

Sofya Kasatskaya

**ORIGIN OF T CELL SUBSETS
STUDIED THROUGH THE LENS OF TCR REPERTOIRES**

Doctoral Thesis





Skolkovo Institute of Science and Technology
Skolkovo Institute of Science and Technology

**ORIGIN OF T CELL SUBSETS
STUDIED THROUGH THE LENS OF TCR REPERTOIRES**

Doctoral Thesis

by

SOFYA KASATSKAYA

DOCTORAL PROGRAM IN LIFE SCIENCES

Supervisor:

Associate Professor Dmitriy Chudakov, Ph.D. D.Sc.

Co-supervisor: Olga Britanova, Ph.D.

Moscow - 2021

© Sofya Kasatskaya 2021

I hereby declare that the work presented in this thesis was carried out by myself at Skolkovo Institute of Science and Technology, Moscow, except where due acknowledgement is made, and has not been submitted for any other degree.

Candidate Sofya A. Kasatskaya
Supervisor Prof. Dmitriy Chudakov
Dr. Olga V. Britanova

TABLE OF CONTENTS

ABSTRACT	6
ACKNOWLEDGEMENTS	8
LIST OF PUBLICATIONS	10
LIST OF ABBREVIATIONS	11
INTRODUCTION	12
1. Role of T cells in human adaptive immunity	
1.1. The diversity of T cells subsets and functions.....	12
1.2. T cell receptor structure and function	15
1.3. T cell receptor repertoires sequencing technology.....	18
1.4. TCR repertoires profiling in system immunology	22
1.5. Individual T cell development and fate	23
1.6. TCR role in T cell lineage selection and survival	26
2. Helper T cell functional subsets diversity	27
3. Gamma/delta T cells and TCR repertoires	
3.1. $\gamma\delta$ T cell biology and function. Functional subsets	29
3.2. $\gamma\delta$ T cell subsets in inflammation	31
3.3. $\gamma\delta$ T cell subsets. VDelta1+ and VDelta2+ subsets	32
3.4. Resident memory T cells' TCR repertoires.....	35
AIMS OF STUDY	38

RESULTS

Chapter I

The changing landscape of naive T cell receptor repertoire with human aging 40

Chapter II

Memory CD4⁺ T cells are generated in the human fetal intestine 54

Chapter III

Functionally specialized human CD4⁺ T cell subsets express physicochemically distinct TCRs 86

Chapter IV

The human V δ 2⁺ T cell compartment comprises distinct innate-like V γ 9⁺ and adaptive V γ 9⁻ subsets 123

Chapter V

Human liver infiltrating $\gamma\delta$ T cells are composed of clonally expanded circulating and tissue-resident populations 146

Chapter VI

$\gamma\delta$ T cell receptors derived from breast cancer-infiltrating T lymphocytes mediate antitumor reactivity 169

CONCLUSIONS 197

REFERENCES..... 201

ABSTRACT

The immune system consists of various cell types, all of which are controlled and directed by signals from helper T cells and functionally similar regulatory cells. Clonal fate and lineage choice of T cells are instructed by costimulatory signals in the moment of priming but may also depend on the mode and duration of T cell receptor interaction with antigen. We investigate how TCR structure predisposes T cells to specific functional and phenotypical states. We are interested both in classical and non-classical T cells, and we study the extent, complexity, and homogeneity of TCR structure influence on the T cell subset selection and survival.

This thesis is structured as a compilation of published articles contributing to the research of alpha/beta and gamma/delta T cell receptor repertoires. Chapter I studies the peripheral selection, survival, and aging among human naive T cells, both CD4+ and CD8+ subsets. In contrast to mice, who produce most of the naive T cells in the thymus, humans replenish their naive T cells pool through homeostatic proliferation due to age-related thymus involution. In order to study the rules for naive T cell survival, we compared the naive TCR repertoires from young and aged donors. Moreover, to study the peripheral selection after thymus exit, we compared RTE and mature naive T cell populations and found high homogeneity in these two subsets. We discovered that age-related changes in TCR repertoires include shortened and less diverse CDR3 regions of TCRs, and the changes are more prominent in the RTE subset. Chapter II continues the research on the selection and survival of naive T cells with a unique dataset. We demonstrate the phenomenon of immunological memory in helper T cells acquired in embryonic intestinal tissue for the first time. This collaborative project describes the distinct subsets of CD4+ Treg and non-Treg memory T cells in the fetal gut through mass cytometry, RNA-seq, and TCR-seq.

Chapter III is the central part of this Ph.D. thesis. It investigates the fundamental questions in helper T cell biology, such as the plasticity of Th1/Th2 and Th17/Treg functional subsets, the stability of acquired clonal programs, and the level of heterogeneity of TCR repertoire features among subsets and unrelated healthy donors. We sorted and profiled intact effector memory helper T cells of the following subsets: Th1, Th2, Th2a, Th17, Th17-1, Th22, Treg, and Tfh. We show that even among unrelated donors, functional T cell subsets share the same patterns of prominent physicochemical TCR features. We demonstrate high plasticity within Th17 and Th22 functional subsets and prominent clonal exchange of both Th17 and Th22 with Th2. Th1-17 subset shares clones with Th1 but not with the Th17 subset. Finally, we investigate relative publicity of effector CD4 subset repertoires across donors. We reveal high publicity of Treg and Tfh and high privacy of Th22 and Th2a repertoires, which corresponds to the number of added N-nucleotides for these subsets, suggesting their earlier fetal and evolutionary origin. Altogether, we provide the first detailed picture of distinct repertoire features, plasticity, and publicity of helper CD4 T cell subsets.

Chapters IV to VI investigate the role of TCR structure in the biology of human gamma/delta T cells in peripheral blood and tissues. The proposed approach to $\gamma\delta$ TCR sequencing and data analysis allows the correction of cross-contaminations and unbiased clonality estimation even in highly expanded T cell samples. Chapters IV and V characterize new innate-like and adaptive-like subsets of blood gamma/delta T cells and a tissue-resident subset in the liver. Chapter VI discusses the role of public T cell receptors in cancer surveillance on an example of gamma/delta TILs in breast cancer samples. Together, this thesis contributes to a more comprehensive model of the TCR repertoire landscape and its changes across functional T cell subsets, across locations of systemic and tissue-specific immunity, and across the human lifespan.

ACKNOWLEDGEMENTS

First of all, I would like to thank all donors who have participated in my research on human immunity. Thank you for agreeing to help in my research.

I am deeply grateful to my colleagues from the Genomics of Adaptive Immunity Laboratory, especially to Dr. Olga Britanova, Mark Izraelson, Dr. Evgeny Egorov, and Dr. Katya Merzlyak. Thank you for the teamwork, inspiration, and good time spent and for being a great example.

I would like to acknowledge my supervisor Prof. Dmitriy Chudakov, who provided me with both guidance, mentorship, and a high degree of freedom and independent exploration. I am grateful for all the big picture in immunology discussions, for the shared knowledge experience, and for the trust in me during all ups and downs of a Ph.D. degree training. Many thanks for the opportunity to teach concepts in immunology at a new and exciting level of MSc students.

I also would like to appreciate the help of Alena Zueva and Nastya Koshenkova, who perform a lot of paperwork for our grants, projects and cherish the laboratory's well-being. Here goes a special acknowledgment to PMI-Skoltech System Biology Fellowship, which supported my Ph.D. research.

I am grateful to my collaborators, primarily Dr. Kristin Ladell, Prof. David Price, Dr. Carrie Willcox and Prof. Ben Willcox, Dr. Jürgen Kuball, and all the teams of the corresponding laboratories in Cardiff, Birmingham, and Utrecht. It was a pleasure working with you and learning from your bright ideas.

Thank you so much to the Skoltech community. Working at Skoltech appeared in my life as an unexpected and fantastic opportunity to dive into a world on edge between academia and innovation-fueled industry. Skoltech students showed me high levels of determination, always positive thinking, pro-active position, and integrity. Moreover, here I found the mentorship and soft skill training I needed. Here I explored my career possibilities and expanded my network. Undoubtedly, the unique, fresh, and bright academic community will continue to inspire me long after I graduate with a Skoltech degree.

Finally, I am thankful for the invaluable community support. Many thanks to my parents, who initially inspired me to study biology and continuously continue to support my interests and choices. I want to acknowledge the community of young scientists around Biomolecula portal and Science FYI community created by Dr. Viktoria Korzhova: these initiatives help a lot to sustain motivation and prevent severe burnouts in academic career. Special thank you to my close friends. You are always there to support and inspire me throughout the long journey of Ph.D. research and hard times of self-retraining to work in bioinformatics.

PUBLICATIONS

1. **Kasatskaya SA***, Ladell K*, Egorov ES, Miners KL, Davydov AN, Metsger M, Staroverov DB, Matveyshina EK, Shagina IA, Mamedov IZ, Izraelson M, Shelyakin PV, Britanova OV, Price DA and Chudakov DM, Functionally specialized human CD4+ T cell subsets express physicochemically distinct TCRs, *eLife*, 2020; 9:e57063. doi: <https://doi.org/10.7554/eLife.57063>.
2. Janssen A, Villacorta Hidalgo J, Beringer DX, Chudakov DM, **Kasatskaya SA**, Frenkel FE, Prinz I, Kuball J et al. $\gamma\delta$ T-cell Receptors Derived from Breast Cancer-Infiltrating T Lymphocytes Mediate Antitumor Reactivity. *Cancer Immunol Res*. 2020;8(4):530-543. doi: 10.1158/2326-6066.CIR-19-0513.
3. Li N, van Unen V, Abdelaal T, Guo N, **Kasatskaya SA**, Ladell K, McLaren JE, Egorov ES, Izraelson M, Chuva de Sousa Lopes SM, Hilt T, Britanova OV, Eggermont J, de Miranda NFCC, Chudakov DM, Price DA, Lelieveldt BPF, Koning F. Memory CD4+ T cells are generated in the human fetal intestine. *Nat Immunol*. 2019 Mar;20(3):301-312. doi: 10.1038/s41590-018-0294-9
4. Egorov ES*, **Kasatskaya SA***, Zubov VN, Izraelson M, Nakonechnaya TO, Staroverov DB, Angius A, Cucca F, Mamedov IZ, Rosati E, Franke A, Shugay M, Pogorelyy MV, Chudakov DM, Britanova OV. The Changing Landscape of Naive T Cell Receptor Repertoire With Human Aging. *Front Immunol*. 2018 Jul 24;9:1618. doi: 10.3389/fimmu.2018.01618.
5. Hunter S, Willcox CR, Davey MS, **Kasatskaya SA**, Jeffery HC, Chudakov DM, Ye H Oo, Willcox BE. Human liver infiltrating $\gamma\delta$ T cells are composed of clonally expanded circulating and tissue-resident populations. *Journal of hepatology* 2018: 69 (3), p.654-665.
6. Davey MS, Willcox CR, Hunter S, **Kasatskaya SA**, Remmerswaal EBM, Salim M, Mohammed F, Bemelman F, Chudakov DM, Ye H Oo, Willcox B. The human V δ 2+ T-cell compartment comprises distinct innate-like V γ 9+ and adaptive V γ 9-subsets. *Nature Communications* 2018 volume 9, Article number: 1760.

* equal contribution

CONFERENCES

1. Synthims: Synthetic and Systems Immunology, 5-8 May 2019, Ascona, Switzerland, poster "TCR repertoires show thymic selection for human naïve Tregs but do not differentiate between RTE and mature naïve T cells".
2. ThymE: Thymus and T cell biology workshop, poster "TCR repertoires show thymic selection for human naïve Tregs but not the CD4 peripheral selection".
3. **Kasatskaya SA**, Ladell K, Egorov ES, Price DA and Chudakov DM, T cell receptor repertoire features display universal rules for selection and plasticity in the functional CD4 T cell subsets. **IMMUNOLOGY2020 conference abstract** collection, *J Immunol* May 1, 2020, 204 (1 Supplement) 230.9 (Special edition https://www.jimmunol.org/content/204/1_Supplement).

LIST OF ABBREVIATIONS

ILC – innate lymphoid cells

NKT cells – T lymphocyte subsets with high expression of NK cell specific receptors

MHC – major histocompatibility complex

pMHC – peptide-MHC complex

RTE – recent thymic emigrants, a stage in T cell development

RNA-seq – RNA (transcriptome) sequencing technology

TCR – T cell receptor

$\gamma\delta$ TCR – gamma/delta T cell receptor

$\gamma\delta$ T cells – T cells which express $\gamma\delta$ TCR

T_{CM} – central memory T cells

T_{EM} – central memory T cells

T_{EMRA} – short-lived effector cells

T_{fh} – follicular helper T cells

Th1 – helper T cells type 1, initiating type 1 immunity

Th2 – helper T cells type 2, initiating type 2 immunity

Th17 – helper T cells type 17, initiating type 3 immunity

Treg – regulatory T cell

T_{RM} – resident memory T cell

T_{SCM} – stem cell memory T cell

TCR-seq – TCR repertoire sequencing via high-throughput technologies

UCB – umbilical cord blood

UMI – unique molecular identifier

INTRODUCTION

1. Role of T cells in human adaptive immunity

1.1. *The diversity of T cells subsets and functions*

T cells occupy the niche of decision-makers in the immune system. All T cells are characterized by the presence of a unique T cell receptor that functions to establish high specificity of antigen recognition in the context of MHC molecules. While the task of cytotoxic CD8⁺ T cells is to compete and create the most specific and effective mode to recognize and destroy the infected host cells, tasks and goals of helper T cells are much more diverse. After TCR-induced activation, helper T cells initiate the massive production of cytokines – signaling molecules in the immune system. Through cytokine signal propagation, helper T cells are able to activate a particular branch of the immune response or suppress it in the case of regulatory T cells (Tregs). Helper T cells can activate macrophages, granulocytes of certain subtypes, or alter the functional states of other innate immune cells [1, **Fig.1**]. Helper T cells license activated killer T cells to ensure the safety of cytotoxicity inside the tissue [1,2]. Dendritic cells' final maturation is dependent on helper T cells-derived cytokine signals as well [3,4].

Helper T cells of a specialized subset – follicular helper T cells (T_{fh}) are required to induce somatic hypermutagenesis in B cell follicles. Thus, helper T cells are indispensable to produce high-affinity antibodies in the course of the adaptive immune response. Finally, helper T cells and Tregs regulate the timing of the onset and termination of the immune response and the transition to the regeneration phase. Altogether, it is hard to imagine either cellular or humoral immune response without helper T cells involvement at multiple stages of the process.

An individual helper T cell cannot perform all of the functions mentioned above, e.g., simultaneously or consequently provide help to B cell maturation and suppress inflammation. On the contrary, T cells but acquire a narrow specialization during their development. The major paths of helper T cell specialization are considered to be the 3 subsets that initiate 3 types of immunity: Th1 helper T cells for Type 1 immunity, Th2 helper T cells for Type 2 immunity, and Th17 helper T cells for Type 3 immunity, along with the fourth important subset of regulatory T cells. These subsets are described in activated effector helper T cells and could be distinguished by cytokine production profile or even by the characteristic surface markers [1,5].

In addition to the four main helper T cell subsets, more functional subsets were described more recently, for example, Th9, Th22, and other subsets with distinct cytokine secretion profiles [1,5,6,7]. How is this functional diversity organized among T cells observed in human blood? If you look at the effector T cells isolated from a normal peripheral blood sample, each type will occupy up to 5-6% of all helper T cells (see Chapter III). What lies in the rest of the helper T cell niche? These are the naive helper T cells with yet undefined functionality and the activated T cells, which by their phenotype occupy an intermediate position that does not fit into the four conventional subsets (**Fig.1**). This observation is discussed in detail in Chapter III of the thesis.

Distinct cytokine production profiles are also observed in innate lymphoid cells (ILC), which resemble T cells in many aspects but do not carry a rearranged T cell receptor. ILCs are rare in blood but are abundant in peripheral tissues, where they can take over some of the helper T cell duties (**Fig.1**). Based on the functional profile, ILCs can also be divided into subsets, such as ILC1, ILC2, and ILC3 [8,9]. Still in central circulating immunity, in blood, lymph nodes, and in spleen, helper T cells are the primary regulators of the inflammatory state and the current balance between type 1/2/3 immunity branches [1,5].

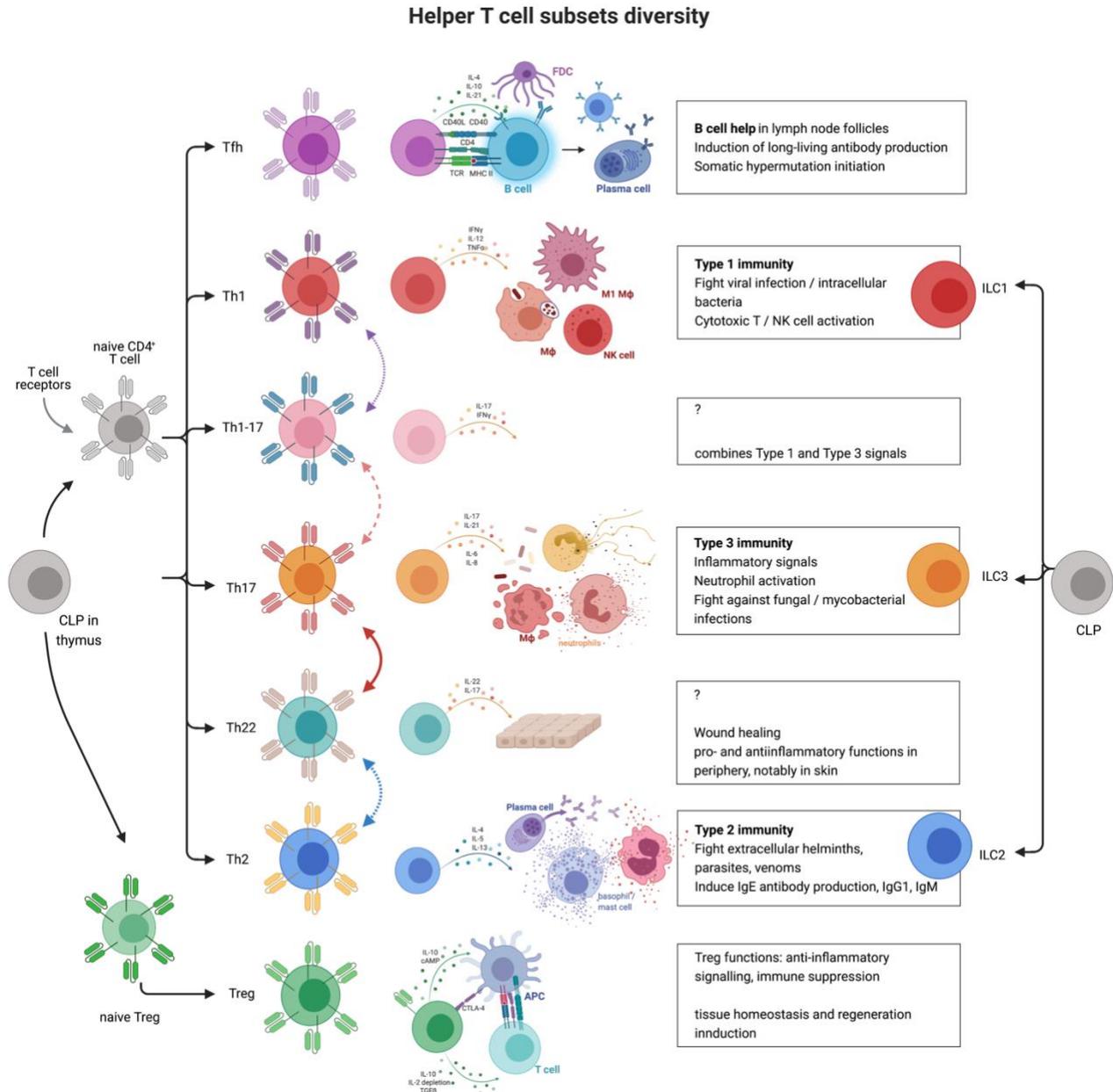


Figure 1. Functional diversity of human CD4⁺ helper T cell subsets. CLP – common lymphoid progenitor, which gives rise both to thymic-derived T cell lineage and thymus-independent innate lymphoid cells (ILC) lineage which lacks the antigen-recognizing T cell receptor. Functional subsets of ILC1, ILC2, ILC3 could mimic the corresponding helper T cell subsets. Illustration was created with BioRender.com.

Another lymphoid cell group takes part in mucosal tissue immunity along with conventional T cells and ILC and fills the spectrum between fully innate and fully adaptive cell types. This group is composed of MAIT T cells, NKT cells, and gamma/delta T cells. These

subsets, for example, gamma/delta T cells, contain innate-like cells and cells with developmental patterns similar to the conventional adaptive immune cells [10, 11, 12]. This is studied in more detail in thesis Chapters IV and V.

Gamma/delta T cells are also present in the peripheral blood where the subset occupies about 1-5% of T cells, similarly to functional subsets of conventional helper T cells (Tfh, Th1, Th2, Th17, etc.). In this regard, it is logical to study various peripheral blood T cell subsets with similar frequency and common regulatory functions in parallel. The current Ph.D. thesis is focused on the research on helper T cells and gamma/delta T cells and their subpopulations through the T cell receptor repertoires structure analysis. Naive CD8⁺ T cells are discussed in Chapter 1, along with naive helper T cells' T cell receptor repertoires study.

1.2. T cell receptor structure and function

T cell receptor is the most important molecular complex expressed on the T cell surface as it determines the specificity to an antigen and the general usefulness of a T cell clone. Each T cell carries on its surface a single specific variant of the T cell receptor typically, although double TCR-bearing T cells were also reported. T cell receptor specificity to the antigen ensures the adequate and balanced response to the presence of infectious or tumor-associated antigens. Two subunits form the T cell receptor complex: α and β , or γ and δ , linked by a disulfide bond and anchored in the surface membrane. In a structural sense, all subunits of T cell receptors belong to the immunoglobulin protein superfamily. Each of the TCR chains includes two domains - the C-terminal, constant (C), which anchors the receptor in the plasma membrane and performs cytoplasmic signal transduction, and the N-terminal, variable (V) domain, responsible for antigen recognition (**Fig. 2**). $\alpha\beta$ T cell receptors recognize antigens only in the form of short peptide

fragments presented on the cell surface as part of a complex glycoprotein complex known as the major histocompatibility complex (MHC), whether $\gamma\delta$ T cell receptors are capable of recognizing non-peptide antigens in the context of non-classical MHC molecules, other receptors, such as EPCR, or in a soluble form [13]. The interaction of the T cell receptor with the MHC complex and the associated antigen occurs through the complementarity determining region (CDR) within the variable domains: three regions from the α -chain and three from the β -chain of the T cell receptor [14].

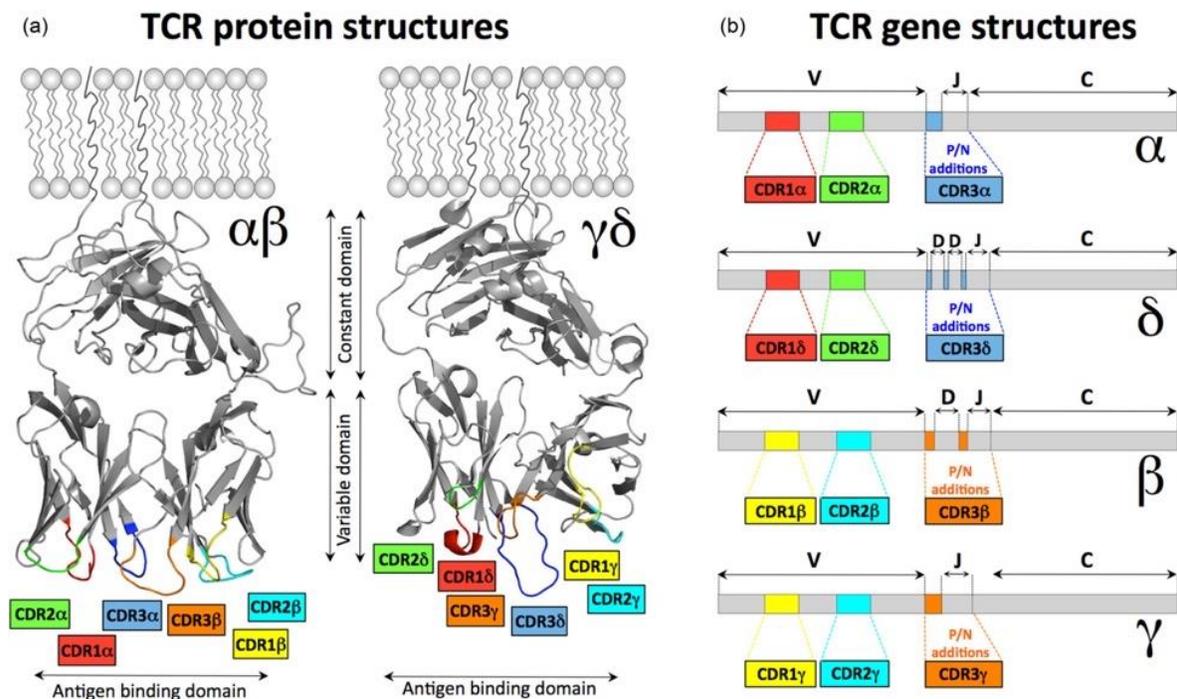


Figure 2. T cell receptor gene and membrane-bound protein complex structure: a scheme. a) Variable and constant domains in the structures of $\alpha\beta$ TCR and $\gamma\delta$ TCR. Complementarity-determining loops 1, 2 and 3 of all chains are color-coded. b) Rearranged TCR gene structure. Positions of variable regions are color-coded as in (a). *Reproduced from [14] under the Creative Commons CC BY license.*

Neither of the genes coding α , β , γ or δ polypeptide chains of the T cell receptor is present in a functional state in a human genome. Instead, these genes are assembled by the sophisticated process of V(D)J recombination of short gene segments in each individual T cell progenitor independently. A complex sequence of reactions has to occur in the TCR locus in the course of V

(D) J recombination during the maturation of T-lymphocytes in the thymus, and a functional gene must always contain 1 V-gene segment, 1 J, and 1 C-gene segment. The correct recombination sequence is ensured by special conserved RSS (recombination signal sequences) that flank V, D, and J gene segments on both sides. The molecular region located at the junction between the V and J gene segments contains the TCR repertoire's greatest diversity and corresponds to the hypervariable CDR3 fragment of the TCR α and TCR β loci. When gene segments are stitched together by ligation, additional diversity is introduced without a template: palindromic P-insertions and random N nucleotides are included in the gene sequence. The P- and N-additions increase the diversity of CDR3 significantly at the VJ junction (for α or γ chains) or the VD and DJ junctions (for β and δ chains) [15, 16]. In delta chains, several D gene segments could be included in the same resulting gene sequence in a row, which allows for more junctional diversity.

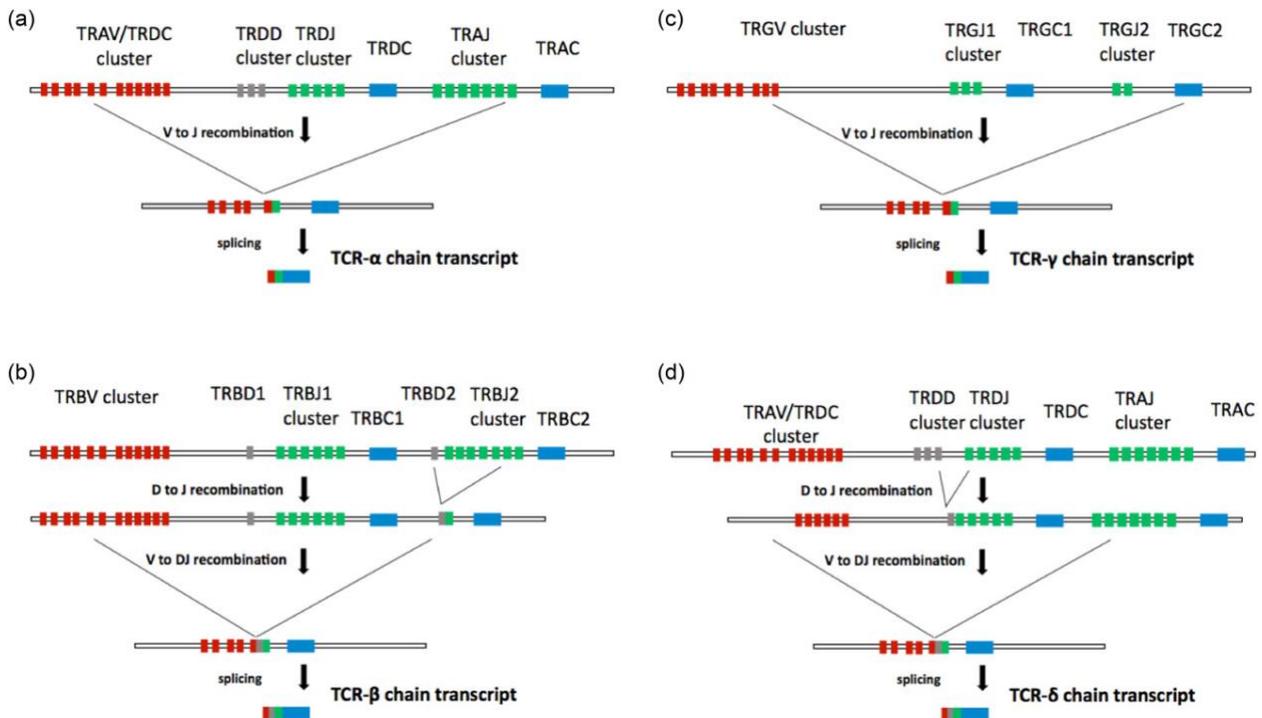


Figure 3. T cell receptor chains are encoded by somatically rearranged gene sequences. Gene segments are quasi-randomly selected and recombined through V(D)J recombination independently for each TCR chain. *Reproduced from [14] under the Creative Commons CC BY license.*

The wide combinatorial and semi-random diversity of the TCR's CDR genetic sequences allows the T cells to recognize millions of various antigens that an immune system encounters during a longer life of a human individual (compared to other multicellular organisms). Theoretically, the repertoire of diverse T cell receptors of the entire human population could be estimated as 10^{15} unique variants [17]. The most significant TCR variability is concentrated in the CDR3 β or CDR3 δ loops, which primarily determine the specificity of binding of the receptor to the antigen in alpha/beta or gamma/delta T cells correspondingly. The CDR1 and CDR2 regions are less specific and are mainly responsible for binding the T cell receptor to MHC [14]. Since the major diversity source is located in the CDR3 region, the sequencing technology to read TCR variability should focus on the shorter region of CDR3; however, it is preferable to sequence full-length molecules instead of CDR3 regions in the case of antibodies.

1.3. T cell receptor repertoires sequencing technology

The strategies for T cell receptor repertoires profiling use the NGS techniques previously developed for other tasks in genomics, but face unique challenges due to the nature of TCR sequences. The field of TCR sequencing is rather new and keeps improving the standards to obtain and interpret the data. The rearrangement of the T cell receptor gene leads to generation of the genetic sequences, which are similar in structure of most of the protein domains; however, they contain hypervariable regions which mostly contribute to the antigen recognition. It means that when sequencing T cell receptors RNA or DNA, we get millions of assembled sequences with high similarity and short, highly dissimilar regions, which creates a special need for accuracy in sequence assembly. Moreover, similar sequences could be generated during the VDJ recombination

process, even from different V-genes and J-genes pairs. Thus, the accuracy of alignment and correct identification of the V- and J-gene segments in the individual TCR sequences is an important goal. If V- or J-gene segments were identified incorrectly, the clonotype assembly from similar and identical sequences would be erroneous, and the estimation of TCR clonality would be further biased [21].

Multiple technologies for TCR profiling exist, both at bulk sequencing and single-cell level. The single-cell technologies are mostly out of scope of the current thesis; bulk TCR sequencing was used in all chapters. The main advantage of single-cell technologies is pairing of the functional phenotype (via scRNA-seq, or CITE-seq) with the TCR clonotype. Another critical advantage is pairing of the TCR chains itself: getting simultaneously both chains of a T cell receptor allow a more straightforward pipeline to testing the antigen specificity of TCR of interest in the downstream *in vitro* experiments. However, the current single-cell technologies still provide less reads per cell than is required for a robust cell identification based on the present cell type-specific transcriptional markers. Expensive reagents add up to the (transient?) disadvantages of the technology.

Bulk TCR sequencing is more productive than single-cell sequencing, but the exact efficiency of the technology may vary (see **Fig.4**). Protocols preparing the TCR libraries from genomic DNA argue that due to DNA stability and the uniform presence of starting material in all T cells, the sequencing protocol is more robust. For the laboratories preparing both TCR and BCR libraries starting from DNA could be an advantage for standardization of the protocols for different kinds of biological samples and adaptive immune cell subsets. DNA-based libraries usually utilize multiplex PCR approach; however, multiplex PCR is also applied to cDNA-based libraries preparation. The disadvantages of multiplex PCR are amplification bias, loss of rare variants, the

prevalence of shorter TCR sequences, and the resulting incomplete sequence of a V-gene.

On the other hand, the TCR libraries preparation with full-length V gene sequencing allows for a more accurate identification of a reference gene during alignment. Moreover, it allows to discover new V gene allelic variants in novel datasets. Using 5'RACE to prepare cDNA libraries (see Fig.4) allows to achieve such accurate full-length sequencing, and to use unique molecular identifiers if needed. Here in the collection of research articles, we used multiplex PCR for sequencing gamma/delta TCRs (Chapters IV and V) and a 5'RACE-based protocol previously developed in Chudakov's lab (also available as kit from Milaboratory LLC) [22].

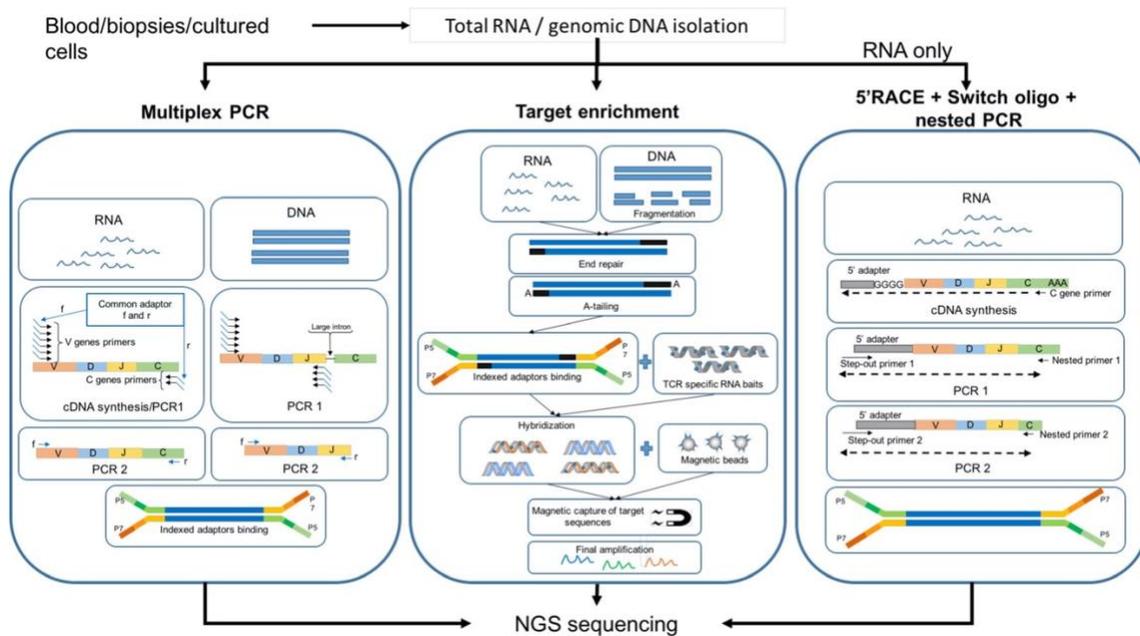


Figure 4. Diversity of molecular approaches for high-throughput sequencing of T cell receptor repertoire libraries. The workflow of three principal methodologies for TCR library preparation are shown: multiplex PCR, basic targeted enrichment and 5'RACE-switch-oligo nested PCR. Multiplex PCR is suitable for both RNA and genomic DNA sequencing. Basic target enrichment includes fragmentation and end-repair of libraries which can affect library recovery. Nested PCR based on the 5'RACE and switch-oligo approach allows UMI incorporation during cDNA synthesis. 5'RACE uses only 1 pair of primers eliminating multiplex PCR bias. Subsequent nested PCRs increases outcome specificity. *Reproduced from [30] with permission.*

In this protocol, each cDNA molecule is marked with a short unique nucleotide sequence at the first stage of the TCR libraries preparation protocol [23, 24, 25]. This allows identifying and removing artificial diversity in the repertoire introduced during library amplification and by sequencing errors [22, 26]. The superior quality of UMI-barcoded data allows using all processed data. On the other hand, other technologies (notably, older iRepertoire protocol) that do not use molecular barcodes usually need to discard lower quartile or even half of the sequencing reads to ensure that the clonotypes observed in the data are “real” and not a result of artificial diversity from sequencing errors [22, 27, 28, 29, 30]. A direct comparison of protocols clearly showed that the UMI-incorporating 5'-RACE approaches resulted in the minimal bias in the V-gene usage compared to other protocols and achieved comparable diversity and sequencing depth [30, **Fig.4**].

Computational analysis of TCR repertoires is the main method used in this dissertation. The alignment and post-processing of T cell receptor sequences were done in accordance with standards in the Genomics of Adaptive Immunity laboratory in Moscow and mostly used software developed in the lab and by Milaboratory LLC. The software is widely used in research and diagnostic/applied R&D laboratories in different countries [18]. Previously developed MiTCR and currently recommended MiXCR software had become the gold standards for immune receptor sequences alignment efficacy and is used in benchmarking [19, 20, 21].

MiXCR software allows to study the alignment in detail and infer VDJ recombination probabilities from specific reference gene segments. For comparative example, the standard output of Immunoseq technology and default data analysis does not provide such information or the intermediate alignment results exporting; the user receives the limited information about clonotypes observed in the sample, and the correct clonality and diversity estimation remains complicated.

To highlight the relevance to author's contribution to original research, it should be mentioned that projects comprising this thesis used a large variety of methods available in the modern laboratory studying molecular immunology. However, in most of the projects, the thesis author's contribution is limited to bioinformatic analysis of TCR repertoires sequencing data. In some projects, the author participated in the experimental part, including healthy donor selection and recruitment, collection and processing of biological samples, flow cytometry antibody panel design, novel sequencing protocols optimization, NGS libraries preparation, and other "wet lab" tasks.

1.4. TCR repertoires profiling in system immunology

Sequencing of immune repertoires is known as RepSeq (repertoires sequencing) or AIRR (adaptive immunity receptors repertoires). This field of immunology, or immunogenomics, generates a large amount of data comparable to genomics and transcriptomics fields in its scale. AIRR promises to be comparable to genomics and transcriptomics in practical importance as well. Genomics is currently expanding in practical applications, including diagnostic confirmation for multiple genetic disorders, and the consumer genomics segment is expanding. Transcriptomics represents an indispensable tool for the majority of molecular biology laboratories focused on fundamental or biomedical research. Immune repertoires sequencing allows us to describe the composition of B cells or T cells immunity in a biological sample, providing the data needed for data-driven decision making in immune-related disorders and corresponding clinical trials; therefore, the AIRR sector's growth is expected in the next decade.

Multiple publications argue that AIRR has a potential to be used in routine diagnostics procedures instead of remaining the status of a fundamental research tool [31, 32, 33]. Therefore,

AIRR data mining will gradually become more useful to create a comprehensive immunome description from a single biological sample derived from a patient. Just recently, an emergency use authorization was issued for the T-Detect test (Adaptive Biotechnologies) for the SARS-CoV-2-specific TCR clonotypes identification from DNA of total peripheral blood T cell samples. This is the first-ever approval of TCRseq analysis as a diagnostic procedure, and this extraordinary event is the primary argument towards prominent changes in the AIRR field [34]. The assignment of function to a T cell receptor sequence remains a complicated problem in the immunology of adaptive immunity: Dr. Shugay and colleagues have developed and maintained the largest database of annotated TCR sequences, but it is not enough to annotate a patient's sample yet [35, 36]. The expanding list of simultaneous identification of transcriptomic state and TCR sequence variants would allow immunologists to speed up the functional annotation of human T cell receptor repertoire [37, 38, 39].

1.5. Individual T cell development and fate

Due to the independent T cell receptor gene rearrangement in each thymocyte, the T cell receptor serves as a barcode or unique identifier for the entire clone of progeny cells of this thymocyte. As a result, when we obtain TCR repertoires sequencing data, we can trace the fate of all descendants from a single naive T cell: progeny could be found at different developmental stages from an early naive T cell that has just exited the thymus after a successful TCR rearrangement (a recent thymus emigrant, or RTE) to a mature naive cell and to memory cell subsets emerging after TCR-pMHC contact and priming.

At most of the developmental stages, T cells require TCR-dependent signaling for survival, selection, or differentiation to the next stage until the terminally differentiated T cell stage or

apoptosis [40]. Testing of T cell receptor specificity at different stages ensures the safety and efficacy of an individual T cell activity in adaptive immunity. Therefore, a functional and structural description of T cell receptor diversity is essential to understand the biology of antigen-specific recognition and T cell-driven immunity regulation. Moreover, understanding the details of T cell selection into functional subsets may promote the development of pharmaceuticals for fine-tuning immune response in patients [15].

T cell receptor gene rearrangement, which was described in 1.1.2, occurs in T cell progenitors at an early stage called DN thymocyte for double-negative, e.g., not expressing either CD4 or CD8 coreceptor. At later DP (double-positive, CD4⁺CD8⁺ cell) and SP (single positive, helper T cell progenitor CD4⁺ or killer T cell progenitor CD8⁺) stages, each T cell progenitor is tested for the functionality of the randomly assembled receptor (positive selection) and the lack of excessively high-affinity host self-antigens (negative selection). The vast majority of thymocytes fail at the thymic selection and undergo apoptosis [14].

Surviving T cells proliferate and exit the thymus into the bloodstream; these are the recent thymic emigrants (RTE), the earliest stage in the naïve T cells [41]. The main function of a naïve T cell is recirculation through major lymph nodes in order to maximize the probability of meeting the cognate antigen through contact with specialized antigen-presenting cells in the T cell zone. In rodents, the naïve T cell pool is maintained with the new portions of naïve T cells being constantly produced in the thymus [42]. In human immunity, however, the naïve T cell pool is believed to be more self-replenishing, which means that naïve T cells slowly proliferate and can survive for multiple years, while the initial thymic naïve T cell production monotonously decreases with age [43, 44, 45].

After meeting with the antigen in the lymph node, the individual T cell's ability to divide

risers dramatically. This initially primed T cell becomes a T_{SCM} - stem cell memory T cell, the earliest memory stage. In line with its name, stem cell memory T cell can give life to various progeny, including long-living central memory cells (T_{CM}), short-lived effector cells that carry out the acute immune response reactions (SLEC or T_{EMRA} cells), and effector memory progenitor cells T_{EM}, which in turn produce T_{EMRA} during division [46]. CD4⁺ T cell subset contains T_{SCM}, T_{CM}, T_{EM} stages, and killer CD8⁺ T cells can belong to T_{SCM}, T_{EM}, and T_{EMRA} differentiation stages. All of these cells leave the lymph node and travel through the blood; activated effector cells can then exit the bloodstream to initiate an immune response in the peripheral tissue close to the pathogen activity location [47].

At the next stages, effector T cells can decide to stay in the peripheral organ and become resident T cells or return to the lymph node-bloodstream constant recirculation. The lymph node-specific chemokines are produced to the lymph node stromal cells and are recognized by the homing receptors CCR7 and CD62L on the T cell surface. Both of these receptors are usually absent on effector cells' surface. Because of this, it has long been a mystery how effector cells can get from the peripheral tissue back to the secondary lymphoid organs - the spleen and lymph nodes, but there is always a second route to enter the lymph node via lymph instead of a blood vessel [47, 48]. Finally, a T cell can die at different stages of its thorny developmental path via intrinsic or instructed apoptosis, long anergy and lack of resources, or via activation-induced death. In most situations, an activated T cell requires T cell receptor signaling for survival, which is called tonic signaling [40, 49, 50, 51]. However, if the cognate foreign antigen is not present in the lymph node or in the tissue, recognizing it T cell clone may not survive for a long time. It is not so clear if TCR signaling is also responsible for the survival and proliferation of the naïve T cell compartment [51]. We decided to investigate if there is a systemic change in the T cell receptor repertoires during

biased survival and peripheral selection among human naïve T cells in Chapter 1.

1.6. TCR role in T cell lineage selection and survival

The understanding of TCR influence on a naïve T cell survival could lead to discoveries of approaches for maintaining the naïve T cell repertoire in aging and therefore prolong adequate adaptive immunity function in elderly patients. The decrease in the number of naïve T cells with age is associated with the rapid aging of the thymus and thymic production of new naïve T cells. Moreover, we thoroughly studied the dynamics of the T cell receptor diversity and found out that TCR repertoire diversity decline is reduced along with the decline in naïve T cell fraction [52, 53]. Both CD4⁺ and CD8⁺ T cells decrease in numbers and TCR diversity in age; we studied more details of TCR repertoires age-related changes in the naïve subsets, specifically in Chapter I of the current thesis. To study the peripheral selection in human naïve T cells, we sorted RTE and mature naïve T cells from healthy donors (Chapter I). We assumed that systemic changes in repertoires would show the strength and extent of peripheral selection and age-related changes.

The research teams from Dr. Chudakov and Dr. Britanova's laboratories have been studying the age-related changes in the T lymphocytes compartment of both human and murine adaptive immunity in the previous years. The total number of T cells is stable throughout a human life after the neonatal period and only slightly decreases with age. But the prominent decrease in the diversity of TCR repertoires is linear and significant, as shown earlier by Dr. Britanova [52, 53]. It is associated with the proliferation of memory T cells and the loss of naïve T cells due to the continuous thymic involution after puberty [42, 44, 48, 53, 54]. Previously, we described the high diversity of T cell receptors in the umbilical cord blood samples and described neonatal immune repertoires' characteristic features [53].

In the aforementioned study of TCR repertoires in human aging [53], we used peripheral blood samples from healthy donors and umbilical cord blood (UCB). The peripheral blood samples are the most available for the human immunology research, whether the research on immunity in organs in tissues requires more complex project planning and samples collection. We cannot infer the rules and patterns of development, selection, or aging in the immune cell subsets occurring in peripheral organs from the limited number of blood samples. From studies made primarily on UCB, other authors and we assumed that newborns have practically no memory T cells, and the T cell repertoire is composed of naive T cells [53, 55, 56]. However, in our next collaboration with Prof. Fritz Köning lab in LUMC, Leiden, the Netherlands, we examined TCR repertoires obtained for T cells from human embryonic intestinal tissue samples. Surprisingly, we found that different populations of memory T cells are generated in a TCR-dependent mechanism quite early in the embryo (see Chapter II), including inflammatory helper T cells and regulatory memory T cells. This phenomenon is difficult to interpret. Memory Treg cells' appearance and the clonal nature of Treg memory may suggest the early mechanisms to develop immune tolerance to the commensal bacteria that will colonize the gut after birth.

We continued to study not only the blood circulating T cells but also those present in peripheral tissues when we had an opportunity in further research. Two different collaborative projects allowed us to study resident gamma/delta T cells in peripheral organs: in the liver (Chapter V) and in breast cancer (Chapter VI).

2. Helper T cell functional subsets diversity

The question of the T cell receptor role in the selection process during the initial T cell priming is more evident than the question of TCR-dependent naïve T cell survival. The TCR-

pMHC interaction is the primary signal for an activated T cell proliferation. However, activation requires the second signal from costimulatory molecules within the same immune synapse and the third signal in the form of cytokines to instruct further T cell differentiation [15, 57].

It is believed, that signal 3 determines the functional state, which would be obtained by a helper T cell after activation. Depending on interleukin composition in the microenvironment, the primed Th0 cell will become either Th1, Th2, or Th17 cell [57-63]. New subsets were added to the list in recent years, such as Th22 and Th9, which are associated with chronic inflammatory conditions [65, 66]. On the other hand, before priming, naive T cells are considered multipotent and able to develop into any of the numerous CD4 T cell phenotypes from the Th0 state. However, thymic regulatory T cells (tTregs) are already differentiated in the thymus and mostly preserve a stable phenotype; the same thymic-born predisposition was shown for other minor T cell subsets such as invariant NKT cells (iNKT) [67-70]. Thymic Tregs are selected based on the high TCR affinity to self-antigen and receive strong survival signaling [68, 72].

We suggested that a single helper T cell could be already predisposed to one of the functional programs before priming in the lymph node, as with thymic Treg cells. Predisposition could result from a combination of transcriptional programs, epigenetic state of a single cell and characteristics of T cell receptor, as well as levels of expression of receptors to differentiation signal, survival signal cytokines, and other factors. Such predisposition could be established in a cell at the different developmental stages:

1. At the T cell progenitor exit from the bone marrow stage;
2. After receiving instructions in the thymus, for example, from antigen-experienced thymus-infiltrating peripheral memory T/B-cells;
3. After extrathymic peripheral selection before encountering a foreign antigen.

To explore the influence of TCR structure on T cell lineage choice, we designed the deep TCR profiling study from major and minor Th subsets in Chapter III. We aimed to extract TCR characteristics specific to functional lineages at the effector cell stage and test if the same characteristics are seen in naïve T cell sorted subsets. Such systemic TCR factors of functional subset choice may help further research an individual's predisposition to inflammatory or autoimmune disorders. Furthermore, in Chapter III, we studied the shared clonotypes in the effector CD4⁺ helper T cell subsets. Shared TCR clonotypes could be regarded of barcodes of T cell transitions from one state to another. However, a common T cell clone may generate progeny with different phenotypes. Thus, the shared TCR sequence between functional subsets can be interpreted with plasticity or common origin explanations. In Chapter III, we showed that the shared sequences are not distributed evenly between pairs of subsets: there are large and common clonal overlaps and there are also subsets with zero overlap in TCR sequences. We believe Chapter III to be the first detailed picture of distinct repertoire features, plasticity, and publicity of TCR repertoires among helper CD4 T cell subsets in human adaptive immunity.

3. Gamma/delta T cells and TCR repertoires

3.1. $\gamma\delta$ T cell biology and unique features.

The second part of this thesis is dedicated to the research on the structure of gamma/delta T cell receptor ($\gamma\delta$ TCR) repertoires. Gamma/delta T cells have diverse functions between innate and adaptive immunity. This subset is a good example of prominent differences between murine and human immunity in circulation and in the peripheral tissues. Gamma/delta T cells carry an alternative assembly of a T cell receptor which consists of gamma and delta TCR chains. This immune cell subset was considered an innate-like population, but it appeared to have adaptive-like lineage development as well in more recent literature [29].

Most of the mature peripheral and circulating $\gamma\delta$ T cells are $CD3^{high}CD4^{-}CD8^{-}$ thus, $\gamma\delta$ TCR recognition is not restricted by MHC molecules [73]. This allows $\gamma\delta$ T-lymphocytes to recognize soluble antigens, as well as non-peptide antigens: phosphoantigens, lipoproteins, lipid, and glycolipid antigens. In human peripheral blood, gamma/delta cells represent a small population of lymphocytes: they make up no more than 15% of T cells in healthy people, usually about 4-11% [74, 75]. Depending on age, gender, and concomitant chronic diseases, the number of gamma/delta T cells in peripheral blood can vary. The absolute number of circulating gamma/delta T cells decreases with age. The decrease occurs due to the disappearance of $V\delta 2^{-}$ T cells (carrying TCR with variable segment 2 (TRDV2⁺) on the delta chain), while $V\delta 1^{+}$ T cells retain their numbers in centenarians [76].

Some $\gamma\delta$ T cells are functionally and phenotypically similar to cytotoxic T cells or NK cells with cytoplasmic granules, a high level of granzymes production, and high cytotoxic potential *in vitro*. Cytotoxicity could either be activated via several mechanisms. 70% of human peripheral blood $\gamma\delta$ T lymphocytes carry the NK cell receptors: NKG2/CD94 dimer on their membrane; 10% express molecules of the KIR family [77, 78]. Activating NKR, antigen-specific $\gamma\delta$ TCR stimulation, superantigen-like mode of TCR activation could lead to the gamma/delta T cell cytotoxicity [79, 80]. The prominent cytotoxic properties of $\gamma\delta$ T cells have attracted the researchers' attention for use in adoptive anticancer cell therapy in recent years. It should be noted that the Prof. Kuball academic collective has advanced this particular applied field significantly [81-83].

$\gamma\delta$ T cells display curious non-classical properties, such as phagocytosis or high expression of TLR receptors and other PRRs compared to conventional T cells [reviewed in 84]. Usually, lymphocytes do not phagocytose, but $\gamma\delta$ T cells are capable of phagocytosis and subsequent

antigen presentation to conventional T cells. Although typically CD16-mediated phagocytosis does not lead to antigen presentation, $\gamma\delta$ T cells are able to effectively present antigens in MHCII molecules, which is unique among T cells [reviewed in 84]. This phenomenon is a non-redundant function of $\gamma\delta$ T cells and can fit as an element in the plethora of not fully studied phenomena of T cell - T cell interactions, along with the killer T cell licensing, quorum sensing decisions among T cells, Treg-helper T cell interactions, and more [85,86].

3.2. $\gamma\delta$ T cell subsets in inflammation.

Other properties bring gamma/delta T cells closer to helper T cells: these lymphocytes are capable of producing a range of proinflammatory and regulatory cytokines and growth factors. Just as helper T cells are divided into functional subpopulations based on the cytokine production profiles and the general direction of the immune response, there are functional subsets among $\gamma\delta$ T cells. Mainly two functional subsets are described: $\gamma\delta$ T17 and IFN γ -producing Th1-like $\gamma\delta$ T cells. $\gamma\delta$ T cell-produced cytokines contribute to dendritic cell maturation and to remodeling of stromal cells of lymphoid organs [87-90]. In barrier and mucosal tissues, resident $\gamma\delta$ T cells affect the level of production of the IgE antibodies [91-93].

In a number of inflammatory diseases and mouse models of infections, $\gamma\delta$ T lymphocytes have been shown to be an essential source of proinflammatory cytokines. Pathological inflammatory effects of $\gamma\delta$ T-promoted inflammation were described in a mouse model of autoimmune encephalomyelitis [94-96]. However, it should be noted that at the later stages of the autoimmune model of encephalitis, $\gamma\delta$ T cells rather play a protective role, despite the high production of the inflammatory agent IL-17, while mice lacking $\gamma\delta$ T17 subset die faster of encephalitis [97,98].

On the other hand, protective role of $\gamma\delta$ T-promoted inflammation is described in

mycobacterial infections. $\gamma\delta$ T cells are considered the primary line of defense and the activators of neutrophils via IL-17 production [99]. The same mechanism is activated in BCG vaccination: $\gamma\delta$ T cells show the most significant reaction to vaccination among all human peripheral blood lymphocyte subsets, with an expansion peak of occurring on day 7 after vaccination, compared with 30 days for alpha/beta T cells [100, 101].

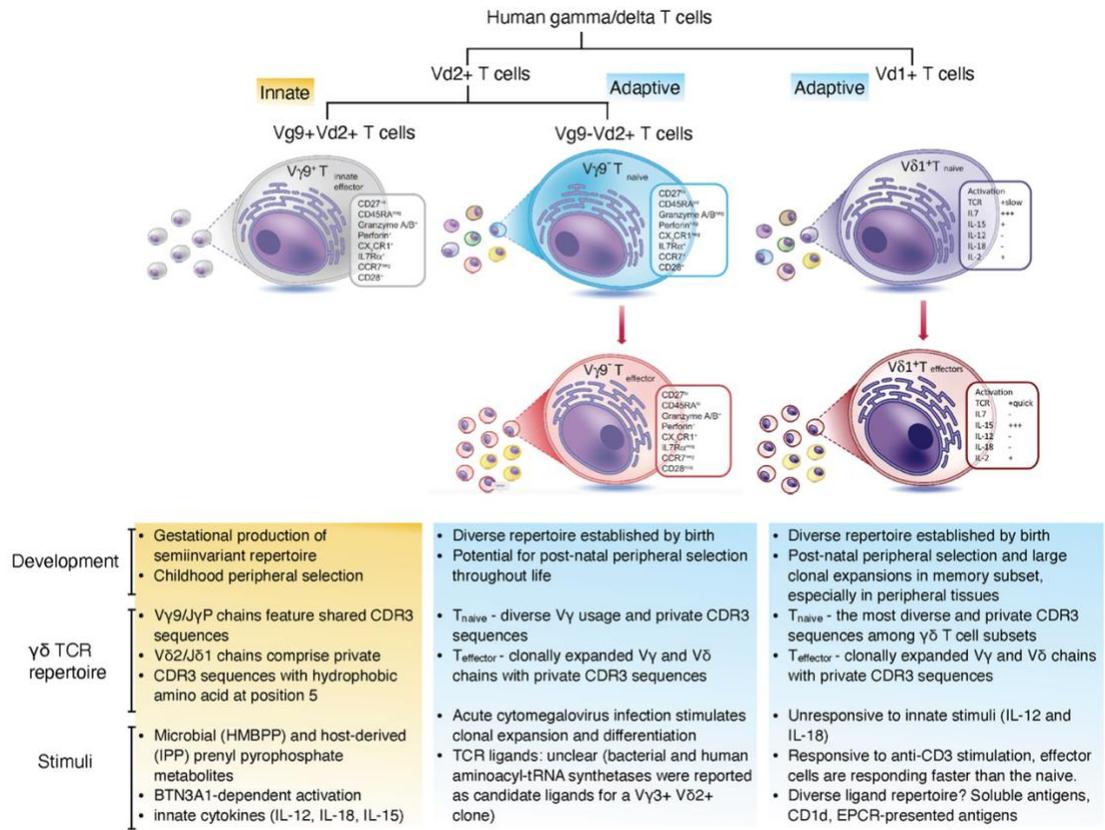
In the experimental infection of macaques with Calmette-Guérin bacilli, it appeared that $\gamma\delta$ T cells react in an oligoclonal way. Reactive $\gamma\delta$ T cell clones were public among macaques' samples and contained the V γ 9 gene segment in the gamma chain. The V γ 9+ $\gamma\delta$ T cells are abundant in primates, but there is no analogous subset in mice. Therefore, the model of tuberculosis infection or BCG vaccination in monkeys appeared much more relevant than rodent models, taking into account the contribution of $\gamma\delta$ T cells to specific immune responses [102].

3.3. $\gamma\delta$ T cell subsets. *VDelta1+* and *VDelta2+* subsets.

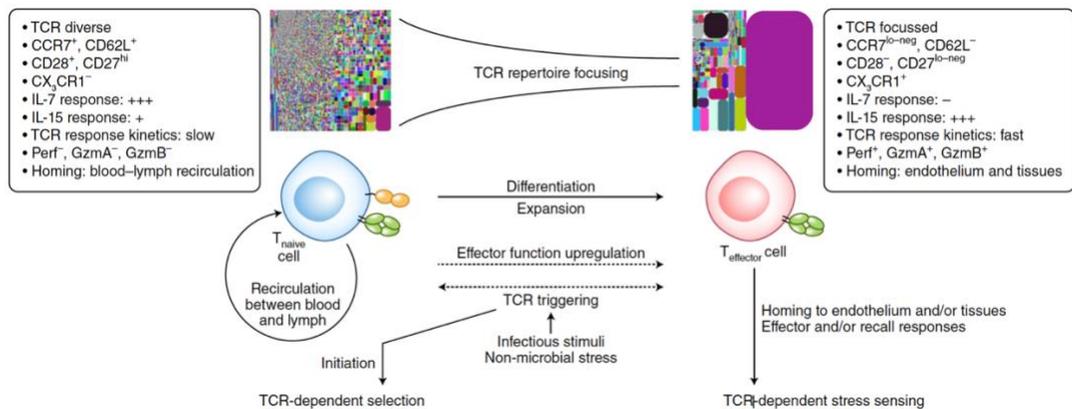
Among gamma/delta T cells in peripheral blood, special attention is paid to the major subset V γ 9+V δ 2+ cells. These are cells with a T cell receptor containing variable gene segments 9 and 2 in rearranged gamma and delta chain genes, respectively (IMGT TCR gene nomenclature). This subset is predominantly found in human peripheral blood in the state of effector cell phenotype and is assigned to the group of innate-like lymphoid cells. Only V δ 2+ but not V δ 1+ cells express NK cell receptors and KIRs on their surface [103, 104]. In general, four subsets of V γ 9+V δ 2+ cells can be distinguished by markers of immune memory CD27 and CD28: long-lived central memory cells T_{CM} (CD27⁺CD28⁺), T_{CM27}⁻ (CD27⁻CD28⁺), effector T_{EM} cells (CD27⁺CD28⁻), and, finally, a subpopulation of T_{EMRA} (CD27⁻CD28⁻), for which the leading role in the (perforin/granzyme B-mediated) cytotoxicity is assumed. A number of published data sources assign the most tremendous cytotoxic potential to this T_{EMRA} V γ 9+V δ 2+ among all $\gamma\delta$ T

cells. The use of a monoclonal antibody for activation of this $\gamma\delta$ T cell subset is currently in a clinical trial for immunotherapy.

In collaboration with the group led by Prof. Willcox from the University of Birmingham, we decided to study how the immune memory of $\gamma\delta$ T cells is formed. Is this process similar to the clonal selection of conventional $\alpha\beta$ T cells (see **Fig.5**)? Do $\gamma\delta$ T cells enter the memory T cell pool when activated through NK cell receptors or when $\gamma\delta$ TCR is triggered with antigen interaction? We succeeded in describing the clonal expansion and memory generation in V δ 1+ subset of $\gamma\delta$ T cells in a high-impact publication [29]. Here we suggested that the V δ 1+ subset should be considered an adaptive immune cell subset [reviewed in 105]. We continued our collaborative research with the next question of adaptive and innate-like characteristics in V δ 2+ subset of $\gamma\delta$ T cells which is presented in the Chapter IV of this thesis and already discussed in major field reviews [106, 107].



A



B

Figure 5. Adaptive and innate-like subsets of gamma/delta T cells. A) Major features of innate subset Vγ9+Vδ2+ and adaptive subsets Vγ9-Vδ2+ and Vδ1+ studied phenotypically and by the TCR repertoire analysis in [29] and Chapter IV. Adaptive γδ T cell subsets develop from naïve T cells to distinct memory T cells with some phenotypical resemblance to conventional cytotoxic T cells. *Adapted from [105] and [106].* B) Phenotypical and biological changes in the adaptive γδ T cell populations along ligand engagement and selection from naïve into memory T cell subset. These apply to both Vδ1+ (more generally Vδ2-) T cells and Vγ9-Vδ2+ T cells. Note the upregulated cytotoxic and cytokine production potential, and enhanced homing to peripheral tissues. Perf - perforin; GzmA - granzyme A; GzmB - granzyme B. *Adapted from [107].*

3.4. Resident memory T cells' TCR repertoires

As mentioned above, chapters V and VI focus primarily on resident memory T cells (T_{RM}) and their T cell receptor repertoires. Where do resident tissue cells first come from? These are the descendants of effector cells that have lost their ability to recirculate and are present in the tissue in a steady state for months or years (**fig. 6**).

Some tissues allow the effector memory T cells to enter freely, and T cells are present there in large quantities. Such T cell-rich tissues are the intestinal mucosa, peritoneal cavity, lung-associated mucosal tissue, etc. Other organs allow a very limited influx of effector T cells and only during the local or systemic inflammatory reaction. The tissues of the second type include those separated from the immune system with a barrier, for example, the brain and spinal cord, as well as many others: peripheral ganglia, mucous membranes of the genital organs, lungs, epidermis, eyes [108, 110]. The difference between the two types of tissues is in the expression of additional homing molecules, which facilitate the entrance for effector T cells, for example, adhesion molecules for penetration into the epithelium MadCAM-1 [109]. Once established in the tissue, T_{RM} helper cells perform immune surveillance of the tissue and regulate its homeostasis.

T_{RM} patrol the tissue, form contacts with tissue-resident macrophages and influence the local microenvironment through pro- and anti-inflammatory cytokines and growth factors, thereby participating in tissue defense and homeostasis, and renewal [110]. Understanding how resident T cells work is essential to fight infections that do not enter the bloodstream immediately but enter the body through barrier tissues - that is, for the vast majority of infections. The rational design of vaccines to protect against this group of infections can and should be aimed precisely at enhancing the first stage of protection with the help of resident immune cells. In an ideal situation, adequate activation of resident antigen-specific cells eliminates the pathogen in the barrier tissue without

triggering the systemic inflammation. Such localization of immune response is beneficial for safety from adverse immune reactions and should be a goal in the next generation of mucosal immunity-oriented vaccines [110-112].

The repertoire of T cell receptors of cells resident in various mucosal membranes is considered partially degenerate and public, that is, identical for many people in the population [10, 11, 114]. However, the accumulated data on TCR repertoires of T_{RM} cells is still very limited; therefore, the notion of publicity of tissue-specific TCR repertoires should be taken with caution [115-117]. There is a need for deeper profiling of human tissue samples to describe the actual tissue-resident T cell repertoire.

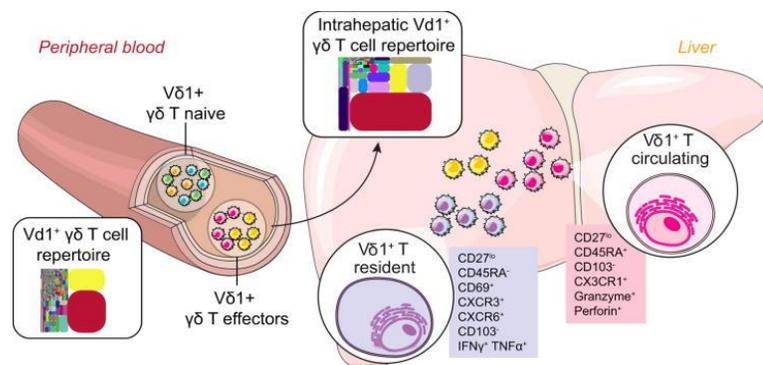


Figure 6. Tissue-resident T cells display a specific TCR repertoire private to an individual. The example of resident gamma/delta T cells in human liver samples shows us the unique potential of this subset to generate unusually large clonal expansions without any public clonotypes involved. Intrahepatic TCR repertoire show little overlap with peripheral blood clonotypes. Figure from Chapter V graphical abstract.

In Chapter V, we describe the first profiling of gamma/delta T cells from human liver samples (**fig. 6**) and perform the comparative analysis of the tissue- and blood-derived TCR clonotypes, indicating a specific novel liver-resident T cell subset. In Chapter VI, we compare the normal blood-derived $\gamma\delta$ TCR repertoire to the $\gamma\delta$ TCR obtained from the biopsies of breast cancer patients to assess the private and public components in the tissue-specific immunity and in cancer surveillance in particular. Determining the patterns of T-lymphocyte homing in certain tissues can provide an advantage in cellular immunotherapy of tumor diseases. In theory, cytotoxic T cells of

the desired specificity for the tumor antigen-activated *in vitro* should kill the patient's tumor cells. However, in practice, such immunotherapy is complicated by the fact that tumor cells are capable of suppressing immune responses, and killer T cells approaching the tumor become anergic. Often, anergic T-lymphocytes, primarily from the T_{RM} pool of this tissue, accumulate in the mass of a growing tumor and around it [118-121]. Of the many active tumor-specific T cells injected into a patient, few will reach the target, and even they may be practically useless in the immunosuppressive tumor microenvironment.

Of course, the data on TCR structure or even TCR specificity is not enough to effectively manipulate or even understand the T cell effector reactions in the tissues. It is necessary to study in detail the factors influencing the tissue-specific T cell subsets and clonal specificities in order to understand the mechanisms of activation of local tissue immunity and induction of T_{RM} tolerance. We need to study the mechanisms of how the local tissue-specific niches for T_{RM} cells are organized and how they are filled up primarily in the early postnatal development of immunity. This initial resident immunity settlement occurs before the child meets a large number of pathogens and, accordingly, before forming a significant pool of effector memory T cells (precursors for the T_{RM} cells). There is also an adjacent question of primary tolerance induction in the barrier tissues and the balance between tolerogenic and protective immunity in tissues.

To answer these questions, we need to go beyond the scope of this thesis. There is a need for simultaneous identification of paired TCR sequence and a T cell's functional phenotype, e.g., TCR-seq and RNA-seq paired datasets. Further development of these technologies will speed up and simplify the development of effective and tissue-specific T cell adoptive immunotherapy.

Aim of study:

The aim of my Ph.D. thesis project was to study the T cell receptor repertoire landscapes across human helper T cell subsets, explore the distinct patterns of TCR-driven selection and the generation of functional memory and early naive T cell subsets, in both conventional and non-classical immune cells.

Specific set of goals was set in the course of Ph.D. thesis research:

1. Describe the TCR landscape in the healthy donors' cohort and explore the peripheral selection in the naive T cells;
2. Study the earliest memory T cell development in tissue-resident fetal immunity (Chapter II);
3. Investigate the role of TCR structure in the functional effector helper T cell subsets, along with transitions between functional subsets (Chapter III);
4. Explore the naive and memory states in the gamma/delta T cell subset development (Chapter IV)
5. Discover the biology of gamma/delta T cell selection and memory formation in tissue resident subset (Chapter V);
6. Explore the role of gamma/delta T cells with specific TCR structure in the cancer immunosurveillance (Chapter VI).

Results

Chapter I

The changing landscape of naive T cell receptor repertoire with human aging

Introduction

Chapter I studies the selection and survival of T cell clones of naive T cell pool across functional subsets and donor age groups. We asked how T cells are selected in the periphery after they exited thymus: which clones survive, proliferate, and persist for long years and which clones are eliminated quickly after thymus emigration? We performed a comparative analysis of TCR repertoires of recent thymic emigrants and mature naive cells in the CD4 subset. Moreover, we followed naive CD4 and CD8 TCR repertoires changes between young and older donors. We showed in this chapter that the peripheral selection is difficult to catch, as recent thymic emigrants (RTE) and more mature naive T cells display similar TCR repertoires. At the same time, RTEs are more prone to aging. Aging patterns in TCR repertoires include shortened and less diverse CDR3 regions of TCRs decreased predicted affinity of interaction, and changes in physicochemical characteristics. Public repertoire fraction increased was observed in CD4 but not the CD8 TCR repertoire with aging.

Contribution

This study utilizes two datasets of TCR repertoires from young and old donors; each contains naive and memory CD4 T cells. The first dataset was produced in our laboratory, whether the second was previously published, and data was made publicly available.

For the first dataset acquisition, I designed the set of molecular markers and a panel of antibodies for flow cytometry-based sorting of T cell subsets of interest. I organized the sample collection from a specific donor cohort and prepared samples for flow cytometry. I performed the DNA libraries preparation for sequencing, performed primary data processing, and participated in data analysis, writing, and discussion of this article. Dr. Evgeny Egorov, an equally contributing author, performed flow cytometry sorting, participated in the data analysis and wrote the manuscript with my and other authors' help.



The Changing Landscape of Naïve T Cell Receptor Repertoire With Human Aging

Evgeny S. Egorov^{1†}, Sofya A. Kasatskaya^{1,2†}, Vasily N. Zubov¹, Mark Izraelson¹, Tatiana O. Nakonechnaya¹, Dmitriy B. Staroverov¹, Andrea Angius³, Francesco Cucca³, Ilgar Z. Mamedov¹, Elisa Rosati⁴, Andre Franke⁴, Mikhail Shugay^{1,2}, Mikhail V. Pogorelyy¹, Dmitriy M. Chudakov^{1,2*} and Olga V. Britanova¹

¹Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia, ²Center of Life Sciences, Skolkovo Institute of Science and Technology, Moscow, Russia, ³Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Monserrato, Italy, ⁴Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany

OPEN ACCESS

Edited by:

Benny Chain,
University College London,
United Kingdom

Reviewed by:

Sian M. Henson,
Queen Mary University of London,
United Kingdom

Encarnita Mariotti-Ferrandiz,
Université Pierre et Marie
Curie, France

*Correspondence:

Dmitriy M. Chudakov
chudakovdm@mail.ru

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted
to T Cell Biology,
a section of the journal
Frontiers in Immunology

Received: 27 April 2018

Accepted: 29 June 2018

Published: 24 July 2018

Citation:

Egorov ES, Kasatskaya SA,
Zubov VN, Izraelson M,
Nakonechnaya TO, Staroverov DB,
Angius A, Cucca F, Mamedov IZ,
Rosati E, Franke A, Shugay M,
Pogorelyy MV, Chudakov DM and
Britanova OV (2018) The Changing
Landscape of Naïve
T Cell Receptor Repertoire With
Human Aging.
Front. Immunol. 9:1618.
doi: 10.3389/fimmu.2018.01618

Human aging is associated with a profound loss of thymus productivity, yet naïve T lymphocytes still maintain their numbers by division in the periphery for many years. The extent of such proliferation may depend on the cytokine environment, including IL-7 and T-cell receptor (TCR) “tonic” signaling mediated by self pMHCs recognition. Additionally, intrinsic properties of distinct subpopulations of naïve T cells could influence the overall dynamics of aging-related changes within the naïve T cell compartment. Here, we investigated the differences in the architecture of TCR beta repertoires for naïve CD4, naïve CD8, naïve CD4⁺CD25⁻CD31⁺ (enriched with recent thymic emigrants, RTE), and mature naïve CD4⁺CD25⁻CD31⁻ peripheral blood subsets between young and middle-age/old healthy individuals. In addition to observing the accumulation of clonal expansions (as was shown previously), we reveal several notable changes in the characteristics of T cell repertoire. We observed significant decrease of CDR3 length, NDN insert, and number of non-template added N nucleotides within TCR beta CDR3 with aging, together with a prominent change of physicochemical properties of the central part of CDR3 loop. These changes were similar across CD4, CD8, RTE-enriched, and mature CD4 subsets of naïve T cells, with minimal or no difference observed between the latter two subsets for individuals of the same age group. We also observed an increase in “publicity” (fraction of shared clonotypes) of CD4, but not CD8 naïve T cell repertoires. We propose several explanations for these phenomena built upon previous studies of naïve T-cell homeostasis, and call for further studies of the mechanisms causing the observed changes and of consequences of these changes in respect of the possible holes formed in the landscape of naïve T cell TCR repertoire.

Keywords: aging, T cell receptor, naïve T cells, immunosequencing, Rep-Seq, CDR3 repertoire

INTRODUCTION

A diverse set of naïve T cell functions (1) and their antigenic receptors—T-cell receptors (TCRs) (2, 3)—protects us from a multitude of infectious and cancer hazards encountered throughout our lifespan. Furthermore, it essentially provides selection of the appropriate amplitude, type, localization, and duration of immune response. Human aging is associated with profound changes in T cell

immunity (2–5), compromising our ability to withstand novel pathogens and manage chronic infections. It also dampens the effect of vaccination (6–8) and can lead to higher cancer susceptibility (9–12). These changes may further result in an imbalanced immune response that can develop into non-specific inflammation, provoking neurodegenerative and cardiovascular disorders, and to the loss of tolerance, leading to autoimmunity (3, 13–15). For the latter, a reduction of regulatory T cell (Treg) diversity (16, 17) could be a one of the causative factors.

With aging, accumulating clonal expansions of memory T cells caused by previously encountered antigens gradually begin to dominate in the available T-cell pool. This leads to a homeostasis characterized by a decreased number of naïve T cells, essentially shrinking the precious reservoir of diverse functions and antigenic specificities (2, 18–23). At the same time, thymus function progressively declines after puberty (24, 25), and drops sharply to a very low level after 40 years of age (4, 26). Along with diminished production of T cell progenitors by the bone marrow (27), this leads to a drop in generation of the so-called recent thymic emigrants (RTE)—the not fully mature (28, 29) form of naïve T cells, and thus in the replenishment of the mature naïve T cell pool (5, 26, 30).

The existing naïve T cells may still support their abundance and diversity for a prolonged period. In humans, both mature naïve T cells and—to a lesser extent—RTE-enriched CD45RA⁺CD31⁺ subset of CD4 T cells (30)—keep ability to proliferate on the periphery (31, 32). However, the number of allowed divisions is not unlimited. Prominent shortening of telomeres is observed in both CD31⁺ and CD31⁻ subsets (30) which eventually leads to a gradual, later avalanche, exhaustion of proliferation capacity and depletion of the naïve T cell pool (20, 33). Additionally, prolonged peripheral proliferation could also be associated with the functional deficiency of naïve T cells that fail to differentiate toward memory phenotype upon a specific antigenic challenge (3), although a recent cytokine profile study suggests that naïve T cells derived from elderly individuals retain their functionality and naiveté (26).

How uniform is the naïve T cell proliferation on the periphery remains questionable. Qi et al. demonstrated that both CD4 and CD8 naïve T cells gated as CCR7⁺CD45RA^{high}CD28⁺ gain clonal expansions by the age of 70–85 years (34). This observation suggests that some of the naïve T cell clones are dividing more prominently than others. Furthermore, the most rapidly dividing ones could exhaust and extinguish more rapidly, while those dividing with a moderate rate could form the observed clonal expansions.

Importantly, the peripheral T cell proliferation may be dependent on the so-called “tonic signaling”—recognition of MHC complexes loaded with self antigens while surveying the peripheral lymphoid organs. Such contacts are transient and do not lead to classic T cell activation, but generate sub-threshold signals required for naïve T cell survival and proliferation (35–38).

The desirable (i.e., required to efficiently recognize foreign antigens within MHC) and allowed (i.e., not leading to self-recognition and autoimmunity) TCR affinity to self peptide–MHC complexes is set in the course of positive and negative thymic selection, respectively. The threshold range of such selection is not that narrow, thus naïve T cells that leave the thymus—initially

as RTE—have a relatively wide range of self-reactivity. The produced pool of naïve T cells is, therefore, subjected to varying degrees of tonic TCR signaling (38). Therefore, peripheral proliferation of naïve T cells could be potentially biased toward preferential exhaustion of naïve T cell clones carrying TCRs with the highest affinity to MHC. Furthermore, naïve T cells bearing high affinity TCRs could also serve as a preferential source of antigen-responding clones (37) thus being the first one to transit from the naïve T cell pool.

Another factor that could contribute to the dynamics of naïve TCR repertoire landscape is the fate of the specific population of T cells produced in fetal period. We have earlier demonstrated that this subset may survive for decades and contribute to adult TCR repertoire (39). Their TCRs are characterized by a low number of nucleotides that are randomly added by TdT enzyme in the course of VDJ recombination (40, 41). Furthermore, these cells originate from a distinct population of hematopoietic stem cells and are characterized with generally higher proliferation potential (42). However, their fate among other naïve T cells in the elder age remains unexplored.

Altogether, there are number of factors that could shape the landscape of naïve T cell TCR repertoire with aging. To shed light on the nature of ongoing changes, we have focused on the comparative analysis of intrinsic characteristics of the TCR repertoires for the bulk naïve CD8⁺, bulk naïve CD4⁺, naïve RTE-enriched CD31⁺CD4⁺, and naïve non-RTE CD4⁺ T cells derived from the peripheral blood of young versus elderly healthy donors, demonstrating that

- 1) Characteristics of TCR beta CDR3 repertoires change in both CD4 and CD8, both RTE-enriched and mature naïve CD4 T cell subsets with age.
- 2) Within the same age group, no significant difference is observed in characteristics of TCR repertoire between RTE-enriched and mature naïve CD4 T cell subsets.
- 3) TRBV and TRBJ gene segment usage also changes prominently and similarly both within RTE-enriched and mature naïve CD4 T cell subsets of different individuals.
- 4) Relative “publicity” (i.e., sharing between individuals) of TCR repertoires grows both within RTE-enriched and mature naïve CD4 T cell subsets with age.

The observed changes suggest functional differences of young versus middle-age/old naïve T cell TCR repertoires with respect of potential range and characteristics of recognized antigens.

MATERIALS AND METHODS

Donors and Cell Sorting

The study was approved by the local ethics committee and conducted in accordance with the Declaration of Helsinki. All donors were informed of the final use of their blood and signed an informed consent document. The cohort included 18 healthy individuals aged 25–88 years. Individuals with previously diagnosed cancer or autoimmune disease were excluded. Peripheral blood (10–20 ml) was collected into a number of EDTA-treated Vacutainer tubes (BD Biosciences, Franklin Lakes,

NJ, USA), PBMCs extracted using Ficoll-Paque (Paneco, Kirov, Russia) density gradient centrifugation with SepMate™ tubes (STEMCELL Technologies, Vancouver, BC, Canada), and stained according to manufacturer's recommendations. Following antibodies were used: CD3-eFluor450 (eBioscience, clone UCHT1), CD45RA-FITC (eBioscience, clone JS-83), CD27-PC5 (Beckman Coulter, clone O323), CD4-PE (Beckman Coulter, clone 13B8.2), CD25-eFluor450 (eBiosciences, clone BC96), and CD31-PC7 (eBiosciences, clone WM59). T cells of interest were sorted using FACS Aria III (BD Biosciences, Franklin Lakes, NJ, USA), directly in 350 µl of RLT buffer (Qiagen) per 100,000 sorted cells. Total RNA was further isolated using RNeasy Micro kit (Qiagen) and completely used for TCR library preparation. 5'-RACE TCR beta cDNA libraries were prepared according to the previously described protocol (43, 44). See also: <https://github.com/repseqio/protocols/blob/master/Human%20TCR%20alpha%20and%20beta%20RNA-based%20RACE%20protocol.md>.

Libraries were sequenced with Illumina HiSeq 2000/2500, paired-end 150 + 150 nucleotides.

TCR Beta Repertoires Profiling and Data Analysis

T-cell receptor beta CDR3 repertoires were extracted using MiXCR software (45), version v2.1.5. Decontamination from memory T cell TCR beta clonotypes and comparative post-analysis were performed using VDJtools software v1.1.7 (46).

Resulting decontaminated TCR beta CDR3 repertoires are available from Figshare:

https://figshare.com/articles/Naive_CD4_CD8_subsets/6548921;
https://figshare.com/articles/naive_RTE_and_non-RTE_CD4_T_cells_subsets/6549059.

The obtained repertoires were further filtered to eliminate out-of-frame and stop codon-containing TCR beta CDR3 variants. Averaged physicochemical properties of amino acid residues in the middle portion (5 amino acid residues) of TCR beta CDR3 were calculated using VDJtools, the following metrics were used: strength (47, 48), hydrophathy, polarity, and volume (values available from: http://www.imgt.org/IMGTeducation/Aide-memoire/_UK/aminoacids/IMGTclasses.html). During calculation, property values were weighted by the frequency of corresponding clonotypes, so the results favor more frequent clonotypes and do not depend on the sequencing/sampling depth (49). See **Table 1** for the values used for each amino acid property. See **Tables 2** and **3** for the counts of sorted T cells, the number of CDR3 containing sequencing reads, and the number of unique TCR beta CDR3 clonotypes in each sample.

Statistical Analysis

For comparison of repertoire properties, one-sided *t*-test with unequal variances (Welch's test) was used. Normality of the distribution of sample means was confirmed by performing Shapiro–Wilk tests, and the decision to reject the null hypothesis was made after adjusting for multiple hypothesis testing as in Benjamini–Hochberg. False discovery rate in normality testing was controlled at a level of 0.05 by setting *p*-adjusted upper bound

TABLE 1 | Values used for CDR3 amino acid properties calculation by VDJtools.

Amino acid	Hydrophathy	Polarity	Volume	Strength
A	1.8	0	67	0
C	2.5	0	86	1
D	-3.5	1	91	0
E	-3.5	1	109	0
F	2.8	0	135	1
G	-0.4	0	48	0
H	-3.2	1	118	0
I	4.5	0	124	1
K	-3.9	1	135	0
L	3.8	0	124	1
M	1.9	0	124	1
N	-3.5	1	96	0
P	-1.6	0	90	0
Q	-3.5	1	114	0
R	-4.5	1	148	0
S	-0.8	1	73	0
T	-0.7	1	93	0
V	4.2	0	105	1
W	-0.9	0	163	1
Y	-1.3	1	141	1

at 0.05. *Z*-score normalization was performed by subtracting the mean value for each TRBV gene segment values and dividing by the SD. Only highly represented TRBV gene segments TRBV9, TRBV7–9, TRBV7–2, TRBV6–5, TRBV29–1, TRBV20–1, and TRBV12–3/12-4, each associated with at least 2% of CDR3 clonotypes in each sample, were taken into analysis.

RESULTS

TCR Repertoires of Both CD4 and CD8 Naïve T Cells Change Properties With Aging

To analyze how the properties of naïve TCR repertoire change with age, we first sorted CD3⁺CD4⁺CD27^{high}CD45RA^{high} and CD3⁺CD4⁺CD27^{high}CD45RA^{high} T cell subsets gated as shown on **Figure 1** from peripheral blood samples of 4 young (25–35 years old) and 7 middle-age/old (51–88 years old) healthy donors (**Table 2**). TCR beta profiling was performed as described in Ref. (43), extraction of CDR3 repertoires was performed using MiXCR (45). To exclude possible contaminations from memory T cell pool that could occur during cell sorting, we also performed TCR beta repertoires profiling for memory T cells sorted from the same donors (**Figure 1**). Naïve TCR beta repertoires were further decontaminated from the clonotypes present in memory subsets using VDJtools “Decontaminate” module with default parameters (20:1 parent-to-child clonotype frequency ratio for contamination filtering). This procedure eliminated from 0.005 to 0.5% of reads and from 0.01 to 0.7% of clonotypes, these numbers did not depend on the donor age group. Despite the low proportion of eliminated reads and clonotypes, such procedure is desirable for accuracy of the whole analysis and general control for cell contamination during sorting.

We complemented our data with multiplex PCR RNA-based TCR profiling data from Ref. (34) for 4 young (20–35 years) and 5 old (70–85 years) healthy donors naïve T cells gated as CD4⁺

TABLE 2 | CD4 and CD8 naïve and memory cell sorting.

Donor	Age	Group	Subset	Replica	Number of sorted cells	Number of CDR3 reads	Number of CDR3 clonotypes
Donor 1	25	Young	CD4	1	500,000	528,824	38,614
				2	500,000	475,392	240,80
			CD8	1	500,000	564,227	26,988
				2	500,000	648,821	19,338
Donor 2	26		CD4	1	135,000	1,069,588	45,098
			CD8	1	251,000	633,854	28,967
Donor 3	35		CD4	1	200,500	380,253	43,526
				2	231,000	223,661	23,038
			CD8	1	350,500	533,091	47,626
Donor 4	32		CD4	1	560,000	17,494,455	165,184
				2	488,000	12,525,047	113,333
				3	395,000	6,077,772	54,546
				4	100,000	4,353,149	24,582
			CD8	1	510,000	9,219,066	101,096
				2	316,000	10,594,890	61,022
Donor 5	51	Old	CD4	1	495,000	13,765,788	183,873
				2	425,000	5,146,822	64,134
				3	437,000	3,832,144	84,458
				4	255,000	4,773,131	90,459
			CD8	1	509,000	7,369,930	125,339
				2	200,000	2,689,594	64,907
				3	343,000	517,265	20,870
				4	120,000	1,945,472	32,259
Donor 6	88		CD4	1	500,000	779,559	65,515
			CD8	1	60,000	142,639	9,991
Donor 7	51		CD4	1	100,000	28,129	5,033
				2	100,000	37,556	5,455
				3	100,000	41,380	6,762
			CD8	1	105,000	51,463	7,434
Donor 8	82		CD4	1	100,000	46,472	8,318
				2	90,000	38,109	6,694
			CD8	1	12,000	6,013	846
Donor 9	55		CD4	1	100,000	82,691	7,768
				2	100,000	27,215	4,740
			CD8	1	100,000	51,742	5,290
				2	100,000	84,532	7,666
Donor 10	77		CD4	1	254,000	677,587	24,652
				2	227,000	963,394	20,087
			CD8	1	50,000	942,430	15,033
				2	105,000	553,867	19,180
					1	101,000	242,248

Donors, replicas, sorted cell counts, and number of extracted T-cell receptor beta CDR3 clonotypes are shown.

CCR7⁺CD45RA^{high}CD28⁺ and CD8⁺CCR7⁺CD45RA^{high}CD28⁺. Repertoire extraction was performed using the same MiXCR version starting from raw data (dbGaP, www.ncbi.nlm.nih.gov/gap, accession no. phs000787.v1.p1). Similarly, we used memory subsets from the same donors in order to decontaminate naïve T cell repertoires from possible contaminations during cell sorting using VDJtools.

Analysis of the normalized Shannon–Wiener diversity index for the joint data confirmed the conclusion by Qi and coauthors

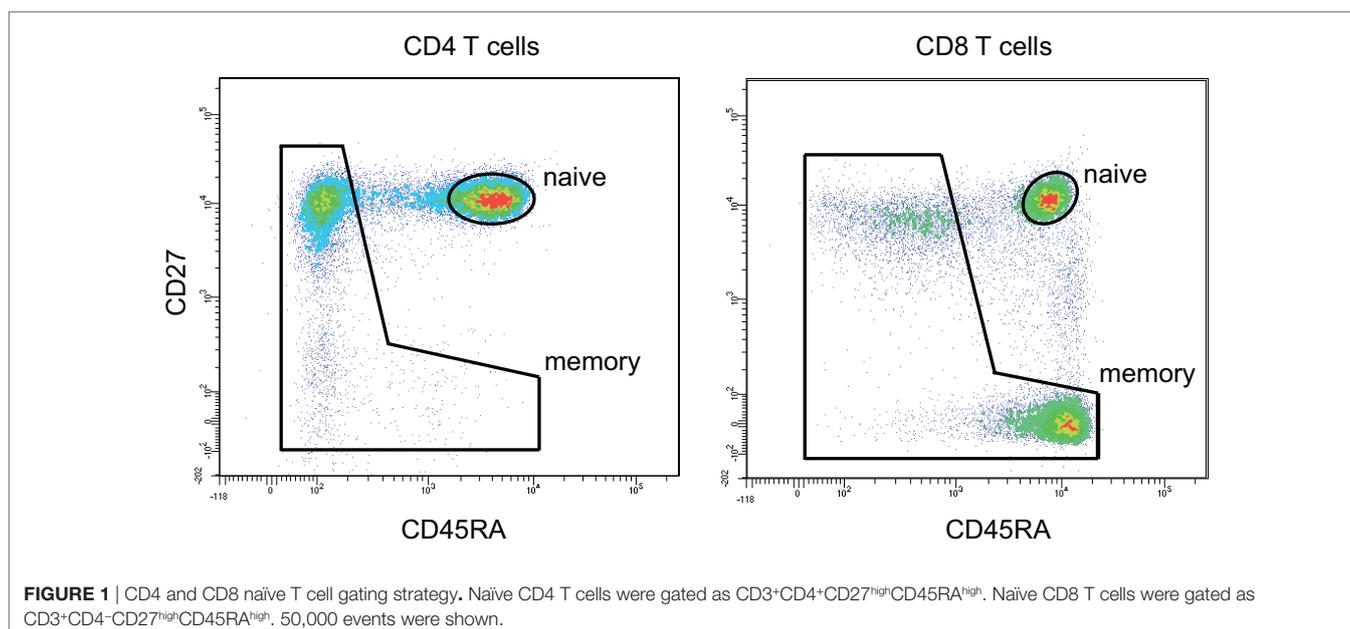
that both CD4 and CD8 naïve T cells accumulate clonal expansion with aging (**Figure 2A**). The accuracy of the results for young individuals generally confirmed the validity of combining the data from both experiments, in spite of the fact that different gating was used for the naïve T cell sorting in the two studies.

Multiplex PCR employed in Qi et al. (34) may cause quantitative biases due to the differing efficiency of primers used to amplify different TRBV segments (50, 51). However, such source of bias does not influence the relative frequency of

TABLE 3 | Recent thymic emigrants (RTEs)-enriched and mature naïve CD4 T cell sorting.

Donor	Age	Group	Subset	Replica	Number of sorted cells	Number of CDR3 reads	Number of CDR3 clonotypes
Donor 12	29	Young	RTE	1	50,000	251,199	27,208
			non-RTE	1	50,000	939,999	33,389
Donor 13	28		RTE	1	100,000	282,998	27,895
			non-RTE	1	100,000	620,320	34,092
Donor 14	31		RTE	1	50,000	144,571	30,139
			non-RTE	1	50,000	622,585	29,253
Donor 15	30		RTE	1	69,000	309,070	33,233
			non-RTE	1	100,000	2,844,397	63,900
Donor 7	51	Old	RTE	1	100,000	14,572	5,306
			non-RTE	1	105,000	17,926	4,030
Donor 8	82		RTE	1	38,000	16,033	2,835
			non-RTE	1	100,000	29,970	6,244
Donor 9	55		RTE	1	89,000	19,622	3,664
			non-RTE	1	189,000	20,891	5,290

Donors, replicas, sorted cell counts, and number of extracted T-cell receptor beta CDR3 clonotypes are shown.



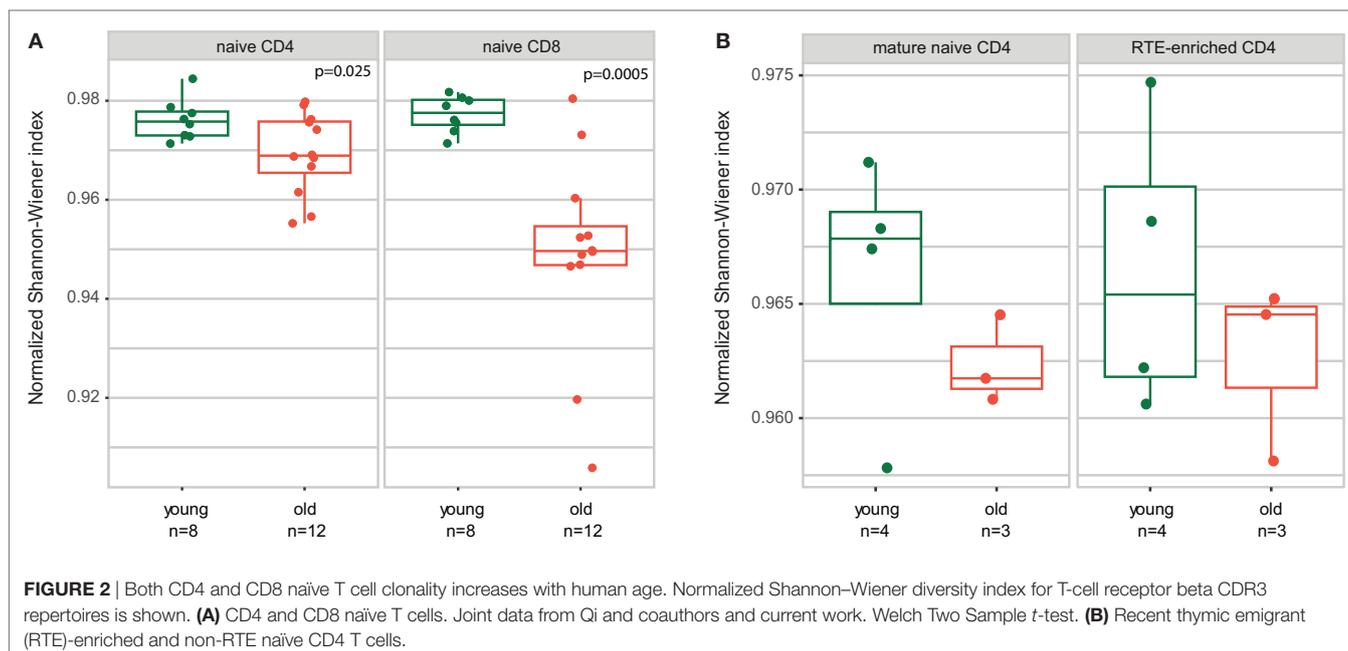
clonotypes within a particular TRBV segment. Therefore, in order to properly join our 5'RACE and multiplex PCR data from Qi et al., we performed further analysis separately for each of the TRBV gene segments that were abundantly represented in the data.

Notably, this approach has two additional benefits. First, different TRBV genes carry distinct CDR1 and CDR2 regions that participate in TCR–pMHC interaction, and, therefore, could differently influence the averaged properties of CDR3 that we analyze below. Separate analysis of TRBV segments allows to neutralize this bias. Second, distinct TRBV genes correspond to distinct T cell subpopulations allowing for independent evaluation of their properties, that provides better statistics for limited donor cohorts. All analyses were

performed “weighted”—per CDR3-covering sequencing read, i.e., accounting for the relative frequency of each clonotype, with Z-score normalization used to combine information from different TRBV segments.

The results of comparative analysis of TRB CDR3 repertoire properties with VDJtools software are shown on **Figure 3**. Notably, dispersion of all parameters grows prominently with age, which already reflects the non-uniform proliferation of naïve T cells with age.

CDR3 length, size of NDN insert, and number of randomly added N nucleotides significantly decrease with age both for CD4 and CD8 naïve T cells (**Figure 3A**). Average characteristics of amino acid residues in the middle of CDR3 also change prominently for CD4 naïve T cells (**Figure 3B**).



Both RTE and Mature Naïve CD4 T Cells Change Their Properties With Aging

To some extent, both CD45RA⁺CD31⁻ mature naïve CD4⁺ T cells and RTE-enriched CD45RA⁺CD31⁺ subsets may support their counts by peripheral division: “CD45RA⁺CD31⁺CD4⁺ subset also undergoes some *in vivo* proliferation without immediate loss of CD31, resulting in an accumulation of CD45RA⁺CD31⁺ proliferative offspring” (30). Nevertheless, counts of CD45RA⁺CD31⁺ naïve CD4⁺ T cell notably decrease with time (5, 30). The CD31⁻ subset is believed to proliferate and support their counts more efficiently than CD31⁺, although the extent of telomere shortening with aging is prominent and comparable for both subsets (30).

Therefore, one could suggest that characteristics of mature naïve CD4⁺CD31⁻ T cells could change more prominently than those of RTE-enriched CD4⁺CD31⁺ T cell pool. The properties of total naïve CD4⁺ T cells could change with aging because of the intrinsic differences between the properties of RTE-enriched and mature naïve CD4 T cell TCR repertoires, and decrease of CD31⁺ cell proportion of all naïve CD4 T cells (5).

To verify the latter hypothesis, we compared TCR beta repertoire characteristics for the sorted CD4⁺CD45RA^{high}CD27^{high}CD31⁺ and CD4⁺CD45RA^{high}CD27^{high}CD31⁻ T cells of 4 young (29–31 years) and 3 elder (aged 51, 55, and 82 years) healthy donors (Table 3). Importantly, to exclude the potential influence of naïve Tregs which characteristics essentially differ from conventional CD4 T cells, here we gated out the CD25⁺ cells from all subsets (Figure 4). It should be noted that this strict gating could also cutoff the CD25^{dull} subset of naïve CD4 T cells that was recently reported to accumulate with aging (52), however, these cells were nearly absent (represented less than 2% of naïve CD4 T cells) in our donors.

Analysis of obtained TCR beta CDR3 repertoires revealed that characteristics of CD4⁺CD45RA^{high}CD27^{high}CD25⁻CD31⁺

and CD4⁺CD45RA^{high}CD27^{high}CD25⁻CD31⁻CD4 T cell TCR repertoires are nearly identical within the same age group, but both prominently differ between the younger and elder donors (Figures 5A,B). It should be noted that, since the average CDR3 length decreases with age, larger portions of TRBV and TRBJ segments could be covered by our analysis of the middle 5 amino acid residues of CDR3, which could in turn influence the result amino acid property averages. However, this influence was not prominent since different TRBV segments behaved similarly in our analysis.

Furthermore, young and old naïve CD4 T cell repertoires were characterized by distinct frequencies of TRBV (Figure 6A), TRBJ (Figure 6B), and paired TRBV–TRBJ (Figure 6C) gene segment usage, without any notable differences observed between the RTE-enriched CD31⁺ and mature naïve CD4 T cell subsets.

Similarly to naïve CD4 and CD8 subsets, RTE-enriched and mature naïve CD4 subsets showed a tendency toward increased clonality in the elder age (Figure 2B).

We concluded that observed changes in the characteristics of naïve CD4 T cell TCR beta CDR3 repertoire with aging affect both RTE-enriched and mature subsets, and do not result from the changes in CD31⁺/CD31⁻ subsets ratio.

Publicity of Naïve CD4 T Cell Repertoire Grows With Aging

Shorter CDR3 length and lower number of randomly added N nucleotides are commonly associated with higher publicity of TCR repertoires (53, 54). To analyze how the relative publicity of naïve CD4 TCR beta repertoires changes with aging, we extracted top-3,000 clonotypes from each dataset, with random sampling of clonotypes having the identical low frequency—a normalization step which is highly desirable to minimize biases in comparison of immune repertoires overlaps. As it could be expected based on

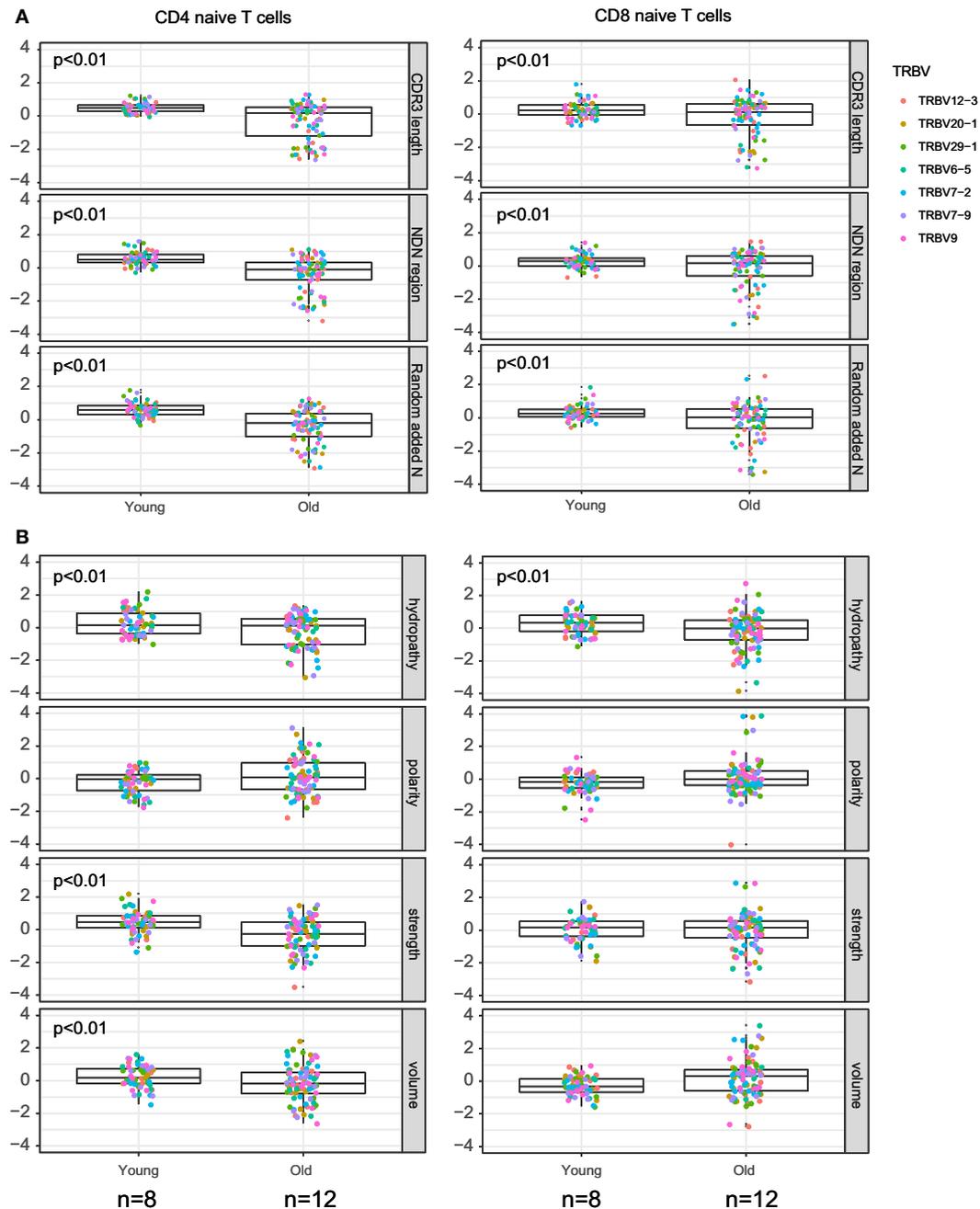
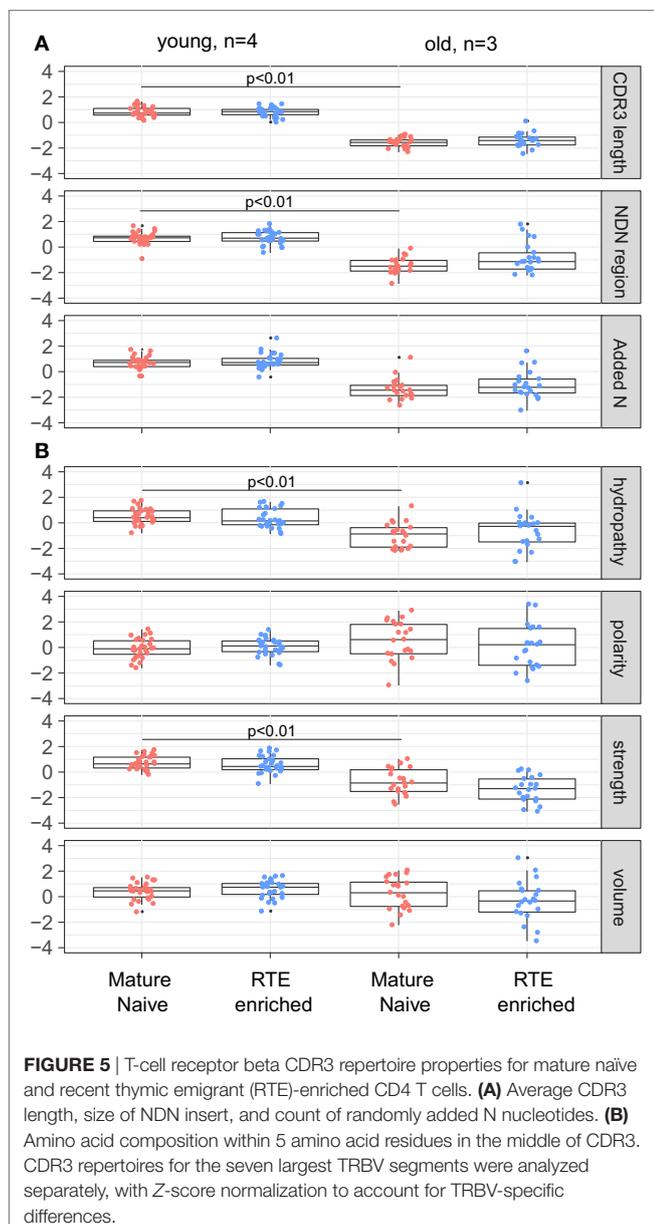
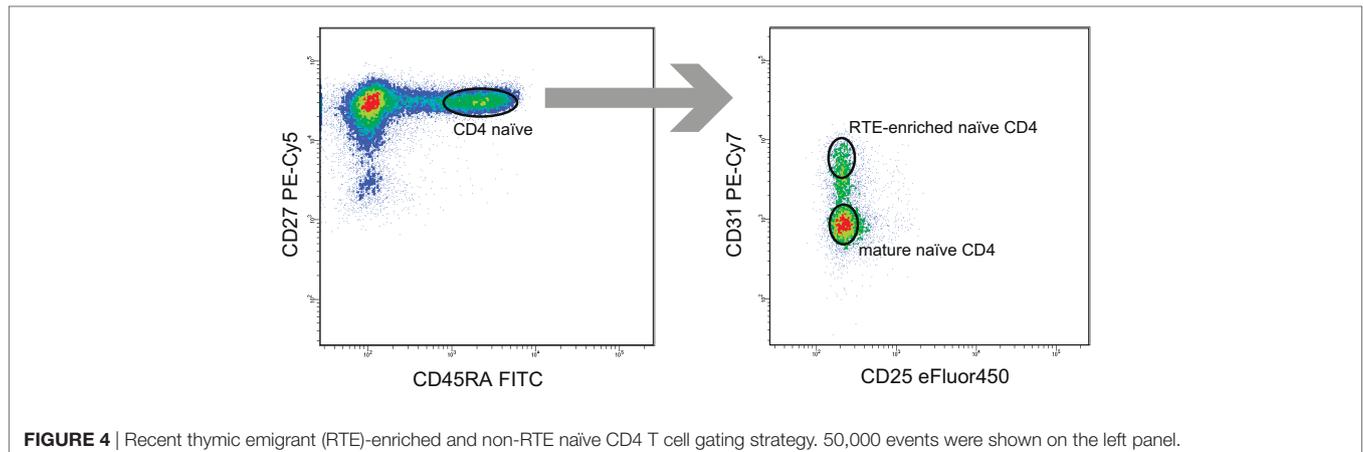


FIGURE 3 | Properties of naive T cell T-cell receptor (TCR) beta CDR3 repertoires and aging. Weighted (accounting for clonotype size) analysis of TCR beta repertoire properties for CD4 and CD8 naive T cells derived from peripheral blood samples of young and old healthy donors. **(A)** Average CDR3 length, size of NDN insert, and count of randomly added N nucleotides. **(B)** Amino acid composition within 5 amino acid residues in the middle of CDR3. Our data and Qi et al. data, $n = 8$ young and 12 old individuals totally. CDR3 repertoires for the seven largest TRBV segments were analyzed separately, with Z-score normalization to account for TRBV-specific differences.

CDR3 characteristics (**Figures 5A** and **7A**), analysis of relative overlaps between TCR beta CDR3 repertoires revealed that relative publicity of total CD4 naive [our data only, excluding the data from Ref. (34)], RTE-enriched CD31⁺ and mature naive CD31⁻ CD4 T cell subsets grows with aging (**Figure 7B**). A moderate overlap was observed between the young and middle-age/old

CD4 naive, RTE-enriched CD31⁺ and mature naive CD31⁻ CD4 T cell subsets. No clear age-related changes in relative publicity were observed for CD8 naive T cells (our data only).

We used CDR3 sequence similarity graph to analyze whether naive TCR repertoires form separate networks in young versus old donors. To build the graph, we selected 3,000 most abundant



clonotypes from each donor and pooled them together to form the set of nodes. We connected two clonotypes with an edge if they had the same VJ-combination and CDR3 differed by a single amino acid substitution. Next, we counted the number of edges connecting clonotypes from donors of different age groups (young versus old) and obtained empirical distributions for these counts by running 1,000 random permutations of age group labels.

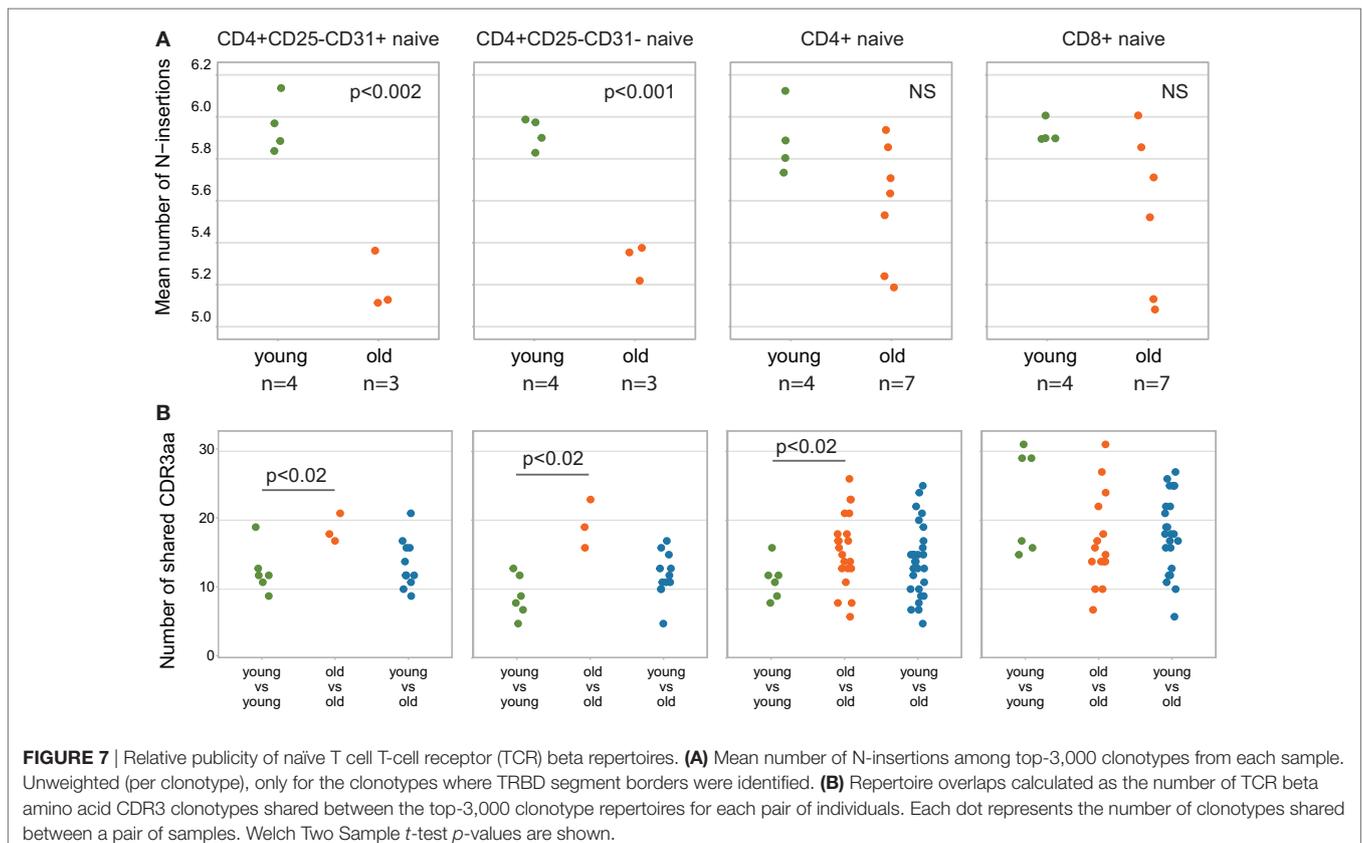
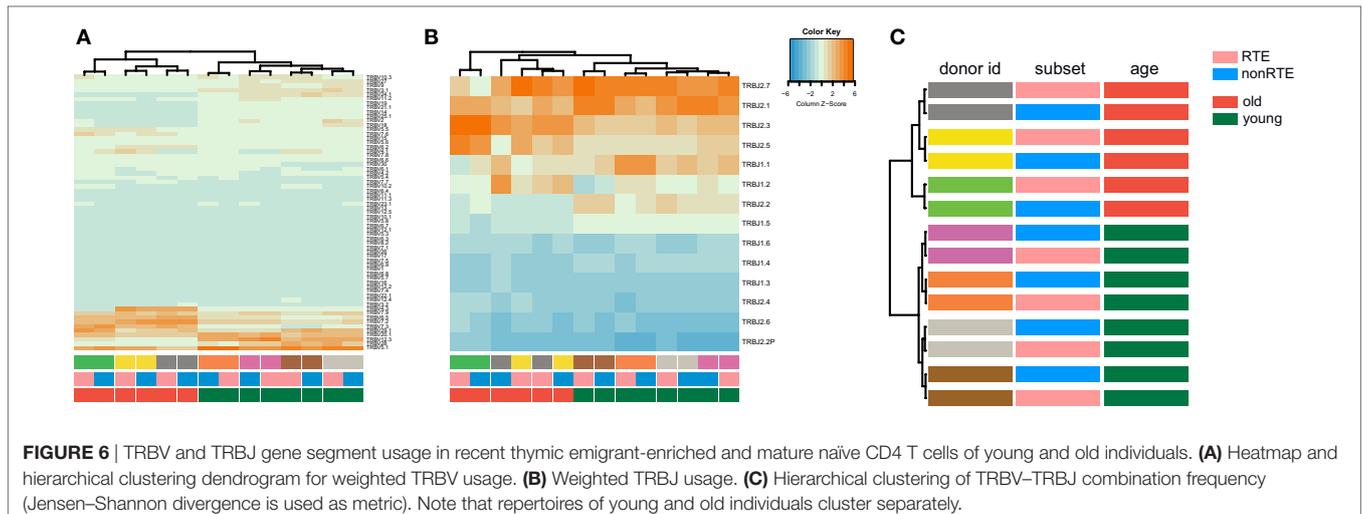
We found, that in CD8 naive repertoires, the number of edges between clonotypes from young and old donors is larger in data than in simulation in 424 donor age group permutations out of 1,000, so there is no evidence for separate CDR3 networks for young and old donors for this subset. In CD4 naive repertoires, however, there was a weak tendency: only in 95 simulations out of 1,000 (empirical *P*-value of 0.095) we found a lower number of edges between donors of different age, than the one observed in real data. This suggests that repertoires of naive CD4 T cells include distinct communities of homologous TCR variants in young and old individuals. However, this effect was only marginally significant and requires further investigation.

DISCUSSION

With aging, decreasing thymic output can not efficiently sustain naive T cell counts, so the homeostatic proliferation becomes the main mechanism to replenish this cell pool in humans. Such proliferation is inevitably associated with certain biases that shape the landscape of naive T cell TCR repertoire and thus affect the spectrum of the antigens they could recognize.

We have utilized immune repertoire sequencing to study the repertoires of naive T cells in young and aged donors and revealed notable changes in human TCR repertoires of both CD4 and CD8 peripheral blood naive T cells with aging:

- (1) We confirm the observation of Ref. (34) that relative clonality reflecting the extent of clonal expansion increases both within CD4 and CD8 naive T cell subsets with age (Figure 2).
- (2) We demonstrate that average CDR3 length, NDN insert length, and number of randomly added N nucleotides significantly decrease with aging in all subsets of naive T cells, including CD4, CD8, CD4 RTE-enriched CD25⁻, and CD4



mature naive CD25⁻ subsets (Figures 3A, 5A and 7A). Interestingly, due to spatial restrictions in TCR-pMHC interaction, the length of CDR3 is inversely related to the length of recognized peptide antigen, which affects the spectrum of recognized pMHCs (Shugay et al., manuscript under consideration). The decrease of CDR3 length with aging could, therefore, reflect the averaged properties of pMHCs that are preferentially recognized by naive T cells in the periphery, and cause better tonic signaling, leading to earlier exhaustion

of proliferation capacity of the cells carrying corresponding TCR variants.

(3) As could be expected based on previous works (53, 54), the abovementioned changes favored higher publicity in CD4 naive T cells (Figure 7B). At the same time, we have not observed clear differences in TCR beta CDR3 repertoire publicity for CD8 compartment. These observations differ from the data from Qi et al. (34) suggesting the decrease of CD8 naive T cell publicity with aging. Further studies

on larger cohorts with thoroughly controlled purity of cell sorting, and proper normalization of the datasets for comparing publicity of repertoires (49) should clarify this point.

- (4) Averaged amino acid characteristics in the middle of CDR3 change prominently in CD4, CD8, CD4 RTE-enriched, and CD4 mature naïve subsets (**Figures 3B** and **5B**). In particular, significant decrease is observed for the “strength” metrics, which represents the count of strongly interacting amino acid residues (47, 48). The “strongly interacting” include F, L, I, M, and V that may form hydrophobic contacts, as well as aromatic residues W and Y that are capable of different types of interactions including offset stacked or edge-to-face interactions, thiol–aromatic interactions, and others (55), and may consist of electrostatic, van der Waals, and hydrophobic forces. Correspondingly, similar changes are observed for the “hydropathy” metrics which counts the number of hydrophobic residues in the middle of CDR3.

The “strength” metric efficiently differentiates functional T cell subpopulations, such as Treg and non-Treg CD4 subsets [see Ref. (49, 56) and our data to be published elsewhere]. This metric can be interpreted as an averaged estimation of TCR repertoire affinity to peptide–MHC complexes and in particular to the antigenic peptide, since the middle portion of CDR3 is often in contact with the presented antigen (**Figure 8**).

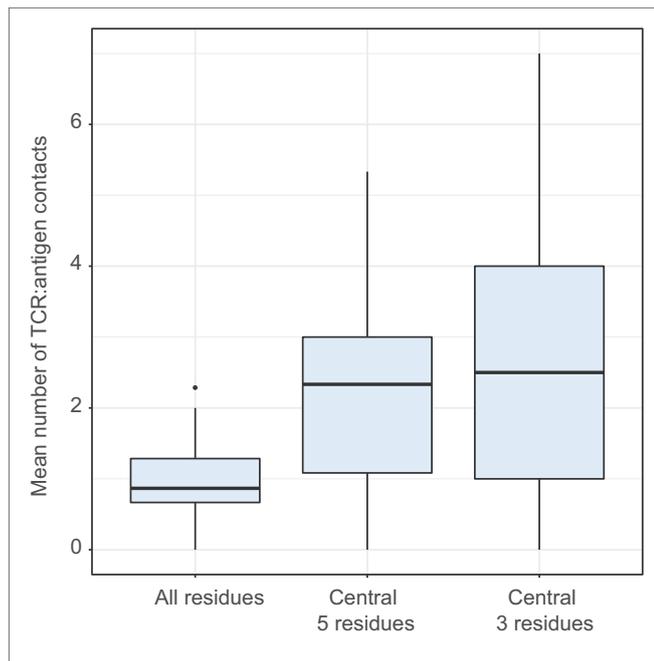


FIGURE 8 | Number of CDR3:antigenic peptide contacts in structural data. Comparing the mean number of contacts for entire CDR3 (All positions) and its central region (central 5 residues and central 3 residues). ANOVA followed by a *post hoc* Tukey test shows significantly higher number of contacts for the central region: $P < 10^{-9}$ when comparing 5 and 3 central residues to all residues, but no difference between 5 and 3 central residues ($P = 0.42$). The analysis was performed for T-cell receptor (TCR) beta chain using 110 human TCR:pMHC complexes from Protein Data Bank.

The decrease of relative abundance of strongly interacting amino acid residues within TCR beta CDR3 repertoire of naïve T cells with aging may, therefore, reflect more rapid depletion of naïve T cell clones with higher affinity to self pMHC. This could result from more efficient tonic signaling and generally faster proliferation, exhaustion of proliferation capacity, and extinction of such naïve T cells (38).

Notably, similar changes were observed within RTE-enriched CD31⁺ and mature naïve CD31⁻ CD4 naïve T cells (**Figures 5–7**). Decrease of the “strength” metric was even more prominent for the RTE-enriched subset (**Figure 5B**), suggesting that the CD31⁺ naïve CD4 T cell clones bearing TCR variants with high affinity to self pMHC are prominently switching to the CD31⁻ phenotype due to more efficient TCR signaling.

Complementary explanation for the changes observed in the naïve T cell TCR repertoire characteristics with aging is that the high affinity variants are washed away from the naïve T cell pool in the course of ongoing immune responses. Both CD4⁺ and CD8⁺ T cells with strong reaction to self and high tonic signaling dominate in responses to foreign antigens (37, 57, 58). Positive selection in thymus thus favors production of more efficiently responding T cells that should be also more rapidly depleted from the naïve T cell pool. If this is the case, the age-related changes are associated with generation of prominent functional holes in the landscape of naïve T cell receptor repertoire.

An additional factor that could contribute to the observed changes in naïve T cell TCR repertoires is the easier conversion of clones with high affinity to self pMHC to the “memory-like” phenotype, as shown in mice models (59, 60), although such observations have not yet found clear confirmation in humans (3).

Altogether, the observed changes could be interpreted as elimination of generally more “sticky”—having higher affinity to self and non-self peptide–MHC complexes—TCR variants from the naïve T cell pool with aging.

However, there is also an alternative explanation which deserves consideration. Shorter CDR3s, lower numbers of randomly added N nucleotides, and higher publicity are characteristic features of the early wave(s) of naïve T cells generated during fetal period (23, 40, 61–63). Such early wave(s) originate from distinct population(s) of hematopoietic stem cells that may have distinct long-term program including higher proliferation potential (39, 42).

Considering the drop in thymic activity that happens in the middle age (4, 26), one could hypothesize that the counts of conventional naïve T cell decrease after exhaustion of their limited proliferation capacity, while the early-wave naïve T cells of fetal origin with prolonged proliferation capacity persist. Such organization of T cell adaptive immunity in the elderly could be beneficial from the point of more predictable innate-like behavior of the T cells carrying a relatively restricted, more germline-encoded TCR repertoire. To some extent, our network analysis of naïve CD4 T cell TCR repertoires supports this concept.

Summing up, our study sheds light on the intrinsic changes in the naïve T cell TCR repertoire structure with aging, and calls

for further functional studies that could clarify the underlying mechanisms.

ETHICS STATEMENT

The study was approved by the local ethics committee and conducted in accordance with the Declaration of Helsinki. All donors were informed of the final use of their blood and signed an informed consent document.

AUTHOR CONTRIBUTIONS

EE and DS performed cell sorting. EE, SK, VZ, TN, MS, and MP analyzed the data. EE, MS, and DC prepared the figures. EE, MI, AA, FC, IM, ER, and AF worked on library preparation and sequencing. DC and OB designed the entire study and wrote the

manuscript. MS, ER, and AF edited the manuscript. All authors reviewed and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Minervina A.A. for the help with figures preparation. Cell sorting experiments were carried out using the equipment provided by the IBCH Core facility (CKP IBCH, supported by Russian Ministry of Education and Science, grant RFMEFI62117 × 0018).

FUNDING

This work was funded by Russian Science Foundation Project 16-15-00149. AF and ER received support from the H2020 EU SYSCID project (grant agreement 733100).

REFERENCES

- Marks BR, Nowyhed HN, Choi JY, Poholek AC, Odegard JM, Flavell RA, et al. Thymic self-reactivity selects natural interleukin 17-producing T cells that can regulate peripheral inflammation. *Nat Immunol* (2009) 10:1125–32. doi:10.1038/ni.1783
- Nikolich-Zugich J. Aging of the T cell compartment in mice and humans: from no naive expectations to foggy memories. *J Immunol* (2014) 193:2622–9. doi:10.4049/jimmunol.1401174
- Goronzy JJ, Fang F, Cavanagh MM, Qi Q, Weyand CM. Naive T cell maintenance and function in human aging. *J Immunol* (2015) 194:4073–80. doi:10.4049/jimmunol.1500046
- Kumar BV, Connors TJ, Farber DL, Human T. Cell development, localization, and function throughout life. *Immunity* (2018) 48:202–13. doi:10.1016/j.immuni.2018.01.007
- van den Broek T, Borghans JAM, van Wijk F. The full spectrum of human naive T cells. *Nat Rev Immunol* (2018) 18(6):363–73. doi:10.1038/s41577-018-0001-y
- Goodwin K, Viboud C, Simonsen L. Antibody response to influenza vaccination in the elderly: a quantitative review. *Vaccine* (2006) 24:1159–69. doi:10.1016/j.vaccine.2005.08.105
- Aspinall R, Del Giudice G, Effros RB, Grubeck-Loebenstien B, Sambhara S. Challenges for vaccination in the elderly. *Immun Ageing* (2007) 4:9. doi:10.1186/1742-4933-4-9
- Pinti M, Appay V, Campisi J, Frasca D, Fulop T, Sauce D, et al. Aging of the immune system: focus on inflammation and vaccination. *Eur J Immunol* (2016) 46:2286–301. doi:10.1002/eji.201546178
- Fulop T, Larbi A, Kotb R, de Angelis F, Pawelec G. Aging, immunity, and cancer. *Discov Med* (2011) 11:537–50.
- Foster AD, Sivarapatna A, Gress RE. The aging immune system and its relationship with cancer. *Aging health* (2011) 7:707–18. doi:10.2217/ahe.11.56
- Fulop T, Larbi A, Witkowski JM, Kotb R, Hirokawa K, Pawelec G. Immunosenescence and cancer. *Crit Rev Oncog* (2013) 18:489–513. doi:10.1615/CritRevOncog.2013010597
- Stronen E, Toebes M, Kelderman S, van Buuren MM, Yang W, van Rooij N, et al. Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* (2016) 352:1337–41. doi:10.1126/science.aaf2288
- Howcroft TK, Campisi J, Louis GB, Smith MT, Wise B, Wyss-Coray T, et al. The role of inflammation in age-related disease. *Aging (Albany NY)* (2013) 5:84–93. doi:10.18632/aging.100531
- Franceschi C, Campisi J. Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. *J Gerontol A Biol Sci Med Sci* (2014) 69(Suppl 1):S4–9. doi:10.1093/gerona/glu057
- Watson N, Ding B, Zhu X, Frisina RD. Chronic inflammation – inflammaging – in the ageing cochlea: a novel target for future presbycusis therapy. *Ageing Res Rev* (2017) 40:142–8. doi:10.1016/j.arr.2017.10.002
- Li MO, Rudensky AY. T cell receptor signalling in the control of regulatory T cell differentiation and function. *Nat Rev Immunol* (2016) 16:220–33. doi:10.1038/nri.2016.26
- Feng Y, van der Veeken J, Shugay M, Putintseva EV, Osmanbeyoglu HU, Dikiy S, et al. A mechanism for expansion of regulatory T-cell repertoire and its role in self-tolerance. *Nature* (2015) 528:132–6. doi:10.1038/nature16141
- Messaoudi I, Lemaoult J, Guevara-Patino JA, Metzner BM, Nikolich-Zugich J. Age-related CD8 T cell clonal expansions constrict CD8 T cell repertoire and have the potential to impair immune defense. *J Exp Med* (2004) 200:1347–58. doi:10.1084/jem.20040437
- Goronzy JJ, Weyand CM. T cell development and receptor diversity during aging. *Curr Opin Immunol* (2005) 17:468–75. doi:10.1016/j.coi.2005.07.020
- Naylor K, Li G, Vallejo AN, Lee WW, Koetz K, Bryl E, et al. The influence of age on T cell generation and TCR diversity. *J Immunol* (2005) 174:7446–52. doi:10.4049/jimmunol.174.11.7446
- Goronzy JJ, Lee WW, Weyand CM. Aging and T-cell diversity. *Exp Gerontol* (2007) 42:400–6. doi:10.1016/j.exger.2006.11.016
- Britanova OV, Putintseva EV, Shugay M, Merzlyak EM, Turchaninova MA, Staroverov DB, et al. Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. *J Immunol* (2014) 192:2689–98. doi:10.4049/jimmunol.1302064
- Britanova OV, Shugay M, Merzlyak EM, Staroverov DB, Putintseva EV, Turchaninova MA, et al. Dynamics of individual T cell repertoires: from cord blood to centenarians. *J Immunol* (2016) 196:5005–13. doi:10.4049/jimmunol.1600005
- Chinn IK, Blackburn CC, Manley NR, Sempowski GD. Changes in primary lymphoid organs with aging. *Semin Immunol* (2012) 24:309–20. doi:10.1016/j.smim.2012.04.005
- Palmer DB. The effect of age on thymic function. *Front Immunol* (2013) 4:316. doi:10.3389/fimmu.2013.00316
- Thome JJ, Grinshpun B, Kumar BV, Kubota M, Ohmura Y, Lerner H, et al. Longterm maintenance of human naive T cells through in situ homeostasis in lymphoid tissue sites. *Sci Immunol* (2016) 1:eaa6506. doi:10.1126/sciimmunol.aah6506
- Kyoizumi S, Kubo Y, Kajimura J, Yoshida K, Imai K, Hayashi T, et al. Age-associated changes in the differentiation potentials of human circulating hematopoietic progenitors to T- or NK-lineage cells. *J Immunol* (2013) 190:6164–72. doi:10.4049/jimmunol.1203189
- Haines CJ, Giffon TD, Lu LS, Lu X, Tessier-Lavigne M, Ross DT, et al. Human CD4+ T cell recent thymic emigrants are identified by protein tyrosine kinase 7 and have reduced immune function. *J Exp Med* (2009) 206:275–85. doi:10.1084/jem.20080996
- Fink PJ. The biology of recent thymic emigrants. *Annu Rev Immunol* (2013) 31:31–50. doi:10.1146/annurev-immunol-032712-100010
- Kilpatrick RD, Rickabaugh T, Hultin LE, Hultin P, Hausner MA, Detels R, et al. Homeostasis of the naive CD4+ T cell compartment during aging. *J Immunol* (2008) 180:1499–507. doi:10.4049/jimmunol.180.10.6437-a

31. Qi Q, Zhang DW, Weyand CM, Goronzy JJ. Mechanisms shaping the naive T cell repertoire in the elderly – thymic involution or peripheral homeostatic proliferation? *Exp Gerontol* (2014) 54:71–4. doi:10.1016/j.exger.2014.01.005
32. den Braber I, Mugwagwa T, Vrisekoop N, Westera L, Mogling R, de Boer AB, et al. Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. *Immunity* (2012) 36:288–97. doi:10.1016/j.immuni.2012.02.006
33. Sauce D, Larsen M, Fastenackels S, Roux A, Gorochov G, Katlama C, et al. Lymphopenia-driven homeostatic regulation of naive T cells in elderly and thymectomized young adults. *J Immunol* (2012) 189:5541–8. doi:10.4049/jimmunol.1201235
34. Qi Q, Liu Y, Cheng Y, Glanville J, Zhang D, Lee JY, et al. Diversity and clonal selection in the human T-cell repertoire. *Proc Natl Acad Sci U S A* (2014) 111:13139–44. doi:10.1073/pnas.1409155111
35. Takeda S, Rodewald HR, Arakawa H, Bluethmann H, Shimizu T. MHC class II molecules are not required for survival of newly generated CD4+ T cells, but affect their long-term life span. *Immunity* (1996) 5:217–28. doi:10.1016/S1074-7613(00)80317-9
36. Azzam HS, Grinberg A, Lui K, Shen H, Shores EW, Love PE. CD5 expression is developmentally regulated by T cell receptor (TCR) signals and TCR avidity. *J Exp Med* (1998) 188:2301–11. doi:10.1084/jem.188.12.2301
37. Stefanova I, Dorfman JR, Germain RN. Self-recognition promotes the foreign antigen sensitivity of naive T lymphocytes. *Nature* (2002) 420:429–34. doi:10.1038/nature01146
38. Myers DR, Zikherman J, Roose JP. Tonic signals: why do lymphocytes bother? *Trends Immunol* (2017) 38:844–57. doi:10.1016/j.it.2017.06.010
39. Pogorelyy MV, Elhanati Y, Marcou Q, Sycheva AL, Komech EA, Nazarov VI, et al. Persisting fetal clonotypes influence the structure and overlap of adult human T cell receptor repertoires. *PLoS Comput Biol* (2017) 13:e1005572. doi:10.1371/journal.pcbi.1005572
40. George JF Jr, Schroeder HW Jr. Developmental regulation of D beta reading frame and junctional diversity in T cell receptor-beta transcripts from human thymus. *J Immunol* (1992) 148:1230–9.
41. Benedict CL, Gilfillan S, Thai TH, Kearney JF. Terminal deoxynucleotidyl transferase and repertoire development. *Immunol Rev* (2000) 175:150–7. doi:10.1111/j.1600-065X.2000.imr017518.x
42. Mold JE, Venkatasubrahmanyam S, Burt TD, Michaelsson J, Rivera JM, Galkina SA, et al. Fetal and adult hematopoietic stem cells give rise to distinct T cell lineages in humans. *Science* (2010) 330:1695–9. doi:10.1126/science.1196509
43. Egorov ES, Merzlyak EM, Shelonov AA, Britanova OV, Sharonov GV, Staroverov DB, et al. Quantitative profiling of immune repertoires for minor lymphocyte counts using unique molecular identifiers. *J Immunol* (2015) 194:6155–63. doi:10.4049/jimmunol.1500215
44. Egorov E. Human TCR alpha and beta RNA-based RACE protocol with unique molecular barcoding. *Protoc Exch* (2016). doi:10.1038/protex.2016.044
45. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods* (2015) 12:380–1. doi:10.1038/nmeth.3364
46. Shugay M, Bagaev DV, Turchaninova MA, Bolotin DA, Britanova OV, Putintseva EV, et al. VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput Biol* (2015) 11:e1004503. doi:10.1371/journal.pcbi.1004503
47. Miyazawa S, Jernigan RL. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J Mol Biol* (1996) 256:623–44. doi:10.1006/jmbi.1996.0114
48. Kosmrlj A, Jha AK, Huseby ES, Kardar M, Chakraborty AK. How the thymus designs antigen-specific and self-tolerant T cell receptor sequences. *Proc Natl Acad Sci U S A* (2008) 105:16671–6. doi:10.1073/pnas.0808081105
49. Izraelson M, Nakonechnaya TO, Moltedo B, Egorov ES, Kasatskaya SA, Putintseva EV, et al. Comparative analysis of murine T-cell receptor repertoires. *Immunology* (2018) 153:133–44. doi:10.1111/imm.12857
50. Bolotin DA, Mamedov IZ, Britanova OV, Zvyagin IV, Shagin D, Ustyugova SV, et al. Next generation sequencing for TCR repertoire profiling: platform-specific features and correction algorithms. *Eur J Immunol* (2012) 42:3073–83. doi:10.1002/eji.201242517
51. Carlson CS, Emerson RO, Sherwood AM, Desmarais C, Chung MW, Parsons JM, et al. Using synthetic templates to design an unbiased multiplex PCR assay. *Nat Commun* (2013) 4:2680. doi:10.1038/ncomms3680
52. Pekalski ML, Ferreira RC, Coulson RM, Cutler AJ, Guo H, Smyth DJ, et al. Postthymic expansion in human CD4 naive T cells defined by expression of functional high-affinity IL-2 receptors. *J Immunol* (2013) 190:2554–66. doi:10.4049/jimmunol.1202914
53. Venturi V, Price DA, Douek DC, Davenport MP. The molecular basis for public T-cell responses? *Nat Rev Immunol* (2008) 8:231–8. doi:10.1038/nri2260
54. Murugan A, Mora T, Walczak AM, Callan CG Jr. Statistical inference of the generation probability of T-cell receptors from sequence repertoires. *Proc Natl Acad Sci U S A* (2012) 109:16161–6. doi:10.1073/pnas.1212755109
55. Chakrabarti P, Bhattacharyya R. Geometry of nonbonded interactions involving planar groups in proteins. *Prog Biophys Mol Biol* (2007) 95:83–137. doi:10.1016/j.pbiomolbio.2007.03.016
56. Bolotin DA, Poslavsky S, Davydov AN, Frenkel FE, Fanchi L, Zolotareva OI, et al. Antigen receptor repertoire profiling from RNA-seq data. *Nat Biotechnol* (2017) 35:908–11. doi:10.1038/nbt.3979
57. Mandl JN, Monteiro JB, Vrisekoop N, Germain RN. T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity* (2013) 38:263–74. doi:10.1016/j.immuni.2012.09.011
58. Fulton RB, Hamilton SE, Xing Y, Best JA, Goldrath AW, Hogquist KA, et al. The TCR's sensitivity to self peptide-MHC dictates the ability of naive CD8(+) T cells to respond to foreign antigens. *Nat Immunol* (2015) 16:107–17. doi:10.1038/ni.3043
59. Rudd BD, Venturi V, Li G, Samadder P, Ertelt JM, Way SS, et al. Nonrandom attrition of the naive CD8+ T-cell pool with aging governed by T-cell receptor:pMHC interactions. *Proc Natl Acad Sci U S A* (2011) 108:13694–9. doi:10.1073/pnas.1107594108
60. Sprent J, Surh CD. Normal T cell homeostasis: the conversion of naive cells into memory-phenotype cells. *Nat Immunol* (2011) 12:478–84. doi:10.1038/ni.2018
61. Feeney AJ. Lack of N regions in fetal and neonatal mouse immunoglobulin V-D-J junctional sequences. *J Exp Med* (1990) 172:1377–90. doi:10.1084/jem.172.5.1377
62. Bogue M, Gilfillan S, Benoist C, Mathis D. Regulation of N-region diversity in antigen receptors through thymocyte differentiation and thymus ontogeny. *Proc Natl Acad Sci U S A* (1992) 89:11011–5. doi:10.1073/pnas.89.22.11011
63. Cherrier M, Cardona A, Rosinski-Chupin I, Rougeon F, Doyen N. Substantial N diversity is generated in T cell receptor alpha genes at birth despite low levels of terminal deoxynucleotidyl transferase expression in mouse thymus. *Eur J Immunol* (2002) 32:3651–6. doi:10.1002/1521-4141(200212)32:12<3651::AID-IMMU3651>3.0.CO;2-D

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Egorov, Kasatskaya, Zubov, Izraelson, Nakonechnaya, Staroverov, Angius, Cucca, Mamedov, Rosati, Franke, Shugay, Pogorelyy, Chudakov and Britanova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Chapter II

Memory CD4⁺ T cells are generated
in the human fetal intestine

Introduction

Chapter I and earlier publications of our research group studied TCR repertoires in naive T cells of young adult donors and older donors [1]. Chapter II follows the research of naive T cells in earlier developmental stages: in fetal immunity. Fetal and neonatal T cells are considered naive as the normally fetal immune system does not encounter foreign antigens before birth. Unexpectedly, in this collaborative study of T cell subsets from fetal intestines, we found that immune memory develops in fetal tissue. This study showed for the first time that CD4⁺ T cells might acquire an immune memory phenotype and display features of resident memory T cells in the gut.

Moreover, we described the T cell expansions both in the helper T cell subset and in the regulatory T cell subset with distinct characteristics of TCRs in naive and memory subsets. The novel finding of early immune memory opened up a field for research into the immune memory activation mechanisms in both helper and regulatory T cells. We continue to explore the patterns and mechanisms of memory formation in CD4⁺ subsets in more detail in Chapter III.

Contribution

Nature Immunology publication “Memory CD4⁺ T cells are generated in the human fetal intestine” used multiple approaches to create a detailed description of phenotypes and probable functionalities of the T cells in fetal intestines. The discovery of memory-like subsets of T cells in fetal samples required confirmation with the TCR structure assessment via TCR sequencing and phenotypical characterization after *in vitro* activation of naive-like and memory-like T cell subsets.

I performed the data analysis of TCR sequencing of sorted T cell subsets, therefore, contributing to the computational analysis part of the publication. TCR sequencing allowed to characterize the clonal expansions and to confirm the memory status of T cell subsets in the fetal intestine. I contributed to the integration of data from different assays, interpretation, and discussion. I participated in the Methods, Results, and Discussion sections of the manuscript. My colleagues Egorov E. and Izraelson M. performed TCR libraries preparation from sorted T cells and, therefore, contributed to the manuscript’s experimental part.

Memory CD4⁺ T cells are generated in the human fetal intestine

Na Li¹, Vincent van Unen¹, Tamim Abdelaal^{2,3}, Nannan Guo¹, Sofya A. Kasatskaya^{4,5}, Kristin Ladell⁶, James E. McLaren⁶, Evgeny S. Egorov⁴, Mark Izraelson⁴, Susana M. Chuva de Sousa Lopes⁷, Thomas Höllt^{2,8}, Olga V Britanova⁴, Jeroen Eggermont⁹, Noel F. C. C. de Miranda¹⁰, Dmitriy M. Chudakov^{4,5,11,12,13,14}, David A. Price¹⁵, Boudewijn P. F. Lelieveldt^{3,9} and Frits Koning^{1*}

The fetus is thought to be protected from exposure to foreign antigens, yet CD45RO⁺ T cells reside in the fetal intestine. Here we combined functional assays with mass cytometry, single-cell RNA sequencing and high-throughput T cell antigen receptor (TCR) sequencing to characterize the CD4⁺ T cell compartment in the human fetal intestine. We identified 22 CD4⁺ T cell clusters, including naive-like, regulatory-like and memory-like subpopulations, which were confirmed and further characterized at the transcriptional level. Memory-like CD4⁺ T cells had high expression of Ki-67, indicative of cell division, and CD5, a surrogate marker of TCR avidity, and produced the cytokines IFN- γ and IL-2. Pathway analysis revealed a differentiation trajectory associated with cellular activation and proinflammatory effector functions, and TCR repertoire analysis indicated clonal expansions, distinct repertoire characteristics and interconnections between subpopulations of memory-like CD4⁺ T cells. Imaging mass cytometry indicated that memory-like CD4⁺ T cells colocalized with antigen-presenting cells. Collectively, these results provide evidence for the generation of memory-like CD4⁺ T cells in the human fetal intestine that is consistent with exposure to foreign antigens.

Adaptive immunity is founded on the selection and expansion of antigen-specific T cells from a clonally diverse pool of naive precursors¹. Naive T cells recirculate among lymph nodes to survey the array of peptide epitopes bound to major histocompatibility complex (MHC) proteins on the surface of antigen-presenting cells (APCs), and functional recognition of a given peptide–MHC molecule is governed by various danger signals and specific engagement via the clonotypically expressed TCR. This triggers a program of differentiation and proliferation that results in the generation of effector T cells, which home to the site of the primary infection and contribute to pathogen clearance, and memory T cells, which remain in the circulation and mediate anamnestic responses to secondary infection. In the last decade, it has also become clear that tissue-resident T cells are commonly present at barrier sites, including the intestine².

Fundamental knowledge of adaptive immunity during early life remains sparse. The infantile intestine is known to harbor clonally expanded T cells³, which have also been identified in the human fetal intestine, but rarely in fetal mesenteric lymph nodes, fetal thymus or fetal spleen, suggesting compartmentalization⁴. In addition, a rare population of CD4⁺ T cells displaying a memory and proinflammatory phenotype has been identified in umbilical cord blood⁵. Although the dogma of a sterile womb has been challenged by reports of bacteria colonization in the placenta^{6,7}, amniotic fluid^{8,9}

and meconium¹⁰, others have questioned these results¹¹. Here we have combined functional studies with mass cytometry, RNA sequencing (RNA-seq) and high-throughput TCR sequencing to perform an in-depth analysis of the fetal intestinal CD4⁺ T cell compartment. Our results provide evidence for memory formation in the human fetal intestine, consistent with in utero exposure to foreign antigens.

Results

Human fetal intestinal CD4⁺ T cells are phenotypically diverse.

To explore the CD4⁺ T cell compartment in the human fetal intestine, we applied a mass cytometry panel comprising 35 antibodies (Supplementary Table 1) that was designed to capture the heterogeneity of the immune system to seven lamina propria samples aged 14–21 gestational weeks¹². After data acquisition, we selected CD45⁺ immune cells (Supplementary Fig. 1a) and mined the dataset via hierarchical stochastic neighbor embedding (HSNE)¹³. At the overview level, HSNE landmarks depicted the general composition of the immune system, with clear separation of the CD4⁺ T cell lineage (Supplementary Fig. 1b). We identified 110,332 CD4⁺ T cells, with an average of 15,761 events per fetal intestine, comprising 47.9 ± 9.6% of all immune cells. We then subjected HSNE-defined CD4⁺ T cells (Supplementary Fig. 1b) to t-distributed stochastic neighbor embedding (t-SNE)¹⁴ in Cytosplore¹⁵ to

¹Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, the Netherlands. ²Leiden Computational Biology Center, Leiden University Medical Center, Leiden, the Netherlands. ³Department of Pattern Recognition and Bioinformatics Group, Delft University of Technology, Delft, the Netherlands. ⁴Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia. ⁵Centre for Data-Intensive Biomedicine and Biotechnology, Skolkovo Institute of Science and Technology, Moscow, Russia. ⁶Division of Infection and Immunity, Cardiff University School of Medicine, Cardiff, UK. ⁷Department of Anatomy and Embryology, Leiden University Medical Center, Leiden, the Netherlands. ⁸Computer Graphics and Visualization Group, Delft University of Technology, Delft, the Netherlands. ⁹Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands. ¹⁰Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands. ¹¹Central European Institute of Technology, Masaryk University, Brno, Czech Republic. ¹²Department of Molecular Technologies, Pirogov Russian National Research Medical University, Moscow, Russia. ¹³MiLaboratory LLC, Skolkovo Innovation Centre, Moscow, Russia. ¹⁴Privolzhsky Research Medical University, Nizhny Novgorod, Russia. ¹⁵Systems Immunity Research Institute, Cardiff University School of Medicine, Cardiff, UK. *e-mail: F.Koning@lumc.nl

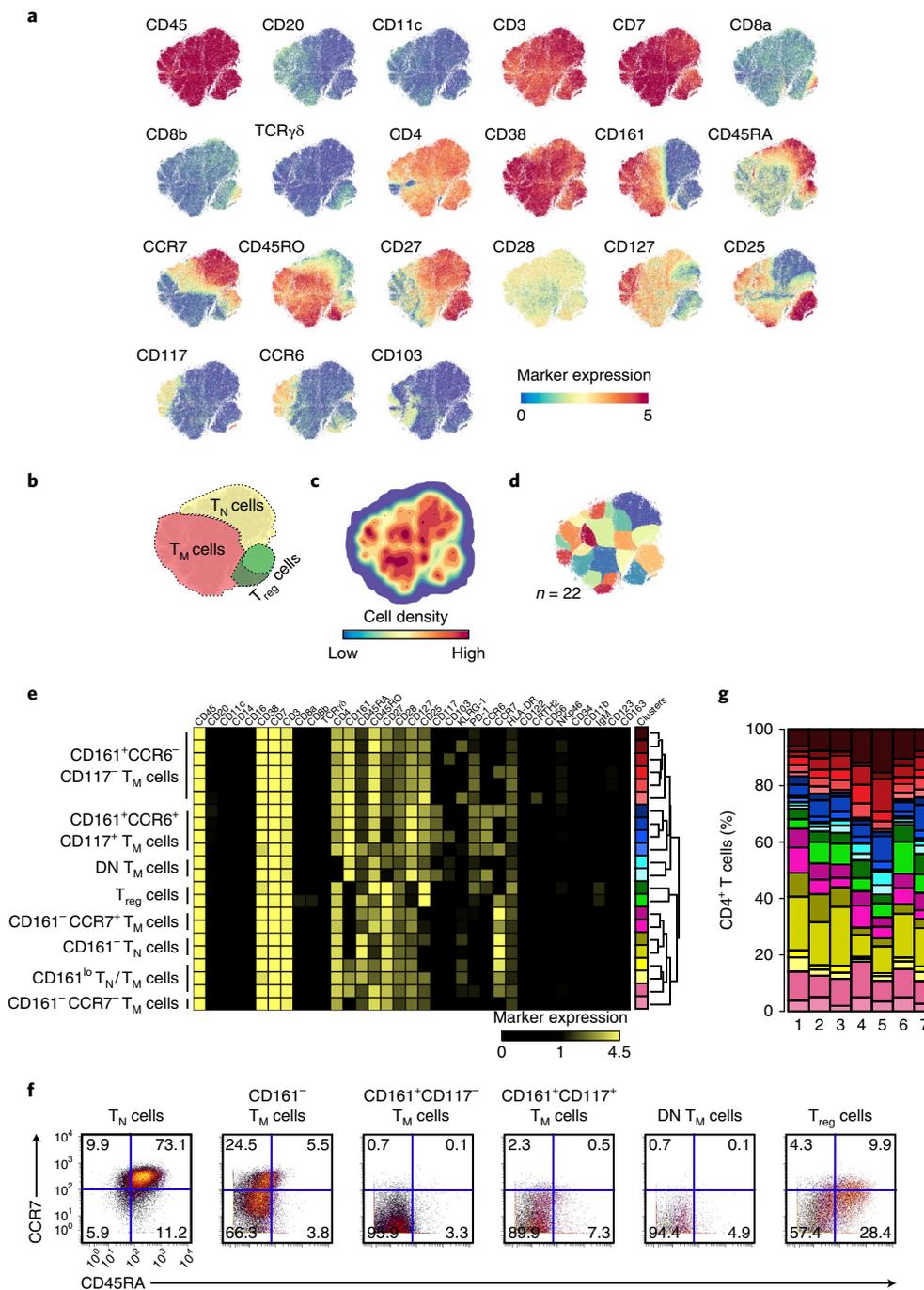


Fig. 1 | Mass cytometric analysis of fetal intestinal CD4⁺ T cells. **a**, t-SNE embedding of all CD4⁺ T cells ($n=110,332$) derived from human fetal intestines ($n=7$). Colors represent the ArcSinh5-transformed expression values of the indicated markers. **b**, t-SNE plot depicting the population cell border for T_N cells (dashed yellow line), T_M cells (dashed red line) and T_{reg} cells (dashed green line). **c**, Density map describing the local probability density of cells, where black dots indicate the centroids of identified clusters using Gaussian mean-shift clustering. **d**, t-SNE plot showing cluster partitions in different colors. **e**, Heat map showing median expression values and hierarchical clustering of markers for the identified subpopulations. **f**, Biaxial plots showing CD45RA and CCR7 expression on the indicated clusters analyzed by mass cytometry. The 22 clusters were merged into six phenotypic groups according to the heat map shown in **e**. **g**, Composition of the CD4⁺ T cell compartment in each fetal intestine represented by vertical bars, where the colored segment lengths represent the proportion of cells as a percentage of all CD4⁺ T cells in the sample. Colors as shown in **e**.

project their marker expression profiles onto a two-dimensional graph (Fig. 1a and Supplementary Fig. 1c). CD4⁺ T cells were characterized as CD45⁺CD3⁺CD4⁺CD7⁺ (Fig. 1a). Moreover, all CD4⁺ T cells were positive for the tissue-resident marker CD38 and approximately 50% of cells expressed CD161. Of the CD4⁺ T cell population, 24.1% coexpressed CD27, CD28, CD45RA and

CCR7, indicative of a naive T cell (T_N) phenotype, whereas 64.5% expressed CD45RO, indicative of a memory T cell (T_M) phenotype (Fig. 1a,b). While all CD45RO⁺ T_M cells were CD28⁺, differential expression of CD25, CD27, CD103, CD117, CD127, CCR6 and CCR7 was observed on these cells (Fig. 1a,b), reflecting substantial phenotypic diversity.

We next applied Gaussian mean-shift clustering to the mass cytometry data using the t-SNE coordinates of the embedded CD4⁺ T cells (Fig. 1a). On the basis of cell density features (Fig. 1c), this identified 22 distinct CD4⁺ T cell clusters (Fig. 1d), each defined by a unique marker expression profile. Hierarchical clustering of the heat map revealed eight major groups (CD161⁺CCR6⁻CD117⁻ T_M cells, CD161⁺CCR6⁺CD117⁺ T_M cells, double negative (DN) T_M cells, T_{reg} cells, CD161⁻CCR7⁺ T_M cells, CD161⁻ T_N cells, CD161^{lo} T_N/T_M cells and CD161⁻CCR7⁻ T_M cells) (Fig. 1e). High expression of CD25 and a lack of CD127 distinguished two regulatory T cell (T_{reg}) clusters, with either a CD45RA⁺ T_N or a CD45RO⁺ T_M phenotype (Fig. 1a,b,e). The CD161⁺CD4⁺ T cells branched into a CCR6⁻CD117⁻CD45RO⁺ T_M and a CCR6⁺CD117⁺CD45RO⁺ T_M cluster (Fig. 1e). Moreover, CD45RA⁺ T_N and CD45RO⁺ T_M cells were detected in both the CD161⁻ and the CD161^{lo} subpopulations. Additional diversity was observed for the expression of several activation markers, including CRTH2, HLA-DR, KLRG-1 and PD-1, the latter especially within the CD45RO⁺ T_M cell clusters (Supplementary Fig. 1c). Of note, a small population of CD4⁺CD8a⁻TCRγδ⁻ (DN) T_M cells clustered among CD4⁺ T cells in both the HSNE and t-SNE plots. Biaxial plots confirmed coexpression of CD45RA and CCR7 on T_N cells (Fig. 1f), whereas the CD161^{lo}-CD45RO⁺ T_M subpopulation contained both CCR7⁺ central memory T (T_{CM}) cells and CCR7⁻ effector memory T (T_{EM}) cells (Fig. 1f). All other CD45RO⁺ T_M subpopulations harbored primarily T_{EM} cells. Quantification of cellular frequencies for the CD4⁺ T cell clusters per fetal intestine revealed highly similar compositions with all CD45RO⁺ T_M clusters detectable in all samples (Fig. 1g). In contrast, parallel analyses of CD4⁺ T cells isolated from three fetal livers and three fetal spleens, from one shared and two additional fetuses aged 16–21 gestational weeks, revealed a predominance of CD45RA⁺ T_N cells (Supplementary Fig. 2a,b). These results delineated a phenotypically diverse array of human fetal intestinal CD4⁺ T cells, most of which displayed features associated with antigen exposure.

Fetal CD4⁺ T cells display a memory gene expression profile.

We next performed single-cell RNA-seq on flow-sorted fetal intestinal CD4⁺ cells from a lamina propria sample that was also included in the mass cytometry analysis. This yielded data for 1,804 CD4⁺ T cells, identifying cell-specific variable expression of 2,174 genes (see Methods), which were further analyzed using the Seurat computational pipeline¹⁶. Unsupervised clustering revealed nine transcriptionally distinct subpopulations, seven of which corresponded to CD3⁺ T cell subsets, while two displayed a gene expression profile matching CD86⁺HLA-DR⁺ APCs. The corresponding gene expression profiles of the seven T cell subsets were projected onto a single graph using t-SNE (Fig. 2a), and the top 20 upregulated genes were displayed in a heat map (Fig. 2b). Five of the seven RNA-seq-identified CD4⁺ T cell subpopulations corresponded to the mass cytometry-defined CD4⁺ T cell major groups: CCR7⁺ T_N with CD45RA⁺ T_N, KLRB1^{lo}-SELL⁻ T_M with CD161^{lo}-CD45RO⁺ T_M, KLRB1⁺CCR6⁻SELL⁻ T_M with CD161⁺CCR6⁻CD45RO⁺ T_M, KLRB1⁺CCR6⁺SELL⁻ T_M with CD161⁺CCR6⁺CD45RO⁺ T_M and FOXP3⁺ T_{reg} cells with CD25⁺CD127^{lo} T_{reg} cells. The mass cytometry-defined CD161⁻ and CD161^{lo} subpopulations (Fig. 1e) could not be discriminated in the RNA-seq dataset. One additional RNA-seq-identified subpopulation corresponded to proliferating cells, on the basis of the expression of genes associated with cell division (CCNB2, CDK1 and MKI67) (Fig. 2a,b).

As CD45RA and CD45RO were not detectable, we used other markers to distinguish T_M from T_N clusters. To compare gene or marker expression among cell clusters, we used violin plots, displaying the mode average as the thickest section (Fig. 2c–e). Consistent with the mass cytometry data, RNA-seq-defined T_N cells were KLRB1⁻CCR7⁺SELL^{+/−} (Fig. 2c,d), the latter confirmed by flow cytometry (Fig. 2e and Supplementary Fig. 3a). In the absence of

T_M-associated markers, SELL⁻ T_M cell populations were identified on the basis of differential expression of KLRB1 and CCR6 (Fig. 2c,d). Consistent with the mass cytometry data, expression of KIT (CD117) was restricted primarily to KLRB1⁺CCR6⁺SELL⁻ T_M cells (Supplementary Fig. 3b). Moreover, the gene expression profile of the IL2RA⁺IL7R^{lo}FOXP3⁺ T_{reg} cell population (Fig. 2c) corresponded to the mass cytometry-defined CD25⁺CD127^{lo} T_{reg} cells (Fig. 2d). In addition, several RNA transcripts, including LAG3, TIGIT, CTLA4 and TNFRSF18 (or glucocorticoid-induced TNFR-related) ascertained the identity of FOXP3⁺ T_{reg} cells (Fig. 2b). Finally, the RNA-seq data revealed an undefined T_M cluster that was not identified by mass cytometry, but expressed genes similar to those detected in the KLRB1⁺CCR6⁺SELL⁻ T_M subpopulation, such as CD69, CCL5 and JAML. Cell population frequencies identified by mass cytometry and RNA-seq were comparable with the exception of mass cytometry-defined CD25⁺CD127^{lo} T_{reg} cells and RNA-seq-defined FOXP3⁺ T_{reg} cells (Supplementary Fig. 3c).

Compared with the CCR7⁺ T_N population, KLRB1^{lo}-SELL⁻ T_M, KLRB1⁺CCR6⁻SELL⁻ T_M, KLRB1⁺CCR6⁺SELL⁻ T_M and undefined T_M subpopulations had high expression of the tissue-resident and activation-associated gene CD69, the differentiation-promoting gene ANXA1 (Annexin A1), the chemokine-like factor CKLF, the cytokine IL32, the proliferation-associated gene JUN (C-Jun) and the adhesion molecule JAML (Fig. 3a). CD40LG (CD154), TNFSF14 and TGFBI were specifically upregulated by KLRB1⁺CCR6⁻SELL⁻ T_M, KLRB1⁺CCR6⁺SELL⁻ T_M and undefined T_M clusters, while CCL5 and MAP3K8 kinase were upregulated by KLRB1⁺CCR6⁻SELL⁻ T_M and undefined T_M subpopulations. Moreover, IL4I1 was specifically expressed by KLRB1⁺CCR6⁺SELL⁻ T_M and undefined T_M cells. In addition, all fetal SELL⁻ T_M subpopulations had high expression of the tissue-resident genes ITGAE (CD103) and/or CD38 (Fig. 3a).

In agreement with the RNA-seq data, flow cytometry indicated that the activation markers CXCR3, CCR4, CD69 and CD226 were highly expressed on CCR7⁻ T_{EM} cells (Fig. 3b). All CD4⁺ T cells expressed CD95, with the highest expression on CD161⁺ T_{EM} cells (Fig. 3b). Expression of CD31, a marker associated with recent thymic emigrants¹⁷, was highest on CD45RA⁺ T_N cells (Fig. 3b). Thus, RNA-seq confirmed the existence of distinct subpopulations of CD4⁺ T cells and indicated that many genes associated with inflammation and tissue residency were upregulated by fetal CD4⁺ T_M cells, consistent with antigen-driven functionality and maturation.

Computational analysis reveals a differentiation pathway of CD4⁺ T cells.

We next visualized the evolution of the t-SNE computation of the mass cytometry and RNA-seq data to reveal the ordering of single cells along putative differentiation trajectories^{12,15}. At the onset of the mass cytometry data computation, where cells are grouped on the basis of major shared features, CD25⁺CD127^{lo} T_{reg} cells clustered separately from the other cells, whereas the other cell clusters were ordered in a linear fashion with the CD45RA⁺ T_N cells next to the CD161^{lo}-CD45RO⁺ T_M cells, followed by the CD161⁺CCR6⁻CD45RO⁺ T_M cells and the CD161⁺CCR6⁺CD45RO⁺ T_M cells, consecutively (Fig. 4a). A similar phenotypic ordering was observed in parallel analyses of the RNA-seq data, although the KLRB1⁺CCR6⁺SELL⁻ T_M subpopulation aligned differently, but remained connected with the KLRB1⁺CCR6⁻SELL⁻ T_M cluster (Fig. 4b). Individual marker expression patterns at the middle of the t-SNE computation validated the ordering of the clusters and the comparability of the mass cytometry and RNA-seq data (Supplementary Fig. 4a,b). Similar patterns were identified using Diffusion map¹⁸, Vortex¹⁹ and principal component analysis (PCA)²⁰ (Supplementary Fig. 4c–f). Therefore, this analysis reveals a putative differentiation pathway leading to T_M formation.

To extend our analysis of the gene expression profiles underlying this putative differentiation trajectory, we used the pseudotime algorithm in the Monocle toolkit^{21,22}, which calculates the ordering

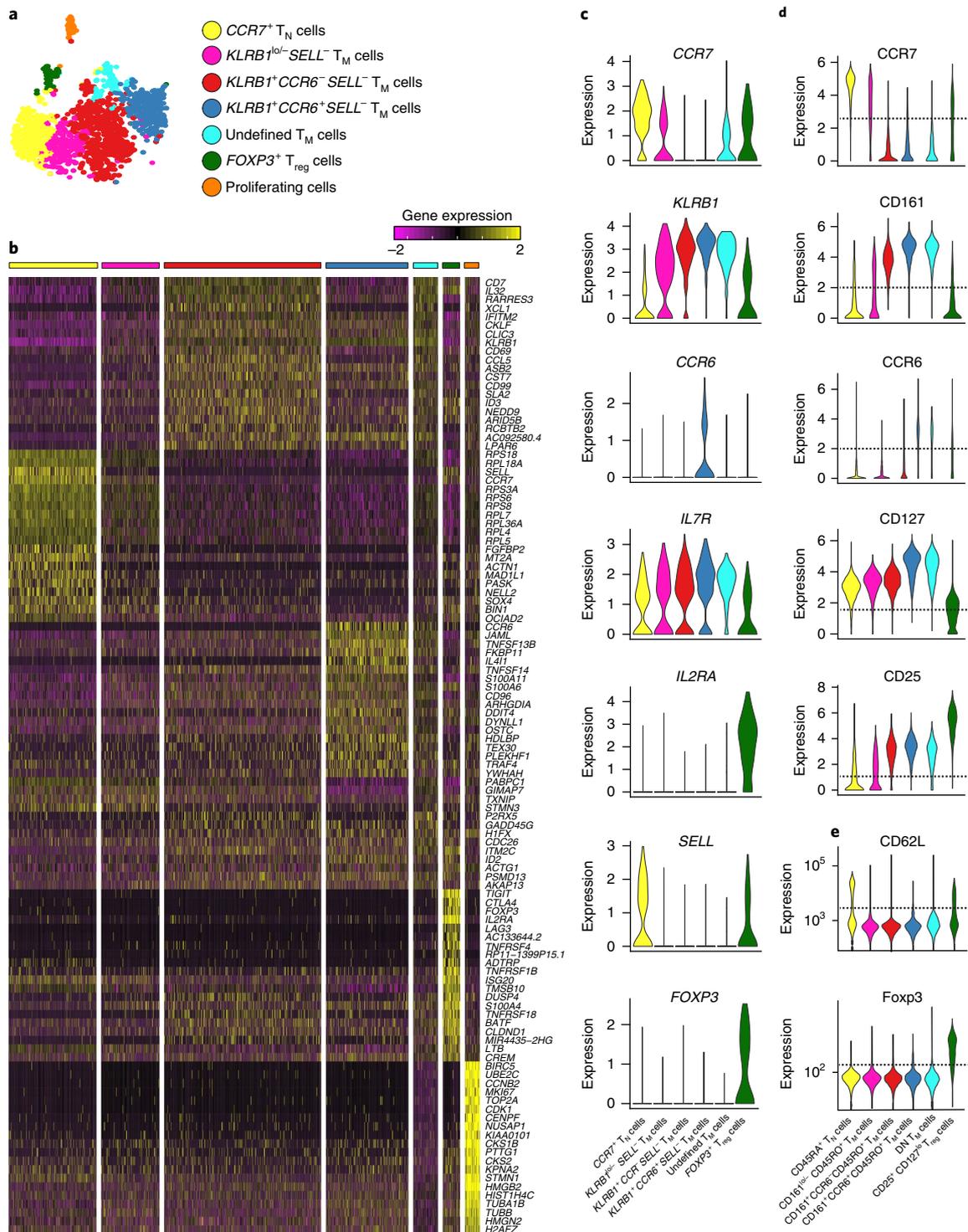


Fig. 2 | Single-cell RNA sequencing of fetal intestinal CD4⁺ T cells. **a**, t-SNE embedding of fetal intestinal CD4⁺ T cells ($n=1,804$) showing seven transcriptionally distinct clusters, including CCR7⁺ T_N cells ($n=358$), KLRB1^{hi}-SELL⁻ T_M cells ($n=237$), KLRB1⁺CCR6⁻SELL⁻ T_M cells ($n=640$), KLRB1⁺CCR6⁺SELL⁻ T_M cells ($n=336$), undefined T_M cells ($n=101$), FOXP3⁺ T_{reg} cells ($n=71$) and proliferating cells ($n=61$). Colors indicate different cell clusters. **b**, Heat map showing the normalized single-cell gene expression value (z score, purple to yellow scale) for the top 20 differentially upregulated genes in each identified cluster. Colors as shown in **a**. **c-e**, Expression of the indicated genes in each identified cluster at the RNA level (log-normalized) (**c**) and the protein level (**d,e**) analyzed by mass cytometry (CyTOF, ArcSinh5-transformed) (**d**) or flow cytometry (**e**), presented as violin plots. Dashed lines indicate background levels. Colors as shown in **a**.

of individual cells on the basis of single-cell expression profiles. On the basis of this analysis, CCR7⁺ T_N cells were separated from SELL⁻ T_M cells (Fig. 4c). When we clustered genes according to expression patterns along the pseudotime trajectory, cell-to-cell transitioning

could be explained by the kinetics of 1,376 variable genes, which formed three large modules (Fig. 4d). The first module contained 540 genes associated with CCR7⁺ T_N cells, including SELL, CCR7, CD27 and CD28 (Fig. 4d). The second module contained 453 genes,

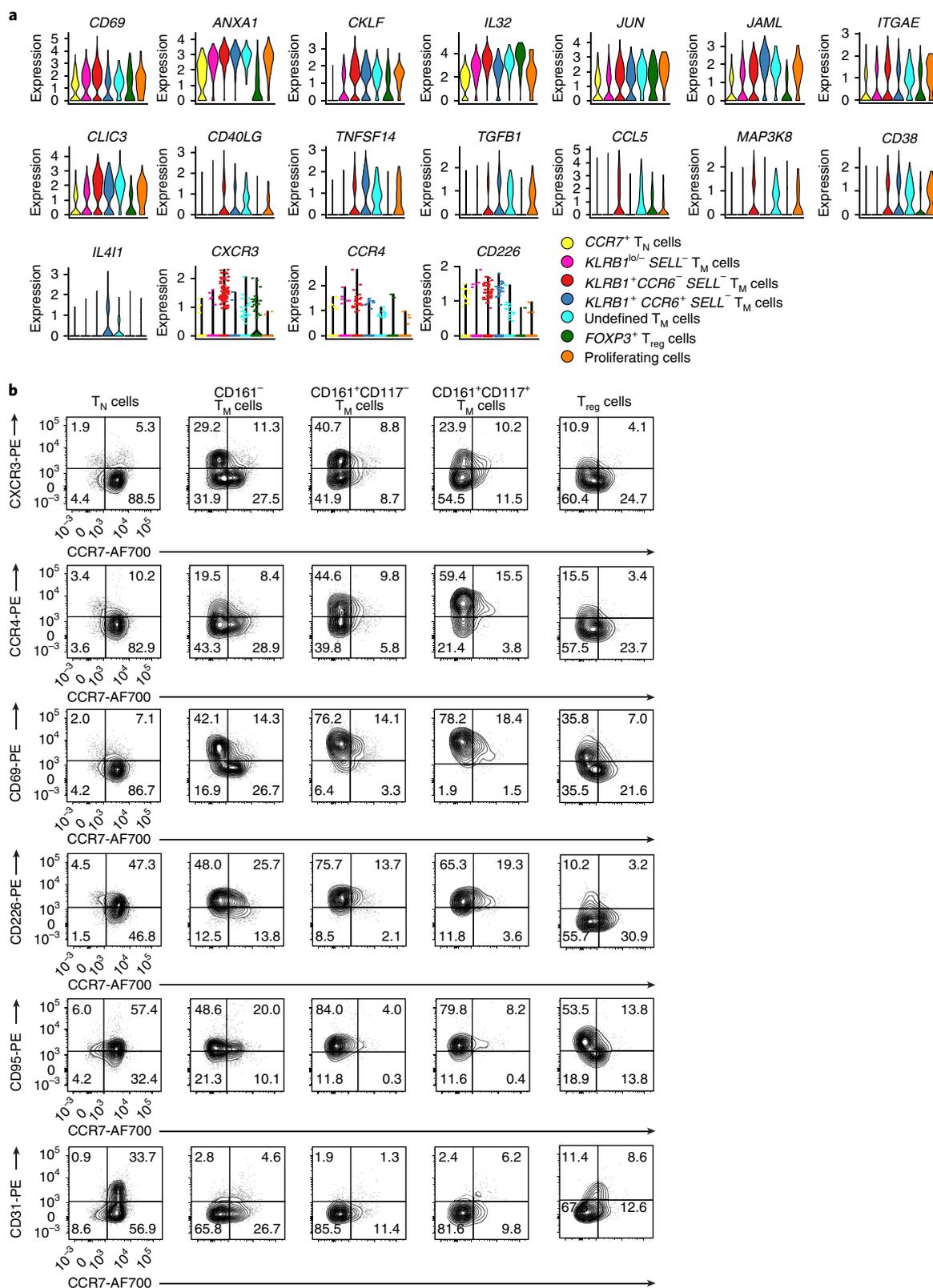


Fig. 3 | Targeted analysis of memory-like fetal intestinal CD4⁺ T cells. a, Expression (log-normalized) of the indicated genes in distinct CD4⁺ T cell clusters as determined by single-cell RNA-seq analysis, presented as violin plots. Colors indicate different cell clusters. **b**, Biaxial plots showing expression of CXCR3, CCR4, CD69, CD226, CD95 and CD31 versus CCR7 on the indicated CD4⁺ T cell clusters analyzed by flow cytometry. Data represent two or three independent experiments.

many of which were associated with an ongoing transcriptional program, such as *RPL21*, *RPS2* and *RPLP1*. The highest activity of this transcriptional gene expression profile coincided with the

transition of cells with a *CCR7*⁺ T_N phenotype into cells with a *SELL*⁻ T_M phenotype (Fig. 4c,d). The third module contained 383 genes (Fig. 4d), 106 of which were associated with cellular activation and

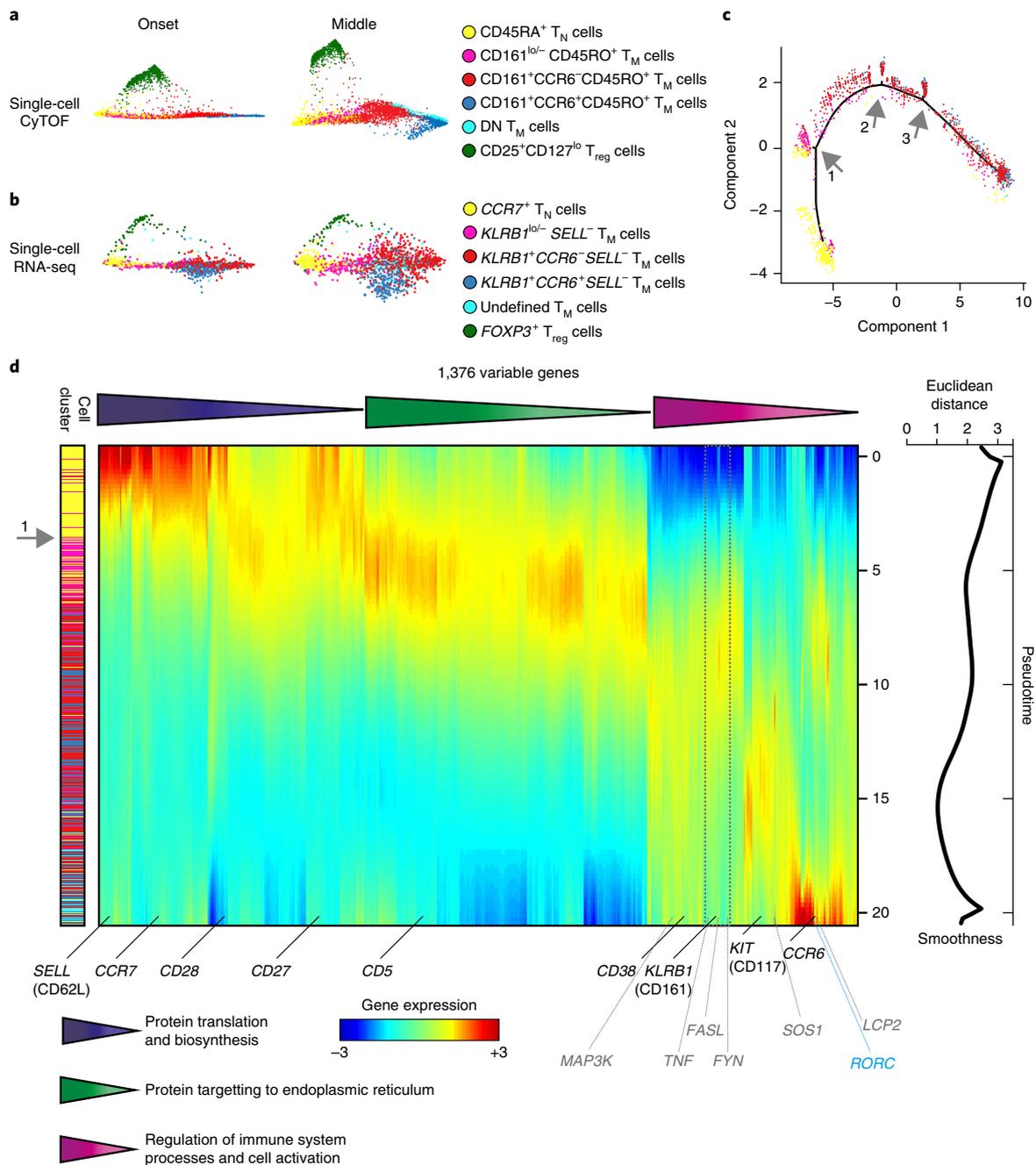


Fig. 4 | Single-cell trajectory analysis of fetal intestinal CD4⁺ T cells. **a, b**, t-SNE embeddings of all fetal intestinal CD4⁺ T cells analyzed by mass cytometry ($n=10,436$) (**a**) and single-cell RNA-seq ($n=1,743$) (**b**) at the onset and at the middle of the t-SNE computation. Colors indicate different cell clusters. **c**, A single-cell trajectory from the RNA-seq data (excluding T_{reg} cells and proliferating cells) recovered by pseudotime analysis. Colors indicate different cell clusters as shown in **b**. Gray arrows indicate three small branches. **d**, Three kinetic modules of pseudotime-dependent genes ($n=1,376$) depicted in a log-variance-stabilized expression heat map, indicating gene-enriched biological processes. Genes confirmed by mass cytometry and flow cytometry are denoted by black labels, and genes involved in TCR signaling are denoted by gray labels. The dashed gray box indicates the coordinated expression profile of *TNF*, *FASL* and *FYN*. Euclidean distance values comparing gene expression profiles for each ordered pair of neighboring cells along the pseudotime trajectory are shown in the graph (right).

regulation of the immune system (Supplementary Fig. 5a), while 133 encoded proteins known to interact physically with each other (Supplementary Fig. 5b). In addition, 23 genes in module 3 could be assigned to cytokine or chemokine receptor pathways, including *CCL20* and its receptor *CCR6*, the interferon receptor *IFNGR1*, TNF family members and IL-1 and IL-17 receptors (Supplementary Fig. 5b). Several signaling cascades were also represented in module 3, including the MAPK, TNF, IL-17 and TCR signaling pathways (*FYN*,

LCP2, *SOS1*, *MAP3K8* kinase, *FASL* and *TNF*) (Fig. 4d). The helper T_H17-associated gene *RORC* was expressed in module 3 (Fig. 4d). In addition, the dynamic expression profiles of *FYN*, *FASL* and *TNF* clustered tightly with *KLRB1* (CD161) (Fig. 4d) at the point in the pseudotime trajectory where *CCR7*⁺ T_N cells were aligned next to *SELL*⁻ T_M cells (Fig. 4d). Finally, we quantified the smoothness of cell-to-cell transitioning on the basis of gene expression changes along the trajectory, which showed that the pseudotime trajectory

was most uncertain at the beginning and toward the end, but quite robust in the middle, where $CCR7^+$ T_N cells were aligned next to $SELL^-$ T_M cells (Fig. 4d). In total, these results identified temporal patterns of gene expression along the single-cell trajectory that are compatible with the transition of cells displaying a T_N phenotype into cells with a T_M phenotype.

TCR analysis reveals clonal expansion of fetal $CD4^+$ T cells. Surface expression of CD5 correlates with TCR avidity^{23–26}. Because CD5 gene expression was upregulated in T_M cells compared with T_N cells, we quantified CD5 expression on all identified fetal intestinal $CD4^+$ T cell subsets using flow cytometry. We observed that all the $CD4^+$ T cell subsets expressed CD5 (Fig. 5a), but that the median fluorescence intensity was higher in $CD161^- T_M$, $CD161^+CD117^- T_M$ and $CD161^+CD117^+ T_M$ cells and lower in $CD25^+CD127^{lo} T_{reg}$ cells and $CD45RA^+ T_N$ cells (Fig. 5a–c), suggesting that cells with a T_M phenotype express TCR with a higher avidity compared with T_N cells.

Next, we evaluated the TCR clonotypic architecture of flow-sorted fetal intestinal $CD45RA^+ T_N$, $CD45RO^+ T_M$ and $CD25^+CD127^{lo} T_{reg}$ subpopulations. Analysis of the TCR β rearrangements in a single fetal intestine indicated limited overlap among the distinct subpopulations, most of which were highly polyclonal (data not shown). Distinct clonotypes were expanded among $CD45RO^+ T_M$ cells compared with $CD45RA^+ T_N$ cells (Supplementary Fig. 6a). We then used a quantitative high-throughput approach for deep sequencing of TCR α and TCR β rearrangements in all identified fetal intestinal $CD4^+$ T cell subsets isolated from two additional fetal intestines (Supplementary Table 2). Post-analysis of the repertoires obtained was conducted using VDJtools²⁷. As expected, all T_M subpopulations showed greater clonality compared with the T_N subpopulation (Fig. 5d,e). The averaged characteristics of CDR3 length, added N nucleotides and physicochemical characteristics of the five amino acid residues located in the middle of the CDR3 loop, which are most likely to contact the peptide–MHC complex²⁸, also differed among all subpopulations (Fig. 5f). The latter analysis included the averaged statistical potential of the CDR3 loop with respect to epitope interactions, comprising the estimated ‘energy’ of the interaction with a random epitope²⁹, the ‘strength’ of the interaction (derivative of ‘energy’, VDJtools²⁷), hydrophobicity (Kidera factor 4)^{30,31} and ‘volume’ (values from http://www.imgt.org/IMGTeducation/Aide-memoire/_UK/aminoacids/IMGTclasses.html). These analyses provided no evidence for clonal expansion of $CD4^+$ T cells as a function of intrinsically high TCR avidities for self-derived peptide–MHC complexes (Fig. 5d–f), suggesting indirectly that antigen-specific interactions triggered clonal selection of $CD4^+$ T cells from the T_N cell pool. Analysis of V–J segment use (Jensen–Shannon divergence; Supplementary Fig. 6b) and overlaps among repertoires in terms of the weighted proportion of shared TCR β clonotypes revealed tightly clustered technical replicates and clearly distinguished all subpopulations of $CD4^+$ T cells (Fig. 5g). At the same time, the $CD161^- T_M$, $CD161^+CD117^- T_M$ and $CD161^+CD117^+ T_M$ cells clustered similarly in each fetus, with minimum overlap with the $CD45RA^+ T_N$ and $CD25^+CD127^{lo} T_{reg}$ cells (Fig. 5g). Analysis of the clonal overlap of amino acid CDR3 repertoires between the same populations in the two fetal intestines revealed that the $CD161^- T_M$, $CD161^+CD117^- T_M$ and $CD161^+CD117^+ T_M$ populations displayed much stronger overlap than the $CD45RA^+ T_N$ and $CD25^+CD127^{lo} T_{reg}$ $CD4^+$ T cells (Supplementary Fig. 6c), which could be explained by TCR selection due to exposure to similar foreign antigens. Finally, although the majority of the TCR repertoire was specific for each population, up to 20% of the T cell clones were shared between the $CD45RA^+ T_N$ and the three $CD45RO^+ T_M$ cell populations (Supplementary Fig. 6d), suggesting a potential clonal relationship between $CD45RA^+ T_N$ and $CD45RO^+ T_M$ cells. These results indicate that avidity-based, clonotype-specific expansion of the T_N pool was associated with

T_M formation and confirmed the close relationship between $CD161^- T_M$, $CD161^+CD117^- T_M$ and $CD161^+CD117^+ T_M$ cells.

Fetal $CD4^+$ T_M cells secrete proinflammatory cytokines. To determine the functional profiles of fetal intestinal $CD4^+$ T cells, we flow-sorted $CD3^+CD4^+$ T cells and measured expression of TNF, IL-2, IFN- γ , IL-4, granzyme B and IL-17A in $CD45RA^+ T_N$ cells and $CD117^-$ and $CD117^+ T_M$ cells after cross-linking CD3 and CD28. The activation marker CD154 (CD40L) was upregulated on all cells analyzed (Fig. 6a,b), indicating efficient stimulation. All three subpopulations secreted large amounts of TNF (Fig. 6a,b), but $CD117^- T_M$ cells and $CD117^+ T_M$ cells displayed the highest median fluorescence intensities (Supplementary Fig. 7a). Moreover, IL-2, IFN- γ , IL-4 and granzyme B were more commonly expressed in $CD117^- T_M$ and $CD117^+ T_M$ cells relative to $CD45RA^+ T_N$ cells (Fig. 6a,b). The majority of cytokine-producing $CD4^+$ T cells did not express Ki-67 (Supplementary Fig. 7b). Importantly, higher frequencies of IL-2⁺IFN- γ ⁺ cells were detected in the $CD117^- T_M$ and $CD117^+ T_M$ cells compared with the $CD45RA^+ T_N$ population (Supplementary Fig. 7c), suggesting greater polyfunctionality. Although the T_H17 -associated *RORC* gene was expressed by 1.3% of $KLRB1^+CCR6^+SELL^- T_M$ cells (Fig. 4d), IL-17A production was undetectable in all T_M cells. Thus, fetal intestinal $CD117^- T_M$ and $CD117^+ T_M$ cells deployed multiple effector functions reminiscent of classical $CD4^+$ T_M cells in response to TCR-mediated signal transduction and costimulation via CD28.

Fetal $CD4^+$ T cells are colocalized with antigen-presenting cells.

The single-cell RNA-seq analysis revealed a *MKI67*⁺ cluster of proliferating cells, together with high expression of the T_M cell-associated markers *KLRB1* (CD161) and *CD69* and low expression of the T_N cell-associated markers *CCR7* and *SELL* (CD62L) (Fig. 7a). Flow cytometry of fetal intestinal $CD4^+$ T cells indicated the presence of Ki-67⁺ cells, predominantly within the $CD45RO^+$ compartment (Fig. 7b). To assess the spatial distribution of $CD4^+$ T cells in situ, we employed a panel of 15 antibodies (Supplementary Table 3) combined with a DNA stain to perform imaging mass cytometry on tissue sections of four human fetal intestinal samples. Stains for collagen I and smooth muscle actin were used to visualize the extracellular matrix of the basement membrane, and the epithelium and lamina propria were distinguished as vimentin⁺E-cadherin⁺ and vimentin⁺E-cadherin⁻, respectively (Fig. 7c,d). Most $CD4^+$ T cells localized to the lamina propria (Fig. 7c,d). Differential expression of *CD45RA* further allowed discrimination of $CD45RA^+ T_N$ (Fig. 7c,d) from $CD45RA^- T_M$ cells in the lamina propria (Fig. 7c,d). In addition, all $CD4^+$ T cells expressed CD38, whereas only some $CD4^+$ T cells expressed CD69 (Fig. 7d). Using a second panel comprising ten antibodies (Supplementary Table 3), we found that $CD4^+$ T cells frequently colocalized with $CD163^+HLA-DR^+$ APCs (Fig. 7e). Moreover, the single-cell RNA-seq analysis of fetal intestinal cells revealed two clusters of cells displaying high expression of gene transcripts encoding HLA-DR, CD74 (HLA-class II invariant chain), inhibitory molecule PD-L1 (CD274), CD80 and CD86, typically found in APCs. Moreover, these APCs expressed gene transcripts encoding CD40, consistent with an activated phenotype (Fig. 7f), whereas stimulated fetal intestinal $CD4^+$ T cells expressed CD40L (CD154) (Fig. 6b). In addition, 25.8% of APCs had high expression of *CCR7*, potentially enabling migration to the mesenteric lymph nodes (Fig. 7f). Collectively, these results indicated the existence of $CD4^+$ T_M cells in the fetal intestine, many of which colocalized in the lamina propria with activated $CD163^+HLA-DR^+$ APCs.

Discussion

Here we used mass cytometry and single-cell RNA-seq to characterize $CD4^+$ T cells in the human fetal intestine. Mass cytometry revealed three major populations of fetal intestinal $CD4^+$ T cells

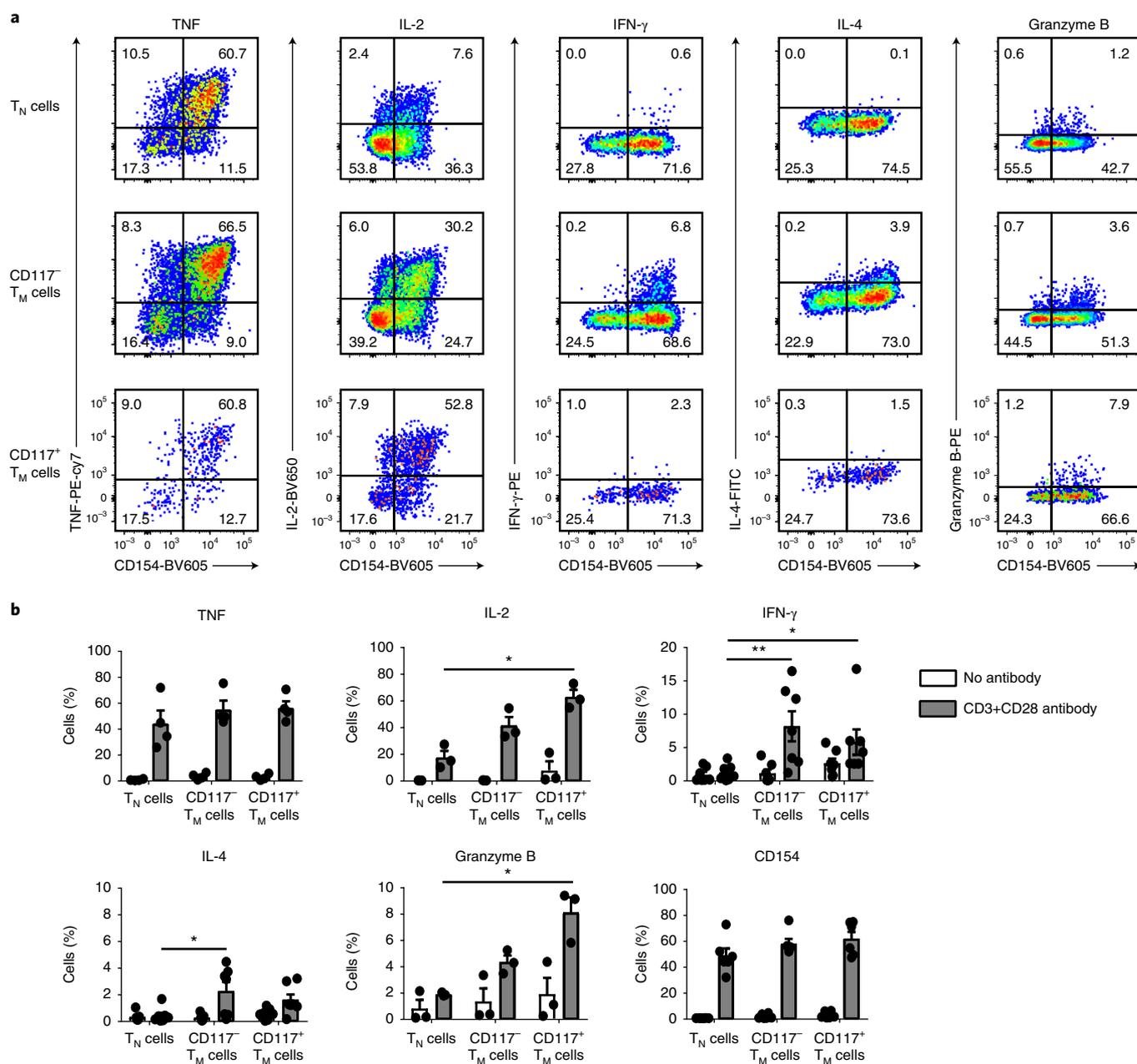


Fig. 6 | Functional profiling of fetal intestinal CD4⁺ T cells. **a, b**, Purified fetal intestinal CD4⁺ T cells were treated with a control antibody or stimulated with anti-CD3 and anti-CD28 for 4 hours. Intracellular expression of TNF, IL-2, IFN- γ , IL-4, granzyme B and CD154 was determined for each subpopulation by flow cytometry. The biaxial plots show data from one representative experiment after stimulation with anti-CD3 and anti-CD28 (**a**) and the bar charts show data for each subpopulation from each fetal intestine (**b**) (TNF: $n=4$ samples in two independent experiments; IL-2 and granzyme B: $n=3$ samples in two independent experiments; IFN- γ , IL-4 and CD154: $n=7$ samples in four independent experiments). Error bars indicate mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, Kruskal-Wallis test with Dunn's test for multiple comparisons.

(T_N, T_M and T_{reg} cells), that could be further distinguished into eight distinct cell clusters that displayed additional heterogeneity. These cell clusters were present in seven human fetal intestines, suggesting a physiologically robust immune composition. Single-cell RNA-seq revealed the presence of seven CD4⁺ T cell subpopulations, five of which displayed phenotypic overlap with the mass cytometry-defined CD4⁺ T cell subpopulations. We used computational tools to construct putative CD4⁺ T cell differentiation trajectories. Using adapted t-SNE³², we obtained remarkably similar trajectories for the mass cytometry and RNA-seq data. We identified three distinct gene expression modules along the differentiation trajectory that corresponded to an increase in gene translation and

subsequent activation of immune-related genes. In addition, high-throughput TCR sequencing indicated clonal expansions within the CD4⁺ T_M cell pool, consistent with the evidence for cell proliferation within the CD45RO⁺ T_M pool that was obtained at both the messenger RNA and protein level. Moreover, CD4⁺ T_M cells secreted higher amounts of proinflammatory cytokines on TCR triggering compared with CD4⁺ T_N cells. Finally, fetal intestinal CD4⁺ T_M cells displayed a tissue-resident profile and were frequently found to colocalize with APCs in the lamina propria. Together, this suggests that clonotype-specific transcriptional programs regulated by antigen encounter underpinned the formation of CD4⁺ T_M cells in the fetal intestine.

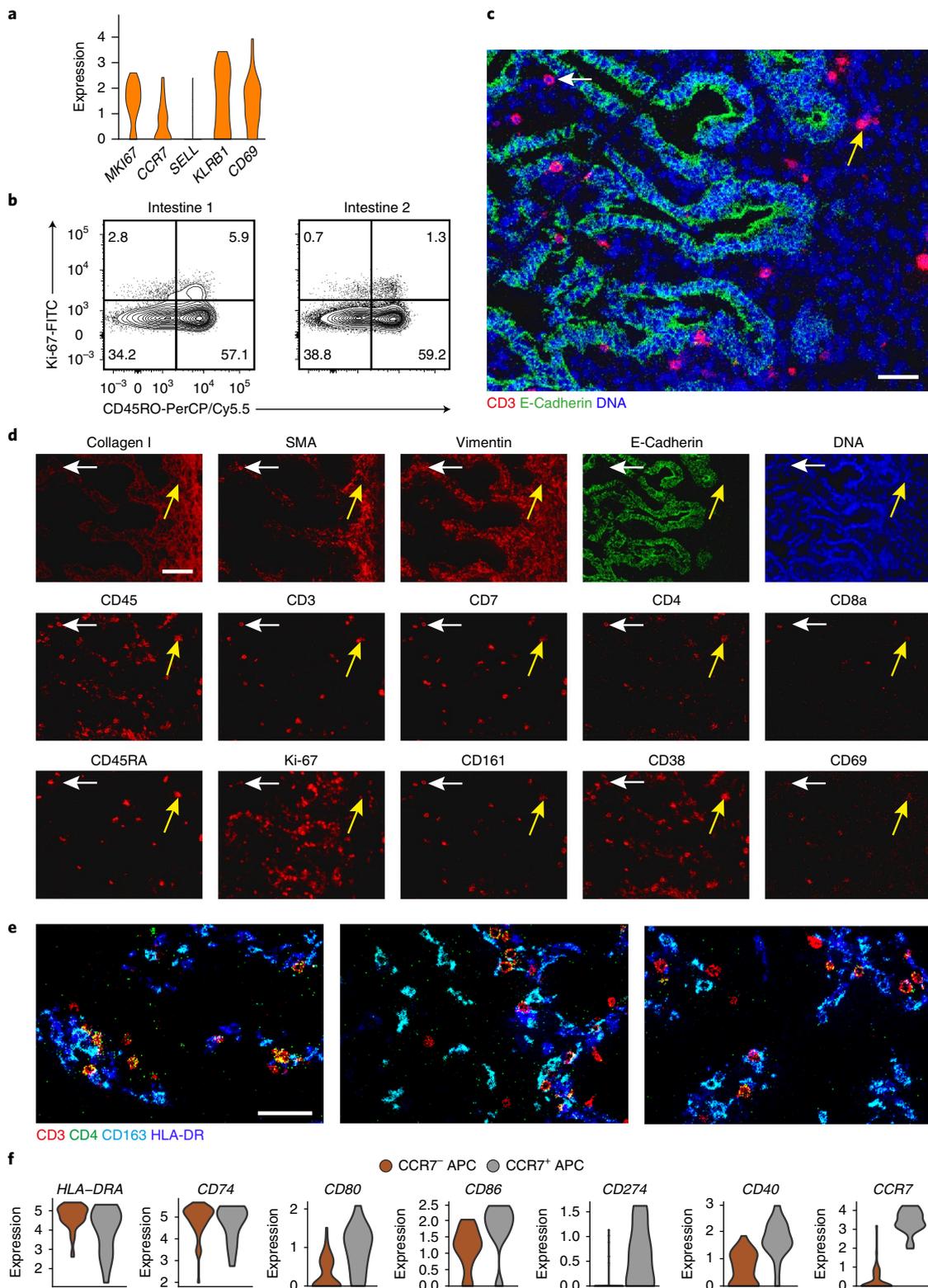


Fig. 7 | Characterization and spatial localization of fetal intestinal CD4⁺ T cells and APCs. **a**, Expression (log-normalized) of the indicated genes in proliferating fetal intestinal CD4⁺ T cells, presented as violin plot. **b**, Biaxial plots showing expression of Ki-67 versus CD45RO in the fetal intestinal CD4⁺ T cell compartment analyzed by flow cytometry. Data represent two independent experiments. **c**, Representative mass cytometry image of a fetal intestine showing the overlay of CD3 (red), E-cadherin (green) and DNA (blue). Scale bar, 100 μ m. **d**, Representative mass cytometry images of fetal intestines showing expression of the indicated stromal markers, immune markers, Ki-67 and DNA by the cells identified in **c**. Scale bar, 200 μ m (applies similarly to all the images in **d**). Yellow arrows: CD4⁺CD45RA⁺ T_N cells; white arrows: CD4⁺CD45RA⁻ T_M cells. **e**, Representative mass cytometry images of a fetal intestine showing the overlay of CD3 (red), CD4 (green), CD163 (cyan) and HLA-DR (blue). Scale bar, 50 μ m. Colors and scale bars are similar in all three panels. Data in **b-d** represent four independent samples in four independent experiments. **f**, Expression (log-normalized) of the indicated genes in two clusters of APCs, presented as violin plots. CCR7⁻ APCs ($n=49$), CCR7⁺ APCs ($n=17$).

T cells in umbilical cord and peripheral blood obtained from infants aged 2 months were reported to display a typical CD45RA⁺T_N phenotype³. The observation herein that a large pool of CD45RO⁺ cells with a tissue-resident profile populated the fetal intestine suggests the compartmentalization of the immune system early in life. In conjunction with the earlier finding that clonally expanded T cells were present in the fetal intestine, but virtually absent in other fetal organs⁴, our results further suggest that memory formation was driven by local exposure to foreign antigens. The observation that there is a substantial overlap in the amino acid CDR3 repertoires of the memory CD4⁺ T cells compartment in the two fetuses analyzed may indicate exposure to similar foreign antigens.

Approximately 50% of all fetal intestinal CD4⁺ T cells were CD161⁺ and transcriptionally distinct from their CD161⁻ counterparts, consistent with a recent study³³. The kinetics of *KLRB1* (CD161) expression was preceded by increased expression of CD5 and coincided with increased expression of several TCR signaling genes, including *FYN*, *FASL* and *TNF*, suggesting a coordinated program of transcription. Of note, CD161 was identified as a costimulatory molecule in the context of TCR stimulation³³.

Although the mass cytometry and RNA-seq data were largely compatible, there were exceptions. For example, coexpression of *CCR6* and *KIT* among *KLRB1*⁺*CCR6*⁺*SELL*⁻ T_M cells was not reflected in the gene expression profiles. Conversely, expression of *ITGAE* (CD103) mRNA was not reflected by protein expression. These anomalies were probably attributable to discordant gene transcription and protein expression³⁴ and may also relate to differences in sensitivity of the techniques used.

The presence of a large population of T_N cells in the fetal intestine is in stark contrast to the predominance of T_M cells in the adult intestine. As the T_N cells expressed relatively high amounts of CD31, which demarcates recent thymic emigrants, our results indicate direct migration of recent thymic emigrants into the intestine^{35,36}. We propose that antigen-specific priming of T_N cells takes place in the mesenteric lymph nodes followed by migration of the resulting T_M cells to the lamina propria leading to a progressive loss of T_N cells. Similarly, memory formation is taking place in the CD8⁺ T cell compartment (not shown).

Distinct subpopulations of fetal intestinal T_{reg} cells were distinguished by several markers, including high expression of CD25 and Foxp3, and a lack of CD127. In line with previous results³⁷, approximately 50% of these cells expressed CD45RO, while the remainder expressed CD45RA. The CD45RA⁺ T_{reg} cells expressed TCRs with longer CDR3β loops, higher numbers of added N nucleotides and distinct physicochemical characteristics, suggesting higher affinities for self-antigens compared with CD45RO⁺ T_{reg} cells³⁸. The presence of oligoclonal T cell expansions in fetuses with autoimmune conditions associated with a genetic absence of T_{reg} cells indicate a key role for these cells in immune suppression in utero³⁹.

In conclusion, our study revealed a putative differentiation trajectory in the fetal intestinal CD4⁺ T cell compartment, consistent with the formation of T_M cells in utero, presumably as a consequence of exposure to foreign antigens. These could include non-inherited maternal HLA molecules⁴⁰ and pathogen-derived ligands, which could be derived from amniotic fluid^{8,9}. We propose that immune priming in the fetal intestine prepares the infant for the massive influx of bacteria that occurs immediately after birth, with anamnestic responses in situ facilitated by the colocalization of CD4⁺ T_M cells with APCs.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41590-018-0294-9>.

Received: 29 May 2018; Accepted: 4 December 2018;

Published online: 21 January 2019

References

- Janeway, C. A. Jr. Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb. Symp. Quant. Biol.* **54**(Pt 1), 1–13 (1989).
- Fan, X. & Rudensky, A. Y. Hallmarks of tissue-resident lymphocytes. *Cell* **164**, 1198–1211 (2016).
- Thome, J. J. C. et al. Early-life compartmentalization of human T cell differentiation and regulatory function in mucosal and lymphoid tissues. *Nat. Med.* **22**, 72–77 (2015).
- Bunders, M. J. et al. Memory CD4(+)CCR5(+) T cells are abundantly present in the gut of newborn infants to facilitate mother-to-child transmission of HIV-1. *Blood* **120**, 4383–4390 (2012).
- Zhang, X. et al. CD4 T cells with effector memory phenotype and function develop in the sterile environment of the fetus. *Sci. Transl. Med.* **6**, 238ra72 (2014).
- Stout, M. J. et al. Identification of intracellular bacteria in the basal plate of the human placenta in term and preterm gestations. *Am. J. Obstet. Gynecol.* **208**, 226.e1–7 (2013).
- Aagaard, K. et al. The placenta harbors a unique microbiome. *Sci. Transl. Med.* **6**, 237ra65 (2014).
- Collado, M. C., Rautava, S., Aakko, J., Isolauri, E. & Salminen, S. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. *Sci. Rep.* **6**, e90784 (2016).
- DiGiulio, D. B. Diversity of microbes in amniotic fluid. *Semin. Fetal Neonatal Med.* **17**, 2–11 (2012).
- Ardissone, A. N. et al. Meconium microbiome analysis identifies bacteria correlated with premature birth. *PLoS ONE* **9**, e90784 (2014).
- Hornef, M. & Penders, J. Does a prenatal bacterial microbiota exist? *Mucosal Immunol.* **10**, 598–601 (2017).
- Li, N. et al. Mass cytometry reveals innate lymphoid cell differentiation pathways in the human fetal intestine. *J. Exp. Med.* **215**, 1383–1396 (2018).
- van Unen, V. et al. Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types. *Nat. Commun.* **8**, 1740 (2017).
- Maaten, L. vander & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
- Höltt, T. et al. Cytosplore: interactive immune cell phenotyping for large single-cell datasets. *Comput. Graph. Forum* **35**, 171–180 (2016).
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
- Kimmig, S. et al. Two subsets of naive T helper cells with distinct T cell receptor excision circle content in human adult peripheral blood. *J. Exp. Med.* **195**, 789–794 (2002).
- Angerer, P. et al. destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241–1243 (2015).
- Samusik, N., Good, Z., Spitzer, M. H., Davis, K. L. & Nolan, G. P. Automated mapping of phenotype space with single-cell data. *Nat. Methods* **13**, 493–496 (2016).
- Jolliffe, I. Principal component analysis. in *International Encyclopedia of Statistical Science* (ed. Lovric, M) Ch. 427, 1094–1096 (Springer, Berlin, 2011).
- Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
- Qiu, X. et al. Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**, 309–315 (2017).
- Azzam, H. S. et al. CD5 expression is developmentally regulated by T cell receptor (TCR) signals and TCR avidity. *J. Exp. Med.* **188**, 2301–2311 (1998).
- Mandl, J. N., Monteiro, J. P., Vriskoop, N. & Germain, R. N. T. Cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity* **38**, 263–274 (2013).
- Persaud, S. P., Parker, C. R., Lo, W.-L., Scott Weber, K. & Allen, P. M. Intrinsic CD4 T cell sensitivity and response to a pathogen are set and sustained by avidity for thymic and peripheral complexes of self peptide and MHC. *Nat. Immunol.* **15**, 266–274 (2014).
- Fulton, R. B. et al. The TCR's sensitivity to self peptide-MHC dictates the ability of naive CD8(+) T cells to respond to foreign antigens. *Nat. Immunol.* **16**, 107–117 (2015).
- Shugay, M. et al. VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput. Biol.* **11**, e1004503 (2015).
- Egorov, E. S. et al. The changing landscape of naive T cell receptor repertoire with human aging. *Front. Immunol.* **9**, 1618 (2018).
- Miyazawa, S. & Jernigan, R. L. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J. Mol. Biol.* **256**, 623–644 (1996).

30. Kidera, A., Konishi, Y., Oka, M., Ooi, T. & Scheraga, H. A. Statistical analysis of the physical properties of the 20 naturally occurring amino acids. *J. Protein Chem.* **4**, 23–55 (1985).
31. Rackovsky, S. Global characteristics of protein sequences and their implications. *Proc. Natl Acad. Sci. USA* **107**, 8623–8626 (2010).
32. Pezzotti, N. et al. Approximated and user steerable tSNE for progressive visual analytics. *IEEE Trans. Vis. Comput. Graph.* **23**, 1739–1752 (2017).
33. Fergusson, J. R. et al. CD161 defines a transcriptional and functional phenotype across distinct human T cell lineages. *Cell Rep.* **9**, 1075–1088 (2014).
34. Edfors, F. et al. Gene-specific correlation of RNA and protein levels in human cells and tissues. *Mol. Syst. Biol.* **12**, 883 (2016).
35. McFarland, R. D., Douek, D. C., Koup, R. A. & Picker, L. J. Identification of a human recent thymic emigrant phenotype. *Proc. Natl Acad. Sci. USA* **97**, 4215–4220 (2000).
36. Staton, T. L. et al. CD8⁺ recent thymic emigrants home to and efficiently repopulate the small intestine epithelium. *Nat. Immunol.* **7**, 482–488 (2006).
37. Michaelsson, J., Mold, J. E., McCune, J. M. & Nixon, D. F. Regulation of T cell responses in the developing human fetus. *J. Immunol.* **176**, 5741–5748 (2006).
38. Feng, Y. et al. A mechanism for expansion of regulatory T-cell repertoire and its role in self-tolerance. *Nature* **528**, 132–136 (2015).
39. Allenspach, E. J. et al. Absence of functional fetal regulatory T cells in humans causes in utero organ-specific autoimmunity. *J. Allergy Clin. Immunol.* **140**, 616–619.e7 (2017).
40. Gomez de Agüero, M. et al. The maternal microbiota drives early postnatal innate immune development. *Science* **351**, 1296–1302 (2016).

Acknowledgements

We thank the Center for Contraception, Abortion and Sexuality (Leiden and The Hague) for collection and provision of fetal material, K. Lodder, T. van Herwaarden, M. Bialecka and F. Wang for dissection of fetal tissues, and S. L. Kloet for assistance with single-cell

RNA sequencing. This research was supported by Leiden University Medical Center (N.L., V.v.U., N.G., F.K.), the Netherlands Organization for Scientific Research (NWO Applied Technical Sciences grant no. 12721 and ZonMW grant no. 91112008) (T.A., T.H., J.E., B.P.F.L.), the Russian Science Foundation (grant no. 16-15-00149) (S.A.K., E.S.E., M.I., O.V.B., D.M.C.), the Wellcome Trust (grant no. 100326/Z/12/Z) (K.L., J.E.M., D.A.P.), the European Commission under an MSCA-ITN award (grant no. 675743/ISPIC) (T.A.), and the China Scholarship Council (N.L. and N.G.).

Author contributions

N.L., V.v.U. and F.K. conceived the study and wrote the manuscript. N.L. performed most of the experiments with help from V.v.U. and N.G. N.L. performed most of the data analyses with help from V.v.U., T.A. and B.P.F.L. S.A.K., K.L., J.E.M., E.S.E., M.I., D.M.C., O.V.B. and D.A.P. performed TCR repertoire analyses. D.A.P. revised the manuscript. N.F.C.d.M. helped with imaging mass cytometry experiments. T.H., V.v.U., J.E. and B.P.F.L. developed Cytosplore. S.M.C.d.S.L. provided human fetal tissues. All authors discussed the results and helped prepare the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41590-018-0294-9>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to F.K.

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Sample processing and cell isolation. Fetal tissues were obtained from elective abortions with informed consent. The gestational age ranged from 14 to 22 weeks. Intestines were separated from mesentery, cut into small pieces, embedded in optimal cutting temperature compound and snap-frozen in isopentane. The remaining intestines were used for single-cell isolation as described previously¹². Briefly, fetal intestines were cleared of meconium, cut into fine pieces, treated with 1 mM dithiothreitol (Fluka) for 2 × 10 min (replacing buffer) at room temperature, and then incubated with 1 mM ethylenediaminetetraacetic acid (Merck) for 2 × 1 h (replacing buffer) at 37 °C under rotation to separate the epithelium from the lamina propria. To obtain single-cell suspensions from the lamina propria, the intestines were rinsed with Hank's balanced salt solution (ThermoFisher Scientific), incubated with 10 U ml⁻¹ collagenase IV (Worthington) and 200 µg ml⁻¹ DNase I grade II (Roche Diagnostics) overnight at 37 °C, and filtered through a 70 µm nylon mesh. Isolated cells were then further purified with a Percoll gradient (GE Healthcare). Fetal liver and spleen tissues were cut into small pieces and filtered through a 70 µm nylon cell strainer and the immune cells were isolated with Ficol-Paque density gradient (provided by the pharmacy of Leiden University Medical Center). All the isolated cells were stored in liquid nitrogen. Study approval was granted by the Medical Ethics Commission of Leiden University Medical Centre (protocol P08.087). All experiments were conducted in accordance with local ethical guidelines and the principles of the Declaration of Helsinki.

Cell suspension mass cytometry. Antibodies used for cell suspension mass cytometry are listed in Supplementary Table 1. Purified antibodies lacking carrier protein were conjugated with metal reporters by using a MaxPar X8 Antibody Labeling Kit (Fluidigm). Procedures for antibody staining and data acquisition were described previously⁴¹. Briefly, cells from fetal intestines were incubated with 5 µM Cell-ID Intercalator-¹⁰⁹Rh (Fluidigm) for 15 min at room temperature and then stained with a cocktail of metal-conjugated antibodies for 45 min at room temperature. After washing, cells were incubated with 125 nM Cell-ID Intercalator-Ir (Fluidigm) in MaxPar Fix and Perm Buffer (Fluidigm) overnight at 4 °C. Data were acquired using a CyTOF 2 mass cytometer (Fluidigm) and normalized using EQ Four Element Calibration Beads with the reference EQ Passport P13H2302 (Fluidigm).

Imaging mass cytometry. Antibodies used for imaging mass cytometry are listed in Supplementary Table 3. Purified antibodies lacking carrier protein were conjugated with metal reporters by using a MaxPar X8 Antibody Labeling Kit (Fluidigm). Snap-frozen human fetal intestinal biopsies were sectioned at a thickness of 5 µm and fixed by incubating with 1% paraformaldehyde for 5 min at room temperature followed by 100% methanol for 5 min at -20 °C. After fixation, tissue sections were washed in Dulbecco's phosphate-buffered saline (ThermoFisher Scientific) containing 0.5% bovine serum albumin and 0.05% Tween, rehydrated in additive-free Dulbecco's phosphate-buffered saline, washed again and blocked with Superblock Solution (ThermoFisher Scientific). Tissue sections were then stained with a cocktail of metal-conjugated antibodies overnight at 4 °C, washed and incubated with 125 nM Cell-ID Intercalator-Ir for 30 min at room temperature. After a further wash, tissue sections were dipped in Milli-Q water (Merck Millipore) for 1 min and dried for 20 min at room temperature. Data were acquired using a Hyperion imaging mass cytometer (Fluidigm) at a resolution of 1 µm, with settings aligned to company guidelines. The ablation frequency was 200 Hz, and the energy was 6 dB. Regions of interest were acquired at a size of 1 × 1 mm². All data were stored as MCD files and txt files.

Single-cell RNA sequencing. Single, live, CD8a⁻TCRγδ⁻CD4⁺ cells from the intestines of fetus no. 6 were sorted using a FACSAria III flow cytometer (BD Biosciences). Post-sort purity was 96.5%. Single-cell RNA sequencing was performed as described previously⁴². Briefly, cells combined with oil, reagents and beads were loaded on a Chromium Single Cell Controller (10x Genomics). Lysis and barcoded reverse transcription of polyadenylated mRNA from single cells were performed inside each gel bead emulsion. Next-generation sequencing libraries were prepared in a single bulk reaction, and transcripts were sequenced using a HiSeq4000 System (Illumina).

Integrated data analysis. For cell suspension mass cytometry, data from single, live, CD45⁺ cells, gated individually using Cytobank as shown in Supplementary Fig. 1a, were sample-tagged and hyperbolic-arcsinh-transformed with a cofactor of 5 using Cytosplore^{+HSNE} software¹³. The major immune lineages shown in Supplementary Fig. 1b were then identified at the overview level by performing a three-level HSNE analysis carried out with default parameters (perplexity: 30; iterations: 1,000). All t-SNE plots and Gaussian mean-shift clustering-derived cell clusters were generated in Cytosplore¹⁵. Hierarchical clustering of the phenotype heat map was created with Euclidean correction and average linkage clustering in Cytosplore^{+HSNE}. Violin plots for cytometry data were generated in R. Diffusion map plots for mass cytometry data were generated using the 'density' package in R. Single-cell force-directed layouts for mass cytometry data were generated using Vortex software¹⁹. For imaging mass cytometry, all images were generated using

MCD Viewer software v1.0.560 (Fluidigm). For single-cell RNA-seq, single-cell transcriptome sequencing data were processed using the single-cell RNA-seq package Seurat in R¹⁶. The Seurat object was generated by following the criteria that each gene was expressed by at least three cells and that at least 200 genes were expressed per cell. Data were further filtered on the basis of the parameters: (1) unique gene count per cell >200 and <2,000; and (2) mitochondrial percentage of all genes <0.05. After log-normalization, a PCA-reduction analysis (pca.compute = 20) was performed on the basis of the 2,174 variable genes across single cells. Next, graph-based clustering detection and a t-SNE algorithm were applied to the top 13 PCA dimensions. The resolution for cluster detection was 0.8. Heat maps, PCA plots, diffusion map plots and violin plots of the RNA-seq data were generated using the Seurat package. The t-SNE plots for RNA-seq data shown in Fig. 4b were generated in Cytosplore^{+HSNE}. Only genes with local standardization (>0.5) across all cells were taken into account. Bar graphs and dot plots (showing mean and s.d.) were generated in Prism (GraphPad). The pseudotime analysis shown in Fig. 4c,d was performed using the Monocle 2 toolkit in R as described previously²², excluding unrelated T_{reg} cells. Briefly, the single-cell trajectory was inferred using the dpFeature unsupervised procedure to identify variable genes, and the dimensions were reduced using t-SNE on the top high-loading principal components. The top 1,000 significant genes were selected as the ordering genes and reduced with the DDRTree method for the single-cell graph shown in Fig. 4c. Variable genes were selected at a significant false discovery rate of <10%, clustered by pseudotemporal expression patterns, and visualized in a heat map in Fig. 4d. We performed gene list enrichment analysis using ToppGene⁴³, gene interaction network analysis using the BioGrid interaction database⁴⁴ and gene pathway analysis using the Kyoto Encyclopedia of Genes and Genomes⁴⁵.

Flow cytometry. For surface staining, cells were incubated with fluorochrome-conjugated antibodies and human Fc block (BioLegend) for 30–45 min at 4 °C. For intracellular cytokine/CD154 staining, cells were stimulated with CD3/CD28-specific (2.5 µg ml⁻¹ each, BioLegend) or control antibodies (5 µg ml⁻¹, BioLegend) for 4 h at 37 °C. Brefeldin A (10 µg ml⁻¹, Sigma) was added for the final 3 h. Cells were then fixed/permeabilized using Fixation Buffer and Intracellular Staining Perm Wash Buffer (BioLegend). For intracellular Foxp3/Ki-67 staining, cells were prepared using a Foxp3 Staining Buffer Set (eBioscience). Electronic compensation was performed using individually stained CompBeads (BD Biosciences). Cells were acquired using an LSR II cytometer (BD Biosciences) or sorted using a FACSAria III flow cytometer (BD Biosciences) as shown in Supplementary Fig. 3a. Data were analyzed with FlowJo software v10 (Tree Star Inc.). The antibodies used in this study are listed in Supplementary Table 4.

TCR repertoire analysis. CD4⁺ T cell subsets were sorted according to the gating strategy shown in Supplementary Fig. 3a. For conventional sequencing, a total of 5,000 cells per subset was sorted directly into RNAlater (Applied Biosystems) using a FACSAria III flow cytometer (BD Biosciences). All expressed TCRβ rearrangements were amplified using a template-switch anchored RT-PCR, sequenced and analyzed as described previously⁴⁶. Gene use was determined according to the ImMunoGeneTics nomenclature⁴⁷.

For high-throughput sequencing, an average of 6,700 ± 2,000 cells per subset was sorted directly into RLT buffer (Qiagen) using a FACSAria III flow cytometer (BD Biosciences). Four volumes of TRIzol (Invitrogen) were then added to the RLT cell lysate. RNA was extracted according to the TRIzol Reagent User Guide. Unique molecular identifier-labeled 5' RACE TCRα and TCRβ complementary DNA libraries were prepared using a Human TCR Profiling Kit (MiLaboratory LLC). All extracted RNA was used for cDNA synthesis, and all synthesized cDNA was used for PCR amplification. Libraries were prepared in parallel using the same number of PCR cycles and sequenced in parallel using a 150 + 150 base pair MiSeq System (Illumina). This approach generated a total of 11,310,000 TCRα and TCRβ sequencing reads (250,000 ± 150,000 reads per library), from which 625,000 unique molecular identifier-labeled TCR cDNA molecules (13,500 ± 7,000 molecules per library) were extracted using MIGEC⁴⁸ and MiXCR⁴⁹ software with a threshold of at least two sequencing reads per unique molecular identifier. Each library contained an average of 3,500 ± 1,300 functional (in-frame, without stop codons) CDR3 nucleotide sequences. Averaged TCR repertoire characteristics weighted by clonotype size were analyzed using VDJtools software²⁷. Gene use was determined according to the ImMunoGeneTics nomenclature⁴⁷.

Statistics. Results are shown as mean ± s.e.m. The statistics tests used were two-tailed Wilcoxon matched-pairs signed-ranks test and Kruskal–Wallis test with Dunn's test for multiple group comparisons, as appropriate (after normality test). *P* ≤ 0.05 was considered to be statistically significant. All statistics were analyzed using GraphPad Prism7 software.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Mass cytometry data are available via Flow Repository (<https://flowrepository.org/id/FR-FCM-ZYRD>). Single-cell RNA-seq data are available via Gene Expression

Omnibus accession code [GSE122846](#). The remaining data that support the findings of this study are available from the corresponding author upon reasonable request.

References

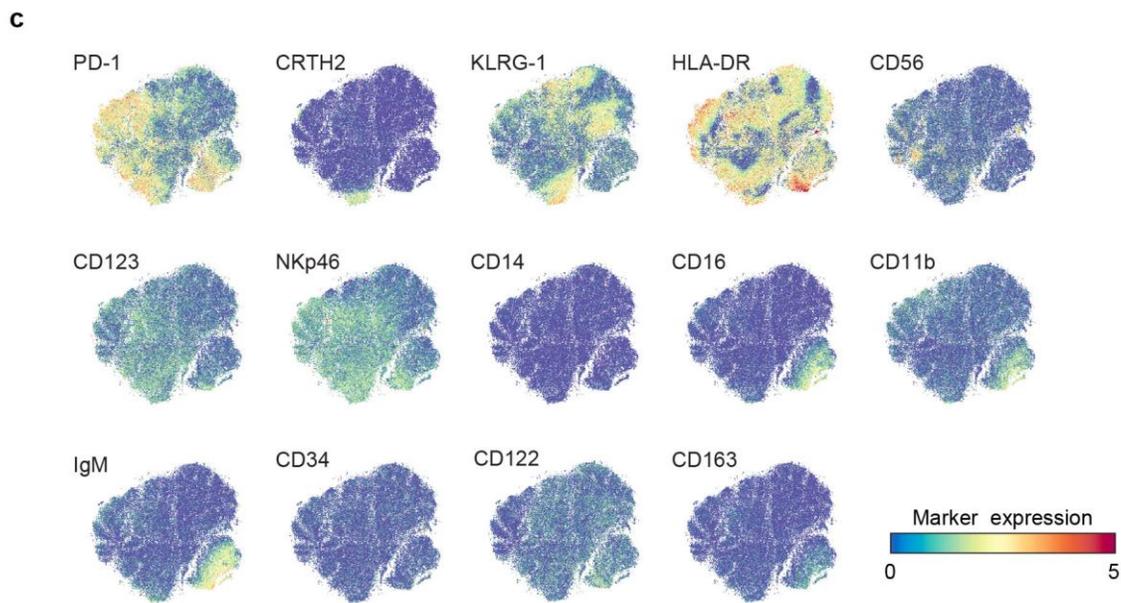
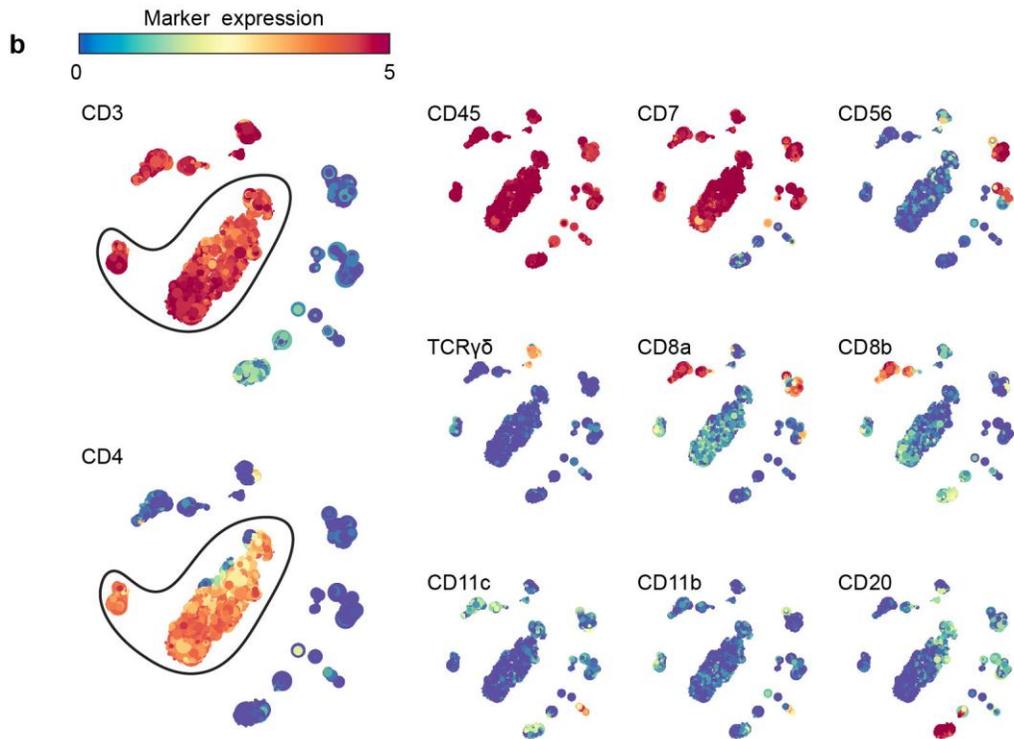
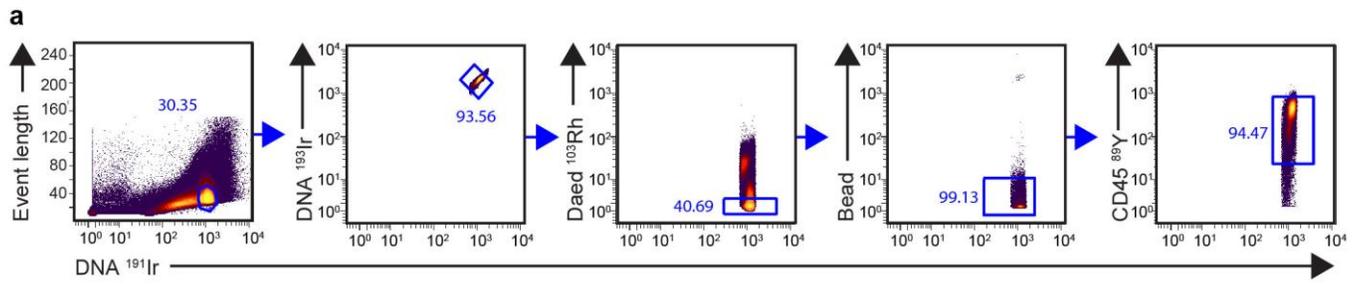
41. van Unen, V. et al. Mass cytometry of the human mucosal immune system identifies tissue- and disease-associated immune subsets. *Immunity* **44**, 1227–1239 (2016).
42. Zheng, G. X. Y. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
43. Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, W305–W311 (2009).
44. Bean, D. M. et al. esyN: network building, sharing and publishing. *PLoS ONE* **9**, e106035 (2014).
45. Ogata, H. et al. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **27**, 29–34 (1999).
46. Quigley, M. F., Almeida, J. R., Price, D. A. & Douek, D. C. Unbiased molecular analysis of T cell receptor expression using template-switch anchored RT-PCR. *Curr. Protoc. Immunol.* **94**, 10.33.1–10.33.16 (2011).
47. Lefranc, M.-P. et al. IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains. *Dev. Comp. Immunol.* **29**, 185–203 (2005).
48. Shugay, M. et al. Towards error-free profiling of immune repertoires. *Nat. Methods* **11**, 653–655 (2014).
49. Bolotin, D. A. et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380–381 (2015).

In the format provided by the authors and unedited.

Memory CD4⁺ T cells are generated in the human fetal intestine

Na Li¹, Vincent van Unen¹, Tamim Abdelaal^{2,3}, Nannan Guo¹, Sofya A. Kasatskaya^{4,5}, Kristin Ladell⁶, James E. McLaren⁶, Evgeny S. Egorov⁴, Mark Izraelson⁴, Susana M. Chuva de Sousa Lopes⁷, Thomas Höllt^{2,8}, Olga V Britanova⁴, Jeroen Eggermont⁹, Noel F. C. C. de Miranda¹⁰, Dmitriy M. Chudakov^{4,5,11,12,13,14}, David A. Price^{6,15}, Boudewijn P. F. Lelieveldt^{3,9} and Frits Koning^{1*}

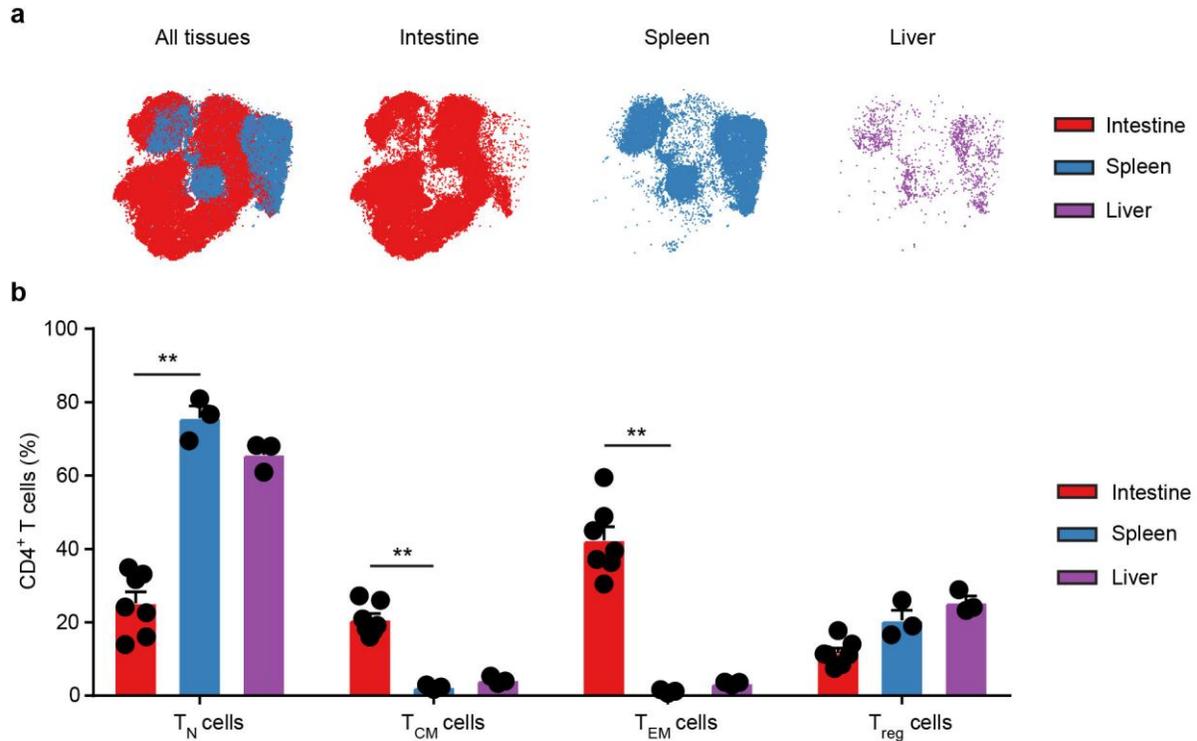
¹Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, the Netherlands. ²Leiden Computational Biology Center, Leiden University Medical Center, Leiden, the Netherlands. ³Department of Pattern Recognition and Bioinformatics Group, Delft University of Technology, Delft, the Netherlands. ⁴Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia. ⁵Centre for Data-Intensive Biomedicine and Biotechnology, Skolkovo Institute of Science and Technology, Moscow, Russia. ⁶Division of Infection and Immunity, Cardiff University School of Medicine, Cardiff, UK. ⁷Department of Anatomy and Embryology, Leiden University Medical Center, Leiden, the Netherlands. ⁸Computer Graphics and Visualization Group, Delft University of Technology, Delft, the Netherlands. ⁹Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands. ¹⁰Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands. ¹¹Central European Institute of Technology, Masaryk University, Brno, Czech Republic. ¹²Department of Molecular Technologies, Pirogov Russian National Research Medical University, Moscow, Russia. ¹³MiLaboratory LLC, Skolkovo Innovation Centre, Moscow, Russia. ¹⁴Privolzhsky Research Medical University, Nizhny Novgorod, Russia. ¹⁵Systems Immunity Research Institute, Cardiff University School of Medicine, Cardiff, UK. *e-mail: F.Koning@lumc.nl



Supplementary Figure 1

Mass cytometric analysis of fetal intestinal immune cells.

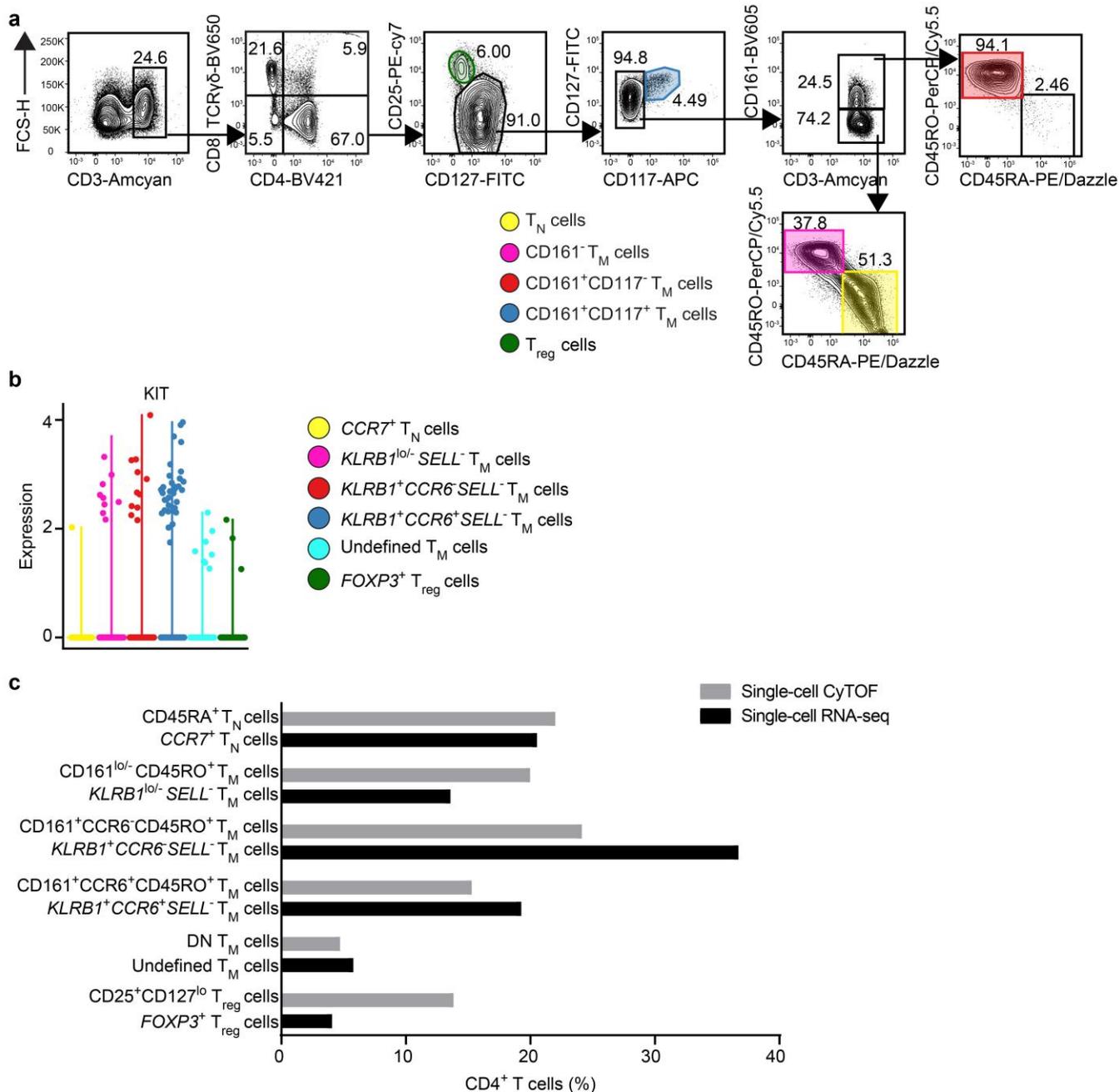
(a) Biaxial plots from one fetal intestine showing the gating strategy for single, live, CD45⁺ cells analyzed by mass cytometry (n = 7). (b) First-level HSNE embedding of CD45⁺ immune cells (n = 224,286) derived from fetal intestines (n = 7). Each dot represents a landmark. The size of each landmark is proportional to the number of cells that the landmark represents. Colors indicate ArcSinh5-transformed expression values. (c) t-SNE embedding showing all CD4⁺ T cells (n = 110,332) derived from fetal intestines (n = 7). Colors indicate the ArcSinh5-transformed expression values.



Supplementary Figure 2

Mass cytometric analysis of the fetal CD4⁺ T cell compartment across tissues.

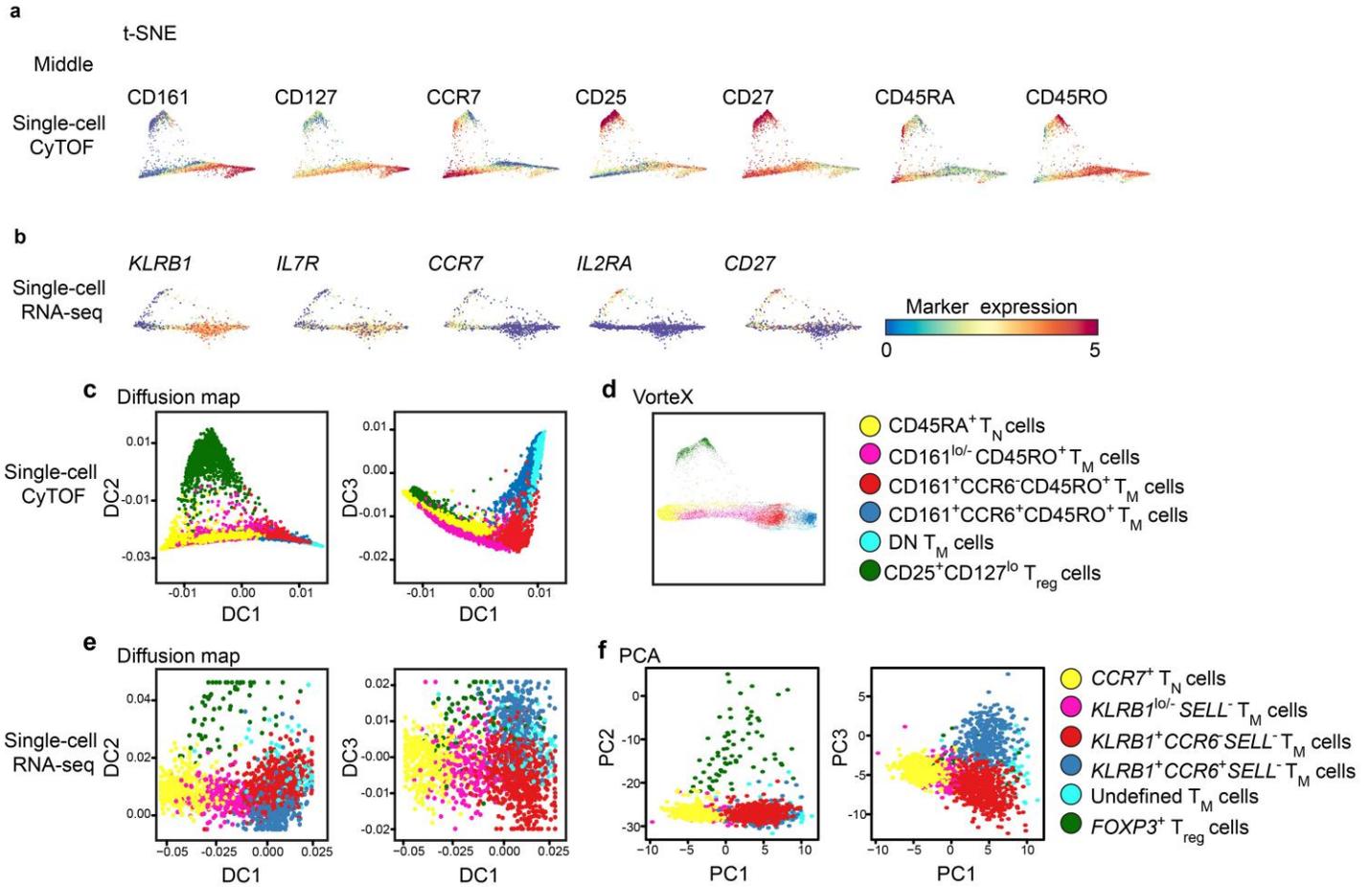
(a) t-SNE embedding of all CD4⁺ T cells ($n = 9.7 \times 10^4$) derived from fetal intestines ($n = 6.6 \times 10^4$ cells from 7 samples), fetal livers ($n = 1,530$ cells from 3 samples), and fetal spleens ($n = 3.0 \times 10^4$ cells from 3 samples). Colors indicate tissue origin. Downsampling was performed for samples with more than 10,000 CD4⁺ T cells. (b) Bar graphs showing the quantification of T_{reg} cells and T_N cells, T_{CM} cells, and T_{EM} cells among fetal CD4⁺ T cells across tissues. T_N cells: CD45RA⁺CCR7⁺; T_{CM} cells: CD45RA⁻CCR7⁺; T_{EM} cells: CD45RA⁻CCR7⁻. Fetal liver and spleen, $n = 3$ independent samples; fetal intestine, $n = 7$ independent samples. Error bars indicate mean \pm SEM. ** $p < 0.01$, Kruskal-Wallis test with Dunn's test for multiple comparisons.



Supplementary Figure 3

Identification of CD4⁺ T cell clusters in fetal intestines.

(a) Representative biaxial plots showing the gating strategy for T_N, CD161⁻ T_M, CD161⁺CD117⁻ T_M, CD161⁺CD117⁺ T_M and T_{reg} cell clusters derived from one fetal intestine analyzed by flow cytometry for expression of CD3, CD4, CD8a, CD25, CD45RA, CD45RO, CD117, CD127, CD161 and TCRγδ (n = 10). (b) Expression (log-normalized) of *KIT* as determined by single-cell RNA-seq analysis, presented as violin plots. Colors indicate different cell clusters. (c) Bar plot depicting cell frequencies for clusters identified by mass cytometry and RNA-seq.



Supplementary Figure 4

Single-cell trajectories of fetal intestinal CD4⁺ T cells.

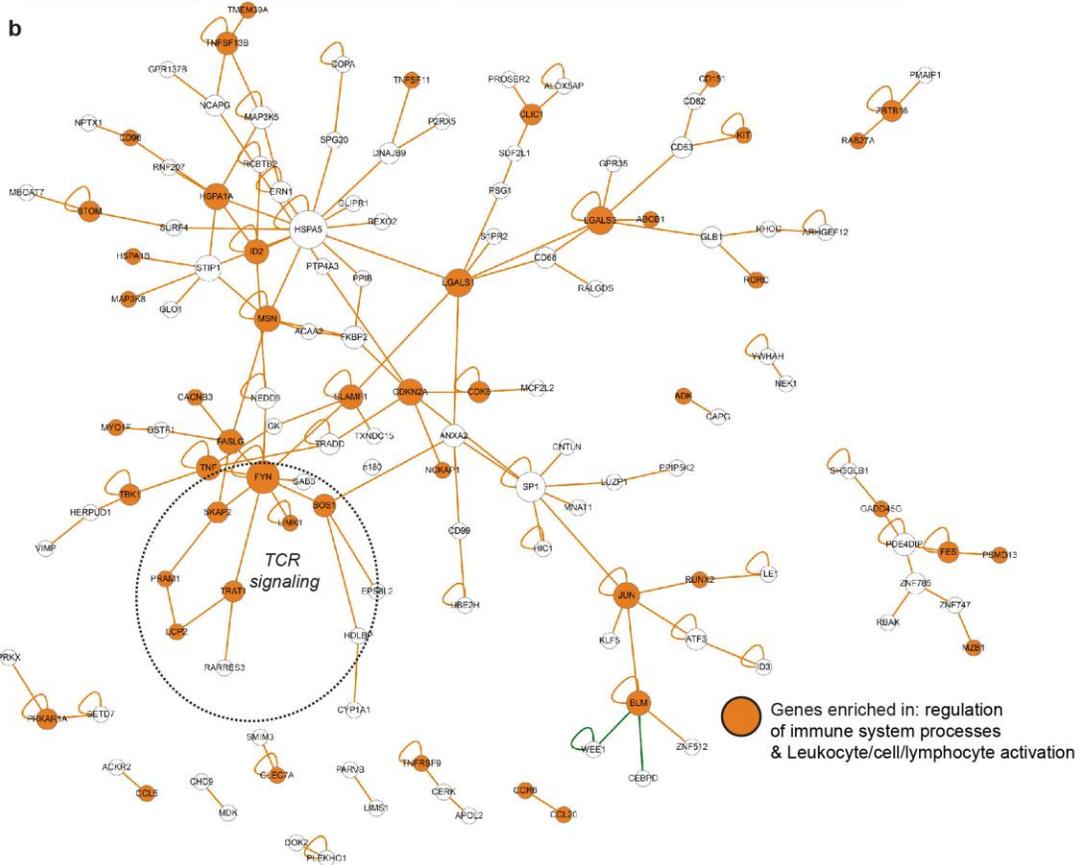
(a-b) t-SNE embeddings of all fetal intestinal CD4⁺ T cells analyzed in (a) the mass cytometry dataset (n = 10,436 cells, 35 proteins) and (b) the RNA-seq dataset (n = 1,743 cells, 300 variable genes) at the middle of the t-SNE computation. Colors indicate marker expression. (c) Diffusion map and (d) Vortex analysis of all fetal intestinal CD4⁺ T cells analyzed in mass cytometry dataset (n = 10,436 cells, 35 proteins). Colors indicate different cell clusters. (e) Diffusion map and (f) PCA analysis of all fetal intestinal CD4⁺ T cells analyzed in the RNA-seq dataset (n = 1,743 cells, 2,174 variable genes). Colors indicate different cell clusters.

a

Nr	Gene Symbol	Gene Name	Original Symbol
1	5243	ABCB1	ATP binding cassette subfamily B member 1
2	101	ADAM8	ADAM metallopeptidase domain 8
3	132	ADK	adenosine kinase
4	8943	AF3D1	adaptor related protein complex 3 delta 1 subunit
5	51676	ASB2	ankyrin repeat and SOCS box containing 2
6	641	BLM	Bloom syndrome RecQ like helicase
7	684	BST2	bone marrow stromal cell antigen 2
8	760	CA2	carbonic anhydrase 2
9	784	CACNB3	calcium voltage-gated channel auxiliary subunit beta 3
10	6364	CCL20	C-C motif chemokine ligand 20
11	6352	CCL5	C-C motif chemokine ligand 5
12	1235	CCR6	C-C motif chemokine receptor 6
13	10803	CCR9	C-C motif chemokine receptor 9
14	977	CD151	CD151 molecule (Raph blood group)
15	952	CD38	CD38 molecule
16	966	CD59	CD59 molecule (CD59 blood group)
17	10225	CD96	CD96 molecule
18	1021	CDK6	cyclin dependent kinase 6
19	1029	CDKN2A	cyclin dependent kinase inhibitor 2A
20	160364	CLEC12A	C-type lectin domain family 12 member A
21	64581	CLEC7A	C-type lectin domain containing 7A
22	1192	CLIC1	chloride intracellular channel 1
23	10695	CNPY3	canopy FGF signaling regulator 3
24	10675	CSPG5	chondroitin sulfate proteoglycan 5
25	9646	CTRF9	CTRF9 homolog, Paf1/RNA polymerase II complex component
26	2833	CXCR3	C-X-C motif chemokine receptor 3
27	23504	DAPK2	death associated protein kinase 2
28	1803	DEFA4	depeptidyl peptidase 4
29	2242	FES	FES proto-oncogene, tyrosine kinase
30	10211	FLOT1	flotillin 1
31	2534	FYN	FYN proto-oncogene, Src family tyrosine kinase
32	10912	GADD45G	grow th arrest and DNA damage inducible gamma
33	10634	GAS2L1	grow th arrest specific 2 like 1
34	2769	GNA15	G protein subunit alpha 15
35	8111	GPR68	G protein-coupled receptor 68
36	3303	HSPA1A	heat shock protein family A (Hsp70) member 1A
37	3304	HSPA1B	heat shock protein family A (Hsp70) member 1B
38	3398	ID2	inhibitor of DNA binding 2
39	3459	IFNGR1	interferon gamma receptor 1
40	3601	L15RA	interleukin 15 receptor subunit alpha
41	8809	L18R1	interleukin 18 receptor 1
42	182	JAG1	jagged 1
43	120425	JAML	junction adhesion molecule like
44	3725	JUN	Jun proto-oncogene, AP-1 transcription factor subunit
45	3815	KIT	KIT proto-oncogene receptor tyrosine kinase
46	3820	KLRB1	killer cell lectin like receptor B1
47	3937	LCR2	lymphocyte cytosolic protein 2
48	3956	LGALS1	galectin 1
49	3958	LGALS3	galectin 3
50	3976	LIF	LIF, interleukin 6 family cytokine
51	3984	LMK1	LIM domain kinase 1
52	7940	LST1	leukocyte specific transcript 1
53	1326	MAP3K8	mitogen-activated protein kinase kinase kinase 8

Nr	Gene Symbol	Gene Name	Original Symbol
54	4478	MSN	mosh
55	4542	MYO1F	myosin F
56	54005	MYO1G	myosin G
57	51237	MZB1	marginal zone B and B1 cell specific protein
58	4681	NBL1	neuroblastoma 1, DAN family BMP antagonist
59	10787	NCKAP1	NCK associated protein 1
60	114548	NLRP3	NLR family pyrin domain containing 3
61	114770	PGLYRP2	peptidoglycan recognition protein 2
62	118788	PKAP1	phosphoinositide-3-kinase adaptor protein 1
63	84106	PRAM1	PML-RARA regulated adaptor molecule 1
64	5573	PRKARIA	protein kinase cAMP-dependent type I regulatory subunit alpha
65	5719	PSMD13	proteasome 26S subunit, non-ATPase 13
66	25191	PTPRC2	protein tyrosine phosphatase, non-receptor type 22
67	5814	RURB	purine rich element binding protein B
68	5873	RAB27A	RAB27A, member RAS oncogene family
69	388	RHOB	ras homolog family member B
70	6095	RORA	RAR related orphan receptor A
71	6097	RORC	RAR related orphan receptor C
72	860	RUNX2	runt related transcription factor 2
73	8935	SKAP2	src kinase associated phosphoprotein 2
74	84174	SLA2	Src like adaptor 2
75	6504	SLAMF1	signaling lymphocytic activation molecule family member 1
76	23657	SLC7A11	solute carrier family 7 member 11
77	6654	SOS1	SOS Ras/Rac guanine nucleotide exchange factor 1
78	2040	STOM	stomatin
79	8464	SUPT3H	SPT3 homolog, SAGA and STAGA complex component
80	29110	TBK1	TANK binding kinase 1
81	55254	TNFRK3A	transmembrane protein 39A
82	7124	TNF	tumor necrosis factor
83	8784	TNFRSF18	TNF receptor superfamily member 18
84	8600	TNFSF11	TNF superfamily member 11
85	10673	TNFSF13B	TNF superfamily member 13b
86	8740	TNFSF14	TNF superfamily member 14
87	50852	TRA11	T cell receptor associated transmembrane adaptor 1
88	28526	TRDC	T cell receptor delta constant
89	53347	UBASH3A	ubiquitin associated and SH3 domain containing A
90	6375	XCL1	X-C motif chemokine ligand 1
91	7704	ZBTB16	zinc finger and BTB domain containing 16
92	80149	ZC3H12A	zinc finger CCHC-type containing 12A
93	3604	TNFRSF9	TNF receptor superfamily member 9
94	5167	BNPP1	ectonucleotide pyrophosphatase/phosphodiesterase 1
95	1075	CTSC	cathepsin C
96	8807	IL18RAP	interleukin 18 receptor accessory protein
97	4217	MAP3K5	mitogen-activated protein kinase kinase kinase 5
98	1115361	GBP4	guanylate binding protein 4
99	6845	XCL2	X-C motif chemokine ligand 2
100	9934	P2RY14	purinergic receptor P2Y14
101	27380	APOBEC3C	apolipoprotein B mRNA editing enzyme catalytic subunit 3C
102	6967	TRGC2	T cell receptor gamma constant 2
103	4938	OAS1	2'-5'-oligoadenylate synthetase 1
104	8530	CST7	cystatin F
105	356	FASLG	Fas ligand
106	976	ADGRE5	adhesion G protein-coupled receptor E5

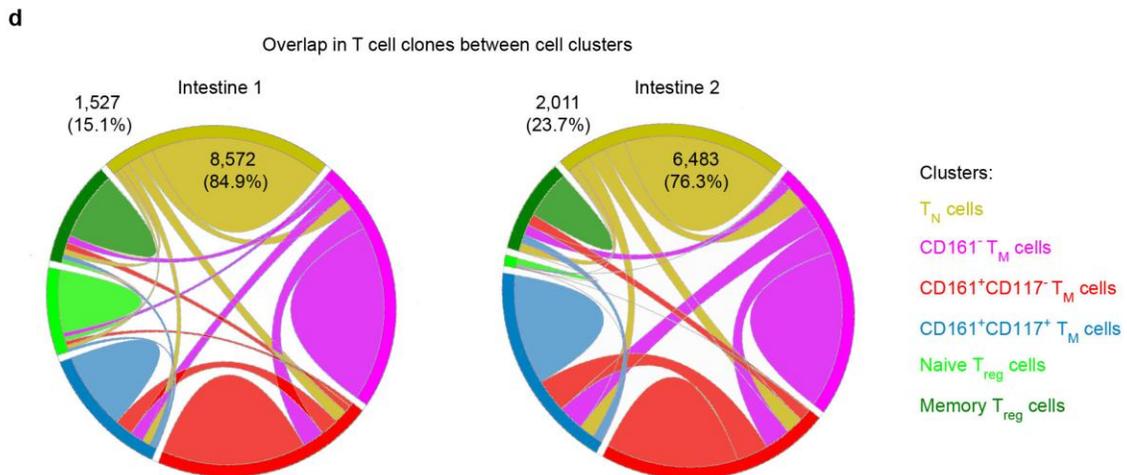
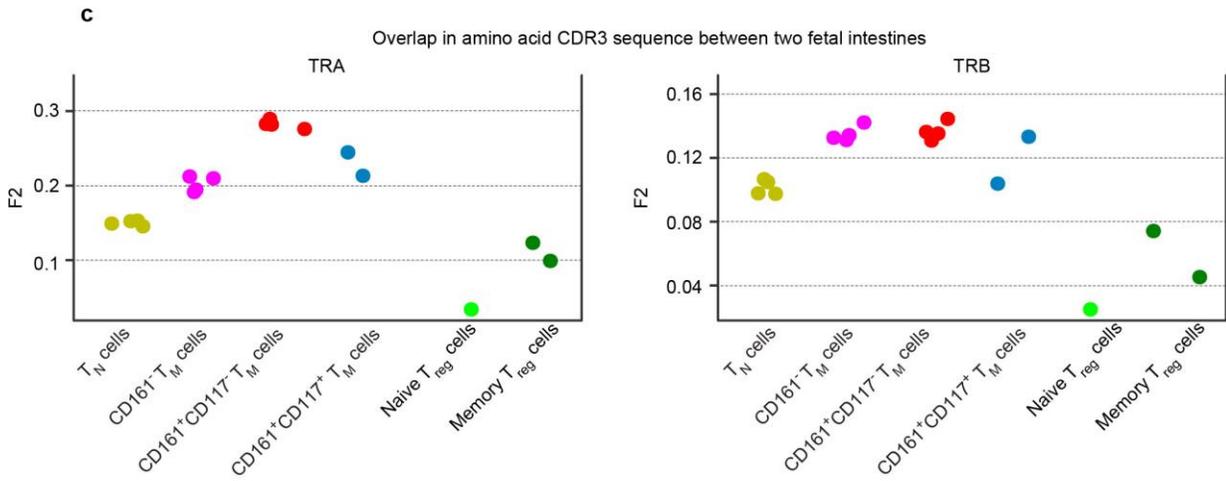
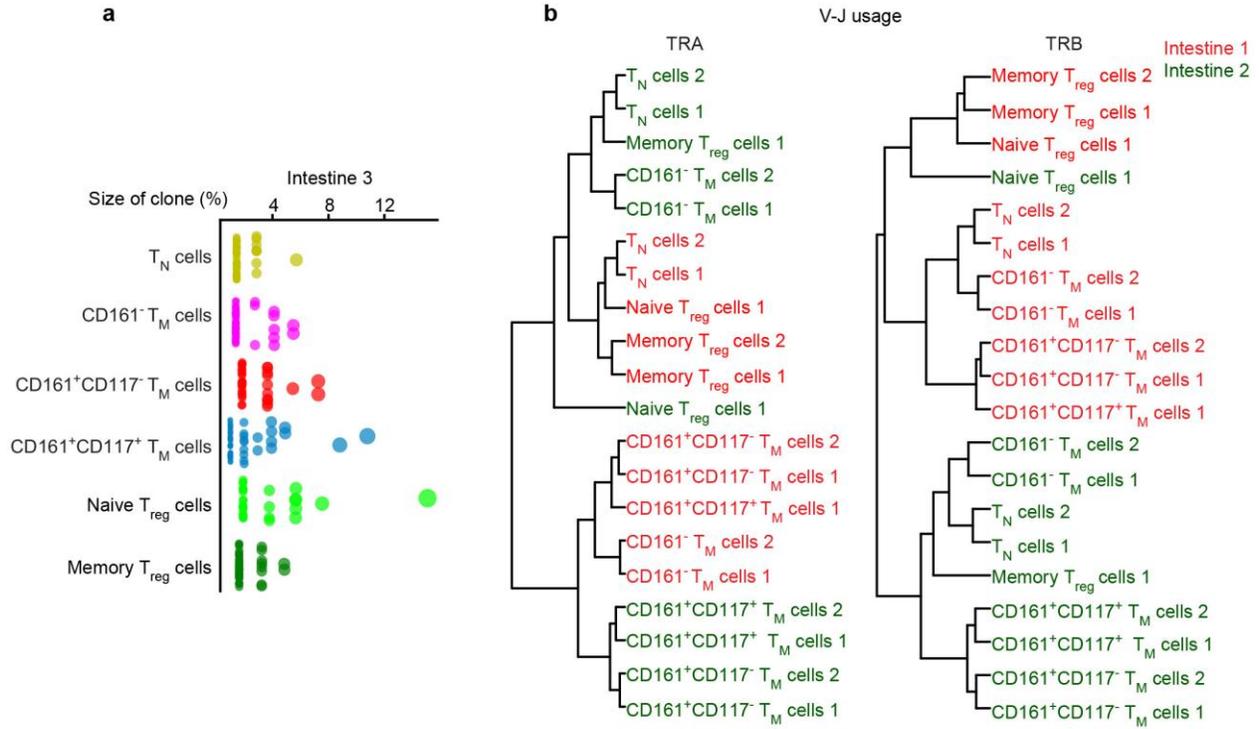
b



Supplementary Figure 5

Immune-related genes in the pseudo-time-dependent gene module.

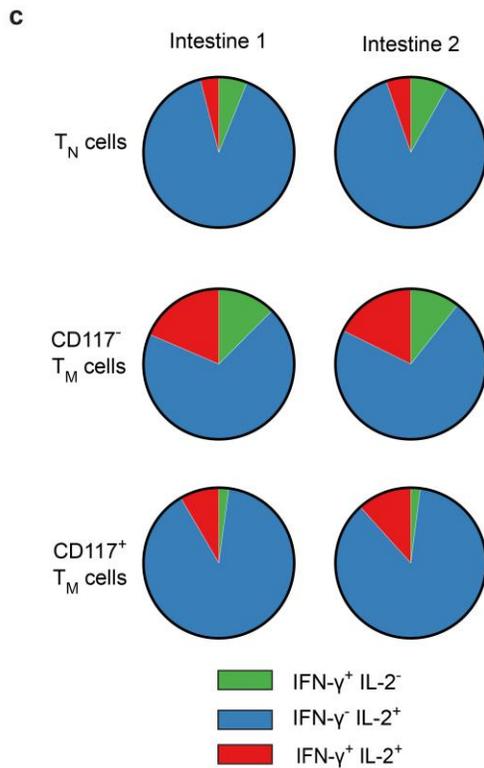
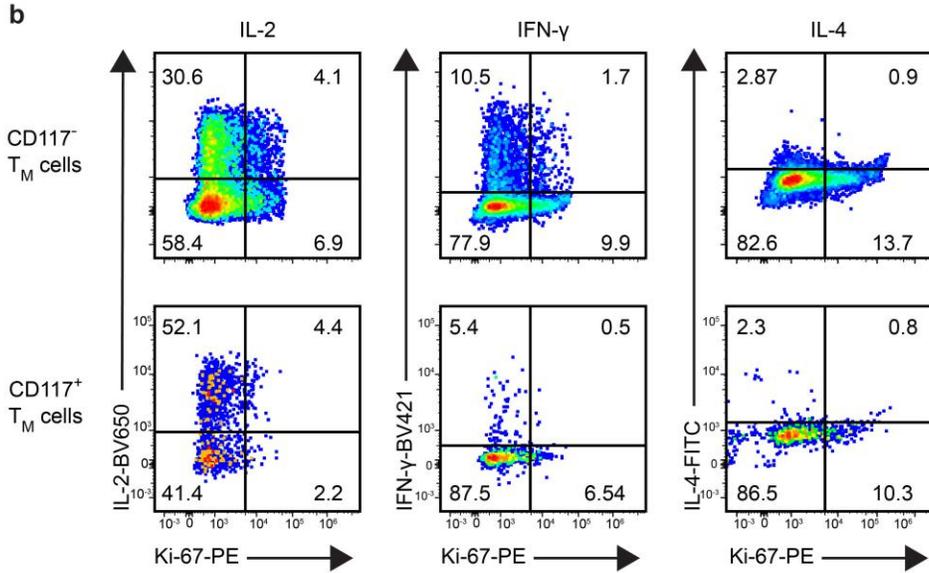
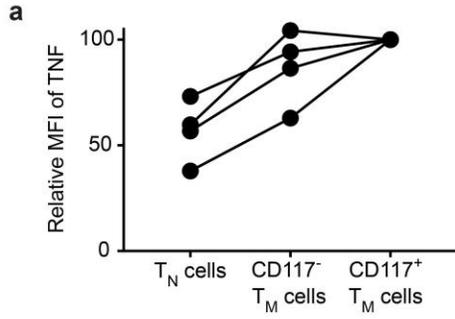
(a) Table listing 106 immune-related genes kinetically enriched in the pseudotime-dependent third gene module (**Fig. 4d**). (b) Interaction network of the pseudotime-dependent third gene module using Esyn. Orange lines indicate literature-described physical interactions of the gene-encoded proteins. Orange nodes indicate genes enriched in immunological processes.



Supplementary Figure 6

TCR sequencing of distinct CD4⁺T cell populations.

(a) Dot plot showing the percentage of reads per unique TCR β sequence in each cluster. Number of unique TCR β sequences per cluster: T_N cells, 59; CD161⁻T_M cells, 54; CD161⁺CD117⁻T_M cells, 37; CD161⁺CD117⁺T_M cells, 51; Naive T_{reg} cells, 29; Memory T_{reg} cells, 50. (b) Hierarchical clustering of CD4⁺T subpopulations based on individual V-J rearrangements in theTCR α (TRA) and TCR β (TRB) datasets. Colors indicate different fetal intestines. (c) Dot plot showing weighted donal overlaps for TCR β amino acid sequences among clusters from two different fetal intestines analyzed using the F2 similarity metric in VDJtools. (d) Chord diagram showing clonal overlaps among different cell dusters. Numbers in the yellow circle indicate the total numbers and frequencies (parentheses) of private TCR β clonotypes among T_N cells. Numbers outside the plot border indicate the total numbers and frequencies (parentheses) of unique TCR β clonotypes shared among T_N cells and other CD4⁺T cell subpopulations. (a) Conventional sequencing, (b-d) high-throughput sequencing.



Supplementary Figure 7

Cytokine production by fetal intestinal CD4⁺T cells.

(a-c) Purified fetal intestinal CD4⁺T cells were treated with a control antibody or stimulated with anti-CD3 and anti-CD28 for 4 h. Intracellular expression of TNF, IFN- γ , IL-2, IL-4 and Ki-67 was determined for each subpopulation by flow cytometry. (a) Dot plot showing the median fluorescence intensity (MFI) of TNF for each cell cluster relative to corresponding CD117⁺ T_M subpopulation in each fetal intestine (n = 4) after stimulation with anti-CD3 and anti-CD28. Data represent two independent experiments. (b) Representative biaxial plots showing coexpression of cytokines vs. Ki-67 for the indicated cell clusters after stimulation with anti-CD3 and anti-CD28. (c) Pie charts depicting coexpression profiles of IFN- γ and IL-2 for the indicated cell clusters from fetal intestines (n = 2) after stimulation with anti-CD3 and anti-CD28. Data represent two independent experiments.

Table S1 CyTOF antibody panel

	Antigen	Tag	Clone	Supplier	Cat.	Final dilution
1	CD127	¹⁶⁵ Ho	AO19D5	Flui	3165008B	1/800
2	CCR6	¹⁴¹ Pr	G034E3	Flui	3141003A	1/200
3	CD8a	¹⁴⁶ Nd	RPA-T8	Flui	3146001B	1/200
4	CD11c	¹⁶² Dy	Bu15	Flui	3162005B	1/200
5	CD38	¹⁷² Yb	HIT2	Flui	3172007B	1/200
6	CD45	⁸⁹ Y	HI30	Flui	3089003B	1/100
7	CD117	¹⁴³ Nd	104D2	Flui	3143001B	1/100
8	CD4	¹⁴⁵ Nd	RPA-T4	Flui	3145001B	1/100
9	CD16	¹⁴⁸ Nd	3G8	Flui	3148004B	1/100
10	CD25	¹⁴⁹ Sm	2A3	Flui	3149010B	1/100
11	CD123	¹⁵¹ Eu	6H6	Flui	3151001B	1/100
12	CD7	¹⁵³ Eu	CD7-6B7	Flui	3153014B	1/100
13	CD163	¹⁵⁴ Sm	GHI/61	Flui	3154007B	1/100
14	CCR7	¹⁵⁹ Tb	G043H7	Flui	3159003A	1/100
15	CD14	¹⁶⁰ Gd	M5E2	Flui	3160001B	1/100
16	CD161	¹⁶⁴ Dy	HP-3G10	Flui	3164009B	1/100
17	CD27	¹⁶⁷ Er	O323	Flui	3167002B	1/100
18	CD45RA	¹⁶⁹ Tm	HI100	Flui	3169008B	1/100
19	CD3	¹⁷⁰ Er	UCHT1	Flui	3170001B	1/100
20	PD-1	¹⁷⁵ Lu	EH 12.2H7	Flui	3175008B	1/100
21	CD56	¹⁷⁶ Yb	NCAM16.2	Flui	3176008B	1/100
22	CD11b	¹⁴⁴ Nd	ICRF44	Flui	3144001B	1/100
23	TCRγδ	¹⁵² Sm	11F2	Flui	3152008B	1/50
24	HLA-DR	¹⁶⁸ Er	L243	BioL	307651	1/200
25	CD20	¹⁶³ Dy	2H7	BioL	302343	1/200
26	CD34	¹⁴² Nd	HIB19	BioL	343531	1/100
27	IgM	¹⁵⁰ Nd	MHM88	BioL	314527	1/100
28	CD103	¹⁵⁵ Gd	Ber-ACT8	BioL	350202	1/100
29	CRTH2	¹⁵⁶ Gd	BM16	BioL	350102	1/100
30	CD28	¹⁷¹ Yb	CD28.2	BioL	302902	1/100
31	CD45RO	¹⁷³ Yb	UCHL1	BioL	304239	1/100
32	CD122	¹⁵⁸ Gd	TU27	BioL	339002	1/50
33	KLRG-1	¹⁶¹ Dy	REA261	MACS	120-014-229	1/50
34	CD8b	¹⁶⁶ Er	SIDI8BEE	ebio	14-5273	1/50
35	NKp46	¹⁷⁴ Yb	9E2	BioL	331902	1/40
Fluidigm (Flui), eBioscience (eBio) and Biolegend (BioL).						

Table S2 Information of TCR repertoire analysis with NGS

Metadata					Basic Statistics					
sample_id	fetus	subset	chain	replicate	cells sorted	Total_reads	cDNA_molecules_UMI	Reads_per_UMI_threshold	cDNA_molecules_UMI_after_filtering	Clonotypes
Intestine_1_CD117T_TRA_1	Intestine 1	CD161+CD117+ Tm	TRA	1	7140	175103	45659	2	11859	3724
Intestine_2_CD117T_TRA_1	Intestine 2	CD161+CD117+ Tm	TRA	1	7507	556612	98407	2	18135	3970
Intestine_2_CD117T_TRA_2	Intestine 2	CD161+CD117+ Tm	TRA	2	3124	281838	50527	2	7599	2026
Intestine_1_memoryCD161neg_TRA_2	Intestine 1	CD161- Tm	TRA	2	7888	333917	66317	2	14426	4915
Intestine_1_memoryCD161neg_TRA_1	Intestine 1	CD161- Tm	TRA	1	7731	254378	54054	2	13125	4713
Intestine_2_memoryCD161neg_TRA_2	Intestine 2	CD161- Tm	TRA	2	7714	554393	115584	2	16868	4541
Intestine_2_memoryCD161neg_TRA_1	Intestine 2	CD161- Tm	TRA	1	7617	197813	44832	2	6826	3073
Intestine_1_memoryCD161pos_TRA_2	Intestine 1	CD161+CD117- Tm	TRA	2	7600	528867	103916	2	13493	4093
Intestine_1_memoryCD161pos_TRA_1	Intestine 1	CD161+CD117- Tm	TRA	1	7556	219828	49602	2	10631	3671
Intestine_2_memoryCD161pos_TRA_2	Intestine 2	CD161+CD117- Tm	TRA	2	7481	741432	175332	2	21492	3714
Intestine_2_memoryCD161pos_TRA_1	Intestine 2	CD161+CD117- Tm	TRA	1	7581	360210	91069	2	17586	3600
Intestine_1_memoryTreg_TRA_2	Intestine 1	Memory Tregs	TRA	2	2343	111156	23571	2	3957	1442
Intestine_1_memoryTreg_TRA_1	Intestine 1	Memory Tregs	TRA	1	7536	119133	40023	2	8774	2757
Intestine_2_memoryTreg_TRA_1	Intestine 2	Memory Tregs	TRA	1	7217	353446	93680	2	15382	2599
Intestine_1_NaiveCD161neg_TRA_1	Intestine 1	Tn	TRA	1	7697	246744	57601	2	9071	4182
Intestine_1_NaiveCD161neg_TRA_2	Intestine 1	Tn	TRA	2	7654	153032	42957	2	9550	4278
Intestine_2_NaiveCD161neg_TRA_2	Intestine 2	Tn	TRA	2	7672	298840	69436	2	11419	4289
Intestine_2_NaiveCD161neg_TRA_1	Intestine 2	Tn	TRA	1	7542	418501	82985	2	8439	3216
Intestine_1_nTreg_TRA_1	Intestine 1	Naïve Tregs	TRA	1	7140	66778	22677	2	5103	2916
Intestine_2_nTreg_TRA_1	Intestine 2	Naïve Tregs	TRA	1	716	31684	8182	2	1267	531
Intestine_1_CD117T_TRB_1	Intestine 1	CD161+CD117+ Tm	TRB	1	7140	219407	53164	2	18482	4555
Intestine_2_CD117T_TRB_1	Intestine 2	CD161+CD117+ Tm	TRB	1	7507	350183	83370	2	25446	4521
Intestine_2_CD117T_TRB_2	Intestine 2	CD161+CD117+ Tm	TRB	2	3124	147361	37373	2	10450	2241
Intestine_1_memoryCD161neg_TRB_2	Intestine 1	CD161- Tm	TRB	2	7888	209909	60182	2	20581	5313
Intestine_1_memoryCD161neg_TRB_1	Intestine 1	CD161- Tm	TRB	1	7731	294147	74522	2	21375	5517
Intestine_2_memoryCD161neg_TRB_2	Intestine 2	CD161- Tm	TRB	2	7714	208020	66757	2	24319	5202
Intestine_2_memoryCD161neg_TRB_1	Intestine 2	CD161- Tm	TRB	1	7617	117422	37026	2	13320	4200
Intestine_1_memoryCD161pos_TRB_2	Intestine 1	CD161+CD117- Tm	TRB	2	7600	260491	64692	2	19719	4885
Intestine_1_memoryCD161pos_TRB_1	Intestine 1	CD161+CD117- Tm	TRB	1	7556	161120	46178	2	16619	4452
Intestine_2_memoryCD161pos_TRB_2	Intestine 2	CD161+CD117- Tm	TRB	2	7481	479970	129609	2	33397	4352
Intestine_2_memoryCD161pos_TRB_1	Intestine 2	CD161+CD117- Tm	TRB	1	7581	274782	79895	2	26100	4292
Intestine_1_memoryTreg_TRB_2	Intestine 1	Memory Tregs	TRB	2	2343	69426	16249	2	5620	1444
Intestine_1_memoryTreg_TRB_1	Intestine 1	Memory Tregs	TRB	1	7536	141814	38439	2	13827	3133
Intestine_2_memoryTreg_TRB_1	Intestine 2	Memory Tregs	TRB	1	7217	237380	72338	2	22032	2870
Intestine_1_NaiveCD161neg_TRB_1	Intestine 1	Tn	TRB	1	7697	391216	77080	2	16174	5306
Intestine_1_NaiveCD161neg_TRB_2	Intestine 1	Tn	TRB	2	7654	128811	42777	2	15246	5040
Intestine_2_NaiveCD161neg_TRB_2	Intestine 2	Tn	TRB	2	7672	139563	53435	2	15449	4928
Intestine_2_NaiveCD161neg_TRB_1	Intestine 2	Tn	TRB	1	7542	226971	59369	2	15403	4017
Intestine_1_nTreg_TRB_1	Intestine 1	Naïve Tregs	TRB	1	7140	70846	30168	2	9278	3812
Intestine_2_nTreg_TRB_1	Intestine 2	Naïve Tregs	TRB	1	716	19528	7090	2	2010	585

Table S3 Imaging-mass cytometry antibody panel

	Antigen	Tag	Clone	Supplier	Cat.	Final dilution	Panel
1	CD45	⁸⁹ Y	HI30	Flui	3089003B	1/50	1 and 2
2	CD3	¹⁷⁰ Er	UCHT1	Flui	3170001B	1/100	1 and 2
3	CD7	¹⁵³ Eu	CD7-6B7	Flui	3153014B	1/100	1 and 2
4	CD4	¹⁴⁵ Nd	RPA-T4	Flui	3145001B	1/50	1 and 2
5	CD38	¹⁷² Yb	HIT2	Flui	3172007B	1/50	1 and 2
6	CD8a	¹⁴⁶ Nd	RPA-T8	Flui	3146001B	1/50	1 and 2
7	CD45RA	¹⁶⁹ Tm	HI100	Flui	3169008B	1/100	1 and 2
8	Ki-67	¹⁶⁶ Er	D3B5	CST	CST9129BF	1/200	1
9	CD161	¹⁶⁴ Dy	HP-3G10	Flui	3164009B	1/50	1 and 2
10	CD69	¹⁴⁴ Nd	FN50	Flui	3149010B	1/50	1
11	CD163	¹⁵⁴ Sm	GHI/61	Flui	3154007B	1/100	1 and 2
12	HLA-DR	¹⁶⁸ Er	L243	BioL	307651	1/300	2
13	Collagen I	¹⁴⁷ Sm	polyclonal	Millipore	AB758	1/100	1
14	Vimentin	¹⁷⁵ Lu	D21H3	CST	CST5741BF	1/200	1
15	SMA	¹⁴⁸ Nd	1A4	CST	CST5685BF	1/200	1
16	E-Cadherin	¹⁵⁰ Nd	2.40E+11	CST	CST3195BF	1/50	1
	Fluidigm (Flui), Cell Signaling Technology (CST) and Biolegend (BioL)						

Table S4 Flow cytometry antibody panel

	Antigen	clone	Flourchrome	Supplier
1	CD3	SK7	Amcyan	BD
2	CD4	RPA-T4	BV421	BioL
3	CD5	L17F12	PE	BD
4	CD8a	RPA-T8	BV650	BD
5	TCR $\gamma\delta$	11F2	BV650	BD
6	CD25	M-A251	PE-cy7	BD
7	CD31	WM59	PE	BD
8	CD45RO	UCHL1	PerCP/Cy5.5	BioL
9	CD45RA	HI100	PE/Dazzle	BioL
10	CD62L	SK11	PE	BD
11	CD69	FN50	PE	BD
12	CD127	A019D5	PE	BioL
13	CD127	A019D5	FITC	BioL
14	CD127	A019D5	PE-cy7	BioL
15	CD117	YB5.B8	APC	BD
16	CD154	24-31	BV605	BioL
17	CD161	DX12	BV605	BD
18	CXCR3	1C6	PE	BD
19	CCR4	1G1	PE	BD
20	DNAM-1	DX11	PE	BD
21	TNF	MAb11	PE-cy7	ebio
22	IFN- γ	4S.B3	PE	BioL
23	IFN- γ	4S.B3	BV421	BioL
24	IL-4	MP4-25D2	FITC	BD
25	IL-17A	BL168	BV421	BioL
26	IL-2	MQ1-17H12	BV650	BioL
27	Ki-67	20Raj1	FITC	eBio
28	Ki-67	20Raj1	PE	eBio
29	Granzyme B	GB11	PE	eBio
30	Foxp3	PCH101	PE	eBio
	eBioscience (eBio), and Biolegend (BioL).			

Chapter III

Functionally specialized human CD4⁺ T cell subsets
express physicochemically distinct TCRs

Introduction

Chapter III represents the central part of my Ph.D. thesis research. In earlier studies in our laboratory, we have noticed that T cell receptor structure differs in regulatory T cells (Tregs) from other CD4⁺ T cells in mice. However, it was not studied in human Treg cells. We approached the question on a systems level and asked 1) which functional CD4⁺ T cell subsets, including Tregs, display distinct characteristics of TCR repertoires and 2) which stage of CD4⁺ T cell development is already TCR-selected? We used deep TCR profiling of T cell subsets from 5 healthy donors and analyzed eight effector CD4 T cell subsets: Th1, Th2, Th2a, Th17, Th17-1, Th22, Treg, and Tfh. This analysis revealed unexpectedly prominent, multivariate subset-specific differences that are highly reproducible across unrelated donors. A similar analysis performed for the sorted naive CD4 T cell subsets showed that specific TCR features of Tregs (short and strongly interacting CDR3 region) are rooted in the selection which takes place in the human thymus.

Furthermore, in this study, we described the shared clonotypes distribution between human effector CD4 subsets for the first time. We demonstrate high clonal overlap within Th17 and Th22 functional subsets and prominent clonal exchange of both Th17 and Th22 with Th2. Finally, we investigate relative publicity of effector CD4 subset repertoires across donors. We reveal high publicity of Treg and Tfh and high privacy of Th22 and Th2a repertoires, which corresponds to the number of added N-nucleotides for these subsets, suggesting their earlier fetal and evolutionary origin. Altogether, we provide the first detailed picture of distinct repertoire features, clonal overlap, and publicity of helper CD4 T cell subsets.

Contribution

The idea of the project was conceived in discussions with Dr. Kristin Ladell, Prof. David Price, and Prof. Dmitry Chudakov. I participated in the experimental study design and prepared the antibody panel design for flow cytometry-based sorting of cell subsets. I performed the TCR libraries preparation with the help of colleagues Dr. Irina Shagina and Dr. Alexey Davydov.

I performed the data analysis and wrote the original manuscript draft. Dr. Kristin Ladell, an equally contributing author, collected the donor samples, performed the FACS sorting and additional *in vitro* phenotyping experiments for functional T cell subsets.

Functionally specialized human CD4⁺ T-cell subsets express physicochemically distinct TCRs

Sofya A Kasatskaya^{1,2†}, Kristin Ladell^{3†}, Evgeniy S Egorov², Kelly L Miners³, Alexey N Davydov⁴, Maria Metsger⁴, Dmitry B Staroverov^{2,5}, Elena K Matveyshina⁶, Irina A Shagina^{2,5}, Ilgar Z Mamedov^{2,5}, Mark Izraelson^{2,5}, Pavel V Shelyakin^{1,2}, Olga V Britanova^{2,5}, David A Price^{3,7‡}, Dmitriy M Chudakov^{1,2,5‡*}

¹Center of Life Sciences, Skolkovo Institute of Science and Technology, Moscow, Russian Federation; ²Genomics of Adaptive Immunity Department, Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russian Federation; ³Division of Infection and Immunity, Cardiff University School of Medicine, Cardiff, United Kingdom; ⁴Adaptive Immunity Group, Central European Institute of Technology, Brno, Czech Republic; ⁵Institute of Translational Medicine, Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Pirogov Russian National Research Medical University, Moscow, Russian Federation; ⁶Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russian Federation; ⁷Systems Immunity Research Institute, Cardiff University School of Medicine, Cardiff, United Kingdom

*For correspondence:
chudakovdm@mail.ru

†These authors also contributed equally to this work

‡These authors also contributed equally to this work

Competing interests: The authors declare that no competing interests exist.

Funding: See page 15

Received: 19 March 2020

Accepted: 04 December 2020

Published: 08 December 2020

Reviewing editor: Armita Nourmohammad, University of Washington, United States

© Copyright Kasatskaya et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract The organizational integrity of the adaptive immune system is determined by functionally discrete subsets of CD4⁺ T cells, but it has remained unclear to what extent lineage choice is influenced by clonotypically expressed T-cell receptors (TCRs). To address this issue, we used a high-throughput approach to profile the $\alpha\beta$ TCR repertoires of human naive and effector/memory CD4⁺ T-cell subsets, irrespective of antigen specificity. Highly conserved physicochemical and recombinatorial features were encoded on a subset-specific basis in the effector/memory compartment. Clonal tracking further identified forbidden and permitted transition pathways, mapping effector/memory subsets related by interconversion or ontogeny. Public sequences were largely confined to particular effector/memory subsets, including regulatory T cells (Tregs), which also displayed hardwired repertoire features in the naive compartment. Accordingly, these cumulative repertoire portraits establish a link between clonotype fate decisions in the complex world of CD4⁺ T cells and the intrinsic properties of somatically rearranged TCRs.

Introduction

Adaptive immunity relies on populations of lymphocytes that express somatically rearranged antigen receptors, including CD4⁺ T cells, which differentiate from the naive pool into functionally and phenotypically distinct effector/memory subsets that determine how the immune system responds to specific challenges. In the classic dichotomy, mycobacterial and viral infections typically elicit T helper 1 (Th1) cells, which produce interferon (IFN)- γ under the control of T-bet, whereas parasitic infections typically elicit Th2 cells, which produce interleukin (IL)-4, IL-5, and IL-13 under the control of GATA3 and STAT6 (*Mosmann and Coffman, 1989*). Many other subsets have been described in the intervening years (*DuPage and Bluestone, 2016; Sallusto, 2016*). The importance of subset

choice as a proximal determinant of response efficacy is apparent from various immune dysregulation syndromes. For example, individuals with Th1 deficiency are predisposed to recurrent bacterial and mycobacterial infections, and individuals with Th17 deficiency are predisposed to chronic mucocandidiasis (McDonald, 2012; Cook and Tangye, 2009; Hernández-Santos et al., 2013). In contrast, systemic autoimmunity is more common in individuals with Th17 overactivity and/or regulatory T-cell (Treg) deficiency (Osnes et al., 2013; Costa et al., 2017; Bonelli et al., 2008; Miyara et al., 2005), and allergy is more common in individuals with a similar imbalance between Th2 cells and Tregs (Bacher and Scheffold, 2018; McGee and Agrawal, 2006; Finotto, 2008). Pathogenic and protective roles have also been described for Th9 and Th22 cells in the context of inflammatory skin diseases and various autoimmune conditions, including type I diabetes (Ryba-Stanisławowska et al., 2016) and multiple sclerosis (Rolla et al., 2014). Similarly, adverse and beneficial outcomes have been associated with the functional attributes of tumor-specific CD4⁺ T cells, consistently linking Th1-like activity with enhanced survival across a range of cancers (Protti et al., 2014). A strictly regulated effector/memory CD4⁺ T-cell profile is therefore essential for immune function and homeostasis.

Subset choice is dictated by the context of antigen presentation (Zhu et al., 2010; Groom et al., 2012; Vroman et al., 2015; Baumjohann and Ansel, 2015; Waickman et al., 2017; Barberis et al., 2018; Eisenbarth, 2019) and potentially by the mode of antigen engagement (Barberis et al., 2018; Adams et al., 2011; Wang and Reinherz, 2012; Hoffmann et al., 2015; Sibener et al., 2018; Constant and Bottomly, 1997; Corse et al., 2011). If the latter supposition is correct, then generic molecular signatures may be present among subset-specific repertoires of expressed T-cell receptors (TCRs). To explore this possibility, we systematically deconvoluted the physicochemical and recombinatorial properties of TCR α and TCR β chains encoded by transcripts isolated from rigorously defined naive and effector/memory subsets of CD4⁺ T cells. These characteristics provide a broad overview of antigen recognition preferences within a given repertoire and help delineate relatedness among distinct subsets based on patterns of clonotype selection.

Each effector/memory subset was characterized by distinct features that were recapitulated across genetically unrelated donors, indicating a predisposition to certain fate decisions at the level of the somatically rearranged TCR. In line with this notion, similar characteristics were observed in some of the corresponding naive repertoires, most notably those derived from Tregs. Repertoire overlaps further identified effector/memory subsets that were related by common ontogenetic and/or permissible transition pathways. Collectively, these findings map the clonal ancestry and organizational complexity of the human CD4⁺ T-cell compartment and demonstrate that subset fate is influenced by the structural topography of clonotypically expressed TCRs.

Results

Experimental logic and study design

We set out to investigate the naive origins and effector/memory relationships of classically defined CD4⁺ T-cell subsets in humans. An overview of the experimental workflow designed to capture these complexities is presented in *Figure 1*.

Effector/memory CD4⁺ T-cell subsets express physicochemically distinct TCRs

To investigate the TCR repertoires of functionally and phenotypically distinct effector/memory CD4⁺ T cells, we used polychromatic flow cytometry to identify and sort the commonly recognized Tfh, Th1, Th1-17, Th17, Th22, Th2a, Th2, and Treg subsets from the peripheral blood of healthy donors (n = 5). The gating strategy is described in *Figure 1—figure supplement 1* and *Table 1*. Subset frequencies are listed in *Table 2*. The corresponding TCR α and TCR β repertoires were obtained from purified mRNA using a high-throughput approach with template switch-based incorporation of unique molecular identifiers (UMIs) as described previously (Egorov et al., 2015).

Statistical analyses of the curated TCR α and TCR β datasets allowed us to describe the somatically rearranged third complementarity-determining region (CDR3) loops in terms of amino acid representation among distinct subsets of effector/memory CD4⁺ T cells. As in previous studies (Bolotin et al., 2017; Izraelson et al., 2018; Egorov et al., 2018; De Simone et al., 2019;

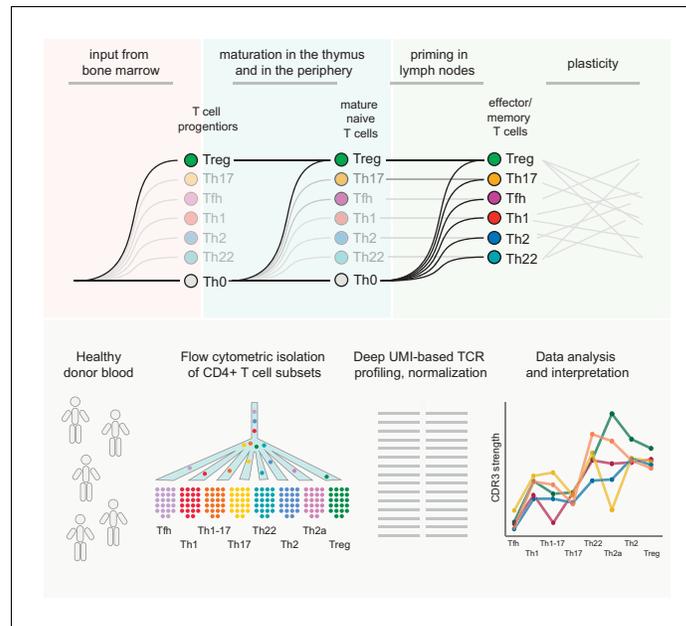


Figure 1. Experimental overview. Top: schematic representation of the general questions addressed in this study. Bottom: schematic representation of the experimental pipeline. Naive and effector/memory CD4⁺ T-cell subsets were flow-sorted from peripheral blood samples obtained from healthy donors. Repertoire characteristics were extracted from normalized datasets obtained from each subset via high-throughput sequence analysis of all expressed TCRs.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Gating strategy for the identification of effector/memory CD4⁺ T-cell subsets.

(Logunova et al., 2020), we focused on amino acid residues located in the middle of the CDR3 loop, which typically dominate contacts with the peptide component of any cognate pMHC (Egorov et al., 2018), and quantified several key physicochemical properties, including hydrophobicity (Kidera et al., 1985) and the predicted energy of TCR interactions averaged across diverse pMHCs (Miyazawa and Jernigan, 1996; Kosmrlj et al., 2008; Kosmrlj et al., 2010). This latter parameter provides a generic measure of interaction strength and depends mainly on the prevalence of aromatic and hydrophobic amino acid residues (Chakrabarti and Bhattacharyya, 2007).

Table 1. Gating strategy for the identification of effector/memory CD4⁺ T-cell subsets.

Gates 1 and 2	Gate 3	Gate 4	Gate 5	Gate 6	Gate 7	Gate 8	Subset
Live single CD3 ⁺ CD14 ⁻ CD19 ⁻ lymphocytes	CD4 ⁺	Exclude CCR7 ⁺ CD45RA ⁺	CD25 ^{high} CD127 ^{low}				Treg
			CD25 ^{low} CD127 ⁺	CXCR5 ⁺			
				CCR10 ⁺			Th22
				CXCR5 ⁻ CCR10 ⁻	CXCR3 ⁺ CCR6 ⁻	CCR4 ⁻	Th1
					CXCR3 ⁻ CCR6 ⁺	CCR4 ⁺	Th17
					CXCR3 ⁺ CCR6 ⁺	CCR4 ⁻	Th1-17
					CXCR3 ⁻ CCR6 ⁻	CCR4 ⁺ CRTh2 ⁻	Th2
						CCR4 ⁺ CRTh2 ⁺	Th2a

See also **Figure 1—figure supplement 1.**

Table 2. Frequencies of sorted effector/memory CD4⁺ T-cell subsets.

Donor	Tfh	Th1	Th1-17	Th17	Th22	Th2a	Th2	Treg
D1	5.44	1.91	1.44	3.06	2.60	1.04	4.86	3.99
D2	5.82	3.29	3.50	3.14	6.64	1.53	9.24	6.92
D3	2.05	0.19	0.31	1.31	0.81	0.26	1.93	1.84
D4	6.70	2.33	2.11	4.22	2.02	0.57	7.19	3.95
D5	4.39	1.16	1.17	3.32	2.12	0.82	3.96	3.99
Mean	4.88	1.78	1.71	3.01	2.84	0.84	5.44	4.14
SD	1.79	1.17	1.19	1.06	2.23	0.48	2.84	1.81

Shown as % of live CD3⁺CD4⁺CD14⁻CD19⁻ non-naive cells. Details in **Figure 1—figure supplement 1**.

Hydrophobicity and the propensity to form strong interactions are common but not necessarily determinative features of highly cross-reactive TCRs (*Kosmrlj et al., 2008; Kosmrlj et al., 2010; Stadinski et al., 2016*).

Although some distinct features, including high scores for hydrophobicity (low Kidera factor 4) and interaction strength in the Treg CDR3 β repertoires, were expected from previous studies in mice (*Bolotin et al., 2017; Izraelson et al., 2018; Logunova et al., 2020; Feng et al., 2015*), more unanticipated characteristics were identified among other subsets of effector/memory CD4⁺ T cells (**Figure 2**). In particular, the Tfh CDR3 β repertoires exhibited the lowest averaged scores for hydrophobicity (high Kidera factor 4; **Figure 2C**), interaction strength (**Figure 2D**), and volume (**Figure 2F**, reflects the number of bulky amino acid residues, namely W, R, K, Y, and F [*Shugay et al., 2015*]), and the highest averaged score for surface (**Figure 2E**, provides an in silico predictive measure of amino acid residues that remain unchanged in terms of accessibility and position in the liganded versus unliganded state [*Martin and Lavery, 2012*]). These exceptional features suggest that selection into the Tfh subset is driven by highly antigen-specific and minimally cross-reactive TCRs. It is tempting to speculate that such defined molecular patterns, which are mirrored in mature antibody repertoires (*Grimsholm et al., 2020*), act to minimize the risk of autoimmunity, given that Tfh cells play a critical role in the development of B-cell responses.

In addition to Tregs, relatively high numbers of strongly interacting amino acid residues were observed in the Th22, Th2a, and Th2 CDR3 β repertoires, which also scored highly in the volume analyses. Of particular note, Th22 cells expressed TCRs with the highest averaged number of random nucleotide (N) additions and the longest averaged CDR3 β length, suggesting a distinct but as yet unknown selection process. Consistent physicochemical differences were also apparent between subsets considered as two distinct groups. In general, amino acid characteristics in the Th1/Th1-17/Th17 group resembled those of Tfh cells, whereas amino acid characteristics in the Th22/Th2a/Th2 group resembled those of Tregs (**Figure 2A–F**). Similar patterns were detected in the corresponding CDR3 α repertoires (**Figure 2—figure supplement 1**). This overall dichotomy at the population level was clearly visualized using principal component analysis of the cumulative CDR3 α and CDR3 β repertoires (**Figure 2G**).

Collectively, these data show that subset fate is associated with the physicochemical properties of amino acids in the middle of the CDR3 α and CDR3 β loops, which typically dominate TCR contacts with the peptide moiety in cognate pMHCs.

Repertoire diversity varies substantially among effector/memory CD4⁺ T-cell subsets

In further analyses, we compared repertoire clonality and diversity across the same phenotypically defined subsets of effector/memory CD4⁺ T cells. Each cloneset was normalized to the lower bound of 16,000 randomly sampled UMI-labeled TCR α or TCR β cDNA molecules (*Izraelson et al., 2018*). Consistent differences in the corresponding metrics were observed among the various subsets (**Figure 3**).

Prominent clonal expansions, reflected by low normalized Shannon-Wiener indices, were apparent in the Th22 and Th2a subsets, indicating focused antigen-specific proliferation. In contrast, the Tfh subset was highly diverse, incorporating approximately 14,500 distinct sequence variants per

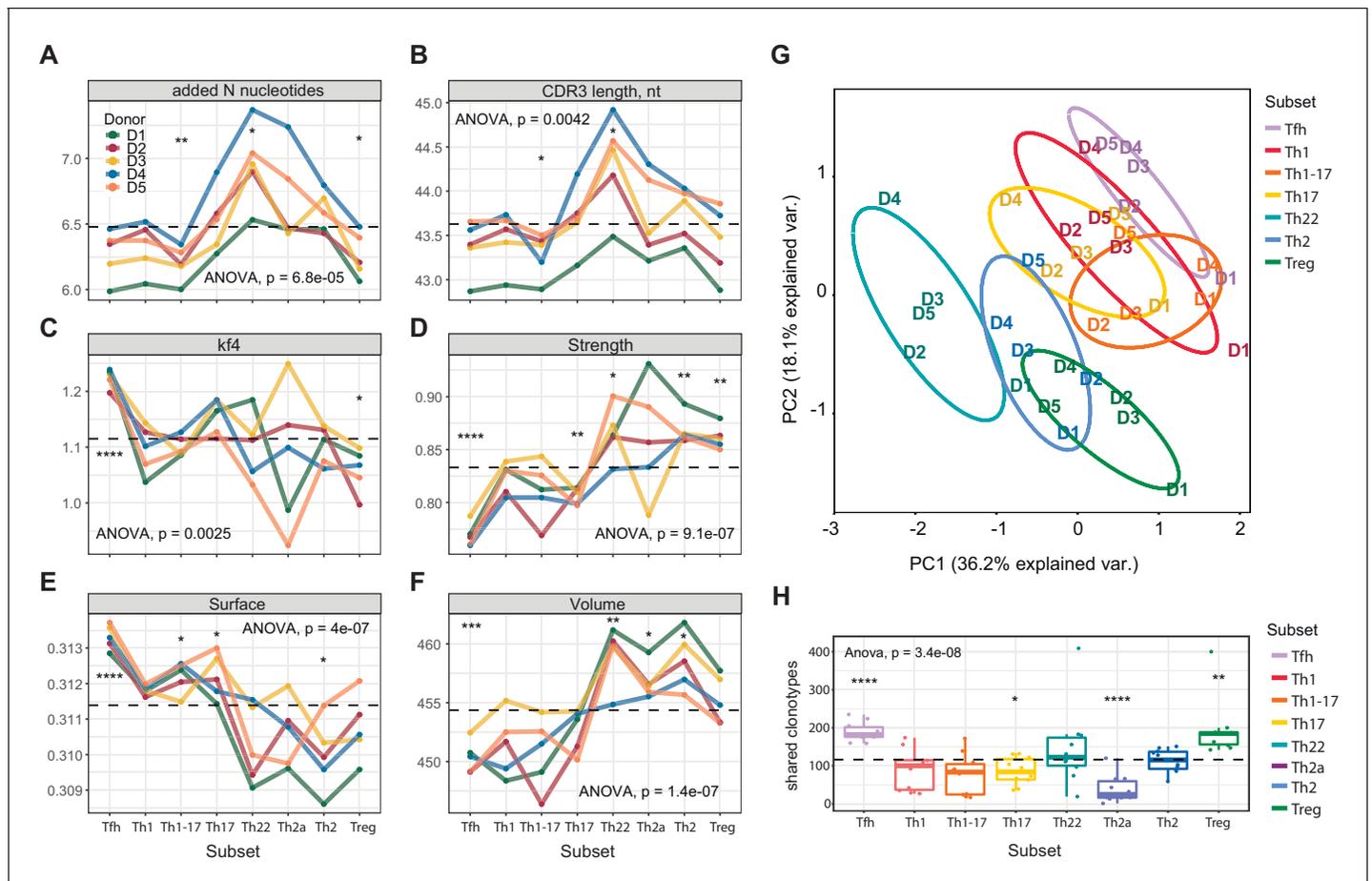


Figure 2. Averaged physicochemical characteristics of CDR3 β repertoires from effector/memory CD4⁺ T-cell subsets. (A–F) Averaged physicochemical characteristics were measured for the five amino acids in the middle of the CDR3 β sequences obtained from each effector/memory CD4⁺ T-cell subset (n = 8) from each healthy donor (n = 5). Calculations were weighted by clonotype frequency. Unweighted analyses yielded similar results (data not shown). (A) Non-germline nucleotide (N) additions. (B) CDR3 β length (nucleotides). (C) Kidera factor 4 (arbitrary scale). (D) Interaction strength (arbitrary scale). (E) Surface (arbitrary scale). (F) Volume (arbitrary scale). (G) Principal component analysis of the cumulative CDR3 α and CDR3 β repertoires from each subset of effector/memory CD4⁺ T cells (n = 28 parameters computed in VDJtools). Top contributing factors to PC1: CDR3 β volume, mJenergy, core, beta, length, number of added nucleotides, strength, and alpha. Top contributing factors to PC2: CDR3 α disorder, CDR3 α Kidera factor 3, CDR3 β disorder, CDR3 α Kidera factor 1, CDR3 α strength, CDR3 β Kidera factors 2, 3, 4, and 10, and CDR3 β charge. (H) Relative publicity measured for each effector/memory CD4⁺ T-cell subset as the number of identical or near-identical (maximum n = 1 mismatch) amino acid residue-defined CDR3 β variants shared between the top 20,000 most frequent clonotypes in the corresponding repertoires from each pair of donors. Dashed lines indicate means. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$ (one-way ANOVA followed by the two-sample Welch t-test with Bonferroni correction for each group versus the mean).

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Averaged physicochemical characteristics of CDR3 α repertoires from effector/memory CD4⁺ T-cell subsets.

16,000 cDNA molecules. Similar levels of diversity have been observed in umbilical cord blood samples, which almost exclusively contain naive T cells (<https://www.biorxiv.org/content/early/2018/09/05/259374>). The absence of large clonal expansions among circulating Tfh cells concurs with the findings of a recent study, which also reported greater clonality among donor-matched samples of tonsil-resident Tfh cells (Brenna et al., 2020). Relatively high levels of diversity were also observed in the Th1, Th17, and Th2 subsets.

Collectively, these results expose substantial variations in clonality and diversity among effector/memory subsets of CD4⁺ T cells, likely reflecting distinct selection processes driven by cognate interactions with distinct arrays of pMHCs.

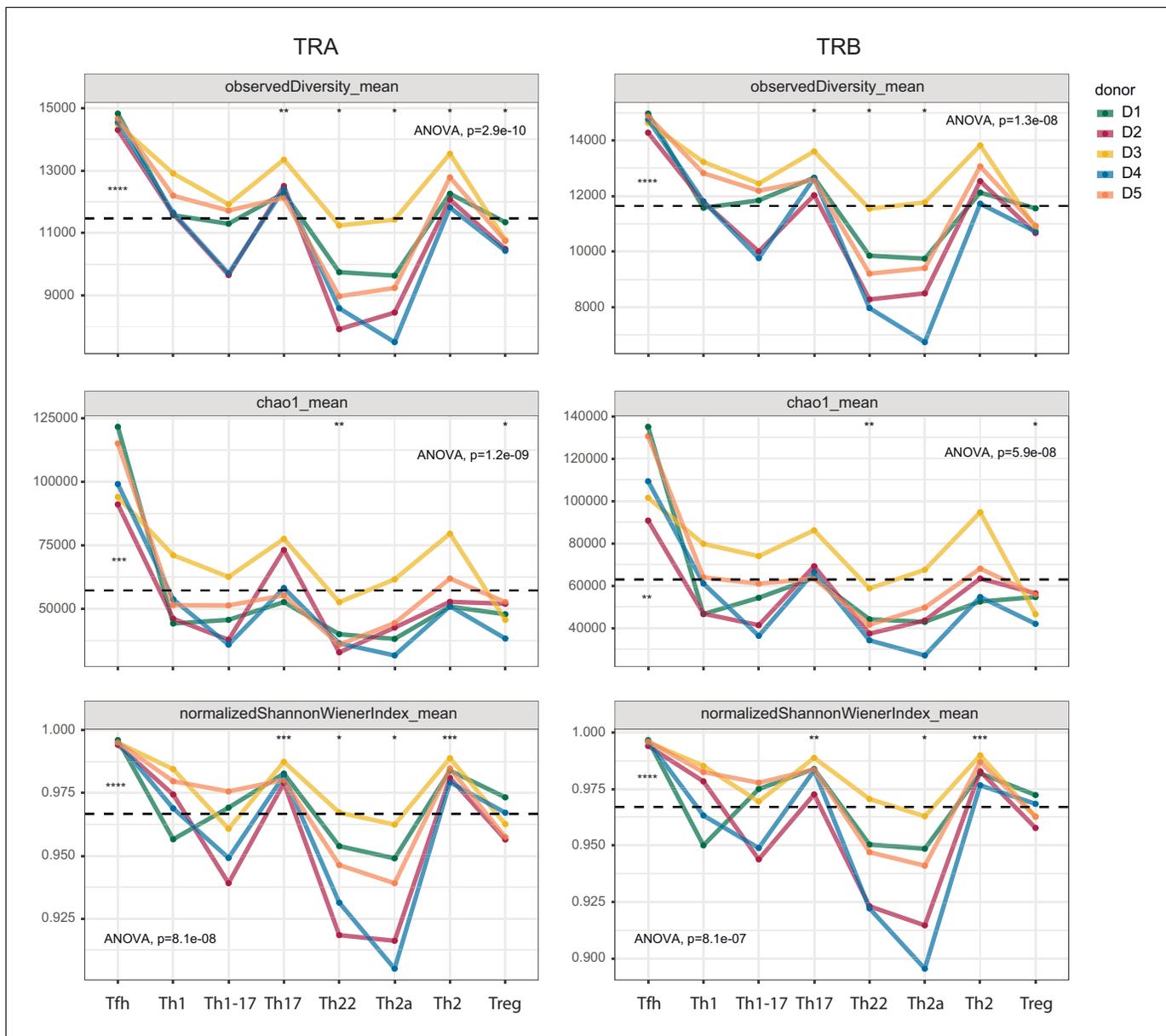


Figure 3. Clonality and diversity of effector/memory CD4⁺ T-cell subsets. Observed diversity (top), the Chao1 estimator (middle), and the normalized Shannon-Wiener index (bottom) were calculated for each TCR α (left) and TCR β repertoire (right) obtained from each effector/memory CD4⁺ T-cell subset ($n = 8$) from each healthy donor ($n = 5$). Dashed lines indicate means. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ (one-way ANOVA followed by the two-sample Welch t-test with Bonferroni correction for each group versus the mean).

Clonal transitions identify related subsets of effector/memory CD4⁺ T cells

Effector/memory CD4⁺ T cells can switch from one functional subset to another, both in vitro, driven by cytokines, and in vivo, driven by changes in the microenvironment. For example, Th2 cells have been shown to adopt a Th1-like phenotype in mice after infection with lymphocytic choriomeningitis virus, which induces type I and type II IFNs (Hegazy et al., 2010). Conversely, Th1 and Th17 cells effectively transitioned into the Th2 subset after transfer into helminth-infected mice, whereas effector Tregs maintained a stable phenotype in the same model (Panzer et al., 2012). Previous studies have also shown that human effector Tregs are relatively stable, with rare transitions to the Th1 phenotype occurring only under extreme conditions (Zhou et al., 2009a; Krebs and Steinmetz, 2016; McClymont et al., 2011). However, experiments conducted in vitro or ex vivo are not sufficient to

allow reliable quantitative estimates of plasticity among human effector/memory CD4⁺ T-cell subsets in vivo.

To address this issue, we measured relative overlap as the number of nucleotide-defined CDR3 β clonotypes shared between each pair of subsets in each donor. Similar analyses were conducted using a weighted metric to account for clonotype frequency. The top 20,000 most frequent clonotypes were selected from each TCR β cloneset to normalize the comparisons (**Figure 4**), and the top 2000 most frequent clonotypes were used to generate the corresponding Cytoscape plots (**Figure 5** and **Figure 5—figure supplements 1–4**). Overall, these analyses revealed prominent clonal exchange among two groups of subsets, namely Th17/Th22/Th2a/Th2 and Th1/Th1-17.

The complementarity and relative functional proximity of the Th17 and Th22 subsets was described previously, albeit without direct evidence of clonal transitions in vivo (**Eyerich et al.**,

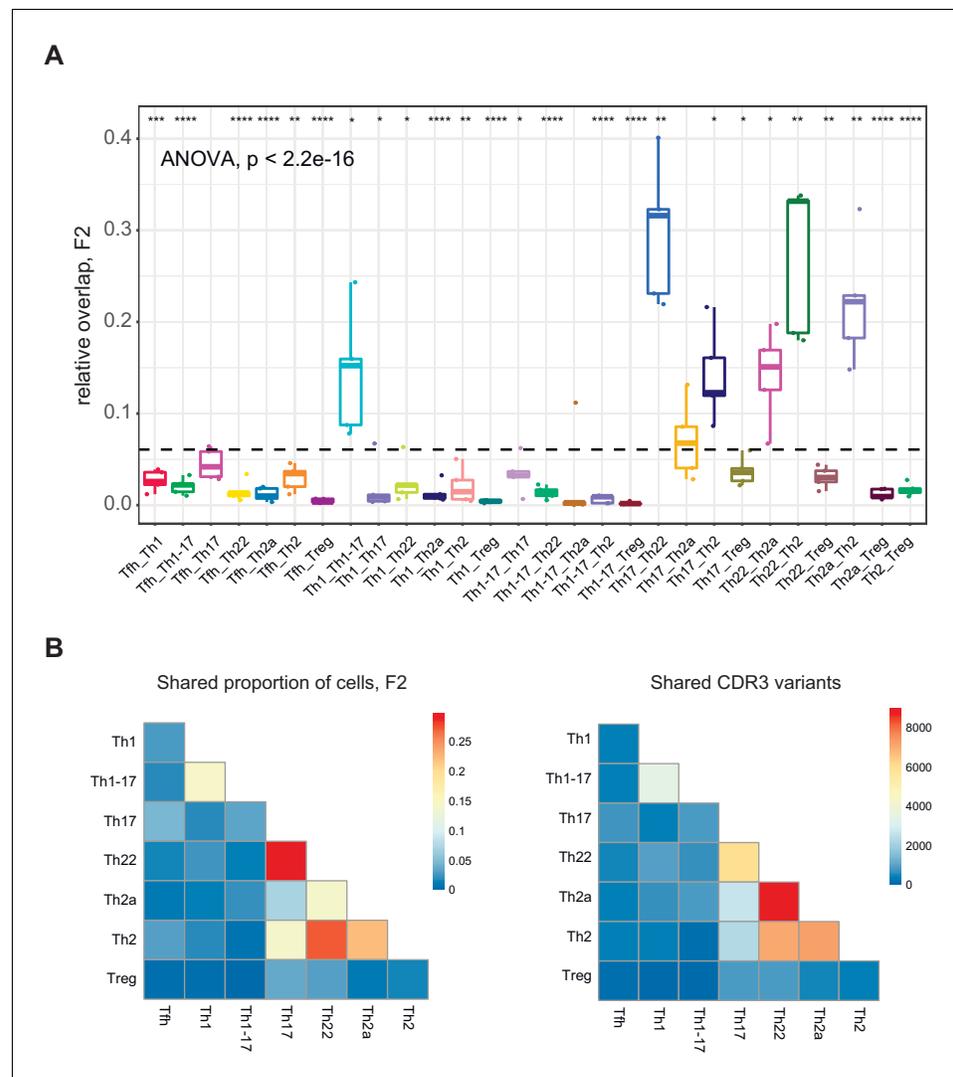


Figure 4. Clonotype overlap among effector/memory CD4⁺ T-cell subsets. (A) Relative overlap between nucleotide-defined CDR3 β repertoires obtained from donor-matched pairs of effector/memory CD4⁺ T-cell subsets. Clonotypes were matched on the basis of identical *TRBV* gene segments and identical CDR3 β sequences. Data were normalized to the top 20,000 most frequent clonotypes and weighted by clonotype frequency (F2 metric in VDJtools). The dashed line indicates the mean (n = 5 donors). *p<0.05, **p<0.01, ***p<0.001, and ****p<0.0001 (one-way ANOVA followed by the two-sample Welch t-test with Bonferroni correction for each group versus the mean). (B) Heatmap representations of the weighted overlap (F2 metric in VDJtools, left) and the estimated relative overlap of nucleotide-defined CDR3 β clonotypes (calculated via the D metric in VDJtools, right) between donor-matched pairs of effector/memory CD4⁺ T-cell subsets.

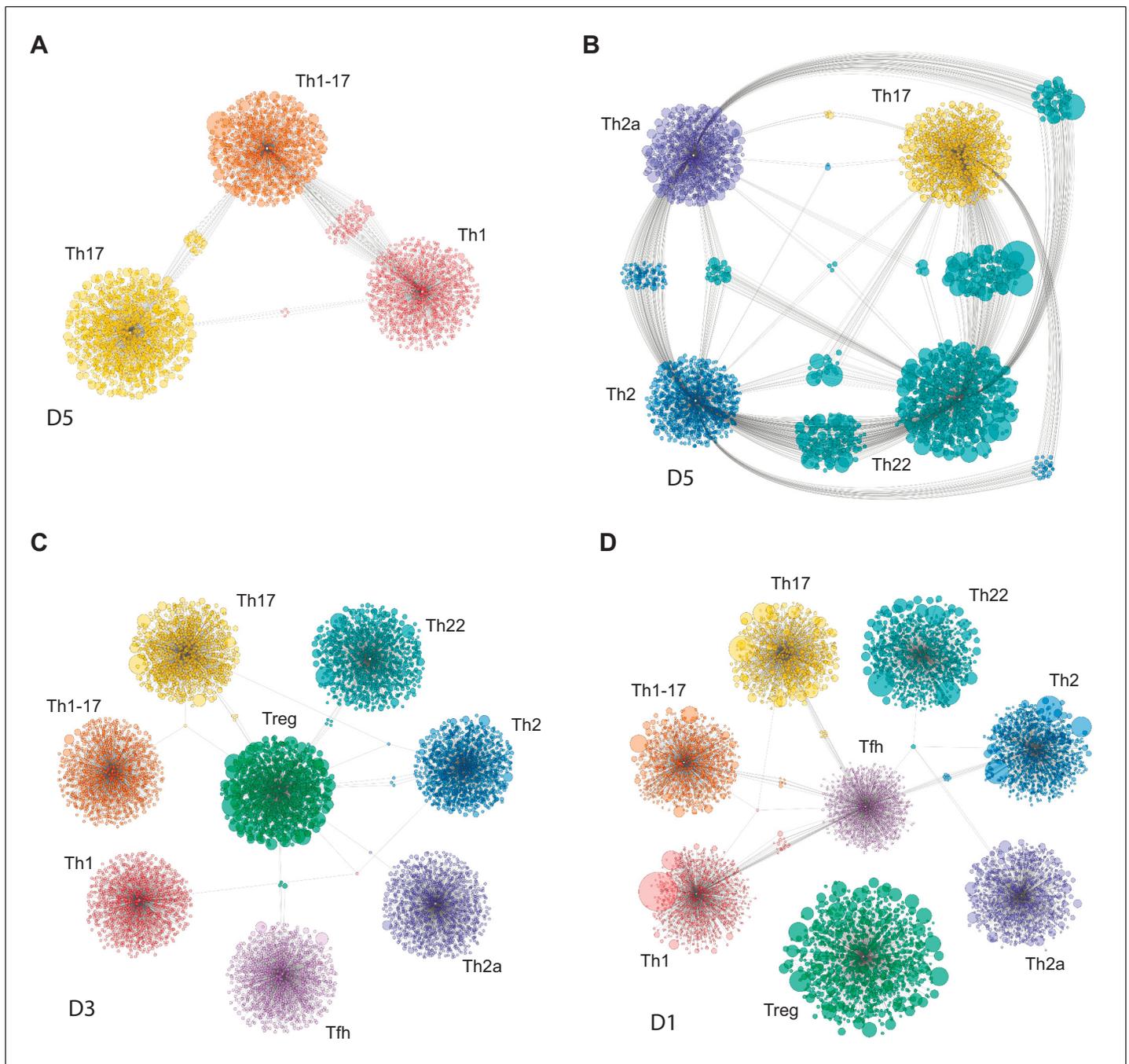


Figure 5. Clonal relatedness among effector/memory CD4⁺ T-cell subsets. Cytoscape network analysis schemes represent the number and size (frequency) of nucleotide-defined clonotype variants shared among the top 2000 most frequent CDR3 β clonotypes in each subset. Each bubble represents one CDR3 β clonotype. The size of each bubble is proportional to the frequency of each CDR3 β clonotype in the corresponding repertoire. Shared clonotypes are depicted as connected clouds among the corresponding subsets. The size of each bubble in these clouds is proportional to the frequency of each CDR3 β clonotype averaged across the maternal subsets. Representative plots were selected for illustrative purposes from donors D1, D3, and D5. (A) Th1/Th1-17/Th17. (B) Th17/Th22/Th2a/Th2. (C) Tregs versus other subsets. Only clonotypes shared with Tregs are shown. (D) Tfh cells versus other subsets. Only clonotypes shared with Tfh cells are shown.

The online version of this article includes the following figure supplement(s) for figure 5:

Figure supplement 1. Clonal relatedness among the Th17, Th22, Th2a, and Th2 subsets of effector/memory CD4⁺ T-cells.

Figure supplement 2. Clonal relatedness among the Th1, Th1-17, and Th17 subsets of effector/memory CD4⁺ T-cells.

Figure supplement 3. Clonal relatedness among Tregs and other subsets of effector/memory CD4⁺ T-cells.

Figure supplement 4. Clonal relatedness among Tfh cells other subsets of effector/memory CD4⁺ T-cells.

2009; Akdis et al., 2012). However, the close relationships between the Th17 and Th2 subsets and between the Th22 and Th2a/Th2 subsets were unforeseen. Of note, several subsets, including Th17 and Th2 cells, shared large clonal expansions with the Th22 subset (Figure 5B and Figure 5—figure supplement 1). This observation appears to conflict with the dogma that Th22 cells are stable (Eyerich et al., 2009; Plank et al., 2017) and suggests that individual clonotypes can seed and/or transition among distinct subsets within the Th17/Th22/Th2a/Th2 group.

It has been suggested previously that Th1-17 cells represent a more mature form of Th17 cells (Muranski and Restifo, 2013). In contrast, our findings suggest that Th1-17 cells are more closely related in terms of clonal proximity to Th1 cells rather than Th17 cells. Repertoire overlap between the Th1-17 and Th17 subsets was nonetheless variable among donors, ranging from zero to levels that approximated those observed between the Th1 and Th1-17 subsets (Figures 4 and 5A, and Figure 5—figure supplement 2).

Collectively, these findings suggest that plasticity is common between certain subsets, such as Th17/Th22 and Th17/Th2, but rare between other subsets, such as Th17/Treg and Th1/Th17 (Maggi et al., 2012). In addition, the Tfh and Treg subsets were largely discrete at the clonal level (Figures 4 and 5, and Figure 5—figure supplements 3 and 4). This latter observation contrasts with previous reports of Treg plasticity (Zhou et al., 2009a) but does not exclude the possibility of transient conversions from the committed Treg phenotype (Yang et al., 2008; Voo et al., 2009).

Publicity is a notable feature of Tfh cells and Tregs

To extend these analyses, we estimated the extent to which amino acid residue-defined CDR3 β clonotypes in each subset were shared among donors, essentially providing a measure of publicity. The top 20,000 most frequent clonotypes were selected from each TCR β cloneset to normalize the comparisons.

Publicity was observed most commonly among Tfh cells and Tregs, the latter in agreement with previous reports (Pacholczyk and Kern, 2008; Lei et al., 2015). In contrast, relatively few CDR3 β clonotypes in the Th22 and Th2a subsets were shared among donors (Figure 2H). These publicity metrics aligned to some extent with subset-specific differences in CDR3 β length and the number of N additions (Figure 2A,B). One possible explanation for the enrichment of public clonotypes in the Tfh and Treg repertoires lies in the nature of the corresponding antigen-driven selection events. In the case of Tfh cells, common foreign antigens presented in a degenerate manner by MHCs may be recognized predominantly by germline-encoded components of the corresponding TCRs, and in the case of Tregs, common self-derived antigens presented and recognized similarly in the thymus may drive the preferential recruitment of different clonotypes bearing germline-like TCRs.

The relative paucity of N additions in these subsets could reflect low levels of terminal deoxynucleotidyl transferase (TdT) activity, especially among Tregs, some of which arise early in life (Tulic et al., 2012; Coutinho et al., 2005; Thiault et al., 2015; Darrigues et al., 2018). A similar phenomenon may likewise explain interindividual differences in publicity, given that all subset-specific effector/memory CD4⁺ T-cell repertoires in one donor were characterized by low numbers of N additions and relatively short CDR3 β loops (Figure 2A,B).

Tregs display similar repertoire features in the naive and effector/memory pools

In general, naive CD4⁺ T cells are thought to be capable of differentiating into any effector/memory subset from the Th0 state, depending on the composite strength of TCR interactions with cognate pMHCs, costimulatory signals, and the cytokine microenvironment (Sad and Mosmann, 1994). However, this paradigm of multipotency has been challenged by the demonstration in several reports that at least some naive CD4⁺ T cells are predisposed to a specific functional program or even committed to a predetermined fate. This phenomenon was first described for thymic Tregs (tTregs), which maintain a largely stable phenotype in the periphery (Silva et al., 2016; Hoffmann et al., 2006). At the early immature double-negative stage, thymocytes are already predisposed to the Treg lineage via epigenetic modifications and increased expression of FoxP3 (Ohkura et al., 2012; Arvey et al., 2015). Other inputs are then required to confirm this commitment, including signals delivered by the IL-2 receptor and intermittent stimulation via high-affinity TCRs (Levine et al., 2014). A similar process of agonist-driven selection has been described for thymic Th17 cells in mice

(Marks et al., 2009). Accordingly, subset fate may be imprinted at the progenitor stage (Feng et al., 2015), during thymic development (Li and Rudensky, 2016), after thymic emigration and before determinative antigen encounter (Fink, 2013), and/or during the key priming event that signals expansion and maturation (Figure 1).

On the basis of these considerations, we reasoned that certain subset-specific repertoire features, at least in the case of Tregs, could be conserved between the corresponding naive and effector/memory pools. To investigate this prediction, we profiled the TCR α and TCR β repertoires of naive CD4⁺ T cells flow-sorted as recent thymic emigrants (RTEs) (Kilpatrick et al., 2008), mature naive T cells, or naive Tregs from the peripheral blood of healthy donors (total, n = 12; twin pairs, n = 5).

The naive Treg CDR3 β repertoires were enriched for bulky, hydrophobic, and strongly interacting amino acid residues compared with the corresponding RTE and mature naive T-cell repertoires (Figure 6A and Figure 6—figure supplement 1). These observations are consistent with potent agonist-driven selection in the thymus (Feng et al., 2015; Jordan et al., 2001). In addition, naive Tregs expressed TCRs with shorter CDR3 α and CDR3 β loops. Similar features were observed in the effector/memory Treg compartment (Figure 2).

To confirm and extend these findings, we conducted similar analyses of naive CD4⁺ T-cell subsets flow-sorted as Th1-like cells (non-Treg CCR4⁻CXCR3⁺), Th2-like cells (non-Treg CCR4⁺CXCR3⁻), and Tregs (CD25^{high}CD127^{low}) from healthy donors (n = 4) matching those shown in Figure 2. The corresponding non-Treg CCR4⁻CXCR3⁻ and non-Treg CCR4⁺CXCR3⁺ populations were analyzed in parallel for comparative purposes.

The naive Treg CDR3 β repertoires were again enriched for bulky, hydrophobic, and strongly interacting amino acid residues compared with the other naive subset-specific CDR3 β repertoires

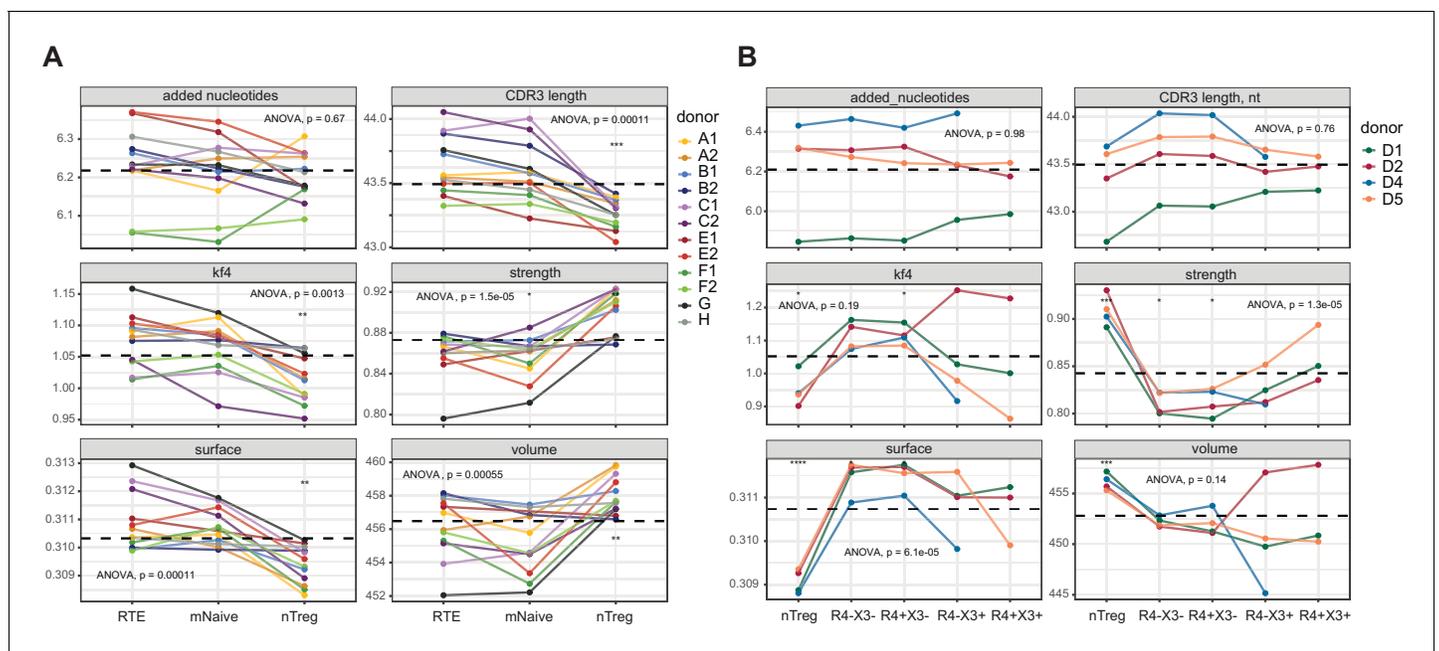


Figure 6. Averaged physicochemical characteristics of CDR3 β repertoires from naive CD4⁺ T-cell subsets. (A) Repertoire analysis of RTEs (CD25⁻CD31⁺), mature naive T cells (mNaive; CD25⁻CD31⁻), and naive Tregs (nTreg; CD25^{high}) from healthy donors (n = 12). Matched letters in the key indicate twin pairs. (B) Repertoire analysis of naive Th1-like cells (non-Treg CCR4⁻CXCR3⁺), naive Th2-like cells (non-Treg CCR4⁺CXCR3⁻), naive Tregs (CD25^{high}CD127^{low}), and the corresponding non-Treg CCR4⁻CXCR3⁻ and non-Treg CCR4⁺CXCR3⁺ populations from healthy donors (n = 4) matching those shown in Figure 2. Averaged physicochemical characteristics were measured for the five amino acids in the middle of the CDR3 β sequences obtained from each naive CD4⁺ T-cell subset. Calculations were weighted by clonotype frequency. Parameter details as in Figure 2. Dashed lines indicate means. *p<0.05, **p<0.01, ***p<0.001, and ****p<0.0001 (one-way ANOVA followed by the two-sample Welch t-test with Bonferroni correction for each group versus the mean). The online version of this article includes the following figure supplement(s) for figure 6:

Figure supplement 1. Averaged physicochemical characteristics of CDR3 α repertoires from naive CD4⁺ T-cell subsets.

Figure supplement 2. Gating strategy for the identification of naive CD4⁺ T-cell subsets.

(Figure 6B). In addition, both naive CXCR3⁺ subsets were characterized by increased numbers of strongly interacting amino acid residues compared with the corresponding naive CXCR3⁻ subsets. A similar dichotomy has been reported for naive CD8⁺ T cells (De Simone et al., 2019). However, the naive Th1-like and naive Th2-like CDR3 β repertoires were generally physicochemically distinct from the corresponding effector/memory CDR3 β repertoires, in contrast to Tregs (Figure 2).

Collectively, these results indicate that distinct repertoire features are hardwired in the Treg lineage during thymic selection, whereas other subset-specific repertoires are generally shaped later in ontogeny, most likely driven by naive CD4⁺ T-cell interactions with cognate pMHCs.

Discussion

In this study, we used polychromatic flow cytometry and an unbiased high-throughput sequencing approach to probe the ontogeny and relatedness of commonly recognized effector/memory CD4⁺ T-cell subsets via in-depth analyses of clonotypically expressed TCRs. We found that each subset-specific repertoire was characterized by distinct physicochemical and recombinatorial features that were highly reproducible across multiple donors. Importantly, these differences were multivariate, such that each subset displayed an array of repertoire characteristics, which in aggregate delineated the spectrum of preferred TCRs.

The CDR3 α and CDR3 β repertoires of effector/memory Tregs contained the highest numbers of hydrophobic and strongly interacting amino acid residues. These features were recapitulated in the corresponding naive Treg repertoires, suggesting that lineage fate was predetermined by selection events in the thymus (Feng et al., 2015; Jordan et al., 2001). Similar physicochemical characteristics have been associated previously with highly cross-reactive TCRs (Kosmrlj et al., 2008; Kosmrlj et al., 2010; Stadinski et al., 2016). However, naive and effector/memory Tregs also expressed TCRs with relatively short CDR3 α and CDR3 β loops, which might limit steric flexibility and thereby enhance the specificity of antigen-driven selection (Li and Rudensky, 2016; Bacher et al., 2016; Su et al., 2016; Spence et al., 2018; Akkaya et al., 2019). Such composite properties are compatible with an inherent predilection for self-derived peptides tempered by a capacity for ligand discrimination. Of note, the effector/memory Treg subset as identified in this study potentially incorporated both thymus-derived and peripherally induced Tregs (Hoffmann et al., 2006). In line with the possibility of mixed origins, higher numbers of N additions were detected in the naive Treg repertoires compared with the effector/memory Treg repertoires, potentially indicating the long-term persistence of early fetal Tregs (Booth et al., 2010).

Substantial heterogeneity is thought to exist in the Treg lineage (Sawant and Vignali, 2014). For example, peripheral interconversion between Th17 cells and Tregs has been observed in the presence of IL-6 and TGF- β 1 (Murphy and Stockinger, 2010), and a loss of Foxp3 expression along with regulatory functions has been observed in the context of lymphopenia (Tang et al., 2008; Lathrop et al., 2008). In a more recent evaluation of this latter phenomenon, however, the ex-Foxp3⁺ cells that accumulated under lymphopenic conditions were not *bona fide* Tregs, but rather descendants of non-Tregs that transiently expressed Foxp3 (Miyao et al., 2012). Partial transition from the Treg subset has also been associated with the Tr1 phenotype, distinguished by high production levels of IL-10 (Häringer et al., 2009). In contrast, we found little evidence of plasticity among effector/memory Tregs, suggesting a largely fixed lineage choice, irrespective of potentially diverse origins.

Our analysis of circulating Tfh (cTfh) cells likely included migratory components of the Th1-like, Th17-like, Th2-like, and follicular regulatory (Tfr) subpopulations of Tfh cells (Bentebibel et al., 2013; Morita et al., 2011; Linterman et al., 2011; Chung et al., 2011; Maceiras et al., 2017; Yang et al., 2019). Unexpectedly in light of this potential heterogeneity, we found that the cTfh CDR3 α and CDR3 β repertoires were characterized by extreme features, including the lowest numbers of bulky, hydrophobic, and strongly interacting amino acid residues, with low dispersion among donors and little overlap with other subsets. These characteristics were further associated with short CDR3 α and CDR3 β loops. Accordingly, cTfh cells formed a distinct cluster in the principal component analysis, closest to the Th1 subset. This configuration suggests a high degree of antigen specificity with minimal cross-reactivity (Kosmrlj et al., 2008; Kosmrlj et al., 2010; Stadinski et al., 2016). It is tempting to speculate that such features are required to prevent the induction of autoantibody responses. In support of this hypothesis, remarkably similar features are acquired

progressively in the B-cell repertoire during the course of affinity maturation, reflecting intense negative selection of cross-reactive antibody variants (*Grimsholm et al., 2020*). Our data also suggest that non-hydrophobic contacts underpin antigen specificity in the context of high-affinity interactions between Tfh cell-expressed TCRs and cognate pMHCs (*Fazilleau et al., 2009a*).

In the periphery, cTfh cells survey multiple tissue sites and respond swiftly to previously encountered antigens, providing a systemic mirror of germinal center reactions after exiting the inceptive lymph node (*Shulman et al., 2013; Vella et al., 2019*). We found no evidence of clonal expansions in the cTfh repertoires of healthy donors, likely reflecting the random nature of recirculation and the consequent sampling of mixed specificities. This interpretation concurs with the findings of a recent study, in which clonality was low among cTfh cells and high among tonsillar Tfh cells (*Brenna et al., 2020*). Network analysis further revealed that cTfh-expressed TCRs were largely subset-specific and rarely exhibited clonal transitions. This observation again concurs with previous work (*Brenna et al., 2020*). Accordingly, cTfh cells appear to represent a distinct lineage rather than a differentiation step in the progressive maturation of other subsets, as proposed in some earlier models (*Fazilleau et al., 2009b; Vinuesa et al., 2016*).

Th22 cells are typically found in the skin, where they play a key role in wound healing (*Alabbas et al., 2018*) and epidermal immunity (*Eyerich et al., 2009*). Pathogenic activity has also been ascribed to this subset in the contexts of multiple sclerosis (*Rolla et al., 2014*), rheumatoid arthritis (*Miyazaki et al., 2018*), and chronic skin graft-versus-host disease (*Gartlan et al., 2018*). On the basis of in vitro studies, Th22 cells are thought to exhibit plasticity with Th1 and possibly with Th2 cells (*Plank et al., 2017*). Our systematic analysis of plasticity in vivo does not support this view. Instead, we found that Th22 cells shared large expansions of unique clonotypes with the Th17, Th2a, and Th2 subsets. This pattern was recapitulated across all donors. Cluster feature analysis nonetheless suggested the existence of clonotypically discrete populations of *bona fide* Th22 cells.

Th1 and Th2 cells are widely considered to be the most stably differentiated subsets of effector/memory CD4⁺ T cells (*Zhou et al., 2009b*), both in vitro and in vivo (*Murphy and Stockinger, 2010; Murphy et al., 1996; Messi et al., 2003; Brown et al., 2015*). However, some central memory Th1 cells can produce large quantities of IL-4 under Th2-polarizing conditions (*Rivino et al., 2004*). Conversely, murine Th2 cells primed in vivo can acquire the ability to produce IFN- γ as well as IL-4 (*Hegazy et al., 2010*), whereas human Th2 cells seem to be more immutable (*Messi et al., 2003*). In functional terms, Th2 cells are clearly defined by the production of IL-4, but in phenotypic terms, the key lineage-defining markers remain a matter of debate, with most laboratories using either CCR4⁺CCR6⁻ or CCR6⁻CRTh2⁺ as the critical parameters. To bypass this controversy, we analyzed CCR4⁺CCR6⁻CRTh2⁻ (Th2) cells and CCR4⁺CCR6⁻CRTh2⁺ (Th2a) cells separately. The core repertoire of the Th2a subset was unique, implying a specialized function, but interestingly, both the Th2a and Th2 subsets shared clonal expansions with the Th22 subset. This unexpected finding nonetheless aligns with current revisions of the classic paradigm toward a more plastic view of Th1 cells (*Leipe et al., 2020*).

In contrast to Th1 and Th2 cells, Th17 cells and Tregs, including naturally occurring and peripherally induced Tregs, are thought to be inherently plastic (*Geginat et al., 2014*), especially in mice (*Cohen et al., 2011*). For example, murine and human Th17 cells differentiated in vivo can be induced to adopt a Th1-like or Th1-17-like phenotype in vitro (*Lee et al., 2009; Annunziato et al., 2007; Hirota et al., 2011*), and human Th17 cells migrating to sites of inflammation can acquire a Th1-like phenotype, characterized by the expression of CD161 as well as CCR6 (*Maggi et al., 2012*). We found that Th17 cells most commonly shared clonal expansions with Th22 cells, which also shared clonal expansions with Th2a and Th2 cells. Little is known about such transitions, in part because TGF- β 1 promotes the development of Th17 cells and inhibits the development of Th2 cells, which are consequently separated in most differentiation schemes (*Muranski and Restifo, 2013*). Further studies are therefore required to interpret these findings in mechanistic terms. Of note, we did not analyze Th9 cells, which appear to derive from Th17 cells under inflammatory conditions (*Beriou et al., 2010*) and are thought to be relatively unstable (*Schlapbach et al., 2014*).

The development of Th17 cells and Tregs in the thymus is linked due to shared microenvironmental factors that favor commitment to both lineages. These cells may also derive from common thymic progenitors (*Yang et al., 2008*). In vivo, Th17 cells have been shown to acquire certain regulatory features, including the ability to produce IL-10 under the influence of IL-12 or IL-27

(Heinemann et al., 2014). However, we found no evidence of interconversion between Th17 cells and Tregs, at least within the effector/memory pool of CD4⁺ T cells.

Th1-17 cells are thought to represent a more mature form of Th17 cells (Muranski and Restifo, 2013). Unexpectedly, we found that Th1-17 cells shared few or no clonotypes with Th17 cells, whereas clonal overlap was common between Th1 and Th1-17 cells. In line with this dichotomy, Th17 cells, but not Th1-17 cells, shared large clonal expansions with Th22 cells. The intermediate nature of Th1-17 cells has been predicted using computational models (Puniya et al., 2018) and observed directly in vitro (Zielinski et al., 2012). Ex vivo, Th1-17 cells are characterized by the coproduction IL-17, IL-22, and IFN- γ (Duhon and Campbell, 2014). Our data suggest that Th1-17 cells are more closely related to Th1 cells rather than Th17 cells, but nonetheless, the core Th1-17 repertoires were largely unique, suggesting that a majority of these cells occupy a distinct lineage and do not simply represent a maturation stage in the development of Th1 or Th17 cells.

Effector/memory CD4⁺ T-cell subsets are classified according to distinct patterns of cytokine production, reflecting differential expression of various master transcription factors. However, these profiles were largely established on the basis of in vitro studies, and consequently, our current understanding of subset phylogeny is most likely an oversimplification (Zhu et al., 2010). Mixed and unexplored subsets therefore almost certainly exist in vivo, reflecting nuances in the epigenetic landscape (Allan et al., 2012) and the relative activities of master regulators (Kanhere et al., 2012; Aune et al., 2009). Greater understanding of these complexities could inform efforts to develop more effective therapies for autoimmune diseases (Ryba-Stanisławowska et al., 2016; Rolla et al., 2014; Walker and von Herrath, 2016) and cancer (Kreiter et al., 2015; Borst et al., 2018; Wei et al., 2017), as well as better targeted vaccines (Misiak et al., 2017). In this context, our data provide an important step on the path to systematic deconvolution of the CD4⁺ T-cell compartment, specifically via the demonstration that subset fate is associated with the non-random selection of clonotypes expressing physicochemically distinct TCRs.

Materials and methods

Samples

Venous blood samples were collected from healthy adult donors ($n = 17$) directly into heparinized syringes or Vacutainer EDTA Tubes (BD Biosciences). Peripheral blood mononuclear cells (PBMCs) were isolated via density gradient centrifugation over Ficoll-Paque (PanEco) or Histopaque-1077 (Sigma-Aldrich). Ethical approval was granted by the institutional review committees at Cardiff University School of Medicine (16/55) and the Pirogov Russian National Research Medical University (2017/52). All donors provided written informed consent in accordance with the Declaration of Helsinki.

Flow cytometric sorting of effector/memory CD4⁺ T-cell subsets

PBMCs were stained immediately after isolation ($n = 5$ donors) with LIVE/DEAD Fixable Aqua (Thermo Fisher Scientific) and the following directly conjugated monoclonal antibodies: anti-CCR6-PE (clone 11A9), anti-CCR7-PE-Cy7 (clone 3D12), anti-CD14-V500 (clone M5E2), anti-CD19-V500 (clone HIB19), and anti-CRTh2-PE-CF594 (clone BM16) from BD Biosciences; anti-CCR4-BV605 (clone L291H4), anti-CD3-APC-Fire750 (clone SK7), anti-CD25-BV711 (clone MA251), anti-CD45RA-PE-Cy5 (clone HI100), anti-CD127-BV421 (clone A019D5), and anti-CXCR5-BV785 (clone J252D4) from BioLegend; anti-CCR10-APC (clone 314305) and anti-CXCR3-FITC (clone 49801.111) from R&D Systems; and anti-CD4-PE-Cy5.5 (clone S3.5) from Thermo Fisher Scientific. The gating strategy is described in **Figure 1—figure supplement 1** and **Table 1**. Subsets were flow-sorted at >98% purity after exclusion of naive CCR7⁺CD45RA⁺ events from the Aqua⁻CD3⁺CD4⁺CD14⁻CD19⁻ gate as Tfh cells (CXCR5⁺), Th1 cells (non-Tfh/Th22/Treg CCR4⁻CCR6⁻CXCR3⁺), Th1-17 cells (non-Tfh/Th22/Treg CCR4⁻CCR6⁺CXCR3⁺), Th17 cells (non-Tfh/Th22/Treg CCR4⁺CCR6⁺CXCR3⁻), Th22 cells (CCR10⁺), Th2a cells (non-Tfh/Th22/Treg CCR4⁺CCR6⁻CRTh2⁺CXCR3⁻), Th2 cells (non-Tfh/Th22/Treg CCR4⁺CCR6⁻CRTh2⁻CXCR3⁻), or Tregs (CD25^{high}CD127^{low}) using a modified FACSAria II (BD Biosciences). All cells ($n = 6,000$ – $150,000$ per subset) were sorted directly into RLT buffer (Qiagen) containing 1% 2-mercaptoethanol (Sigma-Aldrich). Subset frequencies are listed in **Table 2**. Acquisition and post-sort data were analyzed using FlowJo software version 10.6.1 (Tree Star).

Flow cytometric sorting of naive CD4⁺ T-cell subsets

To identify RTEs, mature naive T cells, and naive Tregs in the CD4⁺ lineage, PBMCs were stained immediately after isolation (n = 12 donors) with the following directly conjugated monoclonal antibodies: anti-CD4-PE (clone 13B8.2) and anti-CD27-PE-Cy5 (clone O323) from Beckman Coulter; and anti-CD25-eFluor450 (clone BC96), anti-CD31-PE-Cy7 (clone WM59), and anti-CD45RA-FITC (clone JS-83) from eBioscience. The gating strategy was described previously (Egorov *et al.*, 2018). Subsets were flow-sorted at >98% purity from the CD4⁺CD27⁺CD45RA⁺ gate as RTEs (CD25⁻CD31⁺), mature naive T cells (CD25⁻CD31⁻), or naive Tregs (CD25^{high}) using a FACS Aria III (BD Biosciences). To identify naive Th1-like cells, naive Th2-like cells, and naive Tregs in the CD4⁺ lineage, PBMCs were stained immediately after isolation (n = 4 donors) with LIVE/DEAD Fixable Aqua (Thermo Fisher Scientific) and the following directly conjugated monoclonal antibodies: anti-CCR7-PE-Cy7 (clone 3D12), anti-CD8-V500 (clone RPA-T8), anti-CD14-V500 (clone M5E2), and anti-CD19-V500 (clone HIB19) from BD Biosciences; anti-CCR4-BV605 (clone L291H4), anti-CD3-APC-Fire750 (clone SK7), anti-CD25-BV711 (clone MA251), anti-CD45RA-PE-Cy5 (clone HI100), anti-CD95-PE (clone DX2), and anti-CD127-BV421 (clone A019D5) from BioLegend; anti-CXCR3-FITC (clone 49801.111) from R&D Systems; and anti-CD4-PE-Cy5.5 (clone S3.5) from Thermo Fisher Scientific. The gating strategy is described in **Figure 6—figure supplement 2**. Subsets were flow-sorted at >98% purity from the Aqua⁻CD3⁺CD4⁺CD8⁻CD14⁻CD19⁻CCR7⁺CD45RA⁺CD95⁻ gate as naive Th1-like cells (non-Treg CCR4⁻CXCR3⁺), naive Th2-like cells (non-Treg CCR4⁺CXCR3⁻), or naive Tregs (CD25^{high}CD127^{low}), alongside the corresponding non-Treg CCR4⁻CXCR3⁻ and non-Treg CCR4⁺CXCR3⁺ populations, using a modified FACS Aria II (BD Biosciences). All cells (n = 260-150,000 per subset) were sorted directly into RLT buffer (Qiagen) containing 1% 2-mercaptoethanol (Sigma-Aldrich). Subset frequencies are listed in **Table 3**. Acquisition and post-sort data were analyzed using FlowJo software version 10.6.1 (Tree Star).

TCR sequencing and data analysis

TCR α and TCR β cDNA libraries were prepared using a Human TCR Kit (MiLaboratory LLC) with template switch-based incorporation of UMIs as described previously (Egorov *et al.*, 2015). Libraries were sequenced in paired-end mode (150 + 150 bp) on a NextSeq500 (Illumina). Raw sequence data were analyzed using MIGEC software version 1.2.9 (Shugay *et al.*, 2014). Briefly, UMI sequences were extracted from demultiplexed data using the Checkout utility, yielding sample barcode matches in ~90% of cases. Data were then assembled using the erroneous UMI filtering option in the Assemble utility. For most tasks, the minimum required number of reads per UMI was set at 1. For analyses of overlap and publicity, which are sensitive to even minor cross-sample contaminations, the minimum required number of reads per UMI was set at 3 (Egorov *et al.*, 2015). In-frame TCR α and TCR β repertoires were extracted using MiXCR software version 2.1.1 (Bolotin *et al.*, 2017; Bolotin *et al.*, 2018; Bolotin *et al.*, 2015). At a threshold of 3 reads per UMI, the number of obtained UMI-labeled cDNA molecules per repertoire per sample ranged from 5300 to 303,500, and the number of CDR3 clonotype variants at the nucleotide level per repertoire per sample ranged from 1200 to 83,200. Normalization, data transformation, in-depth analyses, and statistical calculations were performed using R scripts and VDJtools software version 1.2.1 (Shugay *et al.*, 2015). Analyses of averaged CDR3 characteristics were weighted by the abundance of each clonotype in each sample. Basic characteristics included CDR3 length, the number of N additions, interaction strength, hydrophobicity (Kidera factor 4), volume, and surface, which were selected in previous

Table 3. Frequencies of sorted naive CD4⁺ T-cell subsets.

Donor	Th1-like CCR4 ⁻ CXCR3 ⁺	Th2-like CCR4 ⁺ CXCR3 ⁻	CCR4 ⁻ CXCR3 ⁻	CCR4 ⁺ CXCR3 ⁺	Treg CD25 ^{high} CD127 ^{low}
D1	1.75	5.46	44.80	0.23	0.73
D2	0.77	6.77	20.40	0.32	0.57
D3	0.15	5.67	42.60	0.19	1.70
D4	0.16	6.33	33.10	0.05	1.11

Shown as % of live CD3⁺CD4⁺CD8⁻CD14⁻CD19⁻ naive cells. Details in **Figure 6—figure supplement 2**.

analyses of various somatically rearranged lymphocyte receptor datasets (Izraelson *et al.*, 2018; Egorov *et al.*, 2018; Davydov *et al.*, 2018). The amino acid properties used in these analyses can be viewed at https://github.com/mikessh/vdjtools/blob/master/src/main/resources/profile/aa_property_table.txt. The strength feature reflects the predicted sum of interaction affinities between pairs of amino acids at the TCR-pMHC interface, Kidera factor 4 reflects the abundance of hydrophobic amino acids on an inverted scale, and the surface characteristic reflects the relative abundance of amino acids with no predicted changes in accessibility during TCR engagement with cognate pMHCs. Amino acid hierarchies by probability of active involvement at the protein-protein interface or conformational stability relative to the native form in the absence of an interaction were derived from previous work (Martin and Lavery, 2012), in which extensive cross-docking experiments were performed across 198 proteins and 300 partners *in silico* to infer the general roles of amino acids at protein-protein interfaces. Physicochemical characteristics were calculated and averaged for the five amino acid residues located in the middle of each CDR3 loop, which are most likely to contact the peptide epitope in any cognate pMHC (Egorov *et al.*, 2018). Principal component analysis was performed using 28 parameters computed as the average across each CDR3 α and CDR3 β cloneset: Kidera factors ($n = 10$), strength, mjenery, count (CDR3 length), NDN length, number of N insertions, vdins, djins, core, rim, volume, polarity, disorder, surface, alpha, beta, turn, charge, and hydrophathy (VDJtools software version 1.2.1). No significant variations in V/J segment use were detected among subsets (data not shown). Network visualization was performed using Cytoscape (<https://cytoscape.org>). Repertoire overlap was analyzed using the unweighted D (reflecting the proportion of shared clonotypes between paired repertoires) and weighted F2 (reflecting the proportion of shared T cells between paired repertoires) metrics in VDJtools software version 1.2.1.

Quantification and statistical analysis

Statistical analyses were performed on processed datasets in R. Multiple parameter inferences were estimated using ANOVA if the data were distributed normally or the Kruskal-Wallis test if any of the data were not distributed normally. The corresponding p values were calculated using the two-sample Welch t-test or the Wilcoxon rank sum test. The false discovery rate was controlled using Benjamini-Hochberg correction unless stated otherwise. Post-hoc tests were performed using the ggpubr package (<https://CRAN.R-project.org/package=ggpubr>).

Acknowledgements

This work was supported by grants from the Ministry of Science and Higher Education of the Russian Federation (075-15-2019-1789) and the Wellcome Trust (100326/Z/12/Z).

Additional information

Funding

Funder	Grant reference number	Author
Ministry of Science and Higher Education	075-15-2019-1789	Dmitriy M Chudakov
Wellcome Trust	100326/Z/12/Z	David A Price

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Sofya A Kasatskaya, Data curation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Kristin Ladell, Data curation, Investigation, Methodology, Writing - original draft; Evgeniy S Egorov, Kelly L Miners, Maria Metsger, Dmitry B Staroverov, Irina A Shagina, Ilgar Z Mamedov, Investigation, Methodology; Alexey N Davydov, Data curation, Investigation, Methodology; Elena K Matveyshina, Mark Izraelson, Pavel V Shelyakin, Data curation; Olga V Britanova, Data curation, Supervision, Writing - original draft; David A Price, Conceptualization, Supervision, Funding acquisition, Writing - original draft, Writing - review and editing; Dmitriy M Chudakov,

Conceptualization, Supervision, Funding acquisition, Validation, Investigation, Visualization, Writing - original draft, Project administration, Writing - review and editing

Author ORCIDs

Elena K Matveyshina  <http://orcid.org/0000-0003-4641-4906>

David A Price  <https://orcid.org/0000-0001-9416-2737>

Dmitriy M Chudakov  <https://orcid.org/0000-0003-0430-790X>

Ethics

Human subjects: ethical approval was granted by the institutional review committees at Cardiff University School of Medicine (reference number 16/55) and the Pirogov Russian National Research Medical University (protocol number 2017/52), and all donors provided written informed consent in accordance with the Declaration of Helsinki.

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.57063.sa1>

Author response <https://doi.org/10.7554/eLife.57063.sa2>

Additional files

Supplementary files

- Transparent reporting form

Data availability

All extracted repertoires and metadata are available in Figshare: <https://figshare.com/s/2145b1b16c6854445af7> and <https://figshare.com/s/84ec5f412356afb0536d>. Deep TCR profiling data were deposited in the GEO under accession code GSE158848.

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Kasatskaya SA, Ladell K, Miners KL, Davydov AN, Britanova OV, Price DA, Chudakov DM	2020	Human effector/memory CD4 ⁺ T cell subsets: deep TCR profiling.	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE158848	NCBI Gene Expression Omnibus, GSE158848
Kasatskaya SA, Ladell K, Miners KL, Davydov AN, Britanova OV, Price DA, Chudakov DM	2020	TCR repertoires in human CD4 ⁺ T cell subsets: effector/memory subsets, TCR alpha & TCR beta chain sequencing	https://figshare.com/s/2145b1b16c6854445af7	figshare, 10.6084/m9.figshare.2145b1b16c6854445af7
Kasatskaya SA, Ladell K, Egorov ES, Miners KL, Staroverov DB, Britanova OV, Price DA, Chudakov DM	2020	TCR beta repertoires in naive CD4 ⁺ T cell subsets in identical twin donor samples	https://figshare.com/s/84ec5f412356afb0536d	figshare, 10.6084/m9.figshare.84ec5f412356afb0536d

References

- Adams JJ, Narayanan S, Liu B, Birnbaum ME, Kruse AC, Bowerman NA, Chen W, Levin AM, Connolly JM, Zhu C, Kranz DM, Garcia KC. 2011. T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity* **35**:681–693. DOI: <https://doi.org/10.1016/j.immuni.2011.09.013>, PMID: 22101157
- Akdis M, Palomares O, van de Veen W, van Splunter M, Akdis CA. 2012. Th17 and Th22 cells: a confusion of antimicrobial response with tissue inflammation versus protection. *Journal of Allergy and Clinical Immunology* **129**:1438–1449. DOI: <https://doi.org/10.1016/j.jaci.2012.05.003>, PMID: 22657405
- Akkaya B, Oya Y, Akkaya M, Al Souz J, Holstein AH, Kamenyeva O, Kabat J, Matsumura R, Dorward DW, Glass DD, Shevach EM. 2019. Regulatory T cells mediate specific suppression by depleting peptide-MHC class II from

- dendritic cells. *Nature Immunology* **20**:218–231. DOI: <https://doi.org/10.1038/s41590-018-0280-2>, PMID: 30643268
- Alabbas SY**, Begun J, Florin TH, Oancea I. 2018. The role of IL-22 in the resolution of sterile and nonsterile inflammation. *Clinical & Translational Immunology* **7**:e1017. DOI: <https://doi.org/10.1002/cti2.1017>, PMID: 29713472
- Allan RS**, Zueva E, Cammas F, Schreiber HA, Masson V, Belz GT, Roche D, Maison C, Quivy JP, Almouzni G, Amigorena S. 2012. An epigenetic silencing pathway controlling T helper 2 cell lineage commitment. *Nature* **487**:249–253. DOI: <https://doi.org/10.1038/nature11173>, PMID: 22763435
- Annuziato F**, Cosmi L, Santarlasci V, Maggi L, Liotta F, Mazzinghi B, Parente E, Fili L, Ferri S, Frosali F, Giudici F, Romagnani P, Parronchi P, Tonelli F, Maggi E, Romagnani S. 2007. Phenotypic and functional features of human Th17 cells. *Journal of Experimental Medicine* **204**:1849–1861. DOI: <https://doi.org/10.1084/jem.20070663>, PMID: 17635957
- Arvey A**, van der Veecken J, Plitas G, Rich SS, Concannon P, Rudensky AY. 2015. Genetic and epigenetic variation in the lineage specification of regulatory T cells. *eLife* **4**:e07571. DOI: <https://doi.org/10.7554/eLife.07571>, PMID: 26510014
- Aune TM**, Collins PL, Chang S. 2009. Epigenetics and T helper 1 differentiation. *Immunology* **126**:299–305. DOI: <https://doi.org/10.1111/j.1365-2567.2008.03026.x>, PMID: 19178593
- Bacher P**, Heinrich F, Stervbo U, Nienen M, Vahldieck M, Iwert C, Vogt K, Kollet J, Babel N, Sawitzki B, Schwarz C, Bereswill S, Heimesaat MM, Heine G, Gadermaier G, Asam C, Assenmacher M, Kniemeyer O, Brakhage AA, Ferreira F, et al. 2016. Regulatory T cell specificity directs tolerance versus allergy against aeroantigens in humans. *Cell* **167**:1067–1078. DOI: <https://doi.org/10.1016/j.cell.2016.09.050>, PMID: 27773482
- Bacher P**, Scheffold A. 2018. Antigen-specific regulatory T-cell responses against aeroantigens and their role in allergy. *Mucosal Immunology* **11**:1537–1550. DOI: <https://doi.org/10.1038/s41385-018-0038-z>, PMID: 29858582
- Barberis M**, Helikar T, Verbruggen P. 2018. Simulation of stimulation: cytokine dosage and cell cycle crosstalk driving timing-dependent T cell differentiation. *Frontiers in Physiology* **9**:879. DOI: <https://doi.org/10.3389/fphys.2018.00879>, PMID: 30116196
- Baumjohann D**, Ansel KM. 2015. Tracking early T follicular helper cell differentiation in vivo. *Methods in Molecular Biology* **1291**:27–38. DOI: https://doi.org/10.1007/978-1-4939-2498-1_3, PMID: 25836299
- Bentebibel SE**, Lopez S, Obermoser G, Schmitt N, Mueller C, Harrod C, Flano E, Mejias A, Albrecht RA, Blankenship D, Xu H, Pascual V, Banchereau J, Garcia-Sastre A, Palucka AK, Ramilo O, Ueno H. 2013. Induction of ICOS⁺CXCR3⁺CXCR5⁺ TH cells correlates with antibody responses to influenza vaccination. *Science Translational Medicine* **5**:176ra32. DOI: <https://doi.org/10.1126/scitranslmed.3005191>, PMID: 23486778
- Beriou G**, Bradshaw EM, Lozano E, Costantino CM, Hastings WD, Orban T, Elyaman W, Khoury SJ, Kuchroo VK, Baecher-Allan C, Hafler DA. 2010. TGF- β induces IL-9 production from human Th17 cells. *Journal of Immunology* **185**:46–54. DOI: <https://doi.org/10.4049/jimmunol.1000356>, PMID: 20498357
- Bolotin DA**, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, Chudakov DM. 2015. MiXCR: software for comprehensive adaptive immunity profiling. *Nature Methods* **12**:380–381. DOI: <https://doi.org/10.1038/nmeth.3364>, PMID: 25924071
- Bolotin DA**, Poslavsky S, Davydov AN, Frenkel FE, Fanchi L, Zolotareva OI, Hemmers S, Putintseva EV, Obraztsova AS, Shugay M, Ataullakhanov RI, Rudensky AY, Schumacher TN, Chudakov DM. 2017. Antigen receptor repertoire profiling from RNA-seq data. *Nature Biotechnology* **35**:908–911. DOI: <https://doi.org/10.1038/nbt.3979>, PMID: 29020005
- Bolotin DA**, Poslavsky S, Davydov AN, Chudakov DM. 2018. Reply to "Evaluation of immune repertoire inference methods from RNA-seq data". *Nature Biotechnology* **36**:1035–1036. DOI: <https://doi.org/10.1038/nbt.4296>, PMID: 30412204
- Bonelli M**, Savitskaya A, von Dalwigk K, Steiner CW, Aletaha D, Smolen JS, Scheinecker C. 2008. Quantitative and qualitative deficiencies of regulatory T cells in patients with systemic lupus erythematosus (SLE). *International Immunology* **20**:861–868. DOI: <https://doi.org/10.1093/intimm/dxn044>, PMID: 18469329
- Booth NJ**, McQuaid AJ, Sobande T, Kissane S, Agius E, Jackson SE, Salmon M, Falciani F, Yong K, Rustin MH, Akbar AN, Vukmanovic-Stejic M. 2010. Different proliferative potential and migratory characteristics of human CD4⁺ regulatory T cells that express either CD45RA or CD45RO. *Journal of Immunology* **184**:4317–4326. DOI: <https://doi.org/10.4049/jimmunol.0903781>, PMID: 20231690
- Borst J**, Ahrends T, Bąbala N, Melief CJM, Kastenmüller W. 2018. CD4⁺ T cell help in cancer immunology and immunotherapy. *Nature Reviews Immunology* **18**:635–647. DOI: <https://doi.org/10.1038/s41577-018-0044-0>, PMID: 30057419
- Brenna E**, Davydov AN, Ladell K, McLaren JE, Bonaiuti P, Metsger M, Ramsden JD, Gilbert SC, Lambe T, Price DA, Champion SL, Chudakov DM, Borrow P, McMichael AJ. 2020. CD4⁺ T follicular helper cells in human tonsils and blood are clonally convergent but divergent from non-Tfh CD4⁺ cells. *Cell Reports* **30**:137–152. DOI: <https://doi.org/10.1016/j.celrep.2019.12.016>, PMID: 31914381
- Brown CC**, Esterhazy D, Sarde A, London M, Pullabhatla V, Osmá-García I, Al-Bader R, Ortiz C, Elgueta R, Arno M, de Rinaldis E, Mucida D, Lord GM, Noelle RJ. 2015. Retinoic acid is essential for Th1 cell lineage stability and prevents transition to a Th17 cell program. *Immunity* **42**:499–511. DOI: <https://doi.org/10.1016/j.immuni.2015.02.003>, PMID: 25769610
- Chakrabarti P**, Bhattacharyya R. 2007. Geometry of nonbonded interactions involving planar groups in proteins. *Progress in Biophysics and Molecular Biology* **95**:83–137. DOI: <https://doi.org/10.1016/j.pbiomolbio.2007.03.016>, PMID: 17629549

- Chung Y**, Tanaka S, Chu F, Nurieva RI, Martinez GJ, Rawal S, Wang YH, Lim H, Reynolds JM, Zhou XH, Fan HM, Liu ZM, Neelapu SS, Dong C. 2011. Follicular regulatory T cells expressing Foxp3 and Bcl-6 suppress germinal center reactions. *Nature Medicine* **17**:983–988. DOI: <https://doi.org/10.1038/nm.2426>, PMID: 21785430
- Cohen CJ**, Crome SQ, MacDonald KG, Dai EL, Mager DL, Levings MK. 2011. Human Th1 and Th17 cells exhibit epigenetic stability at signature cytokine and transcription factor loci. *Journal of Immunology* **187**:5615–5626. DOI: <https://doi.org/10.4049/jimmunol.1101058>, PMID: 22048764
- Constant SL**, Bottomly K. 1997. Induction of Th1 and Th2 CD4⁺ T cell responses: the alternative approaches. *Annual Review of Immunology* **15**:297–322. DOI: <https://doi.org/10.1146/annurev.immunol.15.1.297>, PMID: 9143690
- Cook MC**, Tangye SG. 2009. Primary immune deficiencies affecting lymphocyte differentiation: lessons from the spectrum of resulting infections. *International Immunology* **21**:1003–1011. DOI: <https://doi.org/10.1093/intimm/dxp076>, PMID: 19651645
- Corse E**, Gottschalk RA, Allison JP. 2011. Strength of TCR-peptide/MHC interactions and in vivo T cell responses. *Journal of Immunology* **186**:5039–5045. DOI: <https://doi.org/10.4049/jimmunol.1003650>, PMID: 21505216
- Costa N**, Marques O, Godinho SI, Carvalho C, Leal B, Figueiredo AM, Vasconcelos C, Marinho A, Moraes-Fontes MF, Gomes da Costa A, Ponte C, Campanillo-Marques R, Córias T, Martins AR, Viana JF, Lima M, Martins B, Fesel C. 2017. Two separate effects contribute to regulatory T cell defect in systemic lupus erythematosus patients and their unaffected relatives. *Clinical & Experimental Immunology* **189**:318–330. DOI: <https://doi.org/10.1111/cei.12991>, PMID: 28542701
- Coutinho A**, Caramalho I, Seixas E, Demengeot J. 2005. Thymic commitment of regulatory T cells is a pathway of TCR-dependent selection that isolates repertoires undergoing positive or negative selection. *Current Topics in Microbiology and Immunology* **293**:43–71. DOI: https://doi.org/10.1007/3-540-27702-1_3, PMID: 15981475
- Darrigues J**, van Meerwijk JPM, Romagnoli P. 2018. Age-dependent changes in regulatory T lymphocyte development and function: a mini-review. *Gerontology* **64**:28–35. DOI: <https://doi.org/10.1159/000478044>, PMID: 28704827
- Davydov AN**, Obratsova AS, Lebedin MY, Turchaninova MA, Staroverov DB, Merzlyak EM, Sharonov GV, Kladova O, Shugay M, Britanova OV, Chudakov DM. 2018. Comparative analysis of B-cell receptor repertoires induced by live yellow fever vaccine in young and middle-age donors. *Frontiers in Immunology* **9**:2309. DOI: <https://doi.org/10.3389/fimmu.2018.02309>, PMID: 30356675
- De Simone G**, Mazza EMC, Cassotta A, Davydov AN, Kuka M, Zanon V, De Paoli F, Scamardella E, Metsger M, Roberto A, Pilipow K, Colombo FS, Tenedini E, Tagliafico E, Gattinoni L, Mavilio D, Peano C, Price DA, Singh SP, Farber JM, et al. 2019. CXCR3 identifies human naive CD8⁺ T cells with enhanced effector differentiation potential. *Journal of Immunology* **203**:3179–3189. DOI: <https://doi.org/10.4049/jimmunol.1901072>, PMID: 31740485
- Duhen T**, Campbell DJ. 2014. IL-1 β promotes the differentiation of polyfunctional human CCR6⁺CXCR3⁺ Th1/17 cells that are specific for pathogenic and commensal microbes. *Journal of Immunology* **193**:120–129. DOI: <https://doi.org/10.4049/jimmunol.1302734>, PMID: 24890729
- DuPage M**, Bluestone JA. 2016. Harnessing the plasticity of CD4⁺ T cells to treat immune-mediated disease. *Nature Reviews Immunology* **16**:149–163. DOI: <https://doi.org/10.1038/nri.2015.18>, PMID: 26875830
- Egorov ES**, Merzlyak EM, Shelenkov AA, Britanova OV, Sharonov GV, Staroverov DB, Bolotin DA, Davydov AN, Barsova E, Lebedev YB, Shugay M, Chudakov DM. 2015. Quantitative profiling of immune repertoires for minor lymphocyte counts using unique molecular identifiers. *Journal of Immunology* **194**:6155–6163. DOI: <https://doi.org/10.4049/jimmunol.1500215>, PMID: 25957172
- Egorov ES**, Kasatskaya SA, Zubov VN, Izraelson M, Nakonechnaya TO, Staroverov DB, Angius A, Cucca F, Mamedov IZ, Rosati E, Franke A, Shugay M, Pogorely MV, Chudakov DM, Britanova OV. 2018. The changing landscape of naive T cell receptor repertoire with human aging. *Frontiers in Immunology* **9**:1618. DOI: <https://doi.org/10.3389/fimmu.2018.01618>, PMID: 30087674
- Eisenbarth SC**. 2019. Dendritic cell subsets in T cell programming: location dictates function. *Nature Reviews Immunology* **19**:89–103. DOI: <https://doi.org/10.1038/s41577-018-0088-1>, PMID: 30464294
- Eyerich S**, Eyerich K, Pennino D, Carbone T, Nasorri F, Pallotta S, Cianfarani F, Odorisio T, Traidl-Hoffmann C, Behrendt H, Durham SR, Schmidt-Weber CB, Cavani A. 2009. Th22 cells represent a distinct human T cell subset involved in epidermal immunity and remodeling. *Journal of Clinical Investigation* **119**:3573–3585. DOI: <https://doi.org/10.1172/JCI40202>, PMID: 19920355
- Fazilleau N**, McHeyzer-Williams LJ, Rosen H, McHeyzer-Williams MG. 2009a. The function of follicular helper T cells is regulated by the strength of T cell antigen receptor binding. *Nature Immunology* **10**:375–384. DOI: <https://doi.org/10.1038/ni.1704>, PMID: 19252493
- Fazilleau N**, Mark L, McHeyzer-Williams LJ, McHeyzer-Williams MG. 2009b. Follicular helper T cells: lineage and location. *Immunity* **30**:324–335. DOI: <https://doi.org/10.1016/j.immuni.2009.03.003>, PMID: 19303387
- Feng Y**, van der Veeken J, Shugay M, Putintseva EV, Osmanbeyoglu HU, Dikiy S, Hoyos BE, Moltedo B, Hemmers S, Treuting P, Leslie CS, Chudakov DM, Rudensky AY. 2015. A mechanism for expansion of regulatory T-cell repertoire and its role in self-tolerance. *Nature* **528**:132–136. DOI: <https://doi.org/10.1038/nature16141>, PMID: 26605529
- Fink PJ**. 2013. The biology of recent thymic emigrants. *Annual Review of Immunology* **31**:31–50. DOI: <https://doi.org/10.1146/annurev-immunol-032712-100010>, PMID: 23121398
- Finotto S**. 2008. T-cell regulation in asthmatic diseases. *Chemical Immunology and Allergy* **94**:83–92. DOI: <https://doi.org/10.1159/000154869>, PMID: 18802339

- Gartlan KH**, Bommiasamy H, Paz K, Wilkinson AN, Owen M, Reichenbach DK, Banovic T, Wehner K, Buchanan F, Varelias A, Kuns RD, Chang K, Fedoriw Y, Shea T, Coghill J, Zaiken M, Plank MW, Foster PS, Clouston AD, Blazar BR, et al. 2018. A critical role for donor-derived IL-22 in cutaneous chronic GVHD. *American Journal of Transplantation* **18**:810–820. DOI: <https://doi.org/10.1111/ajt.14513>, PMID: 28941323
- Geginat J**, Paroni M, Maglie S, Alfen JS, Kastirr I, Gruarin P, De Simone M, Pagani M, Abrignani S. 2014. Plasticity of human CD4 T cell subsets. *Frontiers in Immunology* **5**:630. DOI: <https://doi.org/10.3389/fimmu.2014.00630>, PMID: 25566245
- Grimsholm O**, Piano Mortari E, Davydov AN, Shugay M, Obratsova AS, Bocci C, Marasco E, Marcellini V, Aranburu A, Farroni C, Silvestris DA, Cristofolletti C, Giorda E, Scarsella M, Cascioli S, Barresi S, Lougaris V, Plebani A, Cancrini C, Finocchi A, et al. 2020. The interplay between CD27^{dull} and CD27^{bright} B cells ensures the flexibility, stability, and resilience of human B cell memory. *Cell Reports* **30**:2963–2977. DOI: <https://doi.org/10.1016/j.celrep.2020.02.022>, PMID: 32130900
- Groom JR**, Richmond J, Murooka TT, Sorensen EW, Sung JH, Bankert K, von Andrian UH, Moon JJ, Mempel TR, Luster AD. 2012. CXCR3 chemokine receptor-ligand interactions in the lymph node optimize CD4⁺ T helper 1 cell differentiation. *Immunity* **37**:1091–1103. DOI: <https://doi.org/10.1016/j.immuni.2012.08.016>, PMID: 23123063
- Häringer B**, Lozza L, Steckel B, Geginat J. 2009. Identification and characterization of IL-10/IFN- γ -producing effector-like T cells with regulatory function in human blood. *Journal of Experimental Medicine* **206**:1009–1017. DOI: <https://doi.org/10.1084/jem.20082238>, PMID: 19414553
- Hegazy AN**, Peine M, Helmstetter C, Panse I, Fröhlich A, Bergthaler A, Flatz L, Pinschewer DD, Radbruch A, Löhning M. 2010. Interferons direct Th2 cell reprogramming to generate a stable GATA-3⁺T-bet⁺ cell subset with combined Th2 and Th1 cell functions. *Immunity* **32**:116–128. DOI: <https://doi.org/10.1016/j.immuni.2009.12.004>, PMID: 20079668
- Heinemann C**, Heink S, Petermann F, Vasanthakumar A, Rothhammer V, Doorduyn E, Mitsdoerffer M, Sie C, Prazeres da Costa O, Buch T, Hemmer B, Oukka M, Kallies A, Korn T. 2014. IL-27 and IL-12 oppose pro-inflammatory IL-23 in CD4⁺ T cells by inducing Blimp1. *Nature Communications* **5**:3770. DOI: <https://doi.org/10.1038/ncomms4770>, PMID: 24796719
- Hernández-Santos N**, Huppler AR, Peterson AC, Khader SA, McKenna KC, Gaffen SL. 2013. Th17 cells confer long-term adaptive immunity to oral mucosal *Candida albicans* infections. *Mucosal Immunology* **6**:900–910. DOI: <https://doi.org/10.1038/mi.2012.128>, PMID: 23250275
- Hirota K**, Duarte JH, Veldhoen M, Hornsby E, Li Y, Cua DJ, Ahlfors H, Wilhelm C, Tolaini M, Menzel U, Garafalaki A, Potocnik AJ, Stockinger B. 2011. Fate mapping of IL-17-producing T cells in inflammatory responses. *Nature Immunology* **12**:255–263. DOI: <https://doi.org/10.1038/ni.1993>, PMID: 21278737
- Hoffmann P**, Eder R, Boeld TJ, Doser K, Piseshka B, Andreesen R, Edinger M. 2006. Only the CD45RA⁺ subpopulation of CD4⁺CD25^{high} T cells gives rise to homogeneous regulatory T-cell lines upon in vitro expansion. *Blood* **108**:4260–4267. DOI: <https://doi.org/10.1182/blood-2006-06-027409>, PMID: 16917003
- Hoffmann T**, Krackhardt AM, Antes I. 2015. Quantitative analysis of the association angle between T-cell receptor V α /V β domains reveals important features for epitope recognition. *PLOS Computational Biology* **11**:e1004244. DOI: <https://doi.org/10.1371/journal.pcbi.1004244>, PMID: 26185983
- Izraelson M**, Nakonechnaya TO, Moltedo B, Egorov ES, Kasatskaya SA, Putintseva EV, Mamedov IZ, Staroverov DB, Shemiakina II, Zakharova MY, Davydov AN, Bolotin DA, Shugay M, Chudakov DM, Rudensky AY, Britanova OV. 2018. Comparative analysis of murine T-cell receptor repertoires. *Immunology* **153**:133–144. DOI: <https://doi.org/10.1111/imm.12857>, PMID: 29080364
- Jordan MS**, Boesteanu A, Reed AJ, Petrone AL, Hohenbeck AE, Lerman MA, Naji A, Caton AJ. 2001. Thymic selection of CD4⁺CD25⁺ regulatory T cells induced by an agonist self-peptide. *Nature Immunology* **2**:301–306. DOI: <https://doi.org/10.1038/86302>, PMID: 11276200
- Kanhere A**, Hertweck A, Bhatia U, Gökmen MR, Perucha E, Jackson I, Lord GM, Jenner RG. 2012. T-bet and GATA3 orchestrate Th1 and Th2 differentiation through lineage-specific targeting of distal regulatory elements. *Nature Communications* **3**:1268. DOI: <https://doi.org/10.1038/ncomms2260>, PMID: 23232398
- Kidera A**, Konishi Y, Oka M, Ooi T, Scheraga HA. 1985. Statistical analysis of the physical properties of the 20 naturally occurring amino acids. *Journal of Protein Chemistry* **4**:23–55. DOI: <https://doi.org/10.1007/BF01025492>
- Kilpatrick RD**, Rickabaugh T, Hultin LE, Hultin P, Hausner MA, Detels R, Phair J, Jamieson BD. 2008. Homeostasis of the naive CD4⁺ T cell compartment during aging. *Journal of Immunology* **180**:1499–1507. DOI: <https://doi.org/10.4049/jimmunol.180.3.1499>, PMID: 18209045
- Kosmrlj A**, Jha AK, Huseby ES, Kardar M, Chakraborty AK. 2008. How the thymus designs antigen-specific and self-tolerant T cell receptor sequences. *PNAS* **105**:16671–16676. DOI: <https://doi.org/10.1073/pnas.0808081105>, PMID: 18946038
- Kosmrlj A**, Read EL, Qi Y, Allen TM, Altfeld M, Deeks SG, Pereyra F, Carrington M, Walker BD, Chakraborty AK. 2010. Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature* **465**:350–354. DOI: <https://doi.org/10.1038/nature08997>, PMID: 20445539
- Krebs CF**, Steinmetz OM. 2016. CD4⁺ T cell fate in glomerulonephritis: a tale of Th1, Th17, and novel Treg subtypes. *Mediators of Inflammation* **2016**:5393894. DOI: <https://doi.org/10.1155/2016/5393894>, PMID: 27974866
- Kreiter S**, Vormehr M, van de Roemer N, Diken M, Löwer M, Diekmann J, Boegel S, Schrörs B, Vascotto F, Castle JC, Tadmor AD, Schoenberger SP, Huber C, Türeci Ö, Sahin U. 2015. Mutant MHC class II epitopes

- drive therapeutic immune responses to cancer. *Nature* **520**:692–696. DOI: <https://doi.org/10.1038/nature14426>, PMID: 25901682
- Lathrop SK, Santacruz NA, Pham D, Luo J, Hsieh C-S. 2008. Antigen-specific peripheral shaping of the natural regulatory T cell population. *Journal of Experimental Medicine* **205**:3105–3117. DOI: <https://doi.org/10.1084/jem.20081359>, PMID: 19064700
- Lee YK, Turner H, Maynard CL, Oliver JR, Chen D, Elson CO, Weaver CT. 2009. Late developmental plasticity in the T helper 17 lineage. *Immunity* **30**:92–107. DOI: <https://doi.org/10.1016/j.immuni.2008.11.005>, PMID: 19119024
- Lei H, Kuchenbecker L, Streitz M, Sawitzki B, Vogt K, Landwehr-Kenzel S, Millward J, Juelke K, Babel N, Neumann A, Reinke P, Volk HD. 2015. Human CD45RA[−] FoxP3^{hi} memory-type regulatory T cells show distinct TCR repertoires with conventional T cells and play an important role in controlling early immune activation. *American Journal of Transplantation* **15**:2625–2635. DOI: <https://doi.org/10.1111/ajt.13315>, PMID: 25988290
- Leipe J, Pirronello F, Klose A, Schulze-Koops H, Skapenko A. 2020. Increased plasticity of non-classic Th1 cells toward the Th17 phenotype. *Modern Rheumatology* **30**:930–936. DOI: <https://doi.org/10.1080/14397595.2019.1667473>, PMID: 31512538
- Levine AG, Arvey A, Jin W, Rudensky AY. 2014. Continuous requirement for the TCR in regulatory T cell function. *Nature Immunology* **15**:1070–1078. DOI: <https://doi.org/10.1038/ni.3004>, PMID: 25263123
- Li MO, Rudensky AY. 2016. T cell receptor signalling in the control of regulatory T cell differentiation and function. *Nature Reviews Immunology* **16**:220–233. DOI: <https://doi.org/10.1038/nri.2016.26>, PMID: 27026074
- Linterman MA, Pierson W, Lee SK, Kallies A, Kawamoto S, Rayner TF, Srivastava M, Divekar DP, Beaton L, Hogan JJ, Fagarasan S, Liston A, Smith KG, Vinuesa CG. 2011. Foxp3⁺ follicular regulatory T cells control the germinal center response. *Nature Medicine* **17**:975–982. DOI: <https://doi.org/10.1038/nm.2425>, PMID: 21785433
- Logunova NN, Kriukova VV, Shelyakin PV, Egorov ES, Pereverzeva A, Bozhanova NG, Shugay M, Shcherbinin DS, Pogorelyy MV, Merzlyak EM, Zubov VN, Meiler J, Chudakov DM, Apt AS, Britanova OV. 2020. MHC-II alleles shape the CDR3 repertoires of conventional and regulatory naïve CD4⁺ T cells. *PNAS* **117**:13659–13669. DOI: <https://doi.org/10.1073/pnas.2003170117>, PMID: 32482872
- Maceiras AR, Almeida SCP, Mariotti-Ferrandiz E, Chaara W, Jebbawi F, Six A, Hori S, Klatzmann D, Faro J, Graca L. 2017. T follicular helper and T follicular regulatory cells have different TCR specificity. *Nature Communications* **8**:15067. DOI: <https://doi.org/10.1038/ncomms15067>, PMID: 28429709
- Maggi L, Santarlasci V, Capone M, Rossi MC, Querci V, Mazzoni A, Cimaz R, De Palma R, Liotta F, Maggi E, Romagnani S, Cosmi L, Annunziato F. 2012. Distinctive features of classic and nonclassic (Th17 derived) human Th1 cells. *European Journal of Immunology* **42**:3180–3188. DOI: <https://doi.org/10.1002/eji.201242648>, PMID: 22965818
- Marks BR, Nowyhed HN, Choi JY, Poholek AC, Odegard JM, Flavell RA, Craft J. 2009. Thymic self-reactivity selects natural interleukin 17-producing T cells that can regulate peripheral inflammation. *Nature Immunology* **10**:1125–1132. DOI: <https://doi.org/10.1038/ni.1783>, PMID: 19734905
- Martin J, Lavery R. 2012. Arbitrary protein-protein docking targets biologically relevant interfaces. *BMC Biophysics* **5**:7. DOI: <https://doi.org/10.1186/2046-1682-5-7>, PMID: 22559010
- McClymont SA, Putnam AL, Lee MR, Esensten JH, Liu W, Hulme MA, Hoffmüller U, Baron U, Olek S, Bluestone JA, Brusko TM. 2011. Plasticity of human regulatory T cells in healthy subjects and patients with type 1 diabetes. *Journal of Immunology* **186**:3918–3926. DOI: <https://doi.org/10.4049/jimmunol.1003099>, PMID: 21368230
- McDonald DR. 2012. TH17 deficiency in human disease. *Journal of Allergy and Clinical Immunology* **129**:1429–1435. DOI: <https://doi.org/10.1016/j.jaci.2012.03.034>, PMID: 22554706
- McGee HS, Agrawal DK. 2006. TH2 cells in the pathogenesis of airway remodeling: regulatory T cells a plausible Panacea for asthma. *Immunologic Research* **35**:219–232. DOI: <https://doi.org/10.1385/IR:35:3:219>, PMID: 17172648
- Messi M, Giacchetto I, Nagata K, Lanzavecchia A, Natoli G, Sallusto F. 2003. Memory and flexibility of cytokine gene expression as separable properties of human Th1 and Th2 lymphocytes. *Nature Immunology* **4**:78–86. DOI: <https://doi.org/10.1038/ni872>, PMID: 12447360
- Misiak A, Leuzzi R, Allen AC, Galletti B, Baudner BC, D’Oro U, O’Hagan DT, Pizza M, Seubert A, Mills KHG. 2017. Addition of a TLR7 agonist to an acellular pertussis vaccine enhances Th1 and Th17 responses and protective immunity in a mouse model. *Vaccine* **35**:5256–5263. DOI: <https://doi.org/10.1016/j.vaccine.2017.08.009>, PMID: 28823618
- Miyao T, Floess S, Setoguchi R, Luche H, Fehling HJ, Waldmann H, Huehn J, Hori S. 2012. Plasticity of Foxp3⁺ T cells reflects promiscuous Foxp3 expression in conventional T cells but not reprogramming of regulatory T cells. *Immunity* **36**:262–275. DOI: <https://doi.org/10.1016/j.immuni.2011.12.012>, PMID: 22326580
- Miyara M, Amoura Z, Parizot C, Badoual C, Dorgham K, Trad S, Nochy D, Debré P, Piette JC, Gorochov G. 2005. Global natural regulatory T cell depletion in active systemic lupus erythematosus. *Journal of Immunology* **175**:8392–8400. DOI: <https://doi.org/10.4049/jimmunol.175.12.8392>, PMID: 16339581
- Miyazaki Y, Nakayamada S, Kubo S, Nakano K, Iwata S, Miyagawa I, Ma X, Trimova G, Sakata K, Tanaka Y. 2018. Th22 cells promote osteoclast differentiation via production of IL-22 in rheumatoid arthritis. *Frontiers in Immunology* **9**:2901. DOI: <https://doi.org/10.3389/fimmu.2018.02901>, PMID: 30619268
- Miyazawa S, Jernigan RL. 1996. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *Journal of Molecular Biology* **256**:623–644. DOI: <https://doi.org/10.1006/jmbi.1996.0114>, PMID: 8604144

- Morita R**, Schmitt N, Bentebibel SE, Ranganathan R, Bourdery L, Zurawski G, Foucat E, Dullaers M, Oh S, Sabzghabaei N, Lavecchio EM, Punaro M, Pascual V, Banchereau J, Ueno H. 2011. Human blood CXCR5⁺CD4⁺ T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion. *Immunity* **34**:108–121. DOI: <https://doi.org/10.1016/j.immuni.2010.12.012>, PMID: 21215658
- Mosmann TR**, Coffman RL. 1989. Th1 and Th2 cells: different patterns of lymphokine secretion lead to different functional properties. *Annual Review of Immunology* **7**:145–173. DOI: <https://doi.org/10.1146/annurev.iy.07.040189.001045>, PMID: 2523712
- Muranski P**, Restifo NP. 2013. Essentials of Th17 cell commitment and plasticity. *Blood* **121**:2402–2414. DOI: <https://doi.org/10.1182/blood-2012-09-378653>, PMID: 23325835
- Murphy E**, Shibuya K, Hosken N, Openshaw P, Maino V, Davis K, Murphy K, O'Garra A. 1996. Reversibility of T helper 1 and 2 populations is lost after long-term stimulation. *Journal of Experimental Medicine* **183**:901–913. DOI: <https://doi.org/10.1084/jem.183.3.901>, PMID: 8642294
- Murphy KM**, Stockinger B. 2010. Effector T cell plasticity: flexibility in the face of changing circumstances. *Nature Immunology* **11**:674–680. DOI: <https://doi.org/10.1038/ni.1899>, PMID: 20644573
- Ohkura N**, Hamaguchi M, Morikawa H, Sugimura K, Tanaka A, Ito Y, Osaki M, Tanaka Y, Yamashita R, Nakano N, Huehn J, Fehling HJ, Sparwasser T, Nakai K, Sakaguchi S. 2012. T cell receptor stimulation-induced epigenetic changes and Foxp3 expression are independent and complementary events required for Treg cell development. *Immunity* **37**:785–799. DOI: <https://doi.org/10.1016/j.immuni.2012.09.010>, PMID: 23123060
- Osnes LT**, Nakken B, Bodolay E, Szodoray P. 2013. Assessment of intracellular cytokines and regulatory cells in patients with autoimmune diseases and primary immunodeficiencies — novel tool for diagnostics and patient follow-up. *Autoimmunity Reviews* **12**:967–971. DOI: <https://doi.org/10.1016/j.autrev.2013.02.003>, PMID: 23541481
- Pacholczyk R**, Kern J. 2008. The T-cell receptor repertoire of regulatory T cells. *Immunology* **125**:450–458. DOI: <https://doi.org/10.1111/j.1365-2567.2008.02992.x>, PMID: 19128356
- Panzer M**, Sitte S, Wirth S, Drexler I, Sparwasser T, Voehringer D. 2012. Rapid in vivo conversion of effector T cells into Th2 cells during helminth infection. *Journal of Immunology* **188**:615–623. DOI: <https://doi.org/10.4049/jimmunol.1101164>, PMID: 22156341
- Plank MW**, Kaiko GE, Maltby S, Weaver J, Tay HL, Shen W, Wilson MS, Durum SK, Foster PS. 2017. Th22 cells form a distinct Th lineage from Th17 cells in vitro with unique transcriptional properties and Tbet-dependent Th1 plasticity. *Journal of Immunology* **198**:2182–2190. DOI: <https://doi.org/10.4049/jimmunol.1601480>, PMID: 28100680
- Protti MP**, De Monte L, Di Lullo G. 2014. Tumor antigen-specific CD4⁺ T cells in Cancer immunity: from antigen identification to tumor prognosis and development of the therapeutic strategies. *Tissue Antigens* **83**:237–246. DOI: <https://doi.org/10.1111/tan.12329>, PMID: 24641502
- Puniya BL**, Todd RG, Mohammed A, Brown DM, Barberis M, Helikar T. 2018. A mechanistic computational model reveals that plasticity of CD4⁺ T cell differentiation is a function of cytokine composition and dosage. *Frontiers in Physiology* **9**:878. DOI: <https://doi.org/10.3389/fphys.2018.00878>, PMID: 30116195
- Rivino L**, Messi M, Jarrossay D, Lanzavecchia A, Sallusto F, Geginat J. 2004. Chemokine receptor expression identifies pre-T helper (Th)1, pre-Th2, and nonpolarized cells among human CD4⁺ central memory T cells. *Journal of Experimental Medicine* **200**:725–735. DOI: <https://doi.org/10.1084/jem.20040774>, PMID: 15381728
- Rolla S**, Bardina V, De Mercanti S, Quaglino P, De Palma R, Gned D, Brusa D, Durelli L, Novelli F, Clerico M. 2014. Th22 cells are expanded in multiple sclerosis and are resistant to IFN- β . *Journal of Leukocyte Biology* **96**:1155–1164. DOI: <https://doi.org/10.1189/jlb.5A0813-463RR>, PMID: 25097195
- Ryba-Stanisławowska M**, Werner P, Brandt A, Myśliwiec M, Myśliwska J. 2016. Th9 and Th22 immune response in young patients with type 1 diabetes. *Immunologic Research* **64**:730–735. DOI: <https://doi.org/10.1007/s12026-015-8765-7>, PMID: 26659093
- Sad S**, Mosmann TR. 1994. Single IL-2-secreting precursor CD4 T cell can develop into either Th1 or Th2 cytokine secretion phenotype. *Journal of Immunology* **153**:3514–3522. PMID: 7930573
- Sallusto F**. 2016. Heterogeneity of human CD4⁺ T cells against microbes. *Annual Review of Immunology* **34**:317–334. DOI: <https://doi.org/10.1146/annurev-immunol-032414-112056>, PMID: 27168241
- Sawant DV**, Vignali DA. 2014. Once a Treg, always a Treg? *Immunological Reviews* **259**:173–191. DOI: <https://doi.org/10.1111/imr.12173>, PMID: 24712466
- Schlapbach C**, Gehad A, Yang C, Watanabe R, Guenova E, Teague JE, Campbell L, Yawalkar N, Kupper TS, Clark RA. 2014. Human TH9 cells are skin-tropic and have autocrine and paracrine proinflammatory capacity. *Science Translational Medicine* **6**:219ra8. DOI: <https://doi.org/10.1126/scitranslmed.3007828>, PMID: 24431112
- Shugay M**, Britanova OV, Merzlyak EM, Turchaninova MA, Mamedov IZ, Tuganbaev TR, Bolotin DA, Staroverov DB, Putintseva EV, Plevova K, Linnemann C, Shagin D, Pospisilova S, Lukyanov S, Schumacher TN, Chudakov DM. 2014. Towards error-free profiling of immune repertoires. *Nature Methods* **11**:653–655. DOI: <https://doi.org/10.1038/nmeth.2960>, PMID: 24793455
- Shugay M**, Bagaev DV, Turchaninova MA, Bolotin DA, Britanova OV, Putintseva EV, Pogorelyy MV, Nazarov VI, Zvyagin IV, Kirgizova VI, Kirgizov KI, Skorobogatova EV, Chudakov DM. 2015. VDJtools: unifying post-analysis of T cell receptor repertoires. *PLOS Computational Biology* **11**:e1004503. DOI: <https://doi.org/10.1371/journal.pcbi.1004503>, PMID: 26606115
- Shulman Z**, Gitlin AD, Targ S, Jankovic M, Pasqual G, Nussenzweig MC, Victora GD. 2013. T follicular helper cell dynamics in germinal centers. *Science* **341**:673–677. DOI: <https://doi.org/10.1126/science.1241680>, PMID: 23887872

- Sibener LV**, Fernandes RA, Kolawole EM, Carbone CB, Liu F, McAfee D, Birnbaum ME, Yang X, Su LF, Yu W, Dong S, Gee MH, Jude KM, Davis MM, Groves JT, Goddard WA, Heath JR, Evavold BD, Vale RD, Garcia KC. 2018. Isolation of a structural mechanism for uncoupling T cell receptor signaling from peptide-MHC binding. *Cell* **174**:672–687. DOI: <https://doi.org/10.1016/j.cell.2018.06.017>, PMID: 30053426
- Silva SL**, Albuquerque AS, Serra-Caetano A, Foxall RB, Pires AR, Matoso P, Fernandes SM, Ferreira J, Cheyner R, Victorino RM, Caramalho I, Barata JT, Sousa AE. 2016. Human naïve regulatory T-cells feature high steady-state turnover and are maintained by IL-7. *Oncotarget* **7**:12163–12175. DOI: <https://doi.org/10.18632/oncotarget.7512>, PMID: 26910841
- Spence A**, Purtha W, Tam J, Dong S, Kim Y, Ju CH, Sterling T, Nakayama M, Robinson WH, Bluestone JA, Anderson MS, Tang Q. 2018. Revealing the specificity of regulatory T cells in murine autoimmune diabetes. *PNAS* **115**:5265–5270. DOI: <https://doi.org/10.1073/pnas.1715590115>, PMID: 29712852
- Stadinski BD**, Shekhar K, Gómez-Touriño I, Jung J, Sasaki K, Sewell AK, Peakman M, Chakraborty AK, Huseby ES. 2016. Hydrophobic CDR3 residues promote the development of self-reactive T cells. *Nature Immunology* **17**:946–955. DOI: <https://doi.org/10.1038/ni.3491>, PMID: 27348411
- Su LF**, Del Alcazar D, Stelekati E, Wherry EJ, Davis MM. 2016. Antigen exposure shapes the ratio between antigen-specific Tregs and conventional T cells in human peripheral blood. *PNAS* **113**:E6192–E6198. DOI: <https://doi.org/10.1073/pnas.1611723113>, PMID: 27681619
- Tang Q**, Adams JY, Penaranda C, Melli K, Piaggio E, Sgouroudis E, Piccirillo CA, Salomon BL, Bluestone JA. 2008. Central role of defective interleukin-2 production in the triggering of islet autoimmune destruction. *Immunity* **28**:687–697. DOI: <https://doi.org/10.1016/j.immuni.2008.03.016>, PMID: 18468463
- Thiault N**, Darrigues J, Adoue V, Gros M, Binet B, Peralis C, Leobon B, Fazilleau N, Joffre OP, Robey EA, van Meerwijk JP, Romagnoli P. 2015. Peripheral regulatory T lymphocytes recirculating to the thymus suppress the development of their precursors. *Nature Immunology* **16**:628–634. DOI: <https://doi.org/10.1038/ni.3150>, PMID: 25939024
- Tulic MK**, Andrews D, Crook ML, Charles A, Tourigny MR, Moqbel R, Prescott SL. 2012. Changes in thymic regulatory T-cell maturation from birth to puberty: differences in atopic children. *Journal of Allergy and Clinical Immunology* **129**:199–206. DOI: <https://doi.org/10.1016/j.jaci.2011.10.016>, PMID: 22104606
- Vella LA**, Buggert M, Manne S, Herati RS, Sayin I, Kuri-Cervantes L, Bukh Brody I, O’Boyle KC, Kaprielian H, Giles JR, Nguyen S, Muselman A, Antel JP, Bar-Or A, Johnson ME, Canaday DH, Najj A, Ganusov VV, Laufer TM, Wells AD, et al. 2019. T follicular helper cells in human efferent lymph retain lymphoid characteristics. *Journal of Clinical Investigation* **129**:3185–3200. DOI: <https://doi.org/10.1172/JCI125628>, PMID: 31264971
- Vinuesa CG**, Linterman MA, Yu D, MacLennan IC. 2016. Follicular helper T cells. *Annual Review of Immunology* **34**:335–368. DOI: <https://doi.org/10.1146/annurev-immunol-041015-055605>, PMID: 26907215
- Voo KS**, Wang YH, Santori FR, Boggiano C, Wang YH, Arima K, Bover L, Hanabuchi S, Khalili J, Marinova E, Zheng B, Littman DR, Liu YJ. 2009. Identification of IL-17-producing FOXP3⁺ regulatory T cells in humans. *PNAS* **106**:4793–4798. DOI: <https://doi.org/10.1073/pnas.0900408106>, PMID: 19273860
- Vroman H**, van den Blink B, Kool M. 2015. Mode of dendritic cell activation: the decisive hand in Th2/Th17 cell differentiation. Implications in asthma severity? *Immunobiology* **220**:254–261. DOI: <https://doi.org/10.1016/j.imbio.2014.09.016>, PMID: 25245013
- Waickman AT**, Ligons DL, Hwang S, Park JY, Lazarevic V, Sato N, Hong C, Park JH. 2017. CD4 effector T cell differentiation is controlled by IL-15 that is expressed and presented in trans. *Cytokine* **99**:266–274. DOI: <https://doi.org/10.1016/j.cyto.2017.08.004>, PMID: 28807496
- Walker LS**, von Herrath M. 2016. CD4 T cell differentiation in type 1 diabetes. *Clinical & Experimental Immunology* **183**:16–29. DOI: <https://doi.org/10.1111/cei.12672>, PMID: 26102289
- Wang JH**, Reinherz EL. 2012. The structural basis of $\alpha\beta$ T-lineage immune recognition: TCR docking topologies, mechanotransduction, and co-receptor function. *Immunological Reviews* **250**:102–119. DOI: <https://doi.org/10.1111/j.1600-065X.2012.01161.x>, PMID: 23046125
- Wei SC**, Levine JH, Cogdill AP, Zhao Y, Anang NAS, Andrews MC, Sharma P, Wang J, Wargo JA, Pe’er D, Allison JP. 2017. Distinct cellular mechanisms underlie anti-CTLA-4 and anti-PD-1 checkpoint blockade. *Cell* **170**:1120–1133. DOI: <https://doi.org/10.1016/j.cell.2017.07.024>, PMID: 28803728
- Yang XO**, Nurieva R, Martinez GJ, Kang HS, Chung Y, Pappu BP, Shah B, Chang SH, Schluns KS, Watowich SS, Feng XH, Jetten AM, Dong C. 2008. Molecular antagonism and plasticity of regulatory and inflammatory T cell programs. *Immunity* **29**:44–56. DOI: <https://doi.org/10.1016/j.immuni.2008.05.007>, PMID: 18585065
- Yang BH**, Wang K, Wan S, Liang Y, Yuan X, Dong Y, Cho S, Xu W, Jepsen K, Feng GS, Lu LF, Xue HH, Fu W. 2019. TCF1 and LEF1 control Treg competitive survival and Tfr development to prevent autoimmune diseases. *Cell Reports* **27**:3629–3645. DOI: <https://doi.org/10.1016/j.celrep.2019.05.061>, PMID: 31216480
- Zhou X**, Bailey-Bucktrout S, Jeker LT, Bluestone JA. 2009a. Plasticity of CD4⁺ FoxP3⁺ T cells. *Current Opinion in Immunology* **21**:281–285. DOI: <https://doi.org/10.1016/j.coi.2009.05.007>, PMID: 19500966
- Zhou L**, Chong MM, Littman DR. 2009b. Plasticity of CD4⁺ T cell lineage differentiation. *Immunity* **30**:646–655. DOI: <https://doi.org/10.1016/j.immuni.2009.05.001>, PMID: 19464987
- Zhu J**, Yamane H, Paul WE. 2010. Differentiation of effector CD4 T cell populations. *Annual Review of Immunology* **28**:445–489. DOI: <https://doi.org/10.1146/annurev-immunol-030409-101212>, PMID: 20192806
- Zielinski CE**, Mele F, Aschenbrenner D, Jarrossay D, Ronchi F, Gattorno M, Monticelli S, Lanzavecchia A, Sallusto F. 2012. Pathogen-induced human TH17 cells produce IFN γ or IL-10 and are regulated by IL-1 β . *Nature* **484**:514–518. DOI: <https://doi.org/10.1038/nature10957>, PMID: 22466287



Figures and figure supplements

Functionally specialized human CD4⁺ T-cell subsets express physicochemically distinct TCRs

Sofya A Kasatskaya et al

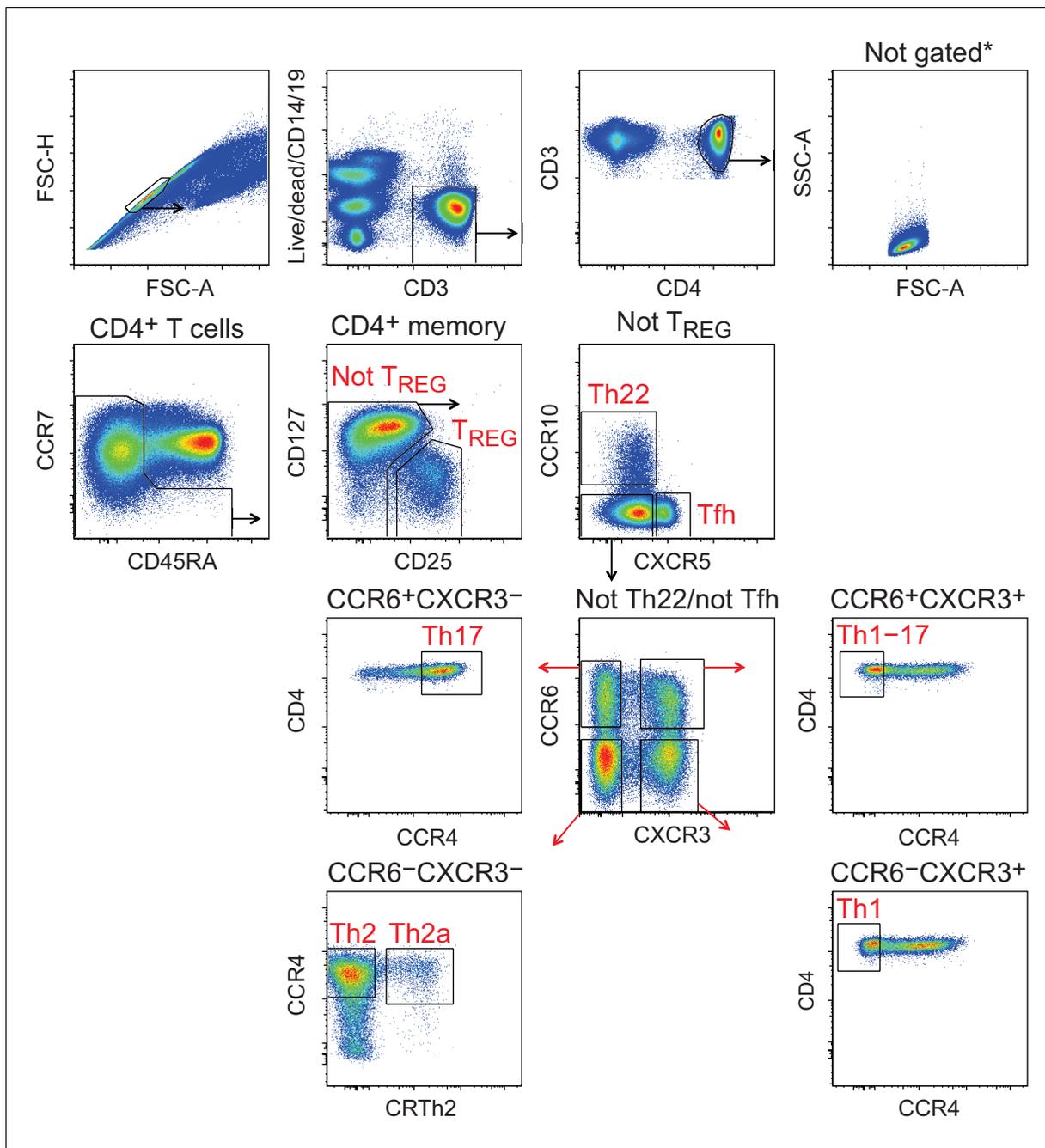


Figure 1—figure supplement 1. Gating strategy for the identification of effector/memory CD4⁺ T-cell subsets. Single lymphocytes were identified in a forward scatter-area (FCS-A) versus forward scatter-height (FSC-H) plot. Viable CD3⁺CD14⁻CD19⁻ cells were gated in the CD4⁺ lineage, and naive cells were excluded as CCR7⁺CD45RA⁺ events. Effector/memory subsets were then sorted as Tfh cells (CXCR5⁺), Th1 cells (non-Tfh/Th22/Treg CCR4⁻CCR6⁻CXCR3⁺), Th1-17 cells (non-Tfh/Th22/Treg CCR4⁻CCR6⁺CXCR3⁺), Th17 cells (non-Tfh/Th22/Treg CCR4⁺CCR6⁺CXCR3⁻), Th22 cells (CCR10⁺), Th2a cells (non-Tfh/Th22/Treg CCR4⁺CCR6⁻CRTh2⁺CXCR3⁻), Th2 cells (non-Tfh/Th22/Treg CCR4⁺CCR6⁻CRTh2⁻CXCR3⁻), or Tregs (CD25^{high}CD127^{low}).

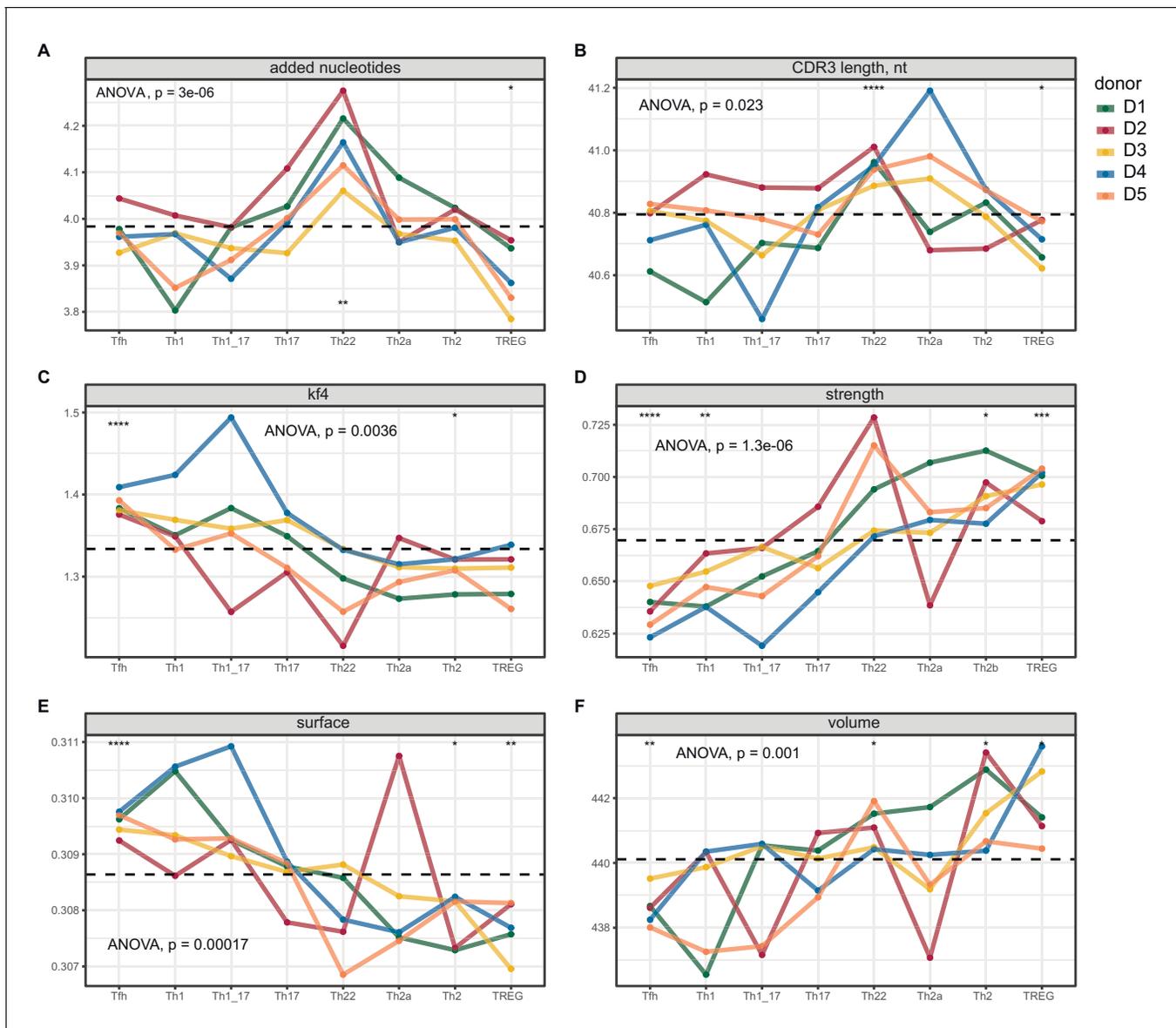


Figure 2—figure supplement 1. Averaged physicochemical characteristics of CDR3 α repertoires from effector/memory CD4⁺ T-cell subsets. (A–F) Averaged physicochemical characteristics were measured for the five amino acids in the middle of the CDR3 α sequences obtained from each effector/memory CD4⁺ T-cell subset ($n = 8$) from each healthy donor ($n = 5$). Calculations were weighted by clonotype frequency. (A) Non-germline nucleotide (N) additions. (B) CDR3 α length (nucleotides). (C) Kidera factor 4 (arbitrary scale). (D) Interaction strength (arbitrary scale). (E) Surface (arbitrary scale). (F) Volume (arbitrary scale). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$ (one-way ANOVA followed by the two-sample Welch t-test with Bonferroni correction for each group versus the mean).

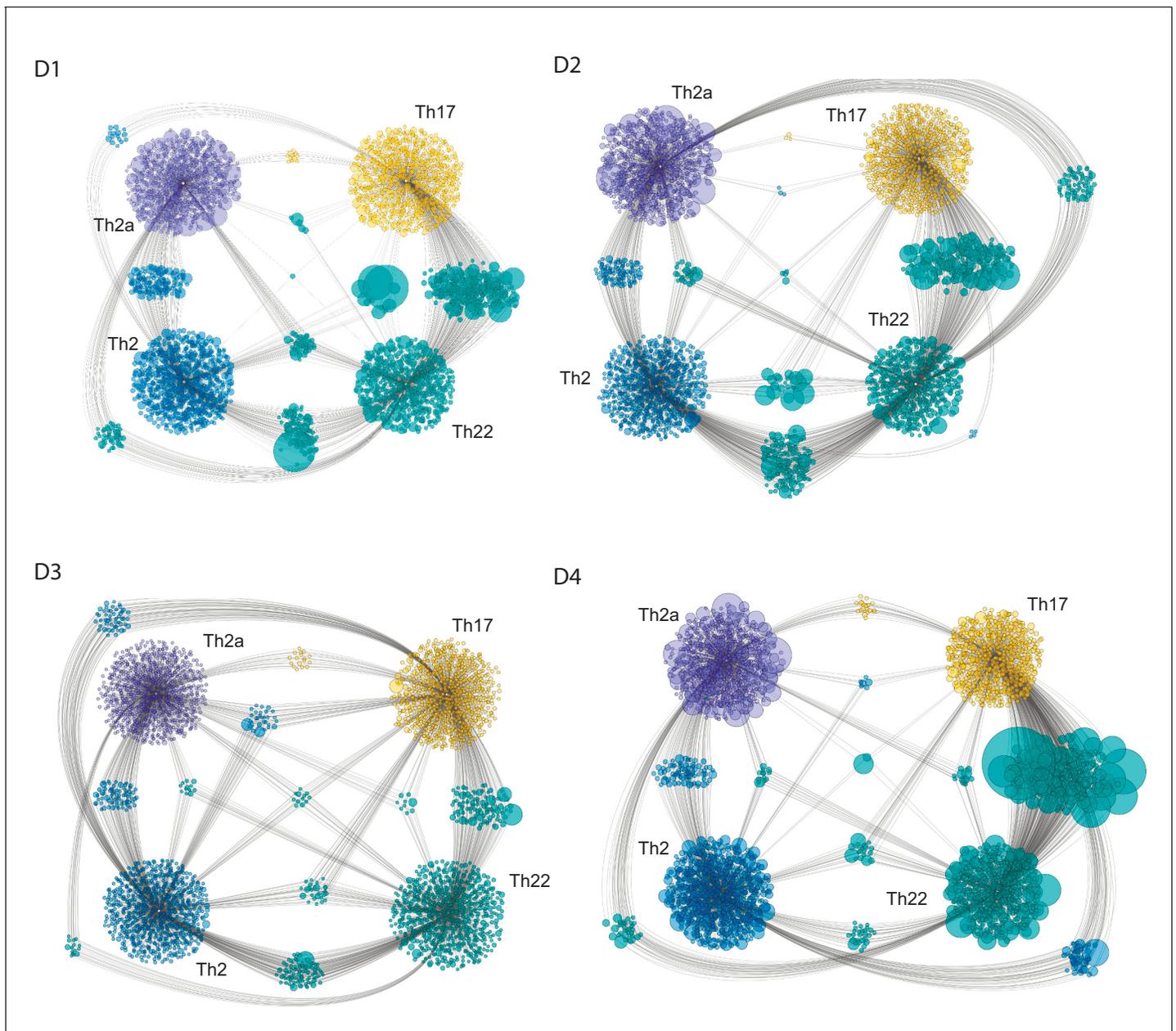


Figure 5—figure supplement 1. Clonal relatedness among the Th17, Th22, Th2a, and Th2 subsets of effector/memory CD4⁺ T-cells. Cytoscape plots for each donor represent the number and size (frequency) of nucleotide-defined clonotype variants shared among the top 2000 most frequent CDR3 β clonotypes in each subset. Details as in **Figure 5**.

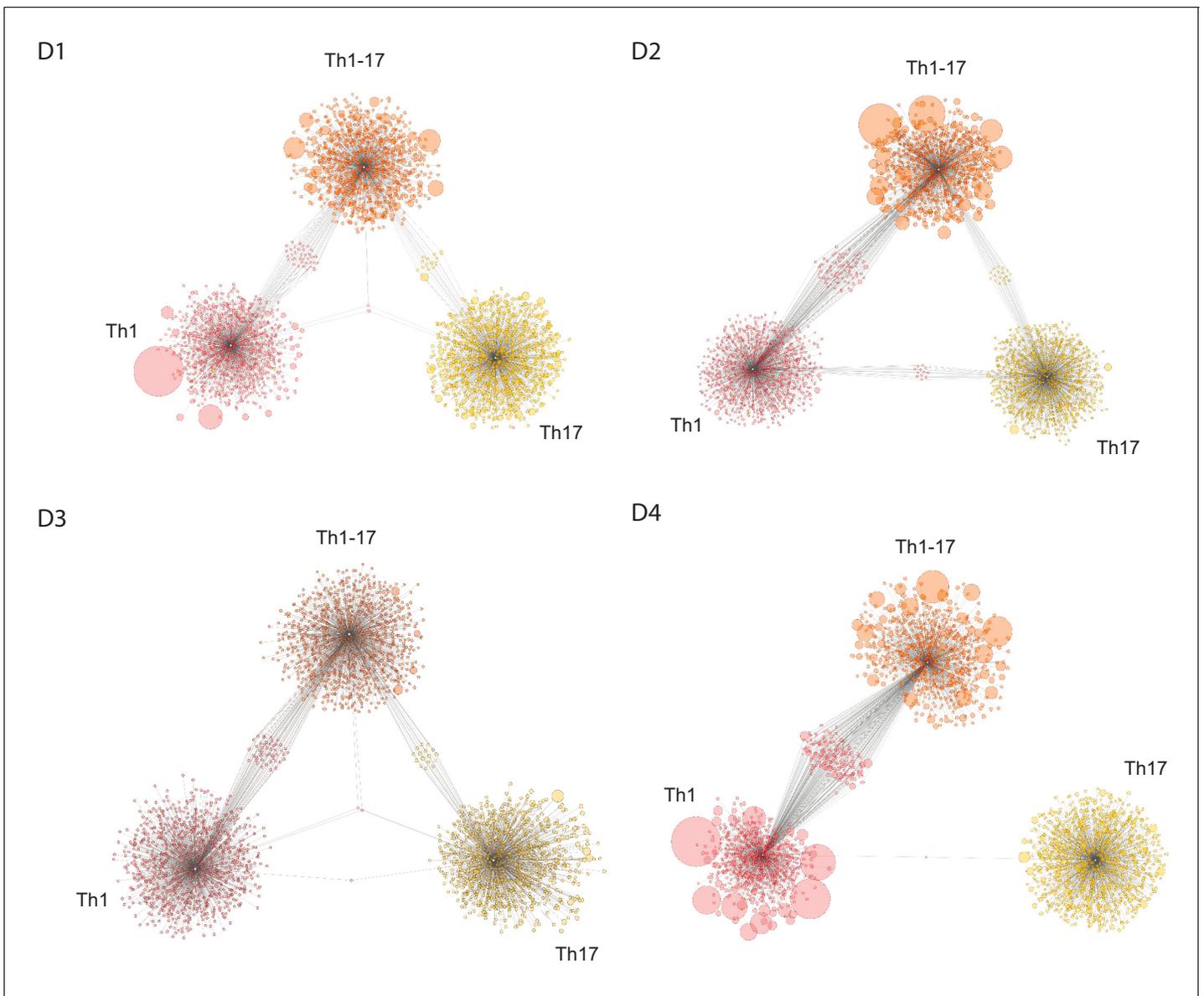


Figure 5—figure supplement 2. Clonal relatedness among the Th1, Th1-17, and Th17 subsets of effector/memory CD4⁺ T-cells. Cytoscape plots for each donor represent the number and size (frequency) of nucleotide-defined clonotype variants shared among the top 2000 most frequent CDR3 β clonotypes in each subset. Details as in **Figure 5**.

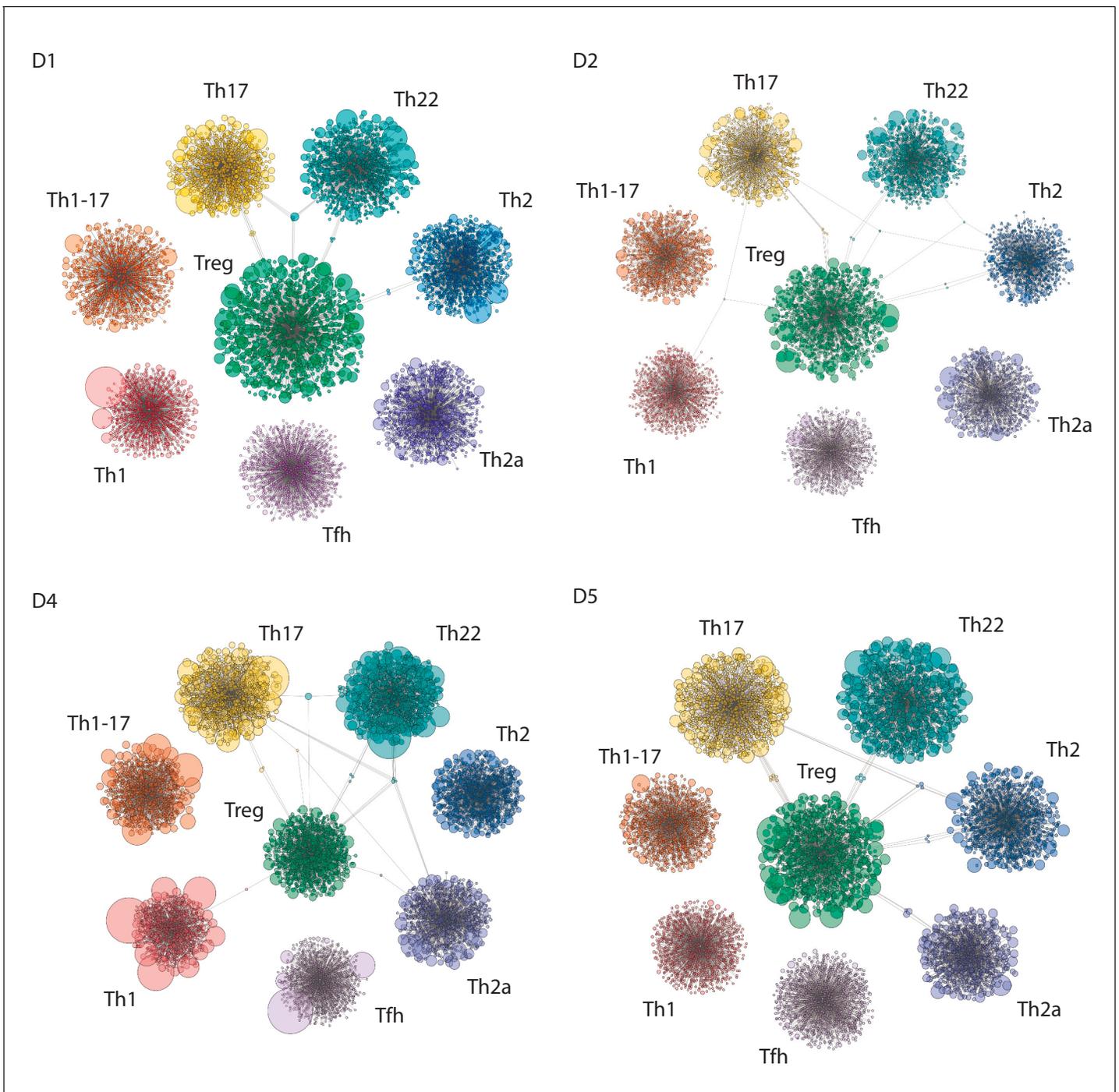


Figure 5—figure supplement 3. Clonal relatedness among Tregs and other subsets of effector/memory CD4⁺ T-cells. Cytoscape plots for each donor represent the number and size (frequency) of nucleotide-defined clonotype variants shared among the top 2000 most frequent CDR3 β clonotypes in each subset. Only clonotypes shared with Tregs are shown. Details as in **Figure 5**.

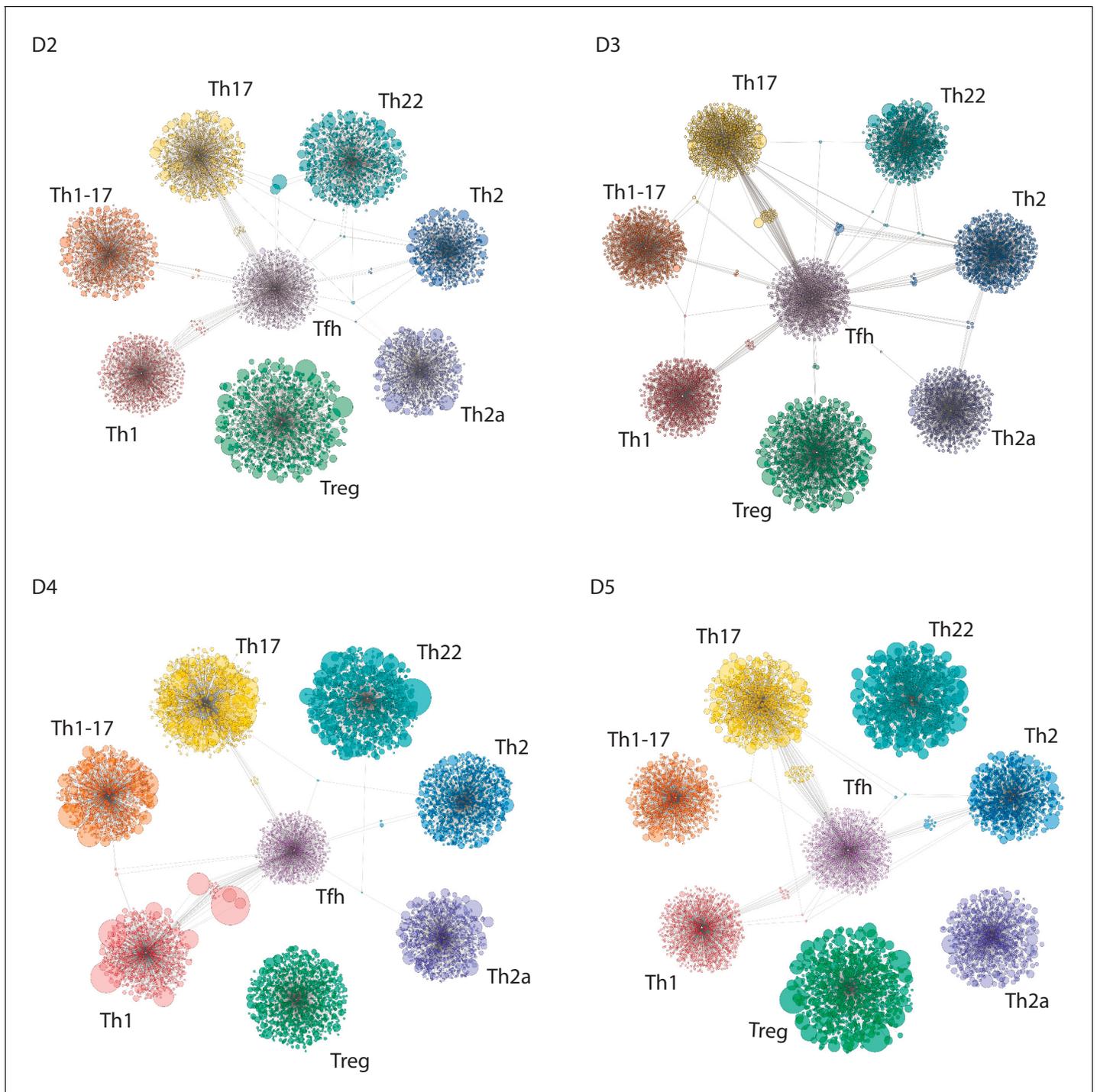


Figure 5—figure supplement 4. Clonal relatedness among Tfh cells other subsets of effector/memory CD4⁺ T-cells. Cytoscape plots for each donor represent the number and size (frequency) of nucleotide-defined clonotype variants shared among the top 2000 most frequent CDR3 β clonotypes in each subset. Only clonotypes shared with Tfh cells are shown. Details as in **Figure 5**.

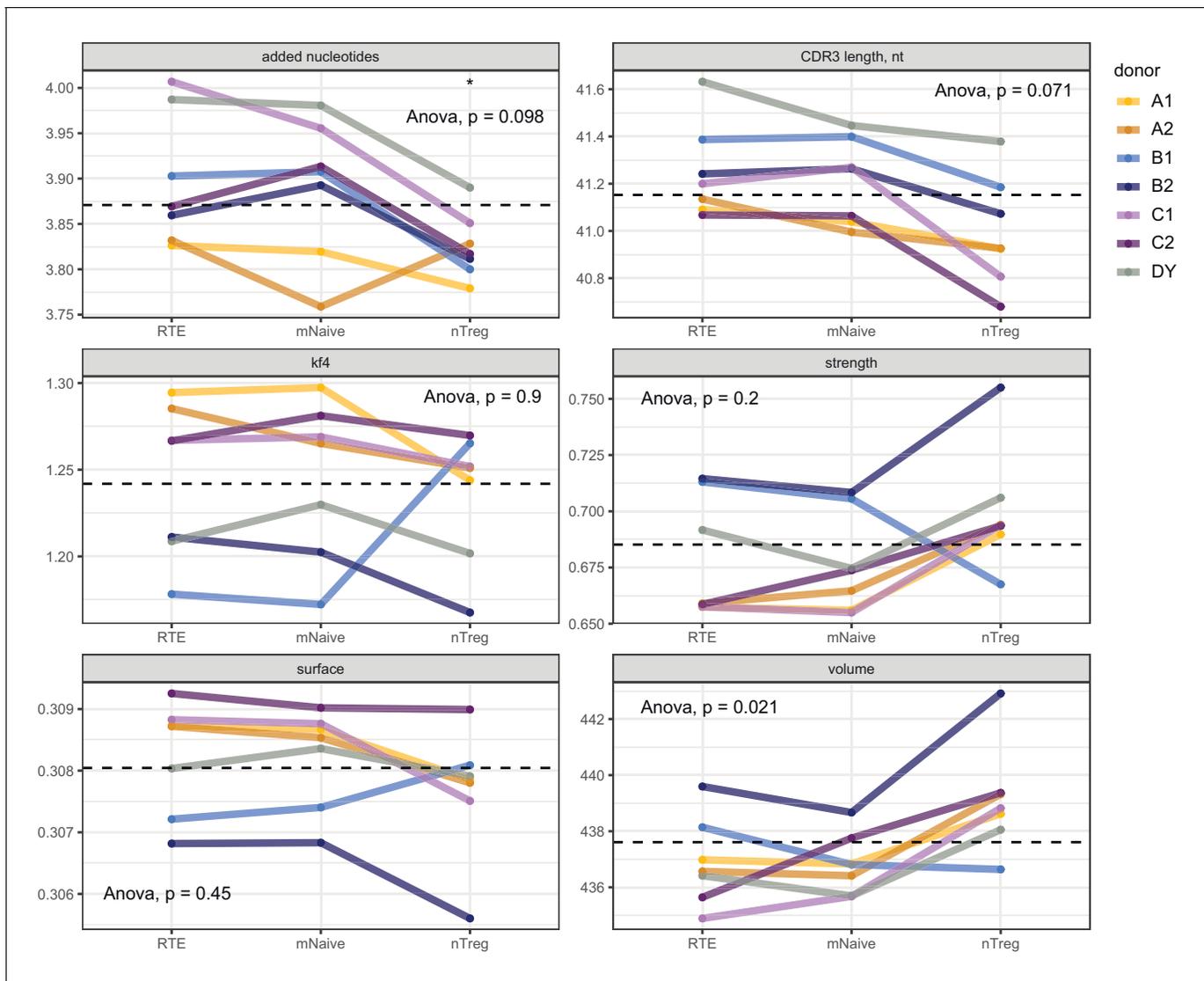


Figure 6—figure supplement 1. Averaged physicochemical characteristics of CDR3 α repertoires from naive CD4⁺ T-cell subsets. Repertoire metrics are shown for RTEs (CD25⁻CD31⁺), mature naive T cells (mNaive; CD25⁻CD31⁻), and naive Tregs (nTreg; CD25^{high}) from healthy donors (n = 7). Matched letters in the key indicate twin pairs. Averaged physicochemical characteristics were measured for the five amino acids in the middle of the CDR3 α sequences obtained from each naive CD4⁺ T-cell subset. Calculations were weighted by clonotype frequency. Parameter details as in **Figure 2**. Dashed lines indicate means. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$ (one-way ANOVA followed by the two-sample Welch t-test with Bonferroni correction for each group versus the mean).

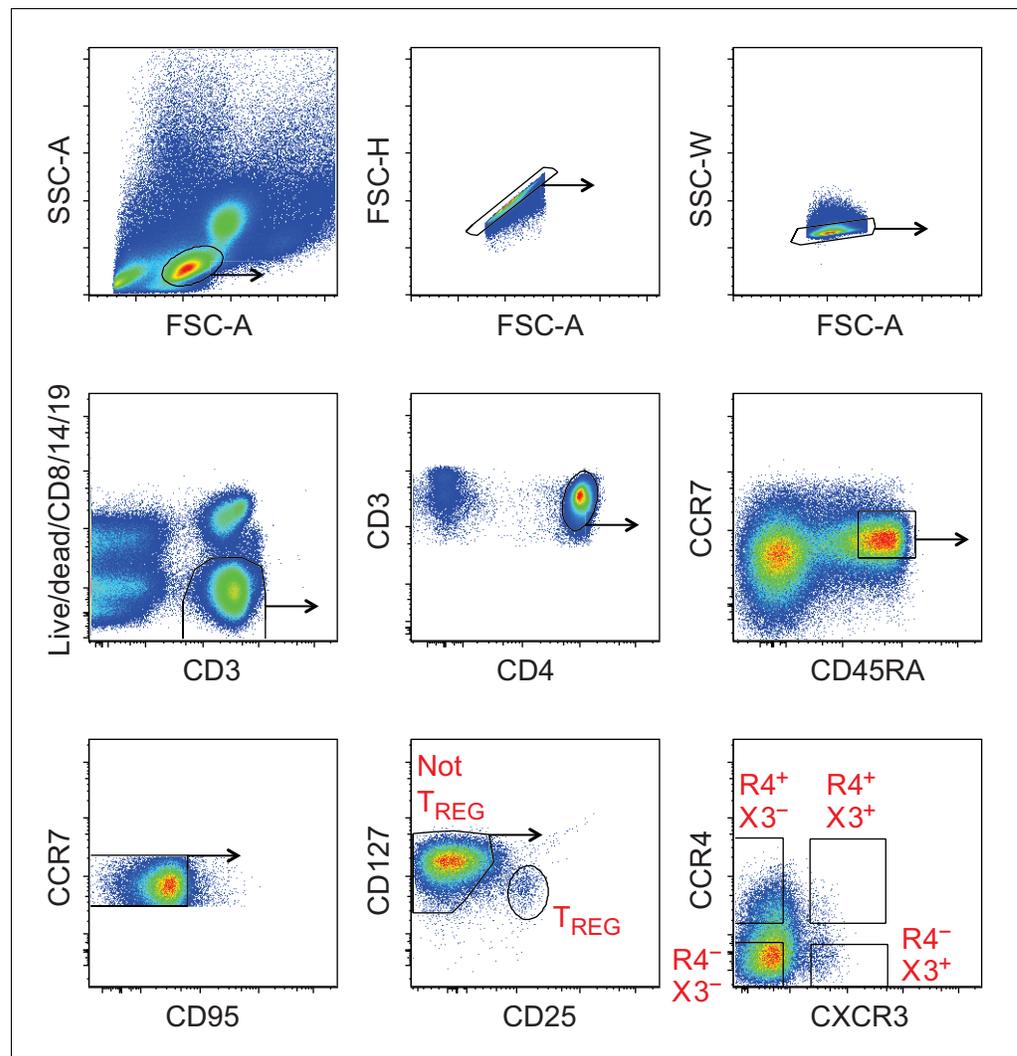


Figure 6—figure supplement 2. Gating strategy for the identification of naive CD4⁺ T-cell subsets. Lymphocytes were identified in a forward scatter-area (FSC-A) versus side scatter-area (SSC-A) plot, and single cells were identified in FSC-A versus forward scatter-height (FSC-H) and FSC-A versus side scatter-width (SSC-W) plots. Naive cells were gated as viable CD3⁺CD4⁺CD8⁻CD14⁻CD19⁻CCR7⁺CD45RA⁺CD95⁻ events, and subsets were sorted as naive Th1-like cells (non-Treg CCR4⁻CXCR3⁺), naive Th2-like cells (non-Treg CCR4⁺CXCR3⁻), or naive Tregs (CD25^{high}CD127^{low}), alongside the corresponding non-Treg CCR4⁻CXCR3⁻ and non-Treg CCR4⁺CXCR3⁺ populations.

Table S1. Gating strategy for the identification of effector/memory CD4⁺ T cell subsets.

Gates 1 & 2	Gate 3	Gate 4	Gate 5	Gate 6	Gate 7	Gate 8	Subset
Live single CD3 ⁺ CD14 ⁻ CD19 ⁻ lymphocytes	CD4 ⁺	Exclude CCR7 ⁺ CD45RA ⁺	CD25 ^{high} CD127 ^{low}				Treg
			CD25 ^{low} CD127 ⁺	CXCR5 ⁺			Tfh
				CCR10 ⁺			Th22
				CXCR5 ⁻ CCR10 ⁻	CXCR3 ⁺ CCR6 ⁻	CCR4 ⁻	Th1
					CXCR3 ⁻ CCR6 ⁺	CCR4 ⁺	Th17
			CXCR3 ⁺ CCR6 ⁺		CCR4 ⁻	Th1-17	
			CXCR3 ⁻ CCR6 ⁻		CCR4 ⁺ CRTh2 ⁻	Th2	
				CCR4 ⁺ CRTh2 ⁺	Th2a		

See also **Figure 1-figure supplement 1**.

Table S2. Frequencies of sorted effector/memory CD4⁺ T cell subsets.

Donor	Tfh	Th1	Th1-17	Th17	Th22	Th2a	Th2	Treg
D1	5.44	1.91	1.44	3.06	2.60	1.04	4.86	3.99
D2	5.82	3.29	3.50	3.14	6.64	1.53	9.24	6.92
D3	2.05	0.19	0.31	1.31	0.81	0.26	1.93	1.84
D4	6.70	2.33	2.11	4.22	2.02	0.57	7.19	3.95
D5	4.39	1.16	1.17	3.32	2.12	0.82	3.96	3.99
Mean	4.88	1.78	1.71	3.01	2.84	0.84	5.44	4.14
SD	1.79	1.17	1.19	1.06	2.23	0.48	2.84	1.81

Shown as % of live CD3⁺CD4⁺CD14⁻CD19⁻ non-naive cells. Details in **Figure 1-figure supplement 1**.

Table S3. Frequencies of sorted naive CD4⁺ T cell subsets.

Donor	Th1-like CCR4 ⁻ CXCR3 ⁺	Th2-like CCR4 ⁺ CXCR3 ⁻	CCR4 ⁻ CXCR3 ⁻	CCR4 ⁺ CXCR3 ⁺	Treg CD25 ^{high} CD127 ^{low}
D1	1.75	5.46	44.80	0.23	0.73
D2	0.77	6.77	20.40	0.32	0.57
D3	0.15	5.67	42.60	0.19	1.70
D4	0.16	6.33	33.10	0.05	1.11

Shown as % of live CD3⁺CD4⁺CD8⁻CD14⁻CD19⁻ naive cells. Details in **Figure 6-figure supplement 2**.

SUPPLEMENTARY FIGURE LEGENDS

Figure 1-figure supplement 1. Gating strategy for the identification of effector/memory CD4⁺ T cell subsets. Single lymphocytes were identified in a forward scatter-area (FCS-A) *versus* forward scatter-height (FSC-H) plot. Viable CD3⁺CD14⁻CD19⁻ cells were gated in the CD4⁺ lineage, and naive cells were excluded as CCR7⁺CD45RA⁺ events. Effector/memory subsets were then sorted as Tfh cells (CXCR5⁺), Th1 cells (non-Tfh/Th22/Treg CCR4⁻CCR6⁻CXCR3⁺), Th1-17 cells (non-Tfh/Th22/Treg CCR4⁻CCR6⁺CXCR3⁺), Th17 cells (non-Tfh/Th22/Treg CCR4⁺CCR6⁺CXCR3⁻), Th22 cells (CCR10⁺), Th2a cells (non-Tfh/Th22/Treg CCR4⁺CCR6⁻CRTh2⁺CXCR3⁻), Th2 cells (non-Tfh/Th22/Treg CCR4⁺CCR6⁻CRTh2⁻CXCR3⁻), or Tregs (CD25^{high}CD127^{low}).

Figure 2-figure supplement 1. Averaged physicochemical characteristics of CDR3 α repertoires from effector/memory CD4⁺ T cell subsets. (A–F) Averaged physicochemical characteristics were measured for the five amino acids in the middle of the CDR3 α sequences obtained from each effector/memory CD4⁺ T cell subset (n = 8) from each healthy donor (n = 5). Calculations were weighted by clonotype frequency. (A) Non-germline nucleotide (N) additions. (B) CDR3 α length (nucleotides). (C) Kidera factor 4 (arbitrary scale). (D) Interaction strength (arbitrary scale). (E) Surface (arbitrary scale). (F) Volume (arbitrary scale). *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001 (one-way ANOVA followed by the two-sample Welch t-test with Bonferroni correction for each group *versus* the mean).

Figure 5-figure supplement 1. Individual donor plasticity among the Th17, Th22, Th2a, and Th2 subsets of effector/memory CD4⁺ T cells. Cytoscape plots for each donor represent the number and size (frequency) of nucleotide-defined clonotype variants shared among the top 2,000 most frequent CDR3 β clonotypes in each subset. Details as in **Figure 5**.

Figure 5-figure supplement 2. Individual donor plasticity among the Th1, Th1-17, and Th17 subsets of effector/memory CD4⁺ T cells. Cytoscape plots for each donor represent the number and size (frequency) of nucleotide-defined clonotype variants shared among the top 2,000 most frequent CDR3 β clonotypes in each subset. Details as in **Figure 5**.

Figure 5-figure supplement 3. Individual donor plasticity among Tregs and other subsets of effector/memory CD4⁺ T cells. Cytoscape plots for each donor represent the number and size (frequency) of nucleotide-defined clonotype variants shared among the top 2,000 most frequent CDR3 β clonotypes in each subset. Only clonotypes shared with Tregs are shown.

Figure 5-figure supplement 4. Individual donor plasticity among Tfh cells other subsets of effector/memory CD4⁺ T cells. Cytoscape plots for each donor represent the number and size (frequency) of nucleotide-defined clonotype variants shared among the top 2,000 most frequent CDR3 β clonotypes in each subset. Only clonotypes shared with Tfh cells are shown. Details as in **Figure 5**.

Figure 6-figure supplement 1. Averaged physicochemical characteristics of CDR3 α repertoires from naive CD4⁺ T cell subsets. Repertoire metrics are shown for RTEs (CD25⁻CD31⁺), mature naive T cells (mNaive; CD25⁻CD31⁻), and naive Tregs (nTreg; CD25^{high}) from healthy donors (n = 7). Matched letters in the key indicate twin pairs. Averaged physicochemical characteristics were measured for the five amino acids in the middle of the CDR3 α sequences obtained from each naive CD4⁺ T cell subset. Calculations were weighted by clonotype frequency. Parameter details as in **Figure 2**. Dashed lines indicate means. *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001 (one-way ANOVA followed by the two-sample Welch t-test with Bonferroni correction for each group *versus* the mean).

Figure 6-figure supplement 2. Gating strategy for the identification of naive CD4⁺ T cell subsets. Lymphocytes were identified in a forward scatter-area (FSC-A) *versus* side scatter-area (SSC-A) plot, and single cells were identified in FSC-A *versus* forward scatter-height (FSC-H) and FSC-A *versus* side scatter-width (SSC-W) plots. Naive cells were gated as viable CD3⁺CD4⁺CD8⁻CD14⁻CD19⁻CCR7⁺CD45RA⁺CD95⁻ events, and subsets were sorted as naive Th1-like cells (non-Treg CCR4⁻CXCR3⁺), naive Th2-like cells (non-Treg CCR4⁺CXCR3⁻), or naive Tregs (CD25^{high}CD127^{low}), alongside the corresponding non-Treg CCR4⁻CXCR3⁻ and non-Treg CCR4⁺CXCR3⁺ populations.

Chapter IV

The human $V\delta 2+$ T-cell compartment comprises distinct innate-like $V\gamma 9+$ and adaptive $V\gamma 9-$ subsets

Introduction

In this and the following chapters, we are dissecting the structure of gamma/delta TCR repertoires from unconventional T cells. Previously in the same collaboration with the University of Birmingham, we analyzed the repertoires of VDelta1+ gamma/delta T-cells of human peripheral blood. We succeeded in showing that the VDelta1+ subset undergoes the same TCR-driven selection and expansion of memory clones as conventional T cells. In this study, we focused on the Vδ2+ T-cell compartment, which is the major gamma/delta T cell population in healthy donors' peripheral blood. Gamma chains of T cell receptors revealed the difference between two smaller subsets. The TCR chains comprised of Vgamma9 display a degenerate public repertoire whether Vgamma9-VDelta2+ gamma-TCR repertoire is less public and holds a substantial diversity of unique clonotypes. Combined with the naive phenotype of Vgamma9-VDelta2+ T cells, such TCR repertoire features contribute to the pattern of adaptive immune cell development. Therefore, by means of TCR sequencing, we described novel innate-like and adaptive-like subsets in unconventional human gamma/delta T cells.

Contribution

In this collaborative project, the experimental part was performed by our colleagues from the University of Birmingham. TCR sequencing was outsourced and performed with kits from iRepertoire (Illumina, USA). The raw sequencing data was transferred and processed in the Moscow laboratory. I performed the data processing, and TCR repertoire analysis, prepared the graphics, and contributed to manuscript preparation and discussion. During this project, I created the optimized pipeline for TCR gamma and TCR delta sequencing data preprocessing and analysis with Milaboratory software. The pipeline was later used in both laboratories in collaboration.

ARTICLE

DOI: 10.1038/s41467-018-04076-0

OPEN

The human $V\delta 2^+$ T-cell compartment comprises distinct innate-like $V\gamma 9^+$ and adaptive $V\gamma 9^-$ subsets

Martin S. Davey¹, Carrie R. Willcox¹, Stuart Hunter^{1,2}, Sofya A. Kasatskaya^{3,4}, Ester B.M. Remmerswaal⁵, Mahboob Salim¹, Fiyaz Mohammed¹, Frederike J. Bemelman⁶, Dmitriy M. Chudakov^{3,4,7,8}, Ye H. Oo² & Benjamin E. Willcox¹

$V\delta 2^+$ T cells form the predominant human $\gamma\delta$ T-cell population in peripheral blood and mediate T-cell receptor (TCR)-dependent anti-microbial and anti-tumour immunity. Here we show that the $V\delta 2^+$ compartment comprises both innate-like and adaptive subsets. $V\gamma 9^+$ $V\delta 2^+$ T cells display semi-invariant TCR repertoires, featuring public $V\gamma 9$ TCR sequences equivalent in cord and adult blood. By contrast, we also identify a separate, $V\gamma 9^-$ $V\delta 2^+$ T-cell subset that typically has a $CD27^{hi}CCR7^+CD28^+IL-7R\alpha^+$ naive-like phenotype and a diverse TCR repertoire, however in response to viral infection, undergoes clonal expansion and differentiation to a $CD27^{lo}CD45RA^+CX_3CR1^+granzymeA/B^+$ effector phenotype. Consistent with a function in solid tissue immunosurveillance, we detect human intrahepatic $V\gamma 9^-$ $V\delta 2^+$ T cells featuring dominant clonal expansions and an effector phenotype. These findings redefine human $\gamma\delta$ T-cell subsets by delineating the $V\delta 2^+$ T-cell compartment into innate-like ($V\gamma 9^+$) and adaptive ($V\gamma 9^-$) subsets, which have distinct functions in microbial immunosurveillance.

¹Cancer Immunology and Immunotherapy Centre, Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham B15 2TT, UK. ²Centre for Liver Research and NIHR Biomedical Research Unit in Liver Disease, Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham B15 2TT, UK. ³Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Science, Moscow 117997, Russia. ⁴Centre for Data-Intensive Biomedicine and Biotechnology, Skolkovo Institute of Science and Technology, Moscow 143026, Russia. ⁵Department of Experimental Immunology, Academic Medical Center, Amsterdam 1105 AZ, The Netherlands. ⁶Renal Transplant Unit, Division of Internal Medicine, Academic Medical Center, Amsterdam 1105 AZ, The Netherlands. ⁷Central European Institute of Technology, Masaryk University, Brno 625 00, Czech Republic. ⁸Pirogov Russian National Research Medical University, Moscow 117997, Russia. These authors contributed equally: Martin S. Davey, Carrie R. Willcox, Stuart Hunter. Correspondence and requests for materials should be addressed to M.S.D. (email: m.davey@bham.ac.uk) or to B.E.W. (email: b.willcox@bham.ac.uk)

$\gamma\delta$ T cells have coevolved alongside B cells and $\alpha\beta$ T cells in the vertebrate immune system for almost 450 million years¹. They provide anti-microbial² and anti-tumour immunity³, but whether they occupy an innate-like or adaptive immunological niche has remained unclear. Notably, $\alpha\beta$ T cells incorporate a group of unconventional T cells, including mucosal-associated invariant T (MAIT) cells and invariant natural killer T (iNKT) cells that recognise antigens in the context of single MHC-like proteins (MR1 and CD1d), and display a semi-invariant T-cell receptor (TCR) repertoire, suggestive of an innate-like biology whereby TCR sensitivity is retained but the $\gamma\delta$ TCR may arguably function as a surrogate pattern recognition receptor⁴. Notably, studies in mice have suggested that innate-like $\gamma\delta$ T-cell development in the thymus can occur via distinct pathways involving agonistic signals⁵. In addition, recently, Wencker et al.⁶ have suggested that following TCR triggering during development, mouse innate-like T cells may transition to a state of TCR hyporesponsiveness in which they preferentially respond to TCR-extrinsic stimuli such as cytokine exposure.

Human $\gamma\delta$ T cells are often delineated into $V\delta 2^+$ and $V\delta 2^-$ subsets⁷. $V\delta 2^-$ $\gamma\delta$ T cells have been directly implicated in anti-viral and anti-tumour immunity^{3,8} and utilise germline-encoded antigen receptors also present on innate-like lymphocytes, including NKG2D and NKP30^{9,10}. However, recent evidence has suggested that they may adopt a TCR-dependent adaptive immunobiology, based on highly clonotypically focused expansions alongside differentiation from a naive to effector phenotype¹¹ and perturbations in clonal expansion upon cytomegalovirus (CMV) infection in post-stem cell transplant patients¹². Conversely, $V\delta 2^+$ T cells are arguably the prototypic unconventional T cell, typically co-expressing $V\gamma 9$ TCR chains and representing the major $\gamma\delta$ subset in adult peripheral blood¹³. $V\gamma 9^+$ $V\delta 2^+$ T cells respond to prenyl pyrophosphate metabolites (phosphoantigens, or P-Ags) produced either by the host mevalonate pathway (isopentenyl pyrophosphate, IPP) or microbial non-mevalonate pathway (*(E)*-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate, HMB-PP)¹⁴, which are sensed in the context of butyrophilin 3A1 (BTN3A1)^{15–17}. They mount important anti-microbial immune responses, including against *Mycobacterium tuberculosis*¹³ and *Plasmodium falciparum*¹⁸, and drive $\alpha\beta$ T-cell responses^{19,20}.

Previously, the extrathymic expansion of $V\gamma 9^+$ $V\delta 2^+$ T cells observed in adult peripheral blood, which is proposed to result from exposure to P-Ag-producing microbes encountered after birth^{21,22} and thought to involve clonotypic expansion^{23,24}, arguably suggests an adaptive immunobiology. However, $V\gamma 9^+$ $V\delta 2^+$ T cells are highly enriched in foetal peripheral blood and display restricted complementarity-determining region (CDR) 3 $\gamma 9$ usage early in gestation²⁵, more consistent with an innate-like pre-natal repertoire of $V\gamma 9^+$ $V\delta 2^+$ T cells.

To better understand human $V\delta 2^+$ $\gamma\delta$ T-cell immunobiology we carried out a dedicated analysis of their clonotypic diversity and cellular phenotype. Our findings suggest $V\delta 2^+$ T cells can be delineated into two discrete subsets: $V\gamma 9^+$ $V\delta 2^+$ T cells adopt a predominantly innate-like biology originating in neonatal development and allowing a degree of clonotypic plasticity, whereas $V\gamma 9^-$ $V\delta 2^+$ T cells adopt a distinct adaptive immunobiology, including focused clonal expansions and differentiation evident both in peripheral blood and solid tissues, and generated in response to acute viral infection. These findings revise our understanding of human $\gamma\delta$ T cells, and prompt further investigation of the adaptive immunity provided by the $V\gamma 9^-$ $V\delta 2^+$ subset.

Results

Adult and neonatal $V\delta 2^+$ T cells have similar TCR diversity. Consistent with previous studies, we confirmed that human cord

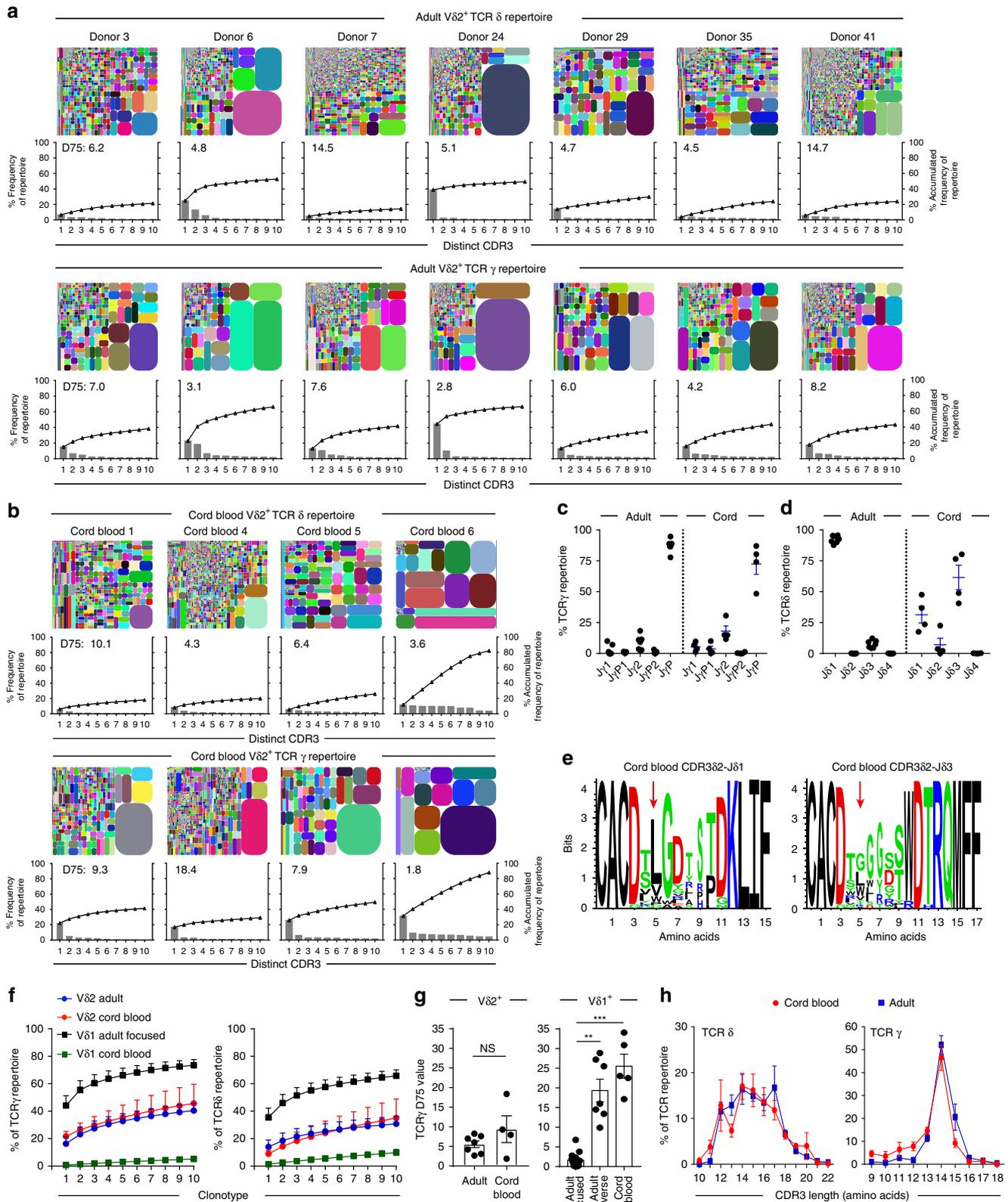
blood $V\delta 2^+$ $\gamma\delta$ T cells were present at relatively low levels, but were increased as a proportion of peripheral T cells by adulthood^{22,26} (Supplementary Fig. 1a), resulting in an adult human $V\delta 2^+$ T-cell repertoire that is uniformly responsive to P-Ag metabolites (Supplementary Fig. 1b). To address whether $V\delta 2^+$ T-cell expansion from the neonatal pool is highly clonally focused, as for $V\delta 1^+$ T cells¹¹, we performed a TCR repertoire analysis restricted to the $V\delta 2^+$ T-cell population (Fig. 1a, b). Indicative of successful $V\delta 2^+$ $\gamma\delta$ T-cell sorting, TCR δ repertoires entirely comprised $V\delta 2$ chain usage (Supplementary Fig. 1c). Consistent with previous findings, TCR γ repertoires were predominantly composed of $V\gamma 9$ chains (Supplementary Fig. 1d) and the joining region $J\gamma P$ (Fig. 1c), and neonatal $V\delta 2$ chains preferentially used joining region $J\delta 3^{27}$, whereas $J\delta 1$ was more commonly utilised in adults²⁸ (Fig. 1d). TCR δ -mediated responses to P-Ag have been associated with hydrophobic amino acid residues (LVW) at position 5 in CDR3 $\delta 2$ sequences²⁹, and this is observed in adult $V\delta 2$ CDR3 repertoires¹¹. We found that although neonatal $V\delta 2$ – $J\delta 3$ sequences generally lacked this motif, and overall were enriched relative to $V\delta 2$ – $J\delta 1$ sequences for neutral residues at position 5, particularly Gly, some $V\delta 2$ – $J\delta 3$ sequences contained the hydrophobic residue Leu, whereas $V\delta 2$ – $J\delta 1$ sequences, although rarer in neonates, were highly enriched for hydrophobic residues at position 5 (Fig. 1e).

We next analysed $V\delta 2^+$ T-cell repertoires using approaches we previously applied to the $V\delta 1^+$ compartment¹¹. Tree plot analysis revealed the presence of some relatively prominent clonotypes in adult $V\delta 2$ TCR γ (between 12 and 47%) and TCR δ (between 1.8 and 39%) repertoires (Fig. 1a). The ten most prevalent TCR γ clonotypes in each donor formed a substantially smaller portion ($P = 0.003$; Mann–Whitney) of the $V\delta 2^+$ T-cell repertoire (mean 40.5% of total $V\delta 2$ reads) than expanded $V\delta 1$ clonotypes (mean 73.42% of total $V\delta 1$ reads¹¹). Unlike the $V\delta 1$ compartment, similarly prominent clonotypes were also present in cord blood $V\delta 2^+$ TCR γ (between 18 and 34%) repertoires (Fig. 1a, b). Analysis of the cumulative frequency curves of the 10 most prevalent CDR3 sequences highlighted both adult and cord blood $V\delta 2^+$ TCR repertoires exhibit an intermediate degree of clonotypic focusing, relative to cord blood $V\delta 1$ (low focusing) and clonally expanded adult $V\delta 1$ repertoires (high focusing; Fig. 1f). Analysis of D75 diversity metrics (the percentage of clonotypes required to occupy 75% of TCR repertoire) highlighted that in $V\delta 2^+$ T-cell repertoires diversity did not generally differ between adult and cord blood $V\delta 2^+$ repertoires (Fig. 1g and Supplementary Fig. 1e). Moreover, $V\delta 2^+$ TCR repertoire D75 metrics were also independent of the level of diversity seen in matched $V\delta 1^+$ TCR repertoires (Supplementary Fig. 1f). Finally, and consistent with a broadly similar $V\delta 2^+$ repertoire in cord and adult blood, the CDR3 length profile of TCR δ and TCR γ repertoires were highly comparable between cord blood and adult (Fig. 1h). Therefore, while some variability in the level of focusing is observed in both settings, the neonatal and adult $V\delta 2^+$ repertoires are broadly similar in diversity, in terms of both accumulated clonotype frequency, D75, and CDR3 length analysis.

Public $V\gamma 9^+$ CDR3s span adult and neonatal $V\delta 2^+$ T cells. Due to the largely similar $V\delta 2^+$ repertoire characteristics found in cord blood and adult, we next assessed the proportion of CDR3 amino acid sequences that were “public”, i.e., detected in more than one donor. Consistent with a high degree of diversity within the TCR δ repertoire, adult TCR $\delta 2$ repertoires contained few shared CDR3 δ sequences (mean 8.7%), although this proportion was increased in cord blood CDR3 δ repertoires (mean 36.8%; Fig. 2a, Supplementary Table 1 and 2). However, a mean 79.3% of

the CDR3 γ amino acid sequences present in a given individual were shared with at least one other adult or cord blood donor (Fig. 2b). Moreover, the majority of sequences in cord blood were directly shared in adult repertoires (Fig. 2c), indicating a core set of CDR3 γ sequences in adult donors essentially indistinguishable from equivalent unrelated neonatal donor sequences.

To further explore this publicity, we examined the 10 most prevalent CDR3 γ sequences from our cohort (see Methods), representing a mean 38.9% of the total CDR3 γ repertoire (Fig. 2d). This highlighted a set of public CDR3 γ sequences that were prevalent in both adult and cord blood repertoires, and at similar frequencies (Fig. 2d). Notably, two simple CDR3 γ sequences were prevalent in all adult and neonatal repertoires



analysed: CALWEVQELGKKIKVF (mean 13.1% in adult, and mean 24.6% in neonates) and CALWEVRELGKKIKVF (mean 4.3% in adults and mean 5.4% in neonates). Two processes explain the high degree of sharing within the CDR3 γ repertoire. Firstly, single-cell TCR sequencing highlighted that CALWEVQELGKKIKVF could be generated by simple recombination of V γ 9 and J γ P gene segments containing no added N nucleotides ("germline") (Table 1). This sequence was common in foetal liver³⁰, and foetal blood²⁵ and observed to persist into adulthood³¹. While a high prevalence in repertoire data might indicate an expanded clonotype, single-cell TCR analysis also revealed that this germline-encoded CDR3 γ sequence in fact was comprised of multiple TCR "pseudoclonotypes", in which the same CDR3 γ was independently generated and paired with different CDR3 δ sequences in the same individual (Table 1). Secondly, CDR3 γ sequences such as CALWEVQELGKKIKVF and CALWEVRELGKKIKVF could also be generated by convergent recombination, whereby variable degrees of exonuclease activity and N nucleotide addition result in the generation of the same CDR3 γ amino acid sequences from different nucleotide sequences (Table 1). Both of these processes explain why CDR3 γ repertoires were less diverse compared to CDR3 δ repertoires (Fig. 2e), and are consistent with the generation of public TCR sequences³², but importantly, these mechanisms are not evident in V δ 1⁺ TCR repertoires¹¹.

Altered V γ 9⁺ V δ 2⁺ T-cell phenotype links to clonal expansion.

We evaluated the memory phenotype of V δ 2⁺ T cells from each donor to establish the relationship with TCR clonotype. Adult V γ 9⁺ V δ 2⁺ T cells from healthy donors (16/18 of total donors, 5/7 TCR repertoire study donors) regularly display a major CD27⁺ CD45RA⁻ phenotype (similar to Central Memory (CM) CD8⁺ $\alpha\beta$ T cells) and a minor CD27⁻ CD45RA⁻ phenotype (similar to Effector Memory (EM) CD8⁺ $\alpha\beta$ T cells) (Fig. 2f), consistent with other studies³³. Importantly, V γ 9⁺ V δ 2⁺ T cells differ from V δ 1⁺ T-cell populations in their expression levels of CD27 (Fig. 2g) and distinct naive and effector T-cell surface marker expression (Fig. 2h).

Although overall our adult cohort displayed similar V γ 9⁺ V δ 2⁺ TCR diversity to cord blood samples, two adult donors (6 and 24) appeared to exhibit a degree of clonotypic focusing, based on their low D75 values for both TCR γ (mean 2.9) and TCR δ (mean 5.0) combined with their higher cumulative frequency of the 10 most prevalent clonotypes (mean 60.5% of TCR γ and mean 50.3% of TCR δ) (Fig. 1a). Consistent with this, each donor displayed individual clonotypes with a similar amplified frequency in both TCR γ and δ repertoires (Fig. 1a). Both donor 6 and 24 each possessed a distinct uncommon but shared sequence, CALWEVRKELGKKIKVF (detected in donor 6 and

also in 2 other donors) or CALWEKMQELGKKIKVF (detected in donor 24 and in 4 other donors), usually outside the top 10 most prevalent shared clonotypes, but now representing 18.5% (donor 6) and 44.6% (donor 24) of the TCR γ repertoire, respectively (Fig. 2d). Also, donor 6 possessed a common sequence found in all donors, CALWETQELGKKIKVF, normally accounting for a mean 0.95% of the TCR γ repertoire, but increased to 22.9% of donor 6's TCR γ repertoire (Fig. 2d). Furthermore, these unusual clonal amplifications were linked to phenotypic changes in the total V γ 9⁺ V δ 2⁺ T-cell population, whereby these donor's cells adopted a predominant CD27⁻ CD45RA⁻ phenotype (Fig. 2i). These data suggest that despite the public repertoire and associated dominant phenotype (CD27⁺ CD45RA⁻) in the majority of donors, there is scope within the V γ 9⁺ V δ 2⁺ TCR repertoire for a degree of specific clonal amplification, alongside concomitant phenotypic changes.

A discrete V γ 9⁻ V δ 2⁺ subset persists from birth into adults.

TCR γ repertoire datasets from V δ 2⁺ T cells were dominated by V γ 9 TCR sequences, but we noted the presence of non-V γ 9 (i.e., V γ 2–8) sequences in adult TCR γ repertoires (mean 4.46%) (Fig. 3a). Although such chain usage could conceivably represent second productive TCR γ rearrangements in V γ 9⁺ V δ 2⁺ T cells, an alternative possibility was that they reflected the existence of an unusual V γ 9⁻ V δ 2⁺ T-cell subset that may persist in adulthood. As V γ 9⁻ TCR usage by neonatal V δ 2⁺ T cells has been previously reported³⁴, we established flow cytometry-based identification of V γ 9⁻ cells in cord blood V δ 2⁺ T cells (Fig. 3b) and quantified their relative frequency (mean 32.41%) (Fig. 3c). Importantly, adult V δ 2⁺ T-cell pools retained this V γ 9⁻ subset at low frequency (mean 4.82%) (Fig. 3c and Supplementary Fig. 2a), but on average occupying a similar frequency within total peripheral blood T cells to that of α -GalCer/CD1d reactive natural killer T cells (NKT) (Fig. 3d and Supplementary Fig. 2b). Notably, the identification of V γ 9⁻ V δ 2⁺ T cells was only possible with the use of an anti-V δ 2 TCR antibody clone, 123R3 (Miltenyi), while this population was undetectable using another TCR V δ 2 antibody clone, B6 (Biolegend), which is likely to be specific for the V γ 9⁺ V δ 2⁺ TCR pairing (Fig. 3e).

V γ 9⁻ V δ 2⁺ T cells and V γ 9⁺ V δ 2⁺ T cells are distinct. As adult peripheral blood V δ 2⁺ T cells are commonly responsive to host and microbial P-Ags, we assessed the proliferative capacity of V γ 9⁻ V δ 2⁺ T cells towards microbial P-Ag (HMB-PP). While CD8⁺, both V δ 2⁺ subsets and V δ 1⁺ T cells all proliferated in response to anti-CD3/CD28 stimulation, only V γ 9⁺ V δ 2⁺ T cells responded to HMB-PP (Fig. 3f). We then compared the memory phenotype of V γ 9⁺ V δ 2⁺ T cells to other human $\gamma\delta$ T-cell

Fig. 1 Adult and cord blood V δ 2⁺ TCR repertoires are broadly similar. **a** Tree maps show each adult donor's V δ 2⁺ TCR repertoire, with each CDR3 clonotype as a coloured segment (each coloured CDR3 segment is chosen randomly and does not match between plots) plotted in relation to the total repertoire size and accompanying clonotype frequency graphs showing the individual clone frequency (left y axis) and the accumulated frequency for the 10 most prevalent clonotypes (right y axis). Inset into each graph are D75 repertoire diversity metrics (measuring the percentage of clonotypes required to occupy 75% of the total TCR repertoire). **b** Tree maps showing TCR γ and δ CDR3 clonotypes, accumulated frequency graphs and D75 metric from cord blood V δ 2⁺ T cells. **c** J γ and **d** J δ segment usage in V δ 2⁺ TCR repertoires from adult peripheral blood ($n = 7$) and cord blood samples ($n = 4$). **e** Logo analysis of amino acid enrichment at each position in neonatal V δ 2–J δ 1 CDR3 δ (left) and V δ 2–J δ 3 CDR3 δ (right) sequences. Analysis was confined to the 10 most abundant CDR3 δ sequences of 13–16 amino acid length. The different amino acids are coloured according to physicochemical properties (acidic (red); basic (blue); hydrophobic (black); and neutral (green)). Red arrows indicate position 5 in the CDR3 sequence (see Methods section). **f** Comparison of accumulated frequency curves generated from the 10 most prevalent TCR γ (left) and δ (right) clonotypes in V δ 2⁺ and V δ 1⁺ TCR repertoires (V δ 1 cohort data analysed from¹¹) from adult peripheral blood (V δ 2⁺, $n = 7$ and V δ 1⁺, $n = 13$) and cord blood (V δ 2⁺, $n = 4$ and V δ 1⁺, $n = 5$). **g** Comparison of TCR γ D75 metrics from adult peripheral blood and cord blood V δ 2⁺ (adult: $n = 7$; cord blood: $n = 4$) and V δ 1⁺ repertoires (adult focused: $n = 13$; adult diverse: $n = 7$; cord blood: $n = 5$). **h** Comparison of the CDR3 length profiles in V δ 2⁺ TCR δ and γ repertoires from adult peripheral blood ($n = 7$) and cord blood ($n = 4$). Error bars indicate means \pm SEM; ** $P < 0.01$; *** $P < 0.001$; p -values were determined by Student's t -test (**g**: left) and Kruskal-Wallis test (ANOVA) with Tukey's post hoc testing (**g**: right). NS not significant

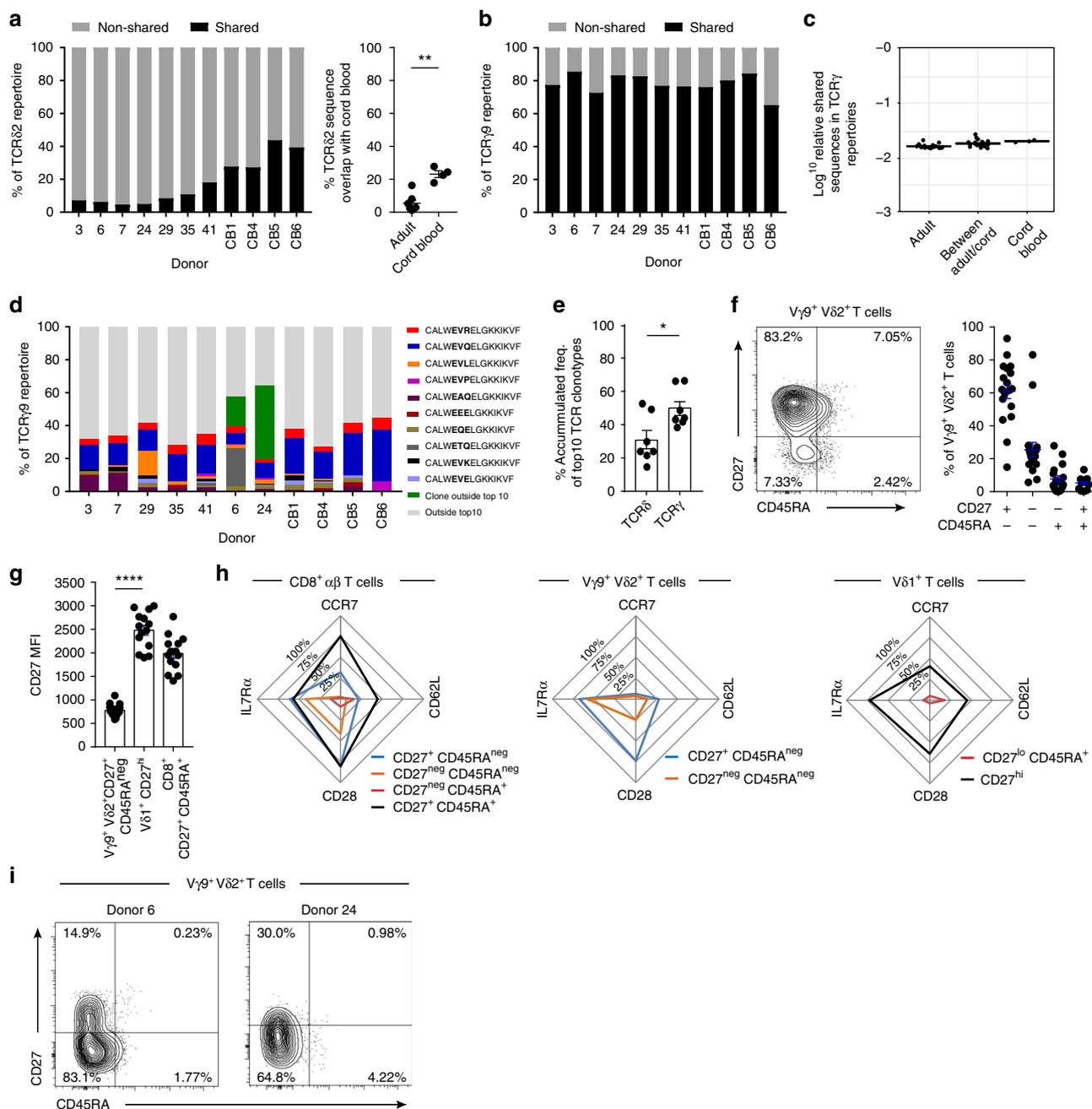


Fig. 2 $V\delta 2^+$ TCR repertoires are formed of public clonotypes. **a** Percentage of CDR3 δ sequences (amino acid) shared between >2 donors within both adults and cord blood (left) and sequences shared with cord blood only (right). **b** Percentage of CDR3 γ sequences (amino acid) shared between >2 donors (public sequences). **c** Comparison of the sequence overlap (relative publicity) of TCR $\gamma 9$ repertoires in adult peripheral blood donors ($n = 7$) or cord blood samples ($n = 3$) and then between both groups. **d** Frequency of each of the 10 most common clonotypes in each donor's TCR $\gamma 9$ repertoire, with the addition of exceptional expanded shared clonotypes usually found outside the top 10 (dark green). **e** Comparison of the accumulated repertoire frequency occupied by the first 10 clonotypes in $V\delta 2^+$ TCR δ and γ ($n = 7$). **f** CD27 and CD45RA T-cell memory marker expression by $V\gamma 9^+$ $V\delta 2^+$ T cells from adult peripheral blood samples ($n = 18$). **g** Comparison of CD27 expression levels (MFI) on CD27 $^+$ CD45RA neg $V\gamma 9^+$ $V\delta 2^+$ T cells ($n = 18$) and CD27 hi $V\delta 1^+$ ($n = 14$) and CD27 $^+$ CD45RA $^+$ CD8 $^+$ T cells ($n = 14$) from adult peripheral blood samples. **h** Summary radar plot data detailing the mean % positive cells for each indicated T-cell marker analysed within each sub-population of $V\gamma 9^+$ $V\delta 2^+$ ($n = 18$), $V\delta 1^+$ ($n = 14$) and CD8 $^+$ $\alpha\beta$ T cells ($n = 14$). **i** Healthy adult peripheral blood donor 6 and 24's expression of CD27 and CD45RA T-cell memory markers on $V\gamma 9^+$ $V\delta 2^+$ T cells. Error bars indicate means \pm SEM; * $P < 0.05$; ** $P < 0.01$; **** $P < 0.0001$; p -values were determined by Student's t -test (**a**, **e**) and one-way ANOVA with Tukey's post hoc testing (**g**)

populations. Cord blood $V\gamma 9^-$ $V\delta 2^+$ T cells expressed a CD27 hi phenotype which extended to the adult population, remarkably similar to naive-like $V\delta 1^+$ T cells (Fig. 3g). We next assessed the expression of naive and effector T-cell markers in paired $\gamma\delta$ T-cell populations from the same donors. $V\gamma 9^-$ $V\delta 2^+$ T cells

commonly displayed lower levels of TCR $V\delta 2$ expression and higher levels of CD27 expression than their $V\gamma 9^+$ counterparts (Fig. 3h). $V\gamma 9^-$ $V\delta 2^+$ T cells expressed the lymphoid homing receptor CCR7, the homeostatic cytokine receptor IL7R α (Fig. 3i) but lacked the cytotoxic effector molecule Granzyme

Table 1 Single-cell TCR sequencing of adult V γ 9⁺ V δ 2⁺ T cells

Donor	CDR3 γ 9 (amino acid)	CDR3 γ 9 nucleotide (nt) sequence	N nt	P nt	CDR3 δ 2 (amino acid)
3	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	Germline	0	CACDSLGLDTPNFDKLF
	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	Germline	0	CACDTGGAAQSWDRQMF
28	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTACAAGAGTTGGGCCAAAAAATCAAGGTATTT	1	0	CACDSLGGTYTDKLF
	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	Germline	0	CACDIILGGQYTDKLF
35	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	1	2	CACESLGPPTGGNPPSSDKLIF
	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTACAAGAGTTGGGCCAAAAAATCAAGGTATTT	1	0	CACDTIHTRTGGPQVTDKLF
42	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	Germline	0	CACDGLGGEYTDKLF
	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	Germline	0	CACDKVGYGSPWDRQMF
231	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	Germline	0	CACDTIPTGGHDPYTDKLF
	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTACAAGAGTTGGGCCAAAAAATCAAGGTATTT	1	0	CACDVTTRDKGADKLF
1014	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTACAAGAGTTGGGCCAAAAAATCAAGGTATTT	1	0	CACDTMGARHTDKLF
	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	1	0	CACDTGFLVQSHGPKRTDKLF
261	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	Germline	0	CACDRLGGNTDKLF
	CALWEVRELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	Germline	0	CACDTPSTGGPKDTDKLF
261	CALWEVQELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	Germline	0	CACDTSGLQRYNSWDRQMF
	CALWEVRELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	3	0	CACDVTGRTVGRDVKLF
	CALWEVRELGKKIKVF	TGTGCCTTGTGGGAGGTGCAAGAGTTGGGCCAAAAAATCAAGGTATTT	1	1	CACDVTGDWGTPLYTDKLF

Public TCR γ 9 clonotypes and their paired TCR δ 2 sequence from seven donors. Sequences were analysed using IMGJ Junction Analysis, which identified V, D and J gene segments used and highlighted N nucleotide addition (bold and italic) and P nucleotide addition (underlined). The table depicts CDR3 γ 9 amino acid sequence, CDR3 γ 9 nucleotide (nt) sequence, germline or N nt addition, P nt addition, CDR3 δ 2 amino acid sequence. Hydrophobic amino acids at position 5 (see Methods section) in the CDR3 δ 2 sequence are highlighted in bold

(Grz) A or the endothelial homing receptor CX₃CR1 (Fig. 3j), with these expression patterns contrasting with V γ 9⁺ V δ 2⁺ T cells (V δ 2^{hi} CD27^{int} CCR7⁻ IL7R α ⁺ Grz A⁺ CX₃CR1⁺) but together highly similar to CD27^{hi} V δ 1⁺ T cells (Fig. 3h–j). Moreover, V γ 9⁻ V δ 2⁺ T cells were responsive to CD3/CD28 stimulation, but like V δ 1 T cells were unresponsive to IL-12/IL-18 (Fig. 3k), whereas V γ 9⁺ V δ 2⁺ T cells were responsive to both CD3/CD28 and IL-12/IL-18¹¹. CD27^{hi} V δ 1⁺ T-cell populations lack expanded clonotypes and possess a diverse and private TCR repertoire¹¹. Analysis of the publicity of adult V γ 9⁻ V δ 2⁺ TCR repertoires indicated a far more private TCR repertoire compared to V γ 9⁺ V δ 2⁺ T cells (Fig. 3l), but equivalent to the highly private V δ 1⁺ TCR repertoire¹¹. This was confirmed by single-cell TCR sequencing of adult V γ 9⁻ V δ 2⁺ T-cell populations, which contained diverse TCR sequences, comprising a range of V γ chains (Fig. 3m), and featuring TCR γ and δ chains that each lacked motifs previously linked to P-Ag reactivity (Table 2). In contrast, V γ 9⁺ V δ 2⁺ T cells sorted from the same donors displayed sequences that included prevalent shared TCR γ sequences (Fig. 3m and Table 1). Collectively, these data suggest that the V γ 9⁻ V δ 2⁺ subset is functionally, phenotypically and clonotypically distinct from its V γ 9⁺ counterparts, but may share a similar biology to that of V δ 1⁺ T cells.

V γ 9⁻ V δ 2⁺ T cells clonally expand into effectors. Within our healthy donor cohort, one individual (“donor X”) had a substantially increased V γ 9⁻ V δ 2⁺ T-cell population, comprising 1.2% of all CD3⁺ T cells (Fig. 4a), far higher than the 17 other healthy donors (mean 0.13% of CD3⁺ T cells). Importantly, and in contrast to the dominant CD27^{hi} phenotype displayed by V γ 9⁻ V δ 2⁺ T cells in all other healthy donors, “donor X’s” V γ 9⁻ V δ 2⁺ T-cell population had downregulated CD27 and retained CD45RA expression (CD27^{lo/neg} phenotype), now accounting for 91.4% of the V γ 9⁻ V δ 2⁺ T-cell population (Fig. 4a). We next carried out single-cell TCR sequencing to determine if any particular $\gamma\delta$ TCR clonotypes correlated with this phenotypic change. Strikingly, a single clone now dominated the V γ 9⁻ V δ 2⁺ T-cell population (35/36 single cells sequenced), composed of a V δ 2 CDR3 (CACGSWWGTYTDKLF) paired with a V γ 8–J γ P1 chain (Fig. 4b), and now represented the single most dominant clonotype within the V δ 2⁺ T-cell repertoire (Fig. 4c). As this dominant clonotype occurred in conjunction with a CD27^{lo/neg} phenotype, we assessed effector T-cell marker expression. “Donor

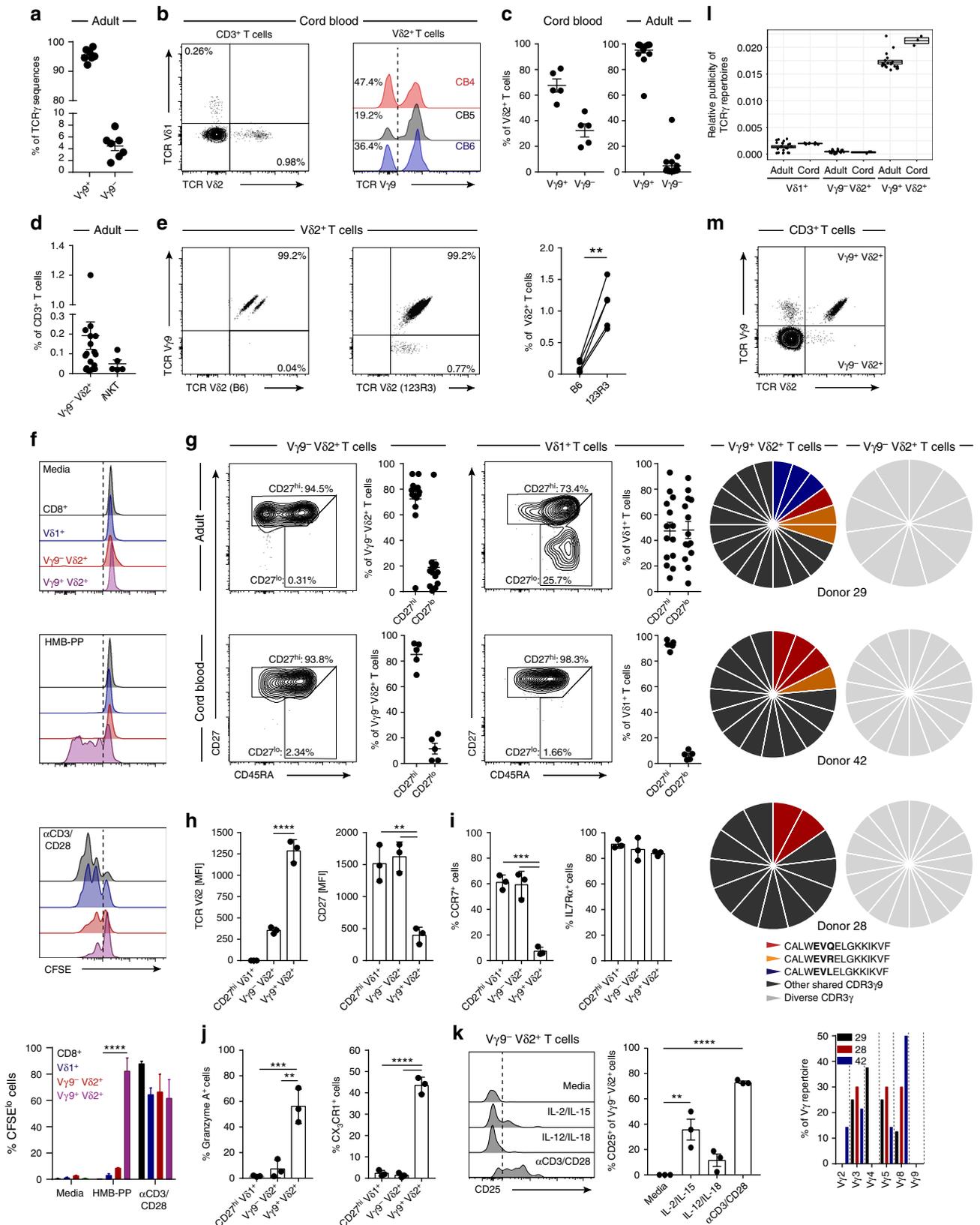
X’s” V γ 9⁻ V δ 2⁺ T-cell population expressed the antibody-mediated cytotoxicity receptor CD16 (Fc γ RIIIa), CX₃CR1, Grz A and had downregulated CCR7, in contrast to naive CD8 T cells, but closely matching CD27^{lo/neg} V δ 1⁺ T cells (Fig. 4d), which also comprised dominant clonotypes¹¹. These results highlight the potential for the V γ 9⁻ V δ 2⁺ T-cell compartment to undergo clonal selection and differentiation, mimicking the transition to clonality and effector status previously observed in V δ 1⁺ T cells¹¹.

Clonally focused V γ 9⁻ V δ 2⁺ T cells infiltrate human liver. V δ 1⁺ T cells are predominantly associated with a role in tissue immunity^{35,36}. To address whether the V γ 9⁻ V δ 2⁺ T-cell population was also represented in solid tissues, we analysed human liver samples for the presence of intrahepatic V γ 9⁻ V δ 2⁺ T cells (Fig. 5a). While the V δ 2⁺ T cells in the liver comprised a lower proportion of total CD3⁺ T cells than in peripheral blood, intrahepatic V δ 2⁺ T cells generally contained a higher fraction of V γ 9⁻ cells (Fig. 5b). Phenotypic analysis of intrahepatic V γ 9⁻ V δ 2⁺ T cells revealed these cells to consistently display a CD27^{lo/neg} phenotype (Fig. 5c, d). Single-cell TCR sequencing analysis, highlighted that a majority of intrahepatic V γ 9⁻ V δ 2⁺ T cells were comprised of a limited set of expanded clonotypes (Fig. 5e and Table 2). Moreover, clonally expanded intrahepatic V γ 9⁻ V δ 2⁺ T cells expressed effector markers CX₃CR1 and Grz A but had downregulated IL7R α (Fig. 5f).

Clonal expansion of V γ 9⁻ V δ 2⁺ T cells in acute CMV infection. The clonotypic and phenotypic parallels between V γ 9⁻ V δ 2⁺ and V δ 1⁺ $\gamma\delta$ T cells, combined with their strong distinction from P-Ag sensing V γ 9⁺ V δ 2⁺ T cells, prompted us to investigate if this subset was reactive to acute cytomegalovirus (CMV) infection, previously shown to drive the numeric³⁷ and clonal¹² expansion of V δ 1⁺ T cells. We examined a cohort of five CMV-seronegative (CMV⁻) patients receiving CMV-seropositive (CMV⁺) kidney transplants who then went on to either develop post-operative acute CMV infection (patient 261, 231 and 1014), or who remained CMV⁻ (patient 279 and 282)³⁸. We analysed peripheral blood samples for V γ 9⁻ V δ 2⁺ T cells before and after post-transplant CMV infection. Prior to CMV infection, at 0–4 weeks post-transplant, V γ 9⁻ V δ 2⁺ T cells represented low frequency populations in all donors (Fig. 6a and Supplementary Fig. 3). However, following CMV infection, occurring between 5–8 weeks post-transplant, we observed the substantial expansion

of the $V\gamma 9^- V\delta 2^+$ T-cell population in donor 261 and 231 (Fig. 6a). Conversely, patients 282 and 279 who did not become infected with CMV, displayed no expansion of $V\gamma 9^- V\delta 2^+$ T cells over a 5-year period after transplantation (Supplementary Fig. 3a). Of note is patient 1014, who became infected with CMV but did not display an expanded $V\gamma 9^- V\delta 2^+$ T-cell population

(Supplementary Fig. 3b), suggesting inter-individual differences in the $V\gamma 9^- V\delta 2^+$ T-cell response to CMV infection consistent with observations in $V\delta 1^+$ T cells from persistently infected CMV⁺ healthy adult donors¹¹. Phenotypic analysis of $V\gamma 9^- V\delta 2^+$ T cells from CMV⁻ patients 279 (Supplementary Fig. 3c) and 261 immediately after transplant (Fig. 6b), indicated a



CD27^{hi} phenotype. Conversely after CMV infection, the expanded V γ 9⁻ V δ 2⁺ T-cell populations in patients 261 and 231 acquired a dominant CD27^{lo/neg} phenotype (Fig. 6b). This prompted us to perform single-cell TCR sequencing of expanded V γ 9⁻ V δ 2⁺ T-cell populations in donors 261 and 231, which identified the selective expansion of dominant clonotypes after CMV infection and evidence of continued clonotypic selection 5 years later (Fig. 6c and Table 2). Finally, we assessed functional T-cell marker expression in V γ 9⁻ V δ 2⁺ T cells from patient 261 and 231 following CMV infection. V γ 9⁻ V δ 2⁺ T cells from healthy donors expressed CCR7, IL7R α and the co-stimulatory molecule CD28, while patient 261 and 231's V γ 9⁻ V δ 2⁺ T cells had downregulated these markers and upregulated CX₃CR1 and Grz A and B (Fig. 6d). Expression of these effector markers occurred concurrently with V γ 9⁻ V δ 2⁺ T-cell clonal expansion at ~1 year (patient 261) and were stable for over 5 years (patient 231) (Fig. 6c, d). These data reveal that as for V δ 1⁺ T cells, V γ 9⁻ V δ 2⁺ T cells exhibit a strong relationship between clonotype and phenotype and are suggestive of a role in adaptive antiviral immunosurveillance.

Discussion

This study sheds light on the biology of V δ 2⁺ T cells, highlighting divergent V γ 9⁺ and V γ 9⁻ subsets. Despite the ~ten-fold increase in V γ 9⁺ V δ 2⁺ T-cell numbers in the first year of life, our data indicate that the extent of clonotypic focusing within the V δ 2-associated V γ 9⁺ repertoire is broadly similar between neonatal cord blood and adulthood. This finding contrasts with the V δ 1⁺ TCR repertoire, which typically displays pronounced clonotypic focusing in adults relative to a highly unfocused neonatal repertoire¹¹. In addition, our analyses reveal an underappreciated degree of public amino acid sequences in the V γ 9 repertoire. This degree of V γ publicity, which relates to both recombination of simple V γ 9 sequences (either germline encoded or with limited N nucleotide addition) and convergent recombination, underlines the semi-invariant nature of the V γ 9⁺ V δ 2⁺ TCR repertoire. Importantly, using single-cell approaches public V γ clonotypes were shown frequently to pair with diverse V δ 2 chains within single individuals, and, therefore, do not usually represent expanded clonotypes. The innate-like biology of the semi-invariant V γ 9⁺ V δ 2⁺ T-cell subset contrasts markedly with the V δ 1⁺ subset, which displays clear hallmarks of adaptive immunity such as clonal amplification and differentiation from an initially unfocused, private TCR repertoire.

Collectively, our repertoire data suggest a model involving a multi-layered selection of V γ 9⁺ V δ 2⁺ T cells from foetal development into adulthood. First, as previously suggested by Dimova et al.²⁵, we propose gestational selection of the V γ 9⁺ V δ 2⁺ subset, potentially via exposure to endogenous P-Ag (e.g., IPP) most likely mediated by BTN3A1^{15–17,39,40}, allowing pre-

programming of a P-Ag reactive V γ 9⁺ V δ 2⁺ TCR repertoire. In keeping with this, we observed that V γ 9 chains are heavily enriched for the J γ P segment, which is not generally present in V γ 9 chains present in V δ 1⁺ TCR repertoires¹¹; similarly the J δ 1⁺ CDR3 δ 2 repertoire in cord blood is enriched for a hydrophobic amino acid at position 5 required for P-Ag recognition^{21,29}, but not in V δ 2 sequences in V γ 9⁻ V δ 2⁺ T cells (see below). Second, in the context of the reported lack of thymic output of V δ 2⁺ T cells after birth⁴¹, our observation that adult and cord blood repertoires contain overlapping canonical P-Ag-sensing V γ 9 sequences strongly indicates there is polyclonal post-natal expansion of this pre-selected repertoire, most likely in response to microbial P-Ag exposure¹¹. However, our observation that cord blood V δ 2 sequences generally utilised J δ 3 segments (also observed in foetal $\gamma\delta$ T cells,⁴²) whereas adult V δ 2 sequences preferentially used J δ 1 segments, suggests that post-natal microbial exposure may impose some constraints on the TCR repertoire that differ from those required in gestation. Of note, V δ 2–J δ 1 sequences typically contain a hydrophobic amino acid previously highlighted as important for V γ 9⁺ V δ 2⁺ T-cell reactivity to microbially-derived P-Ag^{21,29} whereas this is less conserved in V δ 2–J δ 3. This might imply V γ 9⁺ V δ 2⁺ T cells bearing J δ 1 clonotypes preferentially respond to microbial P-Ag after birth compared to J δ 3 clonotypes, a possibility which future studies can address. Finally, although most individuals' V γ 9⁺ V δ 2⁺ T cells were dominated by public V γ 9 sequences and associated with a common central memory phenotype, our observation that a minority of individuals exhibit amplification of unusual public clonotypes, concomitant with adoption of an effector memory phenotype, most likely reflects a third level of selection that imparts a degree of plasticity in this fundamentally innate-like paradigm. This observation extends previous work by Ryan et al.³³ indicating that further functional and transcriptional changes are linked to an effector memory phenotype in V γ 9⁺ V δ 2⁺ T cells. Therefore, the semi-invariant V γ 9⁺ V δ 2⁺ repertoire may provide scope for selective TCR-mediated responses to microbial P-Ag. Importantly, our model shares a number of features with MAIT cells, a semi-invariant population that is also subject to both gestational selection (through interactions with MR1) and post-natal expansion in response to microbial infection/colonisation⁴³, and that displays clonotype-specific responses to metabolite ligands⁴⁴ and microbial species⁴⁵.

Our analyses also highlight a population of V γ 9⁻ V δ 2⁺ T cells that is highly distinct from the V γ 9⁺ V δ 2⁺ P-Ag-reactive subset, and present in all adult peripheral blood and cord blood, where they commonly exist as a CD27^{hi}, TCR diverse subset phenotypically equivalent to naive-like V δ 1⁺ T cells. We note this distinction based on V γ 9 chain usage cannot be applied to the $\gamma\delta$ T-cell compartment as a whole, as the V δ 1⁺ T-cell subset, which appears to exhibit an adaptive immunobiology, contains a

Fig. 3 V γ 9⁻ V δ 2⁺ T cells are clonally and phenotypically distinct from V γ 9⁺ V δ 2⁺ T cells. **a** Frequency of V γ 9 chain usage in V δ 2⁺ TCR repertoire sequencing data from adult peripheral blood ($n = 7$). **b** Identification of V δ 2⁺ T cells in CD3⁺ T cells (left) and V γ 9⁻ cells within V δ 2⁺ T cells (right) from cord blood ($n = 5$). **c** Frequency of V γ 9⁻ and V γ 9⁺ cells in V δ 2⁺ T cells from cord blood (left; $n = 5$) and adult peripheral blood (right; $n = 18$). **d** Frequency of V γ 9⁻ V δ 2⁺ T cells ($n = 18$) and NKT cells (α Galcer/CD1d⁺; $n = 5$) in adult peripheral blood. **e** V γ 9⁻ V δ 2⁺ T cells identified by TCR V δ 2 antibody clones (B6 and 123R3) in matched adult peripheral blood donors ($n = 5$). **f** Proliferation by CFSE dilution of CD8⁺ $\alpha\beta$ and $\gamma\delta$ T-cell populations from PBMC treated with medium alone, 10 nM HMB-PP or 5 μ g/ml anti-CD3/CD28 ($n = 5$). **g** CD27 and CD45RA T-cell memory marker expression profiles on V γ 9⁻ V δ 2⁺ and V δ 1⁺ T cells from adult peripheral blood (top row: V δ 2, $n = 14$; V δ 1, $n = 14$) and cord blood (bottom row: V δ 2 and V δ 1, $n = 5$). **h** V δ 2⁺ TCR and CD27 expression levels, **i** naive and **j** effector T-cell marker expression on donor matched $\gamma\delta$ T-cell populations from adult peripheral blood ($n = 3$). **k** Sorted CD3⁺ T cells were incubated for 72 h with cytokines or anti-CD3/CD28 beads. V γ 9⁻ V δ 2⁺ T cells were then assessed for the upregulation of CD25 ($n = 3$). **l** Comparison of CDR3 γ sequence sharing in $\gamma\delta$ T-cell repertoires from adult peripheral blood (V δ 1, $n = 20$; V δ 2, $n = 7$) and cord blood (V δ 1, $n = 5$; V δ 2, $n = 3$). **m** CDR3 γ sequence analysis of single cell sorted V γ 9⁺ and V γ 9⁻ V δ 2⁺ T cells, public sequences are coloured (black, shared sequences from deep sequencing); graph shows V γ usage of V γ 9⁻ V δ 2⁺ TCR sequences from each donor. Error bars indicate means \pm SEM; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$; p -values were determined by paired t -test (**e**), RM two-way ANOVA with Tukey's post hoc testing (**f**); one-way ANOVA with Tukey's post hoc testing (**h**, **i**) or Dunnett's post hoc testing (**j**, **k**)

Table 2 Single-cell TCR sequencing of V γ 9⁻ V δ 2⁺ T cells

Donor	CDR3 δ 2	D δ	J δ	CDR3 δ length	V γ	J γ	CDR3 γ length
3	CACGS W WGTYTDKLI F	1,3	1	14	8	P1	11
231	CACDT I VGDTLTDKLI F	3	1	15	8	P1	9
	CACVR G GGYPRGADKLI F	3	1	16	8	1/2	14
	CACDT E GEVNTDKLI F	3	1	14	4	P1	11
261	CACDT R FRPGRSARVVRLTAQL F	1	2	22	3	P1	14
Liver 1	CACSR E QAHTDKLI F	3	1	13	3	1/2	10
	CACDT G WGIRGRYTDKLI F	3	1	17	4	1/2	11
Liver 2	CACDR S GARWDKLI F	—	1	13	5	1/2	10
42	CACD A RHYWGISTDKLI F	3	1	16	2	1/2	12
	CACD A RLGEHTDKLI F	3	1	14	ND	ND	ND
	CACDR M GDLPSWDRQMF F	3	3	18	3	1/2	13
	CACGS S WGVSHYTDKLI F	3	1	16	5	1/2	10
	CACD L LGDFTDKLI F	3	1	14	8	1/2	8
	CACSV R GHWGRSTDKLI F	3	1	16	8	1/2	14
	CACDT R GIGDTPDKLI F	3	1	16	8	P1	12
	CACA H YWGTPWDTDKLI F	3	1	17	3	P1	11
	CACDT G DTDTDKLI F	3	1	13	8	P1	13
	CACDR L LLGDTDKLI F	3	1	14	ND	ND	ND
	CACDT R RTGGRDKLI F	3	1	14	5	1/2	9
	CACD K IRWGNPYTDKLI F	3	1	17	8	P2	10
	CACEV P SYEPYWGTKKYTDKLI F	2,3	1	21	3	1/2	11
	CACDT G DWGINTDKLI F	3	1	15	8	P1	14
	CACD Q KYWGGSSTDKLI F	3	1	16	2	P	13

TCR δ sequences of clonally expanded V γ 9⁻ V δ 2⁺ TCRs from donors expressing a CD27^{lo/neg} phenotype (top section). All TCR δ sequences from a representative healthy donor with a CD27^{hi} phenotype (bottom section). Sequences were analysed using IMGT Junction Analysis, which identified V, D and J gene segments used. The table depicts CDR3 δ sequences, D δ usage, J δ usage, CDR3 δ length, V γ usage, J γ usage and CDR3 γ length. Amino acids (aa) at position 5 of the CDR3 sequence (see Methods) are highlighted in bold. ND not determined

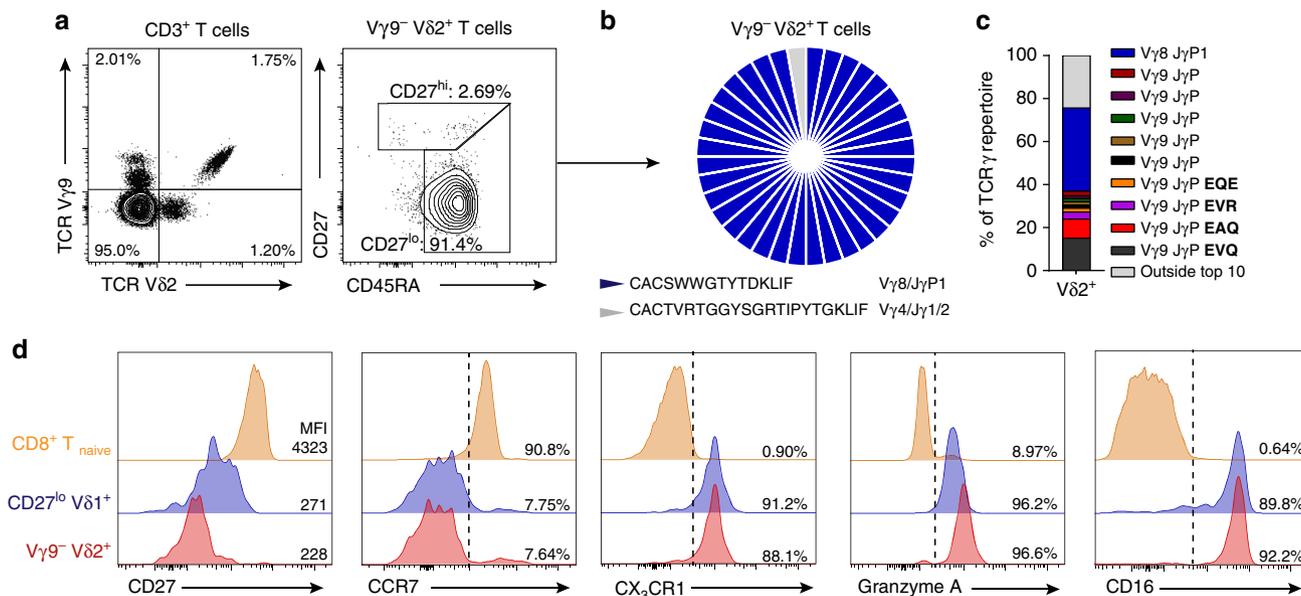


Fig. 4 V γ 9⁻ V δ 2⁺ T cells undergo clonal expansion. **a** Identification of V γ 9⁻ V δ 2⁺ T cells (left) and CD27/CD45RA T-cell memory marker expression on V γ 9⁻ V δ 2⁺ T cells (right) in a peripheral blood sample from “donor X”. **b** Single-cell CDR3 δ sequence analysis and V γ usage by V γ 9⁻ V δ 2⁺ T cells sorted from a peripheral blood sample from “donor X”. **c** TCR γ repertoire analysis of V δ 2⁺ T cells from “donor X”, highlighting the frequency of “donor X’s” V γ 9⁻ V δ 2⁺ T-cell clone (from **b**) within the top 10 most prevalent clonotypes, highlighted are prevalent V γ 9⁺ clonotypes (EVQ, EAQ, EVR and EQE). **d** Analysis of naive and effector T-cell marker expression by V γ 9⁻ V δ 2⁺, CD8⁺ CD27^{hi} CD45RA^{hi} (CD8⁺ T_{naive}) and CD27^{lo/neg} V δ 1⁺ T-cell populations from a peripheral blood sample from “donor X”

sizeable portion of V γ 9⁺ T cells¹¹. Moreover, consistent with a distinct biology for V γ 9 chains in these two contexts, the V γ 9 CDR3 sequences associated with V δ 1 chains did not utilise J γ P sequences and were highly distinct from those associated with V δ 2 chains¹¹. Consistent with expression of TCR γ and TCR δ sequences divergent from the V γ 9⁺ V δ 2⁺ subset, the V γ 9⁻ V δ 2⁺

population was not P-Ag responsive. Thus, V γ 9⁻ V δ 2⁺ T cells represent a universal $\gamma\delta$ subset that is clonotypically, phenotypically and functionally divergent from V γ 9⁺ V δ 2⁺ T cells. Moreover, as evidenced from a single unusual peripheral blood donor and several liver samples, V γ 9⁻ V δ 2⁺ T cells appeared capable of undergoing an adaptive-like programme of highly

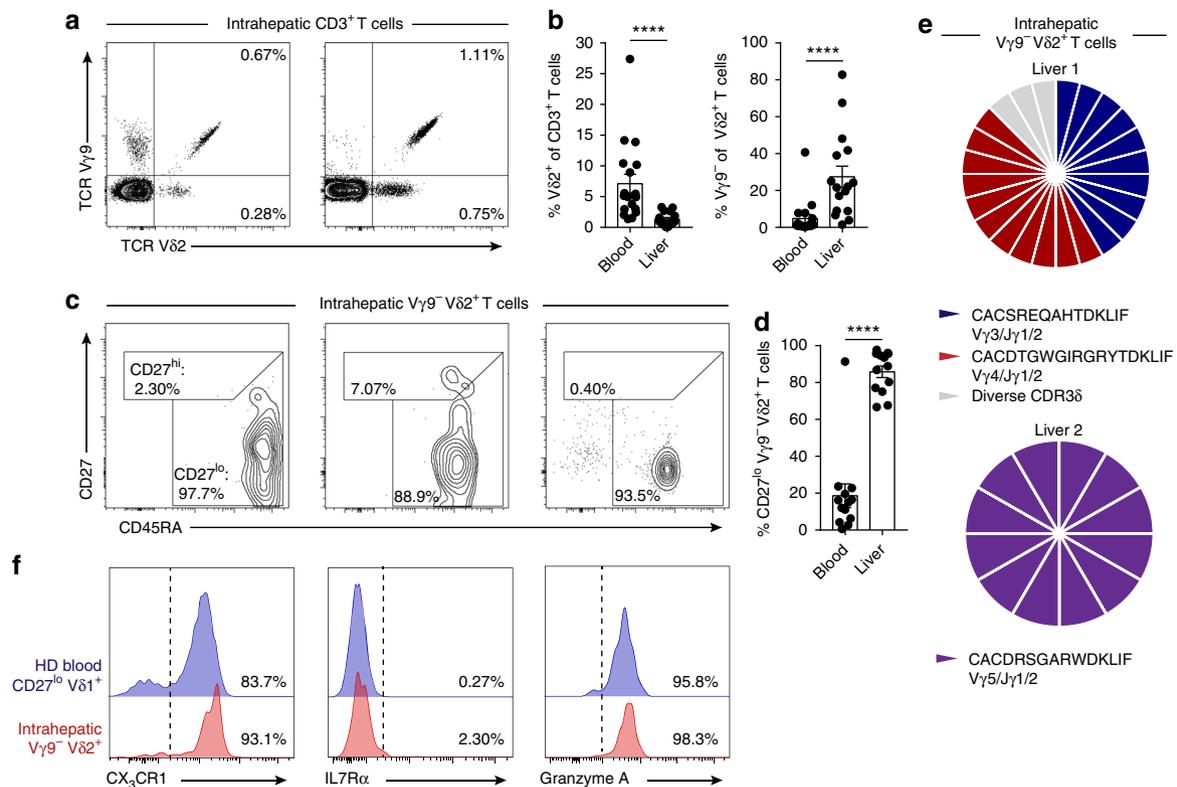


Fig. 5 Clonally expanded $V\gamma 9^- V\delta 2^+$ T cells infiltrate human liver tissue. **a** Representative identification of intrahepatic $V\gamma 9^- V\delta 2^+$ T cells in $CD3^+$ lymphocytes obtained from liver tissue and **b** summary data of the frequency of $V\delta 2^+$ T cells in $CD3^+$ T cells (left) or $V\gamma 9^- V\delta 2^+$ T cells in $V\delta 2^+$ T cells (right) from peripheral blood ($n = 18$) and liver tissue ($n = 16$). **c** Representative intrahepatic $V\gamma 9^- V\delta 2^+$ T-cell expression profiles of CD27 and CD45RA T-cell memory markers and **d** summary data from peripheral blood ($n = 18$) and liver tissue ($n = 16$). **e** Single-cell CDR3 δ sequence analysis and $V\gamma$ usage by intrahepatic $V\gamma 9^- V\delta 2^+$ T cells sorted from two independent liver tissue samples. **f** Representative analysis of CX_3CR1 , $IL7R\alpha$ and Granzyme A expression by intrahepatic $V\gamma 9^- V\delta 2^+$ T cells ($n = 3$) and a healthy donor's (HD) $CD27^{lo/neg} V\delta 1^+$ T-cell population ($n = 8$). Error bars indicate means \pm SEM; **** $P < 0.0001$; p -values were determined by Mann-Whitney t -test (**b, d**)

focused clonotypic expansion and concomitant phenotypic differentiation that closely matches that of $V\delta 1^+$ T cells but is not observed in $V\gamma 9^+ V\delta 2^+$ T cells. This considerably extends the findings of Ravens et al.¹², who detected unusual $V\gamma 9^- V\delta 2^+$ clonotypes in the peripheral blood of one healthy donor and one HSCT recipient.

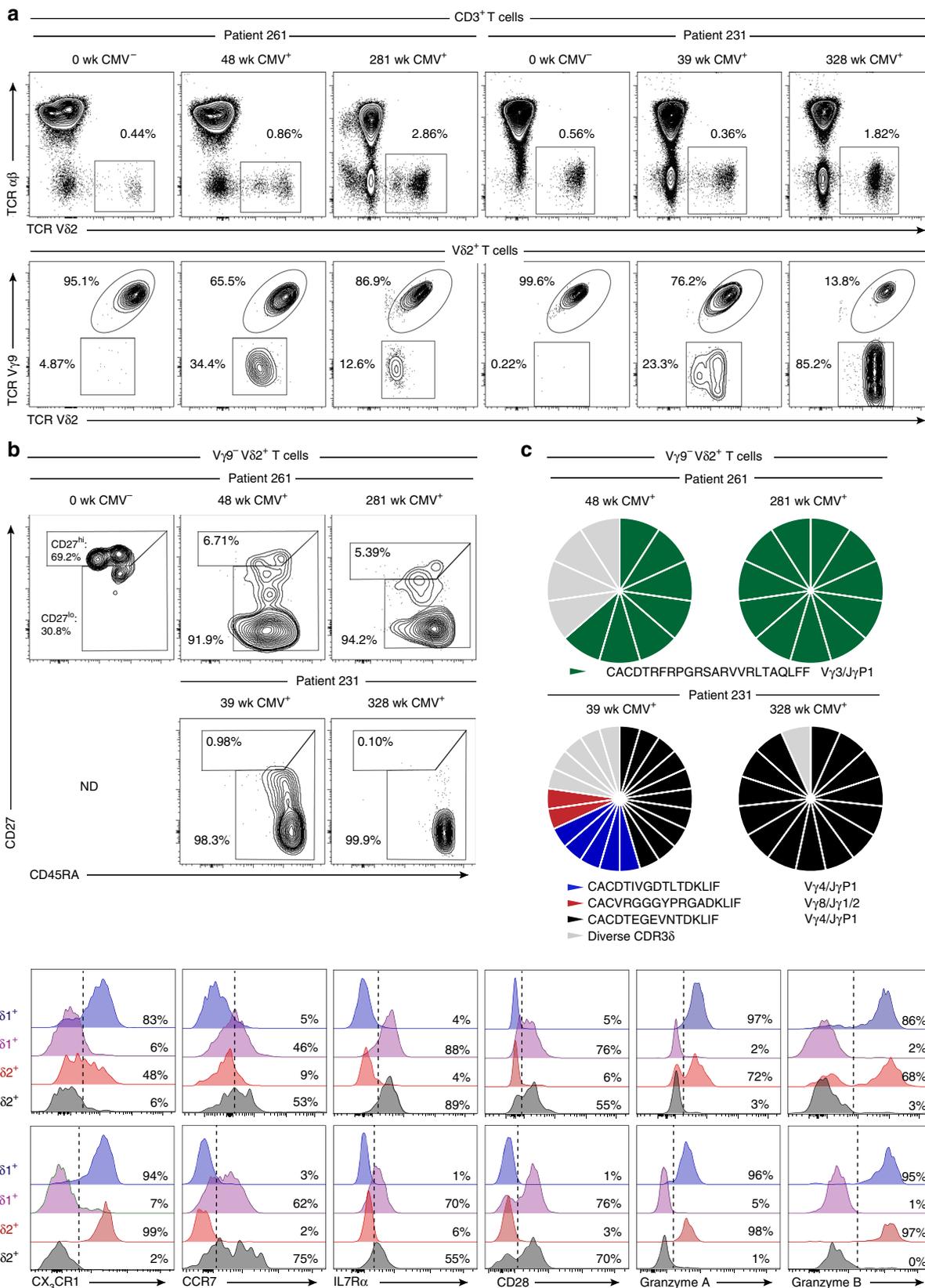
This adaptive biology was confirmed by our analysis of patients following acute CMV infection, which highlighted a striking expansion of the $V\gamma 9^- V\delta 2^+$ subset after infection, linked to the selection of dominant clonotypes, which occurred concomitantly with differentiation from a naive $CD27^{hi}$ to a $CD27^{lo/neg}$ phenotype, and acquisition of CX_3CR1 and cytotoxic granzymes, similar to CMV-specific cytotoxic $CD8^+$ T cells^{46,47}. Interestingly, the single healthy donor who exhibited a clonally expanded $V\gamma 9^- V\delta 2^+$ T-cell subset (“donor X”) was CMV⁺ and unusually, displayed elevated IgG titres, raising the possibility of recent CMV reactivation. Our findings therefore suggest a role for the $V\gamma 9^- V\delta 2^+$ T-cell subset in unconventional immunosurveillance against viral infection, and provide a clear indication that this subset assumes an adaptive immunobiology, highly similar to that suggested for $V\delta 1^+$ T cells^{11,12}. These data suggest clonal selection from a $CCR7^+$, potentially lymphoid tissue homing, naive-like $V\gamma 9^- V\delta 2^+$ T-cell population that expresses a pool of TCRs, enabling amplification of responses to microbial challenges, via differentiation of cells bearing biologically relevant clonotypes to an effector phenotype, resulting in long-lived, functional $\gamma\delta$ T-cell memory. The liver provided an attractive human model in which to examine $V\gamma 9^- V\delta 2^+$ T cells in solid tissues, as it functions as an important site for clearance of both commensals and

pathogens from portal circulation, and has a high proportion of unconventional T cells, such as MAIT cells⁴⁸. Our finding that clonally expanded effector $V\gamma 9^- V\delta 2^+$ T cells can be recruited to the liver suggests this subset may complement the innate-like recognition provided by such hepatic semi-invariant T cells with an unconventional adaptive memory response. Unfortunately, we were unable to obtain CMV serostatus for the patients from whom explanted liver samples were derived, and therefore unable to conclude if the presence of such hepatic $V\gamma 9^- V\delta 2^+$ T cells solely reflected CMV seropositivity or alternatively a wider set of immune challenges. Of relevance, a monoclonally expanded myocytotoxic T-cell population bearing a $V\gamma 3^+ V\delta 2^+$ TCR detected in a patient with polymyositis highlights the potential of the $V\gamma 9^- V\delta 2^+$ T-cell subset to trigger autoimmunity^{49,50}. Further studies are required to define additional microbial and non-microbial challenges that stimulate this unconventional T-cell subset.

In addition to clonal expansion in response to CMV, the finding that peripheral blood $V\gamma 9^- V\delta 2^+$ T cells were TCR responsive but unresponsive to IL-12/IL-18 is another feature that distinguishes them from $V\gamma 9^+ V\delta 2^+$ T cells, and one that again aligns them closely with the $V\delta 1^+$ T-cell subset, which has also been highlighted to display adaptive features^{11,12}. Further studies on such adaptive subsets ($V\delta 1^+$ and $V\gamma 9^- V\delta 2^+$ T cells), ideally ultimately in the context of physiological TCR ligands, are required to explore the full range of functional differences with innate-like subsets. Of note, $V\gamma 9^+ V\delta 2^+$ T cells have previously been demonstrated to exert pleiotropic effector functions in response to cognate P-Ag ligands and different cytokine stimulation⁵¹. Moreover, recent studies in mice highlight that $\gamma\delta$ TCR

triggering of innate-like lymphocytes induced a state of TCR hyporesponsiveness in which T-cell activation could be induced by TCR-extrinsic stimuli⁶. Unfortunately, a systematic comparison of the TCR responsiveness of naive vs. effector V γ 9⁻ V δ 2⁺ T cells within one tissue was not possible, due to the strong dominance of naive V γ 9⁻ V δ 2⁺ T cells in peripheral blood and

effector V γ 9⁻ V δ 2⁺ T cells in liver. However, importantly within the V δ 1⁺ T-cell compartment, CD27^{lo/neg} effectors display a far quicker response to TCR stimulation than CD27^{hi} naive cells¹¹. This observation emphasises the strong distinction between innate-like and adaptive subsets within the $\gamma\delta$ T-cell compartment.



Human V δ ⁺ T cells therefore comprise two separate subsets, which reflect distinct paradigms in human $\gamma\delta$ T-cell biology and also have distinct roles in antimicrobial immunity. V γ 9⁺ V δ ⁺ T cells appear to represent an innate-like subset, most likely originating via gestational selection, which is 'pre-armed' from birth with a semi-invariant TCR repertoire that includes a high level of public V γ TCR chains, permitting polyclonal TCR-mediated responses to microbial P-Ag, as well as to innate cytokines independently of the TCR. However, the V γ 9⁺ V δ ⁺ TCR repertoire diversity may provide the potential for clonotype-specific responses potentially targeted at specific microbial antigenic challenges. In contrast, V γ 9⁻ V δ ⁺ T cells, which are P-Ag-unreactive, represent a previously ill-defined unconventional T-cell compartment that exhibits many of the key hallmarks of an adaptive immunobiology, including clonal expansion and differentiation to effector lymphocytes from an initially naive-like and TCR diverse T-cell pool, including in response to acute viral infection. In doing so, V γ 9⁻ V δ ⁺ T cells most likely contribute to both peripheral blood and tissue immunosurveillance.

Methods

Ethical approval and samples. Peripheral blood samples were obtained from healthy donors who had provided written informed consent for sample collection and subsequent analysis; project approval for this aspect of the study was granted by the NRES Committee West Midlands ethical board (REC reference 14/WM/1254). Samples from patients undergoing renal transplantation were obtained at the Academic Medical Centre, Amsterdam; the medical ethics committee of the Academic Medical Center, Amsterdam, approved this arm of the study and all subjects provided written informed consent in accordance with the Declaration of Helsinki. Longitudinal heparinized blood samples were collected from CMV-seronegative patients who developed primary CMV infection after receiving a renal transplant from a CMV-seropositive donor. As controls, longitudinal PBMC from age matched CMV-seronegative patients who remained CMV-seronegative during the first 5 years following renal transplantation were used. All patients were EBV-seropositive. The course of CMV infection was followed by longitudinal 2-weekly PCR for CMV viral load, and seroconversion was confirmed by detection of CMV-specific IgM and IgG. The patients did not receive CMV prophylaxis, but once the CMV load reached 10⁴ copies/ml, the MMF dose was halved and Valcyte was given. Only patient 1014 experienced CMV-related symptoms: CMV gastritis. All but one patient were treated with a basic immunosuppressive regimen comprising of CD25 mAb (basiliximab) induction treatment, prednisolone, cyclosporine A and MMF. Patient 261, however, was treated with one dose of prednisolone upon transplantation with a kidney from an identical twin, whereafter all immunosuppression was ceased. Liver infiltrating T cells were isolated from explanted diseased human liver tissues from patients undergoing liver transplantation for inflammatory liver diseases, including primary biliary cholangitis, primary sclerosing cholangitis, alcoholic liver disease, and autoimmune hepatitis (Local Research Ethics Committee reference no. 98/CA5192) or normal liver samples from donor liver tissue surplus to clinical requirements (Local Research Ethics Committee reference no. 06/Q2708/11). Umbilical cord blood units were obtained from the Anthony Nolan Cell Therapy Centre Nottingham (ANCTC) under generic tissue bank ethics held by ANCTC and extended to the researchers under a material transfer agreement (MTA).

T-cell isolation, culture and activation. PBMC were isolated from heparinized venous blood by lymphoprep® (Stem Cell Technologies) density gradient centrifugation as per the manufacturer's instructions. For proliferation of T cells, PBMC were labelled with 0.3 μ M carboxyfluorescein succinimidyl ester (CFSE) (eBioscience) and cultured with 10 nM HMB-PP (Sigma) or CD3/CD28 T activator beads (Invitrogen) for 7 days in RPMI-1640 medium (Invitrogen) supplemented with 2 mM L-glutamine, 1% sodium pyruvate, 50 μ g/ml penicillin/streptomycin (Invitrogen) and 10% fetal calf serum (Sigma).

Antibodies and flow cytometry. For total and single-cell sorting of V δ ⁺ populations, PBMC were labelled with anti-CD3 (UCHT1; Biologend; 1:100), TCR $\alpha\beta$ (BW242/412; Miltenyi; 1:200), TCR V δ (123R3; Miltenyi; 1:200), CD27 (M-T271; 1:200), CD45RA (HI100; 1:200); both Biologend, and populations were sorted on an ARIA III Fusion (BD) (Supplementary Fig. 4a). For repertoire analysis, V δ ⁺ T-cell populations were sorted directly into RNeasy (Qiagen). For phenotypic analysis, freshly isolated or frozen PBMC, or cultured cells were labelled with Zombie Aqua viability dye (Biologend), and then subsequently stained (Supplementary Fig. 4b) for cell surface antigens with antibodies directed against CD3 (UCHT1 or HIT3a; 1:100), CD8 (SK1; 1:200), CD45RA (HI100; 1:200), CD27 (M-T271; 1:200), CCR7 (G043H7; 1:50), IL7R α (A019D5; 1:100), CD28 (28.2; 1:100), CX₃CR1 (2A9-1; 1:100), CD16 (3G8; 1:150), CD69 (FN50; 1:100), TCR V δ 2 (B6; 1:100), TCR $\alpha\beta$ (IP26; 1:50); all Biologend. TCR V γ 9 (IMMU360; 1:400); Beckman Coulter. TCR V δ 1 (REA173; 1:200) and TCR V δ 2 (123R3; 1:200); Miltenyi, or α GalCer loaded CD1d dextramers (ProImmune; 1:20). For intracellular staining, after surface antibody staining cells were fixed in IC Fixation buffer (eBioscience) and stained in Permeabilisation Buffer (eBioscience) with antibodies directed against Granzyme A (CBO9; 1:100) and Granzyme B (GB11; 1:100); Biologend). Cells were acquired on an LSR Fortessa X20 (BD) and data analysed with FlowJo V10.2 (TreeStar).

TCR repertoire analysis. RNA was purified from sorted cells (adult V δ ⁺: 25,000 cells; cord blood V δ ⁺: 2,400 - 8,800 cells) using an RNAmicro plus kit (Qiagen) according to the manufacturer's instructions. For high throughput deep sequencing of $\gamma\delta$ TCRs, we used amplicon rescued multiplex (ARM)-PCR and a MiSeq (Illumina) next-generation sequencer (NGS)⁵² to analyse all sorted V δ ⁺ T-cell populations. As detailed in patent WO2009137255A2, a modified version of a protocol devised by Han et al.⁵³ was used, involving initial first-round RT-PCR using high concentrations of gene-specific primers, followed by use of universal primers for the exponential phase of amplification, allowing deep, quantitative and non-biased amplification of TCR γ and TCR δ sequences. All cDNA synthesis, amplification, NGS library preparation and sequencing were performed by iRepertoire, Inc. (Huntsville, USA). We analysed positively sorted $\alpha\beta$ TCR⁻ V δ ⁺ $\gamma\delta$ T cells from 7 healthy donors and 4 umbilical cord blood units (Anthony Nolan Trust, Nottingham).

Single-cell TCR sequencing. PBMC were labelled as above and V δ ⁺ T cells were single cell sorted directly into individual wells in a 96-well plate containing 2 μ l of Superscript VILO cDNA synthesis kit reaction mix (ThermoFisher) containing 0.1% Triton X-100, and incubated according to manufacturer's instructions. TCR γ and TCR δ cDNAs were amplified by two rounds of nested PCR using GoTaq mastermix (Promega) and primers for V δ 2, TCTGGCAGGAGTCAATGT (external) and GAAAGGAGAAGCGATCGGTAAC (internal); for C δ GCAGGA TCAAACCTCTGTATCTTC (external) and TCCTTCACCAGACAAGCGAC (internal); for V γ 1-8 CTGGTACCTACACCAGAGGGGAAGG (external) and TGTGTTGGAATCAGGAVTCAG (internal); for V γ 9 AGAGAGACCTGGTG AAGTCATACA (external) and GGTGGATAGGATACCTGAAACG (internal) and for C γ CTGACGATACATCTGTGTTCTTTG (external) and AATCGTGTG TCTCTTCTTTTCTT (internal). PCR products were separated on 1.2% agarose gels, and products of successful reactions were incubated with ExoSAP-IT PCR cleanup enzyme (Affymetrix) before sequencing with BigDye Terminator v3.1 (Applied Biosystems) following manufacturer's instructions and running on an ABI 3730 capillary sequencer (Functional Genomics Facility, University of Birmingham).

TCR repertoire data analysis. V, D and J gene usage and CDR3 sequences were identified and assigned and tree maps generated using iRweb tools (iRepertoire, Inc, Huntsville, AL, USA)⁵⁴. Tree maps show each unique CDR3 as a coloured rectangle, the size of each rectangle corresponds to each CDR3s abundance within the repertoire and the positioning is determined by the V region usage. To determine the 10 most prevalent shared CDR3 γ 9 sequences, the first 10 most dominant sequences by frequency were filtered from each donor's CDR3 γ 9 protein lists. These sequences were then ordered by frequency for sequences that were shared between >2 donors (to create a hierarchy of 10 common sequences) and uncommon highly amplified sequences were included in the final analysis. For more detailed analysis of the TCR repertoire, datasets were processed using the MiXCR software package⁵⁵ to further correct for PCR and sequencing errors. Diversity metrics, clonotype overlap and gene usage were plotted in R, by VDJTools⁵⁶.

Fig. 6 V γ 9⁻ V δ ⁺ T cells exhibit TCR-specific clonal expansion in response to viral infection. **a** Identification of V δ ⁺ T cells in CD3⁺ T cells (top row) and V γ 9⁻ cells within V δ ⁺ T cells (bottom row) in longitudinal peripheral blood samples from two CMV-seronegative kidney transplant patients developing post-operative acute CMV infection (CMV-seroconversion: patient 231, 8 weeks; patient 261, 7 weeks). **b** CD27 and CD45RA T-cell memory marker expression profiles by detectable V γ 9⁻ V δ ⁺ T cells populations in longitudinal peripheral blood samples from patients 231 and 261. Cytometry data from patient 261 at 281 weeks was acquired on an alternative flow cytometer (**a** and **b**). **c** Single-cell CDR3 δ sequence analysis and V γ usage by V γ 9⁻ V δ ⁺ T cells sorted from patients 231 and 261 peripheral blood samples between 39 and 328 weeks after transplantation. **d** Analysis of the indicated naive and effector T-cell markers by V γ 9⁻ V δ ⁺ T cells from patients 261 (week 48) and 231 (week 328) and in comparison with CD27^{hi}, CD27^{lo/neg} V δ 1⁺ and V γ 9⁻ V δ ⁺ T cells from the peripheral blood of two healthy donors analysed in parallel. ND no detectable population

TCR sequence analyses. The CDR3 sequence was defined as the amino acids between the second Cysteine of the V region and the conserved Phenylalanine of the J region, according to IMGT. CDR3 sequences shown in tables include the conserved Cysteine and Phenylalanine, but only the amino acids between these residues are counted for CDR3 length analysis and for analysis of residues at position 5. N and P nucleotides were identified using the IMGT Junction Analysis tool^{57,58}. Neonatal V δ 2 sequence logos were generated on the Seq2Logo server⁵⁹ in Shannon format without the use of pseudo counts, and give a visual representation of amino acids enriched at different positions in the observed CDR3 δ 2 sequences. The different amino acids are coloured according to physicochemical properties (acidic (DE), red; basic (RKH), blue; hydrophobic (ACFILMPVW), black; and neutral (NSGTY), green). For TCR δ 2 sequences, the ten most abundant clonotypes of 14–16 amino acids using either J δ 3 or J δ 1 from each donor were aligned using Clustal Omega52 with default parameters, before logo generation. Narrower bars in the sequence logo correspond to gaps in the sequences.

Statistical analysis. Each data set was assessed for normality using Shapiro–Wilk normality test. Differences between columns were analysed by two-tailed Student's *t*-tests for normally distributed data and Mann–Whitney for non-parametric data. Differences between groups were analysed using one-way ANOVA with Dunnett's or Tukey's post tests for normally distributed data or with Kruskal–Wallis with Tukey's post tests for non-parametric data and RM two-way ANOVA with Tukey's post tests was used when comparing groups with independent variables. **P* < 0.05, ***P* < 0.01, ****P* < 0.001 and *****P* < 0.0001. Correlation was assessed for non-parametric data by Spearman correlation. Tabulated data were analysed in Graphpad PRISM 7 (Graphpad Software, Inc.).

Data availability. All data are available from the authors upon request. The sequence data that support the findings of this study have been deposited in the NIH NCBI sequence read archive (SRA) database with the primary accession code SRP113556 and SRP096009, for V δ 2⁺ and V δ 1⁺ TCR repertoires respectively.

Received: 16 September 2017 Accepted: 3 April 2018

Published online: 02 May 2018

References

- Rast, J. P. et al. alpha, beta, gamma, and delta T cell antigen receptor genes arose early in vertebrate phylogeny. *Immunity* **6**, 1–11 (1997).
- Zheng, J., Liu, Y., Lau, Y. L. & Tu, W. gammadelta-T cells: an unpolished sword in human anti-infection immunity. *Cell. Mol. Immunol.* **10**, 50–57 (2013).
- Silva-Santos, B., Serre, K. & Norell, H. gammadelta T cells in cancer. *Nat. Rev. Immunol.* **15**, 683–691 (2015).
- Godfrey, D. I., Uldrich, A. P., McCluskey, J., Rossjohn, J. & Moody, D. B. The burgeoning family of unconventional T cells. *Nat. Immunol.* **16**, 1114–1123 (2015).
- Kisielow, J., Tortola, L., Weber, J., Karjalainen, K. & Kopf, M. Evidence for the divergence of innate and adaptive T-cell precursors before commitment to the alphabeta and gammadelta lineages. *Blood* **118**, 6591–6600 (2011).
- Wencker, M. et al. Innate-like T cells straddle innate and adaptive immunity by altering antigen-receptor responsiveness. *Nat. Immunol.* **15**, 80–87 (2014).
- Halary, F. et al. Shared reactivity of V{delta}2(neg) {gamma}{delta} T cells against cytomegalovirus-infected cells and tumor intestinal epithelial cells. *J. Exp. Med.* **201**, 1567–1578 (2005).
- Willcox, C. R. et al. Cytomegalovirus and tumor stress surveillance by binding of a human gammadelta T cell antigen receptor to endothelial protein C receptor. *Nat. Immunol.* **13**, 872–879 (2012).
- Hudspeth, K. et al. Engagement of NKp30 on Vdelta1 T cells induces the production of CCL3, CCL4, and CCL5 and suppresses HIV-1 replication. *Blood* **119**, 4013–4016 (2012).
- Correia, D. V. et al. Differentiation of human peripheral blood Vdelta1 +T cells expressing the natural cytotoxicity receptor NKp30 for recognition of lymphoid leukemia cells. *Blood* **118**, 992–1001 (2011).
- Davey, M. S. et al. Clonal selection in the human Vdelta1 T cell repertoire indicates gammadelta TCR-dependent adaptive immune surveillance. *Nat. Commun.* **8**, 14760 (2017).
- Ravens, S. et al. Human gammadelta T cells are quickly reconstituted after stem-cell transplantation and show adaptive clonal expansion in response to viral infection. *Nat. Immunol.* **18**, 393–401 (2017).
- Morita, C. T., Jin, C., Sarikonda, G. & Wang, H. Nonpeptide antigens, presentation mechanisms, and immunological memory of human Vgamma2Vdelta2 T cells: discriminating friend from foe through the recognition of prenyl pyrophosphate antigens. *Immunol. Rev.* **215**, 59–76 (2007).
- Riganti, C., Massaia, M., Davey, M. S. & Eberl, M. Human gammadelta T-cell responses in infection and immunotherapy: common mechanisms, common mediators? *Eur. J. Immunol.* **42**, 1668–1676 (2012).
- Rhodes, D. A. et al. Activation of human gammadelta T cells by cytosolic interactions of BTN3A1 with soluble phosphoantigens and the cytoskeletal adaptor periplakin. *J. Immunol.* **194**, 2390–2398 (2015).
- Sandstrom, A. et al. The intracellular B30.2 domain of butyrophilin 3A1 binds phosphoantigens to mediate activation of human Vgamma9Vdelta2 T cells. *Immunity* **40**, 490–500 (2014).
- Harly, C. et al. Key implication of CD277/butyrophilin-3 (BTN3A) in cellular stress sensing by a major human gammadelta T-cell subset. *Blood* **120**, 2269–2279 (2012).
- Costa, G. et al. Control of Plasmodium falciparum erythrocytic cycle: gammadelta T cells target the red blood cell-invasive merozoites. *Blood* **118**, 6952–6962 (2011).
- Davey, M. S. et al. Microbe-specific unconventional T cells induce human neutrophil differentiation into antigen cross-presenting cells. *J. Immunol.* **193**, 3704–3716 (2014).
- Tyler, C. J., Doherty, D. G., Moser, B. & Eberl, M. Human Vgamma9/Vdelta2 T cells: Innate adaptors of the immune system. *Cell. Immunol.* **296**, 10–21 (2015).
- Davodeau, F. et al. Peripheral selection of antigen receptor junctional features in a major human gamma delta subset. *Eur. J. Immunol.* **23**, 804–808 (1993).
- Parker, C. M. et al. Evidence for extrathymic changes in the T cell receptor gamma/delta repertoire. *J. Exp. Med.* **171**, 1597–1612 (1990).
- Carding, S. R. & Egan, P. J. Gammadelta T cells: functional plasticity and heterogeneity. *Nat. Rev. Immunol.* **2**, 336–345 (2002).
- Pauza, C. D. & Cairo, C. Evolution and function of the TCR Vgamma9 chain repertoire: It's good to be public. *Cell Immunol.* **296**, 22–30 (2015).
- Dimova, T. et al. Effector Vgamma9Vdelta2 T cells dominate the human fetal gammadelta T-cell repertoire. *Proc. Natl Acad. Sci. USA* **112**, E556–E565 (2015).
- De Rosa, S. C. et al. Ontogeny of gamma delta T cells in humans. *J. Immunol.* **172**, 1637–1645 (2004).
- McVay, L. D., Carding, S. R., Bottomly, K. & Hayday, A. C. Regulated expression and structure of T cell receptor gamma/delta transcripts in human thymic ontogeny. *EMBO J.* **10**, 83–91 (1991).
- Tamura, N. et al. Diversity in junctional sequences associated with the common human V gamma 9 and V delta 2 gene segments in normal blood and lung compared with the limited diversity in a granulomatous disease. *J. Exp. Med.* **172**, 169–181 (1990).
- Wang, H., Fang, Z. & Morita, C. T. Vgamma2Vdelta2 T Cell Receptor recognition of prenyl pyrophosphates is dependent on all CDRs. *J. Immunol.* **184**, 6209–6222 (2010).
- McVay, L. D., Jaswal, S. S., Kennedy, C., Hayday, A. & Carding, S. R. The generation of human gammadelta T cell repertoires during fetal development. *J. Immunol.* **160**, 5851–5860 (1998).
- Sherwood, A. M. et al. Deep sequencing of the human TCRgamma and TCRbeta repertoires suggests that TCRbeta rearranges after alphabeta and gammadelta T cell commitment. *Sci. Transl. Med.* **3**, 90ra61 (2011).
- Venturi, V., Price, D. A., Douek, D. C. & Davenport, M. P. The molecular basis for public T-cell responses? *Nat. Rev. Immunol.* **8**, 231–238 (2008).
- Ryan, P. L. et al. Heterogeneous yet stable Vdelta2(+) T-cell profiles define distinct cytotoxic effector potentials in healthy human individuals. *Proc. Natl Acad. Sci. USA* **113**, 14378–14383 (2016).
- Vermijlen, D. et al. Human cytomegalovirus elicits fetal gammadelta T cell responses in utero. *J. Exp. Med.* **207**, 807–821 (2010).
- Chien, Y. H., Meyer, C. & Bonneville, M. gammadelta T cells: first line of defense and beyond. *Annu. Rev. Immunol.* **32**, 121–155 (2014).
- Bonneville, M., O'Brien, R. L. & Born, W. K. Gammadelta T cell effector functions: a blend of innate programming and acquired plasticity. *Nat. Rev. Immunol.* **10**, 467–478 (2010).
- Pitard, V. et al. Long-term expansion of effector/memory Vdelta2-gammadelta T cells is a specific blood signature of CMV infection. *Blood* **112**, 1317–1324 (2008).
- Klarenbeek, P. L. et al. Deep sequencing of antiviral T-cell responses to HCMV and EBV in humans reveals a stable repertoire that is maintained for many years. *PLoS Pathog.* **8**, e1002889 (2012).
- Palakodeti, A. et al. The molecular basis for modulation of human Vgamma9Vdelta2 T cell responses by CD277/butyrophilin-3 (BTN3A)-specific antibodies. *J. Biol. Chem.* **287**, 32780–32790 (2012).
- Salim, M. et al. BTN3A1 Discriminates gammadelta T cell phosphoantigens from nonantigenic small molecules via a conformational sensor in Its B30.2 domain. *ACS Chem. Biol.* **12**, 2631–2643 (2017).
- McVay, L. D., Hayday, A. C., Bottomly, K. & Carding, S. R. Thymic and extrathymic development of human gamma/delta T cells. *Curr. Top. Microbiol. Immunol.* **173**, 57–63 (1991).

42. McVay, L. D. & Carding, S. R. Extrathymic origin of human gamma delta T cells during fetal development. *J. Immunol.* **157**, 2873–2882 (1996).
43. Koay, H. F. et al. A three-stage intrathymic development pathway for the mucosal-associated invariant T cell lineage. *Nat. Immunol.* **17**, 1300–1311 (2016).
44. Keller, A. N. et al. Drugs and drug-like molecules can modulate the function of mucosal-associated invariant T cells. *Nat. Immunol.* **18**, 402–411 (2017).
45. Gold, M. C. et al. MR1-restricted MAIT cells display ligand discrimination and pathogen selectivity through distinct T cell receptor usage. *J. Exp. Med.* **211**, 1601–1610 (2014).
46. Remmerswaal, E. B. et al. Human virus-specific effector-type T cells accumulate in blood but not in lymph nodes. *Blood* **119**, 1702–1712 (2012).
47. Hertoghs, K. M. et al. Molecular profiling of cytomegalovirus-induced human CD8+T cell differentiation. *J. Clin. Invest.* **120**, 4077–4090 (2010).
48. Jeffery, H. C. et al. Biliary epithelium and liver B cells exposed to bacteria activate intrahepatic MAIT cells through MR1. *J. Hepatol.* **64**, 1118–1127 (2016).
49. Pluschke, G., Ruegg, D., Hohlfeld, R. & Engel, A. G. Autoaggressive myocytotoxic T lymphocytes expressing an unusual gamma/delta T cell receptor. *J. Exp. Med.* **176**, 1785–1789 (1992).
50. Hohlfeld, R., Engel, A. G., Ii, K. & Harper, M. C. Polymyositis mediated by T lymphocytes that express the gamma/delta receptor. *N. Engl. J. Med.* **324**, 877–881 (1991).
51. Vermijlen, D. et al. Distinct cytokine-driven responses of activated blood gammadelta T cells: insights into unconventional T cell pleiotropy. *J. Immunol.* **178**, 4304–4314 (2007).
52. Wang, C. et al. High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proc. Natl Acad. Sci. USA* **107**, 1518–1523 (2010).
53. Han, J. et al. Simultaneous amplification and identification of 25 human papillomavirus types with Tempex technology. *J. Clin. Microbiol.* **44**, 4157–4162 (2006).
54. Yang, Y. et al. Distinct mechanisms define murine B cell lineage immunoglobulin heavy chain (IgH) repertoires. *Elife* **4**, e09083 (2015).
55. Bolotin, D. A. et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380–381 (2015).
56. Shugay, M. et al. VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput. Biol.* **11**, e1004503 (2015).
57. Giudicelli, V. & Lefranc, M. P. IMGT/junction analysis: IMGT standardized analysis of the V-J and V-D-J junctions of the rearranged immunoglobulins (IG) and T cell receptors (TR). *Cold Spring Harb. Protoc.* **2011**, 716–725 (2011).
58. Yousfi Monod, M., Giudicelli, V., Chaume, D. & Lefranc, M. P. IMGT/Junction analysis: the first tool for the analysis of the immunoglobulin and T cell receptor complex V-J and V-D-J junctions. *Bioinformatics* **20**, i379–i385 (2004).
59. Thomsen, M. C. & Nielsen, M. Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Res.* **40**, W281–W287 (2012).

Acknowledgements

We thank all donors and patients who participated in this study, clinical staff at UHB NHS Foundation Trust for recruitment and provision of liver samples, the AMC biobank staff for provision of renal transplant patient samples and the Anthony Nolan Cell Therapy Centre for provision of cord blood samples. We also thank Dr Matthew McKenzie and the University of Birmingham CMD5 Cell Sorting Facility for isolation of $\gamma\delta$ T cells and the University of Birmingham Protein Expression Facility for use of their facilities. The work was supported by a Medical Research Council Clinician Scientist award (G1002552 to Y.H.O.), Russian Foundation for Basic Research grant (17-04-01994 to S.A.K. and 17-54-10018 KO to D.M.C.), Ministry of Education, Youth and Sports of the Czech Republic under the project CEITEC 2020 (LQ1601 to D.M.C.) and Wellcome Trust Investigator award funding, supporting M.S.D., C.R.W., F.M. and M.S. (099266/Z/12/Z to B.E.W.).

Author contributions

M.S.D., C.R.W., B.E.W. conceived and designed experiments; M.S.D. and C.R.W. performed experiments and analysed data; S.A.K. and D.M.C. analysed deep sequencing data; S.H. and Y.H.O. provided liver tissue samples and data analysis; E.B.R. and F.J.B. provided kidney transplant patient samples and critical discussions; M.S. and F.M. provided assistance with single-cell TCR sequencing and critical discussions; B.E.W. supervised the project; M.S.D., C.R.W. and B.E.W. wrote the manuscript and all authors provided critical review.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-018-04076-0>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

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

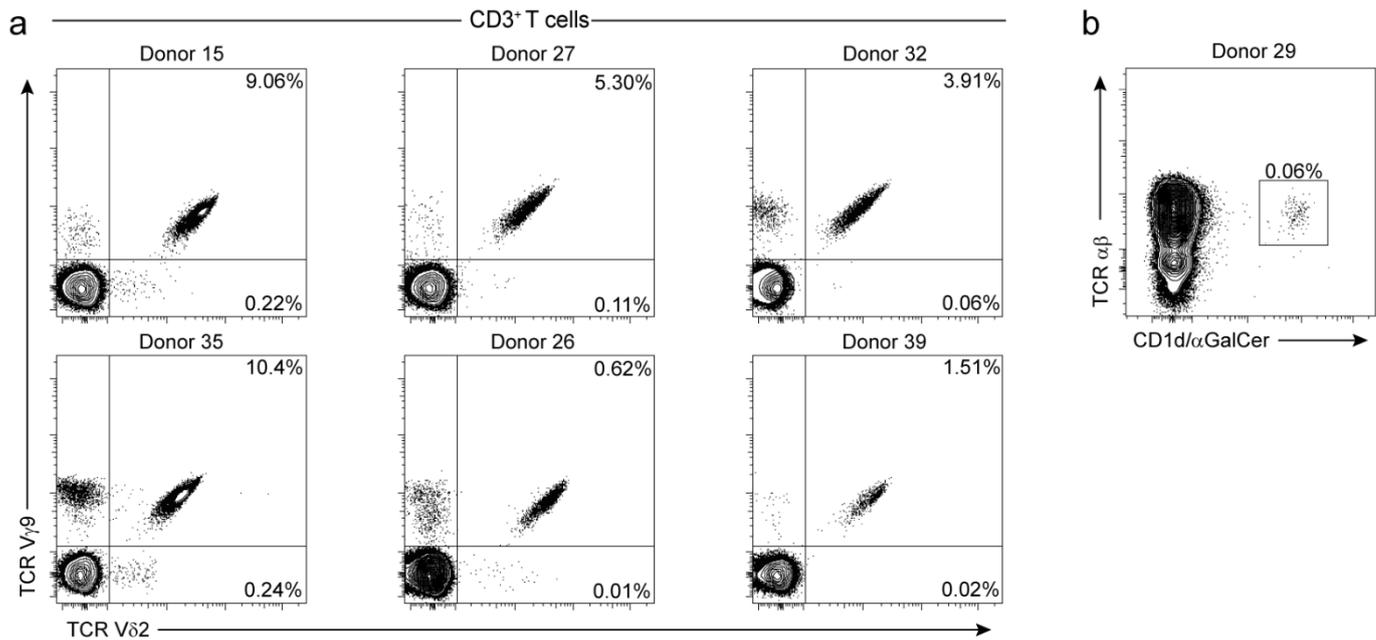


Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative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

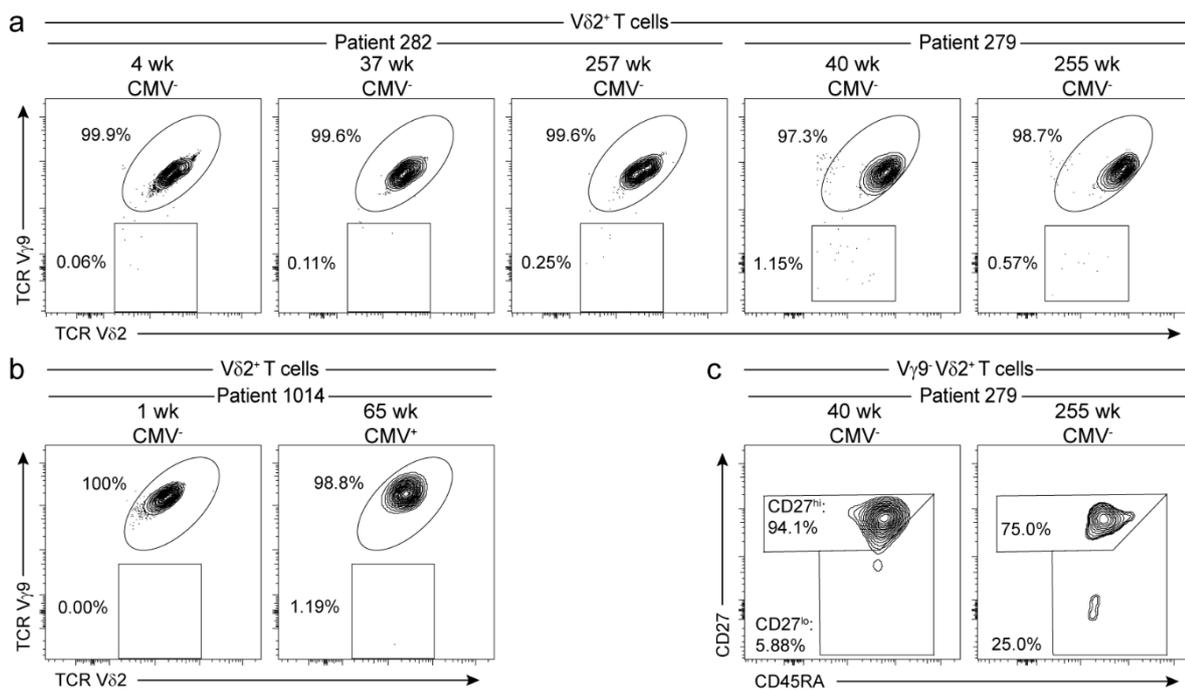
© The Author(s) 2018

The human V δ 2⁺ T cell compartment comprises distinct innate-like V γ 9⁺ and adaptive V γ 9⁻ subsets

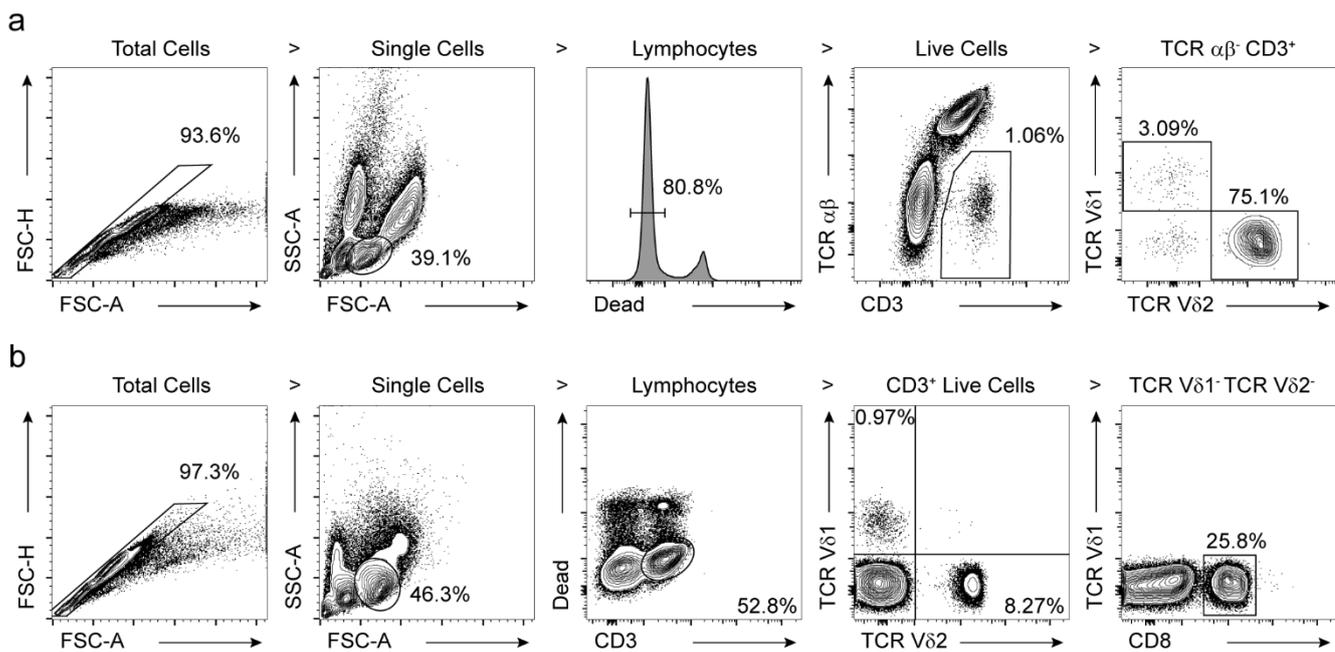
Martin S. Davey, Carrie R. Willcox, Stuart Hunter, Sofya A. Kasatskaya, Ester B. Remmerswaal, Mahboob Salim, Fiyaz Mohammed, Frederike J. Bemelman, Dmitriy M. Chudakov, Ye H. Oo and Benjamin E. Willcox



Supplementary Figure 2. Identification of V γ 9⁺ V δ 2⁺ T cells and NKT cells in healthy adults. (a) Representative identification of V γ 9⁺ V δ 2⁺ T cells by monoclonal antibodies directed against TCR V γ 9 and TCR V δ 2 in CD3⁺ T cells from adult peripheral blood samples (n=18). (b) Representative identification of NKT cells with an α GalCer loaded CD1d-dextramer in CD3⁺ T cells from adult peripheral blood samples (n=5).



Supplementary Figure 3. $V\gamma 9^+ V\delta 2^+$ T cells in kidney transplant patients. (a) Identification of $V\gamma 9^+$ cells within total $V\delta 2^+$ T cells in two CMV-seronegative kidney transplant patients who did not develop post-operative acute-CMV infection and (b) a patient who was infected with CMV and seroconverted at 5 weeks post-transplantation. (c) Analysis of CD27 and CD45RA T cell memory marker expression on detectable $V\gamma 9^+ V\delta 2^+$ T cells populations from CMV-seronegative patient 279 at 40 and 255 weeks post-kidney transplant.



Supplementary Figure 4. Gating strategy for identification and sorting of T cell populations. (a) Representative flow cytometry plots show the gating strategy used to identify single cells>lymphocytes>live cells>CD3⁺ TCR $\alpha\beta$ ⁺ T cells>V δ 2⁺ populations for cell sorting (related to Figure 1, 2 and Supplementary Figure 1). (b) Representative identification of single cells>lymphocytes>live cells CD3⁺>V δ 1⁺, V δ 2⁺ and V δ 1⁻ V δ 2⁻ CD8⁺ T cells from peripheral blood and liver tissue samples (related to Figure 2, 3, 4, 5, 6, Supplementary Figure 2 and 3).

Cord Blood 1	%	Cord Blood 4	%	Cord Blood 5	%	Cord Blood 6	%
CACDWGSSWDTRQMFF	6.2	CACDWGSSWDTRQMFF	8.2	CACDILGDTDKLIF	5.4	CACDSLGDTSPPDKLIF	11.6
CACDILGDTDKLIF	3.1	CACDTGGYSWDTRQMFF	3.4	CACDKLGDTDKLIF	4.1	CACDRGIRRSWDTRQMFF	10.4
CACDVLGDTDKLIF	1.5	CACDILGDTDKLIF	1.9	CACDTVLGDTWDTRQMFF	3.0	CACDGKSTSSWDTRQMFF	9.7
CACDVLGDTAQLFF	1.5	CACDTVLGDTWDTRQMFF	1.4	CACDWGSSWDTRQMFF	2.5	CACDTVYWGIRSSWDTRQMFF	9.7
CACDRGYTDKLIF	1.2	CACDVLGDTDKLIF	1.3	CACDVLGDRHDKLIF	2.2	CACDSLGDTSPPDKLIF	9.7
CACDVLGDLTAQLFF	1.1	CACDYWGSSWDTRQMFF	1.1	CACDVLGDTDKLIF	2.1	CACDTGGRWGIRLWDTRQMFF	9.1
CACDGILTAQLFF	0.9	CACDTWGSSWDTRQMFF	0.9	CACDNTGGSSWDTRQMFF	1.8	CACDTATPLESGGYEVGTDKLIF	7.5
CACDTWGYTDKLIF	0.9	CACDILGDTWDTRQMFF	0.6	CACDTVIGGIRPYTDKLIF	1.7	CACDTGKWDTRQMFF	7.3
CACDSGIWTAQLFF	0.9	CACDWGTWDTRQMFF	0.6	CACDTGGYWDTRQMFF	1.6	CACDTGDTLWDTRQMFF	3.8
CACDILGDTWDTRQMFF	0.8	CACDTWGTDKLIF	0.6	CACDTLGVLTKLIF	1.5	CACDTGWSSWDTRQMFF	3.3

Supplementary Table 1. Top 10 most prevalent TCR δ sequences in cord blood V δ 2⁺ T cells. The table depicts CDR3 δ 2 amino acid sequence and frequency occupied in the total TCR δ repertoire from four cord blood donors.

CDR3 δ 2	CDR3 length	TRDD	TRDJ	P nt	N nt	CB01	CB04	CB05	CB06	Frequency
CACDWGSSWDTRQMFF	14	3	3	1	0	■	■	■		>5%
CACDTAGGYSWDTRQMFF	16	3	3	1	0		■		■	2.5-5%
CACDTWDTRQMFF	11	--	3	0	0		■	■		1-2.5%
CACDWGTWDTRQMFF	13	3	3	0	0		■		■	<1%
CACDTGDTLWDTRQMFF	15	3	3	1	1		■		■	ND
CACDTVLGDTWDTRQMFF	16	3	3	2	0	■	■	■		
CACDNTGGSSWDTRQMFF	15	3	3	1	1	■	■	■		
CACDTGGYWDTRQMFF	14	3	3	0	0		■	■		
CACDTWGSSWDTRQMFF	15	3	3	1	0		■			
CACDYWGSSWDTRQMFF	15	3	3	2	0	■	■	■		
CACDTGGYSWDTRQMFF	15	3	3	0	0	■	■	■		
CACDTAGGSSWDTRQMFF	15	3	3	1	0		■	■		
CACDTAGSSWDTRQMFF	15	3	3	2	1		■	■		
CACDTVGGSSWDTRQMFF	16	3	3	3	2	■	■	■		
CACDTYWGSSWDTRQMFF	16	3	3	2	0	■	■	■		
CACDTGGYSSWDTRQMFF	16	3	3	2	0	■	■	■		
CACDILGDTDKLIF	12	3	1	1	0	■	■	■		
CACDVLGDTDKLIF	12	3	1	2	0	■	■	■		
CACDTLGDTDKLIF	12	3	1	0	0	■	■	■		

Supplementary Table 2. Shared TCR δ sequences in cord blood V δ 2⁺ T cells. The table depicts public CDR3 δ 2 sequences, amino acid length, D δ , and J δ gene segments used, N-nucleotide, P-nucleotide addition and relative frequency within each TCR δ repertoire. Sequences were analysed using IMGT Junction Analysis and are from four cord blood donors. ND = not detected.

Chapter V

Human liver infiltrating $\gamma\delta$ T cells are composed of
clonally expanded circulating
and tissue-resident populations

Introduction

$\gamma\delta$ T cells make up a significant part of the resident lymphocytes of peripheral tissues in both mice and men; however, human gamma/delta T cells make up to only 1-5% of T lymphocytes in blood circulation. Chapter V addresses the question of $\gamma\delta$ T cell functional status in liver tissue. We studied liver samples both from patients with chronic non-carcinogenic liver diseases and liver donor samples. In the liver of both healthy donors and patients, V δ 2- $\gamma\delta$ T cells were predominant. Liver-resident $\gamma\delta$ T cells displayed higher clonal expansions than circulating $\gamma\delta$ T cells.

Unexpectedly, we found that 3-10 particularly dominant memory cell clones can fill up to 75% of repertoire volume. Naive T cells are practically absent in the liver tissue, and memory T cells contain both subsets similar to the circulating memory subset and a unique memory subset with resident cell markers (effector phenotype CD45RA^{lo}CD69⁺CXCR3⁺CXCR6⁺). The liver tissue example shows that tissue-resident gamma/delta TCR repertoires have low diversity and extremely high clonality.

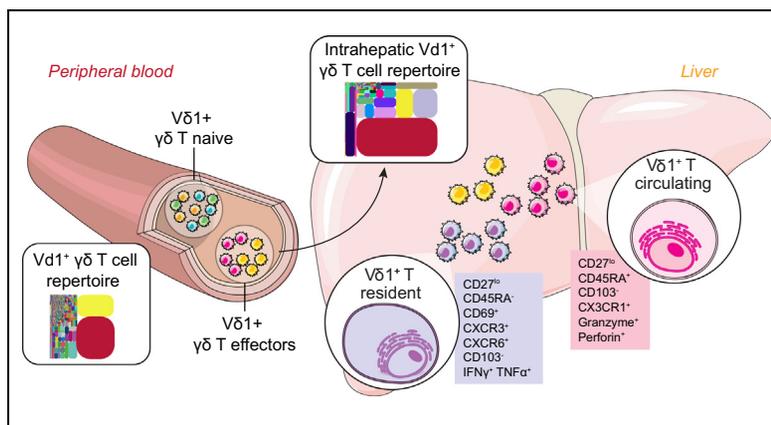
Contribution

We studied the diversity and properties of the T-cell receptor repertoire along with the liver-resident $\gamma\delta$ T-cells phenotype. Collaborators from the University of Birmingham performed immunophenotyping and described the functionality of $\gamma\delta$ infiltrating liver T cells with flow cytometry, in situ hybridization, and immunohistochemistry. TCR sequencing was outsourced to iRepertoire.

I performed TCR sequencing data analysis. The data from tissue samples required a preprocessing step to deal with cross-contamination before comparison of paired samples. I performed multiple comparisons of diversity estimation metrics and proposed the strategy for the characterization of highly expanded T cell populations in tissue samples. TCR repertoire data allowed to describe the difference in clonal composition between paired blood and liver samples. Here for the first time, we characterized clonality and the unique TCR repertoire features of resident $\gamma\delta$ T cells from peripheral tissues.

Human liver infiltrating $\gamma\delta$ T cells are composed of clonally expanded circulating and tissue-resident populations

Graphical abstract



Highlights

- Intrahepatic $V\delta 2^{neg}$ $\gamma\delta$ T cells are clonally focussed and feature private TCR rearrangements.
- Effector $CD27^{lo/neg}$ $V\delta 1^{+}$ T cells are enriched in liver, but naïve $CD27^{hi}$ cells are absent.
- A subset of $V\delta 1^{+}$ T cells is distinct from those in blood and may be liver tissue resident.
- Liver $V\delta 1^{+}$ $\gamma\delta$ T cells are polyfunctional and respond to both TCR and innate stimuli.

Authors

Stuart Hunter, Carrie R. Willcox, Martin S. Davey, ..., Dmitriy M. Chudakov, Ye H. Oo, Benjamin E. Willcox

Correspondence

y.h.oo@bham.ac.uk (Y.H. Oo),
b.willcox@bham.ac.uk (B.E. Willcox)

Lay summary

$\gamma\delta$ T cells are frequently enriched in many solid tissues, however the immunobiology of such tissue-associated subsets in humans has remained unclear. We show that intrahepatic $\gamma\delta$ T cells are enriched for clonally expanded effector T cells, whereas naïve $\gamma\delta$ T cells are largely excluded. Moreover, whereas a distinct proportion of circulating T cell clonotypes was present in both the liver tissue and peripheral blood, a functionally and clonotypically distinct population of liver-resident $\gamma\delta$ T cells was also evident. Our findings suggest that factors triggering $\gamma\delta$ T cell clonal selection and differentiation, such as infection, can drive enrichment of $\gamma\delta$ T cells into liver tissue, allowing the development of functionally distinct tissue-restricted memory populations specialised in local hepatic immunosurveillance.



Human liver infiltrating $\gamma\delta$ T cells are composed of clonally expanded circulating and tissue-resident populations

Stuart Hunter^{1,2,†}, Carrie R. Willcox^{1,†}, Martin S. Davey^{1,†}, Sofya A. Kasatskaya^{3,4}, Hannah C. Jeffery², Dmitriy M. Chudakov^{3,4,5,6}, Ye H. Oo^{2,7,*†}, Benjamin E. Willcox^{1,*†}

¹Cancer Immunology and Immunotherapy Centre, Institute of Immunology and Immunotherapy, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom; ²Centre for Liver Research and National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, Institute of Immunology & Immunotherapy, University of Birmingham, United Kingdom; ³Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Science, Moscow, Russia; ⁴Skolkovo Institute of Science and Technology, Moscow, Russia; ⁵Central European Institute of Technology, Masaryk University, Brno, Czech Republic; ⁶Pirogov Russian National Research Medical University, Moscow, Russia; ⁷University Hospital of Birmingham NHS Foundation Trust, United Kingdom

Background & Aims: $\gamma\delta$ T cells comprise a substantial proportion of tissue-associated lymphocytes. However, our current understanding of human $\gamma\delta$ T cells is primarily based on peripheral blood subsets, while the immunobiology of tissue-associated subsets remains largely unclear. Therefore, we aimed to elucidate the T cell receptor (TCR) diversity, immunophenotype and function of $\gamma\delta$ T cells in the human liver.

Methods: We characterised the TCR repertoire, immunophenotype and function of human liver infiltrating $\gamma\delta$ T cells, by TCR sequencing analysis, flow cytometry, *in situ* hybridisation and immunohistochemistry. We focussed on the predominant tissue-associated V δ 2⁻ $\gamma\delta$ subset, which is implicated in liver immunopathology.

Results: Intrahepatic V δ 2⁻ $\gamma\delta$ T cells were highly clonally focussed, with single expanded clonotypes featuring complex, private TCR rearrangements frequently dominating the compartment. Such T cells were predominantly CD27^{lo/-} effector lymphocytes, whereas naïve CD27^{hi}, TCR-diverse populations present in matched blood were generally absent in the liver. Furthermore, while a CD45RA^{hi} V δ 2⁻ $\gamma\delta$ effector subset present in both liver and peripheral blood contained overlapping TCR clonotypes, the liver V δ 2⁻ $\gamma\delta$ T cell pool also included a phenotypically distinct CD45RA^{lo} effector compartment that was enriched for expression of the tissue tropism marker CD69, the hepatic homing chemokine receptors CXCR3 and CXCR6, and liver-restricted TCR clonotypes, suggestive of intrahepatic tissue residency. Liver infiltrating V δ 2⁻ $\gamma\delta$ cells were capable of polyfunctional cytokine secretion, and unlike peripheral blood subsets, were responsive to both TCR and innate stimuli.

Conclusion: These findings suggest that the ability of V δ 2⁻ $\gamma\delta$ T cells to undergo clonotypic expansion and differentiation is crucial in permitting access to solid tissues, such as the liver, which results in functionally distinct peripheral and liver-resident memory $\gamma\delta$ T cell subsets. They also highlight the inherent functional plasticity within the V δ 2⁻ $\gamma\delta$ T cell compartment and provide information that could be used for the design of cellular therapies that suppress liver inflammation or combat liver cancer.

Lay summary: $\gamma\delta$ T cells are frequently enriched in many solid tissues, however the immunobiology of such tissue-associated subsets in humans has remained unclear. We show that intrahepatic $\gamma\delta$ T cells are enriched for clonally expanded effector T cells, whereas naïve $\gamma\delta$ T cells are largely excluded. Moreover, whereas a distinct proportion of circulating T cell clonotypes was present in both the liver tissue and peripheral blood, a functionally and clonotypically distinct population of liver-resident $\gamma\delta$ T cells was also evident. Our findings suggest that factors triggering $\gamma\delta$ T cell clonal selection and differentiation, such as infection, can drive enrichment of $\gamma\delta$ T cells into liver tissue, allowing the development of functionally distinct tissue-restricted memory populations specialised in local hepatic immunosurveillance. © 2018 European Association for the Study of the Liver. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

$\gamma\delta$ T cells are unconventional lymphocytes enriched in solid tissues, where they are thought to play critical roles in immunosurveillance.¹ Studies of mouse tissue-associated $\gamma\delta$ subsets suggest $\gamma\delta$ T cell function can be predominantly innate-like, involving semi-invariant T cell subsets that enable fast response kinetics without a requirement for clonal selection and differentiation.²⁻⁵ This role may allow for rapid 'lymphoid stress surveillance', limiting damage to host tissues in the face of microbial or non-microbial challenges, prior to full activation of adaptive immunity.^{4,6} As such, $\gamma\delta$ T cells may critically complement the contributions of tissue-resident $\alpha\beta$ subsets, which provide an augmented adaptive response to infections re-encountered at body surfaces,⁷ potentially explaining the retention of $\gamma\delta$ T cells alongside the $\alpha\beta$ T cell and B cell lineage over 450 million years of vertebrate evolution.⁸

Keywords: Gamma delta T cells; T cell receptor; Liver immune surveillance; Liver-resident T cells; Human liver; Immunological memory.

Received 13 October 2017; received in revised form 27 March 2018; accepted 2 May 2018; available online 18 May 2018

* Corresponding authors. Addresses: Centre for Liver Research and NIHR Birmingham Biomedical Research Centre, Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham B15 2TT, United Kingdom. Tel.: +44 (0) 121 414 2246; fax: +44 (0) 121 415 8701 (Y. H. Oo), or Cancer Immunology and Immunotherapy Centre, Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham B15 2TT, United Kingdom. Tel.: +44 0121 414 9533; fax: +44 0121 414 4486 (B. Willcox).

E-mail addresses: y.h.oo@bham.ac.uk (Y.H. Oo), b.willcox@bham.ac.uk (B.E. Willcox).

† These authors contributed equally to the study.



In contrast, the paradigms underlying human $\gamma\delta$ T cell immunobiology are far from clear. In humans, the peripheral blood is dominated by the $V\delta 2^+/V\gamma 9^+$ T cell subset, polyclonally activated by bacterial⁹ and endogenous phospho-antigens,¹⁰ arguably conforming to an innate-like paradigm.¹¹ In contrast, human solid tissues are enriched for $V\delta 2^- \gamma\delta$ T cells, of which the $V\delta 1^+$ subset is the most prevalent. It is far less clear if this dominant human tissue-associated subset also adopts an innate-like biology. Indeed, $V\delta 2^-$ T cells have been linked to recognition of a diverse range of ligands including to date Endothelial Protein C Receptor,¹² CD1 molecules,¹³ Annexin-A2,¹⁴ and even phycoerythrin.¹⁵ Moreover, recent data have provided strong evidence that $V\delta 1^+$ cells display an unconventional adaptive biology, undergoing clonal selection and differentiation from a naïve T cell receptor (TCR)-diverse precursor pool,¹⁶ with viral infection one trigger driving expansion.¹⁷ However, such studies have focussed on the subset of $V\delta 2^- \gamma\delta$ T cells that are retained in peripheral blood. To date, the immunobiology of human tissue-associated $\gamma\delta$ T cells remains relatively unstudied, despite the $V\delta 2^-$ T cell subset representing a considerable proportion of the total T cell infiltration in many human solid tissues, including gut,² lung¹⁸ and liver.¹⁹

To shed light on the function of tissue-associated $\gamma\delta$ T cells and how this relates to peripheral subsets, we characterised human intrahepatic $V\delta 2^-$ T cells. The liver is a site of considerable blood flow, receiving 75% of the total blood in the body every 2 h, with a third of this originating directly from the antigen-rich gut via the portal vein. In addition to providing a generally immunosuppressive microenvironment to facilitate tolerization of T cells toward non-pathogenic antigens present in the portal blood flow, the liver is also home to a large population of innate lymphoid cells, including natural killer (NK) cells, invariant natural killer T (iNKT) cells, mucosal associated invariant T (MAIT) cells²⁰ and $\gamma\delta$ T cells,¹⁹ in addition to CD8⁺ cytotoxic T cells.²¹ This enrichment is believed to balance the need for tolerization with a requirement for rapid identification and elimination of potentially harmful pathogenic entities, for example via pathogen associated molecular pattern receptors and semi-invariant T cell populations.²² To shed light on the immunobiology of $\gamma\delta$ T cells in this context we exploited next generation sequencing (NGS) approaches, allowing us to probe the TCR repertoire, in parallel with immunophenotype, and function.

Our study is the first to define the interconnected clonotypic, phenotypic and functional features of human tissue-associated $\gamma\delta$ T cells. The findings suggest that the liver selectively retains $V\delta 2^-$ T cells that are clonally expanded and adopt an effector phenotype, and which include a subset containing liver-restricted clonotypes that is phenotypically and functionally distinct from those present in peripheral blood.

Material and methods

Ethical approval and samples

Explanted diseased liver tissue and matched blood were obtained from patients who underwent liver transplantation for end-stage liver diseases including primary sclerosing cholangitis (PSC), primary biliary cholangitis (PBC), alcoholic liver disease (ALD), non-alcoholic steatohepatitis (NASH), hepatitis C virus (HCV) and hepatitis B virus (HBV) (Local Research Ethics Committee reference No. 98/CA5192) or normal liver samples from donor liver tissue surplus to clinical requirements

(Local Research Ethics Committee reference No. 06/Q2708/11). Unless otherwise stated (see Fig. 1), all diseased liver tissue analysed was from HCV/HBV-negative donors, and were non-cancerous. Normal liver tissue donors had no known prior history of liver disease or HCV/HBV infection. All diseased livers were Child C decompensated. Adult peripheral blood was obtained from consenting healthy donors (protocol approved by the NRES Committee West Midlands ethical board; REC reference 14/WM/1254).

T cell isolation, culture and activation

Human liver infiltrating lymphocytes were isolated from fresh liver tissue as described previously.²⁰ A whole slice of liver was processed, thereby reducing any effects of heterogeneous disease localisation. Briefly, explanted liver tissue was diced into 5 mm³ cubes, washed with Phosphate Buffered Saline (PBS), and then homogenised in a Seward stomacher 400 circulator (260 rpm, 5 min). The homogenate was filtered through fine (63 μ m) mesh (John Staniar and Co, Manchester, UK) and the lymphocytes were isolated by density gradient separation using Lympholyte (VH Bio, Gateshead, UK) at 800 \times g for 20 min. The lymphocyte layer was collected and washed with PBS. Cell viability was assessed by trypan blue exclusion. Peripheral blood mononuclear cells (PBMCs) were isolated from heparinised venous blood by lymphoprep[®] (Stem Cell Technologies) density gradient centrifugation as per the manufacturer's instructions. The cell culture medium used throughout this study was RPMI-1640 medium (Invitrogen) supplemented with 2 mM L-glutamine, 1% sodium pyruvate, 50 μ g/ml penicillin/streptomycin (Invitrogen) and 10% foetal calf serum (Sigma).

Antibodies and flow cytometry

For total and single-cell sorting of $V\delta 2^-$ and $V\delta 1^+$ $\gamma\delta$ T populations, PBMC were labelled with anti-CD3 (UCHT1; BioLegend), TCR $\gamma\delta$ (BW242/412), TCR $V\delta 2$ (123R3) or TCR $V\delta 1$ (REA173); all Miltenyi, CD27 (M-T271), CD45RA (HI100); BioLegend, and populations were sorted on a MoFlo Astrios (Beckman Coulter) or ARIA III Fusion (BD). For repertoire analysis, $V\delta 2^-$ T cell populations were sorted directly into RNAlater (Sigma). For phenotypic analysis, freshly isolated or frozen PBMCs, or cultured cells were labelled with Zombie Aqua viability dye (BioLegend), and then subsequently stained for cell surface antigens with antibodies directed against CD3 BV421 (UCHT-1, 1:100), CD8 BV650 (SK1; 1:200), CD45RA PeCy7 (HI100; 1:200), CD27 PE/Dazzle 594 (M-T271; 1:200), CCR7 AF647 (G043H7; 1:100), CD62L APC-Cy7 (DREG-56; 1:100), CD28 PE (28.2; 1:80), CD16 PE-Cy7 (3G8; 1:100), CD69 BV605 (FN50; 1:100), CD25 BV421 (2A3; 1:100), CD54 BV421 (HA58; 1:100), TCR $V\delta 2$ PE (B6; 1:100), TCR $\gamma\delta$ PE Cy7 (B1; 1:100), TCR $\alpha\beta$ PE (IP26; 1:50), CXCR3 PE (G025H7); all BioLegend. CXCR6 PE (56811/FAB699P; 1:20) from R&D Systems. Mouse anti-human CX₃CR1-PE (2A9-1; 1:20), CD69 PE (FN50; 1:50) from Immunotools. Mouse anti-human CD127 APC (IM1980U; 1:20), TCR $\gamma\delta$ PE Cy7 (IMMU510; 1:200), TCR $V\delta 3$ FITC and TCR $V\gamma 9$ PE Cy5 (IMMU360; 1:400); Beckman Coulter. TCR $V\delta 1$ PE (TS8.2; 1:100); Fisher Scientific. TCR $V\delta 1$ PE and FITC (REA173; 1:100) and TCR $V\delta 2$ APC (123R3; 1:200); Miltenyi Biotec. For intracellular staining, after surface antibody staining, cells were fixed in Foxp3/Transcription factor fix/perm buffer (eBioscience) and stained in permeabilization buffer (eBioscience) with antibodies directed against Granzyme A FITC (CBO9; 1:100), Granzyme B

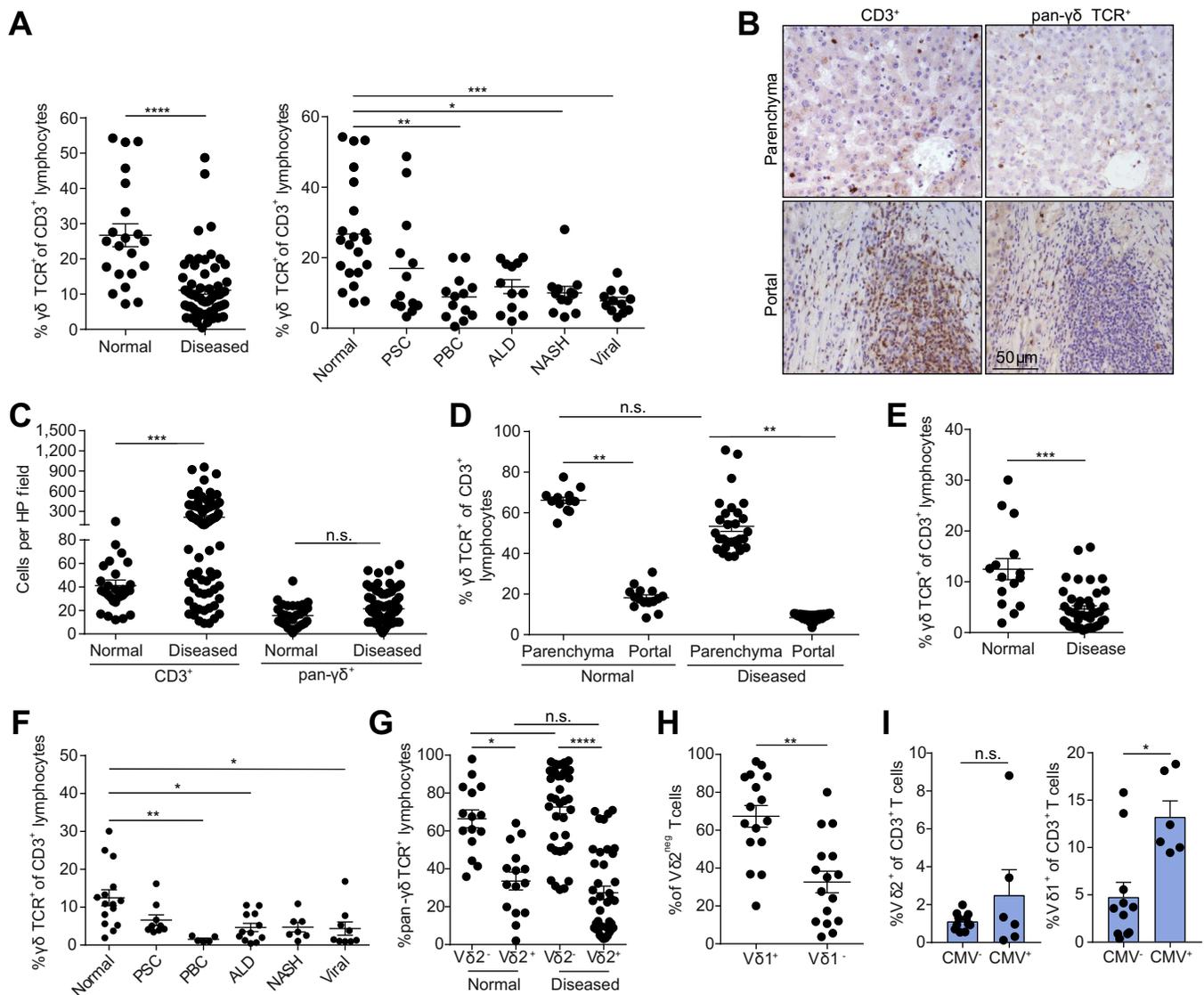


Fig. 1. Normal liver parenchyma is enriched for $\gamma\delta$ T cells. (A) Comparison of $\gamma\delta$ TCR⁺ proportion of CD3⁺ T cells identified by IHC in normal (n = 21) and diseased (n = 62) liver tissue (left) and in CD3⁺ T cells identified by IHC in normal (n = 21), PSC (n = 13), PBC (n = 13), ALD (n = 12), NASH (n = 12) and viral hepatitis (n = 12) liver tissue (right). (B) Representative staining for CD3⁺ (left) and $\gamma\delta$ TCR⁺ (right) cells on sequential FFPE sections from NASH liver tissue viewed at 40 \times magnification. (C) Densities of CD3⁺ and $\gamma\delta$ TCR⁺ cells in normal (n = 21) and diseased (n = 62) liver tissue. (D) Comparison of the $\gamma\delta$ TCR⁺ proportion of CD3⁺ T cells identified by IHC in parenchymal and portal areas of normal (n = 15) and diseased (n = 30) liver tissue. (E) Comparison of the $\gamma\delta$ TCR⁺ proportion of CD3⁺ T cells identified by flow cytometry in normal (n = 15) and diseased (n = 42) liver cell suspensions. (F) Comparison of the $\gamma\delta$ TCR⁺ proportion of CD3⁺ T cells identified by flow cytometry in normal (n = 15) and diseased liver cell suspensions of various aetiologies. (G) Comparison of V δ 2⁺ and V δ 2⁻ proportions in $\gamma\delta$ TCR⁺ cells identified by flow cytometry from normal (n = 15) and diseased (n = 42) liver cell suspensions. (H) Comparison of V δ 1⁺ and V δ 1⁻ proportions in V δ 2⁻ cells from liver cell suspensions (n = 16). (I) Comparison of V δ 2⁺ (left) and V δ 1⁺ (right) proportion of CD3⁺ T cells in CMV⁻ (n = 11) and CMV⁺ donors (n = 6) from diseased livers. Error bars indicate mean \pm SEM; data analysed by Kruskal-Wallis ANOVA with Dunn's post-test comparisons, n.s. $p > 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$. ALD, alcoholic liver disease; CMV, cytomegalovirus; FFPE, formalin-fixed paraffin embedded; IHC, immunohistochemistry; NASH, non-alcoholic steatohepatitis; PBC, primary biliary cholangitis; PSC, primary sclerosing cholangitis; TCR, T cell receptor. (This figure appears in colour on the web.)

APC (GB11; 1:100) and Perforin BV421 (B-D48; 1:80); all BioLegend. For intracellular cytokine staining, antibodies used were interferon- γ (IFN γ) BV421 (340449; 1:200), tumour necrosis factor alpha (TNF α) PE (554512; 1:200); BD Pharmingen, Cells were acquired on a CyAn ADP (Beckman Coulter), LSR II or LSR Fortessa X20 (BD) and data analysed with FlowJo V10.2 (TreeStar) or Summit 4.3 software (Dako Cytomation).

Immunohistochemistry and *in situ* hybridisation

Immunohistochemistry was performed using formalin fixed paraffin embedded (FFPE) sections using standard approaches.

In summary, sections were de-paraffinized, endogenous peroxidase activity was quenched using 0.3% hydrogen peroxide (Sigma Aldrich) in methanol for 20 min, and antigen retrieval carried out, involving boiling sections in 1% EDTA solution for 15 min. After washing and blocking steps, sections were incubated for 1 h in primary antibody (goat polyclonal – anti-human pan-V γ V δ (50 μ g/ml, A-20, Santa Cruz Biotechnology, Santa Cruz, USA) or rabbit polyclonal – anti-human CD3 (2 μ g/ml, ab5690, Abcam, Cambridge, UK) or relevant IgG1 isotype control) diluted in PBS. After washing, sections were incubated with HRP-linked anti-goat or anti-rabbit secondary

antibody (Vector Labs Laboratories) for 30 min at room temperature. Following washing, sections were developed using ImmPACT™ DAB reagent (Vector Laboratories). Excess DAB was then removed by rinsing and sections were counterstained with Mayer's haematoxylin solution (Leica Biosystems). Once dry, slides were mounted using DPX (Cellpath, Newtown Powys, UK) and imaged on a Zeiss Axioskop 40 Microscope. Regions of parenchymal and portal tract tissue were identified and numbers of CD3⁺ or $\gamma\delta$ -TCR⁺ cells were counted per region identified, with five high power fields, selected at random, scored for each section.

For *in situ* hybridisation, TCR chain-specific localisation of gamma delta TCR⁺ cells was performed using two protocols, either the ViewRNA™ ISH Tissue 2-Plex Assay developed by Affymetrix and performed manually, or the RNAscope® 2.5 LS Duplex Assay (ACD). For both protocols, liver slices were cut and immediately fixed in formalin for 24–48 h prior to being embedded in paraffin and mounted. Immediately after which the assay slides were baked at 60 °C for 1 h to immobilise the sections.

TCR repertoire analysis

RNA was purified from sorted cells (intrahepatic V δ 2⁻ T cells: 8,000–50,000 cells) protected in RNAlater (Sigma Aldrich) using an RNAmicro plus kit (Qiagen) according to the manufacturer's instructions. For high throughput deep sequencing of $\gamma\delta$ TCRs, we used amplicon rescued multiplex (ARM)-PCR and a MiSeq (illumina) next generation sequencer to analyse all sorted V δ 2⁻ T cell populations. Following initial first-round RT-PCR using high concentrations of gene-specific primers, universal primers were used for the exponential phase of amplification (Patent: WO2009137255A2), allowing deep, quantitative and non-biased amplification of TCR γ and TCR δ sequences. All cDNA synthesis, amplification, NGS library preparation and sequencing were performed by iRepertoire, Inc. (Huntsville, USA).

Single-cell TCR sequencing

PBMCs were labelled as described above and V δ 1⁺ T cells were single-cell sorted directly into individual wells in a 96 well plate containing 2 μ l of Superscript VILO cDNA synthesis kit reaction mix (ThermoFisher) containing 0.1% Triton X-100, and incubated according to manufacturer's instructions. TCR γ and TCR δ cDNAs were amplified by two rounds of nested PCR using GoTaq mastermix (Promega) and primers for or V δ 1, CAAGCCAGTCATCAG-TATCC (external) and CAACTTCCCAGCAAAGAGATG (internal); for C δ GCAGGATCAAACCTCTGTATCTTC (external) and TCCTTC ACCAGACAAGCGAC (internal); for V δ 3, GGCACGCTGTGTGACAAA (external) and CTGCTCTGCACTTACGACTG (internal); for V γ 1–8 CTGGTACCTACACCAGGAGGGGAAGG (external) and TGT GTTGAATCAGGAVTCAG (internal); for V γ 9 AGAGAGACCTGGT GAAGTCATACA (external) and GGTGGATAGGATACCTGAAACG (internal) and for C γ CTGACGATACATCTGTGTTCTTTG (external) and AATCGTGTGCTCTTCTTTCTT (internal). PCR products were separated on 1.2% agarose gels, and products of successful reactions were incubated with ExoSAP-IT PCR cleanup enzyme (Affymetrix) before sequencing with BigDye Terminator v3.1 (Applied Biosystems) following manufacturer's instructions and running on an ABI 3730 capillary sequencer (Functional Genomics Facility, University of Birmingham).

TCR repertoire data analysis

Sequences data was error corrected and V, D and J gene usage and complementarity-determining region 3 (CDR3) sequences

were identified and assigned, and tree maps generated using iRweb tools (iRepertoire, Inc, Huntsville, AL, USA). Tree maps show each unique CDR3 as a coloured rectangle, the size of each rectangle corresponds to each CDR3s abundance within the repertoire and the positioning is determined by the V region usage. For more detailed analysis and error correction of the TCR repertoire, datasets were processed using the MiXCR software package to further correct for PCR and sequencing errors. Diversity metrics, clonotype overlap and gene usage were plotted in R, by VDJTools.

TCR sequence analyses

The CDR3 length was defined as the number of amino acids between the second cysteine of the V region and the phenylalanine of the J region, according to IMGT. N and P nucleotides were identified using the IMGT Junction Analysis tool.

Statistical analysis

Tabulated data were analysed in Graphpad PRISM 7 (Graphpad Software Inc). Each data set was assessed for normality using Shapiro-Wilk normality test. Differences between columns were analysed by two-tailed Student's *t* tests for normally distributed data and Mann-Whitney for non-parametric data. Differences between groups were analysed using one-way ANOVA with Tukey's post-tests for normally distributed data or with Kruskal-Wallis with Tukey's post-tests for non-parametric data and RM two-way ANOVA with Tukey's post-tests was used when comparing groups with independent variables. **p* <0.05, ***p* <0.01, ****p* <0.001 and *****p* <0.0001.

Data availability

The sequence data that support the findings of this study have been deposited in the NIH NCBI sequence read archive database with the primary accession code SRP113556 and SRP096009, for $\gamma\delta$ TCR repertoires. For more detailed metadata relating to individual samples please contact the authors.

For further details regarding the materials used, please refer to the [CTAT table](#) and [Supplementary information](#).

Results

Human V δ 2⁻ $\gamma\delta$ T cell populations are reportedly tissue tropic in nature, with enrichment of this compartment previously highlighted in diseased human gut²³ and liver.¹⁹ We used immunohistochemistry (IHC) analysis to assess the infiltration and localisation of liver $\gamma\delta$ T cells. Firstly, $\gamma\delta$ T cells were a significantly enriched proportion of infiltrating CD3⁺ T cells in normal livers compared with livers explanted from patients with chronic liver disease (Fig. 1A). Furthermore, we noted the majority of the infiltrating CD3⁺ T cells were localised to portal areas; however, analysis of sequentially stained sections from normal tissue revealed a high proportion of parenchyma-associated CD3⁺ T cells were $\gamma\delta$ TCR⁺ (Fig. 1B). Importantly, while a significant increase in infiltrating CD3⁺ T cells was observed in diseased tissue, $\gamma\delta$ T cell numbers did not significantly change, suggesting that disease drives an increased infiltration of total CD3⁺ T cells but not $\gamma\delta$ TCR⁺ cell infiltration from the periphery (Fig. 1B, C, Fig. S1A). Further analysis of sequentially stained sections from explanted livers confirmed that $\gamma\delta$ TCR⁺ cells were also preferentially associated with the liver parenchyma (Fig. 1D, Fig. S1B). We then examined the TCR δ chain expression of liver infiltrating $\gamma\delta$ T cell populations by flow

cytometry, in homogenised single-cell suspensions of liver tissue from human explanted livers (Fig. S1C). Consistent with our IHC data, a significantly higher proportion of the CD3⁺ T cell compartment was comprised of $\gamma\delta$ T cells in healthy liver tissue compared with disease tissue (Fig. 1E-F), of which the majority were V δ 2⁻ (Fig. 1G), a direct inversion of the predominance of V δ 2⁺ T cells in the peripheral blood.^{24,16} Moreover, the majority of the V δ 2⁻ compartment was made up of V δ 1⁺ $\gamma\delta$ T cells (Fig. 1H, Fig. S1D), with the remainder comprised of other undefined V δ chains. Disease aetiology had no observed impact on this observation (Fig. S1E). Consistent with pan- $\gamma\delta$ T cell IHC, infiltration of V δ 1⁺ $\gamma\delta$ T cells into liver parenchyma was demonstrated using IHC and *in situ* hybridisation; again, IHC staining of sequential sections suggested a high proportion of parenchyma-associated CD3⁺ T cells were V δ 1⁺ (Fig. S1F). Of note, V δ 1⁺ $\gamma\delta$ T cells were significantly enriched as a proportion of intrahepatic T cells in diseased cytomegalovirus (CMV)⁺ liver donors compared with diseased CMV⁻ donors, while V δ 2⁺ T cells were not (Fig. 1I).

We next assessed the TCR repertoire of enriched populations of V δ 2⁻ $\gamma\delta$ T cells from both healthy and diseased liver tissue by amplicon rescued multiplex (ARM)-PCR and deep sequencing (Fig. S2A). Tree plot and clonotype analysis of V δ 2⁻ TCR repertoires indicated that both healthy and diseased liver tissue was generally dominated by a small number of highly prevalent clonotypes (Fig. 2A-C), with the 10 most prevalent CDR3 sequences accounting for >40% of TCR γ and TCR δ sequences in 9 and 8 out of 10 samples, respectively, and one dominant clone representing >50% in 2 of the 10 TCR γ and TCR δ samples (Fig. 2B-C). Comparison with D75 values obtained from adult and cord blood V δ 1⁺ TCR repertoires placed liver V δ 2⁻ TCR repertoires in a comparable range with other highly focussed $\gamma\delta$ TCR repertoires (Fig. 2D). Furthermore, when measuring the number of unique clonotypes detected in the first 10⁵ CDR3 sequences obtained in each sample, an alternative measure of TCR diversity, liver samples displayed a significantly less diverse repertoire than blood $\gamma\delta$ TCR repertoires (Fig. S2B). Comparison of Chao1 diversity metrics revealed no difference in the diversity of clonotypes between healthy and diseased liver TCR repertoires (Fig. S2C). Consistent with a broadly similar TCR repertoire in healthy and diseased tissue, comparison of normalised CDR3 lengths from healthy and diseased samples yielded no discernible difference (Fig. S2D). Previous studies have highlighted that peripheral blood V δ 2⁻ TCR γ repertoires contain few shared sequences.^{16,17} We found that liver V δ 2⁻ TCR γ repertoires were in general more private than blood V δ 2⁻ TCR γ repertoires and had very limited shared sequences between unrelated donors (Fig. S2E).

Consistent with flow cytometry analyses (Fig. 1G-H), V δ chain usage was dominated by V δ 1 (73.96% \pm SEM 8.7) and V δ 3 (24.05% \pm SEM 9.3) chain usage, with little V δ 4, V δ 5 and V δ 8 usage observed (Fig. 2E). Despite dominant clonotypes, V γ chain usage was highly heterogeneous, with all coding V γ chains utilised across our samples (Fig. 2F). Moreover, no significant difference was observed in V δ or V γ chain usage between healthy and diseased samples (Fig. 2E-F), consistent with the similar diversity metrics observed in diseased and healthy liver samples. These TCR sequencing data indicate the overwhelming prevalence of V δ 1⁺ TCR sequences in liver tissue, while confirming previous findings demonstrating a relative enrichment of V δ 3⁺ $\gamma\delta$ T cells in human liver compared to peripheral blood.¹⁹ Next, we assessed individual V δ 1⁺ and V δ 3⁺ TCR repertoires for

evidence of clonal expansion, initially using accumulated frequency curves to measure the 10 most prevalent clonotypes across all samples (Fig. S2F). These analyses provided evidence for clonal dominance in both liver V δ 1⁺ and V δ 3⁺ TCR repertoires, similar to clonotypically focussed peripheral blood V δ 1⁺ TCR repertoires but different from unfocussed cord blood V δ 1⁺ TCR repertoires (Fig. S2F).

This distinctive clonal dominance was unequivocally confirmed by sorting single intrahepatic V δ 1⁺ and V δ 3⁺ T cells and performing single-cell TCR sequencing. This approach highlighted that intrahepatic V δ 1⁺ and V δ 3⁺ (Fig. 3) T cell populations were composed of a small number of dominant clonotypes, using a variety of functional V γ and J γ gene segments. We also confirmed that concurrent clonal focussing can occur in both V δ 1⁺ and V δ 3⁺ TCR repertoires in the same donors (Fig. S3A). Moreover, analysis of CDR3 δ sequences revealed substantial complexity. As in peripheral blood, CDR3 δ 1 were long, frequently using two diversity (D) gene segments and containing extensive non-templated nucleotide (nt) additions (Table. S1). CDR3 δ 3 sequences were generally shorter than CDR3 δ 1 sequences and contained fewer non-templated nt (Table. S2; Fig. S3B), though there was no evidence of CDR3 δ 3 length restriction, in contrast to CDR3 γ 9 sequences in V γ 9⁺/V δ 2⁺ T cells.¹⁶ These data highlight the private nature of expanded clonotypes in intrahepatic V δ 2⁻ TCR repertoires and the broad range of V γ chains that they collectively utilise.

We next assessed the relationship between peripheral blood and intrahepatic V δ 1⁺ TCRs in the same individuals. Flow cytometry analysis of these matched samples indicated the enrichment of $\gamma\delta$ T cells in the liver (Fig. 4A), which occurred alongside the previously noted enrichment of CD8⁺ $\alpha\beta$ T cells (Fig. 4B).²⁵⁻²⁷ Moreover, while V δ 1⁺ T cells were specifically enriched there was an overall reduction in the proportion of infiltrating V δ 2⁺ T cells in the liver compared to the blood (Fig. 4C). Peripheral blood V δ 1⁺ T cells comprise both clonotypically focussed effector and separate TCR-unfocussed naïve sub-compartments, which can be delineated based on distinct CD27^{lo/-} CD45RA⁺ and CD27^{hi} CD45RA^{+/-} expression patterns, respectively.¹⁶ We assessed liver and blood V δ 1⁺ T cells for the expression of CD27 and CD45RA surface markers (Fig. 4D-E); we noted a loss of CD27^{hi} V δ 1⁺ T cells (Fig. 4D) in intrahepatic $\gamma\delta$ T cells, consistent with the lower diversity we observed in liver TCR repertoires than that of peripheral blood. While CD27^{lo/-} CD45RA^{hi} cells were present in both liver and blood, we noted the presence of an intrahepatic CD27^{lo/-} CD45RA^{lo/-} V δ 1⁺ T cell population that was present in all livers to varying degrees, but that was found at only very low levels in peripheral blood (Fig. 4E). The extent of this enrichment in liver was unaffected by liver disease aetiology (Fig. 4E) or CMV infection (Fig. S4).

We then explored the clonality of intrahepatic CD27^{lo/-} CD45RA^{hi} and CD27^{lo/-} CD45RA^{lo/-} populations by single-cell TCR sequencing. In a representative liver sample, sorted intrahepatic CD27^{lo/-} CD45RA^{lo} and CD27^{lo/-} CD45RA^{hi} V δ 1⁺ T cell populations each comprised single prominent, distinct clonotypes using single-cell sort identities (*i.e.* CD45RA^{hi} or ^{lo}), allowing the direct alignment of clonotype to phenotype at the single-cell level (Fig. 5A). Notably, within intrahepatic $\gamma\delta$ T cells, both the CD45RA^{hi} and CD45RA^{lo} populations were predominantly clonally expanded (Fig. 5A; B, left panel). Consistent with previous findings,¹⁶ in blood the CD27^{hi} compartment (reduced in frequency in liver) was polyclonal, whereas the

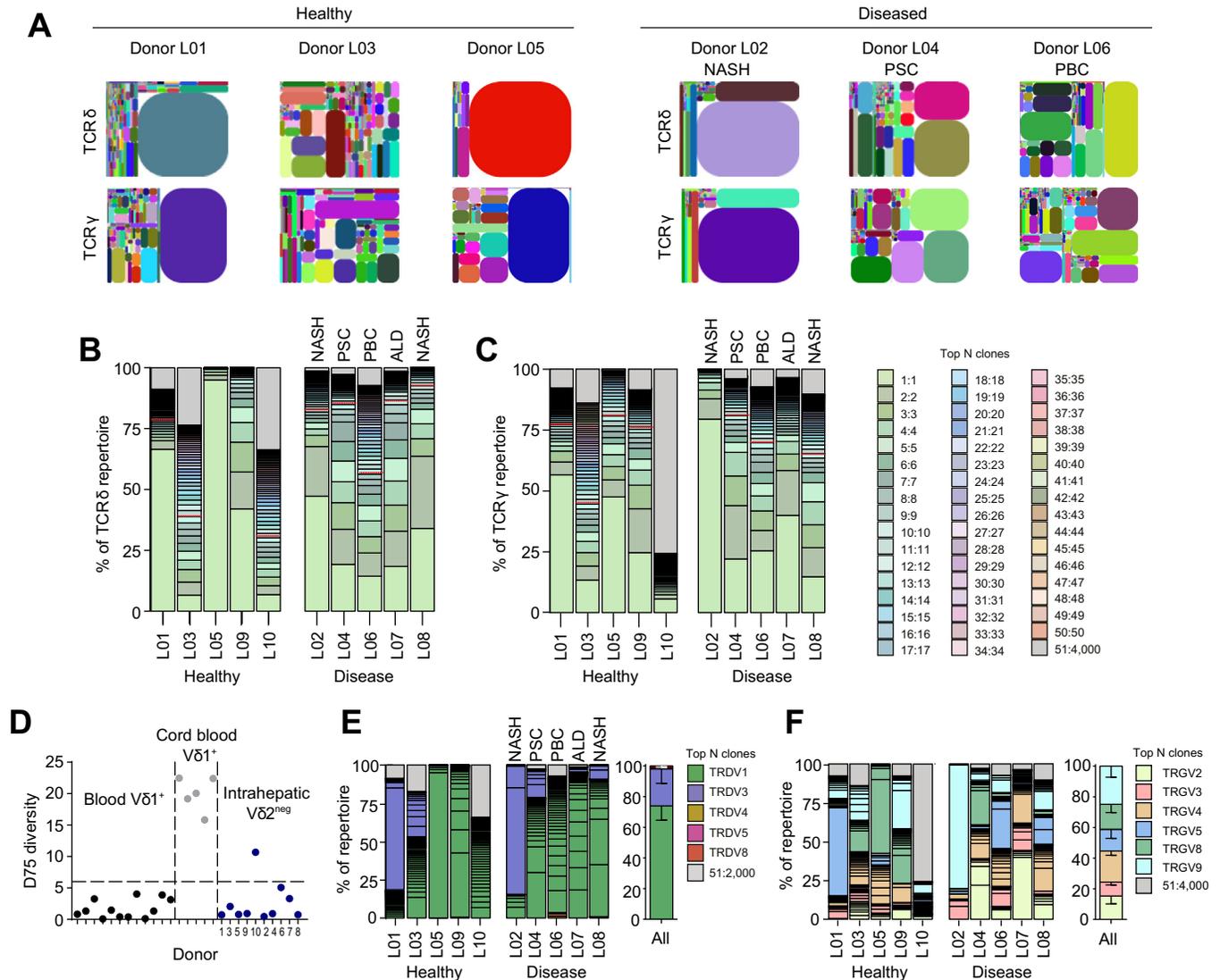


Fig. 2. Intrahepatic $V\delta 2^- \gamma\delta$ T cells are formed of clonally focussed TCR repertoires. (A) Representative tree maps show CDR3 clonotype usage in relation to repertoire size (each CDR3 colour is chosen randomly and does not match between plots) in TCR δ and TCR γ repertoires from $\alpha\beta$ TCR $^+$ $V\delta 2^-$ T cells sorted from normal ($n = 5$) and diseased livers ($n = 5$). Proportion of the total (B) TCR δ and (C) TCR γ repertoire occupied by the 50 most prevalent CDR3 sequences from sorted $V\delta 2^-$ T cells for each sorted liver sample ($n = 10$). The dashed black line denotes the percentage of the repertoire occupied by the ten most frequent clonotypes. (D) Analysis of inter-donor diversity by D75 (percentage of clonotypes required to occupy 75% of the total TCR repertoire) from TCR δ repertoire analyses from 12 healthy donors ($V\delta 1^+$), 5 cord blood donors ($V\delta 1^+$) and 7 liver samples ($V\delta 2^-$) and lowest quartile range plotted (dashed line). (E) $V\delta$ and (F) $V\gamma$ chain usage by the 50 most prevalent $\gamma\delta$ TCR CDR3 sequences from sorted $V\delta 2^-$ T cells from normal and diseased livers with summary plots. Error bars indicate mean \pm SEM. CDR3, complementarity determining region 3; TCR, T cell receptor. (This figure appears in colour on the web.)

$CD27^{lo/-} CD45RA^{hi}$ compartment was dominated by clonal expansions (Fig. 5B, right panel); notably the $CD27^{lo/-} CD45RA^{lo}$ compartment was essentially absent in blood. We then systematically examined the relationship between clonotypic and phenotypic identity from matched pairs of blood and liver $V\delta 1^+ \gamma\delta$ T cells (Fig. 5C). Overall in our paired samples, we identified clonotypes present in both the blood and liver, however importantly we also identified clonotypes unique to either liver or blood (Fig. 5C). The phenotype of clonotypes found only in the blood or shared between blood and liver generally mapped to the $CD27^{lo/-} CD45RA^{hi}$ compartment found both in blood and liver. In contrast, the clonotypes present exclusively in the liver mapped between $CD27^{lo/-} CD45RA^{lo}$ and $CD27^{lo/-} CD45RA^{hi}$ compartments, with a trend towards a $CD27^{lo/-} CD45RA^{lo}$ phenotype (Fig. 5C). As examples, the highly expanded $V\delta 1$ CALGGGGFPQKPGGAGPPTAQLFF and CALGEHPHFLHLIGTIKLIF

clonotypes present in the livers of Donor 0886 and Donor 1421 (both ALD) respectively were $CD27^{lo/-} CD45RA^{hi}$ in phenotype and also present in the respective matched peripheral blood samples, whereas in each case liver-restricted expanded clonotypes were also observed, but predominantly $CD27^{lo/-} CD45RA^{lo}$ (Fig. S5A). Taken together, while considerable clonotypic overlap between liver and blood subsets is observed, we identified a distinct population of intrahepatic $CD27^{lo/-} CD45RA^{lo} V\delta 1^+$ T cells largely absent from the blood, and which frequently contains TCRs restricted to the liver. This paradigm is likely to extend to intrahepatic $V\delta 3^+ \gamma\delta$ T cells, which also exhibited a significant proportion of $CD45RA^{lo}$ cells (Fig. S5B).

We sought to further characterise intrahepatic $CD27^{lo/-} CD45RA^{lo}$ and $CD27^{lo/-} CD45RA^{hi} V\delta 1^+$ T cells for markers associated with tissue retention. Firstly, while the surrogate marker of tissue-resident memory T cells (T_{RM}), CD69, was expressed

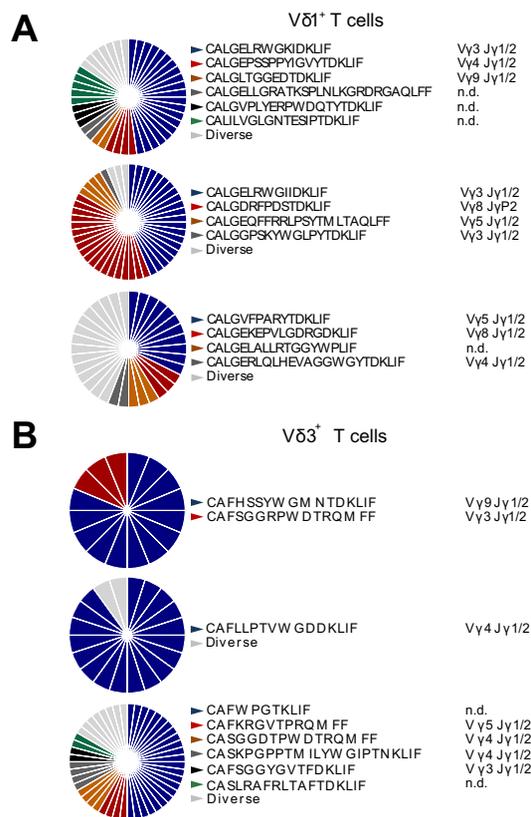


Fig. 3. Single-cell TCR sequencing reveals clonal focussing in Vδ2⁻ γδ T cells. Clonal focussing of intrahepatic (A) Vδ1⁺ and (B) Vδ3⁺ cells determined by single-cell TCR sequencing analysis of CDR3δ. Each colour represents an individual CDR3δ, with clonal sequences labelled beside each chart (from 16–42 single cells per population, as indicated; with each pie chart representing an independent donor). CDR3, complementarity determining region 3; TCR, T cell receptor. (This figure appears in colour on the web.)

widely by Vδ1⁺ T cells, it was markedly higher on CD27^{lo/-} CD45RA^{lo} Vδ1⁺ T cells and comparable to CD45RA^{lo} CD8⁺ αβ T cells (Fig. 6A). In keeping with functional tissue retention, Vδ1⁺ T cells expressed CXCR3 and CXCR6, with expression predominantly associated with the CD27^{lo/-} CD45RA^{lo} population (Fig. 6B). In contrast, the endothelial homing receptor CX₃CR1 (highly expressed by peripheral blood CD27^{lo/-} CD45RA^{hi} Vδ1⁺ T cells¹⁶) was retained on intrahepatic CD27^{lo/-} CD45RA^{hi} cells but was markedly reduced on CD27^{lo/-} CD45RA^{lo} Vδ1⁺ T cells (Fig. 6B). Interestingly, intrahepatic CD45RA^{lo} Vδ1⁺ T cells did not express significantly more CD103 than CD45RA^{hi} Vδ1⁺ T cells, which contrasts with CD8⁺ CD45RA^{lo} T cells isolated from the same livers (Fig. 6B). We next assessed the functionality of intrahepatic Vδ1⁺ T cell populations by *ex vivo* stimulation with recombinant cytokines or by TCR activation. Following TCR stimulation, intrahepatic Vδ1⁺ T cell populations in general strongly upregulated the T cell activation marker CD25, with equivalent responses in CD8⁺ αβ T cells from the same samples, although Vδ1⁺ T cells from some liver samples responded more robustly than others. Importantly, intrahepatic CD27^{lo/-} CD45RA^{lo} Vδ1⁺ T cells displayed a greater sensitivity to innate associated cytokines IL-12 and IL-18, than CD27^{lo/-} CD45RA^{hi} Vδ1⁺ T cells (Fig. 6C). Notably, peripheral blood CD27^{lo/-} CD45RA^{hi} Vδ1⁺ T cells are unresponsive to IL12/IL-18 stimulation.¹⁶ In keeping with a clonally expanded intrahepatic Vδ1⁺ T cell population, significant responses were observed with

IL-15 but not IL-7 cytokines (Fig. 6C). We next assessed effector potential, by analysing intracellular expression of cytolytic granzyme B and perforin. Intrahepatic CD27^{lo/-} CD45RA^{hi} Vδ1⁺ T cells expressed marked levels of both effector molecules while CD27^{lo/-} CD45RA^{lo} Vδ1⁺ T cells had much lower expression (Fig. 6D). Conversely, stimulation of the CD27^{lo/-} CD45RA^{lo} population with PMA and ionomycin produced significantly more of the pro-inflammatory cytokines IFN-γ and TNFα than the CD27^{lo/-} CD45RA^{hi} population (Fig. 6E). These data suggest that intrahepatic CD27^{lo/-} CD45RA^{lo} Vδ1⁺ T cells have a more prominent tissue-associated phenotype than that of the CD27^{lo/-} CD45RA^{hi} Vδ1⁺ T cell population, which are more similar to peripheral blood CD27^{lo/-} CD45RA^{hi} Vδ1⁺ T cells. Moreover, these two populations possess either enhanced cytolytic (CD45RA^{hi}) or pro-inflammatory cytokine (CD45RA^{lo}) responses, suggesting distinct roles in intrahepatic immunity.

Discussion

Tissue-associated T cells are thought to play a critical role in tissue immunosurveillance and homeostasis.^{28–30} In mice, γδ T cells have been implicated in epithelial homeostasis,³¹ cutaneous wound healing³² and maintenance of gut mucosa,³³ and have been highlighted as innate-like, expressing canonical TCRs.³⁴ In humans, solid tissues are known to be enriched for γδ T cells but the immunobiology of the T cells present has remained largely unclear. Recent studies on Vδ1⁺ T cells, the canonical tissue-associated human γδ T cell subset, have revealed an adaptive biology.^{16,17} However, these results were based exclusively on peripheral blood Vδ1⁺ cells, and the immunobiology of solid tissue-associated Vδ1⁺ lymphocytes, often assumed to be innate-like, is of particular interest. We chose to probe these issues by characterising intrahepatic γδ T cells as a human model system.

We used NGS approaches to show the hepatic Vδ2⁻ compartment is comprised of highly clonal, private expansions, based on complex TCR rearrangements. Importantly these were evident in both diseased and healthy livers, with no skewing of the TCR repertoire chain usage observed between the two scenarios. Moreover, the proportion of Vδ2⁻ γδ T cells decreased upon liver inflammation compared with healthy livers, because of an influx of αβ T cells. Therefore, the accumulation of γδ T cells in human liver is not driven by the diseased hepatic microenvironment present in these patients, and may reflect a response to other immune challenges such as infection. Of relevance, CMV infection has recently been highlighted as one of a number of drivers of Vδ2⁻ T cell clonal expansion (specifically of Vδ1⁺ T cells) in peripheral blood.^{16,17} Moreover, studies on murine CMV have highlighted the potential of expanded γδ T cell subsets to populate a range of peripheral tissues, including the liver.^{35,36} These observations raise the significant possibility that the expanded clonotypes that contribute so dominantly to human intrahepatic γδ T cells both in normal and diseased settings have arisen due to previous infections. Consistent with this, Vδ1⁺ γδ T cells were significantly enriched in liver explants from CMV⁺ vs. CMV⁻ donors. Therefore, CMV represents one likely driver of Vδ1⁺ infiltration in the liver. However, it is notable that similar clonotypic focussing and immunophenotypic profiles of intrahepatic Vδ2⁻ T cells were observed in both CMV⁺ and CMV⁻ individuals, consistent with the idea that the Vδ2⁻ subset can mount tissue-localised responses to multiple infections. This mirrors the situation with human Vδ1⁺ T cells

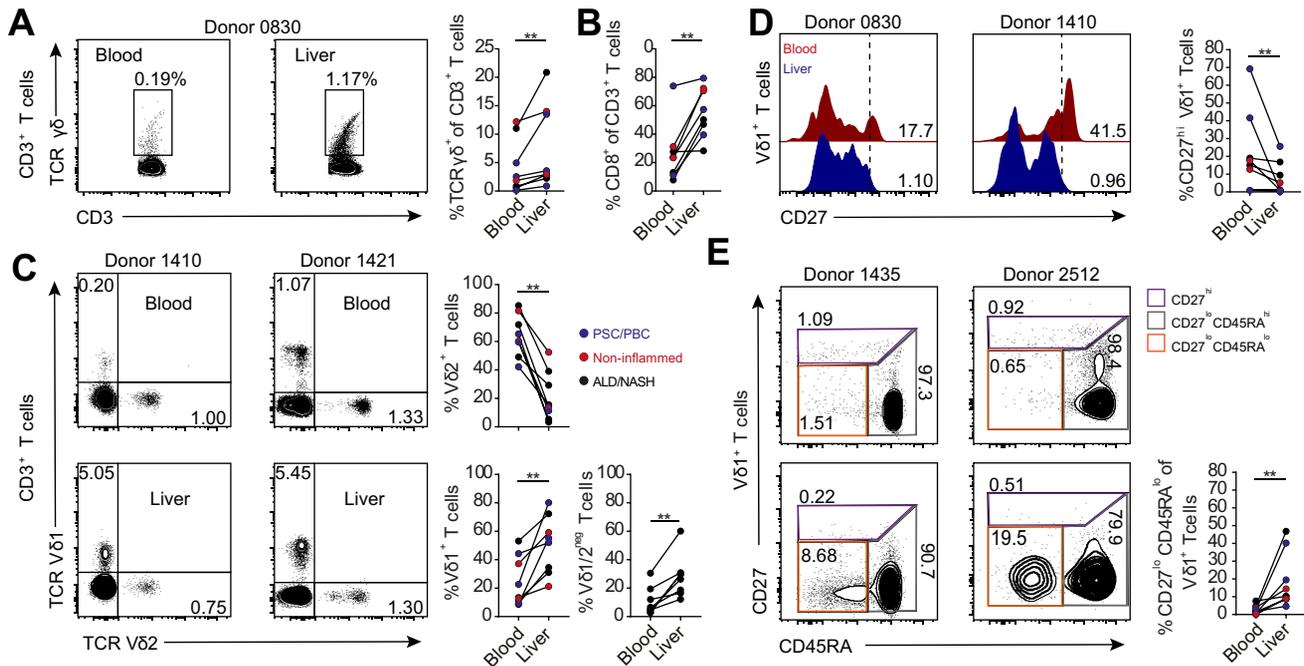


Fig. 4. Intrahepatic $V\delta 1^+$ T cells are phenotypically distinct from those in matched blood. Representative flow cytometry plots and summary data of the enrichment of (A) $\gamma\delta$ TCR⁺ and (B) CD8⁺ cells in donor matched liver and peripheral blood samples (n = 8). (C) Representative flow cytometry plots and summary data of the enrichment of $V\delta 1^+$ and $V\delta 1^-V\delta 2^-$ or contraction of $V\delta 2^+$ T cells in liver (lower plots) and donor matched peripheral blood (upper plots) (n = 8). (D) Representative histograms and summary data of the frequency of CD27^{hi} $V\delta 1^+$ T cells derived from donor matched liver and peripheral blood samples (n = 7). (E) Representative flow cytometry plots and summary data of CD27^{lo/-} CD45RA^{hi} and CD45RA^{lo} populations within donor matched liver (lower panels) and peripheral blood (upper panels) $V\delta 1^+$ T cells (n = 8). Data analysed by Mann-Whitney U test, ***p < 0.01. TCR, T cell receptor. (This figure appears in colour on the web.)

in peripheral blood, where although CMV is linked with an increased proportion of $V\delta 1^+$ T cells^{16,37} and clearly drives clonal expansions of $V\delta 1^+$ clonotypes,¹⁷ such expansions are commonly observed in CMV⁻ individuals, suggestive of other infectious drivers.¹⁶ While the candidate drivers of intrahepatic $V\delta 2^+$ T cell expansion would include HCV/HBV, notably we did not study HCV/HBV-related liver disease, and therefore other non-CMV/HCV/HBV drivers must exist. In principle, an alternative to infection representing a main driver of $V\delta 2^-$ clonal expansion is that intrahepatic $V\delta 2^-$ T cells are populated in the liver during development. However, both their $V\delta 2^-$ chain usage and the highly complex nature of the intrahepatic $V\delta 2^-$ TCR CDR3 regions would argue against this possibility, since foetal $\gamma\delta$ TCRs would be expected to utilise more simple CDR3 sequences and have also been highlighted as predominantly $V\delta 2^+$,³⁸ thereby highlighting post-natal stimuli such as infection as a more likely underlying driver.

Given previous observations regarding peripheral blood $V\delta 1^+$ T cells,¹⁶ which like those in the liver were frequently highly clonal and also featured private expansions based on complex TCR rearrangements, a key question was the extent to which liver $V\delta 2^- \gamma\delta$ T cells mirrored those in the blood. Our study provides compelling evidence that despite the profound link between the liver and the peripheral circulatory system, there is a distinct profile of $V\delta 2^- \gamma\delta$ T cells in each compartment, indicative of compartmentalisation of certain $V\delta 2^-$ subsets.

Comparison of matched liver and blood samples indicated the differentiation status of the $V\delta 2^-$ T cell subset was distinct in each compartment. Strikingly, liver $V\delta 2^-$ T cells were uniformly CD27^{lo/-}, a phenotype previously linked to a clonally expanded effector subset present in peripheral blood, and essentially entirely lacked the CD27^{hi} subset, even when such

populations were relatively prevalent in matched blood. Previously we have shown that CD27^{hi} $V\delta 1^+$ T cells in peripheral blood are TCR-diverse and naïve in phenotype. Consistent with selective exclusion of this clonally diverse CD27^{hi} naïve population, liver $V\delta 2^-$ cells lacked CCR7, CD62L and CD27 present on such naïve populations, and diversity metrics indicated liver $V\delta 2^-$ T cells displayed an even more focussed repertoire in liver than in peripheral blood. Furthermore, the phenotype of liver $V\delta 2^-$ T cells closely matched that of peripheral blood CD27^{lo/-} $V\delta 1^+$ T cells, and there was substantial clonotypic overlap between these two populations. While we cannot exclude the possibility that such hepatic CD27^{lo/-} originated in the liver, these results support the concept that at least some hepatic CD27^{lo/-} cells may derive from those present in peripheral blood. Such a scenario would fit an adaptive model whereby naïve peripheral blood $V\delta 2^-$ CD27^{hi} cells, which express secondary lymphoid homing markers but are devoid of CX₃CR1, recirculate between blood and lymph, whereas the peripheral blood CD27^{lo/-} population, which is clonally expanded and likely antigen-experienced, is capable of accessing solid tissues, potentially because of increased CX₃CR1 expression, and may also upregulate tissue retention markers following liver localisation.

A second indication of compartmentalisation was that in addition to being devoid of CD27^{hi} naïve cells, the hepatic $V\delta 2^-$ T cell compartment comprised both a CD45RA^{hi} and also a distinct CD45RA^{lo} subset. By contrast, the peripheral blood CD27^{lo/-} $V\delta 1^+$ cells are almost entirely CD45RA^{hi}. Importantly, CD45RA^{hi} clonotypes overlapped substantially between blood and liver within individuals. Such cells in the peripheral blood express a high level of the endothelial homing receptor CX₃CR1 as well as increased CD16, low CD27/28, low CD127, and

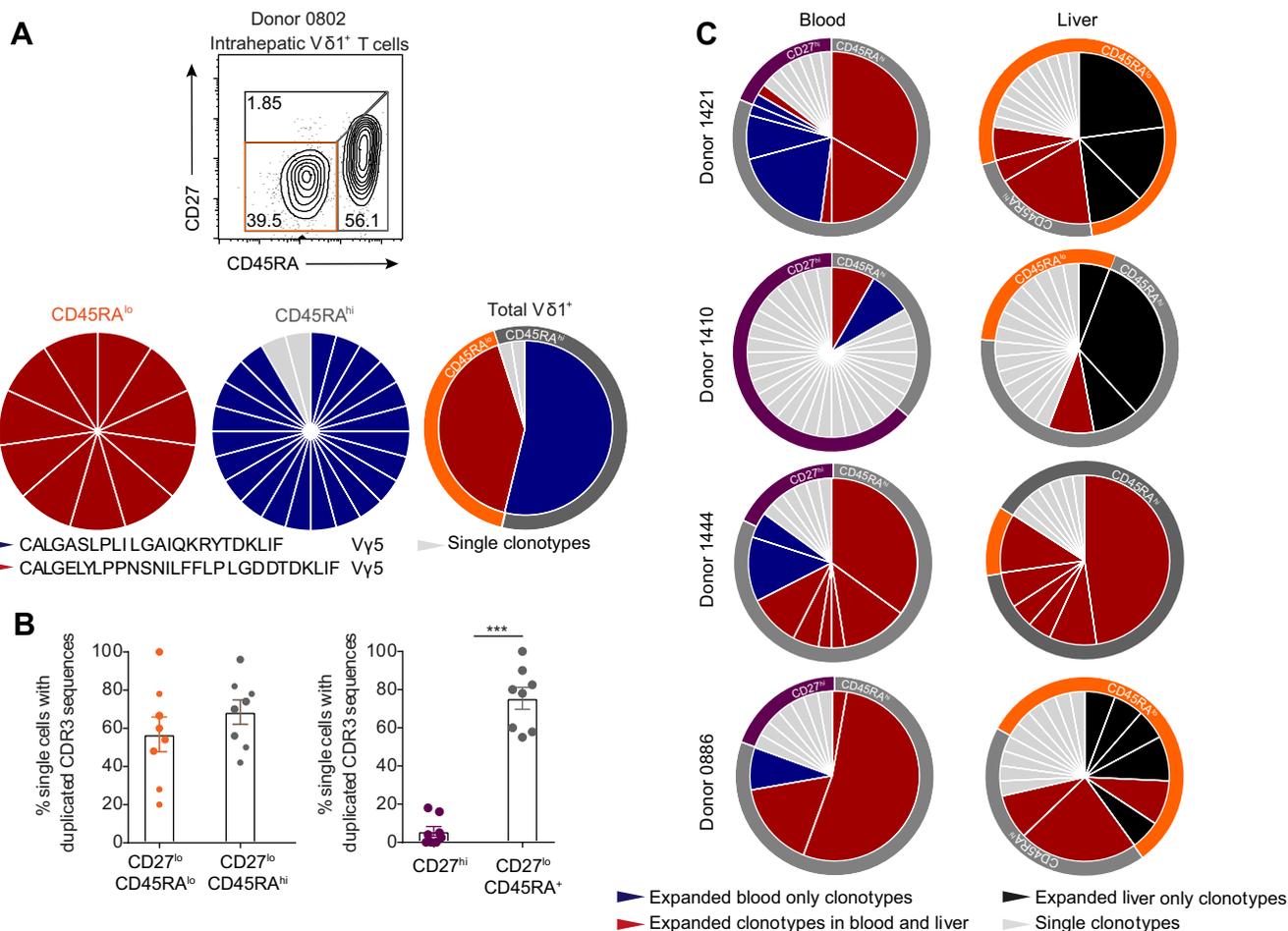


Fig. 5. Intrahepatic Vδ1 T cells contain clonotypes both distinct and overlapping with the blood. (A) Clonal focussing of intrahepatic Vδ1⁺ CD27^{lo/-} CD45RA^{lo} (n = 11 single cells) and CD27^{lo/-} CD45RA^{hi} (n = 24 single cells) cells determined by single-cell TCR sequencing analysis of CDR3δ. Each colour represents an individual CDR3δ, with clonal amino acid sequences labelled below the chart. Total Vδ1⁺: TCR sequence data was combined with flow cytometry data to generate the two layered pie, linking clonotype (inner pie chart) to phenotype (outer pie chart). (B) Assessment of clonality by single-cell TCR sequencing analysis of CD27^{lo/-} CD45RA^{lo}, CD27^{lo}CD45RA^{hi} and CD27^{hi} Vδ1⁺ T cells sorted from liver and donor matched blood (n = 8). (C) Comparison of the relationship between phenotype (outer pie chart) and clonality (inner pie chart) determined by phenotype-linked indexed single-cell TCR sequencing analysis, in donor matched peripheral blood (upper) and liver (lower) Vδ1⁺ T cells, classified according to clone presence within liver and/or blood compartments. Error bars indicate mean ± SEM; data analysed by Mann-Whitney U test, ***p < 0.001. CDR3, complementarity determining region 3; TCR, T cell receptor. (This figure appears in colour on the web.)

enhanced levels of adhesion molecules relative to naïve CD27^{hi} cells.¹⁶ While this could suggest capability of homing from peripheral blood to tissues, alternatively it could imply a vascular association, as has been suggested for effector memory CD8 T cells,³⁹ which include virus-specific CD8⁺⁴⁰ and CD4⁺⁴¹ T cell subsets. The predominantly sinusoidal localisation of these cells identified in this study is consistent with this possibility, and may suggest a role in immunosurveillance at this site, as suggested for NKTs.⁴² In light of the recent report that Vδ1⁺ clonotypes can expand in response to CMV,¹⁷ a virus that infects the endothelial compartment *in vivo*, and our observation here that Vδ1⁺ T cells are enriched in CMV⁺ vs. CMV⁻ liver explants, these findings suggest this subset may contribute to unconventional T cell protection of the vascular niche, including within solid tissues, against chronic viral infection. Moreover, the observation CMV serostatus correlates with an enhanced proportion of intrahepatic Vδ1⁺ T cells but not with a disturbed CD45RA^{hi} vs. CD45RA^{lo} Vδ1⁺ ratio might suggest the potential within both phenotypic sub-compartments to respond to CMV.

In contrast to CD45RA^{hi} clonotypes and consistent with a reduced frequency of CD45RA^{hi} Vδ2⁻ cells in liver compared

to peripheral blood, the same analyses of matched blood/liver samples revealed CD45^{lo} clonotypes were enriched for those restricted to the liver. In addition, this liver CD45RA^{lo} compartment frequently contained clonal expansions. These cells demonstrate striking phenotypic correlation with liver-resident lymphocytes identified in previous studies, including enhanced expression of CD69, CXCR3 and CXCR6, which has been noted in liver-resident NK populations^{43,44} and CD8⁺ αβ populations.²⁵ CD27^{lo/-} CD45RA^{lo} Vδ2⁻ T cells may therefore represent a liver-resident subset, although conceivably they may be able to access other solid tissues. Of note, CD45RA^{lo} Vδ1⁺ T cells exhibited considerably lower expression of CD103 relative to their CD8⁺ counterparts, suggesting other mechanisms may underly their tissue retention. The origin of this subset is unclear. One possibility is that it originates from a subset of blood CD45RA⁺ cells that alter phenotype once in tissues and are retained there, perhaps following activation in the hepatic microenvironment. This route of generation is supported by our detection of liver-restricted clonotypes in both the CD45RA^{lo} and CD45RA^{hi} compartments. In addition, it is possible they may be locally generated. Moreover, recent reports

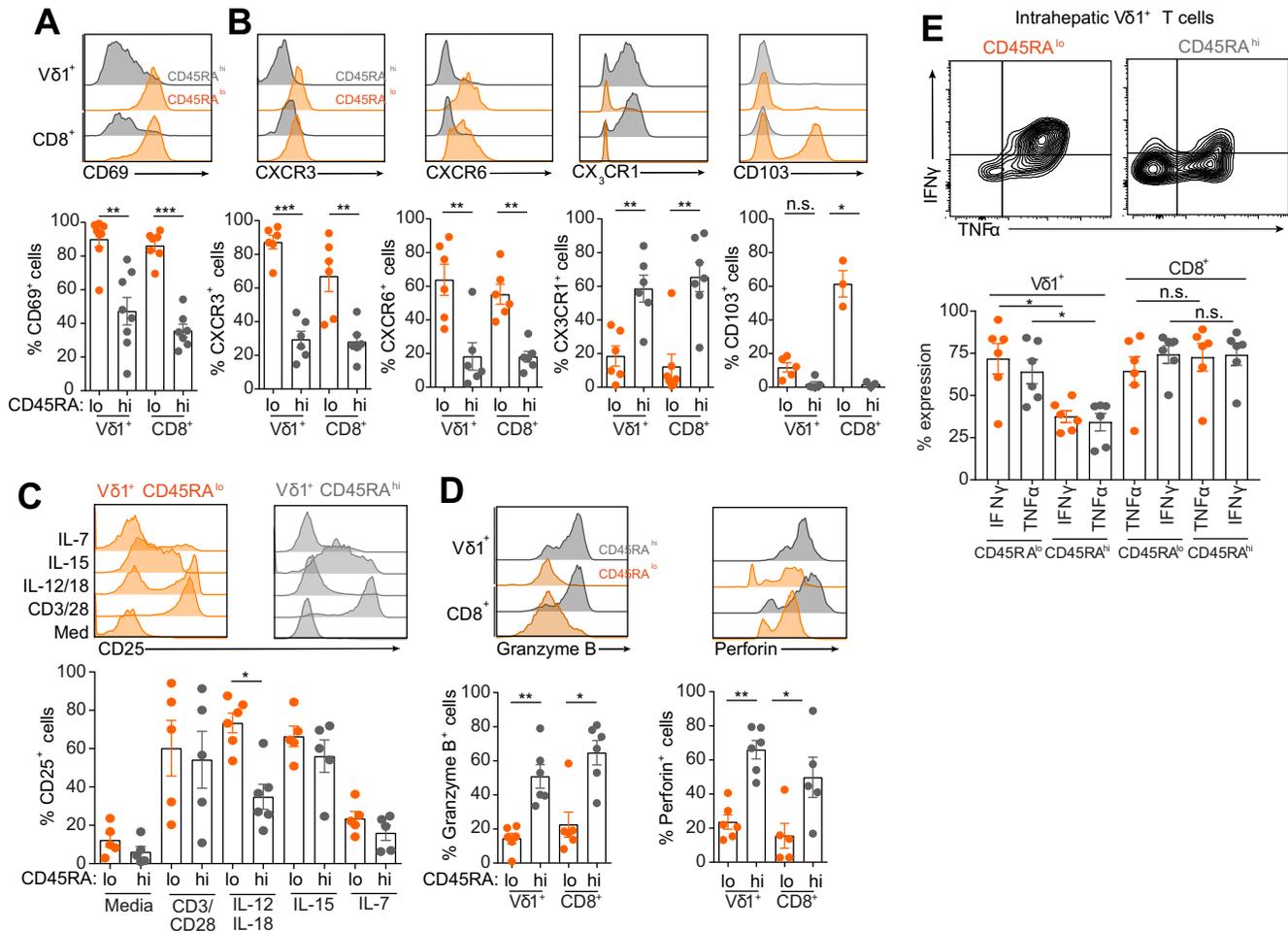


Fig. 6. Intrahepatic Vδ1⁺ T cells segregate into cytokine producing and cytotoxic subsets. (A) Representative histograms from one donor and summary data of CD69 surface expression by CD45RA^{lo} (orange) and CD45RA^{hi} (grey) intrahepatic Vδ1⁺ and CD8⁺ T cells (n = 8). (B) As in (A), but displaying representative histograms and summary data for CXCR3, CXCR6 and CX₃CR1 surface expression by intrahepatic Vδ1⁺ and CD8⁺ T cells (n = 6). (C) Representative histograms and summary data from sorted intrahepatic CD3⁺ T cells were incubated with indicated medium, cytokines or anti-CD3/CD28 beads for 72 h. CD45RA^{lo} (orange) and CD45RA^{hi} (grey) Vδ1⁺ T cells were then assessed for the upregulation of the T cell activation marker CD25 (n = 5–6). (D) Representative histograms and summary data for intracellular granzyme B and perforin expression by CD45RA^{lo} (orange) and CD45RA^{hi} (grey) intrahepatic Vδ1⁺ and CD8⁺ T cells (n = 5–6). (E) Representative flow cytometry plot and summary data of intrahepatic CD3⁺ T cells stimulated with PMA/ionomycin and assessed for the production of intracellular IFNγ and TNFα in CD45RA^{lo} (orange) and CD45RA^{hi} (grey) Vδ1⁺ and CD8⁺ cells (n = 6). Error bars indicate mean ± SEM; data analysed by Kruskal-Wallis ANOVA with Dunn's post-test comparisons, n.s. *p* > 0.05, **p* < 0.05, ***p* < 0.01 and ****p* < 0.001. (A–E) Disease aetiologies analysed included ALD, NASH, PSC, and normal liver no significant differences were observed between different individual disease groups in any of the comparisons highlighted. ALD, alcoholic liver disease; IFN, interferon; NASH, non-alcoholic steatohepatitis; PSC, primary sclerosing cholangitis; TNFα, tumour necrosis factor alpha. (This figure appears in colour on the web.)

highlight that a liver-resident phenotype can be induced in CD8⁺ αβ T cells via IL-15 followed by TGF-β signalling,²⁵ and based on the parallels between Vδ1⁺ and CD8⁺ αβ T cells identified in this study, a similar mechanism may be at work here.

Our results also highlight that hepatic γδ T cells are functionally distinct from equivalent subsets in peripheral blood. While still responsive to TCR stimulation/co-stimulation, compared to blood Vδ2⁻ T cells they displayed markedly increased responsiveness to IL-12/IL-18 in line with CD8⁺ T cells isolated from the same tissue. This responsiveness extended to the liver-restricted CD45RA^{lo} subset, which appeared to display enhanced production of pro-inflammatory cytokines relative to CD45RA^{hi} cells. These observations suggest CD45RA^{hi} and CD45RA^{lo} subsets may have different roles, the former more vascular focussed and cytotoxic, the latter an immunoregulatory tissue-associated subset more focussed on cytokine production and potential induction of a wider T cell response to stress challenges. It is unclear if these distinct features stem

directly from the nature of the clonotypes present and their antigenic targets, or whether they reflect the influence of hepatic microenvironmental factors that may also influence intrahepatic retention.⁴⁵

Importantly, we note several limitations of our study. Firstly, all diseased samples were derived from end-stage liver disease. While the closely matched clonotypic focussing and immunophenotypic profiles present in normal tissue would predict similar profiles at earlier disease stages, we cannot exclude the possibility that disease stage influences the nature of the intrahepatic γδ T cell population, and use of biopsy material from early disease stages with longitudinal follow-up could be an interesting avenue of future investigation. Secondly, while we examined several disease pathologies, these were predominantly restricted to fatty/alcoholic liver disease (ALD, NAFLD) or autoimmune liver disease (AIH, PBC, or PSC). While HCV/HBV+ liver samples showed similar frequencies of γδ T cells, we did not study γδ T cell immunophenotype or clonotypic focussing

in such samples and cannot therefore exclude the possibility that HCV/HBV infection may drive development of distinct intrahepatic $\gamma\delta$ T cell profiles⁴⁷ or clonality, although we hypothesise they would follow broadly similar principles to those observed in this study; moreover, while we did not observe differences between the different disease types we did analyse, conceivably with larger samples sizes differences may have emerged, for example in the extent of $\gamma\delta$ TCR clonotypic focussing or $\gamma\delta$ T cell phenotypes. Finally, a comparison of the data presented here with $\gamma\delta$ T cell clonotype and immunophenotype profiles in other solid tissues, including during chronic inflammation, would shed light on tissue-specific $\gamma\delta$ T cell responses.

Our study establishes that in humans, clonally expanded $\gamma\delta$ T cell effector subsets can be selectively deployed to at least some solid tissues, including the liver, thereby providing ongoing immune surveillance against previously encountered infectious or non-infectious challenges, with CMV infection one likely driver of $V\delta 1^+$ T cell intrahepatic infiltration. Importantly, both $V\delta 1^+$ and $V\delta 3^+$ intrahepatic T cell compartments displayed clonotypic expansion and a CD45RA^{lo} subset, suggesting their immunobiology may be closely aligned. Moreover, the finding that intrahepatic $\gamma\delta$ T cell subsets can be phenotypically, clonotypically and functionally distinct from those in peripheral blood suggests distinct contributions to intrahepatic immune responses, and provides a basis for future investigation of human tissue-resident $\gamma\delta$ T cell populations. Notably, $\gamma\delta$ T cells are of increasing therapeutic interest, due partly to their potential to mount either anti-tumour,^{47–49} or alternatively immunosuppressive⁵⁰ responses, but also their MHC-unrestricted recognition of target cells, which raises the prospect of broad applicability of $\gamma\delta$ T cell-based therapies in patient cohorts. Our finding that there appears to be selective recruitment of $\gamma\delta$ T cell subsets of an effector phenotype into the hepatic pool may inform design of $\gamma\delta$ T cellular therapies that rely on administration/expansion of systemic $\gamma\delta$ T cells. Secondly, the finding that a number of distinct differentiation states exist within the $V\delta 1^+$ compartment (including naïve, circulating effector, tissue-resident effector) indicates a degree of plasticity that could be investigated further and potentially exploited therapeutically, either to increase immunosuppressive functionality during inflammatory liver disease, or for improved anti-tumour effector function in liver cancer. Finally, our finding that CMV infection represents one likely factor driving infiltration of potentially highly inflammatory $V\delta 1^+$ T cells into the liver could have clinical relevance in chronic liver disease and CMV-associated hepatitis. Specifically, future studies correlating CMV titres with biomarkers of liver damage and with $V\delta 1^+$ $\gamma\delta$ T cell frequency may shed light on whether the $\gamma\delta$ T cell response to CMV infection impacts the severity of chronic liver disease.

Financial support

The work was supported by a Medical Research Council funded PhD studentship to S.H., by Wellcome Trust Investigator award funding to B.W., supporting M.D. and C.W. (Grant code: 099266/Z/12/Z), by a Russian Foundation for Basic Research grant 17-04-01994 (S.K.) and 17-54-10018 (D.C.), by a Medical Research Council Clinician Scientist award to Y.H.O (G1002552), and by the Ministry of Education, Youth and Sports of the Czech Republic under the project CEITEC 2020, LQ1601 (D.C.).

Conflict of interest

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Y.O. and B.W. supervised the project; S.H., M.D., C.W., B.W. and Y.O. designed experiments; S.H. performed experiments, prepared liver tissue samples and analysed data; C.W. and M.D. performed experiments and analysed data; H.J. provided liver samples and technical assistance; S.K. and D.C. analysed and interpreted TCR deep sequencing data; S.H., C.W. and M.D. wrote the draft and prepared figures; M.D., C.W. S.H. Y.O and B.W. wrote the final manuscript; and all authors provided critical review of the manuscript.

Acknowledgments

We thank all donors and patients who participated in this study, clinical staff at UHB NHS Foundation Trust for recruitment and provision of liver samples and the Anthony Nolan Cell Therapy Centre for provision of cord blood samples. We also thank Dr Matthew McKenzie and the University of Birmingham CMDS Cell Sorting Facility for isolation of $\gamma\delta$ T cells and the University of Birmingham Protein Expression Facility for use of their facilities, and to Dr Yi Pan, the Cancer Research UK Birmingham Centre Biostatistician, for useful discussions.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.jhep.2018.05.007>.

References

Author names in bold designate shared co-first authorship

- [1] Vantourout P, Hayday A. Six-of-the-best: unique contributions of gammadelta T cells to immunology. *Nat Rev Immunol* 2013;13:88–100.
- [2] Di Marco Barros R et al. Epithelia use butyrophilin-like molecules to shape organ-specific gammadelta T cell compartments. *Cell* 2016;167:203–218, e217.
- [3] Swamy M et al. Intestinal intraepithelial lymphocyte activation promotes innate antiviral resistance. *Nat Commun* 2015;6:7090.
- [4] Wencker M et al. Innate-like T cells straddle innate and adaptive immunity by altering antigen-receptor responsiveness. *Nat Immunol* 2014;15:80–87.
- [5] Strid J, Sobolev O, Zafirova B, Polic B, Hayday A. The intraepithelial T cell response to NKG2D-ligands links lymphoid stress surveillance to atopy. *Science* 2011;334:1293–1297.
- [6] Hayday AC. Gammadelta T cells and the lymphoid stress-surveillance response. *Immunity* 2009;31:184–196.
- [7] Schenkel JM, Masopust D. Tissue-resident memory T cells. *Immunity* 2014;41:886–897.
- [8] Hayday AC. [gamma][delta] cells: a right time and a right place for a conserved third way of protection. *Annu Rev Immunol* 2000;18:975–1026.
- [9] Morita CT, Mariuzza RA, Brenner MB. Antigen recognition by human gamma delta T cells: pattern recognition by the adaptive immune system. *Springer Semin Immunopathol* 2000;22:191–217.
- [10] Tanaka Y et al. Nonpeptide ligands for human gamma delta T cells. *Proc Natl Acad Sci U S A* 1994;91:8175–8179.
- [11] Ryan PL et al. Heterogeneous yet stable Vdelta2(+) T-cell profiles define distinct cytotoxic effector potentials in healthy human individuals. *PNAS* 2016;113:14378–14383.

- [12] Willcox CR, Pitard V, et al. Cytomegalovirus and tumor stress surveillance by binding of a human gammadelta T cell antigen receptor to endothelial protein C receptor. *Nat Immunol* 2012;13:872–879.
- [13] Luoma AM et al. Crystal structure of Vdelta1 T cell receptor in complex with CD1d-sulfatide shows MHC-like recognition of a self-lipid by human gammadelta T cells. *Immunity* 2013;39:1032–1042.
- [14] Marlin R et al. Sensing of cell stress by human gammadelta TCR-dependent recognition of annexin A2. *PNAS* 2017;114:3163–3168.
- [15] Zeng X et al. Gammadelta T cells recognize a microbial encoded B cell antigen to initiate a rapid antigen-specific interleukin-17 response. *Immunity* 2012;37:524–534.
- [16] Davey MS, Willcox CR, et al. Clonal selection in the human V delta 1 T cell repertoire indicates gamma delta TCR-dependent adaptive immune surveillance. *Nature Commun* 2017;8.
- [17] Ravens S et al. Human gammadelta T cells are quickly reconstituted after stem-cell transplantation and show adaptive clonal expansion in response to viral infection. *Nat Immunol* 2017;18:393–401.
- [18] Spinozzi F et al. Increased allergen-specific, steroid-sensitive gamma delta T cells in bronchoalveolar lavage fluid from patients with asthma. *Ann Intern Med* 1996;124:223–227.
- [19] Kenna T et al. Distinct subpopulations of gamma delta T cells are present in normal and tumor-bearing human liver. *Clin Immunol* 2004;113:56–63.
- [20] Jeffery HC et al. Biliary epithelium and liver B cells exposed to bacteria activate intrahepatic MAIT cells through MR1. *J Hepatol* 2016;64:1118–1127.
- [21] Huang LR et al. Intrahepatic myeloid-cell aggregates enable local proliferation of CD8(+) T cells and successful immunotherapy against chronic viral liver infection. *Nat Immunol* 2013;14:574–583.
- [22] Jenne CN, Kubes P. Immune surveillance by the liver. *Nat Immunol* 2013;14:996–1006.
- [23] Dunne MR et al. Persistent changes in circulating and intestinal gamma delta T cell subsets, invariant natural killer T cells and mucosal-associated invariant T cells in children and adults with coeliac disease. *PLoS ONE* 2013;8.
- [24] Casorati G, De Libero G, Lanzavecchia A, Migone N. Molecular analysis of human gamma/delta+ clones from thymus and peripheral blood. *J Exp Med* 1989;170:1521–1535.
- [25] Pallett LJ et al. IL-2(high) tissue-resident T cells in the human liver: sentinels for hepatotropic infection. *J Exp Med* 2017;214:1567–1580.
- [26] Jeffery H et al. Mait cells are enriched in portal tracts and respond to biliary epithelial cells presenting bacterial ligands during liver inflammation. *J Hepatol* 2015;62:S793.
- [27] Parekh K et al. Characterisation of circulating and liver infiltrating mait cells in human inflammatory liver diseases. *Gut* 2014;63:A188.
- [28] Gebhardt T, Mueller SN, Heath WR, Carbone FR. Peripheral tissue surveillance and residency by memory T cells. *Trends Immunol* 2013;34:27–32.
- [29] Chien YH, Meyer C, Bonneville M. Gammadelta T cells: first line of defense and beyond. *Annu Rev Immunol* 2014;32:121–155.
- [30] Bonneville M, O'Brien RL, Born WK. Gammadelta T cell effector functions: a blend of innate programming and acquired plasticity. *Nat Rev Immunol* 2010;10:467–478.
- [31] Jameson JM, Cauvi G, Sharp LL, Witherden DA, Havran WL. Gammadelta T cell-induced hyaluronan production by epithelial cells regulates inflammation. *J Exp Med* 2005;201:1269–1279.
- [32] Jameson J et al. A role for skin gammadelta T cells in wound repair. *Science* 2002;296:747–749.
- [33] Boismenu R, Chen Y, Havran WL. The role of intraepithelial gammadelta T cells: a gut-feeling. *Microbes Infect* 1999;1:235–240.
- [34] Asarnow DM, Goodman T, LeFrancois L, Allison JP. Distinct antigen receptor repertoires of two classes of murine epithelium-associated T cells. *Nature* 1989;341:60–62.
- [35] Khairallah C et al. Gammadelta T cells confer protection against murine cytomegalovirus (MCMV). *PLoS Pathog* 2015;11:e1004702.
- [36] Sell S et al. Control of murine cytomegalovirus infection by gammadelta T cells. *PLoS Pathog* 2015;11:e1004481.
- [37] Pitard V et al. Long-term expansion of effector/memory Vdelta2-gammadelta T cells is a specific blood signature of CMV infection. *Blood* 2008;112:1317–1324.
- [38] Wucherpfennig KW et al. Human fetal liver gamma/delta T cells predominantly use unusual rearrangements of the T cell receptor delta and gamma loci expressed on both CD4+CD8- and CD4-CD8- gamma/delta T cells. *J Exp Med* 1993;177:425–432.
- [39] Gerlach C et al. The chemokine receptor CX3CR1 defines three antigen-experienced CD8 T cell subsets with distinct roles in immune surveillance and homeostasis. *Immunity* 2016;45:1270–1284.
- [40] Hertoghs KM et al. Molecular profiling of cytomegalovirus-induced human CD8+ T cell differentiation. *J Clin Invest* 2010;120:4077–4090.
- [41] Pachnio A et al. Cytomegalovirus infection leads to development of high frequencies of cytotoxic virus-specific CD4+ T cells targeted to vascular endothelium. *PLoS Pathog* 2016;12:e1005832.
- [42] Geissmann F et al. Intravascular immune surveillance by CXCR6+ NKT cells patrolling liver sinusoids. *PLoS Biol* 2005;3:e113.
- [43] Stegmann KA et al. CXCR6 marks a novel subset of T-bet(lo)Eomes(hi) natural killer cells residing in human liver. *Sci Rep* 2016;6:26157.
- [44] Hudspeth K et al. Human liver-resident CD56(bright)/CD16(neg) NK cells are retained within hepatic sinusoids via the engagement of CCR5 and CXCR6 pathways. *J Autoimmun* 2016;66:40–50.
- [45] Stelma F et al. Human intrahepatic CD69 + CD8+ T cells have a tissue resident memory T cell phenotype with reduced cytolytic capacity. *Sci Rep* 2017;7:6172.
- [46] Agrati C et al. Vdelta1 T lymphocytes expressing a Th1 phenotype are the major gammadelta T cell subset infiltrating the liver of HCV-infected persons. *Mol Med* 2001;7:11–19.
- [47] Almeida AR et al. Delta one T cells for immunotherapy of chronic lymphocytic leukemia: clinical-grade expansion/differentiation and preclinical proof of concept. *Clin Cancer Res* 2016;22:5795–5804.
- [48] Gentles AJ et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med* 2015;21:938–945.
- [49] Gomes AQ, Martins DS, Silva-Santos B. Targeting gammadelta T lymphocytes for cancer immunotherapy: from novel mechanistic insight to clinical application. *Cancer Res* 2010;70:10024–10027.
- [50] Wu P et al. GammadeltaT17 cells promote the accumulation and expansion of myeloid-derived suppressor cells in human colorectal cancer. *Immunity* 2014;40:785–800.

Human liver infiltrating $\gamma\delta$ T cells are composed of clonally expanded circulating and tissue-resident populations

Stuart Hunter, Carrie R. Willcox, Martin S. Davey, Sofya A. Kasatskaya, Hannah C. Jeffery, Dmitriy M. Chudakov, Ye H. Oo, Benjamin E. Willcox

Table of contents

Fig. S1.....	2
Fig. S2.....	3
Fig. S3.....	4
Fig. S4.....	5
Fig. S5.....	6
Table S1.....	7
Table S2.....	8

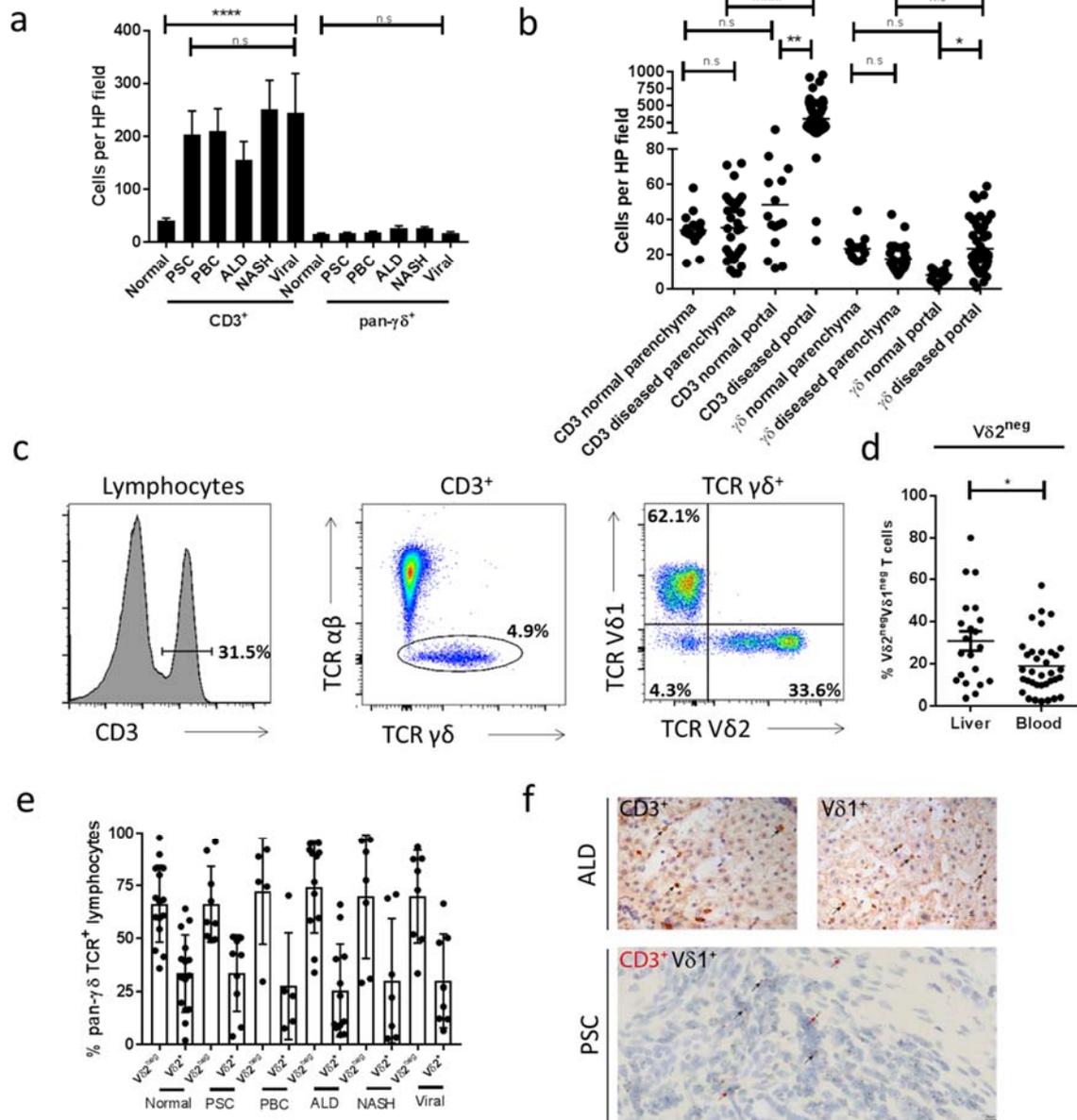


Fig. S1. Intrahepatic $V\delta 2^{neg}$ $\gamma\delta$ T cell proportions in diseased liver tissue.

(a) Densities of $CD3^+$ and $\gamma\delta$ TCR $^+$ cells in normal (n=21) PSC (n=13), PBC (n=13, ALD (n=12), NASH (n=12) and viral hepatitis (n=12) liver tissue. (b) Densities of $CD3^+$ and $\gamma\delta$ TCR $^+$ cells in parenchymal and portal areas from normal (n=15) and diseased (n=30) liver tissue. (c) Representative flow cytometry plots showing gating strategy used to identify intrahepatic $CD3^+$ lymphocytes \Rightarrow $\gamma\delta$ TCR $^+$ $\alpha\beta$ TCR neg \Rightarrow $V\delta 1^+$, $V\delta 2^+$ and $V\delta 1^{neg}$ $V\delta 2^{neg}$ T cells. (d) Comparison of $V\delta 2^{neg}$ $V\delta 1^{neg}$ $\gamma\delta$ T cells in explanted liver (n=21) and peripheral blood (n=35). (e) Comparison of $V\delta 2^+$ and $V\delta 2^{neg}$ proportions in $\gamma\delta$ TCR $^+$ cells identified by flow cytometry from normal (n=15), PSC (n=9), PBC (n=5), ALD (n=12), NASH (n=7) and viral hepatitis (n=8) liver cell suspensions. (f) Representative staining for $CD3^+$ (upper left) and $V\delta 1^+$ TCR $^+$ (upper right) cells on sequential FFPE sections from ALD liver tissue viewed at 20 \times magnification, and $CD3$ (brown) and $V\delta 1$ TCR (red) transcripts by in-situ hybridisation of PSC tissue viewed at 40x magnification. Error bars indicate mean \pm SEM; data analysed by Kruskal–Wallis ANOVA with Dunn’s post-test comparisons (a-c) and Student’s T test (e), n.s. $p>0.05$, * $p<0.05$, ** $p<0.01$ and **** $p<0.0001$.

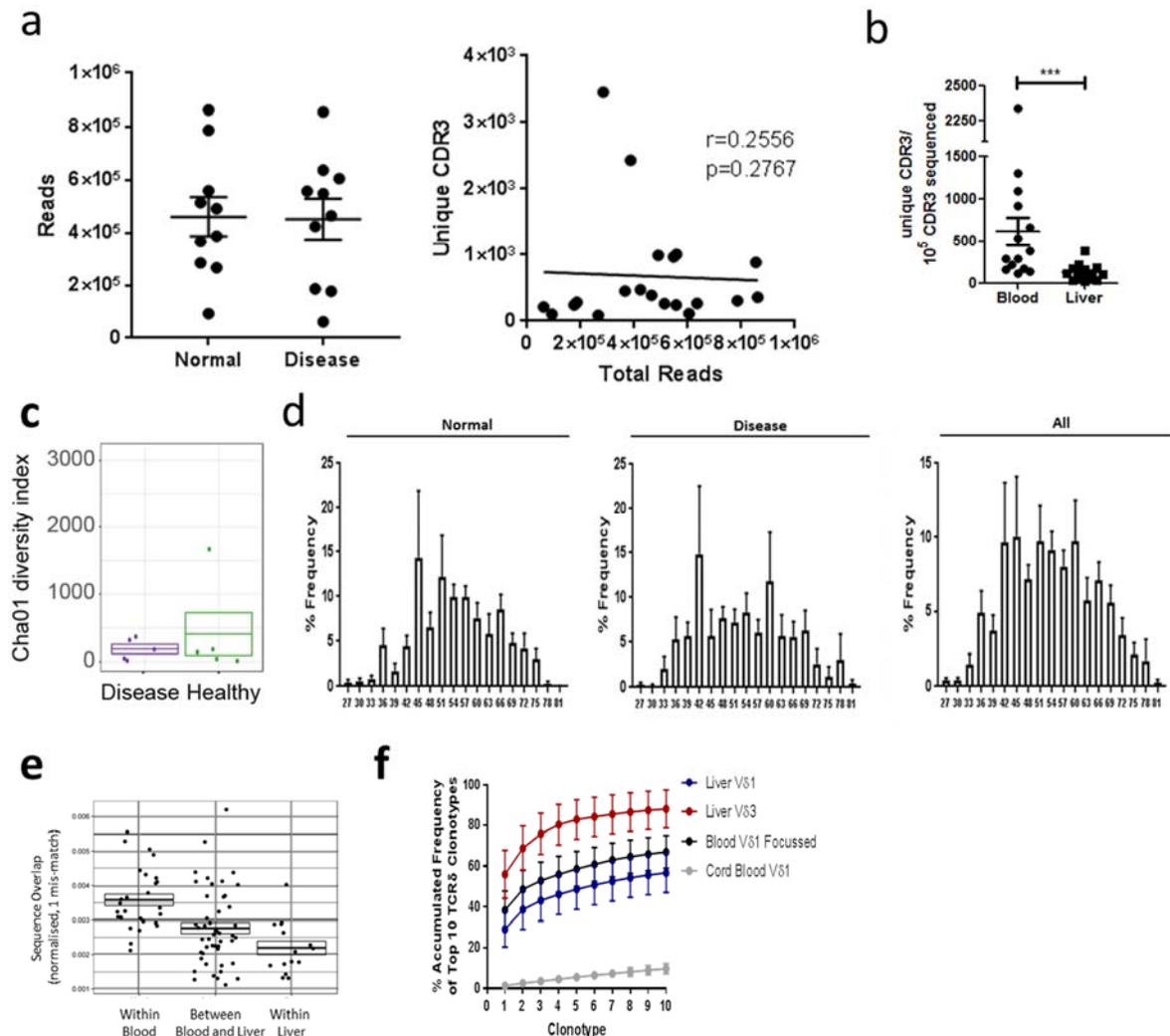


Fig. S2. TCR repertoire sequence analysis of intrahepatic $V\delta 2^{\text{neg}} \gamma\delta$ T cells.

(a) Comparison of combined CDR3 δ and CDR3 γ sequencing reads from $V\delta 2^{\text{neg}}$ T cells isolated from normal (n=5) and diseased (n=5) livers (left) and total sequencing reads against the number of unique CDR3s determined for each donor (n=10). (b) Comparison of the number of unique CDR3 δ and CDR3 γ sequences determined per 10^5 sequences from blood (n=7) and liver (n=10) derived $V\delta 2^{\text{neg}}$ T cells. (c) Comparison of Cha1 total $V\delta 2^{\text{neg}}$ T cell repertoire richness estimation (diversity) from diseased and normal livers. (d) Non-normalised length spectratyping for CDR3 δ in normal (left, n=5), disease (centre, n=5) and combined (right, n=10) intrahepatic $V\delta 2^{\text{neg}}$ T cells. (e) Comparison of the proportion of sequences in the TCR γ repertoire that are observed in more than one donor (allowing up to one amino acid mismatch) within blood samples (n=7), liver samples (n=10), or in between liver and blood samples. (f) Accumulated frequencies occupied by the 10 most prevalent TCR $V\delta 1^+$ and $V\delta 3^+$ clonotypes in liver (n=10) and $V\delta 1^+$ “focussed” blood (n=13) and cord blood (n=5) donors (both from Davey *et al.*, 2017). Error bars indicate mean \pm SEM; data analysed by spearman correlation and student’s T test, ***p<0.001.

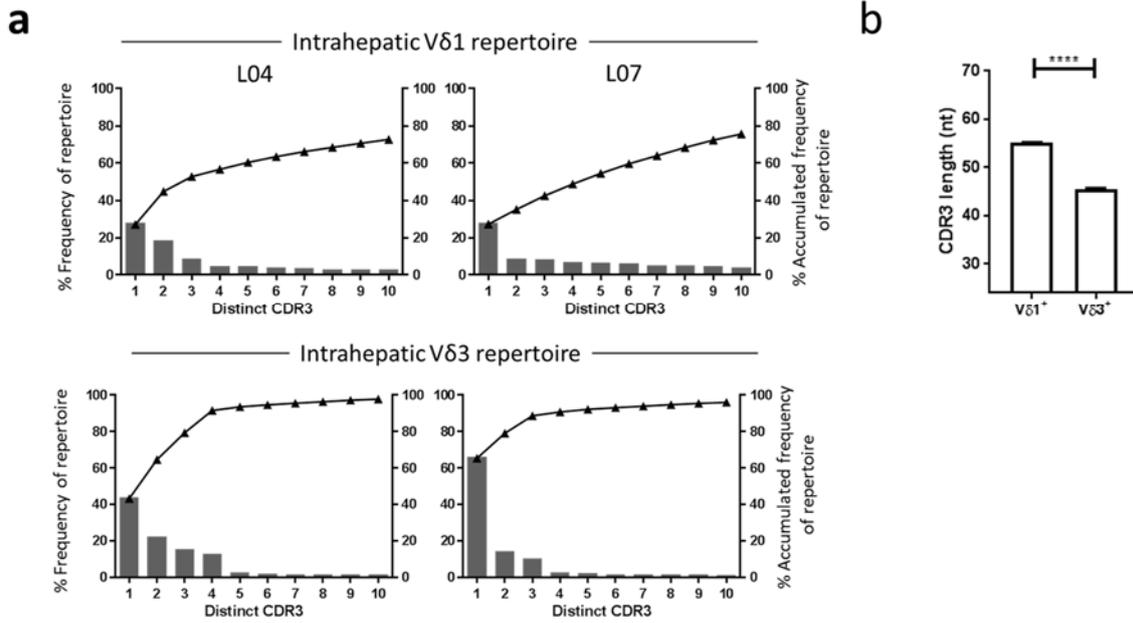


Fig. S3. Concurrent clonal focussing in V δ 2^{neg} $\gamma\delta$ T cell subsets.

(a) Individual clone frequency (left y axis) and accumulated frequency (right y axis) for the 10 most prevalent TCR V δ 1⁺ (upper) and V δ 3⁺ (lower) clonotypes from two liver donors with appreciable V δ 3⁺ T cell populations. (h) Comparison of CDR3 lengths for V δ 1⁺ and V δ 3⁺ TCRs in liver samples (n=10). Error bars indicate mean \pm SEM; data analysed by student's T test, ****p<0.0001.

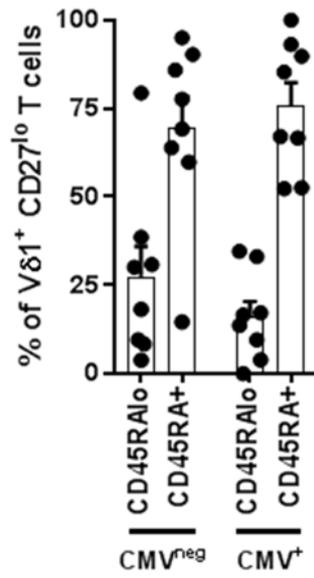


Fig. S4. CMV infection does not influence proportion of CD45RA^{lo} Vδ1⁺ γδ T cells.

Comparison of intrahepatic CD27^{lo} Vδ1⁺ T cell expression of CD45RA according to CMV status of liver donor (n=8 CMV^{neg} & CMV⁺).

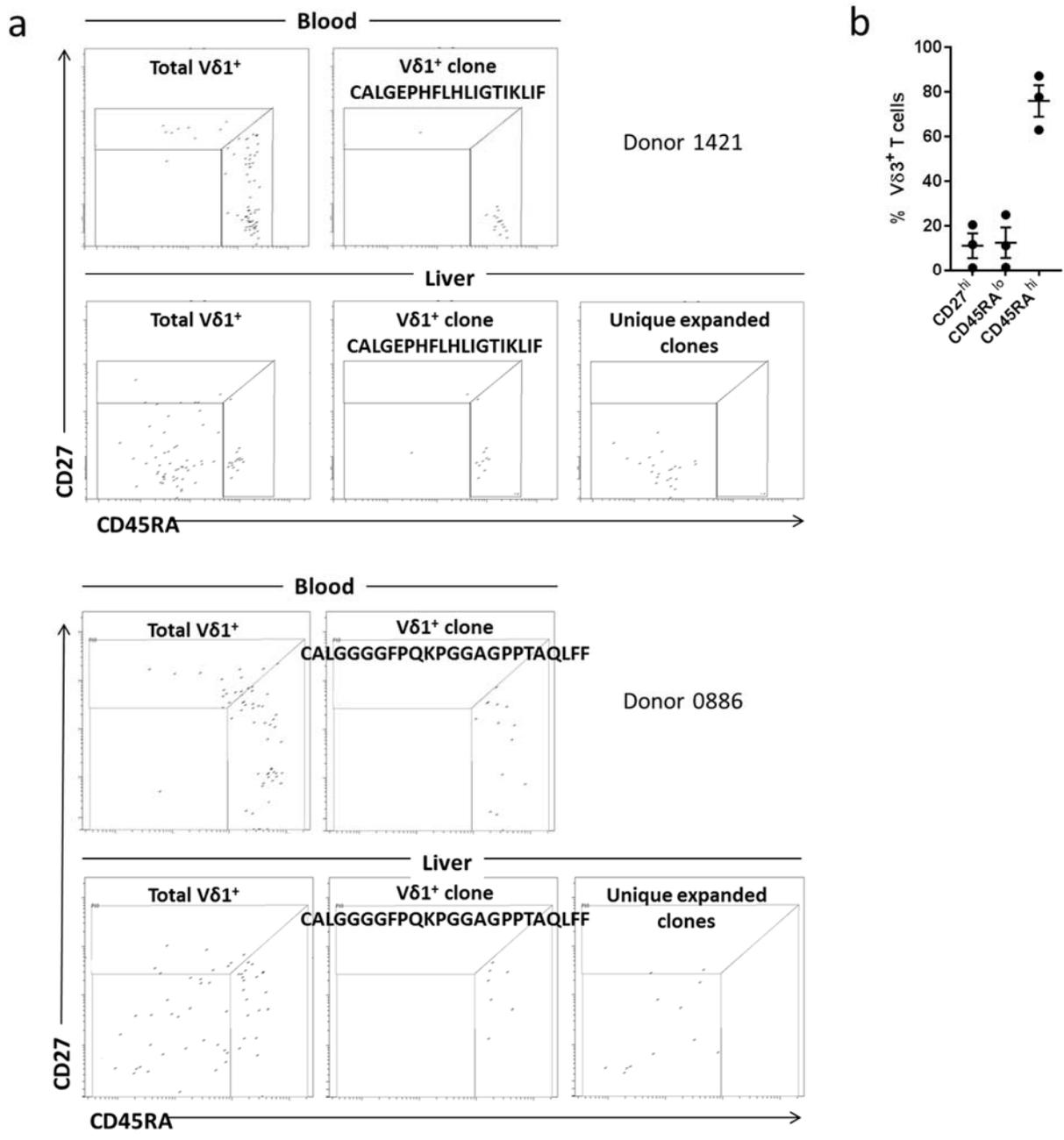


Fig. S5. Phenotype-linked indexed single cell TCR sequencing analysis.

(a) Representative flow cytometry plots from two donors showing expression of CD27 and CD45RA by single Vδ1⁺ T cells sorted from donor matched blood and liver samples. Single cells were subsequently analysed by TCR sequencing and plots are shown for all cells (left), cells with most prevalent CDR3 sequence found in both liver and blood samples (centre) and cells identified only liver samples that displayed CDR3 sequences with at least one other cell (right). (b) Summary data of intrahepatic Vδ3⁺ γδ T cells phenotype according to expression of CD27 and CD45RA (n=3).

Chapter VI

$\gamma\delta$ T-cell receptors derived from breast
cancer-infiltrating T lymphocytes
mediate antitumor reactivity

Introduction

Chapter VI continues the previous chapters with $\gamma\delta$ TCR repertoire research. This study results from a long-term collaboration with Prof. Jurgen Kuball's laboratory, which shares exquisite expertise in fundamental $\gamma\delta$ T cell biology and adoptive T cell therapy approaches. The unique MHC-independent mode of gamma/delta TCR antigen recognition makes it a useful tool for anticancer T cell therapy development. Moreover, suppose the T cell receptor variants with anticancer activity are present in different donors' immune systems. In that case, it is more likely that such T cells would be safer to use in adoptive transfer. Therefore in this study, we explored the public part of $\gamma\delta$ TCR repertoire, which is present in unrelated donors and was observed in independent $\gamma\delta$ TCR datasets. We investigated breast cancer patient's samples for gamma/delta TILs and found a significant proportion of public TCR clonotypes. Chapter VI is an example of how computational TCR repertoires studies could serve at the primary exploratory stages of T cell anticancer therapy development either from TILs or from the donor-derived T lymphocytes.

Contribution

In this collaboration, I adapted the TCRbeta sequencing protocol created earlier in our laboratory to the needs of TCR delta chain sequencing in Dr. Kuball's lab. It made possible the acquisition of TCR delta repertoires from tumor-infiltrating $\gamma\delta$ T cells. In addition to this new dataset, the current study used previously published $\gamma\delta$ TCR sequencing datasets, including those obtained in our collaboration with the University of Birmingham (see Chapters IV and V). As a result, in Chapter VI, we obtained a list of public δ TCR clonotypes with independent validation from different datasets and sequencing platforms. I contributed to the discussion and interpretation of the public clonotypes occurrence. To analyze the long delta TCR chains' rearrangement patterns, I created a script to align CDR3 delta sequences to several D segments in a row since it is a frequent phenomenon in the TCR delta chain. Proposed techniques for more accurate TCR gamma/delta sequencing data analysis and the continuous re-analysis of publicly available datasets both contribute to the development of standards in the immune repertoires research field.

$\gamma\delta$ T-cell Receptors Derived from Breast Cancer-Infiltrating T Lymphocytes Mediate Antitumor Reactivity



Anke Janssen¹, Jose Villacorta Hidalgo^{2,3,4}, Dennis X. Beringer¹, Sanne van Dooremalen¹, Febilla Fernando¹, Eline van Diest¹, Antonela R. Terrizi^{2,4}, Peter Bronsert^{2,4,5,6}, Sylvia Kock^{2,4}, Annette Schmitt-Gräff^{2,4}, Martin Werner^{2,4,5,6}, Kerstin Heise⁷, Marie Follo^{4,8}, Trudy Straetemans¹, Zsolt Sebestyén¹, Dmitry M. Chudakov^{9,10,11}, Sofya A. Kasatskaya^{10,11}, Felix E. Frenkel¹², Sarina Ravens¹³, Eric Spierings¹, Immo Prinz¹³, Ralf Küppers⁷, Miroslav Malkovsky¹⁴, Paul Fisch^{2,4,†}, and Jürgen Kuball^{1,15}

ABSTRACT

$\gamma\delta$ T cells in human solid tumors remain poorly defined. Here, we describe molecular and functional analyses of T-cell receptors (TCR) from tumor-infiltrating $\gamma\delta$ T lymphocytes ($\gamma\delta$ TIL) that were in direct contact with tumor cells in breast cancer lesions from archival material. We observed that the majority of $\gamma\delta$ TILs harbored a proinflammatory phenotype and only a minority associated with the expression of IL17. We characterized TCR γ or TCR δ chains of $\gamma\delta$ TILs and observed a higher proportion of V δ 2⁺ T cells compared with other tumor types. By reconstructing matched V δ 2⁻ TCR δ and TCR δ pairs derived from single-cell

sequencing, our data suggest that $\gamma\delta$ TILs could be active against breast cancer and other tumor types. The reactivity pattern against tumor cells depended on both the TCR γ and TCR δ chains and was independent of additional costimulation through other innate immune receptors. We conclude that $\gamma\delta$ TILs can mediate tumor reactivity through their individual $\gamma\delta$ TCR pairs and that engineered T cells expressing TCR γ and δ chains derived from $\gamma\delta$ TILs display potent antitumor reactivity against different cancer cell types and, thus, may be a valuable tool for engineering immune cells for adoptive cell therapies.

Introduction

Triple-negative breast cancer (TNBC) has a very poor prognosis compared with other breast cancer subtypes. Despite encouraging

results in TNBC patients treated with checkpoint inhibitors (1), the majority of patients at advanced stages of the disease do not respond to therapies, and the substantial genomic heterogeneity of these tumors makes it difficult to identify new therapeutic targets (2–4). One novel opportunity for therapy arises from the observation that $\gamma\delta$ T cells infiltrate various tumors, including TNBCs (5, 6), and these infiltrates appear to be prognostically beneficial (7). However, the number and function of $\gamma\delta$ T cells are substantially diminished in advanced cancer patients (8, 9), suggesting that the possible $\gamma\delta$ T-cell immunosurveillance activity in early cancers may become dysfunctional at the later stages of cancer. Some studies have found that IL17-producing $\gamma\delta$ tumor-infiltrating lymphocytes (TIL) can create a tumor-promoting environment (10, 11). In contrast, another in-depth study suggests that $\gamma\delta$ TILs are not the main producers of IL17 in TNBC and rather support antitumor reactivity in breast cancer through innate-like receptors (ref. 12; for review see refs. 13, 14). However, failures of polyclonal $\gamma\delta$ T cells to induce regressions in advanced cancer patients have been reported (14, 15). To overcome this system failure in advanced cancer patients, we developed the concept of TEGs (T cells engineered to express a defined TCR $\gamma\delta$) for cancer therapies (16–18). For example, TEGs can distinguish between healthy and malignant hematopoietic stem cells by sensing altered lipid pathways present in malignant cells through TCR $\gamma\delta$ 2, and by detecting subtle spatial and conformational changes of CD277 (ref. 19; for review see ref. 14).

To identify tumor-reactive TCR $\gamma\delta$ s, we analyzed TCR sequences in $\gamma\delta$ TILs from TNBCs. We also analyzed TCR δ repertoires in TIL sequencing data sets and in sequencing data sets of peripheral blood from healthy individuals. We combined single-cell sequencing (SCS) of laser microdissected $\gamma\delta$ TILs, targeted high-throughput sequencing (HTS) of TCR $\gamma\delta$ repertoires, and TCR $\gamma\delta$ extraction from bulk tumor RNA sequencing (RNA-seq) data using the MiXCR RNA-seq mode (20). Subsequently, we utilized the TEG format for functional analysis of TCR $\gamma\delta$ s and observed tumor reactivity of matched TCR $\gamma\delta$ pairs against a variety of tumor cells.

¹Laboratory of Translational Immunology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands. ²Institute for Surgical Pathology, University Medical Center, University of Freiburg, Freiburg, Germany. ³Faculty of Biology, University of Freiburg, Freiburg, Germany. ⁴Faculty of Medicine, University of Freiburg, Freiburg, Germany. ⁵German Cancer Consortium (DKTK) and Cancer Research Center (DKFZ), Heidelberg, Germany. ⁶Comprehensive Cancer Center Freiburg, Medical Center – University of Freiburg, Freiburg, Germany. ⁷Institute of Cell Biology (Cancer Research), University of Duisburg-Essen, Essen, Germany. ⁸Department of Medicine I, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany. ⁹Privolzhsky Research Medical University, Nizhny Novgorod, Russia. ¹⁰Center for Data-Intensive Biomedicine and Biotechnology, Skolkovo Institute of Science and Technology, Moscow, Russia. ¹¹Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia. ¹²BostonGene LLC, Lincoln, Nebraska. ¹³Institute of Immunology, Hannover Medical School, Hannover, Germany. ¹⁴UW School of Medicine and Public Health, Madison, Wisconsin. ¹⁵Department of Hematology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands.

Note: Supplementary data for this article are available at Cancer Immunology Research Online (<http://cancerimmunolres.aacrjournals.org/>).

†Deceased.

A. Janssen, J. Villacorta Hidalgo, and D.X. Beringer contributed equally to this article.

P. Fisch and J. Kuball share senior authorship.

Corresponding Author: Jürgen Kuball, UMC Utrecht, Heidelberglaan 100, Utrecht 3508 GA, the Netherlands. Phone: 31-88-755-9030; Fax: 31-88-755-4305; E-mail: j.h.e.kuball@umcutrecht.nl

Cancer Immunol Res 2020;8:530–43

doi: 10.1158/2326-6066.CIR-19-0513

©2020 American Association for Cancer Research.

Materials and Methods

Human subjects

Peripheral blood samples from 13 anonymous healthy donors were obtained via the Dutch blood bank (Sanquin). From the archive of the Institute of Clinical Pathology, Freiburg, 16 formalin-fixed paraffin-embedded tissue specimens with the diagnosis of TNBC were selected after ethics approval by local authorities (Ethics Committee University Medical Center Freiburg) according to the Declaration of Helsinki, and written informed consent was obtained from each patient. The histopathologic diagnoses were established by two independent pathologists (P. Bronsert and A. Schmitt-Gräff) according to the Union for International Cancer Control criteria. All the tumors were grade III using the modified Bloom-Richardson classification (21). Conforming to the recommended evaluation of TILs (22, 23), the hematoxylin and eosin (H&E)-stained samples contained at least 50% tumor infiltration. Subsequently, 11 tumors (5 medullary breast carcinomas and 6 invasive ductal carcinomas) were selected for further analysis based on the pattern of infiltration and the availability of the corresponding frozen tissue samples in the tumor bank of the Comprehensive Cancer Center Freiburg. All samples came from female patients with a median age of 59 years (range, 43–82 years). All samples were classified as having a basal-like subtype based on the expression of CK 5/6 or 14 and EGFR (24). The use of the TNBC patient samples was approved by the ethics committee of the Medical Center University of Freiburg.

Cell lines

Origin and testing of cell lines

The following cell lines were obtained from ATCC between 2010 and 2018: BT549 (HTB-122), Daudi (CCL-213), HCC38 (CRL-2314), HeLa (CCL-2), HEK293T (CRL-3216), HEPG2 (HB-8065), HT29 (HTB-38), K562 (CCL-243), LS123 (CCL-255), MCF-7 (HTB-2), OvCa (HTB-161), Phoenix-AMPHO (ATCC, CRL-3213), Saos2 (HTB-85), SCC9 (CRL-1629), and U937 (CRL-1593.2). Freestyle 293-F cells were obtained from Invitrogen (R790-07). The following cell lines were kindly provided by Ilan Bank (Chaim Sheba Medical Center, Tel Hashomer, Israel) in 2014: BT-474, MDA-MB231, and T47D. The following cell lines were kindly provided by Lisa Wiesmüller (Universitätsfrauenklinik, Ulm, Germany) in 2014: MDA-MB-134, MDA-MB-436, MDA-MB468, and UACC-893. Thordur Oskarsson (Deutschen Krebsforschungszentrum, Heidelberg, Germany) kindly provided cell line MDA-MB157 in 2016. Phil Greenberg (Fred Hutchinson Cancer Research Center, Seattle, USA) kindly provided cell line LCL-TM. Barbara Seliger (University of Halle, Germany) kindly provided cell line MZ1851rc.

The following cell lines were reauthenticated using the Cell Line Authentication (CLA) Test provided by Eurofins Genomics Europe in 2017: Daudi, HEK239T, HeLa, K562, LCL-TM, MDA-MB-231, MZ1851rc, Phoenix-AMPHO, Saos2, SCC9, and T-47D. The following cell lines were reauthenticated using the CLA Test provided by Eurofins Genomics Europe (Ebersberg, Germany) in 2019: BT474, Daudi, HCC38, HT29, K562, LCL-TM, MCF7, MDA-MB-231, MZ1851rc, OvCa, and Phoenix-AMPHO. *Mycoplasma* testing was done regularly in house using the Myco alert Mycoplasma detection kit (Lonza, LT07-318). Cell lines were used in assays from passage 4 up to 25 after thawing.

Cell line culture conditions

The Daudi, K562, U937, and LCL-TM cell lines were cultured in the RPMI-1640 medium (Thermo Fisher, 72400054) with 10% fetal calf

serum (FCS; Sigma-Aldrich, F7524), penicillin (100 IU/mL), and streptomycin (100 μ g/mL; Thermo Fisher, 15140163) at 37°C. The BT474 cell line was cultured in the RPMI-1640 medium with 10% FCS, penicillin (100 IU/mL), and streptomycin (100 μ g/mL), and additionally supplemented with 1 mmol/L sodium pyruvate (Thermo Fisher, 11530396) and bovine insulin (0.01 mg/mL; Sigma-Aldrich, I6634-50MG) at 37°C. The cell lines MZ1851RC, BT549, HCC38, HeLa, HEK293T, HEPG2, HT29, LS123, OvCa, Phoenix-AMPHO, Saos2, and SCC9 were cultured in DMEM with 10% FCS, penicillin (100 IU/mL), and streptomycin (100 μ g/mL) at 37°C. The MCF-7, MDA-MB-134, MDA-MB-157, MDA-MB-436, MDA-MB-468, T-47D, and UACC-893 cell lines were cultured in the DMEM with 10% FCS, penicillin (100 IU/mL), streptomycin (100 μ g/mL), glutamine (2 mmol/L; Thermo Fisher, 25030024), 1% nonessential amino acids (100 mmol/L; Thermo Fisher, 12084947), human insulin (4 μ g/mL; Sigma-Aldrich, I2643), and EGF (10 ng/mL; Sigma-Aldrich, SRP3027) at 37°C. The MDA-MB-231 cell line was cultured in the IMDM (Thermo Fisher, 31980048) with 10% FCS, penicillin (100 IU/mL), streptomycin (100 μ g/mL), and glutamine (2 mmol/L) at 37°C. Freestyle 293-F cells were cultured in FreeStyle 293 Expression Medium (Thermo Fisher, 12338026) at cell densities between 0.3×10^6 and 2.5×10^6 in a shaking incubator at 37°C.

Peripheral blood mononuclear cells (PBMC), T cells engineered to express a defined TCR $\gamma\delta$ (TEG), natural killer T (NKT) cells engineered to express a defined TCR $\gamma\delta$ (NEG), and $\gamma\delta$ T cells engineered to express a defined TCR $\gamma\delta$ (GEG) were cultured in the RPMI-1640 medium supplemented with 5% human serum (Sanquin), penicillin (100 IU/mL), streptomycin (100 μ g/mL), and 50 μ mol/L 2-mercaptoethanol (Merck).

HTS of the TCR δ repertoire in healthy donors

Our sequencing protocol has been modified from Mamedov and colleagues (25). Frozen PBMC samples from healthy donors were thawed and stained with human anti-CD3 eFluor 450 (eBioscience, 16-0037-81), anti-TCR $\alpha\beta$ APC (eBioscience, 17-9986-42), and anti-TCR $\gamma\delta$ PE (Beckman Coulter, IM1349), and the $\gamma\delta$ fractions (CD3⁺ $\gamma\delta$ ⁺) were sorted on a BD FACSAria II cell sorter. RNA was isolated from samples with $\geq 0.5 \times 10^6$ cells with the RNeasy Mini Kit (QIAGEN, 74204) and from samples with $< 0.5 \times 10^6$ cells with the RNeasy Micro Kit (QIAGEN, 74004). TCR δ cDNA was synthesized using SuperScript II Reverse Transcriptase (Thermo Fisher Scientific), utilizing a specific primer at the 3' constant region and a universal template switch adaptor at the 5' end of the V region (for the specific primers, see Supplementary Table S1), and purified using NucleoSpin Gel and PCR Clean-UP (Macherey-Nagel). The samples were amplified in a first round of PCR using Q5 High-Fidelity DNA polymerase (New England BioLabs, Inc.) and a T100 Thermal Cycler (Bio-Rad) under the following cycling conditions: 90 seconds at 98°C, 15 cycles of 7 seconds at 98°C, 20 seconds at 62°C, and 50 seconds at 72°C, followed by 10 minutes at 72°C. A specific nested primer located in the constant region and a step-out primer, which anneals to the switch adaptor, were used (Supplementary Table S1).

The resulting amplicons were loaded onto a 1.5% agarose gel, electrophoresed, and products between 400 and 600 base pairs were size selected. After purification of the gel with NucleoSpin Gel and PCR Clean-UP, the PCR products were used for a second PCR using a reverse nested primer on the constant region and a forward primer which annealed to the switch adaptor (Supplementary Table S1), using the following cycling parameters: 90 seconds at 98°C, 20 cycles of 7 seconds at 98°C, 20 seconds at 62°C, and 50 seconds at 72°C, followed by 10 minutes at 72°C. After purification with NucleoSpin Gel and

PCR Clean-UP, library preparations for HTS were carried out using the HTSgo-LibrX kit with HTSgo-IndX indices following the manufacturer's (Gendx) recommendations. Subsequently, samples were cleaned up (HighPrep PCR beads, GC Biotech), and HTS was performed on a MiSeq System 500 (2×250 bp read lengths), Illumina. Analysis of the HTS data is described in "Analysis of HTS data sets and availability of data sets."

HTS of the TCR δ repertoire in TNBC TILs

RNA samples from TNBC tumors were reverse transcribed into cDNA using Superscript III enzyme (Invitrogen) and oligo(dT) primers according to the manufacturer's protocol. Multiplex primer sequences of *hTRDV1*, *hTRDV2*, *hTRDV3*, *TRDV5/29*, and *hTRDC1* for δ -chain were used to prepare the HTS library (see Supplementary Table S1). Adaptor sequences were added as overhangs to facilitate the Illumina MiSeq analyses. All primers were mixed in equal amounts to achieve a final concentration of 10 μ mol/L. The total volume of each PCR reaction was 20 μ L consisting of 1 \times PCR buffer (without $MgCl_2$; Invitrogen), 1.5 mmol/L $MgCl_2$, 10 mmol/L dNTPs (QIAGEN, 201900), 0.5 μ mol/L forward primer mix, 0.5 μ mol/L reverse primer, and 0.04 units of recombinant Taq DNA polymerase (Invitrogen). Cycling conditions were 3 minutes at 95°C, 30 seconds at 95°C, 30 seconds at 63°C, and 30 seconds at 72°C for 5 cycles; 30 seconds at 95°C and 35 seconds at 72°C for 20 to 25 cycles; and 4 minutes at 72°C. After electrophoresing amplicons on 1% TAE-agarose, 250-bp-sized bands were excised and purified (QIAquick Gel Extraction Kit, Qiagen). For paired-end Illumina sequencing, additional adapters (indicated below), including a flow cell binding site for sequencing on the Illumina MiSeq System and indices for demultiplexing, were introduced to the TCR amplicons. For this, the Nextera Index Kit (Illumina) primers were used in a 50- μ L PCR reaction with the following components: 1 \times Advantage 2 PCR Buffer (Clontech), 1 μ mol/L of Index 1 and Index 2 (Nextera Index Kit; Illumina), 1 \times dNTP Mix (10 mmol/L each; Clontech), 30 to 40 ng DNA template, and 1 \times Advantage 2 Polymerase Mix (Clontech). PCR products were purified using the Agencourt AMPure XP PCR Purification Kit (Beckman Coulter) according to the manufacturer's recommendation. Further preparation and loading of the library to the MiSeq System was done according to the "denature and dilution guide" provided by Illumina. Analysis of the HTS data is described in "Analysis of HTS data sets and availability of data sets."

IHC and immunofluorescence imaging of TNBC samples

Serial FFPE 2- μ m thick sections mounted on SuperFrost plus glass slides (R Langenbrinck GmbH) were dewaxed and rehydrated. After proper antigen retrieval in a pressure cooker with citrate buffer (pH 6) and citrate buffer (pH 6.1; Dako), blocking of nonspecific binding was performed using 5% normal goat serum in PBS (Jackson ImmunoResearch, 005-000-121). Mouse monoclonal anti-TCR γ -chain (Thermo Fisher, TCR1153) and rabbit anti-human cleaved caspase-3 polyclonal antiserum (Cell Signaling Technology, 9664) were used, as we previously reported (6). The human cytomegalovirus (CMV) detection was performed using mouse anti-CMV (clones CCH2 + DDG9, DAKO, and IS752) as described below. Alkaline phosphatase-conjugated and horseradish peroxidase-conjugated detection systems were used to visualize the primary antibodies in separate or sequential protocols for single or double staining tests with red and brown chromogen (DakoREALDetection System Alkaline Phosphatase/RED r/m and EnVisionFLEX Systems, Dako). Acidic hematoxylin was used as a counterstain.

Frozen serial sections of 5- μ m thickness were mounted on SuperFrost plus glass slides, air-dried for 3 hours, and fixed with precooled acetone (-20°C) for 10 minutes. Samples were rinsed with $3 \times$ TBST for 5 minutes, and blocking of nonspecific binding was performed using 5% human serum in PBS for 30 minutes. The samples were incubated with the corresponding primary mouse anti-human mAbs: anti-TCR γ (Thermo Fisher, TCR1153) and anti-CD69 (Leica Microsystems, NCL-CD69), or goat anti-human polyclonal anti-IFN γ (R&D Systems, AF-285-NA), anti-TNF α (Novus bio, NBP1-19532), and anti-IL17 (R&D Systems, AF-317-NA) antisera. Fluorescence-conjugated secondary antibodies used to visualize the primary antibodies were rabbit anti-mouse AlexaFluor 488 (TCR γ ; Thermo Fisher, A-11059), donkey anti-rabbit AlexaFluor 568 (CD4; Abcam, ab175694), donkey anti-goat AlexaFluor 594 (CD69, IFN γ , and TNF α ; Abcam, AB150132), donkey anti-goat AlexaFluor 647 (IL17, pseudocolored; Abcam, AB150131). Samples were mounted using ProLong Diamond Antifade medium with DAPI (Thermo Fisher).

Samples were tile scanned using a Zeiss Axio Observer 7 with Apotome 2 system with an ERc 5s digital camera and analyzed using the AxioVision 4.8 and ZEN BLUE image software (all from Carl Zeiss). Four hundred randomly selected cells positive for TCR γ were counted manually in serial slides and evaluated for colocalization with IFN γ , TNF α , or IL17 with the ImageJ (NIH images) and QuPath (GitHub) softwares and the Bio-Formats, Stack Slicer, and Cell counter plugins (26) by two independent researchers (J. Villacorta Hidalgo and A.R. Terrizi).

CMV detection

The breast cancer cell lines and frozen tumor material were tested for CMV using an IHC-nested PCR according to a previously published protocol (27), and qPCR with the artus CMV TM PCR Kit (Qiagen) in a 7900HT Fast Real-Time PCR cyclers (Applied Biosystems) according to the manufacturer's instructions (for primer sequences, see Supplementary Table S1).

Laser capture microdissection of TNBC samples

Frozen sections (8 μ m thick) were air dried overnight on MembraneSlide 1.0 PEN membrane covered slides (Carl Zeiss) fixed in precooled acetone (-20°C) for 10 minutes, washed twice with TBST, incubated first for 30 minutes with 5% human serum in PBS, and then for an additional 30 minutes at room temperature with mouse anti-human TCR γ (100 μ L, 1:40; Thermo Fisher, TCR1153). To detect the mouse mAb binding, a biotinylated anti-mouse secondary antibody was used from the alkaline phosphatase detection system (Dako REAL Detection System Alkaline P/RED, rabbit-mouse), and the slides were counterstained with Mayer's hematoxylin. Samples were dried at room temperature for 2 hours, examined with an Axiovert light microscope (Carl Zeiss) for TCR γ^+ TIL, and stored at 4°C until processing.

Laser microdissection was performed using an Axiovert microscope equipped with the PALM MicroBeam system (Carl Zeiss). Energy parameters of cutting and catapulting were established individually for each sample. Only infiltrating single lymphocytes in contact with cancer cells were selected and then microdissected. The selection and microdissection processes were performed at 40 \times and 63 \times magnification, respectively. Single cells were catapulted into the cap of an Adhesive Cap 500 opaque 500 μ L tube (Carl Zeiss). Twenty microliters of a 1:10 mix containing Expand High-Fidelity Buffer and proteinase K (20 mg/mL, PCR grade), both from Roche Diagnostics) was then added for digestion. The tubes were incubated with their lids closed for 4 hours at 56°C, centrifuged for 2 minutes at 500 rpm, and heat

inactivated at 95°C for 10 minutes. Additional tubes containing only membrane were used as negative controls. All PCR tubes were overlaid with mineral oil in a laminar flow hood before adding the PCR master mix.

CDR3 $\gamma\delta$ spectratyping of TNBC $\gamma\delta$ TILs

To assess the clonality of $\gamma\delta$ T cells by the length of the CDR3 regions, we used spectratyping analysis. RNA was extracted from frozen tumor tissue using RNeasy Tissue minikit (Qiagen), and cDNA reverse transcribed by Oligo-dT-primers (First-Strand cDNA Synthesis Kit, GE Healthcare) according to the manufacturer's instructions.

A first round of PCR was performed using standard unlabeled primer pairs, followed by a second round of primer extension [run-off (RO)] with a fluorescence-labeled nested antisense oligonucleotide. The PCR reaction conditions for TCR γ and TCR δ were 3 minutes 95°C for initial denaturation, followed by 40 cycles (45 seconds 95°C, 45 seconds 60°C, 60 seconds 72°C), with a final 5 minutes 72°C extension. In the second round, 2 μ L of the PCR products was used for a primer extension reaction with a single fluorescent (5' 6-FAM-labeled) nested antisense primer from the particular constant region (RO primers). The conditions of this "run-off reaction" were 2 minutes at 94°C, then followed by 5 cycles (25 seconds 94°C, 45 seconds 60°C, 45 seconds 72°C), and finally by 5 minutes at 72°C. The products were analyzed using an ABI 3130 XL capillary sequencer (Applied Biosystems) to determine the length distribution of the fluorescent fragments for γ and δ groups. The results were analyzed using the Genescan Analysis software version 3.7 (Applied Biosystems; primer sequences in Supplementary Table S1; ref. 28).

Single-cell PCR of TNBC $\gamma\delta$ TILs

Similar to the single-cell PCR technique previously described for the analysis of rearranged immunoglobulin genes (29), a multiplex, semi-nested, hot-start PCR was set up with 15 newly designed primers (see Supplementary Table S1). For the first round, a master mix was prepared containing 5 μ L dNTP (2 mmol/L), 5 μ L 10 \times PCR buffer (High-Fidelity System), 5 μ L primer mix (2.5 μ mol/L, forward and reverse primers), 3.2 μ L 25 μ mol/L MgCl₂, 6.5 μ L H₂O, and 15 μ L from the DNA digestion. A volume of 0.3 μ L Expand High-Fidelity enzyme mix (3.5 units/ μ L) was added after the first denaturation step to a final volume of 40 μ L. The cyclor program was 95°C 2 minutes, 80°C pause (enzyme added), 72°C 1 minute, 39 cycles at 95°C 50 seconds, 56°C 30 seconds, 72°C 60 seconds, then 72°C 5 minutes and 10°C pause. For the second round, eight master mixes were prepared to detect TCR γ and δ chains: two mixes for TCR γ (V γ 1-8 and V γ 9) and six for TCR δ (V δ 1, V δ 2, V δ 3, V δ 4, V δ 5, and V δ 6) with 2.5 μ L dNTP (2 mmol/L), 2.5 μ L 10 \times PCR buffer, 1.25 μ L of the respective V γ and V δ forward primers, 1.25 μ L of the respective J segment mix primers (see Supplementary Table S1), 2 μ L 25 mmol/L MgCl₂, 12.2 μ L H₂O, 3 μ L of first-round PCR product, and 0.3 μ L Expand High-Fidelity enzyme mix (3.5 units/ μ L). The cyclor program was 95°C 5 minutes, 72°C 1 minute, followed by 35 cycles at 95°C 50 seconds, 55.5°C 30 seconds, 72°C 1 minute, then 72°C 5 minutes, 15°C 5 minutes and 4°C pause. The PCR products were analyzed by 2% agarose gel electrophoresis, and the positive bands were excised and purified from the gel with the QIAEX II Gel Extraction Kit (Qiagen). The DNA was sequenced using the BigDye Terminator 3.1 system (Applied Biosystems). The sequencing reactions consisted of 1 μ L BigDye, 3.75 μ L 5 \times sequencing buffer, 0.75 μ L forward primer (see Supplementary Table S1), 3 to 10 μ L template, and water to a final volume of 20 μ L. The cycling conditions were 96°C 5 minutes, then 24 cycles at 95°C 15 seconds, 50°C 10 seconds, 60°C 4 minutes, followed by a 10°C pause. The sequence

sample was cleaned using the DyeEx 2.0 Spin Kit (Qiagen) and analyzed on the ABI 3130xl capillary sequencer (Applied Biosystems).

Retroviral expression of plasmids for TCR $\gamma\delta$ and endothelial protein C receptor

Codon-optimized DNAs coding for the full lengths of γ and δ TCR chains and endothelial protein C receptor (EPCR) were ordered at BaseClear Inc (see Supplementary Document, "Nucleotide sequences of synthetic DNA constructs"). The synthetic genes were flanked with a 5' NcoI and a 3' BamHI site for subsequent cloning into the retroviral expression vector pBullet (ref. 30; kind gift by Ralph Willemsen, Erasmus Medical Center, Rotterdam, the Netherlands), modified by inserting IRES-neo and IRES-puro cassettes (see Supplementary Document, "Nucleotide sequences of synthetic DNA constructs"). The γ TCR genes were subcloned into the pBullet-IRES-neo, and the δ TCR genes and EPCR were subcloned into the pBullet-IRES-puro as previously performed (31).

Transduction of $\alpha\beta$ T cells, $\gamma\delta$ T cells, NKT cells, and target cells

For the generation of TEGs, PBMCs from healthy donors (Sanquin) were transduced with defined TCR γ and δ chains as described (19, 32, 33). Briefly, retroviral supernatants were produced by Phoenix-AMPHO packaging cells, transfected with retroviral helper vectors, gag-pol (pHIT60; ref. 34) and ENV (pCOLT-GALV; ref. 35; both kind gifts by Ralph Willemsen, Erasmus Medical Center, Rotterdam, the Netherlands), together with pBullet retroviral constructs containing TCR γ and TCR δ of the indicated TCR (Results section) using FuGENE HD (Promega). PBMCs preactivated with anti-CD3 (30 ng/mL; clone OKT3, Janssen-Cilag) and IL2 (50 U/mL; Proleukin, Novartis, 13610880, hospital pharmacy UMCU) were transduced twice (within 48 hours) with viral supernatants in the presence of IL2 (50 U/mL) and polybrene (4 μ g/mL; Sigma-Aldrich). Transduced T cells (TEG) were expanded by stimulation with anti-CD3/anti-CD28 Dynabeads (0.5 \times 10⁶ beads/10⁶ cells; Invitrogen) and IL2 (50 U/mL) and incubated with geneticin (800 μ g/mL; Gibco) and puromycin (5 μ g/mL; Sigma-Aldrich) for 1 week. TEGs were stimulated biweekly with PHA-L (1 μ g/L; Sigma-Aldrich), IL2 (50 U/mL; Proleukin, Novartis), IL15 (5 ng/mL; R&D Systems, 247-IL), and irradiated allogeneic PBMCs (12.5 \times 10⁶, dose: 3,500 cGy), Daudi (2.5 \times 10⁶, dose: 8,000 cGy) and LCL-TM (2.5 \times 10⁶, dose: 8,000 cGy) cells. Fresh IL2 was added twice weekly. Transgenic TCR expression was routinely assessed by flow cytometry on a BD FACSCanto II using anti-TCR $\alpha\beta$ APC (1:10; eBioscience, 17-9986-42), and anti-TCR $\gamma\delta$ PE (1:10; Beckman Coulter, IM1349), anti-CD4-FITC (1:20; eBioscience, 11-0049-42), and anti-CD8-PerCPCy5.5 (1:1,000; BioLegend, 301032).

For the generation of NEG and GEG, NKT and $\gamma\delta$ T cells of healthy donors were transduced using a similar protocol that was used for PBMCs. $\gamma\delta$ and NKT cells were sorted prior to transduction using a MACS TCR γ/δ + T-cell or NKT cell isolation kit (Miltenyi), respectively, a >90% $\gamma\delta$ TCR⁺ cell population and a >80% CD56⁺/CD3⁺ NKT cell population was obtained after MACS. The NEG and GEG were expanded with PHA-L (1 μ g/L; Sigma-Aldrich), IL2 (50 U/mL), IL15 (5 ng/mL; R&D Systems, 247-IL), and irradiated allogeneic PBMCs (12.5 \times 10⁶, dose: 3,500 cGy), Daudi (2.5 \times 10⁶, dose: 8,000 cGy) and LCL-TM (2.5 \times 10⁶, dose: 8,000 cGy) cells. Fresh IL2 was added twice weekly.

HEK293T, HeLa, K562, and U937 cells were transduced with EPCR in a similar fashion as described above. Three days after the second transduction, the transduced cells were selected with puromycin (2.5 μ g/mL), and the transgenic expression of EPCR was assessed by

flow cytometry on a BD FACSCanto II, cells were stained using anti-EPCR (1:50; Abcam, ab81712) and goat-anti-rat IgG AF647 (1:400; Invitrogen, A-21247).

ELISPOT assays

IFN γ ELISPOT assays were performed as previously described (16, 36). Briefly, 15,000 TCR-transduced or mock-transduced T cells and 50,000 target cells (ratio 0.3:1), as indicated in the figures, were cocultured for 24 hours in nitrocellulose-bottomed 96-well plates (Millipore) precoated with anti-IFN γ (5 μ g/mL; clone 1-D1K; Mabtech). Plates were washed and incubated with a second biotinylated anti-IFN γ (1 μ g/mL in PBS/0.5% FCS clone 7-B6-1; Mabtech), followed by streptavidin-HRP (1:500 dilution in PBS; Mabtech). IFN γ spots were visualized with the TMB substrate (Sanquin), and the number of spots was quantified using ELISPOT Analysis Software (Aelvis).

CD107 assay

Two TIL TCR $\gamma\delta$ clones transduced in $\alpha\beta$ T cells, TEG-C132 and TEG-F4, were incubated alone or with target cells at an effector:target (E:T) ratio of 1:1 in the presence of CD107a-PE (BD Biosciences; clone H4A3). $\gamma\delta$ TCR monoclonal antibody/antibodies and a matched isotype control were included during the incubation at a concentration of 20 μ g/mL. For TEG-C132 and TEG-F4, anti- $\gamma\delta$ TCR clone 11F2 (Antibody Chain International B.V.) and isotype control (IgG1 κ clone MOPC-21 (Abcam) were used. For TEG-C132, two additional $\gamma\delta$ TCR monoclonal antibodies were included: clone B1 (BioLegend) and clone 5A6.E9 (Fisher Scientific). After 2 hours of incubation, Golgistop (1:150; BD Biosciences, 554724) was added. After 6 hours, cells were washed in FACS buffer, containing PBS + 1% BSA (fraction V; Sigma-Aldrich, 10735094001) and stained with an antibody mix. For TEG-C132: anti-CD3 eFluor450 (eBioscience; clone OKT3), CD8 PerCP-Cy5.5 (BioLegend; clone RPA-T8), and $\gamma\delta$ TCR APC (BD Biosciences; clone B1). For TEG-F4: V δ 1 FITC (Thermo Scientific; clone TS8.2), CD8 PerCP-Cy5.5 (BioLegend; clone RPA-T8), CD4 PeCy7 (eBioscience; clone RPA-T4), and $\alpha\beta$ TCR APC (eBioscience; clone IP26). Cells were washed two times in FACS buffer and fixed in 1% paraformaldehyde in PBS. Data acquisition was done on FACSCanto and analyzed using FACSDiva software (BD).

Expression, purification, and staining with soluble TCRs

The variable and constant domains of the TCR chains, clones C132, F4, and Zi11, were amplified from synthetic DNA encoding the full-length TCRs using gene-specific primers containing an *AfeI* restriction site for the forward V-gene-specific primers and a *Bam*HI site for the reverse C-gene-specific primers (see Supplementary Table S1). The TCR δ chains were ligated into a modified pBullet vector containing a μ -phosphatase signal peptide at the 5' end and *fos* zipper at the 3' end of the construct (see Supplementary Document, "Nucleotide sequences of synthetic DNA constructs"). The TCR γ chains were ligated in a modified pBullet vector containing a μ -phosphatase signal peptide at the 5' end and at the 3' end, a *jun* zipper followed by a biotin acceptor peptide and a poly-histidine (His) tag (see Supplementary Document, "Nucleotide sequences of synthetic DNA constructs"). Synthetic DNA encoding for the bacterial biotin ligase BirA (Uniprot accession code: P06709; see Supplementary Document, "Nucleotide sequences of synthetic DNA constructs") was also ligated in a pBullet vector containing a signal peptide. Soluble $\gamma\delta$ TCRs were expressed by Freestyle 293-F cells (Thermo Fisher) transiently transfected with plasmids containing TCR δ , TCR γ , and BirA (in a 45:45:10 ratio) using polyethylenimine (PEI) in a 2:3 ratio, 1 μ g of total plasmid

DNA was used to transfect 10^6 cells. Six hours after transfection, the media were supplemented with 1% penicillin/streptomycin (Gibco) and 100 μ mol/L biotin (Sigma-Aldrich; 14400).

The expression media were harvested 5 days after transfection by centrifuging the cultures for 10' at $750 \times g$ to pellet the cells. The supernatant was supplemented with TBS, and loaded on a 1-mL HisTrap Excel column (GE Healthcare). A multistep gradient with increasing concentrations of imidazole (Merck, 1047160250), 10 column volumes (CV) 10 mmol/L imidazole, followed by a linear gradient of 10 to 300 mmol/L imidazole in 20 CV, was used to wash and elute the soluble TCRs from the column. The eluted soluble TCRs were further purified using a 1-mL HiTrapQ column (GE Healthcare) at pH 8.2, soluble TCRs were loaded in TBS-20 mmol/L NaCl and eluted using a linear gradient of 20 to 400 mmol/L NaCl in 30 CV. Fractions were loaded on a 4–20 Mini Protean TGX Gels (Bio-Rad) using Laemmli sample buffer (Bio-Rad, 1610747) after electrophoresis gels were stained with InstantBlue Safe Coomassie Stain (Sigma-Aldrich, ISB1L). Fractions containing the soluble TCR were pooled and concentrated using a Vivaspinn Turbo 4 10-KDa cutoff spin concentrator (Sartorius, VS04T01). Tetramers were prepared from TCR monomers by adding one equivalent of SA-PE (BD Pharmingen, 554061) to five equivalents of TCR in 2 steps over 5 minutes.

MDA-MB436 cells were incubated with TCR tetramers with a streptavidin concentration 3 μ g/mL for 30 minutes at room temperature in the presence of isotype control (IgG1 κ , clone MOPC-21; abcam) or anti- $\gamma\delta$ TCR (clone B1; BioLegend). Cells were washed two times in FACS buffer and fixed in 1% paraformaldehyde in PBS. Data acquisition was done on FACSCanto and analyzed using FACSDiva software (BD).

Analysis of HTS data sets and availability of data sets

Data analysis of HTS data

TCR sequence alignment, assembly, and clonotype extraction were performed using the MiXCR (version-v2.1.1) program for data set 1, 3, 4, and 5 (37). VDJtools (v1.1.4) was utilized for frequency-based correction of clonotypes (38). For data set 1, 3, and 4, only functional reads which passed a frequency filter of 0.1% were used for further analysis. For data set 5 sequences with a read count of ≥ 20 were used for analysis. RNA-seq data of data set 6 were analyzed with the RNA-seq mode of MiXCR. tcr R package (v2.2.1.15) was utilized for TRD repertoire overlap analysis and the estimation of VDJ gene usage.

Availability of HTS data

HTS data in the standard FASTQ format from this study (data sets 1 and 3) are available via the SRA database and can be located using the NCBI BioProject accession number: PRJNA397967. HTS data in the standard FASTQ format from data sets 4 and 5 are publicly available (39, 40). Data set 6 consists of the RNA-seq data of the TCGA database. The following cancer data sets were used: COAD (colon adenocarcinoma), KIRC (kidney clear cell carcinoma), LUAD (lung adenocarcinoma), READ (rectal adenocarcinoma), SKCM (skin cutaneous carcinoma), and TNBC. Data set 6 was supplemented with amino acid CDR3 sequences of publicly available CDR3 amino acid sequences of TILs (41).

Statistical analyses

Unpaired *t* test statistical analysis or a one-way ANOVA followed by Tukey *post hoc* test was performed using GraphPad Prism software (GraphPad Inc, Version 7). Data were expressed as means \pm standard deviation (SD). A value of $P < 0.05$ was considered statistically significant.

Results

$\gamma\delta$ TILs in patients with TNBC are $IFN\gamma^+$, $TNF\alpha^+$, and $IL17^-$

We analyzed $\gamma\delta$ TILs that were in close contact with tumor cells within TNBC frozen sections of 7 patients, 3 with invasive ductal cancer (IDC) and 4 classified as medullary breast cancer (MBC; Supplementary Table S2; ref. 6). In both tumor types, the $\gamma\delta$ TILs were frequently in close contact with apoptotic tumor cells (Fig. 1A and B), and the majority of $\gamma\delta$ TILs stained positive for CD69, $IFN\gamma$, and $TNF\alpha$, whereas less than 20% were positive for $IL17$ ($P < 0.001$, $n = 7$; Fig. 1C–F and H). No difference was seen in the percentage

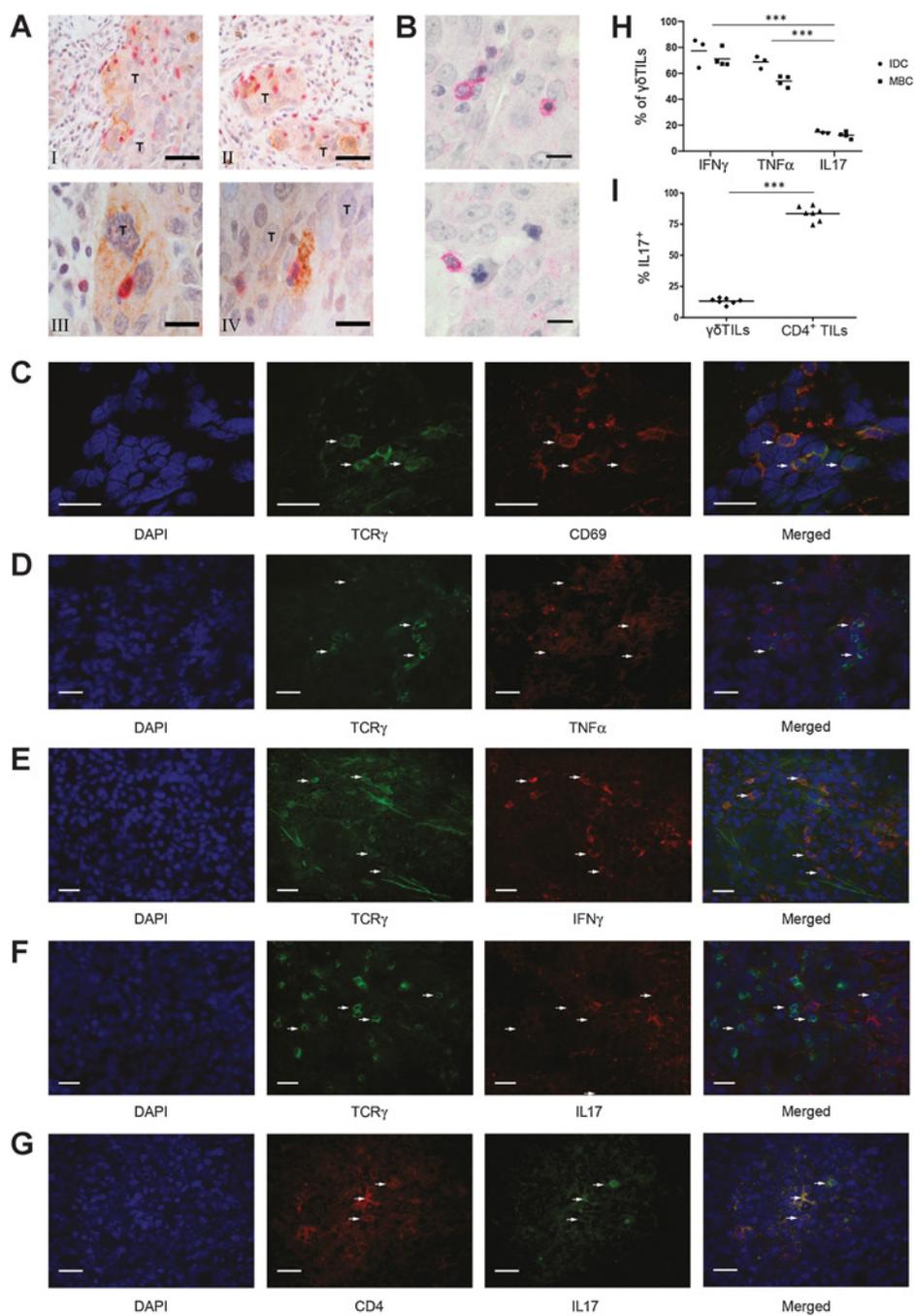
$IFN\gamma^+$ and $IL17^+$ $\gamma\delta$ TILs between MBC and IDC tumors. However, we observed a small decrease in the percentage of $TNF\alpha^+$ $\gamma\delta$ TILs in MBC compared with IDC tumors. Over 80% of $CD4^+TCR\gamma\delta^-$ cells produced $IL17$ and appeared to represent a major source for $IL17$ in the tumor parenchyma (Fig. 1G and I).

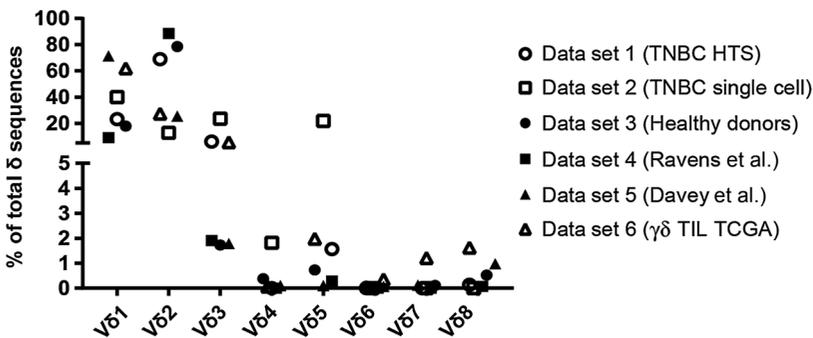
TCR γ and δ repertoire of $\gamma\delta$ TILs in TNBC

Initially, $V\gamma$ and $V\delta$ spectratyping in TNBC patients was performed following RNA extraction from frozen tissue sections (Supplementary Fig. S1A and S1B). These analyses showed polyclonal $\gamma\delta$ T-cell populations where the $V\gamma9$ gene expression was reduced, implying

Figure 1.

$\gamma\delta$ TILs exhibit a cytotoxic phenotype in contact with breast cancer cells. **A**, Representative double IHC staining with TCR γ antibody (red) and cleaved caspase-3 (brown). T, tumor parenchyma. Scale bar, 50 μ m in I and II, and 20 μ m in III and IV. **B**, Representative detail at high magnification of $\gamma\delta$ TILs (red) in active contact, including a pseudopod extension with tumor cells that displays pycnotic nuclear and cytoplasmic characteristics. Scale bar, 20 μ m. **C–G**, Representative colocalization analysis using double immunofluorescence of TCR γ (green) and activation marker CD69 (red) expression (**C**), $TNF\alpha$ (red; **D**), $IFN\gamma$ (red; **E**), and $IL17$ (red; **F**), and CD4 (red) and $IL17$ (green; **G**). DAPI, blue; scale bar, 20 μ m. **H**, Cumulative data obtained from the immunofluorescence analysis (MBC, $n = 4$; IDC, $n = 3$). Percentages of $\gamma\delta$ TILs expressing the $IFN\gamma$, $TNF\alpha$, or $IL17$ (mean and individual data points are indicated). One-way ANOVA followed by Tukey *post hoc* test revealed statistically significant differences (***, $P < 0.001$). **I**, Percentages of $IL17^+$ cells among $CD4^+$ cells and $\gamma\delta$ TILs (mean and individual data points are indicated). An unpaired *t* test revealed statistically significant differences (***, $P < 0.001$).



**Figure 2.**

The V δ gene distribution in different data sets. V δ gene distribution in data sets 1 through 6 (see Materials and Methods for more details on data sets). Filled symbols are used for data sets from healthy donors. Open symbols are used for data sets with $\gamma\delta$ TILs. See also Supplementary Fig. S1. References are for data set 4 (39), data set 5 (40), and data set 6 (41).

that $\gamma\delta$ T cells other than V γ 9/V δ 2 were predominant. HTS of the TNBC samples confirmed the diverse TCR δ repertoire (Fig. 2; data set 1; Supplementary Table S3). In order to characterize the pairs of TCR γ and δ chains of $\gamma\delta$ T cells in contact with tumor cells, we used laser microdissection of single $\gamma\delta$ TILs, followed by a PCR protocol to amplify their CDR3 regions. In total, 530 single $\gamma\delta$ T cells were isolated from 11 tumors. Although some SCS reactions did not generate reliable sequencing data for both of the paired chains, 27 paired TCR γ and δ sequences from 9 different patients were identified (data set 2; Supplementary Table S4). We also obtained 63 TCR γ and 28 TCR δ nonpaired sequences (data set 2; Supplementary Tables S2 and S5). Various non-V γ 9 and non-V δ 2 genes were most prominently represented among the detected sequences from single cells (Fig. 2; data set 2). Out of the 81 unique TCR γ amino acid sequences obtained, 10 different sequences were shared among the patients, 9 of which had the same nucleotide sequence (Supplementary Fig. S2A). One TCR $\gamma\delta$, B9, had a TCR γ chain that has been previously identified in a CD1d-restricted TCR $\gamma\delta$, AU2.3 (42). Four of the 27 SCS-identified paired TCR $\gamma\delta$ used the same V δ 5 amino acid sequence but was always paired with a different TCR γ chain. This V δ 5 sequence, previously found to be associated with CMV and tumor reactivity (43), was identical with 2 of 28 SCS-identified CDR3 δ nonpaired amino acid sequences, classifying this CDR3 δ sequence as coding for a public CDR3 δ chain (Supplementary Fig. S2B). All the nucleotide sequences of this public V δ 5 CDR3 δ chain were identical among different patients (Supplementary Table S4). None of the examined TNBC tissue samples tested positive for CMV (Supplementary Table S2), but the serostatus of CMV in our patients was not determined. One of the 27 SCS-identified paired TCR $\gamma\delta$ s had a shared (public) V δ 1 sequence, and another shared V δ 1 sequence was identified in the pool of the nonpaired TCR δ chains. The shared V δ 1 amino acid sequences also had the same nucleotide sequence. Only four of the 41 SCS-identified unique TCR δ amino acid sequences were present within HTS data belonging to the same TNBC sample, suggesting an initial repertoire focusing (39) at the tumor side, with a subsequent lack of clonal expansion of tumor-interacting $\gamma\delta$ T cells.

Overall usage of TCR δ sequences

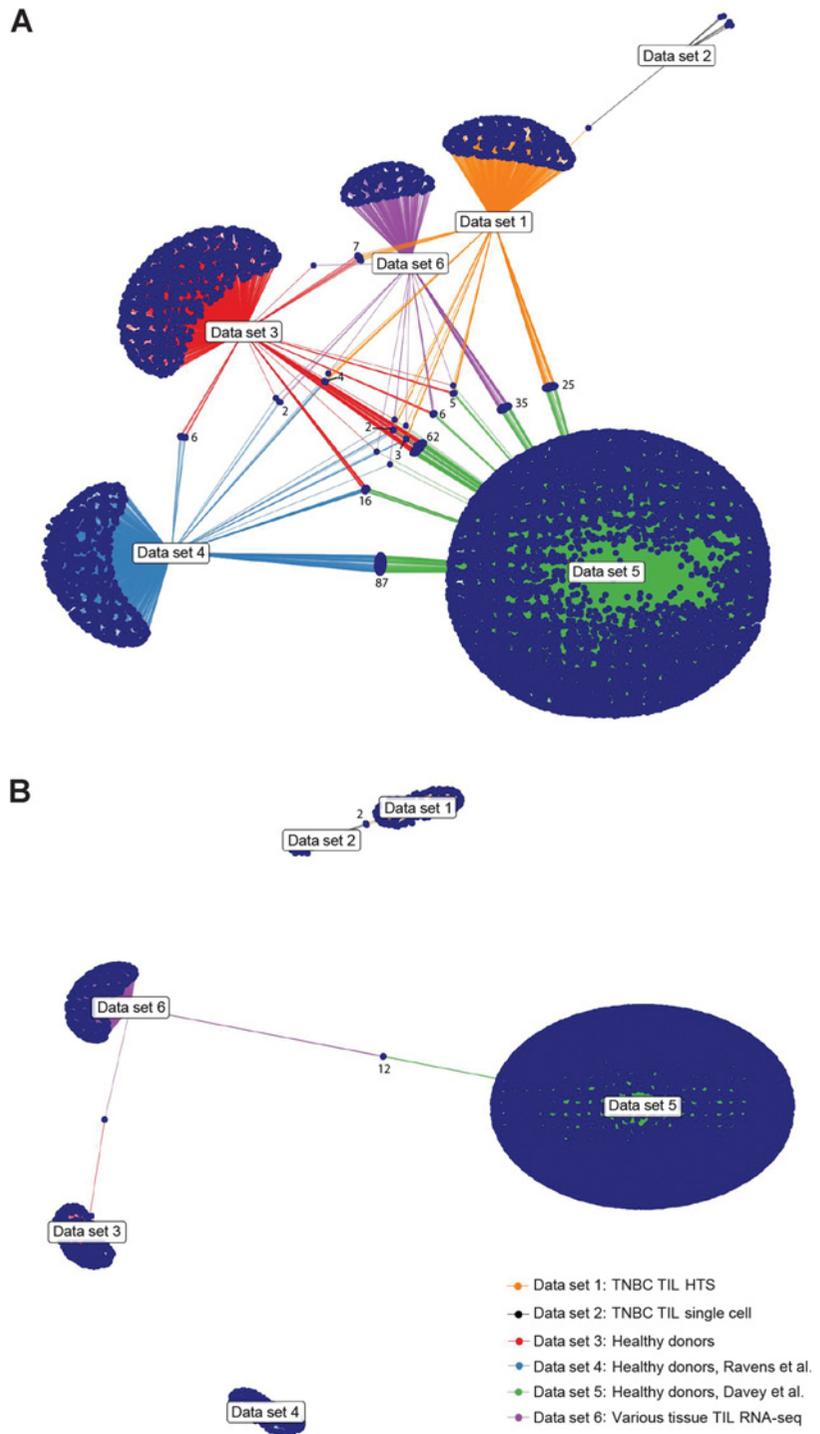
To put our own data of TCR δ chain sequences derived from $\gamma\delta$ TILs in TNBC within the context of natural repertoires observed in the peripheral repertoire, as well as from tumors, we investigated the overall usage of TCR δ chains in the peripheral blood of healthy donors and in other tumor tissues. Therefore, we analyzed the presence of TCR δ amino acid sequences across different healthy and tumor sequencing data sets. First, we analyzed the presence of shared TCR δ chains in the peripheral blood of 13 healthy volunteers by HTS (data set 3; Supplementary Table S3) and studied the most prevalent TCR δ sequences, defined as sequences with a clonal frequency of >0.1%. Of

the 1401 most prevalent amino acid sequences in data set 3, 17 were shared between at least two donors. Next, we analyzed shared sequences of two publicly available TCR δ HTS data sets from peripheral blood of healthy donors. For data set 4 (Supplementary Table S3) of Ravens and colleagues, we again analyzed the sequences with a clonal frequency of >0.1% (39). Data from Davey and colleagues (data set 5; Supplementary Table S3) were pooled in the public space (40), and therefore only sequences with a read count of more than 20 were included in our analysis. The V δ gene distribution of data sets 3 and 4 consisted mainly of V δ 2 sequences representing an unselected peripheral repertoire. Data set 5 was enriched for V δ 1⁺ $\gamma\delta$ T cells, which explains the high percentage of V δ 1 sequences (Fig. 2). Within the different healthy donors' data sets, 186 V δ 2⁺ TCR δ amino acid sequences were shared and, thus, considered public (Fig. 3A; Supplementary Fig. S3A). Contrastingly, no shared V δ 1 sequences were identified in the peripheral blood data sets. Data set 6 (Supplementary Table S3) included TCR δ sequences retrieved from the RNA-seq data of six different cancers from The Cancer Genome Atlas (TCGA) database and from the $\gamma\delta$ TILs CDR3 δ sequences of various cancers published by Li and colleagues (41). The majority (62%) of our analyzed 1,407 $\gamma\delta$ TIL TCR δ sequences were of the V δ 1 origin (Fig. 2). Considering the V δ 2⁺ T-cell proportion in the TIL data set (data set 6), the proportion of V δ 2⁺ in the TNBC HTS data (data set 1) was higher than expected, and more similar to the peripheral blood data sets (Fig. 2). This is compatible with TNBC's higher microvascular density in comparison with other breast tumors (44) or due to the fact that V δ 2⁺ chains can pair with other TCR γ chains than V γ 9 (40). When data set 6 with the TILs was compared with the sequences of all the other data sets, the percentage of V δ 2 TCR δ sequences in tumor tissues appeared relatively low, but 53 shared V δ 2 TIL amino acid sequences were identified within the data sets (Fig. 3A; Supplementary Fig. S3A). Finally, we compared nucleotide sequences across all of the different data sets, except data set 6, where nucleotide sequences were not available from the data of Li and colleagues (41). Across the different data sets, 222 V δ 2 amino acid sequences were shared, and 29 of those (13%) were also identical as nucleotide sequences (Supplementary Table S6). Because the pooling of individual donor data could result in underestimating shared nucleotide sequences, we also analyzed shared nucleotide sequences between different donors within data sets 3 and 4. This analysis revealed higher percentages of shared amino acid sequences with identical nucleotide sequences, i.e., 39% (data set 3) and 45% (data set 4), respectively (Supplementary Table S6).

The percentage of shared V δ 2 amino acid sequences in data set 6 (TILs) was relatively high compared with the shared sequences in peripheral blood of healthy donors [13.8% (53/385) vs. 1.5% (186/12207)]. When peripheral T cells were substantially enriched for V δ 1 $\gamma\delta$ T cells before HTS (data set 5), 13 V δ 1 sequences could be characterized as shared between peripheral blood and $\gamma\delta$ TILs

Figure 3.

Overview of shared TCR δ sequences among data sets. Network plots of the shared sequences among different data sets are shown. Each dot represents 1 CDR3 TCR sequence. When more than 1 sequence is shared between data sets, the actual number of shared sequences is indicated. **A**, Overview of shared V δ 2 sequences among the different data sets. **B**, Overview of shared V δ 1 sequences among the different data sets. See also Supplementary Table S4, Supplementary Fig. S3, and the Materials and Methods section for more information on the tools used to compare the different data sets.



(Fig. 3B; Supplementary Fig. S3B). One V δ 3 sequence, which was present among the $\gamma\delta$ TIL sequences (data set 6), was also present in the TNBC TILs (Supplementary Fig. S3C). Thus, our analyses identified commonly shared V δ amino acid sequences in the peripheral blood and tumor tissues, but the corresponding V δ repertoires differed considerably. In the peripheral blood, the shared sequences were all V δ 2⁺, whereas tumor tissues also contained shared V δ 2⁻ TCR chains.

Tumor reactivity of TCR $\gamma\delta$ sequences derived from $\gamma\delta$ TILs

The TEG format with absence of many coreceptors usually observed on $\gamma\delta$ T cells (19, 32, 33) allowed us to investigate whether tumor reactivity of $\gamma\delta$ TILs was mediated through their individual TCR $\gamma\delta$ s and not by other innate immune receptors usually expressed on $\gamma\delta$ TILs. To this end, we generated a series of 15 TEGs expressing paired TCR γ and δ chains derived from TNBC $\gamma\delta$ TILs (Supplementary Fig. S4A; Supplementary Table S4) to assess whether these chains

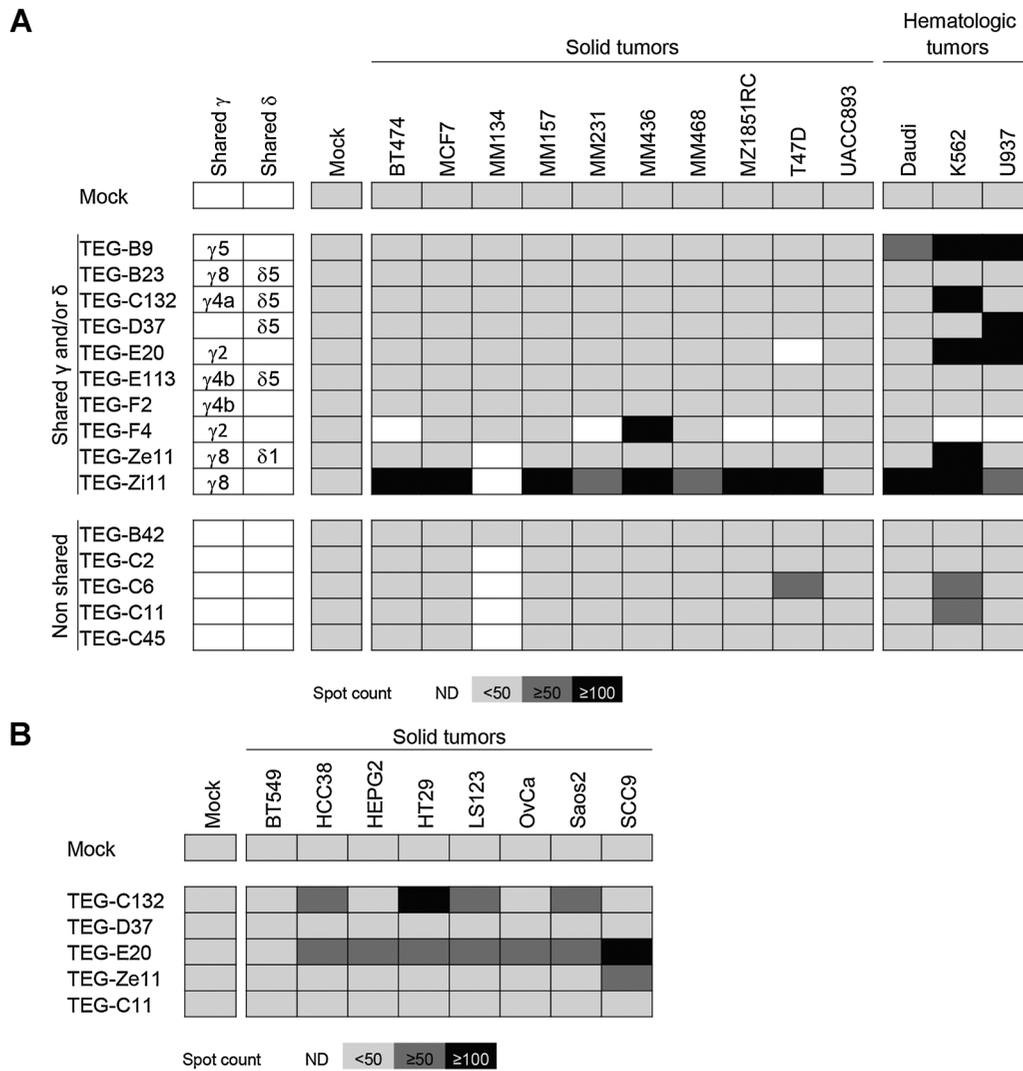


Figure 4.

TEGs engineered with TIL-derived TCR $\gamma\delta$ chains show reactivity against different tumor cell lines. **A** and **B**, Reactivity of $\gamma\delta$ TIL TCRs in TEG format against tumor cell lines with high and low mutational loads was measured by an IFN γ ELISPOT; average spot counts of experiments are shown. As negative control, T cells engineered with nonfunctional TCR $\gamma\delta$ chains were used. Effector and target cells were incubated overnight in a 0.3:1 E:T ratio. TEGs that share either a TCR γ or TCR δ chain within the tested TCRs or with previously published TCR chain indicators are shown in **A**, and the corresponding CDR3 sequences are listed in Supplementary Table S4 and Supplementary Fig. S2. ND, not determined. ND, not determined.

had the potential to mediate antitumor reactivity. Because most human $\gamma\delta$ T cells are not HLA restricted, we used several established cell lines derived from tumors, including breast cancer and leukemia/lymphomas. Activation of TEGs by tumor cell lines was measured using an IFN γ ELISPOT, where we used the following thresholds to quantify reactivity: >50 spots as low reactive and >100 spots as highly reactive. Nine of 15 TEGs recognized at least one of the cancer cell lines (Fig. 4A). Because some of the TEGs recognized tumor cell lines of hematologic origin only, we extended the tumor panel with 8 additional solid tumor cell lines for 5 TEGs. Three of 5 tested TEGs, which showed in the first screen only activity against hematologic tumors, also recognized solid tumor cell lines (Fig. 4B). Two of the 4 TEGs with the shared V $\delta 5$ chain displayed distinct tumor reactivity patterns, indicating that tumor recognition was likely to depend on the specific paired TCR γ chain (Fig. 4A and B). Also, the TCR γ and δ chains from

TEG-C132 were identical with those reported in the EPCR-reactive $\gamma\delta$ T-cell clone (45). The EPCR reactivity of TEG-C132 was confirmed by measuring the response of TEG-C132 against a panel of naturally EPCR-positive or -negative tumor cell lines, or cell lines engineered to overexpress EPCR (Fig. 5A).

$\gamma\delta$ T-cell activation can be influenced by multiple cell receptors (46). To test whether coreceptors expressed by $\gamma\delta$ T cells could modulate the activity of TCR $\gamma\delta$ derived from $\gamma\delta$ TILs, TCR $\gamma\delta$ complexes from TEG-C132 (functional within TEGs) and B23 (nonfunctional within TEGs), as well as the nonfunctional TCR $\gamma\delta$ LM1 (33), were used to construct $\gamma\delta$ T cells with a defined TCR $\gamma\delta$ (GEG). No gain of function was observed for either receptor in the GEG format. In contrast, a slight reduction of activity was detected for GEG-C132 compared with TEG-C132 (Fig. 5A and B), possibly due to the mispairing of introduced TCR γ and δ chains with endogenous TCR γ and δ chains.

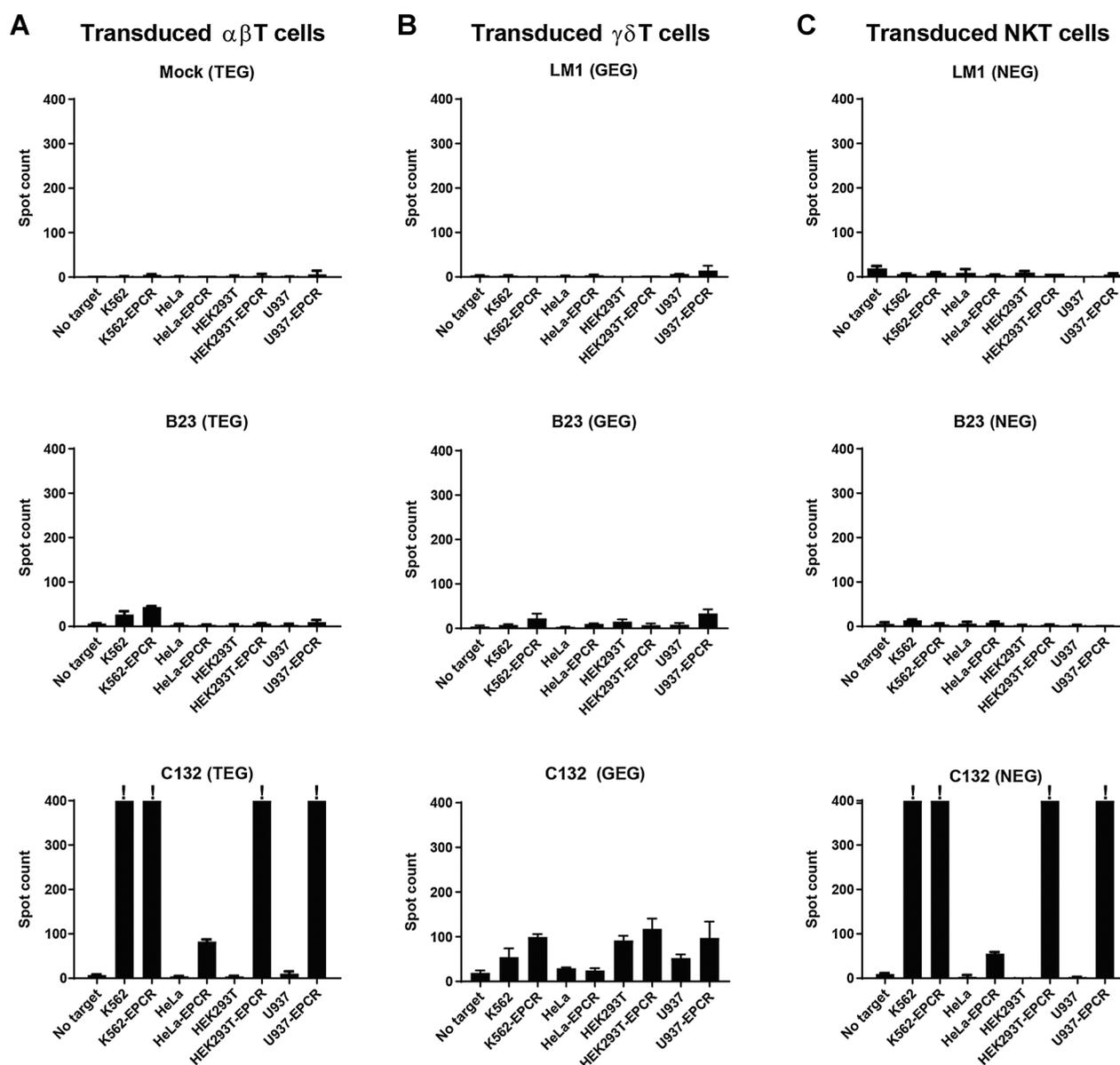


Figure 5. Activity of TCR $\gamma\delta$ when expressed on different carrier cells with varying amounts of NK-like receptors. Activation of TCR $\gamma\delta$ -transduced TEGs (A), GEGs (B), and NEGs (C) by a panel of target cell lines was measured by an IFN γ ELISPOT assay. Effector cells were incubated overnight without or with target cells in a 0.3:1 E:T ratio. Specific settings marked with “!” contained too many spots for accurate assessments. All four cell lines were transduced with retroviral constructs, where the EPCR (CD201)-coding sequence was subcloned into the pBullet-IRES-puro. The transduced cell lines are marked with the affix “-EPCR.” Data, mean \pm SD.

To avoid the possible mispairing with endogenous TCR γ or δ chains, we engineered NKT cells with a defined TCR $\gamma\delta$ (NEG). NKT cells were used because they express coreceptors similar to those present in $\gamma\delta$ T cells (Supplementary Fig. S4B). NEG-C132 showed no increased reactivity compared with TEG-C132, and consistent with this, no gain of function was detected in NEG-B23 cells (Fig. 5C).

To further confirm that the observed tumor reactivity was mediated through the introduced TCR $\gamma\delta$, TEG-C132, TEG-F4, and TEG-Zi11 were selected for CD4⁺ TEGs. CD4⁺ TEGs lack most NK cell receptors (16). CD4⁺ TEGs showed a similar reactivity pattern as the nonsorted TEGs. Again, CD4⁺ TEGs expressing the nonreactive TEG-

LM1 did not show any reactivity (Fig. 6A–C), suggesting that indeed activity was mediated by the introduced TCR γ or δ chains. We next determined whether for selected CD8⁺ TEGs (C132 or F4) degranulation was observed upon stimulation with a recognized target cell line. TEG-C132 showed increased surface expression CD107 after a 6-hour incubation with HT29 cells (Fig. 6D). Activity was selectively associated with cells having the highest expression of TCR γ or δ chains (Supplementary Fig. S4C). In the very same experiments, blocking antibodies have also been added to further substantiate the claim that activity was mediated by the introduced TCR γ or δ chains. In the presence of anti- $\gamma\delta$ TCR, surface CD107 decreased up to 2-fold. Similar

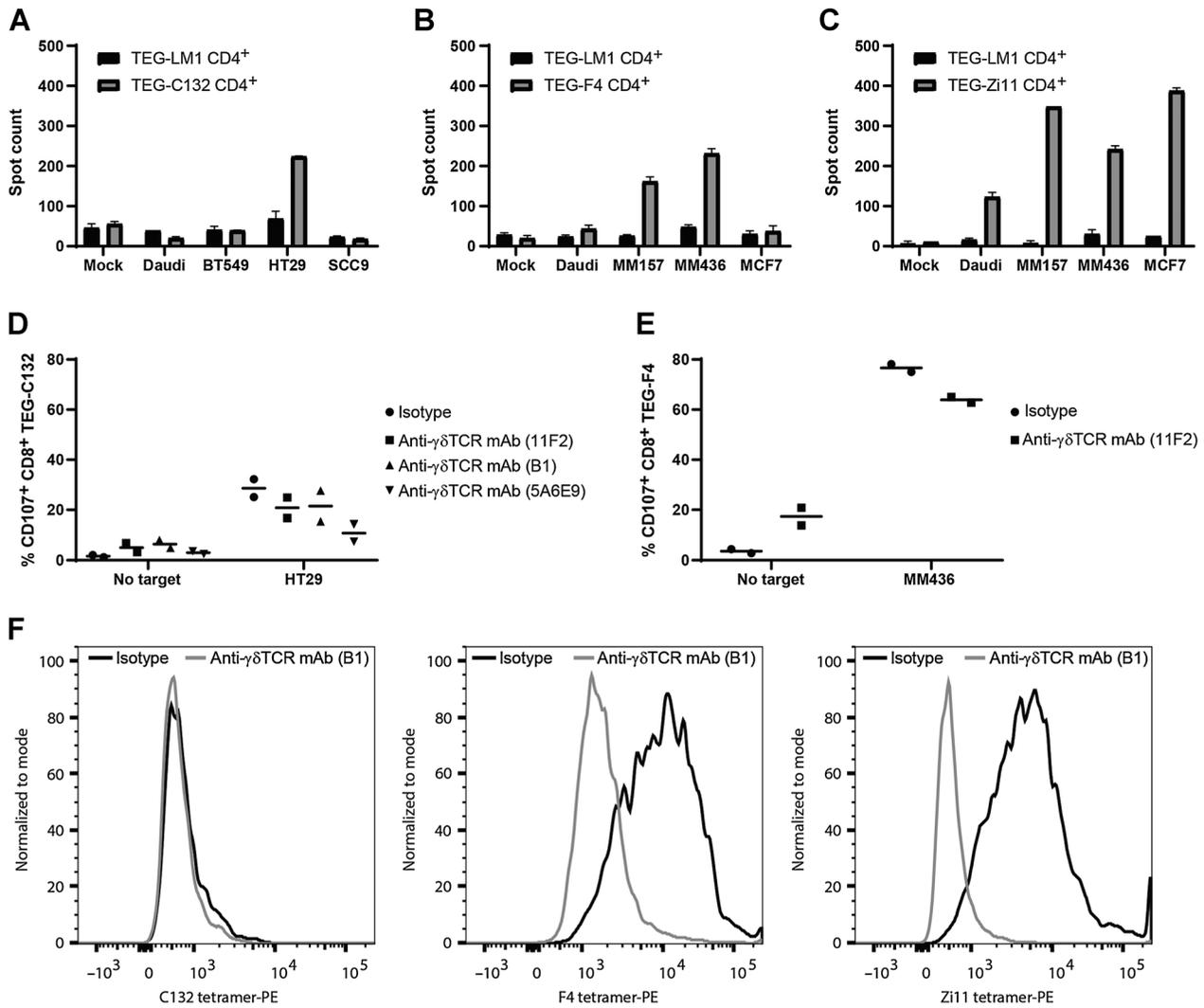


Figure 6. Reactivity against tumor cells is mediated through the TCR $\gamma\delta$. CD4⁺ TEG-C132 (A), TEG-F4 (B), and TEG-Zi11 (C) were incubated overnight in a 0.3:1 E:T ratio with the indicated target cell lines. IFN γ production was measured in an ELISPOT (mean \pm SD). D, Degranulation of CD8⁺ TEG-C132 was assessed by surface expression of CD107 in the absence or presence of several $\gamma\delta$ TCR monoclonal antibodies at a concentration of 20 μ g/mL. TEG-C132 was incubated either without or with tumor cell line HT29 tumor cells for 6 hours. CD107 surface expression was measured by flow cytometry (bars represent mean of two independent experiments; dots represent mean of the independent experiments). E, Degranulation of CD8⁺ TEG-F4 induced by tumor cell line MM436 in the absence or presence of anti- $\gamma\delta$ TCR clone 11F2 as described in D (bar represents mean of two independent experiments; dots represent mean of the independent experiments). F, Tumor cell line MM436 was stained with TCR tetramers (3 μ g/mL streptavidin) in the presence of 20 μ g/mL isotype control antibody or anti- $\gamma\delta$ TCR clone B1. TCR tetramer staining was analyzed by flow cytometry. Representative FACS plots are shown ($n = 4$).

results were obtained for TEG-F4 (Fig. 6E). In the absence of target cells, adding blocking antibodies also associated with a slight increase in CD107, suggesting that functional blocking experiments were partially hampered by a slight activation through the very same antibodies. Therefore, we formally addressed whether the $\gamma\delta$ TCR alone would be sufficient to bind to recognized tumor cell lines by generating fluorescent tetramers from different $\gamma\delta$ TCR that have been coincubated with MM436 as the tumor target. In line with their reactivity in the TEG format, F4 and Zi11 tetramers stained MM436, whereas C132 tetramer did not. $\gamma\delta$ TCR dependency was furthermore confirmed by adding anti- $\gamma\delta$ TCR, which resulted in a complete abrogation of staining through F4 and Zi11 tetramers (Fig. 6F). Thus, we have demonstrated that SCS of $\gamma\delta$ TILs, followed by expression of

their individual TCR $\gamma\delta$ s in lymphocytes, could rapidly confirm the implied and anticipated effectiveness of $\gamma\delta$ TIL-tumor cell interactions, whose antitumor reactivity does not depend on additional innate immune receptors.

Discussion

Our report is a comprehensive analysis of TNBC-infiltrating $\gamma\delta$ TILs, which are in close proximity to TNBC cells. We observed in this “project genesis” from archival material that the majority of $\gamma\delta$ TILs harbored a proinflammatory phenotype and only a minority associated with the expression of IL17. By reconstructing TCR γ or TCR δ from SCS within the TEG format, we provided evidence that paired

TCR γ and TCR δ chains could be active against not only breast cancer cells but also other tumor cell types. However, the lack of autologous viable tumor material did not allow to test the activity in an autologous system, and with the lack of knowledge on most tumor antigens seen by $\gamma\delta$ TILs, we could not formally assess whether frequency of truly tumor-reactive $\gamma\delta$ TILs was equivalent or exceeded the 10% as reported for $\alpha\beta$ TILs (47). We observed a high frequency of reactivity of all characterized TCR γ and TCR δ chains against different and multiple tumor types, which was unique for each TCR pair. Changing one counterpart altered the recognition pattern and was not improved when additional innate coreceptors were expressed in combination with the TCR. Thus, $\gamma\delta$ T cells have, through their individual TCR γ and δ chains, an intrinsic capacity to recognize cancer cells.

Clonal expansions have been reported in neoantigen-specific $\alpha\beta$ TILs (48). The spectratype and HTS analyses of the $\gamma\delta$ TILs in TNBC tumor tissue indicated a polyclonal repertoire, and the SCS detected no clear clonal dominance of tumor-reactive TCR $\gamma\delta$ s. This validates the previously reported absence of V δ 2⁺ TCR-driven lymphocyte expansion in cancer patients (8) and contrasts the expansions of *Mycobacterium tuberculosis*-activated V δ 2⁺ $\gamma\delta$ T lymphocytes (49). This apparent lack of clonal expansions is also in contrast with the potent clonal $\gamma\delta$ T-cell responses to viral infections (39, 40). This dichotomy—i.e., the presence of clonal responses in infections but their absence in progressing cancers—might be caused by a tolerogenic tumor microenvironment. $\gamma\delta$ TILs have been reported to be skewed toward an IL17- rather than an IFN γ -producing phenotype during the progression of neoplastic disease (10). However, in our cohort, only a minority of the TNBC $\gamma\delta$ TILs produced IL17, which has also been observed in colon cancer (50), suggesting that different tolerogenic mechanisms (51) could be involved. Thus, although mouse model studies indicate that the early cancer stages are susceptible to $\gamma\delta$ T-cell immunosurveillance (5, 52), our results imply that the complex tumor microenvironment might prevent an effective clonal expansion of tumor-reactive $\gamma\delta$ TILs.

Our data allow to speculate that there might be public TCR γ or δ chains in tumor tissue which associate with tumor recognition. However, this observation needs to be seen with caution. First, a systematic bias could have occurred because the total number of sequences analyzed in peripheral blood samples was higher. Extracting TCR sequences from RNA-seq data favors shorter sequences because of the read length of the original data, and public sequences tend to be shorter in sequence (39, 40). Finally, because others report that V δ 2⁻ TCRs public chains are a rare event in humans (39, 40), we cannot entirely exclude that in the complex SCS procedure in paraffin-embedded tissues associated with some cross-contamination, and we, therefore, may have overestimated the number of public TCR chains. Regardless of this technical concern, pairing the very same TCR γ with different δ chains and vice versa in our data set allowed to test whether tumor reactivity of TCR γ or δ chains pairs depended on both chains. Different clones with either identical TCR γ or δ chains and matching with a different counterpart associated with a different recognition pattern of tumor cells. This is in particular interesting as we characterized one TCR chain pair that has been initially reported to react against EPCR only (45). However, recognition of tumor cells through V δ 2⁻ $\gamma\delta$ TCRs has most been attributed to a recognition mode where both TCR γ and δ chains can bind to different ligands at the same time (53). For example, the TCR γ chain of the identified clone C132 has been suggested to bind to BTN-family members, whereas other TCR δ chains bind to CD1 family members. Recognition of tumor cells could then depend on varying distributions of both

ligands. This might also explain why the expression of the V δ 5 chain alone was not sufficient to trigger EPCR reactivity. However, such new TCR pairs might also simply recognize only one completely different ligand.

The isolation of an EPCR-reactive V γ 4V δ 5 TCR clone from our TNBC tissues has been characterized by others (45) and reported to be associated with anti-CMV responses. CMV reactivation has been accounted to reduce relapse of leukemia after allogeneic stem cell transplantations (36, 54). The reduced risk of leukemic relapse after CMV reactivation has been attributed to NK and $\gamma\delta$ T cells (36, 55, 56). Within this context, it is tempting to speculate that, in general, the CMV reactivation may not be clinically beneficial only for controlling leukemias (36), but could also confer some protection against certain solid tumors. However, we could not detect CMV in our TNBC sample by IHC, and there was no serum available to test the CMV serostatus of this patient.

Although ligands for most characterized V δ 2⁻ TCR γ or δ chains remain to be defined, testing V δ 2⁻ TCR γ or δ chains in the TEG format emphasized the considerable diversity within the $\gamma\delta$ T-cell repertoire in terms of tumor reactivity, which might also be a major factor contributing to numerous failures of clinical trials using *ex vivo*- and *in vitro*-expanded $\gamma\delta$ T cells (57). Here, we demonstrated that the natural weakness of $\gamma\delta$ TILs could be annulled by extracting their TCRs and created potent antitumor T lymphocytes (TEG) expressing these $\gamma\delta$ TIL-derived TCRs. The absence of HLA restriction of TCR $\gamma\delta$ responses underscores the possibility of TEG-based therapies in allogeneic scenarios and paves a way toward a new plethora of antigens accessible to target solid tumors when utilizing V δ 2⁻ $\gamma\delta$ TCRs (16, 17, 58).

In summary, we demonstrated that $\gamma\delta$ TILs cells frequently harbored TCR γ or δ chains, which could mediate tumor reactivity. The identification of such cancer cell-sensing TCR $\gamma\delta$ s and characterization of their ligands may open novel opportunities for future cell-based cancer therapies.

Disclosure of Potential Conflicts of Interest

J. Villacorta Hidalgo has and P. Fisch had ownership interest in patent applications by the University Medical Center, University of Freiburg. D.X. Beringer and Z. Sebestyen have ownership interest in patent applications by the UMC Utrecht. J. Kuball is a scientific advisor for, reports receiving a commercial research grant from, and has ownership interest (including patents) in Gadeta. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: A. Janssen, J. Villacorta Hidalgo, D.X. Beringer, Z. Sebestyen, E. Spierings, R. Küppers, P. Fisch, J. Kuball

Development of methodology: A. Janssen, J. Villacorta Hidalgo, D.X. Beringer, S. Kock, M. Follo, P. Fisch, J. Kuball

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J. Villacorta Hidalgo, D.X. Beringer, S. van Dooremalen, E. van Diest, A.R. Terrizi, P. Bronsert, S. Kock, M. Werner, K. Heise, M. Follo, T. Straetmans, S. Ravens, R. Küppers, P. Fisch, J. Kuball

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A. Janssen, J. Villacorta Hidalgo, D.X. Beringer, F. Fernando, M. Werner, T. Straetmans, S.A. Kasatskaya, F.E. Frenkel, E. Spierings, M. Malkovsky, P. Fisch, J. Kuball

Writing, review, and/or revision of the manuscript: A. Janssen, J. Villacorta Hidalgo, D.X. Beringer, P. Bronsert, M. Werner, M. Follo, Z. Sebestyen, D.M. Chudakov, E. Spierings, I. Prinz, M. Malkovsky, P. Fisch, J. Kuball

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A.R. Terrizi, P. Fisch, J. Kuball

Study supervision: D.M. Chudakov, P. Fisch, J. Kuball

Other (supervision on diagnostic, planning and analysis as principal pathologist): A. Schmitt-Gräff

Acknowledgments

Funding for this study was provided by ZonMW 43400003 and VIDI-ZonMW 917.11.337, KWF UU 2010-4669, UU 2013-6426, UU 2014-6790 and UU 2015-7601, UU 2018-11393, Vrienden van het UMCU, AICR 10-0736 and 15-0049, and Gadeta to J. Kuball, Lady Tata Memorial Trust and UU 2018-11393 to Z. Sebestyen, and SFB1160 Z2 to P. Fisch. J. Villacorta Hidalgo was supported by a PhD scholarship from Deutscher Akademischer Austauschdienst (DAAD). D.M. Chudakov is supported by grant of the Ministry of Education and Science of the Russian Federation (no. 14.W03.31.0005). We are grateful to Nagesha Appukudige for supporting computational analyses; Guido Kierkels for critical

discussion; Markus Kühs, Katja Gräwe, and Bärbel Weinhold for guidance with tissue processing and IHC; and Elvira Myshkin, Andreas Gaa, and Sabine Glatzel for expert technical assistance.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

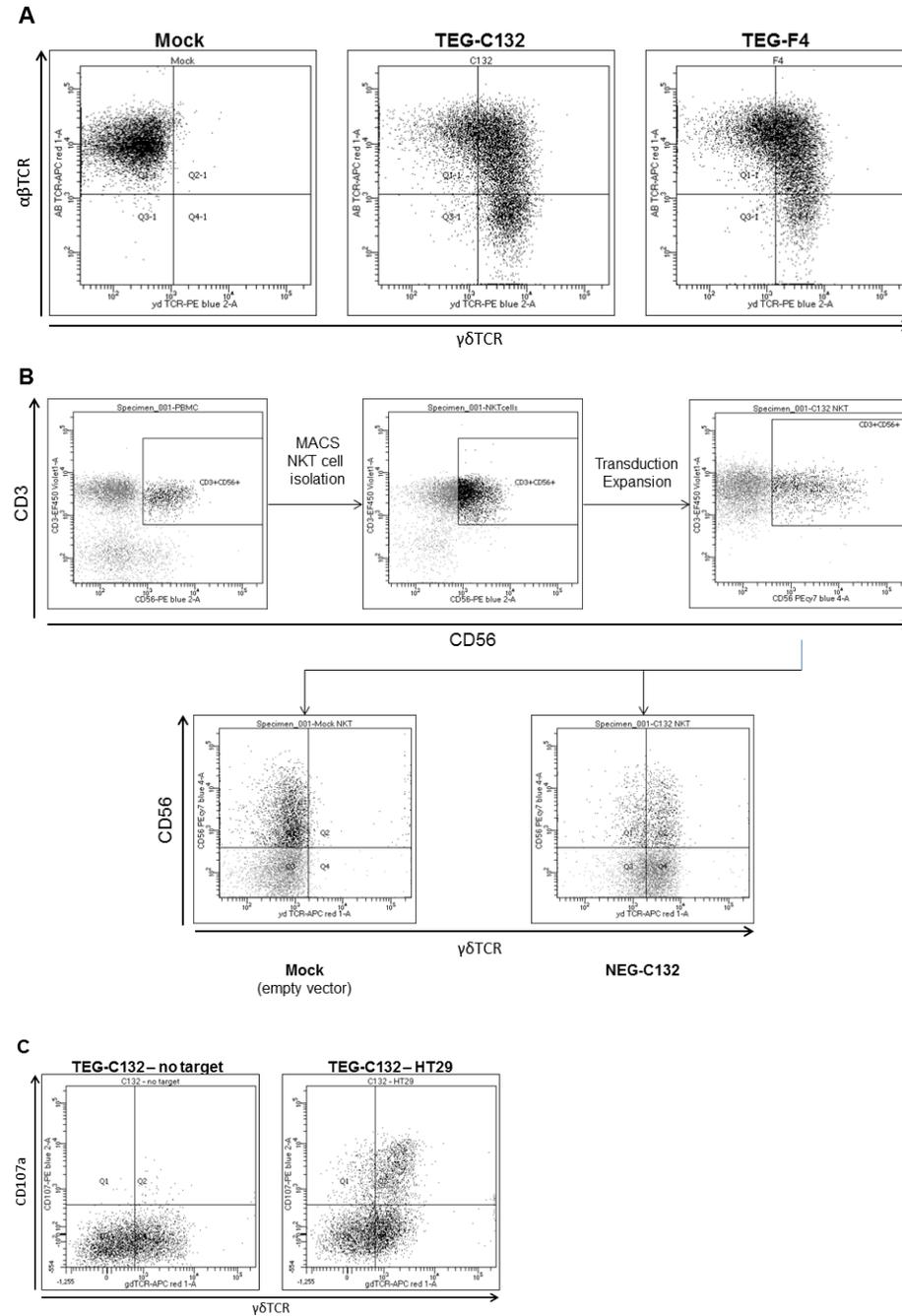
Received July 9, 2019; revised November 25, 2019; accepted January 31, 2020; published first February 4, 2020.

References

- Kwa MJ, Adams S. Checkpoint inhibitors in triple-negative breast cancer (TNBC): where to go from here. *Cancer* 2018;124:2086–103.
- Bareche Y, Venet D, Ignatiadis M, Aftimos P, Piccart M, Rothe F, et al. Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis. *Ann Oncol* 2018;29:895–902.
- Metzger-Filho O, Tutt A, de Azambuja E, Saini KS, Viale G, Loi S, et al. Dissecting the heterogeneity of triple-negative breast cancer. *J Clin Oncol* 2012;30:1879–87.
- Jitarii AA, Cimpean AM, Ribatti D, Raica M. Triple negative breast cancer: the kiss of death. *Oncotarget* 2017;8:46652–62.
- Dadi S, Chhangawala S, Whitlock BM, Franklin RA, Luo CT, Oh SA, et al. Cancer immunosurveillance by tissue-resident innate lymphoid cells and innate-like T cells. *Cell* 2016;164:365–77.
- Hidalgo JV, Bronsert P, Orłowska-Volk M, Diaz LB, Stickeler E, Werner M, et al. Histological analysis of gammadelta T lymphocytes infiltrating human triple-negative breast carcinomas. *Front Immunol* 2014;5:632.
- Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med* 2015;21:938–45.
- Wilhelm M, Kunzmann V, Eckstein S, Reimer P, Weissinger F, Ruediger T, et al. Gammadelta T cells for immune therapy of patients with lymphoid malignancies. *Blood* 2003;102:200–6.
- Coffelt SB, Kersten K, Doornebal CW, Weiden J, Vrijland K, Hau CS, et al. IL-17-producing gammadelta T cells and neutrophils conspire to promote breast cancer metastasis. *Nature* 2015;522:345–8.
- Lo Presti E, Toia F, Oieni S, Buccheri S, Turdo A, Mangiapane LR, et al. Squamous cell tumors recruit gammadelta T cells producing either IL17 or IFN γ depending on the tumor stage. *Cancer Immunol Res* 2017;5:397–407.
- Wu P, Wu D, Ni C, Ye J, Chen W, Hu G, et al. gammadeltaT17 cells promote the accumulation and expansion of myeloid-derived suppressor cells in human colorectal cancer. *Immunity* 2014;40:785–800.
- Wu Y, Kyle-Cezar F, Woolf RT, Naceur-Lombardelli C, Owen J, Biswas D, et al. An innate-like Vdelta1(+) gammadelta T cell compartment in the human breast is associated with remission in triple-negative breast cancer. *Sci Transl Med* 2019;11. pii: eaax9364.
- Morrow ES, Roseweir A, Edwards J. The role of gamma delta T lymphocytes in breast cancer: a review. *Transl Res* 2019;203:88–96.
- Sebestyen Z, Prinz I, Dechanet-Merville J, Silva-Santos B, Kuball J. Translating gammadelta (gammadelta) T cells and their receptors into cancer cell therapies. *Nat Rev Drug Discov* 2019 Sep 6 [Epub ahead of print].
- Deniger DC, Moyes JS, Cooper LJ. Clinical applications of gamma delta T cells with multivalent immunity. *Front Immunol* 2014;5:636.
- Marcu-Malina V, Heijhuurs S, van Buuren M, Hartkamp L, Strand S, Sebestyen Z, et al. Redirecting alpha beta T cells against cancer cells by transfer of a broadly tumor-reactive gammadelta T-cell receptor. *Blood* 2011;118:50–9.
- Bouchie A, DeFrancesco L, Sheridan C, Webb S. Nature Biotechnology's academic spinouts of 2016. *Nat Biotechnol* 2017;35:322–33.
- Kierkels GJJ, Scheper W, Meringa AD, Johanna I, Beringer DX, Janssen A, et al. Identification of a tumor-specific allo-HLA-restricted gammadeltaTCR. *Blood Adv* 2019;3:2870–82.
- Sebestyen Z, Scheper W, Vyborova A, Gu S, Rychnavska Z, Schiffler M, et al. RhoB mediates phosphoantigen recognition by Vgamma9Vdelta2 T cell receptor. *Cell Rep* 2016;15:1973–85.
- Bolotin DA, Poslavsky S, Davydov AN, Frenkel FE, Fanchi L, Zolotareva OI, et al. Antigen receptor repertoire profiling from RNA-seq data. *Nat Biotechnol* 2017; 35:908–11.
- Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991;19:403–10.
- Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruneri G, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol* 2015;26: 259–71.
- Denkert C, Wienert S, Poterie A, Loibl S, Budczies J, Badve S, et al. Standardized evaluation of tumor-infiltrating lymphocytes in breast cancer: results of the ring studies of the international immuno-oncology biomarker working group. *Mod Pathol* 2016;29:1155–64.
- Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 2004;10:5367–74.
- Mamedov IZ, Britanova OV, Zvyagin IV, Turchaninova MA, Bolotin DA, Putintseva EV, et al. Preparing unbiased T-cell receptor and antibody cDNA libraries for the deep next generation sequencing profiling. *Front Immunol* 2013; 4:456.
- Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012;9:671–5.
- Bender C, Zipeto D, Bidoia C, Costantini S, Zamo A, Menestrina F, et al. Analysis of colorectal cancers for human cytomegalovirus presence. *Infect Agent Cancer* 2009;4:6.
- Christopoulos P, Bukatz D, Kock S, Malkovsky M, Finke J, Fisch P. Improved analysis of TCRgammadelta variable region expression in humans. *J Immunol Methods* 2016;434:66–72.
- Brauninger A, Hansmann ML, Strickler JG, Dummer R, Burg G, Rajewsky K, et al. Identification of common germinal-center B-cell precursors in two patients with both Hodgkin's disease and non-Hodgkin's lymphoma. *N Engl J Med* 1999; 340:1239–47.
- Willemsen RA, Weijtens ME, Ronteltap C, Eshhar Z, Gratama JW, Chames P, et al. Grafting primary human T lymphocytes with cancer-specific chimeric single chain and two chain TCR. *Gene Ther* 2000;7: 1369–77.
- Voss RH, Kuball J, Theobald M. Designing TCR for cancer immunotherapy. *Methods Mol Med* 2005;109:229–56.
- Straetemans T, Grunder C, Heijhuurs S, Hol S, Slaper-Cortenbach I, Bonig H, et al. Untouched GMP-ready purified engineered immune cells to treat cancer. *Clin Cancer Res* 2015;21:3957–68.
- Grunder C, van DS, Hol S, Drent E, Straetemans T, Heijhuurs S, et al. gamma9 and delta2CDR3 domains regulate functional avidity of T cells harboring gamma9delta2TCRs. *Blood* 2012;120:5153–62.
- Soneoka Y, Cannon PM, Ramsdale EE, Griffiths JC, Romano G, Kingsman SM, et al. A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res* 1995;23: 628–33.
- Weijtens ME, Willemsen RA, Hart EH, Bolhuis RL. A retroviral vector system 'STITCH' in combination with an optimized single chain antibody chimeric receptor gene structure allows efficient gene transduction and expression in human T lymphocytes. *Gene Ther* 1998;5:1195–203.
- Scheper W, van Dorp S, Kersting S, Pietersma F, Lindemans C, Hol S, et al. gammadeltaT cells elicited by CMV reactivation after allo-SCT cross-recognize CMV and leukemia. *Leukemia* 2013;27:1328–38.
- Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods* 2015;12:380–1.

38. Shugay M, Bagaev DV, Turchaninova MA, Bolotin DA, Britanova OV, Putintseva EV, et al. VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput Biol* 2015;11:e1004503.
39. Ravens S, Schultze-Floreay C, Raha S, Sandrock I, Drenker M, Oberdorfer L, et al. Human gammadelta T cells are quickly reconstituted after stem-cell transplantation and show adaptive clonal expansion in response to viral infection. *Nat Immunol* 2017;18:393–401.
40. Davey MS, Willcox CR, Joyce SP, Ladell K, Kasatskaya SA, McLaren JE, et al. Clonal selection in the human Vdelta1 T cell repertoire indicates gammadelta TCR-dependent adaptive immune surveillance. *Nat Commun* 2017;8:14760.
41. Li B, Li T, Pignon JC, Wang B, Wang J, Shukla SA, et al. Landscape of tumor-infiltrating T cell repertoire of human cancers. *Nat Genet* 2016;48:725–32.
42. Uldrich AP, Le Nours J, Pellicci DG, Gherardin NA, McPherson KG, Lim RT, et al. CD1d-lipid antigen recognition by the gammadelta TCR. *Nat Immunol* 2013;14:1137–45.
43. Lafarge X, Pitard V, Ravet S, Roumanes D, Halary F, Dromer C, et al. Expression of MHC class I receptors confers functional intraclonal heterogeneity to a reactive expansion of gammadelta T cells. *Eur J Immunol* 2005;35:1896–905.
44. Mohammed RA, Ellis IO, Mahmmod AM, Hawkes EC, Green AR, Rakha EA, et al. Lymphatic and blood vessels in basal and triple-negative breast cancers: characteristics and prognostic significance. *Mod Pathol* 2011;24:774–85.
45. Willcox CR, Pitard V, Netzer S, Couzi L, Salim M, Silberzahn T, et al. Cytomegalovirus and tumor stress surveillance by binding of a human gammadelta T cell antigen receptor to endothelial protein C receptor. *Nat Immunol* 2012;13:872–9.
46. Scheper W, Grunder C, Straetmans T, Sebestyen Z, Kuball J. Hunting for clinical translation with innate-like immune cells and their receptors. *Leukemia* 2014;28:1181–90.
47. Scheper W, Kelderman S, Fanchi LF, Linnemann C, Bendle G, de Rooij MAJ, et al. Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers. *Nat Med* 2019;25:89–94.
48. Stevanovic S, Pasetto A, Helman SR, Gartner JJ, Prickett TD, Howie B, et al. Landscape of immunogenic tumor antigens in successful immunotherapy of virally induced epithelial cancer. *Science* 2017;356:200–5.
49. Ding Y, Ma F, Wang Z, Li B. Characteristics of the Vdelta2 CDR3 sequence of peripheral gammadelta T cells in patients with pulmonary tuberculosis and identification of a new tuberculosis-related antigen peptide. *Clin Vaccine Immunol* 2015;22:761–8.
50. Meraviglia S, Lo Presti E, Tosolini M, La Mendola C, Orlando V, Todaro M, et al. Distinctive features of tumor-infiltrating $\gamma\delta$ T lymphocytes in human colorectal cancer. *Oncoimmunology* 2017;6:e1347742.
51. Demoulin S, Herfs M, Delvenne P, Hubert P. Tumor microenvironment converts plasmacytoid dendritic cells into immunosuppressive/tolerogenic cells: insight into the molecular mechanisms. *J Leukoc Biol* 2013;93:343–52.
52. Xiang Z, Liu Y, Zheng J, Liu M, Lv A, Gao Y, et al. Targeted activation of human Vgamma9Vdelta2-T cells controls Epstein-Barr virus-induced B cell lymphoproliferative disease. *Cancer Cell* 2014;26:565–76.
53. Melandri D, Zlatareva I, Chaleil RAG, Dart RJ, Chancellor A, Nussbaumer O, et al. The gammadeltaTCR combines innate immunity with adaptive immunity by utilizing spatially distinct regions for agonist selection and antigen responsiveness. *Nat Immunol* 2018;19:1352–65.
54. Litjens NHR, van der Wagen L, Kuball J, Kwekkeboom J. Potential beneficial effects of cytomegalovirus infection after transplantation. *Front Immunol* 2018;9:389.
55. Airolidi I, Bertaina A, Prigione I, Zorzoli A, Pagliara D, Cocco C, et al. gammadelta T-cell reconstitution after HLA-haploidentical hematopoietic transplantation depleted of TCR-alpha-beta+/CD19+ lymphocytes. *Blood* 2015;125:2349–58.
56. de Witte MA, Sarhan D, Davis Z, Felices M, Vallera DA, Hinderlie P, et al. Early reconstitution of NK and gammadelta T cells and its implication for the design of post-transplant immunotherapy. *Biol Blood Marrow Transplant* 2018;24:1152–62.
57. Scheper W, Sebestyen Z, Kuball J. Cancer immunotherapy using gammadeltaT cells: dealing with diversity. *Front Immunol* 2014;5:601.
58. Straetmans T, Janssen A, Jansen K, Doorn R, Aarts T, van Muyden ADD, et al. TEG001 insert integrity from vector producer cells until medicinal product. *Mol Ther* 2020;28:561–71.

Supplementary Figure 4



Supplementary Figure 4. Examples of TCR $\gamma\delta$ expression of TEGs and NEGs and activation of TEGs. (A) Representative FACS plots show TCR $\alpha\beta$ and TCR $\gamma\delta$ expression on TEGs compared to Mock transduced cells. Related to Figure 4, 5 and 6. **(B)** FACS plots showing the isolation, transduction and expansion of NKT cells used for NEGs, related to Figure 5. **(C)** Activation of CD8+ TEG-132 as measured by CD107a after co-incubation with the tumor target HT-29 or no target. Data shown are gated on CD8+ TEGs.

Supplementary Table S1 Overview of used primers

Oligonucleotides	Sequence
NGS library preparation healthy donors	
cDNA TCR δ reverse primer constant	CTTGGATGACACGAGAT
cDNA template switch adapter with UMI	AAGCAGUGGTAUCAACGCAGAGUNNNUNNNUNNNNUCTT(rG) ⁴
PCR I nested TCR δ primer constant	AACGGATGGTTTGGTATGAG
PCR I forward primer	CACTCTATCCGACAAGCAGTGGTATCAACGCAG
PCR II nested TCR δ primer constant	TTTGGTATGAGGCTGACTTC
PCR II forward primer	CACTCTATCCGACAAGCAGT
PCR Single Cell TNBC	
Single cell PCR V γ 18-F1	AGGGGAAGGCCCCACAGCGTCTTC
Single cell PCR V γ 18-F2	CAGCGTCTTCWGTACTATGAC
Single cell PCR V γ 9-F1	TGACGGCACTGTGAGAAAGGAATC
Single cell PCR V γ 9-F2	TGAGGTGGATAGGATACCTGAAAC
Single cell PCR V δ 1-F1	ATGCAAAAAGTGGTCGCTATTCTG
Single cell PCR V δ 1-F2	CAACTCAAGAAAGCAGCGAAATC
Single cell PCR V δ 2-F1	ATACCGAGAAAAGGACATCTATG
Single cell PCR V δ 2-F2	CAAGGTGACATTGATATTGCAAAG
Single cell PCR V δ 3-F1	GGTTTTCTGTGAAACACATTCTGAC
Single cell PCR V δ 3-F2	CTTCACTTGGTGTCTCTCCAG
Single cell PCR V δ 4-F1	ATGACCAGCAAAATGCAACAGAAG
Single cell PCR V δ 4-F2	CGCTACTCATTGAATTTCCAGAAG
Single cell PCR V δ 5-F1	TACCCTGCTGAAGGTCCTACATTC
Single cell PCR V δ 5-F2	CTGTCTTCTTAAACAAAAGTGCCAAG
Single cell PCR V δ 6-F1	CCCTGCATTATTGATAGCCATACG
Single cell PCR V δ 6-F2	TGCCAAGCAGTTCTCATTGCATATC
Single cell PCR TR γ -JP1/2-R1	TTACCAGGYGAAGTTACTATGAGC
Single cell PCR TR γ -J1/2-R1	AAGTGTTGTTCCACTGCCAAAGAG
Single cell PCR TR γ -JP-R1	AAGCTTTGTTCCGGGACCAAATAC
Single cell PCR J δ 1-R1	TTGGTTCCACAGTCACACGGGTTTC
Single cell PCR J δ 2-R1	CTGGTTCCACGATGAGTTGTGTT
Single cell PCR J δ 3-R1	CAACTCACGGGGCTCCACGAAGAG
Single cell PCR J δ 4-R1	TTGTACCTCCAGATAGGTTCCTTTG
NGS library preparation TNBC samples	
Multiplex TRD δ primer NGS; hTRDV1	TCAAGAAAGCAGCGAAATCC
Multiplex TRD δ primer NGS; hTRDV2	ATTGCAAAGAACCTGGCTGT
Multiplex TRD δ primer NGS; hTRDV3	CGGTTTTCTGTGAAACACATTC
Multiplex TRD δ primer NGS; hTRDV5	ACAAAAGTGCCAAGCACCTC

Multiplex TRD δ primer NGS; hTRDC1	GACAAAAACGGATGGTTTGG
Adaptor sequences for overhang Illumina I	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG
Adaptor sequences for overhang Illumina II	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
Spectratyping	
TRGV1	CAACTTGGAAGGGRGAACRAAGTC
TRGV2	GCAAGCACAAAGGAASAACTTGAG
TRGV3	GTAATATGACGTCTCCACCG
TRGV4	ATGACTCCTACACCTCCAGC
TRGV3V5	CCCAGGAGGTGGAGCTGGAT
TRGV8	GCAAGCACAGGGAAGAGCCTTAA
TRGV9	CGGCACTGTCAGAAAGGAATC
TRGC 1-2 Rev	CAAGAAGACAAAGGTATGTTCCAG
TRGC 1-2 RO	CATCTGCATCAAGTTGTTTATC
TRDV1	ATGCAAAAAGTGGTCGCTATT
TRDV2	ATACCGAGAAAAGGACATCTATG
TRDV3	GTACCGGATAAGGCCAGATTA
TRDV4	ATGACCAGCAAAATGCAACAG
TRDV5	ACCCTGCTGAAGGTCCTACAT
TRDV6	CCCTGCATTATTGATAGCCAT
TRDC Rev	GTAGAATTCCTTACCAGACAAG
TRDC RO	GATGGTTTGGTATGAGGCTGAC
CMV nested PCR	
E-1	TCCAACACCCACAGTACCCGT
E-2	CGGAAACGATGGTGTAGTTCCG
I-1	GTCAAGGATCAGTGGCACAGC
I-2	GTAGCTGGCATTGCGATTGGT
Cloning soluble TCRs	
V γ _ecto_FW (C132 γ / Zi11 γ)	AGCGCTTCCAGCAACCTGGAAGGC
V γ _ecto_FW2 (F4 γ)	AGCGCTTCTAGTAACCTGGAGGGCGG
C γ _ecto_rv (Zi11 γ / F4 γ / C132 γ)	GGATCCGTTGCTCTTGGGGTCCATTGTGATCAC
V δ 5_ecto_FW (C132 γ)	AGCGCTGACCAGCAAGTGAAG
V δ 1_ecto_FW (F4 γ)	AGCGCTCAGAAAGTGACCCAGGCCAG
V δ 1_ecto_FW2 (Zi11 γ)	AGCGCTCAGAAAGTGACACAGGCTCAGAGC
C δ _ecto_rv (C132 γ / F4 γ / Zi11 γ)	GGATCCGCTCTTGCTGGGCTGCTTGG

Supplementary Table 2. Patient Characteristics and overview of sample usage in experiments. CMV status of tumor tissues was tested by nested PCR, qPCR and IHC as described in material and methods.

<u>Patient</u>	<u>Histolog y</u>	<u>Age</u>	<u>Stage</u>	<u>Tumor size (cm)</u>	<u>NeoAdy</u>	<u>CMV</u>	<u>Sub-type</u>	<u>Experiments</u>	<u>Figures</u>
A	MBC	59	G3 / T2N2aM0	2.8	no	Neg	Basal-like	HTS, SCS	
B	MBC	82	G3 / T1cN2M0	1.9	no	Neg	Basal-like	IHC, ST, SCS	IF- TNFa, Fig 1 d
C	IDC	44	G3 / T1cN0M0	1.9	no	Neg	Basal-like	IHC, ST, HTS, SCS	double IHC Fig1 a right, IF-CD69, Fig 1 C
D	MBC	78	G3 / T2N0M0	3.5	no	Neg	Basal-like	IHC, HTS, SCS	double IHC Fig1a left
E	MBC	53	G3 / T2N0M0	2.8	no	Neg	Basal-like	IHC, ST, HTS, SCS	IF-IFNg, Fig 1 e
F	IDC	61	G3 / T2N0M0	2.2	no	Neg	Basal-like	IHC, ST, HTS, SCS	IF-IL-17, Fig 1 f
G	MBC	75	G3 / T2N1M0	3	no	Neg	Basal-like	IHC, ST, HTS, SCS	
H	IDC	49	G3 / T3N1M0	8	no	Neg	Basal-like	IHC, SCS	
I	IDC	43	G3 / T4bN0M0	9	no	Neg	Basal-like	HTS, SCS	
K	IDC	48	G3 / T2N0M0	4.3	no	Neg	Basal-like	SCS	
Z	IDC	74	G3 / T1cN1M0	1.7	no	Neg	Basal-like	SCS	

HTS: high throughput sequencing
 IDC: invasive ductal carcinoma
 IHC: immunohistochemistry/immunofluorescence
 MBC: medullary breast carcinoma
 NeoAdy: neoadjuvant therapy
 SCS: laser microdissection followed by single cell sequencing
 ST: spectratyping

Supplementary Table 3. Overview of sequencing datasets

Dataset	Reference	Material	Sequencing	Analysis	Comment
1	This paper	Bulk TNBC tissue	HTS TCR δ chain	MiXCR, sequences included with clonal frequency of >0.1%.	
2	This paper	Laser microdissected single $\gamma\delta$ TILs in TNBC tissue	Single cell Sanger sequencing TCR γ and δ chain	NA	
3	This paper	Peripheral blood of adult healthy donors	HTS TCR δ chain	MiXCR, sequences included with clonal frequency of >0.1%.	
4	Ravens <i>et al</i> (1)	Peripheral blood of adult healthy donors	HTS TCR δ chain	MiXCR, sequences included with clonal frequency of >0.1%.	
5	Davey <i>et al.</i> (2)	Peripheral blood of adult healthy donors	HTS TCR δ chain	MiXCR, sequences included with read count of ≥ 20	Enriched for V $\delta 1$ cells
6	TCGA database & Li <i>et al.</i> (3)	Various Tumor tissues via TCGA database	TCR δ chain extracted form RNAseq data	MiXCR, RNAseq mode	

Legend: TNBC = Triple negative breast cancer, TIL= Tumor infiltrated lymphocyte, HTS = High throughput sequencing.

Supplementary Table 4. Overview of paired sequences of $\gamma\delta$ TILs from TNBC tissues

Patient	Clone	TRG Gene usage	TRG CDR3	TRD Gene usage	TRD CDR3
A	1	TRGV8*GJ1	CATWDNYKKLF	TRDV2*DJ1*1DD3	CACDTVCGTHDKLIF
A	16	TRGV3*GJ1/2	CATWDRRGKLF	TRDV2*DJ1*1DD3	CACDPLTGGLYTDKLIF
B	9 [#]	TRGV5*GJ1**	CATWDRLYYKLF ^{&}	TRDV1*DD3*DJ1	CALGNGNHIGYWRYTDKLIF
B	23 [#]	TRGV8*GJ1	CATWDNYKKLF	TRAV29/DV5DD3*DJ1*	CAASSPIRGYTGSDKLIF [%]
B	42 [#]	TRGV3*GJ1	CATWDPNNYAKGF	TRDV1*DJ1*01DD3	CALGDHVRROPTDKLIF
C	2 [#]	TRGV9*GJP	CALWEVQGELGKKIKVF	TRDV1*DD2*DJ1	CALGDLTHSTETGWPPAITDKLIF
C	6 [#]	TRGV9*GJP	CALWEVRAELGKKIEVF	TRDV3*DDR3*DJ1	CAFSVGILGDTTDKLIF
C	10	TRGV2*GJ1	CATWDRTDYKKLF	TRDV1*DD2*DJ1	CALGELSHSTETGWPPAITDKLIF
C	11 [#]	TRGV3*GJ1	CATWDRQKLLF	TRDV1*DD2*DJ1	CALGDLTAPPTTEEVIF
C	45 [#]	TRGV4*GJ1	CATWDGTTKKLF	TRDV1*DJ1*01	CALGELGASYTDKLIF
C	132 ^{#&}	TRGV4*GJ1*	CATWDGFYKLF	TRAV29/DV5DJ1DD2DD3*	CAASSPIRGYTGSDKLIF [%]
D	37 [#]	TRGV8*GJ1	CATWDNYMKLF	TRAV29/DV5 *DD2*DD3*DJ1	CAASSPIRGYTGSDKLIF [%]
D	80	TRGV5*GJ2	CATWDSPNYYKLF	TRDV1*DD1*DD2*DD3*DJ1	CALGDYLGDKYPSYDLLGDTTDKLIF
E	20 [#]	TRGV2*J1/2	CATWDGQKKLF	TRDV1*DJ1*DD3	CALGEGVSLGDYTDKLIF
E	42	TRGV9*GJ1	CALKKF	TRDV3*DD3*DJ1	CAFKGYYVGNNTDKLIF
E	49	TRGV8*GJ1	CATWDYGVFLYYKLF	TRDV3*DJ1*DD3	CALVWVILGYTDKLVF
E	113 [#]	TRGV4*GJ1	CATWDGPPYYKLF	TRAV29/DV5 *DD2DD3**DJ1	CAASSPIRGYTGSDKLIF [%]
F	2 [#]	TRGV4*GJ1	CATWDGPPYYKLF	TRDV3 *DD2*DD3*DJ1	CASSYTLKLGDTGRVDRDWKLIF
F	3	TRGV4*GJ1	CATWDDYYNLILF	TRDV1 *DD3*DJ1	CALGEPYSYWGTYTDKLIF
F	4 [#]	TRGV2GJ1	CATWDGQKKLF	TRDV1*DD3*DJ1	CALGELRYWVIVDKLIF
F	80	TRGV9*GJP	CALWELHYLGKKIKVF	TRDV3*DD2*DD3*DJ1	CAFPAFRSRGLTDKLIF
F	103	TRGV8GJ1	CATWDRLYRYKLF	TRDV3DD2DJ1	CAFSPSPYPIRFANLIF
G	9	TRGV8*GJ1	CATWDRSYKLF	TRDV1*DD3*DJ1	CALGDLGDFHYKLF
K	15	TRGV3*01	CATWDRRYKLF	TRDV3*DD3*DJ1	CAFIEPDTGYTDKLIF
Z	12i	TRGV8*GJ1	CATWDRYKLF	TRDV3*DD3*DJ1	CAELWPIRYTDKLIF
Z	e11 [#]	TRGV8*GJ1	CATWDNYKKLF	TRDV1*DD3*01DJ1	CALGDYLGDKYPSYDLLGDTTDKLIF
Z	i11 [#]	TRGV8*GJ1	CATWDNYKKLF	TRDV1*DD1*DD2*DD3*DJ1	CALGELRGQISFLYLLGDTTDKLIF

Table S4. Overview of paired sequences of $\gamma\delta$ TILs found in patients with TNBC. Grey scale colors indicate identical sequences, # indicates that these clones were used to generate TEG's for functional essays, \$ indicates the clone published by Lafarge *et al.* (4), % indicates the δ chain published by Lafarge *et al.* (4) and & indicates the γ chain published by Uldrich *et al.*(5).

Supplementary Table 5. Unique TCR γ and δ sequences in TNBC patients

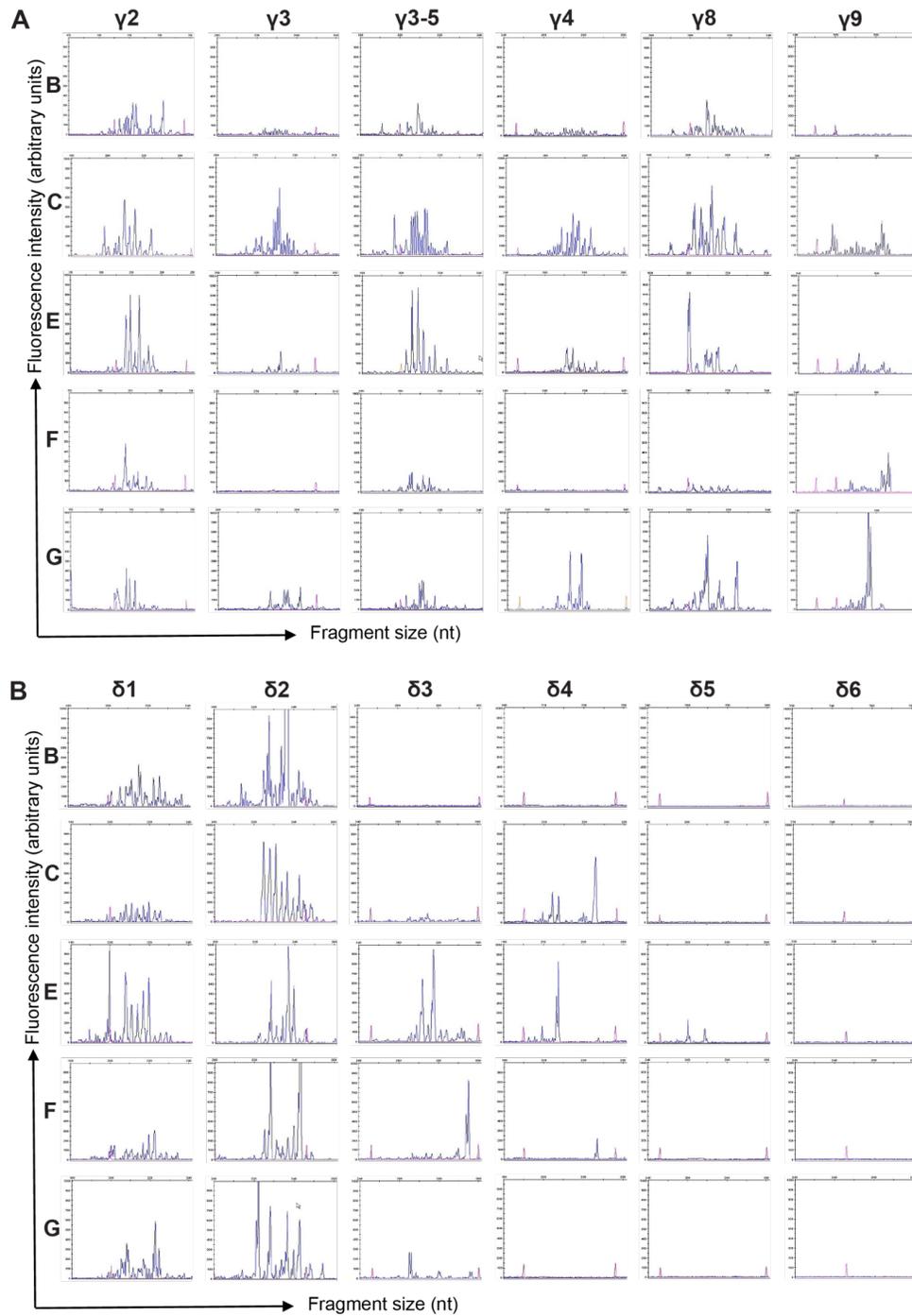
A	CACDTVCGTHDKLIF	A	CATWRQAPGISKLF	D	CATWDNYMKLF
B	CALGNNGHIGYWRYTDKLIF	A	CATWDRPLCISKLF	D	CATWDSLIIKLF
B	CALGDHVRQRPTDKLIF	A	CALWDPEFYKLLF	D	CATWDSPNYKLLF
B	CALGDEDGASRPINSSF	A	CALWEVQELGKKIKVF	E	CALKKLF
B	CACDPLGETSDKLIF	A	CALWEVQMGKRAIYKPTVF	E	CALWELHYLGKKIKVF
B	CACDPLGDRGADKLIF	A	CARMRHYYKLLF	E	CATWDGPPYYKLLF
B	CAASSPIRGYTGSDKLIF	A	CATGRVDWIKTF	E	CATWDGQKLLF
C	CALGELSHSTETGWPPAIDKLIF	A	CATWDEETPLYKLLF	E	CATWDILFHYQKLLF
C	CALGELGASYTDKLIF	A	CATWDGPSYYKLLF	E	CATWDNYKLLF
C	CALGDLTIPRLGWPPRYTDKLIF	A	CATWDNYKLLF	E	CATWDRRYKLLF
C	CALGDLTFHRPGRPPSITDKLIF	A	CATWDRLF	E	CATWDYGVFLYYKLLF
C	CALGDLTAPTEEVIF	A	CATWDRLYKLLF	E	CATWVITYMNYKLLF
C	CAFSVGILGDTTDKLIF	A	CATWDRRGKLF	E	CATWYNLYKLF
C	CAASSPIRGYTGSDKLIF	A	CATWDTLYFILKLF	F	CALWELHYLGKKIKVF
C	CAALRTGGYWLTDDKPIF	A	CATWDVELGKKIKVF	F	CATWDDYNNLILF
D	CALGDYLGDKYPSYDLLGDTTDKLIF	A	CEKLF	F	CATWDGPPYYKLLF
D	CACDALGGFTDGLSTADKILF	B	CALWEVPSNYKLLF	F	CATWDGPRKNYYKLLF
D	CAASSPIRGYTGSDKLIF	B	CATWAFGYKLLF	F	CATWDGQKLLF
E	CAMREDRGG SARLGDTRTDKLIF	B	CATWDDSRVGGKKIKVF	F	CATWDPVYKLLF
E	CALVWGLGYTDKLVF	B	CATWDGFYYKLLF	F	CATWDRLYKLLF
E	CALGEGVSLGDYTDKLIF	B	CATWDKGGYYKLLF	G	CALWEAYYKLLF
E	CALGDLISYWGISELIF	B	CATWDNYKLLF	G	CALWEPWNYKLLF
E	CAFKGYVVGNYNTDKLIF	B	CATWDRKYIMTKLF	G	CALWERYKLLF
E	CACDLLGDPGYTPKLF	B	CATWDRLVKLF	G	CALWEVLCYNKLF
E	CAASSPIRGYTGSDKLIF	B	CATWDRLYKLLF	G	CATWDGPGGKLF
E	CAASLSEDLVLTGDKLIF	B	CATWDRPGYKLLF	G	CATWDRLLYHNKLF
F	CASSYTLKLGDTPGRVRDWKLF	B	CATWDRYNKLF	G	CATWDRSYKLLF
F	CASSKLSNSGDTPGRVRYWKLF	C	CAHREKLF	G	CATWDRVPTGWFKIF
F	CALGEPYWGTYTDKLIF	C	CALWDPEFYKLLF	G	CATWEHYKLLF
F	CALGELRYWGIKLVKLF	C	CALWEGQELAKKSRGF	I	CATWDNYKLLF
F	CAFSPSPYPIRFANLIF	C	CALWELLRYTKLF	I	CATWDRLLYHKLLF
F	CAFSDSNYGLGDAKLF	C	CALWEVQELGKKIKVF	K	CATWDRRYKLLF
F	CAFPFRSRGLTDKLIF	C	CALWEVRAELGKKIEVF	Z	CATWDNYKLLF
F	CAASSPIRGYTGSDKLIF	C	CATWDGEL	Z	CATWDRYKLLF
G	CALGDPRTGGYITDKLIF	C	CATWDGFYYKLLF		
G	CALGDLGDFHYKLF	C	CATWDGLLFYKLF		
G	CAFRRPRVLLGLTDKLIF	C	CATWDGTTKLLF		
G	CACDTVLGSRNGQTQPRVTDKLIF	C	CATWDRAYYCATWTARIKLF		
G	CAASSPIRGYTGSDKLIF	C	CATWDRCATWDRNKKLF		
H	CALGNNGHIGYWRYTDKLIF	C	CATWDRQENLAKKIKVF		
H	CAFRLNDTGYTDKLIF	C	CATWDRQKLLF		
K	CAFIGRYGYTDKLIF	C	CATWDRTDYKLLF		
K	CAFIEPDTGYTDKLIF	C	CATWDVELGKKIKVF		

Z	CALGDYLGDKYPSYDLLGDTTDLIF	D	CALWELQELGKKIKVF		
Z	CALGDLGDKYSSYDLLGDTTDLIF	D	CALWGGMGKKIKVF		
Z	CAELWPIRYTDLIF	D	CATCKSYKKLF		
Z	CALGELRGQISFLYLLGDTTDLIF	D	CATWDGQKKLF		

Supplementary Table 6. Overview of shared amino acid and nucleotide sequences

	Shared amino acid sequences	Shared nucleotide sequences	% aa sequences with exact nucleotide match
Dataset 1, 3, 4 and 5	222	29	13
Dataset 2 (between donors)	3	3	100
Dataset 3 (between donors)	18	7	39
Dataset 4 (between donors)	29	13	45

Supplementary Figure 1.



Supplementary Figure 1. Spectratyping of TCR γ (A) and δ (B) of $\gamma\delta$ TILs from five representative patients. Medullary breast carcinoma: patients B and E. Invasive ductal carcinoma patients C, E, F, G. Related to Figure 2.

Supplementary Figure 2

A

Patient	Identified seq.	Unique seq.
A	18	16
B	13	11
C	16	16
D	8	7
E	12	10
F	8	7
G	9	9
H	0	0
I	2	2
K	1	1
Z	3	2
Total	90	81

Shared TCR γ sequences in triple negative breast cancer											
HD	A	B	C	D	E	F	G	H	I	K	Z
A	16										
B	2	11									
C	2	0	16								
D	0	1	0	7							
E	1	1	0	1	10						
F	0	0	0	1	3	7					
G	0	0	0	0	0	0	9				
H	ND	0									
I	1	1	0	0	1	0	0	ND	2		
K	0	0	0	0	1	0	0	ND	0	1	
Z	1	2	0	0	1	0	0	ND	1	0	2

Shared sequences	V γ	Shared by no of patients
CATWDVELGKKIKVF	V γ 2	2
CATWDRLYYKKLF	V γ 5	2
CATWDNYKKLF	V γ 8	5
CALWDPEFYKKLF	V γ 9	2
CATWDRYKLF	V γ 8	2
CATWDGFYKKLF	V γ 4	2
CATWDGQKLF	V γ 2	3
CATWDRRYKLF	V γ 5	2
CATWDGPPYKLF	V γ 4	2
CALWELHYLGKKIKVF	V γ 9	2

B

Patient	Identified seq.	Unique seq.
A	2	1
B	8	6
C	9	8
D	5	3
E	10	8
F	8	8
G	5	5
H	2	2
I	0	0
K	2	2
Z	4	4
Total	55	47

Shared TCR δ sequences in triple negative breast cancer											
HD	A	B	C	D	E	F	G	H	I	K	Z
A	1										
B	0	6									
C	0	1	8								
D	0	1	1	3							
E	0	1	1	1	8						
F	0	1	1	1	1	8					
G	0	1	1	1	1	1	5				
H	0	1	0	0	0	0	0	2			
I	ND	0									
K	0	0	0	0	0	0	0	0	ND	2	
Z	0	0	0	1	0	0	0	0	ND	0	4

Shared sequences	V δ	Shared by no of patients
CAASSPIRGYTGSDKLIF	V δ 5	6
CALGNNGNHIGYWRYTDKLIF	V δ 1	2
CALGDYLGDKYPSYDLLGDTDKLIF	V δ 1	2

Supplementary Figure 2. Shared TCR δ and γ sequences from $\gamma\delta$ TILs. Overview of shared (A) TCR γ and (B) TCR δ sequences found in tumor samples from TNBC patients (dataset 2). The total identified sequences and the unique sequences in each patient are shown. In the overlap figure, only the unique sequences are depicted. The horizontal axis shows the number of shared sequences between donors. The shared TCR chains were used in TEGs and tested for tumor reactivity in Figure 4.

Supplementary Figure 3

A

Shared V δ 2 TCR sequences between different datasets							
	Source	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
Dataset 1 - TNBC TILs HTS	TNBC tissue	482					
Dataset 2 - TNBC TILs single cell	TNBC tissue	2	7				
Dataset 3 - Healthy donors	Peripheral blood	17	0	1201			
Dataset 4 - Healthy donors Ravens <i>et al.</i>	Peripheral blood	14	0	28	1206		
Dataset 5 - Healthy donors Davey <i>et al.</i>	Peripheral blood	40	0	95	105	9800	
Dataset 6 - $\gamma\delta$ TILs	Various tumor tissue	6	0	11	9	47	385

B

Shared V δ 1 TCR sequences between different datasets							
	Source	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
Dataset 1 - TNBC TILs HTS	TNBC tissue	162					
Dataset 2 - TNBC TILs single cell	TNBC tissue	2	22				
Dataset 3 - Healthy donors	Peripheral blood	0	0	198			
Dataset 4 - Healthy donors Ravens <i>et al.</i>	Peripheral blood	1	0	0	126		
Dataset 5 - Healthy donors Davey <i>et al.</i>	Peripheral blood	0	0	0	0	27368	
Dataset 6 - $\gamma\delta$ TILs	Various tumor tissue	0	0	1	0	12	870

C

Shared V δ 3 TCR sequences between different datasets							
	Source	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
Dataset 1 - TNBC TILs HTS	TNBC tissue	43					
Dataset 2 - TNBC TILs single cell	TNBC tissue	1	13				
Dataset 3 - Healthy donors	Peripheral blood	0	0	22			
Dataset 4 - Healthy donors Ravens <i>et al.</i>	Peripheral blood	0	0	0	26		
Dataset 5 - Healthy donors Davey <i>et al.</i>	Peripheral blood	0	0	0	0	687	
Dataset 6 - $\gamma\delta$ TILs	Various tumor tissue	0	0	0	0	0	78

Supplementary Figure 3. Overview of shared TCR δ sequences among different datasets. (A) Overview of shared V δ 2 sequences among different datasets. **(B)** Overview of shared V δ 1 sequences among different datasets. **(C)** Overview of shared V δ 3 sequences among different datasets. Related to Figure 3.

Chapter VI. Supplementary References

1. Ravens S, Schultze-Florey C, Raha S, Sandrock I, Drenker M, Oberdorfer L, *et al.* Human gammadelta T cells are quickly reconstituted after stem-cell transplantation and show adaptive clonal expansion in response to viral infection. *Nat Immunol* **2017**;18(4):393-401 doi 10.1038/ni.3686.
2. Davey MS, Willcox CR, Joyce SP, Ladell K, Kasatskaya SA, McLaren JE, *et al.* Clonal selection in the human Vdelta1 T cell repertoire indicates gammadelta TCR-dependent adaptive immune surveillance. *Nat Commun* **2017**;8:14760 doi 10.1038/ncomms14760.
3. Li B, Li T, Pignon JC, Wang B, Wang J, Shukla SA, *et al.* Landscape of tumor-infiltrating T cell repertoire of human cancers. *Nature genetics* **2016** doi 10.1038/ng.3581.
4. Lafarge X, Merville P, Cazin MC, Berge F, Potaux L, Moreau JF, *et al.* Cytomegalovirus infection in transplant recipients resolves when circulating gammadelta T lymphocytes expand, suggesting a protective antiviral role. *J Infect Dis* **2001**;184(5):533-41.
5. Uldrich AP, Le Nours J, Pellicci DG, Gherardin NA, McPherson KG, Lim RT, *et al.* CD1d-lipid antigen recognition by the gammadelta TCR. *Nat Immunol* **2013**;14(11):1137-45 doi 10.1038/ni.2713.

Conclusions

This thesis characterized the TCR repertoires in multiple specialized human T cell subsets. Previously developed methods of TCR repertoires sequencing and analysis allowed us to pursue fundamental questions in the human immune system physiology, such as: the earliest emergence of immune memory (Chapter II), the phenomenon of post-thymic selection, and aging in human naïve T cells (Chapter I), the role of the T-cell receptor structure and its biophysical properties on the helper T cell acquired functions (Chapter III). We studied these issues at the systemic level using deep profiling of *ex vivo* derived TCR repertoires instead of individual antigen-specific clones or artificial models used in classical murine immunology. Thus, the series of articles that make up this thesis characterized the human T cell memory formation from a systemic immunology perspective.

Notably, we believe this study to be the first systemic description of helper T cell clonal sharing between subsets (plasticity or common clonal origin) and sharing among healthy donors (publicity of repertoire) (Chapter III). Further research is needed to find out the reason for the common shared TCR clonotypes observed among Th17, Th22 and Th2 cells, and between Th1, Th1-17 cells. Th1-17 cells could be an adjacent phenotype to Th1, with common origin from the naïve T cell state, or these two subsets could be different developmental stages of the same lineage. Also, it could be the frequent event of transition from one functional state to another with limited overlap in functions, at the stage of activated effector cells – such phenomenon we would call plasticity of effector subsets. To distinguish between the mechanisms underlying the observed overlap, we need precise clonal tracking throughout an individual T cell timeline. It is possible with transgenic TCRs or fluorescently tagged T cell clones in mice, but the correct and ethical experimental setting for tracking the fate of human T cells is yet to be designed. For now, from TCR beta chain sequencing we cannot conclude if the common naïve clonal origin or the actual plastic transition is the correct explanation for the shared repertoire.

In Chapter I, we showed that physicochemical characteristics of beta TCR change with aging in the naïve T cell subsets. In the studies of antigen-specific memory T cells, TCRs of adult and older individuals usually display longer CDR3 regions than in neonates or in a child TCR repertoire. This could be explained partially with postnatal thymic development and changes in the thymic rules for TCR rearrangement and selection: neonatal T cells have mostly short TCRs, and later thymus starts to produce more longer TCRs which can get selected in an Ag-specific manner and gain large clonal expansions in the **effector/memory**

T cell pool, especially in the peripheral organs. However, at the repertoire level of analysis, we observed shortening of CDR3 in the naïve T cells with aging, both in RTE and mature naïve T cell subsets. We link this observation with the hypothesis on preferential survival of the **naïve T cell** clones which were generated in fetal and neonatal thymus. In functional analysis of the subsets in Chapter III we observed short CDR3 in Tfh, Th1, Th1-17 and Treg cells and hypothesized that shortening of TCR CDR3 region could decrease the chance of cross-reactivity among TCRs. Therefore, preferential survival of such naïve T cells could also prevent cross-reactivity which can be beneficial in aging and aging-associated inflammatory diseases. Altogether, the observations on the individual antigen-specific TCRs and on the repertoire level can sometimes be contradictory, and the averaged subset-specific patterns are not observed in every single cell belonging to the same subset.

The functional T cell subset-specific features were analyzed in Chapter III in healthy adult donors. In future studies, it could be assessed in aging, similar to the naïve T cell aging in Chapter I. The T cells become more anergic and less functional in aging, and the frequencies of Th subsets may change. The accumulation of Tregs can be a feature of carcinogenic process, but the lack of functional Tregs is linked to various aging-associated inflammatory diseases. Does aging induce specific changes in TCR frequencies and features inside Tregs and other functional subsets? This is yet to be studied. One established aging-associated feature is the decline in TCR diversity. This metric allows to put an individual TCR repertoire on an “immune age” scale. Such estimations are general and do not reflect an actual potential of an individual’s immune system to cope with a specific antigenic challenge. Since the TCR repertoire sequencing is a rather new field, from a bioethical perspective, the TCR repertoire-derived metrics should be used with caution in clinical or diagnostic setting only after extensive validation.

Next, the research presented in this thesis studied unconventional T cells. We expanded and transferred our experience with TCR sequencing of conventional T cells' TCR repertoires to the less studied subset of human $\gamma\delta$ T lymphocytes (Chapters IV-VI). We adapted the sequencing protocol, which allows UMI incorporation to $\gamma\delta$ TCR-seq. Moreover, we adapted the TCR data processing pipeline to work with both in-house developed and other commercially available technologies (e.g., iRepertoire protocol) for gamma/delta TCR-seq. We proposed the approach to deal with contamination and estimate clonality in highly expanded T cell samples obtained from human donor tissue samples. Accurate data engineering allowed us to join efforts with two outstanding research groups in the field of $\gamma\delta$ T cell biology. Together, we succeeded to characterize new subsets of gamma/delta T cells in

human blood and peripheral tissues on examples of liver tissue and biopsy samples from breast cancer patients.

We have carried out deeper sequencing of the gamma/delta TCR repertoires than previously published. Furthermore, we found new fundamental facts about these non-classical T cell subsets. Using clonality and publicity analysis of γ and δ TCR chains repertoires, we have shown that blood-derived populations are divided into innate-like and adaptive-like immune cells. Can we compare the analysis made in this thesis for $\alpha\beta$ and $\gamma\delta$ TCR repertoires? The meta-interpretation is complicated. The observed huge clonal expansions of gamma/delta T cells are larger than any CD8+ $\alpha\beta$ T cell clonal expansion, notwithstanding the smaller CD4+ T cell memory expansions. However, we observed similar changes in transcriptomic profile concurrent with the clonal expansions; therefore, we concluded that $\gamma\delta$ T cells can achieve memory phenotype similarly to conventional CD8+ killer T cells. However, to confirm that specific $\gamma\delta$ T cell subsets undergo clonal selection and memory formation after antigen-specific recognition and activation through TCR, we need to show direct interaction between certain antigens and gamma/delta TCRs. Our repertoire-wide studies of TCRs are not enough to fulfill this purpose, and further investigation of clonal $\gamma\delta$ T cell biology is needed.

Even though we achieved a deeper understanding of functional T cell subsets and transitions between subsets, repertoire analysis does not allow us to predict an individual potential to react to various immune challenges based only on an individual TCR repertoire analysis of β TCR, or δ TCR sequencing, or even based on the paired $\alpha\beta$ and $\gamma\delta$ TCR structural data. Expanding field of TCR specificity prediction software development may help in this goal together with high-throughput assays to screen TCR antigen specificity. However, $\gamma\delta$ TCRs could recognize antigens either in CD1, EPCR, and other membrane-bound molecules as well as soluble antigens. This variability of gamma/delta TCRs recognition modes complicates the specificity identification. There is a need for more high-throughput screening systems for gamma/delta TCRs in parallel to effective transduced APC-systems and pMHC-based assays already available for the $\alpha\beta$ TCR screening.

This thesis achieved the aims formulated at the beginning of the Ph.D. research and allowed for a more comprehensive system of functional T cell diversity in the human adaptive immunity landscape, both in systemic and tissue-specific immunity. We were particularly interested in the public TCR repertoire, which is less MHC restricted and could be safe to use as T cell therapy in patients with different HLA and overall genetic backgrounds. Therefore, we described the publicity of functional helper T cell subsets in Chapter III, assessed the publicity of gamma/delta T cells in Chapter IV in steady-state, and touched the question of

public TCR repertoire function in cancer immunosurveillance in Chapter VI. Interestingly, Treg TCR repertoires display more publicity than other subsets, both in human and murine samples, in the naive and memory T cell subsets. This suggests that Tregs could be a more universal tool than the personalized therapy based on specific CARs of TCRs of killer T cells or other helper T cell subsets. Still, an additional functional analysis is needed for each of the clones to ensure that public TCR-containing T cells could be used in adoptive cell therapy whether the scope of this thesis is only limited to the larger scale view on the repertoire-wide and not clonal-specific features of the TCR structure and its role in T cell functionality.

LIST OF REFERENCES TO THESIS INTRODUCTION

1. Sallusto F. Heterogeneity of human CD4(+) T cells against microbes. *Annu Rev Immunol.* 2016;34:317-334. doi:10.1146/annurev-immunol-032414-112056
2. Lutz MB, Kurts C. Induction of peripheral CD4+ T-cell tolerance and CD8+ T-cell cross-tolerance by dendritic cells. *Eur J Immunol.* 2009;39(9):2325-2330. doi:10.1002/eji.200939548
3. Oh J, Wu N, Barczak AJ, Barbeau R, Erle DJ, Shin J-S. Cd40 mediates maturation of thymic dendritic cells driven by self-reactive cd4+ thymocytes and supports development of natural regulatory t cells. *J Immunol.* 2018;200(4):1399-1412. doi:10.4049/jimmunol.1700768
4. Onishi Y, Fehervari Z, Yamaguchi T, Sakaguchi S. Foxp3+ natural regulatory T cells preferentially form aggregates on dendritic cells *in vitro* and actively inhibit their maturation. *Proc Natl Acad Sci U S A.* 2008;105(29):10113-10118. doi:10.1073/pnas.0711106105
5. DuPage M, Bluestone JA. Harnessing the plasticity of CD4(+) T cells to treat immune-mediated disease. *Nat Rev Immunol.* 2016;16(3):149-163. doi:10.1038/nri.2015.18
6. Muranski P, Restifo NP. Essentials of Th17 cell commitment and plasticity. *Blood.* 2013;121(13):2402-2414. doi:10.1182/blood-2012-09-378653
7. Maggi L, Santarlasci V, Capone M, Rossi MC, Querci V, Mazzoni A, Cimaz R, De Palma R, Liotta F, Maggi E, Romagnani S, Cosmi L, Annunziato F. Distinctive features of classic and nonclassic (Th17 derived) human Th1 cells. *Eur J Immunol.* 2012;42(12):3180-3188. doi:10.1002/eji.201242648
8. Sonnenberg GF, Mjösberg J, Spits H, Artis D. SnapShot: innate lymphoid cells. *Immunity.* 2013;39(3):622-622.e1. doi:10.1016/j.immuni.2013.08.021
9. Bernink J, Mjösberg J, Spits H. Th1- and Th2-like subsets of innate lymphoid cells. *Immunol Rev.* 2013;252(1):133-138. doi:10.1111/imr.12034
10. Greenaway HY, Ng B, Price DA, Douek DC, Davenport MP, Venturi V. NKT and MAIT invariant TCR α sequences can be produced efficiently by VJ gene recombination. *Immunobiology.* 2013;218(2):213-224. doi:10.1016/j.imbio.2012.04.003
11. Venturi V, Rudd BD, Davenport MP. Specificity, promiscuity, and precursor frequency in immunoreceptors. *Curr Opin Immunol.* 2013;25(5):639-645. doi:10.1016/j.coi.2013.07.001
12. Davey MS, Willcox CR, Baker AT, Hunter S, Willcox BE. Recasting human $\nu\delta 1$ lymphocytes in an adaptive role. *Trends Immunol.* 2018;39(6):446-459. doi:10.1016/j.it.2018.03.003
13. Willcox BE, Willcox CR. $\gamma\delta$ TCR ligands: the quest to solve a 500-million-year-old mystery. *Nat Immunol.* 2019;20(2):121-128. doi:10.1038/s41590-018-0304-y
14. Attaf M, Legut M, Cole DK, Sewell AK. The T cell antigen receptor: the Swiss army knife of the immune system. *Clin Exp Immunol.* 2015;181(1):1-18. doi:10.1111/cei.12622
15. Janeway CA, Murphy K, Travers P, Walport M. Janeway's Immunobiology. *Garland Science.* Taylor & Francis Group, New York; 2008.
16. Jamieson BD, Douek DC, Killian S, Hultin LE, Scripture-Adams DD, Giorgi JV, et al. Generation of functional thymocytes in the human adult. *Immunity.* 1999;10: 569–575.

17. Arstila TP, Casrouge A, Baron V, Even J, Kanellopoulos J, Kourilsky P. A Direct Estimate of the Human T Cell Receptor Diversity. *Science*. 1999. pp. 958–961. doi:10.1126/science.286.5441.958
18. Shugay M, Bagaev DV, Turchaninova MA, Bolotin DA, Britanova OV, Putintseva EV, et al. VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLoS Comput Biol*. 2015;11: e1004503.
19. Vander Heiden JA, Marquez S, Marthandan N, Bukhari SAC, Busse CE, Corrie B, Hershberg U, Kleinstein SH, Matsen Iv FA, Ralph DK, Rosenfeld AM, Schramm CA, AIRR Community, Christley S, Laserson U. AIRR community standardized representations for annotated immune repertoires. *Front Immunol*. 2018;9:2206. doi:10.3389/fimmu.2018.02206
20. Bolotin DA, Shugay M, Mamedov IZ, Putintseva EV, Turchaninova MA, Zvyagin IV, Britanova OV, Chudakov DM. MiTCR: software for T-cell receptor sequencing data analysis. *Nat Methods*. 2013;10(9):813-814. doi:10.1038/nmeth.2555
21. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods*. 2015;12: 380–381.
22. Shugay M, Britanova OV, Merzlyak EM, Turchaninova MA, Mamedov IZ, Tuganbaev TR, Bolotin DA, Staroverov DB, Putintseva EV, Plevova K, Linnemann C, Shagin D, Pospisilova S, Lukyanov S, Schumacher TN, Chudakov DM. Towards error-free profiling of immune repertoires. *Nat Methods*. 2014;11(6):653-655. doi:10.1038/nmeth.2960
23. Matz M, Shagin D, Bogdanova E, Britanova O, Lukyanov S, Diatchenko L, et al. Amplification of cDNA ends based on template-switching effect and stepout PCR. *Nucleic Acids Res*. 1999;27: 1558–1560.
24. Mamedov IZ, Britanova OV, Zvyagin IV, Turchaninova MA, Bolotin DA, Putintseva EV, et al. Preparing unbiased T-cell receptor and antibody cDNA libraries for the deep next generation sequencing profiling. *Front Immunol*. 2013;4: 456.
25. Egorov ES, Merzlyak EM, Shelenkov AA, Britanova OV, Sharonov GV, Staroverov DB, et al. Quantitative profiling of immune repertoires for minor lymphocyte counts using unique molecular identifiers. *J Immunol*. 2015;194: 6155–6163.
26. Izraelson M, Nakonechnaya TO, Moltedo B, Egorov ES, Kasatskaya SA, Putintseva EV, Mamedov IZ, Staroverov DB, Shemiakina II, Zakharova MY, Davydov AN, Bolotin DA, Shugay M, Chudakov DM, Rudensky AY, Britanova OV. Comparative analysis of murine T-cell receptor repertoires. *Immunology*. 2018;153(2):133-144. doi:10.1111/imm.12857
27. Simnica D, Akyüz N, Schliffke S, Mohme M, V Wenserski L, Mährle T, Fanchi LF, Lamszus K, Binder M. T cell receptor next-generation sequencing reveals cancer-associated repertoire metrics and reconstitution after chemotherapy in patients with hematological and solid tumors. *Oncoimmunology*. 2019;8(11):e1644110. doi:10.1080/2162402X.2019.1644110
28. Simnica D, Schliffke S, Schultheiß C, Bonzanni N, Fanchi LF, Akyüz N, Gösch B, Casar C, Thiele B, Schlüter J, Lohse AW, Binder M. High-throughput immunogenetics reveals a lack of physiological t cell clusters in patients with autoimmune cytopenias. *Front Immunol*. 2019;10:1897. doi:10.3389/fimmu.2019.01897
29. Davey MS, Willcox CR, Joyce SP, Ladell K, Kasatskaya SA, McLaren JE, Hunter S, Salim M, Mohammed F, Price DA, Chudakov DM, Willcox BE. Clonal selection in the human V δ 1 T cell repertoire indicates $\gamma\delta$ TCR-dependent adaptive immune surveillance. *Nat Commun*. 2017;8:14760. doi:10.1038/ncomms14760

30. 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(1):61. doi:10.1186/s12896-017-0379-9
31. Arnaout RA, Prak ETL, Schwab N, Rubelt F and the Adaptive Immune Receptor Repertoire Community (2021) The Future of Blood Testing Is the Immunome. *Front. Immunol.* 12:626793. doi: 10.3389/fimmu.2021.626793
32. Wauters E, Thevissen K, Wouters C, Bosisio FM, De Smet F, Gunst J, Humblet-Baron S, Lambrechts D, Liston A, Matthys P, Neyts J, Proost P, Weynand B, Wauters J, Tejpar S, Garg AD. Establishing a unified covid-19 “immunome”: integrating coronavirus pathogenesis and host immunopathology. *Front Immunol.* 2020;11:1642. doi:10.3389/fimmu.2020.01642
33. Li Y, Burgman B, McGrail DJ, Sun M, Qi D, Shukla SA, Wu E, Capasso A, Lin S-Y, Wu CJ, Eckhardt SG, Mills GB, Li B, Sahni N, Yi SS. Integrated genomic characterization of the human immunome in cancer. *Cancer Res.* 2020;80(21):4854-4867. doi:10.1158/0008-5472.CAN-20-0384
34. <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-adaptive-biotechnologies-t-detect-covid-test>
35. Shugay M, Bagaev DV, Zvyagin IV, Vroomans RM, Crawford JC, Dolton G, Komech EA, Sycheva AL, Koneva AE, Egorov ES, Eliseev AV, Van Dyk E, Dash P, Attaf M, Rius C, Ladell K, McLaren JE, Matthews KK, Clemens EB, Douek DC, Luciani F, van Baarle D, Kedzierska K, Kesmir C, Thomas PG, Price DA, Sewell AK, Chudakov DM. VDJdb: a curated database of T-cell receptor sequences with known antigen specificity. *Nucleic Acids Res.* 2018 Jan 4;46(D1):D419-D427. doi: 10.1093/nar/gkx760.
36. Bagaev DV, Vroomans RMA, Samir J, Stervbo U, Rius C, Dolton G, Greenshields-Watson A, Attaf M, Egorov ES, Zvyagin IV, Babel N, Cole DK, Godkin AJ, Sewell AK, Kesmir C, Chudakov DM, Luciani F, Shugay M. VDJdb in 2019: database extension, new analysis infrastructure and a T-cell receptor motif compendium. *Nucleic Acids Res.* 2019 Oct 7. pii: gkz874. doi: 10.1093/nar/gkz874.
37. Yost KE, Chang HY, Satpathy AT. Tracking the immune response with single-cell genomics. *Vaccine.* 2020;38(28):4487-4490. doi:10.1016/j.vaccine.2019.11.035
38. Zhang, Z., Xiong, D., Wang, X. *et al.* Mapping the functional landscape of T cell receptor repertoires by single-T cell transcriptomics. *Nat Methods* **18**, 92–99 (2021). <https://doi.org/10.1038/s41592-020-01020-3>
39. Ke-Yue Ma, Alexandra A. Schonnesen, Chenfeng He, Amanda Y. Xia, Eric Sun, EuniseChen, Katherine R Sebastian, Robert Balderas, Mrinalini Kulkarni-Date, Ning Jiang. High-Throughput and High-Dimensional Single Cell Analysis of Antigen-Specific CD8+ T cells. *bioRxiv* 2021.03.04.433914; doi: <https://doi.org/10.1101/2021.03.04.433914>
40. Myers DR, Zikherman J, Roose JP. Tonic signals: why do lymphocytes bother? *Trends Immunol.* 2017;38(11):844-857. doi:10.1016/j.it.2017.06.010
41. Kilpatrick RD, Rickabaugh T, Hultin LE, Hultin P, Hausner MA, Detels R, Phair J, Jamieson BD. Homeostasis of the naive CD4+ T cell compartment during aging. *J Immunol.* 2008;180(3):1499-1507. doi:10.4049/jimmunol.180.3.1499
42. den Braber I, Mugwagwa T, Vrisekoop N, Westera L, M.gling R, de Boer AB, et al. Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. *Immunity.* 2012;36: 288–297.
43. Min H, Montecino-Rodriguez E, Dorshkind K. Reduction in the developmental potential of

- intrathymic T cell progenitors with age. *J Immunol.* 2004;173: 245–250.
44. Naylor K, Li G, Vallejo AN, Lee W-W, Koetz K, Bryl E, et al. The influence of age on T cell generation and TCR diversity. *J Immunol.* 2005;174: 7446-7452
 45. Nikolich-Zugich J. Ageing and life-long maintenance of T-cell subsets in the face of latent persistent infections. *Nat Rev Immunol.* 2008;8: 512–522.
 46. Farber DL, Yudanin NA, Restifo NP. Human memory T cells: generation, compartmentalization and homeostasis. *Nat Rev Immunol.* 2014;14(1):24-35. doi:10.1038/nri3567
 47. Sathaliyawala T, Kubota M, Yudanin N, Turner D, Camp P, Thome JJC, Bickham KL, Lerner H, Goldstein M, Sykes M, Kato T, Farber DL. Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets. *Immunity.* 2013;38(1):187-197. doi:10.1016/j.immuni.2012.09.020
 48. Sprent J, Surh CD. Normal T cell homeostasis: the conversion of naive cells into memory-phenotype cells. *Nat Immunol.* 2011;12: 478–484.
 49. Milam AAV, Bartleson JM, Buck MD, Chang C-H, Sergushichev A, Donermeyer DL, Lam WY, Pearce EL, Artyomov MN, Allen PM. Tonic TCR signaling inversely regulates the basal metabolism of CD4+ t cells. *Immunohorizons.* 2020;4(8):485-497. doi:10.4049/immunohorizons.2000055
 50. Myers DR, Wheeler B, Roose JP. mTOR and other effector kinase signals that impact T cell function and activity. *Immunol Rev.* 2019;291(1):134-153. doi:10.1111/imr.12796
 51. Myers DR, Norlin E, Vercoulen Y, Roose JP. Active tonic MTORC1 signals shape baseline translation in naive T cells. *Cell Rep.* 2019;27(6):1858-1874.e6. doi:10.1016/j.celrep.2019.04.037
 52. Britanova OV, Putintseva EV, Shugay M, Merzlyak EM, Turchaninova MA, Staroverov DB, et al. Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. *J Immunol.* 2014;192: 2689–2698.
 53. Britanova OV, Shugay M, Merzlyak EM, Staroverov DB, Putintseva EV, Turchaninova MA, Mamedov IZ, Pogorelyy MV, Bolotin DA, Izraelson M, Davydov AN, Egorov ES, Kasatskaya SA, Rebrikov DV, Lukyanov S and Chudakov DM. Dynamics of Individual T Cell Repertoires: From Cord Blood to Centenarians. *J Immunol.* 2016;196: 5005–5013.
 54. Haynes BF, Sempowski GD, Wells AF, Hale LP. The human thymus during aging. *Immunol Res.* 2000;22: 253–261.
 55. Crespo M, Martinez DG, Cerissi A, Rivera-Reyes B, Bernstein HB, Lederman MM, Sieg SF, Luciano AA. Neonatal T-cell maturation and homing receptor responses to Toll-like receptor ligands differ from those of adult naive T cells: relationship to prematurity. *Pediatr Res.* 2012;71(2):136-143. doi:10.1038/pr.2011.26
 56. Vakkila J, Äystö S, Saarinen-Pihkala UM, Sariola H. Naive cd4 + t cells can be sensitized with il-7: il-7 primes naive t cells. *Scandinavian Journal of Immunology.* 2001;54(5):501-505. doi:10.1046/j.1365-3083.2001.01001.x
 57. Jameson SC, Masopust D. Understanding subset diversity in T cell memory. *Immunity.* 2018;48(2):214-226. doi:10.1016/j.immuni.2018.02.010
 58. Zhu J, Yamane H, Paul WE. Differentiation of effector CD4 T cell populations. *Annual Review of Immunology* 2010; 28:445–489. DOI: <https://doi.org/10.1146/annurev-immunol-030409-101212>,

PMID: 20192806

59. Groom JR, Richmond J, Murooka TT, Sorensen EW, Sung JH, Bankert K, von Andrian UH, Moon JJ, Mempel TR, Luster AD. CXCR3 chemokine receptor-ligand interactions in the lymph node optimize CD4⁺ T helper 1 cell differentiation. *Immunity* 2012. 37:1091–1103. DOI: <https://doi.org/10.1016/j.immuni.2012.08.016>, PMID: 23123063
60. Vroman H, van den Blink B, Kool M. Mode of dendritic cell activation: the decisive hand in Th2/Th17 cell differentiation. Implications in asthma severity? *Immunobiology* 2015;220:254–261. DOI: <https://doi.org/10.1016/j.imbio.2014.09.016>, PMID: 25245013
61. Baumjohann D, Ansel KM. Tracking early T follicular helper cell differentiation *in vivo*. *Methods in Molecular Biology* 2015;1291:27–38. DOI: https://doi.org/10.1007/978-1-4939-2498-1_3, PMID: 25836299
62. Waickman AT, Ligons DL, Hwang S, Park JY, Lazarevic V, Sato N, Hong C, Park JH. CD4 effector T cell differentiation is controlled by IL-15 that is expressed and presented in trans. *Cytokine* 2017; 99:266–274. DOI: <https://doi.org/10.1016/j.cyto.2017.08.004>, PMID: 28807496
63. Barberis M, Helikar T, Verbruggen P. Simulation of stimulation: cytokine dosage and cell cycle crosstalk driving timing-dependent T cell differentiation. *Frontiers in Physiology* 2018;9:879. DOI: <https://doi.org/10.3389/fphys.2018.00879>, PMID: 30116196
64. Eisenbarth SC. Dendritic cell subsets in T cell programming: location dictates function. *Nature Reviews Immunology* 2019;19:89–103. DOI: <https://doi.org/10.1038/s41577-018-0088-1>, PMID: 30464294
65. Rolla S, Bardina V, De Mercanti S, Quaglino P, De Palma R, Gned D, Brusa D, Durelli L, Novelli F, Clerico M. Th22 cells are expanded in multiple sclerosis and are resistant to IFN- β . *Journal of Leukocyte Biology* 2014;96:1155–1164. DOI: <https://doi.org/10.1189/jlb.5A0813-463RR>, PMID: 25097195
66. Ryba-Stanisławowska M, Werner P, Brandt A, Mys'liwiec M, Mys'liwska J. Th9 and Th22 immune response in young patients with type 1 diabetes. *Immunologic Research* 2016; 64:730–735. DOI: <https://doi.org/10.1007/s12026-015-8765-7>, PMID: 26659093
67. Jordan MS, Boesteanu A, Reed AJ, Petrone AL, Holenbeck AE, Lerman MA, Naji A, Caton AJ. Thymic selection of CD4⁺CD25⁺ regulatory T cells induced by an agonist self-peptide. *Nat Immunol.* 2001;2(4):301-306. doi:10.1038/86302
68. Arvey A, van der Veen J, Samstein RM, Feng Y, Stamatoyannopoulos JA, Rudensky AY. Inflammation-induced repression of chromatin bound by the transcription factor Foxp3 in regulatory T cells. *Nat Immunol.* 2014;15(6):580-587. doi:10.1038/ni.2868
69. Arvey A, van der Veen J, Plitas G, Rich SS, Concannon P, Rudensky AY. Genetic and epigenetic variation in the lineage specification of regulatory T cells. *eLife* 2015;4:e07571. DOI: <https://doi.org/10.7554/eLife.07571>, PMID: 26510014
70. Ohkura N, Hamaguchi M, Morikawa H, Sugimura K, Tanaka A, Ito Y, Osaki M, Tanaka Y, Yamashita R, Nakano N, Huehn J, Fehling HJ, Sparwasser T, Nakai K, Sakaguchi S. T cell receptor stimulation-induced epigenetic changes and Foxp3 expression are independent and complementary events required for Treg cell development. *Immunity* 2012;37:785–799. DOI: <https://doi.org/10.1016/j.immuni.2012.09.010>, PMID: 23123060
71. Marks BR, Nowyhed HN, Choi JY, Poholek AC, Odegard JM, Flavell RA, Craft J. Thymic self-

- reactivity selects natural interleukin 17-producing T cells that can regulate peripheral inflammation. *Nature Immunology* 2009;10:1125–1132. DOI: <https://doi.org/10.1038/ni.1783>, PMID: 19734905
72. Feng Y, van der Veecken J, Shugay M, Putintseva EV, Osmanbeyoglu HU, Dikiy S, Hoyos BE, Moltedo B, Hemmers S, Treuting P, Leslie CS, Chudakov DM, Rudensky AY. A mechanism for expansion of regulatory T-cell repertoire and its role in self-tolerance. *Nature* 2015; 528:132–136. DOI: <https://doi.org/10.1038/nature16141>, PMID: 26605529
 73. Born WK, Aydintug MK, Brien RLO. Diversity of gamma/delta T-cell antigens, *Cell. Mol. Immunol.* 2012; July, 1–8. doi:10.1038/cmi.2012.45
 74. Inghirami G, Zhu BY, Chess L & Knowles DM. Flow cytometric and immunohistochemical characterization of the gamma/delta T-lymphocyte population in normal human lymphoid tissue and peripheral blood. *The American Journal of Pathology*, 1990;136(2), 357–367
 75. Spour EF, Leemhuis T, Janski L, Redmond R, Fillak D, Jansen J. Characterization of normal human CD3+ CD5- and gamma/delta T cell receptor positive T lymphocytes. *Clin Exp Immunol.* 1990;80(1):114-121. doi:10.1111/j.1365-2249.1990.tb06450.x
 76. Argentati K, Re F, Donnini A, Tucci MG, Franceschi C, Bartozzi B, Bernardini G, Provinciali M. Numerical and functional alterations of circulating gammadelta T lymphocytes in aged people and centenarians. *J Leukoc Biol.* 2002;72(1):65-71.
 77. Battistini L, Salvetti M, Ristori G, Falcone M, Raine CS & Brosnan CF. (1995). Gamma delta T cell receptor analysis supports a role for HSP 70 selection of lymphocytes in multiple sclerosis lesions. *Molecular Medicine*, 1995;1(5), 554–62.
 78. French JD, Roark CL, Born WK & O'Brien RL. Gammadelta T lymphocyte homeostasis is negatively regulated by beta2-microglobulin. *Journal of Immunology*, 2009;182(4), 1892–900. doi:10.4049/jimmunol.0803165
 79. Melandri D, Zlatareva I, Chaleil RAG, Dart RJ, Chancellor A, Nussbaumer O, Polyakova O, Roberts NA, Wesch D, Kabelitz D, Irving PM, John S, Mansour S, Bates PA, Vantourout P, Hayday AC. The $\gamma\delta$ TCR combines innate immunity with adaptive immunity by utilizing spatially distinct regions for agonist selection and antigen responsiveness. *Nat Immunol.* 2018;19(12):1352-1365. doi:10.1038/s41590-018-0253-5
 80. Reijneveld JF, Ocampo TA, Shahine A, Gully BS, Vantourout P, Hayday AC, Rossjohn J, Moody DB, Van Rhijn I. Human $\gamma\delta$ T cells recognize CD1b by two distinct mechanisms. *Proc Natl Acad Sci U S A.* 2020;117(37):22944-22952. doi:10.1073/pnas.2010545117
 81. Scheper W, van Dorp S, Kersting S, Pietersma F, Lindemans C, Hol S, Kuball J. $\gamma\delta$ T cells elicited by CMV reactivation after allo-SCT cross-recognize CMV and leukemia. *Leukemia*, 2013;27(6), 1328–38. doi:10.1038/leu.2012.374
 82. Vyborova A, Beringer DX, Fasci D, Karaiskaki F, van Diest E, Kramer L, de Haas A, Sanders J, Janssen A, Straetemans T, Olive D, Lidke KA, Scotet E, Lidke DS, Heck AJ, Sebestyen Z, Kuball J. $\gamma\delta$ 2T cell diversity and the receptor interface with tumor cells. *J Clin Invest.* 2020;130(9):4637-4651. doi:10.1172/JCI132489
 83. Sebestyen Z, Prinz I, Déchanet-Merville J, Silva-Santos B, Kuball J. Translating gammadelta ($\Gamma\delta$) T cells and their receptors into cancer cell therapies. *Nat Rev Drug Discov.* 2020;19(3):169-184.

doi:10.1038/s41573-019-0038-z

84. Vantourout P, Hayday A. Six-of-the-best: unique contributions of $\gamma\delta$ T cells to immunology. *Nat Rev Immunol*. 2013;13(2):88-100. doi:10.1038/nri3384
85. Almeida ARM, Amado IF, Reynolds J, Berges J, Lythe G, Molina-París C, Freitas AA. Quorum-Sensing in CD4(+) T Cell Homeostasis: A Hypothesis and a Model. *Frontiers in Immunology*, 2012;3, 125. doi:10.3389/fimmu.2012.00125
86. Butler TC, Kardar M, Chakraborty AK. Quorum sensing allows T cells to discriminate between self and nonself. *Proceedings of the National Academy of Sciences of the United States of America*, 2013;110(29), 11833–8. doi:10.1073/pnas.1222467110
87. Nakagawa S, Hara M, Seki M, Yagita H, Tagami H, Aiba S. Interaction of cutaneous stromal cells and gamma/delta T cell receptor (Tcr)-positive cells. I. V gamma 5-gamma/delta TcR+ T cells migrating from organ-cultured murine skin proliferate by co-culture with cutaneous stromal cells in the presence of interleukin-2. *Eur J Immunol*. 1993;23(7):1705-1710. doi:10.1002/eji.1830230746
88. Kohlgruber AC, Gal-Oz ST, LaMarche NM, Shimazaki M, Duquette D, Koay H-F, Nguyen HN, Mina AI, Paras T, Tavakkoli A, von Andrian U, Uldrich AP, Godfrey DI, Banks AS, Shay T, Brenner MB, Lynch L. $\gamma\delta$ T cells producing interleukin-17A regulate adipose regulatory T cell homeostasis and thermogenesis. *Nat Immunol*. 2018;19(5):464-474. doi:10.1038/s41590-018-0094-2
89. Ismaili J, Orlislagers V, Poupot R, Fournié J-J, Goldman M. Human gamma delta T cells induce dendritic cell maturation. *Clin Immunol*. 2002;103(3 Pt 1):296-302. doi:10.1006/clim.2002.5218
90. Leslie DS, Vincent MS, Spada FM, Das H, Sugita M, Morita CT, Brenner MB. CD1-mediated gamma/delta T cell maturation of dendritic cells. *J Exp Med*. 2002;196(12):1575-1584. doi:10.1084/jem.20021515
91. Di Fabrizio L, Nassef M, Ware R, Butler VP, Chess L. Human gamma delta T cells amplify IgE production by Epstein-Barr virus-activated B cells. *Trans Assoc Am Physicians*. 1991;104:155-163.
92. McMenamin C, McKersey M, Kühnlein P, Hünig T, Holt PG. Gamma delta T cells down-regulate primary IgE responses in rats to inhaled soluble protein antigens. *J Immunol*. 1995;154(9):4390-4394.
93. Raker V, Stein J, Montermann E, Maxeiner J, Taube C, Reske-Kunz AB, Sudowe S. Regulation of IgE production and airway reactivity by CD4⁻CD8⁻ regulatory T cells. *Immunobiology*. 2015;220(4):490-499. doi:10.1016/j.imbio.2014.10.022
94. Ponomarev ED, Novikova M, Yassai M, Szczepanik M, Gorski J, Dittel BN. Gamma delta T cell regulation of IFN-gamma production by central nervous system-infiltrating encephalitogenic T cells: correlation with recovery from experimental autoimmune encephalomyelitis. *Journal of Immunology* 2004;173(3), 1587–95.
95. Wohler JE, Smith SS, Zinn KR, Bullard DC, Barnum SR. (2009). Gammadelta T cells in EAE: early trafficking events and cytokine requirements. *European Journal of Immunology*, 2009;39(6), 1516–26. doi:10.1002/eji.200839176
96. Wohler JE, Smith SS, Barnum SR. Gammadelta T cells: the overlooked T-cell subset in demyelinating disease. *Journal of Neuroscience Research*, 2010;88(1), 1–6. doi:10.1002/jnr.22176
97. Ponomarev ED, Dittel BN. Gamma delta T cells regulate the extent and duration of inflammation in the central nervous system by a Fas ligand-dependent mechanism. *Journal of Immunology*,

- 2005;174(8), 4678–87.
98. Korn T, Petermann F. Development and function of interleukin 17-producing $\gamma\delta$ T cells. *Annals of the New York Academy of Sciences*, 2012;1247, 34–45. doi:10.1111/j.1749-6632.2011.06355.x
 99. Modlin RL, Pirmez C, et al. Lymphocytes bearing antigen-specific gamma delta T-cell receptors accumulate in human infectious disease lesions. *Nature*, 1989;339(6225), 544–8. doi:10.1038/339544a0
 100. Dieli F, Ivanyi J, Marsh P, Williams A, Naylor I, Sireci G, Salerno A. (2003). Characterization of lung gamma delta T cells following intranasal infection with Mycobacterium bovis bacillus Calmette-Guérin. *Journal of Immunology*, 2003;170(1), 463–9.
 101. Hoft DF, Brown RM, Roodman ST. Bacille Calmette-Guérin vaccination enhances human gamma delta T cell responsiveness to mycobacteria suggestive of a memory-like phenotype. *Journal of Immunology*, 1998;161(2), 1045–54.
 102. Cairo C, Hebbeler AM, Propp N, Bryant JL, Colizzi V, Pauza CD. Innate-like gammadelta T cell responses to mycobacterium Bacille Calmette-Guerin using the public V gamma 2 repertoire in *Macaca fascicularis*. *Tuberculosis*, 2007;87(4), 373–83. doi:10.1016/j.tube.2006.12.004
 103. Cairo C, Armstrong CL, Cummings JS, Deetz CO, Tan M, Lu C, Pauza CD. Impact of age, gender, and race on circulating $\gamma\delta$ T cells. *Human Immunology*, 2010;71(10), 968–75. doi:10.1016/j.humimm.2010.06.014
 104. Puan KJ, Low JSH, Tan TWK, et al. Phenotypic and functional alterations of Vgamma2Vdelta2 T cell subsets in patients with active nasopharyngeal carcinoma. *Cancer Immunology, Immunotherapy*, 2002;CII, 58(7), 1095–107. doi:10.1007/s00262-008-0629-8
 105. Davey MS, Willcox CR, Baker AT, Hunter S, Willcox BE. Recasting human $v\delta 1$ lymphocytes in an adaptive role. *Trends Immunol.* 2018;39(6):446-459. doi:10.1016/j.it.2018.03.003
 106. Davey MS, Willcox CR, Hunter S, Oo YH, Willcox BE. $V\delta 2^+$ t cells—two subsets for the price of one. *Front Immunol.* 2018;9:2106. doi:10.3389/fimmu.2018.02106
 107. Willcox BE, Willcox CR. $\gamma\delta$ TCR ligands: the quest to solve a 500-million-year-old mystery. *Nat Immunol.* 2019;20(2):121-128. doi:10.1038/s41590-018-0304-y
 108. Streilein JW, Takeuchi M, Taylor AW. Immune privilege, T-cell tolerance, and tissue-restricted autoimmunity. *Hum Immunol.* 1997;52(2):138-143. doi:10.1016/S0198-8859(96)00288-1
 109. Schenkel JM, Masopust D. Tissue-resident memory T cells. *Immunity.* 2014;41(6):886-897. doi:10.1016/j.immuni.2014.12.007
 110. Park CO, Kupper TS. The emerging role of resident memory T cells in protective immunity and inflammatory disease. *Nat Med.* 2015;21(7):688-697. doi:10.1038/nm.3883
 111. Misiak A, Leuzzi R, Allen AC, Galletti B, Baudner BC, D’Oro U, O’Hagan DT, Pizza M, Seubert A, Mills KHG. Addition of a TLR7 agonist to an acellular pertussis vaccine enhances Th1 and Th17 responses and protective immunity in a mouse model. *Vaccine.* 2017;35(39):5256-5263. doi:10.1016/j.vaccine.2017.08.009
 112. Kapil P, Merkel TJ. Pertussis vaccines and protective immunity. *Curr Opin Immunol.* 2019;59:72-78. doi:10.1016/j.coi.2019.03.006

113. Chasaide CN, Mills KHG. Next-generation pertussis vaccines based on the induction of protective T cells in the respiratory tract. *Vaccines (Basel)*. 2020;8(4). doi:10.3390/vaccines8040621
114. Godfrey DI, Uldrich AP, McCluskey J, Rossjohn J, Moody DB. The burgeoning family of unconventional T cells. *Nat Immunol*. 2015;16(11):1114-1123. doi:10.1038/ni.3298
115. Farber DL, Yudanin NA, Restifo NP. Human memory T cells: generation, compartmentalization and homeostasis. *Nat Rev Immunol*. 2014;14(1):24-35. doi:10.1038/nri3567
116. Thome JJC, Yudanin N, Ohmura Y, Kubota M, Grinshpun B, Sathaliyawala T, Kato T, Lerner H, Shen Y, Farber DL. Spatial map of human T cell compartmentalization and maintenance over decades of life. *Cell*. 2014;159(4):814-828. doi:10.1016/j.cell.2014.10.026
117. Thome JJC, Grinshpun B, Kumar BV, Kubota M, Ohmura Y, Lerner H, et al. Longterm maintenance of human naive T cells through in situ homeostasis in lymphoid tissue sites. *Sci Immunol*. 2016;1. doi:10.1126/sciimmunol.aah6506
118. Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, Gao R, Kang B, Zhang Q, Huang JY, Konno H, Guo X, Ye Y, Gao S, Wang S, Hu X, Ren X, Shen Z, Ouyang W, Zhang Z. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature*. 2018;564(7735):268-272. doi:10.1038/s41586-018-0694-x
119. Thommen DS, Schumacher TN. T cell dysfunction in cancer. *Cancer Cell*. 2018;33(4):547-562. doi:10.1016/j.ccell.2018.03.012
120. Woroniecka K, Chongsathidkiet P, Rhodin K, Kemeny H, Dechant C, Farber SH, Elsamadicy AA, Cui X, Koyama S, Jackson C, Hansen LJ, Johans TM, Sanchez-Perez L, Chandramohan V, Yu Y-RA, Bigner DD, Giles A, Healy P, Dranoff G, Weinhold KJ, Dunn GP, Fecci PE. T-cell exhaustion signatures vary with tumor type and are severe in glioblastoma. *Clin Cancer Res*. 2018;24(17):4175-4186. doi:10.1158/1078-0432.CCR-17-1846
121. Kallies A, Zehn D, Utzschneider DT. Precursor exhausted T cells: key to successful immunotherapy? *Nat Rev Immunol*. 2020;20(2):128-136. doi:10.1038/s41577-019-0223-7
122. Thesis cover illustration: Immunological synapse by David S. Goodsell, RCSB Protein Data Bank. doi: 10.2210/rcsb_pdb/goodsell-gallery-022