Heterogeneity of CD4+ T Cells

Heterogeneity of CD4+ T cells

A dissertation presented

by

Evgeny Kiner

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

April 2019

Dissertation Advisers: Drs. Christophe Benoist and Diane Mathis Evgeny Kiner

Heterogeneity of CD4+ T cells

Abstract

CD4+ T cells, which consist of T helper cells and T regulatory cells (Tregs) are indispensable for adaptive immunity. T helper cells are key positive regulators of both the humoral and cellular immune responses to infections, whereas Tregs attenuate immune responses and are important for tissue homeostasis. It has been known for over 30 years that CD4+ T cells are a heterogenous population that can be divided into several groups, which were largely defined *in vitro* according to their cytokine production profiles. First, we attempted to unravel the T cell heterogeneity *in vivo* at a single cell resolution. To this end, we performed single cell RNA-seq on colonic CD4+ T cells from healthy mice, as well as mice infected with pathogens that are known to skew the T cells towards particular cytokine responses. Upon infection, effector CD4+ T cells completely change and diversify their transcriptome. However, these transcriptional changes only partially fit with the classic T helper programs. Interestingly, *Il17a*-expressing cells from different conditions are transcriptionally more similar to *Ifng*-expressing cells from the same condition, suggesting that these are not entirely distinct cell subsets. Instead, we identified several overlapping gene modules that orchestrated a continuum of *Ifng* and *Il17a*-expressing cell states. Transcriptional profiling and ELISA of sorted populations from different "poles" of the continuum further validated these data. Moreover, we identified several new CD4+ effector T cells subpopulations, such as a population that

expresses some myeloid genes. Our findings suggest that Th1/17 are not accurate classifications *in vivo,* but that T helper cells are a heterogenous population, which is not defined solely by their cytokine production ability. Second, we developed a set of CRISPR/Cas9 tools to study the effect of loss of function (LOF) mutations on Tregs' heterogeneity *in vivo*. We employed this to validate certain genes that were found to be specifically overexpressed in tumor Tregs, which dampen immune responses to tumors. We found that LOF of *Tnfrsf8*, *Cxcr3*, and *Samsn1* in Tregs led to their depletion in tumors. Thus, these genes may serve as potential targets for cancer immunotherapy.

Acknowledgements

I would like to thank my PhD mentors, Drs. Christophe Benoist and Diane Mathis, for their continuous support, mentoring and for always intellectually challenging me. I would also like to thank the present and past members of the CBDM lab for creating a great environment to work in, continuous help and useful discussions. Special thanks to David Zemmour for advice with bioinformatics, and to Deepshika Ramanan, Martina Sassone-Corsi and Esen Sefik for advice and help with gut experiments. I would also like to thank my bay-mates Jun Cho, Tianli Xiao, Devapragasan Moodley, Lidiya Denu, Sara Haidermota and Allison Cook for creating a fun environment between and during experiments. I would also like to acknowledge my collaborators in the lab- Angela Magnuson, Yanbo Zhang and Brinda Vijaykumar, who I had the pleasure to work with on exciting projects. I am also thankful to Catherine Laplace who helped me with figures and for the thesis.

A special thank you to the Boehringer Ingelheim Fonds for their generous PhD fellowship and continuous support, as well as for connecting me with many new friends. I would also like to thank my Dissertation advisory committee members: Drs. George Church, Judy Lieberman and Nicholas Haining for useful feedback and ideas, as well as my thesis defense committee: Drs. Eric Pamer, Ana Anderson and Jun Huh for taking the time to read and evaluate my dissertation.

I would not reach this point without my previous mentors, Drs. Phillip Sharp, Michael Goldberg, and Michael Boutros, in whose labs I worked and learned a lot, as well as Drs. Nancy

Hopkins, Robert Horvitz and Tyler Jacks for helping me with career advice and for improving my science communication skills. A special thanks to my friends and mentors, Zohreh Amoozgar and Hye-Jung Kim, who have helped me to get through graduate school and make the right decisions at important intersections.

I could not have done it without the unconditional support of my family, loved ones and friends. My parents, Irina and Daniel, my brother, Mark, and my grandparents, Zhanna and Vladimir have always supported me with every decision I made, and have always loved me unconditionally. My aunt and uncle, Olga and Alex, who are my family in Boston and made me feel at home away from home. My friends in Boston, in the graduate program, and outside of it, supported me throughout my PhD. And last but not least, Annie, thank you for all your love and support through graduate school.

Lastly, I would like to dedicate my thesis to my grandfather, Leonid, who passed away, but was the one who engrained in me the love for science. Even though he pursued a PhD in physics, he was unable to finish it, a fact that bothered him till the end of his days. I am proud to finally finish what he had begun.

Table of contents

Chapter 1: Introduction

1.1 The diverse functions of CD4 T cells

CD4 T cells, also known as T helper cells, are indispensable for adaptive immunity as they are key regulators of both the humoral and cellular immune responses. Their importance in the immune system is evident as patients who do not harbor T cells due to either genetic causes (primary T cell deficiency) or HIV infection (which causes decreased counts of CD4 T cells in the blood) easily succumb to bacterial and viral infections, and additionally develop tumors at a higher rate (Geha et al., 2008).

CD4 T cells arise from common lymphoid progenitors (CLPs) that seed the thymus and differentiate into double positive $CD4^{+}CD8^{+}$ T cells (Godfrey et al., 1993). In the thymic cortex they undergo positive selection, in which T cells that can bind MHC class II become CD4-single-positive cells. They then undergo negative selection in the thymic medulla, where autoreactive cells are deleted (Davis et al., 1998). This way, only CD4 T cells that can bind MHC class II, but not at high affinity to MHC class II loaded with self-peptides can exit the thymus. When CD4 T cells egress the thymus, they are naïve and require T cell receptor (TCR) stimulation to become activated perform their role.

CD4 T cells have been shown to perform many diverse functions. In the humoral arm of the immune response, they "help" B cells to properly differentiate, activate and secrete antibodies. Jacques Miller, who initially discovered the immunological function of the thymus, showed that irradiated mice required cells from both the thymus (T cells) and the bone marrow (B cells) to develop an antibody response after immunization (Miller et al., 1961 and Miller et al., 1968). This study, as well as other confirmatory studies with different antigens showed that

thymus-derived lymphocytes help bone marrow-derived cells with antibody production, with the exact term “helper cell” coined in 1970 (Raff, 1970).

CD4 T cells participate in the humoral immune response in several ways. First, they secrete cytokines such as IL-4, IL-5, IL-21 or IL-13 to induce B cell differentiation and proliferation in a cell-cell contact independent manner (Howard et al., 1982). Specifically, IL-4 is important for B cell survival, somatic hypermutation and immunoglobulin class switch recombination to IgE and IgG1 (Crotty et al., 2015). IL-4 knockout mice exhibit lower IgE and IgG1 antibody responses to pathogens but mostly normal germinal center formation and primary B cell responses (Kopf et al, 1995). Conversely, IL-5 enhances IgA secretion by B cells (Harriman et al., 1988). IL-5 can help recruit eosinophils to the site of infection as demonstrated by the fact that monoclonal antibodies to IL-5 but not IL-4 prevent eosinophilia in mice (Coffman et al., 1989). Additionally, IL-21 has been shown to be the most potent cytokine for stimulating plasma cell differentiation as well as induction of AID for somatic hypermutation (Parrish-Novak et al., 2000 and Ettinger et al., 2005).

Second, T helper cells directly activate B cells via interaction between their T cell receptor (T cells) and an appropriate peptide:MHC class II complex on the B cells. Upon TCR recognition of the antigen specific B cell, CD40 ligand on the T cell surface can bind the B-cell surface receptor CD40, which causes the B cell to enter into cell cycle (Armitage et al., 1992).

CD4 T helper cells also have important function in the cellular arm of the immune response. Macrophages get activated and polarized to different states by CD4 T helper cells producing either IFN-γ or IL-4 and IL-13 (Murray et al., 2014). Stout and Bottomly (1989) first demonstrated that only IFN-γ producing Th1s but not IL-4 producing Th2s could activate

effector macrophages in the presence of tumor antigens. CD4 T helper cells can secrete IFN-γ and instruct macrophages to kill parasites (M1) secrete IL-4 and suppress macrophages (M2) (Mills et al., 2000).

CD4 T helper cells are indispensable for the adaptive arm of the cellular response as well. Keene and Forman (1982) showed that helper activity is required to generate CD8+ cytotoxic T lymphocytes (CTLs) *in vivo*. Specifically, Jennings et al. (1991) showed that CD4 T helper cells are required for the generation of a primary CTL response against Herpes Simplex Virus (HSV). Although mice can mount a primary CTL response to highly inflammatory antigens of other bacterial and viral infections such as *Listeria monocytogenes* in the absence of CD4 T cells, they cannot develop protective memory to a secondary challenge (Rahemtulla et al., 1991 and Janssen et al., 2003). CD4 help to CTLs is mediated in several ways. Firstly, CD4 T cells can license dendritic cells (DCs) to present a cognate antigen to CTLs and activate them through the upregulation of CD40 (Schoenberger et al., 1998 and Smith et al., 2004). Secondly, CD4 T cells can recruit CTLs to the sites of infection by induction of local chemokine secretion (Nakanishi et al., 2009). Lastly, cytokines secreted by activated CD4 T cells, especially IL-2 and IFN-γ, are important for survival and proliferation of CTLs (although when CTLs get activated, they can produce IFN-γ in an autocrine manner as well) (Buhlmann et al., 1999, Kumaraguru et al., 2005 and Tewari et al., 2007).

How are CD4 T cells instructed which function to perform? Antigen presenting cells (APCs) and especially Dendritic cells (DCs) can sense which pathogen has infiltrated through a combination of Toll-like receptors (TLRs) , and respond accordingly by instructing T cells how to respond during priming by secreting appropriate cytokines (Zhu et al., 2008). For example, DCs

can secrete IL-12 and thus cause effector CD4 T cells to secrete IFN-γ, and promote the cellular arm of the immune response. There is evidence that the same TCR clone can produce different types of effector cells (Tubo et al., 2013). Th response was also shown to be dependent on infectious dose and duration of exposure to antigen (Rogers et al., 1999). For example, *Trichuris muris* infection at low doses (40 eggs) causes chronic Type I response , whereas high doses (400 eggs) elicit a strong Th2 response that leads to parasite expulsion by day 21 post-infection (Klememtowicz et al., 2012). Similarly, mycobacterial immunization is also dependent on dose: whereas a small dose leads to a Th1 response, a large dose elicits a Th2 response (Power et al., 1998).

CD4 T cells orchestrate the immune responses to different pathogens from intracellular bacteria and eukaryotes to viruses and helminths. This myriad of functions and response abilities against different pathogens has always raised the question whether they are a heterogeneous population and how they can be subset into different populations based on their effector functions.

1.2 CD4 T effector cells and their subsets.

1.2.1 The Th1/Th2 dogma- a historical perspective

Since the discovery of T cells, scientists have strived to explain their multitude of functions by dividing them into particular groups. The first important subdivision of T cells was due to the notion that some cells help the humoral response (T helper cells, which express the surface antigens CD4 and CD5 (originally Ly-4 and Ly-1), while others have cytotoxic functions

(CTLs), which express the surface antigen CD8 (originally Ly-2 and Ly-3) (Cantor et al. 1975 and Kisielow et al., 1975). However, even within the T helper cells, immunologists noticed a division of labor between cells that are responsible for cell-mediated immunity which led to delayed-type hypersensitivity (DTH) and cells that are responsible for helping the humoral response (Liew et al., 1974). Other evidence suggested that one type of T helper cells can help B Cells respond to a specific hapten, whereas another type does not produce a response to the hapten, but responds to other determinants in an antigen-independent manner (Marrack et al., 1975).

Interestingly, the first distinction between Th cell subtypes, as well as their naming, was based on their ability to bind to nylon wool columns (Tada et al, 1978). Tada et al. noticed that the nylon purified cells, which they called Type I helper T cells (Th1) did not need to bind Ia antigen (MHC Class II) to augment the antibody response, whereas the nylon adherent cells (Th2) bound MHCII and were able to augment the antibody response to unrelated antigens. Even though Tada et al. do not provide an explanation to the difference in adherence to the nylon wool, I speculate that the cells that they call “Th2” were tightly bound to MHC Class II on the surface of B cells, thus adhering to the nylon wool and being able to elicit a response to non-cognate antigens.

The formal demonstration that T helper cells come in two different flavors arose from a seminal paper by Mosmann and Coffman in 1986. This paper was also the first one that classified T helper cells by the cytokines they produce, a notion that is largely followed to this day, even amongst other immune cell types such as innate lymphoid cells (ILCs), γδ T cells and natural killer T cells (NKTs). Mosmann et al. (1986) isolated different T cell clones and checked their cytokine (then known as lymphokine) secretion profile by using cell lines that were

responsive to only one or a few cytokines. They found that Th1 cells produced Interferon-gamma, IL-2 and GM-CSF, whereas Th2 cells produced B cell-stimulating factor-1 (BSF-1, now known as IL-4). They were also able to show that Th2 cells, but not Th1 cells cause B cells to undergo class switching to IgE and IgG1. Th1s were later shown to induce class switching to IgG2a (Stevens et al., 1988). Later, they showed that Th1 cells are the key players for delayed-type hypersensitivity through cell-mediated immunity, whereas Th2 cells help other adaptive and innate immune cells such as B cells, mast cells or eosinophils to mount humoral Ab responses or immediate hypersensitivity (Cher et al., 1987).

At the same time (both manuscripts came out in different journals on April 1st), Arthur et al. (1986) detected heterogeneity in rat CD4+ leukocytes by using an antibody to CD45RC (OX-22). They observed that OX-22+ cells produced IL-2 in response to a T cell mitogen, while OX-22- cells produced very little IL-2 but provided help to B cells. They correctly concluded that "helper T cells comprise two functional subsets: one that, on appropriate stimulation, synthesizes high levels of IL-2, and may therefore be presumed to play an important role in cell-mediated immunity, and another that plays an essential role in humoral responses to soluble antigens" (Arthur et al., 1986).

An open question remained- what cues determine whether a naïve CD4 T cell becomes Th1 or Th2? Firstly, two groups found that cytokines produced by Th1 and Th2 cells can inhibit the other cell type and enhance their own proliferation in an autocrine fashion, thus efficiently regulating each other's activity. Gajewski et al. (1988) reported that IFN-γ had an anti-proliferative effect on several Th2 clones, while Fernandez-Botran et al. (1988) showed that IL-4 is proliferative for both cell types, but more so for Th2s. Secondly, several groups identified key

upstream cytokines that regulate Th differentiation. Hsieh et al. (1993) showed that IL-12 produced by macrophages was required for Th1 polarization in a model of *Listeria monocytogenes* infection, while Seder et al. (1992) showed that IL-4 presence during Th priming polarizes the cells to a Th2 phenotype. Both cell types required TCR-MHC II engagement for differentiation. Later, IL-18 was shown to synergize with IL-12 in an antigen independent manner to induce IFN-γ production (Yang et al., 1999). Thirdly, molecular effectors of the polarizing cytokines were identified. IL-12 was shown to exert its Th1-polarization through the phosphorylation of Stat4, while IL-4 was shown to signal through Stat6 (Jacobson et al., 1995; Takeda et al., 1996; Kaplan et al., 1996; Shimoda et al., 1996). Finally, "master regulator" transcription factors that regulate Th1 or Th2 polarization were discovered. The transcription factor Gata3 has been shown to be necessary and sufficient of Th2 cytokine gene expression, while the transcription factor T-bet was shown to upregulate IFN-γ and downregulate IL-4 (Zheng W. et al., 1997; Szabo et al., 2000). At that point, the field moved to subsetting Th cells not only by their cytokine expression or secretion profile but also by their expression of the key transcription factors T-bet and Gata3 (and later on, RORγt).

The notion of Th1/Th2 categorization was widely accepted by immunologists, as they always like to classify cells into categories (see Discussion). Th1 cells were considered to be involved mainly in monocyte and CTL activation (Metchnikoff's cellular immunity), whereas Th2 cells were thought to help B cell antibody production (humoral immunity) (Romagnani, 1996). It was also widely accepted that certain infectious agents polarize T helper cells into either the Th1 lineage (intracellular bacteria such as *Listeria*, *Leishmania* and *Salmonella* as well as parasitic eukaryotes such as *Toxoplasma* gondii) or the Th2 lineage (helminths such as *Trichuris*

muris, *Nippostrongylus brasiliensis* and *Heligmosomoides polygyrus)* depending on the cytokine milieu during priming by APCs (Locksley, 1994 and Bottomly, 1988).

These observations led to the re-examination of many autoimmune and infectious diseases (Liew, 2002). Th1 cells were associated with DTH and auto-immune diseases such as Type I diabetes, multiple sclerosis and Hashimotho's thyroiditis, as well as graft -versus-host-disease (GVHD), whereas Th2 were associated with allergic disorders and various atopic disorders (Romagnani, 1996). The largely oversimplified notion that Th1s are involved in tissue damage, whereas Th2 cells are involved in tissue repair and wound healing appeared (Sandler et al., 2003). This was partially due to the Th1/Th2 model infiltrating the myeloid field, establishing M1 (Th1-activated, inflammatory) and M2 (Th2-activated, suppressive) macrophages, a division that is largely not accepted today (Mills et al., 2000 and Gordon, 2003).

Interestingly, the Th1/Th2 model had survived unrevised for almost 20 years. Perhaps new promising therapies that utilized the basic premise of the Th1/Th2 dogma such as anti-TNF-α for treatment of rheumatoid arthritis and Crohn's disease or anti-IL-4 for treatment of asthma and dermatitis convinced immunologists that there is some truth to this model.

1.2.2 Th17 cells- a revision of the Th1/Th2 hypothesis?

Problems with the Th1/Th2 dogma arose when several groups, thinking that tissue damage is controlled by Th1s, tried to see if administering or blocking gamma-interferon would, respectively, worsen or ameliorate Experimental Autoimmune Encephalomeylitis (EAE) (Billiau et al., 1988 and Ferber et al, 1996). To their surprise, IFN-γ depletion or genetic knockouts of *Ifng* or *Tnf* worsened EAE, while administration of either IFN-γ or TNF-α protected from the disease. This led to the first revision of the Th1/Th2 hypothesis in the early 2000s, when several

seminal papers showed that autoimmune inflammation is orchestrated by IL-23, rather than IL-12, and that IL-23 expanded a T cell subset that secreted IL-17. Cua et al. (2003) showed that IL-$23^{-/-}$ mice were resistant to EAE, while IL-$12^{-/-}$ mice, lacking a cytokine that drives Th1 differentation, were more susceptible to EAE, suggesting a central role for IL-23 in autoimmune T cell mediated inflammation. Interestingly, this group still adhered to the Th1/Th2 model and hypothesized that IL-23 is mainly required for Th1 priming by APCs which allowed persistence of chronic inflammation, rather than having a direct effect on a novel T cell subset. Around the same time, Austin Gurney's group in Genentech published a manuscript based on *in vitro* studies that stated "IL-23 can promote an activation state with features distinct from the well characterized Th1 and Th2 profiles." (Aggarwal et al., 2003). Furthermore, the same group correctly identified IL-17 producing CD4 T cells as downstream targets of IL-23 produced by activated DCs. The Th1/Th2 dogma became Th1/Th2/Th17 dogma when Langrish et al. (2005) showed that naïve CD4 T cells can be stimulated *in vitro* and polarized towards a Th1 phenotype if rIl-12 is added to the media, whereas they could be polarized to Th_{IL-17} phenotype if rIl-23 was added. The *in vitro* stimulated Th1 or Th_{IL-17} were then adoptively transferred to naïve recipient mice and although both were able to infiltrate the brain (as these were TCR-transgenic), only Th_{IL-17} cells were able to induce EAE. These data strongly suggested that a newly identified IL-17 producing T cell type, and not Th1 cells, were responsible for autoimmune inflammation (at least in the context of EAE).

Just like with the Th1s/Th2s, many groups noticed reciprocal interactions between Interferon-gamma and IL-17. For example, addition of Interferon-gamma and IL-4 to in vitro stimulated CD4 T cells decreased the production of Il-17 (Park et al., 2005). These data

suggested that Th1s, Th2s and Th17s are unique and opposite states, which provide each other feedback until they reach equilibrium. Th17s were also found to be reciprocal with regulatory T cells (Tregs), as Bettelli et al. (2006) showed that in vitro stimulation of CD4 T cells with TGF- β (which was previously thought to be anti-inflammatory) together with Il-6 differentiates them into Th17s, while stimulation with TGF-β alone generates FOXP3+ Tregs. Interestingly, they showed that the combination of TGF- β and IL-6 is a more potent Th17 inducer than IL-23, the context in which Th17 were first discovered. This, of course, raised the question of how TGF- β can cause such different outcomes, as Treg and Th17 differentiation. Mucida et al. (2007) and Xiao et al. (2008) showed that the vitamin A metabolite Retinoic Acid (RA) can inhibit the IL-6-mediated induction of proinflammatory Th17 cells and thus regulate the balance between Tregs and Th17s. An alternative pathway to generate Th17s was discovered when IL-6 was knocked out and Tregs deleted (Korn et al., 2007). In that context IL-21 cooperated with TGF-β to induce Th17s in IL-$6^{-/-}$ mice through the IL21-receptor (IL21R).

Soon after, just like with Th1 and Th2 cells, the transcription factor driving Th17 differentiation and function was found to be retinoic orphan receptor gammat (RORγt) (Ivanov et al., 2006). *Rorc*-/- mice were less susceptible to EAE and lacked tissue-infiltrating Th17s. Interferon-regulatory factor 4 (IRF4), which has been shown to be essential for development of Th2 cells, has also been shown to be essential for Th17 differentiation and EAE pathology. Th17s can come in different flavors. When differentiated *in vitro* with IL-23 rather than just TGF-β1 and IL-6, Th17 cells become highly pathogenic, expressing T-bet and effector molecules such as GZMB, CCL4, CCL5 and GM-CSF (Lee et al., 2012). This is due to the induction of TGF-β3 in Th17 cells by IL-23, which then in an autocrine manner become pathogenic. Krausgruber et

al. (2016) also showed that Th17s in transfer colitis models can assume different states, some more pathogenic than others. However, they demonstrate that T-bet is not essential for pathogenicity of Th17s, but is important for modulation and multifunctionality of the IL-23-driven inflammatory response (Krausgruber et al., 2016).

Ciofani et al. (2012) used a combination of techniques such as ChIP-Seq of transcription factors and bulk RNA-Seq to model the transcriptional network in Th17 cells. They created a differentiation model in which upon Th17 cytokine stimulation RORγt, c-Maf and Stat3 are recruited to sites that are occupied by *Irf4* and *Batf* upon TCR activation. This suggests that cytokine signaling and TCR signaling cooperate to drive cells into particular T helper phenotypes. Ciofani et al. (2012) also identified the transcription factors *Bcl11b*, *Etv6*, and *Jmjd3* as positive regulators of Th17 differentiation, and *Fosl2* as a negative regulator of IL-17A production.

Yosef et al. (2013) measured transcriptional profiles of Th17 cells at 18 time points during the course of 72 hours and built a similar model. They identified 1291 genes that were differentially expressed during Th17 differentiation and clustered them into 20 modules. They could then identify which gene module corresponded to each time point and build an interaction network of Th17 differentiation. They could also identify and validate (by siRNA knockdown) several regulators of Th17 differentiation such as *Mina, Fas* and *Pou2af1*, which are positive regulators, as well as the negative regulator *Tsc22d3*. Both of these papers established the intricate transcription factor network in Th17s, which is not only dependent on RORγt.

In the last decade the Th17 field has picked up a lot of steam (Figure 1.1), and Th17s have been shown to be involved in almost every disease, both auto-immune and cancer (Patel et al., 2015). For example, single nucleotide polymorphisms (SNPs) in *IL23R* are associated with strong protection against Crohn's disease (Duerr et al., 2006). Conversely, *IL23R* and two other genes involved in the IL-23 signaling pathway were shown to be associated with psoriasis (Nair et al., 2009). Moreover, SNPs in the *CCR6* locus are associated with increased risk to rheumatoid arthritis (Stahl et al., 2010). Th17 were also found in human tumor tissues and were shown to elicit protective inflammation that promotes activation of CTLs to eliminate the tumor (Martin-Orozco et al., 2009).

Figure 1.1

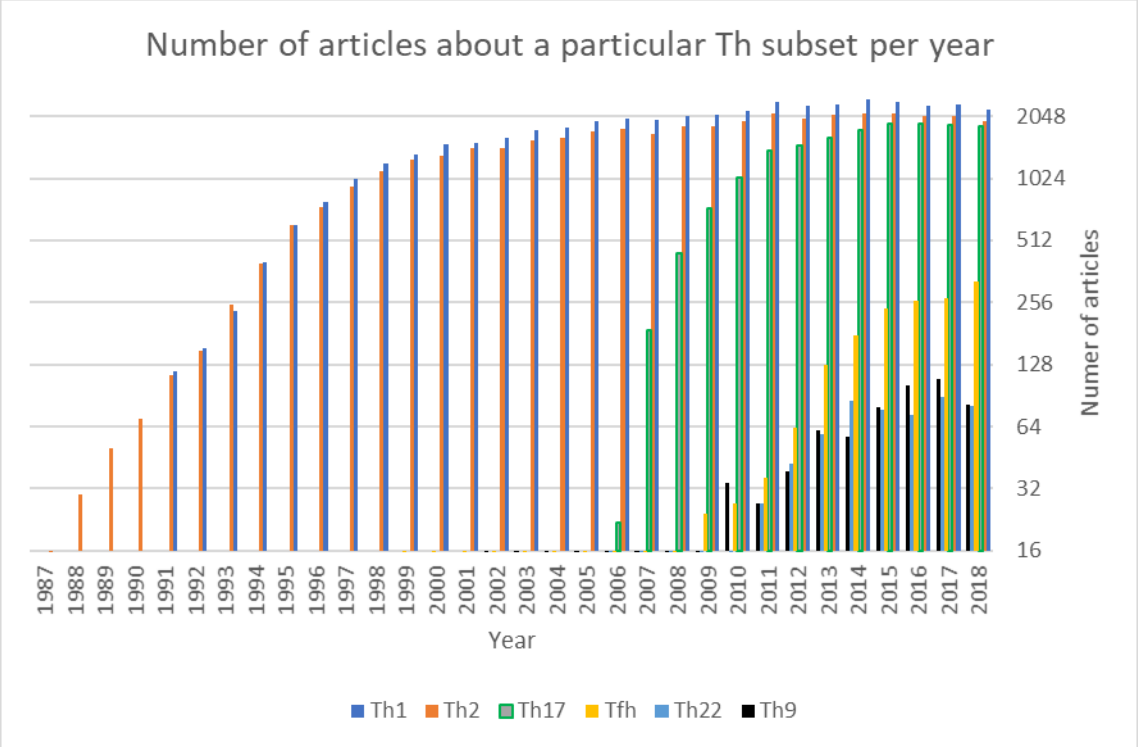


Figure 1.1 Number of PubMed articles by year associated with each "Th" type.

1.2.3 Tfh cells

At around the same time that Th17s were added to the Th1/Th2 paradigm, another T cell subset was born- T follicular helper cells (Tfh). According to the Th1/Th2 paradigm, Th2 cells were thought to be responsible for B cell help, mainly because they produced IL-4, which stimulated B cell proliferation and cell growth (Mossman et al., 1986 and Chtanova et al., 2004). However, the loss of IL-4 in knockout mouse models had no effect on germinal center formation, suggesting that additional cell types and cytokines might play a role in T cell help to B cells other than Th2 cells (Kopf et al., 1995).

Tfh cells were discovered due to the expression of CXCR5 on a subset of human peripheral blood CD4 T cells; As CXCR5 was known to be required for the formation of B cell follicles, Breitfeld et al. (2000) correctly hypothesized and subsequently showed that these cells were responsible for B cell follicle formation. Tfh cells were also found to express the surface markers PD1, ICOS, CD40L, CD200 and BTLA and secrete IL-21 (Chtanova et al., 2004 and Johnston et al., 2009). As with the other Th subtypes, a "master" transcription factor, Bcl6, was identified to be important in their differentiation (Chtanova et al., 2004 and Nurieva et al., 2009). Moreover, Bcl6 overexpression with a retroviral vector was sufficient to polarize naïve TCR transgenic T cells into the Tfh lineage *in vivo* (Johnston et al. ,2009). Interestingly, Bcl6 shuts down all reciprocal Th programs by downregulating Blimp-1, which is associated with other Th programs, as well as other Th cytokines such as IFN-γ (Johnston et al., 2009).

Tfh cells differentiation is driven by TCR signaling as well as IL-6 or IL-21 (Crotty, 2014 and Nurieva et al., 2009). ICOS signaling is important for Tfh development as well, as loss of ICOS inhibitors, Roquin1 and Roquin2 results in spontaneous Tfh development (Pratama et al.,

2013). As is the case with the other Th subsets, cytokines that drive other subsets, such as IL-2, are inhibitory to Tfh (Johnston et al., 2012). These factors elicit Bcl6 expression, which causes the transcription of Cxcr5 and the downregulation of Ccr7. These changes in chemokine receptors allow Tfhs to migrate to the T-B zone border in the germinal center (Crotty, 2014). There, Tfh cells closely interact with follicular B cells which present antigen to them, and drive the somatic hypermutation and class switching of follicular B cells by secreting CXCL-13, IL-21 and IL-4. Interestingly, Reinhardt et al. (2009) showed that Tfh cells and not Th2 cells are, in fact, the major source of IL-4 in lymphoid tissues. IL-4 in Th2 cells and Tfh cells is regulated differently; whereas GATA3 controls its expression in Th2 cells, SAP and PKCθ control it in Tfh cells (Yusuf et al., 2010). Thus, the T helper subsets have been revised yet again to include Tfh cells. Since then, Th1/Th2/Th17/Tfh became the dominant T helper subsets dogma in the current decade (Fig. 1.2).

1.2.4 Th22 and Th9 cells

Th22 and Th9 cells (secreting IL-22 and IL-9, respectively) have been previously described but their existence as a bona fide T helper cell subtype rather than a cell state is more controversial than the "classic" Th subsets. Th22 cells have been first described in human skin memory CD4 T cells (Duhen et al., 2009). Even though the pro-inflammatory IL-22 was previously thought to be a signature cytokine of the Th17 lineage (Liang et al., 2006), Duhen et al. (2009) found many cells that expressed *Il22* but did not express *Il17a* and *Rorc* or any other signature cytokines of other lineages such as IFN-γ. Thus Th22 cells were born. Th22 cells can be differentiated in vivo by adding IL-6 to the culture, while addition of TGF-β to IL-6 polarizes them towards the Th17 phenotype (Zheng et al., 2007).

Figure 1.2

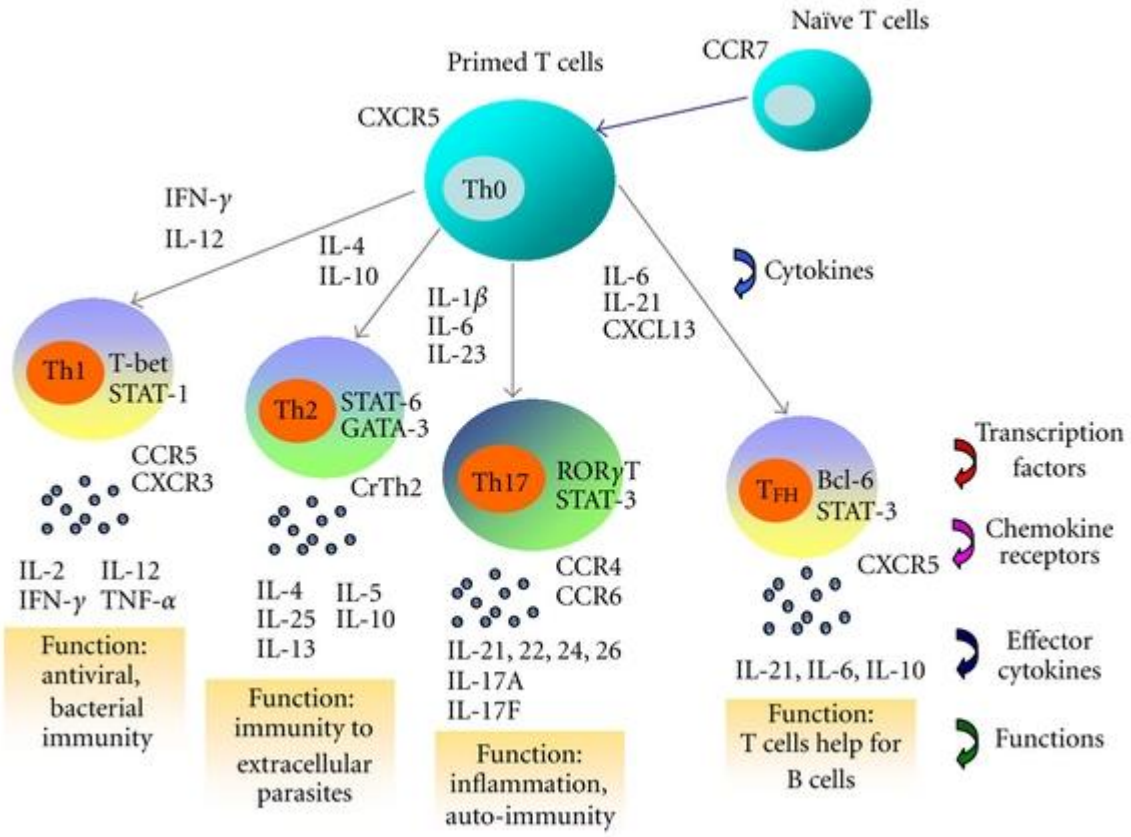


Figure 1.2: The current model of Th subsets (Adapted from Yu et al. 2012)

IL-22 acts mainly on nonhematopoietic tissue cells at barrier surfaces, where it triggers the production of antimicrobial peptides and enhances cell differentiation and survival (Kara et al., 2014). Recently, it has even been shown to protect intestinal stem cells against genotoxic stress (Gronke et al., 2019). Therefore, IL-22 plays an important role in control of bacterial infections such as *K. pneumoniae* and *C. rodentium* (Aujila et al., 2008 and Zheng et al., 2008).

However, many reports suggest that in mice IL-22 is restricted to the Th17 lineage *in vivo* (Kara et al., 2014). Furthermore, even though Aryl-hydrocarbon receptor (Ahr) regulates production of IL-22 from Th22, Ahr is also highly expressed and is functionally important in Th17 cells (Veldhoen et al., 2008). Lastly, Th22 cells share many functions such as induction of anti-microbial peptides and mucus with other Th17-derived cytokines. Recently, it was reported using dual IL-22/IL-17A reporter mice that some Th22 cells (that were differentiated in a novel assay) have never expressed IL-17A, suggesting that these are, indeed, separate subsets (Plank et al., 2017). Thus, it is unclear if Th22 are a real Th subset or just a state of Th17 cells.

Th9, as the name suggests, produce IL-9, one of the most understudied cytokines in T cells. At first, IL-9 was discovered in the context of Th2-mediated responses, as it expanded Th2 cells and protected mice against helminth infections (Gessner et al., 1993 and Faulkner et al., 1998). However, Vijay Kuchroo's and Brigitta Stockinger's group both found that *in vitro* culture of T cells with IL-4 and TGF-b produces T cells that secrete IL-9 and IL-10 (an anti-inflammatory cytokine) (Dardalhon et al., 2008 and Veldhoen et al., 2008). These cells were able to induce colitis and neuritis upon transfer, suggesting an inflammatory phenotype despite the production of IL-10. Licona-Limon et al. (2013) showed that IL-9 producing CD4 T cells protect against helminth infection and caused more potent responses than Th2 that led to helminth expulsion, and increase in mast cells and basophils. Th9 cells require the transcription factor PU.1 (Chang et al., 2010). There is a debate in the field whether Th9 cells are a bona fide Th subset or an auxiliary arm of the Th2 response (Kara et al., 2014). Recently, Micosse' et al (2019) have reported that in humans IL-9 producing cells are a subpopulation of Th2 cells that

transiently upregulate the transcription factor peroxisome proliferator-activated receptor-γ (PPAR- γ).

1.3 CD4 T cells classification according to their antigen experience and homing capacity

At around the same time as the Th1/Th2 dogma started developing, several labs have also found that CD4 T cells can also be classified not only by their cytokine production potential, but also according to their antigen experience, responsiveness to restimulation, as well as homing capacity. Memory T cells are generated from a small subset of effector cells after infection resolution, and are ready to respond to their cognate antigen upon secondary encounter. Whereas naïve T cells usually require TCR stimulation as well as co-stimulatory signals to get activated, memory T cells are activated and can produce cytokines rapidly upon antigenic restimulation (Mackay et al., 1990). Additionally, memory T cells express more adhesion molecules such as CD2, CD44 and LFA-1 that facilitate cell-cell interactions and migration (Sanders et al., 1988 and MacDonald et al. 1990). In humans, memory T cells express an isoform of CD45, CD45RO, whereas naïve cells express CD45RA (Akbar et al., 1988).

A further subdivision of memory cells is into central memory and effector memory. In humans, CCR7, a chemokine receptor that directs T cells to lymph nodes, divides memory cells into effector memory (CCR7-) and central memory (CCR7+) (Sallusto et. al, 1999). T effector memory cells maintain their effector status and can respond very rapidly to pathogens, whereas T central memory cells reside in the secondary lymphoid organs and can only acquire effector activity slowly upon restimulation.

An additional subset of cells is the tissue resident memory population (T_{rm}), which unlike T central memory or effector memory cells, do not recirculate but stay in the non-lymphoid tissues (Schenkel et al., 2014). It is thought that they arise from circulating naïve cells that encounter an antigen in the secondary lymphoid organs and downregulate the transcription factor kruppel-like factor 2 (KLF2), a positive regulator of sphingosinge-1 phosphate receptor 1 (S1PR1). S1PR1 is a recirculation cue and its downregulation prevents T cells from going to the blood stream (Schwab et al., 2007). The naïve cells thus acquire an effector phenotype and can migrate to the site of infection. Effector cells then acquire the C-type lectin CD69 as well as the integrin CD103 in the tissues. After the infection clearance a small number of memory cells remain seeded in the tissue long-term to provide memory in case of re-infection (Schenkel et al.,2014). Parabiosis studies showed that these cells do not recirculate and are not in equilibrium with circulating lymphocytes (Iijima et al., 2014).

1.4 CD4 T cells in the gut and their role in infection

The intestine is one of the largest barrier surfaces in the body, harboring 10^{14} bacteria, fungi, viruses and encountering dietary and other ingested antigens (Shale et al., 2014). Many immune cells, including a large population of CD4 T cells, reside there and are constantly challenged with antigens. It is their function to maintain the immune homeostasis in the intestine. CD4 T cells are important in responding to enteric infections, but can also be responsible for inflammatory bowel disease (IBD) if they react towards self-antigens or symbiotic bacteria.

T cells in the intestine can be classified according to their localization; The first group, intraepithelial lymphocytes (IELs) are located between enterocytes and thus are in immediate proximity to antigens in the intestinal lumen (Ma et al., 2019). IELs consist mostly of γδ T cells and CD8+ (especially CD8αα) rather than CD4+ T cells. In contrast, the second group resides in the lamina propria (LP), the connective tissue beneath the epithelium. Most of the T cells there are CD4 T cells, and they will be the focus of this study.

Naïve T cells migrate to the gut-associated lymphoid tissues (GALT) such as the mesenteric lymph nodes or Peyer's patches through the circulation and are then primed by APCs. Upon priming T cells acquire the ability to migrate to the LP through upregulation of adhesion molecules such as Integrin α4β7 (LPAM-1) or α4β1 (VLA-1) as well as various chemokine receptors, most notably CCR9 for the small intestine (Masopust et al., 2010 and Kunkel et al., 2000).

T effector cells further differentiate in the gut upon encounter with antigens, especially microbial antigens, evidenced by the fact that Germ-Free (GF) mice have fewer IFN-γ and IL-17 producing cells in their intestines (Niess et al., 2008). Intestinal T cells in homeostasis tend to be determined by the composition and properties of the microbiome. in mice Th17 differentiation was found to be mainly driven by adhesion of microbes, such as segmented filamentous bacteria (SFB) and others, to the intestinal epithelial walls (Ivanov et al., 2008 and Atarashi et al., 2015). Th1 cells can be induced by intracellular bacteria such as strains of *Klebsiella* (Atarashi et al., 2017). Th2 cells tend to be induced upon encounter with parasites such as helminths, as mice from colonies without parasitic infection tend to have very few Th2 cells (Maynard et al., 2009).

There are many excellent mouse T cell models for gut infection. *Salmonella enterica* serovar *typhimurium* is a good model for Type I immunity. Usually, wild-type mice succumb to *Salmonella typhimurium* infection, so attenuated strains such as AroAΔ are used (Hoiseth et al., 1981). This strain is auxotrophic for aromatic amino acids, which are not found in the host, and thus does not replicate as rapidly as the wild-type Salmonella. However, TCR αβ deficient and MHC class-II deficient (which do not harbor CD4 T cells) but not MHC class-I deficient mice (which do not harbor CD8 T cells) still succumb to infection, suggesting that CD4 T cells are crucial for control of infection (Hess et al., 1996 and Ravindran et al., 2005). Specifically, Salmonella infection resolution is dependent on CD4 T cells secreting IFN-γ (Hess et al, 1996).

Citrobacter rodentium infection is a useful model to study the role of Th17s in enteric infection. *Citrobacter rodentium* infection is similar to enteropathogenic *Escherichia coli* infection in humans. Infection is usually resolved in wild-type mice within 21 days of infection, with the T cell response peaking at days 12-14 (Mundy et al., 2005). As with *Salmonella* infection, CD4 T cells are the major players in infection clearance, as $CD4^{-/-}$ but not $CD8^{-/-}$ cannot clear infection (Simmons et al., 2003). That said, other cells such as B cells, ILCs and γδ T cells play important roles in infection resolution (Collins et al., 2014). *Citrobacter rodentium* elicits a strong Th17 response one week after infection, leading to increased epithelial barrier integrity via the cytokines IL-17A, IL-17F and IL-22 (Mangan et al., 2006 and Zheng et al., 2008). Although *Citrobacter rodentium* elicits a strong Type III response, Type I response plays a role in infection resolution as well, as IFN-γ-null mice have a compromised ability to limit the systemic dissemination of the bacteria (Simmons et al., 2002).

Nippostrongylus brasiliensis and *Heligmosomoides polygyrus* are helminths that serve as useful models for studying Type II responses in the gut (Camberis et al. ,2003). *Nippostrongylus brasiliensis* has a very convoluted life cycle. It infects the host cutaneously, from where the larvae migrate to the blood, and then settle in the lung. The larvae molt in the lung and migrate through the trachea to the intestine, where they molt again and reach adulthood. *Nippostrongylus brasiliensis* elicits a very strong intestinal Th2 response 7 to 13 days post infection, and via IL-4 induces IgE production 15-fold (Le Gros et al., 1996). IL-13 has been shown to be important for worm explusion as well (Urban et al., 1998). Conversely, *Heligmosomoides polygyrus* infects mice by direct ingestion, but just like *N. brasiliensis* it is a strong Th2 inducer and its explusion is dependent on CD4 T cells production of IL-4 and IL-13 (Urban et al., 1991).

1.5 Th heterogeneity and plasticity

The notion that Th states are terminally differentiated lineages was based on many observations listed above. Long term *in vitro* polarized Th1 and Th2 cells did not reverse their phenotype when restimulated with IL-4 or IL-12, respectively (although some conversion was noticed from a shorter, 1 week stimulation) (Murphy et al., 1996). Even though early studies have postulated mutually exclusive Th1, Th2, Th17 and Tfh programs, and this is still largely the current dogma, some evidence to the contrary emerged. Already in the early days of the Th1/Th2 model, Kelso et al. (1988) reported several IL-4 producing T cell clones that were also producing gamma interferon. However, they were considered either an artifact or a precursor to Th1/Th2 (in a similar way to CD4+CD8+ double positive cells in the thymus that give rise to

CD4+ or CD8+ cells). Thus, Firestein et al. (1989) proposed that in addition to Th1s and Th2s there should be another CD4+ T cell subset with unrestricted pattern of cytokine production, which they called Th0. Heterogeneity was also observed on a single cell level by using immunofluorescence microscopy, when some cells stained for both IL-4 and IFN-γ even after IL-12 was added (Openshaw et al., 1995). Very soon after Th17 cells were characterized, some of them were found to express Th1 cytokines or markers (Zielinski et al. ,2014 and Duhen et al., 2014). In humans, Th17 cells were shown to produce both IL-17A and IL-4 (Cosmi et al., 2010).

Plasticity, the ability to go from one cell state to another, between Th cell types has been shown as well. One of the first papers describing it in T cells showed that influenza hemagglutinin (HA)-specific memory CD4 T cells could mount either Th1 or Th2 responses depending on the recall stimulus (Ahmadzadeh et al., 2002). More recently, Hegazy et al. (2010) showed that lymphocytic choriomeningitis virus (LCMV) infection could reprogram committed Th2 cells to adopt a stable double positive GATA3+T-bet+ IL-4+ IFN- γ+ phenotype. Ex vivo polarized Th1 and Th17 cells can be converted into IL-4-producing Th2 cells in vivo by infection of mice with *Nippostrongylus brasiliensis*. Th17 cells tend to have even more plasticity than Th1 or Th2 cells. For example, fate mapping experiments showed that IL17A-producing cells in chronic infection such as EAE switch to alternative cytokines such as IFN-γ and TNF-α at a late stage of infection (Hirota et al., 2011). Another group has shown that IFN-γ and IL-12 can synergize to convert *in vivo* generated Th17 cells into dual expressors of T-bet and RORγt (Lexberg et al., 2010). In humans, Duhen et al. (2014) identified $CCR6^{+}CXCR3^{+}$ Th1/17 cells pathogenic cells that are driven by IL-1β and IL-12. This plasticity of Th17 cells was attributed to

epigenetic instability of cis-regulatory elements in key cytokine and transcription factor loci (Mukasa et al., 2010).

Th17 and Tregs can also convert to each other as their differentiation pathways are related. iTregs that are stimulated with TGF-β together with IL-6 alone or with IL-1 and IL-23 can downregulate FOXP3 and increase production of IL-17 (Yang et al., 2008). Th17s also have the ability to transdifferentiate into IL-10 producing T regulatory type I (Tr1) cells upon resolution of inflammation in a TGF-β1-dependent manner (Gagliani et al., 2015). They have also been shown to convert to Tfh cells in Peyer's patches, where they induced germinal centers formation and development of IgA responses (Hirota et al., 2013). Interestingly, Th17-deficient mice had impaired antigen-specific IgA following immunization. Conversely, Tfh cells have been shown to express some T-bet and IFN-γ in the context of viral infection (Johnston et al., 2009).

Kenneth Murphy and Brigitta Stockinger have proposed a model which can partially explain plasticity (Murphy and Stockinger, 2010). In their model, unstable T helper states such as iTregs can become more stable upon stimulation with certain cytokines. This model can explain why iTregs can become Th17 and Th17s can become Th1s, but not the Th1-Th2 double positive cells (Hegazy et al., 2010). Taken together, these data suggest that perhaps Th subsets are plastic and are not truly different cell types, but rather transient states of cells responding to antigen stimulation via certain cytokines.

1.6 Transcriptional profiling of different Th subsets

With the advent of Affymetrix gene arrays it became possible to perform full transcriptional profiling of various CD4 T cell subsets and compare them transcriptionally. The

first study to use this powerful technology in CD4 T cells compared resting versus superantigen-activated CD4 T cells *in vivo* (Teague et al., 1999). They found that T cell activation did not change the diversity of expression (number of expressed genes), but rather that the expression of only less than 300 genes was altered (143 increased and 139 decreased 8 hours after activation). Amongst the genes that increased upon activation were *Cxcr5, Tnfrsf4* (OX40), *Ctla4, Ly6e*, *Gzma* and *Gzmb*, whereas the cytokine receptors *Il4ra,Il7r,Il6,* the Kruppel-like transcription factors *Klf2, Klf3, Klf4*, the AP1 complex components *Jun, Junb and Fosb* , as well as other genes such as *, Tsc22d1, Zfp36* and *Itgae* all decreased upon activation. Interestingly, cell cycle genes as well as genes encoding calcium binding proteins such as S100 proteins and Calmodulin were upregulated only 48 hours rather than 8 hours after activation.

The first paper to attempt to transcriptionally profile Th subsets used an Affymetrix oligonucleotide array and compared *in vitro* polarized human Th1 and Th2 (the only Th subsets known at the time) (Rogge et al., 2000). They found 215 genes that were differentially expressed with a fold change of at least 2. As expected, *IFNG*, *IL12RB2* (coding for IL12 receptor) and *IL18R1* (coding for IL18 receptor) were all over-expressed in Th1 cells (though not TBX21 (coding for T-BET), as perhaps the probe was not included in the array at the time). They found that Th1 cells were enriched for cytotoxic and apoptosis-inducing genes such as *PRF1* (perforin), *GZMB* (granzyme B) and TNFSF10 (TRAIL), as well as the chemokines *CCL3*, *CCL4*. However, this study was skewed towards Th1 signature, as most of the differentially expressed genes that were identified were overexpressed in Th1s. Th2 genes were not well represented, as even though *GATA3* was overexpressed in Th2 cells, *IL4*, *IL5* and *IL13* were not. Chtanova et al. (2001) utilized a similar microarray approach to compare in vitro polarized murine Th1 and

Th2 cells but were able to obtain more highly polarized murine Th2 cells and thus able to identify more Th2 signature genes (overall around 100 differentially expressed). Notably, both of these papers identified *Itgb7* (a component of both LPAM-1 and CD103, both important in homing to intestinal sites) as upregulated in Th2 cells, though *Itgα4* (its partner in LPAM-1) was found to be upregulated in Th1 cells. Overall, many of the differentially expressed genes from these two studies did not agree with each other, probably because conditions *in vitro* were variable and the populations were not sorted with the same strategy (in fact, both groups used negative selection to isolate CD4 T cells, which likely introduced some contamination). It is also possible that murine and human cells express different genes in *in vitro* cultures.

The Flavell group subsequently published a kinetic study of *in vitro* Th1 and Th2 differentiation 1,2,3 or 4 days after differentiation (Lu et al., 2004). They found around 300 differentially expressed genes: chemokine receptors, such as *Ccr5* in Th1 cells and *Ccr4* and *Ccr8* in Th2s. Th1s also expressed *Ccl3, Ccl4, Ccl5, Lta* and *Ccl9*, whereas Th2 cells expressed higher *Ccl1* and *Cxcl2*. As previously mentioned, the Littman and Kuchroo labs transcriptionally profiled Th17 cells in an exhaustive manner and built a transcriptional network of Th17 differentiation (Ciofani et al., 2012 and Yosef et al., 2012).

Single cell technology RNA-seq enabled immunologists to look at transcription at a single cell level, and some of the first cells to be profiled by this technology were, indeed T cells. Gaublomme et al. (2015) profiled 976 *in vitro* differentiated Th17 cells or *in vivo* EAE Th17s from the peak of disease. They classified five different subpopulations of Th17s from the CNS and draining lymph nodes of EAE mice: self-renewing, Th17/pre-Th1 effector, Th17/Th1-like effector, Th17/Th1-like memory and Th17 dysfunctional/senescent. *In vitro* generated Th17s

showed varying degrees of pathogenicity depending on their polarization conditions. Additionally, they identified the genes *Toso*, *Gpr65* and *Plzp* to be implicated in Th17 pathogenicity. Importantly, they established that these "Th17 cell states" were on a continuum, rather than discrete states. This study was the first single cell RNA-seq analysis of T cells, but with a few caveats, as the five different *in vivo* Th17 clusters were based on limited signatures and were not actually experimentally shown to be separate subsets.

Another Single cell RNA-Seq study was conducted on Th2 cells from mesenteric and mediastinal lymph nodes of *Nippostrongylus brasiliensis* infected mice (Proserpio et al., 2016). Their conclusions were that differentiated, cytokine-producing cells cycle faster than early activated cells, and thus activation causes acceleration of proliferation and differentiation. Last year, a large single-cell study of a novel subset of human CD4 cytotoxic lymphocyte has been published (Patil et al.,2018). This subset was found to express many cytotoxic genes such as *GZMB*, *PRF1* and *NKG7*.

The notion that T helper cells fall into discrete states has prevailed for over 30 years. However, as I show in this introduction, this notion has some flaws. First, there are many reports describing cells that non-exclusively produce different "Th signature" transcription factors and cytokines. Second, plasticity between different Th subsets was demonstrated almost across all subtypes. Third, most high-throughput transcriptional studies to phenotype Th subsets were largely done *in vitro*, and perhaps did not capture al the states Th cells can acquire *in vivo*. To date, no study has directly compared T effector responses to different "Th"-inducing pathogens *in vivo*. Thus, we attempted to look at the phenotypic states of effector T cells at steady state and during infection at single-cell resolution.

Chapter 2: Cloud model of $CD4^+$ T effector cell phenotypes

Cloud model of CD4$^+$ T effector cell phenotypes

Evgeny Kiner[1,2], Elijah Willie[3], Brinda Vijaykumar[1], Jodie Chandler[4], Graham LeGros[4], Sara Mostafavi[5], Diane Mathis[1,2] and Christophe Benoist[1,2]

[1]Department of Immunology, Harvard Medical School; [2]Evergrande Center for Immunologic Diseases, Harvard Medical School and Brigham and Women's Hospital, Boston, USA; [3]Bioinformatics Program, University of British Columbia, Vancouver, Canada; [4]Malaghan Institute of Medical Research, Wellington, New Zealand; [5]Departments of Statistics and Medical Genetics, University of British Columbia, Vancouver, Canada; Canadian Institute for Advanced Research, Toronto, Canada; Vector Institute, Toronto, Canada.

***Address correspondence to:**Diane Mathis and Christophe Benoist
Department of Immunology
Harvard Medical School
77 Avenue Louis Pasteur, Boston, MA 02115
E-mail: cbdm@hms.harvard.edu
Phone: (617) 432-7741
Fax: (617) 432-7744

Running title: Thelper cell heterogeneity

2.1. Abstract

$CD4^+$ T effector cells (Teffs) are traditionally divided into several subsets that are defined by the cytokines they produce. We attempted to look at their heterogeneity in an unbiased manner by using scRNA-seq. Surprisingly, Teffs, with the exception of Th2 and Tfh, could not be subdivided into parsable clusters, even when challenged with bacteria or helminths that induce strong Teff responses. Th signature transcripts were continuously rather than discretely distributed amongst the cells. We verified these conclusions with robust statistical tools that showed non-discrete distributions of cells. Moreover, the infection type defines the Teff cells more than the cytokine they produce, as IL17-and IFNg-producing cells were transcriptionally more similar to one another than to cells producing the same cytokine from different infections. We identified 16 gene modules that drive the heterogeneity within Teff cells in steady state and upon infection. Lastly, we validated the scRNA-seq data by sorting cells from different “poles” of the continuum and subjecting them to transcriptional and functional analyses. Our data suggest that Th programs are not the major drivers of Teff heterogeneity, and that Teff cells are not discrete, but rather continuous states. Thus, we propose with a model of Th sates that is defined by quantum rather than discrete states.

2.2 Introduction

CD4 effector T cells (Teff) are central regulators of both humoral and cellular immune responses, orchestrating adaptive (antibodies, cytotoxic cells) and innate (macrophage) immune responses to different pathogens, as well as repair of parenchymal tissues. This diversity of functions is primarily rooted in the different cytokines produced by Teff cells. This extraordinary swath of functions has always raised the issue of functional diversity, which was documented by functional assays even prior to molecular identification of MHC molecules, the TCR, or CD4 itself (Bottomly et al., 1986 and Mosmann et al., 1989). The key advance was the demonstration that functional phenotypes in different T cell clones were keyed to the cytokines they produce (Mosmann et al., 1986 and Killar et al., 1987), and the coining of the Th1/Th2 nomenclature. Th1s secrete IFN-γ and mainly support inflammatory and cytotoxic responses; Th2 produce IL4/5/13 and principally help B cells produce antibodies. This division has since been revised multiple times to add more subsets such as IL17-A secreting Th17s, IL9 secreting Th9s or follicular helpers (Tfh; Cua et al., 2003, Kopf et al., 2005, Dardalhon et al, 2008 and Duhen et al., 2009). But the notion that T helper cells belong to discrete and largely stable states that are defined by the cytokines they produce has perdured (Abbas et al. ,1996 and O'Shea et al., 2010). Different types of infectious or allergic challenges preferentially elicit different Th types (Th1 associated with intracellular pathogens; Th2 with helminth parasites, Th17 with bacterial and fungal infections) which also have differential implications in immune-mediated diseases (Zhu et al., 2008).

However, several observations have put this model into question. First, many reports have documented that secretion of IFN-γ, IL-4 or IL-17 is not always mutually exclusive

(reviewed in Geginat et al., 2014 and Zhou et al., 2009). Second, many groups have shown plasticity between Th subtypes, suggesting that these cell states are not as stable or as terminally differentiated as originally inferred from Th clones grown in supra-physiological concentrations of inducer cytokines (O'Shea et al., 2010). In addition, while some cell surface markers could be used for identification of differentiated Th types, these were never completely exclusive, meaning that the formal isolation of an *in vivo*-derived Th cell population remained elusive.

Here, we thought it of value to revisit the spectrum of phenotypic states that effector T cells can adopt, in the unbiased frame that single-cell RNAseq (scRNAseq) can allow (Stubbington et al., 2017). We analysed T cells in the colon lamina propria (LP), a front-line tissue constantly influenced by commensal microbes, and used different infectious agents known to elicit distinctive Teff responses. In such a context, would it be possible to identify discrete Th1, 2 or 17 populations? The results, to the contrary, indicate that Il17- or IFN-γ-producing cells belong to a broad continuum of cell states, one primarily shaped by the infectious context rather than by membership in any Th class.

2.3. Results

A continuum of effector phenotypes in colonic CD4+ effector cells

To probe the transcriptional landscape of CD4 T effector cells in an unbiased manner, we performed single cell RNA sequencing on total CD4+ T cells from the colon LP of conventionally-housed (SPF) C57BL/6 mice. Overall, two independent experiments were performed with droplet-based technique (10X-Genomics), yielding 1782 and 3360 cells passing QC, (Fig. S1A; Table S1A; here and below, datasets were analyzed individually, replicates serving

to confirm conclusions). It was straightforward to parse, with KNN clustering and in the t-SNE projection, cells from these datasets into four main groups of CD4+ T cells (Fig. 2.1A). Regulatory T cells (Treg) and their Rorc+ and Ikzf2+ (Helios) subsets were identified by expression of *Foxp3*, *Il2ra* and *Ikzf2*); among T conventional cells (Tconv), naïve T cells (*Sell*- and *Ccr7*-positive) and Teff cells (*Cd44*+, *Sell*-negative) were present in largely equal proportions. These four groups matched those expected from flow cytometry of the input pool. A similar dataset generated from germ-free mice showed a strong reduction of the Teff component (Fig. S1B), indicating that Teff cells observed in conventional mice were elicited by continuous challenge from the commensal microbiota. Given this project's goals, we focused on and reclustered the Teff population alone (Fig. 2.1B). Here, with the exception of cycling cells, we could not identify in the t-SNE representation any clear partitions between cells, but rather a quasi-continuous cloud. To search for distinctions corresponding to the major recognized Th types, we generated a set of curated signatures, which included the key defining cytokines, driving transcription factors, and other Th-associated genes (Table S1B). These signatures did show biased distributions, albeit to varying degrees: *Ifng*, *Tbx21* and other genes classically associated with Th1 cells were spread widely, while equivalent representations of Th2 or Tfh classes showed more polarized expression (Fig. 2.1B).

Different intestinal infections elicit divergent Teff phenotypes

We hypothesized that under SPF conditions Teff cells are only partially polarized into characteristic subsets because of limited activation by commensal microbes, and that the single-cell data only show "stubs" of more differentiated states that the cells could potentially reach. We thus sought to further polarize the T cell pools to by challenging the mice with

Figure 2.1: scRNA-seq analysis of Teff shows continuous rather than discrete distribution

A) t-SNE representation of single cell data from total colonic LP CD4 T cells of SPF mouse. Right panel colors represent different KNN cell clusters. Lower panel: Expression of key genes.

B) t-SNE representation single cell data zoomed in on effector CD4 T cells. Right panel: overlay of curated Th signatures (Table S2) on the t-SNE.

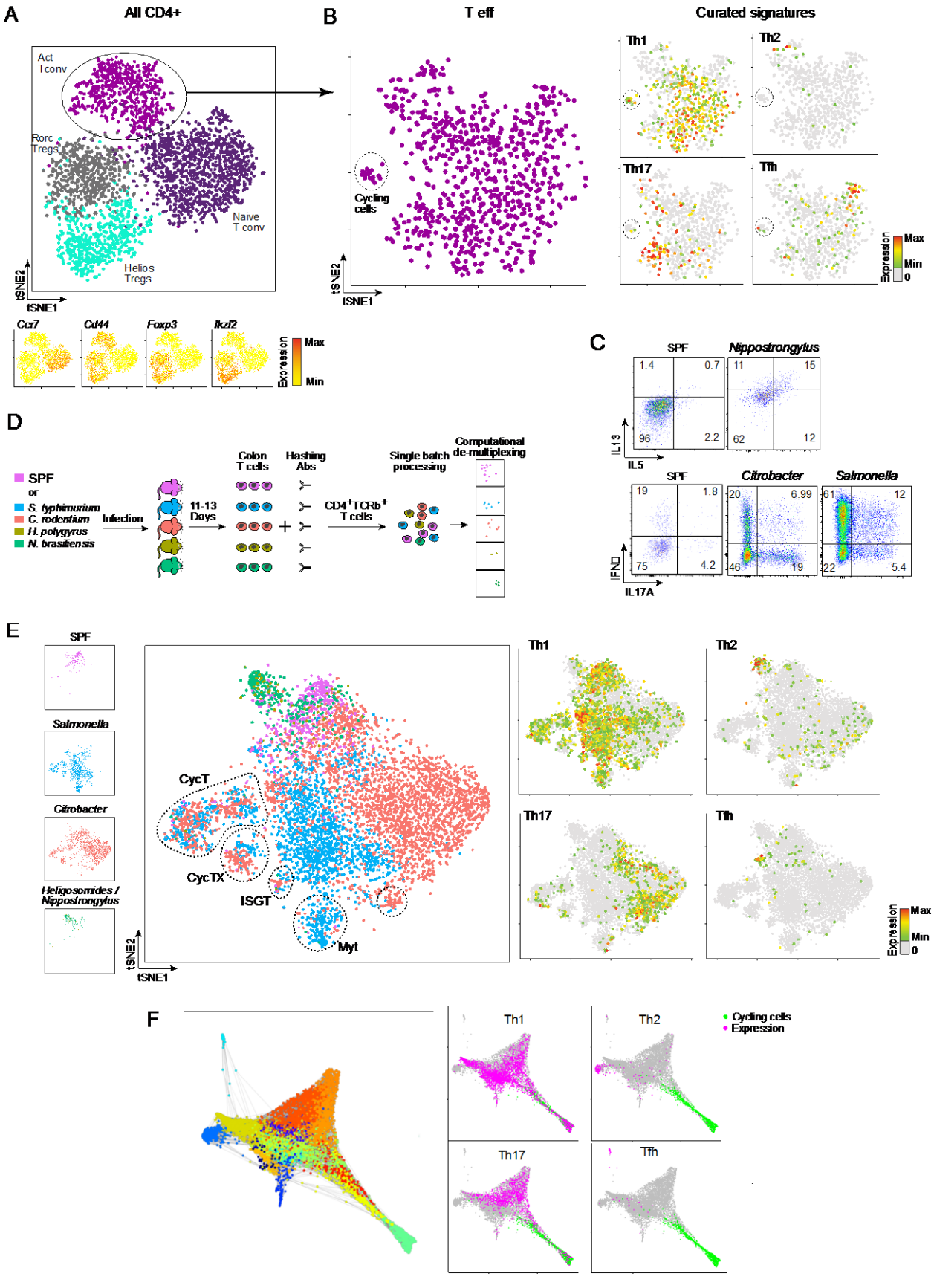
C) Flow cytometry confirmation that the infections elicited strong Th responses. All samples are gated on CD4+ TCRβ+ FOXP3- $CD44^{hi}$. Upper panel: Th2 response (IL5/IL13), lower panel: Th1/17 responses (IL17-A,IFN-γ)

D) Schematic of hashing experiment. Mice were infected with different pathogens, their colonic LP cells extracted, tagged with hashing antibodies, sorted as CD4 T cells and processed as a single batch on the 10X chromium controller. Sample demultiplexing was done computationally.

E) t-SNE representation of single cell data from hashing experiment. Central panel: each color represents a different condition/infection; outlier population are highlighted- CycT-cycling T cells, CycTx-cycling T cells with cytokeletal gene signature, ISGT- T cells expressing type I interferon sensing genes, MyT- T cells expressing myeloid transcripts. Left panel: For ease of visualization, each condition is plotted seperately. Right panel: overlay of curated Th signatures on the hashed t-SNE.

F) SPRING representation. Central panel: Louvain clustering; right panel: overlay of curated Th signatures and cycling cells on the SPRING plot.

Figure 2.1-Cont.



infectious agents known to elicit strong and biased intestinal responses, and infected mice with either (i) AroAΔ *Salmonella typhimurium*, a mutant that elicits IFN-γ-dominated responses in the gut (Hoiseth et al., 1981 and Hess et al., 1996}; (ii) *Citrobacter rodentium*, a strong Il17 inducer (Mangan et al, 2006 and Collins et al, 2014); (iii) *Heligmosomoides polygyrus* and *Nippostrongylus brasiliensis* , two helminths that provoke prototypic Type-2 responses (Urban et al., 1991, LeGros et al., 1996 and Camberis et al, 2003). Cells were harvested well into the infection (day 11+/2) to allow for terminally differentiated T cells to appear, and flow cytometry confirmed that the infections did induce the expected cytokine responses (Fig. 2.1C). As above, two independent experiments were performed for each (Fig. S1C). In one instance, CD4+ cells from each control or infected mouse were tagged with DNA-coded antibodies (Stoeckius et al., 2018) and pooled for sorting and microfluidic bead capture, which made for an ideal comparison of Teff cells, eschewing batch confounders (Fig. 2.1D). As in uninfected mice, CD4 T cells clustered into Tregs, naïve T cells (coming mostly from uninfected mice), and *Cd44*+ Teff cells (Fig. S2A).

When Teff cells were considered on their own (Fig. 2.1E), several strong conclusions emerged from the data. First, most Teff again formed a continuum, but cells from each infected mouse mostly clustered in largely non-overlapping fashion, indicating that the type of infection was the major driver of variability. Second, although there tended to be some segregation of cells expressing the Th1 or Th17 signature sets, we still could not detect cell clusters that were strictly associated with Th1 or Th17 signatures. *Ifng*, *Tbx21*, *Il17a* and *Rorc* were expressed along heterogenous and continuous swaths which encompassed cells from *Citrobacter*- or *Salmonella*-infected and SPF mice (Fig. S2B). Cells from *Nippostrongylus*- and *Heligmosomoides*-

infected mice did cluster together, along with Th2-like cells from SPF mice, suggesting a more specific identity for Il4/5/13-expressing T cells. These conclusions were generally also present in an independent set of colonic CD4+ Teff from mice infected with the same pathogens (Fig. S2C).

Similar conclusions were reached with another dimensionality reduction tool, SPRING, using Louvain clustering (Weinreb et al., 2017; Fig. 2.1F), and further supported by diffusion plots (Haghverdi et al, 2015; Fig. S3A).

Importantly, however, none of the partition clustering algorithms applied to these data were able to break up the cell cloud into zones that overlapped with expression of the Th signature sets. This was true of classic clustering tools, such as KNN clustering (Fig. S3B) or biclustering approaches recently developed to efficiently infer cell-types in single-cell data (Azizi et al., 2018 and Zeisel et al., 2015; Figs. S3C-D).

To account for the high dropout rates in scRNAseq data, we applied an imputation algorithm, but Th1 and Th17 still did not cluster distinctly (Fig. S2D). Furthermore, few genes that dominate the first principal components (PCs) of the datasets corresponded to “Th” associated signatures, implying that “Th” genes were not driving the heterogeneity in this dataset (Fig. S2E). Lastly, several small but distinct Teff subpopulations did emerge. One group corresponds to cycling cells (T-CY), which was expected in conditions of strong T cell stimulation, but others had unexpected characteristics which are detailed further in chapter 3.

The suggestion, from the t-SNE plots of Fig. 1, that Teff cells belong to a continuum of cell states rather than to well-defined Th categories, needed to be treated with caution. Like any dimensionality reduction, t-SNE distorts distances in the multidimensional space, and tends to inflate distances between tightly clustered objects. Clustering approaches are also

inadequate, as such algorithms will find clusters even in fully continuous distributions. To objectively verify the continuity in the distribution of transcriptomes of Teff cells, we used Hartigan's Dip Test of multi-modality (Hartigan et al., 1985) which tests for support of two or more clear sub-populations with clear boundaries based on the distribution of pairwise distances between cells. For robustness, we applied the Dip Test in two ways, by computing (a) pairwise Euclidean distances based on the top 350 most variable genes defined by Dispersion (Butler et al., 2018), or (b) differences in the first principal component (PC). With either approach, discontinuities were identified between entire CD4+ T cells of Fig. 1A, as expected, but not within any of the Teff populations whether from SPF or infected mice when cycling cells were removed (Fig. S3E).

Whither Th modules?

Their expression patterns within the projection plots of Fig. 2.1A indicated that prototypic Th1 or Th17 signature sets do not mark discrete sets of cells, which led us to ask whether they actually are discretely co-regulated entities within activated CD4+ Teff cells of the colon. Gene:gene correlation between individual cells in scRNA-seq data is a powerful means to identify coregulated genesets, as it leverages co-expressions across thousands of individual cells (Macosko et al., 2015 and Tanay et al., 2017). We used the entire dataset of Teff from baseline or infected mice, down-sampled to equivalent numbers of cells to avoid over-representation from any one input, calculated gene:gene correlations within each condition. Some cytokines did show significant positive clustering, as expected (*Il4* or *Il5* and *Il13*; *Il17a* and *Il17f*; Fig. 2.2A). However, no pairs of cytokines were significantly anti-correlated in these data, suggesting that cytokine expression is not strongly exclusive, in accordance with the many

Figure 2.2:

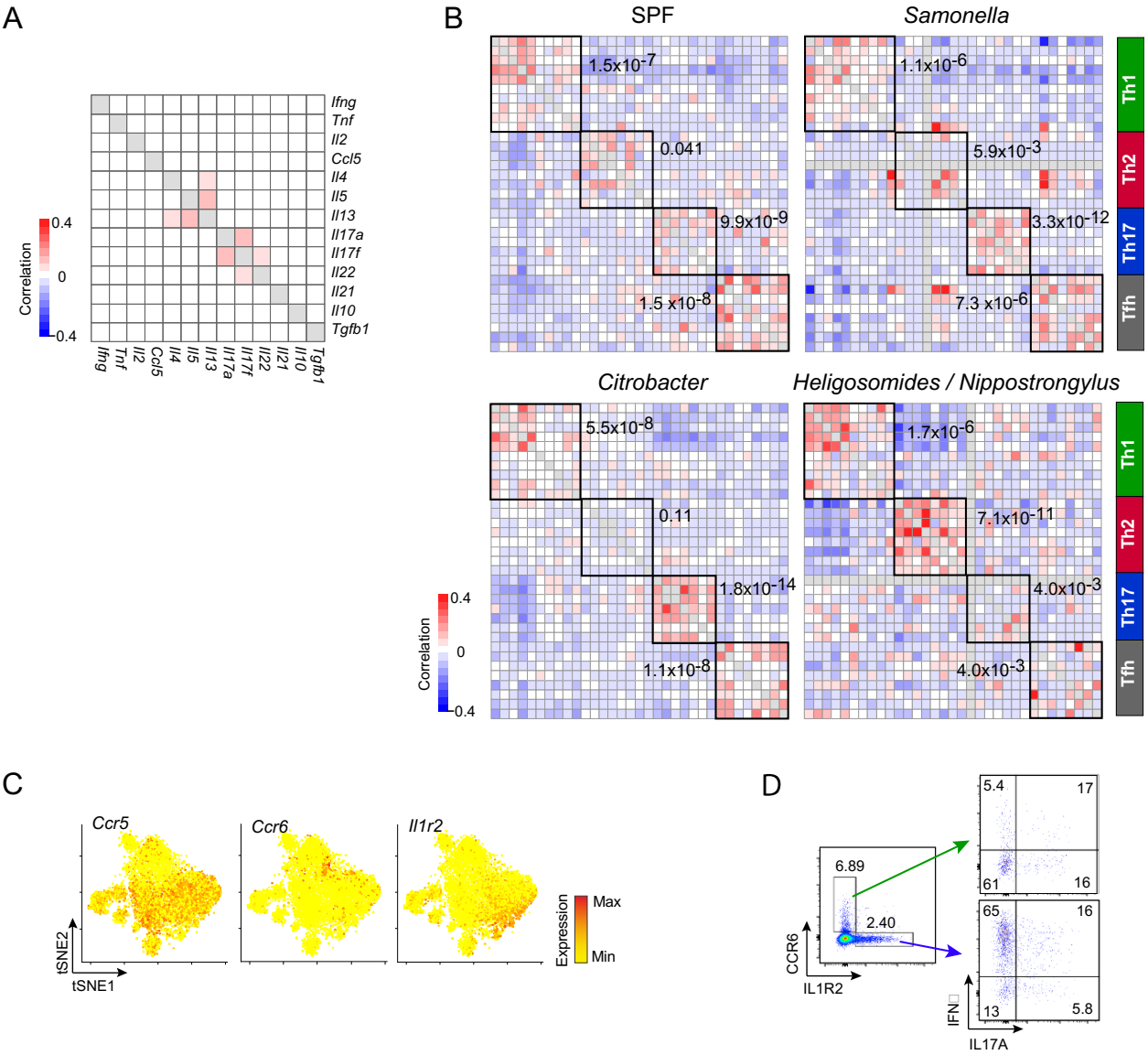


Fig. 2.2: Th modules are not clearly identifiable in the sc-RNAseq data.

A) Co-expression of key cytokines across all samples. Mean Pearson gene:gene correlation of cytokine genes across all samples.Only significantly correlated cytokines are colored($p<0.05$, χ^2 test).

B) Pearson gene:gene correlation plots of curated Th signatures in each single cell sample. Numbers indicate P values of "clustaribility" for each Th cluster compared to the rest by using Mann-Whitney test.

C) Single-cell expression of surface marker genes on the hashed t-SNE.

D) Flow cytometry plot showing CD4 Teff cells (CD4+ TCRβ+ FOXP3- $CD44^{hi}$) from Salmonella-infected mice.

observations of dual expression mentioned above. The signature genesets did not appear like tight gene modules overall (Fig. 2.2B), although each group did show significantly more within-signature correlation than to other signature genes (most Mann-Whitney p values < 10^{-3}). The Th-specific genesets expected to be most highly associated with a particular infection did become more correlated (e.g. the Th17 or Th2 clusters in cells from *Citrobacter* or *Heligmosomoides* infections, respectively), with the exception of the Th1 geneset, which became if anything looser after Salmonella infection. Overall, however, gene:gene correlation within single-cells did not highlight tightly coregulated clusters, in accordance with the absence of clearly identified clusters of cells expressing the conventional Th signatures.

Next, we wanted to see if surface markers that are associated with classic Th subsets and are used for cell sorting (*Ccr5* for Th1; *Ccr6* and *Il1r2* for Th17), show expression in line with their respective Th signatures. Surprisingly, *Ccr6* and *Il1r2* were expressed in a mutually exclusive manner and had only partial overlap with *Il17a* producing cells, whereas *Ccr5* was continuously distributed across the cloud with only partial overlap with the Th1 signature (Fig. 2.2C). Flow cytometric analysis confirmed that CCR6 and IL1R2 were largely exclusive of one another, identifying cells that produced both IFN-γ and Il-17 to different degrees (Fig. 2.2D). Thus, not only are classic Th subsets not clearly identifiable in the transcriptional data, but the markers conventionally used to identify them are poorly matched.

Teff phenotypes distinguished by infecting agents, not by Th type

The distribution of Teff cells in the pooled profiling experiment of Fig. 2.1E suggested that the infecting agent was the major driver of variability. This was also apparent when comparing the overall Euclidean distance between all *Il17a* and *Ifng* expressing cells from the

different conditions. Cells clustered according to the type of infection (or lack thereof), rather than by the cytokine type they expressed (Fig. 2.3A), Ifng or Il17a-expressing Teff from any condition being more similar to each other than to their counterparts in the other mice. Accordingly, analysis of variable transcripts present in *Il17a* and *Ifng*-expressing cells revealed clearly divergent patterns, with blocs of co-expressed transcripts that largely aligned with the infection (Fig. 2.3B). Interestingly, though, a small minority of the Il17a- or Ifng-producing Teff from *Citrobacter*-infected mice did express the transcripts most prevalent after *Salmonella* infection (small arrows in Fig. 2.3B), in accordance with the partial overlap observed in Fig. 2.1E).

To verify this result and exclude technical pitfalls from single-cell RNAseq, we used an Il17A-GFP reporter mouse line and performed conventional low-input RNA-Seq on colonic GFP positive cells at baseline or infection with *Salmonella* or *Citrobacter*, in biological duplicates. Echoing the single cell data, Principal Component Analysis showed that IL17A+ cells from each condition clustered separately from each other (Fig. 2.3C). The direct comparison of IL17A-GFP+ cells from *Salmonella* or *Citrobacter*-infected mice showed that the transcripts with biased representation in the scRNAseq data were similarly biased in the bulk data, for the most part ($p=2X10^{-4}$ for genes upregulated in *Salmonella*, but p=0.24 for genes upregulated in *Citrobacter*). It also revealed genesets differentially elicited by the two infections, with 237 differentially expressed genes (at FC>2 and FDR <0.05; Fig. 2.3D). Most strikingly, the infections elicited many differentially expressed genes in IL17A-GFP+ cells compared to uninfected mice (Citrobacter: 393 genes and Salmonella: 737 genes, with the same cutoffs as above; Fig. 2.3E). Conversely, activated ($CD44^{hi}$ $CXCR6^{hi}$) IL17A-GFP- cells sorted from Salmonella-

Fig. 2.3: Teff phenotypes are distinguishable by infection rather than by Th type

A) Dendogram representing hierarchical clustering of Euclidean distances between cells from single cell RNA data selected by *Ifng* or *Il17a* expression of at least 1 UMI.

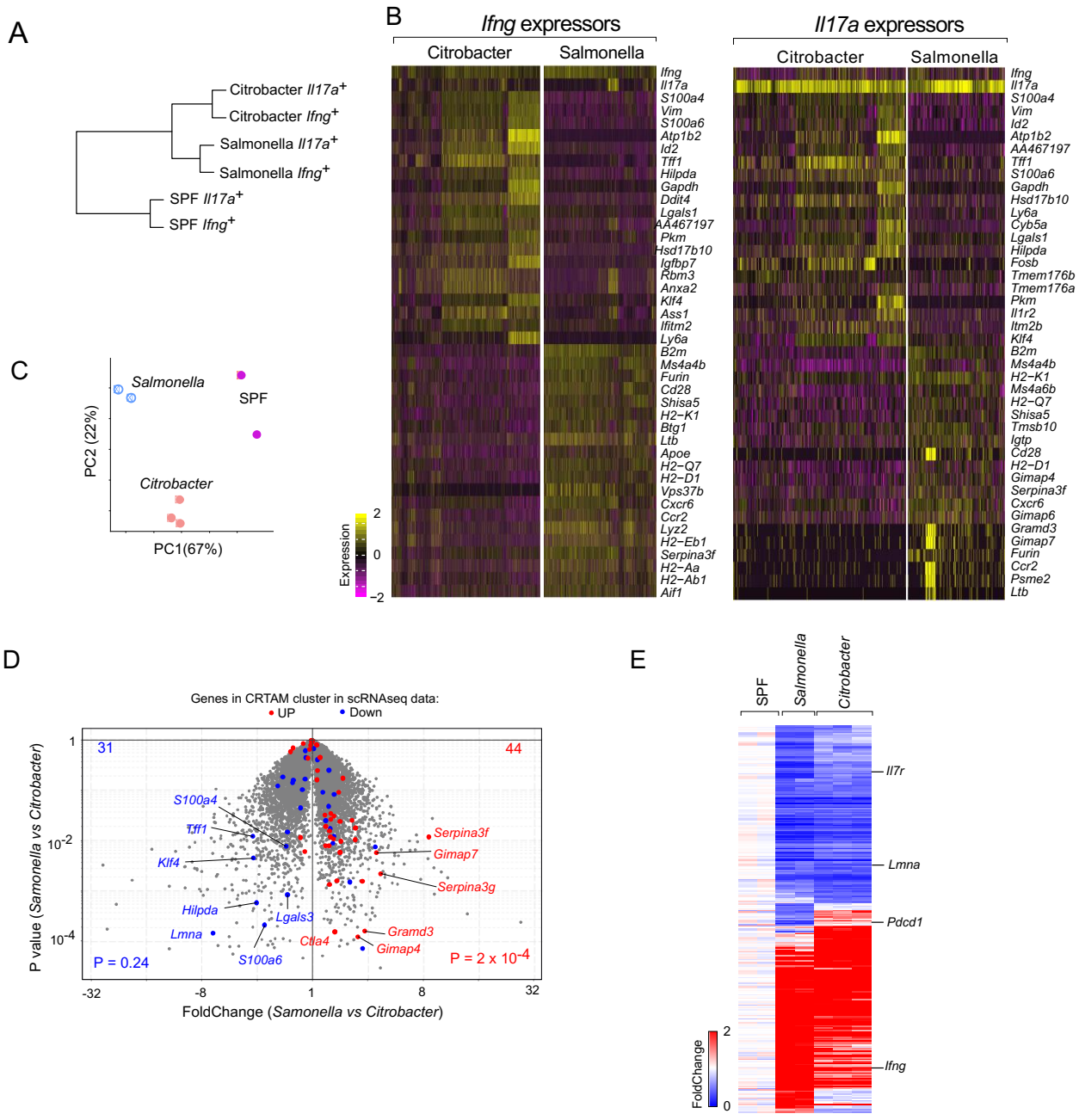
B) Heatmap of top 20 genes differentially expressed between CD4 Tcells from *Citrobacter* or *Salmonella* infected mice in either *Il17a* or *Ifng*-expressing cells. *Ifng* and *Il17a* are in the top two rows.

C) PCA of bulk RNA-sequencing data from IL17-GFP^{+} cells from mice infected with either *Citrobacter* or *Salmonella* as well SPF mice.

D) Volcano plot of bulk RNA-sequencing data from IL17-GFP^{+} cells from mice infected with either *Citrobacter* or *Salmonella.* Red and blue marked genes represent genes that are more highly expressed in CD4 T cells from *Salmonella* or *Citrobacter* infections, respectively.

E) Heatmap of all variable genes in IL17-GFP^{+} cells from *Salmonella*, *Citrobacter* or SPF mice, normalized to mean SPF values, and hierarchically clustered.

Figure 2.3


A
Citrobacter Il17a+
Citrobacter Ifng+
Salmonella Il17a+
Salmonella Ifng+
SPF Il17a+
SPF Ifng+
B
Ifng expressors
Il17a expressors
Citrobacter
Salmonella
Ifng
Il17a
S100a4
Vim
S100a6
Atp1b2
Id2
Tff1
Hilpda
Gapdh
Ddit4
Lgals1
AA467197
Pkm
Hsd17b10
Igfbp7
Rbm3
Anxa2
Klf4
Ass1
Ifitm2
Ly6a
B2m
Ms4a4b
Furin
Cd28
Shisa5
H2−K1
Btg1
Ltb
Apoe
H2−Q7
H2−D1
Vps37b
Cxcr6
Ccr2
Lyz2
H2−Eb1
Serpina3f
H2−Aa
H2−Ab1
Aif1
Expression
2
0
−2
Cyb5a
Fosb
Tmem176b
Tmem176a
Il1r2
Itm2b
Tmsb10
Igtp
Gimap4
Gimap6
Gramd3
Gimap7
Psme2
C
Salmonella
SPF
Citrobacter
PC2 (22%)
PC1(67%)
D
Genes in CRTAM cluster in scRNAseq data:
UP
Down
31
44
P = 0.24
P = 2 x 10-4
P value (Samonella vs Citrobacter)
FoldChange (Samonella vs Citrobacter)
1
10-2
10-4
-32
-8
1
8
32
S100a4
Tff1
Klf4
Hilpda
Lgals3
Lmna
S100a6
Cita4
Gramd3
Gimap4
Serpina3f
Gimap7
Serpina3g
E
SPF
Salmonella
Citrobacter
Il7r
Lmna
Pdcd1
Ifng
FoldChange
2
0

infected mice differentially expressed only 260 genes compared with IL17+ GFP cells from the same mice. Thus, cells that produce the same cytokines, do not represent discrete states, but are rather on a transcriptional continuum that is mostly defined by other factors (such as infection type) other than the cytokine they produce.

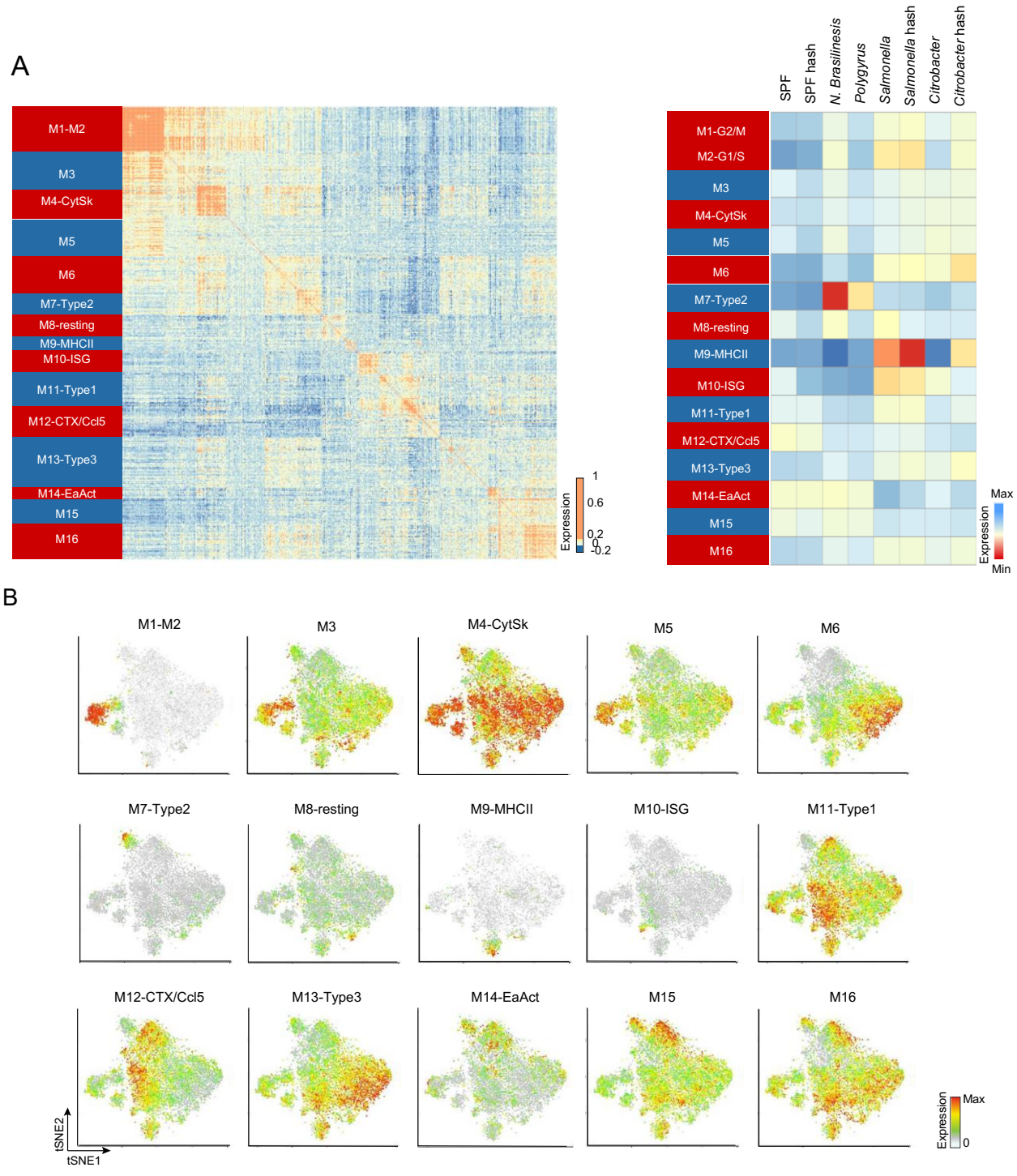
The primary determinants of Teff variability

If cytokine production is not the primary indicator of Teff heterogeneity, what drives it then? We turned again to gene:gene correlation between single cells across the conditions tested, to identify covarying gene modules that constitute the major components of T cell diversification. We used Affinity Propagation to separate coregulated modules in an unsupervised manner (Fig. 2.4A). Several gene modules were related to house-keeping tasks such as Cell cycle (M1-G1/S and M2-G2/M) and cytoskeletal genes that are associated with motility (M4-CytSK). Others were associated with particular T helper modes of action such as Tfh, Type I, II and III responses. Some of the modules were enriched for different signaling pathways such as TNFR2 signaling (M16), IL-2 signaling (M15), JNK signaling (M6), Type I interferon response (M10-ISG), as well as early activation (M14-EaAct). One interesting module (M12-CTX/CCL5) clustered closely to Type I response (M11-TypeI) and included *Ccl5* and cytotoxicity genes (*Gzmb*, *Nkg7* and *Klrk1*, which codes for NKG2D).

Different gene modules were more enriched in some conditions than others. For example, M14-EaAct and M12-CTX/CCL5 programs were enriched in cells from SPF mice, whereas M4-CSK, M11-TypeI, M6 and the cell cycle modules (M1-G1/S and M2-G2/M) were enriched in cells from *Citrobacter* and *Salmonella* infections (Fig. 2.4B). Importantly, when we

Figure 2.4: 16 Gene modules drive Teff heterogeneity across cells and samples
A) Total mean Pearson gene:gene correlation of the most highly variable genes in each sample. 16 clusters were determined by Affinity Propagation.
B) Heatmap representing expression of each of the 16 gene modules from A in each of the single cell samples.
C) Overlay of gene modules on T-SNE from hashed samples.

Figure 2.4


A
M1-M2
M3
M4-CytSk
M5
M6
M7-Type2
M8-resting
M9-MHCII
M10-ISG
M11-Type1
M12-CTX/Ccl5
M13-Type3
M14-EaAct
M15
M16
Expression
1
0.6
0.2
0
-0.2
SPF
SPF hash
N. Brasilinesis
Polygyrus
Salmonella
Salmonella hash
Citrobacter
Citrobacter hash
M1-G2/M
M2-G1/S
Max
Min
B
M1-M2
M3
M4-CytSk
M5
M6
M7-Type2
M8-resting
M9-MHCII
M10-ISG
M11-Type1
M12-CTX/Ccl5
M13-Type3
M14-EaAct
M15
M16
tSNE2
tSNE1
Expression
Max
0

overlayed many of these gene modules on the t-SNE representation, they showed differential representation across cells and were distributed as gradients rather than discrete (Fig. 2.4C).

A functional continuum of CD4+ Teff cells?

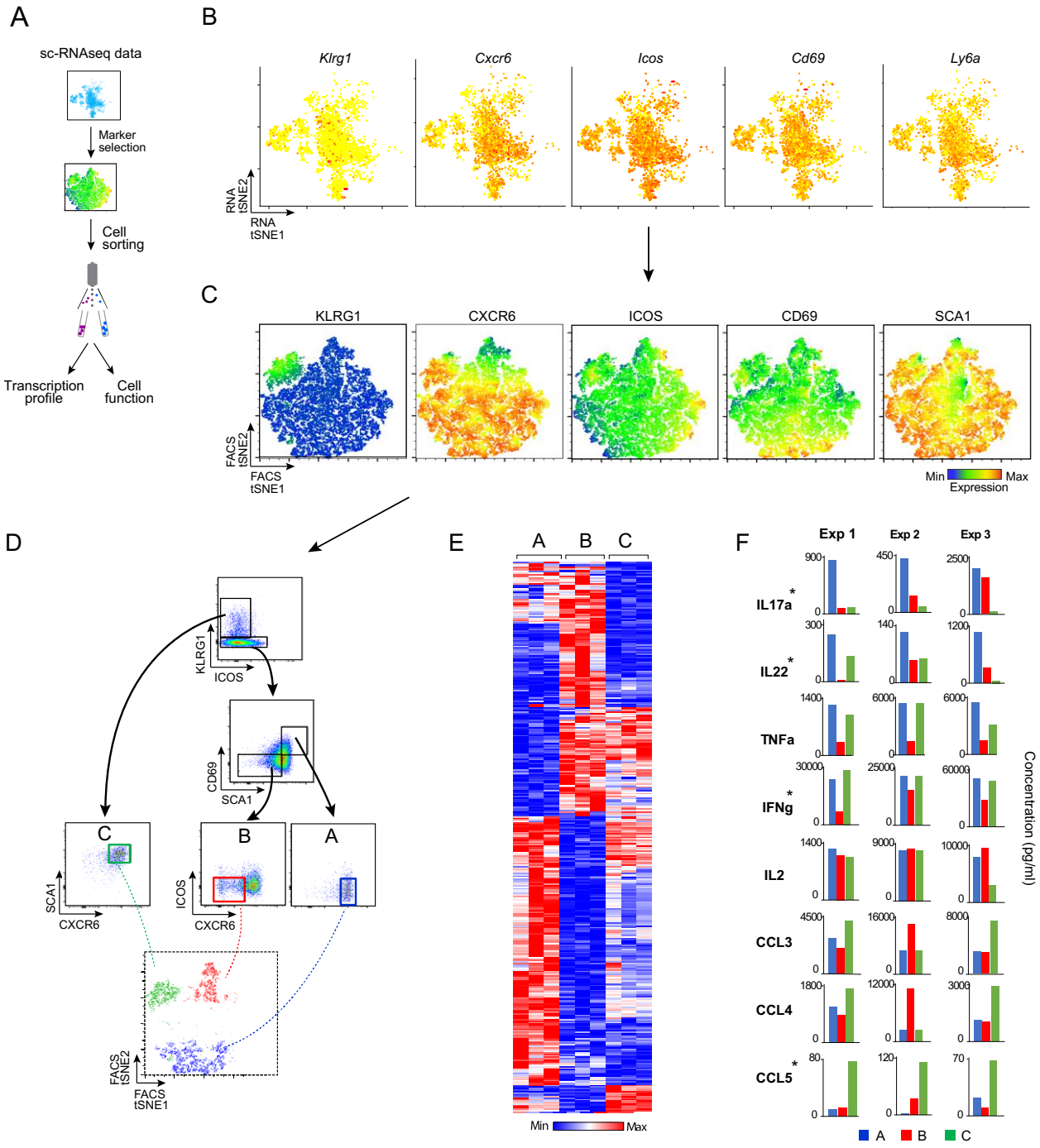
An organization of Teff cells as a continuum of cells in which different functions are distributed along various poles and gradients is more challenging than a framework with discrete and identifiable groups of cells, whether to handle conceptually or as an operational context for experimental design. The "poles" in the data should provide experimentally tractable framework, however. To illustrate the potentialities such a model allows, we followed an integrated multistep strategy (Fig. 2.5A). We first searched for and identified, within the scRNAseq data, several transcripts expressed in Teff cells that showed clear gradients of expression, and encoded cell surface molecules that can be used for flow cytometry with monoclonal antibodies (*Klrg1*, *Cxcr6*, *Icos*, *Cd69*, *Ly6a* (encodes Sca-1); Fig. 2.5B). We then analyzed colon LP cells by flow cytometry with labeled antibodies against these markers, displaying the multiparameter results on a t-SNE plot (Fig. 2.5C). Gates useable to pilot a cell sorter were determined iteratively, to sort cells belonging to specific areas of the cell cloud (Fig. 2.5D), yielding three distinct populations that could be sorted according to specific marker patterns (population A: ICOSloKLRG1loCXCR6hiCD69hiLy6ahi; population B: ICOSloKLRG1loCXCR6loCD69loLy6alo: population C: ICOSloKLRG1hiCXCR6hi). Such cells were sorted from colon LP of Salmonella-infected mice for phenotypic and functional testing.

Conventional RNA-Seq on these sorted population showed a differential transcript representation, with enrichment in populations A, B or C that corresponded well to the

Figure 2.5: Transcriptional and functional validation of Teff continuity

A) Experiment schematic. Continuously distributed surface markers were selected from the sc-RNAseq data. Cells were stained with antibodies to the surface markers, and were analyzed by flow cytometry. Flow t-SNE was then calculated on Teff cells, and gates for sorting were selected to include poles of the flow t-SNE. Cells were sorted for RNA-seq and multiplex ELISA.
B) Gene expression of selected surface markers in CD4 Teffs from *Salmonella* infection.
C)Flow cytometry t-SNE generated from MFIs of CD4 Teffs stained with antibodies against KLRG1, ICOS, CD69, SCA1 and CXCR6.
D) Sorting strategy for bulk RNA-Seq and ELISA. $ICOS^{lo}$ $KLRG1^{lo}$ $CXCR6^{hi}$ $CD69^{hi}$ $SCA1^{hi}$ was sorted as population A, $ICOS^{lo}$ $KLRG1^{lo}$ $CXCR6^{lo}$ $CD69^{lo}$ $SCA1^{lo}$ was sorted as population B and $ICOS^{lo}$ $KLRG1^{hi}$ $CXCR6^{hi}$ was sorted as population C. Sorting gates corresponded to the poles of the flow t-SNE from C.
E)Heatmap comparing differentially expressed genes between populations A, B and C. Hierarchically clustered and row mean normalized.
F) Three independent multiplex ELISA experiments comparing secretion of cytokines and chemokines from populations A,B and C. Asterisks represent p value <0.05.

Figure 2.5 Cont.



predicted signatures derived from the scRNAseq data (Fig. 2.5E and Fig. S4A). These included, in population B, transcripts associated with a more resting state (*Ccr7*, *Sell* and *Tcf7*), while Il17-associated transcripts (*Rorc*, *Il23r*, *Il17re*, *Tmem176a*) were over-represented in Population A. For a test of function, we stimulated the sorted cells and measured cytokine secretion by ELISA (Fig. 2.5F). Distinctive patterns were observed although, as expected, no single pool was associated with exclusive secretion of any one cytokine. Both populations A and C secreted significantly more IFN-γ than population B, whereas population A secreted more Il-17A and IL-22 (Fig. 2.5F) CCL5 levels were slightly higher in population C. These results confirm that the scRNAseq data captures real heterogeneity within the Teff continuum. The t-SNE pole that represents population A expresses more IL-17A and IL-22, and the pole that represents population A expresses more type IFN-γ and TNF-α. But both encompass all potentialities, only in quantitatively different amounts.

Interestingly, we noticed that module M11-Type 1 was overrepresented in population A, whereas module M12-CTX/Ccl5 appeared mostly in populations B and C (Fig. S4B). We went back to the single cell data to see if different cells, in fact, express these genes. Indeed, across both biological replicates of cells from *Salmonella*-infected mice, genes from module M12-CTX/Ccl5 were expressed in a different cell population than the majority of Ifng expressing cells (Fig. S4C). Both of these populations would be previously classified as “Th1” as they were mostly *Tbx21*-positive and *Rorc*-negative. These data further strengthens the notion that T cell heterogeneity is not determined by “Th” subtypes, as we could see large heterogeneity within *Tbx21*-expressing cells.

Lastly, we were interested to see what role TCR activation plays in the T cell heterogeneity that we observed. To this end, we used another droplet-based scRNA-seq platform, InDrop, to obtain both the transcriptome and the TCRαβ variable region sequences from the same single cells as has been previously described in Zemmour et al., 2018. We profiled T cells from a mouse infected with *Citrobacter*, as we observed the most transcriptional heterogeneity in CD4 T cells elicited by this infection (Fig. 2.1E). One particular clone expanded in all clusters, suggesting that cells with the same TCR can express different gene modules and assume different states on the continuous transcriptional space (Fig. S5A-C).

Our data suggest that at least in the case of IL-17 and IFN- γ producing cells, their transcriptional states overlap to a large extent and cannot be parsed out into discrete cell states. We thus propose a new, "quantum states" model for Th heterogeneity (Fig. 2.6). Whereas the classic model describes Th subsets as discrete, finally differentiated states, and the plastic model describes discrete states that can convert to each other, our proposed "cloud model" describes cells on a continuum of partially overlapping states that can be only partially construed by the cytokines they produce.

2.4. Discussion

Our study attempts to study the lingering question of T cell heterogeneity on a transcriptional single-cell level. Since the realization over 40 years ago that T helper cells perform various different functions such as B cell help or macrophage activation, the field has tried to subdivide the T cells into discrete subsets according to their function. CD4 helper function translated to cytokine production, and so T helper cells were divided according to their cytokine production (and later, "master" transcription factor) profile. This labeling has persisted

despite numerous reports of non-exclusive cytokine expression, as well as plasticity between states.

We propose a Th model that is neither discrete nor plastic. We have shown that at least IL17- and IFN-γ-producing cells are not found in discrete states, but plasticity between them has been previously shown by multiple groups. However, our data suggest that Teff cells are on a gradient of gene modules and cell states (with single-positive cells at the poles of the gradient), whereas the plasticity model predicts a binary switch of cell identity that the plasticity model. This model combines and does not exclude neither previously described "distinct" states (albeit at the poles of the continuum) nor "plastic" states (albeit with a continuum of transition states rather than binary state switch). Rather than plasticity between several rigid cell states, we propose fluid transcriptional states, which may allow Teffs to adapt quickly to different pathogens and challenges. This is, in a way, not dissimilar to quantum states, where an electron cannot be found in one discrete state at a particular orbital, but rather within a probability distribution that determines its probability to be observed at a particular state.

Figure 2.6

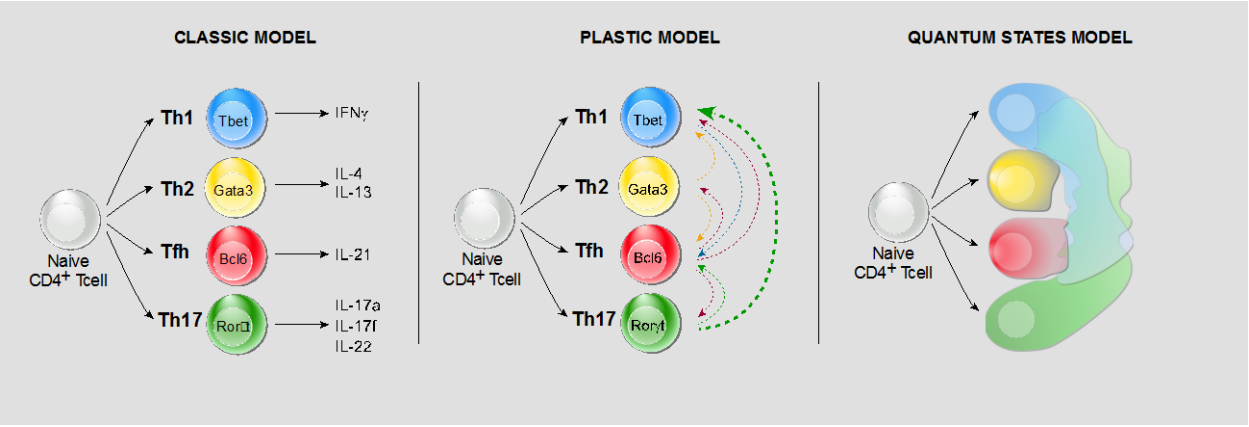


Figure 2.6: A new model for T helper cell classification.

Recently, Bielecki and Riesenfeld et al. (2018) described a similar continuous "cloud" of cells from scRNA-seq data of lung ILCs. This is intriguing, as ILCs are thought to have transcriptional programs that mirror those in T helper cells (ILC1, ILC2 and ILC3). We speculate that Type I and III programs are also found on a continuum in other cell types, such as ILCs, NKTs and γδ T cells.

Interestingly, infections with bacteria or helminths that elicit strong T helper responses, did not produce a better cluster resolution, with the exception of Th2 and Tfh cells. Our results suggest that Th2 and Tfh cells are, indeed, discrete cell states from Th1 and Th17 cells. Overall, Th2 and Tfh were less activated than the other cell types, so we cannot exclude that their discrete states are a consequence of lower inflammation or TCR stimulation. It is possible that upon proper activation, or within germinal centers in draining lymph nodes, these cells would be transcriptionally more similar to activated Th1 and 17 cells. Similarly, the helminth infections that we used are primarily infecting the small intestine, rather than the colon. Thus, it is possible that the Th2 response in the colon is less inflammatory than the responses to Salmonella or Citrobacter, and this is why the Th2 and Tfh responses looks distinct. Published data do support the idea that Th2 is a more stable cell state, but that Th1 cells can produce IL-17 and, conversely, "Th17" cells can produce IFN-γ (Lexberg et al., 2010 and Geginat et al., 2014). Our own data suggest that IL-4, IL-5 and IL-13 production is usually separate from IFN-γ, even though helminths elicit a strong type I response as well.

We were also surprised to see that surface markers that have been used for identification and sorting of particular Th subtypes did not match the expected cytokine profile. The most extreme example was IL1R2, an IL-1 receptor that was shown to be expressed in IL17-

producing cells (Skeppner et al., 2015 and Gaublomme et al., 2015), but in our data matched mostly with IFN-γ producing cells. CCR6+ cells, which are traditionally thought to match with IL17-expressing cells (Acosta-Rodriguez et al., 2007), were also a mix of single or dual Il17 and IFN-γ producers at roughly equal ratios, and missed many IL17-producing cells. This result confirmed the large heterogeneity within the IL17-producing cells. We observed similar heterogeneity within the IL17-negative population, where mostly *Tbx21* positive cells split along two main cell clusters: one that expressed high levels of IFN-γ, and another that expressed high levels of *Ccl5*, *Gzmb* and *Cxcr3*. However, these two cell clusters were still distributed along a continuum, with many cells expressing both *Ccl5* and *Ifng*. This heterogeneity within *Il17a*- or *Tbx21*-expressing cells further suggested that the expression of cytokine genes does not confer a distinct cell identity.

Cells that produce the same cytokine have vastly different transcriptional profiles based on the type of infection they encounter (or lack thereof). This has been previously described *in vitro* in human IL17-producing cells, which produce IFN-γ upon *Candida albicans* infection, and IL-10 upon *Staphylococus aureus* infection (Zielinski et al., 2012). Indeed, we saw some genes that overlapped with whole tissue RNA-Seq of large intestine from *Citrobacter rodentium*-infected mice (Kang et al., 2018). Specifically, we found *Atp1b2 (*a Na^+/K^+ Atpase), *Neto2* (neuronal) and *Uck2* (a nucleotide kinase) to be exclusively expressed in T cells from *Citrobacter rodentium* infection, while Kang et al. (2018) identified them in the top 20 up-regulated differentially expressed genes in *Citrobacter rodentium*-infected colon. This can suggest two things: either that differences in T cells transcriptome during infection are the main changes

that occur in the large intestine, or that the infection causes different cells to express similar gene modules.

Since different pathogens are heterogeneous and employ a myriad of different mechanisms of infection, it is not surprising that the immune response is not limited to few modes of response (against viruses and intracellular bacteria for Th1; helminth for Th2; extracellular bacteria and fungi for Th17). Thus, different pathogens can either directly or, more likely, indirectly (via APCs and other cells) signal to the T cells and thus change their transcriptional profile.

Importantly, we were able to functionally and transcriptionally verify our observations from scRNA-seq by sorting cells not from defined clusters, but from the t-SNE continuum. Sorting cells from different poles of the t-SNE and analyzing their transcriptome by bulk RNA-Seq or testing their functional ability to produce cytokines recapitulated the continuity we previously observed: the cell "cloud" was experimentally tractable; cells did not express one Th signature/cytokine or another, just quantitatively different amounts.

What is the benefit of expressing or having the potential to express different gene modules and cytokines in the same cell? One might argue that cell specialization is a hallmark of immunity, but perhaps under conditions of highly inflammatory infection, it is beneficial for cells to go "Gung-ho" and try to fight the pathogen in any way possible.

In conclusion, our study sheds light on T effector cell heterogeneity at steady-state and infection and establishes a new model for their classification that can perhaps be borrowed into other cell types.

Chapter 3: Novel subsets of CD4+ T effector T cells

In this chapter we focus on the peripheral populations that were outside the main T cell "cloud". One of these populations is the ISGT (Fig. 2.1E), a very small isolated Teff population which was intriguing because it expressed high levels of IFN-I induced transcripts. These cells were again over-represented after *Salmonella* or *Citrobacter* infections, and suggest that a small subset of Teff cells may be uniquely sensitive to IFN stimulation, or that they reside in a small analtomical compartment in which IFN is particularly abundant.

Another intriguing cluster (CycTx, Fig. 2.1E) consists of cells that express many cell cycle genes as well as tubulin-associated genes.Interestingly, cell cycle regression did not merge that cluster with the other populations, suggesting a unique gene expression signature that is not driven solely by cell cycle.Using Ingenuity pathway analysis (IPA) we found enrichment in oxidative phosphorylation, cholesterol biosynthesis and Rho-GTPase signaling, which is important for cell motility and immunological synapse stability (Saoudi et al., 2014).

Another population expressed high levels of the surface markers *Cd160*, *Crtam* and *Lag3,* the neural gene *Nrgn* (encoding neurogranin, a neural protein that binds Ca^{2+}-free calmodulin) as well as the chemokines *Xcl1*, *Ccl1*, *Ccl3* and *Ccl4* (Fig. 3.1A). We thus sorted this population for RNA-Seq as $CRTAM^{+}$ and detected a very similar gene signature from the RNA-seq data, with the chemokines highly expressed in the $CRTAM^{+}$ cells (Fig. 3.1B). IPA revealed Stat3 signaling, prolactin signaling and neuregulin signaling genes to be enriched in this population. Interestingly, many of the most upregulated genes in this population, such as *Crtam*, *Ccl1*, *Nrgn* and *Lag3* and *Tnfrsf9* (encodes 4-1BB) are targets of Egr2, which is suggestive of an anergic state (Zheng et al., 2013).

Another “peripheral” population (MyT, CycTx, Fig. 2.1E), which was mostly coming from *Salmonella*-infected mice, showed high expression of myeloid genes, such as genes coding for MHC class II, as well as *Apoe*, *Lyz2* and *C1qa* (Fig 3.1C). While at first we saw that this was a myeloid:T cell doublet, we became convinced that this is not the case. Firstly, we could detect very few (~0.6%) myeloid cells in the single cell data. Secondly, myeloid cells tend to have more RNA than T cells and thus have more UMI counts per cell, but this was not the case with cluster 8. Thirdly, the expression levels of key CD4 T cell genes such as *Trac*, *Trbc2* and *Cd4* were at the same level as the other clusters. Lastly, this myeloid-like population constituted 3-9% of each sample from *Salmonella typhimurium* infected mice, but they did not appear to be present in other samples.

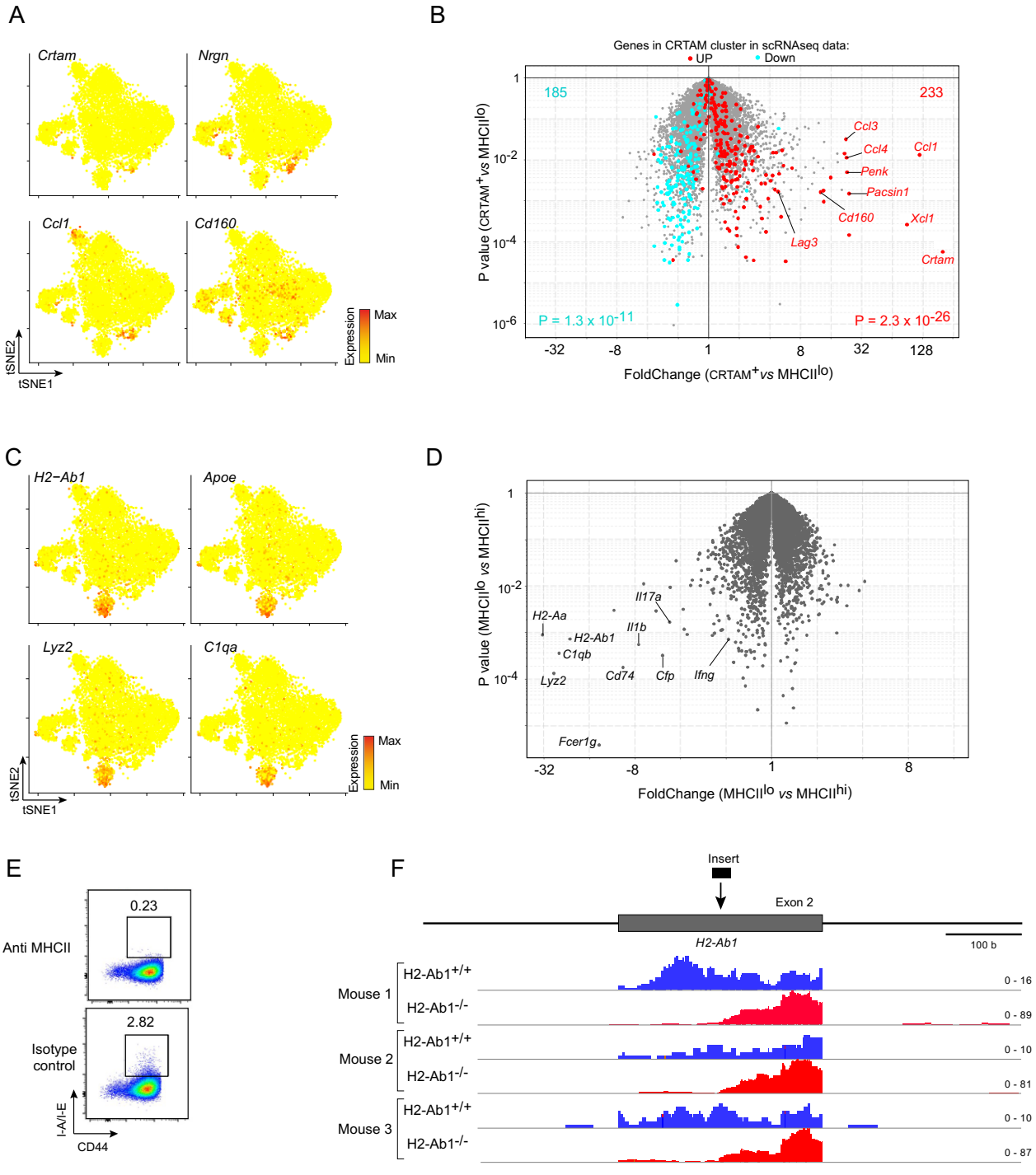
To verify that what we are seeing is real, we stained for MHCII in the colon. Surprisingly, we saw that around 5% of the $CD44^{hi}$ (effector) CD4 T cells but not Tregs expressed MHCII (Fig. 3.1D). We thus sorted $MHCII^{hi}$ and $MHCII^{lo}$ cells for deeper transcriptional profiling. Interestingly, $MHCII^{hi}$ CD4 T cells expressed a very similar set of genes to the ones we observed in the single cell RNA-seq data: *Lyz2*, *Apoe* and other MHCII genes (all except for *H2-Ab1* have null mutations in C57BL/6 mice; Fig. 3.1E). These data suggested that this “myeloid-T” is not an artifact of sorting or droplet-based single cell RNA-sequencing.

To confirm that our observations are due to endogenous expression of myeloid genes and not due to trigocytosis or uptake of foreign myeloid RNA, we created mixed bone marrow chimeras from MHC class II (I-A) knockout (Cosgrove et al., 1991) and congenic wild-type mice. 8 weeks after bone marrow reconstitution the mice were infected with *Salmonella typhimurium* and their colonic CD4 T cells were analyzed by flow cytometry and transcriptionally profiled.

Flow cytometric analysis showed that CD4 T cells from I-A$^{-/-}$ mice still stained positive for MHC class II, suggesting trigocytosis. RNA-seq, on the other hand, showed that all the I-A transcripts were endogenous, i.e. wild-type H2-Ab1 transcripts were only found in wild-type cells and exon 2 deleted H2-Ab1 transcripts were found only in I-A$^{-/-}$ cells (Fig. 3.1F). These data suggested that CD4 T cells endogenously express I-A, but are also able to pick up MHC class II protein through trigocytosis. Surprisingly, many of the transcripts that were upregulated in the T-Myl population as well as the sorted MHC^{hi} population were upregulated in the CD4 T cells from I-A$^{-/-}$ mice, suggesting that H2-Ab2 downregulates these transcripts in a negative feedback loop (Fig. S6). Further experiments will show whether CD4 T cells can present antigen in the context of *Salmonella typhimurium* infection and if these myeloid genes are important for their function.

Figure 3.1:Novel CD4 Teff subsets. A) scRNA-seq expression data of genes in Crtam^{+} cluster (cluster 11).
B) Volcano plot comparing RNA-seq data and scRNA-seq data of Crtam^{+} cells vs total Teff cells (MHCII lo). Over/under-expressed genes from scRNA-seq data are shown in red and blue, respectively, with significance of overlap.
C) scRNA-seq expression data of genes in myT cluster
D) Volcano plot representing RNA-seq data comparing MHCII hi and MHCII lo Teff cells from LP of *Salmonella* infected mice.
E) Flow cytometric analysis of CD4+TCRβ+FOXP3- cells from LP of Salmonella infected mice.
F) RNA-seq reads from WT and H2-Ab1 KO Teff cells from mixed bone marrow chimeras infected with *Salmonella*, aligned to exon2 of H2-Ab1. Neomycin insert that disrupts the gene is shown above.

Figure 3.1


A
Crtam
Nrgn
Ccl1
Cd160
tSNE2
tSNE1
Expression
Max
Min
B
Genes in CRTAM cluster in scRNAseq data:
UP
Down
185
233
Ccl3
Ccl4
Ccl1
Penk
Pacsin1
Cd160
Xcl1
Lag3
Crtam
P = 1.3 x 10-11
P = 2.3 x 10-26
P value (CRTAM+ vs MHCIIlo)
FoldChange (CRTAM+ vs MHCIIlo)
C
H2−Ab1
Apoe
Lyz2
C1qa
D
P value (MHCIIlo vs MHCIIhi)
FoldChange (MHCIIlo vs MHCIIhi)
H2-Aa
H2-Ab1
C1qb
Lyz2
Il17a
Il1b
Cd74
Cfp
Ifng
Fcer1g
E
Anti MHCII
0.23
Isotype control
2.82
I-A/I-E
CD44
F
Insert
Exon 2
H2-Ab1
100 b
Mouse 1
Mouse 2
Mouse 3
H2-Ab1+/+
H2-Ab1-/-
0 - 16
0 - 89
0 - 10
0 - 81
0 - 10
0 - 87

Chapter 4: Tregs and their role in auto-immunity and cancer

4.1.Tregs and their role in auto-immunity and cancer

A hallmark of the adaptive immune system is its ability to tolerate self-antigens and harmless foreign antigens. Without this ability the immune system would attack its own host, its food, and symbiotic microorganisms that live within the host, leading to deleterious effects. Regulatory T cells (Tregs) help to prevent this problem and play an important role in the suppression of immune responses and immune self-tolerance. This rare cell subset (~10% of $CD4^+$ cells) consists of $CD4^+CD25^+$ T cells that express the transcription factor forkhead box P3 (FoxP3). Loss of function mutations of FoxP3 in mice or humans leads to severe organ autoimmunity and inflammation due to lack of Treg cells (Benoist and Mathis, 2012). Ablation of Treg cells in adult mice results in a similar phenotype, while adoptive cell transfer of Tregs ameliorates the disease, suggesting that Tregs exhibit a dominant suppressive phenotype (Kim et al., 2007). Tregs can exert their suppressive function in multiple ways, such as direct killing of APCs or other T cells, production of suppressive cytokines such as TGF-β, IL-10 and IL-35, deprivation of trophic cytokines and nutrients, as well as stripping off stimulatory ligands from APCs (Vignali et al., 2008 and Josefowicz et al., 2012).

As mediators of immune tolerance, Tregs have been implicated in autoimmune diseases such as Type-I diabetes, rheumatoid arthritis, and inflammatory bowel disease (IBD), as well as cancer. Whereas in autoimmune diseases Tregs lose their suppressive function and stability, in cancer Tregs nurture the tumor immunosuppressive microenvironment and prevent the host immune system from eliciting an effective immune response against tumors (Long et al., 2011 and Nishikawa et al., 2014). Adoptive cell transfer (ACT) of Tregs into a mouse model of

rheumatoid arthritis suppressed arthritis, and recruitment of Tregs to pancreatic islets in NOD mice has been shown to protect from autoimmunity (Wright et al., 2009 and Montane et al., 2011).

Treg differentiation in the thymus is not well understood, but current data suggest an interplay between IL-2 levels and TCR signaling in the thymus. $CD4^+$ T cells that experience strong TCR signaling and high levels of IL-2 change their epigenetic landscape and up-regulate CD25 and turn on FoxP3 (Samstein et al., 2012).Treg lineage maintenance highly depends on stable expression of FoxP3, as its loss of its expression leads to inhibition of suppressor function [10]. However, expression of FoxP3 alone is not sufficient for conferring and maintaining the Treg phenotype, as ectopic retroviral transduction of Tconvs with FoxP3 does not induce many Tregsignature genes (Hill et al., 2007 and Kwon et al., 2017). These data suggest involvement of other factors in Treg biology.

Recently, our lab and others have identified tissue-resident subsets of Tregs, which reside in fat, muscle, and intestines, and can play a role in metabolic regulation, tissue repair and interactions with microbiota, respectively (Feuerer et al., 2009, Burzyn et al., 2013, Sefik et al., 2015 and Panduro et al., 2016). An intriguing feature of these tissue Tregs is the adaptation of certain transcriptional programs that mimic some signature genes in their tissue environment (Cipolletta et al., 2012). It is, however, unclear how these Tregs migrate into these peripheral tissues, and furthermore how they change their transcriptome and epigenome.

Overall, many aspects of Treg biology remain unknown. For example, it is unclear which factors are involved in Treg differentiation in the thymus, what maintains their stability, and how their transcriptional program is changed in different contexts and tissues.

4.2. *In vivo* CRISPR/Cas9 screening of genes that affect Tregs

To answer these questions, we set to devise a system that allows us to create loss of function (LOF) mutations in Tregs in a high-throughput mode *in vivo*. To this end, we employed the CRISPR/Cas9 system, which has been employed in many library high-throughput screens (Wang et al., 2014, Shalem et al., 2014, Chen et al., 2015, and Manguso et al. 2017). The principle of these methods is to express the CRISPR/Cas9 endonuclease together with a single guide RNA (sgRNA) that targets DNA cleavage by Cas9, introducing sequence scars by the error-prone non-homologous end joining repair. These sgRNAs are usually transduced as retroviral or lentiviral libraries into a cell population that can be selected for according to a particular function (usually survival or proliferation). If a gene is specifically required for the selected function, sgRNAs that inactivate it will be under-represented relative to the starting pool.

An *in vivo* CRISPR/Cas9 screening system for Tregs involved several caveats. First, Tregs are not easily infected by retro- or lenti-viral vectors, and their transduction efficacy decreases with the size of the expression vector. Thus, we utilized the Cas9 knock-in mouse, which allowed us to "save space" on the vector and express just sgRNAs and a fluorescent marker (Platt et al., 2014). Secondly, *ex-vivo* transduction has been shown to activate T cells and affect their phenotypes (Kurachi et al., 2017). We were able to maintain the T cells "untouched" by transducing hematopoietic progenitor stems cells and allowing them to develop normally in irradiated host mice (Godec et al., 2015). This introduced another problem- only a limited number of CLPs seed the thymus and differentiate into T cells, limiting the coverage of

CRISPR/Cas9 libraries. We solved this problem by limiting the size of libraries to ~300 sgRNAs and using multiple mice per experiment.

Specifically, to screen for genes that affect Treg differentiation, proliferation or survival, we infected Lineage$^-$Sca-1$^+$C-kit$^+$ (LSK) cells from mice that either constitutively express Cas9 (Cst-Cas9), or in which Cas9 is active only in Tregs (FoxP3-CRE) with a lentiviral gRNA vector that expresses the RFP fluorescent marker (Figure 4.1).

After 8-12 weeks to allow for colonization and differentiation of mature T cells, we purified Tregs and Tconvs by flow cytometry and quantified the representation of the sgRNAs in their genomic DNA by PCR and high-throughput sequencing.

Figure 4.1

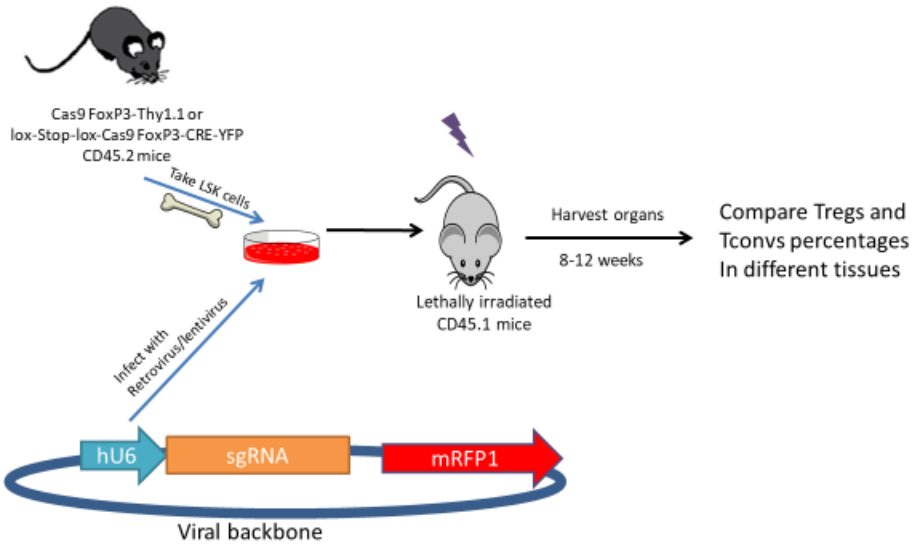


Figure 4.1: Schematic of the *in vivo* CRISPR BM chimera system. Our approach is to infect LSK cells with mRFP1 and sgRNA lentivirus. Only mRFP1 expressing cells carry gRNAs to inactivate the gene of interest. We compare the ratios of cells that are CD4+TCRb+Thy1.1+ (Tregs) and CD4+TCRb+Thy1.1- (Tconvs).

We validated that we achieve high genome editing efficacy with several targets such as FoxP3 (Figure 4.2). We then set to test different genes in a pooled approach. The logic of the experiment is that if a gene is specifically required for the differentiation or homeostatic maintenance of Tregs relative to others, sgRNAs that inactivate it will be under-represented in that cell. However, if a gene is generally necessary for differentiation of all cells, its inactivating sgRNAs will be un-represented everywhere, relative to the starting LSK pool.

We developed a robust tool for gene editing T cells *in vivo*, and for screening different genes for effects on T cells *in vivo*. Several people in the lab are now using this method to study the effects of their genes of interests on Tregs in multiple tissues. Specifically, this approach was used to screen a library of nuclear and chemokine receptors in Tregs (manuscript in preparation).

Figure 4.2

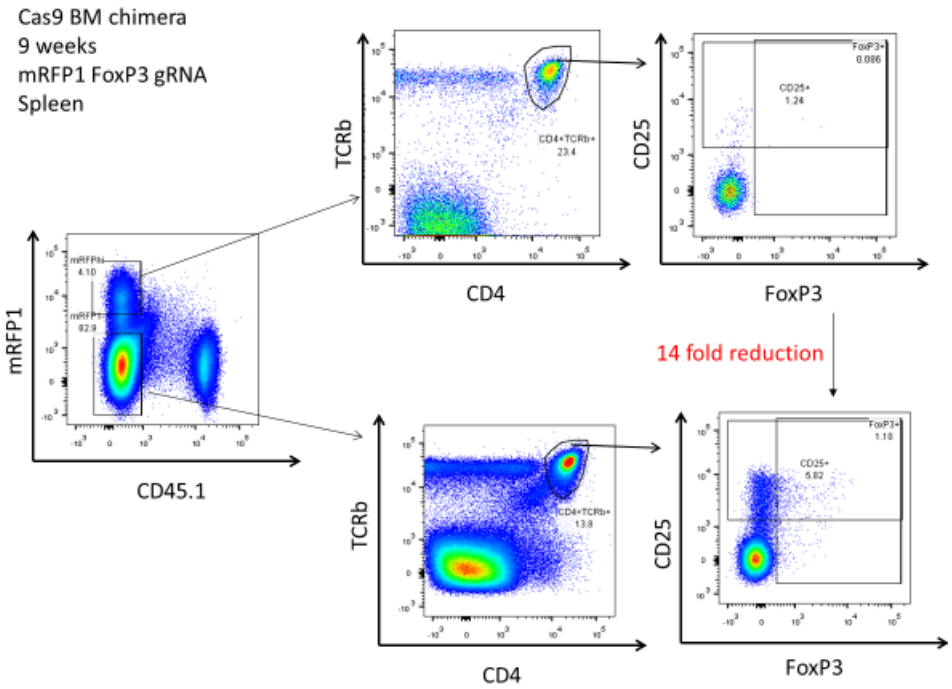


Figure 4.2: The in vivo gene editing is robust. mRFP+ CD4 T cells that have been infected with a gRNA targeting FoxP3 have low levels of FoxP3 and fewer Tregs than the mRFP- cells that have not been infected.

4.3. *In vivo* CRISPR/Cas9 validation of genes that affect tumor-infiltrating Tregs accumulation in tumors

Tregs have been shown to contribute to the immunosuppressive environment of tumors (Galon et al., 2006). Furthermore, several studies showed that when Tregs are depleted from the tumor, tumor progression is deterred (Simpson et al., 2013 and Bos et al., 2013). Our lab has identified genes that are specifically over-expressed in tumor-infiltrating Tregs (TITRs) in both mouse models and human tissues by integrating RNA-seq from mice and humans and TCGA data. To test if these over-expressed genes are functionally important for Tregs accumulation in tumors, we performed a systematic *in vivo* screen of the top-ranked TITR-specific targets by creating LOF mutations in the protein-coding regions by adapting the CRISPR/Cas9 system to specifically edit Treg cells, assessing differential representation in the tumor relative to lymphoid organs (Platt et al., 2014). Tregs were purified from transgenic mice expressing both the Cas9 protein and the FoxP3-Thy1.1 reporter (Liston et al., 2008), were activated *in vitro*, and were transduced with retroviral vectors encoding the targeting single-guide RNAs (sgRNAs) of interest (Figure 3.3 A). Pairs of retroviral vectors expressing two different fluorescent proteins (GFP and RFP) were used, which allowed us to label and compare Tregs carrying targeting or control sgRNAs in the same mouse (and thereby increase the informativity of each mouse). Transduced Tregs were then transferred together with "filler" splenocytes (to avoid homeostatically driven expansion) into alymphoid RAG-deficient mice, which concomitantly received MC38 tumor cells. After 18 days to allow for tumor growth and Treg homing, we analyzed the distributions of Treg cells carrying each of the targeting or control sgRNAs in the tumor relative to lymphoid organs, calculated as a "tumor depletion

index" (Figures 4.3 B and C). By DNA amplification and sequencing we also ensured that the target genes were efficiently edited (estimated as > 50% in all cases). The compiled results showed that several of the LOF mutations had no impact on TITR proportions, but three of them did: *Tnfrsf8* (encodes CD30), *Cxcr3* (a chemokine receptor) and *Samsn1* (an intracellular signaling adapter). LOF of *Tnfrsf8* and *Cxcr3* resulted in consistent effects, which were cross-validated by antibody blockade. This first *Samsn1* LOF result was validated with four independent sgRNAs targeting different regions of *Samsn1*. Each sgRNA decreased Treg accumulation in the tumors (Figure 4.3D), confirming the significance of the effect.

Figure 4.3

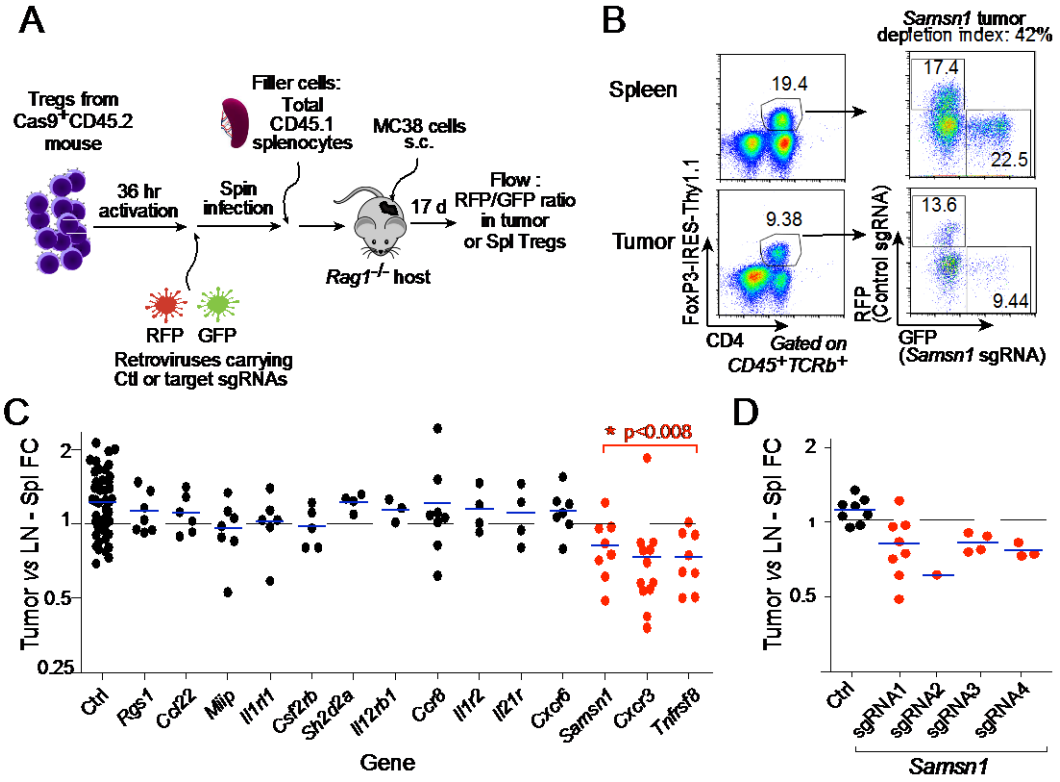


Figure 4.3 CRISPR-based Treg knockouts. A. Schematic depiction of protocol to induce loss of function (LOF) mutations in TITR target genes, specifically in Treg cells, utilizing the CRISPR/Cas9 system. **B**. Exemplar gating used for determining the tumor depletion index- the ratio of the percentage of GFP or RFP+ Tregs in tumor vs. spleen. **C.** Summary of CRISPR-based Treg knockout data. The tumor depletion index was significant in the genes highlighted in red. **D.** Validation of Samsn1 LOF on Treg accumulation in tumors. Four Different sgRNAs showed decreased accumulation of Samsn1 LOF Tregs in the tumors.

4.4: Manuscript (see appendix A)

Collaboration: The work referenced above was initiated by Dr. Angela Magnuson and experiments were conducted and analyzed in collaboration with A. Ergun, J.S. Park, N. Asinovski, A. Ortiz-Lopez, A. Kilcoyne and E. Paoluzzi-Tomada. This study was conducted under the supervision and direction of Christophe Benoist, Diane Mathis and Ralph Weissleder. I initiated, conducted and analyzed the CRISPR knockout validation experiments in Tregs.

Chapter 5: Materials and Methods

Single Cell RNA-seq of CD4

Mice

For single cell experiments, 6-week old male C57BL/6 mice were purchased from Jackson Laboratories, kept for 1 week under specific pathogen-free (SPF) conditions in the animal facility at Harvard Medical School, and then infected with pathogens. IL17A$^{GFP/+}$ mice (Jax:C57BL/6-Il17atm1Bcgen/J) were a gift from Dr. Jun Huh. OT-II TCR-transgenic mice were obtained from Jackson Laboratories (B6.Cg-Tg(TcraTcrb)425Cbn/J). I-A$^{-/-}$ mice were previously generated in our laboratory (Cosgrove et al., 1991). All mice were bred and/or maintained in our specific-pathogen-free facilities at Harvard Medical School (IACUC protocols IS1257, IS187-3, IS2221).

Infections

Salmonella typhimurium: Mice were gavaged with 100 µl of 200 mg/mL streptomycin in water. 24 hours later, they were gavaged with 10^9 *Salmonella enterica* (Serovar Typhimurium) AroAΔ (gift of D. Littman). Unless noted otherwise, mice were sacrificed at day 13 after infection.

Citrobacter rodentium: Mice were gavaged with 5X10^8 *Citrobacter rodentium* and sacrificed 13 days later.

Helminths: Mice were gavaged with 200 L3 larvae of *Heligmosomoides polygyrus* in 200 µl H_2O or subcutaneously injected with 500 L3 larvae of *Nippostrongylus brasiliensis* in 100 ul PBS, and sacrificed 11 days later.

Preparation of mouse lymphocytes for single-cell RNA-seq

Intestinal tissues were treated with RPMI containing 1 mM DTT, 20 mM EDTA and 2% FBS at 37°C for 15 min to remove epithelial cells, minced and dissociated in collagenase solution

(1 mg/ml collagenase VIII (Sigma), 1 µg/ml DNAse and 1%FBS in RPMI) with constant stirring at 37°C for 20min. Single cell suspensions were then filtered and washed with 4% RPMI solution. Single-cell suspensions were stained on ice for 30 min with antibodies against CD4, TCR-β, CD19 and CD45 (Biolegend) and and sorted with Astrios MoFlo (Beckman Coulter)as $Dapi^{-}$ $CD45^{+}$ $CD4^{+}$ $TCR\beta^{+}$ $CD19^{-}$ directly into PBS with the final concentration of 0.04% BSA. For cell hashing, Total Seq A™ antibodies 301-305 were added to each sample individually at the same time as other antibodies. All samples were sorted together directly into RPMI and subsequently spun down and reconstituted in 33 ul of 0.04% PBS. All samples were loaded on the 10X Chromium Controller within 30 minutes of sorting. Libraries were prepared using Chromium Single Cell 3' Reagent Kits v2 according to the manufacturer's protocol (10X Genomics). HTO libraries were prepared as described in Stoeckius et al. (2018). Both libraries were sequenced together on Illumina HiSeq 4000.

Single-Cell RNA-seq Data analysis

Gene counts were obtained by aligning reads to the mm10 genome using CellRanger software (v1.3) (10x Genomics). Hashtag counts were obtained by using Cite-Seq-Count package. Single cell data were analyzed using the Seurat package (Butler et al., 2018). Hashtags were assigned to cells using the HTODemux function and doublets were eliminated from analysis. Cells with less than 1000 UMIs or 400 genes and more than 4000 UMIs were also eliminated from analysis. Tregs and naïve cells were removed from analysis by using the SubsetData function.

Flow cytometry

Cells from colon LP were prepared as previously described (Sefik et el., 2015). Briefly, Intestinal tissues were treated with RPMI containing 1 mM DTT, 20 mM EDTA and 2% FBS at 37°C for 15 min to remove epithelial cells, minced and dissociated in collagenase solution (1.5mg/ml collagenase II (Gibco), 0.5mg/ml Dispase (Gibco) and 1%FBS in RPMI) with constant stirring at 37°C for 45min. Single cell suspensions were then filtered and washed with 4% RPMI solution. For cytokine analysis, cells were treated with RPMI containing 10% FBS, 10ng/ml phorbol 12-myristate 13-acetate (Sigma), 1µM Ionomycin (Sigma) in presence of protein transport inhibitor cocktail (eBioscience) for 3.5 hours. For intracellular staining of cytokines and transcription factors, cells were stained for surface markers and fixed in eBioscience Fix/Perm buffer overnight, followed by permeabilization in eBioscience permeabilization buffer for 45 min in the presence of antibodies. Cells were acquired with BD Symphony and analysis was performed with FlowJo (Tree Star) software. Flow cytometry T-SNEs were generated using FlowJo version 10 (Tree star).

Ultra Low Input RNA-Seq

IL17-GFP mice were gavaged with either *C. Rodentium*, *S. Typhimurium* or PBS, sacrificed at day 13 and double sorted as $CD45^{+}CD4^{+}TCRb^{+}CD44^{+}CD25^{-}GFP^{+}$. For sorting off the t-SNE, cells were first analyzed on the cell sorter, and sorting gates that resulted in the selection of specific regions on the flow cytomotery T-SNE were drawn for subsequent sorting. For sorting MHCII hi vs Lo cells, cells were stained with antibodies against CD4, TCR-β, CD19, CD11b,

CD11c, CD25 and CD44 and sorted on the BD Aria II as $CD11b^{-}$ $CD11c^{-}$ $CD19^{-}$ $CD4^{+}$ TCR-β^{+} $CD44^{hi}$ $CD25^{lo}$ MHC hi or lo.

All cells were double-sorted. For the final sort, 1,000 cells were collected directly into 5ul lysis buffer (TCLBuffer (Qiagen) with 1% 2-Mercaptoethanol), Smart-seq2 libraries were prepared as previously described (Picelli et al, 2014). Briefly, total RNA was captured and purified on RNAClean XP beads (Beckman Coulter). Polyadenylated mRNA was then selected using an anchored oligo(dT) primer and converted to cDNA via reverse transcription. First strand cDNA was subjected to limited PCR amplification followed by transposon-based fragmentation using the Nextera XT DNA Library Preparation Kit (Illumina). Samples were then PCR amplified using barcoded primers such that each sample carries a specific combination of Illumina P5 and P7 barcodes and pooled together prior to sequencing. Paired-end sequencing was performed on an Illumina NextSeq500 using 2 x 25bp reads. Reads were aligned to the mouse genome (GENCODE GRCm38/mm10 primary assembly and gene annotations vM16; https://www.gencodegenes.org/mouse_releases/16.html) or human genome (GENCODE Human Release 27; https://www.gencodegenes.org/releases/27.html; Reference genome sequence: GRCh38/hg38; annotation: GENCODE v27) with STAR 2.5.4a (https://github.com/alexdobin/STAR/releases). The ribosomal RNA gene annotations were removed from GTF (General Transfer Format) file. The gene-level quantification was calculated by featureCounts (http://subread.sourceforge.net/). Raw reads counts tables were normalized by median of ratios method with DESeq2 package from Bioconductor (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) and then converted to GCT and CLS format. Samples with less than 3 million uniquely mapped reads were

automatically excluded from normalization to mitigate the effect of poor quality samples on normalized counts.

Normalized read counts were filtered for robust expression (>10) to avoid confounders from low-level noise, and processed in the Multiplot suite and Morpheus (https://software.broadinstitute.org/morpheus/), or with custom code in R.

Bone Marrow Chimeras

Bone marrow cells were harvested from both femurs and tibias of mice, and treated with ACK buffer to remove red blood cells. CD45.1/CD45.2 mice were irradiated with 10 Gy and reconstituted with equal proportions (~5 million cells each) of CD45.1 and CD45.2 (Ab KO) bone marrow cells on the same day.

Multiplex ELISA

1000-10,000 cells were sorted as previously described (single sort) into 100 ul of T cell medium (RPMI 1640, 10% FBS, 20mM HEPES, 1mM sodium pyruvate, 0.05mM 2-mercaptoethanol, 2mM L-glutamine, 100 mg/ml streptomycin and 100 mg/ml penicillin). Cells were plated in round bottom 96-well plates with 1:1 ratio of anti-CD3/CD28 beads (Miltenyi) and were left at 37C for 24 hours. Supernatants were collected and analyzed with the LegendPlex^{TM} T Helper Cytokine Panel Version 2 and Proinflammatory Chemokine panel kits (Biolegend) per manufacturer's instructions. Samples were acquired via the BD Symphony and analyzed with LegendPlex^{TM} software.

Statistical analysis

Mann-Whitney test was used to assess significance of clusterability of Pearson gene:gene correlation clusters. Chi square test was used to assess significant correlations between cytokines. Paired student t-test was used for significance of multiplex ELISA.

Methods for Treg section

Mice

Rosa26-Cas9 knockin mouse line (Platt et al., 2014) was purchased from the Jackson Laboratory and crossed with Foxp3-IRES-Thy1.1 reporter mouse line (gift of A. Rudensky, Liston et al., 2008) to generate the Rosa26-Cas9;Foxp3-Thy1.1 line (Cst-Cas9). B6.Rag1-/- and C57BL/6 mice were obtained from the Jackson Laboratory. $Foxp3^{IRES\text{-}GFP}$ mice were a gift of T. Chatila (Lin et al., 2007). All mouse breeding and experiments were conducted in an SPF facility at Harvard Medical School according to protocol approved by Harvard Medical School's Institutional Animal Care and Use Committee (IACUC protocols 02954 and 05176).

Tumor induction

CT26 (Wang et al., 1995), MC38 (Corbett et al., 1975) , BRAF.PTEN (Dankort et al., 2009) or B16.F10 cells (Fidler et al., 1975) (obtained from Dr. Arlene Sharpe) were cultured *in vitro* in DMEM supplemented with 10% FCS, 1% Penicillin-Streptomycin, 1% L-Glutamine, 1% HEPES before being injected subcutaneously (0.5 - 1 x 10^6 in 100 µl PBS) in the flank. For MC38 re-challenge, 1 x 10^6 MC38 cells were injected into the opposite flank from the original tumor site. Tumor diameters were measured at regular intervals using digital calipers. Ellipsoid tumor volumes (mm^3) were calculated using the formula: $0.5 \times D \times d^2$ where D is the larger diameter and d is the smaller diameter.

CRISPR-based screen

To induce LOF mutations in the genes of interest, CRISPR sgRNAs were expressed in Tregs from Cas9-expressing transgenic mice by retroviral transduction. The vector also expressed either a GFP or mRFP1 fluorescent reporter to track transduced cells (two colors allowing for Tregs transduced with LOF or non-targeting sgRNA in the same mouse). We then injected these Tregs into tumor-bearing alymphoid mice and compared the migration and retention of these modified Tregs in the tumor by comparing the relative fraction of the fluorescent reporter in the tumor and in lymphoid organs ("tumor depletion index"). Underrepresentation of particular LOF Tregs in the tumors indicated that the gene of interest promoted Treg accumulation in tumors. We validated that there was efficient LOF in Tregs by targeting *Cd4* and observing >95% decrease in protein levels by flow cytometry. *sgRNAs.* sgRNAs were designed using the online tool "CRISPRko" to minimize off-target effects as well as to increase efficacy [https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design (Doench et al., 2016) (Dataset S7)].

Retrovirus production: To express the sgRNAs in Treg cells we cloned them into a retroviral vector (Retro-gRNA) which also expresses the markers eGFP or mRFP1 (for use as internal control in the same recipient mouse). Retro-gRNA was cloned from the MSCV-GFP plasmid (Kamperschroer et al., 2002) by Gibson assembly of U6-gRNA cassette and EF1α promoter from the lentiCRISPR vector (Shalem et al., 2014). mRFP1 was cloned by Gibson assembly in place of eGFP by digesting Retro-gRNA with NcoI and NotI. Cloning of individual gRNAs was done by utilizing BbsI sites, as previously described (Shalem et al., 2014). Retrovirus was prepared by transfecting Platinum-E cells at 70% confluency on a 60 mm plate with 2.3 µg

of Retro-gRNA and 1.5 μg of pCL-Eco using Mirus TransIT-293T reagent (Morita et al., 2000). Medium was changed after 16 hours and the virus (2 ml) was collected and passed through a 0.45 μ filter 48 hours after media change.

Treg culture and transduction: Tregs were isolated from spleens and lymph nodes (pooled axillary, brachial, cervical and inguinal) of 8-10 week old Cst-Cas9 FoxP3-Thy1.1 mice. Cells were first purified through negative selection by combining Biotin Binder Dynabeads (Thermo Fisher) with biotinylated anti-Ter119, CD19, CD11b and CD8, and then by flow cytometry after staining with anti-Thy1.1 and anti-CD4 (all antibodies from Biolegend). Tregs were then activated in culture [10^5 cells per well in a 96 well plate in RPMI-1640 with 10% FCS and 2000 U/ml human IL-2 (Peprotech) with anti-CD3/CD28 beads (Thermo Fisher) at 2:1 beads/cells ratio for ~40 hours]. Tregs were then spin-infected with 200 μl retrovirus supplemented with 8 ug/ml polybrene and 2000 U/ml IL-2 for 2 hours at 2000 RPM at 34^o C. Cells were then cultured for 24 hours at 37^o C in the IL-2-supplemented medium above. After 24 hrs, 2X10^5 Treg cells that were infected with each virus (eGFP- and mRFP1-tagged for target and control sgRNA, respectively) were mixed with ~3X10^6 splenocytes from CD45.1 FoxP3$^{IRES\text{-}GFP}$ mice, to prevent homeostatic expansion of the manipulated Tregs, and injected intravenously into Rag1$^{-/-}$ recipient mice. Tumor induction (s.c. injection) of 5X10^5 MC38 cells was performed at the same time.

Evaluation of differential intratumoral Treg accumulation: To assess the homing and retention of Tregs carrying each sgRNA in MC38 tumors, cell suspensions from tumors, spleens and LNs were prepared and analyzed by flow cytometry (manipulated Treg cells being identified as CD45.2$^+$CD45.1$^-$TCRb$^+$CD4$^+$Thy1.1$^+$), and the proportion of Treg cells expressing the eGFP or

mRFP1 reporters associated with the sgRNAs in each mouse were measured in spleen, LN and tumor-infiltrating T cells. We then divided the percentage of reporter-positive cells in the tumor by the averaged percentage of these cells in the spleen and LNs. We dropped from the analysis samples in which the number of reporter-positive cells was less than 150.

CRISPR scar verification: To verify the efficacy of target gene inactivation by CRISPR scarring, cell suspensions from tumors, spleens and LNs were prepared, and Treg cells sorted as above. 1000 GFP+ Tregs were sorted directly into LoBind Tubes (Eppendorf) containing 6 µl of Arcturus PicoPure DNA extraction buffer (Thermo Scientific), and incubated for 3 hrs at 65°C followed by 10 min at 95°C. We then amplified the DNA using specific primers (Dataset S8) and sequenced the PCR product. To assess DNA editing efficacy we used the TIDE algorithm (Brinkman et al., 2014) to decompose the sequence. We eliminated from the analysis samples in which DNA editing efficacy was less than 50%. Plasmids are available on Addgene under IDs 112914 and 116926.

Lentivirus cloning and production

The U6 promoter-Filler-gRNA scaffold-cPPT-PGK promoter fragment was cloned from LentiCRISPR to pLKO.3G utilizing restriction enzyme sites MluI and BamHI. mRFP1 was used to substitute GFP sequence on pLKO.3G vector by Gibson Assembly. Annealing and phosphorylation of sgRNA were conducted following protocol described by Shalem et al (2014). Lentivirus was packaged using HEK293T cells. Briefly, 3.8 million cells in 10 ml DMEM supplemented with 10% fetal bovine serum were plated into 10 cm dishes 24 hours before transfection. For each dish of cells, 9 µg of psPAX2 plasmids, 1 µg of pVSVG and 9 µg of

pLKO.3G were used to co-transfect cells using TransIT-293 (Mirus Bio) reagent following manufacturer's instruction. Cells were changed with lentivirus collection medium (DMEM supplemented with 10% FBS, 1% BSA, 3 mM L-Glutamine and 100 U/ml penicillin-streptomycin) about 18 hrs post transfection. Supernatant was collected and filtered through a 0.45 μm filter 30 hrs after medium change. Lentivirus was concentrated using PEG-it lentivirus concentration reagent (System Biosciences). The infectious activity of the lentivirus on HEK293T cells was measured before use.

LSK Bone marrow chimeras

Bone marrow cells were prepared from the tibia and fiber bones of the donor mice without using FBS. Lineage$^-$ (anti-CD4, CD8, CD11b, CD11c, CD19, Gr1, NK1.1, Ter119), c-Kit$^+$ and Sca-1$^+$ LSK cells were sorted using a MoFlo Astrios cytometer after enrichment of lineage$^-$ populations using biotinylated antibodies and Dynabeads Biotin Binder (Thermo Fisher Scientific). Sorted LSK cells were cultured overnight in StemPro-34 SFM medium (Gibco), supplemented with 100 ng/mL recombinant murine SCF, Flt3 ligand, IL-7, and TPO (all from PeproTech). Cells were then spin infected with concentrated lentiviruses at 650 g for 90 mins at 37 °C in 96-well non-treated plate coated with 100 μg/mL Retronectin (Takara). Fresh complete medium was used to replace the medium 1 hour after the spin. The next day, 50,000 LSK cells were injected into irradiated (1,000 cGy) recipient mice intravenously via retro-orbital injections. Host mice were given sulfamethoxazole in water during the first two weeks after LSK cell transfer. Mice were sacrificed 10 to 12 weeks after transfer for further analysis.

Statistical Analyses

Statistical significance, indicated by asterisks, was determined by Student t test (two-tailed, unpaired). P values < 0.05 were considered significant: *P < 0.05, **P < 0.01, ***P < 0.001.

Chapter 6: General Discussion

This study looks at two opposite aspects of immunity that mirror each other. On the one hand, heterogeneity in Teff cells developed to better respond to pathogens and tumors. On the other hand, Tregs, which can directly suppress Teff cells to dampen immune responses, needed to adapt to this heterogeneity. Not surprisingly, Tregs can "hijack" some gene modules and express "Th" transcription factors such as T-bet or RORγt (Levine et al., 2017 and Sefik et al., 2015). Our preliminary data (not shown in thesis) suggest that Tregs adapt and change their transcriptome depending on the type of immune response they need to suppress. For example, upon *Salmonella* or *Citrobacter* infections, Tregs acquire many inflammatory genes and Type I or Type III effector molecules that perhaps make them more efficient at suppressing these responses. Conversely, infection with *Nippostrongylus* or *Heligmosomoides* leads to many *Gata3*+ cells that express some Type II effector molecules.

In the first study, we used scRNA-seq to look at CD4 T effector cells heterogeneity *in vivo*. The notion of T helper subsets, i.e. that different subpopulations of T cells perform different effector functions, has existed for more than 30 years, and has been rarely challenged since then. New T cell subsets such as Th17 and Tfh have emerged and added to the plethora of proposed T cell subsets. These T cell subsets were usually discovered based on the cytokines they produced, and their identity was thus inferred from their cytokine production profile. Transcription factors that were unique to each cytokine producing subset and regulated the production of cytokines were soon discovered, which led scientists to believe that these subsets are truly unique and distinct cell types.

Despite many reports on T cell plasticity and T cells that co-express cytokines from different "Th programs" most immunologists still use discrete terms to describe T helper cells,

and these distinctions have even infiltrated other cell types such as macrophages (M1 and M2) and ILCs (ILC1, ILC2 and ILC3). It is only human nature to classify objects into categories (Foroni et al., 2012). However, psychologists showed that the application of category labels distorts the judgements of the objects they describe (Foroni et al.,2012). This is why we think it is important to not assign labels to non-discrete states.

Here, we propose a model that is not based on discrete T cell states, but rather a continuum of states that is defined by contributions of various gene modules that are not necessarily driven by classic Th gene programs. We speculate, that just like subatomic particles in quantum mechanics, a cell is not either a Th1 or a Th17, but is rather on a probability distribution function that defines its probability to be in one state or the other, as well as its velocity, i.e. its potential to become another state (plasticity).

Interestingly, we observed that most T effector cells from different infections clustered separately from the effector T cells from SPF mice, which implied that the "infected" effector T cells are at a different transcriptional state than the "steady-state" effector T cells. It is an interesting observation, because T effector cells are already highly activated (express *Cd44*, *Cd69* and other markers), so this implies a "super-activation" state during infection. Cells that perform the same function under the "Th dogma" such as IL-17A-producing cells are transcriptionally distinct between SPF and infected mice, as they are "super-activated" during infection. We also noticed a difference between IL-17 or IFN-γ producers from different pathogens, implying pathogen-specific gene signature.

It is interesting to note that every infection we tried produced an IFN-γ response. We thus speculate that IFN- γ is basal response to pathogens and microbiota in the gut. At least in

C57BL/6 mice in our colony and in Jackson Laboratories most of the mice are “Th1-like” at steady state (SPF) and predominantly express T-bet and IFN-gamma in the colon with very few IL-17 or IL-4 expressing cells. Since BALB/c mice are known to have a more “Type II”-like skewed immune system, it would be of interest to explore their Teff transcriptional landscape and compare it to the results we observed in C57BL/6 mice (Hsieh et al., 1995). IFN- γ response has also been shown to be dependent on infectious dose. For example, *Trichuris muris* infection at low doses (40 eggs) causes chronic Type I response , whereas high doses (400 eggs) elicit a strong Th2 response that leads to parasite expulsion by day 21 post-infection (Klememtowicz et al., 2012). It would be of interest to see explore how the transcriptional landscape of Teffs changes at different doses of the same pathogen.

We have also found that Th cells diversify their transcriptome upon infection, as determined by the Gini-Simpson index (data not shown). The classic Th subset model might suggest that upon infection with a particular pathogen, T cells get polarized towards a particular response that is defined by a constrained set of cytokines that are helpful in eliminating that particular pathogen (Th1 towards viruses intracellular pathogens, Th2 towards helminths, Th17 towards extracellular bacteria). Our data suggest that, instead, effector T cells diversify their transcriptome, perhaps to make sure that “all bases are covered”. For example, a response to Salmonella elicits cells that produce IL-17A, IFN- γ, CCL5 and cytotoxicity mediators such as granzymes.

We found many genes that are anti-apoptotic (*Bcl2*) or pro-survival (*Zfp36l2*, *Fgl2, Lmna*), mainly found in the gene modules M14-EaAct and M12-CTX/Ccl5, are predominantly expressed in CD4 T cells from SPF mice. Many of these genes are targets of IL-7, so it is possible

that IL-7 is required for maintenance of LP CD4 T cells. We speculate that CD4 T cells in the LP under SPF conditions are activated by either commensal bacteria or food antigens, but do not develop a strong response due to various factors such as Treg inhibition or anergy. However, they require survival signals to ensure that they do not undergo apoptosis despite this partial activation. Interestingly, we found that many genes that are enriched in CD4 T cells from SPF mice such as *Vps37b*, *Junb*, *Fos*, *Zfp36l2* and *Dusp1*, are also enriched in unstimulated ILCs compared to ILCs from IL23-treated mice (Bielecki and Riesenfeld, 2018). This suggests a common quiescence signature.

In addition to the continuum of T cell states, we found some novel and more discrete cell clusters from the scRNA-seq dataset. One particularly intriguing one included Teffs that express MHC class II and other myeloid genes. While MHC class II expression has been extensively shown in human activated T lymphocytes, its expression in mouse T cells is still controversial (Holling et al.,2004). Recently, however, the Tabula Muris consortium (2018) has identified a very small population of T cells that express MHC class II. Moreover, myeloid-specific genes such as *Lyz2* have not been shown previously to be expressed in T cells from any organism.

Human T cells can actually process and present peptides and act as APCs, although they are not as efficient in capturing antigen as APCs are (Lanzavecchia et al.,1988). Additionally, Type III innate lymphoid cells (ILC3s), which are closely related to T cells, were shown to upregulate MHC class II and present antigens in the intestine, thus dampening CD4+ T cell responses to commensal bacteria (Hepworth et al., 2013). Recently, Akkaya et al. (2019) also showed that Tregs can deplete peptide-MHC class II complexes from DCs via trigocytosis.

Under highly inflammatory conditions such as *Salmonella typhimurium* infection, CD4 T cells interact very closely with MHC class II. We hypothesize that some CD4 T cells pick up MHC class II protein from the APCs via trigocytosis, as has been previously shown (Akkaya et al., 2019). Others, or perhaps the same cells, upregulate myeloid genes that are used for antigen presentation or bactericidal genes. Our data suggest that MHC class II may downregulate the expression of other myeloid genes in a negative feedback loop. It is possible that under highly inflammatory conditions effector T cells acquire these non-T cell function as a last resort to fight a potentially deadly infection. MHC class II upregulation can perhaps help with antigen presentation without the help of APCs (which Salmonella can directly infect and kill, thus depleting their pool), while bactericidal proteins can be secreted to directly kill bacteria.

Intriguingly, some leukemias such as Mixed-Phenotype Acute Leukemia (MPAL) express both T cell and myeloid markers, and their cell of origin remains unknown (Charles et al., 2017). It is possible that, if myeloid-like T cells are present in humans after infection, these cells can give rise to MPAL. This population may also correspond to unusual CD3+CD14+ cells reported by Burel et al. (2018) in human blood, which increase upon infection.

In a separate study, we examined Treg heterogeneity between tissues and disease states. We found a set of genes that were differentially expressed between Tregs in tumors and Tregs in other tissues in both mice and men and verified that some of these genes are important for Treg accumulation in tumors. Tregs are a very heterogenous population just like T effector cells, and they can take on different phenotypes depending on environmental cues and the tissue they are in. We found that in the tumor context Tregs are more suppressive and are more “Type I and Type III-like”, as they overexpress suppressive genes such as *Il10* as well as

genes that are important in Type I responses (*Cxcr3* and *Il12rb1)* and Type III responses (*Rorc* and *Il17a)*. Using CRISPR/Cas9 we found 3 novel genes that upon loss of function in Tregs caused their decreased accumulation in tumors: *Tnfrsf8* (encoding for CD30) and *Cxcr3* are both surface receptors that are expressed by activated T cells. *Samsn1*, which encodes a scaffold protein/adaptor, does not have a known immune function and is intracellular. It would be interesting to explore its role in Treg biology by creating a Treg specific knockout mouse model.

These studies could have therapeutic consequences. Currently, auto-immune diseases as well as cancer are classified and treated according to their Th subtype. For example, anti-TNF-α therapy is given to rheumatoid arthritis patients, and anti-IL17 has recently been approved for severe psoriasis. Perhaps new therapies such as combination of biologics that target both "poles" of the Th continuum, or small molecules that target upstream regulatory proteins that drive the T cell continuum, should be considered. For example, IL17-producing cells can also secrete TNF-α and IFN-γ, so targeting IL17 alone may not be as efficient as targeting upstream regulators or both Type I and Type III cytokines.

We also think that targeting the novel genes that we found to be important for Treg acuumulation in tumors could have therapeutic benefits. Monoclonal antibodies targeting CXCR3 or CD30 might prevent Treg accumulation in tumor, and thus be a viable therapeutic strategy in cancer patients. SAMSN1, though an unknown intracellular target, could also be targeted by other means such as small molecule inhibitors or CRISPR/Cas9 gene editing.

References

Abbas, A. K., K. M. Murphy and A. Sher (1996). "Functional diversity of helper T lymphocytes." Nature 383(6603): 787-793.

Acosta-Rodriguez, E. V., L. Rivino, J. Geginat, D. Jarrossay, M. Gattorno, A. Lanzavecchia, F. Sallusto and G. Napolitani (2007). "Surface phenotype and antigenic specificity of human interleukin 17-producing T helper memory cells." Nat Immunol 8(6): 639-646.

Aggarwal, S., N. Ghilardi, M. H. Xie, F. J. de Sauvage and A. L. Gurney (2003). "Interleukin-23 promotes a distinct CD4 T cell activation state characterized by the production of interleukin-17." J Biol Chem 278(3): 1910-1914.

Ahmadzadeh, M. and D. L. Farber (2002). "Functional plasticity of an antigen-specific memory CD4 T cell population." Proc Natl Acad Sci U S A 99(18): 11802-11807.

Akbar, A. N., L. Terry, A. Timms, P. C. Beverley and G. Janossy (1988). "Loss of CD45R and gain of UCHL1 reactivity is a feature of primed T cells." J Immunol 140(7): 2171-2178.

Akkaya, B., Oya, Y., Akkaya, M., Al Souz, J., Holstein, A. H., Kamenyeva, O., ... & Shevach, E. M. (2019). Regulatory T cells mediate specific suppression by depleting peptide–MHC class II from dendritic cells. *Nature immunology*, 1.

Armitage, R. J., W. C. Fanslow, L. Strockbine, T. A. Sato, K. N. Clifford, B. M. Macduff, D. M. Anderson, S. D. Gimpel, T. Davis-Smith, C. R. Maliszewski and et al. (1992). "Molecular and biological characterization of a murine ligand for CD40." Nature 357(6373): 80-82.

Arthur, R. P. and D. Mason (1986). "T cells that help B cell responses to soluble antigen are distinguishable from those producing interleukin 2 on mitogenic or allogeneic stimulation." J Exp Med 163(4): 774-786.

Atarashi, K., W. Suda, C. Luo, T. Kawaguchi, I. Motoo, S. Narushima, Y. Kiguchi, K. Yasuma, E. Watanabe, T. Tanoue, C. A. Thaiss, M. Sato, K. Toyooka, H. S. Said, H. Yamagami, S. A. Rice, D. Gevers, R. C. Johnson, J. A. Segre, K. Chen, J. K. Kolls, E. Elinav, H. Morita, R. J. Xavier, M. Hattori and K. Honda (2017). "Ectopic colonization of oral bacteria in the intestine drives TH1 cell induction and inflammation." Science 358(6361): 359-365.

Atarashi, K., T. Tanoue, M. Ando, N. Kamada, Y. Nagano, S. Narushima, W. Suda, A. Imaoka, H. Setoyama, T. Nagamori, E. Ishikawa, T. Shima, T. Hara, S. Kado, T. Jinnohara, H. Ohno, T. Kondo, K. Toyooka, E. Watanabe, S. Yokoyama, S. Tokoro, H. Mori, Y. Noguchi, H. Morita, Ivanov, II, T. Sugiyama, G. Nunez, J. G. Camp, M. Hattori, Y. Umesaki and K. Honda (2015). "Th17 Cell Induction by Adhesion of Microbes to Intestinal Epithelial Cells." Cell 163(2): 367-380.

Aujla, S. J., Y. R. Chan, M. Zheng, M. Fei, D. J. Askew, D. A. Pociask, T. A. Reinhart, F. McAllister, J. Edeal, K. Gaus, S. Husain, J. L. Kreindler, P. J. Dubin, J. M. Pilewski, M. M. Myerburg, C. A. Mason, Y. Iwakura and J. K. Kolls (2008). "IL-22 mediates mucosal host defense against Gram-negative bacterial pneumonia." Nat Med 14(3): 275-281.

Azizi, E., A. J. Carr, G. Plitas, A. E. Cornish, C. Konopacki, S. Prabhakaran, J. Nainys, K. Wu, V. Kiseliovas, M. Setty, K. Choi, R. M. Fromme, P. Dao, P. T. McKenney, R. C. Wasti, K. Kadaveru, L. Mazutis, A. Y. Rudensky and D. Pe'er (2018). "Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment." Cell 174(5): 1293-1308 e1236.

Benoist, C. and D. Mathis (2012). "Treg cells, life history, and diversity." Cold Spring Harb Perspect Biol 4(9): a007021.

Bettelli, E., Y. Carrier, W. Gao, T. Korn, T. B. Strom, M. Oukka, H. L. Weiner and V. K. Kuchroo (2006). "Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells." Nature 441(7090): 235-238.

Bielecki, Piotr, Samantha J. Riesenfled, Monika S. Kowalczyk, Maria C. Amezcua Vesely, Lina Kroehling, Parastou Yaghoubi, Danielle Dionne et al. (2018). "Skin inflammation driven by differentiation of quiescent tissue-resident ILCs into a spectrum of pathogenic effectors." *bioRxiv* 461228.

Billiau, A., H. Heremans, F. Vandekerckhove, R. Dijkmans, H. Sobis, E. Meulepas and H. Carton (1988). "Enhancement of experimental allergic encephalomyelitis in mice by antibodies against IFN-gamma." J Immunol 140(5): 1506-1510.

Bos, P. D., G. Plitas, D. Rudra, S. Y. Lee and A. Y. Rudensky (2013). "Transient regulatory T cell ablation deters oncogene-driven breast cancer and enhances radiotherapy." J Exp Med 210(11): 2435-2466.

Bottomly, K. (1988). "A functional dichotomy in CD4+ T lymphocytes." Immunol Today 9(9): 268-274.

Breitfeld, D., L. Ohl, E. Kremmer, J. Ellwart, F. Sallusto, M. Lipp and R. Forster (2000). "Follicular B helper T cells express CXC chemokine receptor 5, localize to B cell follicles, and support immunoglobulin production." J Exp Med 192(11): 1545-1552.

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

Brustle, A., S. Heink, M. Huber, C. Rosenplanter, C. Stadelmann, P. Yu, E. Arpaia, T. W. Mak, T. Kamradt and M. Lohoff (2007). "The development of inflammatory T(H)-17 cells requires interferon-regulatory factor 4." Nat Immunol 8(9): 958-966.

Buhlmann, J. E., M. Gonzalez, B. Ginther, A. Panoskaltsis-Mortari, B. R. Blazar, D. L. Greiner, A. A. Rossini, R. Flavell and R. J. Noelle (1999). "Cutting edge: sustained expansion of CD8+ T cells requires CD154 expression by Th cells in acute graft versus host disease." J Immunol 162(8): 4373-4376.

Burel, J. G., Pomaznoy, M., Arlehamn, C. S. L., Weiskopf, D., da Silva Antunes, R., Schulten, V., ... & Premawansa, G. (2018). No cell is an island: circulating T cell: monocyte complexes are markers of immune perturbations. *bioRxiv*, 411553.

Burzyn, D., W. Kuswanto, D. Kolodin, J. L. Shadrach, M. Cerletti, Y. Jang, E. Sefik, T. G. Tan, A. J. Wagers, C. Benoist and D. Mathis (2013). "A special population of regulatory T cells potentiates muscle repair." Cell 155(6): 1282-1295.

Butler, A., P. Hoffman, P. Smibert, E. Papalexi and R. Satija (2018). "Integrating single-cell transcriptomic data across different conditions, technologies, and species." Nat Biotechnol 36(5): 411-420.

Camberis, M., G. Le Gros and J. Urban, Jr. (2003). "Animal model of Nippostrongylus brasiliensis and Heligmosomoides polygyrus." Curr Protoc Immunol Chapter 19: Unit 19 12.

Cantor, H. and E. A. Boyse (1975). "Functional subclasses of T lymphocytes bearing different Ly antigens. II. Cooperation between subclasses of Ly+ cells in the generation of killer activity." J Exp Med 141(6): 1390-1399.

Charles, N. J. and D. F. Boyer (2017). "Mixed-Phenotype Acute Leukemia: Diagnostic Criteria and Pitfalls." Arch Pathol Lab Med 141(11): 1462-1468.

Chen, S., N. E. Sanjana, K. Zheng, O. Shalem, K. Lee, X. Shi, D. A. Scott, J. Song, J. Q. Pan, R. Weissleder, H. Lee, F. Zhang and P. A. Sharp (2015). "Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis." Cell 160(6): 1246-1260.

Cher, D. J. and T. R. Mosmann (1987). "Two types of murine helper T cell clone. II. Delayed-type hypersensitivity is mediated by TH1 clones." J Immunol 138(11): 3688-3694.

Chtanova, T., R. A. Kemp, A. P. Sutherland, F. Ronchese and C. R. Mackay (2001). "Gene microarrays reveal extensive differential gene expression in both CD4(+) and CD8(+) type 1 and type 2 T cells." J Immunol 167(6): 3057-3063.

Chtanova, T., S. G. Tangye, R. Newton, N. Frank, M. R. Hodge, M. S. Rolph and C. R. Mackay (2004). "T follicular helper cells express a distinctive transcriptional profile, reflecting their role as non-Th1/Th2 effector cells that provide help for B cells." J Immunol 173(1): 68-78.

Ciofani, M., A. Madar, C. Galan, M. Sellars, K. Mace, F. Pauli, A. Agarwal, W. Huang, C. N. Parkhurst, M. Muratet, K. M. Newberry, S. Meadows, A. Greenfield, Y. Yang, P. Jain, F. K. Kirigin, C. Birchmeier, E. F. Wagner, K. M. Murphy, R. M. Myers, R. Bonneau and D. R. Littman (2012). "A validated regulatory network for Th17 cell specification." Cell 151(2): 289-303.

Cipolletta, D., M. Feuerer, A. Li, N. Kamei, J. Lee, S. E. Shoelson, C. Benoist and D. Mathis (2012). "PPAR-gamma is a major driver of the accumulation and phenotype of adipose tissue Treg cells." Nature 486(7404): 549-553.

Coffman, R. L., B. W. Seymour, S. Hudak, J. Jackson and D. Rennick (1989). "Antibody to interleukin-5 inhibits helminth-induced eosinophilia in mice." Science 245(4915): 308-310.

Collins, J. W., K. M. Keeney, V. F. Crepin, V. A. Rathinam, K. A. Fitzgerald, B. B. Finlay and G. Frankel (2014). "Citrobacter rodentium: infection, inflammation and the microbiota." Nat Rev Microbiol 12(9): 612-623.

Corbett, T. H., D. P. Griswold, Jr., B. J. Roberts, J. C. Peckham and F. M. Schabel, Jr. (1975). "Tumor induction relationships in development of transplantable cancers of the colon in mice for chemotherapy assays, with a note on carcinogen structure." Cancer Res 35(9): 2434-2439.

Cosgrove, D., D. Gray, A. Dierich, J. Kaufman, M. Lemeur, C. Benoist and D. Mathis (1991). "Mice lacking MHC class II molecules." Cell 66(5): 1051-1066.

Cosmi, L., L. Maggi, V. Santarlasci, M. Capone, E. Cardilicchia, F. Frosali, V. Querci, R. Angeli, A. Matucci, M. Fambrini, F. Liotta, P. Parronchi, E. Maggi, S. Romagnani and F. Annunziato (2010). "Identification of a novel subset of human circulating memory CD4(+) T cells that produce both IL-17A and IL-4." J Allergy Clin Immunol 125(1): 222-230 e221-224.

Crotty, S. (2014). "T follicular helper cell differentiation, function, and roles in disease." Immunity 41(4): 529-542.

Crotty, S. (2015). "A brief history of T cell help to B cells." Nat Rev Immunol 15(3): 185-189.

Cua, D. J., J. Sherlock, Y. Chen, C. A. Murphy, B. Joyce, B. Seymour, L. Lucian, W. To, S. Kwan, T. Churakova, S. Zurawski, M. Wiekowski, S. A. Lira, D. Gorman, R. A. Kastelein and J. D. Sedgwick (2003). "Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain." Nature 421(6924): 744-748.

Dankort, D., D. P. Curley, R. A. Cartlidge, B. Nelson, A. N. Karnezis, W. E. Damsky, Jr., M. J. You, R. A. DePinho, M. McMahon and M. Bosenberg (2009). "Braf(V600E) cooperates with Pten loss to induce metastatic melanoma." Nat Genet 41(5): 544-552.

Dardalhon, V., A. Awasthi, H. Kwon, G. Galileos, W. Gao, R. A. Sobel, M. Mitsdoerffer, T. B. Strom, W. Elyaman, I. C. Ho, S. Khoury, M. Oukka and V. K. Kuchroo (2008). "IL-4 inhibits TGF-beta-induced Foxp3+ T cells and, together with TGF-beta, generates IL-9+ IL-10+ Foxp3(-) effector T cells." Nat Immunol 9(12): 1347-1355.

Davis, M. M., J. J. Boniface, Z. Reich, D. Lyons, J. Hampl, B. Arden and Y. Chien (1998). "Ligand recognition by alpha beta T cell receptors." Annu Rev Immunol 16: 523-544.

Duerr, R. H., K. D. Taylor, S. R. Brant, J. D. Rioux, M. S. Silverberg, M. J. Daly, A. H. Steinhart, C. Abraham, M. Regueiro, A. Griffiths, T. Dassopoulos, A. Bitton, H. Yang, S. Targan, L. W. Datta, E. O. Kistner, L. P. Schumm, A. T. Lee, P. K. Gregersen, M. M. Barmada, J. I. Rotter, D. L. Nicolae and J. H. Cho (2006). "A genome-wide association study identifies IL23R as an inflammatory bowel disease gene." Science 314(5804): 1461-1463.

Duhen, T. and D. J. Campbell (2014). "IL-1beta promotes the differentiation of polyfunctional human CCR6+CXCR3+ Th1/17 cells that are specific for pathogenic and commensal microbes." J Immunol 193(1): 120-129.

Duhen, T., R. Geiger, D. Jarrossay, A. Lanzavecchia and F. Sallusto (2009). "Production of interleukin 22 but not interleukin 17 by a subset of human skin-homing memory T cells." Nat Immunol 10(8): 857-863.

Ettinger, R., G. P. Sims, A. M. Fairhurst, R. Robbins, Y. S. da Silva, R. Spolski, W. J. Leonard and P. E. Lipsky (2005). "IL-21 induces differentiation of human naive and memory B cells into antibody-secreting plasma cells." J Immunol 175(12): 7867-7879.

Fernandez-Botran, R., V. M. Sanders, T. R. Mosmann and E. S. Vitetta (1988). "Lymphokine-mediated regulation of the proliferative response of clones of T helper 1 and T helper 2 cells." J Exp Med 168(2): 543-558.

Feuerer, M., L. Herrero, D. Cipolletta, A. Naaz, J. Wong, A. Nayer, J. Lee, A. B. Goldfine, C. Benoist, S. Shoelson and D. Mathis (2009). "Lean, but not obese, fat is enriched for a unique population of regulatory T cells that affect metabolic parameters." Nat Med 15(8): 930-939.

Fidler, I. J. (1975). "Biological behavior of malignant melanoma cells correlated to their survival in vivo." Cancer Res 35(1): 218-224.

Firestein, G. S., W. D. Roeder, J. A. Laxer, K. S. Townsend, C. T. Weaver, J. T. Hom, J. Linton, B. E. Torbett and A. L. Glasebrook (1989). "A new murine CD4+ T cell subset with an unrestricted cytokine profile." J Immunol 143(2): 518-525.

Foroni, F. and M. Rothbart (2013). "Abandoning a label doesn't make it disappear: The perseverance of labeling effects." J Exp Soc Psychol 49(1): 126-131.

Gagliani, N., M. C. Amezcua Vesely, A. Iseppon, L. Brockmann, H. Xu, N. W. Palm, M. R. de Zoete, P. Licona-Limon, R. S. Paiva, T. Ching, C. Weaver, X. Zi, X. Pan, R. Fan, L. X. Garmire, M. J. Cotton, Y. Drier, B. Bernstein, J. Geginat, B. Stockinger, E. Esplugues, S. Huber and R. A. Flavell (2015). "Th17 cells transdifferentiate into regulatory T cells during resolution of inflammation." Nature 523(7559): 221-225.

Gajewski, T. F. and F. W. Fitch (1988). "Anti-proliferative effect of IFN-gamma in immune regulation. I. IFN-gamma inhibits the proliferation of Th2 but not Th1 murine helper T lymphocyte clones." J Immunol 140(12): 4245-4252.

Galon, J., A. Costes, F. Sanchez-Cabo, A. Kirilovsky, B. Mlecnik, C. Lagorce-Pages, M. Tosolini, M. Camus, A. Berger, P. Wind, F. Zinzindohoue, P. Bruneval, P. H. Cugnenc, Z. Trajanoski, W. H. Fridman and F. Pages (2006). "Type, density, and location of immune cells within human colorectal tumors predict clinical outcome." Science 313(5795): 1960-1964.

Gaublomme, J. T., N. Yosef, Y. Lee, R. S. Gertner, L. V. Yang, C. Wu, P. P. Pandolfi, T. Mak, R. Satija, A. K. Shalek, V. K. Kuchroo, H. Park and A. Regev (2015). "Single-Cell Genomics Unveils Critical Regulators of Th17 Cell Pathogenicity." Cell 163(6): 1400-1412.

Geginat, J., Paroni, M., Maglie, S., Alfen, J. S., Kastirr, I., Gruarin, P., ... & Abrignani, S. (2014). “Plasticity of human CD4 T cell subsets”. *Frontiers in immunology*, *5*, 630.

Geha, R. S., L. D. Notarangelo, J. L. Casanova, H. Chapel, M. E. Conley, A. Fischer, L. Hammarstrom, S. Nonoyama, H. D. Ochs, J. M. Puck, C. Roifman, R. Seger, J. Wedgwood and C. International Union of Immunological Societies Primary Immunodeficiency Diseases Classification (2007). "Primary immunodeficiency diseases: an update from the International Union of Immunological Societies Primary Immunodeficiency Diseases Classification Committee." J Allergy Clin Immunol 120(4): 776-794.

Godec, J., G. S. Cowley, R. A. Barnitz, O. Alkan, D. E. Root, A. H. Sharpe and W. N. Haining (2015). "Inducible RNAi in vivo reveals that the transcription factor BATF is required to initiate but not maintain CD8+ T-cell effector differentiation." Proc Natl Acad Sci U S A 112(2): 512-517.

Godfrey, D. I. and A. Zlotnik (1993). "Control points in early T-cell development." Immunol Today 14(11): 547-553.

Gordon, S. (2003). "Alternative activation of macrophages." Nat Rev Immunol 3(1): 23-35.

Gronke, K., P. P. Hernandez, J. Zimmermann, C. S. N. Klose, M. Kofoed-Branzk, F. Guendel, M. Witkowski, C. Tizian, L. Amann, F. Schumacher, H. Glatt, A. Triantafyllopoulou and A. Diefenbach (2019). "Interleukin-22 protects intestinal stem cells against genotoxic stress." Nature 566(7743): 249-253.

Haghverdi, L., Buettner, F., & Theis, F. J. (2015). “Diffusion maps for high-dimensional single-cell analysis of differentiation data.” *Bioinformatics*, *31*(18), 2989-2998.

Harriman, G. R., D. Y. Kunimoto, J. F. Elliott, V. Paetkau and W. Strober (1988). "The role of IL-5 in IgA B cell differentiation." J Immunol 140(9): 3033-3039.

Hartigan, J.A. and Hartigan, P.M., 1985. The dip test of unimodality. *The annals of Statistics*, *13*(1), pp.70-84.

Hegazy, A. N., M. Peine, C. Helmstetter, I. Panse, A. Frohlich, A. Bergthaler, L. Flatz, D. D. Pinschewer, A. Radbruch and M. Lohning (2010). "Interferons direct Th2 cell reprogramming to

generate a stable GATA-3(+)T-bet(+) cell subset with combined Th2 and Th1 cell functions." Immunity 32(1): 116-128.

Hepworth, M. R., L. A. Monticelli, T. C. Fung, C. G. Ziegler, S. Grunberg, R. Sinha, A. R. Mantegazza, H. L. Ma, A. Crawford, J. M. Angelosanto, E. J. Wherry, P. A. Koni, F. D. Bushman, C. O. Elson, G. Eberl, D. Artis and G. F. Sonnenberg (2013). "Innate lymphoid cells regulate CD4+ T-cell responses to intestinal commensal bacteria." Nature 498(7452): 113-117.

Hess, J., C. Ladel, D. Miko and S. H. Kaufmann (1996). "Salmonella typhimurium aroA- infection in gene-targeted immunodeficient mice: major role of CD4+ TCR-alpha beta cells and IFN-gamma in bacterial clearance independent of intracellular location." J Immunol 156(9): 3321-3326.

Hill, J. A., M. Feuerer, K. Tash, S. Haxhinasto, J. Perez, R. Melamed, D. Mathis and C. Benoist (2007). "Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature." Immunity 27(5): 786-800.

Hirota, K., J. H. Duarte, M. Veldhoen, E. Hornsby, Y. Li, D. J. Cua, H. Ahlfors, C. Wilhelm, M. Tolaini, U. Menzel, A. Garefalaki, A. J. Potocnik and B. Stockinger (2011). "Fate mapping of IL-17-producing T cells in inflammatory responses." Nat Immunol 12(3): 255-263.

Hirota, K., J. E. Turner, M. Villa, J. H. Duarte, J. Demengeot, O. M. Steinmetz and B. Stockinger (2013). "Plasticity of Th17 cells in Peyer's patches is responsible for the induction of T cell-dependent IgA responses." Nat Immunol 14(4): 372-379.

Hoiseth, S. K. and B. A. Stocker (1981). "Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines." Nature 291(5812): 238-239.

Holling, T. M., E. Schooten and P. J. van Den Elsen (2004). "Function and regulation of MHC class II molecules in T-lymphocytes: of mice and men." Hum Immunol 65(4): 282-290.

Howard, M., J. Farrar, M. Hilfiker, B. Johnson, K. Takatsu, T. Hamaoka and W. E. Paul (1982). "Identification of a T cell-derived b cell growth factor distinct from interleukin 2." J Exp Med 155(3): 914-923.

Hsieh, C. S., S. E. Macatonia, C. S. Tripp, S. F. Wolf, A. O'Garra and K. M. Murphy (1993). "Development of TH1 CD4+ T cells through IL-12 produced by Listeria-induced macrophages." Science 260(5107): 547-549.

Hsieh, C. S., Macatonia, S. E., O'Garra, A., & Murphy, K. M. (1995). T cell genetic background determines default T helper phenotype development in vitro. *Journal of Experimental Medicine*, *181*(2), 713-721.

Iijima, N. and A. Iwasaki (2014). "T cell memory. A local macrophage chemokine network sustains protective tissue-resident memory CD4 T cells." Science 346(6205): 93-98.

Ivanov, II, L. Frutos Rde, N. Manel, K. Yoshinaga, D. B. Rifkin, R. B. Sartor, B. B. Finlay and D. R. Littman (2008). "Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine." Cell Host Microbe 4(4): 337-349.

Ivanov, II, B. S. McKenzie, L. Zhou, C. E. Tadokoro, A. Lepelley, J. J. Lafaille, D. J. Cua and D. R. Littman (2006). "The orphan nuclear receptor RORgammat directs the differentiation program of proinflammatory IL-17+ T helper cells." Cell 126(6): 1121-1133.

Jacobson, N. G., S. J. Szabo, R. M. Weber-Nordt, Z. Zhong, R. D. Schreiber, J. E. Darnell, Jr. and K. M. Murphy (1995). "Interleukin 12 signaling in T helper type 1 (Th1) cells involves tyrosine phosphorylation of signal transducer and activator of transcription (Stat)3 and Stat4." J Exp Med 181(5): 1755-1762.

Janssen, E. M., E. E. Lemmens, T. Wolfe, U. Christen, M. G. von Herrath and S. P. Schoenberger (2003). "CD4+ T cells are required for secondary expansion and memory in CD8+ T lymphocytes." Nature 421(6925): 852-856.

Jennings, S. R., R. H. Bonneau, P. M. Smith, R. M. Wolcott and R. Chervenak (1991). "CD4-positive T lymphocytes are required for the generation of the primary but not the secondary CD8-positive cytolytic T lymphocyte response to herpes simplex virus in C57BL/6 mice." Cell Immunol 133(1): 234-252.

Johnston, R. J., Y. S. Choi, J. A. Diamond, J. A. Yang and S. Crotty (2012). "STAT5 is a potent negative regulator of TFH cell differentiation." J Exp Med 209(2): 243-250.

Johnston, R. J., A. C. Poholek, D. DiToro, I. Yusuf, D. Eto, B. Barnett, A. L. Dent, J. Craft and S. Crotty (2009). "Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation." Science 325(5943): 1006-1010.

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

Kang, E., G. Zhou, M. Yousefi, R. Cayrol, J. Xia and S. Gruenheid (2018). "Loss of disease tolerance during Citrobacter rodentium infection is associated with impaired epithelial differentiation and hyperactivation of T cell responses." Sci Rep 8(1): 847.

Kara, E. E., I. Comerford, K. A. Fenix, C. R. Bastow, C. E. Gregor, D. R. McKenzie and S. R. McColl (2014). "Tailored immune responses: novel effector helper T cell subsets in protective immunity." PLoS Pathog 10(2): e1003905.

Keene, J. A. and J. Forman (1982). "Helper activity is required for the in vivo generation of cytotoxic T lymphocytes." J Exp Med 155(3): 768-782.

Kelso, A. and N. M. Gough (1988). "Coexpression of granulocyte-macrophage colony-stimulating factor, gamma interferon, and interleukins 3 and 4 is random in murine alloreactive T-lymphocyte clones." Proc Natl Acad Sci U S A 85(23): 9189-9193.

Killar, L., G. MacDonald, J. West, A. Woods and K. Bottomly (1987). "Cloned, Ia-restricted T cells that do not produce interleukin 4(IL 4)/B cell stimulatory factor 1(BSF-1) fail to help antigen-specific B cells." J Immunol 138(6): 1674-1679.

Kim, J. M., J. P. Rasmussen and A. Y. Rudensky (2007). "Regulatory T cells prevent catastrophic autoimmunity throughout the lifespan of mice." Nat Immunol 8(2): 191-197.

Kisielow, P., J. A. Hirst, H. Shiku, P. C. Beverley, M. K. Hoffman, E. A. Boyse and H. F. Oettgen (1975). "Ly antigens as markers for functionally distinct subpopulations of thymus-derived lymphocytes of the mouse." Nature 253(5488): 219-220.

Klementowicz, J. E., M. A. Travis and R. K. Grencis (2012). "Trichuris muris: a model of gastrointestinal parasite infection." Semin Immunopathol 34(6): 815-828.

Kopf, M., G. Le Gros, A. J. Coyle, M. Kosco-Vilbois and F. Brombacher (1995). "Immune responses of IL-4, IL-5, IL-6 deficient mice." Immunol Rev 148: 45-69.

Korn, T., E. Bettelli, W. Gao, A. Awasthi, A. Jager, T. B. Strom, M. Oukka and V. K. Kuchroo (2007). "IL-21 initiates an alternative pathway to induce proinflammatory T(H)17 cells." Nature 448(7152): 484-487.

Krausgruber, T., C. Schiering, K. Adelmann, O. J. Harrison, A. Chomka, C. Pearson, P. P. Ahern, M. Shale, M. Oukka and F. Powrie (2016). "T-bet is a key modulator of IL-23-driven pathogenic CD4(+) T cell responses in the intestine." Nat Commun 7: 11627.

Kumaraguru, U., K. Banerjee and B. T. Rouse (2005). "In vivo rescue of defective memory CD8+ T cells by cognate helper T cells." J Leukoc Biol 78(4): 879-887.

Kunkel, E. J., J. J. Campbell, G. Haraldsen, J. Pan, J. Boisvert, A. I. Roberts, E. C. Ebert, M. A. Vierra, S. B. Goodman, M. C. Genovese, A. J. Wardlaw, H. B. Greenberg, C. M. Parker, E. C. Butcher, D. P. Andrew and W. W. Agace (2000). "Lymphocyte CC chemokine receptor 9 and epithelial thymus-expressed chemokine (TECK) expression distinguish the small intestinal immune compartment: Epithelial expression of tissue-specific chemokines as an organizing principle in regional immunity." J Exp Med 192(5): 761-768.

Kurachi, M., J. Kurachi, Z. Chen, J. Johnson, O. Khan, B. Bengsch, E. Stelekati, J. Attanasio, L. M. McLane, M. Tomura, S. Ueha and E. J. Wherry (2017). "Optimized retroviral transduction of mouse T cells for in vivo assessment of gene function." Nat Protoc 12(9): 1980-1998.

Kwon, H. K., H. M. Chen, D. Mathis and C. Benoist (2017). "Different molecular complexes that mediate transcriptional induction and repression by FoxP3." Nat Immunol 18(11): 1238-1248.

Langrish, C. L., Y. Chen, W. M. Blumenschein, J. Mattson, B. Basham, J. D. Sedgwick, T. McClanahan, R. A. Kastelein and D. J. Cua (2005). "IL-23 drives a pathogenic T cell population that induces autoimmune inflammation." J Exp Med 201(2): 233-240.

Lanzavecchia, A., E. Roosnek, T. Gregory, P. Berman and S. Abrignani (1988). "T cells can present antigens such as HIV gp120 targeted to their own surface molecules." Nature 334(6182): 530-532.

Le Gros, G., N. Schultze, S. Walti, K. Einsle, F. Finkelman, M. H. Kosco-Vilbois and C. Heusser (1996). "The development of IgE+ memory B cells following primary IgE immune responses." Eur J Immunol 26(12): 3042-3047.

Lee, Y., A. Awasthi, N. Yosef, F. J. Quintana, S. Xiao, A. Peters, C. Wu, M. Kleinewietfeld, S. Kunder, D. A. Hafler, R. A. Sobel, A. Regev and V. K. Kuchroo (2012). "Induction and molecular signature of pathogenic TH17 cells." Nat Immunol 13(10): 991-999.

Levine, A. G., Mendoza, A., Hemmers, S., Moltedo, B., Niec, R. E., Schizas, M., ... & Fujisawa, S. (2017). Stability and function of regulatory T cells expressing the transcription factor T-bet. *Nature*, *546*(7658), 421.

Lexberg, M. H., A. Taubner, I. Albrecht, I. Lepenies, A. Richter, T. Kamradt, A. Radbruch and H. D. Chang (2010). "IFN-gamma and IL-12 synergize to convert in vivo generated Th17 into Th1/Th17 cells." Eur J Immunol 40(11): 3017-3027.

Liang, S. C., X. Y. Tan, D. P. Luxenberg, R. Karim, K. Dunussi-Joannopoulos, M. Collins and L. A. Fouser (2006). "Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides." J Exp Med 203(10): 2271-2279.

Licona-Limon, P., J. Henao-Mejia, A. U. Temann, N. Gagliani, I. Licona-Limon, H. Ishigame, L. Hao, D. R. Herbert and R. A. Flavell (2013). "Th9 Cells Drive Host Immunity against Gastrointestinal Worm Infection." Immunity 39(4): 744-757.

Liew, F. Y. (2002). "T(H)1 and T(H)2 cells: a historical perspective." Nat Rev Immunol 2(1): 55-60.

Liew, F. Y. and C. R. Parish (1974). "Lack of a correlation between cell-mediated immunity to the carrier and the carrier-hapten helper effect." J Exp Med 139(3): 779-784.

Lin, W., D. Haribhai, L. M. Relland, N. Truong, M. R. Carlson, C. B. Williams and T. A. Chatila (2007). "Regulatory T cell development in the absence of functional Foxp3." Nat Immunol 8(4): 359-368.

Liston, A., K. M. Nutsch, A. G. Farr, J. M. Lund, J. P. Rasmussen, P. A. Koni and A. Y. Rudensky (2008). "Differentiation of regulatory Foxp3+ T cells in the thymic cortex." Proc Natl Acad Sci U S A 105(33): 11903-11908.

Locksley, R. M. (1994). "Th2 cells: help for helminths." J Exp Med 179(5): 1405-1407.

Long, S. A. and J. H. Buckner (2011). "CD4+FOXP3+ T regulatory cells in human autoimmunity: more than a numbers game." J Immunol 187(5): 2061-2066.

Lu, B., P. Zagouras, J. E. Fischer, J. Lu, B. Li and R. A. Flavell (2004). "Kinetic analysis of genomewide gene expression reveals molecule circuitries that control T cell activation and Th1/2 differentiation." Proc Natl Acad Sci U S A 101(9): 3023-3028.

Ma, H., W. Tao and S. Zhu (2019). "T lymphocytes in the intestinal mucosa: defense and tolerance." Cell Mol Immunol.

MacDonald, H. R., R. C. Budd and J. C. Cerottini (1990). "Pgp-1 (Ly 24) as a marker of murine memory T lymphocytes." Curr Top Microbiol Immunol 159: 97-109.

Mackay, C. R., W. L. Marston and L. Dudler (1990). "Naive and memory T cells show distinct pathways of lymphocyte recirculation." J Exp Med 171(3): 801-817.

Macosko, E. Z., A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev and S. A. McCarroll (2015). "Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets." Cell 161(5): 1202-1214.

Mangan, P. R., L. E. Harrington, D. B. O'Quinn, W. S. Helms, D. C. Bullard, C. O. Elson, R. D. Hatton, S. M. Wahl, T. R. Schoeb and C. T. Weaver (2006). "Transforming growth factor-beta induces development of the T(H)17 lineage." Nature 441(7090): 231-234.

Manguso, R. T., H. W. Pope, M. D. Zimmer, F. D. Brown, K. B. Yates, B. C. Miller, N. B. Collins, K. Bi, M. W. LaFleur, V. R. Juneja, S. A. Weiss, J. Lo, D. E. Fisher, D. Miao, E. Van Allen, D. E. Root, A. H. Sharpe, J. G. Doench and W. N. Haining (2017). "In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target." Nature 547(7664): 413-418.

Marrack, P. C. and J. W. Kappler (1975). "Antigen-specific and nonspecific mediators of T cell/B cell cooperation. I. Evidence for their production by different T cells." J Immunol 114(3): 1116-1125.

Martin-Orozco, N., P. Muranski, Y. Chung, X. O. Yang, T. Yamazaki, S. Lu, P. Hwu, N. P. Restifo, W. W. Overwijk and C. Dong (2009). "T helper 17 cells promote cytotoxic T cell activation in tumor immunity." Immunity 31(5): 787-798.

Masopust, D., D. Choo, V. Vezys, E. J. Wherry, J. Duraiswamy, R. Akondy, J. Wang, K. A. Casey, D. L. Barber, K. S. Kawamura, K. A. Fraser, R. J. Webby, V. Brinkmann, E. C. Butcher, K. A. Newell and R. Ahmed (2010). "Dynamic T cell migration program provides resident memory within intestinal epithelium." J Exp Med 207(3): 553-564.

Maynard, C. L. and C. T. Weaver (2009). "Intestinal effector T cells in health and disease." Immunity 31(3): 389-400.

Micosse, C., L. von Meyenn, O. Steck, E. Kipfer, C. Adam, C. Simillion, S. M. Seyed Jafari, P. Olah, N. Yawlkar, D. Simon, L. Borradori, S. Kuchen, D. Yerly, B. Homey, C. Conrad, B. Snijder,

M. Schmidt and C. Schlapbach (2019). "Human "TH9" cells are a subpopulation of PPAR-gamma(+) TH2 cells." Sci Immunol 4(31).

Miller, J. F. (1961). "Immunological function of the thymus." Lancet 2(7205): 748-749.

Miller, J. F. and G. F. Mitchell (1968). "Cell to cell interaction in the immune response. I. Hemolysin-forming cells in neonatally thymectomized mice reconstituted with thymus or thoracic duct lymphocytes." J Exp Med 128(4): 801-820.

Mills, C. D., K. Kincaid, J. M. Alt, M. J. Heilman and A. M. Hill (2000). "M-1/M-2 macrophages and the Th1/Th2 paradigm." J Immunol 164(12): 6166-6173.

Montane, J., L. Bischoff, G. Soukhatcheva, D. L. Dai, G. Hardenberg, M. K. Levings, P. C. Orban, T. J. Kieffer, R. Tan and C. B. Verchere (2011). "Prevention of murine autoimmune diabetes by CCL22-mediated Treg recruitment to the pancreatic islets." J Clin Invest 121(8): 3024-3028.

Morita, S., T. Kojima and T. Kitamura (2000). "Plat-E: an efficient and stable system for transient packaging of retroviruses." Gene Ther 7(12): 1063-1066.

Mosmann, T. R., H. Cherwinski, M. W. Bond, M. A. Giedlin and R. L. Coffman (1986). "Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins." J Immunol 136(7): 2348-2357.

Mosmann, T. R. and R. L. Coffman (1989). "TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties." Annu Rev Immunol 7: 145-173.

Mucida, D., Y. Park, G. Kim, O. Turovskaya, I. Scott, M. Kronenberg and H. Cheroutre (2007). "Reciprocal TH17 and regulatory T cell differentiation mediated by retinoic acid." Science 317(5835): 256-260.

Mukasa, R., A. Balasubramani, Y. K. Lee, S. K. Whitley, B. T. Weaver, Y. Shibata, G. E. Crawford, R. D. Hatton and C. T. Weaver (2010). "Epigenetic instability of cytokine and transcription factor gene loci underlies plasticity of the T helper 17 cell lineage." Immunity 32(5): 616-627.

Mundy, R., T. T. MacDonald, G. Dougan, G. Frankel and S. Wiles (2005). "Citrobacter rodentium of mice and man." Cell Microbiol 7(12): 1697-1706.

Murphy, E., K. Shibuya, N. Hosken, P. Openshaw, V. Maino, K. Davis, K. Murphy and A. O'Garra (1996). "Reversibility of T helper 1 and 2 populations is lost after long-term stimulation." J Exp Med 183(3): 901-913.

Murphy, K. M. and B. Stockinger (2010). "Effector T cell plasticity: flexibility in the face of changing circumstances." Nat Immunol 11(8): 674-680.

Murray, P. J., J. E. Allen, S. K. Biswas, E. A. Fisher, D. W. Gilroy, S. Goerdt, S. Gordon, J. A. Hamilton, L. B. Ivashkiv, T. Lawrence, M. Locati, A. Mantovani, F. O. Martinez, J. L. Mege, D. M. Mosser, G. Natoli, J. P. Saeij, J. L. Schultze, K. A. Shirey, A. Sica, J. Suttles, I. Udalova, J. A. van Ginderachter, S. N. Vogel and T. A. Wynn (2014). "Macrophage activation and polarization: nomenclature and experimental guidelines." Immunity 41(1): 14-20.

Nair, R. P., K. C. Duffin, C. Helms, J. Ding, P. E. Stuart, D. Goldgar, J. E. Gudjonsson, Y. Li, T. Tejasvi, B. J. Feng, A. Ruether, S. Schreiber, M. Weichenthal, D. Gladman, P. Rahman, S. J. Schrodi, S. Prahalad, S. L. Guthery, J. Fischer, W. Liao, P. Y. Kwok, A. Menter, G. M. Lathrop, C. A. Wise, A. B. Begovich, J. J. Voorhees, J. T. Elder, G. G. Krueger, A. M. Bowcock, G. R. Abecasis and P. Collaborative Association Study of (2009). "Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways." Nat Genet 41(2): 199-204.

Nakanishi, Y., B. Lu, C. Gerard and A. Iwasaki (2009). "CD8(+) T lymphocyte mobilization to virus-infected tissue requires CD4(+) T-cell help." Nature 462(7272): 510-513.

Niess, J. H., F. Leithauser, G. Adler and J. Reimann (2008). "Commensal gut flora drives the expansion of proinflammatory CD4 T cells in the colonic lamina propria under normal and inflammatory conditions." J Immunol 180(1): 559-568.

Nishikawa, H. and S. Sakaguchi (2014). "Regulatory T cells in cancer immunotherapy." Curr Opin Immunol 27: 1-7.

Nurieva, R. I., Y. Chung, D. Hwang, X. O. Yang, H. S. Kang, L. Ma, Y. H. Wang, S. S. Watowich, A. M. Jetten, Q. Tian and C. Dong (2008). "Generation of T follicular helper cells is mediated by interleukin-21 but independent of T helper 1, 2, or 17 cell lineages." Immunity 29(1): 138-149.

Nurieva, R. I., Y. Chung, G. J. Martinez, X. O. Yang, S. Tanaka, T. D. Matskevitch, Y. H. Wang and C. Dong (2009). "Bcl6 mediates the development of T follicular helper cells." Science 325(5943): 1001-1005.

Openshaw, P., E. E. Murphy, N. A. Hosken, V. Maino, K. Davis, K. Murphy and A. O'Garra (1995). "Heterogeneity of intracellular cytokine synthesis at the single-cell level in polarized T helper 1 and T helper 2 populations." J Exp Med 182(5): 1357-1367.

Panduro, M., C. Benoist and D. Mathis (2016). "Tissue Tregs." Annu Rev Immunol 34: 609-633.

Panzer, M., S. Sitte, S. Wirth, I. Drexler, T. Sparwasser and D. Voehringer (2012). "Rapid in vivo conversion of effector T cells into Th2 cells during helminth infection." J Immunol 188(2): 615-623.

Park, H., Z. Li, X. O. Yang, S. H. Chang, R. Nurieva, Y. H. Wang, Y. Wang, L. Hood, Z. Zhu, Q. Tian and C. Dong (2005). "A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17." Nat Immunol 6(11): 1133-1141.

Parrish-Novak, J., S. R. Dillon, A. Nelson, A. Hammond, C. Sprecher, J. A. Gross, J. Johnston, K. Madden, W. Xu, J. West, S. Schrader, S. Burkhead, M. Heipel, C. Brandt, J. L. Kuijper, J. Kramer, D. Conklin, S. R. Presnell, J. Berry, F. Shiota, S. Bort, K. Hambly, S. Mudri, C. Clegg, M. Moore, F. J. Grant, C. Lofton-Day, T. Gilbert, F. Rayond, A. Ching, L. Yao, D. Smith, P. Webster, T. Whitmore, M. Maurer, K. Kaushansky, R. D. Holly and D. Foster (2000). "Interleukin 21 and its receptor are involved in NK cell expansion and regulation of lymphocyte function." Nature 408(6808): 57-63.

Patel, D. D. and V. K. Kuchroo (2015). "Th17 Cell Pathway in Human Immunity: Lessons from Genetics and Therapeutic Interventions." Immunity 43(6): 1040-1051.

Patil, V. S., A. Madrigal, B. J. Schmiedel, J. Clarke, P. O'Rourke, A. D. de Silva, E. Harris, B. Peters, G. Seumois, D. Weiskopf, A. Sette and P. Vijayanand (2018). "Precursors of human CD4(+) cytotoxic T lymphocytes identified by single-cell transcriptome analysis." Sci Immunol 3(19).

Picelli, S., O. R. Faridani, A. K. Bjorklund, G. Winberg, S. Sagasser and R. Sandberg (2014). "Full-length RNA-seq from single cells using Smart-seq2." Nat Protoc 9(1): 171-181.

Plank, M. W., G. E. Kaiko, S. Maltby, J. Weaver, H. L. Tay, W. Shen, M. S. Wilson, S. K. Durum and P. S. Foster (2017). "Th22 Cells Form a Distinct Th Lineage from Th17 Cells In Vitro with Unique Transcriptional Properties and Tbet-Dependent Th1 Plasticity." J Immunol 198(5): 2182-2190.

Platt, R. J., S. Chen, Y. Zhou, M. J. Yim, L. Swiech, H. R. Kempton, J. E. Dahlman, O. Parnas, T. M. Eisenhaure, M. Jovanovic, D. B. Graham, S. Jhunjhunwala, M. Heidenreich, R. J. Xavier, R. Langer, D. G. Anderson, N. Hacohen, A. Regev, G. Feng, P. A. Sharp and F. Zhang (2014). "CRISPR-Cas9 knockin mice for genome editing and cancer modeling." Cell 159(2): 440-455.

Power, C. A., G. Wei and P. A. Bretscher (1998). "Mycobacterial dose defines the Th1/Th2 nature of the immune response independently of whether immunization is administered by the intravenous, subcutaneous, or intradermal route." Infect Immun 66(12): 5743-5750.

Pratama, A., R. R. Ramiscal, D. G. Silva, S. K. Das, V. Athanasopoulos, J. Fitch, N. K. Botelho, P. P. Chang, X. Hu, J. J. Hogan, P. Mana, D. Bernal, H. Korner, D. Yu, C. C. Goodnow, M. C. Cook and C. G. Vinuesa (2013). "Roquin-2 shares functions with its paralog Roquin-1 in the repression of mRNAs controlling T follicular helper cells and systemic inflammation." Immunity 38(4): 669-680.

Proserpio, V., A. Piccolo, L. Haim-Vilmovsky, G. Kar, T. Lonnberg, V. Svensson, J. Pramanik, K. N. Natarajan, W. Zhai, X. Zhang, G. Donati, M. Kayikci, J. Kotar, A. N. McKenzie, R. Montandon, O. Billker, S. Woodhouse, P. Cicuta, M. Nicodemi and S. A. Teichmann (2016).

"Single-cell analysis of CD4+ T-cell differentiation reveals three major cell states and progressive acceleration of proliferation." Genome Biol 17: 103.

Raff, M. C. (1970). "Role of thymus-derived lymphocytes in the secondary humoral immune response in mice." Nature 226(5252): 1257-1258.

Rahemtulla, A., W. P. Fung-Leung, M. W. Schilham, T. M. Kundig, S. R. Sambhara, A. Narendran, A. Arabian, A. Wakeham, C. J. Paige, R. M. Zinkernagel and et al. (1991). "Normal development and function of CD8+ cells but markedly decreased helper cell activity in mice lacking CD4." Nature 353(6340): 180-184.

Ravindran, R. and S. J. McSorley (2005). "Tracking the dynamics of T-cell activation in response to Salmonella infection." Immunology 114(4): 450-458.

Reinhardt, R. L., H. E. Liang and R. M. Locksley (2009). "Cytokine-secreting follicular T cells shape the antibody repertoire." Nat Immunol 10(4): 385-393.

Rogers, P. R. and M. Croft (1999). "Peptide dose, affinity, and time of differentiation can contribute to the Th1/Th2 cytokine balance." J Immunol 163(3): 1205-1213.

Rogge, L., E. Bianchi, M. Biffi, E. Bono, S. Y. Chang, H. Alexander, C. Santini, G. Ferrari, L. Sinigaglia, M. Seiler, M. Neeb, J. Mous, F. Sinigaglia and U. Certa (2000). "Transcript imaging of the development of human T helper cells using oligonucleotide arrays." Nat Genet 25(1): 96-101.

Romagnani, S. (1996). "Th1 and Th2 in human diseases." Clin Immunol Immunopathol 80(3 Pt 1): 225-235.

Sallusto, F., D. Lenig, R. Forster, M. Lipp and A. Lanzavecchia (1999). "Two subsets of memory T lymphocytes with distinct homing potentials and effector functions." Nature 401(6754): 708-712.

Samstein, R. M., S. Z. Josefowicz, A. Arvey, P. M. Treuting and A. Y. Rudensky (2012). "Extrathymic generation of regulatory T cells in placental mammals mitigates maternal-fetal conflict." Cell 150(1): 29-38.

Sanders, M. E., M. W. Makgoba, S. O. Sharrow, D. Stephany, T. A. Springer, H. A. Young and S. Shaw (1988). "Human memory T lymphocytes express increased levels of three cell adhesion molecules (LFA-3, CD2, and LFA-1) and three other molecules (UCHL1, CDw29, and Pgp-1) and have enhanced IFN-gamma production." J Immunol 140(5): 1401-1407.

Sandler, N. G., M. M. Mentink-Kane, A. W. Cheever and T. A. Wynn (2003). "Global gene expression profiles during acute pathogen-induced pulmonary inflammation reveal divergent roles for Th1 and Th2 responses in tissue repair." J Immunol 171(7): 3655-3667.

Saoudi, A., S. Kassem, A. Dejean and G. Gaud (2014). "Rho-GTPases as key regulators of T lymphocyte biology." Small GTPases 5.

Schenkel, J. M. and D. Masopust (2014). "Tissue-resident memory T cells." Immunity 41(6): 886-897.

Schoenberger, S. P., R. E. Toes, E. I. van der Voort, R. Offringa and C. J. Melief (1998). "T-cell help for cytotoxic T lymphocytes is mediated by CD40-CD40L interactions." Nature 393(6684): 480-483.

Schwab, S. R. and J. G. Cyster (2007). "Finding a way out: lymphocyte egress from lymphoid organs." Nat Immunol 8(12): 1295-1301.

Seder, R. A., W. E. Paul, M. M. Davis and B. Fazekas de St Groth (1992). "The presence of interleukin 4 during in vitro priming determines the lymphokine-producing potential of CD4+ T cells from T cell receptor transgenic mice." J Exp Med 176(4): 1091-1098.

Sefik, E., N. Geva-Zatorsky, S. Oh, L. Konnikova, D. Zemmour, A. M. McGuire, D. Burzyn, A. Ortiz-Lopez, M. Lobera, J. Yang, S. Ghosh, A. Earl, S. B. Snapper, R. Jupp, D. Kasper, D. Mathis and C. Benoist (2015). "MUCOSAL IMMUNOLOGY. Individual intestinal symbionts induce a distinct population of RORgamma(+) regulatory T cells." Science 349(6251): 993-997.

Shale, M., C. Schiering and F. Powrie (2013). "CD4(+) T-cell subsets in intestinal inflammation." Immunol Rev 252(1): 164-182.

Shalem, O., N. E. Sanjana, E. Hartenian, X. Shi, D. A. Scott, T. Mikkelson, D. Heckl, B. L. Ebert, D. E. Root, J. G. Doench and F. Zhang (2014). "Genome-scale CRISPR-Cas9 knockout screening in human cells." Science 343(6166): 84-87.

Simmons, C. P., S. Clare, M. Ghaem-Maghami, T. K. Uren, J. Rankin, A. Huett, R. Goldin, D. J. Lewis, T. T. MacDonald, R. A. Strugnell, G. Frankel and G. Dougan (2003). "Central role for B lymphocytes and CD4+ T cells in immunity to infection by the attaching and effacing pathogen Citrobacter rodentium." Infect Immun 71(9): 5077-5086.

Simmons, C. P., N. S. Goncalves, M. Ghaem-Maghami, M. Bajaj-Elliott, S. Clare, B. Neves, G. Frankel, G. Dougan and T. T. MacDonald (2002). "Impaired resistance and enhanced pathology during infection with a noninvasive, attaching-effacing enteric bacterial pathogen, Citrobacter rodentium, in mice lacking IL-12 or IFN-gamma." J Immunol 168(4): 1804-1812.

Simpson, T. R., F. Li, W. Montalvo-Ortiz, M. A. Sepulveda, K. Bergerhoff, F. Arce, C. Roddie, J. Y. Henry, H. Yagita, J. D. Wolchok, K. S. Peggs, J. V. Ravetch, J. P. Allison and S. A. Quezada (2013). "Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma." J Exp Med 210(9): 1695-1710.

Skepner, J., M. Trocha, R. Ramesh, X. A. Qu, D. Schmidt, E. Baloglu, M. Lobera, S. Davis, M. A. Nolan, T. J. Carlson, J. Hill, S. Ghosh, M. S. Sundrud and J. Yang (2015). "In vivo regulation

of gene expression and T helper type 17 differentiation by RORgammat inverse agonists." Immunology 145(3): 347-356.

Smith, C. M., N. S. Wilson, J. Waithman, J. A. Villadangos, F. R. Carbone, W. R. Heath and G. T. Belz (2004). "Cognate CD4(+) T cell licensing of dendritic cells in CD8(+) T cell immunity." Nat Immunol 5(11): 1143-1148.

Stahl, E. A., S. Raychaudhuri, E. F. Remmers, G. Xie, S. Eyre, B. P. Thomson, Y. Li, F. A. Kurreeman, A. Zhernakova, A. Hinks, C. Guiducci, R. Chen, L. Alfredsson, C. I. Amos, K. G. Ardlie, B. Consortium, A. Barton, J. Bowes, E. Brouwer, N. P. Burtt, J. J. Catanese, J. Coblyn, M. J. Coenen, K. H. Costenbader, L. A. Criswell, J. B. Crusius, J. Cui, P. I. de Bakker, P. L. De Jager, B. Ding, P. Emery, E. Flynn, P. Harrison, L. J. Hocking, T. W. Huizinga, D. L. Kastner, X. Ke, A. T. Lee, X. Liu, P. Martin, A. W. Morgan, L. Padyukov, M. D. Posthumus, T. R. Radstake, D. M. Reid, M. Seielstad, M. F. Seldin, N. A. Shadick, S. Steer, P. P. Tak, W. Thomson, A. H. van der Helm-van Mil, I. E. van der Horst-Bruinsma, C. E. van der Schoot, P. L. van Riel, M. E. Weinblatt, A. G. Wilson, G. J. Wolbink, B. P. Wordsworth, Y. Consortium, C. Wijmenga, E. W. Karlson, R. E. Toes, N. de Vries, A. B. Begovich, J. Worthington, K. A. Siminovitch, P. K. Gregersen, L. Klareskog and R. M. Plenge (2010). "Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci." Nat Genet 42(6): 508-514.

Stevens, T. L., A. Bossie, V. M. Sanders, R. Fernandez-Botran, R. L. Coffman, T. R. Mosmann and E. S. Vitetta (1988). "Regulation of antibody isotype secretion by subsets of antigen-specific helper T cells." Nature 334(6179): 255-258.

Stoeckius, M., S. Zheng, B. Houck-Loomis, S. Hao, B. Z. Yeung, W. M. Mauck, 3rd, P. Smibert and R. Satija (2018). "Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics." Genome Biol 19(1): 224.

Stout, R. D. and K. Bottomly (1989). "Antigen-specific activation of effector macrophages by IFN-gamma producing (TH1) T cell clones. Failure of IL-4-producing (TH2) T cell clones to activate effector function in macrophages." J Immunol 142(3): 760-765.

Stubbington, M. J. T., O. Rozenblatt-Rosen, A. Regev and S. A. Teichmann (2017). "Single-cell transcriptomics to explore the immune system in health and disease." Science 358(6359): 58-63.

Suzuki, Y., M. A. Orellana, R. D. Schreiber and J. S. Remington (1988). "Interferon-gamma: the major mediator of resistance against Toxoplasma gondii." Science 240(4851): 516-518.

Szabo, S. J., S. T. Kim, G. L. Costa, X. Zhang, C. G. Fathman and L. H. Glimcher (2000). "A novel transcription factor, T-bet, directs Th1 lineage commitment." Cell 100(6): 655-669.

Tabula Muris, C., c. Overall, c. Logistical, c. Organ, processing, p. Library, sequencing, a. Computational data, a. Cell type, g. Writing, g. Supplemental text writing and i. Principal (2018).

"Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris." Nature 562(7727): 367-372.

Tada, T., T. Takemori, K. Okumura, M. Nonaka and T. Tokuhisa (1978). "Two distinct types of helper T cells involved in the secondary antibody response: independent and synergistic effects of Ia- and Ia+ helper T cells." J Exp Med 147(2): 446-458.

Tanay, A. and A. Regev (2017). "Scaling single-cell genomics from phenomenology to mechanism." Nature 541(7637): 331-338.

Teague, T. K., D. Hildeman, R. M. Kedl, T. Mitchell, W. Rees, B. C. Schaefer, J. Bender, J. Kappler and P. Marrack (1999). "Activation changes the spectrum but not the diversity of genes expressed by T cells." Proc Natl Acad Sci U S A 96(22): 12691-12696.

Tewari, K., Y. Nakayama and M. Suresh (2007). "Role of direct effects of IFN-gamma on T cells in the regulation of CD8 T cell homeostasis." J Immunol 179(4): 2115-2125.

Tubo, N. J., A. J. Pagan, J. J. Taylor, R. W. Nelson, J. L. Linehan, J. M. Ertelt, E. S. Huseby, S. S. Way and M. K. Jenkins (2013). "Single naive CD4+ T cells from a diverse repertoire produce different effector cell types during infection." Cell 153(4): 785-796.

Urban, J. F., Jr., I. M. Katona, W. E. Paul and F. D. Finkelman (1991). "Interleukin 4 is important in protective immunity to a gastrointestinal nematode infection in mice." Proc Natl Acad Sci U S A 88(13): 5513-5517.

Urban, J. F., Jr., N. Noben-Trauth, D. D. Donaldson, K. B. Madden, S. C. Morris, M. Collins and F. D. Finkelman (1998). "IL-13, IL-4Ralpha, and Stat6 are required for the expulsion of the gastrointestinal nematode parasite Nippostrongylus brasiliensis." Immunity 8(2): 255-264.

Veldhoen, M., K. Hirota, A. M. Westendorf, J. Buer, L. Dumoutier, J. C. Renauld and B. Stockinger (2008). "The aryl hydrocarbon receptor links TH17-cell-mediated autoimmunity to environmental toxins." Nature 453(7191): 106-109.

Veldhoen, M., C. Uyttenhove, J. van Snick, H. Helmby, A. Westendorf, J. Buer, B. Martin, C. Wilhelm and B. Stockinger (2008). "Transforming growth factor-beta 'reprograms' the differentiation of T helper 2 cells and promotes an interleukin 9-producing subset." Nat Immunol 9(12): 1341-1346.

Vignali, D. A., L. W. Collison and C. J. Workman (2008). "How regulatory T cells work." Nat Rev Immunol 8(7): 523-532.

Wang, M., V. Bronte, P. W. Chen, L. Gritz, D. Panicali, S. A. Rosenberg and N. P. Restifo (1995). "Active immunotherapy of cancer with a nonreplicating recombinant fowlpox virus encoding a model tumor-associated antigen." J Immunol 154(9): 4685-4692.

Wang, T., J. J. Wei, D. M. Sabatini and E. S. Lander (2014). "Genetic screens in human cells using the CRISPR-Cas9 system." Science 343(6166): 80-84.

Weinreb, C., S. Wolock and A. M. Klein (2018). "SPRING: a kinetic interface for visualizing high dimensional single-cell expression data." Bioinformatics 34(7): 1246-1248.

Wright, G. P., C. A. Notley, S. A. Xue, G. M. Bendle, A. Holler, T. N. Schumacher, M. R. Ehrenstein and H. J. Stauss (2009). "Adoptive therapy with redirected primary regulatory T cells results in antigen-specific suppression of arthritis." Proc Natl Acad Sci U S A 106(45): 19078-19083.

Xiao, S., H. Jin, T. Korn, S. M. Liu, M. Oukka, B. Lim and V. K. Kuchroo (2008). "Retinoic acid increases Foxp3+ regulatory T cells and inhibits development of Th17 cells by enhancing TGF-beta-driven Smad3 signaling and inhibiting IL-6 and IL-23 receptor expression." J Immunol 181(4): 2277-2284.

Yang, J., T. L. Murphy, W. Ouyang and K. M. Murphy (1999). "Induction of interferon-gamma production in Th1 CD4+ T cells: evidence for two distinct pathways for promoter activation." Eur J Immunol 29(2): 548-555.

Yang, X. O., R. Nurieva, G. J. Martinez, H. S. Kang, Y. Chung, B. P. Pappu, B. Shah, S. H. Chang, K. S. Schluns, S. S. Watowich, X. H. Feng, A. M. Jetten and C. Dong (2008). "Molecular antagonism and plasticity of regulatory and inflammatory T cell programs." Immunity 29(1): 44-56.

Yosef, N., A. K. Shalek, J. T. Gaublomme, H. Jin, Y. Lee, A. Awasthi, C. Wu, K. Karwacz, S. Xiao, M. Jorgolli, D. Gennert, R. Satija, A. Shakya, D. Y. Lu, J. J. Trombetta, M. R. Pillai, P. J. Ratcliffe, M. L. Coleman, M. Bix, D. Tantin, H. Park, V. K. Kuchroo and A. Regev (2013). "Dynamic regulatory network controlling TH17 cell differentiation." Nature 496(7446): 461-468.

Yu, S. L., W. P. Kuan, C. K. Wong, E. K. Li and L. S. Tam (2012). "Immunopathological roles of cytokines, chemokines, signaling molecules, and pattern-recognition receptors in systemic lupus erythematosus." Clin Dev Immunol 2012: 715190.

Yusuf, I., R. Kageyama, L. Monticelli, R. J. Johnston, D. Ditoro, K. Hansen, B. Barnett and S. Crotty (2010). "Germinal center T follicular helper cell IL-4 production is dependent on signaling lymphocytic activation molecule receptor (CD150)." J Immunol 185(1): 190-202.

Zeisel, A., A. B. Munoz-Manchado, S. Codeluppi, P. Lonnerberg, G. La Manno, A. Jureus, S. Marques, H. Munguba, L. He, C. Betsholtz, C. Rolny, G. Castelo-Branco, J. Hjerling-Leffler and S. Linnarsson (2015). "Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq." Science 347(6226): 1138-1142.

Zemmour, D., R. Zilionis, E. Kiner, A. M. Klein, D. Mathis and C. Benoist (2018). "Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR." Nat Immunol 19(3): 291-301.

Zheng, W. and R. A. Flavell (1997). "The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells." Cell 89(4): 587-596.

Zheng, Y., P. A. Valdez, D. M. Danilenko, Y. Hu, S. M. Sa, Q. Gong, A. R. Abbas, Z. Modrusan, N. Ghilardi, F. J. de Sauvage and W. Ouyang (2008). "Interleukin-22 mediates early host defense against attaching and effacing bacterial pathogens." Nat Med 14(3): 282-289.

Zheng, Y., Y. Zha, R. M. Spaapen, R. Mathew, K. Barr, A. Bendelac and T. F. Gajewski (2013). "Egr2-dependent gene expression profiling and ChIP-Seq reveal novel biologic targets in T cell anergy." Mol Immunol 55(3-4): 283-291.

Zhou, L., M. M. Chong and D. R. Littman (2009). "Plasticity of CD4+ T cell lineage differentiation." Immunity 30(5): 646-655.

Zhu, J. and W. E. Paul (2008). "CD4 T cells: fates, functions, and faults." Blood 112(5): 1557-1569.

Zielinski, C. E., F. Mele, D. Aschenbrenner, D. Jarrossay, F. Ronchi, M. Gattorno, S. Monticelli, A. Lanzavecchia and F. Sallusto (2012). "Pathogen-induced human TH17 cells produce IFN-gamma or IL-10 and are regulated by IL-1beta." Nature 484(7395): 514-518.

Appendix A

Publication

"***Identification and validation of a tumor-infiltrating Treg transcriptional signature conserved across species and tumor types.***"

Angela M Magnuson, Evgeny Kiner, [Ψ]Ayla Ergun, Jun Seok Park, Natasha Asinovsk[1], Adriana Ortiz-Lopez, Aoife Kilcoyne, Elisa Paoluzzi-Tomada, Ralph Weissleder, Diane Mathis and Christophe Benoist

Abstract

FoxP3^{+} T regulatory (Treg) cells are a central element of immunological tolerance. They are abundant in many tumors, where they restrict potentially favorable anti-tumor responses. To identify genes related to Treg presence and function in the tumor micro-environment, we used a three-pronged strategy. Gene-expression profiles were generated from tumor-infiltrating Tregs (TITR) of both human and mouse tumors, and were compared with those from Tregs of lymphoid organs or normal tissue from the same individuals. A computational deconvolution of whole-tumor datasets from the Cancer Genome Atlas (TCGA) was performed to identify transcripts specifically associated with Tregs across thousands of tumors from different stages and locations. We identified a set of TITR-differential transcripts, with striking reproducibility between tumor types in mice, between mice and humans, and between different human patients spanning different tumor stages. Many of the TITR-preferential transcripts were shared with "tissue-Tregs" that reside in non-lymphoid tissues, but a tumor-preferential segment could be identified. Many of these TITR signature transcripts were confirmed by mining of TCGA datasets, which also brought forth transcript modules likely representing parenchymal attraction of, or response to, tumor-Tregs. Importantly, the TITR signature included several genes encoding effective targets of tumor immunotherapy. A number of other targets were validated by CRISPR-based gene inactivation in mouse Tregs. These results confirmed the validity of the signature, generating a wealth of leads for understanding the role of Tregs in tumor progression, and implicating potential targets for cancer immunotherapy.

Significance Statement

FOXP3+ Treg cells dampen immune responses in many environments, and in particular tumors, where they contribute to cancer's resistance to immunological defenses. This very broad analysis of tumor-infiltrating Tregs has identified a set of genes that are preferentially expressed by these Tregs in different species, tumor models or cohorts, and types or stages of tumors. This striking commonality suggests that there are core mechanisms that tumors use to attract and mold Tregs, whose perturbation should unleash anti-tumor immunity. Experimental validation by genome editing provides a proof-of-concept for the relevance of these genes in TITRs.

Introduction

Regulatory T cells (Tregs) characterized by the transcription factor FoxP3 are critical for maintaining immunologic homeostasis, enforcing tolerance to self, and preventing runaway immune responses (1, 2), in both mice (3, 4) and humans (5). Tregs regulate the activation and differentiation of conventional $CD4^+$ T cells (Tconv) but also of many other cell lineages within the innate and adaptive immune systems, through a variety of effector mechanisms (reviewed in (6)). There is also an increasing realization that Tregs have extra-immunologic roles in the homeostasis of several tissues, controlling the noxious side-effects of inflammation ensuring effective tissue repair, and otherwise promoting homeostasis (7, 8).

Tregs are often found at elevated frequency in tumors relative to blood or lymphoid organs in human cancer patients and mouse models (9, 10). For a number of cancers, but not all, a high density of Tregs correlates with poor prognosis (reviewed in (9, 11, 12)). However, correlative analyses of this nature can be misleading because the abundance of Tregs in a locale tends to track with the extent of overall immunocyte infiltration. Causal involvement of Tregs in tumor progression was first demonstrated in mice, where their depletion via administration of anti-CD25 antibody inhibited or reversed several tumors (13). These and other studies (14-19) showed that Treg depletion increased the number of $CD4^+$ and/or $CD8^+$ effector T cells (Teffs) in the tumor, associated with robust tumor-specific killing activity. Thus, tumor eradication in these settings was due, at least in part, to removal of Treg-mediated suppression of the anti-tumor immune response. Similarly, the success of several FDA-approved immunotherapies for cancer (e.g. anti-CTLA-4 and anti-PD-1) may be attributable to their effects on Tregs in addition to promoting Teff killing (14, 20-22).

There are several indications that the phenotypes of TITRs are distinct from their more generic Treg counterparts found in lymphoid organs. For instance, human TITRs express the cell surface receptor NRP1, which is absent from Tregs in blood or lymphoid organs (23). Analyses of tumor-infiltrating Tregs in colorectal and breast tumors showed that they are highly suppressive, expressing a number of markers associated with "activated Tregs" (aTregs) (19, 24-28). These aTreg characteristics are also over-represented in tissue-Tregs (8).

At present, immunotherapies approved for use in human cancers often have significant side-effects. Some are of short duration but can be life-threatening, and require intensive-care management; others lead to long-lasting autoimmune and auto-inflammatory pathologies, of the types that might be predicted to result from Treg depletion or incapacitation (reviewed in (29, 30)). The ideal therapy to release the inhibition imparted by Treg cells on anti-tumor immune responses would be tumor-selective in its effect, target tumor-infiltrating Tregs (TITR) while not affecting Tregs in general to avoid autoimmune consequences, or tissue Tregs to avoid homeostatic perturbations in other organ systems.

Our goal was to identify genes differentially expressed in TITRs relative to Tregs found in secondary lymphoid organs and normal non-lymphoid tissues, in both human patients and mouse models. We followed a three-pronged strategy: gene-expression profiling was performed on fresh Tregs from human or mouse tumors, and profiles were compared with those of Tregs from lymphoid organs or normal tissue. In parallel, we conducted a bioinformatic analysis of large datasets from whole tumors in TCGA, to ferret out transcripts specifically associated with Treg cells across large numbers of tumors. The combined analysis identified a number of candidate genes encoding plausible targets of anti-tumor immunotherapy. Direct

perturbation of some of these proteins by CRISPR-based genetic ablation validated our approach, and provided exciting leads for cancer immunotherapy.

Results

We followed the multi-pronged approach described in Fig. 1, which provided three orthogonal and cross-validating datasets. First, we purified and profiled TITRs from three transplantable tumors in mice, comparing these profiles with those of Tregs and other lymphocytes from various lymphoid and non-lymphoid organs. Second, we purified and profiled TITRs from patients with colorectal tumors, in comparison with Treg cells from normal colon tissue of the same donors. Both of these approaches should identify transcripts and pathways that distinguish TITRs from both standard, lymphoid-organ Tregs and non-lymphoid-organ Tregs. Third, we broadened the scope of the data by mining whole-tumor datasets from several different tumor types, generated by TCGA ((31); TCGA Research Network: http://cancergenome.nih.gov/.), to identify genes whose expression correlated specifically with that of the gene encoding the Treg-defining factor FoxP3. Each of these datasets was analyzed alone, and so their intersection defined high-confidence predictions, conserved across species and tumor-types, of gene signatures specifically active in tumor-infiltrating Tregs.

Transcripts specific to mouse tumor-Tregs.

We primarily used three mouse models of cancer, namely the transplantable MC38 and CT26 colon carcinomas and B16 melanoma. In each case, tumors were inoculated subcutaneously into immunocompetent FoxP3$^{IRES\text{-}GFP}$ reporter mice (32), such that Treg cells could be sharply and uniquely identified by the fluorescent GFP reporter (Fig. 2A). After establishment and growth of these tumors (21 days), immunocytes were isolated from tumors

and spleens, and Tregs were highly purified (cytometric double-sorting; Fig. 2A) for gene expression profiling (in biological triplicates).

As illustrated in Fig. 2B, several hundred transcripts were significantly altered in Treg cells of each tumor compared with corresponding splenic Tregs. Some differences were very pronounced, greater than 20-fold. The majority were over-expressed in TITRs relative to spleen Tregs, suggesting inductive events operating in the tumors. Moreover, the vast majority of genes differentially expressed in TITRs from one tumor model were similarly affected in others (Fig. 2C, Dataset S1). That the tumor-Treg-specific signature is quasi-identical in tumors as different as melanoma and colon carcinoma suggests that these inductive events represented a recurrent response of Tregs to tumor micro-environments.

We then asked which of these differential transcripts were specific to Treg cells, or might be activated in all lymphocytes found in the tumor micro-environment. Comparing the tumor/spleen ratio in CD8+ T or Tconv cells to that seen for Tregs (Fig. 2D) indicated that many of the transcripts altered in TITRs were also changed in these other intra-tumoral lymphocytes. This trend was not absolute, however, 10 to 15% of the transcripts being uniquely induced in TITRs, but not in the other T cells tested. To better delineate TITR-specific transcripts, as opposed to those equally expressed in other tumor-infiltrating lymphocytes, we performed a direct comparison of expression levels between tumor-infiltrating lymphocytes. A large fraction of the transcripts that distinguished TITRs from spleen Tregs (red highlights in Fig. S1A) were indeed over-expressed in TITRs relative to CD8+ or Tconv cells.

As introduced above, there is growing awareness that Tregs have an extra-immunologic role in several tissues, regulating the harmful side-effects of inflammation and promoting

effective tissue repair and otherwise ensuring tissue homeostasis. Given our goal to identify molecules specific for TITRs, as opposed to Tregs in other locations, and because some of the transcripts overexpressed in TITRs were reminiscent of non-lymphoid-tissue Tregs, we compared the transcriptomes of these two Treg classes. For Fig. 2E, we calculated the mean of fold-changes in several tissue Tregs relative to spleen (data from visceral adipose tissue, injured muscle and colonic lamina propria; (33)) and plotted these vs the mean of fold-changes between TITRs and corresponding splenic Tregs (averaged from the tumor models described above). Several points can be made. First, and as previously noted (27), there was a strong relationship between TITRs and non-lymphoid-tissue Tregs, with a strong correlation overall, and many genes that showed very comparable induction or repression in the two cases (green highlights in Fig. 2E). This correspondence is not surprising, as one might expect that some of the cues that drive Tregs to normal tissues might also apply to tumors. Second, there were, however, transcripts with a stronger bias in Tregs within normal tissues (blue in Fig. 2E, e.g. *Il10*, *Fos*, or *Areg*) or tumors (red). The latter included *Il12rb2, Tnfrsf9* and *Cd274.* To better separate tissue-Treg- and TITR-specific transcripts, we used a Principal Components Analysis (PCA) approach. As illustrated in Fig. S1B, the first PC was strongly related to the TITR/spleen differential, but also brought out clearly the transcripts that fell off the main line because of over-expression in tissue-Tregs (e.g. *Il10*, *Areg*, *Ctla2a* and *Ccr5*). Further, when we generated new data to compare the transcriptomes of TITR and colon-resident Treg cells, versus splenic Tregs, from MC38 tumor-bearing mice (Fig. S2), we observed a similar distribution of transcripts as that shown in Fig. 2E. Moreover, when the gene-set from Dataset S1 (described below) was

highlighted on these data, it was apparent that while certain transcripts were shared between TITR and tissue Tregs, others were indeed preferentially expressed in TITR.

These comparisons highlighted a large number of transcripts differentially expressed in TITRs from several tumor types relative to generic Treg cells from lymphoid organs, but with significant overlap with other cells (tissue-Tregs, other tumor-infiltrating T cells). Since our intent was to identify a gene-set that most specifically characterized TITRs, we selected 139 transcripts based on a combination of criteria: >3-fold over-expression in TITRs relative to spleen Tregs for MC38, B16 or CT26 (with $p < 10^{-2}$), over-expression in TITR relative to tumor CD4+ or CD8+ T cells (>2), high value in the PC that distinguished tumor- from tissue-Tregs (Dataset S1). We also discarded a group of differential transcripts typical of tumor-infiltrating myeloid cells (identified using myeloid cell signatures from the ImmGen database; this contamination was difficult to completely eliminate from the sorts in spite of stringent negative controls and flow sorting). Interestingly, our gene-selection pathway led us to transcripts encoding proteins already recognized for their role in Treg function and/or costimulation (*Ctla4, Tigit, Tnfrsf9* (encoding 4-1BB). We also noted that genes like *Cd274*, *Icos* and *Lag3* narrowly missed the selection criteria. The presence of such genes in our TITR gene-set lends credence to their lesser known counterparts in the same gene-set, and argues that they might constitute effective targets for Treg modulation.

Transcripts specific for human colorectal tumor Tregs.

In the second arm of our study, we purified Tregs from twelve surgically resected and cryopreserved human colorectal carcinomas (CRC) or normal colon tissue, often from the same

donors. These tumors spanned a range of stages from early (stage I, IIA) to late (stage IIIB) CRCs. The usual criteria ($CD25^{hi}CD127^{lo}$) were used to identify and double-sort Tregs to high purity. As in the mouse models, and as previously reported in human tumors (27, 34-37), Tregs represented a higher fraction of tumor CD4+ T cells (20.0 ± 4.8) than in normal colon (7.7 ± 4.9**;** Fig. 3A). RNAseq was performed on these purified Treg samples. Differential gene-expression analysis, comparing all tumor-Tregs with all colon Tregs as a group, revealed a significant bias in the transcriptome of human TITRs (Fig. 3B). As in the mouse models, these differential transcripts were mostly over-expressed in TITRs relative to Tregs from normal tissues, and included several encoding proteins previously recognized to have costimulatory function (*TNFRSF4, TNFRSF9, TNFRSF18* (aka OX40, 4-1BB, GITR, respectively). Interestingly, *FOXP3*, itself, was over-expressed relative to colonic Tregs, suggesting that the TITR environment specifically activated the *FOXP3* locus. Some transcripts were down-regulated in TITRs, which surprisingly included genes encoding proteins known to be involved in Treg function, such as *CXCR5* and *IL10* (38-40). CCR7 is typical of resting "rTregs" (41), and its reduction points to an increase in activated "aTreg" phenotypes among TITRs. The drop in *IL10* suggests that different suppressor mechanisms may come into play in Treg cells from tumors relative to normal colon.

From these data, we selected a set of 408 genes (355 UP, 73 DN; Dataset S2) that distinguished TITR from normal-tissue Tregs based on a combination of criteria (mean overall fold-change, or high fold-change in at least two patients; see Methods). Importantly, these differences in gene expression were very reproducible between individuals (Fig. 3C), irrespective of the stage of CRC examined (namely stages I, IIA, IIIB), a pattern that was not necessarily expected given the heterogeneity of human genetics and the various tumor stages

represented in our samples. To further test the generality of our TITR signature, we compared the overlap with recently published Treg datasets from non-small-cell lung cancer (NSCLC) (28), breast cancer (27), and hepatocellular carcinoma (37). Indeed, there was marked overlap, as illustrated for breast cancer TITRs in Fig. 3D. Thus, the TITR signature is indeed a general one shared among multiple cancer types.

We then investigated the composition of these UP and DN signatures. GeneOntology analysis was not informative, yielding mainly generic categories. As for the corresponding selection for mouse TITRs, a number of transcripts encoding proteins implicated in aTreg differentiation (e.g. several members of the TNFR superfamily (GITR, 4.1BB, OX40 and CD30), or chemokine receptors (CCR1, CCR5, CCR8)) were present. Accordingly, a number of these transcripts were associated with the cell cycle (Dataset S2), and Enrichr motif analysis brought out FOXM1 as a transcription factor likely to be controlling a portion of these TITR-specific transcripts, particularly those associated with the cell cycle (Fig. 3E). The set also included a cluster typical of Th17 cells (IL17A, IL17F, IL22), in keeping with previous results showing RORγt-dependent expression of these cytokines in human colon tumors (42).

Lastly, we asked whether the genes over-expressed in TITRs within mouse tumors were also differentially represented in human TITRs. Indeed, there was significant overlap of the tumor-Treg signature between the mouse and human data-sets (Fig. 3F): 77 of 90 transcripts up-regulated >4-fold in TITR vs splenic Tregs in the mouse were also over-represented in human TITRs relative to normal colonic Tregs (chisq $p < 10^{-10}$). Conversely, highlighting the human TITR-upregulated signature on the mouse tumor/tissue comparison (per Fig. 2E) showed a strong bias towards over-expression in tumors, mainly tumor-preferential (Fig. 3G). These results

clearly demonstrate conservation of the tumor Treg signature across species, and provide validation for the relevance of the mouse results.

Genes correlated with Tregs in TCGA tumor datasets.

In the third arm of this work, the analytical logic was different. FoxP3 is the key transcription factor that defines Tregs and determines much of their transcriptional identity (1). Thus, genes that are specifically expressed in TITR, relative to other tumor-infiltrating immunocytes, should exhibit a tight correlation with *FOXP3* transcripts across panels of gene-expression profiles generated from whole tumors. To this end, we took advantage of TCGA, a publically available database of gene expression in 33 types of cancer from more than 11,000 patients. We selected the RNA Sequencing data for 4 types of cancer, representing different frequencies of mutational load, and hence likelihood of anti-tumor immune response (43) (colon, breast, pancreas and lung, with 285, 1093, 178 and 501 tumor samples, respectively). First attempts at correlation with *FOXP3* transcripts in these datasets brought out a large number of genes, many of which were typical of non-Tregs (e.g. immunoglobulin transcripts from B cells). This observation indicated that the proportion of Tregs in the tumor samples was parallel to the overall degree of infiltration by immune cells in general, which is known to vary quite widely between individual tumors (9).

To account for, and computationally remove, this confounder, we mathematically regressed out the general degree of infiltration, in an approach algorithmically similar to those used in estimating the abundance of tumor-infiltrating cells (44, 45). First, we curated prototypical signature gene-sets corresponding to the major types of immunocytes that can be

found in a tumor (B cells, dendritic cells, eosinophils, mast cells, macrophages, neutrophils, NK cells and T cells; Dataset S3). For each sample, we calculated an average infiltrate index for each cell type *k* using these marker genes, and then used the cell-type indices as covariates in a linear model for each gene *i* ($Y_i=\beta_0+\Sigma\beta_k$ *$Cell\_Index_k+\varepsilon$). We used the residual of this fit as a measure of gene-expression independent of the degree of infiltration into each tumor, and correlated each gene's residual expression with *FOXP3* expression levels (using bootstrap permutation of samples or of signature gene-sets to demonstrate the robustness of the procedure, and compute 95% confidence intervals; Fig. S3, Dataset S4). Fig. 4A displays genome-wide correlations with *FOXP3* expression before and after removal of the infiltrate component (averaged across the 4 tumor types; individual tumor plots in Fig. S4A). In keeping with our hypothesis, the correlation with *FOXP3* expression essentially disappeared for most of the transcripts, although for a minority of genes the correlation with *FOXP3* remained high (validity supported by disappearance in randomly permuted datasets, Fig. S4B). These transcripts included some encoding well-known Treg-specific proteins such as IL2RA or those encoding the costimulatory molecules CTLA4, and ICOS. These transcripts had high initial correlation, which largely persisted in the regressed data (area 1 in Fig. 4A). Another group of transcripts showed low initial correlation with *FOXP3* expression, which increased after the correction (area 2). Importantly, these correlation patterns were similar in different types of tumors (Fig. 4B, S4), again indicating that the TITR-specific transcriptome was largely shared between human tumor types. In addition, these *FOXP3*-correlated transcripts were not simply a rediscovery of the classic Treg signature (46): only a limited fraction of the Treg-up signature

was positively correlated with *FOXP3* expression, as illustrated for the breast and lung tumor datasets in Fig. 4C.

To assess the uniformity of expression of *FOXP3*-correlated transcripts across individual tumors, we selected 219 with the highest post-regression correlation, and investigated their distribution in the four tumor data-sets (Fig. 4D, Dataset S5). Their levels were clearly not uniform across tumors. Some gene modules varied in lockstep between individual tumors, in patterns that carried across tumor types, underscoring the robustness of these relationships and hinting at a common and fundamental mechanism of interaction amongst tumors and the immune system.

In principle, the correlation between these 219 transcripts and *FOXP3* transcripts could have three roots: (i) expression in TITRs, themselves; the classic Treg-associated transcripts of area 1 clearly belong to this class. (ii) expression connected to mechanisms that draw or retain Tregs into tumors and control their abundance; or (iii) expression that results from Treg presence and activity. To best resolve the co-regulated modules observed in Fig. 4D, we computed a gene-gene correlation matrix from the expression of these 219 genes in each tumor (averaged heatmap in Fig. 4E, individual tumors in Fig. S5). Several strong modules emerged, which were almost identical in all four tumor types, as predictable from Fig. 4D. These modules were of very different composition, as evidenced by their expression levels in the tissue data-sets from the GTEx database (47) (Fig. 4F). Correlated Set (CS)1 was composed of transcripts primarily encoding collagen or matrix metalloprotease of the ADAM family, which are not expressed in Treg cells or lymphoid cells, but in fibroblasts or connective tissues. CS2 corresponds to a subset of cell-cycle-related transcripts (high in leukemic cells and testis). The

previously observed TITR transcripts (*IL2RA, CCR8, TIGIT, CD80, ICOS*, etc) grouped in a more loosely correlated region (CS3). Transcripts of CS4 were less distinctive, but included number of histone/protein modifiers (*HCFC1*, *BRD4*, *SETD1A/B*). Thus, correlation with *FOXP3* expression brought forth, in addition to TITR-specific transcripts, distinct correlated gene modules whose expression segregates in many tumors, and which may be causally related to Treg presence or activity.

Data integration

Next, we integrated these three inputs in order to compare their outcomes and to select the set of transcripts most specific for TITRs across different tumors. Since the actual data and their distribution were different, we opted to combine the rankings of orthologous genes in the two species, rather than the expression metrics. For the human TITR data-sets, each gene was ranked according to its expression in TITRs of each tumor sample relative to batch-matched Tregs from normal colon tissue, and then these ranks were summed. For the mouse TITR data-sets, genes were similarly ranked by their over-expression in TITRs within the three tumor-types relative to matched splenic Tregs, and these ranks were also summed. Fig. 5A displays the comparison of these overall scores for human and mouse TITRs. Consistent with the trans-species conservation of TITR-specific gene expression described above, there was a strong correlation between these rankings ($r=0.34$, $p<10^{-16}$). A group of transcripts was on top of the ranking for tumor-specificity in both species, and the gene list selected on the basis of mouse tumor data (Dataset S1) included many transcripts with high rank in the human data-sets (Fig. S6A). Some genes led the ranking for over-expression in human TITRs but not in mice (and vice-

versa), suggesting that at least some were species-specific (or might be tied to "natural" vs transplanted tumors). Genes that correlated with *FOXP3* in the TCGA data were also high in the rankings for over-expression in human and mouse TITRs (Fig. S6B), as more directly shown in Fig. 5B. Many of the highest ranked Treg transcripts also showed the highest association with *FOXP3* transcripts in the TCGA datasets (e.g. *CCR8*, *TNFRSF9*, *IL21R*), although this was not always the case (e.g. *DUSP4*). From these three orthogonal inputs, also integrating TITR:tissue Treg and TITR:CD8 differential data, we assembled a list of 108 genes most differentially expressed and best correlated with tumor Tregs (Fig. S6C, Dataset S6).

Experimental validation of TITR-specific targets

It was of importance to validate the relevance of these gene sets to TITR physiology, both to understand how they might mechanistically modulate TITR activity and to serve as potential therapeutic targets. We examined the importance of *C3AR1, IL12RB2 and IL1RL1* (ST2) using available strains of conventional knockout (KO) mice. MC38 tumors were induced in KO mice and their wild type (WT) littermates, and tumor growth was measured over time (Fig. S7). There was no significant difference in tumor growth between the KO and wild-type (WT) littermates for the genes examined, except for the Treg-specific *IL1RL1* (ST2) KO group.

Simultaneously, we performed a systematic *in vivo* screen of the top-ranked TITR-specific targets by creating loss of function (LOF) mutations in the protein-coding regions by adapting the CRISPR/Cas9 system (48) to specifically edit Treg cells, assessing differential representation in the tumor relative to lymphoid organs (Fig. 6A, Dataset S7). Tregs were purified from transgenic mice expressing both the Cas9 protein (48) and the FoxP3-Thy1.1

reporter (49)**,** activated *in vitro*, and were transduced with retroviral vectors encoding the targeting single-guide RNAs (sgRNAs) of interest. Pairs of retroviral vectors expressing two different fluorescent proteins (GFP and RFP) were used, which allowed us to label and compare Tregs carrying targeting or control sgRNAs in the same mouse (and thereby increase the informativity of each mouse). Transduced Tregs were then transferred together with "filler" splenocytes (to avoid homeostatically driven expansion) into alymphoid RAG-deficient mice, which concomitantly received MC38 tumor cells. After 18 days to allow for tumor growth and Treg homing, we analyzed the distributions of Treg cells carrying each of the targeting or control sgRNAs in the tumor relative to lymphoid organs, calculated as a "tumor depletion index" (Figs. 6B and C). By DNA amplification and sequencing we also ensured that the target genes were efficiently edited (estimated as > 50% in all cases). The compiled results showed that several of the LOF mutations had no impact on TITR proportions, but three of them did: *Tnfrsf8* (encodes CD30), *Cxcr3* (a chemokine receptor) and *Samsn1* (an intracellular signaling adapter). This first *Samsn1* LOF result was validated with four independent sgRNAs targeting different regions of *Samsn1*. Each sgRNA decreased Treg accumulation in the tumors (Fig. 6D), confirming the effect.

Overall, these proof-of-concept results indicated that some of the TITR transcripts highlighted in Dataset S6 encoded functionally relevant proteins and might, indeed, serve as valuable targets for cancer immunotherapy.

Discussion

Our goal was to identify genes differentially expressed in TITRs from various origins, relative to generic Tregs from secondary lymphoid organs or Tregs from non-lymphoid tissues, as a means to better understand the unique biology of Treg cells in tumors and to identify potential targets for immunotherapy. A similar motivation has driven previous studies, but we adopted a broader, multi-pronged strategy that spanned species, harnessed the large TCGA datasets, and performed proof-of-concept experimental validation. Our complementary approaches also rested on different principles, namely differential expression between Tregs in tumors and non-lymphoid tissues and transcript correlations with *FOXP3* expression. This combination strategy allowed us to identify a TITR-specific signature of unexpected constancy**.** There was a very strong overlap between the transcriptomes of Treg cells infiltrating tumors and non-lymphoid tissues, which is perhaps not surprising as some of the signals that permit Tregs to accumulate in tissues might also allow their residence in tumors. These commonalities did present a challenge for defining TITR-specific transcripts, but we were ultimately able to distil a subset of transcripts with marked tumor preference. The power of our strategy was validated by the "re-discovery" of several established targets of tumor immunotherapy (*CTLA4*, *TIGIT*; and *PD-L1* and *LAG3* barely missing our thresholds). Such validation provides credibility for the lesser-known genes in our TITR signature being potential targets for preferentially modulating TITRs.

A striking convergence of TITR signatures was apparent between tumor types in mouse models, between mouse and human TITRs, and between different patients with the same tumor. It also extended to the large sets of tumors in TCGA data-sets, for which careful

computational parsing uncovered the same core signature of *FOXP3*-correlated genes was observed in every tumor-type tested. In addition, there was a strong overlap with differentially expressed transcripts recently reported in more focused studies of Tregs from human lung, colorectal (28), breast (27) and liver (37) tumors. This convergence implies that a fundamental mechanism must be at play to allow such diverse tumors to elicit the same response. From a practical standpoint, it means that the same pathways might be effectively targeted by immunotherapy to relieve the dampening of anti-tumor responses by TITRs in many tumor types.

While the bulk of the TITR signature was indeed shared by different cancers, we were able to identify a handful of TITR transcripts specific to CRC when we compared our data with previously reported profiles from breast cancer Tregs (Fig. 3D). These CRC TITR-specific genes included loci encoding cytokines typical of Th17 cells (*IL17A*, *IL17F*, *IL22*) and *RORC,* the major driver of Th17 cell differentiation. This observation is consistent with a previous report of RORγt-dependent expression of Th17 cytokines in Tregs from human colon tumors (42), and with the microbiome-dependent RORγt+ Tregs that are found electively in the colon (33, 50). Thus, beyond a core TITR signature shared in all tumors, individual locations can imprint an additional local component on TITRs. Similarly, some transcripts were noted as differential only in one species, such as *Il21R* which had a strong differential score in human but not mouse TITRs.

Finally this multipronged exploration, and particularly the TCGA arm, revealed sets of transcript modules that correlated with TITR levels, but were not TITR transcripts themselves. It

will be interesting to explore the leads opened here, for instance by assessing how ADAM and collagen gene expression might help Tregs accumulate in the tumors.

Beyond providing a detailed landscape of the strikingly constant phenotype adopted by TITRs in the tumor environment, these explorations yielded sets of transcripts encoding proteins that might be effective targets for immunotherapy through relief of the brakes that Treg cells impose on anti-tumor responses. There could be several roots for the over-expression of a particular molecule in TITRs, which condition the actual therapeutic strategy. Some might be over-expressed because they are required by TITRs to survive specifically in this environment: chemokine receptors, for instance, are needed for TITRs to accumulate in tumors. Blocking them or reducing their expression could exclude TITRs from tumors. Other molecules, in contrast, might instead denote negative feedback loops that exist in all biological systems, and instead act to limit TITR over-expansion; their elimination might actually benefit TITRs. Still others might be upregulated in response to stimuli in the tumor microenvironment but play no role in Treg survival or function therein.

Given this diversity, we used parallel genetic approaches to test the significance of the genes revealed by these parallel genomic comparisons. A few could be tested in knockout mice, but we leveraged editing by CRISPR/Cas9 to assess systematically how their inactivation affected Treg accumulation in the tumor. This approach was particularly valuable for intracellular targets such as *Samsn1*, which cannot be easily probed by mAb infusion. In this way, 3 of the 14 targets tested showed a significant effect of editing. One limitation of the genetic approach, with regards to potential therapeutic application, is that it is blind to effects of blocking the encoded protein, or of killing TITRs by complement or ADCC mechanisms, which

could be revealed by mAb treatment and therapeutically valuable. For example, editing *CCR8* in Tregs did not lower their accumulation in tumors, but our preliminary explorations show that mAb engagement of CCR8 caused a tumor mass reduction in treated mice. Thus, CCR8 may not be uniquely indispensable for TITR recruitment or survival in tumors (redundancy is not uncommon in chemokine networks), but it might serve as a worthwhile mAb target. Moreover, the other targets for which LOF mutations had no impact on TITR proportions should not necessarily be excluded from exploration as therapeutic targets by other means of perturbation.

In conclusion, this multi-pronged study has opened many perspectives and leads on genomic aspects of Tregs in tumors, pointing to specific transcriptional programs that could be harnessed to lift the immune inhibition by Tregs in multiple cancer contexts.

Acknowledgements

We thank Drs. A. Sharpe, M. Pittet, V. Kuchroo for insightful discussion or materials, K. Hattori, C. Araneo, A. Rhoads, A. Cook and L. Yang for help with mice, cell sorting, and profiling, and the Broad BTL for library construction and sequencing. This work was funded by the Immune Tolerance Fund at HMS, the Blavatnik Biomedical Accelerator at Harvard University; AM was supported by T32CA079443, EK by a PhD Fellowship from Boehringer Ingelheim Fonds.

References

1. Josefowicz SZ, Lu,LF, Rudensky,AY (2012) Regulatory T cells: mechanisms of differentiation and function. *Annu. Rev. Immunol.* 30 531-564.

2. Benoist, C. & Mathis, D. (2013) in *Immune Tolerance*, eds. Mathis, D. & Rudensky, A. (Cold Spring Harbor Press, Cold Spring Harbor, NY), pp. 31-44.

3. Brunkow ME, et al. (2001) Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. *Nat. Genet.* 27(1): 68-73.

4. Fontenot JD, Gavin,MA, Rudensky,AY (2003) Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nat. Immunol* 4(4): 330-336.

5. Bennett CL, et al. (2001) The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nat. Genet.* 27 20-21.

6. Vignali DA, Collison,LW, Workman,CJ (2008) How regulatory T cells work. *Nat. Rev. Immunol.* 8(7): 523-532.

7. Feuerer M, et al. (2009) Lean, but not obese, fat is enriched for a unique population of regulatory T cells that affect metabolic parameters. *Nat Med* 15(8): 930-939.

8. Panduro M, Benoist,C, Mathis,D (2016) Tissue Tregs. *Annu. Rev Immunol* 34 609-633.

9. Fridman WH, Pages,F, Sautes-Fridman,C, Galon,J (2012) The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer* 12(4): 298-306.

10. Roychoudhuri R, Eil,RL, Restifo,NP (2015) The interplay of effector and regulatory T cells in cancer. *Curr. Opin. Immunol* 33 101-111.

11. deLeeuw RJ, Kost,SE, Kakal,JA, Nelson,BH (2012) The prognostic value of FoxP3+ tumor-infiltrating lymphocytes in cancer: a critical review of the literature. *Clin Cancer Res.* 18(11): 3022-3029.

12. Shang B, Liu,Y, Jiang,SJ, Liu,Y (2015) Prognostic value of tumor-infiltrating FoxP3+ regulatory T cells in cancers: a systematic review and meta-analysis. *Sci Rep* 5 15179.

13. Shimizu J, Yamazaki S, Sakaguchi S (1999) Induction of tumor immunity by removing CD25+CD4+ T cells: a common basis between tumor immunity and autoimmunity. *J Immunol* 163(10): 5211-5218.

14. Quezada SA, Peggs,KS, Curran,MA, Allison,JP (2006) CTLA4 blockade and GM-CSF combination immunotherapy alters the intratumor balance of effector and regulatory T cells. *J. Clin. Invest* 116(7): 1935-1945.

15. Liakou CI, et al. (2008) CTLA-4 blockade increases IFNgamma-producing CD4+ICOShi cells to shift the ratio of effector to regulatory T cells in cancer patients. *Proc Natl Acad Sci U S A* 105(39): 14987-14992.

16. Klages K, et al. (2010) Selective depletion of Foxp3+ regulatory T cells improves effective therapeutic vaccination against established melanoma. *Cancer Res.* 70(20): 7788-7799.

17. Bos PD, et al. (2013) Transient regulatory T cell ablation deters oncogene-driven breast cancer and enhances radiotherapy. *J. Exp. Med.* 210(11): 2435-2466.

18. Joshi NS, et al. (2015) Regulatory T Cells in Tumor-Associated Tertiary Lymphoid Structures Suppress Anti-tumor T Cell Responses. *Immunity* 43(3): 579-590.

19. Grinberg-Bleyer Y, et al. (2017) NF-kB c-Rel Is Crucial for the Regulatory T Cell Immune Checkpoint in Cancer. *Cell* 170(6): 1096-1108.

20. Simpson TR, et al. (2013) Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma. *J. Exp. Med.* 210(9): 1695-1710.

21. Peggs KS, et al. (2009) Blockade of CTLA-4 on both effector and regulatory T cell compartments contributes to the antitumor activity of anti-CTLA-4 antibodies. *J. Exp. Med.* 206(8): 1717-1725.

22. Wang W, et al. (2009) PD1 blockade reverses the suppression of melanoma antigen-specific CTL by CD4+ CD25(Hi) regulatory T cells. *Int. Immunol.* 21(9): 1065-1077.

23. Overacre-Delgoffe AE, et al. (2017) Interferon-gamma Drives Treg Fragility to Promote Anti-tumor Immunity. *Cell* 169(6): 1130-1141.

24. Darrasse-Jeze G, et al. (2009) Tumor emergence is sensed by self-specific CD44hi memory Tregs that create a dominant tolerogenic environment for tumors in mice. *J Clin. Invest* 119(9): 2648-2662.

25. Saito T, et al. (2016) Two $FOXP3^{+}CD4^{+}$ T cell subpopulations distinctly control the prognosis of colorectal cancers. *Nat Med* 22(6): 679-684.

26. Luo CT, et al. (2016) Graded Foxo1 activity in Treg cells differentiates tumour immunity from spontaneous autoimmunity. *Nature* 529(7587): 532-536.

27. Plitas G, et al. (2016) Regulatory T Cells Exhibit Distinct Features in Human Breast Cancer. *Immunity* 45(5): 1122-1134.

28. De Simone M, et al. (2016) Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells. *Immunity* 45(5): 1135-1147.

29. Michot JM, et al. (2016) Immune-related adverse events with immune checkpoint blockade: a comprehensive review. *Eur. J. Cancer* 54 139-148.

30. Postow MA, Sidlow,R, Hellmann,MD (2018) Immune-Related Adverse Events Associated with Immune Checkpoint Blockade. *N. Engl. J Med* 378(2): 158-168.

31. Grossman RL, et al. (2016) Toward a Shared Vision for Cancer Genomic Data. *N. Engl. J Med* 375(12): 1109-1112.

32. Lin W, et al. (2007) Regulatory T cell development in the absence of functional Foxp3. *Nat. Immunol* 8(4): 359-368.

33. Sefik E, et al. (2015) Individual intestinal symbionts induce a distinct population of RORg^{+} regulatory T cells. *Science* 349 993-997.

34. Fu J, et al. (2007) Increased regulatory T cells correlate with CD8 T-cell impairment and poor survival in hepatocellular carcinoma patients. *Gastroenterology* 132(7): 2328-2339.

35. Salama P, et al. (2009) Tumor-infiltrating FOXP3+ T regulatory cells show strong prognostic significance in colorectal cancer. *J Clin. Oncol.* 27(2): 186-192.

36. Tang Y, et al. (2014) An increased abundance of tumor-infiltrating regulatory T cells is correlated with the progression and prognosis of pancreatic ductal adenocarcinoma. *PLoS One* 9(3): e91551.

37. Zheng C, et al. (2017) Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* 169(7): 1342-1356.

38. Murai M, et al. (2009) Interleukin 10 acts on regulatory T cells to maintain expression of the transcription factor Foxp3 and suppressive function in mice with colitis. *Nat Immunol.* 10(11): 1178-1184.

39. Schneider MA, et al. (2007) CCR7 is required for the in vivo function of CD4+ CD25+ regulatory T cells. *J. Exp. Med.* 204(4): 735-745.

40. Chung Y, et al. (2011) Follicular regulatory T cells expressing Foxp3 and Bcl-6 suppress germinal center reactions. *Nat Med* 17(8): 983-988.

41. Zemmour D, et al. (2018) Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nat Immunol* 19(3): 291-301.

42. Blatner NR, et al. (2012) Expression of RORgammat marks a pathogenic regulatory T cell subset in human colon cancer. *Sci Transl. Med.* 4(164): 164ra159.

43. Schumacher TN & Schreiber,RD (2015) Neoantigens in cancer immunotherapy. *Science* 348(6230): 69-74.

44. Li B, et al. (2016) Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol.* 17(1): 174.

45. Aran D, Hu,Z, Butte,AJ (2017) xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18(1): 220.

46. Ferraro A, et al. (2014) Interindividual variation in human T regulatory cells. *Proc Natl Acad Sci U S A* 111(12): E1111-E1120.

47. Mele M, et al. (2015) Human genomics. The human transcriptome across tissues and individuals. *Science* 348(6235): 660-665.

48. Platt RJ, et al. (2014) CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 159(2): 440-455.

49. Liston A, et al. (2008) Differentiation of regulatory Foxp3+ T cells in the thymic cortex. *Proc Natl. Acad Sci U S. A.* 105(33): 11903-11908.

50. Ohnmacht C, et al. (2015) The microbiota regulates type 2 immunity through $ROR\gamma^{+}$ T cells. *Science* 349(6251): 989-993.

Figures

Figure 1

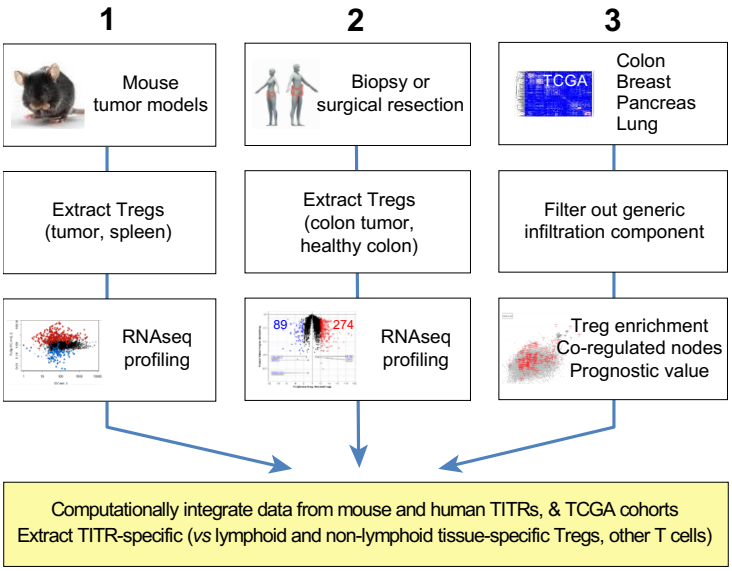


Figure 1. Schematic of multi-pronged work flow. This flow chart describes the generation of our three independent and cross-confirming datasets. **1.** Purification and profiling of Treg cells infiltrating three different transplantable tumors in immunocompetent mice. **2.** Purification of TITR cells from patients with colorectal tumors, and comparison of their gene expression profiles with those of Treg cells purified from normal human colon (many from the same donors). **3.** Mining of large datasets from TCGA for genes whose expression correlated with that of the Treg-defining factor *FOXP3*. Ultimately, these three datasets were combined to identify genes specifically overexpressed in TITRs.

Figure 2

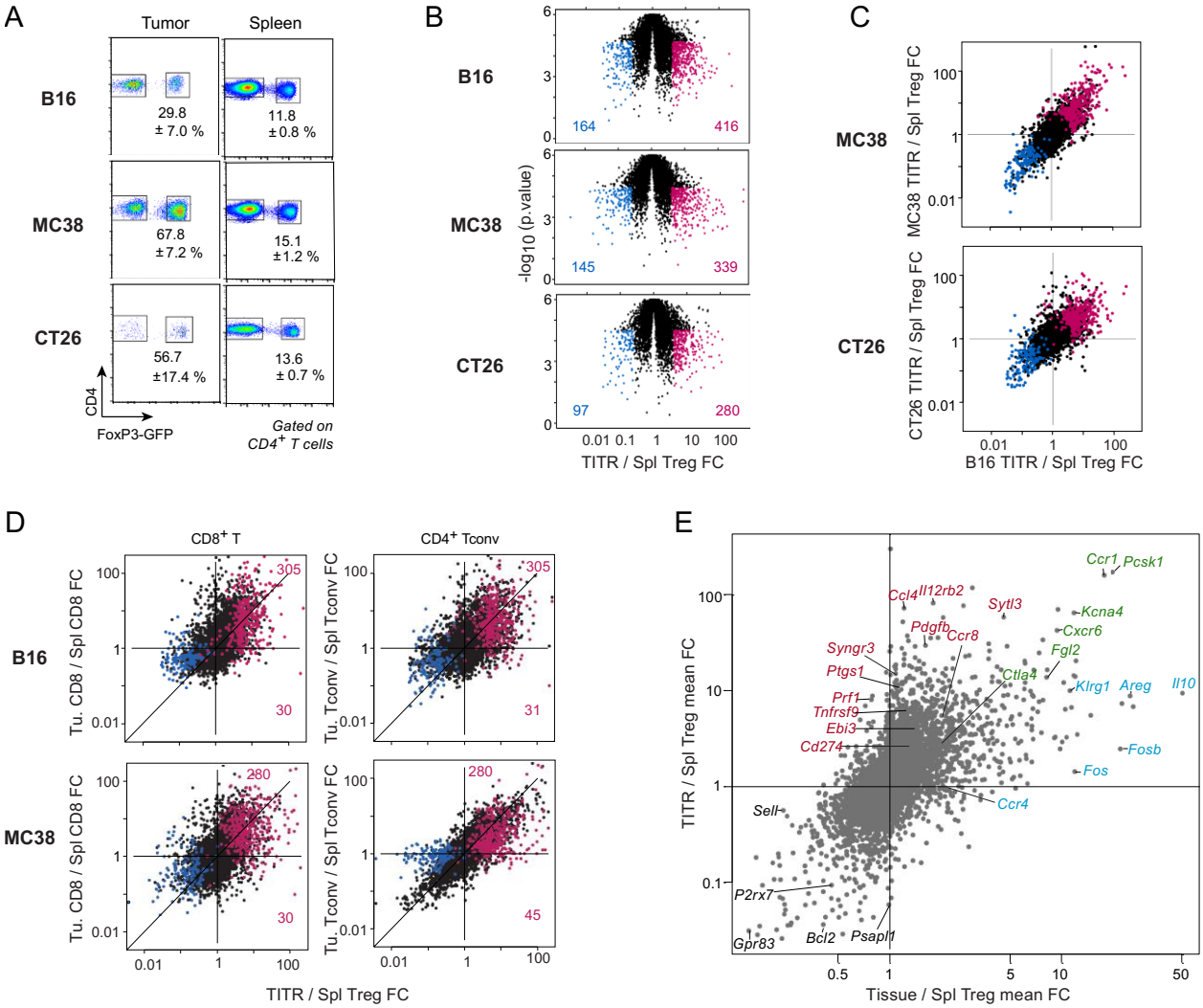


Figure 2. Identification of TITR signature. **A.** Exemplar gating used for FACSorting Tregs from three different mouse tumor models (B16, MC38, CT26). Mean ± SD of Treg population size (as % total CD4$^+$ T cells) is noted. **B.** Volcano plots showing the FoldChange (FC) in gene expression between TITR and splenic Tregs for each of the mouse tumor models examined. Genes with FC in expression ≥ 3 (red) or ≤ -3 (blue) in TITRs versus splenic Tregs are highlighted and enumerated. **C.** FC x FC plots depicting the FC in expression of genes in tumor versus spleen for one tumor type versus another tumor type (left) B16 x MC38, (right) B16 x CT26. Additive, filtered gene-sets (genes with FC in expression ≥ 3 (red) or ≤ -3 (blue) in TITRs versus splenic Tregs, for each of the three transplantable tumor models are highlighted. **D.** Comparison of tumor/spleen FC in CD8+ T or CD4+ Tconv with that seen for Treg cells. The additive TITR gene sets described in C are highlighted here. **E.** Comparison of transcriptomes between TITR and tissue-resident Treg cells. Mean FC in several tissue Tregs (visceral adipose tissue, injured muscle and colonic lamina propria) relative to splenic Tregs (x-axis) versus mean FC in TITR relative to splenic Tregs (average of tumor models noted above; y-axis) is shown.

Figure 3. Conservation of TITR signature across species and individual human colon cancer patients. A. Exemplar gating used for FACSorting Tregs from patient samples. Mean ± SD of Treg population size (as % total $CD4^+$ T cells) is noted. **B.** Transcriptomic profile of human colon tumor Tregs versus normal colonic mucosa Tregs. Plot shows FC and p values for the expression of each gene in tumor/normal colonic Tregs. Annotated genes include some known to be involved in Treg activity and/or costimulation. **C.** 408 genes (335 UP, 73 DN) were selected that distinguish TITR from normal colon Tregs. FC (TITR/normal colon Tregs) values for these modulated genes for individual patients are presented in the heatmap. **D.** Overlap of our TITR transcriptome (408 genes) with recently published Treg dataset from breast cancer. Mean FC in breast cancer TITR relative to normal breast parenchyma (NBP) Tregs (x-axis) versus mean FC in CRC TITR relative to normal colonic Tregs (y-axis) is shown. **E.** Top 10 hits from Enrichr motif analysis of gene set preferentially upregulated in TITR compared to normal colonic Tregs (335 UP genes). Overlap denotes the number of transcripts from our signature/number of transcripts known to be associated with a given transcription factor. **F.** Mouse TITR signature is highlighted in red on human tumor/ normal colonic Treg data. **G.** Human TITR transcripts are highlighted in red on mouse tumor/tissue Treg comparison (per Fig. 2E).

Figure 3-Cont.

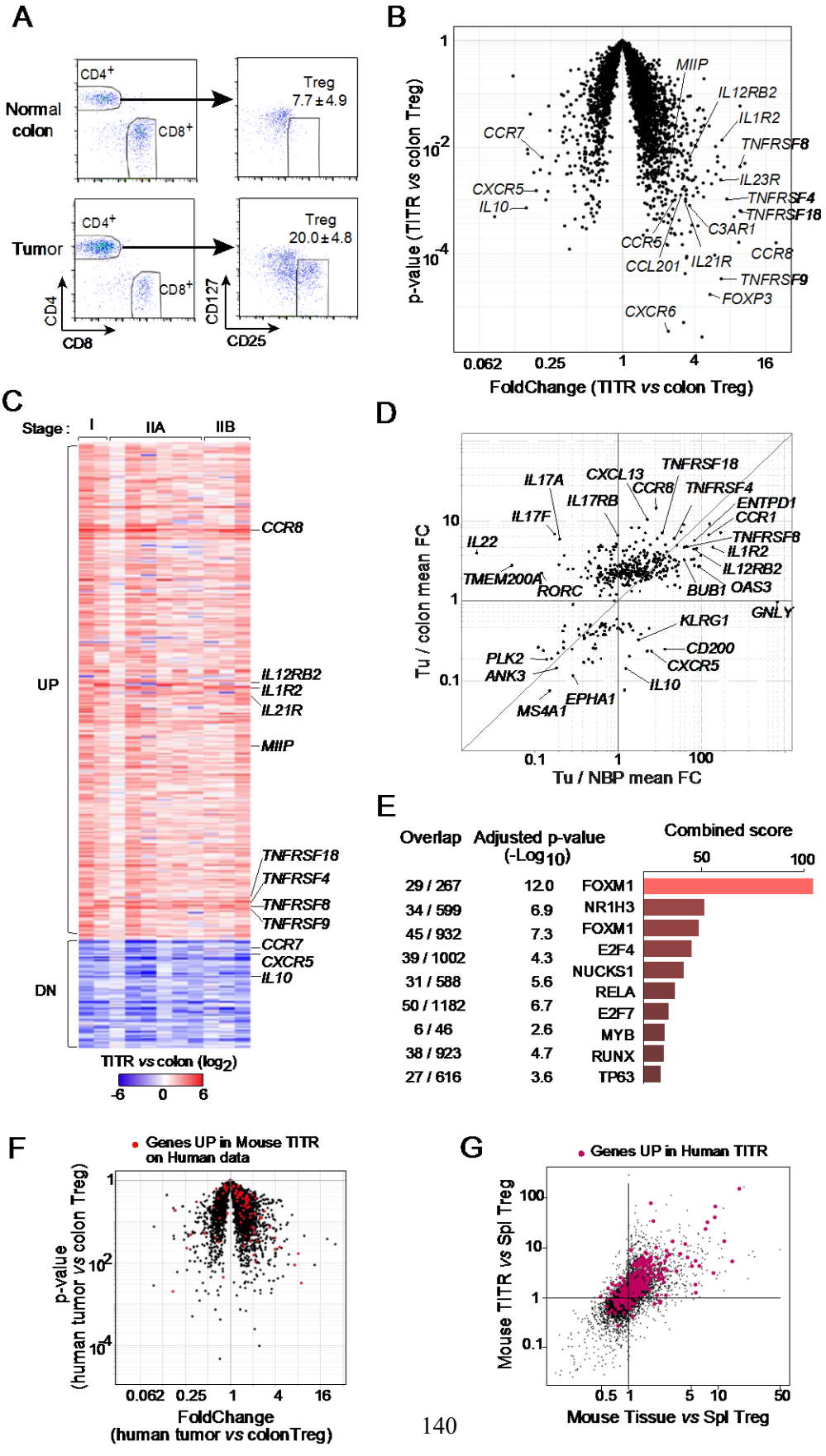


Figure 4

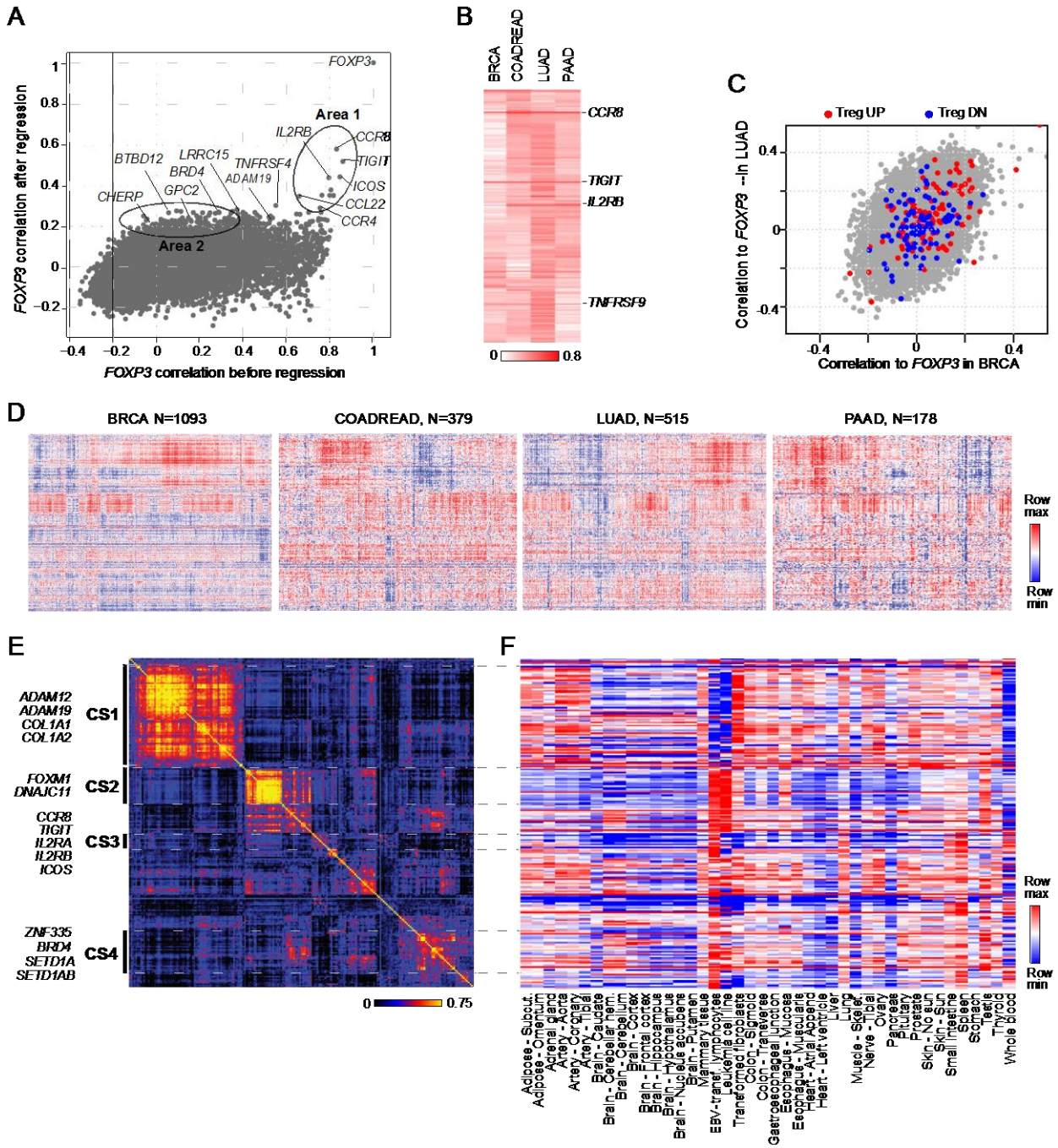


Figure 4. Correlation to *FOXP3* in TCGA. A. Correlation to *FOXP3* before and after removal of immune cell component. Genes with higher correlation to *FOXP3* after immune filtrate removal are highlighted. **B.** Correlation to *FOXP3* after immune infiltrate removal across four TCGA cancer datasets. **C.** Canonical Up and Down Treg signature highlighted on post-regression correlation with *FOXP3* on LUAD vs BRCA datasets. **D.** Variability of the 219 transcripts with the highest post-regression correlation in the four TCGA datasets. **E.** Co-expression matrix of 219 transcripts averaged across four tumors. Four gene-sets are highlighted. **F.** Tissue Expression distribution of 219 transcripts in GTEx.

Figure 5

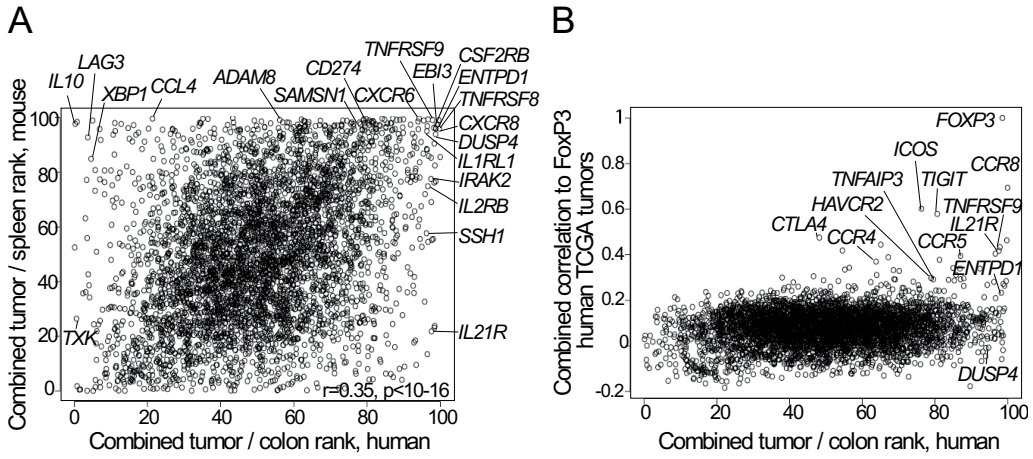


Figure 5. Combinatorial data integration. A. Comparison of overall scores for human (x-axis) and mouse (y-axis) TITRs. Highlighted/annotated are genes either at the top of the ranking in both species, with high scores in the mouse and still in top 10% of differential transcripts in human, or highest in the human ranking, but not in the mouse. **B.** Depiction of overall score for human TITRs (x-axis) versus average correlation to Foxp3 score derived from the whole-tumor TCGA datasets (y-axis).

Figure 6

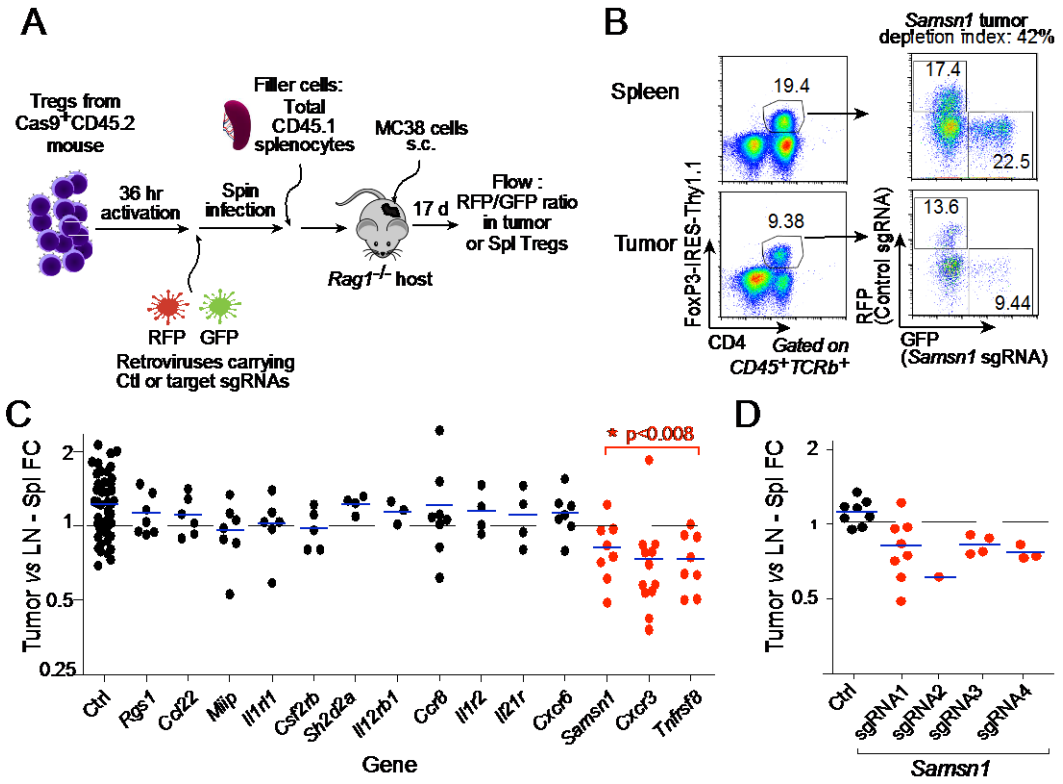
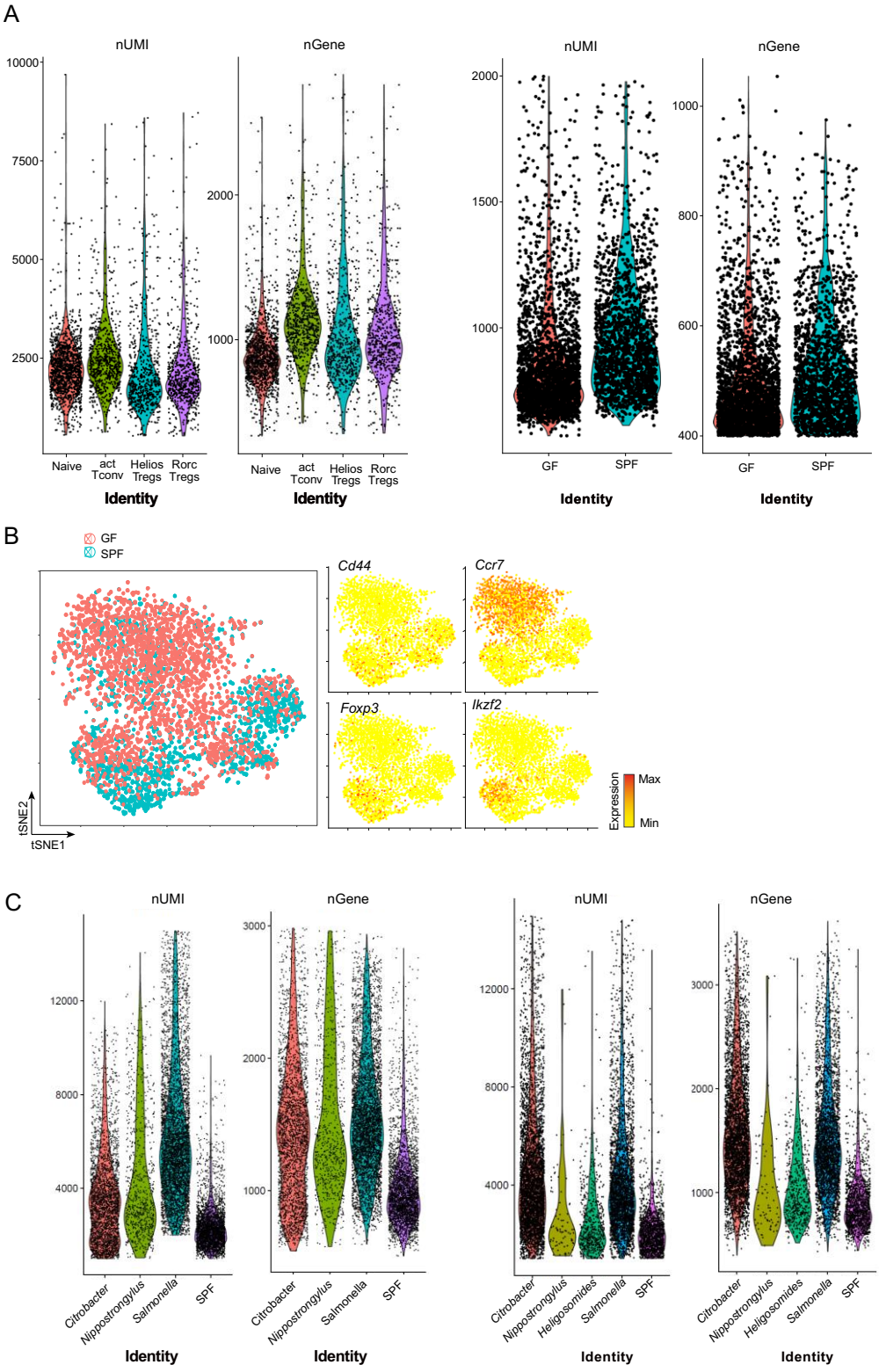


Figure 6. CRISPR-based Treg knockouts. A. Schematic depiction of protocol to induce loss of function (LOF) mutations in TITR target genes, specifically in Treg cells, utilizing the CRISPR/Cas9 system. **B**. Exemplar gating used for determining the tumor depletion index- the ratio of the percentage of GFP or RFP+ Tregs in tumor vs. spleen. **C.** Summary of CRISPR-based Treg knockout data. The tumor depletion index was significant in the genes highlighted in red. **D.** Validation of Samsn1 LOF on Treg accumulation in tumors. Four Different sgRNAs showed decreased accumulation of Samsn1 LOF Tregs in the tumors.

Appendix B: Supplementary Figures and Tables

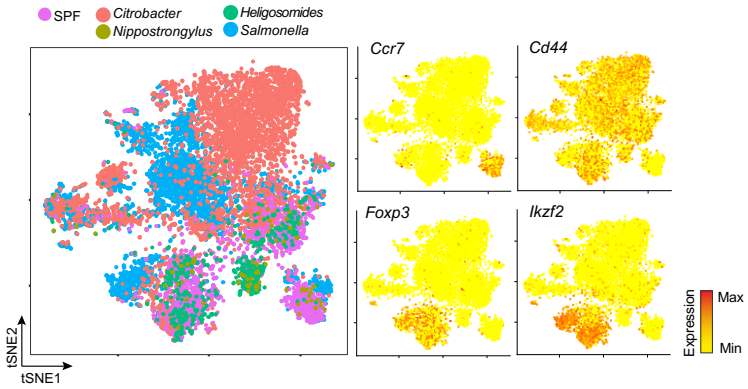
Figure S1: A) QC of SPF scRNA-seq data, representing the number of UMI and Genes expressed in the dataset. Right panel- QC for SPF and GF scRNA-seq data from 10X V1.
B) t-SNE representation of total LP CD4 T cells from SPF and GF mice. Right panel: gene expression of key genes that identify the main clusters.
C) Left panel: QC of Hashed scRNA-seq data. Right panel: QC of non-hashed replicates scRNA-seq data. Numbers of UMI and Genes expressed in the dataset are shown as violin plots.

Supplementary figure 1

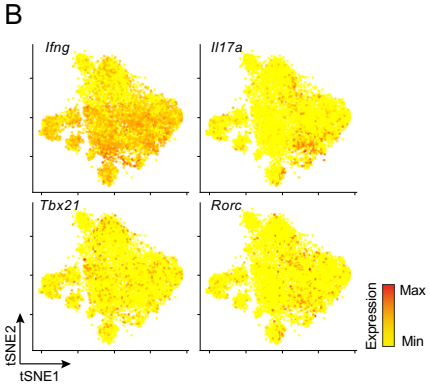


Supplementary figure 2 (1)

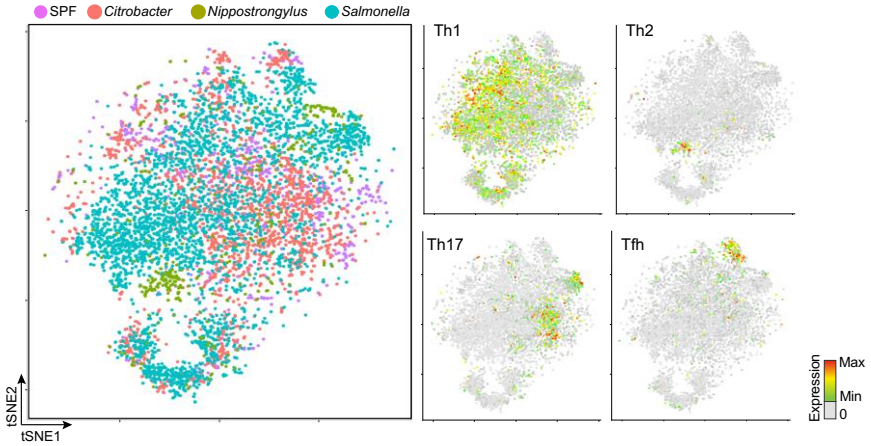
A



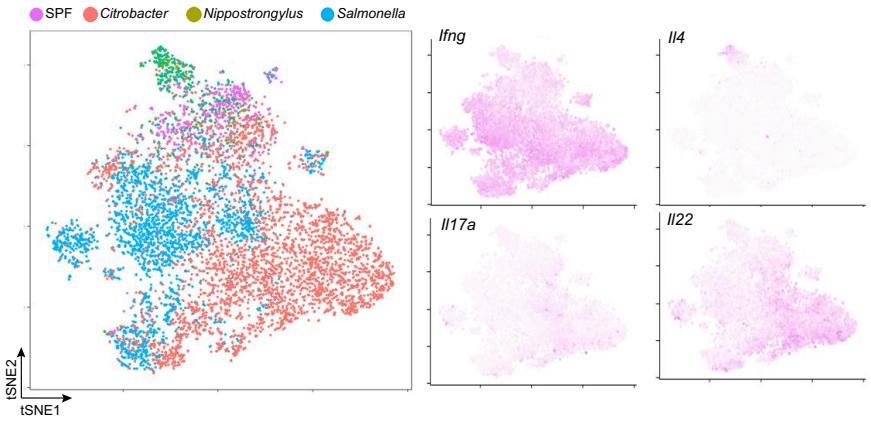
B



C



D



Supplementary figure 2 (2)

E

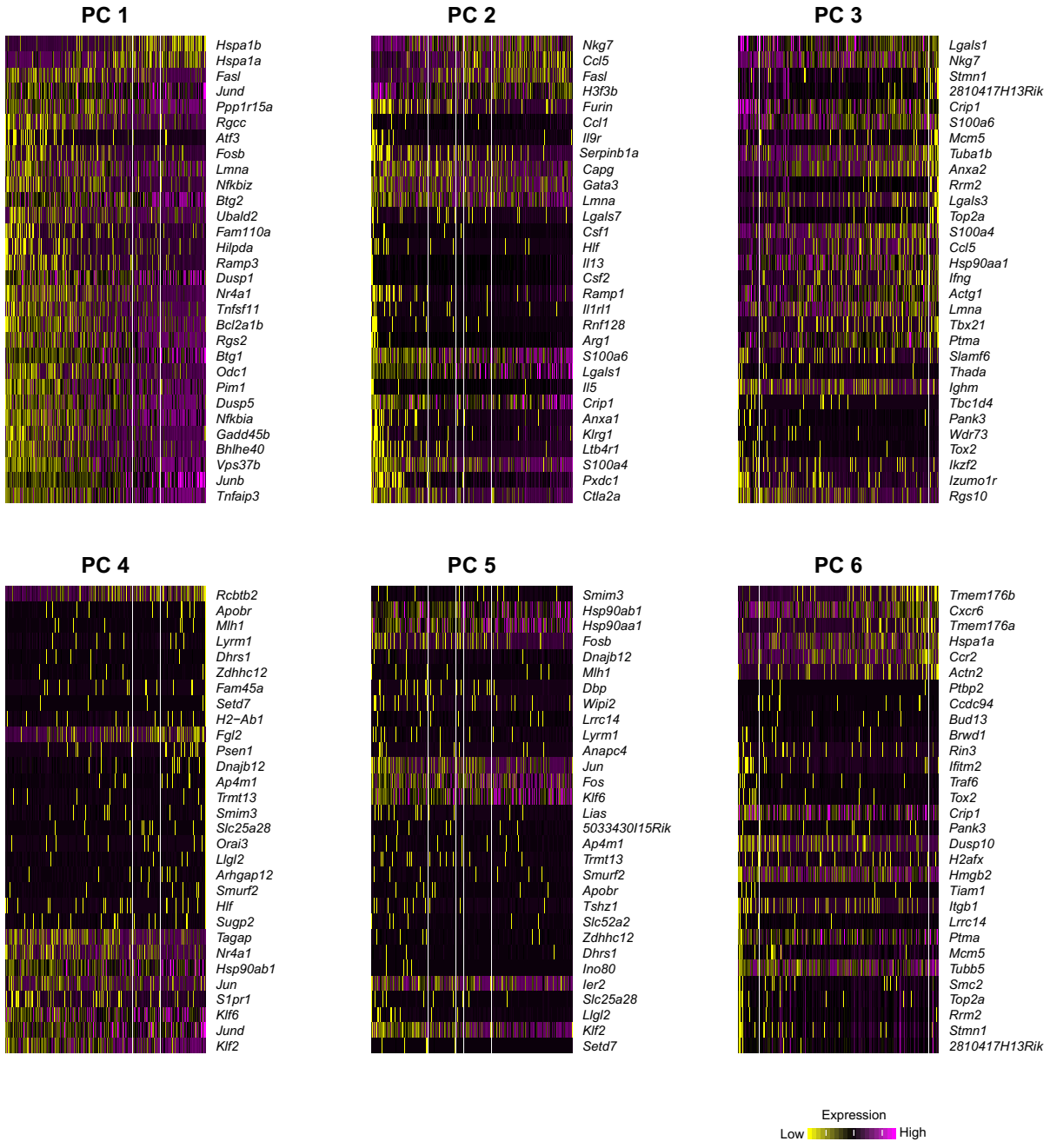


Figure S2: A) t-SNE representation of all CD4 T cells from hashed experiment by sample.
B) Gene expression of key cytokines and TFs in Teff cells
C) CCA overlay of non-hashed single-cell replicates. T-SNE represents pre-selected Teff cells. Right panel represents expression of curated Th signatures in each cell.
D) t-SNE calculated by imputed values of variable genes. Right panel: Gene expression of key cytokines and TFs in Teff cells.
E) First 6 PCs of scRNA-seq data from hashed sample.

Figure S3: Clustering approaches. A) Diffusion Map of hashed data, colored by sample. Right panel: Overlay of Th signatures on diffusion map.

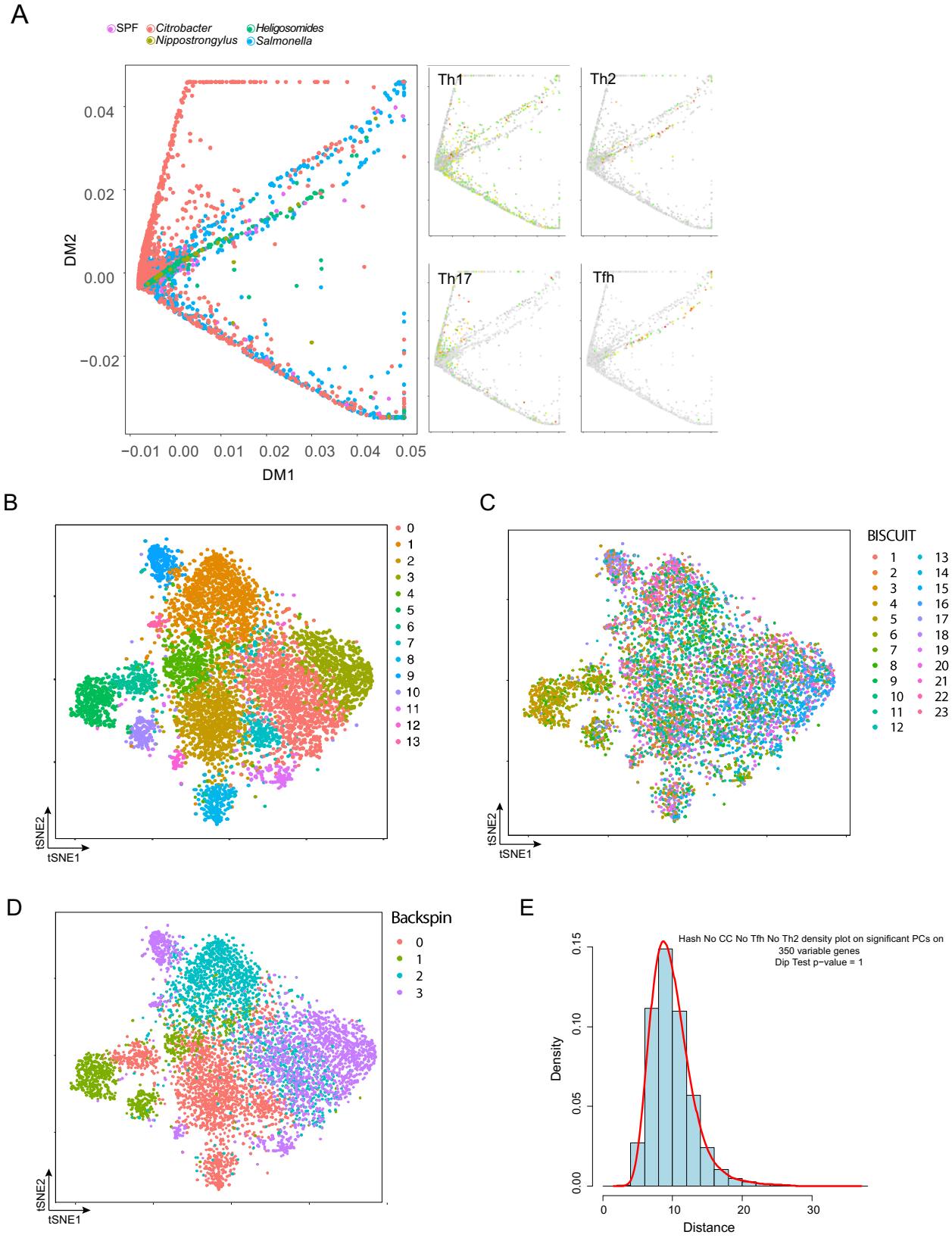
B)KNN clustering of hashed data.

C) BISCUIT clutsering of hashed data.

D) Backspin clustering of hashed data.

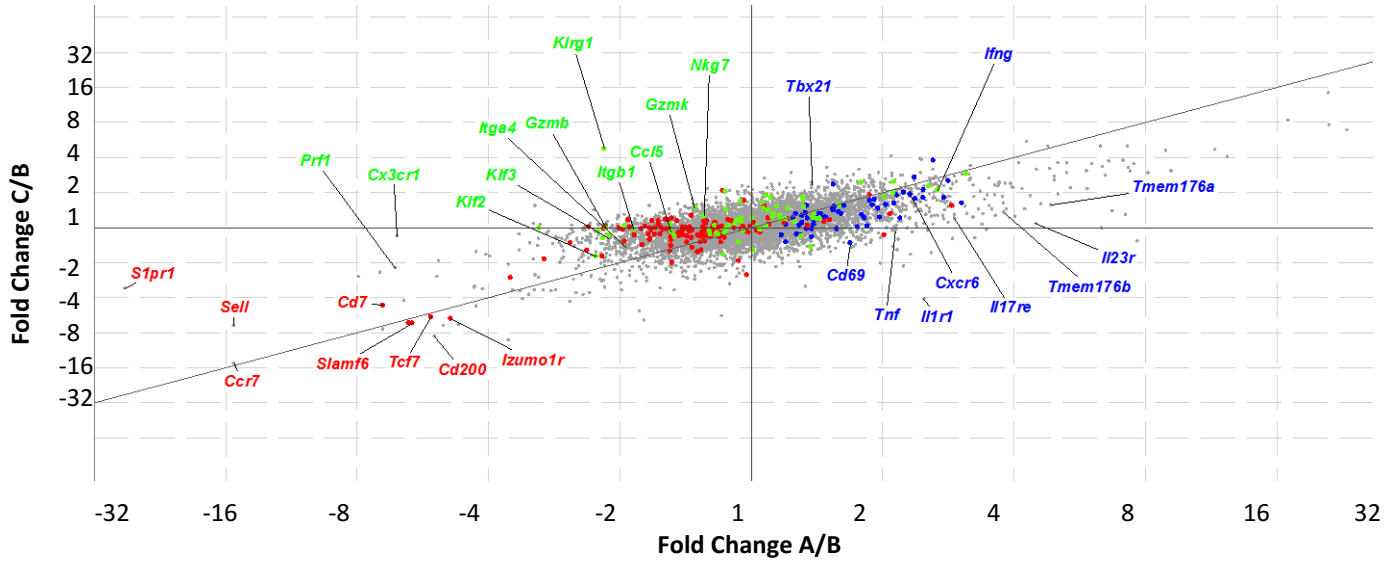
E) Hartigan's dip test, done on 350 most variable genes in hashed data from "main cloud" (clusters 0,1,2,3,4,9,11)

Supplementary figure 3

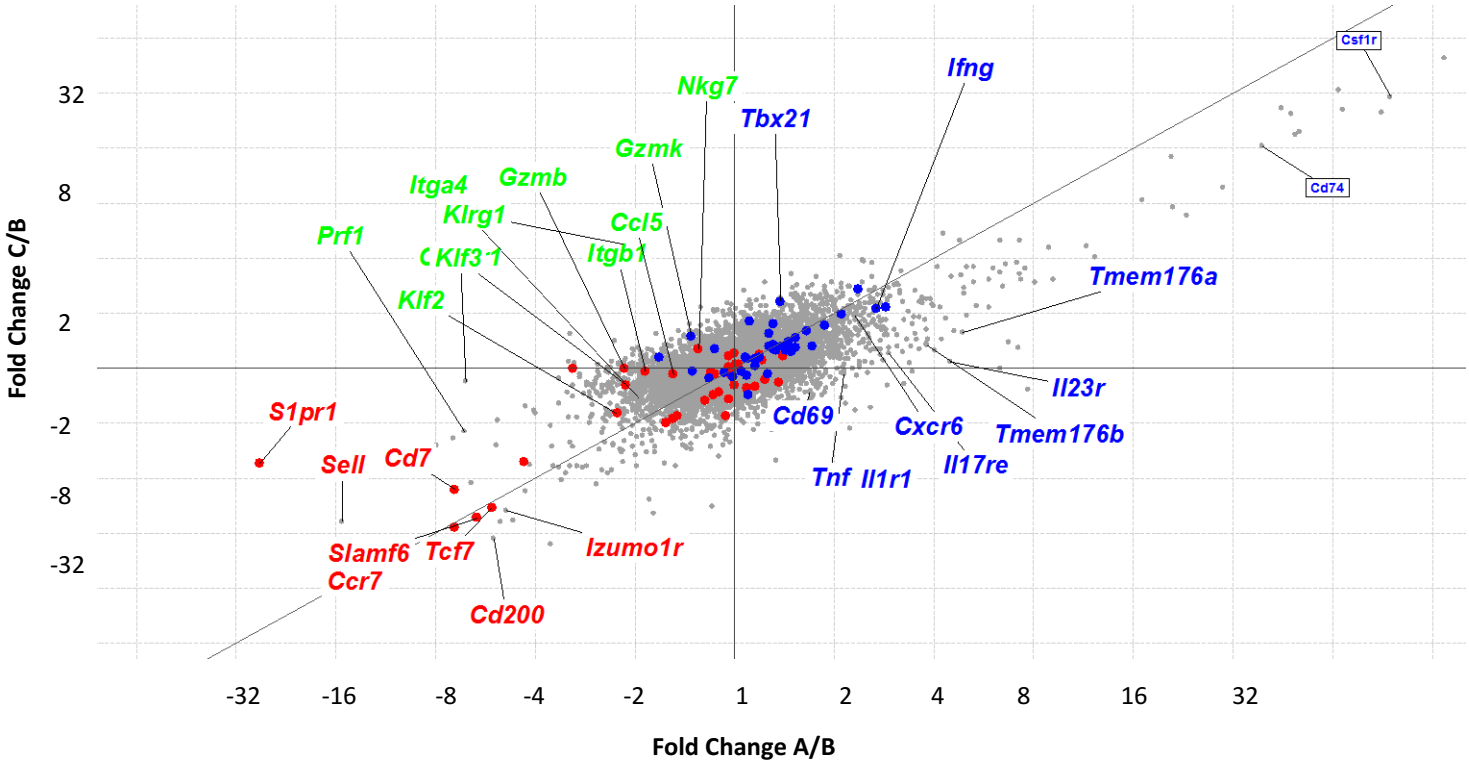


Supplementary Figure 4

A



B



Supplementary Figure 4-cont.

C

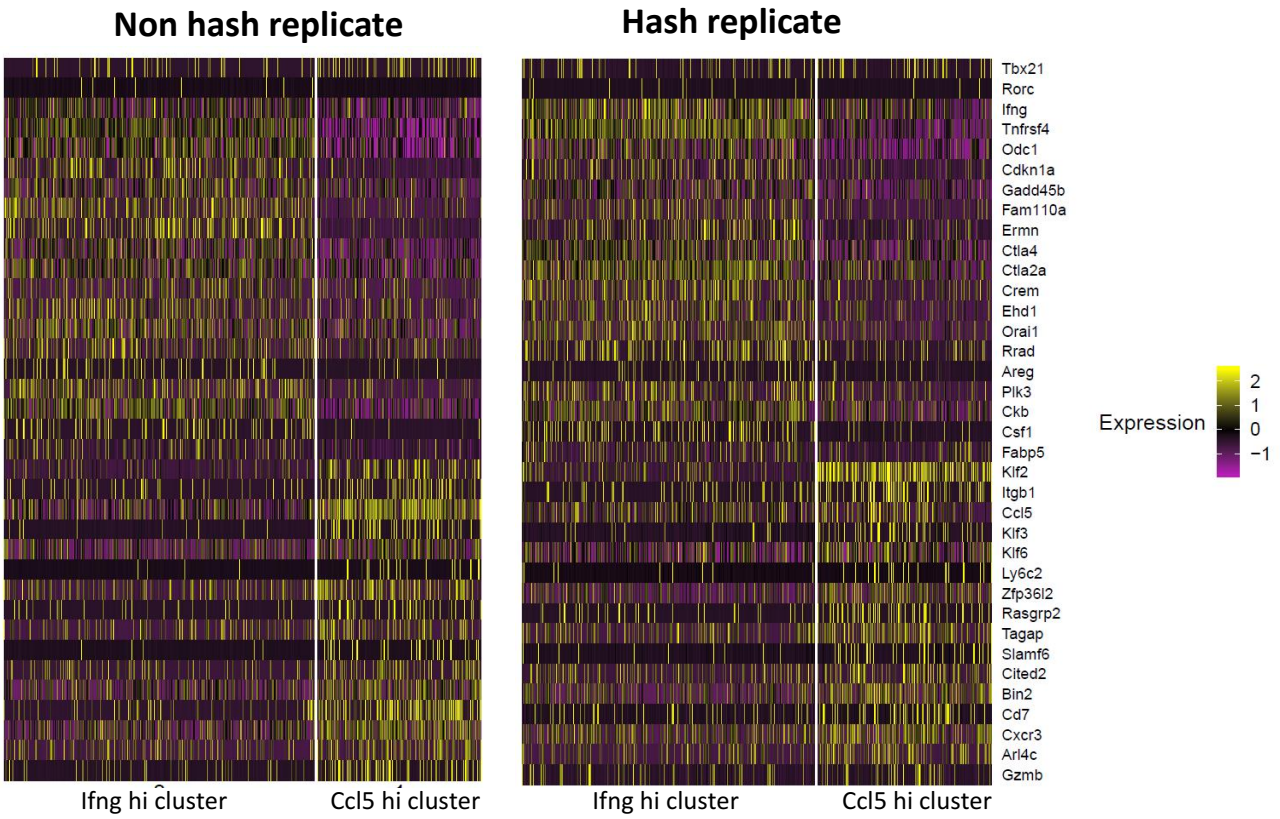


Figure S4: A) Fold change/fold change plot of RNA-Seq data comparing between populations A, B and C.

Blue, red, green dots- genes high in scRNA-seq data of populations A, B, C respectively.

B) Same as A but blue dots represent genes that are in module M11-Type1, while red dots represent genes in M12-CTX/Ccl5.

C) Heatmap of differentially expressed genes between *Ifng* hi and *Ccl5* hi cell clusters in both hashed and non-hashed Salmonella replicates. *Tbx21* and *Rorc* are in the top rows to show that these cells are in the "Th1-pole".

Supplementary figure 5

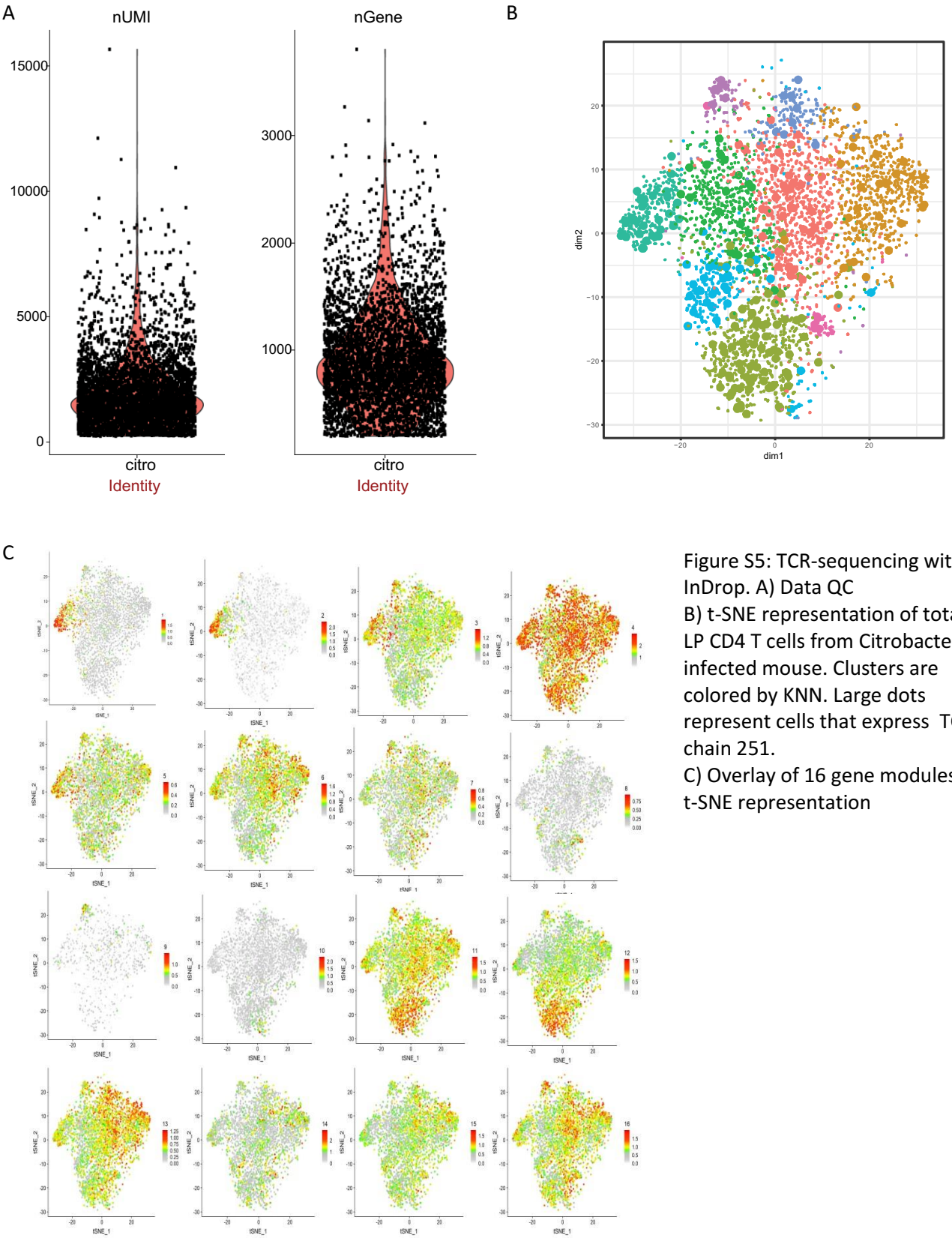


Figure S5: TCR-sequencing with InDrop. A) Data QC
B) t-SNE representation of total LP CD4 T cells from Citrobacter infected mouse. Clusters are colored by KNN. Large dots represent cells that express TCR chain 251.
C) Overlay of 16 gene modules on t-SNE representation

Supplementary figure 6

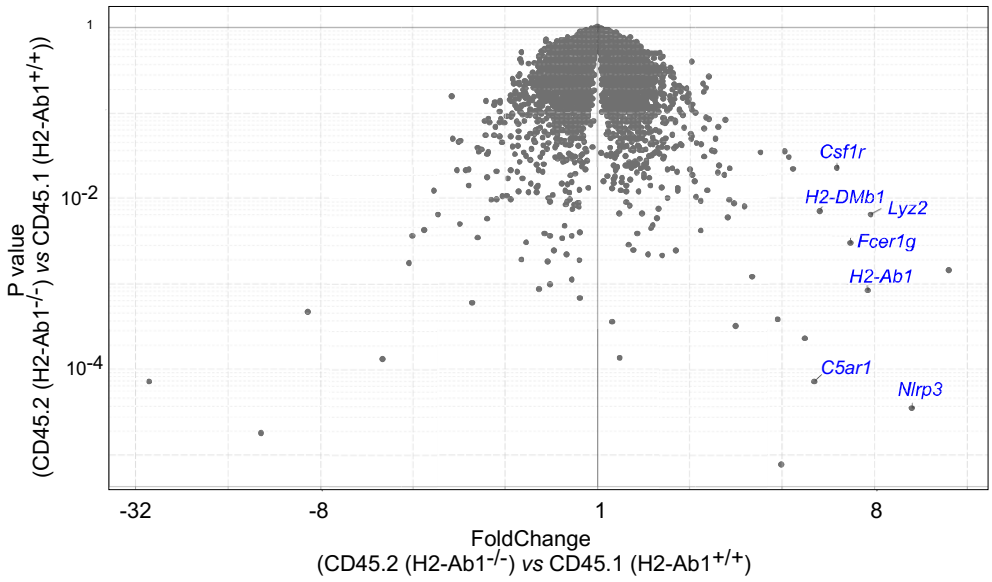


Figure S6: Volcano plot of RNA-seq data comparing LP CD4 T cells from mixed bone marrow chimera mice (CD45.2 H2-Ab1-/- and CD45.1 H2-Ab1+/+). Cells were sorted as CD45.1+ or CD45.2+ CD4+TCRβ+CD44hi CD25lo.

Table S1: scRNA-seq datasets used

	Infection/mouse	Platform	Batch #	Total cells passing QC	Total Teff	QC threshold UMI/cell	QC threshold genes/cell	Median UMI/cell	Median genes/cell
Hash:				9348	7107	1000	400	3394	1345
	SPF	10X V2	3	1390	512	1000	400	1978	858
	S. typhimurium	10X V2	3	3193	2514	1000	400	3721	1433
	C. rodentium	10X V2	3	4143	3685	1000	400	4002	1494
	H. polygyrus	10X V2	3	542	348	1000	400	2181	976
	N. brasiliensis	10X V2	3	80	48	1000	400	2186	905
Non-Hash	SPF	10X V2	1	3360	826	600	400	2137	951
	S. typhimurium	10X V2	2	5454	3916	1000	500	5870	1386
	C. rodentium	10X V2	1	3051	2311	540	320	3238	1521
	N. brasiliensis	10X V2	3	875	538	1000	500	3847	1426
	GF?	10X V2	1	3517	91	500	300	1790	804
10X V1	SPF	10X V1	4	1782	680	600	400	910	495
	GF	10X V1	4	2644	630	570	400	793	464
Indrop	*C. rodentium*	Indrop	5	1715	NA	250	200	797	502

Table S2: Curated Th signatures

Th1	**Th2**	**Th17**	**Tfh**
Ifng	Gata3	Rorc	Cxcr5
Tbx21	Il4	Il17a	Izumo1r
Cxcr3	Il5	Il17f	Nt5e
Nkg7	Il13	Il22	Bcl6
Fasl	Il10	Ccr6	Slamf6
Il2	Areg	Il23r	Cd200
Tnf	Maf	Tmem176a	Ccr7
Il12rb1	Il17rb		Tcf7
Il12rb2			
Gzmb			