Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Deep immunophenotyping identifies a population of tissue-resident naive-like CD4$^+$ T cells expanded in livers with primary sclerosing cholangitis

Dissertation

to obtain the degree

Doctor rerum naturalium (Dr. rer. nat.)

at the Faculty of Mathematics, Informatics and Natural Sciences
of the University of Hamburg

submitted by

Tobias Poch

Hamburg, August 2020

This study was performed in the period from June 2016 to August 2020 in the research laboratories of the I. Department of Medicine of the University Medical-Center Hamburg-Eppendorf. Supervision during this time was thankfully taken by Prof. Dr. Christoph Schramm and Dr. Dorothee Schwinge. The project is part of the Clinical Research Unit 306 (CRU306) and therefore funded by the Deutsche Forschungsgemeinschaft (DFG, Project no. 278045702).

Date of thesis defense and approval for publication: 13.11.2020

REFEREES:

Prof. Dr. Christoph Schramm
I. Department of Medicine
University Medical Center Hamburg-Eppendorf
Martinistrasse 52
20251 Hamburg

Prof. Dr. Wolfram Brune
Department of Virus-Host-Interaction
Heinrich Pette Institute
Martinistrasse 52
20251 Hamburg

There's very little difference in people.

But that little difference makes a big difference.

That little difference is attitude.

The big difference is whether it's positive or negative.

H_2O. "Black Sheep". *Use Your Voice*, Bridge Nine Records, 2015.

TABLE OF CONTENTS

DECLARATION ON OATH IV

EIDESSTATTLICHE VERSICHERUNG IV

PUBLICATIONS V

CONFERENCE PARTICIPATIONS VIII

ABBREVIATIONS IX

GENE NOMENCLATURE XII

ABSTRACT XV

ZUSAMMENFASSUNG XVII

1. INTRODUCTION 1

1.1. T CELLS 2

1.1.1. NAIVE T CELLS 3

1.1.2. T HELPER CELLS 4

1.1.3. REGULATORY T CELLS 5

1.1.4. CYTOTOXIC T CELLS 6

1.1.5. UNCONVENTIONAL T CELLS 7

1.2. AUTOIMMUNITY 8

1.3. AUTOIMMUNE LIVER DISEASES 9

1.3.1. PRIMARY SCLEROSING CHOLANGITIS 10

2. AIM OF THE STUDY 12

3. RESULTS 13

3.1. THE PERIPHERAL ADAPTIVE IMMUNE SYSTEM BEARS A DISTINCT COMPOSITION IN BILIARY LIVER DISEASE 13

3.2. THE RATIO BETWEEN $CD4^+$ AND $CD8^+$ T CELLS IN LIVERS WITH PSC RESEMBLES THE PHENOTYPE IN PERIPHERAL BLOOD 17

3.3. GENERATION OF AN ATLAS OF INTRAHEPATIC T CELLS IN PSC BY SINGLE-CELL SEQUENCING METHODS 19

3.4. IDENTIFICATION OF A POPULATION OF INTRAHEPATIC NAIVE-LIKE $CD4^+$ T CELLS 24

3.5. INTRAHEPATIC NAIVE-LIKE CD4+ T CELLS ARE EXPANDED IN LIVERS WITH PSC COMPARED TO CONTROL LIVER DISEASE 25
3.6. NAIVE-LIKE CD4+ T CELLS DISPLAY FEATURES OF BOTH, PERIPHERAL AND TISSUE-RESIDENT T CELLS 30
3.7. TRAJECTORY INFERENCE IMPLIES DEVELOPMENTAL POTENTIAL OF NAIVE-LIKE CD4+ T CELLS TO DEVELOP INTO TH17-POLARIZED EFFECTOR CELLS 34
3.8. FUNCTIONAL EXPERIMENTS REVEALED INCREASED CAPACITY OF NAIVE CD4+ T CELLS TO DEVELOP TH17-LIKE EFFECTOR FUNCTION IN PSC 38
4. DISCUSSION 41
5. CONCLUSIONS AND OUTLOOK 48
6. METHODS 49
6.1. PATIENT SELECTION AND ACQUISITION OF BLOOD AND TISSUE SAMPLES 49
6.2. ISOLATION OF MONONUCLEAR CELLS FROM WHOLE BLOOD 51
6.3. ISOLATION OF MONONUCLEAR CELLS FROM LIVER TISSUE 51
6.4. THAWING OF LYMPHOCYTES 51
6.5. MAGNETIC-ACTIVATED CELL SORTING (MACS) 52
6.6. IN VITRO PROLIFERATION ASSAY OF NAIVE CD4+ T CELLS 52
6.7. IN VITRO TH17 POLARIZATION OF NAIVE CD4+ T CELLS 52
6.8. LEGENDPLEX ASSAY 53
6.9. STAINING OF LYMPHOCYTES WITH CONJUGATED ANTIBODIES FOR FLOW CYTOMETRY 53
6.10. CONJUGATION OF ANTIBODIES FOR CITE-SEQ 57
6.11. FACS-SORTING OF LIVING T CELLS FOR CITE- AND SCRNA-SEQ 60
6.12. STAINING OF T CELLS WITH CONJUGATED ANTIBODIES FOR CITE-SEQ 60
6.13. FACS-SORTING OF LIVING NAIVE CD4+ T CELLS FOR SCTCR-SEQ 61
6.14. PREPARATION OF CDNA-LIBRARIES FOR SINGLE-CELL SEQUENCING METHODS 62
6.15. SINGE-CELL RNA SEQUENCING 62
6.16. SINGLE-CELL TCR SEQUENCING 63
6.17. DATA ANALYSES 63
6.17.1. SUPERVISED ANALYSIS OF MULTI-PARAMETER FLOW CYTOMETRY DATA 63
6.17.2. UNSUPERVISED ANALYSIS OF MULTI-PARAMETER FLOW CYTOMETRY DATA 63

6.17.3. SINGLE-CELL SEQUENCING DATA ANALYSIS 64
6.17.3.1. SCRNA-SEQ DATA PROCESSING 64
6.17.3.2. QUALITY CONTROL AND NORMALIZATION 64
6.17.3.3. ADT NORMALIZATION 65
6.17.3.4. BATCH CORRECTION 65
6.17.3.5. DIMENSION REDUCTION, CLUSTERING AND EXPRESSION ANALYSIS 66
6.17.3.6. CELL SCORE CALCULATION 66
6.17.3.7. TRAJECTORY INFERENCE 67
6.17.3.8. COMPARISON OF INTRAHEPATIC AND PERIPHERAL BLOOD T CELLS 67
6.17.3.9. SCTCR-SEQ DATA PROCESSING 68
7. MATERIAL 69
8. SAFETY 77
9. REFERENCES 79
10. SCIENTIFIC COOPERATIONS 90
11. DANKSAGUNGEN 91

DECLARATION ON OATH

I hereby certify, on oaths, that I have written this dissertation myself and that I have not used any aids other than those specified. The submitted written version corresponds to that on the electronic storage medium. I confirm that this dissertation was not submitted in a previous doctoral procedure.

EIDESSTATTLICHE VERSICHERUNG

Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.

Date, Signature

PUBLICATIONS

(1) Poch, T., Krause, J., Casar, C., Liwinski, T., Glau, L., Kaufmann, M., Ahrenstorf, A. E., Hess, L. U., Ziegler, A. E., Martrus, G., Lunemann, S., Sebode, M., Li, J., Schwinge, D., Krebs, C. F., Franke, A., Friese, M. A., Fischer, L., Altfeld, M., Lohse, A. W., Huber, S., Tolosa, E., Gagliani, N., Schramm, C. (2020). Tissue-resident naive-like CD4+ T cells are prone to effector function in patients with primary sclerosing cholangitis. *Journal of hepatology* (under review)

(2) Kunzmann, L. K., Schoknecht, T., Poch, T., Henze, L., Stein, S., Kriz, M., Grewe, I., Preti, M., Hartl, J., Pannicke, N., Peiseler, M., Sebode, M., Zenouzi, R., Horvatits, T., Böttcher, M., Petersen, B. S., Weiler-Normann, C., Hess, L. U., Elise Ahrenstorf, A., Lunemann, S., … Schwinge, D. (2020). Monocytes as potential mediators of pathogen-induced Th17 differentiation in patients with primary sclerosing cholangitis (PSC). *Hepatology (Baltimore, Md.)*, 10.1002/hep.31140. Advance online publication. https://doi.org/10.1002/hep.31140

(3) Hess, L. U., Martrus, G., Ziegler, A. E., Langeneckert, A. E., Salzberger, W., Goebels, H., Sagebiel, A. F., Hagen, S. H., Poch, T., Ravichandran, G., Koch, M., Schramm, C., Oldhafer, K. J., Fischer, L., Tiegs, G., Richert, L., Bunders, M. J., Lunemann, S., & Altfeld, M. (2020). The Transcription Factor Promyelocytic Leukemia Zinc Finger Protein Is Associated With Expression of Liver-Homing Receptors on Human Blood CD56^{bright} Natural Killer Cells. *Hepatology communications*, *4*(3), 409–424. https://doi.org/10.1002/hep4.1463

(4) Ravichandran, G., Neumann, K., Berkhout, L. K., Weidemann, S., Langeneckert, A. E., Schwinge, D., Poch, T., Huber, S., Schiller, B., Hess, L. U., Ziegler, A. E., Oldhafer, K. J., Barikbin, R., Schramm, C., Altfeld, M., & Tiegs, G. (2019). Interferon-γ-dependent immune responses contribute to the pathogenesis of sclerosing cholangitis in mice. *Journal of hepatology*, *71*(4), 773–782. https://doi.org/10.1016/j.jhep.2019.05.023

(5) Martrus, G., Goebels, H., Langeneckert, A. E., Kah, J., Flomm, F., Ziegler, A. E., Niehrs, A., Löbl, S. M., Russu, K., Hess, L. U., Salzberger, W., Poch, T., Nashan, B., Schramm, C., Oldhafer, K. J., Dandri, M., Koch, M., Lunemann, S., & Altfeld, M. (2019). CD49a Expression Identifies a Subset of Intrahepatic Macrophages in Humans. *Frontiers in immunology*, *10*, 1247. https://doi.org/10.3389/fimmu.2019.01247

(6) Lunemann, S., Langeneckert, A. E., Martrus, G., Hess, L. U., Salzberger, W., Ziegler, A. E., Löbl, S. M., Poch, T., Ravichandran, G., Sauter, J., Schmidt, A. H., Schramm, C., Oldhafer, K. J., Altfeld, M., & Körner, C. (2019). Human liver-derived CXCR6^{+} NK cells are predominantly educated through NKG2A and show reduced cytokine production. *Journal of leukocyte biology*, *105*(6), 1331–1340. https://doi.org/10.1002/JLB.1MA1118-428R

(7) Langeneckert, A. E., Lunemann, S., Martrus, G., Salzberger, W., Hess, L. U., Ziegler, A. E., Poch, T., Ravichandran, G., Matschl, U., Bosse, J. B., Tiegs, G., Fischer, L., Koch, M., Herkel, J., Oldhafer, K. J., Schramm, C., & Altfeld, M. (2019). CCL21-expression and accumulation of CCR7^{+} NK cells in livers of patients with primary sclerosing cholangitis. *European journal of immunology*, *49*(5), 758–769. https://doi.org/10.1002/eji.201847965

(8) Lunemann, S., Schöbel, A., Kah, J., Fittje, P., Hölzemer, A., Langeneckert, A. E., Hess, L. U., Poch, T., Martrus, G., Garcia-Beltran, W. F., Körner, C., Ziegler, A. E., Richert, L., Oldhafer, K. J., Schulze Zur Wiesch, J., Schramm, C., Dandri, M., Herker, E., & Altfeld, M. (2018). Interactions Between KIR3DS1 and HLA-F Activate Natural Killer Cells to Control HCV Replication in Cell Culture. *Gastroenterology*, *155*(5), 1366–1371.e3. https://doi.org/10.1053/j.gastro.2018.07.019

(9) Salzberger, W., Martrus, G., Bachmann, K., Goebels, H., Heß, L., Koch, M., Langeneckert, A., Lunemann, S., Oldhafer, K. J., Pfeifer, C., Poch, T., Richert, L., Schramm, C., Wahib, R., Bunders, M. J., & Altfeld, M. (2018). Tissue-resident NK cells differ in their expression profile of the nutrient transporters Glut1, CD98 and CD71. *PloS one*, *13*(7), e0201170. https://doi.org/10.1371/journal.pone.0201170

(10) Martrus, G., Kautz, T., Lunemann, S., Richert, L., Glau, L., Salzberger, W., Goebels, H., Langeneckert, A., Hess, L., Poch, T., Schramm, C., Oldhafer, K. J., Koch, M., Tolosa, E., Nashan, B., & Altfeld, M. (2017). Proliferative capacity exhibited by human liver-resident CD49a+CD25+ NK cells. *PloS one*, *12*(8), e0182532. https://doi.org/10.1371/journal.pone.0182532

CONFERENCE PARTICIPATIONS

(1) European association for the study of the liver (EASL), The international liver congress 2018, Paris, France, Participation with conference paper:
T. Poch *et al.,* Deep immunophenotyping of liver infiltrating lymphocytes reveals distinct differences in the T cell compartment of patients with primary sclerosing cholangitis. *Journal of hepatology*. **68**, S461-S462 (2018), doi:10.1016/S0168-8278(18)31168-1.

(2) German association of the study of the liver (GASL), Annual meeting 2020, Mainz, Germany, Participation with conference paper:
T. Poch *et al., in 36. Jahrestagung der Deutschen Arbeitsgemeinschaft zum Studium der Leber* (Georg Thieme Verlag KG, 2020).

ABBREVIATIONS

Abbreviation	Full form
ADT	antibody-derived tag
AIH	autoimmune hepatitis
AILD	autoimmune liver diseases
ALD	alcoholic liver disease
ALP	alkaline phosphatase
ALT	alanine transaminase
APC	antigen-presenting cell
AZA	azathioprine
BILI	Bilirubin
BSA	bovine serum albumin
CCR	C-C motif chemokine receptor
CD	cluster of differentiation
CD	Crohn’s disease
CITE-Seq	cellular indexing of transcriptomes and eptitopes by sequencing
CTLA	cytotoxic T lymphocyte-associated protein
CXCL	C-X-C motif chemokine
CXCR	C-X-C motif chemokine receptor
DC	dendritic cell
DMSO	dimethyl sulfoxide
e.g.	example given
EDTA	ethylenediaminetetraacetic acid
FACS	fluorescence-activated cell sorting
FCS	fetal calf serum
FDR	false discovery rate
Fig.	figure
GWAS	genome-wide association study
HCV	chronic hepatitis C virus infection
HD	healthy donor
HLA	human leukocyte antigen

Abbreviation	Full form
i.e.	id est
IBD	inflammatory bowel disease
IFNγ	interferon-gamma
IL	interleukin
ILC	innate lymphoid cell
KCl	potassium chloride
KH_2PO_4	monopotassium phosphate
LPS	lipopolysaccharide
LRM	liver resection margin
LTX	liver transplantation
MACS	magnetic-activated cell sorting
MAIT	mucosal-associated invariant T cell
MHC	major histocompatibility complex
Na_2HPO_4	disodium phosphate
NaCl	sodium chloride
NaN_3	sodium azide
NASH	non-alcoholic steatohepatitis
NK	natural killer
NKT	natural killer T cell
P/S	penicillin/streptomycin
PBC	primary biliary cholangitis
PBMC	peripheral blood mononuclear cells
PBS	phosphate buffered saline
PD	programmed cell death protein
PMA	phorbol 12-myristate 13-acetate
PRR	pattern recognition receptor
PSC	primary sclerosing cholangitis
RNA	ribonucleic acid
scRNA-Seq	single-cell RNA sequencing
scTCR-Seq	single-cell T cell receptor sequencing
SNP	single nucleotide polymorphism

Abbreviation	Full form
STAT	signal transducer and activator of transcription
T_C	cytotoxic T cell
T_{CM}	central memory T cell
TCR	T cell receptor
T_{EM}	effector memory T cell
T_{FH}	T follicular Helper
TGF	transforming growth factor
T_H	T Helper
TNF	tumor necrosis factor
T_{REG}	regulatory T cells
t-SNE	t-distributed stochastic neighbor embedding
UC	ulcerative colitis
UDCA	ursodeoxycholic acid
UMAP	uniform manifold approximation and projection

GENE NOMENCLATURE

Gene	**Protein**
AREG	Amphiregulin
CCL20	C-C Motif Chemokine Ligand 20
CCL4	C-C Motif Chemokine Ligand 4
CCL5	C-C Motif Chemokine Ligand 5
CCR6	C-C Motif Chemokine Receptor 6
CCR7	C-C Motif Chemokine Receptor 7
CD27	CD27 Molecule
CD4	CD4 Molecule
CD63	CD63 Molecule
CD69	CD69 Molecule
CD74	CD74 Molecule
CD8A	CD8a Molecule
CD8B	CD8b Molecule
CTLA4	Cytotoxic T-Lymphocyte Associated Protein 4
CXCR3	C-X-C Motif Chemokine Receptor 3
CXCR4	C-X-C Motif Chemokine Receptor 4
CXCR6	C-X-C Motif Chemokine Receptor 6
FCER1G	Fc Fragment Of IgE Receptor Ig
FCGR3A	Fc Fragment Of IgG Receptor IIIa
FGFBP2	Fibroblast Growth Factor Binding Protein 2
FOS	Fos Proto-Oncogene, AP-1 Transcription Factor Subunit
FOXP1	Forkhead Box P1
FOXP3	Forkhead Box P3
GNLY	Granulysin
GPR183	G Protein-Coupled Receptor 183
GZMA	Granzyme A
GZMB	Granzyme B
GZMH	Granzyme H
GZMK	Granzyme K

Gene	**Protein**
ICOS	Inducible T Cell Costimulator
IFNG	Interferon Gamma
IL10	Interleukin 10
IL12RB1	Interleukin 12 Receptor Subunit Beta 1
IL17A	Interleukin 17A
IL23R	Interleukin 23 Receptor
IL2RA	Interleukin 2 Receptor Subunit Alpha
IL7R	Interleukin 7 Receptor
ITGA1	Integrin Subunit Alpha 1
ITGB1	Integrin Subunit Beta 1
JUN	Jun Proto-Oncogene, AP-1 Transcription Factor Subunit
KLF2	Kruppel Like Factor 2
KLRB1	Killer Cell Lectin Like Receptor B1
KLRC1	Killer Cell Lectin Like Receptor C1
KLRD1	Killer Cell Lectin Like Receptor D1
LEF1	Lymphoid Enhancer Binding Factor 1
LTB	Lymphotoxin Beta
LYST	Lysosomal Trafficking Regulator
MAF	MAF BZIP Transcription Factor
MAL	Mal, T Cell Differentiation Protein
NKG7	Natural Killer Cell Granule Protein 7
PIK3R1	Phosphoinositide-3-Kinase Regulatory Subunit 1
PRF1	Perforin 1
RORA	RAR Related Orphan Receptor A
RORC	RAR Related Orphan Receptor C
RPS17	Ribosomal Protein S17
RPS4Y1	Ribosomal Protein S4 Y-Linked 1
SELL	L-Selectin
SOCS3	Suppressor Of Cytokine Signaling 3

Gene	**Protein**
TBX21	T-Box Transcription Factor 21
TCF7	Transcription Factor 7
TIGIT	T Cell Immunoreceptor With Ig And ITIM Domains
TNF	Tumor Necrosis Factor
TNFRSF18	TNF Receptor Superfamily Member 18
TNFRSF4	TNF Receptor Superfamily Member 4
TRDC	T Cell Receptor Delta Constant
TRGC1	T Cell Receptor Gamma Constant 1
TSHZ2	Teashirt Zinc Finger Homeobox 2
TXNIP	Thioredoxin Interacting Protein
ZBTB16	Zinc Finger And BTB Domain Containing 16
ZNF683	Zinc Finger Protein 683

ABSTRACT

Deciphering the single-cell atlas of the parenchymal liver cells has improved our knowledge of human liver physiology by unravelling the zonation of the different hepatic cell types. However, little is known on the composition of intrahepatic immune cells and more importantly, on their contribution to immune-mediated liver diseases, which are of rising incidence. One such disease is primary sclerosing cholangitis (PSC), an enigmatic and progressive liver disease, still lacking effective therapy. Strong HLA-associations indicate involvement of the adaptive immune system, and in particular of the T cell compartment in PSC pathogenesis.

By combining different single-cell RNA sequencing methods and multi-parameter flow cytometry, we conducted an in-depth analysis of the peripheral and intrahepatic immunophenotypes in PSC. We thereby provide comprehensive flow cytometric data on several immune cell subsets from peripheral blood and the liver of patients with PSC, confirming a distinct composition of the T cell compartment. Based on these findings, we generated the first single-cell atlas of intrahepatic T cells from livers with PSC and thereby provide the first atlas of T cells in liver disease. Therein, we identified a previously unrecognized population of intrahepatic naive-like $CD4^+$ T cells, which displayed a transcriptomic signature of tissue residency. Even though we confirmed the presence of this population across different liver diseases, we found it to be particularly expanded in PSC patients. This finding provides evidence for naive T cells residing in non-lymphoid tissue, a cell population which has been subject to speculation, but never substantiated in liver diseases. Interestingly, the transcriptome of the tissue-resident naive-like $CD4^+$ T cell population was similar to their counterparts from peripheral blood, except for genes associated with tissue residency. However, trajectory inference revealed a propensity of tissue-resident naive-like $CD4^+$ T cells to develop towards effector T cells with a T_H17 polarization-state. This developmental propensity could be confirmed in functional experiments. Importantly, T_H17 cells have previously been described to cause autoimmune diseases and microbiota-induced T_H17 cells have been implicated in PSC pathogenesis.

In summary, the findings described herein provide both a comprehensive data source on immune-mediated liver diseases, particularly PSC, and evidence for naive T cells

resident in non-lymphoid tissue. These findings advance the understanding of tissue-resident T cells in immune-mediated liver diseases. In addition, by deciphering this dataset, we identified tissue-resident naive-like $CD4^+$ T cells as potential precursors of T_H17 cells in livers with shed light on a new population presumably involved in the pathogenesis of PSC.

ZUSAMMENFASSUNG

Die Entschlüsselung des Einzelzellatlas der parenchymalen Leberzellen hat unser Wissen über die Physiologie der menschlichen Leber verbessert, indem die lokalen Genexpressionsprofile der verschiedenen Leberzelltypen entschlüsselt wurden. Über die Zusammensetzung intrahepatischer Immunzellen und vor allem über ihren Beitrag zu immunvermittelten Lebererkrankungen, die eine steigende Inzidenz aufweisen, ist jedoch wenig bekannt. Die primär sklerosierende Cholangitis (PSC) ist eine chronische und fortschreitende Erkrankung der Gallengänge, die zu ebendiesen immunvermittelten Lebererkrankungen zählt und für die es keine kausale Therapie gibt. Starke HLA-Assoziationen weisen auf eine Beteiligung des adaptiven Immunsystems und insbesondere des T-Zell-Kompartiments an der PSC-Pathogenese hin.

Durch die Kombination verschiedener Einzelzell-RNA-Sequenzierungsmethoden und der Multiparameter-Durchflusszytometrie wurde im Rahmen dieser Arbeit eine eingehende Analyse der peripheren und intrahepatischen Immunphänotypen der PSC durchgeführt. Damit liefern wir umfassende durchflusszytometrische Daten zu mehreren Immunzell-Untergruppen aus peripherem Blut und der Leber von Patienten mit PSC, die eine besondere Zusammensetzung des T-Zell-Kompartiments bestätigten. Basierend auf diesen Erkenntnissen haben wir den ersten Einzelzellatlas intrahepatischer T-Zellen aus Lebern mit PSC erstellt und somit den ersten Atlas von T-Zellen bei immunvermittelten Lebererkrankungen generiert. Darin identifizierten wir eine zuvor nicht bekannte Population von intrahepatischen naiven $CD4^+$ T-Zellen, die eine transkriptomische Signatur der Geweberesidenz zeigten. Obwohl wir das Vorhandensein dieser Population bei verschiedenen Lebererkrankungen bestätigt haben, stellten wir fest, dass sie bei PSC-Patienten besonders ausgeprägt ist. Dieses Ergebnis liefert Hinweise auf naive T-Zellen, die sich in nicht-lymphatischen Geweben befinden. Hierüber wurde bereits seit geraumer Zeit spekuliert, jedoch konnte dies bisher bei Lebererkrankungen nicht belegt werden. Interessanterweise war das Transkriptom der gewebsresidenten naiven $CD4^+$ T-Zellpopulation ähnlich wie das des peripheren Blutes, mit Ausnahme der Gene, die mit der Geweberesidenz assoziiert sind. Die Analyse des Entwicklungspotenzials der intrahepatischen Zellen zeigte jedoch eine Neigung der gewebsresidenten naiven $CD4^+$ T-Zellen, sich zu Effektor-T-

Zellen mit einem T_H17-polarisierten Zustand zu entwickeln. Diese Neigung konnte in funktionellen Experimenten *in vitro* bestätigt werden. Hier ist es wichtig zu erwähnen, dass zuvor beschrieben wurde, dass T_H17-Zellen Autoimmunerkrankungen antreiben können. Des Weiteren konnte gezeigt werden, dass Mikrobiota-induzierte T_H17-Zellen eine Rolle in der PSC-Pathogenese spielen.

Zusammenfassend liefern die hier beschriebenen Ergebnisse sowohl eine umfassende Datenquelle zu immunvermittelten Lebererkrankungen, insbesondere zu PSC, als auch Hinweise auf naive T-Zellen, die sich resident in nicht-lymphoiden Geweben befinden. Diese Ergebnisse fördern das Verständnis von T-Zellen im Gewebe bei immunvermittelten Lebererkrankungen. Darüber hinaus identifizierten wir durch Entschlüsselung dieses Datensatzes gewebsresidente naive $CD4^+$ T-Zellen als potenzielle Vorläufer von T_H17-Zellen in der erkrankten Leber, die auf eine neue Population hindeuten, die vermutlich an der Pathogenese von PSC beteiligt ist.

1. INTRODUCTION

The immune system comprises a multitude of effector cells and molecules, separated into two major parts, i.e. the innate and the adaptive immune system, that mediate defence against microorganisms and toxins. These mechanisms, in combination with physical and chemical barriers, are critical for an organism to maintain tissue integrity. The innate immune system is a rapid defence mechanism that is activated minutes to hours after detection of a potential pathogen (*1*). The effector cells of the innate immune system thereby act in an unspecific manner to quickly delete the invading microorganisms. These cells include e.g. monocytes/macrophages, granulocytes and dendritic cells (DC) that detect microorganisms via pattern recognition receptors (PRR) and delete them through phagocytosis and antimicrobial peptides (*2*). In addition, these phagocytic cells secrete cytokines and chemokines upon pathogen recognition and thereby attract other innate immune cells, e.g. natural killer (NK) and innate lymphoid cells (ILC), and induce their differentiation into effector cells to support the clearance of infection. Moreover, the secretion of cytokines and chemokines additionally recruits cells of the adaptive immune system (*3*). Adaptive immune cells express highly specific receptors on their cell surface for manifold peptide antigens. Since these cells act highly specific, the adaptive immune system takes hours to days to get activated. The adaptive immune cells comprise B and T cells, which express antigen-specific B and T cell receptors to detect a vast amount of different antigens (*4*). Upon antigen encounter, the respective cell gets activated and starts to proliferate. This process is referred to as clonal expansion. B cells that undergo this process then differentiate into plasma cells that secrete antibodies against their antigen and thereby enhance other immune cells to detect and delete the microorganism coated by these antibodies. T cells on the other hand differentiate into one of several types of effector T cells, e.g. T Helper cells, cytotoxic T cells or regulatory T cells, which will be addressed in the following chapters. Most of these effector cells die after successfully clearing the infection. A small pool of cells develops into memory cells though and thereby forms a long lasting memory against the respective microorganism (*5*–*7*).

1.1. T CELLS

T lymphocytes or T cells are the major part of the adaptive immune system and therefore play a central role within the immune response. They are characterized by the expression of antigen-specific T cell receptors (TCR), consisting of two different chains, which are TCRα and TCRβ in 95 % of T cells. However, a small fraction of T cells bears a TCR consisting of TCRγ and TCRδ. The different TCR chains are closely connected to different chains of the TCR co-receptor CD3, which altogether form the functional TCR-complex (Fig. 1).

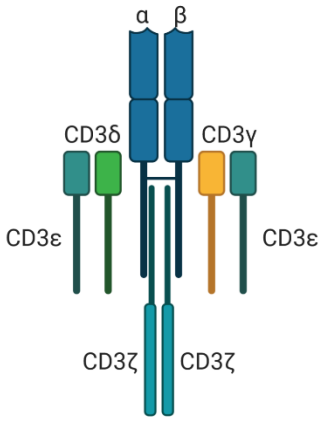

Figure 1: Schematic structure of the T cell receptor complex.
The classical TCR, consisting of alpha- and beta-chains, is depicted. The different signalling components of the CD3 co-receptor are added, which in combination form the functional TCR-complex. The scheme was created with BioRender.com.

The TCR is a highly variable receptor, which allows the recognition of a vast amount of different antigens. This variability is a result of V(D)J recombination, which, in short, randomly re-organizes the genes for the different chains of the TCR during the early developmental stages of T cells. Thereby, approximately $3x10^{11}$ combinations of the gene segments, resulting in different antigen-specific TCRs, can be generated (*5*). To disable T cells with a TCR specific for antigens deriving from own tissues, these cells undergo maturation within the thymus, where autoreactive T cells are deleted (*6*).
The functional TCR binds its specific ligands presented on major histocompatibility complex (MHC) molecules, which are present on all cells of the organism. To properly induce activation of T cells, the TCR complex needs accessory signalling molecules,

i.e. co-receptors. The major co-receptor for this process is CD28, which binds to its ligands CD80/CD86 mainly expressed on professional antigen-presenting cells (APC). Without concomitant binding of CD28 to its ligands, the TCR complex alone only induces parts of the signalling pathways associated with T cell activation, which leads to the cessation of proliferation. This improper activation is referred to as anergy and is an irreversible phenomenon in most cases. MHC molecules present antigens deriving from either endogenous or exogenous proteins. Endogenous proteins are constantly degraded within all types of cells and fragments of these proteins are presented on the cellular surface. These fragments are presented on all cells throughout the organism via MHC class I molecules. Exogenous proteins, deriving from e.g. invading microorganisms, are presented by innate immune cells via MHC class II molecules. These APC, i.e. monocytes/macrophages and dendritic cells, phagocyte microorganisms and degrade them intracellularly to then present protein fragments on their surface. The binding of the TCR to MHC molecules is also highly dependent on co-receptors. The successful binding of TCRs to MHC class I molecules requires the co-receptor CD8, which is present only on a subset of T cells. These T cells have the capacity to induce apoptosis, a controlled form of cell death, in their target cells and are therefore termed cytotoxic T cells. In contrast, MHC class II molecules bind to a different co-receptor, which is called CD4 and present on a different subset of T cells. Since MHC class II molecules are mostly expressed on phagocytic cells, $CD4^+$ T cells are directly involved in the initiation of an adaptive immune response against pathogens. Therefore, these cells were termed T Helper cells. The different subsets of T cells will be further described in the following chapters. T cells that express both CD4 and CD8 is a very small fraction of unknown relevance so far and will therefore not be further discussed (*5*).

1.1.1. NAIVE T CELLS

After maturation, the antigen-unexperienced or naive T cells egress from the thymus and enter the circulation between secondary lymphoid organs and blood. The process of homing to the secondary lymphoid organs is mediated by CCR7 and CD62L (*8*), which are hallmarks of naive T cells. In addition, naive T cells express the CD45RA isoform of the leukocyte common antigen (LCA) isoform on their surface, which changes to CD45RO, a shorter isoform, upon activation. The process of entering

secondary lymphoid organs and egressing back to the blood is a recurring process until the cells recognize their specific antigen. Human naive T cells are therefore long-lived with lifespans up to 6-8 years. A T cell is considered to be naive until it encounters its specific antigen and initiates proliferation and clonal expansion, followed by differentiation into effector and memory T cells, that are able to migrate to peripheral tissues and locally identify and deplete their antigen (*9*).

1.1.2. T HELPER CELLS

Throughout clonal expansion and inflammation, $CD4^+$ T cells encounter diverse microenvironments that specifically modulate the differentiation of naive cells into effector cells. Upon activation, APC secrete high amounts of different cytokines and chemokines that in turn recruit and activate other immune cells. The interplay of these cells shape this microenvironment and promote the differentiation into the subtypes of T Helper (T_H) cells. Since the number of T_H cell subsets is increasing, only the major subtypes will be depicted here.

T Helper 1 (T_H1) cells are induced by large amounts of interleukin (IL) -12, which mostly derives from activated dendritic cells. Since IL-12 stimulates NK cells as well, these cells begin to secrete interferon-γ (IFNγ). This in turn promotes the expression of signal transducer and activator of transcription (STAT) 1 in T cells, which results in the expression of T-bet (*TBX21*), a hallmark transcription factor in T_H1 cells. Downstream signalling of T-bet involves IFNγ, which is a signature cytokine for T_H1 cells, and IL-12Rβ2, which closes the loop for expansion of these cells through IL-12 (*10, 11*). T_H1 cells are additionally characterized by surface expression of the chemokine receptor CXCR3 and are considered as a pro-inflammatory cell type in general.

The differentiation of naive $CD4^+$ T cells into T Helper 2 (T_H2) cells *in vivo* is discussed controversially. The inducing cytokines and signature markers are known and were proven *in vitro*. IL-4 is the key cytokine for induction of T_H2 differentiation. Upon binding of IL-4 to its receptor, the expression of STAT6 is induced in T cells. As a result of STAT6 activation, expression of GATA3 is induced. This in turn leads to expression of IL-4 itself, which is the signature cytokine of T_H2 cells. Dendritic cells, granulocytes and innate lymphoid cells are the major candidates for induction of T_H2 differentiation. However, the source of IL-4 *in vivo* remains incompletely understood (*12*). For

phenotyping, the chemokine receptor CCR4 is used for identification of T_H2 cells. This type of T cells was mainly associated to allergy-related inflammatory processes but evidence is growing that these cells are involved in pathogenic pro-inflammatory environments, too (*13*).

The third major subset of T_H cells are T Helper 17 (T_H17) cells. The major signalling component for differentiation into T_H17 cells is STAT3. This transcription factor induces the expression of the hallmark transcription factor in T_H17 cells, which is RORγt (*RORC*). In addition, T_H17 cells are identified by surface expression of chemokine receptor CCR6 and CD161 (*KLRB1*). The activation of STAT3 may happen via different cytokines. IL-6 was identified as the most important cytokine and strongest stimulus, which is secreted by activated innate immune cells. Other STAT3-inducing cytokines, i.e. IL-23 and IL-21, were identified to play a role for T_H17 cells as well. They are accessory and rather maintaining these cells than differentiating them. Interestingly, TGF-β1, a cytokine mainly associated with regulatory T cells, was shown to be strongly involved in the differentiation into T_H17 cells. However, the exact role of TGF-β1 in this process is unclear (*10, 14*).

1.1.3. REGULATORY T CELLS

Immune responses need to be strictly controlled and terminated after clearance of the infection, which is the task of regulatory T cells (T_{REG}). The hallmarks of T_{REG} are the expression of IL-10 and Foxp3 (*FOXP3*), which is the master transcription factor for these cells, as well as surface expression of CD25 (*IL2RA*), accompanied with low expression of CD127. The expression of Foxp3 is induced via STAT5 by binding of the cytokines TGF-β1 and IL-2, which is mostly produced by activated non-regulatory T cells. This is of interest, since this induction generates a negative feedback loop by expanding T_{REG} that in turn suppress activated non-regulatory T cells. Upon antigen encounter within the secondary lymphoid organs T_{REG} get activated and suppress proliferation and differentiation of naive T cells. Of note, activation of T_{REG} requires 10-100-fold less antigen compared to naive T cells. T_{REG} exert suppression of other cells by a variety of mechanisms, e.g. absorption of cytokines, suppressing IL-2 production by degrading extracellular ATP to adenosine via CD73/CD39, as well as modulating

APC via CTLA4 (*15*). A defective function of T_{REG} thus leads to exceeding immune responses, which contributes to development of diseases.

Differentiation of naive $CD4^+$ T cells into the different subsets mentioned above is summarized in Figure 2.

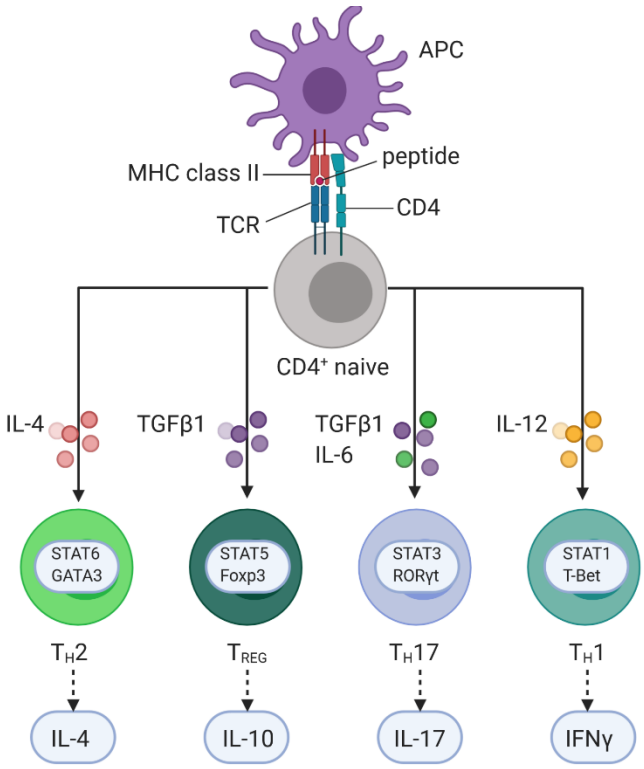

Figure 2: Schematic differentiation of naive $CD4^+$ T cells into subsets of T_H cells. *Schematic illustration of cytokines directing the fate of naive $CD4^+$ T cells via their key signalling pathways upon TCR-stimulation by APC. Only the hallmarks of differentiation for each subset are depicted. The scheme was created with BioRender.com*

1.1.4. CYTOTOXIC T CELLS

The vast majority of $CD8^+$ T cells are thought to be cytotoxic lymphocytes (T_C) that delete aberrant, e.g. infected or transformed, cells. Naive $CD8^+$ T cells are primed within the secondary lymphoid organs by APC. Upon antigen binding, these cells undergo massive expansion which is accompanied by a metabolic switch that allows to produce proteins and nucleic acids while generating enough energy. The cytokines

IL-2 and IL-12 support the expansion and differentiation of $CD8^+$ T cells, which comes along with expression of the transcriptional repressor Blimp-1. This in turn induces gene expression of *TBX21* and *ID2*, which are thought to be important transcription factors for effector cell differentiation. However, the differentiation of naive $CD8^+$ T cells into effector cells is far from fully understood. After leaving the secondary lymphoid organs, $CD8^+$ T cells migrate to peripheral sites of infection mainly via chemokine receptor CXCR3, which binds to a range of chemokines. At the site of infection, effector $CD8^+$ T cells release effector molecules, e.g. IFNγ, TNFα, granzymes and perforins, to deplete their target cells. In addition to this major function, $CD8^+$ T cells were shown to bear regulatory function to prevent tissue damage. This regulatory function is mainly accomplished by transient secretion of IL-10 (*16*).

1.1.5. UNCONVENTIONAL T CELLS

Most T cell subsets are listed among the adaptive immune cells. However, some T cell populations exhibit innate-like properties and thus cover both the adaptive and innate immune system. A shared feature of these populations is an invariant TCR, which only includes a limited number of different TCR chains and combinations of these.

One example for these cells are mucosal-associated invariant T (MAIT) cells, expressing an invariant alpha-chain of the TCR (Vα7.2-Jα33), which binds to antigens presented via the highly conserved MHC class I-associated molecule MR1. Thereby, MAIT cells are able to recognize bacterial metabolites. These cells are characterized by high expression of CD161 (*KLRB1*), irrespective of CD4 and CD8 expression. In addition, they express the chemokine receptor CCR6 and cytokine receptors IL-12R and IL-18R, which reflects an MR1-independent way of activation. Upon activation, MAIT cells secrete the pro-inflammatory cytokines IL-17, IFNγ and TNFα, as well as cytotoxic granzymes. In addition, an activated phenotype, characterized by expression of classical activation markers as CD69, CD38 and PD-1 (*PDCD1*), is rapidly acquired by these cells. Since this subset of T cells is highly abundant within liver tissue and at mucosal sites, it is thought to be first line of defence against infections (*17*).

Another rather small population of invariant T cells shares phenotypical features with NK cells. Therefore, this T cell subset is referred to as NKT cells, which is also

characterized by expression of an invariant TCR alpha-chain (Vα2.4-Jα1.8). Due to the conserved TCR, these cells are restricted to lipid-antigens presented via CD1d, a protein related to non-classical MHC class I molecules. In addition to the TCR alpha-chain, surrogate markers for NKT cells are CD161 (*KLRB1*) and CD56 (*NCAM1*), which is generally expressed on NK cells. Upon activation, these cells secrete high levels of IL-4 as well as IFNγ, which drove the hypothesis for NKT cells supporting T_H2-related immune responses. Since these cells appear at strongly varying frequencies in humans, their function is not fully understood yet (*18, 19*).

The third subtype of unconventional T cells does not express a TCR consisting of alpha- and beta-chains, as all subsets mentioned above, but a TCR consisting of gamma- and delta-chains. Therefore, this subset is termed γδ T cells. The γδTCR recognizes phosphorylated non-peptide molecules, mainly deriving from metabolites of the isoprenoid synthesis, which is used by many microorganisms. This T cell subset exhibits a large cytotoxic potential, which is mediated by expression of NKG2D (*KLRK1*) and secretion of effector molecules as granzymes (*GZMA*-*GZMK*), perforin (*PRF1*) and granulysin (*GNLY*). In addition, these cells can deplete their target cells in an indirect way of antibody-dependent cell-mediated cytotoxicity via CD16. Upon activation, these cells secrete large amounts of TNFα, IL-17 and cytotoxic effector molecules, which renders them potent effector cells against invading microorganisms. γδ T cells are mainly located within epithelial tissues and detectable only at low frequencies in peripheral blood. Due to this fact, these cells are thought, as already mentioned for MAIT cells, to take part in the first line of defence against infections (*20, 21*).

1.2. AUTOIMMUNITY

As outlined above, the immune system comprises a variety of different effector mechanisms to destroy exogenous microorganisms. It is therefore critical that these effector mechanisms are strictly controlled to avoid destruction of own tissues (*1*). The process ensuring this effect is referred to as self-tolerance and has therefore been extensively studied. Self-tolerance is known to be accomplished by presentation of self- and non-self-antigens to immature cells of the adaptive immune system. This procedure is called negative selection or central tolerance and takes part in primary

lymphoid organs, e.g. the thymus. Most cells of the adaptive immune system, namely B and T cells, that recognize self-antigens are deleted by central tolerance (*22*). Cells escaping central tolerance and egressing from the primary lymphoid organs will be deleted by mechanisms of so called peripheral tolerance in most cases. Peripheral tolerance is accomplished by either deletion of the reactive cell, conversion into T_{REG} or suppression by mature T_{REG} and induction of anergy (*23*). These mechanisms are occasionally disturbed by e.g. genetic variants as single-nucleotide polymorphisms (SNP). SNPs located within gene segments that are involved in presentation and recognition of antigens may lead to defective tolerance mechanisms and thereby support the development of autoimmune disorders (*24*). There are nearly 100 different listed autoimmune diseases and conditions with varying prevalence and incidence. These diseases can occur at any age, however, each disease has its characteristic age of onset. The risk to develop autoimmune disorders is increased within first degree relatives and affects more women than men in most cases. In general, the incidence of several autoimmune diseases increased within the western world over the last decades.

1.3. AUTOIMMUNE LIVER DISEASES

Autoimmune liver diseases (AILD) mainly comprise three major categories, namely autoimmune hepatitis (AIH), primary biliary cholangitis (PBC) and primary sclerosing cholangitis (PSC) (*25–27*). All three are complex disorders, characterized by a progressive disease course that leads to cirrhosis, liver failure and need for liver transplantation (LTX), if left untreated. Due to different pathogenetic mechanisms, the treatment options of the AILD vary. AIH is treated with immunosuppressive agents such as azathioprine (AZA) and corticosteroids, which is effective in approximately 80 % of patients. In contrast, the diseases of the bililary tract, i.e. PSC and PBC, are both mainly treated with ursodeoxycholic acid (UDCA), whereas only the latter shows a favourable response to this bile acid regarding long-term survival.

Overall, treatment success of AILD is still limited or missing, resulting in a range of patients that do no not respond to the current options and therefore have a poor prognosis (*28*). This is mainly a result of so far unknown pathogenetic disease mechanisms. These facts highlight the urgent and unmet need to identify effective

treatment targets in AILD. mechanisms. These facts highlight the urgent and unmet need to identify effective treatment targets in AILD.

1.3.1. PRIMARY SCLEROSING CHOLANGITIS

Primary sclerosing cholangitis (PSC) is a progressive cholestatic liver disease with persistent inflammation and obliteration of the bile ducts. In contrast to the related biliary disease PBC, PSC affects both the intra- and extrahepatic bile ducts. In comparison to other autoimmune diseases, PSC affects more males than females (1.5:1) and the disease is of increasing incidence in the western countries. The median onset of disease occurs at 41 years of age and most patients with PSC (approximately 70 %) have concomitant inflammatory bowel disease (IBD). This comorbidity increases the risk for the development of both cholangio- and colorectal carcinoma. In general, PSC is slowly progressing, however, there is no reliable biomarker that predicts the pace of progression and course of disease. Importantly, there is no medical treatment with a proven effect on disease progression. The median survival without liver transplantation is within 12 - 20 years after diagnosis (*29–31*).

The knowledge on PSC disease pathogenesis is largely unknown but several mechanisms have been considered as part of a complex and multifactorial disease origin. These include e.g. alterations of the gut microbiota and dysregulated immune responses. The microbiome is an emerging field of research in PSC. It has been shown that the diversity of intestinal microbiota was markedly reduced in patients with PSC. This finding was independent from the presence of associated colitis, and the abundance of potentially pro-inflammatory bacterial species was found to be increased among different cohorts (*32, 33*). Moreover, bile from patients with PSC was recently identified to harbour an altered microbiome as well (*34*). This is of interest since the bile ducts are the site of injury in PSC. Consequently, biliary epithelial cells or cholangiocytes are thought to be critically involved in disease pathogenesis. The findings of altered microbiota within intestine in patients with PSC, combined with concomitant IBD in most cases, established the hypothesis of a ‘leaky gut’. This concept involves disruption of the intestinal barrier which leads to increased permeability and translocation of bacteria or bacterial products into portal blood (*35*). In addition to an altered microbiome, bile is considered to play an important role in the pathogenesis of PSC, since biliary stasis and disruption of the biliary epithelial barrier

aggravates the injury of the surrounding tissue. Therefore, reducing exposure of cholangiocytes to toxic bile acids is one strategy to slow down disease progression. To that end, the farnesoid X receptor (FXR) and G protein-coupled bile acid receptor 1 (TGR5) are considered to be the most promising candidates for interference with bile acid synthesis and signalling (*36*).

Genome-wide association studies (GWAS) have revealed strong associations of PSC to polymorphisms within the human leukocyte antigen (HLA; resembles MHC in humans) class II gene segments (*37–39*). These associations suggest the adaptive immune cells to be involved in PSC. In addition, associated risk loci outside of the HLA-region include genes that are potentially related to the function of T cells (*30*). Indeed, T cells were shown to be the predominant immune cell type within the portal inflammatory infiltrate in liver biopsies from patients with PSC (*30*) and interaction of cholangiocytes, expressing HLA molecules as well as non-classical receptors such as CD1d and MR1 with T cells could be shown (*40, 41*). Due to its rarity, studies of intrahepatic T cells in PSC are limited. However, published studies described intrahepatic T cells co-expressing CCR9 and integrins $\alpha4\beta7$, which are markers associated with lymphocyte-homing to the intestine. In addition to this, expression of the ligand of CCR9 and integrin $\beta7$, i.e. CCL25 and MAdCAM-1, respectively, was detected in these livers (*42–44*). These findings in combination with altered intestinal microbiota fueled the hypothesis on a process referred to aberrant ‘gut lymphocyte homing’, which involves the activation of T cells within the gut, which are erroneously guided to the liver and cause hepatic inflammation (*35*). Further evidence for this process was shown by the detection of T cells with common clonal origin in gut and liver of patients with PSC (*45*).

In peripheral blood, dysregulated T cell responses, e.g. reduced regulatory T cell (T_{REG}) numbers and increased IL-17 production of peripheral blood $CD4^+$ T cells, have been reported in PSC (*46–50*). Interestingly, microbiota derived from patients with PSC were shown to stimulate T cell differentiation into T_H17 cells (*51*) and may thereby drive liver inflammation in mouse models of sclerosing cholangitis (*52*).

2. AIM OF THE STUDY

Dysregulated T cells, particularly T_{REG} and T_H17 cells, are considered to be involved in the pathogenesis of PSC. However, due to the rarity of PSC and the difficulties in obtaining sufficient human liver tissue, most studies were conducted on cells derived from peripheral blood or in translational mouse models. Little is known about the composition and function of human intrahepatic T cells and how they may contribute to disease pathogenesis. Within this study, we aimed to decipher in unprecedented depth the peripheral and intrahepatic immunophenotypes in PSC. Our data will provide new insights into disease pathogenesis and will help to impact future clinical management and potential immunotherapeutic treatment options.

3. RESULTS

3.1. THE PERIPHERAL ADAPTIVE IMMUNE SYSTEM BEARS A DISTINCT COMPOSITION IN BILIARY LIVER DISEASE

Within this study, we performed deep immunophenotyping via flow cytometry on a well-characterized cohort of patients with autoimmune liver diseases (AILD) as well as healthy donors (HD) to obtain a comprehensive overview of a multitude of immune cell populations across different conditions. To that end, we included a total of 118 patients (PSC: n=33; PBC: n=31; AIH: n=24; HD: n=30) and analysed frequencies of several immune cell subsets *ex vivo* in peripheral blood. Clinical parameters of the patient cohort are summarized in Table 4. We determined the frequency of the respective cell type within total $CD45^+$ cells and separated the data by innate and adaptive origin and calculated Z-Scores to get an impression of overall differences among the different conditions.

The score was calculated with the following equation:

$$z_i = \frac{x_i - \overline{x}}{\sigma}$$

with $\overline{x}$: mean of the population and σ: standard deviation of the population (*53*).

Within the innate immune system, we identified eight different subtypes of immune cells such as NK cells, different APCs and innate lymphoid cells (ILC) (Fig. 3a). NK cells were further subdivided into $CD56^{bright}$ and $CD56^{dim}$ NK cells, based on the expression of the classical NK cell markers CD56 and CD16 (gating cf. Fig. 19) (*54, 55*). The subsets of NK cells showed only slight differences across the different patient cohorts. $CD56^{bright}$ NK cells were detected at similar frequencies, whereas $CD56^{dim}$ NK cells were decreased in patients with AIH ($p<0.0001$, Fig. 3b). Three subtypes of antigen-presenting cells (APC) were gated, including monocytes, plasmacytoid dendritic cells (pDC) and myeloid dendritic cells (mDC). Monocytes were identified based on expression of CD14 and CD16 (*56, 57*). After excluding monocytes, subtypes of dendritic cells were gated based on expression of CD123, CD11c and CD1c (gating cf. Fig. 19) (*55, 58*). Monocytes and pDC were detected at similar frequencies across

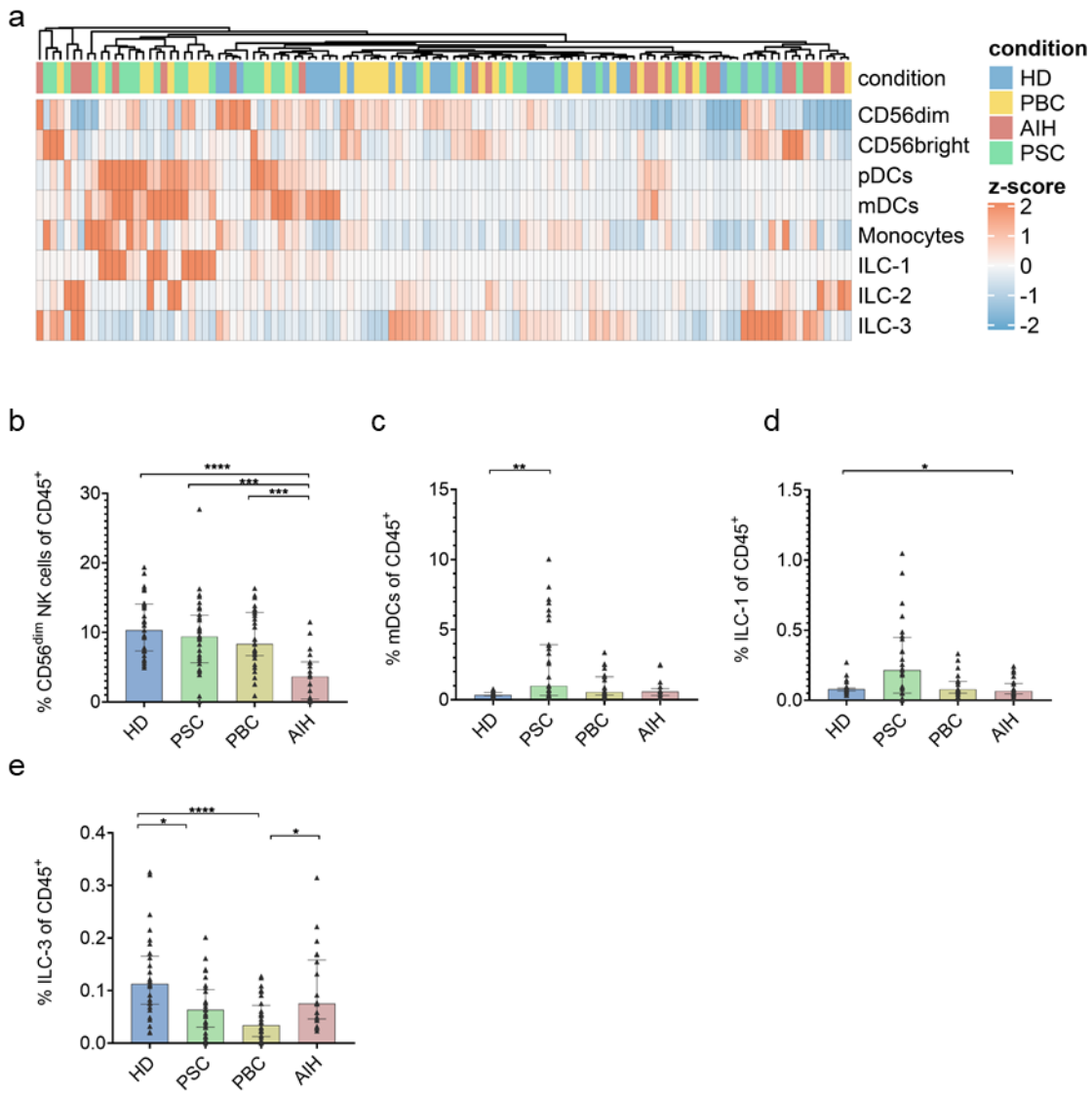

Figure 3: Frequencies of innate immune cells in AILD compared to HD.

***a**, Frequencies of each cell type were analysed by flow cytometric analysis and calculated within total CD45⁺ cells. Data is presented as Z-Score for each population among all conditions to gain an impression about overall differences. Populations that significantly differed between the conditions are shown in detail. **b**,**c**,**d**,**e**, Frequency of CD56^{dim} NK cells (**b**), mDCs (**c**), ILC-1 (**d**) and ILC-3 (**e**) within total CD45⁺ cells.*

*HD, healthy donor (n=30); PSC, primary sclerosing cholangitis (n=33); PBC, primary biliary cholangitis (n=31); AIH, autoimmune hepatitis (n=24); NK, natural killer, mDCs, myeloid dendritic cells; ILC, innate lymphoid cells. Data in (**b**,**c**,**d**,**e**) are presented as median with interquartile-range. *: p<0.05; **: p<0.01; ***: p<0.001; ****: p<0.0001 as assessed by Kruskal-Wallis test.*

all conditions, whereas mDC were increased in all diseases with PSC reaching statistical significance, when compared to HD (p=0.0046, Fig. 3c). The remaining three cell types all belong to ILC, which were separated based on different expression of CD117 and CD294 (gating cf. Fig. 20) (*55, 59*).

ILC Type 1 and Type 3 cells were found to differ between the conditions, the former being increased in PSC compared to AIH (p=0.0394) and the latter being decreased in all diseases compared to HD (p<0.0001, Fig. 3d-e).

Although detecting differences between the different AILD, the overall composition of these subtypes of innate immune cells showed only slight differences when comparing AILD to healthy donors (Fig. 3).

We next aimed to decipher the distribution of immune cell population within the adaptive immune system. First, B cells were detected by expression of high levels of CD19 and varying levels of CD20 (gating cf. Fig. 19) (*55*). The frequency of B cells was decreased in patients with AIH only, when compared to PSC (p=0.0076, Fig. 4).

First, three subtypes of innate-like immune cells, i.e. $\gamma\delta$ T cells, MAIT, NKT cells, were analysed. $\gamma\delta$ T cells express the gamma and delta chains of the TCR; mucosal-associated invariant T (MAIT) cells express high levels of CD161 as well as a specific alpha chain of the TCR (Vα 7.2); and NKT cells express the NK cell marker CD56 as well as a specific alpha chain of the TCR (Vα 2.4) (gating cf. Fig. 20 + 21) (*55*). NKT and $\gamma\delta$ T cells showed only slight differences, as only $\gamma\delta$ T cells showed an increase in PSC compared to PBC (p=0.0196, Fig. 4). Interestingly, MAIT cells were decreased in PBC and AIH, but not PSC, when compared to HD (p<0.0001, Fig. 4). The decrease of peripheral MAIT cells in AILD could be shown in other studies (*60, 61*). After extracting the innate-like T cells, irrespective of their expression of costimulatory molecules, from total $CD3^+$ cells, we analysed the classical $CD4^+$ and $CD8^+$ T cells. Frequencies of $CD4^+$ T cells, which contain several subsets of different T Helper (T_H) cells and $CD8^+$ T cells, which mainly function as cytotoxic cells, were determined (gating cf. Fig. 21). Interestingly, $CD4^+$ T cells were detected at significantly reduced frequencies within the biliary diseases, but not AIH, when compared to HD (p<0.0001, Fig. 4). In turn, $CD8^+$ T cells were markedly increased in these patients (p<0.0001, Fig. 4).

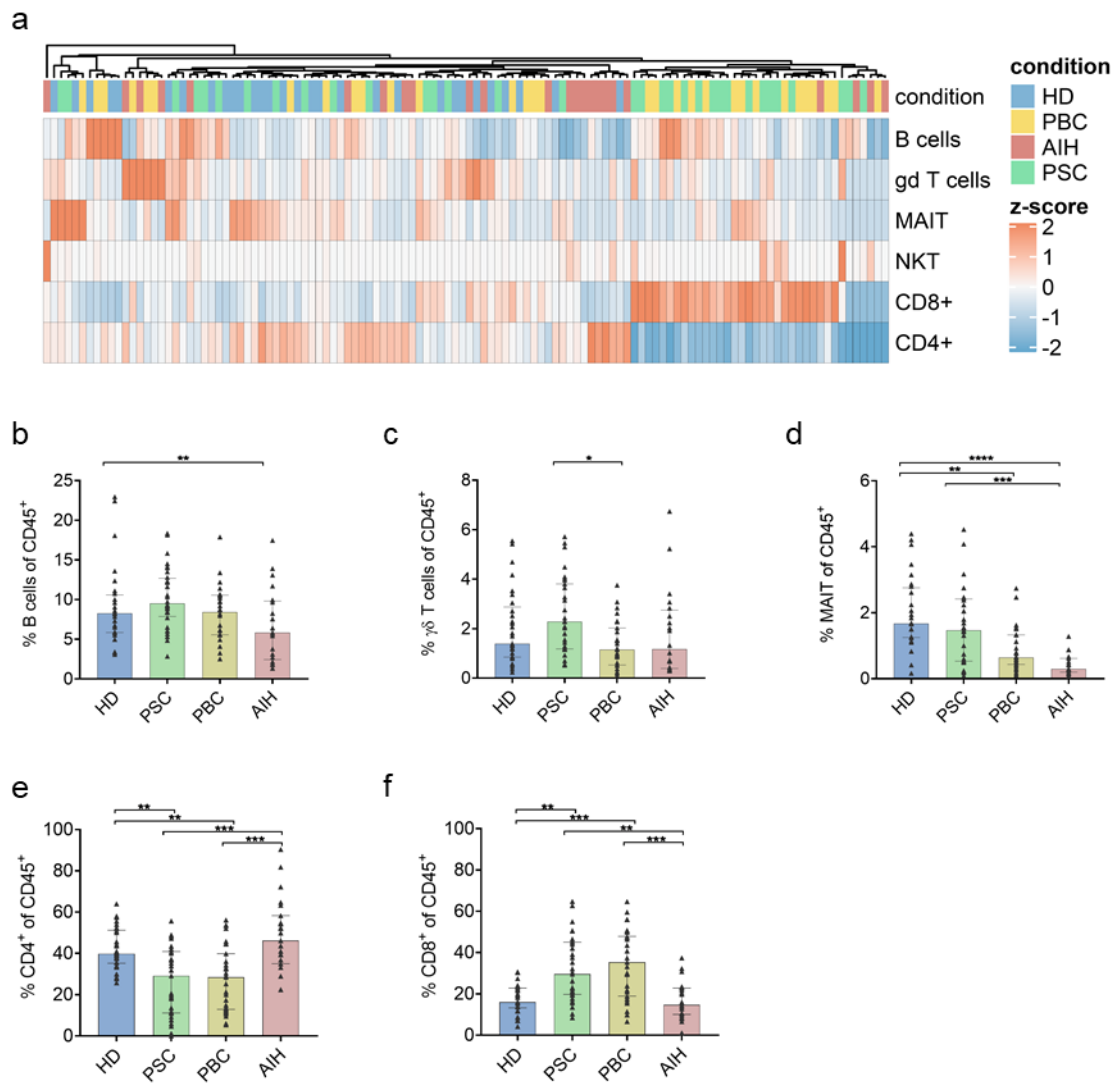

Figure 4: Frequencies of adaptive immune cells in AILD compared to HD.

***a**, Frequencies of each cell type were analysed by flow cytometric analysis and calculated within total CD45⁺ cells. Data is presented as Z-Score for each population among all conditions to gain an impression about overall differences. Populations that significantly differed between the conditions are shown in detail. **b**,**c**,**d**,**e**,**f**, Frequency of B cells (**b**), γδ T cells (**c**), MAIT cells (**d**), CD4⁺ (**e**) and CD8⁺ (**f**)T cells within total CD45⁺ cells. HD, healthy donor (n=30); PSC, primary sclerosing cholangitis (n=33); PBC, primary biliary cholangitis (n=31); AIH, autoimmune hepatitis (n=24); MAIT, mucosal-associated invariant T. Data in (**b**,**c**,**d**,**e**,**f**) are presented as median with interquartile-range. *: p<0.05; **: p<0.01; ***: p<0.001; ****: p<0.0001 as assessed by Kruskal-Wallis test.*

When calculating the ratio between $CD4^+$ and $CD8^+$ T cells per patient, this effect was even more pronounced in biliary disease (PSC: 0.46; PBC: 0.76) (Fig. 5).
In sum, our immunophenotyping approach revealed the adaptive immune system to bear a distinct composition of immune cells, particularly T cell subtypes, within the biliary diseases PSC and PBC, when compared to HD and patients with AIH.

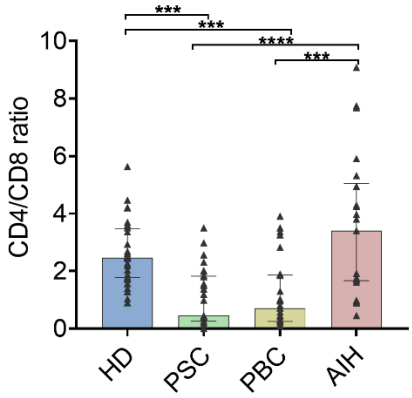

Figure 5**: **Frequencies of CD4⁺ and CD8⁺ T cells in peripheral blood of patients with AILD compared to HD.
*Ratio between $CD4^+$ and $CD8^+$ T cells in peripheral blood of patients with AILD and HD. Frequencies of each cell type within total $CD45^+$ cells were calculated. HD, healthy donor (n=30); PSC, primary sclerosing cholangitis (n=33); PBC, primary biliary cholangitis (n=31); AIH, autoimmune hepatitis (n=24). Data are presented as median with interquartile-range. *: p<0.05; **: p<0.01; ***: p<0.001; ****: p<0.0001 as assessed by Kruskal-Wallis test.*

Since we detected the strongest differences regarding PSC within the T cell compartment, we decided to further investigate these differences.

3.2. THE RATIO BETWEEN $CD4^+$ AND $CD8^+$ T CELLS IN LIVERS WITH PSC RESEMBLES THE PHENOTYPE IN PERIPHERAL BLOOD

Since characterisation of immune cell populations in PBMC only scratches the surface of a disease affecting a specific organ, we next investigated whether the findings regarding the $CD4^+$ and $CD8^+$ T cells were true for intrahepatic T cells, too. To that end, T cells were isolated from liver explant tissue and analysed by flow cytometry. We

included PSC (n=16), chronic hepatitis C virus infection (HCV, n=5), non-alcoholic steatohepatitis (NASH, n=3) and alcoholic liver disease (ALD, n=11) related cirrhosis as control diseases. To exclude cirrhosis-related effects, non-cirrhotic liver resection margins (LRM, n=9) were used as controls. The frequencies of liver-derived $CD4^+$ T cells differed significantly between patients with PSC and ALD only (p=0.009). Importantly, the other immune-mediated liver disease, i.e. HCV, and LRM showed similar trends as PSC, although not reaching statistical significance (Fig. 6a). The lower frequencies of $CD4^+$ T cells in PSC and other immune-mediated diseases was accompanied by a significantly increased frequency of $CD8^+$ T cells in PSC patients when compared to ALD patients (p=0.0412). Of note, the LRM samples showed by far the highest frequency of $CD8^+$ T cells (53.99 % vs. 45.10 % vs. 28.85 % vs. 48.11 % vs. 32.12 % of $CD3^+$, respectively), which might be due to strong infiltration of tumour specific $CD8^+$ T cells in tumour surrounding tissue (*62*).

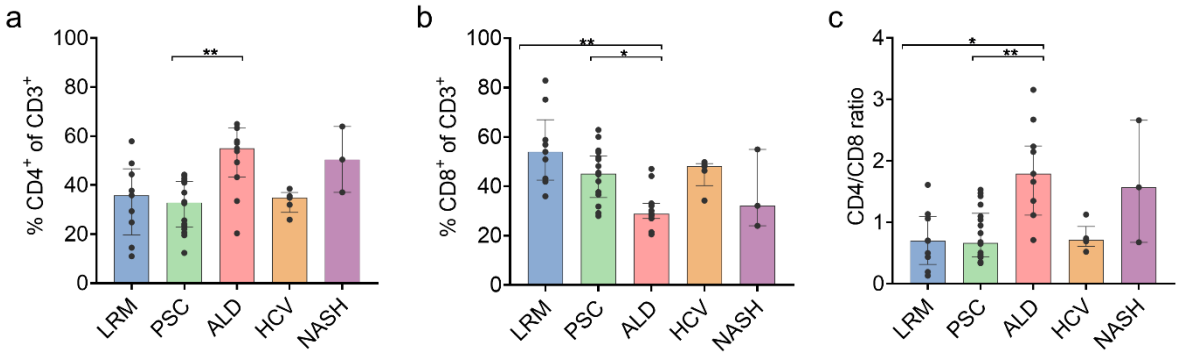

Figure 6: Frequencies of CD4⁺ and CD8⁺ T cells in cirrhotic and non-cirrhotic liver tissue among different diseases.

a, *Frequencies of intrahepatic $CD4^+$ T cells in patients with PSC, HCV, ALD and NASH related cirrhosis compared to non-cirrhotic LRM.* ***b****, Frequencies of intrahepatic $CD8^+$ T cells.* ***c****, Ratio between intrahepatic $CD4^+$ and $CD8^+$ T cells. Frequencies of each cell type within total $CD3^+$ cells were calculated. LRM, liver resection margin (n=9); PSC, primary sclerosing cholangitis (n=18); ALD, alcoholic liver disease (n=11); HCV, chronic hepatitis C virus infection (n=5); NASH, non-alcoholic steatohepatitis (n=3). Data are presented as median with interquartile-range. *: p<0.05; **: p<0.01 as assessed by Kruskal-Wallis test.*

Of note, the metabolic disorders, i.e. NASH and ALD, showed similar trends towards higher frequencies of $CD4^+$ than $CD8^+$ T cells and vice versa was detected the immune-mediated diseases, i.e. PSC and HCV (Fig. 6a-b).

The differences in frequencies of $CD4^+$ and $CD8^+$ T cells were even more pronounced by calculating the ratio of these cells. Thus, PSC showed the lowest ratio between $CD4^+$ and $CD8^+$ T cells, which was strongly decreased compared to ALD (0.66 vs. 1.79, p=0.0042) (Fig. 6c) and was therefore used as control for further experiments.

3.3. GENERATION OF AN ATLAS OF INTRAHEPATIC T CELLS IN PSC BY SINGLE-CELL SEQUENCING METHODS

We next aimed to decipher the PSC-associated T cell immunophenotype in more detail. Therefore, we generated a single-cell atlas of intrahepatic T cells isolated from PSC livers, which included cirrhotic tissue from explanted livers with PSC (n=9) as well as non-cirrhotic liver tissue from resection of PSC-associated cholangiocarcinoma (CCA, n=2) and thereby covered different stages of disease. We here combined single-cell RNA sequencing (scRNA-Seq) and Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-Seq) on FACS-sorted T cells (gating cf. Fig. 23) from freshly processed liver tissue as well as stored intrahepatic lymphocytes (Fig. 7).

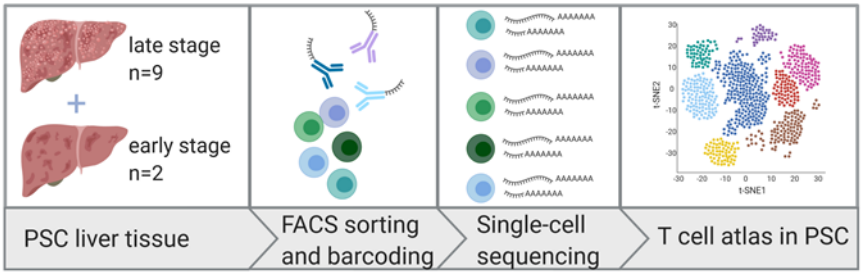

Figure 7: Graphical abstract of the workflow for generation of the single-cell atlas of intrahepatic T cells in PSC. *Liver tissue was hashed immediately after surgery. After successive filtration through cell strainers of different sizes, lymphocytes were purified through density centrifugation (see methods section) and labelled with antibodies for FACS-sorting and CITE-Seq. The scheme was created with BioRender.com*

We identified 13 distinct subsets of intrahepatic T cells (Fig. 8) by assembling differentially expressed genes (DEG) and protein expression via CITE-Seq antibodies (antibody-derived tags, ADT) (Fig. 9a-b). These subsets included four clusters of $CD8^+$ T cells and five clusters of $CD4^+$ T cells. Furthermore, we identified four clusters that exhibited heterogeneous expression of CD4 and CD8, consisting of low expression of either CD4 or CD8, and a small fraction of double positive $CD4^+CD8^+$ cells and double negative $CD4^-CD8^-$ cells. Cells within these subsets did not cluster according to their expression pattern of CD4 and CD8 (Fig. 8 + Fig. 9b), but to the expression of innate-like genes. One of the clusters was assigned to γδ T cells, expressing genes such as *TRDC*, *TRGC1,* and genes associated with cytotoxicity, including *GNLY*, and innate immune response genes such as *FCER1G*.

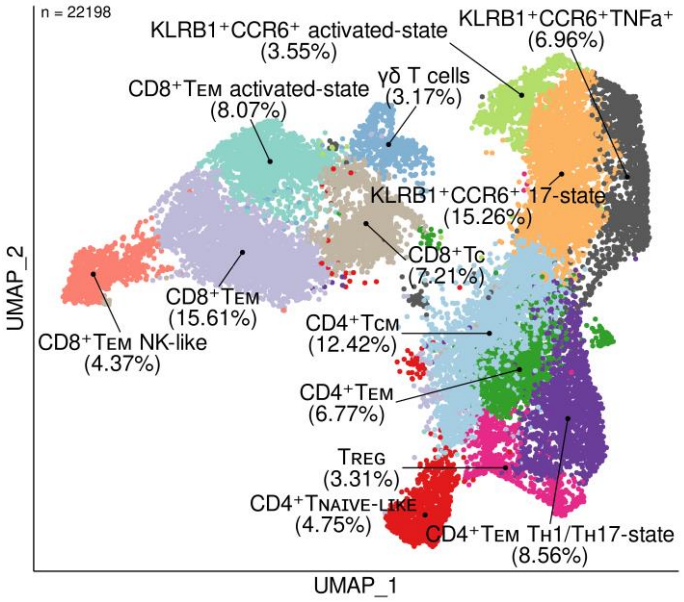

Figure 8: Single-cell atlas of intrahepatic T cells in PSC.

Atlas of 22,196 intrahepatic T cells from patients with PSC (n=11), showing 13 distinct clusters. T_{CM}, central memory cell; T_{EM}, effector memory cell; T_C, cytotoxic T lymphocyte; T_{REG}, regulatory T cell; T_{EM} T_H1/T_H17-state, effector memory cell with a T_H1 and T_H17 polarization-state. (in cooperation with Dr. Jenny Krause and Christian Casar)

The remaining three clusters were characterized by high expression of *KLRB1* and *CCR6,* and surface markers CCR6 and CD56, which are known hallmarks of MAIT cells (*63–65*). These three clusters exhibited distinct functional phenotypes: (i) T_H17-

associated properties, e.g. by expressing the transcription factor *RORA*; (ii) a pro-inflammatory cellular state, characterized by expressing high levels of *TNF*; (iii) an activated cellular state, by expression of chemokines *CCL20*, *CCL4* and *CD69* (Fig. 9a-b).Among the four $CD8^+$ T cell clusters, we identified effector memory cells ($CD8^+$ T_{EM}), characterized by low expression of CCR7 and low expression of granzymes; $CD8^+$ T cells with an activated cellular state ($CD8^+$ activated), characterized by the expression of effector molecules such as *IFNG*, *CCL4*, *CCL5, and GZMK*; cytotoxic $CD8^+$ T cells ($CD8^+$ T_C) with cytotoxicity-related genes such as *GNLY*, *GZMB*, *CD63,* and tissue residency gene *ZNF683 (HOBIT)*; and $CD8^+$ T cells, with features of natural killer (NK) cells expressing genes such as *NKG7*, *KLRD1*, *FGFBP2,* and *FCGR3A*, which we refer to as '$CD8^+$ NK-like' (Fig. 9a-b). The $CD4^+$ T cell population consisted of effector memory ($CD4^+$ T_{EM}) and central memory ($CD4^+$ T_{CM}) cell populations, characterized by the typical expression of surface markers (i.e. CD127, CD45RA and CCR7) and signature genes (*MAL*, *LTB*, *IL7R*) (Fig. 9a-b). Furthermore, we identified $Foxp3^+$ regulatory T cells (T_{REG}) by the expression of signature genes *IL2RA* and *FOXP3* as well as co-inhibitory molecules *TIGIT* and *CTLA4,* and the expression of surface markers such as CD25 (Fig. 9a-b). Interestingly, we identified a cluster of $CD4^+$ T cells that was characterized by the expression of a naive-like gene signature such as *SELL*, *CCR7*, *TSHZ2*, *LEF1,* and *TCF7* (*7, 9, 66, 67*), which was accompanied by surface expression of CCR7 and CD45RA, which are typical markers for naive T cells (*7, 9*) (Fig. 9a-b). The last cluster of $CD4^+$ T cells was identified to exhibit properties of both T Helper 1 (T_H1) and T Helper 17 (T_H17) cells (*CCR6*, *KLRB1*, *MAF* and *CXCR6)* (Fig. 9a-b). We further elucidated this cluster by analysing expression of signature genes for T_H1 (*IFNG*, *TNF*, *TBX21*, *CXCR3* and *IL12RB1*) as well as T_H17 cells (*IL17A*, *RORC*, *RORA*, *IL23R* and *CCR6*) and confirmed the presence of cells with these polarization-states (Fig. 10). Therefore, we termed the cluster "$CD4^+$ effector memory cells with a T_H1 and T_H17 polarization-state" ($CD4^+$ T_{EM} T_H1/T_H17-state).

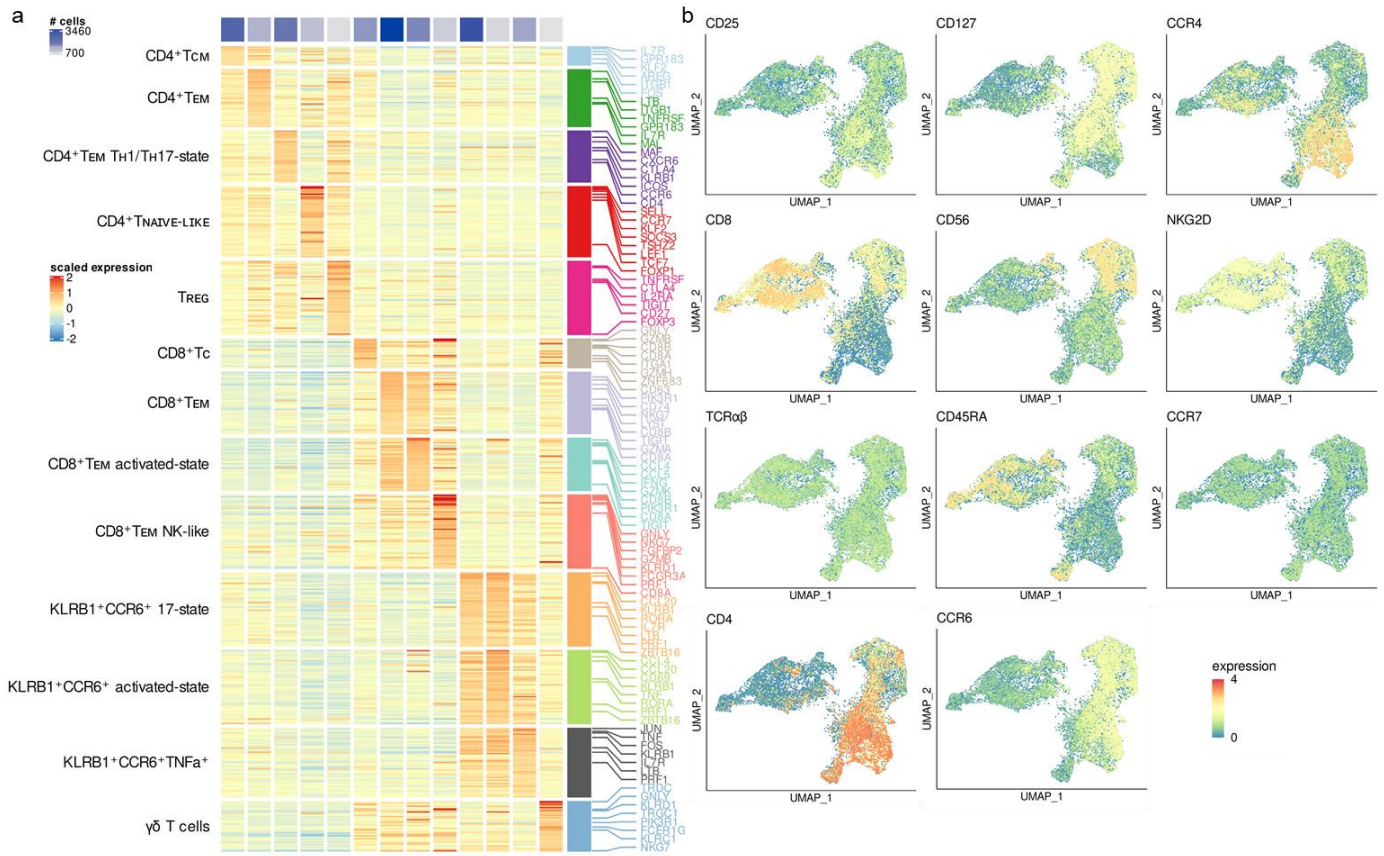

Figure 9: Assignment of cell types to each cluster of the single-cell atlas of intrahepatic T cells.

***a**, Cluster-Heatmap highlighting signatures of differentially expressed genes (DEG) for each cluster of the atlas. **b**, ADT expression on intrahepatic T cells. Signals were used in combination with gene expression data for determination of different T cell subtypes. T_{CM}, central memory cell; T_{EM}, effector memory cell; T_C, cytotoxic T lymphocyte; T_{REG}, regulatory T cell; T_{EM} T_H1/T_H17-state, effector memory cell with a T_H1 and T_H17 polarization-state. (in cooperation with Dr. Jenny Krause and Christian Casar)*

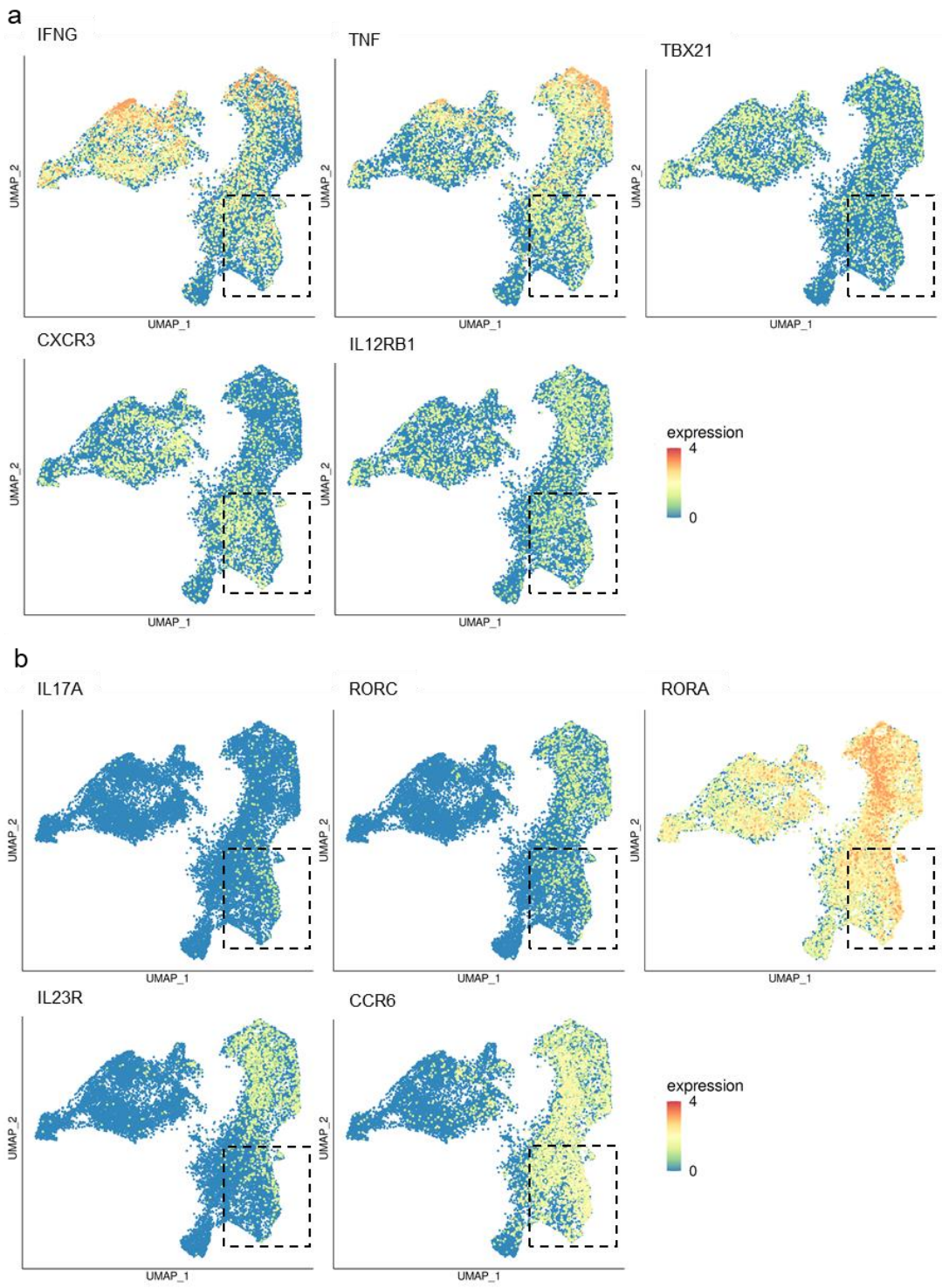

Figure 10: Assignment of a cluster of CD4⁺ T cells with properties of T_H17 and T_H1 cells.

***a**,**b**, Gene expression in the edged area, including the CD4⁺ T cells with T_H1 and T_H17 associated properties from the atlas, was analysed. Relative expression of genes associated with T_H1 (**a**) and T_H17 (**b**) cells, respectively. (in cooperation with Dr. Jenny Krause and Christian Casar)*

3.4. IDENTIFICATION OF A POPULATION OF INTRAHEPATIC NAIVE-LIKE CD4⁺ T CELLS

We next examined the unexpected abundance of intrahepatic T cells with a naive-like phenotype within the atlas described above. Therefore, we examined surface expression of CCR7 and CD45RA, which are typical markers of naive T cells (*9*), by hierarchical gating on ADT-derived data which was corresponding to the samples included in the atlas. By projection of the manually gated cells onto the atlas, we identified a significant population of naive-like CD4⁺ T cells (CCR7⁺CD45RA⁺) that matched the cluster of naive-like CD4⁺ T cells from the atlas (cf. Fig. 8) and thereby confirmed the abundance of these cells inside the inflamed liver (Fig. 11a-b). Consequently, we wondered whether these cells were truly naive and had not yet undergone clonal expansion through antigen-recognition. Therefore, we performed single-cell TCR sequencing (scTCR-Seq) on corresponding blood- and liver-derived naive-like CD4⁺ T cells from a subgroup of patients with PSC from the atlas (n=3).

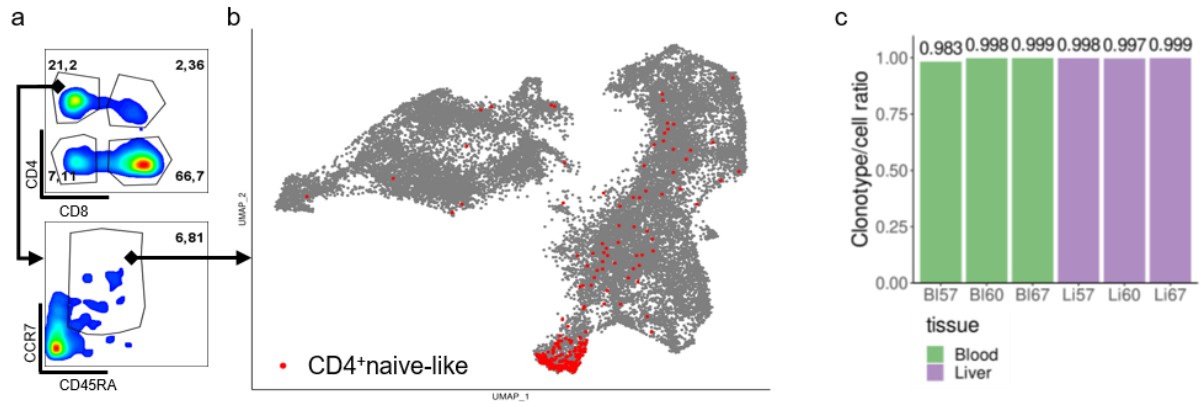

Figure 11: Single-cell TCR sequencing confirmed naive cellular state of naive-like CD4⁺ T cells in PSC.

***a**, Hierarchical gating of CD4⁺ T cells, using the ADT for CCR7 and CD45RA. **b**, Cells gated in (**a**) are highlighted in the UMAP (Atlas) containing intrahepatic T cells from patients with PSC (n=11). Gating confirms abundance of T cells with a naive phenotype. **c**, scTCR-Seq of naive-like CD4⁺ T cells from corresponding blood and liver of patients with PSC (n=3) showing overall high clonotype per cell ratios. (in cooperation with Dr. Jenny Krause and Christian Casar)*

The heterogeneity of the TCR repertoire showed a high clonotype to cell ratio (approximately 1 clonotype per cell) and was similar between intrahepatic naive-like $CD4^+$ T cells and peripheral bona fide naive $CD4^+$ T cells, confirming that the intrahepatic naive-like $CD4^+$ T cells had not undergone clonal expansion, which is a hallmark of naive T cells (Fig. 11c). We therefore termed these cells 'intrahepatic naive-like $CD4^+$ T cells'.

3.5. INTRAHEPATIC NAIVE-LIKE $CD4^+$ T CELLS ARE EXPANDED IN LIVERS WITH PSC COMPARED TO CONTROL LIVER DISEASE

Since naive T cells are usually located within lymphoid tissue, we decided to further explore this population of intrahepatic naive-like $CD4^+$ T cells. In particular, we wondered whether the presence of this population within the liver was a unique feature of PSC, or a feature common to other liver inflammatory related conditions, such as cirrhosis due to ALD (n=11), HCV (n=5), NASH (n=3) as well as non-cirrhotic LRM (n=9) as control.

We determined the frequencies of naive-like $CD4^+$ T cells from liver tissue of a different cohort of patients with PSC (n=16) and controls using a flow cytometry-based targeted approach (Fig. 12). Naive-like $CD4^+$ T cells were detected in all livers examined (LRM: 1.15 %; PSC: 3.59 %; ALD: 0.98 %; HCV: 1.39 %; NASH: 1.15 % of $CD4^+$ T cells) (Fig. 12b). Interestingly, the highest frequency was detected in patients with PSC, which reached statistical significance, compared to ALD ($p=0.019$). In line with this, the frequency of bona fide naive $CD4^+$ T cells was also increased in PSC in peripheral blood, when compared to the respective controls (ALD: $p=0.0018$; NASH: $p=0.0398$) (Fig. 12c).

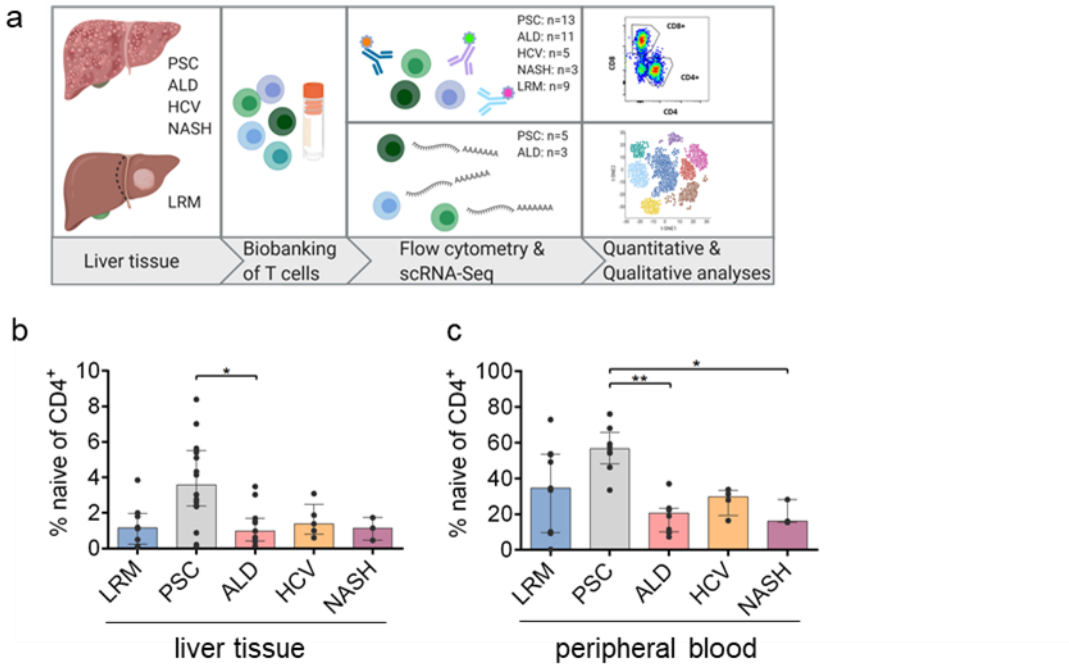

Figure 12: Naive-like CD4⁺ T cells are expanded in patients with PSC.

***a**, Graphical abstract of the workflow. **b**, Frequencies of intrahepatic naive-like CD4⁺ T cells determined by classical hierarchical gating for CCR7 and CD45RA. **c**, Frequencies of naive-like CD4⁺ T cells in peripheral blood, respectively, determined by classical hierarchical gating for CCR7 and CD45RA.*

*Data in (**b**,**c**) is presented as median with interquartile-range. *: p<0.05 as assessed by Kruskal-Wallis test. PSC, primary sclerosing cholangitis; ALD, alcoholic liver disease; HCV, chronic infection with hepatitis C virus; NASH, non-alcoholic steatohepatitis; LRM, liver resection margin.*

We next aimed to examine the enrichment of naive-like CD4⁺ T cells in PSC by an unsupervised approach. Therefore, we used the hierarchically gated populations from our flow cytometry-based approach and analysed them with different algorithms. First, we performed a principle component analysis (PCA) considering all intrahepatic CD4⁺ T cell populations, e.g. naive, effector and memory, and found a difference within the CD4⁺ T cell compartment among PSC and ALD patients (p=0.001; R2=8.4 %) (Fig. 13a). Second, to determine which T cell populations accounted for the differences between PSC and ALD, we utilized random forest machine learning, using all 24 markers defining CD4⁺ T cell populations and subpopulations in our flow cytometry panels (Table 1). An accurate and statistically significant classification performance

was obtained that distinguished CD4⁺ T cells between PSC and ALD patients (p<0.0001; AUC=0.805; Sensitivity=0.857; Specificity=0.786) (Fig. 13b).

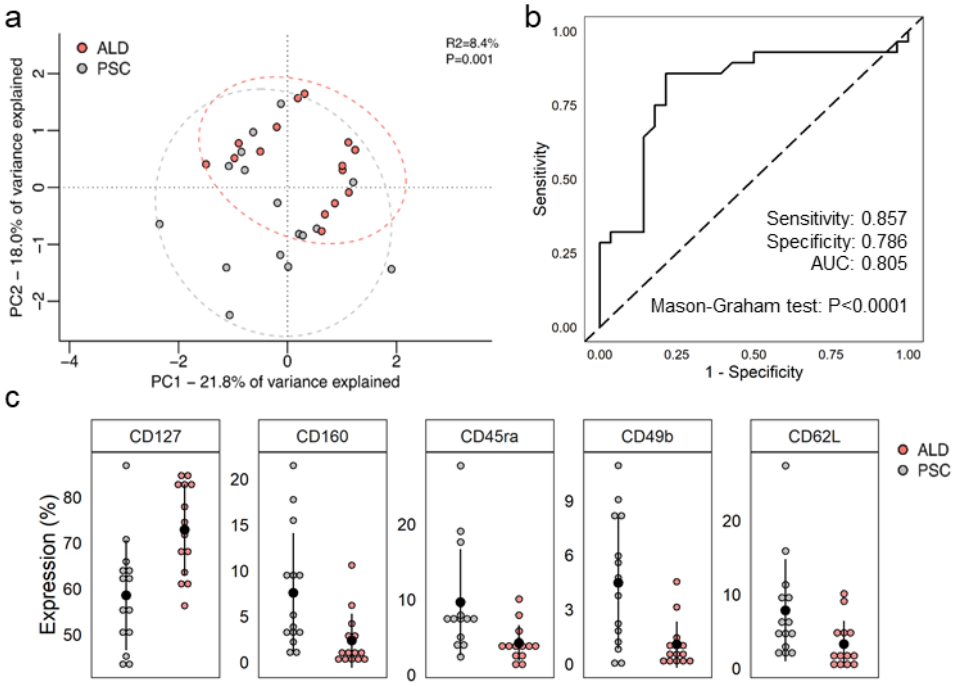

Figure 13: Unsupervised random forest machine learning confirms expansion of intrahepatic naive-like CD4⁺ T cells in patients with PSC.

***a**, PCA of multi-parameter flow cytometry data of intrahepatic CD4⁺ T cells from late stage PSC (grey) and ALD (red). Permutational analysis of variance on Euclidean distance was performed to determine statistical differences between the groups. P<0.05 was considered statistically significant. **d**, Random forest classifier for intrahepatic CD4⁺ T cells of PSC and ALD. Discrimination was assessed by means of area under the receiver operating characteristic curve (AUC), specificity and sensitivity. Statistical significance of the classifier performance was tested by Mason's and Graham's non-parametric test. P<0.05 was considered statistically significant. **e**, Populations of intrahepatic CD4⁺ T cells identified to best discriminate between PSC and ALD, identified by iteratively removing the features, which are proved by a statistical test to be less relevant than random samples generated by permutation of the original variable values. PSC, primary sclerosing cholangitis; ALD, alcoholic liver* disease. *(in cooperation with Dr. Timur Liwinski)*

Table 1: Antigens included for machine learning CD4⁺ T cells.

Cluster of differentiation	***Alias***
CD25	Interleukin-2 Receptor Subunit Alpha
CD27	CD27 antigen
CD28	T-cell-specific surface glycoprotein CD28
CD39	Ectonucleoside triphosphate diphosphohydrolase 5
CD57	Beta-1,3-Glucuronyltransferase 1
CD69	Early Activation Antigen CD69
CD73	5'-Nucleotidase Ecto
CD103	Integrin Subunit Alpha E
CD127	Interleukin-7 Receptor Subunit Alpha
CD160	Natural Killer Cell Receptor BY55
CD161	Killer Cell Lectin Like Receptor B1
CD183	C-X-C Motif Chemokine Receptor 3
CD194	C-C Motif Chemokine Receptor 4
CD196	C-C Motif Chemokine Receptor 6
CD197	C-C Motif Chemokine Receptor 7
CD223	Lymphocyte Activation Gene 3 Protein
CD244	Natural Killer Cell Receptor 2B4
CD272	B- And T-Lymphocyte Attenuator
CD279	Programmed Cell Death Protein 1
CD45RA	Receptor-type tyrosine-protein phosphatase C
CD49a	Integrin alpha-1
CD49b	Integrin beta-1
CD62L	L-selectin
HLA-DR	Class II HLADR-beta 1 chain

The classifier identified five surface markers on CD4⁺ T cells, including CD45RA and CD62L, which were strongly increased and are classical markers of naive T cells (*9*) (Fig. 13c).

In addition, the significant differences in frequencies of CD127, which is also a classical marker included in distinguishing different T cell memory subtypes (*7, 68*), as well as CD49b and CD160 on intrahepatic $CD4^+$ T cells were detected to be meaningful for discriminating between PSC and ALD. We next asked whether naive-like $CD4^+$ T cells bear qualitative differences among liver diseases. Therefore, we performed scRNA-Seq on biobanked T cells isolated from the livers of patients with ALD (n=3), which showed the strongest quantitative differences compared to PSC, and integrated the data with scRNA-Seq data from patients with PSC (n=5) from the atlas. We again detected intrahepatic naive-like $CD4^+$ T cells, located within a distinct cluster but containing cells from both PSC and ALD samples, by alignment of the respective gene signature (partly shown in Table 2) extracted from the single-cell atlas (Fig. 14a).

Table 2: DEG signature of naive-like T cells from the atlas of intrahepatic T cells in PSC.

Gene	**Protein**
SELL	L-Selectin
CCR7	C-C motif chemokine receptor 7
KLF2	Kruppel like factor 2
SOCS3	Suppressor of cytokine signaling 3
AREG	Amphiregulin
TSHZ2	Teashirt zinc finger homeobox 2
LDHB	Lactate dehydrogenase B
LEF1	Lymphoid enhancer binding factor 1
MAL	Mal, T cell differentiation protein
TCF7	Transcription factor 7
FOXP1	Forkhead box P1

Comparing the naive-like $CD4^+$ T cells between PSC and ALD, we found that the gene expression profiles of intrahepatic naive-like $CD4^+$ T cells did not show significantly different gene expression between the conditions, indicating high overall similarity of these cells across different conditions (Fig. 14b).

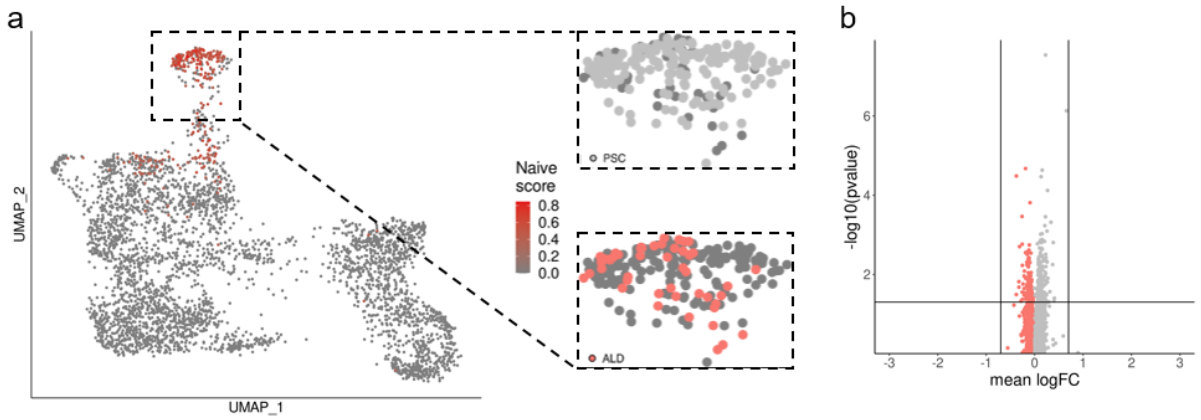

Figure 14: Intrahepatic naive-like CD4⁺ T cells are similar among liver diseases. ***a****, UMAP projection of 4,072 intrahepatic T cells of late stage PSC and ALD. Projection of naive-like DEG core signature, extracted from the atlas, is shown. Cells with matching transcriptomes are highlighted in red. Contribution of underlying disease to the cluster of naive-like CD4⁺ T cells is shown in the respective colour.* ***b****, Volcano plot of DEG between intrahepatic naive-like CD4⁺ T cells from late stage PSC and ALD, indicating similarity of gene expression. Lines indicate cut-off of statistical significance (p<0.05) and logarithmic fold change of expression (|logFC|>0.7). (in cooperation with Dr. Jenny Krause and Christian Casar)*

Taken together, these data show the presence of a naive-like CD4⁺ T cell population which was detectable among liver diseases of different aetiologies and showed comparable gene expression irrespective of the underlying disease. However, this population was particularly expanded not only within the liver, but also in peripheral blood of patients with PSC.

3.6. NAIVE-LIKE CD4⁺ T CELLS DISPLAY FEATURES OF BOTH, PERIPHERAL AND TISSUE-RESIDENT T CELLS

We next wondered if the intrahepatic naive-like CD4⁺ T cells are tissue-resident or merely blood-deriving cells circulating through the liver. Therefore, we compared these cells to circulating bona fide naive CD4⁺ T cells from peripheral blood within the same patients. We performed CITE-Seq on T cells deriving from peripheral blood of patients

with PSC (n=2) and integrated the corresponding data of intrahepatic T cells from the atlas (Fig. 15).

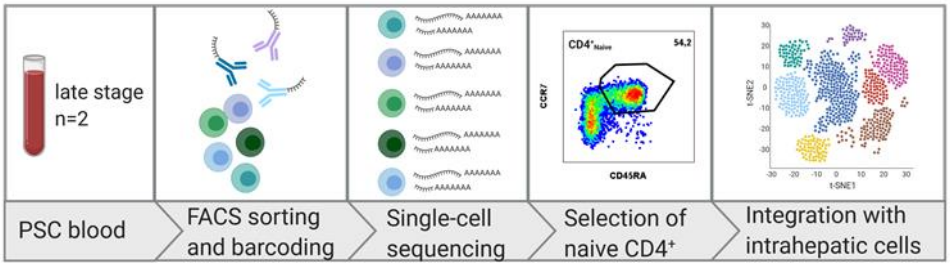

Figure 15: Graphical abstract of the workflow for examination of tissue residency. *Mononuclear cells were isolated from peripheral blood by density centrifugation, followed by FACS-sorting for T cells. Single-cell sequencing was performed on T cells. Naive CD4⁺ T cells were selected by expression of the naive DEG signature from the atlas and expression of CCR7 and CD45RA. The scheme was created with BioRender.com*

First, we tested the accordance of all sequenced T cells to the naive DEG signature extracted from the atlas (cf. Table 2) to identify naive-like CD4⁺ T cells. Naive-like CD4⁺ T cells deriving from both tissues were detected, which did only partially cluster based on their origin, indicating overall similarity of their gene expression profiles. To evaluate the degree of similarity between the naive-like CD4⁺ T cells originating from blood and liver, we assessed differentially expressed genes between liver- and blood-derived cells. Interestingly, when directly comparing the transcriptomes between the different origins, the liver-derived cells showed upregulation of genes associated with tissue residency, i.e. *CD69* and *CXCR4* (Fig. 16b), the latter supporting T cell homing to the epithelial surface of the liver (*69, 70*). Consequently, we aimed to further evaluate the hints of tissue residency, and tested the accordance of these cells to a literature-based gene signature of CD69⁺ tissue-resident T cells (*71*). To that end, we used the published gene signature and added *CD69* to that list (Table 3).

Table 3: Core gene signature associated with CD69$^+$ tissue-resident T cells.

Gene	**Protein**
ITGA1	Integrin α-1
ITGA6	Integrin α-6
IL2	Interleukin-2
IL10	Interleukin-10
CD69	Early Activation Antigen CD69
CXCR6	C-X-C chemokine receptor type 6
CXCL13	Chemokine (C-X-C motif) ligand 13
KCNK5	Potassium channel subfamily K member 5
RGS1	Regulator of G-protein signaling 1
CRTAM	Cytotoxic and regulatory T cell molecule
DUSP6	Dual specificity phosphatase 6
PDCD1	Programmed cell death protein 1
IL23R	Interleukin-23 receptor

Interestingly, a tissue-resident signature was seen in liver-derived naive-like CD4$^+$ T cells, which was not seen for cells from peripheral blood. The expression of that gene signature was comparable levels to CD4$^+$ T_{EM} T_H1/T_H17-state from the same dataset and significantly higher compared to the blood-derived cells ($p<0.001$, Fig. 16c). However, part of the liver-derived cells did not display a tissue-resident signature. These cells likely account for blood-derived cells within the liver. The population of CD4$^+$ T_{EM} T_H1/T_H17-state from the same dataset was selected as for the atlas (cf. Fig. 10) and served as a positive control for tissue residency. To quantify the cells with tissue-resident features, we applied a classification implemented in the *Seurat* package (*72*). Based on expression of the gene signature, cells were either categorized as resident or circulating. With this, we identified 38 % of the intrahepatic naive-like CD4$^+$ T cells to be indeed tissue-resident (Fig. 16d).

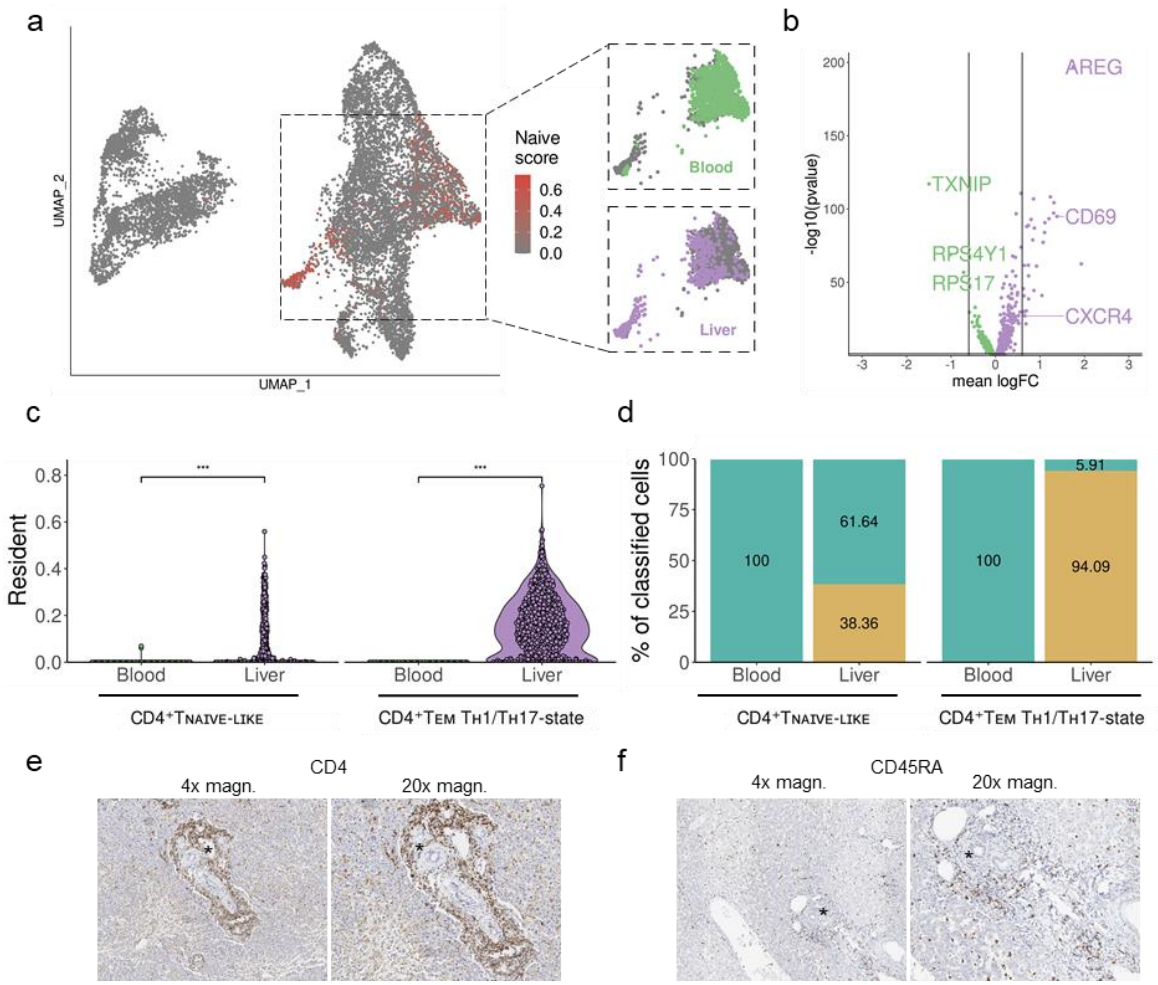

Figure 16: Intrahepatic naive-like CD4⁺ T cells are alike peripheral blood bona fide naive CD4⁺ T cells but display a signature of tissue residency.

***a**, UMAP projection of 10,671 T cells deriving from corresponding blood and liver of patients with PSC (n=2). Accordance to the naive DEG signature extracted from the atlas is highlighted. Cells with a positive score were selected as naive cells for further analyses. Enlarged areas highlight tissue origin of selected naive-like CD4⁺ T cells. **b**, Volcano plot of DEG between intrahepatic and peripheral blood naive-like CD4⁺ T cells, indicating high overall similarity of gene expression. Lines indicate logarithmic fold change of expression (|logFC|>0.6). **c**, Violin plot quantifying accordance with gene signature for tissue residency (cf. Table 3). CD4⁺ T_{EM} T_H1/T_H17-state were included as positive control. **d**, Quantification cells that were categorized as resident or circulating based on the residency-score in (**c**). **e**,**f**, Immuno-histochemistry staining of CD4 (**e**) and CD45RA (**f**), respectively, in a liver with PSC. Bile ducts are indicated by asterisks. (**c**) *: p<0.05; **: p<0.01; ***: p<0.001 as determined by Wilcoxon signed-rank test. (in cooperation with Dr. Jenny Krause and Christian Casar)*

Finally, to confirm that there are naive-like $CD4^+$ T cells within tissue, we performed immunohistochemistry. $CD4^+$ and $CD45RA^+$ T cells were found to be rather located in the portal tracts of livers with PSC, which is the site of periductular inflammation and fibrosis (*29, 30*), than within the liver sinusoids (Fig. 16e-f).

These data show that overall gene expression of intrahepatic naive-like $CD4^+$ T cells is, except of genes associated with tissue residency, comparable bona fide naive $CD4^+$ T cells from peripheral blood. Altogether, this paragraph reveals a significant fraction of the population of intrahepatic naive-like $CD4^+$ T cells to be indeed tissue-resident. Therefore, we termed these cells tissue-resident naive-like $CD4^+$ T cells.

3.7. TRAJECTORY INFERENCE IMPLIES DEVELOPMENTAL POTENTIAL OF NAIVE-LIKE $CD4^+$ T CELLS TO DEVELOP INTO T_H17-POLARIZED EFFECTOR CELLS

After identifying a significant fraction of tissue-resident naive-like $CD4^+$ T cells within the inflamed liver of patients with PSC, we aimed to assess their functionality. In order to study the fate of these cells and thus gain functional insight into this population, we utilized trajectory inference algorithms.

Therefore, we extracted and re-clustered the intrahepatic $CD4^+$ T cells from our atlas of PSC patient-derived samples, of which we obtained CITE-Seq data (n=6). Hereby, we identified eight distinct clusters of $CD4^+$ T cells including $CD4^+$ T_{EM}, $CD4^+$ T_{CM}, T_{REG} and naive-like cell populations as stated above for the atlas (Fig. 17a). In addition, we identified a population of $CD4^+$ T cells that expressed the surface marker CCR4 and low amounts of CD25, as well as *CD27* and several integrins, e.g. *ITGB1*, *ITGB2* and genes associated with calcium signalling. Since gene expression hinted adhesion and potential migration, we assigned a transitional cellular state to these cells, which is why we refer to them as "$CD4^+$ transition-state". Cytotoxic effector $CD4^+$ T cells ($CD4^+$ T_C) were identified by expression of granzymes (*GZMK*, *GZMA*) as well as *IL10* and chemokine receptors *CXCR4* and *CXCR6* (Fig. 18a-b). The two remaining clusters were assigned to effector memory cells with a T_H1 (*IFNG*, *TNF*, *TBX21*) and a T_H17 (*CCR6*, *IL17A*, *RORA*, *RORC*) polarization-state ($CD4^+$ T_{EM} T_H1-state and $CD4^+$ T_{EM}

T_H17-state, respectively), which we were likely separated here since we only clustered $CD4^+$ T cells, compared to the atlas.

After identifying the different subsets of $CD4^+$ T cells, we employed Slingshot (*73*) to infer the developmental fate of the tissue-resident naive-like $CD4^+$ T cells, which were defined as the origin of this trajectory analysis.

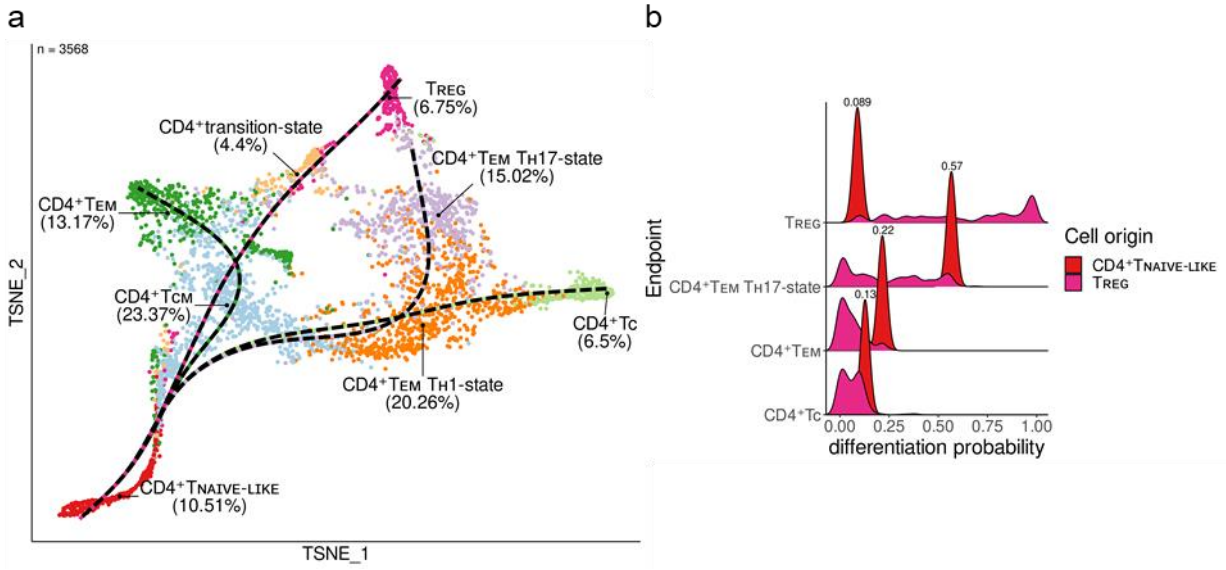

Figure 17: Trajectory inference implies naive-like $CD4^+$ T cells to develop into T_H17-polarized effector cells in livers with PSC.

***a**, Diffusion map based UMAP projection of 3,568 extracted intrahepatic $CD4^+$ T cells from the CITE-Seq dataset of patients with PSC (n=6). Slingshot was applied with naive-like $CD4^+$ T cells defined as starting point and $CD4^+$ effector memory cells in a polarized state towards T_H17 cells ($CD4^+$ T_{EM} T_H17-state), $CD4^+$ effector memory ($CD4^+$ T_{EM}), cytotoxic $CD4^+$ ($CD4^+$ T_C) T cells and regulatory T cells (T_{REG}) as endpoints. **b**, Palantir analysis of the data from (**a**). Endpoints were determined following the results obtained from (**a**). Probabilities of naive-like $CD4^+$ T cells to develop into the defined endpoints were calculated. T_{REG} were included as negative control for differentiation into the other endpoints. Mean probabilities are shown. (in cooperation with Dr. Jenny Krause and Christian Casar)*

Four major developmental trajectories were generated; a trajectory towards $CD4^+$ T_{EM} T_H17-state was found, as well as trajectories towards $CD4^+$ T_{EM}, $CD4^+$ T_C and T_{REG}. Of note, the cluster of $CD4^+$ T_{EM} T_H17-state cells, which has recently been implicated in

PSC pathogenesis (*46, 50–52*), represented 15 % of intrahepatic CD4$^+$ T cells in PSC livers (Fig. 17a). To estimate if the naive-like CD4$^+$ T cells preferably differentiate along one of these trajectories, we further analysed the trajectories by employing Palantir's algorithm (*74*), a Markov chain based approach of estimating differentiation probabilities. Therefore, we selected the endpoint of each trajectory (i.e. CD4$^+$ T_{EM}, CD4$^+$ T_C, CD4$^+$ T_{EM} T_H17-state, T_{REG}) and determined the naive-like CD4$^+$ T cells as developmental starting point. Interestingly, we observed a higher probability for intrahepatic naive-like CD4$^+$ T cells to evolve into cells with a T_H17 polarization-state than into the other endpoints (CD4$^+$ T_{EM}: 12.8 % vs. CD4$^+$ T_C: 21.6 % vs. CD4$^+$ T_{EM} T_H17-state: 56.6 % vs. T_{REG}: 8.94 %) (Fig. 17b).This is of specific interest for PSC, due to recent the implication of T_H17 cells mentioned above. T_{REG} were included in the Palantir analysis as negative control for differentiation capacity, which was determined to be low by the algorithm (Fig. 19a). In line with the findings of preferably developing a T_H17 polarization state, we detected substantial expression of the transcription factors *STAT3* and *ARID5A* within the tissue-resident naive-like CD4$^+$ T cells in PSC, which were previously shown to support the differentiation into T_H17 cells (*75*) (Fig. 19b-c). These results illustrated the potential cell fates of the intrahepatic naive-like CD4$^+$ T cells and suggest that these cells are particularly prone to evolve towards effector cells with a T_H17 polarization-state.

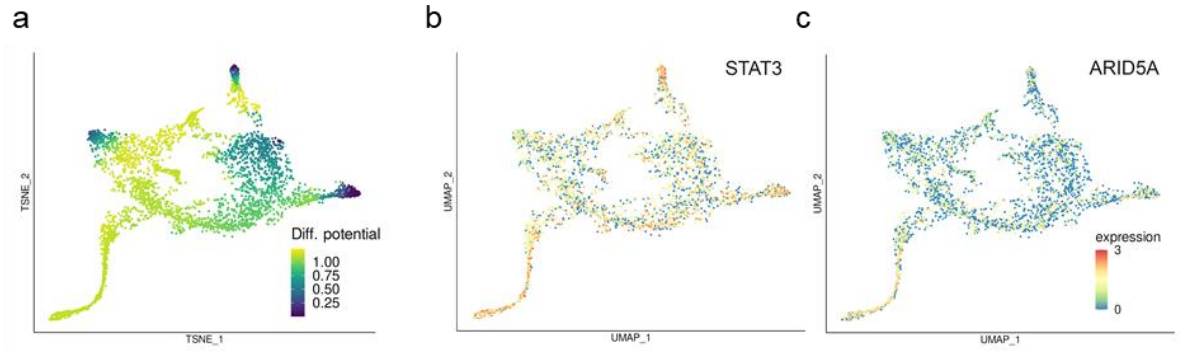

Figure 18: Naive-like CD4$^+$ T cells express transcription factors supporting T_H17 polarization.

***a**, UMAP projection of re-clustered CD4$^+$ T cells displaying estimated differentiation potential of each cell. **b**,**c**, Relative expression of STAT3 (**b**) and ARID5A (**c**) in re-clustered intrahepatic CD4$^+$ T cells. (in cooperation with Dr. Jenny Krause and Christian Casar)*

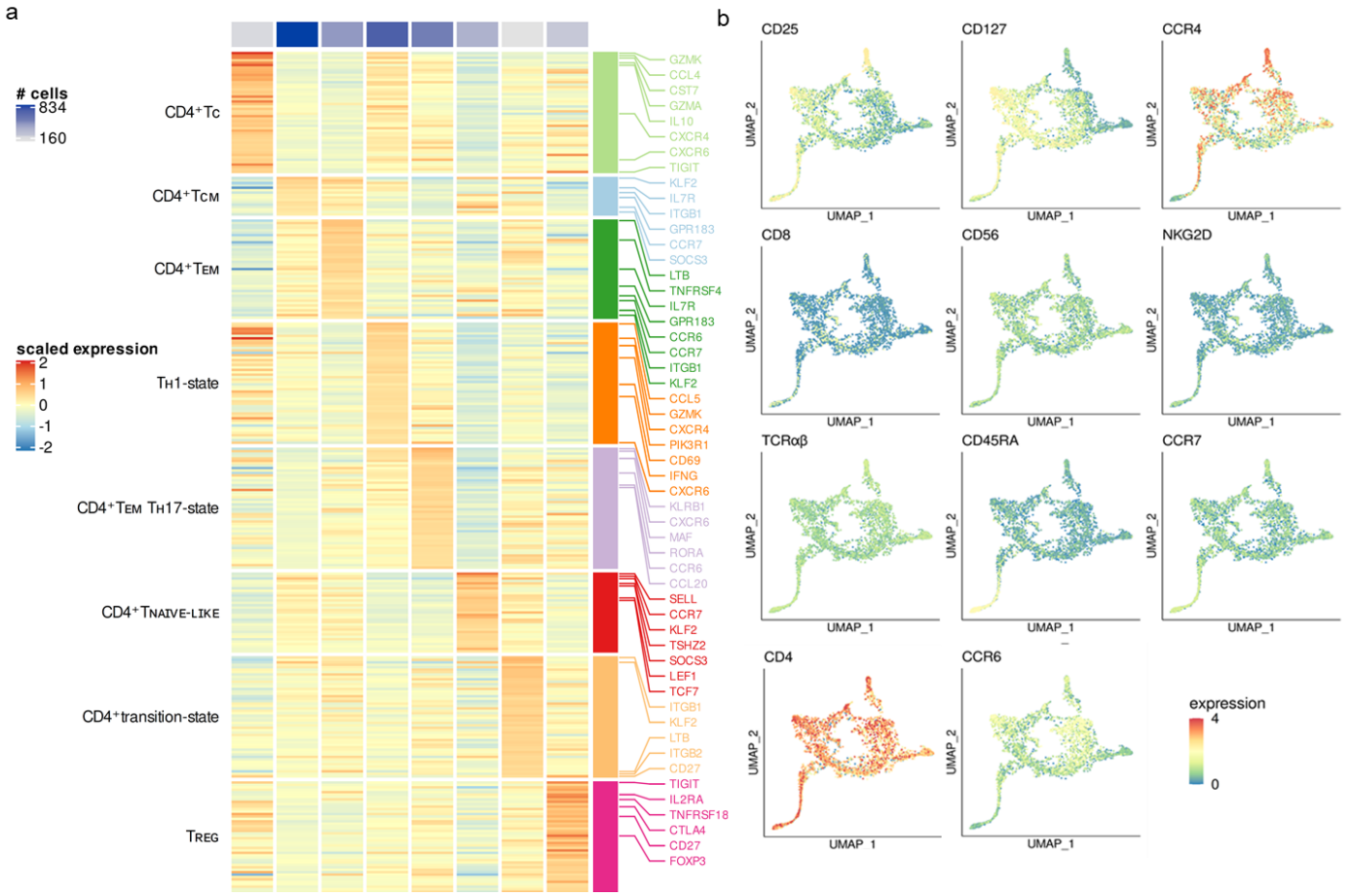

Figure 19: Assignment of cell types to re-clustered CD4+ T cells extracted from the atlas.

***a**, Cluster-Heatmap highlighting signatures of differentially expressed genes (DEG) for each cluster of the atlas. **b**, ADT expression on intrahepatic T cells. Signals were used in combination with gene expression data for determination of different T cell subtypes. T_{CM}, central memory cell; T_{EM}, effector memory cell; T_C, cytotoxic T lymphocyte; T_{REG}, regulatory T cell; T_{EM} T_H1/T_H17-state, effector memory cell with a T_H1 and T_H17 polarization-state. (in cooperation with Dr. Jenny Krause and Christian Casar)*

3.8. FUNCTIONAL EXPERIMENTS REVEALED INCREASED CAPACITY OF NAIVE CD4$^+$ T CELLS TO DEVELOP T_H17-LIKE EFFECTOR FUNCTION IN PSC

To investigate whether the predisposition of naive-like CD4$^+$ T cells to develop into T_H17 cells is even increased in the pathological context of PSC, we conducted functional experiments *in vitro*. Therefore, we isolated naive-like CD4$^+$ T cells from patients with PSC (n=8) and analysed their capacity to differentiate into T_H17 cells compared to sex- and age-matched healthy donors (n=10). Since intrahepatic cells are difficult to isolate in sufficient numbers for *in vitro* experiments, we used blood-derived cells for the functional assays. Importantly, peripheral blood naive CD4$^+$ T cells closely resembled the tissue-resident naive-like CD4$^+$ T cells as shown above by gene expression analysis (cf. Fig. 16c). First, we determined the frequencies of bona fide blood-derived naive CD4$^+$ T cells *ex vivo* and found them to be slightly increased in PSC compared to healthy donors (56.53 % vs. 42.09 %), though not reaching statistical significance (Fig. 20).

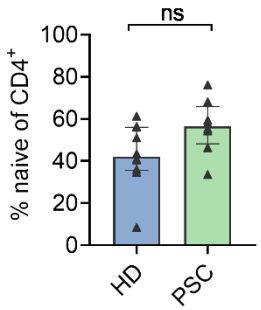

Figure 20: Frequency of peripheral blood naive CD4$^+$ T cells in patients with PSC and HD. *Frequency of naive CD4$^+$ T cells was determined by flow cytometry ex vivo. Data is presented as median with interquartile-range. ns: p>0.05 as assessed by Mann-Whitney U test.*

Second, we investigated the proliferation and differentiation capacity of these cells *in vitro* (Fig. 21a). Interestingly, peripheral blood naive CD4$^+$ T cells from PSC patients showed a higher rate of proliferation compared to healthy donors (47.65 % vs. 25.55 %, p=0.0231), when stimulating with non-polarizing conditions via αCD3/αCD28 alone (Fig. 21b). Furthermore, these cells secreted higher amounts of IL-22 compared to

controls (747.8 pg/ml vs. 570.3 pg/ml ; p=0.0429), a cytokine characteristic of T_H17 cells (*76, 77*) (Fig. 21d). Third, we assessed the capacity of naive $CD4^+$ T cells from PSC patients and healthy donors to develop into T_H17 cells, when stimulated with T_H17-polarizing conditions (αCD3/αCD28 and e.g. IL-1ß, IL-6).

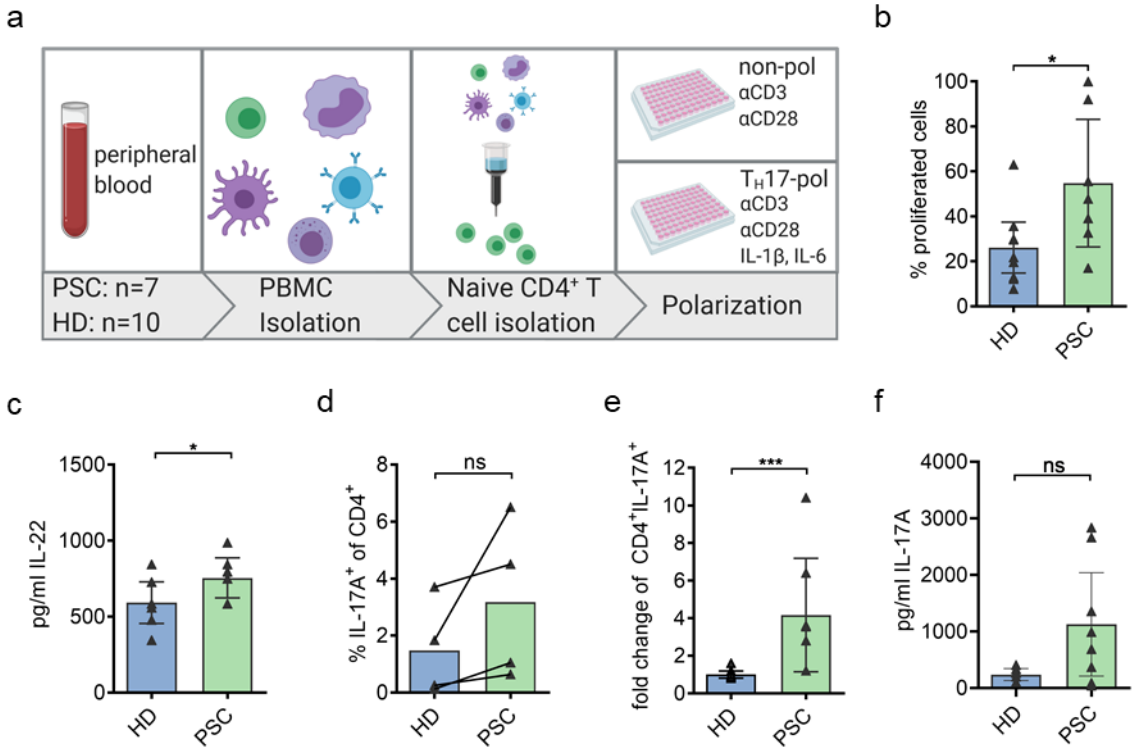

Figure 21: Bona fide naive $CD4^+$ T cells exhibit a qualitative advantage to develop into T_H17-polarized effector cells in PSC.

***a**, Graphical abstract of the workflow of in vitro experiments. The scheme was created with BioRender.com. **b**,**c**, Cell proliferation of naive-like $CD4^+$ T cells (PSC: n=7; HD: n=10). Proliferation (**b**) was determined by flow cytometry and secretion of IL-22 (**c**) analysed by multiplex ELISA on day 7. **d**,**e**,**f**, In vitro culture of naive-like $CD4^+$ T cells under T_H17-polarizing conditions for 12 d (PSC: n=7; HD: n=10). Fold change of frequencies of $CD4^+IL\text{-}17A^+$ cells normalized to the respective controls of each experiment (n=4). (**d**) and non-normalized frequencies of $CD4^+IL\text{-}17A^+$ cells as determined by flow cytometry (**e**). Lines connect respective controls and samples from each experiment. Secretion of IL-17A (**f**) was analysed by multiplex ELISA. Data in (**b**,**c**,**d**,**e**,**f**) is presented as median with interquartile-range. ns: p>0.05; *: p<0.05; **: p<0.01; ***: p<0.001 as assessed by Mann-Whitney U test.*

We detected an increased frequency of IL-17A producing $CD4^+$ T cells in cells deriving from PSC patients compared to cells isolated from healthy controls (fold-change: 3.545 vs. 0.900; p=0.0007). However, we observed a high variance within each group, which is why we normalized each experiment to the respective controls (Fig. 21d-e). In accordance, increased levels of IL-17A were detected in supernatants from PSC-derived cells, however not reaching statistical significance (Fig. 21f). In contrast, the frequencies of IFNγ-producing $CD4^+$ T cells and levels of IFNγ in supernatants were similar between PSC and controls (Data not shown).

Taken together these results confirmed that naive-like $CD4^+$ T cells in patients with PSC are prone to develop a T_H17-polarized phenotype, as suggested by the trajectory inference from liver-derived T cells.

4. DISCUSSION

PSC is a rare progressive cholestatic liver disease of presumed autoimmune pathogenesis that lacks causal treatment options. Several lines of evidence point to the involvement of T cells in disease pathogenesis, including strong HLA associations and an imbalance in regulatory (T_{REG}) and pro-inflammatory (T_H17 cells) cell subsets (*30, 48, 51, 52, 78*). Understanding autoimmune diseases requires the analysis of affected tissue, which harbours infiltrating and tissue-resident immune cells as well as their target cells, and thus displays a disease specific microenvironment. However, few studies analysed the intrahepatic T cell compartment in PSC (*43, 44, 47, 79, 80*), mostly due to the rarity of the disease and the limited availability of tissue specimens.

Within this study, we first applied a comprehensive immunophenotyping approach based on flow cytometry on peripheral blood-derived immune cells. Thereby, we determined the frequencies of several populations belonging to the innate and adaptive parts of the immune system. In this first step, the T cell compartment in particular was detected to bear strong differences among patients with autoimmune liver disease and healthy donors. The biliary diseases PSC and PBC showed a strongly decreased ratio of $CD4^+$ and $CD8^+$ T cells, which is usually ranging between 1.5 - 2.5 (*81, 82*). In contrast, patients with AIH, which is of known autoimmune origin (*83*), showed an enriched ratio between $CD4^+$ and $CD8^+$ T cells. This underlines the distinct phenotype of the biliary diseases, that are both considered to be of autoimmune origin as well (*84*). Of Note, patients with AIH generally showed distinct frequencies of immune cells, when compared to the other AILD. This observation may be due to the common treatment with immunosuppressive azathioprine in these patients (cf. Table 4, which was shown to have an effect on the frequency of T cell subsets in Crohn's disease (*85, 86*).

Since our research focus is PSC and published studies implied T cells as being critical for disease pathogenesis, combined with our findings that peripheral T cells are distinct in biliary disease, we decided to focus on this subset of immune cells for our in-depth analyses.

By combining different, but complementary unbiased approaches such as scRNA-Seq, CITE-Seq and multi-parameter flow cytometry, we provide an in-depth analysis and

comprehensive atlas of the intrahepatic T cell compartment in early and late stage PSC.

We report for the first time on tissue-resident $CD4^+$ T cells with a naive-like phenotype in human inflamed liver. These cells express CCR7 and CD45RA on their surface, which are commonly accepted markers for the identification of naive T cells (*7, 9*) and show high expression of *SELL*, *CCR7*, *TSHZ2*, *TCF7* and *LEF1*, which are genes that are associated with naive T cells (*66, 67, 87, 88*). The naive phenotype of the intrahepatic cells was further confirmed when we demonstrated the lack of clonal expansion.

Within this study, we detected an expansion of the naive-like $CD4^+$ T cells in PSC-livers, when compared to control diseases. However, abundance of these cells was confirmed in the liver tissue of all tested diseases. A major characteristic of bona fide naive T cells is the surface expression of high levels of CCR7 (*9*). Interestingly, it has recently been shown that CCL21, the ligand of CCR7 is highly expressed in livers of patients with PSC (*89*). Of note, the overall expression of CCL21 was increased within PSC, compared to ALD, and particularly strong within the portal tracts of livers with PSC (*79*). CCR7 is associated with migration into secondary lymphoid organs and it has been shown that this receptor promotes retention within lymphoid tissue (*90*). In addition, CCR7 influences the motility and migration speed of T cells into these tissues (*91*). Interestingly, the CCR7/CCL21-axis was identified to be strongly involved in homing of lymphocytes to non-lymphoid tissues, i.e. the lung, and to be critical for recruitment of effector cells to peripheral tissues (*92*). These findings, particularly the correlating expression of CCL21 to increased numbers of naive-like $CD4^+$ T cells in PSC compared to ALD, stimulate speculation on the CCR7/CCL21-axis being involved within homing of these cells to non-lymphoid tissue.

When comparing the overall gene expression of intrahepatic naive-like $CD4^+$ T cells to bona fide naive $CD4^+$ T cells derived from peripheral blood of the same patients, only slight differences could be detected. Notably, what distinguished the intrahepatic cells from the blood-deriving naive T cells was a gene signature typical of tissue residency (*71*), including expression of the accepted hallmark *CD69* (*93*). However, not all intrahepatic naive-like $CD4^+$ T cells displayed accordance to this gene signature, which are likely representing peripheral blood cells that circulate through the liver. Tissue

residency is a feature that is usually assigned to tissue-resident memory T (T_{RM}) cells and is established during the local immune response within non-lymphoid tissues (*94*). Upon antigen encounter, naive T cells expand and gain the function to migrate to non-lymphoid tissues and lose their ability to enter the secondary lymphoid organs (*95*). In contrast to memory T cells, naive T cells usually circulate between blood and lymph nodes and whether they are able to reside in non-lymphoid tissues such as the liver has never been substantiated in humans (*6*). Interestingly, one such population of lymphocytes with a naive phenotype ($CCR7^+CD45RA^+$) was detected before within livers with PSC, but not analysed in-depth (*79*). Within the gene signature associated with tissue residency, we appreciated expression of the chemokine receptor *CXCR6* in intrahepatic naive-like $CD4^+$ T cells, which was shown to support T cell homing to the epithelial surface of the liver, i.e. the biliary epithelium, via its ligand CXCL16 (*69, 70*). In addition, *CXCR4* was also significantly higher expressed in the intrahepatic cells and its ligand, CXCL12, was shown to be expressed by cholangiocytes in inflammatory liver diseases (*96*). This is of specific interest for PSC, since biliary fibrosis is one of the hallmarks of disease and subsets of $CD4^+$ T cells were detected around the bile ducts in livers with PSC (*80*).

In line with this, we detected the population of $CD45RA^+$ cells to be mainly located within the tissue of the portal tracts, including the bile ducts, of livers with PSC. Consequently, we termed these cells tissue-resident naive-like $CD4^+$ T cells. However, since a co-staining in immunohistochemistry was not feasible, we only could show that the $CD4^+$ and $CD45RA^+$ staining was in proximity. To address this issue in the future, we will connect immunohistochemistry to our scRNA-Seq data by spatial transcriptomics, which would enable to align the gene signatures from our single-cell atlas to transcriptomic data *in situ*.

Portal tissue was shown to be the main source of CCL21 in livers with PSC (*79*) and expression of MAdCAM-1, which is the ligand of CD62L, another hallmark of naive T cells, was found within endothelium of portal vessels in livers with PSC (*42*). The findings regarding CCL21 and CD62L combined might partly explain why naive T cells are more frequent in livers with PSC than other liver diseases. In addition to the intrahepatic expansion of T cells with a naive phenotype in PSC, we detected elevated frequencies of bona fide naive $CD4^+$ T cells in the corresponding peripheral blood of

these patients. These data combined indicate a relatively high frequency of naive $CD4^+$ T cells, which was a rather unexpected finding in end-stage liver disease.

One of the PSC-associated SNP outside the HLA-region was detected within the *FOXP1* gene (rs80060485). Interestingly, this SNP is associated with PSC exclusively and occurs at a frequency of approximately 10 % in humans (*97*). However, functional data on this SNP is missing. The transcription factor Foxp1 was shown to be a key molecule in quiescence and maintenance of naive T cells by repressing IL-7 related proliferation and thereby inhibiting differentiation. The transcription factor was therefore termed ‚gatekeeper' of naive T cells (*87, 98, 99*). In line with this, a different study detected Foxp1 to be constitutively repressed in patients with lymphoproliferative disease (*100*). Interestingly, we detected downregulation of the IL-7 receptor (CD127), which is associated with IL-7-mediated stimulation (*101, 102*), on intrahepatic $CD4^+$ T cells from patients with PSC, compared to $CD4^+$ T cells from patients with ALD. In addition to downregulation of CD127, upregulation of CD160 and CD49b, which are markers of activation (*103–105*), was detected to discriminate between PSC- and ALD-derived intrahepatic $CD4^+$ T cells. Taken together, these data indicate a higher frequency of activated and proliferated $CD4^+$ T cells in PSC. In line with this, we detected increased proliferation of bona fide naive $CD4^+$ T cells from peripheral blood of patients with PSC *in vitro*. Beyond interfering with IL-7 signalling, it could be shown that Foxp1 is involved in Foxp3-mediated transcription by supporting binding to the respective loci and is thereby critically involved in the proper development and function of T_{REG} (*106*). Since an imbalance of T_{REG} and T_H17 cells was described in patients with PSC (*30, 48, 51, 52, 78*), it is tempting to speculate on dysfunctional Foxp1 being involved in the described T cell phenotype.

To gain functional insight into the expanded population of naive $CD4^+$ T cells, we utilized trajectory inference. This *in silico* analysis was performed on $CD4^+$ T cells resident in the inflamed liver. Since the current dogma affirms naive T cells to be primed and expanded within lymphoid organs, followed by polarization towards effector cells depending on the microenvironment (*9, 94, 95, 107*), studies on the fate of naive T cells are mostly confined to peripheral blood. We here calculated the propensity of the tissue-resident naive-like $CD4^+$ T cells to develop into the respective endpoints of their trajectories. Although admitting that manually choosing the developmental

endpoints and including all naive-like $CD4^+$ T cells, irrespective of their tissue residency score, might have created bias within the analysis, we detected a strong propensity of these cells to rather develop a T_H17 polarization-state than to differentiate into T_{REG}. Even though fitting well into the published literature on PSC pathogenesis, these results were generated *in silico* and therefore are of limited validity. However, we detected substantial expression of the transcription factors *STAT3* and *ARID5A* within liver-resident naive-like $CD4^+$ T cells of PSC. *ARID5A* was shown to stabilize *STAT3*, a major transcription factor for the differentiation into T_H17 cells, and thereby determining the fate of naive T cells (*75*). This potentially contributes to the finding of liver-resident naive-like $CD4^+$ T cells being prone to T_H17-polarized effector function.

To test the calculated propensity of $CD4^+$ T cells with a naive phenotype to acquire a T_H17 polarization-state, we isolated these cells from peripheral blood and performed functional experiments *in vitro*. The suggested propensity of naive $CD4^+$ T cells from patients with PSC to develop a T_H17-polarized effector function could be confirmed *in vitro*, compared to healthy donors. These results showed a intragroup variance, which is common for human specimens and often overcome with larger patient cohorts, as we were able to show for this experimental setup (*46*). Interestingly, the increased capacity of $CD4^+$ T cells to develop a T_H17 polarization-state was shown not only compared to healthy donors, but to patients with the closely related disease PBC (*46, 51*). One unique predisposition of PSC is the SNP within *FOXP1* described above. These findings stimulate should further research into the potential role of this SNP in the described phenotype. Therefore, the abundance of this SNP should be determined retrospectively and then be correlated to frequencies of naive as well as T_H17-polarized cells and thereby pave the way to functional experiments on proliferation and polarization of naive $CD4^+$ T cells from patients with and without this condition.

Due to the limited numbers of tissue-resident naive-like $CD4^+$ T cells from explant liver tissue, we were forced to use cells deriving from peripheral blood for these experiments. Albeit showing only slight differences in their transcriptomes, cells from peripheral blood are clearly not identical to those residing in the liver, which is another caveat of the study. However, limited numbers of cells due to rare populations and underlying diseases is a frequently faced problem of studies based on solely human samples.

It has previously been shown that microbiota induce higher secretion of IL-17 by PSC-derived $CD4^+$ T cells (*51*), and it has also been shown that T_H17 cells are located around the bile ducts in livers with PSC (*46, 51, 80*). The origin of these cells, however, remained unclear. The diversity of intestinal microbiota was shown to be markedly reduced in patients with PSC, independently from the presence of associated colitis, and the abundance of potentially pro-inflammatory bacterial species was found to be increased (*32, 33*). This is of particular interest, since T_H17 cells have recently been shown to be induced by PSC-derived intestinal microbiota and to contribute to disease severity in mouse models of sclerosing cholangitis (*52*). In line with this, bile from patients with PSC has recently been identified to harbour an altered microbiome with an increase of potential pathogenic bacterial strains, e.g. *Enterococcus faecalis* (*34*), which was shown to induce T_H17 polarization in a previous study (*51*). Further, the interplay between pathogen-stimulated monocytes and cholangiocytes or biliary epithelial cells (BEC) was shown to support a T_H17-polarizing microenvironment, comprising IL-6 and IL-1β, in PSC (*46*). It is therefore tempting to speculate that the PSC-associated microbiota, deriving from bile or gut, shape the T cell response in these patients.

In our experiments, we show that naive-like $CD4^+$ T cells bear a qualitative advantage to develop into T_H17-polarized effector cells, when exposed to an environment containing IL-1β and IL-6. Admitting that this is vague speculation and needs further research, the interplay between BEC and APC mentioned above could in turn promote the polarization of liver-resident naive-like $CD4^+$ T cells within the liver. This is an unconventional concept, since naive T cells usually travel to secondary lymphoid organs or structures and screen APC for their specific antigen. However, evidence for this process was found in studies on liver transplantation in mice (*108*). Hence, these cells could support the inflammatory environment in PSC. Evidence for resident $CD4^+$ T cells promoting hypersensitivity reactions at barrier sites was given (*95*) and commensal microbiota, which are detectable within the bile ducts in PSC (*34, 51*), are widely considered to induce T_H17 cells and thereby promote autoimmunity (*109*).

In summary, this work reports on the first atlas of intrahepatic T cells in PSC and the identification of a previously unrecognised population of tissue-resident naive-like $CD4^+$ T cells in the inflamed liver.

These cells were expanded in PSC and exhibited a propensity to develop into T_H17-polarized effector cells, which renders these cells likely contributors to disease pathogenesis. The single-cell atlas opens the door to understand the contribution of intrahepatic T cells subsets for disease pathogenesis and may reveal novel treatment targets in the future.

5. CONCLUSIONS AND OUTLOOK

Within this work, three important findings were shown that will be highlighted here.
First, by utilizing different single-cell sequencing methods, we provided the first atlas of intrahepatic T cells in the diseased liver. This atlas provides a major data source to the field and needs to be further explored in the future. Consequently, this dataset will generate new hypotheses beyond our findings described here.
Second, we showed evidence for a population of $CD4^+$ T cells with a naive phenotype within our dataset. These cells were shown to reside in non-lymphoid tissue, i.e. the inflamed liver. Although the population of tissue-resident naive T cells has been part of speculations for some time, it never substantiated in human tissue. The evidence provided here should stimulate further research into this interesting immune cell population, which could very well contribute to disease pathogeneses and therefore might be an interesting target for interference with local inflammation processes as present in e.g. AILD.
Third, we detected the population of T cells with a naive phenotype to be expanded in patients with PSC, compared to other liver diseases. This was particularly true for the tissue-resident within the liver, but also detectable within peripheral blood. This finding stimulates into this population with specific regard to PSC, since this disease was associated with several SNP, that are involved in differentiation and maturation of T cells, e.g. in the *FOXP1* gene. As mentioned above, strong evidence for T_H17 cells to be major contributors to PSC pathogenesis was previously shown. The combination of an increased frequency of naive-like $CD4^+$ T cells, the SNP within *FOXP1*, altered biliary microbiota and imbalance of T_H17 and T_{REG} cells forms a basis for future experiments on the interplay of these conditions.

6. METHODS

6.1. PATIENT SELECTION AND ACQUISITION OF BLOOD AND TISSUE SAMPLES

A total of 182 patients who attended the specialized outpatient clinic of the I. Department of Medicine, University Medical Centre Hamburg-Eppendorf (UKE), were included in this study. Diseases were diagnosed and treated following the guidelines of the European Association for the Study of the Liver (EASL) (*110*). Patient characteristics are summarized in Table 4.

We conducted a single-centre cross-sectional study in adult patients (> 18 years of age) undergoing liver transplantation (LTX) for PSC (n=16), ALD (n=16), HCV (n=5), NASH (n=3) or resection for PSC-associated cholangiocarcinoma (PSC-CCA, n=2) and metastasis (LRM, n=9) at the University Medical Center Hamburg-Eppendorf (UKE, Hamburg, Germany). Liver tissue and whole blood were obtained from each patient at the time of surgery. Clinical parameters were recorded immediately before surgery.

AILD patients with early disease stages (PSC: n=41, PBC: n=32, AIH: n=24) were recruited for phenotyping and functional experiments from peripheral blood. Sex and age matched healthy blood donors (HD, n=32) were used as additional controls.

All patients and donors participating in this study provided written informed consent according to the ethical guidelines of the Institutional Review Board of the medical faculty of the University of Hamburg (PV4081).

Table 4: Patient characteristics.

Data is presented as median with interquartile range. HD, healthy donor; LRM, liver resection margin; AIH, autoimmune hepatitis; PBC, primary biliary cholangitis; PSC, primary sclerosing cholangitis; ALD, alcoholic liver disease; HCV, chronic hepatitis C virus infection; NASH, non-alcoholic steatohepatitis; outp., outpatient clinic; LTX, liver transplantation; ALT, alanine aminotransferase; ALP, alkaline phosphatase; BR, Bilirubin; IBD, inflammatory bowel disease; UC, ulcerative colitis; CD, Crohn's disease; UDCA, ursodeoxycholic acid; AZA, azathioprine.

Cohort	Sex (m/f)	Age [a]	Duration of illness [a]	ALT [U/l]	ALP [U/l]	Total BR [mg/dl]	IBD (yes/no)	UC / CD	UDCA / AZA
HD	15/17	47.8 ±6.1	-	-	-	-	0/32	0/0	0/0
LRM	3/6	73 ±6.3	-	25.5 ±7.9	109 ±33.2	0.4 ±0.1	0/9	0/0	0/0
AIH outp.	3/21	66.5 ±8.8	11 ±4	28.5 ±11.1	42 ±33.5	0.6 ±0.3	0/24	0/0	3/13
PBC outp.	2/30	64.5 ±7.6	7 ±5	28 ±9	84 ±30	0.4 ±0.1	0/32	0/0	26/1
PSC outp.	25/16	49.5 ±8.1	7.5 ±2.5	38 ±22.3	120.5 ±75.8	0.8 ±0.4	23/18	5/0	39/2
PSC LTX	12/4	47.5 ±9.2	16 ±7	124.5 ±32.8	428 ±167.5	10.6 ±8.9	10/6	8/2	10/2
ALD LTX	11/5	59.5 ±4.8	5 ±3	30 ±5.5	126 ±37.5	3.4 ±1.7	0/16	0/0	0/0
HCV LTX	5/0	58.0 ±10.3	10 ±0	68 ±47.0	306 ±343.0	1.7 ±15.1	0/5	0/0	0/0
NASH LTX	2/1	56.2 ±6.7	-	29.0 ±117.0	111.0 ±33.0	5.3 ±21.7	0/3	0/0	0/0

6.2. ISOLATION OF MONONUCLEAR CELLS FROM WHOLE BLOOD

Peripheral blood mononuclear cells (PBMC) were isolated from whole blood using Biocoll (Biochrom, Germany) density centrifugation. In brief, whole blood treated with EDTA was diluted approximately 1/1 with 1x PBS. Biocoll was layered under the diluted whole blood with a dilution of 1/3. Centrifugation was carried out at 600g for 20min without brake. Afterwards, PBMC were extracted from the density gradient and washed two times with 1x PBS. Cells were counted and frozen for long-term storage at a defined concentration in FCS containing 10 % DMSO or directly used for further experiments.

6.3. ISOLATION OF MONONUCLEAR CELLS FROM LIVER TISSUE

Liver tissue was cut into small pieces, covered with RPMI medium containing 10 % foetal calf serum (FCS) and 1 % penicillin/streptomycin (P/S).

The pieces were transferred into gentleMACS™ C Tubes (Miltenyi Biotec, Germany) and covered with RPMI 1640 containing 10 % FCS and 1 % P/S, and hashed using the gentleMACS™ Octo Dissociator (Miltenyi Biotec, Germany). The suspension obtained was successively filtered using EASYstrainer™ Cell Sieves (Greiner Bio-One, Germany) with mesh sizes of 300 µm, 200 µm, 100 µm, 70 µm and 40 µm. The purified cells were counted and frozen for long-term storage at a defined concentration in FCS containing 10 % DMSO.

For the purpose of scRNA-Seq, mononuclear cells were further purified by density centrifugation with OptiPrep™ Density Gradient Medium (60 % solution) (Sigma Aldrich, Germany) and afterwards frozen as described above.

6.4. THAWING OF LYMPHOCYTES

Frozen mononuclear cells were taken from long-term storage (LN2) and directly put into a water bath of 37 °C until they were completely thawed. After thawing, cells were quickly diluted in RPMI 1640 containing 10 % FCS and 1 % P/S. After two steps of

washing with the same media, cells were counted and viability was controlled by staining with trypan blue (Sigma Aldrich, Germany) for further experiments.

6.5. MAGNETIC-ACTIVATED CELL SORTING (MACS)

For the purpose of *in vitro* experiments, we used the *Naive CD4⁺ T cell Isolation Kit II* (Miltenyi Biotec, Germany) to enrich naive $CD4^+$ T cells from freshly isolated PBMC. All steps were conducted strictly following the manufacturer's protocol.

6.6. *IN VITRO* PROLIFERATION ASSAY OF NAIVE $CD4^+$ T CELLS

Freshly MACS isolated naive $CD4^+$ T cells ($1x10^6$ cells/ml) were labelled with CellTrace Violet (ThermoFisher Scientific, USA) according to the manufacturer's protocol and cultured with 2 µg/ml plate-bound αCD3 (clone OKT3, Biolegend, USA) and 2 µg/ml soluble αCD28 (clone CD28.2, Biolegend, USA) in RPMI1640 medium (ThemoFisher Scientific, USA), supplemented with 10 % FCS and 1 % P/S (non-polarizing conditions) for up to 7 days. On day 3, 50 % of the cell culture medium including the cytokines and antibodies was substituted. On day 7, dilution of CellTrace Violet in viable cells was analyzed via flow cytometry and multiplex cytokine measurements were conducted (see 7.6).

6.7. *IN VITRO* T_H17 POLARIZATION OF NAIVE $CD4^+$ T CELLS

Freshly MACS isolated naive $CD4^+$ T cells ($1x10^6$ cells/ml) were cultured with 2 µg/ml plate-bound αCD3 (clone OKT3, Biolegend, USA) and 2 µg/ml soluble αCD28 (clone CD28.2, Biolegend, USA) in RPMI1640 medium (ThermoFisher Scientific, USA), supplemented with 10 % FCS and 1 % P/S for 12 days. To polarize the naive cells towards T_H17 phenotype, 30 ng/ml IL-6, 30 ng/ml IL-1β, 30 ng/ml IL-23 and 2.25 ng/ml TGF-β1 (all Miltenyi Biotec, Germany) were supplemented. Additionally, 1 µg/ml αIL-12 (Peprotech, Germany) and 1 µg/ml αIFNγ and 2.5 µg/ml αIL-4 antibodies (both Miltenyi Biotec, Germany) were added to the culture. On days 5 and 8, 50 % of the cell culture medium including the cytokines and antibodies was substituted. On day 12,

supernatants and cells were harvested. Cell were re-stimulated and stained for flow cytometric analysis. Amount of IL-17A expressing $CD4^+$ T cells was determined by flow cytometry with a BD LSRII cytometer (BD Biosciences, USA) and multiplex cytokine measurements were conducted (see 7.6). Discrimination of viable and dead cells was achieved by using either Alexa Fluor™ 750 NHS-Ester or Pacific Orange NHS-Ester (both ThermoFisher Scientific, USA) in a dilution of 1:1000 v/v.

6.8. LEGENDPLEX ASSAY

For multiplex cytokine measurements of cell culture supernatants, we applied the LEGENDplex™ Human Th Panel (13-plex). All steps were conducted following the manufacturer's protocol. Flow cytometry measurements were performed with a BD LSRII cytometer (BD Biosciences, USA).

6.9. STAINING OF LYMPHOCYTES WITH CONJUGATED ANTIBODIES FOR FLOW CYTOMETRY

Cells were counted and split onto a defined number of tubes, washed with 1x PBS and then stained with antibody cocktails for 30min at room temperature. Discrimination of viable and dead cells was achieved using either Alexa Fluor™ 750 NHS-Ester or Pacific Orange NHS-Ester (both ThermoFisher Scientific, USA) in a dilution of 1:1000 v/v.

Afterwards, cells were fixed with 1x RBC Lysis/Fixation Solution (Biolegend, USA) and subsequently washed two times with PBS. After centrifugation, surface-only stained cells were suspended in PBS containing 2 % FCS and 0.01 % NaN_3. The antibodies used are included in Table 6. Samples were acquired using a BD LSRFortessa™ and manual data analysis by classical hierarchical gating was performed using BD FACSDiva™ software version 8 and FlowJo™ v10 (BD, USA). Gating strategies for major innate and adaptive immune cell populations are shown in Figures 22-25.

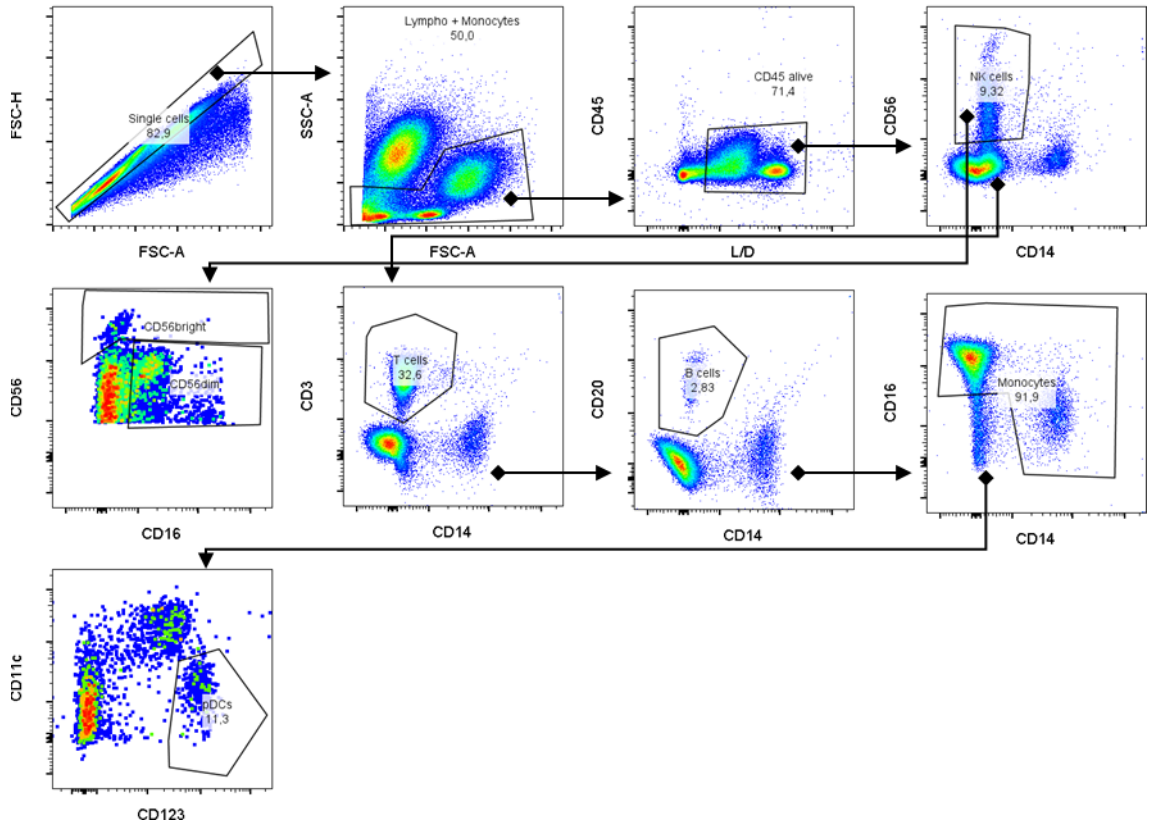

Figure 22: Gating strategy for NK cells, B cells, monocytes and dendritic cells. *After exclusion of cell doublets and debris by Forward and Side Scatter parameters, viable leukocytes were determined by CD45 and live/dead staining (L/D). Subsequently, different populations of immune cells were identified by hierarchical gating.*

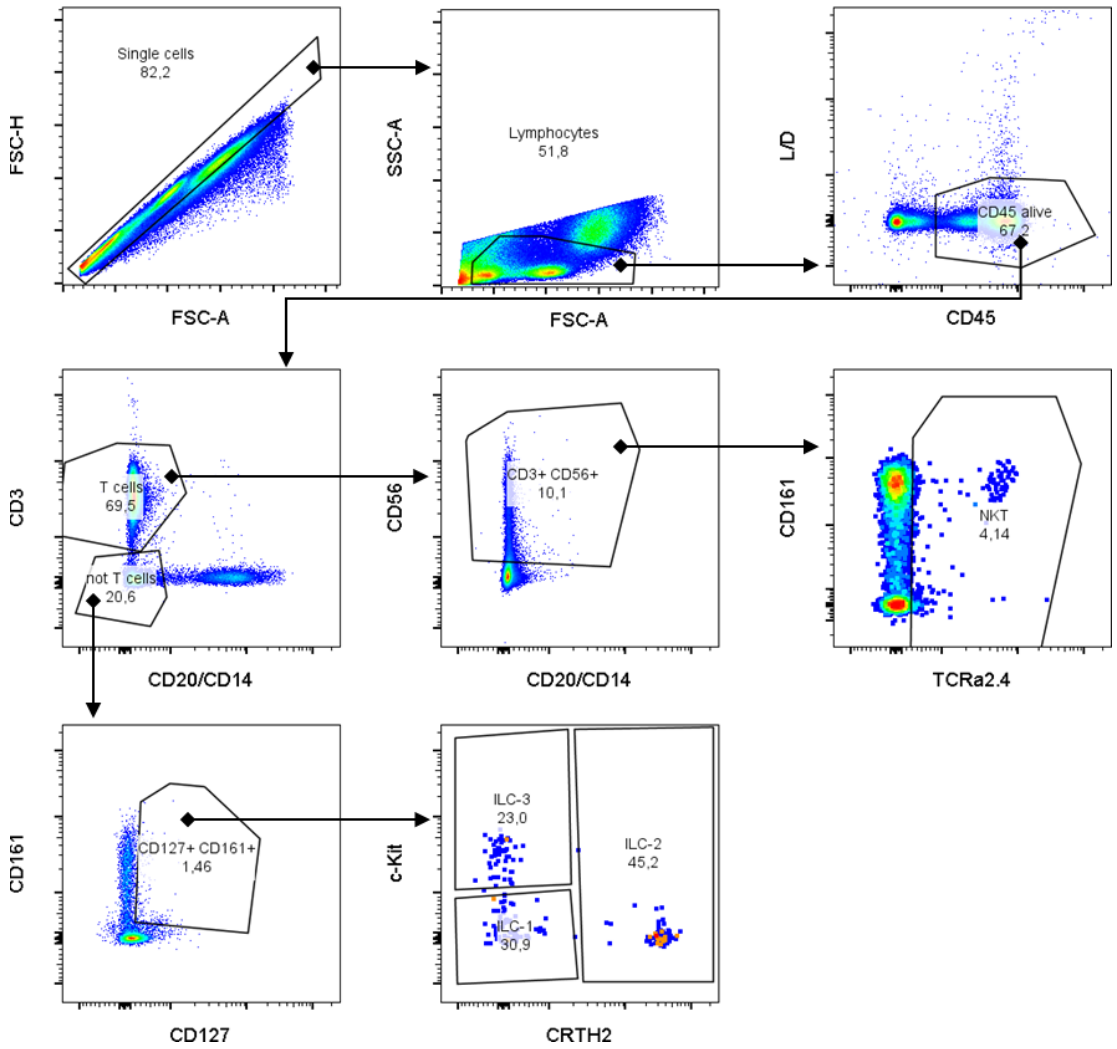

Figure 23: Gating strategy for NKT and innate lymphoid cells.

After exclusion of cell doublets and debris by Forward and Side Scatter parameters, viable leukocytes were determined by CD45 and live/dead staining (L/D). Subsequently, different populations of immune cells were identified by hierarchical gating.

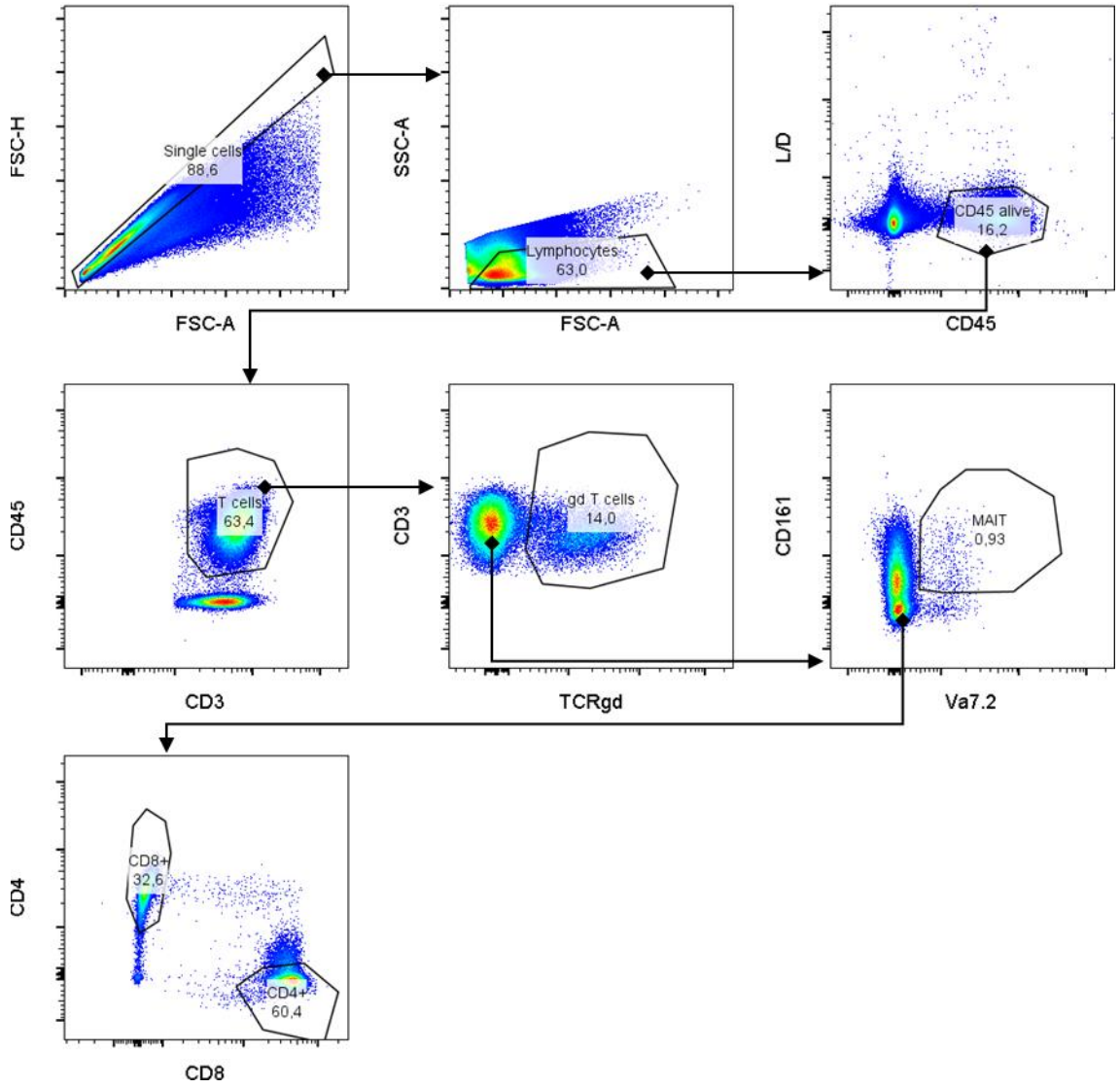

Figure 24: Gating strategy for γδ T cells, MAIT, CD4⁺ and CD8⁺ T cells.

After exclusion of cell doublets and debris by Forward and Side Scatter parameters, viable leukocytes were determined by CD45 and live/dead staining (L/D). Subsequently, different populations of immune cells were identified by hierarchical gating.

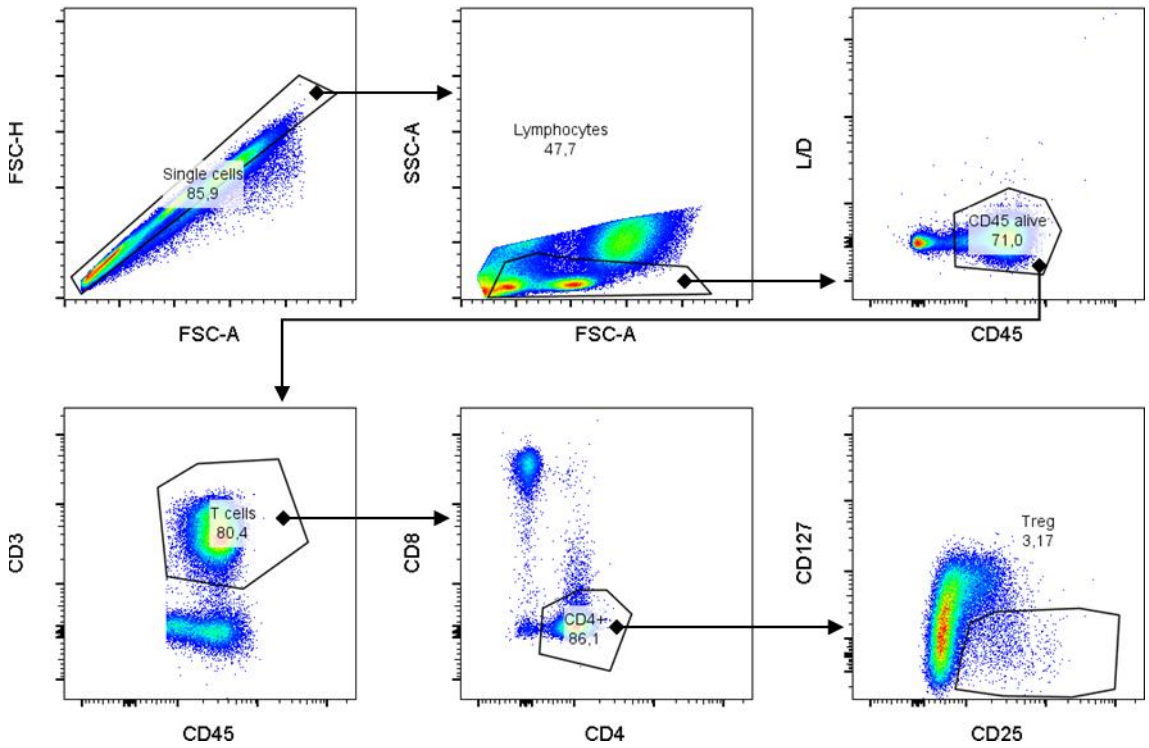

Figure 25: Gating strategy for T_{REG}.

After exclusion of cell doublets and debris by Forward and Side Scatter parameters, viable leukocytes were determined by CD45 and live/dead staining (L/D). Subsequently, different populations of immune cells were identified by hierarchical gating.

6.10. CONJUGATION OF ANTIBODIES FOR CITE-SEQ

The conjugation of antibodies with barcoding oligonucleotides was performed as originally described by Stoeckius *et al*. (*111*). Briefly, 15 µg of each unconjugated antibody (Table 5) was resolved in PBS pH 8.5 and transferred to an Amicon ultra-0.5 centrifugal filter unit (UFC505024) (Millipore, USA) for centrifugal concentration to a volume of 30 µl.

The concentrated antibody solution was mixed with 3 µl of modifier reagent from the LYNX Rapid Streptavidin Antibody Conjugation Kit (LNK163STR) (Bio-Rad, USA) and transferred onto 10 µg of lyophilized streptavidin from the same kit. After overnight incubation at room temperature 3 µl of quencher reagent, 4 µl of 5 M NaCl and 4 µl of

Tween (0.1 % in H_2O) were added, followed by mixture with 800 pmol of biotin-labeled barcoding oligonucleotides (Table 6).

Table 5: Antibodies that were conjugated for CITE-Seq.

Antigen	Clone	Cat. No.	Company
CD196	G034E3	353401	Biolegend
TCR$\alpha\beta$	IP26	306702	Biolegend
CD366	F38-2E2	345001	Biolegend
IL-17A	BL168	512301	Biolegend
CD45	2D1	368502	Biolegend
CD57	HNK-1	359602	Biolegend

The conjugation reaction was again incubated at room temperature overnight. The following day, conjugated antibodies were washed seven times on Amicon filter units with PBS containing 0.5 M NaCl. For storage conjugated antibodies were kept in PBS supplemented with 1 µg/µl BSA and 0.05 % NaN_3. Successful conjugation was confirmed by measuring antibodies on the Qubit 3 fluorometer (Invitrogen, USA) with the dsDNA HS Assay (Q32854) (ThermoFisher, USA).

Conjugation of antibodies (Table 5) for CITE-Seq was performed in cooperation with Dr. Max Kaufmann in the lab of Prof. Manuel Friese (Hamburg, Germany).

Table 6: Barcode sequences of conjugated antibodies for CITE-Seq.

Antigen	Barcode	Sequence (5' - 3')
CD196	CAAAAG	Biotin-CCTTGGCACCCGAGAATTCCACAAAAGBAAAAAAAAAAAAAAAAAAAAAAAAAAAA*A*A
TCRαβ	CAACTA	Biotin-CCTTGGCACCCGAGAATTCCACAACTABAAAAAAAAAAAAAAAAAAAAAAAAAAAA*A*A
CD366	CACCGG	Biotin-CCTTGGCACCCGAGAATTCCACACCGGBAAAAAAAAAAAAAAAAAAAAAAAAAAAA*A*A
IL-17A	CACGAT	Biotin-CCTTGGCACCCGAGAATTCCACACGATBAAAAAAAAAAAAAAAAAAAAAAAAAAAA*A*A
CD45	CAGGCG	Biotin-CCTTGGCACCCGAGAATTCCACAGGCGBAAAAAAAAAAAAAAAAAAAAAAAAAAAA*A*A
CD57	CATGGC	Biotin-CCTTGGCACCCGAGAATTCCACATGGCBAAAAAAAAAAAAAAAAAAAAAAAAAAAA*A*A

6.11. FACS-SORTING OF LIVING T CELLS FOR CITE- AND SCRNA-SEQ

Mononuclear cells were stained with αCD3 for 30 min at 4 °C. Discrimination of living and dead cells was performed at the same time, using Fixable Viability Dye eFluor™ 506 (eBioscience, USA) diluted 1:2000 in PBS. After washing the cells with PBS, cells were resuspended in cold PBS containing EDTA and FCS and immediately FACS-sorted. All steps were carried out at 4 °C or on ice. FACS-sorting was performed on a BD Aria™ Fusion (BD, USA). After gating on single cells via forward- and sideward-scatters, $CD3^+$/eFluor™ 506^- cells were sorted at a flow rate of 1 (Fig. 26).

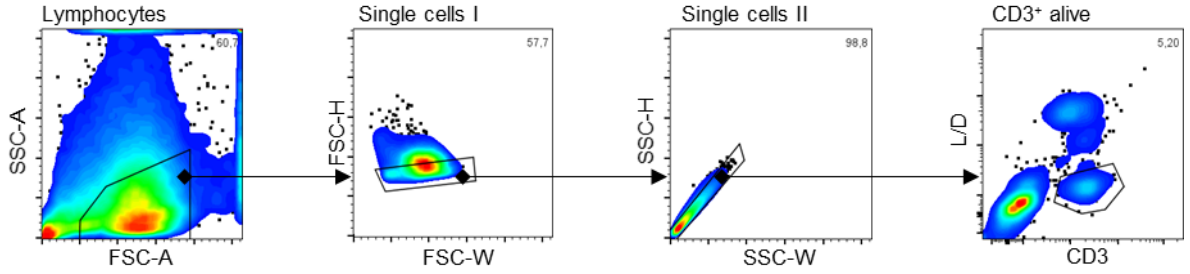

Figure 26: Gating strategy for FACS-sorting of intrahepatic T cells.
After exclusion of cell doublets and debris by Forward and Side Scatter parameters, viable T cells were determined by CD3 and live/dead staining (L/D) and selected for sorting.

6.12. STAINING OF T CELLS WITH CONJUGATED ANTIBODIES FOR CITE-SEQ

In parallel to staining with fluorochrome-conjugated antibodies for FACS-sorting, cells were stained with antibodies for CITE-Seq using TotalSeq A antibodies, (Biolegend, USA) (Table 7) for simultaneous detection of epitopes, following the same protocol as described above.

Table 7: Antibodies used for CITE-Seq.

Antigen	Clone	Cat. No.	Company
CD25	BC96	302643	Biolegend
CD127	A019D5	351352	Biolegend
CD194	L291H4	359423	Biolegend
CD183	G025H7	353745	Biolegend
CD279	EH12.2H7	329955	Biolegend
CD223	11C3C65	369333	Biolegend
CD152	BNI3	369619	Biolegend
TIGIT	A15153G	372725	Biolegend
CD4	SK3	344649	Biolegend
CD8	SK1	344751	Biolegend
HLA-DR	L243	307659	Biolegend
TCR$\gamma\delta$	B1	331229	Biolegend
CD56	5.1H11	362557	Biolegend
CD314	1D11	320835	Biolegend
CD45RA	HI100	304157	Biolegend
CD197	G043H7	353247	Biolegend
CD49b	P1E6-C5	359311	Biolegend

6.13. FACS-SORTING OF LIVING NAIVE $CD4^+$ T CELLS FOR SCTCR-SEQ

Mononuclear cells were stained with αCD3, αCD4, αCD197 and αCD45RA for 30 min at 4 °C. Discrimination of living and dead cells was performed at the same time, using eBioscience™ Fixable Viability Dye eFluor™ 780 (eBioscience, USA) at a dilution of 1:1000 in PBS. After washing the cells with PBS, cells were resuspended in cold Pre-sort buffer and immediately FACS-sorted into cold Post-sort buffer. All steps were carried out at 4 °C or on ice. FACS-sorting was performed on a BD Aria™ Fusion (BD, USA). After gating on single cells via forward- and sideward- scatters, $CD3^+$/eFluor™ 780^- cells were gated on $CD4^+$ and then $CD197^+$/$CD45RA^+$. Cells were sorted with a BD Aria™ Fusion (BD, USA) at a flow rate of 1.

6.14. PREPARATION OF CDNA-LIBRARIES FOR SINGLE-CELL SEQUENCING METHODS

The scRNA-Seq library was prepared using the 10x V3 Single Cell Kit, according to the manufacturer's instructions. FACS-sorted cells were washed once with PBS containing 0.04 % bovine serum albumin (BSA) and resuspended in PBS. 10,000 to 20,000 cells were used for GEM generation through the 10x Chromium Controller using the 10x V3 B Chip GEM generation kit. Droplet preparation was followed by reverse transcription and cell barcoding, the emulsions were resolved and cDNA was purified using Dynabeads MyOne SILANE followed by a PCR amplification (98°C for 45s; 13–18 cycles of 98°C for 20s, 67°C for 30s, 72°C for 1 min; 72°C for 1 min), with additional use of a primer for amplification of antibody derived tags (ADTs). Amplified cDNA was then used for 3′ gene expression library and ADT library construction. For gene expression library construction, amplified cDNA was fragmented and end-repaired, double-sided size-selected using SPRIselect beads. cDNA was further amplified by PCR with sample index primers (98°C for 45s; 14–16 cycles of 98°C for 20s, 54°C for 30s, 72°C for 20s; 72°C for 1 min), and double-sided size-selected with SPRIselect beads. For ADT library construction, antibody derived transcripts were enriched after size selection aliquoting resulting in the initial cDNA amplification step, according to the CITE-seq protocol published by Stoeckius *et al.* (*111*).

Preparation of cDNA-libraries was performed in cooperation with Dr. Jenny Krause in the lab of Prof. Nicola Gagliani (Hamburg, Germany).

6.15. SINGE-CELL RNA SEQUENCING

The scRNA libraries were sequenced on an Illumina NextSeq or HiSeq 4000 to a minimum sequencing depth of 25,000 reads per cell using read lengths of 26 bp read 1, 8 bp i7 index, 98 bp read 2.

Sequencing was performed at Novogene Co., Ltd. (Beijing, China).

6.16. SINGLE-CELL TCR SEQUENCING

The single-cell TCR libraries were sequenced on NovaSeq6000 to a minimum sequencing depth of 5,000 reads per cell using read lengths of 150 bp read 1, 8 bp i7 index, 150 bp read 2.

Sequencing was performed at the Competence Centre for Genomic Analysis (CCGA, Kiel, Germany)

6.17. DATA ANALYSES

6.17.1. SUPERVISED ANALYSIS OF MULTI-PARAMETER FLOW CYTOMETRY DATA

Multi-parameter flow cytometry data was analysed by classical hierarchical gating as well as in an unsupervised manner (see below). Exported frequencies from gating are displayed as median with interquartile range. Tests for statistical significance were conducted by Mann-Whitney U and Kruskal-Wallis tests.

6.17.2. UNSUPERVISED ANALYSIS OF MULTI-PARAMETER FLOW CYTOMETRY DATA

Immune cells with significant differential relative distribution between PSC patients and controls were identified using the random forests based high dimensional feature selection method *Boruta* (*112*). This algorithm iteratively removes the features, which are proved by a statistical test to be less relevant than random probes generated by permutation of the original variable values. The robustness of the identified immune cell signature was assessed by training a random forest classifier using 3-fold cross validation. Discrimination was assessed by means of area under the receiver operating characteristic curve (AUC), specificity and sensitivity. The *ranger* (*113*) package was used via the *caret* machine learning interface (*114*). Statistical significance of the classifier performance was tested by Mason's and Graham's non-parametric test as implemented in the *verification* package (*115, 116*).

Random forest machine learning was performed in cooperation with Dr. Timur Liwinski in the lab of Prof. Eran Elinav (Rehovot, Israel).

6.17.3. SINGLE-CELL SEQUENCING DATA ANALYSIS

All following parts of processing and analysing single-cell sequencing data were performed in cooperation with Christian Casar in the Bioinfomatics Core of the University Medical Center Hamburg-Eppendorf (Hamburg, Germany).

6.17.3.1. SCRNA-SEQ DATA PROCESSING

For the scRNA-seq data, after demultiplexing, reads were aligned against the GRCh38 human reference genome (release 93) and summarized using the Cellranger pipeline (version 3.0.2, 10x Genomics, USA). Further analysis steps were performed in R (version 3.6.2; The R Foundation for Statistical Computing, Vienna, Austria). Due to the use of cell hashing for the frozen ALD and PSC samples additional sample demultiplexing based on the ADT for CD45-BC32 and CD45-MK33 was performed. After centered log ratio transformation, we used the HTODemux function implemented in the Seurat package (version 3.1) (*117*) to separate each sample into donor specific count matrices. The HTODemux function first performs K-medoid clustering (K = #samples + 1) on the normalized ADT expression data. Then for each CD45-tag the cluster with the lowest average expression was determined and a negative binomial distribution was fit to this cluster. Based on these distributions and a quantile of 0.975 each cell was classified into being positive or negative for each CD45-tag. Cells positive for more than one tag were classified as doublets and discarded.

6.17.3.2. QUALITY CONTROL AND NORMALIZATION

For each sample genes not observed in at least 1 % of all cells were dropped. Low quality, damaged or possible doublet cells were excluded using a combination of multiple sample dependent quality measures: minimum UMI count (mean 1234, range 800-2000), minimum and maximum number of expressed genes (mean 756, range

450-1000, and mean 2625, range 2500-3000 respectively), and mitochondrial transcript percentage (mean 6 %, range 5-7.5 %).

For normalization we used Seurat's SCTransform (*118*) function. SCTransform combines the usual scRNA workflow of data normalization, identification of highly variable genes (HVG) and data scaling. Briefly, for each gene a regularized negative binomial regression is performed and the resulting pearson residuals (regression residuals divided by expected standard deviation) represent a variance-stabilizing transformation of the expression data that can be used as normalization for downstream analysis. HVGs are selected based on the highest variance in the pearson residuals.

We opted not to correct for possible cell cycle influence during this process since visual inspection of the first two principal components on cell cycle genes (cc.genes included in Seurat) did not show an obvious bias towards these genes.

These steps resulted in overall 16 samples and 28,301 cells, consisting of 2,427 cells from biobanked PBMC of PSC patients (n=2), 1,654 cells from biobanked intrahepatic T cells of ALD (n=3) and 2,952 cells of biobanked PSC samples (n=5) as well as 14,643 cells from freshly processed liver tissue of patients with PSC (n=4) and 6,625 cells from freshly processed liver tissue of patients with non-cirrhotic PSC (n=2).

6.17.3.3. ADT NORMALIZATION

For each sample where ADT expression data was available, we applied feature-wise centred log ratio transformation implemented in Seurat's NormalizeData function. After rescaling the normalized values to values between -4 and 4 we exported the ADT matrices to Flow Cytometry Standard (FCS) files using the flowCore package (version 1.52.1) for manual gating in FlowJo™ v10 (BD, USA).

6.17.3.4. BATCH CORRECTION

To perform joint analysis on the various combinations of samples dedicated batch correction was necessary since concatenation of the sample data followed by PCA and

UMAP embedding showed obvious batch effects resulting in cell clustering by donor instead of cell types. To compensate for this bias we applied Seurat's gene anchor based batch correction method, using the 2,000 genes that appear in the HVG set of the maximum number of samples.

6.17.3.5. DIMENSION REDUCTION, CLUSTERING AND EXPRESSION ANALYSIS

The batched corrected counts for all integration anchors were used as input for PCA. For UMAP embedding and graph-based shared nearest neighbour clustering we used up to 20 principle components. We iteratively increased the clustering resolution parameter until no additional clusters with biological meaningful cluster markers were detected. Cluster marker detection was performed by differential gene expression analysis for each cluster against all remaining cells using logistic regression on the sample-wise normalized RNA matrices. We included the donor variable as covariate in the regression model and tested for significant differential expression against a null model with likelihood ratio test.

Cluster labels were assigned using a combination of detected marker genes, ADT expression and UMAP overlay of the cell populations identified by manual gating of the ADT expression.

6.17.3.6. CELL SCORE CALCULATION

Cell scores calculated for naive T cells, tissue residency and cell circulation were based on the cluster markers of naive-like $CD4^+$ T cells and the gene sets reported by Kumar *et al.* (*71*). Seurat's AddModuleScore function scores each cell by averaging the expression of each gene set within each cell and subtracting the aggregated expression of a randomly chosen control gene set. For visualization, scores <0 were set to zero 0.

6.17.3.7. TRAJECTORY INFERENCE

Trajectory inference analyses were based on the batch corrected combination of all PSC liver tissue samples (n=6), which were processed immediately after surgery. First, $CD4^+$ cell types were extracted from the whole dataset and the clustering analysis was repeated to define a more refined sub-clustering.

Two dimensional t-SNE embedding based on a dimension reduction by diffusion maps of the batch corrected gene expression as input for slingshot (16) (version 1.5.0) were used to determine cell lineages providing a randomly selected naive-like $CD4^+$ T cell as starting point. As endpoints, the clusters of $CD4^+$ T_{EM}, $CD4^+$ T_{EM} T_H17-state, $CD4^+$ T_C, and T_{REG} were defined.

To estimate if these cells preferably differentiate along one of these trajectories, a Markov chain based approach of estimating differentiation probabilities implemented in the Palantir (*74*) (version 0.2.2) package was used. In brief, a randomly selected naive-like $CD4^+$ T cell was set as starting point for the estimation and random cell of each lineage exclusive cell population identified by slingshot was set as endpoint.

6.17.3.8. COMPARISON OF INTRAHEPATIC AND PERIPHERAL BLOOD T CELLS

For the single-cell expression data from the blood cells we repeated the previously mentioned data processing steps and integrated them with their paired samples from the liver samples. After dimension reduction and clustering we investigated the shared $CD4^+$ naive-like cluster more closely by calculating conserved genes between blood and liver cells within this cluster using Seurat's FindConservedMarkers function. We used the shared population of $CD4^+_{EM}$ T_H17-state cells as comparison group to focus on the contrast to an activated cellular state. The resulting lists of conservedly up- and down-regulated genes within the shared $CD4^+$ naive-like cluster were separately submitted to WebGestalt (*119*) for gene overrepresentation analysis against GO terms with an FDR cutoff < 0.05. For visualization purposes only those gene sets with the 10 highest enrichment ratios for each list were kept.

6.17.3.9. scTCR-SEQ DATA PROCESSING

TCR-Seq data annotation was performed using cellranger vdj (version 3.1.0, 10x Genomics, USA). On average 38,210 reads were sequenced per cell and 88.43 % of the reads were mapped to either the TRA or TRB loci in each cell. 83.29 % of the detected T cells were assigned a TCR. On average, we detected 4,470 unique clonotypes in each patient (range 867 - 8,209), resulting in a mean of 0.97 (range 0.87 - 0.99) for the ratio of inverse Simpson Index to number of cells with productive V-J spanning pairs. Overall, we detected 26,821 clonotypes in 26,892 cells, clonotype sizes ranged from 1 to 11 cells.

7. MATERIAL

Table 8: List of consumables used in this study.

Article	Cat. No.	Company
Tubes 1.5 ml	72.690.001	Sarstedt AG, Germany
Tubes 2 ml	72.691	Sarstedt AG, Germany
Tubes 5 ml	72.701	Sarstedt AG, Germany
Tubes 15 ml	188271	Greiner Bio-One, Germany
Tubes 50 ml	227261	Greiner Bio-One, Germany
FACS Tubes	55.1579	Sarstedt AG, Germany
Cryotubes 1.8 ml	379	Sarstedt AG, Germany
Pipette tips 1 ml	762	Sarstedt AG, Germany
Pipette tips 0.2 ml	760012	Sarstedt AG, Germany
Pipette tips 0.01 ml	1130100	Sarstedt AG, Germany
Pipettes 25 ml	1685020	Sarstedt AG, Germany
Pipettes 10 ml	1688010	Sarstedt AG, Germany
Pipettes 5 ml	1687010	Sarstedt AG, Germany
96-Well Plates	83.3925.500	Sarstedt AG, Germany
gentleMACS C Tubes	130-093-237	Miltenyi Biotec, Germany

Table 9: List of devices used in this study.

Article	Cat. No.	Company
Eppendorf 5920R	5948000010	Eppendorf, Germany
Eppendorf 5810R	5810000010	Eppendorf, Germany
Eppendorf 5427R	5409000010	Eppendorf, Germany
MSC 1.8 workbench	51025430	ThermoFisher, USA
MCO-18AIC CO_2 Incubator	SA-MCO18	Sanyo, Japan
MCO-19AIC CO_2 Incubator	SA-MCO19	Sanyo, Japan
Shaker LS500	11-0002	C. Gerhardt GmbH, Germany
GFL 1083 water bath	1083	GFL GmbH, Germany
Leica DM IRB		Leica Microsystems, Germany
T100 Thermal Cycler	1861096	Bio-Rad, USA
Chromium Controller		10x Genomics, USA
BD™ LSR II		BD, USA
BD LSRFortessa™		BD, USA
BD FACSAria™ Fusion		BD, USA
Pipetboy Acu 2	155000	Integra Biosciences, Germany
Eppendorf Research® plus 0.1 – 2 µl	3123000012	Eppendorf, Germany
Eppendorf Research® plus 0.5 – 10 µl	3124000016	Eppendorf, Germany
Eppendorf Research® plus 10 – 200 µl	3124000083	Eppendorf, Germany
Eppendorf Research® plus 100 – 1000 µl	3124000121	Eppendorf, Germany
Vortex-Genie 2	SI-0236	Scientific Industries, USA
Eppendorf ThermoMixer® Comfort	5382000015	Eppendorf, Germany
gentleMACS Octo Dissociator with Heaters	130-096-427	Miltenyi Biotec, Germany

Table 10: List of antibodies used in this study.

Antigen	Cat. No.	Company
CD103	350218	Biolegend, USA
CD117	313206	Biolegend, USA
CD11c	301641	Biolegend, USA
CD123	306010	Biolegend, USA
CD127	351326	Biolegend, USA
CD127	351322	Biolegend, USA
CD14	301838	Biolegend, USA
CD14	301804	Biolegend, USA
CD16	302018	Biolegend, USA
CD16	302039	Biolegend, USA
CD160	341210	Biolegend, USA
CD161	339916	Biolegend, USA
CD161	339904	Biolegend, USA
CD183	353732	Biolegend, USA
CD19	302252	Biolegend, USA
CD194	359410	Biolegend, USA
CD196	353406	Biolegend, USA
CD197	353218	Biolegend, USA
CD199	358911	Biolegend, USA
CD1c	331520	Biolegend, USA
CD20	302322	Biolegend, USA
CD20	302330	Biolegend, USA
CD20	302303	Biolegend, USA
CD223	369304	Biolegend, USA
CD24	311116	Biolegend, USA
CD25	302630	Biolegend, USA
CD268	316914	Biolegend, USA
CD27	302816	Biolegend, USA
CD272	344511	Biolegend, USA
CD279	329924	Biolegend, USA

Antigen	Cat. No.	Company
CD28	302926	Biolegend, USA
CD294	350125	Biolegend, USA
CD3	317324	Biolegend, USA
CD3	317336	Biolegend, USA
CD38	303512	Biolegend, USA
CD39	328212	Biolegend, USA
CD4	317426	Biolegend, USA
CD4	300548	Biolegend, USA
CD43	343208	Biolegend, USA
CD44	338808	Biolegend, USA
CD45	304036	Biolegend, USA
CD45	304048	Biolegend, USA
CD45RA	304138	Biolegend, USA
CD49a	328312	Biolegend, USA
CD49b	359306	Biolegend, USA
CD49d	304303	Biolegend, USA
CD56	318328	Biolegend, USA
CD57	359619	Biolegend, USA
CD62L	304843	Biolegend, USA
CD69	310914	Biolegend, USA
CD69	310926	Biolegend, USA
CD73	344004	Biolegend, USA
CD8	301028	Biolegend, USA
CD8	301048	Biolegend, USA
CD8	300930	Biolegend, USA
Foxp3	320214	Biolegend, USA
Granzyme B	515403	Biolegend, USA
HLA-DR	307642	Biolegend, USA
HLA-DR	307604	Biolegend, USA
IFNγ	502530	Biolegend, USA
IgD	348203	Biolegend, USA

Antigen	Cat. No.	Company
TCR Vα24-J18	342921	Biolegend, USA
TNFα	502938	Biolegend, USA
TCRγδ	655410	BD, USA
β7-integrin	121010	Biolegend, USA
IL-17A	512328	Biolegend, USA
IL-10	501404	Biolegend, USA
IL-8	511411	Biolegend, USA
IL-6	501112	Biolegend, USA
IL-4	500824	Biolegend, USA
IL-1β	511710	Biolegend, USA
IL-12/IL-23 p40	501822	Biolegend, USA
Vα7.2	351708	Biolegend, USA
CD25	302643	Biolegend, USA
CD127	351352	Biolegend, USA
CD194	359423	Biolegend, USA
CD183	353745	Biolegend, USA
CD279	329955	Biolegend, USA
CD223	369333	Biolegend, USA
CD152	369619	Biolegend, USA
TIGIT	372725	Biolegend, USA
CD4	344649	Biolegend, USA
CD8	344751	Biolegend, USA
HLA-DR	307659	Biolegend, USA
TCR gd	331229	Biolegend, USA
CD56	362557	Biolegend, USA
CD314	320835	Biolegend, USA
CD45RA	304157	Biolegend, USA
CD197	353247	Biolegend, USA
CD49b	359311	Biolegend, USA

Table 11: List of media used in this study.

Name	Cat. No.	Company
RPMI 1640	11875085	ThermoFisher, USA
X-Vivo™ 15	881024	Biozym, Germany
Biocoll	L6113	Biochrom, Germany
OptiPrep™	D1556	Sigma-Aldrich, USA

Table 12: List of commercially available solutions used in this study.

Name	Cat. No.	Company
RBC Lysis/Fixation Solution (10X)	422401	Biolegend, USA
Monensin Solution (1000X)	420701	Biolegend, USA
Brefeldin A Solution (1000X)	00-4506-51	eBioscience, USA
Alexa Fluor™ 750 NHS-Ester	A20011	ThermoFisher, USA
Pacific Orange NHS-Ester	P30253	ThermoFisher, USA
Fixable Viability Dye eFluor™ 506	65-0866-14	eBioscience, USA
CellTrace™ Violet Cell Proliferation Kit	C34557	ThermoFisher, USA

Table 13: List of cell culture supplements used in this study.

Article	Cat. No.	Company
αCD3	317302	Biolegend, USA
αCD28	302902	Biolegend, USA
αIL-4	130-095-753	Biolegend, USA
αIFNγ	130-113-492	Biolegend, USA
αIL-12	500-M12	PeproTech, Germany
IL-23	130-095-757	Miltenyi Biotec, Germany
IL-6	130-095-365	Miltenyi Biotec, Germany
IL-1β	130-093-898	Miltenyi Biotec, Germany
TGF-β1	130-095-066	Miltenyi Biotec, Germany
FCS	A15-101	PAA Laboratories, USA
P/S	15140122	ThermoFisher, USA
DMSO	D4540	Sigma-Aldrich, USA

Table 14: List of commercially available Kits used in this study.

Article	Cat. No.	Company
Naive $CD4^+$ T cell Isolation Kit II, human	130-094-131	Miltenyi Biotec, Germany
LEGENDplex™ Human Th Panel (13-plex)	740722	Biolegend, USA
Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3	PN-1000075	10x Genomics, USA
Chromium Single Cell 5' Library Construction Kit	PN-1000020	10x Genomics, USA

Table 15: Software used in this study.

Software	Company
FlowJo™ v10	BD, USA
BD FACSDiva™ v8	BD, USA
R v3.6.2	The R Foundation for Statistical Computing, Austria
GraphPad PRISM v6	GraphPad Software, USA
GraphPad PRISM v8	GraphPad Software, USA
LEGENDplex™ Software	Biolegend, USA

Table 16: List of buffers and solutions used in this study.

Name	Ingredients	Concentration	Company
PBS	NaCl	137 mM	Carl Roth, Germany
	KCl	2.7 mM	Merck, Germany
	Na_2HPO_4	6.5 mM	Carl Roth, Germany
	KH_2PO_4	1.5 mM	Merck, Germany
FACS-Buffer	EDTA	2 mM	Carl Roth, Germany
	FCS	2 %	Carl Roth, Germany
	NaN_3	0.01 %	Merck, Germany
		in PBS	
MACS-Buffer	EDTA	2 mM	Carl Roth, Germany
	BSA	1 %	Carl Roth, Germany
		in PBS	
Pre-Sort buffer	HEPES	25 mM	Carl Roth, Germany
	EDTA	1 mM	Carl Roth, Germany
	FCS	2 %	
	pH 7.2	in PBS	
Post-sort buffer	HEPES	25 mM	Carl Roth, Germany
	EDTA	1 mM	Carl Roth, Germany
	FCS	20 %	
	pH 7.2	in PBS	
Saponin buffer	Saponin	0.3 %	Sigma-Aldrich, USA
	BSA	0.1 %	Carl Roth, Germany
		in PBS	
PMA-solution	PMA	1 mg/ml	Sigma-Aldrich, USA
		in DMSO	
Ionomycin-solution	Ionomycin	1 mg/ml	Sigma-Aldrich, USA
		in DMSO	
LPS-solution	LPS	5 mg/ml	Sigma-Aldrich, USA
		in PBS	

8. SAFETY

The execution of all experiments for this study was performed in laboratories with biosafety level 1 or 2. Processing of explanted liver tissue was performed in laboratories with biosafety level 3. The usage and disposition of all chemicals accorded to their Material Safety Data Sheets (Table 17).

Table 17: List of hazardous substances used in this study.

Substance	**GHS classification**	**CAS-number**	**H-statements**	**P-statements**
RBC Lysis/Fixation Solution (10X)	Danger	111-46-6 50-00-0	331-302-319-315-317-350-341-335-373	201-202-280-271-260-270-264-272
Monensin Solution (1000X)	Warning	64-17-5	226-319-335-336-373	280-210-241-242-243-233-271-260-264
Brefeldin A Solution (1000X)	Danger	67-56-1	225-301-311-331-370	210-264-280-261-260
CellTrace™ Violet Cell Proliferation Kit	Warning	67-68-5	227	210-280
DMSO	Warning	67-68-5	227	210-280

Substance	***GHS classification***	***CAS-number***	***H-statements***	***P-statements***
HEPES	Warning	7365-45-9	315-319-335	261-264-270-271-280-301+312-304+340-305+351+338-312-321-322-330-332+313-337+313-362-363-403+233-405-501
EDTA	Warning	60-00-4	319	305+351+338
NaN_3	Danger	26628-22-8	300-310-400-410	260-280-301+310-501

9. REFERENCES

1. D. D. Chaplin, Overview of the immune response. *The Journal of allergy and clinical immunology*. **125**, S3-23 (2010), doi:10.1016/j.jaci.2009.12.980.
2. L. B. Nicholson, The immune system. *Essays in biochemistry*. **60**, 275–301 (2016), doi:10.1042/EBC20160017.
3. C. V. Jakubzick, G. J. Randolph, P. M. Henson, Monocyte differentiation and antigen-presenting functions. *Nature reviews. Immunology*. **17**, 349–362 (2017), doi:10.1038/nri.2017.28.
4. J. S. Marshall, R. Warrington, W. Watson, H. L. Kim, An introduction to immunology and immunopathology. *Allergy, asthma, and clinical immunology : official journal of the Canadian Society of Allergy and Clinical Immunology*. **14**, 49 (2018), doi:10.1186/s13223-018-0278-1.
5. K. Murphy, C. Weaver, *Janeway Immunologie* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2018).
6. B. V. Kumar, T. J. Connors, D. L. Farber, Human T Cell Development, Localization, and Function throughout Life. *Immunity*. **48**, 202–213 (2018), doi:10.1016/j.immuni.2018.01.007.
7. Y. D. Mahnke, T. M. Brodie, F. Sallusto, M. Roederer, E. Lugli, The who's who of T-cell differentiation. *European journal of immunology*. **43**, 2797–2809 (2013), doi:10.1002/eji.201343751.
8. N. D. Pennock *et al.,* T cell responses: naive to memory and everything in between. *Advances in Physiology Education*. **37**, 273–283 (2013), doi:10.1152/advan.00066.2013.
9. T. van den Broek, J. A. M. Borghans, F. van Wijk, The full spectrum of human naive T cells. *Nature reviews. Immunology*. **18**, 363–373 (2018), doi:10.1038/s41577-018-0001-y.
10. J. Zhu, W. E. Paul, CD4 T cells. *Blood*. **112**, 1557–1569 (2008), doi:10.1182/blood-2008-05-078154.
11. S. J. Szabo, B. M. Sullivan, S. L. Peng, L. H. Glimcher, M OLECULAR M ECHANISMS R EGULATING T H 1 I MMUNE R ESPONSES. *Annual review of immunology*. **21**, 713–758 (2003), doi:10.1146/annurev.immunol.21.120601.140942.

12. J. A. Walker, A. N. J. McKenzie, TH2 cell development and function. *Nat Rev Immunol*. **18**, 121–133 (2018), doi:10.1038/nri.2017.118.
13. T. Nakayama *et al.,* Th2 Cells in Health and Disease. *Annual review of immunology*. **35**, 53–84 (2017), doi:10.1146/annurev-immunol-051116-052350.
14. S. Bhaumik, R. Basu, Cellular and Molecular Dynamics of Th17 Differentiation and its Developmental Plasticity in the Intestinal Immune Response. *Front. Immunol.* **8**, 254 (2017), doi:10.3389/fimmu.2017.00254.
15. S. Sakaguchi, T. Yamaguchi, T. Nomura, M. Ono, Regulatory T cells and immune tolerance. *Cell*. **133**, 775–787 (2008), doi:10.1016/j.cell.2008.05.009.
16. N. Zhang, M. J. Bevan, CD8(+) T cells. *Immunity*. **35**, 161–168 (2011), doi:10.1016/j.immuni.2011.07.010.
17. A. Toubal, I. Nel, S. Lotersztajn, A. Lehuen, Mucosal-associated invariant T cells and disease. *Nature reviews. Immunology*. **19**, 643–657 (2019), doi:10.1038/s41577-019-0191-y.
18. D. I. Godfrey, K. J.L. Hammond, L. D. Poulton, M. J. Smyth, A. G. Baxter, NKT cells. *Immunology Today*. **21**, 573–583 (2000), doi:10.1016/S0167-5699(00)01735-7.
19. D. I. Godfrey, S. P. Berzins, Control points in NKT-cell development. *Nat Rev Immunol*. **7**, 505–518 (2007), doi:10.1038/nri2116.
20. M. M. Nielsen, D. A. Witherden, W. L. Havran, γδ T cells in homeostasis and host defence of epithelial barrier tissues. *Nat Rev Immunol*. **17**, 733–745 (2017), doi:10.1038/nri.2017.101.
21. M. Lawand, J. Déchanet-Merville, M.-C. Dieu-Nosjean, Key Features of Gamma-Delta T-Cell Subsets in Human Diseases and Their Immunotherapeutic Implications. *Front. Immunol.* **8**, 467 (2017), doi:10.3389/fimmu.2017.00761.
22. L. Klein, B. Kyewski, P. M. Allen, K. A. Hogquist, Positive and negative selection of the T cell repertoire. *Nature reviews. Immunology*. **14**, 377–391 (2014), doi:10.1038/nri3667.
23. D. L. Mueller, Mechanisms maintaining peripheral tolerance. *Nat Immunol*. **11**, 21–27 (2010), doi:10.1038/ni.1817.
24. P. K. Gregersen, L. M. Olsson, Recent advances in the genetics of autoimmune disease. *Annual review of immunology*. **27**, 363–391 (2009), doi:10.1146/annurev.immunol.021908.132653.

25. L. Wang, F.-S. Wang, M. E. Gershwin, Human autoimmune diseases. *Journal of internal medicine*. **278**, 369–395 (2015), doi:10.1111/joim.12395.
26. J. H. Tabibian, A. H. Ali, K. D. Lindor, Primary Sclerosing Cholangitis, Part 1. *Gastroenterology & hepatology*. **14**, 293–304 (2018).
27. M. K. Washington, Autoimmune liver disease. *Mod Pathol*. **20**, S15-S30 (2007), doi:10.1038/modpathol.3800684.
28. M. Carbone, J. M. Neuberger, Autoimmune liver disease, autoimmunity and liver transplantation. *Journal of hepatology*. **60**, 210–223 (2014), doi:10.1016/j.jhep.2013.09.020.
29. K. N. Lazaridis, N. F. LaRusso, Primary Sclerosing Cholangitis. *The New England journal of medicine*. **375**, 1161–1170 (2016), doi:10.1056/NEJMra1506330.
30. T. H. Karlsen, T. Folseraas, D. Thorburn, M. Vesterhus, Primary sclerosing cholangitis - a comprehensive review. *Journal of hepatology*. **67**, 1298–1323 (2017), doi:10.1016/j.jhep.2017.07.022.
31. R. Liberal, C. R. Grant, Cirrhosis and autoimmune liver disease. *World journal of hepatology*. **8**, 1157–1168 (2016), doi:10.4254/wjh.v8.i28.1157.
32. M. Kummen *et al.,* The gut microbial profile in patients with primary sclerosing cholangitis is distinct from patients with ulcerative colitis without biliary disease and healthy controls. *Gut*. **66**, 611–619 (2017), doi:10.1136/gutjnl-2015-310500.
33. M. C. Rühlemann *et al.,* Faecal microbiota profiles as diagnostic biomarkers in primary sclerosing cholangitis. *Gut*. **66**, 753–754 (2017), doi:10.1136/gutjnl-2016-312180.
34. T. Liwinski *et al.,* Alterations of the bile microbiome in primary sclerosing cholangitis. *Gut*. **69**, 665–672 (2020), doi:10.1136/gutjnl-2019-318416.
35. V. Pontecorvi, M. Carbone, P. Invernizzi, The "gut microbiota" hypothesis in primary sclerosing cholangitis. *Annals of translational medicine*. **4**, 512 (2016), doi:10.21037/atm.2016.12.43.
36. E. Halilbasic, C. Fuchs, H. Hofer, G. Paumgartner, M. Trauner, Therapy of Primary Sclerosing Cholangitis--Today and Tomorrow. *Digestive diseases (Basel, Switzerland)*. **33 Suppl 2**, 149–163 (2015), doi:10.1159/000440827.
37. E. Melum *et al.,* Genome-wide association analysis in primary sclerosing cholangitis identifies two non-HLA susceptibility loci. *Nature genetics*. **43**, 17–19 (2011), doi:10.1038/ng.728.

38. T. H. Karlsen *et al.,* Genome-Wide Association Analysis in Primary Sclerosing Cholangitis. *Gastroenterology*. **138**, 1102–1111 (2010), doi:10.1053/j.gastro.2009.11.046.
39. J. Z. Liu *et al.,* Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nature genetics*. **45**, 670–675 (2013), doi:10.1038/ng.2616.
40. J. M. Banales *et al.,* Cholangiocyte pathobiology. *Nat Rev Gastroenterol Hepatol*. **16**, 269–281 (2019), doi:10.1038/s41575-019-0125-y.
41. X.-M. Chen, S. P. O'Hara, N. F. LaRusso, The immunobiology of cholangiocytes. *Immunology and cell biology*. **86**, 497–505 (2008), doi:10.1038/icb.2008.37.
42. A. J. Grant, P. F. Lalor, S. G. Hübscher, M. Briskin, D. H. Adams, MAdCAM-1 expressed in chronic inflammatory liver disease supports mucosal lymphocyte adhesion to hepatic endothelium (MAdCAM-1 in chronic inflammatory liver disease). *Hepatology (Baltimore, Md.)*. **33**, 1065–1072 (2001), doi:10.1053/jhep.2001.24231.
43. B. Eksteen *et al.,* Hepatic endothelial CCL25 mediates the recruitment of CCR9+ gut-homing lymphocytes to the liver in primary sclerosing cholangitis. *The Journal of experimental medicine*. **200**, 1511–1517 (2004), doi:10.1084/jem.20041035.
44. B. Eksteen *et al.,* Gut homing receptors on CD8 T cells are retinoic acid dependent and not maintained by liver dendritic or stellate cells. *Gastroenterology*. **137**, 320–329 (2009), doi:10.1053/j.gastro.2009.02.046.
45. E. K. K. Henriksen *et al.,* Gut and liver T-cells of common clonal origin in primary sclerosing cholangitis-inflammatory bowel disease. *Journal of hepatology*. **66**, 116–122 (2017), doi:10.1016/j.jhep.2016.09.002.
46. L. K. Kunzmann *et al.,* Monocytes as potential mediators of pathogen-induced Th17 differentiation in patients with primary sclerosing cholangitis (PSC). *Hepatology (Baltimore, Md.)* (2020), doi:10.1002/hep.31140.
47. E. Liaskou *et al.,* Loss of CD28 expression by liver-infiltrating T cells contributes to pathogenesis of primary sclerosing cholangitis. *Gastroenterology*. **147**, 221-232.e7 (2014), doi:10.1053/j.gastro.2014.04.003.
48. M. Sebode *et al.,* Reduced FOXP3(+) regulatory T cells in patients with primary sclerosing cholangitis are associated with IL2RA gene polymorphisms. *Journal of hepatology*. **60**, 1010–1016 (2014), doi:10.1016/j.jhep.2013.12.027.

49. R. Liberal *et al.,* Regulatory T-cell conditioning endows activated effector T cells with suppressor function in autoimmune hepatitis/autoimmune sclerosing cholangitis. *Hepatology*. **66**, 1570–1584 (2017), doi:10.1002/hep.29307.
50. R. Liberal *et al.,* CD39 mediated regulation of Th17-cell effector function is impaired in juvenile autoimmune liver disease. *Journal of autoimmunity*. **72**, 102–112 (2016), doi:10.1016/j.jaut.2016.05.005.
51. J. Katt *et al.,* Increased T helper type 17 response to pathogen stimulation in patients with primary sclerosing cholangitis. *Hepatology (Baltimore, Md.)*. **58**, 1084–1093 (2013), doi:10.1002/hep.26447.
52. N. Nakamoto *et al.,* Gut pathobionts underlie intestinal barrier dysfunction and liver T helper 17 cell immune response in primary sclerosing cholangitis. *Nature microbiology*. **4**, 492–503 (2019), doi:10.1038/s41564-018-0333-1.
53. K. A. Aho, *Foundational and applied statistics for biologists using R* (CRC Press, Boca Raton, FL, 2014).
54. A. M. Abel, C. Yang, M. S. Thakar, S. Malarkannan, Natural Killer Cells. *Front. Immunol.* **9**, 1869 (2018), doi:10.3389/fimmu.2018.01869.
55. A. Cossarizza *et al.,* Guidelines for the use of flow cytometry and cell sorting in immunological studies. *European journal of immunology*. **47**, 1584–1797 (2017), doi:10.1002/eji.201646632.
56. M. Guilliams, A. Mildner, S. Yona, Developmental and Functional Heterogeneity of Monocytes. *Immunity*. **49**, 595–613 (2018), doi:10.1016/j.immuni.2018.10.005.
57. R. D. Abeles *et al.,* CD14, CD16 and HLA-DR reliably identifies human monocytes and their subsets in the context of pathologically reduced HLA-DR expression by CD14(hi) /CD16(neg) monocytes. *Cytometry. Part A : the journal of the International Society for Analytical Cytology*. **81**, 823–834 (2012), doi:10.1002/cyto.a.22104.
58. M. Collin, N. McGovern, M. Haniffa, Human dendritic cell subsets. *Immunology*. **140**, 22–30 (2013), doi:10.1111/imm.12117.
59. E. Vivier *et al.,* Innate Lymphoid Cells. *Cell*. **174**, 1054–1066 (2018), doi:10.1016/j.cell.2018.07.017.
60. K. Böttcher *et al.,* MAIT cells are chronically activated in patients with autoimmune liver disease and promote profibrogenic hepatic stellate cell activation. *Hepatology*. **68**, 172–186 (2018), doi:10.1002/hep.29782.

61. E. Seth *et al.,* Primary sclerosing cholangitis leads to dysfunction and loss of MAIT cells. *Eur. J. Immunol.* **48**, 1997–2004 (2018), doi:10.1002/eji.201847608.
62. M. Egeblad, E. S. Nakasone, Z. Werb, Tumors as Organs. *Developmental Cell*. **18**, 884–901 (2010), doi:10.1016/j.devcel.2010.05.012.
63. A. Kurioka, L. J. Walker, P. Klenerman, C. B. Willberg, MAIT cells. *Clinical & translational immunology*. **5**, e98 (2016), doi:10.1038/cti.2016.51.
64. L. J. Walker *et al.,* Human MAIT and CD8αα cells develop from a pool of type-17 precommitted CD8+ T cells. *Blood*. **119**, 422–433 (2012), doi:10.1182/blood-2011-05-353789.
65. J. E. Ussher, P. Klenerman, C. B. Willberg, Mucosal-Associated Invariant T-Cells. *Front. Immunol.* **5**, 1907 (2014), doi:10.3389/fimmu.2014.00450.
66. P. A. Szabo *et al.,* Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nature communications*. **10**, 4706 (2019), doi:10.1038/s41467-019-12464-3.
67. S. Yu *et al.,* The TCF-1 and LEF-1 transcription factors have cooperative and opposing roles in T cell development and malignancy. *Immunity*. **37**, 813–826 (2012), doi:10.1016/j.immuni.2012.08.009.
68. F. Sallusto, J. Geginat, A. Lanzavecchia, Central memory and effector memory T cell subsets. *Annual review of immunology*. **22**, 745–763 (2004), doi:10.1146/annurev.immunol.22.012703.104702.
69. P. Ramachandran *et al.,* Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature*. **575**, 512–518 (2019), doi:10.1038/s41586-019-1631-3.
70. M. Heydtmann *et al.,* CXC chemokine ligand 16 promotes integrin-mediated adhesion of liver-infiltrating lymphocytes to cholangiocytes and hepatocytes within the inflamed human liver. *Journal of immunology (Baltimore, Md. : 1950)*. **174**, 1055–1062 (2005), doi:10.4049/jimmunol.174.2.1055.
71. B. V. Kumar *et al.,* Human Tissue-Resident Memory T Cells Are Defined by Core Transcriptional and Functional Signatures in Lymphoid and Mucosal Sites. *Cell reports*. **20**, 2921–2934 (2017), doi:10.1016/j.celrep.2017.08.078.
72. R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, A. Regev, Spatial reconstruction of single-cell gene expression data. *Nature biotechnology*. **33**, 495–502 (2015), doi:10.1038/nbt.3192.

73. K. Street *et al.,* Slingshot. *BMC genomics*. **19**, 477 (2018), doi:10.1186/s12864-018-4772-0.
74. M. Setty *et al.,* Characterization of cell fate probabilities in single-cell data with Palantir. *Nature biotechnology*. **37**, 451–460 (2019), doi:10.1038/s41587-019-0068-4.
75. K. Masuda *et al.,* Arid5a regulates naive CD4+ T cell fate through selective stabilization of Stat3 mRNA. *The Journal of experimental medicine*. **213**, 605–619 (2016), doi:10.1084/jem.20151289.
76. I. Backert *et al.,* STAT3 activation in Th17 and Th22 cells controls IL-22-mediated epithelial host defense during infectious colitis. *Journal of immunology (Baltimore, Md. : 1950)*. **193**, 3779–3791 (2014), doi:10.4049/jimmunol.1303076.
77. Y. Chung *et al.,* Expression and regulation of IL-22 in the IL-17-producing CD4+ T lymphocytes. *Cell research*. **16**, 902–907 (2006), doi:10.1038/sj.cr.7310106.
78. T. H. Karlsen, M. Vesterhus, K. M. Boberg, Review article. *Alimentary pharmacology & therapeutics*. **39**, 282–301 (2014), doi:10.1111/apt.12581.
79. A. J. Grant *et al.,* Hepatic Expression of Secondary Lymphoid Chemokine (CCL21) Promotes the Development of Portal-Associated Lymphoid Tissue in Chronic Inflammatory Liver Disease. *The American Journal of Pathology*. **160**, 1445–1455 (2002), doi:10.1016/S0002-9440(10)62570-9.
80. Y. H. Oo *et al.,* CXCR3-dependent recruitment and CCR6-mediated positioning of Th-17 cells in the inflamed liver. *Journal of hepatology*. **57**, 1044–1051 (2012), doi:10.1016/j.jhep.2012.07.008.
81. J. A. McBride, R. Striker, Imbalance in the game of T cells: What can the CD4/CD8 T-cell ratio tell us about HIV and health? *PLoS Pathogens*. **13**, e1006624 (2017), doi:10.1371/journal.ppat.1006624.
82. A. Amadori *et al.,* Genetic control of the CD4/CD8 T-cell ratio in humans. *Nature medicine*. **1**, 1279–1283 (1995), doi:10.1038/nm1295-1279.
83. K.-H. zum Meyer Büschenfelde, Autoimmunity and Liver Disease. *Seminars in liver disease*. **11**, iii–v (1991), doi:10.1055/s-2008-1040489.
84. K. Boonstra, U. Beuers, C. Y. Ponsioen, Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *Journal of hepatology*. **56**, 1181–1188 (2012), doi:10.1016/j.jhep.2011.10.025.

85. A. Cossu, L. Biancone, M. Ascolani, F. Pallone, M. Boirivant, "In vitro" azathioprine-induced changes in peripheral T cell apoptosis and IFN-γ production associate with drug response in patients with Crohn's Disease. *Journal of Crohn's & colitis*. **7**, 441–450 (2013), doi:10.1016/j.crohns.2012.06.020.
86. N. E. McCarthy *et al.,* Azathioprine therapy selectively ablates human Vδ2$^+$ T cells in Crohn's disease. *The Journal of clinical investigation*. **125**, 3215–3225 (2015), doi:10.1172/JCI80840.
87. P. Durek *et al.,* Epigenomic Profiling of Human CD4+ T Cells Supports a Linear Differentiation Model and Highlights Molecular Regulators of Memory Development. *Immunity*. **45**, 1148–1161 (2016), doi:10.1016/j.immuni.2016.10.022.
88. T. Willinger *et al.,* Human naive CD8 T cells down-regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in vivo. *Journal of immunology (Baltimore, Md. : 1950)*. **176**, 1439–1446 (2006), doi:10.4049/jimmunol.176.3.1439.
89. A. E. Langeneckert *et al.,* CCL21-expression and accumulation of CCR7+ NK cells in livers of patients with primary sclerosing cholangitis. *European journal of immunology*. **49**, 758–769 (2019), doi:10.1002/eji.201847965.
90. T. H.M. Pham, T. Okada, M. Matloubian, C. G. Lo, J. G. Cyster, S1P1 Receptor Signaling Overrides Retention Mediated by Gαi-Coupled Receptors to Promote T Cell Egress. *Immunity*. **28**, 122–133 (2008), doi:10.1016/j.immuni.2007.11.017.
91. R. Förster, A. C. Davalos-Misslitz, A. Rot, CCR7 and its ligands. *Nature reviews. Immunology*. **8**, 362–371 (2008), doi:10.1038/nri2297.
92. J. C. Lo *et al.,* Differential regulation of CCL21 in lymphoid/nonlymphoid tissues for effectively attracting T cells to peripheral tissues. *J. Clin. Invest.* **112**, 1495–1505 (2003), doi:10.1172/JCI19188.
93. P. A. Szabo, M. Miron, D. L. Farber, Location, location, location. *Sci. Immunol.* **4**, eaas9673 (2019), doi:10.1126/sciimmunol.aas9673.
94. J. M. Schenkel, D. Masopust, Tissue-resident memory T cells. *Immunity*. **41**, 886–897 (2014), doi:10.1016/j.immuni.2014.12.007.

95. D. Masopust, A. G. Soerens, Tissue-Resident T Cells and Other Resident Leukocytes. *Annual review of immunology*. **37**, 521–546 (2019), doi:10.1146/annurev-immunol-042617-053214.
96. O. Wald *et al.,* Involvement of the CXCL12/CXCR4 pathway in the advanced liver disease that is associated with hepatitis C virus or hepatitis B virus. *Eur. J. Immunol.* **34**, 1164–1174 (2004), doi:10.1002/eji.200324441.
97. S.-G. Ji *et al.,* Genome-wide association study of primary sclerosing cholangitis identifies new risk loci and quantifies the genetic relationship with inflammatory bowel disease. *Nature genetics*. **49**, 269–273 (2017), doi:10.1038/ng.3745.
98. X. Feng *et al.,* Transcription factor Foxp1 exerts essential cell-intrinsic regulation of the quiescence of naive T cells. *Nature immunology*. **12**, 544–550 (2011), doi:10.1038/ni.2034.
99. H. Wang *et al.,* The transcription factor Foxp1 is a critical negative regulator of the differentiation of follicular helper T cells. *Nat Immunol*. **15**, 667–675 (2014), doi:10.1038/ni.2890.
100. S. Garaud *et al.,* FOXP1 is a regulator of quiescence in healthy human CD4+ T cells and is constitutively repressed in T cells from patients with lymphoproliferative disorders. *European journal of immunology*. **47**, 168–179 (2017), doi:10.1002/eji.201646373.
101. F. Carrette, C. D. Surh, IL-7 signaling and CD127 receptor regulation in the control of T cell homeostasis. *Seminars in immunology*. **24**, 209–217 (2012), doi:10.1016/j.smim.2012.04.010.
102. A. Vranjkovic, A. M. Crawley, K. Gee, A. Kumar, J. B. Angel, IL-7 decreases IL-7 receptor alpha (CD127) expression and induces the shedding of CD127 by human CD8+ T cells. *International immunology*. **19**, 1329–1339 (2007), doi:10.1093/intimm/dxm102.
103. G. Cai *et al.,* CD160 inhibits activation of human CD4+ T cells through interaction with herpesvirus entry mediator. *Nature immunology*. **9**, 176–185 (2008), doi:10.1038/ni1554.
104. G. Cai, G. J. Freeman, The CD160, BTLA, LIGHT/HVEM pathway. *Immunological reviews*. **229**, 244–258 (2009), doi:10.1111/j.1600-065X.2009.00783.x.

105. G. Kassiotis, D. Gray, Z. Kiafard, J. Zwirner, B. Stockinger, Functional specialization of memory Th cells revealed by expression of integrin CD49b. *Journal of immunology (Baltimore, Md. : 1950)*. **177**, 968–975 (2006), doi:10.4049/jimmunol.177.2.968.

106. C. Konopacki, Y. Pritykin, Y. Rubtsov, C. S. Leslie, A. Y. Rudensky, Transcription factor Foxp1 regulates Foxp3 chromatin binding and coordinates regulatory T cell function. *Nature immunology*. **20**, 232–242 (2019), doi:10.1038/s41590-018-0291-z.

107. F. Sallusto, Heterogeneity of Human CD4(+) T Cells Against Microbes. *Annual review of immunology*. **34**, 317–334 (2016), doi:10.1146/annurev-immunol-032414-112056.

108. I. Klein, I. N. Crispe, Complete differentiation of CD8+ T cells activated locally within the transplanted liver. *The Journal of experimental medicine*. **203**, 437–447 (2006), doi:10.1084/jem.20051775.

109. W. E. Ruff, M. A. Kriegel, Autoimmune host-microbiota interactions at barrier sites and beyond. *Trends in molecular medicine*. **21**, 233–244 (2015), doi:10.1016/j.molmed.2015.02.006.

110. EASL Clinical Practice Guidelines. *Journal of hepatology*. **51**, 237–267 (2009), doi:10.1016/j.jhep.2009.04.009.

111. M. Stoeckius *et al.,* Simultaneous epitope and transcriptome measurement in single cells. *Nature methods*. **14**, 865–868 (2017), doi:10.1038/nmeth.4380.

112. M. B. Kursa, W. R. Rudnicki, Feature Selection with the Boruta Package. *J. Stat. Soft.* **36** (2010), doi:10.18637/jss.v036.i11.

113. M. N. Wright, A. Ziegler, ranger. *J. Stat. Soft.* **77** (2017), doi:10.18637/jss.v077.i01.

114. M. Kuhn, Building Predictive Models in R Using the caret Package. *J. Stat. Soft.* **28** (2008), doi:10.18637/jss.v028.i05.

115. N. Manubens *et al.,* An R package for climate forecast verification. *Environmental Modelling & Software*. **103**, 29–42 (2018), doi:10.1016/j.envsoft.2018.01.018.

116. S. J. Mason, N. E. Graham, Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves. *Q. J. R. Meteorol. Soc.* **128**, 2145–2166 (2002), doi:10.1256/003590002320603584.

117. T. Stuart *et al.,* Comprehensive Integration of Single-Cell Data. *Cell*. **177**, 1888-1902.e21 (2019), doi:10.1016/j.cell.2019.05.031.

118. C. Hafemeister, R. Satija, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome biology*. **20**, 296 (2019), doi:10.1186/s13059-019-1874-1.

119. Y. Liao, J. Wang, E. J. Jaehnig, Z. Shi, B. Zhang, WebGestalt 2019. *Nucleic Acids Research*. **47**, W199-W205 (2019), doi:10.1093/nar/gkz401.

10. SCIENTIFIC COOPERATIONS

This study was thankfully supported by several scientific cooperations, which is why the contributions of each cooperation is specified below.
Sample acquisition and sample processing for scRNA-Seq as well as execution of scRNA-Seq experiments was performed in cooperation with Dr. Jenny Krause[1]. Labelling of antibodies for CITE-Seq was performed in cooperation with Dr. Max Kaufmann[2]. Bioinformatic analysis of scRNA-Seq data was performed in cooperation with Christian Casar[1,3]. Bioinformatic analysis of flow cytometry data was performed in cooperation with Dr. Timur Liwinski[4].

Affiliations:

[1]I. Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[2]Institute of Neuroimmunology and Multiple Sclerosis, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[3]Bioinformatics Core, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[4]Immunology Department, Weizmann Institute of Science, Rehovot, Israel

11. DANKSAGUNGEN

Zuerst möchte ich Christoph Schramm und Dorothee Schwinge für die sehr gute Betreuung, das sehr gute Arbeitsklima, die konstante Unterstützung und die konstruktive Kritik über die letzten vier Jahre danken. Ich hätte mir keinen besseren Ort für die Durchführung dieser Arbeit aussuchen können und freue mich sehr darüber, noch etwas bei euch zu bleiben. Lieber Christoph, liebe Doro, Danke für alles! :)

Als nächstes möchte ich mich bei Johannes Herkel und Antonella Carambia bedanken. Danke für die offenen Ohren, das stetige Feedback und das nette Miteinander! Für das nette Miteinander danke ich natürlich allen Doktoranden, technischen Angestellten und Wissenschaftlern im schönen dritten Obergeschoss von O58.

Für die Unterstützung und Hilfe, gerade in der Anfangszeit dieser Arbeit, vor allem beim Etablieren unserer umfangreichen FACS-panels, möchte ich mich herzlich bei Eva Tolosa bedanken.

Für die großartige wissenschaftliche Kooperation und das dabei sehr nette Umfeld bedanke ich mich ganz herzlich bei Jenny Krause, Christian Casar und Nicola Gagliani. For Nico I switch to english here. Thank you for the constant questions, comments and for teaching me the meaning of resilience. This really pushed the project to a new level!

Wissenschaftliche Kooperation ist das Stichwort, um dem *liver team* der Arbeitsgruppe von Marcus Altfeld zu danken. Danke für die sehr produktive und angenehme Zusammenarbeit zu sehr unangenehmen Uhrzeiten!

Da nun ich nun hoffentlich alle Arbeitskollegen erwähnt habe, möchte ich den Dank an meine kleine Familie richten. Liebe Annika, vielen Dank für die konstante Unterstützung, das Verständnis für durcharbeitete Nächte, die offenen Ohren für Unmut und Euphorie, das Einbringen von Ideen und den starken Rückhalt. Lieber Malte, vielen Dank, dass du deinen Papa ständig auf Trab hältst und dadurch von der Arbeit ablenkst, sobald er Zuhause ist. Ich habe euch sehr lieb!

Zu guter Letzt möchte ich noch meinem Großvater für die nette Wohngemeinschaft über die letzten Jahre danken.

In diesem Sinne: Vielen Dank, Thanks a lot, Grazie mille, Muchas gracias!