

# **Computational characterization of T cells in inflammatory diseases**

**DISSERTATION**

for the degree of *Doctor rerum naturalium* (Dr. rer. nat.)

Institute for Medical Systems Biology

Center for Molecular Neurobiology Hamburg

Faculty of Mathematics, Informatics and Natural Sciences

University of Hamburg

*submitted by*

**Yu Zhao**

from Wuwei, China

November 2021

Supervisor: Prof. Dr. Stefan Bonn  
Institute for Medical Systems Biology, Center for Molecular Neurobiology Hamburg

Co-supervisor: Prof. Dr. Thomas Oertner  
Institute for Synaptic Physiology, Center for Molecular Neurobiology Hamburg

Reviewers:

Prof. Dr. Stefan Bonn  
Institute for Medical Systems Biology, Center for Molecular Neurobiology Hamburg

Prof. Dr. Tobias Lenz  
Institute of Zoology, University of Hamburg

Prof. Dr. Immo Prinz  
Institute for Systems Immunology, Center for Molecular Neurobiology Hamburg

Date of disputation: May 19th, 2022

## Table of Contents

Summary .....	1
Zusammenfassung .....	3
1. Chapter 1: Introduction.....	5
1.1 Immune cells .....	5
1.2 T cell immunity .....	6
1.2.1 T cell development in the thymus .....	6
1.2.2 T cell activation .....	7
1.2.3 CD4 <sup>+</sup> T cell differentiation .....	8
1.2.4 Memory T cell compartmentalization.....	11
1.2.5 Function of tissue resident memory T cells .....	13
1.2.6 Summary of T cell subtypes .....	14
1.3 Inflammatory diseases .....	15
1.3.1 Glomerulonephritis .....	15
1.3.2 NAFLD and NASH .....	17
1.3.3 COVID-19 .....	18
1.4 Single cell RNA sequencing technique and data analysis.....	21
1.4.1 Single cell RNA sequencing technique.....	21
1.4.2 Pre-processing and quality control .....	22
1.4.3 Normalization and integration .....	23
1.4.4 Dimensionality reduction and visualization.....	25
1.4.5 Clustering and cell type annotation .....	26
1.4.6 Downstream analysis .....	28
Chapter 2: Publication in <i>Cell and Tissue Research</i> , June 2021 .....	31
Single-cell biology to decode the immune cellular composition of kidney inflammation .....	31
Chapter 3: Publication in <i>Science Immunology</i> , August 2020.....	43
Pathogen-induced tissue-resident memory T <sub>H</sub> 17 (T <sub>RM</sub> 17) cells amplify autoimmune kidney disease .....	43
Chapter 4: Unpublished Manuscript, revision in <i>JCI Insights</i> .....	85
Multi cytokine producing liver CD4 <sup>+</sup> T cells characterize fibrosis in NASH patients .....	85
Chapter 5: Publication in <i>Science Immunology</i> , February 2021 .....	125
Clonal expansion and activation of tissue-resident memory-like T <sub>H</sub> 17 cells expressing GM-CSF in the lungs of severe COVID-19 patients .....	125
Chapter 6. Discussion.....	157
6.1 Infection and autoimmunity.....	157
6.1.1 Bacterial infection and kidney autoimmunity .....	157
6.1.2 Gut-liver axis and liver inflammation.....	158
6.1.3 SARS-CoV-2 infection and autoimmunity .....	159
6.1.4 Mechanisms of infection in autoimmune diseases.....	161
6.2 Characterization and functions of TRMs.....	164
6.3 Interaction of TRMs with other immune/non-immune cells.....	166
6.4 Tissue differences of TRMs .....	168
6.5 Clonal expansion of T cells in inflammatory diseases.....	169

6.6 Interpretation of expanded TCR sequences.....	170
6.7 Therapeutically target pathogenic TRMs.....	172
6.7.1 Inactivate or deplete pathogenic TRMs .....	173
6.7.2 Target cytokines produced by TRMs .....	174
6.8 Conclusions and outlook .....	177
Reference.....	179
Acknowledgements .....	191
List of publications .....	193
Abbreviations .....	195
Eidesstattliche Versicherung.....	197

## Summary

T cells play critical roles in adaptive immunity. The diversity of T cell receptor (TCR) sequences enables T cells to recognize antigens derived from various pathogens in a specific way. Antigen stimulation can induce the formation of memory T cells, which maintain long-term immunity against previously encountered pathogens. Tissue resident memory T cells (TRMs) reside in peripheral organs without recirculating through the bloodstream. They can rapidly respond to local alarming signals and mediate inflammatory reactions. Especially the recent advent of single cell RNA sequencing (scRNA-seq) and its computational analysis enables the unbiased characterization of complex cell types from different tissues. The combination of scRNA-seq with cell surface protein measurement (CITE-seq) and TCR sequencing (TCR-seq) results in deep insights into T cell identities, clonality, and functions.

While recent research characterized many different T cell types and their potential roles in immunity, a deep understanding of the role of T cell subtypes, especially TRMs, in different organs and inflammatory diseases remains elusive. In this thesis, we characterized T cell subtypes and interpreted their functions in distinct organs and diseases using single cell sequencing and computational analysis.

In the first project, we profiled T cells from blood and kidney biopsies of both controls and patients with ANCA-associated glomerulonephritis using scRNA-seq combined with CITE-seq. We demonstrated that T helper 17 (Th17) TRMs are present in control and diseased human kidneys. Experimental mouse models showed that bacteria-induced Th17 TRMs exacerbates autoimmune kidney disease.

In the second project, we performed scRNA-seq on memory T cells sorted from liver samples from subjects with non-alcoholic fatty liver disease (NAFLD) or non-alcoholic steatohepatitis (NASH). We identified multi-cytokine producing TRMs that are enriched in

NASH and contribute to the progression of fibrosis. We also demonstrated the correlation between *Bacteroides* in the small intestine and hepatic multi-cytokine producing TRMs in NASH patients.

In the third project, we profiled immune cell populations in the bronchoalveolar fluid and peripheral blood specimens obtained from patients with severe COVID-19 or bacterial pneumonia. Using simultaneous scRNA-seq, CITE-seq and TCR-seq, we found that a TRM-like Th17 cell population expressing GM-CSF and IL-17A was expanded in bronchoalveolar fluid from COVID-19 patients. Interactome analysis showed that these Th17 cells interact with macrophages and cytotoxic T cells. We also found that high levels of serum GM-CSF and IL-17A correlated with severe COVID-19.

Overall, this thesis provides a systematic characterization of T cells in different inflammatory diseases covering different organs. The computational analysis addressed important aspects of the pathogenic role of TRMs in renal autoimmunity, chronic liver inflammation and the current COVID-19 pandemic. Our results suggest a potential link between bacterial/viral infection and autoimmune-like inflammatory responses. We also provide insights that might open novel avenues for therapeutic intervention for inflammatory diseases.

## Zusammenfassung

T-Zellen spielen eine entscheidende Rolle bei der adaptiven Immunität. Die Vielfalt der T-Zell-Rezeptor (TCR)-Sequenzen ermöglicht es T-Zellen, Antigene von verschiedenen Krankheitserregern auf spezifische Weise zu erkennen. Die Stimulierung durch ein Antigen kann die Bildung von Gedächtnis-T-Zellen auslösen, die eine langfristige Immunität gegen bereits bekannte Krankheitserreger aufrechterhalten. Gewebsresidente T-Gedächtniszellen („tissue-resident memory T cells“, TRMs) befinden sich in peripheren Organen, ohne dass sie über den Blutkreislauf rezirkulieren. Sie können schnell auf lokale Alarmsignale reagieren und lokale Entzündungsreaktionen auslösen. Vor allem die jüngsten Entwicklungen in der Einzelzell-RNA-Sequenzierung (scRNA-seq) und deren computergestützte Analyse ermöglichen die unvoreingenommene Charakterisierung komplexer Zelltypen aus verschiedenen Geweben. Die Kombination von scRNA-seq mit der Messung von Zelloberflächenproteinen (CITE-seq) und TCR-Sequenzierung (TCR-seq) ermöglicht tiefe Einblicke in die Identität, Klonalität und Funktionen von T-Zellen.

Während in letzter Zeit viele verschiedene T-Zelltypen und ihre potenzielle Rolle in der Immunität charakterisiert wurden, bleibt das Verständnis der Rolle von T-Zell-Subtypen, insbesondere TRMs, in verschiedenen Organen und entzündlichen Erkrankungen unzureichend. In dieser Arbeit wurden T-Zell-Subtypen charakterisiert und ihre Funktionen in verschiedenen Organen und Krankheiten mit Hilfe von Einzelzellsequenzierung und computergestützter Analyse interpretiert.

Im ersten Projekt wurden T-Zellen aus Blut und Nierenbiopsien von Kontrollpersonen und Patienten mit ANCA-assoziiierter Glomerulonephritis mittels scRNA-seq in Kombination mit CITE-seq untersucht. Wir konnten nachweisen, dass T-Helfer 17 (Th17) TRMs in gesunden und entzündeten menschlichen Nierengewebe vorhanden sind. Experimentelle Mausmodelle zeigten, dass bakterieninduzierte Th17-TRMs die Autoimmunerkrankung der Niere verschlimmern.

Im zweiten Projekt führten wir scRNA-seq an Gedächtnis-T-Zellen durch, die aus Leberproben von Patienten mit nichtalkoholischer Fettlebererkrankung (NAFLD) oder nichtalkoholischer Steatohepatitis (NASH) per FACS-Sortierung isoliert wurden. Wir identifizierten TRMs, die mehrere Zytokine produzieren, diese sind in NASH angereichert und tragen zum Fortschreiten der Fibrose bei. Außerdem konnte der Zusammenhang zwischen *Bacteroides* im Dünndarm und hepatischen Multizytokin produzierenden TRMs bei NASH-Patienten nachgewiesen werden.

Im dritten Projekt wurde ein Profil der Immunzellpopulationen in der bronchoalveolären Flüssigkeit und in peripheren Blutproben von Patienten mit schwerer COVID-19 oder bakterieller Lungenentzündung erstellt. Mithilfe von gleichzeitiger scRNA-seq, CITE-seq und TCR-seq konnte dargestellt werden, dass eine TRM-ähnliche Th17-Zellpopulation, die GM-CSF und IL-17A exprimiert, in der bronchoalveolären Flüssigkeit von COVID-19-Patienten vermehrt vorkommt. Die Interaktomanalyse zeigte, dass diese Th17-Zellen mit Makrophagen und zytotoxischen T-Zellen interagieren. Wir konnten auch eine Korrelation von hohen GM-CSF- und IL-17A-Konzentrationen im Serum mit schweren COVID-19-Verläufen aufzeigen.

Insgesamt bietet diese Arbeit eine systematische Charakterisierung von T-Zellen bei verschiedenen entzündlichen Erkrankungen in unterschiedlichen Organen. Die computergestützte Analyse befasste sich mit wichtigen Aspekten der pathogenen Rolle von TRMs bei Autoimmunerkrankungen der Niere, chronischer Leberentzündung und der aktuellen COVID-19-Pandemie. Die Ergebnisse deuten auf einen möglichen Zusammenhang zwischen bakteriellen/viralen Infektionen und immun-vermittelten Entzündungsreaktionen hin. Darüberhinaus weisen die Erkenntnisse neue Wege für therapeutische Interventionen bei Entzündungskrankheiten auf.

# 1. Chapter 1: Introduction

## 1.1 Immune cells

The immune system protects the host from external pathogens and consists of a complex network of different cell types. In general, the vertebrate immune system can be divided into two arms: innate immunity and adaptive immunity.

The innate system serves as a frontline barrier against pathogens and tries to clear the pathogens or the infected cells before severe damage occurs. It consists of physical barriers (such as skin and mucosa), inflammation molecules (such as enzymes and complements), and innate immune cells (Netea *et al.*, 2020). Innate immune cells can distinguish non-self pathogens by molecular structures called pathogen-associated molecular patterns (PAMPs) presented by viruses, bacteria and fungi. Apart from clearing invading pathogens, innate immune cells can also sense signals from damaged host cells by recognition of damage-associated molecular patterns (DAMPs). Typical innate immune cells include macrophages, neutrophils, monocytes, mast cells, dendritic cells (DCs), natural killer cells (NK cells) and innate lymphoid cells (ILCs) (Gasteiger *et al.*, 2017). These cells are not pathogen-specific and can fight against a wide range of invading pathogens within a short time period.

In contrast to the innate immune system, the key features of the adaptive immune system are the specificity and memory formation to different pathogens. The components of the adaptive immune system are B cells and T cells. Both B and T cells originate in the bone marrow but T cells maturation happens in the thymus. B cells are the producer of antibodies and T cells can activate and regulate other immune cells by secreting cytokines. T cells can also directly kill pathogen-infected cells.

B cells and T cells both work in an antigen-specific manner by expressing B cell receptors (BCRs) or T cell receptors (TCRs). During their development, each B cell/T cell generates a unique BCR/TCR by V(D)J recombination, a mechanism of randomly rearranging variable (V), diversity (D) and joining (J) gene segments at the gene loci for BCR/TCRs. The diversity of TCR/BCR determines their specificity to various antigens. Human T cells exhibit a highly diverse TCR and can contain up to 100 million different antigen specificities (Qi *et al.*, 2014). Antigens are typically peptide or protein structures from pathogens or self-body and they can be presented by the antigen presenting cells (APCs) such as DCs to the BCR/TCRs. Antigen presentation is the key step of adaptive immunity and APCs serve as the bridge between innate and adaptive immune system.

B cells and T cells can also differentiate into memory cells after specific antigen stimulation. These memory B/T cells allow the host to respond rapidly when exposed to the same antigen in the future. The combination of specificity and memory make the adaptive immune system extremely powerful.

The major purpose of this thesis is to understand the heterogeneity of T cells and their roles in different inflammatory diseases. In the following introduction sections, the focus will be T cell immunity and functional subsets of T cells.

## **1.2 T cell immunity**

### **1.2.1 T cell development in the thymus**

The thymus is the organ where T cell progenitors migrate to and undergo maturation and selection (Kumar *et al.*, 2018). In the thymus, T cells develop mainly into two distinct lineages: CD4<sup>+</sup> T-helper cells and CD8<sup>+</sup> cytotoxic T cells. When the precursors of T cells just enter the thymus from the bone marrow, they are negative for both CD4 and CD8 surface receptors. Here, these CD4<sup>-</sup>CD8<sup>-</sup> cells undergo TCR gene rearrangement (V(D)J recombination) and form either  $\alpha\beta$ -TCR or  $\gamma\delta$ -TCR. The cells will further develop into

CD4 and CD8 double positive cells. The CD4<sup>+</sup>CD8<sup>+</sup> cells then go through a selection process and give rise to CD4 single positive or CD8 single positive cells. During the selection process, self-antigens are presented by APCs with the major histocompatibility complexes (MHCs). The cells with too strong or too weak MHC-TCR binding affinity die at this step. This process is called central tolerance (Hogquist *et al.*, 2005), aiming to eliminate self-reactive T cells. The CD4 single positive cells recognize the antigen presented on MHC class II molecule and will become T helper cells, the CD8 single positive cells recognize MHC class I-presented antigen and will give rise to cytotoxic T cells. The surviving CD4<sup>+</sup> T cells and CD8<sup>+</sup> T cells leave the thymus and enter the periphery as naive T cells, meaning they have no experience of encountering foreign antigens. The naive T cells have diverse TCR sequences and only a small fraction of naive T cells is specific for a given antigen.

### 1.2.2 T cell activation

Naive T cells circulate through the periphery blood to lymphoid organs such as lymph nodes. In lymph nodes, APCs can present the peptides (antigens) from the processed pathogens to naive T cells. The naive T cells with matched TCRs can be activated via the TCR signaling pathways (Fooksman *et al.*, 2010). Activated T cells start to proliferate (a process called clonal expansion, resulting in offspring T cells with same TCR sequences) and differentiate into effector and memory subsets. The effector CD4<sup>+</sup> T helper cells can produce cytokines that further activate APCs and augment antigen presentation. This positive feedback loop will rapidly activate the immune response. The cytokines can also stimulate B cells to produce large amounts of antibodies to fight against the pathogens. The effector cytotoxic CD8<sup>+</sup> T cells can kill infected cells. The memory CD4<sup>+</sup> and CD8<sup>+</sup> T cells can live for a long period and when they encounter pathogens next time, they can be activated and have effector functions in a short time.

### 1.2.3 CD4<sup>+</sup> T cell differentiation

The activation and differentiation of naive CD4<sup>+</sup> T cells leads to distinct subsets of T helper cells. The T helper cells can mainly be characterized by their key transcription factors and specific cytokines (Fig 1). The major CD4<sup>+</sup> T helper cell types are: T helper 1 (Th1); T helper 2 (Th2); T helper 17 (Th17); T follicular helper cells (Tfh); regulatory T cells (Treg); and Type 1 regulatory T cells (Tr1).

#### 1) Th1

The earliest discovery of CD4<sup>+</sup> T helper cells was the description of Th1 and Th2 cells in the late 1980s (Mosmann and Coffman, 1989). Th1 differentiation is induced by Interleukin 12 (IL-12) from the innate immune cells and can be further regulated by the lineage transcription factor T-bet. Th1 cells produce interferon- $\gamma$  (IFN- $\gamma$ ) as their key cytokine (Liew, 2002). The IFN- $\gamma$  produced by Th1 cells can activate and enhance the function of macrophages during antibacterial and antiviral responses. Therefore, Th1 cells are critical for cellular immunity.

#### 2) Th2

Th2 differentiation is promoted by IL-25 and IL-33, both secreted by epithelial cells. The lineage-specific transcription factor for Th2 cells is GATA-3. GATA-3 promotes the expression of IL-4, IL-5 and IL-13, key signature cytokines of Th2 cells. IL4 can drive B cell proliferation and antibody production (Liew, 2002). At the same time, GATA-3 can also repress the Th1 differentiation. Th2 cells play important roles in allergic and parasite infection such as asthma and worm infection (Braga *et al.*, 2019).

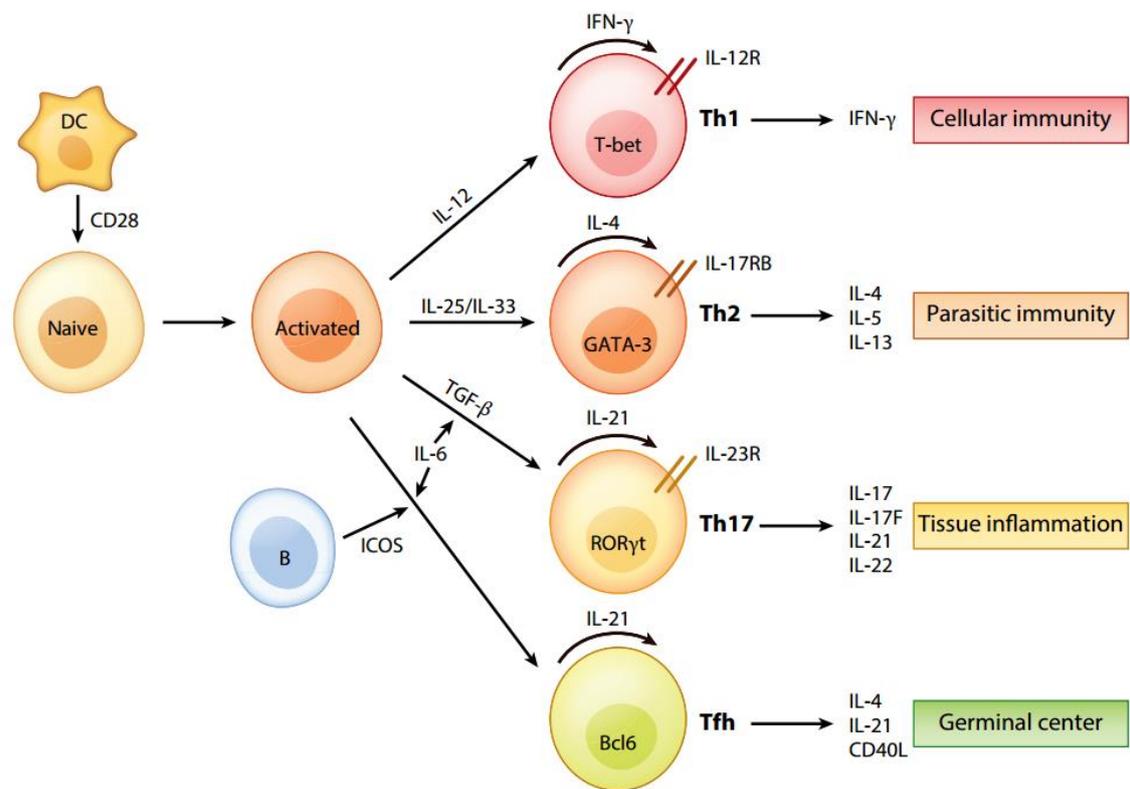
#### 3) Th17

In 2005, a new subtype of CD4<sup>+</sup> T help cells were identified (Harrington *et al.*, 2005; Park *et al.*, 2005). This new subtype of T cells produces IL-17A and can be induced at the early stage by the combination of two cytokines: IL-6 and transforming growth factor beta (TGF- $\beta$ ). The master transcription factor of Th17 cells is Retineic acid receptor related

orphan nuclear receptor gamma (ROR $\gamma$ t). In addition to ROR $\gamma$ t, the transcription factors ROR $\alpha$  and STAT3 are also important for Th17 differentiation. Although IL-17A is the key cytokine, Th17 cells also produce IL-17F, IL-21, IL-22, and GM-CSF (Granulocyte-macrophage colony-stimulating factor) (Korn *et al.*, 2009). These cytokines can serve to recruit neutrophils and monocytes to the site of infection and enhance inflammation. Th17 cells contribute to eliminating bacteria and fungi such as *Staphylococcus aureus* and *Candida albicans* (Huang *et al.*, 2004; Ye *et al.*, 2001). In addition, they are also involved in tissue inflammation and multiple autoimmune diseases including inflammatory bowel disease (IBD), psoriasis and multiple sclerosis (Duerr *et al.*, 2006; Liu *et al.*, 2008).

#### 4) Tfh

T follicular helper cells (Tfh) are a special subset of CD4<sup>+</sup> T helper cells. They are located in the germinal center structure in lymphoid organs such as lymph nodes and can facilitate B cell activation and antibody production (Vinuesa *et al.*, 2016). Tfh can be induced by IL-6 and ICOS-L (Inducible costimulator-ligand) produced by DCs. The signature transcription factor of Tfh cells is B cell lymphoma 6 (BCL6). Tfh cells produce IL-4 and IL-21 which can drive B cell proliferation. Tfh cells can also give co-stimulation to the B cells by CD40-ligand (CD40L) interacting with CD40 expressed on the B cell surface (King *et al.*, 2010).



**Fig 1. T helper cells differentiation.** Naive  $CD4^+$  T cells activate upon stimulation from antigen presenting cells such as dendritic cells (DCs). Activated T cells further differentiate into distinct helper T cell subtypes induced by different cytokines. In the center of Th1, Th2, Th17 and Tfh cells, the key lineage-specific transcription factors are shown. The cell type-specific receptors are also shown on the surface of the cells. T helper cells produce different cytokines which further facilitate the differentiation and mediate different types of immunity or inflammation functions. Adapted from Dong (Dong, 2021).

## 5) Treg

Regulatory T cells (Tregs) are essential for the regulation of the immune cells and prevent the immune system from over-activation. In other words, they suppress the other immune cells to keep the balance state of the system. The key signature transcription factor of Tregs is Forkhead box P3 (FOXP3). CD25 is an important surface marker of Tregs. Tregs also express CTLA4 (Cytotoxic T-lymphocyte-associated Protein 4) on the surface, which can interact with B7 surface proteins expressed by APCs. This CTLA-4/B7 interaction can inhibit the activation of APCs. Tregs can also produce anti-inflammatory soluble mediators such as IL-10, TGF $\beta$  and IL-35 (Raffin *et al.*, 2020).

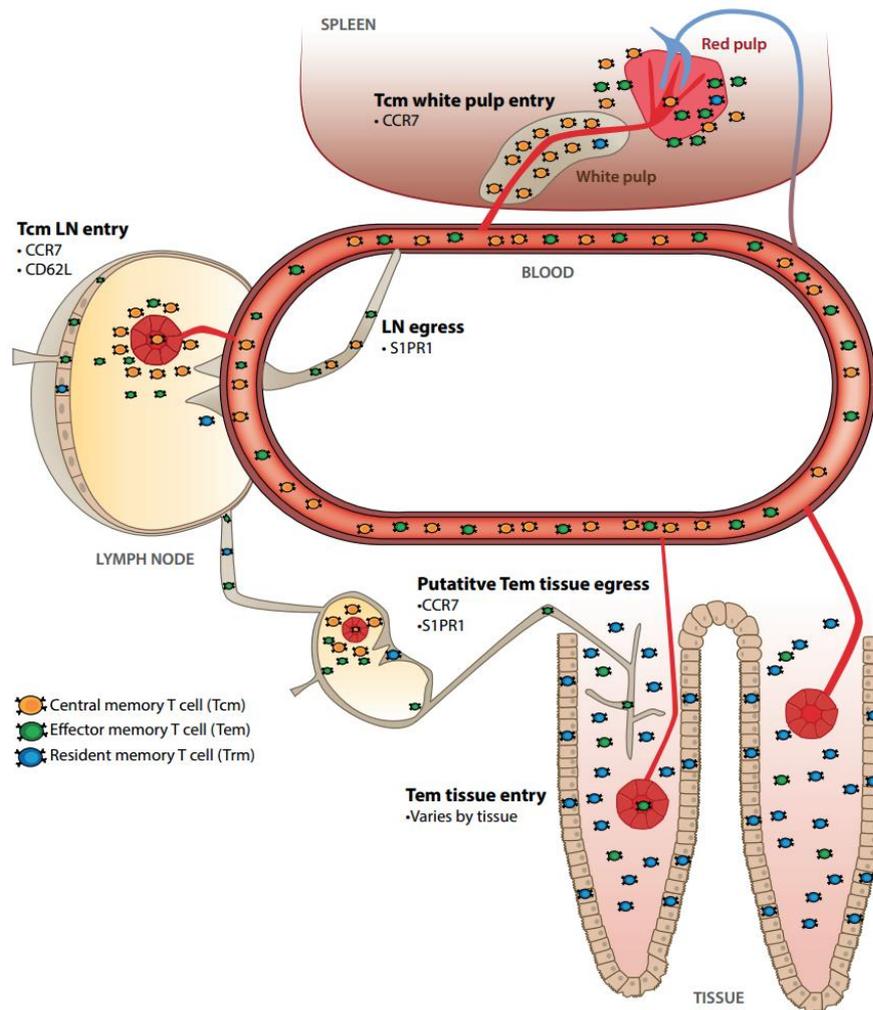
## 6) Tr1

Type 1 regulatory T cells (Tr1) are a subset of CD4<sup>+</sup> T helper cells that prevent and downregulate inappropriate immune responses. Compared to the conventional FOXP3<sup>+</sup> Tregs described above, Tr1 cells do not constitutively express the transcription factor FOXP3. Functionally, FOXP3<sup>+</sup> Treg cells are important for the initial tolerance induction at the inflamed organs, while Tr1 cells are critical to maintain long-term tolerance (Battaglia *et al.*, 2006). Tr1 cells secrete predominantly cytokines IL-10 and TGF- $\beta$ . IL-10 can downregulate co-stimulatory, MHC II and pro-inflammatory cytokine expression of APCs. TGF- $\beta$  can also downregulate APC functions and inhibit T cell proliferation and cytokine production (Roncarolo *et al.*, 2014). The most important surface markers of Tr1 cells are CD49b (or ITGA2, integrin alpha2 subunit) and LAG-3 (lymphocyte-activation gene 3) (Gagliani *et al.*, 2013). In addition to these markers, the expression of PD-1, CTLA-4, TIM-3, TIGIT, CD226, ICOS, and CCR5, has also been associated with Tr1 cells (Roncarolo *et al.*, 2018). Unlike the previously described other T helper cell subsets, no master transcription factor defining Tr1 cells has been determined so far. Several transcription factors including Eomes, IRF4, Blimp-1, c-Maf and AhR, can directly regulate IL10 gene expression underlying Tr1 development (Roncarolo *et al.*, 2018).

### 1.2.4 Memory T cell compartmentalization

Antigen presentation and T cell activation lead to clonal expansion of pathogen-specific T cells. These T cells differentiate into effector T cells including CD4<sup>+</sup> T helper cells and CD8<sup>+</sup> cytotoxic T cells. After pathogen clearance mediated by these expanded effector T cells, most of them (more than 90%) die through programmed cell death; a fraction of the T cells, however, can form memory T cells and will keep the ability to fight against the same pathogen upon second exposure (Gourley *et al.*, 2004). This long-term immunologic memory is also the foundation for vaccination.

Memory T cells classically determined by CD45RO isoform expression and lack of CD45RA isoform (Farber *et al.*, 2014). The CD45RO<sup>+</sup>CD45RA<sup>-</sup> memory T cells can be further divided into 3 main subtypes: central memory T cells (TCMs), effector memory T cells (TEMs), and tissue resident memory T cells (TRMs) (Fig 2).



**Fig 2. Compartmentalization of memory T cells.** TCMs recirculate through the bloodstream and migrate through lymph nodes and spleen. TEMs can also migrate to different tissues. TRMs reside at the tissue and do not participate in blood circulation. The key proteins required for tissue entry or egress are labeled. Adapted from Masopust and Soerens (Masopust and Soerens, 2019).

Both of TCMs and TEMs can circulate through the body via blood stream. TCMs predominantly enter secondary lymphoid organs such as lymph nodes in the absence of

inflammation. They constitutively enter lymph nodes via high endothelial venules. TCMs are easy to proliferate, but have no effector functions.

In contrast to TCMs, TEMs can still rapidly execute effector functions. TEMs can recirculate through the bloodstream and they are abundant in non-lymphoid tissues (Rosato *et al.*, 2017). However, TEMs do not enter uninflamed lymph nodes through high endothelial venules (Masopust and Soerens, 2019).

TRMs are non-recirculating memory T cells that stably reside in non-lymphoid tissues and do not participate in blood circulation. The major function of the TRMs is that they can rapidly recruit other immune cells specifically to the site of infection; therefore, TRMs serve as immune sentinels in different tissues.

The main criteria to distinguish these memory T cell subtypes is their surface protein expression. CCR7 can bind to its ligand CCL21 expressed on high endothelial venules, therefore, CCR7 is critical for TCMs to enter resting lymph nodes. TCMs are CD45RA<sup>+</sup>CCR7<sup>+</sup>, whereas TEMs and TRMs are CD45RA<sup>-</sup>CCR7<sup>-</sup>.

The majority of CD4<sup>+</sup> and CD8<sup>+</sup> TRMs in multiple organs express the early T cell activation marker CD69 (Kumar *et al.*, 2017). CD69 promotes T cell tissue retention by downregulating transcription factor KLF2 (Kruppel like factor 2) and S1PR1 (sphingosine-1-phosphate receptor), which is required for tissue exit (Krebs *et al.*, 2016; Mueller and Mackay, 2016). For CD8<sup>+</sup> TRMs particularly at mucosal sites and skin, the  $\alpha$ E integrin (CD103) is also an important surface marker (Kumar *et al.*, 2018).

### 1.2.5 Function of tissue resident memory T cells

The conventional immune cell analysis focused on the blood and lymphoid tissues. Because of their location, TRMs are less accessible and have been neglected by

immunology researchers for many years. The total number of TRMs distributed at different organs is much higher than all the other T cell subtypes (Mueller and Mackay, 2016). TRMs are local specialists in immune defense and play essential roles in infection, allergy, cancer and autoimmunity.

When TRMs encounter cognate antigens, they can initiate and organize the immune response in several forms. First, TRMs can produce cytokines such as IFN- $\gamma$ , TNF- $\alpha$ , and IL-2. Second, TRMs can activate endothelium and recruit other immune cells such as TEMs, TCMs, DCs, NK cells. Activation of all these immune cells further strengthens the response to infection. Third, some CD8<sup>+</sup> TRMs can directly kill the infected cells via cytotoxic mechanisms. Finally, TRMs can undergo proliferation and therefore amplify the immune response (Masopust and Soerens, 2019).

Promoting local immune response by TRMs at the infection sites is beneficial to fight against reinfection, however, when TRMs recognize self-antigens or get activated inappropriately, they can also cause severe damage to the tissue (Sasson *et al.*, 2020). Recent studies have shown the pathogenic role of TRMs in various autoimmune diseases such as inflammatory bowel disease, psoriasis, and multiple sclerosis (Cheuk *et al.*, 2017; Kleinschek *et al.*, 2009; Sasaki *et al.*, 2014). Due to the ability of local proliferation, cytokine production, cytotoxicity and immune cell recruitment, TRMs can be the key accelerator of the inflammation in autoimmunity and the trigger of the relapses. Characterization and functional investigation of TRMs in different inflammatory diseases is crucial to understand the disease pathology and to develop potential therapeutic methods.

### 1.2.6 Summary of T cell subtypes

In this section, the different T cell subtypes were introduced. From the functional perspective, T cells include CD4<sup>+</sup> T helper cells and CD8<sup>+</sup> cytotoxic T cells. CD4<sup>+</sup> T helper

cells contain mainly Th1; Th2; Th17; Tfh; Treg and Tr1 subtypes. Of note, in addition to the classic CD4<sup>+</sup> or CD8<sup>+</sup> T cell, there are also  $\gamma\delta$ T cells and innate-like T cells such as NKT (Natural killer T) cells and MAIT (mucosal-associated invariant T) cells.  $\gamma\delta$ T cells express a unique TCR composed of one  $\gamma$  and one  $\delta$  chain. They are less frequent than  $\alpha\beta$ T cells, but they are present in tissues at high risk of infection and can mediate immune responses in healthy and disease conditions (Ribot *et al.*, 2021).

From the activation state perspective, T cells can be classified into naive, effector and memory subtypes. Memory T cells can be further divided into TCMs, TEMs and TRMs according to the migration properties. Understanding the role of distinct T cell subtypes is crucial to investigate the pathogenic process of inflammatory diseases.

### **1.3 Inflammatory diseases**

In this section, I will introduce the representative inflammatory diseases investigated in this thesis, including glomerulonephritis (GN), nonalcoholic fatty liver disease (NAFLD), nonalcoholic steatohepatitis (NASH) and the current Coronavirus disease 2019 (COVID-19). The major organs involved in these diseases are kidney, liver and lung. In all of the diseases, T cells play critical roles in the inflammatory responses.

#### **1.3.1 Glomerulonephritis**

The major function of the kidney is to maintain body homeostasis (e.g. electrolyte and water balance) by purifying toxic metabolic waste and acid from blood. The functional units of the kidney are nephrons, each consisting of one glomerulus and one tubule structure. Thousands of glomeruli filter blood into the tubular compartment and eventually form urine.

Glomerulonephritis (GN) is a group of inflammatory diseases of the glomeruli. In general, GN results from invasion of the kidney by autoantibodies, immune cells and complexes. These immune components either come from blood circulation or they form in situ at the kidney (Suárez-Fueyo *et al.*, 2017). The inflammation disrupts the function of other glomerular cells and tubular cells or even causes direct tissue damage.

The most aggressive form of GN is called rapid progressive glomerulonephritis (RPGN). In RPGN patients, strong renal inflammation leads to severe proteinuria and significant loss of kidney function. Clinically, a crescent-shaped structure is formed in Bowman's capsule. Thus, RPGN is also named crescentic glomerulonephritis (cGN). The crescents in RPGN result from proliferation of glomerular parietal epithelial cells and infiltrating immune cells (Anguiano *et al.*, 2020). RPGN includes 3 different forms based on the pathogenic mechanisms (Couser, 1988): anti-glomerular basement membrane glomerulonephritis (anti-GBM-GN); immune complex-mediated RPGN and anti-neutrophil cytoplasmic antibody associated GN (ANCA-GN).

In anti-GBM-GN, antibodies are produced against the  $\alpha 3$  chain of type IV collagen, a component of the glomerular basement membrane. This will lead to deposits on the basement membrane and cause dysfunction of the blood filtering (Cui and Zhao, 2011).

In immune complex-mediated RPGN, antibody-antigen immune complexes are deposited at the basement membrane. Typical examples are IgA nephropathy and lupus nephritis. The immune complexes can mediate the recruitment of other immune cells (e.g. neutrophils) and further induce tissue injury (Mayadas *et al.*, 2009).

The most common form of RPGN is called ANCA-GN, an inflammation of the blood vessels (vasculitis) in which the kidney is often also involved (Netea *et al.*, 2020). In ANCA-GN, autoantibodies are formed against components of neutrophils. Two autoantigens have been identified: myeloperoxidase (MPO) and proteinase 3 (PR3). These two autoantigens and their autoantibodies are not kidney specific, however, in areas of the

kidney with a rich neutrophil environment, they can form immune complexes (Jennette and Falk, 2014).

T cells play a critical role during various renal autoimmune responses (Martinez Valenzuela *et al.*, 2019; Suárez-Fueyo *et al.*, 2016). On the one hand, T cells can boost the survival, proliferation and autoantibody production of self-reactive B cells. On the other hand, T cells can also produce cytokines to promote inflammation and use their cytotoxic function to directly injure renal tissue in different types of GN (Suárez-Fueyo *et al.*, 2017; Tipping and Holdsworth, 2006).

### 1.3.2 NAFLD and NASH

The liver plays essential roles in metabolism, nutrient storage and detoxification. Nutrients and lipids absorbed from the gut can enter the liver via portal vein. However, together with the nutrients, many gut-derived pathogens can also enter the liver. Liver also serves an important role as a frontline immune organ, in addition to its role in metabolic functions. Large immune related-molecules can be produced in the liver, such as complements, cytokines, chemokines, and acute phase proteins (Crispe, 2009). Liver also contains the largest population of macrophages in the human body (Kubes and Jenne, 2018).

Excess glucose in the blood can be transformed into fat and stored in the liver. When the fat accumulation exceeds 5% of the liver weight, it is termed to be “steatosis” (fat increase). If the process is in the absence of significant alcohol consumption, it will be described as nonalcoholic fatty liver disease (NAFLD) (Rau *et al.*, 2016). NAFLD is tightly linked to the worldwide increase in overweight and obesity. Around 20-30% of the NAFLD patients can develop into a more severe form of the disease, called nonalcoholic steatohepatitis (NASH). Histologically, NASH is defined by the combination of liver steatosis; parenchymal damage (hepatocyte apoptosis and ballooning); lobular and/or portal inflammation and a variable degree of fibrosis (Rinella, 2015). In NASH patients, the

hepatocyte apoptosis and inflammation further activate hepatic stellate cells. These cells produce collagen fibers and form scar tissue. Up to one-third of the NASH patients can progress to advanced fibrosis or cirrhosis. NASH cirrhosis is a rising cause of hepatocellular carcinoma (Baffy *et al.*, 2012). Due to the prevalence of obesity worldwide and lack of treatment strategy, NASH has become the major cause of the end-stage liver failure and liver transplantation (Kazankov *et al.*, 2019).

The hepatic inflammatory response is a key driver of NAFLD-NASH disease progression (Koyama and Brenner, 2017). Numerous previous studies focused on the innate immune responses in NASH (Friedman *et al.*, 2018; Tilg and Moschen, 2010). The central players of the NASH innate immunity are the resident Kupffer cells, which represent around one-third of the non-parenchymal cells in the liver (Bilzer *et al.*, 2006). The activation of Kupffer cells leads to production of pro-inflammatory cytokines and recruitment of more neutrophils and monocytes to the liver.

Emerging evidence from recent studies, however, suggests the role of the lymphocyte-mediated adaptive immunity in NASH progression (Sutti and Albano, 2020). More than 50% of the NASH patients have focal aggregates of B cells and T cells. And the size of the aggregates correlates with the inflammation and fibrosis scores (Bruzzì *et al.*, 2018). Understanding the adaptive responses, especially functions of the distinct T cell subtypes, is crucial to elucidate the mechanisms underlying inflammation in the NAFLD and NASH patients. The interplay of T cells and innate immune cells is also an important research direction.

### 1.3.3 COVID-19

The lung is the organ responsible for gas exchange. Air flows in through the mouth and nose and reaches the lungs via the respiratory tract. Inside the lungs, the airway ends in clusters of saccular structures called alveoli, surrounded by blood vessels. Alveoli is the

place where oxygen and carbon dioxide exchange occurs. As a barrier tissue exposed to the external environment, the lung also plays an essential role in mucosal immunity. Some pathogens, including bacteria and viruses, however, can infect lung cells and cause lung inflammation, namely pneumonia.

Coronavirus disease 2019 (COVID-19) is a type of lung infection caused by the worldwide rapid spread of the severe acute respiratory syndrome coronavirus 2 (SARS -CoV-2) since December 2019. COVID-19 has affected 252,137,836 patients, with 5,083,512 deaths globally (source: Johns Hopkins University, as of November 12, 2021). We are currently still in this unprecedented pandemic which brings enormous threat to health care and economic systems globally.

COVID-19 patients range from asymptomatic or mild to severe cases that require hospitalization. Those hospitalized severe patients usually have severe pneumonia (lung inflammation) and some develop acute respiratory distress syndrome, septic shock and even multiple organ failure (Huang *et al.*, 2020; Matthay *et al.*, 2020; Xu *et al.*, 2020). The typical symptoms of the COVID-19 patients are: fever, shortness of breath, expectoration, fatigue, dry cough, and myalgia (Qin *et al.*, 2020). Around 20-30% hospitalized patients need to be treated in intensive care units with respiratory support such as extracorporeal membrane oxygenation (Chen *et al.*, 2020; Huang *et al.*, 2020).

SARS-CoV-2 is mainly transmitted through respiratory droplets and infects cells expressing the surface receptors ACE2 (angiotensin- converting enzyme 2) and TMPRSS2 (transmembrane serine protease 2) (Hoffmann *et al.*, 2020). These cells include airway epithelial cells, alveolar epithelial cells and vascular endothelial cells in the lung. SARS-CoV-2 replicates in these host cells and the released copies will infect more cells. The damaged host cells can release DAMPs, which activate the innate immune cells such as alveolar macrophages (Tay *et al.*, 2020). The activated innate immune cells can further promote the inflammation by producing pro-inflammatory cytokines and recruiting other immune cells. If the adaptive immune cells are able to establish a specific and appropriate

response, the virus and the infected cells can be rapidly cleared. Then the inflammatory response will be tuned down. This would be an ideal case for a healthy immune response in COVID-19.

However, in severe COVID patients, substantially elevated circulating levels of pro-inflammatory cytokines were detected. The cytokines mainly include: IL-1, IL-2, IL-6, IFN- $\gamma$ , TNF- $\alpha$ , IL-17A and GM-CSF (granulocyte- macrophage colony-stimulating factor) (Qin *et al.*, 2020; Tan *et al.*, 2020; Xu *et al.*, 2020). This overproduction of pro-inflammatory cytokines is termed cytokine storm (Mangalmurti and Hunter, 2020). One severe effect of the cytokine storm is: the vascular integrity is lost, more leukocytes infiltrate into the lung, fluid builds up in the alveoli. The alveoli fluid blocks oxygen exchange in the lung and leads to respiratory failure (Teuwen *et al.*, 2020). The cytokine storm can also circulate to different organs and induce multi-organ failure (Tay *et al.*, 2020).

It is crucial to characterize the cytokine-producing cells at the site of infection. Due to the important role of T cells in cytokine production and organization of adaptive response, a detailed investigation of the COVID-19-specific T cells at the lung is essential to understand the pathogenicity of severe COVID-19.

In summary, this section introduced different inflammatory diseases investigated in this thesis. To understand the function of T cells in each disease, it is important to first characterize the heterogeneous subtypes of T cells. Previous studies about T cell functions in these diseases can be found in the introduction part of chapter 3, 4, 5 respectively. In the next section, I will introduce the single cell RNA sequencing technique and computational analysis approaches to characterize T cells in different inflammatory diseases. Regarding the biological motivation of single cell sequencing and its application in interpreting immune functions, please see my latest review article “Single cell biology to decode the immune cellular composition of kidney inflammation” in chapter 2.

## **1.4 Single cell RNA sequencing technique and data analysis**

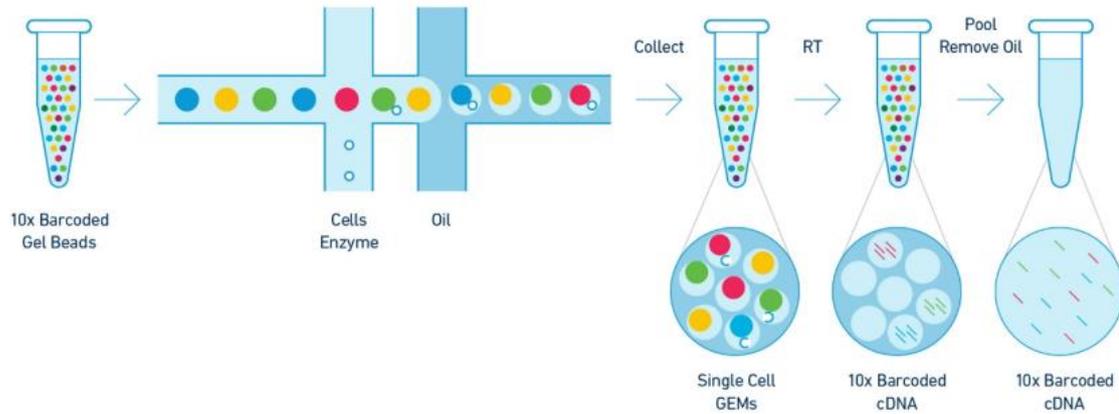
### **1.4.1 Single cell RNA sequencing technique**

Cellular identity and function is reflected by its gene expression. High-throughput RNA sequencing (RNA-seq) enables the simultaneous measurement of the expression level of tens of thousands of genes. The conventional RNA-seq approach is termed bulk RNA-seq, meaning the gene expression levels are measured on an entire sample. Bulk RNA-seq only reveals the average transcriptome across a mixed group of cells. Therefore, the heterogeneity of different cell populations in a tissue is hidden in this approach.

Single cell RNA sequencing (scRNA-seq) measures the gene expression of the single cells instead of the population level. This new technique allows us to characterize and identify heterogeneous cell populations in an unbiased way. Cellular barcoding is the core method of the single cell sequencing technology. Two main strategies of cellular barcoding are droplet-based (cells captured in nanoliter droplets) and plate-based (cells sorted into multi-well plates). Comparisons of plate and droplet-based single cell sequencing methods have been reviewed in depth previously (Papalexi and Satija, 2018).

Here I briefly introduce the droplet-based approach to achieve single cell isolation, using the 10x Genomics Chromium platform as an example since all projects in this thesis were performed on this platform. As illustrated in Fig 3, a microfluidic device containing different channels is used for single cell isolation. Dissociated cells (digested from tissue or further enriched by fluorescence-activated cell sorting) are injected to the microfluidic device from one channel. The oligonucleotides-barcoded gel beads are injected from another channel. Cells are partitioned into nanoliter-scale GEMs (Gel bead-in-Emulsions). Each cell is captured together with a barcoded gel bead into one oil droplet. Afterwards, the cells are lysed inside the droplets and the released mRNA from the same cell are attached to the same barcode sequence. The mRNAs are then transformed to cDNA by reverse transcription. The cDNA sequences can later be PCR-amplified and sequenced

together after breaking the oil droplets. Each cDNA sequence contains the original oligonucleotides sequence of the cell barcode, therefore, each sequence can be traced back to its original cell.



**Fig 3. Illustration of droplet-based single cell capture of 10x Genomics Chromium.** Barcoded beads, dissociated cells, as well as reacting enzymes are injected to the microfluidic device from different channels. Each oil drop captures one cell and one bead, forming single cell GEMs. Cells are lysed within the GEMs and mRNAs are released and attached to barcoded beads. After reverse transcription and oil removal, the cDNAs can be pooled as a library and further sequenced (Adapted from 10xgenomics.com).

Of note, scRNA-seq can be also combined with surface protein measurement (CITE-seq) and BCR/TCR sequencing (VDJ-seq). These techniques were also employed in this thesis and further introduced in more detail in Chapter 2. In the remaining part of the section, the scRNA-seq analysis workflow will be introduced stepwise.

### 1.4.2 Pre-processing and quality control

The starting point of scRNA-seq data analysis is to obtain a gene-cell count matrix from raw sequencing reads files. This step can be achieved by alignment pipelines such as Cell Ranger from the 10x Genomics company. Each sequencing read needs to be assigned to a cell barcode and aligned to the reference genome. Then the expression of each gene in each cell will be counted. In the output gene-cell count matrix, each row is a gene; each column

is a cell; each value in the matrix represents the counts of one gene detected in one particular cell.

After obtaining the gene-cell count matrix, it is essential to perform a quality control step in order to exclude the low quality cells from the downstream analysis. Low quality cells include several situations: dying cells; damaged cells; not fully dissociated cells i.e. multiple cells were tagged with the same barcode (doublets) (McGinnis *et al.*, 2019). Dying and damaged cells typically have low total gene counts per cell and high mitochondrial gene counts per cell, since their cytoplasmic mRNAs leaked out. Doublets typically have aberrantly high total gene counts per cell. Therefore, setting a threshold for total gene counts and mitochondrial gene fraction per cell is a standard step to filter out low quality cells. The threshold settings depend on each dataset due to different protocols and experimental conditions.

### 1.4.3 Normalization and integration

Since each cell is captured independently, the count depth can vary a lot even for identical cells. In order to avoid sample effects and make the count data among different cells comparable, a normalization step is necessary after quality control.

A common single cell normalization method is global count depth scaling (Vallejos *et al.*, 2017). Different gene counts for each cell were divided by total counts of the cell, then scaled with a size factor, typically 10,000. The results will be further log-transformed as the output normalized count values. The log-transformation step can reduce the skewness of the zero-inflated distribution (due to miss detection of many weakly expressed genes) of the single cell count data and make it easier for downstream analysis such as differential expression (Luecken and Theis, 2019).

Another important challenge of analyzing single cell data is to deal with the batch effects of different samples. Batch effects refer to the technical variation between samples processed in distinct batches. In the single cell data context, it can be samples from different patients of one group, samples from different research groups, or even samples from different single cell methods. Many computational tools have been developed to integrate data from different batches.

Popular strategies of single cell data integration include CCA (canonical correlation analysis) and MNN (mutual nearest neighbors). CCA identifies a multi-dimensional subspace consistent across datasets (Butler *et al.*, 2018). It aims to find projections of two datasets where the gene-gene correlation between the two projections is maximized. The mathematical expression for CCA is

$$\max_{u,v} u^T X^T Y v \text{ subject to } u^T X^T X u \leq 1, v^T Y^T Y v \leq 1$$

where  $X, Y$  are two gene expression matrices. The goal is to find projection vectors  $u$  and  $v$  such that the correlation between the two indices  $Xu$  and  $Yv$  is maximized. With the projection vectors, two datasets could be further aligned together.

MNN calculates the distance of cells between datasets and tries to identify the cells that share a neighbourhood (Haghverdi *et al.*, 2018). MNN identifies mutual nearest neighbours by computing the Euclidean distance between the normalized expression profiles of pairs of cells. The Euclidean distance between two cells is formulated as

$$distance(a, b) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

where  $a, b$  represent two cells, each as a point in Euclidean space.  $a_i$  and  $b_i$  are the normalized expression values of gene  $i$  in cell  $a$  and cell  $b$ . In total  $n$  genes are calculated. If a pair of cells from different batches have small distances to each other, these cells are considered as mutual nearest neighbours. Similar cell types from different batches should have mutual nearest neighbours. The systematic differences between cells in MNN pairs should reflect the batch effect.

In different single cell projects of this thesis, the anchor method from the Seurat package was used for sample integration (Stuart *et al.*, 2019). This method combines CCA and MNN, and builds “anchor” cells that have shared features. There are also other algorithms for data integration and batch removal, such as Harmony (Korsunsky *et al.*, 2019). Harmony projects cells from different datasets into a shared embedding and groups them by cell type instead of batch conditions. Integration methods based on different machine learning approaches were reviewed in detail by Oller-Moreno *et al* (Oller-Moreno *et al.*, 2021).

#### 1.4.4 Dimensionality reduction and visualization

Human scRNA-seq count matrices can contain up to 30,000 genes. Each cell can be regarded as a point in a 30,000 dimensional space. It would be a heavy computational burden to compute the distance of the points with such huge dimensions. Therefore, dimensional reduction is a necessary step for scRNA-seq data processing.

One straightforward approach to reduce dimensions is to filter out some genes that do not vary across different cells. The expression level and variability of each gene can be calculated and the highly variable genes (HVGs) are helpful to distinguish different cell types (Brennecke *et al.*, 2013). The typical number of HVGs for downstream computation is 1000-5000. The selected HVGs and their expression values will be used for further dimensional reduction analysis.

Single cell sequencing data analysis usually contains both linear and non-linear dimensionality reduction approaches. The most commonly used linear approach is PCA (Principal component analysis) (Townes *et al.*, 2019). PCA calculates principal components (PCs), each PC is a linear combination of the genes. Different PCs are ranked according to their contribution to the variance of the data. In other words, the first PC

explains the largest amount of variation in the dataset. The top 10-50 PCs, which capture the majority of variation between cells, are usually taken for downstream analysis.

Since the scRNA-seq count matrix is zero-inflated and has a non-linear structure, PCA can not serve as an optimized method for visualization. The two most popular non-linear dimensional reduction approaches for single cell data visualization are t-SNE (t-distributed Stochastic Neighbour Embedding) (van der Maaten and Hinton, 2008) and UMAP (Uniform Approximation and Projection) (McInnes *et al.*, 2020). t-SNE keeps the local similarity while overlooking the global structure. Cells from the same group are very close at the low dimensional space, but the distance between different groups is exaggerated (Luecken and Theis, 2019). UMAP is a more recent graph-based dimensional reduction approach. Compared to t-SNE, UMAP can better preserve the global data structure and the continuity of the cell subsets (Becht *et al.*, 2019).

#### 1.4.5 Clustering and cell type annotation

Clustering is an unsupervised machine learning approach that aims to group similar cells together based on the gene expression and to assign them as different populations. The similarity of cells can be measured by distance metrics such as Euclidean distance (Kiselev *et al.*, 2019).

K-means is a very classical clustering approach where k cluster centers are determined and each cell can be assigned to the closest cluster center. K-means approach has been implemented by many single cell clustering methods, such as SC3 (Kiselev *et al.*, 2017) and SIMLR (Wang *et al.*, 2017a).

Another popular single cell clustering approach is based on K-Nearest Neighbor graph (KNN). This approach represents cells graphically as nodes and the most similar K nodes will be determined as nearest neighbors. Neighboring cells are forming different clusters.

The Louvain algorithm (Blondel, 2008) is based on KNN approach and detects neighborhoods on the graph. The output cluster number can be determined by a resolution parameter. This algorithm has been employed in the most popular single cell analysis tools Seurat (Butler *et al.*, 2018) and Scanpy (Wolf *et al.*, 2018).

Cells can be distinguished as different groups by the unsupervised clustering algorithms, however, to interpret the function of each group, further comparison between clusters need to be performed. This step aims to annotate the cell type for each cluster.

A common practice is to run differential gene expression (DEG) analysis between a single cluster versus all other clusters and iterate the step for each cluster. In the end, genes that are highly expressed in each cluster (with high fold change and small p value in the DEG analysis) are obtained as cluster marker genes. Interpreting the cell types according to the cluster marker genes is a crucial but difficult step. Clustering is an unbiased process which does not guarantee the output marker genes corresponding to canonical cell type markers. Sometimes, the cluster marker genes also reflect cell states such as proliferation or stress. It is important to check the previous studies of the same potential cell types to annotate the clusters properly.

Computational automated cluster annotation methods have been developed recently, such as Garnett (Pliner *et al.*, 2019), scmap (Kiselev *et al.*, 2018), and scID (Boufeva *et al.*, 2020). These tools usually refer to the public databases and ontologies of cell type markers or correlates to other labeled single cell datasets (Pasquini *et al.*, 2021). Our major object in this thesis are T cells from different tissues of patients. However, due to the complexity of subtypes and lack of single cell level study of disease-specific tissue T cells, the automated annotation methods can give limited interpretation. Since different surface markers are helpful to distinguish T cell subtypes, we employed CITE-seq technique for part of our data. For the ANCA-GN (Chapter 3) and COVID-19 (Chapter 5) T cell datasets, we use the combination of cluster marker genes and CITE-seq surface protein information to assign

the cell types. The advantages were also summarized and discussed in my review article in Chapter 2.

#### 1.4.6 Downstream analysis

After the annotation to the distinct cell clusters, further functional interpretation of the cell types can be performed through downstream steps. Here the main downstream analysis methods performed in the projects of this thesis are introduced, including differential gene expression, trajectory and interactome analysis.

##### 1) Differential gene expression analysis

Differential gene expression (DEG) analysis is commonly used when we compare different cell clusters to obtain cluster marker genes. Additionally, for a given cluster, DEG analysis between distinct conditions such as healthy and disease is crucial to understand cell type-specific changes. The popular DEG tools which implement different statistical testing are DESeq2 (Love *et al.*, 2014), EdgeR (Robinson *et al.*, 2009) and MAST (Finak *et al.*, 2015). The Wilcox rank test is also commonly used in single cell packages such as Seurat. DEG analysis typically obtains fold change and p values of each gene and the results can be visualized by volcano plots (see chapter 4, Fig 3B and chapter 5, Fig 2E).

To interpret the function of top DEGs, gene enrichment analysis is usually performed after DEG analysis. In multiple databases such as KEGG and Gene Ontology, biological functions of diverse gene sets are annotated. Enrichment analysis tools typically use the Fisher exact test to find the over-represented functional terms for the top DEG gene sets. As a result, the disease-associated functions or pathways of a certain cell type can be revealed. Commonly used gene enrichment analysis tools include Enrichr (Kuleshov *et al.*, 2016) and Webgestalt (Wang *et al.*, 2017b).

##### 2) Trajectory analysis

Current scRNA-seq technology provides snapshots of dynamic biological processes such as T cell development introduced in previous sections. Trajectory analysis aims to order cells along a continuous trajectory between different states. This strategy is also called pseudotime analysis. The trajectories are built by finding paths through the cells that minimize transcriptional changes between neighboring cells. Trajectory analysis methods have been reviewed in depth previously (Saelens *et al.*, 2019). Popular trajectory algorithms are Monocle (Trapnell *et al.*, 2014), Slingshot (Street *et al.*, 2018), and PAGA (Wolf *et al.*, 2019). In Chapter 4, Slingshot was used to build the trajectories between T cell subtypes in NAFLD and NASH patients.

### 3) Interactome analysis

Another important usage of scRNA-seq data is to decode the cell-cell communications via transcriptional profile, especially through the expression of ligand-receptor genes. Ligand-receptor interactions trigger cellular signaling events and therefore, serve as key to understand communication between different cells in the local environment. In inflammatory conditions, the communication between different immune cell populations via cytokines and their receptors are critical for mediating inflammation response. Several computational tools have been developed to address the interactome using single cell RNA-seq data, such as cellphoneDB (Efremova *et al.*, 2020), CellChat (Jin *et al.*, 2021) and ICELLNET (Noël *et al.*, 2021). For example, cellphoneDB software uses a curated database of ligands and receptors and computes enriched ligand-receptor pairs between two clusters based on ligand gene expression of one cluster and receptor gene expression of the other (Efremova *et al.*, 2020). In Chapter 5, cellphoneDB was used to identify the interaction between T cells and other cell populations in COVID-19 patients.

In summary, this section introduced the scRNA-seq technology as well as representative computational analysis methods. In the following chapters, the different publications will demonstrate how we used these computational methods to address the heterogeneity and functions of T cells in different inflammatory diseases.



## Chapter 2: Publication in *Cell and Tissue Research*, June 2021

### Single-cell biology to decode the immune cellular composition of kidney inflammation

#### Specific contributions:

The following review paper has been published in *Cell and Tissue Research*. I am the first author and corresponding author of the paper. I wrote the manuscript with editing from the other authors.

I, Stefan Bonn, agree with the above statements as the direct supervisor.

Signature ,

Date & Place 15.11.2021 Hamburg





# Single-cell biology to decode the immune cellular composition of kidney inflammation

Yu Zhao<sup>1,2,3,4</sup> · Ulf Panzer<sup>1,4</sup> · Stefan Bonn<sup>1,2,3</sup> · Christian F. Krebs<sup>1,4</sup> 

Received: 22 March 2021 / Accepted: 3 June 2021  
© The Author(s) 2021

## Abstract

Single-cell biology is transforming the ability of researchers to understand cellular signaling and identity across medical and biological disciplines. Especially for immune-mediated diseases, a single-cell look at immune cell subtypes, signaling, and activity might yield fundamental insights into the disease etiology, mechanisms, and potential therapeutic interventions. In this review, we highlight recent advances in the field of single-cell RNA profiling and their application to understand renal function in health and disease. With a focus on the immune system, in particular on T cells, we propose some key directions of understanding renal inflammation using single-cell approaches. We detail the benefits and shortcomings of the various technological approaches outlined and give advice on potential pitfalls and challenges in experimental setup and computational analysis. Finally, we conclude with a brief outlook into a promising future for single-cell technologies to elucidate kidney function.

**Keywords** Single-cell biology · Single-cell RNA profiling · Renal function

## Introduction

The immune system is a complex network composed of various cell types that interact with each other and with parenchymal cells in the tissue. Its function or dysfunction is pronounced in inflammatory diseases, where various immune cells can play a central role in disease pathogenesis. It is the cross-talk between many types of cells that in fact mediates

immune processes (Hewitt and Lloyd 2021). Depending on the specific micro-environmental context, the immune cells communicate with each other and with parenchymal cells of the particular organ (Masopust and Soerens 2019). Characterizing the participating cell types, their cellular networks, unique pathways, genes, and interactions might be key to understand the immune-pathogenesis of immune-mediated kidney diseases. This can provide the basis for manipulating the immune system in a targeted approach.

During the past 20 years, sequencing technologies have revealed a detailed picture of the human genome (McGuire et al. 2020). The roles of genes and transcripts in the development of organisms and disease have been intensively investigated (Rahman et al. 2020). Genome-wide transcriptional profiling paves the way for comprehensive measurements of the molecular state of cells, in lieu of strategies based on selected markers (Yofe et al. 2020). Up until recently, comprehensive genomic analyses relied either on pooling heterogeneous mixtures of cells or on sorting and then profiling subpopulations (Kulkarni et al. 2019). While bulk profiling can provide expression averages, which enables the identification of group differences between a healthy state and disease, it cannot differentiate between changes in cell proportions and cell type-specific gene expression changes (Papalexi and Satija 2018). While recent approaches in cell deconvolution algorithms

---

Stefan Bonn and Christian F. Krebs contributed equally to this work.

✉ Yu Zhao  
yu.zhao@zmnh.uni-hamburg.de

✉ Christian F. Krebs  
c.krebs@uke.de

- <sup>1</sup> Hamburg Center for Translational Immunology (HCTI), University Medical Center Hamburg-Eppendorf, Hamburg, Germany
- <sup>2</sup> Institute of Medical Systems Biology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
- <sup>3</sup> Center for Biomedical AI, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
- <sup>4</sup> Translational Immunology, III. Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

allow for approximate estimation of cell proportions from bulk profiling data, they still lack in accuracy and cell type-specific gene expression determination (Menden et al. 2020).

In the past 5 years, technical progress has enabled the high-throughput analysis of single cells. It is now possible to simultaneously measure thousands of genes and transcripts across thousands of individual cells using microfluidic approaches (Papalexi and Satija 2018). This is made possible by trapping single cells in water droplets in an oil phase. Recent technical advances allowed for single-cell RNA sequencing (scRNA-seq) of very small samples, enabling the profiling of e.g., human biopsies in clinical settings (Braga et al. 2019; Haber et al. 2017; Krebs et al. 2020a; Zheng et al. 2017). In contrast to FACS-sorting and plate-based scRNA-seq techniques, microdroplet-based scRNA-seq approaches cannot directly link protein and transcriptional expression information. Since cell surface proteins are a common means to define cell types and RNA for these proteins might be lacking, purely scRNA-seq-based cell type detection is often difficult. To link the surface protein and transcription profile at the single-cell level for the microdroplet-based approaches, CITE-seq (Stoeckius et al. 2017) was developed and has been widely used in the immune single-cell studies.

B cells and T cells are the adaptive arm of the immune system, and B cells produce antibodies that can neutralize or opsonize pathogens. These antibodies are also present on the cell surface known as B cell receptor (BCR). T cells are defined by the T cell receptor (TCR), which mediates recognition of pathogen-associated epitopes through interactions with peptide and major histocompatibility complexes (pMHCs) (Peters et al. 2020). BCRs/TCRs are generated by genomic rearrangement of the germline BCR/TCR locus, a process termed V(D)J recombination, that has the potential to generate marked diversity of BCRs/TCRs (estimated to exceed  $10^{15}$  possible receptors) (De Simone et al. 2018). Using paired B/T cell receptor sequencing to study V(D)J recombination at the single-cell level (Stubbington et al. 2016) enables researchers to assess BCR/TCR-based clonality and cell migration, while providing deep insights into cellular function and activation.

In this review, we highlight recent advances in scRNA-seq technology and their application to elucidate kidney function in health and disease, with a special focus on immune cells. Combining single-cell transcriptome, BCR/TCR, and CITE-seq information will provide deep mechanistic insights into kidney inflammation and highlight potential novel cell type and organ-specific therapeutic avenues.

## Single-cell expression profiling basics

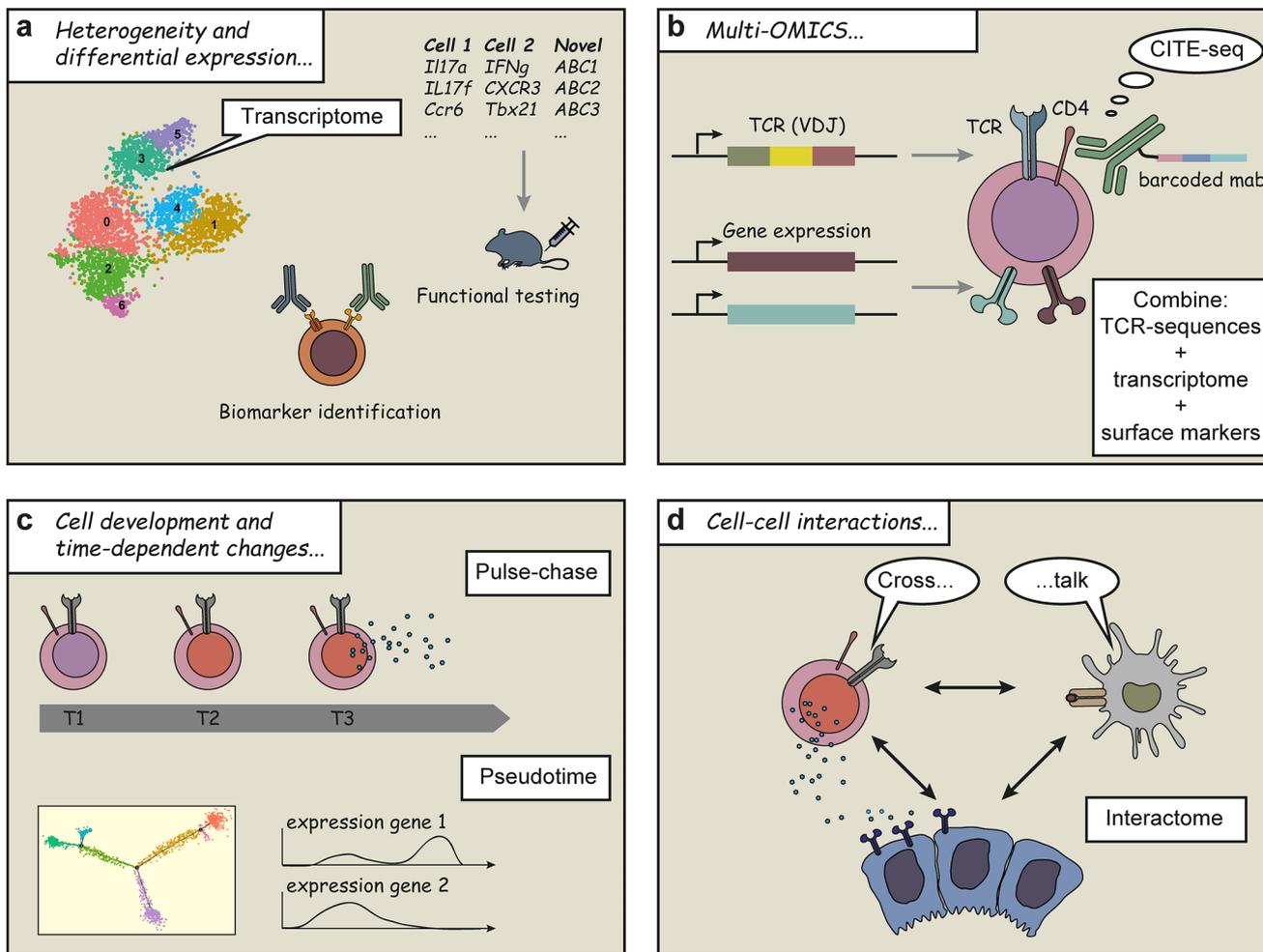
The earliest scRNA-seq study was conducted in 2009 by Tang et al. (2009), in which the transcriptome of a single-cell was analyzed. After around 5 years of technical

improvement, the throughput and quality of the scRNA-seq was dramatically improved. Current standard scRNA-seq protocols include tissue dissociation, single-cell isolation, cell lysis, and reverse transcription, followed by PCR amplification and sequencing (Wu and Humphreys 2017). The key method of isolating single cells is the introduction of the cellular barcoding technique (Stewart et al. 2020). Cells are sorted into multi-well plates or captured in nanoliter droplets. In each small compartment, the single cell will be lysed and the mRNAs from this particular cell will be labeled with a unique cell barcode during the cDNA synthesis. After pooling of the cDNA, the cell barcodes can later be used to trace back the cell origin of each mRNA transcript (Macosko et al. 2015). Detailed comparisons of plate and droplet-based scRNA-seq methods have been reviewed before (Papalexi and Satija 2018; Potter 2018). In brief, plate-based approaches can capture full-length mRNAs and usually capture more transcripts per cell but suffer from relatively low throughput and higher cost (Papalexi and Satija 2018). On the contrary, droplet-based methods are more cost efficient and allow high throughput of up to millions of cells but do not sequence the full length of the transcript.

The choice of sequencing platform has actual ramification for the subsequent information obtained. If deep molecular and splicing information per cell is of essence, then maybe, a plate-based assay should be taken. If it is important to capture many cells and maybe highlight sparse cellular subpopulations with a limited budget, droplet-based assays are favorable.

## Single-cell transcriptome atlases of the kidney in health and disease

Single-cell technologies can be employed to uncover the cellular heterogeneity of cells within the kidney (Fig. 1). One of the first single-cell atlases of mouse kidney was reported in 2018 by Park and colleagues (Park et al. 2018). They performed scRNA-seq of 57,979 murine kidney cells and identified 21 major tubular and glomerular cell types. Besides the cell type identification, they also addressed the cell type specificity of the kidney disease GWAS genes (genome-wide association study) using the expression matrix. Other studies that profiled human kidney single-cell expression were published around the same time (Liao et al. 2020; Sivakamasundari et al. 2017; Wilson and Humphreys 2019). In general, human and murine kidney cell compositions are quite similar in healthy individuals, while human kidney biopsies from allografts (Wu et al. 2018), tumors (Young et al. 2018),



**Fig. 1** Different approaches to high-dimensional analysis of cells by single-cell techniques. This figure gives an overview of some of the many possible applications of single-cell expression profiling. The heterogeneity of cells can be uncovered by gene expression analysis at the single-cell level. This can result in the identification of new biomarkers or in the generation of new hypothesis that can be tested for example in animal models (a). Multi-OMIC approaches

can be performed by combining gene expression analysis with genetic modifications (T or B cell receptor rearrangement) and protein identification in individual cells (b). Developmental trajectories can be investigated by pseudotime analysis (c). Cell-cell interactions can be scrutinized by identifying ligand and receptor matches on different cells (d)

and other kidney diseases such as diabetic kidney disease (Wilson et al. 2019), IgA nephropathy (Zheng et al. 2020), and lupus nephritis (LN) showed differences to healthy murine kidney (Der et al. 2019). These studies provided first insights into renal cell heterogeneity and cell type-specific responses to disease. However, in these first kidney cell atlases, tubular and endothelial cells constituted the vast majority of the observed cell types, while rare cell types, such as immune cells, were hardly detected in healthy and diseased human and murine samples. To fully grasp the impact of the immune cells in renal pathology, a detailed kidney immune landscape would be essential.

### Immune landscape of the diseased human kidney

A recent study explored kidney immune cell heterogeneity in lupus nephritis (LN) patient biopsies (Arazi et al. 2019). The researchers sorted leukocytes with flow cytometry and then performed scRNA-seq. After clustering analysis, they identified 5 macrophage clusters, 7 T cell clusters (including NK cells), and 4 B cell clusters. After normalization and comparison with one living donor control, they showed that the type I interferon response score is higher in LN patients. The IFN signaling pathways were suggested to be potential prognostic markers of LN.

Another interesting recent publication by Stewart et al. profiled the positioning of cells in the kidney (Stewart et al. 2019). By investigating cells from the human kidney and using machine learning to reconstruct the spatial information, they inferred gross anatomical positioning of a cell in the renal tissue. This approach has been summarized and discussed in more detail recently (Krebs et al. 2020b).

## Using CITE-seq to investigate renal T cell subtypes

Approaches that solely rely on profiling single-cell transcripts yield good results when used for the quantification and characterization of cell types that can be profiled in sufficient numbers, e.g., T cell, B cell, and myeloid cell types. For the detection and clustering of rare subpopulations of immune cells, however, the molecular information obtained usually does not suffice to reliably classify these cell types. This problem can be solved by increasing the cell number, which easily gets prohibitively expensive, by presorting of specific cell types using magnetic- or FACS-sorting strategies, or by using CITE-seq antibody-based labeling of cells (Stoeckius et al. 2017).

Three potential challenges of T cell subtype investigation using scRNA-seq, only, are as follows: First, not all the surface markers which define T cell subpopulation have high mRNA expression. In our in-house human renal biopsy scRNA-seq datasets, we noticed that CD4 mRNA expression is low, compared to the detection of surface CD4 protein in e.g., FACS analysis (Krebs et al. 2020c). This observation is in line with publications from other research groups (Ding et al. 2020; Zemmour et al. 2020). Second, some surface protein markers share the same gene origin and are the products of alternative splicing. The most relevant example might be the gene *PTPRC*, which encodes for CD45, CD45RA, and CD45RO. CD45 is expressed on all leukocytes and CD45RA is a marker for naïve T cells whereas CD45RO marks memory T cells (Masopust and Soerens 2019). The expression of *PTPRC* obtained by scRNA-seq approaches cannot easily distinguish the CD45RA and CD45RO T cells if not full-length techniques are employed (Ntranos et al. 2019). Third, single-cell dissociation and other experimental procedures might induce stress and other responses in select cells, resulting in gene expression that might obscure cell identity (O'Sullivan et al. 2019). CD69 protein is expressed in most of the tissue resident memory T cells (Trm cells), but it is also a marker of early activation (Kumar et al. 2017). In our in-house human renal biopsy scRNA-seq datasets, we detected CD69 mRNA almost in all the T cells, although the flow cytometry data shows that only a subset of cells is CD69<sup>+</sup>.

These three potential challenges can be resolved by simultaneous measurement of scRNA-seq and surface proteins, a technique named CITE-seq (cellular indexing of transcriptomes and epitopes by sequencing) (Stoeckius et al. 2017). The core idea of CITE-seq is conjugating polyadenylated DNA barcodes to antibodies targeting cell surface proteins so that the surface proteins can be translated into sequenceable information. The DNA barcode can be captured together with the mRNA from target cells. After the reverse transcription step, a separate cell surface protein-specific library can be obtained and sequenced. The same cell barcode for both RNA and protein antibody will later enable overlap of CITE-seq and scRNA-seq data. Compared to the flow cytometry technique, the advantage of CITE-seq is that it can measure tens of proteins, while the number of fluorescence labels for antibodies in flow cytometry is much more limited.

We performed CITE-seq and scRNA-seq together in our antineutrophil cytoplasmic antibody (ANCA)-associated glomerulonephritis (GN) patients' renal biopsy T cells (Krebs et al. 2020c). Indeed, our CITE-seq data showed robust signal of CD4 surface proteins and could easily distinguish CD45, CD45RA, and CD45RO. The CD69 protein expression is more restricted to a subgroup of T cells and we identified CD69<sup>+</sup>/CCR6<sup>+</sup> Trm17 cells in the datasets from the kidney. We also further validated the tissue resident signature of the Trm clusters by overlapping them with previously reported human Trm signatures (Kumar et al. 2017). By performing functional experiments in animal models, we have identified a T cell subset (Trm17 cells) that can be induced by bacterial infections and reactivated in unrelated inflammation to produce IL-17A and contribute to tissue damage.

## Single-cell VDJ-seq to understand clonal expansion in the kidney and across tissues

In viral or bacterial infections, the antigen-specific naïve B/T cells can divide and expand themselves profoundly (Tu et al. 2019). B/T cell infiltration and expansion have been linked to multiple autoimmune diseases such as multiple sclerosis (MS) (Pappalardo et al. 2020; Arneth 2019), inflammatory bowel disease (Smillie et al. 2019; Mizoguchi et al. 2017), and glomerulonephritis (Krebs et al. 2017; Schrezenmeier et al. 2018), while it is still unclear whether auto-antigen triggered B/T cell expansion occurs in immune-mediated kidney disease (Kitching et al. 2020).

To understand the expansion of B/T cell antigen receptor (BCR/TCR)-specific cell clones, it is essential to sequence variable regions of the BCRs/TCRs that confer antigen specificity, as well as the gene expression of the corresponding cells, a technique called single-cell BCR/TCR

sequencing (O’Sullivan et al. 2019). BCR/TCR sequencing can also serve as a natural barcode to trace B/T cell migration between the kidney, lymph nodes, and peripheral blood.

In our very recent study of severe COVID-19 patients (Zhao et al. 2021), we identified clonally expanded tissue-resident memory-like Th17 cells (Trm17 cells) in the lungs by single-cell sequencing of TCRs and RNA from sorted T cells. In fact, this is the first use of this technique to trace the T cell clones across tissues in these patients. Our analysis further shows these clonally expanded Trm17 cells express high levels of cytokines such as GM-CSF and IL-17A, molecules both implicated in cytokine storms observed in patients with severe COVID-19.

To our knowledge, similar studies about single-cell BCR/TCR sequencing in the autoimmune kidney diseases have not been reported yet, while it is tempting to speculate that renal resident B/T cells might expand upon stimulation and contribute to the pathogenic process in kidney inflammation. Another unanswered question is whether the origin of inflammation resides in the kidney or other organs. In particular in systemic vasculitis like ANCA-associated GN (Kitching et al. 2020), tracking the clones across tissues might be a potential way to shed light on the cellular origin and relations in kidney inflammation.

## Immune cell interactome

Investigating different roles of each immune cell population is crucial to decode the inflammation; however, the immune cells also interact with each other and with parenchymal cells of the kidney via chemokines, cytokines, and their respective receptors (Fig. 1). Using known receptor–ligand interactions, scRNA-seq data can be used to computationally derive potential cellular crosstalk (Arazi et al. 2019; Stewart et al. 2019; Wu et al. 2018). For example, in the LN study by Arazi et al., the authors analyzed the chemokine- and cytokine-mediated cellular networks between the characterized immune cell types and suggested CXCR4 and CX3CR1 as potential therapeutic targets. A more convenient computational tool and database, called CellPhoneDB, has been made to explore ligand–receptor interactions using single-cell data (Efremova et al. 2020). In our study of patients with COVID-19, CellPhoneDB enabled us to obtain a detailed interactome of lung immune cells. Our cell cross-talk analysis suggested that Trm17 cells could potentially interact with other cells associated with COVID-19 severity and lung damage, such as lung macrophages and CD8<sup>+</sup> killer T cells (Zhao et al. 2021). In a kidney setting, computational analysis using CellPhoneDB or similar information could be used to understand the interaction of podocytes and T cells, for example (Fig. 2).

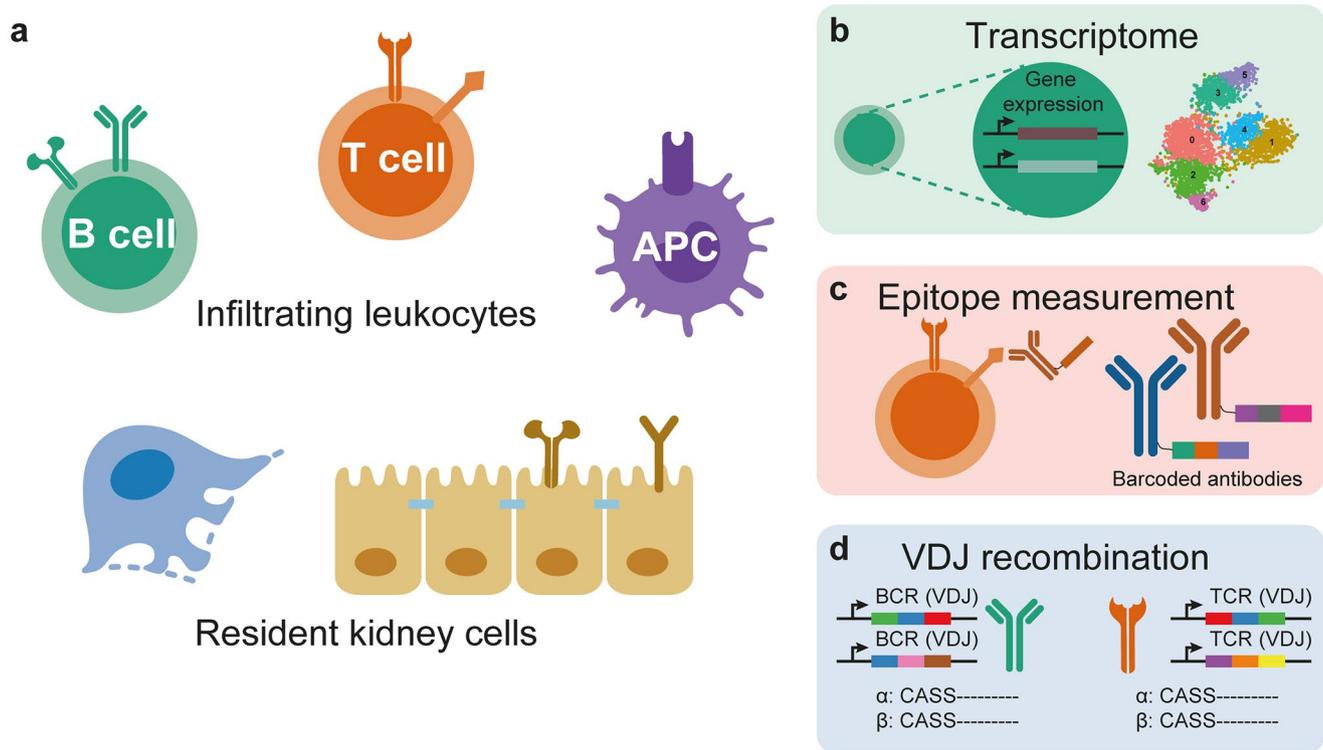
## Key challenges in kidney single-cell immunology

### Sample availability and batch effects

The investigation of immune and tissue cell signaling in solid tissues is complicated by the necessary dissociation of the sample prior to scRNA-seq. The acquisition of fresh samples, the dissociation of the sample into single cells, and the timely further extraction and processing of the RNA for scRNA-seq pose significant challenges and room for error. Over-digestion of samples to extract single cells, for example, will result in stressed and dead cells and bad quality scRNA-seq data. The cell number, isolation efficiency, and experimental bias might all influence the downstream analysis (Saelens et al. 2019). Lots of efforts have been made by the bioinformatics community to remove the batch effects and integrate the different datasets (Stuart et al. 2019). The major challenge in the batch correction field is to remove the technical bias while maintaining the biological differences between samples. Over-correction of samples will hamper the biological interpretation of datasets. The major batch correction methods have evolved from Bayesian algorithms, such as Limma (Smyth and Speed 2003) and ComBat (Johnson et al. 2007), to PCA (principal component analysis), CCA (canonical correlation analysis), MNN (mutual nearest neighbors), and deep learning-based approaches, such as Harmony (Korsunsky et al. 2019), Seurat v2 (Butler et al. 2018), Scanorama (Hie et al. 2019), Seurat v3 (Stuart et al. 2019), LIGER (Welch et al. 2019), and BERMUDA (Wang et al. 2019). Detailed comparisons between multiple batch correction approaches have been summarized recently (Oller-Moreno et al. 2021; Li et al. 2020; Tran et al. 2020). For the kidney immune single-cell datasets, especially the patient biopsy datasets, suitable batch correction methods need to be evaluated within the community in the future.

### Cell subtype identification

Cell type identification is the most crucial step after data quality control and integration. Many downstream interpretation steps rely on the accurate classification of cell types. Determining cell types for individual cells is currently very difficult due to noise and technical zeros (dropouts) in single-cell RNA sequencing. As a result, most of the current cell type identification approaches use common expression patterns of marker genes to identify cell types (Oller-Moreno et al. 2021). The common unsupervised clustering algorithms include partitioning, hierarchical clustering, or graph-based clustering (Petegrosso et al. 2020; Zheng and Wang 2019). Meanwhile, with the increasing of more annotated single-cell datasets, computational biologists also developed supervised methods based on machine learning



**Fig. 2** The combination of single-cell transcriptome sequencing with surface protein measurement and VDJ sequencing. The renal tissue is composed of resident kidney cells, including epithelial cells (podocytes, tubular epithelial cells), infiltrating leukocytes (such as B cells,

T cells, and myeloid cells), and others (a). Single-cell technologies can be used to combine transcriptome sequencing (b), epitope measurement of cell surface molecules (c), and V(D)J recombination of the T and B cell receptors (d)

or deep learning, such as scPred (Alquicira-Hernandez et al. 2019), MARS (Brbić et al. 2020), and rCASC (Alessandri et al. 2019). The amount of single-cell datasets containing detailed characterization of immune cells from the kidney is still limited. Therefore, we need more efforts in the future to apply the supervised algorithms to identify the immune cell subtypes in renal inflammatory diseases.

### Temporal dynamics

All the biological events in the body are complicated dynamic processes such as T cell polarization and activation upon antigen stimuli. However, current single-cell methods are all snapshots of these dynamic processes. It is technically difficult to obtain human renal tissues at different time points and if possible, there is bias given by a different piece of tissue. The computational biologists developed multiple pseudotime analysis algorithms such as PAGA (Wolf et al. 2019), Monocle (Trapnell et al. 2014), Slingshot (Street et al. 2018), and single-cell RNA velocity (Bergen et al. 2020) to infer the continuous processes. The processes are reconstructed by finding paths through cellular space that minimize transcriptional changes between neighboring cells. The performance of different pseudotime algorithms can be

very variable across datasets. A detailed benchmarking on those pseudotime methods has been performed by Saelens et al. (2019). Since the immune process in the tissue is not clearly studied, how much of the temporary dynamics can be reflected through the transcriptional similarity is still an uncertain question. This should also be addressed systematically in different experimental animal models.

### Spatial organization

While the single-cell sequencing techniques described so far capture molecular profiles of single cells at unprecedented depth, they usually do not confer spatial information of where in the tissue the single cells originate from or which cells they interacted with. This information, however, is critically relevant in the kidney and other organs, as different kidney compartments have variable micro-environment conditions such as sodium and oxygen concentration gradients in the cortex and medulla (Stewart et al. 2020). It is therefore quite likely that immune and tissue cells in different macro- and micro-environments expose different gene regulation, signaling, and activity states. Unsurprisingly, most of the current single-cell kidney data do not capture spatial information, while some recent studies provide insights into the

immune topology of the human kidney (Stewart et al. 2019). The recent development of single-cell spatial transcriptomic technologies, such as MERFISH (multiplexed error-robust fluorescence in situ hybridization) (Moffitt et al. 2018) and STARmap (spatially resolved transcript amplicon readout mapping) (Wang et al. 2018), paves the way to spatial single-cell transcriptomic experiments on renal tissues, to leverage information contained in local cell interactions (Andersson et al. 2020). How to apply these methods to different kidney disease samples will also be of major importance to dissect the renal immune spatial organization.

## Epigenetic landscape of the genome

While the gene expression program is quick to respond to external and internal stimuli per se, epigenetic changes of the chromatin can restrict, expand, or change the repertoire of expression changes a cell can make. Especially in the context of mid- to long-term activation and signaling changes of cell types, such as tissue resident memory cells in the kidney, it is important to profile and understand underlying epigenetic changes. Two methods of choice that allow for single-cell epigenetic profiling are scATAC-seq (single-cell assay for transposase-accessible chromatin using sequencing) (Buenrostro et al. 2015) and scChIP-seq (single-cell chromatin immunoprecipitation followed by sequencing) (Rotem et al. 2015). Especially scATAC-seq is a cost-effective and rather reliable technology for single-cell epigenetic profiling, as exemplified in a scATAC-seq study on murine kidney (Cusanovich et al. 2018). While studies on single-cell epigenetic regulation (of immune cells) in healthy or diseased kidneys are still rare, we expect to see a surge in studies investigating epigenetic mechanisms of gene regulation in conjunction with scRNA-seq data soon.

## Concluding remarks

Single-cell genomics serves as a molecular microscope for the observation of cellular landscapes of different tissues and cell types (Giladi and Amit 2017). The first kidney single-cell atlases have been established by the latest advancements in the field of scRNA-seq and analysis of many millions of cells from renal tissue. Combining simultaneous epitope measurement with gene expression data gives additional power to identify subtypes and states of immune cells. The VDJ-seq technology will provide further insights into lymphocytic clonal expansion and lineages. Ligand–receptor analysis based on single-cell data reveals the interactome across cell populations. We listed several major challenges in the renal immunology research. In the future, we expect to observe spatial single-cell sequencing technology applied to kidney research, potentially augmented by temporal

information via multi-label (life) imaging (Zimmermann et al. 2021). Spatial and clonal information might pave the way to understand localized immune action in its tissue context to grasp mechanisms of renal inflammation and injury. We believe that single-cell techniques will become routine methods to elucidate the mechanisms that underlie kidney inflammation, paving the way for novel treatment strategies.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This article was supported by grants from the *Deutsche Forschungsgemeinschaft*, DFG (SFB 1192 to U.P., S.B., and C.F.K.).

## Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alessandri L, Cordero F, Beccuti M, Arigoni M, Olivero M, Romano G, Rabellino S, Licheri N, De Libero G, Pace L et al (2019) RCASC: reproducible classification analysis of single-cell sequencing data. *Gigascience* 8:1–8
- Alquicira-Hernandez J, Sathe A, Ji HP, Nguyen Q, Powell JE (2019) ScPred: accurate supervised method for cell-type classification from single-cell RNA-seq data. *Genome Biol* 20:264
- Andersson A, Bergensträhle J, Asp M, Bergensträhle L, Jurek A, Fernández Navarro J, and Lundeberg J (2020) Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Commun Biol* 3
- Arazi A, Rao DA, Berthier CC, Davidson A, Liu Y, Hoover PJ, Chicoine A, Eisenhaure TM, Jonsson AH, Li S et al (2019) The immune cell landscape in kidneys of patients with lupus nephritis. *Nat Immunol* 20:902–914
- Arneth BM (2019) Impact of B cells to the pathophysiology of multiple sclerosis. *J Neuroinflammation* 16:128
- Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ (2020) Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* 38:1408–1414
- Braga FAV, Kar G, Berg M, Carpaij OA, Polanski K, Simon LM, Brouwer S, Gomes T, Hesse L, Jiang J et al (2019) Cell states in health and in asthma. *Nat Med* 25
- Brbić M, Zitnik M, Wang S, Pisco AO, Altman RB, Darmanis S, Leskovec J (2020) MARS: discovering novel cell types across heterogeneous single-cell experiments. *Nat Methods* 17:1200–1206
- Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ (2015) Single-cell chromatin

- accessibility reveals principles of regulatory variation. *Nature* 523:486–490
- Butler A, Hoffman P, Smibert P, Papalexi E, Satija R (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 36:411–420
- Cusanovich DA, Hill AJ, Aghamirzaie D, Daza RM, Pliner HA, Berletch JB, Filippova GN, Huang X, Christiansen L, DeWitt WS et al (2018) A single-cell atlas of in vivo mammalian chromatin accessibility. *Cell* 174:1309–1324.e18
- Der E, Suryawanshi H, Morozov P, Kustagi M, Goilav B, Ranabathou S, Izmirly P, Clancy R, Belmont HM, Koenigsberg M et al (2019) Tubular cell and keratinocyte single-cell transcriptomics applied to lupus nephritis reveal type I IFN and fibrosis relevant pathways. *Nat Immunol* 20
- Ding J, Smith SL, Orozco G, Barton A, Eyre S, Martin P (2020) Characterisation of CD4+ T-cell subtypes using single cell RNA sequencing and the impact of cell number and sequencing depth. *Sci Rep* 10:1–11
- Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R (2020) Cell PhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat Protoc* 15:1484–1506
- Giladi A, Amit I (2017) Immunology, one cell at a time. *Nature* 547:27–29
- Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K, Smillie C, Burgin G, Delorey TM, Howitt MR, Katz Y et al (2017) A single-cell survey of the small intestinal epithelium. *Nature* 551:333–339
- Hewitt RJ, and Lloyd CM (2021). Regulation of immune responses by the airway epithelial cell landscape. *Nat. Rev. Immunol.* 1–16.
- Hie B, Bryson B, Berger B (2019) Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat Biotechnol* 37:685–691
- Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Bio-statistics* 8:118–127
- Kitching AR, Anders HJ, Basu, N, Brouwer E, Gordon J, Jayne DR, Kullman J, Lyons PA, Merkel PA, Savage COS et al (2020) ANCA-associated vasculitis. *Nat Rev Dis Prim* 6
- Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh Pru, Raychaudhuri S (2019) Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 16:1289–1296
- Krebs CF, Schmidt T, Riedel JH, Panzer U (2017) T helper type 17 cells in immune-mediated glomerular disease. *Nat Rev Nephrol* 13:647–659
- Krebs CF, Reimers D, Zhao Y, Paust HJ, Bartsch P, Nuñez S, Roseblatt MV, Hellmig M, Kilian C, Borchers A et al (2020a). Pathogen-induced tissue-resident memory T H 17 (TRM 17) cells amplify autoimmune kidney disease (*Science Immunology*).
- Krebs CF, Schlitzer A, Kurts C (2020b) Drawing a single-cell landscape of the human kidney in (pseudo)-space and time. *Kidney Int* 97:842–844
- Krebs CF, Reimers D, Zhao Y, Paust HJ, Bartsch P, Nuñez S, Roseblatt MV, Hellmig M, Kilian C, Borchers A et al (2020c) Pathogen-induced tissue-resident memory TH17 (TRM17) cells amplify autoimmune kidney disease. *Sci Immunol* 5:4163
- Kulkarni A, Anderson AG, Merullo DP, Konopka G (2019) Beyond bulk: a review of single cell transcriptomics methodologies and applications. *Curr Opin Biotechnol* 58:129–136
- Kumar BV, Ma W, Miron M, Granot T, Guyer RS, Carpenter DJ, Senda T, Sun X, Ho SH, Lerner H et al (2017) Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. *Cell Rep* 20:2921–2934
- Li J, Yu C, Ma L, Wang J, Guo G (2020) Comparison of Scanpy-based algorithms to remove the batch effect from single-cell RNA-seq data. *Cell Regen* 9:10–10
- Liao J, Yu Z, Chen Y, Bao M, Zou C, Zhang H, Liu D, Li T, Zhang Q, Li J et al (2020) Single-cell RNA sequencing of human kidney. *Sci Data* 7:4
- Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM et al (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161:1202–1214
- Masopust D, Soerens AG (2019) Tissue-resident T cells and other resident leukocytes. *Annu Rev Immunol* 37:521–546
- McGuire AL, Gabriel S, Tishkoff SA, Wonkam A, Chakravarti A, Furlong EEM, Treutlein B, Meissner A, Chang HY, López-Bigas N et al (2020) The road ahead in genetics and genomics. *Nat Rev Genet* 21:581–596
- Menden K, Marouf M, Oller S, Dalmia A, Magruder DS, Kloiber K, Heutink P, and Bonn S (2020) Deep learning–based cell composition analysis from tissue expression profiles. *Sci Adv* 6, eaba2619
- Mizoguchi A, and Bhan AK (2017) Immunobiology of B cells in inflammatory bowel disease. In *Crohn’s disease and ulcerative colitis: from epidemiology and immunobiology to a rational diagnostic and therapeutic approach: Second Edition*, (Springer International Publishing), pp. 111–116
- Moffitt JR, Bambah-Mukku D, Eichhorn SW, Vaughn E, Shekhar K, Perez JD, Rubinstein ND, Hao J, Regev A, Dulac C et al (2018) Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* 362:792
- Ntranos V, Yi L, Melsted P, Pachter L (2019) A discriminative learning approach to differential expression analysis for single-cell RNA-seq. *Nat Methods* 16:163–166
- O’Sullivan ED, Mylonas KJ, Hughes J, Ferenbach DA (2019) Complementary roles for single-nucleus and single-cell RNA sequencing in kidney disease research. *J Am Soc Nephrol* 30:712–713
- Oller-Moreno S, Kloiber K, Machart P, and Bonn S (2021) Algorithmic advances in machine learning for single cell expression analysis. *Curr Opin Syst Biol.* <https://doi.org/10.1016/j.coisb.2021.02.002>
- Papalexi E, Satija R (2018) Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol* 18:35–45
- Pappalardo JL, Zhang L, Pecsok MK, Perlman K, Zografou C, Raddassi K, Abulaban A, Krishnaswamy S, Antel J, van Dijk D et al (2020) Transcriptomic and clonal characterization of T cells in the human central nervous system. *Sci Immunol* 5:8786
- Park J, Shrestha R, Qiu C, Kondo A, Huang S, Werth M, Li M, Barasch J, Suszták K (2018) Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* 360:758–763
- Petegrosso R, Li Z, Kuang R (2020) Machine learning and statistical methods for clustering single-cell RNA-sequencing data. *Brief Bioinform* 21:1209–1223
- Peters B, Nielsen M, Sette A (2020) T cell epitope predictions. *Annu Rev Immunol* 38:123–145
- Potter SS (2018) Single-cell RNA sequencing for the study of development, physiology and disease. *Nat Rev Nephrol* 14:479–492
- Rahman RU, Liebhoff AM, Bansal V, Fiosins M, Rajput A, Sattar A, Magruder DS, Madan S, Sun T, Gautam A et al (2020) SEAwEB: the small RNA expression Atlas web application. *Nucleic Acids Res* 48:D204–D219
- Rotem A, Ram O, Shores N, Sperling RA, Goren A, Weitz DA, Bernstein BE (2015) Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol* 33:1165–1172
- Saelens W, Cannoodt R, Todorov H, Saeys Y (2019) A comparison of single-cell trajectory inference methods. *Nat Biotechnol* 37:547–554
- Schrezenmeier E, Jayne D, and Dörner T (2018) Targeting B cells and plasma cells in glomerular diseases: Translational perspectives. *J Am Soc Nephrol* 29, 741–758

- De Simone M, Rossetti G, Pagani M (2018) Single cell T cell receptor sequencing: techniques and future challenges. *Front Immunol* 9:1638
- Sivakamasundari V, Bolisetty M, Sivajothi S, Bessonett S, Ruan D, and Robson P (2017). Comprehensive cell type specific transcriptomics of the human kidney. *BioRxiv* 238063.
- Smillie CS, Biton M, Ordovas-Montanes J, Sullivan KM, Burgin G, Graham DB, Herbst RH, Rogel N, Slyper M, Waldman J et al (2019) Intra- and Inter-cellular rewiring of the human colon during ulcerative colitis. *Cell* 178:714–730.e22
- Smyth GK, Speed T (2003) Normalization of cDNA microarray data. *Methods* 31:265–273
- Stewart BJ, Ferdinand JR, Young MD, Mitchell TJ, Loudon KW, Riding AM, Richoz N, Frazer GL, Staniforth JUL, Braga FAV et al (2019) Spatiotemporal immune zonation of the human kidney. *Science* 365:1461–1466
- Stewart BJ, Ferdinand JR, Clatworthy MR (2020) Using single-cell technologies to map the human immune system — implications for nephrology. *Nat Rev Nephrol* 16:112–128
- Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R, Smibert P (2017) Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 14:865–868
- Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, Purdom E, Dudoit S (2018) Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* 19:1–16
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R (2019) Comprehensive integration of single-cell data. *Cell* 177:1888–1902.e21
- Stubbington MJT, Lönnberg T, Proserpio V, Clare S, Speak AO, Dougan G, Teichmann SA (2016) T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods* 13:329–332
- Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A et al (2009) mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 6:377–382
- Tran HTN, Ang KS, Chevrier M, Zhang X, Lee NYS, Goh M, Chen J (2020) A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol* 21:12
- Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 32:381–386
- Tu AA, Gierahn TM, Monian B, Morgan DM, Mehta NK, Ruitter B, Shreffler WG, Shalek AK, Love JC (2019) TCR sequencing paired with massively parallel 3' RNA-seq reveals clonotypic T cell signatures. *Nat Immunol* 20:1692–1699
- Wang T, Johnson TS, Shao W, Lu Z, Helm BR, Zhang J, Huang K (2019) BERMUDA: a novel deep transfer learning method for single-cell RNA sequencing batch correction reveals hidden high-resolution cellular subtypes. *Genome Biol* 20:165
- Wang X, Allen WE, Wright MA, Sylwestrak EL, Samusik N, Vesuna S, Evans K, Liu C, Ramakrishnan C, Liu J et al (2018) Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* 361:380
- Welch JD, Kozareva V, Ferreira A, Vanderburg C, Martin C, Macosko EZ (2019) Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* 177:1873–1887.e17
- Wilson PC, Humphreys BD (2019) Single-cell genomics and gene editing: implications for nephrology. *Nat Rev Nephrol* 15:63–64
- Wilson PC, Wu H, Kirita Y, Uchimura K, Ledru N, Rennke HG, Welling PA, Waikar SS, Humphreys BD (2019) The single-cell transcriptomic landscape of early human diabetic nephropathy. *Proc Natl Acad Sci U S A* 116:19619–19625
- Wolf FA, Hamey FK, Plass M, Solana J, Dahlin JS, Göttgens B, Rajewsky N, Simon L, Theis FJ (2019) PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol* 20:1–9
- Wu H, Humphreys BD (2017) The promise of single-cell RNA sequencing for kidney disease investigation. *Kidney Int* 92:1334–1342
- Wu H, Malone AF, Donnelly EL, Kirita Y, Uchimura K, Ramakrishnan SM, Gaut JP, Humphreys BD (2018) Single-cell transcriptomics of a human kidney allograft biopsy specimen defines a diverse inflammatory response. *J Am Soc Nephrol* 29:2069–2080
- Yofe I, Dahan R, Amit I (2020) Single-cell genomic approaches for developing the next generation of immunotherapies. *Nat Med* 26:171–177
- Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, Collord G, Botting RA, Popescu DM, Loudon KW et al (2018) Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* 361:594–599
- Zemmour D, Kiner E, Benoist C (2020) CD4+ teff cell heterogeneity: the perspective from single-cell transcriptomics. *Curr Opin Immunol* 63:61–67
- Zhao Y, Kilian C, Turner JE, Bosurgi L, Roedel K, Bartsch P, Gnirck AC, Cortesi F, Schultheiß C, Hellmig M et al (2021) Clonal expansion and activation of tissue-resident memory-like Th17 cells expressing GM-CSF in the lungs of severe COVID-19 patients. *Sci Immunol* 6, eabf6692
- Zheng J, Wang K (2019) Emerging deep learning methods for single-cell RNA-seq data analysis. *Quant Biol* 7:247–254
- Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, Kang B, Hu R, Huang JY, Zhang Q et al (2017) Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. *Cell* 169:1342–1356.e16
- Zheng Y, Lu P, Deng Y, Wen L, Wang Y, Ma X, Wang Z, Wu L, Hong Q, Duan S et al (2020). Single-cell transcriptomics reveal immune mechanisms of the onset and progression of IgA nephropathy. *Cell Rep* 33, 108525
- Zimmermann M, Bonn S, Puelles VG, Zimmermann AM, Klaus M, Wong MN, Thebille A, Wulf S, Wiech T, Panzer U et al (2021) Deep learning-based molecular morphometrics for kidney biopsies. *JCI Insight*. <https://doi.org/10.1172/jci.insight.144779>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## Chapter 3: Publication in *Science Immunology*, August 2020

### Pathogen-induced tissue-resident memory T<sub>H</sub>17 (T<sub>RM</sub>17) cells amplify autoimmune kidney disease

#### Specific contributions:

The following research article has been published in *Science Immunology*. I am the second author of the paper. I processed and analyzed the single cell RNA-seq and CITE-seq data of T cells sorted from human kidney biopsy and paired blood samples. These results are explicitly displayed in Figures: 1B, C, D, E, G; 2B, C, J, K, L and supplementary Figures: 1B; 2A, B, C; 3A, B, C; 4A, B, C, D. I also processed and analyzed the single cell RNA sequencing data of sorted Th17 cells from mouse kidneys. This part is represented in Figures: 8C, D; supplementary Figures: 8A, B, C; and supplementary Table 4. I contributed to the manuscripts proofreading before the first submission and during the revisions.

I, Stefan Bonn, agree with the above statements as the direct supervisor.

Signature

,

Date & Place

15.11.2021 Hamburg



## AUTOIMMUNITY

Pathogen-induced tissue-resident memory T<sub>H</sub>17 (T<sub>RM</sub>17) cells amplify autoimmune kidney disease

Christian F. Krebs<sup>1,2,\*†</sup>, Daniel Reimers<sup>3†</sup>, Yu Zhao<sup>4</sup>, Hans-Joachim Paust<sup>1</sup>, Patricia Bartsch<sup>1</sup>, Sarah Nuñez<sup>5</sup>, Mariana V. Roseblatt<sup>6</sup>, Malte Hellmig<sup>1</sup>, Christoph Kilian<sup>1</sup>, Alina Borchers<sup>1</sup>, Leon U. B. Enk<sup>1</sup>, Michael Zinke<sup>1</sup>, Martina Becker<sup>7</sup>, Joanna Schmid<sup>3</sup>, Stefanie Klinge<sup>3</sup>, Milagros N. Wong<sup>7</sup>, Victor G. Puelles<sup>7,8</sup>, Constantin Schmidt<sup>3</sup>, Tabea Bertram<sup>3</sup>, Natascha Stumpf<sup>9</sup>, Elion Hoxha<sup>7</sup>, Catherine Meyer-Schwesinger<sup>10</sup>, Maja T. Lindenmeyer<sup>7</sup>, Clemens D. Cohen<sup>7,11</sup>, Michael Rink<sup>12</sup>, Christian Kurts<sup>9</sup>, Sören Franzenburg<sup>13</sup>, Friedrich Koch-Nolte<sup>2,3</sup>, Jan-Eric Turner<sup>2,7</sup>, Jan-Hendrik Riedel<sup>1</sup>, Samuel Huber<sup>2,14</sup>, Nicola Gagliani<sup>2,14,15</sup>, Tobias B. Huber<sup>2,7</sup>, Thorsten Wiech<sup>6</sup>, Holger Rohde<sup>17</sup>, Maria Rosa Bono<sup>18</sup>, Stefan Bonn<sup>2,4,19,20</sup>, Ulf Panzer<sup>1,2†</sup>, Hans-Willi Mittrücker<sup>2,3,\*†</sup>

Although it is well established that microbial infections predispose to autoimmune diseases, the underlying mechanisms remain poorly understood. After infection, tissue-resident memory T (T<sub>RM</sub>) cells persist in peripheral organs and provide immune protection against reinfection. However, whether T<sub>RM</sub> cells participate in responses unrelated to the primary infection, such as autoimmune inflammation, is unknown. By using high-dimensional single-cell analysis, we identified CD4<sup>+</sup> T<sub>RM</sub> cells with a T<sub>H</sub>17 signature (termed T<sub>RM</sub>17 cells) in kidneys of patients with ANCA-associated glomerulonephritis. Experimental models demonstrated that renal T<sub>RM</sub>17 cells were induced by pathogens infecting the kidney, such as *Staphylococcus aureus*, *Candida albicans*, and uropathogenic *Escherichia coli*, and persisted after the clearance of infections. Upon induction of experimental glomerulonephritis, these kidney T<sub>RM</sub>17 cells rapidly responded to local proinflammatory cytokines by producing IL-17A and thereby exacerbate renal pathology. Thus, our data show that pathogen-induced T<sub>RM</sub>17 cells have a previously unrecognized function in aggravating autoimmune disease.

## INTRODUCTION

Infections have been associated with the development and aggravation of autoimmune diseases (1), such as antineutrophil cytoplasmic antibody (ANCA)-associated glomerulonephritis (GN) (2), multiple sclerosis (3), or inflammatory bowel diseases (IBDs) (4). Molecular mimicry, enhanced presentation of autoantigens during infection, and bystander activation of T cells are all considered to be underlying causes. However, the cellular and molecular mechanisms of how infections create a permissive environment for autoimmune diseases

in sites like the kidney, brain, or gut are far from being fully understood (1, 5).

GN represents a heterogeneous group of renal autoimmune diseases that are characterized by a pathogenic immune response against renal autoantigens or by manifestations of systemic autoimmunity in the kidney (6). GNs are a major cause of chronic kidney diseases, accounting for about 10% of all patients on dialysis. The most severe form, rapidly progressive or crescentic GN, is morphologically defined by extensive glomerular crescent formation, which is formed by proliferation of glomerular parietal epithelial cells and infiltrating leukocytes. This is typically accompanied by tubulointerstitial inflammation and immune cell infiltration. Clinically, crescentic GN is characterized by a nephritic syndrome (hematuria, proteinuria, and often arterial hypertension) rapidly progressing to end-stage renal disease (7, 8). Crescentic GN is a typical feature of renal involvement in ANCA-associated vasculitis and of anti-glomerular basement membrane GN (anti-GBM-GN) (6). There is strong evidence for a major role of CD4<sup>+</sup> T cells in crescentic GN. GNs are associated with distinct human leukocyte antigen (HLA) class II haplotypes (7, 9, 10), renal biopsies from patients with crescentic GN show substantial CD4<sup>+</sup> T cell infiltration (11), and CD4<sup>+</sup> T cells reactive to autoantibody targets can be identified during active disease (7). Rodent crescentic GN models confirm the role of CD4<sup>+</sup> T cells and particularly T<sub>H</sub>17 cells in disease development (12–14). These studies indicate that T cells promote renal damage predominantly by local production of cytokines, such as interleukin-17A (IL-17A), IL-17F, and interferon- $\gamma$  (IFN- $\gamma$ ) (6, 8, 15). For example, CD4<sup>+</sup> T<sub>H</sub>17 cell-derived IL-17A induces the renal expression of CXCL1 and CXCL5 and thereby drives the infiltration of neutrophils and other leukocyte subsets, which promote kidney pathology in crescentic GN (16).

<sup>1</sup>III. Department of Medicine, Division of Translational Immunology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>2</sup>Hamburg Center for Translational Immunology (HCTI), University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>3</sup>Institute for Immunology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>4</sup>Institute of Medical Systems Biology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>5</sup>Fundacion Ciencia Vida, Santiago, Chile. <sup>6</sup>Faculty of Medicine and Science, Universidad San Sebastian, Santiago, Chile. <sup>7</sup>III. Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>8</sup>Department of Nephrology, Monash Health, and Center for Inflammatory Diseases, Monash University, Melbourne, VIC, Australia. <sup>9</sup>Institutes of Molecular Medicine and Experimental Immunology (IMMEI), Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany. <sup>10</sup>Institute for Cellular and Integrative Physiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>11</sup>Nephrological Center, Medical Clinic and Policlinic IV, University of Munich, Munich, Germany. <sup>12</sup>Department of Urology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>13</sup>Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany. <sup>14</sup>Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>15</sup>Immunology and Allergy Unit, Department of Medicine, Solna, Karolinska Institute and University Hospital, 17176 Stockholm, Sweden. <sup>16</sup>Institute for Pathology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>17</sup>Institute for Medical Microbiology, Virology and Hygiene, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>18</sup>Department of Biology, Faculty of Sciences, Universidad de Chile, Santiago, Chile. <sup>19</sup>German Center for Neurodegenerative Diseases, Tübingen, Germany. <sup>20</sup>Center for Biomedical AI, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

\*Corresponding author. Email: c.krebs@uke.de (C.F.K.); h.mittrucker@uke.de (H.-W.M.)

†These authors contributed equally to this work.

Tissue-resident memory T ( $T_{RM}$ ) cells are a recently found distinct memory T cell subset mainly located at barrier sites such as skin or mucosal tissues.  $T_{RM}$  cells are defined by their exclusion from circulation and can act as local sentinels for pathogens that have been encountered before (17–20). In terms of phenotype and expression profile,  $T_{RM}$  cells differ substantially from other memory T cell subsets.  $T_{RM}$  cells are characterized by high expression of CD69 and down-regulation of the sphingosine 1 phosphate receptor and the transcription factor Krueppel-like factor 2 (KLF2), which is essential for their long-term tissue residency (21–23). Most studies so far focused on CD8<sup>+</sup>  $T_{RM}$  cells and their relevance for immediate protection against pathogens upon local reinfection. In contrast, the role of CD4<sup>+</sup>  $T_{RM}$  cells in peripheral tissues is less well characterized (24). Moreover, the implications of  $T_{RM}$  cells for the development and progression of autoimmunity, such as GNs, are largely unknown.

Here, we analyzed the composition of human renal T cells in biopsies from patients with ANCA-GN and from healthy kidney tissue from tumor nephrectomies and identified different T cell subsets including CD4<sup>+</sup>  $T_{RM}$  cells with  $T_{H1}$  or  $T_{H17}$  expression signatures. On the basis of this finding, we hypothesized that aberrant activation of pathogen-induced  $T_{RM}$  cells promotes renal pathology in patients with GN. To test this, we combined mouse infection models with models for GN and provide evidence that infection-induced  $T_{RM17}$  cells can rapidly react to local inflammatory cytokines and aggravate renal autoimmune disease. These results support a new concept for the predisposing role of microbial infections in aggravating  $T_{H17}$  cell-driven autoimmune diseases.

## RESULTS

### Human renal tissue displays high abundance of CD4<sup>+</sup> $T_{RM}$ cells

T cells are present in the kidney under homeostatic conditions, but there is limited information on the composition and function of these cells (25). To elucidate their phenotype and potential function, we performed single-cell RNA sequencing (scRNA-seq) (26) in combination with epitope measurement using barcoded antibodies [cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq)] (27, 28) of CD3<sup>+</sup> T cells from the human kidney (Fig. 1A). Because biopsies from individuals without kidney diseases were not available, we used tissue from the healthy parts of kidneys after tumor nephrectomy (table S1). We identified 5905 cells and 16,148 genes from three renal samples (Fig. 1B and figs. S1 and S2). Unsupervised clustering analysis revealed 12 distinct populations for renal T cells. CITE-seq considerably improved the analysis by adding important information about CD4, CD8, and CD69 surface expression (Fig. 1C and fig. S2). Clusters C1 to C3 in kidneys (representing 7.5, 11.8, and 14.0% of analyzed cells, respectively) were composed of CD4<sup>+</sup> CD69<sup>high</sup> CD45RA<sup>low</sup> CCR7<sup>low</sup> T cells. CD69 expression on CD4<sup>+</sup> T cells is indicative of a  $T_{RM}$  phenotype of these cells. Comparison of expression profiles of T cells from our datasets with a published signature gene set for human CD4<sup>+</sup>  $T_{RM}$  cells (21) revealed a  $T_{RM}$  profile of renal CD4<sup>+</sup> CD69<sup>+</sup> T cells but not of renal CD4<sup>+</sup> CD69<sup>low</sup> T cells or blood T cells (Fig. 1D). The expression profiles of CD69<sup>high</sup> clusters (C1 to C3) included genes associated with  $T_{RM}$  cells, e.g., high expression of *CD69*, *RGS1*, or *CXCR6* and low expression of *KLF2* or *SELL* (Fig. 1E) (21). Flow cytometry confirmed the presence of renal CD45RA<sup>−</sup> CCR7<sup>−</sup> CD69<sup>+</sup> CD4<sup>+</sup>  $T_{RM}$  cells in the kidney (Fig. 1F).

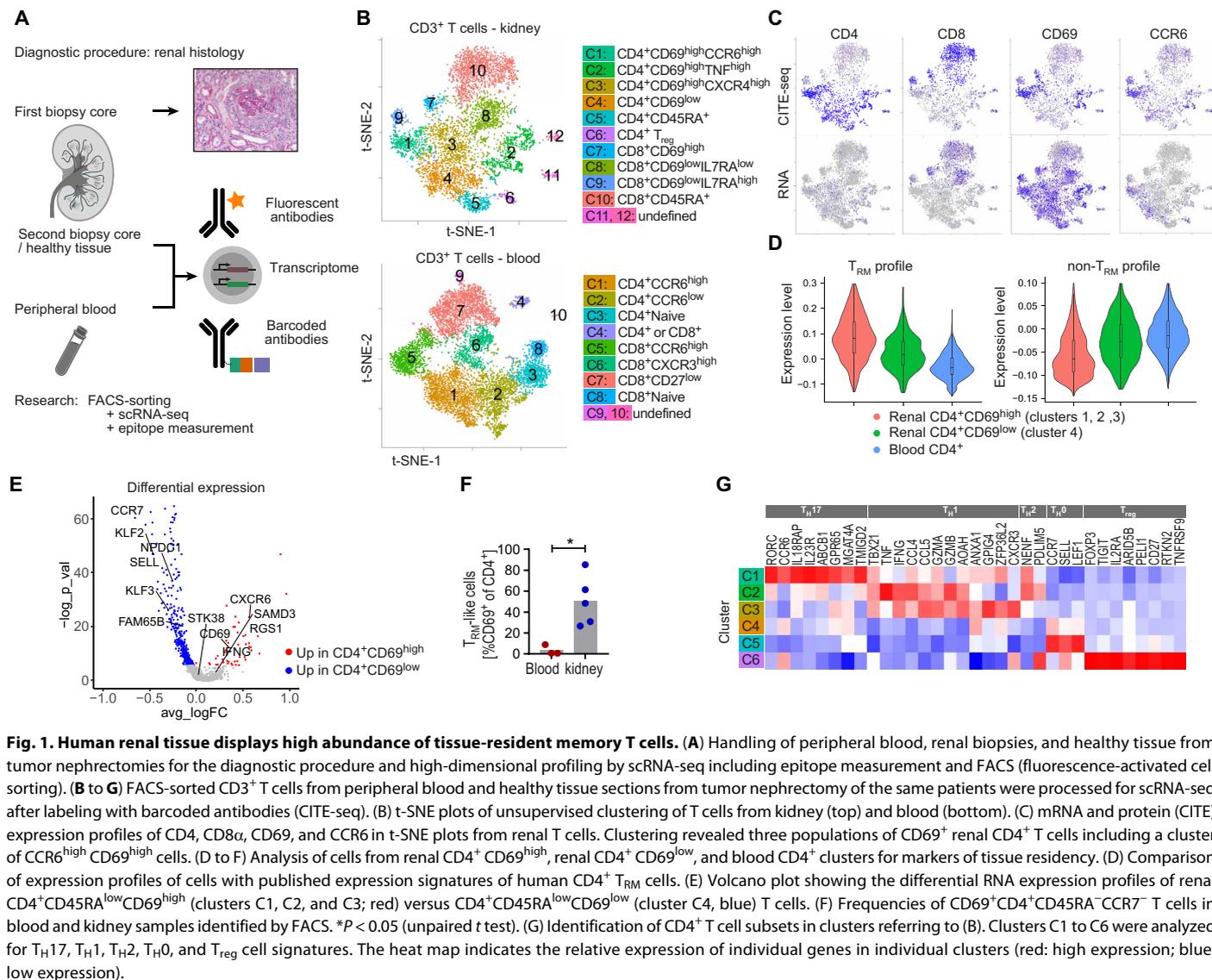
The clusters of CD4<sup>+</sup> T cells from renal tissue (C1 to C6) were analyzed for the expression of sets of signature genes of  $T_{H17}$ ,  $T_{H1}$ ,  $T_{H2}$ ,  $T_{H0}$ , and regulatory T ( $T_{reg}$ ) cells (Fig. 1G) (29). CD4<sup>+</sup> CD69<sup>high</sup> CCR6<sup>high</sup> T cells of C1 presented with a  $T_{H17}$  signature, cells in C2 expressed signature genes of  $T_{H1}$  and  $T_{H2}$  cells, and cells in C3 expressed signature genes of  $T_{H1}$  cells. CD45RA<sup>+</sup> cells in C5 expressed *CCR7*, *SELL*, and *LEF1* and most likely presented naïve cells, and cells in C6 were  $T_{reg}$  cells. Last, we detected IL-17A and IFN- $\gamma$  production in a subset of these CD4<sup>+</sup> T cells after phorbol 12-myristate 13-acetate/ionomycin stimulation (fig. S2D).

In conclusion, our results demonstrate that, under homeostatic conditions, human kidneys harbor a substantial population of CD4<sup>+</sup>  $T_{RM}$  cells, including major subsets with a  $T_{H1}$  and  $T_{H17}$  cell phenotype ( $T_{RM1}$  and  $T_{RM17}$  cells).

### $T_{RM}$ cells are enriched in patients with GN

To address a potential contribution of  $T_{RM}$  cells to autoimmune kidney diseases, we analyzed renal expression levels of CD4<sup>+</sup>  $T_{RM}$  cell-associated genes in different glomerular diseases using datasets from the European Renal cDNA Bank database (30). Diseases included minimal change disease, systemic lupus erythematosus, IgA (immunoglobulin A) nephritis, focal segmental glomerulosclerosis, membranous GN, and ANCA-associated GN. Expression levels of the  $T_{RM}$  cell-associated genes were highest in the glomeruli of patients with ANCA-associated GN (Fig. 2A). Renal biopsies from patients with active ANCA-GN were further studied by scRNA-seq and CITE-seq. To this end, 6829 cells and 15,360 genes from six samples could be analyzed (table S2). T cells from the kidney are distributed to 10 clusters that differed from those observed in matched blood from these patients (Fig. 2B and figs. S3 and S4).  $T_{RM}$  cell-associated genes were enriched in renal CD69<sup>+</sup> T cells of clusters C1 to C3, representing 3.4, 17.4, and 13.4% of analyzed cells (fig. S4, C and D). In line with observations under control conditions, signature gene sets for  $T_H$  cell subtypes could be associated with distinct clusters of CD69<sup>+</sup> T cells.  $T_{H17}$  signature genes were enriched in C1, and  $T_{H1}$  and  $T_{H2}$  signature genes were enriched in C2 and C3 (Fig. 2C). CD4<sup>+</sup> CD69<sup>+</sup> cells were also detected by immunofluorescence histology in the renal cortex of patients with ANCA-GN (Fig. 2D). Flow cytometry confirmed the presence of renal CD45RA<sup>−</sup> CCR7<sup>−</sup> CD69<sup>+</sup> CD4<sup>+</sup>  $T_{RM}$  cells (Fig. 2, E and F). We observed limited expression of CD103 on these cells, consistent with a restriction of CD103 to mucosal  $T_{RM}$  cells (31). To further characterize these renal T cells, we made use of chemokine receptor staining. CD69<sup>+</sup> CD4<sup>+</sup>  $T_{RM}$  cells included CXCR3<sup>+</sup> and CCR6<sup>+</sup> cells representing  $T_{H1}$  and  $T_{H17}$  cells, respectively, as well as a large proportion of CXCR3<sup>+</sup> CCR6<sup>+</sup> cells, which have been described as  $T_{H1}^*/exT_{H17}$  cells (32) (Fig. 2, G and H).

In comparison with healthy renal tissue, biopsies from patients with ANCA-GN contained substantially increased numbers of CD45RA<sup>−</sup> CCR7<sup>−</sup> CD69<sup>+</sup> CD4<sup>+</sup>  $T_{RM}$  cells (Fig. 2I). To compare the expression profiles of T cells from renal biopsies of ANCA-GN and control biopsies derived from nephrectomy patients, scRNA-seq and CITE-seq datasets from both groups were combined, and unsupervised clustering was performed (Fig. 2J). T cells separated into 11 clusters, and T cells from both groups of samples were found in all clusters (Fig. 2K). CD69<sup>+</sup> CD4<sup>+</sup> T cells were allocated to C1 (CD4<sup>+</sup> CD69<sup>high</sup> CCR6<sup>high</sup>) and C2 (CD4<sup>+</sup> CD69<sup>high</sup>). For pathway analysis of differentially expressed genes of CD69<sup>+</sup> CD4<sup>+</sup> T cells (C1 and C2) from control and ANCA-GN samples, we used a web-based



**Fig. 1. Human renal tissue displays high abundance of tissue-resident memory T cells.** (A) Handling of peripheral blood, renal biopsies, and healthy tissue from tumor nephrectomies for the diagnostic procedure and high-dimensional profiling by scRNA-seq including epitope measurement and FACS (fluorescence-activated cell sorting). (B to G) FACS-sorted CD3<sup>+</sup> T cells from peripheral blood and healthy tissue sections from tumor nephrectomy of the same patients were processed for scRNA-seq after labeling with barcoded antibodies (CITE-seq). (B) t-SNE plots of unsupervised clustering of T cells from kidney (top) and blood (bottom). (C) mRNA and protein (CITE) expression profiles of CD4, CD8, CD69, and CCR6 in t-SNE plots from renal T cells. Clustering revealed three populations of CD69<sup>+</sup> renal CD4<sup>+</sup> T cells including a cluster of CCR6<sup>high</sup> CD69<sup>high</sup> cells. (D to F) Analysis of cells from renal CD4<sup>+</sup>CD69<sup>high</sup>, renal CD4<sup>+</sup>CD69<sup>low</sup>, and blood CD4<sup>+</sup> clusters for markers of tissue residency. (D) Comparison of expression profiles of cells with published expression signatures of human CD4<sup>+</sup> T<sub>RM</sub> cells. (E) Volcano plot showing the differential RNA expression profiles of renal CD4<sup>+</sup>CD45RA<sup>low</sup>CD69<sup>high</sup> (clusters C1, C2, and C3; red) versus CD4<sup>+</sup>CD45RA<sup>low</sup>CD69<sup>low</sup> (cluster C4, blue) T cells. (F) Frequencies of CD69<sup>+</sup>CD4<sup>+</sup>CD45RA<sup>-</sup>CCR7<sup>-</sup> T cells in blood and kidney samples identified by FACS. \*P < 0.05 (unpaired t test). (G) Identification of CD4<sup>+</sup> T cell subsets in clusters referring to (B). Clusters C1 to C6 were analyzed for T<sub>H</sub>17, T<sub>H</sub>1, T<sub>H</sub>2, T<sub>H</sub>0, and T<sub>reg</sub> cell signatures. The heat map indicates the relative expression of individual genes in individual clusters (red: high expression; blue: low expression).

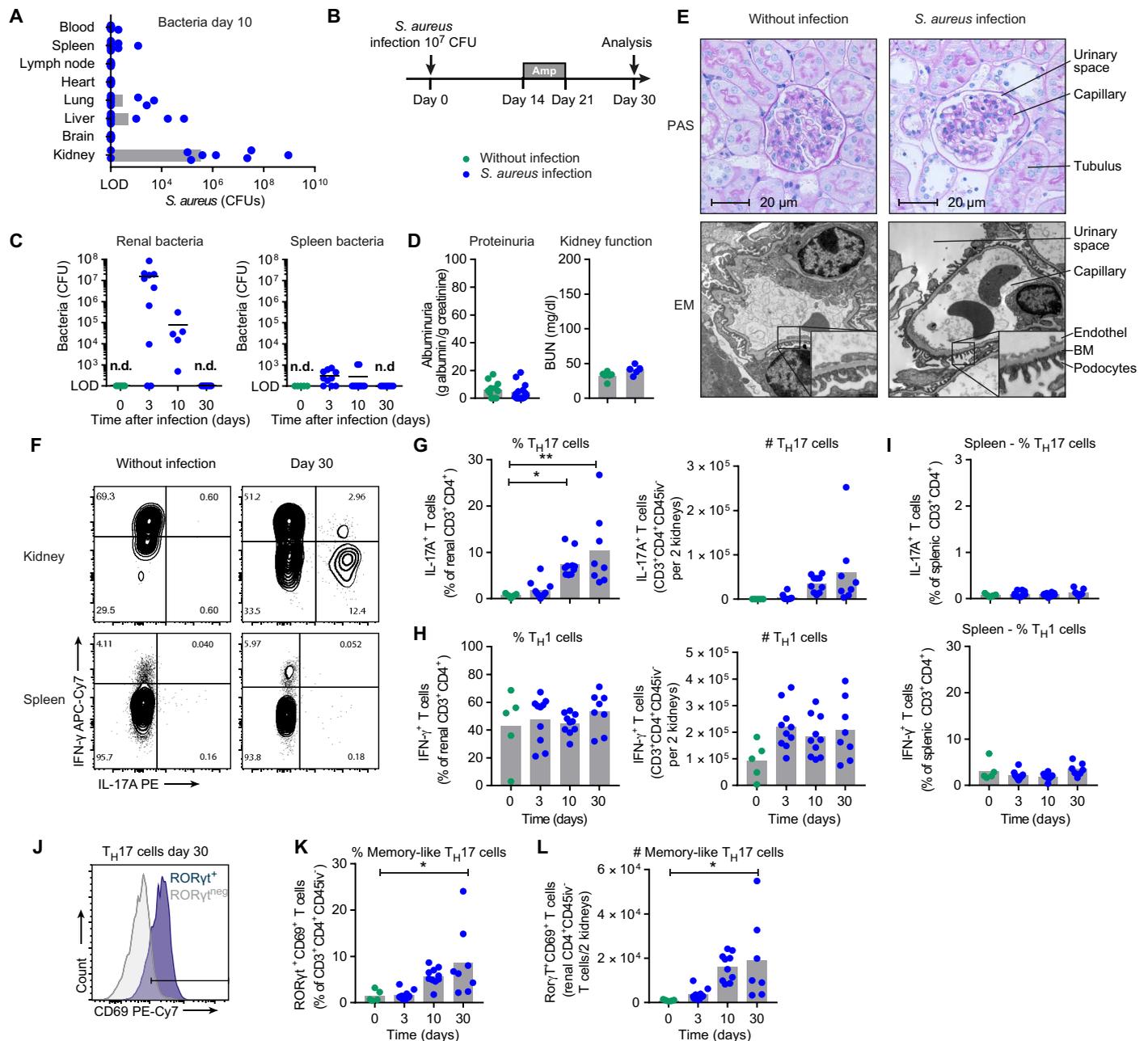
enrichment analysis tool (Enrichr) (33). This revealed up-regulation of genes involved in proliferation, cell activation, and cytokine signaling pathways in T cells from patients with ANCA-GN (Fig. 2L). Scoring of renal biopsies from patients with active ANCA-GN further revealed that increased numbers of CD69<sup>+</sup> mononuclear cells correlated with reduced renal filtration rate, as indicated by high serum creatinine concentration (Fig. 2M and table S3). Together, these observations show that kidneys from ANCA-GN show massive accumulation of CD4<sup>+</sup> T<sub>RM</sub> cells including cells with the phenotype and expression signature of T<sub>H</sub>1 and T<sub>H</sub>17 cells compared with control samples, and the number of CD69<sup>+</sup> renal cells in acute disease correlates directly with impairment of kidney function.

**Systemic *Staphylococcus aureus* infection induces renal CD4<sup>+</sup> T<sub>H</sub>17 cells**

To test whether CD4<sup>+</sup> T<sub>RM</sub> cells influence renal autoimmune disease, we combined mouse infection models with models for crescentic GN. Previous findings from mouse models and human studies suggest that T<sub>H</sub>17 cells play a role in crescentic GN (8). Kidneys of healthy

specific pathogen-free mice contain only low frequencies of T<sub>H</sub>17 cells (11). *S. aureus* infection induces and is controlled by T<sub>H</sub>17 cell responses (32, 34) and has been associated with relapses of granulomatosis with polyangiitis, a form of ANCA-associated vasculitis (35). Intravenous infection of mice with *S. aureus* induced high bacterial titers transiently in the kidney but not in other lymphoid and non-lymphoid tissues (Fig. 3A). Mice were treated with ampicillin to curb chronic infection (Fig. 3, B and C), and mice had completely recovered from infection and showed normal renal function and morphology by day 30 (Fig. 3, D and E). Recovered mice had normal urinary albumin and blood urea concentrations and showed normal kidney pathology. No electron-dense depositions along the GBM were detected, excluding postinfectious GN that is observed occasionally after bacterial infections (Fig. 3E). Renal T cells were isolated after intravenous injection of fluorochrome-labeled anti-CD45 monoclonal antibody (mAb) to identify vascular cells and polyclonally stimulated to induce cytokine expression. Kidneys of naïve mice contained only marginal frequencies of intraparenchymal (CD45<sup>iv</sup><sup>-</sup>) IL-17A-secreting CD4<sup>+</sup> T<sub>H</sub>17 cells but a substantial population of IFN-γ-secreting T<sub>H</sub>1





**Fig. 3. Systemic *S. aureus* infection results in high bacterial titers in the kidney and induces renal CD4<sup>+</sup> TH17 cells.** (A) C57BL/6 mice were infected with *S. aureus*. To analyze the organ involvement of the infection, different lymphoid and nonlymphoid organs as indicated were analyzed for *S. aureus* [limit of detection (LOD)]. Data are representative of three independent experiments. (B to E) Bacteria were cleared with oral ampicillin and kidneys were analyzed at the indicated time points. (C) Bacterial titers in kidney and spleen. (D) At day 30, renal function with normal urinary protein and regular blood urea nitrogen (BUN) concentrations were detected. (E) Regular glomerular morphology (PAS staining) and absence of electron-dense deposits along the glomerular capillaries [electron microscopy (EM)] at day 30. (F) Flow cytometry of stimulated renal and splenic leukocytes at indicated time points after infection. CD45iv<sup>-</sup>CD3<sup>+</sup>CD4<sup>+</sup>T cells were assessed for IFN-γ and IL-17A production. Quantification of relative and absolute IL-17A<sup>+</sup> and IFN-γ<sup>+</sup> CD45iv<sup>-</sup>CD4<sup>+</sup>T cells in the kidney (G and H) and spleen (I). (J to L) CD69 and RORγt expression of renal CD45iv<sup>-</sup>T cells. Data are representative of three to four independent experiments. \**P* < 0.05 and \*\**P* < 0.01 [one-way analysis of variance (ANOVA) with multiple comparison test].

cells. *S. aureus* infection caused increased proportions of IL-17A<sup>+</sup> CD4<sup>+</sup> T cells as well as total numbers of these cells in the kidney, which remained elevated after clearance of *S. aureus* (Fig. 3, F and G, and fig. S5, A and B). In contrast, increase of frequencies and numbers of renal IFN-γ<sup>+</sup> CD4<sup>+</sup> TH1 cells was less evident and infection resulted in only marginal changes of splenic TH1 and TH17 cells

(Fig. 3, H and I). Ex vivo staining for RORγt confirmed the accumulation of TH17 cells in kidneys from infected mice (fig. S5C). At day 30, the majority of renal RORγt<sup>+</sup> CD4<sup>+</sup> T cells had a CD69<sup>+</sup> phenotype of T<sub>RM</sub> cells (Fig. 3, J to L, and fig. S5C). Together, these findings indicate that infection with *S. aureus* induces renal TH17 cells with the phenotype resembling T<sub>RM</sub> cells.

### ***S. aureus*-induced T<sub>H</sub>17 cells persist in the kidney**

T<sub>RM</sub> cells are defined by their exclusion from recirculation. To test whether *S. aureus*-induced RORγt<sup>+</sup>CD69<sup>+</sup>CD4<sup>+</sup> T cells persisted in the kidney, we performed parabiosis of CD45.1 mice after *S. aureus* infection with CD45.2 control mice (Fig. 4A). After 4 weeks of parabiosis, splenic CD4<sup>+</sup> T cells showed equilibrium between host- and donor-derived cells. In contrast, CD69<sup>+</sup>RORγt<sup>+</sup> CD4<sup>+</sup> T cells in the kidney of the previously infected mice were still host-derived (Fig. 4, B to E). We also used *Kaede* transgenic mice that ubiquitously express the *Kaede* protein, which is permanently converted from green to red fluorescence by light exposure (Fig. 4F) (11). *Kaede* transgenic mice were infected with *S. aureus*, and after clearance of infection, one kidney was exposed to light to induce photoconversion (*Kaede* red<sup>+</sup>) of renal CD4<sup>+</sup> T cells (Fig. 4, G to K). Because of the limited depth of light penetration, only a fraction of renal cells can be photoconverted. Four weeks later, the light-exposed kidney still contained 10% *Kaede* red<sup>+</sup> CD4<sup>+</sup> T cells, indicating that these cells had persisted in the tissue. Converted CD4<sup>+</sup> T cells contained an enlarged fraction of up to 60% RORγt<sup>+</sup> cells, and in contrast to *Kaede* green<sup>+</sup> RORγt<sup>+</sup> cells, the vast majority of *Kaede* red<sup>+</sup> RORγt<sup>+</sup> cells were CD69<sup>+</sup>. Thus, CD69<sup>+</sup> T<sub>H</sub>17 cells represented a major population of persisting cells. Photoconverted cells were not detected in the spleen or the contralateral kidney. These experiments indicate that *S. aureus* infection induces bona fide T<sub>RM</sub> cells with a T<sub>H</sub>17 signature in the kidney, referred to as T<sub>RM</sub>17 cells.

### ***S. aureus* aggravates crescentic GN and increases the renal T<sub>H</sub>17 cell response**

To investigate whether infection-induced T<sub>RM</sub>17 cells can amplify renal autoimmune disease, we used an established model of crescentic GN (6). After passive transfer, anti-GBM sheep Igs bind within the glomerulus and provoke a T cell response against the deposited antigen. This response results in T<sub>H</sub>1 cell- and T<sub>H</sub>17 cell-driven formation of glomerular crescents, tubulointerstitial damage, and loss of renal function, resembling multiple aspects of human crescentic GN (6, 8, 11). Ten days after sheep Ig application, control mice showed proteinuria and crescent formation in about 20% of glomeruli (Fig. 5, A to D). Treatment of mice with resolved *S. aureus* infection resulted in significantly enhanced numbers of glomerular crescents, demonstrating a more severe course of crescentic GN. Induction of crescentic GN was associated with increased frequencies of renal T<sub>H</sub>1 and T<sub>H</sub>17 cells (Fig. 5, E and F). Aggravated crescentic GN after *S. aureus* infection was accompanied by even greater increases in the frequency of renal T<sub>H</sub>17 cells as determined by IL-17A staining after polyclonal stimulation and of CD69<sup>+</sup> RORγt<sup>+</sup> T<sub>RM</sub>17 cells (Fig. 5G). In contrast, renal IFN-γ<sup>+</sup> T<sub>H</sub>1 cells were not further increased in mice with prior *S. aureus* infection. The more pronounced T<sub>H</sub>17 response but similar T<sub>H</sub>1 response in mice with prior *S. aureus* infection was also evident when total numbers of renal T<sub>H</sub>1 and T<sub>H</sub>17 cells were calculated (Fig. 5F). mRNA analysis corroborated the enhanced T<sub>H</sub>17 response (Fig. 5H) because we observed higher renal expression of the T<sub>H</sub>17-associated cytokines *Il17a*, *Cxcl5*, *Il6*, and *Il23a* in mice after *S. aureus* infection and crescentic GN induction. Compared with mice that did not have a prior *S. aureus* infection, renal *Il10* levels were not reduced, indicating that aggravated crescent formation and enhanced T<sub>H</sub>17 cell responses in mice with prior infection were not due to impaired IL-10 production. IL-17A fate reporter mice (*Il17a*<sup>Cre</sup> × *R26*<sup>eYFP</sup>) (36), in which T cells that had produced IL-17A, constitutively expressed enhanced

yellow fluorescent protein (eYFP), allowed for the localization of T<sub>H</sub>17 cells in the kidney (Fig. 5I). eYFP<sup>+</sup> cells could not be detected in kidneys of control mice but were frequently found after recovery of *S. aureus* infection with or without subsequent induction of crescentic GN. In these mice, eYFP<sup>+</sup> cells were mainly found in the tubulointerstitium surrounding the glomeruli. In conclusion, *S. aureus* infection before crescentic GN induction results in aggravated disease, which is associated with an enhanced renal T<sub>H</sub>17 cell response.

### ***S. aureus*-induced renal T<sub>RM</sub>17 cells drive crescentic GN**

Aggravated crescentic GN after *S. aureus* infection could be caused by persisting renal damage not detected in our extensive analysis or by alterations in the local composition and differentiation status of innate immune cells. To exclude T cell-independent mechanisms, *Rag1*<sup>-/-</sup> mice were infected with *S. aureus*, and after 7 days, bacteria were cleared with ampicillin. After recovery, mice were reconstituted with purified CD4<sup>+</sup> T cells from naïve mice and crescentic GN was induced (Fig. 5J). Reconstituted mice with and without prior infection showed similar severity of crescentic GN (Fig. 5, K and L), indicating that aggravation of disease in immunocompetent mice was likely due to *S. aureus*-induced changes in the renal T cell composition.

To link the aggravated crescentic GN directly with the enhanced T<sub>H</sub>17 cell response in mice with prior *S. aureus* infection, we used an IL-17A neutralization approach. One group of mice was infected with *S. aureus* and a control group remained uninfected, and all mice were subsequently treated with ampicillin. Two months after infection, crescentic GN was induced in the presence of anti-IL-17A mAb (Fig. 6A). Blockade of IL-17A was associated with low levels of glomerular crescent formation, as described previously (37). Neutralization of IL-17A abolished the difference in crescent formation between mice without and with prior *S. aureus* infection (Fig. 6, B and C). After IL-17A neutralization, kidneys of *S. aureus*-infected mice still contained enhanced frequencies of renal IL-17A<sup>+</sup> and RORγt<sup>+</sup> T<sub>H</sub>17 cells, whereas frequencies of IL-17A<sup>+</sup> or RORγt<sup>+</sup> T<sub>H</sub>17 cells in the spleen were not affected (Fig. 6, D and E). Both groups of mice showed similar renal mRNA expression of the IL-17A target gene *Cxcl5* (Fig. 6F).

In an alternative approach, we used mice with a diphtheria toxin receptor (DTR) constitutively induced by *Il17a* expression (*Il17a*<sup>Cre</sup> × *R26*<sup>DTR</sup> × *R26*<sup>eYFP</sup>) (36, 38). In these mice, injection of DT resulted in effective depletion of T<sub>H</sub>17 cells (Fig. 6G). Mice with resolved *S. aureus* infection and control mice were treated three times with DT to remove T<sub>H</sub>17 cells, including those induced by *S. aureus*. Two weeks later, crescentic GN was induced (Fig. 6, H to K), and both groups showed similarly low levels of crescent formation, which indicates that DT treatment abolished the exacerbated kidney damage acquired by prior *S. aureus* infection. We also observed similar frequencies of renal IL-17A<sup>+</sup> T<sub>H</sub>17 cells in both groups of mice, suggesting that DT had effectively removed preexisting T<sub>H</sub>17 cells including those induced by *S. aureus* infection.

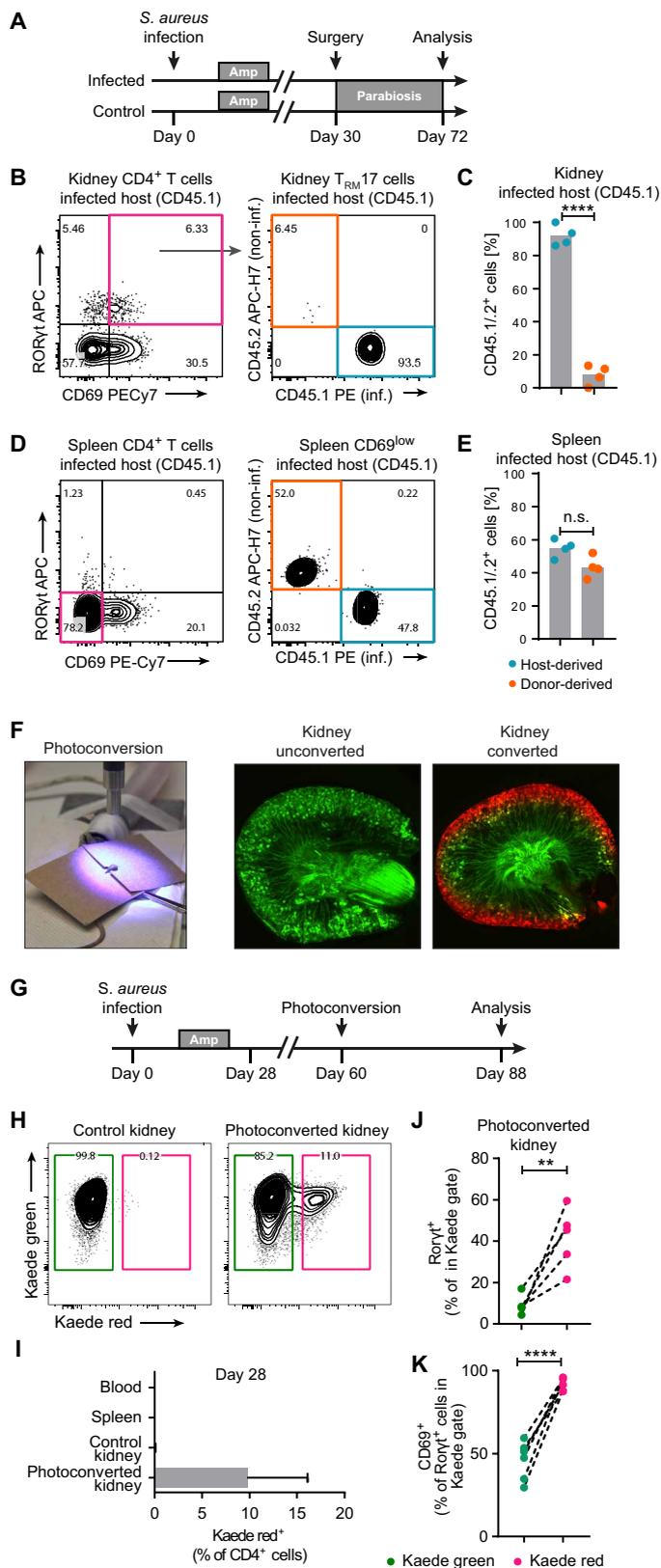
In addition to IL-17A-producing T<sub>RM</sub>17 cells, IFN-γ<sup>+</sup> CD4<sup>+</sup> T<sub>RM</sub> cells could be detected in the kidney of mice with prior *S. aureus* infection (Fig. 5, E and F). We neutralized IFN-γ to directly test whether IFN-γ-producing T<sub>RM</sub> cells could amplify crescentic GN. In contrast to IL-17A blockade, IFN-γ neutralization did not prevent more pronounced crescent formation in mice with prior *S. aureus* infection (Fig. 6, L to O). Together, these results indicate that *S. aureus*-induced renal T<sub>RM</sub>17 cells play a role and significantly contributed to the aggravated crescentic GN.

**Fig. 4. *S. aureus* induces tissue-resident memory  $T_{H17}$  cells in the kidney.** (A to E) Parabiosis of *S. aureus*-infected and noninfected mice. (A) CD45.1<sup>+</sup> mice were infected with *S. aureus*. Two weeks later, infected and CD45.2<sup>+</sup> control mice were treated for 1 week with ampicillin in the drinking water. Thirty days after infection, the circulation of CD45.1<sup>+</sup> and CD45.2<sup>+</sup> mice was surgically connected. Spleen and kidney T cells were analyzed 28 days later by flow cytometry. (B) Identification of donor-derived (CD45.2<sup>+</sup>) and host-derived (CD45.1<sup>+</sup>) CD4<sup>+</sup> ROR $\gamma$ <sup>+</sup> CD69<sup>+</sup> T<sub>RM</sub> cells in kidneys of previously infected host. (C) Frequencies of CD4<sup>+</sup> ROR $\gamma$ <sup>+</sup> CD69<sup>+</sup> T<sub>RM</sub> cells in kidneys of infected hosts. (D) Representative staining and (E) quantification of CD4<sup>+</sup> ROR $\gamma$ <sup>+</sup> CD69<sup>-</sup> of spleen cells from previously infected CD45.1<sup>+</sup> mice. Data are combined from two independent experiments. \*\*\*\**P* < 0.0001 (unpaired *t* test). n.s., not significant. (F to I) Use of photoconvertible *Kaede* transgenic mice to investigate tissue residency of renal T cells. (F) Photoconversion and fluorescence microscopy before and after renal photoconversion. (G) Experimental setup. Sixty days after *S. aureus* infection, one kidney was photoconverted and analyzed 4 weeks later. (H) *Kaede* red<sup>+</sup> T cells 4 weeks after photoconversion. (I) Frequency of photoconverted *Kaede* red<sup>+</sup> CD4<sup>+</sup> T cells in blood, spleen, and kidneys of *Kaede* mice 28 days after photoconversion (mean  $\pm$  SD, *n* = 3). (J) Quantification of ROR $\gamma$ <sup>+</sup> cells among *Kaede* red<sup>+</sup> and *Kaede* green<sup>+</sup> CD4<sup>+</sup> T cells 4 weeks after photoconversion. (K) CD69 expression of *Kaede* red<sup>+</sup> and *Kaede* green<sup>+</sup> CD4<sup>+</sup> ROR $\gamma$ <sup>+</sup> T cells. Data are representative of three independent experiments. \*\**P* < 0.01 and \*\*\*\**P* < 0.0001 (unpaired *t* test).

To investigate whether *S. aureus*-induced T cells residing in the kidney become reactivated and contribute to the T<sub>H17</sub> cell response during subsequent crescentic GN, we made use of *Kaede* transgenic mice. Mice were infected with *S. aureus*, and crescentic GN was induced after recovery from infection. Four days after crescentic GN induction, which is before the main wave of T<sub>H17</sub> cell recruitment to the kidney occurs (39), cells in one kidney were converted (*Kaede* red<sup>+</sup>) (Fig. 6P). Three days later, a large population of *Kaede* red<sup>+</sup> CD4<sup>+</sup> T cells that expressed IL-17A was detected in the kidney of mice with preceding infection. In contrast, the vast majority of IL-17A-expressing CD4<sup>+</sup> T cells in kidneys of control mice were *Kaede* green<sup>+</sup> (Fig. 6, Q and P). This indicates that the majority of T<sub>H17</sub> cells in control mice were recruited to the kidney, but mice with prior infection had a large fraction of these cells derived from *S. aureus*-induced renal-resident T<sub>H17</sub> cells.

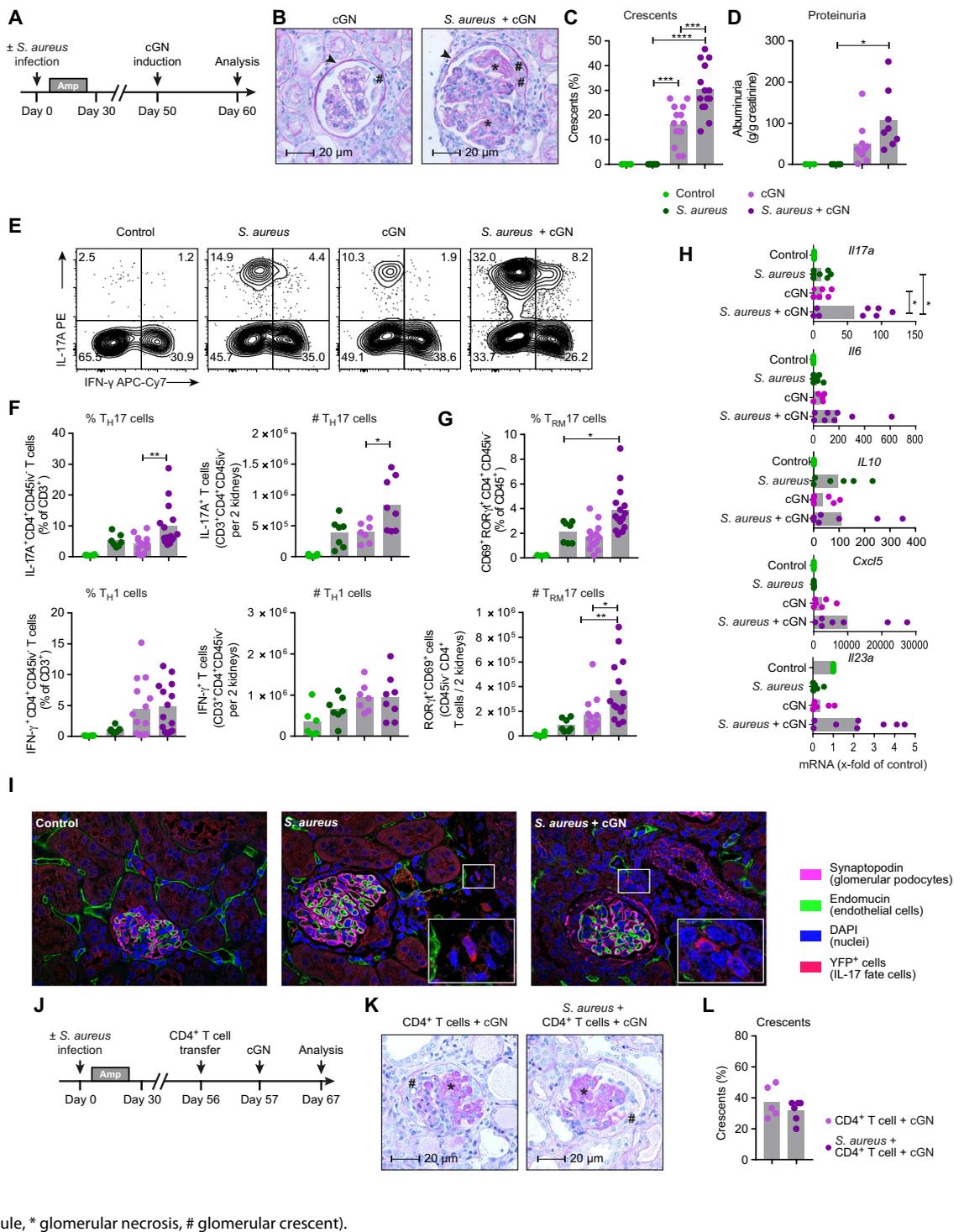
We tested the consequence of *S. aureus* infection-induced renal T<sub>RM17</sub> cells in a second model of crescentic GN, experimental autoimmune GN (EAG), a mouse model of Goodpasture syndrome (40). In this model, repeated immunization of DBA/1 mice with a fragment of type IV collagen induces antibodies as well as T<sub>H1</sub> and T<sub>H17</sub> cells specific to the collagen matrix of the GBM. Over a period of several weeks, mice develop severe crescentic GN and eventually succumb to renal failure (14). DBA/1 mice were infected with *S. aureus* and EAG was induced after recovery from infection (Fig. 7, A to D). In contrast to immunized control mice, a substantial number of previously infected mice succumbed to renal disease. After 10 weeks, surviving mice showed more severe crescentic GN in terms of enhanced crescent formation compared with mice that were not infected. Mice had increased frequencies of renal IL-17A<sup>+</sup> T<sub>H17</sub> cells, but similar frequencies of IFN- $\gamma$ <sup>+</sup> T<sub>H1</sub> cells. Overall, results from the EAG model were consistent with the results observed after transfer of anti-GBM sheep Ig.

*Listeria monocytogenes* transiently infects the kidney and induces strong CD4<sup>+</sup> T<sub>H1</sub> cell responses (41). After induction of crescentic GN, mice with and without preceding *L. monocytogenes* infection showed similar levels of kidney injury and equal frequencies of renal ROR $\gamma$ <sup>+</sup> T<sub>H17</sub> cells (fig. S6). Thus, in contrast to *S. aureus*



infection, which caused aggravation of renal damage in different models, *L. monocytogenes* infection did not change renal disease. These results indicate that the course and severity of renal autoimmunity

**Fig. 5. *S. aureus* aggravates crescentic GN and exacerbates the renal  $T_H17$  immune response.** (A) Mice were infected as described in Fig. 3. After 8 weeks, experimental crescentic GN (cGN) was induced, and kidneys were analyzed 10 days later. (B) Preceding *S. aureus* infection causes enhanced glomerular crescent formation. PAS staining of renal tissue sections 10 days after cGN induction. (C) Quantification of glomerular crescents and (D) proteinuria. (E) Flow cytometry of renal  $CD45iv^- CD4^+$  T cells after intracellular staining for IL-17A and IFN- $\gamma$ . (F) Quantification of renal  $CD45iv^- T_H1$  and  $T_H17$  cells by their cytokine production as well as (G) renal tissue-resident  $T_H17$  cells by ROR $\gamma t$  and CD69 expression. (H) RT-PCR analysis of cytokines in the renal cortex at day 10 of cGN as indicated. Data are representative of four independent experiments. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , and \*\*\*\* $P < 0.0001$  (one-way ANOVA with multiple comparison test). (I) Localization of  $T_{RM}17$  cells analyzed by renal immunofluorescence microscopy. Data are representative of two independent experiments ( $n = 3$  to 4 mice per group). (J to L) Analysis of the contribution of the innate immune response to *S. aureus*-induced aggravation of cGN. (J) *Rag1*<sup>-/-</sup> mice were infected with *S. aureus*. After bacterial clearing,  $CD4^+$  T cells were transferred and cGN was induced subsequently. (K) Renal histology and (L) quantification of glomerular crescents of mice with and without preceding *S. aureus* infection. Data are representative of two independent experiments (unpaired *t* test) (► Bowman's capsule, \* glomerular necrosis, # glomerular crescent).



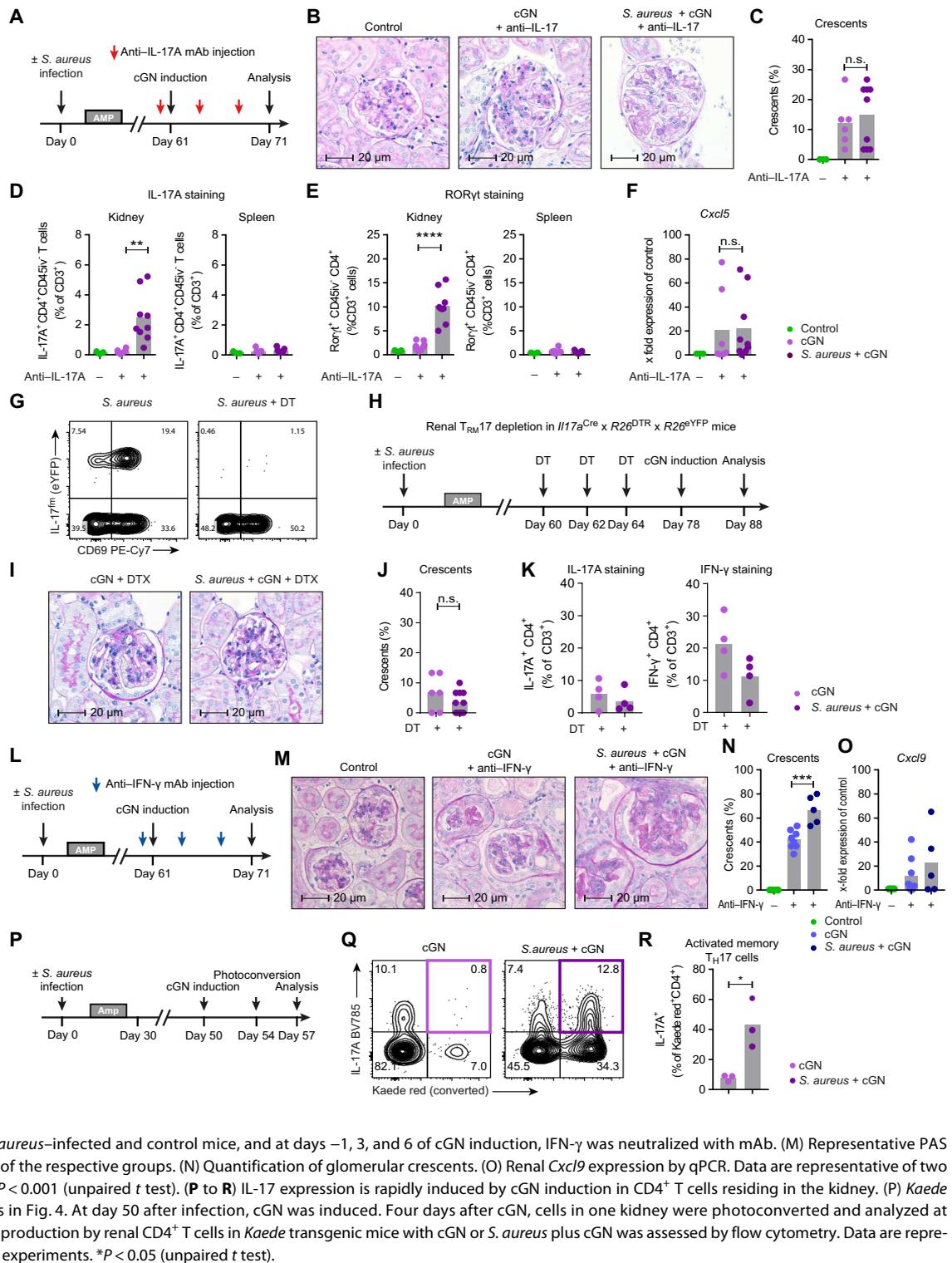
depend on the type of prior infection and the renal T cell composition elicited during the antipathogen response and suggest a unique pathogenic role of  $T_{RM}17$  cells in this context.

**Infection-induced  $T_{RM}17$  cells did not promote autoimmunity at distant sites**

After systemic infection, *S. aureus* disseminated into peripheral tissues, particularly into the kidney, which was associated with renal

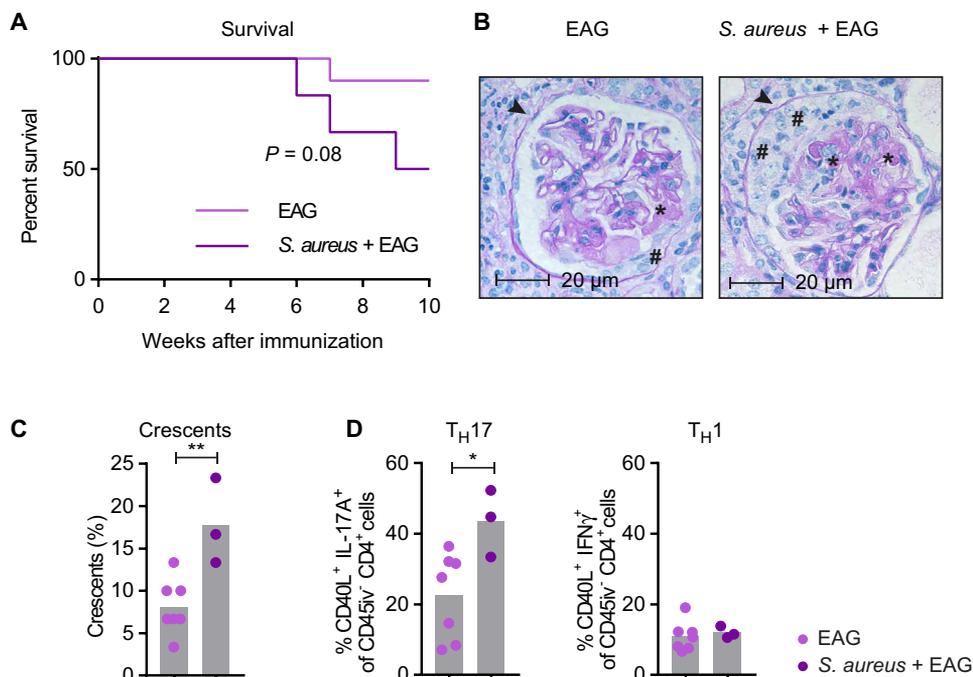
expansion of the  $T_{RM}17$  cell population (Fig. 3, A, K, and L). To test whether infection at distant sites also resulted in the generation of renal  $T_{RM}17$  cells, mice received intradermal infection with *S. aureus* (fig. S7A) (42). Analysis after antibiotic treatment and recovery revealed that local skin infection with *S. aureus* did not induce  $T_{RM}17$  cells in the kidney (fig. S7, B and C), suggesting that the presence of pathogens in tissue is a prerequisite for local  $T_{RM}17$  generation.

**Fig. 6. *S. aureus*-induced renal tissue-resident memory  $T_{H17}$  cells drive crescentic GN.** (A to F) Inhibition of IL-17 signaling by monoclonal antibodies (mAb).



We also tested whether systemic *S. aureus* promoted autoimmunity in the skin using an imiquimod-induced dermal inflammation, which is a well-established mouse model of IL-17-driven psoriasis (43). Mice were intravenously infected with *S. aureus* and imiquimod was applied to one ear of each mouse after recovery from infection. Ears were scored for thickness and inflammation (fig. S7, D to G). In

contrast to the GN models, preceding *S. aureus* infection did not aggravate dermal inflammation. These results suggest that *S. aureus* infection does not cause a generally increased susceptibility to autoimmune diseases. Aggravation of disease may depend on local generation of  $T_{RM17}$  cells that become reactivated during an autoimmune response at the site.



**Fig. 7. Impact of *S. aureus* infection on experimental crescentic GN.** (A) EAG was induced in DBA/1 mice. Survival of mice during the course of EAG. (B) PAS staining of glomeruli in EAG mice with and without preceding bacterial infection. (C) Quantification of renal damage (crescents) and (D) renal  $T_H1$  and  $T_H17$  cells by their cytokine production. Data are representative of two independent experiments \* $P < 0.05$  and \*\* $P < 0.01$  (unpaired  $t$  test) (► Bowman's capsule, \* glomerular necrosis, # glomerular crescent).

### Rapid IL-17 response of renal $T_{RM17}$ cells is cytokine-driven

To directly assess reactivation of renal  $T_{RM17}$  cells, *Kaede* mice that were also transgenic for the fluorescent protein Katushka under the control of the *Il17a* promoter [*Kaede*  $\times$  *Il17a*<sup>Kat</sup> mice (11, 44)] were infected with *S. aureus*. After recovery, one kidney was photoconverted and crescentic GN was induced 1 day later, and CD4<sup>+</sup> T cells from both kidneys were analyzed 2 days later (Fig. 8, A and B). In the nonexposed kidneys, about 1% of CD4<sup>+</sup> T cells were Kat<sup>+</sup>, indicating that cells actively produced IL-17A. In the light-exposed kidney, Kat<sup>+</sup> cells were mainly found in the photoconverted CD4<sup>+</sup> T cell population when compared with nonconverted CD4<sup>+</sup> T cells. This result demonstrates that IL-17A production in the early phase of crescentic GN was mainly due to the activation of kidney-resident T cells and not to rapid recruitment of cells to the kidney. Because photoconversion of renal cells is incomplete, a fraction of IL-17A<sup>+</sup> cells in the nonconverted *Kaede* green<sup>+</sup> CD4<sup>+</sup> T cell population was most likely also derived from kidney-resident cells, which suggests that IL-17A production was almost exclusively due to renal-resident CD4<sup>+</sup> T cells.

It is unlikely that *S. aureus*-specific  $T_{RM17}$  cells generated during infection recognize and become activated by unrelated antigens in two distinct models of crescentic GN. The rapid IL-17A response of *S. aureus*-induced renal T cells after crescentic GN induction suggests that activation is antigen independent. IL-17A fate reporter mice were infected to identify activation pathways of renal  $T_{RM17}$  cells. After recovery from infection, crescentic GN was induced in one group of mice. At day 2 of crescentic GN, which is a very early time point preceding renal T cell recruitment and substantial tissue injury, scRNA-seq of renal IL-17A fate<sup>+</sup> CD4<sup>+</sup> T cells was performed. Unsupervised

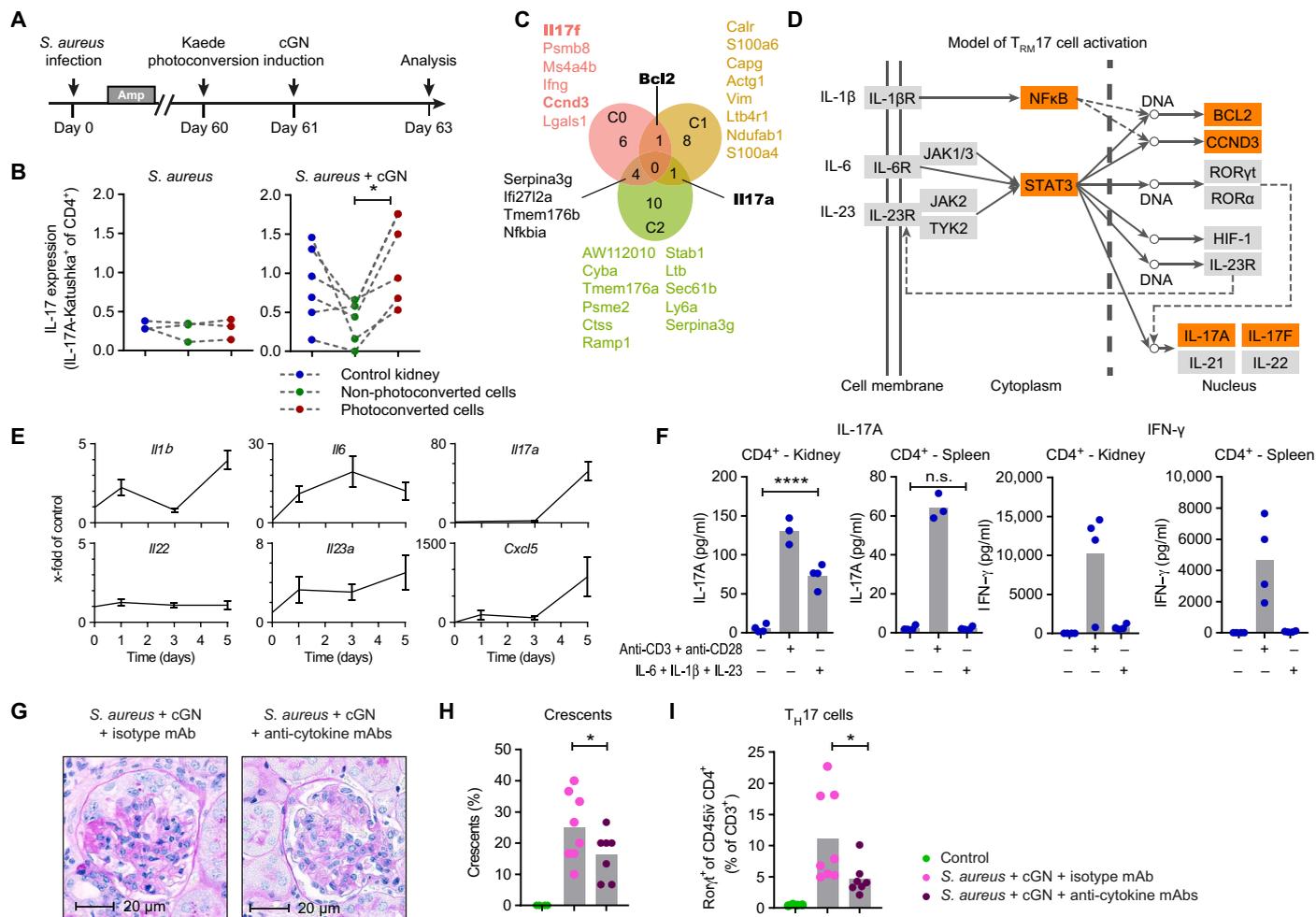
clustering of renal IL-17A fate<sup>+</sup> CD4<sup>+</sup> T cells from mice with and without crescentic GN defined eight distinct populations. The T cell composition from both groups of mice was relatively similar, and T cells from both groups were found in all clusters. This excluded extensive recruitment of  $T_H17$  cells from extrarenal sites at this time point (fig. S8, A to C). In the most prominent clusters, C0 to C2, cells from mice with crescentic GN displayed up-regulation of a number of genes, including *Il17a*, as compared with cells from control animals (Fig. 8C and fig. S8C). To uncover activation pathways that were triggered in these T cells by induction of crescentic GN, we made use of the TRRUSTv2 database, which was generated to study transcriptional regulation based on gene expression profiles (45). The TRRUSTv2 analysis tool identified several transcription factors that might be active in these cells with signal transducer and activator of transcription 3 (STAT3) and nuclear factor  $\kappa$ B (NF $\kappa$ B) (p50/Nfkb1) as top candidates (table S4). On the basis of these analyses and Kyoto Encyclopedia of Genes and Genomes reference pathways for STAT3 and NF $\kappa$ B (46), we proposed a model of  $T_{RM17}$  cell

activation by inflammatory cytokines such as IL-1 $\beta$ , IL-6, and IL-23 (Fig. 8D), which is also consistent with the observation that  $T_H17$  cell can be activated by inflammatory cytokines (47).

Analyses of renal tissues revealed increased expression of *Il1b*, *Il6*, and *Il23a* mRNA at 1 day after crescentic GN induction (Fig. 8E). To test whether these cytokines induce IL-17A, CD4<sup>+</sup> T cells were sorted from the spleen and kidney of mice previously infected with *S. aureus* and were stimulated with a cocktail of IL-1 $\beta$ , IL-6, and IL-23, or with anti-CD3 and anti-CD28 mAb as a control (Fig. 8F), and cytokines in supernatants were determined. Anti-CD3/CD28 stimulation induced IL-17A and IFN- $\gamma$  in T cells from the spleen and kidney. Renal CD4<sup>+</sup> T cells also responded to inflammatory cytokines in the absence of T cell receptor (TCR) stimulation by producing IL-17A, indicating that inflammatory cytokines can activate renal  $T_{RM17}$  cells in the early phase of crescentic GN. On the basis of these results, mice were infected with *S. aureus*, and after recovery, crescentic GN was induced in the presence of neutralizing mAb against IL-1 $\beta$ , IL-6, and IL-23p40 (Fig. 8, G to I). Anti-cytokine mAb-treated mice demonstrated attenuated crescent formation when compared with mice treated with control IgG (isotype) (Fig. 8, G and H). We also observed reduced frequencies of renal ROR $\gamma$ <sup>+</sup>  $T_H17$  cells, most likely as a consequence of reduced renal inflammation (Fig. 8I). Collectively, these results demonstrate that resident renal  $T_H17$  cells can be activated by inflammatory cytokines and thereby cause aggravated crescentic GN.

### DISCUSSION

In contrast to the role of  $T_{RM}$  cells in the defense against pathogens at interfaces with the environment, such as skin, gut, and airways,



**Fig. 8. Rapid IL-17 response in tissue-resident memory T cells is cytokine-driven.** (A and B) cGN drives rapid IL-17 expression in  $T_{RM}17$  cells in the kidney. (A) *Kaede*  $\times$  *Il17a*<sup>Kat</sup> mice were infected with *S. aureus*. After bacterial clearance, one kidney was photoconverted and cGN was induced. Two days after cGN, Katushka expression (indicating IL-17A production) was determined in renal CD45iv<sup>-</sup> CD4<sup>+</sup> T cells by flow cytometry. (B) Quantification of in vivo IL-17A production in kidneys at day 2 of cGN in *Kaede*  $\times$  *Il17a*<sup>Kat</sup> mice. Data are representative of two independent experiments. \**P* < 0.05 (one-way ANOVA with multiple comparison test). (C and D) Gene expression profiles induced by cGN in  $T_{RM}17$  cells. (C) IL-17A fate reporter mice with prior *S. aureus* infection were subjected to cGN and control treatment. After 2 days, sorted renal CD4<sup>+</sup> eYFP<sup>+</sup>  $T_{RM}17$  cells were analyzed by scRNA-seq. In infected mice, three clusters with up-regulation of *Il17a* or *Il17f* were identified (C0 to C2). Genes up-regulated after cGN in cells of the respective clusters are shown. (D) Proposed model of the potential regulatory network of  $T_{RM}17$  cell activation based on scRNA profiles, TRRUSTv2 analysis of differentially expressed genes of cells from mice with and without cGN (clusters C0 to C2) and KEGG reference pathways (orange: proteins and transcription factors identified by scRNA-seq and TRRUSTv2). (E and F)  $T_{RM}17$  cells can be activated by cytokines in the absence of TCR signaling. (E) Proinflammatory cytokines in the kidney were analyzed during the early phase of cGN by qPCR (mean  $\pm$  SEM, *n* = 5 to 11). (F) FACS-sorted CD4<sup>+</sup> T cells from spleen and kidney were stimulated with proinflammatory cytokines for 72 hours. Cytokines were measured in the supernatants. Data are representative of three independent experiments. \*\*\*\**P* < 0.0001 (unpaired *t* test). (G to I) After prior *S. aureus* infection, cGN was induced in mice. Cytokines (IL-1 $\beta$ , IL-6, and IL-23p40) were neutralized using mAb at days 0, 3, and 6 of cGN. (G and H) Glomerular crescents and (I) frequencies of  $T_H17$  cells were determined at day 10 of cGN. Data are representative of two independent experiments. \**P* < 0.05 (one-tailed *t* test).

their function in autoimmune diseases is currently not well understood (18–20). The long-term persistence of  $T_{RM}$  cells in tissues often affected by chronic relapsing and remitting autoimmune diseases such as skin (psoriasis) (48), gut (IBD) (49, 50), and brain (multiple sclerosis) (51) suggests a potential contribution of  $T_{RM}$  cells to these diseases.  $T_{RM}$  cells are detectable in active skin lesions of patients with psoriasis and persist in these areas after successful treatment (48), which might explain the recurrence of treated and resolved psoriasis lesions at the same site. The accumulation of CD4<sup>+</sup>CD69<sup>+</sup>CD103<sup>+</sup> T cells in the mucosa of patients with IBD was predictive for the development of disease flares, and in experimental IBD, de-

pletion of  $T_{RM}$  cells attenuated colitis activity, suggesting a critical role of  $T_{RM}$  cells in the pathogenesis of IBD (50). Steinbach *et al.* (52) recently reported that a transient viral infection of the mouse brain in early life induces brain  $T_{RM}$  cells that precipitate CNS autoimmune disease.

We detected T cells with a CD4<sup>+</sup> CD45RA<sup>low</sup> CCR7<sup>low</sup> CD69<sup>+</sup> phenotype and the signature transcriptional profile of CD4<sup>+</sup>  $T_{RM}$  cells (21) in the human kidney. When compared with healthy renal tissue from tumor nephrectomies, tissue from patients with ANCA-GN had substantially increased numbers of CD4<sup>+</sup>  $T_{RM}$  cells. In addition, in biopsies of a larger cohort of patients with active

ANCA-associated GN, high numbers of renal CD69<sup>+</sup> cells are correlated with impaired kidney function. In terms of mRNA expression profile and chemokine receptor surface expression, major subsets of renal CD4<sup>+</sup>CD69<sup>+</sup> T<sub>RM</sub> cells showed features of T<sub>H1</sub> or T<sub>H17</sub> cells and therefore were termed T<sub>RM1</sub> and T<sub>RM17</sub> cells, respectively. Direct comparison of CD4<sup>+</sup> T<sub>RM</sub> cells derived from renal tissue of patients with ANCA-GN with CD4<sup>+</sup> T<sub>RM</sub> cells from healthy renal tissue from patients with tumor nephrectomy revealed up-regulation of genes involved in activation, proliferation, and cytokine signaling. This expression profile likely reflects activation of T<sub>RM</sub> cells in the context of an acute renal immune response in the patients with ANCA-GN.

*S. aureus* induces T<sub>H17</sub> cell responses, and *S. aureus*-specific T<sub>H17</sub> cells are relatively abundant in human peripheral blood (34, 53). On the basis of our experimental data from mouse models, we hypothesized that T<sub>RM17</sub> cells specific for *S. aureus* could amplify an autoimmune reaction in the kidney. Patients with ANCA-associated vasculitis that carry *S. aureus* on their nasal mucosa show higher relapse rates and antibiotic treatment with trimethoprim and sulfamethoxazole (cotrimoxazole) prevents relapses (35, 54). *S. aureus*-specific T<sub>RM17</sub> cells could be a link between *S. aureus* and GN in ANCA-vasculitis. Overt or subclinical systemic infections (55) could boost renal T<sub>RM17</sub> cells, which subsequently amplify a smoldering autoimmune reaction. Other pathogens that infect the human kidney are uropathogenic *Escherichia coli* strains and *Candida albicans*. We infected mice with uropathogenic *E. coli*, which induces a transient form of acute pyelonephritis (56), and with *C. albicans*, a well-known trigger of T<sub>H17</sub> cell response (57). Analysis after clearance of the renal infections revealed that both pathogens induced T<sub>RM17</sub> cell formation in the kidneys of mice (fig. S9). These results suggest that in addition to *S. aureus*, *C. albicans* and uropathogenic *E. coli* but also other so far unknown pathogens can induce T<sub>RM</sub> cells in human kidneys.

It is currently unknown to which extent pathogen-specific T cells and particularly T cells specific for *S. aureus*, *E. coli*, or *C. albicans* contribute to CD4<sup>+</sup> T<sub>RM17</sub> cell populations in the human kidney. Because all three pathogens are inducers of T<sub>H17</sub> cells in humans (34, 58), it is likely that a fraction of human renal T<sub>RM17</sub> cells is directed against these pathogens. CD4<sup>+</sup> T cells specific for these three pathogens are found in the blood of healthy individuals in the range of 1 in 10<sup>3</sup> to 10<sup>4</sup> CD4<sup>+</sup> T cells (53, 59). We therefore estimate that at least 10<sup>6</sup> purified CD4<sup>+</sup> T cells would be required for the reliable detection of these pathogen-specific CD4<sup>+</sup> T cells. Using currently available approaches, it is not possible to obtain these quantities of CD4<sup>+</sup> T cells from renal biopsies or from tissue of tumor nephrectomies.

Our infection studies in mice support the concept that renal T<sub>RM17</sub> cells are generated in response to kidney infection because we found high numbers of T<sub>RM17</sub> cells that persisted in the kidney after resolution of infections. In the *S. aureus* infection model, parabiosis proved that the tissue residency of infection induced T<sub>RM17</sub> cells. Photoconversion of kidneys in *Kaede* transgenic mice confirmed this result and might provide a tool to circumvent ethically controversial parabiosis experiments in the future. It is therefore possible that T<sub>RM17</sub> cells identified in the healthy human kidney represent the memory of preceding infections.

Resolved *S. aureus* infection caused aggravation of renal disease in two different crescentic GN models. We carefully excluded persisting glomerular damage as cause for aggravated crescentic GN. *Rag1*<sup>-/-</sup> mice reconstituted with CD4<sup>+</sup> T cells after *S. aureus* infec-

tion and reconstituted controls showed similar crescentic GN development. Comparable severity of disease was also observed after depletion of T<sub>RM17</sub> cells or IL-17A neutralization in mice with and without prior infection. In contrast, neutralization of IFN- $\gamma$  did not prevent the more pronounced crescent formation in mice with prior *S. aureus* infection. These experiments define the enlarged T<sub>H17</sub> cell population as a mediator of worsened disease, which is consistent with the detrimental function of T<sub>H17</sub> cells in both crescentic GN models (11, 14). Labeling of cells in *Kaede* mice before crescentic GN revealed that renal T<sub>RM17</sub> cells become locally activated in the kidney. Renal T<sub>H17</sub> cells produced IL-17A after stimulation with IL-1 $\beta$ , IL-6, and IL-23, and transcriptional analyses indicated a response to inflammatory cytokines in renal T<sub>H17</sub> cells in the early stage of crescentic GN (60). Together, these results suggest that T<sub>RM17</sub> cells can respond to the local inflammation by production of IL-17A and most likely further cytokines that promote renal damage. Antigen-independent activation by inflammatory cytokines, including IL-1 family members, IL-6, IL-12, and IL-23, has been described not only for innate T cells, such as natural killer T cells (61), but also for conventional CD4<sup>+</sup> and CD8<sup>+</sup> effector T cells (62–66). The bystander activation of conventional T cells could support control of infection but could also contribute to tissue damage in autoimmune responses. In the model of experimental autoimmune encephalomyelitis, Lee *et al.* (47) could recently show that accumulation of T<sub>H17</sub> cells with irrelevant specificity in the inflamed central nervous system and bystander activation of these cells by IL-1 $\beta$  and IL-23 contribute to local injury. Our data indicate that T<sub>RM17</sub> cells induced by previous unrelated infections could function in a similar manner and promote tissue damage after activation by inflammatory cytokines.

Whether the generation and persistence of CD4<sup>+</sup> T<sub>RM17</sub> cells in nonlymphoid tissues in response to T<sub>H17</sub> cell-inducing microbes, e.g., *S. aureus*, *C. albicans*, or *E. coli*, are a common amplifier for T<sub>H17</sub> cell-driven autoimmune diseases in humans remain to be elucidated. In addition to T<sub>RM17</sub> cells, we detected T<sub>RM</sub> cells of the T<sub>H1</sub> or T<sub>H2</sub> phenotype as well as CD69<sup>+</sup> CD8<sup>+</sup> T cells in human kidneys. It would be of interest to investigate whether these T cells can act in a similar way in amplifying immune-mediated disorders. In the murine crescentic GN model, infection with *L. monocytogenes* that induces a local CD4<sup>+</sup> T<sub>H1</sub> cell response (67) did not affect the course of renal disease. Moreover, IFN- $\gamma$  neutralization did not prevent the more pronounced crescent formation in mice with prior *S. aureus* infection. However, these results do not exclude the idea that T<sub>RM1</sub> cells interfere with responses in other autoimmune diseases.

Together, we provide a concept for the predisposing role of microbial infections for T<sub>H17</sub>-driven autoimmune diseases and identify T<sub>RM17</sub> cells as a potential therapeutic target (fig. S10), indicating that T<sub>RM17</sub> cell depletion rather than transient cytokine (IL-17) neutralization might be a causal treatment strategy in chronic relapsing and remitting autoimmune diseases.

## MATERIALS AND METHODS

### Study design

The aims of the study were to characterize T cell subsets in the kidney and to determine their potential role in renal autoimmunity. T cells from human kidney biopsies from controls and from individuals with ANCA-associated GN were analyzed by scRNA-seq, CITE-seq, flow cytometry, and immune histology. These assays revealed renal CD4<sup>+</sup>

T<sub>RM</sub> cells with phenotype and mRNA expression profile of T<sub>H</sub>17 cells. For the functional characterization of renal CD4<sup>+</sup> T<sub>RM</sub> cells, we combined mouse infection models (*S. aureus*, *L. monocytogenes*, *E. coli*, and *C. albicans*) with models of crescentic GN. Transgenic cytokine reporter mice and *Kaede* mice that allowed identification and tracking of cells in vivo were included in these experiments. Mice were analyzed using standard assays for renal function, as well as scRNA-seq, quantitative polymerase chain reaction, flow cytometry, and histology. Detailed methods are given below and in the Supplementary Materials.

## Human studies

Human biopsies for flow cytometry and scRNA-seq were obtained either from the healthy part of kidneys after tumor nephrectomy or from renal biopsies of patients with ANCA-associated GN. Human kidney sections and clinical parameters were analyzed from patients included in the Hamburg GN Registry. For some of the analyses, matched blood samples from the same patient were used. These studies were approved by the Ethik-Kommission der Ärztekammer Hamburg, local ethics committee of the chamber of physicians in Hamburg (Registration numbers PV 5026 and PV 5822), and were conducted in accordance with the ethical principles stated by the Declaration of Helsinki. Informed consent was obtained from all participating patients. Information on the patient cohorts is provided in tables S1 to S3. Details on the analysis of human renal biopsies and of human T cells are given in the Supplementary Materials.

## Mice

Mouse experiments were carried out in accordance with the national guidelines. The protocols were approved by local ethics committees. Mice were housed under specific pathogen-free conditions in individually ventilated cages with standard food and water ad libitum. During experiments, mice were monitored daily and mice with signs of severe disease were euthanized with an O<sub>2</sub>/CO<sub>2</sub> mixture to minimize suffering. All experiments were conducted with age-matched male mice on C57BL/6 background, with the exception of EAG that was induced in male DBA/1 mice. The following transgenic mouse strains were used: *Kaede* mice (11), *Il17a*<sup>Kat</sup> mice (44), *Il17a*<sup>Cre</sup> × *R26*<sup>eYFP</sup> mice (36, 39), *Il17a*<sup>Cre</sup> × *R26*<sup>DTR</sup> × *R26*<sup>eYFP</sup> mice (36, 38), CD45.1 congenic mice (B6.SJL-*Ptprca*<sup>a</sup> *Pepc*<sup>b</sup>/*BoyJ*), and *Rag1*<sup>-/-</sup> mice (B6.129S7-*Rag1*<sup>tm1Mom</sup>/*J*). Details on mouse experiments are given in the Supplementary Materials.

## Statistics

Statistical analysis was performed using GraphPad Prism (La Jolla, CA). The results are shown as mean ± SEM or mean ± SD when presented as a bar graph or as single data points with the mean in a scatter dot plot. Differences between two individual groups were compared using a two-tailed *t* test. In the case of three or more groups, a one-way analysis of variance (ANOVA) with Bonferroni's multiple comparisons test was used.

## SUPPLEMENTARY MATERIALS

immunology.sciencemag.org/cgi/content/full/5/50/eaba4163/DC1

Material and Methods

Fig. S1. scRNA-seq analysis of T cells from human blood and healthy kidney tissue from tumor nephrectomies.

Fig. S2. Profiling of T cells from human blood and healthy kidney tissue from tumor nephrectomies.

Fig. S3. scRNA-seq analysis of T cells from blood and kidney of patients with ANCA-GN.

Fig. S4. Profiling of T cells from kidney and blood of patients with ANCA-GN.

Fig. S5. Flow cytometry gating strategy for murine T cell analysis.

Fig. S6. Prior *L. monocytogenes* infection does not affect crescentic GN.

Fig. S7. Effect of local *S. aureus* skin infection on renal T<sub>RM</sub>17 cell generation and effect of systemic *S. aureus* infection on imiquimod-induced skin inflammation.

Fig. S8. Gene expression analysis of renal CD4<sup>+</sup> T cells from nephritic mice previously infected with *S. aureus*.

Fig. S9. The T<sub>H</sub>17-associated pathogens *C. albicans* and *E. coli* cause renal T<sub>RM</sub>17 cell formation.

Fig. S10. Conceptual link of pathogen-induced tissue resident memory T cells with autoimmunity.

Table S1. Baseline characteristics of controls at time of biopsy.

Table S2. Baseline characteristics of patients with ANCA-GN at the time of biopsy.

Table S3. Clinical characteristics and ANCA-serotype of patients from the Hamburg GN Registry.

Table S4. List of transcription factors identified by TRRUSTv2 analysis.

Data file S1.

References (68–79)

[View/request a protocol for this paper from Bio-protocol.](#)

## REFERENCES AND NOTES

1. K. W. Wucherpfennig, Mechanisms for the induction of autoimmunity by infectious agents. *J. Clin. Invest.* **108**, 1097–1104 (2001).
2. K. N. Konstantinov, C. J. Ulf-Møller, A. H. Tzamaloukas, Infections and antineutrophil cytoplasmic antibodies: Triggering mechanisms. *Autoimmun. Rev.* **14**, 201–203 (2015).
3. A. J. Steelman, Infection as an environmental trigger of multiple sclerosis disease exacerbation. *Front. Immunol.* **6**, 520 (2015).
4. E. A. Mann, S. A. Saeed, Gastrointestinal infection as a trigger for inflammatory bowel disease. *Curr. Opin. Gastroenterol.* **28**, 24–29 (2012).
5. W. G. Couser, R. J. Johnson, The etiology of glomerulonephritis: Roles of infection and autoimmunity. *Kidney Int.* **86**, 905–914 (2014).
6. C. Kurts, U. Panzer, H. J. Anders, A. J. Rees, The immune system and kidney disease: Basic concepts and clinical implications. *Nat. Rev. Immunol.* **13**, 738–753 (2013).
7. W. G. Couser, Basic and translational concepts of immune-mediated glomerular diseases. *J. Am. Soc. Nephrol.* **23**, 381–399 (2012).
8. C. F. Krebs, T. Schmidt, J. H. Riedel, U. Panzer, T helper type 17 cells in immune-mediated glomerular disease. *Nat. Rev. Nephrol.* **13**, 647–659 (2017).
9. B. R. Lúdviksson, M. C. Sneller, K. S. Chua, C. Talar-Williams, C. A. Langford, R. O. Ehrhardt, A. S. Fauci, W. Strober, Active Wegener's granulomatosis is associated with HLA-DR<sup>+</sup> CD4<sup>+</sup> T cells exhibiting an unbalanced Th1-Type T cell cytokine pattern: Reversal with IL-10. *J. Immunol.* **160**, 3602–3609 (1998).
10. J. D. Ooi, J. Petersen, Y. H. Tan, M. Huynh, Z. J. Willett, S. H. Ramarathnam, P. J. Eggenhuizen, K. L. Loh, K. A. Watson, P. Y. Gan, M. A. Alikhan, N. L. Dudek, A. Handel, B. G. Hudson, L. Fugger, D. A. Power, S. G. Holt, P. T. Coates, J. W. Gregersen, A. W. Purcell, S. R. Holdsworth, N. L. La Gruta, H. H. Reid, J. Rossjohn, A. R. Kitching, Dominant protection from HLA-linked autoimmunity by antigen-specific regulatory T cells. *Nature* **545**, 243–247 (2017).
11. C. F. Krebs, H.-J. Paust, S. Krohn, T. Koyro, S. R. Brix, J.-H. Riedel, P. Bartsch, T. Wiech, C. Meyer-Schwesinger, J. Huang, N. Fischer, P. Busch, H.-W. Mittrücker, U. Steinhoff, B. Stockinger, L. Garcia Perez, U. O. Wenzel, M. Janneck, O. M. Steinmetz, N. Gagliani, R. A. K. Stahl, S. Huber, J.-E. Turner, U. Panzer, Autoimmune renal disease is exacerbated by S1P-receptor-1-dependent intestinal Th17 cell migration to the kidney. *Immunity* **45**, 1078–1092 (2016).
12. H.-J. Paust, J.-E. Turner, O. M. Steinmetz, A. Peters, F. Heymann, C. Höltscher, G. Wolf, C. Kurts, H.-W. Mittrücker, R. A. K. Stahl, U. Panzer, The IL-23/Th17 axis contributes to renal injury in experimental glomerulonephritis. *J. Am. Soc. Nephrol.* **20**, 969–979 (2009).
13. S. A. Summers, O. M. Steinmetz, M. Li, J. Y. Kausman, T. Semple, K. L. Edgton, D. B. Borza, H. Braley, S. R. Holdsworth, A. R. Kitching, Th1 and Th17 cells induce proliferative glomerulonephritis. *J. Am. Soc. Nephrol.* **20**, 2518–2524 (2009).
14. S. Hünemörder, J. Treder, S. Ahrens, V. Schumacher, H. J. Paust, T. Menter, P. Matthyis, T. Kamradt, C. Meyer-Schwesinger, U. Panzer, H. Hopfer, H.-W. Mittrücker, T<sub>H</sub>1 and T<sub>H</sub>17 cells promote crescent formation in experimental autoimmune glomerulonephritis. *J. Pathol.* **237**, 62–71 (2015).
15. A. Suarez-Fueyo, S. J. Bradley, D. Klatzmann, G. C. Tsokos, T cells and autoimmune kidney disease. *Nat. Rev. Nephrol.* **13**, 329–343 (2017).
16. E. M. Disteldorf, C. F. Krebs, H. J. Paust, J. E. Turner, G. Nouailles, A. Tittel, C. Meyer-Schwesinger, G. Stege, S. Brix, J. Velden, T. Wiech, U. Helmchen, O. M. Steinmetz, A. Peters, S. B. Bannstein, A. Kaffke, C. Llanto, S. A. Lira, H. W. Mittrücker, R. A. K. Stahl, C. Kurts, S. H. E. Kaufmann, U. Panzer, CXCL5 drives neutrophil recruitment in T<sub>H</sub>17-mediated GN. *J. Am. Soc. Nephrol.* **26**, 55–66 (2015).
17. S. N. Mueller, L. K. Mackay, Tissue-resident memory T cells: Local specialists in immune defence. *Nat. Rev. Immunol.* **16**, 79–89 (2016).

18. O. Gaide, R. O. Emerson, X. Jiang, N. Gulati, S. Nizza, C. Desmarais, H. Robins, J. G. Krueger, R. A. Clark, T. S. Kupper, Common clonal origin of central and resident memory T cells following skin immunization. *Nat. Med.* **21**, 647–653 (2015).
19. P. A. Szabo, M. Miron, D. L. Farber, Location, location, location: Tissue resident memory T cells in mice and humans. *Sci. Immunol.* **4**, eaas9673 (2019).
20. M. C. Amezcua Vesely, P. Pallis, P. Bielecki, J. S. Low, J. Zhao, C. C. D. Harman, L. Kroehling, R. Jackson, W. Bailis, P. Licona-Limón, H. Xu, N. Iijima, P. S. Pillai, D. H. Kaplan, C. T. Weaver, Y. Kluger, M. S. Kowalczyk, A. Iwasaki, J. P. Pereira, E. Esplugues, N. Gagliani, R. A. Flavell, Effector T<sub>H</sub>17 cells give rise to long-lived T<sub>RM</sub> cells that are essential for an immediate response against bacterial infection. *Cell* **178**, 1176–1188.e15 (2019).
21. B. V. Kumar, W. Ma, M. Miron, T. Granot, R. S. Guyer, D. J. Carpenter, T. Senda, X. Sun, S. H. Ho, H. Lerner, A. L. Friedman, Y. Shen, D. L. Farber, Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. *Cell Rep.* **20**, 2921–2934 (2017).
22. P. Hombrik, C. Helbig, R. A. Backer, B. Piet, A. E. Oja, R. Stark, G. Brassler, A. Jongejan, R. E. Jonkers, B. Nota, O. Basak, H. C. Clevers, P. D. Moerland, D. Amsen, R. A. W. van Lier, Programs for the persistence, vigilance and control of human CD8<sup>+</sup> lung-resident memory T cells. *Nat. Immunol.* **17**, 1467–1478 (2016).
23. S. C. Jameson, D. Masopust, Understanding subset diversity in T cell memory. *Immunity* **48**, 214–226 (2018).
24. J. M. Schenkel, D. Masopust, Tissue-resident memory T cells. *Immunity* **41**, 886–897 (2014).
25. J.-E. Turner, M. Becker, H.-W. Mittrücker, U. Panzer, Tissue-resident lymphocytes in the kidney. *J. Am. Soc. Nephrol.* **29**, 389–399 (2018).
26. E. Z. Macosko, A. Basu, R. Satija, J. Nemes, 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, S. A. McCarroll, Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
27. M. Stoeciuk, C. Hafemeister, W. Stephenson, B. Houck-Loomis, P. K. Chattopadhyay, H. Swerdlow, R. Satija, P. Smibert, Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
28. E. P. Mimitou, A. Cheng, A. Montalbano, S. Hao, M. Stoeciuk, M. Legut, T. Roush, A. Herrera, E. Papalex, Z. Ouyang, R. Satija, N. E. Sanjana, S. B. Koralov, P. Smibert, Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019).
29. F. A. Vieira Braga, G. Kar, M. Berg, O. A. Carpaiz, K. Polanski, L. M. Simon, S. Brouwer, T. Gomes, L. Hesse, J. Jiang, E. S. Fasouli, M. Efreanova, R. Vento-Tormo, C. Talavera-López, M. R. Jonker, K. Affleck, S. Palit, P. M. Strzelecka, H. V. Firth, K. T. Mahubani, A. Cvejic, K. B. Meyer, K. Saeb-Parsy, M. Luinge, C. A. Brandsma, W. Timens, I. Angelidis, M. Strunz, G. H. Koppelman, A. J. van Oosterhout, H. B. Schiller, F. J. Theis, M. van den Berge, M. C. Nawijn, S. A. Teichmann, A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat. Med.* **25**, 1153–1163 (2019).
30. C. D. Cohen, K. Frach, D. Schlondorff, M. Kretzler, Quantitative gene expression analysis in renal biopsies: A novel protocol for a high-throughput multicenter application. *Kidney Int.* **61**, 133–140 (2002).
31. L. S. Cauley, L. Lefrançois, Guarding the perimeter: Protection of the mucosa by tissue-resident memory T cells. *Mucosal Immunol.* **6**, 14–23 (2013).
32. F. Sallusto, Heterogeneity of human CD4<sup>+</sup> T cells against microbes. *Annu. Rev. Immunol.* **34**, 317–334 (2016).
33. E. Y. Chen, C. M. Tan, Y. Kou, Q. Duan, Z. Wang, G. Vaz Meirelles, N. R. Clark, A. Ma'ayan, Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
34. C. E. Zielinski, F. Mele, D. Aschenbrenner, D. Jarrossay, F. Ronchi, M. Gattorno, S. Monticelli, A. Lanzavecchia, F. Sallusto, Pathogen-induced human T<sub>H</sub>17 cells produce IFN- $\gamma$  or IL-10 and are regulated by IL-1 $\beta$ . *Nature* **484**, 514–518 (2012).
35. C. A. Stegeman, J. W. Tervaert, P. E. de Jong, C. G. Kallenberg; Dutch Co-trimoxazole Wegener Study Group, Trimethoprim-sulfamethoxazole (co-trimoxazole) for the prevention of relapses of Wegener's granulomatosis. *N. Engl. J. Med.* **335**, 16–20 (1996).
36. K. Hirota, J. H. Duarte, M. Veldhoen, E. Hornsby, Y. Li, D. J. Cua, H. Ahlfors, C. Wilhelm, M. Tolaini, U. Menzel, A. Gafeflaky, A. J. Potocnik, B. Stockinger, Fate mapping of IL-17-producing T cells in inflammatory responses. *Nat. Immunol.* **12**, 255–263 (2011).
37. H.-J. Paust, J.-E. Turner, J.-H. Riedel, E. Disteldorf, A. Peters, T. Schmidt, C. Krebs, J. Velden, H. W. Mittrücker, O. M. Steinmetz, R. A. K. Stahl, U. Panzer, Chemokines play a critical role in the cross-regulation of Th1 and Th17 immune responses in murine crescentic glomerulonephritis. *Kidney Int.* **82**, 72–83 (2012).
38. T. Buch, F. L. Heppner, C. Tertilt, T. J. A. J. Heinen, M. Kremer, F. T. Wunderlich, S. Jung, A. Waisman, A Cre-inducible diphtheria toxin receptor mediates cell lineage ablation after toxin administration. *Nat. Methods* **2**, 419–426 (2005).
39. C. F. Krebs, J.-E. Turner, H.-J. Paust, S. Kapffer, T. Koyno, S. Krohn, F. Ufer, M. A. Friese, R. A. Flavell, B. Stockinger, O. M. Steinmetz, R. A. K. Stahl, S. Huber, U. Panzer, Plasticity of Th17 cells in autoimmune kidney diseases. *J. Immunol.* **197**, 449–457 (2016).
40. H. Hopfer, J. Holzer, S. Hünemörder, H. J. Paust, M. Sachs, C. Meyer-Schwesinger, J. E. Turner, U. Panzer, H.-W. Mittrücker, Characterization of the renal CD4<sup>+</sup> T-cell response in experimental autoimmune glomerulonephritis. *Kidney Int.* **82**, 60–71 (2012).
41. M. Kursar, K. Bonhagen, A. Köhler, T. Kamradt, S. H. E. Kaufmann, H.-W. Mittrücker, Organ-specific CD4<sup>+</sup> T cell response during *Listeria monocytogenes* infection. *J. Immunol.* **168**, 6382–6387 (2002).
42. J. Tinteln, F. Ufer, J. B. Engler, H. Winkler, K. Lücke, H. W. Mittrücker, M. A. Friese, Arg3.1 defines dendritic cells and Langerhans cells with superior migratory ability independent of phenotype and ontogeny in mice. *Eur. J. Immunol.* **49**, 724–736 (2019).
43. L. van der Fits, S. Mourits, J. S. A. Voerman, M. Kant, L. Boon, J. D. Laman, F. Cornelissen, A.-M. Mus, E. Florencia, E. P. Prens, E. Lubberts, Imiquimod-induced psoriasis-like skin inflammation in mice is mediated via the IL-23/IL-17 axis. *J. Immunol.* **182**, 5836–5845 (2009).
44. N. Gagliani, M. C. Amezcua Vesely, A. Iseppon, L. Brockmann, H. Xu, N. W. Palm, M. R. de Zoete, P. Licona-Limón, 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, R. A. Flavell, Th17 cells transdifferentiate into regulatory T cells during resolution of inflammation. *Nature* **523**, 221–225 (2015).
45. H. Han, J.-W. Cho, S. Lee, A. Yun, H. Kim, D. Bae, S. Yang, C. Y. Kim, M. Lee, E. Kim, S. Lee, B. Kang, D. Jeong, Y. Kim, H. N. Jeon, H. Jung, S. Nam, M. Chung, J. H. Kim, I. Lee, TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* **46**, D380–D386 (2018).
46. M. Kanehisa, S. Goto, KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
47. H.-G. Lee, J.-U. Lee, D.-H. Kim, S. Lim, I. Kang, J.-M. Choi, Pathogenic function of bystander-activated memory-like CD4<sup>+</sup> T cells in autoimmune encephalomyelitis. *Nat. Commun.* **10**, 709 (2019).
48. I. Gallais Serezal, C. Classon, S. Cheuk, M. Barrientos-Somarrivas, E. Wadman, E. Martini, D. Chang, N. X. Landén, M. Ehrström, S. Nylén, L. Eidsmo, Resident T cells in resolved psoriasis steer tissue responses that stratify clinical outcome. *J. Invest. Dermatol.* **138**, 1754–1763 (2018).
49. S. Omenetti, C. Bussi, A. Metidji, A. Iseppon, S. Lee, M. Tolaini, Y. Li, G. Kelly, P. Chakravarty, S. Shoaie, M. G. Gutierrez, B. Stockinger, The intestine harbors functionally distinct homeostatic tissue-resident and inflammatory Th17 cells. *Immunity* **51**, 77–89.e6 (2019).
50. S. Zundler, E. Becker, M. Spocinska, M. Slawik, L. Parga-Vidal, R. Stark, M. Wiendl, R. Atreya, T. Rath, M. Leppkes, K. Hildner, R. López-Posadas, S. Lukassen, A. B. Eicki, C. Neufert, I. Atreya, K. P. J. M. van Gisbergen, M. F. Neurath, Hobit- and Blimp-1-driven CD4<sup>+</sup> tissue-resident memory T cells control chronic intestinal inflammation. *Nat. Immunol.* **20**, 288–300 (2019).
51. K. Steinbach, I. Vincenti, M. Kreutzfeldt, N. Page, A. Muschwack, I. Wagner, I. Drexler, D. Pinschewer, T. Korn, D. Merkler, Brain-resident memory T cells represent an autonomous cytotoxic barrier to viral infection. *J. Exp. Med.* **213**, 1571–1587 (2016).
52. K. Steinbach, I. Vincenti, K. Egervari, M. Kreutzfeldt, F. van der Meer, N. Page, B. Klimek, I. Rossitto-Borlat, G. Di Liberto, A. Muschwack, I. Wagner, K. Hammad, C. Stadelmann, T. Korn, O. Hartley, D. D. Pinschewer, D. Merkler, Brain-resident memory T cells generated early in life predispose to autoimmune disease in mice. *Sci. Transl. Med.* **11**, eaav5519 (2019).
53. P. Bacher, F. Heinrich, U. Stervbo, M. Nienen, M. Vahldeick, C. Iwer, K. Vogt, J. Kollet, N. Babel, B. Sawitzki, C. Schwarz, S. Bereswill, M. M. Heimesaat, G. Heine, G. Gadermaier, C. Asam, M. Assenmacher, O. Kniemeyer, A. A. Brakhage, F. Ferreira, M. Wallner, M. Worn, A. Scheffold, Regulatory T cell specificity directs tolerance versus allergy against aeroantigens in humans. *Cell* **167**, 1067–1078.e16 (2016).
54. C. A. Stegeman, J. W. Tervaert, W. J. Sluiter, W. L. Manson, P. E. de Jong, C. G. Kallenberg, Association of chronic nasal carriage of *Staphylococcus aureus* and higher relapse rates in Wegener granulomatosis. *Ann. Intern. Med.* **120**, 12–17 (1994).
55. N. B. Parahitiyawa, L. J. Jin, W. K. Leung, W. C. Yam, L. P. Samaranyake, Microbiology of odontogenic bacteremia: Beyond endocarditis. *Clin. Microbiol. Rev.* **22**, 46–64 (2009).
56. M. Schiwon, C. Weisheit, L. Franken, S. Gutweiler, A. Dixit, C. Meyer-Schwesinger, J. M. Pohl, N. J. Maurice, S. Thiebes, K. Lorenz, T. Quast, M. Fuhrmann, G. Baumgarten, M. J. Lohse, G. Opendakker, J. Bernhagen, R. Bucala, U. Panzer, W. Kolanus, H. J. Gröne, N. Garbi, W. Kastenmüller, P. A. Knolle, C. Kurts, D. R. Engel, Crosstalk between sentinel and helper macrophages permits neutrophil migration into infected uroepithelium. *Cell* **156**, 456–468 (2014).
57. A. Vrablikova, L. Czernekova, R. Cahlikova, Z. Novy, M. Petrik, S. Imran, Z. Novak, M. Krupka, V. Cеровsky, J. Turanek, M. Raska, Lasioglossins LIII affect the morphogenesis of *Candida albicans* and reduces the duration of experimental vaginal candidiasis in mice. *Microbiol. Immunol.* **61**, 474–481 (2017).
58. K. Atarashi, 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. I. Yokoyama, S. Tokoro, H. Mori, Y. Noguchi, H. Morita, I. I. Ivanov, T. Sugiyama, G. Nuñez, J. G. Camp, M. Hattori, Y. Umesaki, K. Honda,

- Th17 cell induction by adhesion of microbes to intestinal epithelial cells. *Cell* **163**, 367–380 (2015).
59. P. Bacher, T. Hohnstein, E. Beerbaum, M. Röcker, M. G. Blango, S. Kaufmann, J. Röhmel, P. Eschenhagen, C. Grehn, K. Seidel, V. Rickerts, L. Lozza, U. Stervbo, M. Nienen, N. Babel, J. Milleck, M. Assenmacher, O. A. Cornely, M. Ziegler, H. Wisplinghoff, G. Heine, M. Worm, B. Siegmund, J. Maul, P. Creutz, C. Tabeling, C. Ruwwe-Glösenkamp, L. E. Sander, C. Knosalla, S. Brunke, B. Hube, O. Kniemeyer, A. A. Brakhage, C. Schwarz, A. Scheffold, Human anti-fungal Th17 immunity and pathology rely on cross-reactivity against *Candida albicans*. *Cell* **176**, 1340–1355.e15 (2019).
  60. M. J. McGeachy, D. J. Cua, S. L. Gaffen, The IL-17 family of cytokines in health and disease. *Immunity* **50**, 892–906 (2019).
  61. P. J. Brennan, M. Brigl, M. B. Brenner, Invariant natural killer T cells: An innate activation scheme linked to diverse effector functions. *Nat. Rev. Immunol.* **13**, 101–117 (2013).
  62. R. E. Berg, E. Crossley, S. Murray, J. Forman, Memory CD8<sup>+</sup> T cells provide innate immune protection against *Listeria monocytogenes* in the absence of cognate antigen. *J. Exp. Med.* **198**, 1583–1593 (2003).
  63. L. Guo, G. Wei, J. Zhu, W. Liao, W. J. Leonard, K. Zhao, W. Paul, IL-1 family members and STAT activators induce cytokine production by Th2, Th17, and Th1 cells. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13463–13468 (2009).
  64. J. E. Kohlmeier, T. Cookenham, A. D. Roberts, S. C. Miller, D. L. Woodland, Type I interferons regulate cytolytic activity of memory CD8<sup>+</sup> T cells in the lung airways during respiratory virus challenge. *Immunity* **33**, 96–105 (2010).
  65. B. Becher, S. Spath, J. Goverman, Cytokine networks in neuroinflammation. *Nat. Rev. Immunol.* **17**, 49–59 (2017).
  66. T. Chu, A. J. Tyznik, S. Roepke, A. M. Berkley, A. Woodward-Davis, L. Pattacini, M. J. Bevan, D. Zehn, M. Prlc, Bystander-activated memory CD8 T cells control early pathogen load in an innate-like, NKG2D-dependent manner. *Cell Rep.* **3**, 701–708 (2013).
  67. M. Kursar, K. Bonhagen, A. Köhler, T. Kamradt, S. H. E. Kaufmann, H. W. Mittrücker, Antigen-specific CD8<sup>+</sup> T cell responses in intestinal tissues during murine listeriosis. *Microbes Infect.* **6**, 8–16 (2004).
  68. M. J. Horsburgh, J. L. Aish, I. J. White, L. Shaw, J. K. Lithgow, S. J. Foster,  $\sigma B$  modulates virulence determinant expression and stress resistance: Characterization of a functional *rsbU* strain derived from *Staphylococcus aureus* 8325-4. *J. Bacteriol.* **184**, 5457–5467 (2002).
  69. K. E. Foulds, L. A. Zenewicz, D. J. Shedlock, J. Jiang, A. E. Troy, H. Shen, Cutting edge: CD4 and CD8 T cells are intrinsically different in their proliferative responses. *J. Immunol.* **168**, 1528–1532 (2002).
  70. J. D. Ooi, J.-H. Jiang, P. J. Eggenhuizen, L. L. Chua, M. van Timmeren, K. L. Loh, K. M. O'Sullivan, P. Y. Gan, Y. Zhong, K. Tsyganov, L. R. Shochet, J. Ryan, C. A. Stegeman, L. Fugger, H. H. Reid, J. Rossjohn, P. Heeringa, S. R. Holdsworth, A. Y. Peleg, A. R. Kitching, A plasmid-encoded peptide from *Staphylococcus aureus* induces anti-myeloperoxidase nephritogenic autoimmunity. *Nat. Commun.* **10**, 3392 (2019).
  71. K. T. Bæk, D. Frees, A. Renzoni, C. Barras, N. Rodriguez, C. Manzano, W. L. Kelley, Genetic variation in the *Staphylococcus aureus* 8325 strain lineage revealed by whole-genome sequencing. *PLOS ONE* **8**, e77122 (2013).
  72. A. Louie, G. L. Drusano, P. Banerjee, Q. F. Liu, W. Liu, P. Kaw, M. Shayegani, H. Taber, M. H. Miller, Pharmacodynamics of fluconazole in a murine model of systemic candidiasis. *Antimicrob Agents Chemother.* **42**, 1105–1109 (1998).
  73. P. Kamran, K. I. Sereti, P. Zhao, S. R. Ali, I. L. Weissman, R. Ardehali, Parabiosis in mice: A detailed protocol. *J. Vis. Exp.* **80**, 50556 (2013).
  74. K. G. Anderson, K. Mayer-Barber, H. Sung, L. Beura, B. R. James, J. J. Taylor, L. Qunaj, T. S. Griffith, V. Vezys, D. L. Barber, D. Masopust, Intravascular staining for discrimination of vascular and tissue leukocytes. *Nat. Protoc.* **9**, 209–222 (2014).
  75. B. Rissiek, M. Lukowiak, F. Raczkowski, T. Magnus, H. W. Mittrücker, F. Koch-Nolte, In vivo blockade of murine ARTC2.2 during cell preparation preserves the vitality and function of liver tissue-resident memory T cells. *Front. Immunol.* **9**, 1580 (2018).
  76. M. Marouf, P. Machart, V. Bansal, C. Kilian, D. S. Magruder, C. F. Krebs, S. Bonn, Realistic in silico generation and augmentation of single-cell RNA-seq data using generative adversarial networks. *Nat. Commun.* **11**, 166 (2020).
  77. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck III, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).
  78. V. G. Tusher, R. Tibshirani, G. Chu, Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5116–5121 (2001).
  79. J. Tao, L. Mariani, S. Eddy, H. Maecker, N. Kambham, K. Mehta, J. Hartman, W. Wang, M. Kretzler, R. A. Lafayette, JAK-STAT signaling is activated in the kidney and peripheral blood cells of patients with focal segmental glomerulosclerosis. *Kidney Int.* **94**, 795–808 (2018).

**Acknowledgments:** FACS sorting was performed at the UKE FACS sorting core facility. Next-generation sequencing was performed at the Deutsche Forschungsgemeinschaft (DFG)-funded "Competence Centre for Genome Analysis Kiel" at the University of Kiel.

**Funding:** This study was supported by grants from the DFG (SFB1192 to C.F.K., F.K.-N., J.-E.T., S.H., N.G., V.G.P., E.H., C.M.-S., T.W., C.Ku., T.H., U.P., and H.-W.M.; SFB1286 to S.B.; SFBTR57 to C.Ku.), grants from the Deutsche Nierenstiftung and Deutsche Gesellschaft für Nephrologie to C.F.K., a grant from the Werner Otto Stiftung to L.U.B.E., a grant from the eMed Consortia "Fibromap" from the Federal Ministry of Education and Research to V.G.P., and grants Conicyt Proyecto AFB 170004 and Conicyt/FONDEQUIP/EQM140016 to M.R.B. The ERCB was supported by the Else Kröner-Fresenius Foundation. **Author contributions:** C.F.K., D.R., U.P., and H.-W.M. wrote the manuscript and designed figures. C.F.K., D.R., H.-J.P., H.R., C.Ku., F.K.-N., M.R.B., U.P., and H.-W.M. designed experiments. C.F.K., D.R., H.-J.P., P.B., S.N., M.V.R., M.H., A.B., L.U.B.E., M.Z., M.B., J.S., M.R., S.F., T.W., S.K., M.N.W., V.G.P., C.S., T.B., N.S., C.M.-S., and E.H. conducted experiments and acquired data. C.F.K., D.R., Y.Z., H.-J.P., P.B., C.Ki., C.D.C., M.T.L., and M.H. analyzed data. C.F.K., D.R., Y.Z., T.B.H., S.B., U.P. and H.-W.M. interpreted data. C.F.K. and H.-W.M. did statistical analysis. J.-E.T., J.-H.R., S.H., and N.G. edited the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The datasets generated and analyzed during the current study are available from the corresponding authors on reasonable request. All datasets used for figures are provided as Supplementary Materials. The single-cell gene expression count and metadata tables containing clustering and quality control metrics for each cell are available at FigShare: <https://figshare.com/s/7912de1afc7fd5bbefd4>.

Submitted 2 December 2019  
Resubmitted 11 May 2020  
Accepted 16 July 2020  
Published 7 August 2020  
10.1126/sciimmunol.aba4163

**Citation:** C. F. Krebs, D. Reimers, Y. Zhao, H.-J. Paust, P. Bartsch, S. Nuñez, M. V. Roseblatt, M. Hellmig, C. Kilian, A. Borchers, L. U. B. Enk, M. Zinke, M. Becker, J. Schmid, S. Klinge, M. N. Wong, V. G. Puelles, C. Schmidt, T. Bertram, N. Stumpf, E. Hoxha, C. Meyer-Schwesinger, M. T. Lindenmeyer, C. D. Cohen, M. Rink, C. Kurts, S. Franzenburg, F. Koch-Nolte, J.-E. Turner, J.-H. Riedel, S. Huber, N. Gagliani, T. B. Huber, T. Wiech, H. Rohde, M. R. Bono, S. Bonn, U. Panzer, H.-W. Mittrücker, Pathogen-induced tissue-resident memory T<sub>H</sub>17 (T<sub>RM</sub>17) cells amplify autoimmune kidney disease. *Sci. Immunol.* **5**, eaba4163 (2020).

## Pathogen-induced tissue-resident memory T<sub>H</sub>17 (T<sub>RM</sub>17) cells amplify autoimmune kidney disease

Christian F. Krebs, Daniel Reimers, Yu Zhao, Hans-Joachim Paust, Patricia Bartsch, Sarah Nuñez, Mariana V. Roseblatt, Malte Hellmig, Christoph Kilian, Alina Borchers, Leon U. B. Enk, Michael Zinke, Martina Becker, Joanna Schmid, Stefanie Klinge, Milagros N. Wong, Victor G. Puelles, Constantin Schmidt, Tabea Bertram, Natascha Stumpf, Elion Hoxha, Catherine Meyer-Schwesinger, Maja T. Lindenmeyer, Clemens D. Cohen, Michael Rink, Christian Kurts, Sören Franzenburg, Friedrich Koch-Nolte, Jan-Eric Turner, Jan-Hendrik Riedel, Samuel Huber, Nicola Gagliani, Tobias B. Huber, Thorsten Wiech, Holger Rohde, Maria Rosa Bono, Stefan Bonn, Ulf Panzer and Hans-Willi Mittrücker

*Sci. Immunol.* **5**, eaba4163.  
DOI: 10.1126/sciimmunol.aba4163

### Autoimmunity promotor

Tissue-resident memory T (T<sub>RM</sub>) cells are involved in peripheral immunity against reinfection, but their role in autoimmunity is unclear. Krebs *et al.* examine the contribution of T<sub>RM</sub> cells in patients with antineutrophil cytoplasmic antibody (ANCA)-dependent glomerulonephritis (GN). They identified multiple T cell subsets in healthy kidney tissue biopsies, but a marked increase in CD4<sup>+</sup> T<sub>RM</sub> cells was seen in kidney biopsies from patients with ANCA-GN. They infected mice with *Staphylococcus aureus*, which induced renal T<sub>H</sub>17 cells that had a T<sub>RM</sub> cell phenotype and persisted in kidney tissue. In a mouse model of crescentic GN, *S. aureus* infection aggravated kidney pathology and appeared to drive localized renal autoimmune responses. These findings provide critical insight into the role of CD4<sup>+</sup> T<sub>RM</sub> cells in contributing to autoimmune disease.

### ARTICLE TOOLS

<http://immunology.sciencemag.org/content/5/50/eaba4163>

### SUPPLEMENTARY MATERIALS

<http://immunology.sciencemag.org/content/suppl/2020/08/03/5.50.eaba4163.DC1>

### REFERENCES

This article cites 79 articles, 19 of which you can access for free  
<http://immunology.sciencemag.org/content/5/50/eaba4163#BIBL>

Use of this article is subject to the [Terms of Service](#)

*Science Immunology* (ISSN 2470-9468) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Immunology* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

## Supplementary Materials for

### Pathogen-induced tissue-resident memory T<sub>H</sub>17 (T<sub>RM</sub>17) cells amplify autoimmune kidney disease

Christian F. Krebs\*, Daniel Reimers, Yu Zhao, Hans-Joachim Paust, Patricia Bartsch, Sarah Nuñez, Mariana V. Roseblatt, Malte Hellmig, Christoph Kilian, Alina Borchers, Leon U. B. Enk, Michael Zinke, Martina Becker, Joanna Schmid, Stefanie Klinge, Milagros N. Wong, Victor G. Puelles, Constantin Schmidt, Tabea Bertram, Natascha Stumpf, Elion Hoxha, Catherine Meyer-Schwesinger, Maja T. Lindenmeyer, Clemens D. Cohen, Michael Rink, Christian Kurts, Sören Franzenburg, Friedrich Koch-Nolte, Jan-Eric Turner, Jan-Hendrik Riedel, Samuel Huber, Nicola Gagliani, Tobias B. Huber, Thorsten Wiech, Holger Rohde, Maria Rosa Bono, Stefan Bonn, Ulf Panzer, Hans-Willi Mittrücker\*

\*Corresponding author. Email: c.krebs@uke.de (C.F.K.); h.mittruecker@uke.de. (H.-W.M.)

Published 7 August 2020, *Sci. Immunol.* 5, eaba4163 (2020)

DOI: 10.1126/sciimmunol.aba4163

#### The PDF file includes:

##### Material and Methods

Fig. S1. scRNA-seq analysis of T cells from human blood and healthy kidney tissue from tumor nephrectomies.

Fig. S2. Profiling of T cells from human blood and healthy kidney tissue from tumor nephrectomies.

Fig. S3. scRNA-seq analysis of T cells from blood and kidney of patients with ANCA-GN.

Fig. S4. Profiling of T cells from kidney and blood of patients with ANCA-GN.

Fig. S5. Flow cytometry gating strategy for murine T cell analysis.

Fig. S6. Prior *L. monocytogenes* infection does not affect crescentic GN.

Fig. S7. Effect of local *S. aureus* skin infection on renal T<sub>RM</sub>17 cell generation and effect of systemic *S. aureus* infection on imiquimod-induced skin inflammation.

Fig. S8. Gene expression analysis of renal CD4<sup>+</sup> T cells from nephritic mice previously infected with *S. aureus*.

Fig. S9. The T<sub>H</sub>17-associated pathogens *C. albicans* and *E. coli* cause renal T<sub>RM</sub>17 cell formation.

Fig. S10. Conceptual link of pathogen-induced tissue resident memory T cells with autoimmunity.

Table S1. Baseline characteristics of controls at time of biopsy.

Table S2. Baseline characteristics of patients with ANCA-GN at the time of biopsy.

Table S3. Clinical characteristics and ANCA-serotype of patients from the Hamburg GN Registry.

Table S4. List of transcription factors identified by TRRUSTv2 analysis.

References (68–79)

**Other Supplementary Material for this manuscript includes the following:**

(available at [immunology.sciencemag.org/cgi/content/full/5/50/eaba4163/DC1](http://immunology.sciencemag.org/cgi/content/full/5/50/eaba4163/DC1))

Data file S1 (Microsoft Excel format).

## Material and Methods

### *S. aureus*, *L. monocytogenes*, *C. albicans* and *E. coli* infection of mice

Mice were infected with *Staphylococcus aureus* strain (SH1000) (68), an ovalbumin-recombinant *Listeria monocytogenes* strain (LmOVA) (69), or with *Candida albicans* (CEC1293) (57). The SH1000 strain does not code for the 6-phosphogluconate dehydrogenase protein which is found in some *S. aureus* isolates and can induce T and B cell responses in mice cross reactive to myeloperoxidase (70, 71). Mice received  $1 \times 10^7$  staphylococci in 200  $\mu$ l sterile PBS via the lateral tail vein. For *S. aureus* skin infection,  $1 \times 10^6$  staphylococci were injected intradermally into the right flank (42). Bacterial inocula were controlled by plating serial dilution onto LB agar plates. To clear the *S. aureus* infection, infected mice and corresponding uninfected controls were treated for one week with 1 mg/ml ampicillin in the drinking water. *L. monocytogenes* ( $2 \times 10^4$  in 200  $\mu$ l sterile PBS) was applied via the lateral tail vein. For quantification of bacterial titers, organs were homogenized in PBS and serial dilutions of suspensions were plated on LB agar. Plates were incubated at 37°C and colony forming units (CFU) were counted the next day.

*C. albicans* was cultivated on Sabouraud dextrose agar. For infection, fresh YPD medium was inoculated with an overnight culture and incubated for 3h at 27°C, 100 rpm. Cultures in log phase were harvested and  $1 \times 10^3$  or  $1 \times 10^4$  fungi were administered in sterile PBS by tail vein injection. The mice were treated twice with fluconazole (7.5 mg/kg i.p.) to clear fungal infection (72).

For urinary tract infection, mice were infected with uropathogenic *E. coli* (strain 536). This bacterial strain originates from a patient undergoing urinary tract infection. Mice were anesthetized by Ketamin/Rompun and infected transurethral via catheter with  $1 \times 10^{10}$  *E. coli* (56).

### Reconstitution of *Rag1*<sup>-/-</sup> with CD4<sup>+</sup> T cells

Lymphocytes were isolated from spleen and inguinal lymph nodes of naïve C57BL/6 mice and CD4<sup>+</sup> T cells were enriched with the EasySep™ Mouse Naïve CD4 kit (StemCell Technologies, Vancouver, Canada) according to the manufacturer's protocol. Enrichment of CD4<sup>+</sup> T cells was controlled with flow cytometry and consistently reached >97%. *Rag1*<sup>-/-</sup> were reconstituted i.v. with  $10^6$  CD4<sup>+</sup> T cells.

### Experimental crescentic glomerulonephritis

The nephrotoxic nephritis model was used to induce crescentic glomerulonephritis. Male mice were i.p. injected with sheep serum directed against the glomerular basement membrane (2.5 mg/g bodyweight), as previously described (11). Unless indicated otherwise, mice were analyzed 10 days later. Urinary albumin excretion was determined by standard ELISA analysis (Mice-Albumin Kit, Bethyl, Montgomery, TX), while urinary creatinine, and serum BUN were measured using standard laboratory methods.

IL-17A or IFN- $\gamma$  was neutralized by i.p. injection of 500  $\mu$ g of anti-IL-17A mAb (clone: 17F3, BioXCell, West Lebanon, NH) or 400  $\mu$ g anti-IFN- $\gamma$  (clone XMG1.2, BioXCell) in 200  $\mu$ l PBS on days -1, 3, and 6 after induction of crescentic GN with serum. Inflammatory cytokines were neutralized by i.p. injection of a combination of mAbs against IL-1 $\beta$  (200  $\mu$ g, clone B122, BioXCell), IL-6 (300  $\mu$ g, clone MP5-20F3, BioXCell) and IL-12/23p40 (300  $\mu$ g, clone R1-5D9, BioXCell) in 200  $\mu$ l PBS on days 0, 3, and 6 after induction of crescentic GN. Control mice received control IgG (clone MOPC-21, BioXCell) at the same time points and with similar concentrations. *Il17a*<sup>Cre</sup> $\times$ *R26*<sup>DTR</sup> $\times$ *R26*<sup>eYFP</sup> mice were treated with 0.25  $\mu$ g

diphtheria toxin (Merck, Darmstadt, Germany) in 100 µl PBS i.p. on days 60, 62, and 64 after *S. aureus* infection. Efficacy of DT treatment was controlled in the kidney (Fig. 6G).

#### *Experimental Autoimmune Glomerulonephritis*

To induce experimental autoimmune glomerulonephritis (EAG), male DBA/1 mice were immunized s.c. with 30 µg of the murine non-collagenous domain of the  $\alpha 3$ -chain of type IV collagen (m- $\alpha 3$ IVNC1) in complete Freund's adjuvant and boosted after 3, 5 and 7 weeks with 30 µg of the protein in incomplete Freund's adjuvant. The production of recombinant  $\alpha 3$ IVNC1 has been described previously (40).

#### *Photoconversion of intra-renal cells in Kaede mice*

Transgenic *Kaede* mice were anaesthetized with isoflurane (4 vol.-%) and buprenorphine (0.1 mg/kg body weight). The left kidney was accessed via the left flank and gently located on sterile drapes. The anterior and posterior sides of the kidney were each exposed to near UV-A light (385 nm) from 4 cm distance for 30 seconds using an LED light source (Blue Wave LED Prime UVA, Dymax, Wiesbaden, Germany). This corresponds to an intensity of 11.2 J/cm<sup>2</sup>. Subsequently, the kidney was reposed, the wound was closed with sutures and mice received post-interventional analgesia with tramadol per drinking water *ad libitum* (12.5 mg/100 ml).

#### *Surgical connection of the circulation system of mice (parabiosis)*

CD45.1 mice were infected with *S. aureus* and after 2 weeks infected and non-infected CD45.2 C57BL/6 control mice were treated for one week with ampicillin in the drinking water. On day 30 post infection, two mice (CD45.1 and CD45.2) were surgically joined according to standard protocols (73). After 28 days, mice were sacrificed and analyzed separately.

#### *Imiquimod-induced psoriasis*

Imiquimod (IMQ)-induced skin inflammation is regarded as an IL-17 dependent mouse model for psoriasis (43). Mice were infected with *S. aureus* and after 2 weeks infected and non-infected control mice were treated for one week with ampicillin in the drinking water. On day 30 post infection, one ear was treated daily with 5 mg IMQ cream (5% IMQ, MEDA Pharma, Bad Homburg, Germany). The other ear was treated with Vaseline (control). Thickness of ears was measured daily with a caliper. From this a score was calculated with: < 20% above baseline :0, < 35%: 1, < 50%: 2, < 65 %: 3, > 65%: 4. In addition, ears were daily scored for erythema and desquamation. Both parameters were independently scored with: 0: no, 1: mild, 2: moderate, 3: strong and 4: severe symptoms. The psoriasis score represents the mean of all three scores.

#### *Histology, immunofluorescence and electron microscopy*

Glomerular crescent formation was assessed in a blinded fashion in 30 glomeruli per mouse in periodic acid–Schiff (PAS)–stained paraffin sections (11).

For the evaluation of *Kaede* photoconversion in cells of the kidney, tissue sections were subjected to confocal imaging without further preparation. Kidney samples were snap-frozen in Tissue-Tec (Sakura Finetek, Alphen aan den Rijn, Netherlands) and stored at -80°C. All slides were evaluated under an Axioskop light microscope and photographed with an

Axiocam HRc (Zeiss, Jena, Germany) using the ZEN software or by confocal microscopy with an A1R using NIS-Element software (Nikon Instruments, Amsterdam, Netherlands).

For Immunohistochemistry human paraffin-embedded kidney sections (2  $\mu\text{m}$ ) from renal biopsies of patients with ANCA-GN were stained with an antibody directed against CD69 (clone 8B6, ThermoFischer, USA). Renal CD69<sup>+</sup> cells in 3 high-power fields per renal biopsy (magnification  $\times 400$ ) were counted.

Immunofluorescence microscopy was performed in 1  $\mu\text{m}$  paraffin-embedded sections, following 15-minute antigen retrieval with pH9 antigen retrieval solution (Agilent, Santa Clara, CA) and incubation with primary anti-CD4 mAb (ab133616, Abcam, Cambridge, UK), and anti-CD69 Ab (ab202909, Abcam). Images were captured using a laser confocal microscope (LSM800, Zeiss, Jena, Germany). Indirect immunofluorescence was performed using primary antibodies against GFP (ab290, Abcam), synaptopodin (163004, Synaptic Systems, Göttingen, Germany), and endomucin (sc-65495, Santa Cruz Biotechnology, Dallas, TX).

For electron microscopy the part of the mouse kidney selected for electron microscopy was transferred from 4% formaldehyde into a cacodylate buffer with sucrose for 10 min at 80°C. Afterwards, osmium tetroxyde was applied for 2 h. The specimen was washed in cacodylate buffer plus sucrose two times for 5 min. Subsequently, the sample was contrasted with uranyl acetate for 1 h. The specimen was put into ethanol baths with rising ethanol concentrations for 5 min each, followed by bathing in Methyl tert-butyl ether (MTBE) plus epoxide mixture (in a 1:3 dilution) twice for 5 min each. Afterwards, the specimens were embedded in an epoxide mixture at 60°C for 48 h and then at 100°C for 11½ h. Semithin and ultrathin sections were cut on a Leica Microsystems microtome (Wetzlar, Germany). Grids were purchased from Polyscience (Cham, Germany). The grids were then analyzed using electron microscopes (EM 109 and EM 902, Zeiss, Jena, Germany) equipped with digital electron microscope cameras (Tröndle, Moorenweis, Germany). 4 glomeruli from each mouse were analyzed.

#### *Quantitative rt-PCR of renal mRNA*

Total RNA from the renal cortex was isolated with the NucleoSpin Kit (Macharey-Nagel, Düren, Germany) in accordance with the manufacturer's protocol. RNA was reverse transcribed with the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher, Waltham, MA). Real-time qPCR was performed using specific primers or QuantiTect Primer assay for *Il17a* (QT00103278, Qiagen, Hilden Germany). *Il22* was determined with TaqMan PCR (primers: Mm00444241, Life Technologies, Carlsbad, CA). The measurement was performed on a StepOnePlus Real-Time PCR system (Thermo Fisher, Waltham, MA) with the 18S rRNA as a housekeeping gene.

#### *Primers*

Primer (sequence 5' > 3')

*18S frw*: CAC GGC CGG TAC AGT GAA AC

*18S rev*: AGA GGA GCG AGC GAC CAA A

*Cxcl1 frw*: GCA CCC AAA CCG AAG TCA TAG

*Cxcl1 rev*: CAA GGG AGC TTC AGG GTC AA

*Cxcl2 frw*: CAC TGC GCC CAG ACA GAA

*Cxcl2 rev*: CAG GGT CAA GGC AAA CTT TTT G

*Cxcl5 frw*: TGG CAT TTC TGT TGC TGT TCA

*Cxcl5 rev*: TTG CGG CTA TGA CTG AGG AA

*Cxcl9 frw*: CCC AAG CCC CAA TTG CA

*Cxcl9 rev:* GCA GGT TTG ATC TCC GTT C  
*Il-1 $\beta$  frw:* CCT TCC AGG ATG AGG ACA TGA  
*Il-1 $\beta$  rev:* TCA TCC CAT GAG TCA CAG AGG AT  
*Il6 frw:* TGG AAA TCG TGG AAA TGA GAA  
*Il6 rev:* AAG TGC ATC ATC GTT GTT CTA CA  
*Il23a frw:* TCC CTA CTA GGA CTC AGC CAA C  
*Il23a rev:* TGG GCA TCT GTT GGG TCT  
*Rorc frw:* ACC TCT TTT CAC GGG AGG A  
*Rorc rev:* TCC CAC ATC TCC CAC ATT G

#### *Isolation and flow cytometric analysis of human and murine lymphocytes*

Single-cell suspensions were obtained from human kidney biopsies by enzymatic digestion in RPMI 1640 medium with collagenase D (0.4 mg/ml, Roche, Mannheim, Germany) and DNase I (10  $\mu$ g/ml, Sigma-Aldrich, Saint Louis, MO) at 37°C followed by dissociation with gentleMACS (Miltenyi Biotec, Bergisch Gladbach, Germany). Leukocytes from blood samples were separated using Leucosep tubes (Greiner Bio-One, Kremsmünster, Austria). Samples were filtered through a 30  $\mu$ m filter (Partec, Görlitz, Germany) before antibody staining and flow cytometry.

For the analysis of murine cells, intravascular (CD45<sup>iv+</sup>) cells were marked by i.v. injection of fluorochrome-conjugated anti-CD45 mAb (clone: 30-F11, 2.5  $\mu$ g per mouse) 3 min before sacrificing the mice (74). Cells from murine spleens were isolated by squashing the organ through a 70  $\mu$ m cell strainer. Erythrocytes were lysed using lysis buffer (155 mM NH<sub>4</sub>Cl, 10 mM KHCO<sub>3</sub>, 10  $\mu$ M EDTA, pH 7.2). Kidneys were enzymatically digested with 400  $\mu$ g/ml collagenase D (Roche, Mannheim, Germany) and 10 U/ml DNase I (Sigma-Aldrich, St. Louis, MO) for 45 min at 37°C. Subsequently, leukocytes were isolated by density gradient centrifugation using 37% Easycoll (Merck Millipore, Darmstadt, Germany) and a filtration step using a 30  $\mu$ m cell strainer (Partec, Görlitz, Germany).

For intranuclear staining, samples were processed using the Foxp3-Kit for intranuclear staining according to the manufacturer's instructions (Foxp3-Kit; eBioscience). To induce cytokines, spleen or kidney cells were incubated in RPMI 1640 medium supplemented with 5% FCS, L-glutamine, pyruvate, gentamicin and 2-mercaptoethanol with phorbol 12-myristate 13-acetate (PMA, 50 ng/ml, Sigma Aldrich, St. Louis, MO) and ionomycin (1  $\mu$ M, Sigma Aldrich). After 30 min, brefeldin A (10  $\mu$ g/ml, Sigma Aldrich) was added to prevent cytokine secretion. After further 3.5 h of culture, cytokine expression was determined by intracellular staining and flow cytometry.

To minimize unspecific antibody binding, mouse cells were incubated with 10  $\mu$ g/ml 2.4G2 (anti-Fc $\gamma$ RII/III; BioXcell, West Lebanon, NH) and 1:100 rat serum in PBS. Staining was performed on ice with fluorochrome-conjugated mAbs (see antibody table). Cells were surface stained with mAb and a fixable dead cell stain (pacific orange succinimidyl ester or live/dead<sup>TM</sup> fixable near-IR dead cell stain kit; Life Technologies, Carlsbad, CA) to exclude dead cells from analysis. For staining of cytokines and transcription factors, cells were fixed and permeabilized using the Foxp3 transcription factor staining buffer kit (eBioscience, San Diego, CA). Cells were incubated in permeabilization buffer with 1:100 mouse serum. After 10 min, mAb were added and after a further 30 min at 4°C, cells were washed with permeabilization buffer.

#### *Flow cytometry and cell sorting*

Samples were measured with a FACSCanto II, FACSCelesta, FACS LSR II, or LSRFortessa flow cytometer (all BD Biosciences, San Jose, CA). Data analysis was performed using the FlowJo software (Treestar, Ashland, OR) or the FACSDiva software (BD Biosciences). Representative gating strategies are given in fig. S5A and B.

FACS-sorting was performed on a FACS AriaFusion or AriaIIIu (BD Biosciences). For pre-enrichment of murine CD4<sup>+</sup> T cells MACS (Miltenyi Biotech, Bergisch Gladbach, Germany) or EasySep (StemCell Technologies, Vancouver, Canada) isolation kits were used according to the manufacturers' protocols. In both cases, negative enrichment without staining of T cells or CD4<sup>+</sup> T cells was used.

#### *In vitro stimulation T cells and measurement of cytokines*

Anti-Art2b nanobodies (clone: S+16, 50 µg in 100 µl PBS) were injected i.p. 30 min before sacrificing the mice to prevent NAD-induced cell death (75). CD4<sup>+</sup> T cells from spleen and kidney of mice were isolated and purified using magnetic enrichment and FACS sorting as described before.  $6 \times 10^3$  CD4<sup>+</sup> T cells were incubated in a volume of 40 µl of IMDM medium containing 10% FCS, streptomycin, penicillin, 2-mercaptoethanol and mIL-2 (10 ng/ml) and mIL-7 (10 ng/ml). All cytokines were purchased from Biolegend, San Diego, CA. Cells were stimulated with either anti-CD3 mAb (2 µg/ml, clone: 145-2C11, Biolegend) and anti CD28 mAb (1 µg/ml, clone: 37.51, Biolegend) or with a cocktail of mIL-1β (20 ng/ml), mIL-6 (20 ng/ml) and mIL-23 (40 ng/ml). After 3 days, cytokines in the supernatant were determined using the Legendplex assay with the mouse inflammation panel (Biolegend) according to the manufacturer's protocol.

*Table: Antibodies used in this study*

Antigen	Clone	Fluorochrome	Provider
human CD3ε	OKT3	BV785	Biolegend
human CD4	RPA-T4	BV650	Biolegend
human CD4	ab133616	- (Immunohistology)	Abcam
human CD8α	RPA-T8	AF700, APC	BD
human CD45	HI30	BV510	Biolegend
human CD45RA	HI100	BUV737	BD
human CD69	FN50	BUV395	Biolegend
human IL-17A	KL168	BV421	Biolegend
human IFN-γ	4S.B3	AF488	Biolegend
human CD69	ab202909	- (Immunohistology)	Abcam
human CD69	8B6	- (Immunohistology)	ThermoFisher
human CD103	Ber-ACT8	FITC	Biolegend
human CD183/CXCR3	G025H7	PE-Dazzle	Biolegend
human CD196/CCR6	G034E3	PerCP-Cy5.5	Biolegend
human CD197/CCR7	G043H7	BV605	Biolegend
mouse CD3ε	eBio500A2	eFluor450, BV510	eBioscience
mouse CD4	RM4-5	BV605, BV650 AF700, BV785, APC	Biolegend, eBioscience
mouse CD8α	53-6.7	BV650, BV711	Biolegend
mouse CD16+CD32	2.4G2	- (Fc block)	BioXCell
mouse CD44	IM7	APC	eBioscience
mouse CD45	30-F11	PerCP, AF700, APCCy7	Biolegend
mouse CD69	H1.2F3	BV785, PECy7	Biolegend
mouse CD103	M290	AF700	BD

mouse CD40L	MR1	APC	Biolegend
mouse IFN- $\gamma$	XMG1.2	APCCy7	BD
mouse IL-17A	TC11-18H10.1	BV785, PE	Biolegend
mouse ROR $\gamma$ t	Q31-378	PE, APC	BD
mouse FOXP3	FJK-16S	FITC	eBioscience
mouse Tbet	4B10	BV421	Biolegend
mouse CD3 $\epsilon$	145-2C11	- (stimulation)	Biolegend
mouse CD28	37.51	- (stimulation)	Biolegend
mouse IL-1 $\beta$	B122	- (neutralizing)	BioXCell
mouse IL-6	MP5-20F3	- (neutralizing)	BioXCell
mouse IL-12B/p40	R1-5D9	- (neutralizing)	BioXCell
mouse IL-17A	17F3	- (neutralizing)	BioXCell
mouse IFN- $\gamma$	XMG1.2	- (neutralizing)	BioXCell
mouse IgG	MOPC-21	- (isotype control)	BioXCell
Synaptodin	Polyclonal	- (Immunohistology)	Synaptic Systems
Endomucin	V.7C7	- (Immunohistology)	Santa Cruz
GFP/YFP	Polyclonal	- (Immunohistology)	Abcam

### *Single cell RNA sequencing and epitope measurement*

To perform scRNA-seq, single cell suspensions were obtained from human and mouse samples as described above. Cells were simultaneously labelled with fluorescence-conjugated mAb and barcoded mAb (all Biolegend, San Diego, CA). For a list of barcoded antibodies used for this study see fig. S2 and S4. For scRNA-seq, FACS-sorted CD3<sup>+</sup> T cells underwent droplet-based single cell analysis and transcriptome library preparation using the Chromium Single Cell 3' Reagent Kits v2 according to the manufacturer's protocols (10x Genomics, Pleasanton, CA). Library construction for epitope measurement (CITE-seq) was performed according to Total-seq protocol (Biolegend, San Diego, CA). The generated scRNA-seq libraries were sequenced using a NovaSeq6000 system (100 cycles) (Illumina, San Diego, CA).

### *Quality control and pre-processing of single-cell RNA-sequencing data*

Quality control and scRNA-seq pre-processing was performed as previously described (76). In brief, the Cell Ranger software pipeline (version 2.2.0, 10x Genomics, Pleasanton, CA) was used to demultiplex cellular barcodes and map reads to the reference genome (refdata-cellranger-hg19-1.2.0) (function cellranger count). The matrices of cells and the unique molecular identifier (UMI) count were obtained and further processed by the R package Seurat (version 3.0.0.9000). As a quality-control (QC) step, we first filtered out genes detected in less than 3 cells and those cells in which less than 200 genes had nonzero counts. To remove potential doublets (two or more cells per droplet), we excluded cells with total number of detected genes more than 2000. Following visual inspection of the distribution of cells by the percentage of mitochondrial genes expressed, we further removed low-quality cells with more than 10% mitochondrial genes of all detected genes. In order to avoid potential contamination of non-T cells during sorting, only cells with an average count of CD3D, CD3E and CD3G of 1 or higher were kept for further analysis.

### *Dimensionality reduction and clustering*

The Seurat package (version 3.0.0.9000) was used to perform unsupervised clustering analysis on scRNA-seq data (77). In brief, gene counts for cells that passed QC were normalized by library size and log-transformed (function `NormalizeData`, `normalization.method = "LogNormalize"`, `scale.factor = 10000`). Then, highly variable genes were detected (function `FindVariableFeatures`, `selection.method = "vst"`, `nfeatures = 2000`). To reduce batch effects, we apply the integration method implemented in the latest Seurat version 3 (function `FindIntegrationAnchors` and `IntegrateData`, `dims = 1:30`). The integrated matrix was then scaled by `ScaleData` function (default parameters). Principal component analysis was performed on the scaled data (function `RunPCA`, `npcs = 30`) in order to reduce dimensionality. 30 principal components were determined by using the `ElbowPlot` function and used to compute the KNN graph based on the euclidean distance (function `FindNeighbors`), which then generated cell clusters using function `FindClusters` (`resolution = 0.6`). T-distributed Stochastic Neighbor Embedding (tSNE) was used to visualize clustering results. The top differential expressed genes in each cluster were found using the `FindAllMarkers` function (`min.pct = 0.1`) and ran Wilcoxon rank sum tests. The differential expression between clusters were calculated by `FindMarkers` function (`min.pct = 0.1`), which also ran Wilcoxon rank sum tests.

#### *CITE-seq data processing and integration*

Each monoclonal antibody had a corresponding barcode (16 nucleotides) (Biolegend, San Diego, CA). A pseudo-reference genome for CITE-seq was built with `cellranger mkref` function in Cell Ranger (version 2.2.0). We used the name of the antibody target protein as pseudo-gene name and 16 nucleotide barcode as pseudo-gene sequence. Then the CITE-seq raw data were aligned to this pseudo-reference genome using Cell Ranger function `cellranger count`. The output files are cells and ADT (Antibody-derived tags) count matrices. ADT data were integrated to single cell RNA-seq data by creating another assay slots in Seurat (function `CreateAssayObject`). The raw ADT counts were normalized using centered log-ratio (CLR) method (Seurat function `NormalizeData`, `method = "CLR"`) (27, 77). After normalization, the ADT counts were further scaled with `ScaleData` function (default parameters).

#### *Tissue resident memory scores and T subset marker genes*

Average expression levels of  $T_{RM}$  gene set ( $T_{RM}$  score) and non- $T_{RM}$  gene sets (non- $T_{RM}$  score) were calculated by Seurat function `AddModuleScore` (`nbin = 24`, `ctrl = 100`). Tissue resident memory ( $T_{RM}$ ) signature genes and non- $T_{RM}$  genes were obtained from a study of sorted  $CD4^+CD69^+$  and  $CD4^+CD69^{neg}$  T cells of lung, spleen and blood (Fig. 2C of Kumar et al., 2017).

$T_{RM}$  gene set: *CD69, CA10, IL17F, IL2, CDHR1, IL21, IL10, IL23R, CXCL13, CXCR6, KCNK5, ITGA1, JAG2, SRGAP3, TOX2, CH25H, NEK10, TMEM200A, MYO1B, PLXDC1, IKZF3, GFOD1, CRTAM, DUSP6, RGS1, TP53I11, GFII, IFNG, SLC7A5, GCNT4.*

Non- $T_{RM}$  gene set: *FAM65B, STK38, GRASP, KLF3, SAMD3, GABBR1, FRY, ARHGEF11, VIPR1, BAIAP3, MFGES8, SBK1, HAPLN3, TTC16, CX3CR1, USP46, PLXNA4, D4S234E, DSEL, CNTNAP1, VSIG1, RGMB, TTYH2, EPHA4, TNFRSF11A, MUC1, CR1, E2F2, KLF2, EDA, KRT73, ZNF462, RAP1GAP2, SIPR1, NPDC1, FLJ13197, ISM1, TSPAN18, KCTD15, KRT72, SEMA5A, WNT7A, SOX13, FUT7, PTGDS, P116, SEMA3G, SYT4.*

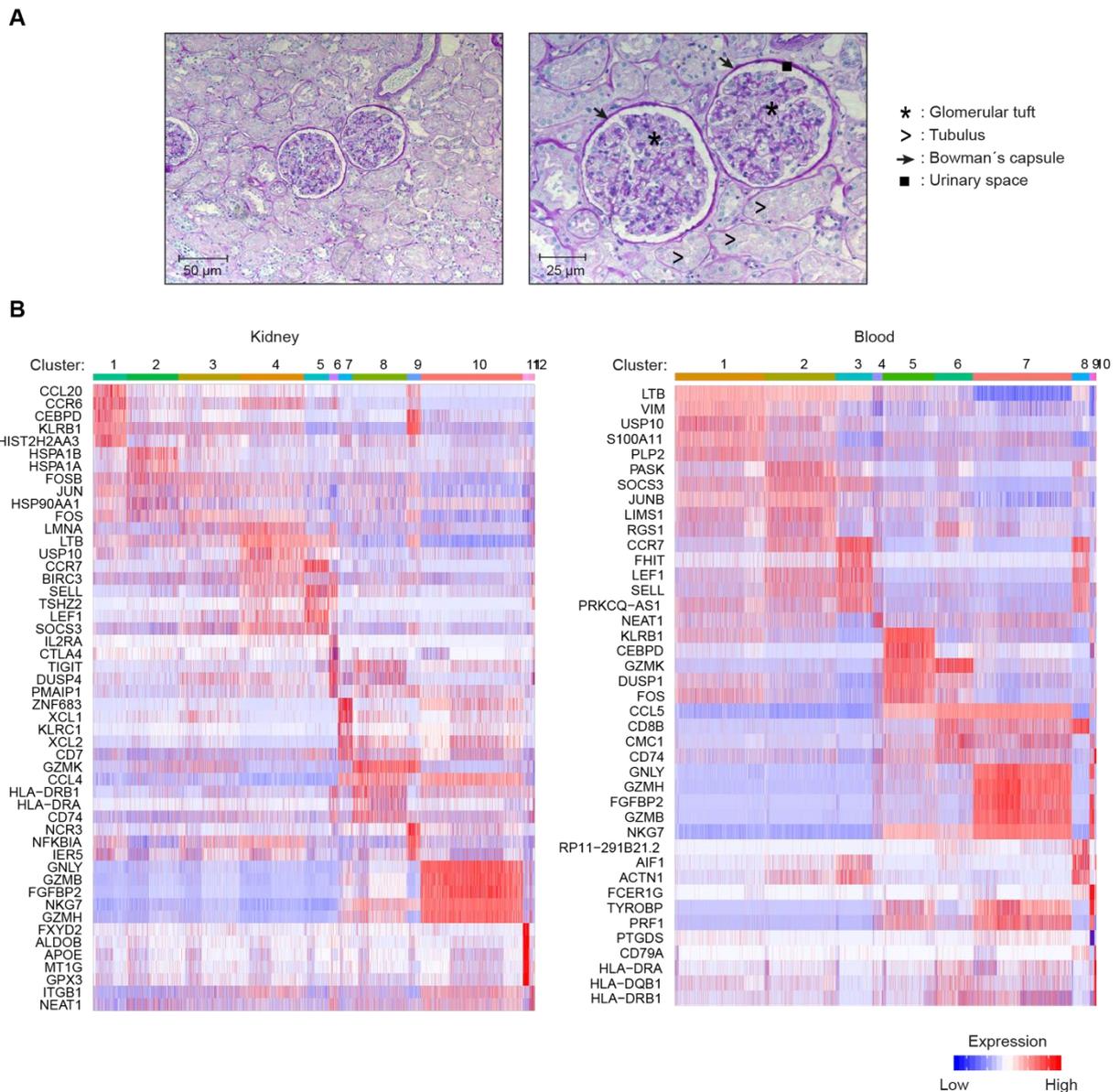
The marker gene list for  $CD4^+$  T cell subsets of  $T_H17$ ,  $T_H1$ ,  $T_H2$ ,  $T_H0$  and Treg cells) were curated based on a previous single cell study of human T cell subtypes (29). Low expressed genes (detected in less than 5% of the cells/cluster) were removed from the marker gene list.

Seurat functions `AverageExpression` and `DoHeatmap` were used to visualize the expression of the T cell subtype marker genes in each CD4<sup>+</sup> T cell cluster.

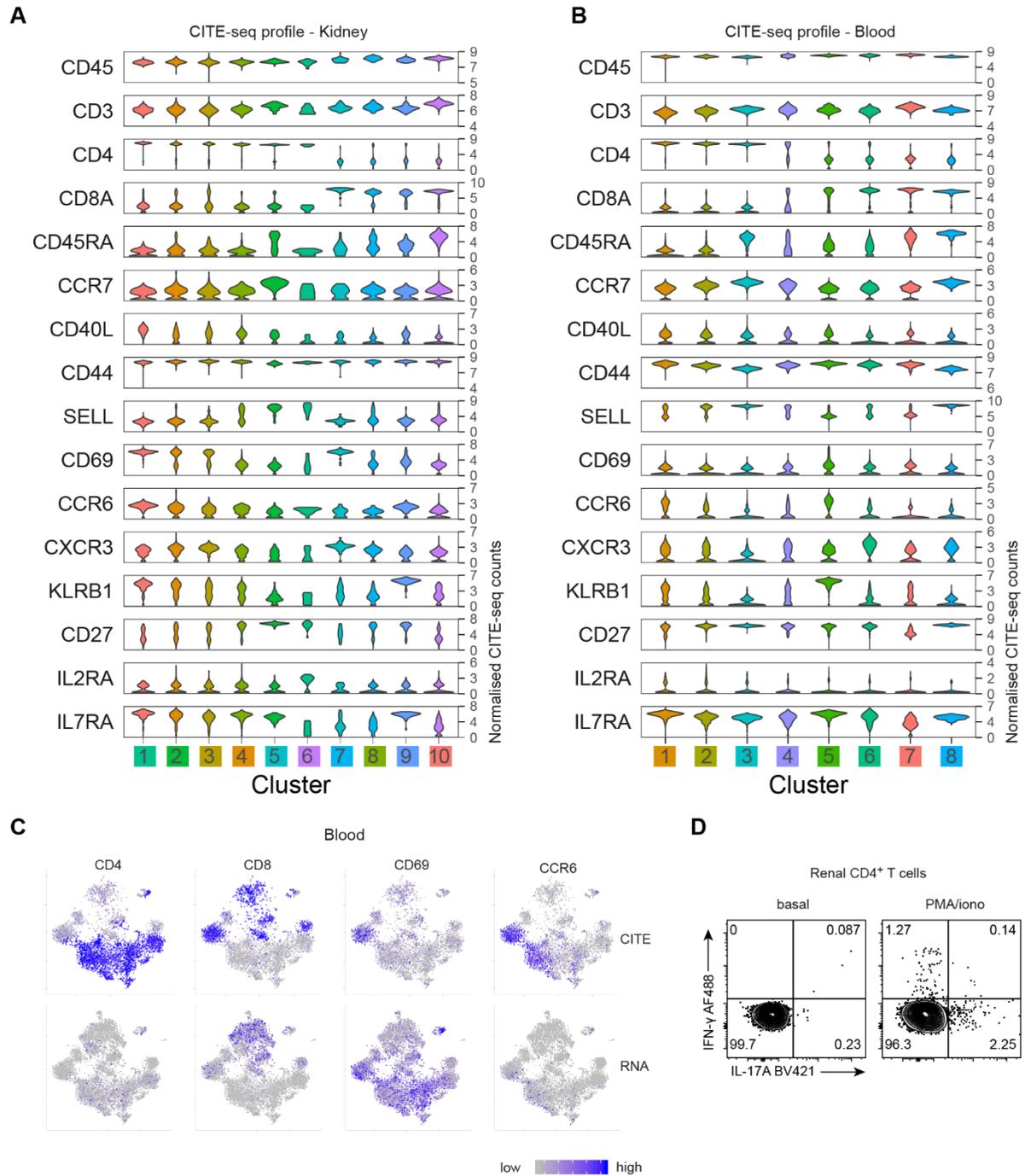
### *Human Microarray Analysis*

Human renal biopsy specimens and Affymetrix microarray expression data were obtained within the framework of the European Renal cDNA Bank - Kröner-Fresenius Biopsy Bank (30). Biopsies were obtained from patients after informed consent and with approval of the local ethics committees. Following renal biopsy, the tissue was transferred to RNase inhibitor and glomeruli were micro-dissected. Total RNA was isolated, reverse transcribed and amplified. Fragmentation, hybridization, staining and imaging were performed according to the Affymetrix Expression Analysis Technical Manual (Affymetrix, Santa Clara, CA, USA). Published datasets of glomerular samples were analyzed for mRNA expression levels of CD4<sup>+</sup> T<sub>RM</sub> cell signature genes (21). Analysis included datasets from patients with minimal change disease (MCD; n=14), focal segmental glomerulosclerosis (FSGS; n=23), membranous nephropathy (MGN; n=21), IgA nephropathy (n=27), lupus nephritis (SLE; n=32) and ANCA-associated glomerulonephritis (ANCA; n=23) as well as controls (living donors; n=42) (GSE 99340, GSE32591, GSE35489, GSE37463). CEL file normalization was performed with the Robust Multichip Average method using RMAExpress (Version 1.0.5) and the human Entrez Gene custom CDF annotation from Brain Array version 18 (<http://brainarray.mbni.med.umich.edu/Brainarray/default.asp>). To identify differentially expressed genes, the SAM (Significance Analysis of Microarrays) method (78) was applied using SAM function in Multiple Experiment Viewer (TiGR MeV, Version 4.9). A q-value below 5% was considered to be statistically significant.

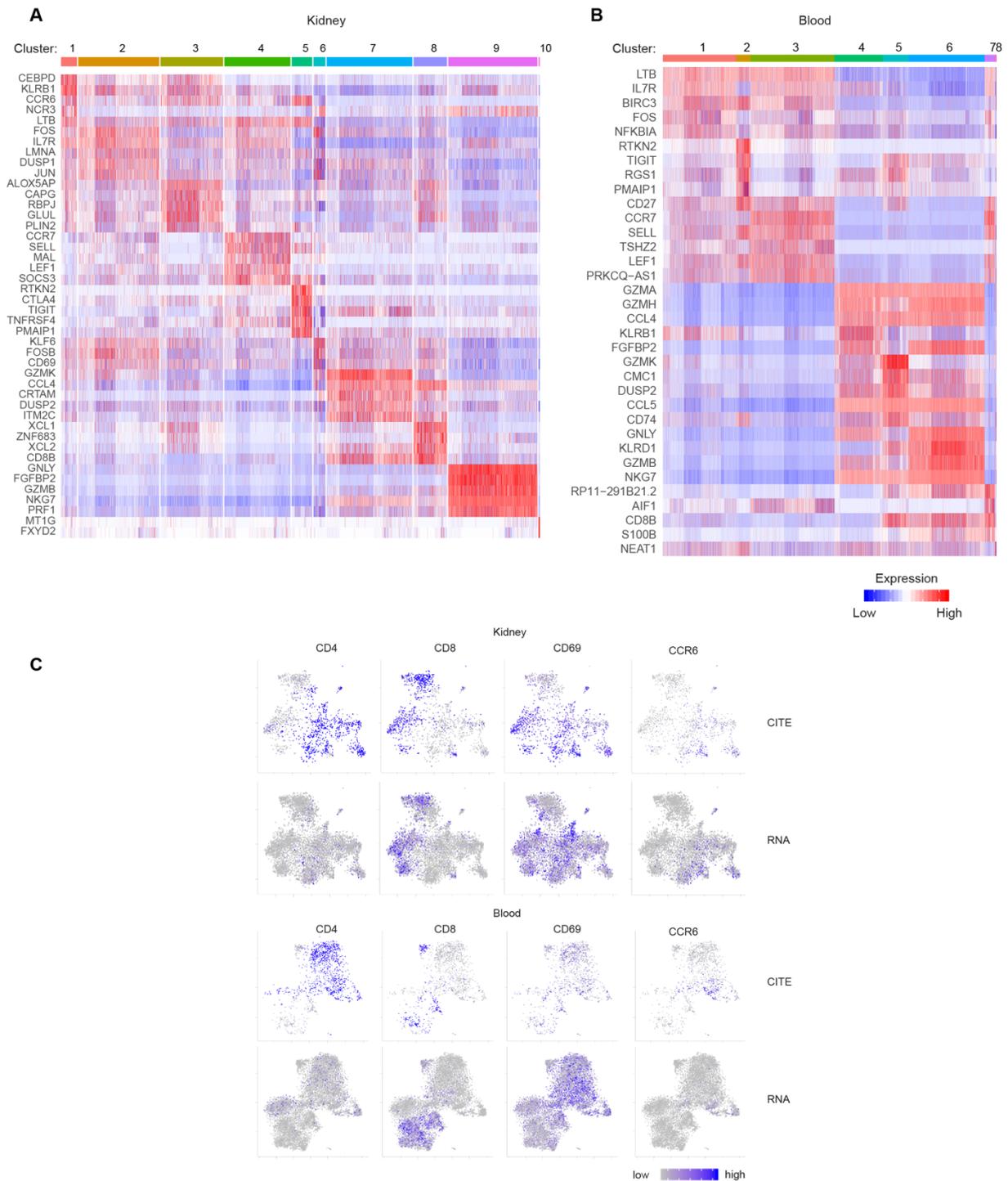
A T<sub>RM</sub> signature score was calculated as described previously (79). Of the 30 genes in the T<sub>RM</sub> gene set, 17 genes were above background. These gene sets were used to calculate a T<sub>RM</sub> signature score for each compartment. The T<sub>RM</sub> signature score was generated by transforming log<sub>2</sub> expression profiles into Z-scores, and averaging Z-scores of the T<sub>RM</sub> gene sets to generate a signature score for each patient sample. Activation scores for different renal disease represent the fold change of their mean scores compared to the mean score of controls (living donors).



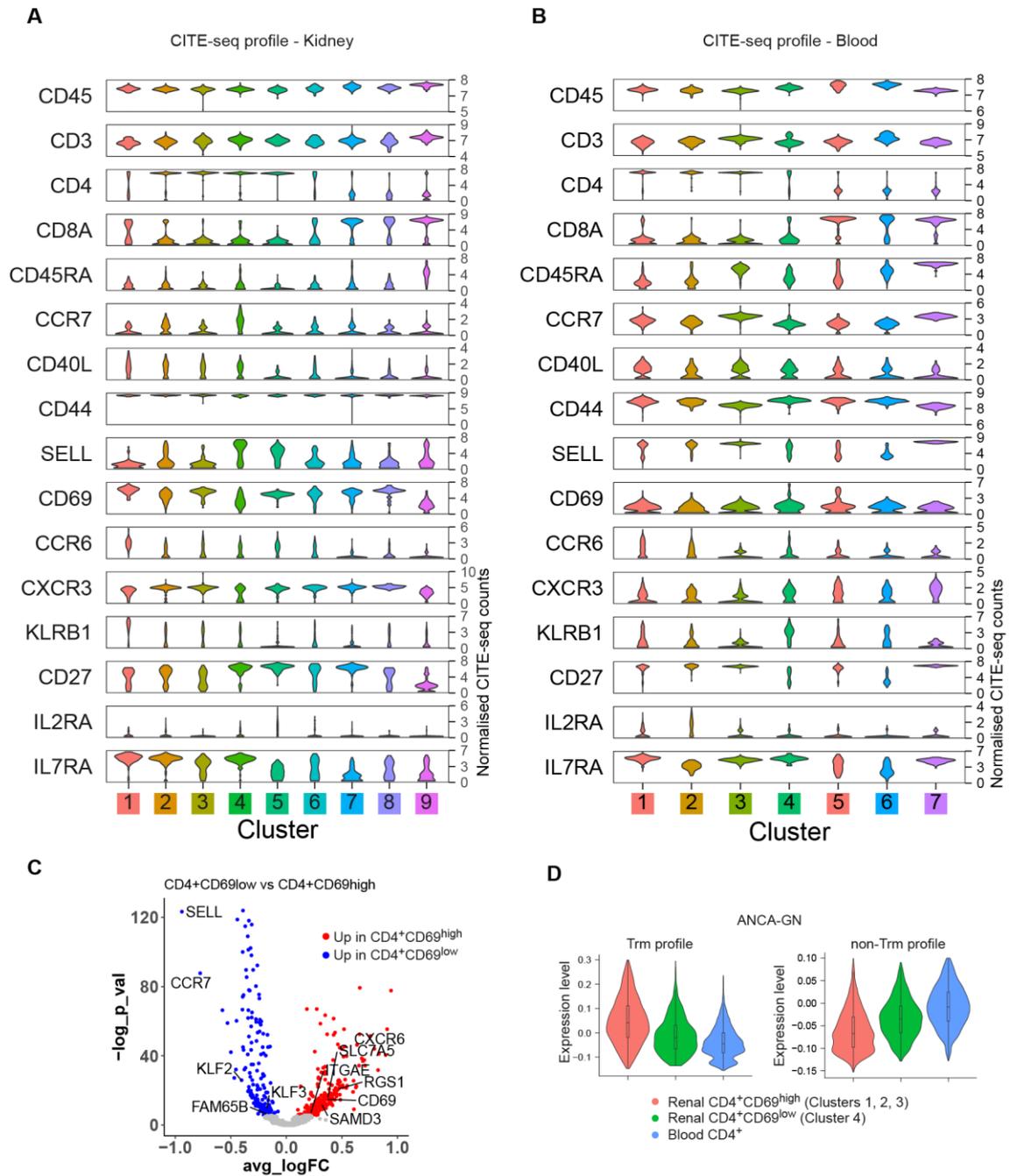
**Figure S1: scRNA-seq analysis of T cells from human blood and healthy kidney tissue from tumor nephrectomies.** (A) PAS staining of non-affected tissue from the kidney of a patient undergoing tumor nephrectomy that was used for isolation of human renal T cells. Histologic inspection excluded tumor infiltration into this part of the kidney and revealed regular tissue architecture and regular morphology of the glomeruli and the tubulointerstitial compartment. (B) Human CD3<sup>+</sup> T cells were sorted from renal biopsies and blood samples and analyzed by scRNA-seq and CITE-seq. Heat maps show clustering of T cells according to their profile of differentially expressed genes. Clusters correspond to clusters in main Fig. 1B.



**Figure S2: Profiling of T cells from human blood and healthy kidney tissue from tumor nephrectomies.** (A, B) Human CD3<sup>+</sup> T cells were sorted from renal biopsies and blood samples and analyzed by scRNA-seq and CITE-seq as described in fig. S1. Violin blots give expression profiles of surface molecules identified by barcoded antibodies (CITE-seq profile) on CD3<sup>+</sup> T cells from different T cell clusters in kidney (A) and blood (B). Clusters correspond to clusters in main Fig. 1B. (C) t-SNE plots for mRNA and protein (CITE) expression of CD4, CD8 $\alpha$ , CD69 and CCR6 of CD3<sup>+</sup> T cells from peripheral blood. (D) T cells from human kidney were stimulated with PMA/ionomycin and analyzed for intracellular IFN- $\gamma$  and IL-17A expression.

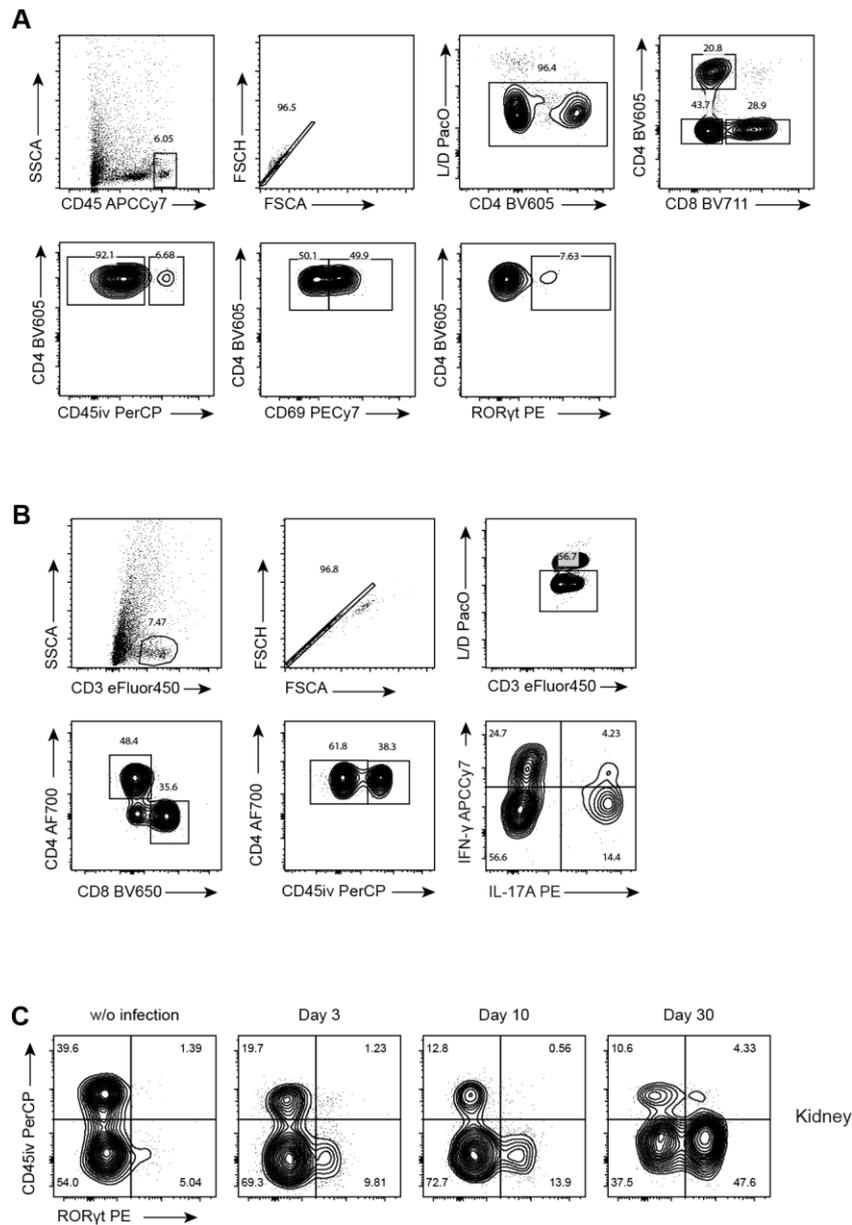


**Figure S3: scRNA-seq analysis of T cells from blood and kidney of patients with ANCA-GN.** Human CD3<sup>+</sup> T cells were sorted from renal biopsies and blood samples from ANCA-GN patients and analyzed by scRNA-seq and CITE-seq. (A, B) Heat maps show clustering and gene expression profiles of T cells according to their profile of differentially expressed genes and surface proteins. (C) t-SNE plots for mRNA and protein (CITE) expression of CD4, CD8 $\alpha$ , CD69 and CCR6 of CD3<sup>+</sup> T cells from kidney biopsies (upper panels) and peripheral blood (lower panels) of ANCA-GN patients.

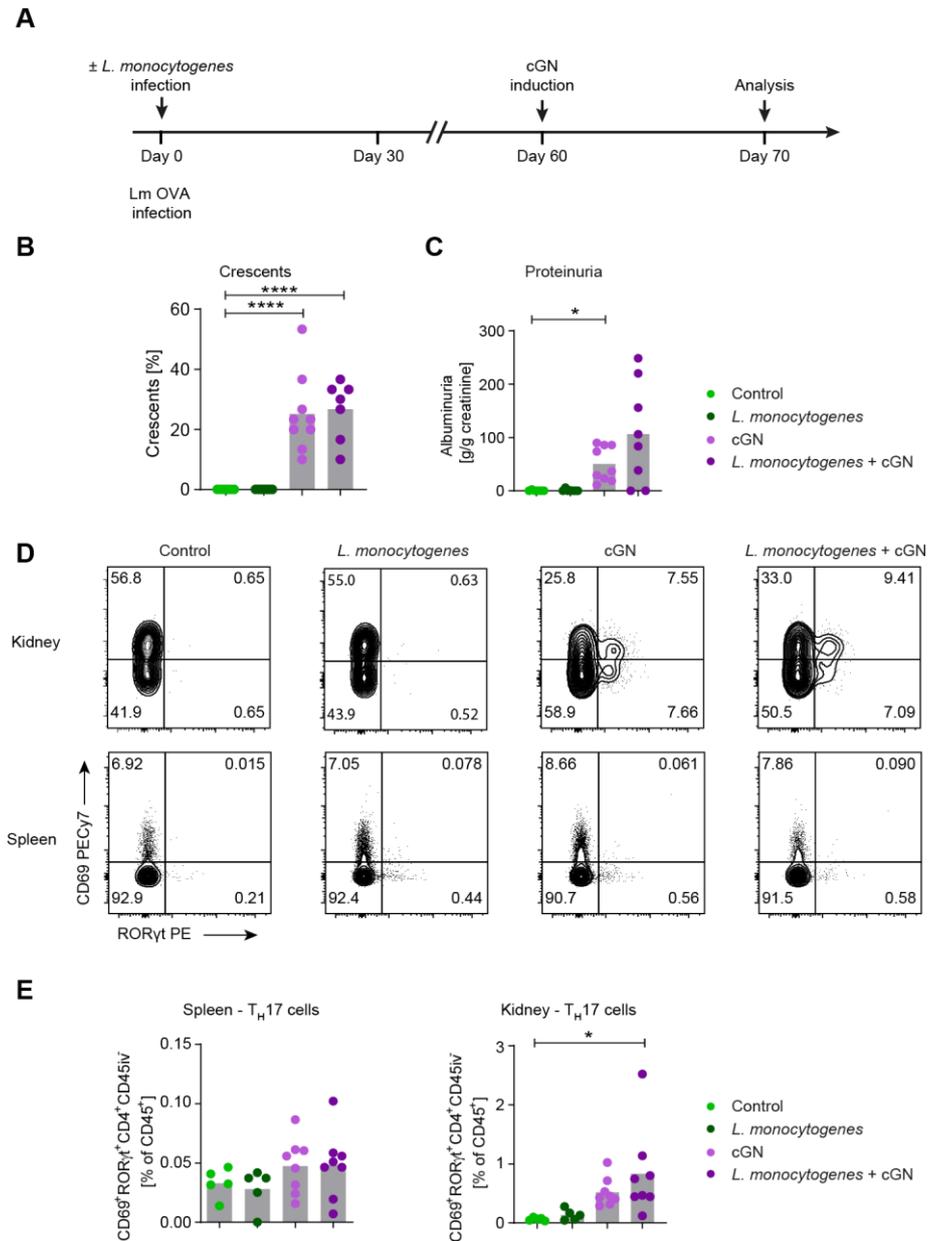


**Figure S4: Profiling of T cells from kidney and blood of patients with ANCA-GN.**

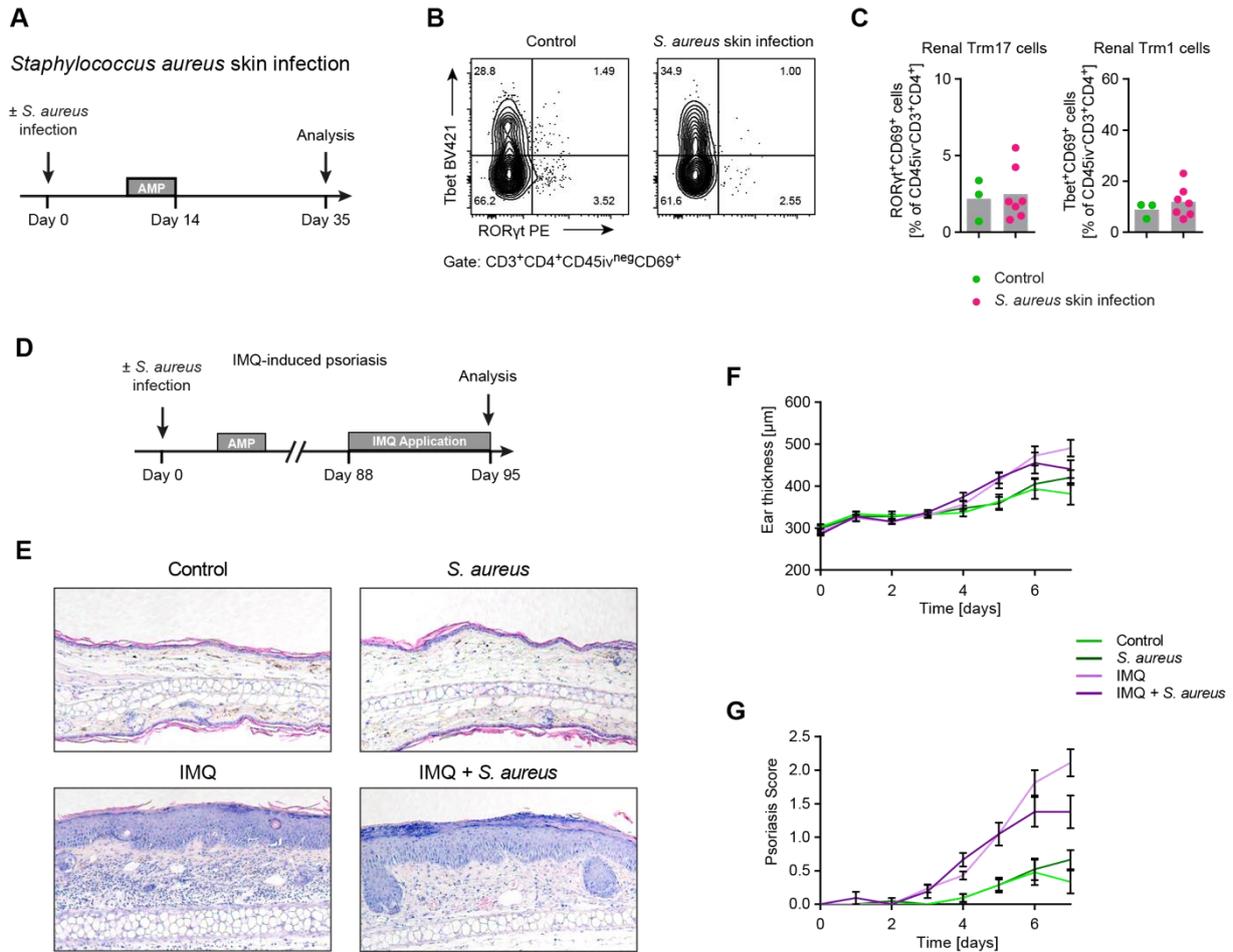
Human CD3<sup>+</sup> T cells were sorted from renal biopsies and blood samples from ANCA-GN patients and analyzed by scRNA-Seq and CITE-seq as described in fig. S3. Violin blots give normalized CITE-seq counts of surface molecules identified by barcoded antibodies (CITE-seq) on CD3<sup>+</sup> T cells from different T cell clusters in kidney (A) and blood (B). Clusters correspond to clusters in Fig. 2B and fig. S3. (C) Volcano plot showing the differential RNA expression profiles of renal CD4<sup>+</sup>CD45RA<sup>low</sup>CD69<sup>high</sup> (Cluster 1, 2, 3; red) vs. CD4<sup>+</sup>CD45RA<sup>low</sup>CD69<sup>low</sup> (Cluster 4, blue) T cells. (D) Comparison of expression profiles of CD69<sup>+</sup> T cells with published expression signatures of human CD4<sup>+</sup> T<sub>RM</sub> cells (Kumar et al. 2017).



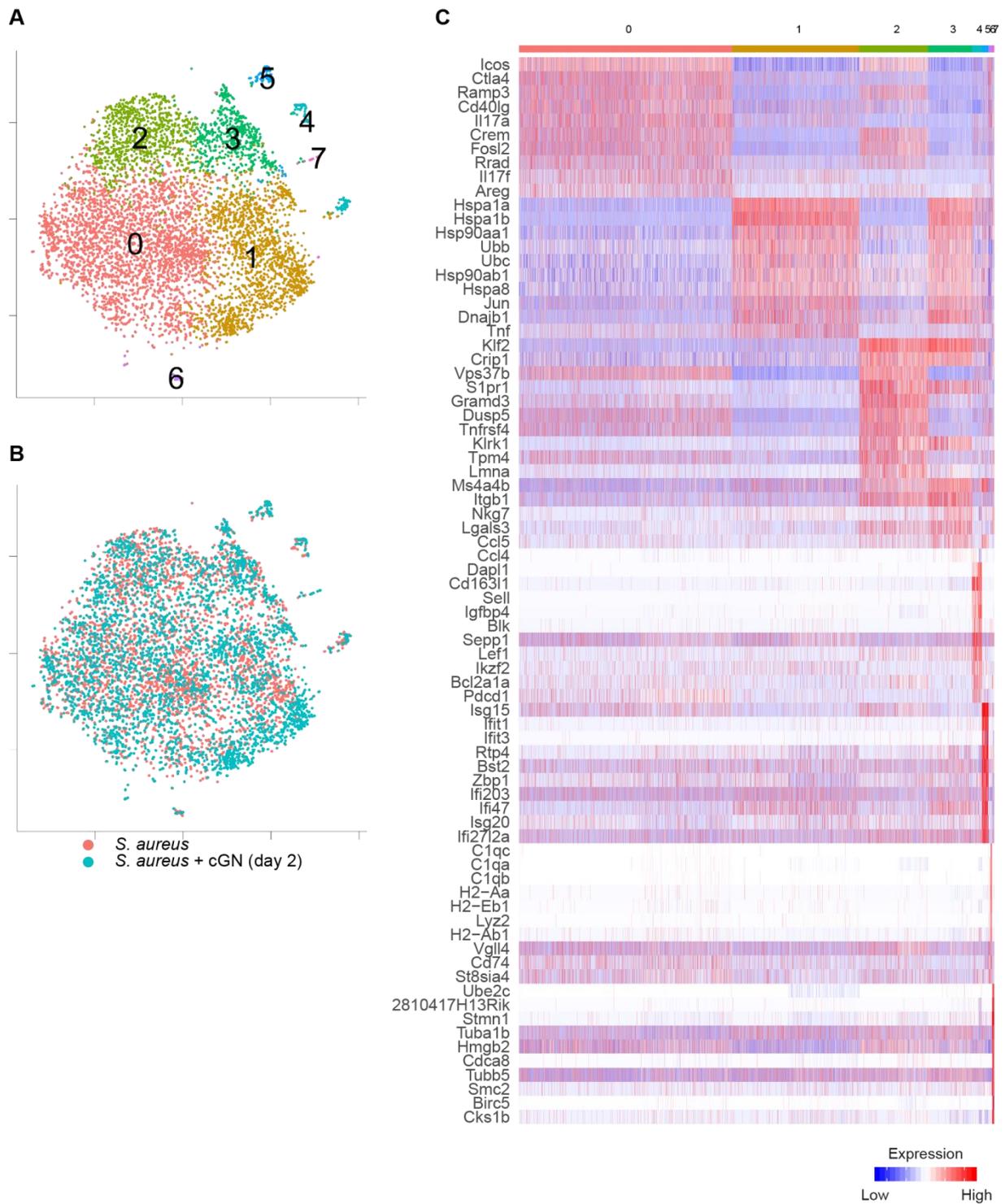
**Figure S5: Flow cytometry gating strategy for murine T cell analysis.** (A) General gating strategy for murine kidney cells after extracellular and intracellular mAb staining. (B) General gating strategy for murine kidney cells after stimulation with PMA/ionomycin and extracellular and intracellular mAb staining. (C) To exclude intravascular T cells from further analysis, fluorescent-labelled CD45 antibody was administered intravenously prior to sacrificing the mice. Renal CD4<sup>+</sup> T cells were analyzed by flow cytometry for ROR $\gamma$ t expression and CD45iv staining (related to main Fig. 3).



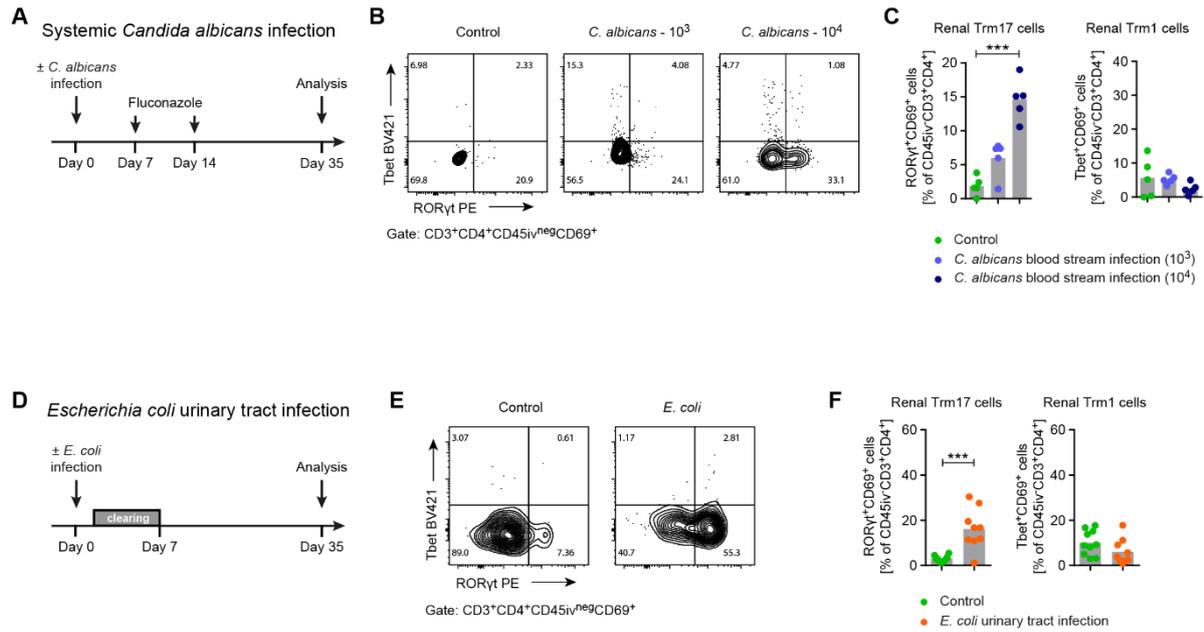
**Figure S6: Prior *L. monocytogenes* infection does not affect crescentic GN.** (A) Mice were infected with *L. monocytogenes*. Since C57BL/6 mice normally eradicate listeria within 10-14 days, ampicillin treatment was omitted. After 60 days, crescentic glomerulonephritis (cGN) was induced with nephrotoxic sheep serum. Kidneys were analyzed 10 days later. (B) Kidney damage (crescents). (C) Kidney function (proteinuria). (D) Analysis of CD69 and RORγt expression in renal and splenic CD45<sup>iv-</sup> CD4<sup>+</sup> T cells. (E) Prior *L. monocytogenes* infection did not enhance the frequencies of renal T<sub>H</sub>17 cells upon cGN induction. Data are representative of two independent experiments (one-way ANOVA with multiple comparison test).



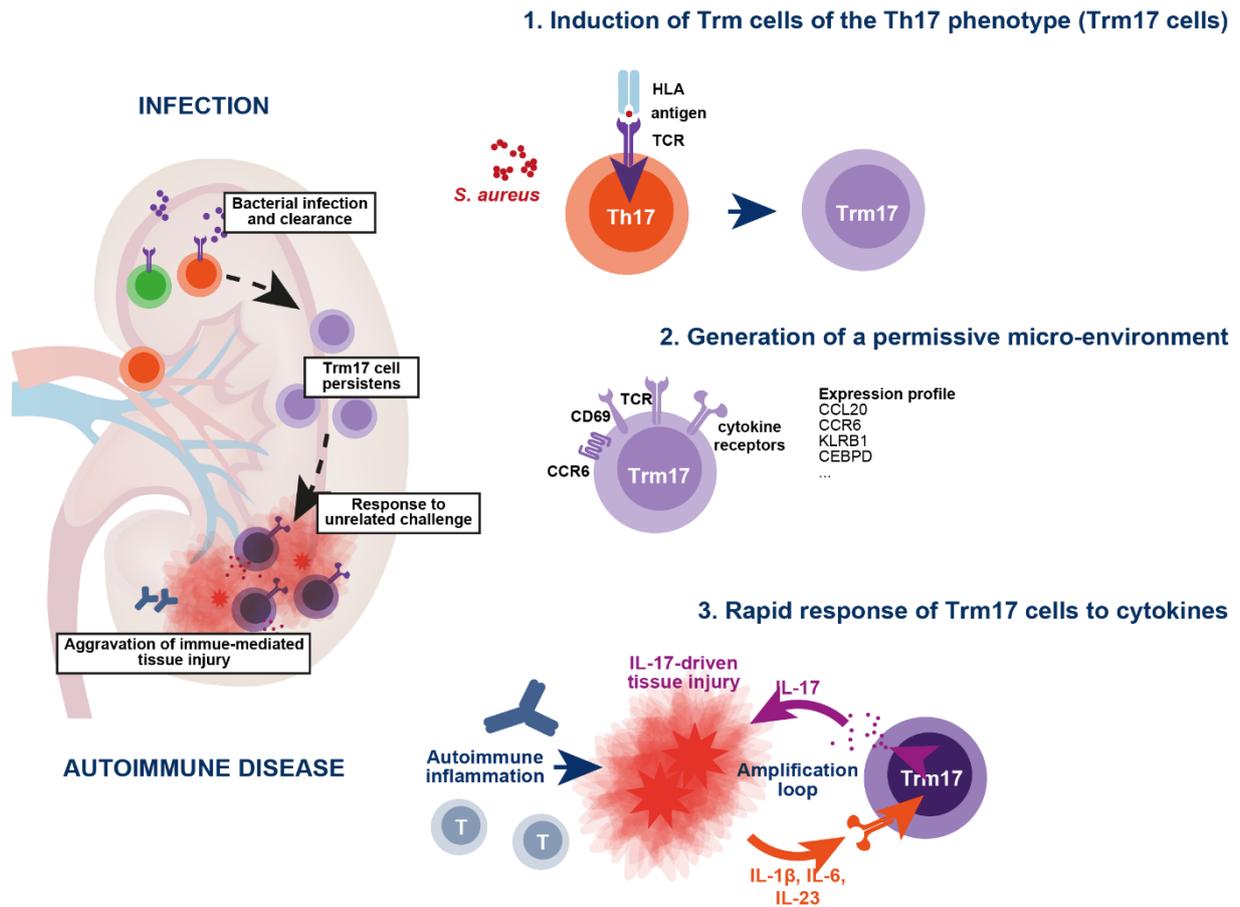
**Figure S7: Effect of local *S. aureus* skin infection on renal T<sub>RM17</sub> cell generation and effect of systemic *S. aureus* infection on imiquimod-induced skin inflammation.** (A) Mice received intradermal *S. aureus* ( $10^6$  CFUs) or PBS and were treated with ampicillin for complete bacterial clearance. (B and C) Analysis of renal CD69<sup>+</sup> Tbet<sup>+</sup> T<sub>RM1</sub> and CD69<sup>+</sup> Roryt<sup>+</sup> T<sub>RM17</sub> cells. (unpaired t-test) Data are representative of two independent experiments. (D) Imiquimod-induced psoriasis was elicited in mice by daily treatment of one ear with 15 mg Imiquimod cream (5% IMQ) 2 months after *S. aureus* infection or in control mice. The other ear was treated with Vaseline (control). (E) H&E staining of ears of mice subjected to systemic *S. aureus* infection and IMQ as indicated. (F) Ear thickness and (G) psoriasis score (0-4) were determined in the course of the disease in all groups as indicated. (mean ± SEM, n=7) Data are representative of two independent experiments.



**Figure S8: Gene expression analysis of renal CD4<sup>+</sup> T cells from nephritic mice previously infected with *S. aureus*.** IL-17A fate reporter mice (*Il17a*<sup>Cre</sup>*xR26*<sup>eYFP</sup>) were infected with *S. aureus*. After 2 weeks, mice were treated with ampicillin. At day 30, cGN was induced in one group of mice. Two days later eYFP<sup>+</sup> CD4<sup>+</sup> T<sub>RM</sub>17 cells from the kidney of mice with and without cGN were sorted and analyzed by scRNA-seq. (A) t-SNE plot showing the unsupervised clustering. (B) Cells from mice with cGN and without contributed to all clusters. (C) Heat map of the highest expressed genes in each cluster. (related to main Fig. 8C and D). To identify differentially expressed genes by cGN stimulation, only clusters 0-2 were compared (see Fig. 8C).



**Figure S9: The T<sub>H</sub>17-associated pathogens *C. albicans* and *E. coli* cause renal T<sub>RM</sub>17 cell formation.** (A) Mice were treated with and without systemic *C. albicans* (10<sup>3</sup> and 10<sup>4</sup> CFU, i.v.) and received fluconazole at days 7 and 14. Data are representative of two independent experiments. \*\*\*P < 0.001 (one-way ANOVA with multiple comparison test). (B and C) After 35 days, kidneys were analyzed for T<sub>RM</sub>1 and T<sub>RM</sub>17 cells by transcription factor staining as indicated. (D) Urinary tract infection was induced with *E. coli*. (E and F) After 35 days, kidneys were analyzed for T<sub>RM</sub>1 and T<sub>RM</sub>17 cells by transcription factor staining as indicated. Data are representative of two independent experiments \*\*\*P < 0.001 (unpaired t-test).



**Figure S10: Conceptual link of pathogen-induced tissue resident memory T cells with autoimmunity.** Infection-induced  $CD4^+$   $T_{RM}17$  cells in peripheral tissues can become activated by inflammatory cytokines to produce IL-17A and thereby aggravate autoimmunity.

	Sex (f/m)	Age (years)	Creatinine (mg/dl)	Trace proteinuria	RCC	FACS / scRNA
Number	3/3	49.6±16.5	0.93±0.13	2	6	4/3
%	50/50			33.3	100	

**Table S1: Baseline characteristics of controls at time of biopsy.** Blood samples and renal biopsies were used for analyses in main Fig. 1 and 2I-L. RCC = Renal cell carcinoma.

Nr.	Age (years)	Sex (f/m)	ANCA antigen	Creatinine (mg/dl)	IS prior to biopsy	FACS	scRNA-seq
1	72	m	MPO	3.4	none	N	Y
2	58	m	MPO	6.3	Steroids	N	Y
3	80	m	MPO	4.3	Steroids	N	Y
4	67	m	MPO	6.8	Steroids	Y	Y
5	63	m	MPO	4.7	Steroids	Y	Y
6	64	m	MPO	2.1	Steroids	Y	Y
7	54	m	PR3	1.7	Steroids	Y	N
8	66	f	PR3	1.2	Steroids	Y	N
9	82	m	MPO	2.2	Steroids	Y	N
10	58	f	PR3	0.8	none	Y	N
11	64	m	PR3	1.9	Steroids	Y	N
12	19	m	MPO	6.6	Steroids	Y	N
13	67	m	MPO	4.3	MMF	Y	N

**Table S2: Baseline characteristics of patients with ANCA-GN at the time of biopsy.** Blood samples and renal biopsies were used for analyses in main Fig. 1I-K. IS = immunosuppressive therapy; ANCA-GN = anti-neutrophil cytoplasmic antibody-associated glomerulonephritis; MPO = Myeloperoxidase; PR3 = Proteinase 3; MMF = Mycophenolate mofetil.

**Table S3. Clinical characteristics and ANCA-serotype of patients from the Hamburg GN Registry.**

This table summarizes gender, creatinine level at time of biopsy, ANCA-serotype and glomeruli identified per renal biopsy of the patients included in the analysis of renal CD69<sup>+</sup> cells and kidney function in Figure 2. PR3 = proteinase 3, MPO = myeloperoxidase;

	Gender (f/m)	Age (years)	Creatinine (mg/dl)	ANCA antigen (PR3/MPO/neg.)	Glomeruli per biopsy
Number	19/27	61.9±16.5	3.47±2.4	15/29/2	15.6±5.8
%	41/59			33/63/4	

**Table S4: List of transcription factors identified by TRRUSTv2 analysis.** Differentially expressed genes of T cell clusters 0-2 from Supplementary Figure 7 were analyzed with the TRRUST version 2 database for upstream transcription factor identification (related to main Figures 8C and D).

Key TF	Description	# of overlapped genes	P value	Q value	List of overlapped genes
<b>Stat3</b>	signal transducer and activator of transcription 3	4	4.01e-06	3.92e-05	Ccnd3,Ly6a,Bcl2,Il17f
<b>Nfkb1</b>	nuclear factor of kappa light polypeptide gene enhancer in B cells 1, p105	5	7.13e-06	3.92e-05	Nfkbia,Ifng,Ccnd3,Psme2,Bcl2
<b>Etv5</b>	ets variant 5	2	2.73e-05	0.0001	Il17a,Il17f
<b>Foxp3</b>	forkhead box P3	2	0.00012	0.000329	Ifng,Il17a
<b>Ezh2</b>	enhancer of zeste homolog 2 (Drosophila)	2	0.000165	0.000362	Il17a,Il17f
<b>Bcl6</b>	B cell leukemia/lymphoma 6	2	0.000276	0.000484	Il17a,Ifng
<b>Nr3c1</b>	nuclear receptor subfamily 3, group C, member 1	2	0.000308	0.000484	Ccnd3,Nfkbia
<b>Rela</b>	v-rel reticuloendotheliosis viral oncogene homolog A (avian)	3	0.000603	0.000829	Psme2,Ifng,Bcl2
<b>Ppara</b>	peroxisome proliferator activated receptor alpha	2	0.00182	0.00223	Nfkbia,Bcl2
<b>Jun</b>	jun proto-oncogene	2	0.0139	0.0153	Ccnd3,Ifng
<b>Sp1</b>	trans-acting transcription factor 1	2	0.0529	0.0529	Lgals1,Ccnd3

## Chapter 4: Unpublished Manuscripts, revision in *JCI Insights*

### Multi cytokine producing liver CD4<sup>+</sup> T cells characterize fibrosis in NASH patients

#### Specific contributions:

The following research article is currently in revision in *JCI Insights*. I am the shared first author of the paper. I designed research studies together with the other co-first authors and corresponding author. I processed and analyzed all the single cell RNA-seq data of CD4<sup>+</sup> T cells sorted from NAFLD and NASH patients' liver samples. These results are demonstrated in Figures: 1B, C, D, E, F, G; 2B, C, D, E, F, G; 3A, B, D, F; 5B, C, D and supplementary Figures: 1B, C, D; 2A, B, C; 3A, B, D, E. I contributed to edit the manuscripts.

I, Stephan Bonn, agree with the above statements as the direct supervisor.

Signature ,

Date & Place 15.11.2021 Hamburg



# Multi cytokine producing liver CD4<sup>+</sup> T-cells and intestinal *Bacteroides* characterize fibrosis in NASH patients

## Authors

Anna Woestemeier<sup>1†</sup>, Pasquale Scognamiglio<sup>1†</sup>, Yu Zhao<sup>5†</sup>, Jonas Wagner<sup>1</sup>, Fanziska Muscate<sup>2</sup>, Christian Casar<sup>1,3</sup>, Leonie Konczalla<sup>1</sup>, Ramez Wahib<sup>1</sup>, Karl-Frederick Karstens<sup>1</sup>, Anastasios D. Giannou<sup>1,2</sup>, Anna Duprée<sup>1</sup>, Stefan Wolter<sup>1</sup>, Agata Bielecka<sup>6</sup>, Vikas Bansal<sup>5</sup>, Oliver Mann<sup>1</sup>, Ansgar W. Lohse<sup>2</sup>, Jakob R Izbicki<sup>1</sup>, Noah Palm<sup>6</sup>, Stefan Bonn<sup>5</sup>, Samuel Huber<sup>2</sup>, Nicola Gagliani<sup>1, 2, 4</sup>

## Affiliations

<sup>1</sup>Department for General, Visceral and Thoracic Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>2</sup> I Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>3</sup>Bioinformatics Core, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>4</sup>Immunology and Allergy Unit, Department of Medicine, Solna, Karolinska Institute and University Hospital, 17176 Stockholm, Sweden

<sup>5</sup>Institute of Medical Systems Biology, baiome Center for Biomedical AI, University Medical Center Hamburg-Eppendorf, Germany

<sup>6</sup> Department of Immunobiology, School of Medicine, Yale University, New Haven, 06520, USA.

† Authors contributed equally

## Keywords

Single-cell sequencing, gut-liver axis, microbiota, NAFLD, lymphocytes

**Corresponding author**

Nicola Gagliani Ph.D

Martinistraße 52

20246 Hamburg

Tel.: +49741055406

Mail: n.gagliani@uke.de

**Financial support**

Deutsche Forschungsgemeinschaft, DFG, SFB841 to S.H., N.G., SB. ; European Research Council (715271 to N.G.); RU 306 and LFF-VF78 to YZ, VB, SB.

The authors have declared that no conflict of interest exists

## Abstract

A role of CD4<sup>+</sup> T-cells in development of NASH has been suggested, but which specific subset characterize this disease and which the development of fibrosis, remains unclear. In addition, CD4<sup>+</sup> T-cells are connected to intestinal microbiota which is link to liver pathology. A gut-liver axis has been suggested to play a role in NASH, but the role of CD4<sup>+</sup> T-cells is unclear.

We provide the first cell atlas of liver infiltrating CD4<sup>+</sup> T-cells in NAFLD and NASH patients, showing that NASH is characterized by a unique population of multi cytokine producing CD4<sup>+</sup> T-cells. Moreover, only IL-17A producing CD74<sup>+</sup> CD4<sup>+</sup> T-cells showed an enrichment in patients with advanced fibrosis.

Next, we show intestinal microbiota to be present in the liver of NAFLD and NASH patients, with an enrichment of *Bacteroides* infiltrating the mucosal layer of the small intestine of NASH patients and a positive correlation between the frequency of liver T-cells co-producing cytokines and *Bacteroides*.

We deliver the first CD4<sup>+</sup> T-cell atlas of NAFLD and NASH providing the rational to target liver CD4<sup>+</sup> T-cells to block fibrosis in NASH patients. Finally, our data suggest that these multiple cytokine producing CD4<sup>+</sup> T-cells could be part of the gut-liver axis characterizing NASH pathogenesis.

## Introduction

Non-alcoholic fatty liver disease (NAFLD) is the leading cause of chronic liver disease in the Western world (1)(2). NAFLD can progress to non-alcoholic steatohepatitis (NASH) (3) and approximately 40% of NASH patients inevitably develop severe complications such as fibrosis, cirrhosis and hepatocellular carcinoma (HCC) (1)(2). However, the etiology underlying the progression from NAFLD to NASH and ultimately to fibrosis, a major determinant of mortality (2), remains largely unclear. Moreover, there is not one single drug approved by the Food and Drug Administration (FDA) or European Medicines Agency (EMA) for the treatment of NASH. Previous mouse studies suggest a role of CD4<sup>+</sup> T-cells in NASH (4) (5). Human studies associated the frequency of different CD4<sup>+</sup> T-cell populations, such as T<sub>H</sub>1, T<sub>H</sub>2, T<sub>H</sub>17 and Foxp3<sup>+</sup> Treg cells to the immunopathogenesis of NASH (6)(7)(8). While these previous studies provided the fundamental indication of a role of CD4<sup>+</sup> T-cells in NASH, they were based on a supervised analysis which classifies CD4<sup>+</sup> T on the basis of a small number of pre-selected cytokines. Therefore, they did not fully describe the complexity of human T-cell biology and their potential implication in NASH.

It is known that the liver CD4<sup>+</sup> T-cell composition can be altered by the intestinal microbiota. There is increasing evidence not only that the intestinal microbiota might influence intrahepatic T-cells, but also that it plays a crucial role in the pathogenesis of liver diseases, including NASH (10)(11)(12). By profiling the stool microbiome, it was shown that the abundance of bacterial species such as *Bacteroides* was more abundant in NASH patients than in healthy individuals (12)(13). Moreover, bacterial DNA was found in the liver of NAFLD patients (14) and the liver microbial composition differed between morbidly obese and non-morbidly obese NAFLD patients (14).

On the basis of all this, a connection between the intestine and the liver, i.e. the gut-liver axis, has been proposed for NASH. However, a study analyzing the immune system and the tissue specific microbiota composition within the same patient cohort remains to be performed. Besides this, the previous studies were limited in that they compared either NASH or NAFLD

patients with healthy individuals and were thus unable to describe the developmental spectrum covering NAFLD to NASH progression (12)(13)(14).

Here, combining single cell RNA sequencing, multi-parameters flow cytometry analysis and tissue derived microbiome sequencing from liver and intestinal samples, our study reveals the complete CD4<sup>+</sup> T-cell landscape and the microbial composition in liver and intestine of NASH patients.

## Results

### Characterization of 8 distinct subsets of liver infiltrating CD4<sup>+</sup> T-cells in NAFLD patients

To understand CD4<sup>+</sup> T-cells infiltrating the liver of NAFLD patients in detail, we performed scRNA-seq (n= 4421 cells) on FACS sorted memory CD45RA<sup>-</sup> CD4<sup>+</sup> T-cells isolated from 3 NAFLD patients (Fig.1A, Suppl. Fig.1A). Our analysis revealed 8 distinct T-cell populations (Fig.1B) (23) (17)(24) (25). We found populations of CD4<sup>+</sup> T-cells with either a T<sub>H1</sub> (*IFNG*) or a T<sub>H17</sub> (*IL17A*, *RORA*) cell polarization state. We also found cells with an anti-inflammatory T<sub>R1</sub> (*IL10*, *TIGIT*, *LAG-3*, *CTLA4*) and Foxp3<sup>+</sup> Treg (*FOXP3*, *CTLA4*, *IL2RA*) resembling transcriptome. Central memory cells (*CCR7*, *SELL*), heat shock protein (HSP)<sup>+</sup> cells and cytotoxic (*GZMB*, *PRF1*) cells were also identified. Finally, we found a small contamination of Kupffer cells (*CD163*, *CD14*) (Fig.1B,C).

Regarding the central memory cells, we found two distinct-cell populations. Central memory 2 cell population show a less-active cell signature expressing high levels of *LEF1*, *CD7*, *SELL* and *CCR7* (26), compared to the population 1 which is characterized by a high expression of *ITGB1*, *TNFRSF4*, *KLRB1*, *IL32* and *RORA* (Fig.1D), suggesting very active memory cells with possible pathogenic activity. We did not identify specific cell populations resembling T<sub>H2</sub>, T<sub>H9</sub> and T<sub>FH</sub> subsets (Suppl Fig.1B).

Besides circulating central memory and effector memory cells, tissue-resident memory (T<sub>RM</sub>) cells play a crucial role in orchestrating both protective and pathological tissue specific immune responses (27)(28). Therefore, we wondered which of the above-mentioned populations of CD4<sup>+</sup> T-cells display a T<sub>RM</sub> transcriptomic signature. By using a literature curated T<sub>RM</sub> signature score (16), we observed that the T helper cell clusters, namely T<sub>H1</sub>, T<sub>H17</sub> and T<sub>R1</sub>, reflect a T<sub>RM</sub> gene profile (Fig.1E). Thus, these T<sub>RM</sub> cells comprehend a heterogenous group of T helper cell subsets, probably reflecting their effector precursors.

It has been shown that an apparently distinct CD4<sup>+</sup> T-cell subset, such as T<sub>H17</sub>, might instead acquire the phenotype of another subset such as T<sub>H1</sub> or T<sub>R1</sub> (29)(30)(31). To further investigate this in our liver CD4<sup>+</sup> T-cell subsets, we first performed a sub-clustering analysis of T<sub>H1</sub>, T<sub>R1</sub>, and T<sub>H17</sub> cells, since these types of cells seemed already more similar to each other and

shared interconnected cell states. We found six different sub-clusters reflecting different cell states (Fig.1F, Suppl. Fig.1C) with combined  $T_{H1}$ ,  $T_{R1}$ , and  $T_{H17}$  features and partial overlaps between the different states. Next, we inferred differentiation cell states performing pseudotime analysis (Fig.1G, Suppl. Fig.1D). This analysis revealed a potential cell trajectory connecting  $T_{H17}$  and  $T_{H1}$  and further  $T_{R1}$  cell states. Hence, we postulate that some of these T helper cells, characterized by a  $T_{RM}$  phenotype, might represent a unique cellular spectrum composed by different cellular states, instead of being distinct lineages as originally proposed.

Overall, our analysis reveals the landscape of the  $CD4^+$  T-cell populations in the liver of NAFLD patients, including one population of cytotoxic effector memory T-cells, two distinct central memory populations and  $T_{RM}$ -like cells characterized by a heterogenous and a continuum of different cell states, namely  $T_{H1}$ ,  $T_{H17}$  and  $T_{R1}$ .

#### Characterization of 8 distinct subsets of liver infiltrating $CD4^+$ T-cells in NASH patients

Next we wondered whether the above populations of liver  $CD4^+$  T-cells were specific for NAFLD patients or also present in NASH. Therefore, we performed scRNA seq (n= 6435) on  $CD45RA^- CD4^+$  T-cells FACS sorted from the liver of three NASH patients (Fig.2A) and identified 8 different T-cell subsets (Fig.2B,C), including four distinct central memory populations,  $Foxp3^+$  Treg cells and a heterogenous population of  $T_{RM}$ -like cells composed of cells with  $T_{H1}$ ,  $T_{H1\_T_{H17}}$  and  $T_{R1}$  polarization states (Fig.2E).

In addition, our analysis revealed four different cell states of central memory cells, all expressing *SELL* and *CCR7*. While  $T_{CM} 2$  and 4 were mainly present in a single patient (Suppl. Fig.3B),  $T_{CM} 1$  appear to be later activated (*KLF6*, *CXCR4*) compared to  $T_{CM} 1$  (*TXNIP*) (Fig. 2D). In accordance with the NAFLD map, we did not identify specific clusters of  $T_{H2}$ ,  $T_{H9}$  and  $T_{FH}$  subsets using a supervised analysis (Suppl Fig.2A).

To further characterize the  $T_{RM}$  cells and their potential transitional states, we subclustered the  $T_{RM}$  cells and identified six different sub-populations (Fig.2F, Suppl. Fig.2B) and again, found partial overlaps between the different states. However, in contrast to the NAFLD map, the pseudotime analysis showed more chaotic paths with different potentially cellular endpoints

including the a potentially cytotoxic  $T_{R1}$  sub-cluster 3 (*IL-10*, *TIGIT*, *GZMH*, *GZMB*, *PRF1*, *GZMK*) (Fig.2g, Suppl. Fig.2C). This chaotic path suggests a less delineated T effector cell trajectory to be present in NASH compared to NAFLD patients.

#### Differences revealed between liver $CD4^+$ T-cells in NAFLD and NASH patients

To further investigate differences between NAFLD and NASH, we decided to compare and contrast the two  $CD4^+$  T-cell profiles. (Fig.3A, Suppl. Fig.3A). We found that each cluster contained cells from NASH and NAFLD patients (Suppl. Fig.3B). Overall, nearly all clusters were equally represented by all specimens and proportionally distributed across patients.

Next we wondered whether the transcriptional profiles of these clusters were different between NAFLD and NASH. Differential gene expression analysis of IL-10-producing  $Foxp3^{neg}$   $CD4^+$  T-cells revealed that IL-10 is more expressed in the cells from the NAFLD patients, while the chemokine receptor *CXCR3* was significantly overexpressed in IL-10-producing  $Foxp3^{neg}$   $CD4^+$  T-cells deriving from NASH patients. Of note, *CXCR3* was shown to be associated with an overproduction of cytokines (32) (Fig.3B). In line with that, our recently published data show that IL-10-producing  $Foxp3^{neg}$   $CD4^+$  T-cells are a functional heterogeneous population of cells with both anti-inflammatory function and potential pathogenic activity (17). Therefore we created a  $T_{R1}$  score based on our previous published data (17) and found that the IL-10-producing  $Foxp3^{neg}$   $CD4^+$  T-cells in NASH had a lower  $T_{R1}$  score compared to those in the liver of NAFLD patients (Suppl. Fig.3D). This suggests an enrichment of non  $T_{R1}$  - IL-10-producing  $Foxp3^{neg}$   $CD4^+$  T-cells in NASH.

Next, we focused on  $T_{H17}$  cells since these have been suggested to play a role in NASH development (6) (33). We performed a differential gene expression analysis between  $T_{H17}$  cells found in NAFLD and NASH liver samples. *IL23R*, *IL32* and *CD74* were upregulated in the cells isolated from NASH patients compared to those from NAFLD, suggesting a possible pro-inflammatory  $T_{H17}$  cell state in NASH patients (Fig.3B). Moreover, when the differential gene expression analysis was also performed on  $T_{H1}$  and  $Foxp3^+$  Treg cells, we observed that *CD74*, was more expressed in the cells isolated from NASH patients compared to NAFLD

patients (Suppl. Fig.3E). To validate this finding, we extend our analysis to 25 NAFLD and 32 NASH patients by testing *CD74* expression by qPCR analysis. We found that the expression of *CD74* was significantly higher in NASH patients (Fig.3C).

Since the transitional cell states,  $T_{H1}$ ,  $T_{H17}$  and  $T_{H1\_TH17}$  also displayed a  $T_{RM}$  phenotype, we further focused on these cells and found a higher expression of the  $T_{RM}$  transcriptomic signature in the liver cells isolated from the NAFLD patients (Fig.3D). In order to validate these findings, we performed flow cytometry analysis and confirmed the enrichment of  $CD69^+ CD4^+$  T-cells in NAFLD patients (Fig.3E, Suppl. Fig.3C). This suggests that more circulatory T helper cells are present in the liver of NASH patients.

Finally, since our differential gene expression analysis of  $T_{R1}$  and  $T_{H17}$  cells revealed genes connected with the activation of T-cells and an increase in cytokine production, such as *CXCR3*, *IL23R* and *CD74*, we analyzed our single cell sequencing data set for the frequency of cytokine producing  $CD4^+$  T-cells. We found a significantly higher percentage of cells co-expressing a combination of *TNF*, *IFNG*, *IL17A* and *IL10* (Fig.3F) in NASH compared to NAFLD.

In conclusion, we found that  $T_{H17}$  cells, with a pathogenic transcriptome, and IL-10 producing T-cells with a non-regulatory signature characterize NASH patients. In addition,  $CD4^+$  T-cells found in the liver of NASH patients are possibly circulatory cells, and beside the different clusters, they have an overall pro-inflammatory transcriptomic signature characterized by the co-expression of many pro-inflammatory cytokines.

#### Liver $CD4^+$ T-cells in NASH have a distinct cytokine profile compared to NAFLD

To validate the observation regarding the increased frequency of liver  $CD4^+$  T-cells able to co-produce multiple cytokines in NASH patients, we performed intracellular cytokine FACS staining on  $CD4^+$  T-cells freshly isolated from the liver of 27 NAFLD and 40 NASH patients (Suppl. Fig.4A). Considering that many cytokine combinations can possibly occur, we performed a viSNE analysis to identify which of them are indeed present in the liver of at least either NAFLD or NASH patients (Fig.4B). Furthermore, we found that different populations of

unconventional multi cytokine producing CD4<sup>+</sup> T-cells are significantly enriched in the liver of NASH compared to NAFLD patients (Fig.4C), namely cells co-producing TNF- $\alpha$ /IFN- $\gamma$ , IFN- $\gamma$ /IL-17A, IFN- $\gamma$ /IL-10, TNF- $\alpha$ /IL-17A, TNF- $\alpha$ /IL-10, TNF- $\alpha$ /IFN- $\gamma$ /IL-10 and TNF- $\alpha$ /IFN- $\gamma$ /IL-17A.

Then, we wanted to further test the relevance of the above identified multi-cytokine producing cells to distinguish between NASH and NAFLD using a random forest classifier. As reference, we used clinical parameters, such as transaminase levels, since they are known to discriminate between NASH and NAFLD, but still with a low sensitivity and specificity. By random forest machine learning, we identified the most important features to discriminate between NAFLD and NASH (Fig.4D). Visualization by tSNE analysis indeed showed that NAFLD and NASH patients segregated apart from each other (Fig.4E). In line with this, using the selected features, a classification performance of 89% was obtained that distinguished between NAFLD and NASH patients (AUC=0.89; sensitivity 0.84; specificity=0.83) (Fig.4F,G). As a control, we used the clinical parameters alone (sex, age, BMI and transaminase levels) or cytokine producing cells alone, and both showed less accuracy (AUC=0.79 and AUC=0.69; Suppl Fig.4B,C) than when combined together.

Taken together, we found populations of unconventional multi cytokine producing CD4<sup>+</sup> T-cells which are significantly enriched in the liver of the NASH patients and are able to distinguish between NASH and NAFLD patients in combination with clinical parameters.

#### IL-17A producing CD4<sup>+</sup> T are enriched in the liver of NASH patients with fibrosis

We wondered whether the frequency of the liver multi cytokine producing CD4<sup>+</sup> T-cells correlate with the fibrosis score in NASH patients. Among the different multi cytokine producing CD4<sup>+</sup> T-cells, we found that CD4<sup>+</sup> T-cells single producing IL-17A are significantly enriched in the liver of NASH patients with fibrosis compared to no fibrosis (Fig.5A). Furthermore, in trend there is a correlation between the frequency of IL-17A producing cells and fibrosis score (Fig.5B).

These findings were confirmed with our single cell data set, showing that the expression of *IL17A* correlated positively with the fibrosis score in NASH patients (Fig.5C) and there was a positive correlation between the fibrosis score the pro-fibrotic gene *COL1A1* (Fig.5D).

Since we found the expression of *CD74* to be significantly higher within the cluster of *IL17A* expressing cells in NASH patients (Fig.3B), we wondered whether *CD74* is also enriched in patients with fibrosis. In a different cohort of patients, analysis of *CD74* expression by qPCR indeed revealed a significantly higher expression in NASH patients with fibrosis compared to no fibrosis (Fig.5E).

Taken together, we found IL-17A producing CD4<sup>+</sup> T-cells and *CD74* to be enriched in NASH patients with fibrosis compared to no fibrosis. While unconventional multi cytokine producing CD4<sup>+</sup> T-cells might play a role in the progression from NAFLD to NASH, our data suggest that IL-17A producing CD4<sup>+</sup> T-cells might be key for the further development of fibrosis.

#### Translocation of microbiota into the liver of NASH patients

Considering the link among CD4 T-cell function and intestinal microbiota, our finding regarding the enrichment of multi cytokine producing CD4<sup>+</sup> T-cells in NASH patients led us to argue in favor of a possible liver-gut axis as a reason of the hyper activation of the liver T-cells. To support this, instead of analyzing the fecal microbiome, we tested (16s rDNA seq) directly the tissue-specific microbiome of both the liver (n=50) and the small intestine (n=19) (Fig.6A).

Then we examined differences of the microbiome structure between the liver and small intestine across NAFLD and NASH patients by principal coordinate analysis (PcoA) of Bray-Curtis distance (Suppl. Fig.5A). The analysis revealed similar microbial communities within NASH and NAFLD patients ( $p > 0.1$ ). Furthermore, the within-sample diversity ( $\alpha$ -diversity; Shannon's diversity) did not differ between different disease conditions (Suppl. Fig.5B). To adjust for inter-patient variability, we next analyzed paired liver and small intestine samples, and observed that even though there was a clear separation between the liver and small intestine tissue microbiome ( $p < 0.001$ ; Fig.6B), the microbial composition between NAFLD and NASH patients still did not differ ( $p > 0.1$ ).

Next, we explored the community structure on genus level and found *Bacteroides* to be significantly enriched in the small intestine of some of the NASH patients (Fig.6C). *Bacteroides* were also present in the liver tissue, but there were no significant differences in genus abundances between NAFLD and NASH in morbidly obese patients (Fig.6C).

Finally, we wondered whether the elevated *Bacteroides* level correlates with the cytokine production within patients. First, we performed intracellular cytokine FACS staining on CD4<sup>+</sup> T-cells from the small intestine of 15 NAFLD and 10 NASH patients. We found that in trend, there is also an enrichment of cells co-expressing cytokines in the intestine of NASH patients compared to NAFLD (Suppl. Fig.5C,D). Subsequently, we tested for a possible correlation between T-cells producing cytokines in the liver and small intestine and *Bacteroides* abundance in NASH patients (Fig.6e, Suppl. Fig.5E). We found that there is a positive correlation between the frequency of liver (Fig.6E) and small intestine (Suppl. Fig.5E) T-cells co-producing TNF- $\alpha$ /IFN- $\gamma$  or IFN- $\gamma$ /IL-10 and intestinal *Bacteroides* abundance.

In short, we show intestinal microbiota to be present in the liver of both NAFLD and NASH patients. Intestinal *Bacteroides* appear to be enriched in the mucosal layer of the small intestine of NASH patients, and to correlate with T-cells co-producing TNF- $\alpha$ /IFN- $\gamma$  or IFN- $\gamma$ /IL-10 in the liver and small intestine.

## Discussion

By providing an unsupervised analysis, our study shows for the first time, that the developmental process from NAFLD to NASH is not associated with only classical  $T_H$  cell subsets as previously suggested (6)(33)(40)(41). But rather, with a larger spectrum of different subsets including unconventional multi-cytokine producing cells which might represent an in-between developmental stage.

Furthermore, during progression to fibrosis, multi-cytokine producing  $CD4^+$  T-cells with a  $T_H17$  polarization state are the main enriched subset. Interestingly, cells producing  $IFN\gamma$ , IL-10 and TNF- $\alpha$  were not enriched in patients with fibrosis (Fig.5A), suggesting that these cytokines might be relevant during acute inflammation, but not for the progression to fibrosis. In mice the inhibition of IL-17A showed reduced mRNA levels of profibrogenic genes, such as *COL1A1*, and a reduced degree of liver fibrosis (41).

Using fate-mapping mouse models, we have previously shown that  $T_H17$  cells can either have a pro-inflammatory fate participating in a pathological immune response or an anti-inflammatory fate contributing to the resolution of the immune response (30)(31). By combining an unsupervised transcriptional profiling and FACS analysis, we describe a dynamic interconnection among  $T_H1$ ,  $T_H17$  and  $T_R1$  cell states. Moreover, we show that there is a more chaotic developmental path with different pathogenic endpoints and an increase of multiple cytokine production in NASH patients that could be one of the causes of the chronic liver inflammation typical of these patients.

Within the  $T_H17$ ,  $T_H1$ ,  $T_R1$ , and  $Foxp3^+$  Treg clusters we found *CD74* to be enriched in NASH patients. *CD74* is usually not known to be expressed on T-cells, but it was previously reported to regulate proliferation, survival and secretion of inflammatory cytokines, after activation by MIF (34)(35)(42). In a mouse model of chronic liver injury, *CD74*-deficient mice were protected from liver fibrosis (42). Therefore, *CD74* could be a potential novel target to disrupt NASH progression and development of fibrosis by anti-*CD74* antibody hLL1 milatuzumab (36)(37). The role of *CD74* in T-cells is still unclear.

The role of T<sub>RM</sub> cells in the development of NASH is largely unknown. We show here that the percentage of cells with a T<sub>RM</sub> gene signature and phenotype is higher in NAFLD than in NASH patients. This suggests the presence of more circulatory T helper cells in the liver of NASH patients, that might have been previously activated in different tissues, such as mucosal associated lymphoid tissue (MALT) or in the intestine and then migrate through the liver. In line with this, emerging evidence indicates that intestinal dysbiosis has a significant role in the pathogenesis of human liver diseases. Dysbiosis was shown to associate with intestinal permeability (43) and consequently lead to translocation of microorganisms and microorganism-derived molecules into the liver (44)(14). Using 16S ribosomal RNA gene sequencing of fecal samples, Boursier *et al.* found that the bacterial genera *Bacteroides* and *Ruminococcus* were substantially higher, and *Prevotella* was lower in patients with NASH compared with those without NASH (12). Moreover, Sookian *et al.* showed that the liver tissue of NAFLD patients contains a diverse repertoire of bacterial DNA that differs between morbidly obese and non-morbidly obese NAFLD patients (14). Here, we investigated for the first time, the tissue-specific microbiome of both the liver and small intestine within NASH and NAFLD morbidly obese patients. Our data confirm the presence of intestinal microbiota in the liver of both NAFLD and NASH patients. However, within the liver tissue of morbidly obese NAFLD and NASH patients, the microbial communities were similar. Since obesity was shown to be an important driver of liver microbial DNA composition (14), this might explain the difference between - but not within - morbidly obese and non-morbidly obese patients. These results also suggest that the liver microbiome is a premise to NASH development, but does not seem to be the main driver. In contrast, investigating the microbial composition of the intestinal mucosal layer, we found *Bacteroides* to be enriched in the small intestine of some morbidly obese NASH patients. This is in line with studies showing an increase in *Bacteroides* abundance in stool samples of NASH patients (12). Interestingly, patients with elevated levels of *Bacteroides* also displayed elevated percentages of cytokine producing CD4<sup>+</sup> T-cells, above median levels. Since T-cells have been shown to react to *Bacteroides* polysaccharide A by secreting cytokines (45), we speculate that in NASH patients, T-cells are pre-activated by the microbiota

(e.g. *Bacteroides*) in the intestine or in MALTs and subsequently migrate to the liver where they find a similar overall microbiota composition which unleash their predisposition to secrete many cytokines simultaneously. Finally, the persistence of IL-17A producing CD4<sup>+</sup> T-cells are key in the development of fibrosis.

In short, our data suggest that a gut-liver axis might be present in NASH patients, but in contrast to what has been suggested so far, the pre-activation of T-cells might occur outside the liver, probably in the intestine or MALTs where the microbiota differ between NASH and NAFLD patients. Nevertheless, this currently remains a working hypothesis and further studies are required to fully confirm this. Indeed, one of the main limitations of our study is the lack of functional experiments using for example, NASH mouse models.

In conclusion, targeting multi cytokine producing CD4<sup>+</sup> T-cells and the intestinal microbiota can open new therapeutic strategies for NASH patients, including antibiotic or probiotic treatments.

## Methods

### Patient samples

Fresh samples used in the study were obtained from patients undergoing bariatric surgery and simultaneous liver biopsy at the University Hospital Hamburg-Eppendorf.

### Flow cytometry

Briefly, after immune cell isolation using a Percoll gradient (GE Healthcare), cells were re-stimulated with PMA/Ionomycin for 3 h and stained (Suppl. Table 5). Samples were acquired on an LSR II flow cytometer (BD Bioscience) and analyzed using Flowjo. Additionally, data were analyzed using the Cytobank platform (viSNE analysis).

### RNA analysis

Total RNA was extracted from liver tissue using TRIzol reagent (Invitrogen). The high-capacity cDNA synthesis Kit (Applied Biosystems) was used for cDNA synthesis. Real-time PCR was performed using the Kapa Probe Fast qPCR Master Mix (Kapa Biosystems) on the StepOnePlus system (Applied Biosystems). Probes used were HPRT1: Hs02800695\_m1 (Thermo Fisher) and CD74: HS00269961\_m1 (Thermo Fisher). Relative RNA expression was normalized to HPRT and calculated using the  $2^{-\Delta\Delta Ct}$  method.

### 10x sample processing and cDNA library preparation

The sorted cellular suspension was loaded on a 10x Genomics Chromium instrument to generate single-cell gel beads in emulsion (GEMs). Single-cell RNA-Seq libraries were prepared as described by the 10x Genomics Single Cell 3' v2 Reagent Kit user guide and using the following Reagent Kits: Chromium Single Cell 3' Library & Gel Bead Kit v2 (PN-120237), Chromium Single Cell A Chip Kit (PN-120236) and Chromium i7 Multiplex Kit (PN-120262).

### Pre-processing and quality control of single-cell RNA-sequencing data

10x Genomics raw sequencing data were processed using CellRanger software (version 2.1.0, 10x Genomics, Pleasanton, CA) and the 10x human genome GRCh38 1.2.0 release as the reference (function `cellranger count`). The matrices of cells and the unique molecular identifier (UMI) count were obtained and further processed by the R package Seurat (version 3.0.2) (15). In order to avoid potential contamination of non-T-cells during sorting, only cells with an average count of CD3D, CD3E and CD3G higher than 0 were kept for further analysis.

#### Dimensionality reduction and clustering

The Seurat R package (version 3.0.2) was used to perform unsupervised clustering analysis on scRNA-seq data. T-distributed Stochastic Neighbor Embedding (tSNE) was used to visualize clustering results. The top differential expressed genes in each cluster were found using the `FindAllMarkers` function (`min.pct = 0.25`, `logfc.threshold = 0.25`) that ran Wilcoxon rank sum tests. The differential expression between conditions (NASH vs NAFLD) were calculated by `FindMarkers` function (`min.pct = 0.1`), which also ran Wilcoxon rank sum tests.

#### $T_{RM}$ and Tr1 scores

Average expression levels of  $T_{RM}$  gene set ( $T_{RM}$  score) and Tr1 gene sets (Tr1 score) (Suppl. Table. 6) were calculated by Seurat function `AddModuleScore` (`nbin = 24`, `ctrl = 100`).  $T_{RM}$  signature genes were obtained from the study of sorted CD4+CD69+ and CD4+CD69neg T-cells of lung, spleen and blood (16). Tr1 signature genes were first selected from the study of mouse Tr1 cells (17) and then converted to corresponding human genes using `getLDS` function from the `biomaRt` package (version 2.40.3).

#### Trajectory analysis

The `slingshot` R package (version 1.3.2) (18) was used for trajectory analysis. The seurat objects of the selected clusters were first converted to `SingleCellExperiment` object (function `as.SingleCellExperiment`). Then function `slingshot` (`reducedDim = "PCA"`) was used to identify the global lineage structure with a cluster-based minimum spanning tree (MST) and fit

simultaneous principal curves to describe each lineage. The pseudotime of each cell was also obtained in this step and then visualized by ggplot R package (version 3.2.1).

#### Random forest

To determine the set of most discriminative features, we used random forest (RF) binary classification algorithm, implemented in the 'randomForest' R package (v4.6-14). Performance measures like AUC, sensitivity, specificity, recall and precision were calculated using 'caret' (v6.0-79) and 'ROCR' (v1.0-7) R packages. The analysis was performed using R version 3.4.1.

#### 16s rRNA gene metagenomic sequencing of liver and small intestine tissues

First, a host DNA depletion by osmotic lysis in distilled water, followed by PMA treatment was conducted (according to Marotz et al.) (19). Next, bacterial DNA was extracted using the PowerMag Microbiome DNA isolation kit following the manufacturer's instructions. DNA libraries were constructed, and high-throughput sequencing was performed on an MiSeq platform (Illumina). A negative control was used to test the potential presence of contaminant DNA; the sample showed no product.

#### 16S rRNA sequence data analysis

Raw fastq files were checked for overall data quality using FastQC (v0.11.5). All further preprocessing and analysis steps were performed in R (version 3.6.0). We used the dada2 (20) (v1.12.1) package to generate amplicon sequence variants (ASV) by following the authors recommended processing steps for paired sequencing data. Taxonomic assignment was performed using the DECIPHER (21) (v2.12.0) package against the SILVA SSU database (release 132).

All further analysis was performed within the phyloseq framework (version 1.32.0). Alpha diversities were calculated using the estimate\_richness function. Bray-Curtis distances were calculated using the distance function and differences between the resulting distance matrices were analyzed using PERMANOVA with the adonis function from the vegan package (version

2.5.6), again using categorized age and BMI covariates. Differential abundance in ASV counts were assessed using the metagenomeSeq package (22) (version 1.30.0). The differential abundance threshold was set to  $FDR < 0.05$  and an absolute  $\log_{2}FC > 0.5$ .

### Statistics

For non-transcriptomic data, two-tailed Mann–Whitney U test, paired t-test, Wilcoxon test, or one-way ANOVA (with Tukey post-hoc test) test were used to calculate significance, where applicable. P values below 5% were considered significant. Statistical calculation was performed using Prism program 5.0 (GraphPad Software).

### Study approval and ethics statements

The study was approved by the Ethics Committee of Hamburg (ID: PV4889) and the study protocol was conformed to the ethical guidelines of the 1975 Declaration of Helsinki. Informed consent was obtained from all patients and experiments were carried out in accordance with all relevant ethical regulations.

**Author contributions**

A W: designing research studies, conducting experiments, acquiring data, writing the manuscript

P S: designing research studies, conducting experiments, acquiring data, analyzing data

Y Z : designing research studies, analyzing data

J W: conducting experiments, acquiring data

F M: conducting experiments, acquiring data

C C : analyzing data

L K: conducting experiments, acquiring data

R W: conducting experiments, acquiring data

KF K: acquiring data

AD G: conducting experiments, acquiring data

A D: providing samples

S W: providing samples

A B: performing and supervising the microbiota analysis

V B: performing the random forest analysis

O M: providing samples

AW L: designing research studies

JR I: providing samples

N P: performing and supervising the microbiota analysis

S B: supervising the bioinformatic analysis and providing the infrastructure

S H: designing research studies

N G: designing research studies, supervising the project

A W, P S and Y Z designed the study together and cooperated to create and analyse the necessary data, contributing equally to the development of the project

**Acknowledgments**

This project was conducted thanks to the financial support of the “Deutsche Forschungsgemeinschaft, DFG”, SFB841 to S.H., N.G., SB, the “European Research Council” (715271 to N.G.), “RU 306” and “LFF-VF78” to YZ, VB, SB.

## References

1. **Younossi Z.M, Koenig AB, Abdelatif D et al.** *Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes.* *Hepatology.*2016.
2. **Adams L A, Lymp JF, St Sauver J et al.** *The natural history of nonalcoholic fatty liver disease: a population-based cohort study.* *Gastroenterology.*2005.
3. **Kleiner DE, Brunt EM, Van Natta M et al.** *Design and validation of a histological scoring system for nonalcoholic fatty liver disease.* *Hepatology.*2005.
4. **Winer S, Chan Y, Paltser G et al.** *Normalization of obesity-associated insulin resistance through immunotherapy.* *Nat Med.*2009.
5. **Gomes AL, Teijeiro A, Burén S et al.** *Metabolic inflammation-associated IL-17A causes non-alcoholic steatohepatitis and hepatocellular carcinoma.* *Cancer Cell.*2016.
6. **Rau M, Schilling AK, Meertens J et al.** *Progression from Nonalcoholic Fatty Liver to Nonalcoholic Steatohepatitis Is Marked by a Higher Frequency of Th17 Cells in the Liver and an Increased Th17/Resting Regulatory T Cell Ratio in Peripheral Blood and in the Liver.* *J Immunol.*2016.
7. **Inzaugarat ME, Ferreyra Solari NE, Billordo LA et al.** *Altered phenotype and functionality of circulating immune cells characterize adult patients with nonalcoholic steatohepatitis.* *J Clin Immunol.* 2011.
8. **Vonghia L, Magrone T, Verrijken A et al.** *Peripheral and Hepatic Vein Cytokine Levels in Correlation with Non-Alcoholic Fatty Liver Disease (NAFLD)-Related Metabolic, Histological, and Haemodynamic Features.* *PLoS One.*2015.
9. **Tedesco D, Thapa M, Chin CY, et al.** *Alterations in Intestinal Microbiota Lead to Production of Interleukin 17 by Intrahepatic  $\gamma\delta$  T-Cell Receptor-Positive Cells and Pathogenesis of Cholestatic Liver Disease.* *Gastroenterology.*2018.
10. **Tilg H, Cani PD, Mayer EA.** *Gut microbiome and liver diseases.* *Gut.*2016.

11. **Tripathi A, Debelius J, Brenner DA et al.** *The gut-liver axis and the intersection with the microbiome. Nat Rev Gastroenterol Hepatol.*2018.
12. **Boursier J, Mueller O, Barret M et al.** *The severity of nonalcoholic fatty liver disease is associated with gut dysbiosis and shift in the metabolic function of the gut microbiota. Hepatology.*2016.
13. **Zhu L, Baker SS, Gill C, et al.** *Characterization of gut microbiomes in nonalcoholic steatohepatitis (NASH) patients: a connection between endogenous alcohol and NASH. Hepatology.*2013.
14. **Sookoian S, Salatino A, Castaño GO et al.** *Intrahepatic bacterial metataxonomic signature in non-alcoholic fatty liver disease. Gut.*2020.
15. **Stuart T, Satija R.** *Integrative single-cell analysis. Nat Rev Genet.*2019.
16. **Kumar BV, Ma W, Miron M et al.** *Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. Cell Rep.*2017.
17. **Brockmann L, Soukou S, Steglich B et al.** *Molecular and functional heterogeneity of IL-10-producing CD4+ T cells. Nat Commun.*2018.
18. **Street K, Risso D, Fletcher RB et al.** *Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics.*2018.
19. **Marotz CA, Sanders JG, Zuniga C et al.** *Improving saliva shotgun metagenomics by chemical host DNA depletion. Microbiome.*2018.
20. **Callahan BJ., McMurdie PJ, Rosen MJ et. al.** *DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods.*2016.
21. **Murali A, Bhargava A, Wright ES.** *IDTAXA: a novel approach for accurate taxonomic classification of microbiome sequences. Microbiome.*2018.
22. **Paulson JN, Stine OC, Bravo HC et al.** *Differential abundance analysis for microbial marker-gene surveys. Nat Methods.*2013.
23. **Brummelman J, Pilipow K, Lugli E.** *The Single-Cell Phenotypic Identity of Human CD8+ and CD4+ T Cells. Int Rev Cell Mol Biol.*2018.

24. **Aizarani N, Saviano A, Sagar et al.** *A human liver cell atlas reveals heterogeneity and epithelial progenitors. Nature.2019.*
25. **MacParland SA, Liu JC, Ma XZ et al.** *Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat Commun.2018.*
26. **Steinke FC, Yu S, Zhou X et al.** *TCF-1 and LEF-1 act upstream of Th-POK to promote the CD4(+)T cell fate and interact with Runx3 to silence Cd4 in CD8(+)T cells. Nat Immunol.2014.*
27. **Szabo PA, Miron M, Farber DL.** *Location, location, location: Tissue resident memory T cells in mice and humans. Sci Immunol.2019.*
28. **Amezcuca Vesely MC, Pallis P et al.** *Effector TH17 Cells Give Rise to Long-Lived TRM Cells that Are Essential for an Immediate Response against Bacterial Infection. Cell.2019.*
29. **Hirota K, Duarte JH, Veldhoen M, et al.** *Fate mapping of IL-17-producing T cells in inflammatory responses. Nature Immunology.2011.*
30. **Gagliani N, Amezcuca Vesely MC, Iseppon A et al.** *Th17 cells transdifferentiate into regulatory T cells during resolution of inflammation. Nature.2015.*
31. **Xu H, Agaloti T, Zhao J, et al.** *The induction and function of the anti-inflammatory fate of TH17 cells. Nat Commun.2020.*
32. **Weinreich MA, Takada K, Skon C et al.** *KLF2 transcription-factor deficiency in T cells results in unrestrained cytokine production and upregulation of bystander chemokine receptors. Immunity.2009.*
33. **Zhao L, Qiu DK, Ma X.** *Th17 cells: The emerging reciprocal partner of regulatory T cells in the liver. J Digest Dis.2010.*
34. **Mensali N, Grenov A, Pati NB et al.** *Antigen-delivery through invariant chain (CD74) boosts CD8 and CD4 T cell immunity. Oncoimmunology.2019.*
35. **Su H, Na N, Zhang X et al.** *The biological function and significance of CD74 in immune diseases. Inflamm Res.2017.*

36. **Haran M, Mirkin V, Braester A et al.** *A phase I-II clinical trial of the anti-CD74 monoclonal antibody milatuzumab in frail patients with refractory chronic lymphocytic leukaemia: A patient based approach. Br J Haematol.*2018.
37. **Martin P, Furman RR, Rutherford S et al.** *Phase I study of the anti-CD74 monoclonal antibody milatuzumab (hLL1) in patients with previously treated B-cell lymphomas. Leuk Lymphoma.*2015.
38. **Schuppan D, Surabattula R, Wang XY.** *Determinants of fibrosis progression and regression in NASH. J Hepatol.*2018.
39. **Honda K, Littman DR.** *The microbiota in adaptive immune homeostasis and disease. Nature.*2016.
40. **Bertola A, Bonnafous S, Anty R et al.** *Hepatic expression patterns of inflammatory and immune response genes associated with obesity and NASH in morbidly obese patients. PLoS ONE.*2010.
41. **Fabre T, Molina MF, Soucy G et al.** *Type 3 cytokines IL-17A and IL-22 drive TGF- $\beta$ -dependent liver fibrosis. Sci Immunol.*2018.
42. **Heinrichs D, Knauel M, Offermanns C et al.** *Macrophage migration inhibitory factor (MIF) exerts antifibrotic effects in experimental liver fibrosis via CD74. Proc Natl Acad Sci U S A.*2011.
43. **Leclercq S, Cani PD, Neyrinck AM et al.** *Role of intestinal permeability and inflammation in the biological and behavioral control of alcohol-dependent subjects. Brain. Behav. Immun.*2012.
44. **Filliol A, Piquet-Pellorce C, Raguénès-Nicol C et al.** *RIPK1 protects hepatocytes from Kupffer cells-mediated TNF-induced apoptosis in mouse models of PAMP-induced hepatitis. J Hepatol.*2017.
45. **Ramakrishna C, Kujawski M, Chu H et al.** *Bacteroides fragilis polysaccharide A induces IL-10 secreting B and T cells that prevent viral encephalitis. Nat Commun.*2019.



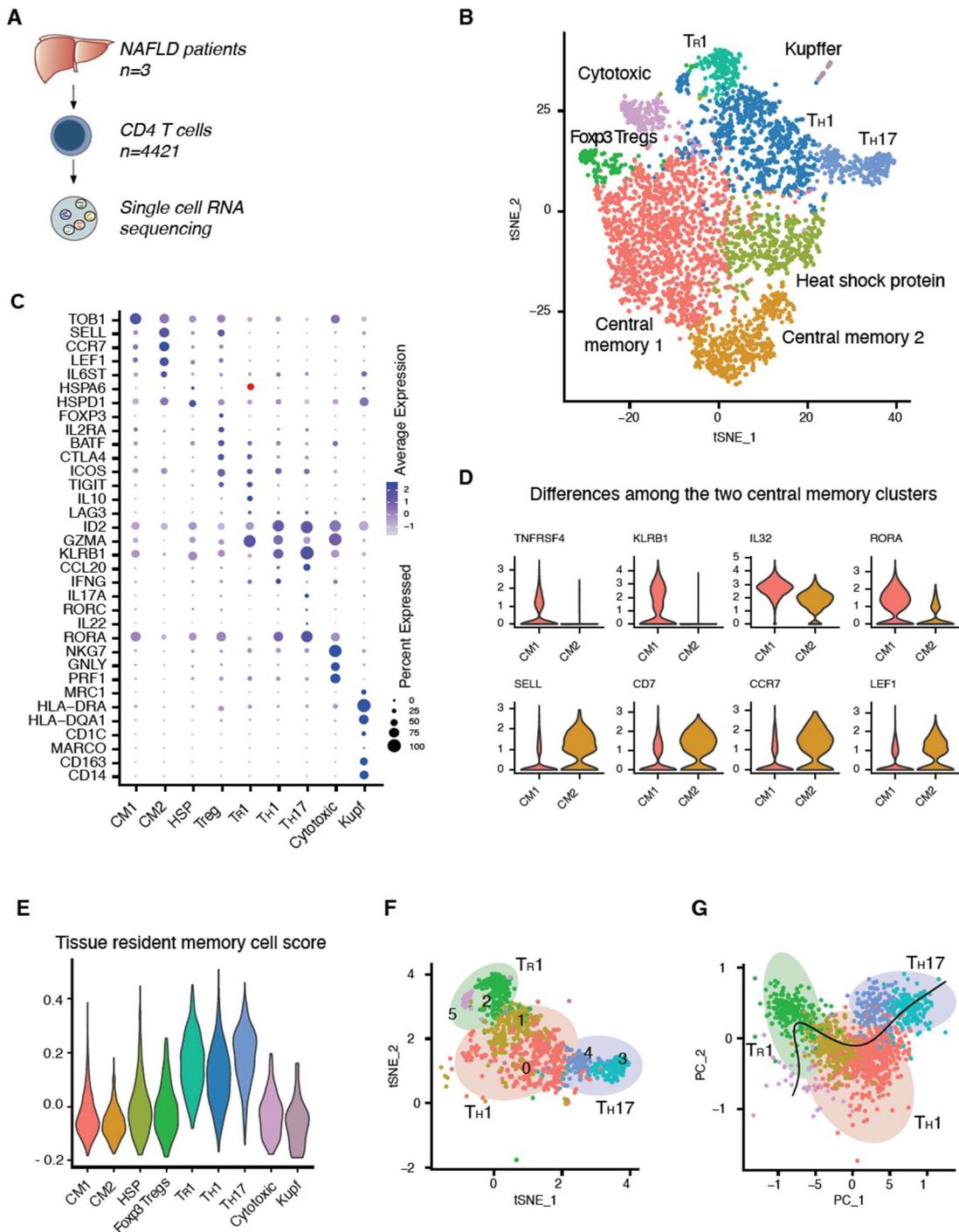


Figure 1

**Cellular map of CD4<sup>+</sup> T cells in NAFLD patients** (A) Schematic of experimental setup: FACS-sorted CD4<sup>+</sup> T cells from liver tissue of NAFLD were processed for scRNA-seq (B) t-SNE plot of unsupervised clustering of T cells (C) Dotplot of CD4<sup>+</sup> T cell clusters with unique signature genes (D) Violin plots comparing the mRNA expression profile of different genes in the central memory clusters (E) Comparison of expression profiles of cells with published expression signatures of human CD4<sup>+</sup> TRM cells. (F) t-SNE plot of subclustering of T helper cell clusters TR1, TH1 and TH17 (G) PC-plot of pseudotime developmental trajectory analysis.

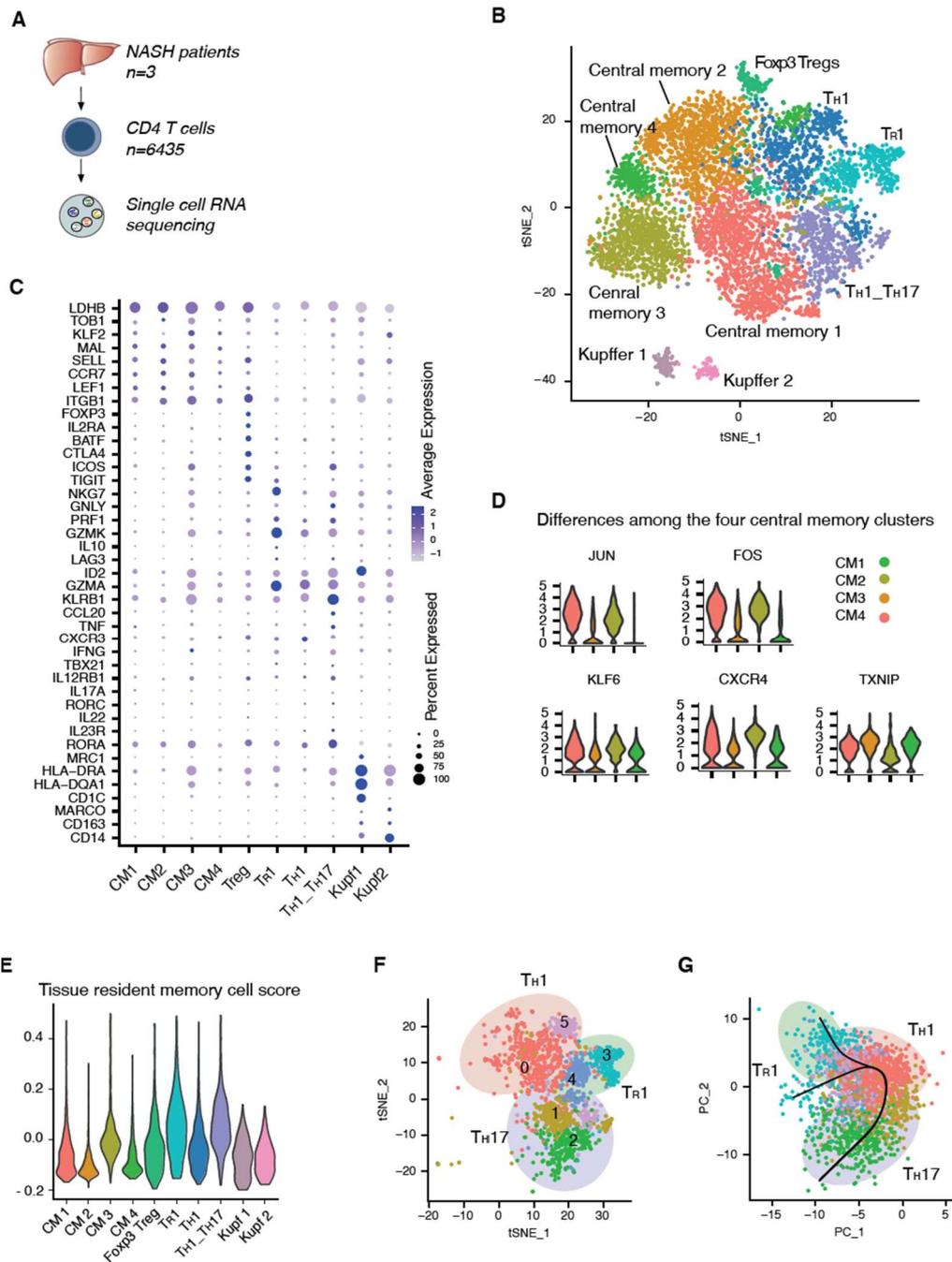


Figure 2

**Cellular map of CD4<sup>+</sup> T cells in NASH patients (A)** Schematic of experimental setup: FACS-sorted CD4<sup>+</sup> T cells from liver tissue of NASH were processed for scRNA-seq **(B)** t-SNE plot of unsupervised clustering of T cells **(C)** Dotplot of CD4<sup>+</sup> T cell clusters with unique signature genes **(D)** Violin plots comparing the mRNA expression profile of different genes in the central memory clusters **(E)** Comparison of expression profiles of cells with published expression signatures of human CD4<sup>+</sup> TRM cells. **(F)** t-SNE plot of subclustering of T helper cell clusters TR1, TH1 and TH17 **(G)** PC-plot of pseudotime developmental trajectory analysis.

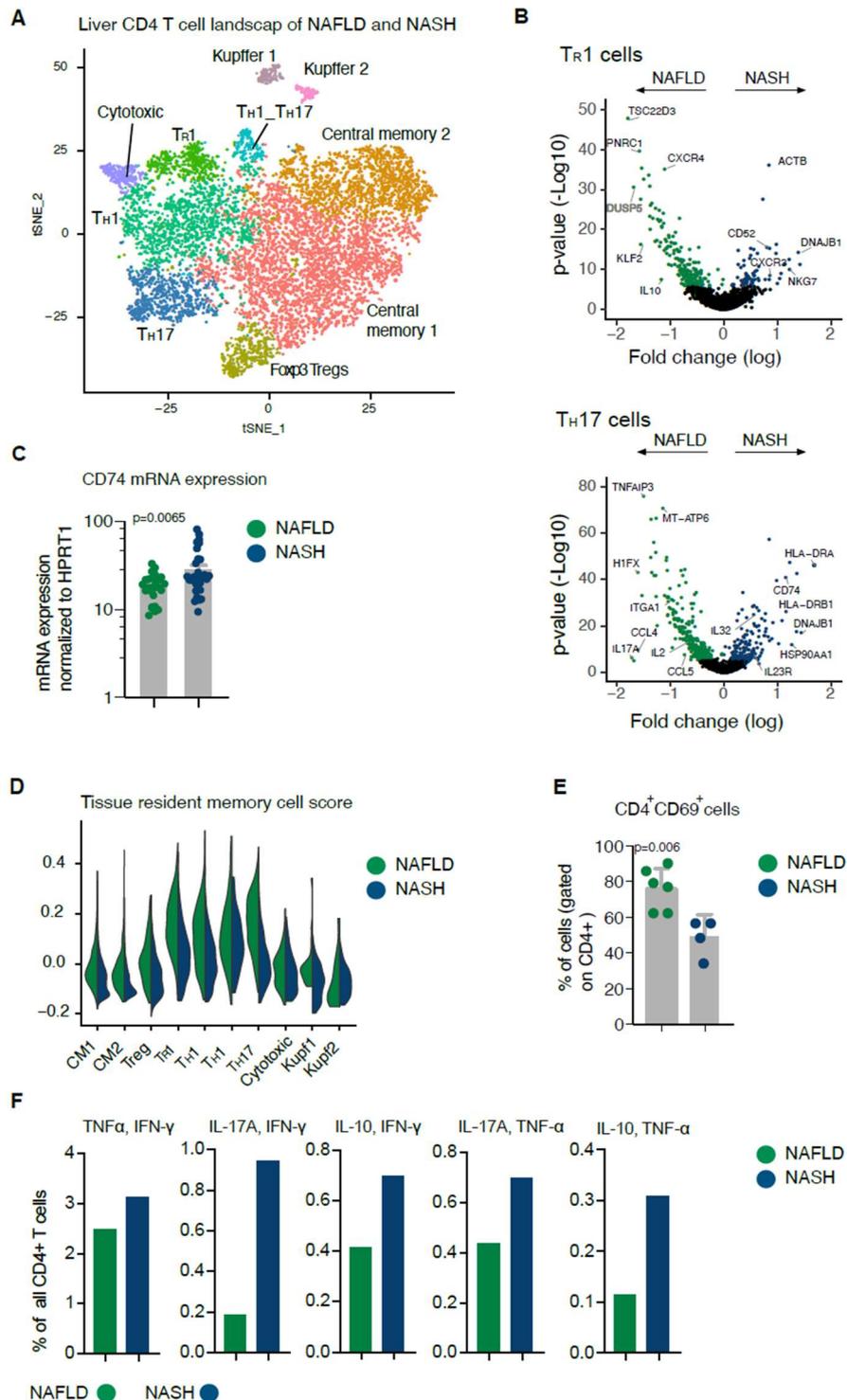


Figure 3

**Differences between liver CD4<sup>+</sup> T cells in NAFLD and NASH patients (A)** combined t-SNE plot of unsupervised clustering of T cells of NAFLD and NASH **(B)** Volcano plot showing the differential RNA expression profiles of liver TR1 and TH17 cells **(C)** Relative mRNA expression of CD74 level in liver tissues from NAFLD (n=25) and NASH (n=32) patients. Data are presented as mean  $\pm$  SEM **(D)** Comparison of expression profiles of NAFLD and NASH cells with published expression signatures of human CD4<sup>+</sup> TRM cells **(E)** Frequencies of CD69<sup>+</sup> cells within CD4<sup>+</sup> T cells (unpaired t-test) **(F)** Frequencies of indicated cytokine combinations within CD4<sup>+</sup> T cells based on RNA level

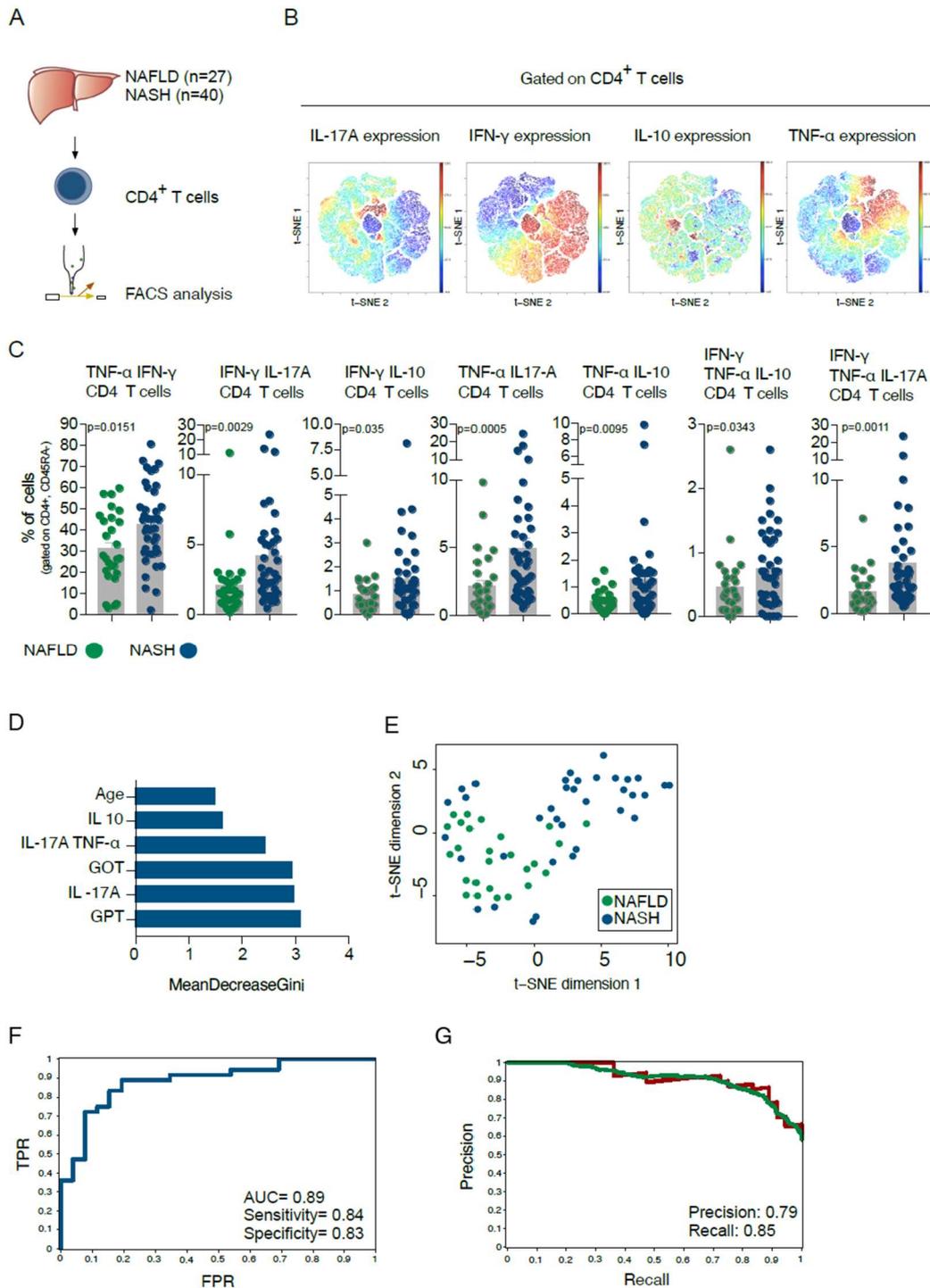


Figure 4

**Liver CD4<sup>+</sup> T cells in NASH have a distinct cytokine profile compared to NAFLD (A)** Schematic of experimental setup: CD4<sup>+</sup>, CD45RA<sup>-</sup> T cells freshly isolated from liver tissue of NAFLD and NASH patients were analyzed by FACS analysis (B) Representative tSNE analysis of liver infiltrating memory T cells of 5 NASH patients. (C) Frequencies of indicated cytokine combinations within liver CD4<sup>+</sup> T cells based on protein level (unpaired t-test) (D) Top 6 features for NAFLD/NASH prediction (E) tSNE visualization of patients on the basis of indicated top 6 features of prediction (F) Receiver-operating characteristic (ROC) curve showing true- and false-positive rates for the discrimination between NAFLD and NASH prediction based on the top 6 features of prediction (G) Precision and Recall (PR) Plot for prediction based on indicated top 6 features of prediction.

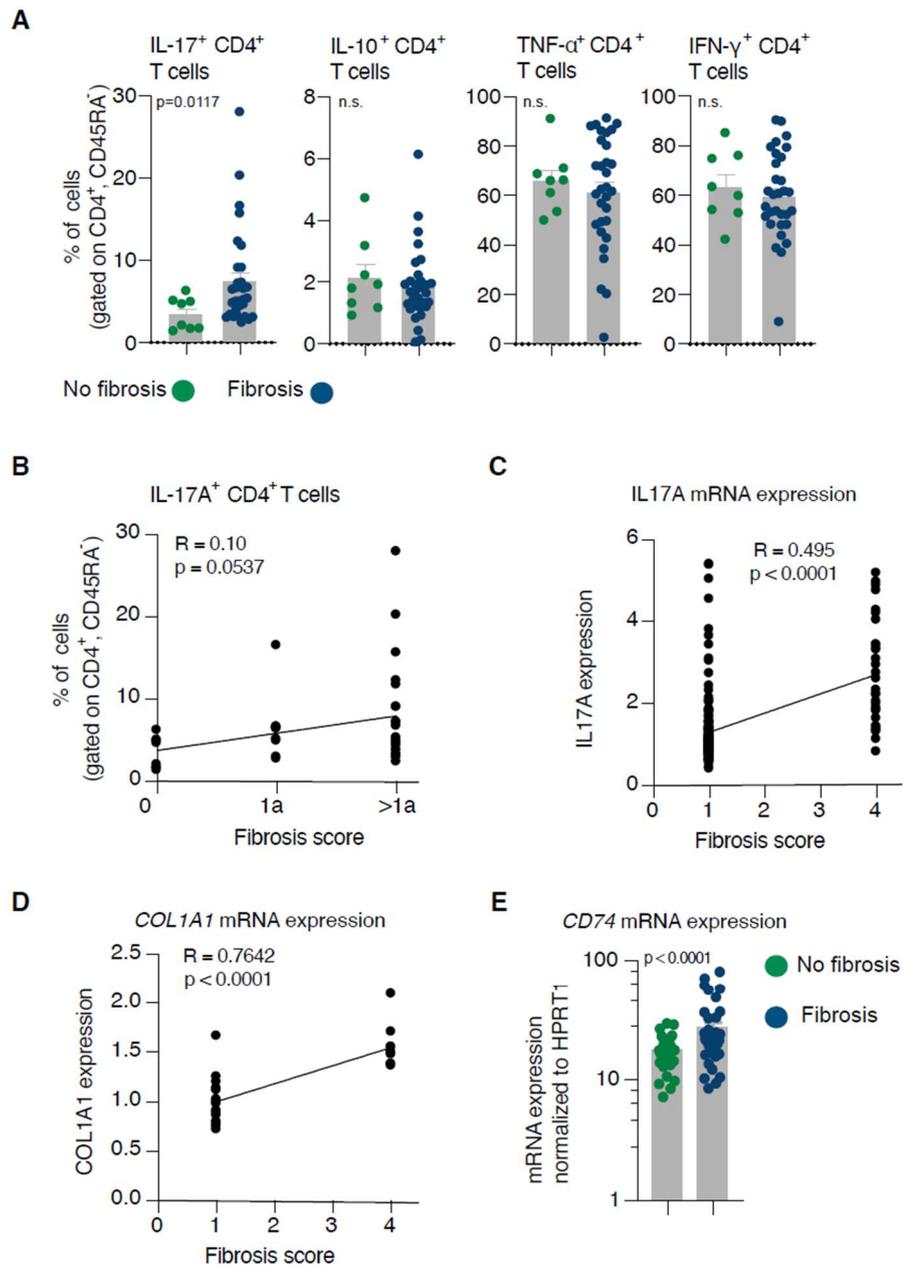


Figure 5

**IL-17A producing CD4<sup>+</sup> T are enriched in the liver of NASH patients with fibrosis (A)**

Frequencies of indicated cytokine combinations within liver CD4<sup>+</sup> T cells based on protein level (unpaired t-test). Data are presented as mean  $\pm$  SEM (B) Correlation between Fibrosis score and frequency of IL-17A producing T cells in the liver (C) Correlation between Fibrosis score and *IL17A* expression in the liver of NASH patients (D) Correlation between Fibrosis score and *COL1A1* expression in the liver of NASH patients (E) Relative mRNA expression of CD74 level in liver tissues of NASH patients. Data are presented as mean  $\pm$  SEM

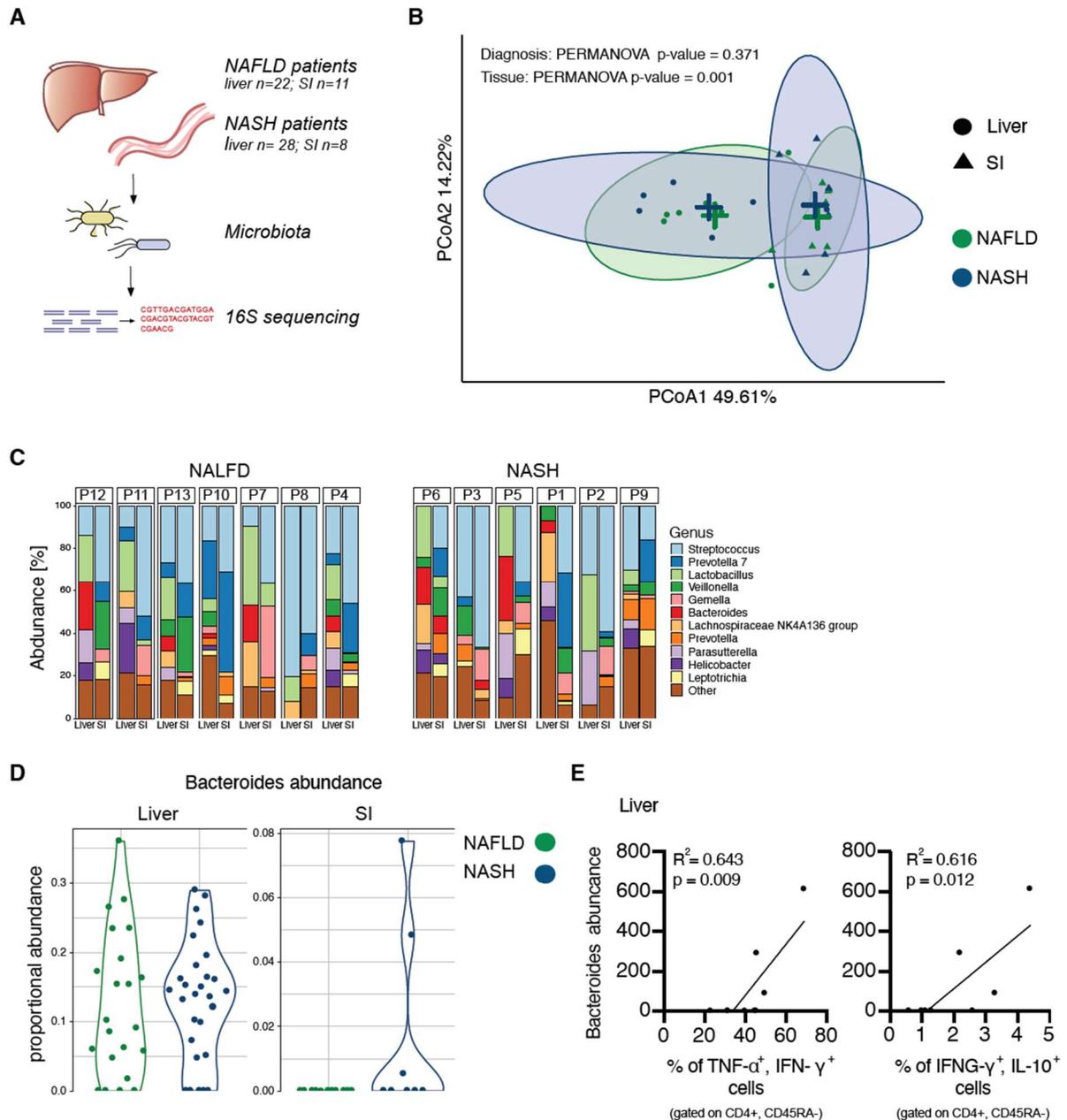
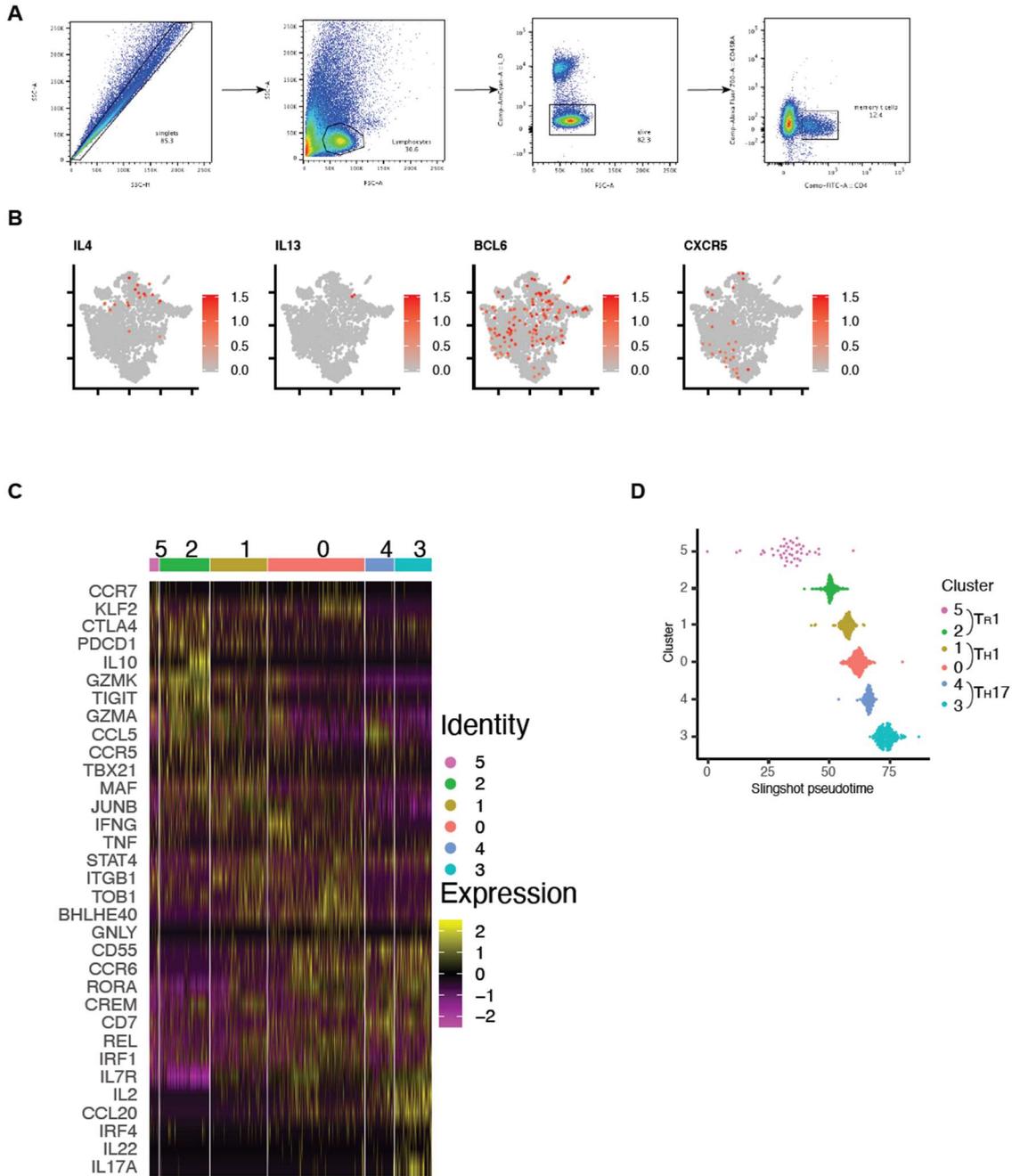
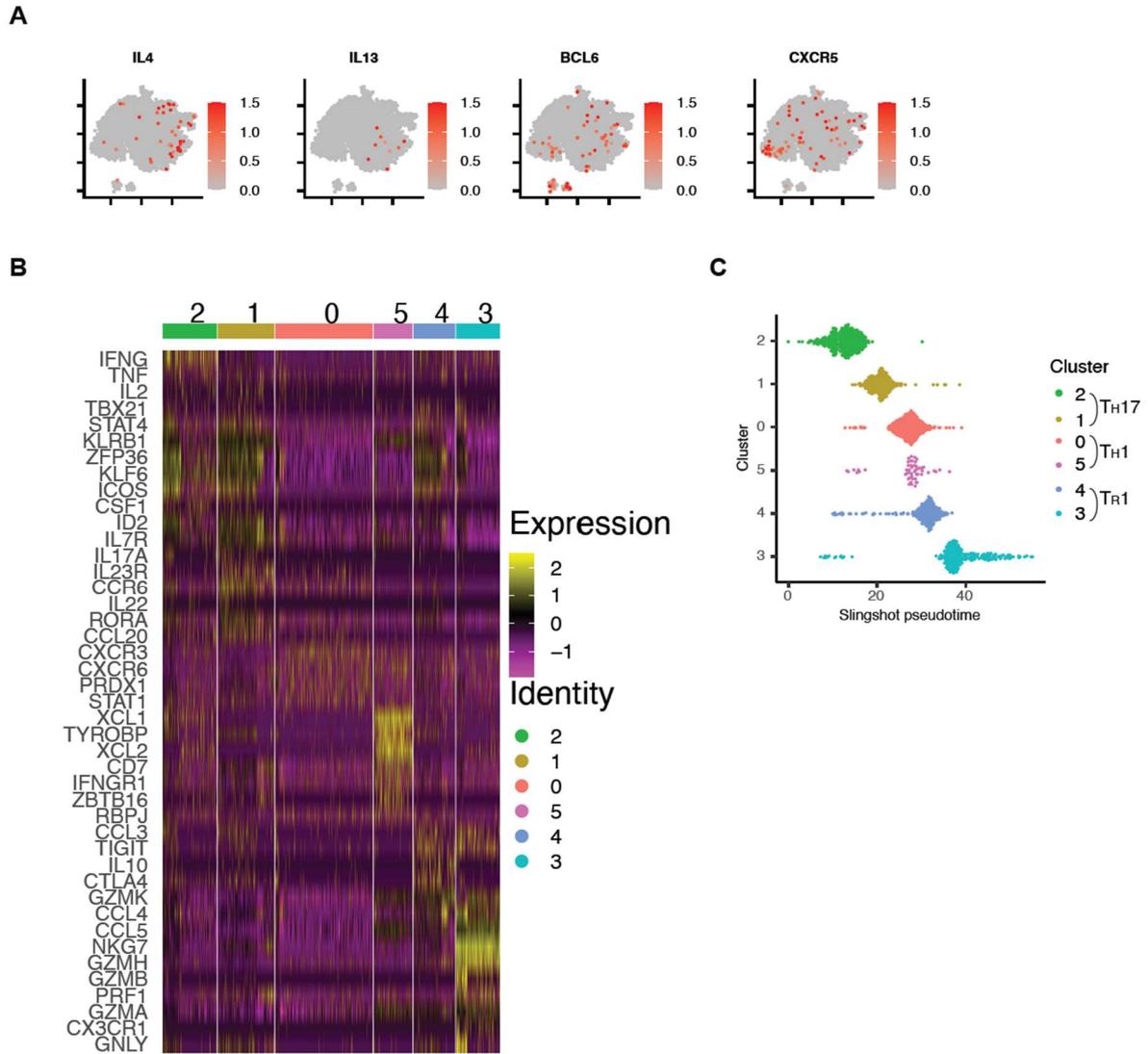


Figure 6

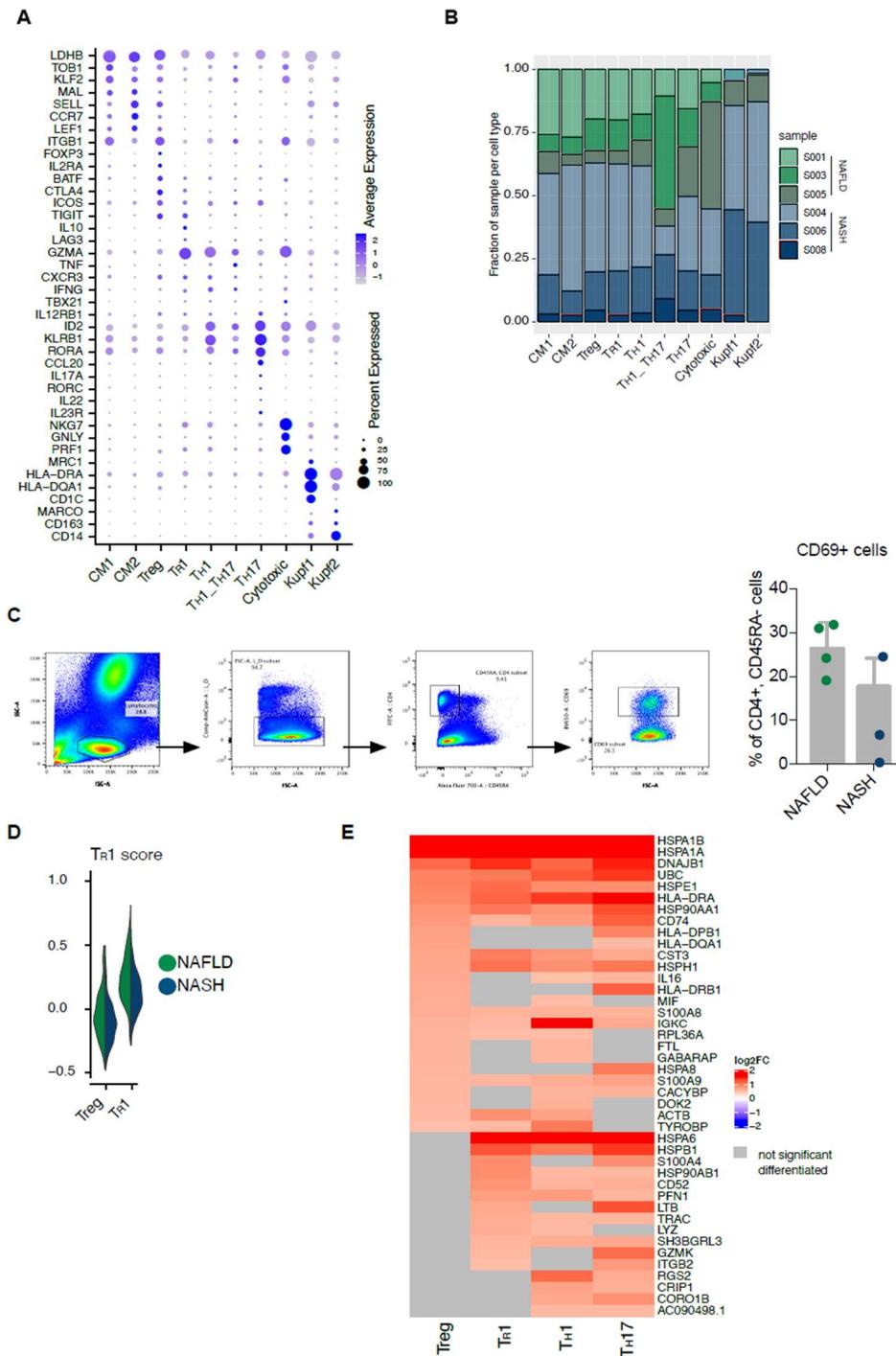
**Translocation of microbiota into the liver of NASH patients (A)** Schematic of experimental setup: Microbiota isolated from liver and small intestine tissue was subjected to 16S rRNA sequencing **(B)** PCoA plot showing the differences of microbial  $\beta$ -diversity between NAFLD and NASH, liver and small intestine samples **(C)** Differential abundance of *Bacteroides* between NAFLD and NASH patients in liver and small intestine tissue **(D)** Correlation between *Bacteroides* abundance in the small intestine and frequency of cytokine producing T cells in the liver



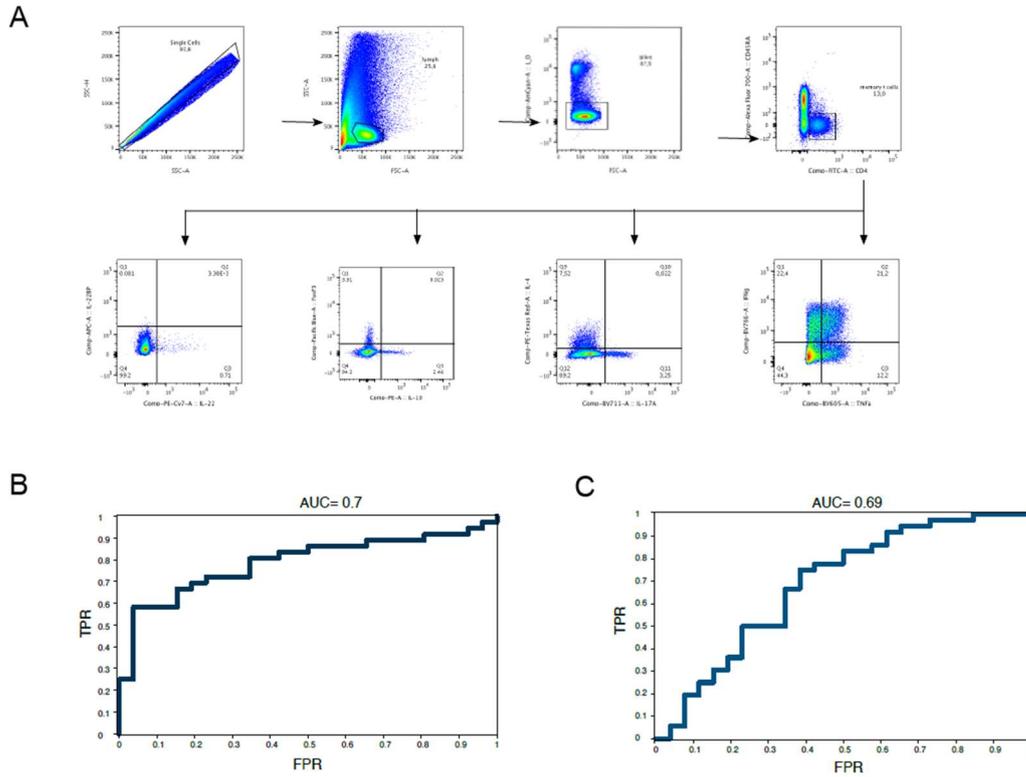
**Suppl. Figure 1. (A)** Sorting strategy of  $CD4^+$ ,  $CD45^-$  cells for single cell sequencing **(B)** tSNE map with mRNA expression profiles of  $T_H2$ ,  $T_H9$  and  $T_{FH}$  genes **(C)** Heat map depicting the average expression levels per subcluster of differentially expressed markers in each cluster **(D)** Slingshot pseudotime analysis of T helper cell subclusters



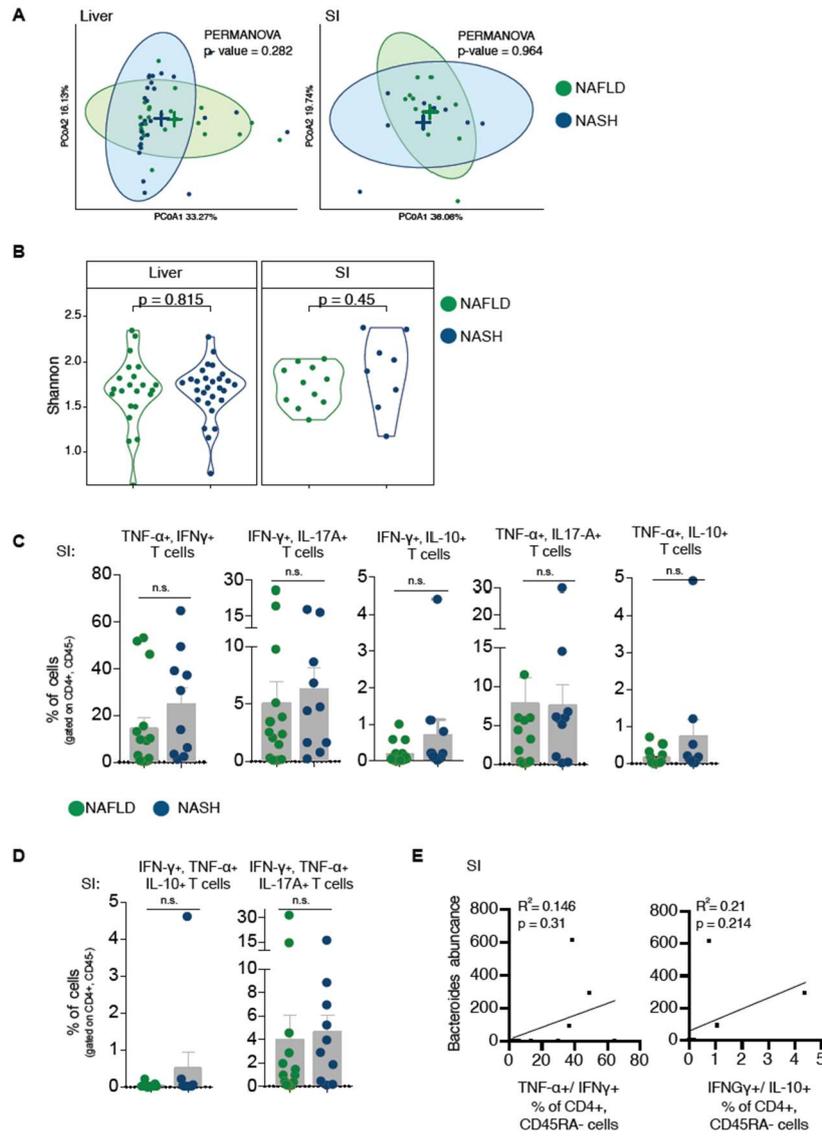
**Suppl. Figure 2. (A)** tSNE map with mRNA expression profiles of  $T_{H2}$ ,  $T_{H9}$  and  $T_{FH}$  genes **(B)** Heat map depicting the average expression levels per subcluster of differentially expressed markers in each cluster **(C)** Slingshot pseudotime analysis of T helper cell subclusters



**Suppl. Figure 3. (A)** Dotplot of CD4<sup>+</sup> T cell clusters with unique signature genes **(B)** The fractions of NAFLD and NASH samples per cell type **(C)** Flow cytometry analysis of unstimulated liver samples (left) and fraction of CD69<sup>+</sup> cells within CD4<sup>+</sup>, CD45<sup>-</sup> cells (right) **(D)** Comparison of expression profiles of NAFLD and NASH cells with Tr1 score **(E)** Heatmap showing the differential RNA expression analysis across T cell clusters



**Suppl. Figure 4. (A)** Flow cytometry analysis of stimulated liver samples **(B)** ROC curve showing true- and false-positive rates for the discrimination between NAFLD and NASH prediction based on sex, age, BMI, transaminase levels **(C)** ROC curve showing true- and false-positive rates for the discrimination between NAFLD and NASH prediction based on cytokine levels



**Suppl. Figure 5. (A)** Differences of microbial  $\beta$ -diversity between NAFLD and NASH patients within liver and small intestine samples.  $\beta$ -diversity was visualised by constrained analysis of principal coordinates using Bray-Curtis distance. **(B)** Violin plots for Shannon  $\alpha$ -diversity within liver and small intestine tissue **(C)** Frequencies of indicated cytokine combinations within small intestinal CD4<sup>+</sup> T cells based on protein level (unpaired t-test) **(D)** Frequencies of indicated cytokine combinations within small intestinal CD4<sup>+</sup> T cells based on protein level (unpaired t-test) **(E)** Correlation between *Bacteroides* abundance in the small intestine and frequency of cytokine producing T cells in the small intestine



## Chapter 5: Publication in *Science Immunology*, February 2021

### Clonal expansion and activation of tissue-resident memory-like T<sub>H</sub>17 cells expressing GM-CSF in the lungs of severe COVID-19 patients

#### Specific contributions:

The following research article has been published in *Science Immunology*. I am the first author of the paper. I processed and analyzed the single cell RNA-seq, CITE-seq and TCR-seq data of immune cells sorted from lung and blood of COVID-19 and bacterial pneumonia patients. These results are explicitly displayed in Figures: 1B, C, D, E, G; 2A, B, C, D, E, F, G, H, I, J; 3A, B, C; 5B, C and supplementary Figures: 1C, D, E; 2A, B, C; 5A, B, C, D, E; 6B, C; 7A, B, C, D, E, F, G, H; 8A, B, C; 10A. I participated in the discussion of data interpretation and plans for experimental validation. I contributed to the editing of the manuscripts before the first submission and during the revisions.

I, Stefan Bonn, agree with the above statements as the direct supervisor.

Signature

,

Date & Place

15.11.2021 Hamburg



## CORONAVIRUS

# Clonal expansion and activation of tissue-resident memory-like T<sub>H</sub>17 cells expressing GM-CSF in the lungs of patients with severe COVID-19

Yu Zhao<sup>1,2,3,4\*</sup>, Christoph Kilian<sup>1\*</sup>, Jan-Eric Turner<sup>3,5\*</sup>, Lidia Bosurgi<sup>6,7\*</sup>, Kevin Roedi<sup>8\*</sup>, Patricia Bartsch<sup>1</sup>, Ann-Christin Gnirck<sup>3,5</sup>, Filippo Cortesi<sup>6</sup>, Christoph Schultheiß<sup>9</sup>, Malte Hellmig<sup>1</sup>, Leon U.B. Enk<sup>1</sup>, Fabian Hausmann<sup>2</sup>, Alina Borchers<sup>1</sup>, Milagros N. Wong<sup>5</sup>, Hans-Joachim Paust<sup>1</sup>, Francesco Siracusa<sup>6</sup>, Nicola Scheibel<sup>6</sup>, Marissa Herrmann<sup>6</sup>, Elisa Rosati<sup>10</sup>, Petra Bacher<sup>10,11</sup>, Dominik Kyllies<sup>5</sup>, Dominik Jarczak<sup>8</sup>, Marc Lütgehetmann<sup>12</sup>, Susanne Pfefferle<sup>12</sup>, Stefan Steurer<sup>13</sup>, Julian Schulze zur Wiesch<sup>6</sup>, Victor G. Puelles<sup>5</sup>, Jan-Peter Sperhake<sup>14</sup>, Marylyn M. Addo<sup>15,16</sup>, Ansgar W. Lohse<sup>6</sup>, Mascha Binder<sup>9</sup>, Samuel Huber<sup>3,6</sup>, Tobias B. Huber<sup>3,5</sup>, Stefan Kluge<sup>8</sup>, Stefan Bonn<sup>2,3,4</sup>, Ulf Panzer<sup>1,3†</sup>, Nicola Gagliani<sup>3,6,17,18†‡</sup>, Christian F. Krebs<sup>1,3†‡</sup>

Hyperinflammation contributes to lung injury and subsequent acute respiratory distress syndrome with high mortality in patients with severe coronavirus disease 2019 (COVID-19). To understand the underlying mechanisms involved in lung pathology, we investigated the role of the lung-specific immune response. We profiled immune cells in bronchoalveolar lavage fluid and blood collected from patients with severe COVID-19 and patients with bacterial pneumonia not associated with viral infection. By tracking T cell clones across tissues, we identified clonally expanded tissue-resident memory-like T<sub>H</sub>17 cells (T<sub>RM</sub>17 cells) in the lungs even after viral clearance. These T<sub>RM</sub>17 cells were characterized by a potentially pathogenic cytokine expression profile of *IL17A* and *CSF2* (GM-CSF). Interactome analysis suggests that T<sub>RM</sub>17 cells can interact with lung macrophages and cytotoxic CD8<sup>+</sup> T cells, which have been associated with disease severity and lung damage. High IL-17A and GM-CSF protein levels in the serum of patients with COVID-19 were associated with a more severe clinical course. Collectively, our study suggests that pulmonary T<sub>RM</sub>17 cells are one potential orchestrator of the hyperinflammation in severe COVID-19.

## INTRODUCTION

On 11 March 2020, the World Health Organization (WHO) communicated that the spread of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) had reached pandemic status. By the end of 2020, there were more than 80 million confirmed cases including 1.7 million deaths (1). These epidemiological data highlight

the need to rapidly develop therapies for treating COVID-19 that reduce the high case fatality rate. The promising results of the clinical trial RECOVERY, in which dexamethasone was administered to 2104 patients with COVID-19 (2), suggest that one of the causes of the acute respiratory distress syndrome and ultimately death of patients with COVID-19 is the hyperactivation of the immune system. Supporting the pathogenic role of immune hyperactivation, the use of neutralizing antibodies, blocking, for example, granulocyte-macrophage colony-stimulating factor (GM-CSF) and interleukin-1β (IL-1β), has shown encouraging clinical results (3–6). The efficacy of a therapy blocking IL-6 has not yet been broadly recognized (7), but one recent study showed that tocilizumab reduced disease progression in patients with COVID-19 not receiving mechanical ventilation (8).

Considering that peripheral blood myeloid cells appear not to be able to produce high amounts of proinflammatory cytokines (9) and the numbers of blood T cells are reduced in patients with COVID-19 (10, 11), the lungs may serve as a reservoir for cells producing these cytokines. However, additional investigation into the lung-specific cellular source of the proinflammatory cytokines typical of severe COVID-19, including IL-6, tumor necrosis factor-α (TNF-α), IL-1β, and IL-17A, is needed. Using single-cell RNA sequencing (scRNA-seq) methods, it has been shown that proinflammatory macrophages expressing *IL6*, *IL1B*, and *TNF* and CD8<sup>+</sup> T cells with a tissue-resident cytotoxic signature are present in the bronchoalveolar lavage (BAL) and upper respiratory tract of patients with COVID-19 (12, 13). The accumulation of interferon-γ (IFN-γ)-producing CD4<sup>+</sup> T cells in the BAL of patients with COVID-19 has also recently been described (14). Still, the role of CD4<sup>+</sup> T cells and of their different polarization states, especially at the site of infection, needs to be further elucidated.

<sup>1</sup>III. Department of Medicine, Division of Translational Immunology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>2</sup>Institute of Medical Systems Biology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>3</sup>Hamburg Center for Translational Immunology (HCTI), University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>4</sup>Center for Biomedical AI, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>5</sup>III. Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>6</sup>I. Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>7</sup>Protozoa Immunology, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany. <sup>8</sup>Department of Intensive Care Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>9</sup>Department of Internal Medicine IV, Oncology/Hematology, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany. <sup>10</sup>Institute of Clinical Molecular Biology, Christian-Albrechts University of Kiel, Kiel, Germany. <sup>11</sup>Institute of Immunology, Christian-Albrechts-University of Kiel and UKSH Schleswig-Holstein, Kiel, Germany. <sup>12</sup>Institute for Medical Microbiology, Virology and Hygiene, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>13</sup>Institute for Pathology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>14</sup>Department of Legal Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>15</sup>I. Department of Medicine, Division of Infectious Diseases, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>16</sup>German Center for Infection Research, Partner Site Hamburg-Lübeck-Borstel-Riems, Hamburg, Germany. <sup>17</sup>Department of General, Visceral and Thoracic Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>18</sup>Immunology and Allergy Unit, Department of Medicine, Solna, Karolinska Institute and University Hospital, Stockholm, Sweden. \*These authors contributed equally to this work. †These authors jointly supervised this work. ‡Corresponding author (alphabetically ordered). Email: n.gagliani@uke.de (N.G.); c.krebs@uke.de (C.F.K.)

CD4<sup>+</sup> T cells orchestrate the immune response for example, by affecting on macrophage function and activation of cytotoxic CD8<sup>+</sup> T cells. To mediate these different functions, naïve CD4<sup>+</sup> T cells differentiate into effector cells characterized by different polarization states such as T helper cell 1 (T<sub>H1</sub>) and T<sub>H17</sub>. We have shown that T<sub>H17</sub> cells can infiltrate the lungs and acquire a tissue-resident phenotype during bacterial infection (15). Upon infectious stimuli, these tissue-resident long-lived cells can reacquire the original cytokine profile, for example, IL-17A/F, or switch toward the production of IFN- $\gamma$ , the signature cytokine of the T<sub>H1</sub> polarization state. Although these long-lived tissue-resident T<sub>H17</sub>, referred to here as T<sub>RM17</sub> cells, usually exert a protective function (15), we have also recently shown that these cells can contribute to immune-mediated inflammatory diseases (16). Whether T<sub>RM17</sub> cells are present in the lungs of patients with COVID-19 and how they interact with other potentially pathogenic immune cells of these patients remains to be studied.

Here, we identified two populations of T<sub>H17</sub> cells in the BAL fluid (BALF) of patients with COVID-19. One of these was mainly resident in the lung, characterized by the expression of GM-CSF and shared clones with T<sub>H1</sub> cells. Moreover, we provide a lung-specific immune cell-cell interaction map showing the potential role of T<sub>RM17</sub> cells in support of the already known pathological immune players, such as proinflammatory and profibrotic macrophages, and cytotoxic CD8<sup>+</sup> T cells (12, 13). These data provide support for continuing to test anticytokine therapies including those that have undergone preliminary clinical testing, for example, GM-CSF neutralization (17, 18), or those now under consideration, for example, anti-IL-17A/F treatment (19).

## RESULTS

### Immune profile of T cells and myeloid cells in the blood and BALF of patients with COVID-19

To provide a detailed analysis of the lung-specific and peripheral immune responses in COVID-19, BALF and peripheral blood were taken from nine patients with severe COVID-19. The major goal of the study was to analyze the tissue-specific immune response in patients with COVID-19, with a particular focus on T cells. In addition, we also included BALF and peripheral blood mononuclear cells (PBMCs) of five patients with bacterial pneumonia, not associated with viral infection. All samples were analyzed by flow cytometry and T cells were fluorescence-activated cell sorting (FACS)-sorted and subjected to scRNA and T cell receptor (TCR) sequencing, as well as to sequencing-based epitope measurement [cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq)]. From BAL samples, CD3<sup>-</sup> non-T cells, including mainly myeloid cells, were also analyzed by scRNA-seq and CITE-seq (56,735 cells from the BAL and 77,457 cells from the peripheral blood) (Fig. 1A and fig. S1, A and B).

All patients were treated on the intensive care unit at the University Medical Center Hamburg-Eppendorf. Eight of nine patients with COVID-19 and all patients with bacterial pneumonia were on mechanical ventilation at time of sampling, indicating the severity of disease, further reflected by high mortality in both groups (Fig. 1B). Detailed patient characteristics of patients with SARS-CoV-2 and patients with bacterial pneumonia, including comorbidities and relevant medications, are presented in table S1. In eight of the nine patients with COVID-19, symptomatic SARS-CoV-2 infection was diagnosed

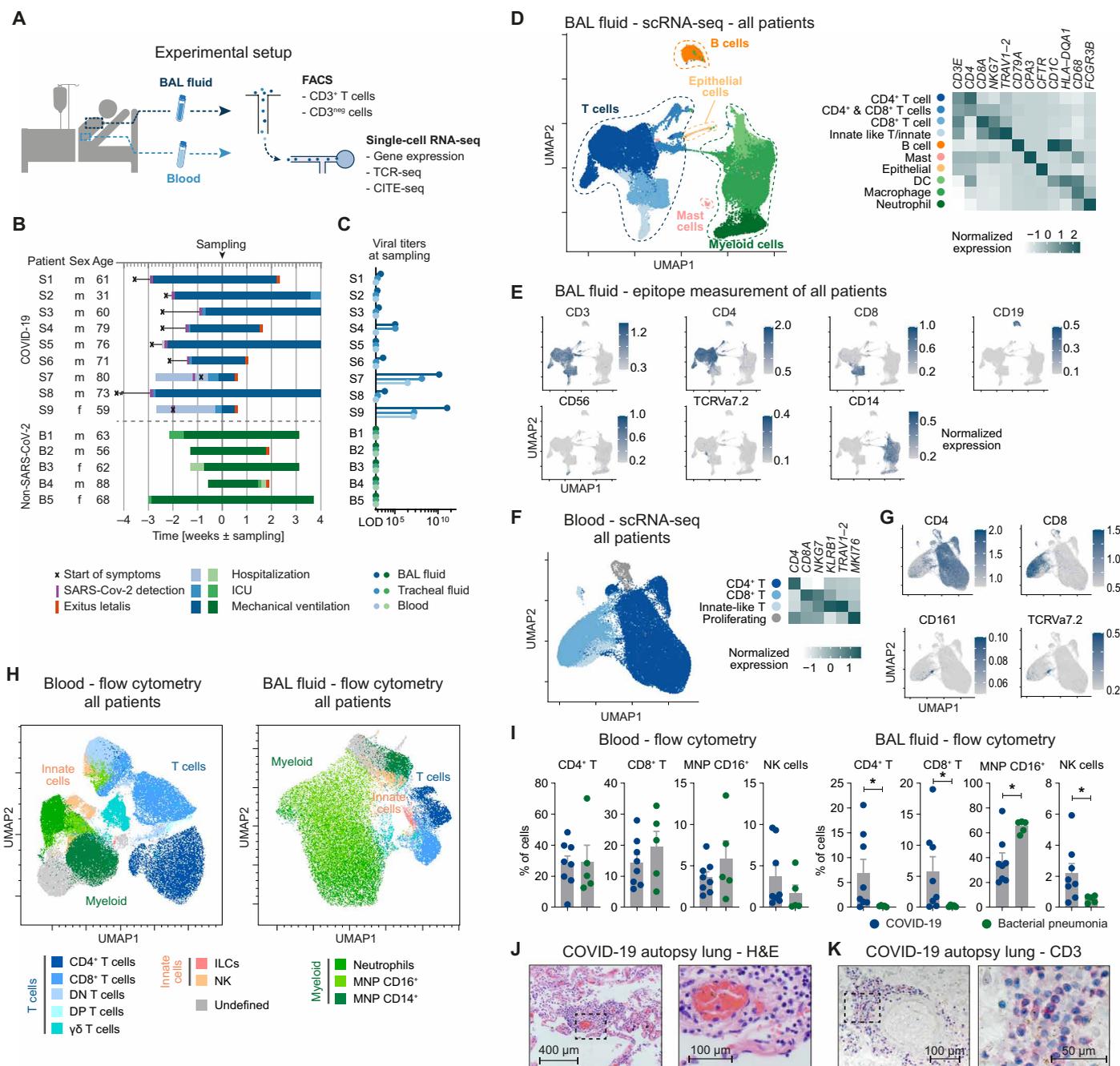
>2 weeks before BAL sampling, whereas in the remaining patients, SARS-CoV-2 was detected 8 days before sampling (Fig. 1B). In line with this, viral titers in BALF, tracheal secretion fluid, and blood of the majority of patients were very low or negative at time of sampling (Fig. 1C).

To examine the immune profile in the lungs of patients with COVID-19 and, at the same time, achieve robust clustering, we integrated the scRNA-seq datasets of patients with COVID-19 and bacterial pneumonia. The analysis revealed the presence of five main clusters in the BALF based on key signature genes and standard surface markers (T cells, B cells, mast cells, myeloid cells, and epithelial cells) represented in the Uniform Manifold Approximation and Projection (UMAP) dimension reduction (Fig. 1D). The T cell cluster could be further subdivided into CD4<sup>+</sup> T cells CD8<sup>+</sup> T cells, and innate-like T cells/innate lymphocytes (*NKG7* and *TRAV1-2*). Also, the myeloid cluster consisted of several subsets that were identified as macrophages (*CD68*), neutrophils (*FCGR3B*), and dendritic cells (DCs) (*HLA-DQA1*) (Fig. 1, D and E, and fig. S1C). All subsets identified were present in each sample of the two patient groups (fig. S1, D and E). We also combined RNA-, TCR-, and CITE-seq of peripheral blood T cells from patients with COVID-19 and bacterial pneumonia and yielded a sufficient number of CD4<sup>+</sup>, CD8<sup>+</sup>, and innate-like T cells for further clonal and expression analyses, allowing for a detailed comparison of peripheral and lung-specific T cell responses (Fig. 1, F and G, and fig. S2).

Next, we quantified the respective lymphoid and myeloid cell subsets in BALF and PBMC of patients with COVID-19 and bacterial pneumonia using flow cytometry data (Fig. 1H and fig. S3). Whereas in peripheral blood, there was not any obvious difference in the frequencies of cells analyzed comparing patients with COVID-19 and bacterial pneumonia, T cell subsets and natural killer (NK) cells showed significantly increased frequencies in the BALF of patients with COVID-19 compared with bacterial pneumonia (Fig. 1I). We confirmed the presence of T cells and mononuclear cells in the lung parenchyma of patients with COVID-19 by hematoxylin and eosin staining (Fig. 1J) and by CD3 staining (Fig. 1K) on lung autopsy tissue. An accumulation of T cells was predominantly identified in the perivascular space (fig. S4). Together, these data show an accumulation of T cells in the lungs of patients with COVID-19.

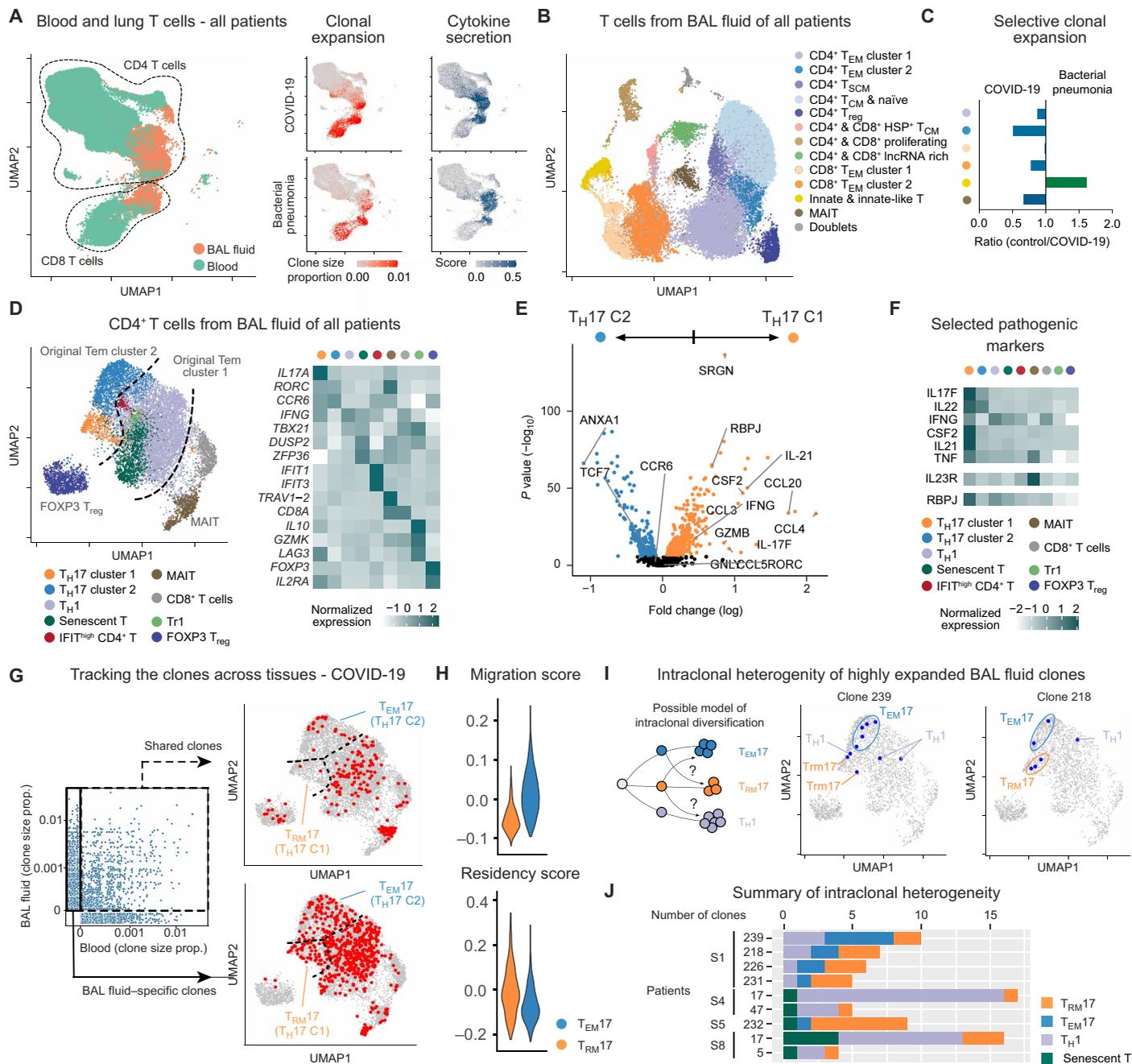
### Tracking T cell clonality in blood and BAL identifies tissue-specific T<sub>RM17</sub> cells

Considering the observation of the perivascular accumulation of T cells in the lungs of patients with COVID-19, we decided to further investigate the tissue-specific T cell response. We therefore integrated the blood and BALF T cell datasets to examine tissue-specific clonal expansion and activation of T cells. We examined T cell activation by analyzing the concomitant expression of a chosen pool of proinflammatory cytokines, i.e., *IL2*, *TNF*, *IL17A*, *IL17F*, *IFNG*, and *IL22*. Clonal expansion was analyzed by quantifying similar T cell clones on the basis of the TCR sequence information. We observed that both CD4<sup>+</sup> and CD8<sup>+</sup> T cells mainly underwent clonal expansion and activation predominantly in the BALF in both groups of patients (Fig. 2A). Clonal expansion does not necessarily reflect Ag specificity; therefore, we next investigated whether SARS-CoV-2-specific T cell clones are present in the BALF by comparing the TCRs identified in our study with those of two publicly available datasets of SARS-CoV-2-specific TCR sequences obtained from peripheral blood (25, 26). The frequency of shared clones was higher in the



**Fig. 1. Immune landscape of severe COVID-19 and bacterial pneumonia.** (A) Schematic representation of experimental setup. (B) Overview of baseline characteristics and clinical course of patients with COVID-19 and patients with bacterial pneumonia (LOD, limit of detection; ICU, intensive care unit). (C) Virus titers measured by quantitative polymerase chain reaction from BALF, tracheal fluid, and peripheral blood at time of sampling. (D) UMAP dimensionality reduction embedding of all cells from BALF ( $n = 56,735$  cells,  $n = 8$  for COVID-19, and  $n = 4$  for bacterial pneumonia, samples of patients S6 and B1 were excluded for technical reasons) colored according to cell type assessed by gene expression and (E) epitope measurement using CITE-seq of key markers (scale bars indicate normalized expression). (F) Single-cell analysis of CD3<sup>+</sup> T cells from peripheral blood of all patients ( $n = 77,457$  cells,  $n = 7$  for COVID-19, and  $n = 4$  for bacterial pneumonia). (G) CITE-seq information of cluster-defining epitopes (scale bars indicate normalized expression). (H) Flow cytometry of peripheral blood and BALF of patients with COVID-19 ( $n = 8$ ) and bacterial pneumonia ( $n = 5$ ). Per patient, an equal number of viable CD45<sup>+</sup> cells were exported for analysis and concatenated together before calculating the UMAP (total cells in peripheral blood = 129,141; in BALF = 114,927). Cell types were defined according to cell surface expression profiles by manual gating. Patient S9 was excluded from the statistical analysis due to low cell numbers. (I) Comparison of cell frequencies as measured by flow cytometry of cells from patients with COVID-19 and bacterial pneumonia ( $*P < 0.05$ ). (J) Hematoxylin and eosin staining (H&E) and (K) CD3 staining of lung autopsy tissue of one representative of seven patients.

Downloaded from <https://www.science.org> at Staats und Universitätsbibliothek Hamburg on September 05, 2021



**Fig. 2. T cell clonality in pulmonary inflammation.** (A) Blood-lung activation map of T cells from blood and BALF of all patients: UMAP dimensionality reduction embedding of T cells (left); clone size proportion (clone count divided by number of cells per sample) of T cells (middle), and the cytokine secretion score of T cells (right) from COVID-19 and bacterial pneumonia as indicated. (B) UMAP presentation of T cells from BALF of all patients. Clusters were annotated according to gene expression and epitopes measurement of key markers. T<sub>CM</sub>, T central memory; T<sub>SCM</sub>, T stem cell-like memory; lncRNA, long noncoding RNA. (C) Ratio of clonal expansion of bacterial pneumonia versus COVID-19 for the major expanded BALF T cell clusters. (D) Subclustering analysis of clonally expanded CD4<sup>+</sup> T cells of all patients. Clusters were annotated according to gene expression presented in the heatmap. (E) Volcano plot showing differential gene expression between T<sub>H</sub>17 clusters 1 and 2 of all patients. Genes were considered significant with adjust  $P < 0.05$ . Nonsignificant genes are shown in black. (F) Heatmap of selected pathogenic gene markers of T<sub>H</sub>17 cells of all patients in comparison with other T cell clusters. (G) Clone size proportion of T cells in peripheral blood and BALF of patients with COVID-19 and presentation of high abundant clones (clone size > 5) that are shared between BALF and blood and BAL-specific clones as indicated. (H) CD4 migration and tissue residency score of T<sub>H</sub>17 cluster1 (T<sub>RM</sub>17) and 2 (T<sub>EM</sub>17) from all patients. (I) Possible model of intracлонаl diversification of CD4<sup>+</sup> T cell subsets (left); distribution of two representative BALF clones from a patient with COVID-19 (patient S1 clone239 and clone218) on the UMAP (middle and right). (J) Bar plot of top expanded BALF clones containing T<sub>RM</sub>17 cells from patients with COVID-19. COVID-19:  $n = 8$  for BALF and  $n = 7$  for blood; bacterial pneumonia:  $n = 4$  for BALF and  $n = 4$  for blood.

COVID-19 cohort compared with the patients with bacterial pneumonia (fig. S6A). In addition, we observed higher frequencies of shared clones in T cells from the BALF as compared with peripheral blood.

Considering this tissue-specific activation, we examined BALF T cells with more granularity. By combining CITE-seq and differentially expressed genes (DEG), we identified five major CD4<sup>+</sup> T cell clusters, including a population of Foxp3<sup>+</sup> regulatory T (T<sub>reg</sub>) cells and two CD8<sup>+</sup> T cell clusters (Fig. 2B and fig. S5). We also found three clusters composed of both CD4<sup>+</sup> and CD8<sup>+</sup> T cells and characterized by heat shock proteins, genes associated with proliferation (*MKI67* and *STMN1*) and expression of long noncoding RNAs (*MALAT1* and *NEAT1*). Last, we identified a distinct cluster formed by MAIT cells (TCRVa7.2/*TRA V1-2*) and one by innate and innate-like T cells (CD56 and *NKG7*).

We then quantified the clonal expansion for each population (fig. S6, B and C). We observed high clonal expansion in two CD4<sup>+</sup> T cell clusters [i.e., CD4<sup>+</sup> T effector memory (T<sub>EM</sub>) cells, clusters 1 and 2], in the two main CD8<sup>+</sup> T cell populations, in MAIT cells, and in the other innate-like T cell cluster. We next wondered whether these cells were also expanded in a different type of infection. By comparing the clone size proportion of the above indicated populations between the two patient groups (Fig. 2C), we found that the CD4<sup>+</sup> T<sub>EM</sub> cell cluster 2 was most selectively expanded in patients with COVID-19 compared with patients with bacterial pneumonia. Therefore, we decided to further analyze this cluster and, as controls, we used the other CD4<sup>+</sup> T<sub>EM</sub> cell cluster (i.e., cluster 1), MAIT, and Foxp3<sup>+</sup> T<sub>reg</sub> cells (Fig. 2D and fig. S7). We observed that the original cluster 2 was enriched for genes typical of T<sub>H</sub>17 polarization states, whereas the original cluster 1 primarily contained CD4<sup>+</sup> T<sub>EM</sub> cells expressing genes associated with a T<sub>H</sub>1 polarization state. We then tested which of these clusters were selectively expanded in patients with COVID-19 and found that although T<sub>H</sub>1 cells are expanded in all patients, both T<sub>H</sub>17 clusters were only expanded in patients with COVID-19 (fig. S7, E and F).

Differential expression analysis revealed that although both T<sub>H</sub>17 clusters express similar levels of *RORC* and *CCR6*, T<sub>H</sub>17 cell cluster 1 is enriched for genes associated with cytotoxicity (*SRGN*, *GZMB*, and *GZMY*) and for genes translating for proinflammatory cytokines (*IL21*, *IL17F*, *IFNG*, and *CSF2/GM-CSF*) and chemokines (*CCL3*, *CCL4*, and *CCL5*) (Fig. 2E and fig. S7G). We also observed that this cluster has high expression of the transcriptional factor *RBPJ*, which has been shown to be fundamental for the pathogenicity of T<sub>H</sub>17 cells in an experimental autoimmune encephalomyelitis mouse model (20). We next compared the expression of some of these DEG, in addition to other genes associated with T<sub>H</sub>17 cell pathogenicity, among all the subclusters of the CD4<sup>+</sup> T cells isolated from the BALF. We found that among all, *CSF2* (GM-CSF) and *IL21* were the most selective genes expressed by the cluster of the potentially pathogenic T<sub>H</sub>17 cells (Fig. 2F).

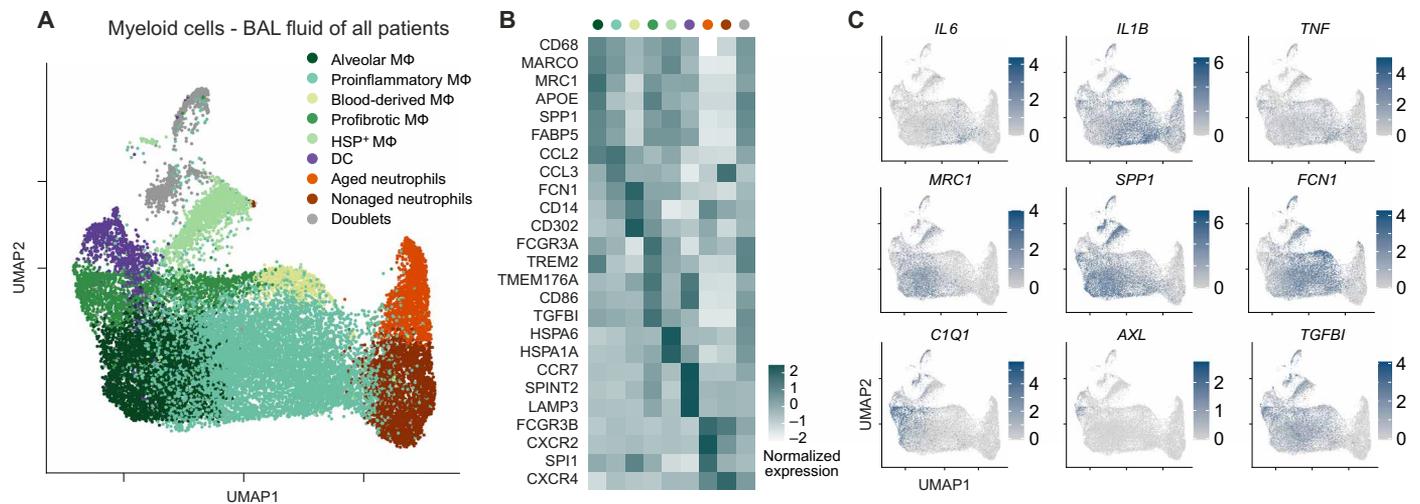
We have recently shown that T<sub>H</sub>17 cells can acquire a tissue-resident phenotype in the lungs (15). We therefore used TCR sequences as markers to test whether this population of T<sub>H</sub>17 cells is mainly found in the BALF and not in the circulation, suggesting possible resident behavior. We found that, virtually, none of the highly expanded T<sub>H</sub>17 cluster 1 cells in the lungs shared clones with T cells in the blood, supporting the notion that these cells are resident in the lung. In contrast, the other T<sub>H</sub>17 cell cluster (i.e., cluster 2) and the T<sub>H</sub>1 cell populations are composed of a mixture of resident and circulatory clones (Fig. 2G). Then, we used a literature-based residency and migratory

scores and found that the potentially pathogenic T<sub>H</sub>17 cell cluster 1 has on average a lower migratory and higher residency score compared with the other effector T<sub>H</sub>17 cell cluster 2 (Fig. 2H). These data suggest that the T<sub>H</sub>17 cell cluster 1 is enriched for pathogenic and resident cells compared with the T<sub>H</sub>17 cell cluster 2. To simplify the classification of these clusters and reflect their features, we named cluster 1 as tissue-resident memory-like T<sub>H</sub>17 cells (T<sub>RM</sub>17) and cluster 2 as effector memory T<sub>H</sub>17 cells (T<sub>EM</sub>17). We observed that the TCR sequences shown to be specific for SARS-CoV-2 in other studies (21, 22) can also be identified in the cluster of T<sub>RM</sub>17 cells (fig. S7H). Because it has been shown that T<sub>RM</sub>17 cells still retain a certain degree of plasticity, especially toward T<sub>H</sub>1 cells (23), we wondered whether this was the case in COVID-19. We used the TCR sequences as natural lineage barcodes to follow the origin/fate of some of the T<sub>RM</sub>17 cell clones expanded in patients with COVID-19. We observed that sister clones of the T<sub>RM</sub>17 cells were also found to express other T cell phenotypes, such as the T<sub>H</sub>1 phenotype (Fig. 2, I and J). This intraclonal diversification (24) suggests that some of the T<sub>RM</sub>17 cells have a dynamic developmental trajectory in common with other types of tissue-specific CD4<sup>+</sup> T cell populations, in particular with T<sub>H</sub>1 cells that, as expected, are the dominant expanded clones in term of quantity. In summary, we identified lung-specific T<sub>RM</sub>17 cells in the BALF of patients with COVID-19 on the basis of cytokine expression profiles and clonal expansion.

### Different types of myeloid cells identified in the BALF of COVID-19

Next, we set out to examine the different populations of myeloid cells in BALF that were identified in our scRNA-seq analysis (Fig. 1D). As above, to achieve robust clustering, we included myeloid cells from all patients (with COVID-19 and bacterial pneumonia) in this analysis. Subclustering of macrophages and neutrophils revealed the heterogeneity of macrophage polarization status and stages in neutrophil maturation (Fig. 3A). In particular, alveolar macrophages were defined by the gene expression of class A scavenger receptor *MARCO*, the mannose receptor *MRC1*, and the intracellular lipid transporter *FABP5*. They also express high levels of the profibrotic gene *SPP1* (Fig. 3, B and C) (12). We also detected high levels of *TREM2*, a surface receptor able to prevent macrophage apoptosis upon viral replication (25). Next, we defined proinflammatory macrophages as cells expressing high levels of *CCL2* and *CCL3*, chemokines involved in recruitment of adaptive and innate cells to sites of infection, and which were also characterized by the expression *IL6*, *IL1B*, and *TNF* (Fig. 3, A and C). Blood-derived macrophages have been defined on the basis of high expression of *CD14*. Their proinflammatory signature is mirrored by the high expression of *FCN1*, as previously described in patients with COVID-19 (12), by the high expression of *CD302*, a C-type lectin receptor induced in vitro upon lipopolysaccharide stimulation, and by the expression of alarmins such as *S100A8* and *S100A12*, calcium-binding proteins, and danger-associated molecular patterns (DAMPs), whose expression is regulated by proinflammatory molecules such as IFN- $\gamma$  and TNF- $\alpha$ , and that can lead to the secretion of IL-6 and IL-8 (Fig. 3B and fig. S8A).

We defined a population of cells with a tissue-remodeling signature as profibrotic macrophages, which, under a persistent inflammatory trigger, might acquire a profibrotic function (26). These profibrotic macrophages expressed higher levels of *APOE*, *TGFBI*, *TMEM176A*, and *CD86* and were enriched in complement components (*C1QB*, *C1QA*, and *C1QC*) (Fig. 3, B and C, and fig. S8A), in line with what



**Fig. 3. Landscape of myeloid cells in the lung.** (A) UMAP dimensionality reduction embedding of myeloid cells from BALF of all patients from our study (COVID-19  $n = 8$  and bacterial pneumonia  $n = 4$ ). (B) Heatmap of key marker gene expression of the indicated clusters. (C) UMAP plots showing expression of genes mirroring key features of macrophage polarization and function (scale bars indicate normalized expression).

was previously described for profibrotic macrophages in the context of SARS-CoV-2 infection (12). This macrophage cluster showed a profibrotic signature mostly similar to the alveolar macrophage subcluster, potentially indicating that the two macrophage populations have a similar biological function. However, the profibrotic macrophages were also characterized by the expression of high levels of *FCGR3A* (CD16) and intermediate/low levels of *CD14*, together with the expression of genes associated with antigen presentation (fig. S8A), therefore suggesting them as a subcluster of cells potentially originating from a nonclassical/intermediate monocyte population. Profibrotic macrophages also expressed *AXL*, a receptor tyrosine kinase that is required for resolution of lung inflammatory disease upon viral infection, induced by GM-CSF, and associated with development of tissue fibrosis in mouse models (27, 28). A population highly enriched in heat shock protein (*HSPA6*, *HSPA1A*, *HSPH1*, *DNAJB1*, and *HSPA1B*) was also observed and named as heat shock protein positive (HSP<sup>+</sup>) macrophages.

Last, expression of *CXCR2* and *CXCR4*, which define neutrophil maturation stages and regulate their trafficking from bone marrow, was detected in the two neutrophil clusters identified in patients with COVID-19 (aged and nonaged neutrophils) (Fig. 3, A and B). All subsets described were reproducibly found in both patients with COVID-19 and bacterial pneumonia (fig. S8, B and C). The identification of myeloid cell populations in our dataset provided us with a foundation for investigating the interactions between immune cells in the lungs of patients with COVID-19.

### T<sub>RM17</sub> cell interactome with other pathological cell types in COVID-19

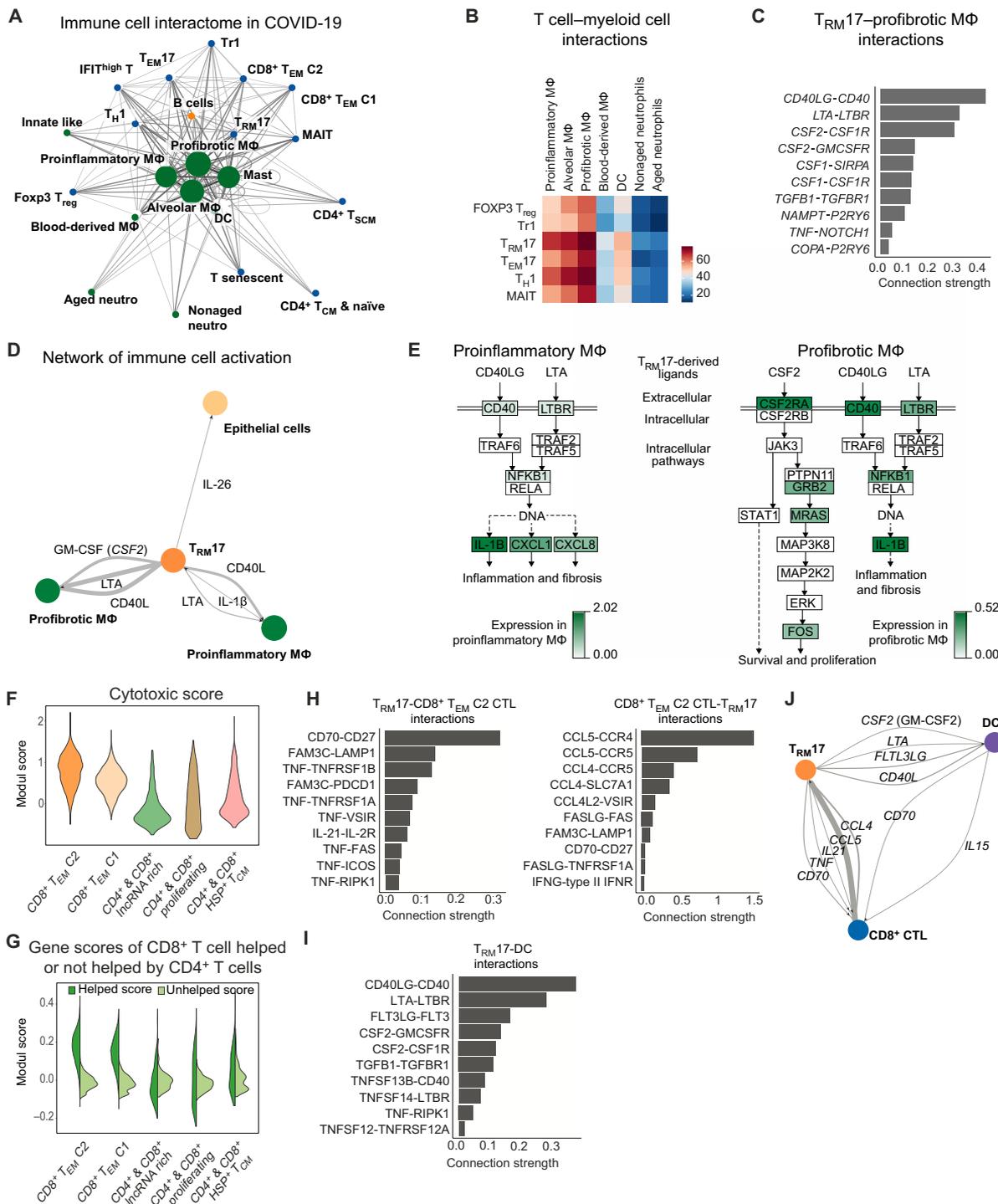
Once the landscape of the myeloid and lymphoid compartment was clarified, we investigated the cell-cell interactions of T<sub>RM17</sub> cells with the other immune cells, in particular myeloid and cytotoxic CD8<sup>+</sup> T cells, which are known to correlate with lung damage in patients with COVID-19 (12, 13).

We constructed the interactome of all immune cells found in the BALF of patients with COVID-19 using the T cell and the myeloid cell subclusters defined in Figs. 2 and 3, respectively. Then, we performed

a network analysis that is based on transcriptomic levels of ligand-receptor interactions between cell types. We identified different clusters of macrophages (proinflammatory, profibrotic, and alveolar macrophages) at the center of our communication network, having the highest number of different ligand-receptor pairs with other cell populations (Fig. 4A). We also observed that the cells of the main T cell populations (T<sub>H1</sub>, T<sub>RM17</sub>, T<sub>EM17</sub>, FOXP3 T<sub>reg</sub>, Tr1, and MAIT) interact more with tissue macrophages than with other myeloid cell clusters, such as blood-derived macrophages, DCs, or neutrophils (Fig. 4B). Among the T cells, T<sub>RM17</sub>, T<sub>EM17</sub>, and T<sub>H1</sub> appear to have more ligand/receptor interactions with macrophages than the other T cell clusters do (Fig. 4, A and B). To further explore the interactions between lung T<sub>RM17</sub> that we previously found to be clonally expanded and macrophages, we selected the 10 most specific ligand-receptor pairs of both populations on the basis of rank calculated using CellPhoneDB (29). To calculate the connection strength of these interactions, we multiplied the average ligand expression with the proportion of cells expressing the receptor from the respective clusters. We found that CD40LG/CD40, LTA/LTBR (lymphotoxin-alpha/lymphotoxin-beta-receptor), and GM-CSF/GM-CSFR had the highest connection strength between T<sub>RM17</sub> and profibrotic macrophages (Fig. 4C). In addition, we identified CD40LG/CD40, LTA-LTBR, and CSF2-CSF1R to demonstrate the most selective and strongest interactions of T<sub>RM17</sub> with proinflammatory macrophages (fig. S9A).

To understand how T<sub>RM17</sub> cells directly interact with the lung epithelium, we analyzed the interactions of T<sub>RM17</sub> cells with epithelial cells. Here, we identified the cytokines IL-26, IL-17A, and IL-17F to be among the top 10 interactions (fig. S9B). IL-26 interacts with the IL-26R on epithelial cells (fig. S9C), potentially playing a role in antiviral response (30). Because of the low number of epithelial cells in our dataset, we confirmed the expression of IL-26R, a heterodimer consisting of IL-10RB and IL-20RA, using the transcriptomic data on epithelial cells recently published in the context of SARS-CoV-2 infection (13) (fig. S9D).

Then, focusing on the most relevant interactions on the basis of rank and connection strength between T<sub>RM17</sub> cells, profibrotic macrophages, proinflammatory macrophages, and epithelial cells,



**Fig. 4. Interactome of T cells and myeloid cells in the lungs of patients with COVID-19.** (A) Interaction network of all BALF clusters based on the number of ligand-receptor interaction (>30 edges) based on Fruchterman-Reingold force-directed algorithm from patients with COVID-19 ( $n = 8$ ). (B) Adjacency map of T cell-myeloid cell interactions. (C) Ligand and receptor interaction strength ([mean ligand expression] × [proportion of receptor expression per cluster]) of T<sub>RM17</sub> cells (ligands) and profibrotic macrophages (receptors). Interactions were filtered for cytokines and, for specificity, based on rank scoring. (D) Supervised interaction map of potential key players in sustaining lung inflammation in patients with COVID-19. Line width correlates with interaction strength. (E) Pathway analysis of CD40L (CD40LG), LTA, and GM-CSF (CSF2) signaling in proinflammatory and profibrotic macrophages indicating the log<sub>2</sub> fold change in COVID-19 versus bacterial pneumonia. (F) Cytotoxic module scores in all clusters which include CD8<sup>+</sup> cells using proinflammatory and cytotoxic mediator genes in CTL from (13). (G) Module scores in the indicated clusters using the highest 50 differential expressed genes of CD8<sup>+</sup> T cells receiving help or no help from CD4<sup>+</sup> cells, respectively, according to (37). (H and I) Ligands and receptor interaction strength (H) between T<sub>RM17</sub> and CD8<sup>+</sup> T<sub>EM</sub> CTL Cluster 2 and (I) between T<sub>RM17</sub> and DC. Interactions were filtered according to their rank score. (J) Supervised interaction map of T<sub>RM17</sub>, CD8<sup>+</sup> T<sub>EM</sub> CTL, and DCs with annotated ligands. Line width correlates with interaction strength.

we constructed a smaller curated interaction map that depicts that  $T_{RM17}$  can act on both profibrotic and proinflammatory macrophages as well as epithelial cells (Fig. 4D). In return, proinflammatory macrophages, by secreting IL-1 $\beta$ , may act on  $T_{RM17}$  cells (Fig. 4D and fig. S9E) and additionally through secretion of various chemokines like CCL2, CCL3, and CCL20 target the respective chemokine receptors on  $T_{RM17}$  cells (fig. S9E).

Next, we aimed to gain additional insight into intracellular signaling induced by CD40L, LTA, and GM-CSF in proinflammatory and profibrotic macrophages. For this, we used Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and annotated the DEG to the respective pathway by color-coding genes up-regulated in macrophages from patients with COVID-19 versus bacterial pneumonia. In proinflammatory macrophages, IL-1 $\beta$ , CXCL1, and CXCL8 might be produced because of signals transmitted by CD40, LTBR, and NFKB1. A similar signaling cascade could be induced in profibrotic macrophages. Furthermore, GM-CSF was associated with a pathway capable of triggering survival and proliferation signals in profibrotic macrophages (Fig. 4E).

Last, because it is known that  $CD4^+$  T cells are necessary for regulating the magnitude and quality of the cytotoxic  $CD8^+$  T cell response, we investigated the molecular mechanisms by which  $T_{RM17}$  might regulate the  $CD8^+$  T cell cytotoxic response in patients with COVID-19. The cytotoxic  $CD8^+$  T cell response has been proposed to mediate lung tissue damage in these patients (12, 13). To identify highly cytotoxic  $CD8^+$  T cell clusters among the ones found (Fig. 2B), we created a cytotoxic scoring using proinflammatory genes that are expressed by cytotoxic T lymphocyte (CTL) in critical patients with COVID-19 as described by Chua *et al.* (13) and applied this dataset to the  $CD8^+$  clusters that we identified in our analysis. The highest cytotoxic signature was observed in  $CD8^+$   $T_{EM}$  cluster 2 and  $CD8^+$   $T_{EM}$  cluster 1 (Fig. 4F). We then calculated a second scoring to investigate whether  $CD8^+$  T cell clusters from the BALF of patients with COVID-19 might receive help from  $CD4^+$  T cells. To this end, we used the top 50 DEG identified by Ahrends *et al.* (31) as characteristic of  $T_{EM}$  cells receiving help from  $CD4^+$  T cells or not and applied this information to the different  $CD8^+$  T cell clusters from our dataset. We identified  $CD8^+$   $T_{EM}$  cluster 2 to display the highest help module score (Fig. 4G). Next, to gain insight on how  $T_{RM17}$  cells and  $CD8^+$   $T_{EM}$  affect each other, we determined the most specific and strongest interactions according to rank and connection strength. We identified CD70-CD27 and CCL5-CCR4 as the pathways highly engaged in this cell-to-cell interaction (Fig. 4H). Because  $CD4^+$ - $CD8^+$  T cell interaction occurs in a spatiotemporally organized interaction with DCs (32), we further dissected ligand-receptors interaction between  $T_{RM17}$  cells with DCs and CTL  $CD8^+$  T cells with DCs (Fig. 4I and fig. S9F).  $T_{RM17}$  cells had the potential to affect DCs via CD40L, FTL3LG, and GM-CSF (Fig. 4I). On the basis of all these data, we explored how the three cell populations could be connected and observed potential connections among  $T_{RM17}$ ,  $CD8^+$  CTL, and DCs (Fig. 4J).

In short, these data show the potential interaction of  $T_{RM17}$  cells and other tissue-specific immune cells, namely, macrophages and CTL  $CD8^+$ , which have been associated with disease severity of COVID-19.

### The cellular map of GM-CSF-expressing cells

To test whether GM-CSF and IL-17A correlate with the severity of COVID-19, we first measured these two cytokines in the serum

of patients with COVID-19 and of healthy blood donors and observed increased GM-CSF concentration in the patients (Fig. 5A). We excluded patient S9 from the Hamburg cohort because this patient received intravenous cytokine treatment. Second, we analyzed a different cohort, obtained from the University of Halle, which included patients with moderate and severe COVID-19 (33). We observed that GM-CSF and IL-17A appear to differentiate moderate versus severe disease (Fig. 5A).

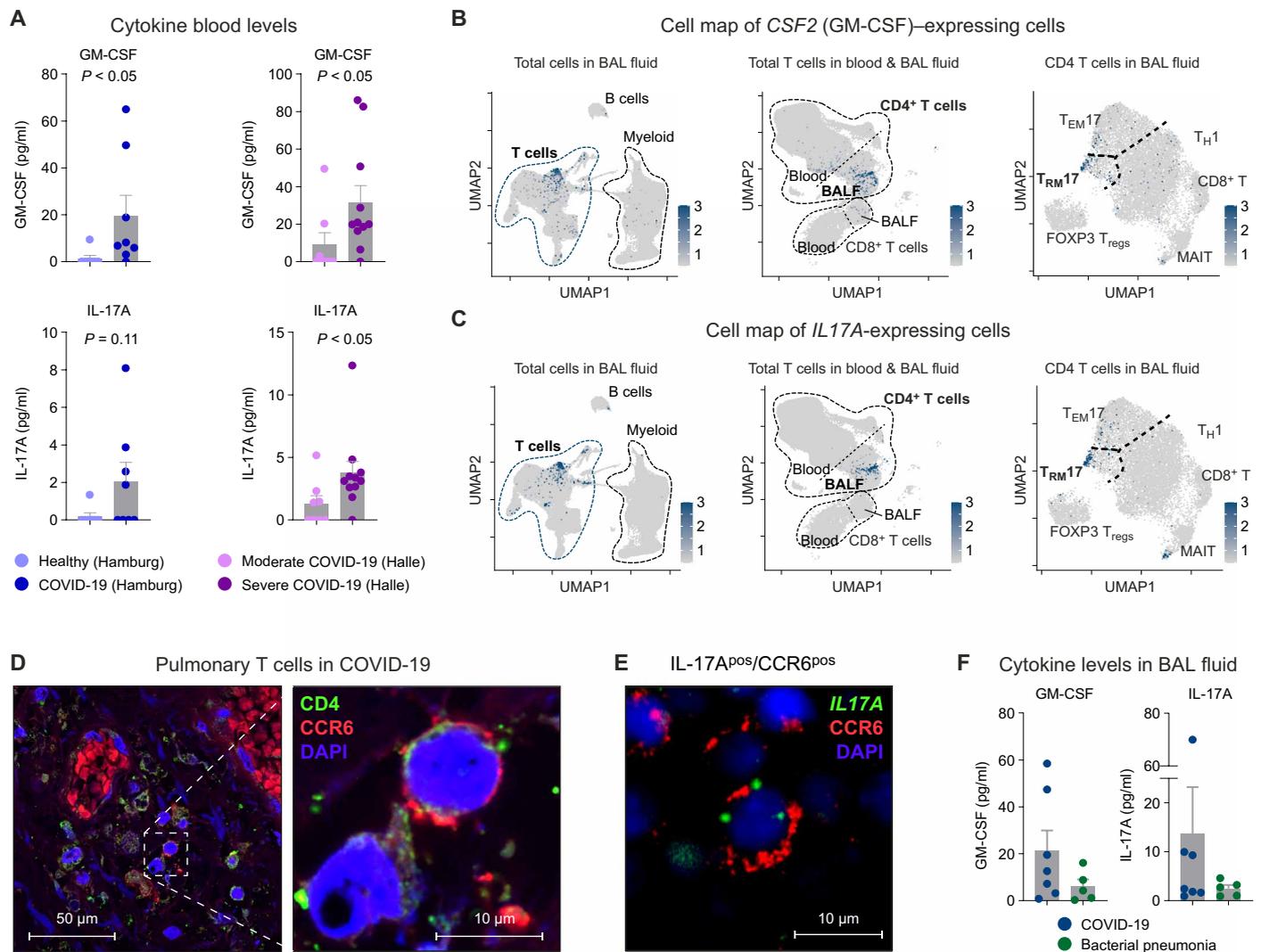
To examine potential cellular sources of GM-CSF and IL-17A, we mapped *CSF2* (GM-CSF)- and *IL17A*-expressing cells on three UMAPs with different granularity and tissues (Fig. 5, B and C): all cells in the BALF (left), all T cells in BALF and peripheral blood (center), and  $CD4^+$  T cells in the BALF (right) and displayed gene expression in a violin/dot plot (overview, lung and blood; fig. S10A). These data show that gene expression of both cytokines is mainly restricted to the  $T_{RM17}$  cell subset from BALF.

Last, to further support the presence of  $T_{RM17}$  cells in the lungs of patients with COVID-19, we first performed immunofluorescence staining showing the presence of  $CD4^+$  cells expressing CCR6 in the perivascular infiltrate of the lungs (Fig. 5D and fig. S10B). Second, in a proof-of-principle analysis, the combination of immunofluorescence and fluorescence in situ hybridization (FISH) of CCR6 and *IL17A*, respectively, shows the presence of CCR6 and *IL17A* coexpressing cells in COVID-19 (Fig. 5E and fig. S10C). Third, we found high concentrations of GM-CSF and IL-17A as well as IFN- $\gamma$  and IL-6 in the BALF of patients with COVID-19 (Fig. 5F and fig. S10D). These findings show that  $T_{RM17}$  cells are one potential source of the cytokines GM-CSF and IL-17A, which are prototypical of the hyperinflammation and are present locally within the lungs and in circulation in severe COVID-19.

### DISCUSSION

Here, we report a comprehensive single-cell transcriptional and TCR landscape of  $CD4^+$  T cells collected from the BALF and the peripheral blood of patients with severe COVID-19. We observed clonal expansion and characterized the activation profile of tissue-resident memory-like  $CD4^+$  T cells in the lungs of patients with COVID-19 that persist even after clearance of the virus. These cells express high amounts of the genes encoding the proinflammatory cytokines IL-17A/F and GM-CSF. Cell-cell interactome analysis uncovered a pathogenic network in the lung, involving GM-CSF-expressing  $T_{RM17}$  cells, IL-1 $\beta$ -expressing macrophages with a proinflammatory phenotype, macrophages expressing the GM-CSF receptor and genes associated with fibrosis, and cytotoxic  $CD8^+$  T cells. The relevance of our findings is further supported by the fact that serum protein levels of GM-CSF and IL-17A were elevated in a cohort of patients with severe COVID-19.

It has been speculated that reduction of T cells observed in the peripheral blood of patients with COVID-19 might be due to the recruitment of T cells to inflamed tissues (34). In support of this hypothesis, our flow cytometry data revealed an increased frequency of T cells in the lungs of these patients compared with patients with bacterial pneumonia while confirming reduced peripheral lymphocyte numbers. Moreover, activation status and cytokine expression were higher in lung T cells compared with peripheral T cells. By analyzing TCR clonality, we found a robust expansion of  $CD8^+$  T cells in the lungs of patients with COVID-19, as demonstrated by previous studies (12). However, the comparison with patients with bacterial



**Fig. 5. Cytokine secretion profile and cellular source of GM-CSF.** (A) GM-CSF and IL-17A protein in serum of patients with COVID-19 ( $n = 8$ ) and healthy controls ( $n = 7$ ) from Hamburg and of patients with moderate ( $n = 8$ ) or severe COVID-19 ( $n = 11$ ) from Halle as indicated. Cell map of (B) CSF2 (GM-CSF) expressing and (C) IL17A expressing cells (scale bars indicate normalized expression). Three different UMAPs with different cellular granularity showing the respective gene expression of in total cells of the BALF (left), total T cells of blood and BALF (middle), and total T cells in BALF (right) from all patients. (D) Immunofluorescence of CD4<sup>+</sup> (green) CCR6<sup>+</sup> (red) T<sub>RM</sub>17 cells in the lungs of a deceased patient with COVID-19 infection [nuclear staining 4',6-diamidino-2-phenylindol (DAPI), blue] (two additional samples are presented in fig. S10B). (E) Combined immunofluorescence (CCR6) and FISH (IL17A) of lung samples from one patient with COVID-19. (F) Concentrations of the indicated cytokines in the BALF of patients with COVID-19 and bacterial pneumonia.

pneumonia showed that this clonal expansion of CD8<sup>+</sup> T cell subsets was a general hallmark of mild to severe lung inflammation. In contrast to this, CD4<sup>+</sup> T cells and, in particular, those displaying a T<sub>H</sub>17 polarization state mainly expanded in the BALF of patients with COVID-19. One of these T<sub>H</sub>17 clusters expressed high levels of cytokines that have previously been associated with pathogenic activation of the immune system, such as GM-CSF (35, 36) and IL-17A (37), as well as other known markers of T<sub>H</sub>17 cell pathogenicity, such as the transcriptional factor RBPJ (20).

Moreover, using TCR clonality analysis across compartments, we showed that clonally expanded cells in this T<sub>H</sub>17 cell cluster were almost exclusively present in BALF samples, but not in the peripheral blood. These data were supported by an enrichment of genes usually expressed by resident T cells and together provided evidence

that these cells represent tissue-resident lung T<sub>H</sub>17 cells that are probably responsive to SARS-CoV-2-related antigens. In a previous study, we identified and characterized T<sub>RM</sub>17 cells in the lungs of mice and showed that they play a critical role in protecting from experimental *Klebsiella pneumoniae* infection (15). More recently, we found a correlation of kidney T<sub>RM</sub>17 cells and severity of immune-mediated kidney disease. Then, using mouse models, we demonstrated that T<sub>RM</sub>17 cells persist in the tissue after bacterial infection and can rapidly respond to inflammatory stimuli, such as IL-1 $\beta$  by producing IL-17A, which ultimately aggravates immune-mediated tissue injury (16). These two studies suggested that T<sub>RM</sub>17 cells can orchestrate a protective function against extracellular pathogens or fungi, but they can also participate in tissue damage if overactivated by inflammatory stimuli, such as IL-1 $\beta$ . In particular, because the

$T_H17/IL-17$  axis has not been linked to protective antiviral immunity, we propose that overproduction of IL-17A and GM-CSF by overactivated  $T_{RM17}$  cells is a feature of severe COVID-19 that might be involved in the immunopathology. However, our data do not rule out the possibility that  $T_{RM17}$  cells could provide a certain degree of protection at an early phase of the infection or in asymptomatic patients. Preclinical animal models of SARS-CoV-2 infection allowing the activity of  $T_{RM17}$ -derived cytokines to be blocked are needed to fully address their role at different time points after the infection.

To provide a detailed view of how the different immune cell populations interact in COVID-19, we performed unbiased interactome analysis. Here,  $T_{RM17}$  cells were among the T cell subsets showing the strongest interaction with different myeloid cell subsets and  $CD8^+$  T cells. By ranking the interactions of  $T_{RM17}$  cells with myeloid cell subsets according to specificity and interaction strength, we identified GM-CSF, CD40L, and lymphotoxin- $\alpha$  as the most important effector pathways used by these cells to induce proinflammatory cytokine and chemokine production such as IL-1 $\beta$ , CXCL1, and CXCL8 in macrophages. Release of IL-1 $\beta$  by proinflammatory macrophages, in turn, could signal back to  $T_{RM17}$  cells to increase their pathogenicity (16, 38). One of the most evident features of  $T_{RM17}$  cells was the expression of GM-CSF (see Figs. 2F and 5B), and the interaction with its receptor was among the top hits in the unsupervised interactome analysis of  $T_{RM17}$  cells with myeloid cell subsets. T cell-derived GM-CSF can result in activation and differentiation of myeloid cells (39). GM-CSF has further been shown to promote inflammatory tissue damage in a mouse model of Kawasaki disease, which is characterized by hyperinflammation that may share some features with severe COVID-19 (40). Enhanced frequencies of GM-CSF/IFN- $\gamma$  coproducing T cells have been found in the blood of patients with COVID-19 and seemed to correlate with disease activity (41). Our data indicate that CSF2/GM-CSF-expressing cells are found in the lungs and coexpress *IL17A*. These data, in addition to the presence of  $CD4^+CCR6^+$  T cells in the lung tissue as well as GM-CSF and IL-17A in the BALF, provide clinical evidence that  $T_{RM17}$  cell-associated cytokines are present in patients with severe COVID-19.

The major conclusions of this study derive from the comparison between cells taken from the blood and BALF of the patients with COVID-19. Nevertheless, in comparing patients with COVID-19 with patients with bacterial pneumonia, we observed that  $CD4^+$  T cells and, in particular,  $T_{RM17}$  cells were more clonally expanded in the BALF of the virally infected group. This comparison, however, also poses a key limitation of our study because we were unable to conclude whether the clonal expansion of  $T_{RM17}$  cells is specific to patients with COVID-19 or a common feature of severe viral infection. To address this point, we would have benefitted from having the BALFs from other viral infections such as influenza, in which the type of immune cells engaged is overall similar to a SARS-CoV-2 infection. A secondary limitation is our use of the term  $T_{RM17}$  cell, which was used on the basis of the expression profile and the reduced shared clonality between lung and blood of this population. However, the conclusion that these cells reside in the lungs is not definitive, as determination of tissue residency in human tissues remains challenging. Another limitation is the limited sample size of our study, and therefore, the results of this study need to be further validated in a larger cohort of patients, in which  $T_{RM17}$  cell associations with diseases severity are also examined. Almost all

patients had severe COVID-19 according to the WHO classification in our study.

Last, on the basis of our data, we propose a model in which  $T_{RM17}$  cells are activated or reactivated as part of an ongoing cytokine storm, during which they can start producing proinflammatory cytokines such as GM-CSF. This could lead to further activation of macrophages and  $CD8^+$  T cells, which others have linked to the severity of the disease (12), and lastly mediate lethal lung damage (36, 39). Two small pilot studies have indicated that targeting GM-CSF in patients with severe COVID-19 lung diseases using anti-GM-CSF receptor monoclonal antibodies mavrilimumab or lenzilumab, respectively, may be a strategy for improving clinical outcomes (3, 4), although larger controlled clinical trials would be needed to determine efficacy and biological impact of such approaches. This network of tissue-resident cells may persist in the lungs even after the initiating event, e.g., viral infection, has been cleared, contributing to chronic lung pathology. In conclusion, our study provides a snapshot analysis of  $CD4^+$  T cells in the lungs of patients with severe COVID-19 and identifies  $T_{RM17}$  cells as one of the components of the lung-specific immune response. In addition, our data provide a rationale for investigating therapeutic approaches targeting  $T_{RM17}$  cells and the GM-CSF network in the search for urgently needed therapies for treating COVID-19 pneumonia.

## MATERIALS AND METHODS

### Study design

Patients with SARS-CoV-2 infection can develop a severe COVID-19 course with pulmonary involvement and high mortality. Because the adaptive immune system may play a major role in COVID-19 pathogenesis, we sought to investigate the immune response in the lungs of these patients by scRNA-seq with a focus on T cells and their cytokines. To this end, we planned simultaneous gene expression, TCR repertoire sequencing, and cell surface protein analyses. Because this is only possible from live cells, we obtained BALF from the lungs of patients with COVID-19 and from patients with bacterial pneumonia, which served as a control. We included nine patients with COVID-19 and five patients with bacterial pneumonia in the comprehensive scRNA-seq analysis at the University Medical Center Hamburg (tables S1 and S2). To compare blood cytokine levels in patients with moderate and severe COVID-19, we analyzed patients from the University of Halle, Germany (table S3).

### Cell isolation

Human BALF and peripheral blood for flow cytometry and scRNA-seq were both obtained from patients undergoing BAL. The indication and performance of bronchoscopy were in accordance with the current guideline recommendations (42). These studies were approved by the Ethik-Kommission der Ärztekammer Hamburg, local ethics committee of the chamber of physicians in Hamburg, and were conducted in accordance with the ethical principles stated by the Declaration of Helsinki. Informed consent was obtained from all participating patients or legal representatives. Single-cell suspensions were obtained from BALF by washing with phosphate-buffered saline followed by filtering through 100-, 70-, 40- (Greiner Bio-One, Kremsmünster, Austria), and 30- $\mu$ m cell strainers (Partec, Görlitz, Germany). Leukocytes from blood samples were separated from red blood cells using BD Vacutainer CPT tubes with an integrated FICOLL gradient (BD Biosciences, San Jose, CA, USA). Samples were filtered

through a 30- $\mu$ m filter (Partec, Görlitz, Germany) before antibody staining and flow cytometry.

To minimize unspecific antibody binding, cells were incubated with Human BD FC Block (BD Biosciences) for 10 min. Next, cells were surface stained with fluorochrome conjugated antibodies [CD45 (clone HI30), CD3 (OKT3), CD4 (RPA-T4), CD8 (RPA-T8), CD56 (MEM-188),  $\gamma\delta$ -TCR 8(B1), CD31 (WM59), CD326 (9C4), CD14 (HCD 14), CD7 (CD7-6B7), CD16 (PC3G8), CD19 (HIB19), and CD324 (DECM-1); BioLegend and BD Biosciences], barcode-labeled antibodies (BioLegend) for 15 min (see table S4 for a complete list of antibodies and barcodes). Subsequently, a fixable dead cell stain (LIVE/DEAD Fixable Near-IR Dead Cell Stain Kit; Life Technologies, Carlsbad, CA) to exclude dead cells from analysis was used according to the manufacturer's instructions. Cells were analyzed and sorted on a BD Biosciences FACS AriaFusion.

### Histology

For immunohistochemistry, human paraffin-embedded lung sections (2  $\mu$ m) from patients with SARS-CoV-2 infection were stained with an antibody directed against CD3 (polyclonal rabbit anti-human, ref. A0452, DAKO, Glostrup, Denmark). Immunofluorescence microscopy was performed in 1- $\mu$ m paraffin-embedded sections, after 15-min antigen retrieval with pH 9 antigen retrieval solution (Agilent, Santa Clara, CA, USA) and incubation with polyclonal primary goat anti-CD4 antibody (R&D Systems, Minneapolis, MN, USA, AF-379) and rabbit anti-CCR6 antibody (Abcam, Cambridge, UK, ab140768). Images were captured using a laser confocal microscope (LSM800, Zeiss, Jena, Germany).

For combined detection of *IL17A* mRNA and CCR6 FISH was performed on formalin-fixed paraffin-embedded human lung samples using RNAscope Technology as previously described (43) in accordance with the directions from Advanced Cell Diagnostics. The RNAscope Hs-IL17A-C3 probe from Advanced Cell Diagnostics (Advanced Cell Diagnostics, 310931-C3) was used as the target probe to detect IL-17a mRNA. Fluorescent labeling of the target probe was performed using OPAL 690 dye (dilution, 1:1000; Akoya Biosciences, FP1497001KT). Subsequent immunofluorescence labeling was performed with an antibody against CCR6 (dilution: 1:200; OriGene Technologies, TA316610) in the same sections after completing the FISH protocol. Epifluorescence imaging was performed using the THUNDER Imager 3D Live Cell and 3D Cell Culture (Leica Microsystems).

### Multiplex

We used a bead-based immunoassay technology (LEGENDplex, BioLegend) to quantify the concentration of cytokines in the serum and BALF for each sample. The premixed Human Anti-Virus Response Panel (catalog no. 740349) and the Human Essential Immune Response Panel (catalog no. 740929) were applied to analyze the relevant cytokines following the manufacturer's protocol. Values below the limit of detection were considered zero. Collection of the Halle cohort was performed under institutional review board approval numbers 2020-039 and 11/17. This cohort is partially published (33).

### Cell sorting, library preparation, and next-generation sequencing

To enrich for T cells from the BALF, we FACS-sorted T cells, alveolar macrophages, monocytes, CD45<sup>high</sup>CD3<sup>neg</sup> cells (including innate lymphoid cells), and CD45<sup>neg</sup> cells (lung cells) according to the

gating strategy presented in fig. S1A. From peripheral blood, we FACS-sorted CD3<sup>pos</sup> T cells. Subsequent scRNA-seq using the 10X Chromium Controller (10X Genomics, Pleasanton, CA, USA) was loaded with the following proportions: first lane: 100% BALF T cells; second lane: BALF 17% alveolar macrophages, 17% monocytes, 33% CD45<sup>neg</sup> cells, and 33% CD45<sup>high</sup>CD3<sup>neg</sup> cells; third lane: 100% blood T cells (cell numbers were based the FACS information).

Single-cell libraries were generated with the 10X Genomics Chromium Single Cell 5'v1.1 reagents kit according to the manufacturer's instructions. Fifty-nanometer cDNA was used for gene expression library construction. Quality control (QC) was performed with hsDNA Qubit (Thermo Fisher Scientific, Waltham, MA, USA) and BioAnalyzer (Agilent). The libraries were sequenced on an Illumina NovaSeq 6000 system (S4 flow cell) with 150 base pairs and paired-end configurations.

### Preprocessing of single-cell RNA-seq and CITE-seq data

The Cell Ranger software pipeline (v3.1.0, 10X Genomics) was used to demultiplex cellular barcodes and map reads to the human reference genome (refdata-cellranger-GRCh38-3.0.0) (command `cellranger count`). The CITE-seq antibody and barcode information was included in a feature reference csv file and passed to the `cellranger count` command. As the output, we obtained the feature-barcode matrix that contains gene expression counts alongside CITE-seq counts for each cell barcode. The feature-barcode matrices for all the sample were further processed by the R package Seurat (v3.1.4) (44). As a QC step, we first filtered out the cells in which less than 200 genes were detected in the BALF samples and less than 500 genes were detected in the blood samples. To remove potential doublets, we excluded cells with total number of detected genes more than 5000. After visual inspection of the distribution of cells by the percentage of mitochondrial genes expressed, we further removed low-quality cells with more than 5% mitochondrial genes of all detected genes. We used LogNormalize method in Seurat to normalize the scRNA-seq and CITE-seq counts for the cells passed the QC.

### Sample aggregation and integration

For the BALF cell analysis, we first aggregated the BALF CD3<sup>+</sup> sample and CD3<sup>-</sup> sample for each patient using the function `merge` in Seurat. (We excluded the CD3<sup>-</sup> samples of patients S2 and B3 and the CD45<sup>-</sup> sample of patient S9 due to low sequence quality. For patient S9, we merged the CD45<sup>+</sup> sample and EpCAM<sup>+</sup> sample). After we obtained the merged BALF Seurat object for each patient, to remove the batch effects across different patients, we applied the integration method implemented in Seurat (function `FindIntegrationAnchors` and `IntegrateData`, `dims = 1:30`). For the blood CD3<sup>+</sup> cell analysis, we directly applied the integration to the samples of all patients. For the combined analysis of BALF and blood T cells, we selected T cell clusters identified in the BALF samples (as described below) and aggregated with corresponding blood CD3<sup>+</sup> samples for each patient using the `merge` function. Integration was then applied to the merged objects for all the patients.

### Dimensionality reduction and clustering

For each integrated object, the integrated matrix was scaled by `ScaleData` function (default parameters) and highly variable genes were detected (function `FindVariableFeatures`, `selection.method = "vst"`, `nfeatures = 2000`). Principal components analysis was performed on the scaled data (function `RunPCA`, `npcs = 30`) to reduce dimensionality.

Thirty principal components were used to compute the  $k$ -nearest neighbor graph on the basis of the Euclidean distance (function FindNeighbors), which then generated cell clusters using function FindClusters. The resolution parameter of the FindClusters function for each dataset was also determined by exploration of top marker genes of each cluster. UMAP was used to visualize clustering results. The top DEG in each cluster was found using the FindAllMarkers function (min.pct = 0.25 and logfc.threshold = 0.25) that ran Wilcoxon rank sum tests. Seurat functions AverageExpression and DoHeatmap were used to visualize the expression of the top marker genes or CITE-seq protein expression in each cell cluster. The top marker genes as well as the CITE-seq expression patterns were then used to determine the cell type of each cluster. The differential expression between selected clusters was calculated by the FindMarkers function (min.pct = 0.1), which also ran Wilcoxon rank sum tests.

### BALF T cells, myeloid cells reintegration, and subclustering

For the separate analysis of BALF T cells and BALF myeloid cells, we selected the clusters identified in the total BALF samples and reintegrated them by patients. Reclustering was performed after integration as described above and a detailed cell-type annotation was obtained after exploring the top marker genes and the CITE-seq expression profiles of clusters. For the subclustering analysis of CD4<sup>+</sup> cells, reintegration by patients and reclustering were also performed respectively before the identification of the cell subtypes.

### Processing of TCR-seq data and integration

TCR-seq data for each sample were assembled by the Cell Ranger software (v3.1.0, 10X Genomics) with the command cellranger vdj using the reference genome (refdata-cellranger-vdj-GRCh38-alt-ensembl-3.1.0). For each sample, Cell Ranger generated an output file, filtered\_contig\_annotations.csv, containing TCR- $\alpha$  chain and TCR- $\beta$  chain CDR3 nucleotide sequences for single cells that were identified by barcodes. The R package scRepertoire (v1.2.1) (45) was used to further combine the contig\_annotation data of different samples to a single list object (function combineTCR). The combined TCR contig list file was then integrated with the corresponding Seurat object of the scRNA-seq data using the function combineExpression (cloneCall = "gene+nt"). Only the cells with both TCR and scRNA-seq data were kept for downstream clonotype analysis. The clonotype was defined according to the genes comprising the TCR and the nucleotide sequence of the CDR3 region. The frequency of the each clonotype in each patient was then calculated as clone count. To get a normalized clone count size for each clonotype, we also calculated the clone size proportion (clone count divided by number of cells per patient). The clone count and clone size proportion were added to metadata of the single-cell matrices.

### Calculation of gene signature scores

Signature scores of gene sets were calculated by Seurat function AddModuleScore (nbin = 24 and ctrl = 100). The cytokine secretion gene set includes major proinflammatory cytokines produced by T cells. The residency and migration gene sets were obtained from a core list of up-regulated and down-regulated genes by CD4 tissue-resident T cells (table S5) (16, 46).

### Cell-cell interaction analysis

We applied CellphoneDB's statistical analysis method (2.1.2) and receptor-ligand database (2.0.0) to calculate statistically enriched

cell-cell interactions (<https://github.com/Teichlab/cellphonedb>). We used the log-normalized RNA assay of our BALF dataset containing all samples and selected cells either from patients with COVID-19 or patients with bacterial pneumonia to gain the count matrix. Because we have different levels of subclustering, we annotated each cell according to its cluster in its deepest level and used it as metadata input (clustering level: all BALF samples > T cells > CD4<sup>+</sup> T cells; all BALF samples > myeloid cells). We ran CellphoneDB with the default parameters. In total, CellphoneDB returned 13,034 significant ( $P < 0.05$ ) interactions. The rank of a ligand-receptor pair was calculated by CellphoneDB dividing its total number of significant  $P$  values by the number of cluster-cluster comparisons. For downstream analysis, we excluded integrin ligand-receptor pairs and interactions being annotated as not secreted. Moreover, we excluded CCL20-CXCR3 interactions (Id\_cp\_interaction "CPI-SS0F8C664D9") because of a lack of evidence in the literature.

### Connection strength

The connection strength of a specific interaction between two clusters was calculated by multiplying the mean expression of the ligand in the ligand cluster by the proportion of cells expressing the receptor in the receptor cluster. The receptor was expressed if the log-normalized expression value was greater than 0. In case of a receptor complex, the receptor component expressed in the least cells was used.

### Pathway analysis

The differential expression of SARS-CoV-2-infected patients was calculated using patients with bacterial pneumonia as control group. With the differential expression data, enriched pathways were determined using Gene Ontology terms and KEGG pathways. This reveals commonly known pathways (e.g., Janus kinase-signal transducer and activator of transcription signaling and phosphatidylinositol 3-kinases pathway). From these results, relevant parts of the pathways were curated and combined to the final pathway, using the KEGG Markup Language schema ([www.kegg.jp/kegg/xml/docs/](http://www.kegg.jp/kegg/xml/docs/)). Coherent components of these enriched pathways were combined to a single representative pathway for each subset of macrophages. The log fold change of DEG was added as color code to the elements of the pathway.

### Network plots

Network plots were created using the R package igraph (1.2.5) (<https://github.com/igraph>). The layout of the network in Fig. 4A was calculated using the Fruchterman-Rheingold algorithm (function layout\_with\_fr, niter = 5000). The weight parameter was set to the number of interactions between two clusters. The clusters "CD8\_Tcells," "M1\_HSP," and "epithelial cells" were excluded. The vertex size was set by using graph strength, which sums up the edge weights for each vertex. The curated interaction map in Fig. 4D was made using the R package igraph (<https://github.com/igraph>) without applying the Fruchterman-Reingold algorithm. The thickness of the lines correlates with the respective connection strength ([average ligand expression]  $\times$  [proportion of cells expressing the receptor]).

### Statistics

Statistical analysis was performed using GraphPad Prism (La Jolla, CA). The results are shown as single data points with the means  $\pm$  SEM in a scatter dot plot. Differences between two individual groups were compared using a Mann-Whitney test. In the case of three or more groups, Wilcoxon test was used.

## SUPPLEMENTARY MATERIALS

immunology.sciencemag.org/cgi/content/full/6/56/eabf6692/DC1

Fig. S1. Sorting strategy and clustering information of BALF immune cells from all patients.

Fig. S2. Clustering information of blood T cells from all patients.

Fig. S3. Flow cytometry of cells from peripheral blood and BALF.

Fig. S4. Location of T cells in the lungs of COVID-19.

Fig. S5. Clustering information of BALF T cells from all patients.

Fig. S6. Clonal expansion analysis of BALF T cells from all patients.

Fig. S7. Subclustering of clonally expanded BALF CD4 T cells from all patients.

Fig. S8. Clustering information of BALF myeloid cells from all patients.

Fig. S9. Cell-cell interaction of T cells with myeloid cells and epithelial cells.

Fig. S10. Cytokine levels in BALF.

Table S1. Baseline characteristics and disease-related parameters of patients with COVID-19 and controls.

Table S2. Relevant medication of patients with COVID-19 and controls.

Table S3. Baseline characteristics of the Halle cohort.

Table S4. CITE-seq antibodies and barcodes (Excel file).

Table S5. Gene sets (Excel file).

Table S6. Raw data file (Excel file).

[View/request a protocol for this paper from Bio-protocol.](#)

## REFERENCES AND NOTES

- E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
- RECOVERY Collaborative Group, Dexamethasone in hospitalized patients with Covid-19. *N. Engl. J. Med.* **384**, 639–704 (2021).
- Z. Temesgen, M. Assi, F. N. U. Shweta, P. Vergidis, S. A. Rizza, P. R. Bauer, B. W. Pickering, R. R. Razonable, C. R. Libertin, C. D. Burger, R. Orenstein, H. E. Vargas, R. Palraj, A. S. Dababneh, G. Chappell, D. Chappell, O. Ahmed, R. Sakemura, C. Durrant, S. S. Kenderian, A. D. Badley, GM-CSF neutralization with lenzilumab in severe COVID-19 pneumonia: A case-control study. *Mayo Clin. Proc.* **95**, 2382–2394 (2020).
- G. De Luca, G. Cavalli, C. Campochiaro, E. Della-Torre, P. Angelillo, A. Tomelleri, N. Boffini, S. Tentori, F. Mette, N. Farina, P. Rovere-Querini, A. Ruggeri, T. D'Aliberti, P. Scarpellini, G. Landoni, F. De Cobelli, J. F. Paolini, A. Zangrillo, M. Tresoldi, B. C. Trapnell, F. Ciceri, L. Dagna, GM-CSF blockade with mavrilumab in severe COVID-19 pneumonia and systemic hyperinflammation: A single-centre, prospective cohort study. *Lancet Rheumatol* **2**, e465–e473 (2020).
- T. Huet, H. Beaussier, O. Voisin, S. Jouvesshomme, G. Dauriat, I. Lazareth, E. Sacco, J.-M. Naccache, Y. Bézie, S. Laplanche, A. L. Berre, J. L. Pavéc, S. Salmeron, J. Emmerich, J.-J. Mourad, G. Chatellier, G. Hayem, Anakinra for severe forms of COVID-19: A cohort study. *Lancet Rheumatol* **2**, e393–e400 (2020).
- C. Ucciferri, A. Auricchio, M. Di Nicola, N. Potere, A. Abbate, F. Cipollone, J. Vecchiet, K. Falasca, Canakinumab in a subgroup of patients with COVID-19. *Lancet Rheumatol* **2**, e457–e458 (2020).
- A. J. Stone, M. J. Frigault, N. J. Serling-Boyd, A. D. Fernandes, L. Harvey, A. S. Foulkes, N. K. Horick, B. C. Healy, R. Shah, A. M. Bensaci, A. E. Woolley, S. Nikiforow, N. Lin, M. Sagar, H. Schragar, D. S. Huckins, M. Axelrod, M. D. Pincus, J. Fleisher, C. A. Sacks, M. Dougan, C. M. North, Y.-D. Halvorsen, T. K. Thurber, Z. Dagher, A. Scherer, R. S. Wallwork, A. Y. Kim, S. Schoenfeld, P. Sen, T. G. Neilan, C. A. Peruginio, S. H. Unizony, D. S. Collier, M. A. Matza, J. M. Yin, K. A. Bowman, E. Meyerowitz, A. Zafar, Z. D. Drobni, M. B. Bolster, M. Kohler, K. M. D'Silva, J. Dau, M. M. Lockwood, C. Cubbison, B. N. Weber, M. K. Mansour; BACC Bay Tocilizumab Trial Investigators, Efficacy of tocilizumab in patients hospitalized with Covid-19. *N. Engl. J. Med.* **383**, 2333–2344 (2020).
- C. Salama, J. Han, L. Yau, W. G. Reiss, B. Kramer, J. D. Neidhart, G. J. Criner, E. Kaplan-Lewis, R. Baden, L. Pandit, M. L. Cameron, J. Garcia-Diaz, V. Chávez, M. Mekebeb-Reuter, F. L. de Menezes, R. Shah, M. F. González-Lara, B. Assman, J. Freedman, S. V. Mohan, Tocilizumab in patients hospitalized with Covid-19 pneumonia. *N. Engl. J. Med.* **384**, 20–30 (2021).
- A. J. Wilk, A. Rustagi, N. Q. Zhao, J. Roque, G. J. Martínez-Colón, J. L. McKechnie, G. T. Ivison, T. Ranganath, R. Vergara, T. Hollis, L. J. Simpson, P. Grant, A. Subramanian, A. J. Rogers, C. A. Blish, A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat. Med.* **26**, 1070–1076 (2020).
- W.-J. Guan, Z.-Y. Ni, Y. Hu, W.-H. Liang, C.-Q. Ou, J.-X. He, L. Liu, H. Shan, C.-L. Lei, D. S. C. Hui, B. Du, L.-J. Li, G. Zeng, K.-Y. Yuen, R.-C. Chen, C.-L. Tang, T. Wang, P.-Y. Chen, J. Xiang, S.-Y. Li, J.-L. Wang, Z.-J. Liang, Y.-X. Peng, L. Wei, Y. Liu, Y.-H. Hu, P. Peng, J.-M. Wang, J.-Y. Liu, Z. Chen, G. Li, Z.-J. Zheng, S.-Q. Qiu, J. Luo, C.-J. Ye, S.-Y. Zhu, N.-S. Zhong, W.-J. Guan, Z.-Y. Ni, Y. Hu, W.-H. Liang, C.-Q. Ou, J.-X. He, L. Liu, H. Shan, C.-L. Lei, D. S. C. Hui, B. Du, L.-J. Li, G. Zeng, K.-Y. Yuen, R.-C. Chen, C.-L. Tang, T. Wang, P.-Y. Chen, J. Xiang, S.-Y. Li, J.-L. Wang, Z.-J. Liang, Y.-X. Peng, L. Wei, Y. Liu, Y.-H. Hu, P. Peng, J.-M. Wang, J.-Y. Liu, Z. Chen, G. Li, Z.-J. Zheng, S.-Q. Qiu, J. Luo, C.-J. Ye, S.-Y. Zhu, N.-S. Zhong; China Medical Treatment Expert Group for Covid-19, Clinical characteristics of Coronavirus disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020).
- C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).
- M. Liao, Y. Liu, J. Yuan, Y. Wen, G. Xu, J. Zhao, L. Cheng, J. Li, X. Wang, F. Wang, L. Liu, I. Amit, S. Zhang, Z. Zhang, Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
- R. L. Chua, S. Lukassen, S. Trump, B. P. Hennig, D. Wendisch, F. Pott, O. Debnath, L. Thümann, F. Kurth, M. T. Völker, J. Kazmierski, B. Timmermann, S. Twardziok, S. Schneider, F. Machleidt, H. Müller-Redetzky, M. Maier, A. Krannich, S. Schmidt, F. Falzer, J. Liebig, J. Loske, N. Suttrop, J. Eils, N. Ishaque, U. G. Liebert, C. von Kalle, A. Hocke, M. Witzernath, C. Goffinet, C. Drosten, S. Laudi, I. Lehmann, C. Conrad, L.-E. Sander, R. Eils, COVID-19 severity correlates with airway epithelium-immune cell interactions identified by single-cell analysis. *Nat. Biotechnol.* **38**, 970–979 (2020).
- R. A. Grant, L. Morales-Nebreda, N. S. Markov, S. Swaminathan, M. Querrey, E. R. Guzman, D. A. Abbott, H. K. Donnelly, A. Donayre, I. A. Goldberg, Z. M. Klug, N. Borkowski, Z. Lu, H. Kihshen, Y. Politanska, L. Sichizya, M. Kang, A. Shilatfard, C. Qi, J. W. Lomasney, A. C. Argento, J. M. Kruser, E. S. Malsin, C. O. Pickens, S. B. Smith, J. M. Walter, A. E. Pawlowski, D. Schneider, P. Nannapaneni, H. Abdala-Valencia, A. Bharat, C. J. Gottardi, G. R. S. Budinger, A. V. Misharin, B. D. Singer, R. G. Wunderink; The NU SCRIPST Study Investigators, Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. *Nature* **590**, 635–641 (2021).
- M. C. Amezcua Vesely, P. Pallis, P. Bielecki, J. S. Low, J. Zhao, C. C. D. Harman, L. Kroehling, R. Jackson, W. Bailis, P. Licona-Limón, H. Xu, N. Iijima, P. S. Pillai, D. H. Kaplan, C. T. Weaver, Y. Kluger, M. S. Kowalczyk, A. Iwasaki, J. P. Pereira, E. Esplugues, N. Gagliani, R. A. Flavell, Effector T<sub>H</sub>17 cells give rise to long-lived T<sub>RM</sub> cells that are essential for an immediate response against bacterial infection. *Cell* **178**, 1176–1188.e15 (2019).
- C. F. Krebs, D. Reimers, Y. Zhao, H.-J. Paust, P. Bartsch, S. Nuñez, M. V. Roseblatt, M. Hellmig, C. Kilian, A. Borchers, L. U. B. Enk, M. Zinke, M. Becker, J. Schmid, S. Klinge, M. N. Wong, V. G. Puelles, C. Schmidt, T. Bertram, N. Stumpf, E. Hoxha, C. Meyer-Schwesinger, M. T. Lindenmeyer, C. D. Cohen, M. Rink, C. Kurts, S. Franzenburg, F. Koch-Nolte, J.-E. Turner, J.-H. Riedel, S. Huber, N. Gagliani, T. B. Huber, T. Wiech, H. Rohde, M. R. Bono, S. Bonn, U. Panzer, H.-W. Mittrücker, C. F. Krebs, D. Reimers, Y. Zhao, H.-J. Paust, P. Bartsch, S. Nuñez, M. V. Roseblatt, M. Hellmig, C. Kilian, A. Borchers, L. U. B. Enk, M. Zinke, M. Becker, J. Schmid, S. Klinge, M. N. Wong, V. G. Puelles, C. Schmidt, T. Bertram, N. Stumpf, E. Hoxha, C. Meyer-Schwesinger, M. T. Lindenmeyer, C. D. Cohen, M. Rink, C. Kurts, S. Franzenburg, F. Koch-Nolte, J.-E. Turner, J.-H. Riedel, S. Huber, N. Gagliani, T. B. Huber, T. Wiech, H. Rohde, M. R. Bono, S. Bonn, U. Panzer, H.-W. Mittrücker, Pathogen-induced tissue-resident memory T<sub>H</sub>17 (T<sub>RM</sub>17) cells amplify autoimmune kidney disease. *Sci. Immunol.* **5**, eaba4163 (2020).
- G. Dimopoulos, Q. de Mast, N. Markou, M. Theodorakopoulou, A. Komnos, M. Mouktaroudi, M. G. Netea, T. Spyridopoulos, R. J. Verheggen, J. Hoogerwerf, A. Lachana, F. L. van de Veerdonk, E. J. Giamarellos-Bourboulis, Favorable Anakinra responses in severe covid-19 patients with secondary hemophagocytic lymphohistiocytosis. *Cell Host Microbe* **28**, 117–123.e1 (2020).
- P. Mehta, J. C. Porter, J. J. Manson, J. D. Isaacs, J. P. M. Openshaw, I. B. McInnes, C. Summers, R. C. Chambers, Therapeutic blockade of granulocyte macrophage colony-stimulating factor in COVID-19-associated hyperinflammation: Challenges and opportunities. *Lancet Respir. Med.* **8**, 822–830 (2020).
- O. Pacha, M. A. Sallman, S. E. Evans, COVID-19: A case for inhibiting IL-17? *Nat. Rev. Immunol.* **20**, 345–346 (2020).
- G. M. Z. Horste, C. Wu, C. Wang, L. Cong, M. Pawlak, Y. Lee, W. Elyaman, S. Xiao, A. Regev, V. K. Kuchroo, RBPJ controls development of pathogenic Th17 cells by regulating IL-23 receptor expression. *Cell Rep.* **16**, 392–404 (2016).
- P. Bacher, E. Rosati, D. Esser, G. R. Martini, C. Saggau, E. Schiminsky, J. Dargviniene, I. Schröder, I. Wieters, Y. Khodamoradi, F. Eberhardt, M. J. G. T. Vehreschild, H. Neb, M. Sonntagbauer, C. Conrad, F. Tran, P. Rosenstiel, R. Markewitz, K.-P. Wandler, M. Augustin, J. Rybnikier, M. Kochanek, F. Leybold, O. A. Cornely, P. Koehler, A. Franke, A. Scheffold, Low-avidity CD4<sup>+</sup> T cell responses to SARS-CoV-2 in unexposed individuals and humans with severe COVID-19. *Immunity* **53**, 1258–1258.e5 (2020).
- B. J. Meckliff, C. Ramírez-Suastegui, V. Fajardo, S. J. Chee, A. Kusnadi, H. Simon, S. Eschweiler, A. Grifoni, E. Pelosi, D. Weiskopf, A. Sette, F. Ay, G. Seumois, C. H. Ottensmeier, P. Vijayanand, Imbalance of regulatory and cytotoxic SARS-CoV-2-reactive CD4<sup>+</sup> T cells in COVID-19. *Cell* **183**, 1340–1353.e16 (2020).
- M. DuPage, J. A. Bluestone, Harnessing the plasticity of CD4<sup>+</sup> T cells to treat immune-mediated disease. *Nat. Rev. Immunol.* **16**, 149–163 (2016).
- F. Sallusto, A. Cassotta, D. Hoces, M. Foglierini, A. Lanzavecchia, Do memory CD4 T cells keep their cell-type programming: Plasticity versus fate commitment? T-cell

- heterogeneity, plasticity, and selection in humans. *Cold Spring Harb. Perspect. Biol.* **10**, a029421 (2018).
25. K. Wu, D. E. Byers, X. Jin, E. Agapov, J. Alexander-Brett, A. C. Patel, M. Cella, S. Gilfilan, M. Colonna, D. L. Kober, T. J. Brett, M. J. Holtzman, TREM-2 promotes macrophage survival and lung disease after respiratory viral infection. *J. Exp. Med.* **212**, 681–697 (2015).
  26. T. A. Wynn, T. R. Ramalingam, Mechanisms of fibrosis: Therapeutic translation for fibrotic disease. *Nat. Med.* **18**, 1028–1040 (2012).
  27. N. Fujino, O. J. Brand, D. J. Morgan, T. Fujimori, A. M. Grabiec, C. P. Jagger, R. A. Maciewicz, M. Yamada, K. Itakura, H. Sugiura, M. Ichinose, T. Hussell, Sensing of apoptotic cells through Axl causes lung basal cell proliferation in inflammatory diseases. *J. Exp. Med.* **216**, 2184–2201 (2019).
  28. A. Zagorska, P. G. Traves, L. Jimenez-Garcia, J. D. Strickland, J. Oh, F. J. Tapia, R. Mayoral, P. Burrola, B. L. Cople, G. Lemke, Differential regulation of hepatic physiology and injury by the TAM receptors Axl and Mer. *Life Sci. Alliance* **3**, e202000694 (2020).
  29. M. Efremova, M. Vento-Tormo, S. A. Teichmann, R. Vento-Tormo, CellPhoneDB: Inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
  30. E. Stephen-Victor, H. Fickenscher, J. Bayry, IL-26: An emerging proinflammatory member of the IL-10 cytokine family with multifaceted actions in antiviral, antimicrobial, and autoimmune responses. *PLOS Pathog.* **12**, e1005624 (2016).
  31. T. Ahrends, J. Busselaar, T. M. Severson, N. Babala, E. de Vries, A. Bovens, L. Wessels, F. van Leeuwen, J. Borst, CD4<sup>+</sup> T cell help creates memory CD8<sup>+</sup> T cells with innate and help-independent recall capacities. *Nat. Commun.* **10**, 5531 (2019).
  32. J. L. Hor, P. G. Whitney, A. Zaid, A. G. Brooks, W. R. Heath, S. N. Mueller, Spatiotemporally distinct interactions with dendritic cell subsets facilitates CD4<sup>+</sup> and CD8<sup>+</sup> T cell activation to localized viral infection. *Immunity* **43**, 554–565 (2015).
  33. C. Schultheiß, L. Paschold, D. Simnica, M. Mohme, E. Willscher, L. von Wenserski, R. Scholz, I. Wieters, C. Dahlke, E. Tolosa, D. G. Sedding, S. Ciesek, M. Addo, M. Binder, Next-generation sequencing of T and B cell receptor repertoires from COVID-19 patients showed signatures associated with severity of disease. *Immunity* **53**, 442–455.e4 (2020).
  34. Z. Chen, E. John Wherry, T cell responses in patients with COVID-19. *Nat. Rev. Immunol.* **20**, 529–536 (2020).
  35. B. Becher, S. Tugues, M. Greter, GM-CSF: From growth factor to central mediator of tissue inflammation. *Immunity* **45**, 963–973 (2016).
  36. J. A. Hamilton, GM-CSF in inflammation. *J. Exp. Med.* **217**, e20190945 (2020).
  37. B. Stockinger, S. Omenetti, The dichotomous nature of T helper 17 cells. *Nat. Rev. Immunol.* **17**, 535–544 (2017).
  38. C. E. Zielinski, F. Mele, D. Aschenbrenner, D. Jarrossay, F. Ronchi, M. Gattorno, S. Monticelli, A. Lanzavecchia, F. Sallusto, Pathogen-induced human TH17 cells produce IFN- $\gamma$  or IL-10 and are regulated by IL-1 $\beta$ . *Nature* **484**, 514–518 (2012).
  39. F. M. Lang, K. M.-C. Lee, J. R. Teijaro, B. Becher, J. A. Hamilton, GM-CSF-based treatments in COVID-19: Reconciling opposing therapeutic approaches. *Nat. Rev. Immunol.* **20**, 507–514 (2020).
  40. A. T. Stock, J. A. Hansen, M. A. Sleeman, B. S. McKenzie, I. P. Wicks, GM-CSF primes cardiac inflammation in a mouse model of Kawasaki disease. *J. Exp. Med.* **213**, 1983–1998 (2016).
  41. Y. Zhou, B. Fu, X. Zheng, D. Wang, C. Zhao, Y. Qi, R. Sun, Z. Tian, X. Xu, H. Wei, Pathogenic T cells and inflammatory monocytes incite inflammatory storms in severe COVID-19 patients. *Natl. Sci. Rev.* **7**, 998–1002 (2020).
  42. S. Kluge, U. Janssens, T. Welte, S. Weber-Carstens, G. Marx, C. Karagiannidis, German recommendations for critically ill patients with COVID19. *Med. Klin. Intensivmed. Notfmed.* **115**, 111–114 (2020).
  43. F. Wang, J. Flanagan, N. Su, L. C. Wang, S. Bui, A. Nielson, X. Wu, H. T. Vo, X. J. Ma, Y. Luo, RNAscope: A novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagn.* **14**, 22–29 (2012).
  44. T. Stuart, A. Butler, R. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck III, Y. Hao, M. Stoekius, P. Smibert, R. Satija, Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).
  45. N. Borcherding, N. L. Bormann, G. Kraus, scRepertoire: An R-based toolkit for single-cell immune receptor analysis. *F1000Res* **9**, 47 (2020).
  46. B. V. Kumar, W. Ma, M. Miron, T. Granot, R. S. Guyer, D. J. Carpenter, T. Senda, X. Sun, S. H. Ho, H. Lerner, A. L. Friedman, Y. Shen, D. L. Farber, Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. *Cell Rep.* **20**, 2921–2934 (2017).

**Acknowledgments:** We sincerely thank healthy donors, patients, and their families for participating. FACS sorting was performed at the UKE FACS sorting core facility. We thank E. Hussey for editing the paper. **Funding:** This study was supported by grants from the Bundesministerium für Bildung und Forschung, BMBF (IT-COVID-19 to N.G. and C.F.K.); Deutsche Forschungsgemeinschaft (DFG) (SFB1192 to J.-E.T., V.G.P., S.H., T.B.H., S.B., U.P., N.G., and C.F.K.; SFB841 to S.H. and N.G.; KFO296 to N.G.; and SFB1286 to S.B.); Swedish Research Council VR grant 2017-02932 to N.G.; European Research Council (Diet-namic, 715271); Deutsche Zentrum für Infektionsforschung (DZIF) to N.G.; grants from the Deutsche Nierenstiftung and Deutsche Gesellschaft für Nephrologie to C.F.K.; a grant from the Werner Otto Stiftung to L.U.B.E.; and a grant from the Else Kröner-Fresenius-Stiftung (EKFS) (iPRIME). **Author contributions:** Study initiation: S.H., T.B.H., S.K., U.P., N.G., and C.F.K. Conceptualization: J.-E.T., P. Bacher, J.S.Z.W., M.M.A., A.W.L., M.B., S.H., T.B.H., S.K., S.B., U.P., N.G., and C.F.K. Methodology and analysis: Y.Z., C.K., L.B., K.R., P. Bartsch, A.-C.G., F.C., C.S., M. Hellmig, L.U.B.E., F.H., A.B., M.N.W., H.-J.P., F.S., N.S., M. Herrmann, E.R., D.K., D.J., M.L., S.P., S.S., J.-P.S., S.B., N.G., and C.F.K. Writing: J.-E.T., L.B., U.P., N.G., and C.F.K. Visualization: Y.Z., C.K., J.-E.T., L.B., V.G.P., S.S., S.B., N.G., and C.F.K. Supervision: U.P., N.G., and C.F.K. **Competing interests:** M.L. has the following financial relation relationships: Roche Diagnostics: member of an advisory board, received speaking fees and research funding; DiaSorin: received speaking fees; and bioMérieux: received speaking fees. **Data and materials availability:** The raw and processed sequencing data are deposited at Gene Expression Omnibus under accession number GSE1167118. All other data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

Submitted 13 November 2020

Accepted 18 February 2021

Published First Release 23 February 2021

Final published 19 March 2021

10.1126/sciimmunol.abf6692

**Citation:** Y. Zhao, C. Kilian, J.-E. Turner, L. Bosurgi, K. Roedel, P. Bartsch, A.-C. Gnirck, F. Cortesi, C. Schultheiß, M. Hellmig, L. U. B. Enk, F. Hausmann, A. Borchers, M. N. Wong, H.-J. Paust, F. Siracusa, N. Scheibel, M. Herrmann, E. Rosati, P. Bacher, D. Kyliès, D. Jarczak, M. Lütgehetmann, S. Pfefferle, S. Steurer, J. Schulze zur Wiesch, V. G. Puelles, J.-P. Spherhake, M. M. Addo, A. W. Lohse, M. Binder, S. Huber, T. B. Huber, S. Kluge, S. Bonn, U. Panzer, N. Gagliani, C. F. Krebs, Clonal expansion and activation of tissue-resident memory-like TH17 cells expressing GM-CSF in the lungs of patients with severe COVID-19. *Sci. Immunol.* **6**, eabf6692 (2021).

## Clonal expansion and activation of tissue-resident memory-like T17 cells expressing GM-CSF in the lungs of patients with severe COVID-19

Yu Zhao Christoph Kilian Jan-Eric Turner Lidia Bosurgi Kevin Roedl Patricia Bartsch Ann-Christin Gnirck Filippo Cortesi Christoph Schultheiß Malte Hellmig Leon U.B. Enk Fabian Hausmann Alina Borchers Milagros N. Wong Hans-Joachim Paust Francesco Siracusa Nicola Scheibel Marissa Herrmann Elisa Rosati Petra Bacher Dominik Kyllies Dominik Jarczak Marc Lütgehetmann Susanne Pfefferle Stefan Steurer Julian Schulze zur Wiesch Victor G. Puelles Jan-Peter Sperhake Marylyn M. Addo Ansgar W. Lohse Mascha Binder Samuel Huber Tobias B. Huber Stefan Kluge Stefan Bonn Ulf Panzer Nicola Gagliani Christian F. Krebs

*Sci. Immunol.*, 6 (56), eabf6692.

### TH17 cells in severe COVID-19

Generation of T helper 17 (T17) cells has been associated with immunopathogenesis in multiple autoimmune diseases. Using integrated single-cell transcriptome and TCR repertoire profiling, Zhao *et al.* showed that a population of T17 cells with features of tissue-resident memory T cells was clonally expanded in bronchoalveolar lavage (BAL) fluid collected from the lungs of patients with severe COVID-19, but not in samples from patients with bacterial pneumonia. Lung tissue-resident memory-like T17 cells were the primary immune cell type in BAL expressing the cytokine GM-CSF, which was also elevated in serum from a cohort of patients with severe COVID-19 compared with those with moderate disease. These results provide insight into specific T cell responses associated with severe COVID-19 pneumonia and identify a potential cellular target of GM-CSF-neutralizing therapies.

### View the article online

<https://www.science.org/doi/10.1126/sciimmunol.abf6692>

### Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

*Science Immunology* (ISSN 2470-9468) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Immunology* is a registered trademark of AAAS.

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).



## Supplementary Materials for

### Clonal expansion and activation of tissue-resident memory-like T<sub>H</sub>17 cells expressing GM-CSF in the lungs of patients with severe COVID-19

Yu Zhao, Christoph Kilian, Jan-Eric Turner, Lidia Bosurgi, Kevin Roedl, Patricia Bartsch, Ann-Christin Gnirck, Filippo Cortesi, Christoph Schultheiß, Malte Hellmig, Leon U.B. Enk, Fabian Hausmann, Alina Borchers, Milagros N. Wong, Hans-Joachim Paust, Francesco Siracusa, Nicola Scheibel, Marissa Herrmann, Elisa Rosati, Petra Bacher, Dominik Kyllies, Dominik Jarczak, Marc Lütgehetmann, Susanne Pfefferle, Stefan Steurer, Julian Schulze-zur-Wiesch, Victor G. Puelles, Jan-Peter Sperhake, Marylyn M. Addo, Ansgar W. Lohse, Mascha Binder, Samuel Huber, Tobias B. Huber, Stefan Kluge, Stefan Bonn, Ulf Panzer, Nicola Gagliani\*, Christian F. Krebs\*

\*Corresponding author. Email: n.gagliani@uke.de (N.G.); c.krebs@uke.de (C.F.K.)

Published 23 February 2021, *Sci. Immunol.* **6**, eabf6692 (2021)  
DOI: 10.1126/sciimmunol.abf6692

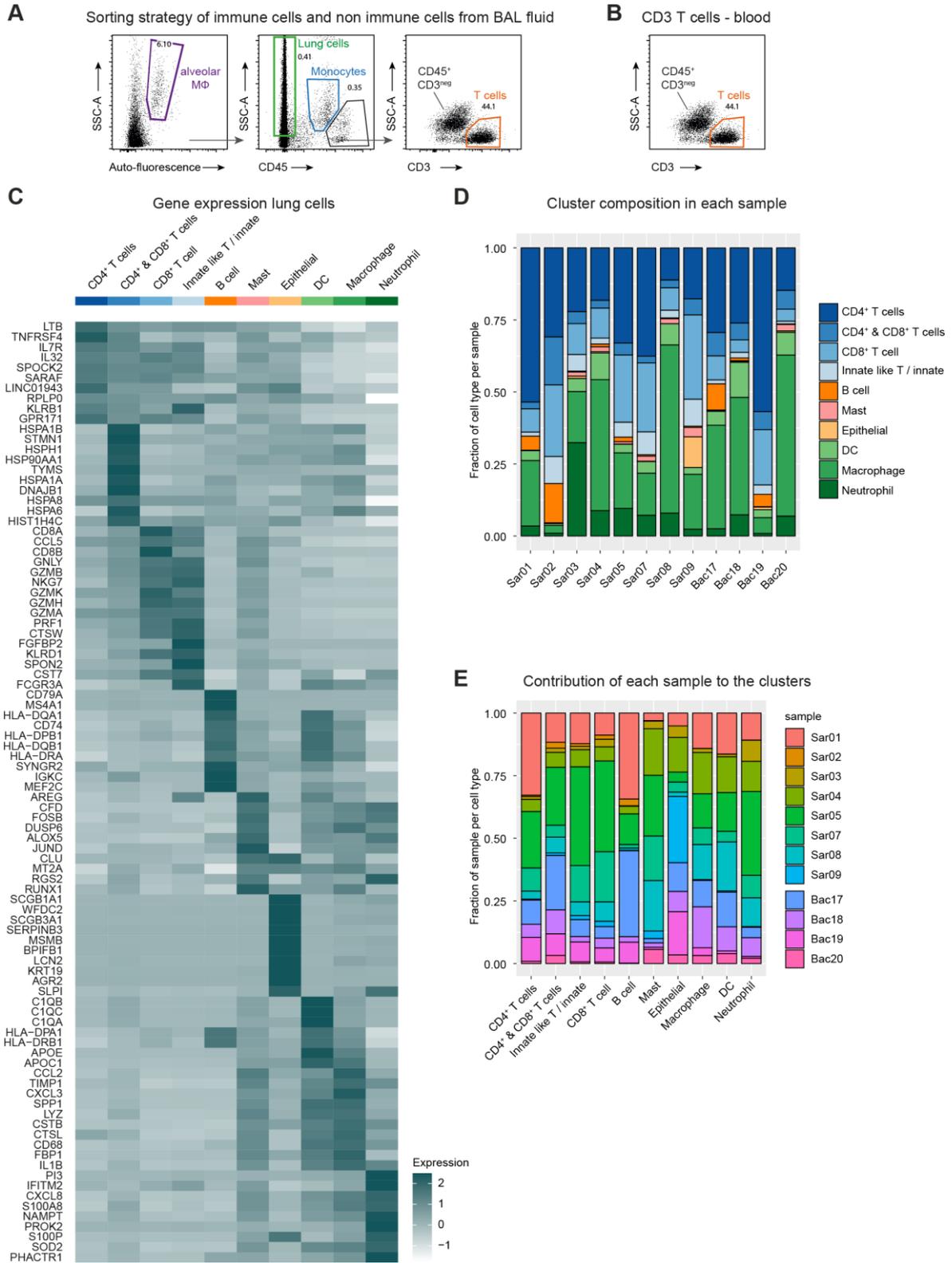
#### The PDF file includes:

- Fig. S1. Sorting strategy and clustering information of BALF immune cells from all patients.
- Fig. S2. Clustering information of blood T cells from all patients.
- Fig. S3. Flow cytometry of cells from peripheral blood and BALF.
- Fig. S4. Location of T cells in the lungs of COVID-19.
- Fig. S5. Clustering information of BALF T cells from all patients.
- Fig. S6. Clonal expansion analysis of BALF T cells from all patients.
- Fig. S7. Subclustering of clonally expanded BALF CD4 T cells from all patients.
- Fig. S8. Clustering information of BALF myeloid cells from all patients.
- Fig. S9. Cell-cell interaction of T cells with myeloid cells and epithelial cells.
- Fig. S10. Cytokine levels in BALF.
- Table S1. Baseline characteristics and disease-related parameters of patients with COVID-19 and controls.
- Table S2. Relevant medication of patients with COVID-19 and controls.
- Table S3. Baseline characteristics of the Halle cohort.

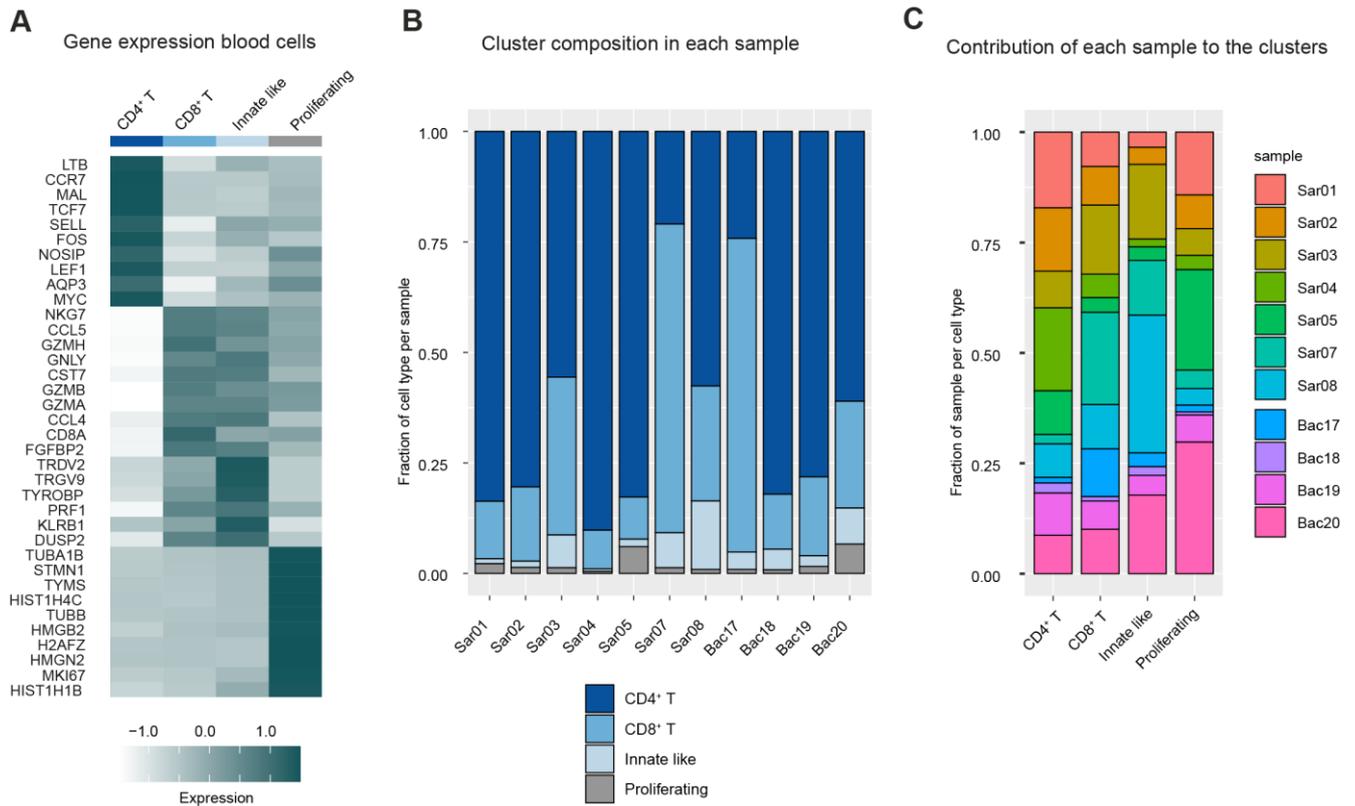
#### Other Supplementary Material for this manuscript includes the following:

(available at [immunology.sciencemag.org/cgi/content/full/6/57/eabf6692/DC1](https://immunology.sciencemag.org/cgi/content/full/6/57/eabf6692/DC1))

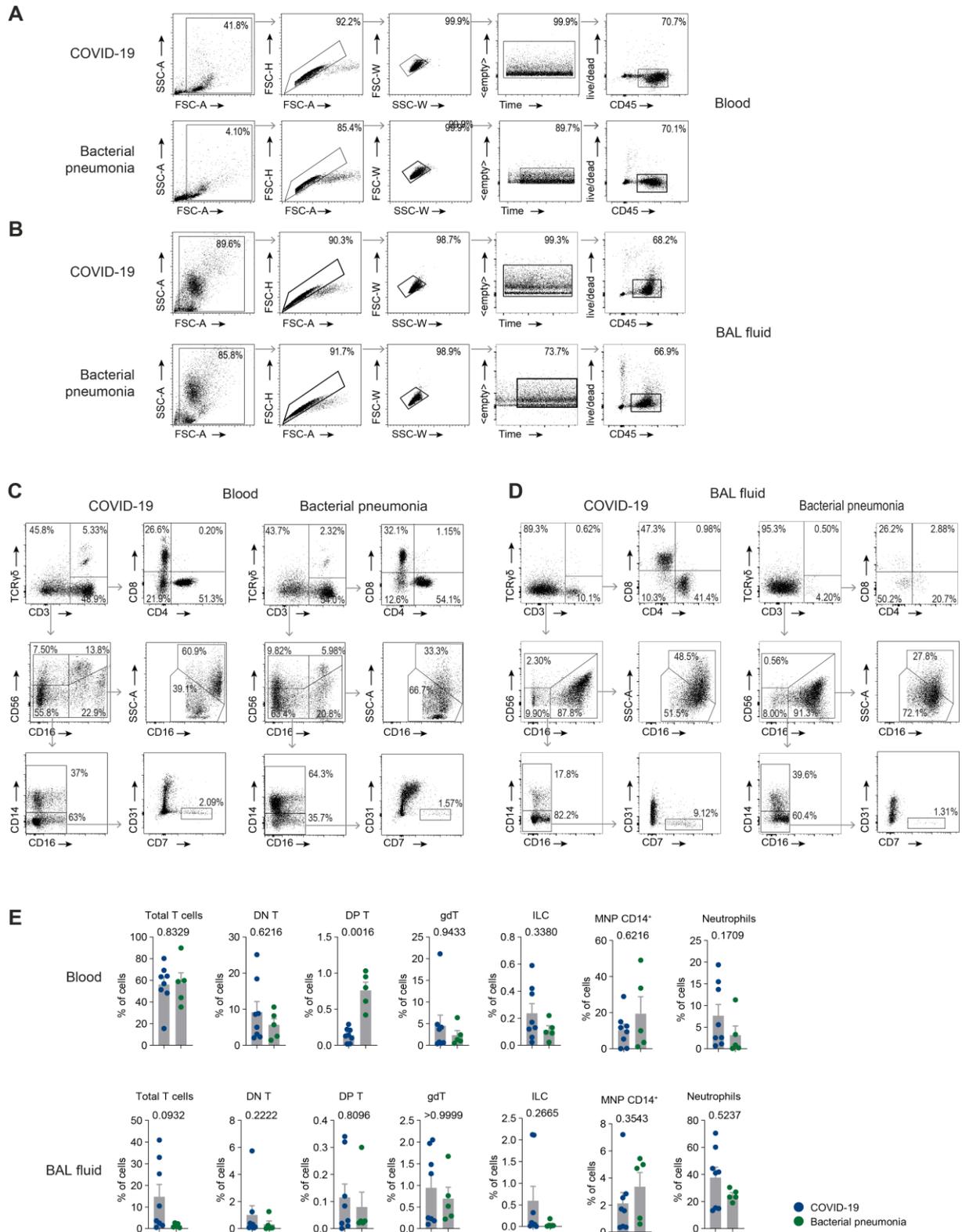
- Table S4. CITE-seq antibodies and barcodes (Excel file).
- Table S5. Gene sets (Excel file).
- Table S6. Raw data file (Excel file).



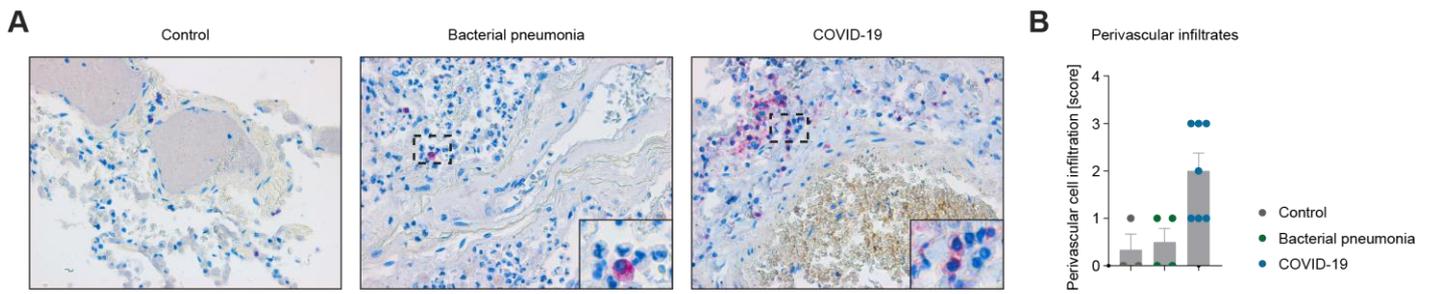
**Fig. S1. Sorting strategy and clustering information of BALF immune cells from all patients.** T cells and CD3 negative cells from BAL fluid (A) and T cells from the peripheral blood (B) were FACS sorted accordingly. (C) Heat map showing top 10 marker genes of each cell type of all lung cells. (D) Bar plot showing the composition of the clusters in each patient. (E) Bar plot showing the relative contribution to the clusters by each patient.



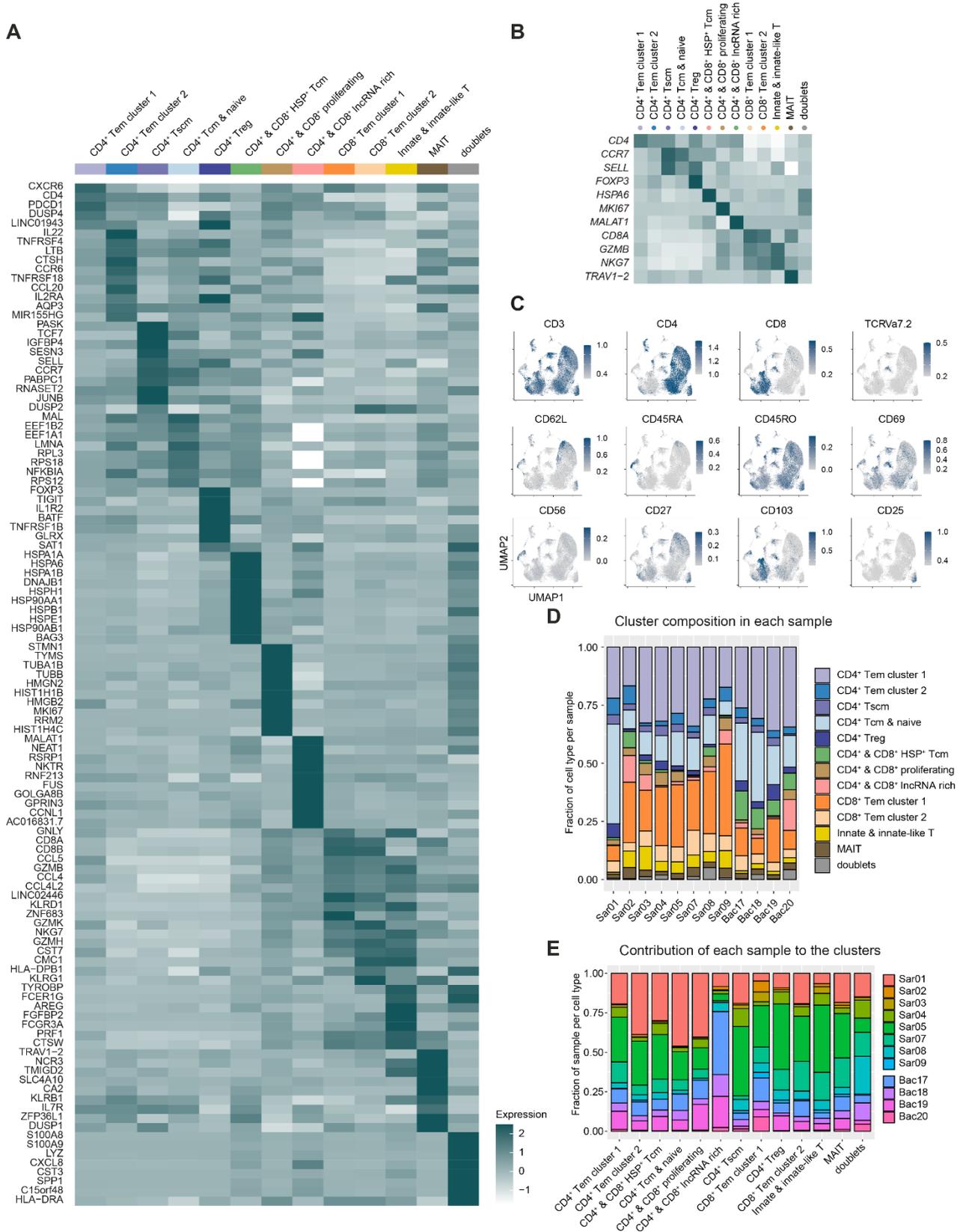
**Fig. S2. Clustering information of blood T cells from all patients. (A)** Heat map of top 10 marker genes of each cell type. **(B)** Bar plot showing the composition of the clusters in each patient. **(C)** Bar plot showing the relative contribution to the clusters by each patient.



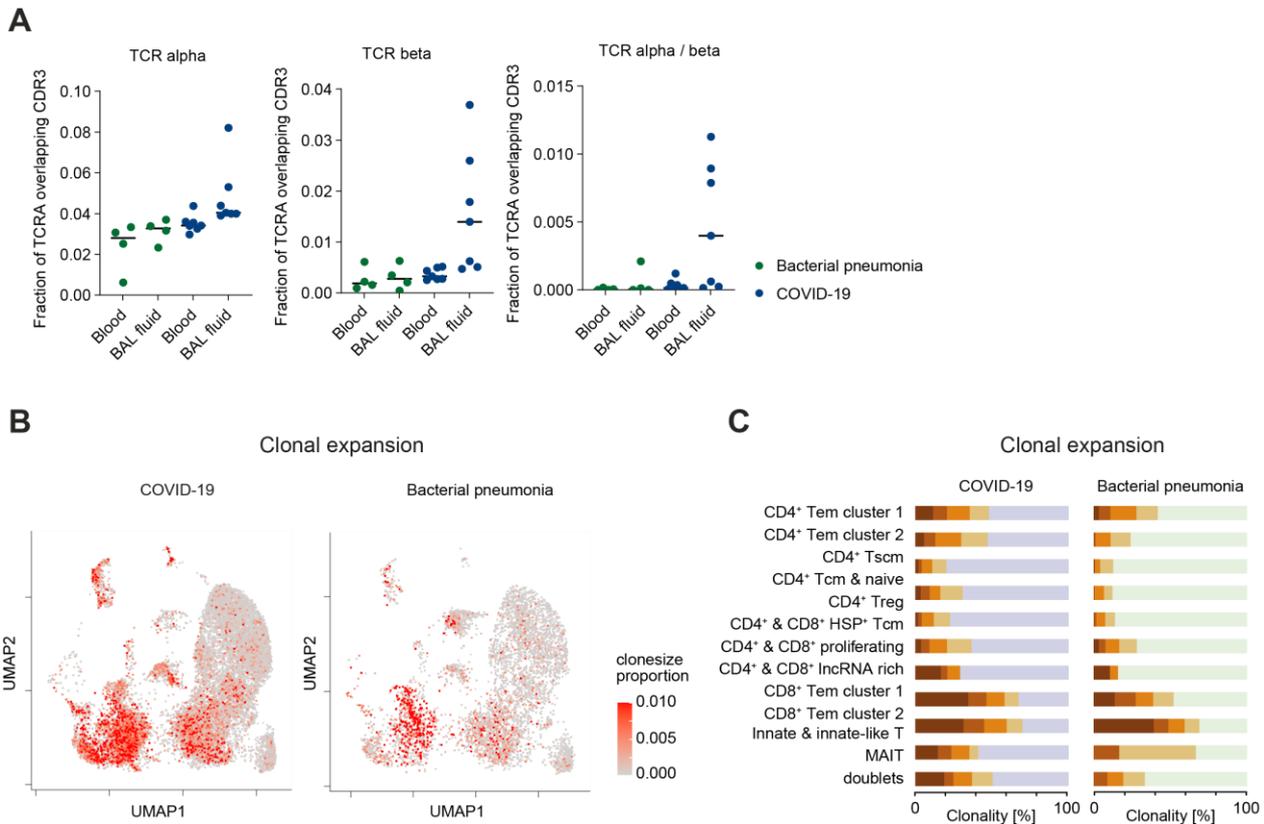
**Fig. S3. Flow cytometry of cells from peripheral blood and BALF.** Gating strategy of live CD45+ cells from the peripheral blood (A) and BAL fluid (B) of patients with bacterial pneumonia and COVID-19 as indicated. Gating strategy for cell type identification of peripheral blood cells (C) and cells from the BAL fluid (D) from patients with bacterial pneumonia and COVID-19 related to Figure 1H. (E) Comparison of cell frequencies as measured by flow cytometry of cells from COVID-19 and bacterial pneumonia patients as indicated.



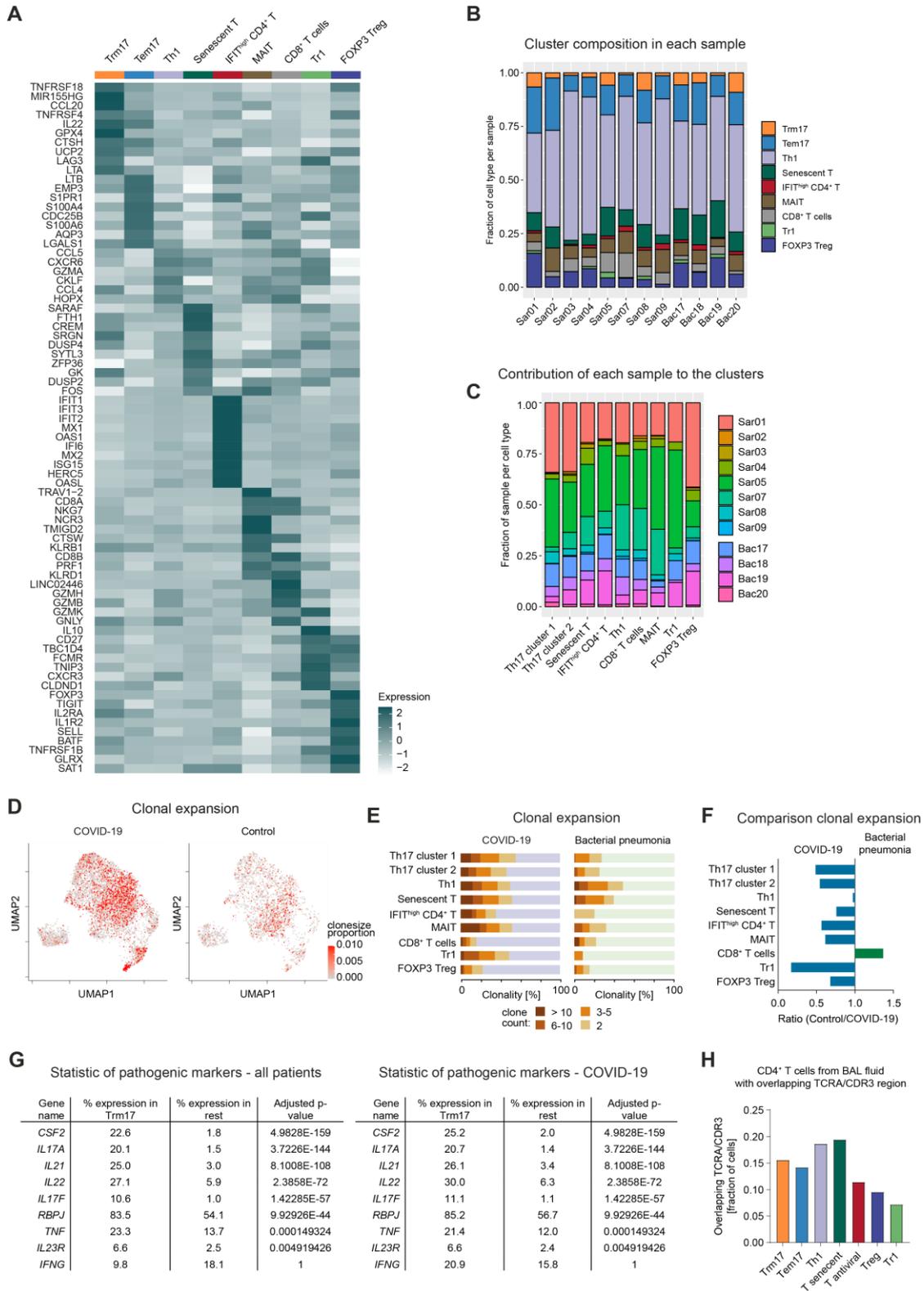
**Fig. S4. Location of T cells in the lungs of COVID-19.** (A) Lung autopsy tissue from non-pneumonia control (n=3), bacterial pneumonia (n=4) and COVID-19 (n=7) was analyzed by CD3 immunohistochemistry. (B) Scoring of perivascular infiltration of CD3 positive cells in the respective groups.



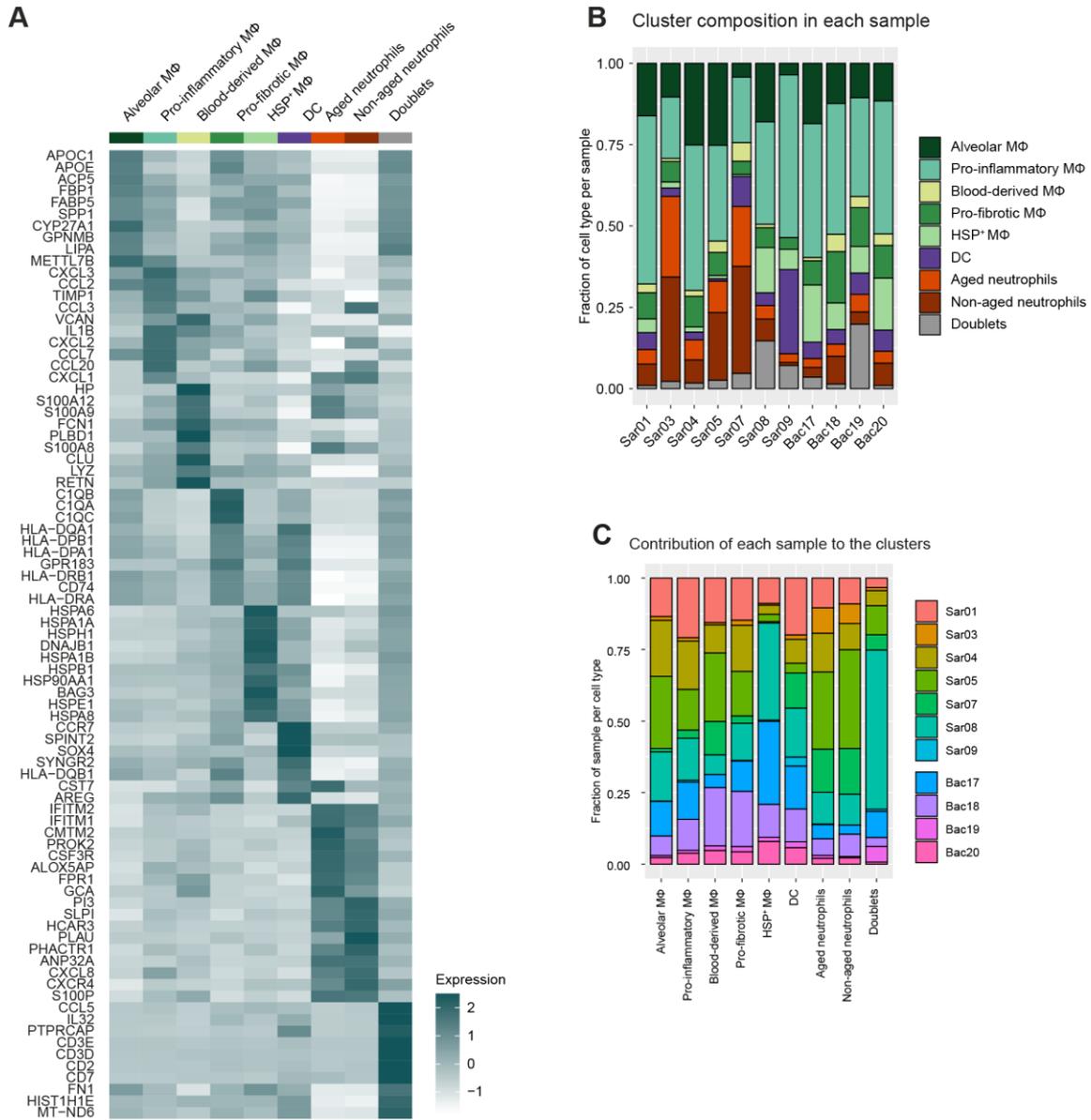
**Fig. S5. Clustering information of BALF T cells from all patients. (A)** Heat map of top 10 marker genes of each cell type. **(B)** Heat map showing the expression of representative genes to identify the cell type of each clusters. **(C)** CITE-seq expression profile of cluster-defining epitopes (scale bars indicate normalized expression). **(D)** Bar plot showing the cluster composition in each patient. **(E)** Bar plot showing the relative contribution to the clusters by each patient.



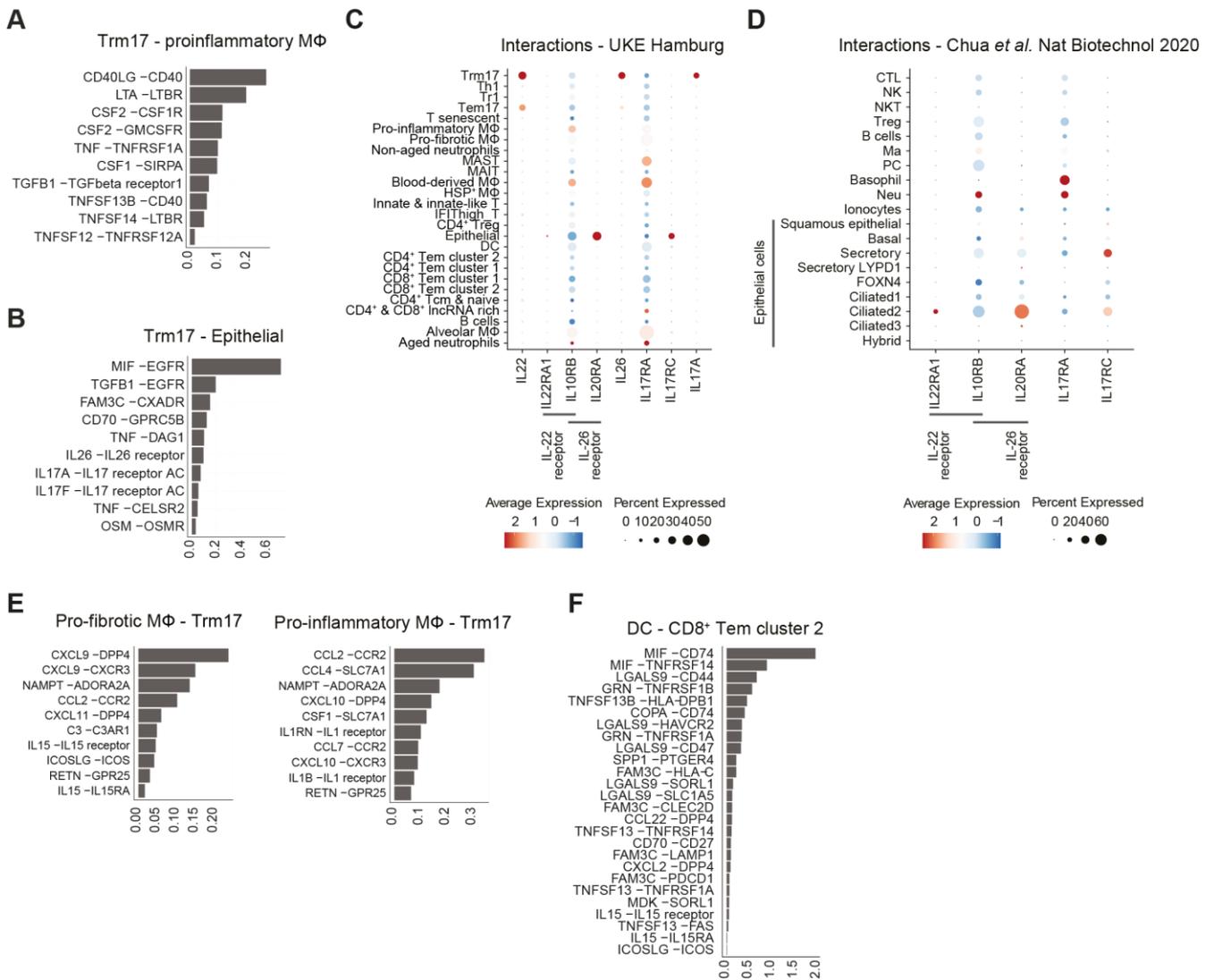
**Fig. S6. Clonal expansion analysis of BALF T cells from all patients.** The CDR3 region of the TCR sequences of CD4<sup>+</sup> T cells from our dataset were compared with two publicly available datasets of SARS-CoV-2-specific TCR sequences (Meckiff et al., Cell 2020 and Bacher et al., Immunity 2020). **(A)** Fraction of cells from COVID-19 and bacterial pneumonia patients with overlapping TCR sequences between our dataset and public data. **(B)** UMAP projection of clonally expanded T cells from COVID-19 and bacterial pneumonia patients. **(C)** Bar plot showing the clone size distribution for each cluster.



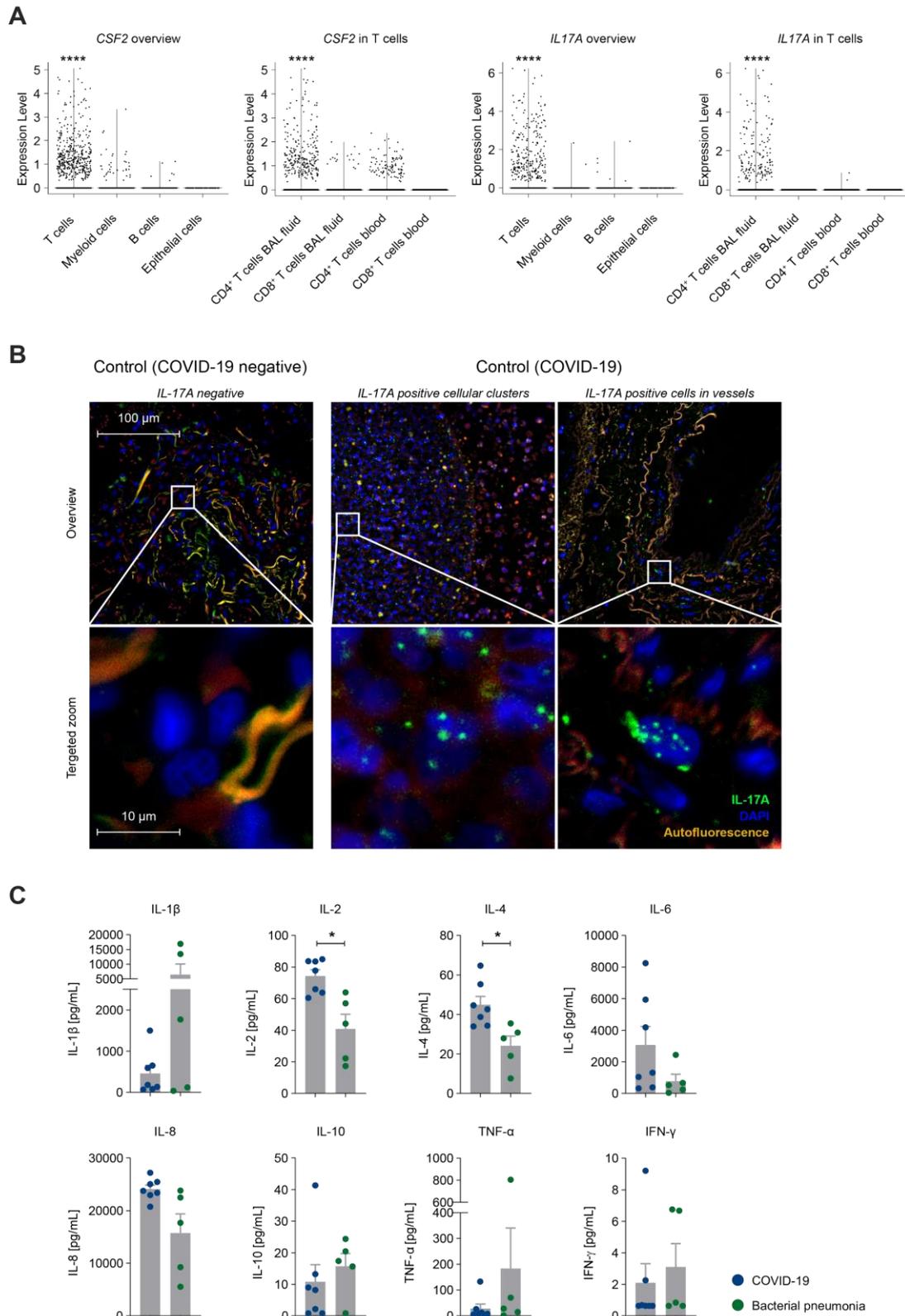
**Fig. S7. Subclustering of clonally expanded BALF CD4 T cells from all patients.** (A) Heat map of top 10 marker genes of each cell type. (B) Bar plot showing the cluster composition in each patient. (C) Bar plot showing the relative contribution to the clusters by each patient. (D) UMAP projection of clonal expanded T cells from COVID-19 and bacterial pneumonia patients. (E) Bar plot showing the clone size distribution in each cluster as indicated. (F) Ratio of expanded clones of bacterial pneumonia versus COVID-19 per cluster. (G) Statistical analysis of pathogenic markers expression in the Trm17 cell cluster versus all other clusters from all patients (bacterial pneumonia and COVID-19) and COVID-19 samples. (H) Fraction of cells with overlapping sequences in the respective cluster from our dataset compared with two publicly available datasets of SARS-CoV-2-specific TCR sequences (Meckiff et al., Cell 2020 and Bacher et al., Immunity 2020).



**Fig. S8. Clustering information of BALF myeloid cells from all patients. (A)** Heat map of top 10 marker genes of each cell type. **(B)** Bar plot showing the cluster composition in each patient. **(C)** Bar plot showing the relative contribution to the clusters by each patient.



**Fig. S9. Cell-cell interaction of T cells with myeloid cells and epithelial cells.** (A) Ligand and receptor interaction strength (mean ligand expression x proportion of receptor expression per cluster) of Trm17 cells (ligands) and pro-inflammatory macrophages (receptors). Interactions were filtered for cytokines and for specificity based on rank scoring. (B) Ligand and receptor interaction strength of Trm17 cells (ligands) and epithelial cells (receptors). Interactions were filtered for cytokines and for specificity based on rank scoring. (C) Ligand and receptor expression of cytokine signaling as indicated from our dataset (UKE Hamburg) (D) and from Chua et al, Nat Biotechnol 2020). (E) Ligand and receptor interaction strength of pro-fibrotic macrophages (ligands) and Trm17 cells (receptors) (left panel) and pro-inflammatory macrophages (ligands) and Trm17 cells (receptors) (right panel). Interactions were filtered for cytokines and for specificity based on rank scoring. (F) Ligand and receptor interaction strength of dendritic cells (ligands) and CD8<sup>+</sup> Tem cluster 2 cells (receptors).



**Fig. S10. Cytokine levels in BALF.** (A) Expression of *CSF2* (GM-CSF) and *IL17A* was analyzed in different cell types (T cells, myeloid cells, B cells and epithelial cells) and in T cell compartments (CD4 and CD8; BAL fluid and blood) (differential expression of each cluster against the average of other clusters using FindAllMarker function in Seurat; \*\*\*\*:  $p < 0.0005$ ). (B) Autopsy tissue from deceased patients with COVID-19 were analyze by fluorescence *in situ* hybridization (FISH) for IL-17A mRNA. Control samples were obtained from COVID-19 negative patients that died of non-infectious cause. (C) Cytokine levels in the BAL fluid of COVID-19 and bacterial pneumonia patients as indicated (t-Test; \*:  $p < 0.05$ ).

**Table S1. Baseline characteristics and disease-related parameters of patients with COVID-19 and controls.**

Baseline characteristics						Disease-related parameters at time of sampling						
Patient	Diagnosis	Sex	Age [y]	BMI [kg/m <sup>2</sup> ]	Relevant comorbidities	SOFA score	ARDS (Berlin definition)	Extracorporeal treatment	CrP [mg/dL; <5]	PCT [µg/L; <0.5]	WBC [10 <sup>9</sup> /L; 3.8-11.0]	Lymphocytes [10 <sup>9</sup> /L; 1.1-3.4]
S1	COVID-19 pneumonia	m	61	32	Type 2 diabetes mellitus, sarcoidosis, art. HTN	18	severe	ECMO, RRT	177	1.8	5.5	0.4
S2	COVID-19 pneumonia	m	31	41	Morbid obesity	10	severe	ECMO	30	3.0	13.2	3.2
S3	COVID-19 pneumonia	m	60	22	None	9	severe	no	160	7.9	10.0	0.7
S4	COVID-19 pneumonia	m	79	25	Atrial fibrillation, lymphocytic colitis	10	severe	no	206	1.4	10.6	0.8
S5	COVID-19 pneumonia	m	76	31	Type 2 diabetes mellitus, atrial fibrillation, hemochromatosis, rheumatoid arthritis, history of intracranial bleeding, art. HTN	15	severe	RRT	212	3.1	11.6	0.9
S6	COVID-19 pneumonia	m	71	26	Metastatic prostate cancer, pleural asbestosis, coronary artery disease, third-degree AV block with pacemaker, history of cerebral insult, art. HTN	15	severe	RRT	223	4.6	16.9	0.7
S7	COVID-19 pneumonia	m	80	25	Myelofibrosis, pyoderma gangrenosum, coronary artery disease, history of tuberculosis, history of pulmonary embolism, art. HTN	12	severe	no	291	2.2	10.3	2.0
S8	COVID-19 pneumonia	m	73	26	Atrial fibrillation, art. HTN	17	moderate	RRT	229	10.1	15.8	1.6
S9	COVID-19 pneumonia	f	59	25	Light chain myeloma	13	moderate	no	265	7.4	0.0	0.0
B1	Bacterial pneumonia	m	63	16	Type 2 diabetes mellitus, coronary artery disease, chronic kidney disease, history of papillary thyroid carcinoma	13	moderate	no	67	6.9	11.5	ND
B2	Bacterial infection / mediastinitis	m	56	26	Gastric adenocarcinoma, history of esophageal carcinoma with esophagectomy, art. HTN	15	moderate	RRT	259	2.7	14.7	ND
B3	Bacterial pneumonia	f	62	35	Type 2 diabetes mellitus, ulcerative colitis, history of gastrointestinal stromal tumor of the small intestine, history of cerebral insult, art. HTN	15	no	no	145	0.9	8.7	ND
B4	Bacterial pneumonia	m	88	31	History of rectal carcinoma, art. HTN	7	moderate	no	102	ND	8.0	ND
B5	Bacterial pneumonia	f	68	22	Pulmonary aspergillosis, atrial fibrillation, esophageal carcinoma, emphysema	10	no	no	123	0.3	9.9	ND

BMI = body mass index, art. HTN = arterial hypertension, SOFA = sequential organ failure assessment, ARDS = acute respiratory distress syndrome, ECMO = extracorporeal membrane oxygenation, RRT = renal replacement therapy, CrP = C-reactive protein, PCT = procalcitonin, WBC = white blood cell count. Units and reference range for laboratory parameters are given in brackets. ND = not done.

**Table S2. Relevant medication of patients with COVID-19 and controls.**

Patient	Immune modulators / chemotherapy (at time of sampling)	Immune modulators / chemotherapy (within 14 days prior to sampling)	Antivirals (at time of sampling)	Antivirals (within 14 days prior to sampling)	Antibacterials (at time of sampling)	Antifungals (at time of sampling)	Permanent Medication (ACE-Inhibitor / Angiotensin receptor blocker)
S1	no	Glucocorticoids	no	Lopinavir/ritonavir	Ceftazidim, vancomycin	no	yes
S2	no	no	Aciclovir	Aciclovir	Levofloxacin	no	no
S3	no	no	no	no	Meropenem	no	no
S4	no	Glucocorticoids	no	no	Meropenem, vancomycin	no	no
S5	Methotrexate Prednisolone (2,5mg/d)	Methotrexate Glucocorticoids	Aciclovir	Aciclovir	Vancomycin, levofloxacin	Caspofungin	yes
S6	no	Glucocorticoids	no	no	Meropenem	no	yes
S7	Ruxolitinib Prednisolone (40mg/d)	Ruxolitinib Glucocorticoids	no	no	Meropenem	no	no
S8	no	Glucocorticoids	Aciclovir	Aciclovir	Cefotaxim, levofloxavin	no	yes
S9	Salvage chemotherapie	Salvage chemotherapy Intravenous immunoglobulines	Aciclovir	Aciclovir	Atovaquone, vancomycin, meropenem	Posaconazol	no
B1	no	no	no	no	no	Voriconazol	no
B2	no	no	no	no	Meropenem, vancomycin	Caspofungin	no
B3	no	no	no	no	Piperacillin/tazobactam	no	yes
B4	no	no	no	no	Ampicillin/sulbactam	no	yes
B5	Radiochemotherapie	Radiochemotherapie	no	no	Piperacillin/tazobactam	Voriconazol	no

**Table S3. Baseline characteristics of the Halle cohort.**

Patient	Diagnosis	Sex	Age [y]	BMI [kg/m <sup>2</sup> ]	ARDS (Berlin definition)	Relevant comorbidities
S-09H	COVID-19 pneumonia	F	68	22.2	Moderate	AV block I, art HTN, hypothyroidism, adrenal adenoma
S-10H	COVID-19 pneumonia	F	70	24.7	Moderate	CHE, Barrett's oesophagus, CKD, art. HTN, diabetes mellitus, steatosis hepatitis
S11H	COVID-19 pneumonia	M	29	N/A	Moderate	Hodgkin's disease
S12H	COVID-19 pneumonia	M	30	N/A	Moderate	N/A
S-13H	COVID-19 pneumonia	F	23	N/A	Moderate	N/A
S-14H	COVID-19 pneumonia	M	68	37.0	Moderate	CAD, MI, art. HTN, diabetes mellitus, stroke
S-15H	COVID-19 pneumonia	M	67	25.9	Severe	Art. HTN, diabetes mellitus
S-16H	COVID-19 pneumonia	M	66	25.7	Severe	AML, prostate cancer
S-17H	COVID-19 pneumonia	M	61	35.8	Severe	AML, art. HTN, cardiomyopathy, art. HTN
S-18H	COVID-19 pneumonia	M	76	23.3	Severe	NHL, AIHA, myocardial infarction, rheumatoid arthritis,
S-19H	COVID-19 pneumonia	M	41	41.4	Severe	Art. HTN, Factor XIII deficiency
S-20H	COVID-19 pneumonia	M	68	30.4	Severe	Art. HTN, atrial fibrillation
S-21H	COVID-19 pneumonia	M	68	25.9	Severe	Chronic obstructive pulmonary disease
S-22H	COVID-19 pneumonia	M	65	25.8	Severe	Myasthenia gravis
S-23H	COVID-19 pneumonia	M	74	27.8	Severe	N/A
S-24H	COVID-19 pneumonia	F	76	25.0	Moderate	Art. HTN, CKD
S-25H	COVID-19 pneumonia	F	62	32.4	Moderate	Renal cell carcinoma
S-26H	COVID-19 pneumonia	M	54	N/A	Severe	Art. HTN
S-27H	COVID-19 pneumonia	M	70	N/A	Severe	CKD, cardiac transplantation, art. HTN, diabetes mellitus

Art. HTN = arterial hypertension, CHE = cholecystectomy, CKD = chronic kidney disease, CAD = coronary artery disease, MI = myocardial infarction, AML = acute myeloid leukemia, NHL = Non-Hodgkin lymphoma, AIHA = autoimmune hemolytic anemia

## Chapter 6. Discussion

### 6.1 Infection and autoimmunity

Infection can trigger memory cell formation and therefore change the immune system in the long term. This would enable rapid immune responses when the system is re-exposed to known pathogens. Autoimmune diseases in general arise from disturbed tolerance of the immune system, but in many cases detailed mechanisms are unclear. In infectious diseases, the resident immune cells are protective and fight against the pathogens. In contrast, in autoimmune disease, the immune cells are pathogenic and trigger reactions against their own body structures. In this thesis, the results of the computational characterization of T cells in different diseases suggest the potential connection between infection and autoimmunity.

#### 6.1.1 Bacterial infection and kidney autoimmunity

In the ANCA-GN project, our experimental mouse model demonstrates a strong link between bacterial infection and autoimmunity. The key mediator is the TRM17 cells. Bacterial infection induces the persistence of TRM17 cells and they can activate and aggravate autoimmunity. In human ANCA-GN, our single cell analysis characterized TRM17 cells using signature genes and these cells also show upregulation of proliferation, cytokine signaling and cell activation genes. Our current research could not demonstrate the origin of TRMs detected in ANCA-GN patients. However, the increased relapse frequency of ANCA vasculitis in patients carrying *S. aureus* on their nasal mucosa indicates a link between bacterial infection and autoimmune response (Stegeman *et al.*, 1994). More interestingly, *S. aureus*-reactive T-cell clones were found to be able to recognize PR3 antigen (one of the auto-antigens of ANCA-GN) (Mayet *et al.*, 1999). The microbial factors have also been suggested as key environmental factors of other autoimmune kidney diseases, such as lupus nephritis (Qiu *et al.*, 2019).

It is tempting to hypothesize that the TRMs we observed in our ANCA-GN patients' kidneys were induced by previous bacterial infection. Of course, we can not rule out the possibility that these TRMs were triggered directly by autoantigens. In order to further address the origin of the kidney TRMs in humans, the TCR sequencing techniques could potentially be used. The relevant questions that need to be answered are: is there T cell clonal expansion in the ANCA-GN patients? If there are indeed clonal expanded T cells in the ANCA-GN kidneys, are they specific to certain bacterial antigens or autoantigens?

### 6.1.2 Gut-liver axis and liver inflammation

The key purpose of the NAFLD/NASH project was to understand which components contribute to the transition from pure steatosis (NAFLD) to NASH. We characterized multi-cytokine producing liver T cells as a distinguished feature of NASH patients. One major limitation of our NASH study is the lack of direct evidence to show if the multi-cytokine producing liver T cells in NASH patients are related to intestinal and liver microbiota, however, recent studies have suggested a potential link between hepatic T cell activation and gut-liver axis infection.

Portal vein contributes up to 70% of the liver blood supply, meaning that the liver is exposed to gut-derived bacterial components (Wiest *et al.*, 2017). The main hypothesis of gut-liver axis infection is that the dysfunction of the gut barrier results in bacterial translocation and further triggers liver inflammation (Schuster *et al.*, 2018). Dysbiosis of the gut microbiota is evident in some NASH patients (Bashiardes *et al.*, 2016; Brandl and Schnabl, 2017). In this context, excess bacterial products such as endotoxin can enter the liver through portal circulation (Farhadi *et al.*, 2008). These bacterial products can disturb the liver immune tolerance by activating liver immune cells such as Kupffer cells or CD8<sup>+</sup> T cells (Farhadi *et al.*, 2008). Whether liver CD4<sup>+</sup> T cells are responsive to the bacterial products in the NASH patients remains to be explored in the future.

NAFLD/NASH are conventionally not considered as autoimmune disease. Typical autoimmune liver diseases include for instance AIH (autoimmune hepatitis) and PBC (primary biliary cholangitis). However, the comorbidity of NAFLD, AIH and PBC makes the diagnosis of chronic liver diseases challenging (Dalekos *et al.*, 2020). In a clinical study performed by Tsuneyama *et al.* (Tsuneyama *et al.*, 2013), they examined 54 NASH patients proven by histology and around half of them were positive for autoimmune serology testing. These patients also have clinical features of AIH or PBC. In another study, 182 out of 864 patients with NAFLD have significantly elevated titers of serum autoantibodies (Vuppalanchi *et al.*, 2012). These studies suggest that at least part of the NAFLD/NASH patients have autoimmune features.

The function of liver TRMs in NAFLD/NASH patients is still not clear. How they contribute to the autoimmune features in some patients remains to be investigated. In the future, the clonal expansion of the liver T cells can be accessed using single cell TCR-seq. Ideally, the sequencing should be performed on paired liver and gut samples to further compare the tissue similarities and differences. If possible, the paired microbiota data from the same patients can be analyzed simultaneously to further infer the correlation of the gut-liver axis and liver inflammation.

### 6.1.3 SARS-CoV-2 infection and autoimmunity

COVID-19 results clearly from SARS-CoV-2 virus infection but the increased inflammatory cytokine levels of severe COVID-19 patients suggests autoimmune-like features. Our single cell analysis clearly demonstrated the pathogenic role of T cells in the hyperinflammation of severe COVID-19 patients. Numerous studies performed during the past pandemic year indicated the link between SARS-CoV-2 infection and autoimmunity. This idea is supported by several perspectives.

First, multiple COVID-19 symptoms resemble autoimmune-like pathology. The most typical symptom is excessive production of inflammatory cytokines such as TNF- $\alpha$ , IFN- $\gamma$ , IL1B, GM-CSF. Both Th1 and Th17 responses were observed in severe COVID-19 patients, suggested by our data and also previous studies (Chen and John Wherry, 2020; Zhao *et al.*, 2021). Macrophage, neutrophil, T cell and B cell activation all accompany the cytokine storm. The uncontrolled autoimmune-like responses further induce damage to different organs and systems such as the lung, kidney, nervous system and cardiovascular system (Liu *et al.*, 2021).

In addition to the autoimmune-like symptoms, autoantibodies have also been detected in COVID-19 patients. In a very recent study published in May, 2021, Wang *et al.* used the rapid extracellular antigen profiling technique to detect the autoantibody in 172 COVID-19 patients and 22 healthy controls (Wang *et al.*, 2021). They discovered dramatic increases of autoantibodies in COVID-19 patients, and these autoantibodies directly target extracellular and secreted proteins which are involved in immunomodulatory functions. This study provided strong evidence to support the role of autoimmunity in COVID-19 pathogenesis. The authors of this study also believe these autoantibodies might explain the long-lasting symptoms for weeks or months after first infection of some COVID-19 patients. Another very important study reported the detection of autoantibodies against type I interferons in about 10% of 987 hospitalized COVID-19 patients (Bastard *et al.*, 2020). Strikingly, no one in the 663 mild/asymptomatic COVID-19 patients and only 4 of 1227 healthy individuals had the same autoantibodies detected. Since interferons play a critical role in virus infection defense, the anti-interferon antibodies might prevent a normal virus immune response and lead to severe disease outcome.

Last but not least, various autoimmune diseases have been diagnosed in a subgroup of SARS-CoV-2-infected patients. The autoimmune diseases secondary to COVID-19 include Guillian-Barré syndrome (Toscano *et al.*, 2020); Immune thrombocytopenic purpura (Zulfiqar *et al.*, 2020); and Kawasaki disease (Verdoni *et al.*, 2020). More interestingly, there were case studies reporting ANCA vasculitis secondary to COVID-19 (Izci Duran *et*

*al.*, 2021; Powell *et al.*, 2021). Circulating autoantibodies of these autoimmune diseases have been detected in COVID-19 patients (Halpert and Shoenfeld, 2020).

#### 6.1.4 Mechanisms of infection in autoimmune diseases

Bacterial or viral infections have been also associated with the development and aggravation of other autoimmune diseases, such as multiple sclerosis (Steelman, 2015) or inflammatory bowel diseases (IBDs) (Mann and Saeed, 2012). One of the well studied associations between a specific virus and an autoimmune disease is that between Epstein–Barr virus and multiple sclerosis (Mann and Saeed, 2012).

The underlying cellular and molecular mechanisms of how infection enhances autoimmunity are still not fully understood. There are a couple of theories, including molecular mimicry, bystander T cell activation and epitope spreading (Ercolini and Miller, 2009). Molecular mimicry refers to a situation in which a pathogen antigen resembles self proteins in the host body. When the pathogen derived epitopes are presented to T cells by APCs, activated T cells might organize the attack to both the pathogen and self tissue (Toscano *et al.*, 2020). A very classic example is T cells from rheumatic heart inflammation patients that recognize the bacteria *Streptococcus pyogenes* and host cardiac muscle protein (Guilherme *et al.*, 1995). Bystander T cell activation means that T cells get activated in a non-specific, TCR-independent manner (Pacheco *et al.*, 2019). In epitope spreading (Root-Bernstein and Fairweather, 2014), self-tissue damage caused by persistent pathogen infection might reveal self-antigens which are normally unseen by the self immune system. These self-antigens released from destructed tissue could become targets of APCs and T cells.

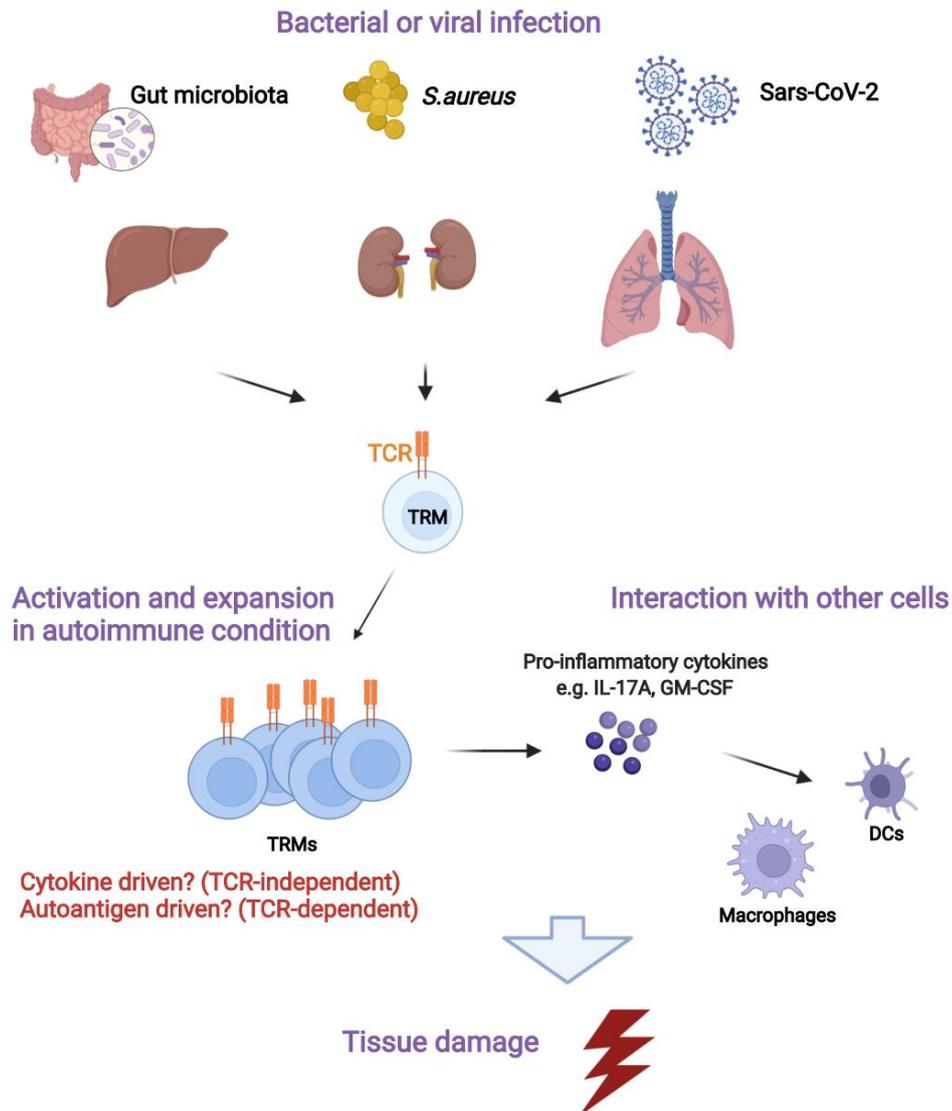
Since the bacterial infection in the ANCA-GN mouse model did not cause direct damage or strong autoimmune response at the kidney, it is less likely to be explained by molecular mimicry and epitope spreading mechanisms. However, our experiments suggest that the

bacteria-induced TRMs become bystander T cells and later in autoimmune condition, the change of the local environment activates the bystander TRM cells and they could further damage the tissue. Since NAFLD/NASH is not a clearly defined autoimmune disease, I will not discuss the above autoimmune mechanisms in detail.

The multi-organ damage in COVID-19 could be potentially explained by molecular mimicry (Angileri, 2020). Angileri *et al.* reported the similarity of SARS-CoV-2 protein and 3 human proteins: OR7D4 (Odorant Receptor 7D4, expressed on olfactory neurons); PARP9 (Poly ADP-Ribose Polymerase Family Member 9, expressed by B cell and macrophages) and SLC12A6 (Solute Carrier Family 12 Member 6, expressed by endothelial cells). The molecular mimicry of these proteins might be linked to the anosmia, leukopenia, and vascular damage observed in COVID-19 patients. The epitope spreading hypothesis has also been proposed as a potential mechanism underlying COVID-19, since SARS-CoV-2 has been detected in multiple organs and tissue damage was observed (Gupta and Weaver, 2021). The self-antigens revealed after the tissue damage might cause further autoimmune responses. Our sequencing data analysis of severe COVID-19 patients clearly demonstrates the activation and clonal expansion of T cells. As our patients were infected by SARS-CoV-2 virus, one can easily expect that expanded TCR clones are specific to SARS-CoV-2 antigens. However, there is a possibility that bystander T cells are activated and expanded during SARS-CoV-2 infection.

In summary, our single cell analysis and other studies suggest the strong association between bacterial/viral infection and autoimmune-like inflammatory responses (Fig 4). The central idea is that pathogen infections trigger a sustained immune response and attack the own body structures of the hosts. TRMs might play a bridging role between infection and autoimmunity due to their tissue resident feature and long lasting memory. The time interval between the infection and the autoimmune onset could be very long, therefore, it is hard to build direct connections between the two. However, the COVID-19 pandemic turns out to be a special circumstance. The link between viral infection and autoimmunity has been pushed into the center of public healthcare focus by COVID-19. This also suggests

immunosuppression to be an important strategy to treat COVID-19 severe patients since there are available immunosuppressive drugs in autoimmune diseases. The treatment strategies will be further discussed in section 6.7.



**Fig 4. Conceptual link between infection and autoimmunity.** Chronic or acute infection can trigger the formation of TRMs in different organs such as the liver, kidney and lung. These TRMs might be activated and even clonally expanded in autoimmune or autoimmune-like conditions. It is still unclear whether the activation and expansion are cytokine driven or autoantigen driven. Both TCR-independent and TCR-dependent mechanisms are possible. This can be further addressed by analyzing the TCR sequences and predicting potential autoantigens. Activated TRMs can produce pro-inflammatory cytokines and interact with other cells. The interaction of TRMs and other local cells could aggravate inflammation and lead to further tissue damage.

## 6.2 Characterization and functions of TRMs

Humans are constantly exposed to the external environment that contains a plethora of different pathogens, leading to the establishment of numerous TRMs. For the experimental animal model, in the clean laboratory environment, the TRMs are less abundant but can be induced by controlled bacterial infection such as *S.aureus*. TRMs are characterized by their long residence time and low division rate in different organs. In mouse models, several experimental approaches can be applied to examine the tissue residency of T cells. Parabiosis experiments are regarded as the gold standard for the investigation of TRMs. The blood circulations of two mice are connected over weeks and subsequently the distribution of the studied cells between the mice can be investigated (Kamran *et al.*, 2013). In addition to parabiosis, we also used *in situ* cell labeling (photo converted transgenic Kaede mice) and *in vivo* intravascular antibody staining to access the tissue residency. In human tissues, it is impossible to perform these experimental approaches and therefore, the characterization of TRMs relies largely on the key surface markers such as CD69 or CD103 and transcriptional profiling (Szabo *et al.*, 2019). In most of the recent single cell studies of human T cells from different tissues, characterization of TRMs relies purely on the TRM core transcriptional signature based on previous bulk-seq studies of sorted CD69 and CD103 T cells (Braga *et al.*, 2019; Kozlenkov *et al.*, 2018; Savas *et al.*, 2018). We used the same strategy to identify TRMs in the NASH project. However, our single cell analysis combining RNA and CITE-seq of T cells from ANCA-GN and COVID-19 patients enables a more accurate characterization of TRMs in kidney and lung. More importantly, the paired TCR-seq not only helps to identify clonal expansion, but also indicates the tissue migration or residency of the expanded T cells.

TRMs have both positive functions and negative functions. They can rapidly clear infection locally but they can also activate in inflammatory diseases and lead to increased immunopathology.

The pathogenic roles of TRMs have been demonstrated in various autoimmune diseases including psoriasis, autoimmune hepatitis and rheumatoid arthritis (Wu *et al.*, 2018). In psoriasis patients, skin lesions frequently recur at the same locations, suggesting the potential role of memory cells. IL-17A-producing CD4<sup>+</sup> TRMs have been later demonstrated to be a key mediator of psoriasis (Ryan *et al.*, 2021). Similar to psoriasis, ANCA-GN patients also frequently suffer from relapse (Jennette and Nachman, 2017). Our experimental mouse model clearly demonstrated that TRM17 cells can aggravate kidney autoimmunity, giving further evidence of the TRMs' pathogenic roles. Due to the human sample quality variation and different treatment conditions, we did not correlate the TRM17 cells of ANCA-GN patients with disease severity. This needs further investigation in a larger patient cohort. Our single cell analysis of NAFLD/NASH patients characterized the multi-cytokine T cells as a pathogenic feature and these cells also have a TRM profile. The pathogenic role of liver resident T cells has been also explored in other groups. For example, in a recent NASH single cell sequencing study, Dudek *et al.* characterized a group of liver resident CXCR6<sup>+</sup> CD8<sup>+</sup> T cells in NASH patients and a mouse model of the disease (Dudek *et al.*, 2021). This group of CD8<sup>+</sup> TRMs is capable of killing hepatocytes in an antigen-independent way.

Our research and most other studies suggest pathogenicity of TRMs in many inflammatory diseases. However, in COVID-19, the protective/pathogenic role of TRMs has been a subject of considerable debate (Altmann and Boyton, 2020). We speculate that the time of action is the real discriminating factor, in other words, TRMs are protective at early phase but inappropriate constant activation can be pathogenic. However, in our study, we did not have data from the early phase of the infection or asymptomatic patients. Peripheral blood is the most commonly sampled tissue, assessing TRMs from lung tissue is more difficult, especially for asymptomatic or early phase patients. Interestingly, some studies reported the pre-existing cross-reactive memory T cells in old blood samples stored before COVID-19. For example, Grifoni *et al.* and Braun *et al.* both detected SARS-CoV-2 cross-reactive CD4<sup>+</sup> T cells in the blood samples from around 35–60% of unexposed individuals (Braun *et al.*, 2020; Grifoni *et al.*, 2020). These cross-reactive memory cells are

likely to be formed due to previous infection of other coronaviruses (Lipsitch *et al.*, 2020). Although these blood T cells are likely to be TCMs or TEMs instead of TRMs, the above findings indicate the existence of SARS-CoV-2 cross-reactive TRMs in unexposed individuals. It is still not clear whether the cross-reactive memory T cells can give additional protection or they can suppress a *de novo* primary response upon SARS-CoV-2 infection or even mediate the hyperinflammatory responses (Jarjour *et al.*, 2021). In order to understand the protective and pathogenic roles of TRMs in COVID-19 patients, further investigation is required to address the dynamic and composition of the TRMs.

### **6.3 Interaction of TRMs with other immune/non-immune cells**

To better understand the function of TRMs, the focus should not be only on TRMs. Interaction of TRMs with other immune or non-immune cells must be taken into consideration. In our COVID-19 project, we used interactome analysis to infer the link between TRM17 cells and myeloid cells. This interaction analysis further suggests how TRM17 cells mediate fibrosis and tissue damage in severe COVID-19. Due to the limited amount of epithelial cells acquired from the BAL fluid in our patient cohort, the interaction between T cells and epithelial cells was not addressed in detail. In another COVID-19 single cell study, the interaction between immune cells and epithelial cells were investigated (Chua *et al.*, 2020). In this study, they collected immune and non-immune cells from the upper and lower respiratory tracts. They also inferred the intercellular communication by ligand-receptor analysis using CellPhoneDB. They found that the interactions between immune cells and epithelial cells are higher in severe COVID-19 patients compared to moderate patients. Since lung epithelial cells express SARS-CoV-2 receptors ACE2 and TMPRSS2 and the virus first infects these cells, the damage of epithelial cells is likely to happen before the inflammatory responses. However, in ANCA-GN and NASH, there are no defined pathogens causing initial damage to epithelial cells, so the interaction of immune cells and epithelial cells might be more important to understand the origin of tissue damage.

In a previous human kidney single cell study (Stewart *et al.*, 2019), the ligand-receptor interactions between myeloid cells and nephron cell types were reported. The epithelial immune crosstalk was clearly demonstrated, however, the interaction of renal T cells and other cell types were not described in the study. There are some potential mechanisms describing how the interaction of Th17 cells and resident kidney cells might contribute to renal damage. For example, glomerular mesangial cells and tubular epithelial cells can express IL-17R (receptor of IL17A produced by Th17 cells). After receiving IL17A signals, these resident kidney cells can release CCL20 and CXCL5, which further recruit Th17 cells and neutrophils, causing severe tissue injury (Krebs *et al.*, 2017). However, an unbiased interactome based on single cell analysis in the inflamed kidney of ANCA-GN patients is still required. In the future, the myeloid cells and non-immune cells in the kidney biopsy of ANCA-GN patients should be also profiled. The interactions between the resident T cells and myeloid or parenchymal cell compartments will be crucial to understand the molecular mechanism of glomerular disease pathology.

In a recent liver cirrhosis single cell study, Ramachandran *et al.* characterized fibrosis associated macrophages, endothelial and mesenchymal cells in human liver (Ramachandran *et al.*, 2019). The multi-lineage interaction analysis between these fibrosis-associated cell types highlights TNFRSF12A, PDGFRA and Notch signalling as important regulators in the liver fibrotic niche. How TRMs interact with these fibrosis-associated liver cells remains elusive, and in our NAFLD/NASH study, we indicated TRMs might be involved in the fibrosis development. One interesting detail of our single cell data is that we had 2 Kupffer cell clusters potentially due to insufficient isolation of cells. Kupffer cells might form immune synapses with T cells, leading to potential direct cell-cell interactions. A very recent approach called PIC-seq (physically interacting cell sequencing) specifically focuses on these doublets in single cell analysis (Giladi *et al.*, 2020). They used this technique to investigate the T cell-myeloid interactions in an *in vivo* infection model. It is tempting to apply this strategy to study the crosstalk between TRMs and their neighboring cells in different organs.

Of note, although single cell sequencing gives a detailed characterization of different cell types and the ligand-receptor interactome analysis provides information about the communication between various cell types, the spatial information about cellular interactions is lost during the tissue digestion and single cell isolation process. In order to better study cellular interactions in situ, spatial transcriptomics combined with imaging techniques were developed (Rodrigues *et al.*, 2019; Ståhl *et al.*, 2016). Integration of single cell RNA-seq data and spatial transcriptomics will give high resolution maps of tissue dynamics and cellular communication (Longo *et al.*, 2021).

#### **6.4 Tissue differences of TRMs**

Memory T cells include different subtypes distinguished by their functions and locations. TRMs reside at the peripheral organs such as kidney, pancreas as well as surface tissue to the external environment such as skin or mucous membranes (Szabo *et al.*, 2019). These TRMs can provide a site-specific immune response upon reinfection. TRMs are distributed across multiple organs, therefore, they need to adapt to different local environments under transcriptional regulation (Mackay and Kallies, 2017). In this thesis, although TRMs from three different organs (kidney, liver, lung) were analyzed, we did not make direct comparisons across the tissues. It will be very important to know the tissue/organ-specific features of the TRMs in order to target them in a site-specific way when considering the therapeutic strategy to treat TRM-mediated inflammatory diseases.

Miragaia *et al.* explored Tregs tissue adaptation using single cell sequencing (Miragaia *et al.*, 2019). By comparing the transcriptome signature of T cells from mouse skin, colon, lymph nodes and spleen, they identified the TNFRSF-NF- $\kappa$ B pathway as key to barrier tissue adaptation of Tregs. In another very recently published study, Christo *et al.* compared mouse CD8<sup>+</sup> TRMs development across multiple organs including skin, liver

and salivary gland (Christo *et al.*, 2021). They identified TGF- $\beta$  as a major driver of tissue specificity of TRMs.

In human studies, TRMs from different tissues are usually generated in separate studies from different laboratories and are usually subject to strong inter-study batch effects. The difficulty of removing unwanted batch effects might be reduced by new computational single cell data integration methods (Oller-Moreno *et al.*, 2021). However, the most ideal case is to profile TRMs from different tissues of the same patients in the future. This is surely difficult for living donors, however, it is possible to collect tissues such as liver, small intestine, fat and peripheral blood from the same patients if they receive bariatric surgery. Single cell sequencing of the T cells from these samples would give great insights into tissue specificity of TRMs.

## **6.5 Clonal expansion of T cells in inflammatory diseases.**

In our COVID-19 project, we used the combination of single cell RNA-seq, CITE-seq and TCR-seq to obtain a comprehensive landscape of immune cells in severe COVID-19 patients. By tracking T cell clones across tissues, we identified clonally expanded Trm17 cells in the lungs even after viral clearance.

The ANCA-GN and NAFLD/NASH projects started before the COVID-19 pandemic and at that time, our paired single cell transcriptome, surface protein and TCR sequencing workflow was not established. Benefiting from the experience of performing the COVID-19 patient single cell analysis, we were able to establish a robust pipeline to simultaneously profile single-cell transcriptome with surface protein measurement and VDJ recombination as introduced in details in Chapter 2. It is very natural to ask: are the T cells in ANCA-GN and NAFLD/NASH clonally expanded? Actually, my colleagues started to perform single cell TCR-seq for tissue samples from ANCA-GN and NAFLD/NASH patients and our preliminary analysis already detected the expansion of T cells in both diseases.

The expansion of T cells has been explored also in other autoimmune diseases such as MS and IBD. scRNA-seq analysis combined with TCR profiling has been recently used in a cohort of pediatric patients with colitis and IBD (Huang *et al.*, 2019). This study showed the preferential clonal expansion of two subsets of CD8<sup>+</sup> T cells expressing *ITGAE* and in CD8<sup>+</sup> T cells expressing *GZMK*. In another recent study, Pappalardo *et al.* used paired scRNA-seq and TCR sequencing to characterize the T cells from the blood and cerebrospinal fluid of MS patients and healthy individuals (Pappalardo *et al.*, 2020). Differential gene expression analysis revealed a gene set associated with MS. Many of these genes were related to T cell activation (*CD69*, *RAC2*, *IL32*), cytotoxicity (*GZMA*, *GZMK*, *SRGN*, *CST7*, *PLEK*, *LITAF* and *PFN1*), and an effector phenotype (*CD74*, *EOMES*, *KLRG1*).

T cell expansion seems to be a general feature of different inflammatory diseases. However, in each particular disease, do the expanded T cells have similar functions or do they have specific roles? Do the expanded T cells have tissue resident memory features? These are still open questions that need to be addressed in the near future. It will be informative to explore whether the clonal expanded T cells in ANCA-GN kidney or NAFLD/NASH liver samples have similar transcriptional signatures with expanded T cells in the other autoimmune diseases. The published data sets can be downloaded and we can evaluate similarity by calculating a correlation coefficient of the gene expression values in the expanded T cells from in-house and published studies. There is potential to identify common gene markers shared by clonal expanded T cells among different autoimmune diseases and to find kidney/liver-specific gene signatures. This would be an essential step towards disease-specific treatment of autoimmunity and inflammation.

## **6.6 Interpretation of expanded TCR sequences.**

Single cell TCR-seq proves to be extremely powerful to understand the T cell clonal expansion in inflammatory responses. Most of the current studies focus on the

transcriptomic features of expanded T cells. However, TCR sequence can serve as a natural barcode to trace the T cells across tissues. Additionally, TCR sequences also contain valuable information for the prediction of the antigens.

In our COVID-19 data analysis, we focused on the BAL fluid specific clones but we indeed also identified a decent amount of shared clones between BAL fluid and blood T cells (Fig 2G in Chapter 5). Interestingly, shortly after we published our study, Simnica *et al.* had a deeper look into our TCR-seq data and overlapped with TCR-seq of brain tissue from 5 fatal COVID-19 patients (Simnica *et al.*, 2021). From our BAL fluid and blood data, they identified 39 T cell clones with high sequence similarity to their brain-derived TCR sequences. Of these 39 clones, 9 were mapped to the BAL fluid CD4<sup>+</sup> T cell clones; all can be mapped to the blood CD4<sup>+</sup> T cell clones but none can be mapped to the BAL/blood CD8<sup>+</sup> T cell clones. This suggests expanded T cells might migrate across different tissues and this tissue sharing pattern might be different for helper T cells and cytotoxic T cells.

Although we sequenced the TCR genes and inferred the clonal expansion in the severe COVID-19 patients, we did not fully address the antigen specificity of the clonal expanded T cells. In the Supplementary Fig 6 of Chapter 5, we showed the shared TCR  $\alpha/\beta$  sequences between our patients and other published datasets of SARS-CoV-2-specific T cells (Meckiff *et al.*, 2020; Schultheiß *et al.*, 2020). The fractions of shared TCR sequences to the public datasets were higher in our COVID-19 patients than our bacterial pneumonia patients. Within our COVID-19 patients, we also observed higher frequencies of shared clones in T cells from the BAL fluid as compared to peripheral blood. However, this still does not mean all the expanded T cells in COVID-19 were SARS-CoV-2 specific. A more accurate experimental method to address this question would be *in vitro* stimulation of BAL fluid cells with SARS-CoV-2 peptides and following measurement of the TCR-sequences for expanded T cells. This will require a larger patient cohort to provide enough cells.

The TCR-antigen specificity in most of the autoimmune diseases remains largely unknown. Classifying the TCR sequences of the expanded T cells in the inflammatory diseases would be of great relevance to the prediction of the potential antigen which triggered the T cell response. In autoimmune diseases such as ANCA-GN, it will be very crucial to figure out whether the antigens are from certain pathogens or from the host.

We can further try to cluster TCRs based on the similarity of CDR3 sequences obtained from the single cell TCR-seq data. There are several TCR clustering algorithms developed for bulk TCR-seq analysis, such as GLIPH (Glanville *et al.*, 2017) and TCRdist (Dash *et al.*, 2017). More recently, a deep-learning based approach, DeepTCR, was presented to perform antigen-specific TCR clustering for single cell TCR-seq data (Sidhom *et al.*, 2021). Using these methods, in the future, we can try to group TCRs of common specificity from different tissues of the same donor (intra-individual) or even from the different donors (inter-individual) in autoimmune diseases. There are also attempts in the computational biology field to predict antigen specificity of single T cells based on CDR3 sequences such as TcellMatch (Fischer *et al.*, 2020). This field still needs more known TCR-antigen data from different diseases to make meaningful predictions (Joglekar and Li, 2020).

## **6.7 Therapeutically target pathogenic TRMs**

Currently, no medication has been approved by the FDA to treat NAFLD/NASH. The main strategy is losing weight by diet control and exercise. The current treatment strategies for ANCA-GN and other autoimmune kidney diseases are mainly corticosteroids and cytotoxic agents, which suppress the whole immune system (Panzer and Huber, 2021). These strategies are not specific and have severe side effects (Floege and Amann, 2016). Specific treatment strategies are required for the future investigation in autoimmune diseases. For COVID-19 severe patients, controlling cytokine storm and hyper-inflammatory responses is a key strategy for treatment. Our computational characterization of T cells from ANCA-GN, NAFLD/NASH and severe COVID-19

patients demonstrate the pathogenicity of TRMs in these inflammatory diseases. Combined together, our findings could potentially reveal new therapeutic strategies in inflammatory diseases. Here I would like to discuss the strategies of targeting pathogenic TRMs. There are several potential directions to target TRMs: inactivating or depleting TRMs directly and/or neutralizing inflammatory cytokines produced by TRMs.

### 6.7.1 Inactivate or deplete pathogenic TRMs

In the ANCA-GN project, the cGN mouse experiment model demonstrates that rapid IL-17 response of renal TRM17 cells is cytokine-driven. In other words, TRMs can be activated after induction of cGN *in vivo* by a TCR-independent mechanism. Based on the identification of this TCR-independent activation pathway, a therapeutic approach in cGN could be investigated. We can use monoclonal antibodies to neutralize the pro-inflammatory cytokines used for activation *in vivo*, thereby preventing TRMs activation. The anti-IL-1b, anti-IL-6, and anti-IL-23 cocktails in our experiment showed protective effects in the cGN mice. Monoclonal antibodies targeting these cytokines or their receptors could serve as potential treatment for ANCA-GN.

Direct depletion of TRMs at the inflamed tissue can be technically challenging. The first prerequisite is to have tissue-specific surface markers of targeted TRMs. The second prerequisite is to penetrate the tissue to access TRMs. The tissue-specific surface markers could be further derived from our current single cell RNA and CITE-seq analysis but need to be further confirmed by *in vivo* animal experiments as well as imaging staining of the human tissue. Once the surface markers are confirmed, nanobodies could be used to access the cells deep in the tissue due to their small size compared to normal antibodies (Muyldermans, 2021). Nanobodies can be fused to anti-CD16 to bring NK cells (express CD16 on the surface) to the target TRMs. This approach is based on antibody-dependent cell cytotoxicity (González *et al.*, 2019). Combination of several nanobodies with different

surface markers is also possible (Muyldermans, 2021). This idea is still very preliminary and should be explored under careful experiments in the mouse experimental models.

### 6.7.2 Target cytokines produced by TRMs

Our single cell analysis of severe COVID-19 BAL samples show pathogenic TRM17 cells with expression of GM-CSF and IL-17A. High serum GM-CSF and IL-17A levels correlates with severe clinical courses. There are several other recent studies addressing the role of GM-CSF and IL-17A, some of them reported after we published our work in February 2021.

#### 1) GM-CSF

In the study performed by Thwaites and his colleagues, published in March, 2021 (Thwaites *et al.*, 2021), they collected plasma samples from 471 hospitalized severe COVID-19 patients, 39 mild COVID-19 patients and 20 influenza patients. They found circulating GM-CSF protein level is elevated in hospitalized COVID-19 patients compared to mild patients. Within hospitalized COVID-19 patients, the GM-CSF level is associated with the disease severity score. In addition, GM-CSF levels in fatal COVID-19 patients are much higher than the fatal influenza. In contrast, the other cytokine IL-6 is high in both COVID-19 and influenza. Their study used influenza as control and our study used bacterial pneumonia as control, both suggesting the specific role of GM-CSF in COVID-19.

Another study published in July, 2021 (Kreutmair *et al.*, 2021) argued that GM-CSF producing T cells are a shared feature of both severe COVID-19 and hospital-acquired pneumonia. This study used single-cell spectral cytometry technique to investigate a selection of 50 markers in the blood of 121 COVID-19 patients and 25 hospital-acquired pneumonia patients. They stimulated the cells *in vitro* before spectral cytometry. They identified a group of TEMRA cells (CCR7<sup>-</sup>, CD45RA<sup>+</sup>) producing GM-CSF cytokines,

although this group of cells were a mixture of CD4<sup>+</sup> or CD8<sup>+</sup>. They demonstrated that the frequency of these GM-CSF positive cells positively correlates with COVID-19 severity. This finding further supports our conclusion. In addition to their own spectral cytometry analysis, this study also re-analyzed our single cell dataset of the COVID-19 patients' blood cells described in Chapter 5 (Zhao *et al.*, 2021). They specifically separated *CSF2* (the gene of GM-CSF) high and low CD4<sup>+</sup> T cells and performed differential gene expression analysis. The *CSF2* high cells also have a higher expression level of other proinflammatory genes such as *TNF*, *IL21*, *GZLY*, *CD40LG*, *CCL20*. Together with our published results, their further analysis suggests GM-CSF as a key driver of hyperinflammation in COVID-19 severe patients.

Although the mechanism of GM-CSF mediating severe COVID-19 remains elusive, the promising therapeutic target role of GM-CSF has already been strongly indicated. There are several ongoing clinical trials aiming at GM-CSF blockade in COVID-19 subjects. The most promising trial so far came from the Phase 3 trial carried out by Humanigen company using anti-GM-CSF monoclonal antibody Lenzilumab. Their results were published in May, 2021 (Temesgen *et al.*, 2021). This study collected 520 COVID-19 patients and gave Lenzilumab administration to 261 of them. The remaining 259 patients received placebo treatment. After 28 days, the results demonstrated that Lenzilumab significantly improved the likelihood of survival without ventilation in hospitalized COVID-19 patients. Subjects with the blood CRP level smaller than 150 mg/L (C-reactive protein, a marker of inflammation) and age < 85 years showed an improvement in survival and had the greatest benefit from Lenzilumab. Another GM-CSF inhibitor Namilumab was investigated in a Phase 2 study together with TNF inhibitor Infliximab. Their results published in June, 2021 demonstrated that Namilumab instead of Infliximab showed reduction of inflammation in hospitalized COVID-19 patients (Fisher *et al.*, 2021). Another earlier Phase 2 study using anti-GM-CSF monoclonal antibody Otilimab showed substantial benefit in COVID-19 patients aged > 70 (Patel *et al.*, 2021). In summary, the above trials all support the benefits of GM-CSF blockade in severe COVID-19 patients. This reinforced our conclusion about the pathogenic role of GM-CSF producing TRM cells.

## 2) IL17A

IL-17A plays a critical role in the cytokine storm of severe acute respiratory distress syndrome (Wiche Salinas *et al.*, 2020). A very recent study published in September, 2021 observed higher IL-17A levels in critically ill COVID-19 patients compared to critically ill non-COVID-19 patients and mild COVID-19 patients (Jahaj *et al.*, 2021). Interestingly, in another recent mouse study, Lin *et al.* demonstrated that the SARS-CoV-2 coding protein open reading frame 8 (ORF8) could induce lung inflammation in mice by activating IL-17 signaling pathway (Lin *et al.*, 2021). Inhibition of IL-17 pathway by antibody could decrease the ORF8-induced inflammation. These studies, together with our results, suggest the rationale of IL-17A blockade in COVID-19.

IL-17A blockade strategy has been used in other autoimmune diseases. There are two anti-IL-17A humanized monoclonal antibodies: Ixekizumab and Secukinumab, both of them have been used in psoriasis and psoriatic arthritis (Patel *et al.*, 2013). Interestingly, a retrospective observational study reported that psoriasis patients receiving IL-17A inhibitor therapy did not develop severe COVID-19 (Gisoni *et al.*, 2020).

In a very recent preprint published in July, 2021, Resende *et al.* performed a phase 2 study using the IL-17A blocker Secukinumab in hospitalized COVID-19 patients in Brazil (Resende *et al.*, 2021). This randomized trial included 50 patients, half of them received Secukinumab plus standard care and the other half only had standard care.

The clinically relevant thromboembolic events were significantly less in Secukinumab-treated patients. In severe COVID-19 patients, the prevalence of clinically relevant thromboembolic events is above 75% (Giannis *et al.*, 2021). This suggests the potential role of IL-17A in pulmonary thromboembolism. However, there was no decrease in the number of ventilator-free days at day-28 (VFD-28, the primary endpoint) in the Secukinumab group. Therefore, this study did not show significant benefits of Secukinumab in ameliorating the pulmonary failure in severe COVID-19 patients. The

other IL-17A blocker Ixekizumab is still being investigated for the treatment of COVID-19 (Liu *et al.*, 2020). No current study of IL17A blockade has been conducted in ANCA-GN or NASH yet. However, neutralizing IL-17A should be explored in the pre-clinical animal models for these diseases.

## **6.8 Conclusions and outlook**

T cells play center roles in adaptive immune responses and dysfunction of T cell regulation can cause inflammatory diseases. Fundamental advances in single cell sequencing technology enabled unbiased computational characterization of T cell subsets. Using high-dimensional single cell analysis, in this thesis we systematically explored the heterogeneity and function of T cells in different inflammatory diseases involving the kidney, liver and lung. The major focus of the study was characterization of TRMs, an important subset of T cells. The role of TRMs in inflammatory diseases is largely unexplored, especially for human patient samples. The projects provide deep insights of pathogenic TRMs in inflammatory diseases and indicate their bridging role of connecting infection and autoimmunity. The study also highlights the potential treatment strategy of targeting TRMs or their products.

Several main future directions were also discussed in detail. Firstly, the interaction of TRMs and other resident immune and non-immune cells should be dissected based on single cell sequencing data. This can be even better investigated by combining spatial transcriptomics techniques. Secondly, tissue specific characteristics of the TRMs should be further addressed. This will be crucial for tissue specific targeted therapy. Thirdly, the transcriptional features as well as the TCR sequences of clonal expanded T cells in inflammatory disease should be explored in detail. This will help to understand the disease pathogenicity and predict the potential autoantigens. Last but not least, therapeutic strategies of targeting pathogenic TRMs need to be further validated in various disease models.



## Reference

- Altmann, D.M., and Boyton, R.J. (2020). SARS-CoV-2 T cell immunity: Specificity, function, durability, and role in protection. *Sci. Immunol.* 5, e6160.
- Angileri, F. (2020). Molecular mimicry may explain multi-organ damage in COVID-19. *Autoimmun. Rev.* 8, e102591.
- Anguiano, L., Kain, R., and Anders, H.J. (2020). The glomerular crescent: Triggers, evolution, resolution, and implications for therapy. *Curr. Opin. Nephrol. Hypertens.* 29, 302–309.
- Baffy, G., Brunt, E.M., and Caldwell, S.H. (2012). Hepatocellular carcinoma in non-alcoholic fatty liver disease: An emerging menace. *J. Hepatol.* 56, 1384–1391.
- Bashiardes, S., Shapiro, H., Rozin, S., Shibolet, O., and Elinav, E. (2016). Non-alcoholic fatty liver and the gut microbiota. *Mol. Metab.* 5, 782–794.
- Bastard, P., Rosen, L.B., Zhang, Q., Michailidis, E., Hoffmann, H.H., Zhang, Y., Dorgham, K., Philippot, Q., Rosain, J., Béziat, V., *et al.* (2020). Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science.* 370, e4585.
- Battaglia, M., Gregori, S., Bacchetta, R., and Roncarolo, M.-G. (2006). Tr1 cells: From discovery to their clinical application. *Semin. Immunol.* 18, 120–127.
- Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* 37, 38–47.
- Bilzer, M., Roggel, F., and Gerbes, A.L. (2006). Role of Kupffer cells in host defense and liver disease. *Liver Int.* 26, 1175–1186.
- Blondel (2008). Fast unfolding of communities in large networks. *J. Stat. Mech* 10008.
- Boufea, K., Seth, S., and Batada, N.N. (2020). scID Uses Discriminant Analysis to Identify Transcriptionally Equivalent Cell Types across Single-Cell RNA-Seq Data with Batch Effect. *IScience* 23, 100914.
- Braga, F.A.V., Kar, G., Berg, M., Carpaij, O.A., Polanski, K., Simon, L.M., Brouwer, S., Gomes, T., Hesse, L., Jiang, J., *et al.* (2019). Cell States in Health and in Asthma. *Nat. Med.* 25.
- Brandl, K., and Schnabl, B. (2017). Intestinal microbiota and nonalcoholic steatohepatitis. *Curr. Opin. Gastroenterol.* 33, 128–133.
- Braun, J., Loyal, L., Frentsch, M., Wendisch, D., Georg, P., Kurth, F., Hippenstiel, S., Dingeldey, M., Kruse, B., Fauchere, F., *et al.* (2020). SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* 587, 270–274.
- Brennecke, P., Anders, S., Kim, J.K., Kołodziejczyk, A.A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S.A., Marioni, J.C., *et al.* (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* 10, 1093–1098.
- Bruzzì, S., Sutti, S., Giudici, G., Burlone, M.E., Ramavath, N.N., Toscani, A., Bozzola, C., Schneider, P., Morello, E., Parola, M., *et al.* (2018). B2-Lymphocyte responses to oxidative stress-derived antigens contribute to the evolution of nonalcoholic fatty liver disease (NAFLD). *Free Radic. Biol. Med.* 124, 249–259.
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating

single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* *36*, 411–420.

Chen, Z., and John Wherry, E. (2020). T cell responses in patients with COVID-19. *Nat. Rev. Immunol.* *20*, 529–536.

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., *et al.* (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* *395*, 507–513.

Cheuk, S., Schlums, H., Gallais S  rezal, I., Martini, E., Chiang, S.C., Marquardt, N., Gibbs, A., Detlofsson, E., Introini, A., Forkel, M., *et al.* (2017). CD49a Expression Defines Tissue-Resident CD8+ T Cells Poised for Cytotoxic Function in Human Skin. *Immunity* *46*, 287–300.

Christo, S.N., Evrard, M., Park, S.L., Gandolfo, L.C., Burn, T.N., Fonseca, R., Newman, D.M., Alexandre, Y.O., Collins, N., Zamudio, N.M., *et al.* (2021). Discrete tissue microenvironments instruct diversity in resident memory T cell function and plasticity. *Nat. Immunol.* *22*, 1140–1151.

Chua, R.L., Lukassen, S., Trump, S., Hennig, B.P., Wendisch, D., Pott, F., Debnath, O., Th  rmann, L., Kurth, F., V  lker, M.T., *et al.* (2020). COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. *Nat. Biotechnol.* *38*, 970–979.

Couser, W.G. (1988). Rapidly Progressive Glomerulonephritis: Classification, Pathogenetic Mechanisms, and Therapy. *Am. J. Kidney Dis.* *11*, 449–464.

Crispe, I.N. (2009). The Liver as a Lymphoid Organ. *Annu. Rev. Immunol.* *27*, 147–63.

Cui, Z., and Zhao, M.H. (2011). Advances in human anti-glomerular basement membrane disease. *Nat. Rev. Nephrol.* *7*, 697–705.

Dalekos, G.N., Gatselis, N.K., Zachou, K., and Koukoulis, G.K. (2020). NAFLD and autoimmune hepatitis: Do not judge a book by its cover. *Eur. J. Intern. Med.* *75*, 1–9.

Dash, P., Fiore-Gartland, A.J., Hertz, T., Wang, G.C., Sharma, S., Souquette, A., Crawford, J.C., Clemens, E.B., Nguyen, T.H.O., Kedzierska, K., *et al.* (2017). Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature* *547*, 89–93.

Dong, C. (2021). Cytokine Regulation and Function in T Cells. *Annu. Rev. Immunol.* *39*, 51–76.

Dudek, M., Pfister, D., Donakonda, S., Filpe, P., Schneider, A., Laschinger, M., Hartmann, D., H  ser, N., Meiser, P., Bayerl, F., *et al.* (2021). Auto-aggressive CXCR6+ CD8 T cells cause liver immune pathology in NASH. *Nature* *592*, 444–449.

Duerr, R.H., Taylor, K.D., Brant, S.R., Rioux, J.D., Silverberg, M.S., Daly, M.J., Steinhardt, A.H., Abraham, C., Regueiro, M., Griffiths, A., *et al.* (2006). A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* *314*, 1461–1463.

Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* *15*, 1484–1506.

Ercolini, A.M., and Miller, S.D. (2009). The role of infections in autoimmune disease. *Clin. Exp. Immunol.* *155*, 1–15.

Farber, D.L., Yudanin, N.A., and Restifo, N.P. (2014). Human memory T cells: Generation, compartmentalization and homeostasis. *Nat. Rev. Immunol.* *14*, 24–35.

Farhadi, A., Gundlapalli, S., Shaikh, M., Frantzides, C., Harrell, L., Kwasny, M.M., and Keshavarzian, A. (2008). Susceptibility to gut leakiness: a possible mechanism for endotoxaemia in non-alcoholic steatohepatitis. *Liver Int.* 28, 1026–1033.

Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A.K., Slichter, C.K., Miller, H.W., McElrath, M.J., Prlic, M., *et al.* (2015). MAST: A flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* 16, 1–13.

Fischer, D.S., Wu, Y., Schubert, B., and Theis, F.J. (2020). Predicting antigen specificity of single T cells based on TCR CDR 3 regions. *Mol. Syst. Biol.* 16, 1–14.

Fisher, B., Veenith, T., and Slade, D. (2021). Namilumab or infliximab compared to standard of care in hospitalised patients with COVID-19 (CATALYST): a phase 2 randomised adaptive trial (preprint). *Medrxiv* 19, 2021.06.02.21258204.

Floege, J., and Amann, K. (2016). Primary glomerulonephritides. *Lancet (London, England)* 387, 2036–2048.

Fooksman, D.R., Vardhana, S., Vasiliver-shamis, G., Liese, J., Blair, D.A., Waite, J., Sacrist, C., Victora, G.D., Zanin-zhorov, A., and Dustin, M.L. (2010). Functional Anatomy of T Cell Activation and Synapse Formation. *Annu. Rev. Immunol.* 28, 79-105.

Friedman, S.L., Neuschwander-Tetri, B.A., Rinella, M., and Sanyal, A.J. (2018). Mechanisms of NAFLD development and therapeutic strategies. *Nat. Med.* 24, 908–922.

Gagliani, N., Magnani, C.F., Huber, S., Gianolini, M.E., Pala, M., Licona-Limon, P., Guo, B., Herbert, D.R., Bulfone, A., Trentini, F., *et al.* (2013). Coexpression of CD49b and LAG-3 identifies human and mouse T regulatory type 1 cells. *Nat. Med.* 19, 739–746.

Gasteiger, G., D’osualdo, A., Schubert, D.A., Weber, A., Bruscia, E.M., and Hartl, D. (2017). Cellular Innate Immunity: An Old Game with New Players. *J. Innate Immun.* 9, 111–125.

Giannis, D., Barish, M.A., Goldin, M., Cohen, S.L., Kohn, N., Gianos, E., Chatterjee, S., Lesser, M., Coppa, K., Hirsch, J.S., *et al.* (2021). Incidence of Venous Thromboembolism and Mortality in Patients with Initial Presentation of COVID-19. *J. Thromb. Thrombolysis* 51, 897–901.

Giladi, A., Cohen, M., Medaglia, C., Baran, Y., Li, B., Zada, M., Bost, P., Blecher-Gonen, R., Salame, T.M., Mayer, J.U., *et al.* (2020). Dissecting cellular crosstalk by sequencing physically interacting cells. *Nat. Biotechnol.* 38, 629–637.

Gisoni, P., Facheris, P., Dapavo, P., Piaserico, S., Conti, A., Naldi, L., Cazzaniga, S., Malagoli, P., and Costanzo, A. (2020). The impact of the COVID-19 pandemic on patients with chronic plaque psoriasis being treated with biological therapy: the Northern Italy experience. *Br. J. Dermatol.* 183, 373–374.

Glanville, J., Huang, H., Nau, A., Hatton, O., Wagar, L.E., Rubelt, F., Ji, X., Han, A., Krams, S.M., Pettus, C., *et al.* (2017). Identifying specificity groups in the T cell receptor repertoire. *Nature* 547, 94–98.

González, C., Chames, P., Kerfelec, B., Baty, D., Robert, P., and Limozin, L. (2019). Nanobody-CD16 Catch Bond Reveals NK Cell Mechanosensitivity. *Biophys. J.* 116, 1516–1526.

Gourley, T.S., Wherry, E.J., Masopust, D., and Ahmed, R. (2004). Generation and maintenance of immunological memory. *Semin. Immunol.* 16, 323–333.

Grifoni, A., Weiskopf, D., Ramirez, S.I., Mateus, J., Dan, J.M., Moderbacher, C.R., Rawlings, S.A., Sutherland, A., Premkumar, L., Jardi, R.S., *et al.* (2020). Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* *181*, 1489-1501.e15.

Guilherme, L., Cunha-Neto, E., Coelho, V., Snitcowsky, R., Pomerantzeff, P.M., Assis, R. V, Pedra, F., Neumann, J., Goldberg, A., and Patarroyo, M.E. (1995). Human heart-infiltrating T-cell clones from rheumatic heart disease patients recognize both streptococcal and cardiac proteins. *Circulation* *92*, 415–420.

Gupta, M., and Weaver, D.F. (2021). COVID-19 as a Trigger of Brain Autoimmunity. *ACS Chem. Neurosci.* *12*, 2558–2561.

Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* *36*, 421–427.

Halpert, G., and Shoenfeld, Y. (2020). SARS-CoV-2, the autoimmune virus. *Autoimmun. Rev.* *19*, 102695.

Harrington, L.E., Hatton, R.D., Mangan, P.R., Turner, H., Murphy, T.L., Murphy, K.M., and Weaver, C.T. (2005). Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nat. Immunol.* *6*, 1123–1132.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A., *et al.* (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* *181*, 271-280.e8.

Hogquist, K.A., Baldwin, T.A., and Jameson, S.C. (2005). Central tolerance: Learning self-control in the thymus. *Nat. Rev. Immunol.* *5*, 772–782.

Huang, B., Chen, Z., Geng, L., Wang, J., Liang, H., Cao, Y., Chen, H., Huang, W., Su, M., Wang, H., *et al.* (2019). Mucosal Profiling of Pediatric-Onset Colitis and IBD Reveals Common Pathogenics and Therapeutic Pathways. *Cell* *179*, 1160-1176.e24.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., *et al.* (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* *395*, 497–506.

Huang, W., Na, L., Fidel, P.L., and Schwarzenberger, P. (2004). Requirement of Interleukin-17A for Systemic Anti-Candida albicans Host Defense in Mice. *J. Infect. Dis.* *190*, 624–631.

Izci Duran, T., Turkmen, E., Dilek, M., Sayarlioglu, H., and Arik, N. (2021). ANCA-associated vasculitis after COVID-19. *Rheumatol. Int.* *41*, 1523–1529.

Jahaj, E., Vassiliou, A.G., Keskinidou, C., Gallos, P., Vrettou, C.S., Tsipilis, S., Mastora, Z., Orfanos, S.E., Dimopoulou, I., and Kotanidou, A. (2021). Evaluating the role of the interleukin-23/17 axis in critically ILL COVID-19 patients. *J. Pers. Med.* *11*.

Jarjour, N.N., Masopust, D., and Jameson, S.C. (2021). T Cell Memory: Understanding COVID-19. *Immunity* *54*, 14–18.

Jennette, J.C., and Falk, R.J. (2014). Pathogenesis of antineutrophil cytoplasmic autoantibody-mediated disease. *Nat. Rev. Rheumatol.* *10*, 463–473.

Jennette, J.C., and Nachman, P.H. (2017). ANCA glomerulonephritis and vasculitis. *Clin. J. Am. Soc. Nephrol.* *12*, 1680–1691.

Jin, S., Guerrero-Juarez, C.F., Zhang, L., Chang, I., Ramos, R., Kuan, C.H., Myung, P., Plikus, M. V., and Nie, Q. (2021). Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* *12*, 1–20.

Joglekar, A. V, and Li, G. (2020). T cell antigen discovery. *Nat. Methods.* *18*, 873–880.

Kamran, P., Sereti, K.-I., Zhao, P., Ali, S.R., Weissman, I.L., and Ardehali, R. (2013). Parabiosis in Mice: A Detailed Protocol. *J. Vis. Exp.* *80*, 50556.

Kazankov, K., Jørgensen, S.M.D., Thomsen, K.L., Møller, H.J., Vilstrup, H., George, J., Schuppan, D., and Grønbaek, H. (2019). The role of macrophages in nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Nat. Rev. Gastroenterol. Hepatol.* *16*, 145–159.

King, C., Tangye, S.G., and Mackay, C.R. (2010). T Follicular Helper ( T FH ) Cells in Normal and Dysregulated Immune Responses. *Annu. Rev. Immunol.* *26*, 741–66.

Kiselev, V.Y., Kirschner, K., Schaub, M.T., Andrews, T., Yiu, A., Chandra, T., Natarajan, K.N., Reik, W., Barahona, M., Green, A.R., *et al.* (2017). SC3: Consensus clustering of single-cell RNA-seq data. *Nat. Methods* *14*, 483–486.

Kiselev, V.Y., Yiu, A., and Hemberg, M. (2018). Scmap: Projection of single-cell RNA-seq data across data sets. *Nat. Methods* *15*, 359–362.

Kiselev, V.Y., Andrews, T.S., and Hemberg, M. (2019). Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat. Rev. Genet.* *20*, 273–282.

Kleinschek, M.A., Boniface, K., Sadekova, S., Grein, J., Murphy, E.E., Turner, S.P., Raskin, L., Desai, B., Faubion, W.A., De Malefyt, R.W., *et al.* (2009). Circulating and gut-resident human Th17 cells express CD161 and promote intestinal inflammation. *J. Exp. Med.* *206*, 525–534.

Korn, T., Bettelli, E., Oukka, M., and Kuchroo, V.K. (2009). IL-17 and Th17 cells. *Annu. Rev. Immunol.* *27*, 485–517.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P. ru, and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* *16*, 1289–1296.

Koyama, Y., and Brenner, D.A. (2017). Liver inflammation and fibrosis. *J. Clin. Invest.* *127*, 55–64.

Kozlenkov, A., Li, J., Apontes, P., Hurd, Y.L., Byne, W.M., Koonin, E. V., Wegner, M., Mukamel, E.A., and Dracheva, S. (2018). A unique role for DNA (hydroxy)methylation in epigenetic regulation of human inhibitory neurons. *Sci. Adv.* *4*.

Krebs, C.F., Paust, H.J., Krohn, S., Koyro, T., Brix, S.R., Riedel, J.H., Bartsch, P., Wiech, T., Meyer-Schwesinger, C., Huang, J., *et al.* (2016). Autoimmune Renal Disease Is Exacerbated by S1P-Receptor-1-Dependent Intestinal Th17 Cell Migration to the Kidney. *Immunity* *45*, 1078–1092.

Krebs, C.F., Schmidt, T., Riedel, J.H., and Panzer, U. (2017). T helper type 17 cells in immune-mediated glomerular disease. *Nat. Rev. Nephrol.* *13*, 647–659.

Kreutmair, S., Unger, S., Núñez, N.G., Ingelfinger, F., Alberti, C., De Feo, D., Krishnarajah, S., Kauffmann, M., Friebel, E., Babaei, S., *et al.* (2021). Distinct immunological signatures discriminate severe COVID-19 from non-SARS-CoV-2-driven critical pneumonia. *Immunity* *54*, 1578-1593.e5.

Kubes, P., and Jenne, C. (2018). Immune Responses in the Liver. *Annu. Rev. Immunol.* *36*,

247–77.

Kuleshov, M. V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., *et al.* (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* *44*, 90–97.

Kumar, B. V., Ma, W., Miron, M., Granot, T., Guyer, R.S., Carpenter, D.J., Senda, T., Sun, X., Ho, S.H., Lerner, H., *et al.* (2017). Human Tissue-Resident Memory T Cells Are Defined by Core Transcriptional and Functional Signatures in Lymphoid and Mucosal Sites. *Cell Rep.* *20*, 2921–2934.

Kumar, B. V., Connors, T.J., and Farber, D.L. (2018). Human T Cell Development, Localization, and Function throughout Life. *Immunity* *48*, 202–213.

Liew, F.Y. (2002). *Nat Rev Immunol.* *2*, 55–60.

Lin, X., Fu, B., Yin, S., Li, Z., Liu, H., Zhang, H., Xing, N., Wang, Y., Xue, W., Xiong, Y., *et al.* (2021). ORF8 contributes to cytokine storm during SARS-CoV-2 infection by activating IL-17 pathway. *IScience* *24*, 102293.

Lipsitch, M., Grad, Y.H., Sette, A., and Crotty, S. (2020). Cross-reactive memory T cells and herd immunity to SARS-CoV-2. *Nat. Rev. Immunol.* *20*, 709–713.

Liu, P., Huang, Z., Yin, M., Liu, C., Chen, X., Pan, P., and Kuang, Y. (2020). Safety and Efficacy of Ixekizumab and Antiviral Treatment for Patients with COVID-19: A structured summary of a study protocol for a Pilot Randomized Controlled Trial. *Trials* *21*, 20–22.

Liu, Y., Helms, C., Liao, W., Zaba, L.C., Duan, S., Gardner, J., Wise, C., Miner, A., Malloy, M.J., Pullinger, C.R., *et al.* (2008). A Genome-Wide Association Study of Psoriasis and Psoriatic Arthritis Identifies New Disease Loci. *PLOS Genet.* *4*, e1000041.

Liu, Y., Sawalha, A.H., and Lu, Q. (2021). COVID-19 and autoimmune diseases. *Curr. Opin. Rheumatol.* *33*, 155–162.

Longo, S.K., Guo, M.G., Ji, A.L., and Khavari, P.A. (2021). Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* *22*, 627–644.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* *15*, 1–21.

Luecken, M.D., and Theis, F.J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.* *15*.

van der Maaten, L., and Hinton, G. (2008). Visualizing Data using t-SNE. *J. Mach. Learn. Res.* *9*, 2579–2605.

Mackay, L.K., and Kallies, A. (2017). Transcriptional Regulation of Tissue-Resident Lymphocytes. *Trends Immunol.* *38*, 94–103.

Mangalmurti, N., and Hunter, C.A. (2020). Cytokine Storms: Understanding COVID-19. *Immunity* 1–7.

Mann, E.A., and Saeed, S.A. (2012). Gastrointestinal infection as a trigger for inflammatory bowel disease. *Curr. Opin. Gastroenterol.* *28*, 24–29.

Martinez Valenzuela, L., Bordignon Draibe, J., Fulladosa Oliveras, X., Bestard Matamoros, O., Cruzado Garrit, J.M., and Torras Ambrós, J. (2019). T-lymphocyte in ANCA-associated vasculitis: What do we know? A pathophysiological and therapeutic approach. *Clin. Kidney J.* *12*, 503–511.

Masopust, D., and Soerens, A.G. (2019). Tissue-Resident T Cells and Other Resident Leukocytes. *Annu. Rev. Immunol.* *37*, 521–546.

Matthay, M.A., Aldrich, J.M., and Gotts, J.E. (2020). Treatment for severe acute respiratory distress syndrome from COVID-19. *Lancet Respir. Med.* *8*, 433–434.

Mayadas, T.N., Tsokos, G.C., and Tsuboi, N. (2009). Mechanisms of immune complex-mediated neutrophil recruitment and tissue injury. *Circulation* *120*, 2012–2024.

Mayet, W.J., Märker-Hermann, E., Schlaak, J., and Meyer Zum Büschenfelde, K.H. (1999). Irregular cytokine pattern of CD4+ T lymphocytes in response to *Staphylococcus aureus* in patients with Wegener’s granulomatosis. *Scand. J. Immunol.* *49*, 585–594.

McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* *8*, 329–337.e4.

McInnes, L., Healy, J., and Melville, J. (2020). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. [arXiv:1802.03426](https://arxiv.org/abs/1802.03426).

Meckiff, B.J., Ramírez-Suástegui, C., Fajardo, V., Chee, S.J., Kusnadi, A., Simon, H., Eschweiler, S., Grifoni, A., Pelosi, E., Weiskopf, D., *et al.* (2020). Imbalance of Regulatory and Cytotoxic SARS-CoV-2-Reactive CD4+ T Cells in COVID-19. *Cell* *183*, 1340–1353.e16.

Miragaia, R.J., Gomes, T., Chomka, A., Jardine, L., Riedel, A., Hegazy, A.N., Whibley, N., Tucci, A., Chen, X., Lindeman, I., *et al.* (2019). Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation. *Immunity* *50*, 493–504.e7.

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

Mueller, S.N., and Mackay, L.K. (2016). Tissue-resident memory T cells: Local specialists in immune defence. *Nat. Rev. Immunol.* *16*, 79–89.

Muyldermans, S. (2021). Applications of Nanobodies. *Annu. Rev. Anim. Biosci.* *9*, 401–421.

Netea, M.G., Domínguez-Andrés, J., Barreiro, L.B., Chavakis, T., Divangahi, M., Fuchs, E., Joosten, L.A.B., van der Meer, J.W.M., Mhlanga, M.M., Mulder, W.J.M., *et al.* (2020). Defining trained immunity and its role in health and disease. *Nat. Rev. Immunol.* *20*, 375–388.

Noël, F., Massenet-Regad, L., Carmi-Levy, I., Cappuccio, A., Grandclaude, M., Trichot, C., Kieffer, Y., Mechta-Grigoriou, F., and Soumelis, V. (2021). Dissection of intercellular communication using the transcriptome-based framework ICELLNET. *Nat. Commun.* *12*, 1–16.

Oller-Moreno, S., Kloiber, K., Machart, P., and Bonn, S. (2021). Algorithmic advances in Machine Learning for single cell expression analysis. *Curr. Opin. Syst. Biol.*

Pacheco, Y., Acosta-Ampudia, Y., Monsalve, D.M., Chang, C., Gershwin, M.E., and Anaya, J.M. (2019). Bystander activation and autoimmunity. *J. Autoimmun.* *103*, 102301.

Panzer, U., and Huber, T.B. (2021). Immune-mediated glomerular diseases: new basic concepts and clinical implications. *Cell Tissue Res.* 2–4.

Papalex, E., and Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* *18*, 35–45.

Pappalardo, J.L., Zhang, L., Pecsok, M.K., Perlman, K., Zografou, C., Raddassi, K., Abulaban, A., Krishnaswamy, S., Antel, J., van Dijk, D., *et al.* (2020). Transcriptomic and clonal characterization of T cells in the human central nervous system. *Sci. Immunol.* *5*, e8786.

Park, H., Li, Z., Yang, X.O., Chang, S.H., Nurieva, R., Wang, Y.H., Wang, Y., Hood, L., Zhu, Z., Tian, Q., *et al.* (2005). A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. *Nat. Immunol.* *6*, 1133–1141.

Pasquini, G., Rojo Arias, J.E., Schäfer, P., and Busskamp, V. (2021). Automated methods for cell type annotation on scRNA-seq data. *Comput. Struct. Biotechnol. J.* *19*, 961–969.

Patel, D.D., Lee, D.M., Kolbinger, F., and Antoni, C. (2013). Effect of IL-17A blockade with secukinumab in autoimmune diseases. *Ann. Rheum. Dis.* *72 Suppl 2*, ii116-23.

Patel, J., Beishuizen, A., Ruiz, X.B., Boughanmi, H., Cahn, A., Criner, G.J., Davy, K., de-Miguel-Díez, J., Fernandes, S., François, B., *et al.* (2021). A Randomized Trial of Otilimab in Severe COVID-19 Pneumonia (OSCAR). *MedRxiv* 2021.04.14.21255475.

Pliner, H.A., Shendure, J., and Trapnell, C. (2019). Supervised classification enables rapid annotation of cell atlases. *Nat. Methods* *16*, 983–986.

Powell, W.T., Campbell, J.A., Ross, F., Jiménez, P.P., Rudzinski, E.R., and Dickerson, J.A. (2021). Acute ANCA vasculitis and asymptomatic COVID-19. *Pediatrics* *147*.

Qi, Q., Liu, Y., Cheng, Y., Glanville, J., Zhang, D., Lee, J.Y., Olshen, R.A., Weyand, C.M., Boyd, S.D., and Goronzy, J.J. (2014). Diversity and clonal selection in the human T-cell repertoire. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 13139–13144.

Qin, C., Zhou, L., Hu, Z., Zhang, S., Yang, S., Tao, Y., Xie, C., Ma, K., Shang, K., Wang, W., *et al.* (2020). Dysregulation of immune response in patients with coronavirus 2019 (COVID-19) in Wuhan, China. *Clin. Infect. Dis.* *71*, 762–768.

Qiu, C.C., Caricchio, R., and Gallucci, S. (2019). Triggers of Autoimmunity: The Role of Bacterial Infections in the Extracellular Exposure of Lupus Nuclear Autoantigens. *Front. Immunol.* *10*, 1–15.

Raffin, C., Vo, L.T., and Bluestone, J.A. (2020). Treg cell-based therapies: challenges and perspectives. *Nat. Rev. Immunol.* *20*, 158–172.

Ramachandran, P., Dobie, R., Wilson-Kanamori, J.R., Dora, E.F., Henderson, B.E.P., Luu, N.T., Portman, J.R., Matchett, K.P., Brice, M., Marwick, J.A., *et al.* (2019). Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* *575*, 512-518.

Rau, M., Schilling, A.-K., Meertens, J., Hering, I., Weiss, J., Jurowich, C., Kudlich, T., Hermanns, H.M., Bantel, H., Beyersdorf, N., *et al.* (2016). Progression from Nonalcoholic Fatty Liver to Nonalcoholic Steatohepatitis Is Marked by a Higher Frequency of Th17 Cells in the Liver and an Increased Th17/Resting Regulatory T Cell Ratio in Peripheral Blood and in the Liver. *J. Immunol.* *196*, 97–105.

Resende, G.G., Lage, R. da C., Lobê, S.Q., Medeiros, A.F., Silva, A.D.C. e, Sá, A.T.N., Oliveira, A.J. de A., Sousa, D., Guimarães, H.C., Gomes, I.C., *et al.* (2021). Blockade of Interleukin Seventeen (IL-17A) with Secukinumab in Hospitalized COVID-19 patients – the BISHOP study. *MedRxiv* 28, 2021.07.21.21260963.

Ribot, J.C., Lopes, N., and Silva-Santos, B. (2021).  $\gamma\delta$  T cells in tissue physiology and surveillance. *Nat. Rev. Immunol.* *21*, 221–232.

Rinella, M.E. (2015). Nonalcoholic fatty liver disease a systematic review. *JAMA - J. Am.*

Med. Assoc. 313, 2263–2273.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.

Rodrigues, S.G., Stickels, R.R., Goeva, A., Martin, C.A., Murray, E., Vanderburg, C.R., Welch, J., Chen, L.M., Chen, F., and Macosko, E.Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 363, 1463–1467.

Roncarolo, M.G., Gregori, S., Bacchetta, R., and Battaglia, M. (2014). Tr1 cells and the counter-regulation of immunity: Natural mechanisms and therapeutic applications. *Curr. Top. Microbiol. Immunol.* 380, 39–68.

Roncarolo, M.G., Gregori, S., Bacchetta, R., Battaglia, M., and Gagliani, N. (2018). The Biology of T Regulatory Type 1 Cells and Their Therapeutic Application in Immune-Mediated Diseases. *Immunity* 49, 1004–1019.

Root-Bernstein, R., and Fairweather, D. (2014). Complexities in the relationship between infection and autoimmunity. *Curr. Allergy Asthma Rep.* 14, 407.

Rosato, P.C., Beura, L.K., and Masopust, D. (2017). Tissue resident memory T cells and viral immunity. *Curr. Opin. Virol.* 22, 44–50.

Ryan, G.E., Harris, J.E., and Richmond, J.M. (2021). Resident Memory T Cells in Autoimmune Skin Diseases. *Front. Immunol.* 12, 14–16.

Saelens, W., Cannoodt, R., Todorov, H., and Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* 37, 547–554.

Sasaki, K., Bean, A., Shah, S., Schutten, E., Huseby, P.G., Peters, B., Shen, Z.T., Vanguri, V., Liggitt, D., and Huseby, E.S. (2014). Relapsing–Remitting Central Nervous System Autoimmunity Mediated by GFAP-Specific CD8 T Cells. *J. Immunol.* 192, 3029–3042.

Sasson, S.C., Gordon, C.L., Christo, S.N., Klenerman, P., and Mackay, L.K. (2020). Local heroes or villains: tissue-resident memory T cells in human health and disease. *Cell. Mol. Immunol.* 17, 113–122.

Savas, P., Virassamy, B., Ye, C., Salim, A., Mintoff, C.P., Caramia, F., Salgado, R., Byrne, D.J., Teo, Z.L., Dushyanthen, S., *et al.* (2018). Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat. Med.* 24, 986–993.

Schultheiß, C., Paschold, L., Simnica, D., Mohme, M., Willscher, E., Wenserski, L. von, Scholz, R., Wieters, I., Dahlke, C., Tolosa, E., *et al.* (2020). Next Generation Sequencing of T and B cell receptor repertoires from COVID-19 patients showed signatures associated with severity of disease. *Immunity* 53, 442–455.

Schuster, S., Cabrera, D., Arrese, M., and Feldstein, A.E. (2018). Triggering and resolution of inflammation in NASH. *Nat. Rev. Gastroenterol. Hepatol.* 15, 349–364.

Sidhom, J.W., Larman, H.B., Pardoll, D.M., and Baras, A.S. (2021). DeepTCR is a deep learning framework for revealing sequence concepts within T-cell repertoires. *Nat. Commun.* 12, 1–12.

Simnica, D., Schultheiß, C., Mohme, M., Paschold, L., Willscher, E., Matschke, J., Ciesek, S., Sedding, D.G., Fitzek, A., and Klaus, P. (2021). Landscape of T-cell repertoires with public COVID-19-associated T-cell receptors in pre-pandemic risk cohorts. *Immunity* 53, 1–16.

- Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., *et al.* (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* (80- ). 353, 78–82.
- Steelman, A.J. (2015). Infection as an environmental trigger of multiple sclerosis disease exacerbation. *Front. Immunol.* 6.
- Stegeman, C.A., Cohen Tervaert, J.W., Sluiter, W.J., Manson, W.L., De Jong, P.E., and Kallenberg, C.G.M. (1994). Association of chronic nasal carriage of *Staphylococcus aureus* and higher relapse rates in Wegener granulomatosis. *Ann. Intern. Med.* 120, 12–17.
- Stewart, B.J., Ferdinand, J.R., Young, M.D., Mitchell, T.J., Loudon, K.W., Riding, A.M., Richoz, N., Frazer, G.L., Staniforth, J.U.L., Braga, F.A.V., *et al.* (2019). Spatiotemporal immune zonation of the human kidney. *Science* (80- ). 365, 1461–1466.
- Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* 19, 1–16.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902.e21.
- Suárez-Fueyo, A., Bradley, S.J., and Tsokos, G.C. (2016). T cells in Systemic Lupus Erythematosus. *Curr. Opin. Immunol.* 43, 32–38.
- Suárez-Fueyo, A., Bradley, S.J., Klatzmann, D., and Tsokos, G.C. (2017). T cells and autoimmune kidney disease. *Nat. Rev. Nephrol.* 13, 329–343.
- Sutti, S., and Albano, E. (2020). Adaptive immunity: an emerging player in the progression of NAFLD. *Nat. Rev. Gastroenterol. Hepatol.* 17, 81–92.
- Szabo, P.A., Miron, M., and Farber, D.L. (2019). Location, location, location: Tissue resident memory T cells in mice and humans. *Sci. Immunol.* 4, 9673.
- Tan, M., Liu, Y., Zhou, R., Deng, X., Li, F., Liang, K., and Shi, Y. (2020). Immunopathological characteristics of coronavirus disease 2019 cases in Guangzhou, China. *Immunology* 160, 261–268.
- Tay, M.Z., Poh, C.M., Rénia, L., MacAry, P.A., and Ng, L.F.P. (2020). The trinity of COVID-19: immunity, inflammation and intervention. *Nat. Rev. Immunol.* 20, 363–374.
- Temesgen, Z., Burger, C.D., Baker, J., Polk, C., Libertin, C., Kelley, C., Marconi, V.C., Orenstein, R., Durrant, C., Chappell, D., *et al.* (2021). Lenzilumab Efficacy and Safety in Newly Hospitalized Covid-19 Subjects: Results From the Live-Air Phase 3 Randomized Double-Blind Placebo-Controlled Trial. *MedRxiv Prepr. Serv. Heal. Sci.* 1–36.
- Teuwen, L.A., Geldhof, V., Pasut, A., and Carmeliet, P. (2020). COVID-19: the vasculature unleashed. *Nat. Rev. Immunol.* 20, 389–391.
- Thwaites, R.S., Uruchurtu, A.S.S., Siggins, M.K., Liew, F., Russell, C.D., Moore, S.C., Fairfield, C., Carter, E., Abrams, S., Short, C.E., *et al.* (2021). Inflammatory profiles across the spectrum of disease reveal a distinct role for GM-CSF in severe COVID-19. *Sci. Immunol.* 6.
- Tilg, H., and Moschen, A.R. (2010). Evolution of inflammation in nonalcoholic fatty liver disease: The multiple parallel hits hypothesis. *Hepatology* 52, 1836–1846.
- Tipping, P.G., and Holdsworth, S.R. (2006). T cells in crescentic glomerulonephritis. *J. Am.*

Soc. Nephrol. *17*, 1253–1263.

Toscano, G., Palmerini, F., Ravaglia, S., Ruiz, L., Invernizzi, P., Cuzzoni, M.G., Franciotta, D., Baldanti, F., Daturi, R., Postorino, P., *et al.* (2020). Guillain-Barré Syndrome Associated with SARS-CoV-2. *N. Engl. J. Med.* *382*, 2574–2576.

Townes, F.W., Hicks, S.C., Aryee, M.J., and Irizarry, R.A. (2019). Feature selection and dimension reduction for single-cell RNA-Seq based on a multinomial model. *Genome Biol.* *20*, 1–16.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* *32*, 381–386.

Tsuneyama, K., Baba, H., Kikuchi, K., Nishida, T., Nomoto, K., Hayashi, S., Miwa, S., Nakajima, T., Nakanishi, Y., Masuda, S., *et al.* (2013). Autoimmune features in metabolic liver disease: A single-center experience and review of the literature. *Clin. Rev. Allergy Immunol.* *45*, 143–148.

Vallejos, C.A., Risso, D., Scialdone, A., Dudoit, S., and Marioni, J.C. (2017). Normalizing single-cell RNA sequencing data: Challenges and opportunities. *Nat. Methods* *14*, 565–571.

Verdoni, L., Mazza, A., Gervasoni, A., Martelli, L., Ruggeri, M., Ciuffreda, M., Bonanomi, E., and D'Antiga, L. (2020). An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *Lancet* *395*, 1771–1778.

Vinuesa, C.G., Linterman, M.A., Yu, D., and MacLennan, I.C.M. (2016). Follicular Helper T Cells. *Annu. Rev. Immunol.* *34*, 335–368.

Vuppalanchi, R., Gould, R.J., Wilson, L.A., Unalp-Arida, A., Cummings, O.W., Chalasani, N., and Kowdley, K. V. (2012). Clinical significance of serum autoantibodies in patients with NAFLD: Results from the nonalcoholic steatohepatitis clinical research network. *Hepatology* *6*, 379–385.

Wang, B., Zhu, J., Pierson, E., Ramazzotti, D., and Batzoglou, S. (2017a). Visualization and analysis of single-cell rna-seq data by kernel-based similarity learning. *Nat. Methods* *14*, 414–416.

Wang, E.Y., Mao, T., Klein, J., Dai, Y., Huck, J.D., Jaycox, J.R., Liu, F., Zhou, T., Israelow, B., Wong, P., *et al.* (2021). Diverse functional autoantibodies in patients with COVID-19. *Nature* *595*, 283–288.

Wang, J., Vasaikar, S., Shi, Z., Greer, M., and Zhang, B. (2017b). WebGestalt 2017: A more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* *45*, 130–137.

Wiche Salinas, T.R., Zheng, B., Routy, J.-P., and Ancuta, P. (2020). Targeting the interleukin-17 pathway to prevent acute respiratory distress syndrome associated with SARS-CoV-2 infection. *Respirology* *25*, 797–799.

Wiest, R., Albillos, A., Trauner, M., Bajaj, J.S., and Jalan, R. (2017). Targeting the gut-liver axis in liver disease. *J. Hepatology* *67*, 1084–1103.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY : large-scale single-cell gene expression data analysis. *Genome Biology* *19*, 15-19.

- Wolf, F.A., Hamey, F.K., Plass, M., Solana, J., Dahlin, J.S., Göttgens, B., Rajewsky, N., Simon, L., and Theis, F.J. (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* *20*, 1–9.
- Wu, H., Liao, W., Li, Q., Long, H., Yin, H., Zhao, M., Chan, V., Lau, C.-S., and Lu, Q. (2018). Pathogenic role of tissue-resident memory T cells in autoimmune diseases. *Autoimmun. Rev.* *17*, 906–911.
- Xu, Z., Shi, L., Wang, Y., Zhang, J., Huang, L., Zhang, C., Liu, S., Zhao, P., Liu, H., Zhu, L., *et al.* (2020). Pathological findings of COVID-19 associated with acute respiratory distress syndrome. *Lancet Respir. Med.* *8*, 420–422.
- Ye, P., Rodriguez, F.H., Kanaly, S., Stocking, K.L., Schurr, J., Schwarzenberger, P., Oliver, P., Huang, W., Zhang, P., Zhang, J., *et al.* (2001). Requirement of interleukin 17 receptor signaling for lung CXC chemokine and granulocyte colony-stimulating factor expression, neutrophil recruitment, and host defense. *J. Exp. Med.* *194*, 519–527.
- Zhao, Y., Kilian, C., Turner, J.E., Bosurgi, L., Roedl, K., Bartsch, P., Gnirck, A.C., Cortesi, F., Schultheiß, C., Hellmig, M., *et al.* (2021). Clonal expansion and activation of tissue-resident memory-like Th17 cells expressing GM-CSF in the lungs of severe COVID-19 patients. *Sci. Immunol.* *6*, e6692.
- Zulfiqar, A.-A., Lorenzo-Villalba, N., Hassler, P., and Andrès, E. (2020). Immune Thrombocytopenic Purpura in a Patient with Covid-19. *N. Engl. J. Med.* *382*, e43.

## Acknowledgements

First, I would like to express my deepest gratitude to my supervisor Prof. Stefan Bonn. I moved together with your group from Göttingen to Hamburg in 2017 and started my bioinformatics training during the master thesis project in the group. My PhD started shortly after the master thesis in 2018. During the past 4 years, you motivated me with your great academic enthusiasm and rigorous research attitude. You also taught me how to communicate with collaborators and how to handle the pressure in academia. I want to thank my joint supervisor Prof. Ulf Panzer. Thank you for giving me the opportunities to engage in different research directions in the SFB1192 consortium. I am thankful to the collaborating group leaders including Dr. Christian Krebs, Prof. Nicola Gagniali, and Prof. Hans-Willi Mittrücker. You offered me valuable opportunities to participate in various exciting projects. More importantly, you shared a lot of knowledge and experience in the immunology research field with me. I would also like to thank my co-supervisor Prof. Thomas Oertner. Thank you for your support to my PhD administration and thesis.

Many, many thanks to all members of the Bonn lab. First, I would mention Dr. Vikas Bansal, who supervised my master thesis. Your guidance was fundamentally important for my PhD projects. Then I would like to thank Dr. Sergio Oller, who gave me many help in data visualization and IT management, especially during the COVID-19 pandemic time. I would like to thank Ting and Can for all the discussion about single cell analysis and biological insights. In addition, I would thank Fabian for supporting the analysis of the COVID-19 project. I am also grateful to Dr. Sonja Hänzelmann for helping to proofread my thesis draft. Apart from the scientific support, I would also like to express sincere gratitude to all my colleagues who supported my life and shared all the great moments together over these years. In particular, I would thank Sabine, Marina, Emma, Marouf, Pierre, Raza, Anna, Sven, Hannes, Tianran and Robin. We had a wonderful time together and I will be grateful for my entire life.

I am very fortunate to work with many collaborators from different groups. I would like to thank Daniel from the Mittrücker group; Hajo and Ning from the Panzer group; Alina, Christoph, and Malte from the Krebs group; Anna and Pasquale from the Gagliani group. Your tremendous efforts in wet lab experiments provided me with all the sequencing data and post-analysis validation. A special thanks to all the other co-authors of my publications. Without all your dedication, these research achievements would have been impossible.

Last but not least, I would like to say thank you to all my family and friends. Thank you for your love, support and encouragement all the time.

## List of publications

**Zhao, Y.**, Kilian, C., Turner, J.E., Bosurgi, L., Roedl, K., ..., Bonn, S., Panzer, U., Gagliani N., Krebs, C.F. (2021). Clonal expansion and activation of tissue-resident memory-like Th17 cells expressing GM-CSF in the lungs of severe COVID-19 patients. *Science Immunology*. 6, e6692.

**Zhao, Y.**, Panzer, U., Bonn, S., Krebs, C.F. (2021). Single-cell biology to decode the immune cellular composition of kidney inflammation. *Cell Tissue Research*. 385, 435-443.

Krebs, C.F., Reimers, D., **Zhao, Y.**, Paust, H.J., ..., Bonn, S., Panzer, U., Mittrücker, H.M. (2020). Pathogen-induced tissue-resident memory TH17 (TRM17) cells amplify autoimmune kidney disease. *Science Immunology*. 5, e4163.

Woestemeier, A.\*, Scognamiglio, P.\*, **Zhao, Y.\***, Wagner, J., Muscate, F., ..., Bonn, S., Huber, S., Gagliani N. Multi cytokine producing liver CD4+ T cells characterize fibrosis in NASH patients. *JCI insights*. Manuscript in revision. (\* equal contribution)

Liu, S.\*, Lu, S.\*, **Zhao, Y.\***, Lindenmeyer, M.T., Nair, V., ..., Bonn, S., Huber, T.B. Single-cell transcriptome profiling of early diabetic nephropathy highlights mechanoresponsive signaling as a driver of glomerulopathy. Manuscript prepared. (\* equal contribution)

Solagna, F., Tezze, C., Lindenmeyer, M.T., Lu, S., Wu, G., Liu, S., **Zhao, Y.**, ..., Puelles, V.G., Patel, K., Huber, T.B. (2021). Pro-cachectic factors link experimental and human chronic kidney disease to skeletal muscle wasting programs. *Journal of Clinical Investigation*. 131, e135821.

Berghoff, S.A., Spieth, L., Sun, T., Hosang, L., Depp, C., Sasmita, A.O., Vasileva, M.H., Scholz, P., **Zhao, Y.**, ..., Edgar, J.M., Saher, G. (2021). Neuronal cholesterol synthesis is essential for repair of chronically demyelinated lesions in mice. *Cell Reports*. 37, e109889.

Simnica, D., Schultheiß, C., Mohme, M., Paschold, L., Willscher, E., Matschke, J., Ciesek, S., Sedding, D.G., Fitzek, A., Klaus, P., Matschke, J., Ciesek, S., Sedding, D.G., **Zhao, Y.**, ..., Schulze-zur-Wiesch, J., Binder, M. (2021). Landscape of T-cell repertoires with public COVID-19-associated T-cell receptors in pre-pandemic risk cohorts. *Clinical & Translational Immunology*. 10, e1340.



## Abbreviations

AIH	autoimmune hepatitis
ANCA-GN	anti-neutrophil cytoplasmic antibody associated glomerulonephritis
APC	antigen presenting cell
BCR	B cell receptor
CCA	canonical correlation analysis
cGN	crescentic glomerulonephritis
CITE-seq	cellular indexing of transcriptomes and epitopes by sequencing
COVID-19	Coronavirus disease 2019
DAMP	damage-associated molecular pattern
DC	dendritic cell
DEG	differential gene expression
GN	glomerulonephritis
HVG	highly variable gene
IBD	inflammatory bowel disease
ILC	innate lymphoid cell
KNN	K-nearest neighbor graph
MAIT	mucosal-associated invariant T cell
MHC	major histocompatibility complex
MNN	mutual nearest neighbor
MS	multiple sclerosis
NAFLD	non-alcoholic fatty liver disease
NASH	non-alcoholic steatohepatitis
NK	natural killer cell
PAMP	pathogen-associated molecular pattern
PBC	primary biliary cholangitis
PCA	principal component analysis

RPGN	rapid progressive glomerulonephritis
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
scRNA-seq	single cell RNA sequencing
TCM	central memory T cell
TCR	T cell receptor
TEM	effector memory T cell
Tfh	T follicular helper cell
Th1	T helper 1 cell
Th2	T helper 2 cell
Th17	T helper 17 cell
Tr1	Type 1 regulatory T cell
Treg	regulatory T cells
TRM	tissue resident memory T cell
t-SNE	t-distributed stochastic neighbour embedding
UMAP	Uniform approximation and projection

# Eidesstattliche Versicherung

Declaration on oath

**Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.**

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hamburg, den 15.11.2021

Unterschrift 