A fetal wave of human type-3 $\gamma\delta$ T cells with restricted TCR diversity persists into adulthood

Authors: Likai Tan^{1,7}, Alina Suzann Fichtner^{1,7}, Anja Bubke¹, Ivan Odak¹, Christian Schultze-Florey^{1,2}, Christian Koenecke^{1,2}, Reinhold Förster^{1,6}, Michael Jarek³, Constantin von Kaisenberg⁴, Alina Borchers⁵, Ulf Panzer⁵, Christian Krebs⁵, Sarina Ravens^{1,6,8}, Immo Prinz^{1,6,8*}

¹ Institute of Immunology, Hannover Medical School, Hannover, Germany

² Department of Hematology, Hemostasis, Oncology and Stem Cell Transplantation, Hannover Medical School, Hannover, Germany

³ Genome Analytics, Helmholtz Centre for Infection Research, Braunschweig, Germany

⁴ Department of Obstetrics, Gynecology and Reproductive Medicine, Hannover Medical School, Hannover, Germany

⁵ Translational Immunology, III. Department of Medicine, Hamburg Center for Translational Immunology (HCTI), University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁶ Cluster of Excellence RESIST (EXC 2155), Hannover Medical School, Hannover, Germany

^{7,8} These authors contributed equally to this work

*Correspondence should be addressed to Immo Prinz, Institute of Immunology, Hannover Medical School, Hannover, Germany, Phone 0049 511 532 9739, Email prinz.immo@mh-hannover.de

Keywords: Human $\gamma\delta$ T cells, paired $\gamma\delta$ TCR sequencing, single-cell RNA sequencing, $\gamma\delta$ T cell development, $\gamma\delta$ T17 cells

Abstract

Accumulating evidence suggests that the human embryonic thymus produces distinct waves of innate effector $\gamma\delta$ T cells. However, it is unclear whether this process comprises a dedicated subset of IL-17-producing $\gamma\delta$ T ($\gamma\delta$ T17) cells, like reported in mice. Here we present a novel protocol for high-throughput paired $\gamma\delta$ TCR-sequencing, which in combination with single-cell RNA-sequencing revealed a high heterogeneity of effector $\gamma\delta$ T cell clusters. While immature $\gamma\delta$ T cell clusters displayed mixed and diverse TCR, effector cell types in neonatal and adult blood segregated according to $\gamma\delta$ TCR usage. In adult samples, mature $V\delta 1^+$ T cells segregated into exhausted PD-1^{hi} and active PD-1^{low} clusters. Among $V\gamma 9V\delta 2^+$ T cell subsets, we identified distinct PLZF-positive effector $\gamma\delta$ T cell clusters with innate type-1 and type-3 T cell signatures that were already detectable in a public dataset of early embryonic thymus organogenesis. Together, this suggests that functionally distinct waves of human innate effector $\gamma\delta$ T cells including CCR6⁺ $\gamma\delta$ T17 cells develop in the early fetal thymus and persist into adulthood.

Introduction

$\gamma\delta$ T cells are an evolutionary conserved subset of T lymphocytes that can respond to microbial stimuli and provide tissue-surveillance independent of MHC-peptide recognition. Due to their seemingly intrinsic cytotoxicity towards a large array of tumors, $\gamma\delta$ T cells are the subject of many current efforts to design anti-cancer immunotherapies¹. Although generally regarded as rather innate T cells, $\gamma\delta$ T cells display a great level of phenotypic heterogeneity².

The most abundant $\gamma\delta$ T lymphocytes in human adult blood are defined by a semi-invariant T cell receptor (TCR) composed of a TCR γ chain (*TRG*) using the variable (V) segment $V\gamma 9$ (*TRGV9*) rearranged to the joining (J) segment JP (*TRGJP*), and a TCR δ chain (*TRD*) using a $V\delta 2$ segment³. Such $V\gamma 9V\delta 2^+$ T cells still display a considerable TCR diversity, but collectively respond to the intracellular accumulation of microbial and tumor-derived metabolites called phosphoantigens⁴, which somehow support binding of canonical $V\gamma 9V\delta 2^+$ T cells to butyrophilin-molecules BTN3A1⁵ and BTN2A1^{6,7}. Furthermore, $V\gamma 9V\delta 2^+$ T cells share common features with other human innate-like T cells like invariant natural killer T (NKT) and mucosa-associated invariant T (MAIT) cells, including expression of PLZF⁸. The remaining non- $V\gamma 9V\delta 2^+$ human $\gamma\delta$ T cell subsets express diverse *TRG* rearrangements, preferentially paired with $V\delta 1$ ³. Information about specific ligands binding to $V\delta 1^+$ TCRs is still fragmentary, however, recent findings showed expansion of individual $V\delta 1^+$ clones in response to environmental, e.g. viral stimuli^{9,10}.

In this setting, the correlation of $\gamma\delta$ TCR usage, TCR-specific activation and acquisition of $\gamma\delta$ T cell effector function is in the central focus of the field. In mice, distinct waves of effector $\gamma\delta$ T cells with restricted TCR diversity develop sequentially in the fetal thymus¹¹. In particular, IL-17-producing effector $\gamma\delta$ T ($\gamma\delta$ T17) cells are exclusively generated before birth and persist throughout the entire life as self-renewing, tissue-resident T cells in a diverse range of tissues^{12,13}. There are parallels in humans, as V γ 9V δ 2⁺ T cells with restricted TCR diversity are preferentially generated before birth, expand postnatally and persist into adulthood^{14–17}. However, if and to which extent human $\gamma\delta$ T cells acquire effector function already in the thymus is a matter of debate. Postnatal thymi contained $\gamma\delta$ T cells with immense TCR diversity¹⁸, but these were not equipped with effector functions such as IFN- γ or IL-17 production¹⁹. In contrast, the human fetal thymus generates invariant effector $\gamma\delta$ T cells that express invariant germline-encoded CDR3 γ and CDR3 δ repertoires²⁰. Furthermore, a recent study compared human V δ 1⁺ and V δ 2⁺ $\gamma\delta$ T cells by single-cell RNA sequencing (scRNA-seq) and concluded that both displayed parallel maturation trajectories, starting with expression of maturation-defining genes including *CCR7*, *IL7R*, and *CD27* and gradually acquiring cytotoxicity-related genes such as *NKG7*, *PRF1*, and *FCGR3A*²¹.

To directly assess the role of the TCR in guiding the activation and differentiation of individual $\gamma\delta$ T cell clones, we sorted total $\gamma\delta$ T cells from cord and adult blood and combined scRNA-seq with single-cell sequencing of paired *TRG* and *TRD* rearrangements (scTCR-seq). We found a functional heterogeneity of immature and effector $\gamma\delta$ T cells, that segregated according to $\gamma\delta$ TCR usage. Notably, we identified distinct $\gamma\delta$ T17 and Th1-like and cytotoxic effector ($\gamma\delta$ T1) subsets of V γ 9V δ 2⁺ T cells that were already present in neonatal blood and provide evidence that these might be exclusively generated in early fetal waves. In addition, we established a minimal flow cytometric panel comprising three TCR-specific and nine surface markers that reproduces the phenotypic map of human $\gamma\delta$ T cell subsets obtained by scRNA-seq.

Results

scRNA-seq reveals high heterogeneity among human $\gamma\delta$ T cells

To investigate transcriptional programs of human $\gamma\delta$ T cells at birth and in adulthood, total $\gamma\delta$ T cells were isolated from two unrelated neonatal cord blood samples (CB_donor1 and CB_donor2) and two unrelated adult peripheral blood samples (PB_donor1 and PB_donor2) via flow cytometric sorting for scRNA-seq (**Supplementary Fig. 1**). After quality control, a total of 23,760 $\gamma\delta$ T cells from all respective donors (median gene number = 1,307; median counts = 3,266) were considered for post-analysis (**Supplementary Fig. 2a**). Based on transcriptomic profiles, these $\gamma\delta$ T cells were projected by Uniform Manifold Approximation

and Projection (UMAP) and unsupervised clustering identified 12 cell clusters (c1 – c12) (**Fig. 1a**). A polarized age-dependent distribution was observed: Neonatal $\gamma\delta$ T cells dominated c1 – c5 and adult $\gamma\delta$ T cells formed the large majority of c10 – c12 (**Fig. 1b**). Clusters in the middle of the UMAP (c6 – c9) were more heterogeneous. Notably, 90.2% of c5 cells were from CB_donor1, and 96.2% of the large cluster c11 originated from PB_donor2, indicating variances of phenotypes on the individual's level (**Fig. 1c**). However, all four donors actually contributed to all 12 clusters (**Supplementary Fig. 2b**). The map of the Top-10 differentially expressed genes (DEGs) in each cluster demonstrates the distinct gene expression profiles of each cluster (**Fig. 1d**). Among the 1,934 DEGs, we found T cell-related regulatory transcription factors (TFs), cytokine receptors, cytotoxic genes, T cell activation genes, and natural killer (NK) cell receptors (**Supplementary Fig. 2c**).

Next, we wondered how the heterogeneous gene expression patterns of the identified clusters reflected their potential functional commitment. Interestingly, a proportion of DEGs was expressed on more than one identified cluster (**Fig. 1d and Supplementary Fig. 2c**), suggesting overlapping traits. To understand how the 12 identified clusters were interwoven, all DEGs were subjected to unsupervised clustering based on their average expression per cluster. This analysis identified seven gene co-expression modules (GM) (**Fig. 2a and Supplementary Table S1**), which were annotated on the basis of gene ontology enrichment analysis (**Supplementary Fig. 3**). *Vice versa*, modular scoring revealed the differential expression of these GMs among the 12 $\gamma\delta$ T cell clusters defined above (**Fig. 2b**). The gene module A (GM_A) represented an immature naïve state of T cell differentiation and was enriched within the majority of neonatal $\gamma\delta$ T cells (c1, c3). GM_A included genes such as *LEF1* and *TCF7* that are key regulators of naïve T cells^{22,23} as well as known regulatory TFs of murine immature $\gamma\delta$ T cells (e.g. *SOX4*)^{13,24} (**Fig. 2a**). The innate T cell differentiation gene module GM_B was present in a variety of $\gamma\delta$ T cell clusters (c4, c5, c6, c8, and c10) and comprised the TFs *GATA3* and *ZBTB16* (PLZF), described to be critical for the development of invariant NKT cells and mucosal-associated T cells^{25,26}. GM_C represented cell proliferation-associated genes in c2, while GM_E was highly enriched in interferon (IFN) type I signaling-related genes in c9. The module GM_G indicated acute T cell activation in c2 and c7. Of note, the type-3 immunity gene module GM_D, characterized by gene signatures connected to interleukin-17 (IL-17) producing T cells^{13,26–28}, was exclusively enriched within c8, which contained neonatal as well as adult $\gamma\delta$ T cells (**Fig. 2 and Supplementary Fig. 2b**). Finally, the major adult-derived $\gamma\delta$ T cell clusters (c11 and c12) expressed genes related to NK cells (*KLRC1*, *KLRC1*), cytotoxicity (*FCGR3A*, *BATF*, *GNLY*), and granzymes (*GZMA*, *GZMB*) that confined them to the cytotoxic T lymphocyte (CTL) response gene module (GM_F) (**Fig. 2 and Supplementary Fig. 3**).

Altogether, scRNA-seq of > 20,000 human $\gamma\delta$ T cells from cord blood and adult peripheral blood revealed a high functional heterogeneity with common and distinct cell transcriptional patterns among neonatal and adult $\gamma\delta$ T cells. Detailed analysis of DEGs identified key core gene co-expression modules that defined a naïve phenotype of $\gamma\delta$ T cells, their activation, differentiation to functional subsets, as well as type-3 and CTL effector functions.

Expression of *TRD* clones clusters immature, $V\delta 1^+$, and $V\delta 2^+$ T cells

Human $\gamma\delta$ T cells are typically classified according to their expressed $V\delta$ -chain³. To link the single-cell transcriptome profiles described above to individual $\gamma\delta$ TCR clones, we next PCR-amplified and sequenced full-length V(D)J segments in a two-step enrichment process using gene-specific primers for the *TRDC* and *TRGC* genes as described in the methods section. Cellular barcodes allowed mapping of approximately 45 to 70% of single-cell transcriptomes to individual paired *TRG* and *TRD* clones, respectively (**Fig. 3a,b and Supplementary Fig. 4a**). Focusing on the more diverse *TRD* repertoires, the identified 12 clusters largely delineated the distribution of *TRDV* genes into three main sections. An immature section included neonatal $\gamma\delta$ T cell clusters (c1 – c3) consisting of a heterogeneous mix of *TRDV1*, *TRDV2*, *TRDV3*, and a few other *TRDV* genes. The adult $\gamma\delta$ T cell clusters c11 and c12 as well as the mixed neonatal/adult cluster c7 were mainly *TRDV1*⁺, while the *TRDV2*⁺ section was composed by cord blood-derived clusters c4 and c5, mixed clusters c6, c8, c9, and adult cluster c10 (**Fig. 3a,b**). As expected, those $V\delta 2^+$ clones preferentially paired with $V\gamma 9$ chains (**Supplementary Fig. 4b**). Next, we examined $\gamma\delta$ T cell clonal expansions based on the distribution of individual paired CDR3 γ and CDR3 δ sequences (**Fig. 3c**) and corresponding Gini diversity indices (**Fig. 3d**). In general, adult $V\delta 1^+$ and $V\delta 2^+$ T cells showed a focused oligoclonal *TRD* repertoire (high Gini indices), while neonatal $V\delta 1^+$ and $V\delta 2^+$ *TRD* repertoires were rather polyclonal (lower Gini indices)^{9,10,29}. Specifically, immature neonatal $\gamma\delta$ T cells of the c1 – c3 section, which were regulated by gene modules of naïve or immature T cell (GM_A), showed no clonal expansions and Gini indices in c1 – c3 were close to zero (**Fig. 3c,d**). Strikingly, c11 cells, which mainly originated from the adult PB_donor2 (**Fig. 1b**), were largely constituted by one single *TRGV5/TRDV1* clone (**Supplementary Fig. 4a-c**), indicating that these cells accumulated due to antigen-specific clonal responses. However, we do not have any detailed information on the medical record of anonymous adult donors 1 and 2, so we can only speculate that these cells may have expanded in response to a viral infection such as CMV^{9,10,30}. In sum, scTCR-seq introduced an additional layer of information and segregated the 12 $\gamma\delta$ T cell clusters into immature, *TRDV1*⁺, and *TRDV2*⁺ sections that differed in their clonal compositions.

$V\delta 1^+$ $\gamma\delta$ T cells mature into PD-1hi and PD-1low cells

To understand the effector differentiation associated with clonal expansion of adult V δ 1⁺ T cells, we analyzed their transition from immature to effector V δ 1⁺ T cells at the transcriptional level. While V δ 1⁺ T cells of c1 were dominated by the gene module of naïve and immature T cells (GM_A), both neonatal and adult V δ 1⁺ T cells in c7 showed gene expression profiles of acute activation (GM_G). Adult effector V δ 1⁺ T cells (c11, c12) were controlled by the gene core module of CTL responses (GM_F) including *TBX21*, *EOMES*, and granzyme genes (**Fig. 2b and 4a**). Intriguingly, cells within the clonal cluster c11 expressed exhaustion signature genes such as *IKZF2*, *PDCD1*, *LAG3*, and *HLA-DRA*, whereas c12 was enriched for NK receptors, e.g. *KLRD1* and *KLRF1*, and cytotoxic genes, e.g. *FCGR3A* (CD16) and *TYROBP* (**Fig. 4b**). Based on these signatures, and supported by gene set enrichment analysis, adult V δ 1⁺ T cells of PD-1^{hi} c11 and PD-1^{low} c12 clusters can thus be classified into exhausted and active subsets, respectively (**Fig. 4c**).

Mapping of unique paired CDR3 γ /CDR3 δ sequences from the two adult donors to the V δ 1⁺ clusters revealed that the clones of c7, c11, and c12 were highly overlapping in each donor but distinct from immature cells in c1 (**Fig. 4d**). Altogether, these data support a trajectory of V δ 1⁺ T cells from an immature/ naïve state in c1 over an acute activation in c7 to differentiate into fully activated V δ 1⁺ $\gamma\delta$ CTLs (c12, PD1^{low}) or exhausted V δ 1⁺ $\gamma\delta$ CTLs (c11, PD1^{hi}) in adult individuals.

Identification of a distinct IL-17-committed V δ 2⁺ $\gamma\delta$ T cell subset

Next, we investigated the heterogeneity of neonatal and adult V δ 2⁺ T cell subsets in the *TRDV2*-dominated clusters c4 – 6 and c8 – 10. Interestingly, when comparing the 3.764 individual *TRDV2* gene clones from this scTCR-seq analysis to a large number of *TRD* repertoires obtained from 80 independent donors^{9,16} by bulk TCR sequencing of V γ 9V δ 2⁺ T cells, we noticed that the majority of all V δ 2⁺ T cells displayed public *TRDV2* gene rearrangements, i.e. exactly the same *TRDV2*-*CDR3* sequences were shared between individual samples (**Fig 5a and Supplementary Fig. 5a**). This may have been a consequence of an expansion of public fetal-derived V γ 9V δ 2 T cell in the perinatal period^{31,32}. In contrast, basically all *TRDV1* clones were private, i.e. not sharing their unique *TRDV1*-*CDR3* between individuals. Of note, V δ 2⁺ T cell clones of CB samples^{16,33} and of c8 showed high levels of public abundance (**Supplementary Fig. 5b**). In addition to their high level of shared *TRDV2* clones, these V δ 2⁺ T cells had *TRGV9*/*TRGJP* rearrangements (**Supplementary Fig. 5c**) and expressed two genes that guide lineage commitment of innate-like T cells including canonical V γ 9V δ 2⁺ T cells^{15,34}, namely *ZBTB16* (encoding PLZF) and CD161-encoding *KLRB1* (**Fig. 5b**).

Nevertheless, a number of other genes were differentially expressed among the V δ 2⁺ T cell clusters (**Fig. 5c,d**). The mostly CB-derived V δ 2⁺ T cells of c4 and c5 expressed *CCR4* and *CD40LG*, related to type-2 immunity³⁵. In contrast, the mostly adult-derived V δ 2⁺ T cells in c9 expressed type-1 immunity signature genes such as *TBX21* and genes related to cytotoxicity, e.g. *FCGR3A* (CD16), *PRF1*, and *GZMA*, suggesting differential roles of V δ 2⁺ T cells in the neonatal and adult immune system. However, clusters c6, c8, and c10 were composed of V δ 2⁺ T cells derived from both cord and adult blood donors. Of those, c6 cells displayed gene expression profiles (*CCR7*, *CD27*) of naïve T cells, and c10 cells were enriched in IFN-induced genes. Interestingly, c8 cells expressed type-3 immunity/ Th17 signature genes including *CCR6*, *RORC*, *IL23R*, and *DPP4* (**Fig. 5c,d and Supplementary Fig. 5d**), thereby identifying c8 V δ 2⁺ T cells as IL17-committed $\gamma\delta$ T cells ($\gamma\delta$ T17) that were present in neonates and adults.

To investigate the clonal relationship of naïve (c6), type-1 IFN-related (c9) or $\gamma\delta$ T17 (c8) V δ 2⁺ T cell subsets in adults with the $\gamma\delta$ T1 cells in c10, we calculated the overlap of paired CDR3 γ /CDR3 δ sequences. Strikingly, naïve (c6) and effector (c9, c10) V δ 2 T cells shared >50% to 100% of their clones in both adult samples while $\gamma\delta$ T17 cell (c8)-derived CDR3 δ sequences were using distinct TCR clonotypes/sequences (**Fig. 6a**). Furthermore, none of the five most expanded V γ 9V δ 2⁺ TRD clones in c8 were found in c6, c9, and c10 clusters of the respective donors (**Fig. 6b**). Together, this establishes a close developmental relationship of IFN effector (c9) and adult $\gamma\delta$ T1 V δ 2 T cells (c10), while $\gamma\delta$ T17 cells (c8) show distinct/ different yet public TCR repertoires. In combination with the type-3 immunity-related gene expression profile of c8 cells, this supports the hypothesis that human $\gamma\delta$ T17 cells may have a distinct early ontogenetic origin.

Human $\gamma\delta$ T17 cells exclusively originate from the earliest embryonic thymocytes

Murine $\gamma\delta$ T17 cells emerge during a defined window of embryonic development and are prewired to produce IL-17^{12,28}. To test whether human $\gamma\delta$ T17 cells arise similarly during early human ontogeny, we took advantage of a recent scRNA-seq study that resolved human thymus organogenesis in the embryonic weeks 8 to 10 (Ewk8 - 10)³⁶. Re-analyzing this valuable dataset (GEO:GSE133341), we identified four thymic $\gamma\delta$ T cell clusters (gdT_1 to 4) according to *CD3D* and *TCR* gene expression (**Fig. 7a and Supplementary Fig. 6**). Most strikingly, two smaller $\gamma\delta$ T cell clusters (labeled gdT_3 and gdT_4) exclusively originated from Ewk8 and Ewk9 thymi, while the two larger $\gamma\delta$ T cells cluster (labeled gdT_1 and gdT_2) mainly arose from Ewk10 thymocytes (**Fig. 7b**). Interestingly, applying the gene modules identified above (**Fig. 2**), the majority of Ewk10 $\gamma\delta$ T cells (gdT_1 & 2) as well as the $\alpha\beta$ T cell clusters were distinguished by gene modules of naïve and immature T cells (GM_A) (**Fig. 7c**). In contrast, $\gamma\delta$ T cells enriched in Ewk8 and Ewk9 (gdT_3 and 4), showed strong

expression of the innate T cell differentiation gene module (GM_B). Moreover, the Ewk8 and Ewk9 clusters were separated by CTL response ($\gamma\delta$ T_4) and type-3 immunity (gdT_3) gene modules (**Fig. 7c**). The heat map in **Fig. 7d** highlights the differential expression of cytotoxicity-related genes (e.g. *KLRD1*, *PLAC8*, *NKG7*, *TYROBP*, *GZMA*) on cluster gdT_4 and of $\gamma\delta$ T17 cell signature genes (e.g. *RORC*, *CCR6*, *IL23R*) on cluster gdT_3. Both subsets expressed *ZBTB16* (encoding PLZF) and *KLRB1* (encoding CD161), two genes previously described on fetal-blood V γ 9V δ 2 T cells¹⁵. Expression of *CD44* and *CD69* suggests that these two $\gamma\delta$ T cell subsets are already activated in the embryonic thymus (**Fig. 7d**). In addition, re-analysis of embryonic thymocytes from the human cell atlas³⁷ also revealed a small group of $\gamma\delta$ T cells from embryonic week 7 that expressed GM_B: Innate T cell differentiation and GM_D: Type-3 immunity (not shown). However, the low number of $\gamma\delta$ T cells in that dataset made it less evident than re-analysis of the Zeng et al. resource³⁶.

In summary, human $\gamma\delta$ T17 cells were exclusively present in the human thymus at embryonic week 8 and 9, but not week 10. Their transcriptional patterns point towards a functional pre-commitment for IL-17 production during early thymic development that is conserved and persists in adult peripheral blood $\gamma\delta$ T17 cells.

Delineating V δ 1⁺ and V δ 2⁺ effector subsets via flow cytometry

The combination of scRNA-seq and scTCR-seq identified a number of distinct phenotypes of human $\gamma\delta$ T cells, including $\gamma\delta$ T17 cells, in cord blood and adult peripheral blood. To validate and apply these results, we selected nine differentially expressed indicator surface markers that represented the phenotypes and gene modules of $\gamma\delta$ T cells for validation by flow cytometry (**Supplementary Fig. 7**). In addition to antibodies specific for V δ 1⁺, V δ 2⁺, and pan- $\gamma\delta$ TCR, the panel included nine antibodies directed against CCR7 and CD127 (*IL7R*) for central memory and naïve phenotypes, CCR4 to identify Th2-like V δ 2⁺ T cells, CD161 (*KLRB1*), CCR6, and CD26 (*DPP4*) as markers for $\gamma\delta$ T17 cells^{38,39}, CD94 (*KLRD1*) and CD16 (*FCGR3A*) representing CTL activity, and PD-1 (*PDCD1*) indicating exhaustion. With this simplified approach, we could successfully reproduce the presence of $\gamma\delta$ T17 and other $\gamma\delta$ T cell subsets in independent cord and adult blood donors. Based on the combined flow cytometric analysis of five cord blood and six adult blood samples by UMAP and unsupervised clustering, we identified ten distinct clusters (named FACS1-10) that largely corresponded with the twelve clusters identified by scRNA-seq (**Fig. 8a,b and Supplementary Fig. 8a-c**). Notably, although dimensional reduction was performed based on only nine selected surface markers, the clustering obtained from this analysis segregated sections of immature (mixed), V δ 1⁺ and V δ 2⁺ T cells (**Fig. 8c**) similarly to total scRNA-seq analyses (**Fig. 3a**). Also, $\gamma\delta$ T cells allocated to distinct clusters depending on their origin

from neonatal or adult donors (**Fig. 8d and Supplementary Fig. 8a**). A prominent exception to this rule was cluster FACS6, representing CCR6⁺CD26⁺CD161⁺Vδ2⁺ γδT17 cells. To address whether these cells comprised in cluster c8/FACS6 could be generated from hematopoietic stem cells in a postnatal thymus, we applied our FACS panel to analyze the peripheral blood γδ T cells of an adult *IL2RG*-deficient severe combined immunodeficiency (SCID) patient who was transplanted with *IL2RG*-sufficient T cell-depleted bone marrow at the age of one (**Supplementary Fig. 8d**). In line with previous reports⁴⁰, frequencies of Vγ9⁺CCR6⁺ T cell frequencies were highly heterogeneous in healthy control donors, however basically undetectable in the reconstituted SCID patient.

The phenotype, composition, key markers and interrelatedness of γδ T cell clusters defined by flow cytometry versus scRNA-seq is summarized in **Fig. 8e**. In brief, the flow cytometry panel comprising three γδTCR and nine surface molecule markers not only validated the prior scRNA-seq data, but provides a simple strategy to monitor the diverse γδ T cell subsets identified in neonatal and adult blood.

DISCUSSION

In this work, scRNA-seq of sorted adult and neonatal γδ T cells allowed us to establish a comprehensive map of γδ T cell phenotypes and to associate them to paired γδTCR sequences. In essence, functional diversity of human γδ T cells, as monitored by DEGs, largely correlated with the type of γδTCRs expressed in each cell. While polyclonal neonatal cells dominated the more “naïve” first half of clusters (c1 to c6), characterized by expression of genes involved in innate T cell differentiation, the other half of clusters (c7 to c12) showed distinct patterns of T cell activation, proliferation, and lineage-specific differentiation. Furthermore, we extracted seven GMs^{23,24} that explained the naïve, proliferating, type-3 immunity, cytotoxic response, IFN-induced response, NKT differentiation, and acute activation phenotypes in human γδ T cells. Thereby, compositions of naïve, IFN-induced, and cytotoxic GMs resembled those identified from human αβ T cells²³. Most importantly, our GMs allowed us to assess and interpret the phenotypes of embryonic γδ thymocytes from Zeng et al³⁶, and we suggest that these GMs will serve as a universal tool to understand the transcriptional programming of human γδ T cells. Furthermore, we provide a FACS panel that, in addition to TCR-specific mAbs, requires only nine surface makers to largely reproduce the phenotypic map obtained by scRNA-seq.

A previous single cell analysis study²¹ identified γδ T cells in sets of human circulating and tumor-infiltrating cells. They suggested that Vδ1⁺ and Vδ2⁺ subsets form close yet distinct subclusters that similarly undergo a “cytotoxic maturation”, i.e. a gradual acquisition of CTL

capacity. Our single cell analysis of large numbers of sorted $\gamma\delta$ T cells and simultaneous analysis of their paired *TRG/TRD* repertoires now adds several levels of resolution to the understanding of human $\gamma\delta$ T cell subsets.

Firstly, comparison of cord and adult blood samples was consistent with findings that neonatal $\gamma\delta$ T cells are more immature and polyclonal^{16,33,41}, whereas adult $\gamma\delta$ T cells are more activated and show restricted clonality^{9,10,42}. At the same time, cord blood and early fetal thymus contained innate $\gamma\delta$ T17 and $\gamma\delta$ T1 effector populations that presumably persisted into adulthood, sharing some features with Lin28b-dependent effector $\gamma\delta$ T cells, recently discovered in week 17-19 fetal thymus²⁰.

Regarding the distribution of clonotypes, cord (and adult) blood-derived $\gamma\delta$ T cells in the immature subsets generally displayed a diverse V δ 1-dominated repertoire, with a high frequency of TCRs containing *TRDV3* and other non-*TRDV1* or *TRDV2* rearrangements. The high expression of *SOX4*, *ID2*, and *ID3* thereby suggests that these cells recently emigrated from thymus^{10,18,41}. From the naïve state, we found a stepwise activation trajectory of V δ 1⁺ T cells that goes hand-in-hand with clonal expansion and expression of genes associated with CTL response, but also acquisition of exhaustion markers like PD-1 similar to CD8⁺ CTLs after intensive antigen-exposure⁴³. Together with the notion that repertoire focusing and proportions of V δ 1⁺ CTLs vary drastically among adults¹⁰, this fits well with the hypothesis that naive V δ 1⁺ T cells differentiate into V δ 1⁺ CTLs and clonally expand after antigen-specific TCR-dependent activation.

In contrast, canonic V γ 9V δ 2⁺ T cells rather grouped with each other but separated into three effector types with strong parallels to lineage-specific differentiation of other PLZF⁺ innate lymphocytes such as NKT-1, NKT-2, and NKT-17 cells⁴⁴. As an important disclaimer, certainly not all V δ 2-chains must pair with V γ 9-chains and *vice versa*^{3,9,20,29}. However, besides the expression of PLZF, an abundance of public TCR clones, polyclonal expansion after phosphoantigen stimulation, and the lack of an exhausted phenotype suggest an innate rather than adaptive phenotype of V γ 9V δ 2⁺ cells¹⁶.

While V γ 9V δ 2⁺ cells with type-2 immunity-related markers including *CCR4* and *CD40LG* were rather restricted to the neonatal samples, we observed committed $\gamma\delta$ T17 and $\gamma\delta$ T1 cells already in the earliest wave of thymus organogenesis, as well as in cord blood and adult blood samples. Notably, this is consistent with the hypothesis that human innate V γ 9V δ 2⁺ effector $\gamma\delta$ T cells develop in early fetal waves and persist into adulthood, as it has been described for mouse $\gamma\delta$ T17 and $\gamma\delta$ T1 cells^{12,45}. So why was this not described before? Timing might be critical for their observation, since V γ 9V δ 2⁺ cells are certainly present very early in T cell ontogeny^{14,15} but hard to detect in later fetal thymus (week 17-19)²⁰ and

absent in postnatal thymus¹⁹. In conclusion, the earliest innate $\gamma\delta$ T17 and $\gamma\delta$ T1 cells probably leave the human fetal thymus and home to distant tissues already after gestational week 9, in line with our detection of these cells in thymi of week 8 and 9, but not week 10. Such a strict fetal origin of effector cells is literally a feature of innateness and is in parallel to the ontogeny of IL-17-producing lymphoid tissue-inducer cells that seed and persist in secondary lymphoid tissues⁴⁶. Likewise, PLZF expression maps the early stages of ILC1 lineage development⁴⁷.

Furthermore, there is currently no uniform definition of human $\gamma\delta$ T17 cells. Several reports identified that subsets of $V\delta 2^+$ cells expressing IL-23R⁴⁸, CD161^{38,49} or CCR6^{38,40,50,51} are the ones prone to IL-17 production. However, these reports concluded that peripheral differentiation of $V\gamma 9V\delta 2^+$ T cells into IL-17 producers required a highly inflammatory milieu. Here we define human type-3 $V\gamma 9V\delta 2^+$ T cells as $\gamma\delta$ T17 cells by expression of canonical $V\gamma 9V\delta 2^+$ TCR and the GM type-3 immunity comprising *IL23R*, *RORC*, and *BLK*, and by FACS as CCR6⁺, DPP4 (CD26)⁺, KLRB1 (CD161)⁺, and negative for KLRD1 (CD94) and FC γ R3a (CD16). However, in sum it emerges that CCR6, like in mice⁵², is the most stringent single marker for human $\gamma\delta$ T17 cells.

A very recent report⁵³ designated CD26⁺CD94⁻ $V\delta 2^+$ cells as a distinct population with a certain overlap to previously described CD28⁺CD16⁻ $V\delta 2^+$ cells⁴⁰. Notably, both studies concluded that the earliest neonatal $V\delta 2^+$ cells are enriched in cells with an CD26⁺CD28⁺CCR6⁺ phenotype, but would later transition into a more mature profile of $V\delta 2^+$ cells^{40,53}. However, from our scRNA-seq and FACS data it is evident that CD26⁺CD94⁻ $V\delta 2^+$ cells are actually a heterogeneous population comprising CCR6⁺ and CCR6⁻ cells. Consistent with the finding of Ryan et al.⁴⁰, we found that there are high levels of phenotypic variability between donors, with some profiles dominated by CCR6⁺ $V\delta 2^+$ cells and others that displayed only very few CCR6⁺ $V\delta 2^+$ cells, i.e. around 1%. However, we would speculate that these profiles are already established very early in ontogeny, perhaps in a neonatal window of opportunity, and persist and influence immune responses during an individual's entire life. Only the analysis of larger long-term cohorts monitoring individuals from infancy to adulthood will be able to address this issue.

Notwithstanding, we propose the obvious, namely that human CCR6⁺ $\gamma\delta$ T17, similarly to mouse CCR6⁺ $\gamma\delta$ T17 cells¹², are preferentially or even exclusively generated early before birth. This hypothesis is supported by i) the finding that an adult *IL2RG*-deficient SCID patient who was transplanted with *IL2RG*-sufficient T cell-depleted bone marrow at the age of one was lacking CCR6⁺ $V\gamma 9^+$ T cells, ii) the direct identification of these cells in published dataset³⁶ of fetal thymus Ewk8-9, but not Ewk10, and iii) the presence of the type-3

immunity-biased cluster c8 cells shared in samples from cord and adult blood, and iv) paired TCR analyses revealing that c8 clusters showed unique TCR profiles different to the other V γ 9V δ 2⁺ cell-dominated clusters (c6, c9, c10) with more overlapping TCR repertoires. Nevertheless, the physiological functions of human $\gamma\delta$ T17 cells remain to be established.

Taken together, we present a comprehensive map of functional human $\gamma\delta$ T cell subsets and provide moderately complex FACS panel that can be easily implemented to reproduce this classification. This should contribute to decipher the factors that drive the individual TCR and functional subset composition of individuals early in life and to understand whether these parameters would influence or predict how individuals will cope with neoplastic and infectious challenges.

Acknowledgements

The work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2155 “RESIST” – Project ID 39087428 to R.F., S.R. and I.P.; SFB900 – Project ID 158989968 to R.F., C.Könecke, S.R. and I.P. and DFG-funded research group FOR2799 to S.R. (Project ID RA3077/1-1) and I.P. (Project ID PR727/11-1). We would like to thank Matthias Ballmeier of the MHH cell sorting facility and the Genomics Core Unit of Hannover Medical School.

Author contributions

L.T., A.F., S.R. and I.P. designed the study and experiments; L.T. analyzed data; A.F. contributed in data analysis; A.F., I.O., A.B. and S.R. conducted experiments; M.J. provided bioinformatics analysis tools; C.S.-F., C.Könecke and C.v.K. recruited and coordinated study participants; R.F., C. Krebs and U.P. provided technical support and intellectual input and L.T., A.F., S.R. and I.P. wrote the manuscript.

Declaration of competing interests

The authors declare no conflict of interest.

References

1. 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.* **19**, 169–184 (2020).
2. Vantourout, P. & Hayday, A. Six-of-the-best: unique contributions of $\gamma\delta$ T cells to immunology. *Nat. Rev. Immunol.* **13**, 88–100 (2013).
3. Fichtner, A. S., Ravens, S. & Prinz, I. Human $\gamma\delta$ TCR Repertoires in Health and Disease. *Cells* **9**, 800 (2020).
4. Constant, P. *et al.* Stimulation of human gamma delta T cells by nonpeptidic mycobacterial ligands. *Science (80-.)*. **264**, 267–270 (1994).
5. Halary, F. *et al.* Shared reactivity of V δ 2neg $\gamma\delta$ T cells against cytomegalovirus-infected cells and tumor intestinal epithelial cells. *J. Exp. Med.* **201**, 1567–1578 (2005).
6. Karunakaran, M. M. *et al.* Butyrophilin-2A1 Directly Binds Germline-Encoded Regions of the V γ 9V δ 2 TCR and Is Essential for Phosphoantigen Sensing. *Immunity* **52**, 487-498.e6 (2020).
7. Rigau, M. *et al.* Butyrophilin 2A1 is essential for phosphoantigen reactivity by $\gamma\delta$ T cells. *Science (80-.)*. **367**, eaay5516 (2020).
8. Provine, N. M. *et al.* Unique and Common Features of Innate-Like Human V δ 2+ $\gamma\delta$ T Cells and Mucosal-Associated Invariant T Cells. *Front. Immunol.* **9**, (2018).
9. Ravens, S. *et al.* Human $\gamma\delta$ 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).
10. Davey, M. S. *et al.* Clonal selection in the human V δ 1 T cell repertoire indicates $\gamma\delta$ TCR-dependent adaptive immune surveillance. *Nat. Commun.* **8**, 1–15 (2017).
11. Prinz, I., Silva-Santos, B. & Pennington, D. J. D. J. Functional development of $\gamma\delta$ T cells. *Eur. J. Immunol.* **43**, 1988–1994 (2013).
12. Haas, J. D. J. D. *et al.* Development of Interleukin-17-Producing $\gamma\delta$ T Cells Is Restricted to a Functional Embryonic Wave. *Immunity* **37**, 48–59 (2012).
13. Tan, L. *et al.* Single-Cell Transcriptomics Identifies the Adaptation of Scart1+ V γ 6+ T Cells to Skin Residency as Activated Effector Cells. *Cell Rep.* **27**, 3657–3671 (2019).

14. McVay, L. D. & Carding, S. R. Extrathymic origin of human gamma delta T cells during fetal development. *J. Immunol.* (1996).
15. Dimova, T. *et al.* Effector V γ 9V δ 2 T cells dominate the human fetal $\gamma\delta$ T-cell repertoire. *Proc. Natl. Acad. Sci.* **112**, E556–E565 (2015).
16. Fichtner, A. S. *et al.* TCR repertoire analysis reveals phosphoantigen-induced polyclonal proliferation of V γ 9V δ 2 T cells in neonates and adults. *J. Leukoc. Biol.* **107**, 1023–1032 (2020).
17. Vermijlen, D. & Prinz, I. Ontogeny of innate T lymphocytes - some innate lymphocytes are more innate than others. *Front. Immunol.* **5**, (2014).
18. Di Lorenzo, B., Ravens, S. & Silva-Santos, B. High-throughput analysis of the human thymic V δ 1+ T cell receptor repertoire. *Sci. data* **6**, 115 (2019).
19. Ribot, J. C., Ribeiro, S. T., Correia, D. V., Sousa, A. E. & Silva-Santos, B. Human $\gamma\delta$ Thymocytes Are Functionally Immature and Differentiate into Cytotoxic Type 1 Effector T Cells upon IL-2/IL-15 Signaling. *J. Immunol.* **192**, 2237–2243 (2014).
20. Tieppo, P. *et al.* The human fetal thymus generates invariant effector $\gamma\delta$ T cells. *J. Exp. Med.* **217**, (2019).
21. Pizzolato, G. *et al.* Single-cell RNA sequencing unveils the shared and the distinct cytotoxic hallmarks of human TCRV δ 1 and TCRV δ 2 $\gamma\delta$ T lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 11906–11915 (2019).
22. Xing, S. *et al.* Tcf1 and Lef1 transcription factors establish CD8+ T cell identity through intrinsic HDAC activity. *Nat. Immunol.* **17**, 695–703 (2016).
23. Szabo, P. A. *et al.* Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat. Commun.* **10**, 4706 (2019).
24. Sagar *et al.* Deciphering the regulatory landscape of fetal and adult $\gamma\delta$ T-cell development at single-cell resolution. *EMBO J.* (2020)
doi:10.15252/embj.2019104159.
25. Gérart, S. *et al.* Human iNKT and MAIT cells exhibit a PLZF-dependent proapoptotic propensity that is counterbalanced by XIAP. *Blood* **121**, 614–623 (2013).
26. Narayan, K. *et al.* Intrathymic programming of effector fates in three molecularly distinct $\gamma\delta$ T cell subtypes. *Nat. Immunol.* **13**, 511–518 (2012).

27. Venken, K. *et al.* ROR γ t inhibition selectively targets IL-17 producing iNKT and $\gamma\delta$ -T cells enriched in Spondyloarthritis patients. *Nat. Commun.* **10**, 1–15 (2019).
28. Spidale, N. A. *et al.* Interleukin-17-Producing $\gamma\delta$ T Cells Originate from SOX13 + Progenitors that Are Independent of $\gamma\delta$ TCR Signaling. *Immunity* **49**, 857-872.e5 (2018).
29. Davey, M. S. *et al.* The human V δ 2+ T-cell compartment comprises distinct innate-like V γ 9+ and adaptive V γ 9- subsets. *Nat. Commun.* **9**, 1760 (2018).
30. Pitard, V. *et al.* Long-term expansion of effector/memory V δ 2– $\gamma\delta$ T cells is a specific blood signature of CMV infection. *Blood* **112**, 1317–1324 (2008).
31. Papadopoulou, M. *et al.* Fetal public V γ 9V δ 2 T cells expand and gain potent cytotoxic functions early after birth. *Proc. Natl. Acad. Sci. U. S. A.* (2020) doi:10.1073/pnas.1922595117.
32. Ravens, S. *et al.* Microbial exposure drives polyclonal expansion of innate $\gamma\delta$ T cells immediately after birth. *Proc. Natl. Acad. Sci.* (2020) doi:10.1073/pnas.1922588117.
33. Papadopoulou, M. *et al.* TCR Sequencing Reveals the Distinct Development of Fetal and Adult Human V γ 9V δ 2 T Cells. *J. Immunol.* **203**, 1468–1479 (2019).
34. Lu, Y., Cao, X., Zhang, X. & Kovalovsky, D. PLZF Controls the Development of Fetal-Derived IL-17 + V γ 6 + $\gamma\delta$ T Cells. *J. Immunol.* **195**, 4273–4281 (2015).
35. Engel, I. *et al.* Innate-like functions of natural killer T cell subsets result from highly divergent gene programs. *Nat. Immunol.* **17**, 728–739 (2016).
36. Zeng, Y. *et al.* Single-Cell RNA Sequencing Resolves Spatiotemporal Development of Pre-thymic Lymphoid Progenitors and Thymus Organogenesis in Human Embryos. *Immunity* **51**, 930-948.e6 (2019).
37. Park, J.-E. *et al.* A cell atlas of human thymic development defines T cell repertoire formation. *Science (80-.).* **367**, eaay3224 (2020).
38. Caccamo, N. *et al.* Differentiation, phenotype, and function of interleukin-17-producing human v γ 9v δ 2 T cells. *Blood* **118**, 129–138 (2011).
39. Bengsch, B. *et al.* Human Th17 Cells Express High Levels of Enzymatically Active Dipeptidylpeptidase IV (CD26). *J. Immunol.* (2012) doi:10.4049/jimmunol.1103801.
40. Ryan, P. L. *et al.* Heterogeneous yet stable V δ 2(+) T-cell profiles define distinct

- cytotoxic effector potentials in healthy human individuals. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 14378–14383 (2016).
41. Heiden, M. *et al.* Characterization of the $\gamma\delta$ T-cell compartment during infancy reveals clear differences between the early neonatal period and 2 years of age. *Immunol. Cell Biol.* **98**, 79–87 (2020).
42. Kallemeijn, M. J. *et al.* Next-generation sequencing analysis of the human TCR $\gamma\delta$ + T-cell repertoire reveals shifts in V γ - and V δ -usage in memory populations upon aging. *Front. Immunol.* (2018) doi:10.3389/fimmu.2018.00448.
43. Wherry, E. J. T cell exhaustion. *Nature Immunology* vol. 12 492–499 (2011).
44. Lee, Y. J. *et al.* Lineage-Specific Effector Signatures of Invariant NKT Cells Are Shared amongst $\gamma\delta$ T, Innate Lymphoid, and Th Cells. *J. Immunol.* **197**, 1460–1470 (2016).
45. Ribot, J. C. *et al.* CD27 is a thymic determinant of the balance between interferon- γ - and interleukin 17-producing $\gamma\delta$ T cell subsets. *Nat. Immunol.* (2009) doi:10.1038/ni.1717.
46. Cupedo, T. *et al.* Human fetal lymphoid tissue–inducer cells are interleukin 17–producing precursors to RORC+ CD127+ natural killer–like cells. *Nat. Immunol.* **10**, 66–74 (2009).
47. Constantinides, M. G. *et al.* PLZF expression maps the early stages of ILC1 lineage development. *Proc. Natl. Acad. Sci. U. S. A.* (2015) doi:10.1073/pnas.1423244112.
48. Moens, E. *et al.* IL-23R and TCR signaling drives the generation of neonatal V 9V 2 T cells expressing high levels of cytotoxic mediators and producing IFN- and IL-17. *J. Leukoc. Biol.* (2011) doi:10.1189/jlb.0910501.
49. Maggi, L. *et al.* CD161 is a marker of all human IL-17-producing T-cell subsets and is induced by RORC. *Eur. J. Immunol.* **40**, 2174–2181 (2010).
50. Singh, S. P., Zhang, H. H., Foley, J. F., Hedrick, M. N. & Farber, J. M. Human T Cells That Are Able to Produce IL-17 Express the Chemokine Receptor CCR6. *J. Immunol.* (2008) doi:10.4049/jimmunol.180.1.214.
51. Laggner, U. *et al.* Identification of a Novel Proinflammatory Human Skin-Homing V γ 9V δ 2 T Cell Subset with a Potential Role in Psoriasis. *J. Immunol.* **187**, 2783–2793 (2011).

52. Haas, J. D. *et al.* CCR6 and NK1.1 distinguish between IL-17A and IFN- γ -producing $\gamma\delta$ effector T cells. *Eur. J. Immunol.* **39**, 3488–3497 (2009).
53. Wragg, K. M. *et al.* High CD26 and Low CD94 Expression Identifies an IL-23 Responsive V δ 2+ T Cell Subset with a MAIT Cell-like Transcriptional Profile. *Cell Rep.* **31**, 107773 (2020).

Figure 1: scRNA-seq reveals high heterogeneity among human $\gamma\delta$ T cells.

Single-cell transcriptome libraries were generated from FACS-sorted $\gamma\delta$ T cells of neonatal cord blood (CB_donor1 and CB_donor2; $n = 2$) or adult peripheral blood (PB_donor1 and PB_donor2; $n = 2$). **(a,b)** Uniform Manifold Approximation and Projection (UMAP) was adopted for visualization of cells. Each point represents a single cell. **(a)** Individual cells of all four donors are colored by clusters (c1-c12) that were obtained from un-supervised clustering. **(b)** Individual cells are color-coded by the respective four donors. **(c)** Bar plot reveals fractions of absolute cell numbers that contributed from each donor to c1-c12. Total cell numbers of donors that were considered for analysis are indicated in the figure legend. **(d)** Dot plot shows the Top-10 upregulated differentially expressed genes (DEGs, rows) per cluster. Each column reflects a single cell, the color bars above represent c1-c12. For visualization, cell numbers were downsampled to max. 800 cells per cluster. Gene expression values were scaled to a Log2-fold change (logFC). DEGs were defined by being expressed on more than 10% of cells in at least one cluster ($\text{min.pct} \geq 10\%$), with absolute $\text{logFC} \geq 0.25$ and adjusted p-value ($\text{adj.p} \leq 0.01$) (bimod test provided by Seurat package).

Figure 2: Identification of gene co-expression modules from single-cell transcriptomes.

All 1,934 DEGs ($\text{logFC} \geq 0.25$, $\text{adj.p} \leq 0.01$, $\text{min.pct} \geq 10\%$, filtered ribosomal and mitochondrial genes) were subjected to unsupervised clustering based on the DEG's average expression per cluster (c1-c12) and seven gene co-expression modules (GMs) were identified. Annotation of the seven GMs was supported by gene ontology enrichment analysis (Suppl. Fig3). **(a)** For data visualization, the seven GMs were enwrapped by UMAP and color-coded. Selected key genes were further labeled on UMAP. Each dot represents a single cell that originates from the respective donors in Fig1a-b. **(b)** The heat map depicts the modula scores of each cell. Each column reflects a single cell, and the color bars above represent c1-c12. Cell numbers were downsampled to max. 600 cells per cluster.

Figure 3: Single-cell TCR repertoire shows the correlation between transcriptome and TCR.

Full-length $\gamma\delta$ TCR V(D)J segments were amplified from cDNA of scRNA-seq libraries for NGS analysis ($n = 4$). Via cellular barcodes single-cell $\gamma\delta$ TCR repertoires were mapped to matching single-cell transcriptomes of the respective four donors described in Fig1a. **(a)** UMAP of $\gamma\delta$ T cells visualizes the *TRD* repertoires' V gene usage (*TRDV*), while dots represent single-cells. Cells without productive *TRD* gene rearrangements are shown in grey. Contours outline cluster c1-c12 defined in Fig.1a. **(b)** Bar plots depict *TRDV* gene

compositions in each cluster per donor. Cells without a productive *TRDV* gene were excluded. The total number of mapped *TRDV* genes per cluster are provided on top of each bar. (c) The frequency of each identified paired $\gamma\delta$ TCR clone was calculated and visualized within the UMAP of $\gamma\delta$ T cells. Cells without a paired $\gamma\delta$ TCR were excluded. (d) Single-cell paired TCR repertoire diversities were estimated by Gini indices that range from 0 (completely polyclonal) to 1 (completely monoclonal). Bar plot shows Gini indices of each donor within the respective cluster (c1-c12) and total single-cell transcriptomes (all clusters per donor). Clusters containing less than 20 paired TCR were excluded from this analysis and defined as N.D (no data).

Figure 4: $V\delta 1^+$ $\gamma\delta$ CTLs mature into $PD-1^{hi}$ and $PD-1^{low}$ cells.

Cell clusters with high *TRDV1* gene expression (from Fig. 3a; c1, c7, c11-c12) were further analyzed. (a) Heat map describes selected DEGs among cluster c1, c7 and c11-c12. (b) Expression levels of DEGs that are key markers of exhausted (*IKF2*, *PDCD1*) or activated (*KLRD1*, *FCGR3A*) T cells in c11-c12 were plotted by violin scatter plots. (c) Gene set enrichment analysis (GSEA) was performed on gene expression profiles of c11 versus c12 referring to the Immunological gene sets of Molecular Signatures Database (MSigDb; www.gsea-msigdb.org). The enriched terms were from an analysis of $CD8^+$ T cells, GEO (GSE26495). Lower panel: Genes are arranged from being enriched in c11 to being enriched in c12 based on their logFC values. Middle panel: Each bar represents one hit gene from the enriched term on the gene list of c11 versus 12. Upper panel: Curves depict the running enrichment score of the selected enriched terms. (d) Paired TCR repertoire analysis of $V\delta 1^+$ T cells (c1, c7, c11 – c12) from PB donor_1 and PB donor_2. All paired TCR clones of the respective cluster were subjected to overlap analysis. Venn diagrams describe shared and non-shared TCR clones among clusters within PB donor_1 (left side) or PB donor_2 (right side). The size of each ellipse reflects the number of unique $V\delta 1^+$ clones in the respective cluster, further displayed in numbers.

Figure 5: $V\delta 2^+$ $\gamma\delta$ T cells are heterogeneous innate-like T cells.

Combined scTCR-seq and scRNA-seq analysis focusing on neonatal and adult $V\delta 2^+$ T cell clusters defined in Fig. 3. (a) UMAP of $V\delta 2^+$ $\gamma\delta$ T cells with delineation of respective clusters (from Fig. 1a) projects the publicity of *TRD* repertoires per cluster. Bulk *TRD* repertoires from independent study populations (ranging from infants to adults, n = 80) were compared to the four scTCR-seq libraries of this study and devoted to define a list of most public, common public, public, rare public and private $V\delta 2^+$ T cell clones by being present in the respective number of donors, indicated in brackets of the legend. (b) Visualization of *ZBTB16* (encoding PLZF) and *KLRB1* (encoding CD161) expression on $\gamma\delta$ T cells (adopted from Figure S2c). (c) Cell clusters dominated by *TRDV2* expression (from Fig. 3a, c4 – c6, c8 – c10) were further

investigated. Expression levels of selected DEGs were plotted by heat map. **(d)** Key DEGs of innate-like T cells (*KLRB1*), Th-17 (RORC, CCR6) and Th-1 (TBX21, FCGR3A, KLRD1) immunity are illustrated by violin scatter plots within c8 cells (cord blood and adult) and c10 cells (adult only).

Figure 6: $\gamma\delta$ T17 cells show distinct TCR repertoire profiles.

Paired TCR repertoires of $V\delta 2^+$ T clusters either being committed to a Th17-like (c8) or Th1-like (c6, c10, c9) response were investigated in adult PB_donor1 and PB_donor2. **(a)** Overlap analysis of paired TCR clones obtained from scTCR-seq was performed for PB_donor1 and PB_donor2. The amount of overlapping clones among clusters is visualized and described in venn diagrams, while the size of the ellipse reflects the number of unique TCR clones in each cluster. **(b)** The five most expanded paired TCR clones (Top-5) of each *TRDV2* dominated cluster (c8, c6, c10, c9) were selected for overlap analysis. The colored bands between two columns represent shared clones among clusters. The order of clusters was adjusted for visualizing. Results are described for PB_donor1 on the left side and for PB_donor2 on the right side. **(a,b)** Cluster c4 and c5 were excluded from analysis as few cells originated from adult donors.

Figure 7: $\gamma\delta$ T cells acquire effector phenotypes during the early fetal thymic development.

Single-cell transcriptome dataset of thymi from embryonic week 8 to 10 was acquired from Gene Expression Omnibus (Zeng et al., 2019; GEO: GSE133341). T cells (defined by *CD3D* expression in Fig.S6a) were subjected to this re-analysis. **(a)** UMAP illustrates unsupervised clustering results of fetal-thymus T cells. Names of clusters were determined by the expression of TCR genes (Figure S6b). **(b)** UMAP describes the origin in embryonic weeks of fetal-thymus T cells via respective color codes. **(c)** Aggregated expression values of GMs (defined in Fig. 2) were projected on the UMAP of fetal-thymus T cells. Cluster gdT_3 and gdT_4 are magnified in the black box for clearer visualization. **(d)** Representative key DEGs among fetal-thymus T cells are presented in heat map. Cells were downsampled to max. 200 cells per cluster.

Figure 8: Flow cytometry analysis delineates $V\delta 1^+$ and $V\delta 2^+$ naïve and effector $\gamma\delta$ T cell subsets.

$\gamma\delta$ T cells from five cord blood and six adult peripheral blood donors were subjected to multicolor FACS analysis. CB_donor4_2 was an independent replicate of CB_donor4 that serves as an internal control for batch effects, while PB_donor1 and PB_donor2 were identical to adult donors of scRNA-seq data. Neonatal and adult $\gamma\delta$ T cell were stained by pan- $\gamma\delta$ TCR, $V\delta 1$ and $V\delta 2$ antibodies and nine antibodies for key surface markers to

distinguish different phenotypes were included **(a)** FACS data of $\gamma\delta$ T cells was subject to unsupervised clustering that was based on those nine key surface markers and enwrapped by UMAP. Ten FACS clusters were identified. Each point represents one single-cell and is color-coded by FACS cluster. **(b)** Expression values of the nine lineage markers were scaled to a range between 0 and 1 and projected on UMAP. **(c,d)** Individual cells are color-coded by (c) corresponding TCR δ chains or (d) donor ID. Gating of V δ 1⁺ and V δ 2⁺ T cells is shown in **Supplementary Fig. 8b**. **(e)** Overview of the corresponding clusters from multicolor FACS analysis and scRNA-seq, phenotypes, sources, TCR δ usage and corresponding signature genes are summarized in the table. The nine lineage markers selected for FACS validation are underscored.

Figure 1

a

b

e

FACS Cluster	scRNA Cluster	Phenotype	TCRD	Main Source	Signature Genes
FACS1	c1	naive	Mixed	Cord blood	<i>CCR7</i> ; <i>IL7R</i> ; <i>TCF7</i> ; <i>SOX4</i>
FACS2	c3	naive	Mixed	Cord blood	<i>IL7R</i> ; <i>TCF7</i> ; <i>SOX4</i>
FACS3	c12	Active CTL	TRDV1	Adult	<i>FCGR3A</i> ; <i>GNLY</i> ; <i>BATF</i>
FACS4	c11	Exhausted CTL	TRDV1	Adult	<i>PDCD1</i> ; <i>TIGIT</i> ; <i>IKZF2</i>
FACS5	c6	Innate-naïve	TRDV2	Mixed	<i>IL7R</i> ; <i>CCR7</i> ; <i>GZMK</i>
FACS6	c8	Th-17 like	TRDV2	Mixed	<i>CCR6</i> ; <i>KLRB1</i> ; <i>DPP4</i>
FACS7	c10	Th-1 like (IL7R+)	TRDV2	Adult	<i>IL7R</i> ; <i>KLRD1</i>
FACS8	c10	Th-1 like (CD16+)	TRDV2	Adult	<i>KLRD1</i> ; <i>FCGR3A</i>
FACS9	c4; c5	Th-2 like	TRDV2	Cord blood	<i>CCR4</i> ; <i>CCR7</i>
FACS10	N.A.	unidentified	other	Mixed	<i>CCR6</i> ; <i>CCR7</i>
undetected	c2	Replicating naive	Mixed	Cord blood	<i>CD69</i> ; <i>CREM</i> ; <i>LMNA</i>
Undetected	c7	Acutely activated	TRDV1	Mixed	<i>CD69</i> ; <i>NR4A1</i> ; <i>TNFRSF4</i>
Undetected	c9	IFN-induced	TRDV2	Mixed	<i>BST2</i> ; <i>IRF7</i> ; <i>IFI6</i>