Supplementary Materials and Methods

Single-cell RNA-seq data preprocessing

First, we performed quality control to remove poor-quality cells and badly detected genes on CD4+ cells. In detail, we loaded into the scanpy environment all count data for the 7 patients and their relative TCR data, we matched them through the scirpy *pp.merge_with_ir* function, and then filtered out cells with a mitochondrial read rate >0.6 and expressed <1250 genes detected in less than 10 cells, except for patient 3 for which PB sample was missing, so we excluded it from our dataset. We preferred to preserve cells with such a high mitochondrial rate, since the sample contains populations with a high respiratory activity or with an exhausted phenotype. To handle the risk of doublets, we applied the scrublet algorithm [Wolock S.L. et al., 2019] which identified and marked a small number of doublets. Cell-specific biases were normalized by dividing the measured counts by the size factor obtained through the scran *computeSumFactors* method, which implements the deconvolution strategy for scaling normalization [Lun A.T. et al., 2016]. Next, all counts were log-transformed after addition of a pseudocount of 1. Later, we mitigated the batch effect through the matching mutual nearest neighbors (MNN) algorithm. In detail, we took the union of highly variable genes whose expression was common across all samples, thus resulting in 4000 genes that were used for the MNN batch correction [Haghverdi L. et al., 2018], performed through the python mnnpy package [https://github.com/chriscainx/mnnpy access on 10/12/2022]. Then, we used these genes to perform the dimension reduction on the MNN batch corrected data through principal component analysis (PCA) and the first 50 principal components (PCs) were used to construct a neighborhood graph of observations through the *pp.neighbors* function, which relies on the UMAP algorithm to estimate connectivity of data points [McInnes et al., 2018]. Next, we clustered data by *tl.louvain* function at multiple resolutions (0.5, 0.7, 1) and 0.7 revealed to better capture functional states, producing seven clusters of CD4+ cells. To annotate them, we defined their markers by running the tl.rank_gene_groups function and adopting the Wilcoxon rank-sum method. By looking at cluster

markers we observed presence of monocytes, defined by CD14 and ITGAM co-expression, so we removed them from dataset and recalculated the neighborhood graph on the latent space in order to cluster and annotate the resulting cleaned data. Next, we evaluated the composition of our dataset by looking for the expression of canonical markers from key immune cell types, states and pathways, already used by Julie et al. [11] and genesets adopted by Mashmeyer at al. [10], in detail NAIVE_VS_EFF_MEMORY_CD4_TCELL_DN (GSE11057) containing genes whose expression have been found to be downregulated in naive CD4+ T cells as opposed to effector memory T cells, the curated REACTOME_TCR_SIGNALING (M15381) and REACTOME_CELL_CYCLE (M543) gene sets, from the Reactome database, containing genes that are associated with TCR signaling and with the cell cycle, respectively, and TRM_CELLS_UP containing genes that were found to be upregulated in TRM cells (CD69, ITGAE, RUNX3, ITGA1, CXCR6, KCNK5, RGS1, DUSP6, PD-1) [Siracusa F. et al., 2019 and Kumar B.V. et al.,2017]. To show the presence of Th1- and Th17- like cells, we defined them looking for the expression of typical markers already used in previous works [Yost K.E. et al., 2019, McCluskey D. et al., 2022, Andreatta M. et al., 2021], and calculated their distribution density through the *tl.embedding_density* scanpy function. To define the PDCD1-TIGIT subtypes, we resambled the previous process but, first, we looked for cells expressing PDCD1 and TIGIT exclusively or in combination and we annotated them as “both expressed” if the expression of both genes was co-occurent over a defined treshold (0.5), “Neither expressed” if was under for both, “TIGIT not PDCD1” and “PDCD1 not TIGIT” if only one gene passed the threshold.

Next, we proceded with the TCR analysis through the scirpy module. After categorizing cells on the basis of the detection of productive antigen receptor chains (*tl.chain_qc* function), we selected cells with a single pair of productive chains for further analysis. T cell clonotypes were defined at the amino acid level (*pp.ir_dist*, *tl.define_clonotypes* and *tl.clonotype_network* functions), considering

both receptor chains. TCR diversity and TCR clonal size were estimated using *tl.alpha_diversity* and *tl.clonal_expansion* functions, respectively.

Finally, we used the tool Slingshot to infer trajectories. In detail, to better define the development trajectories, we first regressed-out the cell-cycle effect from our dataset [Vento-Tormo R. et al., 2018, Büttner M. et al., 2018] through the scanpy pp.regress_out function, passing as input the previously calculated S and G2M scores. Cell cycle analysis was performed by creating two lists of genes associated to the S and G2/M phases based on cell cycle genes previously defined, passed to the *tl.score_cell_cycle_genes* function to score S and G2/M phases.

Once the cell cycle was regressed-out, we recalculated the neighborhood graph on the latent space in order to cluster and annotate the resulting data which was then prepared for slingshot analysis (keeping only the 4000 highly variable genes) conducted through the *slingshot* function, using UMAP as visualization method and choosing Tcm as starting cluster, based on two aspects. First, the distribution of naïve-like cells compared to the memory- and effector-like ones, second, the clonal distribution. Tcm cluster revealed to be rich in naïve-like cells and poor in highly expanded clones. Finally, to evaluate the presence of shared clones along trajectories, we binned clones basing on clone-size intervals, keeping only clones with size > 5, we evaluated their expansion in each cluster and we plotted this information as a matrix showing the dataset clusters on an axis and the clone size bins on the other, in order to know the clone id, its size based on the belonging bin and its sharing level.

The same analysis was conducted on CD8+ cells. We collected and merged data for 6 patients (we excluded patient 3 to be consistent with CD4+ analyses) we filtered out cells based on previous thresholds, normalized, batch corrected through MNN and then visualized. Next we clustered, through the *tl.louvain* function, at multiple resolutions and we preferred 0.7 as for CD4+ cells. After cluster marker definition we looked for contaminant cell lines and we found again monocytes

(cluster 11), so we removed them along with two other cluster (clusters 9 and 11) containing cells with no TCR information. Using canonical gene-sets (NAIVE_VS_EFF_CD8_TCELL_DN, REACTOME_CELL_CYCLE, REACTOME_TCR_SIGNALING, TRM_CELLS_UP) we defined three effector and two central memory signature clusters, along with a Trm, an IFN stimulated and an interestingly cluster resembling some already described gene signature of NKT cells [15, 16] and of a particular subset of innate-like CD8+ T cells [17] that we called NKT-KLRB1. We also observed that at this cluster resolution we identified two clusters that showed a very similar transcriptome except for the tissue of origin. For this reason, we decided to merge these population in a unique cluster (CD8-TIGIT+).

Also in this case we defined the distribution density of PDCD1-TIGIT populations, and next proceeded with TCR and trajectory analysis, respectively, regressing-out cell-cycle effect also in this case. Finally, we generated a clone type heatmap as for CD4+ cells, in order to evaluate clone sharing among clusters.

BIBLIOGRAFY

Andreatta M, Corria-Osorio J, Müller S, Cubas R, Coukos G, Carmona SJ. Interpretation of T cell states from single-cell transcriptomics data using reference atlases. Nat Commun. 2021 May 20;12(1):2965. doi: 10.1038/s41467-021-23324-4. PMID: 34017005; PMCID: PMC8137700.

Büttner M, Miao Z, Wolf FA, Teichmann SA, Theis FJ. A test metric for assessing single-cell RNA-seq batch correction. Nat Methods. 2019 Jan;16(1):43-49. doi: 10.1038/s41592-018-0254-1. Epub 2018 Dec 20. PMID: 30573817.

Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat Biotechnol. 2018 Jun;36(5):421-427. doi: 10.1038/nbt.4091. Epub 2018 Apr 2. PMID: 29608177; PMCID: PMC6152897.

Lun AT, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Res. 2016 Aug 31;5:2122. doi: 10.12688/f1000research.9501.2. PMID: 27909575; PMCID: PMC5112579.

Kumar BV, Ma W, Miron M, Granot T, Guyer RS, Carpenter DJ, Senda T, Sun X, Ho SH, Lerner H, Friedman AL, Shen Y, Farber DL. Human Tissue-Resident Memory T Cells Are Defined by Core Transcriptional and Functional Signatures in Lymphoid and Mucosal Sites. Cell Rep. 2017 Sep 19;20(12):2921-2934. doi: 10.1016/j.celrep.2017.08.078. PMID: 28930685; PMCID: PMC5646692.

McCluskey D, Benzian-Olsson N, Mahil SK, Hassi NK, Wohnhaas CT; APRICOT and PLUM study team, Burden AD, Griffiths CEM, Ingram JR, Levell NJ, Parslew R, Pink AE, Reynolds NJ, Warren RB, Visvanathan S, Baum P, Barker JN, Smith CH, Capon F. Single-cell analysis implicates TH17-to-TH2 cell plasticity in the pathogenesis of palmoplantar pustulosis. J Allergy Clin Immunol. 2022 Oct;150(4):882-893. doi: 10.1016/j.jaci.2022.04.027. Epub 2022 May 12. PMID: 35568077.

McInnes et al., (2018). UMAP: Uniform Manifold Approximation and Projection. Journal of Open Source Software, 3(29), 861, https://doi.org/10.21105/joss.00861

Siracusa F, Durek P, McGrath MA, Sercan-Alp Ö, Rao A, Du W, Cendón C, Chang HD, Heinz GA, Mashreghi MF, Radbruch A, Dong J. CD69+ memory T lymphocytes of the bone marrow and spleen express the signature transcripts of tissue-resident memory T lymphocytes. Eur J Immunol. 2019

Jun;49(6):966-968. doi: 10.1002/eji.201847982. Epub 2019 Jan 30. PMID: 30673129; PMCID: PMC6563480.

Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park JE, Stephenson E, Polański K, Goncalves A, Gardner L, Holmqvist S, Henriksson J, Zou A, Sharkey AM, Millar B, Innes B, Wood L, Wilbrey-Clark A, Payne RP, Ivarsson MA, Lisgo S, Filby A, Rowitch DH, Bulmer JN, Wright GJ, Stubbington MJT, Haniffa M, Moffett A, Teichmann SA. Single-cell reconstruction of the early maternal-fetal interface in humans. Nature. 2018 Nov;563(7731):347-353. doi: 10.1038/s41586-018-0698-6. Epub 2018 Nov 14. PMID: 30429548; PMCID: PMC7612850.

Wolock SL, Lopez R, Klein AM. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst. 2019 Apr 24;8(4):281-291.e9. doi: 10.1016/j.cels.2018.11.005. Epub 2019 Apr 3. PMID: 30954476; PMCID: PMC6625319.

Yost KE, Satpathy AT, Wells DK, Qi Y, Wang C, Kageyama R, McNamara KL, Granja JM, Sarin KY, Brown RA, Gupta RK, Curtis C, Bucktrout SL, Davis MM, Chang ALS, Chang HY. Clonal replacement of tumor-specific T cells following PD-1 blockade. Nat Med. 2019 Aug;25(8):1251-1259. doi: 10.1038/s41591-019-0522-3. Epub 2019 Jul 29. PMID: 31359002; PMCID: PMC6689255.

Supplementary Table 1: Clinical features of JIA patients.

JIA Patients	Gender (M/F)	Median age in Year	JIA subtype (Oligoarticular/Polyarticular RF-neg)	Therapies
22	6/16	10,5 SD: 5,9	12/10	n=14: at the onset of disease n=2: anti-TNF monoclonal antibody n=6: MTX

RF: rheumatoid factor; MTX: methotrexate

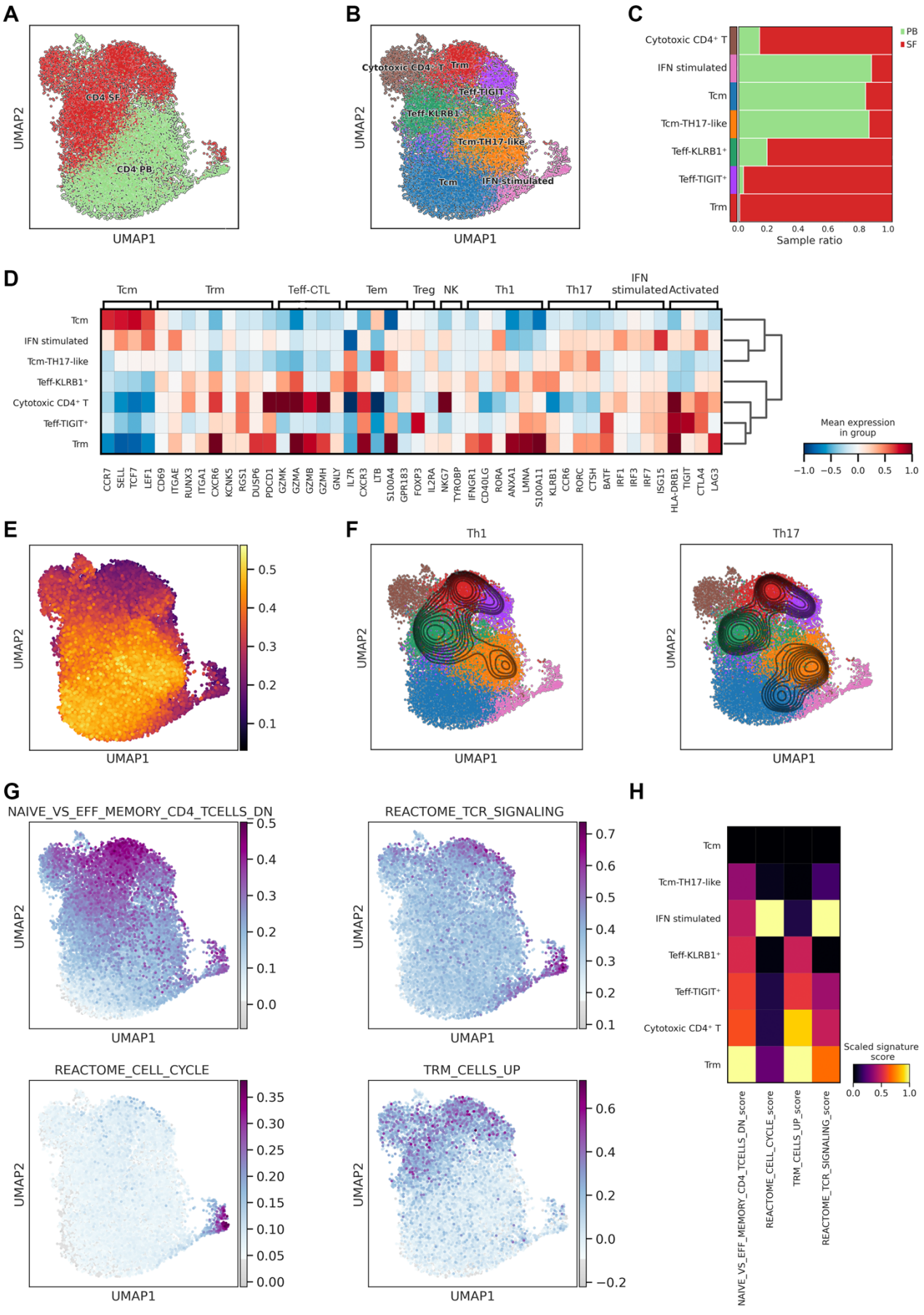
Supplementary Table 2. List of all fluorochrome mAbs used for flow cytometric immunophenotyping of circulating and SF T cells.

Antigen	**Fluorochrome**	**Clone**	**Company**
CD103	FITC	Ber ACT8	BDBioscience
CD69	PE	L78	BDBioscience
CD3	PerCP	SK7	BDBioscience
PD-1	PE-Cy7	EH12.2H7	BioLegend
CD8	Super Bright 600	SK1	eBioscience™
TIGIT	APC	MBSA43	eBioscience™
CD4	eFlour506	RPA-T4	eBioscience™

Supplementary Table 3. List of all fluorochrome mAbs used for flow cytometric evaluation of T cells cytokine production.

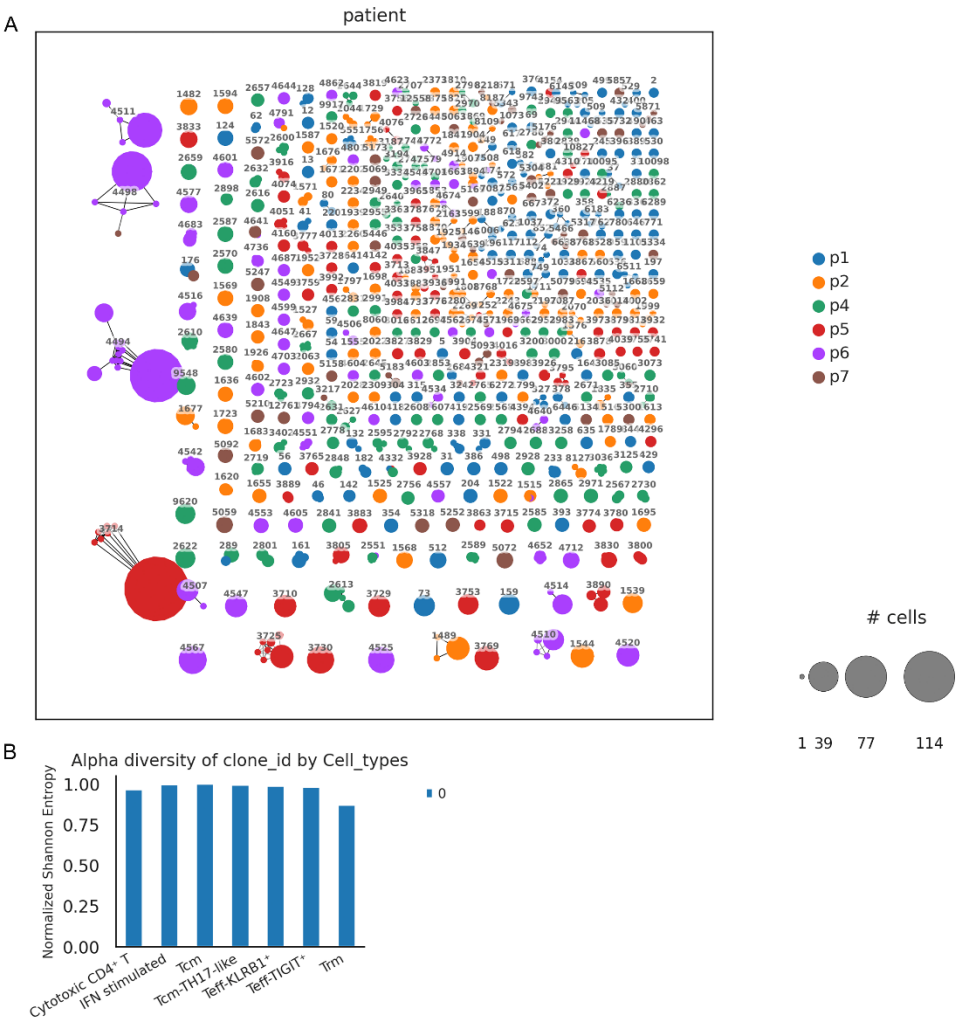
Antigen	**Fluorochrome**	**Clone**	**Company**
IFNγ	FITC	25723.11	BDBioscience
IL2	PE	5344.111	BDBioscience
TNFα	FITC	6401.1111	BDBioscience
GM-CSF	PE	BVD2-21C11	BDBioscience
IL17	FITC	eBio64DEC17	eBioscience™
IL10	PE	JES3-9D7	Milteniy Biotec
CD3	Super Bright 702	UCTH1	eBioscience™
CD8	Super Bright 600	SK1	eBioscience™
TIGIT	APC	MBSA43	eBioscience™

CD4	eFlour506	RPA-T4	eBioscience™
PD-1	Pacific Blue	EH12.2H7	BioLegend
L/D	Fixable Viability Stain 780		BDBioscience


A
CD4 SF
CD4 PB
UMAP2
UMAP1
B
Cytotoxic CD4⁺ T
Trm
Teff-TIGIT⁺
Teff-KLRB1⁺
Tcm-TH17-like
Tcm
IFN stimulated
C
PB
SF
Sample ratio
D
Tcm
Trm
Teff-CTL
Tem
Treg
NK
Th1
Th17
IFN stimulated
Activated
Mean expression in group
E
F
Th1
Th17
G
NAIVE_VS_EFF_MEMORY_CD4_TCELLS_DN
REACTOME_TCR_SIGNALING
REACTOME_CELL_CYCLE
TRM_CELLS_UP
H
Scaled signature score
NAIVE_VS_EFF_MEMORY_CD4_TCELLS_DN_score
REACTOME_CELL_CYCLE_score
TRM_CELLS_UP_score
REACTOME_TCR_SIGNALING_score

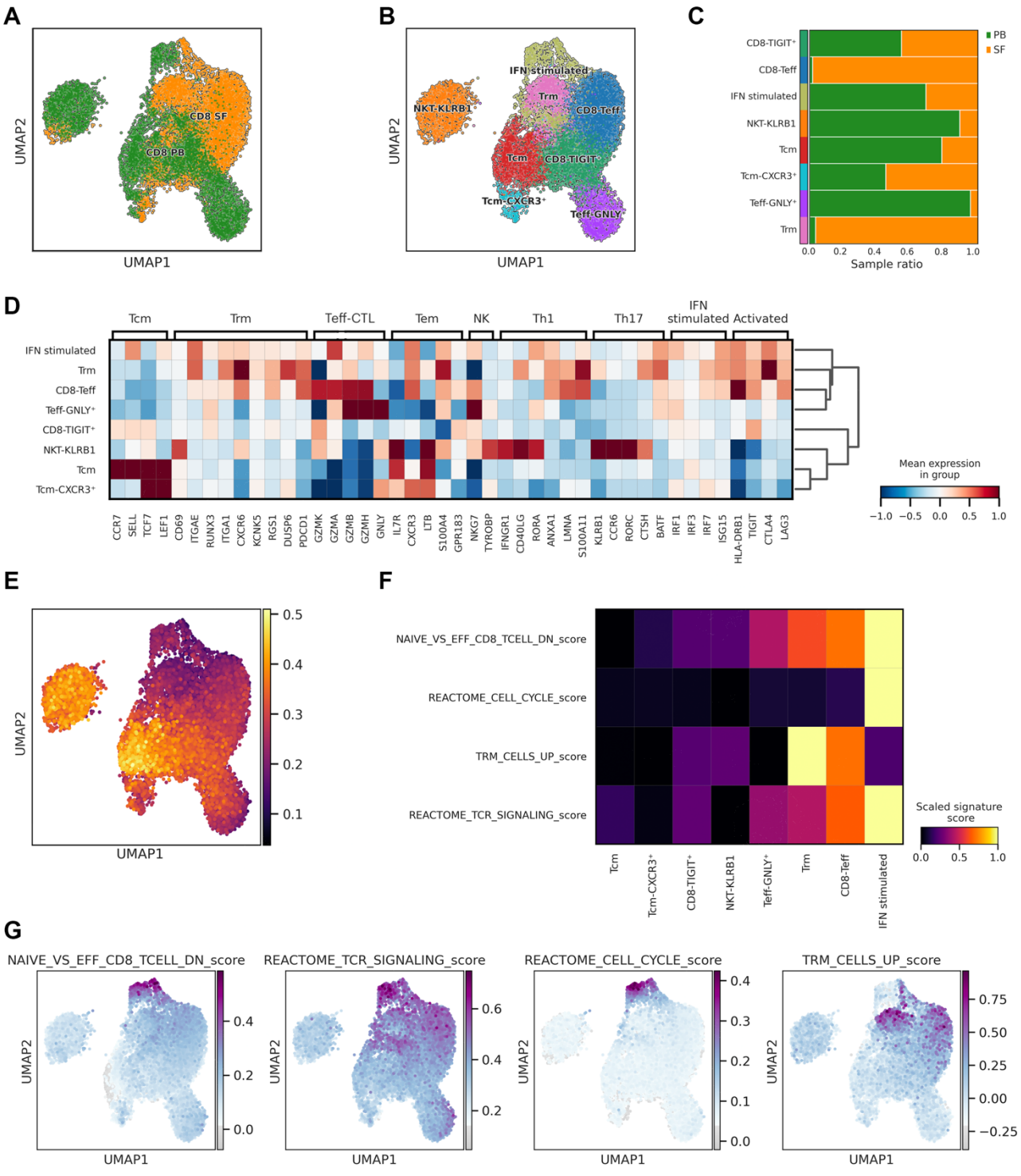
Supplementary Figure S1. scRNA-seq analysis of CD4⁺ T lymphocytes.

(A) UMAP representation of CD4+CD45RO+ T cells from PB and inflamed joints (SF) of 6 JIA patients colored by tissue origin. (B) UMAP projection of CD4+ cells grouped into 7 distinct Louvain clusters. (C) Barplot showing the PB or SF tissue composition of each cluster. (D) Heatmap showing the mean expression of functionally defined gene sets in each cluster. For each gene the color code illustrates the average level of expression of all cells in the cluster. (E) UMAP visualization of ribosomal genes mean expression. (F) UMAP representation of Th1 and Th17 localization expressed in terms of density. (G) Modules score of gene sets associated to previously described T cell functions projected on UMAP maps and (H) comparison between clusters illustrated as heatmap.



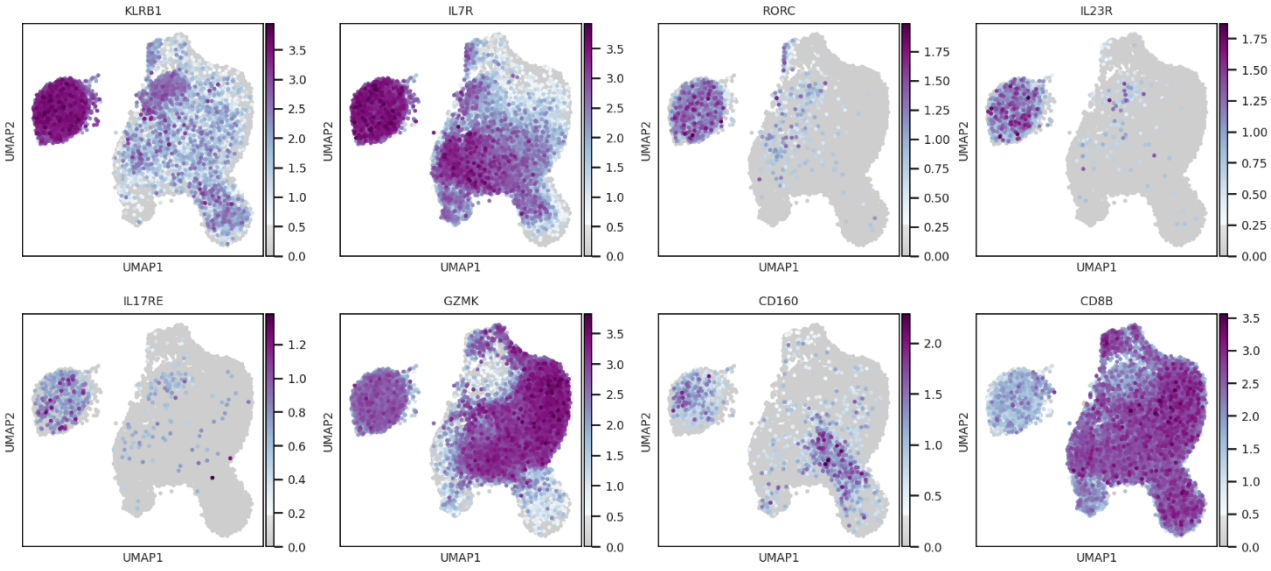
Supplementary Figure S2. CD4+ T cells TCRs diversity and clonal distribution among analyzed patients.

(A) Network clustering of CD4+ T cells clones based on TCR sequence similarities. For each clones, the size of the dot is proportional to the percentage of cells with the same TCR, and the color code illustrates from which patient the clone belongs. (B) Bar plot showing the Alpha diversity of clones in each cluster, bars correspond to the normalized Shannon Entropy index.



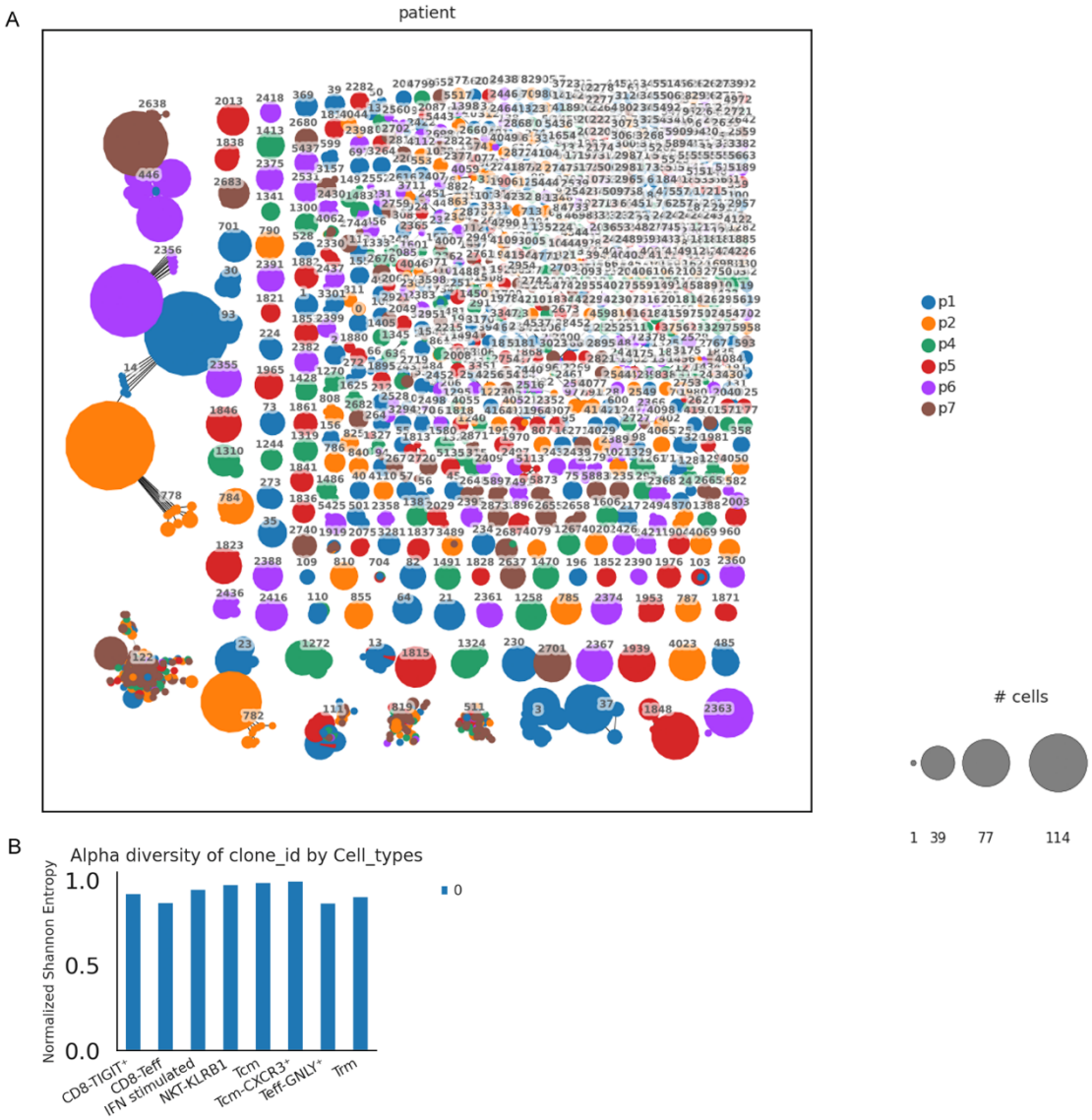
Supplementary Figure S3. scRNA-seq analysis of CD4⁺ T lymphocytes.

(A) CD8+CD45RO+ T cells from PB and inflamed joints (SF) of 7 JIA patients were sequenced, clustered and projected onto a UMAP representation of the cluster distributions among tissue samples. (B) Clustering separated them into 8 different subsets. (C) Bar plot showing the PB or SF tissue composition of each cluster. (D) Heatmap showing the mean expression of functionally defined gene sets in each cluster. For each gene the color code illustrates the average level of expression of all cells in the cluster. (E) UMAP visualization of ribosomal genes mean expression. (F) Clusters expression of modules score of gene set associated to previously described T cell functions illustrated as heatmap and (G) projected on UMAP maps.



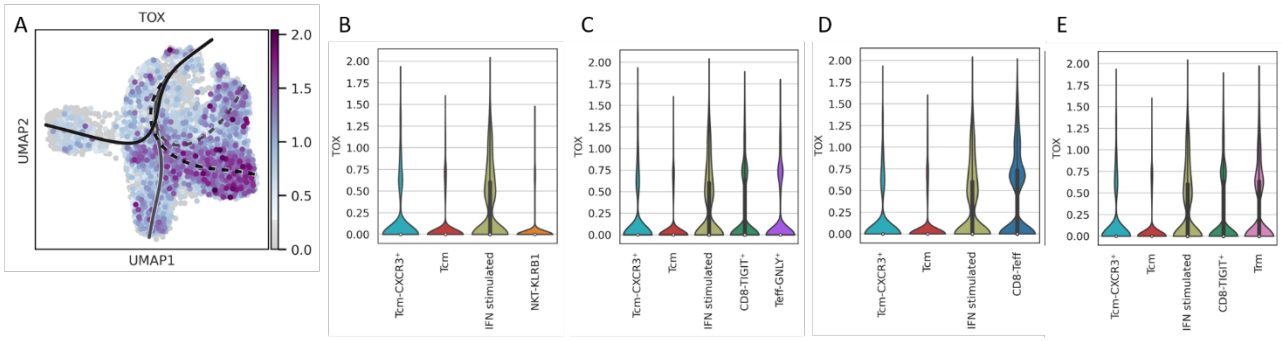
Supplementary Figure S4. Characteristic markers of CD8+ T cells of cluster 1.

UMAP visualization of gene expression characterizing cluster 1 cells of CD8+CD45RO+ cells. In particular, expression levels of KLRB1, IL7R, RORC, IL23R, IL17RE, GZMK, CD160 and CD8B are shown.



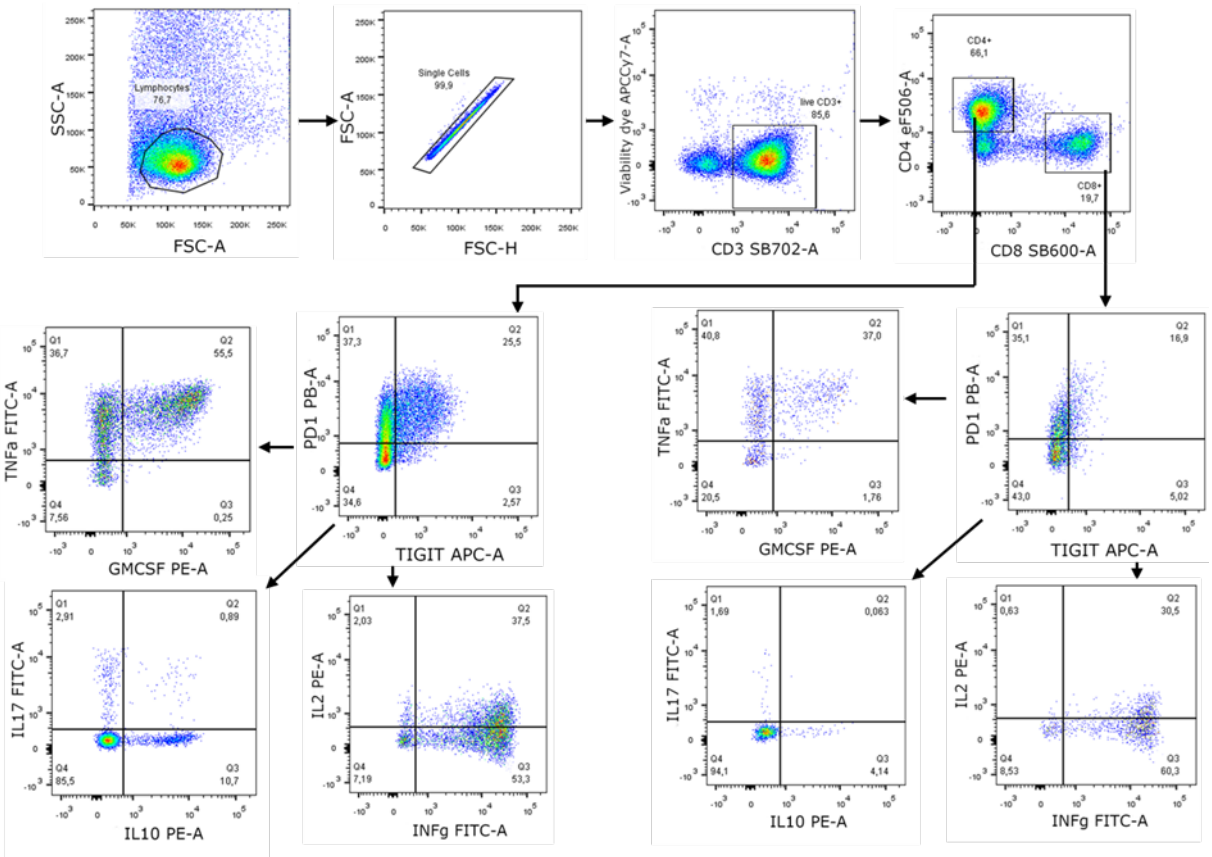
Supplementary Figure S5. CD8+ T cells TCRs diversity and clonal distribution among analyzed patients.

(A) Network clustering of CD8+ T cells clones based on TCR sequence similarities. For each clones, the size of the dot is proportional to the percentage of cells with the same TCR, and the color code illustrates from which patient the clone belongs. (B) Bar plot showing the Alpha diversity of clones in each cluster, bars correspond to the normalized Shannon Entropy index.



Supplementary Figure S6. Expression of *TOX* along CD45RO+CD8+ T cells pseudotime trajectory.

(A) UMAP visualization of TOX expression and slingshot trajectory in CD8+ T cells. (B-E) TOX gene expression reported for each cluster as Violin plot. Clusters are divided based on the trajectory lineage: Branch 1 (B), Branch 2 (C), Branch 3 (D), Branch 4 (E).



Supplementary Figure S7. Gating strategy for the flow cytometric identification of PD-1 and TIGIT expressing T cells and their differential cytokines production in PB and SF.

Lymphocytes were gated based on physical parameters (FSC-SSC), then doublets were removed using FSC-A and FSC-H parameters. Dead cells were excluded using viability stain and T cells were identified as CD3+. We then identified CD8+ and CD4+ T cells and on both these populations we evaluated TIGIT and PD-1 expression. On the four subpopulations identified based on TIGIT and PD-1 expression (TIGIT-PD-1-, TIGIT-PD-1+, TIGIT+PD-1+, TIGIT+PD-1-) we evaluated the production of IFN-γ, IL-2, TNF-α, GM-CSF, IL-17, IL-10. Cytokines are conjugated with the same fluorochrome FITC and PE, indeed three different staining were performed in order to evaluate all the listed cytokines.