

DISS. ETH NO. 29969

Multimodal single-cell profiling of T lymphocytes in infection and cancer

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES
(Dr. sc. ETH Zurich)

presented by

Florian Bieberich
M.Sc. Heidelberg University

born on 5.12.1995

accepted on the recommendation of
Prof. Dr. Sai Reddy
Prof. Dr. Barbara Treutlein
Prof. Dr. Alfred Zippelius

2024

“I don’t pretend we have all the answers. But the questions are certainly worth thinking about.” - Arthur Charles Clarke

Abstract

The adaptive immune system is highly complex and can target foreign antigens with high specificity. Central to this system, T and B cells enable the specified targeting through expression of unique antigen receptors which are produced by genetic recombination. This process creates an immense diversity of antigen receptors with a theoretical diversity of T cell receptors of 10^{18} , enabling the adaptive immune system to recognize and respond to almost any non-self-protein. While B cells target a broad spectrum of extracellular protein structures, T cells recognize intracellularly processed proteins that are presented on the surface of almost every nucleated cell in the body, enabling them to detect and eliminate cells infected by viruses or transformed by oncogenic mutations. Through applying multimodal single-cell sequencing and functional reactivity testing, we uncover novel insights into adaptive immune cell dynamics, focusing on COVID-19 convalescence and T cell reactivity to lung cancer.

Chapter 1 of this thesis introduces the principles of T cell receptor generation, how they can bind to their cognate target and, upon target recognition, T cell activation and subsequent cellular differentiation. Consecutively, we highlight the role of T cells in infection and cancer as well as cancer immunotherapeutic applications of T cells. Finally, we introduce single-cell sequencing and its coupling with functional T cell reactivity.

In Chapter 2, we use single-cell sequencing to study adaptive immune responses in convalescent COVID-19 patients of varying ages. By analyzing the T and B cell receptor repertoires, paired with the cells' transcriptomes at the single-cell level, we identify differences in B cell receptors and T cell differentiation dynamics linked to disease duration and patient age. Further, this analysis allows us to pinpoint distinct clonal expansion patterns and identify potential SARS-CoV-2 reactive T cells that are persisting post-recovery through *in silico* reactivity matching.

Chapter 3 focuses on the transcriptome-based identification of tumor-reactive T cells. We develop a workflow for the unbiased discovery of reactivity, combining Cas9-mediated T cell receptor knockin, fluorescence-activated cell sorting, and next-generation sequencing. Using this workflow with T cells from treatment-naive lung cancer patients, we functionally screened ca. 140 T cell receptors for their reactivity. We identified differentially expressed genes in tumor-reactive T cells by integrating this data with multimodal single-cell sequencing.

Subsequently, the marker genes enabled the creation of a "tumor score," enhancing our ability to enrich tumor-reactive T cells. This thesis shows the potential of combining multimodal single-cell sequencing with innovative genetic technologies for therapeutic development and personalized medicine.

Zusammenfassung

Das adaptive Immunsystem ist sehr komplex und kann fremde Antigene mit hoher Spezifität erkennen. T- und B-Zellen sind zentral für diese Spezifität, da sie einzigartige Antigenrezeptoren exprimieren, die durch genetische Rekombination entstehen. Dieser Prozess führt zu einer immensen Diversität an Antigenrezeptoren mit einer theoretischen Vielfalt von z.B. 10^{18} T-Zell-Rezeptoren, die es dem adaptiven Immunsystem ermöglichen, nahezu jede fremde Proteinstruktur zu erkennen und darauf zu reagieren. Während B-Zellen ein breites Spektrum extrazellulärer Proteinstrukturen erkennen, identifizieren T-Zellen intrazellulär prozessierte Proteine, die auf fast jeder kernhaltigen Zelle im Körper vorkommen, so dass sie von Viren infizierte oder durch onkogene Mutationen veränderte Zellen erkennen und eliminieren können. Durch die Anwendung multimodaler Einzelzell-Sequenzierung und funktionaler Reaktivitätstests erhalten wir in dieser Arbeit neue Erkenntnisse über die Dynamik adaptiver Immunzellen, wobei wir uns auf die Genesung von COVID-19 und die T-Zell Reaktivität in Lungenkrebs fokussieren.

In Kapitel 1 dieser Arbeit werden die Prinzipien der T-Zell-Rezeptor-Generierung, das Erkennen und Binden der T-Zelle an ihr passendes Ziel und die dadurch induzierte T-Zell-Aktivierung, sowie die anschließende zelluläre Differenzierung vorgestellt. Anschließend wird die Rolle der adaptiven Immunzellen und vor allem von T-Zellen bei Infektionen und Krebs sowie die immuntherapeutischen Anwendungen von T-Zellen für die Krebsbekämpfung beleuchtet. Schließlich stellen wir die Einzelzell-Sequenzierung und ihre Kombination mit der funktionellen T-Zell-Reaktivität vor.

In Kapitel 2 verwenden wir die Einzelzell-Sequenzierung zur Untersuchung der adaptiven Immunantwort bei genesenen COVID-19-Patienten unterschiedlichen Alters. Durch die Analyse des T- und B-Zell Repertoires, gepaart mit dem Transkriptom der Zellen auf Einzelzell-Ebene, identifizieren wir Unterschiede in den B-Zell-Rezeptoren und der T-Zell-Differenzierungsdynamik, die mit der Krankheitsdauer und dem Alter der Patienten zusammenhängen. Darüber hinaus ermöglicht uns diese Analyse, unterschiedliche klonale Expansionsmuster zu erkennen und durch *In-silico*-Reaktivitätsanalysen mögliche SARS-CoV-2-reaktive T-Zellen zu identifizieren, die nach der Genesung weiter bestehen bleiben.

Kapitel 3 konzentriert sich auf die transkriptom-basierte Identifizierung von tumor-reaktiven T-Zellen. Wir entwickeln einen Arbeitsablauf für die unverfälschte Entdeckung der T-Zell-Reaktivität, welcher Cas9-induzierten T-Zell-Rezeptor-Knockin, fluoreszenz-aktivierte Zellsortierung und Hochdurchsatz-Sequenzierung kombiniert. Unter Verwendung dieses Ansatzes mit T-Zellen von unbehandelten Lungenkrebspatienten haben wir ca. 140 T-Zell-Rezeptoren funktionell auf ihre Reaktivität untersucht. Durch die Integration dieser Daten mit multimodaler Einzelzell-Sequenzierung konnten wir Gene identifizieren, die spezifisch in tumor-reaktiven T-Zellen exprimiert werden. Diese Gene ermöglichten anschließend die Erstellung eines "Tumor-Scores", welcher uns bei der Selektion tumor-reaktiver T-Zellen hilft. Diese Arbeit zeigt das Potenzial der Kombination multimodaler Einzelzell-Sequenzierung mit innovativen Gentechnologien für die Identifikation von neuen Biomarkern mit therapeutischer Anwendung und in der personalisierten Medizin.

Acknowledgements

In this journey of academic and personal growth, I owe a debt of gratitude to numerous individuals whose guidance, support, and inspiration have been pivotal.

First and foremost, my thanks go to Sai Reddy for not only allowing me to pursue my interests in his lab but also for maintaining an open door policy. Our discussions about translational science have been invaluable for me and my career.

I extend my gratitude to Rodrigo, who not only supervised and co-led our first project but also demonstrated the importance of perseverance in scientific endeavors. His journey of translating research into a startup has been a source of great inspiration to me.

I am deeply thankful to Barbara Treutlein and Alfred Zippelius for their invaluable mentorship as doctoral committee members. Our discussions around single cell technologies and the clinical implications of my research results have profoundly impacted my work.

My appreciation also goes to Kwankwan Zhu, Miral Tariq, Huixin Jin, Sandro Fliri, Elena Boschet, and Xu Yangguang, whom I had the pleasure of supervising as my master students. Their enthusiasm and dedication have been a constant source of motivation and enabled me to grow.

I am grateful to Bastian Wagner, Rocio Castellanos-Rueda, Kai-Lin Hong, and Roy Ehling for their collaboration on various projects and our enriching discussions.

A special thank you to the single cell and genomics facility team - Christian, Ina, Mariangela, Chiara, Renan, and Ola - for their indispensable assistance with my FACS and NGS projects.

The entire Reddy lab deserves my deepest appreciation for their tremendous support, both inside and outside the lab. Special mentions to Rocio, Basti, Federica, Rita, Mason and Fabrice for their helpful project discussions and memorable post-retreat hikes.

Reflecting on my journey, I must acknowledge my physics and biology class teachers, Jan Maschinsky and Rita Hergenröder. Their passion for scientific discoveries and engaging class setups steered me unknowingly towards a career in natural sciences.

My time at the University of Cambridge, under John Marioni and Aaron Lun, sparked my interest in single cell data analysis. Similarly, my projects at Harvard Medical School with Pamela Silver and Jeffrey Way, and at Osaka University with Tomohiro Kurosaki and Takeshi Inoue, ignited my passion for the intersection of synthetic biology and immunology. I am also thankful to my master thesis supervisors, Rienk Offringa and Isabel Poschke at DKFZ in

Heidelberg, for inspiring me with the potential of tumor-reactive T cells in cancer immunotherapy.

My thanks also to Dimitri and Margaux for their enthusiasm in building our student organization and for their inspiring volunteer spirit.

A special shoutout to all my friends, especially Lukas, Lovis, Fritz, Jakob, Guido, Dimitri, Rocio, Margaux, Nils and many others - for their unwavering support, engaging discussions, introducing new sports into my life, and shared adventures in Switzerland and the world, which have made the past four years an incredible time to look back at. - *“A friend is one soul abiding in two bodies” - Aristotle*

I express my deepest gratitude to my parents Patricia and Manfred for instilling in me a fascination with nature, the joy of experimentation and support in every situation, and to my grandparents Irene and Heinrich for their continuous excitement and support.

Finally, thank you cc, for being my calm and my inspiration.

Contributions to this dissertation

Dr. Rodrigo Vazquez-Lombardi is co-first author of the research project: “A single-cell atlas of lymphocyte adaptive immune repertoires and transcriptomes reveals age-related differences in convalescent COVID-19 patients”. Further, he supervised and helped with the experimental design of the functional T cell reactivity screening work using TnT cells as part of the research project: “Multimodal single-cell profiling of T cell specificity and reactivity in lung cancer”

Huixin Jin analyzed single-cell multiome data as part of the research project: “Multimodal single-cell profiling of T cell specificity and reactivity in lung cancer”.

Kai-Lin Hong genetically engineered HEK cells to express patient autologous HLA receptors as part of the research project: “Multimodal single-cell profiling of T cell specificity and reactivity in lung cancer”.

Abbreviations.....	1
Chapter 1: Introduction.....	2
1.1 The human adaptive immune system.....	2
1.1.1 T cells.....	2
1.2 T cell dynamics in infection and cancer.....	5
1.2.1 T cell activation and cytotoxicity.....	6
1.2.2 T cell differentiation.....	7
1.2.3 Cancer-immune interaction.....	7
1.2.4 Tumor infiltrating lymphocytes.....	8
1.2.5 Cancer immunotherapy.....	9
1.3 Single-cell sequencing for T cell analysis.....	11
1.4 Linking functional reactivity to cell state.....	13
1.5 References of Chapter 1.....	15
Hypothesis and Objectives.....	22
Chapter 2: Single cell lymphocyte analysis of convalescent COVID-19 patients.....	24
2.1 Summary.....	24
2.2 Introduction.....	25
2.3 Results.....	28
2.4 Discussion.....	43
2.5 Methods.....	46
2.6 References of Chapter 2.....	53
Chapter 3: Multimodal single-cell profiling of T cell specificity and reactivity in lung cancer.....	57
3.1 Summary.....	57
3.2 Introduction.....	59
3.3 Results.....	62
3.4 Discussion.....	80
3.5 Methods.....	83
3.6 References of Chapter 3.....	94
Chapter 4: General discussion.....	101
4.1 Conclusion.....	101
4.2 Contributions of this thesis.....	102
4.3 Future directions.....	105
4.3.1 Adaptive immune state assessment for personalized medicine.....	105
4.3.2 Tumor reactive T cell discovery for personalized therapies.....	106
4.3.3 TCR identification for other diseases.....	109
4.4 References of Chapter 4.....	111

Appendix.....	117
A.1 Supplementary information of Chapter 2.....	117
A.2 Supplementary information of Chapter 3.....	128
Curriculum Vitae.....	134

Abbreviations

APC	Antigen-presenting cell
ATAC	Assay for Transposase-Accessible Chromatin
ACT	Adoptive cell transfer
BCR	B cell receptor
CAR	Chimeric antigen receptor
cDNA	complementary DNA
CDR	Complementarity determining region
CMV	Cytomegalovirus
CTLA-4	Cytotoxic T lymphocyte associated protein 4
CXCL13	C-X-C motif chemokine 13
DGE	Differential Gene Expression
EBV	Epstein-Barr virus
FACS	Fluorescence-activated cell sorting
GEX	Gene expression
HLA	Human leukocyte antigen
ICB	Immune checkpoint blockade
Ig	Immunoglobulin
ITAM	Immunoreceptor-tyrosine-based-activation-motif
MHC	Major Histocompatibility complex
NFAT	Nuclear factor of activated T cells
NGS	Next generation sequencing
NSCLC	Non-small cell lung cancer
PD1	Programmed cell death 1
RT	Reverse transcription
sc	Single-cell
scSeq	Single-cell sequencing
sgRNA	Single guide RNA
SHM	Somatic hypermutation
TAA	Tumor-associated antigen
TCR	T cell receptor
TF	Transcription factor
TIL	Tumor infiltrating lymphocyte
TME	Tumor microenvironment
UMAP	Uniform manifold approximation and projection

Chapter 1: Introduction

1.1 The human adaptive immune system

The adaptive immune system consists of T and B lymphocytes, which are essential for the human immune response due to their precise targeting of a broad spectrum of foreign antigens. Each naïve lymphocyte expresses a distinct antigen receptor, enabling recognition and binding to antigens, thereby creating a diverse and extensive repertoire of T and B lymphocytes. A critical feature of this system is its capacity for immunological memory, a process that ensures a more rapid and potent immune response upon subsequent encounters with a previously recognised foreign antigen ¹.

1.1.1 T cells

T lymphocytes or T cells possess a highly variable receptor called the T cell receptor (TCR), which consists of an α - and a β -chain. Each naïve T cell expresses a unique TCR with an exclusive specificity for various antigens. The high diversity of TCRs is achieved through the irreversible recombination of gene segments. In more detail, the VDJ recombination process allows the somatic recombination of V-, D- and J-gene segments in developing T cells (a similar process exists in B lymphocytes) (**Figure 1**) ². The diversity emerging from such genetic recombination events is further increased through nucleotide insertions/deletions in the junction regions of the V-, D- and J-gene segments. This process is analogously taking place for the α - (70 V-, 0 D- and 61 J-genes) and the β - (52 V-, 2 D- and 13 J-genes) chain with an estimated theoretical TCR diversity of 10^{18} ³. Since the VDJ recombination process occurs in the genomic DNA of each developing T cell, the progeny of each T cell will inherit the same TCR sequence. Therefore, each T cell can be identified by its TCR sequence and T cells with the same TCR are termed T cell clones or clonal T cells.

The V-gene and the junctional region between V-, (D-), and J-genes account for the highest diversity and are also called complementarity determining regions (CDRs). CDR1 and CDR2 regions are part of the V-gene segment and have a lower variability, while CDR3 has the highest variability and describes the junction in which the V-, (D-), and J-regions are combined. The entire collection of TCRs in a human being at any given time is defined as this individual's TCR

repertoire. Analysis of an individual's repertoire diversity or T cell subpopulations (i.e. TCR repertoire in the tumor), as well as clonal sizes, can give insights into the state of infection, cancer or age⁴⁻⁶.

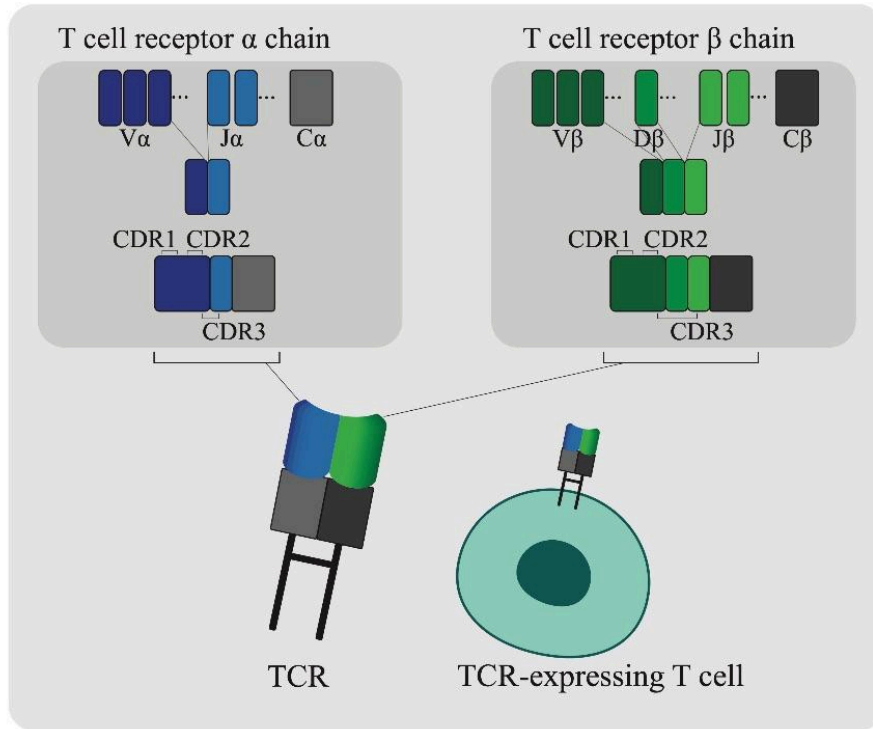


Figure 1: Schematic of the T cell receptor VDJ recombination.

The TCR interacts with the major histocompatibility complex (MHC), also called the human leukocyte antigen (HLA) receptor, to bind to and recognise a foreign antigen. This membrane-bound receptor presents intracellularly processed peptides in its peptide-binding groove. Due to the high impact of the HLA groove sequence on the selection of peptide residues that can interact with the HLA groove, the HLA gene locus represents the area with the highest degree of polymorphism in the human genome^{7,8}. On a population level, the high degree of HLA diversity allows part of the population to present specific peptides better than the others. For instance, during a pandemic, this enables a subgroup of people to present a virus's peptides more effectively and, in turn, have a higher chance of a successful immune response. The TCR's CDR3 regions most strongly interact with the bound peptide, whereas CDR1 and CDR2 mainly interact with the HLA and have a stabilizing effect⁹⁻¹². HLA receptors can be divided into two

subclasses, HLA class I and HLA class II, differing in their structure and eliciting different peptide presentation patterns.

One of two T cell co-ligands, CD4 or CD8, engages with the HLA complex to strengthen and prolong the interaction between TCR and peptide-HLA. Most T cells express CD4 or CD8 and are therefore termed CD4⁺ or CD8⁺ T cells. Since CD4 interacts with HLA class II and CD8 interacts with HLA class I, the CD4⁺ and CD8⁺ polarization also determines if the TCR interacts with HLA class II or HLA class I, respectively ¹³⁻¹⁵.

HLA class I is expressed on almost all nucleated cells, whereas HLA class II is only expressed on professional antigen-presenting cells (APCs) such as macrophages, dendritic cells, and B cells. Thus, depending on the CD4/CD8 polarization, T cells can interact with different cells across the body.

The peptides binding to the HLA groove are mostly 8-10 amino acids in length for HLA class I, whereas HLA class II can bind to peptides of variable lengths, often exceeding 10 amino acids ¹⁶⁻¹⁸. HLA class I can be subdivided into three subtypes: HLA A, B and C. Due to the bi-allelic nature of our genome, humans carry two versions of each HLA class I subtype. A similar diversity exists for HLA class II. The work described in this thesis will mainly focus on CD8⁺ T cells and their interactions with HLA class I.

Owing to the high diversity of TCRs and the resulting number of peptides that can be targeted, there is a theoretically high potential for TCR reactivity against peptides resulting from intracellular processing of healthy self-proteins. To prevent TCRs from recognising such self-peptide-presenting cells, thymic selection occurs during the early stages of T cell development in the thymus ¹⁹. In short, TCRs are tested for their binding strength against self-HLAs that present self-peptides. Strong and non-binding T cells undergo targeted apoptosis, with only weak-binding TCRs selected for maturation ²⁰⁻²². This ensures a minimal amount of T cells eliciting autoimmune responses and a preferred selection of T cells that can interact with self-HLAs.

To understand the importance of T cells as part of the adaptive immune system, it helps to consider the systems immunology perspective on the processes happening during infection and cancer. B cells and their B cell receptors (BCRs), which in soluble form are called antibodies, can bind to surface proteins such as surface antigens of viruses or bacteria that are able to pass physical barriers and enter the body. T cells, on the other hand, can bind to intracellularly

processed peptides. These intracellularly processed peptides arise from ubiquitinated (a posttranslational modification that marks a protein for degradation) proteins that are processed into peptides through proteasomal degradation^{23–25}. The proteasome continuously processes proteins, whose peptides are subsequently presented on the cell's HLA. During infection, pathogens such as viruses can enter their target cell and use the intracellular transcriptional and translational machinery to produce viral copies of themselves. In cancer, oncogenic mutations that allow the cancer cell to proliferate more rapidly are one example of how the normal protein sequence can be altered. Thus, the process of intracellular peptide presentation allows T cells to detect foreign peptides from viruses and mutation-altered proteins in cancer cells²⁶. The following section will detail the interaction between T cells and antigen-presenting cells.

1.2 T cell dynamics in infection and cancer

During a viral infection, the intracellularly replicating virus typically fuses with the host cell's membrane and hijacks the cellular machinery to replicate itself. This process leads to viral RNA and protein production inside the host cell, keeping it undetected from the immune system. Similar to self-proteins, a fraction of viral proteins will undergo proteasomal degradation into short peptides. Suppose a peptide, also called an epitope, results from this degradation process that has the correct affinity to co-locate into the binding groove of the HLA receptor. In this case, the peptide-HLA receptor complex will translocate to the surface, presenting itself extracellularly to be potentially recognised by a peptide-HLA-specific (cognate) TCR²⁷.

Due to the TCR's selection process in the thymus, as described earlier, its affinity to foreign peptides is higher than to self.

For antigen-inexperienced (naïve) T cells to get activated, they undergo a process called T cell priming, where they need to encounter their cognate peptide-HLA presented by a professional APC, i.e., a dendritic cell^{28,29}. The need for priming to occur through a professional APC is because three different signals are involved in the priming of naïve T cells to ensure proper activation and avoid reactivity to self-antigens. In addition to the (I) TCR-peptide-HLA binding signal, these are (II) costimulatory-receptor binding, usually the T cell-specific surface glycoprotein CD28 that is triggered through the same APC, and (III) secreted cytokines from the APC to further support the differentiation of the primed T cell^{30,31}.

Most viruses exhibit a tropism to infect a specific cell type, and these often exclude professional APCs. However, professional APCs have evolved a dual mechanism to ensure that even if uninfected, they can still prime naïve T cells effectively. Not only can professional APCs present intracellular peptides on their HLA, but they can also take up extracellular peptides (such as those released from infected apoptotic cells) and initiate cross-priming to present such peptides on their HLA for T cells to recognize ³².

1.2.1 T cell activation and cytotoxicity

Once a TCR binds to its cognate peptide-HLA, with the help of the CD4 or CD8 co-receptor, the TCR transmits a signal to intracellularly associated components. These so-called CD3 molecules, co-associated with the TCR, contain immunoreceptor tyrosine-based activation motifs (ITAMs) that, in turn, get phosphorylated ^{33–35}. Together with the CD28 co-receptor activation and cytokine signaling, this triggers further downstream events that lead to a cytosolic influx of Ca^{2+} , and together with other kinases and phosphatases, initiates the activation of a family of transcription factors (TFs), one of them being called nuclear factor of activated T cells (NFAT) ³⁶. Ultimately, this results in the TF-induced expression of genes, such as Interleukin-2 (IL-2) and CD69, responsible for survival, proliferation, and differentiation into effector and memory cells, initiating rapid cell division that, for T cells, is also called clonal expansion ^{37,38}. While CD8^+ T cells mainly differentiate into cytotoxic and memory T cells ^{39,40} that can target and kill cognate peptide-HLA presenting cells, CD4^+ T cells can differentiate into various subsets, performing different effector and regulatory functions ⁴¹. For the scope of this thesis, we will focus on the differentiation and role of CD8^+ T cells.

After undergoing clonal expansion and differentiating into an effector state, CD8^+ T cells can elicit a response without co-stimulatory signals ^{42,43}. This independence from professional APCs is essential for responding to, i.e., virus-infected cells. Additionally, an altered expression of cellular adhesion molecules causes primed T cells to stop circulating in the lymph nodes and instead egress to sites of inflammation ^{44,45}. Upon binding of the TCR to its cognate peptide-HLA, the formation of the immunological synapse between the T cell and the target cell is initiated ⁴⁶. The immunological synapse allows the targeted transport of effector molecules such as perforin and granzymes into the target cell, where they induce cell death. In addition, triggered effector T cells also initiate the production of other cytokines ⁴⁷, such as $\text{IFN}\gamma$ and

TNF α , that lead to the recruitment and activation of other immune cells as well as, specifically for IFN γ , to the upregulation of MHC class I and II on surrounding cells ⁴⁸, collectively enabling a prolonged and strengthened immune response.

1.2.2 T cell differentiation

Immunological memory is one of the major pillars of adaptive immunity. After a T cell recognises its cognate peptide-HLA on a professional APC and subsequent clonal expansion, it does not only give rise to effector T cells but also to memory T cells, all expressing the same TCR ^{49,50}. These memory T cells are generally long-lasting and more protected from reinfection ⁵¹. With memory T cells circulating across lymphoid and normal tissues, an immune response can be more readily mounted upon subsequent exposures to a previously encountered pathogen ⁵². Effector T cells, on the other hand, rapidly undergo apoptosis upon antigen clearance because of a lack of stimulatory survival signals ⁴⁰.

To avoid the overshooting of an immune response to a foreign antigen, T cells have suppressive checkpoints that allow them to be tuned down from their effector state ^{36,53}. This happens by competition of checkpoint receptors with stimulatory co-receptors, such as CD28, for ligand binding, while also exhibiting a higher affinity for their common target ligand ⁵⁴. These checkpoints include receptors such as cytotoxic T lymphocyte associated protein 4 (CTLA-4) and programmed cell death 1 (PD1), which we will discuss in more detail later.

1.2.3 Cancer-immune interaction

As described above for viral peptides, cancer cells also generate non-self-peptides induced during oncogenesis. In this process, tumor suppressor genes can accumulate loss-of-function mutations, and oncogenes can accumulate gain-of-function mutations. Other possibilities include I) Demethylation of genes that are usually only expressed during fetal development, exposing novel antigens that TCRs were not negatively selected against during thymic development, and II) Increase of random mutations due to loss of DNA damage repair mechanisms, leading to additional new non-self-peptides, so-called neo-antigens ⁵⁵⁻⁵⁸. With the complex combination of cellular mutations leading to cancerous cells, shifting them from benign to malignant growth, the immune system and cancer interaction can be divided into three phases: Elimination, equilibrium, and escape ^{59,60}.

In the elimination phase, the immune system recognizes the foreign antigens expressed by self-tissue and destroys such oncogenically transformed cells.

The equilibrium phase is shaped by cancer immunoediting, in which the immune system deletes transformed cancer cells but simultaneously, similar to evolutionary theory, selects cancer cell variants for survival.

This ultimately leads to the escape phase, in which the selected cancer cells gain the ability to escape immune control.

Along these three phases, the tumor microenvironment (TME) gets established, consisting of various cell types which are mainly fibroblasts and innate and adaptive immune cells. The TME gains more immunosuppressive features throughout cancer development and disrupts the anti-tumor immune response⁶¹⁻⁶³.

1.2.4 Tumor infiltrating lymphocytes

T cells relocating to the tumor, searching for cognate peptide-HLAs and supported by inflammatory cytokines and chemokines, are termed tumor infiltrating lymphocytes (TILs). Once they identify their cognate target, they rapidly differentiate into cytotoxic T cells (mainly for CD8⁺ T cells), killing the targeted tumor cell and further contributing to the inflammation and recruitment of other components of the immune system. Depending on whether the number of TILs is high or low, tumor subtypes have been termed hot or cold tumors, respectively. The number of TILs has also been correlated with better patient survival, further indicating their importance for anti-tumor protection and tumor regression⁶⁴⁻⁶⁶.

Unfortunately, in late-stage cancers, TILs often fail to elicit strong enough responses to resolve the tumor burden fully. This has been attributed to various reasons; here, we will highlight the two most common ones. Firstly, T cells that are continuously exposed to a high antigen load can become overstimulated. Upon repeated stimulation, a T cell differentiates into a state of dysfunction or exhaustion. While a similar process happens in chronic viral infection due to continued exposure to foreign antigens, the immunosuppressive environment accelerates and strengthens this process. Immune checkpoint receptors such as CTLA-4 and PD-1 get increasingly expressed throughout T cell activation and exhaustion⁶⁷. These receptors are targeted by cancer cells and cells in the TME to increase the T cell's exhaustion level further. This renders the T cells hypo-responsive, leading to the inability to exhibit any proliferative or

functional capacity, even when faced with their cognate peptide-HLA target ⁶⁸. Secondly, cancer immunoediting and evolution lead to the evolutionary selection of cancer cell clones that have not been killed by cognate T cells, among other reasons, by not displaying antigens that can be recognised by TILs or by having reduced expression of HLA molecules ^{69,70}.

1.2.5 Cancer immunotherapy

This chapter is an author-produced adaptation of a research grant proposal submitted and approved for funding from PHRT.

The emergence and clinical success of immunotherapy over the past decades have established it as the fourth pillar of cancer therapy, in addition to surgery, radiotherapy, and chemotherapy ⁷¹. This has been possible through basic and clinical research that revealed the role of the immune system on cancer development and regression ^{55,72,73}. One of the most important discoveries was the ability of T cells to recognise and respond to tumor antigens that include over-expressed and/or mutated self-antigens ⁵⁵. Harnessing the specificity and cytotoxicity of T cells against tumor-antigens represents the fundamental mode of action used for a diverse set of immunotherapies.

As introduced earlier, immune checkpoints strongly contribute to immunosuppression and lead to dysfunctional TILs, rendering them unresponsive to the tumor. Immune checkpoint blockade (ICB) therapy is blocking the immune checkpoint molecules expressed on T cells by using antibodies, preventing or reversing exhaustion. A recent study showed that tumor mutational burden (number of mutations per megabase in the cancer genome), i.e., melanoma having the highest, is correlated with response to ICB, highlighting the importance of tumor-reactive T cell functionality in inducing cancer regression ^{74,75}. Moreover, personalized cellular immunotherapies that consist of autologous adoptive cell transfer (ACT) of T cells, such as chimeric antigen receptor T cells (CAR-T), tumor-infiltrating lymphocytes (TILs) and T cell receptor-T cells (TCR-T) have all shown promising results with several therapies in late-stage clinical trials or clinically approved ⁷⁶⁻⁷⁸ (**Figure 3**).

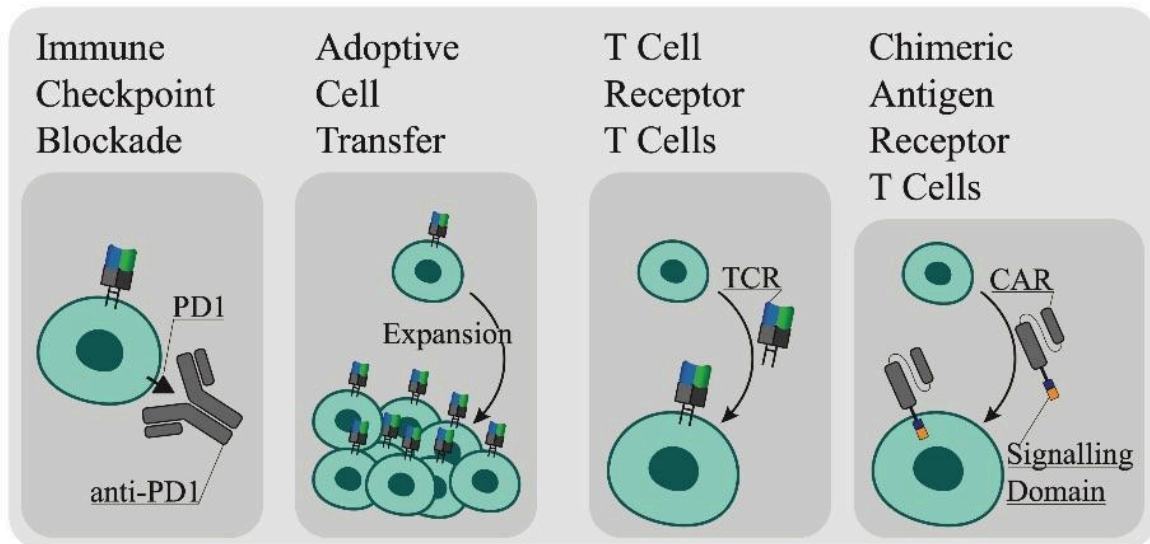


Figure 3: Schematic of the four types of cancer immunotherapies. From left to right: ICB, ACT of TILs, TCR-T and CAR-T therapy. ACT of TILs, TCR-T and CAR-T therapy are getting infused into the patient upon generation of the cell product through expansion and/or genomic knock-in of the TCR or CAR of interest.

Despite its recent therapeutic success and remarkable activity in a subset of patients with advanced cancers, ICB fails to induce durable responses in the majority of treated patients ^{79,80}. ACT of TILs is one of the most effective personalized immunotherapies against solid tumors to date and is based on the extraction of TILs from a tumor biopsy, *ex vivo* rapid expansion, and reinfusion of a large T cell population (ca. 10^9 - 10^{11} cells) back into the patient ^{81,82}. In a recent clinical trial in advanced-stage melanoma patients who were ICB (anti-PD1) treatment-refractory, ACT of TILs was shown to be highly effective and has led to a median increase of 4.1 months of progression-free survival, representing more than double the progression-free survival compared to the alternative treatment (anti-CTLA4 ICB) ⁸³. By co-culturing extracted TILs with autologous tumor cells and co-stimulatory factors such as IL-2, it is attempted to ensure the tumor-specificity of TILs, which subsequently get rapidly expanded. However, in the current setting, while some TILs get activated because of their specificity to the tumor cells, others get activated through nonspecific stimuli, possibly dependent on the current cell state and suitability to *in vitro* culturing conditions. Because tumor regression is associated

with the number of infused tumor-reactive T cells, it is an active area of research to increase the number of tumor-reactive TILs in the reinfusion product ^{84,85}.

Another potential breakthrough in personalized immunotherapy is the development of TCR-T therapies, another form of ACT that relies on viral transduction of autologous patient T cells with constructs encoding tumor-reactive TCRs, followed by their expansion and reinfusion (a similar process is used for CAR-T therapy). Since TCRs can recognize intracellular antigen targets in the form of peptides presented on a HLA, they are an attractive modality for treating solid tumors, in which tumor-specific surface antigens have been difficult to identify or patient T cells exhibit no native tumor-reactivity. Ultimately, this increases the number of tumor-targeting T cells, while allowing for more efficient infiltration into solid tumors, an area in which CAR-T cells have shown underwhelming efficacy ^{86,87}. Small trials of TCR-T therapies in solid tumors, targeting the tumor associated antigens NY-ESO-1 (melanoma, synovial sarcoma, and myeloma) and, more recently, MAGE-A4 (NSCLC, head and neck cancer, melanoma, synovial sarcoma, and others) have shown remarkable response rates in patients with advanced refractory cancers ^{88,89}. Advances in single-cell sorting and sequencing technologies enable more efficient identification of tumor-reactive TCRs. Accordingly, several recent studies have exploited these technologies to demonstrate the feasibility of identifying candidates for personalized TCR-T therapies ⁹⁰⁻⁹².

1.3 Single-cell sequencing for T cell analysis

Analyzing a cell's whole mRNA transcriptome gives insights into its cellular state. This can be done by first capturing the mRNA and subsequently performing reverse transcription (RT) and amplification polymerase chain reaction (PCR) of the resulting complementary DNA (cDNA). In the past, measurements of whole transcriptomes of cells have been done from bulk populations of cells, meaning that all cellular transcripts were analyzed together, resulting in the average expression across the bulk population, limiting insights of specific cell types and cell state dynamics. A decade ago, single-cell sequencing was awarded *Method of the Year* ⁹³ after being capable of analyzing small numbers of whole transcriptomes of single cells (scRNA) to profile cellular subpopulations ^{94,95}. Another breakthrough in 2015 led to the development of

high-throughput sequencing of single cells using droplets ⁹⁶, which opened up possibilities of extensive and affordable sequencing experiments (**Figure 4**).

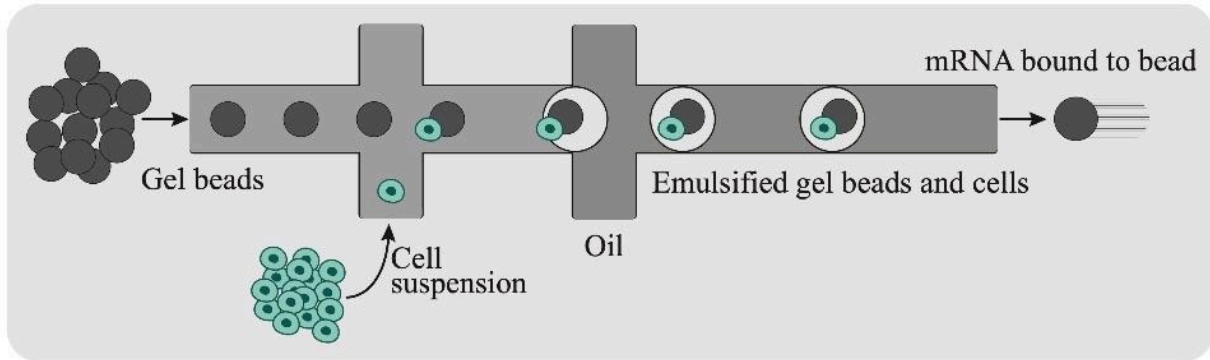


Figure 4: Schematic of single-cell droplet generation. Gel beads and cells form water-in-oil droplets. Upon cell lysis, cellular mRNA gets bound to the gel bead through oligonucleotide adaptors. Subsequently RT-PCR and cDNA amplification can be performed.

Given the high complexity of single-cell transcriptomic data, a substantial amount of research in the past decade has concentrated on the precise analysis and interpretation of insights derived from such rich datasets. To correctly analyze a large number of genes, cells and potentially multiple replicates or samples, tools for quality control, normalization, and batch effect removal have been developed ⁹⁷⁻⁹⁹. To facilitate the clustering and dimensionality reduction of high dimensional single-cell data for identifying cell populations or differences between samples, algorithms such as principal component analysis ¹⁰⁰ and uniform manifold approximation and projection (UMAP) ¹⁰¹ were developed. At this point, scRNA sequencing has been extensively employed to profile the gene-expression signatures of T cells. This approach has been instrumental in identifying novel cellular subsets and phenotypes, as well as in understanding their dynamics during vaccination, infection, and cancer ¹⁰²⁻¹⁰⁵.

Furthermore, the analysis of adaptive immune cells, specifically T cells, experienced significant advancements through the development of high-throughput scRNA sequencing methods due to the ability to link the transcriptome of a cell to the natively paired TCR α - and β -chains. This was not previously possible on this scale with the standard bulk sequencing methods of lymphocytes because the two receptor chains are expressed as unique transcripts from separate chromosomes. Coupling the TCR sequence to the transcriptome within an individual cell enables the phenotypic

analyses of a clonal population of T lymphocytes, allowing effective barcoding of each cell
106–108

1.4 Linking functional reactivity to cell state

Parts of this chapter are an author-produced adaptation of a research grant proposal submitted and approved for funding from PHRT.

Testing T cell reactivity through functional co-culture assays allows the identification of TCR-peptide-HLA pairs. The co-culture setup mainly consists of T cells of interest, APCs that express patient autologous HLAs, and the target antigen of interest. After co-culture, the activation markers in T cells can be detected using methods such as enzyme-linked immunosorbent assay or fluorescence-activated cell sorting (FACS). By performing scRNA sequencing of a subset of the T cells before co-culture, their cell state can be assessed without co-culture-induced changes. Post-co-culture functional reactivity results and cell state can be linked to allow for the correlation of a T cell's cell state and its reactivity. Due to the limited knowledge of a T cell's target peptides, functional reactivity testing is often laborious and challenging. Especially for human tumors, there are only a few known public antigen targets of T cells^{56,109–111}. Typically, whole exome sequencing of the tumor cells is performed to identify mutations that can give rise to neo-antigens. However, only 0.2% of tested tumor-specific antigens could elicit a response in TILs, with a large fraction recognising viral antigens, indicating that a significant portion of TILs are bystanders¹¹². While for melanoma, it is possible to enrich tumor-targeting T cells based on the clonal size in the TIL repertoire, this seems to be not applicable to other cancer types¹¹³. Enhancing the number of tumor-reactive T cells that can infiltrate tumors could significantly boost the efficacy of personalized immunotherapies⁸⁴. Thus, identification of markers for the rapid selection of tumor-reactive T cells could be useful for making treatment decisions and bringing improvements to current therapies. As mentioned earlier, continuous antigen exposure and recognition within the TME can be a hallmark and driver for T cell exhaustion. Thus, the presence of exhausted or dysfunctional T cells in the TME is often used as a proxy for the abundance of tumor-specific T cells, with studies concluding that

the abundance of dysfunctional TILs is correlated with the positive response to ICB ¹¹⁴⁻¹¹⁶. Further, studies have shown that surface markers of exhaustion, such as Ectonucleoside triphosphate diphosphohydrolase 1 (ENTPD1, also often referred to as CD39), can potentially enrich tumor-reactive T cells ^{112,117}.

Recent work evaluated patient peripheral blood mononuclear cell (PBMC)-derived T cell and TIL tumor-reactivity by screening against lung cancer-derived neo-antigens and identified tumor-reactive T cells in anti-PD1 treated patients ¹¹⁸. By linking these reactivities to TIL transcriptomes derived from scRNA sequencing data, tumor-reactivity marker genes such as ENTPD1 and C-X-C motif chemokine 13 (CXCL13), mainly expressed in resident memory and exhausted T cells, could be identified ^{119,120}. Another study from the Rosenberg lab generated a neo-antigen reactivity score that enabled the enrichment of tumor-reactive TILs in metastatic cancer samples ¹²¹.

1.5 References of Chapter 1

1. Hammarlund, E., Lewis, M. W., Hansen, S. G., Strelow, L. I., Nelson, J. A., Sexton, G. J., Hanifin, J. M. & Slifka, M. K. Duration of antiviral immunity after smallpox vaccination. *Nat. Med.* **9**, 1131–1137 (2003).
2. Davis, M. M. & Bjorkman, P. J. T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395–402 (1988).
3. Murphy, K. & Weaver, C. *Janeway's Immunobiology*. (Garland Science, 2016).
4. Chang, Y. M., Wieland, A., Li, Z.-R., Im, S. J., McGuire, D. J., Kissick, H. T., Antia, R. & Ahmed, R. T Cell Receptor Diversity and Lineage Relationship between Virus-Specific CD8 T Cell Subsets during Chronic Lymphocytic Choriomeningitis Virus Infection. *J. Virol.* **94**, (2020).
5. Valpione, S., Mundra, P. A., Galvani, E., Campana, L. G., Lorigan, P., De Rosa, F., Gupta, A., Weightman, J., Mills, S., Dhomen, N. & Marais, R. The T cell receptor repertoire of tumor infiltrating T cells is predictive and prognostic for cancer survival. *Nat. Commun.* **12**, 4098 (2021).
6. Sun, X., Nguyen, T., Achour, A., Ko, A., Cifello, J., Ling, C., Sharma, J., Hiroi, T., Zhang, Y., Chia, C. W., Wood, W., 3rd, Wu, W. W., Zukley, L., Phue, J.-N., Becker, K. G., Shen, R.-F., Ferrucci, L. & Weng, N.-P. Longitudinal analysis reveals age-related changes in the T cell receptor repertoire of human T cell subsets. *J. Clin. Invest.* **132**, (2022).
7. Falk, K., Rötzschke, O., Stevanović, S., Jung, G. & Rammensee, H. G. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **351**, 290–296 (1991).
8. Mungall, A. J., Palmer, S. A., Sims, S. K., Edwards, C. A., Ashurst, J. L., Wilming, L., Jones, M. C., Horton, R., Hunt, S. E., Scott, C. E., Gilbert, J. G. R., Clamp, M. E., Bethel, G., Milne, S., Ainscough, R., Almeida, J. P., Ambrose, K. D., Andrews, T. D., Ashwell, R. I. S., Babbage, A. K., Bagguley, C. L., Bailey, J., Banerjee, R., Barker, D. J., Barlow, K. F., Bates, K., Beare, D. M., Beasley, H., Beasley, O., Bird, C. P., Blakey, S., Bray-Allen, S., Brook, J., Brown, A. J., Brown, J. Y., Burford, D. C., Burrill, W., Burton, J., Carder, C., Carter, N. P., Chapman, J. C., Clark, S. Y., Clark, G., Clee, C. M., Clegg, S., Cobley, V., Collier, R. E., Collins, J. E., Colman, L. K., Corby, N. R., Coville, G. J., Culley, K. M., Dhami, P., Davies, J., Dunn, M., Earthrowl, M. E., Ellington, A. E., Evans, K. A., Faulkner, L., Francis, M. D., Frankish, A., Frankland, J., French, L., Garner, P., Garnett, J., Ghori, M. J. R., Gilby, L. M., Gillson, C. J., Glithero, R. J., Grafham, D. V., Grant, M., Gribble, S., Griffiths, C., Griffiths, M., Hall, R., Halls, K. S., Hammond, S., Harley, J. L., Hart, E. A., Heath, P. D., Heathcote, R., Holmes, S. J., Howden, P. J., Howe, K. L., Howell, G. R., Huckle, E., Humphray, S. J., Humphries, M. D., Hunt, A. R., Johnson, C. M., Joy, A. A., Kay, M., Keenan, S. J., Kimberley, A. M., King, A., Laird, G. K., Langford, C., Lawlor, S., Leongamornlert, D. A., Leversha, M., Lloyd, C. R., Lloyd, D. M., Loveland, J. E., Lovell, J., Martin, S., Mashreghi-Mohammadi, M., Maslen, G. L., Matthews, L., McCann, O. T., McLaren, S. J., McLay, K., McMurray, A., Moore, M. J. F., Mullikin, J. C., Niblett, D., Nickerson, T., Novik, K. L., Oliver, K., Overton-Larty, E. K., Parker, A., Patel, R., Pearce, A. V., Peck, A. I., Phillimore, B., Phillips, S., Plumb, R. W., Porter, K. M., Ramsey, Y., Ranby, S. A., Rice, C. M., Ross, M. T., Searle, S. M., Sehra, H. K., Sheridan, E., Skuce, C. D., Smith, S., Smith, M., Spraggon, L., Squares, S. L., Steward, C. A., Sycamore, N., Tamlyn-Hall, G., Tester, J., Theaker, A. J., Thomas, D. W., Thorpe, A., Tracey, A., Tromans, A., Tubby, B., Wall, M., Wallis, J. M., West, A. P., White, S. S., Whitehead, S. L., Whittaker, H., Wild, A., Willey, D. J., Wilmer, T. E., Wood, J. M., Wray, P. W., Wyatt, J. C., Young, L., Younger, R. M., Bentley, D. R., Coulson, A., Durbin, R., Hubbard, T., Sulston, J. E., Dunham, I., Rogers, J. & Beck, S. The DNA sequence and analysis of human chromosome 6. *Nature* **425**, 805–811 (2003).
9. Feng, D., Bond, C. J., Ely, L. K., Maynard, J. & Garcia, K. C. Structural evidence for a germline-encoded T cell receptor-major histocompatibility complex interaction ‘codon’. *Nat. Immunol.* **8**, 975–983 (2007).
10. La Gruta, N. L., Gras, S., Daley, S. R., Thomas, P. G. & Rossjohn, J. Understanding the drivers of MHC restriction of T cell receptors. *Nat. Rev. Immunol.* **18**, 467–478 (2018).
11. Rossjohn, J., Gras, S., Miles, J. J., Turner, S. J., Godfrey, D. I. & McCluskey, J. T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* **33**, 169–200 (2015).
12. Garcia, K. C. & Adams, E. J. How the T cell receptor sees antigen—a structural view. *Cell* **122**, 333–336 (2005).
13. Buslepp, J., Wang, H., Biddison, W. E., Appella, E. & Collins, E. J. A correlation between TCR Valpha docking on MHC and CD8 dependence: implications for T cell selection. *Immunity* **19**, 595–606 (2003).
14. Wang, R., Natarajan, K. & Margulies, D. H. Structural basis of the CD8 alpha beta/MHC class I interaction: focused recognition orients CD8 beta to a T cell proximal position. *J. Immunol.* **183**, 2554–2564 (2009).
15. Zamoyska, R. CD4 and CD8: modulators of T-cell receptor recognition of antigen and of immune responses?

- Curr. Opin. Immunol.* **10**, 82–87 (1998).
16. Sette, A., Adorini, L., Colon, S. M., Buus, S. & Grey, H. M. Capacity of intact proteins to bind to MHC class II molecules. *J. Immunol.* **143**, 1265–1267 (1989).
 17. Yewdell, J. W. & Bennink, J. R. Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Annu. Rev. Immunol.* **17**, 51–88 (1999).
 18. Trolle, T., McMurtrey, C. P., Sidney, J., Bardet, W., Osborn, S. C., Kaever, T., Sette, A., Hildebrand, W. H., Nielsen, M. & Peters, B. The Length Distribution of Class I-Restricted T Cell Epitopes Is Determined by Both Peptide Supply and MHC Allele-Specific Binding Preference. *J. Immunol.* **196**, 1480–1487 (2016).
 19. Kisielow, P., Teh, H. S., Blüthmann, H. & von Boehmer, H. Positive selection of antigen-specific T cells in thymus by restricting MHC molecules. *Nature* **335**, 730–733 (1988).
 20. Anderson, M. S., Venanzi, E. S., Klein, L., Chen, Z., Berzins, S. P., Turley, S. J., von Boehmer, H., Bronson, R., Dierich, A., Benoist, C. & Mathis, D. Projection of an immunological self shadow within the thymus by the aire protein. *Science* **298**, 1395–1401 (2002).
 21. Ashton-Rickardt, P. G., Bandeira, A., Delaney, J. R., Van Kaer, L., Pircher, H. P., Zinkernagel, R. M. & Tonegawa, S. Evidence for a differential avidity model of T cell selection in the thymus. *Cell* **76**, 651–663 (1994).
 22. Zerrahn, J., Held, W. & Raulet, D. H. The MHC reactivity of the T cell repertoire prior to positive and negative selection. *Cell* **88**, 627–636 (1997).
 23. Basler, M., Kirk, C. J. & Groettrup, M. The immunoproteasome in antigen processing and other immunological functions. *Curr. Opin. Immunol.* **25**, 74–80 (2013).
 24. Grommé, M. & Neefjes, J. Antigen degradation or presentation by MHC class I molecules via classical and non-classical pathways. *Mol. Immunol.* **39**, 181–202 (2002).
 25. Schubert, U., Antón, L. C., Gibbs, J., Norbury, C. C., Yewdell, J. W. & Bennink, J. R. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* **404**, 770–774 (2000).
 26. Clark, R. E., Dodi, I. A., Hill, S. C., Lill, J. R., Aubert, G., Macintyre, A. R., Rojas, J., Bourdon, A., Bonner, P. L., Wang, L., Christmas, S. E., Travers, P. J., Creaser, C. S., Rees, R. C. & Madrigal, J. A. Direct evidence that leukemic cells present HLA-associated immunogenic peptides derived from the BCR-ABL b3a2 fusion protein. *Blood* **98**, 2887–2893 (2001).
 27. Madden, D. R., Garboczi, D. N. & Wiley, D. C. The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2. *Cell* **75**, 693–708 (1993).
 28. Schlienger, K., Craighead, N., Lee, K. P., Levine, B. L. & June, C. H. Efficient priming of protein antigen-specific human CD4(+) T cells by monocyte-derived dendritic cells. *Blood* **96**, 3490–3498 (2000).
 29. Guermónprez, P., Valladeau, J., Zitvogel, L., Théry, C. & Amigorena, S. Antigen presentation and T cell stimulation by dendritic cells. *Annu. Rev. Immunol.* **20**, 621–667 (2002).
 30. Jenkins, M. K. & Johnson, J. G. Molecules involved in T-cell costimulation. *Curr. Opin. Immunol.* **5**, 361–367 (1993).
 31. Curtsinger, J. M., Lins, D. C. & Mescher, M. F. Signal 3 determines tolerance versus full activation of naive CD8 T cells: dissociating proliferation and development of effector function. *J. Exp. Med.* **197**, 1141–1151 (2003).
 32. Jung, S., Unutmaz, D., Wong, P., Sano, G.-I., De los Santos, K., Sparwasser, T., Wu, S., Vuthoori, S., Ko, K., Zavala, F., Pamer, E. G., Littman, D. R. & Lang, R. A. In vivo depletion of CD11c+ dendritic cells abrogates priming of CD8+ T cells by exogenous cell-associated antigens. *Immunity* **17**, 211–220 (2002).
 33. Call, M. E., Pyrdol, J., Wiedmann, M. & Wucherpfennig, K. W. The organizing principle in the formation of the T cell receptor-CD3 complex. *Cell* **111**, 967–979 (2002).
 34. Klausner, R. D. & Samelson, L. E. T cell antigen receptor activation pathways: the tyrosine kinase connection. *Cell* **64**, 875–878 (1991).
 35. Iwashima, M., Irving, B. A., van Oers, N. S., Chan, A. C. & Weiss, A. Sequential interactions of the TCR with two distinct cytoplasmic tyrosine kinases. *Science* **263**, 1136–1139 (1994).
 36. Chen, L. & Flies, D. B. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nat. Rev. Immunol.* **13**, 227–242 (2013).
 37. Acuto, O. & Michel, F. CD28-mediated co-stimulation: a quantitative support for TCR signalling. *Nat. Rev. Immunol.* **3**, 939–951 (2003).
 38. Cebrián, M., Redondo, J. M., López-Rivas, A., Rodríguez-Tarduchy, G., De Landázuri, M. O. & Sánchez-Madrid, F. Expression and function of AIM, an activation inducer molecule of human lymphocytes, is dependent on the activation of protein kinase C. *Eur. J. Immunol.* **19**, 809–815 (1989).
 39. Wong, P. & Pamer, E. G. CD8 T cell responses to infectious pathogens. *Annu. Rev. Immunol.* **21**, 29–70 (2003).

40. Wherry, E. J. & Ahmed, R. Memory CD8 T-cell differentiation during viral infection. *J. Virol.* **78**, 5535–5545 (2004).
41. Jenkins, M. K., Khoruts, A., Ingulli, E., Mueller, D. L., McSorley, S. J., Reinhardt, R. L., Itano, A. & Pape, K. A. In vivo activation of antigen-specific CD4 T cells. *Annu. Rev. Immunol.* **19**, 23–45 (2001).
42. Schweitzer, A. N. & Sharpe, A. H. Studies using antigen-presenting cells lacking expression of both B7-1 (CD80) and B7-2 (CD86) show distinct requirements for B7 molecules during priming versus restimulation of Th2 but not Th1 cytokine production. *J. Immunol.* **161**, 2762–2771 (1998).
43. Gudmundsdottir, H., Wells, A. D. & Turka, L. A. Dynamics and requirements of T cell clonal expansion in vivo at the single-cell level: effector function is linked to proliferative capacity. *J. Immunol.* **162**, 5212–5223 (1999).
44. Matloubian, M., Lo, C. G., Cinamon, G., Lesneski, M. J., Xu, Y., Brinkmann, V., Allende, M. L., Proia, R. L. & Cyster, J. G. Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on S1P receptor 1. *Nature* **427**, 355–360 (2004).
45. Schwab, S. R. & Cyster, J. G. Finding a way out: lymphocyte egress from lymphoid organs. *Nat. Immunol.* **8**, 1295–1301 (2007).
46. Bossi, G., Trambas, C., Booth, S., Clark, R., Stinchcombe, J. & Griffiths, G. M. The secretory synapse: the secrets of a serial killer. *Immunol. Rev.* **189**, 152–160 (2002).
47. Dobrzanski, M. J., Reome, J. B., Hollenbaugh, J. A. & Dutton, R. W. Tc1 and Tc2 effector cell therapy elicit long-term tumor immunity by contrasting mechanisms that result in complementary endogenous type 1 antitumor responses. *J. Immunol.* **172**, 1380–1390 (2004).
48. Zhou, F. Molecular mechanisms of IFN-gamma to up-regulate MHC class I antigen processing and presentation. *Int. Rev. Immunol.* **28**, 239–260 (2009).
49. Butz, E. A. & Bevan, M. J. Massive expansion of antigen-specific CD8+ T cells during an acute virus infection. *Immunity* **8**, 167–175 (1998).
50. Lanzavecchia, A. & Sallusto, F. Understanding the generation and function of memory T cell subsets. *Curr. Opin. Immunol.* **17**, 326–332 (2005).
51. Baumann, C., Fröhlich, A., Brunner, T. M., Holeccka, V., Pinschewer, D. D. & Löhning, M. Memory CD8 T Cell Protection From Viral Reinfection Depends on Interleukin-33 Alarmin Signals. *Front. Immunol.* **10**, 1833 (2019).
52. Woodland, D. L. & Kohlmeier, J. E. Migration, maintenance and recall of memory T cells in peripheral tissues. *Nat. Rev. Immunol.* **9**, 153–161 (2009).
53. Chambers, C. A., Kuhns, M. S., Egen, J. G. & Allison, J. P. CTLA-4-mediated inhibition in regulation of T cell responses: mechanisms and manipulation in tumor immunotherapy. *Annu. Rev. Immunol.* **19**, 565–594 (2001).
54. Wing, K., Onishi, Y., Prieto-Martin, P., Yamaguchi, T., Miyara, M., Fehervari, Z., Nomura, T. & Sakaguchi, S. CTLA-4 control over Foxp3+ regulatory T cell function. *Science* **322**, 271–275 (2008).
55. van der Bruggen, P., Traversari, C., Chomez, P., Lurquin, C., De Plaen, E., Van den Eynde, B., Knuth, A. & Boon, T. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science* **254**, 1643–1647 (1991).
56. Schumacher, T. N., Scheper, W. & Kvistborg, P. Cancer Neoantigens. *Annu. Rev. Immunol.* **37**, 173–200 (2019).
57. Caballero, O. L. & Chen, Y.-T. Cancer/testis (CT) antigens: potential targets for immunotherapy. *Cancer Sci.* **100**, 2014–2021 (2009).
58. Boyse, E. A., Old, L. J., Stockert, E. & Shigeno, N. Genetic origin of tumor antigens. *Cancer Res.* **28**, 1280–1287 (1968).
59. Quezada, S. A., Peggs, K. S., Simpson, T. R. & Allison, J. P. Shifting the equilibrium in cancer immunoediting: from tumor tolerance to eradication. *Immunol. Rev.* **241**, 104–118 (2011).
60. Dunn, G. P., Old, L. J. & Schreiber, R. D. The three Es of cancer immunoediting. *Annu. Rev. Immunol.* **22**, 329–360 (2004).
61. Zou, W. Immunosuppressive networks in the tumour environment and their therapeutic relevance. *Nat. Rev. Cancer* **5**, 263–274 (2005).
62. Whiteside, T. L. The tumor microenvironment and its role in promoting tumor growth. *Oncogene* **27**, 5904–5912 (2008).
63. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
64. Fridman, W. H., Pagès, F., Sautès-Fridman, C. & Galon, J. The immune contexture in human tumours: impact on clinical outcome. *Nat. Rev. Cancer* **12**, 298–306 (2012).
65. Gooden, M. J. M., de Bock, G. H., Leffers, N., Daemen, T. & Nijman, H. W. The prognostic influence of tumour-infiltrating lymphocytes in cancer: a systematic review with meta-analysis. *Br. J. Cancer* **105**, 93–103

- (2011).
66. Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., Tosolini, M., Camus, M., Berger, A., Wind, P., Zinzindohoué, F., Bruneval, P., Cugnenc, P.-H., Trajanoski, Z., Fridman, W.-H. & Pagès, F. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* **313**, 1960–1964 (2006).
 67. Wherry, E. J. T cell exhaustion. *Nat. Immunol.* **12**, 492–499 (2011).
 68. Schietinger, A., Philip, M., Krisnawan, V. E., Chiu, E. Y., Delrow, J. J., Basom, R. S., Lauer, P., Brockstedt, D. G., Knoblaugh, S. E., Hämmerling, G. J., Schell, T. D., Garbi, N. & Greenberg, P. D. Tumor-Specific T Cell Dysfunction Is a Dynamic Antigen-Driven Differentiation Program Initiated Early during Tumorigenesis. *Immunity* **45**, 389–401 (2016).
 69. Matsushita, H., Vesely, M. D., Koblodt, D. C., Rickert, C. G., Uppaluri, R., Magrini, V. J., Arthur, C. D., White, J. M., Chen, Y.-S., Shea, L. K., Hundal, J., Wendl, M. C., Demeter, R., Wylie, T., Allison, J. P., Smyth, M. J., Old, L. J., Mardis, E. R. & Schreiber, R. D. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* **482**, 400–404 (2012).
 70. Restifo, N. P., Esquivel, F., Asher, A. L., Stötter, H., Barth, R. J., Bennink, J. R., Mulé, J. J., Yewdell, J. W. & Rosenberg, S. A. Defective presentation of endogenous antigens by a murine sarcoma. Implications for the failure of an anti-tumor immune response. *J. Immunol.* **147**, 1453–1459 (1991).
 71. McCune, J. S. Rapid Advances in Immunotherapy to Treat Cancer. *Clin. Pharmacol. Ther.* **103**, 540–544 (2018).
 72. Kaplan, D. H., Shankaran, V., Dighe, A. S., Stockert, E., Aguet, M., Old, L. J. & Schreiber, R. D. Demonstration of an interferon gamma-dependent tumor surveillance system in immunocompetent mice. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 7556–7561 (1998).
 73. Hill, W., Lim, E. L., Weeden, C. E., Lee, C., Augustine, M., Chen, K., Kuan, F.-C., Marongiu, F., Evans, E. J., Jr, Moore, D. A., Rodrigues, F. S., Pich, O., Bakker, B., Cha, H., Myers, R., van Maldegem, F., Boumelha, J., Veeriah, S., Rowan, A., Naceur-Lombardelli, C., Karasaki, T., Sivakumar, M., De, S., Caswell, D. R., Nagano, A., Black, J. R. M., Martínez-Ruiz, C., Ryu, M. H., Huff, R. D., Li, S., Favé, M.-J., Magness, A., Suárez-Bonnet, A., Priestnall, S. L., Lüchtenborg, M., Lavelle, K., Pethick, J., Hardy, S., McDonald, F. E., Lin, M.-H., Troccoli, C. I., Ghosh, M., Miller, Y. E., Merrick, D. T., Keith, R. L., Al Bakir, M., Bailey, C., Hill, M. S., Saal, L. H., Chen, Y., George, A. M., Abbosh, C., Kanu, N., Lee, S.-H., McGranahan, N., Berg, C. D., Sasieni, P., Houlston, R., Turnbull, C., Lam, S., Awadalla, P., Grönroos, E., Downward, J., Jacks, T., Carlsten, C., Malanchi, I., Hackshaw, A., Litchfield, K., TRACERx Consortium, DeGregori, J., Jamal-Hanjani, M. & Swanton, C. Lung adenocarcinoma promotion by air pollutants. *Nature* **616**, 159–167 (2023).
 74. Couzin-Frankel, J. Breakthrough of the year 2013. Cancer immunotherapy. *Science* **342**, 1432–1433 (2013).
 75. Quezada, S. A. & Peggs, K. S. Exploiting CTLA-4, PD-1 and PD-L1 to reactivate the host immune response against cancer. *Br. J. Cancer* **108**, 1560–1565 (2013).
 76. Dudley, M. E., Wunderlich, J. R., Robbins, P. F., Yang, J. C., Hwu, P., Schwartzentruber, D. J., Topalian, S. L., Sherry, R., Restifo, N. P., Hubicki, A. M., Robinson, M. R., Raffeld, M., Duray, P., Seipp, C. A., Rogers-Freezer, L., Morton, K. E., Mavroukakis, S. A., White, D. E. & Rosenberg, S. A. Cancer regression and autoimmunity in patients after clonal repopulation with antitumor lymphocytes. *Science* **298**, 850–854 (2002).
 77. Maude, S. L., Frey, N., Shaw, P. A., Aplenc, R., Barrett, D. M., Bunin, N. J., Chew, A., Gonzalez, V. E., Zheng, Z., Lacey, S. F., Mahnke, Y. D., Melenhorst, J. J., Rheingold, S. R., Shen, A., Teachey, D. T., Levine, B. L., June, C. H., Porter, D. L. & Grupp, S. A. Chimeric antigen receptor T cells for sustained remissions in leukemia. *N. Engl. J. Med.* **371**, 1507–1517 (2014).
 78. Strønen, E., Toebes, M., Kelderman, S., van Buuren, M. M., Yang, W., van Rooij, N., Donia, M., Bösch, M.-L., Lund-Johansen, F., Olweus, J. & Schumacher, T. N. Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* **352**, 1337–1341 (2016).
 79. Antonia, S. J., Borghaei, H., Ramalingam, S. S., Horn, L., De Castro Carpeño, J., Pluzanski, A., Burgio, M. A., Garassino, M., Chow, L. Q. M., Gettinger, S., Crinò, L., Planchard, D., Butts, C., Drilon, A., Wojcik-Tomaszewska, J., Otterson, G. A., Agrawal, S., Li, A., Penrod, J. R. & Brahmer, J. Four-year survival with nivolumab in patients with previously treated advanced non-small-cell lung cancer: a pooled analysis. *Lancet Oncol.* **20**, 1395–1408 (2019).
 80. Hiltbrunner, S., Cords, L., Kasser, S., Freiburger, S. N., Kreutzer, S., Toussaint, N. C., Grob, L., Opitz, I., Messerli, M., Zoche, M., Soltermann, A., Rechsteiner, M., van den Broek, M., Bodenmiller, B. & Curioni-Fontecedro, A. Acquired resistance to anti-PD1 therapy in patients with NSCLC associates with immunosuppressive T cell phenotype. *Nat. Commun.* **14**, 5154 (2023).
 81. Delorme, E. J. & Alexander, P. TREATMENT OF PRIMARY FIBROSARCOMA IN THE RAT WITH

IMMUNE LYMPHOCYTES. *Lancet* **2**, 117–120 (1964).

82. Rosenberg, S. A. & Restifo, N. P. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science* **348**, 62–68 (2015).
83. Rohaan, M. W., Borch, T. H., van den Berg, J. H., Met, Ö., Kessels, R., Geukes Foppen, M. H., Stoltenberg Granhøj, J., Nuijen, B., Nijenhuis, C., Jedema, I., van Zon, M., Scheij, S., Beijnen, J. H., Hansen, M., Voermans, C., Noringriis, I. M., Monberg, T. J., Holmstroem, R. B., Wever, L. D. V., van Dijk, M., Griepink-Ongering, L. G., Valkenet, L. H. M., Torres Acosta, A., Karger, M., Borgers, J. S. W., Ten Ham, R. M. T., Retèl, V. P., van Harten, W. H., Lalezari, F., van Tinteren, H., van der Veldt, A. A. M., Hospers, G. A. P., Stevensen-den Boer, M. A. M., Suijkerbuijk, K. P. M., Aarts, M. J. B., Piersma, D., van den Eertwegh, A. J. M., de Groot, J.-W. B., Vreugdenhil, G., Kapiteijn, E., Boers-Sonderen, M. J., Fiets, W. E., van den Berkmortel, F. W. P. J., Ellebaek, E., Hölmich, L. R., van Akkooi, A. C. J., van Houdt, W. J., Wouters, M. W. J. M., van Thienen, J. V., Blank, C. U., Meerveld-Eggink, A., Klobuch, S., Wilgenhof, S., Schumacher, T. N., Donia, M., Svane, I. M. & Haanen, J. B. A. G. Tumor-Infiltrating Lymphocyte Therapy or Ipilimumab in Advanced Melanoma. *N. Engl. J. Med.* **387**, 2113–2125 (2022).
84. Andersen, R., Donia, M., Ellebaek, E., Borch, T. H., Kongsted, P., Iversen, T. Z., Hölmich, L. R., Hendel, H. W., Met, Ö., Andersen, M. H., Thor Straten, P. & Svane, I. M. Long-Lasting Complete Responses in Patients with Metastatic Melanoma after Adoptive Cell Therapy with Tumor-Infiltrating Lymphocytes and an Attenuated IL2 Regimen. *Clin. Cancer Res.* **22**, 3734–3745 (2016).
85. Bianchi, V., Harari, A. & Coukos, G. Neoantigen-Specific Adoptive Cell Therapies for Cancer: Making T-Cell Products More Personal. *Front. Immunol.* **11**, 1215 (2020).
86. Ahmed, N., Brawley, V. S., Hegde, M., Robertson, C., Ghazi, A., Gerken, C., Liu, E., Dakhova, O., Ashoori, A., Corder, A., Gray, T., Wu, M.-F., Liu, H., Hicks, J., Rainusso, N., Dotti, G., Mei, Z., Grilley, B., Gee, A., Rooney, C. M., Brenner, M. K., Heslop, H. E., Wels, W. S., Wang, L. L., Anderson, P. & Gottschalk, S. Human Epidermal Growth Factor Receptor 2 (HER2) -Specific Chimeric Antigen Receptor-Modified T Cells for the Immunotherapy of HER2-Positive Sarcoma. *J. Clin. Oncol.* **33**, 1688–1696 (2015).
87. Thistlethwaite, F. C., Gilham, D. E., Guest, R. D., Rothwell, D. G., Pillai, M., Burt, D. J., Byatte, A. J., Kirillova, N., Valle, J. W., Sharma, S. K., Chester, K. A., Westwood, N. B., Halford, S. E. R., Nabarro, S., Wan, S., Austin, E. & Hawkins, R. E. The clinical efficacy of first-generation carcinoembryonic antigen (CEACAM5)-specific CAR T cells is limited by poor persistence and transient pre-conditioning-dependent respiratory toxicity. *Cancer Immunol. Immunother.* **66**, 1425–1436 (2017).
88. Robbins, P. F., Kassim, S. H., Tran, T. L. N., Crystal, J. S., Morgan, R. A., Feldman, S. A., Yang, J. C., Dudley, M. E., Wunderlich, J. R., Sherry, R. M., Kammula, U. S., Hughes, M. S., Restifo, N. P., Raffeld, M., Lee, C.-C. R., Li, Y. F., El-Gamil, M. & Rosenberg, S. A. A pilot trial using lymphocytes genetically engineered with an NY-ESO-1-reactive T-cell receptor: long-term follow-up and correlates with response. *Clin. Cancer Res.* **21**, 1019–1027 (2015).
89. Sun, Q., Zhang, X., Wang, L., Gao, X., Xiong, Y., Liu, L., Wei, F., Yang, L. & Ren, X. T-cell receptor gene therapy targeting melanoma-associated antigen-A4 by silencing of endogenous TCR inhibits tumor growth in mice and human. *Cell Death Dis.* **10**, 475 (2019).
90. Zhang, S.-Q., Ma, K.-Y., Schonnesen, A. A., Zhang, M., He, C., Sun, E., Williams, C. M., Jia, W. & Jiang, N. High-throughput determination of the antigen specificities of T cell receptors in single cells. *Nat. Biotechnol.* (2018). doi:10.1038/nbt.4282
91. Oliveira, G., Stromhaug, K., Klaeger, S., Kula, T., Frederick, D. T., Le, P. M., Forman, J., Huang, T., Li, S., Zhang, W., Xu, Q., Cieri, N., Clauser, K. R., Shukla, S. A., Neuberger, D., Justesen, S., MacBeath, G., Carr, S. A., Fritsch, E. F., Hacohen, N., Sade-Feldman, M., Livak, K. J., Boland, G. M., Ott, P. A., Keskin, D. B. & Wu, C. J. Phenotype, specificity and avidity of antitumour CD8 T cells in melanoma. *Nature* **596**, 119–125 (2021).
92. Levy, P. L. & Gros, A. Fast track to personalized TCR T cell therapies. *Cancer Cell* **40**, 447–449 (2022).
93. Method of the year 2013. *Nat. Methods* **11**, 1 (2014).
94. Dalerba, P., Kalisky, T., Sahoo, D., Rajendran, P. S., Rothenberg, M. E., Leyrat, A. A., Sim, S., Okamoto, J., Johnston, D. M., Qian, D., Zabala, M., Bueno, J., Neff, N. F., Wang, J., Shelton, A. A., Visser, B., Hisamori, S., Shimono, Y., van de Wetering, M., Clevers, H., Clarke, M. F. & Quake, S. R. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* **29**, 1120–1127 (2011).
95. Wu, A. R., Neff, N. F., Kalisky, T., Dalerba, P., Treutlein, B., Rothenberg, M. E., Mburu, F. M., Mantalas, G. L., Sim, S., Clarke, M. F. & Quake, S. R. Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* **11**, 41–46 (2014).
96. Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A. & McCarroll, S. A.

- Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).
97. Luecken, M. D. & Theis, F. J. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.* **15**, e8746 (2019).
 98. Stegle, O., Teichmann, S. A. & Marioni, J. C. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* **16**, 133–145 (2015).
 99. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-R. & Raychaudhuri, S. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
 100. Hotelling, H. *Analysis of a Complex of Statistical Variables Into Principal Components*. (1933).
 101. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software* **3**, 861 (2018).
 102. Singer, M., Wang, C., Cong, L., Marjanovic, N. D., Kowalczyk, M. S., Zhang, H., Nyman, J., Sakuishi, K., Kurtulus, S., Gennert, D., Xia, J., Kwon, J. Y. H., Nevin, J., Herbst, R. H., Yanai, I., Rozenblatt-Rosen, O., Kuchroo, V. K., Regev, A. & Anderson, A. C. A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells. *Cell* **166**, 1500–1511.e9 (2016).
 103. Brummelman, J., Mazza, E. M. C., Alvisi, G., Colombo, F. S., Grilli, A., Mikulak, J., Mavilio, D., Alloisio, M., Ferrari, F., Lopci, E., Novellis, P., Veronesi, G. & Lugli, E. High-dimensional single cell analysis identifies stem-like cytotoxic CD8 T cells infiltrating human tumors. *J. Exp. Med.* **215**, 2520–2535 (2018).
 104. Cohn, L. B., da Silva, I. T., Valieris, R., Huang, A. S., Lorenzi, J. C. C., Cohen, Y. Z., Pai, J. A., Butler, A. L., Caskey, M., Jankovic, M. & Nussenzweig, M. C. Clonal CD4 T cells in the HIV-1 latent reservoir display a distinct gene profile upon reactivation. *Nat. Med.* **24**, 604–609 (2018).
 105. Buettner, F., Natarajan, K. N., Casale, F. P., Proserpio, V., Scialdone, A., Theis, F. J., Teichmann, S. A., Marioni, J. C. & Stegle, O. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.* **33**, 155–160 (2015).
 106. Stubbington, M. J. T., Lönnberg, T., Proserpio, V., Clare, S., Speak, A. O., Dougan, G. & Teichmann, S. A. T cell fate and clonality inference from single-cell transcriptomes. *Nat. Methods* **13**, 329–332 (2016).
 107. van der Leun, A. M., Thommen, D. S. & Schumacher, T. N. CD8 T cell states in human cancer: insights from single-cell analysis. *Nat. Rev. Cancer* **20**, 218–232 (2020).
 108. Lönnberg, T., Svensson, V., James, K. R., Fernandez-Ruiz, D., Sebina, I., Montandon, R., Soon, M. S. F., Fogg, L. G., Nair, A. S., Liligeto, U., Stubbington, M. J. T., Ly, L.-H., Bagger, F. O., Zwiessle, M., Lawrence, N. D., Souza-Fonseca-Guimaraes, F., Bunn, P. T., Engwerda, C. R., Heath, W. R., Billker, O., Stegle, O., Haque, A. & Teichmann, S. A. Single-cell RNA-seq and computational analysis using temporal mixture modelling resolves Th1/Tfh fate bifurcation in malaria. *Sci Immunol* **2**, (2017).
 109. Pearlman, A. H., Hwang, M. S., König, M. F., Hsiue, E. H.-C., Douglass, J., DiNapoli, S. R., Mog, B. J., Bettgowda, C., Pardoll, D. M., Gabelli, S. B., Papadopoulos, N., Kinzler, K. W., Vogelstein, B. & Zhou, S. Targeting public neoantigens for cancer immunotherapy. *Nat Cancer* **2**, 487–497 (2021).
 110. Leko, V. & Rosenberg, S. A. Identifying and Targeting Human Tumor Antigens for T Cell-Based Immunotherapy of Solid Tumors. *Cancer Cell* **38**, 454–472 (2020).
 111. Parkhurst, M. R., Robbins, P. F., Tran, E., Prickett, T. D., Gartner, J. J., Jia, L., Ivey, G., Li, Y. F., El-Gamil, M., Lalani, A., Crystal, J. S., Sachs, A., Groh, E., Ray, S., Ngo, L. T., Kivitz, S., Pasetto, A., Yossef, R., Lowery, F. J., Goff, S. L., Lo, W., Cafri, G., Deniger, D. C., Malekzadeh, P., Ahmadzadeh, M., Wunderlich, J. R., Somerville, R. P. T. & Rosenberg, S. A. Unique Neoantigens Arise from Somatic Mutations in Patients with Gastrointestinal Cancers. *Cancer Discov.* **9**, 1022–1035 (2019).
 112. Simoni, Y., Becht, E., Fehlings, M., Loh, C. Y., Koo, S.-L., Teng, K. W. W., Yeong, J. P. S., Nahar, R., Zhang, T., Kared, H., Duan, K., Ang, N., Poidinger, M., Lee, Y. Y., Larbi, A., Khng, A. J., Tan, E., Fu, C., Mathew, R., Teo, M., Lim, W. T., Toh, C. K., Ong, B.-H., Koh, T., Hillmer, A. M., Takano, A., Lim, T. K. H., Tan, E. H., Zhai, W., Tan, D. S. W., Tan, I. B. & Newell, E. W. Bystander CD8 T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* **557**, 575–579 (2018).
 113. Scheper, W., Kelderman, S., Fanchi, L. F., Linnemann, C., Bendle, G., de Rooij, M. A. J., Hirt, C., Mezzadra, R., Slagter, M., Dijkstra, K., Kluin, R. J. C., Snaebjornsson, P., Milne, K., Nelson, B. H., Zijlmans, H., Kenter, G., Voest, E. E., Haanen, J. B. A. G. & Schumacher, T. N. Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers. *Nat. Med.* **25**, 89–94 (2019).
 114. Thommen, D. S., Koelzer, V. H., Herzog, P., Roller, A., Trefny, M., Dimeloe, S., Kiialainen, A., Hanhart, J., Schill, C., Hess, C., Savic Prince, S., Wiese, M., Lardinois, D., Ho, P.-C., Klein, C., Karanikas, V., Mertz, K. D., Schumacher, T. N. & Zippelius, A. A transcriptionally and functionally distinct PD-1 CD8 T cell pool with

- predictive potential in non-small-cell lung cancer treated with PD-1 blockade. *Nat. Med.* **24**, 994–1004 (2018).
115. Sade-Feldman, M., Yizhak, K., Bjorgaard, S. L., Ray, J. P., de Boer, C. G., Jenkins, R. W., Lieb, D. J., Chen, J. H., Frederick, D. T., Barzily-Rokni, M., Freeman, S. S., Reuben, A., Hoover, P. J., Villani, A.-C., Ivanova, E., Portell, A., Lizotte, P. H., Aref, A. R., Eliane, J.-P., Hammond, M. R., Vitzthum, H., Blackmon, S. M., Li, B., Gopalakrishnan, V., Reddy, S. M., Cooper, Z. A., Paweletz, C. P., Barbie, D. A., Stemmer-Rachamimov, A., Flaherty, K. T., Wargo, J. A., Boland, G. M., Sullivan, R. J., Getz, G. & Hacohen, N. Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. *Cell* **175**, 998–1013.e20 (2018).
 116. Daud, A. I., Loo, K., Pauli, M. L., Sanchez-Rodriguez, R., Sandoval, P. M., Taravati, K., Tsai, K., Nosrati, A., Nardo, L., Alvarado, M. D., Algazi, A. P., Pampaloni, M. H., Lobach, I. V., Hwang, J., Pierce, R. H., Gratz, I. K., Krummel, M. F. & Rosenblum, M. D. Tumor immune profiling predicts response to anti-PD-1 therapy in human melanoma. *J. Clin. Invest.* **126**, 3447–3452 (2016).
 117. Duhon, T., Duhon, R., Montler, R., Moses, J., Moudgil, T., de Miranda, N. F., Goodall, C. P., Blair, T. C., Fox, B. A., McDermott, J. E., Chang, S.-C., Grunkemeier, G., Leidner, R., Bell, R. B. & Weinberg, A. D. Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat. Commun.* **9**, 2724 (2018).
 118. Caushi, J. X., Zhang, J., Ji, Z., Vaghasia, A., Zhang, B., Hsiue, E. H.-C., Mog, B. J., Hou, W., Justesen, S., Blosser, R., Tam, A., Anagnostou, V., Cottrell, T. R., Guo, H., Chan, H. Y., Singh, D., Thapa, S., Dykema, A. G., Burman, P., Choudhury, B., Aparicio, L., Cheung, L. S., Lanis, M., Belcaid, Z., El Asmar, M., Illei, P. B., Wang, R., Meyers, J., Schuebel, K., Gupta, A., Skaist, A., Wheelan, S., Naidoo, J., Marrone, K. A., Brock, M., Ha, J., Bush, E. L., Park, B. J., Bott, M., Jones, D. R., Reuss, J. E., Velculescu, V. E., Chaft, J. E., Kinzler, K. W., Zhou, S., Vogelstein, B., Taube, J. M., Hellmann, M. D., Brahmer, J. R., Merghoub, T., Forde, P. M., Yegnasubramanian, S., Ji, H., Pardoll, D. M. & Smith, K. N. Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. *Nature* **596**, 126–132 (2021).
 119. Hanada, K.-I., Zhao, C., Gil-Hoyos, R., Gartner, J. J., Chow-Parmer, C., Lowery, F. J., Krishna, S., Prickett, T. D., Kivitz, S., Parkhurst, M. R., Wong, N., Rae, Z., Kelly, M. C., Goff, S. L., Robbins, P. F., Rosenberg, S. A. & Yang, J. C. A phenotypic signature that identifies neoantigen-reactive T cells in fresh human lung cancers. *Cancer Cell* **40**, 479–493.e6 (2022).
 120. Liu, B., Zhang, Y., Wang, D., Hu, X. & Zhang, Z. Single-cell meta-analyses reveal responses of tumor-reactive CXCL13 T cells to immune-checkpoint blockade. *Nat Cancer* **3**, 1123–1136 (2022).
 121. Lowery, F. J., Krishna, S., Yossef, R., Parikh, N. B., Chatani, P. D., Zacharakis, N., Parkhurst, M. R., Levin, N., Sindiri, S., Sachs, A., Hitscherich, K. J., Yu, Z., Vale, N. R., Lu, Y.-C., Zheng, Z., Jia, L., Gartner, J. J., Hill, V. K., Copeland, A. R., Nah, S. K., Masi, R. V., Gasmi, B., Kivitz, S., Paria, B. C., Florentin, M., Kim, S. P., Hanada, K.-I., Li, Y. F., Ngo, L. T., Ray, S., Shindorf, M. L., Levi, S. T., Shepherd, R., Toy, C., Parikh, A. Y., Prickett, T. D., Kelly, M. C., Beyer, R., Goff, S. L., Yang, J. C., Robbins, P. F. & Rosenberg, S. A. Molecular signatures of antitumor neoantigen-reactive T cells from metastatic human cancers. *Science* **375**, 877–884 (2022).

Hypothesis and Objectives

In this thesis, I) the cellular dynamics of T and B lymphocytes in convalescent COVID-19 patients and T lymphocytes in lung cancer patients are assessed, and II) T lymphocytes from lung cancer patients are functionally profiled to identify markers of reactivity (**Figure 2**). This work combines single-cell high-throughput technologies, next-generation sequencing and CRISPR-based technologies to derive cellular insights.

Chapter 2 & 3: Previous discoveries have shown how T and B lymphocytes can mount an immune reaction against cognate targets. In infection and cancer, the cell state of lymphocytes and their clonal expansion level can give insights into their involvement in the immune response and disease progression. By employing state-of-the-art single-cell multi-omic technologies, we assess the cell state of T and B cells from the blood of convalescent COVID-19 patients and T cells from the tumor and adjacent non-tumor tissue of lung cancer patients that we received in collaboration with the Swiss Military and the UniBasel Spital, respectively. This allowed us to identify the persisting circulating lymphocyte cell states across young and old convalescent COVID-19 patients with mild-to-moderate disease progression. Further, single-cell analysis of T cells in lung cancer tissue showed a large variety of cell states with a marked progression towards exhaustion. Comparison between tumor tissue and adjacent non-tumor tissue gave insights into shared T cell clones and different gene-expression modules in shared clones between tumor and non-tumor tissue.

Chapter 3: The theoretical diversity of T cell receptors is 10^{18} , each with a considerable number of potential target peptide-HLAs. This vast diversity significantly contributes to the challenge of identifying reactive T cells and their specific targets. Especially in the case of cancer immunology, the identification of tumor-reactive T cells poses a challenge for the development of better therapies. Recent work has shown the potential of linking T cell reactivity to cell state using functional T cell profiling and single-cell technologies. Due to potential hypo-responsiveness in exhausted TILs, reactivity profiling remains challenging and biased towards responsive TILs. Using Cas9-enabled TCR transfer of TILs, we were able to develop an unbiased screen for the functional reactivity of TILs. We identified novel gene expression

markers across different reactivities that may help in the *in silico* enrichment of tumor-reactive TILs.

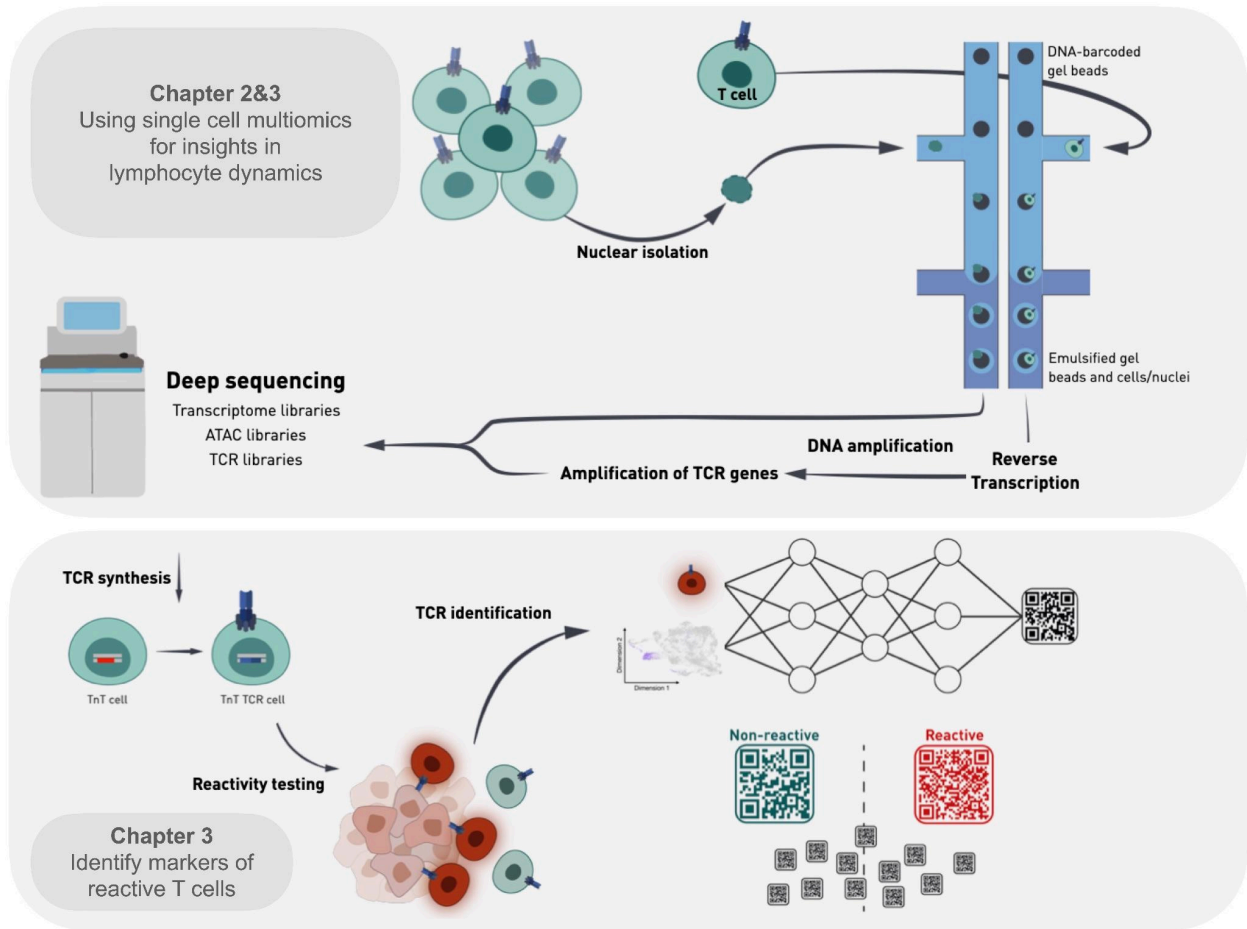


Figure 2: Content and connection of the thesis chapters 2 and 3. This figure is an author produced adaptation of a research grant proposal submitted and approved for funding from PHRT

Chapter 2: Single cell lymphocyte analysis of convalescent COVID-19 patients

This is an author-produced version of an article published in Frontiers in Immunology.

A single-cell atlas of lymphocyte adaptive immune repertoires and transcriptomes reveals age-related differences in convalescent COVID-19 patients

[Florian Bieberich*, Rodrigo Vazquez-Lombardi*, Alexander Yermanos, Roy A. Ehling, Derek M. Mason, Bastian Wagner, Edo Kapetanovic, Raphael Brisset Di Roberto, Cédric R. Weber, Miodrag Savic, Fabian Rudolf and Sai T. Reddy; *These authors contributed equally to this work]

FB, RV-L, AY, and SR designed the study. FB, RV-L, RE, DM, BW, EK, and RB performed experiments; FB, RV-L, AY, and CW analyzed data (note for shared first authors: FB analyzed the majority of the single-cell RNA and VDJ data, including GLIPH2 analysis. RV-L performed Graphpad Prism-based statistical analysis of age correlations and analyzed the HLA NGS data.) FB, RV-L, AY, and SR wrote the manuscript with input from all authors. All authors contributed to the article and approved the submitted version. Shared first author contributions in detail:

2.1 Summary

COVID-19 disease outcome is highly dependent on adaptive immunity from T and B lymphocytes, which play a critical role in the control, clearance and long-term protection against SARS-CoV-2. To date, there is limited knowledge on the composition of the T and B cell immune receptor repertoires [T cell receptors (TCRs) and B cell receptors (BCRs)] and transcriptomes in convalescent COVID-19 patients of different age groups. Here, we utilize single-cell sequencing (scSeq) of lymphocyte immune repertoires and transcriptomes to quantitatively profile the adaptive immune response in COVID-19 patients of varying age. We discovered highly expanded T and B cells in multiple patients, with the most expanded clonotypes coming from the effector CD8⁺ T cell population. Highly expanded CD8⁺ and CD4⁺

T cell clones show elevated markers of cytotoxicity (CD8: PRF1, GZMH, GNLY; CD4: GZMA), whereas clonally expanded B cells show markers of transition into the plasma cell state and activation across patients. By comparing young and old convalescent COVID-19 patients (mean ages = 31 and 66.8 years, respectively), we found that clonally expanded B cells in young patients were predominantly of the IgA isotype and their BCRs had incurred higher levels of somatic hypermutation than elderly patients. In conclusion, our scSeq analysis defines the adaptive immune repertoire and transcriptome in convalescent COVID-19 patients and shows important age-related differences implicated in immunity against SARS-CoV-2.

2.2 Introduction

T and B lymphocytes are crucial for protection from SARS-CoV-2 infection, viral clearance and the formation of persisting antiviral immunity^{1,2}. Yet, adaptive immune responses have also been implicated in contributing to immunopathology during COVID-19, with higher mortality rates in elderly individuals³⁻⁶. However, the exact determinants of a successful adaptive immune response against SARS-CoV-2 and its variability between different age groups remain to be fully elucidated.

Lymphocytes express either T cell receptors (TCR) or B cell receptors (BCR), which possess a highly diverse pair of variable chains [variable alpha ($V\alpha$) and beta ($V\beta$) for TCR and variable light (V_L) and heavy (V_H) for BCR] that are able to directly engage with antigen (e.g., viral proteins or peptides). Diversity in these variable chains are generated by somatic recombination of V-, D- and J-gene germline segments and along with combinatorial receptor chain pairing and somatic hypermutation (BCR only) results in an estimated human TCR and BCR diversity of 10^{18} and 10^{13} , respectively^{7,8}. Upon encountering cognate antigens, lymphocytes are phenotypically activated and undergo massive proliferation, also referred to as clonal selection and expansion. Deep sequencing of TCRs and BCRs has become a powerful strategy to profile the diversity of immune repertoires and to reveal insights on clonal selection, expansion and evolution (somatic hypermutation in B cells)⁹⁻¹² and has been instrumental in studying long term effects following vaccination, infection and aging¹³⁻¹⁷. In the context of COVID-19, immune

repertoire sequencing has shown diminished TCR repertoire diversity and BCR isotype switching and respective expansion during early disease onset ¹⁸.

In recent years, single-cell sequencing (scSeq) of transcriptomes has progressed substantially through the development and integration of technologies such as cell sorting, microwells and droplet microfluidics ^{19,20}; most notably commercial systems like those of 10X Genomics have been established and are providing standardized protocols for wider implementation of scSeq. To find interactions across multiple genes and cells, analysis and visualization of this high dimensional single-cell data is facilitated by clustering and nonlinear dimensionality reduction algorithms [e.g., t-distributed stochastic neighbor embedding (t-SNE) or Uniform Manifold Approximation and Projection (UMAP)] ^{21,22}. scSeq of transcriptomes has been used extensively to profile the gene expression signatures of T and B cells to identify novel cellular subsets and phenotypes as well as their response to vaccination, infection and cancer ²³⁻²⁶. Furthermore, clustering with scSeq data enables the unbiased identification of cellular states and analyses of the broad continuum of T and B cell populations as well as their differentiation trajectories ²⁷. In the context of patients with severe symptoms of COVID-19, scSeq has revealed a dysfunctional T cell response of interferon expression combined with elevated levels of exhaustion ²⁸.

In addition to transcriptome sequencing, a major advantage of scSeq is that it also enables information on the native pairing of TCR V α and V β chains and BCR V_L and V_H chains ²⁹⁻³², which was not previously possible with the standard bulk sequencing of lymphocytes as these receptor chains are expressed as unique transcripts from separate chromosomes ³³. Coupling TCR or BCR sequence to the transcriptome within an individual cell enables phenotypic analyses of a clonal population of lymphocytes and their dynamics ³⁴⁻³⁶. scSeq of transcriptomes and immune repertoires in COVID-19 patients with severe symptoms has shown a high level of clonal expansion in specific T cell subsets (Th1, Th2, and Th17) and preferential germline gene usage in clonally expanded B cells ^{28,34,37}; while a more recent study found a positive correlation between clonal expansion of effector-like CD8⁺ T cells and disease severity ³⁸.

An important question that remains to be answered is whether there are age-related differences in mounting a successful adaptive immune response against SARS-CoV-2. Here, we perform scSeq

on the immune repertoires and transcriptomes of T and B cells derived from eight convalescent COVID-19 patients of two different age groups (mean ages = 31 and 66.8 years) at one month of convalescence following mild to moderate disease. We observed preferential clonal expansion of effector CD8⁺ T cells across all patients, although a significantly higher CD8-to-CD4 T cell ratio was detected in young patients of our cohort. Further, clonally expanded B cells in young patients displayed significantly higher levels of somatic hypermutation and an increased immunoglobulin (Ig) class-switching compared to clonally expanded B cells from older patients. Our analyses serve as a valuable resource for future scSeq characterization of SARS-CoV-2 adaptive immunity and highlight important age-related differences in the adaptive immune status of convalescent COVID-19 patients.

2.3 Results

Study design and single-cell profiling of convalescent COVID-19 patient lymphocytes

We performed scSeq of immune receptor repertoires and transcriptomes of lymphocytes from convalescent COVID-19 patients to characterize the adaptive immune response against SARS-CoV-2. For this purpose, we selected eight patients enrolled in the SERO-BL-COVID-19 clinical study ³⁹, all of which fully recovered from COVID-19 without requiring hospitalization or the administration of supplemental oxygen. Patients tested positive for the presence of SARS-CoV-2 after RT-PCR of naso/oropharyngeal swab samples (day 0), displayed COVID-19 symptoms for 4-14 days, and showed positive seroconversion at the time of blood collection (mean sample collection time = 32.5 ± 4.1 days post-symptom onset) (**Fig. 1a and Supplemental Table 1**). Since COVID-19 often affects older patients more severely ⁴⁰, subjects were divided into two groups according to their age, namely Group 1 (mean = 66.75 ± 6.9 years) and Group 2 (mean = 31 ± 5.9 years), with the aim of investigating potential differences in their responses against SARS-CoV-2. In addition to older age, significant differences in Group 1 versus Group 2 included elevated IgA/IgG SARS-CoV-2-specific antibody titers and an increased duration of COVID-19 symptoms (**Fig. 1b and Supplemental Fig.1**). Patients from Group 1 also experienced an increased severity of COVID-19 symptoms relative to Group 2 (**Supplemental Table 1**). Despite increased symptom duration in the older cohort, correlation of this parameter with age was only modest ($R^2 = 0.4647$), likely reflecting the small sample size (**Supplemental Fig. 1c**).

To profile patient lymphocytes, we isolated peripheral blood mononuclear lymphocytes (PBMC) from blood and purified T cells and B cells by negative immunomagnetic enrichment. Plasma cells (PCs) were depleted from PBMC samples prior to this step for scSeq in a companion study (Ehling et al., manuscript in preparation), and thus were excluded from our analyses. After purification, T cells and B cells underwent the 10X genomics protocol for scSeq 5' library preparation, which included gel encapsulation single-cell barcoding of mRNA, followed by cDNA generation through polydT reverse transcription. Finally, after full-length V(D)J segment enrichment, construction of TCR and BCR V(D)J and transcriptome sequencing libraries was done according to the V(D)J enrichment and 5' library construction kits, respectively. Deep

sequencing of immune repertoires and transcriptomes was performed using the Illumina NovaSeq with paired-end 26 x 91 bp cycles per read. For TCR and BCR V(D)J and transcriptome libraries, we recovered on average 20.000 and 10.000 reads per cell, respectively (Fig. 1c).

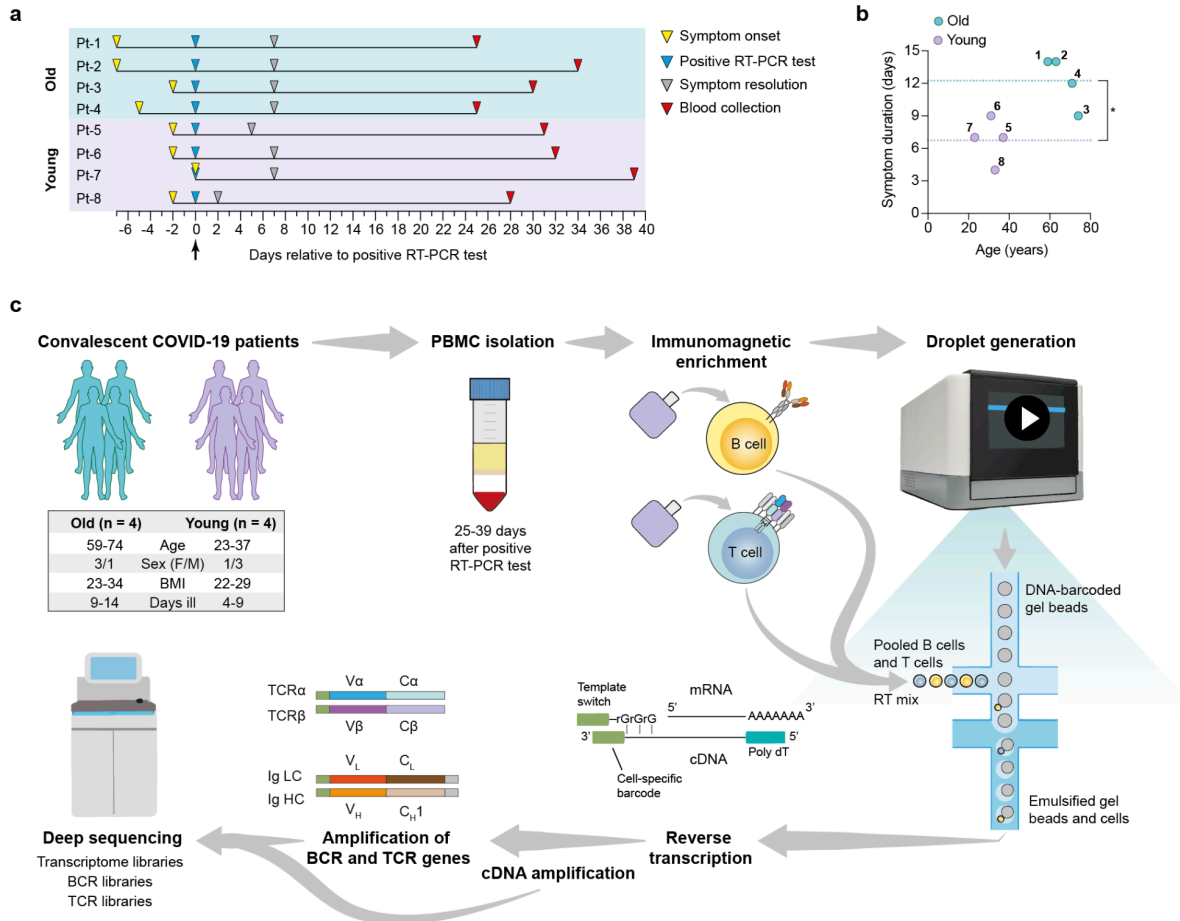


Figure 1. Overview of single-cell transcriptome and immune receptor profiling of convalescent COVID-19 patient lymphocytes. Convalescent COVID-19 patients enrolled in the SERO-BL-COVID-19 study were selected according to their age for single-cell sequencing analysis of their T cells and B cells. **a**, Timeline illustrates symptom onset, symptom resolution and collection of blood samples from individual patients relative to the time of positive SARS-CoV-2 RT-PCR test (day 0). **b**, Graph displays the ages and duration of COVID-19 symptoms in individual patients. Dotted lines show the mean duration of symptoms in the young ($y = 6.75$ days) and old ($y = 12.25$ days) groups. A significant difference in symptom duration between groups is indicated with an asterisk ($p = 0.0127$; unpaired t-test). **c**, Single-cell sequencing protocol. Whole blood was collected following the resolution of COVID-19 symptoms and subjected to density gradient separation for isolation of PBMC. T cells and B cells from individual patients were purified from PBMC using negative immunomagnetic enrichment, pooled (intra-patient) and prepared for droplet generation using the 10x Genomics Chromium system. Single cells were emulsified with

DNA-barcoded gel beads and mRNA transcripts were reverse-transcribed within droplets, resulting in the generation of first-strand cDNA molecules labeled with cell-specific barcodes at their 3' ends (added by template switching). Emulsions were disrupted and cDNA was amplified by means of PCR for further processing of transcriptome libraries. Transcriptome libraries from individual patients were indexed and multiplexed for deep sequencing using the Illumina NovaSeq platform. Targeted enrichment of recombined V(D)J transcripts was performed by PCR and the resulting products were processed for the generation of BCR and TCR libraries, which were then indexed, multiplexed and deep-sequenced.

Single-cell transcriptome analysis defines major T and B cell subsets

Bioinformatic filtering was performed to exclude the following: Cell doublets, cells with a very low or high number of genes, and T cells with no detectable expression of CD8 and CD4 (see Methods), which resulted in the identification of 30,096 cells in total from all eight patients. Cells were then split into CD8⁺ T cell (**Fig. 2a**, 7,353 cells), CD4⁺ T cell (**Fig. 2b**, 8,334 cells) and B cell (**Fig. 2c**, 14,409 cells) datasets. In order to reduce the dimensionality of the data, while preserving the global structure, we used UMAP for better visualisation and interpretation purposes ⁴¹. UMAP and unsupervised clustering of these subgroups led to the identification of eleven dominant cell subsets (**Figs. 2a-c**). CD8⁺ T cells clustered into naïve (SELL⁺, TCF7⁺), memory (IL7R⁺, CD40LG⁺) and effector cells (GZMB⁺, NKG7⁺) (**Figs. 2a and 2d**), which also encompassed exhausted CD8⁺ T cells (**Supplemental Fig. 2**). We identified four different CD4⁺ T cell subsets, namely naïve (SELL⁺, LEF1⁺), memory (S100A4⁺), effector (CCL5⁺, GZMK⁺) and regulatory cells (FOXP3⁺) (**Figs. 2b and 2e**). The B cell compartment consisted of naïve (CD23A⁺), marginal zone (MZ) (FCRL3⁺, CD1C⁺), activated (CD83⁺) and memory cells (CD27⁺, TACI⁺) (**Figs. 2c and 2f**). Notably, clustering of single cells based on transcriptome data revealed a trajectory that reflected a progression in lymphocyte differentiation from naïve to effector (or activated) subsets (**Figs. 2a-c**). Pseudotime analysis of the dataset supports this differentiation trajectory in CD8⁺ and CD4⁺ T cells (**Figs. 2g and 2h**). Interestingly, pseudotime analysis of B cell data not only showed naïve-MZ-memory and naïve-MZ-activated trajectories, but also a third MZ-memory-activated trajectory that suggests the presence of reactivated memory B cells, possibly through antigen encounter (**Fig. 2i**).

Having defined the major T cell and B cell subsets from pooled patient data, we next compared their proportions across patients (**Figs. 2j-l**) and between different age groups (**Supplemental Fig. 3**). We found that young patients had a significantly higher CD8-to-CD4 T cell ratio relative to older ones (**Supplemental Fig. 3a**), which may reflect a previously reported age-dependent

difference ⁴². Interestingly, despite this reduction, there was a trend that older patients had a higher proportion of effector CD8⁺ T cells relative to their younger counterparts (**Supplemental Fig. 3b**). While this difference was not significant, it is consistent with the increased symptom severity experienced by older patients (**Supplemental Table 1**), a feature that has been associated with elevated proportions of effector/exhausted CD8⁺ T cells in the periphery ²⁸. Of note, we found that older patients had a small but significant increase in CD4⁺ Tregs compared to young patients (**Supplemental Fig. 3h**), and that increased proportions of MZ B cells occurred in two of the older patients (**Supplemental Fig. 3l**). Taken together, our data highlights the diversity of elevated responses in specific patients across age groups, as exemplified by individuals with a high abundance of effector CD8⁺ T cells (e.g., Pt-2 and Pt-3) and/or activated B cells (e.g., Pt-2, Pt-5 and Pt-8).

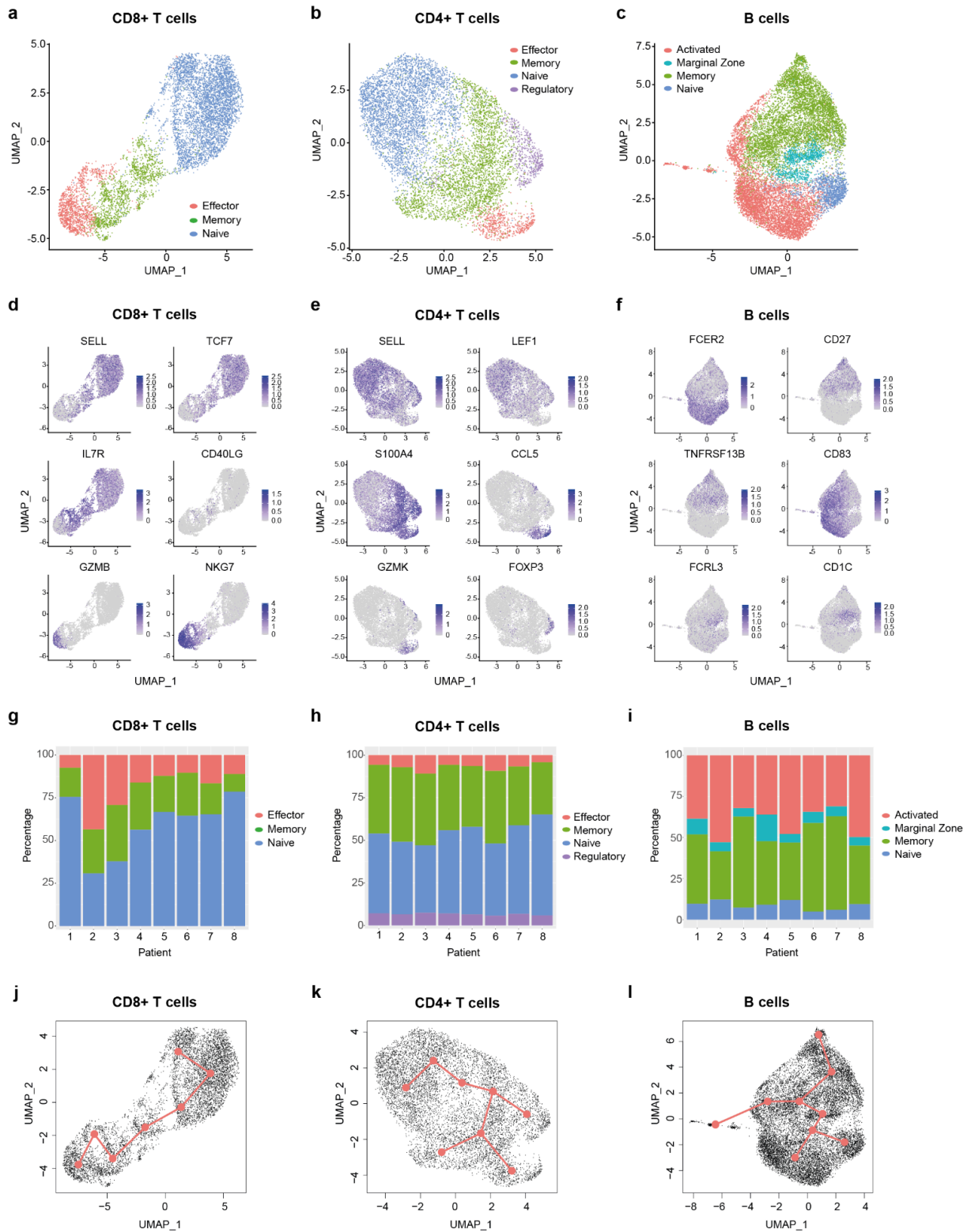


Figure 2. Single-cell transcriptomic analysis delineates major T and B cell subsets. a-c, Uniform manifold approximation and projection (UMAP) plots of major cellular subsets identified within the CD8⁺ T cell (a), CD4⁺ T cell (b) and B cell (c) populations. Cells from all patients are displayed in each plot. d-f, UMAP plots showing the

expression levels of selected genes used to delineate cellular subsets within the CD8⁺ T cell (**d**), CD4⁺ T cell (**e**) and B cell (**f**) populations. Cells from all patients are displayed in each plot. **g-i**, Graphs display pseudotime and trajectory inference analysis applied to CD8⁺T cell (**g**), CD4⁺ T cell (**h**) and B cell (**i**) clusters. **j-l**, Bar graphs show the proportions of identified cellular subsets within the CD8⁺ T cell (**j**), CD4⁺ T cell (**k**) and B cell (**l**) populations in each patient. CCL5 = C-C Motif Chemokine Ligand 5; CD27 = TNFRSF7; CD40LG = CD40 ligand; FCER2 = Fc Fragment of IgE Receptor II (also: CD23a); FCRL3 = Fc Receptor Like 3; FOXP3 = Forkhead Box Protein P3; GZMB = Granzyme B; GZMK = Granzyme K; IL7R = Interleukin-7 Receptor; LEF1 = Lymphoid Enhancer Binding Factor 1; NKG7 = Natural Killer Cell Granule Protein 7; S100A4 = S100 Calcium Binding Protein A4; SELL = Selectin L; TCF7 = Transcription Factor 7; TNFRSF13B = TNF Receptor Superfamily Member 13B.

Single-cell profiling of immune receptor repertoires identifies highly expanded TCR and BCR clonotypes

We next determined the clonal expansion levels of T cells and B cells in convalescent COVID-19 patients by quantifying the number of cells expressing unique TCRs (**Fig. 3a**) or BCRs (**Fig. 3c**) (clonotype definition in the methods section). We found substantial heterogeneity in T cell clonal expansion levels across patients, with the highest number of expanded TCR clonotypes occurring in four patients, namely Pt-2 and Pt-3 (Group 1), Pt-7 and Pt-8 (Group 2). Within these patients, Pt-2 displayed the largest amount of expanded TCR clonotypes, which is consistent with the high abundance of effector CD8⁺ T cells in this subject (**Fig. 2g**). Analysis of TCR α and TCR β germline V-gene usage in the ten most expanded clonotypes per patient revealed a frequent occurrence of TRBV20-1 (7 out of 8 patients) and TRAV-29/DV5 genes (5 out of 8 patients), though pairing of these germline genes was not observed (**Fig. 3b and Supplemental Fig. 4**). In agreement with the overall higher expansion of CD8⁺ effector over CD4⁺ effector T cell subsets (**Figs. 2g and 2h**), we found that the vast majority (85%) of the ten most expanded TCR clonotypes per patient originated from CD8⁺ T cells (**Supplemental Fig. 5**). Based on this observation, we genotyped patient HLA class I alleles by means of amplicon deep sequencing (**Supplemental Table 2**). We found that the two patients with the highest levels of T cell clonal expansion (i.e., Pt-2 and Pt-8) shared the HLA-A*0201 allele, as well as a number of TRBV and TRAV genes in their most expanded clonotypes, which indicates a possible convergence towards germlines that may be related to SARS-CoV-2 specificity. Analysis of single-cell BCR repertoire sequencing data revealed that highly expanded BCR clonotypes occurred more frequently in younger patients, for example in Pt-5, Pt-6 and Pt-8 (**Fig. 3c**). This is an unexpected finding, particularly as older patients in our cohort displayed significantly higher SARS-CoV-2-specific IgA/IgG titers in serum (**Supplemental Figs. 1e and 1f**). Thus, this suggests that older patients

may harbor a higher diversity of relatively unexpanded SARS-CoV-2-specific B cells. Supporting this observation, we found that older patients had a wider range of heavy chain complementarity determining region 3 (CDR3H) lengths relative to younger ones, indicating a possible larger degree of variability in their antibody paratopes (**Supplemental Fig. 6**). Analysis of heavy chain and light chain germline V-gene pairing in the ten most expanded BCR clonotypes per patient revealed a frequent occurrence of the IGHV-3-23 / IGKV-3-20 pairing (7 out of 8 patients) (**Fig. 3d and Supplemental Fig. 7**). However, as this pairing is the most frequently found in healthy cohorts ^{43,44}, such antibodies may not necessarily be enriched for SARS-CoV-2 specificity.

To further characterize BCR repertoires across different levels of clonal expansion we divided clonotypes into additional subsets: unexpanded (1 cell per clonotype), expanded (2-4 cells per clonotype) and highly expanded (≥ 5 cells per clonotype) and assessed their levels of somatic hypermutation (SHM). We found that the degree of SHM largely correlated with clonal expansion, with expanded and highly expanded clonotypes having higher SHM (i.e., more divergent from their germline V-genes) than unexpanded ones (**Fig. 3e**). Strikingly, highly expanded BCR clonotypes from young patients had significantly higher SHM levels compared to older patients, potentially indicating more efficient affinity maturation had occurred in response to SARS-CoV-2 antigens (**Fig. 3f**). Finally, we examined the distribution of Ig isotypes across clonal expansion groups. As expected, IgM was the most frequent isotype in unexpanded BCR clonotypes, with the proportion of this isotype being reduced in expanded BCR clonotypes (2-4 cells) of all patients. Conversely, the proportions of IgG and IgA isotypes in expanded clonotypes increased for all patients, thus indicating class-switching in response to clonal expansion. Analysis of Ig isotype distribution in highly expanded BCR clonotypes (≥ 5 cells) revealed that a subset of patients harbored a vast majority of class-switched IgA (Pt-5, Pt-6 and Pt-7) or IgG (Pt-4). Notably, we found that Ig isotype class-switching in highly expanded clonotypes was correlated with SHM levels across patients (**Fig. 3e**), highlighting the temporal connection between affinity maturation and class-switching processes in the germinal center ⁴⁵.

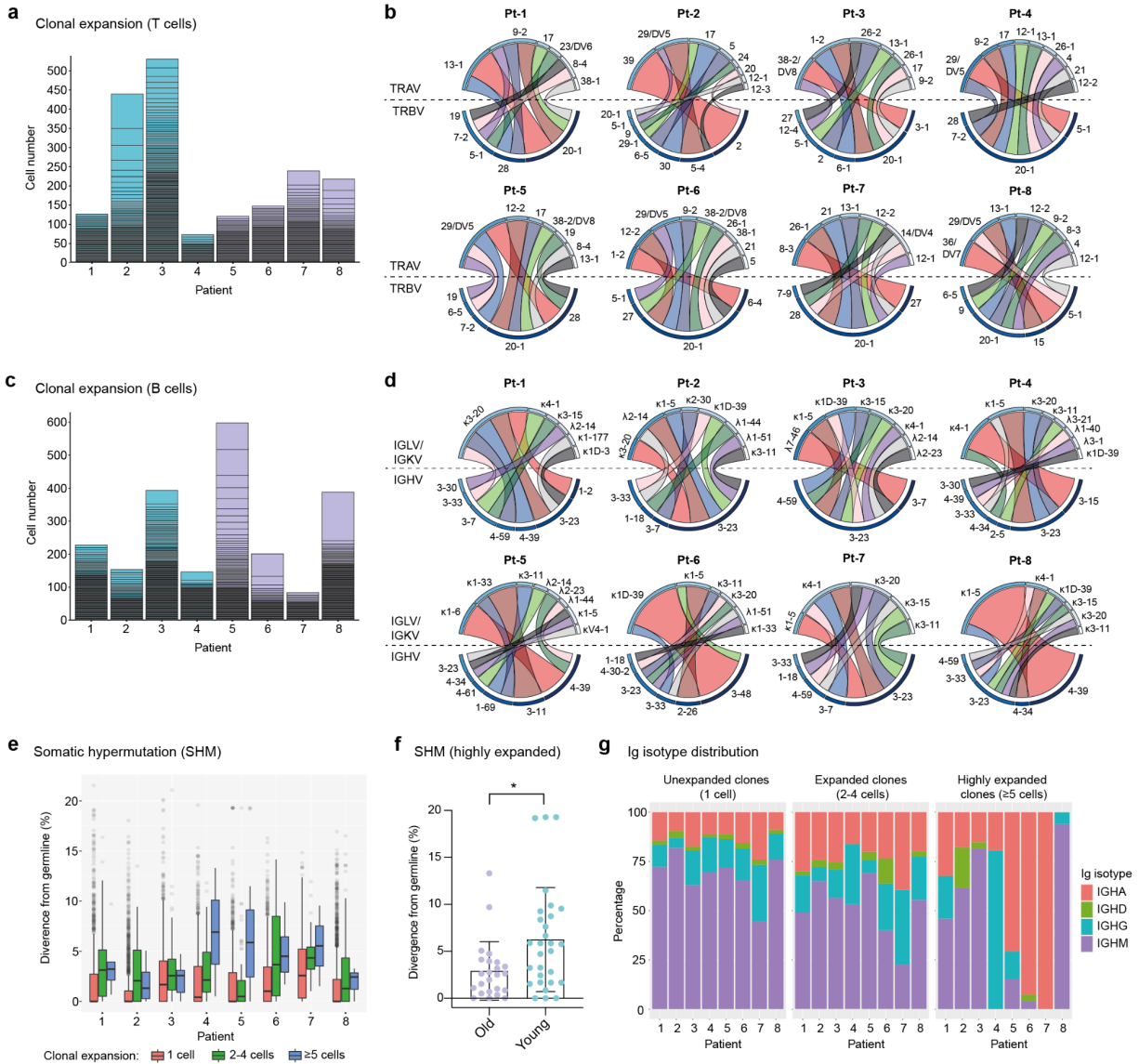


Figure 3. Single-cell profiling of immune repertoires highlights differential levels of inter-patient T cell and B cell clonal expansion. **a-b**, Analysis of T cell clonal expansion in convalescent COVID-19 patients. **a**, Bar graphs show T cell clonal expansion, as determined by the number of cells identified per TCR clonotype. Each box represents the size of individual TCR clonotypes. TCR clonotypes present in more than one cell are shown. **b**, Circos plots display V-gene usage in the top ten most expanded TCR clonotypes for each patient. The size and color (dark to light) of outer bars reflect the relative abundance of T cells expressing specific V-genes on a per class basis (top: TCR α chain, bottom: TCR β chain). **c-e**, Analysis of B cell clonal expansion in convalescent COVID-19 patients. **c**, Bar graphs show B cell clonal expansion, as determined by the number of cells identified per BCR clonotype. Each box represents the size of individual BCR clonotypes. BCR clonotypes present in more than one cell are shown. **d**, Circos plots display V-gene usage in the top ten most expanded BCR clonotypes for each patient. The size and color (dark to light) of outer bars reflect the relative abundance of B cells expressing specific V-genes on a per class basis (top: Ig light chain, bottom: Ig heavy chain). **e**, Graph displays the levels of somatic hypermutation (SHM) in unexpanded (1 cell), expanded (2-4 cells) and highly expanded (5 cells) BCR clonotypes across patients. SHM levels are based on the percentage similarity between Ig heavy chain V-genes and their corresponding germlines. Data are displayed as median \pm IQR. **f**, Graph displays SHM levels in highly expanded

BCR clones (≥ 5 cells) of old ($n = 24$ clones) and young ($n = 29$ clones) patients. Asterisks indicate a significant difference in SHM levels between groups ($p = 0.0085$; unpaired t-test). Data are displayed as median \pm IQR. **g.** Bar graphs show the Ig isotype distribution in unexpanded (1 cell), expanded (2-4 cells) and highly expanded (≥ 5 cells) BCR clonotypes across patients.

Single-cell transcriptome and TCR profiling reveals predominant cytotoxic programs in highly clonally expanded CD4⁺ and CD8⁺ T cells

We next investigated the patterns of clonal expansion in different T cell subsets by mapping single-cell TCR sequencing data onto individual CD8⁺ and CD4⁺ T cells visualized by UMAP (**Figs. 4a and 4b**). For this analysis, we identified a total of 4,730 CD8⁺ T cells and 5,509 CD4⁺ T cells with available TCR clonotype and transcriptome information. Both CD8⁺ and CD4⁺ T cells showed increased levels of clonal expansion when progressing from naïve to effector phenotypes, with highly expanded TCR clonotypes (≥ 5 cells) almost exclusively expressed by effector T cells (**Figs. 4a and 4b, Figs. 2a and 2b**). As previously observed (**Supplemental Figs. 3 and 5**), CD8⁺ T cells showed substantially higher levels of clonal expansion relative to CD4⁺ T cells, in which highly expanded clonotypes were rare. Notably, young patients (Group 2) had a markedly higher abundance of unexpanded CD8⁺ T cell clonotypes compared to older patients (Group 1), which could indicate an ongoing resolution of their CD8⁺ T cell response at the analyzed timepoint. Patients with high levels of CD8⁺ T cell clonal expansion, however, were identified across age groups (i.e., Pt-2, Pt-3, Pt-7 and Pt-8), with Pt-2 (Group 1) showing the highest abundance of highly expanded clonal T cells. We further explored the relationship between clonal expansion and T cell phenotype by performing differential gene expression analysis in unexpanded, expanded and highly expanded T cell clonotypes (**Figs. 4c and 4d**). CD8⁺ T cells with high clonal expansion displayed elevated cytotoxicity (PRF1, GZMH, GNLY), activation (NKG7, CCL5), inflammation (NFKBIA, S100A4, S100A6) and type I interferon-induced (IFITM2) markers in all patients. Additionally, components of MHC class I (HLA-A, HLA-B, HLA-C and B2M) were also increased in this subgroup, indicating increased IFN γ -induced activation⁴⁶. Conversely, unexpanded CD8⁺ T cell clonotypes across age groups displayed upregulated markers found in naïve and memory CD8⁺ T cell subsets (IL7R, LTB)⁴⁷, as well as markers likely associated with homeostatic proliferation (LDHB, NOSIP, EEF1B2, NPM1, TPT1, PABPC1).

Despite the low levels of clonal expansion observed in CD4⁺ T cells, we identified distinct gene expression signatures in the highly expanded clonotypes that were present in 5 of 8 patients. Similar to CD8⁺ T cells, highly expanded CD4⁺ T cell clonotypes showed upregulation of genes related to activation (CCL5), cytotoxicity (GZMA), inflammation (IL32, CD99, NFKBIA) and MHC class I molecules (HLA-A, HLA-B, HLA-C and B2M), while unexpanded CD4⁺ T cells displayed markers of naïve T cells (SELL) and proliferation markers (LDHB, NOSIP, PABPC1). Taken together, integration of TCR sequencing and transcriptome data reveals clonal expansion as a hallmark of effector T cell subsets, and highlights a dominant role of CD8⁺ T cells in possible clonal responses against SARS-CoV-2 in convalescent COVID-19 patients.

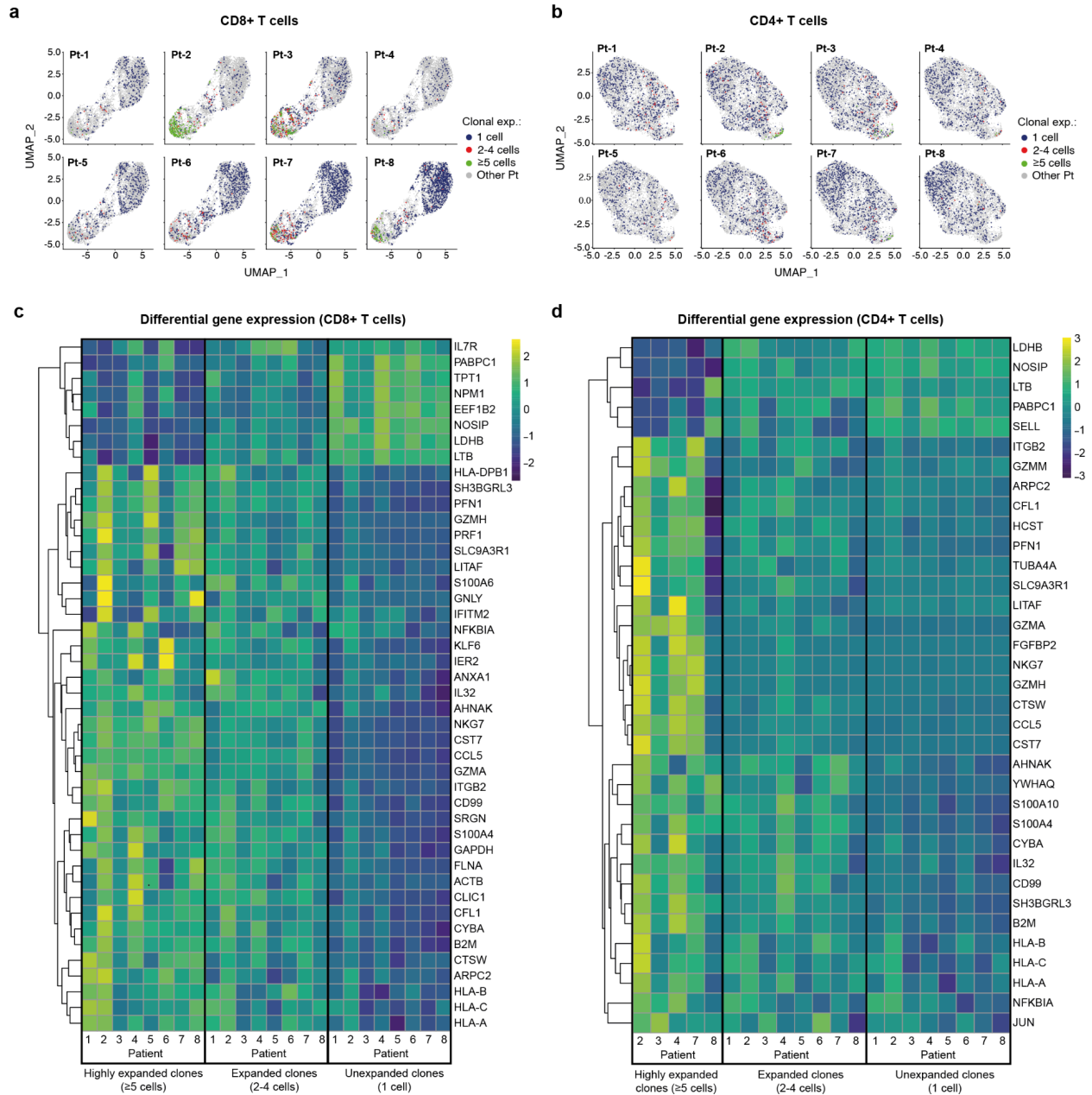


Figure 4. Single-cell transcriptome and TCR sequencing reveals preferential clonal expansion in effector T cells. **a-b** , UMAP plots display CD8⁺ (**a**) and CD4⁺ (**b**) T cells from specific patients according to their clonal expansion levels. T cells from other patients in each individual plot are shown in grey. **c-d**, Heatmaps show differential gene expression (DGE) in unexpanded, expanded and highly expanded CD8⁺ (**c**) or CD4⁺ (**d**) T cells. Genes were filtered to include those with detectable expression in at least 50% of cells and that had a minimum 50% fold-change in expression level between groups.

Transcriptomic and BCR profiling of single B cells reveals plasma cell transition, class-switching and SHM patterns

We next integrated single-cell BCR sequencing data onto the B cell transcriptional landscape to relate clonal expansion, SHM and isotype distribution to different B cell phenotypes. For this analysis we identified 11,227 individual B cells with available BCR and transcriptomic information. We observed a generally low level of B cell clonal expansion, with preferential localization of expanded B cell clonotypes (2-4 cells) to the memory and MZ B cell regions and a rare occurrence of highly expanded (≥ 5 cells) clonotypes in most patients (**Fig. 5a**). Analysis of differential gene expression showed that expanded clonotypes had increased expression of genes involved in cytoskeleton reorganization (VIM) and genes associated with the unfolded protein response (HSBP90, CALR, PPIB), indicating B cell activation and transition into plasma cells, respectively ⁴⁸⁻⁵⁰ (**Fig. 5b**). Furthermore, expanded B cell clonotypes across patients showed downregulation of MHC class II genes (CD74, HLA-DR, -DQA1, -DRB1), further supporting their trajectory towards antibody-producing plasma cells ⁵¹.

We next analyzed single-cell BCR sequencing data to assess SHM levels in different B cell subsets. BCRs from memory B cells showed the highest levels of SHM across patients, while BCRs extracted from naïve and activated B cells displayed similarly low median SHM values. Notably, however, activated B cells expressed a larger number of high-SHM outliers than naïve B cells, suggesting ongoing affinity maturation in this subset (**Fig. 5c**). Mapping of Ig isotype information onto the B cell transcriptional UMAP space, revealed an even distribution of IgM expression across B cell subsets, rare occurrence of IgD-expressing B cells and, most notably, confinement of class-switched IgG- and IgA-expressing B cells to the memory and MZ B cell regions (**Fig. 5d**). Finally, assessment of Ig isotype distribution across patients and B cell subsets revealed that as expected the vast majority of naïve B cells expressed the IgM isotype (**Fig. 5e**), with minimal levels of class-switching observed in activated B cells but prominent class-switching to IgG and IgA isotypes in the memory B cell compartment across all patients (**Fig. 5e**).

Together our results indicate ongoing transition of clonally-expanded B cells into antibody-producing plasma cells, as well as high levels of SHM and Ig isotype class-switching in the memory B cells of convalescent COVID-19 patients.

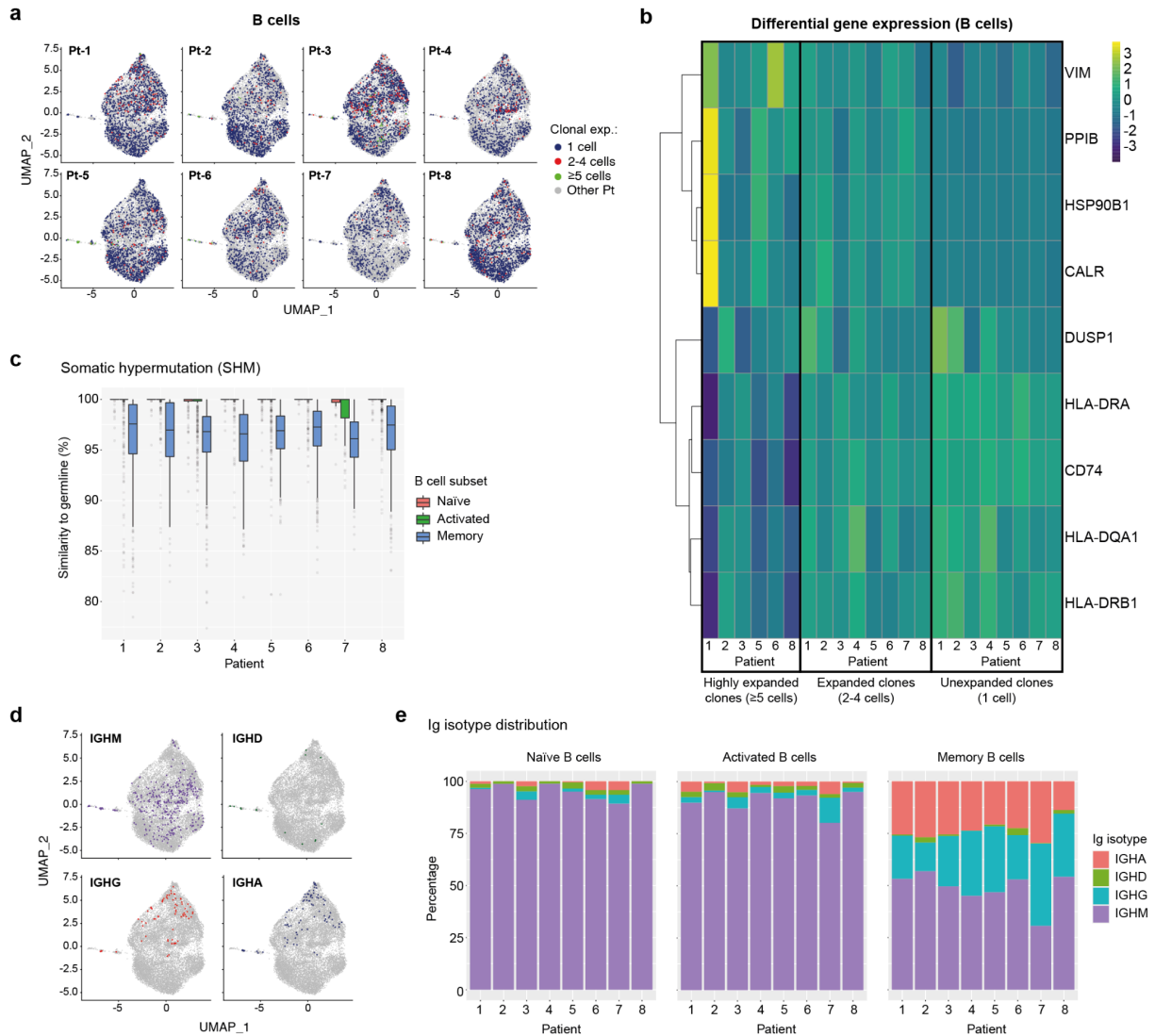


Figure 5. Single-cell transcriptome and BCR profiling reveals elevated class-switching and somatic hypermutation levels in memory B cells. **a**, UMAP plots display B cells from specific patients according to their clonal expansion levels. B cells from other patients in each individual plot are shown in gray. **b**, Heatmap shows differential gene expression (DGE) in unexpanded, expanded and highly expanded B cells. Genes were filtered to include those with detectable expression in at least 50% of cells and that had a minimum 50% fold-change in expression level between groups. **c**, Graph displays the levels of somatic hypermutation (SHM) in the BCRs of naïve, activated and memory B cells across patients. SHM levels are based on the percentage similarity between BCR heavy chain V-gene and its corresponding germline. Data are displayed as median \pm IQR. **d**, Graph shows the distribution of B cells expressing specific Ig isotypes relative to their location in transcriptome UMAP plots. B cells from all patients are shown. **e**, Bar graphs show Ig isotype distribution of BCRs found in naïve, activated and memory B cells across patients. CALR = Calreticulin; CD74 = HLA class II Histocompatibility Antigen Gamma Chain; DUSP1 = Dual Specificity Phosphatase 1; HLA-D = Major Histocompatibility Complex, Class II; HSP90B1 = Heat Shock Protein 90 Beta Family Member 1; PPIB = Peptidylprolyl Isomerase B; VIM = Vimentin.

Computational prediction of shared specificity identifies candidate SARS-CoV-2-specific TCRs

Motivated by the high levels of CD8⁺ T cell clonal expansion and activation observed in convalescent COVID-19 patients, we further analyzed single-cell TCR repertoires for potential SARS-CoV-2 specificity. To this end, we applied GLIPH2, an algorithm developed by M. Davis and colleagues that clusters TCRs with a high probability of recognizing the same epitope into specificity groups (based on conserved motifs and similarity levels in CDR3 β)⁵². In addition, the provision of HLA typing data enables the prediction of HLA restriction in specific TCR clusters. Analysis of 23,010 paired TCR α and TCR β sequences derived from the CD8⁺ T cells of eight patients led to the identification of a total of 552 specificity groups with attributed HLA restriction (seven alleles). We observed distinct proportions of shared specificity groups between pairs of patients, with Pt-1:Pt-8, Pt-3:Pt-7 and Pt-1:Pt-7 showing the highest overlap (**Fig. 6a**). Furthermore, the vast majority of clusters were attributed with HLA-A*01:01, HLA-A*03:01, HLA-B*13:02 or HLA-C*03:04 restriction, and HLA-A*02:01, A*24:02 and C*04:01 were attributed to less than 15% of TCR clusters (**Fig. 6b**). While some of these clusters may be defined by SARS-CoV-2 specificity, it is difficult to exclude reactivity to common human viruses (e.g., CMV, EBV). To further investigate potential for SARS-CoV-2 specificity, we analyzed the sequences of known HLA-A*02:01-restricted SARS-CoV-2-specific, CMV-specific and EBV-specific TCRs alongside those derived from patients expressing the HLA-A*02:01 allele (i.e., Pt-2, Pt-4 and Pt-8) (**Supplemental Tables 3 and 4**). GLIPH2 identified 35 unique patient TCR sequences that clustered together with known SARS-CoV-2-specific TCRs, with the great majority originating from Pt-8 (**Fig. 6c and Table 1**). Thus, such TCRs represent candidates for mediating CD8⁺ T cell immunity against SARS-CoV-2 infection.

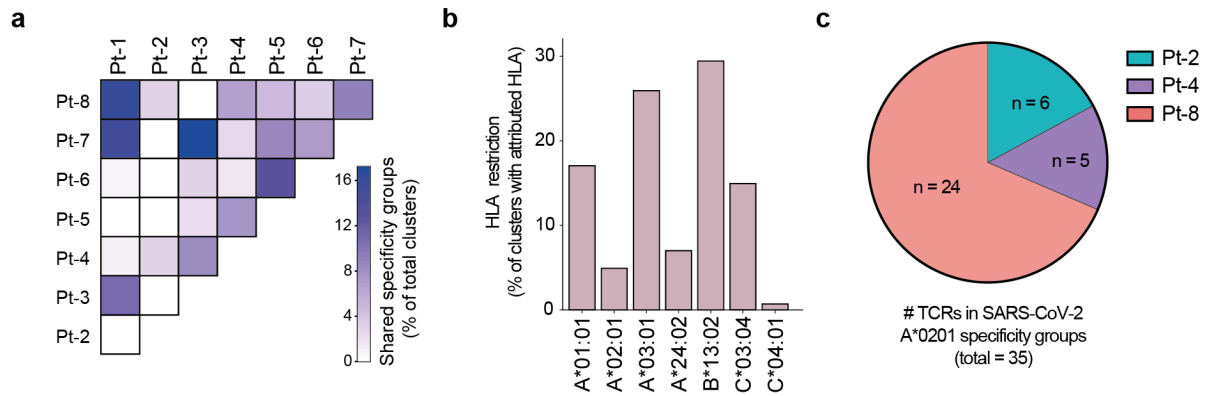


Figure 6. GLIPH2 analysis of single-cell paired TCR repertoires reveals candidate SARS-CoV-2-specific TCRs. **a**, Heatmap shows the proportion of TCR specificity groups containing sequences from specific pairs of patients, as determined by GLIPH2 analysis (total TCR clusters = 552). **b**, Bar plot displays the proportions of predicted HLA class I alleles in HLA-attributed TCR specificity groups (total TCR clusters = 552). **c**, Graph displays the proportions and numbers of candidate SARS-CoV-2-specific TCRs derived from HLA-A*02:01-positive patients, as determined by GLIPH2 clustering with known SARS-CoV-2-specific TCR sequences.

Limitations of the study

Our sample size of eight patients (four per age group) is small and reduces the number of conclusions that we can confidently make from the observed data. Nevertheless, single-cell data offers a deeper characterization of each patient than normal bulk transcriptome and repertoire studies. Furthermore, while the time between symptom onset and sample collection is highly uniform across patients, the time between symptom resolution and sample collection is significantly shorter in the old patient group due to prolonged symptom duration. This is an important variable that should be considered when interpreting the findings presented here.

2.4 Discussion

Here we apply scSeq for in-depth immune repertoire and transcriptomic analysis of T cells and B cells derived from non-severe COVID-19 patients at one month of convalescence. Our analyses of transcriptomic data defined eleven T cell and B cell subsets, of which the effector CD8⁺ T cell subset (GZMB, NKG7) showed the highest levels of expansion in specific patients, both in terms of proportion and clonality. These findings are in agreement with recent scSeq studies of convalescent COVID-19 patients^{28,53}. For example, effector tissue-resident CD8⁺ T cells from bronchoalveolar lavage fluid were found to be highly clonally expanded in convalescent COVID-19 patients that experienced moderate but not severe infection⁵³. In addition, scSeq of PBMCs revealed increases in cytotoxic effector CD4⁺ and CD8⁺ T cell subsets in non-severe convalescent COVID-19 patients only²⁸. In our study, we observed high levels of clonal expansion of the effector CD8⁺ T cell subset but less evident expansion of CD4⁺ T cell subsets. Importantly, however, differential gene expression analysis revealed that both highly clonally expanded CD8⁺ and CD4⁺ T cells had elevated markers of cytotoxicity (CD8: PRF1, GZMH, GNLY; CD4: GZMA).

A recent study of functional T cell responses against SARS-CoV-2 reported significantly higher CD8⁺ T cell responses directed at spike, M/NP and ORF/Env epitopes in convalescent COVID-19 patients experiencing moderate symptoms compared to those recovering from severe infection⁵⁴. Further evidence supporting a potential role of CD8⁺ T cells in rapid viral clearance comes from the occurrence of SARS-CoV-2-specific T cells in asymptomatic seronegative family members of COVID-19 patients⁵⁵, as well as in samples from asymptomatic seronegative control subjects obtained during the COVID-19 pandemic but not prior to it^{55,56}. In line with these findings, our analysis of clonal expansion in lymphocyte subsets suggest a key role of CD8⁺ effector T cells in the clearance and protection against SARS-CoV-2 in patients with moderate disease. Notably, the fact that younger patients in our cohort had significantly higher CD8-to-CD4 T cell ratios might have contributed to reduced symptom duration relative to older patients. In this context, we predict that methods for the identification of CD8-derived SARS-CoV-2-specific TCRs including functional assays^{54,55} and the application of motif clustering⁵⁶ or machine learning⁵⁷ to TCR repertoire data, as well as methods for the

identification of their corresponding epitopes^{58,59} will become increasingly important tools for monitoring SARS-CoV-2 immunity.

In contrast to T cells, we found modest levels of B cell clonal expansion that were not exclusively restricted to a particular subset but spanned activated, memory and MZ B cells. These low levels of B cell clonal expansion are in contrast to the high IgG and IgA serum titers found in all patients of our cohort, particularly those in the older patient group. This suggests that the B cell compartment may experience clonal contraction at one month of convalescence, with the disconnect from serum titers of IgG likely explained by the long half-lives of secreted IgG (~3 weeks)⁶⁰. It should also be noted that antibody-producing plasma cells, which were not included in our analysis, may display clonal expansion levels that are more congruent with the observed levels of SARS-CoV-2-specific antibodies in serum. Although rare, highly expanded B cell clones showed elevated markers of plasma cell transition and activation, thus indicating possible ongoing differentiation into plasma cells, albeit at low levels, at one month after symptom onset.

Analysis of BCR repertoires and transcriptomes revealed that the highest levels of SHM and Ig class-switching occurred in the memory B cell subset. This finding is consistent with the generation of germinal center-derived memory B cells following antigen encounter^{61,62}. Similarly, SHM and Ig class-switching levels were directly correlated with B cell clonal expansion. Remarkably, we found that highly expanded B cell clones from the young patient group had significantly higher SHM levels (median = 6.7% divergence from germline) than those derived from the old patient group (median = 2.5% divergence from germline). This occurred despite significantly higher serum titers of SARS-CoV-2-specific antibodies in the old patient group, and suggests that more effective affinity maturation may occur in younger patients. Initial studies characterizing SARS-CoV-2-specific antibodies reported low levels of SHM, with median divergence from germline ranging from 0.7-2% in convalescent patients of varied symptom severity analyzed at 20-40 days following symptom onset⁶³⁻⁶⁵. Notably, a recent report has described ongoing affinity maturation of SARS-CoV-2-specific antibodies at six months following symptom onset in non-severe patients, with SHM levels rising to 3% divergence from germline⁶⁶. Interestingly, the cited study provides evidence of viable SARS-CoV-2 antigen in the gut of such patients, which has been proposed as a source of antigen

for ongoing affinity maturation linked to elevated IgA serum titers. It is thus unclear whether increased SHM levels found in the young patients of our cohort are the result of a better capacity for affinity maturation upon initial antigen encounter or of ongoing affinity maturation resulting from longer exposure to antigen following symptom resolution. The clear dominance of IgA class-switched BCRs expressed by highly expanded B cells in 3 out of 4 patients in the young group appears to support the latter.

In conclusion, our in-depth characterization integrating single-cell immune repertoire and transcriptome profiling of T and B cells represents a valuable resource to better understand the adaptive immune response and age-related differences in convalescent COVID-19 patients with moderate disease. Furthermore, it serves as an important point of reference for future single-cell characterization of lymphocytes at later time points of convalescence, or of lymphocytes isolated from patients experiencing long-term COVID-19 sequelae and can help in defining markers for clinical monitoring of disease progression.

Supplementary information

Supplemental Figures 1-7

Supplemental Tables 1-5

2.5 Methods

Patient samples

Patients were participants of the SERO-BL-COVID-19 study sponsored by the Department of Health, Canton Basel-Landschaft, Switzerland. All analyzed patients tested positive for SARS-CoV-2 after RT-PCR of naso- and oropharyngeal swab samples and experienced a resolution of COVID-19 symptoms without requiring hospitalization. Whole blood was collected 25 to 39 days following a positive RT-PCR test and subjected to density gradient centrifugation using the Ficoll Paque Plus reagent (GE Healthcare, #17-1440-02). After separation, the upper plasma layer was collected for ELISA detection of IgG and IgA SARS-CoV-2-specific antibodies (Euroimmun Medizinische Labordiagnostika, #EI2668-9601G, #EI2606-9601A). Peripheral blood mononuclear cells (PBMC) were collected from the interphase, resuspended in freezing medium (RPMI 1640, 10%(v/v) FBS, 10%(v/v) dimethyl sulfoxide) and cryopreserved in liquid nitrogen. Point-of-care lateral flow immunoassays assessing the presence of IgG and IgM SARS-CoV-2-specific antibodies (Qingdao Hightop Biotech, #H100) were performed at the time of blood collection.

Immunomagnetic isolation of B cells and T cells

PBMC samples were thawed, washed in complete media (RPMI 1640, 10%(v/v) FBS) and pelleted by centrifugation. Cells were resuspended in 0.5 mL complete media, counted and treated with 10 U ml⁻¹ DNase I (Stemcell Technologies, #) for 15 min at RT in order to prevent cell clumping. After DNase I digestion, cells were washed twice in complete media, pelleted by centrifugation and resuspended in 0.5 mL flow cytometry buffer (PBS, 2%(v/v) FBS, 2 mM EDTA). The cell suspension was filtered through a 40 µM cell strainer prior to immunomagnetic isolation. As a first step, plasma cells were isolated using the EasySep Human CD138 Positive Selection Kit II (Stemcell Technologies, #17877) for analysis in a companion study (manuscript in preparation). The negative fraction of the above selections was divided into two aliquots that were subjected to negative immunomagnetic isolation of either B cells (EasySep Human Pan-B cell Enrichment Kit, Stemcell Technologies, #19554) or T cells (EasySep Human T cell Isolation Kit, Stemcell Technologies, #17951). After isolation, B cells and T cells were pelleted by

centrifugation, resuspended in PBS, 0.4%(v/v) BSA, filtered through a 40 μ M cell strainer and counted. T cells and B cells originating from the same patient were pooled in equal numbers and the final suspension was counted and assessed for viability using a fluorescent cell counter (Cellometer Spectrum, Nexcelom). Whenever possible, cells were adjusted to a concentration of 1×10^6 live cells/mL in PBS, 0.04%(v/v) BSA before proceeding with droplet generation.

Single cell droplet generation and preparation of sequencing libraries

Encapsulation of lymphocytes and DNA-barcoded gel beads was performed using the Chromium controller (10x Genomics, PN-110203). Briefly, 1.4×10^4 to 1.7×10^4 cells (in reverse transcription mix) were loaded per channel for a targeted recovery of 8×10^3 to 1×10^4 cells per sample. Reverse transcription and preparation of single-cell transcriptome, BCR and TCR libraries was performed according to the manufacturer's instructions (CG000086 manual, RevM, 10x Genomics) and using the following kits: Chromium Single Cell 5' Library & Gel Bead Kit (PN-1000006), Chromium Single Cell 5' Library Construction Kit (PN-1000020), Chromium Single Cell V(D)J Enrichment Kit, Human T Cell (PN-1000005), Chromium Single Cell V(D)J Enrichment Kit, Human B Cell (PN-1000016), Chromium Single Cell A Chip Kit (PN-1000009), Chromium i7 Multiplex Kit (PN-120262).

Deep sequencing

The quality and concentrations of transcriptome (i.e., cDNA), TCR and BCR libraries were determined using a fragment analyzer (Agilent Bioanalyzer) at specific steps of library preparation, as recommended in the 10x Genomics scSeq protocol (CG000086 manual, RevM). Following multiplexing (Chromium i7 Multiplex Kit, #PN-120262, 10x Genomics), transcriptome libraries were treated with free adapter blocking (FAB) reagent to prevent index switching (#20024144, Illumina). Paired-end sequencing of multiplexed transcriptome libraries was performed using a NovaSeq 6000 sequencer (Illumina) and SP100-cycle kit (#20027464, Illumina). TCR and BCR libraries were multiplexed, FAB-treated and paired-end-sequenced using a second SP100-cycle kit in a separate run.

HLA class I typing

HLA class I transcripts were amplified and deep-sequenced from two overlapping RT-PCR reactions flanking exons 2 (PCR 1) or 3 (PCR 2) using barcoded primers designed to target conserved regions (**Supplemental Table 5**)⁶⁷. Total RNA was extracted from patient PBMCs by resuspension in TRIzol reagent (Invitrogen, # 15596018), and column-purified using the PureLink RNA Mini kit (Invitrogen, #12183025). For reverse transcription, 100 pmol of oligo dT, 10 nmol of each dNTP, 40 ng RNA and sufficient nuclease-free water for a final 14 μ L volume were mixed, incubated at 65°C for 5 min and chilled on ice for 5 min. This was followed by addition of 4 μ L of 5X RT buffer, 40 units of RiboLock RNase inhibitor (Thermo Fisher, #EO0381) and 200 units of Maxima H-minus reverse transcriptase (Thermo Fisher, #EP0751) and mixing. Reverse transcription was performed at 50°C for 30 min, followed by inactivation at 85°C for 5 min. 5 μ L of the resulting cDNA-containing reverse transcription mixes were then used as templates for 25 μ L PCR reactions using the KAPA HiFi PCR kit with GC buffer (Roche Diagnostics, #07958846001) and the following thermal cycling conditions: 95°C for 3 min; 35 cycles of 98°C for 20 s, 61°C for 15 s, 72°C for 15 s; and final extension at 72°C for 30 s. HLA amplicons were purified by gel-extraction (QIAquick gel extraction kit, Qiagen #28704) and submitted for Illumina paired-end deep-sequencing (Amplicon-EZ, Genewiz). Unique sequences originating from specific patients were identified from their respective DNA barcodes and aligned using the ClustalOmega tool to cluster sequences arising from the same allele. Sequences with the highest amount of reads in each cluster were used as input for the basic local alignment search tool (BLAST; Nucleotide collection, *Homo sapiens*). Sequences returning matching or highly similar alleles across PCR 1 and PCR 2 in each patient were then assembled and queried against the IMGT/HLA database for final validation.

Transcriptome scSeq alignment and quality control (QC)

Reads from transcriptome scSeq (FASTQ format) were aligned to the GRCh38 reference human genome and output as filtered gene expression matrices using the 10x Genomics Cell Ranger software (version 3.1.0). Subsequent data QC and analysis was performed using R (version 3.6.2) and the Seurat package (version 3.1.5). QC steps consisted of the exclusion of TCR and BCR genes (prevention of clonotype influence on subsequent clustering), the exclusion of cells with lower than 150 or greater than 3500 genes (low quality cells), and the exclusion of cells in

which more than 20% of UMIs were associated with mitochondrial genes (reduction of freeze-thaw metabolic effects) ⁶⁸.

Dataset normalization and integration of multiple datasets

Patient datasets were merged into a Seurat object list using the *merge* and *SplitObject* function. Each patient dataset was then separately normalized using *SCTransform*. Variable integration features (3,000) were calculated using the *SelectIntegrationFeatures* function from the R package Seurat ⁶⁹ and setting them as variable features after merging the normalized patient datasets. Principal component analysis for dimensionality reduction was performed using the *RunPCA* function with up to 50 principal components. Potential batch effects between patient samples were addressed with the Harmony R package (version 1.0) using the *RunHarmony* function ⁷⁰. Finally, unsupervised clustering was performed using the *FindNeighbours* and *FindClusters* functions. Non-linear dimensionality reduction using the *RunUMAP* function was performed using the first 50 principal components to generate the final UMAP visualization of cell clusters.

Dataset subsetting of CD8⁺ T cells, CD4⁺ T cells and B cells

Initial T and B cell separation was performed by mapping of TCR and BCR (VDJ) cell-specific barcodes onto the scSeq transcriptome dataset. Double attribution of TCR and BCR to the same cell (i.e., barcode) was used to identify and exclude doublets. Separation of CD8⁺ and CD4⁺ T cells was performed using the *WhichCells* function, from the R package Seurat, based on the singular expression of CD8A and CD4, respectively. Additional filtering of B cells was done by discarding all B cells that showed expression of CD3E or SDC1 as well as excluding B cells whose cellular barcodes occurred in the plasma cell BCR (VDJ) cell barcodes (data not shown).

Cell state annotation and marker identification

The expression of specific markers in identified clusters was determined using the *FindAllMarkers* function using the Wilcoxon Rank Sum test. Cluster-specific markers were thresholded by having a log₂(fold-change) greater than 0.25 between cells in the respective cluster and remaining cells; with marker expression occurring in at least 25% of cells in the

cluster. Clusters were then attributed with specific cell states based on the expression of canonical markers

Differential gene expression analysis

Differentially expressed genes between two groups of cells were identified using the *FindMarkers* function. Genes were thresholded by being expressed in more than 50% of the cells and by having a $\log_2(\text{fold-change})$ greater than 0.5 between cells of the different groups using the Wilcoxon Rank Sum test.

Paired TCR and BCR (VDJ) single-cell sequencing alignment and QC

TCR and BCR reads in FASTQ format were aligned with the VDJ-GRCh38-alts-ensembl reference using the 10x Genomics Cell Ranger VDJ software (version 3.1.0). This generated single-cell VDJ sequences and annotations such as gene usage, clonotype frequency and cell-specific barcode information. As a QC step, only cells with one productive α - and one productive β -chain (T cells) or with one productive heavy and one productive light chain (B cells) were retained for downstream analysis.

Paired TCR and BCR (VDJ) analysis

Clonotype definition was adjusted to count all sequences as clonal if they met the following criteria: (1) Same V and J gene usage in both chains, (2) Same CDR3 length in both chains and (3) 80% amino acid sequence similarity in the CDR3 region of the TCR β (T cells) or BCR heavy chain (B cells). Shared cellular barcode information between TCR/BCR (VDJ) scSeq and transcriptome scSeq data was used to project TCR and BCR clonotypes onto the UMAP plots (color-coded by clonal expansion level).

Somatic hypermutation analysis

SHM levels in individual BCR clonotypes were determined using the change-o toolkit from the Immcantation portal as a wrapper to run IgBlast on the Cell Ranger VDJ output. The IgBlast output enabled assessment of germline similarity of single-cell BCR (VDJ) sequences. Germline identity was used as a proxy for somatic hypermutation levels and was calculated from alignments of BCR clonotypes with their corresponding VH and VL germline sequences.

TCR specificity group identification using GLIPH2

GLIPH2 clusters TCRs into specificity groups predicted to share the same antigen specificity based on sequence similarity⁵². We used this algorithm to cluster TCRs from HLA-A*0201 patients (Pt-2, Pt-4 and Pt-8) together with known SARS-CoV-2 binders as well as CMV and EBV binders, which were also from HLA-A*0201 background (obtained from VDJdb database). Specificity groups that were reported by GLIPH2, were filtered for groups that were significant according to the Fisher's Exact test (significance level < 5%) and contained at least one patient TCR and one TCR of known specificity (i.e., SARS-CoV-2, EBV or CMV). Specificity groups were identified with either global (0-1 amino acid differences in CDR3 β) or local similarities (CDR3 β share a common motif that is rare in the reference dataset). CD4 expressing clonotypes were filtered out.

Pseudotime analysis

Pseudotime and trajectory inference was applied to scSeq transcriptome data using the *slingshot* function with default parameters from the Slingshot package in R⁷¹. The naive cluster from each CD8⁺ T cell, CD4⁺ T cell and B cell subgroup was set as the starting point for the minimum spanning tree. The previously generated UMAP clustering was set as the cellular embedding on which Slingshot performed trajectory inference computation.

Code availability

The data analysis pipeline followed the standard procedures as outlined in Cell Ranger and Seurat documentations. Custom scripts and functions for easier downstream analysis and visualization purposes are available upon request.

List of utilized R packages

Biobase (2.46.0), BiocGenerics (0.32.0), BiocParallel (1.20.1), Cell Ranger (3.1.0), Change-O (1.0.0), circlize (0.4.10), data.table (1.12.8), DelayedArray (0.12.3), dplyr (0.8.5), GenomeInfoDb (1.22.1), GenomicRanges (1.38.0), ggplot2 (3.3.2.9000), harmony (1.0), pheamap (1.0.12), printrcurve (2.1.5), RColorBrewer (1.1-2), matrixStats (0.56.0), sctransform (0.2.1), Seurat (3.1.5), slingshot (1.4.0), stringdist (0.9.5.5), stringr (1.4.0), tibble (3.0.3), tidyr (1.1.0), tidyverse (1.3.0).

Acknowledgements

We thank the Genomics Facility Basel (D-BSSE, ETH Zurich), in particular Ms. Ina Nissen and Dr. Christian Beisel, for their excellent support with the scSeq protocol and for performing deep sequencing runs.

Funding

This study is supported by funding from the Personalized Health and Related Technologies Postdoctoral Fellowship (to R.V.-L), the NCCR Molecular Systems Engineering (to S.T.R.), Helmut Horten Stiftung (to S.T.R.), Botnar Research Centre for Child Health (to S.T.R.).

Declaration of interests

The authors declare no competing interests

Data and Materials Availability

The raw FASTQ files from deep sequencing that support the findings of this study are deposited under E-MTAB-10169. Additional data that support the findings of this study are available from the corresponding author upon reasonable request.

2.6 References of Chapter 2

1. Dan JM, Mateus J, Kato Y, Hastie KM, Yu ED, Faliti CE, Grifoni A, Ramirez SI, Haupt S, Frazier A, et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* (2021) doi:10.1126/science.abf4063
2. Gaebler C, Wang Z, Lorenzi JCC, Muecksch F, Finkin S, Tokuyama M, Cho A, Jankovic M, Schaefer-Babajew D, Oliveira TY, et al. Evolution of antibody immunity to SARS-CoV-2. *Nature* (2021) doi:10.1038/s41586-021-03207-w
3. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, Rawlings SA, Sutherland A, Premkumar L, Jadi RS, et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* (2020) **181**:1489–1501.e15. doi:10.1016/j.cell.2020.05.015
4. Liu L, Wang P, Nair MS, Yu J, Rapp M, Wang Q, Luo Y, -W. Chan JF, Sahi V, Figueroa A, et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* (2020) **584**:450–456. doi:10.1038/s41586-020-2571-7
5. Juno JA, Tan H-X, Lee WS, Reynaldi A, Kelly HG, Wragg K, Esterbauer R, Kent HE, Batten CJ, Mordant FL, et al. Humoral and circulating follicular helper T cell responses in recovered patients with COVID-19. *Nat Med* (2020) **26**:1428–1434.
6. Altmann DM. Adaptive immunity to SARS-CoV-2. *Oxford Open Immunology* (2020) **1**: doi:10.1093/oxfimm/iqaa003
7. Rees AR. Understanding the human antibody repertoire. *MAbs* (2020) **12**:1729683.
8. Rosati E, Dowds CM, Liaskou E, Henriksen EKK, Karlsen TH, Franke A. Overview of methodologies for T-cell receptor repertoire analysis. *BMC Biotechnol* (2017) **17**:61.
9. Briney B, Inderbitzin A, Joyce C, Burton DR. Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature* (2019) **566**:393–397. doi:10.1038/s41586-019-0879-y
10. Soto C, Bombardi RG, Branchizio A, Kose N, Matta P, Sevy AM, Sinkovits RS, Gilchuk P, Finn JA, Crowe JE. High frequency of shared clonotypes in human B cell receptor repertoires. *Nature* (2019) **566**:398–402. doi:10.1038/s41586-019-0934-8
11. DeWitt WS, Smith A, Schoch G, Hansen JA, Matsen FA, Bradley P. Human T cell receptor occurrence patterns encode immune history, genetic background, and receptor specificity. *eLife* (2018) **7**: doi:10.7554/elife.38358
12. Greiff V, Menzel U, Miho E, Weber C, Riedel R, Cook S, Valai A, Lopes T, Radbruch A, Winkler TH, et al. Systems Analysis Reveals High Genetic and Antigen-Driven Predetermination of Antibody Repertoires throughout B Cell Development. *Cell Reports* (2017) **19**:1467–1478. doi:10.1016/j.celrep.2017.04.054
13. Emerson RO, DeWitt WS, Vignali M, Gravley J, Hu JK, Osborne EJ, Desmarais C, Klinger M, Carlson CS, Hansen JA, et al. Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire. *Nat Genet* (2017) **49**:659–665.
14. Galson JD, Trück J, Fowler A, Clutterbuck EA, Münz M, Cerundolo V, Reinhard C, van der Most R, Pollard AJ, Lunter G, et al. Analysis of B Cell Repertoire Dynamics Following Hepatitis B Vaccination in Humans, and Enrichment of Vaccine-specific Antibody Sequences. *EBioMedicine* (2015) **2**:2070–2079.
15. Egorov ES, Kasatskaya SA, Zubov VN, Izraelson M, Nakonechnaya TO, Staroverov DB, Angius A, Cucca F, Mamedov IZ, Rosati E, et al. The Changing Landscape of Naive T Cell Receptor Repertoire With Human Aging. *Front Immunol* (2018) **9**:1618.
16. Reddy ST, Ge X, Miklos AE, Hughes RA, Kang SH, Hoi KH, Chrysostomou C, Hunnicke-Smith SP, Iverson BL, Tucker PW, et al. Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nat Biotechnol* (2010) **28**:965–969.
17. Vollmers C, Sit RV, Weinstein JA, Dekker CL, Quake SR. Genetic measurement of memory B-cell recall using antibody repertoire sequencing. *Proc Natl Acad Sci U S A* (2013) **110**:13463–13468.
18. Niu X, Li S, Li P, Pan W, Wang Q, Feng Y, Mo X, Yan Q, Ye X, Luo J, et al. Longitudinal Analysis of T and B Cell Receptor Repertoire Transcripts Reveal Dynamic Immune Response in COVID-19 Patients. *Frontiers in Immunology* (2020) **11**: doi:10.3389/fimmu.2020.582010
19. Macosko EZ, Basu A, Satija R, Nemes J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* (2015) **161**:1202–1214.
20. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, Saadatpour A, Zhou Z, Chen H, Ye F, et al. Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell* (2018) **173**:1307.

21. Maaten L van der, van der Maaten L, Hinton G. Visualizing non-metric similarities in multiple maps. *Machine Learning* (2012) **87**:33–55. doi:10.1007/s10994-011-5273-4
22. McInnes L, Healy J, Saul N, Großberger L. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software* (2018) **3**:861. doi:10.21105/joss.00861
23. Singer M, Wang C, Cong L, Marjanovic ND, Kowalczyk MS, Zhang H, Nyman J, Sakuishi K, Kurtulus S, Gennert D, et al. A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells. *Cell* (2017) **171**:1221–1223.
24. Brummelman J, Mazza EMC, Alvisi G, Colombo FS, Grilli A, Mikulak J, Mavilio D, Alloisio M, Ferrari F, Lopci E, et al. High-dimensional single cell analysis identifies stem-like cytotoxic CD8 T cells infiltrating human tumors. *J Exp Med* (2018) **215**:2520–2535.
25. Cohn LB, da Silva IT, Valieris R, Huang AS, Lorenzi JCC, Cohen YZ, Pai JA, Butler AL, Caskey M, Jankovic M, et al. Clonal CD4 T cells in the HIV-1 latent reservoir display a distinct gene profile upon reactivation. *Nat Med* (2018) **24**:604–609.
26. Chen J, Tan Y, Sun F, Hou L, Zhang C, Ge T, Yu H, Wu C, Zhu Y, Duan L, et al. Single-cell transcriptome and antigen-immunoglobulin analysis reveals the diversity of B cells in non-small cell lung cancer. *Genome Biol* (2020) **21**:152.
27. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* (2014) **32**:381–386.
28. Zhang J-Y, Wang X-M, Xing X, Xu Z, Zhang C, Song J-W, Fan X, Xia P, Fu J-L, Wang S-Y, et al. Single-cell landscape of immunological responses in patients with COVID-19. *Nat Immunol* (2020) **21**:1107–1118.
29. DeKosky BJ, Kojima T, Rodin A, Charab W, Ippolito GC, Ellington AD, Georgiou G. In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat Med* (2015) **21**:86–91.
30. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R, Smibert P. Simultaneous epitope and transcriptome measurement in single cells. *Nature Methods* (2017) **14**:865–868. doi:10.1038/nmeth.4380
31. Howie B, Sherwood AM, Berkebile AD, Berka J, Emerson RO, Williamson DW, Kirsch I, Vignali M, Rieder MJ, Carlson CS, et al. High-throughput pairing of T cell receptor α and β sequences. *Sci Transl Med* (2015) **7**:301ra131.
32. Setliff I, Shiakolas AR, Pilewski KA, Murji AA, Mapengo RE, Janowska K, Richardson S, Oosthuysen C, Raju N, Ronsard L, et al. High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. *Cell* (2019) **179**:1636–1646.e15. doi:10.1016/j.cell.2019.11.003
33. Friedensohn S, Khan TA, Reddy ST. Advanced Methodologies in High-Throughput Sequencing of Immune Repertoires. *Trends Biotechnol* (2017) **35**:203–214.
34. Han A, Glanville J, Hansmann L, Davis MM. Corrigendum: Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat Biotechnol* (2015) **33**:210.
35. Stubbington MJT, Lönnberg T, Proserpio V, Clare S, Speak AO, Dougan G, Teichmann SA. T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods* (2016) **13**:329–332.
36. Lönnberg T, Svensson V, James KR, Fernandez-Ruiz D, Sebina I, Montandon R, Soon MSF, Fogg LG, Nair AS, Liligeto U, et al. Single-cell RNA-seq and computational analysis using temporal mixture modelling resolves Th1/Tfh fate bifurcation in malaria. *Sci Immunol* (2017) **2**: doi:10.1126/sciimmunol.aal2192
37. Xu G, Qi F, Li H, Yang Q, Wang H, Wang X, Liu X, Zhao J, Liao X, Liu Y, et al. The differential immune responses to COVID-19 in peripheral and lung revealed by single-cell RNA sequencing. *Cell Discov* (2020) **6**:73.
38. Su Y, Chen D, Yuan D, Lausted C, Choi J, Dai CL, Voillet V, Duvvuri VR, Scherler K, Troisch P, et al. Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19. *Cell* (2020) **183**:1479–1495.e20.
39. Kaltenbach H-M, Rudolf F, Linnik J, Deichmann J, Ruf T, Altamura R, Kapetanovic E, Mason D, Wagner B, Goetz T, et al. Initial characterisation of ELISA assays and the immune response of the clinically correlated SARS-CoV-2 biobank SERO-BL-COVID-19 collected during the pandemic onset in Switzerland. doi:10.1101/2020.07.05.20145888
40. Davies NG, CMMID COVID-19 working group, Klepac P, Liu Y, Prem K, Jit M, Eggo RM. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine* (2020) **26**:1205–1211. doi:10.1038/s41591-020-0962-9
41. Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, Ginhoux F, Newell EW. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* (2018) doi:10.1038/nbt.4314

42. Yan J, Greer JM, Hull R, O'Sullivan JD, Henderson RD, Read SJ, McCombe PA. The effect of ageing on human lymphocyte subsets: comparison of males and females. *Immun Ageing* (2010) **7**:4.
43. Galson JD, Schaetzle S, Bashford-Rogers RJM, Raybould MIJ, Kovaltsuk A, Kilpatrick GJ, Minter R, Finch DK, Dias J, James LK, et al. Deep Sequencing of B Cell Receptor Repertoires From COVID-19 Patients Reveals Strong Convergent Immune Signatures. *Front Immunol* (2020) **11**:605170.
44. DeKosky BJ, Ippolito GC, Deschner RP, Lavinder JJ, Wine Y, Rawlings BM, Varadarajan N, Giesecke C, Dörner T, Andrews SF, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol* (2013) **31**:166–169.
45. Wabl M, Steinberg C. Affinity maturation and class switching. *Current Opinion in Immunology* (1996) **8**:89–92. doi:10.1016/s0952-7915(96)80110-5
46. Rosa FM, Fellous M. Regulation of HLA-DR gene by IFN-gamma. Transcriptional and post-transcriptional control. *J Immunol* (1988) **140**:1660–1664.
47. Upadhyay V, Fu Y-X. Lymphotoxin signalling in immune homeostasis and the control of microorganisms. *Nat Rev Immunol* (2013) **13**:270–279.
48. Tsui C, Maldonado P, Montaner B, Borroto A, Alarcon B, Bruckbauer A, Martinez-Martin N, Batista FD. Dynamic reorganisation of intermediate filaments coordinates early B-cell activation. *Life Sci Alliance* (2018) **1**:e201800060.
49. Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, Lee A-H, Qian S-B, Zhao H, Yu X, Yang L, Tan BK, Rosenwald A, et al. XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* (2004) **21**:81–93.
50. Gass JN, Gunn KE, Sriburi R, Brewer JW. Stressed-out B cells? Plasma-cell differentiation and the unfolded protein response. *Trends Immunol* (2004) **25**:17–24.
51. Shimoda M, Li T, Pihkala JPS, Koni PA. Role of MHC Class II on Memory B Cells in Post-Germinal Center B Cell Homeostasis and Memory Response. *The Journal of Immunology* (2006) **176**:2122–2133. doi:10.4049/jimmunol.176.4.2122
52. Huang H, Wang C, Rubelt F, Scriba TJ, Davis MM. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat Biotechnol* (2020) **38**:1194–1202.
53. Liao M, Liu Y, Yuan J, Wen Y, Xu G, Zhao J, Cheng L, Li J, Wang X, Wang F, et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat Med* (2020) **26**:842–844.
54. Peng Y, Mentzer AJ, Liu G, Yao X, Yin Z, Dong D, Dejnirattisai W, Rostron T, Supasa P, Liu C, et al. Broad and strong memory CD4 and CD8 T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat Immunol* (2020) **21**:1336–1345.
55. Sekine T, Perez-Potti A, Rivera-Ballesteros O, Strålin K, Gorin J-B, Olsson A, Llewellyn-Lacey S, Kamal H, Bogdanovic G, Muschiol S, et al. Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19. *Cell* (2020) **183**:158–168.e14.
56. Shomuradova AS, Vagida MS, Sheetikov SA, Zornikova KV, Kiryukhin D, Titov A, Peshkova IO, Khmelevskaya A, Dianov DV, Malasheva M, et al. SARS-CoV-2 epitopes are recognized by a public and diverse repertoire of human T cell receptors. *Immunity* (2020) doi:10.1016/j.immuni.2020.11.004
57. Shoukat MS, Foers AD, Woodmansey S, Evans SC, Fowler A, Soilleux E. Use of machine learning to identify a T cell response to SARS-CoV-2. *Cell Reports Medicine* (2021) doi:10.1016/j.xcrm.2021.100192
58. Ferretti AP, Kula T, Wang Y, Nguyen DMV, Weinheimer A, Dunlap GS, Xu Q, Nabils N, Perullo CR, Cristofaro AW, et al. Unbiased Screens Show CD8 T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside outside the Spike Protein. *Immunity* (2020) **53**:1095–1107.e3.
59. Tarke A, Sidney J, Kidd CK, Dan JM, Ramirez SI, Yu ED, Mateus J, da Silva Antunes R, Moore E, Rubiro P, et al. Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. *Cell Reports Medicine* (2021) doi:10.1016/j.xcrm.2021.100204
60. Vidarsson G, Dekkers G, Rispens T. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* (2014) **5**:520.
61. Suan D, Sundling C, Brink R. Plasma cell and memory B cell differentiation from the germinal center. *Curr Opin Immunol* (2017) **45**:97–102.
62. Laidlaw BJ, Cyster JG. Transcriptional regulation of memory B cell differentiation. *Nat Rev Immunol* (2020) doi:10.1038/s41577-020-00446-2
63. Robbiani DF, Gaebler C, Muecksch F, Lorenzi JCC, Wang Z, Cho A, Agudelo M, Barnes CO, Gazumyan A, Finkin S, et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* (2020) **584**:437–442.

64. Rogers TF, Zhao F, Huang D, Beutler N, Burns A, He W-T, Limbo O, Smith C, Song G, Woehl J, et al. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* (2020) **369**:956.
65. Brouwer PJM, Caniels TG, van der Straten K, Snitselaar JL, Aldon Y, Bangaru S, Torres JL, Okba NMA, Claireaux M, Kerster G, et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* (2020) **369**:643–650.
66. Gaebler C, Wang Z, Lorenzi JCC, Muecksch F, Finkin S, Tokuyama M, Cho A, Jankovic M, Schaefer-Babajew D, Oliveira TY, et al. Evolution of antibody immunity to SARS-CoV-2. *Nature* (2021) doi:10.1038/s41586-021-03207-w
67. Lank SM, Golbach BA, Creager HM, Wiseman RW, Keskin DB, Reinherz EL, Brusic V, O'Connor DH. Ultra-high resolution HLA genotyping and allele discovery by highly multiplexed cDNA amplicon pyrosequencing. *BMC Genomics* (2012) **13**:378.
68. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol* (2019) **15**:e8746.
69. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive Integration of Single-Cell Data. *Cell* (2019) **177**:1888–1902.e21.
70. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh P-R, Raychaudhuri S. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* (2019) **16**:1289–1296.
71. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, Purdom E, Dudoit S. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* (2018) **19**:477.

Chapter 3: Multimodal single-cell profiling of T cell specificity and reactivity in lung cancer

This is an author-produced version of an article under review in Immunity.

Multimodal single-cell profiling of T cell specificity and reactivity in lung cancer

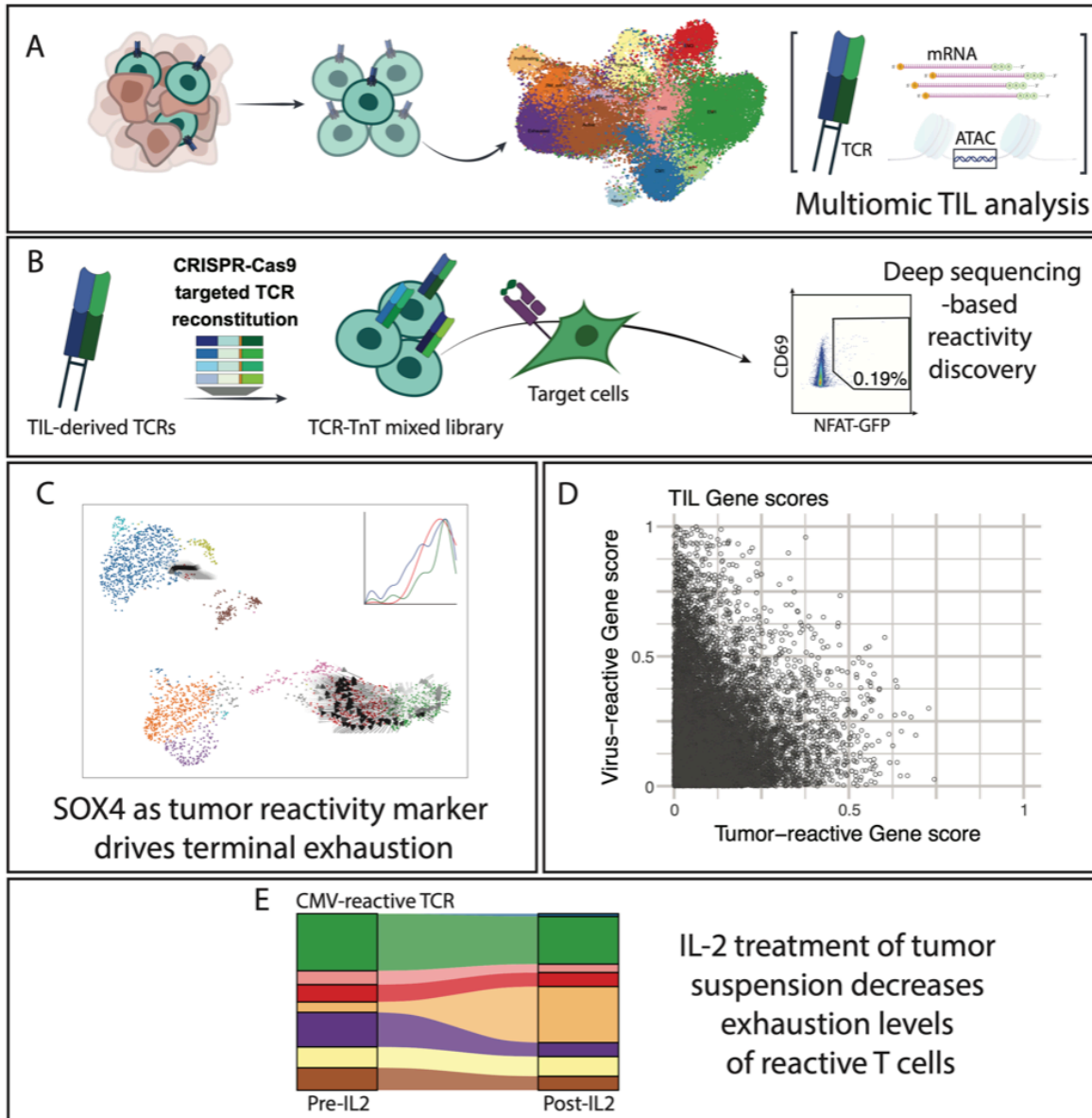
[Florian Bieberich, Rodrigo Vazquez-Lombardi, Huixin Jin, Kai-Lin Hong, Petra Herzig, Marcel Trefny, Marta Trüb, Heinz Läubli, Didier Lardinois, Kirsten Mertz, Matthias S. Matter, Alfred Zippelius and Sai T. Reddy]

F.B., R.V.-L. and S.T.R. designed the study; F.B., R.V.-L., M.T., M.T., M.M., A.Z. and S.T.R. contributed to experimental design; F.B., R.V.-L., K.-L.H., P.H. performed experiments; F.B. and H.J. analyzed data; F.B. and S.T.R. wrote the manuscript with input from all authors.

3.1 Summary

Adoptive transfer of autologous tumor-infiltrating lymphocyte T cells (TILs) offers one of the most promising approaches for cancer immunotherapy. However, high variability in patient responses highlight the need for an enhanced understanding of the transcriptional phenotypes of TILs and reactivity of their T cell receptors (TCR). Here, we employ single-cell multiomics approaches and TCR functional screening to investigate TILs from treatment-naive non-small cell lung cancer patients. This comprehensive analysis integrates scRNA-seq, scTCR-seq, and scATAC-seq, enabling a high-resolution examination of TILs within lung cancer tissue, as well as the adjacent non-tumor tissue. We apply a cellular functional screening platform to identify reactive TCRs that represent >1,000 TILs and have specificity towards a multitude of targets, including primary tumor cells, neoantigens, tumor-associated antigens, and viral antigens. Tumor-reactive TILs were primarily associated with dysfunctional phenotypes, whereas viral antigen-reactive TCRs were found in effector phenotype clusters. Key marker genes were identified and used to construct a tumor or viral reactivity score. Comparing clones shared in tumor and non-tumor tissue, a higher fraction of exhausted cells was observed in the tumor tissue, whereas non-tumor adjacent tissue possessed more effector cells, thus providing insight

into potential sources for therapeutic T cells. Elucidating the specific T cell populations within TILs and their associated TCRs may support strategies to enhance the efficacy of TIL-based therapies.



Graphical Abstract: Multimodal single cell profiling and reactivity testing of TILs.

(A) CD8⁺ T cells of treatment naive non-small cell lung cancer patients and adjacent lung tissue were isolated by fluorescence-activated cell sorting (FACS) and were then subjected to scRNA-seq + scTCR-seq or scATAC-seq. (B) TCRs were functionally screened using a cellular platform (TnT cells) and target cells (tumor cells, antigen-pulsed antigen-presenting cells, PBMCs) by flow cytometry and deep sequencing. (C) scRNA-seq + scATAC-seq allowed trajectory inference of transcription factors and genes along pseudotime. (D) Gene scores for tumor- and virus-reactivity were developed by combining functional reactivity and transcriptomic profiling for each CD8⁺ T

cell. (E) TIL scRNA-seq pre and post IL-2 treatment in tumor suspension displayed as alluvial plot shows change of clonal cell state composition.

3.2 Introduction

Lung cancer stands as the primary contributor to cancer-related fatalities worldwide and is characterized¹⁻⁴ by a high tumor mutational burden, typically exhibiting a substantial amount of immune cell infiltration within the tumor microenvironment⁵⁻⁷. Somatic mutations in the genome of cancer cells leads to the generation of neoantigens, which are mutated proteins that are processed and presented as peptides on human leukocyte antigen / major histocompatibility class I receptors (HLA-I / MHC-I) on tumor cells. Together with tumor-associated antigens (TAA), neoantigens form the core mechanism for tumor cell recognition by cytotoxic CD8 T cells via their T cell receptors (TCR)^{8,9}. Despite these defense mechanisms, a significant proportion of T cells present in a tumor (also referred to as tumor-infiltrating lymphocytes, TILs) exist in a dysfunctional state, rendering them ineffective in halting tumor progression^{10,11,12-15}.

TIL therapy consists of an expanded autologous cell product and when applied for treatment of melanoma has been associated with complete and lasting responses in patients¹⁶⁻²⁰. This outcome is largely due to a subset of tumor-reactive T cells that are capable of driving tumor eradication²¹⁻²³. However, TIL therapy is only beneficial for a limited number of patients, as often the expansion and reactivity of the TIL products are insufficient, and there is a high prevalence of dysfunctional cell states²⁴. This is likely due to the condition of T cells upon resection and an imbalanced clonal expansion during the expansion phase²⁵, underlining the correlation between T cell phenotypes and the effectiveness of TIL therapy.

T cells reacting to neoantigens provide a highly specific way to target tumor cells while obviating healthy tissue²⁶⁻²⁸. However, identifying specific peptide-HLA-TCR (pHLA-TCR) pairs is a complex task due to the high diversity of peptides and TCRs within the tumor tissue, as well as the relatively weak (low affinity) interactions between TAA pHLA-TCR pairs²⁹⁻³¹. Tumor cells can present a large number of peptides that can potentially be recognized by TCRs, including unique neoantigens and over-expression of TAA, which arise from direct (mutation) or indirect (i.e., post-translational changes) changes in the cancer genome. Besides the large number

of potential pHLA-TCR that need to be screened, the low affinity pHLA-TCR interactions pose the need for highly sensitive assays^{32,33}. While computational methods have been used to predict peptide-HLA binding (e.g., netMHC)^{34,35} and decrease the number of peptides that need to be screened, discovery of reactive TCRs remains challenging^{30,36}. Current methods for TCR reactivity discovery rely on different selection methods, such as binding based detection through T cell labeling with soluble and fluorescently-tagged pHLA multimers^{29,37,38} or, more recently through co-culturing T cells with antigen presenting cells (APCs) that display tumor antigens through peptide pulsing or minigene expression^{39,40} and using T cell activation markers or proliferation as readouts⁴¹⁻⁴³. Essential for the high-throughput nature of these assays are TCR repertoire profiling methods such as single-cell and bulk deep sequencing, which facilitate the sequencing of paired ($V\alpha$ and $V\beta$) or single TCR chains, respectively^{44,45}. The TCR repertoire is the collection of TCRs present in an individual, tissue, or co-culture setting and can be associated with tumor mutational load and clinical outcome⁴⁶.

Single-cell RNA sequencing (scRNA-seq) offers the ability to analyze the transcriptional profiles of TILs and combine it with single-cell sequencing of TCRs (scTCR-seq). Innovative methods, such as single-cell multiomics (sc-Multiomics) (simultaneous scRNA-seq and single-cell Assay for Transposase-Accessible Chromatin sequencing [scATAC-seq]), allow for a more detailed assessment of cell state dynamics and regulators⁴⁷. This involves a combined analysis of transcription factor (TF) transcriptional expression and the genomic accessibility of the TF binding site as T cells progress towards an exhausted state⁴⁸⁻⁵². Consequently, these methods offer a comprehensive insight into TIL cell state dynamics and TCR clonal repertoires and enable functional reactivity assays.

Recent work using single-cell sequencing and functional reactivity testing methods has underlined previous discoveries that neoantigen-reactive TILs are enriched in exhaustion and resident memory cell states. For example, in the studies by Caushi et al. and Lowery et al. it was discovered that only a small proportion of TILs are reactive to neoantigens and selectively express CD103, CD39, CXCL13, TOX2 and CXCR6 among others as part of their reactivity signatures^{53,54}. Further, these and similar discoveries by others, have inspired the development of tumor reactivity signatures based on differential expression of various marker genes on both RNA and protein levels^{43,55,56}. TILs often exhibit biased reactivity signaling, which can

frequently lead to false negatives in reactivity testing ⁵⁷. Therefore, it is not yet determined what percentage of the actual tumor-reactive TILs can be classified using such molecular signatures, given that most existing gene expression profiles are based on a sparse number of TIL clones. For example, a study by Krishna et al. revealed that patients who responded to TIL therapy retained a pool of stem-like CD39-negative TILs, thus contradicting parts of the previously mentioned work and emphasizing the need to explore additional markers of T cell reactivity against tumors ^{58,59}.

Here, we profile CD8⁺ TILs from nine patients with primary non-small cell lung cancer resections that had not been previously subjected to systemic therapies (from now on mentioned as treatment naive). We achieve this through a comprehensive single-cell multiomics approach combining scRNA-seq, scTCR-seq and scATAC-seq of TILs from non-small cell lung cancer and non-tumor adjacent tissue. For four patients, we selected a total of 138 TCRs for tumor reactivity testing using a previously established functional cellular screening platform ⁶⁰. T cell reactivity screening was performed against a panel of potential antigens, including common viral antigens, TAAs and neoantigens and led to the identification of 18 reactive TCRs that were represented by transcriptomes of over 1,000 TILs. Further, we identified cell state differences in shared clones in tumor and non-tumor tissue and generated a gene regulatory network that identified drivers of exhaustion. Through short-term IL-2 treatment of tumor suspension we were able to compare the potential cell state change of virus- and tumor-reactive TCR clones, possibly proving valuable for therapeutic decision making. This study combines TIL cell state dynamics with T cell reactivity and provides a helpful resource for the targeted selection of reactive TILs in lung cancer and non-tumor adjacent tissue.


3.3 Results

CD8⁺ T cell transcriptomes in treatment naive lung cancer patients

Tissue was obtained from nine patients resected for lung tumors who had not undergone therapy prior to resection. Seven patients were males and two females. Patient average age was 67.4 years. Five patients had a squamous cell carcinoma and four an adenocarcinoma (**Table 1, Table S1**).

Table 1: Single cell sequencing methods used in this study for each patient

Patients	TIL scRNA-seq	TIL scTCR-seq	TIL scRNA + scATAC-seq (Multiome)	Jux T cell scRNA-seq	Jux T cell scTCR-seq
BS476	Yes	Yes	No	Yes	Yes
BS631	Yes	Yes	No	No	No
BS833	Yes	Yes	No	No	No
BS867	No	No	Yes	No	No
BS956	No	No	Yes	No	No
BS980	Yes	Yes	No	No	No
BS1047	Yes	Yes	No	No	No
BS1064	Yes	Yes	No	No	No
BS1140	Yes	Yes	Yes	Yes	Yes
LI033	Yes	Yes	No	Yes	Yes



We analyzed CD8⁺ T cells (hereafter referred to as TILs) from tumor samples of eight patients, using scRNA-seq and scTCR-seq. For three of the patients (BS476, BS1140 and LI033), T cells from non-tumoral adjacent lung tissue were additionally sequenced for comparative analysis. For another two patients (BS867 and BS1140), TILs were subjected to single-cell multiomics sequencing (scMultiome: scRNA-seq and scATAC-seq from the same cell).

Across all patients and tissue samples, scRNA-seq profiles were generated from 57,432 T cells; and for scTCR-seq, data was generated from 40,228 cells across seven patients. 12 TIL sub-cell

states were identified by using a combination of marker genes previously described^{58,61,62} (**Fig. 1A, B**). Unified manifold approximation and projection (UMAP) analysis was performed and identified exhausted T cell subsets (*Exhausted*, *RM_exh1*) that shared expression of markers such as CXCL13, ENTPD1 (CD39) and PDCD1, as well as identification of a proximally located “*Proliferating*” cluster associated with MKI67 expression. Interestingly, expression of innate-like/NK cell markers were present in the *Innate_like_T* and *RM_exh2* cluster, suggesting an NKT-like state next to effector and exhausted cell states. A gene set consisting of markers that were previously shown to correlate with tumor-reactive T cells could be identified in the *Exhausted*, *RM_exh1*, *RM_exh2* and *Proliferating* clusters^{63,64}. A stem-like T cell marker gene set was highly expressed by cells in the *EM2* and *Innate_like_T* clusters, with similar co-location of a recent activation marker gene set. An effector gene set showed broad expression with increased levels in the *Exhausted*, *Innate_like_T* and *EM3* clusters (**Fig. 1D**). Pseudotime analysis indicates that *RM_Exh1* and *Proliferating* clusters contain cells with the most recent gene expression signature and are the most actively cycling cells (similar as shown by Gueguen et al.⁶²) (**Fig. 1E**). Further comparison of cluster frequencies across patients between tumor and adjacent tissue shows that the combined area of clusters *Proliferating*, *RM_exh1*, *RM_exh2* and *Exhausted* is greater than two times as much in TILs from tumor compared to adjacent (non-tumor) tissue, both across patients and by patient pair (**Fig. 1C**).

Identifying reactive TILs by functional TCR screening

Next, we identified TILs that are reactive to tumor-specific mutations as well as to antigens from common viruses, such as cytomegalovirus (CMV) and Epstein-Barr virus (EBV). To achieve this we utilized a recently established TCR functional screening platform based on an engineered human T cell line (TnT cells)⁶⁰. Briefly, TnT cells were created by immunogenomic engineering of Jurkat cells to introduce CD8 (CD8A, CD8B), Cas9 and a reporter of T cell activation (NFAT-GFP), as well as knock-out of beta-2-microglobulin (B2M) and the endogenous TCR α -chain. TnT cells thus serve as cellular recipients for genomic integration of recombinant TCRs, including libraries based on scTCR-seq of TILs. Further, TnT cells provide a functional screening platform for TCR activation by expression of NFAT-GFP or surface expression of CD69 following co-culture with cells (APCs or tumor cells) expressing cognate antigen (pHLA). Compared to primary T cells or TILs, which can have altered cell states and phenotypes that can compromise their activation (e.g., exhaustion), the synthetic TnT cells provide an unbiased functional cell state that can be used for TCR screening, including engineering TCR specificity to target antigens⁶⁰.

In order to be able to screen a multitude of TCRs and antigens against each other, we leveraged deep sequencing data following selection (FACS) of activated TnT cells that were co-cultured with APCs. To validate our approach, we generated a TCR test library, which included as positive controls two TCR clones (A3WT and A3-05)⁶⁰ with known reactivity against a TAA of the MAGE-A3 protein (HLA-A*0101-restricted MAGE-A3₁₆₈₋₁₇₆ peptide: EVDPIGHLIY) along with an additional 23 TCRs (extracted from patient BS833 and expressed in TnT cells) of unknown reactivity that are not expected to be reactive to MAGE-A3-derived antigens. As target cells, we used the MAGE-A3-expressing EJM tumor cell line or non-MAGE-A3-expressing COLO205 tumor cell line, which both have HLA-A*0101 expression. After co-culture of the TCR-TnT library cells with target tumor cells, FACS was performed to isolate TCR-TnT cells based on the activation markers CD69 and NFAT-GFP (**Fig 2A**). Following sorting, TCR-TnT cells were subjected to targeted deep sequencing of the transgenic TCR locus. Sequencing read TCR frequencies of cells co-cultured with EJM, COLO205 or prior to co-culture (background) revealed a >20-fold enrichment of the control MAGEA3-specific TCRs (A3WT and A3-05) compared to the TCRs with unknown reactivity (**Fig S1**).

Leveraging scRNA-seq and scTCR-seq data, we rationally selected TCR clones to screen in TnT cells. Here, we define TCR clones as T cells that have identical V(D)J germline usage and have 100% amino acid sequence identity of CDR3 α and CDR3 β sequences. The TCR clones selected for screening were based on several parameters that may be associated with reactivity: i) highest clonal abundance in a given patient, ii) the ten most abundant clones in each of the following clusters: *Proliferating*, *Exhausted*, *RM_exh1* or *RM_exh2*, iii) clones that were co-abundant in both *EM1* and one or more of the previously mentioned clusters (see **Methods**; **Fig 2B**). The selected TCRs were integrated in TnT cells by CRISPR-Cas9 homology-directed repair (HDR) and the resulting TnT-TCR libraries were screened as a pool with co-cultures of autologous tumor cell suspension or co-culture with patient-specific HLA-presenting cells that were pulsed with either neoantigens, multiple TAA linked to T cell reactivity in lung cancer (polyTAA: WT1/WT33, CEA, NY-ESO-1, MAGE-A3) ⁶⁵ or viral antigens derived from EBV and CMV. Neo-antigens were selected based on *in silico* predicted binding strength of mutated peptides from exome sequencing data to patient-specific HLAs (see **STAR Methods**). TnT-TCR cells co-cultured with patient-derived PBMCs that did not receive any peptide were used as a negative control. The bacterial superantigen phytohemagglutinin (PHA) was used as a positive control to detect nonspecific activation.

TCR reactivity data for identification of cognate pHLA was generated as follows. FACS was used to isolate TCR-TnT cells expressing above-background levels of the dual activation markers NFAT-GFP and CD69 (**Fig S2**). Next, targeted deep sequencing was performed on the transgenic TCR regions of TnT cells. Deep sequencing data was analyzed to determine TCR clones that were enriched under the different co-culture conditions. TCR clone frequencies were compared between target and control co-culture conditions (pre co-culture, PBMC co-culture, PHA-activated co-culture). Through statistical analysis, each TCR clone per target co-culture received a fold change in comparison to the highest control co-culture. A reactivity threshold could be determined based on the density plot of fold-change values across all TCR clones and target co-cultures. TCR clones with a fold-change value > 2.5 were classified as reactive (see **Methods**; **Fig S3**; **Table S2**).

This led to the identification of 18 reactive TCRs (out of 138 tested TCRs). Multiple TCRs showed similar fold change reactivities for each duplicate against CMV (BS833, clone 6 and 8),

EBV (BS1064, clone 11) and pTAA (BS833, clone 119), showing the robustness of this functional screening approach. One TCR showed cross reactivity between neoantigen/self-antigen and CMV (BS1140, clone 14); with another one showing reactivity against tumor suspension and neoantigen (BS1064, clone 10) (**Fig 2C**). Genes that harbored mutations for neoantigen generation and were associated with TCR reactivity were mutated in 2.5%-10%⁶⁶⁻⁶⁸ of lung cancer patients and included: KMT2D, ATR and SORCS1.

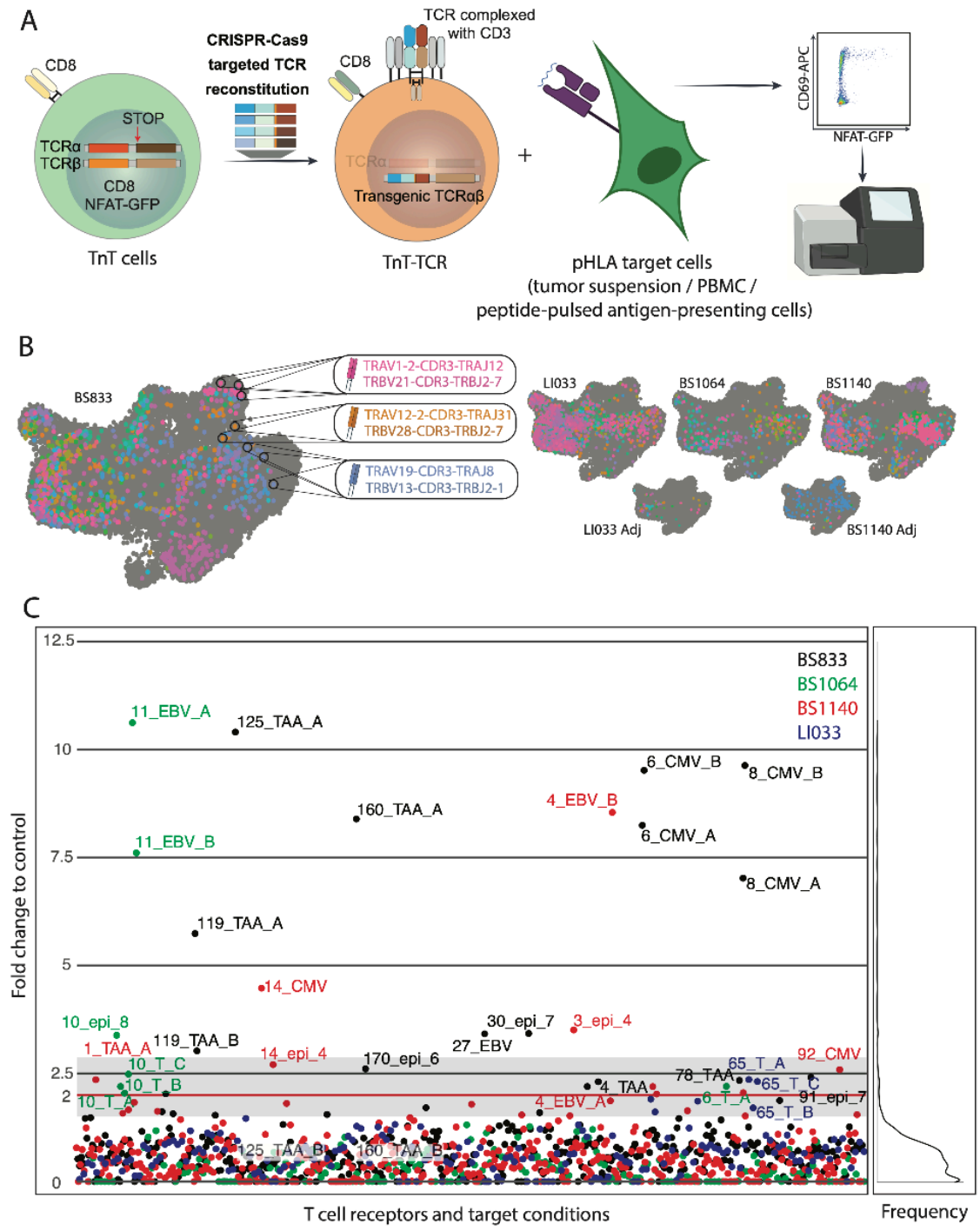


Figure 2: Reactivity profiling of selected TCRs by functional screening and deep sequencing.

(A) TCR libraries are genomically integrated into TnT cells by CRISPR-Cas9, which are then co-cultured with pHLA target cells such as tumor suspension, patient PBMC or peptide pulsed antigen presenting cells (see STAR Methods). Fluorescence activated cell sorting (FACS) and deep sequencing are performed to select and identify reactive TCR-TnT cells. (B) T cells with TCR clones selected for functional screening are shown with their

respective transcriptomes on UMAP. (C) Left, dotplot shows deep sequencing results following TCR-TnT functional screening. Fold-change represents TCR sequencing reads of each target reactivity condition compared to the highest control value, across all control conditions for the same TCR, color corresponding to the patient. Reactive TCRs are designated as those with fold-change values >2.5. Right, density plot of deep sequencing results for all TCRs and their target conditions across patients. X-axis represents the frequency of all TCR-condition values per fold-change value that is shown as in (left) on the y-axis.

Phenotypes of reactive T cells in treatment naive lung cancer biopsies

Following identification of 18 TCRs with reactivity to tumor cells and viral antigens, we next analyzed the associated transcriptional phenotypes that consisted of over 1,000 corresponding TILs. Tumor- and polyTAA-reactive TILs showed high abundance in clusters associated with (pre-) dysfunctional phenotypes, whereas CMV and EBV viral antigen-reactive TCRs were mainly present in effector phenotype clusters (**Fig. 3B, C**). Patient tissues showed absence of expression of EBV or CMV, indicating that these virus-reactive TCRs respond to systemic antigens rather than tissue-specific antigens (not tumor-specific) (**Table S3**). We subsequently determined key marker genes in TILs associated with tumor or viral reactivity by differential gene expression analysis between these two subgroups. The strongest marker for tumor reactivity is CXCL13, which corroborates a previous study⁶⁴ (**Fig. 3A**). SOX4 is a transcription factor that facilitates the development of CXCL13-producing T cells⁶⁹ and is also among the five genes with the highest differential gene expression for tumor reactivity, further suggesting the importance of the CXCL13 pathway for TIL activation. Other tumor reactivity markers included genes involved in co-stimulation (CD7)⁷⁰, repeated antigen stimulation (DUSP4)⁷¹ and immune checkpoints (TIGIT)⁷². For virus reactivity, cytotoxic markers such as GZMK and GZMH showed high abundance in EBV- and CMV-reactive TILs, respectively. Analysis of T cell reactivities across clusters revealed a strong enrichment of tumor reactivity in the *Exhausted*, *RM_exh1*, *RM_exh2* and *Proliferating* clusters. It is important to consider that since non-reactive T cells showed abundance in clusters that were enriched for CMV-reactive and tumor-reactive T cells, they may still possess reactivity to antigens that were not included in the co-culture screening (**Fig 3C**).

To infer likelihood of reactivity in untested cells, we developed a reactivity score based on gene expression similarity to the experimentally validated tumor- and virus-reactive TILs. This is

accomplished by creating a gene module score, which is a linear combination of gene expression levels, derived from the differentially expressed genes observed in tumor- or virus-reactive T cells. (**Fig 3D**). When visualized on the UMAP, these two module scores show *Proliferating*, *Exhausted* and *RM_exh1-2* or *EMI-3* as enriched areas for locations of tumor- or virus-reactive T cells, respectively (**Fig 3E**). In agreement with previous studies that showed little to no correlation between T cell clonal expansion level and detected reactivity³², we did not observe enrichment of reactive TCRs in the most expanded clones (**Fig S4**).

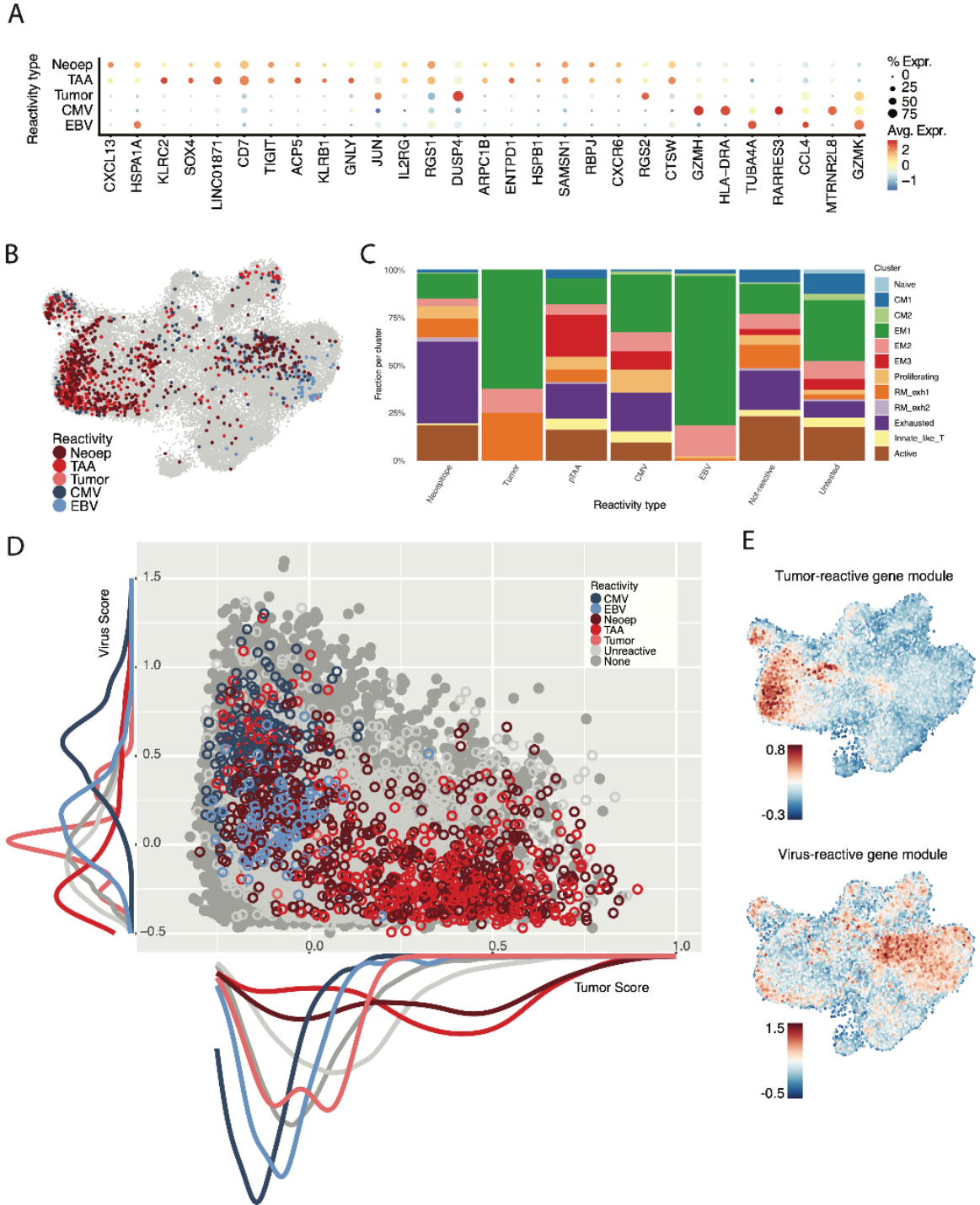


Figure 3: Characterisation of transcriptional programs of reactive TILs.

(A) Dotplot map showing each identified marker gene for tumor and virus reactive T cells. Rows correspond to the reactivity groups of T cells. Color indicates average expression, dot size indicates fraction (%) of cells in a reactivity group. (B) Reactive T cells highlighted on the UMAP according to observed reactivity to cognate antigen. (C) Fraction of cells per cluster sorted by each reactivity group of T cells. Unreactive T cells are defined as TCRs that were functionally screened but did not show reactivity. Untested are all TCRs that were not subjected to functional

screening. (D) Scatter plot of tumor- and virus-reactive gene scores, whereby each dot represents a cell and is color-coded according to observed reactivity type. Gene scores were derived by combining differentially expressed gene expression levels (see STAR Methods). Density lines on the x- and y-axis are colored by observed reactivity type. Each density line is normalized on the cells contained in its subgroup. (E) Tumor (top) and virus (bottom) gene module score expression on UMAP; score range is relative to each gene module score.

TCR clonal sharing in treatment naive lung cancer and non-tumor adjacent tissue

The combination of scRNA-seq and scTCR-seq is able to connect transcriptomes with full TCR sequences (paired $V\alpha$ and $V\beta$) and thus elucidate cell state diversity across T cell clones. From the 40,228 cells with scTCR-seq data, 26,018 had corresponding scRNA-seq data (**Fig S5**). Among these, 12,284 scRNA-seq and scTCR-seq cells were in patients with both tumor (8,854) and adjacent (3,430) tissue (**Fig 4A**).

The scTCR-seq repertoire analysis revealed major differences in T cell clonal expansion levels across tissues. T cells from adjacent tissue show larger clonal expansion levels and in turn also reduced diversity of clones compared to T cells in the tumor counterpart, especially for the most expanded clones (**Fig 4A**). Across TIL TCR repertoires, the largest five clones accounted for more than 25% of the cells. Expansion level, based on clonal size, describes the number of cells identified per clone and can be caused by extensive proliferation through TCR activation upon cognate pHLA binding. Transcriptome analysis associated with clonal expansion levels (here, shown as percentage of expanded clones that is ordered by rank of clonal size) revealed preferential phenotypes between T cells in tumor and adjacent (non-tumor) lung tissue. In the tumor tissue, unexpanded clones (1 cell per clone detected) are mostly located in the *Naive* and *CMI-2* clusters, and moderately expanded clones (*Frequency 6-99%*) are evenly distributed across *CMI-2*, *EMI-3*, *Exhausted* and *Proliferating* clusters. Whereas the highest expanded clones (*Frequency 0-5%*) show enrichment in the *Exhausted*, *RM_exh1*, *RM_exh2*, *Proliferating* clusters (**Fig 4B, left**). Interestingly, the highest expanded clones that reside in the adjacent tissue are almost exclusively located in the *EMI-3* clusters (**Fig 4B, right**). Compared to the tumor tissue, this shows vastly different T cell phenotypes for cells with the same expansion levels. To discern which TCRs are predominantly located in the tumor as opposed to the adjacent non-tumor lung tissue, we contrasted the frequencies and cellular states of matched TCRs (clones observed within both the tumor and adjacent tissue) (**Fig S6**). The largest matched clones

identified were generally found in greater abundance within the tumor. However, among the top 100 matched clones, over 10% appear to be in greater abundance within the surrounding tissue, suggesting potential circulation from the adjacent tissue into the tumor. Conversely, clones larger in the tumor may be migrating outward into the adjacent tissue (**Fig 4C**). This T cell residency is also reflected in a differing phenotype that could be identified when comparing cell states of T cell clones that were more than 3-fold enriched in one over the other tissue type (tumor- or lung (adjacent)-residency). There was a marked increase of effector T cell states when contrasted with the resident memory, exhaustion, and proliferating cell states found in tumor-resident T cells (**Fig 4D**). Further analysis of the 20 largest shared clones across tissues revealed higher instances of exhaustion in the clones present in the tumor. This suggests the adjacent lung tissue could be a more likely source of effector-like T cells (**Fig 4E**).

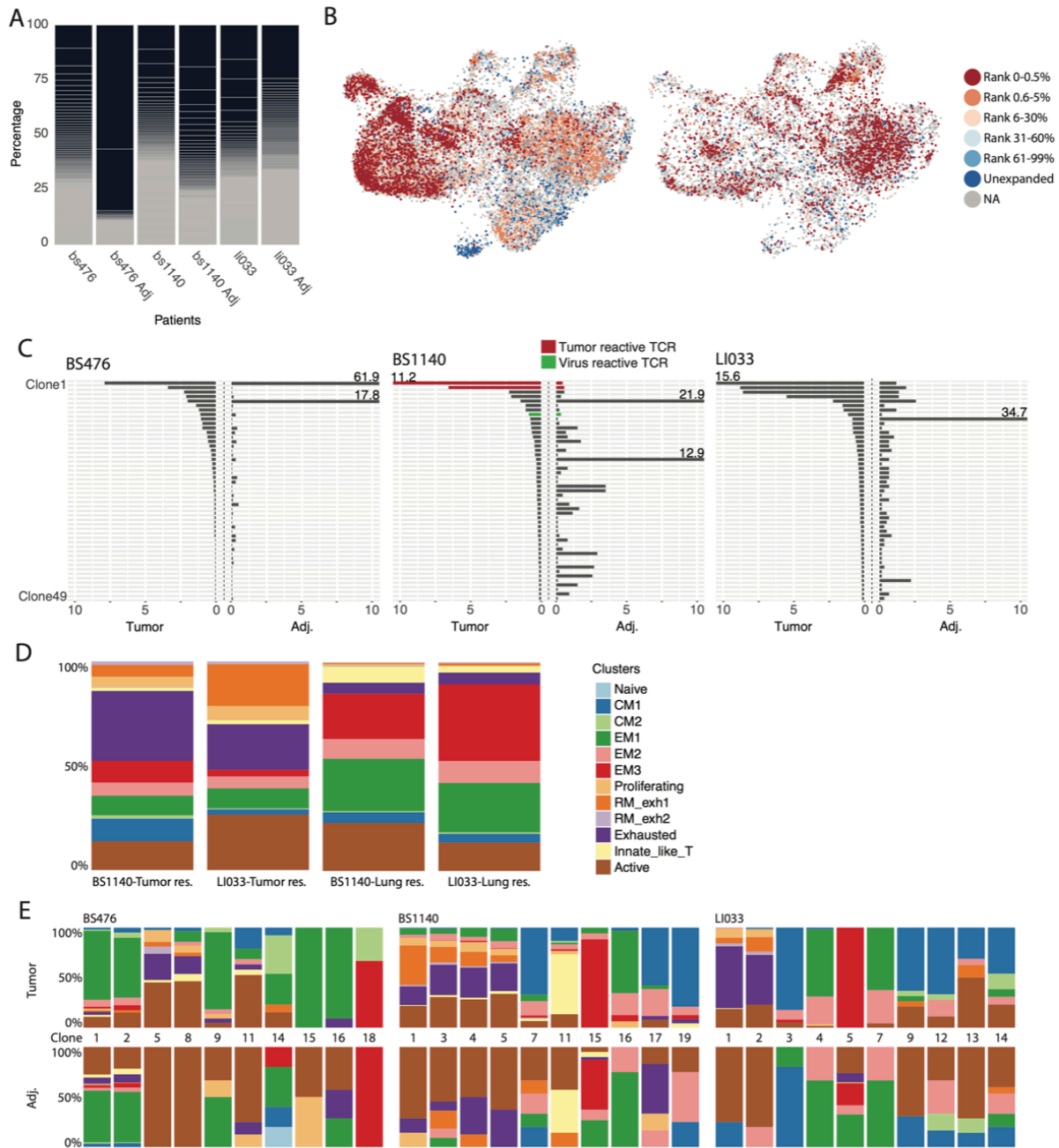


Figure 4: Differential TIL phenotype of shared clones in tumor and non-tumor adjacent lung tissue.

(A) The TCR clonal distribution as percentage of the observed repertoire. Grey lines separate clones. (B) Clonal expansion levels for TCRs mapped onto the UMAP. Left, TCRs of T cells that were identified in patient tumor tissue. Right, TCRs of T cells that were identified in patient non-tumor adjacent (adj.) lung tissue. Color represents clonal fractions according to their expansion level. Unexpanded clones are defined as TCRs found in only one cell. (C) The 49 largest clones shared between tumor (left) and adjacent (adj.) (right) tissue. Each line represents one clone. X-axis shows the percentage of TCRs in the corresponding patient repertoire. If the percentage exceeded the x-axis limits, the number indicates the clonal percentage. Lines are colored by reactivity to tumor (red) or virus (green). (D) Cluster distribution of all T cells that are resident by majority in tumor or adjacent lung tissue by

patient, according to the previous tissue abundance measurements. (E) Transcriptome of the 10 largest shared clones across both tissue modalities. Shown as percentage of clone per cluster.

Single-cell multiomic analysis reveals new drivers of a T cell terminal exhaustion phenotype

The state of T cells is significantly influenced by their environment and the stimuli they receive through cytokines and TCR signaling. To study the dynamics of T cell states in untreated, primary human lung adenocarcinoma, we used a sc-Multiomic analysis that integrates scATAC-seq with scRNA-seq. This approach allowed us to study both chromatin accessibility and gene expression in the same cell (**Fig 5A**). With scATAC-seq, it is possible to identify DNA motifs that act as enhancers or repressors through TF binding. scMultiomics also provides gene expression measurements of TF-encoding genes, TF-target genes and accessibility of TF-motifs; the combination of these three features is referred to as a Regulon⁷³. Regulons were identified across all T cells and used for a new clustering analysis, resulting in 10 clusters (**Fig 5C**). Regulon analysis also revealed a major difference in clustering of exhausted and proliferative cells (*clusters 2 and 3 in Fig 5C*) in comparison to using only the scRNA-seq or scATAC-seq data (**Fig 5B, Fig S7, Fig S8**). Consistent with the previously shown analysis (**Fig 1E**), pseudotime in the Regulon UMAP displays exhausted and proliferating clusters as the latest cells along the trajectory (**Fig 5F**). Further, analysis of TCRs from sc-Multiomic sequencing shows comparable transcriptomes for the same TCR clone in scRNA-seq data (**Fig S9**). Differential Regulon analysis identified multiple known TFs implicated in specific T cell states such as LEF1 and BACH2 in the naive, EOMES in the effector and PRDM1 in the exhausted state. We also identified multiple TFs that are uniquely implicated in the exhausted state such as SOX4, RFX2, *ETS1*, *ETV1* and *STAT1* (**Fig 5D**). To contextualize Regulons within cellular differentiation, we constructed a gene regulatory network (GRN)⁷⁴. We marked TFs and their encoding genes with names that are associated with certain cell states or that have a high connectedness in the network (centrality). The lower area shows enrichment of genes implicated in exhausted T cells, such as TOX, CD39 and CXCL13. Genes that cluster together with such may have an effect on T cell persistence and exhaustion (**Fig 5E**). For example, it was just recently shown that SNX9 deletion alleviates T cell exhaustion⁷⁵. Detailed analysis of TFs that are strongly associated with

driving exhaustion (RUNX1 and PRDM1), as well as being a marker for tumor-reactive TILs (SOX4) visualizes their impact on T cell differentiation towards the exhausted state (**Fig 5G**). In line with this, markers represented in the tumor-reactive signature that are represented in the GRN are clustered almost uniquely in the exhausted area (**Fig 5E**). While TF expression and target gene expression steadily rise along the exhaustion trajectory, the target region accessibility abruptly increases in the last exhaustion stage, possibly corresponding with the transition of early (reversible) to late (irreversible) dysfunction (**Fig 5G**).

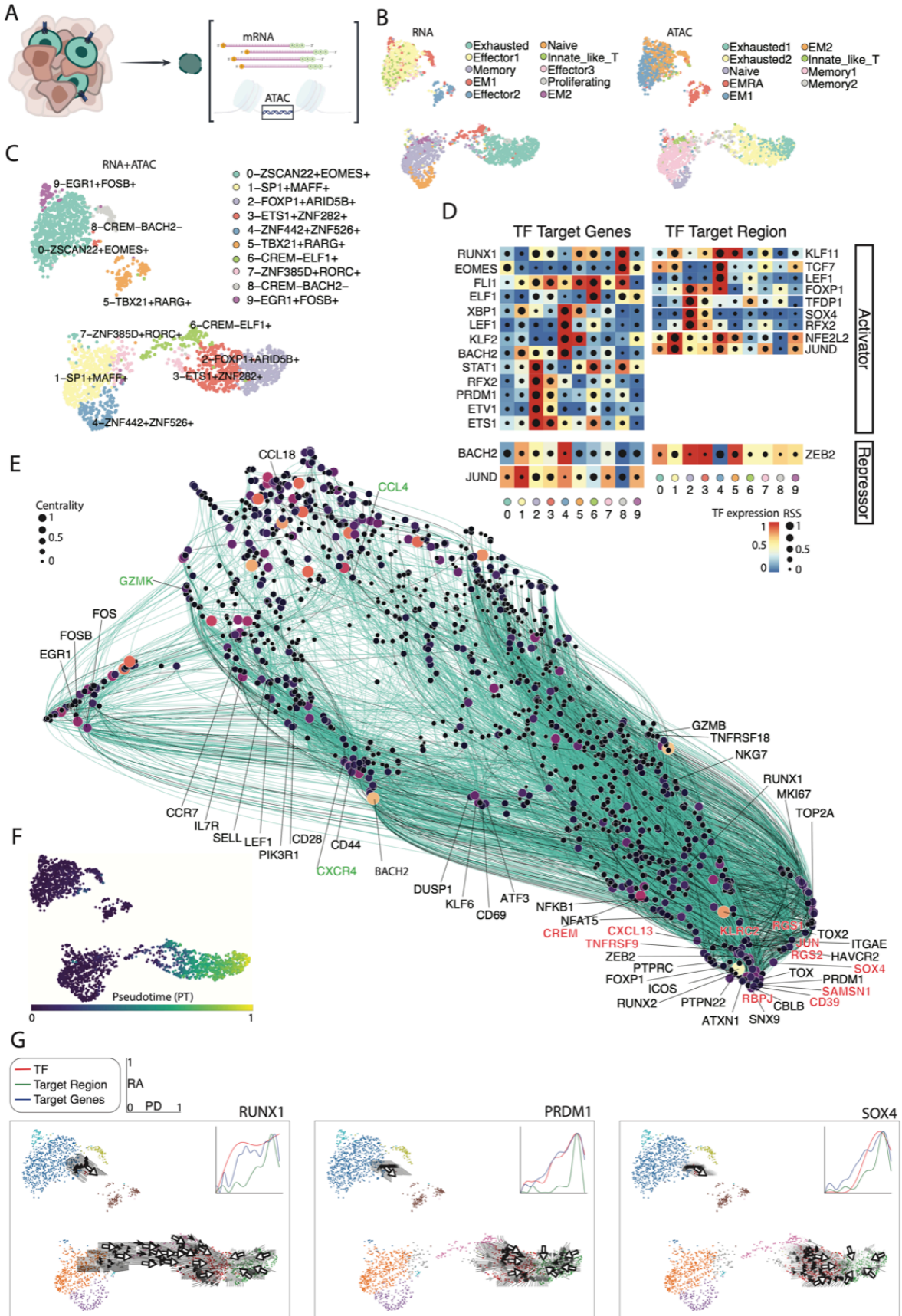


Figure 5: Applying scMultiome sequencing for gene regulatory network analysis of TILs.

(A) Schematic of workflow for scMultiome sequencing (scRNA-seq and scATAC-seq). (B and C) Clustering was done via SCENIC+. (B) Left, UMAP cluster coloring based on scRNA-seq data alone. Right, UMAP cluster coloring based on scATAC-seq data alone. (C) Multiome UMAP cluster coloring based on Regulon information for each cell. Regulon is a combination of scATAC- and scRNA-seq data for TF expression, TF target-gene expression and TF-binding site accessibility. Each cluster was named by the most differentially expressed and active Regulon. (D) Differentially expressed transcription factors (TFs) and their target genes (RNA-based) (left) or target regions (ATAC-based) (right) according to the scMultiome UMAP. (E) Gene regulatory network (GRN) using Regulon-information across all cells from patient BS1140. Color and size correspond to centrality (large and bright = high centrality). Line color indicates enhancing (turquoise) or repressing (gray) interaction between genes. (F) Pseudotime of Multiome UMAP. (G) Arrow- and Curve-plots of TFs. Arrow plots indicate differentiation force and direction, based on scMultiome-seq data. Curve plots show TF expression level (red), target gene expression level (blue) as well as specific region accessibility (green) along the pseudotime (Fig 5F) from naive (cluster 4) to exhaustion (cluster 2 and 3).

Tumor suspension IL-2 co-culture leads to a decreased exhaustion cell state in tumor- and virus-reactive TILs

Having discovered gene reactivity markers and clones with associated reactivities that are also present in the *Proliferating* cluster, we determined how these clonal phenotypes change in short-term (2 days) culture conditions in order to simulate tumor suspension-based stimulation of reactive T cells. We thus chose normal concentrations of IL-2 (100 IU/mL for 2 days) as the culture condition of TILs and used a tumor single cell suspension from patient BS833 that contained TIL clones with various reactivities. Analysis of scRNA-seq from TILs, both pre- and post-IL-2 treatment, revealed a 4-fold larger fraction of TILs in the *Proliferating* cluster; crucially, there was a stronger exchange of RNA velocity vectors between the *RM_exh1* and *Proliferating* clusters, suggesting active transfer between these two cell states (**Fig 6A**). Comparison of reactive T cells from pre- and post-IL-2 treatment showed clear enrichment of virus-reactive (CMV) and neoantigen-reactive T cells in the *Proliferating* cluster. Neoantigen-reactive T cells are driven into an *Innate_like_T* cell state and reduce their *Exhausted* cell state upon IL-2 treatment (**Fig 6B**). The 20 largest shared clones in pre- and post-IL-2 treatment data showed overall increases in *Proliferating*, *Innate_like_T* and *EM1* states and decreases in *Exhausted*, *CM1-2* and *EM2* states post-IL-2 treatment (**Fig 6C**).

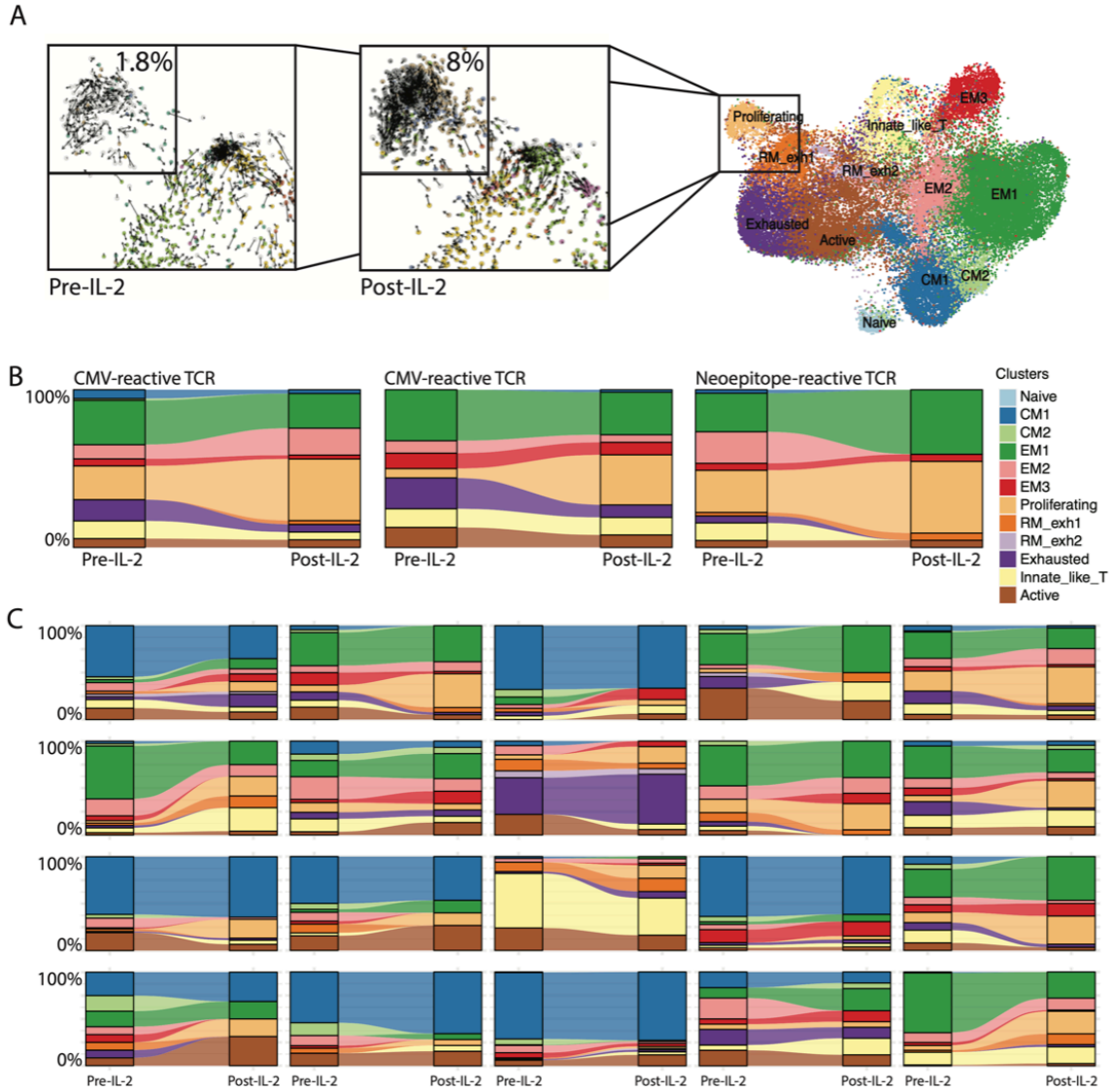


Figure 6: TIL cell state changes upon low IL-2 treatment while in tumor suspension.

(A) UMAP showing transcriptome data of cells of patient BS833. Zoom windows show RNA velocity representations of the *Proliferating* cluster pre- and post-IL-2 treatment. The smaller window represents the percentage of cells being part of the *Proliferating* cluster for both timepoints. (B and C) Alluvial-plots showing the change of cell-state distribution per T cell clone pre- and post-IL-2 treatment as percentage of clusters. (B) Shown are three TCR clones with known reactivities. (C) Shown are the 20 largest shared clones for pre- and post-IL-2 treatment data (from left to right and top to bottom).

3.4 Discussion

In this study, we profiled CD8⁺ TILs from untreated primary lung cancer patients by using single-cell multiomics, which combined scRNA-seq, scTCR-seq, and scATAC-seq and was followed by functional screening to identify reactive TCRs.

Importantly, by using a previously established TCR functional screening platform based on an engineered human T cell line ⁶⁰, we were able to avoid the heterogeneity of cellular phenotypes present in primary TILs, which are often used for screening endogenous or transgenic TCRs ^{43,54} and can be a source for confounding results. For example, dysfunctional TILs upregulate CISH, a key regulator of T cell functional avidity, thus leading to an increased activation threshold upon cognate pMHC binding and in turn this may result in higher false negative rates of TILs in co-culture screens ⁵⁷. While functional screening was able to identify reactive TCRs, it may still be possible that some tumor-reactive TILs are already dysfunctional at the time of scRNA-seq and thus did not upregulate gene activation markers of proliferation upon cognate (tumor-) antigen binding. Furthermore, the selection of antigens for reactivity testing is critical for interrogation of the tumor-reactive TIL population. Therefore, we included a diverse mix of neoantigens, TAA, and patient-derived tumor cells. While other studies have primarily focused on neoantigens (minigenes expressing proteins with tumor-specific mutations identified by whole exome sequencing) ^{53,54}, their sole use may underestimate the number of reactive TILs ⁷⁶. In future studies, it may be beneficial to couple TCR reactivity screening with de-orphaning of antigen specificity through high-throughput pHLA antigen screening systems ⁷⁷⁻⁸⁰.

Our study identified a distinct transcriptomic signature in tumor-reactive T cells. This signature highlights gene expression patterns linked to repeated antigen stimulation ^{61,81}, emphasizing its connection to recently discovered markers of exhaustion such as SNX9, as revealed by our GRN ^{53,75}. Central to this tumor reactivity signature is the exertion of immune cell recruitment via the CXCL13 pathway, including SOX4, and the upregulation of immune checkpoints (e.g., ITGAE), T cell differentiation markers (e.g., CD7 and DUSP4), as well as a distribution among clusters of T cell exhaustion, stemness, resident memory and proliferation ⁶⁹⁻⁷², possibly being the result of asymmetric cell division upon antigen encounter ^{82,83}. Notably, CXCL13 emerged as a robust

marker for tumor-reactivity, corroborated by findings in other tumor types and has also been linked to tertiary lymphoid structure formation ^{54,64,84,85}. Furthermore, markers such as CD7 and DUSP4, suggestive of repeated antigen stimulation, alongside resident memory marker ITGAE, underscore the heightened interactions between T cells and tumor cells within the tumor microenvironment .

The importance of T cell state of stemness for TIL adoptive cell therapies has been demonstrated previously ⁵⁸, however neoantigen-reactive TILs have been observed to lack such stemness phenotypes ^{43,53}. Here, we find a substantial overlap of the tumor-reactive gene module and stemness in TILs, possibly owing to our inclusion of polyTAA and tumor suspension as antigen targets, as well as focusing on treatment naive non-small cell lung cancer patients. Interestingly, for shared TCR clones (cells present in tumor and the non-tumor adjacent tissue), the T cells from adjacent tissue display lower levels of dysfunction, indicating a bifurcation of T cell states of the same clonal origin, possibly through lowered antigen exposure or suppressive signal in the adjacent tissue. Here, we observed a significant clonal T cell overlap between tumor and non-tumor adjacent tissues, corroborating previous findings ⁶². Moreover, through our functional reactivity analysis, we determined that tumor-reactive T cells not only migrate from the tumor to the adjacent non-tumor tissue but also exhibit a pronounced effector-like phenotype within this adjacent milieu.

In addition to the identification of a tumor-reactive signature, the scMultiomic approach may also support the preferential enrichment of tumor-reactive T cells through a gene network of exhaustion markers, likely caused by repeated antigen stimulation. The identification of TCRs could be used or engineered for molecular or cell therapies ⁸⁶⁻⁸⁸. Recent findings demonstrate that single TCRs could be harnessed as therapeutics for multiple cancer types by exploiting their cross-reactivity against multiple tumor antigens ⁸⁹; this underlines the importance of identifying tumor-reactive TCRs, potentially through methods such as gene signature enrichment.

Moreover, using non-tumor adjacent lung tissue as an additional source for reactive T cells may enable the extraction of more effector-like tumor-reactive T cells, which may help with expanding a sufficient number of T cells for autologous therapy ⁹⁰. As IL-2 is an integral part of many TIL-rapid expansion protocols, we could also show that short-term culture with low-dose

IL-2 and tumor suspension leads to varied changes of clonal cell state diversity and could serve as a predictor for patients' response to immunotherapies, as well as offering potential new strategies for IL-2-based expansion ⁹¹ or targeting of T cells ⁹². Long-term, it will be critical to understand the relationship between TIL cell states and therapeutic outcome, as is currently being done for chimeric-antigen receptor T cell therapies ⁹³.

Supplementary information

Supplemental Figures 1-9

Supplemental Tables 1 and 2

3.5 Methods

Key resources table

Reagent or Resource	Source	Identifier
Antibodies		
Anti-human CD3 PE/Cy7 (clone UCHT1)	Biolegend	#300419
Anti-human CD4 APC (clone RPA-T4)	Biolegend	#300514
Anti-human CD8 BV421 (clone RPA-T8)	Biolegend	#301035
Anti-human CD45 PE (clone 2D1)	Biolegend	#368509
Anti-human CD69 APC (clone FN50)	Biolegend	#314845
Anti-human CD28 antibody (clone CD28.2)	Biolegend	#11150591
TruStain FcX	Biolegend	#422301
Biological Samples		
Patient tumor, adjacent non-tumor lung tissue and peripheral blood samples	Unispital Basel, Switzerland	N/A
Chemicals, peptides, and recombinant proteins		
PHA	Invivogen	#inh-phap
Recombinant human IFN γ	Peptotech	#300-02
Recombinant human IL-2	Peptotech	#200-02
MAGEA3 peptide pool	JPT	#PM-MAGEA3
CEA peptide pool	JPT	#PM-CEA
NY-ESO peptide pool	JPT	#PM-NYE
WT1/WT33 peptide pool	JPT	#PM-WT1
EBV peptide pool	Miltenyi Biotech	#130-103-462
CMV peptide pool	Mabtech	#3619-1
Alt-R HDR enhancer	IDT	#1081073
DAPI viability dye	ThermoFischer	#62248
TRIZol	Invitrogen	#15596018
RiboLock RNase inhibitor	ThermoFischer	#EO0381
Maxima H-minus reverse transcriptase	ThermoFischer	#EP0753
Custom peptide synthesis	Genscript	N/A
eBioscience Cell Stimulation Cocktail (81 nM PMA, 1.34 μ M ionomycin)	ThermoFischer	#00497093
Critical commercial assays		

Chromium Single Cell V(D)J Enrichment Kit, Human T Cell	10X Genomics	PN-1000005
Chromium Single Cell 5' Library & Gel Bead Kit	10X Genomics	PN-1000006
Chromium Single Cell A Chip Kit	10X Genomics	PN-1000009
Chromium i7 Multiplex Kit	10X Genomics	PN-120262
Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Reagent Bundle	10X Genomics	PN-1000283
Chromium Next GEM Chip J Single Cell Kit	10X Genomics	PN-1000230
Single Index Kit N Set A	10X Genomics	PN-1000212
Dual Index Kit TT Set A	10X Genomics	PN-1000215
Arcturus PicoPure RNA Isolation Kit	Life Technologies	#KIT0204
Oncomine™ Comprehensive Assay Plus	ThermoFischer	#A49667
KAPA HiFi PCR kit with GC buffer	Roche Diagnostics	#07958846001
KAPA2G Fast ReadyMix kit	Sigma Aldrich	#KK5102
SE Cell Line 4D-Nucleofector X Kit L	Lonza	#V4XC-1024
Deposited Data		
Tumor and non-tumor tissue T-cell single cell analysis	This Paper	N/A
Oligonucleotides		
anti-TCR-CDR3b-TnT crRNA: tcgacctgttcggctaacta	https://doi.org/10.1016/j.immuni.2022.09.004	N/A
anti-HLA-ABC-HEK crRNA: gatgtaatccttgccgtcgt	This paper	N/A
anti-GFP crRNA: caactacaagaccgcgccc	This paper	N/A
RVL127 primer: 5'-gcatgcctctgtgccaacag-3'	https://doi.org/10.1016/j.immuni.2022.09.004	N/A
RVL128 primer: 5'-ttttatctgtcatggccgtgaccg-3'	https://doi.org/10.1016/j.immuni.2022.09.004	N/A
RVL144 primer: 5'-gaggagaaccctggacctatg-3'	https://doi.org/10.1016/j.immuni.2022.09.004	N/A
RVL145 primer: 5'-ctagagacccccagccttacc-3'	https://doi.org/10.1016/j.immuni.2022.09.004	N/A
Oligo-dT primer	https://doi.org/10.1038/nprot.2014.006	N/A
45 TRBV gene primer set	https://doi.org/10.1016/j.xpro.2022.101391	
Alt-R tracrRNA	IDT	#1072534
Software and algorithms		
Cellranger	10X Genomics	
FlowJo X	FlowJo, LLC	https://www.flowjo.com/solutions/flowjo
R		version 4.2.2
Seurat package	R project	Version 4.3.0

Signac	R project	version 1.9.0
Loupe VDJ Browser	10X Genomics	Version 4.0
SCENIC+	R project	version 1.0.0
ArchR	R project	version 1.0.1
Pando	R project	version 1.0.0
ChromVAR	R project	version 1.22.1
Biostrings	R project	version 2.68.1
tidyverse	R project	version 1.3.2
pheatmap	R project	version 1.0.12
ggplot2	R project	version 3.4.0
ggalluvial	R project	version 0.12.5
Original code	This paper	Github: Analysis of scRNA and reactivity data: https://github.com/LS-SI-ETH/TIL_scRNA_profiling_lung_cancer_Bieberich ; Analysis of scMultiome data: https://github.com/LS-SI-ETH/TIL_multiome_profiling_lung_cancer_Bieberich
Other		
BD LSRFortessa	BD Biosciences	N/A
BD FACSAria Fusion	BD Biosciences	N/A
4D-Nucleofector, including LV unit	Lonza	N/A
Chromium controller	10X Genomics	PN-110203

Patient information and biospecimen processing

Patient samples were collected from individuals with lung cancer that had undergone primary resections and no systemic therapy (i.e. no chemotherapy, targeted therapy or immunotherapy) at Basel University Hospital between 2016 and 2019. The patients described in this study provided written informed consent. All bio-specimens were obtained from patients with stage II-III lung cancer. Detailed patient characteristics are provided in **Table S1**.

Primary resections of lung cancer lesions were immediately processed into single-cell suspensions through mechanical dissociation and enzymatic digestion using accutase,

collagenase IV, hyaluronidase and DNase type IV (Stemcell). Single-cell suspensions of samples were cryopreserved until experimental use.

HLA, tumor exome sequencing and viral exposure verification

HLA sequencing of peripheral blood mononuclear cells was performed by the pathology department at Basel University Hospital. Tumor exome sequencing was done using the commercially available OncoPanel platform (OncoPrint™ Comprehensive Assay Plus - DNA). Sequencing and analysis were performed by the Pathology Unit, led by Matthias S. Matter at Basel University Hospital. EBV and CMV expression in patient tumor tissue was tested using the QuantStudio 7 Pro System and was performed by the Pathology Unit, led by Kirsten Mertz at Kantonsspital Liestal (data is available upon reasonable request).

T cell sorting

Tumor single-cell suspension was thawed, washed, resuspended in blocking solution (Fc-block agent 1:50 in FACS buffer) for 15 min at 4 °C. Staining solution was added and suspension was incubated for 30 min at 4 °C (fluorescently-labeled antibodies were diluted in FACS buffer to 1 ug/mL). Cells were washed twice and cell sorting was performed using a BD FACS Fusion in FACS buffer (DPBS with 1% FBS). The fluorescently-labeled antibody staining consisted of anti-CD8 (BV421, Biolegend), anti-CD4 (APC, Biolegend), anti-CD3 (PE/Cy7, Biolegend), anti-CD45 (PE, Biolegend) and DAPI. TILs (CD8⁺, CD4⁻ and CD3⁺) were immediately subjected to single-cell droplet encapsulation.

Single-cell sequencing of TILs

Sorted TILs were counted using a hemocytometer and diluted at the desired cell concentration used in for the 10x single-cell sequencing protocol (1,000 cells/uL). The single-cell 5' VDJ and 5' GEX kits (10X Genomics) were used to capture transcriptomic information as well as allowing for a target enrichment step to capture immune repertoire information of single-cell paired TCR α - and β - chains. Encapsulation of lymphocytes and DNA-barcoded gel beads was performed using the Chromium controller (10x Genomics, PN-110203). Briefly, 10'000 - 20'000 cells were loaded per channel for a targeted recovery of 5'000-10'000 droplet encapsulated cells. Reverse transcription and preparation of single-cell transcriptome and TCR libraries was done

according to the manufacturer's instructions (CG000086 Rev M, 10X Genomics) and using the following kits: PN-1000006 and PN-1000005. Libraries were sequenced at 5000 (VDJ), 20000 (Transcriptome) and 25000 (ATAC) paired-end reads per cell.

For single-cell Multiome sequencing library preparation, the following changes were made to the protocol: Sorted TILs were subjected to nuclei isolation according to the manufacturer's instructions (CG000365). Isolated nuclei were adjusted to the desired concentration (2,500 cells/uL) and incubated in a Transposition Mix before also loading them onto a channel for encapsulation. Preparation of transcriptome and chromatin accessibility libraries were done according to the manufacturer's protocol (CG000338).

Single-cell data pre-processing and quality control

CellRanger software version 5 was used to demultiplex fastq reads from transcriptome, ATAC and TCR scSeq libraries and align them to the GRCh38 reference human genome. Subsequent data QC and analysis was performed using R (version 4.2.2) and the Seurat (version 4.3.0)⁹⁴, Signac (version 1.9.0)⁹⁵ and ArchR (version 1.0.1)⁹⁶ packages. QC steps consisted of the exclusion of TCR and BCR genes (prevention of clonotype influence on subsequent clustering), the exclusion of cells with lower than 150 or greater than 4500 genes (low quality cells), and the exclusion of cells in which more than 10% of UMIs were associated with mitochondrial genes (reduction of freeze-thaw metabolic effects).

Single-cell data integration and clustering

Single-cell RNA sequencing data was read into R using *Read10X*. A Seurat object was generated for each patient (*CreateSeuratObject*), all patients were merged into the same object (*merge*) and QC steps as mentioned before were applied. The Seurat object was split into individual object (by patient) and normalization (*SCTransform*) was applied and mitochondrial genes were regressed out. Variable features were selected from the list of normalized patient Seurat object (*SelectIntegrationFeatures*) and Seurat objects were merged again (*merge*). Previously extracted variable features were set and principal component analysis (PCA) was performed (*RunPCA*) on the normalized assay data. To reduce batch effects, we used the Harmony (version ...) package (*RunHarmony*). Finally, UMAP was calculated, and clusters were generated (*RunUMAP*, *FindNeighbors*, *FindClusters*).

scTCR-seq data analysis

T cell receptor data was analyzed using R Studio and Loupe VDJ Browser 4.0. In brief, TCR information was read into R after Cellranger analysis of the respective fastq files to generate the `filtered_contig_annotations.csv` file. Using R, clonal definitions were made based on the abundance of an exact match of the amino acid CDR3- α and - β sequence. Single chains were filtered out and clones that matched with multiple chains (more than one V α or V β) were separated from clones that matched with exactly one V α and V β chain.

Single-cell multiome analysis

Signac: Multiome analysis performed with the Signac package consisted of the following 7 steps: 1. Read in multiome data and create multimodal Seurat object (*Read10x_h5*, *CreateSeuratObject*, *CreateChromatinAssay*, *NucleosomeSignal*, *TSSEnrichment*, *CallPeaks*, *keepStandardChromosomes*, *subsetByOverlaps*, *FeatureMatrix*, *CreateChromatinAssay*); ATAC QC settings as in Signac documentation. 2. Create a common set of peaks (*reduce*, *width*, *combined.peaks*, *FeatureMatrix*, *CreateChromatinAssay*, *RunTFIDF*, *FindTopFeatures*, *RunSVD*). 3. Merge ATAC (*merge*). 4. Integrate ATAC via harmony (*RunHarmony*). 5. Normalize RNA (*SCTransform*). 6. Integrate SCT via PCA, harmony and UMAP (*SelectIntegrationFeatures*, *merge*, *RunPCA*, *RunHarmony*, *RunUMAP*, *FindNeighbors*, *FindClusters*). 7. Merge ATAC and RNA (*FindMultiModalNeighbors*, *RunUMAP*, *FindClusters*).

Single-cell multiome (scRNA and scATAC) inferences for gene regulatory network building are supported by several bioinformatic tools such as SCENIC+, ArchR, Pando and Signac.

ChromVAR, ArchR and SCENIC+ were used according to the official documentation.

Query mapping of scRNA-seq data onto scMultiome data

Query mapping was performed using the integrated functions of the Signac package. Transfer anchors between reference (multiome data) and query (RNA data) were identified (*FindTransferAnchors*). These were then used to map the query onto the reference locations (*MapQuery*).

TCR selection and synthesis for reactivity testing

Analysis of the single-cell RNA + TCR sequencing data, allowed for a guided selection of clonotypes. Initially, the top 10 largest clones were already selected for testing. Additional clones were selected using the following criteria: 1. Select barcodes from clusters of interest (here *Exhausted*, *RM_exh1*, *RM_exh2* and *Proliferating*). 2. Check clonal frequency in that cluster subset. 3. Filter clones by applying the following selection criteria: (A) Percentage of clonotype in that cluster needs to be >5%. (B) Select top 10 hits from each cluster subset. Based on these criteria the clones with highest abundance in one of the respective clusters were selected. Further clones were selected that were abundant in both the EM1 as well as in one of the previously mentioned clusters.

Paired TCR sequences were selected from the sequencing information and similarly designed and ordered as mentioned in the TCR-Engine pipeline⁶⁰. In brief, the sequence design had the following order of components: VDJalpha-TRAC-P2A-VDJbeta-TRBC exon 1 (with splicing domain to splice with the endogenous TRBC exon 2). Sequences were ordered through GenScript.

CRISPR-based generation of TCR libraries in TnT cells for reactivity testing

We used the same CRISPR-Cas9-based knock-in procedure as mentioned in the TCR-Engine publication. In brief, we PCR amplified the HDR template from the received plasmids using Primers RVL127 and RVL128. We mixed up to 43 different HDR templates together at equimolar concentration of 300 ng per plasmid. We then used the Lonza Nucleofector with the SE kit and the “Jurkat E6.1 new” setting with 1×10^5 TnT cells, 250 ng of HDR template mix and 0.7 μ L of anti-TCR-CDR3b-TnT crRNA. After electroporation, TnT cells were cultured in ATCC-modified RPMI-1640 (Thermo Fisher, #A1049101) medium and HDR enhancer (Alt-R HDR Enhancer, IDT) was added at 1:100 for 16 hours. One week post electroporation, TnT cells were FACS sorted using DAPI and anti-CD3 (PE/Cy7, Biolegend) according to the previously mentioned staining protocol. CD3⁺ TnT cells were cryobanked until further use.

Peptides for co-culture

Peptides and peptide pools used in targeted reactivity testing, were diluted in DMSO and added to the co-culture setup at a final concentration of 1 μ g/mL (EBV and TAA peptide pools, and

custom synthesized epitopes) or 2 ug/mL (CMV peptide pool) per peptide according to the manufacturer's instructions.

Co-culture and FACS of reactive TnT cells

Target to effector cells were co-cultured for 16 hours at a 1:1 ratio. Anti-CD28 (final concentration 2 ug/mL, Life technologies) and recombinant human IL-2 (final concentration 50 U/mL, Peprotech, #200-02) were added at the start of the co-culture. As positive control, PMA+Ionomycin (1:500, Life technologies) or PHA (10 ug/mL, Invivogen, #inh-phap) were used. Target cells, expressing patient autologous HLA class I's were generated as described below. Patient material co-culture with TnT cells: One day before co-culture start, tumor suspension and PBMCs were thawed and plated with recombinant human IFN γ at 200 ng/mL (Peprotech, #300-02) (only for tumor suspension) overnight in X-vivo 15 medium (Lonza, #02-060Q) and washed in PBS prior to the co-culture start. Co-culture was started by combining TnT cells and tumor suspension/PBMC cells at a ratio of 1:1. For PBMC co-cultures, peptides were added at above mentioned concentrations.

After 16 hours of co-culture, reactive TnT cells were sorted via anti-CD3 (PE/Cy7, Biolegend), anti-CD69 (APC, Biolegend) and NFAT-GFP (staining protocol was performed as mentioned above). Cells were sorted into 200 uL of FACS buffer, supplemented with 800 uL of TRIZOL (Sigma-Aldrich, #T9424) and frozen at -80 °C.

Library preparation of sorted TnT cells, deep sequencing and data analysis

RNA from sorted and frozen TnT cells in TRIZOL, was extracted using the Arcturus PicoPure RNA Isolation Kit (Life Technologies, #KIT0204), according to the manufacturer's protocol. RNA quantity was determined using a NanoDrop device (ThermoFischer).

cDNA was generated via reverse transcription of an established protocol used in the TCR-Engine publication. In brief, RNA (200 ng) was mixed with Oligo dT Primer (Picelli et al., SMARTSeq2), dNTPs in a total of 36 uL of water. Mix was incubated for 3 min at 72 °C. Following incubation, Maxima RT (2.5 uL, ThermoFischer, #EP0752), Ribolock (1.5 uL, Thermo Scientific, #EO0381) and 5x RT buffer (10 uL, ThermoFischer, #EP0752) were added and the total mix was incubated for 30 min at 50 °C, then 5 min at 85 °C.

RT-PCR was performed using established primers (RVL144 and RVL145) as described in TCR-Engine, using the KAPA Hifi HotStart ReadyMix (Roche, #KK2602). Annealing temperature was set to 62 °C for 25 cycles.

RT-PCR product was QC'ed using the DNA high sensitivity Bioanalyzer Chip (Agilent Technologies). Sample preparation for Illumina deep sequencing was done using the KAPA HyperPrep Kit, PCR-free (Roche, #KK8503) with KAPA Unique-Dual Indexed (UDI) Adapters (Roche, #8861919702). Sequencing was performed on a MiSeq 250PE flow cell (Illumina).

Libraries were sequenced on a MiSeq sequencer (Illumina) for 2 × 250 cycles using MiSeq PE cluster generation kits and MiSeq SBS Kit sequencing reagents (Illumina).

Data analysis was performed in R using the Biostrings package (version 2.68.1).

TCR reactivity classification

TCR clone frequencies were compared for all control co-culture conditions (pre co-culture, PBMC co-culture, PHA-activated co-culture) and the highest frequency per TCR selected. To generate a fold-change for each target condition per TCR clone, each TCR clone target condition frequency was compared to the previously determined, highest scoring, control. TCR reactivity threshold was selected based on the value density plot across all TCR-condition fold change values, which showed a separation from the low-fold changes at 2.08 (as can be seen in Fig 2C right). To determine the uncertainty in our deep sequencing results and experimental conditions, we calculated the difference between replicates (i.e. Patient BS833, clone 6, CMV_A and CMV_B) in the $\log_2()$ space: $\log_2(\text{Fold-change clone 6 CMV_A}) - \log_2(\text{Fold-change clone 6 CMV_B})$, by doing this for all high scoring replicates, we calculated a mean $\log_2()$ -based uncertainty value of 0.47 that could be combined with our previously determined threshold of 2.08 using this formula: $2^{(\log_2(2.08) \pm 0.47)}$. This resulted in upper and lower uncertainty fold-change values of 2.9 and 1.5, rendering all values above 2.9 as highly significantly reactive.

TCR clones with a fold-change value of >2.5 were termed reactive by taking the 50% of the uncertainty ($2.9 - 2.08 = 0.82$) on top of the chosen threshold of 2.08 into account ($2.08 + 0.82/2 = 2.49$).

Neoantigen identification and synthesis

OncoPanel platform (OncoPrint™ Comprehensive Assay Plus) - based analysis identified tumor-specific mutations. The surrounding 9 amino acids (AA) on each side (19 AA total) were selected to generate 9 AA long peptides (11 different peptides per mutation). netMHCpan 4.1 was used to screen peptide sequences for binding affinity to autologous HLAs. The upper 30% of peptides, sorted based on HLA-binding affinity, were selected for synthesis by Genscript.

Generation of patient-specific HLA-expressing cell line

HEK-Blue cells were electroporated with a guide RNA (gRNA) targeting the CCR5 genomic locus in complex with Cas9 recombinant protein (rCas9) and a double-stranded DNA homology-directed repair (HDR) template containing the following elements: (i) left and right homology arms of 924 and 906 bp, respectively mapping to the CCR5 genomic locus; (ii) constitutive chicken beta-actin promoter; (iii) Cas9 encoding gene; (iv) nuclear localization signal; (v) T2A peptide; (vi) puromycin resistance gene; (vii) bGH promoter; (viii) GFP encoding gene. FACS was performed to isolate cells based on GFP expression and a clonal cell line was genotyped and phenotyped to confirm integration and expression of Cas9 from the CCR5 locus. The resulting cell line expressing Cas9 and GFP was subjected to a second step of CRISPR-Cas9 genome editing. Second step, we use a guide RNA (gRNA) targeting the native HLA-ABC genomic locus to knock out the native HLA expression. In a third step of genome editing, we replaced the GFP transgene for a gene encoding the HLA allele of interest. For this purpose, HEK Gen-1 cells were electroporated with a gRNA targeting the GFP transgene located in the CCR5 genomic locus and a double-stranded DNA homology-directed repair (HDR) template containing the following elements: (i) left and right homology arms of 796 and 935 bp (ii) HLA-A*0201, (iii) sv40 terminator. A clonal cell line was genotyped and phenotyped to confirm replacement of GFP with the HLA allele of interest at the CCR5 locus. The resulting cell line expressed Cas9 and the HLA allele of interest.

TCR amplification, sequencing and analysis from single-cell multiome cDNA

An equimolar primer pool of 45 TRBV gene primers was ordered and generated⁹⁷. Primer pool and partial read 1 primer (10x Genomics barcode binding) together with KAPA HiFi ReadyMix were used at 59 °C annealing temperature for 35 cycles, according to the described protocol from

Hudson et al. Following PCR clean up, library preparation was done as mentioned before. Sequencing analysis was done as previously described for TnT sequencing.

Data and Code availability

- Raw FASTQ files from deep sequencing that support the findings of this study will be uploaded to SRA (NCBI) and will be publicly available as of the date of publication, see resources table.
- Original code for the single cell RNA, multiome and deep sequencing analysis has been deposited on Github and will be uploaded to Zenodo, see resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Acknowledgements

We thank Dr. Jack Kuipers for his advice on the statistical analysis of TCR reactivity deep sequencing data. We thank the ETH Zurich D-BSSE Single-Cell Facility for their assistance with FACS, especially Mariangela Di Tacchio and Chiara Cavallini and the ETH Zurich D-BSSE Genomics Facility Basel for their assistance with single cell and deep sequencing experiments, especially Ina Nissen and Dr. Christian Beisel. This study is supported by funding from the Personalized Health and Related Technologies Grant (to S.T.R.) and the NCCR Molecular Systems Engineering (to S.T.R.).

Declaration of interests

R.V.-L and S.T.R. are co-founders and hold shares of Engimmune Therapeutics.. S.T.R. holds shares of Alloy Therapeutics. S.T.R. and A.Z. are on the scientific advisory board of Engimmune Therapeutics. S.T.R is on the scientific advisory board of Alloy Therapeutics.

3.6 References of Chapter 3

1. Coukos, G. TIL Therapy Entering the Mainstream. *N. Engl. J. Med.* **387**, 2185–2186 (2022).
2. Dafni, U., Michielin, O., Lluesma, S. M., Tsourti, Z., Polydoropoulou, V., Karlis, D., Besser, M. J., Haanen, J., Svane, I.-M., Ohashi, P. S., Kammula, U. S., Orcurto, A., Zimmermann, S., Trueb, L., Klebanoff, C. A., Lotze, M. T., Kandalafi, L. E. & Coukos, G. Efficacy of adoptive therapy with tumor-infiltrating lymphocytes and recombinant interleukin-2 in advanced cutaneous melanoma: a systematic review and meta-analysis. *Ann. Oncol.* **30**, 1902–1913 (2019).
3. Rosenberg, S. A. & Dudley, M. E. Adoptive cell therapy for the treatment of patients with metastatic melanoma. *Curr. Opin. Immunol.* **21**, 233–240 (2009).
4. van den Berg, J. H., Heemskerk, B., van Rooij, N., Gomez-Eerland, R., Michels, S., van Zon, M., de Boer, R., Bakker, N. A. M., Jorritsma-Smit, A., van Buuren, M. M., Kvistborg, P., Spits, H., Schotte, R., Mallo, H., Karger, M., van der Hage, J. A., Wouters, M. W. J. M., Pronk, L. M., Geukes Foppen, M. H., Blank, C. U., Beijnen, J. H., Nuijen, B., Schumacher, T. N. & Haanen, J. B. A. G. Tumor infiltrating lymphocytes (TIL) therapy in metastatic melanoma: boosting of neoantigen-specific T cell reactivity and long-term follow-up. *J Immunother Cancer* **8**, (2020).
5. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2018. *CA Cancer J. Clin.* **68**, 7–30 (2018).
6. Chalmers, Z. R., Connelly, C. F., Fabrizio, D., Gay, L., Ali, S. M., Ennis, R., Schrock, A., Campbell, B., Shlien, A., Chmielecki, J., Huang, F., He, Y., Sun, J., Tabori, U., Kennedy, M., Lieber, D. S., Roels, S., White, J., Otto, G. A., Ross, J. S., Garraway, L., Miller, V. A., Stephens, P. J. & Frampton, G. M. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **9**, 34 (2017).
7. Willis, C., Fiander, M., Tran, D., Korytowsky, B., Thomas, J.-M., Calderon, F., Zyczynski, T. M., Brixner, D. & Stenhejm, D. D. Tumor mutational burden in lung cancer: a systematic literature review. *Oncotarget* **10**, 6604–6622 (2019).
8. Renkvist, N., Castelli, C., Robbins, P. F. & Parmiani, G. A listing of human tumor antigens recognized by T cells. *Cancer Immunol. Immunother.* **50**, 3–15 (2001).
9. McGranahan, N., Furness, A. J. S., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S. K., Jamal-Hanjani, M., Wilson, G. A., Birkbak, N. J., Hiley, C. T., Watkins, T. B. K., Shafi, S., Murugaesu, N., Mitter, R., Akarca, A. U., Linares, J., Marafioti, T., Henry, J. Y., Van Allen, E. M., Miao, D., Schilling, B., Schadendorf, D., Garraway, L. A., Makarov, V., Rizvi, N. A., Snyder, A., Hellmann, M. D., Merghoub, T., Wolchok, J. D., Shukla, S. A., Wu, C. J., Peggs, K. S., Chan, T. A., Hadrup, S. R., Quezada, S. A. & Swanton, C. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469 (2016).
10. Philip, M. & Schietinger, A. CD8⁺ T cell differentiation and dysfunction in cancer. *Nat. Rev. Immunol.* **22**, 209–223 (2021).
11. Schietinger, A. & Greenberg, P. D. Tolerance and exhaustion: defining mechanisms of T cell dysfunction. *Trends Immunol.* **35**, 51–60 (2014).
12. Renkvist, N., Castelli, C., Robbins, P. F. & Parmiani, G. A listing of human tumor antigens recognized by T cells. *Cancer Immunol. Immunother.* **50**, 3–15 (2001).
13. ElTanbouly, M. A. & Noelle, R. J. Rethinking peripheral T cell tolerance: checkpoints across a T cell’s journey. *Nat. Rev. Immunol.* **21**, 257–267 (2021).
14. Speiser, D. E., Ho, P.-C. & Verdeil, G. Regulatory circuits of T cell function in cancer. *Nat. Rev. Immunol.* **16**, 599–611 (2016).
15. Zippelius, A., Batard, P., Rubio-Godoy, V., Bioley, G., Liénard, D., Lejeune, F., Rimoldi, D., Guillaume, P., Meidenbauer, N., Mackensen, A., Rufer, N., Lubenow, N., Speiser, D., Cerottini, J.-C., Romero, P. & Pittet, M. J. Effector function of human tumor-specific CD8 T cells in melanoma lesions: a state of local functional tolerance. *Cancer Res.* **64**, 2865–2873 (2004).
16. Coukos, G. TIL Therapy Entering the Mainstream. *N. Engl. J. Med.* **387**, 2185–2186 (2022).
17. Dafni, U., Michielin, O., Lluesma, S. M., Tsourti, Z., Polydoropoulou, V., Karlis, D., Besser, M. J., Haanen, J., Svane, I.-M., Ohashi, P. S., Kammula, U. S., Orcurto, A., Zimmermann, S., Trueb, L., Klebanoff, C. A., Lotze, M. T., Kandalafi, L. E. & Coukos, G. Efficacy of adoptive therapy with tumor-infiltrating lymphocytes and recombinant interleukin-2 in advanced cutaneous melanoma: a systematic review and meta-analysis. *Ann. Oncol.* **30**, 1902–1913 (2019).
18. Rosenberg, S. A. & Dudley, M. E. Adoptive cell therapy for the treatment of patients with metastatic melanoma. *Curr. Opin. Immunol.* **21**, 233–240 (2009).

19. van den Berg, J. H., Heemskerk, B., van Rooij, N., Gomez-Eerland, R., Michels, S., van Zon, M., de Boer, R., Bakker, N. A. M., Jorritsma-Smit, A., van Buuren, M. M., Kvistborg, P., Spits, H., Schotte, R., Mallo, H., Karger, M., van der Hage, J. A., Wouters, M. W. J. M., Pronk, L. M., Geukes Foppen, M. H., Blank, C. U., Beijnen, J. H., Nuijen, B., Schumacher, T. N. & Haanen, J. B. A. G. Tumor infiltrating lymphocytes (TIL) therapy in metastatic melanoma: boosting of neoantigen-specific T cell reactivity and long-term follow-up. *J Immunother Cancer* **8**, (2020).
20. Rohaan, M. W., Borch, T. H., van den Berg, J. H., Met, Ö., Kessels, R., Geukes Foppen, M. H., Stoltenberg Granhøj, J., Nuijen, B., Nijenhuis, C., Jedema, I., van Zon, M., Scheij, S., Beijnen, J. H., Hansen, M., Voermans, C., Noringriis, I. M., Monberg, T. J., Holmstroem, R. B., Wever, L. D. V., van Dijk, M., Grijpink-Ongering, L. G., Valkenet, L. H. M., Torres Acosta, A., Karger, M., Borgers, J. S. W., Ten Ham, R. M. T., Retèl, V. P., van Harten, W. H., Lalezari, F., van Tinteren, H., van der Veldt, A. A. M., Hospers, G. A. P., Stevense-den Boer, M. A. M., Suijkerbuijk, K. P. M., Aarts, M. J. B., Piersma, D., van den Eertwegh, A. J. M., de Groot, J.-W. B., Vreugdenhil, G., Kapiteijn, E., Boers-Sonderen, M. J., Fiets, W. E., van den Berkmortel, F. W. P. J., Ellebaek, E., Hölmich, L. R., van Akkooi, A. C. J., van Houdt, W. J., Wouters, M. W. J. M., van Thienen, J. V., Blank, C. U., Meerveld-Eggink, A., Klobuch, S., Wilgenhof, S., Schumacher, T. N., Donia, M., Svane, I. M. & Haanen, J. B. A. G. Tumor-Infiltrating Lymphocyte Therapy or Ipilimumab in Advanced Melanoma. *N. Engl. J. Med.* **387**, 2113–2125 (2022).
21. Kvistborg, P., Shu, C. J., Heemskerk, B., Fankhauser, M., Thru, C. A., Toebes, M., van Rooij, N., Linnemann, C., van Buuren, M. M., Urbanus, J. H. M., Beltman, J. B., Thor Straten, P., Li, Y. F., Robbins, P. F., Besser, M. J., Schachter, J., Kenter, G. G., Dudley, M. E., Rosenberg, S. A., Haanen, J. B. A. G., Hadrup, S. R. & Schumacher, T. N. M. TIL therapy broadens the tumor-reactive CD8(+) T cell compartment in melanoma patients. *Oncoimmunology* **1**, 409–418 (2012).
22. Kristensen, N. P., Heeke, C., Tvingsholm, S. A., Borch, A., Draghi, A., Crowther, M. D., Carri, I., Munk, K. K., Holm, J. S., Bjerregaard, A.-M., Bentzen, A. K., Marquard, A. M., Szallasi, Z., McGranahan, N., Andersen, R., Nielsen, M., Jönsson, G. B., Donia, M., Svane, I. M. & Hadrup, S. R. Neoantigen-reactive CD8+ T cells affect clinical outcome of adoptive cell therapy with tumor-infiltrating lymphocytes in melanoma. *J. Clin. Invest.* **132**, (2022).
23. Creelan, B. C., Wang, C., Teer, J. K., Toloza, E. M., Yao, J., Kim, S., Landin, A. M., Mullinax, J. E., Saller, J. J., Saltos, A. N., Noyes, D. R., Montoya, L. B., Curry, W., Pilon-Thomas, S. A., Chiappori, A. A., Tanvetyanon, T., Kaye, F. J., Thompson, Z. J., Yoder, S. J., Fang, B., Koomen, J. M., Sarnaik, A. A., Chen, D.-T., Conejo-Garcia, J. R., Haura, E. B. & Antonia, S. J. Tumor-infiltrating lymphocyte treatment for anti-PD-1-resistant metastatic lung cancer: a phase 1 trial. *Nat. Med.* **27**, 1410–1418 (2021).
24. Ma, Y., Ou, J., Lin, T., Chen, L., Wang, J., Qiao, D., Lai, S., Duan, C., Cheng, Y., Chang, R., Zhang, C. & Wang, M. Phenotypic analysis of tumor-infiltrating lymphocytes from non-small cell lung cancer and their potential application for adoptive cell therapy. *Immunopharmacol. Immunotoxicol.* **42**, 319–329 (2020).
25. Poschke, I. C., Hassel, J. C., Rodriguez-Ehrenfried, A., Lindner, K. A. M., Heras-Murillo, I., Appel, L. M., Lehmann, J., Lövgren, T., Wickström, S. L., Lauenstein, C., Roth, J., König, A.-K., Haanen, J. B. A. G., van den Berg, J., Kiessling, R., Bergmann, F., Flossdorf, M., Strobel, O. & Offringa, R. The Outcome of TIL Expansion Is Highly Influenced by Spatial Heterogeneity of the Tumor T-Cell Repertoire and Differences in Intrinsic Growth Capacity between T-Cell Clones. *Clin. Cancer Res.* **26**, 4289–4301 (2020).
26. Tran, E., Robbins, P. F. & Rosenberg, S. A. ‘Final common pathway’ of human cancer immunotherapy: targeting random somatic mutations. *Nat. Immunol.* **18**, 255–262 (2017).
27. Rosenberg, S. A., Yang, J. C., Sherry, R. M., Kammula, U. S., Hughes, M. S., Phan, G. Q., Citrin, D. E., Restifo, N. P., Robbins, P. F., Wunderlich, J. R., Morton, K. E., Laurencot, C. M., Steinberg, S. M., White, D. E. & Dudley, M. E. Durable complete responses in heavily pretreated patients with metastatic melanoma using T-cell transfer immunotherapy. *Clin. Cancer Res.* **17**, 4550–4557 (2011).
28. Lu, Y.-C., Yao, X., Li, Y. F., El-Gamil, M., Dudley, M. E., Yang, J. C., Almeida, J. R., Douek, D. C., Samuels, Y., Rosenberg, S. A. & Robbins, P. F. Mutated PPP1R3B is recognized by T cells used to treat a melanoma patient who experienced a durable complete tumor regression. *J. Immunol.* **190**, 6034–6042 (2013).
29. Khodadoust, M. S., Olsson, N., Wagar, L. E., Haabeth, O. A. W., Chen, B., Swaminathan, K., Rawson, K., Liu, C. L., Steiner, D., Lund, P., Rao, S., Zhang, L., Marceau, C., Stehr, H., Newman, A. M., Czerwinski, D. K., Carlton, V. E. H., Moorhead, M., Faham, M., Kohrt, H. E., Carette, J., Green, M. R., Davis, M. M., Levy, R., Elias, J. E. & Alizadeh, A. A. Antigen presentation profiling reveals recognition of lymphoma immunoglobulin neoantigens. *Nature* **543**, 723–727 (2017).
30. Simoni, Y., Becht, E., Fehlings, M., Loh, C. Y., Koo, S.-L., Teng, K. W. W., Yeong, J. P. S., Nahar, R., Zhang, T., Kared, H., Duan, K., Ang, N., Poidinger, M., Lee, Y. Y., Larbi, A., Khng, A. J., Tan, E., Fu, C., Mathew, R.,

- Teo, M., Lim, W. T., Toh, C. K., Ong, B.-H., Koh, T., Hillmer, A. M., Takano, A., Lim, T. K. H., Tan, E. H., Zhai, W., Tan, D. S. W., Tan, I. B. & Newell, E. W. Bystander CD8 T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* **557**, 575–579 (2018).
31. Cole, D. K., Pumphrey, N. J., Boulter, J. M., Sami, M., Bell, J. I., Gostick, E., Price, D. A., Gao, G. F., Sewell, A. K. & Jakobsen, B. K. Human TCR-binding affinity is governed by MHC class restriction. *J. Immunol.* **178**, 5727–5734 (2007).
 32. Scheper, W., Kelderman, S., Fanchi, L. F., Linnemann, C., Bendle, G., de Rooij, M. A. J., Hirt, C., Mezzadra, R., Slagter, M., Dijkstra, K., Kluin, R. J. C., Snaebjornsson, P., Milne, K., Nelson, B. H., Zijlmans, H., Kenter, G., Voest, E. E., Haanen, J. B. A. G. & Schumacher, T. N. Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers. *Nat. Med.* **25**, 89–94 (2019).
 33. Tan, M. P., Gerry, A. B., Brewer, J. E., Melchiori, L., Bridgeman, J. S., Bennett, A. D., Pumphrey, N. J., Jakobsen, B. K., Price, D. A., Ladell, K. & Sewell, A. K. T cell receptor binding affinity governs the functional profile of cancer-specific CD8+ T cells. *Clin. Exp. Immunol.* **180**, 255–270 (2015).
 34. Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **48**, W449–W454 (2020).
 35. Harndahl, M., Rasmussen, M., Roder, G., Dalgaard Pedersen, I., Sørensen, M., Nielsen, M. & Buus, S. Peptide-MHC class I stability is a better predictor than peptide affinity of CTL immunogenicity. *Eur. J. Immunol.* **42**, 1405–1416 (2012).
 36. Strønen, E., Toebes, M., Kelderman, S., van Buuren, M. M., Yang, W., van Rooij, N., Donia, M., Bösch, M.-L., Lund-Johansen, F., Olweus, J. & Schumacher, T. N. Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* **352**, 1337–1341 (2016).
 37. Altman, J. D., Moss, P. A., Goulder, P. J., Barouch, D. H., McHeyzer-Williams, M. G., Bell, J. I., McMichael, A. J. & Davis, M. M. Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94–96 (1996).
 38. Cohen, C. J., Gartner, J. J., Horovitz-Fried, M., Shamalov, K., Trebska-McGowan, K., Bliskovsky, V. V., Parkhurst, M. R., Ankri, C., Prickett, T. D., Crystal, J. S., Li, Y. F., El-Gamil, M., Rosenberg, S. A. & Robbins, P. F. Isolation of neoantigen-specific T cells from tumor and peripheral lymphocytes. *J. Clin. Invest.* **125**, 3981–3991 (2015).
 39. Parkhurst, M., Gros, A., Pasetto, A., Prickett, T., Crystal, J. S., Robbins, P. & Rosenberg, S. A. Isolation of T-Cell Receptors Specifically Reactive with Mutated Tumor-Associated Antigens from Tumor-Infiltrating Lymphocytes Based on CD137 Expression. *Clin. Cancer Res.* **23**, 2491–2505 (2017).
 40. Lu, Y.-C., Yao, X., Crystal, J. S., Li, Y. F., El-Gamil, M., Gross, C., Davis, L., Dudley, M. E., Yang, J. C., Samuels, Y., Rosenberg, S. A. & Robbins, P. F. Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions. *Clin. Cancer Res.* **20**, 3401–3410 (2014).
 41. Cattaneo, C. M., Battaglia, T., Urbanus, J., Moravec, Z., Voogd, R., de Groot, R., Hartemink, K. J., Haanen, J. B. A. G., Voest, E. E., Schumacher, T. N. & Scheper, W. Identification of patient-specific CD4 and CD8 T cell neoantigens through HLA-unbiased genetic screens. *Nat. Biotechnol.* (2023). doi:10.1038/s41587-022-01547-0
 42. Danilova, L., Anagnostou, V., Caushi, J. X., Sidhom, J.-W., Guo, H., Chan, H. Y., Suri, P., Tam, A., Zhang, J., Asmar, M. E., Marrone, K. A., Naidoo, J., Brahmer, J. R., Forde, P. M., Baras, A. S., Cope, L., Velculescu, V. E., Pardoll, D. M., Housseau, F. & Smith, K. N. The Mutation-Associated Neoantigen Functional Expansion of Specific T Cells (MANAFEST) Assay: A Sensitive Platform for Monitoring Antitumor Immunity. *Cancer Immunol Res* **6**, 888–899 (2018).
 43. Hanada, K.-I., Zhao, C., Gil-Hoyos, R., Gartner, J. J., Chow-Parmer, C., Lowery, F. J., Krishna, S., Prickett, T. D., Kivitz, S., Parkhurst, M. R., Wong, N., Rae, Z., Kelly, M. C., Goff, S. L., Robbins, P. F., Rosenberg, S. A. & Yang, J. C. A phenotypic signature that identifies neoantigen-reactive T cells in fresh human lung cancers. *Cancer Cell* **40**, 479–493.e6 (2022).
 44. Robins, H. S., Campregheer, P. V., Srivastava, S. K., Wachter, A., Turtle, C. J., Kahsai, O., Riddell, S. R., Warren, E. H. & Carlson, C. S. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114**, 4099–4107 (2009).
 45. Redmond, D., Poran, A. & Elemento, O. Single-cell TCRseq: paired recovery of entire T-cell alpha and beta chain transcripts in T-cell receptors from single-cell RNAseq. *Genome Med.* **8**, 80 (2016).
 46. Yusko, E., Vignali, M., Wilson, R. K., Mardis, E. R., Hodi, F. S., Horak, C., Chang, H., Woods, D. M., Robins, H. & Weber, J. Association of Tumor Microenvironment T-cell Repertoire and Mutational Load with Clinical Outcome after Sequential Checkpoint Blockade in Melanoma. *Cancer Immunol Res* **7**, 458–465 (2019).
 47. Riegel, D., Romero-Fernández, E., Simon, M., Adenugba, A. R., Singer, K., Mayr, R., Weber, F., Kleemann, M., Imbusch, C. D., Kreutz, M., Brors, B., Ugele, I., Werner, J. M., Siska, P. J. & Schmidl, C. Integrated

- single-cell profiling dissects cell-state-specific enhancer landscapes of human tumor-infiltrating CD8+ T cells. *Mol. Cell* **83**, 622–636.e10 (2023).
48. Gennert, D. G., Lynn, R. C., Granja, J. M., Weber, E. W., Mumbach, M. R., Zhao, Y., Duren, Z., Sotillo, E., Greenleaf, W. J., Wong, W. H., Satpathy, A. T., Mackall, C. L. & Chang, H. Y. Dynamic chromatin regulatory landscape of human CAR T cell exhaustion. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
 49. Shin, H. M., Kim, G., Kim, S., Sim, J. H., Choi, J., Kim, M., Kwon, M., Ye, S.-K., Lee, D.-S., Cho, S. W., Kim, S. T., Lee, J. & Kim, H.-R. Chromatin accessibility of circulating CD8 T cells predicts treatment response to PD-1 blockade in patients with gastric cancer. *Nat. Commun.* **12**, 975 (2021).
 50. Satpathy, A. T., Granja, J. M., Yost, K. E., Qi, Y., Meschi, F., McDermott, G. P., Olsen, B. N., Mumbach, M. R., Pierce, S. E., Corces, M. R., Shah, P., Bell, J. C., Jhutti, D., Nemec, C. M., Wang, J., Wang, L., Yin, Y., Giresi, P. G., Chang, A. L. S., Zheng, G. X. Y., Greenleaf, W. J. & Chang, H. Y. Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat. Biotechnol.* **37**, 925–936 (2019).
 51. Chen, C., Liu, J., Chen, Y., Lin, A., Mou, W., Zhu, L., Yang, T., Cheng, Q., Zhang, J. & Luo, P. Application of ATAC-seq in tumor-specific T cell exhaustion. *Cancer Gene Ther.* **30**, 1–10 (2023).
 52. Philip, M., Fairchild, L., Sun, L., Horste, E. L., Camara, S., Shakiba, M., Scott, A. C., Viale, A., Lauer, P., Merghoub, T., Hellmann, M. D., Wolchok, J. D., Leslie, C. S. & Schietinger, A. Chromatin states define tumour-specific T cell dysfunction and reprogramming. *Nature* **545**, 452–456 (2017).
 53. Lowery, F. J., Krishna, S., Yossef, R., Parikh, N. B., Chatani, P. D., Zacharakis, N., Parkhurst, M. R., Levin, N., Sindiri, S., Sachs, A., Hitscherich, K. J., Yu, Z., Vale, N. R., Lu, Y.-C., Zheng, Z., Jia, L., Gartner, J. J., Hill, V. K., Copeland, A. R., Nah, S. K., Masi, R. V., Gasmi, B., Kivitz, S., Paria, B. C., Florentin, M., Kim, S. P., Hanada, K.-I., Li, Y. F., Ngo, L. T., Ray, S., Shindorf, M. L., Levi, S. T., Shepherd, R., Toy, C., Parikh, A. Y., Prickett, T. D., Kelly, M. C., Beyer, R., Goff, S. L., Yang, J. C., Robbins, P. F. & Rosenberg, S. A. Molecular signatures of antitumor neoantigen-reactive T cells from metastatic human cancers. *Science* **375**, 877–884 (2022).
 54. Caushi, J. X., Zhang, J., Ji, Z., Vaghasia, A., Zhang, B., Hsiue, E. H.-C., Mog, B. J., Hou, W., Justesen, S., Blosser, R., Tam, A., Anagnostou, V., Cottrell, T. R., Guo, H., Chan, H. Y., Singh, D., Thapa, S., Dykema, A. G., Burman, P., Choudhury, B., Aparicio, L., Cheung, L. S., Lanis, M., Belcaid, Z., El Asmar, M., Illei, P. B., Wang, R., Meyers, J., Schuebel, K., Gupta, A., Skaist, A., Wheelan, S., Naidoo, J., Marrone, K. A., Brock, M., Ha, J., Bush, E. L., Park, B. J., Bott, M., Jones, D. R., Reuss, J. E., Velculescu, V. E., Chaff, J. E., Kinzler, K. W., Zhou, S., Vogelstein, B., Taube, J. M., Hellmann, M. D., Brahmer, J. R., Merghoub, T., Forde, P. M., Yegnasubramanian, S., Ji, H., Pardoll, D. M. & Smith, K. N. Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. *Nature* **596**, 126–132 (2021).
 55. Oliveira, G., Stromhaug, K., Klaeger, S., Kula, T., Frederick, D. T., Le, P. M., Forman, J., Huang, T., Li, S., Zhang, W., Xu, Q., Cieri, N., Clauser, K. R., Shukla, S. A., Neuberg, D., Justesen, S., MacBeath, G., Carr, S. A., Fritsch, E. F., Hacohen, N., Sade-Feldman, M., Livak, K. J., Boland, G. M., Ott, P. A., Keskin, D. B. & Wu, C. J. Phenotype, specificity and avidity of antitumour CD8+ T cells in melanoma. *Nature* **596**, 119–125 (2021).
 56. Zheng, C., Fass, J. N., Shih, Y.-P., Gunderson, A. J., Sanjuan Silva, N., Huang, H., Bernard, B. M., Rajamanickam, V., Slagel, J., Bifulco, C. B., Piening, B., Newell, P. H. A., Hansen, P. D. & Tran, E. Transcriptomic profiles of neoantigen-reactive T cells in human gastrointestinal cancers. *Cancer Cell* **40**, 410–423.e7 (2022).
 57. Palmer, D. C., Webber, B. R., Patel, Y., Johnson, M. J., Kariya, C. M., Lahr, W. S., Parkhurst, M. R., Gartner, J. J., Prickett, T. D., Lowery, F. J., Kishton, R. J., Gurusamy, D., Franco, Z., Vodnala, S. K., Diers, M. D., Wolf, N. K., Slipek, N. J., McKenna, D. H., Sumstad, D., Viney, L., Henley, T., Bürckstümmer, T., Baker, O., Hu, Y., Yan, C., Meerzaman, D., Padhan, K., Lo, W., Malekzadeh, P., Jia, L., Deniger, D. C., Patel, S. J., Robbins, P. F., McIvor, R. S., Choudhry, M., Rosenberg, S. A., Moriarity, B. S. & Restifo, N. P. Internal checkpoint regulates T cell neoantigen reactivity and susceptibility to PD1 blockade. *Med* **3**, 682–704.e8 (2022).
 58. Krishna, S., Lowery, F. J., Copeland, A. R., Bahadiroglu, E., Mukherjee, R., Jia, L., Anibal, J. T., Sachs, A., Adebola, S. O., Gurusamy, D., Yu, Z., Hill, V., Gartner, J. J., Li, Y. F., Parkhurst, M., Paria, B., Kvistborg, P., Kelly, M. C., Goff, S. L., Altan-Bonnet, G., Robbins, P. F. & Rosenberg, S. A. Stem-like CD8 T cells mediate response of adoptive cell immunotherapy against human cancer. *Science* **370**, 1328–1334 (2020).
 59. Gattinoni, L., Klebanoff, C. A. & Restifo, N. P. Paths to stemness: building the ultimate antitumour T cell. *Nat. Rev. Cancer* **12**, 671–684 (2012).
 60. Vazquez-Lombardi, R., Jung, J. S., Schlatter, F. S., Mei, A., Mantuano, N. R., Bieberich, F., Hong, K.-L., Kucharczyk, J., Kapetanovic, E., Aznauryan, E., Weber, C. R., Zippelius, A., Läubli, H. & Reddy, S. T. High-throughput T cell receptor engineering by functional screening identifies candidates with enhanced

- potency and specificity. *Immunity* **55**, 1953–1966.e10 (2022).
61. Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., Wu, N., Zhang, N., Zheng, H., Ouyang, H., Chen, K., Bu, Z., Hu, X., Ji, J. & Zhang, Z. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* **374**, abe6474 (2021).
 62. Gueguen, P., Metoikidou, C., Dupic, T., Lawand, M., Goudot, C., Baulande, S., Lameiras, S., Lantz, O., Girard, N., Seguin-Givelet, A., Lefevre, M., Mora, T., Walczak, A. M., Waterfall, J. J. & Amigorena, S. Contribution of resident and circulating precursors to tumor-infiltrating CD8 T cell populations in lung cancer. *Sci Immunol* **6**, (2021).
 63. Duhén, T., Duhén, R., Montler, R., Moses, J., Moudgil, T., de Miranda, N. F., Goodall, C. P., Blair, T. C., Fox, B. A., McDermott, J. E., Chang, S.-C., Grunkemeier, G., Leidner, R., Bell, R. B. & Weinberg, A. D. Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat. Commun.* **9**, 2724 (2018).
 64. Liu, B., Zhang, Y., Wang, D., Hu, X. & Zhang, Z. Single-cell meta-analyses reveal responses of tumor-reactive CXCL13+ T cells to immune-checkpoint blockade. *Nature Cancer* **3**, 1123–1136 (2022).
 65. Safi, S., Yamauchi, Y., Rathinasamy, A., Stamova, S., Eichhorn, M., Warth, A., Rauch, G., Dienemann, H., Hoffmann, H. & Beckhove, P. Functional T cells targeting tumor-associated antigens are predictive for recurrence-free survival of patients with radically operated non-small cell lung cancer. *Oncoimmunology* **6**, e1360458 (2017).
 66. AACR Project GENIE Consortium. AACR Project GENIE: Powering Precision Medicine through an International Consortium. *Cancer Discov.* **7**, 818–831 (2017).
 67. Deng, Y., Zhang, X., Li, D. & Xu, H. Pan-cancer analysis of the oncogenic effect of SorCS1 in human tumors and its correlation with LRP2 protein. *Research Square* (2022). doi:10.21203/rs.3.rs-2186632/v1
 68. Ardeshir-Larijani, F., Bhateja, P., Lipka, M. B., Sharma, N., Fu, P. & Dowlati, A. KMT2D Mutation Is Associated With Poor Prognosis in Non-Small-Cell Lung Cancer. *Clin. Lung Cancer* **19**, e489–e501 (2018).
 69. Yoshitomi, H., Kobayashi, S., Miyagawa-Hayashino, A., Okahata, A., Doi, K., Nishitani, K., Murata, K., Ito, H., Tsuruyama, T., Haga, H., Matsuda, S. & Toguchida, J. Human Sox4 facilitates the development of CXCL13-producing helper T cells in inflammatory environments. *Nat. Commun.* **9**, 3762 (2018).
 70. Stillwell, R. & Bierer, B. E. T cell signal transduction and the role of CD7 in costimulation. *Immunol. Res.* **24**, 31–52 (2001).
 71. Bignon, A., Régent, A., Klipfel, L., Desnoyer, A., de la Grange, P., Martinez, V., Lortholary, O., Dalloul, A., Mouthon, L. & Balabanian, K. DUSP4-mediated accelerated T-cell senescence in idiopathic CD4 lymphopenia. *Blood* **125**, 2507–2518 (2015).
 72. Yu, X., Harden, K., Gonzalez, L. C., Francesco, M., Chiang, E., Irving, B., Tom, I., Ivelja, S., Refino, C. J., Clark, H., Eaton, D. & Grogan, J. L. The surface protein TIGIT suppresses T cell activation by promoting the generation of mature immunoregulatory dendritic cells. *Nat. Immunol.* **10**, 48–57 (2009).
 73. Bravo González-Blas, C., De Winter, S., Hulselmans, G., Hecker, N., Matetovici, I., Christiaens, V., Poovathingal, S., Wouters, J., Aibar, S. & Aerts, S. SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat. Methods* (2023). doi:10.1038/s41592-023-01938-4
 74. Badia-I-Mompel, P., Wessels, L., Müller-Dott, S., Trimbour, R., Ramirez Flores, R. O., Argelaguet, R. & Saez-Rodriguez, J. Gene regulatory network inference in the era of single-cell multi-omics. *Nat. Rev. Genet.* (2023). doi:10.1038/s41576-023-00618-5
 75. Trefny, M. P., Kirchhammer, N., Auf der Maur, P., Natoli, M., Schmid, D., Germann, M., Fernandez Rodriguez, L., Herzig, P., Löttscher, J., Akrami, M., Stinchcombe, J. C., Stanczak, M. A., Zingg, A., Buchi, M., Roux, J., Marone, R., Don, L., Lardinois, D., Wiese, M., Jeker, L. T., Bentires-Alj, M., Rossy, J., Thommen, D. S., Griffiths, G. M., Läubli, H., Hess, C. & Zippelius, A. Deletion of SNX9 alleviates CD8 T cell exhaustion for effective cellular cancer immunotherapy. *Nat. Commun.* **14**, 86 (2023).
 76. Haen, S. P., Löffler, M. W., Rammensee, H.-G. & Brossart, P. Towards new horizons: characterization, classification and implications of the tumour antigenic repertoire. *Nat. Rev. Clin. Oncol.* **17**, 595–610 (2020).
 77. Kula, T., Dezfulian, M. H., Wang, C. I., Abdelfattah, N. S., Hartman, Z. C., Wucherpfennig, K. W., Lyerly, H. K. & Elledge, S. J. T-Scan: A Genome-wide Method for the Systematic Discovery of T Cell Epitopes. *Cell* **178**, 1016–1028.e13 (2019).
 78. Sharma, G., Rive, C. M. & Holt, R. A. Rapid selection and identification of functional CD8+ T cell epitopes from large peptide-coding libraries. *Nat. Commun.* **10**, 4553 (2019).
 79. Gee, M. H., Han, A., Lofgren, S. M., Beausang, J. F., Mendoza, J. L., Birnbaum, M. E., Bethune, M. T., Fischer, S., Yang, X., Gomez-Eerland, R., Bingham, D. B., Sibener, L. V., Fernandes, R. A., Velasco, A., Baltimore, D., Schumacher, T. N., Khatri, P., Quake, S. R., Davis, M. M. & Garcia, K. C. Antigen Identification

- for Orphan T Cell Receptors Expressed on Tumor-Infiltrating Lymphocytes. *Cell* **172**, 549–563.e16 (2018).
80. Lee, M. N. & Meyerson, M. Antigen identification for HLA class I- and HLA class II-restricted T cell receptors using cytokine-capturing antigen-presenting cells. *Sci. Immunol.* **6**, eabf4001 (2021).
 81. Tietscher, S., Wagner, J., Anzeneder, T., Langwieder, C., Rees, M., Sobottka, B., de Souza, N. & Bodenmiller, B. A comprehensive single-cell map of T cell exhaustion-associated immune environments in human breast cancer. *Nat. Commun.* **14**, 98 (2023).
 82. Borsa, M., Barnstorf, I., Baumann, N. S., Pallmer, K., Yermanos, A., Gräbnitz, F., Barandun, N., Hausmann, A., Sandu, I., Barral, Y. & Oxenius, A. Modulation of asymmetric cell division as a mechanism to boost CD8 T cell memory. *Sci Immunol* **4**, (2019).
 83. Gill, A. L., Wang, P. H., Lee, J., Hudson, W. H., Ando, S., Araki, K., Hu, Y., Wieland, A., Im, S., Gavora, A., Medina, C. B., Freeman, G. J., Hashimoto, M., Reiner, S. L. & Ahmed, R. PD-1 blockade increases the self-renewal of stem-like CD8 T cells to compensate for their accelerated differentiation into effectors. *Sci Immunol* **8**, eadg0539 (2023).
 84. Thommen, D. S., Koelzer, V. H., Herzig, P., Roller, A., Trefny, M., Dimeloe, S., Kiialainen, A., Hanhart, J., Schill, C., Hess, C., Savic Prince, S., Wiese, M., Lardinois, D., Ho, P.-C., Klein, C., Karanikas, V., Mertz, K. D., Schumacher, T. N. & Zippelius, A. A transcriptionally and functionally distinct PD-1 CD8 T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. *Nat. Med.* **24**, 994–1004 (2018).
 85. Luoma, A. M., Suo, S., Wang, Y., Gunasti, L., Porter, C. B. M., Nabils, N., Tadros, J., Ferretti, A. P., Liao, S., Gurer, C., Chen, Y.-H., Criscitiello, S., Ricker, C. A., Dionne, D., Rozenblatt-Rosen, O., Uppaluri, R., Haddad, R. I., Ashenberg, O., Regev, A., Van Allen, E. M., MacBeath, G., Schoenfeld, J. D. & Wucherpfennig, K. W. Tissue-resident memory and circulating T cells are early responders to pre-surgical cancer immunotherapy. *Cell* **185**, 2918–2935.e29 (2022).
 86. Baulu, E., Gardet, C., Chuvin, N. & Depil, S. TCR-engineered T cell therapy in solid tumors: State of the art and perspectives. *Sci Adv* **9**, eadf3700 (2023).
 87. Malviya, M., Aretz, Z. E. H., Molvi, Z., Lee, J., Pierre, S., Wallisch, P., Dao, T. & Scheinberg, D. A. Challenges and solutions for therapeutic TCR-based agents. *Immunol. Rev.* (2023). doi:10.1111/imr.13233
 88. Yang, X., Nishimiya, D., Löchte, S., Jude, K. M., Borowska, M., Savvides, C. S., Dougan, M., Su, L., Zhao, X., Piehler, J. & Others. Facile repurposing of peptide--MHC-restricted antibodies for cancer immunotherapy. *Nat. Biotechnol.* 1–12 (2023).
 89. Dolton, G., Rius, C., Wall, A., Szomolay, B., Bianchi, V., Galloway, S. A. E., Hasan, M. S., Morin, T., Caillaud, M. E., Thomas, H. L., Theaker, S., Tan, L. R., Fuller, A., Topley, K., Legut, M., Attaf, M., Hopkins, J. R., Behiry, E., Zabkiewicz, J., Alvares, C., Lloyd, A., Rogers, A., Henley, P., Fegan, C., Ottmann, O., Man, S., Crowther, M. D., Donia, M., Svane, I. M., Cole, D. K., Brown, P. E., Rizkallah, P. & Sewell, A. K. Targeting of multiple tumor-associated antigens by individual T cell receptors during successful cancer immunotherapy. *Cell* **186**, 3333–3349.e27 (2023).
 90. Wang, S., Sun, J., Chen, K., Ma, P., Lei, Q., Xing, S., Cao, Z., Sun, S., Yu, Z., Liu, Y. & Li, N. Perspectives of tumor-infiltrating lymphocyte treatment in solid tumors. *BMC Med.* **19**, 140 (2021).
 91. Mo, F., Yu, Z., Li, P., Oh, J., Spolski, R., Zhao, L., Glassman, C. R., Yamamoto, T. N., Chen, Y., Golebiowski, F. M., Hermans, D., Majri-Morrison, S., Picton, L. K., Liao, W., Ren, M., Zhuang, X., Mitra, S., Lin, J.-X., Gattinoni, L., Powell, J. D., Restifo, N. P., Garcia, K. C. & Leonard, W. J. An engineered IL-2 partial agonist promotes CD8 T cell stemness. *Nature* **597**, 544–548 (2021).
 92. Codarri Deak, L., Nicolini, V., Hashimoto, M., Karagianni, M., Schwalie, P. C., Lauener, L., Varypataki, E. M., Richard, M., Bommer, E., Sam, J., Joller, S., Perro, M., Cremasco, F., Kunz, L., Yanguez, E., Hüsser, T., Schlenker, R., Mariani, M., Tosevski, V., Herter, S., Bacac, M., Waldhauer, I., Colombetti, S., Gueripel, X., Wullschleger, S., Tichet, M., Hanahan, D., Kissick, H. T., Leclair, S., Freimoser-Grundschober, A., Seeber, S., Teichgräber, V., Ahmed, R., Klein, C. & Umaña, P. PD-1-cis IL-2R agonism yields better effectors from stem-like CD8 T cells. *Nature* **610**, 161–172 (2022).
 93. Sarén, T., Ramachandran, M., Gammelgård, G., Lövgren, T., Mirabello, C., Björklund, Å. K., Wikström, K., Hashemi, J., Freyhult, E., Ahlström, H., Amini, R.-M., Hagberg, H., Loskog, A., Enblad, G. & Essand, M. Single-cell RNA analysis reveals cell-intrinsic functions of CAR-T cells correlating with response in a phase II study of lymphoma patients. *Clin. Cancer Res.* (2023). doi:10.1158/1078-0432.CCR-23-0178
 94. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., 3rd, Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., Hoffman, P., Stoeckius, M., Papalexi, E., Mimitou, E. P., Jain, J., Srivastava, A., Stuart, T., Fleming, L. M., Yeung, B., Rogers, A. J., McElrath, J. M., Blish, C. A., Gottardo, R., Smibert, P. & Satija, R. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
 95. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with

- Signac. *Nat. Methods* **18**, 1333–1341 (2021).
96. Granja, J. M., Corces, M. R., Pierce, S. E., Bagdatli, S. T., Choudhry, H., Chang, H. Y. & Greenleaf, W. J. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).
 97. Hudson, W. H. & Sudmeier, L. J. Localization of T cell clonotypes using the Visium spatial transcriptomics platform. *STAR Protoc* **3**, 101391 (2022).

Chapter 4: General discussion

4.1 Conclusion

The adaptive immune system, with its high complexity, plays a key role in the body's defense mechanisms against infections and cancer. Since its introduction, single-cell sequencing has emerged as a crucial tool for its analysis, offering an unprecedented, unbiased, and detailed view of complex systems like the adaptive immune system. As a result, it has facilitated a better understanding of the variations in the immune system across individuals and diseases ¹. The integration of additional modalities, such as VDJ receptor sequencing and chromatin accessibility analysis, with single-cell transcriptomic analysis has enabled the detailed single-cell assessment of T cell clones (cells sharing the same VDJ receptor). Further, it facilitated the discrimination of cell differentiation states that are not discernible through transcriptomic data alone. This allowed for deeper insights into the immune cell interaction in disease contexts and elucidated the cell state of reactive cells during or after disease by integrating functional reactivity testing ²⁻⁴.

This thesis applied multimodal single-cell sequencing technologies and a combination with functional reactivity testing to human adaptive immune cells upon resolved COVID-19 infection and to CD8⁺ T cells from lung cancer patient biopsies, respectively.

For the immune response to COVID-19, prior work established the difference in immune responses to an influenza infection as contributor to varying disease severities ⁵⁻⁷ and identified age as a cause for a differing immune response ⁸⁻¹⁰. However, to our knowledge this is the first time that the convalescent COVID-19 immune state has been analyzed for individuals of differing age and disease duration groups.

Regarding the immune response to cancer, a large body of prior work discovered T cells' causal and contributing role to cancer regression and applications in therapy ¹¹⁻¹³. Additionally, recent studies that identified tumor-reactive T cells to be enriched in exhausted cell states ¹⁴, lead us to ask if the efficient enrichment of tumor-reactive T cells based on gene-expression patterns is possible.

The work conducted in this thesis could show the bifurcation of patient immune responses to resolved mild-to-moderate COVID-19 infection and its imprint into adaptive immune clonality and cell states, while also differing based on age. Further, using an unbiased functional reactivity testing platform, we could identify tumor- and virus-reactive T cells in human lung cancer biopsies. This ultimately enabled us to develop a tumor-reactive T cell signature that shows potential for enriching tumor-reactive T cells.

4.2 Contributions of this thesis

Chapter 2: Adaptive immune repertoire of convalescent COVID-19 patients

By performing single-cell sequencing of the transcriptome and VDJ receptors of T and B cells from convalescent COVID-19 patients, we identified differences in the lasting immune response between young and old patients, with patient age being correlated with disease duration.

An in-depth analysis of B cell dynamics showed that young patients had a higher rate of B cell expansion and somatic hypermutations in the BCRs. Although younger patients exhibited higher levels of IgA switching, older patients demonstrated greater levels of SARS-CoV-2-specific IgA/G in serum. This is likely explained by the fact that old patients had a broader range of CDR3 heavy chain lengths, indicating that old patients harbor a higher diversity of unexpanded SARS-CoV-2-specific B cells.

T cell analysis revealed a higher CD8 to CD4 ratio in young patients, whereas old patients had higher proliferation marker expression in the memory T cell compartment. This indicates an earlier T cell activation and expansion, especially of the cytotoxic T cells, in young patients and a more recent viral clearance in old patients.

When investigating the TCR repertoire in combination with the associated clonal cell states, we identified a frequent occurrence of TRBV20-1 in 7 out of 8 patients, hinting at clonal convergence (possibly against SARS-CoV-2). Further, two patients (one old and one young) that had the highest clonal expansion level and shared the HLA-A allele A*0201 had multiple overlapping TRBV and TRAV genes in their most expanded clonotypes, indicating a possible germline convergence that may be related to SARS-CoV-2 specificity. *In silico* matching of TCR sequences with known SARS-CoV-2-reactive TCRs showed a shared reactivity group across all three HLA A*0201 patients and known SARS-CoV-2-reactive TCRs.

This approach shows the potential of understanding adaptive immune cell status upon disease convalescence, showcasing single-cell immune profiling as a meaningful way to identify long-term immune state changes and remnants that can give clues on the prior disease state. Combining transcriptomic information and paired VDJ receptor sequences offers a deeper understanding of clonal convergence, reactivities and clonal phenotypes.

Due to the difficulty and scarcity of sample acquisition at the early stage of the pandemic, this study also comes with limitations. In our study, patient age is correlated with disease duration/severity, which makes it challenging to attribute the differences among patient groups to patient age or disease duration. To fully understand the disease dynamics, more comprehensive studies are required, including healthy control groups, as this will help determine whether the observed effects between disease duration and age are linked. In the last part of the study, we identified reactivity groups between patient TCRs and known SARS-CoV-2-specific TCRs. While this is a strong hint at the TCR's reactivity, functional validation is necessary to confidently attribute reactivity to these TCRs, work that has been the focus of chapter 3 of this thesis.

Chapter 3: Unbiased identification of tumor-reactive T cells in lung cancer

Transcriptome-based identification of reactive T cells has recently been shown to be possible across different cancers and infectious diseases with varying enrichment potential^{15–20}. Tumor-reactive T cells aid in tumor regression and correlate with survival. Thus, there is a need for accelerated discovery of tumor-reactive T cells in a personalized manner. By developing a workflow for the unbiased discovery of tumor-reactive T cells and combining it with multimodal single-cell sequencing technologies, we were able to specify the phenotypic markers of tumor-reactive T cells and develop a T cell tumor-reactivity score for untreated lung cancer patient samples.

Due to the high exhaustion levels in TILs and the potential hypo-responsivity, direct utilization of TILs for reactivity testing can introduce biases for cells that have the ability to get activated when encountering their cognate target antigen. Thus, we first developed an unbiased reactivity testing platform that allowed us to test large numbers of TCRs while enabling all TCRs to have the same cellular activation threshold upon cognate target antigen encounter.

By using TnT cells that were previously developed in the lab and combining them with FACS-based T cell activation sorting and NGS readout of the sorted TCRs, we could test ca. 140 TCRs across four patients for their reactivity. This led to the identification of 18 TCRs that exerted reactivity against neo-antigens, tumor-associated antigens and viral antigens. The low number of neo-antigen-reactive T cells matches the literature, with most solid tumors having observed neo-antigen-reactive TIL frequencies of 0-0.5% ²¹⁻²³. Further the abundance of virus-reactive TILs is in line with previous work showing a considerable amount of bystander TILs that elicit reactivity against viral antigens ^{24,25}.

Combining functional reactivity readout with previously obtained single-cell sequencing data, allowed us to identify the clonal phenotype for each reactive, unreactive, and untested TCR. By comparing the phenotypes of tumor- and virus-reactive clones, we identified differentially expressed genes specific to tumor-reactive TILs. Combining these marker genes into a tumor score and applying it to our generated data allowed us to, in theory, enrich tumor-reactive T cells ca. 3-fold compared to the unbiased selection we used initially. This paves the way towards a potential *in silico* selection of tumor-reactive TCRs without the need for laborious functional reactivity validation.

Of the genes identified as differentially expressed in tumor-reactive T cells, approximately one-third overlapped with marker genes documented in previous research ¹⁶⁻²⁰. What is important to note here is that the current most used tumor-reactivity markers CXCL13 ²⁶ and ENTPD1 ¹⁴ are part of our set of 22 genes, highlighting both the agreement with previous work and the identification of novel marker genes. The differences in marker gene sets could be due to the unbiased approach that we chose to perform or the usage of virus-reactive T cells as the group of T cells to compare the tumor-reactive T cells against. Using virus-reactive T cells as comparison was chosen because it is notoriously challenging to characterize a T cell as non-tumor-reactive because tumor antigens can be derived from most genomic mutations as well as from changes in chromatin remodeling and post-transcriptional processes ²⁷⁻³⁰. In addition, clonal editing and cancer evolution causes previously expressed targets to be no longer abundant or expressed ³¹, making potential targets difficult to identify and often too many to screen. Since the virus-reactive T cells showed sole reactivity to our viral antigens, we thus used these as a non-tumor reactive group, hypothesizing that their infiltration was due to their previous viral antigen encounter in the body and their cell state unrelated to tumor-reactivity.

The work we elaborated on in Chapter 3 introduced an efficient way to screen TCRs for tumor-reactivity. The combination of the screening method and multimodal single-cell sequencing allowed for the identification of marker genes and the development of a tumor-reactivity score that has shared abundances in the proliferation, exhaustion and resident memory clusters. However, it is important to note that our tumor score was developed based on a limited number of patient samples and TCRs. While initial data suggests the score's effectiveness in identifying tumor-reactive T cells, its validity and applicability across diverse patient groups and tumor types necessitate further evaluation on a more extensive dataset of TCRs and patient samples. Eventually, enhancing the speed and reducing the cost of identifying tumor-reactive T cells are expected to significantly improve the efficacy of therapies and pave the way for a more accessible and personalized cancer treatment landscape.

4.3 Future directions

4.3.1 Adaptive immune state assessment for personalized medicine

Monitoring the immune state prior to, during and after infection can give insights into the therapeutic importance of specific immune cell types and identify diagnostic markers. For example, in COVID-19, an overreaction of the immune system and T cell apoptosis strongly contributed to the fatality rate upon SARS-CoV-2 contraction ^{32,33}.

In Chapter 2, we show that patients with convalescent COVID-19 exhibited a continued imprint of their immune response both at cell state and at immune receptor repertoire levels, correlating with disease duration and age. Additional studies have identified differential immune responses between mild and severe cases that allowed for deeper insights into the disease and immune mechanisms, such as higher antibody affinity maturation and TCR repertoire diversity being associated with mild COVID-19 ^{34,35}, and lymphocyte exhaustion with severe COVID-19 ³⁶.

These findings suggest that monitoring of a patient's immune receptor repertoire can give insights into the course of the disease with potential predictive power. While more data is needed to identify clear biomarkers that are unbiased from the patient's metadata (i.e., age, medication, etc.), better integration of next-generation sequencing in the clinical setting could lead to more accurate diagnosis and personalized treatments. To illustrate, a recent study introduced the

possibility of NGS-based tumor classification during surgery, highlighting that NGS-based technologies are feasible and complementary for certain diseases and procedures ³⁷.

An additional hurdle for applying biomedical research findings obtained through single-cell sequencing technologies ³⁸ in the clinic, is scalability. Therefore, future biomarker identification needs to combine the high resolution of single-cell sequencing with the scalability of FACS or bulk-RNA sequencing technologies to generate actionable insights for patients. In addition, current developments in single-cell sequencing technologies may lead to a drastic reduction in price, similar to what we have witnessed with next-generation sequencing, which in turn would also allow for better scalability.

Thus, a more detailed assessment of the adaptive immune state and the use of scalable technologies can help to rapidly understand the disease and immune response mechanisms and improve the development of not only novel therapies but also i.e., vaccines by facilitating comparative analyses of immune system alterations between natural and vaccine-induced infection ³⁹⁻⁴³.

The interest in deciphering disease mechanisms through the systematic profiling of every human cell type has led to the initiation of the Human Cell Atlas ⁴⁴. This ambitious project seeks to accomplish its goals by conducting large-scale single-cell sequencing experiments across various research groups and computationally integrating the resulting data. Due to the complex interactions of the human immune system, similar work is being done to systematically map the immune system and evaluate cytokine function in order to understand the differences between health and disease ^{45,46}.

4.3.2 Tumor reactive T cell discovery for personalized therapies

Cancer immunotherapy has led to significant breakthroughs and improvements in treatment methods, since its emergence over two decades ago. Most recently, adoptive cell transfer (ACT) of TILs has been shown to significantly enhance overall survival rates by more than double in late-stage melanoma patients resistant to anti-PD1 ICB, compared to the standard of care, anti-CTLA4 ICB ⁴⁷.

Complementing this, previous work has shown that increasing the fraction of tumor-reactive T cells in the ACT-TIL infusion product enhances therapeutic efficacy, emphasizing the importance of these cells in effective cancer immunotherapy ^{14,26,48-51}. This has spurred growing interest in

the rapid identification and selection of tumor-reactive T cells for personalized treatment approaches.

In Chapter 3, we demonstrated the ability to unbiasedly discover tumor-reactive T cells and improve their selection using gene-expression markers. Concurrently, several studies have identified tumor-reactive T cell marker genes across various cancer types, each employing a different set of methodologies. Across studies, there is a significant overlap of identified marker genes which gives hope towards finding a universal set of markers with the highest predictive power across tumor entities. However, given the substantial variability in patient backgrounds, tumor types, and stages, the task of developing personalized tumor-reactive T cell marker sets presents significant challenges.

To overcome this, a large-scale multi-center effort, similar to the human cell atlas, with the following objectives is needed: I) Collection of comprehensive patient metadata, similar to the approach of the UK Biobank, including whole exome sequencing and long-term follow-up; II) Integrating existing datasets across various tumor types; III) Validating predicted tumor-reactive TCRs using standardized reactivity testing methods.

Making this data widely available for machine learning applications could not only aid in developing personalized tumor-reactive T cell biomarkers but also enhance the criteria for stratifying patient subgroups, focusing on TIL cell states and repertoire dynamics⁵².

To efficiently carry out the proposed reactivity testing of TCRs, recently developed high-throughput *in vitro*^{53–55} and *in silico*^{56–58} methods that make use of library screening and structural TCR-peptide features, respectively, have to be employed. In addition to verifying the TCR selection procedure, these technologies will facilitate the creation of a database of tumor-reactive TCR-peptide-HLA pairs (similar to VDJdb⁵⁹) that can be applied in personalized clinical strategies, which we will delve into in the subsequent sections. In conclusion, the identification of tumor-reactive T cells holds potential to impact both diagnostic and therapeutic areas.

In a diagnostic context, evaluating the size of the tumor-reactive T-cell fraction and their associated cell states could inform treatment choices, like determining the need to introduce off-the-shelf TCR-T cells or to rejuvenate (reverse exhaustion) existing T cells. Another use case could be the monitoring of treatment response through continuous screening of tumor-reactive T

cell abundance. More precisely, this strategy can be implemented by initiating the identification and tracking of a patient's tumor-reactive T cells *in vitro* or *in silico* prior to treatment, followed by a continuous assessment of the tumor-reactive T cell fraction ^{60,61}.

In a therapeutic context, the cancer therapy field is moving towards genome-centric treatments, with mutation-targeting therapeutics being cross-approved for different cancer entities ⁶² and in turn shifting away from previous approaches that were tumor tissue-specific. While TCRs are acting in a genome-centric way, due to recognition of non-self peptides which are derived from i.e. genomic mutations, the accessibility and immunosuppression of the target tissue still need to be taken into account for personalized applications.

To improve the effectiveness of TCR therapies, it is essential to comprehend not only the abundance and availability of the TCR's peptide-HLA target but also how various T cell reactivities and their corresponding targets influence treatment outcomes. An interesting observation is the clonal convergence towards dominant low-affinity TCRs during chronic infections, which may be a mechanism to balance targeting efficacy with immunopathology. This phenomenon could also indicate the essential role of low-affinity TCRs in mounting an effective and safe immune response ^{63,64}. It was also shown that bystander T cells' inflammatory response upon intratumoral injection of pathogenic peptides can lead to higher immune infiltration and anti-tumor response ⁶⁵. Thus, it needs to be systematically investigated how bystander T cells, and low- and high-affinity tumor-reactive T cells each by themselves and in combination can contribute to tumor regression ²⁵. A comprehensive database comprising characterized tumor-reactive and bystander TCRs would facilitate a systematic analysis of the effects that various TCR combinations have on tumor regression *in vivo*. This could pave the way towards identifying the most efficacious combination of TCR reactivities for personalized therapeutic use.

Compared to antibodies, TCRs require a matching peptide and HLA to exert their function, making HLA information a crucial patient identifier for treatment with TCR therapies. The recent approval of Tebentafusp, a soluble TCR-bispecific T cell engager, which necessitates the expression of HLA-A*0201 for its functionality and administration, underscores the importance

of routinely obtaining patient HLA information ⁶⁶. Further development of tumor-reactive TCRs into off-the-shelf TCR therapeutics that are tailored to a patient's tumor mutational signature and corresponding HLA type could be made feasible through the previously mentioned database of tumor-reactive TCRs and associated peptide-HLA complexes.

While most TCR therapeutics are currently being applied as monoclonal therapeutics, targeting only a single peptide-HLA complex, systematic characterisation of tumor-reactive TCRs may also allow for polyclonal efforts in the future, targeting more than one peptide-HLA complex, potentially reducing therapeutic resistance through loss of antigen or HLA expression ^{67,68}. The high cost and turnaround time necessary for the generation of an autologous T cell product, an approach mainly used for CAR T cell and TIL therapy, have led to a growing interest in transitioning into off-the-shelf strategies ⁶⁹.

Emerging alternatives such as soluble TCRs and allogeneic cell therapy hold promise to revolutionize the treatment landscape and have recently gained first approvals or are still undergoing clinical trials, respectively ^{66,70}. However, each of these approaches comes with its own set of challenges and risks that go beyond identifying TCRs with known reactivity and sufficient affinity.

4.3.3 TCR identification for other diseases

While the main focus of TCR discovery in the past has been on virus and tumor-reactive TCRs, one can readily envision that accelerated discovery of reactive TCRs will also impact other disease areas. To focus on one, the class of autoimmune disorders is often characterized by targeted tissue pathology, with frequent T cell involvement, and has long struggled to identify pathogenic TCRs and their cognate peptide-HLA targets. This has made it challenging to understand disease mechanisms and develop treatments, with current therapeutic approaches mainly focusing on general immunosuppression. Future therapies must be more targeted to the pathogenic cellular subsets to alleviate the detrimental side effects of immunosuppressive therapies, such as increased cancer risk⁷¹ and the inability to fight off infections. Recent breakthroughs have been achieved with an anti-CD19 CAR-T cell therapy and targeted depletion of TRBV9-expressing T cells to treat systemic lupus erythematosus and ankylosing spondylitis, respectively ^{72,73}, giving hope that targeted therapies can have lasting and systemic effects. The advent of previously mentioned high-throughput *in vitro* and *in silico* TCR-peptide-HLA

screening platforms, combined with reactive T cell phenotyping will enable the accelerated identification of autoreactive TCRs and their cognate target peptide-HLAs. Such findings will not only help the understanding of the detailed disease mechanism but will also allow an improved subgrouping of patients, based on common sets of target antigens and autoreactive T cell states, for a more customized treatment.

In addition, regulatory T cells (Tregs) are known for their immunosuppressive functions, making them a viable option for treating autoimmune disorders. However, the ability of Treg cells to induce immunosuppression in a non-specific manner raises concerns about their potential for systemic immunosuppression ⁷⁴. In contrast, antigen-specific Treg cells are able to selectively relocate to the target tissue to exert their immunosuppressive function ⁷⁵, showing how improved identification of autoreactive TCRs could also improve Treg therapies.

4.4 References of Chapter 4

1. Stubbington, M. J. T., Rozenblatt-Rosen, O., Regev, A. & Teichmann, S. A. Single-cell transcriptomics to explore the immune system in health and disease. *Science* **358**, 58–63 (2017).
2. Stubbington, M. J. T., Lönnberg, T., Proserpio, V., Clare, S., Speak, A. O., Dougan, G. & Teichmann, S. A. T cell fate and clonality inference from single-cell transcriptomes. *Nat. Methods* **13**, 329–332 (2016).
3. Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing, R., Gao, R., Zhang, L., Dong, M., Hu, X., Ren, X., Kirchhoff, D., Roeder, H. G., Yan, T. & Zhang, Z. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat. Med.* **24**, 978–985 (2018).
4. Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., Wu, N., Zhang, N., Zheng, H., Ouyang, H., Chen, K., Bu, Z., Hu, X., Ji, J. & Zhang, Z. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* **374**, abe6474 (2021).
5. McElhaney, J. E., Verschoor, C. P., Andrew, M. K., Haynes, L., Kuchel, G. A. & Pawelec, G. The immune response to influenza in older humans: beyond immune senescence. *Immun. Ageing* **17**, 10 (2020).
6. Dunning, J., Blankley, S., Hoang, L. T., Cox, M., Graham, C. M., James, P. L., Bloom, C. I., Chaussabel, D., Banchereau, J., Brett, S. J., Moffatt, M. F., O’Garra, A., Openshaw, P. J. M. & MOSAIC Investigators. Progression of whole-blood transcriptional signatures from interferon-induced to neutrophil-associated patterns in severe influenza. *Nat. Immunol.* **19**, 625–635 (2018).
7. Arunachalam, P. S., Wimmers, F., Mok, C. K. P., Perera, R. A. P. M., Scott, M., Hagan, T., Sigal, N., Feng, Y., Bristow, L., Tak-Yin Tsang, O., Wagh, D., Coller, J., Pellegrini, K. L., Kazmin, D., Alaaeddine, G., Leung, W. S., Chan, J. M. C., Chik, T. S. H., Choi, C. Y. C., Huerta, C., Paine McCullough, M., Lv, H., Anderson, E., Edupuganti, S., Upadhyay, A. A., Bosinger, S. E., Maecker, H. T., Khatri, P., Roupheal, N., Peiris, M. & Pulendran, B. Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science* **369**, 1210–1220 (2020).
8. Martinez-Jimenez, C. P., Eling, N., Chen, H.-C., Vallejos, C. A., Kolodziejczyk, A. A., Connor, F., Stojic, L., Rayner, T. F., Stubbington, M. J. T., Teichmann, S. A., de la Roche, M., Marioni, J. C. & Odom, D. T. Aging increases cell-to-cell transcriptional variability upon immune stimulation. *Science* **355**, 1433–1436 (2017).
9. Kasler, H. & Verdin, E. How inflammaging diminishes adaptive immunity. *Nat Aging* **1**, 24–25 (2021).
10. Weng, N.-P. Aging of the immune system: how much can the adaptive immune system adapt? *Immunity* **24**, 495–499 (2006).
11. Fridman, W. H., Pagès, F., Sautès-Fridman, C. & Galon, J. The immune contexture in human tumours: impact on clinical outcome. *Nat. Rev. Cancer* **12**, 298–306 (2012).
12. Gooden, M. J. M., de Bock, G. H., Leffers, N., Daemen, T. & Nijman, H. W. The prognostic influence of tumour-infiltrating lymphocytes in cancer: a systematic review with meta-analysis. *Br. J. Cancer* **105**, 93–103 (2011).
13. Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., Tosolini, M., Camus, M., Berger, A., Wind, P., Zinzindohoué, F., Bruneval, P., Cugnenc, P.-H., Trajanoski, Z., Fridman, W.-H. & Pagès, F. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* **313**, 1960–1964 (2006).
14. Duhén, T., Duhén, R., Montler, R., Moses, J., Moudgil, T., de Miranda, N. F., Goodall, C. P., Blair, T. C., Fox, B. A., McDermott, J. E., Chang, S.-C., Grunkemeier, G., Leidner, R., Bell, R. B. & Weinberg, A. D. Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat. Commun.* **9**, 2724 (2018).
15. Fischer, D. S., Ansari, M., Wagner, K. I., Jarosch, S., Huang, Y., Mayr, C. H., Strunz, M., Lang, N. J., D’Ippolito, E., Hammel, M., Mateyka, L., Weber, S., Wolff, L. S., Witter, K., Fernandez, I. E., Leuschner, G., Milger, K., Frankenberger, M., Nowak, L., Heinig-Menhard, K., Koch, I., Stoleriu, M. G., Hilgendorff, A., Behr, J., Pichlmair, A., Schubert, B., Theis, F. J., Busch, D. H., Schiller, H. B. & Schober, K. Single-cell RNA sequencing reveals ex vivo signatures of SARS-CoV-2-reactive T cells through ‘reverse phenotyping’. *Nat. Commun.* **12**, 4515 (2021).
16. Hanada, K.-I., Zhao, C., Gil-Hoyos, R., Gartner, J. J., Chow-Parmer, C., Lowery, F. J., Krishna, S., Prickett, T. D., Kivitz, S., Parkhurst, M. R., Wong, N., Rae, Z., Kelly, M. C., Goff, S. L., Robbins, P. F., Rosenberg, S. A. & Yang, J. C. A phenotypic signature that identifies neoantigen-reactive T cells in fresh human lung cancers. *Cancer Cell* **40**, 479–493.e6 (2022).
17. Lowery, F. J., Krishna, S., Yossef, R., Parikh, N. B., Chatani, P. D., Zacharakis, N., Parkhurst, M. R., Levin, N.,

- Sindiri, S., Sachs, A., Hitscherich, K. J., Yu, Z., Vale, N. R., Lu, Y.-C., Zheng, Z., Jia, L., Gartner, J. J., Hill, V. K., Copeland, A. R., Nah, S. K., Masi, R. V., Gasmi, B., Kivitz, S., Paria, B. C., Florentin, M., Kim, S. P., Hanada, K.-I., Li, Y. F., Ngo, L. T., Ray, S., Shindorf, M. L., Levi, S. T., Shepherd, R., Toy, C., Parikh, A. Y., Prickett, T. D., Kelly, M. C., Beyer, R., Goff, S. L., Yang, J. C., Robbins, P. F. & Rosenberg, S. A. Molecular signatures of antitumor neoantigen-reactive T cells from metastatic human cancers. *Science* **375**, 877–884 (2022).
18. Caushi, J. X., Zhang, J., Ji, Z., Vaghasia, A., Zhang, B., Hsiue, E. H.-C., Mog, B. J., Hou, W., Justesen, S., Blosser, R., Tam, A., Anagnostou, V., Cottrell, T. R., Guo, H., Chan, H. Y., Singh, D., Thapa, S., Dykema, A. G., Burman, P., Choudhury, B., Aparicio, L., Cheung, L. S., Lanis, M., Belcaid, Z., El Asmar, M., Illei, P. B., Wang, R., Meyers, J., Schuebel, K., Gupta, A., Skaist, A., Wheelan, S., Naidoo, J., Marrone, K. A., Brock, M., Ha, J., Bush, E. L., Park, B. J., Bott, M., Jones, D. R., Reuss, J. E., Velculescu, V. E., Chaft, J. E., Kinzler, K. W., Zhou, S., Vogelstein, B., Taube, J. M., Hellmann, M. D., Brahmer, J. R., Merghoub, T., Forde, P. M., Yegnasubramanian, S., Ji, H., Pardoll, D. M. & Smith, K. N. Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. *Nature* **596**, 126–132 (2021).
 19. Zheng, C., Fass, J. N., Shih, Y.-P., Gunderson, A. J., Sanjuan Silva, N., Huang, H., Bernard, B. M., Rajamanickam, V., Slagel, J., Bifulco, C. B., Piening, B., Newell, P. H. A., Hansen, P. D. & Tran, E. Transcriptomic profiles of neoantigen-reactive T cells in human gastrointestinal cancers. *Cancer Cell* **40**, 410–423.e7 (2022).
 20. Oliveira, G., Stromhaug, K., Klaeger, S., Kula, T., Frederick, D. T., Le, P. M., Forman, J., Huang, T., Li, S., Zhang, W., Xu, Q., Cieri, N., Clauser, K. R., Shukla, S. A., Neuberger, D., Justesen, S., MacBeath, G., Carr, S. A., Fritsch, E. F., Hacohen, N., Sade-Feldman, M., Livak, K. J., Boland, G. M., Ott, P. A., Keskin, D. B. & Wu, C. J. Phenotype, specificity and avidity of antitumour CD8 T cells in melanoma. *Nature* **596**, 119–125 (2021).
 21. Khodadoust, M. S., Olsson, N., Wagar, L. E., Haabeth, O. A. W., Chen, B., Swaminathan, K., Rawson, K., Liu, C. L., Steiner, D., Lund, P., Rao, S., Zhang, L., Marceau, C., Stehr, H., Newman, A. M., Czerwinski, D. K., Carlton, V. E. H., Moorhead, M., Faham, M., Kohrt, H. E., Carette, J., Green, M. R., Davis, M. M., Levy, R., Elias, J. E. & Alizadeh, A. A. Antigen presentation profiling reveals recognition of lymphoma immunoglobulin neoantigens. *Nature* **543**, 723–727 (2017).
 22. Cohen, C. J., Gartner, J. J., Horovitz-Fried, M., Shamalov, K., Trebska-McGowan, K., Bliskovsky, V. V., Parkhurst, M. R., Ankri, C., Prickett, T. D., Crystal, J. S., Li, Y. F., El-Gamil, M., Rosenberg, S. A. & Robbins, P. F. Isolation of neoantigen-specific T cells from tumor and peripheral lymphocytes. *J. Clin. Invest.* **125**, 3981–3991 (2015).
 23. McGranahan, N., Furness, A. J. S., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S. K., Jamal-Hanjani, M., Wilson, G. A., Birkbak, N. J., Hiley, C. T., Watkins, T. B. K., Shafi, S., Murugaesu, N., Mitter, R., Akarca, A. U., Linares, J., Marafioti, T., Henry, J. Y., Van Allen, E. M., Miao, D., Schilling, B., Schadendorf, D., Garraway, L. A., Makarov, V., Rizvi, N. A., Snyder, A., Hellmann, M. D., Merghoub, T., Wolchok, J. D., Shukla, S. A., Wu, C. J., Peggs, K. S., Chan, T. A., Hadrup, S. R., Quezada, S. A. & Swanton, C. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469 (2016).
 24. Simoni, Y., Becht, E., Fehlings, M., Loh, C. Y., Koo, S.-L., Teng, K. W. W., Yeong, J. P. S., Nahar, R., Zhang, T., Kared, H., Duan, K., Ang, N., Poidinger, M., Lee, Y. Y., Larbi, A., Khng, A. J., Tan, E., Fu, C., Mathew, R., Teo, M., Lim, W. T., Toh, C. K., Ong, B.-H., Koh, T., Hillmer, A. M., Takano, A., Lim, T. K. H., Tan, E. H., Zhai, W., Tan, D. S. W., Tan, I. B. & Newell, E. W. Bystander CD8 T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* **557**, 575–579 (2018).
 25. Meier, S. L., Satpathy, A. T. & Wells, D. K. Bystander T cells in cancer immunology and therapy. *Nat Cancer* **3**, 143–155 (2022).
 26. Liu, B., Zhang, Y., Wang, D., Hu, X. & Zhang, Z. Single-cell meta-analyses reveal responses of tumor-reactive CXCL13 T cells to immune-checkpoint blockade. *Nat Cancer* **3**, 1123–1136 (2022).
 27. Zhang, M., Fritsche, J., Roszik, J., Williams, L. J., Peng, X., Chiu, Y., Tsou, C.-C., Hoffgaard, F., Goldfinger, V., Schoor, O., Talukder, A., Forget, M. A., Haymaker, C., Bernatchez, C., Han, L., Tsang, Y.-H., Kong, K., Xu, X., Scott, K. L., Singh-Jasuja, H., Lizee, G., Liang, H., Weinschenk, T., Mills, G. B. & Hwu, P. RNA editing derived epitopes function as cancer antigens to elicit immune responses. *Nat. Commun.* **9**, 3919 (2018).
 28. Peng, X., Xu, X., Wang, Y., Hawke, D. H., Yu, S., Han, L., Zhou, Z., Mojumdar, K., Jeong, K. J., Labrie, M., Tsang, Y. H., Zhang, M., Lu, Y., Hwu, P., Scott, K. L., Liang, H. & Mills, G. B. A-to-I RNA Editing Contributes to Proteomic Diversity in Cancer. *Cancer Cell* **33**, 817–828.e7 (2018).
 29. Marijt, K. A. & van Hall, T. To TAP or not to TAP: alternative peptides for immunotherapy of cancer. *Curr. Opin. Immunol.* **64**, 15–19 (2020).

30. Ebstein, F., Textoris-Taube, K., Keller, C., Golnik, R., Vigneron, N., Van den Eynde, B. J., Schuler-Thurner, B., Schadendorf, D., Lorenz, F. K. M., Uckert, W., Urban, S., Lehmann, A., Albrecht-Koepke, N., Janek, K., Henklein, P., Niewianda, A., Kloetzel, P. M. & Mishto, M. Proteasomes generate spliced epitopes by two different mechanisms and as efficiently as non-spliced epitopes. *Sci. Rep.* **6**, 24032 (2016).
31. Matsushita, H., Vesely, M. D., Koboldt, D. C., Rickert, C. G., Uppaluri, R., Magrini, V. J., Arthur, C. D., White, J. M., Chen, Y.-S., Shea, L. K., Hundal, J., Wendl, M. C., Demeter, R., Wylie, T., Allison, J. P., Smyth, M. J., Old, L. J., Mardis, E. R. & Schreiber, R. D. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoeediting. *Nature* **482**, 400–404 (2012).
32. André, S., Picard, M., Cezar, R., Roux-Dalvai, F., Alleaume-Butaux, A., Soundaramourty, C., Cruz, A. S., Mendes-Frias, A., Gotti, C., Leclercq, M., Nicolas, A., Tauzin, A., Carvalho, A., Capela, C., Pedrosa, J., Castro, A. G., Kundura, L., Loubet, P., Sotto, A., Muller, L., Lefrant, J.-Y., Roger, C., Claret, P.-G., Duvnjak, S., Tran, T.-A., Racine, G., Zghidi-Abouzid, O., Nioche, P., Silvestre, R., Droit, A., Mammano, F., Corbeau, P. & Estaquier, J. T cell apoptosis characterizes severe Covid-19 disease. *Cell Death Differ.* **29**, 1486–1499 (2022).
33. Yao, C., Bora, S. A., Parimon, T., Zaman, T., Friedman, O. A., Palatinus, J. A., Surapaneni, N. S., Matusov, Y. P., Cerro Chiang, G., Kassar, A. G., Patel, N., Green, C. E. R., Aziz, A. W., Suri, H., Suda, J., Lopez, A. A., Martins, G. A., Stripp, B. R., Gharib, S. A., Goodridge, H. S. & Chen, P. Cell-Type-Specific Immune Dysregulation in Severely Ill COVID-19 Patients. *Cell Rep.* **34**, 108590 (2021).
34. Schultheiß, C., Paschold, L., Simnica, D., Mohme, M., Willscher, E., von Wenserski, L., Scholz, R., Wieters, I., Dahlke, C., Tolosa, E., Sedding, D. G., Ciesek, S., Addo, M. & Binder, M. Next-Generation Sequencing of T and B Cell Receptor Repertoires from COVID-19 Patients Showed Signatures Associated with Severity of Disease. *Immunity* **53**, 442–455.e4 (2020).
35. Ravichandran, S., Lee, Y., Grubbs, G., Coyle, E. M., Klenow, L., Akasaka, O., Koga, M., Adachi, E., Saito, M., Nakachi, I., Ogura, T., Baba, R., Ito, M., Kiso, M., Yasuhara, A., Yamada, S., Sakai-Tagawa, Y., Iwatsuki-Horimoto, K., Imai, M., Yamayoshi, S., Yotsuyanagi, H., Kawaoka, Y. & Khurana, S. Longitudinal antibody repertoire in ‘mild’ versus ‘severe’ COVID-19 patients reveals immune markers associated with disease severity and resolution. *Sci Adv* **7**, (2021).
36. Zhang, J.-Y., Wang, X.-M., Xing, X., Xu, Z., Zhang, C., Song, J.-W., Fan, X., Xia, P., Fu, J.-L., Wang, S.-Y., Xu, R.-N., Dai, X.-P., Shi, L., Huang, L., Jiang, T.-J., Shi, M., Zhang, Y., Zumla, A., Maeurer, M., Bai, F. & Wang, F.-S. Single-cell landscape of immunological responses in patients with COVID-19. *Nat. Immunol.* **21**, 1107–1118 (2020).
37. Vermeulen, C., Pagès-Gallego, M., Kester, L., Kranendonk, M. E. G., Wesseling, P., Verburg, N., de Witt Hamer, P., Kooi, E. J., Dankmeijer, L., van der Lugt, J., van Baarsen, K., Hoving, E. W., Tops, B. B. J. & de Ridder, J. Ultra-fast deep-learned CNS tumour classification during surgery. *Nature* **622**, 842–849 (2023).
38. Arvaniti, E. & Claassen, M. Sensitive detection of rare disease-associated cell subsets via representation learning. *Nat. Commun.* **8**, 14825 (2017).
39. Zhu, L., Yang, P., Zhao, Y., Zhuang, Z., Wang, Z., Song, R., Zhang, J., Liu, C., Gao, Q., Xu, Q., Wei, X., Sun, H.-X., Ye, B., Wu, Y., Zhang, N., Lei, G., Yu, L., Yan, J., Diao, G., Meng, F., Bai, C., Mao, P., Yu, Y., Wang, M., Yuan, Y., Deng, Q., Li, Z., Huang, Y., Hu, G., Liu, Y., Wang, X., Xu, Z., Liu, P., Bi, Y., Shi, Y., Zhang, S., Chen, Z., Wang, J., Xu, X., Wu, G., Wang, F.-S., Gao, G. F., Liu, L. & Liu, W. J. Single-Cell Sequencing of Peripheral Mononuclear Cells Reveals Distinct Immune Response Landscapes of COVID-19 and Influenza Patients. *Immunity* **53**, 685–696.e3 (2020).
40. Kramer, K. J., Wilfong, E. M., Voss, K., Barone, S. M., Shiakolas, A. R., Raju, N., Roe, C. E., Suryadevara, N., Walker, L. M., Wall, S. C., Paulo, A., Schaefer, S., Dahunsi, D., Westlake, C. S., Crowe, J. E., Jr, Carnahan, R. H., Rathmell, J. C., Bonami, R. H., Georgiev, I. S. & Irish, J. M. Single-cell profiling of the antigen-specific response to BNT162b2 SARS-CoV-2 RNA vaccine. *Nat. Commun.* **13**, 3466 (2022).
41. Sureshchandra, S., Lewis, S. A., Doratt, B. M., Jankeel, A., Coimbra Ibraim, I. & Messaoudi, I. Single-cell profiling of T and B cell repertoires following SARS-CoV-2 mRNA vaccine. *JCI Insight* **6**, (2021).
42. Zhang, B., Upadhyay, R., Hao, Y., Samanovic, M. I., Herati, R. S., Blair, J. D., Axelrad, J., Mulligan, M. J., Littman, D. R. & Satija, R. Multimodal single-cell datasets characterize antigen-specific CD8 T cells across SARS-CoV-2 vaccination and infection. *Nat. Immunol.* **24**, 1725–1734 (2023).
43. Waickman, A. T., Victor, K., Li, T., Hatch, K., Rutvisuttinunt, W., Medin, C., Gabriel, B., Jarman, R. G., Friberg, H. & Currier, J. R. Dissecting the heterogeneity of DENV vaccine-elicited cellular immunity using single-cell RNA sequencing and metabolic profiling. *Nat. Commun.* **10**, 3666 (2019).
44. Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., Clevers, H., Deplancke, B., Dunham, I., Eberwine, J., Eils, R., Enard, W., Farmer, A., Fugger, L., Göttgens, B., Hacohen, N., Haniffa, M., Hemberg, M., Kim, S., Klenerman, P., Kriegstein, A.,

- Lein, E., Linnarsson, S., Lundberg, E., Lundeberg, J., Majumder, P., Marioni, J. C., Merad, M., Mhlanga, M., Nawijn, M., Netea, M., Nolan, G., Pe'er, D., Phillipakis, A., Ponting, C. P., Quake, S., Reik, W., Rozenblatt-Rosen, O., Sanes, J., Satija, R., Schumacher, T. N., Shalek, A., Shapiro, E., Sharma, P., Shin, J. W., Stegle, O., Stratton, M., Stubbington, M. J. T., Theis, F. J., Uhlen, M., van Oudenaarden, A., Wagner, A., Watt, F., Weissman, J., Wold, B., Xavier, R., Yosef, N. & Human Cell Atlas Meeting Participants. The Human Cell Atlas. *Elife* **6**, (2017).
45. Suo, C., Dann, E., Goh, I., Jardine, L., Kleshchevnikov, V., Park, J.-E., Botting, R. A., Stephenson, E., Engelbert, J., Tuong, Z. K., Polanski, K., Yayon, N., Xu, C., Suchanek, O., Elmentaite, R., Domínguez Conde, C., He, P., Pritchard, S., Miah, M., Moldovan, C., Steemers, A. S., Mazin, P., Prete, M., Horsfall, D., Marioni, J. C., Clatworthy, M. R., Haniffa, M. & Teichmann, S. A. Mapping the developing human immune system across organs. *Science* **376**, eabo0510 (2022).
 46. Cui, A., Huang, T., Li, S., Ma, A., Pérez, J. L., Sander, C., Keskin, D. B., Wu, C. J., Fraenkel, E. & Hacohen, N. Dictionary of immune responses to cytokines at single-cell resolution. *Nature* (2023). doi:10.1038/s41586-023-06816-9
 47. Rohaan, M. W., Borch, T. H., van den Berg, J. H., Met, Ö., Kessels, R., Geukes Foppen, M. H., Stoltenberg Granhøj, J., Nuijen, B., Nijenhuis, C., Jedema, I., van Zon, M., Scheij, S., Beijnen, J. H., Hansen, M., Voermans, C., Noringriis, I. M., Monberg, T. J., Holmstroem, R. B., Wever, L. D. V., van Dijk, M., Grijpink-Ongering, L. G., Valkenet, L. H. M., Torres Acosta, A., Karger, M., Borgers, J. S. W., Ten Ham, R. M. T., Retèl, V. P., van Harten, W. H., Lalezari, F., van Tinteren, H., van der Veldt, A. A. M., Hospers, G. A. P., Stevensen-den Boer, M. A. M., Suijkerbuijk, K. P. M., Aarts, M. J. B., Piersma, D., van den Eertwegh, A. J. M., de Groot, J.-W. B., Vreugdenhil, G., Kapiteijn, E., Boers-Sonderen, M. J., Fiets, W. E., van den Berkmortel, F. W. P. J., Ellebaek, E., Hölmich, L. R., van Akkooi, A. C. J., van Houdt, W. J., Wouters, M. W. J. M., van Thienen, J. V., Blank, C. U., Meerveld-Eggink, A., Klobuch, S., Wilgenhof, S., Schumacher, T. N., Donia, M., Svane, I. M. & Haanen, J. B. A. G. Tumor-Infiltrating Lymphocyte Therapy or Ipilimumab in Advanced Melanoma. *N. Engl. J. Med.* **387**, 2113–2125 (2022).
 48. Andersen, R., Donia, M., Ellebaek, E., Borch, T. H., Kongsted, P., Iversen, T. Z., Hölmich, L. R., Hendel, H. W., Met, Ö., Andersen, M. H., Thor Straten, P. & Svane, I. M. Long-Lasting Complete Responses in Patients with Metastatic Melanoma after Adoptive Cell Therapy with Tumor-Infiltrating Lymphocytes and an Attenuated IL2 Regimen. *Clin. Cancer Res.* **22**, 3734–3745 (2016).
 49. Reuben, A., Zhang, J., Chiou, S.-H., Gittelman, R. M., Li, J., Lee, W.-C., Fujimoto, J., Behrens, C., Liu, X., Wang, F., Quek, K., Wang, C., Kheradmand, F., Chen, R., Chow, C.-W., Lin, H., Bernatchez, C., Jalali, A., Hu, X., Wu, C.-J., Eterovic, A. K., Parra, E. R., Yusko, E., Emerson, R., Benzeno, S., Vignali, M., Wu, X., Ye, Y., Little, L. D., Gumbs, C., Mao, X., Song, X., Tippen, S., Thornton, R. L., Cascone, T., Snyder, A., Wargo, J. A., Herbst, R., Swisher, S., Kadara, H., Moran, C., Kalhor, N., Zhang, J., Scheet, P., Vaporciyan, A. A., Sepesi, B., Gibbons, D. L., Robins, H., Hwu, P., Heymach, J. V., Sharma, P., Allison, J. P., Baladandayuthapani, V., Lee, J. J., Davis, M. M., Wistuba, I. I., Futreal, P. A. & Zhang, J. Comprehensive T cell repertoire characterization of non-small cell lung cancer. *Nat. Commun.* **11**, 603 (2020).
 50. Tumei, P. C., Harview, C. L., Yearley, J. H., Shintaku, I. P., Taylor, E. J. M., Robert, L., Chmielowski, B., Spasic, M., Henry, G., Ciobanu, V., West, A. N., Carmona, M., Kivork, C., Seja, E., Cherry, G., Gutierrez, A. J., Grogan, T. R., Mateus, C., Tomasic, G., Glaspy, J. A., Emerson, R. O., Robins, H., Pierce, R. H., Elashoff, D. A., Robert, C. & Ribas, A. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014).
 51. Yeong, J., Suteja, L., Simoni, Y., Lau, K. W., Tan, A. C., Li, H. H., Lim, S., Loh, J. H., Wee, F. Y. T., Nerurkar, S. N., Takano, A., Tan, E. H., Lim, T. K. H., Newell, E. W. & Tan, D. S. W. Intratumoral CD39CD8 T Cells Predict Response to Programmed Cell Death Protein-1 or Programmed Death Ligand-1 Blockade in Patients With NSCLC. *J. Thorac. Oncol.* **16**, 1349–1358 (2021).
 52. Andreatta, M., Corria-Osorio, J., Müller, S., Cubas, R., Coukos, G. & Carmona, S. J. Interpretation of T cell states from single-cell transcriptomics data using reference atlases. *Nat. Commun.* **12**, 2965 (2021).
 53. Dezfulian, M. H., Kula, T., Pranzatelli, T., Kamitaki, N., Meng, Q., Khatri, B., Perez, P., Xu, Q., Chang, A., Kohlgruber, A. C., Leng, Y., Jupudi, A. A., Joachims, M. L., Chiorini, J. A., Lessard, C. J., Darise Farris, A., Muthuswamy, S. K., Warner, B. M. & Elledge, S. J. TScan-II: A genome-scale platform for the de novo identification of CD4 T cell epitopes. *Cell* **186**, 5569–5586.e21 (2023).
 54. Kula, T., Dezfulian, M. H., Wang, C. I., Abdelfattah, N. S., Hartman, Z. C., Wucherpfennig, K. W., Lyerly, H. K. & Elledge, S. J. T-Scan: A Genome-wide Method for the Systematic Discovery of T Cell Epitopes. *Cell* **178**, 1016–1028.e13 (2019).
 55. Joglekar, A. V. & Li, G. T cell antigen discovery. *Nat. Methods* **18**, 873–880 (2021).

56. Dash, P., Fiore-Gartland, A. J., Hertz, T., Wang, G. C., Sharma, S., Souquette, A., Crawford, J. C., Clemens, E. B., Nguyen, T. H. O., Kedzierska, K., La Gruta, N. L., Bradley, P. & Thomas, P. G. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature* **547**, 89–93 (2017).
57. Huang, H., Wang, C., Rubelt, F., Scriba, T. J. & Davis, M. M. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat. Biotechnol.* **38**, 1194–1202 (2020).
58. Textor, J., Buytenhuijs, F., Rogers, D., Gauthier, È. M., Sultan, S., Wortel, I. M. N., Kalies, K., Fähnrich, A., Pagel, R., Melichar, H. J., Westermann, J. & Mandl, J. N. Machine learning analysis of the T cell receptor repertoire identifies sequence features of self-reactivity. *Cell Syst* (2023). doi:10.1016/j.cels.2023.11.004
59. Goncharov, M., Bagaev, D., Shcherbinin, D., Zvyagin, I., Bolotin, D., Thomas, P. G., Minervina, A. A., Pogorelyy, M. V., Ladell, K., McLaren, J. E., Price, D. A., Nguyen, T. H. O., Rowntree, L. C., Clemens, E. B., Kedzierska, K., Dolton, G., Rius, C. R., Sewell, A., Samir, J., Luciani, F., Zornikova, K. V., Khmelevskaya, A. A., Sheetikov, S. A., Efimov, G. A., Chudakov, D. & Shugay, M. VDJdb in the pandemic era: a compendium of T cell receptors specific for SARS-CoV-2. *Nat. Methods* **19**, 1017–1019 (2022).
60. Holm, J. S., Funt, S. A., Borch, A., Munk, K. K., Bjerregaard, A.-M., Reading, J. L., Maher, C., Regazzi, A., Wong, P., Al-Ahmadie, H., Iyer, G., Tamhane, T., Bentzen, A. K., Herschend, N. O., De Wolf, S., Snyder, A., Merghoub, T., Wolchok, J. D., Nielsen, M., Rosenberg, J. E., Bajorin, D. F. & Hadrup, S. R. Neoantigen-specific CD8 T cell responses in the peripheral blood following PD-L1 blockade might predict therapy outcome in metastatic urothelial carcinoma. *Nat. Commun.* **13**, 1935 (2022).
61. Au, L., Hatipoglu, E., Robert de Massy, M., Litchfield, K., Beattie, G., Rowan, A., Schnidrig, D., Thompson, R., Byrne, F., Horswell, S., Fotiadis, N., Hazell, S., Nicol, D., Shepherd, S. T. C., Fendler, A., Mason, R., Del Rosario, L., Edmonds, K., Lingard, K., Sarker, S., Mangwende, M., Carlyle, E., Attig, J., Joshi, K., Uddin, I., Becker, P. D., Sunderland, M. W., Akarca, A., Puccio, I., Yang, W. W., Lund, T., Dhillon, K., Vasquez, M. D., Ghorani, E., Xu, H., Spencer, C., López, J. I., Green, A., Mahadeva, U., Borg, E., Mitchison, M., Moore, D. A., Proctor, I., Falzon, M., Pickering, L., Furness, A. J. S., Reading, J. L., Salgado, R., Marafioti, T., Jamal-Hanjani, M., PEACE Consortium, Kassiotis, G., Chain, B., Larkin, J., Swanton, C., Quezada, S. A., Turajlic, S. & TRACERx Renal Consortium. Determinants of anti-PD-1 response and resistance in clear cell renal cell carcinoma. *Cancer Cell* **39**, 1497–1518.e11 (2021).
62. Doebele, R. C., Drlon, A., Paz-Ares, L., Siena, S., Shaw, A. T., Farago, A. F., Blakely, C. M., Seto, T., Cho, B. C., Tosi, D., Besse, B., Chawla, S. P., Bazhenova, L., Krauss, J. C., Chae, Y. K., Barve, M., Garrido-Laguna, I., Liu, S. V., Conkling, P., John, T., Fakih, M., Sigal, D., Loong, H. H., Buchsacher, G. L., Jr, Garrido, P., Nieva, J., Steuer, C., Overbeck, T. R., Bowles, D. W., Fox, E., Riehl, T., Chow-Maneval, E., Simmons, B., Cui, N., Johnson, A., Eng, S., Wilson, T. R., Demetri, G. D. & trial investigators. Entrectinib in patients with advanced or metastatic NTRK fusion-positive solid tumours: integrated analysis of three phase 1-2 trials. *Lancet Oncol.* **21**, 271–282 (2020).
63. Schober, K., Voit, F., Grassmann, S., Müller, T. R., Eggert, J., Jarosch, S., Weißbrich, B., Hoffmann, P., Borkner, L., Nio, E., Fanchi, L., Clouser, C. R., Radhakrishnan, A., Mihatsch, L., Lückemeier, P., Leube, J., Dössinger, G., Klein, L., Neuenhahn, M., Oduro, J. D., Cicin-Sain, L., Buchholz, V. R. & Busch, D. H. Reverse TCR repertoire evolution toward dominant low-affinity clones during chronic CMV infection. *Nat. Immunol.* **21**, 434–441 (2020).
64. Martinez, R. J. & Evavold, B. D. Lower Affinity T Cells are Critical Components and Active Participants of the Immune Response. *Front. Immunol.* **6**, 468 (2015).
65. Rosato, P. C., Wijeyesinghe, S., Stolley, J. M., Nelson, C. E., Davis, R. L., Manlove, L. S., Pennell, C. A., Blazar, B. R., Chen, C. C., Geller, M. A., Vezys, V. & Masopust, D. Virus-specific memory T cells populate tumors and can be repurposed for tumor immunotherapy. *Nat. Commun.* **10**, 567 (2019).
66. Nathan, P., Hassel, J. C., Rutkowski, P., Baurain, J.-F., Butler, M. O., Schlaak, M., Sullivan, R. J., Ochsenreither, S., Dummer, R., Kirkwood, J. M., Joshua, A. M., Sacco, J. J., Shoushtari, A. N., Orloff, M., Piulats, J. M., Milhem, M., Salama, A. K. S., Curti, B., Demidov, L., Gastaud, L., Mauch, C., Yushak, M., Carvajal, R. D., Hamid, O., Abdullah, S. E., Holland, C., Goodall, H., Piperno-Neumann, S. & IMCgp100-202 Investigators. Overall Survival Benefit with Tebentafusp in Metastatic Uveal Melanoma. *N. Engl. J. Med.* **385**, 1196–1206 (2021).
67. Dolton, G., Rius, C., Wall, A., Szomolay, B., Bianchi, V., Galloway, S. A. E., Hasan, M. S., Morin, T., Caillaud, M. E., Thomas, H. L., Theaker, S., Tan, L. R., Fuller, A., Topley, K., Legut, M., Attaf, M., Hopkins, J. R., Behiry, E., Zabkiewicz, J., Alvares, C., Lloyd, A., Rogers, A., Henley, P., Fegan, C., Ottmann, O., Man, S., Crowther, M. D., Donia, M., Svane, I. M., Cole, D. K., Brown, P. E., Rizkallah, P. & Sewell, A. K. Targeting of multiple tumor-associated antigens by individual T cell receptors during successful cancer immunotherapy. *Cell*

- 186**, 3333–3349.e27 (2023).
68. Foy, S. P., Jacoby, K., Bota, D. A., Hunter, T., Pan, Z., Stawiski, E., Ma, Y., Lu, W., Peng, S., Wang, C. L., Yuen, B., Dalmas, O., Heeringa, K., Sennino, B., Conroy, A., Bethune, M. T., Mende, I., White, W., Kukreja, M., Gunturu, S., Humphrey, E., Hussaini, A., An, D., Litterman, A. J., Quach, B. B., Ng, A. H. C., Lu, Y., Smith, C., Campbell, K. M., Anaya, D., Skrdlant, L., Huang, E. Y.-H., Mendoza, V., Mathur, J., Dengler, L., Purandare, B., Moot, R., Yi, M. C., Funke, R., Sibley, A., Stallings-Schmitt, T., Oh, D. Y., Chmielowski, B., Abedi, M., Yuan, Y., Sosman, J. A., Lee, S. M., Schoenfeld, A. J., Baltimore, D., Heath, J. R., Franzusoff, A., Ribas, A., Rao, A. V. & Mandl, S. J. Non-viral precision T cell receptor replacement for personalized cell therapy. *Nature* **615**, 687–696 (2023).
 69. Klebanoff, C. A., Chandran, S. S., Baker, B. M., Quezada, S. A. & Ribas, A. T cell receptor therapeutics: immunological targeting of the intracellular cancer proteome. *Nat. Rev. Drug Discov.* **22**, 996–1017 (2023).
 70. Benjamin, R., Jain, N., Maus, M. V., Boissel, N., Graham, C., Jozwik, A., Yallop, D., Konopleva, M., Frigault, M. J., Teshima, T., Kato, K., Boucaud, F., Balandraud, S., Gianella-Borradori, A., Binlich, F., Marchiq, I., Dupouy, S., Almendra-Carrasco, M., Pannaux, M., Fouliard, S., Brissot, E., Mohty, M. & CALM Study Group. UCART19, a first-in-class allogeneic anti-CD19 chimeric antigen receptor T-cell therapy for adults with relapsed or refractory B-cell acute lymphoblastic leukaemia (CALM): a phase 1, dose-escalation trial. *Lancet Haematol* **9**, e833–e843 (2022).
 71. Polesie, S., Gillstedt, M., Schmidt, S. A. J., Egeberg, A., Pottegård, A. & Kristensen, K. Use of methotrexate and risk of skin cancer: a nationwide case-control study. *Br. J. Cancer* **128**, 1311–1319 (2023).
 72. Mackensen, A., Müller, F., Mougiakakos, D., Böltz, S., Wilhelm, A., Aigner, M., Völkl, S., Simon, D., Kleyer, A., Munoz, L., Kretschmann, S., Kharboutli, S., Gary, R., Reimann, H., Rösler, W., Uderhardt, S., Bang, H., Herrmann, M., Ekici, A. B., Buettner, C., Habenicht, K. M., Winkler, T. H., Krönke, G. & Schett, G. Anti-CD19 CAR T cell therapy for refractory systemic lupus erythematosus. *Nat. Med.* **28**, 2124–2132 (2022).
 73. Britanova, O. V., Lupyry, K. R., Staroverov, D. B., Shagina, I. A., Aleksandrov, A. A., Ustyugov, Y. Y., Somov, D. V., Klimentko, A., Shostak, N. A., Zvyagin, I. V., Stepanov, A. V., Merzlyak, E. M., Davydov, A. N., Izraelson, M., Egorov, E. S., Bogdanova, E. A., Vladimirova, A. K., Iakovlev, P. A., Fedorenko, D. A., Ivanov, R. A., Skvortsova, V. I., Lukyanov, S. & Chudakov, D. M. Targeted depletion of TRBV9 T cells as immunotherapy in a patient with ankylosing spondylitis. *Nat. Med.* **29**, 2731–2736 (2023).
 74. Bittner, S., Hehlhans, T. & Feuerer, M. Engineered Treg cells as putative therapeutics against inflammatory diseases and beyond. *Trends Immunol.* **44**, 468–483 (2023).
 75. Yang, S. J., Singh, A. K., Drow, T., Tappen, T., Honaker, Y., Barahmand-Pour-Whitman, F., Linsley, P. S., Cerosaletti, K., Mauk, K., Xiang, Y., Smith, J., Mortensen, E., Cook, P. J., Sommer, K., Khan, I., Liggitt, D., Rawlings, D. J. & Buckner, J. H. Pancreatic islet-specific engineered T exhibit robust antigen-specific and bystander immune suppression in type 1 diabetes models. *Sci. Transl. Med.* **14**, eabn1716 (2022).

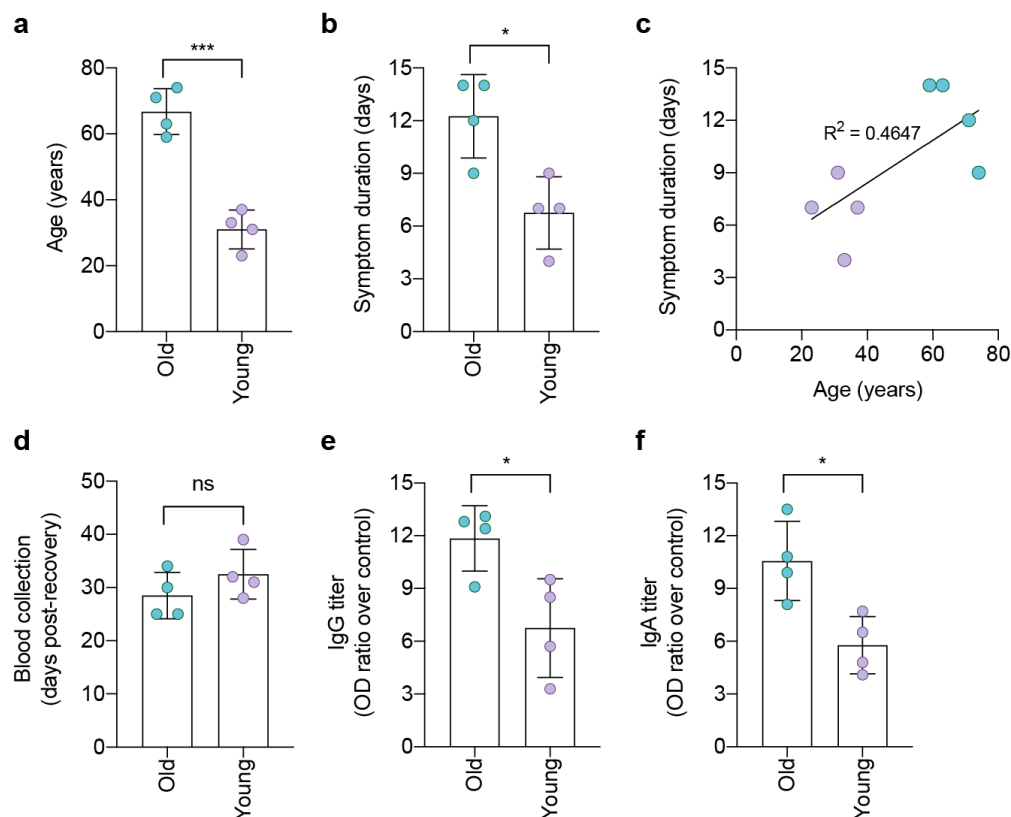
Appendix

A.1 Supplementary information of Chapter 2

Due to the extensive length of some supplementary material, not the entire supplementary material is presented here. It can be found in the online version of the published article.

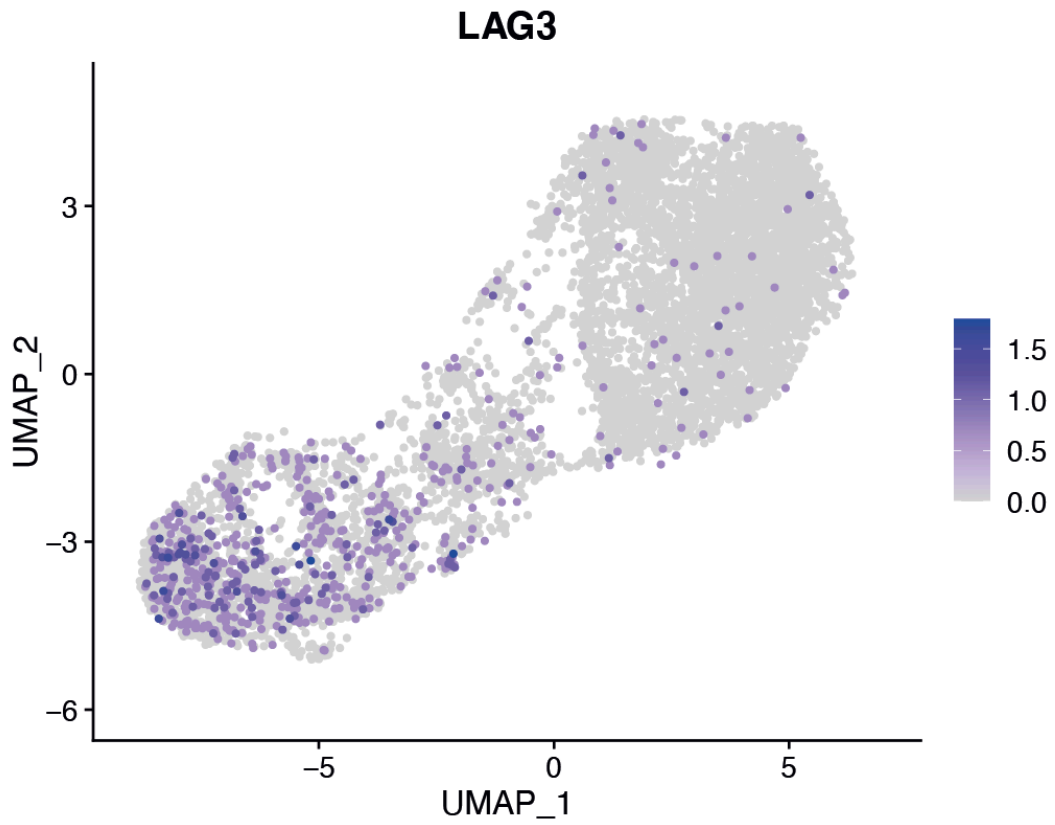
Supplemental Figures 1-7

Supplemental Tables 1-5

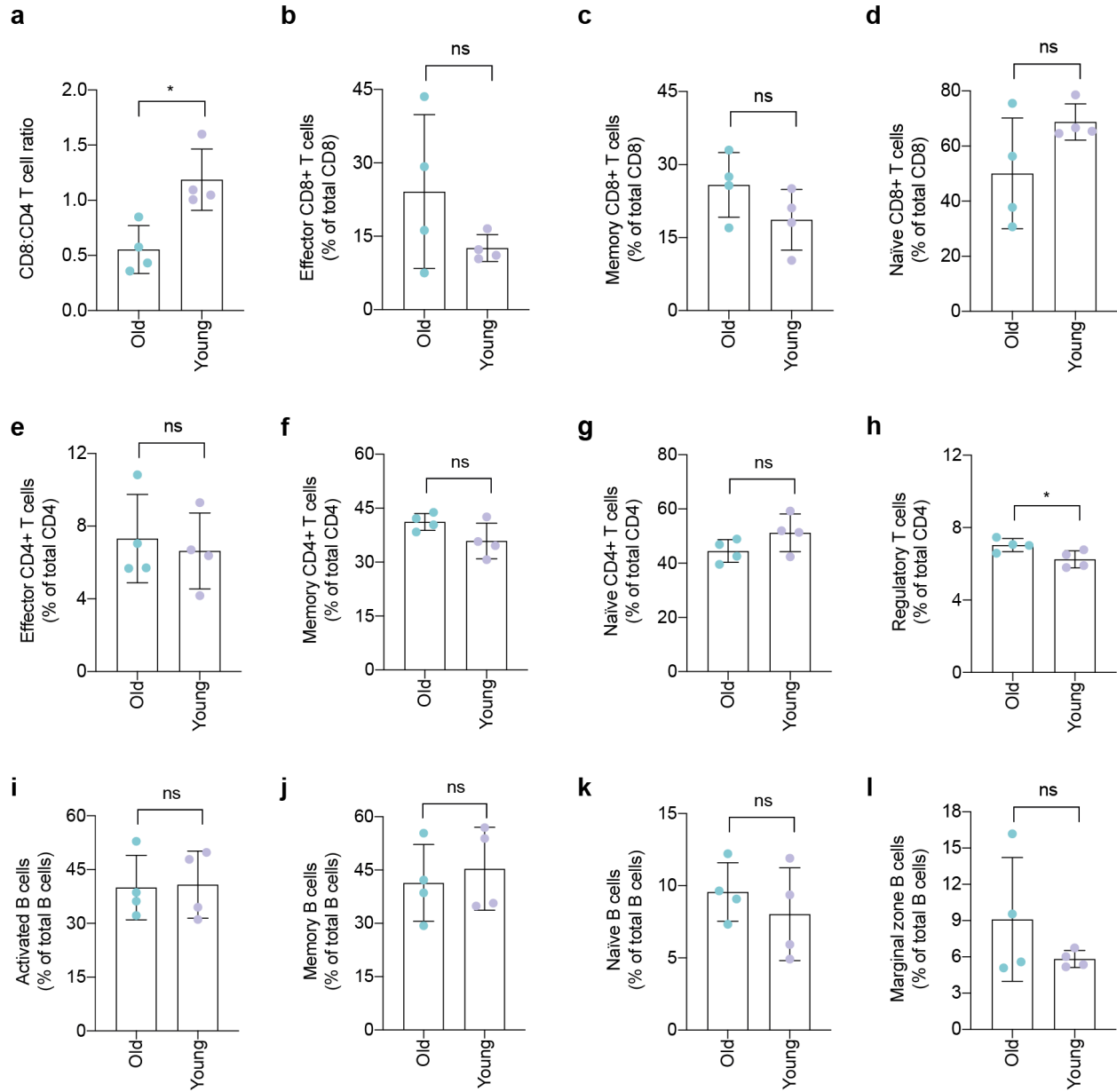


Supplemental Figure 1. Clinical and serological characteristics of convalescent COVID-19 patients.

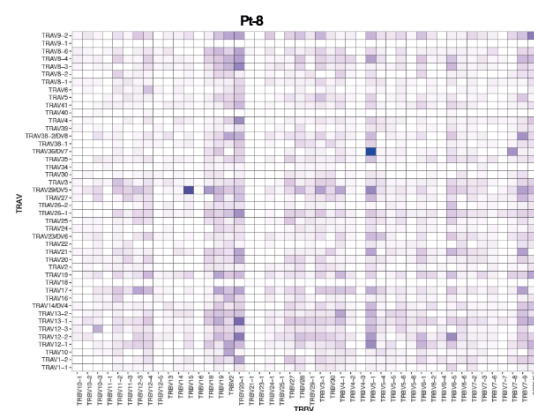
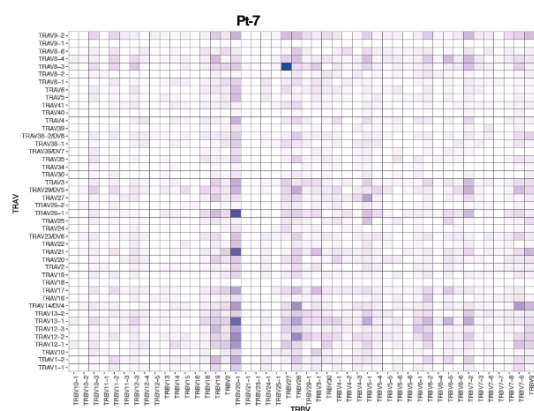
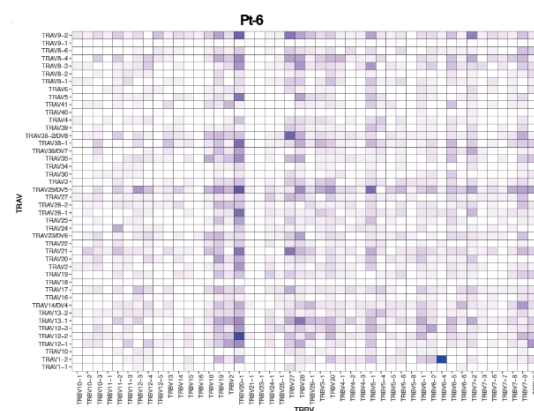
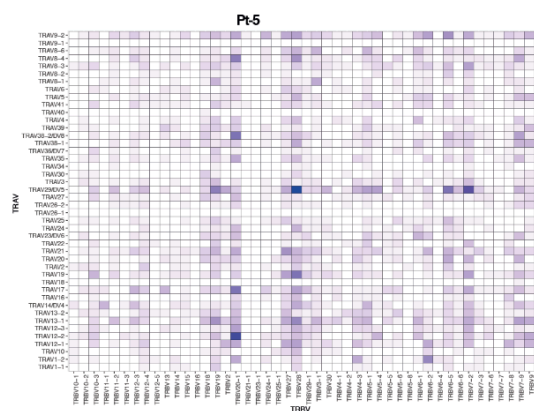
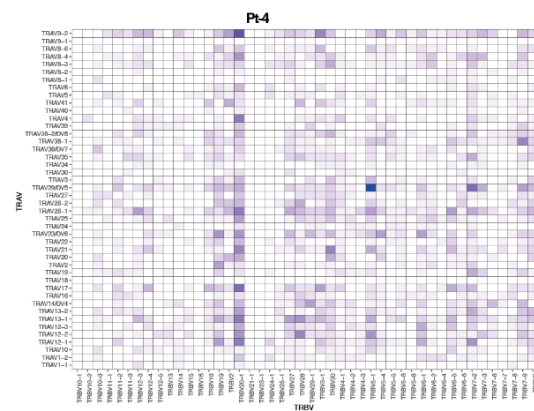
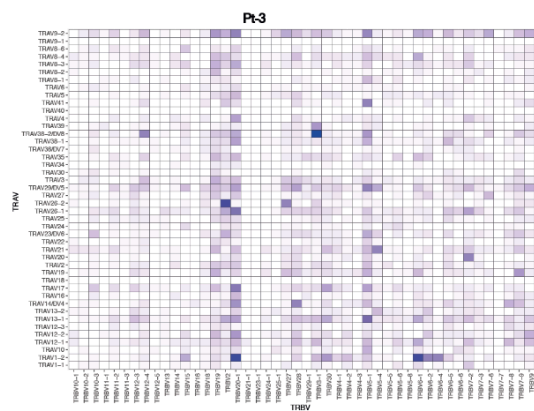
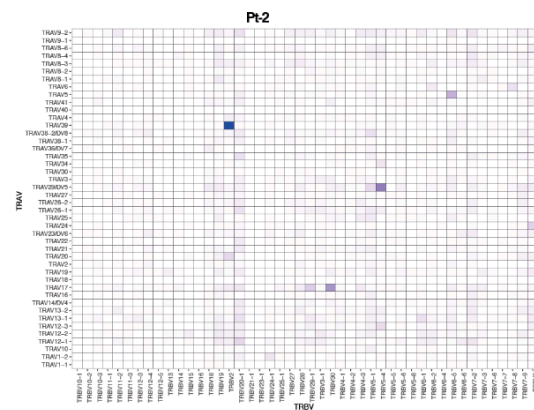
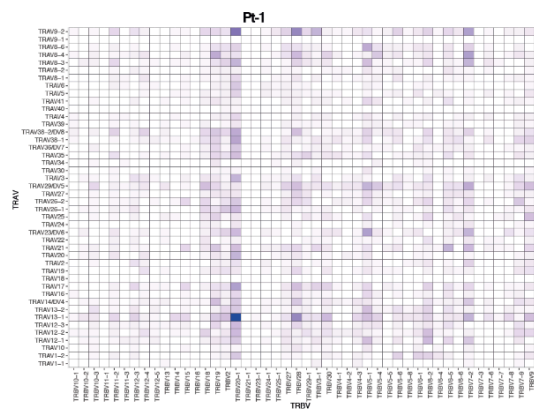
a-b, Age and COVID-19 symptom duration in patients from the old and young groups. **c**, Linear regression analysis of age and symptom duration in convalescent COVID-19 patients. **d-f**, Serological analysis of convalescent COVID-19 patients displaying time of blood collection (**d**), SARS-CoV-2-specific IgG antibody titers (**e**) and SARS-CoV-2-specific IgA antibody titers (**f**), as determined by ELISA. Data are displayed as mean \pm SD. Asterisks indicate significant differences between groups. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, ns = not significant.



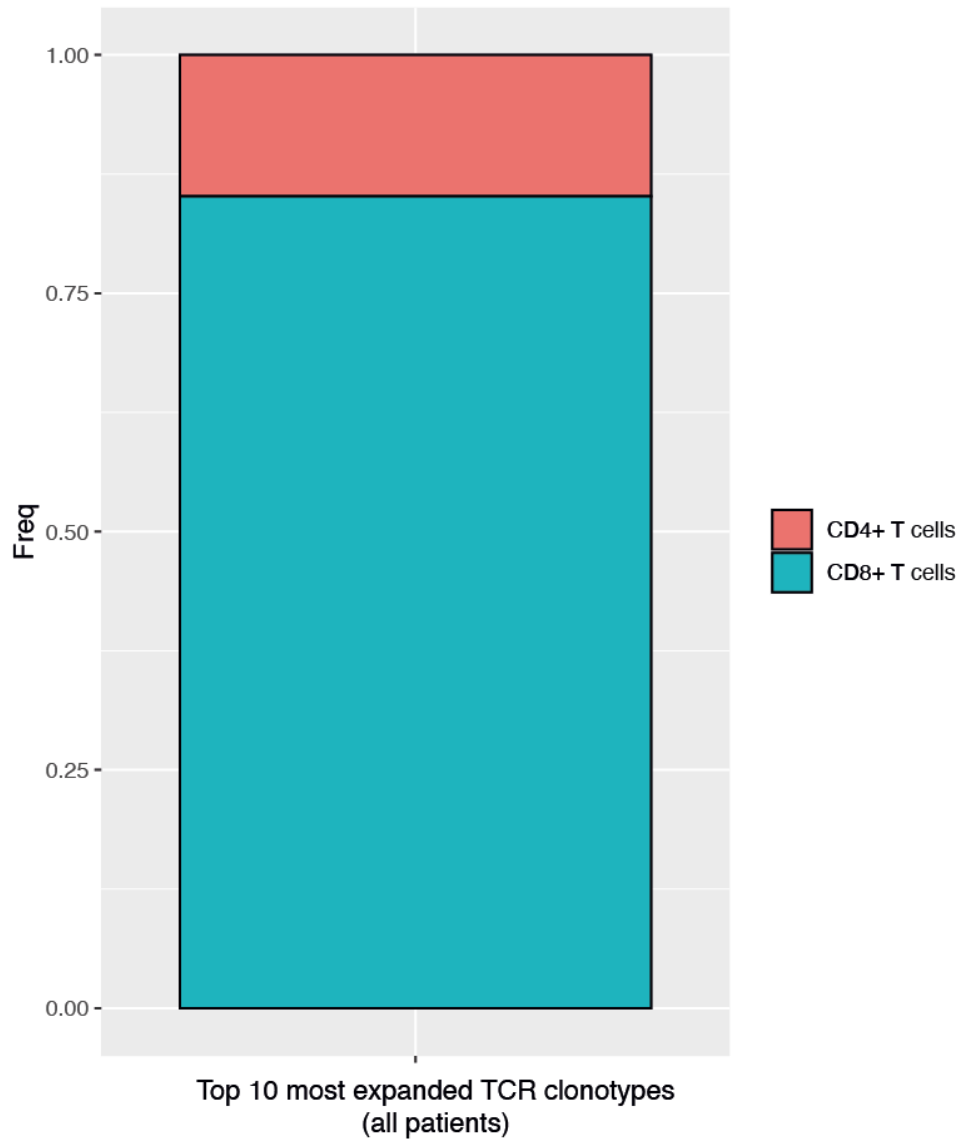
Supplemental Figure 2. Exhausted CD8+ T cells predominantly localize to the effector CD8+ T cell region. UMAP plot shows expression of the LAG-3 exhaustion marker by individual CD8+ T cells from all patients.



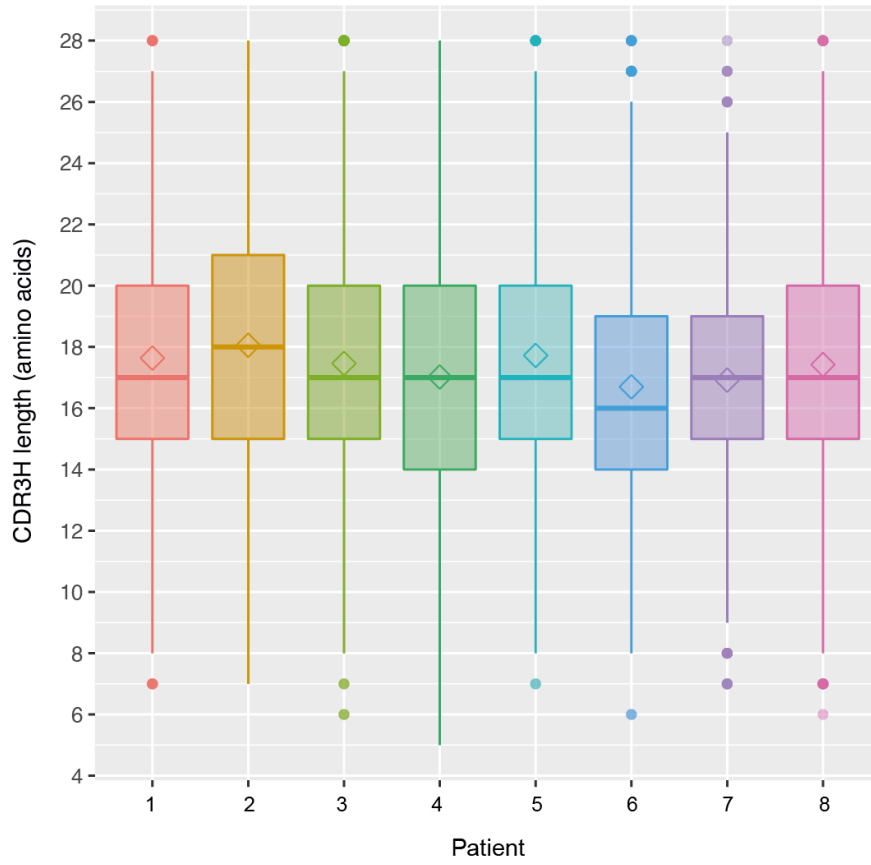
Supplemental Figure 3. Proportions of lymphocyte subsets across studied age groups. Bar graphs compare the proportions of T cell and B cell subsets in the old and young groups. Data are displayed as mean \pm SD. Asterisks indicate significant differences between groups. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, ns = not significant.



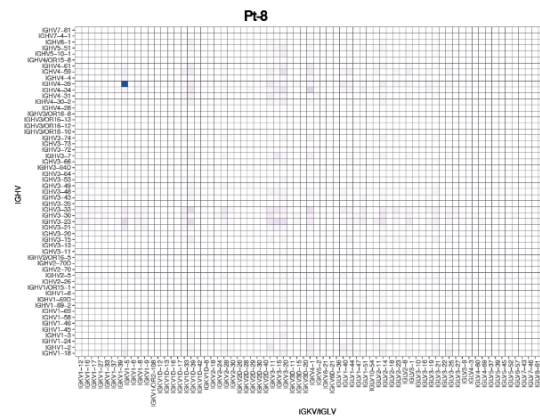
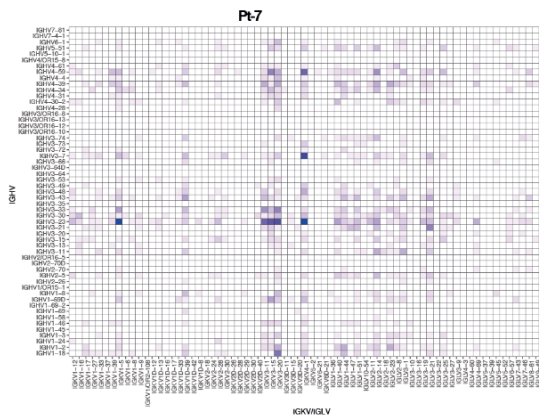
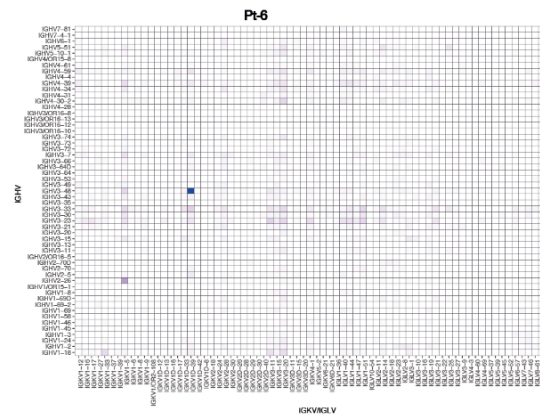
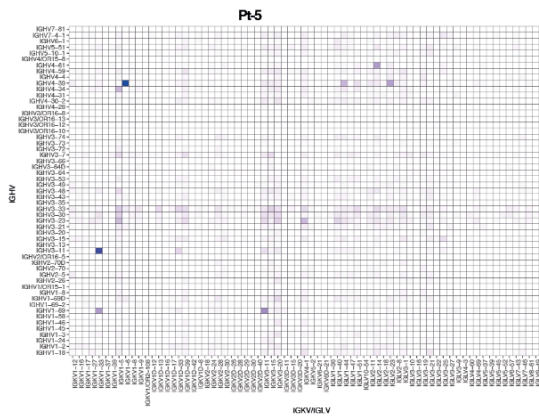
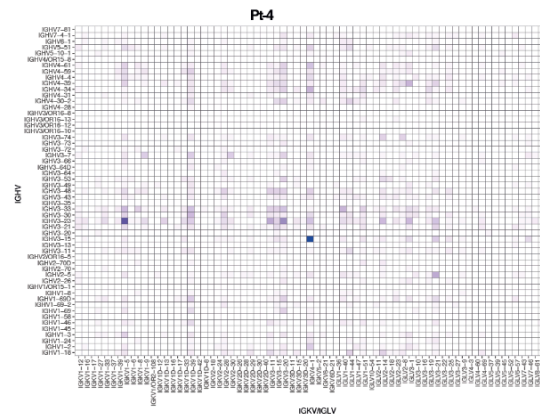
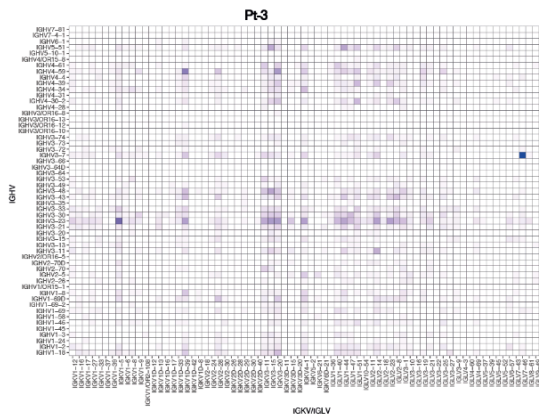
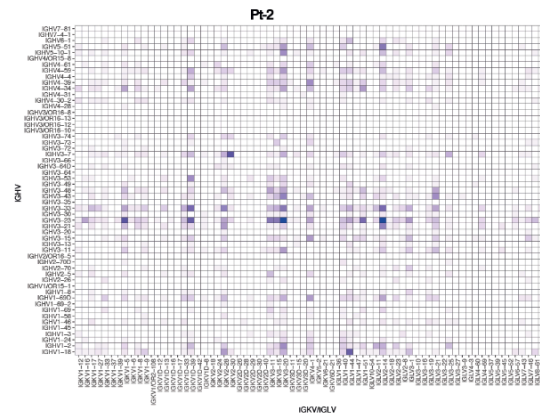
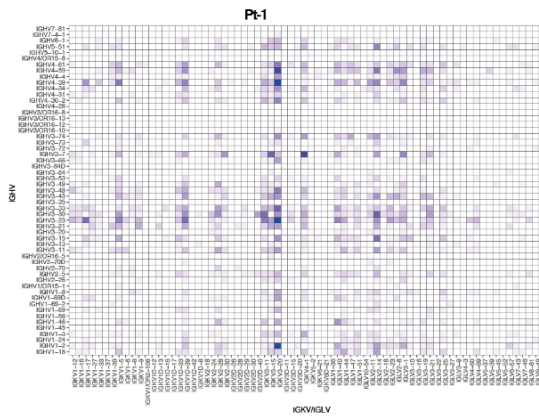
Supplemental Figure 4. TCR V-gene pairings in convalescent COVID-19 patients. Heatmaps display the relative frequencies of specific TCR V-gene pairings in the T cells of individual patients. Data is normalized to the most frequent pairing found in each patient.



Supplemental Figure 5. T cell clonal expansion in convalescent COVID-19 patients is dominated by CD8+ subsets. Bar graph shows the proportions of CD8+ T cells versus CD4+ T cells within the top ten most expanded TCR clonotypes from each patient.



Supplemental Figure 6. BCR heavy chain CDR3 length distribution in convalescent COVID-19 patients.



Supplemental Figure 7. BCR V-gene pairings in convalescent COVID-19 patients. Heatmaps display the relative frequencies of specific BCR V-gene pairings in the B cells of individual patients. Data is normalized to the most frequent pairing found in each patient.

Supplemental Table 1. Characteristics of convalescent COVID-19 patients analyzed in this study.

Patient ID	Age	Sex	BMI (kg/m ²)	RT-PCR test	Days ill (before PCR / after PCR / total)	Restrictions	Sample collection (days after symptom onset)	Rapid test (LFA)	ELISA test	IgA titer (OD ratio)	IgG titer (OD ratio)
Pt-1	59	F	33.9	Positive	7 / 7 / 14	Bedridden	32	IgM+ / IgG+	IgA+ / IgG+	10.8	13.1
Pt-2	63	M	24.4	Positive	7 / 7 / 14	Bedridden	41	IgM+ / IgG+	IgA+ / IgG+	13.5	12.8
Pt-3	74	F	23.4	Positive	2 / 7 / 9	Help needed	32	IgM+ / IgG+	IgA+ / IgG+	8.1	9.1
Pt-4)	71	F	28.9	Positive	5 / 7 / 12	Bedridden	30	IgM+ / IgG+	IgA+ / IgG+	9.9	12.4
Pt-5	37	M	28.7	Positive	2 / 5 / 7	No restrictions	33	IgM- / IgG+	IgA+ / IgG+	7.7	3.3
Pt-6	31	M	23.2	Positive	2 / 7 / 9	No restrictions	34	IgM- / IgG+	IgA+ / IgG+	6.5	8.5
Pt-7	23	F	22	Positive	0 / 7 / 7	No restrictions	39	IgM+ / IgG+	IgA+ / IgG+	4.8	5.7
Pt-8	33	M	26	Positive	2 / 2 / 4	No restrictions	30	IgM+ / IgG+	IgA+ / IgG+	4.1	9.5

Supplemental Table 2. HLA class I alleles of studied convalescent COVID-19 patients.

Patient	HLA-A	HLA-B	HLA-Cw
Pt-1	A*03:01, A*30:01	B*13:02, B*40:02	C*02:02, C*06:02
Pt-2	A*02:01, A*26:01	B*38:01, B*44:02	C*05:01, C*12:03
Pt-3	A*01:01, A*24:02	B*07:02, B*15:29	C*07:02, C*07:04
Pt-4	A*02:01, A*02:01	B*15:01, B*35:01	C*03:04, C*04:01
Pt-5	A*11:01, A*11:01	B*18:03, B*40:01	C*03:04, C*07:01
Pt-6	A*01:01, A*68:01	B*07:02, B*35:03	C*04:01, C*07:02
Pt-7	A*03:01, A*24:02	B*18:01, B*35:02	C*04:01, C*07:01
Pt-8	A*02:01, A*24:02	B*13:02, B*35:02	C*04:01, C*06:02

[not listed here] Supplemental Table 3. GLIPH2 analysis of patient TCRs and known HLA-A*02:01-restricted SARS-CoV-2-specific TCRs.

[not listed here] Supplemental Table 4. GLIPH2 analysis of patient TCRs and known HLA-A*02:01-restricted CMV- and EBV-specific TCRs.

Supplemental Table 5. Primer sequences for HLA class I typing.

Primer Name	Restriction site (HindIII or BamHI)	Index	Target specific sequence	Final sequence (5' - 3')
PCR1_fwd-1	TACGCCAAGCTT	ACGAGTGCCT	TGGCCCTGACCSAGACCTG	TACGCCAAGCTTACGAGTGCCTTGGCCCTGACCSAGACCTG
PCR1_fwd-2	TACGCCAAGCTT	ACGCTCGACA	TGGCCCTGACCSAGACCTG	TACGCCAAGCTTACGCTCGACATGGCCCTGACCSAGACCTG
PCR1_fwd-3	TACGCCAAGCTT	AGACGCACTC	TGGCCCTGACCSAGACCTG	TACGCCAAGCTTAGACGCACTCTGGCCCTGACCSAGACCTG
PCR1_fwd-4	TACGCCAAGCTT	AGCACTGTAG	TGGCCCTGACCSAGACCTG	TACGCCAAGCTTAGCACTGTAGTGGCCCTGACCSAGACCTG
PCR1_fwd-5	TACGCCAAGCTT	ATCAGACACG	TGGCCCTGACCSAGACCTG	TACGCCAAGCTTATCAGACACGTGGCCCTGACCSAGACCTG
PCR1_fwd-6	TACGCCAAGCTT	ATATCGCGAG	TGGCCCTGACCSAGACCTG	TACGCCAAGCTTATATCGCGAGTGGCCCTGACCSAGACCTG
PCR1_fwd-7	TACGCCAAGCTT	CGTGTCTCTA	TGGCCCTGACCSAGACCTG	TACGCCAAGCTTCGTGTCTCTATGGCCCTGACCSAGACCTG
PCR1_fwd-8	TACGCCAAGCTT	CTCGCGTGTC	TGGCCCTGACCSAGACCTG	TACGCCAAGCTTCTCGCGTGCTGGCCCTGACCSAGACCTG
PCR1_rev	ACCCGGGGATCC	-	GKCCTCGCTCTGGTTGTAGT	ACCCGGGGATCCGKCCTCGCTCTGGTTGTAGT
PCR2_fwd-1	TACGCCAAGCTT	TAGTATCAGC	ACTACAACCAGAGCGAGGMC	TACGCCAAGCTTTAGTATCAGCACTACAACCAGAGCGAGGMC
PCR2_fwd-2	TACGCCAAGCTT	TCTCTATGCG	ACTACAACCAGAGCGAGGMC	TACGCCAAGCTTTCTCTATGCGACTACAACCAGAGCGAGGMC
PCR2_fwd-3	TACGCCAAGCTT	TGATACGTCT	ACTACAACCAGAGCGAGGMC	TACGCCAAGCTTTGATACGTCTACTACAACCAGAGCGAGGMC
PCR2_fwd-4	TACGCCAAGCTT	TACTGAGCTA	ACTACAACCAGAGCGAGGMC	TACGCCAAGCTTTACTGAGCTAACTACAACCAGAGCGAGGMC
PCR2_fwd-5	TACGCCAAGCTT	CATAGTAGTG	ACTACAACCAGAGCGAGGMC	TACGCCAAGCTTCATAGTAGTACTACAACCAGAGCGAGGMC
PCR2_fwd-6	TACGCCAAGCTT	CGAGAGATAC	ACTACAACCAGAGCGAGGMC	TACGCCAAGCTTCGAGAGATACACTACAACCAGAGCGAGGMC
PCR2_fwd-7	TACGCCAAGCTT	ATACGACGTA	ACTACAACCAGAGCGAGGMC	TACGCCAAGCTTATACGACGTAATACTACAACCAGAGCGAGGMC
PCR2_fwd-8	TACGCCAAGCTT	TCACGTACTA	ACTACAACCAGAGCGAGGMC	TACGCCAAGCTTTCACGTACTAACTACAACCAGAGCGAGGMC
PCR2_rev	ACCCGGGGATCC	-	TGCCAGGTCAGTGTGATCTC	ACCCGGGGATCCTGCCAGGTCAGTGTGATCTC

A.2 Supplementary information of Chapter 3

Due to the extensive length of some supplementary material, not the entire supplementary material is presented here. If needed, it can be requested from the authors.

Supplemental Tables 1-2

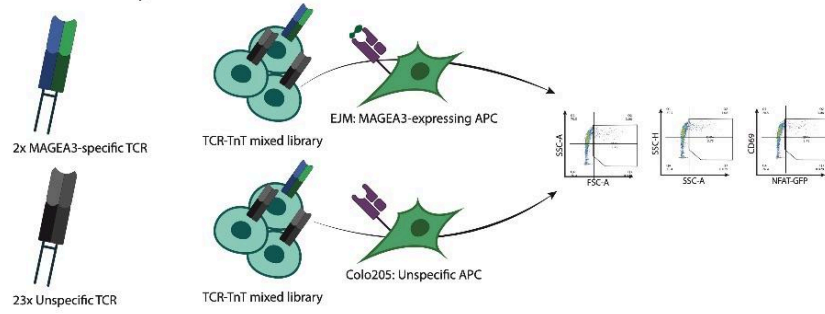
Supplemental Figures 1-9

Supplemental Table 1: Sequencing data generated per patient and meta data on patient characteristics and lung cancer subtype.

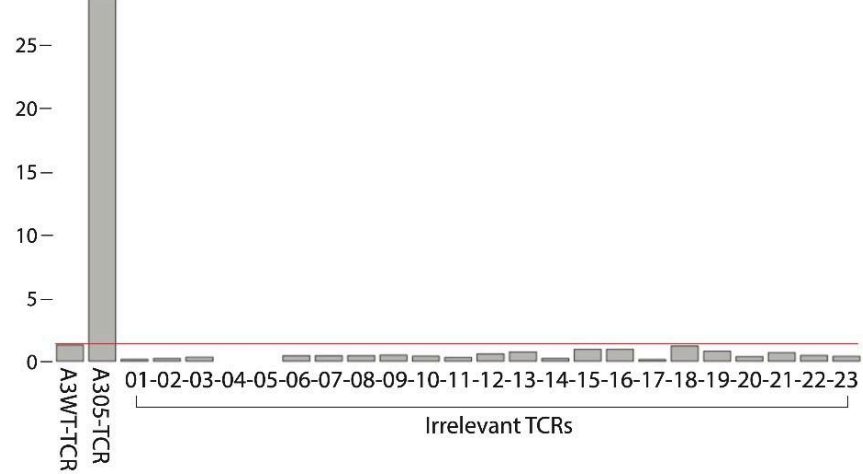
Patients	Age	Diagnosis	Gender
BS476	76	Adenocarcinoma	male
BS631	75	Squamous cell carcinoma	male
BS833	60	Squamous cell carcinoma	male
BS867	69	Adenocarcinoma	male
BS980	74	Squamous cell carcinoma	female
BS1047	60	Adenocarcinoma	male
BS1064	75	Squamous cell carcinoma	male
BS1140	63	Squamous cell carcinoma	male
LI033	55	Adenocarcinoma	female

[Not listed here] **Supplemental Table 2:** P-value and fold change (FC)-value for each TCR clonotype and co-culture condition and patient.

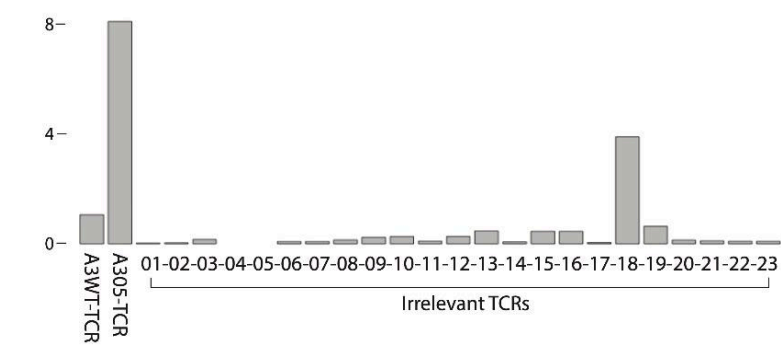
A Summary



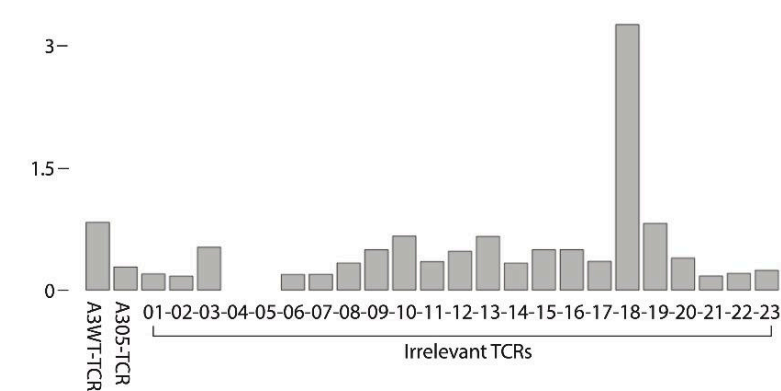
B Co-culture with EJM (MAGEA3-expressing APC) vs Colo205 (unspecific APC)



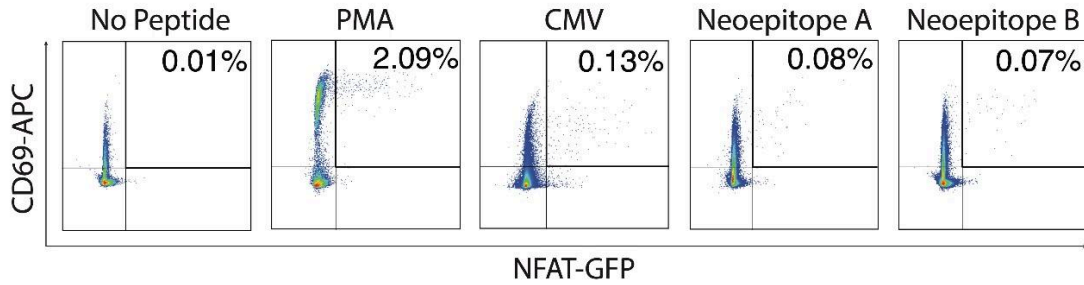
C Co-culture with EJM vs pre-Co-culture



D Co-culture with Colo205 vs pre-Co-culture

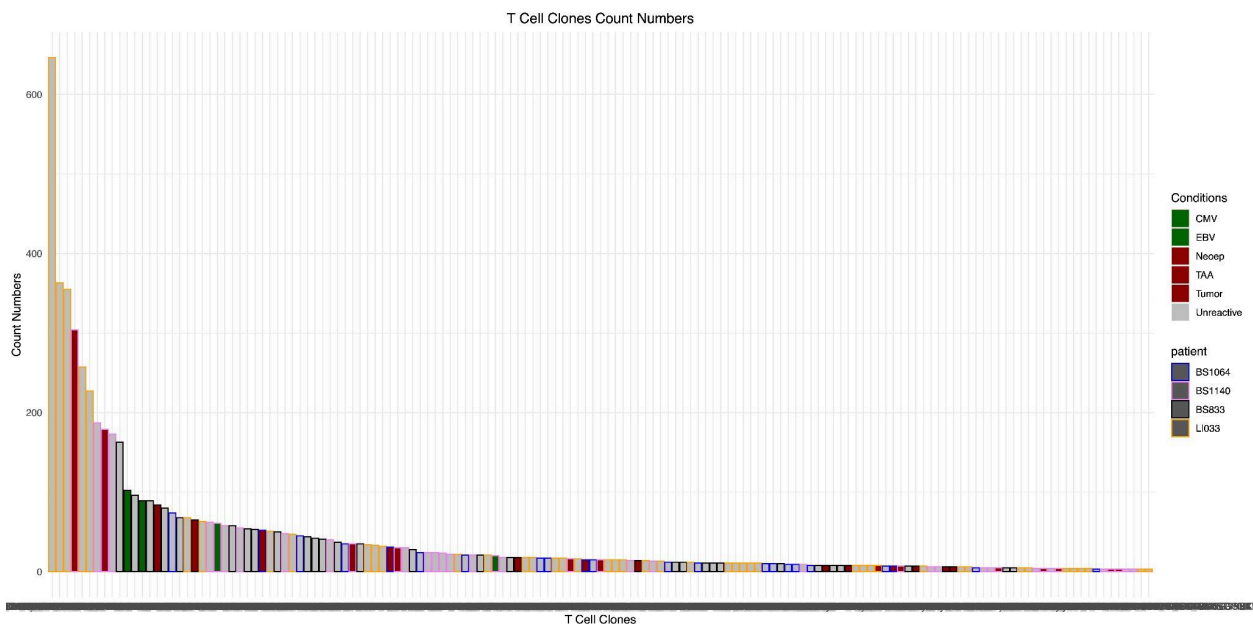


Supplemental Figure 1: (A) Schematic of TCRs with either MAGE-A3 specificity or unspecific TCRs, expressed as mixed library in TnT cells and co-cultured with MAGE-A3 expressing or unspecific (no MAGE-A3 expression) APCs, followed by FACS-based readout. (B) Relative fold-change frequency per TCR after co-culture with MAGE-A3 expressing APCs compared to unspecific APCs. (C) Relative fold-change frequency per TCR after co-culture with MAGE-A3 expressing APCs compared to pre-coculture TCR frequencies. (C) Relative fold-change frequency per TCR after co-culture with unspecific APCs compared to pre-coculture TCR frequencies.

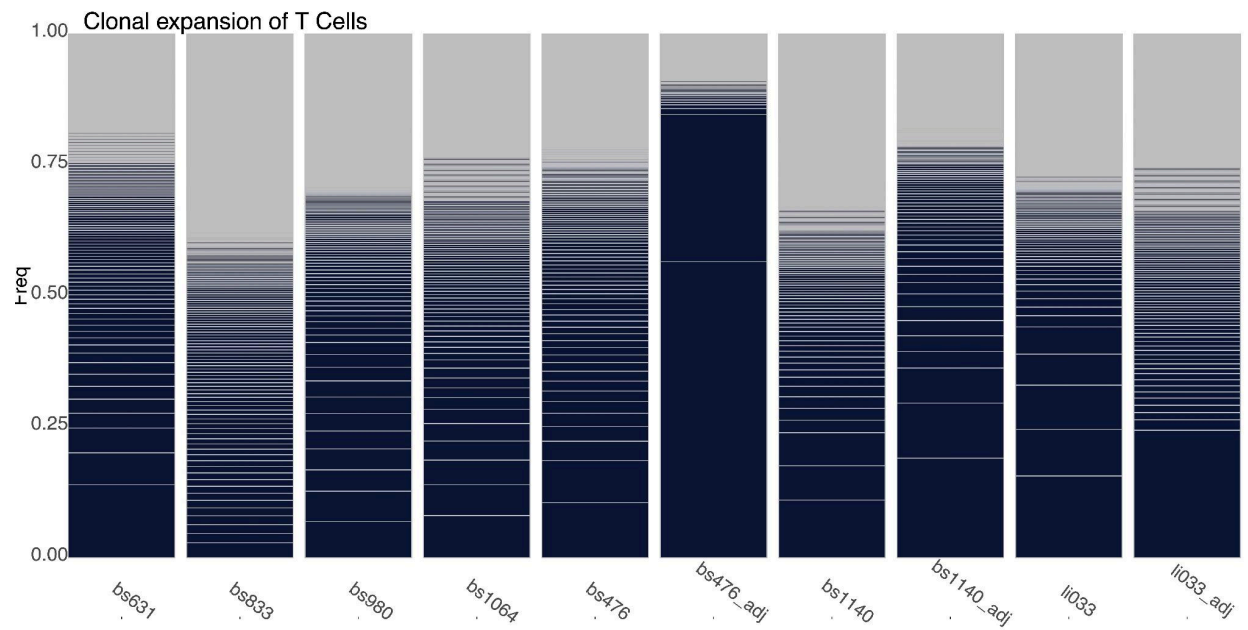


Supplemental Figure 2: Example FACS results for different types of co-cultures of TCR-TnT libraries and control conditions or peptide conditions.

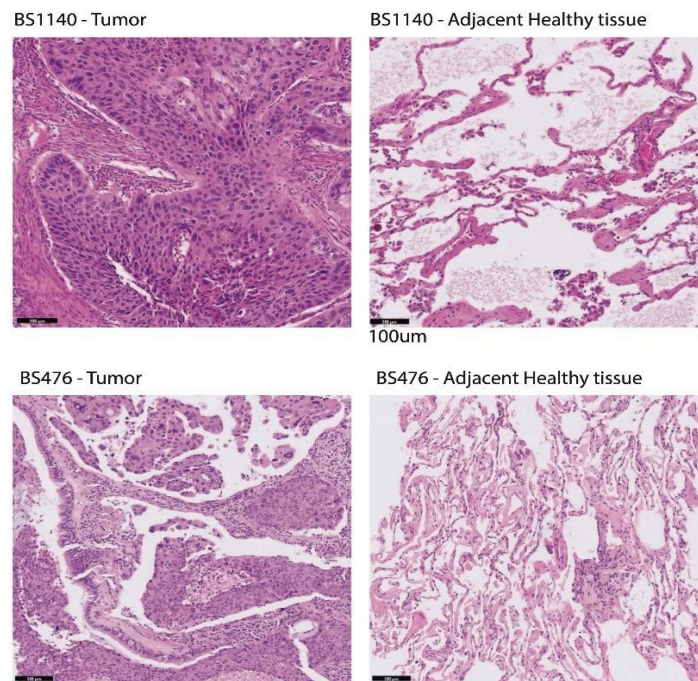
[Not listed here] **Supplemental Figure 3:** Plots for relative frequency per TCR across co-culture conditions per patient and screening experiment.



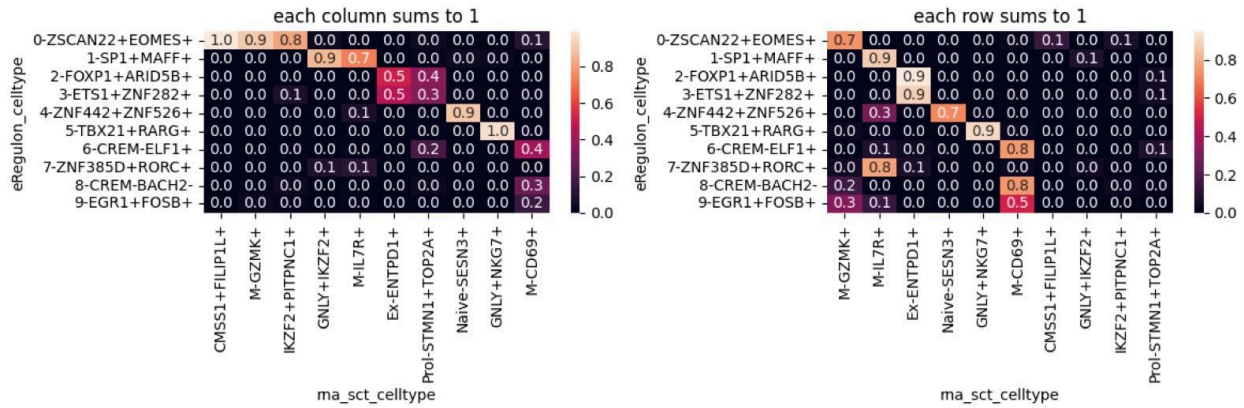
Supplemental Figure 4: T cell clone sizes across patients with color-coded functional reactivity.



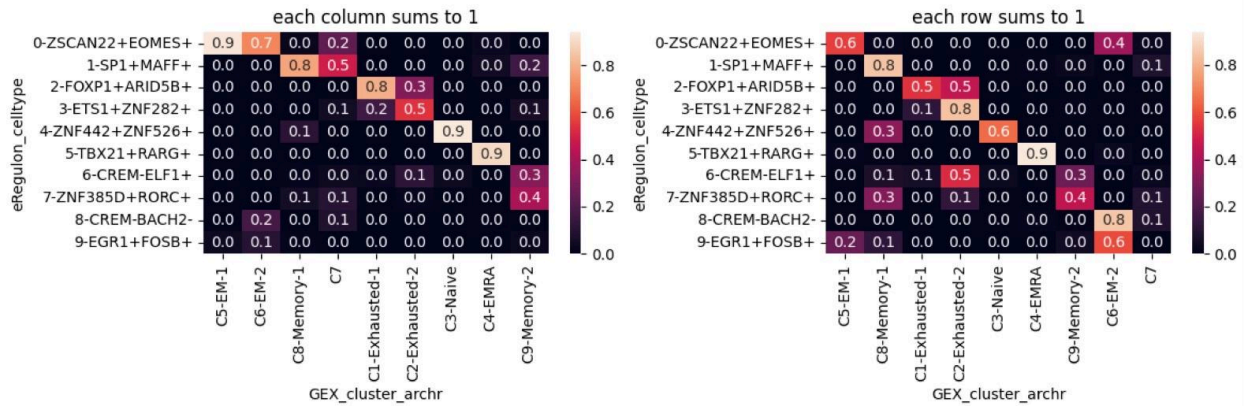
Supplemental Figure 5: TCR repertoire clonal expansion for each patient. Clones are separated by grey lines and expansion-levels are normalized to 100%.



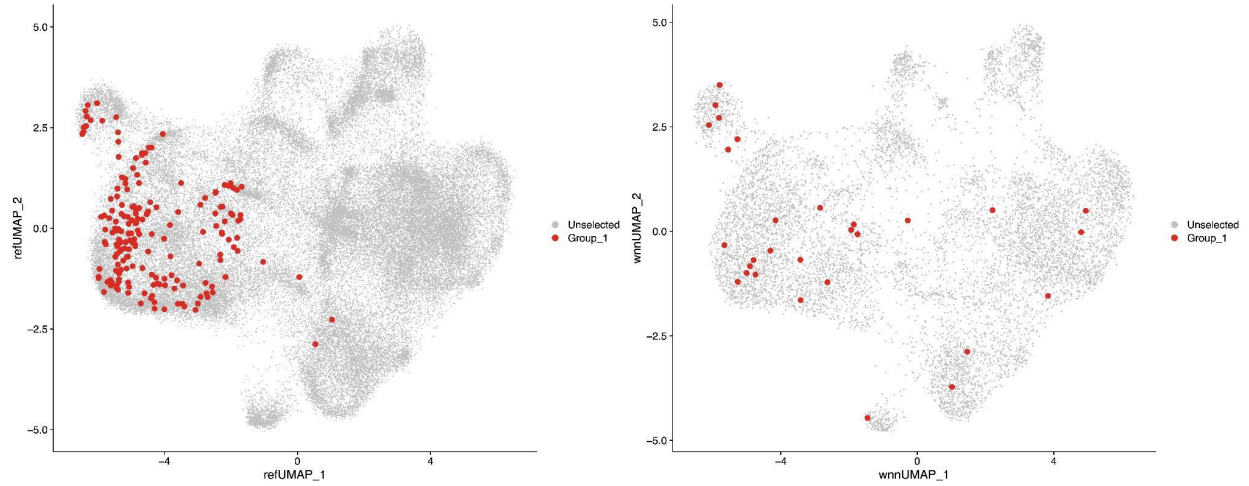
Supplemental Figure 6: HE-slides for tumor and adjacent tissue samples that were used in this study.



Supplemental Figure 7: Shared cells across clusters generated by RNA-modality, compared to eRegulon-modality. Left, cells in RNA-generated clusters spread across eRegulon-generated clusters. Right, cells in eRegulon-generated clusters spread across RNA-generated clusters.



Supplemental Figure 8: Shared cells across clusters generated by ATAC-modality, compared to eRegulon-modality. Left, cells in ATAC-generated clusters spread across eRegulon-generated clusters. Right, cells in eRegulon-generated clusters spread across ATAC-generated clusters.



Supplemental Figure 9: T cells that share expression of the same TCR clone “CASSSSYEQYF_CAVRYPGGDDKIIF” (CDR3B_CDR3A) with reactivity against an epitope of the mutated SORCS1 gene, identified in tumor exome sequencing data. Left, TCR clone in scRNA-seq data of patient BS1140. Right, TCR clone in scMultiomic data of patient BS1140.