Influenza infection recruits distinct waves of regulatory T cells to the lung that limit lung resident IgA+ B cells

Louisa E. Sjaastad¹²³, David L. Owen¹²³, Sookyong Joo¹²³, Todd P. Knutson⁴, Christine H. O’Connor⁴, Braedan McCluskey⁴, Rebecca S. LaRue⁴, Ryan A. Langlois¹⁵, and Michael A. Farrar¹²³*.

Affiliations:
¹Center for Immunology, ²Masonic Cancer Center, ³Department of Laboratory Medicine and Pathology, ⁴Minnesota Supercomputing Institute, ⁵Department of Microbiology and Immunology, University of Minnesota, Minneapolis, MN 55455

*Address correspondence to MAF at: farra005@umn.edu
Abstract

The role of regulatory T cells (Tregs) in limiting responses to pathogens in tissues remains poorly described. We used scRNA-Seq and a newly generated Foxp3-lineage reporter line (Foxp3-iDTR mice) to track Tregs in the lungs and peripheral blood following infection with influenza virus. Few Tregs of any type were found in the lung at steady-state. Following influenza infection Tregs expressing a strong interferon-stimulated gene signature (ISG-Tregs) appeared by day 3, peaked by day 7, and largely disappeared by day 21 post-infection. A second diverse wave of tissue-repair-like Tregs (TR-Tregs) appeared by day 10 and were maintained through day 21 post-infection. These two distinct Treg subsets had different gene expression patterns and distinct TCR repertoires. To establish the role of Tregs during influenza infection, we acutely ablated Tregs at day 6 post-infection; this resulted in a significant increase in IgA+ B cells in the lung. To determine whether distinct Tregs subsets could also be observed in response to respiratory viral infections in humans we analyzed scRNA-Seq datasets of patients with COVID-19. Peripheral blood from healthy human volunteers had multiple Treg subsets defined by unique gene expression patterns, but few ISG-Tregs. In contrast, two distinct Tregs subsets were expanded in COVID-19 patients - ISG-Tregs and IL32 expressing Tregs (16-fold and 2-fold increased, respectively). ISG-Tregs were present at significantly higher levels in patients with mild versus severe COVID-19, while IL32 expressing Tregs showed the opposite pattern. Thus, the Treg response to respiratory viruses in humans is also diverse and correlates with disease outcome.
Introduction

Regulatory T-cells (Tregs) protect against autoimmunity and inappropriate immune responses to commensal microorganisms and dampen immune responses to pathogens (Josefowicz et al., 2012; Sjaastad et al., 2021; Veiga-Parga et al., 2013). Although Tregs were initially described as a monomorphic population it soon became evident that they exhibited considerable heterogeneity. Tregs were initially broken down into a population of quiescent central Tregs that circulate through secondary lymphoid tissues and depend on IL-2 for survival, and a second population of effector Tregs that exhibits an activated phenotype, is prevalent within tissues, and requires signaling through ICOS (Smigiel et al., 2014). Subsequent studies demonstrated that during an active immune response, Tregs can express transcription factors that define distinct T-helper cell lineages and enable these Tregs to limit the corresponding T-helper immune response (Harrison et al., 2019; Koch et al., 2009; Sefik et al., 2015; Zheng et al., 2009). Mice bearing Tregs that are deficient in these transcription factors develop uncontrolled inflammation in multiple contexts (Chaudhry et al., 2009; Harrison et al., 2019; Levine et al., 2017; Wohlfert et al., 2011). More recently, single-cell RNA-sequencing (scRNA-seq) studies have allowed the examination of Treg heterogeneity in more depth (Lu et al., 2020; Miragaia et al., 2019; Owen et al., 2022; Owen et al., 2019). These studies demonstrated transcriptional heterogeneity in Tregs within the spleen, lymph nodes, thymus, lung and colon at steady-state, and that Tregs undergo a step-wise acquisition of a tissue-like transcriptome as they move from the lymph nodes to the tissue itself (Delacher et al., 2020; Miragaia et al., 2019). While these studies have profiled Treg diversity at steady state, less is known about the heterogeneity of Tregs within tissues in the context of an activated immune response. In particular, there is limited understanding of how infection impacts Treg heterogeneity within the infected tissue. Herein we demonstrate that Tregs infiltrate the lungs following influenza virus infection and that they do so in distinct waves. An initial wave consisting of Tregs with an IFN-stimulated gene expression signature (ISG-
Tregs) and distinct TCR repertoire arises within 7 days post infection before being replaced by a second wave of highly diverse Tregs that express numerous genes associated with tissue repair (TR-Tregs) such as Areg and Tgfb1. Collectively, Tregs play an important role in limiting the number of IgA+ expressing B cells in the lung. Distinct Treg subsets including the ISG-Treg subset are also observed in human patients following infection with SARS COV2. Healthy volunteers have very few ISG-Tregs but they expand greatly in response to infection with SARS-CoV-2 as does a distinct IL32-expressing subset. Moreover, ISG-Tregs were ~50% more abundant in patients with mild versus severe COVID-19; conversely, IL32+ Tregs were selectively increased in patients with severe COVID-19. Thus, our findings describe diverse subsets of Tregs that respond rapidly to respiratory viral infections, and which correlate with disease severity in human patients.

Results

During infection with influenza A virus, Tregs infiltrate and expand within the lung tissue. To examine this in more detail, WT mice were infected with influenza A and analyzed at several time points post-infection. Prior to tissue harvest we injected mice iv with an antibody to CD4 to identify CD4⁺ cells in the blood versus those in the lung tissue (Anderson et al., 2014). In uninfected mice the majority of Tregs isolated from the lungs are in the blood (~90%) while at 10 days post infection the majority of Tregs are in the lung tissue (~85%) (Fig. S1A). Over the course of the infection, the number of Tregs within the lungs increased substantially and then contracted following virus clearance (Fig. 1A). This was primarily due to an increased number and frequency of Tregs within the lung tissue (Fig. 1A, 1B). At 10 days post infection, nearly 50% of Tregs within the lung tissue expressed the Th1-associated transcription factor, TBET, while expression of the Th2 and Th17-associated transcription factors, GATA3 and RORyt respectively, remained much lower (Fig. 1C). This result fits with past observations that influenza A virus infection stimulates a strong type I immune response. TBET-expressing Tregs
have been identified in several other studies and play an important role in limiting Th1-driven inflammation. While this result demonstrated that TBET+ Tregs expand in the lungs during influenza A infection, the degree of Treg heterogeneity during an influenza infection and their functions within the lungs during the infection remained unclear.

To interrogate the phenotype of Tregs responding to the lungs during influenza infection, we performed single-cell RNA-sequencing (scRNA-seq). Tc11b+ Tcrav1× Foxp3-GFP mice received temporally staggered infections with influenza A virus (PR8 strain) (Fig. 2A). Prior to tissue harvest we injected mice iv with a CITE-Seq antibody to CD4 to identify CD4+ cells in the blood (Ab+) versus those in the lung (Ab−) (Anderson et al., 2014). We used oligonucleotide-conjugated tags (“hashtag” antibodies) to identify cells obtained at different timepoints after infection, as well as mark distinct biological replicates. CD4+FOXP3+ Tregs were isolated from the lung by FACS and single cell libraries were generated. After quality control and removal of doublets, transcriptomes from 17,938 combined single cells were analyzed. A dimensional reduction with graph-based clustering approach was applied to combined data from all five timepoints to identify distinct Treg subsets; this resulted in a final UMAP plot with thirteen distinct clusters (Fig. 2B, Table S1). Tregs derived from experimental replicates at each timepoint were phenotypically reproducible (Fig. S1B). Distinct Treg subsets were present within the lung tissue and blood fractions and arose with different kinetics (Fig. 2B, 2C, S1C). A distinct cluster of ISG-Tregs (cluster 3) emerged at 7 days post infection (dpi) and contracted upon resolution of the infection (10 and 21 dpi) (Fig. S1C, S1D, 2C). In the ISG-cluster, all the top differentially expressed genes were ISGs. In addition to the ISG-Treg cluster, this analysis also demonstrated intriguing dynamic heterogeneity within the lung Tregs across the infection time course. Cluster 1 Tregs were present primarily in the blood and expressed transcripts found in central Tregs, such as Ccr7 (Fig. S1D). Cluster 4 Tregs appear to represent a transitional early
effector Treg subset that branches off into three distinct lineages (Fig. 2D) that we will
tentatively refer to as tissue-resident-like Tregs (cluster 0), tissue-repair-like (TR)-Tregs
(clusters 2, 5, 6, 8, 10, and 12) and ISG-Tregs. Cluster 0 was defined by transcripts indicative of
early effector-like Tregs such as $S100a10$, $S100a4$, and $S100a6$ (Fig. S1D) (Permanyer et al.,
2021; Zemmour et al., 2018). Transcripts for $Areg$ were highly expressed in cluster 2, indicative
of Tregs with tissue-repair functions (Fig. S1D). This is consistent with studies demonstrating
that deletion of $Areg$ in all Tregs leads to substantial defects in lung function in mice infected
with influenza (Arpaia et al., 2015). Cluster 2 was closely linked to clusters 5, 6, 8, 10 and 12,
which we collectively refer to as TR-Tregs. Tregs within cluster 5 express gene transcripts
associated with activation including $Ccl5$ and $Nkg7$ (Fig. S1D). Cluster 6 represents granzyme
B-expressing Tregs (Fig. S1D), which have a previously described role in limiting lung
inflammation during RSV viral infection (Loebbermann et al., 2012). Cluster 8 Tregs uniquely
expressed transcripts for $Plac8$ (Fig. S1D), a placental protein that has been observed in Tregs
but whose function is unclear. Cluster 10 Tregs expressed transcripts for $Penk$ (Fig. S1D), an
endogenous opioid, that was recently shown to be expressed in Tregs from UVB-irradiated skin
that promote wound healing (Shime et al., 2020). Finally, Tregs within cluster 12 expressed $Tff1$
(Fig. S1D), a member of the trefoil factor protein family, which has been shown to promote
repair of gut epithelium (Playford et al., 1996). Transcripts for $Tff1$ have been found in IL18R+
Tregs within the thymus (Peligero-Cruz et al., 2020) while accessible chromatin at the $Tff1$ locus
has been identified in ST2-expressing Tregs from the colon and skin (Delacher et al., 2020).
Thus, Tregs in clusters 2, 5, 6, 8, 10, and 12 all express transcripts that have been associated
with distinct aspects of tissue-repair.

ISG-Tregs are distinct from other Treg lineages that migrate to the lung during influenza virus
infection. Using slingshot trajectory analysis, we observed that central Tregs (cluster 1) progress
into cluster 4 Tregs (an early effector Treg-state) and then branch off into 3 distinct lineages: ISG-Tregs, TR-Tregs (clusters 2, 5, 6, 8, 10, 12), and tissue-resident-like Tregs (cluster 0) (Fig. 2D). In support of this conclusion, TCR repertoire analysis revealed a high degree of similarity between Treg TCRs from biological replicate mice within the ISG-Treg cluster (Fig. 2E).

Likewise, there was considerable overlap between TCRs from TR-Tregs between independent mice. In contrast, much lower similarity between TCRs was observed when comparing ISG-Tregs with the TR-Treg subset, even when these populations came from the same mouse (Fig. 2E). This exclusive relationship held even when comparing ISG-Treg TCRs with TCRs from Tregs in each of the other clusters alone (Fig. S2A). Finally, TR-Treg clusters expressed Itgav and Itgb8, which pair to form integrin αvβ8 (Fig. S2B). These Tregs also express Tgfβ1 mRNA (Fig. S2B). Previous studies have shown that integrin αvβ8 on Tregs is important for the cleavage and activation of TGFβ and subsequent suppression of effector T-cell responses during inflammation (Worthington et al., 2015). Thus, TR-Tregs may be involved in lung remodeling following influenza virus infection. This pattern of transcript expression was not characteristic of all integrins, as transcripts for integrin β1 were preferentially expressed in cluster 0 Tregs (Fig. S2B). TR-Treg clusters also expressed higher levels of transcripts for chemokine receptors, Cxcr6 and Ccr8, as well as activation induced markers, Klrγ1, Il1r2, and Pdcd1 (Fig. S2B). In contrast, ISG-Tregs had limited expression of markers characteristic of TR-Tregs, but instead uniquely expressed high amounts of the chemokine Cxcl10 (Fig. S2B). Thus, ISG-Tregs are a unique Treg lineage that appear early in the lung tissue during influenza virus infection and are phenotypically distinct from other Treg lineages that arrive later.

We have not found reliable surface markers to specifically identify ISG-Tregs. Therefore, to interrogate the dynamics and function of ISG-Tregs, we designed a mouse model that allows one to track distinct Treg lineages using CRE-reporter mice. Specifically, we used
CRISPR/CAS9 based approaches to insert a LoxP-flanked STOP cassette upstream of the IRES-DTR-GFP sequence of Foxp3-DTR mice (Kim et al., 2007) to enable CRE-inducible expression of DTR and GFP (Fig. S3A). Insertion of the floxed STOP cassette was confirmed by PCR and whole genome sequencing (Fig. S3B, S3C). In the absence of CRE the DTR-GFP cassette is not expressed. In contrast, Cd4-Cre x Foxp3-iDTR/WT heterozygous mice resulted in ~50% of splenic Tregs (CD4+CD25+ T-cells) expressing GFP (Fig. S3D). Foxp3-iDTR mice also allow for inducible expression of DTR/GFP when crossed with mice expressing a tamoxifen-inducible CRE such as CD4-CREERT2 (Fig. S3E). Finally, we confirmed that diphtheria toxin (DT) administration deletes labeled GFP+ Tregs in Cd4-CreERT2 x Foxp3-iDTR mice just as efficiently as in the original Foxp3-DTR mouse (Fig. S3F).

To identify ISG-Tregs we crossed Foxp3-iDTR mice to Mx1-Cre mice, as Mx1 is a well-characterized interferon-stimulated gene (Kühn et al., 1995). At steady state, 6.8± 2.5% of Tregs within thymus, and 4.0±0.4% of Tregs in spleen and mesenteric lymph nodes expressed GFP (Fig. S3G). The frequency of ISG-Tregs identified in the Mx1-Cre x Foxp3-iDTR mice was similar to that which we observed when reanalyzing a previous scRNA-seq dataset for thymic Tregs and Treg progenitors (Owen et al., 2019), in which ISG-Tregs comprised ~8% of thymic Tregs and Treg progenitors. Subsequent scRNA-seq studies found a similar population of ISG-Tregs in the spleen (~5%) (data not shown). This population of Tregs is similar to one identified in the spleen, lung, gut, and skin at steady state and following treatment with IL-2 mutein (Lu et al., 2020; Miragaia et al., 2019). Thus, the percentage of Tregs labeled in Mx1-Cre x Foxp3-iDTR mice closely mimics that detected in the spleen and thymus by scRNA-seq.

To determine the origin and fate of ISG-Tregs during influenza infection, we utilized the Mx1-Cre x Foxp3-iDTR mouse model. Single-cell RNA-seq data demonstrated that ISG-Tregs peak in the lung 7 days post-infection and then contract upon clearance of the infection (Fig. 2C).
However, it is unclear whether ISG-Tregs give rise to the subsequent TR-Treg lineage or not. In Mx1-Cre x Foxp3-iDTR mice, ISG-Tregs are permanently labeled with GFP and DTR. In naïve mice, GFP⁺ ISG-Tregs represent ~4.0% of Tregs within the lung tissue; this expanded to ~22.1% GFP⁺ at 7 days post infection (Fig. 3A, 3B). While the frequency of GFP⁺ Tregs remained high at 10 days post infection, this decreased precipitously by 21 days post infection (Fig. 3B). A similar pattern of expansion and contraction of GFP⁺ ISG-Tregs occurred in the mediastinal LN (Fig. 3B). Thus, the second wave of Tregs in the lung are not derived from ISG-Tregs.

Since a subset of ISG-Tregs exists in the spleen and lymph nodes of naïve mice, we examined whether ISG-Tregs in the lung following influenza infection arise exclusively from this pre-existing subset, or whether they can be induced to adopt this phenotype. Two days prior to infection we depleted ISG-Tregs with DT. Depletion in the lungs and spleen was quite effective (Fig. 3C). ISG-Tregs in the spleen were depleted for the entire time course of infection (Fig. 3C). In contrast, although ISG-Tregs were effectively depleted in the lungs at day 0, they had fully recovered by day 3 post-infection and remained at equivalent levels to that seen in undepleted mice out to at least day 7 post-infection (Fig. 3C). It is unclear whether this reflects rapid expansion of a small set of pre-existing ISG-Tregs that were not effectively depleted or whether non ISG-Tregs are recruited to the lung upon deletion of the pre-existing ISG-Tregs, and these Tregs then acquire an ISG-signature. As ISG-Tregs contract in the lungs following clearance of influenza, we assessed whether they could be recalled during a secondary infection, and if this too was independent on the pre-existing population or merely reflected IFN abundance in the infected tissue. To examine this issue Mx1-Cre x Foxp3-iDTR mice were given a primary infection with influenza X31. The ISG-Tregs were then depleted at 28 days post infection via DT administration. One day after ISG-Treg depletion the mice were rechallenged.
with influenza PR8 and then analyzed 7 days following the secondary infection. The frequency of ISG-Tregs during the secondary infection was significantly diminished in mice depleted of ISG-Tregs prior to infection compared to control mice (Fig. 3D). This finding demonstrates that ISG-Tregs seen in a secondary infection are dependent on the pre-existing ISG-Treg subset and do not simply arise because they migrate to a tissue with significant IFN production.

To understand the function of Tregs during influenza A infection, we analyzed whole lung transcriptomes of Treg-depleted mice. Mx1-Cre x Foxp3-iDTR and Foxp3-DTR mice were infected with influenza A virus and depleted of ISG- or total Tregs, respectively, at days 6, 7, and 8 post infection. RNA-seq of whole lung demonstrated that depletion of ISG-Tregs in Mx1-Cre x Foxp3-iDTR mice had minimal effect on gene transcription (data not shown). This may reflect our inability to effectively deplete this population in the lungs. However, analysis of Treg-depleted Foxp3-DTR mice revealed an increase in transcripts for immunoglobulin genes (Fig. 4A). In particular, transcripts for Iga and Igj (J-chain) were highly upregulated in the absence of Tregs in depleted Foxp3-DTR mice (Fig. 4A). To validate this result, we evaluated the lungs of Treg-depleted Foxp3-DTR mice by flow cytometry and found an increased frequency of IgA-expressing B-cells (Fig. 4B). Previous studies have supported a protective role for secreted IgA during influenza infection (van Riet et al., 2012). Our results suggest that Tregs present during peak anti-viral immune activity against influenza A virus serve to limit IgA production by B-cells within the lungs.

The role of IFNs in the severity of COVID-19 is controversial with some studies supporting a protective role and others implicating them in progression to severe disease (Lee and Shin, 2020). To examine this in more detail and to determine whether ISG-Tregs exist in humans with
respiratory viral infections and correlate with COVID-19 severity, we reanalyzed a scRNA-seq meta-analysis of PBMCs obtained from 111 healthy donors and patients with mild or severe COVID-19 (Mukund et al., 2021). We extracted data for CD4⁺ and CD8⁺ T cells and identified cluster 9 as Tregs based on high expression of FOXP3 and IL2RA transcripts but low expression for IL7RA (Fig. S4A). Upon reclustering the Treg subset, we identified 6 distinct clusters; cluster 4 expressed an IFN gene signature similar to that observed in mice (Fig. 5A, Table S2). Additional clusters included ones characterized by the inflammatory cytokine IL32 (cluster 0), the gene S100A9 (cluster 1), an activated subset expressing immediate early genes such as FOS and JUN (cluster 2), a subset expressing of TCF7 and PLAC8 (cluster 3), and a subset expressing GZMA, GZMK, CCL5 and NKG7 (cluster 5). We separated Tregs based on disease status of the patients from whom they were collected; we found that healthy donors had a very small population of ISG-Tregs. In contrast, patients with both mild and severe COVID-19 exhibited a significant increase in ISG-Tregs (Fig. 5B). Notably, ISG-Tregs were >50% more abundant in patients with mild COVID-19 as opposed to those with severe disease (1.6-fold↑) (Fig. 5B, 5C). There was no significant difference in the time after disease onset when samples were collected (Fig. S4B). The opposite relationship was observed for the IL32-expressing Treg cluster, as patients with severe COVID-19 showed a significant increase in IL32-expressing Tregs relative to healthy donors (2.9-fold↑) or patients with mild disease (1.9-fold↑) (Fig. 5B, 5C). Thus, ISG-Tregs correlate with better outcomes while IL32-Tregs correlate with worse outcomes in COVID-19 patients.

Discussion

In this study, we profile Treg heterogeneity within the lung tissue during infection with influenza A virus and introduce a new mouse model to track and deplete Treg subsets. The Foxp3-iDTR mouse model allows one to simultaneously track and delete distinct Treg subset. Previous systems to study Treg subsets have relied on deletion of specific floxed target genes within
Tregs using Foxp3-Cre mice or deletion of Foxp3 using Foxp3-floxed mice. While useful, these models are limited as they do not actually delete the Treg subset of interest but rather convert them into either a different type of Treg or an effector T-cell, respectively. Thus Foxp3-iDTR mice should be an invaluable tool for studying the roles of particular Treg subsets in a variety of tissues and disease settings. Using scRNA-Seq and our Foxp3-iDTR lineage tracking mice we demonstrate that in response to influenza A infection at least two transcriptionally distinct waves of Tregs populate the lungs: an initial ISG-Treg population that emerges at day 7 post-infection and contracts during clearance of influenza A, and a second wave of TR-Tregs that emerge during and following virus clearance. Importantly, the ISG-Treg and TR-Treg subsets have distinct TCR repertoires, suggesting that they recognize different antigens. The TR-Treg subset can be further broken down into a number of distinct transcriptional subsets, although these subsets have an overlapping TCR repertoire, suggesting that they recognize related antigens. Thus, the Treg response to influenza exhibits substantial diversity in kinetics of response, TCR repertoire, and transcriptome.

A number of key questions remain. First, it is unclear how the distinct Treg subsets observed during influenza A infection arise. A small subset of thymic (~7%) and splenic Tregs (~5%) expressing the ISG-signature can be found in the steady-state. These cells may be recruited to the lungs following viral infection. However, depletion of the vast majority of these pre-existing ISG-Tregs had no impact on the influx of ISG-Tregs to the lungs during a primary infection. It is unclear whether the ISG-Treg population in the lungs that arise after depletion of the pre-existing ISG-Treg repertoire comes from extremely rapid proliferation of the few ISG-Tregs that escape deletion, or possibly ISG-Tregs that fail to express sufficient Mx1 to activate our reporter and thus are not deleted in our model. Alternatively, these could arise from Tregs that never encountered IFN previously, and hence did not express an ISG-signature initially, but acquired
an ISG-signature during infection. Interestingly, a similar effect was not observed during a secondary infection with influenza, in which prior depletion of ISG-Tregs did result in a significant reduction in this subset of Tregs in the lung. One potential explanation is that prior to influenza infection only a fraction of the potential ISG-Treg pool acquires this gene signature during thymic development or at steady-state, but that after a primary infection all Tregs with this potential are labeled, allowing for their efficient depletion prior to a secondary infection.

A second question is what regulates the distinct temporal migration of ISG-Tregs and TR-Tregs. It is clear that these Treg subsets arrive in the lung with different kinetics but how this happens is not obvious. Likewise, where these Treg subsets localize within the lungs, and whether ISG-Tregs and TR-Tregs are found in different locations in the lungs, is also not clear.

Finally, the function of these distinct subsets of Tregs is unclear. For example, it remains unclear why distinct Treg subsets exist that uniquely express genes encoding endogenous opioids (Penk⁺), trefoil factor proteins (Tff1), or Plac8. A recent study demonstrated that activation of sensory neurons in the skin stimulated a Th17 response, indicative of crosstalk between the nervous and the immune systems (Cohen et al., 2019). PENK produced by Tregs may be acting on opioid receptors expressed by neurons in the lung to modulate pain perception. TFF1 is associated with mucus production in the gut (Järvå et al., 2020), although the function of TFF1-producing Tregs during IAV infection is unclear. The exact function of the ISG-Treg lineage also remains unclear although the observation that they correlate with improved response to COVID-19 suggests they play an important role. One possibility is that ISG-Tregs may be enriched for TCRs that recognize peptides derived from IFN stimulated genes, which are abundant during viral infections (Spencer et al., 2015). In this scenario, ISG-
Tregs are drawn to an IFN-rich niche within the tissue and help to limit IFN-induced inflammation. In support of this possibility, recent studies have shown that limiting type I and type III IFN production following viral infections of the lung is required for tissue regeneration (Broggi et al., 2020; Major et al., 2020). Likewise, a recent study showed that depletion of Tregs during imiquimod-induced psoriasis led to uncontrolled IFNα and subsequent transcription of IFN-induced genes (Stockenhuber et al., 2018). Finally, a separate report demonstrated that house-dust mite (HDM)-reactive CD4+ T-cells and Tregs that express an IFN-signature are more frequent in asthmatic individuals without HDM allergy than those with HDM allergy (Seumois et al., 2020). Thus, ISG-Tregs are expanded in multiple inflammatory conditions. Identifying the antigens recognized by TR-Tregs and ISG-Tregs would likely provide key insights into their functional roles. Finally, in our analysis of Tregs in human patients with COVID-19, we identified a Treg subset expressing the cytokine IL32. As IL32 is not expressed in rodents (Kim, 2014) it was not identified in our murine scRNA-Seq studies. An important question is what causes the reciprocal relation between ISG-Tregs and IL32-Tregs in patients with mild versus severe COVID-19. The answer to that question could provide important insights into mechanisms leading to more severe disease in COVID-19 patients.

Whether the distinct Treg subsets in the lung are functionally redundant or play non-overlapping roles during recovery from influenza A virus infection is also unclear. Using Foxp3-DTR mice we established that Tregs play an important role in limiting the frequency of IgA+ B cells in the lungs. Previous studies have shown that secreted IgA protects against influenza virus infection (van Riet et al., 2012). It is unclear why Tregs would limit this process, but they may play a role in ensuring preferential selection for high-affinity anti-viral B cells. Alternatively, limiting IgA+ B cells in the lung may be necessary to allow for proper repair of lung tissue post-
infection. In conclusion, depletion of Tregs at the peak of the anti-influenza response does play an important role in subsequent immune function.
Acknowledgements

We thank G. Hubbard, A. Rost, N. Keller, M. Dinh, T. Maiers, and J. Fiege for technical assistance, T. Martin, J. Motl and P. Champoux for cell sorting and maintenance of the Flow Cytometry Core Facility (UFCR), B. Ruis with the Genome Engineering Shared Resource (GESR) and Y. You with the Mouse Genetics Lab (MGL) for assistance developing the Foxp3-iDTR mice, the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing computational resources. and Drs. Mukund and Subramaniam for providing the quality-filtered non-integrated human meta-analysis dataset. LES was supported by an immunology training grant (T32 AI997313) and an individual predoctoral F31 fellowship from the NIH (1F31AI154706-01A1). MAF was supported by NIH R01 AI124512 and R01 AI159554. This work was supported in part by NIH P30 CA77598 utilizing the following Masonic Cancer Center, University of Minnesota shared resource(s): UFCR, GESR and MGL.


Declaration of interests. The authors declare no competing interests.

Data Availability Statement
All scRNA-Seq data is available at GEO under the following accession numbers: GSE123067, GSE195488, and GSE196909. Flow cytometry data is available upon request.
Figure legends.

**Figure 1. Tregs expand in the lungs during influenza A virus infection.** (A) Number of Tregs within the lungs and divided between the IV+ and IV- fractions over the course of influenza A infection. (B) Frequency of Tregs within the IV+ and IV- fractions of the lungs over the course of influenza A infection. (C) Frequency of Tregs within the IV- fraction of the lungs expressing TBET, RORyt, and GATA3. N=4-8 per time point. Horizontal bars represent means and vertical bars represent standard deviations.

**Figure 2. Treg heterogeneity during influenza A virus infection.** (A) Experimental scheme for single-cell RNA-seq of lung Tregs from influenza A virus infected mice. Infections were staggered and IV-labeled Tregs were sorted from the lung and hashtagged prior to sequencing. (B) UMAP of Tregs clustered at resolution 0.5 and separated by IV-labeling status. (C) UMAP of Tregs separated by days post infection and IV-labeling status. (D) Slingshot trajectory analysis. (E) Morisita-horn similarity index of Treg TCRs among ISG-Tregs (cluster 3) and TR-Tregs (clusters 2, 5, 6, 8, 10, and 12) from individual mice.

**Figure 3. Tracking ISG-Tregs using Mx1-Cre x Foxp3-iDTR mice.** (A) ISG-Tregs in the lungs of Mx1-Cre x Foxp3-iDTR mice separated by time point post infection and IV-status. (B) Frequency and numbers of ISG-Tregs in the lungs and mediastinal LN of Mx1-Cre x Foxp3-iDTR mice across time course. For the lungs, n=5 for uninfected and n=6 for all other time points. For mediastinal LN n=3 for uninfected and n=6 for all other time points. Horizontal bars represent means and vertical bars represent standard deviations, samples are from 2-independent experiments. (C) Frequency of ISG-Tregs in the lungs and spleen following pre-infection depletion of ISG-Tregs. Data is representative of 2 experiments, n=3-4 mice per time point. Horizontal bars represent the means and vertical bars represent the standard deviations. **P<0.01, ****P<0.0001, determined by unpaired t-test with Holm-Sidak multiple comparisons correction. (D) Frequency of GFP+ Tregs within the lung 7 days following secondary influenza infection. Mice were infected with 1000pfu X31 and rechallenged with 15pfu of PR8 after 28 days. Open and closed symbols represent samples from 2 separate experiments. For Mx1-Cre x Foxp3-iDTR +DT n=10, and for Mx1-Cre x Foxp3-iDTR no DT n=7. Horizontal bars represent means and vertical bars represent standard deviations. P-value determined by Mann-Whitney test.

**Figure 4. Increased IgA+ cells in the lungs of influenza A infected mice following depletion of Tregs.** (A) Volcano plot showing changes in gene expression between Treg depleted Foxp3-DTR mice and control mice. (B) Frequency of IgA+ B-cells following depletion of Tregs in Foxp3-DTR mice. Gated on live>CD45+>Thy1.2->B220+IgG H+L->GL7- cells. N=8 for each treatment group. Open and closed symbols represent samples from 2 separate experiments. Horizontal bars represents means and vertical bars represents standard deviations. ****p<0.0001, determined by student’s t-test.
Figure 5. ISG-Tregs in COVID-19. (A) UMAP of Tregs from the blood of healthy donors and COVID-19 patients (resolution 0.3). Heatmap of top differentially-expressed genes in each cluster. (B) UMAP of Tregs separated by disease status. (C) Proportion of Tregs within each cluster separated by disease status of source patients. Error bars represent 95% confidence intervals determined using a bootstrapping approach.
References


Materials and Methods

**Mice.** *Tclib x TCRα+/- x Foxp3-GFP* (Jax stock #006772) mice were bred in house as previously described (Hsieh et al., 2004). *Tclib x TCRα+/- x Foxp3-RFP* (Jax stock #008374) mice were bred in house. *Foxp3-iDTR* mice were created as described below and bred to *Mx1-Cre* (B6.Cg-Tg(Mx1-cre)1Cgn/J, Jax stock # 003556), *CD4-Cre* (B6.Cg-Tg(Cd4-cre)1Cwi/BfluJ, Jax stock # 022071) and *CD4-CreERT2* (B6(129X1)-Tg(Cd4-cre/ERT2)11Gnri/J, Jax stock #022356) mice.

**Foxp3-iDTR mice.** A double-stranded DNA template was designed to contain a floxed stop cassette flanked by 800 base pair sequences homologous to the *Foxp3-DTR* locus. The DNA template, guide RNA, and CAS9 protein were micro-injected into *Foxp3-DTR* zygotes and zygotes were implanted into pseudopregnant female mice that had received prior injection of pregnant mare serum gonadotropin and human chorionic gonadotropin. Resulting pups were screened for the insertion using PCR primers 5’-CAAGTGCTCCAATCCCTGC and 5’-CTTGCATTCCTTTGGCGAGA. Whole genome sequencing of tail DNA confirmed that the *Loxp-STOP-Loxp* cassette was integrated correctly and that there were no detectable off-target integrations or CAS9-induced mutations.

**Infections.** Mice were anesthetized with a weight-determined dose of ketamine/xylazine and infected intranasally with 15 plaque-forming units (pfu) of influenza A virus (strain A/Puerto Rico/8/1934 H1N1). For depletion experiments, mice were administered 100ng of intraperitoneal DT on days 6, 7, and 8 post infection and 10ng of intranasal DT on day 7 post infection. Experiments involving mice were performed as dictated by the University of Minnesota Institutional Animal Care and Use Committee.

**IV antibody staining.** IV antibody staining was performed as previously described (Anderson et al., 2014). Briefly, mice received an intravenous injection of 3ug of anti-CD4 (RM4-4, Biolegend 116022) antibody and were euthanized 3 minutes after injection.

**Isolation of lymphocytes.** Mice were euthanized and the lymphoid tissues harvested into FACS buffer (1X PBS/2% FBS/0.05% sodium azide/2 mM EDTA, pH 7.4). Tissues were mashed between glass slides and filtered through 70um filter mesh to prepare single-cell suspensions. RBCs were lysed with ACK lysis buffer (150mM NH4Cl/10mM KHCO3/0.1mM Na2EDTA) for 5 minutes at room temperature.

**Isolation of lung lymphocytes.** Lung lymphocytes were isolated as previously described (Fiege et al., 2019).

**Flow cytometry.** Single-cell suspensions were washed with FACS buffer (1X PBS/2% FBS/2mM EDTA/0.05% sodium azide) and stained with ghost dye red 780 viability dye (Tonbo
Biosciences 13-0865-T100) and antibodies against TCRβ (H57-597, BD Biosciences 742485), CD4 (RM4-5, Biolegend 100559), CD8 (53-6.7, Tonbo Biosciences 20-0081-U100), CD25 (PC61.5, Tonbo Biosciences 65-0251-U100), CD73 (TY/11.8, ThermoFisher Scientific 25-0731-82), B220 (RA3-6B2, BD Biosciences 563793), CD45 (30-F11, Biolegend 103155), CD90.2 (53.21 BD Biosciences 564365), GL7 (GL7, Biolegend 144604), and IgA (11-44-2, ThermoFisher Scientific 12-5994-81). Following surface staining, cells were fixed with the FOXP3 fixation and permeabilization kit and stained with antibodies against FOXP3 (FJK-16s ThermoFisher Scientific 17-5773-82, 48-5773-82), GFP (Rockland, 600-102-215), and IgG H+L (polyclonal, eBioscience 17-4012-82).

**Lung harvest, cell sorting, and cell hashtagging for single-cell RNA-seq.** Mice received an intravenous injection of 3ug anti-CD4-biotin (RM4-4, Biolegend 116010) antibody 3 minutes prior to euthanasia. Lymphocytes from lungs were isolated into sort buffer (1XPBS/2% FBS/2mM EDTA) and processed as described above. Single-cell suspensions were stained with 2.5ug of anti-CD4-APC (RM4-5, Biolegend 100516) antibody in 1mL of sort buffer and incubated at 4°C for 30 minutes. Following wash, cells were resuspended in 50ul of anti-APC microbeads (Miltenyi) and 350ul of sort buffer. Cells were then washed and run over an LS column (Miltenyi). The bound fraction was collected and stained with ghost dye red 780 viability dye (Tonbo Biosciences 13-0865-T100), antibodies against TCRβ (H57-597, Biolegend 109229), CD8a (53-6.7, ThermoFisher Scientific 47-0081-82), Gr1 (RB6-8C5, ThermoFisher Scientific 47-5931-82), NK1.1 (PK136, ThermoFisher Scientific 47-5941-82), CD11b (M1/70, ThermoFisher Scientific 47-0112-82), CD11c (N418, ThermoFisher Scientific 47-0114-82), F4/80 (BM8, ThermoFisher Scientific 47-4801-82), and Ter119 (ThermoFisher Scientific 47-5921-82), 0.25ug total-seq C streptavidin-PE (C0951, Biolegend 405261), and 1ul total-seq C hashtag antibody (C0301-C0310, Biolegend) and incubated for 30 minutes at 4°C. Following wash, all samples were combined, resuspended in sort buffer, and run on the cell sorter.

**Single-cell RNA-sequencing of splenic Tregs.** Spleens from Tc11 x TCRα”” x Foxp3-RFP mice were dissociated as previously described (Owen et al., 2019) and stained with biotin-linked antibodies against CD8a and B220 as well as Ter119. After staining with streptavidin magnetic beads, cells were run over LS columns. The unbound fraction was used to sort CD4⁺CD8⁻ TCRVα2⁺FOXP3⁺ splenocytes into catch buffer. Antibodies against CD4 (GK1.5, BD Biosciences 563331), CD8a (53-6.7, ThermoFisher Scientific 47-0081-82), and TCRVα2 (B20.1, ThermoFisher Scientific 17-5812-82) were used. Cells were then centrifuged and resuspended in 1X PBS/10% FBS for single-cell capture. The 10X Genomics Chromium Next GEM Single Cell 5’ Kit (v2) was used to measure (1) mRNA transcript expression, and (2) VDJ TCR repertoire. The library was sequenced on the NovaSeq 6000 with pair-end reads (26 bp for read 1 and 91 bp for read 2) (Illumina). Data available at accession number GSE195488.

**Single-cell RNA-sequencing of lung Tregs.** To isolate Tregs, we sorted Dump⁻ TCRb⁺CD4⁺FOXP3⁺-GFP⁺ cells into catch buffer (1XPBS/50% FBS). Cells were then centrifuged and resuspended in 1X PBS/10% FBS for single-cell capture. The 10X Genomics Chromium Next GEM Single Cell 5’ Kit (v2) was used to measure (1) mRNA transcript expression, (2) hashtag oligos (HTO), (3) CITE-seq oligos (ADT) for expression of cell-surface proteins and (4)
VDJ TCR repertoire. To enhance TCRα chain CDR3 detection, given the fixed TCRβ allele, we only amplified the TCRα sequence using the following primer sets:

**Mouse T cell Mix 1:**

Forward Primer (2µM final concentration): 5’- AATGATACGGCCACCCGATCTACACTTTCTTCCCTACACGACGCTC-3’

Reverse Outer Primer (0.5µM final concentration): 5’-CTGGTTTCCTCCAGGCAATGG-3’

**Mouse T cell Mix 2:**

Forward Primer (0.5µM final concentration): 5’-AATGATACGGCCACCCGATCTC-3’

Reverse Inner Primer (0.5µM final concentration): 5’-AGTCAAAGTCGGTGAACAGGCA-3’

Following library preparation, quality was assessed using a bioanalyzer and preliminary RNA-seq to determine approximate cell number and quality was performed on a MiSeq. After quality control, the libraries were sequenced on a NovaSeq 6000 with 2 x 150bp paired end reads. Raw and processed data are available through the GEO accession (GSE195488).

**Bioinformatics analysis of thymus and spleen single-cell RNA-sequencing data.** The 10x Genomic Cell Ranger pipeline (version 3.0.0) was used to map reads to the mm10 reference (provided by 10X Genomics, version 3.0.0) and generate counts for each cell. Raw and processed data have been deposited at Gene Expression Omnibus and are available via GEO accession (see accession numbers above). Thymic and splenic T regulatory single cell expression data was analyzed from data initially collected in (Owen et al., 2019) (series GSE123067, dataset GSM3494565). Raw count data for both splenic and thymic single cell datasets were individually analyzed in R (version 3.5.0) using the Seurat package (version 3.1.1) and ggplot2 package (version 3.0.0). Each dataset was filtered for cells that expressed greater than 200 genes but less than 3200 genes. The remaining cells counts were normalized by a centered-log ratio method and principal component analysis (PCA) was performed on the 2000 most variable genes. Cells were clustered using the top 13 PC vectors for the splenic T cells and top 11 PC vectors for the thymic T cells using the FindNeighbors and FindClusters functions. Two-dimensional representations were generated using the RunTSNE and RunUMAP functions. Differential expression (wilcox) of clusters at different resolutions was performed to define biological significance. The interferon signature was determined to be best defined in the splenic T cell and thymic T cell cells at a resolution of 0.4 (cluster 3) and 0.35 (cluster 3) respectively.

**Bioinformatics analysis of lung Treg single-cell RNA-sequencing data.** Raw sequencing data was processed using the 10X Genomics cellranger software (ver. 6.0.0) mkfastq function to demultiplex the Illumina libraries into gene expression (GEX), CITEseq expression (HTO and/or ADT), and T-cell receptor (TCR) datasets. The cellranger count function was used to align reads to the mouse reference genomes (refdata-gex-mm10-2020-A, refdata-cellranger-vcj-GRCm38-alts-ensembl-5.0.0; provided by 10X Genomics). See Tables 1 and 2 below.
Raw count tables were loaded into R (ver. 4.0.3) and analyzed with the Seurat (ver. 4.0.1) or tidyverse (1.3.1) R packages. The GEX dataset was filtered to include only gelbeads in emulsion (GEMs, which are oil droplets containing uniquely barcoded beads that ideally contain one individual cell) expressing more than 300 genes (counts > 0) and genes expressed in more than 3 GEMs (counts > 0). The proportion of mitochondrial RNA in each GEM was calculated and GEMs with extreme levels (top 0.5% of all GEMs) were removed from the analysis. For the remaining GEMs, the raw HTO count table was supplied to GMM-demux software (ver. 0.2.1.3) (Xin et al., 2020) to classify which HTO tags were detected in each GEM. GEMs containing multiple HTOs (i.e. doublets or multiplets) were removed from downstream analysis. Initial analysis revealed no differences in HTO clustering between technical replicates and they were combined for all downstream analyses.

Raw single cell surface protein expression data (ADT counts) were normalized according to the centered-log-ratio method using Seurat. Raw GEX counts were normalized and transformed using the Seurat SCTransform function (Hafemeister and Satija, 2019) including the percentage of mitochondria expression as a regression factor. Upon further analysis, we determined that four T-cell receptor genes (Trbv1, Trbv16, Trbv31 and Trav4-4-dv10) were driving the clustering of small unique clusters. The normalized gene expression levels for these genes were also used as regression factors in the final dataset. After regression the cells in these 4 TCR driven clusters clusters redistributed relatively uniformly throughout the UMAP. Each cell was classified according to its expression of canonical cell cycle genes using the Seurat CellCycleScoring function (S-phase and G2/M-phase gene sets provided by Seurat were originally developed by (Tirosh et al., 2016). Principal components analysis (PCA) was performed using the normalized, mean-centered, and scaled SCT dataset (RunPCA function). Two-dimensional projections were generated using the top 30 PCA vectors as input to RunUMAP function in Seurat. Cells were clustered using the FindNeighbors (top 30 PCA vectors) and FindClusters functions (testing a range of possible resolutions: 0.3, 0.4, 0.5, 0.6, 0.7). A final clustering resolution of 0.5 was selected.

Pairwise differential expression (DE) testing (Wilcoxon rank-sum) using the Seurat FindMarkers function was performed between all clusters. DE genes were significant based on log2-fold-change (≥ 0.25) and BH adjusted p-value (≤ 0.01). Figures were generated using the ggplot2 R package (ver. 2.3.3.3) (Wickham, 2016). The distribution of normalized GEX and ADT expression levels were displayed for cells/clusters using color heatmaps on UMAPs, dot plots, and tile plots.

A pseudotime cell trajectory analysis was completed using the Slingshot R package (ver. 1.8.0) (Street et al., 2018). The Seurat-based R object was converted for use with Slingshot using the Seurat function as.SingleCellExperiment. A trajectory was inferred with the slingshot function using cluster labels, UMAP coordinates, and the central Tregs (cluster 1) as the root of the trajectory. Cell groups were compared (e.g. uninfected vs. infected) along each of the slingshot lineages using a differential topology approach. To determine statistical significance, a permutation test was performed. For each cell group, the weighted mean of pseudotime values was calculated using the slingshot curve weights. An initial test-statistic was calculated as the difference in weighted means between the two cell groups. A null distribution was generated by
randomization of cell group labels and resampling 10,000 times. P-values were calculated as the number of test-statistics generated under the null distribution that were as or more extreme than the initial test-statistic (i.e., using real cell group labels).

Raw and processed data have been deposited at the Gene Expression Omnibus (GEO) database (accession number: GSE196909).

Table 1: Hashtag (HTO) antibodies

<table>
<thead>
<tr>
<th>Dataset</th>
<th>BioLegend Product Number</th>
<th>Sample (uninfected or infected, days post infection, replicate letter)</th>
<th>Barcoded Read</th>
<th>Barcode pattern</th>
<th>Barcode Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse lung</td>
<td>C0301</td>
<td>u0a</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>ACCCACCAGTAAGAC</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0302</td>
<td>u0b</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>GGTGAGAGCATTCA</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0303</td>
<td>i3a</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>CTTGCCGCATGTCAT</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0304</td>
<td>i3b</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>AAAGCATTCTTCACG</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0305</td>
<td>i7a</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>CTTTGCTTTGTGAG</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0306</td>
<td>i7b</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>TATGCTGCACCGTA</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0307</td>
<td>i10a</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>GAGTCTGCCAGTATC</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0308</td>
<td>i10b</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>TATAGAACGCCAGGC</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0309</td>
<td>i21a</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>TGCCTATGAAACAAG</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0310</td>
<td>i21b</td>
<td>R2</td>
<td>5PNNNNNNNNNNN(BC)</td>
<td>CCGATTGTAACAGAC</td>
</tr>
<tr>
<td>Dataset</td>
<td>BioLegend Product Number</td>
<td>Protein Target</td>
<td>Barcoded Read</td>
<td>Barcode pattern</td>
<td>Barcode Sequence</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------</td>
<td>----------------</td>
<td>---------------</td>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0195</td>
<td>TNFRSF4</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>CTCACCTACCTATGG</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0097</td>
<td>IL2RA</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>ACCATGAGACACAGT</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0004</td>
<td>PDCD1</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>GAAAGTCAAAGCCT</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0003</td>
<td>HAVCR2</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>ATGGGCACTCAGATG</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0378</td>
<td>LAG3</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>ATTCGTCCTCAAGG</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0848</td>
<td>TIGIT</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>GAAAGTCGCCCAACAG</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0228</td>
<td>CXCR3</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>GTTCAGCGCGTGTTA</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0833</td>
<td>CCR4</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>TTTCATGTGGTGTGC</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0846</td>
<td>CXCR5</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>ACGTAGTCACCTAGT</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0077</td>
<td>CD73</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>ACACTTAACGTCTGG</td>
</tr>
<tr>
<td>Mouse lung</td>
<td>C0837</td>
<td>IL1RL1</td>
<td>R2</td>
<td>5PNNNNNNNNNN(BC)</td>
<td>GCATGGAGCATGTT</td>
</tr>
</tbody>
</table>

Table 2: CITEseq (ADT) antibodies
Whole Lung RNA-seq. Whole lung was added to 3mL of buffer RLT with β-mercaptoethanol and reagent DX and digested in M-tubes on the GentleMacs RNA setting. After a quick spin to remove debris, lung homogenate was aliquoted and 300ul was processed through the Qiagen RNeasy Mini Kit to extract total RNA. The RNA-seq library was prepared using the TruSeq Stranded kit. Samples were sequenced on the NovaSeq 6000 using 150 base-pair paired-end reads for a total of 20 million reads per sample.

Bulk RNAseq libraries were run through the CHURP pipeline as previously described (https://dl.acm.org/doi/10.1145/3332186.3333156). Briefly, this wrapper quality trims reads with Trimmmomatic (Bolger et al., 2014), aligns to the human reference genome (GRCh38) using HISAT2 (Kim et al., 2019), and converts alignments to feature counts using SAMTools (Li et al., 2009) and featureCounts() from the Rsubread package (Liao et al., 2019). Expression of genes present in the Ensembl GRCh38 v100 annotation set and longer than 300 bases were further quantified using edgeR (Robinson et al., 2010). Features present in two or more samples at more than 1 count per million were selected to calculate normalization factors with method "RLE" using calcNormFactors(). Dispersion was then estimated using estimateDisp() and differentially expressed genes were identified using glmQLFTest() with default parameters. Volcano plots were generated using ggplot2 (Wickham, 2016).

GSEA analyses were performed on genes lists consisting of genes with at least a 2-fold expression difference and a false-discovery corrected p-value less than 0.1. Genes were rank ordered based direction of difference followed by the FDR-corrected p-value of the comparison. These gene lists were analyzed using the R package clusterProfiler (Wu et al., 2021) and compared against human HALLMARK and GO term gene lists using an q-value cut off of 0.1. Dotplots were generated through ClusterProfiler using the dotplot() function from the R package enrichplot (Wu et al., 2021).
Bioinformatics analysis of human COVID-19 dataset. The dataset described in section 4.3 Mukund et al. (2021) was made available to us by the paper's authors. The dataset was already filtered for the criteria described in section 4.3 but not integrated. Using available metadata, the dataset was separated into the 5 original datasets and processed and integrated using Seurat's (v. 4.0.5) IntegrateData function. The same parameters were used as described in section 4.3 of Mukund et al. 2021. Following integration, cells annotated by Mukund et al. (Mukund et al., 2021) as CD4 or CD8 T-cells were subsampled from the original dataset and processed as described in section 4.4 of Mukund et al., 2021, and re-clustered at a range of resolutions (0.8 – 1.2). A resolution value of 0.8 was used and cluster 9 (CD4⁺FOXP3⁺CD25⁺IL7R<sub>lo</sub> T cells) was subsampled for further analysis. The CD4⁺FOXP3⁺CD25⁺IL7R<sub>lo</sub> cell subset was processed using the same parameters described in section 4.4 of Mukund et al., 2021. Clustering was performed at a range of resolutions (0.1 – 1) and the clusters identified at a resolution value of 0.3 were used in further analyses. Confidence bars were calculated for the resolution 0.3 clusters using single cell differential composition (scDC) analysis (v. 0.1.0) (Cao et al., 2019). The scDC_noClustering function with the bias corrected and accelerated (BCa) confidence interval calculation method was used. All work was done in R (v. 4.1.2).

Statistics

In the graphs horizontal bars represents means and vertical bars represent standard deviations. In the text, +/- values represent standard deviations. Normality was assessed using Shapiro-Wilks test. For comparisons of two groups, an unpaired t-test was used for normal data and Mann-Whitney test was used for non-normal data. For comparisons of two groups across multiple time points, correction for multiple comparisons was performed using the Holm-Sidak method. A two-way ANOVA was used to compare three or more groups containing two factors.
References


Figure 2

A

IAV-PR8 21 days IV-labeled Lung Tregs

10 days

7 days

3 days Naive

B

Blood Tissue

Cluster
0 - Non-lymphoid tissue-like
1 - Central
2 - Areg/Tissue repair
3 - ISG
4 - Non-lymphoid tissue-like
5 - Gzmb
6 - Activated/Suppressive
7 - Proliferating
8 - Plac8
9 - Central ISG
10 - Penk
11 - RNA-processing
12 - Trf1

C

day 0 day 3 day 7 day 10 day 21

Blood

Tissue

D

E

Morisita-Horn Similarity Index

rep1_Repairclusters2_5_6_8_10_12
rep2_ISGcluster3
rep2_Repairclusters2_5_6_8_10_12
rep1_ISGcluster3
rep1_Repairclusters2_5_6_8_10_12

0.098
0.203
0.063
0.248
0.152
0.062

value

0.0 0.098 0.10 0.15 0.2 0.248 0.203 0.203 0.098 0.152 0.063 0.062 0.1

Available under a CC-BY-NC-ND 4.0 International license (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a Creative Commons license (Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International, CC-BY-NC-ND).
Figure 4

A

B
Figure 5

A

B

C

Percent of cells per disease status category (median value from bootstrap analysis, error bars represent 95% CI)
Supplemental Figure 1

(A) IV-labeling of lung Tregs in naïve mice and mice at 10 days post influenza A infection. Cells gated on live>TCRβ+>CD4+>Foxp3+. (B) UMAPs of replicate mice at each time point post infection. (C) Frequency of Tregs from each time point that are within each cluster. (D) Feature plots of defining genes for each cluster.
Supplemental Figure 2. (A) Morisita-horn similarity index of TCRs from each UMAP cluster. (B) Feature plots of transcripts defining IFNsig-Treg cluster and tissue-repair Treg clusters.
Supplemental Figure 3. (A) Model of Foxp3-iDTR mouse. (B) PCR amplification of the floxed-STOP cassette in the Foxp3-iDTR mice. Red arrows represent PCR primers. (C) Whole genome sequencing reads from the Foxp3-DTR and Foxp3-iDTR mice. (D) Expression of GFP in CD4-Cre x Foxp3-iDTR mice. Representative of 6 mice. (E) Expression of GFP in CD4-CreERT2 x Foxp3-iDTR mice pre- and post-tamoxifen treatment. Representative of 7 mice. (F) Expression of GFP pre- and post-DT treatment. Representative of 4 mice. (G) Frequency of ISG-Tregs in the spleen, mesenteric LN, and thymus of Mx1-Cre x Foxp3-iDTR mice at the steady state. The horizontal bars represent the mean and vertical bars represent the standard deviation.
Supplemental Figure 4. (A) UMAP of PBMCs from healthy people and COVID-19 patients. Feature plots of FOXP3, IL2RA, and IL7R. (B) Sampling day of mild and severe patients' PBMCs in each study included in meta-analysis. Horizontal bars represent means and vertical bars represent standard deviations. 2-way ANOVA showed that there was no statistical difference between the sampling days of mild and severe patients (p=0.7304) or the sampling days of the studies included in the meta-analysis (p=0.0660). For mild patients, n=43 and for severe patients, n=68.