

Single cell sequencing reveals expanded cytotoxic CD4⁺ T cells and two states of peripheral helper T cells in synovial fluid of ACPA+ RA patients

Alexandra Argyriou^{1*}, Marc H. Wadsworth II^{2*}, Adrian Lendvai¹, Stephen M. Christensen², Aase H. Hensvold^{1,3}, Christina Gerstner¹, Kellie Kravarik², Aaron Winkler², Vivianne Malmström¹, Karine Chemin¹

¹Division of Rheumatology, Center for Molecular Medicine, Department of Medicine, Solna, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden

²Inflammation & Immunology Research Unit, Pfizer Inc., Cambridge, MA 02139, United States

³Center for Rheumatology, Academic Specialist Center, Stockholm Health Region, Stockholm, Sweden

*Equal contribution

Abstract

Rheumatoid arthritis is an autoimmune disease affecting the synovial joints where different subsets of CD4⁺ T cells are suspected to play a pathogenic role. So far, our understanding of the contribution of cytotoxic CD4⁺ T cells is incomplete, particularly in the context of the recently described peripheral helper T-cell subset (T_{PH}). Here, using single cell sequencing and multi-parameter flow cytometry, we show that cytotoxic CD4⁺ T cells are enriched in synovial fluid of anti-citrullinated peptides antibody (ACPA)-positive RA patients. We identify two distinct T_{PH} states differentially characterized by the expression of *CXCL13* and *PRDM1*, respectively. Our data reveal that the adhesion G-Protein Coupled Receptor 56 (GPR56), a marker of circulating cytotoxic cells, delineates the synovial T_{PH} CD4⁺ T-cell subset. At the site of inflammation, GPR56⁺CD4⁺ T cells expressed the tissue-resident memory markers LAG-3, CXCR6 and CD69. Further, TCR clonality analysis revealed that most expanded clones in SF are contained within the cytotoxic and the CXCL13⁺ T_{PH} CD4⁺ T-cell populations. Finally, the detection of common TCRs between the two T_{PH} and cytotoxic CD4⁺ T-cell clusters suggest a shared differentiation. Our study provides comprehensive immunoprofiling of the heterogenous T-cell subsets at the site of inflammation in ACPA+ RA and suggests GPR56 as a therapeutic target to modulate T_{PH} cells and cytotoxic CD4⁺ T cell function.

Correspondence to:

Karine Chemin (karine.chemin@ki.se), Division of Rheumatology, Department of Medicine, KI, Stockholm

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Introduction

CD4⁺ T cells with cytotoxic functions (CD4⁺ CTL) have gained attention in recent years and accumulating evidence supports their importance in the defense against human viral infections such as CMV¹, EBV², dengue³, HIV^{4,5} and SARS-CoV-2⁶. They share common functional characteristics with NK and CD8⁺ T cells, including the expression of cytolytic proteins PRF1 (perforin-1), GZMB (granzyme B)⁷ and NKG7 (Natural Killer Cell Granule Protein 7)⁵. CD4⁺ CTL can originate from Th1 cells⁸ but can also differentiate from a distinct precursor expressing the class I-restricted T cell-associated molecule (CRTAM)⁹. CD4⁺ CTL differentiation processes are still unclear but antigen dose¹⁰, IL-2¹¹ and type 1 IFN¹² signaling are suspected to contribute. 4-1BB (CD137) co-receptor triggering induces CD4⁺ CTL differentiation through the induction of the transcription factor Eomesodermin (Eomes)¹³ and CD4⁺ CTL cytotoxic capacity is increased after IL-15 exposure¹⁴. Eomes cooperates with Runx3 to induce the transcription of *Prfl* in mouse CD8⁺ T cells¹⁵. Together with Blimp-1, the transcription factor Hobit sustains the transcription of *Gzmb*¹⁶. Hobit and Blimp are also involved in the maintenance of tissue resident memory T cells (T_{RM})¹⁷. Both Eomes^{18,19} and Hobit²⁰ are expressed in circulating human cytotoxic CTL, including CD4⁺ CTL, suggesting that they convey similar functions in a human setting. GPR56 is an adhesion G-protein coupled receptor encoded by the *ADGRG1* gene expressed on human circulating NK, CD8⁺ and CD4⁺ CTL²¹. Many of the molecular regulators of CD4⁺ CTL have been identified in mouse models or in circulating human blood. However, CD4⁺ CTL have not been investigated in great detail at the site of inflammation in human inflammatory disorders, especially in the context of T_{RM}. In this environment, it is still unclear from which precursors they differentiate from.

Rheumatoid arthritis (RA), which affects around 1% of the global population, is an autoimmune disease characterized by articular bone erosion leading to physical disability, pain and decreased quality of life. Bone damage is particularly exacerbated in the subset of RA patients who present with antibodies against citrullinated proteins (ACPA)^{22,23}. Although the use of biologics has revolutionized the treatment of RA, 30-40% of patients fail to respond to treatment stressing the need for novel therapeutic targets²⁴. The role of CD4⁺ T cells in RA pathogenesis is evidenced by the HLA-DRB1 association with ACPA+ RA²⁵. Expanded patient-specific T-cell clones have been identified in the synovium and synovial fluid (SF) of RA patients using bulk TCR β repertoire analysis²⁶. Over the years, several CD4⁺ T cell subsets, including Th1 and Th17 cells have been described in the synovial joints of RA patients²⁷. Recently, PD-1^{high} HLA-Class II⁺ peripheral helper T cells (T_{PH}) were identified in the synovial fluid and tissue of seropositive RA patients^{28,29} where they are proposed to facilitate B-cell recruitment and activation through CXCL13 and IL-21 production; however, there is still no satisfying marker that defines the T_{PH} subset. It is also unclear how this population is generated in the synovial joint. We have recently shown that CD4⁺ T cells with cytotoxic features are observed in SF of RA patients carrying the *PTPN22* 1858T risk allele variant³⁰. This finding, as well as previous studies reporting expanded CD4⁺ CD28^{null} cytotoxic T cells in the blood of RA patients^{31,32}, suggests that CD4⁺ CTL might contribute to RA. However, a transcriptomic and clonality analysis of CD4⁺ CTL, in the context of previously described CD4⁺ T cell subsets, is still missing. The identification of T-cell effector functions is deemed important since it can lead to new therapeutic strategies.

Here, we performed single cell sequencing in combination with 5' TCR $\alpha\beta$ sequencing on peripheral blood (PB) and synovial fluid (SF) from ACPA+ RA patients to investigate CD4⁺ CTL. We demonstrate that clonal expansions are prominent amongst CD4⁺ CTL in SF of ACPA+ RA. We also describe two T_{PH} CD4⁺ T-cell clusters differentially characterized by the expression of *PRDM1* and *CXCL13*, suggestive of two different stages of T_{PH} differentiation. Finally, we find that T_{PH} cells express the receptor GPR56 and the T_{RM} markers LAG-3, CXCR6 and CD69 implicating that these T cells are maintained in the ACPA+ RA synovial joint. Our data provide a comprehensive immunoprofiling map of pathogenic CD4⁺ T cells at the site of inflammation in ACPA+ RA.

Results

CD4⁺ T cells expressing cytotoxic effector molecules are enriched in synovial fluid of ACPA+ RA patients.

SFMC from ACPA- (n=9) and ACPA+ (n=12) RA patients were screened for the expression of cytotoxic effector molecules and transcription factors in both CD4⁺ and CD8⁺ T cells by flow cytometry (Fig. 1, supplementary Fig. 1). ACPA+ SFMC presented with a significantly increased frequency of GZMB⁺ PRF1⁺ (p=0.0072), Hobit⁺ (p=0.0018), NKG7^{high} (p=0.0036) and GPR56⁺ (p=0.0007) CD4⁺ cells (Fig. 1a-b). No significant difference was observed for the expression of GZMA and Eomes. Amongst CD8⁺ T cells, Hobit expression was increased in ACPA+ SFMC (p=0.0170) whereas a similar trend was observed for the expression of GZMA, GZMB and NKG7 without reaching statistical significance (supplementary Fig. 2). No significant difference was observed for the expression of PRF1 and the frequency of GPR56 in

CD8⁺ T cells in ACPA⁺ versus ACPA⁻ RA. The frequency of GZMB⁺ PRF1⁺ CD4⁺ T cells in SF positively correlated with CCP (cyclic citrullinated peptides) antibody titers ($r=0.5701$, $p=0.0070$) (Fig. 1c). In ACPA⁺ synovial fluid, GPR56 identified a distinct CD4⁺ T cell population with a frequency that also correlated with CCP antibody titers ($r=0.7871$, $p<0.0001$) (Fig. 1d).

Single cell sequencing identifies a cluster of cytotoxic CD4⁺ T cells and two states of T_{PH} CD4⁺ T cells in ACPA⁺ SFMC.

To further characterize cytotoxic CD4⁺ T cells in SF of ACPA⁺ RA, we performed 10X single cell sequencing in combination with 5' TCR $\alpha\beta$ sequencing on purified CD4⁺ T cells from paired SFMC and PBMC from seven ACPA⁺ RA patients (Fig. 2a, supplementary table 1). After quality control (supplementary Fig. 3a), we obtained transcriptomic data from 66,360 CD4⁺ T cells in SFMC and 69,373 CD4⁺ T cells in PBMC. Unsupervised analysis of the transcriptomic data from all CD4⁺ T cells ($n=135,733$) from PB and SF using the Seurat package³³ generated a total of 16 clusters annotated based on known and high scoring genes (Fig. 2b-c, supplementary Fig. 3 and 4). All 16 clusters were detected in both PB and SF from all patients although with different frequencies (Fig. 2d, supplementary table 2). A Treg cluster (cluster 15) was annotated based on the expression of *FOXP3* and *CTLA-4*³⁴. Resting naïve or central memory (CM) CD4⁺ T cells express *SELL*, *CCR7* and *IL7R*³⁵ (cluster 10). Two proliferating T-cell subsets expressing *MKI67* and *HELLS* were identified (clusters 13 and 14). Based on the expression of *CXCL13*, *MAF*, *TOX*, *PRDM1*, *TIGIT* and *PDCD-1*^{28,36,37}, two T_{PH} cell clusters were subclassified as T_{PH} *CXCL13*⁺ (cluster 3) and T_{PH} *PRDM1*⁺ (cluster 12). Cluster 4 defines cytotoxic CD4⁺ T cells expressing cytotoxicity related genes including *NKG7*, *GZMH* and *PRF1*. This population was present in SF of

all ACPA+ RA patients with a frequency ranging from 1.9 to 55% (Fig. 2d, supplementary Fig. 3b, supplementary table 2). In general, the frequency of T_{PH} *PRDM1*⁺, T_{PH} *CXCL13*⁺, Tregs and and cytotoxic CD4⁺ T cells was increased in SF as compared to PB (Fig. 2e, supplementary table 2). Hence, single cell transcriptomic data identify a subset of cytotoxic CD4⁺ T cells and show that the subset of T_{PH} cells is composed of two distinct clusters in ACPA+ RA patients.

GPR56 delineates peripheral helper T cells in SF of ACPA+ RA patients

Since GPR56 expression was upregulated on 26% of CD4⁺ T cells in SF of ACPA+ RA patients (Fig. 1a-b), we further investigated *ADGRG1* (encoding GPR56) in the single cell dataset. In PB, *ADGRG1* expression was expressed in T_{PH} cells (cluster 3 and 12), Humanin CD4⁺ T cells (cluster 7) and cytotoxic CD4⁺ T cells (cluster 4), the latter being characterized by the expression of *GZMA*, *GZMB*, *NKG7*, *PRF-1* and *Hobit* (supplementary Fig. 5a-b). At the protein level, flow cytometry experiments validated that, in circulating CD4⁺ T cells, GPR56 correlated with the expression of cytotoxic effector molecules PRF1, GZMB, GZMA, NKG7 and the transcription factors Eomes and Hobit (Supplementary Fig. 5c-d) confirming a previous report²¹. Still, a smaller subset of GPR56⁺ non-cytotoxic CD4⁺ T cells could also be detected in blood (data not shown). Interestingly, in ACPA+ RA SF, *ADGRG1* expression was detected in 3 subsets: T_{PH} *PRDM1* (cluster 12), T_{PH} *CXCL13*⁺ (cluster 3) and cytotoxic CD4⁺ T cells (cluster 4) with the highest expression in cluster 12 (Fig. 3a, supplementary Fig. 6). We further examined the expression of known T_{PH} associated markers^{28,36,37} and observed that *ADGRG1* expression overlapped with the expression of *PD-1*, *CXCL13*, *HLA-DRB1*, *MAF*, *ICOS*, *PDRM1* and *TOX* (Fig 3a). The two T_{PH} clusters were characterized by different levels of *ADGRG1*, *PD-1*, *CXCL13*, *HLA-DRB1*, *MAF* and *PDRM1*. Flow cytometry experiments confirmed that, in SF of ACPA+ RA patients,

GPR56 expression defines the subset of T_{PH} cells (Fig. 3b-c) as illustrated by the increased expression of PD-1 ($p < 0.0001$) and MHC-II ($p < 0.0001$) in GPR56⁺ CD4⁺ T cells when compared to GPR56⁻ CD4⁺ T cells (Fig. 3d). GPR56 was not detected on cytotoxic CD4⁺ T cells in SF (Supplementary Fig. 6c-d) as opposed to *ADGRG1* (Fig 3a, 6a-b), highlighting the difference between transcriptomic and protein expression. These data indicate that, while GPR56 is a marker for cytotoxic CD4⁺ T cells in PB, it also delineates the subsets of PD-1^{high} MHC-Class-II⁺ T_{PH} cells in ACPA+ RA SF.

GPR56⁺ CD4⁺ T cells display distinct tissue-resident memory receptors

Besides being an inhibitory receptor, PD-1 is also included in the signature associated to T-cell tissue residency³⁸ and implicated in follicular helper T cells maintenance in germinal centers³⁹. We therefore assessed the expression of known tissue resident memory (T_{RM}) T-cell markers: *LAG3*, *ITGAI* (CD49a), *CD69*, *CXCR6* but also *CX3CR1* which is downregulated on T_{RM} T cells³⁸. In SF, we found that *LAG-3* was mainly expressed on T_{PH} cells (cluster 3 and 12) and proliferating T cells (clusters 13 and 14) (Fig. 4a). A similar tendency was also observed for *CXCR6*. These genes were not expressed on cytotoxic CD4⁺ T cells, consistent with their recirculation in the blood. Since we previously observed that GPR56 delineates the subset of T_{PH} cells, we assessed the expression of tissue-resident memory markers in the context of GPR56 by flow cytometry (Supplementary Fig.7). *LAG-3* ($p < 0.0001$) and *CXCR6* ($p = 0.0052$) frequencies as well as *CD69* ($p = 0.0002$) expression was enriched on synovial GPR56⁺ CD4⁺ T cells (Fig. 4b-c). In contrast, *CD49* frequency was higher on GPR56-negative CD4⁺ T cells ($p < 0.0001$). In general, *LAG-3* frequency was also significantly increased on CD4⁺ T cells in ACPA+ SF as compared to ACPA- SF ($p = 0.0005$) (Fig. 4d) and correlated with CCP titers ($r = 0.8088$, $p < 0.0001$) (Fig. 4e). In CD8⁺ T cells, *CD49a* was the only marker for which an increased frequency was observed in ACPA+ SF

($p=0.0343$) (supplementary Fig. 8). These data indicate that, in ACPA+ SF, GPR56 expression correlates with the concomitant expression of a distinct set of tissue-resident memory markers on CD4⁺ T cells: LAG3, CXCR6 and CD69.

Clonally expanded cytotoxic and CXCL13⁺ T_{PH} CD4⁺ T cells are identified in synovial fluid of ACPA+ RA patients

Next, we examined the expanded TCR clones in relation to the aforementioned CD4⁺ T-cell clusters. A clone was defined by at least 2 cells sharing the same CDR3 amino acid sequence, either from TCR α , TCR β or the combination of both. The 10 most expanded clones were primarily identified in the cytotoxic CD4⁺ T-cell subset and to a lesser extent in the Humanin⁺ CD4⁺ T-cell subset in ACPA+ RA SF (Fig. 5a, right panel and supplementary Table 3 and 4). Some of the expanded clones were shared between PB and SF (Fig. 5a, left panel). These 10 expanded T-cell clones were originating from 4 RA patients (Supplementary Figure 9) of which three identical clones were identified between patient 5 and 6. The analysis of synovial fluid T-cell clones composed of at least 4 cells in each patient revealed that they mainly originate from cytotoxic CD4⁺ T cells, CXCL13⁺ T_{PH} and to a lesser extent PRDM1⁺ T_{PH} and Humanin CD4⁺ T cells (Fig. 5b, supplementary Table 3 and 4). Further, we examined overlapping CDR3 sequences between SF CD4⁺ T-cell subsets that would be indicative of a possible shared differentiation pathway (Fig. 5c). In SF, the CXCL13⁺ T_{PH} cluster shared TCR sequences with the PRDM1⁺ T_{PH}, Humanin⁺ CD4⁺ and proliferating CD4⁺ T-cell clusters. Common CDR3s were also identified between cytotoxic CD4⁺ T cells and CXCL13⁺ T_{PH} cells. Of note, in SF, Tregs shared only few CDR3 sequences with other CD4⁺ T-cell subsets, suggesting that they could be thymus-derived Tregs⁴⁰ and not induced at the site of inflammation from conventional CD4⁺ T cells. We then further

focused our TCR analysis on the clusters of interest (cytotoxic CD4⁺ T cells, PRDM1⁺ T_{PH}, CXCL13⁺ T_{PH} and proliferating CD4⁺ T cells 1 and 2) in PB and SF (Fig. 5d). We observed that some CDR3 sequences were shared between cytotoxic CD4⁺ T cells in PB and SF. Similarly, CXCL13⁺ T_{PH} cells had common CDR3s in PB and SF. Overall, these data show that, in ACPA⁺ synovial fluid, expanded clones are mainly found within the cytotoxic and the CXCL13⁺ T_{PH} CD4⁺ T-cell subsets and that shared CDR3 sequences are identified between clusters in particular between the two T_{PH} clusters and between CXCL13⁺ T_{PH} and cytotoxic CD4⁺ T cells.

Discussion

Cytotoxic CD4⁺ CD28^{null} T cells are expanded in the blood of patients with RA³¹ but their presence and clonality at the site of inflammation in ACPA⁺ RA was still not deeply characterized. Leveraging single cell technology, we identified the presence of cytotoxic CD4⁺ T cells in ACPA⁺ RA SF that are clonally expanded in synovial fluid and, for some patients, in their corresponding peripheral blood. We further define the T_{PH} CD4⁺ T-cell subset at the single cell level showing that it is composed of two independent clusters with different *CXCL13* and *PRDM1* (Blimp-1) expression levels. So far, the T_{PH} CD4⁺ T-cell subset was only defined by high expression of PD-1 and MHC-Class-II and more accurate markers were still lacking. We show that GPR56 clearly delineates the T_{PH} subset and associates with the expression of tissue resident memory markers LAG-3, CXCR6 and CD69. Together, these data provide a refined characterization of cytotoxic and T_{PH} CD4⁺ T cells, two pathogenic CD4⁺ T-cell subsets in ACPA⁺ RA (Fig. 6).

In synovial tissue of ACPA⁺ RA patients, T_{PH} CD4⁺ cells produce CXCL13^{28,36,41}, localize with B cells and have the ability to induce *in vitro* plasma cell differentiation

through IL-21 production²⁸. T_{PH} cells are defined by the high expression of PD-1 and MHC-Class II, the cytokines CXCL13, IL-21 and IL-10, and the transcription factors Blimp-1, TOX, MAF and SOX4^{28,36,37}. T_{PH} CD4⁺ T-cell differentiation is induced by TGFβ in IL-2 neutralizing conditions³⁷. Single cell sequencing identified two distinct clusters of T_{PH} cells: PRDM1⁺ T_{PH} cells and CXCL13⁺ T_{PH} cells. We postulate that these two clusters correspond to two states of differentiation: PRDM1⁺ T_{PH} CD4⁺ T cells could be the precursors of the more differentiated cytokine producing CXCL13⁺ T_{PH} CD4⁺ T cells. Indeed, Blimp-1 encoded by *PRDM1* negatively regulates the expression of CXCL13³⁷. In line with this hypothesis, common TCR sequences observed between PRDM1⁺ T_{PH} cells and CXCL13⁺ T_{PH} cells clusters suggest of a common origin between these two subsets.

The impact of CD4⁺ CTL in the synovial joint is less well characterized. Expansions of CD4⁺ CTL are frequently observed in chronic viral infections and CD4⁺ CTL are capable of directly killing EBV-infected B cells in a MHC-class II dependent manner². Whether CD4⁺ CTL are protective or pathogenic in viral infections is still a matter of debate and their role may vary depending on the type of infections⁵. Recently, single cell sequencing data has also identified infiltrating CD4⁺ CTL in liver cancer⁴² and in bladder cancer where they provided anti-tumor killing in the context of MHC-Class II⁴³. Interestingly, a CD4⁺ CTL signature in bladder tumors predicted clinical response to anti-PD-L1⁴³. Furthermore, expansions of cytotoxic CD28^{null} CD4⁺ T cells have been repeatedly reported in the blood of patients with rheumatic diseases such as RA³¹, myositis⁴⁴ and vasculitis⁴⁵ as well as in cardiovascular diseases^{46,47}. PRF1⁺ CD4⁺ T cells were also identified by immunofluorescence in synovial tissue of RA patients⁴⁸. Our data shows that cytotoxic CD4⁺ T cells are enriched in SF of ACPA⁺ RA patients and correlates with CCP titers showing that cytotoxic CD4⁺ T cells are mainly

implicated in the subset of autoantibody-positive RA. Purified perforin-1⁴⁹ as well as supernatants from IL-15 stimulated synovial CD8⁺ T cells⁵⁰ induces histone citrullination in neutrophils suggesting that cytotoxic effector molecules and citrullination are tightly linked in RA. In our dataset, cytotoxic CD8⁺ T cells also tended to be more cytotoxic in ACPA+ RA and could contribute to a perforin-dependent citrullination mechanism. The presence of CD4⁺ CTL only in SF of ACPA+ RA suggest that they have a pathogenic function, but further studies are warranted to understand the functional consequences of their interactions with antigen presenting cells in the synovial joint.

The differentiation processes that lead to the differentiation of cytotoxic CD4⁺ T cells are not completely elucidated. IL-2, IL-15 and type I IFNs, which have been implicated in the differentiation of cytotoxic CD4⁺ T cells, are also expressed in the synovial joint possibly contributing to their development. We show that the transcription factor Hobit, which has been shown to regulate *Gzmb* expression in mouse¹⁶, was expressed in GZMB/PRF-1-expressing CD4⁺ T cells in both PB and SF suggesting that Hobit also drive CD4⁺ CTL differentiation in ACPA+ RA. We have previously shown that PRF-1-expressing Eomes⁺ CD4⁺ T cells are increased in SF of *PTPN22* risk carriers³⁰. In dengue viral infections, an increased frequency of CD4⁺ CTL was observed in HLA-DRB1*0401+ individuals⁵¹. Hence, RA-associated genetic polymorphisms might contribute to increase the differentiation of CD4⁺ CTL through mechanisms that remain to be investigated.

Expanded TCR clones have been identified in synovial tissues and fluid from RA patients using bulk TCR β sequencing²⁶. In blood, expanded CD4⁺ CD8^{null} cells present with a bias in their TCRV β usage^{52,53}. In our dataset, the most expanded T-cell clones

identified in SF of ACPA+ RA presented with cytotoxic or CXCL13⁺ T_{PH} signatures. The differentiation towards the cytotoxic CD4⁺ T cell lineage has been proposed to be linked to repeated antigen stimulation as illustrated in the case of chronic viral infections such as CMV¹⁰. Expanded cytotoxic CD4⁺ T cell clones in SF might therefore be the final stage of differentiation of autoreactive CD4⁺ T cells which have been exposed to several rounds of antigen stimulation. Several studies have shown that CMV reactivity is found in the subset of CD4⁺ CD8^{null} cell subset but so far, their autoreactive properties have not been investigated in the context of citrulline immunity. In our data set, some TCR sequences were shared between T_{PH} cells and cytotoxic CD4⁺ T cells and we can speculate that T_{PH} cells, after repeated antigenic stimulation led to cytotoxic CD4⁺ T cell expanded clones.

GPR56 is an adhesion G-protein coupled receptor encoded by the *ADGRG1* gene implicated in numerous migration/adhesion processes including neuron migration⁵⁴ and tumor growth inhibition⁵⁵. In NK cells, GPR56 is an inhibitory receptor binding to the tetraspanin CD81⁵⁶. We confirmed a previous report showing that GPR56 is also a marker of circulating cytotoxic CD4⁺ cells²¹. At the site of inflammation, however, we discovered that GPR56 was highly expressed on T_{PH} cells in ACPA+ RA patients and could be used as a new marker to identify T_{PH} CD4⁺ T cells. In SF CD4⁺ T cells, GPR56 expression correlated with the expression of PD-1, LAG-3, CXCR6 and CD69 expressed on tissue-resident memory T cells³⁸. Similarly, in multiple sclerosis, CD8⁺ T cells from active lesions exhibit a tissue resident memory phenotype (CD69, CD103, PD-1, CD49a) associated with an intermediate expression of GPR56 without granzyme B expression⁵⁷. In mouse, Blimp-1 and Hobit mediates the transcriptional program associated with T-cell tissue residency¹⁷. In ACPA+ SF, Blimp-1 was expressed on T_{PH} CD4⁺ cells presenting with tissue resident memory markers whereas Hobit was only

expressed on cytotoxic CD4⁺ T cells suggesting that, Blimp-1 could drive tissue residency of T_{PH} CD4⁺ T cells in the synovial joint. In human NK cells, *Hobit* triggers *ADGRG1* expression⁵⁶ but, in SF, *Hobit* was not expressed in *ADGRG1*⁺ CD4⁺ T cells suggesting that a different transcription factor could regulate GPR56. GPR56 might be implicated in the migration and/or maintenance of T_{PH} CD4⁺ T cells in the synovial joints where its ligand still remains to be determined.

In summary, we identify clonally expanded cytotoxic CD4⁺ T cells and two clusters of T_{PH} CD4⁺ T cells in ACPA⁺ SF. We show that GPR56 is a marker that should be used to identify T_{PH} CD4⁺ cells. Future studies aiming at understanding the function of GPR56 in T_{PH} CD4⁺ and cytotoxic CD4⁺ T cells will contribute to evaluate the feasibility of targeting this receptor in ACPA⁺ RA.

Methods

Synovial fluid from ACPA⁻ (n=9) and ACPA⁺ (n=12) RA patients were collected at the Rheumatology Clinic of Karolinska University Hospital, Stockholm (Supplementary Table 1). In n=8 ACPA⁻ and n=8 ACPA⁺ RA patients, paired PBMCs were also available. All patients had established disease, based on the 1987 criteria from the American College of Rheumatology. This study was approved by the research ethics committee of Karolinska University Hospital and all patients signed informed consent according to the Declaration of Helsinki. PBMC and SFMC were isolated by Ficoll separation (GE Healthcare) and cryopreserved in fetal bovine serum (FBS) with 10% DMSO in liquid nitrogen.

Flow Cytometry

1.10⁶ PBMC and 1.10⁶ SFMC were labeled with LIVE DEAD Fixable NEAR IR Dead Cell Stain Kit (Invitrogen) and further stained with fluorescently labeled antibodies (Supplementary table 5). ACPA- and ACPA+ SF and PB were always run in the same experiment. Depending on the number of recovered cells, 2 to 4 flow cytometry panels were run (Supplementary table 6). For intracellular stainings, cells were fixed and permeabilized using the FOXP3 permeabilization buffer kit (ebioscience) according to the manufacturer's protocol. Samples were run on the BD LSR Fortessa and flow cytometry data was analyzed with the FlowJo software (version X 10.07r2).

Statistical Analysis

Flow cytometry results were analyzed with the Prism 7 software (GraphPad, San Diego, CA, USA) using two-tailed Mann-Whitney U test. Statistical tests performed are indicated in the figure legends.

Single Cell Processing

CD4⁺ T cells were enriched from cryopreserved PBMC and SFMC samples using the EasySep human CD4 Positive Selection Kit II (STEMCELL Technologies, Catalog# 17852). Following CD4⁺ T-cell isolation, cells were counted, hashed with TotalSeq-C0251 (ref: 394661) and TotalSeq-C0253 (ref: 394665) anti-human antibodies (Biolegend) and pooled per patient (i.e., pool 1 = PB and SF from patient 1, pool 2 = PB and SF from patient 2, etc.). The pooled cells were then counted again and immediately loaded onto the chromium chip G using the standard protocol for the Chromium single cell 5' kit v1.1 (10x Genomics, Inc). Following Gel Bead-in Emulsion (GEM) generation, samples were processed using the standard manufacturer's protocol. Once sequencing libraries passed standard quality control

metrics, the libraries were sequenced on Illumina NextSeq500/550 high output 75-cycle v2.5 kits with the following read structure: read1: 26, read2: 58, index 1: 8. Libraries were sequenced to obtain a read depth greater than 20,000 reads/cell for the gene-expression (GEX) libraries and greater than 5,000 reads/cell for the V(D)J-enriched T Cell libraries.

Alignment and Preprocessing

Following sequencing, the pooled GEX libraries were demultiplexed using bc2fastq (v2.17) with modified parameters (`--minimum-trimmed-read-length=10 --mask-short-adapter-reads=10 --ignore-missing-positions --ignore-missing-controls`) to generate FASTQ files for each patient. FASTQ files were then uploaded to Terra (www.app.terra.bio) where the raw sequencing data were mapped and quantified using STAR within the 10X Genomics Inc software package Cell Ranger on Cumulus (<https://cumulus.readthedocs.io/en/latest/cellranger.html>, snapshot 15, Cell Ranger 5.0.1, default parameters) with GRCh38-2020-A (<https://cumulus.readthedocs.io/en/latest/cellranger.html>) to generate UMI-collapsed, gene-by-cell expression matrices. Aligned matrices were first filtered to remove low-quality barcodes, keeping only those with greater than 200 UMIs, less than 4,000 UMIs, and less than 20% mitochondrial reads. Cell doublets were then computationally removed using DoubletFinder⁵⁸ with a calculated doublet rate of 14% and default parameters.

Unsupervised Analysis and Cell Type Annotation

Using the Seurat⁵⁹ (v4.0.0) package, filtered gene-by-cell matrices were merged and then processed using a standard unsupervised workflow (i.e., normalization, scaling, dimensionality reduction, batch correction, cell clustering, and differential gene expression analysis). First, the merged matrix was normalized and log-transformed

(scaling factor = 10,000). The top 3000 highly variable genes were identified and used for scaling the data. During scaling, two sources of unwanted variation, cumulative UMI capture and percent mitochondrial reads, were regressed out. Dimensionality reduction was then performed with principal component analysis (PCA) over the top 3000 variable genes and, using the JackStraw function in Seurat, we selected 53 statistically significant principal components (PCs, $p < 0.01$). Using these significant PCs, we corrected for compartment and patient-specific batch effects with the harmony⁶⁰ package, and then proceeded with constructing the nearest neighbor graph and Uniform Manifold Approximation and Projection (UMAP) plots. Cells were clustered using Louvain clustering and the package clustree⁶¹ was used to generate a clustering tree and identify which resolution achieved stability. Differential gene expression analysis was then performed using the FindAllMarkers Seurat function with test.use set to 'MAST'⁶². We also used Nebulosa⁶³ to visualize the joint density approximations of known T-cell phenotypic markers. We attempted to deconvolute the multiplexed PBMC and SFMC cell populations using the cell hashing antibodies, however, due to sub-optimal tagging we were unable to separate the cell populations. To address this issue, we generated a compartment-specific reference using PBMC and SFMC samples from two more volunteers and then reference-mapped our mixed samples to assign compartment membership⁵⁹. Only barcodes with a predicted id score greater than 0.5 were kept and used for downstream analyses. Finally, non-T cells were filtered out by removing barcodes with zero gene expression for CD3 genes. Taking both the results of the differential expression analysis and joint density approximations of known phenotypic markers, we identified 16 unique T cell subsets: 1. Activated CD4⁺ T Cells (CD38⁺ HLA-DRA⁺); 2. CD4⁺ T Cells (IL7R⁺); 3. CXCL13⁺ Tph (CXCL13⁺ CD3D⁺); 4. Cytotoxic CD4⁺ T Cells (GNLY⁺ GMZB⁺ PRF1⁺ GZMA⁺

NKG7⁺); 5. Effector CD4⁺ T Cells (CCR6⁺ KLRB1⁺ CD27⁺); 6. EGR1⁺ CD4⁺ T Cells (EGR1⁺); 7. Humanin⁺ CD4⁺ T Cells (MTRNR2L12⁺ MTRNR2L8⁺); 8. IL3RA⁺ CD4⁺ T Cells (IL3RA⁺); 9. LAMP3⁺ CD4⁺ T Cells (LAMP3⁺); 10. Naïve and Central Memory (CM) T Cells (IL7R⁺, CCR7⁺, SELL⁺); 11. Mitochondrial High CD4⁺ T cells (MT-CO1⁺MT-ATP8⁺MT-ND4⁺); 12. PRDM1⁺ Tph (PRDM1⁺ ADGRG1⁺); 13. Proliferating CD4⁺ T Cells 1 (STMN1⁺); 14. Proliferating CD4⁺ T Cells 2 (KI67⁺); 15. Tregs (TIGIT⁺ FOXP3⁺ IL2RA⁺); 16. TTN⁺ CD4⁺ T cells (TTN⁺).

TCR Mapping

Following sequencing, the V(D)J pooled libraries were demultiplexed using bc2fastq (v2.17.1.14) to generate FASTQ files for each patient, uploaded to terra, and mapped and quantified using the 10X Genomics Inc software package CellRanger on Cumulus (<https://cumulus.readthedocs.io/en/latest/cellranger.html>, snapshot 15, Cellranger 5.0.1, default parameters) to generate consensus annotation files for each patient. Using the scRepertoire⁶⁴ package, patient-specific consensus annotation files were consolidated into a list of TCR sequencing results and then integrated with the Seurat object using the combineExpression function. Clonotypes were called based on the CDR3 gene and clonotype expansion were assigned based on the absolute frequency. Finally, the interconnectivity between specific cell types, both within and between tissue compartments, were visualized using chord diagrams.

LEGENDS

Figure 1. Cytotoxic CD4⁺ T cell frequency in synovial fluid of ACPA- and ACPA+ RA patients. (a) Representative flow cytometry dot plot staining of effector molecules, receptors and transcription factors associated with cytotoxic functions in CD4⁺ T cells from synovial fluid (SF) from ACPA- (upper panel) and ACPA+ (lower panel) RA patients. quantified in (b),

(ACPA-, n=7-9) (ACPA+, n=10-12). Line represents median, two-tailed Mann-Whitney U test, ns: not significant. **(c)** Correlation between the frequency of GZMB⁺ PRF1⁺ in CD4⁺ T cells in SF and the level of serum anti-CCP (cyclic citrullinated peptides) antibodies levels, n=21, Spearman two-tailed test. **(d)** Correlation between the frequency of GPR56 in CD4⁺ T cells in SF and the level of anti-CCP (cyclic citrullinated peptides) antibodies levels, n=21, Spearman two-tailed test. **(a-d)** Data are from a pool of nine independent experiments where a circle is a single replicate. Blue dots indicate ACPA- RA SF and red dots indicate ACPA+ RA SF.

Figure 2. Single cell RNA sequencing of CD4⁺ T cells from SF and PB of ACPA+ RA patients. **a)** Technical workflow including CD4⁺ T-cell enrichment and 10X 5' single cell sequencing coupled to TCR $\alpha\beta$ sequencing on 7 paired PB and SF from ACPA+ RA patients. **b)** UMAP plots of 135,733 cells colored by cell type cluster split by combined compartments (left), SF-only (middle) and PB-only (right panel). **c)** Heatmap showing gene expression values for a curated list of CD4⁺ T-cell subset-defining genes. **d)** Stacked barplots displaying cell type composition for each ACPA+ RA patient. **e)** Stacked barplots depicting comparative frequency of the CD4⁺ T-cell subsets found in PB and SF.

Figure 3. GPR56 expression on peripheral helper CD4⁺ T cells in ACPA+ RA SF.

a) Nebulosa density (left panel) and gene expression (right panel) single cell plots of *ADGRG1* and T_{PH} CD4⁺ T-cell markers in ACPA+ RA SF. **b)** Representative flow cytometry dot plot of PD-1 and MHC-Class II staining in combination with GPR56 in ACPA- and ACPA+ RA. **(c)** Representative T-SNE flow cytometry plots showing CD4, GPR56, PD-1 and MHC-Class II expression in one ACPA- and one ACPA+ RA SF (gated on CD3⁺ T cells). **(d)** Mean fluorescence intensity (MFI) of PD-1 (n=12) and MHC-II (n=11) in GPR56-negative and positive CD4⁺ T cells in ACPA+ RA SF. Line represents median, two-tailed Mann-Whitney U test. Data are from a pool of nine independent experiments where a circle is a single replicate. White dots indicate GPR56⁻ CD4⁺ T cells and black dots indicate GPR56⁺ CD4⁺ T cells.

Figure 4. Tissue resident memory markers (T_{RM}) on $CD4^+$ T cells in ACPA+ RA SF.

a) 2-D dot plots showing the expression of selected genes in the different $CD4^+$ T-cell clusters in ACPA+ SF (circle size indicates cell frequency, color intensity indicates average expression). **b)** Representative flow cytometry dot plots showing the expression of LAG-3, CXCR6, CD69, CD49a and CX3CR1 on $GPR56^+$ and $GPR56^-$ $CD4^+$ T cells in ACPA+ RA, quantified in **c)** ($n=10$, LAG3, CXCR6, CD69, CD49a) ($n=8$, CX3CR1). White dots indicate $GPR56^-$ $CD4^+$ T cells and blackdots indicate $GPR56^+$ $CD4^+$ T cells. **(d)** Expression of LAG-3, CXCR6, CD69, CD49a ($n=8$ ACPA-, $n=10$ ACPA+) and CX3CR1 ($n=7$ ACPA-, $n=8$ ACPA+) on $CD4^+$ T cells in SF. Blue dots indicate ACPA- RA SF and red dots indicate ACPA+ RA SF. **c-d)** Line represents median, two-tailed Mann-Whitney U test. Data are from a pool of eight independent experiments where a circle is a single replicate. **(e)** Correlation between the frequency of LAG-3 in $CD4^+$ T cells in RA SF and the level of anti-CCP (cyclic citrullinated peptides) antibodies levels ($n=18$), Spearman two-tailed test.

Figure 5. Clonally expanded $CD4^+$ T cells in ACPA+ RA. **a)** UMAP plots showing the top 10 expanded clones in ACPA+ RA PB and SF (Each dot displays a cell). Each individual clone sharing the same CDR3 (from $TCR\alpha$, $TCR\beta$ or both) amino acid sequences is highlighted by a different color. **b)** Stacked barplots displaying the phenotype of the SF expanded clones ($n>4$ cells) in each ACPA+ RA patient, quadrant represents individual clone. **c)** Chord diagram showing the interconnectivity between $CD4^+$ T cells sharing the same CDR3 (from $TCR\alpha$, $TCR\beta$ or both) amino acid sequences within $CD4^+$ T-cell clusters in SF of 7 ACPA+ RA patients. **d)** Chord diagram showing the interconnectivity between $CD4^+$ T cells sharing the same CDR3 (from $TCR\alpha$, $TCR\beta$ or both) amino acid sequences within selected $CD4^+$ T-cell clusters: $PDRM1^+$ T_{PH} , $CXCL13^+$ T_{PH} , cytotoxic $CD4^+$ T cells, proliferating $CD4^+$ T cells in paired PB and SF.

Figure 6. Graphical abstract showing a model for the differentiation of two states of peripheral helper T cells and cytotoxic CD4⁺ T cells in ACPA+ RA SF.

Acknowledgments

We thank the patients who donated samples and the medical staff at the Rheumatology Clinic of Karolinska University Hospital. Julia Boström, Gloria Rostvall, and Susana Hernandez Machado are acknowledged for organizing the sampling, storage, and administration of biomaterial. We thank Leonid Padyukov and Barbro Larsson for HLA-DR genotyping of the samples and Lena Israelsson for technical assistance with ACPA measurement. This study is supported by grants from Dr. Margaretha Nilssons, the Nanna Svartz, the Börje Dahlin and the Ulla and Gustaf af Ugglas foundations as well as the Swedish association against rheumatism. This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 777537 (RTCure). The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA.

Author contributions

A.A., A.L. and C.G. performed and analysed the flow cytometry experiments. M.H.W performed the 10X experiments and led the single cell analysis. S.M.C. performed preliminary 10X experiments and single cell data analysis. A.H.H. diagnosed RA patients and collected clinical information. K.K. and A.W. contributed to the design of the single cell experiments. V.M. contributed to the design of the study and the RA sample collection. A.A., M.H.W and K.C. generated the figures. K.C designed the study, oversaw the analyses and wrote the paper. All authors read, edited and approved the paper.

Competing interest

A.A, A.L., C.G., A.H.H, V.M. and K.C. have no competing interest to declare. M.H.W., S.M.C., K.K. and A.W. are employees of Pfizer, Inc, Cambridge, MA 02139, United States.

References

- 1 Casazza, J. P. *et al.* Acquisition of direct antiviral effector functions by CMV-specific CD4⁺ T lymphocytes with cellular maturation. *J Exp Med* **203**, 2865-2877, doi:10.1084/jem.20052246 (2006).
- 2 Landais, E. *et al.* Direct killing of Epstein-Barr virus (EBV)-infected B cells by CD4 T cells directed against the EBV lytic protein BHRF1. *Blood* **103**, 1408-1416, doi:10.1182/blood-2003-03-0930 (2004).
- 3 Kurane, I., Meager, A. & Ennis, F. A. Dengue virus-specific human T cell clones. Serotype crossreactive proliferation, interferon gamma production, and cytotoxic activity. *J Exp Med* **170**, 763-775, doi:10.1084/jem.170.3.763 (1989).
- 4 Appay, V. *et al.* Characterization of CD4(+) CTLs ex vivo. *J Immunol* **168**, 5954-5958, doi:10.4049/jimmunol.168.11.5954 (2002).
- 5 Juno, J. A. *et al.* Cytotoxic CD4 T Cells-Friend or Foe during Viral Infection? *Front Immunol* **8**, 19, doi:10.3389/fimmu.2017.00019 (2017).
- 6 Meckiff, B. J. *et al.* Imbalance of Regulatory and Cytotoxic SARS-CoV-2-Reactive CD4(+) T Cells in COVID-19. *Cell* **183**, 1340-1353 e1316, doi:10.1016/j.cell.2020.10.001 (2020).
- 7 Swain, S. L., McKinstry, K. K. & Strutt, T. M. Expanding roles for CD4(+) T cells in immunity to viruses. *Nat Rev Immunol* **12**, 136-148, doi:10.1038/nri3152 (2012).
- 8 Qui, H. Z. *et al.* CD134 plus CD137 dual costimulation induces Eomesodermin in CD4 T cells to program cytotoxic Th1 differentiation. *J Immunol* **187**, 3555-3564, doi:10.4049/jimmunol.1101244 (2011).
- 9 Takeuchi, A. *et al.* CRTAM determines the CD4⁺ cytotoxic T lymphocyte lineage. *J Exp Med* **213**, 123-138, doi:10.1084/jem.20150519 (2016).
- 10 Derhovanessian, E. *et al.* Infection with cytomegalovirus but not herpes simplex virus induces the accumulation of late-differentiated CD4⁺ and CD8⁺ T-cells in humans. *J Gen Virol* **92**, 2746-2756, doi:10.1099/vir.0.036004-0 (2011).
- 11 Brown, D. M., Kamperschroer, C., Dilzer, A. M., Roberts, D. M. & Swain, S. L. IL-2 and antigen dose differentially regulate perforin- and FasL-mediated cytolytic activity in antigen specific CD4⁺ T cells. *Cell Immunol* **257**, 69-79, doi:10.1016/j.cellimm.2009.03.002 (2009).
- 12 Hua, L. *et al.* Cytokine-dependent induction of CD4⁺ T cells with cytotoxic potential during influenza virus infection. *J Virol* **87**, 11884-11893, doi:10.1128/JVI.01461-13 (2013).
- 13 Curran, M. A. *et al.* Systemic 4-1BB activation induces a novel T cell phenotype driven by high expression of Eomesodermin. *J Exp Med* **210**, 743-755, doi:10.1084/jem.20121190 (2013).
- 14 Mucida, D. *et al.* Transcriptional reprogramming of mature CD4(+) helper T cells generates distinct MHC class II-restricted cytotoxic T lymphocytes. *Nat Immunol* **14**, 281-289, doi:10.1038/ni.2523 (2013).
- 15 Cruz-Guilloty, F. *et al.* Runx3 and T-box proteins cooperate to establish the transcriptional program of effector CTLs. *J Exp Med* **206**, 51-59, doi:10.1084/jem.20081242 (2009).
- 16 Kragten, N. A. M. *et al.* Blimp-1 induces and Hobit maintains the cytotoxic mediator granzyme B in CD8 T cells. *Eur J Immunol* **48**, 1644-1662, doi:10.1002/eji.201847771 (2018).

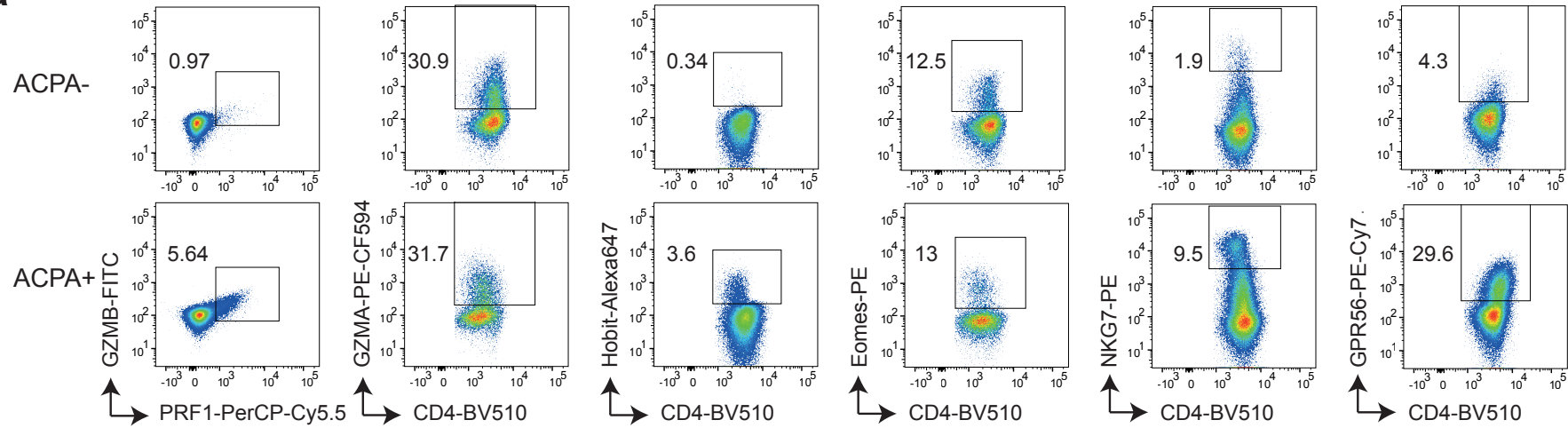
- 17 Mackay, L. K. *et al.* Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science* **352**, 459-463, doi:10.1126/science.aad2035 (2016).
- 18 Knox, J. J., Cosma, G. L., Betts, M. R. & McLane, L. M. Corrigendum: Characterization of T-bet and Eomes in Peripheral Human Immune Cells. *Front Immunol* **7**, 337, doi:10.3389/fimmu.2016.00337 (2016).
- 19 Raveney, B. J. *et al.* Eomesodermin-expressing T-helper cells are essential for chronic neuroinflammation. *Nat Commun* **6**, 8437, doi:10.1038/ncomms9437 (2015).
- 20 Oja, A. E. *et al.* The Transcription Factor Hobit Identifies Human Cytotoxic CD4(+) T Cells. *Front Immunol* **8**, 325, doi:10.3389/fimmu.2017.00325 (2017).
- 21 Peng, Y. M. *et al.* Specific expression of GPR56 by human cytotoxic lymphocytes. *J Leukoc Biol* **90**, 735-740, doi:10.1189/jlb.0211092 (2011).
- 22 Kroot, E. J. *et al.* The prognostic value of anti-cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis. *Arthritis Rheum* **43**, 1831-1835, doi:10.1002/1529-0131(200008)43:8<1831::AID-ANR19>3.0.CO;2-6 (2000).
- 23 Malmstrom, V., Catrina, A. I. & Klareskog, L. The immunopathogenesis of seropositive rheumatoid arthritis: from triggering to targeting. *Nat Rev Immunol* **17**, 60-75, doi:10.1038/nri.2016.124 (2017).
- 24 Smolen, J. S. *et al.* Rheumatoid arthritis. *Nat Rev Dis Primers* **4**, 18001, doi:10.1038/nrdp.2018.1 (2018).
- 25 Padyukov, L. *et al.* A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis. *Ann Rheum Dis* **70**, 259-265, doi:10.1136/ard.2009.126821 (2011).
- 26 Musters, A. *et al.* In Rheumatoid Arthritis, Synovitis at Different Inflammatory Sites Is Dominated by Shared but Patient-Specific T Cell Clones. *J Immunol* **201**, 417-422, doi:10.4049/jimmunol.1800421 (2018).
- 27 Chemin, K., Gerstner, C. & Malmstrom, V. Effector Functions of CD4+ T Cells at the Site of Local Autoimmune Inflammation-Lessons From Rheumatoid Arthritis. *Front Immunol* **10**, 353, doi:10.3389/fimmu.2019.00353 (2019).
- 28 Rao, D. A. *et al.* Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. *Nature* **542**, 110-114, doi:10.1038/nature20810 (2017).
- 29 Zhang, F. *et al.* Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat Immunol* **20**, 928-942, doi:10.1038/s41590-019-0378-1 (2019).
- 30 Chemin, K. *et al.* EOMES-positive CD4(+) T cells are increased in PTPN22 (1858T) risk allele carriers. *Eur J Immunol* **48**, 655-669, doi:10.1002/eji.201747296 (2018).
- 31 Schmidt, D., Goronzy, J. J. & Weyand, C. M. CD4+ CD7- CD28- T cells are expanded in rheumatoid arthritis and are characterized by autoreactivity. *J Clin Invest* **97**, 2027-2037, doi:10.1172/JCI118638 (1996).
- 32 Fasth, A. E., Cao, D., van Vollenhoven, R., Trollmo, C. & Malmstrom, V. CD28nullCD4+ T cells--characterization of an effector memory T-cell population in patients with rheumatoid arthritis. *Scand J Immunol* **60**, 199-208, doi:10.1111/j.0300-9475.2004.01464.x (2004).

- 33 Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* **33**, 495-502, doi:10.1038/nbt.3192 (2015).
- 34 Zemmour, D. *et al.* Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nat Immunol* **19**, 291-301, doi:10.1038/s41590-018-0051-0 (2018).
- 35 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, doi:10.1038/s41467-019-12464-3 (2019).
- 36 Kobayashi, S. *et al.* A distinct human CD4⁺ T cell subset that secretes CXCL13 in rheumatoid synovium. *Arthritis Rheum* **65**, 3063-3072, doi:10.1002/art.38173 (2013).
- 37 Yoshitomi, H. *et al.* Human Sox4 facilitates the development of CXCL13-producing helper T cells in inflammatory environments. *Nat Commun* **9**, 3762, doi:10.1038/s41467-018-06187-0 (2018).
- 38 Kumar, B. V. *et al.* Human Tissue-Resident Memory T Cells Are Defined by Core Transcriptional and Functional Signatures in Lymphoid and Mucosal Sites. *Cell Rep* **20**, 2921-2934, doi:10.1016/j.celrep.2017.08.078 (2017).
- 39 Shi, J. *et al.* PD-1 Controls Follicular T Helper Cell Positioning and Function. *Immunity* **49**, 264-274 e264, doi:10.1016/j.immuni.2018.06.012 (2018).
- 40 Bluestone, J. A. & Abbas, A. K. Natural versus adaptive regulatory T cells. *Nat Rev Immunol* **3**, 253-257, doi:10.1038/nri1032 (2003).
- 41 Manzo, A. *et al.* Mature antigen-experienced T helper cells synthesize and secrete the B cell chemoattractant CXCL13 in the inflammatory environment of the rheumatoid joint. *Arthritis Rheum* **58**, 3377-3387, doi:10.1002/art.23966 (2008).
- 42 Zheng, C. *et al.* Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* **169**, 1342-1356 e1316, doi:10.1016/j.cell.2017.05.035 (2017).
- 43 Oh, D. Y. *et al.* Intratumoral CD4(+) T Cells Mediate Anti-tumor Cytotoxicity in Human Bladder Cancer. *Cell* **181**, 1612-1625 e1613, doi:10.1016/j.cell.2020.05.017 (2020).
- 44 Fasth, A. E. *et al.* T cell infiltrates in the muscles of patients with dermatomyositis and polymyositis are dominated by CD28null T cells. *J Immunol* **183**, 4792-4799, doi:10.4049/jimmunol.0803688 (2009).
- 45 Moosig, F., Csernok, E., Wang, G. & Gross, W. L. Costimulatory molecules in Wegener's granulomatosis (WG): lack of expression of CD28 and preferential up-regulation of its ligands B7-1 (CD80) and B7-2 (CD86) on T cells. *Clin Exp Immunol* **114**, 113-118, doi:10.1046/j.1365-2249.1998.00695.x (1998).
- 46 Sato, K. *et al.* TRAIL-expressing T cells induce apoptosis of vascular smooth muscle cells in the atherosclerotic plaque. *J Exp Med* **203**, 239-250, doi:10.1084/jem.20051062 (2006).
- 47 Dumitriu, I. E. The life (and death) of CD4⁺ CD28(null) T cells in inflammatory diseases. *Immunology* **146**, 185-193, doi:10.1111/imm.12506 (2015).
- 48 Griffiths, G. M., Alpert, S., Lambert, E., McGuire, J. & Weissman, I. L. Perforin and granzyme A expression identifying cytolytic lymphocytes in rheumatoid arthritis. *Proc Natl Acad Sci U S A* **89**, 549-553, doi:10.1073/pnas.89.2.549 (1992).

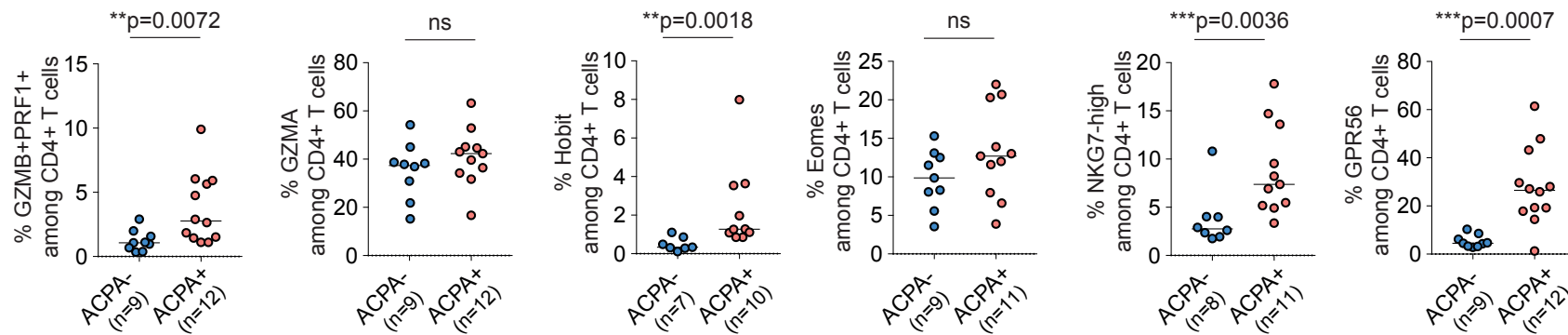
- 49 Romero, V. *et al.* Immune-mediated pore-forming pathways induce cellular hypercitrullination and generate citrullinated autoantigens in rheumatoid arthritis. *Sci Transl Med* **5**, 209ra150, doi:10.1126/scitranslmed.3006869 (2013).
- 50 Jung, J. *et al.* Synovial fluid CD69(+)CD8(+) T cells with tissue-resident phenotype mediate perforin-dependent citrullination in rheumatoid arthritis. *Clin Transl Immunology* **9**, e1140, doi:10.1002/cti2.1140 (2020).
- 51 Weiskopf, D. *et al.* Dengue virus infection elicits highly polarized CX3CR1+ cytotoxic CD4+ T cells associated with protective immunity. *Proc Natl Acad Sci U S A* **112**, E4256-4263, doi:10.1073/pnas.1505956112 (2015).
- 52 Schmidt, D., Martens, P. B., Weyand, C. M. & Goronzy, J. J. The repertoire of CD4+ CD28- T cells in rheumatoid arthritis. *Mol Med* **2**, 608-618 (1996).
- 53 Wagner, U. *et al.* Clonally expanded CD4+CD28null T cells in rheumatoid arthritis use distinct combinations of T cell receptor BV and BJ elements. *Eur J Immunol* **33**, 79-84, doi:10.1002/immu.200390010 (2003).
- 54 Piao, X. *et al.* G protein-coupled receptor-dependent development of human frontal cortex. *Science* **303**, 2033-2036, doi:10.1126/science.1092780 (2004).
- 55 Yang, L., Friedland, S., Corson, N. & Xu, L. GPR56 inhibits melanoma growth by internalizing and degrading its ligand TG2. *Cancer Res* **74**, 1022-1031, doi:10.1158/0008-5472.CAN-13-1268 (2014).
- 56 Chang, G. W. *et al.* The Adhesion G Protein-Coupled Receptor GPR56/ADGRG1 Is an Inhibitory Receptor on Human NK Cells. *Cell Rep* **15**, 1757-1770, doi:10.1016/j.celrep.2016.04.053 (2016).
- 57 Fransen, N. L. *et al.* Tissue-resident memory T cells invade the brain parenchyma in multiple sclerosis white matter lesions. *Brain* **143**, 1714-1730, doi:10.1093/brain/awaa117 (2020).
- 58 McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337 e324, doi:10.1016/j.cels.2019.03.003 (2019).
- 59 Hao Y, H. S., Andersen-Nissen E, III WMM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zagar M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LB, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P, Satija R Integrated analysis of multimodal single-cell data. *bioRxiv* (2020).
- 60 Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289-1296, doi:10.1038/s41592-019-0619-0 (2019).
- 61 Zappia, L. & Oshlack, A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience* **7**, doi:10.1093/gigascience/giy083 (2018).
- 62 McDavid, A., Finak, G. & Yajima, M. MAST: Model-based Analysis of Single Cell Transcriptomics. *R package version 1.16.0*, <https://github.com/RGLab/MAST/>. (2020).
- 63 Alquicira-Hernandez, J. & Powell, J. E. Nebulosa recovers single cell gene expression signals by kernel density estimation. *Bioinformatics*, doi:10.1093/bioinformatics/btab003 (2021).
- 64 Borcharding, N., Bormann, N. L. & Kraus, G. scRepertoire: An R-based toolkit for single-cell immune receptor analysis. *F1000Res* **9**, 47, doi:10.12688/f1000research.22139.2 (2020).

Figure 1

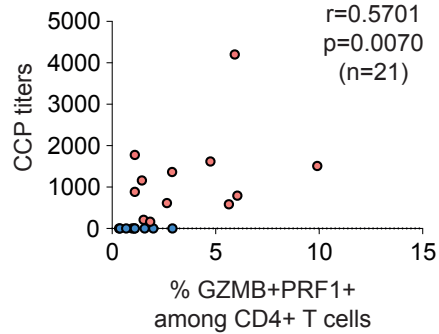
a



b



c



d

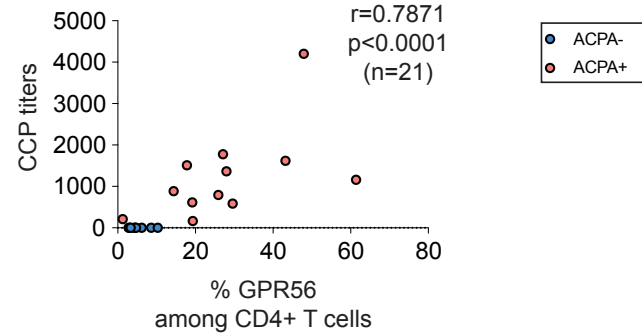
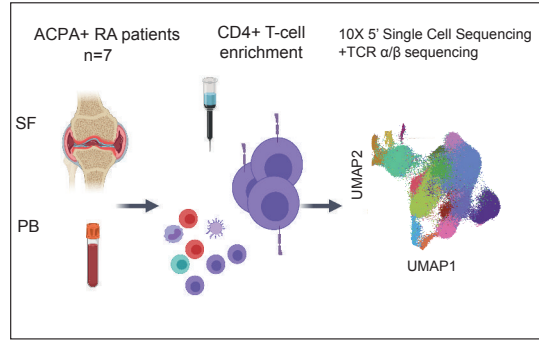


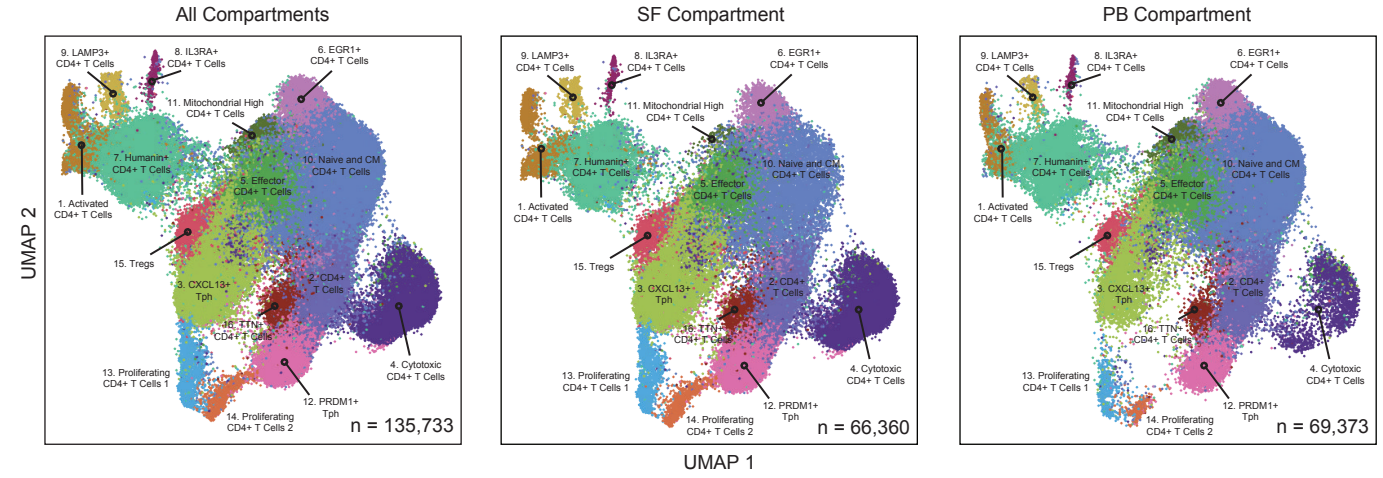
Figure 1. Cytotoxic CD4+ T-cell frequency in synovial fluid of ACPA- and ACPA+ RA patients. **a**) Representative flow cytometry dot plot staining of effector molecules, receptors and transcription factors associated with cytotoxic functions in CD4+ T cells from synovial fluid (SF) from ACPA- (upper panel) and ACPA+ (lower panel) RA patients, quantified in **b**) (ACPA-, n=7-9) (ACPA+, n=10-12). Line represents median, two-tailed Mann-Whitney U test, ns: not significant. **c**) Correlation between the frequency of GZMB+ PRF1+ in CD4+ T cells in SF and the level of serum anti-CCP (cyclic citrullinated peptides) antibodies levels, n=21, Spearman two-tailed test. **d**) Correlation between the frequency of GPR56 in CD4+ T cells in SF and the level of anti-CCP (cyclic citrullinated peptides) antibodies levels, n=21, Spearman two-tailed test. **a-d**) Data are from a pool of nine independent experiments where a circle is a single replicate. Blue dots indicate ACPA- RA SF and red dots indicate ACPA+ RA SF.

Figure 2

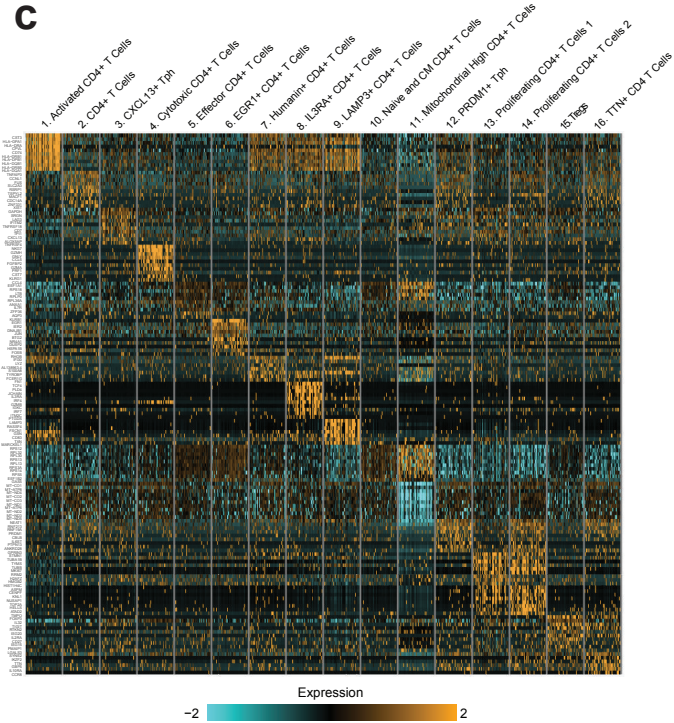
a



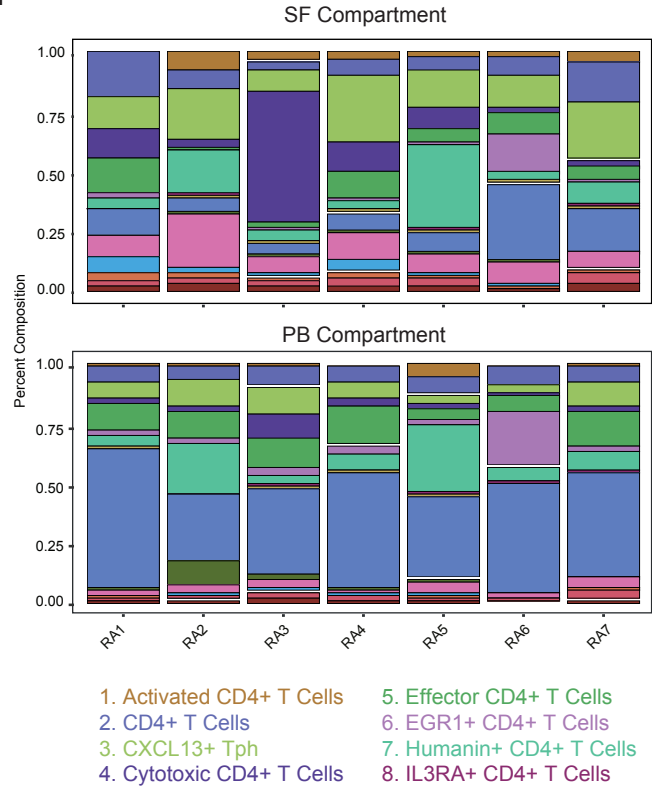
b



c



d



e

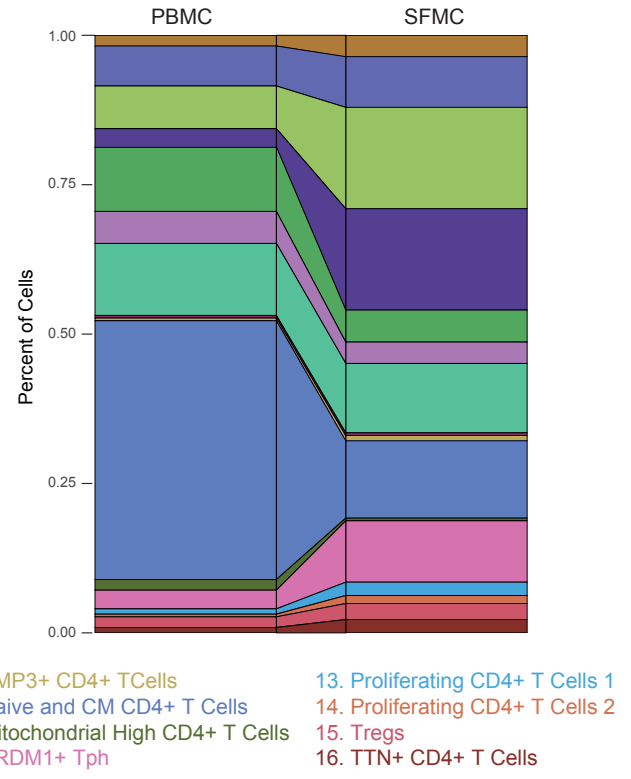


Figure 2. Single cell RNA sequencing of CD4+ T cells from SF and PB of ACPA+ RA patients. a) Technical workflow including CD4+ T-cell enrichment and 10X 5' single cell sequencing coupled to TCR α/β sequencing on 7 paired PB and SF from ACPA+ RA patients. b) UMAP plot of 135,733 cells colored by cell type cluster split by combined compartments (left), SF-only (middle) and PB-only (right panel) c) Heatmap showing gene expression values for a curated list of CD4+ T cell subset-defining genes d) Stacked barplots displaying cell type composition for each ACPA+ RA patient. e) Stacked barplots depicting comparative frequency of the CD4+ T-cell subsets found in PB and SF.

Figure 3

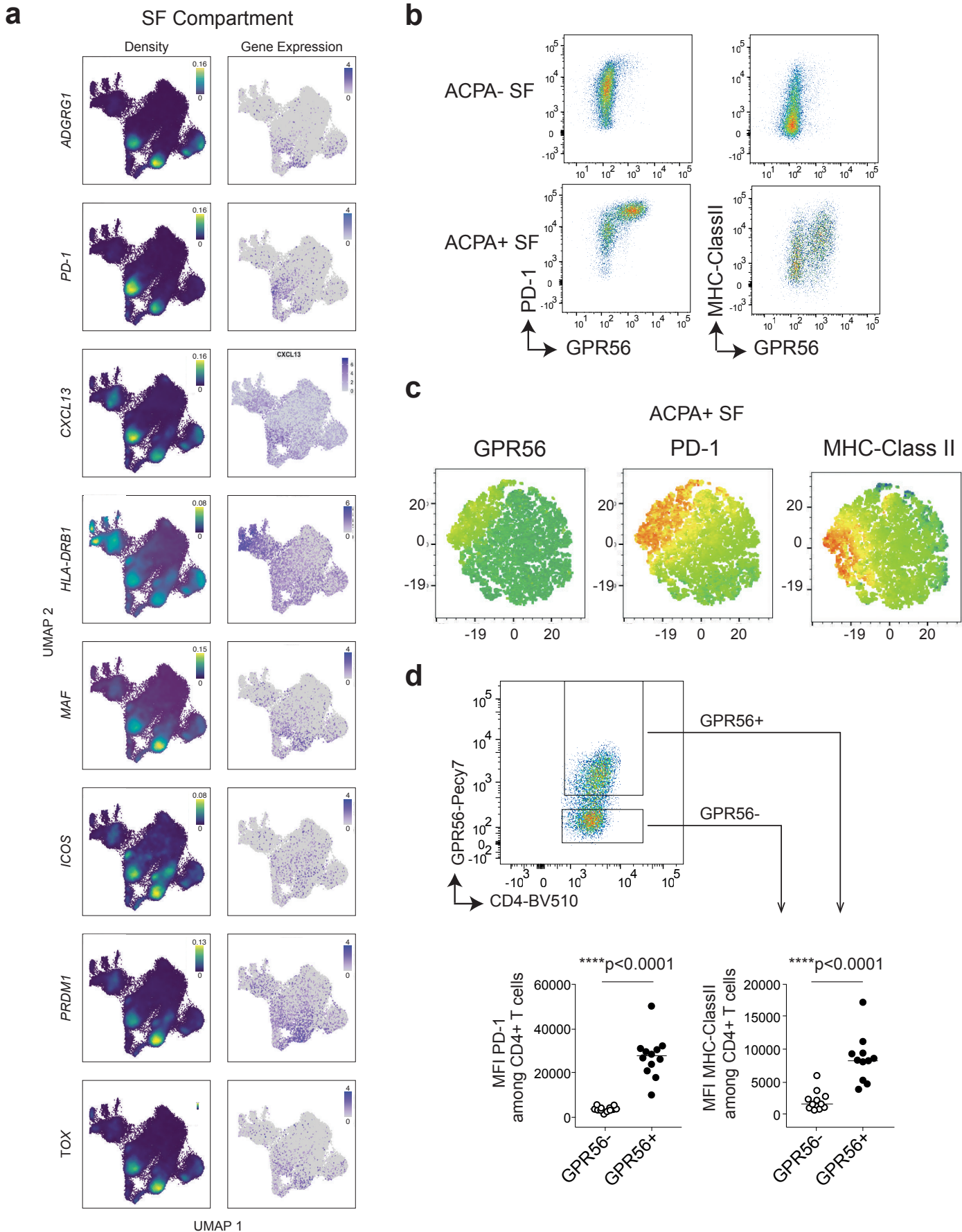


Figure 3. GPR56 expression on peripheral helper CD4+ T cells in ACPA+ RA SF. **a**) Nebulosa density (left panel) and gene expression (right panel) single cell plots of ADGRG1 and Tph CD4+ T-cell markers in ACPA+ RA SF. **b**) Representative flow cytometry dot plot of PD-1 and MHC-Class II staining in combination with GPR56 in ACPA- and ACPA+ RA SF. **c**) Representative T-SNE flow cytometry plots showing CD4, GPR56, PD-1 and MHC-Class II expression in one ACPA- and one ACPA+ RA SF (gated on CD3+ T cells). **d**) Mean fluorescence intensity (MFI) of PD-1 (n=12) and MHC-II (n=11) in GPR56-negative and positive CD4+ T cells in ACPA+ RA SF. Line represents median, two-tailed Mann-Whitney U test. Data are from a pool of nine independent experiments where a circle is a single replicate. White dots indicate GPR56- CD4+ T cells and black dots indicate GPR56+ CD4+ T cells.

Figure 4

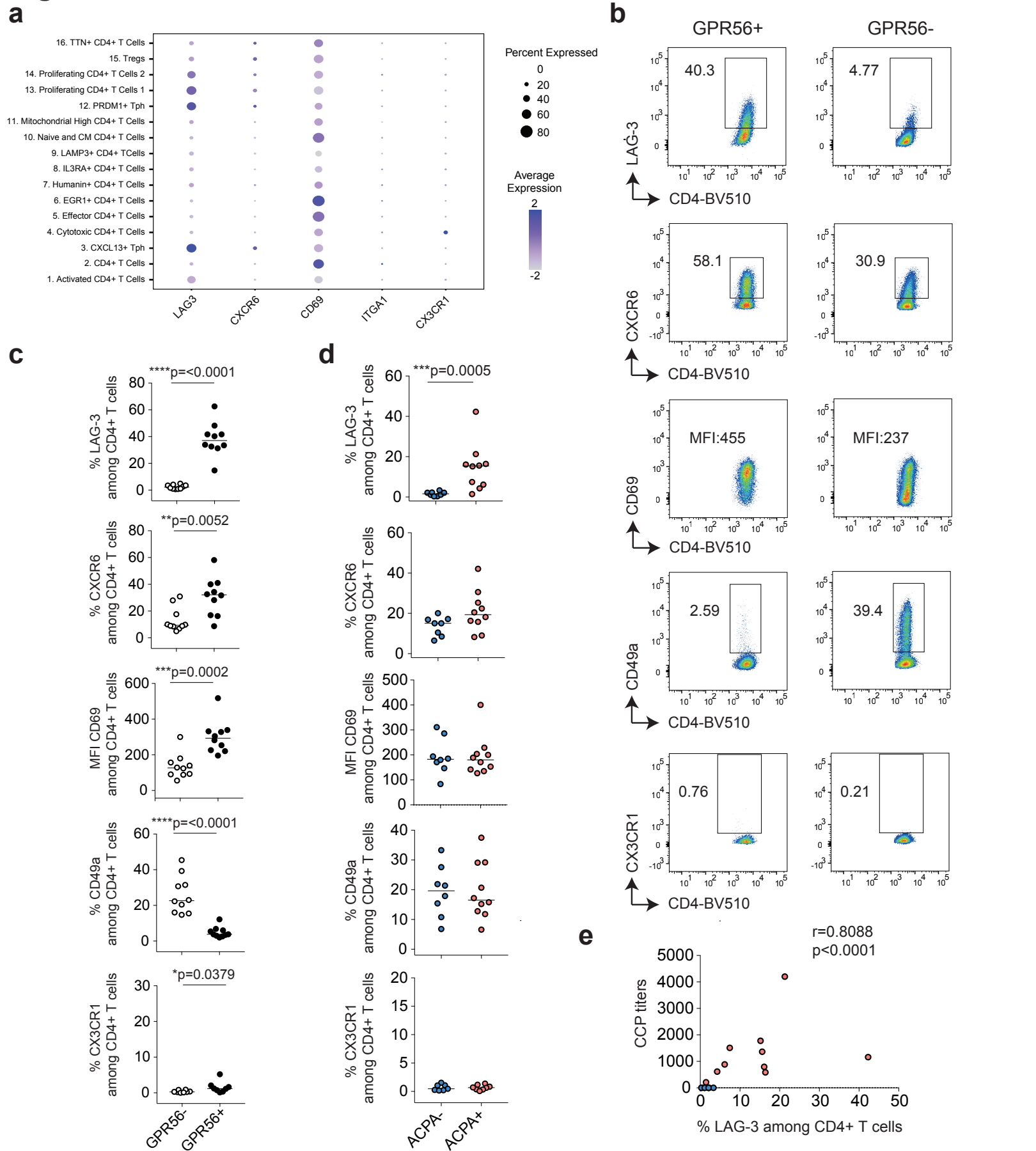
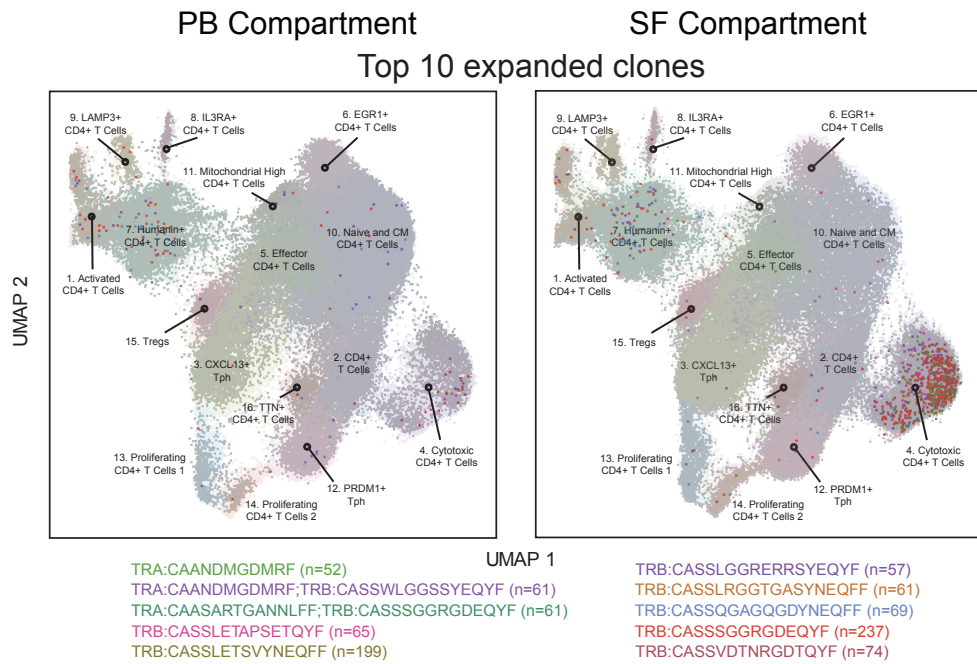


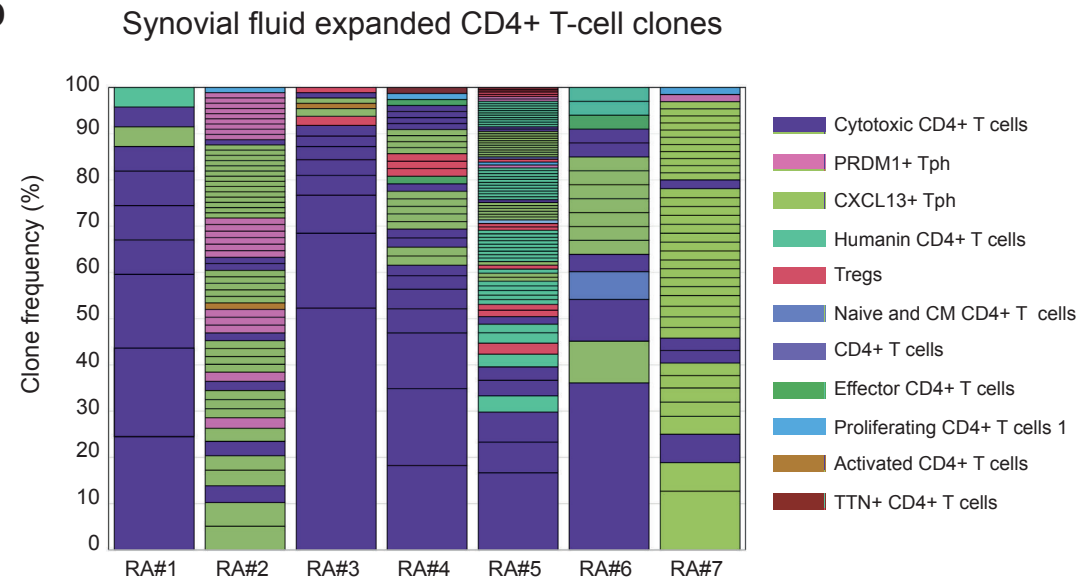
Figure 4. Tissue resident memory markers (TRM) on CD4+ T cells in ACPA+ RA SF. **a**) 2-D dot plots showing the expression of selected genes in the different CD4+ T-cell clusters in ACPA+ SF (circle size indicates cell frequency, color intensity indicates average expression). **b**) Representative flow cytometry dot plots showing the expression of LAG-3, CXCR6, CD69, CD49a and CX3CR1 on GPR56+ and GPR56- CD4+ T cells in ACPA+ RA, quantified in **c**) ($n=10$, LAG3, CXCR6, CD69, CD49a) ($n=8$, CX3CR1). White dots indicate GPR56- CD4+ T cells and black dots indicate GPR56+ CD4+ T cells. **d**) Expression of LAG-3, CXCR6, CD69, CD49a ($n=8$ ACPA-, $n=10$ ACPA+) and CX3CR1 ($n=7$ ACPA-, $n=8$ ACPA+) on CD4+ T cells in SF. Blue dots indicate ACPA- RA SF and red dots indicate ACPA+ RA SF. **c-d**) Line represents median, two-tailed Mann-Whitney U test. Data are from a pool of eight independent experiments where a circle is a single replicate. **e**) Correlation between the frequency of LAG-3 in CD4+ T cells in RA SF and the level of anti-CCP (cyclic citrullinated peptides) antibodies levels ($n=18$), Spearman two-tailed test.

Figure 5

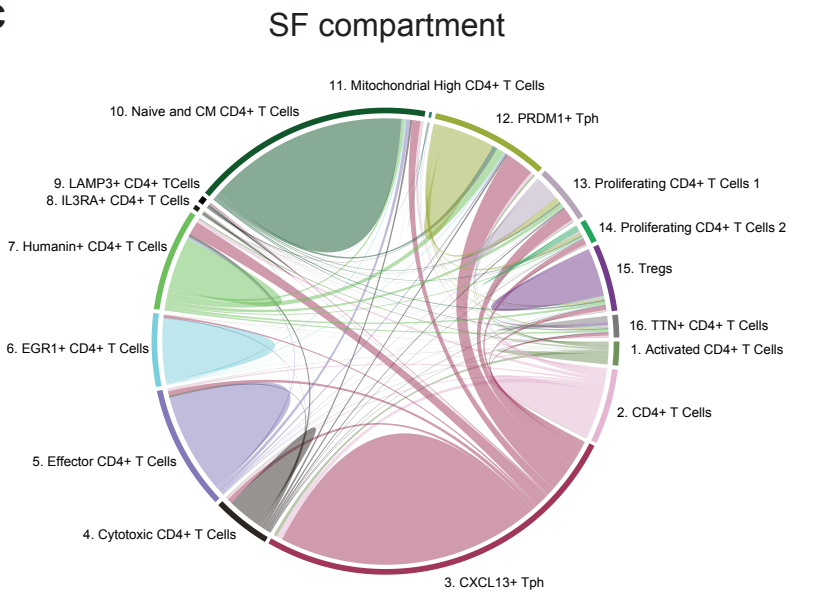
a



b



c



d

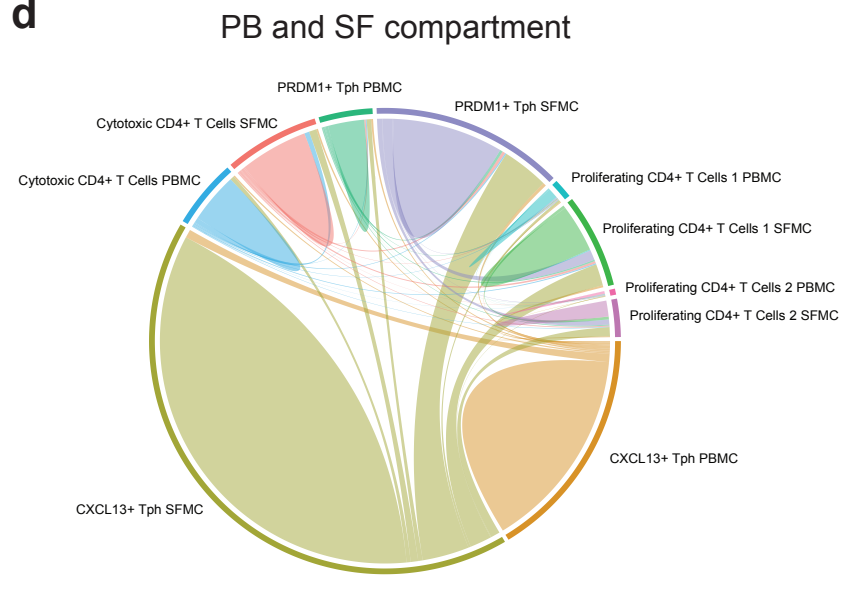


Figure 5. Clonally expanded CD4+ T cells in ACPA+ RA. a) UMAP plots showing the top 10 expanded clones in ACPA+ RA PB and SF (Each dot displays a cell). Each individual clone sharing the same CDR3 (from TCR α , TCR β or both) amino acid sequences is highlighted by a different color. **b)** Stacked barplots displaying the phenotype of the SF expanded clones (n>4 cells) in each ACPA+ RA patient, quadrant represents individual clone. **c)** Chord diagram showing the interconnectivity between CD4+ T cells sharing the same CDR3 (from TCR α , TCR β or both) amino acid sequences within CD4+ T-cell clusters in SF of 7 ACPA+ RA patients. **d)** Chord diagram showing the interconnectivity between CD4+ T cells sharing the same CDR3 (from TCR α , TCR β or both) amino acid sequences within selected CD4+ T-cell clusters: PDRM1+ Tph, CXCL13+ Tph, cytotoxic CD4+ T cells, proliferating CD4+ T cells in paired PB and SF.

Figure 6

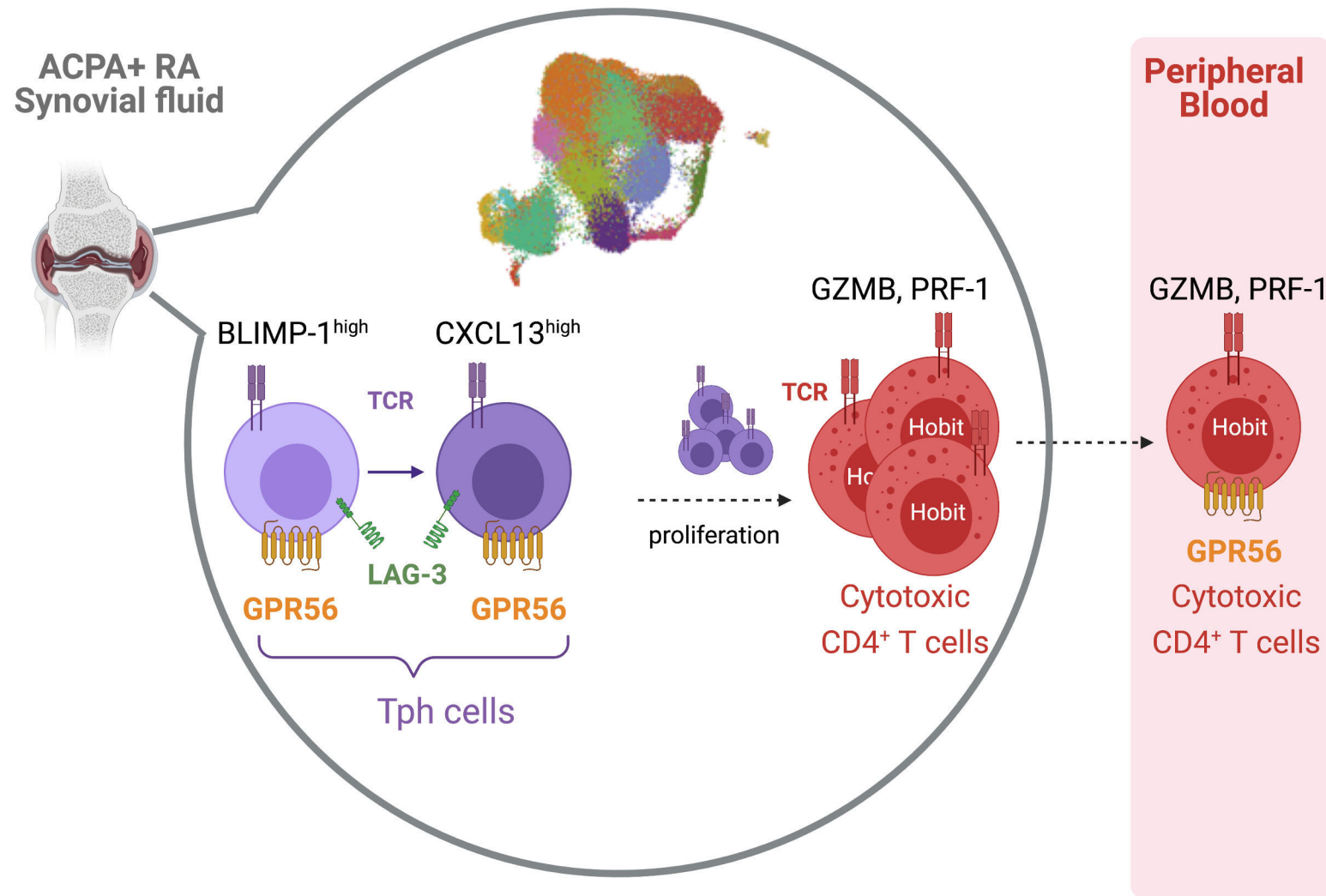


Figure 6. Graphical abstract showing a model for the differentiation of two states of peripheral helper T cells and cytotoxic CD4⁺ T cells in ACPA+ RA SF.