-II(Δ*Runx3*) mice (**Figure 2.12f, g, k, l**). These data indicate that a balance in ThPOK levels regulates inflammatory, CD4$^+$ T cell-mediated responses to dietary antigens. Additionally, these results suggest that CD4-IELs exert a cell-extrinsic control of local intestinal inflammation.

**Figure 2.12 CD4-IELs have a regulatory role in local inflammatory response towards dietary antigens.** (**a-d**) OVA-specific TCR transgenic mice (*Rag1*$^{-/-}$ background) were fed an OVA-containing diet for 7 days. *Thpok*$^{GFP}$ OT-II(Δ*Runx3*) and control *Thpok*$^{GFP}$ OT-II mice were analyzed. Bar graphs represent frequency of CD8$\alpha^+$ (**a**), ThPOK$^{high}$ (**b**), intracellular Foxp3 (**c**) and intracellular IL-17A and IFN-$\gamma$ (**d**) among TCRV$\alpha$2$^+$CD4$^+$CD8$\beta^-$ cells from spleen (spl), mesenteric lymph node (mLN), small and large intestine epithelium (IE) and lamina propria (LP), analyzed by flow cytometry. (**e**) Hematoxylin and eosin (H&E) staining of the small intestine jejunum (SI) (top) and the large intestine colon (LI) (bottom). Original magnification, 40x. Graphs represent histological scores of the SI (top) and the LI (bottom) (each symbol represents one mouse). (**f-g**) Sorted naïve OVA-specific TCR transgenic cells (TCRV$\alpha$2$^+$CD4$^+$CD62L$^+$CD44$^{low}$) from wild-type OT-II or tamoxifen-treated iOT-II(Δ*Thpok*) were transferred to host OT-II(Δ*Runx3*) prior to an OVA-containing diet for 7 days. (**f**) Frequency of diarrhea-free mice after oral OVA challenge. (**g**) Quantification of fecal Lipocalin-2. (**h-j**) Tamoxifen-treated iOT-II(Δ*Thpok*) and control OT-II mice treated as in (a) were analyzed. Bar graphs represent frequency of CD8$\alpha^+$ (**h**) and Foxp3$^+$ (**i**) among TCRV$\alpha$2$^+$CD4$^+$ cells in the indicated tissues, analyzed by flow cytometry. (**j**) H&E staining of the small intestine (top) and large intestine (bottom). Original magnification, 40x. Graphs represent histological scores of the SI and the LI (each symbol represents one mouse). Data are expressed as median ± interquartile range. (**k,l**) Sorted naïve OVA-specific TCR transgenic cells (TCRV$\alpha$2$^+$CD4$^+$CD62L$^+$CD44$^{low}$) from wild-type OT-II CD45.1 congenic mice (**k**) or tamoxifen-treated iOT-II(Δ*Thpok*) mice (**l**) were transferred to host OT-II(Δ*Runx3*) mice prior to treatment as in (a). Bar graphs represent frequency of CD8$\alpha^+$ (left k,l) and Foxp3$^+$ (right k,l) among TCRV$\alpha$2$^+$CD4$^+$CD8$\beta^-$ derived from transferred (CD45.1) and host cells (**k**) or total cells (**l**) in indicated tissues.


a
OT-II
OT-II(ΔRunx3)
CD8α+ (% in CD4+)
spl
mLN
IE
LP
small
large
b
ThPOKhigh (% in CD4+)
c
Foxp3+ (% in CD4+)
ns
d
IL-17+ (%)
IFN-γ+ (%)
small IE
e
OT-II
OT-II(ΔRunx3)
SI
LI
Histological score (SI)
Histological score (LI)
f
Diarrhea free (%)
Time (days)
OT-II
OT-II(ΔRunx3)
OT-II(ΔRunx3) + OT-II(CD45.1)
OT-II(ΔRunx3) + OT-II(ΔThPOK)
g
Lipocalin-2 (ng/g)
h
Control
iOT-II(ΔThpok)
sIE
sLP
i
Foxp3+ (% in CD4+)
j
Control
iOT-II(ΔThpok)
Histological score
SI
LI
k
OT-II(CD45.1)
OT-II(ΔRunx3)
l
OT-II(ΔRunx3) + OT-II(ΔThPOK)
OT-II(ΔRunx3)

To specifically address the possibility that iTregs and CD4-IELs play complementary anti-inflammatory roles in the intestine, we compared T cell responses to dietary OVA using BALB/c background monoclonal OVA-specific TCR strains, carrying either wild-type Foxp3 or a *scurfy* mutation (Foxp3$^{sf}$)[133], which results in a Foxp3 loss of function [153]. In contrast to OT-II mice (C57BL/6 background), TBmc Foxp3$^{wt}$ mice fed an OVA diet showed a high rate of iTreg induction in all tissues examined, but less ThPOK loss and CD4-IELs in the epithelium (**Figure 2.13a-c**). Conversely, TBmc Foxp3$^{sf}$ showed a high degree of ThPOK loss and increased CD4-IEL development in the small intestine, in a frequency that mirrored the relative amounts of iTregs in TBmc Foxp3$^{wt}$ mice (**Figure 2.13a-e**). However, no inflammatory phenotype or diarrhea was observed even in the absence of functional Foxp3 in this monoclonal model (**Figure 2.13f-h**). To directly address whether exaggerated CD4-IEL differentiation could compensate for the absence of iTregs in TBmc Foxp3$^{sf}$ mice, we depleted CD4-IELs using anti-CD8$\alpha$ antibodies during OVA feeding (**Figure 2.13d, e**). We found that TBmc Foxp3$^{sf}$ mice treated with anti-CD8$\alpha$ antibodies, but not TBmc Foxp3$^{wt}$ treated with anti-CD8$\alpha$ antibodies or TBmc Foxp3$^{sf}$ treated with control antibodies, showed severe intestinal inflammation and diarrhea (**Figure 2.13f-i**). These results support a model in which CD4-IELs and iTregs cooperate in the regulation of local intestinal inflammation.

**Figure 2.13 Complementary roles for iTregs and CD4-IELs in regulating local inflammatory response towards dietary antigens.** (**a-c**) TBmc *Foxp3*$^{sf}$ (*Scurfy*) and TBmc control mice were analyzed. Bar graphs represent frequency of intracellular Foxp3$^{+}$ (**a**), CD8$\alpha^{+}$ (TL-tetramer) (**b**) and intracellular ThPOK$^{high}$ (**c**) cells among KJ1.26$^{+}$CD4$^{+}$CD8$\beta^{-}$ cells in the indicated tissues, analyzed by flow cytometry. (**d,c**) TBmc Foxp3$^{sf}$ (*Scurfy*) and TBmc control fed an OVA-containing diet for 7 days and injected with isotype control or anti-CD8$\alpha$ depleting antibody. Representative flow cytometry contour plots for surface CD8$\alpha$ (TL-tetramer) and intracellular Foxp3 (**d**) or ThPOK (**e**) expression among KJ1.26$^{+}$CD4$^{+}$CD8$\beta^{-}$ small intestine IELs. Bar graphs represent frequency of Foxp3$^{+}$ and CD8$\alpha\alpha^{+}$ (**d**) or ThPOK$^{high}$ (**e**) cells among KJ1.26$^{+}$CD4$^{+}$CD8$\beta^{-}$ small intestine IELs. (**f-h**) TBmc Foxp3$^{sf}$ (*scurfy*) and TBmc control fed an OVA-containing diet for 7 days and injected with isotype control or anti-CD8$\alpha$ depleting antibody. (**f**) Hematoxylin and eosin (H&E) staining of the SI (top) and the LI (bottom). Original magnification, 40x. Graphs represent histological scores of the SI (top) and the LI (bottom) (each symbol represents one mouse). (**g**) Frequency of diarrhea-free mice after oral OVA challenge. (**h**) Quantification of fecal Lipocalin-2. (**i**) H&E staining of the small intestine (SI) and large intestine (LI). Original magnification, 40x. Data expressed as mean ± SD, representative of at least two independent experiments (n = 3 - 8 per group). Scale bar, 200 $\mu$m. *$P$<0.05, **$P$<0.01, ***$P$<0.001; ns, not significant (Student's *t* test, or one-way ANOVA with Tukey post test).


a
b
c
d
e
TBmc
TBmc^sf
Foxp3^+ (%)
CD4^+CD8α^+ (%)
ThPOK^high (%)
Foxp3^+ (% in CD4^+)
CD8α^+ (% in CD4^+)
ThPOK^high (% in CD4^+)
spl
mLN
small LP
large IE
large LP
small IE
ns
TBmc+anti-CD8α
TBmc^sf+anti-CD8α
f
TBmc+anti-CD8α
TBmc^sf+anti-CD8α
SI
LI
Histological score (SI)
Histological score (LI)
g
Diarrhea free (%)
Time (days)
h
Lipocalin-2 (ng/g)
n.s
TBmc
TBmc^sf
TBmc +anti-CD8α
TBmc^sf +anti-CD8α
i
TBmc+isotype
TBmc^sf+isotype
SI
LI

## 2.7 Conclusion

The single-layered intestinal epithelium constitutes a uniquely challenging location for immune regulatory processes, given its proximity to highly stimulatory luminal contents and limited spatial organization. It is currently thought that Tregs utilize several redundant and complementary mechanisms to suppress inflammatory responses, and their capacity to sense specific environmental cues plays a major role in their function [154, 155, 156]. The physiological instability that we observe in the Treg lineage within the intestinal epithelium may represent an important modulation of regulatory activity that is coordinated by this particular environment [95, 123, 137, 154, 157]. The data presented here supports a cell-extrinsic suppressive role for CD4-IELs, although the likely epithelium-specific mechanism employed by these cells to actively regulate or prevent tissue inflammation remains unclear. However, a role for CD4-IELs in triggering inflammation via their cytotoxic activity, under specific contexts, is conceivable [66, 158]. Nevertheless, the observation that particular environmental cues, such as the microbiota, induce plasticity of seemingly stable lymphocyte lineages may contribute to the understanding of how specialized tissue-adaptation pathways function to balance efficient immune protective responses with tissue tolerance.

# Chapter 3: Stepwise chromatin and transcriptional acquisition of an intraepithelial lymphocyte program

## 3.1 Introduction

The hallmarks of CD4-IEL conversion include co-expression of CD4 and the CD8$\alpha\alpha$ homodimer which is induced by a switch from conventional CD4-to-CD8 lineage–defining transcription factors. CD4-IELs downregulate ThPOK (encoded by *Zbtb7b)*, expressed by all conventional mature CD4$^+$ T cells and upregulate the long form of Runx3, expressed by mature CD8$^+$ T cells[1, 66, 67]. While environmental cues and transcription factors involved in the differentiation of naïve CD4$^+$ T cells into gut–adapted T cell subsets have been described in the past decade, the transcriptional and chromatin changes that accompany such tissue–specific imprinting have not been elucidated.

We used genetic fate-mapping and gene ablation mouse models coupled with single-cell RNA-, ATAC- and ChIP-sequencing approaches to uncover the molecular mechanisms of how the intestinal tissue can imprint T-cell fate decisions on migrating mLN CD4$^+$ T cells. We found that gut-associated CD4$^+$T cell transcriptional profiles largely segregate by tissue location, indicating that upon leaving the gut-draining lymph nodes (LNs), migrating cells quickly adapt to either the LP or IE compartments; IE-adapted cells followed a stepwise acquisition of an IEL profile through a distinct pre-IEL stage. We specifically followed how the generally stable Treg phenotype is destabilized upon T cell migration to the IE compartment. We found that the Treg program is first downregulated at a pre-IEL stage before a cytotoxic IEL program is made accessible at the chromatin levels and then subsequently transcribed. Finally, we showed that while natural ThPOK downmodulation marks the pre-IEL stage, premature ThPOK loss in Tregs allows for the expression of the IEL profile before the Treg program is fully shut down. Our studies uncovered wide, tissue-specific and stepwise chromatin and transcriptional changes in T cells upon transitioning from tissue-draining LNs to tissue sites and revealed specific roles for lineage-defining transcription factors in driving this process.

## 3.2 Intestinal tissue sites imprint a unique program on migrating CD4$^+$ T cells

Both peripheral Tregs and Tconv CD4$^+$ T cells further differentiate into CD4-IELs upon migration to the intestinal epithelium[1, 66, 67, 159]. To characterize CD4$^+$ T cell heterogeneity in the gut along the draining lymph nodes -tissue axis, we combined single-cell RNA-sequencing (scRNA-seq) with fate-mapping approaches. First, to simultaneously study the transcriptional changes of Tregs and Tconv as they convert to CD4-IELs, we crossed the i*Foxp3*$^{Tomato}$ Treg-fate mapping strain[131] with *Zbtb7b*$^{GFP}$ reporter mice[21], the latter representing a strategy to follow IEL development using ThPOK loss as a marker[1]. We administered tamoxifen to label

Tregs and follow their fates (as Tomato$^+$) in three different sites along the LN-tissue axis: gut-draining mesenteric lymph nodes (mLN), lamina propria (LP) and epithelium (IE). We sorted tomato negative (Tom$^-$) or positive (Tom$^+$) CD4$^+$ T cells from the three locations and pooled the cells per tissue at approximately a 2:1 (Tom$^-$: Tom$^+$) ratio in order to achieve a better representation of current Tregs (Tom$^+$GFP$^+$), former Tregs (Tom$^+$GFP$^-$) and non-Tregs (Tom$^-$). We performed droplet-based scRNA-seq using the Chromium 10X (10X Genomics) platform on the three libraries, allowing us to explore the tissue-dependent relative heterogeneities of 6668 cells (**Figure 3.1a, b**). As expected from our previous studies[66, 67], CD4$^+$ T cell populations in the epithelium showed low levels of ThPOK primarily in the IE, which correlated with the acquisition of CD8$\alpha$ (**Figure 3.1a, b**).

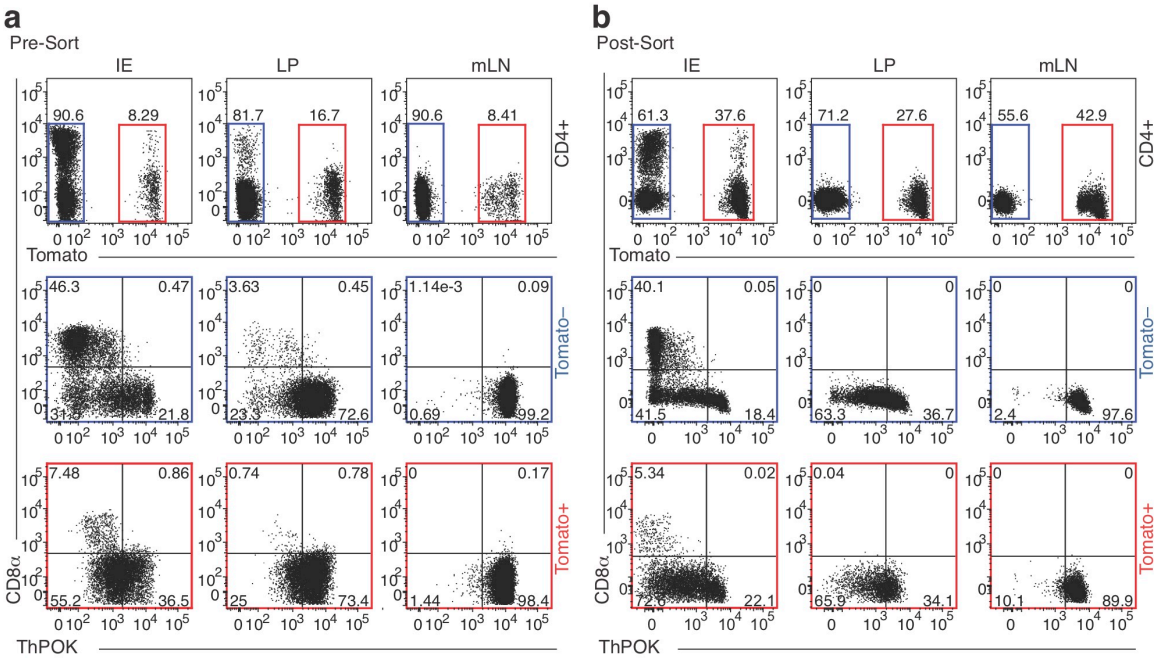


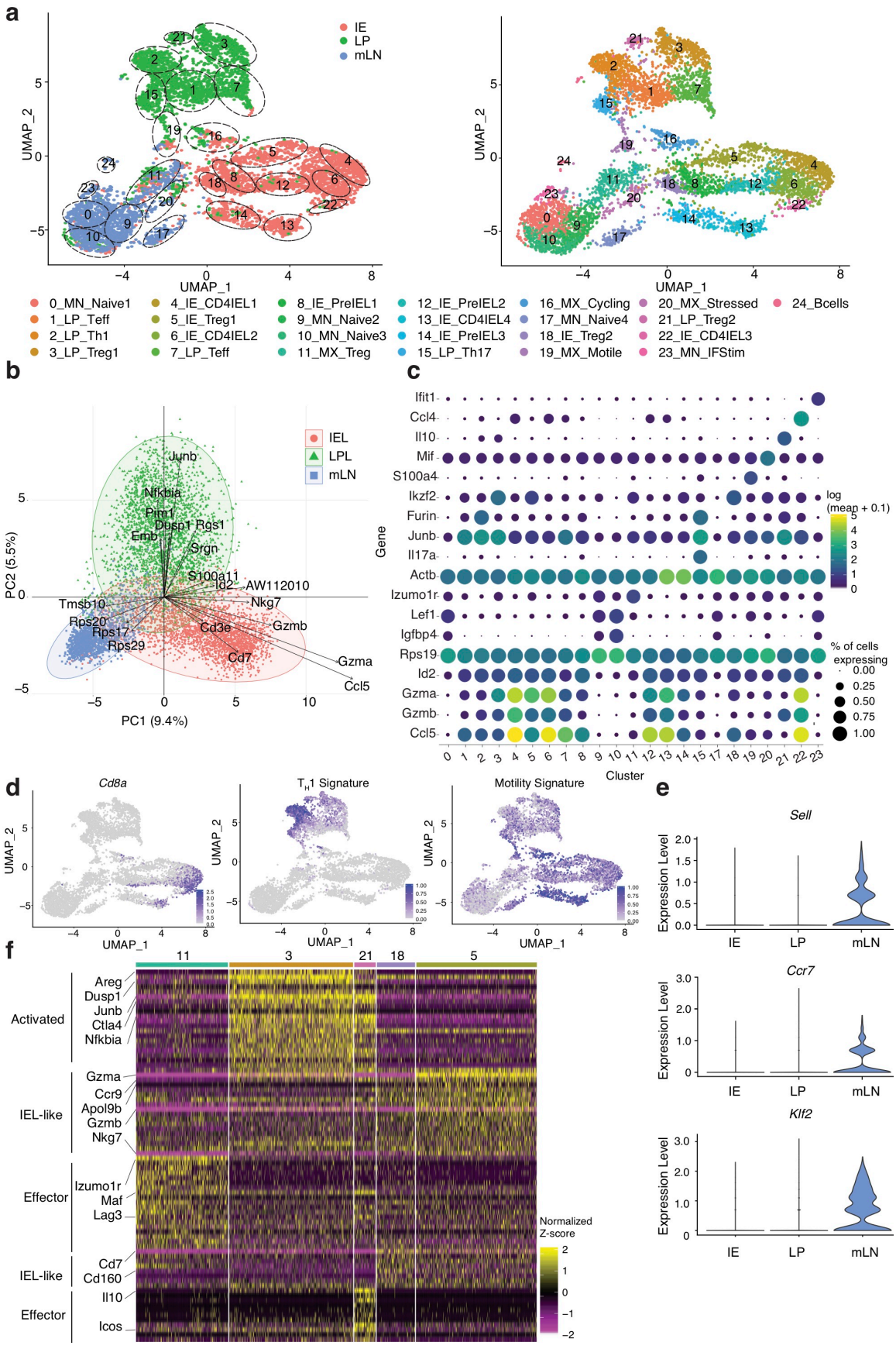

**Figure 3.1 Sorting strategy for single cell RNA-sequencing of CD4$^+$ T cells in three intestine-associated tissues.** (**a-b**) i*Foxp3*$^{Tom}$ThPOK$^{GFP}$ mice were treated with tamoxifen for 10 weeks, and Tomato$^-$ and Tomato$^+$ CD4$^+$ T cells from mesenteric lymph nodes (mLN), lamina propria (LP) and intestinal epithelium (IE) were sorted for scRNA-seq by the 10X Genomics platform. Sorted Tomato$^-$ (blue gates) and Tomato$^+$ (red gates) cells were pooled in a 2:1 ratio per tissue, resulting in 3 separate libraries. Protein levels of CD8$\alpha$, ThPOK and Tomato of CD4$^+$ T cell populations from mLN, LP and IE before sorting (**a**) and after sorting (**b**).

Cells were positioned based on gene expression similarities by uniform manifold approximation and projection (UMAP)[160] and assigned to 25 unbiased clusters (0-24, numbered by size order), including a small cluster of contaminating B cells (cluster 24). The vast majority of cells were segregated by tissue, suggesting that different microenvironments, including those within the gut, play a major role in defining the gene expression programs (**Figure 3.2a, b**). We defined the clusters

based on their top differentially–expressed genes (**Figure 3.2a-c**). Four clusters (11, 16, 19 and 20) contained cells from at least two different locations **(Figure 3.2a)**, and we classified them as mixed clusters (MX). Cluster 20 was defined by genes encoding mitochondrial localized proteins (*Atad3a, Cox7b*), snoRNAs regulators (*Nhp2, Gar1*), and genes related to ER stress such as heat shock proteins (*Hspd1, Hspa9*), and was thus defined as "stressed" cells. Cluster 16 was characterized by expression of genes related to cell cycle, including *Cdc's*, *Cdk's* and *Ccn's* and was inferred to contain cycling cells. Cells of cluster 19 expressed high levels of integrins (*Itga4, Itgb1, Itgb7*) as well as *Vim* and *Actg1*, suggesting these are motile cells (**Figure 3.2d**). The largest mixed tissue cluster, cluster 11, was partly defined by effector Treg genes (*Izumo1r, Pdcd1, Lag3, Icos, Tnfrsf4*) (**Figure 3.2c**). The three largest mLN clusters (MN_0,9,10) were defined by naïve markers (*Sell, Ccr7, Klf2*), as expected given their location (**Figure 3.2e**). One of the smallest mLN clusters, cluster 17, was defined by expression of *Actb*, whereas cluster 23 consists of multiple genes encoding interferon-induced proteins (Ifits) (**Figure 3.2c,)**. In general, cells isolated from mLN expressed more ribosomal protein-encoding genes in comparison to LP and IE cells, possibly due to their more naïve state and thus lower expression of cell state-defining markers (**Figure 3.2b**). As a whole, T cells from the LP expressed genes related to activation such as those in the Jun, Dusp, and NF-κb families (**Figure 3.2b, c**). Clusters 3 and 21 consisted of cells with relatively higher levels of *tdTomato, Ctla4, Icos,* and other effector Treg genes[61, 161] (**Figure 3.2f, Figure 3.3a,**). Cells in LP clusters 2 and 15 expressed $T_H1$ (*Il12rb2, Il18r, Stat4*) and $T_H17$ (Il*17a, Il22, Rora*)-associated genes, respectively (**Figure 3.2a, c, d**). Clusters 1 and 7 contained cells from LP that did not express the typical $T_H1$ or $T_H17$ signatures but had the effector phenotype. Cluster 7 was polarized towards an IEL-profile, with higher expression of *Nkg7*, *Gzmk*, and *Ccl5* (**Figure 3.2a, c, Figure 3.3b, c**). We considered all clusters in the IE as *tdTomato*-expressing Tregs (clusters 5,18), *Cd8a*-expressing CD4-IELs (clusters 4,6,13,22) or cells that shared multiple genes with CD4-IELs but lacked *Cd8a expression* (clusters 8, 12, 14) (**Figure 3.2a, f, Figure 3.3a-f**).

To gain insight into how the gut environment influences Treg intra-tissue adaptation, we further compared all Treg clusters (11, 3, 21, 18, 5) within the different gut-associated sites (**Figure 3.2f, Figure 3.3b**). The two Treg clusters (3,21) present in the LP displayed a more activated profile *(Ctla4, Areg, Dusp1, Junb, Nfkbia)*. Cluster 21 consisted of Tregs expressing *Icos* and *Il10*, while those of cluster 3 expressed higher levels of *Gzma*. Both IE Treg clusters (5, 18) displayed an activated/effector-like phenotype (*Cd3e*, *Cd3g*, *Tnfrsf9*, *Tnfrsf18*), and were also more CD8-like (*Cd160, Cd7*). Additionally, cells in cluster 5 expressed elevated levels of granzymes and *Nkg7,* suggestive of cytotoxic potential. Cells in cluster 11, the only cluster to include mLN Tregs, had lower expression of effector Treg markers and of the gut homing receptor *Ccr9* compared to enteric tissue Tregs (**Figure 3.2f, Figure 3.3b**). These results point to a previously unappreciated level of Treg heterogeneity between closely related, yet

**Figure 3.2 Transcriptional profiles of intestinal CD4$^+$ T cells segregate by tissue. (a-f**) scRNA-seq analysis of Tomato$^-$ and Tomato$^+$ CD4$^+$ T cells from mesenteric lymph nodes (mLN), lamina propria (LP) and intestinal epithelium (IE) from i*Foxp3*$^{Tom}$ThPOK$^{GFP}$ mice after tamoxifen treatment. (**a**) Uniform manifold approximation and projection (UMAP) representation of sequenced cells, color coded by tissue of origin with outlined clusters (left) or colored by cluster (right). Cluster names and numbers correspond to colors throughout the figure as indicated. (**b**) Principle component (PC) analysis of cells per tissue, exhibiting top variable features. (**c**) Top expressed genes per cluster. Size represents proportion of cells per cluster expressing the indicated gene and color represents expression level. (**d**) Expression levels of *Cd8a*, the $T_H1$ signature, and the motility signature of all sequenced cells. (**e**) Expression levels of *Sell* (top)*, Ccr7* (middle)*,* and *Klf2* (bottom) of cells from indicated tissues. (**f**) Heatmap of top differentially expressed genes in Treg clusters represented by the normalized Z-score.

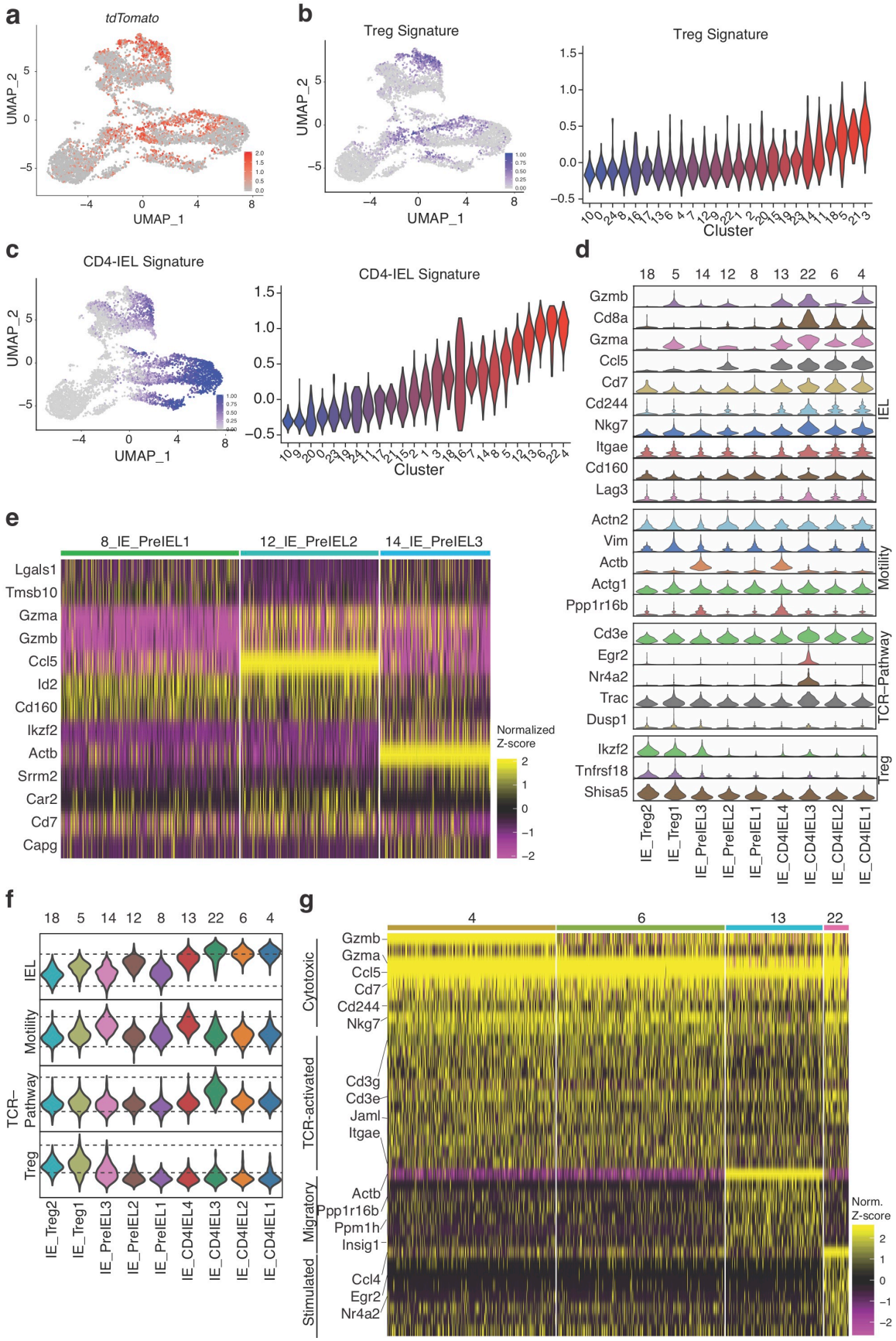

a
IE
LP
mLN
UMAP_1
UMAP_2
0_MN_Naive1
1_LP_Teff
2_LP_Th1
3_LP_Treg1
4_IE_CD4IEL1
5_IE_Treg1
6_IE_CD4IEL2
7_LP_Teff
8_IE_PreIEL1
9_MN_Naive2
10_MN_Naive3
11_MX_Treg
12_IE_PreIEL2
13_IE_CD4IEL4
14_IE_PreIEL3
15_LP_Th17
16_MX_Cycling
17_MN_Naive4
18_IE_Treg2
19_MX_Motile
20_MX_Stressed
21_LP_Treg2
22_IE_CD4IEL3
23_MN_IFStim
24_Bcells
b
IEL
LPL
mLN
PC1 (9.4%)
PC2 (5.5%)
Junb
Nfkbia
Pim1
Dusp1
Rgs1
Emb
Srgn
S100a11
Id2
AW112010
Nkg7
Tmsb10
Rps20
Rps17
Rps29
Cd3e
Gzmb
Cd7
Gzma
Ccl5
c
Gene
Cluster
log (mean + 0.1)
% of cells expressing
Ifit1
Ccl4
Il10
Mif
S100a4
Ikzf2
Furin
Junb
Il17a
Actb
Izumo1r
Lef1
Igfbp4
Rps19
Id2
Gzma
Gzmb
Ccl5
d
Cd8a
TH1 Signature
Motility Signature
e
Sell
Ccr7
Klf2
Expression Level
IE
LP
mLN
f
Normalized Z-score
Activated
Areg
Dusp1
Junb
Ctla4
Nfkbia
IEL-like
Gzma
Ccr9
Apol9b
Gzmb
Nkg7
Effector
Izumo1r
Maf
Lag3
IEL-like
Cd7
Cd160
Effector
Il10
Icos

distinct tissues, with a skew towards the IEL/cytotoxic program within the epithelium.

Next, we further characterized the heterogeneity of epithelial CD4$^+$ T cells. We defined gene-expression signatures for Tregs, TCR-stimulated cells, motile cells, and CD4-IELs, based on the expression profiles of the different clusters (**Figure 3.2f, Figure 3.3b-g**). Although both Treg clusters (5 and 18) expressed the highest levels of Treg-signature genes, cluster 14 also expressed the same genes at a lower level. CD4-IEL cluster 22 displayed the highest level of genes in the TCR pathway (**Figure 3.3d, f**). Clusters 13 and 14, positioned next to each other, expressed the highest levels of motility-related genes, suggesting increased migratory capacity. Finally, while three out of the four CD4-IEL clusters (4, 6 and 22) displayed equally high CD4-IEL signatures, cluster 13 was slightly less polarized towards the CD4-IEL phenotype (**Figure 3.2c, d, f**). The three non-CD4-IEL and non-Treg clusters in the IE (8, 12, 14) nevertheless displayed a high level of the CD4-IEL signature (**Figure 3.3c**).

These clusters were characterized by the presence of cells expressing equal or slightly lower levels of IEL markers such as *Itgae*, *Cd160*, *Nkg7* compared to *bona fide* IELs (**Figure 3.3e, g**), suggesting that these constitute a pre-IEL population. Among these pre-IEL clusters, which were very similar to each other, cluster 12 was the most similar to CD4-IELs expressing higher levels of *Ccl5* and *Gzmb* (**Figure 3.3c-f**). Likewise, the CD4-IEL clusters were separated by only minor differences, including upregulation of genes in the TCR pathway in cluster 22 (*Egr2* and *Nr4a2*) and of a motility signature in cluster 13 (*Actb* and *Ppp1r16b*). All CD4-IELs expressed *Cd8a, Cd244, Gzma, Cd7, Nkg7, Jaml* as well as *Ccl5* **(Figure 3.3g**). Of note, tomato positive and negative CD4-IELs were dispersed among all CD4-IEL clusters, indicating that CD4-IELs derived from Tregs or from Tconvs did not display major differences capable of overriding their CD4-IEL signature similarities (**Figure 3.3a**). Together, our data indicate that CD4$^+$ T cells in the IE are placed on a gradient leading to the full expression of a CD4-IEL signature.

**Figure 3.3 Intestinal epithelium imprints a cytotoxic program on migrating CD4$^{+}$ T cells. (a-g**) Single cell RNA-seq analysis of Tomato$^{-}$ and Tomato$^{+}$ CD4$^{+}$ T cells from mesenteric lymph nodes (mLN), lamina propria (LP) and intestinal epithelium (IE) from i*Foxp3*$^{Tom}$ThPOK$^{GFP}$ mice after tamoxifen treatment. (**a**) *tdTomato* gene expression in 6668 sequenced cells. (**b-c**) Expression levels of Treg (**b**) and CD4-IEL(**c**) signatures in sequenced cells (left) and per cluster of cells (right). (**d**) Expression levels of indicated genes among IE clusters. (**e)** Expression heatmap of selected genes in pre-IEL clusters. (**f**) Expression levels of indicated signatures among IE clusters as indicated. (**g**) Heatmap of top differentially expressed genes in CD4-IEL clusters by the normalized Z-score.


a
tdTomato
UMAP_2
UMAP_1
b
Treg Signature
Treg Signature
Cluster
c
CD4-IEL Signature
CD4-IEL Signature
Cluster
d
18 5 14 12 8 13 22 6 4
Gzmb
Cd8a
Gzma
Ccl5
Cd7
Cd244
Nkg7
Itgae
Cd160
Lag3
IEL
Actn2
Vim
Actb
Actg1
Ppp1r16b
Motility
Cd3e
Egr2
Nr4a2
Trac
Dusp1
TCR–Pathway
Ikzf2
Tnfrsf18
Shisa5
Treg
IE_Treg2
IE_Treg1
IE_PreIEL3
IE_PreIEL2
IE_PreIEL1
IE_CD4IEL4
IE_CD4IEL3
IE_CD4IEL2
IE_CD4IEL1
e
8_IE_PreIEL1
12_IE_PreIEL2
14_IE_PreIEL3
Lgals1
Tmsb10
Gzma
Gzmb
Ccl5
Id2
Cd160
Ikzf2
Actb
Srrm2
Car2
Cd7
Capg
Normalized Z-score
f
18 5 14 12 8 13 22 6 4
IEL
Motility
TCR–Pathway
Treg
IE_Treg2
IE_Treg1
IE_PreIEL3
IE_PreIEL2
IE_PreIEL1
IE_CD4IEL4
IE_CD4IEL3
IE_CD4IEL2
IE_CD4IEL1
g
4
6
13
22
Gzmb
Gzma
Ccl5
Cd7
Cd244
Nkg7
Cytotoxic
Cd3g
Cd3e
Jaml
Itgae
TCR-activated
Actb
Ppp1r16b
Ppm1h
Insig1
Migratory
Ccl4
Egr2
Nr4a2
Stimulated
Norm. Z-score

To understand how the gradient of transcriptional changes culminates in tissue adaptation of $CD4^+$ T cells within the lymphoid to non-lymphoid tissue axis, we performed pseudotime analysis that produce cell trajectory inferences using two independent but complementary methods, Monocle3 and Slingshot (**Figure 3.4a**). While Monocle3 constructs a minimum spanning tree utilizing the individual cells and projects it onto the UMAP embedding, Slingshot uses the cluster centers making it less sensitive to outliers. Color-coding cells along differentiation paths showed that cells making up the clusters along the edges of the LP and IE compartments in the UMAP, including all CD4-IEL clusters, are the most "terminally-differentiated" compared to mLN cells. (**Figure 3.4a**). The trajectory branches leading to LP versus IE clusters bifurcated within the mixed-tissue cluster 11 and no further connections were detected between LP and IE (**Figure 3.4a**), suggesting that once cells acquire a LP profile they likely remain there. In agreement with the UMAP, pre-IELs were positioned as intermediate stages in the trajectories from naïve $CD4^+$ T cells of the mLN to CD4-IELs (**Figure 3.4a**). Based on the gradient towards the IEL profile within the epithelium as well as trajectory analysis, we chose three combinations of clusters along the mLN-IE axis, and ordered their cells based on principal component analysis (PC1). The diagonal distribution of cells in all combinations corroborated the differentiation pattern from least differentiated mLN cells to terminally differentiated CD4-IELs (**Figure 3.4b**). This progression consisted of the gradual upregulation of genes such as *Ccl5*, granzymes, *Nkg7 and Cd8a*, until completion of their terminal IEL differentiation (**Figure 3.4b, c**). These genes were also the main drivers of the IE-signature described in figure 1, further suggesting that the IEL profile defines epithelial imprinting. Our scRNA-seq revealed that while the lamina propria allows for the maintenance of distinct $CD4^+$ T cell subsets, this heterogeneity is not seen in the epithelium, suggesting that localization to the epithelium is responsible for acquisition of the IEL profile.

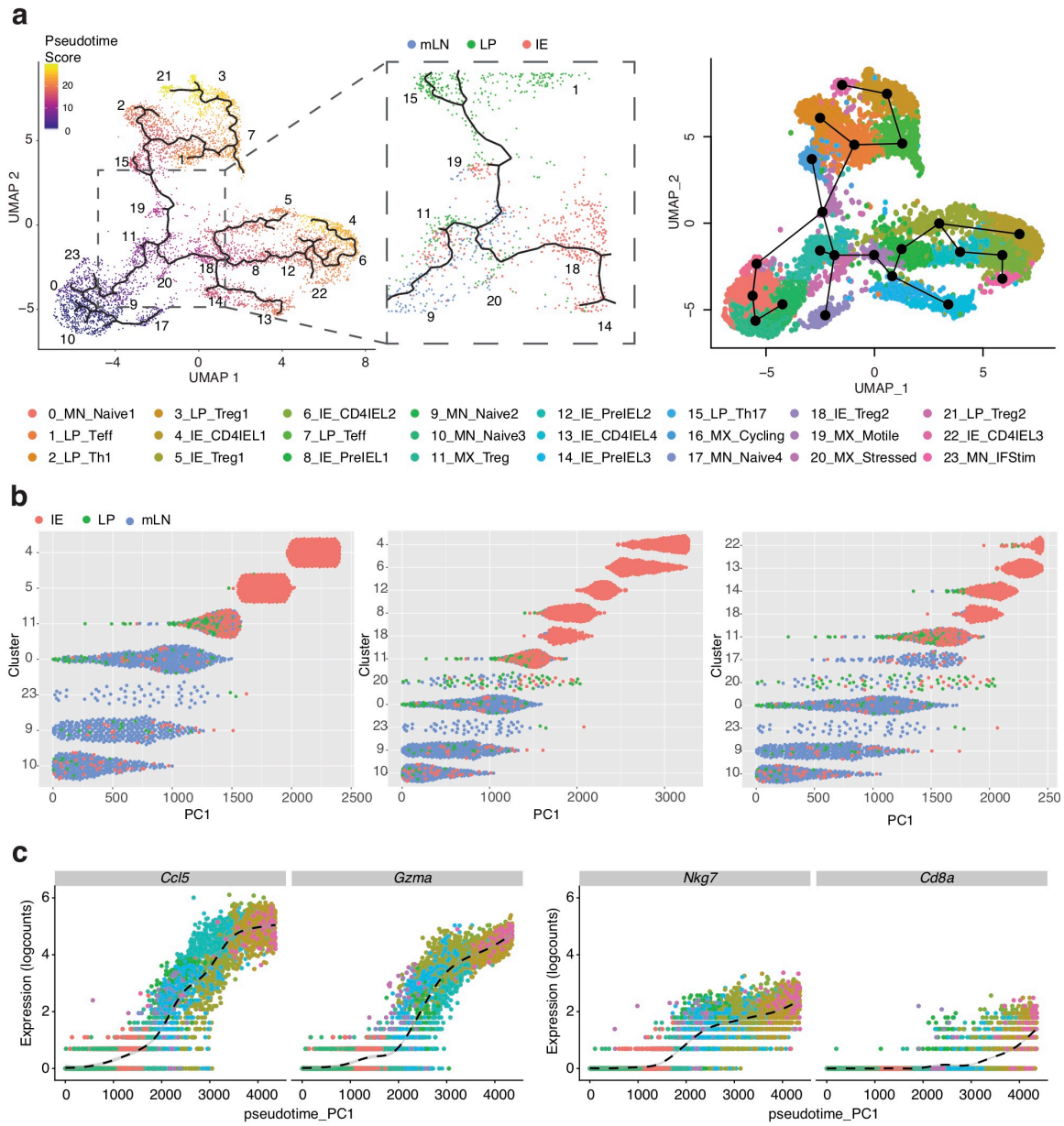



**Figure 3.4 Tissue segregation precedes acquisition of an IEL profile. (a-c)** Single cell RNA-seq analysis of Tomato$^{-}$ and Tomato$^{+}$ CD4$^{+}$ T cells from mesenteric lymph nodes (mLN), lamina propria (LP) and intestinal epithelium (IE) from iFoxp3$^{Tom}$ThPOK$^{GFP}$ mice after tamoxifen treatment. (**a**) Pseudotime analysis of sequenced cells using Monocle3 with cells color-coded by pseudotime gradient (left) and by Slingshot with cells color-coded by cluster (right). Numbers indicate the UMAP clusters. Middle panel displays a point of bifurcation in the Monocle3 trajectory analysis, cells colored by tissue. Numbers indicate the UMAP clusters. (**b**) Cells ordered according to principle component 1 (PC1). (**c**) Scatter plot of gene expression ranked based on PC1.

## 3.3 Coordinated transcriptional and chromatin changes during T cell adaptation to the epithelium

We next sought to couple the stepwise changes in transcription with those of chromatin accessibility as Tregs destabilized to become CD4-IELs. As noted in previous studies[66, 67], including in the chapter above, ThPOK downmodulation precedes CD8α-acquisition. Our scRNA-seq analysis using the ThPOK$^{GFP}$ reporter mouse revealed intermediate stages within the IE, and thus we refer to ThPOK$^{low}$CD8α$^{-}$ cells in the following studies as pre-IELs. We performed an assay for transposase-accessible chromatin followed by sequencing (ATAC-seq) and bulk RNA sequencing on induced Tregs (iTreg; neuropilin1$^{low}$Tom$^{+}$GFP$^{high}$CD8α$^{-}$), pre-IELs coming from Tregs (Tom$^{+}$GFP$^{low}$ CD8α$^{-}$), Treg-derived CD4-IELs (exTreg-IEL; Tom$^{+}$GFP$^{low}$ CD8α$^{+}$) and Tconv-derived CD4-IELs (Tom$^{-}$ GFP$^{low}$CD8α$^{+}$) sorted from the IE of i*Foxp3*$^{Tomato}$ThPOK$^{GFP}$ mice (**Figure 3.5a, b**). Although the Tom$^{-}$ CD4-IELs in our system are not derived from Tom$^{+}$ Tregs, we use them as a phenotypic endpoint to define CD4-IELs, and to compare any differences associated with Treg- versus Tconv-derived IELs.

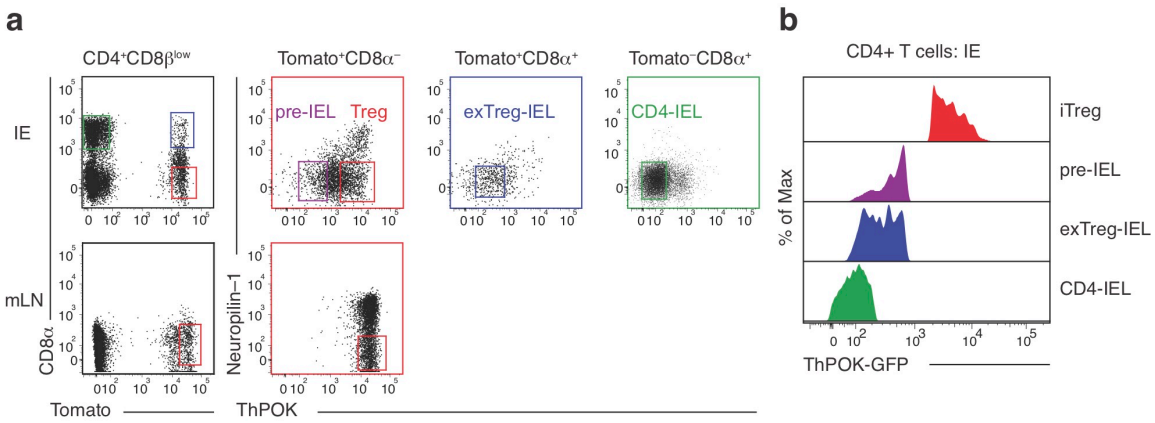
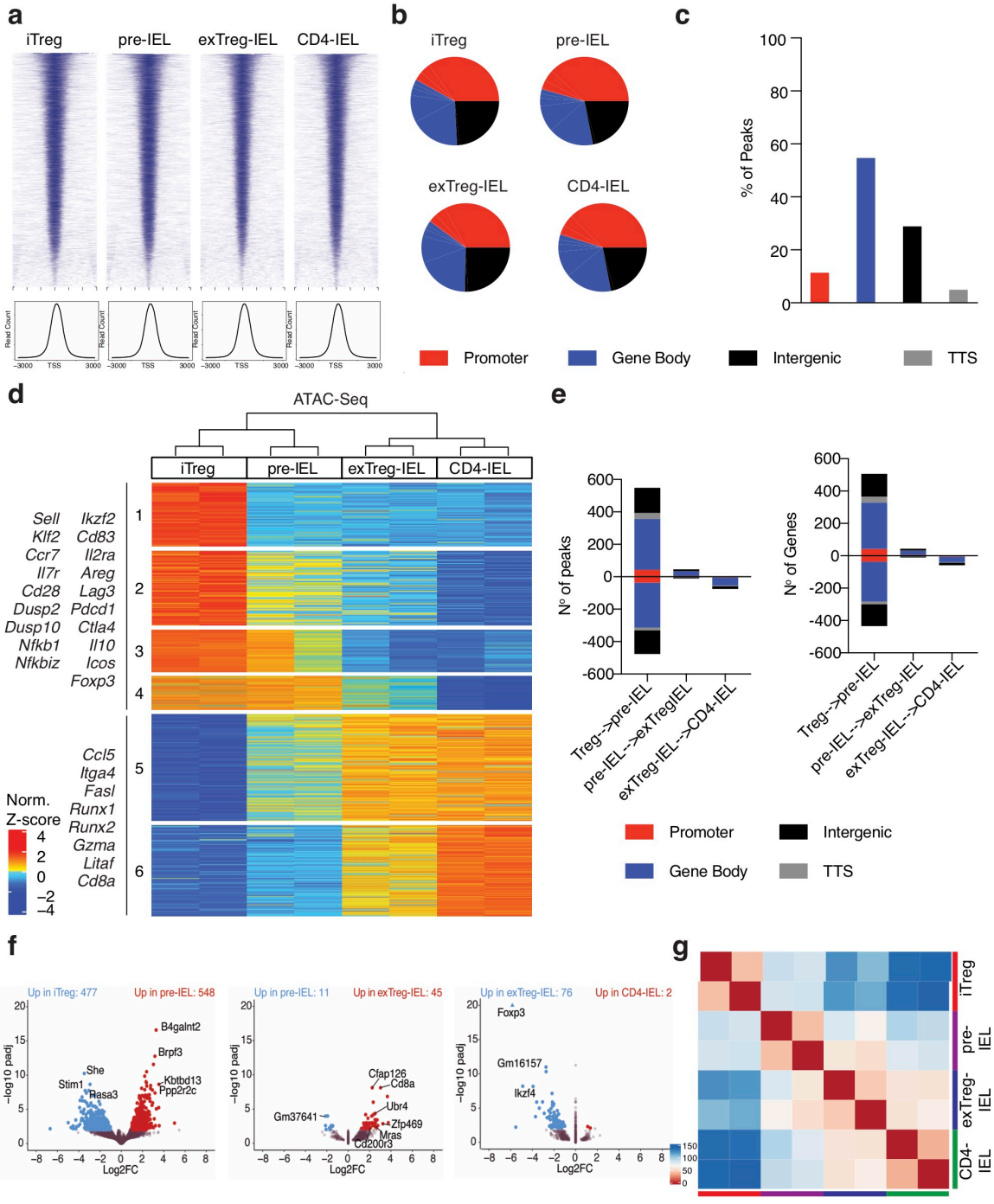



**Figure 3.5 Sorting strategy for ATAC- and RNA- sequencing.**
(**a-b**) i*Foxp3*$^{Tom}$ThPOK$^{GFP}$ mice were treated with tamoxifen for 10 weeks and induced Tregs (iTreg; CD4$^{+}$Tomato$^{+}$GFP$^{High}$neuropilin-1$^{-}$CD8α$^{-}$), pre-IELs (CD4$^{+}$Tomato$^{+}$ GFP$^{Low}$CD8α$^{-}$), exTreg-IELs (CD4$^{+}$Tomato$^{+}$GFP$^{Low}$CD8α$^{+}$), and CD4-IELs (CD4$^{+}$ Tomato$^{-}$GFP$^{Low}$CD8α$^{+}$) were sorted in bulk from the IE. Assay for transposase-accessible chromatin (ATAC) or RNA libraries were prepared followed by sequencing of indicated populations. iTregs were also sorted from the mLN for RNA-seq. (**a**) Sorting strategy of indicated populations. CD8α and Tomato expression among CD4$^{+}$ CD8β$^{low}$ T cells (left) from IE (top) and mLN (bottom). ThPOK and neuropilin-1 expression among Tomato$^{+}$CD8α$^{-}$ (red), Tomato$^{+}$CD8α$^{+}$ (purple) and Tomato$^{-}$CD8α$^{+}$ (green) cells from IE (top) and mLN (bottom). (**b**) Levels of ThPOK expression in each sorted population in the IE.

In agreement with previous ATAC-seq data[162], the majority of accessible chromatin peaks occurred close to transcriptional start sites (TSS) (**Figure 3.6a**). All populations revealed a similar profile of chromatin accessibility relative to genome regions, with the highest frequency of accessibility at promoters (**Figure 3.6b**). However, out of the genes with differentially–accessible chromatin regions (DACR) between the populations surveyed, the majority falls within gene bodies, and not promoter regions, suggesting non-random chromatin changes (**Figure 3.6c**). The likelihood ratio test (LRT) showed that DACRs were divided into 6 clusters; regions in clusters 1-4 gradually decreased in accessibility, while those in clusters 5-6 become more accessible as Tregs differentiate into CD4-IELs (**Figure 3.6d**). Both activation (*Cd28*, Dusp and Nf-κb families) and naïve-related genes (*Sell, Klf2, Ccr7, Il7r*) contained chromatin regions which decrease in accessibility as Tregs destabilized, suggesting a shift in activation state during cell-type progression. DACRs in this progression also included *Ifng* and *Stat4,* suggesting a role for interferon regulation in the Treg to IEL progression, as previously suggested[127]. As expected, classic Treg (*Foxp3*, *Il10, Cd83, Il2ra, Areg)* and effector (*Lag3, Pdcd1, Ctla4, Il10, Icos*) genes decreased in accessibility as Tregs developed into IELs. Of note, while most Treg markers appeared in multiple clusters (1-4), *Foxp3* was detected only in the 4$^{th}$ cluster of regions that decreased in accessibility as pre-IELs transitioned to the IEL phenotype. This suggests that regulation of other Treg-signature genes occurs before *Foxp3*, in agreement with previous reports showing that *Foxp3* expression does not always correlate with Treg profile and function[61]. Chromatin regions in clusters 5 and 6, which began to become more accessible at the pre-IEL stage, and peaked in accessibility in IELs, included typical CD8-IEL (*Runx1/2/3, Gzma, Litaf, Fasl)* and homing (*Ccl5, Itga4)* genes. Cluster 5 contained regions that remained equally accessible in IELs derived from Tregs (exTreg-IEL) and Tconv (CD4-IEL), providing further evidence suggesting that the IEL signature overrides previous lineage or stage characteristics. Chromatin regions in cluster 6, which become more accessible at the exTreg-IEL stage, were more open in CD4-IELs. In addition to regions in typical IEL genes, cluster 6 included *Cd8a*, suggesting that it is not until the IEL program is made accessible that *Cd8a* is upregulated (**Figure 3.6d**). Moreover, Wald pairwise comparisons (**Figure 3.6e, f**) as well as Euclidean distance analysis (**Figure 3.6g**) showed that the largest changes in chromatin accessibility occurred as iTregs destabilized towards the ThPOK-low pre-IEL stage. The majority of the DACRs in this step fell on gene bodies (**Figure 3.6e**). A relatively minor number of regions, including that of *Cd8a*, became more accessible as pre-IELs transitioned into CD8α-expressing exTreg-IELs, which were the most similar to CD4-IELs; however, the former did retain a number of accessible regions at Treg genes, including *Foxp3*, when compared to CD4-IELs, suggesting that the chromatin landscape of Treg genes is not completely lost upon transition to the IEL phenotype (**Figure 3.6f**). Overall, it is not until the Treg program is set to become less accessible that the chromatin of the IEL program becomes more accessible, suggesting that ThPOK loss at the

**Figure 3.6 The chromatin landscape changes in a stepwise manner as Tregs convert to CD4-IELs through a pre-IEL stage.** (**a-g**) ATAC-seq analysis of iTregs, pre-IELs, exTreg-IELs and CD4-IELs of the intestinal epithelium of i*Foxp3*$^{Tom}$ThPOK$^{GFP}$ mice post tamoxifen treatment. (**a**) Mapped accessible chromatin regions relative to transcriptional start sites (TSS) of all genes in ATAC-seq data (top) and their relative read counts per genomic regions (bottom). (**b**) ATAC-seq peak annotations as indicated per cell type. (**c**) Percent of total differentially accessible chromatin regions as follows: 5'UTR and promoters (Promoter; red), 3'UTR with exons and introns (Gene Body; blue), transcriptional termination site (TTS; grey) and intergenic (black). (**d**) Heatmap of likelihood ratio test (LRT) of all differentially accessible chromatin regions (DACR) of indicated bulk populations. (**e**) Number of DACR (left) and the genes on which they are positioned on (right) between cell types in sequential progression as performed by Wald pairwise test. Colored by chromosome region as in (c). (**f**) Volcano plots representing DACRs between cell types in sequential progression as performed by Wald pairwise test. (**g**) Euclidean distance correlation of chromatin accessibility profiles of all samples. Significant DACR $p<0.01$

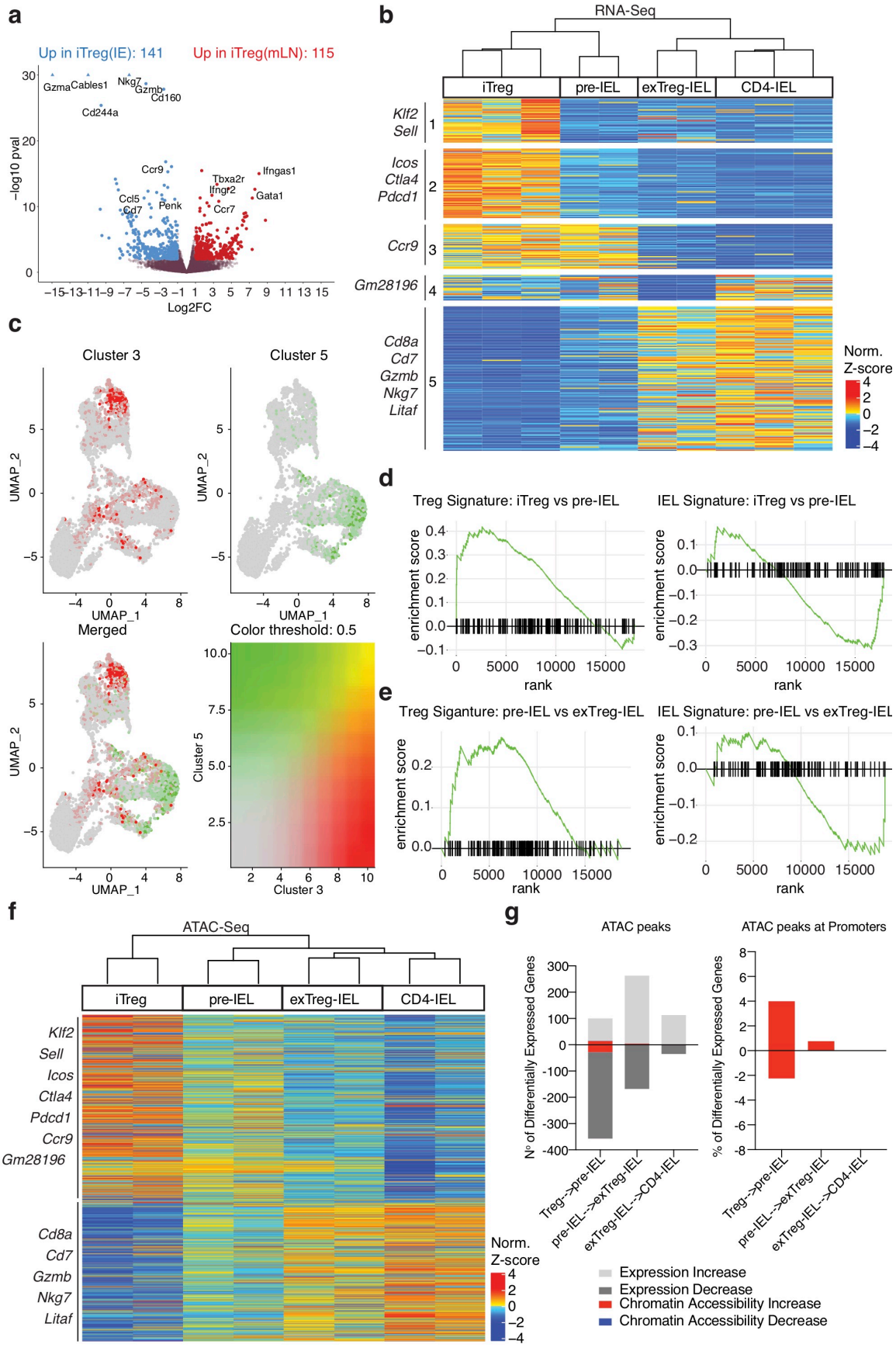

a
iTreg
pre-IEL
exTreg-IEL
CD4-IEL
Read Count
−3000
TSS
3000
b
iTreg
pre-IEL
exTreg-IEL
CD4-IEL
c
% of Peaks
100
80
60
40
20
0
Promoter
Gene Body
Intergenic
TTS
d
ATAC-Seq
iTreg
pre-IEL
exTreg-IEL
CD4-IEL
Sell
Klf2
Ccr7
Il7r
Cd28
Dusp2
Dusp10
Nfkb1
Nfkbiz
Ikzf2
Cd83
Il2ra
Areg
Lag3
Pdcd1
Ctla4
Il10
Icos
Foxp3
Ccl5
Itga4
Fasl
Runx1
Runx2
Gzma
Litaf
Cd8a
1
2
3
4
5
6
Norm.
Z-score
4
2
0
−2
−4
e
Nº of peaks
600
400
200
0
-200
-400
-600
Treg-->pre-IEL
pre-IEL-->exTregIEL
exTreg-IEL-->CD4-IEL
Nº of Genes
Treg-->pre-IEL
pre-IEL-->exTreg-IEL
exTreg-IEL-->CD4-IEL
Promoter
Intergenic
Gene Body
TTS
f
Up in iTreg: 477
Up in pre-IEL: 548
B4galnt2
Brpf3
She
Stim1
Kbtbd13
Ppp2r2c
Rasa3
−log10 padj
Log2FC
Up in pre-IEL: 11
Up in exTreg-IEL: 45
Cfap126
Cd8a
Ubr4
Gm37641
Zfp469
Mras
Cd200r3
Up in exTreg-IEL: 76
Up in CD4-IEL: 2
Foxp3
Gm16157
Ikzf4
g
iTreg
pre-IEL
exTreg-IEL
CD4-IEL

pre-IEL stage, or the events that lead to it, initiate the Treg plasticity observed at the intestinal epithelium.

To define how changes in chromatin accessibility relate to differentially expressed genes (DEG) in IEL populations, we performed bulk RNA-seq on the same IE subsets. We also sequenced mLN iTregs in order to compare their transcriptional profile to that of iTregs in the IE (**Figure 3.7a**). Comparing mLN to IE iTregs confirmed that transcriptional differences segregated by tissue, the latter expressing more activation, intestinal and CD8–associated genes (*Runx3, Cd160, Cbfb, Itgae)*, whereas the former expressed genes associated with resting LN Tregs (*Ccr7, Sell, Klf2*) (**Figure 3.7a**). DEGs between all sequenced IE populations revealed 5 unbiased clusters, highlighting the changes in transcriptional profile as Tregs convert to IELs (**Figure 3.7b**). At the pre-IEL stage when ThPOK is downmodulated, expression of effector Treg genes (Cluster2; *Icos, Ctla4, and Pdcd1)* was decreased. The clusters of genes down-regulated at the exTreg-IEL stage (Clusters 3 and 4) included multiple pseudogenes and lincRNAs, which could be implicated in cell maintenance and identity integrity, serving as a hallmark of cell differentiation at the exTreg-IEL stage. Finally, cluster 5, which is upregulated by both Treg- and Tconv- derived IELs (exTreg-IELs and CD4-IELs), includes many CD8- and IEL-associated genes (*Cd8a, Cd7, Gzmb, Nkg7, Litaf*). Although these genes begin to be slightly upregulated in the pre-IEL stage, they reach their full levels only when T cells become IELs, as marked by CD8$\alpha$ protein and gene expression (**Figure 3.5a, Figure 3.7b**). Whereas the signature of cluster 3 maps to the Treg clusters in all tissues in the scRNA-seq data set, that of cluster 5 is primarily expressed by IE cells, particularly CD4-IELs (**Figure 3.7c**). Although there is very low expression of this signature by a few LP cells, this CD4-IEL program is gained in the IE, after cells have left the LP, again suggesting that they are not polarized towards this program prior to entering the IE compartment. Gene set enrichment analysis (GSEA) using Treg and IEL signatures from our scRNA-seq (Clusters 21 and 6, respectively) revealed that expression of the Treg signature decreases as the IEL signature increases and cells progress from the Treg to pre-IEL stage (**Figure 3.7d**). In contrast, the pre-IEL to exTreg-IEL progression displayed only minimal changes in the Treg signature (**Figure 3.7e**), since Treg-derived pre-IELs no longer expressed most of the *bona-fide* Treg genes. Overall, the gene expression changes during IEL differentiation suggest that the IEL program begins during the pre-IEL stage, before the full Treg program is shut down but after the downregulation of ThPOK.

To further understand how changes in gene transcription correlate with changes in chromatin accessibility, we plotted the levels of chromatin accessibility of genes that changed in transcription in the same order of the heatmap representing RNA-seq (**Figure 3.7b**). We found that in general, the transcriptional profile correlated with chromatin accessibility of those genes in iTregs and CD4-IELs (**Figure 3.7f**). Genes that were transcriptionally downregulated after the iTreg stage remained

**Figure 3.7 Treg program shutdown precedes IEL program acquisition through a pre-IEL stage.** (**a-g**) ATAC- and RNA-seq analysis of iTregs, pre-IELs, exTreg-IELs and CD4-IELs of intestinal epithelium of i*Foxp3*$^{Tom}$ThPOK$^{GFP}$ mice post tamoxifen treatment. (**a**) Volcano representation of differentially expressed genes between IE iTregs (higher expression in blue) and mLN iTregs (higher expression in red), performed by Wald pairwise comparison test, $p_{adj}<0.05$ values were considered significant. (**b**) Heatmap of LRT of differentially expressed genes (DEG) between indicated populations. (**c**) Treg signature from cluster 3 and IEL signature from cluster 5 of the bulk RNA-seq heatmap (b) overlaid onto the scRNA-seq UMAP from figure 3.2a. (**d-e**) Gene set enrichment analysis of Treg and CD4-IEL signatures from scRNA-seq (top panels, cluster 21 and bottom panels, cluster 6, respectively as shown in *figure 3.2*) in iTreg to pre-IEL (**d**) and pre-IEL to exTreg-IEL (**e**) progressions. (**f**) Levels of chromatin accessibility among DEG, in the same order as in (b). (**g**) Numbers of DEGs (left) (light grey for increase, dark grey for decrease) and percent of DACR at promoters among DEG (right) in between cell types in sequential progression as performed by Wald pairwise test as indicated. DACR within those genes (red for increase, blue for decrease in accessibility).


a
Up in iTreg(IE): 141
Up in iTreg(mLN): 115
−log10 pval
Log2FC
Gzma
Cables1
Nkg7
Gzmb
Cd160
Cd244a
Ccr9
Tbxa2r
Ifngr2
Ifngas1
Gata1
Ccl5
Cd7
Penk
Ccr7
b
RNA-Seq
iTreg
pre-IEL
exTreg-IEL
CD4-IEL
Klf2
Sell
Icos
Ctla4
Pdcd1
Ccr9
Gm28196
Cd8a
Cd7
Gzmb
Nkg7
Litaf
Norm. Z-score
c
Cluster 3
Cluster 5
UMAP_1
UMAP_2
Merged
Color threshold: 0.5
d
Treg Signature: iTreg vs pre-IEL
IEL Signature: iTreg vs pre-IEL
enrichment score
rank
e
Treg Siganture: pre-IEL vs exTreg-IEL
IEL Signature: pre-IEL vs exTreg-IEL
f
ATAC-Seq
g
ATAC peaks
ATAC peaks at Promoters
N° of Differentially Expressed Genes
% of Differentially Expressed Genes
Treg-->pre-IEL
pre-IEL-->exTreg-IEL
exTreg-IEL-->CD4-IEL
Expression Increase
Expression Decrease
Chromatin Accessibility Increase
Chromatin Accessibility Decrease

slightly more accessible in the pre-IEL stage and their chromatin slowly closes as they become IELs. On the other hand, genes that were upregulated at the IEL stage were more accessible already at the pre-IEL stage. This data further consolidates pre-IELs as a *bona-fide* transitional stage between Tregs and exTreg-IELs. Moreover, in each Wald pairwise comparison, we analyzed how chromatin accessibility changed in genes that increased or decreased in transcription levels (**Figure 3.7g**) We found that concomitant changes in transcription and chromatin accessibility occurred primarily at the iTreg-to-pre-IEL stage, only with about 6% of changes occurring at promoter regions (**Figure 3.7g**). Transcriptional changes in the following stages were largely not accompanied by changes in chromatin accessibility. Comparison of differentially-accessible chromatin and gene expression as Tregs become IELs indicates that the Treg program begins to shut down and IEL program is initiated only at the pre-IEL stage.

The transitioning pre-IEL stage is not only marked by ThPOK loss, but also by a transcriptional shut down of the Treg program, which marks the start of chromatin accessibility of the IEL program. It is not until CD8$\alpha$ is expressed at the exTreg-IEL stage that the IEL-genes are concomitantly transcribed. Once Tregs fully transition to IELs, they display a very similar transcriptional profile to that of Tconv–derived CD4-IELs, although the former retains a small level of accessible chromatin of the Treg program. In conclusion, once $CD4^+$ T cells reach the epithelium, ThPOK is inevitably downmodulated followed by the commencement of the IEL program.

## 3.4 ThPOK downmodulation together with the epithelial environment is required for IEL differentiation

Our characterization using single-cell and bulk approaches strongly indicated that ThPOK downmodulation occurs at the crucial step of Treg destabilization to the pre-IEL stage. To mechanistically understand how ThPOK modulation allows for Tregs to proceed to the pre-IEL stage, we assessed possible genes modulated by ThPOK binding using ThPOK chromatin immunoprecipitation followed by sequencing (ChIP-seq). Due to antibody and cell number limitations for this assay, we performed ThPOK ChIP-seq on *in vivo* expanded splenic Tregs. We found that the mouse ThPOK binding motif was similar to the published human motif[163], and the protein preferentially bound to promoter regions, followed by introns (**Figure 3.8a**).

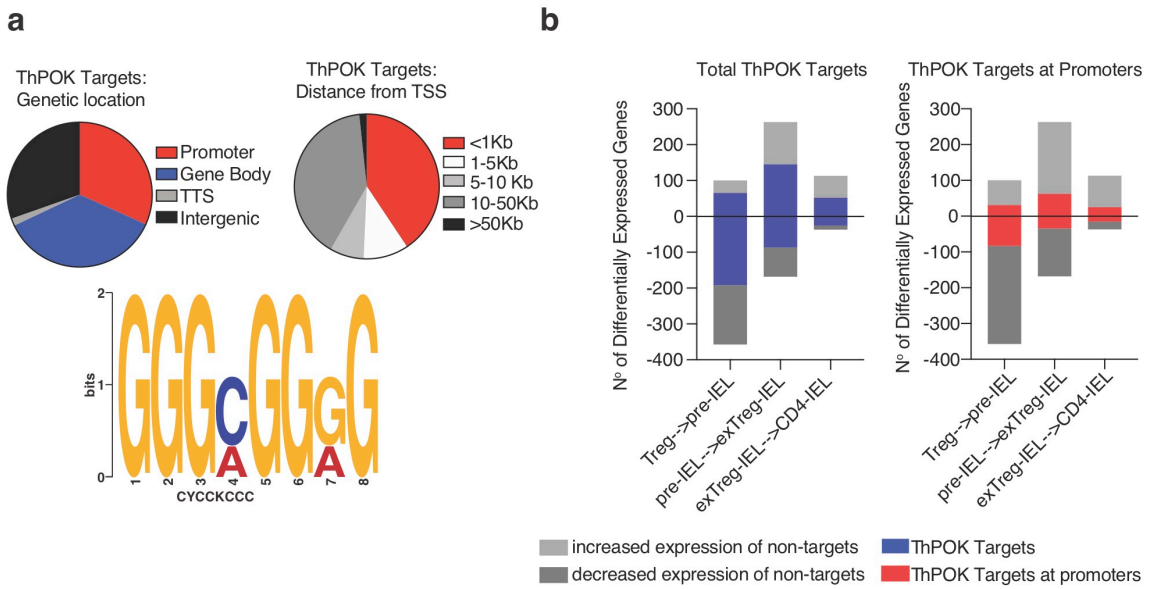



**Figure 3.8: ThPOK preferentially binds to promoter regions in Tregs.** (**a-b**) ThPOK ChIP-seq analysis of *in vivo*-expanded splenic Tregs from *Foxp3*$^{RFP}$ mice coupled with RNA-seq of IE iTregs, pre-IELs, exTreg-IELs, and CD4-IELs from i*Foxp3*$^{Tom}$ThPOK$^{GFP}$ mice. (**a**) Frequency of ChIP peaks at indicated regions (top left) and at indicated distances from transcriptional start sites TSS (top right) with *de novo* ThPOK binding motif determined by MEME-ChIP (bottom). (**b**) Numbers of differentially expressed genes (light grey for increase, dark grey for decrease) with total putative ThPOK targets (blue; left) or ThPOK targets at promoters (red, right).

Overall, genes with ThPOK binding sites were dispersed throughout the Treg to IEL progression (**Figure 3.8b**). We assessed the incidence of Runx3, ThPOK (from our analysis) and Foxp3 binding motifs in differentially-accessible chromatin regions during the iTreg to IEL progression. We separately performed this analysis on DACRs from each cluster of the ATAC-seq heatmap. We found that the Runx3 binding motif was centrally enriched in all clusters, with a marked increase during the acquisition of the IEL program (clusters 5 and 6, **Figure 3.9a**). While the ThPOK binding motif was centrally enriched in differentially-accessible chromatin regions only of cluster 1, it was accessible in other clusters, suggesting that ThPOK may play a role in both the expression of Treg and repression of IEL genes, as suggested by previous studies targeting ThPOK in mature CD4$^+$ T cells[1, 67, 164]. The Foxp3 binding motif was not centrally enriched in any cluster, implying that it may no longer play a functional role in transitioning epithelial Tregs (**Figure 3.9a**). Furthermore, to broadly investigate other possible TF regulation that may be correlated with the changes in chromatin accessibility during T cell adaptation to the IE environment, we performed TF motif analysis using MEME-ChIP[165] on each cluster of the ATAC data (**Figure 3.6d, Figure 3.9b**). Out of the top 15 significantly enriched TF motifs per cluster, those bound by the RUNX family stood out.

Accessible chromatin regions with the RUNX binding motif increased during the iTreg to IEL progression, reaching peak significance in clusters 5 and 6. ThPOK downmodulation may leave space for other TFs to bind in adjacent DNA regions, a possibility analogous to previously described roles of Runx3 to the CD8-lineage[166, 167]. For instance, Runx3 can bind to ThPOK silencer regions that in turn allows for the expression of a CD8 program in thymocytes[23].

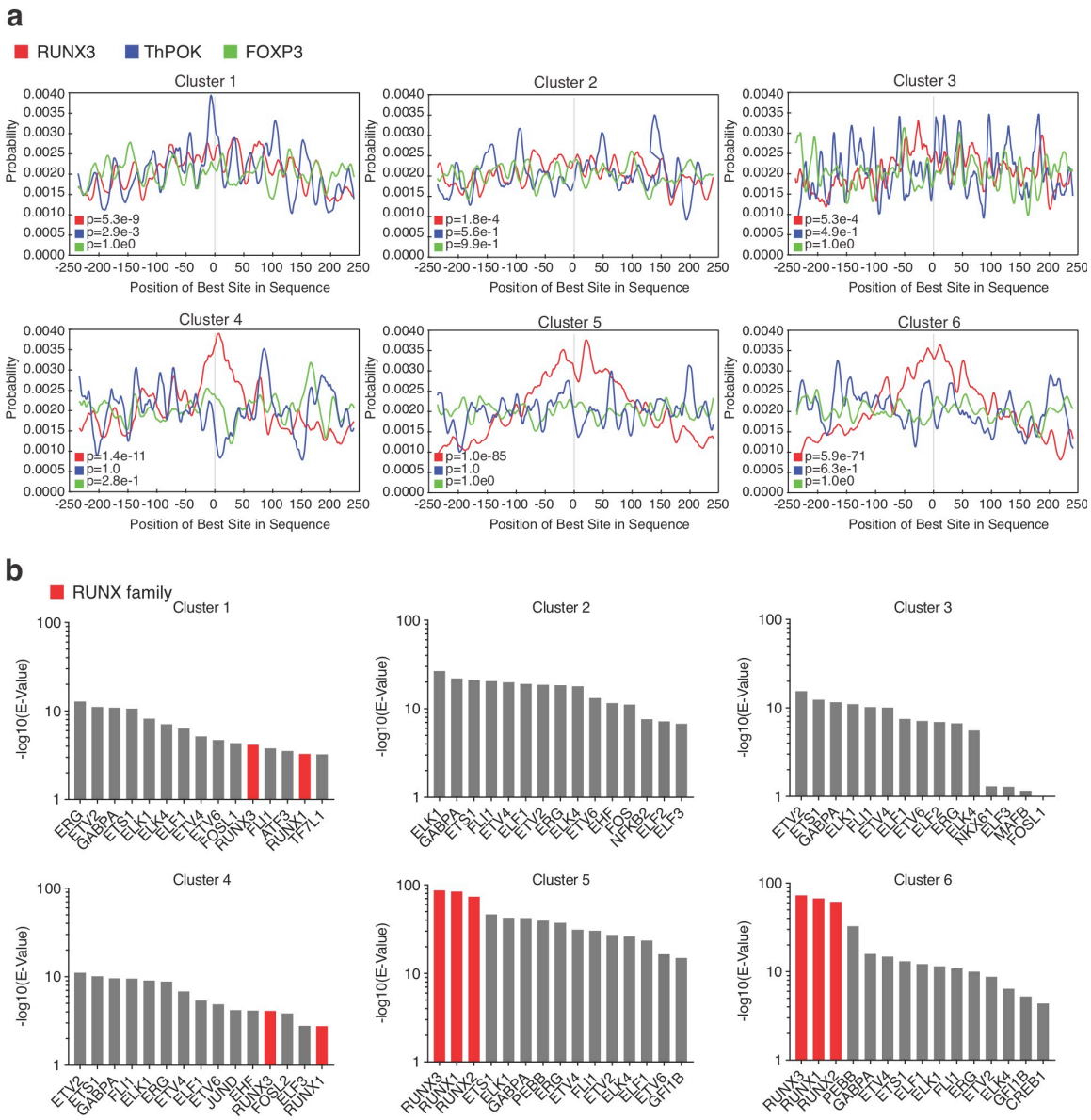
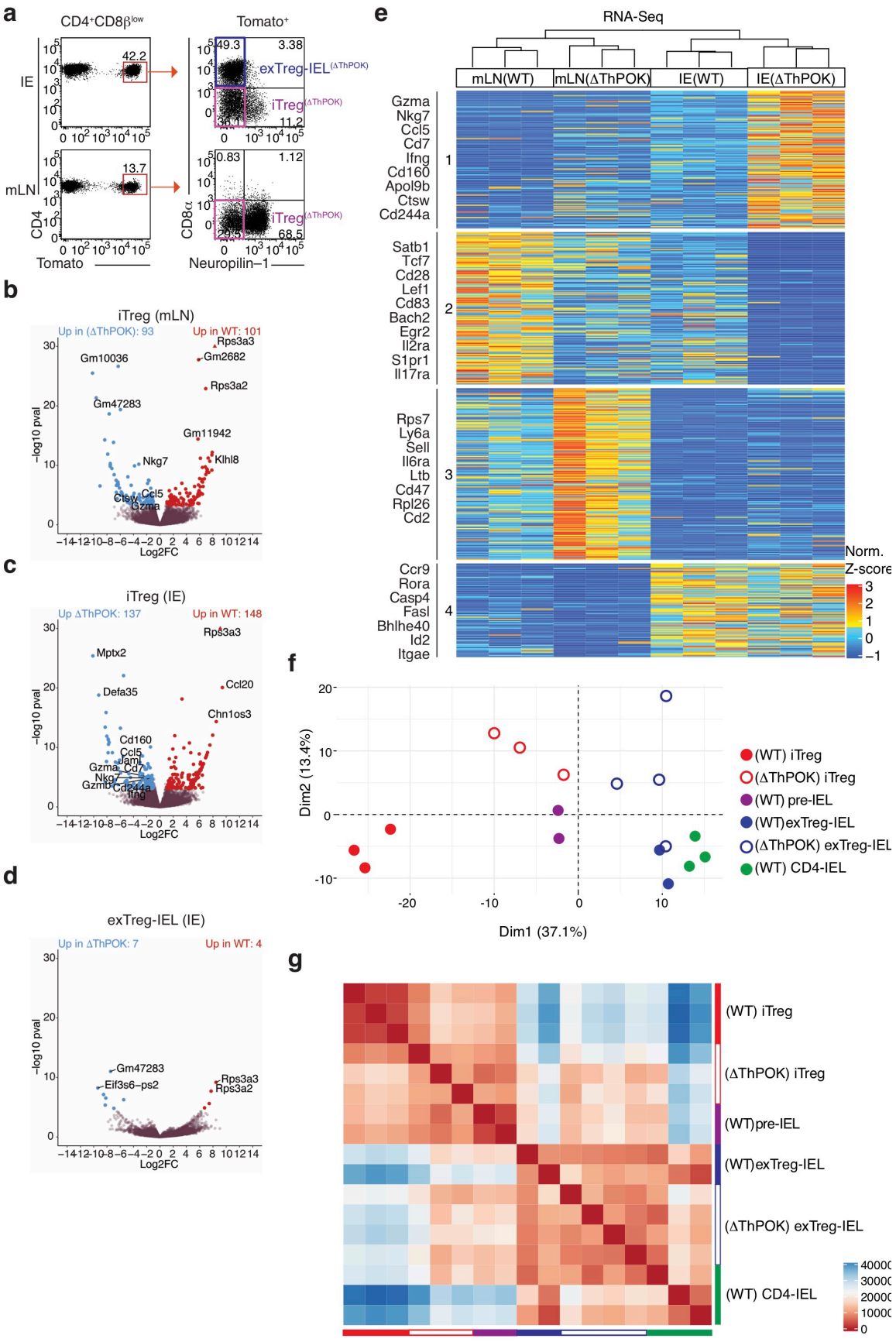



**Figure 3.9: Runx binding motifs are increasingly accessible during the iTreg to IEL progression.** (**a-b**) ThPOK ChIP-seq analysis of *in vivo*-expanded splenic Tregs from *Foxp3*$^{RFP}$ mice coupled with ATAC-seq of iTregs, pre-IELs, exTreg-IELs, and CD4-IELs from the IE of i*Foxp3*$^{Tom}$ThPOK$^{GFP}$ mice. (**a**) Transcription factor (TF) motif analysis of indicated TFs, collapsed around transcriptional start site. (**b**) TF motif analysis per cluster of ATAC-seq LRT heatmap of Figure 3.6d, performed by MEME-ChIP. Runx family TFs indicated in red.

Our sequencing data reveled that the Treg and IEL programs begin to turn off and on, respectively, at the pre-IEL stage, a transition marked by ThPOK loss. Likewise, our *in* vivo data showed that ablation of ThPOK at the Treg stage led to premature conversion of Tregs to CD4-IELs. To understand the underlying mechanisms set by ThPOK during this progression, we performed RNA and ATAC sequencing on iTregs and exTreg-IELs after tamoxifen-induced deletion of ThPOK in Tregs from i*Foxp3*$^{tomato}$(ΔThPOK) mice (**Figure 3.10a**). In agreement with our *in vivo* data in the previous chapter, expression of *Foxp3* and related Treg genes was not affected by ThPOK deletion (**Figure 3.10b-d**). Transcriptional profiling of WT and ΔThPOK mLN Tregs revealed that ThPOK loss anticipated the expression of a number of IEL-related transcripts (*Ccl5, Nkg7, Ctsw)* outside the gut environment (**Figure 3.10b, e**). ThPOK-deficient iTregs within the epithelium displayed increased expression of additional IEL genes (*Ccl5, Cd7, Jaml, Cd244a, Nkg7, Cd160, Gzma, Gzmb)* relative to their WT counterparts (**Figure 3.10c, e).** However, only a few gene expression changes between WT and ΔThPOK exTreg-IELs were detected (**Figure 3.10d**), an expected result given that ThPOK is naturally downregulated in exTreg-IELs. Principle component analysis (PCA) and Euclidean distances comparing the transcriptional profiles of all sequenced WT and ΔThPOK cell types in the IE confirmed that abrogation of ThPOK has the most effect at the iTreg stage. The profiles of both iTregs and ex-Treg-IELs in the IE of ΔThPOK mice more closely resembled the WT pre-IELs than their WT counterparts, indicating that forced ThPOK loss at the Treg stage prematurely pushes cells to differentiation to IELs before the Treg program is shut down (**Figure 3.10f, g**). This was also suggested by the increase in Foxp3$^{+}$CD8$\alpha^{+}$CD4$^{+}$ T cells we observed in the IE of mice transferred with ΔThPOK iTregs.

**Figure 3.10 Abrogation of ThPOK in Tregs anticipates progression to IELs at the transcriptional level.** (**a-g**) Induced Tregs (iTregs; Tomato$^{+}$CD8$\alpha^{-}$neuropilin-1$^{-}$) and Treg-derived CD4-IELs (exTreg-IELs; Tomato$^{+}$CD8$\alpha^{+}$) were sorted from *Zbtb7b*$^{fl/fl}$x*Runx3*$^{fl/+}$x*Rosa26*$^{lsltdTomato}$ x*Foxp3*$^{CreER}$ (i*Foxp3*$^{(\Delta ThPOK)}$) mice after 10 weeks of tamoxifen administration followed by RNA-sequencing from IE or mLN. (**a**) ΔThPOK T cell populations sequenced. (**b-d**) Volcano representation of differentially expressed genes (DEG) between WT and ΔThPOK of the same cell type as indicated, performed by Wald pairwise comparison test, $p_{adj} < 0.05$. (**e**) Heatmap of DEGs between indicated bulk populations. Expression values represents the normalized Z-score of gene abundances (TPM) (**f, g**) Principle component analysis (**f**) and corresponding Euclidean distance analysis (**g**) of DEG of all WT and ΔThPOK cell types from the IE. (n=2-4 mice).


a
CD4+CD8βlow
Tomato+
IE
mLN
42.2
13.7
49.3
3.38
exTreg-IEL(ΔThPOK)
iTreg(ΔThPOK)
36.1
11.2
0.83
1.12
68.5
CD4
CD8α
Tomato
Neuropilin−1
b
iTreg (mLN)
Up in (ΔThPOK): 93
Up in WT: 101
Rps3a3
Gm10036
Gm2682
Rps3a2
Gm47283
Gm11942
Klhl8
Nkg7
Ctsw
Ccl5
Gzma
−log10 pval
Log2FC
c
iTreg (IE)
Up ΔThPOK: 137
Up in WT: 148
Rps3a3
Mptx2
Ccl20
Defa35
Chn1os3
Cd160
Ccl5
Jaml
Gzma
Cd7
Nkg7
Gzmb
Cd244a
Ifng
−log10 pval
Log2FC
d
exTreg-IEL (IE)
Up in ΔThPOK: 7
Up in WT: 4
Gm47283
Eif3s6−ps2
Rps3a3
Rps3a2
−log10 pval
Log2FC
e
RNA-Seq
mLN(WT)
mLN(ΔThPOK)
IE(WT)
IE(ΔThPOK)
Gzma
Nkg7
Ccl5
Cd7
Ifng
Cd160
Apol9b
Ctsw
Cd244a
1
Satb1
Tcf7
Cd28
Lef1
Cd83
Bach2
Egr2
Il2ra
S1pr1
Il17ra
2
Rps7
Ly6a
Sell
Il6ra
Ltb
Cd47
Rpl26
Cd2
3
Ccr9
Rora
Casp4
Fasl
Bhlhe40
Id2
Itgae
4
Norm.
Z-score
3
2
1
0
−1
f
Dim2 (13.4%)
Dim1 (37.1%)
(WT) iTreg
(ΔThPOK) iTreg
(WT) pre-IEL
(WT)exTreg-IEL
(ΔThPOK) exTreg-IEL
(WT) CD4-IEL
g
(WT) iTreg
(ΔThPOK) iTreg
(WT)pre-IEL
(WT)exTreg-IEL
(ΔThPOK) exTreg-IEL
(WT) CD4-IEL
40000
30000
20000
10000
0

Pairwise comparisons of accessible chromatin between WT and ΔThPOK counterparts revealed minimal effects of ThPOK ablation at the Treg stage, corroborating our *in vivo* data and suggesting that the effects of ThPOK loss primarily result in transcriptional changes relative to chromatin accessibility (**Figure 3.11a, b**). About 30-40% of all differences in chromatin accessibility between the genotypes occur at promoter regions, a proportion in line with that of ThPOK binding sites (**Figure 3.11c**). Similar to the RNA-seq analysis, PCA and corresponding Euclidean distances of the ATAC-seq data also revealed that the ablation of ThPOK at the Treg stage results in both iTregs and exTreg-IELs from ΔThPOK mice to more closely resemble WT pre-IELs (**Figure 3.11d, e**). Taken together, our RNA and ATAC-sequencing data show that the untimely downmodulation of ThPOK at the Treg stage results in increased chromatin accessibility and transcription of the IEL program prior to the shutting down of the Treg program, resulting in IELs that resemble destabilized Tregs at the pre-IEL stage. This suggests that the natural IE-induced downmodulation of ThPOK allows for the full loss of the Treg program and acquisition of the IEL program in succession. Taken together, these analyses suggest that while ThPOK downmodulation initiate the acquisition of an IEL program already outside the gut, the epithelial environment is still critical for full IEL differentiation.

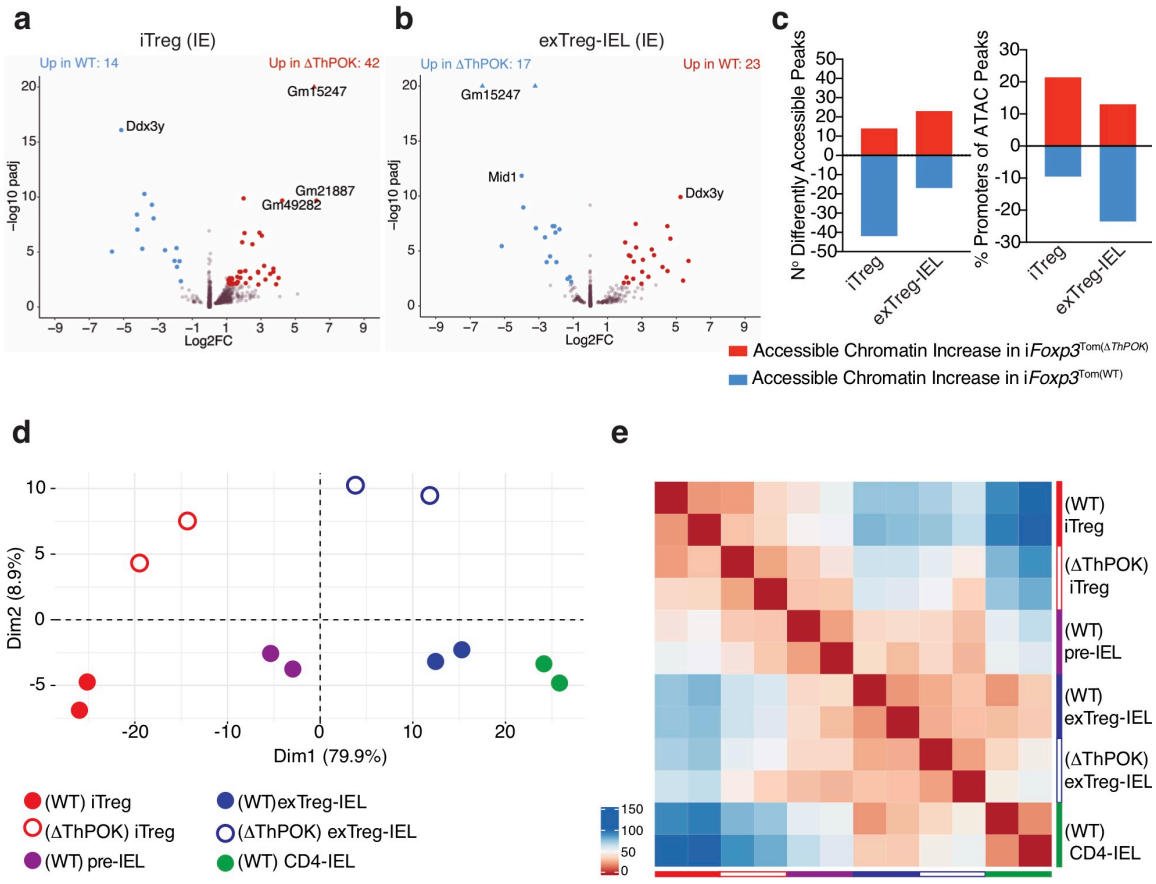



**Figure 3.11: Abrogation of ThPOK in Tregs enhances progression to IELs at the chromatin level.** (**a-e**) Induced Tregs (iTregs; Tomato$^{+}$CD8$\alpha^{-}$neuropilin-1$^{-}$) and Treg-derived CD4-IELs (exTreg-IELs; Tomato$^{+}$CD8$\alpha^{+}$) were sorted from *Zbtb7b*$^{fl/fl}$x*Runx3*$^{fl/+}$x*Rosa26*$^{lsltdTomato}$ x*Foxp3*$^{CreER}$ (i*Foxp3*$^{(\Delta ThPOK)}$) mice after 10 weeks of tamoxifen administration followed by ATAC-sequencing from IE or mLN. (**a, b**) Volcano representation of differentially accessible chromatin regions (DACR) between WT and ΔThPOK of the same cell type as indicated, performed by Wald pairwise comparison test, $p_{adj}$<0.01. (**c**) Data of accessible chromatin increase (red) and decrease (blue) between WT and ΔThPOK of the same cell type as indicated. Numbers of DACR (left) and percent of DACRs at promoter regions (right). (**d-e**) Principle component analysis (**d**) and corresponding Euclidean distance heatmap (**e**) of DACR of all WT and ΔThPOK cell types from the IE. (n=2 samples of 2-3 pooled mice).

## 3.5 Conclusion

In this study, we coupled genetic fate-mapping and gene ablation mouse models with RNA and ATAC sequencing to elucidate how the intestinal tissue induces T cell plasticity from the gut-draining LNs to the intestinal wall. Our unbiased survey

of transcriptional profiles of CD4$^+$ T cells in three enteric-associated sites revealed a strong inter- and intra- tissue adaptation, which included chromatin and transcriptional changes towards an IEL program within the gut epithelium.

We uncovered a stepwise process imprinted on T cells upon migrating to the epithelium, regardless of the subset of origin. Despite a very similar transcriptional profile between IELs derived from Tregs and conventional CD4$^+$ T cells, exTreg-IELs retained low levels of the chromatin profile from their Treg precursors. Our characterization of the intermediate pre-IEL stage reveled the gradual transcriptional and chromatin changes a cell undergoes upon adapting to the intestinal epithelium. We found that CD4$^+$ T cells shut down key features of transcriptional programs in place prior to the acquisition of the full IEL program. This coordinated process is in part due to the timely downmodulation of the CD4-lineage defining transcription factor ThPOK.

# Chapter 4: Local T cell receptor stimulation is required for intraepithelial lymphocyte differentiation, but not maintenance

## 4.1 Introduction

We sought to address how T cell receptor (TCR) properties and signaling modulate the location and plasticity of CD4$^+$ T cells in the intestine. We combined TCR repertoire analysis with single-cell transcriptomics using a fate-mapping strategy that allowed us to track Tregs and conventional T cells (Tconv) as they migrate to the intestinal epithelium and differentiate into CD8$\alpha\alpha$-expressing IELs. Reduced TCR diversity was associated with terminal differentiation of CD4$^+$ T cells into CD4-IELs; Tregs and intermediate stages were more diverse, whereas fully differentiated CD4-IELs were clonally restricted. Using *in vivo* genetic tools, we showed that ablation of surface TCR complexes on Tregs and other activated CD4$^+$ T cells impaired CD4-IEL differentiation, suggesting that TCR expression is required for terminal T cell differentiation at the intestinal epithelium. Inducible deletion of MHC class II on intestinal epithelial cells (IEC) also prevented CD4-IEL differentiation, resulting in Treg accumulation at the epithelium. However, TCR ablation in fully differentiated CD4-IELs had little, if any, impact on their accumulation or on the maintenance of the IEL transcriptional program. Our findings indicate that TCR expression and local MHC class II on IECs are required for T cell plasticity at the intestinal epithelium, but not for the maintenance of the IEL program.

## 4.2 Clonal expansion of intraepithelial CD4$^+$ T cells

Specific TCR usage has been reported in both natural- and peripherally-induced Tregs located in distinct tissues, including the intestine [168, 169, 170, 171, 172]. While in most sites Tregs are thought to stably express Foxp3, we showed that a fraction of Tregs loses Foxp3 and acquires an IEL program, including CD8$\alpha\alpha$ expression, upon migrating to the gut epithelium [1]. To define to what extent such plasticity is associated with specific TCR usage, we first analyzed the diversity of CD4$^+$ T cells in the small intestine epithelium in wild-type (WT) mice using inducible Foxp3 fate-mapping coupled with Foxp3-reporter approaches. By crossing i*Foxp3*$^{Tom}$ with *Foxp3*$^{GFP}$ mice (i*Foxp3*$^{Tom}$ *Foxp3*$^{GFP}$), we were able to specifically analyze the TCR$\alpha\beta$ repertoire of four distinct subsets of IELs upon continuous tamoxifen administration: conventional CD4$^+$ T cells (CD8$\alpha\alpha^-$, Tomato$^-$, GFP$^-$; Tconv), current Tregs (CD8$\alpha\alpha^-$, GFP$^+$) and CD4-IELs originated from either Tconv (CD8$\alpha\alpha^+$, Tomato$^-$) or from Tregs (CD8$\alpha\alpha^+$, Tomato$^+$; ex-Treg CD4-IEL) (**Figure 4.1a**). We compared the IEL TCR repertoire diversity with Tregs isolated from gut-draining mesenteric lymph nodes (mLN) and lamina propria from the same animals by single-cell TCR$\alpha\beta$ sequencing (scTCR-seq; **Figure 4.1a**). Cells with identical

TCRβ CDR3 nucleotide sequences were considered as the same clones; additionally, clonality based on TCRβ CDR3s was confirmed by sequencing the TCRα of the same expanded cells (**Figure 4.1b**).

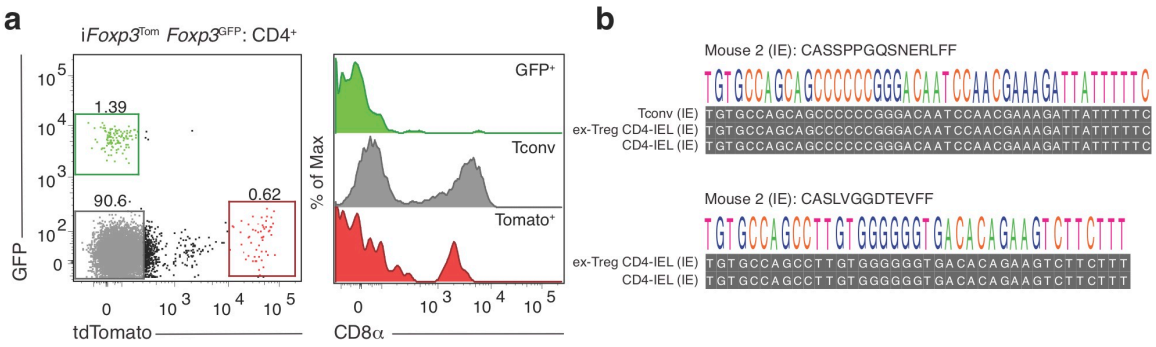
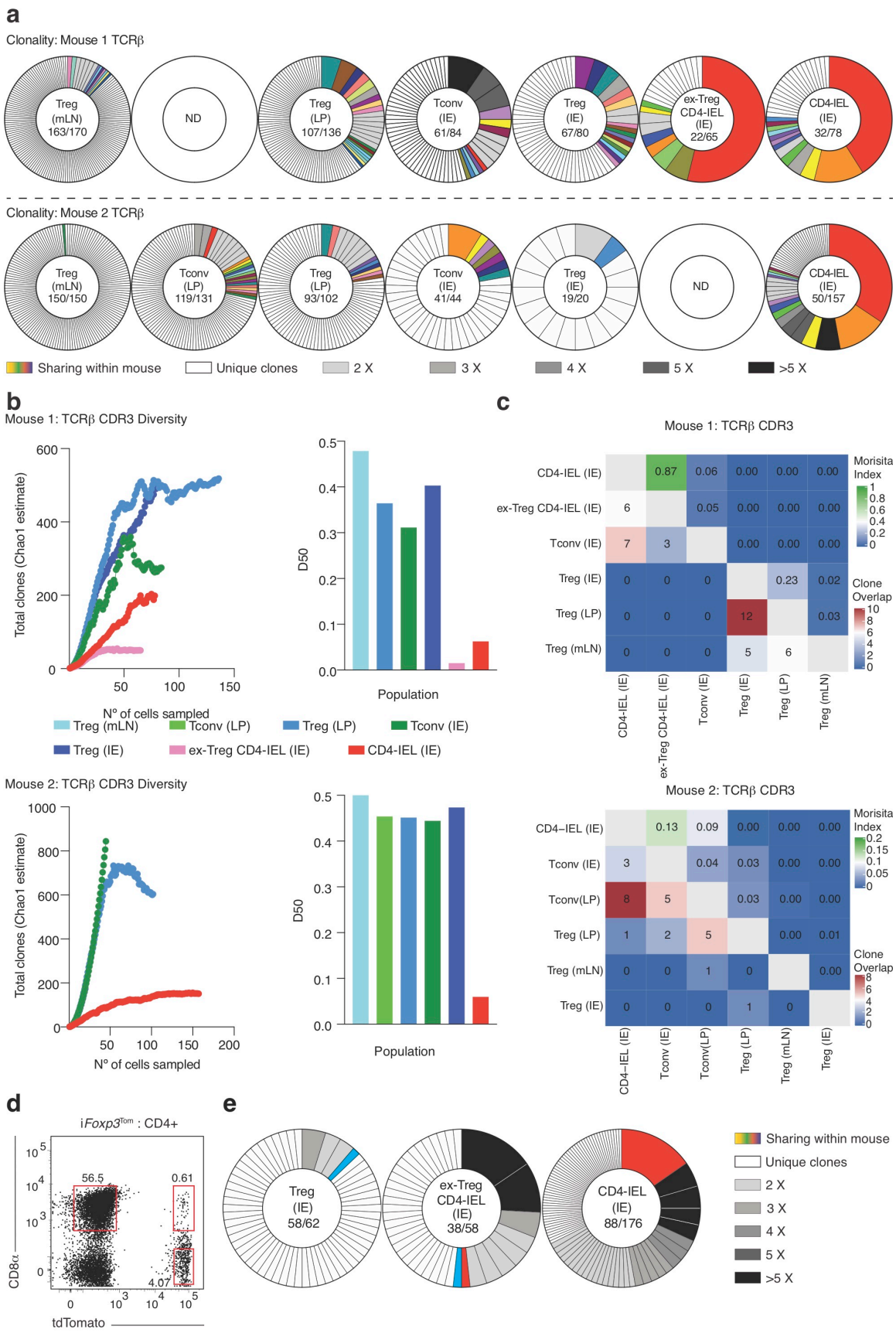



**Figure 4.1 Sorting strategy to assess intraepithelial lymphocyte TCR diversity.** (**a,b**) *Foxp3* $^{eGFP\text{-}Cre\text{-}ERT2}$ x *Rosa26* $^{lsl\text{-}tdTomato}$ x *Foxp3* $^{IRES\text{-}GFP}$ (i*Foxp3* $^{Tom}$ *Foxp3* $^{GFP}$) mice were treated with tamoxifen for 10 weeks, and CD4$^+$ T cells from mesenteric lymph nodes (mLN), lamina propria (LP) and intestinal epithelium (IE) were sorted as follows: CD4$^+$ Conventional (Tconv; GFP$^-$Tomato$^-$CD8α$^-$), regulatory T cells (Treg; GFP$^+$ or Tomato$^+$CD8α$^-$), ex-Treg CD4-IEL (Tomato$^+$CD8α$^+$) and CD4-IEL (GFP$^-$Tomato$^-$CD8α$^+$). TCRβ (and TCRα of expanded TCRβ clones) were sequenced via the MiSeq platform. (**a**) Representative dot plot of GFP$^+$ (green), tdTomato$^+$ (red) and double-negative (grey) CD4$^+$ T cells in the IE (left) and their corresponding CD8α expression (histogram, right). (**b**) Representative nucleotide sequence alignment of TCRβ CDR3 of 2 expanded shared clones.

We observed a diverse repertoire among Tregs isolated from mLN; from 320 sequenced cells, we retrieved 313 unique clones (**Figure 4.2a, b**). TCR repertoire of Tregs and Tconv isolated from LP and IE was also diverse (**Figure 4.2a, b**). In contrast, we observed large clonal expansions and reduced TCR diversity among CD4-IELs and ex-Treg CD4-IELs (**Figure 4.2a, b**). Additionally, we detected clonal sharing between LP Tregs or Tconv and all IEL subsets analyzed: Tconv, CD4-IEL or ex-Treg CD4-IEL (**Figure 4.2a, c**), suggesting that peripheral CD4$^+$ T cells first migrate to the lamina propria before entering the epithelium or that the same T cell clones migrate simultaneously to both locations. Moreover, several expanded clones were shared between ex-Treg CD4-IELs (Tomato$^+$) and CD4-IELs (Tomato$^-$), including some with Tconv, raising the possibility that a single Tconv precursor can differentiate into Tregs and then to CD4-IELs, in addition to directly converting to CD4-IELs (**Figure 4.2a, c**). Analysis of an additional i*Foxp3* $^{Tom}$ mouse showed similar results (**Figure 4.2d, e**). Because of the low proliferation capacity of CD4-IELs[66], these results indicate that potential precursors migrate to the epithelium and proliferate before they can fully differentiate into CD4-IELs.

**Figure 4.2 CD4-IELs are clonally expanded with decreased TCR diversity.**
(**a-c**) *Foxp3*$^{eGFP\text{-}Cre\text{-}ERT2}$ x *Rosa26*$^{lsl\text{-}tdTomato}$ x *Foxp3*$^{IRES\text{-}GFP}$ (i*Foxp3*$^{Tom}$ *Foxp3*$^{GFP}$) mice were treated with tamoxifen for 10 weeks, and CD4$^+$ T cells from mesenteric lymph nodes (mLN), lamina propria (LP) and intestinal epithelium (IE) were sorted as follows: CD4$^+$ conventional (Tconv; GFP$^-$Tomato$^-$CD8$\alpha^-$), regulatory T cells (Treg; GFP$^+$ or Tomato$^+$CD8$\alpha^-$), ex-Treg CD4-IEL (Tomato$^+$CD8$\alpha^+$) and CD4-IEL (GFP$^-$Tomato$^-$CD8$\alpha^+$). TCR$\beta$ were sequenced via the MiSeq platform. TCR$\alpha$ of expanded TCR$\beta$ clones were also sequenced to confirm clonality. (**a**) TCR$\beta$ clonal diversity of indicated populations from two separate mice. Each slice represents a distinct TCR$\beta$ CDR3. Colored clones represent sharing within each mouse. White slices represent unique clones and grey-scale slices represent expanded clones at indicated populations. The numbers enclosed in each graph indicates number of clones (numerator) and total number of cells (denominator) per indicated population. Empty graphs with "ND" indicate no data for corresponding cell types. (**b**) Diversity estimated by Chao1 estimate (left) and D50 (right) of indicated cells in two separate mice based on their TCR$\beta$ CDR3s. (**c**) Normalized Morisita index (top right) and number of shared clones (bottom left) of TCR$\beta$ CDR3s per cell type of both mice. (**d**) Dot plot depicting CD8$\alpha$ and tdTomato expression by intraepithelial CD4$^+$ T cells of an i*Foxp3*$^{Tom}$ mouse. Sorted Tregs (Tomato$^+$ CD8$\alpha^-$), ex-Treg CD4-IELs (Tomato$^+$CD8$\alpha^+$) and CD4-IELs (Tomato$^-$CD8$\alpha^+$) indicated by red boxes were sequenced by Sanger sequencing. (**e**) TCR$\beta$ clonal diversity of indicated IE populations. Each slice represents a distinct TCR$\beta$ CDR3. Colored clones represent sharing between populations. White slices represent unique clones and grey-scale slices represent expanded clones at indicated numbers. TCR$\alpha$ of expanded and/or shared clones were sequenced to confirm clonality.


a
Clonality: Mouse 1 TCRβ
Treg (mLN) 163/170
ND
Treg (LP) 107/136
Tconv (IE) 61/84
Treg (IE) 67/80
ex-Treg CD4-IEL (IE) 22/65
CD4-IEL (IE) 32/78
Clonality: Mouse 2 TCRβ
Treg (mLN) 150/150
Tconv (LP) 119/131
Treg (LP) 93/102
Tconv (IE) 41/44
Treg (IE) 19/20
ND
CD4-IEL (IE) 50/157
Sharing within mouse
Unique clones
2 X
3 X
4 X
5 X
>5 X
b
Mouse 1: TCRβ CDR3 Diversity
Total clones (Chao1 estimate)
N° of cells sampled
D50
Population
Treg (mLN)
Tconv (LP)
Treg (LP)
Tconv (IE)
Treg (IE)
ex-Treg CD4-IEL (IE)
CD4-IEL (IE)
Mouse 2: TCRβ CDR3 Diversity
c
Mouse 1: TCRβ CDR3
CD4-IEL (IE)
ex-Treg CD4-IEL (IE)
Tconv (IE)
Treg (IE)
Treg (LP)
Treg (mLN)
Morisita Index
Clone Overlap
Mouse 2: TCRβ CDR3
Tconv(LP)
d
iFoxp3Tom : CD4+
CD8α
tdTomato
56.5
0.61
4.07
e
Treg (IE) 58/62
ex-Treg CD4-IEL (IE) 38/58
CD4-IEL (IE) 88/176

## 4.3 Clonal distribution follows the trajectory of CD4-IEL differentiation

Our live imaging and fate-mapping studies (described in the previous chapters) suggest that emigrating CD4+ T cells quickly acquire an IEL program at the gut epithelium while losing hallmarks of peripheral CD4+ T cells or Tregs, including the expression of ThPOK and Foxp3, respectively [1, 67]. To concomitantly address intestinal epithelium-induced CD4+ T cell plasticity and specific TCR features in this process, we performed 5' single-cell RNA sequencing (scRNA-seq) coupled to TCR-seq analysis using the Chromium Single Cell V(D)J platform (10X Genomics). This strategy also allowed analysis of intra-mouse TCR sharing between current Tregs and ex-Treg CD4-IELs, which was not possible in our dual Foxp3 reporter-fate-mapping strategy used above given that *Foxp3* is X-linked. We analyzed tamoxifen-treated i*Foxp3*$^{Tom}$ mice by sorting all Tomato+ (library 1) or Tomato− (library 2) CD4+CD8β− T cells, therefore examining the whole spectrum of heterogeneity of CD4+ T cells that gained access to the epithelium. We confirmed the presence of CD4-IELs and Tregs in sorted cells by expression of CD8α and Foxp3, respectively (**Figure 4.3**). From the two libraries, we obtained a total of 1,294 scRNA-seq profiles (898 for Tomato+ library 1 and 396 for Tomato− library 2) with paired αβTCR sequences for 952 cells (651 for Tomato+ library 1 and 301 for Tomato− library 2).

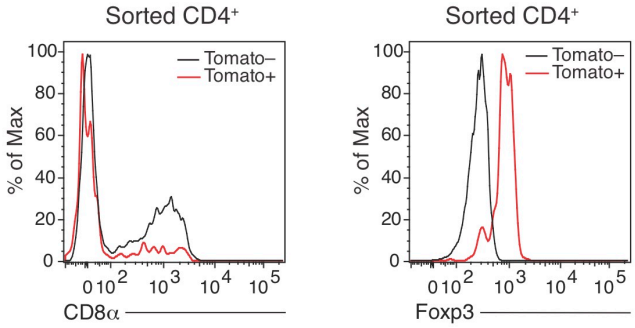
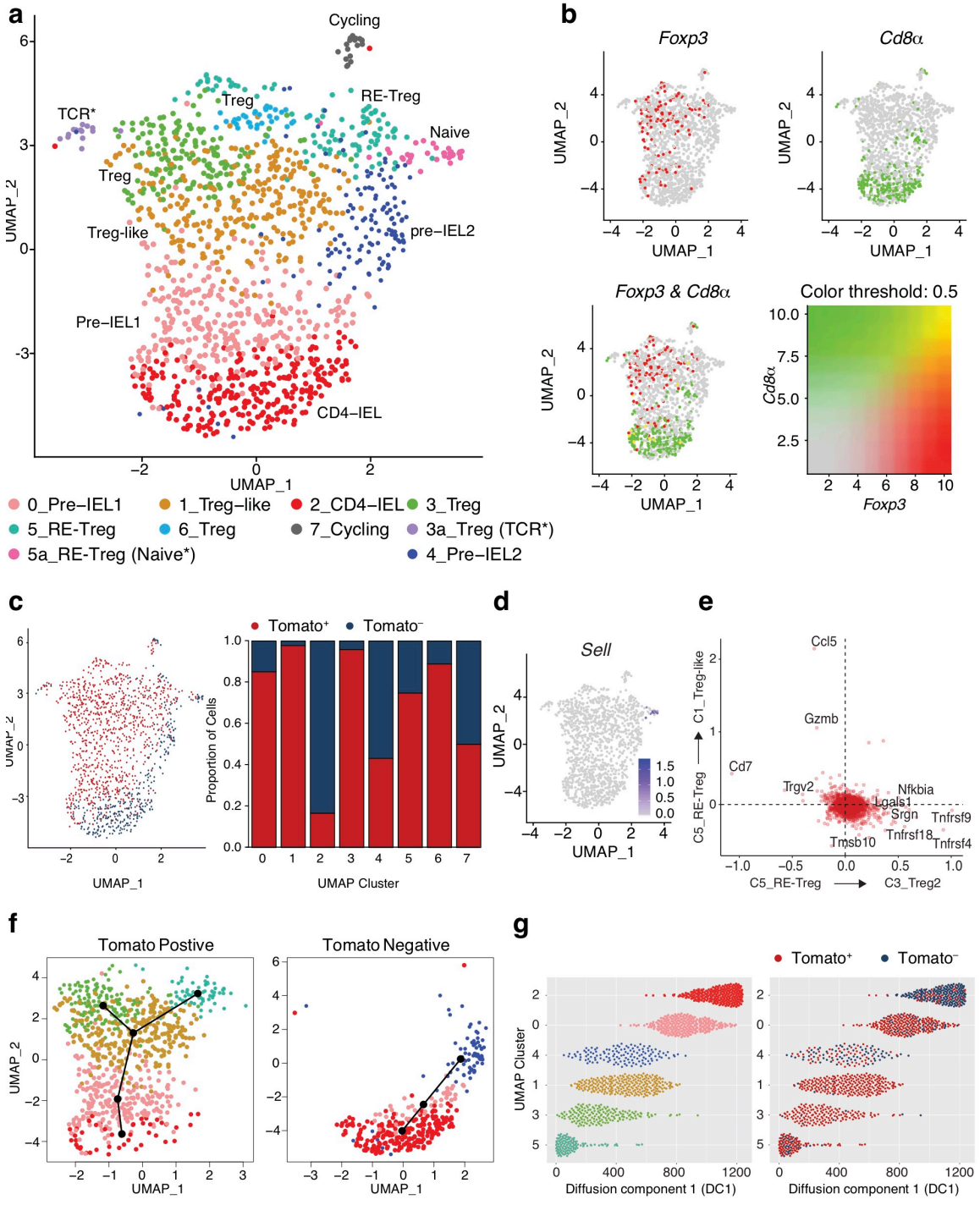



**Figure 4.3 Sorting strategy to assess intraepithelial lymphocyte clonal distribution.** i*Foxp3*$^{Tom}$ mice were treated with tamoxifen for 10 weeks, and Tomato+ (library 1) and Tomato− (library 2) CD4+ T cells from the intestinal epithelium (IE) were sorted for scRNA-seq using 10X Genomics platform. Surface CD8α and intranuclear Foxp3 expression of sequenced CD4+ Tomato+ (red) and Tomato− (black) cells.

We identified 8 clusters ordered by cell number (0-7) visualized by UMAP (**Figure 4.4a**), comparable to what we identified in the previous chapter describing the molecular mechanisms of the differentiation towards CD4-IELs. The cluster of cycling cells (7) was excluded from most of the downstream analysis due to low number of cells and strong proliferation gene signature that segregated them from all the other clusters. We identified three Treg clusters (3, 5 and 6) composed primarily of Tomato+ cells that express genes ascribed to Tregs (**Figure 4.4b, c**). Of note, cluster 5 contains cells with a profile of resting Tregs of lymphoid origin

(*Tcf7, Il7r* and *Ccr7)*, includes a small subset of *Sell*-expressing cells (Naïve*) and was classified as recent emigrant Treg (RE-Treg; **Figure 4.4d**). In addition to *bona-fide* Treg genes (*Foxp3*, *Ikzf2 Capg*), cells within Treg cluster 3 also express a profile associated to non-lymphoid tissue Tregs (*Tnfrsf4, 9, 18* and *Tigit)* [161]. A small fraction of cells in cluster 3 (TCR*) expressed high levels of *Nr4a* and *Egr* family members, related to increased TCR stimulation and activation [61]. Treg-like (cluster 1) was ascribed to cells expressing a Treg profile, yet to a lower extent than the Tregs of cluster 3, while also expressing some IEL genes (*Cd7* and *Gzmb*) (**Figure 4.4e**). Finally, we identified three non-Treg clusters, one of them composed mostly of Tomato$^+$ cells (cluster 0), indicating Treg origin, while the other two (clusters 2 and 4) contained a mix of Tomato$^+$ and Tomato$^-$ cells (**Figure 4.4a-c**). Cluster 2 was rather homogeneous containing CD8$\alpha$-expressing cells with a "full IEL program" [65, 173], which included the expression of *Gzma, Gzmb, Cd244 (2B4),*and *Itgae* (CD103) (**Figure 4.4a, b**). Trajectory analysis allowed the inference of a peripheral CD4$^+$ T cell differentiation hierarchy as they gain access to the epithelium (**Figure 4.4f-g**). We observed three different potential trajectories, all leading to the differentiation of CD4-IELs from either Tregs (Tomato$^+$) or from Tconv (Tomato$^-$) (**Figure 4.4f-g**). Clusters 0 and 4 directly preceded the CD4-IEL cluster 2 and we refer to them as "pre-IEL1" (cluster 0) and "pre-IEL2" (cluster 4). Together, these findings suggest that CD4-IELs represent a final stage of differentiation, and that peripheral CD4$^+$ T cells acquire a similar IEL program regardless of their origin (Tomato$^+$ or Tomato$^-$).

**Figure 4.4 Conventional and regulatory CD4$^+$ T cells transition through pre-IEL stages in CD4-IEL conversion.** (**a-g**) scRNA-seq analysis of Tomato$^+$ and Tomato$^-$CD4$^+$ T cells from the intestinal epithelium (IE) of a i*Foxp3*$^{Tom}$ mouse. (**a**) UMAP clustering of 8 (0-7) distinct populations of single cells, including sub-clusters (3a, 5a). Sub-cluster 3a_Treg (TCR*) indicates Tregs expressing TCR-stimulated genes. Sub-clusters 5a_RE-Treg (Naïve*) indicates *Sell*-expressing Tregs. Cluster names correspond to cell colors throughout the figure as indicated. (**b**) Expression levels of *Foxp3* (red), *Cd8a* (green), or both (yellow) by all analyzed cells. **(c**) Proportion of cells per UMAP cluster from library 1 (Tomato$^+$, red) and library 2 (Tomato$^-$, blue). (**d**) Expression levels of *Sell* by analyzed cells. (**e**) 2D volcano plot comparing clusters 5 (RE-Treg), 3 (Treg) and 1 (Treg-like). (**f**) Pseudotime trajectory analysis of Tomato$^+$ (left) and Tomato$^-$ (right) cells. (**g**) Cells on UMAP clusters ordered along diffusion component 1 (DC1), colored according to cluster (left) or tomato expression (right).


a
Cycling
Treg
RE-Treg
TCR*
Naive
Treg
Treg-like
pre–IEL2
Pre–IEL1
CD4–IEL
UMAP_1
UMAP_2
0_Pre–IEL1
1_Treg–like
2_CD4–IEL
3_Treg
5_RE-Treg
6_Treg
7_Cycling
3a_Treg (TCR*)
5a_RE-Treg (Naive*)
4_Pre–IEL2
b
Foxp3
Cd8α
Foxp3 & Cd8α
Color threshold: 0.5
c
Tomato⁺
Tomato⁻
Proportion of Cells
UMAP Cluster
d
Sell
e
Ccl5
Gzmb
Cd7
Trgv2
Nfkbia
Lgals1
Srgn
Tnfrsf9
Tnfrsf18
Tnfrsf4
Tmsb10
C1_Treg-like
C5_RE-Treg
C3_Treg2
f
Tomato Postive
Tomato Negative
g
UMAP Cluster
Diffusion component 1 (DC1)

Overlaying the TCR$\alpha\beta$ repertoire information obtained from each cell on the UMAP clusters revealed variable levels of clonal expansion and spreading across the identified clusters (**Figure 4.5a-f**). Consistent with our previous scTCR-seq analysis, the CD4-IEL cluster was the least diverse cluster, as measured by the Chao1 index. Interestingly, although pre-IEL1 cluster 0 is not as diverse, it displayed a similar level of clonal expansion, or dominance, as the CD4-IELs, as indicated by the low Diversity 50 (D50) score. Likewise, Treg and Treg-like clusters 3 and 1, respectively, displayed the next highest levels of clonal dominance (**Figure 4.5c**). Combination of pseudotime analysis (**Figure 4.4f, g**) with clonal distribution (**Figure 4.5a, e**) and clonal diversity and dominance scores (**Figure 4.5c**), suggested that clonal expansion increased as cells developed towards CD4-IELs, but did not become homogeneous until the final CD4-IEL development stage. Furthermore, this analysis revealed a series of pathways through which different expanded clones differentiate from one state to another. This was especially true for fate-mapped cells, for which our tamoxifen labeling strategy served as a timestamp as Treg clones entered the epithelium, differentiated into pre-IELs and then into ex-Treg CD4-IELs. All of the top expanded clones that were present among the CD4-IELs were also found in at least one more cluster, with distributions that tended to follow the pseudotime trajectories (**Figure 4.5a, b, d**), suggesting expanded CD4$^+$ T cells undergo differentiation at the epithelium. Our clonal composition analysis also indicates that terminal differentiation of CD4-IELs directly correlates with reduction in TCR diversity (**Figure 4.5c**). For example, the top expanded Treg-derived Tomato$^+$ clones (clones 1TP and 7TP) were present in Treg, Treg-like, pre-IEL1 and CD4-IEL clusters, while some other expanded clones (clones 3TP, 4TP, 6TP, 9TP,10TP, 12TP) were not found among CD4-IELs but were shared between Tregs, Treg-like and pre-IEL1 clusters. Of note, the top expanded Tconv-derived Tomato$^-$ clones were found within pre-IEL2 and CD4-IEL clusters (**Figure 4.5d**). It is possible that a fraction of T cell precursors did not receive sufficient signals to convert into CD4-IELs, or that the time required for conversion was longer than the timeframe of our analysis. Overall, our findings reveal a high degree of TCR sharing between rather heterogeneous gut CD4$^+$ T cell populations, with reduced diversity as cells differentiate into IELs.

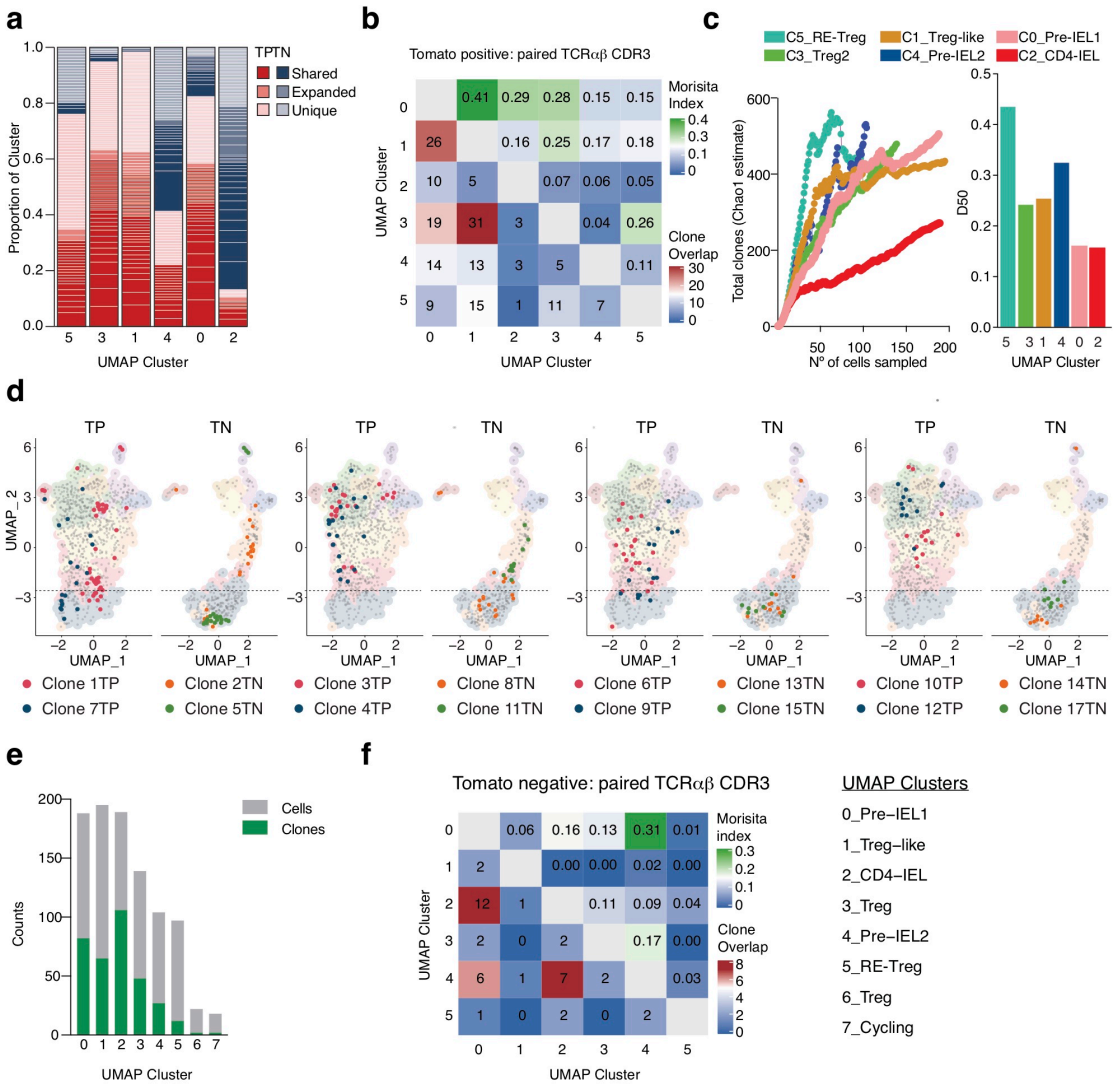
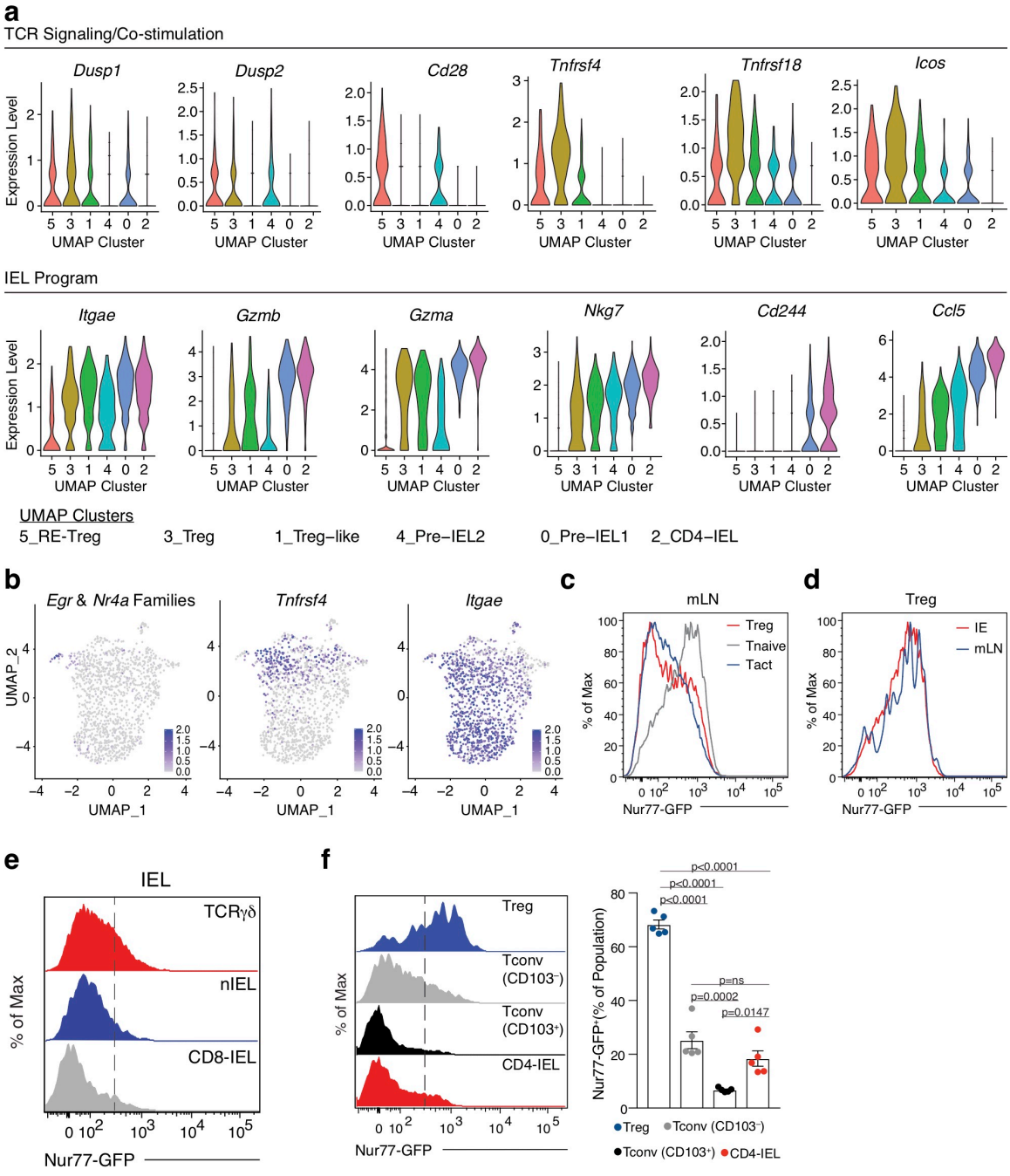



**Figure 4.5 Clonal distribution of intraepithelial CD4⁺ T cells follows single-cell transcriptional trajectories.** (**a-f**) scRNA-seq analysis of CD4⁺ T cells from a i*Foxp3*$^{Tom}$ mouse. (**a**) Paired αβTCR CDR3 of cells per UMAP cluster ordered by pseudotime trajectories of Tomato⁺ (TP, red) and Tomato⁻ (TN, blue) cells. Lightest shades indicate unique clones, intermediate shades indicate expanded, but not shared clones within TP or TN. Darkest shades indicate αβTCR sharing between clusters per TP or TN, respectively. (**b**) Normalized Morisita index (top right) and absolute number of shared clones (bottom left) of paired αβTCR per UMAP cluster among Tomato⁺ cells. (**c**) Diversity estimated by Chao1 (left) and D50 (right) of cells based on paired αβTCR per cluster. (**d**) Top expanded clones per Tomato⁺ (TP, red and blue) and Tomato⁻ (TN, orange and green) indicated in UMAP clusters separated by tomato expression. Dashed line indicates top limit of CD4-IEL cluster 2. (**e**) Total number of cells with paired αβTCR sequences (grey) and total number of clones (green) within each UMAP cluster. (**f**) Normalized Morisita index (top right) and number of shared clones (bottom left) of paired αβTCR per UMAP cluster among Tomato⁻ cells.

## 4.4 Decreased TCR signaling precedes IEL differentiation

Analysis of the scRNA-seq data showed that acquisition of IEL markers, such as *Itgae* (CD103) by $CD4^+$ T cells in the epithelium, is inversely correlated with expression of genes downstream of TCR signaling and co-stimulation, such as *Nr4a1* (Nur77) and *Tnfrsf4* (OX40). As pre-IELs acquire IEL and cytotoxic markers, they downmodulate TCR signaling molecules (**Figure 4.6a, b**). To confirm that downstream TCR signaling is associated with peripheral IEL differentiation, we analyzed Nur77 expression along differentiating peripheral IELs using *Nur77*$^{GFP}$ *Foxp3*$^{RFP}$ reporter mice (**Figure 4.5c-e**). CD4-IELs and $CD4^+CD103^+$ cells, which are enriched in pre-IELs, express lower levels of Nur77 when compared to recently-emigrated $CD4^+CD103^-$ cells or to Tregs found in the epithelium (**Figure 4.6f**). Thus, Nur77 expression is inversely associated with acquisition of CD103 and CD8$\alpha\alpha$ by $CD4^+$ T cells in the epithelium, suggestive of a role for TCR signaling in IEL differentiation from peripheral $CD4^+$ T cells.

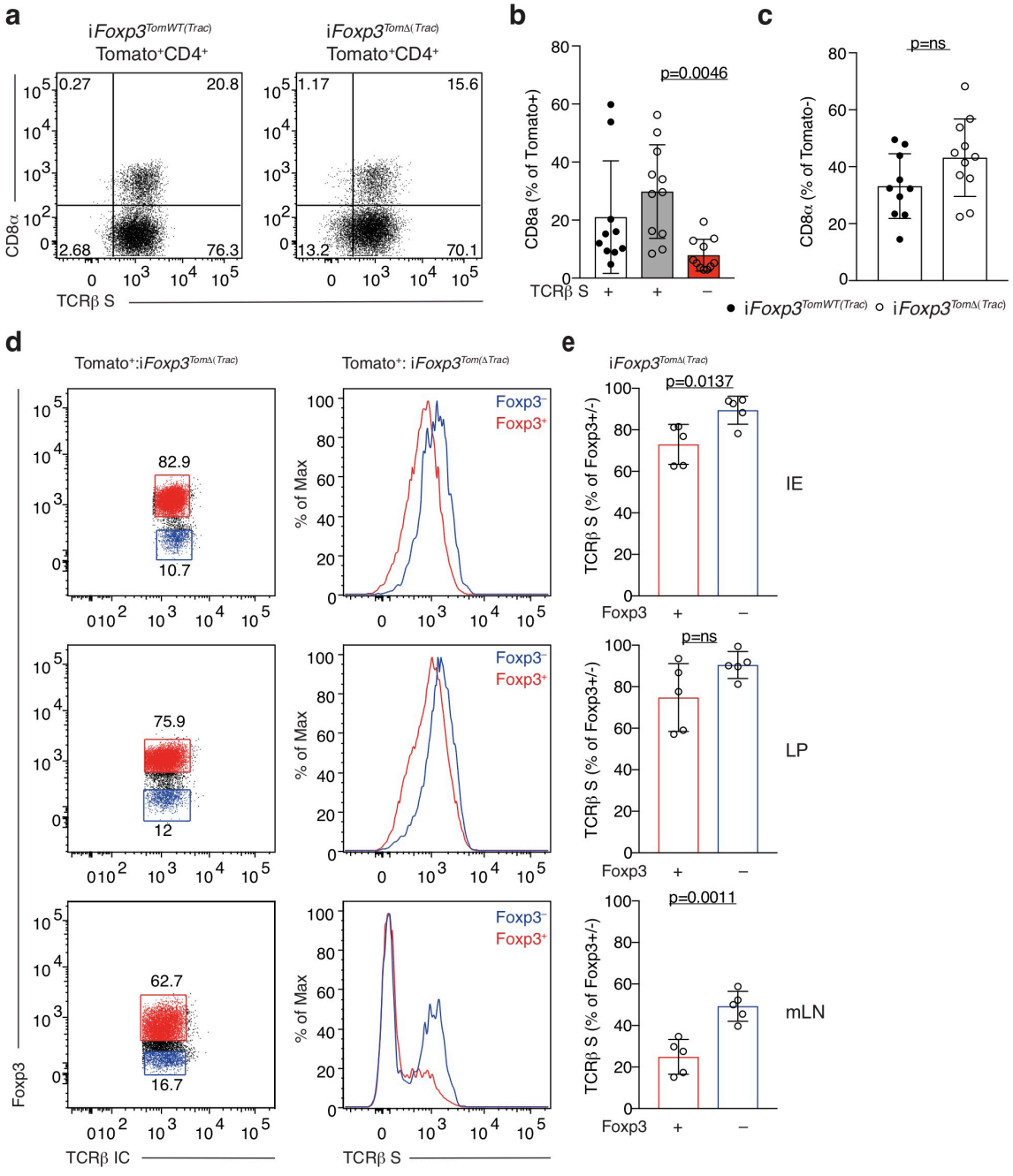
**Figure 4.6 Intraepithelial CD4$^+$ T cell subsets show an inverse correlation between TCR signaling and IEL program. (a-b**) scRNA-seq analysis of Tomato$^+$ and Tomato$^-$CD4$^+$ T cells from the intestinal epithelium (IE) of a i*Foxp3*$^{Tom}$ mouse. (**a**) Expression levels of genes related to TCR signaling (top) or IEL program (bottom) in each UMAP cluster ordered by pseudotime trajectories. (**b**) Expression levels of *Egr* and *Nr4a* families (left), *Tnfrsf4* (OX40, middle) and *Itgae* (CD103, right) in all sequenced cells. (**c-f**) *Nur77*-GFP expression levels by indicated cell types from *Nur77*$^{GFP}$ *Foxp3*$^{RFP}$ double-reporter mice. (**c**) *E*xpression among CD4$^+$ Foxp3$^+$ regulatory T cells (Treg, red), CD62L$^{high}$CD44$^{low}$ naïve T cells (Tnaïve, grey), and CD62L$^{low}$CD44$^{high}$ Foxp3$^-$ activated T cells (Tact, blue) in the mesenteric lymph nodes (mLN). (**d**) *E*xpression among Foxp3$^+$ Tregs in the intestinal epithelium (IE, red) and mLN (blue). (**e**) Expression among TCR$\gamma\delta$-IELs (red), CD8$\alpha\alpha^+$ CD8$\beta^-$CD4$^-$TCR$\alpha\beta^+$ natural IELs (nIEL, blue) and CD8$\alpha\alpha^+$ CD8$\beta^+$ TCR$\alpha\beta^+$ IELs (CD8-IEL, grey) in the small intestine epithelium. (**f**) Nur77-GFP fluorescence expression levels (left) and frequencies (right) among Foxp3$^+$ regulatory T cells (RFP$^+$, Treg, blue), conventional CD4$^+$ T cells (RFP$^-$, CD8$\alpha^-$, Tconv) CD103$^-$ (grey) or CD103$^+$ (black) cells, and CD4-IELs (RFP$^-$, CD8$\alpha^+$TL-Tetramer$^+$, red) in the intestinal epithelium.


a
TCR Signaling/Co-stimulation
Dusp1
Dusp2
Cd28
Tnfrsf4
Tnfrsf18
Icos
Expression Level
UMAP Cluster
IEL Program
Itgae
Gzmb
Gzma
Nkg7
Cd244
Ccl5
UMAP Clusters
5_RE-Treg
3_Treg
1_Treg–like
4_Pre–IEL2
0_Pre–IEL1
2_CD4–IEL
b
Egr & Nr4a Families
Tnfrsf4
Itgae
UMAP_2
UMAP_1
c
mLN
Treg
Tnaive
Tact
% of Max
Nur77-GFP
d
Treg
IE
mLN
e
IEL
TCRγδ
nIEL
CD8-IEL
f
Treg
Tconv (CD103⁻)
Tconv (CD103⁺)
CD4-IEL
p<0.0001
p<0.0001
p<0.0001
p=ns
p=0.0002
p=0.0147
Nur77-GFP⁺(% of Population)
Treg
Tconv (CD103⁻)
Tconv (CD103⁺)
CD4-IEL

## 4.5 TCR signaling is required for CD4-IEL development

To directly assess the requirement of the TCR for CD4-IEL differentiation, we employed multiple Cre mediated TCR ablation strategies targeting different stages or subsets of CD4$^+$ T cells. First, we addressed whether TCR signaling is required for Treg plasticity in the gut epithelium and subsequent CD4-IEL differentiation by crossing *Trac*$^{f/f}$ mice to i*Foxp3*$^{Tom}$ mice (to generate the i*Foxp3*$^{Tom\Delta(Trac)}$ strain), which allowed for the tracking of ex-Tregs as they lose surface TCR expression (**Figure 4.7a, b, d**). Whereas CD4-IEL differentiation from TCR-expressing (TCRβ S$^+$; Tomato$^-$ or Tomato$^+$) cells was similar between i*Foxp3*$^{Tom}$ and i*Foxp3*$^{Tom\Delta(Trac)}$ mice, it was significantly impaired among TCR-deficient (TCRβ S$^-$ Tomato$^+$) cells of i*Foxp3*$^{Tom\Delta(Trac)}$ mice (**Figure 4.7a-c**). Our scRNA-seq data showed that *Tnfrsf4* (OX40) expression changed in a developmentally controlled manner, peaking in Treg cluster 3 and then rapidly decaying by pre-IEL clusters 4 and 0 (**Figure 4.6a, b**). This pattern of expression allowed us to use the well-established OX40$^{Cre}$ driver [67] to delete the TCR from IEL precursors but not from IELs themselves. We confirmed TCR deletion in activated CD4$^+$ T cells and Tregs in the mLN of *Trac*$^{f/f}$ x OX40$^{Cre}$ (OX40$^{\Delta(Trac)}$) mice by flow cytometry (**Figure 4.8a-e**). In the epithelium, while the total frequency of CD4$^+$ T cells remained the same in OX40$^{WT(Trac)}$ and OX40$^{\Delta(Trac)}$ mice (**Figure 4.8f**), TCR ablation significantly decreased the CD4-IEL population (**Figure 4.8g-i**). We further compared the frequencies of CD4-IELs among TCR-sufficient and TCR-deficient CD4$^+$ T cells in OX40$^{\Delta(Trac)}$ mice based on surface TCRβ expression. CD4-IELs were only observed among TCR-sufficient cells (**Figure 4.8j**). Taken together, our data indicate that the differentiation of both Tregs and Tconvs to CD4-IELs requires TCR expression.

**Figure 4.7 TCR signaling is required for CD4-IEL differentiation from Tregs.** (**a-e**) Flow cytometry analysis of the intestinal epithelium (IE) of i*Foxp3*$^{Tom WT (Trac)}$ (*Trac*$^{+/+}$ i*Foxp3*$^{Tom}$) or i*Foxp3*$^{Tom\Delta (Trac)}$ (*Trac*$^{f/f}$ i*Foxp3*$^{Tom}$) mice 8-12 weeks after tamoxifen administration. (**a**) Representative dot plots of surface CD8$\alpha$ and TCR$\beta$ of Tomato$^+$ CD4$^+$ T cells in i*Foxp3*$^{Tom WT (Trac)}$ (left) and i*Foxp3*$^{Tom\Delta (Trac)}$ (right) animals. (**b**) Frequencies of CD8$\alpha^+$ cells among Tomato$^+$ CD4$^+$ T cells within TCR-sufficient cells from i*Foxp3*$^{Tom WT (Trac)}$ (white bar) or i*Foxp3*$^{Tom\Delta (Trac)}$ (grey bar) mice, or TCR-deficient cells from i*Foxp3*$^{Tom\Delta (Trac)}$ (red bar) mice. (**c**) Frequency of CD8$\alpha$-expressing CD4$^+$ T (CD4-IEL) cells among tomato$^-$ cells in the IE of i*Foxp3*$^{Tom WT (Trac)}$ or i*Foxp3*$^{Tom\Delta (Trac)}$ mice. **(d**) Representative dot plots (left) for intracellular TCR$\beta$ and Foxp3, or histograms (right) of surface TCR$\beta$ expression among tomato$^+$ Foxp3$^+$ Tregs (red) and tomato$^+$ Foxp3$^-$ CD4$^+$ T cells (blue) in the intestinal epithelium (IE, top), lamina propria (LP, middle) and mesenteric lymph nodes (mLN, bottom). (**e**) Frequencies of surface TCR$\beta$-expressing cells among Foxp3$^+$ or Foxp3$^-$ among tomato$^+$ CD4$^+$ T cells from i*Foxp3*$^{Tom\Delta (Trac)}$ IE (top), LP (middle) or mLN (bottom).


a
iFoxp3TomWT(Trac)
Tomato+CD4+
iFoxp3TomΔ(Trac)
Tomato+CD4+
0.27
20.8
2.68
76.3
1.17
15.6
13.2
70.1
CD8α
TCRβ S
b
p=0.0046
CD8a (% of Tomato+)
TCRβ S + + −
c
p=ns
CD8α (% of Tomato-)
iFoxp3TomWT(Trac) iFoxp3TomΔ(Trac)
d
Tomato+:iFoxp3TomΔ(Trac)
Tomato+: iFoxp3Tom(ΔTrac)
82.9
10.7
75.9
12
62.7
16.7
Foxp3
TCRβ IC
Foxp3−
Foxp3+
% of Max
TCRβ S
e
iFoxp3TomΔ(Trac)
p=0.0137
IE
p=ns
LP
p=0.0011
mLN
TCRβ S (% of Foxp3+/-)
Foxp3 + −

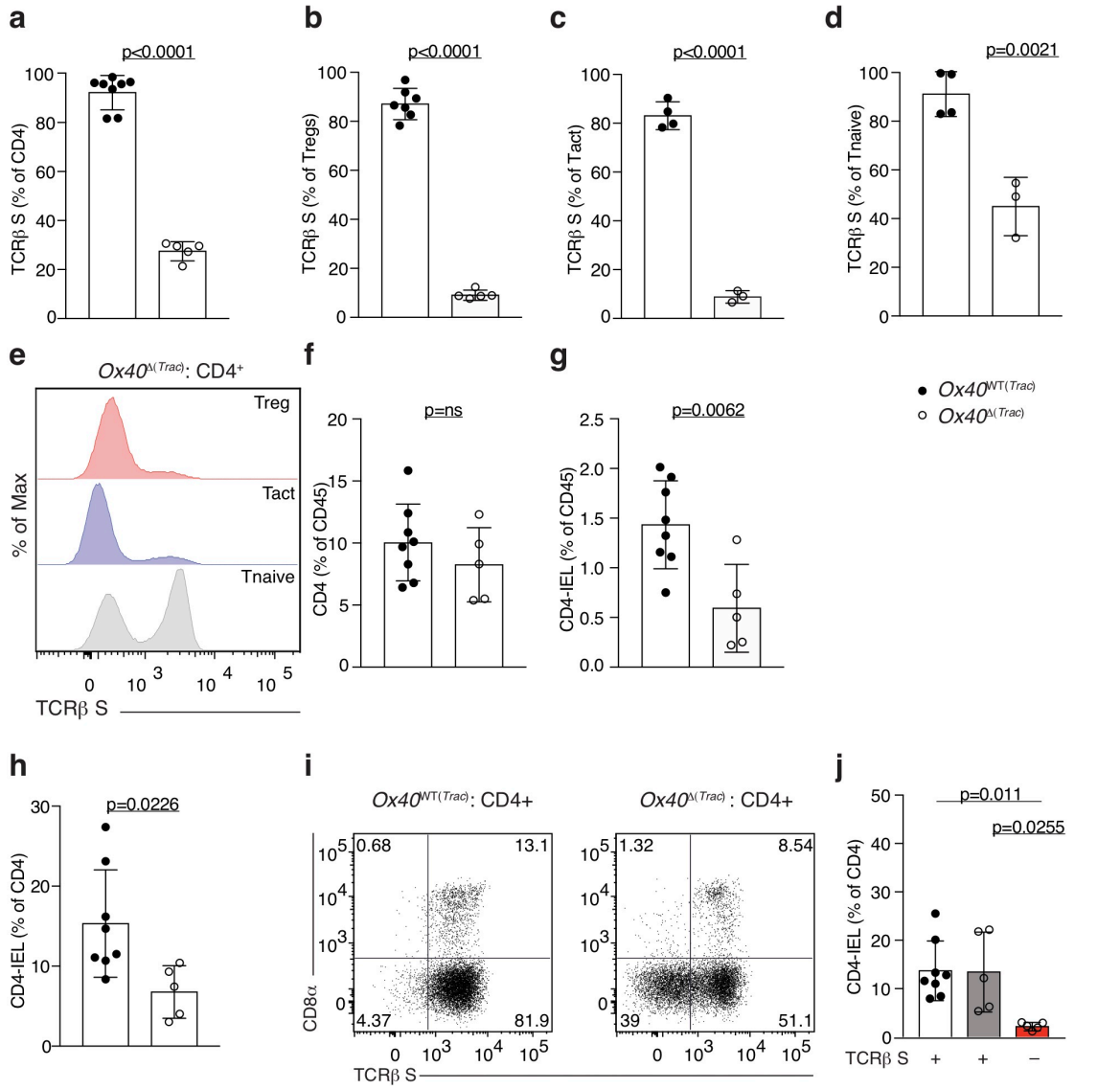
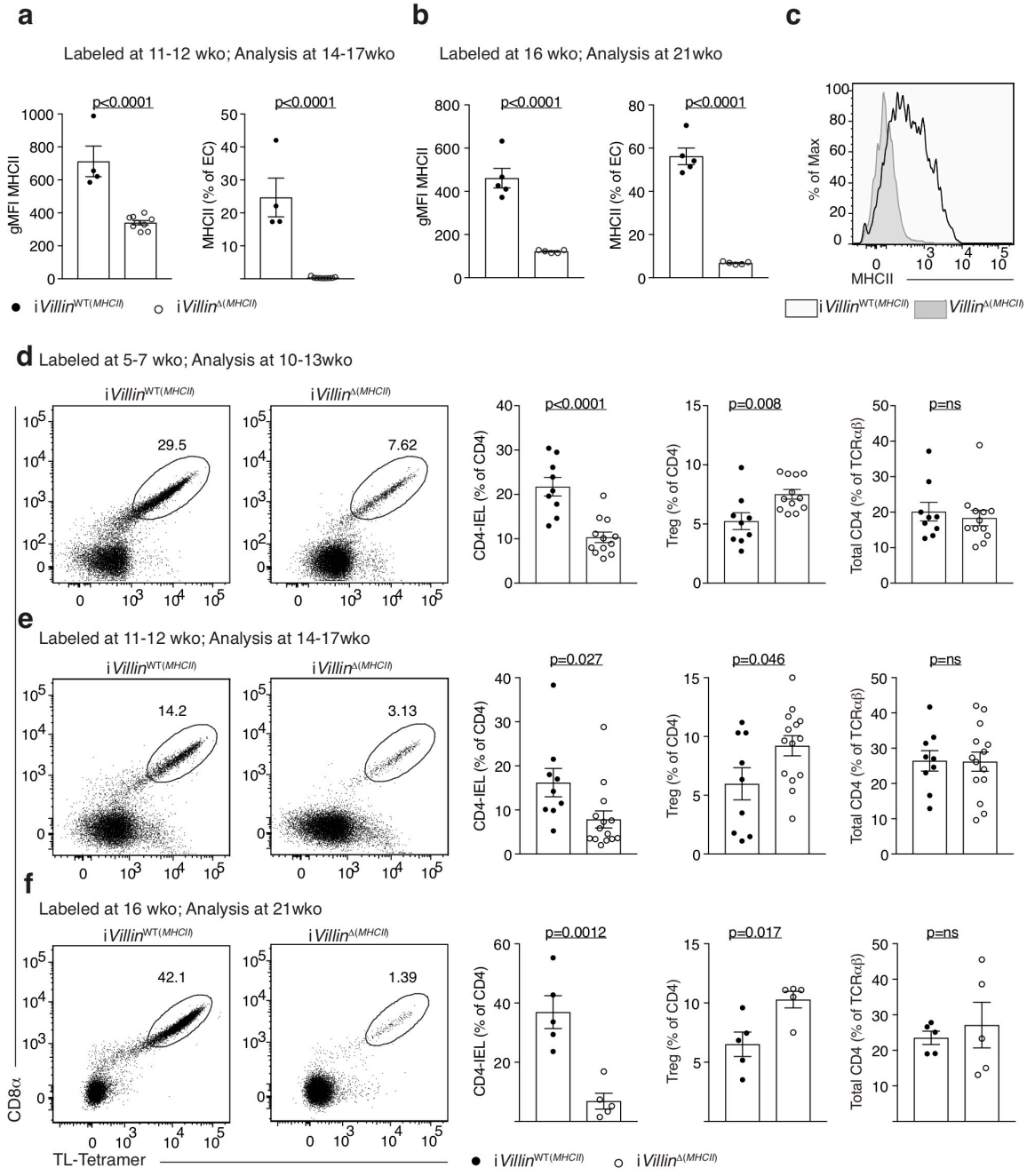



**Figure 4.8 TCR signaling is required for CD4-IEL differentiation from regulatory and conventional CD4⁺ T cells.** (**a-j**) Flow cytometry analysis of IE and mLN cells of 9-12-week-old OX40$^{WT(Trac)}$ (*Trac*$^{+/+}$ OX40Cre$^{+/-}$ or *Trac*$^{f/f}$ OX40Cre$^{-/-}$) or OX40$^{\Delta(Trac)}$ (*Trac*$^{f/f}$ OX40Cre$^{+/-}$) mice. (**a-d**) Frequencies of surface TCRβ expression among total CD4⁺ T cells (**a**), Tregs (**b**), CD4⁺ activated (Tact, Foxp3$^{-}$CD44$^{high}$CD62L$^{low}$) (**c**) and CD4⁺ Naïve (Tnaïve, Foxp3$^{-}$CD44$^{low}$CD62L$^{high}$) (**d**) cells in the mLN. (**e**) Surface TCRβ expression among Tregs (red), Tact (blue) and Tnaïve (grey) CD4⁺ T cells in the mLN. (**f-g**) Frequencies among total CD45⁺ cells in the IE of CD4⁺ cells (**f**) and CD4-IELs (**g**). (**h**) Frequency of CD4-IELs among CD4⁺ T cells. (**i**) Representative plots of surface CD8α and TCRβ of CD4⁺ T cells. (**j**) Frequencies of CD4-IELs among conventional CD4⁺ T cells (Tconv, CD4⁺CD8α$^{-}$Foxp3$^{-}$) within TCR-sufficient cells from OX40$^{WT(Trac)}$ (white bar) or OX40$^{\Delta(Trac)}$ (grey bar) mice, or TCR-deficient cells from OX40$^{\Delta(Trac)}$ (red bar) mice. Data are expressed as mean +/- SEM of individual mice (n=5-8). Significant p values as indicated [student's t test or one-way ANOVA and Bonferroni].

## 4.6 MHC-II expression on epithelial cells modulates CD4-IEL differentiation

Recent studies have suggested a role for intestinal epithelial cell (IEC)-mediated antigen presentation via MHC-II in the regulation of intestinal $CD4^+$ T cell function [105, 174]. We therefore asked whether local MHC-II expression by IECs is required for CD4-IEL differentiation or maintenance. We targeted MHC class II expression exclusively on IECs by crossing $Villin^{CreERT2}$ mice to *H2-Ab1*$^{f/f}$ mice ($Villin^{\Delta(MHCII)}$) (**Figure 4.9a-c**). Tamoxifen treatment of $Villin^{\Delta(MHCII)}$ mice starting at 5-7 weeks of age (prior to the appearance of CD4-IELs in the epithelium) did not affect the frequency of total $CD4^+$ T cells in the epithelium 5-6 weeks later; however, it led to a significant reduction in the frequency of CD4-IELs, which was accompanied by an increase in Treg frequency (**Figure 4.9d**). Tamoxifen treatment of 11, 12- or 16-week-old mice, which typically carry a sizable population of CD4-IELs, also significantly impacted CD4-IEL frequencies and conversely led to an accumulation of Tregs in the epithelium (**Figure 4.9e, f**), suggesting that MHC-II expression on IECs is required for continuous differentiation into CD4-IELs in adult mice.

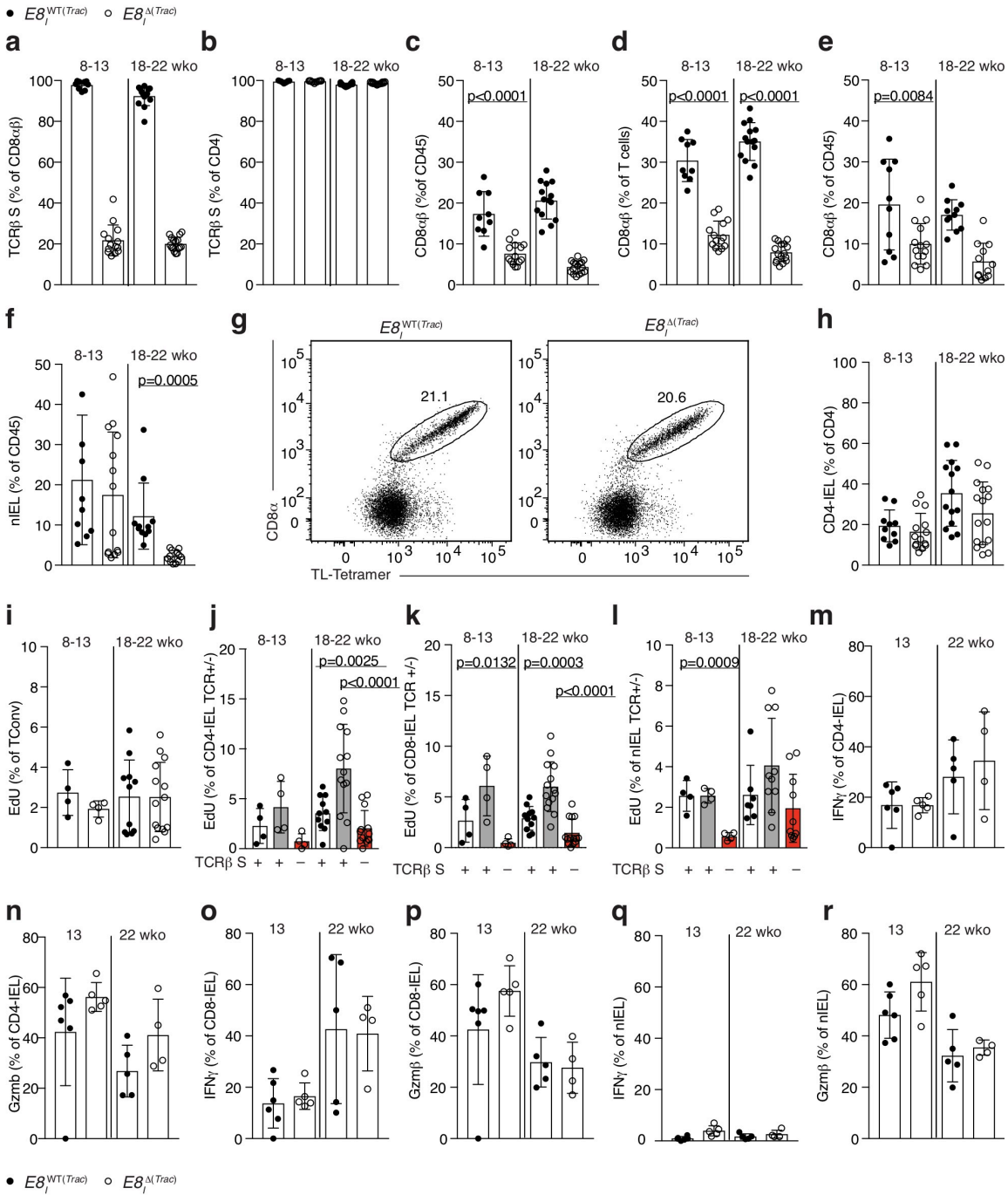
**Figure 4.9 MHC-II expression by epithelial cells is required for CD4-IEL conversion).** (**a-f**) Flow cytometry analysis of the intestinal epithelium (IE) of i*Villin*$^{WT(MHCII)}$ (*H2-Ab1*$^{+/+}$ Villin$^{CreERT2+/-}$ or *H2-Ab1*$^{f/f}$ Villin$^{CreERT2-/-}$) or i*Villin*$^{\Delta(MHCII)}$ (*H2-Ab1*$^{f/f}$ Villin$^{CreERT2+/+}$) mice after tamoxifen administration. (**a-b**) Geometric mean fluorescence intensity (gMFI) (left) and frequency (right) of MHC-II expression by epithelial cells of 14-17-week-old mice 3-5 weeks after tamoxifen administration (**a**) and of 21-week-old mice 5 weeks after tamoxifen administration (**b**). (**c**) Representative histogram of MHC-II expression by EpCAM$^+$ epithelial cells in 20-week-old i*Villin*$^{WT(MHCII)}$ (white) or i*Villin*$^{\Delta(MHCII)}$ (grey) mice 4 weeks after tamoxifen administration. (**d-f**) Representative dot plots of surface CD8$\alpha$ and TL-Tetramer among CD4$^+$ T cells (left) and frequencies of CD4-IELs (CD4$^+$CD8$\alpha^+$TL-Tetramer$^+$) or Foxp3$^+$ regulatory cells (Tregs) among CD4$^+$ T cells (middle), and total CD4$^+$ T cells among TCR$\alpha\beta^+$ cells (right).) (**d)** Tamoxifen administration to 5-7-week-old mice, analyzed at 10-13 weeks, (**e**) to 11-12-week-old mice, analyzed at 14-17 weeks of age, (**f**) to 16-week-old mice, analyzed at 21 weeks of age. Data are expressed as mean +/- SEM of individual mice (n=5-14). Significant p values as indicated [student's t test].


a
Labeled at 11-12 wko; Analysis at 14-17wko
p<0.0001
p<0.0001
gMFI MHCII
MHCII (% of EC)
b
Labeled at 16 wko; Analysis at 21wko
p<0.0001
p<0.0001
c
% of Max
MHCII
d Labeled at 5-7 wko; Analysis at 10-13wko
29.5
7.62
p<0.0001
p=0.008
p=ns
CD4-IEL (% of CD4)
Treg (% of CD4)
Total CD4 (% of TCRαβ)
e
Labeled at 11-12 wko; Analysis at 14-17wko
14.2
3.13
p=0.027
p=0.046
p=ns
f
Labeled at 16 wko; Analysis at 21wko
42.1
1.39
p=0.0012
p=0.017
p=ns
CD8α
TL-Tetramer

## 4.7 TCR signaling is largely dispensable for IEL program maintenance

The transition to an IEL program includes upregulation of NK- and cytolytic-molecules as well as CD8αα [65, 173], with concomitant downregulation of TCR signaling. To evaluate the role of the TCR signaling in establishing the IEL program, and in the maintenance of differentiated IELs, we crossed *Trac*$^{f/f}$ mice to those expressing Cre under the enhancer I of the *Cd8a* gene (E8I$^{\Delta(Trac)}$). E8I is required for CD8α expression on mature T cells and CD8αα IELs, but not required for CD8αβ or CD8αα expression in developing thymocytes [175]. In E8I$^{\Delta(Trac)}$ mice, TCR deletion is accompanied by a decrease in CD8α– expressing TCRαβ$^+$ cells in the mLN and IE, including CD4$^-$CD8β$^+$CD8α$^+$ T cells and CD4$^-$CD8β$^-$CD8αα$^+$ TCRαβ$^+$ natural IELs (nIEL) (**Figure 4.10a-f**). In the case of peripheral CD4$^+$ T cells, which only express CD8αα in the final stages of their differentiation into CD4-IEL, this model allowed us to selectively address the role of the TCR in the maintenance of differentiated CD4-IELs. We did not observe any significant changes in the accumulation of CD4-IELs in E8I$^{\Delta(Trac)}$ mice when compared to WT littermates (E8I$^{WT(Trac)}$), even in older animals (**Figure 4.10g, h**), suggesting that upon terminal differentiation, CD4-IELs do not rely on the TCR for their maintenance in the epithelium. Whereas proliferation, measured by Edu incorporation, was similar among Tconv from E8I$^{WT(Trac)}$ and E8I$^{\Delta(Trac)}$ mice, remaining TCR-expressing CD4-IELs in E8I$^{\Delta(Trac)}$ mice proliferated more than those of E8I$^{WT(Trac)}$ mice, but only in older animals (**Figure 4.10i, j**). Therefore, similar frequencies of CD4-IELs cannot be exclusively attributed to differential proliferative abilities of TCR-sufficient and TCR-deficient cells. Similar proliferation rates were observed in other CD8αα-expressing IEL subsets, including CD8αβ$^+$CD8αα$^+$ TCRαβ$^+$ (CD8-IELs) and CD8αα$^+$ nIELs (**Figure 4.10k, l**). Common IEL functional readouts such as IFNγ production and granzyme B expression showed that hallmarks of the CD4-IEL phenotype are maintained in the absence of the TCR (**Figure 4.10m, n**). Analysis of CD8-IELs and nIELs also revealed intact IFNγ and granzyme B production, despite TCR loss (**Figure 4.10o-r**). Taken together, our results suggest that CD4-IEL accumulation and at least some of the characteristic features of IELs are maintained in the absence of TCR signaling.

**Figure 4.10 TCR signaling is not essential for CD4-IEL maintenance.**
**(a-r)** Flow cytometry analysis of 8-22-week-old E8I$^{WT(Trac)}$ (*Trac*$^{+/+}$ E8ICre$^{+}$ or *Trac*$^{+/+}$ E8ICre$^{-}$ or *Trac*$^{f/f}$ E8ICre$^{-}$) or E8I$^{\Delta(Trac)}$ (*Trac*$^{f/f}$ E8ICre$^{+}$) mice, grouped by age as indicated. Frequencies of surface TCRβ-expressing cells among CD8αβ cells (**a**) and CD4 (**b**) T cells in the mesenteric lymph nodes (mLN). Frequency of CD8αβ T cells among total CD45$^{+}$ cells (**c**) and intracellular TCRβ-expressing T cells (**d**) in the mLN. Frequencies of CD8αβ (**e**) and natural IELs (nIEL, CD4$^{-}$ CD8αα$^{+}$CD8β$^{-}$TL-Tetramer$^{+}$) (**f**) among CD45$^{+}$ cells in the intestinal epithelium (IE). (**g**) Representative dot plots of surface CD8α and TL-Tetramer among CD4$^{+}$ T cells in E8I$^{WT(Trac)}$ (left) or E8I$^{\Delta(Trac)}$ (right) mice. (**h**) Frequency of CD4-IELs (CD8α$^{+}$TL-Tetramer$^{+}$) among CD4$^{+}$ T cells in E8I$^{WT(Trac)}$ and E8I$^{\Delta(Trac)}$ mice. (**I,j**) Frequency of proliferation (measured by EdU incorporation) of CD4$^{+}$CD8α$^{-}$Foxp3$^{-}$ cells (Tconv) (**i**) among cells with or without surface TCR expression (**j**). Mice were injected with Edu 16 and 4 hours prior to analysis. (**k,l**) Proliferation (EdU incorporation) of CD8-IELs (**k**) or nIELs (**l**) with or without surface TCRβ expression after EdU injection 16 and 4 hours prior to analysis. (**m,n**) Frequencies of IFNγ (**m**) and Gzmb (**n**) production upon PMA/Ionomycin *ex-vivo* stimulation among CD4-IELs. (**o-r**) Frequencies of IFNγ (**o,q**) and Gzmb (**p,r**) production upon PMA/Ionomycin *ex-vivo* stimulation among CD8-IELs (**o,p**) and nIELs (**q,r**). Data are expressed as mean +/- SEM of individual mice. (n=4-16). Significant p values as indicated [student's t test or one-way ANOVA and Bonferroni].


• E8I WT(Trac) ○ E8I Δ(Trac)
a
8-13
18-22 wko
TCRβ S (% of CD8αβ)
b
TCRβ S (% of CD4)
c
p<0.0001
CD8αβ (%of CD45)
d
p<0.0001
p<0.0001
CD8αβ (% of T cells)
e
p=0.0084
CD8αβ (% of CD45)
f
p=0.0005
nIEL (% of CD45)
g
E8I WT(Trac)
E8I Δ(Trac)
21.1
20.6
CD8α
TL-Tetramer
h
CD4-IEL (% of CD4)
i
EdU (% of TConv)
j
p=0.0025
p<0.0001
EdU (% of CD4-IEL TCR+/-)
TCRβ S + + – + + –
k
p=0.0132
p=0.0003
p<0.0001
EdU (% of CD8-IEL TCR +/-)
l
p=0.0009
EdU (% of nIEL TCR+/-)
m
13
22 wko
IFNγ (% of CD4-IEL)
n
Gzmb (% of CD4-IEL)
o
IFNγ (% of CD8-IEL)
p
Gzmβ (% of CD8-IEL)
q
IFNγ (% of nIEL)
r
Gzmβ (% of nIEL)

To comprehensively address the extent to which the IEL program can be affected or maintained in the absence of TCR signaling, we performed bulk RNA-seq on TCR-expressing CD4-IELs from $E8I^{WT(Trac)}$ mice and on TCR-deficient CD4-IELs from $E8I^{\Delta(Trac)}$ mice (**Figure 4.11a**). While genes related to TCR activation or signaling, cell cycle or proliferation, anti-apoptosis and protein transport showed reduced expression in cells lacking their TCR, most IEL-related genes were not significantly changed in $E8I^{\Delta(Trac)}$ mice, suggesting the IEL program is maintained in the absence of TCR expression (**Figure 4.11b-d**). Gene ontology (GO) enrichment analysis showed that only a few pathways, such as protein phosphorylation, were upregulated in TCR-deficient cells (**Figure 4.11c**). Likewise, IEL genes were not changed among CD8-IELs or nIELs in $E8I^{\Delta(Trac)}$ mice, but genes related to TCR signaling were significantly decreased in TCR-deficient cells (**Figure 4.11e, f**). Of note, Lag3 (Lymphocyte-activation gene 3), a TCR inhibitory co-receptor, was downmodulated in all 3 IEL subsets lacking TCR expression (**Figure 4.11b**).

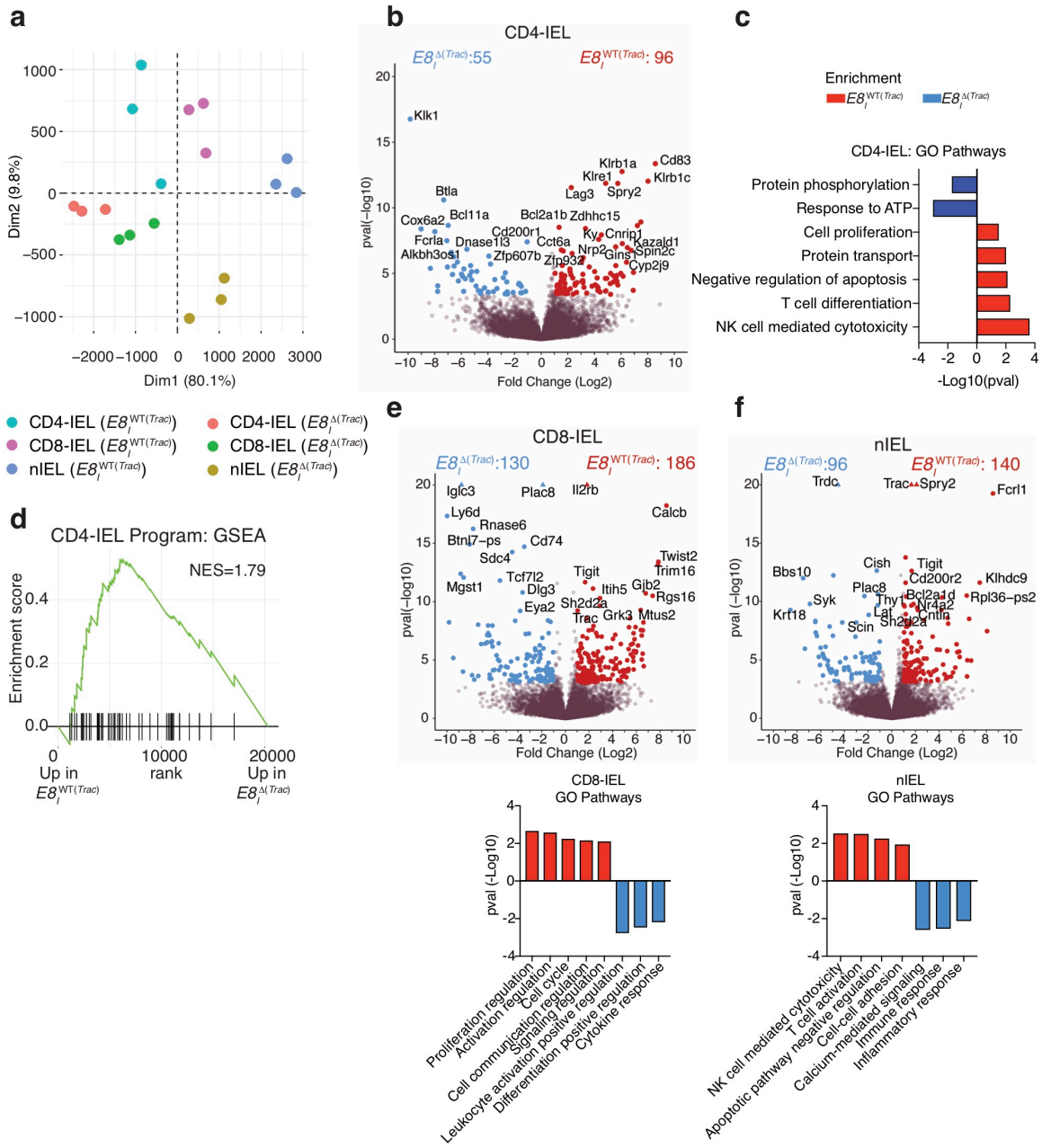



**Figure 4.11 TCR signaling is not essential for CD4-IEL program maintenance.** (**a-f**) Bulk RNA-sequencing was performed on TCR$^{+}$ CD4-IELs, CD8-IELs and nIELs from E8I$^{WT(Trac)}$ and TCR$^{-}$ CD4-IELs from E8I$^{\Delta(Trac)}$ mice. (**a**) Principal component analysis of indicated cell populations (**b-d**) Volcano plot of differentially expressed genes between indicated populations ($p<0.05$, in color) (**b**), selected differentially-enriched gene ontology (GO) pathways between groups (**c**), and gene set enrichment analysis (GSEA) of CD4-IEL program as defined by the top differentially-expressed genes in the CD4-IEL cluster in our single cell RNA-sequencing from Figure 4.4 (**d**). (**e, f**) Volcano plots of differentially expressed genes between indicated populations ($p<0.05$, in color) (top), and selected differentially-enriched gene ontology (GO) pathways between them (bottom).

To further evaluate whether the TCR confers any competitive advantage to IELs, we reconstituted sub-lethally irradiated *Rag1*$^{-/-}$ mice with a 1:1 mix of congenic-marked bone marrow from WT CD45.1 and E8I$^{\Delta(Trac)}$ CD45.2 mice (**Figure 4.12a**). Whereas the ratio of WT to E8I$^{\Delta(Trac)}$ remained approximately 1:1 in non-αβT cells, this ratio was significantly skewed towards WT among CD8-IELs (**Figure 4.12a, b**), indicating that TCR-deficient CD8-IELs were outcompeted by TCR-expressing cells. In contrast, within both nIELs and CD4-IELs, the WT: E8I$^{\Delta(Trac)}$ ratio was much less skewed, and surface-TCRβ$^{-}$ cells were readily detected among these populations (**Figure 4.12a**). Thus, CD4-IELs are much less sensitive than CD8-IELs to TCR loss. Together, our data establish an important role of TCR in IEL differentiation, particularly in the final stages within the epithelium, but also indicate that TCR signaling is mostly dispensable for CD4-IEL maintenance.

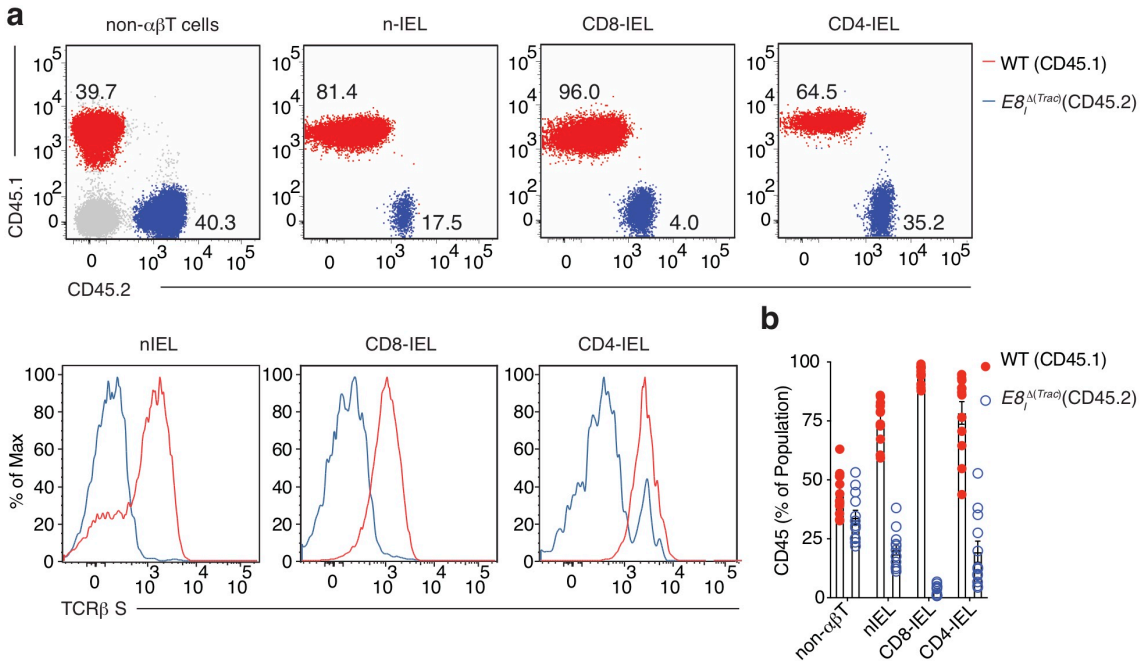



**Figure 4.12 CD4-IELs are less sensitive to TCR loss than CD8-IELs in a competitive setting.** (**a, b**) Flow cytometry analysis of bone marrow chimeras in sub-lethally irradiated *Rag1*$^{-/-}$ hosts reconstituted with 1/1 ratio of WT CD45.1 and E8I$^{\Delta(Trac)}$ CD45.2 cells, analyzed 12-16 weeks after reconstitution. Representative dot plots and histograms (**a**) and frequency (**b**) of WT CD45.1 (red) versus E8I$^{\Delta(Trac)}$ CD45.2 (blue) among non-αβT cells (includes γδT cells and non-T cells), natural IELs (nIEL, CD4$^{-}$ CD8αα$^{+}$ CD8β$^{-}$ TL-Tetramer$^{+}$), CD8-IELs (CD4$^{-}$ CD8αα$^{+}$ CD8β$^{+}$ TL-Tetramer$^{+}$), and CD4-IELs in the IE. Bottom panels, representative histograms of surface TCRβ levels in WT CD45.1 (red line) and E8I$^{\Delta(Trac)}$ CD45.2 (blue line) of indicated cell populations.

## 4.8 Conclusion

Our study revealed a previously unappreciated degree of TCR sharing between largely distinct CD4$^+$ T cell subsets in the periphery, distinct levels of clonal expansion and TCR diversity along differentiating IELs, and a specific requirement for TCR signaling during the early stages of the IEL differentiation process.

Coupling TCR repertoire to single cell transcriptional analysis revealed that expanded clones followed the trajectory analysis while displaying intra-clonal plasticity. The homogenous group of CD4-IELs included expanded clones while the heterogenous subsets contained less expanded clones. We used *in vivo* fate-mapping, reporter, and gene ablation models to uncover the requirements of the TCR during T cell adaptation to the epithelium. We found that Nur77, which marks TCR stimulation strength, is higher in migrating CD4$^+$ T cells but is progressively downmodulated as they differentiate into pre-IELs and remains low in the majority of CD4-IELs. Indeed, we demonstrated that the expression of the TCR complex is crucial for CD4-IEL development, but not maintenance. Intestinal epithelial cell specific ablation of MHC-II precluded the accumulation of CD4-IELs, suggesting that antigen presentation by IECs is important for their development. Together, our data suggests that local recognition of antigens is an important signal for the adaptation, but not maintenance of CD4$^+$ T cells in the epithelium.

# Chapter 5: Discussion

## 5.1 Overview of the thesis

The chronic exposure to a large amount of diverse and potentially stimulatory antigens derived from the diet and microbiota is thought to result in a "physiologically inflamed" state of the intestine. To maintain homeostasis, the intestinal immune system must be able to both mount an effective yet controlled immune response against potential pathogens while tolerating innocuous or beneficial substances. Failure to do so can lead to adverse health effects, such as inflammatory bowel disease, celiac disease, colorectal cancer, infections and food allergies. Achieving harmonized tolerance responses towards harmless antigens and controlled effector responses directed at pathogenic stimuli requires the collective coordination and orchestration of various components within the intestine, including immune cells, epithelial cells, and a multitude of environmental factors, such as cytokines and microbiota- or food-derived metabolites. To this end, many processes are in place for the adaptation of lymphocytes to distinct environments within the intestinal tissues. However, these mechanisms of adaptation as well as their direct functional consequences are not completely understood.

The work presented in this thesis provides mechanistic insight into how $CD4^+$T cells adapt to tissue environments within the mammalian intestine - which serves as the largest tissue reservoir of lymphocytes. Specifically, we addressed the sequence of events leading to intestinal tissue-specific signatures within different niches and inflammatory contexts. To this end, we characterized the levels of heterogeneity and plasticity of $CD4^+$ T cell populations within the intestinal epithelium (IE), underlying lamina propria (LP), and draining mesenteric lymph nodes (mLN). Although closely related, these tissues harbor distinct compositions and signatures, in part due to differential proximity to the intestinal lumen. We found that $CD4^+$ regulatory T cells (Tregs), primarily those induced in the periphery, de-stabilize and convert to $CD8\alpha\alpha$-expressing CD4-intraepithelial lymphocytes (CD4-IELs) upon migrating to the epithelium. To elucidate the precise mechanism by which this occurs, we used genetic fate-mapping and gene-ablation mouse models and in part attributed this plasticity to the loss of the CD4-lineage transcription factor ThPOK. Coupling RNA- and ATAC- sequencing with these *in vivo* models uncovered a temporal program, whereby cessation of the Treg profile is followed by the subsequent increase in accessibility and expression of the CD4-IEL program at the epithelium. Interestingly, this unique series of events is partially reversed by the premature ablation of *Thpok* in Tregs, which leads to the acquisition of the IEL program before the complete shutdown of the Treg program.

In addition to revealing the stepwise transcriptional and chromatin accessibility changes associated with $CD4^+$T cell plasticity, we also uncovered the role of the T

cell receptor (TCR) in this adaptation to the intestinal epithelium. The potential extent of TCR diversity and the role of TCR signal strength is well characterized in thymocyte development. However, much less is known about how TCR signaling and repertoire impact tissue-induced plasticity. We found that while terminally differentiated CD4-IELs are clonally restricted, Tregs and intermediate stages are more diverse. Furthermore, we used *in vivo* mouse models to ablate the TCR complexes on activated CD4$^+$ T cells, revealing that TCR expression is required for differentiation into CD4-IELs at the intestinal epithelium. CD4-IEL differentiation is also precluded upon the deletion of MHC-II on intestinal epithelial cells (IECs). However, TCR ablation on CD4-IELs results in little impact on their maintenance. Taken together, these data suggest that local MHC-II expression, and likely antigen presentation, are required for T cell adaptation, but not maintenance of the IEL program, at the intestinal epithelium.

## 5.2 Regulatory T cell adaptation to the intestinal epithelium

The recent availability of ATAC- and single cell RNA (scRNA) - sequencing data has been instrumental in providing insight not only into the heterogeneity of T cell populations, but also diversity of the same populations between distinct tissues. Such studies highlight the importance of inter-tissue adaptation for local homeostasis[70, 74, 161, 176, 177, 178, 179, 180, 181, 182]. Coupling these two types of technologies revealed that Tregs from visceral adipose, skeletal muscle, skin, and the colonic lamina propria differed in both their chromatin landscapes and transcriptional profiles[177, 183]. Likewise, our unbiased scRNA-seq profiling of CD4$^+$T cells in three gut-associated tissues (mLN, LP, and IE) revealed that the presumably stable T cell subset of Tregs drastically differs in its transcriptional profile between location. Strikingly, the transcription profiles of these Tregs more closely resembled other T cell subtypes in the same tissues than Tregs in neighboring tissues, suggesting that tissue-dependent imprinting often overrides cell subtype specification. This finding is corroborated by the shared core set of tissue- but not cell type-dependent genes between regulatory and memory CD4$^+$T cells within the same tissue (skin or colon)[161].

In contrast to literature noting Treg adaptation without the loss of its core signature and identity[61, 161, 177], we describe the destabilization of Tregs with the concomitant loss of the Treg-lineage defining factor, Foxp3, upon adaptation to the intestinal epithelium. This finding is also at odds with extensive studies showing Treg stability, even during inflammatory and lymphopenia conditions[131]. Several mechanisms are in place to ensure Foxp3 expression and stability, the most well described of which are features of the conserved noncoding sequences (CNS) 1-3[184]. The de-methylated, and thus accessible, state of the CpG-rich CNS2 allows for the binding of multiple transcription factors (such as STAT5, Cbfb-Runx1 and Foxp3)[184, 185, 186], leading to Foxp3 maintenance. Despite certain inflammatory conditions (such as IL-6 and IL-4) which respectively recruit STAT3 and STAT6

binding on CNS2, resulting in its methylation and silencing[187, 188, 189], other histone modifications can nevertheless reinforce Foxp3 transcription[190]. A similar interplay of USP7 mediated de-ubiquitination with inflammation- or hypoxia-induced ubiquitination by Stub1 functions to counteract the Foxp3 pro-degradation aspects of the latter, further supporting Treg identity [191, 192, 193, 194]. Other factors indirectly maintaining the Treg program, including Blimp1, Foxp1, Pak2, and *miR-10a*, function by inhibiting the expression of alternative fates[195, 196, 197, 198, 199, 200, 201] in established Tregs. Together, these mechanisms suggest that Foxp3 stability is an active process. On the other hand, Foxp3 deletion does not lead to major changes in enhancer landscapes of Tregs[202], suggesting either Foxp3-independent stability or that the independent formation of these elements are exploited upon Foxp3 expression, a more passive process. As similar factors, such as TGFβ and retinoic acid (RA) are required for both iTreg and CD4-IEL differentiation, it is possible that they may be directed towards the acquisition of the IEL fate instead of Treg stabilization in combination with other features in the epithelium[127]. Indeed, IFNγ and IL-27, both of which are abundant in the epithelium, can suppress TGFβ and RA mediated Foxp3 induction[127]. Furthermore, TCR stimulation by microbiota-derived peptide antigen may drive Treg plasticity[59, 60]. Indeed, we found that iTregs, but not nTregs, preferentially destabilize in the IE, in line with the notion that the latter are self-reactive while iTregs recognize foreign derived peptide-antigen. The reliance on different CNS regions[184] between nTregs and iTregs, and the more methylated and thus less stabilized, state of the Foxp3 locus in iTregs[131, 203, 204], suggests that a combination of cell-intrinsic and extrinsic factors drive its destabilization within the epithelium.

Natural and induced Tregs may also have distinct functions in the intestine. Tolerance towards diet and microbiota is thought to be mediated by induced, but not natural Tregs[51, 205, 206, 207], while central nervous system protection in the model of experimental autoimmune encephalomyelitis (EAE) is mediated by nTregs, which are self-reactive[206]. However, as previously demonstrated, we observed that both iTregs and nTregs were able to prevent disease in our transfer model of colitis. Our monoclonal *in vivo* experiments suggest that in response to oral antigen, CD4-IELs may play a tolerogenic regulatory role, similar to Tregs, corroborated by the IL-10 and TGFβ mediated function of CD4-IELs in the inhibition of $T_H1$-induced intestinal inflammation[208]. Although these studies did not distinguish conventional $CD4^+$ T cell (Tconv) from Treg-derived CD4-IELs, our sequencing data reveals their similar transcriptional profiles, suggestive of shared function. One hypothesis is that in order to maintain Treg function at the chronically activated epithelial environment, the acquisition of the TCR-dampening CD8αα homodimer prevents exhaustion and enables their survival. However, similar to the vast array of mechanisms and types of function displayed by Tregs in different tissues[196, 209, 210, 211, 212] CD4-IELs may also acquire distinct functions from their precursors. Examining the differences between the epithelium and adjacent lamina propria and draining mLN, in the context of known regulation mechanisms, may

provide insight into both CD4-IEL function and requirement for Treg plasticity. For instance, Tregs can mitigate inflammation by inhibiting antigen presenting cells (APCs) which would otherwise activate naïve T cells[53,54]. While this may be effective in the LP, which contains APCs, and the mLN, where APCs and naïve T cells interact, it may prove useless in the epithelium, which is primarily composed of activated T cells. Furthermore, the anti-inflammatory cytokines produced by Tregs in other tissues may be diluted or ineffective in the highly stimulated epithelium. Indeed, CD4-IEL differentiation is accompanied by a cohort of genes normally expressed by $CD8^+$ T cells, indicating that they may also have cytotoxic function. It would therefore be interesting to assess if CD4-IELs can kill pathogen-infected cells. Given their proximity to IECs, which can be in close contact with potential luminal pathogens, CD4-IELs may have a role in barrier protection, similar to TCR$\gamma\delta$-IELs[76]. Furthermore, Tregs can assume non-canonical functions upon adaptation to various tissues, such as supporting hair follicle and intestinal stem cell niches and contributing to regeneration and repair of central nervous system, lung, skin and muscle[196, 209, 210, 211, 212]. Taken together, it is conceivable that in the epithelium, environmental cues drive iTreg destabilization, and subsequent development into CD4-IEL leads to functional adaptation.

## 5.3 Mechanisms of T cell adaptation to intestinal tissues

In addition to Treg heterogeneity and plasticity, our scRNA-seq data showed broad distinctions between $CD4^+$T cells populations in the closely related IE, LP, and mLN environments. We speculate that the relative inter-tissue heterogeneity is likely a result of the environment, and not just of the subtypes themselves. Corroborating this notion, although HSV-specific memory T cells segregate between skin dermis and epidermis based on their broad lineage, CD4 and CD8, respectively, the differences in key features such as motility between the two tissues is not directly dependent on cell subtype[74]. In general, cells within the epithelium express a cytotoxic profile, paralleling the more homogenous population of $CD8^+$T cells within the IE in comparison with other locations[79, 213, 214]. However, the lamina propria consist of multiple distinct subtypes of $CD4^+$ T cells displaying a TCR-stimulated signature, in line with the highly stimulated intestinal environment. Relative to the intestinal tissues, cells in the draining mLNs display a more naïve transcriptional profile.

It is unclear if $CD4^+$ T cells in the mLN are epigenetically naïve as well or poised for certain programs upon the right factors. Differences between the two states have been noted in Tregs[177] as well as in LCMV specific destabilized tissue resident memory (ex-$T_{RM}$) cells, which despite their transcriptionally de-differentiated state, remained epigenetically poised for migration and re-differentiation to their original tissue and state upon encounter of their cognate antigen and proper cytokine milieu[179]. Thus, it would be interesting to couple the transcriptional and epigenetic levels of the naïve (or not) states of mLN T cells.

One study found an additive combination of accessible chromatin regions from different non lymphoid tissue Tregs in splenic Tregs, which included super-enhancers and particular histone marks[177]. This study suggests that splenic Tregs are simultaneously poised to adopt multiple tissue signatures, which in turn rely on distinct transcription factor regulation. A similar possibility of epigenetic "poise" may exist in mLN lymphocytes as well, and it would be interesting to note if epigenetic distinctions exist between cells destined for the LP versus the IE.

Our trajectory analysis of T cell transcriptional profiles suggests that cells which end up in the epithelium, likely by migrating through the lamina propria upon mLN emigration, do not first acquire a LP signature. In line with a core shared signature of memory and regulatory T cells undergoing tissue-draining lymph node to respective tissue trajectory, irrespective of trajectory[161], our data imply that the acquisition of environment-dependent profiles occurs after decision of tissue location. The differentiation of long-lived memory $CD8^+$ T cells in skin or gut[70] upon acquiring tissue homing molecules in inguinal or mesenteric LNs, respectively, corroborates this finding and also shows that migration imprinting occurs in the draining lymph nodes[70, 180, 213]. It is conceivable that differences in peptide antigen presentation to naïve T cells, both in terms of TCR-peptide antigen specificity and APC-secreted factors, drive this decision. This hypothesis could be in part addressed by coupling APCs bearing specific pre-loaded peptide antigens[215] or TCR transgenic models with the "*Labeling Immune Partnership by SorTagging Intercellular Contacts*" (LIPSTIC) mouse model[216].

After priming in the lymph nodes and subsequent tissue migration, cells must adapt to local survival cues, such as cytokines[217, 218], as exemplified by the differential dependence of $T_{RM}$ cells on IL-15 for maintenance between skin dermis and draining lymph nodes[68, 219]. The predominant expression of *Ccl5* in $CD4^+$T cells within the epithelium may contribute to, or result from, a differential cytokine milieu or receptor Ccr5 expression in the IE. T cells in *Ccl5*-deficient mice have impaired capacities for proliferation and cytokine production upon cognate peptide-antigen stimulation[220], suggesting that differential *Ccl5* expression may modulate the local environment. Furthermore, only cells in the IE are polarized with a gradient of TGFβ- dependent *Itgae* (encoding CD103) expression[68, 221, 222, 223 66, 67, 79, 224]. This is in contrast to memory $CD8^+$ T cells which express similar levels of CD103 within the LP and IE, although it is only required for their retention in the latter[79]. As TGFβ has multiple origins, its expression levels may vary in different niches, thereby differentially influencing local lymphocytes. Taken together, analysis of the local cytokine milieu may provide insight into the functional mechanisms of tissue adaptation.

A key predictor of the ability of a T cell to adapt to different environments may be the extent of its epigenetic accessibility or poise. For example, $CD8^+$ T cells which adopt a $T_C17$ profile upon skin *Staphylococcus epidermidis* colonization, have a

poised type 2 profile, the realization of which is released upon the addition of pro-inflammatory cytokine IL-18[225]. Although we did not find epithelial Tregs to display an accessible IEL profile prior to their destabilization, it is possible that a level of the IEL program was masked when compared to other subsets in the already polarized IE. Perhaps an unappreciated level of IEL poise would be revealed in ATAC-seq and ChIP-seq assays, particularly for histone modifications, comparing peripheral and epithelial Tregs with CD4-IELs. For instance, ChIP-seq of both repressive and permissive histone modifications, H3K27me3 and H3K4me respectively, revealed the poised state of *Tbx21* (encoding T-bet) in cell types that do not canonically express it ($T_H2$ and $T_H17$ cells)[226] unless stimulated to advance to expression by inflammatory signals[227].

Regardless of potential epigenetic predisposition for plasticity, our identification of a pre-IEL signature, marked by the loss of key features of conventional CD4 and/or Treg programs, suggests that during tissue adaptation, a cell must first shut down transcriptional programs in place (or decouple transcription factors bound to its DNA targets) that may prevent tissue imprinting, a suggestion corroborated by previous studies describing roles of Runx3 to the CD8-lineage or tissue-resident memory T cells ($T_{RM}$)[166, 167]. While a similar level of transcriptional changes occurs prior and post the pre-IEL stage, the majority of chromatin changes occurred in the initial step of pre-IEL formation, suggesting that *in situ* epigenetic changes likely set the stage for IEL differentiation. Performing histone modification ChIP analysis at the different stages of IEL-differentiation would thus help us understand the mechanisms involved. A recent study also describes a stepwise process taken by Tregs, in which peripheral precursors gradually acquire chromatin accessibility and reprograming towards nonlymphoid-tissue Tregs. The authors identified the transcription factor BATF as a main driver of the molecular tissue programing in the precursors[228]. These observations parallel our identification of the role of ThPOK down-modulation in epithelial imprinting on iTregs, driving their plasticity towards the IEL fate at the epithelium. It remains to be determined, like the BATF differentiation described, whether lamina propria or epithelium Tregs share a common precursor in the mLN prone to tissue differentiation in the gut.

### ***5.3.1 Mechanisms of T cell adaptation: The role of ThPOK***

Previous studies demonstrated that CD4-IEL development from conventional $CD4^+$ T cells is in part dependent on the reciprocal down- and up- regulation of *Thpok* and *Runx3*, respectively[66, 67]. We conclude that ThPOK, but not Runx3 plays a similar role in the development of exTreg-IELs. Although Runx3 did not reveal a stabilizing role for Foxp3 as described[150, 151, 229], it is possible that it may have set the stage for IEL development in Tregs prior to its ablation in our *in vivo* systems. In our polyclonal *in vivo* experiments, both Runx3 and ThPOK seemed dispensable to Treg function, while Foxp3 expression was only modestly affected upon *Thpok* ablation. While natural *Thpok* down-modulation led to the coordinated

stepwise destabilization of the Treg program followed by the acquisition of the IEL program, *Thpok* ablation at the Treg stage prematurely led to this acquisition. We therefore concluded that, while *Thpok* had little effect on Treg function, its timely down-modulation was important in the correct progression towards the CD4-IEL program. Similarly, while *Thpok* deletion in activated $CD4^+$ T cells does not affect viral clearance or the expansion of antigen-specific cells following LCMV infection, it plays a role in the regulation of $CD4^+$ T cell differentiation, particularly towards $T_H1$ cells, which share some expression profile with interferon-producing CD4-IELs[164]. One caveat of our broad conclusions regarding the role of ThPOK and Runx3 in Treg programming is that our sequencing analyses revealed that *Foxp3* is one of the last canonical Treg genes to be down-modulated during the IEL transition, which reinforces previous studies suggesting that Foxp3 is a late-acting transcription factor which binds to pre-established accessible enhancer regions[202]. Therefore, it is plausible that the Treg program may be modulated independently of Foxp3; in the case of tissue adaptation events addressed here, this may preclude the use of Foxp3 as a Treg marker without considering the entire Treg profile. Additionally, it is possible that ThPOK acts in conjunction with other transcription factors, including Runx3[150, 151, 229], to regulate Treg function. By inactivating both ThPOK and its homolog LRF in differentiated Tregs, a previous study concluded that these transcription factors redundantly support Foxp3 function[230].

Although our work elucidated the events following ThPOK downmodulation in Tregs, the exact tissue-dependent triggers leading to ThPOK downmodulation is not completely known[66, 67, 127]. A microbiota induced aryl-hydrocarbon receptor requirement has been proposed to drive ThPOK downmodulation[134], and it would be interesting to understand this and additional mechanisms in greater detail. Several factors, such as retinoic acid, IFNγ, IL-27, and IL-15 have been implicated in the development of IELs, although their mechanisms are incompletely understood[66, 67, 127]. Network analysis of downstream genes of these factors in the progression of Tregs or Tconv to pre-IEL development may elucidate the environmental factors driving *Thpok* down-modulation. The physiological hypoxic state of the IE (with half the oxygen in the underlying LP[231, 232, 233] ) may also play a role in CD4-IEL differentiation and is thus an avenue worth investigating. For instance, hypoxia inducible factor 1-deficient IECs led to a decrease of *Il7* and *Il15* mRNA, resulting in changes in IEL frequencies as well as microbiota composition[234], displaying that multiple factors may be at play.

## 5.4 T cell receptor stimulation in $CD4^+$T cell adaptation at the intestinal epithelium

In addition to the down-modulation of ThPOK as Tregs differentiate to pre-IELs, we find that Nur77, which is expressed proportionally to the strength of the TCR stimulation[235, 236], is also progressively down-modulated. This is similar to the high

level of Nur77 displayed by agonist-selected T cells such as iNKT cells during thymic selection, which subsequently decays upon migration of these cells to the spleen or liver [237]. It would be particularly interesting to analyze the relative dynamics of ThPOK and Nur77 downregulation to understand if TCR stimulation plays a role in driving ThPOK down-modulation. On the other hand, despite being "chronically activated" as defined by the expression of activation markers such as CD69 and CD44 [66], CD4-IELs do not show signs of strong TCR activation. While the majority of CD4-IELs express low levels of Nur77, a small fraction of them maintained Nur77 expression, suggesting that TCR re-engagement can be modulated within the epithelium.

Indeed, availability of antigens presented by intestinal epithelial cells could function in the late-stage of IEL differentiation, possibility supported by the decreased CD4-IEL population upon MHC-II targeting on IECs. The modulation of MHC-II expression by IECs has been linked to the capacity of microbes to attach to the epithelium, and to IFN$\gamma$ production in both human and murine models [238, 239, 240]. MHC-II expression by IECs has been recently associated with a variety of physiological, such as regulation of the stem cell niche, and pathological functions, such as CD4$^+$ T cell-mediated inflammation during graft-versus-host disease, in part through interacting with gut resident T cells that provide cytokines [105, 106, 174]. Our data suggests a model in which antigen presentation by IECs is instrumental for the differentiation of CD4-IELs, which may further enhance MHC-II expression by IECs via IFN$\gamma$ production. Whether this is exclusively dependent on antigen presentation by IECs needs further demonstration. The transfer of naïve CD4$^+$ TCR transnuclear (TN) T cells[241], specific for *Parabacteroides Goldsteinii*-derived $\beta$-hexosaminidase, into mice lacking MHC-II expression by intestinal epithelial cells may provide additional insight. Upon *in vivo* cognate peptide administration, these cells preferentially develop into CD4-IELs at the epithelium in SPF conditions. If our hypothesis of the role of antigen presentation by IECs is correct, then the lack of MHC-II expression on IECs in this system would preclude CD4-IEL development. Of note, preliminary experiments involving cognate peptide administration to germ-free mice with these transferred naïve CD4$^+$ TCR transgenic T cells does not lead to CD4-IEL development, suggesting that other antigen-independent microbiota factors are required. Additionally, the TCR complex itself, and presumably antigenic stimulation, is important for CD4-IEL differentiation as demonstrated by TCR ablation on OX40-expressing cells as well as on Tregs, which then precluded CD4-IEL differentiation.

However, our results show that TCR ablation from CD4-IELs does not impair their persistence in the epithelium nor the production of IFN$\gamma$ and granzyme B. This is consistent with the decreased antigen sensitivity and increased threshold for TCR activation in cells expressing CD8$\alpha\alpha$ homodimers [85]. Furthermore, our RNA-seq analysis suggests that despite the down-modulation of genes downstream of TCR activation in TCR-deficient CD4-IELs, the maintenance of an IEL program may not

depend on continuous TCR signaling. This is in contrast to the requirement of TCR expression on mature Tregs for their suppressive function and maintenance of an effector Treg program [170]. The extent to which CD4-IELs depend on TCR signaling for specific functions and whether it is needed in discrete modules during different stages of differentiation from Tregs to CD4-IELs remains to be determined. In this line, it would be interesting to see if the TCR is also dispensable for CD4-IEL maintenance upon perturbation, such as in the context of an immunological challenge like during an intestinal infection.

## 5.5 T cell receptor repertoire in CD4$^+$T cells in the intestinal epithelium

Although surface expression of T cell receptors did not significantly impact CD4-IEL maintenance, they displayed a restricted TCR repertoire. Likewise, early analyses of TCR diversity in nIEL subsets revealed a restriction in TCR repertoire, referred to as "oligoclonal repertoire" [64, 242, 243, 244]. More recently, work using a TCR$^{mini}$ mouse model harboring restricted TCR$\alpha$ and TCR$\beta$ repertoires showed a substantial TCR overlap between CD4-IELs and Tregs [245]. In addition to the agonist selection of thymic IEL precursors [246, 247], studies that generated transgenic mouse strains carrying existing IEL $\alpha\beta$TCRs strongly suggested that TCR specificity may be sufficient to drive IEL fate [248, 249]. It remained unclear, however, how specific TCRs correlate with IEL differentiation. Our scRNA-seq and trajectory analyses coupled to TCR repertoire allowed us to unbiasedly define the relationship between TCR diversity and CD4$^+$ T cell plasticity during migration and differentiation towards IELs. Expanded clones followed pseudotime trajectory analysis and displayed intra-clonal plasticity: less expanded clones spread among heterogenous subsets, while highly expanded clones were found in the homogenous cluster of CD4-IELs. A possible explanation for this finding is that the less expanded clones lacked additional signals, such as TCR ligands or environmental components, required for full differentiation into CD4-IELs [66, 67, 127, 134, 250]. The clonal analyses presented here, including the TCR sharing of CD4-IELs derived from Treg or from Tconv, corroborate previous studies suggesting a lineage relationship between Tregs and IELs [1, 241]. Additionally, their clonal distribution suggest that IEL differentiation may favor particular TCR specificities, perhaps in a process analogous to peripheral Treg differentiation, where TCR recognition in a context-dependent manner leads to Foxp3 expression [251, 252].

Due to their limited TCR repertoire, low TCR-stimulation, and phenotypic resemblance to tissue resident memory T cells, CD4-IELs may be specific for a restricted set of antigens and display $T_{RM}$-like function upon their encounter. $T_{RM}$ cells reside tissues even in the absence of their cognate antigen-peptide but can exert effector-like function in their presence. To test this possibility, one could use an ovalbumin (OVA)-specific TCR transgenic mouse model (OT-II) to generate OVA-specific CD4-IELs in the presence of a pathogen engineered to expresses OVA. Cytokine production and pathogen clearance will then be assessed after

initial pathogen clearance as well as following subsequent re-infection. Furthermore, analysis of the CD4-IEL TCR repertoire and that of their precursors, in various controlled bacterial and diet derived antigen conditions, such as colonization with the defined oligo-mouse microbiota (OMM[12])[253] bacterial consortium or single-antigen diets, may provide functional clues to the type of stimuli these cells may respond to. An expanded cohort of specific TCR-cognate peptide-antigen pairs, similar to OTII-OVA peptide and TN-β-hexosaminidase, which can lead to CD4-IEL generation in different contexts[241], can provide the tools to decouple environmental from TCR-specificity requirements for pre-IEL and CD4-IEL development.

## 5.6 Concluding remarks and outstanding questions

Overall, this work presents how T lymphocytes adapt to the unique environments of the lamina propria and intestinal epithelium. The $CD4^+$ T cell populations within the LP displayed a TCR-stimulated signature but remained adherent to canonical subtypes. However, those in the epithelium were polarized with a gradient of a cytotoxic IEL profile, an imprinting that overrode thymically established lineage specificity to some extent. Furthermore, as cells within the epithelium are not marked by a lamina propria-derived signature, our data suggests that tissue imprinting occurs after the decision of tissue location. We showed that in adapting to the epithelium, $CD4^+$ T cells, including induced regulatory T cells, shut down lineage-defining transcriptional programs in place before acquiring the CD4-IEL profile. TCR stimulation and accompanying local recognition of possibly a limited set of antigens is also essential in this T cell adaptation. We identified an intermediate “pre-IEL” stage, marked by the loss of CD4-lineage defining transcription factor, ThPOK and the gradual upregulation of cytotoxic markers characteristic of $CD8^+$ T cells. Although *Thpok* abrogation in Tregs in the mesenteric lymph nodes anticipated the expression of some IEL-related genes, the combination of the epithelium environment with the timely downmodulation of *Thpok* at the pre-IEL stage were required for proper T cell adaptation. Taken together, our work shows that the combination of genetic, environmental, and TCR triggers is crucial in driving T cell plasticity and adaptation to the intestinal epithelium.

Although our work, and that of others, provide examples of transcriptional and epigenetic changes within the progression of T cell adaptation, many mechanisms remain unknown, such as driving forces behind certain modulations described in our work. Furthermore, the developmental consequences of only a limited set of TCR and cognate peptide-antigens pairs have not been established. Given the vast antigenic diversity within the intestine, further characterization of specificities and consequences of the TCR repertoire, which becomes more limited in the progression of adaptation towards the IEL fate, remains to be elucidated. It would be interesting to de-couple the environmental and TCR-stimulation effects as T

cells adapt to the epithelium. Finally, the function of CD4-IELs remains poorly characterized, particularly in humans. Although they have also been suggested to contribute to pro-inflammatory damage in celiac disease[90], reduced frequencies of CD4-IELs have been noted in human patients with chronic intestinal inflammation, suggesting an anti-inflammatory role[93, 94]. Similarly, our studies point to a regulatory role of CD4-IELs in the context of diet-derived antigen in mice. Although much is left to be discovered, out work sheds light on the mechanisms of lymphocyte adaptation to intestinal tissues, a process crucial for the maintenance of homeostasis.

# Materials and Methods

## Animals

Animal care and experimentation were consistent with the NIH guidelines and were approved by the Institutional Animal Care and Use Committee at the Rockefeller University. C57BL/6J (000664), actin-mRFP1 (005584), *Nr4a1*$^{EGFP/Cre}$ (Nur77$^{GFP}$, 016617), OT-II TCR transgenic (004194), *Rag1*$^{-/-}$ (002216), TCR$\gamma\delta^{GFP}$ (016941), *Zbtb7b*$^{eGFP}$(027663), *Zbtb7b*$^{fl/fl}$ [20](009369), *Cd4*$^{Cre}$ (017336), *Cd4*$^{Cre\text{-}ERT2}$(022356), OX40$^{IRES\text{-}Cre}$ (012839), *Foxp3*$^{IRES\text{-}mRFP}$(008374), *Rosa26*$^{lsltdTomato}$(007914), *Foxp3*$^{eGFP\text{-}Cre\text{-}ERT2}$(016961) and CD45.1(B6.SJL *Ptprc*$^{a}$, 002014) mice were purchased from the Jackson Laboratories and maintained in our facilities. BALB/c, TBmc and BALB/c TBmc Foxp3$^{sf}$ mice were generated and provided by J. Lafaille (NYU). *E8I*$^{Cre}$ (Jax 008766) were generated by I. Taniuchi (Riken) and kindly provided by H. Cheroutre (LJAI) and were used to target cells with an active *Cd8a* enhancer I, including mature CD8$^{+}$ T cells and also CD8$\alpha\alpha^{+}$ CD4-IELs. *Cd4*$^{Cre\text{-}ER}$ mice were previously described [138], and kindly provided by S. Kuralov (NYU) and were used to target mature CD4$^{+}$ T cells and in recent thymic emigrants upon tamoxifen administration. *Runx3*$^{fl/fl}$ mice were kindly provided by T. Egawa (Wash. U.). *Tbx21*$^{fl/fl}$ were kindly provided by S. Reiner (Columbia U.). *Foxp3*$^{Cre\text{-}YFP}$ and *Rosa26*$^{lsl\text{-}DsRed}$ mice were kindly provided by J. Lafaille (NYU). *Trac*$^{f/f}$ mice were kindly provided by A. Rudensky (MSKCC). *Villin*$^{CreERT2}$ mice were generated by [254] and kindly provided by D. Artis (Cornell). *Foxp3*$^{IRES\text{-}GFP}$ mice were provided by V. Kuchroo (Harvard) and *H2-Ab1*$^{f/f}$ mice were provided by M. Nussenzweig (Jax 013181). Several of these lines were interbred in our facilities to obtain the final strains described in the text. Genotyping was performed according to the protocols established for the respective strains by Jackson Laboratories. Mice were maintained at the Rockefeller University animal facilities under specific pathogen-free (SPF) or germ-free (GF) conditions. Germ-free mice were obtained from Sarkis Mazmanian and bred and maintained in germ-free isolators in our facilities. GF status was confirmed by plating feces as well as by qPCR analysis (16S rRNA). Mice were used at 8-10 weeks of age for most experiments, unless otherwise stated.

## Antibodies and flow cytometry analysis.

Fluorescent dye–conjugated antibodies were purchased from BD Biosciences, Biolegend or Ebioscience (Thermofisher). The following clones were used: anti-CD45.1, A20; anti-CD45.2, anti-Foxp3, FJK-16s; anti-CD4, RM4-5; anti-CD45.2, 104; anti-CD8$\alpha$, 53-6.7; anti-CD8$\beta$, YTS 156.7.7; anti-CD44, IM7; anti-CD45, 30-F11; anti-CD62L, MEL-14; G8.8; anti-TCR$\beta$, H57-597; anti-TCR$\gamma\delta$, eBioG23; anti-CD25, PC61.5; anti-CD19, eBio1D3; anti-IFN-$\gamma$, XMG1.2; 104; anti–I-A/I-E, M5/114.15.2; anti-Granzyme B, NGZB; anti-CD103, 2E7; anti-EpCAM, G8.8;. Live/dead fixable dye Aqua and Edu (ThermoFisher Scientific) was used according to manufacturer's instructions. Intracellular staining of Foxp3 was conducted using Foxp3 Mouse Regulatory T Cell Staining Kit (eBioscience, USA). For flow

cytometric analysis of cytokine-secreting cells in chapter 2, cells were incubated at 37°C in the presence of 100ng/ml phorbol 12-myristate 13-acetate (PMA, Sigma), 500ng/ml Ionomycin (Sigma) and 10μg/ml Brefeldin A (Sigma) for 3 hours prior to staining. Brefeldin A was added after 30 minutes of incubation. For analysis of cytokine-secretion in chapter 4, IELs were plated in 48-well plates and incubated at 37°C with 100ng/mL PMA and 200ng/mL ionomycin (Sigma) for 4 hours. Monesin (2μM, Sigma) was added 1h after PMA and ionomycin. Cell populations were first stained with antibodies against the indicated cell surface markers, followed by permeabilization in Fix/Perm buffer, and intracellular staining in Perm/Wash buffer (BD Pharmingen, USA) as per kit instructions. TL-Tetramer was first kindly provided by Hilde Cheroutre (LJAI, San Diego) and then was obtained from NIH tetramer facility. Anti-ThPOK ChIP antibody was kindly provided by T. Egawa (Wash. U.). Flow cytometry data was acquired on a LSR-II flow cytometer (Becton Dickinson, USA) and analyzed using FlowJo software package (Tri-Star, USA).

**Absolute cell numbers**

AccuCheck Counting Beads (Thermo Fisher, USA) were used for counting of absolute cell numbers.

**Cell sorting**

Lymphocytes were sorted on a FACS Aria II instrument as indicated in the figure legends.

**Tamoxifen treatment**

For *in vitro* treatment, sorted naïve TCRβ$^{+}$CD45RB$^{high}$CD25$^{-}$CD4$^{+}$CD8α$^{-}$ were cultured in standard T cell media in the presence of 1 μl/ml tamoxifen (Sigma, USA) in the time course specified. RNA was extracted to confirm treatment efficiency. For *in vivo* treatment of Foxp3$^{CreERT2}$ and CD4$^{CreERT2}$ animals, mice were intragastrically administered with 5 mg of tamoxifen (Sigma) dissolved in corn oil (Sigma) and 10% ethanol at 50 mg/ml. Tamoxifen was administered to mice starting at 6-7 weeks old, 4 times in the first week and then 2 times every week (3 days apart) every other week for 8-10 weeks, unless otherwise indicated. For *in* vivo treatment of *Villin*$^{CreERT2}$ animals, five doses of Tamoxifen (1mg/dose) was administered intraperitonially at 10mg/mL within one week, with 2 extra 1mg boosts 3 days apart 2 weeks before analysis, when analysis was more than 4 weeks after initial dosing, as indicated in figure legends. To label cells for transfer, tamoxifen was administered twice, 2 days apart, 2 days prior to sorting. 2 days after transfer to *Rag1*$^{-/-}$ mice, animals were administered 100uL of 10mg/mL Tamoxifen intraperitonially.

**Antibiotic treatment**

0.5 g/L of Vancomycin (Acros, USA), 1.0 g/L of Ampicillin (Fisher, USA), 1.0 g/L of Neomycin (Sigma) and 0.5 g/L of Metronidazole (Sigma) cocktail was administered

in drinking water for 4 weeks. Drinking water for antibiotic treated and control mice contained 10 g/ml of sucralose-based artificial sweetener (Splenda®, McNeil, USA).

**Oral OVA treatment**

OT-II or TBmc animals were fed with regular diet containing 1% chicken-ovalbumin (OVA) for 7 days and analyzed 3 days later. To measure diarrhea in OVA fed OT-II mice, after OVA ingestion, animals were challenged twice a day by gavage with 50 mg of OVA solution (250 mg/ml) until diarrhea was detected. Mice were classified as positive when signs of diarrhea were detected two consecutive times after challenge. Tissue inflammation was histologically scored, according to a standardized scoring scheme with the following combination of parameters: extension of inflammation (0 - none, 1 - lamina propria, 2 - submucosal, 3 - transmural), degree of inflammation (0 - none, 1 - mild, 2 - moderate, 3 - severe), abnormal crypt morphology (0 - 3), neutrophil infiltrate (0 - 4), goblet cell loss (0 - none, 1 - moderate, 2 - severe), mucosal erosion or ulcer (0 or 1), crypt abscess (0 or 1) with a maximum score possible of 17. Scoring was conducted in a double-blinded fashion by at least 4 independent scorers.

**Cell Transfer**

For oral antigen colitis experiments, OT-II naïve T cells ($CD45^+CD4^+CD8\alpha^-CD62L^+CD44^{low}$) cells were sorted from the donor mice as indicated and $10^6$ transferred to the respective host mice. Host animals were gavaged with 50 mg of OVA solution (250 mg/ml) 18 hours after transfer. After transfer host animals were subjected to oral OVA treatment as described above. For polyclonal colitis experiments, induced (neuropilin-$1^-$) and natural (neuropilin-$1^+$) Tregs ($CD45.2^+TCR\beta^+CD4^+CD8^-CD25^+Tomato^+$) from spleen and mLN of donor CD45.2 mice were sorted 2 days after mice were intragastrically given 5mg of tamoxifen prepared as described above for 2 days and co-transferred with naïve $CD4^+$ T cells ($CD45.1^+TCR\beta^+CD4^+CD8^-CD62L^{high}CD44^{low}$) sorted from CD45.1 donor mice and 100,000 $CD45.2^+$ Tregs with 400,000 $CD45.1^+$ naïve T cells were intravenously transferred to *Rag1*$^{-/-}$ hosts. Body weight and fecal lipocalin-2 levels were monitored until terminal analysis.

**Generation of mixed bone marrow chimeras**

Bone marrow cells were harvested from WT CD45.1 or *Trac*$^{f/f}$ E8I$^{Cre}$ CD45.2 donors and depleted of T cell precursors using CD90.2 beads (Miltenyi) according to manufacturer's instructions. An equal mix of $5x10^6$ total cells from WT CD45.1 and *Trac*$^{f/f}$ E8I$^{Cre}$ CD45.2 donors was intravenously injected into sub-lethally irradiated (6 Gy) *Rag1*$^{-/-}$ hosts. Mice were analyzed 12-16 weeks after reconstitution.

**ELISA for Lipocalin-2**

Lipocalin-2 was analyzed by using Lcn-2 ELISA kit (R&D, MN) as described by [255].

**Edu treatment and Detection**

1mg EdU was injected intravenously at 5mg/mL in PBS 16 and 4 hours prior to analysis. Detection was performed using the Click-iT™ Plus EdU Flow Cytometry Assay kit (Thermo Fisher Scientific, C10632), according to manufacturer's instructions.

***In vivo* administration of antibodies**

Antibodies against CD8α (53-6.72) were administered by intraperitoneal (i.p.) injection every other day. All antibodies and isotype control rat anti-mouse IgG were obtained from BioXCell (USA).

**Isolation of intestinal T cells.**

Intraepithelial and lamina propria lymphocytes were isolated as previously described [67, 241]. Briefly, small intestines were harvested and washed in PBS and 1mM dithiothreitol (DTT) followed by 30 mM EDTA. Intraepithelial cells were recovered from the supernatant of DTT and EDTA washes and mononuclear cells were isolated by gradient centrifugation using Percoll. Lamina propria lymphocytes were obtained after collagenase digestion of the tissue. Single-cell suspensions were then stained with fluorescently labeled antibodies for 20min at 4ºC prior to downstream flow cytometry (analysis or sorting) as specified in figure legends.

**Multi-photon intravital imaging and cell tracking.**

For $TCR\gamma\delta_{IEL}$ imaging, $TCR\gamma\delta^{GFP}$ mice were used. For $T_{reg}$ imaging, *Foxp3*$^{CreER}$:*Rosa26*$^{tdTomato}$ mice (i*Foxp3*$^{Tomato}$) were i.p. injected with 4mg of tamoxifen 24 hours prior to intravital imaging. For imaging of ThPOK $(GFP)^{high}$ and ThPOK $(GFP)^{low}$ cells, $CD4^+$ cells from OT-II (RFP *Thpok*$^{GFP}$) mice were sorted and transferred to *Rag1*$^{-/-}$ recipient mice. Recipient animals were fed OVA-containing diet for 7 days before intravital imaging. Nuclei of IEC were stained with Hoechst prior to imaging. Animals were anesthetized with 15μl per gram of 2.5% tribromoethanol (Avertin) and kept on 1% isoflurane oxygen mixture throughout the procedure. A loop of the terminal ileum was exposed through a small ventral incision and the intestine was cut open using a high temperature disposable cautery pen (Bovie Medical, USA) to expose the mucosal side. The tissue was immobilized in thermal play dough (Laird Technologies, USA) using a lamina attached to a plastic ring containing metal pins. Image was acquired on a FV1000MPE Twin upright multiphoton system (Olympus, Japan) for an average of 30 minutes each. Data were analyzed using *Imaris* software (Bitplane, UK). Target cells were identified using co-localization for appropriate channel intensities. Auto-regressive tracking algorithms were performed using default settings. The scoring of intra-epithelial (IE) or lamina propria (LP) compartment localization and LP to and from IE migration dynamics was performed computationally in an unbiased manner. Tracking of cells was done using the same algorithm for all samples included with manual verification of correct lymphocyte identification. IE or LP compartment localization of cells over time was based on algorithms comparing

cell location versus distance to intestinal epithelial cells (IEC, as identified with Hoechst staining) to establish IE and LP localization as well as movement dynamics. IE or LP compartment residency was defined as spending >80% of tracked time in the respective compartment. Migratory cells were defined as non-resident cells moving from IE to LP or vice versa in the tracked time. Cell tracks affected by peristalsis (defined as multiple compartment switches in a short time frame) were excluded automatically (mean % of cells excluded over all samples was $5.4\%^{\pm 2.4\%}$). Distance was normalized for each movie based on the rendered size of Hoechst-stained IEC nuclei; approximately 1.5x the rendered size was set as the maximum distance from the IECs for the IE compartment. Where possible, data from multiple *villi* in the same movie were pooled.

**RNA isolation and Real-Time qPCR**

Isolation of total RNA using Trizol reagent (Ambien, USA) and subsequent generation of cDNA using the iScript kit (BioRad, USA) was carried out according to the manufacturer's instructions. Quantitative PCR was performed using SYBR® Green (Roche, Switzerland) reagent and a QuantStudio® 6 Flex Real-Time PCR system (Applied Biosystems, USA).
Primers used were:
*Rpl32* forward, 5'-ACAATGTCAAGGAGCTGGAG-3',
*Rpl32* reverse, 5'-TTGGGATTGGTGACTCTGATG-3';
*Thpok* forward, 5'-ATGGGATTCCAATCAGGTCA-3',
*Thpok* reverse, 5'-TTCTTCCTACACCCTGTGCC-3';

**ATAC-sequencing**

ATAC-seq was performed as previously described[162, 256] on 5,000-40,000 FACS-purified cells from 2-9 mice. In brief, cells were lysed in lysis buffer for 1 minute and transposed with Tagment DNA Enzyme 1 (Illumina) for 30 minutes. DNA was cleaned up using a MinElute DNA purification Kit (Qiagen), followed by barcoding and library preparation by the Nextera DNA Library preparation kit (Illumina) according to manufacturer's guidelines and sequenced on an Illumina NextSeq500.

**ChIP-sequencing**

Cells were fixed in 1% formaldehyde for 20min, quenched with 0.15M glycine and washed in PBS. $2x10^6$ were then sorted and lysed for 30min at 4ºC. Cells were then sonicated for 22 minutes at 30 seconds on/off using the Bioruptor sonicator (Diagenode) and spun down. 10% was frozen for input, the remaining 90% was incubated with anti-ThPOK antibody bound to goat anti-rabbit M280 magnetic beads (Invitrogen) overnight at 4ºC prior to washing with RIPA buffer and overnight decrosslinking at 65ºC. DNA was then eluted off beads into TE buffer and, along with input, purified using the Zymogen DNA clean & concentrator kit. DNA was sequenced using NextSeq2500. To obtain Tregs, $Foxp3^{RFP}$ reporter mice were injected with IL-2/$\alpha$IL-2 complex as previously described for 3 consecutive days

for Treg expansion[257, 258]. CD4+ T cells were isolated from the spleen and mLNs and sorted as Thy1+TCRβ+CD4+CD8α−RFP+.

**Bulk RNA-sequencing**

Sorted cells (300-800) were lysed in TCL buffer (Qiagen, 1031576) supplemented with 1% β-mercaptoethanol. RNA was isolated using RNAClean XP beads (Agentcourt, A63987), reversibly transcribed, and amplified as described[259]. Uniquely barcoded libraries were prepared using Nextera XT kit (Illumina) following manufacturer's instructions. Sequencing was performed on an Illumina NextSeq550.

**Single-cell TCR sequencing**

Single cells were sorted using a FACS Aria into 96-well plates containing 5μL of lysis buffer (TCL buffer, Qiagen 1031576) supplemented with 1% β-mercaptoethanol) and frozen in -80°C prior to RT-PCR. RNA and RT-PCRs for TCRα and TCRβ were prepared as previously described [260]. PCR products for TCRα and TCRβ were either subjected to Sanger sequencing or multiplexed with barcodes and subjected to MiSeq sequencing [261] using True Seq Nano kit (Illumina). For Miseq data, Fastaq files were de-multiplexed and paired-end sequences assembled using PANDAseq [262] and FASTAX toolkit. Demultiplexed and collapsed reads were assigned to wells according to barcodes. Fasta files from both Sanger and Miseq sequences were aligned and analyzed on IMGT (imgt.org/HighV-QUEST) [263]. Cells with identical TCRβ CDR3 nucleotide sequences were considered as the same clones. Clonality was confirmed by sequencing TCRα of the expanded clones as assessed by TCRβ sequencing.

**Single cell RNA-seq library preparation**

IELs were sorted, counted for viability and immediately subjected to library preparation. The scRNA-seq and scTCR-seq libraries were prepared using the 10x Single Cell Immune Profiling Solution Kit, according to the manufacturer's instructions at the Genomics core of Rockefeller University. The scRNA libraries were sequenced on an Illumina NextSeq550 to a minimum sequencing depth of 50,000 reads per cell using read lengths of 26 bp read 1, 8 bp i7 index, 98 bp read 2. The single-cell TCR libraries were sequenced on an Illumina NextSeq550 to a minimum sequencing depth of 5,000 reads per cell using read lengths of 150 bp read 1, 8 bp i7 index, 150 bp read 2.

**ATAC-seq analysis**

The raw bulk ATAC-seq files were processed using the ENCODE-DCC pipeline (https://github.com/ENCODE-DCC/atac-seq-pipeline) automated through Cromwell (https://github.com/broadinstitute/cromwell). Shortly, adapter sequences were trimmed (Cutadapt) prior to genome mapping (Bowtie2)[264] and filtered subsequently filtered (Samtools)[265]. Next, sequencing enriched regions, associated to chromatin accessible peaks were called (MACS2) and filtered (IDR)

by running Irreproducible Discovery Rate algorithm with a 10% threshold, meaning 10% chance of being an irreproducible peak. Later, filtered peak regions were annotated by using HOMER software tools. Comprehensive motif screening was performed with MEME-ChIP[165] by using promoter originated peaks and, length-normalized reads to a 500 kb to a fixed window of 500bp. Enriched transcriptional factor motifs on ATAC-Seq peaks were detected with Centrimo[266] by providing known Human Runx3 (MA0684) and Foxp3 (MA0850) as published in JASPAR 2020 [267, 268]. As indicative of the presence of Zbtb7b (Thpok) we used the enriched motif detected in our Thpok ChIP-Seq peaks.

**ChIP-seq analysis**

The raw bulk ChIP-seq files were processed using the ENCODE-DCC pipeline (https://github.com/ENCODE-DCC/chip-seq-pipeline2) automated through Cromwell (https://github.com/broadinstitute/cromwell). Shortly, adapter sequences were trimmed (Cutadapt) prior to genome mapping (BWA) and filtered subsequently (Samtools)[265, 269]. Next, sequencing enriched regions, associated to chromatin accessible peaks were called (SPP) and filtered (IDR) by running Irreproducible Discovery Rate algorithm with a 5% threshold, meaning 5% chance of being an irreproducible peak[270, 271]. Later, filtered peak regions were annotated by the HOMER software suite of tools[272]. Comprehensive motif screening was performed with MEME-ChIP by using promoter originated, length-normalized reads to a 500 kb.

**Bulk RNA-seq data analysis**

The raw fastq sequencing files were processed together with the gencode mouse annotation database (v. M21), by running kallisto (v. 0.46.0) to calculate the transcript abundances[273]. Next, the abundance files were submitted to the sleuth (v. 0.30.0) pipeline based on R, for transcript abundance normalization, gene expression analysis and statistics[274, 275]. The quantified transcripts were combined into genes for the downstream analysis. Batch effects were evaluated through principal component analysis and removed with Limma[276]. To capture significantly expressed genes we performed a likelihood ratio test between a null model and our experimental design using an adjusted p-value of 0.05 as threshold. The selected features were clustered using k-means and represented as a heatmap to define group gene signatures. Pairwise comparisons between groups were performed using the Wald-test and significant genes were considered for downstream analysis by having an adjusted p-value smaller than 0.05 and a log2 fold-change of 1. Pre-ranked gene set enrichment analysis (GSEA) were performed with the fgsea package[277] by comparing gene lists sorted by their log2 fold-change with gene signatures from the clusters 21 (Activated-Treg) and 6 (CD4-IEL) of the single cell dataset (10x Genomics).

**Data processing of single cell RNA-seq and single cell TCR-seq libraries**
Raw fastq files derived from our RNA-seq libraries were processed with cellranger count (v3.1.0) using the 10x Genomics prebuilt mouse reference (v3.0.0 mm10). Our libraries were processed independently and merged into a single experiment at the analysis level using Seurat (v3.1.1) [278]. Quality control was performed by removing cells with high (> 5%) mitochondrial UMI content. Cells with more than 4000 or less than 200 genes were excluded from our analysis. TCR contigs and annotation were performed with the Cellranger vdj workflow. Paired TCR clonotypes were defined by the V, (D), J and CDR3 nucleotide composition for alpha and beta chains. Cells in which only one of the TCR sequences was recovered were excluded from the paired TCR clonal composition analysis. TCR clonotype sharing was assigned to cells expressing identical V, (D), J, and CDR3 nucleotide and amino acid sequences. Further processing and statistical analysis were performed using various R packages as described [279].

**Single cell RNAseq normalization and statistical analysis**
The raw UMI counts were normalized by applying a regression model with negative binomial error distribution, available through the SCTransform function in the Seurat (v3.1.1) package [280]. The top 3000 variable genes were first used for dimensional reduction by PCA using the scaled data. The first 30 principal components were further used on the clustering algorithm and UMAP embedding for two dimensional visualization by using the Seurat workflow [278, 280].

**Diffusion map and pseudotime analysis**
To infer cell differentiation trajectories based on the expression data, we used a diffusion map algorithm adapted for sc-RNAseq analysis and implemented through the package destiny [281]or the Seurat dataset was converted into a 'Single Cell Experiment' object and used within Slingshot and Monocle3[282]. Both Slingshot and Monocle3 algorithm was performed by setting the cluster 10 (naïve cells from mLN) as the root. Normalized values were used as input for the Diffusion Map function. Cells were ordered based on the first diffusion component. To visualize lineage differentiation within our UMAP embedding and find differentially expressed genes over pseudotime, we used the slingshot package [283].

**Statistical Analyses**
Statistical analysis was carried out using GraphPad Prism v.8. Flow cytometry analysis was carried out using FlowJo software. Data in graphs show mean +/- SEM and p values <0.05 were considered significant. The Chao1 index [284], a measure of alpha diversity, was calculated using EstimateS software [285]. Diversity 50 was calculated on Excel as the fraction of dominant clones that account for the cumulative 50% of the total paired CDR3s identified in each UMAP cluster. CDR3 similarity was calculated using the Morisita-horn overlap index by using the divo package [286]. GraphPadPrism v.8 was used for graphs and Adobe Illustrator 2019 used to assemble and edit figures.

# References


1. Sujino, T. *et al.* Tissue adaptation of regulatory and intraepithelial CD4(+) T cells controls gut inflammation. *Science (New York, N.Y.)* **352**, 1581-1586 (2016).

2. Zachariah, M.A. & Cyster, J.G. Neural crest-derived pericytes promote egress of mature thymocytes at the corticomedullary junction. *Science (New York, N.Y.)* **328**, 1129-1135 (2010).

3. Berzins, S.P., Godfrey, D.I., Miller, J.F. & Boyd, R.L. A central role for thymic emigrants in peripheral T cell homeostasis. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9787-9791 (1999).

4. McCaughtry, T.M., Wilken, M.S. & Hogquist, K.A. Thymic emigration revisited. *The Journal of experimental medicine* **204**, 2513-2520 (2007).

5. Gascoigne, N.R., Rybakin, V., Acuto, O. & Brzostek, J. TCR Signal Strength and T Cell Development. *Annu Rev Cell Dev Biol* **32**, 327-348 (2016).

6. Starr, T.K., Jameson, S.C. & Hogquist, K.A. Positive and negative selection of T cells. *Annual review of immunology* **21**, 139-176 (2003).

7. Davis, M.M. & Bjorkman, P.J. T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395-402 (1988).

8. Murugan, A., Mora, T., Walczak, A.M. & Callan, C.G., Jr. Statistical inference of the generation probability of T-cell receptors from sequence repertoires. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 16161-16166 (2012).

9. Sethna, Z. *et al.* Insights into immune system development and function from mouse T-cell repertoires. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 2253-2258 (2017).

10. Garcia, K.C. & Adams, E.J. How the T cell receptor sees antigen--a structural view. *Cell* **122**, 333-336 (2005).

11. Hennecke, J. & Wiley, D.C. T cell receptor-MHC interactions up close. *Cell* **104**, 1-4 (2001).

12. Spits, H. Development of alphabeta T cells in the human thymus. *Nat Rev Immunol* **2**, 760-772 (2002).

13. Ellmeier, W., Sawada, S. & Littman, D.R. The regulation of CD4 and CD8 coreceptor gene expression during T cell development. *Annual review of immunology* **17**, 523-554 (1999).

14. Singer, A., Adoro, S. & Park, J.H. Lineage fate and intense debate: myths, models and mechanisms of CD4- versus CD8-lineage choice. *Nat Rev Immunol* **8**, 788-801 (2008).

15. Egawa, T., Tillman, R.E., Naoe, Y., Taniuchi, I. & Littman, D.R. The role of the Runx transcription factors in thymocyte differentiation and in homeostasis of naive T cells. *The Journal of experimental medicine* **204**, 1945-1957 (2007).

16. Woolf, E. *et al.* Runx3 and Runx1 are required for CD8 T cell development during thymopoiesis. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 7731-7736 (2003).

17. Sato, T. *et al.* Dual functions of Runx proteins for reactivating CD8 and silencing CD4 at the commitment process into CD8 thymocytes. *Immunity* **22**, 317-328 (2005).

18. He, X. *et al.* The zinc finger transcription factor Th-POK regulates CD4 versus CD8 T-cell lineage commitment. *Nature* **433**, 826-833 (2005).

19. Sun, G. *et al.* The zinc finger protein cKrox directs CD4 lineage differentiation during intrathymic T cell positive selection. *Nat Immunol* **6**, 373-381 (2005).

20. Egawa, T. & Littman, D.R. ThPOK acts late in specification of the helper T cell lineage and suppresses Runx-mediated commitment to the cytotoxic T cell lineage. *Nat Immunol* **9**, 1131-1139 (2008).

21. Setoguchi, R. *et al.* Repression of the transcription factor Th-POK by Runx complexes in cytotoxic T cell development. *Science (New York, N.Y.)* **319**, 822-825 (2008).

22. Kappes, D.J., He, X. & He, X. Role of the transcription factor Th-POK in CD4:CD8 lineage commitment. *Immunol Rev* **209**, 237-252 (2006).

23. Muroi, S. *et al.* Cascading suppression of transcriptional silencers by ThPOK seals helper T cell fate. *Nat Immunol* **9**, 1113-1121 (2008).

24. Taniuchi, I. CD4 Helper and CD8 Cytotoxic T Cell Differentiation. *Annual review of immunology* **36**, 579-601 (2018).

25. Kappes, D.J. Expanding roles for ThPOK in thymic development. *Immunol Rev* **238**, 182-194 (2010).

26. Mondino, A., Khoruts, A. & Jenkins, M.K. The anatomy of T-cell activation and tolerance. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 2245-2252 (1996).

27. Kambayashi, T. & Laufer, T.M. Atypical MHC class II-expressing antigen-presenting cells: can anything replace a dendritic cell? *Nat Rev Immunol* **14**, 719-730 (2014).

28. Zhu, J., Yamane, H. & Paul, W.E. Differentiation of effector CD4 T cell populations (*). *Annual review of immunology* **28**, 445-489 (2010).

29. Jenkins, M.K. *et al.* In vivo activation of antigen-specific CD4 T cells. *Annual review of immunology* **19**, 23-45 (2001).

30. Panea, C. *et al.* Intestinal Monocyte-Derived Macrophages Control Commensal-Specific Th17 Responses. *Cell reports* **12**, 1314-1324 (2015).

31. Bogunovic, M. *et al.* Origin of the lamina propria dendritic cell network. *Immunity* **31**, 513-525 (2009).

32. Ginhoux, F. *et al.* The origin and development of nonlymphoid tissue CD103+ DCs. *The Journal of experimental medicine* **206**, 3115-3130 (2009).

33. Neutra, M.R., Mantis, N.J. & Kraehenbuhl, J.P. Collaboration of epithelial cells with organized mucosal lymphoid tissues. *Nat Immunol* **2**, 1004-1009 (2001).

34. Niess, J.H. *et al.* CX3CR1-mediated dendritic cell access to the intestinal lumen and bacterial clearance. *Science (New York, N.Y.)* **307**, 254-258 (2005).

35. Schulz, O. *et al.* Intestinal CD103+, but not CX3CR1+, antigen sampling cells migrate in lymph and serve classical dendritic cell functions. *The Journal of experimental medicine* **206**, 3101-3114 (2009).

36. DuPage, M. & Bluestone, J.A. Harnessing the plasticity of CD4(+) T cells to treat immune-mediated disease. *Nat Rev Immunol* **16**, 149-163 (2016).

37. Till, J.E., McCulloch, E.A. & Siminovitch, L. A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells. *Proceedings of the National Academy of Sciences of the United States of America* **51**, 29-36 (1964).

38. Gattinoni, L. *et al.* A human memory T cell subset with stem cell-like properties. *Nat Med* **17**, 1290-1297 (2011).

39. Graef, P. *et al.* Serial transfer of single-cell-derived immunocompetence reveals stemness of CD8(+) central memory T cells. *Immunity* **41**, 116-126 (2014).

40. Ma, F. *et al.* The microRNA miR-29 controls innate and adaptive immune responses to intracellular bacterial infection by targeting interferon-gamma. *Nat Immunol* **12**, 861-869 (2011).

41. O'Connell, R.M. *et al.* MicroRNA-155 promotes autoimmune inflammation by enhancing inflammatory T cell development. *Immunity* **33**, 607-619 (2010).

42. Duhen, R. *et al.* Cutting edge: the pathogenicity of IFN-gamma-producing Th17 cells is independent of T-bet. *Journal of immunology (Baltimore, Md. : 1950)* **190**, 4478-4482 (2013).

43. Voo, K.S. *et al.* Identification of IL-17-producing FOXP3+ regulatory T cells in humans. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 4793-4798 (2009).

44. Pratama, A., Schnell, A., Mathis, D. & Benoist, C. Developmental and cellular age direct conversion of CD4+ T cells into RORgamma+ or Helios+ colon Treg cells. *The Journal of experimental medicine* **217** (2020).

45. Bilate, A.M. & Lafaille, J.J. Induced CD4+Foxp3+ regulatory T cells in immune tolerance. *Annual review of immunology* **30**, 733-758 (2012).

46. Sugimoto, N. *et al.* Foxp3-dependent and -independent molecules specific for CD25+CD4+ natural regulatory T cells revealed by DNA microarray analysis. *Int Immunol* **18**, 1197-1209 (2006).

47. Hill, J.A. *et al.* Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature. *Immunity* **27**, 786-800 (2007).

48. Whibley, N., Tucci, A. & Powrie, F. Regulatory T cell adaptation in the intestine and skin. *Nat Immunol* **20**, 386-396 (2019).

49. Godfrey, V.L., Wilkinson, J.E., Rinchik, E.M. & Russell, L.B. Fatal lymphoreticular disease in the scurfy (sf) mouse requires T cells that mature in a sf thymic environment: potential model for thymic education. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 5528-5532 (1991).

50. Bennett, C.L. *et al.* The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nature genetics* **27**, 20-21 (2001).

51. Mucida, D. *et al.* Oral tolerance in the absence of naturally occurring Tregs. *J Clin Invest* **115**, 1923-1933 (2005).

52. Mottet, C., Uhlig, H.H. & Powrie, F. Cutting edge: cure of colitis by CD4+CD25+ regulatory T cells. *Journal of immunology (Baltimore, Md. : 1950)* **170**, 3939-3943 (2003).

53. Qureshi, O.S. *et al.* Trans-endocytosis of CD80 and CD86: a molecular basis for the cell-extrinsic function of CTLA-4. *Science (New York, N.Y.)* **332**, 600-603 (2011).

54. Liang, B. *et al.* Regulatory T cells inhibit dendritic cells by lymphocyte activation gene-3 engagement of MHC class II. *Journal of immunology (Baltimore, Md. : 1950)* **180**, 5916-5926 (2008).

55. Klein, M. & Bopp, T. Cyclic AMP Represents a Crucial Component of Treg Cell-Mediated Immune Regulation. *Front Immunol* **7**, 315 (2016).

56. Collison, L.W. *et al.* The inhibitory cytokine IL-35 contributes to regulatory T-cell function. *Nature* **450**, 566-569 (2007).

57. Asseman, C., Mauze, S., Leach, M.W., Coffman, R.L. & Powrie, F. An essential role for interleukin 10 in the function of regulatory T cells that inhibit intestinal inflammation. *The Journal of experimental medicine* **190**, 995-1004 (1999).

58. Chaudhry, A. *et al.* Interleukin-10 signaling in regulatory T cells is required for suppression of Th17 cell-mediated inflammation. *Immunity* **34**, 566-578 (2011).

59. Atarashi, K. *et al.* Induction of colonic regulatory T cells by indigenous Clostridium species. *Science (New York, N.Y.)* **331**, 337-341 (2011).

60. Xu, M. *et al.* c-MAF-dependent regulatory T cells mediate immunological tolerance to a gut pathobiont. *Nature* **554**, 373-377 (2018).

61. Zemmour, D. *et al.* Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nature immunology* **19**, 291-301 (2018).

62. Nishio, J. *et al.* Requirement of full TCR repertoire for regulatory T cells to maintain intestinal homeostasis. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 12770-12775 (2015).

63. Solomon, B.D. & Hsieh, C.S. Antigen-Specific Development of Mucosal Foxp3+RORgammat+ T Cells from Regulatory T Cell Precursors. *Journal of immunology (Baltimore, Md. : 1950)* **197**, 3512-3519 (2016).

64. Guy-Grand, D. *et al.* Two gut intraepithelial CD8+ lymphocyte populations with different T cell receptors: a role for the gut epithelium in T cell differentiation. *The Journal of experimental medicine* **173**, 471-481 (1991).

65. Cheroutre, H., Lambolez, F. & Mucida, D. The light and dark sides of intestinal intraepithelial lymphocytes. *Nat Rev Immunol* **11**, 445-456 (2011).

66. Mucida, D. *et al.* Transcriptional reprogramming of mature CD4(+) helper T cells generates distinct MHC class II-restricted cytotoxic T lymphocytes. *Nature immunology* **14**, 281-289 (2013).

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

68. Mackay, L.K. *et al.* The developmental pathway for CD103(+)CD8+ tissue-resident memory T cells of skin. *Nat Immunol* **14**, 1294-1301 (2013).

69. Gebhardt, T. *et al.* Memory T cells in nonlymphoid tissue that provide enhanced local immunity during infection with herpes simplex virus. *Nat Immunol* **10**, 524-530 (2009).

70. Masopust, D. *et al.* Dynamic T cell migration program provides resident memory within intestinal epithelium. *The Journal of experimental medicine* **207**, 553-564 (2010).

71. Schenkel, J.M. & Masopust, D. Tissue-resident memory T cells. *Immunity* **41**, 886-897 (2014).

72. Masopust, D., Vezys, V., Marzo, A.L. & Lefrancois, L. Preferential localization of effector memory cells in nonlymphoid tissue. *Science (New York, N.Y.)* **291**, 2413-2417 (2001).

73. Ariotti, S., Haanen, J.B. & Schumacher, T.N. Behavior and function of tissue-resident memory T cells. *Adv Immunol* **114**, 203-216 (2012).

74. Gebhardt, T. *et al.* Different patterns of peripheral migration by memory CD4+ and CD8+ T cells. *Nature* **477**, 216-219 (2011).

75. Thompson, E.A. *et al.* Interstitial Migration of CD8alphabeta T Cells in the Small Intestine Is Dynamic and Is Dictated by Environmental Cues. *Cell reports* **26**, 2859-2867 e2854 (2019).

76. Hoytema van Konijnenburg, D.P. *et al.* Intestinal Epithelial and Intraepithelial T Cell Crosstalk Mediates a Dynamic Response to Infection. *Cell* **171**, 783-794 e713 (2017).

77. Mackay, L.K. *et al.* Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science (New York, N.Y.)* **352**, 459-463 (2016).

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

79. Casey, K.A. *et al.* Antigen-independent differentiation and maintenance of effector-like resident memory T cells in tissues. *Journal of immunology (Baltimore, Md. : 1950)* **188**, 4866-4875 (2012).

80. Freeman, B.E., Hammarlund, E., Raue, H.P. & Slifka, M.K. Regulation of innate CD8+ T-cell activation mediated by cytokines. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 9971-9976 (2012).

81. Veillette, A., Bookman, M.A., Horak, E.M. & Bolen, J.B. The CD4 and CD8 T cell surface antigens are associated with the internal membrane tyrosine-protein kinase p56lck. *Cell* **55**, 301-308 (1988).

82. Cawthon, A.G., Lu, H. & Alexander-Miller, M.A. Peptide requirement for CTL activation reflects the sensitivity to CD3 engagement: correlation with

CD8alphabeta versus CD8alphaalpha expression. *Journal of immunology (Baltimore, Md. : 1950)* **167**, 2577-2584 (2001).

83. Hershberg, R. *et al.* Expression of the thymus leukemia antigen in mouse intestinal epithelium. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 9727-9731 (1990).

84. Leishman, A.J. *et al.* T cell responses modulated through interaction between CD8alphaalpha and the nonclassical MHC class I molecule, TL. *Science (New York, N.Y.)* **294**, 1936-1939 (2001).

85. Cheroutre, H. & Lambolez, F. Doubting the TCR coreceptor function of CD8alphaalpha. *Immunity* **28**, 149-159 (2008).

86. Dalton, J.E. *et al.* Intraepithelial gammadelta+ lymphocytes maintain the integrity of intestinal epithelial tight junctions in response to infection. *Gastroenterology* **131**, 818-829 (2006).

87. Ismail, A.S. *et al.* Gammadelta intraepithelial lymphocytes are essential mediators of host-microbial homeostasis at the intestinal mucosal surface. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 8743-8748 (2011).

88. Sheridan, B.S. *et al.* gammadelta T cells exhibit multifunctional and protective memory in intestinal tissues. *Immunity* **39**, 184-195 (2013).

89. Offit, P.A. & Dudzik, K.I. Rotavirus-specific cytotoxic T lymphocytes appear at the intestinal mucosal surface after rotavirus infection. *J Virol* **63**, 3507-3512 (1989).

90. Jabri, B. *et al.* Selective expansion of intraepithelial lymphocytes expressing the HLA-E-specific natural killer receptor CD94 in celiac disease. *Gastroenterology* **118**, 867-879 (2000).

91. Faria, A.M.C., Reis, B.S. & Mucida, D. Tissue adaptation: Implications for gut immunity and tolerance. *The Journal of experimental medicine* **214**, 1211-1226 (2017).

92. Meresse, B. *et al.* Coordinated induction by IL15 of a TCR-independent NKG2D signaling pathway converts CTL into lymphokine-activated killer cells in celiac disease. *Immunity* **21**, 357-366 (2004).

93. Carton, J., Byrne, B., Madrigal-Estebas, L., O'Donoghue, D.P. & O'Farrelly, C. CD4+CD8+ human small intestinal T cells are decreased in coeliac

patients, with CD8 expression downregulated on intra-epithelial T cells in the active disease. *Eur J Gastroenterol Hepatol* **16**, 961-968 (2004).

94. Senju, M., Wu, K.C., Mahida, Y.R. & Jewell, D.P. Coexpression of CD4 and CD8 on peripheral blood T cells and lamina propria T cells in inflammatory bowel disease by two colour immunofluorescence and flow cytometric analysis. *Gut* **32**, 918-922 (1991).

95. Hooper, L.V. & Macpherson, A.J. Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat Rev Immunol* **10**, 159-169 (2010).

96. Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol* **14**, e1002533 (2016).

97. Vaishnava, S., Behrendt, C.L., Ismail, A.S., Eckmann, L. & Hooper, L.V. Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 20858-20863 (2008).

98. Fahlgren, A., Hammarstrom, S., Danielsson, A. & Hammarstrom, M.L. Increased expression of antimicrobial peptides and lysozyme in colonic epithelial cells of patients with ulcerative colitis. *Clin Exp Immunol* **131**, 90-101 (2003).

99. Bevins, C.L. & Salzman, N.H. Paneth cells, antimicrobial peptides and maintenance of intestinal homeostasis. *Nat Rev Microbiol* **9**, 356-368 (2011).

100. Muniz, L.R., Knosp, C. & Yeretssian, G. Intestinal antimicrobial peptides during homeostasis, infection, and disease. *Front Immunol* **3**, 310 (2012).

101. Salzman, N.H. *et al.* Enteric defensins are essential regulators of intestinal microbial ecology. *Nat Immunol* **11**, 76-83 (2010).

102. Smith, K., McCoy, K.D. & Macpherson, A.J. Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota. *Semin Immunol* **19**, 59-69 (2007).

103. Lu, P. *et al.* Intestinal epithelial Toll-like receptor 4 prevents metabolic syndrome by regulating interactions between microbes and intestinal epithelial cells in mice. *Mucosal Immunol* **11**, 727-740 (2018).

104. Ramanan, D., Tang, M.S., Bowcutt, R., Loke, P. & Cadwell, K. Bacterial sensor Nod2 prevents inflammation of the small intestine by restricting the expansion of the commensal Bacteroides vulgatus. *Immunity* **41**, 311-324 (2014).

105. Biton, M. *et al.* T Helper Cell Cytokines Modulate Intestinal Stem Cell Renewal and Differentiation. *Cell* **175**, 1307-1320 e1322 (2018).

106. Ladinsky, M.S. *et al.* Endocytosis of commensal antigens by intestinal epithelial cells regulates mucosal T cell homeostasis. *Science (New York, N.Y.)* **363** (2019).

107. Davis, C.P. & Savage, D.C. Habitat, succession, attachment, and morphology of segmented, filamentous microbes indigenous to the murine gastrointestinal tract. *Infect Immun* **10**, 948-956 (1974).

108. Holscher, H.D. Dietary fiber and prebiotics and the gastrointestinal microbiota. *Gut Microbes* **8**, 172-184 (2017).

109. Park, J.H. *et al.* Promotion of Intestinal Epithelial Cell Turnover by Commensal Bacteria: Role of Short-Chain Fatty Acids. *PLoS One* **11**, e0156334 (2016).

110. Schneider, C. *et al.* A Metabolite-Triggered Tuft Cell-ILC2 Circuit Drives Small Intestinal Remodeling. *Cell* **174**, 271-284 e214 (2018).

111. Kanaya, T. & Ohno, H. The Mechanisms of M-cell Differentiation. *Biosci Microbiota Food Health* **33**, 91-97 (2014).

112. von Moltke, J., Ji, M., Liang, H.E. & Locksley, R.M. Tuft-cell-derived IL-25 regulates an intestinal ILC2-epithelial response circuit. *Nature* **529**, 221-225 (2016).

113. Nakamura, Y., Kimura, S. & Hase, K. M cell-dependent antigen uptake on follicle-associated epithelium for mucosal immune surveillance. *Inflamm Regen* **38**, 15 (2018).

114. Andrews, C., McLean, M.H. & Durum, S.K. Cytokine Tuning of Intestinal Epithelial Function. *Front Immunol* **9**, 1270 (2018).

115. Hu, M.D. *et al.* Epithelial IL-15 Is a Critical Regulator of gammadelta Intraepithelial Lymphocyte Motility within the Intestinal Mucosa. *Journal of immunology (Baltimore, Md. : 1950)* **201**, 747-756 (2018).

116. Ihara, S., Hirata, Y. & Koike, K. TGF-beta in inflammatory bowel disease: a key regulator of immune cells, epithelium, and the intestinal microbiota. *J Gastroenterol* **52**, 777-787 (2017).

117. Oliveira, L.M., Teixeira, F.M.E. & Sato, M.N. Impact of Retinoic Acid on Immune Cells and Inflammatory Diseases. *Mediators Inflamm* **2018**, 3067126 (2018).

118. Tsilingiri, K., Fornasa, G. & Rescigno, M. Thymic Stromal Lymphopoietin: To Cut a Long Story Short. *Cell Mol Gastroenterol Hepatol* **3**, 174-182 (2017).

119. Mucida, D. *et al.* Reciprocal TH17 and regulatory T cell differentiation mediated by retinoic acid. *Science (New York, N.Y.)* **317**, 256-260 (2007).

120. Hadley, G.A., Bartlett, S.T., Via, C.S., Rostapshova, E.A. & Moainie, S. The epithelial cell-specific integrin, CD103 (alpha E integrin), defines a novel subset of alloreactive CD8+ CTL. *Journal of immunology (Baltimore, Md. : 1950)* **159**, 3748-3756 (1997).

121. Iwata, M. *et al.* Retinoic acid imprints gut-homing specificity on T cells. *Immunity* **21**, 527-538 (2004).

122. Shale, M., Schiering, C. & Powrie, F. CD4(+) T-cell subsets in intestinal inflammation. *Immunological reviews* **252**, 164-182 (2013).

123. Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446-450 (2013).

124. Smith, P.M. *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 569-573 (2013).

125. Arpaia, N. *et al.* Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* **504**, 451-455 (2013).

126. Atarashi, K. *et al.* Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* **500**, 232-236 (2013).

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

128. Coombes, J.L. *et al.* A functionally specialized population of mucosal CD103+ DCs induces Foxp3+ regulatory T cells via a TGF-beta and retinoic

acid-dependent mechanism. *The Journal of experimental medicine* **204**, 1757-1764 (2007).

129. Hall, J.A., Grainger, J.R., Spencer, S.P. & Belkaid, Y. The role of retinoic acid in tolerance and immunity. *Immunity* **35**, 13-22 (2011).

130. Sun, C.M. *et al.* Small intestine lamina propria dendritic cells promote de novo generation of Foxp3 T reg cells via retinoic acid. *The Journal of experimental medicine* **204**, 1775-1785 (2007).

131. Rubtsov, Y.P. *et al.* Stability of the regulatory T cell lineage in vivo. *Science (New York, N.Y.)* **329**, 1667-1671 (2010).

132. Edelblum, K.L. *et al.* Dynamic migration of gammadelta intraepithelial lymphocytes requires occludin. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 7097-7102 (2012).

133. Materials and methods are available as supplementary materials on *Science* Online.

134. Cervantes-Barragan, L. *et al.* Lactobacillus reuteri induces gut intraepithelial CD4(+)CD8alphaalpha(+) T cells. *Science (New York, N.Y.)* **357**, 806-810 (2017).

135. Umesaki, Y., Setoyama, H., Matsumoto, S. & Okada, Y. Expansion of alpha beta T-cell receptor-bearing intestinal intraepithelial lymphocytes after microbial colonization in germ-free mice and its independence from thymus. *Immunology* **79**, 32-37 (1993).

136. Langille, M.G. *et al.* Microbial shifts in the aging mouse gut. *Microbiome* **2**, 50 (2014).

137. Atarashi, K. *et al.* Induction of Colonic Regulatory T Cells by Indigenous Clostridium Species. *Science* (2010).

138. Sledzinska, A. *et al.* TGF-beta signalling is required for CD4(+) T cell homeostasis but dispensable for regulatory T cell function. *PLoS biology* **11**, e1001674 (2013).

139. Klinger, M. *et al.* Thymic OX40 expression discriminates cells undergoing strong responses to selection ligands. *J Immunol* **182**, 4581-4589 (2009).

140. Maekawa, Y. *et al.* Notch2 integrates signaling by the transcription factors RBP-J and CREB1 to promote T cell cytotoxicity. *Nat Immunol* **9**, 1140-1147 (2008).

141. Rubtsov, Y.P. *et al.* Regulatory T cell-derived interleukin-10 limits inflammation at environmental interfaces. *Immunity* **28**, 546-558 (2008).

142. Miyao, T. *et al.* Plasticity of Foxp3(+) T cells reflects promiscuous Foxp3 expression in conventional T cells but not reprogramming of regulatory T cells. *Immunity* **36**, 262-275 (2012).

143. Tanoue, T., Atarashi, K. & Honda, K. Development and maintenance of intestinal regulatory T cells. *Nat Rev Immunol* **16**, 295-309 (2016).

144. Kim, K.S. *et al.* Dietary antigens limit mucosal immunity by inducing regulatory T cells in the small intestine. *Science* (2016).

145. Naoe, Y. *et al.* Repression of interleukin-4 in T helper type 1 cells by Runx/Cbf beta binding to the Il4 silencer. *The Journal of experimental medicine* **204**, 1749-1755 (2007).

146. Polansky, J.K. *et al.* DNA methylation controls Foxp3 gene expression. *Eur J Immunol* **38**, 1654-1663 (2008).

147. Powrie, F., Leach, M.W., Mauze, S., Caddle, L.B. & Coffman, R.L. Phenotypically distinct subsets of CD4+ T cells induce or protect from chronic intestinal inflammation in C. B-17 scid mice. *Int Immunol* **5**, 1461-1471 (1993).

148. Powrie, F., Correa-Oliveira, R., Mauze, S. & Coffman, R.L. Regulatory interactions between CD45RBhigh and CD45RBlow CD4+ T cells are important for the balance between protective and pathogenic cell-mediated immunity. *The Journal of experimental medicine* **179**, 589-600 (1994).

149. Maloy, K.J. *et al.* CD4+CD25+ T(R) cells suppress innate immune pathology through cytokine-dependent mechanisms. *The Journal of experimental medicine* **197**, 111-119 (2003).

150. Kitoh, A. *et al.* Indispensable role of the Runx1-Cbfbeta transcription complex for in vivo-suppressive function of FoxP3+ regulatory T cells. *Immunity* **31**, 609-620 (2009).

151. Klunker, S. *et al.* Transcription factors RUNX1 and RUNX3 in the induction and suppressive function of Foxp3+ inducible regulatory T cells. *J Exp Med* **206**, 2701-2715 (2009).

152. Li, M.O., Wan, Y.Y. & Flavell, R.A. T cell-produced transforming growth factor-beta1 controls T cell tolerance and regulates Th1- and Th17-cell differentiation. *Immunity* **26**, 579-591 (2007).

153. Curotto de Lafaille, M.A. *et al.* Adaptive Foxp3+ regulatory T cell-dependent and -independent control of allergic inflammation. *Immunity* **29**, 114-126 (2008).

154. Josefowicz, S.Z., Lu, L.F. & Rudensky, A.Y. Regulatory T cells: mechanisms of differentiation and function. *Annu Rev Immunol* **30**, 531-564 (2012).

155. Ohnmacht, C. *et al.* The microbiota regulates type 2 immunity through RORgammat(+) T cells. *Science* **349**, 989-993 (2015).

156. Sefik, E. *et al.* Individual intestinal symbionts induce a distinct population of RORgamma(+) regulatory T cells. *Science* **349**, 993-997 (2015).

157. Bollrath, J. & Powrie, F.M. Controlling the frontier: regulatory T-cells and intestinal homeostasis. *Seminars in immunology* **25**, 352-357 (2013).

158. Jabri, B. & Abadie, V. IL-15 functions as a danger signal to regulate tissue-resident T cells and tissue destruction. *Nat Rev Immunol* **15**, 771-783 (2015).

159. Reis, B.S. *et al.* Leptin receptor signaling in T cells is required for Th17 differentiation. *Journal of immunology* **194**, 5253-5260 (2015).

160. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nature biotechnology* (2018).

161. Miragaia, R.J. *et al.* Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation. *Immunity* **50**, 493-504 e497 (2019).

162. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. & Greenleaf, W.J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213-1218 (2013).

163. Badis, G. *et al.* Diversity and complexity in DNA recognition by transcription factors. *Science* **324**, 1720-1723 (2009).

164. Ciucci, T. *et al.* The Emergence and Functional Fitness of Memory CD4(+) T Cells Require the Transcription Factor Thpok. *Immunity* **50**, 91-105.e104 (2019).

165. Machanick, P. & Bailey, T.L. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics* **27**, 1696-1697 (2011).

166. Milner, J.J. *et al.* Runx3 programs CD8(+) T cell residency in non-lymphoid tissues and tumours. *Nature* **552**, 253-257 (2017).

167. Shan, Q. *et al.* The transcription factor Runx3 guards cytotoxic CD8(+) effector T cells against deviation towards follicular helper T cell lineage. *Nature immunology* **18**, 931-939 (2017).

168. Zhou, X. *et al.* Instability of the transcription factor Foxp3 leads to the generation of pathogenic memory T cells in vivo. *Nat Immunol* **10**, 1000-1007 (2009).

169. Liu, X. *et al.* T cell receptor CDR3 sequence but not recognition characteristics distinguish autoreactive effector and Foxp3(+) regulatory T cells. *Immunity* **31**, 909-920 (2009).

170. Levine, A.G., Arvey, A., Jin, W. & Rudensky, A.Y. Continuous requirement for the TCR in regulatory T cell function. *Nat Immunol* **15**, 1070-1078 (2014).

171. Fan, X. & Rudensky, A.Y. Hallmarks of Tissue-Resident Lymphocytes. *Cell* **164**, 1198-1211 (2016).

172. Lathrop, S.K. *et al.* Peripheral education of the immune system by colonic commensal microbiota. *Nature* **478**, 250-254 (2011).

173. McDonald, B.D., Jabri, B. & Bendelac, A. Diverse developmental pathways of intestinal intraepithelial lymphocytes. *Nat Rev Immunol* **18**, 514-525 (2018).

174. Koyama, M. *et al.* MHC Class II Antigen Presentation by the Intestinal Epithelium Initiates Graft-versus-Host Disease and Is Influenced by the Microbiota. *Immunity* (2019).

175. Ellmeier, W., Sunshine, M.J., Losos, K., Hatam, F. & Littman, D.R. An enhancer that directs lineage-specific expression of CD8 in positively selected thymocytes and mature T cells. *Immunity* **7**, 537-547 (1997).

176. Masopust, D. & Soerens, A.G. Tissue-Resident T Cells and Other Resident Leukocytes. *Annual review of immunology* **37**, 521-546 (2019).

177. DiSpirito, J.R. *et al.* Molecular diversification of regulatory T cells in nonlymphoid tissues. *Science immunology* **3** (2018).

178. Fonseca, R. *et al.* Developmental plasticity allows outside-in immune responses by resident memory T cells. *Nature immunology* (2020).

179. Fonseca, R. *et al.* Developmental plasticity allows outside-in immune responses by resident memory T cells. *Nat Immunol* **21**, 412-421 (2020).

180. Campbell, D.J. & Butcher, E.C. Rapid acquisition of tissue-specific homing phenotypes by CD4(+) T cells activated in cutaneous or mucosal lymphoid tissues. *The Journal of experimental medicine* **195**, 135-141 (2002).

181. Cohen, M. *et al.* Lung Single-Cell Signaling Interaction Map Reveals Basophil Role in Macrophage Imprinting. *Cell* **175**, 1031-1044 e1018 (2018).

182. Ricardo-Gonzalez, R.R. *et al.* Tissue signals imprint ILC2 identity with anticipatory function. *Nat Immunol* **19**, 1093-1099 (2018).

183. Tibbitt, C.A. *et al.* Single-Cell RNA Sequencing of the T Helper Cell Response to House Dust Mites Defines a Distinct Gene Expression Signature in Airway Th2 Cells. *Immunity* **51**, 169-184 e165 (2019).

184. Zheng, Y. *et al.* Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate. *Nature* **463**, 808-812 (2010).

185. Burchill, M.A., Yang, J., Vogtenhuber, C., Blazar, B.R. & Farrar, M.A. IL-2 receptor beta-dependent STAT5 activation is required for the development of Foxp3+ regulatory T cells. *Journal of immunology (Baltimore, Md. : 1950)* **178**, 280-290 (2007).

186. Kim, H.P. & Leonard, W.J. CREB/ATF-dependent T cell receptor-induced FoxP3 gene expression: a role for DNA methylation. *The Journal of experimental medicine* **204**, 1543-1551 (2007).

187. Feng, Y. *et al.* Control of the inheritance of regulatory T cell identity by a cis element in the Foxp3 locus. *Cell* **158**, 749-763 (2014).

188. Ogawa, C. *et al.* TGF-beta-mediated Foxp3 gene expression is cooperatively regulated by Stat5, Creb, and AP-1 through CNS2. *Journal of immunology (Baltimore, Md. : 1950)* **192**, 475-483 (2014).

189. Zhang, Q. *et al.* STAT3- and DNA methyltransferase 1-mediated epigenetic silencing of SHP-1 tyrosine phosphatase tumor suppressor gene in malignant T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 6948-6953 (2005).

190. Li, C. *et al.* MeCP2 enforces Foxp3 expression to promote regulatory T cells' resilience to inflammation. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E2807-2816 (2014).

191. Chen, Z. *et al.* The ubiquitin ligase Stub1 negatively modulates regulatory T cell suppressive activity by promoting degradation of the transcription factor Foxp3. *Immunity* **39**, 272-285 (2013).

192. van Loosdregt, J. *et al.* Stabilization of the transcription factor Foxp3 by the deubiquitinase USP7 increases Treg-cell-suppressive capacity. *Immunity* **39**, 259-271 (2013).

193. Dang, E.V. *et al.* Control of T(H)17/T(reg) balance by hypoxia-inducible factor 1. *Cell* **146**, 772-784 (2011).

194. Hsiao, H.W. *et al.* Deltex1 antagonizes HIF-1alpha and sustains the stability of regulatory T cells in vivo. *Nature communications* **6**, 6353 (2015).

195. Sugiyama, H. *et al.* Dysfunctional blood and target tissue CD4+CD25high regulatory T cells in psoriasis: mechanism underlying unrestrained pathogenic effector T cell proliferation. *Journal of immunology (Baltimore, Md. : 1950)* **174**, 164-173 (2005).

196. Korn, T. & Muschaweckh, A. Stability and Maintenance of Foxp3(+) Treg Cells in Non-lymphoid Microenvironments. *Front Immunol* **10**, 2634 (2019).

197. Hirahara, K. *et al.* The majority of human peripheral blood CD4+CD25highFoxp3+ regulatory T cells bear functional skin-homing receptors. *Journal of immunology (Baltimore, Md. : 1950)* **177**, 4488-4494 (2006).

198. Koch, M.A. *et al.* The transcription factor T-bet controls regulatory T cell homeostasis and function during type 1 inflammation. *Nat Immunol* **10**, 595-602 (2009).

199. Chaudhry, A. *et al.* CD4+ regulatory T cells control TH17 responses in a Stat3-dependent manner. *Science (New York, N.Y.)* **326**, 986-991 (2009).

200. Zheng, Y. *et al.* Regulatory T-cell suppressor program co-opts transcription factor IRF4 to control T(H)2 responses. *Nature* **458**, 351-356 (2009).

201. Takahashi, H. *et al.* TGF-beta and retinoic acid induce the microRNA miR-10a, which targets Bcl-6 and constrains the plasticity of helper T cells. *Nat Immunol* **13**, 587-595 (2012).

202. Samstein, R.M. *et al.* Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell* **151**, 153-166 (2012).

203. Sakaguchi, S., Vignali, D.A., Rudensky, A.Y., Niec, R.E. & Waldmann, H. The plasticity and stability of regulatory T cells. *Nat Rev Immunol* **13**, 461-467 (2013).

204. Floess, S. *et al.* Epigenetic control of the foxp3 locus in regulatory T cells. *PLoS Biol* **5**, e38 (2007).

205. Kim, K.S. *et al.* Dietary antigens limit mucosal immunity by inducing regulatory T cells in the small intestine. *Science (New York, N.Y.)* **351**, 858-863 (2016).

206. Josefowicz, S.Z. *et al.* Extrathymically generated regulatory T cells control mucosal TH2 inflammation. *Nature* **482**, 395-399 (2012).

207. Campbell, C. *et al.* Extrathymically Generated Regulatory T Cells Establish a Niche for Intestinal Border-Dwelling Bacteria and Affect Physiologic Metabolite Balance. *Immunity* **48**, 1245-1257 e1249 (2018).

208. Das, G. *et al.* An important regulatory role for CD4+CD8 alpha alpha T cells in the intestinal epithelial layer in the prevention of inflammatory bowel disease. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5324-5329 (2003).

209. Ulges, A., Schmitt, E., Becker, C. & Bopp, T. Context- and Tissue-Specific Regulation of Immunity and Tolerance by Regulatory T Cells. *Adv Immunol* **132**, 1-46 (2016).

210. Kuswanto, W. *et al.* Poor Repair of Skeletal Muscle in Aging Mice Reflects a Defect in Local, Interleukin-33-Dependent Accumulation of Regulatory T Cells. *Immunity* **44**, 355-367 (2016).

211. Ali, N. & Rosenblum, M.D. Regulatory T cells in skin. *Immunology* **152**, 372-381 (2017).

212. Li, C. *et al.* TCR Transgenic Mice Reveal Stepwise, Multi-site Acquisition of the Distinctive Fat-Treg Phenotype. *Cell* **174**, 285-299 e212 (2018).

213. Masopust, D., Vezys, V., Wherry, E.J., Barber, D.L. & Ahmed, R. Cutting edge: gut microenvironment promotes differentiation of a unique memory CD8 T cell population. *Journal of immunology (Baltimore, Md. : 1950)* **176**, 2079-2083 (2006).

214. Watanabe, R. *et al.* Human skin is protected by four functionally and phenotypically discrete populations of resident and recirculating memory T cells. *Sci Transl Med* **7**, 279ra239 (2015).

215. Rosskopf, S. *et al.* Creation of an engineered APC system to explore and optimize the presentation of immunodominant peptides of major allergens. *Sci Rep* **6**, 31580 (2016).

216. Pasqual, G. *et al.* Monitoring T cell-dendritic cell interactions in vivo by intercellular enzymatic labelling. *Nature* **553**, 496-500 (2018).

217. Masopust, D. & Schenkel, J.M. The integration of T cell migration, differentiation and function. *Nat Rev Immunol* **13**, 309-320 (2013).

218. Mueller, S.N., Gebhardt, T., Carbone, F.R. & Heath, W.R. Memory T cell subsets, migration patterns, and tissue residence. *Annual review of immunology* **31**, 137-161 (2013).

219. Schenkel, J.M., Fraser, K.A. & Masopust, D. Cutting edge: resident memory CD8 T cells occupy frontline niches in secondary lymphoid organs. *Journal of immunology (Baltimore, Md. : 1950)* **192**, 2961-2964 (2014).

220. Makino, Y. *et al.* Impaired T cell function in RANTES-deficient mice. *Clin Immunol* **102**, 302-309 (2002).

221. Lee, Y.T. *et al.* Environmental and antigen receptor-derived signals support sustained surveillance of the lungs by pathogen-specific cytotoxic T lymphocytes. *J Virol* **85**, 4085-4094 (2011).

222. Sheridan, B.S. *et al.* Oral infection drives a distinct population of intestinal resident memory CD8(+) T cells with enhanced protective function. *Immunity* **40**, 747-757 (2014).

223. Zhang, N. & Bevan, M.J. Transforming growth factor-beta signaling controls the formation and maintenance of gut-resident memory T cells by regulating migration and retention. *Immunity* **39**, 687-696 (2013).

224. El-Asady, R. *et al.* TGF-{beta}-dependent CD103 expression by CD8(+) T cells promotes selective destruction of the host intestinal epithelium during graft-versus-host disease. *The Journal of experimental medicine* **201**, 1647-1657 (2005).

225. Harrison, O.J. *et al.* Commensal-specific T cell plasticity promotes rapid tissue adaptation to injury. *Science (New York, N.Y.)* **363** (2019).

226. Zhou, Y., Kim, J., Yuan, X. & Braun, T. Epigenetic modifications of stem cells: a paradigm for the control of cardiac progenitor cells. *Circ Res* **109**, 1067-1081 (2011).

227. Joshi, N.S. *et al.* Inflammation directs memory precursor and short-lived effector CD8(+) T cell fates via the graded expression of T-bet transcription factor. *Immunity* **27**, 281-295 (2007).

228. Delacher, M. *et al.* Precursors for Nonlymphoid-Tissue Treg Cells Reside in Secondary Lymphoid Organs and Are Programmed by the Transcription Factor BATF. *Immunity* (2020).

229. Bruno, L. *et al.* Runx proteins regulate Foxp3 expression. *J Exp Med* **206**, 2329-2337 (2009).

230. Carpenter, A.C. *et al.* Control of Regulatory T Cell Differentiation by the Transcription Factors Thpok and LRF. *Journal of immunology (Baltimore, Md. : 1950)* **199**, 1716-1728 (2017).

231. Carreau, A., El Hafny-Rahbi, B., Matejuk, A., Grillon, C. & Kieda, C. Why is the partial oxygen pressure of human tissues a crucial parameter? Small molecules and hypoxia. *J Cell Mol Med* **15**, 1239-1253 (2011).

232. Fisher, E.M., Khan, M., Salisbury, R. & Kuppusamy, P. Noninvasive monitoring of small intestinal oxygen in a rat model of chronic mesenteric ischemia. *Cell Biochem Biophys* **67**, 451-459 (2013).

233. Taylor, C.T. & Colgan, S.P. Hypoxia and gastrointestinal disease. *J Mol Med (Berl)* **85**, 1295-1300 (2007).

234. Sun, L. *et al.* Intestinal Epithelial Cells-Derived Hypoxia-Inducible Factor-1alpha Is Essential for the Homeostasis of Intestinal Intraepithelial Lymphocytes. *Front Immunol* **10**, 806 (2019).

235. Osborne, B.A. *et al.* Identification of genes induced during apoptosis in T lymphocytes. *Immunol Rev* **142**, 301-320 (1994).

236. Moran, A.E. *et al.* T cell receptor signal strength in Treg and iNKT cell development demonstrated by a novel fluorescent reporter mouse. *The Journal of experimental medicine* **208**, 1279-1289 (2011).

237. Moran, A.E. *et al.* T cell receptor signal strength in Treg and iNKT cell development demonstrated by a novel fluorescent reporter mouse. *J Exp Med* **208**, 1279-1289 (2011).

238. Umesaki, Y., Okada, Y., Matsumoto, S., Imaoka, A. & Setoyama, H. Segmented filamentous bacteria are indigenous intestinal bacteria that activate intraepithelial lymphocytes and induce MHC class II molecules and fucosyl asialo GM1 glycolipids on the small intestinal epithelial cells in the ex-germ-free mouse. *Microbiol Immunol* **39**, 555-562 (1995).

239. Ivanov, I.I. *et al.* Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485-498 (2009).

240. Panja, A., Goldberg, S., Eckmann, L., Krishen, P. & Mayer, L. The regulation and functional consequence of proinflammatory cytokine binding on human intestinal epithelial cells. *Journal of immunology* **161**, 3675-3684 (1998).

241. Bilate, A.M. *et al.* Tissue-specific emergence of regulatory and intraepithelial T cells from a clonal T cell precursor. *Sci Immunol* **1**, eaaf7471 (2016).

242. Rocha, B., Vassalli, P. & Guy-Grand, D. The V beta repertoire of mouse gut homodimeric alpha CD8+ intraepithelial T cell receptor alpha/beta + lymphocytes reveals a major extrathymic pathway of T cell differentiation. *J Exp Med* **173**, 483-486 (1991).

243. Regnault, A., Cumano, A., Vassalli, P., Guy-Grand, D. & Kourilsky, P. Oligoclonal repertoire of the CD8 alpha alpha and the CD8 alpha beta TCR-

alpha/beta murine intestinal intraepithelial T lymphocytes: evidence for the random emergence of T cells. *J Exp Med* **180**, 1345-1358. (1994).

244. Regnault, A. *et al.* The expansion and selection of T cell receptor alpha beta intestinal intraepithelial T cell clones. *Eur J Immunol* **26**, 914-921 (1996).

245. Wojciech, L. *et al.* Non-canonicaly recruited TCRalphabetaCD8alphaalpha IELs recognize microbial antigens. *Sci Rep* **8**, 10848 (2018).

246. Yamagata, T., Mathis, D. & Benoist, C. Self-reactivity in thymic double-positive cells commits cells to a CD8 alpha alpha lineage with characteristics of innate immune cells. *Nat Immunol* **5**, 597-605 (2004).

247. Leishman, A.J. *et al.* Precursors of functional MHC class I- or class II-restricted CD8alphaalpha(+) T cells are positively selected in the thymus by agonist self-peptides. *Immunity* **16**, 355-364 (2002).

248. Mayans, S. *et al.* alphabetaT cell receptors expressed by CD4(-)CD8alphabeta(-) intraepithelial T cells drive their fate into a unique lineage with unusual MHC reactivities. *Immunity* **41**, 207-218 (2014).

249. McDonald, B.D., Bunker, J.J., Ishizuka, I.E., Jabri, B. & Bendelac, A. Elevated T cell receptor signaling identifies a thymic precursor to the TCRalphabeta(+)CD4(-)CD8beta(-) intraepithelial lymphocyte lineage. *Immunity* **41**, 219-229 (2014).

250. Cortez, V.S. *et al.* CRTAM controls residency of gut CD4+CD8+ T cells in the steady state and maintenance of gut CD4+ Th17 during parasitic infection. *J Exp Med* **211**, 623-633 (2014).

251. Mucida, D. *et al.* Oral tolerance in the absence of naturally occurring Tregs. *J Clin Invest* **115**, 1923-1933 (2005).

252. Curotto de Lafaille, M.A., Lino, A.C., Kutchukhidze, N. & Lafaille, J.J. CD25-T cells generate CD25+Foxp3+ regulatory T cells by peripheral expansion. *J Immunol* **173**, 7259-7268 (2004).

253. Brugiroux, S. *et al.* Genome-guided design of a defined mouse microbiota that confers colonization resistance against Salmonella enterica serovar Typhimurium. *Nat Microbiol* **2**, 16215 (2016).

254. el Marjou, F. *et al.* Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. *Genesis* **39**, 186-193 (2004).

255. Chassaing, B. *et al.* Fecal lipocalin 2, a sensitive and broadly dynamic non-invasive biomarker for intestinal inflammation. *PloS one* **7**, e44328 (2012).

256. Buenrostro, J.D., Wu, B., Chang, H.Y. & Greenleaf, W.J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21 29 21-29 (2015).

257. Boyman, O., Kovar, M., Rubinstein, M.P., Surh, C.D. & Sprent, J. Selective stimulation of T cell subsets with antibody-cytokine immune complexes. *Science (New York, N.Y.)* **311**, 1924-1927 (2006).

258. Webster, K.E. *et al.* In vivo expansion of T reg cells with IL-2-mAb complexes: induction of resistance to EAE and long-term acceptance of islet allografts without immunosuppression. *The Journal of experimental medicine* **206**, 751-760 (2009).

259. Trombetta, J.J. *et al.* Preparation of Single-Cell RNA-Seq Libraries for Next Generation Sequencing. *Curr Protoc Mol Biol* **107**, 4 22 21-17 (2014).

260. Dash, P. *et al.* Paired analysis of TCRalpha and TCRbeta chains at the single-cell level in mice. *J Clin Invest* **121**, 288-295 (2011).

261. Han, A., Glanville, J., Hansmann, L. & Davis, M.M. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat Biotechnol* **32**, 684-692 (2014).

262. Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G. & Neufeld, J.D. PANDAseq: PAired-eND Assembler for Illumina sequences. *Bmc Bioinformatics* **13** (2012).

263. Brochet, X., Lefranc, M.P. & Giudicelli, V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res* **36**, W503-508 (2008).

264. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).

265. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

266. Bailey, T.L. & Machanick, P. Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Res* **40**, e128 (2012).

267. Jolma, A. *et al.* DNA-binding specificities of human transcription factors. *Cell* **152**, 327-339 (2013).

268. Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic acids research* **48**, D87-D92 (2020).

269. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

270. Kharchenko, P.V., Tolstorukov, M.Y. & Park, P.J. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat Biotechnol* **26**, 1351-1359 (2008).

271. Li, Q., Brown, J.B., Huang, H. & Bickel, P.J. Measuring reproducibility of high-throughput experiments. *Ann. Appl. Stat.* **5**, 1752-1779 (2011).

272. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010).

273. Bray, N.L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nature biotechnology* **34**, 525-527 (2016).

274. Pimentel, H., Bray, N.L., Puente, S., Melsted, P. & Pachter, L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature methods* **14**, 687-690 (2017).

275. Team, R.C. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing* (2020).

276. Ritchie, M.E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).

277. Sergushichev, S.A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv* (2016).

278. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-+ (2019).

279. RCoreTeam. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing* (2019).

280. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *bioRxiv*, 576827 (2019).

281. Angerer, P. *et al.* destiny: diffusion maps for large-scale single cell data in R. *Bioinformatics* **32**, 1241-1243 (2016).

282. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496-502 (2019).

283. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *Bmc Genomics* **19** (2018).

284. Chao, A. Nonparametric-Estimation of the Number of Classes in a Population. *Scand J Stat* **11**, 265-270 (1984).

285. Colwell, R.K. *et al.* Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *J Plant Ecol* **5**, 3-21 (2012).

286. Rempala, G.A. & Seweryn, M. Methods for diversity and overlap analysis in T-cell receptor populations. *J Math Biol* **67**, 1339-1368 (2013).