

Characterization of CCR8⁺ regulatory T cells in human tumors and healthy tissue

Dissertation

Zur Erlangung des Grades

Dr. rer. nat.

Am Fachbereich Biologie
der Johannes Gutenberg-Universität Mainz

vorgelegt von

Kathrin Luise Braband

geb. am 25.05.1992 in Bad Bergzabern

Mainz, 2024

Dekan:

Prof. Dr. Eckhard Thines

1. Berichterstatter:

Prof. Dr. Michael Delacher

2. Berichterstatter:



Tag der mündlichen Prüfung:

6. Mai 2024

Declaration:

Parts of this work have been published or are currently being prepared for publication: We have published one review (1), which part of the introduction is inspired by, and which **Figure 3** was printed from. scATAC-seq and scRNA/TCR-seq workflows established in our labs have been published in methods articles (2, 3), which explanations of the respective methods, as well as **Figures 5 and 6** in this dissertation are based on. We are further planning on publishing data presented in the results section of this dissertation.

INDEX

1. INTRODUCTION	1
1.1 The innate and adaptive immune system work in synergy	1
1.2 Regulatory T cells are the guardians of peripheral tolerance	2
1.3 Description of a Treg cell lineage that promotes tissue regeneration.....	3
1.4 CCR8 ⁺ Treg cells can promote tissue regeneration	6
1.4.1 Tissue Treg cells promote tissue homeostasis under steady state in the VAT, skin, and liver.....	7
1.4.2 Tissue Treg cells repair tissue damage	8
1.4.3 Treg cells are present in the tumor microenvironment.....	10
1.5 Aim of the project.....	11
1.6 Methods.....	13
1.6.1 Investigating chromatin accessibility on single-cell level using scATAC-seq.....	13
1.6.2 Investigating gene expression on single-cell level combined with T cell receptor sequences using scRNA/TCR-seq.....	14
1.7 Description of tumor types studied in this thesis	15
1.7.1 Clear cell renal carcinoma	15
1.7.2 Renal angiomyolipoma	16
1.7.3 Hepatocellular carcinoma	16
1.7.4 Colorectal carcinoma	16
1.7.5 Lung carcinoma	17
1.7.6 Gastric adenocarcinoma	17
1.7.7 Oral cavity squamous cell carcinoma (OCSCC).....	17
1.7.8 Urothelial carcinoma	18
1.7.9 Seminoma.....	18
1.7.10 Liver metastases.....	18
2. MATERIALS	19
2.1 Equipment.....	19
2.2 Consumables.....	19
3. METHODS.....	22
3.1 Human tissue.....	22

3.2	Isolation of immune cells from human tissues.....	23
3.3	Restimulation for scRNA/TCR-seq experiments.....	25
3.4	Cell sorting.....	25
3.5	scATAC-seq.....	26
3.5.1	scATAC-seq library preparation.....	26
3.5.2	Sequencing.....	28
3.5.3	Computation.....	28
3.5.4	Alignment of reads.....	28
3.5.5	Pre-processing of scATAC-seq data.....	29
3.5.5.1	Creating the count matrix.....	29
3.5.5.2	Quality filtering.....	30
3.5.5.3	Dimensionality reduction and clustering.....	32
3.5.5.4	Doublet filtering.....	34
3.5.6	scATAC-seq data analysis.....	35
3.5.6.1	Cell type annotation.....	35
3.5.6.2	Differential peaks and motif enrichment.....	39
3.5.6.3	Browser tracks.....	41
3.6	scRNA/TCR-seq.....	42
3.6.1	scRNA/TCR-seq library preparation.....	42
3.6.2	Sequencing.....	42
3.6.3	Computation.....	43
3.6.4	Alignment of reads.....	43
3.6.5	Pre-processing of scRNA/TCR-seq data.....	44
3.6.5.1	Creating the count matrix.....	44
3.6.5.2	Quality filtering.....	46
3.6.5.3	Doublet filtering.....	48
3.6.5.4	Dimensionality reduction and clustering.....	49
3.6.6	scRNA/TCR-seq data analysis.....	52
3.6.6.1	Cell type annotation.....	52
3.6.6.2	Differential gene expression analysis.....	54
3.6.6.3	TCR clonotype analysis.....	55
4.	RESULTS.....	58
4.1	CCR8 ⁺ Treg cells are present in human tumor- and healthy tissue.....	58
4.2	CCR8 ⁺ Treg cells isolated from kidney tumors share molecular characteristics with CCR8 ⁺ Treg cells isolated from NAT.....	60

4.3	CCR8 ⁺ Treg cells isolated from liver tumors share molecular characteristics with CCR8 ⁺ Treg cells isolated from NAT	64
4.4	CCR8 ⁺ Treg cells isolated from a variety of solid tumors share molecular characteristics with CCR8 ⁺ Treg cells isolated from the corresponding NAT	66
4.5	CCR8 ⁺ Treg cells isolated from metastases to the liver share molecular characteristics with CCR8 ⁺ Treg cells isolated from healthy liver tissue	69
4.6	CCR8 ⁺ Treg cells isolated from CRC and metastasis to the liver in the same patient share molecular characteristics	73
4.7	CCR8 ⁺ Treg cells from tumor and NAT share gene expression profiles	75
5.	DISCUSSION	78
5.1	Treg cells exert a non-canonical function in tissue physiology	78
5.2	Treg cell infiltration into the tumor correlates with decreased overall survival.....	79
5.3	The relevance of CCR8 ⁺ Treg cells across tumor types is unknown.....	79
5.4	Current approaches targeting “tumor Treg” cells.....	80
5.5	CCR8 ⁺ Treg cells are present in tumor and normal adjacent tissue.....	81
5.6	CCR8 ⁺ Treg cells from tumor and tumor-free tissue share epigenetic features	82
5.7	CCR8 ⁺ Treg cells from tumor and tumor-free tissue share gene expression	83
5.8	Understanding CCR8 ⁺ Treg cell dynamics	83
5.9	Implications for cancer therapy targeting CCR8 ⁺ Treg cells	84
6.	ABSTRACT	86
7.	ZUSAMMENFASSUNG.....	88
8.	APPENDIX	I
8.1	References	I
8.2	Figures.....	XII
8.3	Tables	XIII
8.4	Code	XIII

9. CV.....	II
10. DANKSAGUNG.....	VI

Abbreviations

AC	Adenocarcinoma
AEC2	Type II alveolar epithelial cell
AhR	Aryl hydrocarbon receptor
AML	Angiomyolipoma
AMP	Adenosine monophosphate
APC	Antigen presenting cells
Areg	Amphiregulin
ASC	Adenosquamous carcinoma
ATAC-seq	Assay for transposase-accessible chromatin with sequencing
ATP	Adenosine triphosphate
BATF	Basic leucine zipper transcription factor, ATF-like
bcl	Binary base call
Bcl2	B-cell lymphoma 2
BCR	B cell receptor
bp	Base pairs
bZIP	Basic leucine zipper
CAF	Cancer-associated fibroblasts
cAMP	Cyclic adenosine monophosphate
CCL	C-C motif chemokine ligand
CCR	C-C motif chemokine receptor
CCRC	Clear cell renal carcinoma
CD	Cluster of differentiation
CRC	Colorectal cancer
Ctla-4	Cytotoxic T-lymphocyte associated protein 4
CX3CR	C-X3-C motif chemokine receptor
CXCL	C-X-C motif chemokine ligand
DC	Dendritic cell
DEG	Differentially expressed gene
DNA	Desoxyribonucleic acid
EAE	Experimental autoimmune encephalomyelitis
EGFR	Epidermal growth factor receptor
ELF2	E74-like factor 2
EMT	Epithelial-mesenchymal transition
ENTPD1	Ectonucleoside triphosphate diphosphohydrolase 1
EOMES	Eomesodermin

FACS	Fluorescence activated cell sorting
FAP	Fibro/ adipogenic progenitors
Fas-L	Fas ligand
FDR	False discovery rate
FGFBP	Fibroblast growth factor binding protein
FGFR3	Fibroblast growth factor receptor 3
FOXP3	Forkhead box protein 3
GAC	Gastric adenocarcinoma
Gata3	GATA binding protein 3
GEX	Gene expression
GI	Gastrointestinal
GZMB	Granzyme B
HAVCR2	Hepatitis A virus cellular receptor 2
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HFSC	Hair follicle stem cell
HIF	Hypoxia-inducible factor
HPV	Human papilloma virus
ICOS	Inducible T cell costimulator
Id3	Inhibitor of DNA binding 3
IFN	Interferon
Ikzf	IKAROS family zinc finger
IL	Interleukin
Il1lr1	Interleukin 1 receptor-like 1
Il2ra	Interleukin 2 receptor subunit alpha
ILC2	Type 2 innate lymphoid cells
IPEX	Immunodysregulation polyendocrinopathy enteropathy X-linked
Irf	Interferon regulatory factor
iTreg	<i>In vitro</i> induced Treg
Jag1	Jagged1
Kgf	Keratinocyte growth factor
KLF	Kruppel-like factor
Klrg1	Killer cell lectin-like receptor G1
kNN	k-nearest neighbor
LAYN	Layilin

LSI	Latent semantic indexing
LTA	Lymphotoxin alpha
mAb	Monoclonal antibody
MAD	Median absolute deviation
Maf	Musculoaponeurotic fibrosarcoma
MDSC	Myeloid-derive suppressor cells
MET	Metastasis
MIBC	Muscle-invasive bladder cancer
mRNA	Messenger ribonucleic acid
mTOR	Mechanistic target of rapamycin
NASH	Non-alcoholic steatohepatitis
NFKB	Nuclear factor 'kappa-light-chain-enhancer' of activated B-cells
NR4A1	Nuclear receptor subfamily 4 group A member 1
NSCLC	Non-small cell lung cancer
OCSCC	Oral cavity squamous cell carcinoma
PENK	Proenkephalin
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
PRDM1	PR domain zinc finger protein 1
PCR	Polymerase chain reaction
PRR	Pattern recognition receptor
Pdcd1	Programmed cell death 1
PMA	Phorbol-12-myristate-13-acetate
pTreg	Peripherally induced Treg
QC	Quality control
RNA	Ribonucleic acid
RNA/TCR-seq	RNA sequencing with TCR profiling
ROR	RAR-related orphan receptor
RUNX	Runt-related transcription factor
Satb1	SATB homeobox 1
SCC	Squamous cell carcinoma
SCE	SingleCellExperiment object
Sell	Selectin L
sc	Single-cell
Sparc	Secreted acidic cysteine-rich glycoprotein
ST2	Suppression of tumorigenicity 2

STAT	Signal transducer and activator of transcription
TAM	Tumor-associated macrophage
TBR	T-box brain transcription factor
TBX	T-box transcription factor
TCF	T cell factor
Tconv	Conventional T (cell)
TCR	T cell receptor
Teff	Effector T (cell)
Texh	Exhausted T (cell)
TERT	Telomerase reverse transcriptase
TGCT	Testicular germ cell tumor
Th	T helper
TF	Transcription factor
TF	Term frequency
Tfh	Follicular helper T (cell)
TF-IDF	Term frequency – inverse document frequency
TIM3	Transmembrane immunoglobulin and mucin domain 3
TME	Tumor microenvironment
Tnf	Tumor necrosis factor
Tnfsf	TNF superfamily
TNFRSF	TNF receptor superfamily
TOX	Thymocyte selection-associated high mobility group box protein
TP53	Tumor protein p53
TRAIL	TNF-related apoptosis-inducing ligand
Treg	Regulatory T (cell)
TSS	Transcription start site
tTreg	Thymus-derived Treg
TWGBS	Tagmentation-based whole-genome bisulfite sequencing
UC	Urothelial carcinoma
UMAP	Uniform manifold approximation and projection
UMI	Unique molecular identifier
UV	Ultraviolet
VAT	Visceral adipose tissue
VHL	Von Hippel-Lindau
ZEB	Zinc finger E-box binding homeobox

1. INTRODUCTION

1.1 The innate and adaptive immune system work in synergy

The immune system is a complex network of cells and tissues that work together to defend the body against harmful invaders, such as bacteria or viruses, fungi, and other pathogens. Its primary function is to recognize and eliminate these foreign substances, as well as endogenous transformed or malignant cells, thereby maintaining the body's overall health and integrity.

The immune system can be broadly categorized into two main branches: innate immunity and adaptive immunity. Innate immunity serves as the body's first line of defense and provides immediate, non-specific protection against a wide range of pathogens. It includes physical barriers like the skin and mucosa, as well as various immune cells such as neutrophils and macrophages, which are able to recognize surface structures common to many bacteria via pattern recognition receptors (PRR) and destroy the invading pathogens. They are also able to recruit additional immune cells to the site of infection. However, some bacteria, which have evolved to evade the innate response by shielding such structures from the immune system, as well as many viruses, cannot be detected in this first line of defense. Dendritic cells (DC), which are innate immune cells specialized in surveying peripheral tissues for pathogens, are often responsible for initiating an adaptive immune response by presenting antigens of engulfed pathogens to T cells. These, in turn, can mount a more sophisticated and specific immune response (4).

T cells, which belong to the adaptive immune system, recognize specific antigens. During their development, T cells undergo random re-arrangement of their T cell receptor (TCR) variable segments. Although each T cell has a single specificity, this results in a pool of T cells that can recognize virtually any antigen. After TCR re-arrangement, T cells undergo thymic selection, where self-antigens are presented to the T cells, and self-reactive T cells are eliminated. A similar process produces B cells with specific B cell receptors (BCR) in the bone marrow. After this selection process, the mature lymphocytes (T- and B cells) migrate to the lymphoid tissue, ready to be activated. Upon activation of a mature T cell by a DC, it clonally expands to produce many T cells with exactly the same TCR and thus antigen specificity. Depending on

the type of T cell, it can then set out to directly destroy the pathogen (cluster of differentiation (CD)8⁺ T cells) or orchestrate the immune response by activating other cells like B cells, which in turn produce antibodies to neutralize the pathogen, and by secreting specific cytokines which determine the specific type of immune response that will be mounted (CD4⁺ T cells). A special subset of CD4⁺ T cells are regulatory T (Treg) cells, which, amongst other functions, are important for preventing excessive immune responses (5).

1.2 Regulatory T cells are the guardians of peripheral tolerance

Regulatory T (Treg) cells are an immune-regulatory subset of CD4⁺ T cells that is essential for maintaining peripheral tolerance to self-antigens and for restraining excessive inflammation (6). Treg cells are characterized by the expression of the transcription factor forkhead box protein 3 (FOXP3), which is essential for orchestrating the immunosuppressive program in Treg cells, as well as a specific methylation pattern. This methylation pattern is required for the stable expression of an array of genes required for Treg cell function, including, but not limited to, FOXP3, and thus ensures a stable phenotype (7).

FOXP3 is specifically expressed in Treg cells and is essential for their development (8, 9). The importance of Treg cells is demonstrated by loss-of-function mutations in the *FOXP3* gene, which cause uncontrolled inflammation and multi-system autoimmunity, leading to the scurfy phenotype in mice (8, 10) and the immunodysregulation polyendocrinopathy enteropathy X-linked (IPEX) syndrome in humans (11-13). FOXP3 expression alone, however, is not sufficient for establishing the Treg cell phenotype. Treg cell specific gene expression, full suppressive capacity and lineage stability additionally require a specific methylation pattern, which is established independently of FOXP3 expression. This methylation pattern is induced in the thymus and continued in the periphery, and is dependent on TCR signaling (14).

Treg cells can be subdivided into thymus-derived Treg (tTreg) cells, which develop in the thymus, and peripherally induced Treg (pTreg) cells, which are induced from conventional T (Tconv) cells in the periphery upon antigen exposure. *In vitro*, Treg cells are often induced from Tconv cells using TGF- β . However, although these cells

(termed iTreg cells) express FOXP3 and FOXP3-associated genes, they lack the Treg cell specific methylation pattern and thus are both less suppressive and less stable (14).

There are several ways in which Treg cells can suppress an immune response. These include the inhibition of Tconv cells via the secretion of anti-inflammatory cytokines interleukin (IL)-10, IL-35 and TGF- β ; the induction of apoptosis by Perforin and Granzyme-A and -B secretion or cell-cell interaction through TNF-related apoptosis-inducing ligand (TRAIL), Fas ligand (Fas-L), or Galectin-9/ transmembrane immunoglobulin and mucin domain (TIM)-3; and the interference with Tconv cell metabolism via cyclic adenosine monophosphate (cAMP) production or CD39-mediated conversion of adenosine triphosphate (ATP) to adenosine monophosphate (AMP). Treg cells are also able to suppress Tconv cells indirectly by interacting with antigen presenting cells (APC) via cytotoxic T-lymphocyte associated protein (CTLA)-4 (reviewed in (15)).

1.3 Description of a Treg cell lineage that promotes tissue regeneration

In addition to tTreg cells that can be found in lymphoid organs and mainly function in keeping inflammatory responses in check, Treg cells have been also been found in non-lymphoid tissues. The latter, which are also thymus-derived, express the IL-33 receptor suppression of tumorigenicity 2 (ST2) and have been described to exert tissue-regenerative functions.

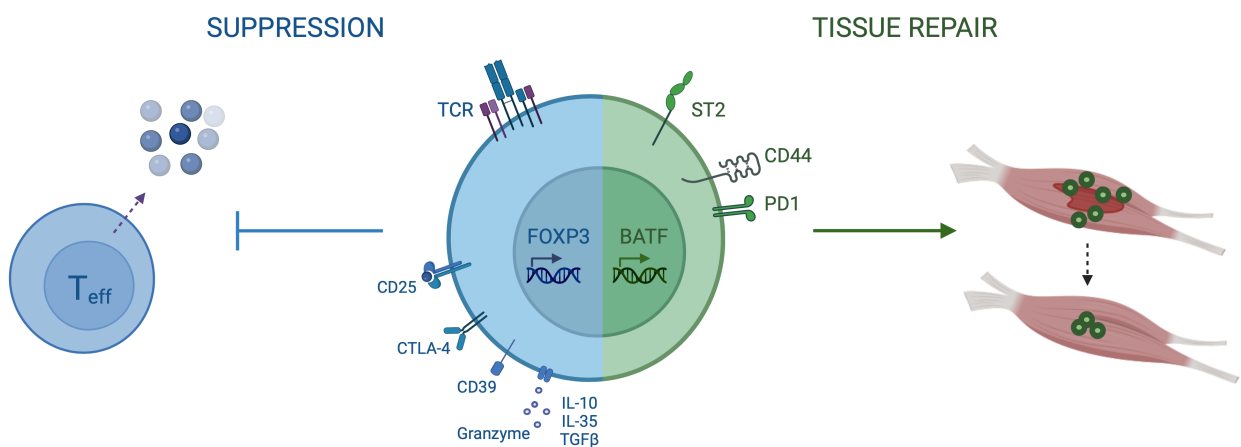


Figure 1: CCR8⁺ Treg cell. Left side: “classical” suppressive function: Treg cell suppresses effector T cells via different mechanisms including depletion of the environment of IL-2 via the IL-2 receptor CD25,

INTRODUCTION

the secretion of anti-inflammatory cytokines IL-10, IL-35 and TGF- β , the induction of apoptosis by granzyme secretion, interference with effector T cell metabolism via CD39, or indirectly via CTLA-4. Right side: Tissue regenerative program mediated by the transcription factor BATF: Treg cell in non-lymphoid tissue promotes tissue regeneration via Areg in mouse and an unknown factor in human. Figure elements created with BioRender.

In the murine system, tagmentation-based whole-genome bisulfite sequencing (TWGBS) of tissue-resident Treg cells from visceral adipose tissue (VAT) and skin, and lymphoid tissue-derived Treg and Tconv cells, revealed a great similarity of VAT and skin Treg cell methylation patterns, which were distinct from the ones of lymphoid Treg and Tconv cells. In addition to the Treg cell-specific epigenetic signature, including hypomethylation of the *Ctla4*, *IKAROS family zinc finger (Ikzf)2*, *Ikzf4* and *interleukin 2 receptor subunit alpha (Il2ra)* loci (14,16), a tissue Treg cell signature shared between VAT and skin tissue-resident Treg cells was identified. This signature comprised hypomethylation of genes encoding for the T helper type 2 (Th2) cell-associated transcription factors *GATA Binding Protein (Gata)3*, *musculoaponeurotic fibrosarcoma (Maf)* and *interferon regulatory factor (Irf)4*, and further *basic leucine zipper transcription factor, ATF-like (Batf)*; *Killer cell lectin-like receptor G1 (Klrg1)*, and *interleukin 1 receptor-like 1 (Il1rl1)*, encoding for ST2. Further characterization of this Treg cell subtype classified them as CD44^{hi}, suggesting previous activation, and revealed elevated expression of IL-10 and amphiregulin (Areg) (16, 17), an epidermal growth factor receptor (EGFR) ligand known for its tissue-regenerative function (18, 19). Although the methylomes of VAT and skin-resident Treg cells showed large overlaps, several tissue-specific characteristics were also found. One example is the hypomethylation of *Ahr*, which encodes for the transcription factor aryl hydrocarbon receptor (AhR), specifically in skin-resident Treg cells (16). Beside the VAT and skin, where tissue-resident Treg cells constitute around 80-90% and 50-60% of all Treg cells, respectively, Treg cells with this common signature were found in other tissues including the colon (30%), lungs, liver, and bone marrow (10-20%), and a small fraction in the spleen (<5%) (16, 20). Together, this data identifies a distinct subset of Treg cells in all tissues investigated in these studies, with a common tissue-resident Treg cell specific methylation pattern on top of the Treg cell specific epigenetic imprint, and some tissue-specific differences. This suggests a common developmental path for tissue-resident Treg cells.

Indeed, pseudotime analysis of single cell (sc)RNA-seq data from Treg cells present in VAT, skin, lung, liver, bone marrow, blood, lymph nodes and spleen, and from putative tissue Treg cell precursor populations, identified two common precursor stages of tissue-resident Treg cells: Klrp1⁻Nfil3⁺ and Klrp1⁺Nfil3⁺ Treg cells residing in lymphoid tissues. Sequential differentiation of Foxp3⁺Klrp1⁻Nfil3⁻ via Foxp3⁺Klrp1⁻Nfil3⁺ and Foxp3⁺Klrp1⁺Nfil3⁺ to tissue Treg cells was further supported by RNA velocity analysis of scRNA-seq data and adoptive transfer experiments (17).

In search of factors driving this tissue Treg cell program, the transcription factor Batf was found to be specifically enriched in tissue-resident Treg cells and their lymphoid precursors. In a scRNA-seq dataset of tissue Treg cells, more than 60 Batf-correlated or -anti-correlated genes were identified, including *TNF superfamily member 4 (Tnfsf4)*, *Klrp1*, *programmed cell death 1 (Pdc1)* (correlated) and *B-cell lymphoma 2 (Bcl2)*, *selectin L (Sell)*, *Ccr7*, *SATB homeobox 1 (Satb1)* and *inhibitor of DNA binding 3 (Id3)* (anti-correlated). The role of Batf as key transcription factor of tissue-resident Treg cells is further supported by the fact that Batf-deficient animals have a severely reduced frequency of tissue-resident Treg cells, and Klrp1⁻ Treg cells from Batf^{-/-} animals were unable to give rise to Klrp1⁺ tissue Treg precursor cells in adoptive transfer experiments (17).

Therefore, beside their function in immune regulation, which is dependent on Foxp3, tissue resident Treg cells are additionally involved in tissue homeostasis and regeneration, dependent on the transcription factor Batf.

Treg cells with a comparable tissue regeneration program have recently been identified in human tissues, based on single cell chromatin accessibility data of murine and human Treg cells from blood, skin and fat tissue. This human tissue-resident Treg cell subset can be identified by the expression of the transcription factor BATF and the chemokine receptor CCR8, and a BATF-associated tissue-repair Treg cell chromatin signature (20). Comparable to murine Ccr8⁺ Treg cells, human CCR8⁺ Treg cells are at the final step of differentiation from naïve via memory-type Treg cells. While murine Ccr8⁺ Treg cells express several genes associated with Th2 cells (16, 17, 21), human CCR8⁺ Treg cells have a follicular helper T (Tfh)-like signature, and a tissue-repair Treg cell-like phenotype can be induced from naïve Treg cells *in vitro* using typical

Tfh-inducing cytokines. Both in a wound healing assay and a reconstructed human epidermis model, supernatant from *in vitro* induced tissue-resident Treg cells promoted wound closure, highlighting the tissue-regenerative potential of this cell type (20).

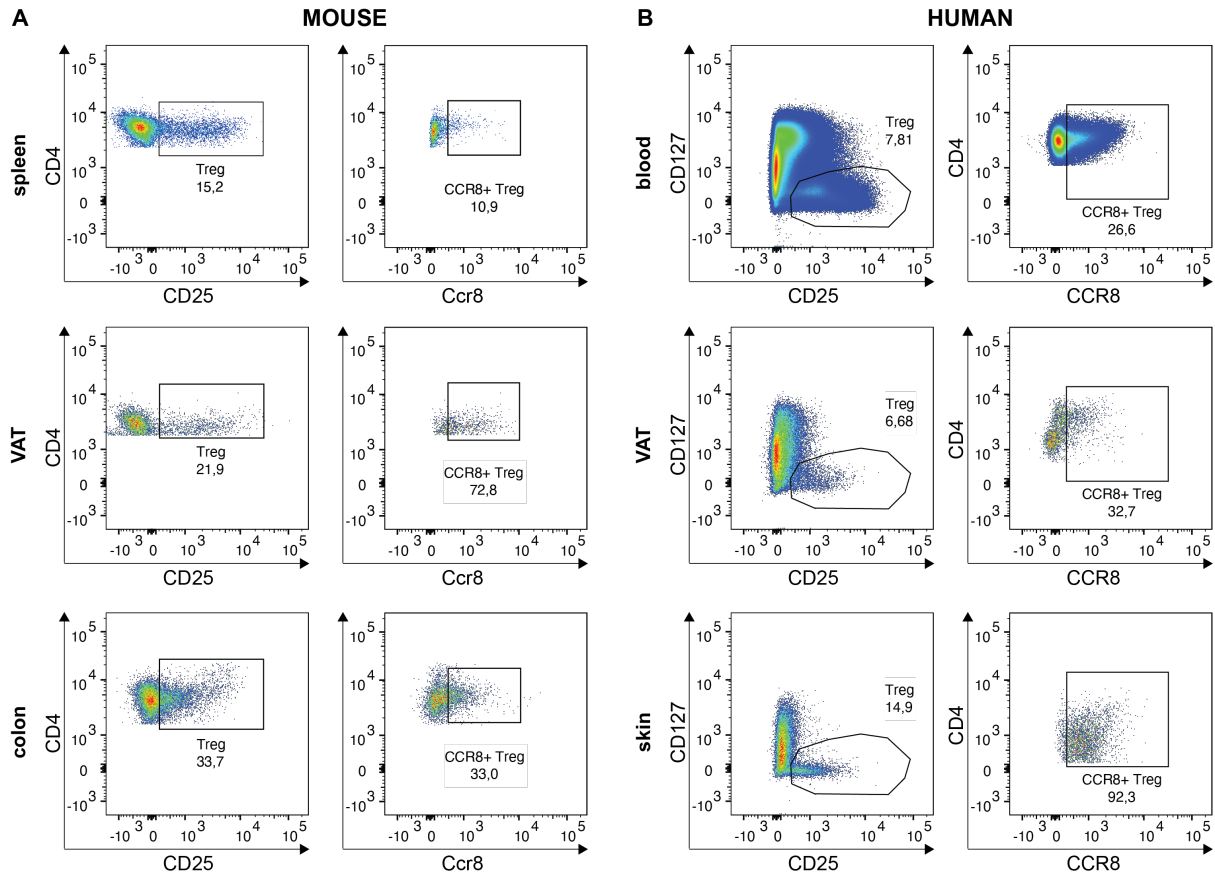


Figure 2: CCR8⁺ Treg cells in murine and human tissues. **A.** A fraction of murine CD25⁺ Treg cells expresses Ccr8: Treg cells are shown as CD4⁺CD25⁺ (left) and Ccr8⁺ Treg cells are shown as fraction of total Treg cells (right). **B.** Also in the human system, a subset of Treg cells expresses CCR8: Treg cells are shown as CD127⁺CD25⁺ (left) and CCR8⁺ Treg cells are shown as a fraction of total Treg cells (right).

1.4 CCR8⁺ Treg cells can promote tissue regeneration

In addition to their Foxp3-dependent classical function, CCR8⁺ Treg cells have tissue regenerative potential, for which they depend on the transcription factor BATF. Regarding this regenerative function, they have been reported to be involved in organ homeostasis, tissue regeneration upon different types of damage, like inflammation and infection, and in cancer development.

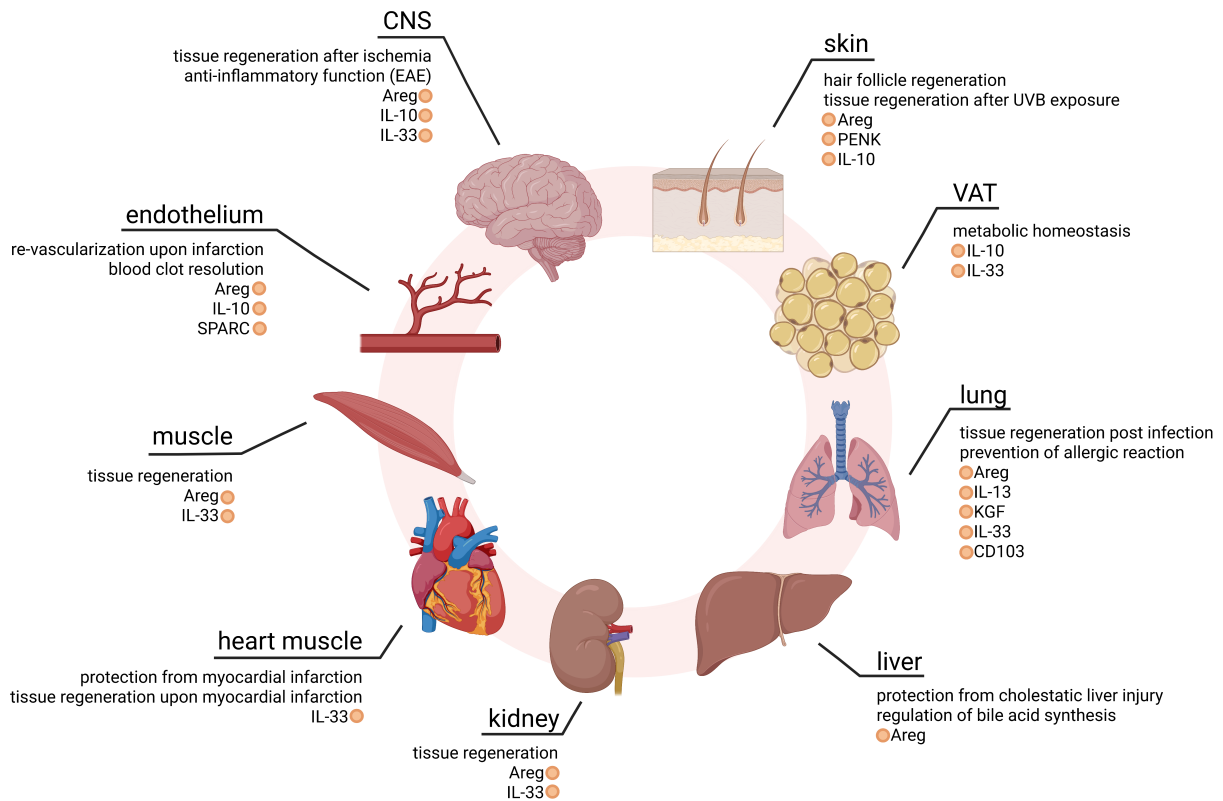


Figure 3: Treg cell-mediated tissue regeneration. Tissues for which a role of Treg cells in homeostasis or tissue damage control has been reported in mouse and/or human. Factors promoting these non-classical Treg functions are marked in orange. Printed from Braband *et al.* (1).

1.4.1 Tissue Treg cells promote tissue homeostasis under steady state in the VAT, skin, and liver

In the context of homeostasis, Treg cells found in non-lymphoid tissues are best known for their function in metabolic control. In mice, Treg cells residing in the VAT have a phenotype distinct from lymphoid organ Treg cells, with enhanced expression of IL-10 and early activation markers, amongst others. This special Treg cell subset is strikingly reduced in the VAT of obese mice, and insulin resistance is directly correlated with VAT-resident Treg cell number. Treg cells in the VAT are located in close proximity to other immune cells like macrophages, and may control an excessive inflammatory immune response and the downstream consequences like insulin resistance via the IL-33 – ST2 axis. VAT Treg cells expressing ST2 have also been found in the human system (22-24).

Murine tissue Treg cells moreover are known to regulate hair follicle stem cell (HFSC) differentiation. In the skin, Treg cells are mainly localized in close proximity of HFSCs, and their activation status correlates with the hair follicle cycle. Depletion of Treg cells using a *Foxp3^{DTR}* mouse line markedly reduced HFSC proliferation and the telogen-to-anagen transition, demonstrating their importance in hair follicle cycling. Mechanistically, transcriptome analysis revealed enhanced expression of the Notch-ligand family member *Jagged1* (*Jag1*) in activated skin-resident Treg cells, as compared to Treg cells from skin-draining lymph nodes, which is crucial for the induction of HFSC proliferation via the *Jag1*-Notch pathway (25). Alopecia areata, which is a form of autoimmune hair loss in humans, has further been associated with impaired Treg function in the skin (26, 27).

The accumulation of bile acid in the liver is known to cause hepatocellular injury. Treg cells contribute to homeostasis in the liver by regulating bile acid synthesis in an Areg-dependent manner, and thus protect from cholestatic liver injury. *Areg^{-/-}* mice display more severe liver injury in a bile duct ligation and alpha-naphthylisothiocyanate (ANIT) gavage liver injury model, and recombinant Areg protected from liver injury (28).

1.4.2 Tissue Treg cells repair tissue damage

Beside homeostatic functions, murine regulatory T cells in the skin are further involved in skin regeneration upon exposure to ultraviolet (UV)B radiation. Upon exposure, a specific subset of Treg cells expands, with a distinct TCR repertoire and gene expression pattern. This Treg cell subset exhibits high expression levels of proenkephalin (PENK), which can induce keratinocyte proliferation, and the tissue-regenerative factor Areg (29). Treg cells also accumulate in the skin upon injury. These Treg cells help wound healing, as was shown by a markedly attenuated wound closure in Treg cell-ablated mice. In this setting, Treg cells reduce the accumulation of pro-inflammatory (M1) macrophages by inhibiting interferon (IFN) γ production of Th1 cells (30). Treg cells have further been shown to regulate Th17 cells in the skin, inhibiting their expression of IL-17A and subsequent C-X-C motif chemokine ligand (CXCL)5-dependent neutrophil attraction to the site of injury. Treg cell-mediated regulation of

excessive inflammation, in turn, allows for the differentiation of HFSCs into epithelial cells (31).

Similarly, a specific population of Treg cells accumulates in skeletal muscle upon injury, as has been shown both in models of acute muscle injury and a muscular dystrophy model of chronic injury (18, 32, 33). Treg cells found in the skeletal muscle after injury display a typical “muscle-Treg” phenotype, including enhanced expression of *Klrg1*, *Ccr8*, *Il-10*, *Areg*, and *ST2* (18). Upon injury, Treg cells seem to be recruited from the circulating T cell pool, and expand in the muscle in an IL-33 dependent manner (32). Injury-induced expansion of this cell type decreases with age, dependent on IL-33 expression by fibro/ adipogenic progenitors (FAP) in the tissue, and is essential for effective repair of the skeletal muscle (32). The tissue-regenerative effect of tissue resident Treg cells in this setting is largely mediated by their expression of *Areg*. *Areg* can activate satellite cells, which have a very low turnover in healthy muscle, and thus mediate the formation of new myofibers (18, 33). Genes induced in the muscle of Treg-ablated mice in a model of chronic muscle injury have been found to correspond well with genes repressed in the muscle of *Areg*-treated mice. Further, gene expression in Treg-ablated mice was largely normalized upon treatment with *Areg* (18). Mice lacking *ST2* expression specifically on Treg cells showed impaired expansion of Treg cells upon injury, and were less successful in clearing the muscle infiltrate, demonstrating the importance of this tissue-specific subset of Treg cells in the regeneration of muscle tissue (32). Tissue resident Treg cells further protect from, and promote tissue regeneration after myocardial infarction, depending also on the tissue regenerative factor secreted acidic cysteine-rich glycoprotein (*Sparc*) (34-36).

In the lungs, as well, Treg cells have been shown to promote tissue repair in an IL-33 – dependent manner by enhancing type II alveolar epithelial cell (AEC2) proliferation (37). A tissue repair factor identified here beside *Areg* is keratinocyte growth factor (*Kgf*), which is important for AEC2 proliferation after acute lung injury. *KGF* has further been found to be expressed by human lung Treg cells (38). Treg cells seem to rely on the integrin CD103 for this effector function (39). IL-13, secreted by Treg cells upon IL-33 stimulation, further seems to have a regenerative effect on lung tissue (40). Treg cells in the lungs further play a role in reducing fibrosis by impairing fibrocyte recruitment, as demonstrated in an acute lung injury mouse model (41).

Tissue Treg cells found in the kidney were reported to ameliorate renal injury via the IL-33 – ST2 axis, in an Areg-dependent manner (42).

Finally, in the CNS, Treg cells are important for limiting brain damage after ischemic stroke. Treg cells are essential for limiting brain damage after experimental brain ischemia, for which they rely on IL-10, which dampens the activation of inflammatory cells, and on Areg, for the suppression of neurotoxic astrogliosis. Depletion of Treg cells in this model drastically increased the severity of the brain damage (43, 44). The activation of Treg cells in the brain occurs via the alarmin IL-33 (45). In a mouse model for multiple sclerosis, the experimental autoimmune encephalomyelitis (EAE) model, Treg cells proved crucial for promoting oligodendrocyte differentiation and remyelination (46), dependent on ST2 signaling (47).

1.4.3 Treg cells are present in the tumor microenvironment

Tumors arise from cells, in which defects in cellular pathways regulating proliferation and homeostasis have accumulated. Hanahan and Weinberg have defined eight hallmarks of cancer, which allow for the development of tumors, and which are common to most cancer types. These include self-sufficiency in growth signals, insensitivity to anti-growth signals, evasion of apoptosis, the potential for limitless proliferation, sustained angiogenesis, the ability to invade tissues and to form metastases, metabolic deregulation, and avoiding immune destruction (48, 49). Genome instability is an important enabling characteristic for the acquisition of abovementioned hallmarks through genomic mutations, as is an inflammatory cell state (**Figure 4**). Additionally, epigenetic reprogramming may also lead to deregulation of said pathways (50).

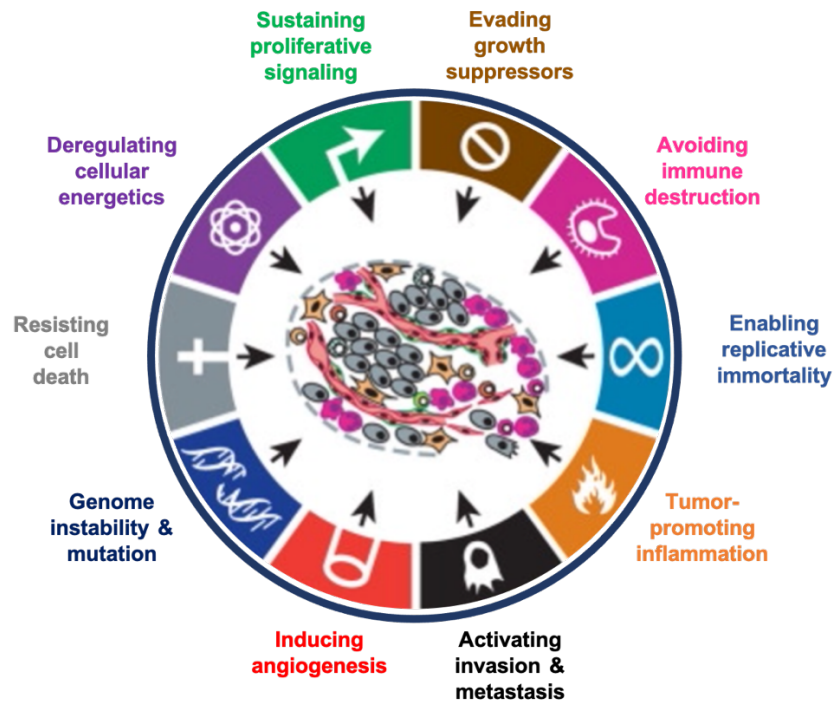


Figure 4: The hallmarks of cancer. The eight hallmarks of cancer as defined by Hanahan and Weinberg comprise sustaining proliferative signaling, evading growth suppression, enabling replicative immortality, resisting cell death, inducing angiogenesis, activating invasion and metastasis, deregulating cellular energetics and avoiding immune destruction. These hallmarks are enabled by genome instability and mutations, and by tumor-promoting inflammation. Adapted from (49).

Avoiding immune destruction is one hallmark of cancer. The immune system monitors the body not only for pathogens, but also for deregulated or transformed cells. Evading recognition by the immune system thus is a strategy crucial for the development of tumors. Immune evasion strategies include the downregulation of antigen presentation, the limitation of nutrient availability to immune cells, the secretion of immunosuppressive factors, and the recruitment of myeloid-derived suppressor cells (MDSC) and Treg cells to the tumor microenvironment, amongst other mechanisms. Treg cells employed by the tumor suppress the effector function of tumor-infiltrating lymphocytes (TIL), actively preventing the killing of tumor cells (49, 51).

1.5 Aim of the project

Treg cells have many positive effects on tissue homeostasis and repair, yet they might be detrimental in the tumor setting due to their immunosuppressive and potentially also their tissue-remodeling properties. Treg cells in the tumor, often referred to as “tumor Treg” cells, have been attributed additional functions beside

immunosuppression, based on their unique transcriptional signature (52). One study suggests that these “tumor Treg” cells express CCR8 and the transcription factor BATF, which orchestrates the effector program of this cell type (53). Many studies, however, compare Treg cells isolated from tumor tissue to Treg cells isolated from the circulation (53-56). This makes it impossible to distinguish differences between tumor- and healthy tissue from differences arising from lymphoid vs non-lymphoid tissue. As described earlier, CCR8⁺ Treg cells have important functions in tissue homeostasis and damage repair. For the development of “tumor Treg” cell-targeting cancer therapies it is therefore of utmost importance to deconvolute discrepancies in transcriptional programs and thus targetable differences between Treg cells in tumor vs healthy tissue from discrepancies in transcriptional programs between lymphoid vs non-lymphoid Treg cells. Therefore, in this thesis, we ask the following questions:

- 1) Can we find human CCR8⁺ Treg cells in the tumor microenvironment as well as in tumor-free human tissues?
- 2) Is the tissue regeneration program active in tumor-resident CCR8⁺ Treg cells?
- 3) Do tumor-resident CCR8⁺ Treg cells share their epigenetic and transcriptional programs with CCR8⁺ Treg cells in normal tissue?
- 4) Are there differences in the tissue regeneration program between primary tumors and metastasis?

For a detailed characterization of the transcriptional programs governing this cell type, we assayed the single-cell resolution chromatin accessibility profiles with scATAC-seq. We further performed scRNA/TCR-seq to analyze gene expression, complementing the information from scATAC-seq. Analyzing the TCR repertoire of CCR8⁺ Treg cells in tumor and healthy tissue further allowed us to speculate on CCR8⁺ Treg cells expanding with the growing tumor versus CCR8⁺ Treg cells being actively recruited by the tumor at a later stage.

As discussed above, CCR8⁺ Treg cells isolated from tumor tissue have been compared to lymphoid tissue Treg cells in various studies. This did not make it possible to distinguish tumor-vs-healthy tissue from lymphoid-vs-nonlymphoid tissue differences. Comparing the tumor tissue with tumor-free tissue of the same patient in this study allowed for the deconvolution of these differences.

1.6 Methods

1.6.1 Investigating chromatin accessibility on single-cell level using scATAC-seq

Desoxyribonucleic acid (DNA), or chromatin, is organized around histones. Dependent on whether or not genes encoded on a certain part of the DNA are actively translated, this part of the DNA is loosely organized and thus accessible for transcription machinery, or tightly wrapped around a histone and thus inaccessible. How the chromatin is organized around histones is dependent on certain modifications on the histones and the DNA itself, including histone methylation, acetylation and phosphorylation, and DNA methylation (57, 58).

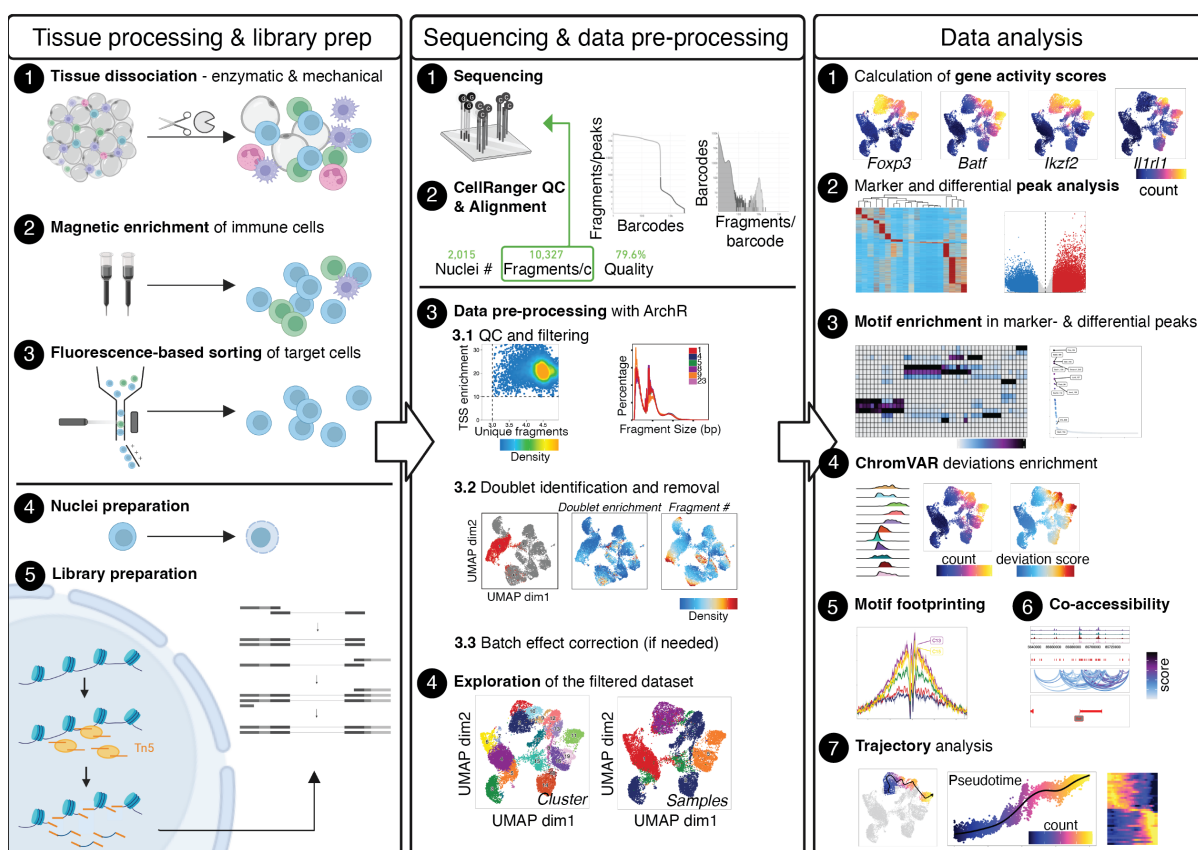


Figure 5: Overview scATAC-seq. This figure describes the process from isolating cells of interest from fresh tissue, via library preparation and sequencing, to scATAC-seq data pre-processing and analysis. Left pane: Fresh tissue is mechanically and enzymatically dissociated (1) followed by pre-enrichment of the immune cell type of interest (2) and fluorescence activated cell sorting (FACS) (3). For the generation of a scATAC-seq library, cells are subjected to nuclei preparation (4) and transposition, which yields fragments for the preparation of a sequencing library. Cells containing fragments are then barcoded for the single-cell application, followed by library preparation (5). Middle pane: scATAC-seq library is sequenced (1) and fragments are aligned using Cell Ranger ATAC (2). Prior to analysis, quality

metrics are applied to the data, and only informative, high-quality cells are kept in the dataset (3, 4). Right pane: Analysis of scATAC-seq data can include cell type annotation based on gene activity scores or pre-defined signatures, (1) peak calling and differential analysis (2), motif enrichment analysis within differential peaks (3, 4), motif footprinting (5), co-accessibility analysis (6) and pseudotime analysis (7). Printed from Braband *et al.* (2).

In the Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq), a library for sequencing is generated only from the accessible regions of the DNA, using a hyperactive Tn5 transposase. The Tn5 transposase cannot access closed chromatin due to sterical hinderance, but binds to open chromatin, inserting adapters for library generation (59, 60) (**Figure 5**, left pane, step 5). Identifying accessible chromatin by sequencing allows us to infer transcriptional programs that are active in the cells. ATAC-seq is further possible on the single-cell level (scATAC-seq), which is especially useful for characterizing heterogeneous populations in more detail (61-63).

1.6.2 Investigating gene expression on single-cell level combined with T cell receptor sequences using scRNA/TCR-seq

Messenger ribonucleic acid (mRNA) is transcribed from accessible chromatin regions as a kind of “construction plan” for generating proteins. Sequencing mRNA gives us a snapshot of which genes are being translated into proteins at the given timepoint, which is a valuable complement to the information obtained from scATAC-seq data. RNA-seq is also available as a single-cell technique (64). TCR sequences can further be analyzed using combined scRNA/TCR-seq, where variable regions of the TCR are being amplified and subjected to library preparation and sequencing. The TCR can be used as a “biological barcode” of the cell. Since each T cell that leaves the thymus has an individual TCR, and T cells undergoing clonal expansion in the periphery generate daughter cells with a TCR identical to their own, information from TCR-seq lets us infer the clonal relationships of T cells.

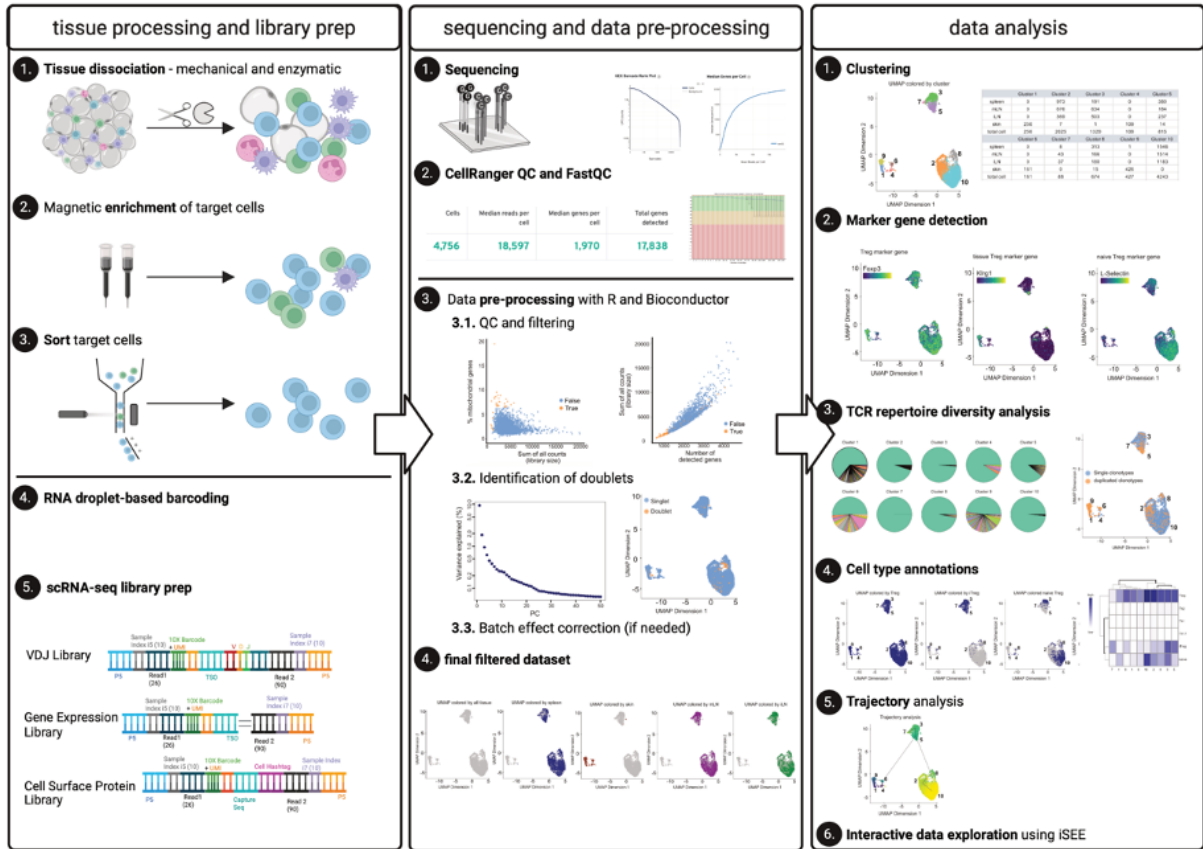


Figure 6: Overview scRNA/TCR-seq. This figure describes the workflow from isolating single cells from tissues, via library preparation and sequencing, to data pre-processing and analysis. Left pane: Fresh tissues are mechanically and enzymatically dissociated to yield a single cell suspension (1). The cell type of interest is then isolated from this suspension via magnetic enrichment (2) followed by FACS (3). Single cells are barcoded (4) and gene expression, V(D)J, and optionally cell surface protein libraries are prepared (5). Middle Pane: After sequencing (1), fragments are aligned using Cell Ranger (2). Data pre-processing is performed in R using various packages from CRAN and Bioconductor, and includes quality filtering and doublet removal (3, 4). Right pane: the filtered dataset is then subjected to data analysis, which can include clustering (1), the identification of marker genes for certain clusters (2), TCR repertoire analysis (3), manual cell type annotation (4) and pseudotime analysis (5). Data can further be interactively explored using the R package iSEE (6). Printed from Nedwed *et al.* (3).

1.7 Description of tumor types studied in this thesis

1.7.1 Clear cell renal carcinoma

Clear cell renal carcinoma (CCRC) is the most prevalent renal malignancy, making up about 80% of all renal carcinomas (65). CCRC is characterized by a loss of chromosome 3p, both in sporadic and hereditary disease, as well as mutation or loss of the *von Hippel-Lindau (VHL)* gene. In hereditary disease, this is in the form of germline loss, whereas in sporadic disease, *VHL* may be mutated or somatically lost.

Chromosome 3p loss as well as *VHL* loss or mutation are initiators of CCRC and are ubiquitously present. Secondary mutations include mutations of *SET2*, *PBRM1* and *BAP1*, and copy number changes in chromosomes 9p and 14q (66). Inactivation of the *VHL* gene leads to upregulation of hypoxia-inducible factor (HIF), which in turn promotes angiogenesis. Besides therapies targeting angiogenesis, CCRC also respond to immunotherapy, including high-dose IL-2, type I interferons, and the immune checkpoint inhibitor anti-PD1 (67).

1.7.2 Renal angiomyolipoma

Renal angiomyolipoma (AML) is a benign tumor of the kidney, which consists of myocytes, adipocytes and blood vessels. Most AML occur sporadically, while about 30% of the patients have underlying mutations leading to the activation of the mechanistic target of rapamycin (mTOR) pathway, which controls cell growth and metabolism. The main clinical concern with AML is life-threatening hemorrhage, and AML may be treated by resection, arterial embolization or mTOR inhibitors (68).

1.7.3 Hepatocellular carcinoma

Hepatocellular carcinoma (HCC) is the most common malignancy of the liver. Risk factors for HCC include hepatitis B virus (HBV) or hepatitis C virus (HCV) infection, excessive alcohol consumption leading to cirrhosis, and obesity-related non-alcoholic steatohepatitis (NASH). The inflammatory environment enables mutations, leading to the development of HCC. Driver mutations in HCC often affect the *telomerase reverse transcriptase (TERT)* gene and induce telomerase activation, but also include mutations which result in the activation of Wnt-signaling or the alteration of cell cycle control. Depending on the tumor stage, HCC may be treated by surgical resection or systemic therapies such as tyrosine kinase inhibitors and immune checkpoint blockade (reviewed in (69)).

1.7.4 Colorectal carcinoma

Colorectal carcinoma (CRC) is the third most common malignancy world-wide. CRC can develop sporadically, with environmental risk factors being diet and lifestyle, or it can be hereditary or colitis-associated. CRC develop from epithelial cells in the colon or rectum and generally have a high mutational burden (70). Depending on the tumor

stage, standard therapies are surgical resection, chemo-, radio-, or immunotherapy, which CRC of later stages often develop resistance to (71).

1.7.5 Lung carcinoma

Lung cancer is the malignancy with the highest prevalence world-wide, and non-small cell lung cancer (NSCLC) is the most frequent form of lung cancer. NSCLC comprise adenocarcinoma (AC), squamous cell carcinoma (SCC), and adenosquamous carcinoma (ASC), dependent on the cell type of origin. The most well-known risk factor for NSCLC is smoking, however also air pollution and other carcinogens are important environmental risk factors. Genetic risk factors include mutations in genes involved in DNA repair, cell cycle control, apoptotic pathways, inflammation and telomere length. NSCLC belongs to the malignancies with the highest mutational burden (72). Treatment options include resection, chemo-, radio-, and immunotherapy, as well as targeted drugs for specific mutations (73).

1.7.6 Gastric adenocarcinoma

Gastric adenocarcinoma (GAC) is the most common gastric malignancy, and is genotypically highly heterogeneous. A common risk factor for GAC is infection with *helicobacter pylori* (74). GAC is very treatable at early stages, where resection is sufficient. Treatments for advanced GAC include chemotherapy, monoclonal antibodies preventing angiogenesis, or immune checkpoint blockade, however median survival for advanced GAC is below one year (75).

1.7.7 Oral cavity squamous cell carcinoma (OCSCC)

Squamous cell carcinoma of the oral cavity (OCSCC) arises from the mucosal epithelium of the oral cavity and is a common type of head and neck cancer. Risk factors for OCSCC include environmental risk factors such as smoking and alcohol consumption, human papilloma virus (HPV) infection, and genetic mutations. OCSCC is usually treated by surgical resection combined with chemo-, radio- or immunotherapy. Depending on the underlying genetic alterations, targeted therapies are possible (76, 77).

1.7.8 Urothelial carcinoma

Urothelial carcinoma (UC) is a malignancy with a relatively high mutational burden, especially in muscle invasive disease. Common mutations are in the *TERT* gene, which is mutated in 70-80% of all cases, and also in *tumor protein p53 (TP53)*, involved in cell cycle control, *phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA)* and *fibroblast growth factor receptor 3 (FGFR3)*, involved in proliferative signaling, and in genes involved in chromatin modification. Non-muscle invasive UC can usually be treated by surgical resection. Treatment options for muscle invasive UC are resection combined with chemotherapy, and immunotherapy such as immune checkpoint blockade (78, 79).

1.7.9 Seminoma

Seminoma is a malignant testicular germ cell tumor (TGCT) and occurs predominantly in younger men between 15 and 35 years of age. Defective maturation of gonocytes in the fetus is thought to be the underlying cause for germ cell neoplasia, which then develops into TGCT after puberty. TGCT have a comparably low mutational burden. Seminoma is treated by resection and cisplatin-based chemotherapy, with an excellent response rate (80).

1.7.10 Liver metastases

Metastases arise from single tumor cells which have entered the circulation, seeded to another tissue, and successfully proliferate in the target tissue. The liver, as a highly perfused organ, is a common site for metastases. This includes metastases from CRC, pancreatic cancer, gastric cancer, breast cancer, cutaneous melanoma and lung cancer. Gastrointestinal (GI) tumors metastasize to the liver especially frequently owing to the blood flow pattern from the GI tract to the liver via the portal vein (81). When the tumor cell arrives at the target organ, the local microenvironment supports the growth of the metastasis through interaction of the resident cells and recruited immune cells with the tumor cells, a concept known as the “seed-and-soil” hypothesis (82). Treatment of metastatic cancer depends on the tissue of origin and is generally more challenging than early-stage localized cancer, requiring a systemic approach.

2. MATERIALS

2.1 Equipment

Device	Catalog number	Company
Bioanalyzer 2100	-	Agilent
Chromium Controller	-	10X Genomics
FACS Aria II	-	BD
GentleMACS	130-096-427	Miltenyi
NextSeq 500/550	-	Illumina

2.2 Consumables

Antibodies	Catalog number	Company
CCR8-BV421	566379	BD Biosciences
CD127-BV711	351328	Biologend
CD19-APC	302212	Biologend
CD206-APC	321110	Biologend
CD25-PE	302606	Biologend
CD3-BV785	317330	Biologend
CD39-BV605	328236	Biologend
CD4-R718	566352	BD Biosciences
CD45-BUV737	748719	BD Biosciences
CD45RA-BV510	304142	Biologend
CD45RO-PE-Cy7	304230	Biologend
CD8-BB700	566452	BD Biosciences
PD1-Viobright515	130-120-386	Miltenyi
TotalSeq™-C0251 anti-human Hastag 1 Antibody	394661	Biologend
TotalSeq™-C0252 anti-human Hastag 2 Antibody	394663	Biologend
TotalSeq™-C0253 anti-human Hastag 3 Antibody	394665	Biologend

Chemicals	Catalog number	Company
Cell Stimulation Cocktail	00-4975-03	eBioscience
Digitonin 5%	BN2006	Thermo Fisher
MgCl ₂ , 1M	M1028	Sigma

MATERIALS

NaCl, 5M	59222C	Sigma
NP-40	98379-10ML-F	Sigma
TRIS-HCL pH7.4	T2194	Sigma
Tween-20	1662404	Biorad

Enzymes	Catalog number	Company
Collagenase II	C6885-1G	Sigma
Collagenase IV	C5138-1G	Sigma
DNase I	11284932001	Roche

Kits	Catalog number	Company
5' Feature Barcode Kit	PN-1000256	10X Genomics
Chromium Next GEM Chip H Single Cell Kit	PN-1000161	10X Genomics
Chromium Next GEM Chip K Single Cell Kit	PN-1000286	10X Genomics
Chromium Next GEM Single Cell 5' Reagent Kit v2	PN-1000263	10X Genomics
Chromium Next GEM Single Cell ATAC Library & Gel Bead Kit	PN-1000175	10X Genomics
Chromium Single Cell Human TCR Amplification Kit	PN-1000252	10X Genomics
Dual Index Kit TN Set A	PN-1000250	10X Genomics
Dual Index Kit TT Set A	PN-1000215	10X Genomics
High Sensitivity DNA Kit	5067-4626	Agilent
Library Construction Kit	PN-1000190	10X Genomics
NextSeq 500/550 High Output Kit v2.5 (75 Cycles)	20024906	Illumina
NextSeq 500/550 High Output Kit v2.5 (150 Cycles)	20024907	Illumina
Single Index Kit N, Set A	PN-1000212	10X Genomics

Media and buffers	Catalog number	Company
BSA	P06-1402500	PAN-Biotech

MATERIALS

DMEM	10938-025	Gibco
FCS	26140079	Gibco
HEPES	P05-01100	PAN-Biotech
nuclei buffer	PN-2000207	10X Genomics
RPMI	11875093	Gibco

Other	Catalog number	Company
45 um filter net	2183468DCCC	Labomedic
70 um strainers	542070	Greiner
100 um strainers	542000	Greiner
Nuclease-free water	W4502-10X50ML	Sigma
TIL microbeads, human	130-118-780	Miltenyi
Zombie (NIR)	423106	Biologend

3. METHODS

3.1 Human tissue

Fresh human tissue was obtained from patients undergoing tumor resection at the University Medical Center Mainz. Tumor tissue and normal tissue adjacent to the tumor of patients who had not been treated with chemo-, radio-, or immunotherapy prior to resection was characterized by a pathologist and was processed separately. **Tables 2 and 3** show details on human tissue samples processed for scATAC-seq and scRNA/TCR-seq, respectively. Experiments were approved by the Landesärztekammer Rheinland-Pfalz, ethics vote 2021-15834.

Table 1: Tissue samples processed for scATAC-seq. TNM Grading: T, tumor, grade 1-4 relates to size and degree of tissue infiltration; N, lymph nodes, grade 0 = no lymph nodes affected, grade 1-3 relates to degree of tumor cells infiltrating lymph nodes; M, metastasis, grade 0 = no metastases found, grade 1 = metastases found; G: grade of differentiation, G1-4, degree of differentiation is anti-proportional, malignancy is proportional to the value of the number; L, lymph vessels affected (0/1); V, blood vessels affected (0/1); Pm, distant metastases.

Sample	Age	Gender	Diagnosis	T	N	M	G	L	V	Pm
AML	62	w	Renal angiomyolipoma							
CCRC1	64	m	Clear cell renal carcinoma	1a	X		2			
CCRC2	71	w	Clear cell renal carcinoma	3a	0		2	0	1	0
CRC + liver MET	48	w	Colorectal adenocarcinoma with metastasis to the liver	3	2b	1b	3	1	0	1
GAC	84	m	Gastric adenocarcinoma (intestinal type)	3	0		2	0	0	0
TGCT1	42	m	Seminoma							
TGCT2	27	M	Seminoma							
HCC1	62	m	Hepatocellular carcinoma	2	1		2	1	1	0
HCC2	84	m	Hepatocellular carcinoma	1b	X		2	0	0	0
HCC3	70	m	Hepatocellular carcinoma	3	X		3	0	0	0
MET1	45	w	Liver metastasis from colorectal adenocarcinoma							
MET2	49	m	Liver metastasis from colorectal adenocarcinoma							
MET3	57	w	Liver metastasis from colorectal adenocarcinoma							

METHODS

Lung AC	62	m	Lung adenocarcinoma	3	2	2	1	0	0
Lung ASC	61	m	Lung adenosquamous carcinoma	1c	0	3	0	0	0
OCSCC	62	m	Oral cavity squamous cell carcinoma	4a		2			
UC1	69	w	Urothelial carcinoma	3	0	2	0	0	0
UC2	54	m	Urothelial carcinoma	2b	0	2	0	0	0

Table 2: Tissue samples in the scRNA/TCR-seq dataset. TNM Grading: T, tumor, grade 1-4 relates to size and degree of tissue infiltration; N, lymph nodes, grade 0 = no lymph nodes affected, grade 1-3 relates to degree of tumor cells infiltrating lymph nodes; M, metastasis, grade 0 = no metastases found, grade 1 = metastases found; G: grade of differentiation, G1-4, degree of differentiation is anti-proportional, malignancy is proportional to the value of the number; L, lymph vessels affected (0/1); V, blood vessels affected (0/1); Pm, distant metastases.

Sample	Age	Gender	Diagnosis	T	N	M	G	L	V	Pm
CCRC 3	57	w	Clear cell renal carcinoma	1b	X		1	0	0	0

3.2 Isolation of immune cells from human tissues

Immune cells were isolated from human tumor tissue and NAT separately, in order to compare their immune cell type composition and -identity. To this end, tissue was placed in 5 mL digestion buffer (for digestion buffer composition for each tissue type see **Tables 4-9**) and was mechanically dissociated using scissors prior to enzymatic digestion on a GentleMACS with the following programs: the Miltenyi GentleMACS program “h_tumor_01_01” followed by 60 min at 37°C and 45 rpm for liver tissue; the Miltenyi GentleMACS program “37_C_Multi_H” for tissue from oral floor; 40 min at 37°C and -20 rpm for kidney, testes, bladder, pancreas and lung tissue; and 60 min at -20rpm and 37°C for colon tissue. Stomach tissue was digested for 15 min at 37°C on a rotating shaker. After digestion, cell suspension was filtered through a 100 µm strainer followed by filtration through a 40 µm strainer. The cell suspension was then magnetically enriched for CD45⁺ immune cells by positive selection, using TIL microbeads. Enrichment was performed according to the manufacturer’s protocol.

METHODS

Table 3: Digestion buffer for liver, kidney, testes and bladder tissue

Compound	Concentration
DMEM	
Collagenase Type IV	1 mg/mL
DNase I	50 µg/mL
HEPES	10 mM
FCS	10 %

Table 4: Digestion buffer for colon tissue

Compound	Concentration
RPMI	
Collagenase Type II	0.5 mg/mL
DNase I	50 µg/mL
HEPES	10 mM
FCS	7.5 %

Table 5: Digestion buffer for lung tissue

Compound	Concentration
DMEM	
Collagenase Type IV	1 mg/mL
DNase I	50 µg/mL
BSA	5 mg/mL

Table 6: Digestion buffer for stomach tissue

Compound	Concentration
DMEM	
Collagenase Type IV	2 mg/mL
DNase I	20 µg/mL
HEPES	10 mM
BSA	5 mg/mL

Table 7: Digestion buffer for pancreas tissue

Compound	Concentration
DMEM	
Collagenase Type IV	2 mg/mL

METHODS

DNase I	50 µg/mL
HEPES	10 mM
BSA	5 mg/mL

Table 8: Digestion buffer for oral floor tissue

Compound	Concentration
DMEM	
Collagenase Type IV	4 mg/mL
DNase I	50 µg/mL
HEPES	10 mM
FCS	2%

3.3 Restimulation for scRNA/TCR-seq experiments

After cell isolation from fresh tissues as described above, T cells were restimulated prior to sorting for scRNA/TCR-seq for investigating their effector transcriptional profile. For restimulation, we incubated total CD45⁺ T cells with a commercial Cell Stimulation Cocktail (eBioscience) containing phorbol-12-myristate-13-acetate (PMA) and Ionomycin for stimulation, and Brefeldin A and Monensin as transport inhibitors. Stimulation was done according to the manufacturer's instructions. The tumor sample, from which we are usually able to isolate more cells than from the NAT, was split into two, and one part was incubated with transport inhibitors only as a control. After adding the stimulation mix plus transport inhibitors, or transport inhibitors only, cells were incubated for 4h in a cell incubator (37°C, 5% CO₂).

3.4 Cell sorting

For scATAC-seq we sorted total CD3⁺ T cells, for scRNA/TCR-seq we sorted Treg cells, and filled up with total CD4⁺ T cells. To this end, cell suspensions pre-enriched for CD45⁺ immune cells were stained using the following antibodies: CD45-BUV737, CD3-BV786, CD19-APC, CD206-APC, CD4-R718, CD8-BB700, CD25-PE, CD127-BV711, CD45RA-BV510, CD45RO-PE-Cy7, CCR8-BV421, CD39-BV605, PD1-Viobright515, and Zombie NIR live/dead dye. The gating strategy is shown in **Figure 7**.

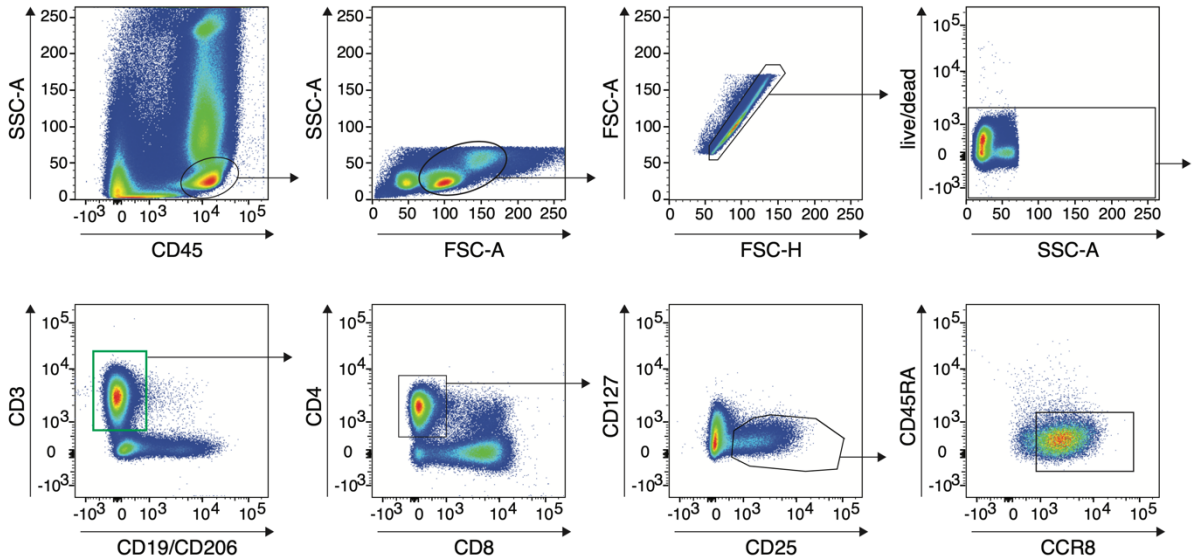


Figure 7: Gating strategy for sorting CD3⁺ T cells. CD45⁺ immune cells were gated, followed by FSC-A vs SSC-A for size and granularity. After doublet exclusion via FSC-A vs FSC-H, live cells were gated, followed by CD3⁺ T cells against a dump channel including CD19 and CD206. For scATAC-seq, total CD3⁺ T cells were sorted (sort gate marked in green). For analysis, we further gated CD4⁺ T cells against CD8⁺ T cells, and from CD4⁺ T cells we analysed Treg cells (CD127⁻ CD25^{hi}) and CCR8⁺ Treg cells (CD45RA⁻ CCR8⁺). For scRNA/TCR-seq, Treg cells were sorted, along with total CD4⁺ T cells.

3.5 scATAC-seq

3.5.1 scATAC-seq library preparation

For scATAC-seq library preparation, nuclei isolation was performed immediately after cell sorting. Cells were centrifuged (300 g, 5 min, 4°C), and were washed once with 100 μL PBS 0.04% BSA (300 g, 5 min, 4°C). Cells were re-suspended in 45 μL lysis buffer (**Table 10**) at 4°C, and lysis was performed for 2 min at 4°C. Lysis was stopped by adding 50 μL wash buffer (**Table 11**). Cells were centrifuged (500 g, 5 min, 4°C) and washed with diluted nuclei buffer pre-cooled to 4°C (500 g, 5 min, 4°C). Nuclei were then re-suspended in 7 μL diluted nuclei buffer, of which 1 μL was used for cell counting and 5 μL were used for transposition. Transposition and library prep were performed according to the manufacturer's protocol (Chromium Next GEM Single Cell ATAC Reagent Kits v1.1 User Guide, CG000209 Rev G).

METHODS

Table 9: Nuclei prep lysis buffer

Compound	Concentration
Nuclease-free water	1X
TRIS-HCL pH7.4	10 mM
NaCl	10 mM
MgCl ₂	3 mM
Tween-20	0.1%
NP-40	0.1%
Digitonin	0.01%
Bovine Serum Albumin	1%

Table 10: Nuclei prep wash buffer

Compound	Concentration
Nuclease-free water	1X
TRIS-HCL pH7.4	10 mM
NaCl	10 mM
MgCl ₂	3 mM
Tween-20	0.1%
Bovine Serum Albumin	1%

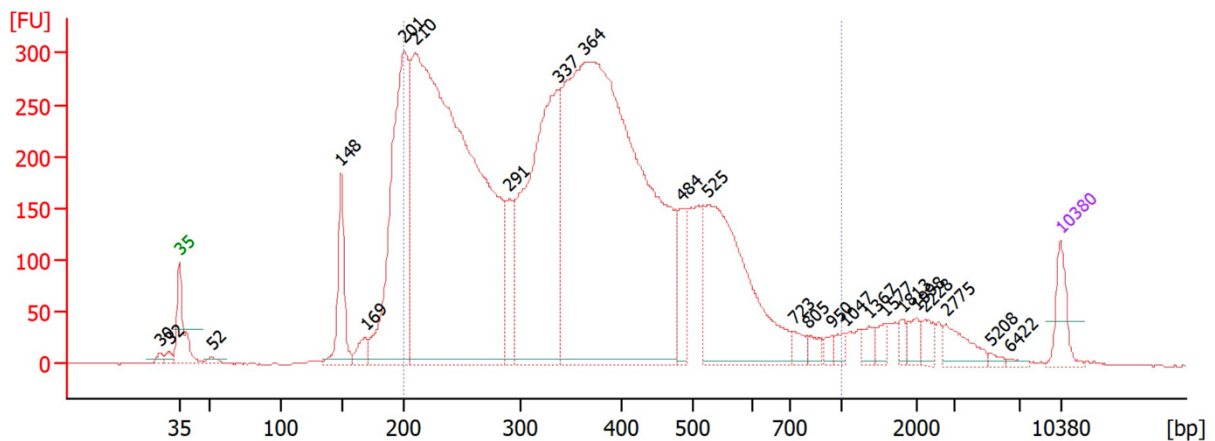


Figure 8: scATAC-seq library fragment length distribution as measured on a Bioanalyzer 2100 using a High Sensitivity DNA Assay. The fragment length distribution shows nucleosomal periodicity, which is an indication of a high-quality sample: Since fragments yielded by the Tn5 transposase may have no, one, or more nucleosomes in between, fragment sizes are enriched in 150 bp-steps, which is the circumference of a nucleosome.

3.5.2 Sequencing

scATAC-seq libraries were sequenced on a NextSeq 500/550 with a 75-cycle high-output cartridge. We sequenced the libraries paired-end, with 34 nucleotides from each side (PE 34-8-16-34) to a sequencing depth of 10,000 median high-quality fragments per cell.

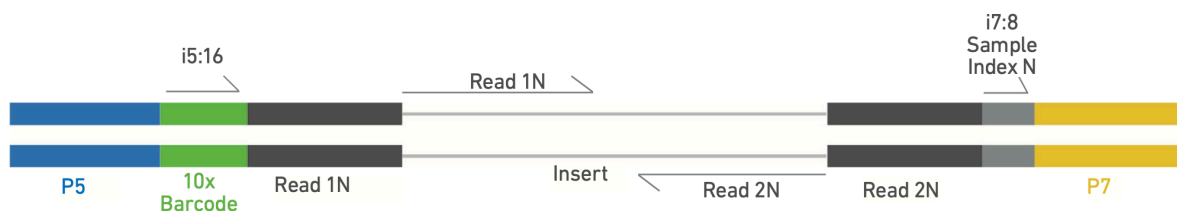


Figure 9: Schematic of scATAC-seq library fragment. P5 and P7 are complementary to Illumina flow cell adapters and are needed for bridge amplification. Read i5 is the 10X barcode and is 16 bp long. Read i7 is the sample index and is 8 bp long. Read 1N and read 2N cover the insert, into which we read 34 bp from each side with the PE 34-8-16-34 sequencing strategy. Figure printed from the Chromium Single Cell ATAC Reagent Kits User Guide (10X Genomics, CG000209).

3.5.3 Computation

scATAC-seq data analysis was performed on the computing cluster of the German Cancer Research Center (DKFZ) Heidelberg, in collaboration with Prof. Dr. Benedikt Brors, division of Applied Bioinformatics.

3.5.4 Alignment of reads

After conversion of bcl raw data to the FASTQ format using Illumina's *bcl2fastq()* (**Box 1**), reads were aligned to the reference genome using Cell Ranger ATAC (10X Genomics, v2), (**Box 2**). Reads were aligned against the reference genome GRCh38, and were filtered for the following quality control (QC) criteria: fragments had to map to the reference genome with a MAPQ > 30 on both reads, they further needed to map non-chimerically, were not allowed to be mitochondrial, and they were required to map to a primary contig. Further, duplicate fragments resulting from polymerase chain reaction (PCR) amplification were identified, and all fragments but one were discarded for each group of duplicates.

Box 1: bcl2fastq for the conversion of bcl format raw data to the FASTQ format

```
# run bcl2fastq
$ bcl2fastq --use-bases-mask=Y34,I8,Y16,Y34 \
  --create-fastq-for-index-reads \
  --minimum-trimmed-read-length=8 \
  --mask-short-adaptor-reads=8 \
  --ignore-missing-positions \
  --ignore-missing-controls \
  --ignore-missing-filter \
  --ignore-missing-bcls \
  -r 6 -w 6 \
  -R /media/raw_data/NextSeq/name_of_the_run \
  --output-dir=/media/raw_data/NextSeq/name_of_the_run/fastq \
  --sample-sheet=/media/raw_data/NextSeq/name_of_the_run/SampleSheet.csv \
  --no-lane-splitting
```

Box 2: Cell Ranger ATAC for aligning reads to the reference genome

```
# download the appropriate reference data from
# https://support.10xgenomics.com/single-cell-atac/software/downloads/latest

# move to the directory you want the output to be written to and prepend cellranger-
atac
$ cd ~/directory
$ export PATH=./path/to/cellranger-atac-2.0.0:$PATH

#run cellranger-atac count for several files (fastq files from multiple runs)
$ for x in samplename1 samplename2 samplename3;
  do cellranger-atac count\
    --localcores=10\
    --id=name_your_sample_"$x"\
    --reference=./path/to/reference/data\           #download see 10X documentation
    --fastqs=./path/to/fastq/files,./path/to/more/fastq/files\
    --sample="$x";
  done
```

3.5.5 Pre-processing of scATAC-seq data

Pre-processing of scATAC-seq data was performed following the ArchR (v1.0.1, (83)) workflow, and entailed creating the count matrix, filtering for high-quality cells, dimensionality reduction, clustering, and doublet identification and removal.

3.5.5.1 Creating the count matrix

The count matrix, which all further analysis is based on, was created from the *fragments.tsv* file output by Cell Ranger ATAC. The resultant count matrix contains fragments per genomic region (500 bp tiles) per cell. In ArchR, this is achieved by creating an *arrow file* for each sample, and samples that will be analyzed together are subsequently combined into an ArchRProject (**Boxes 3, 4**).

Box 3: Creating arrow files

```

# read in fragments files
inputFiles = vector()
n = 1
for(X in c(2,5:8,10:29,32:47)) {
  inputFiles[n] <- paste0("../..../02 CellRanger out/scATAC_",X,"
outs/fragments.tsv.gz")
  n = n+1
}

# name input files:
names(inputFiles) = vector()
n = 1
for(X in c(2,5:8,10:29,32:47)) {
  names(inputFiles)[n] <- paste0("scATAC_",X,"")
  n = n+1
}

# create arrow files
# Evaluate different thresholds depending on your data:
# - minTSS: Start with 0 to see all cells, afterwards evaluate which threshold
#           works for all of the samples
# - minFrag: Recommended to set >= 1000, otherwise the analysis might not be
#           robust enough

ArrowFiles <- createArrowFiles(
  inputFiles = inputFiles,
  sampleNames = names(inputFiles),
  minTSS = 0,
  minFrag = 1000,
  addTileMat = TRUE,
  addGeneScoreMat = TRUE,
  force = TRUE
)

```

Box 4: Creating the ArchRProject (exemplary case of kidney sample CCRC1)

```

# Define arrow files
ArrowFiles = c("../..../ArchRProject_raw/ArrowFiles/scATAC_14.arrow",
"../..../ArchRProject_raw/ArrowFiles/scATAC_15.arrow")

# Create ArchRProject
proj_kidney_14_15 <- ArchRProject(
  ArrowFiles = ArrowFiles,
  copyArrows = TRUE
)

```

3.5.5.2 Quality filtering

Cells containing less than 1000 unique fragments per cell were discarded, as were cells with a transcription start site (TSS) enrichment < 10 (**Box 5**). The TSS enrichment is calculated as the average accessibility of each TSS region (50 bp) divided by the average accessibility of TSS-flanking regions, and is a measure of sample quality: Large protein complexes at TSS render the chromatin more accessible compared to non-TSS regions. A low TSS enrichment suggests random transposition across the

genome, which occurs in dead or dying cells that have lost the integrity of their chromatin. We further filtered cells for barcodes that were identified as cell barcode by Cell Ranger ATAC, keeping cells with “is_cell_barcode = 1” in the *singlecell.csv* file output by *cellranger ATAC count*, excluding gel bead doublets, barcode multiplets and low-targeting barcodes (**Box 6**). If samples that were analyzed together (tumor and NAT sample from the same patient) had a significantly different number of unique fragments per cell, the sample with the greater sequencing depth was downsampled to match the median unique fragments per cell of the sample with lower sequencing depth.

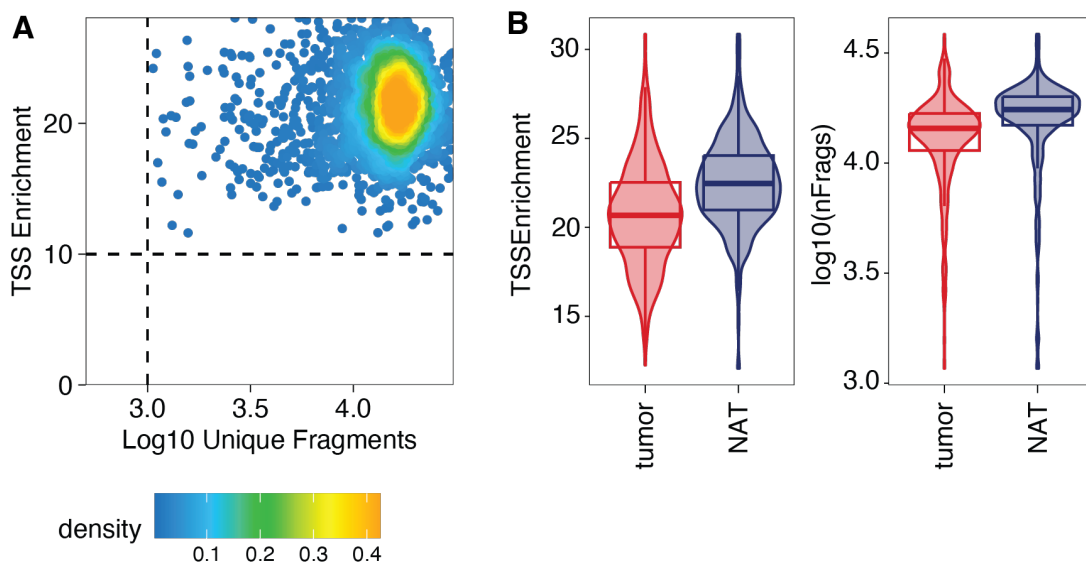


Figure 10: Quality filtering. **A.** TSS enrichment vs log10(fragments). Cells with less than 1000 unique fragments per cell as well as cells with a TSS enrichment < 10 are discarded. **B.** TSS enrichment and number of fragments for each sample separately.

Box 5: QC filtering step 1

```
# Filter for cells passing the TSS enrichment cut-off determined above
proj_kidney_14_15 <-
proj_kidney_14_15[proj_kidney_14_15@cellColData$TSSEnrichment >= 10, ]
```

Box 6: QC filtering step 2

```
# read in files
singlecell_14 <-
read.csv("/omics/groups/OE0436/internal/Regensburg/Mainz/02_CellRanger_out/scATAC_14_outs/singlecell.csv")
singlecell_15 <-
read.csv("/omics/groups/OE0436/internal/Regensburg/Mainz/02_CellRanger_out/scATAC_15_outs/singlecell.csv")
```

METHODS

```
# make an extra column for the singlecell_X df with the cell barcode with the
"scATAC_X#" in front as in proj_scATAC_X@cellColData
# make vector with "scATAC_14_15#" of length of the table
scATAC_14_vector <- c(rep("scATAC_14#", nrow(singlecell_14)))
scATAC_15_vector <- c(rep("scATAC_15#", nrow(singlecell_15)))

# merge scATAC_X_vector with singlecell_X
singlecell_14_bc <- cbind(scATAC_14_vector, singlecell_14)
singlecell_15_bc <- cbind(scATAC_15_vector, singlecell_15)

# create a single column merging the columns "scATAC_X_vector" and "barcode"
singlecell_14_fullbc <- singlecell_14_bc %>% unite("full_barcode",
scATAC_14_vector:barcode, remove = FALSE, sep = "")
singlecell_15_fullbc <- singlecell_15_bc %>% unite("full_barcode",
scATAC_15_vector:barcode, remove = FALSE, sep = "")

# rowbind the two dataframes
DT1 <- singlecell_14_fullbc
DT2 <- singlecell_15_fullbc
l <- list(DT1,DT2)
singlecell_14_15_fullbc <- rbindlist(l, use.names = FALSE, fill = FALSE)

# extract rownames that are also in the ArchRProject
rownames_archr <- rownames(proj_kidney_14_15)
subset_singlecell_14_15_fullbc <-
singlecell_14_15_fullbc[singlecell_14_15_fullbc$full_barcode %in% rownames_archr,
]

# extract is_cell_barcode column from singlecell.csv and give it barcodes as
rownames
df_is_cell_barcode <-
as.data.frame(subset_singlecell_14_15_fullbc$is_cell_barcode)
rownames(df_is_cell_barcode) <- subset_singlecell_14_15_fullbc$full_barcode

# order is_cell_barcode the way the ArchRProject is ordered and create filter
is_cell_barcode <- df_is_cell_barcode[order(match(rownames(df_is_cell_barcode),
rownames_archr)), ]
filter_archr <- is_cell_barcode==1

# filter ArchRProject for is_cell_barcode==1
proj_14_15_cellsPass <- proj_kidney_14_15[filter_archr, ]
```

3.5.5.3 Dimensionality reduction and clustering

Due to the sparsity of scATAC-seq data, an iterative latent semantic indexing (LSI) approach is implemented in ArchR for dimensionality reduction, which is performed on genomic tiles 500 bp wide (83). The term frequency (TF) was first calculated with depth normalization for each single cell. In order to penalize accessible regions that are not specific, the term frequency – inverse document frequency (TF-IDF) was calculated, followed by log-transformation of the resulting TF-IDF matrix. In order to minimize batch effects, the first LSI transformation identifies the most accessible features, on which the most variable features are identified in the following iterations (**Box 7**). After dimensionality reduction, data can be visualized as Uniform Manifold Approximation and Projection (UMAP). Clustering was performed using the Louvain algorithm (84)

(**Box 8**), and clustering resolution was adjusted after manual cell type identification based on gene scores of marker genes (see section “Cell type annotation”).

Box 7: Dimensionality reduction using iterative LSI

```
# LSI dimensionality reduction
proj_14_15_cellsPass <- addIterativeLSI(
  ArchRProj = proj_14_15_cellsPass,
  useMatrix = "TileMatrix",
  name = "IterativeLSI",
  iterations = 2,
  clusterParams = list(
    resolution = c(0.2),
    sampleCells = 10000,
    n.start = 10
  ),
  varFeatures = 25000,
  dimsToUse = 1:30,
  force = TRUE
)
```

Box 8: Clustering using the Louvain algorithm

```
# clustering:
proj_14_15_cellsPass <- addClusters(
  input = proj_14_15_cellsPass,
  reducedDims = "IterativeLSI",
  method = "Seurat",
  name = "Clusters",
  resolution = 0.8, #adjust later on according to manual cell type identification
  force = TRUE
)
```

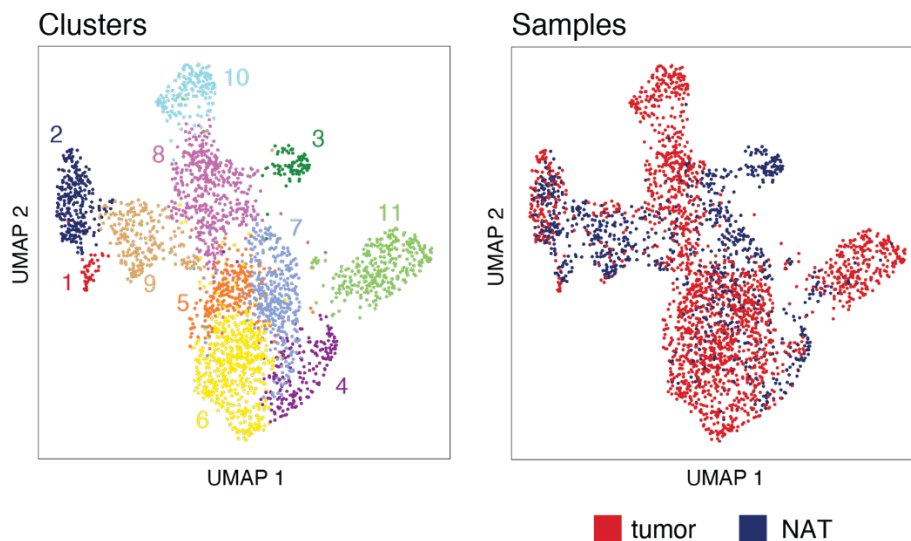


Figure 11: Data visualized as UMAP plot after dimensionality reduction and clustering. Left: UMAP encoding cluster membership. Right: UMAP encoding sample information.

3.5.5.4 Doublet filtering

In ArchR, doublet identification is performed as follows (83): Synthetic doublets of any two cells in the dataset are calculated and projected onto the UMAP embedding, and their nearest neighbors are identified using the k-nearest neighbor (kNN) algorithm. Doublet enrichments are calculated based on the similarity of a cell to synthetic doublets, compared to what would be expected assuming a uniform distribution (**Box 9**). The filter ratio to apply to the dataset was then chosen taking into account the expected doublets depending on the number of nuclei loaded onto the chip (Chromium Next GEM Single-cell ATAC Reagent Kits v1.1 User Guide CG000209). Since some cells might naturally have a higher doublet score (e.g. intermediate cell states), additional parameters such as the number of unique fragments per cell may also be taken into account. We further calculated gene scores of marker genes for certain cell types and overlaid them on the UMAP prior to filtering doublets, in order to make a biologically informed decision.

Box 9: Doublet filtering

```
# Calculate doublet scores on the ArchRProject
proj_14_15_cellsPass <- addDoubletScores(
  input = proj_14_15_cellsPass,
  k = 10,
  knnMethod = "UMAP",
  LSIMethod = 1,
  force = TRUE
)

# After trying out different filter ratios, create new ArchRProject and filter
doublets with a filterRatio of 1
proj_14_15_doubfilter <- filterDoublets(
  proj_14_15_cellsPass,
  filterRatio = 1,
)
```

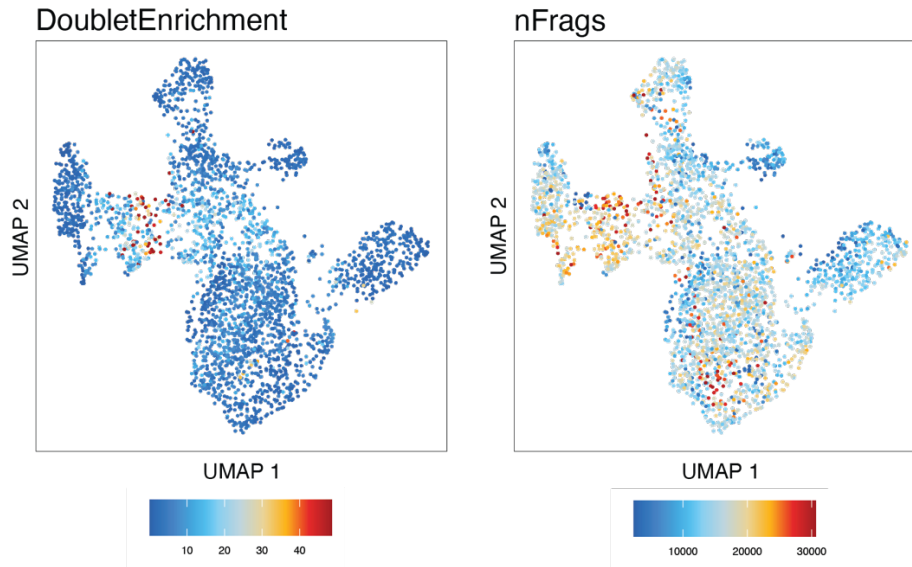


Figure 12: UMAP encoding doublet enrichment score and number of fragments per cell. Cells with a high doublet enrichment score are not distributed evenly across the UMAP but lie between clusters indicate doublets (left). Doublets further are generally expected to have a higher number of fragments per cell (right).

3.5.6 scATAC-seq data analysis

scATAC-seq data analysis comprised general immune cell type annotation and identification of CCR8⁺ Treg cells using a pre-defined signature, peak calling and the identification of differential peaks, and transcription factor (TF) motif enrichment within these differential peaks. Further, chromatin accessibility of specific gene regions in different cell types was analyzed using the integrative genomics viewer (IGV).

3.5.6.1 Cell type annotation

Cell type annotation was performed manually based on the gene scores of pre-defined marker genes. Gene scores are calculated as a proxy for gene expression based on the accessibility of gene regions. The function for gene score calculation implemented in ArchR (83) takes into account the tiles that cover the gene body as well as the promoter region, and that do not intersect with additional genes. Since large genes tend to have more peaks than small genes, the latter get larger weights for normalization. Since our dataset contains total CD3⁺ T cells, CD4 and CD8A were used as markers for T helper and cytotoxic T cells. CD4⁺ T cells were further subdivided into Treg cells (*IL2RA*, *FOXP3*, *ectonucleoside triphosphate diphosphohydrolase 1 (ENTPD1)*, *IRF4*, *nuclear factor 'kappa-light-chain-enhancer' of*

METHODS

activated B-cells (*NFKB1*), nuclear receptor subfamily 4 group A member 1 (*NR4A1*), *CTLA4*, inducible T cell costimulator (*ICOS*), tumor necrosis factor receptor superfamily (*TNFRSF4*), *TNFRSF9*, *TNFRSF18*, *CD28*, T cell factor (*TCF7*), *IL10*, granzyme B (*GZMB*), PR domain zinc finger protein 1 (*PRDM1*), *IL2*, CCR8⁺ Treg cells (*BATF*, *CCR8*), pTreg cells (*RAR-related orphan receptor C (ROR)C*, *IKZF2*), and Tconv cells (Th1 cells: *signal transducer and activator of transcription (STAT)1*, *STAT4*, *T-box transcription factor (TBX)21*, *IFNG*, *TNF*, *lymphotoxin alpha (LTA)*; Th2 cells: *GATA3*, *IRF4*, *STAT5A*, *STAT6*, *IL4*, *IL5*, *IL9*, *IL10*, *IL13*, *IL21*; Th17 cells: *IRF4*, *RORA*, *RORC*, *STAT3*, *C-C motif chemokine ligand (CCL)20*, *IL17A*, *IL17F*, *IL22*, *IL26*), and CD8⁺ T cells were further subdivided into effector (*TBX21*, *C-X3-C motif chemokine receptor (CX3CR)1*, *Zinc finger E-box binding homeobox (ZEB)2*, *Kruppel-like factor (KLF)2*, *fibroblast growth factor binding protein (FGFBP)2*), exhausted (*thymocyte selection-associated high mobility group box protein (TOX)*, *ENTPD1*, *hepatitis A virus cellular receptor 2 (HAVCR2)*, *CTLA4*, *Layilin (LAYN)*, *PDCD1*, *TNFRSF9*) and PD1⁺CD39⁻ (*PDCD1*, *ENTPD1*) cells. Gene scores were overlaid on the UMAP embedding with Markov affinity-based graph imputation of cells (MAGIC) smoothing (85) in order to facilitate visual interpretation and manual cluster annotation (**Box 10**). As for our cell type of interest, CCR8⁺ Treg cells, a peak signature derived from healthy human skin and fat was available, along with a core Treg cell signature (20), we calculated z-scores using the *addDeviationMatrix* function from ArchR to confirm CCR8⁺ Treg cells identified with the marker gene approach (**Box 11**).

Box 10: Cell type annotation based on gene scores

```
# Overlay gene scores on UMAP embedding of proj, use MAGIC smoothing
# Define which genes to plot
markerGenes <- c(
  "CD4", "CD8", #divide in T helper and cytotoxic T cells
  "IL2RA", "FOXP3", "ENTPD1", "IRF4", "NFKB1", "NR4A1", "CTLA4", "ICOS", "TNFRSF4",
  "TNFRSF9", "TNFRSF18", "CD28", "TCF7", "IL10", "GZMB", "PRDM1", "IL2", #Treg
  "BATF", "CCR8", #CCR8+ Treg
  "RORC", "IKZF2", #pTreg
  "STAT1", "STAT4", "TBX21", "IFNG", "TNF", "LTA", #Th1
  "GATA3", "IRF4", "STAT5A", "STAT6", "IL4", "IL5", "IL9", "IL10", "IL13", "IL21", #Th2
  "IRF4", "RORA", "RORC", "STAT3", "CCL20", "IL17A", "IL17F", "IL22", "IL26", #Th17
  "TBX21", "CX3CR1", "ZEB2", "KLF2", "FGFBP2", #CD8eff
  "TOX", "ENTPD1", "HAVCR2", "CTLA4", "LAYN", "PDCD1", "TNFRSF9" #CD8exh
)

# Impute weights for smoothing via MAGIC
proj_14_15_doubfilter <- addImputeWeights(proj_14_15_doubfilter)

# Overlay gene scores on UMAP for visual interpretation
magic_genes <- plotEmbedding(
```

METHODS

```
ArchRProj = proj_14_15_doubfilter,  
colorBy = "GeneScoreMatrix",  
name = markerGenes,  
embedding = "UMAP",  
plotAs = "points",  
imputeWeights = getImputeWeights(proj_14_15_doubfilter),  
size = 0.5  
)
```

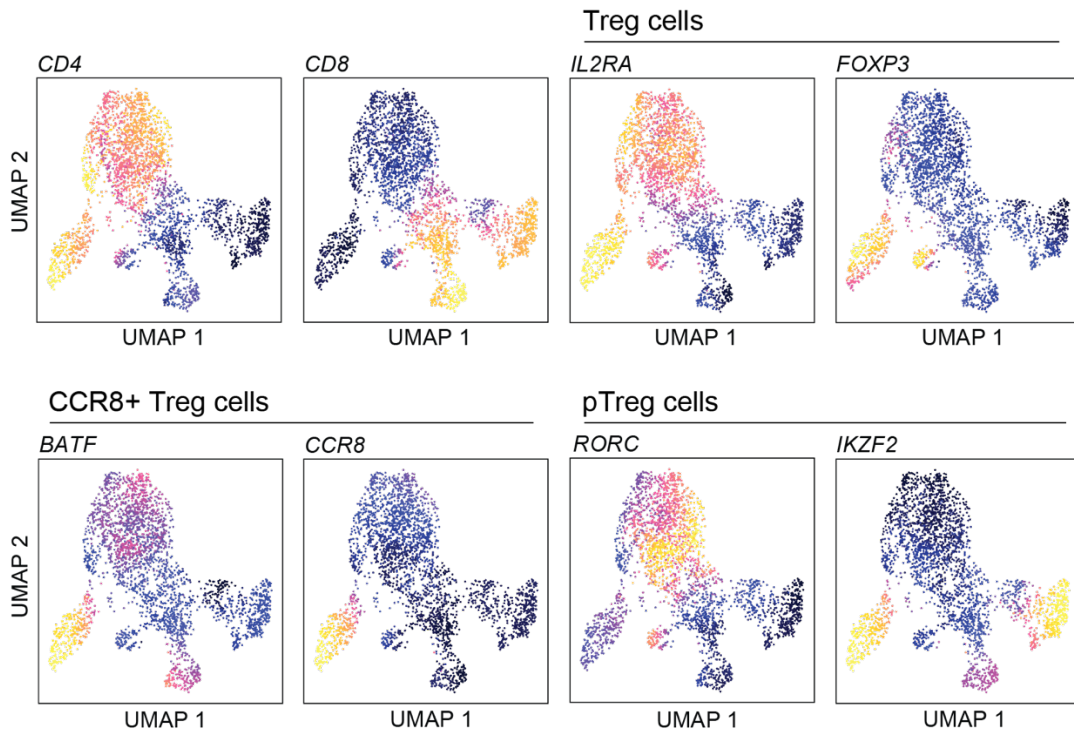


Figure 13: UMAP plots encoding gene scores with MAGIC smoothing applied. Overlaying gene score information on the UMAP plots greatly facilitates visual interpretation and manual cluster-based cell type annotation. IL2RA and FOXP3 shown as exemplary marker genes for Treg cells, BATF and CCR8 for CCR8⁺ Treg cells, and RORC and IKZF2 for identifying pTreg cells.

Box 11: Calculate signature z-scores for identifying CCR8⁺ Treg cells

```
archr_add_peak_signatures = function(proj, signature_list, signature_name){  
  #signature_list: list of granges  
  #signature_name: name string for the set of signatures  
  add_df_to_cellcoldata = function(pro, pheno_df, force=FALSE){  
    stopifnot(identical(rownames(pro@cellColData), rownames(pheno_df)))  
    cnames = colnames(pheno_df)  
    for(i in 1:ncol(pheno_df)){  
      pro = addCellColData(ArchRProj = pro, data=pheno_df[, i], name =  
cnames[i],  
                           cells = rownames(pro@cellColData), force = force)  
    }  
    return(pro)  
  }  
  if(length(signature_list)<2){  
    stop('Currently, only works if at least two signatures are provided')  
  }  
  for(i in seq_along(signature_list)){  
    names(signature_list[[i]]) = NULL  
  }  
}
```

METHODS

```
}
proj <- addPeakAnnotations(ArchRProj = proj,
                          regions = signature_list,
                          name = signature_name,
                          force=T)

method_use = "chromVAR" #does only work with fixed width peaks
if(any(sapply(signature_list, function(x) length(unique(width(x)))) > 1)){
  method_use = 'ArchR'
}

proj = addBgdPeaks(proj, force = T, method=method_use)

proj = addDeviationsMatrix(
  ArchRProj = proj,
  peakAnnotation = signature_name,
  binarize=TRUE,
  bgdPeaks = getBgdPeaks(proj, method = method_use),
  force = TRUE
)

dr_df = as.data.frame(proj@cellColData)
sig_se = getMatrixFromProject(proj, paste0(signature_name, 'Matrix'))
z_score_mat = t(assays(sig_se)[['z']])
z_score_mat = z_score_mat[match(rownames(dr_df), rownames(z_score_mat)), ]
colnames(z_score_mat) = paste0('z_', colnames(z_score_mat))
stopifnot(identical(rownames(z_score_mat), rownames(dr_df)))
dev_score_mat = t(assays(sig_se)[['deviations']])
dev_score_mat = dev_score_mat[match(rownames(dr_df), rownames(dev_score_mat)), ]
colnames(dev_score_mat) = paste0('dev_', colnames(dev_score_mat))
stopifnot(identical(rownames(dev_score_mat), rownames(dr_df)))

proj = add_df_to_cellcoldata(proj, z_score_mat, force=T)
proj = add_df_to_cellcoldata(proj, dev_score_mat, force=T)
return(proj)
}

signature_list <- list(tisTreg_sig = tisTregGR, core_Treg_sig = core_TregGR)

proj_14_15_doubfilter <- archr_add_peak_signatures(proj_14_15_doubfilter,
signature_list, "signatures")
```

CCR8⁺ Treg signature

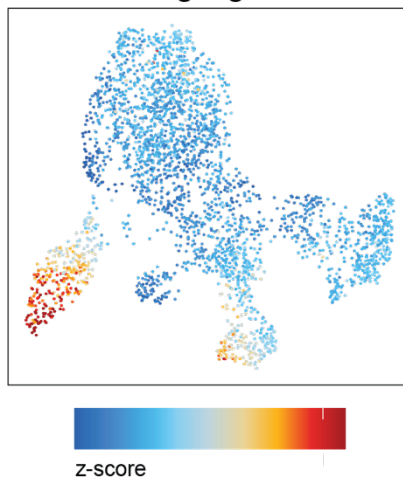


Figure 14: UMAP encoding CCR8⁺ Treg signature z-score. Z-score was calculated for each cell to determine the similarity to CCR8⁺ Treg cells using a published signature. This confirms the cell type annotation for CCR8⁺ Treg cells done manually using gene scores.

3.5.6.2 Differential peaks and motif enrichment

Peak calling was performed on pseudobulk replicates which were generated with the approach implemented in ArchR (83). Pseudobulk data was generated on the clusters in a sample-aware fashion by adjusting the *minCells* and *minReplicates* parameters accordingly (**Box 12**). Peak calling was performed using the iterative peak merging algorithm implemented in MACS2 (86) (**Box 13**). Differential peaks between certain clusters, as well as marker peaks for a certain cluster (i.e. differential peaks for the cluster of interest against all other clusters), were calculated using a Wilcoxon rank-sum test (87) with Benjamini-Hochberg multiple hypothesis correction (88). Cutoffs for the false discovery rate (FDR) and effect size were set to $FDR \leq 0.01$ and $\text{Log}_2FC \geq 2$, respectively (**Box 14**). TF motifs enriched in marker- or differential peaks were identified based on motif annotations (**Box 15**). The accessibility of TF motifs can further be evaluated using the ChromVAR package (89), where the deviation of the accessibility of each motif from the expected motif accessibility is calculated on a per-cell basis.

Box 12: Pseudobulk replicates

```
# The key parameter here is groupBy, which defines the groups for which pseudo-
bulk replicates should be made
proj_14_15_doubfilter <- addGroupCoverages(
  ArchRProj = proj_14_15_doubfilter,
  groupBy = "Clusters",
  minCells = 40,
  maxCells = 500,
  minReplicates = 2,
  maxReplicates = 5,
  sampleRatio = 0.8
)
```

Box 13: Peak calling using iterative peak merging

```
# create peak matrix using MACS2
pathToMacs2 = findMacs2()

proj_14_15_doubfilter <- addReproduciblePeakSet(
  ArchRProj = proj_14_15_doubfilter,
  groupBy = "Clusters",
  pathToMacs2 = pathToMacs2
)

# add peak matrix to ArchRProject
proj_14_15_doubfilter <- addPeakMatrix(proj_14_15_doubfilter)
```

METHODS

Box 14: Extract differential peaks e.g. between CCR8⁺ Treg cells and all other CD3⁺ T cells

```
# calculate differential peaks between Treg cluster and CCR8+ Treg cluster
markerTest <- getMarkerFeatures(
  ArchRProj = proj_14_15_doubfilter,
  useMatrix = "PeakMatrix",
  groupBy = "Clusters_tisTreg",
  testMethod = "wilcoxon",
  bias = c("TSSEnrichment", "log10(nFrag)",
  useGroups = "tisTreg",
  bgdGroups = "other"
)

# visualize as volcano plot:
pv2 <- markerPlot(seMarker = markerTest,
  name = "tisTreg",
  cutOff = "FDR <= 0.01 & abs(Log2FC) >= 2",
  plotAs = "Volcano"
)
```

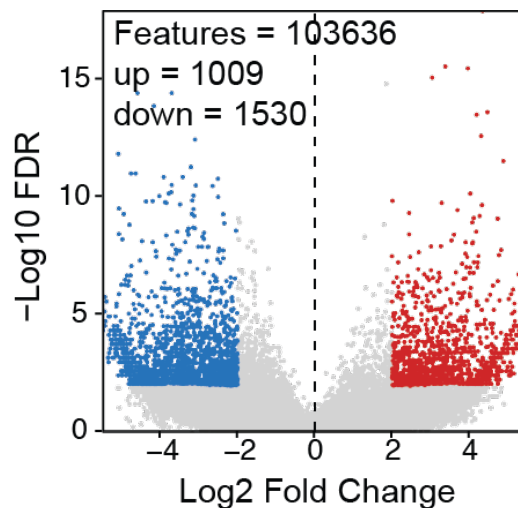


Figure 15: Differential peaks between CCR8⁺ Treg cells and all other CD3⁺ T cells presented as volcano plot. Effect size (log₂ fold change) is shown on the x-axis and significance (-log₁₀ of the FDR adjusted p-value) is shown on the y-axis. Out of 103636 peaks in total, 1009 peaks are up-regulated in CCR8⁺ Treg cells as compared to Treg cells, and 1530 peaks are down-regulated in CCR8⁺ Treg cells as compared to all other CD3⁺ T cells.

Box 15: TF motif enrichment in differential peaks e.g. between CCR8⁺ Treg cells and all other CD3⁺ T cells

```
# add motif annotations to the ArchRProject
proj_14_15_doubfilter <- addMotifAnnotations(ArchRProj = proj_14_15_doubfilter,
  motifSet = "cisbp", name = "Motif")

# calculate TF motifs enriched in CCR8+ Treg cells vs CD3+ T cells
motifsUp <- peakAnnoEnrichment(
  seMarker = markerTest, #defined above as CCR8 Treg vs all other CD3+ T cells
  ArchRProj = proj_14_15_doubfilter,
  peakAnnotation = "Motif",
  cutOff = "FDR <= 0.01 & Log2FC >= 2"
```

METHODS

```
)  
  
# calculate TF motifs enriched in all other CD3+ T cells vs CCR8+ Treg cells  
motifsDo <- peakAnnoEnrichment(  
  seMarker = markerTest,  
  ArchRProj = proj_14_15_doubfilter,  
  peakAnnotation = "Motif",  
  cutOff = "FDR <= 0.01 & Log2FC <= -2"  
)
```

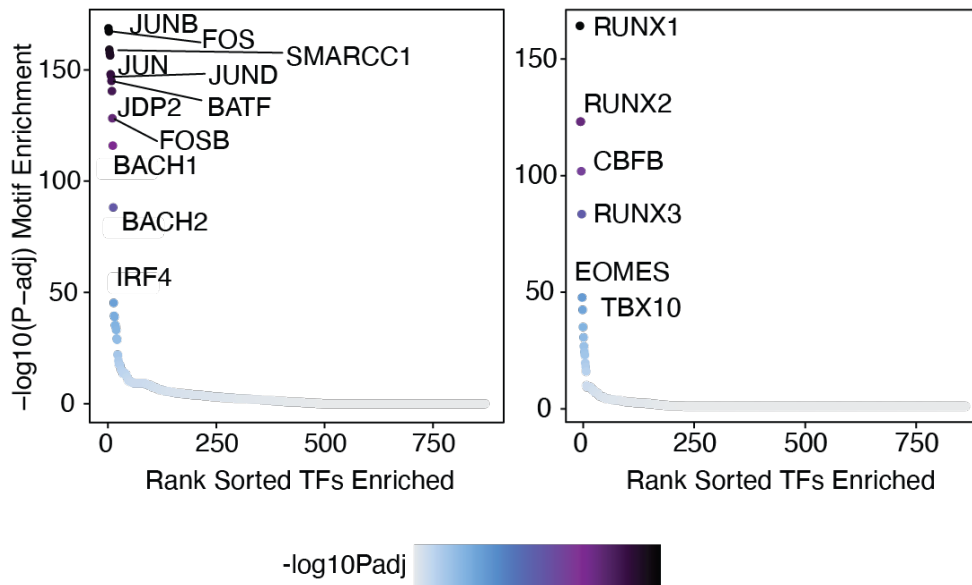


Figure 16: TF motif enrichment in differential peaks between CCR8⁺ Treg cells and all other T cells. Left: TF motifs enriched in peaks upregulated in CCR8⁺ Treg cells compared to all other CD3⁺ T cells. Motifs are sorted by p value rank. Right: TF motifs enriched in peaks upregulated in all other CD3⁺ T cells compared to CCR8⁺ Treg cells.

3.5.6.3 Browser tracks

Browser tracks were created using IGV v2.17.2 (90). Separate bigwig files for each annotated cell type for a patient were exported from ArchR as shown in **Box 16** and were loaded into IGV.

Box 16: Export bigwig files for analysis using IGV

```
# export bigwig files by cell type  
getGroupBW(  
  ArchRProj = proj_14_15_doubfilter,  
  groupBy = "Clusters_anno"  
)
```

METHODS

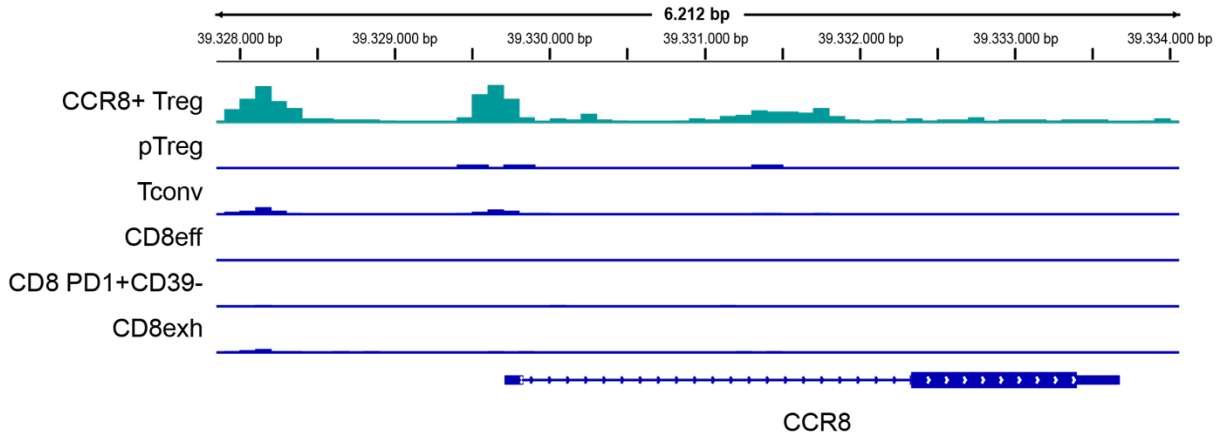


Figure 17: IGV browser tracks showing the CCR8 gene for one kidney patient by cell type. Track heights are normalized across all cell types.

3.6 scRNA/TCR-seq

3.6.1 scRNA/TCR-seq library preparation

Library preparation for scRNA/TCR-seq was performed immediately after cell sorting. Cells were first loaded on a 10X Chromium Next GEM Chip K (10X Genomics #1000287), followed gene expression (GEX) library prep using the Chromium Next GEM Single Cell 5' Reagent Kit v2 and library prep of the hashtag library using the 5' Feature Barcode Kit. V(D)J amplification was performed on the GEX library using Chromium Single Cell Human TCR Amplification Kit, followed by V(D)J library prep. Libraries were prepared according to the manufacturer's protocol (Chromium Next GEM Single Cell 5' Reagent Kits v2 (Dual Index) User Guide, CG000330 Rev A).

3.6.2 Sequencing

scRNA/TCR-seq libraries were sequenced on a NextSeq 500/550 using a 150-cycle high-output cartridge, with a PE 26-10-10-90 sequencing strategy.

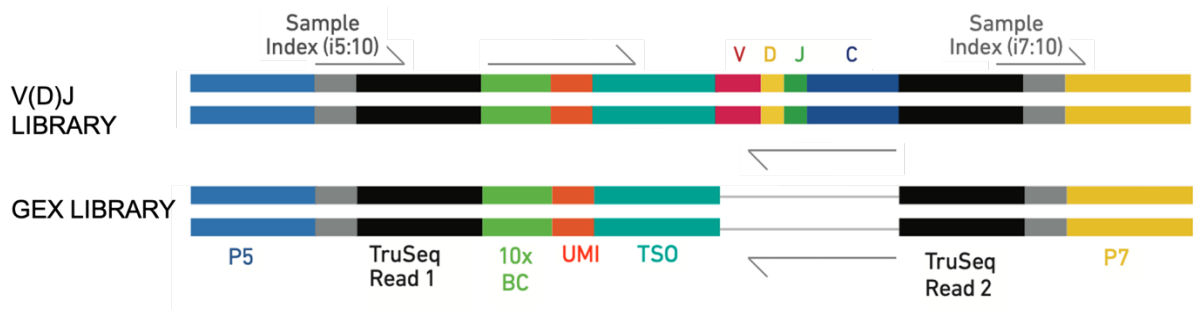


Figure 18: Schematic of scRNA-seq gene expression (GEX) and V(D)J library fragments. P5 and P7 are complementary to Illumina flow cell adapters and are needed for bridge amplification. Read i5

and i7 are sample indices and are 10 bp long. The 10X barcode and the unique molecular identifier (UMI) are covered by read 1N, which is 26 bp in our sequencing strategy. Read 2N is 90 bp and covers the insert, which in the V(D)J library are the V, D and J regions of the TCR. Figure adapted from the Chromium Next GEM Single Cell 5' Reagent Kits User Guide (10X Genomics, CG000331).

3.6.3 Computation

scRNA/TCR-seq data analysis was performed on the MOGON NHR computing cluster in Mainz, through the affiliation with the SFB TRR355.

3.6.4 Alignment of reads

After creating FASTQ files from the raw bcl output (described in detail in scATAC-seq section), Cell Ranger (10X Genomics, v7.1) was used for aligning scRNA-seq reads to the reference genome GRCh38. *Cellranger multi* was used to first assign cells to samples based on the hashtag library. For each sample, *Cellranger multi* was run again on the FASTQ files created after demultiplexing the samples, integrating the GEX and V(D)J library information (**Box 17**).

Box 17: Cell Ranger for aligning reads to the reference genome

```
$ sh CellRanger.sh samplename date

$ cat CellRanger.sh
SAMPLE=$1
DATE=$2

PROJECTPATH=/path/to/output/folder

#move to output folder
cd ${PROJECTPATH}

#run Cellranger multi to combine GEX and FB
${PROJECTPATH}/../../../../CellRanger/cellranger-7.1.0/bin/cellranger multi --
jobmode local --disable-ui --id ${DATE}_${SAMPLE} --csv
${PROJECTPATH}/csv_files/configmulti.csv

cd ${PROJECTPATH}/${DATE}_${SAMPLE}
mkdir bamtofastq

# now run for each tissue (=sample-tagged sample)
for TISSUE in Tumor_restim Tumor NAT_restim
do
    #run bamtofastq
    ${PROJECTPATH}/../../../../CellRanger/cellranger-7.1.0/lib/bin/bamtofastq --
reads-per-fastq=1000000000
outs/per_sample_outs/${TISSUE}/count/sample_alignments.bam bamtofastq/${TISSUE}

    #run CellRangermulti to include VDJ
    ${PROJECTPATH}/../../../../CellRanger/cellranger-7.1.0/bin/cellranger multi -
-jobmode local --disable-ui --id ${DATE}_${SAMPLE}_${TISSUE} --csv
${PROJECTPATH}/csv_files/${SAMPLE}_${TISSUE}.csv
```

```

done

$ cat configmulti.csv
[gene-expression]
reference,/path/to/CellRanger/refdata-gex-GRCh38-2020-A
cmo-set,/path/to/cmo.csv
check-library-compatibility,false

[libraries]
fastq_id,fastqs,feature_types
Kidney1,/path/to/FeaturBarcode/FASTQs,Multiplexing Capture
Kidney1,/path/to/GEX/FASTQs,Gene Expression

[samples]
sample_id,cmo_ids
Tumor_restim,C1
Tumor,C2
NAT_restim,C3

$ cat cmo.csv
id,name,read,pattern,sequence,feature_type
C1,HTO_C0301,R2,5PNNNNNNNNNNN (BC) NNNNNNNNNN,GTCAACTCTTTAGCG,Multiplexing Capture
C2,HTO_C0302,R2,5PNNNNNNNNNNN (BC) NNNNNNNNNN,TGATGGCCTATTGGG,Multiplexing Capture
C3,HTO_C0303,R2,5PNNNNNNNNNNN (BC) NNNNNNNNNN,TTCCGCCTCTCTTG,Multiplexing Capture

```

3.6.5 Pre-processing of scRNA/TCR-seq data

Pre-processing and analysis of scRNA/TCR-seq data was performed using a set of R packages from CRAN (<https://www.r-project.org>) and Bioconductor (91). Pre-processing of scRNA/TCR-seq data included filtering for high-quality cells containing TCR information, doublet removal, as well as dimensionality reduction and clustering.

3.6.5.1 Creating the count matrix

The count matrix for each sample was built from the filtered feature-barcode matrix output by Cell Ranger using the *Read10X* function from the Seurat package (92). The filtered feature-barcode matrix only contains barcodes that were identified by Cell Ranger as cell barcodes. Then, the count matrix was converted into a SingleCellExperiment (SCE) object, which all further pre-processing and analysis steps are based on, using the SingleCellExperiment package (93). To the SCE object, TCR information was added, which is stored in the *clonotypes.csv* file output by Cell Ranger (**Box 18**). The AnnotationHub package (94) was used to convert gene identifiers in to more easily readable gene names (**Box 19**).

Box 18: Creating the SCE object

```

# Function to read in the data
# provide all the filepaths to the count data as a list
# as well as a list of the respective tissues
readDataset = function(filepath_list, tissue) {
  sceRNA = list()
  # iterate over each sample of the input data
  for (i in 1:length(filepath_list)) {
    # read the count data
    counts = Read10X(filepath_list[[i]])
    # generate a SingleCellExperiment object
    sce = SingleCellExperiment(assays = list(counts = counts#`Gene Expression`
                                           ))
    # Add the tissue type information as meta data
    sce$tissue = rep(tissue[[i]], ncol(sce))
    sceRNA = c(sceRNA, sce)
  }
  # return the list of SingleCellExperiment objects
  return(sceRNA)
}

# supply paths to filtered feature barcode matrices
filepaths <- c(
  "path/to/tissue1/sample_filtered_feature_bc_matrix",
  "path/to/tissue2/sample_filtered_feature_bc_matrix",
  "path/to/tissue3/sample_filtered_feature_bc_matrix"
)

sceRNA <- readDataset(filepaths,
  tissue = c("Kidney1_NAT_restim", "Kidney1_Tumor_restim", "Kidney1_Tumor"))

# name objects in the list
names(sceRNA) <- c("Kidney1_NAT_restim", "Kidney1_Tumor_restim", "Kidney1_Tumor")

```

Box 19: Converting gene identifiers to gene names

```

sce <- sceRNA$Kidney1_NAT_restim

# set up the annotation hub
ah = AnnotationHub()

# extract the identifiers and names for human data
query(ah, c("Homo sapiens", "Ensembl", "EnsDb"))
AH109606 = ah[["AH109606"]]
genes(AH109606)[, 2]

# search for mitochondrial genes
is.mito <- grepl("^MT-", rownames(sce))

chr.loc <- mapIds(
  AH109606,
  keys = rownames(sce),
  keytype = "GENENAME",
  column = "SEQNAME"
)
is.mito <- which(chr.loc == "MT")

```

3.6.5.2 Quality filtering

Quality- and doublet filtering was performed on each sample separately. In a first filtering step, cells which had no TCR information or ambiguous TCR information (cell doublets) were discarded (**Box 20**). For quality filtering, the library size as well as the amount of mitochondrial DNA were considered. For setting cut-offs for these parameters, the *addPerCellQC* function from the *scater* package (95) was used (**Box 21**). For each quality control parameter, the function calculates the median across all cells, as well as the median absolute deviation (MAD) for each cell. Cells differing more than 3 MAD from the median in one or more QC parameters are excluded from analysis.

Box 20: QC filtering step 1

```
addTCRMetaData = function(sce, tcr_filepath, clonotypes_filepath) {
  # Read in TCR information
  tcr = read.csv(tcr_filepath)
  clonotypes = read.csv(clonotypes_filepath)

  # Remove duplicated barcodes
  tcr = tcr[!duplicated(tcr$barcode),]

  # Subset to only keep barcode and raw clonotype columns
  tcr = tcr[, c("barcode", "raw_clonotype_id")]

  # Rename column to match to the clonotypes file
  names(tcr)[names(tcr) == "raw_clonotype_id"] = "clonotype_id"

  # Extract the TCR chain information from the clonotypes file
  tcr = merge(tcr, clonotypes[, c("clonotype_id", "cdr3s_aa")])

  # Re-order columns, set barcodes as rownames
  tcr = tcr[, c(2, 1, 3)]
  rownames(tcr) <- tcr[, 1]
  tcr[, 1] <- NULL

  # Add the TCR chain and clonotype information as metadata to the data
  clonotype =
    tcr$clonotype_id[match(colnames(sce), rownames(tcr))]
  sce$clonotype = clonotype
  cdr3s_aa = tcr$cdr3s_aa[match(colnames(sce), rownames(tcr))]
  sce$cdr3s_aa = cdr3s_aa

  # filter out cells without a clonotype
  sce = sce[!,is.na(sce$clonotype)]
  return(sce)
}

# Add the information of the TCR chains and the clonotypes to our data
sceRNA$Kidney1_NAT_restim <- addTCRMetaData(
  sce = sceRNA$Kidney1_NAT_restim,
  tcr_filepath = "path/to/ NAT_restim/vdj_t/filtered_contig_annotations.csv",
  clonotypes_filepath = "path/to/ NAT_restim/vdj_t/clonotypes.csv"
)
sceRNA$Kidney1_NAT_restim

sceRNA$Kidney1_Tumor_restim <- addTCRMetaData(
```

METHODS

```
sce = sceRNA$Kidney1_Tumor_restim,
tcr_filepath = "path/to/Tumor_restim/vdj_t/filtered_contig_annotations.csv",
clonotypes_filepath = " path/to/Tumor_restim/vdj_t/clonotypes.csv"
)
sceRNA$Kidney1_Tumor_restim

sceRNA$Kidney1_Tumor <- addTCRMetaData(
  sce = sceRNA$Kidney1_Tumor,
  tcr_filepath = "path/to/Tumor/vdj_t/filtered_contig_annotations.csv",
  clonotypes_filepath = "path/to/Tumor/vdj_t/clonotypes.csv"
)
sceRNA$Kidney1_Tumor
```

Box 21: QC filtering step 2

```
# per-sample filtering for percentage of mitochondrial reads
Kidney1_NAT_restim <- sceRNA$Kidney1_NAT_restim

rowData(Kidney1_NAT_restim)$gene_name <- rownames(Kidney1_NAT_restim)
rowData(Kidney1_NAT_restim)$location <- chr.loc
Kidney1_NAT_restim <- addPerFeatureQC(Kidney1_NAT_restim)

rowData(Kidney1_NAT_restim)

Kidney1_NAT_restim <- addPerCellQC(Kidney1_NAT_restim, subsets = list(Mito =
is.mito))
qcstats <- perCellQCMetrics(Kidney1_NAT_restim, subsets = list(Mito = is.mito))
filtered <- quickPerCellQC(qcstats, percent_subsets = "subsets_Mito_percent")
colSums(as.data.frame(filtered))

table(filtered$low_n_features, filtered$high_subsets_Mito_percent)

# Flag the low quality cells as discard
Kidney1_NAT_restim$discard <- filtered$discard

# Assign the data back to the object
sceRNA$Kidney1_NAT_restim <- Kidney1_NAT_restim

# Filter out cells
sceRNA$Kidney1_NAT_restim <- Kidney1_NAT_restim[, !Kidney1_NAT_restim$discard]
```

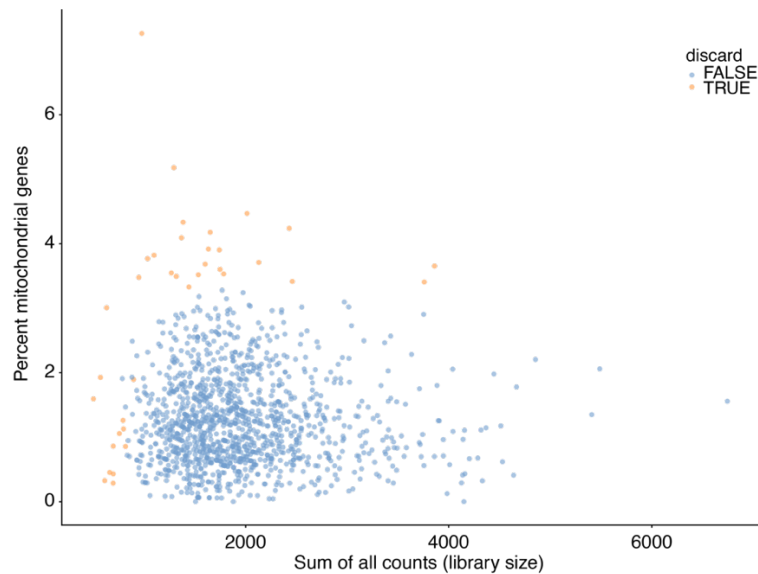


Figure 19: Quality filtering. Cells that will be discarded are marked in orange. Cells which have a small library size and/or a high fraction of mitochondrial reads, which indicates dead or dying cells, are discarded as low-quality cells.

3.6.5.3 Doublet filtering

For the removal of cell doublets, the R package `scDblFinder` (96) was used (**Box 22**). The `scDblFinder` function creates artificial doublets by combining the expression profiles of any two cells in the dataset. A PCA is performed with these artificial doublets and the actual cells in the dataset, and nearest neighbors are identified. Based on the vicinity of a cell to artificial doublets, a doublet score is calculated, which reflects the similarity of each cell to simulated doublets, and which is used for filtering out the top 1% of cells with the highest doublet score.

Box 22: Doublet filtering

```
# doublet filtering on a per-sample basis

# doublet detection
Kidney1_NAT_restim <- sceRNA$Kidney1_NAT_restim
Kidney1_NAT_restim <- scDblFinder(Kidney1_NAT_restim)

# print statistics table
table(Kidney1_NAT_restim$scDblFinder.class)

# assign the object back without cells marked as doublets
sceRNA$Kidney1_NAT_restim <-
Kidney1_NAT_restim[Kidney1_NAT_restim$scDblFinder.class == "singlet"]

# it can also make sense to assign the complete object back and filter out cells
# marked as doublet later, after seeing how they cluster
```

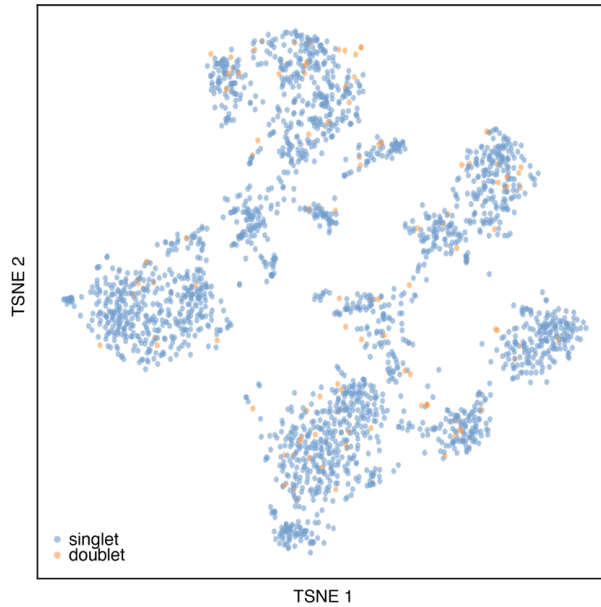


Figure 20: t-SNE plot encoding doublet information. Visualization in reduced dimension manifolds helps to assess the distribution of cells detected as doublets. Cells identified as doublets (marked in orange) are distributed relatively heterogeneously across the t-SNE plot.

3.6.5.4 Dimensionality reduction and clustering

Prior to dimensionality reduction and clustering, expression values were log-scaled using *scran*'s *logNormCounts* function. The *modelGeneVar* function was then used for calculating the variability in gene expression and the 10% of genes showing the highest variability were extracted using the *getTopHVGs* function and were used for downstream analysis. Dimensionality reduction was then performed on the merged SCE containing all samples in the dataset, which was built using the *multiBatchNorm* function of the *batchelor* package (97). For dimensionality reduction, principal component analysis (PCA) on the 10% top highly variable genes was performed using the *runPCA* function from the *BiocSingular* package (91) (**Box 23**). Based on the scree plot, the top 30 PCs were used for clustering. Graph-based clustering was performed using *scran*'s *buildSNNGraph* function and the *cluster_walktrap* function of the *igraph* package (**Box 24**).

Box 23: Dimensionality reduction

```
# log-scaling of the expression values
sceRNA <- lapply(sceRNA, logNormCounts)

# feature selection
# calculate variability in gene expression
all.dec <- lapply(sceRNA, modelGeneVar)
```

METHODS

```
# extract top 10% of variable genes for downstream analysis
all.hvgs <- lapply(all.dec, getTopHVGs, prop = 0.1)

# normalize counts across the samples
rescaled <- multiBatchNorm(sceRNA)

# extract the individual samples
Kidney1_NAT_restim <- rescaled$Kidney1_NAT_restim
Kidney1_Tumor_restim <- rescaled$Kidney1_Tumor_restim
Kidney1_Tumor <- rescaled$Kidney1_Tumor

# combine the selected features
combined.dec <- combineVar(all.dec)
chosen.hvgs <- combined.dec$bio > 0
sum(chosen.hvgs)

# synchronize metadata
rowData(Kidney1_NAT_restim) <-
  rowData(Kidney1_NAT_restim)[, c("gene_name", "location")]
rowData(Kidney1_Tumor_restim) <-
  rowData(Kidney1_Tumor_restim)[, c("gene_name", "location")]
rowData(Kidney1_Tumor) <-
  rowData(Kidney1_Tumor)[, c("gene_name", "location")]

# merge individual objects into one final object
sce_merged <- cbind(
  Kidney1_NAT_restim,
  Kidney1_Tumor_restim,
  Kidney1_Tumor
)

# dimensionality reduction
sce_merged <- runPCA(sce_merged,
  subset_row = chosen.hvgs,
  BSPARAM = BiocSingular::RandomParam()
)
```

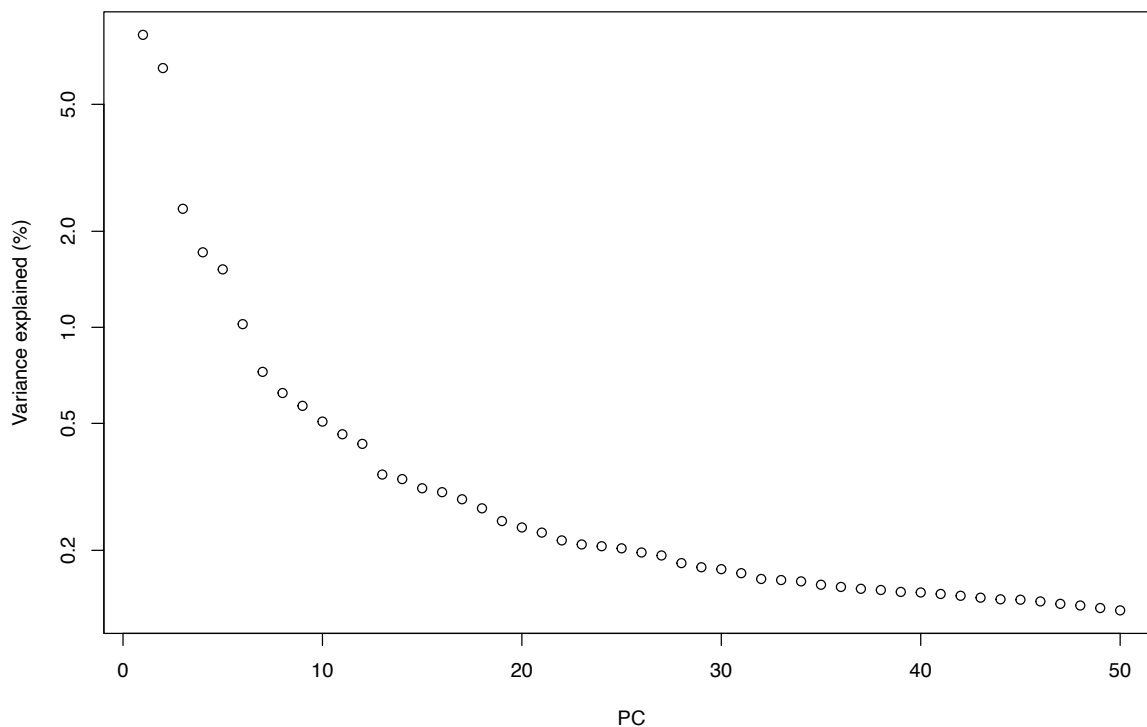


Figure 21: Scree plot showing the variance explained by each principal component (PC). Based on this plot, the number of PCs for clustering is chosen and passed to the “k” argument of the *buildSNNGraph* function (see **Box 24**).

Box 24: Clustering

```
# calculate the clusters
snn.gr <- buildSNNGraph(sce_merged,
  k = 30, # number of PCs to use
  use.dimred = "PCA"
)

clusters <- igraph::cluster_walktrap(snn.gr)$membership

# sample distribution in clusters
tab <- table(Cluster = clusters, Batch = sce_merged$tissue)
tab

# set clusters as colLabels of the sce
colLabels(sce_merged) <- factor(clusters)
```

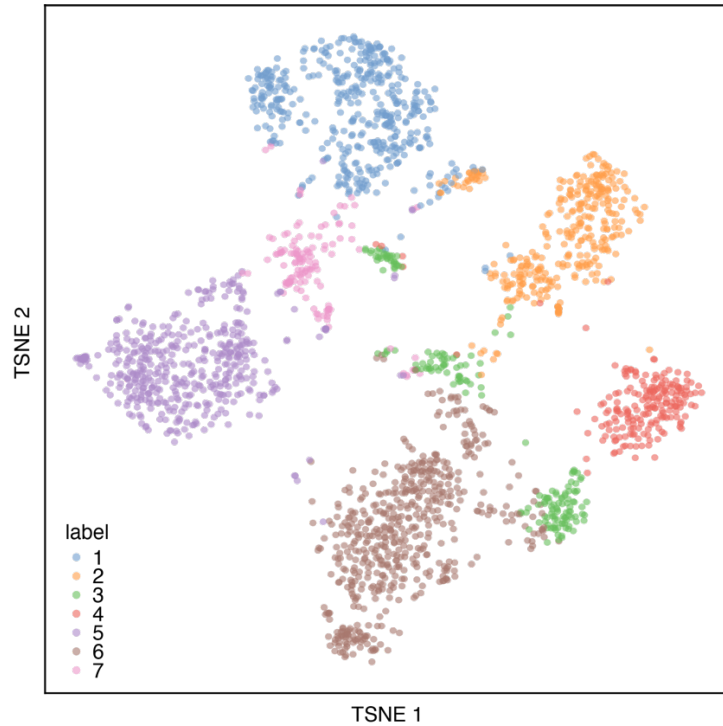


Figure 22: t-SNE plot encoding cluster membership

3.6.6 scRNA/TCR-seq data analysis

scRNA/TCR-seq analysis comprised cell type annotation, marker gene detection, differential analysis, and the analysis of TCR repertoire diversity as well as its overlap between tumor and healthy tissue.

3.6.6.1 Cell type annotation

Cell type annotation was done manually with the use of pre-defined marker genes (**Box 25**). In analogy to scATAC-seq data analysis, the following genes were used for distinguishing cell types: *IL2RA*, *FOXP3*, *ENTPD1*, *IRF4*, *NFKB1*, *NR4A1*, *CTLA4*, *ICOS*, *TNFRSF4*, *TNFRSF9*, *TNFRSF18*, *CD28*, *TCF7*, *IL10*, *GZMB*, *PRDM1*, *IL2* for Treg cells, *BATF*, *CCR8* for CCR8⁺ Treg cells, *RORC*, *IKZF2* for pTreg cells, *STAT1*, *STAT4*, *TBX21*, *IFNG*, *TNF*, *LTA* (Th1), *GATA3*, *IRF4*, *STAT5A*, *STAT6*, *IL4*, *IL5*, *IL9*, *IL10*, *IL13*, *IL21* (Th2), *IRF4*, *RORA*, *RORC*, *STAT3*, *CCL20*, *IL17A*, *IL17F*, *IL22*, *IL26* (Th17) for Tconv cells, *TBX21*, *CX3CR1*, *ZEB2*, *KLF2*, *FGFBP2* for effector CD8⁺ T cells, *TOX*, *ENTPD1*, *HAVCR2*, *CTLA4*, *LAYN*, *PDCD1*, *TNFRSF9* for exhausted CD8⁺ T cells and *PDCD1*, *ENTPD1* for PD1⁺CD39⁻ CD8⁺ T cells. The *scoreMarkers* function of the *scran* package was further used for identifying the top differentially expressed genes for each cluster (**Box 26**).

Box 25: Cell type annotation by gene expression

```
# set up a list of known marker genes for certain cell types
# Treg cells
Treg <- c(
  "CD4", "IL2RA", "FOXP3", "ENTPD1", "IRF4", "NFKB1", "NR4A1", "CTLA4", "ICOS", "TNFRSF4",
  "TNFRSF9", "TNFRSF18", "CD28", "TCF7", "IL10", "GZMB", "PRDM1", "IL2")
CCR8_Treg = c("BATF", "CCR8")
pTreg = c("RORC", "IKZF2")

# overlay expression values on TSNE for visual interpretation, e.g. FOXP3
plotExpression(sce_merged, features = "FOXP3",
  x = "label", colour_by = "label")
plotTSNE(sce_merged, color_by = "FOXP3", order_by = "FOXP3")
```

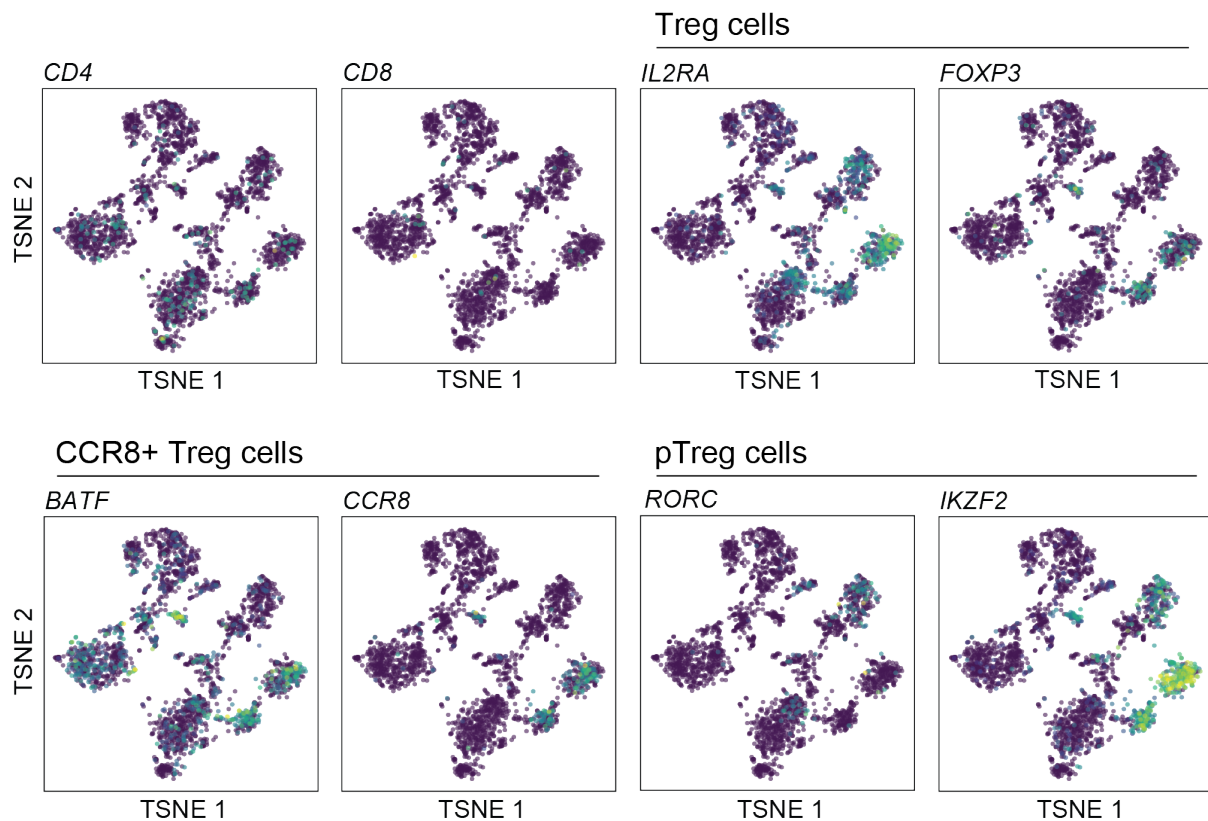


Figure 23: t-SNE plots encoding gene expression. IL2RA and FOXP3 shown as exemplary marker genes for Treg cells, BATF and CCR8 for CCR8⁺ Treg cells, and RORC and IKZF2 for identifying pTreg cells.

Box 26: Identifying top differentially expressed genes for each cluster

```
# score the marker genes between the individual pairs of clusters
markerGenes <- scoreMarkers(sce_merged, colLabels(sce_merged))

# extract marker genes for clusters e.g. cluster4
markerGenes_cluster4 <- as.data.frame(markerGenes[[4]])
markerGenes_cluster4 <-
  markerGenes_cluster4[order(markerGenes_cluster4$mean.logFC.detected,
    decreasing = T), ]
```

METHODS

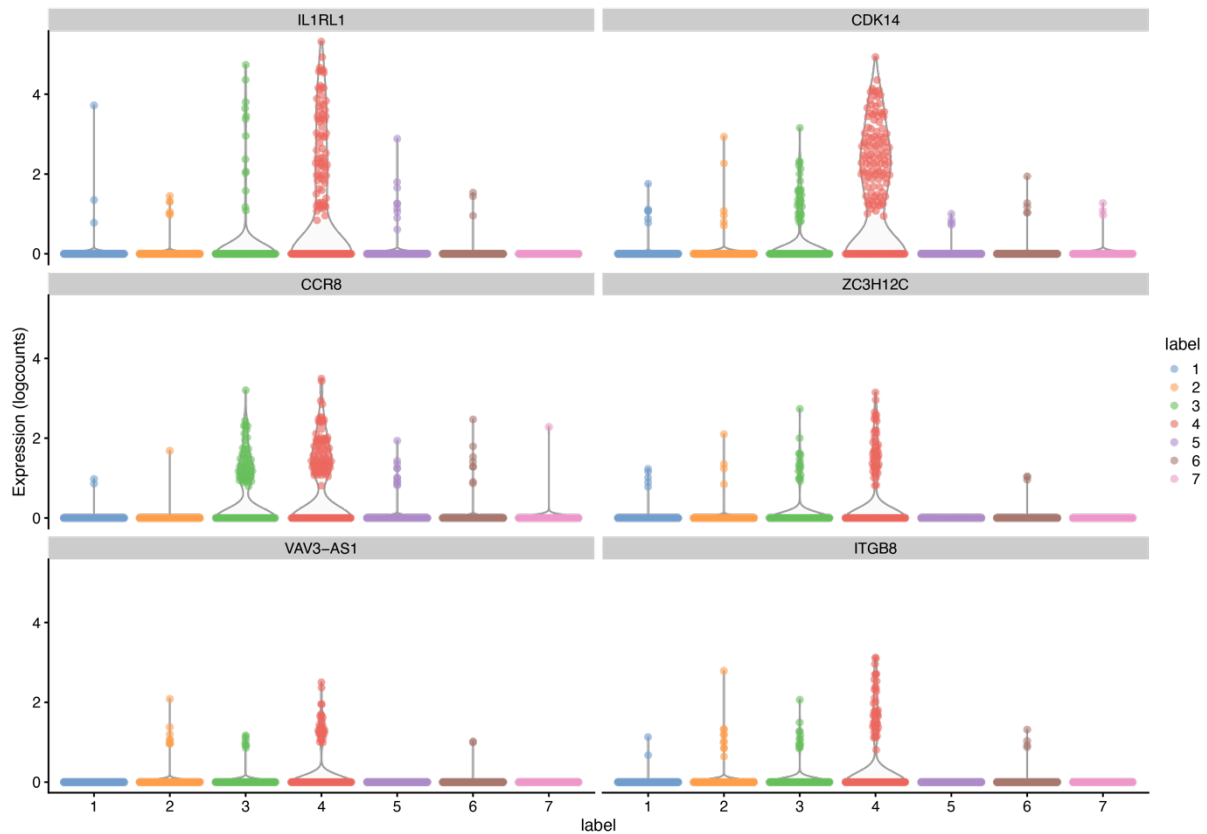


Figure 24: Top differentially expressed genes for cluster 4. Expression (log counts) shown for all clusters.

3.6.6.2 Differential gene expression analysis

Analysis of differentially expressed genes (DEG) was performed using the `FindMarkers()` function from Seurat (92), which uses a Wilcoxon rank sum test (87) with Bonferroni correction for multiple testing, for comparing the specified groups of cells. For this purpose, the SCE object was first converted to a Seurat object (**Box 27**).

Box 27: Differential gene expression analysis between cells from tumor compared to NAT

```
# convert to Seurat object
sce_de <- as.Seurat(sce_de)

# Normalize the data: Feature counts for each cell are divided by the total
counts for that cell and multiplied by the scale.factor
sce_de <- NormalizeData(sce_de)

# Set identity classes to an existing column in metadata
Idents(sce_de) <- "tissue2"

# Find DE features between Kidney1_Tumor and Kidney1_NAT
DE_markers_tu_vs_NAT <- FindMarkers(sce_de, ident.1 = "Kidney1_Tumor", ident.2 =
"Kidney1_NAT")
```

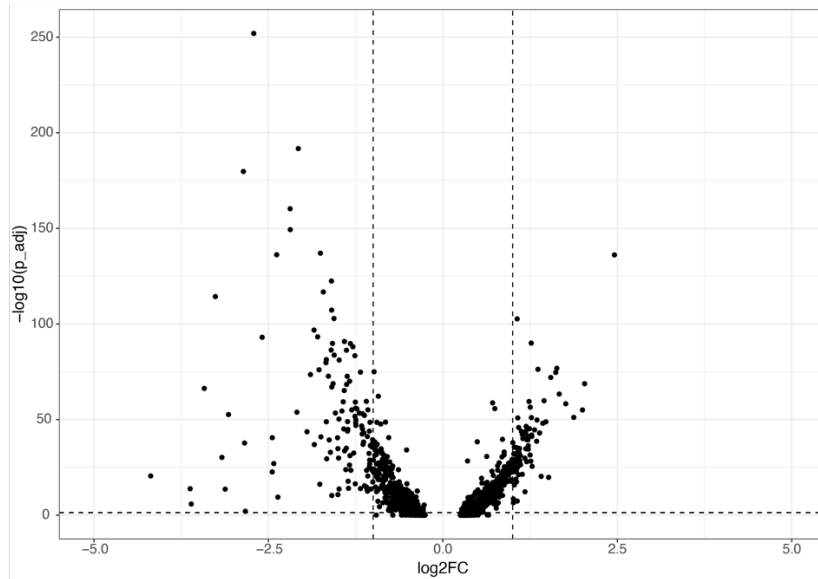


Figure 25: Volcano plot showing DEGs between cells from the tumor and the NAT. Effect size is shown on the x-axis as the log2 fold change, and significance is shown on the y-axis as $-\log_{10}$ of the p-value with Bonferroni correction.

3.6.6.3 TCR clonotype analysis

Clonotype counts derived from TCR information previously added to the SCE were used to analyze clonal expansion within the tissue, as well as TCR repertoire overlap between healthy tissue and tumor (**Boxes 28, 29**).

Box 28: TCR repertoire diversity analysis

```
# function to build dataframes containing TCR clonotype information and frequency
# per cluster
buildDFClonality = function(sce, cluster = NULL){
  if(!is.null(cluster)){
    sce = sce[, colLabels(sce) == cluster]
  }

  n_occur = data.frame(table(sce$cdr3s_aa))
  singletons = 0
  n_occur_sum = as.data.frame(t(data.frame(singleton = c("placeholder", 0))))
  colnames(n_occur_sum) = c("Var1", "Freq")

  for(i in 1:nrow(n_occur)){
    if(n_occur[i, ]$Freq == 1){
      singletons = singletons + 1
    }else{
      n_occur_sum = rbind(n_occur_sum, n_occur[i, ])
    }
  }
  n_occur_sum["singleton", ] = c("Singletons", singletons)
  colnames(n_occur_sum) = c("TCR_chain", "Freq")
  return(n_occur_sum)
}

# using the function defined above, build dataframes containing TCR clonotypes
```


METHODS

```
chain = c(names(chains_duplicated[i]),
          unique_entries,
          paste(chains_frequency[[i]], "(", sum(entries ==
unique_entries[[1]]), "/", sum(entries == unique_entries[[2]]), ")"), sep = ""))
tissue_shared = rbind(tissue_shared, chain)
}
}
}

colnames(tissue_shared) <- c("TCR", "tissue 1", "tissue 2", "Frequency overall
(tissue 1/tissue 2)")

one_tissue <- tissue_shared[tissue_shared$`tissue 2` == "", ]
one_tissue <- one_tissue[, !names(one_tissue) %in% c("tissue 2")]
colnames(one_tissue) <- c("TCR", "tissue", "Frequency overall")
tissue_shared = tissue_shared[!(tissue_shared$TCR %in% one_tissue$TCR), ]

DT::datatable(one_tissue[order(one_tissue$`Frequency overall`, decreasing =
TRUE),])
DT::datatable(tissue_shared[order(tissue_shared$`Frequency overall (tissue
1/tissue 2)`, decreasing = TRUE),])
```

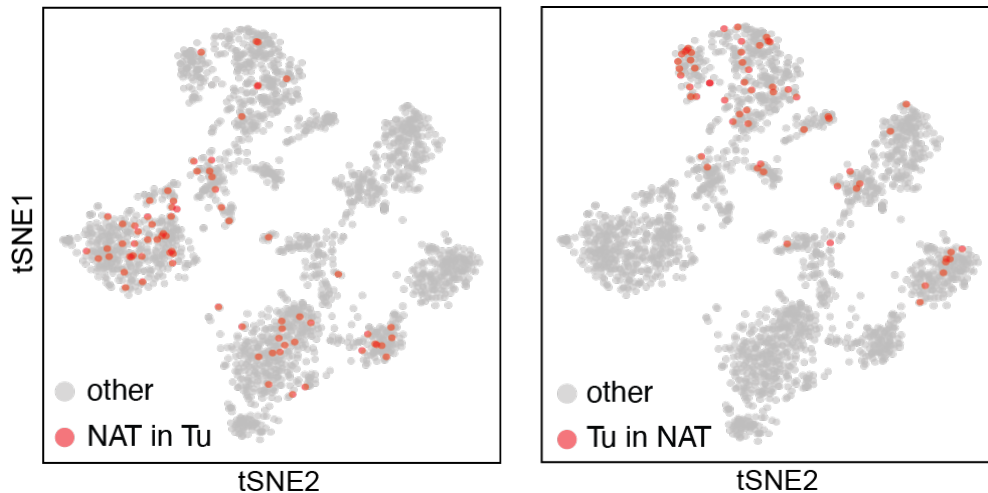


Figure 27: t-SNE plot encoding clonotype information. Left: Tumor cells with TCRs also found in cells from the NAT are marked in red. Right: NAT cells with TCRs also found in tumor cells are marked in red.

4. RESULTS

4.1 CCR8⁺ Treg cells are present in human tumor- and healthy tissue

The vast majority of data on CCR8⁺ Treg cells comes from the murine system, and only recently has this cell type been identified in human skin and VAT tissue (20). We therefore first set out to confirm that CCR8⁺ Treg cells are indeed present in various human tissues, including the liver, lungs, kidney, pancreas, stomach, colon, bladder, and testes (**Figure 28A**). We isolated CD45⁺ immune cells from fresh tumor and NAT from these tissues and analyzed them via FACS. **Figure 28B** shows a representative sample from the kidney, pre-gated on CD4⁺ T cells. In this example, Treg cells (CD127⁻CD25^{hi}) are present both in the tumor and the NAT, as are CCR8⁺ Treg cells (classified as CD45RA⁻CCR8⁺). We analyzed tumor and NAT from kidney (n=21), liver (n=29), colon and rectum (n=17), lungs (n=7), bladder (n=6) and pancreas (n=18), as well as samples from the thyroid gland, esophagus, ovary, breast, oral floor, small intestine, bile duct, stomach, and tongue (n=1-4 per tissue, data not shown). We could find CCR8⁺ Treg cells in each of these tissues, in the tumor as well as the healthy tissue. As shown in **Figure 28C**, the frequency of CCR8⁺ Treg cells of all Treg cells was higher in the tumor compared to the NAT, as was the total cell count of CCR8⁺ Treg cells (data not shown). The mean frequency (with SD) for the kidney was 45% (SD=0.30) (tumor) and 20% (SD=0.21) (NAT), for the liver 49% (SD=0.31) (tumor) and 24% (SD=0.25) (NAT), for colorectal samples 34% (SD=0.25) (tumor) and 14% (SD=0.24) (NAT), for the lungs 57% (SD=0.24) (tumor) and 28% (SD=0.23) (NAT), for the bladder 56% (SD=0.16) (tumor) and 21% (SD=0.11) (NAT) and for pancreas 46% (SD=0.34) (tumor). The Treg cell frequency of CD4⁺ T cells also had a tendency to be higher in the tumor as compared to the NAT, with 13% (SD=0.12) (tumor) and 10% (SD=0.05) (NAT) for the kidney, 28% (SD=0.25) (tumor) and 14% (SD=0.26) (NAT) for the liver, 25% (SD=0.23) (tumor) and 12% (SD=0.23) (NAT) for colon and rectum, 24% (SD=0.15) (tumor) and 6.5% (SD=0.04) (NAT) for the lungs, 31% (SD=0.08) (tumor) and 16% (SD=0.05) (NAT) for the bladder, and 22% (SD=0.15) (tumor) for pancreas. For the pancreas, only cells isolated from the tumor tissue are shown, since CD45⁺ immune cells are not usually present in healthy pancreatic tissue.

RESULTS

To conclude, we identified CCR8⁺ Treg cells in the tumor as well as NAT from a wide variety of tissues, with generally a higher frequency and number of CCR8⁺ Treg cells in the tumor compared to the NAT.

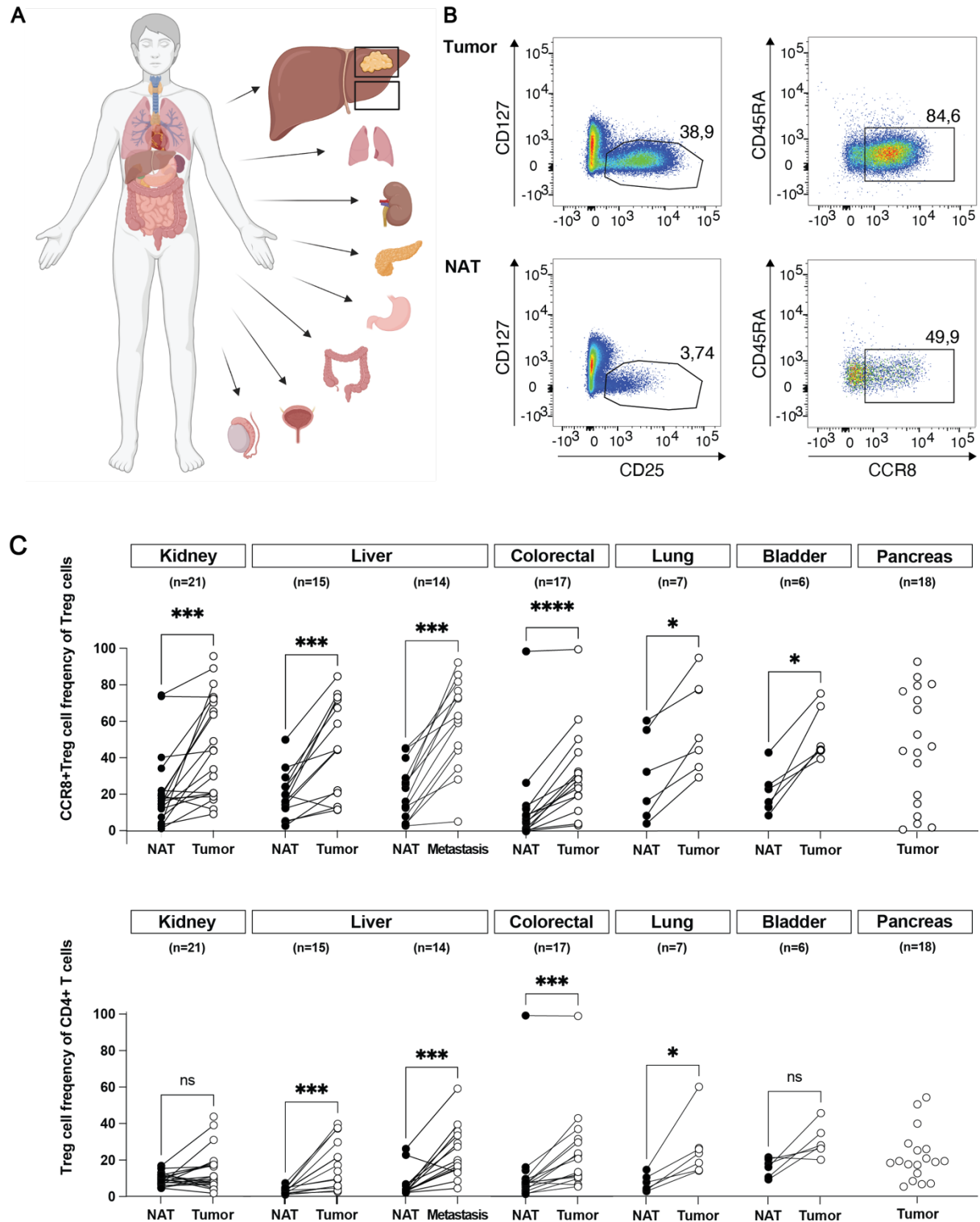


Figure 28: CCR8⁺ Treg cells are present in human tumor- and healthy tissue. **A.** Schematic showing which tissues were analyzed for the dataset presented. Beside tumor tissue, NAT was

analyzed to assess the T cell landscape in tumor-free tissue. **B.** FACS plots of a representative kidney tumor and NAT sample. Pre-gated on CD4⁺ T cells. Treg cells are CD127⁻ CD25^{hi}, CCR8⁺ Treg cells are CD45RA⁻ CCR8⁺. **C.** CCR8⁺ Treg cell frequency of total Treg cells (top) and Treg cell frequency of CD4⁺ T cells (bottom) for tumor and healthy tissue isolated from the kidney (n=21), liver (primary tumor n= 15, metastases n=14), colon and rectum (n=17), lung (n=7), bladder (n=6), and pancreas (n=18). Elements of this figure were created with BioRender.

4.2 CCR8⁺ Treg cells isolated from kidney tumors share molecular characteristics with CCR8⁺ Treg cells isolated from NAT

Since we were able to identify CCR8⁺ Treg cells in human tumors and NAT, we wanted to characterize the molecular programs active in this cell type in both disease states. We therefore performed scATAC-seq of CD3⁺ T cells isolated from tumor and NAT of the kidney, liver, colon, stomach, oral floor, lungs, and testes. Sample pre-processing, including quality filtering, as well as downstream analysis, was performed using ArchR (see methods section for detailed workflow).

For the kidney, a CCR8⁺ Treg cell cluster was identified for all three patients (two patients with CCRC and one patient with AML). Cell type annotation was performed using gene scores of pre-defined marker genes, as demonstrated on the example of the patient CCRC1 in **Figure 29A-C**. Total CD3⁺ T cells were first sub-divided into the two major subsets, CD4⁺ T cells and CD8⁺ T cells with the gene scores for *CD4* and *CD8*. From CD4⁺ T cells, Treg cells were identified as cells with high gene scores for *IL2RA*, *FOXP3*, as well as genes encoding for Treg cell effector molecules such as *ENTPD1*, *CTLA4*, *GZMB* and *IL-10*. For the distinction between tTreg and pTreg, we used *RORC*, an elevated gene score of which is expected in pTreg cells, as well as *IKZF2*, which is expressed by tTreg, but not pTreg cells. CCR8⁺ Treg cells, as a subset of tTreg cells, were further identified using the marker genes *BATF* and *CCR8*. Tconv cells include Th1, Th2 and Th17 cells and were thus identified using marker genes for these subtypes, including *TBX21*, *IFNG*, and *TNF* for Th1 cells, *GATA3*, *IL4*, *IL5*, and *IL13* for Th2 cells, and *RORC*, *IL-17A* and *IL-17F* for Th17 cells. On the CD8⁺ T cell side, we divided cells into CD8 effector cells (elevated gene scores for *TBX21*, *FGFBP2*, *ZEB2* and *KLF2*), exhausted CD8 T cells (elevated gene scores for *TOX*, *CTLA4*, *HAVCR2*, *PDCD1*, *ENTPD1*), and PD1⁺CD39⁻ CD8⁺ T cells (elevated gene scores for *PDCD1*, but not for *ENTPD1*), in which we have previously observed similar

RESULTS

tissue regenerative functions as are described for CCR8⁺ Treg cells (129). In agreement with the data from our FACS analyses, the cell type composition differs between the tumor and the NAT of the same patient, with a tendency for a higher frequency of CCR8⁺ Treg cells in the tumor as compared to the NAT. In the CD8⁺ T cell population, the frequency of exhausted CD8⁺ T (Texh) cells was usually increased in the tumor, and more effector T (Teff) cells were found in the NAT, which was also the case for CCRC1 (**Figure 29B**).

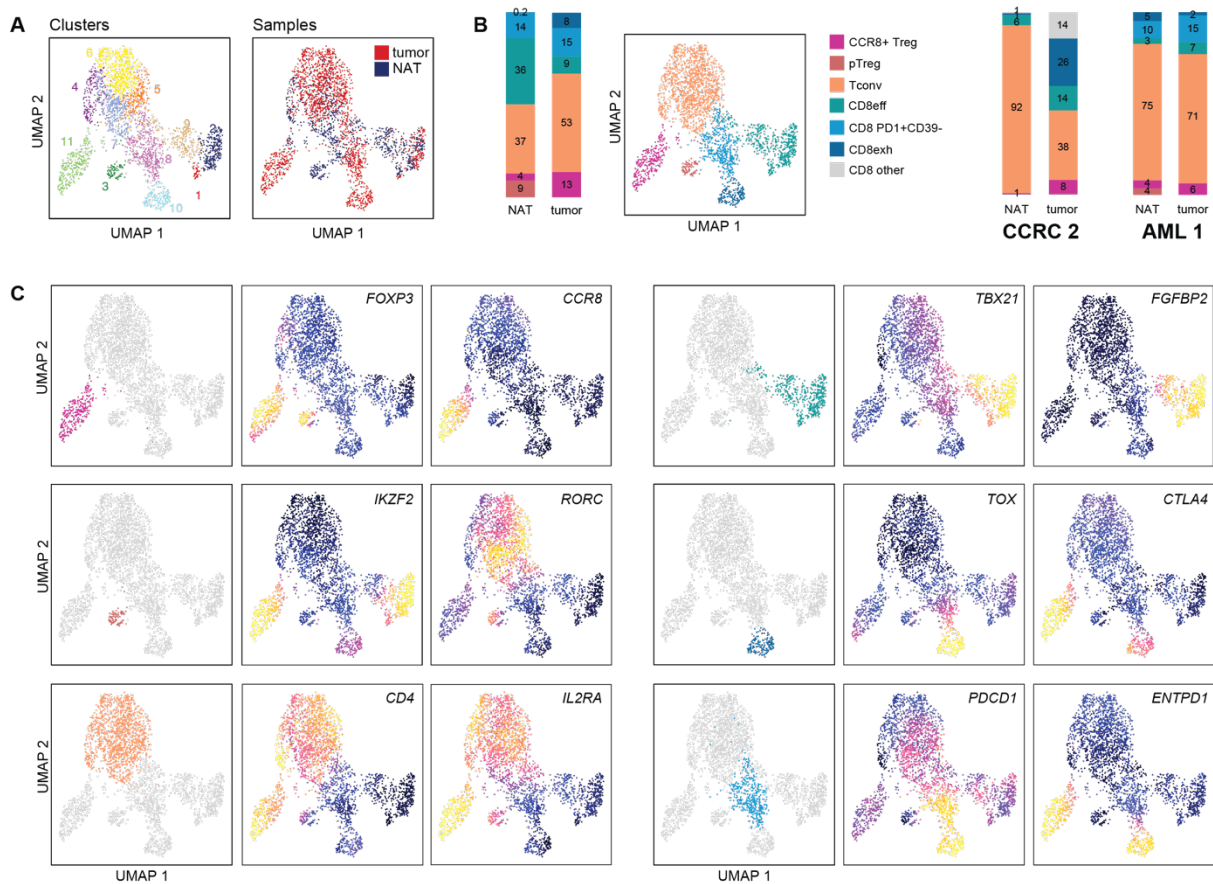


Figure 29: scATAC-seq of kidney tumors (CCRC and AML). **A.** UMAP plot of CCRC1 encoding cluster membership (left) and sample information (right). **B.** Cell type composition of total CD3⁺ T cells isolated from NAT and tumor presented as stacked bar graph and cell type information overlaid on the UMAP embedding (left). Cell type composition of CCRC2 and AML1 presented as stacked bar graphs (right). **C.** Cell type annotation supported by gene scores: per cell type, smoothed gene scores of two exemplary marker genes are overlaid on the UMAP embedding.

The identity of the CCR8⁺ Treg cell cluster was confirmed by calculating the z-scores for the core Treg cell signature and the CCR8⁺ Treg cell signature for each cell (**Figure 30A, B**). The CCR8⁺ Treg cell signature was constructed on CCR8⁺ Treg cells

RESULTS

isolated from tumor-free VAT and skin tissue samples from abdominoplasties compared to CD45RA⁺ Treg cells isolated from the blood of healthy donors. The core Treg cell signature was constructed on CD45RA⁺ Treg cells compared to CD45RA⁺ Tconv cells from human blood of healthy donor. Signatures were extracted from Delacher *et al.* (20). The z-score positively correlates with the similarity of the cell to the signature cell type. **Figure 30A and B** show that the CCR8⁺ Treg cell cluster of CCRC1 as defined via gene scores has high z-scores (indicated by warm colors) for both the core Treg cell signature and the CCR8⁺ Treg cell signature. The same is true for CCRC2 and AML1. We can conclude from this that the “classical” suppressive mechanisms as well as the tissue repair program are active in CCR8⁺ Treg cells. **Figure 30B** further shows the footprints of the transcription factor BATF in the CCR8⁺ Treg cell cluster compared to a “classical” Treg cell cluster or, if absent, a Tconv cell cluster. We can appreciate a stronger BATF footprint in the CCR8⁺ Treg cell cluster as compared to the reference for all three kidney patients, meaning that more binding sites for BATF are accessible, which indicates a stronger dependence of this cell type on the transcription factor BATF.

After confirming the identity of the CCR8⁺ Treg cell cluster, we wanted to investigate the differential regulation of this cell type in the tumor compared to the NAT. We therefore performed differential analyses on the accessible peaks of CCR8⁺ Treg cells isolated from the tumor compared to CCR8⁺ Treg cells isolated from the NAT. Strikingly, we observed no differential peaks out of 103636, 127957, and 101263 features for CCRC1, CCRC2, and AML1, respectively (**Figure 30C**). This indicates that there might not be a “tumor Treg” cell type at all, but that CCR8⁺ Treg cells found in the tumor have the same non-lymphoid tissue phenotype as CCR8⁺ Treg cells found in the NAT. Differential analysis between CCR8⁺ Treg cells and all other CD3⁺ T cells in the dataset resulted in 1009 (0.97%) up- and 1530 (1.48%) downregulated peaks for CCRC1, 67 (0.06%) up- and 491 (0.38%) downregulated peaks for CCRC2, and 1510 (1.49%) up- and 2499 (2.47%) downregulated peaks for AML1, out of 103636, 127957, and 101263 features in total, respectively (**Figure 30C**). TF motifs enriched in differential peaks include BATF, which belongs to the activator protein 1 (AP-1) superfamily of transcription factors, AP-1 TF subunits JUN, JUNB, JDP2, FOS, and FOSB, as well as CNC-bZip family transcription factors BACH1 and BACH2, IRF4, and SMARCC1 (enriched), and runt-related transcription factor (RUNX)1, its co-factor

RESULTS

CBFB, RUNX2, RUNX3, eomesodermin (EOMES), and TBX10 (negatively enriched) (**Figure 30D**). These TF motifs are also amongst the top 25 motifs in differential peaks in patients CCR2 and AML1 (data not shown). **Figure 30E** shows browser tracks for each cell type for CCRC1 at the CCR8 locus. CCR8⁺ Treg cells show higher accessibility at the promoter as well as in the gene body compared to all other cell types in the sample. The same was observed for the remaining kidney patients (data not shown).

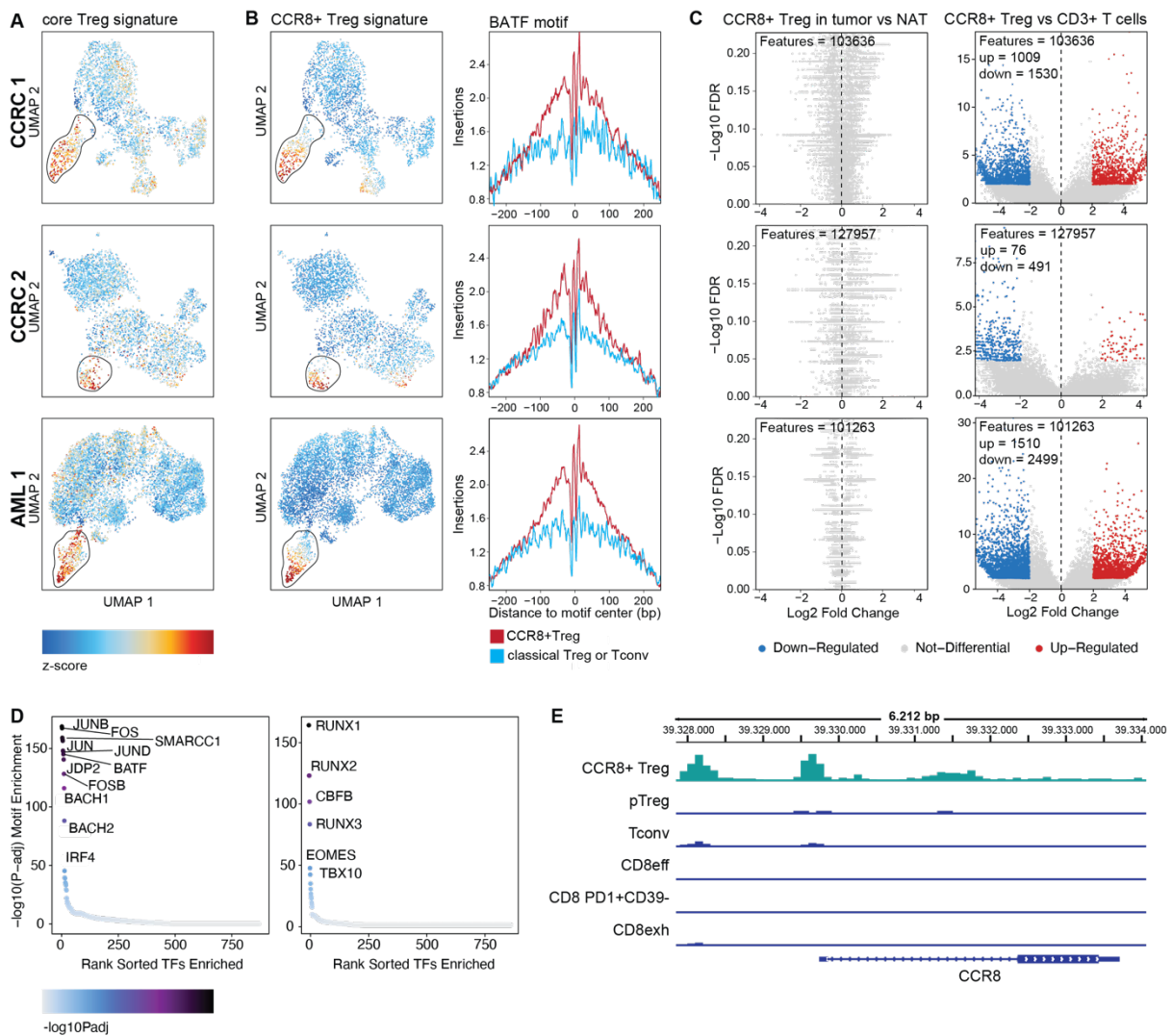


Figure 30: scATAC-seq of kidney tumors (CCRC and AML). **A.** UMAP encoding z-scores for the core Treg signature for CCRC1, CCRC2 and AML. CCR8⁺ Treg cell cluster is circled. **B.** UMAP encoding z-scores for the CCR8⁺ Treg signature for CCRC1, CCRC2 and AML, along with footprints of the transcription factor BATF in the CCR8⁺ Treg cell cluster compared, if present, to another Treg cell cluster, otherwise to a Tconv cell cluster. **C.** Differential analysis of peaks in CCR8⁺ Treg cells from the tumor compared to the NAT (left), and of peaks in CCR8⁺ Treg cells as compared to all other CD3⁺ T cells. **D.** TF motif enrichment in differential peaks between CCR8⁺ Treg cells and all remaining CD3⁺ T

cells for CCRC1. TFs are sorted by rank (significance) along the x-axis, and the y-axis shows the $-\log_{10}$ of the adjusted p-value. Top TFs are annotated in the plot. **E.** IGV browser tracks for CCRC1 by cell type at the CCR8 gene.

4.3 CCR8⁺ Treg cells isolated from liver tumors share molecular characteristics with CCR8⁺ Treg cells isolated from NAT

If “tumor Treg” cells were actually the same cell type as CCR8⁺ Treg cells in normal tissue, this would be a very important finding with implications in “tumor Treg” cell-targeting cancer therapies. We therefore set out to confirm our findings from the kidney in a different entity, hepatocellular carcinoma (HCC).

Comparable to the kidney, we can see that the cell type composition changes drastically between tumor and NAT, with a relative increase CD4⁺ T cells in general, but also in CCR8⁺ Treg cells in the tumor as compared to the NAT in HCC1-3, as well as an increase in exhausted CD8⁺ T cells in HCC1 and HCC2 (**Figure 31B**). CCR8⁺ Treg cells in all three patients show high z-scores for the core Treg cell signature, indicating suppressive capacity, as well as the CCR8⁺ Treg cell signature, indicating tissue remodeling capacity (**Figure 31C, D**). The latter is complemented by a stronger footprint of the transcription factor BATF, which orchestrates the tissue remodeling program, in CCR8⁺ Treg cell clusters compared to classical Treg cell or Tconv cell clusters (**Figure 31D**). Differential analysis of accessible regions in CCR8⁺ Treg cells isolated from the tumor compared to the NAT yielded no differential feature out of 210017 and 175431 features in HCC1 and HCC3, respectively (**Figure 31E**). This is in line with the findings from the kidney, and indicates shared transcriptional programs between CCR8⁺ Treg cells from the tumor and the NAT. For HCC2, only 5 CCR8⁺ Treg cells were found in the NAT, which did not allow us to perform differential analysis for this patient. Comparing CCR8⁺ Treg cells compared to all other CD3⁺ T cells yielded 5942 (2.83%) up- and 4643 (2.21%) downregulated peaks for HCC1, 56 (0.05%) up- and 70 (0.07%) downregulated peaks for HCC2, and 139 (0.08%) up- and 532 (0.30%) downregulated peaks for HCC3, out of 210017, 103258 and 175431 features, respectively (**Figure 31E**). TF motif enrichment analysis in differential peaks between CCR8⁺ Treg cells and all remaining CD3⁺ T cells of patient HCC1 revealed enrichment for the BATF motif, and further the motifs of the transcription factors JUN, JUNB, JUND, JDP2, FOSB, FOSL1, BACH1 and BACH2, SMARCC1, and IRF4.

RESULTS

Negative enrichment was observed for RUNX1, RUNX2, RUNX3 and CBF β transcription factor motifs, amongst others (**Figure 31F**). Similar TF motif enrichment was found for HCC2 and HCC3. It is noteworthy that TF motifs enriched in differential peaks for CCR8⁺ Treg cells in HCC patients were also found as enriched TF motifs in renal tumors. Comparable to the kidney samples, CCR8⁺ Treg cells show higher accessibility in the CCR8 gene region compared to all other cell types in this dataset in all HCC patients, as shown for HCC1 in **Figure 31G**.

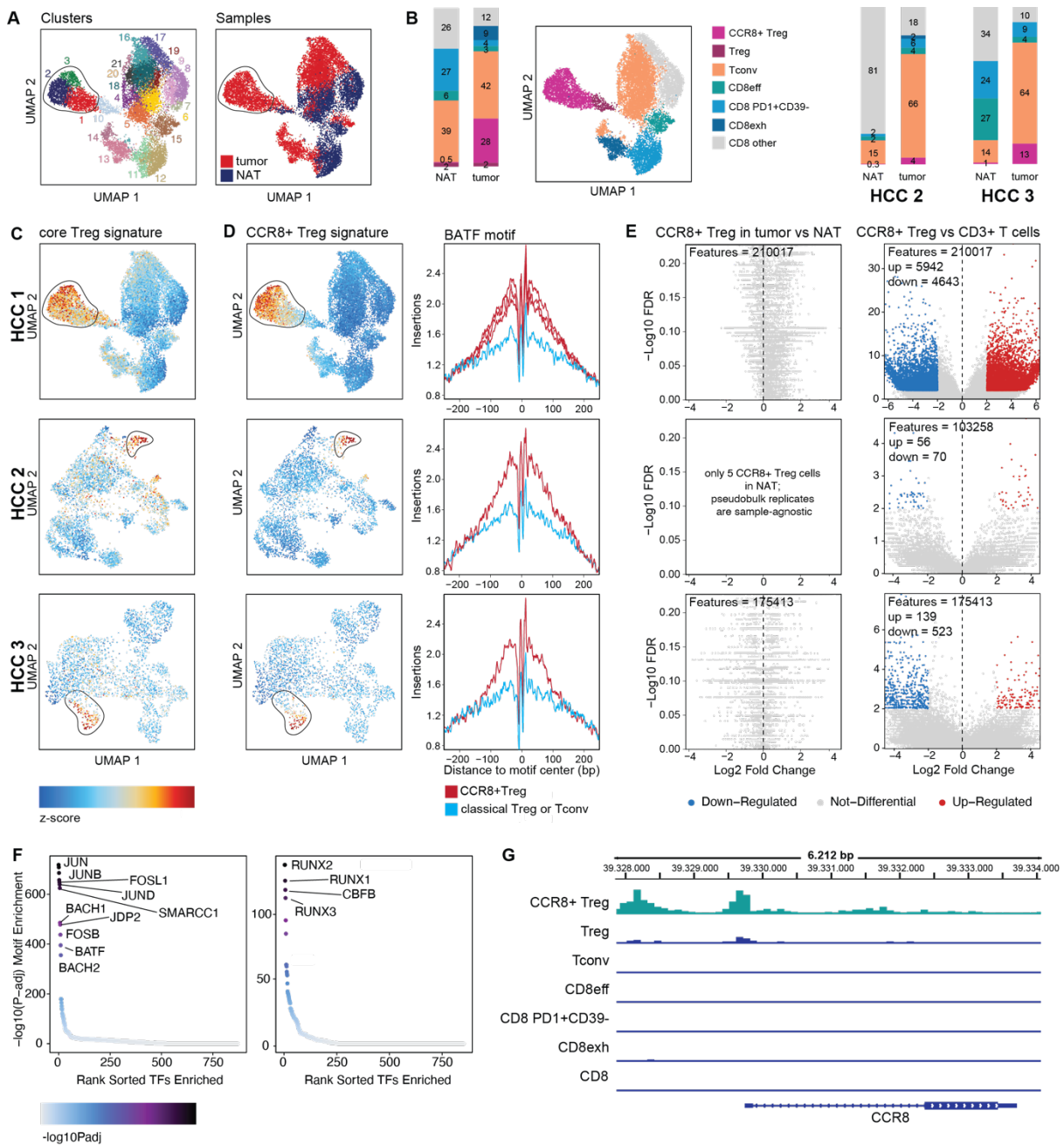


Figure 31: scATAC-seq of liver tumors (HCC). **A.** UMAP plot of HCC1 encoding cluster membership (left) and sample information (right). **B.** CD3⁺ cell type composition of HCC1 presented as a stacked

bar graph and overlaid on the UMAP (left). Cell type composition of CD3⁺ T cells for patients HCC2 and HCC3 are shown as stacked bar graphs on the right. **C.** UMAP plots of HCC1, HCC2, and HCC3 encoding core Treg cell signature z-score. CCR8⁺ Treg cell cluster is circled. **D.** UMAP plots of HCC1, HCC2, and HCC3 encoding CCR8⁺ Treg cell signature z-score (left) and BATF motif footprint of the CCR8⁺ Treg cell cluster(s) compared to another Treg cell cluster, if present, otherwise to a Tconv cluster. **E.** Differential analysis of peaks in CCR8⁺ Treg cells from tumor vs NAT (left), and of peaks in CCR8⁺ Treg cells compared to all other CD3⁺ T cells (right). **F.** TF motif enrichment in differential peaks between CCR8⁺ Treg cells and all remaining CD3⁺ T cells for HCC1. TFs are sorted by rank along the x-axis, the y-axis shows the -log₁₀ of the adjusted p-value. **G.** IGV browser tracks for each CD3⁺ T cell subset identified for patient HCC1 at the CCR8 gene.

4.4 CCR8⁺ Treg cells isolated from a variety of solid tumors share molecular characteristics with CCR8⁺ Treg cells isolated from the corresponding NAT

In order to get a broader picture, we analyzed tumors and NAT from the bladder (UC), the oral floor (OCSCC), lungs (lung AC and ASC), stomach (GAC), and testes (TGCT). In all patients analyzed we could identify a CCR8⁺ Treg cell cluster with high z-scores for the core Treg cell signature and the CCR8⁺ Treg cell signature (**Figure 32A-C**). In line with the high z-scores for the CCR8⁺ Treg cell signature, we also observe a stronger BATF motif footprint in all CCR8⁺ Treg cell clusters compared to a classical Treg cell or Tconv cell cluster (**Figure 32C**). Differential analysis for patients UC1, lung ASC, lung AC and TGCT1 yielded no differential peaks out of 99156, 172412, 77230 and 120377 features in total, respectively. For UC2 and OCSCC there was not a sufficient amount of CCR8⁺ Treg cells in the NAT to do a differential analysis, and for TGCT2 we only processed the tumor, since there was no NAT (**Figure 32A**), and thus could not do this comparison. We do get some differential peaks for CCR8⁺ Treg cells in GAC vs NAT, however the cells from the NAT within the CCR8⁺ Treg cell cluster do not have a high z-score for the CCR8⁺ Treg cell signature (**Figure 32A, C**), and thus calls for sub-clustering rather than reflecting true differences within the CCR8⁺ Treg cell population. Differential peaks in CCR8⁺ Treg cells compared to all other CD3⁺ T cells could be identified for all patients but lung AC and TGCT1 (**Figure 32D**), with 377 (0.38%) up- and 209 (0.21%) downregulated peaks out of 99156 peaks in total for UC1, 29 (0.01%) up- and 94 (0.05%) downregulated peaks out of 203798 peaks in total for UC2, 2995 (1.97%) up- and 3147 (2.07%) downregulated peaks out of 151729 peaks in total for OCSCC, 36 (0.02%) up- and 16 (0.01%) downregulated peaks out of 172412 peaks in total for lung ASC, 0 up- and 3 (0%) downregulated

peaks out of 77230 peaks in total for lung AC, 2697 (2.79%) up- and 3711 (3.84%) downregulated peaks out of 96636 peaks in total for lung GAC, 0 up- and 3 (0%) downregulated peaks out of 120377 peaks in total for lung TGCT1, and 692 (0.80%) up- and 1232 (1.43%) downregulated peaks out of 86301 peaks in total for lung TGCT2 (**Figure 32D**). That for some tissues we only see few or no differential peaks between CCR8⁺ Treg cells and the remaining CD3⁺ T cells might be attributed to a suboptimal tissue digest, since especially for the lungs, testes and bladder tissue we do not seem to be able to differentiate CCR8⁺ Treg cells based on peaks. Also, low activity of the Tn5 transposase during the transposition reaction could lead to an effect like this.

Taken together, we were able to identify CCR8⁺ Treg cells with both the core Treg program and the tissue repair program in all tissue entities analyzed.

RESULTS

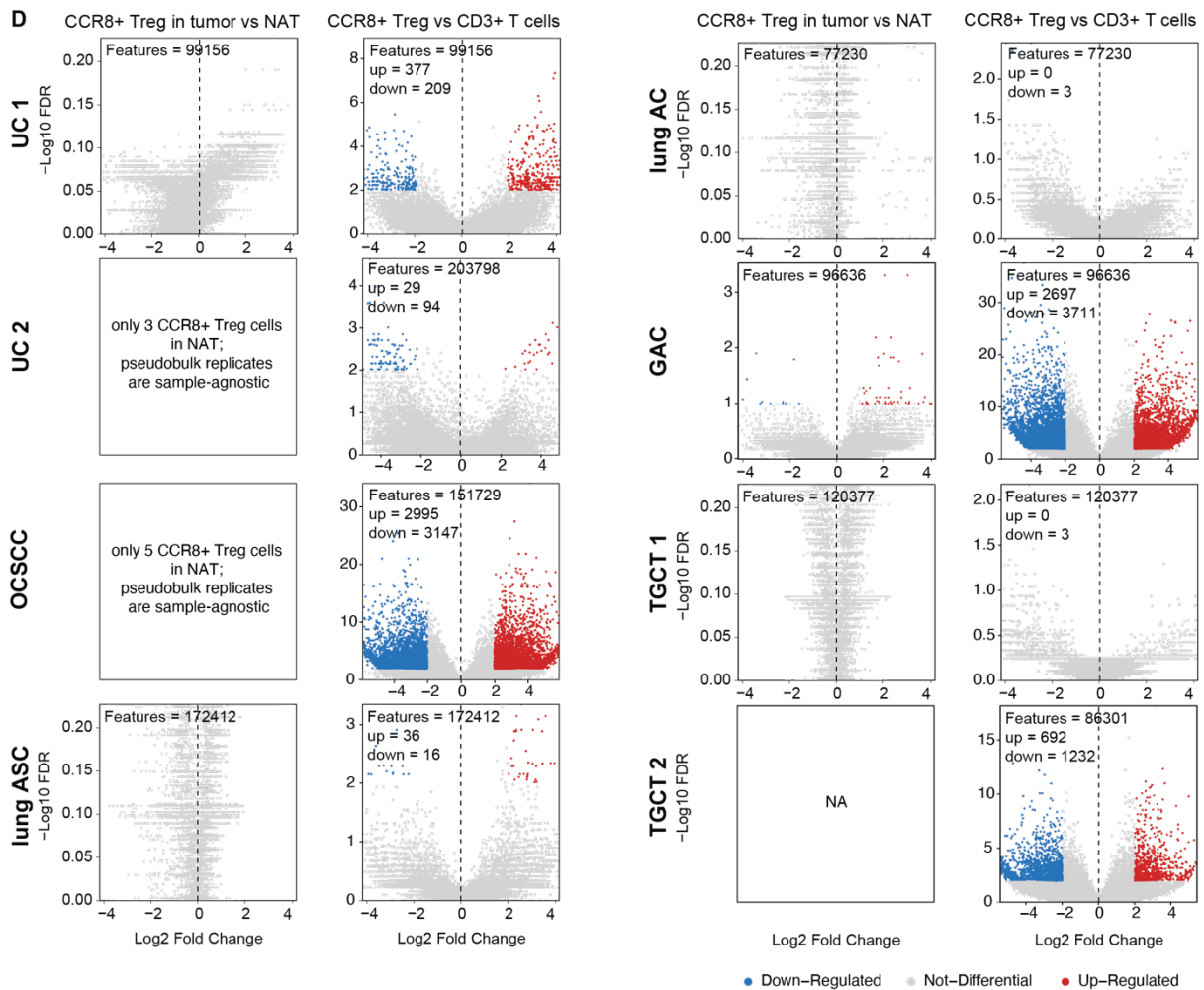


Figure 32: scATAC-seq of tumors of the bladder, oral floor, lungs, stomach and testes. A. UMAP plots encoding cluster membership (left) and sample information (right) for urothelial carcinoma (UC), oral cavity squamous cell carcinoma (OCSCC), lung adeno- and adenosquamous carcinoma (AC, ASC), gastric adenocarcinoma (GAC), and testicular germ cell tumor (TGCT). CCR8⁺ Treg cell clusters are circled. **B.** Z-scores for the core Treg cell signature overlaid on the UMAP embedding. **C.** Z-scores for the CCR8⁺ Treg cell signature overlaid on the UMAP embedding (left), BATF motif footprints for the CCR8⁺ Treg cell clusters compared to “classical” Treg cell or Tconv cell clusters. **D.** Differential peak analysis of CCR8⁺ Treg cells isolated from the tumor compared to CCR8⁺ Treg cells isolated from NAT (left) and of CCR8⁺ Treg cells compared to all other CD3⁺ T cells.

4.5 CCR8⁺ Treg cells isolated from metastases to the liver share molecular characteristics with CCR8⁺ Treg cells isolated from healthy liver tissue

Tissue in which a tumor arises is often inflamed or otherwise not entirely healthy. This is especially the case for HCC, which is preceded by hepatitis (viral or non-viral) or cirrhosis (69), but also CCRC is associated with an inflamed tissue environment (98). It is therefore possible that what we are observing in CCRC and HCC patients is

RESULTS

that CCR8⁺ Treg cells in the tumor are identical to CCR8⁺ Treg cells in inflamed tissue, with regard to regulatory programs. In order to compare the tumor to the corresponding healthy tissue, we analyzed metastases to the liver. Since the malignancy did not arise in the liver but, in the case of our three patients, in the colon, the liver NAT is affected by the metastasis very little: the lobular structure of the liver parenchyma is maintained, there is a neglectable degree of inflammatory infiltrate and no apparent steatosis, as shown in **Figure 33B** for patient MET3. In contrast, the characteristic lobular structure is lost in the liver NAT of our HCC patients. We can further observe portal as well intralobular inflammatory infiltrate and steatosis, as shown for patient HCC1 in **Figure 33A**. The liver NAT from a metastasis patient thus is the closest we can get to a healthy corresponding tissue.

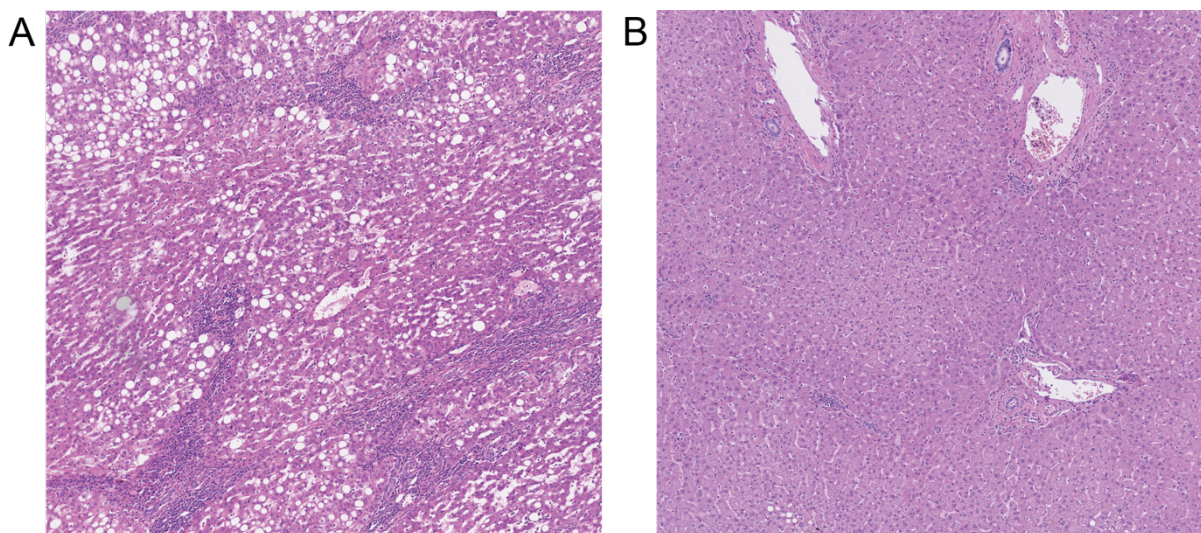


Figure 33: H&E staining of the NAT of an HCC patient and the NAT of a patient with metastasis to the liver. A. NAT of patient HCC1 shows inflammatory infiltrate (portal and intralobular) as well as steatosis. The lobular structure is lost. **B.** NAT of patient MET3 shows lobular structured liver parenchyma with a neglectable degree of inflammatory infiltrate and no steatosis.

Comparable to the kidney and primary liver tumor samples, we also observe a difference in cell type composition between the metastasis and the corresponding liver NAT in the three liver metastasis patients. The total CD4⁺ T cell frequency as well as the frequency of CCR8⁺ Treg cells is increased in the metastasis compared to the NAT, as is the frequency of exhausted CD8⁺ T cells. We can further observe a decrease of CD8⁺ effector T cells in the metastasis compared to the NAT (**Figure 34B**). CCR8⁺ Treg cell clusters from all three patients are characterized by a high z-

RESULTS

score for the core Treg cell signature, which identifies the “classical” suppressive program (**Figure 34C**) and the CCR8⁺ Treg cell signature, which indicates that also the tissue repair program is active (**Figure 34D**). The latter is underlined by a stronger BATF footprint in the CCR8⁺ Treg cells cluster compared to other Treg or Tconv cell clusters (**Figure 34D**). For patient MET1, we were not able to isolate a sufficient number of CCR8⁺ Treg cells from the NAT for performing differential analyses. For MET2, differential analysis yielded no differential peaks out of 99233 features in total, in agreement with data from primary kidney and liver tumors. We identified 5 out of 127493 differential peaks for MET3 (0.004%), however with a very low significance (adjusted p-value) of $-\text{Log}_{10} \text{FDR} < 2$ (**Figure 34E**). Differential analysis of CCR8⁺ Treg cells compared to all other CD3⁺ T cells revealed 1926 (1.79%) up- and 1333 (1.24%) downregulated peaks for MET1, 2543 (2.56%) up- and 3648 (3.68%) downregulated peaks for MET2, and 3403 (2.67%) up- and 1840 (1.44%) downregulated peaks for MET3, out of 107573, 99233, and 127493 features, respectively (**Figure 34E**). Amongst the top TF motifs identified in differential peaks between CCR8⁺ Treg cells and all other CD3⁺ T cells for patient MET1 are BATF, JUN, JUNB, JUND, JDP2, FOS, FOSB, FOSL1, FOSL2, BACH1, BACH2 and SMARCC1 (enriched), and RUNX1, RUNX2, RUNX3, CBFEB, and EOMES, as well as TBX4, TBX5, TBX10, T-box brain transcription factor (TBR)1, MGA, and E74-like factor (ELF)2 (negatively enriched) (**Figure 34F**). These motifs are also amongst the top 25 enriched motifs in peaks of patients MET2 and MET3. In accordance with the kidney and HCC samples, CCR8⁺ Treg cells from MET1-3 show increased accessibility at the CCR8 gene compared to other CD3⁺ T cell subsets (**Figure 34G**).

Taken together, the results from the metastasis patients suggest common molecular characteristics in CCR8⁺ Treg cells from the metastasis compared even to healthy tissue.

RESULTS

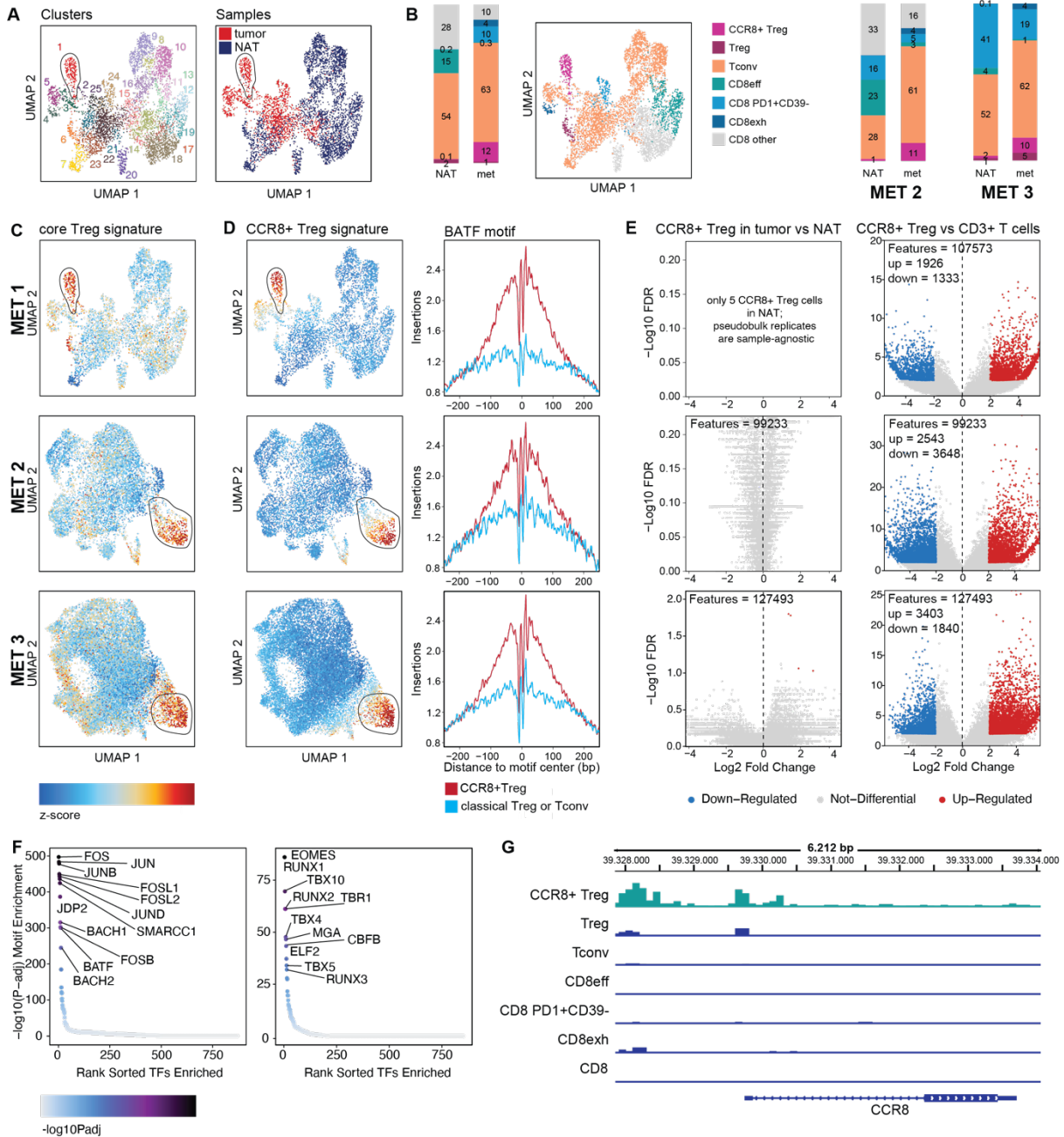


Figure 34: scATAC seq of metastases to the liver. A. UMAP plot of MET1 encoding cluster membership (left) and sample information (right). **B.** Cell type composition of CD3⁺ T cells in MET1 as stacked bar graph and overlaid on the UMAP embedding (left). Stacked bar graphs showing CD3⁺ cell type composition in MET2 and MET3 (right). **C.** Core Treg cell signature z-score overlaid on the UMAP embedding for MET1, MET2, and MET3. CCR8⁺ Treg cell cluster is circled. **D.** CCR8⁺ Treg cell signature z-scores for MET1, MET2, and MET3 overlaid on the corresponding UMAP (left) and BATF motif footprints for the CCR8⁺ Treg cell cluster and a “classical” Treg cell or Tconv cell cluster. **E.** Differential peaks between CCR8⁺ Treg cells from the tumor compared to CCR8⁺ Treg cells from the NAT (left), and between CCR8⁺ Treg cells and all remaining CD3⁺ T cells (right). **F.** TF motif enrichment in differential peaks between CCR8⁺ Treg cells and all remaining CD3⁺ T cells for patient MET1. TFs are sorted by rank along the x-axis, the y-axis shows the -log₁₀ of the adjusted p-value. **G.** IGV browser tracks showing the CCR8 gene for each CD3⁺ T cell subset identified for patient MET1.

4.6 CCR8⁺ Treg cells isolated from CRC and metastasis to the liver in the same patient share molecular characteristics

Now that we have established that CCR8⁺ Treg cells in tumor (primary or metastasis) and the NAT are governed by the same transcriptional programs, we wanted to analyze whether there are any differences in the chromatin accessibility of CCR8⁺ Treg cells between the primary tumor and the corresponding metastasis. We therefore processed a CRC sample with the corresponding colon NAT along with the metastasis to the liver and the corresponding liver NAT, all from the same patient (**Figure 35A**). The cell type composition of total CD3⁺ T cells differs between the tumor and colon NAT, as well as between the liver metastasis and liver NAT, with a higher frequency of CCR8⁺ Treg cells in the malignancy as compared to the NAT. Regarding CD4⁺ T cell- and CCR8⁺ Treg cell frequency, tumor and metastasis are comparable (**Figure 35C**). Two CCR8⁺ Treg cell clusters were identified, one mostly containing cells from the CRC, the other mostly containing cells from the liver metastasis. Both clusters have high z-scores for the core Treg cell signature as well as the CCR8⁺ Treg cell signature, along with a stronger footprint for the transcription factor BATF compared to Tconv cell clusters (**Figure 35D, E**). Out of the 128949 peaks in total, no differential peaks in CCR8⁺ Treg cells from the tumor compared to CCR8⁺ Treg cells from the metastasis were identified. Differential analysis of CCR8⁺ Treg cells compared to the remaining CD3⁺ T cells yielded 3922 up- and 2641 downregulated peaks out of 128949 total features (**Figure 35F**). TF motifs enrichment analysis within these differential peaks identified BATF, JUN, JUNB, JUND, JDP2, FOS, FOSB, FOSL1, FOSL2, BACH1, BACH2, SMARCC1, as well as NFkB1 and RELA as enriched motifs, and RUNX1, RUNX2, RUNX3, CBFb and ELF2 as negatively enriched motifs (**Figure 35G**). Browser tracks across all cell types identified in this patient revealed chromatin accessibility in the CCR8 gene region solely in CCR8⁺ Treg cells (**Figure 35H**).

RESULTS

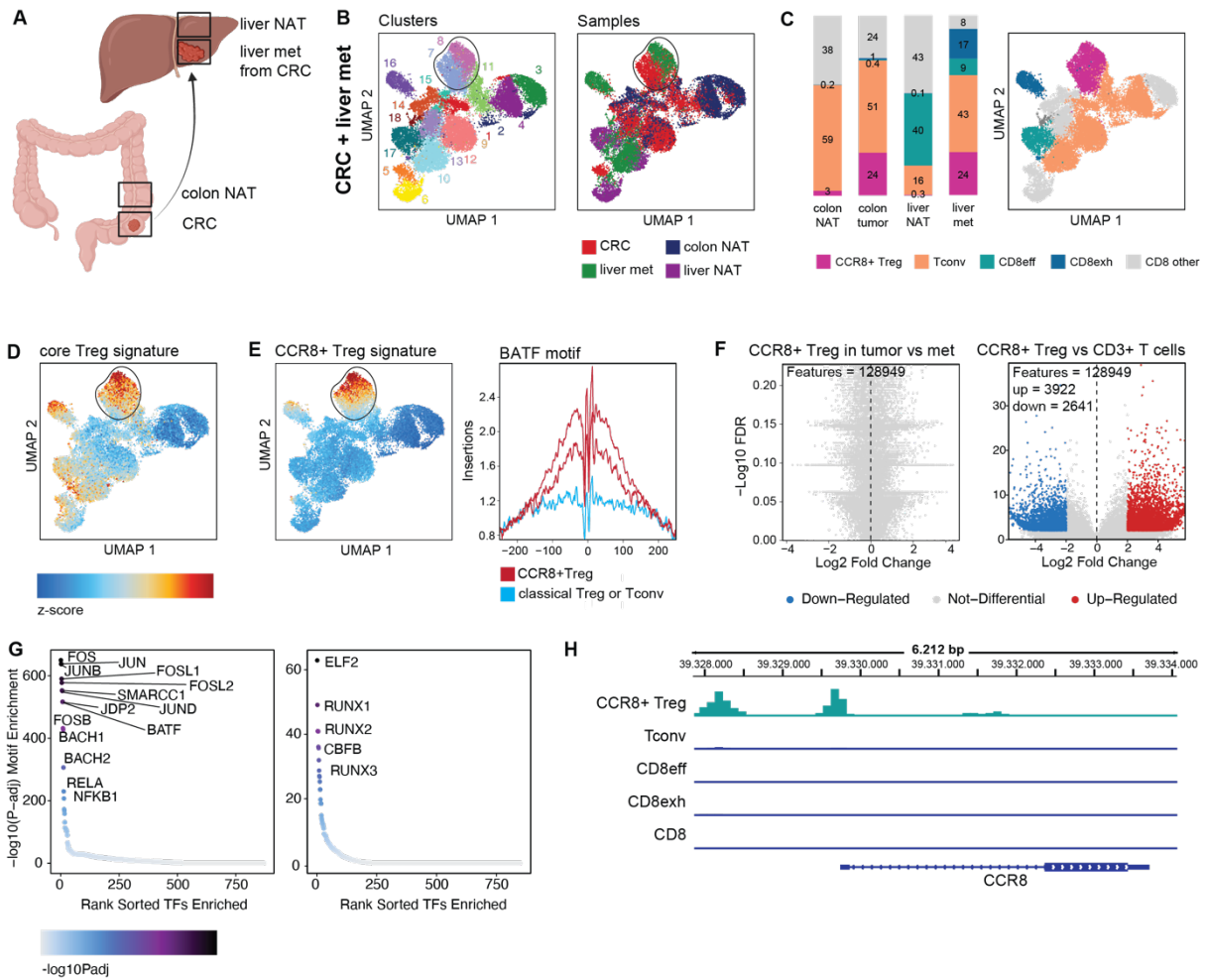


Figure 35: scATAC-seq of CRC with metastasis to the liver. **A.** Scheme of the samples analyzed. The primary tumor is a colorectal carcinoma (CRC), which has metastasized to the liver. Samples analyzed here are the CRC with colon NAT, as well as the metastasis to the liver and liver NAT. **B.** UMAP plot containing all four samples encoding cluster membership (left) and sample information (right). **C.** Cell type composition of colon NAT, colon tumor, liver NAT and liver metastasis presented as stacked bar graphs (left) and overlaid on the UMAP embedding (right). **D.** UMAP encoding core Treg signature z-score. CCR8⁺ Treg cell cluster is circled. **E.** UMAP encoding CCR8⁺ Treg cell signature z-score (left) and BATF motif footprints of the CCR8⁺ Treg clusters compared to a Tconv cluster (right). **F.** Differential analysis of peaks in CCR8⁺ Treg cells in the primary tumor as compared to the metastasis (left) and of peaks in CCR8⁺ Treg cells as compared to all remaining CD3⁺ T cells (right). **G.** TF motif enrichment in differential peaks between CCR8⁺ Treg cells and all remaining CD3⁺ T cells isolated from the patient. TFs are sorted by rank along the x-axis, the y-axis shows the $-\log_{10}$ of the adjusted p-value. **H.** IGV browser tracks for each CD3⁺ T cell subset identified in patient samples. Tracks are shown at the CCR8 gene region. Elements of this figure were created with BioRender.

In conclusion, our scATAC-seq analyses suggest that “tumor Treg” cells are the same cell type as CCR8⁺ Treg cells in the NAT, sharing both the core Treg program

governing suppressive properties and the tissue regenerative program governing tissue remodeling and repair. This, however, does not exclude differences on the gene expression level between CCR8⁺ Treg cells in the tumor and the NAT. We therefore decided to investigate this cell type further using scRNA/TCR-seq.

4.7 CCR8⁺ Treg cells from tumor and NAT share gene expression profiles

In order to investigate whether we find differences between CCR8⁺ Treg cells from the tumor and the NAT on the gene expression level, we processed tumor and NAT of a CCRC patient (CCRC3) and performed scRNA/TCR-seq. We sorted CD4⁺ T cells and enriched for Treg cells as our cell type of interest. Prior to sorting, cells were restimulated in order to investigate the potentially different effector program between CCR8⁺ Treg cells from the tumor compared to the NAT. Bioinformatic analysis was performed using a set of R packages from CRAN (<https://www.r-project.org>) and Bioconductor (91), and is described in detail in the Methods section.

As shown in **Figure 36A**, cells from both the restimulated tumor and the restimulated NAT clustered together, whereas tumor cells that were not restimulated separated in t-SNE space. This was also the case for CCR8⁺ Treg cells, which were identified by the expression of a set of marker genes as clusters 3 (restimulated CCR8⁺ Treg cells from tumor and NAT) and 4 (non-restimulated CCR8⁺ Treg cells) (**Figure 36B**). Differential analysis between restimulated CCR8⁺ Treg cells from the tumor compared to the NAT resulted in no differential features out of 4059 features in total (**Figure 36C**). This indicates that even after restimulation, CCR8⁺ Treg cells in the tumor do not display a deviating gene expression profile to CCR8⁺ Treg cells from the NAT.

We further analyzed the TCR repertoire in this dataset to get some insight into the clonality of CCR8⁺ Treg cells. As shown in **Figure 36D**, clonal expansion was generally rather low with about 85% of CCR8⁺ Treg cells in cluster 3 (CCR8⁺ Treg cells from the non-restimulated tumor sample) and about 75% of CCR8⁺ Treg cells in cluster 4 (CCR8⁺ Treg cells from re-stimulated tumor and NAT samples) having unique TCRs. TCRs were shared not only between CCR8⁺ Treg cells within one condition, but also between CCR8⁺ Treg cells from the tumor and the NAT (**Figure 36E**). Out of

RESULTS

the 235 CCR8⁺ Treg cells from the NAT and 155 CCR8⁺ Treg cells from the tumor, 8 cells from each condition shared the same TCR, i.e. are clonally related.

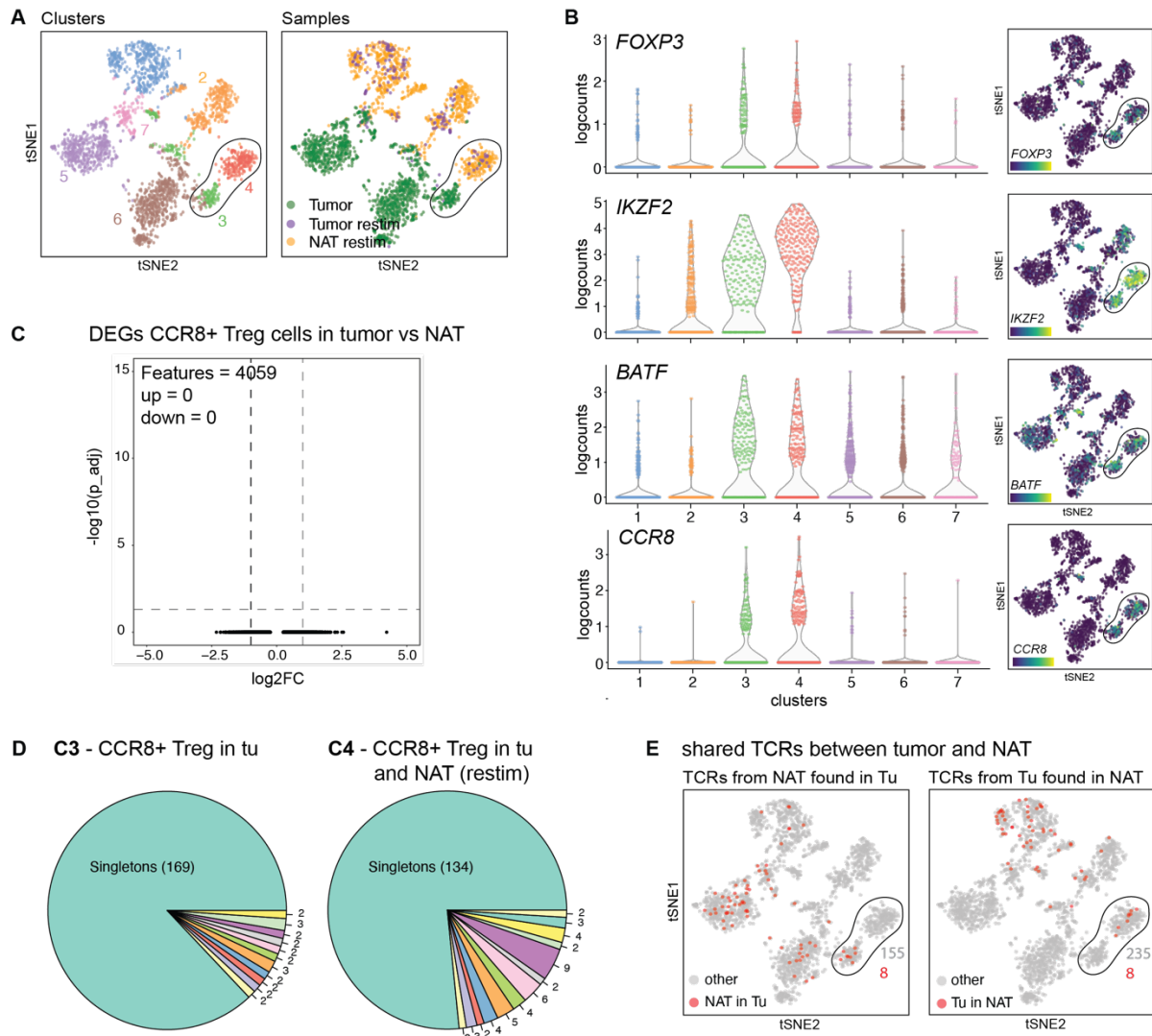


Figure 36: scRNA/TCR-seq of CCRC. **A.** tSNE plots encoding cluster- (left) and sample membership (right). **B.** Gene expression of CCR8⁺ Treg cell marker genes as violin plots (left) and overlaid on tSNE plots (right). **C.** Differential analysis of CCR8⁺ Treg cells from the tumor compared to CCR8⁺ Treg cells from the NAT presented as volcano plot, with the effect size (log₂ fold change) on the x-axis and the significance (-log₁₀ of the adjusted p-value) on the y-axis. **D.** Pie charts showing TCR usage in clusters 3 (CCR8⁺ Treg cells from the non-restimulated tumor sample) and 4 (CCR8⁺ Treg cells from restimulated tumor and NAT samples). Singletons are shown together in one slice, duplicated TCRs are shown in separate slices with the number of their occurrence annotated. **E.** UMAP coloured by TCRs from the NAT that are also found in cells from the tumor (left) and by TCRs from the tumor that are also found in the NAT (right). Numbers in gray denote total cells in the respective sample (tumor on the left, NAT on the right), and numbers in red denote the cells with shared TCR between tumor and NAT.

In conclusion, we have shown that CCR8⁺ Treg cells in the tumor and the NAT do not show any differential chromatin accessibility and are thus governed by the same molecular programs. Additionally, we could not observe any differences in gene expression in this cell type between the two disease conditions, suggesting that “tumor Treg” cells could actually be the same cell type that can be found in healthy non-lymphoid tissues.

5. DISCUSSION

5.1 Treg cells exert a non-canonical function in tissue physiology

Treg cells found in non-lymphoid tissues have been described to promote tissue regeneration and homeostasis, a non-canonical function in addition to their “classical” role in immune-regulation which is largely dependent on the transcription factor FOXP3. This tissue regeneration and remodeling function has been identified in murine as well as human Treg cells isolated from healthy non-lymphoid tissues, and is orchestrated by the transcription factor BATF. The non-lymphoid tissue Treg cell subset has elevated expression of FOXP3, indicating a highly suppressive phenotype, as well increased BATF expression, indicating active tissue regeneration properties. On the surface, both human and mouse FOXP3⁺BATF⁺ tissue Treg cells express, amongst other surface receptors, the chemokine receptor CCR8 (17, 20).

CCR8⁺ Treg cells have been shown to promote tissue homeostasis and -repair in a variety of tissues, data mostly based on murine model systems: These include the skin, where they regulate HFSC differentiation in the steady state (25) and promote tissue repair upon UVB-induced skin damage (29) or mechanical injury (30); the VAT, where they exert functions in metabolic control (22-24); the liver, where they contribute to homeostasis by regulating bile acid synthesis (28); the kidney, where they promote tissue repair after injury (42); the lungs, where they also play a role in tissue regeneration and the reduction of fibrosis (37, 40, 41); skeletal muscle and the heart, where they have regenerative functions in dystrophic muscle (18, 32, 33) and after myocardial infarction (34-36), respectively; and the CNS, where they have been shown to limit brain damage after ischemic stroke (43, 44) and promote oligodendrocyte differentiation and re-myelination in a mouse model of multiple sclerosis (46). All these models indicate an important role of tissue Treg cells for regeneration and homeostasis of tissue, but this also raises the question whether CCR8⁺ Treg cells also have tissue repair or -remodeling functions in the tumor - and how this affects tumor growth and prognosis.

5.2 Treg cell infiltration into the tumor correlates with decreased overall survival

In contrast to their proposed beneficial role in tissue homeostasis and repair, Treg cells are generally considered detrimental in the tumor setting. Meta-analysis of malignancies across a variety of tissues showed that a high level of Treg cell infiltration into the tumor correlates with decreased overall patient survival in most of the cancer types investigated, including melanoma, HCC, and cervical, mammary, renal, and gastric cancers (99). Also in this study, conversely, elevated Treg cell infiltration correlated with better outcomes in colorectal, head and neck, and esophageal cancers, pointing towards the need to better understand Treg cell heterogeneity in different tumors. Specifically in CRC patients, for example, the Treg cell population in the tumor consists of two sub-populations: a suppression-competent FOXP3^{hi} Treg cell subset, and a non-suppressive FOXP3^{lo} Treg cell subset, with opposing effects on patient survival (100). However, besides evaluating Treg cell infiltration alone, the Treg:CD8 effector T cell ratio has proven to be a superior prognostic marker for cancer survival, with the balance tilted toward the CD8 effector T cell side being associated with more favorable outcomes (101).

5.3 The relevance of CCR8⁺ Treg cells across tumor types is unknown

There is only a small body of literature specifically describing CCR8⁺ Treg cells in the context of human cancers, however there is a consensus on their detrimental effect. In a breast cancer study, for example, a high CCR8⁺ Treg cell frequency of total Treg cells was correlated with a poor prognosis (102), as was a high expression of CCR8 in CRC and NSCLC samples, where tumor-infiltrating CCR8⁺ Treg cells were described as highly suppressive (103). In another study on NSCLC, single-cell profiling of Treg cells identified an IRF4-expressing population of tumor-infiltrating, highly suppressive Treg cells, reminiscent of CCR8⁺ Treg cells that also express high levels of IRF4, and the presence of this cell type was further described to correlate with a poor prognosis in various human cancer types (104). Retrospective analysis of samples from more than 250 patients with muscle-invasive bladder cancer (MIBC) explored the prognostic value of CCR8⁺ Treg cell levels on responsiveness to chemotherapy and overall survival. They found that the abundance of CCR8⁺ Treg cell in MIBC correlated inversely with responsiveness and overall survival (56). But not only their elevated suppressive capacity renders CCR8⁺ Treg cells undesirable in the

context of cancer, also the tissue regenerative program seems to promote cancer progression. In human, the factor(s) promoting tissue regeneration remain to be identified, however it is known that in the murine system, the EGFR-ligand Areg plays a central role in tissue repair (18, 19, 28, 44). Areg expression and by extension tissue regeneration have detrimental effects in a variety of tumor models: in a transplantable lung tumor mouse model, Areg deficiency in Treg cells led to delayed tumor progression (105), and Areg produced by Treg cells has further been shown to actively promote tumor cell proliferation, invasion, and the formation of metastases in the lungs, through EGFR signaling (106). Areg has been described to facilitate epithelial-mesenchymal transition (EMT) in a pancreatic cancer model, where Areg silencing led to a decrease in tumor growth and the capability to metastasize (107). In addition to their ability to reprogram the cancer cell transcriptome, leading to the upregulation of angiogenesis and EMT, Areg has been found to induce PD-L1 expression in cancer cells, promoting an immunosuppressive tumor microenvironment (TME) (108). Together, these studies suggest that the tissue regeneration program of CCR8⁺ Treg cells adds to their pro-tumor function in the TME. Areg-mediated tissue repair is not restricted to CCR8⁺ Treg cells, however. Other immune cell types have been found capable of Areg-mediated tissue regeneration. These mainly include cells involved in the type 2 immune response, such as type 2 innate lymphoid cells (ILC2) (109-111), basophils (112, 113), eosinophils (114), but also gdT cells (115).

5.4 Current approaches targeting “tumor Treg” cells

Taking together the pro-tumor effects Treg cells in the TME exert through their suppressive properties as well as their tissue regenerative function, it becomes clear why efforts are put into targeting this cell type for cancer treatment. Several studies have made an effort to better characterize this subset of highly suppressive Treg cells in the tumor, often termed “tumor Treg”, however differential analysis is often performed on Treg cells isolated from the tumor compared to Treg cells from peripheral blood (53-56, 102, 116, 117). Based on these findings, therapies targeting “tumor Treg” are being developed, with some promising results: CCR8-specific monoclonal antibodies (mAb) designed to deplete CCR8⁺ Treg cells induce CD8⁺ effector T cell infiltration into the tumor and lead to pro-inflammatory responses in a variety of tumor mouse models (54, 116, 118). In a mouse model of colorectal carcinoma, CCR8⁺ Treg cell depletion, beside an increase in CD8⁺ effector T cells in

the TME, was further associated with a decrease of exhausted CD8⁺ T cells in the TME, characterized by high expression levels of TOX. In human tumors, high levels of CCR8⁺ Treg cells in the TME also frequently correlate with elevated TOX expression in CD8⁺ T cells, which gives hope for efficacy of this therapy also in these patients (119). Combination of CCR8⁺ Treg cell targeting mAb with anti-PD1 therapy has been shown to potentiate the anti-tumor effect in murine tumor models (55, 116, 118). However, some degree of autoimmune inflammation has been observed upon CCR8⁺ Treg cell depletion, albeit at lower levels than would be induced by complete Treg cell depletion (116). This makes sense, since CCR8⁺ Treg cells are also found in healthy human tissue, where they exert important functions in peripheral tolerance as well as tissue homeostasis and regeneration. Therefore, it remains to be seen whether the non-lymphoid tissue-specific elimination of Treg cells in patients could a) boost anti-tumor immunity, leading to tumor rejection, while b) still preventing systemic autoimmune adverse events while tolerating some level of local autoimmunity in the tissue, and c) whether this level of local autoimmunity is tolerable for the duration of the treatment. Whether CCR8 is an attractive target to boost cancer immunotherapy in different human tumors, and whether this treatment strategy would also compromise neighboring, non-tumor-bearing tissues, is the primary question of this thesis.

5.5 CCR8⁺ Treg cells are present in tumor and normal adjacent tissue

Differential analysis based on abovementioned comparison, namely Treg cells isolated from the tumor compared to lymphoid tissue Treg cells isolated from peripheral blood, does not make it possible to deconvolute tumor vs healthy from lymphoid tissue vs non-lymphoid tissue effects. We therefore performed a detailed characterization of CCR8⁺ Treg cells isolated from human tumors and the corresponding NAT. We found CCR8⁺ Treg cells to be present both in the tumor and the NAT in all tissues investigated, including the liver, lungs, kidney, pancreas, stomach, colon, bladder, testes, thyroid gland, esophagus, ovary, breast, oral floor, small intestine, bile duct, stomach, and tongue. The CCR8⁺ Treg cell frequency of total Treg cells was significantly increased in tumors from kidney, liver (primary and metastases), colon and rectum, lungs, and bladder compared to the NAT. From the other tissues, we did not have enough samples for a solid statistical analysis. The Treg

cell frequency of total CD4⁺ T cells followed the same trend, with a significant increase in the tumors from the liver (primary and metastases), colorectal cancer, and the lungs.

5.6 CCR8⁺ Treg cells from tumor and tumor-free tissue share epigenetic features

In order to assess transcriptional programs active in CCR8⁺ Treg cells isolated from tumor and NAT, we performed scATAC-seq of samples from liver, lungs, kidney, stomach, colon, bladder, and testes. We found that CCR8⁺ Treg cells from both disease states shared the same molecular profiles, including the core Treg suppressive program and the tissue regenerative program. This suggests that “tumor Treg” cells are not a tumor-specific cell type, but that CCR8⁺ Treg cells found in tumor and healthy tissue are the same cell type with the same tissue-regenerative properties on top of the classical suppressive Treg cell program. Differential analysis of chromatin accessibility yielded not a single differential peak, which argues for CCR8⁺ Treg cells in the tumor and healthy tissue being one cell type. To our knowledge, there is no study to date that analyzed CCR8⁺ Treg cells in the tumor compared to CCR8⁺ Treg cells from the corresponding healthy tissue. However, Itahashi *et al.* (53) did a study on human lung cancer, where they identified BATF as a key transcription factor for Treg cells in the TME comparing Treg cells isolated from the tumor to peripheral Treg cells isolated from blood. From our previous studies we know, that BATF also orchestrates the tissue regenerative program in CCR8⁺ Treg cells in healthy tissues (17, 20). CCR8⁺ Treg cells with identical transcriptional programs were also found in inflamed tissue (liver NAT from HCC patients), which agrees with a study describing CCR8⁺ Treg cells in arthritic joint inflammation being transcriptionally and epigenetically similar to tumor-infiltrating Treg cells (120). Further, we could not see any differences in chromatin accessibility of CCR8⁺ Treg cells isolated from the primary tumor compared to the metastasis. Taken together, we did not observe any differences in chromatin accessibility of CCR8⁺ Treg cells isolated from a variety of disease states, including healthy tissue, inflamed tissue, primary tumor and metastasis. Due to the expected technical variability between donors, resulting from different tissue digest protocols and sample processing on different days, patients could not be directly compared amongst each other. However, independent of tissue type and patient, CCR8⁺ Treg cell clusters in all patients showed high z-scores for the core Treg- and tissue repair Treg signature, which would not be the case if significant

differences existed. This also agrees with recent findings from the Liston lab: They systematically characterized murine Treg cell populations across lymphoid and non-lymphoid tissues using scRNA/TCR-seq, and found that Treg cell subsets were shared between many tissues, including the CCR8⁺ Treg cell population being shared (and being transcriptionally comparable) especially between non-lymphoid tissues (121).

5.7 CCR8⁺ Treg cells from tumor and tumor-free tissue share gene expression

We further performed scRNA/TCR-seq of tumor and NAT of a CCRC patient to investigate whether the missing differences in chromatin accessibility translate to missing differences between CCR8⁺ Treg cells from the tumor and the NAT on the transcriptional level. Indeed, we could not identify any differentially expressed genes (DEG) between CCR8⁺ Treg cells isolated from the tumor compared to CCR8⁺ Treg cells isolated from the NAT. In conclusion, it is quite probable that cells described as “tumor Treg” cells are actually the same CCR8⁺ Treg cell type which is also commonly found in healthy human tissues. This would have implications for “tumor Treg” cell targeting therapies, since it would mean that also CCR8⁺ Treg cells in healthy tissues are being depleted along with CCR8⁺ Treg cells in the tumor. And again, this would explain the low levels of autoimmunity induced by CCR8⁺ Treg cell-depleting therapies in first tests in mice (116).

5.8 Understanding CCR8⁺ Treg cell dynamics

In order to develop an informed therapy for targeting CCR8⁺ Treg cells in cancer, it is also important to understand the dynamics of this cell type. Do CCR8⁺ Treg cells reside in the healthy tissue and expand with the tumor? Are they actively recruited by the tumor? Is it a combination of both, or are we dealing with a more dynamic system in which CCR8⁺ Treg cells frequently leave and enter their target tissue or even different non-lymphoid tissues? To take a first step towards answering these questions, we performed scRNA-TCR-seq of CCR8⁺ Treg cells from the tumor and NAT of a CCRC patient and investigated TCR clonotypes. Since every T cell that leaves the thymus has a different TCR, this directly translates to T cells with the same TCR being clonally expanded. In our dataset, we find some shared TCRs between CCR8⁺ Treg cells from the tumor and the NAT. This could hint towards a system where CCR8⁺ Treg cells are either already present in the healthy tissue and then expand with the tumor, or towards a more dynamic system, where CCR8⁺ Treg cells have clonally

expanded at some point during their lifetime and frequently exit and re-enter peripheral tissues, including the NAT as well as the tumor. To conclusively answer this question, one would have to analyze TCRs of CCR8⁺ Treg cells across a multitude of tissues from the same patient. To our knowledge, a study like this has not been published to date, and it would be important to perform these experiments in order to get a better insight into the dynamics of human CCR8⁺ Treg cells. In the murine system, we know that CCR8⁺ Treg cells from different tissues share TCR clones, and parabiosis experiments revealed that CCR8⁺ Treg cells frequently leave and re-enter non-lymphoid tissues in a non-tissue-specific manner, with dwell times of about 3 weeks for CCR8⁺ Treg cells in most non-lymphoid tissues. Together, these findings suggest a very dynamic model of tissue residency of CCR8⁺ Treg cells, with multi-tissue homing in mice (121). Our data from human tumor and NAT would fit well into this model, however the human system needs further investigation.

5.9 Implications for cancer therapy targeting CCR8⁺ Treg cells

CCR8 might be the target of choice when targeting CCR8⁺ Treg cells for tumor therapy, as we found it to be present in the top differentially accessible genes between CCR8⁺ Treg cells and all other CD3⁺ T cells across all tissues investigated. Amongst all CD3⁺ T cells, we can further only detect CCR8 protein expression on CCR8⁺ Treg cells, which indicates that a CCR8-targeting therapy would not affect any other T cell type. Against the prevailing belief that Tregs in the tumor specifically and exclusively express CCR8, we found that CCR8⁺ Treg cells are present in both the tumor and the healthy tissue, where they supposedly exert important functions in peripheral tolerance as well as tissue homeostasis and regeneration. These findings have to be taken into consideration when targeting CCR8 Treg cells for cancer therapy.

An alternative to directly targeting CCR8⁺ Treg cells could be to interfere with their recruitment to the tumor. If the dynamics observed in the murine system, namely a relatively short dwell time in non-lymphoid tissues of about 3 weeks, holds true for the human system, this might be a promising strategy. Known ligands of human CCR8 are CCL1 (122, 123), CCL16 (124), and CCL18 (125). CCL1 has already been described in the context of cancer, where it drives chemotaxis of CCR8⁺ Treg cells into tumors. It was further suggested that CCL1 potentiates Treg cell suppressive function via the induction of CCR8, FOXP3, CD39, Granzyme B and IL10 expression

(122, 123). In a breast cancer study, CCL1 expression in the tumor correlated with the number of infiltrating Treg cells and further with high tumor grade scores (123). The cell type expressing CCL1 are intra-tumoral myeloid cells, as described in a breast cancer study. This is also true for CCL18, another CCR8 ligand (102). CCL18 has further been described to be expressed by B cells, plasma cells, and T cells, as well as on cancer-associated fibroblasts (CAF) and tumor cells themselves (126-128). In our chromatin accessibility data, we see high gene scores for both CCL18 and CCL16 especially in exhausted CD8⁺ T cells. This indicates that there might be a self-enhancing mechanism by which CCR8⁺ Treg cells in the tumor induce exhaustion of CD8⁺ T cells, which in turn recruit more CCR8⁺ Treg cells to the TME. To see whether this holds true, and to better understand the recruitment of CCR8⁺ Treg cells to the tumor and healthy tissue, we are currently performing spatial transcriptomics on the same kidney and liver patient samples we processed for analyzing chromatin accessibility. With this approach we will be able to identify the cells that produce CCR8 ligands via RNA probes, and further shed some light on the interaction of CCR8⁺ Treg cells with other cell types in the TME and healthy tissue. Also, further experiments are needed to identify how CCR8⁺ Treg cells respond to either of those 3 ligands. This will allow a simulation of CCR8⁺ Treg recruitment, proliferation and survival in the TME and open potential avenues on how to best target CCR8⁺ Treg cells for cancer (immuno)therapy.

6. ABSTRACT

Regulatory T (Treg) cells are well known for their immune-regulatory functions, namely restraining excessive inflammation and mediating peripheral tolerance to self-antigens, for which they rely on the transcription factor FOXP3. A highly-activated, effector-like subset of Treg cells can also be found in non-lymphoid tissues. In addition to the immunosuppressive program, this subset further has tissue regenerative properties and is characterized by the expression of the transcription factor BATF and the surface receptor CCR8.

While pivotal for maintaining tissue homeostasis, the tissue-regenerative function of Treg cells could be deleterious in the context of cancer. It is still unclear, whether this tissue-specific CCR8⁺ Treg cell subset can also be found in the tumor microenvironment. In the literature, Treg cells found in tumors, often referred to as “tumor Treg” cells, are described to have a similar phenotype to CCR8⁺ Treg cells, and efforts are put into therapeutically targeting this cell type. One drawback of the majority of these studies, however, is that Treg cells isolated from the tumor are compared to Treg cells isolated from peripheral blood. This does not allow for the deconvolution of tumor-vs-healthy and lymphoid-vs-non-lymphoid differences.

In this study, we performed a molecular characterization with single-cell resolution of human T cells from a variety of healthy tissues, inflamed tissue, and tumors, with the aim of comparing CCR8⁺ Treg cells from the tumor to their non-lymphoid tissue counterparts as well as other cell types in the tissues. We analyzed chromatin accessibility genome-wide (via single-cell ATAC sequencing) to infer transcriptional programs, as well as gene expression profiles, using single-cell RNA sequencing. Our investigation reveals shared molecular programs in human CCR8⁺ Treg cells, irrespective of (non-lymphoid) tissue type or disease state (healthy, inflamed, primary tumor, and metastasis). Moreover, CCR8⁺ Treg cells from tumor did not show significant differences from CCR8⁺ Treg cells from healthy tissue at the transcriptional level. T cell receptor (TCR) analysis of CCR8⁺ Treg cells from the tumor and healthy tissue revealed shared clones between the disease states, identifying a clonal relationship.

ABSTRACT

Taken together, our results argue against a special “tumor Treg” cell subset, but rather suggest a common CCR8⁺ Treg cell type common to the tumor and healthy tissue. This finding calls for caution regarding the development of “tumor Treg” cell-targeting therapies, since the presence of CCR8⁺ Treg cells in healthy tissue, essential for peripheral tolerance and tissue homeostasis, suggests potential unintended consequences of such interventions.

7. ZUSAMMENFASSUNG

Regulatorische T Zellen, kurz Treg, sind bekannt für ihre immunregulatorischen Funktionen, z.B. die Begrenzung von übermäßiger Entzündung oder die Vermittlung peripherer Immuntoleranz gegenüber Selbstantigenen. Diese besonderen Eigenschaften werden durch den Transkriptionsfaktor FOXP3 vermittelt. Ein hochaktivierter, Effektor-ähnlicher Subtyp von Treg-Zellen kann aber auch in nicht-lymphoidem Gewebe gefunden werden. Neben dem immunsuppressiven Programm besitzt dieser Subtyp geweberenerative Eigenschaften und zeichnet sich durch die Expression des Transkriptionsfaktors BATF und des Oberflächenrezeptors CCR8 aus.

Obwohl für die Aufrechterhaltung der Gewebshomöostase entscheidend, könnte die geweberenerative Funktion von Treg-Zellen im Kontext von Krebserkrankungen nachteilig sein. Es ist allerdings immer noch unklar, ob dieser gewebespezifische CCR8⁺ Treg-Zell Subtyp überhaupt in der Tumormikroumgebung zu finden ist. In der Literatur beschriebene Treg-Zellen aus dem Tumor, oft „Tumor Treg“-Zellen genannt, haben einen ähnlichen Phänotyp wie CCR8⁺ Treg-Zellen. Basierend auf dieser Literatur werden aktuell Krebstherapien entwickelt, welche auf „Tumor Treg“-Zellen abzielen. Ein Nachteil dieser Studien ist jedoch, dass Treg-Zellen, die aus dem Tumor isoliert wurden, mit Treg-Zellen aus dem peripheren Blut verglichen werden. Dies lässt keine Rückschlüsse darauf zu, welche der beobachteten Unterschiede von der Variablen „Tumor gegen gesund“ und welche von der Variablen „lymphoid gegen nicht-lymphoid“ stammen.

In dieser Arbeit haben wir eine molekulare Charakterisierung mit Einzelzellauflösung von humanen T-Zellen aus verschiedenen gesunden Geweben, entzündetem Gewebe und Tumoren durchgeführt, mit dem Ziel, CCR8⁺ Treg-Zellen aus dem Tumor mit CCR8⁺ Treg-Zellen aus nicht-lymphoidem Gewebe sowie mit anderen Zelltypen im Gewebe zu vergleichen. Wir haben die genomweite Chromatinverfügbarkeit mithilfe von Einzelzell-ATAC-Sequenzierung analysiert, um transkriptionelle Programme abzuleiten. Außerdem haben wir Genexpressionsprofile anhand von Einzelzell-RNA-Sequenzierung erstellt. Unsere Untersuchung zeigt identische molekulare Programme in menschlichen CCR8⁺ Treg-Zellen, unabhängig vom (nicht-lymphoidem) Gewebetyp oder Krankheitszustand (gesund, entzündet,

primärer Tumor und Metastase). Darüber hinaus zeigten CCR8⁺ Treg-Zellen aus dem Tumor keine signifikanten Unterschiede zu CCR8⁺ Treg-Zellen aus gesundem Gewebe auf transkriptioneller Ebene. Die T-Zell-Rezeptor-Analyse von CCR8⁺ Treg-Zellen aus Tumor und gesundem Gewebe identifizierte außerdem gemeinsame Klone zwischen gesunden und erkrankten Geweben, was Rückschlüsse auf eine klonale Beziehung zwischen Tumor-residenten und nicht-Tumor CCR8⁺ Treg zulässt.

Zusammenfassend sprechen unsere Ergebnisse gegen einen speziellen "Tumor-Treg"-Zell Subtyp, sondern legen eher einen gemeinsamen CCR8⁺ Treg-Zelltyp nahe, der sowohl im Tumor als auch im gesunden Gewebe vorkommt. Diese Erkenntnis muss bei der Entwicklung von Immuntherapien bedacht werden, die gezielt auf "Tumor Treg"-Zellen abzielen. Diese neuen Therapiemodelle könnten aller Wahrscheinlichkeit nach CCR8⁺ Treg nicht nur im Tumor, sondern auch in gesunden Geweben vernichten. Da dieser Zelltyp aber für periphere Toleranz und die Gewebshomöostase wichtig ist, könnte dessen systemische Eliminierung unvorhergesehene Nebenwirkungen in gesunden Geweben hervorrufen.

8. APPENDIX

8.1 References

1. Braband KL, Kaufmann T, Floess S, Zou M, Huehn J, Delacher M. Stepwise acquisition of unique epigenetic signatures during differentiation of tissue Treg cells. *Front Immunol.* 2022;13:1082055.
2. Braband KL, Nedwed AS, Helbich SS, Simon M, Beumer N, Brors B, et al. Using single-cell chromatin accessibility sequencing to characterize CD4+ T cells from murine tissues. *Front Immunol.* 2023;14:1232511.
3. Nedwed AS, Helbich SS, Braband KL, Volkmar M, Delacher M, Marini F. Using combined single-cell gene expression, TCR sequencing and cell surface protein barcoding to characterize and track CD4+ T cell clones from murine tissues. *Front Immunol.* 2023;14:1241283.
4. Janeway CA, Jr., Medzhitov R. Innate immune recognition. *Annu Rev Immunol.* 2002;20:197-216.
5. Bonilla FA, Oettgen HC. Adaptive immunity. *J Allergy Clin Immunol.* 2010;125(2 Suppl 2):S33-40.
6. Sakaguchi S, Yamaguchi T, Nomura T, Ono M. Regulatory T cells and immune tolerance. *Cell.* 2008;133(5):775-87.
7. Morikawa H, Sakaguchi S. Genetic and epigenetic basis of Treg cell development and function: from a FoxP3-centered view to an epigenome-defined view of natural Treg cells. *Immunol Rev.* 2014;259(1):192-205.
8. Fontenot JD, Gavin MA, Rudensky AY. Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nat Immunol.* 2003;4(4):330-6.
9. Hori S, Nomura T, Sakaguchi S. Control of regulatory T cell development by the transcription factor Foxp3. *Science.* 2003;299(5609):1057-61.
10. Brunkow ME, Jeffery EW, Hjerrild KA, Paepfer B, Clark LB, Yasayko SA, et al. Disruption of a new forkhead/winged-helix protein, scurf, results in the fatal lymphoproliferative disorder of the scurfy mouse. *Nat Genet.* 2001;27(1):68-73.
11. Park JH, Lee KH, Jeon B, Ochs HD, Lee JS, Gee HY, et al. Immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome: A systematic review. *Autoimmun Rev.* 2020;19(6):102526.
12. Wildin RS, Ramsdell F, Peake J, Faravelli F, Casanova JL, Buist N, et al. X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. *Nat Genet.* 2001;27(1):18-20.
13. Bennett CL, Christie J, Ramsdell F, Brunkow ME, Ferguson PJ, Whitesell L, et al. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nat Genet.* 2001;27(1):20-1.

14. Ohkura N, Hamaguchi M, Morikawa H, Sugimura K, Tanaka A, Ito Y, et al. T cell receptor stimulation-induced epigenetic changes and Foxp3 expression are independent and complementary events required for Treg cell development. *Immunity*. 2012;37(5):785-99.
15. Schmidt A, Oberle N, Krammer PH. Molecular mechanisms of treg-mediated T cell suppression. *Front Immunol*. 2012;3:51.
16. Delacher M, Imbusch CD, Weichenhan D, Breiling A, Hotz-Wagenblatt A, Trager U, et al. Genome-wide DNA-methylation landscape defines specialization of regulatory T cells in tissues. *Nat Immunol*. 2017;18(10):1160-72.
17. Delacher M, Imbusch CD, Hotz-Wagenblatt A, Mallm JP, Bauer K, Simon M, et al. Precursors for Nonlymphoid-Tissue Treg Cells Reside in Secondary Lymphoid Organs and Are Programmed by the Transcription Factor BATF. *Immunity*. 2020;52(2):295-312 e11.
18. Burzyn D, Kuswanto W, Kolodin D, Shadrach JL, Cerletti M, Jang Y, et al. A special population of regulatory T cells potentiates muscle repair. *Cell*. 2013;155(6):1282-95.
19. Arpaia N, Green JA, Moltedo B, Arvey A, Hemmers S, Yuan S, et al. A Distinct Function of Regulatory T Cells in Tissue Protection. *Cell*. 2015;162(5):1078-89.
20. Delacher M, Simon M, Sanderink L, Hotz-Wagenblatt A, Wuttke M, Schambeck K, et al. Single-cell chromatin accessibility landscape identifies tissue repair program in human regulatory T cells. *Immunity*. 2021;54(4):702-20 e17.
21. Delacher M, Schmidl C, Herzig Y, Breloer M, Hartmann W, Brunk F, et al. Rbpj expression in regulatory T cells is critical for restraining TH2 responses. *Nat Commun*. 2019;10(1):1621.
22. Cipolletta D, Feuerer M, Li A, Kamei N, Lee J, Shoelson SE, et al. PPAR-gamma is a major driver of the accumulation and phenotype of adipose tissue Treg cells. *Nature*. 2012;486(7404):549-53.
23. Feuerer M, Herrero L, Cipolletta D, Naaz A, Wong J, Nayer A, et al. Lean, but not obese, fat is enriched for a unique population of regulatory T cells that affect metabolic parameters. *Nat Med*. 2009;15(8):930-9.
24. Vasanthakumar A, Moro K, Xin A, Liao Y, Gloury R, Kawamoto S, et al. The transcriptional regulators IRF4, BATF and IL-33 orchestrate development and maintenance of adipose tissue-resident regulatory T cells. *Nat Immunol*. 2015;16(3):276-85.
25. Ali N, Zirak B, Rodriguez RS, Pauli ML, Truong HA, Lai K, et al. Regulatory T Cells in Skin Facilitate Epithelial Stem Cell Differentiation. *Cell*. 2017;169(6):1119-29 e11.

26. Petukhova L, Duvic M, Hordinsky M, Norris D, Price V, Shimomura Y, et al. Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. *Nature*. 2010;466(7302):113-7.
27. Shin BS, Furuhashi T, Nakamura M, Torii K, Morita A. Impaired inhibitory function of circulating CD4+CD25+ regulatory T cells in alopecia areata. *J Dermatol Sci*. 2013;70(2):141-3.
28. Santamaria E, Rodriguez-Ortigosa CM, Uriarte I, Latasa MU, Urtasun R, Alvarez-Sola G, et al. The Epidermal Growth Factor Receptor Ligand Amphiregulin Protects From Cholestatic Liver Injury and Regulates Bile Acids Synthesis. *Hepatology*. 2019;69(4):1632-47.
29. Shime H, Odanaka M, Tsuiji M, Matoba T, Imai M, Yasumizu Y, et al. Proenkephalin(+) regulatory T cells expanded by ultraviolet B exposure maintain skin homeostasis with a healing function. *Proc Natl Acad Sci U S A*. 2020;117(34):20696-705.
30. Nosbaum A, Prevel N, Truong HA, Mehta P, Ettinger M, Scharschmidt TC, et al. Cutting Edge: Regulatory T Cells Facilitate Cutaneous Wound Healing. *J Immunol*. 2016;196(5):2010-4.
31. Mathur AN, Zirak B, Boothby IC, Tan M, Cohen JN, Mauro TM, et al. Treg-Cell Control of a CXCL5-IL-17 Inflammatory Axis Promotes Hair-Follicle-Stem-Cell Differentiation During Skin-Barrier Repair. *Immunity*. 2019;50(3):655-67 e4.
32. Kuswanto W, Burzyn D, Panduro M, Wang KK, Jang YC, Wagers AJ, et al. Poor Repair of Skeletal Muscle in Aging Mice Reflects a Defect in Local, Interleukin-33-Dependent Accumulation of Regulatory T Cells. *Immunity*. 2016;44(2):355-67.
33. Castiglioni A, Corna G, Rigamonti E, Basso V, Vezzoli M, Monno A, et al. FOXP3+ T Cells Recruited to Sites of Sterile Skeletal Muscle Injury Regulate the Fate of Satellite Cells and Guide Effective Tissue Regeneration. *PLoS One*. 2015;10(6):e0128094.
34. Xia N, Lu Y, Gu M, Li N, Liu M, Jiao J, et al. A Unique Population of Regulatory T Cells in Heart Potentiates Cardiac Protection From Myocardial Infarction. *Circulation*. 2020;142(20):1956-73.
35. Zacchigna S, Martinelli V, Moimas S, Colliva A, Anzini M, Nordio A, et al. Paracrine effect of regulatory T cells promotes cardiomyocyte proliferation during pregnancy and after myocardial infarction. *Nat Commun*. 2018;9(1):2432.
36. Weirather J, Hofmann UD, Beyersdorf N, Ramos GC, Vogel B, Frey A, et al. Foxp3+ CD4+ T cells improve healing after myocardial infarction by modulating monocyte/macrophage differentiation. *Circ Res*. 2014;115(1):55-67.
37. Tan W, Zhang B, Liu X, Zhang C, Liu J, Miao Q. Interleukin-33-Dependent Accumulation of Regulatory T Cells Mediates Pulmonary Epithelial

- Regeneration During Acute Respiratory Distress Syndrome. *Front Immunol.* 2021;12:653803.
38. Dial CF, Tune MK, Doerschuk CM, Mock JR. Foxp3(+) Regulatory T Cell Expression of Keratinocyte Growth Factor Enhances Lung Epithelial Proliferation. *Am J Respir Cell Mol Biol.* 2017;57(2):162-73.
 39. Mock JR, Garibaldi BT, Aggarwal NR, Jenkins J, Limjunyawong N, Singer BD, et al. Foxp3+ regulatory T cells promote lung epithelial proliferation. *Mucosal Immunol.* 2014;7(6):1440-51.
 40. Liu Q, Dwyer GK, Zhao Y, Li H, Mathews LR, Chakka AB, et al. IL-33-mediated IL-13 secretion by ST2+ Tregs controls inflammation after lung injury. *JCI Insight.* 2019;4(6).
 41. Garibaldi BT, D'Alessio FR, Mock JR, Files DC, Chau E, Eto Y, et al. Regulatory T cells reduce acute lung injury fibroproliferation by decreasing fibrocyte recruitment. *Am J Respir Cell Mol Biol.* 2013;48(1):35-43.
 42. Sakai R, Ito M, Komai K, Iizuka-Koga M, Matsuo K, Nakayama T, et al. Kidney GATA3(+) regulatory T cells play roles in the convalescence stage after antibody-mediated renal injury. *Cell Mol Immunol.* 2021;18(5):1249-61.
 43. Liesz A, Suri-Payer E, Veltkamp C, Doerr H, Sommer C, Rivest S, et al. Regulatory T cells are key cerebroprotective immunomodulators in acute experimental stroke. *Nat Med.* 2009;15(2):192-9.
 44. Ito M, Komai K, Mise-Omata S, Iizuka-Koga M, Noguchi Y, Kondo T, et al. Brain regulatory T cells suppress astrogliosis and potentiate neurological recovery. *Nature.* 2019;565(7738):246-50.
 45. Gadani SP, Walsh JT, Smirnov I, Zheng J, Kipnis J. The glia-derived alarmin IL-33 orchestrates the immune response and promotes recovery following CNS injury. *Neuron.* 2015;85(4):703-9.
 46. Dombrowski Y, O'Hagan T, Dittmer M, Penalva R, Mayoral SR, Bankhead P, et al. Regulatory T cells promote myelin regeneration in the central nervous system. *Nat Neurosci.* 2017;20(5):674-80.
 47. Hemmers S, Schizas M, Rudensky AY. T reg cell-intrinsic requirements for ST2 signaling in health and neuroinflammation. *J Exp Med.* 2021;218(2).
 48. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000;100(1):57-70.
 49. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646-74.
 50. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 2022;12(1):31-46.

51. Kim SK, Cho SW. The Evasion Mechanisms of Cancer Immunity and Drug Intervention in the Tumor Microenvironment. *Front Pharmacol.* 2022;13:868695.
52. Chao JL, Savage PA. Unlocking the Complexities of Tumor-Associated Regulatory T Cells. *J Immunol.* 2018;200(2):415-21.
53. Itahashi K, Irie T, Yuda J, Kumagai S, Tanegashima T, Lin YT, et al. BATF epigenetically and transcriptionally controls the activation program of regulatory T cells in human tumors. *Sci Immunol.* 2022;7(76):eabk0957.
54. Villarreal DO, L'Huillier A, Armington S, Mottershead C, Filippova EV, Coder BD, et al. Targeting CCR8 Induces Protective Antitumor Immunity and Enhances Vaccine-Induced Responses in Colon Cancer. *Cancer Res.* 2018;78(18):5340-8.
55. Van Damme H, Dombrecht B, Kiss M, Roose H, Allen E, Van Overmeire E, et al. Therapeutic depletion of CCR8(+) tumor-infiltrating regulatory T cells elicits antitumor immunity and synergizes with anti-PD-1 therapy. *J Immunother Cancer.* 2021;9(2).
56. Wang T, Zhou Q, Zeng H, Zhang H, Liu Z, Shao J, et al. CCR8 blockade primes anti-tumor immunity through intratumoral regulatory T cells destabilization in muscle-invasive bladder cancer. *Cancer Immunol Immunother.* 2020;69(9):1855-67.
57. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet.* 2009;10(5):295-304.
58. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res.* 2011;21(3):381-95.
59. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods.* 2013;10(12):1213-8.
60. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol.* 2015;109:21 9 1- 9 9.
61. Buenrostro JD, Wu B, Litzenger UM, Ruff D, Gonzales ML, Snyder MP, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature.* 2015;523(7561):486-90.
62. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science.* 2015;348(6237):910-4.
63. Mezger A, Klemm S, Mann I, Brower K, Mir A, Bostick M, et al. High-throughput chromatin accessibility profiling at single-cell resolution. *Nat Commun.* 2018;9(1):3647.

64. Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol.* 2018;18(1):35-45.
65. Escudier B, Porta C, Schmidinger M, Rioux-Leclercq N, Bex A, Khoo V, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2019;30(5):706-20.
66. Jonasch E, Walker CL, Rathmell WK. Clear cell renal cell carcinoma ontogeny and mechanisms of lethality. *Nat Rev Nephrol.* 2021;17(4):245-61.
67. Kase AM, George DJ, Ramalingam S. Clear Cell Renal Cell Carcinoma: From Biology to Treatment. *Cancers (Basel).* 2023;15(3).
68. Restrepo JCA, Millan DAC, Sabogal CAR, Bernal AFP, Donoso WD. New Trends and Evidence for the Management of Renal Angiomyolipoma: A Comprehensive Narrative Review of the Literature. *J Kidney Cancer VHL.* 2022;9(1):33-41.
69. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers.* 2021;7(1):6.
70. Hossain MS, Karuniawati H, Jairoun AA, Urbi Z, Ooi J, John A, et al. Colorectal Cancer: A Review of Carcinogenesis, Global Epidemiology, Current Challenges, Risk Factors, Preventive and Treatment Strategies. *Cancers (Basel).* 2022;14(7).
71. Kumar A, Gautam V, Sandhu A, Rawat K, Sharma A, Saha L. Current and emerging therapeutic approaches for colorectal cancer: A comprehensive review. *World J Gastrointest Surg.* 2023;15(4):495-519.
72. Chalela R, Curull V, Enriquez C, Pijuan L, Bellosillo B, Gea J. Lung adenocarcinoma: from molecular basis to genome-guided therapy and immunotherapy. *J Thorac Dis.* 2017;9(7):2142-58.
73. Araghi M, Mannani R, Heidarnejad Maleki A, Hamidi A, Rostami S, Safa SH, et al. Recent advances in non-small cell lung cancer targeted therapy; an update review. *Cancer Cell Int.* 2023;23(1):162.
74. Ajani JA, Lee J, Sano T, Janjigian YY, Fan D, Song S. Gastric adenocarcinoma. *Nat Rev Dis Primers.* 2017;3:17036.
75. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. *Lancet.* 2020;396(10251):635-48.
76. Chamoli A, Gosavi AS, Shirwadkar UP, Wangdale KV, Behera SK, Kurrey NK, et al. Overview of oral cavity squamous cell carcinoma: Risk factors, mechanisms, and diagnostics. *Oral Oncol.* 2021;121:105451.
77. Tan Y, Wang Z, Xu M, Li B, Huang Z, Qin S, et al. Oral squamous cell carcinomas: state of the field and emerging directions. *Int J Oral Sci.* 2023;15(1):44.

78. Dyrskjot L, Hansel DE, Efstathiou JA, Knowles MA, Galsky MD, Teoh J, et al. Bladder cancer. *Nat Rev Dis Primers*. 2023;9(1):58.
79. Tran L, Xiao JF, Agarwal N, Duex JE, Theodorescu D. Advances in bladder cancer biology and therapy. *Nat Rev Cancer*. 2021;21(2):104-21.
80. Batool A, Karimi N, Wu XN, Chen SR, Liu YX. Testicular germ cell tumor: a comprehensive review. *Cell Mol Life Sci*. 2019;76(9):1713-27.
81. Tsilimigras DI, Brodt P, Clavien PA, Muschel RJ, D'Angelica MI, Endo I, et al. Liver metastases. *Nat Rev Dis Primers*. 2021;7(1):27.
82. Brodt P. Role of the Microenvironment in Liver Metastasis: From Pre- to Prometastatic Niches. *Clin Cancer Res*. 2016;22(24):5971-82.
83. Granja JM, Corces MR, Pierce SE, Bagdatli ST, Choudhry H, Chang HY, et al. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet*. 2021;53(3):403-11.
84. Vincent D Blondel J-LG, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *IOPscience*. 2008.
85. van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, et al. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell*. 2018;174(3):716-29 e27.
86. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.
87. Wilcoxon F. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*. 1945;Vol. 1, No. 6(Dec., 1945):80-3.
88. Yoav Benjamini YH. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1995.
89. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods*. 2017;14(10):975-8.
90. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-6.
91. Amezquita RA, Lun ATL, Becht E, Carey VJ, Carpp LN, Geistlinger L, et al. Orchestrating single-cell analysis with Bioconductor. *Nat Methods*. 2020;17(2):137-45.
92. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-87 e29.

93. Aaron Lun DR, Keegan Korthauer, Kevin Rue-Albrecht. SingleCellExperiment: S4 Classes for Single Cell Data.
94. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 2015;12(2):115-21.
95. McCarthy DJ, Campbell KR, Lun AT, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics*. 2017;33(8):1179-86.
96. Germain PL, Lun A, Garcia Meixide C, Macnair W, Robinson MD. Doublet identification in single-cell sequencing data using scDbfFinder. *F1000Res*. 2021;10:979.
97. Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol*. 2018;36(5):421-7.
98. Liu G, Xiong D, Che Z, Chen H, Jin W. A novel inflammation-associated prognostic signature for clear cell renal cell carcinoma. *Oncol Lett*. 2022;24(3):307.
99. Shang B, Liu Y, Jiang SJ, Liu Y. Prognostic value of tumor-infiltrating FoxP3+ regulatory T cells in cancers: a systematic review and meta-analysis. *Sci Rep*. 2015;5:15179.
100. Saito T, Nishikawa H, Wada H, Nagano Y, Sugiyama D, Atarashi K, et al. Two FOXP3(+)CD4(+) T cell subpopulations distinctly control the prognosis of colorectal cancers. *Nat Med*. 2016;22(6):679-84.
101. Elkoshi Z. On the Prognostic Power of Tumor-Infiltrating Lymphocytes - A Critical Commentary. *Front Immunol*. 2022;13:892543.
102. Plitas G, Konopacki C, Wu K, Bos PD, Morrow M, Putintseva EV, et al. Regulatory T Cells Exhibit Distinct Features in Human Breast Cancer. *Immunity*. 2016;45(5):1122-34.
103. De Simone M, Arrigoni A, Rossetti G, Gruarin P, Ranzani V, Politano C, et al. Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells. *Immunity*. 2016;45(5):1135-47.
104. Alvisi G, Brummelman J, Puccio S, Mazza EM, Tomada EP, Losurdo A, et al. IRF4 instructs effector Treg differentiation and immune suppression in human cancer. *J Clin Invest*. 2020;130(6):3137-50.
105. Green JA, Arpaia N, Schizas M, Dobrin A, Rudensky AY. A nonimmune function of T cells in promoting lung tumor progression. *J Exp Med*. 2017;214(12):3565-75.
106. Halvorsen EC, Franks SE, Wadsworth BJ, Harbourne BT, Cederberg RA, Steer CA, et al. IL-33 increases ST2(+) Tregs and promotes metastatic tumour

- growth in the lungs in an amphiregulin-dependent manner. *Oncoimmunology*. 2019;8(2):e1527497.
107. Wang L, Wang L, Zhang H, Lu J, Zhang Z, Wu H, et al. AREG mediates the epithelial-mesenchymal transition in pancreatic cancer cells via the EGFR/ERK/NF-kappaB signalling pathway. *Oncol Rep*. 2020;43(5):1558-68.
 108. Xu Q, Long Q, Zhu D, Fu D, Zhang B, Han L, et al. Targeting amphiregulin (AREG) derived from senescent stromal cells diminishes cancer resistance and averts programmed cell death 1 ligand (PD-L1)-mediated immunosuppression. *Aging Cell*. 2019;18(6):e13027.
 109. Salimi M, Barlow JL, Saunders SP, Xue L, Gutowska-Owsiak D, Wang X, et al. A role for IL-25 and IL-33-driven type-2 innate lymphoid cells in atopic dermatitis. *J Exp Med*. 2013;210(13):2939-50.
 110. Monticelli LA, Sonnenberg GF, Abt MC, Alenghat T, Ziegler CG, Doering TA, et al. Innate lymphoid cells promote lung-tissue homeostasis after infection with influenza virus. *Nat Immunol*. 2011;12(11):1045-54.
 111. Monticelli LA, Osborne LC, Noti M, Tran SV, Zaiss DM, Artis D. IL-33 promotes an innate immune pathway of intestinal tissue protection dependent on amphiregulin-EGFR interactions. *Proc Natl Acad Sci U S A*. 2015;112(34):10762-7.
 112. Qi Y, Operario DJ, Oberholzer CM, Kobie JJ, Looney RJ, Georas SN, et al. Human basophils express amphiregulin in response to T cell-derived IL-3. *J Allergy Clin Immunol*. 2010;126(6):1260-6 e4.
 113. Meulenbroeks C, van Weelden H, Schwartz C, Voehringer D, Redegeld FAM, Rutten V, et al. Basophil-derived amphiregulin is essential for UVB irradiation-induced immune suppression. *J Invest Dermatol*. 2015;135(1):222-8.
 114. Matsumoto K, Fukuda S, Nakamura Y, Saito H. Amphiregulin production by human eosinophils. *Int Arch Allergy Immunol*. 2009;149 Suppl 1:39-44.
 115. du Halgouet A, Bruder K, Peltokangas N, Darbois A, Obwegs D, Salou M, et al. Multimodal profiling reveals site-specific adaptation and tissue residency hallmarks of gammadelta T cells across organs in mice. *Nat Immunol*. 2024;25(2):343-56.
 116. Kidani Y, Nogami W, Yasumizu Y, Kawashima A, Tanaka A, Sonoda Y, et al. CCR8-targeted specific depletion of clonally expanded Treg cells in tumor tissues evokes potent tumor immunity with long-lasting memory. *Proc Natl Acad Sci U S A*. 2022;119(7).
 117. Liu S, Tao Z, Lou J, Li R, Fu X, Xu J, et al. CD4(+)CCR8(+) Tregs in ovarian cancer: a potential effector Tregs for immune regulation. *J Transl Med*. 2023;21(1):803.
 118. Weaver JD, Stack EC, Bugge JA, Hu C, McGrath L, Mueller A, et al. Differential expression of CCR8 in tumors versus normal tissue allows specific depletion of

- tumor-infiltrating T regulatory cells by GS-1811, a novel Fc-optimized anti-CCR8 antibody. *Oncoimmunology*. 2022;11(1):2141007.
119. Ueyama A, Nogami W, Nashiki K, Haruna M, Miwa H, Hagiwara M, et al. Immunotherapy Targeting CCR8+ Regulatory T Cells Induces Antitumor Effects via Dramatic Changes to the Intratumor CD8+ T Cell Profile. *J Immunol*. 2023;211(4):673-82.
 120. Mijnheer G, Lutter L, Mokry M, van der Wal M, Scholman R, Fleskens V, et al. Conserved human effector Treg cell transcriptomic and epigenetic signature in arthritic joint inflammation. *Nat Commun*. 2021;12(1):2710.
 121. Oliver Burton OB, Samar Tareen, Vaclav Gergelits, Simon Andrews, Carlos P. Roca, Carly Whyte, Steffie Junius, Aleksandra Brajic, Emanuela Pasciuto, Magda Ali, Pierre Lemaitre, Susan M. Schlenner, Harumichi Ishigame, Brian D. Brown, James Dooley, Adrian Liston. The tissue-resident regulatory T cell pool is shaped by transient multi-tissue migration and a conserved residency program. *bioRxiv*. 2023.
 122. Barsheshet Y, Wildbaum G, Levy E, Vitenshtein A, Akinseye C, Griggs J, et al. CCR8(+)FOXP3(+) T(reg) cells as master drivers of immune regulation. *Proc Natl Acad Sci U S A*. 2017;114(23):6086-91.
 123. Kuehnemuth B, Piseddu I, Wiedemann GM, Lauseker M, Kuhn C, Hofmann S, et al. CCL1 is a major regulatory T cell attracting factor in human breast cancer. *BMC Cancer*. 2018;18(1):1278.
 124. Howard OM, Dong HF, Shirakawa AK, Oppenheim JJ. LEC induces chemotaxis and adhesion by interacting with CCR1 and CCR8. *Blood*. 2000;96(3):840-5.
 125. Islam SA, Ling MF, Leung J, Shreffler WG, Luster AD. Identification of human CCR8 as a CCL18 receptor. *J Exp Med*. 2013;210(10):1889-98.
 126. Cinier J, Hubert M, Besson L, Di Roio A, Rodriguez C, Lombardi V, et al. Recruitment and Expansion of Tregs Cells in the Tumor Environment-How to Target Them? *Cancers (Basel)*. 2021;13(8).
 127. Peng Q, Zhao L, Hou Y, Sun Y, Wang L, Luo H, et al. Biological characteristics and genetic heterogeneity between carcinoma-associated fibroblasts and their paired normal fibroblasts in human breast cancer. *PLoS One*. 2013;8(4):e60321.
 128. Ma L, Wang H, Li Z, Geng X, Li M. Chemokine (C-C motif) ligand 18 is highly expressed in glioma tissues and promotes invasion of glioblastoma cells. *J Cancer Res Ther*. 2019;15(2):358-64.
 129. Delacher M, Schmidleithner L, Simon M, Stüve P, Sanderink L, Hotz-sWagenblatt A, Wuttke M, Schambeck K, Ruhland B, Hofmann V, Bittner S, Ritter U, Pant A, Helbich SS, Voss M, Lemmermann NA, Bessiri-Schake L, Bohn T, Eigenberger A, Menevse AN, Gebhard C, Strieder N, Abken H, Rehli M, Huehn J, Beckhove P, Hehlhans T, Junger H, Geissler EK, Prantl L, Werner

JM, Schmidl C, Brors B, Imbusch CD, Feuerer M. The effector program of human CD8 T cells supports tissue remodeling. J Exp Med. 2024 Feb 5;221(2):e20230488. s

8.2 Figures

Figure 1: CCR8 ⁺ Treg cell.....	3
Figure 2: CCR8 ⁺ Treg cells in murine and human tissues	6
Figure 3: Treg cell-mediated tissue regeneration	7
Figure 4: The hallmarks of cancer	11
Figure 5: Overview scATAC-seq.....	13
Figure 6: Overview scRNA/TCR-seq	15
Figure 7: Gating strategy	26
Figure 8: scATAC-seq library fragment length distribution	27
Figure 9: Schematic of scATAC-seq library fragment.....	28
Figure 10: Quality filtering	31
Figure 11: Data visualized as UMAP plot	33
Figure 12: UMAP encoding doublet enrichment score and number of fragments per cell	35
Figure 13: UMAP plots encoding gene scores.....	37
Figure 14: UMAP encoding CCR8 ⁺ Treg signature z-score.....	38
Figure 15: Differential peaks	40
Figure 16: TF motif enrichment in differential peaks.....	41
Figure 17: IGV browser tracks	42
Figure 18: Schematic of scRNA-seq gene expression (GEX) and V(D)J library fragments. .	42
Figure 19: Quality filtering	48
Figure 20: t-SNE plot encoding doublet information	49
Figure 21: Scree plot.....	51
Figure 22: t-SNE plot	52
Figure 23: t-SNE plots encoding gene expression.....	53
Figure 24: Top differentially expressed genes	54
Figure 25: Volcano plot.....	55
Figure 26: TCR repertoire piechart	56
Figure 27: t-SNE plot encoding clonotype information.....	57
Figure 28: CCR8 ⁺ Treg cells are present in human tumor- and healthy tissue.....	59
Figure 29: scATAC-seq of kidney tumors (CCRC and AML).....	61
Figure 30: scATAC-seq of kidney tumors (CCRC and AML).....	63
Figure 31: scATAC-seq of liver tumors (HCC).....	65
Figure 32: scATAC-seq of tumors of the bladder, oral floor, lungs, stomach and testes.....	69
Figure 33: H&E staining of the NAT of an HCC patient and the NAT of a patient with metastasis to the liver.	70
Figure 34: scATAC seq of metastases to the liver.....	72
Figure 35: scATAC-seq of CRC with metastasis to the liver.....	74

Figure 36: scRNA/TCR-seq of CCRC.....76

8.3 Tables

Table 2: Tissue samples processed for scATAC-seq.....22
 Table 3: Tissue samples in the scRNA/TCR-seq dataset.....23
 Table 4: Digestion buffer for liver, kidney, testes and bladder tissue.....24
 Table 5: Digestion buffer for colon tissue.....24
 Table 6: Digestion buffer for lung tissue24
 Table 7: Digestion buffer for stomach tissue.....24
 Table 8: Digestion buffer for pancreas tissue24
 Table 9: Digestion buffer for oral floor tissue25
 Table 10: Nuclei prep lysis buffer.....27
 Table 11: Nuclei prep wash buffer27

8.4 Code

Box 1: bcl2fastq29
 Box 2: Cell Ranger ATAC29
 Box 3: Creating arrow files.....30
 Box 4: Creating the ArchRProject.....30
 Box 5: QC filtering step 131
 Box 6: QC filtering step 231
 Box 7: Dimensionality reduction.....33
 Box 8: Clustering.....33
 Box 9: Doublet filtering.....34
 Box 10: Cell type annotation36
 Box 11: Calculate signature z-scores37
 Box 12: Pseudobulk replicates.....39
 Box 13: Peak calling39
 Box 14: Extract differential peaks40
 Box 15: TF motif enrichment.....40
 Box 16: Export bigwig files.....41
 Box 17: Cell Ranger.....43
 Box 18: Creating the SCE object45
 Box 19: Converting gene identifiers to gene names45
 Box 20: QC filtering step 146
 Box 21: QC filtering step 247

APPENDIX

Box 22: Doublet filtering.....	48
Box 23: Dimensionality reduction.....	49
Box 24: Clustering.....	51
Box 25: Cell type annotation.....	53
Box 26: Identifying top differentially expressed genes.....	53
Box 27: Differential gene expression analysis.....	54
Box 28: TCR repertoire diversity analysis.....	55
Box 29: Identify shared clonotypes.....	56

9. CV

10. Danksagung

Danksagung
