Single-cell transcriptome and TCR analysis of human Cytomegalovirus (hCMV)-specific T cell memory reveals effector and pre-effectors of CD8$^+$ and CD4$^+$ cytotoxic T cells

Raunak Kar[1], Somdeb Chattopadhyay[1], Anjali Sharma[2,#], Kirti Sharma[1,#], Gopalakrishnan Aneeshkumar Arimbasseri[1,*], Veena S. Patil[1,*,$]

[1]National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi

[2]VMMC and Safdarjung Hospital, New Delhi.

[#] Equal contribution

[*]Correspondence can be addressed to GAA (aneesh@nii.ac.in) or VSP (veena@nii.ac.in)

[$]Lead contact

Abstract

The latent human Cytomegalovirus (hCMV) infection can pose a serious threat of reactivation and disease occurrence in immune-compromised individuals, as well as burdens the immune system in immune-competent individuals. Though, T cells are at the core of the protective immune response to hCMV infection, a detailed characterization of different T cell subsets involved in protection against the hCMV infection is lacking. Here we analyzed the single-cell transcriptomes and the single-cell T cell antigen receptor (TCR) repertoires of over 8000 hCMV-reactive peripheral T cells isolated from different memory compartments. The hCMV-reactive T cells were highly heterogeneous and consisted of different developmental memory and functional T cell subsets such as, the long-term memory precursors and effectors, T helper-17, T regulatory cells (T_{REGs}) and cytotoxic T lymphocytes (CTLs). The hCMV-antigen specific T_{REGs} were enriched for molecules linked to their suppressive function and interferon response genes. The CTLs were of two types, the pre-effector and effector like. Of particular interest was the mixture of both CD4-CTLs and CD8-CTLs in both the pre-effector and effector cytotoxic clusters, suggesting that both CD4-CTLs and CD8-CTLs share transcriptomic signatures. The huge TCR clonal expansion of both the cytotoxic clusters imply their predominant role in protective immune response to CMV. Further the clonotype sharing between the CTL clusters and the long-term memory clusters, indicate potential progenitors of CD4-CTLs. Together our study has identified many subsets of hCMV-specific memory T cells that may have implication in better understanding the hCMV-specific T cell immunity to design vaccination strategies and therapeutics.

Introduction

Immune response to infections is a well-choreographed event involving various innate and adaptive immune cell types through the course of infection. The acquisition of immunological memory to infections is majorly the characteristic of adaptive immune cells such as B cells and T cells. During this evolutionarily conserved process of T cell immunological memory development, the naive T cells that have not previously encountered antigen, differentiate during the primary infection, into memory T cells that have specialized functions in immune defence, to a subsequent infection with the same pathogen. $CD4^+$ memory T cells are classified based on their functional properties and additionally, more recently, on their transcriptomes, such as but not limited to, T helper (T_H)1, T_H2, T_H17, $T_H1/17$, T follicular helper (T_{FH}), and T regulatory (T_{REG}) (Arlehamn *et al*, 2014; Crotty, 2019; O'Shea & Paul, 2010; Sallusto, 2016; Schmiedel *et al*, 2018). For example, T_{FH} cells are the specialized providers of help to B cells, and T_{REGs}, the regulators of the immune response with their immunosuppressive properties (Crotty, 2019; Sakaguchi *et al*, 1995; Scheinecker *et al*, 2020). Furthermore, each of those subtypes can exhibit great functional diversity and plasticity (Sallusto, 2016). The peripheral $CD4^+$ as well as $CD8^+$ memory T cells can also be classified based on their developmental stages as long-lived precursor memory and short-lived effector memory such as, stem cell memory T cells (T_{SCM}), central memory T cells (T_{CM}), effector memory T cells (T_{EM}), and effector memory T cells expressing CD45RA (T_{EMRA}) (Mahnke *et al*, 2013). The $CD8^+$ T cells are known for their cytotoxic functions, and the $CD4^+$ T cells, majorly, for their helper function where they provide help to other immune cell types including $CD8^+$ T cells. However, others and we have shown that $CD4^+$ T cells can also function as cytotoxic T cells and are enriched in the CD4-T_{EMRA} compartment and show marked TCR clonal expansion similar to the CD8-T_{EMRA} cells in many acute and chronic viral infections including the cytomegalovirus (CMV) (Cheroutre & Husain, 2013; Kusnadi *et al*, 2021; Lyu *et al*, 2021; Meckiff *et al*, 2020; Patil *et al*, 2018; van Leeuwen *et al*, 2004; Weiskopf *et al*, 2015).

After the initial infection, human CMV (hCMV) establishes persistence or latency in immune-competent individuals, although intermittent viral shedding can occur (Gandhi & Khanna, 2004). In absence of effective immune response especially in immune-compromised individuals, the virus can reactivate and cause the disease. Hence, reactivation of hCMV poses a higher risk for individuals undergoing transplant or chemotherapy and new-borns of seropositive mothers (Gandhi & Khanna, 2004). Both $CD4^+$ and $CD8^+$ T cells have shown to play a role in immune response to CMV. While hCMV-specific $CD8^+$ T cells have extensively been studied, studies on hCMV-specific $CD4^+$ T cells are sparse (Lim *et al*, 2020). Further, most of the studies have only analysed a fraction of hCMV reactive T cells of either CD4 or CD8 origin reactive to a single immune-dominant peptide or peptide pool derived from a single immunodominant protein. Hence, there is a limited understanding about the overall T cell immune response to hCMV infection or latency. Identifying and characterizing the components of hCMV-specific T cell immune response in CMV seropositive individuals who are otherwise healthy, will enable us to identify the corelates of protective immune response to hCMV.

The diversity and the temporal nature of expression of co-stimulatory molecules and cytokines as well as the specificity of HLA molecules to individual human donors has restricted the researchers to explore only a fraction of antigen-specific T cell repertoire based on either a specific cytokine, co-stimulatory molecule/s, or donors with specific HLA and immunodominant peptide. Hence, in this study, independent of their cytokine profile, we explored the spatio-temporal nature of the surface expression of co-stimulatory molecules CD154 and CD137 upon TCR dependent activation of T cells across different $CD4^+$ and $CD8^+$ memory subsets (T_N, T_{CM}, T_{EM}, T_{EMRA}) using both polyclonal stimulation and stimulation using hCMV peptide pools. We observed that all $CD8^+$ T cell memory subsets predominantly expressed CD137 at both 6 and 24 hours (hrs), while $CD4^+$ T cell memory subsets showed a more dynamic expression pattern where both CD154 and CD137 expression was observed across memory subsets though at varying levels. Hence to understand the heterogeneity and TCR clonal diversity of hCMV-specific T cells from different memory subsets, we isolated and performed single-cell transcriptomics analysis of over 8000 hCMV-specific T cells, isolated based on the surface expression of $CD154^+$ and/or $CD137^+$ across 10 donors, after stimulation with pool of hCMV peptides derived from not just the dominant pp65 but also IE-1 proteins, at both 6 hrs and 24 hrs. The unbiased clustering based on the transcriptomes identified nine distinct clusters of various functional subsets such as T_{REGs}, T_H17, T_H1 and $T_H1/17$ subsets, cytotoxic cells of both CD4 and CD8 origin (CD4-CTLs and CD8-CTLs) and actively cycling cells. The T_{REGs} were enriched for molecules known for suppressive functions, and the interferon (IFN) response signature genes. The transcriptomics analysis combined with TCR repertoire analysis revealed that the cytotoxic cells are majorly of 2 types; pre-effectors and effectors and are highly clonally expanded. The transcriptomes of pre-effectors and effectors of CD4-CTLs and CD8-CTLs were largely indistinguishable. Together our study has identified many subsets of hCMV-specific memory T cells which may have implications for understanding the reactivation of hCMV in immune-compromised individuals. This knowledge base can be an important resource for designing vaccines and therapeutics.

Results

Expression kinetics of T cell activation markers CD154 and CD137 across T cell memory subsets

To isolate antigen-specific T cells independent of their cytokine profile and HLA (Human Leukocyte Antigen)-restriction, the surface expression of co-stimulatory markers such as CD154 and CD137 upon the engagement of T cell antigen receptor (TCR) with cognate pMHC (peptide Major Histocompatibility Complex) complex has been widely used (Kusnadi *et al.*, 2021; Meckiff *et al.*, 2020; Poloni *et al*, 2023). The surface expression of both CD154 and CD137 upon TCR activation is both time- and T cell type-dependent (Poloni *et al.*, 2023). The surface expression of CD137 is observed between 16-24 hours (hrs) and CD154 at ~6 hrs after TCR-pMHC complex engagement (Kusnadi *et al.*, 2021; Meckiff *et al.*, 2020; Poloni *et al.*, 2023). We assessed the kinetics of surface expression of CD137 and CD154 in total $CD4^+$ and $CD8^+$ T cell subsets as well as within each individual memory

compartments at both 6 and 24 hrs after polyclonal stimulation with αCD3/CD28 beads in the presence of anti-CD154 antibody to minimise the receptor internalization (Fig. 1A-B). Total $CD8^+$ T cells predominantly expressed CD137 at both 6 and 24 hrs (Fig. 1B-C). However, a higher percentage of the $CD4^+$ T cells expressed CD154 compared to CD137 at both 6 and 24 hrs, although the difference was markedly reduced at 24 hrs (Fig. 1B-C). Overall, both CD154 and CD137 showed higher expression at 24 hrs compared to 6 hrs in both $CD4^+$ and $CD8^+$ T cell compartments (Fig. 1B-C). Similarly, we next assessed the surface expression kinetics of CD154 and CD137 across different memory subsets within $CD4^+$ and $CD8^+$ T cell compartments: naïve T cells (T_N; $CD45RA^+CCR7^+$), central memory T cells (T_{CM}; $CD45RA^-CCR7^+$), effector memory T cells (T_{EM}; $CD45RA^-CCR7^-$), and effector memory T cells expressing CD45RA (T_{EMRA}; $CD45RA^+CCR7^-$) (Mahnke *et al.*, 2013). While the $CD8^+$ T cells showed significantly higher expression of CD137 across all 4 memory compartments, the $CD4^+$ T cell memory subsets expressed both CD154 and CD137, but the expression was both time- and memory subset-dependent (Fig. 1B and D). Within $CD4^+$ T cell memory compartments, the expression of CD137 was higher in effector memory subsets (T_{EM} and T_{EMRA}) compared to long-term memory (T_{CM}) and $T_{N,}$ but the expression of CD154 between memory subsets within $CD4^+$ T cell subsets did not change significantly at 6 hrs (Fig. 1B and D). At 24 hrs the expression of CD154 and CD137 were comparable within each memory compartment of $CD4^+$ T cells (Fig. 1B and D). Overall, a combined analysis of 6 and 24 hrs, revealed significantly higher percentage of cells expressing CD137 than CD154 in all four $CD8^+$ T cell memory subsets, and CD4-$T_{EMRA,}$ but not in the other three $CD4^+$ T cell subsets (T_N, T_{CM}, T_{EM}) (Fig. 1E). Similarly, when we stimulated the PBMCs isolated from human Cytomegalovirus (hCMV)-seropositive donors, with hCMV peptide pools derived from pp65 and IE-1 proteins, we observed memory subtype- and time-dependent expression of CD154 and CD137 in $CD4^+$ memory T cell compartment while CD8 memory T cells predominantly expressed CD137 (Fig. S1A-B and Fig. 1F). Together, these results suggest that the expression of CD137 and CD154 is dynamic and depends on the duration of stimulation, T cell types and memory sub-type. Hence in subsequent experiments, to capture all activated cells independent of time- and memory-subset-specific expression, cells that expressed either CD154 or CD137 in each memory compartment, in response to TCR stimulation, were considered as activated cells (Fig. 1G and Fig. S1B).

Single-cell transcriptomic analysis of CMV-specific memory T cells

To identify and understand the heterogeneity and clonal diversity of hCMV-specific T cells across different memory subsets, we performed single-cell transcriptomic (scRNA-Seq) and single-cell TCR repertoire analysis (scTCR-Seq) of >8000 hCMV-specific T cells (Fig. 2A). To isolate, hCMV-reactive T cells for scRNA-Seq, we stimulated the PBMCs, with hCMV-peptide pools (pp65 and IE-1) for a duration of 6 and 24 hrs, from 10-seropositive individuals that showed comparatively higher response to peptide stimulation (10 out of 16 donors) (Fig. 1G, Fig. 2A and Fig. S1A-B). The responding cells from different $CD4^+$ memory T cell compartments including naïve (T_N, T_{CM}, T_{EM} and T_{EMRA}) were

isolated using florescence activated cell sorting (FACS) based on the surface expression of CD154$^+$ and/or CD137$^+$ (Fig. 1B and F-G, and Fig. 2A). CD4-T_{EMRA} subset has been shown to be enriched for CD4-CTLs (cytotoxic T lymphocytes), hence to compare the hCMV-specific CD4-CTLs and CD8-CTLs we also sorted hCMV-specific T cells from CD8-CTLs (CD8$^+$CCR7$^-$ (CD8-T_{EM} and -T_{EMRA})) (Fig. 2A) (Patil *et al.*, 2018). To have representation from each memory subset, the cells from each of these CD4 and CD8 memory subsets were pooled in desired ratios before performing scRNA-seq and scTCR-Seq assays using 10X genomics platform (Fig. 2A). An unbiased clustering of 6899 T cells (out of 8062 cells) that passed our quality control parameters (QC; please see methods section), identified 9 distinct clusters based on transcriptomes (Fig. 2B), and was not influenced by donors or the experimental batch (Fig. S2A-B and Table S1). Different developmental T cell memory subsets were distinguished based on the surface expression of oligo-tagged antibodies (antibody derived tags; adt) against CD45RA, CCR7, CD95 and IL7Rα as well as *CCR7*, *CD27*, *CD28* transcripts (Fig. 2C-D and Table S2) (Mahnke *et al.*, 2013; Patil *et al.*, 2018). Transcripts for *CD4* were distributed across all clusters while *CD8B* transcripts were majorly found in clusters 2 and 4 (Fig. 2E). Cells in cluster 2 and 4 were mostly T_{EMRA} cells as they expressed CD45RA (adt CD45RA) and lacked the expression of CCR7 protein (adt CCR7), and transcripts for *CCR7*, *CD28* and *CD27* (Fig. 2C-D) (Mahnke *et al.*, 2013; Patil *et al.*, 2018). T_N cells were very few and were mostly found in cluster 6 (CD45RA$^+$CCR7$^+$IL7Rα$^+$) (Fig. 2C-D). T_{CM} and T_{EM} cells were distributed across multiple clusters (Clusters 0, 1, 3, 5, 6, 7) (Fig. 2C-D). These observations were independently confirmed by the enrichment of subset-specific gene signatures from T_{CM}, T_{EM}, and T_{EMRA}, as well as common gene signatures from T_{CM}-T_{EM}, and T_{EM}-T_{EMRA} (Patil et al, 2018) (Fig. 2F and Fig. S2C, Table S3). Similar to the flow cytometry analysis (Fig. 1), we observed that T_{EMRA} cells (clusters 2 and 4) majorly showed higher expression of *CD137* while other T_{CM} and T_{EM} clusters (0, 3, 6) showed higher expression of *CD154* (Fig. 2G). Interestingly many of the T cell activation markers showed dynamic expression amongst the clusters; *CD154* and *CD137* showed opposing expression and similar observations were made for *CD69* and *IL2RA* (Fig. 2D and G). Further, co-expression analysis of various T cell activation markers in combination, confirmed the variability in expression (Fig. 2G). Most strikingly, the *CD137*^{hi}*CD154*^{lo} cells showed low *CD69* expression, while the *CD154* and *CD137* co-expressing cells expressed similar amounts of *TNFRSF4* transcripts (encoding OX40) (Fig. 2G). Together, these results show that our data well represents various memory subsets that show differential expression of T cell activation markers, and hence validating the strategy used in our study to isolate antigen-specific T cells based on the expression of CD154 and/or CD137 from both 6 and 24 hrs after stimulation.

hCMV-specific memory T cells are heterogeneous and can be classified into multiple functional T cell subsets

The differential expression analysis of transcripts amongst the nine clusters of hCMV-specific T cells, identified cluster specific expression of transcripts, revealing many functional T cell subsets (Fig. 3A-

B and Table S4). Multiple clusters showed enrichment for long-term memory markers such as *AQP3*, *CCR7*, *TCF7*, *IL7R*, *LTB*, *JUNB*, *SELL* (Fig. 3A-B) (Meckiff *et al.*, 2020; Patil *et al.*, 2018). Cluster 1 showed enrichment for *FOXP3*, *BATF*, *TIGIT* and *TNFRSF1B* transcripts that are known to be expressed in T regulatory cells (T_{REG}) and transcripts encoding interferon (IFN) response signature genes (*ISG20*, *ISG15*, *OAS1*, *IFIT3*, *MX1*) (Fig. 3B). In an independent analysis, cluster 1 cells showed a higher enrichment signature score for overall T_{REG} and IFN response signature gene sets (Fig. 3C and Table S3), suggesting that cells in cluster 1 are T_{REG} cells and are possibly responding to IFN signalling. Cells with active cell cycle were found in cluster 3, one of the clusters enriched for T_{CM} signatures (Fig. 3B-C and Table S3). Cells in cluster 2 and 4 that showed enrichment for T_{EMRA}-specific gene sets, expressed transcripts encoding cytolytic genes such as *GZMB*, *PRF1* and *NKG7* and showed enrichment for cytotoxicity gene signatures (Fig. 2F, Fig. 3A-C and Fig. S2C and Tables S3 and S4) (Patil *et al.*, 2018), suggesting that clusters 2 and 4 are the cytotoxic clusters.

We next interrogated the single-cell transcriptome of hCMV-specific T cells for functional T helper (T_H) cell subtypes such as T_H1, T_H17 and $T_H1/17$, that are known to have role in viral infections (Bahr *et al*, 2013; Meckiff *et al.*, 2020). Multiple clusters showed expression of T_H subset-specific transcripts: T_H17 (*CCR4*, *CCR6*, *IL4I1*, *CTSH*), T_H1 (*CXCR3*, *TNF*, *IFNG*, *CSF2*) and $T_H1/17$ (*CXCR3*, *DPP4*) (Arlehamn *et al.*, 2014; Schmiedel *et al.*, 2018) (Fig. 3D and Table S4). Interestingly, we found several clusters co-expressing transcripts from different T_H subsets, suggesting each cluster may have cells belonging to more than one T_H subtype. Independently, module enrichment analysis using T_H subset-specific gene sets corroborated these results (Fig. S3A and Table S3). The T_H1-specific gene sets were enriched in cluster 2 and 4, the cytotoxic clusters (Fig. S3A). Considering there are several overlapping genes between T_H1 and cytotoxic gene sets, this observation is in alignment with the known literature (Arlehamn *et al.*, 2014; Patil *et al.*, 2018; Schmiedel *et al.*, 2018). T_H17 signature genes showed highest enrichment in cluster 0, followed by clusters 7, 3 and 1. $T_H1/17$ signatures were mostly enriched in cluster 0 followed by clusters 5 and 7 (Fig. S3A) (Arlehamn *et al.*, 2014; Schmiedel *et al.*, 2018). Together these observations indicate that each hCMV-specific T_H subset is at different developmental stages of the memory (long-term memory and effector memory).

TCR clonally expanded hCMV-specific T cells are majorly observed in cytotoxic clusters.

To understand the TCR clonal diversity amongst the hCMV-specific T cells we analysed the TCR-Seq data generated from the same cells. We recovered TCRs from 83.56% (5765 cells of 6899 cells) of the total QC-ed cells in our data set, of which 21.40% (1234 of 5765 cells) of the cells were clonally expanded (Fig. 3E and Table S5). Several clusters shared TCR clonotypes, indicating the developmental lineage sharing between these clusters (Fig. S3B). A maximum of 14 clonotypes were shared between clusters 2 and 4, followed by clusters 0 and 3, sharing 10 clonotypes (Fig. S3B). Cells in cluster 2 and 4 showed highest clonal expansion with ~78% and ~76% of the cells being expanded respectively (Fig. 3E and Table S5). As high as 76% of the total expanded cells (938 of 1234) were found in cluster 2

(50%) and cluster 4 (~26%) (Table S5). A few top expanded clonotypes were found in as many as 144, 79 and 51 cells from clusters 2 and 4, indicating the preferential expansion of a few specific clonotypes (Fig. 3F and S3C). Interestingly, close to 58% of the expanded cells carried only 20 clonotypes (Fig. 3F). Both CD4-CTLs and CD8-CTLs (T_{EMRA}) are shown to be clonally expanded in several diseases including CMV in both humans and animal models (Klenerman & Oxenius, 2016; Lim *et al.*, 2020; Meckiff *et al.*, 2020; Patil *et al.*, 2018; Schober *et al*, 2020). Clonal expansion and enrichment for cytotoxic signature in clusters 2 and 4 emphasizes the role for cytotoxic cells in hCMV-specific protective immune response.

Single-cell transcriptome analysis of hCMV-specific T_{REG} cells

T_{REGs} are known for their immunosuppressive function and are characterised by the higher expression of CD25 (IL2RA) and the transcription factor FOXP3, but a lower expression of IL7Rα (CD127) in humans (Seddiki *et al*, 2006). CMV-specific T_{REGs} have been reported in both mouse models and humans, however their detailed signatures have not been well described (Almanan *et al*, 2017; Lim *et al.*, 2020; Lyu *et al.*, 2021; Tovar-Salazar & Weinberg, 2020; Velaga *et al*, 2013). The transcript encoding the T_{REG} specific transcription factor FOXP3 was differentially expressed in clusters 1 (85% of cells) and 7 (Fig. 3B, Fig. 4A-B and Table S4), however though not differentially expressed, a portion of cells from cluster 5 showed high *FOXP3* expression (41% of cells) (Fig. 4A-B). The cells co-expressing *FOXP3*, *IL2RA*, and *FOXP3*, *IKZF2* were enriched in cluster 1 and portion of cluster 5, and as expected they showed lower expression of IL7Rα (Fig. 4A and Fig. 2C) (Seddiki *et al.*, 2006). Clusters 1 and 5 were also characterized by the higher expression of other T_{REG}-specific transcripts such as *TIGIT*, *BATF*, *TNFRSF1B* and the Ikaros Zinc Finger (IkZF) transcription factor (TF) family members *IKZF1* encoding Ikaros, *IKZF2* encoding Helios, *IKZF3* encoding Aiolos, *IKZF4* encoding Eos (Fig. 3B and Fig. 4C). The overall T_{REG} gene signature was enriched in cluster 1 and fraction of cluster 5 (Fig. 3C). Together these results show that T_{REG} cells are predominantly in cluster 1 and portion of cluster 5. We next interrogated if the hCMV antigen-specific T_{REGs} in our data set were indeed memory T_{REGs}. Based on the expression of CD45RA, the $FOXP3^{+}$ T_{REGs} have been classified as resting or naïve T_{REGs} ($FOXP3^{lo}CD45RA^{+}$) and effector T_{REGs} ($FOXP3^{hi}CD45RA^{-}$) (Miyara et al 2009, Valmori et al 2005). Further the effector T_{REGs} have been classified as central and effector memory cells based on the expression of CCR7, CD27 and CD28 (Matsuki et al, 2013). The T_{REGs} in our data set were characterized by the lack of expression of CD45RA, and a moderate expression of CCR7, *CD27* and *CD28*, indicating that they are mostly central and effector memory T_{REGs} rather than naïve T_{REGs} (Fig. 2C-D).

Although, T_{REGs} expressing FOXP3, are known for their suppressive function, many reports have shown the presence of T cells expressing FOXP3 transiently, in response to TCR activation that are non-suppressive in nature. Hence, we assessed if the T_{REGs} in our hCMV peptide stimulated T cell data set are suppressive in nature. Our data provided several evidence for the suppressive nature of

T_{REGs} found in cluster 1 and 5, while the non-suppressive nature of cells expressing low levels of *FOXP3* observed in clusters 0 and 3 (Fig. 4A). The T_{REGs} in clusters 1 and 5 were *CD137*$^{+}$*CD154*^{lo}, the activation induced characteristic of T_{REGs} implicative of epigenetically stable antigen-activated T_{REGs} that retain the high suppressive potential in humans after *in vitro* and *ex vivo* activation (Fig. 2D and G) (Nowak *et al*, 2018). These T_{REGs} in clusters 1 and 5, showed higher expression of *TIGIT*, *TNFRSF18* (GITR) and *CTLA-4*, genes that have been shown to be associated with the T_{ERGs} containing high affinity TCRs, further confirming their antigen-specificity (Fig 3B and Fig. 4C). The T_{REGs} in our data set also showed higher expression of *ENTPD1* (CD39), an ectoenzyme ATP apyrase, *CCR4*, *CTLA4*, Helios (*IKZF2*) and FAS (CD95), molecules that have been shown to be associated with effector and suppressive function (Fig. 2C, Fig. 3D, Fig. 4C) (Deaglio *et al*, 2007; Jayaratne *et al*, 2018; Raffin *et al*, 2013). Previous reports have suggested the suppressive function of Helios in T_{REGs} where Helios^{+}FOXP3^{+} T_{REG} cells produced lower amounts of T_H17-specific cytokine, IL17, compared to Helios^{-}FOXP3^{+} T_{REG} cells, hence promoting the formation of T_{REGs} rather than T_H17 subtype (Raffin *et al.*, 2013). Consistent with these observations, we found that clusters 0 and 3 that expressed *IL17F* and *IL17A*, expressed low levels of *FOXP3*, but lack the expression of *IKZF2* (Helios) and these cells showed higher expression of CD154 and lower expression of CD137 (Fig. 4A and C, Fig. 3B and D) (Nowak *et al.*, 2018; Raffin *et al.*, 2013). Together these observations suggest that the low *FOXP3* expressing cells in clusters 0 and 3 are potentially the induced/transient T_{REGs} that are possibly expressing the FOXP3 transiently in response to activation and are non-suppressive in nature. Conversely, T_{REG} cells in cluster 1 and 5 are Helios^{+}FOXP3^{+}, and lack the expression of *IL17A* and *IL17F*, indicative of suppressive nature (Fig. 4A and C and Fig. 3B and D). The lack of expression of exhaustion markers *PDCD1* and *LAG3* further suggests these are not associated with anergy or exhaustion (Fig. 4C). Together these results provide multiple evidence to prove that the T_{REGs} in clusters 1 and 5 are indeed effector T_{REGs} and possess suppressive ability. One of the other mechanisms through which T_{REGs} execute their suppressive function on effector T cells is through suppressor cytokines such as IL10, TGFβ and IL35 (reviewed in (Shevach, 2009)). The T_{REGs} in cluster 1 lacked or showed low expression of these suppressor cytokine transcripts *TGFB1*, *IL10* and *IL35*, indicating that cells in cluster 1 may not be actively suppressing the hCMV-specific effector T cells (Fig. 4C and Table S4).

We next interrogated the T_{REG} subsets in our data set for their nature of suppressive role in hCMV-specific immunity. Studies using mouse models have shown that T_{REGs} can show remarkable functional plasticity and execute their suppressive function by expressing the TF and the chemokine receptor of the target T_H subtypes (Koch *et al*, 2009; Scheinecker *et al.*, 2020; Sefik *et al*, 2015; Wang *et al*, 2011; Zheng *et al*, 2009). For example, the T_{REGs} suppressive to T_H1 will express T-bet (encoded by *TBX21*) and CXCR3 (T_H1-T_{REGs}), while the ones toward T_H2 will express GATA3 and CCR4 (T_H2-T_{REGs}), albeit at a reduced level compared to the T_H subtype itself, without necessarily themselves making the T_H subtype specific cytokines (Koch *et al.*, 2009; Scheinecker *et al.*, 2020; Wang *et al.*, 2011; Zheng *et al.*, 2009). Interestingly the differential gene expression analysis revealed that the T_{REGs}

in cluster 1 and 5, are characterized by the higher expression of T_H2-specific TF *GATA3* and chemokine receptor *CCR4* along with *FOXP3* (Fig. 4D-F). GATA3 has been shown to be important for FOXP3 expression and T_{REG} functions (Wang *et al.*, 2011). Hence, we independently, examined the overall T_H2 subset-specific gene signatures using module enrichment score. We found that both cluster 1 and 5 are enriched for T_H2 subset-specific gene signatures (Fig. S4A), suggesting the T_{REGs} in cluster 1 and 5 are T_H2-like T_{REGs}. A recent study in cancer has also shown the presence of T_H2- like T_{REGs} that were marked by a higher expression of *GATA3* (Chu *et al*, 2023). Clusters 1 and 5 were also characterised by the enrichment of IFN response signature genes (Fig. 3B-C). In asthma and house dust mite allergy, it has been shown that both T_H and T_{REG} subsets enriched for IFN response gene sets showed dampened T_H2 response in individuals with no prior history of allergy or asthma (Seumois *et al*, 2020). Interestingly, in our data set a joint enrichment analysis using the house dust mite allergy specific T_{REG} gene signature and T_H2-specific gene signature showed enrichment in cluster 1 and 5 (Fig. S4B) (Seumois *et al.*, 2020). Together these results indicate that T_{REGs} in cluster 1 and 5 are T_H2-T_{REGs} with IFN response signatures that are potentially suppressive towards T_H2. T follicular regulatory cells (T_{FRs}) are the T_{REGs} that are known to suppress T_{FH} responses and are characterized by the expression of IL1R2 and CCR8 (Eschweiler *et al*, 2021; Meckiff *et al.*, 2020; Xie *et al*, 2019). The hCMV-reactive T_{REG} cells in cluster 1 and 5 co-expressed *IL1R2* and *CCR8* transcripts along with *FOXP3* transcripts, hence, may represent T_{FR} cells (Fig. 4C-F) (Eschweiler *et al.*, 2021; Meckiff *et al.*, 2020; Xie *et al.*, 2019). These interesting observations warrants further investigation to understand the relevance T_H2-T_{REGS} and T_{FRs} in CMV immunity, especially when hCMV-reactive conventional (non-T_{REGs}) T_H2 and T_{FH} cells were absent in our data set.

hCMV reactive CD4-CTLs and CD8-CTLs share similar transcriptome profile

Majority of the cells in clusters 2 and 4 and fraction of cells in clusters 0 and 7 were characterized by the high expression of transcripts encoding cytolytic molecules such as Granzyme B (*GZMB*), Perforin (*PRF1*), Granulysin (*GNLY*), and transcription factors *HOPX* and *ZEB2* and showed enrichment for cytotoxicity and T_{EMRA}-specific signature genes (Fig. 2F, Fig. 3B-C and Fig. 5A-B) (Meckiff *et al.*, 2020; Patil *et al.*, 2018). Interestingly, these cytotoxic clusters could not be distinguished based on the expression of *CD4* or *CD8A* and *CD8B* transcripts, but by other cytotoxicity related molecules (Fig. 2E, and Fig. 5A-C). The cytolytic granules have been shown to contain chemokines such as CCL3, CCL4 and CCL5 in addition to the classical mediators of cytotoxicity (GZMB, PRF1, GNLY), in both viral and bacterial infections, indicating their role in cytolysis (Caccamo *et al*, 2002; Wagner *et al*, 1998). Cells in cluster 2 showed higher expression of the transcripts encoding the cytokines such as *TNF* and *IFNG*, chemokines such as *CCL3*, *CCL4*, *CCL4L2*, *XCL1*, *XCL2*, while cluster 4 showed higher expression of *CCL5* (Fig. 5B-C). The chemokine receptor *CCR5*, the receptor for chemokines *CCL3*, *4* and *5*, and *IFNGR1*, encoding the receptor for IFNG, was expressed by the cytotoxic cells in cluster 0, suggesting these cells may be responding to the cytokines and chemokines made by the cells

in cytotoxic clusters 2 and 4 and are potentially intermediate populations (Fig. 5A-B and Table S4) (Eberlein *et al*, 2020). This is further supported by the single-cell trajectory analysis, where on the pseudo time scale cluster 0 appears before clusters 2 and 4 (Fig. 5D). These results suggest that the cytotoxic cells in cluster 0 are potentially an intermediate population in the CTL lineage and these cells are mostly T_{EM} cells as they mostly lack the expression of CD45RA and CCR7 (Fig. 2C).

Compared to cluster 4, cluster 2 cells showed higher expression of *SLAMF7*, *ZEB2*, *GZMB*, *GPR18*, *CD72*, *PDCD1* (PD1), *CRTAM* that are known to be expressed in terminally differentiated effector cells in viral infections and various cancers (Awwad *et al*, 2021; Correa-Rocha *et al*, 2018; Meckiff *et al.*, 2020; Omilusik *et al*, 2015; Sumida & Cyster, 2018) (Fig. 3B, Fig. 4C and Fig. 5B and D). On the contrary, cluster 4 was characterized by the higher expression of *GNLY*, *EOMES*, *ADGRG1* (GPR56), *KLRD1*, *UCP2* and *CRIP1* (Fig. 3B and Fig. 5B and D), that are not associated with terminally differentiated effector T cells. For example, CRIP1 has been shown to be anti-apoptotic in cancers and the antigen-specific cells expressing UCP2 have been shown to survive better in HIV infection through metabolic reprogramming (Chaudhuri *et al*, 2016; Zhang *et al*, 2019), suggesting the possible role of these genes in providing survival advantage to cells in cluster 4. A coordinated expression of CCL5 followed by GNLY and PRF1 has been shown to provide host-defence against *Mycobacterium tuberculosis* (Mtb) infection, where the expression of CCL5 attracted the Mtb infected macrophages that in-turn triggered the expression of GNLY and PRF1 in CD8 T cells (Stegelmann *et al*, 2005). Further, a delayed onset of cells expressing GNLY has been shown in multiple infections (Pena *et al*, 1997; Pena & Krensky, 1997). These results suggest that the cytotoxic cells could be at different stages of their development or effector function. The differential expression of many of these cytolytic molecules between clusters 2 and 4, imply that the cells in cluster 2 are potentially the terminal effector cells while those in cluster 4 are pre-effector cells. This observation is further strengthened by the TCR analysis (Fig. 5E-F, Fig. 3E-F and Fig. S3B-C). The 25 most expanded clonotypes were found majorly in cluster 2 and 4 (Fig. 3F and Fig. S3C). Cluster 2 is relatively more clonally expanded than cluster 4, where over 67% of the cells harbouring the 20 most expanded clonotypes were found in cluster 2 (Fig. 5F). The trajectory analysis also showed that on a pseudo-time scale, cluster 2 and 4 appears to bifurcate from a common origin of cluster 0 (Fig. 5D). Further, the pathway enrichment analysis of genes differentially expressed between clusters 2 and 4 revealed an enrichment for cytokine and chemokine response pathways in cluster 4, while cluster 2 showed enrichment of pathways associated with pathogen induced cytokine storm (Fig. S5A-B). The effector T cells have been shown to utilize glycolysis, as the predominant choice of metabolic pathway and play an important role for cytokine secretion and lytic nature of these effector cells (Cham & Gajewski, 2005; Gubser *et al*, 2013; Rangel Rivera *et al*, 2021). The cells in cluster 2 showed enrichment for many metabolic pathways associated effector cells such as glycolysis and HIF1α signalling as well as pathways related to anti-viral responses (Fig. S5A). Together, these combined analysis of the transcriptome, TCR clonotype

expansion and sharing along with the trajectory analysis provides the evidence to support the idea that cluster 4 cells are pre-effectors while cluster 2 cells are effectors.

Next to asses if the observed differences between the CTL clusters 2 and 4 are influenced by the proportion and the characteristics of CD4-CTLs and CD8-CTLs, we compared the expression of these transcripts between CD4 and CD8 cells within clusters 2 and 4. Firstly, the proportions of CD4 and CD8 T cells within clusters 2 and 4 were comparable (Fig. 5G), indicating that the disproportionate distribution of $CD4^+$ and $CD8^+$ T cells is not the cause of the observed variation between clusters 2 and 4. Overall, the transcriptome profile of CD4, CD8 or cells that could not be assigned as CD4 or CD8 T cells (NA) within each cluster did not vary except for *CD8A* and *CD8B* transcripts (Fig. 5H-I). Transcripts such as *PRF1* and *NKG7* that showed comparable expression between cluster 2 and 4, followed the same pattern even when these two clusters were further classified based on the CD4 and CD8 T cell sub-type. The transcripts that showed higher expression in cluster 2 (*XCL1*, *ZEB2*, *CRTAM* and *GZMB*) and cluster 4 (*CCL5*, *GNLY*, *EOMES*), followed similar pattern of expression in CD4 and CD8 T cells within each of these effector and pre-effector clusters (Fig. 5H-I). Both CD4 and CD8 T cells from cluster 4 showed higher expression of IFN response genes *ISG20*, *IFI6* while those in cluster 2 showed higher expression of cytokines *TNF* and *IFNG* (Fig. 5H-I). Overall, we observed that the cells in clusters 2 and 4 have distinct gene expression pattern that is not influenced by either CD4 or CD8 T cell proportion (Fig. 5H-I). Overall, our data has revealed that both hCMV-reactive CD4-CTLs and CD8-CTLs share similar transcriptome profile.

Discussion

A significant world population carries the latent hCMV infection, without causing any disease. However, latent infection poses a huge risk of reactivation in immune-compromised individuals such as those undergoing transplant or infants born to CMV seropositive mothers. Although there is very little risk of the disease occurrence in immunocompetent individuals, the latent CMV infection puts a huge burden on the immune system especially the elderly, where immune cells are actively involved in maintaining the latency. Identifying and understanding the characteristics of the CMV-specific T cell immune memory components in individuals carrying virus in latent form, representing protective immune response, can be of high value to design therapy or vaccines to eradicate CMV.

A small number of antigen-specific memory T cells can have an enormous impact in directing or misdirecting an immune response upon secondary infection. Our understanding of *ex vivo*, antigen-specific human T cell responses to pathogens such as hCMV remains limited, due to the rarity of the cells, the complexity of the biology, and technical challenges of identifying overall antigen-specific $CD4^+$ memory T cells, since they can be extremely heterogenous in the expression of activation associated markers and cytokines. An added layer of complexity comes from HLA-restriction of individuals and the prior knowledge about the immuno-dominant peptide, limiting the use to pMHC multimers. Most studies have focused on isolating the total CD4 or CD8 T cells reactive to hCMV,

without considering the differences in memory proportions that can inadvertently bias the analysis towards analysing the predominant memory compartment within $CD4^+$ T cells (T_{CM} or T_{EM}) and $CD8^+$ T cells (T_{EMRA}). Our stimulation strategy that combined the kinetic analysis of activation markers across different memory subsets followed by high resolution single-cell RNA-Seq paired with single-cell TCR analysis allowed us to interrogate and compare a larger pool of hCMV (pp65 and IE-1)-reactive cells from different memory subsets across 10 seropositive and peptide-responsive donors. This unbiased approach enabled us to overlay the heterogeneity in T cell memory subsets classified based on developmental stages and functional subsets along with their TCR dependent activation markers. The expression of CD154 and CD137 in different set of cells that further differentially co-expressed one or more T cell activation associated markers such as CD69, CD25, OX40, warrants for careful selection of activation markers while analysing antigen-specific cells *ex vivo*. In this study, we observed that though activated (based on the expression of *CD154*, *CD137*, *CD69*, *CD25*, *OX40*), most cells except effector cells did not actively make cytokines and those that made also showed heterogeneity for the cytokine profile. Many studies previously have also relied on cytokine profiles to analyse *ex vivo* isolated antigen-specific cells and hence have mostly analysed effector memory T cells.

The single-cell RNA-Seq of a large number of hCMV reactive cells isolated from 10 different human subjects, identified several interesting functional subsets; T_H17, T_{REGs}, CD4-CTLs and CD8-CTLs. Of particular interest were CD4- and CD8-CTLs, that clustered together. The two cytotoxic clusters observed could not be distinguished based on the CD4- and CD8-specific transcripts *CD4* or *CD8A*/*CD8B*, but on the basis of other cytotoxic molecules that were linked to different states of effector functions (Omilusik *et al.*, 2015). The cluster that was predominantly making cytokine and chemokine (*IFNG*, *TNF*, *XCL1*, *XCL2*, *CCL3*, *CCL4*) also expressed transcription factor *ZEB2*, the TF known to be associated with late effector function (Omilusik *et al.*, 2015; Patil *et al.*, 2018). The cytotoxic cells in both cluster 4 and 0 appeared to respond to the cytokines and chemokines by expressing their receptors (*IFNGR1*, *CCR5*). However, the cytotoxic cells across these clusters expressed molecules linked to cytolytic granules (*PRF1*, *GZMB*, *GNLY*), and transcription factor *HOPX*. The differential expression of many of the molecules linked to cytotoxicity and effector functions, suggests the effector and pre-effector like subsets. These subsets, possibly at different developmental states, might exist to ensure a continuity in immune response during the entire course of the infection.

Another interesting discovery in our data set is the identification of T_{REGs} with suppressive ability that showed enrichment for the IFN response gene signatures which has been shown to be associated with a dampened T_H2 response in allergy (Seumois *et al.*, 2020). The enrichment of T_H2 signatures as well as the expression of *GATA3* and *CCR4*, the T_H2 associated transcription factor and chemokine receptor by the T_{REGs} further suggested that these could be suppressive towards the T_H2 subset. Although a very recent study in cancer has also identified this subset, to the best of our knowledge, presence of T_{REGs} enriched for IFN response genes has not been reported earlier in viral infections in humans (Chu *et al.*, 2023). Based on these observations, we hypothesize that during the

initial stages of the infection where the expansion of effector cells is necessary, T_{REGs} may not be executing their suppressive function towards these effectors, instead, suppressing other subtypes of no relevance to the disease. However, further kinetic studies in both animal model and human cohorts of different types of diseases such as viral and bacterial infections, allergy, cancers are necessary to verify or refute this hypothesis.

In conclusion, using single-cell multi-omics (transcriptome, TCR repertoire and expression of a few selected proteins) our study has identified an unprecedented amount of heterogeneity in hCMV-specific CD4+ T cell subsets. The knowledge base can serve as an important resource for designing vaccines and therapeutic strategies to eradicate CMV infections, especially in immunocompromised individuals.

Material and methods

Study subjects

Study subjects were healthy adult donors who donated blood at the Safdarjung Blood Bank during 2019-20. The donors consented for use of the buffy coat samples for research. Donors were HIV-negative and had no history of Hepatitis C infection. The median age was 30.5 (24-39) and all these subjects were male. Approval for the use of this material for research was obtained from the institutional human ethics committees from both National Institute of Immunology and VMMC and Safdarjung Hospital, New Delhi, India.

PBMC processing

Peripheral Blood Mononuclear Cells (PBMCs) were isolated from buffy coat samples from healthy blood donors obtained from Safdarjung hospital blood bank by density gradient centrifugation using Ficoll-Paque Premium (GE Healthcare Biosciences). PBMCs were cryopreserved in 90% fetal bovine serum (FBS) supplemented with 10% di-methyl sulfoxide (DMSO).

CMV IgG analysis

ELISA (Enzyme Linked Immunosorbent Assay) was performed on the plasma samples from the donors obtained from the blood bank using Human Anti-Cytomegalovirus IgG ELISA kit (abcam, #ab108724) according to the manufacturer's protocol to check for the donors' sero-status. Donors with the absorbance values at least10% higher than that of the cut-off control supplied in the kit were considered positive.

PBMC stimulation

hCMV-specific PepTivators (Miltenyi Biotec) derived from pp65 and IE-1 or αCD3/CD28 DynaBeads (Invitrogen) were used to stimulate PBMCs from CMV-seropositive donors, according to

the manufacturer's protocol. Briefly, PBMCs were thawed and treated with 50 U/mL Benzonase (Sigma) in RPMI before resting overnight in serum-free TexMACS medium (Miltenyi Biotec) supplemented with 1% Penicillin/Streptomycin at 37°C in a 96-well plate. Cells were stimulated at a density of 10 million/mL either by the addition of hCMV peptide pools (pp65 and IE-1 PepTivators) at a final concentration of 1 µg/mL or with 1:1 ratio of αCD3/CD28 DynaBeads Human T cell activator (Invitrogen), or water (as unstimulated control) in the presence of a blocking CD40 antibody (1 µg/mL; Miltenyi Biotec) and CD28 antibody (1 µg/mL; eBioscience). After 18 hours of activation, fresh rested, either CTV stained or unstained PBMCs (250,000 per well) were carefully added to the stimulated wells without disturbing the stimulated cells in the well. At the end of 24 hours (24 hours for the initially activated cells and 6 hours for the freshly added cells), cells were washed and stained with cocktail of antibodies (Table S2A) and analysed by flow cytometry. All flow cytometry data was analysed using FlowJo software (v10.8).

For single-cell RNA-seq using 10X genomics, PBMCs from 10 donors that showed highest response to peptide stimulation in preliminary experiments, were stimulated with pp65 and IE-1 PepTivators for 6 and 24 hrs. The stimulated cells from each donor were washed and individually stained with Cell-hashtag (HTO:Hash Tag Oligo) TotalSeq C antibody (Biolegend) (0.5 µg/condition, Table S2B) as per manufacturer's recommendations for 30 minutes on ice. The stained and washed PBMCs from all 10 donors were then pooled and stained with oligo-tagged TotalSeq C antibody for cell surface proteins for T cell memory markers (CD45RA, CCR7, CD95 and IL7Rα) (Biolegend) (0.5 µg/condition) (Table S2B) for 30 minutes on ice. Following washing, the pooled cells were further stained with fluorescence-conjugated surface antibodies to enable FACS sorting of antigen-specific cells (Table S2A). Cells were sorted using FACSAria Fusion Cell Sorter (Becton Dickinson) in LoBind 1.5 mL microcentrifuge (Eppendorf) tubes with 50% FBS supplemented with recombinant RNase inhibitor (1:100, Takara). The live singlet gated $CD3^+$ T cells were further gated as per the gating strategy shown in Fig. 1B.

Single-cell RNA-Seq

The sorted antigen-specific cells from various memory compartments were pooled in desired ratios and used for single-cell RNA-Seq and TCR-Seq assay using 10X Genomics kit, 5′ V(D)J with gene expression and cell surface protein expression v1.1 as per the manufacturer's recommendations. Initial amplification of cDNA libraries and final libraries were performed for 15 cycles and 14 cycles respectively for gene expression library. V(D)J and cell surface protein libraries were generated corresponding to each 5′ gene expression library using 9 cycles of amplification. Libraries were sequenced on Illumina NovaSeq6000 sequencing platform for PE150 reads.

Single-cell RNA-Seq analysis

Sequenced reads from the 5′ single-cell gene expression and associated V(D)J library were mapped to the GRCh38 reference genome using 10X genomics' cellranger multi (v6.1.1). The cell surface protein quantification was run as part of the same pipeline. The Seurat package (v4.3) in R (v4.2.2) was used to perform all further downstream analyses (Hao *et al*, 2021).

For downstream analysis, only cells that passed the QCs were retained; cells with UMI counts >500, number of genes >350 and percent mitochondrial genes <15% were considered as good quality cells. Further, cells were assigned as doublets, singlets or negatives based on the HTO counts using Demux with default parameters in Seurat package (v4.3) (Stoeckius *et al*, 2018). Singlets were assigned to their respective donor based on HTO count. Doublets were removed from downstream analysis. After applying all the quality control parameters, we retained 6899 cells out of 8062 cells for downstream analysis (85.57% cells passed the quality parameters). Cells from experiment 1 (3 donors) and experiment 2 (10 donors) were combined using Seurat package after applying the QCs (Table S1).

For each cell, normalized UMI counts for each gene were obtained by dividing the raw count by the total count across all genes. This was multiplied by a scaling factor of 10,000 and log transformed to yield the final normalized values. Prior to clustering, we filtered genes by selecting for the top 2000 with highest variance across cells, using *FindVariableFeatures* with *selection.method*="*vst*". This data was then centered and scaled using *ScaleData*, and the top 50 principal components were calculated. Top 18 PCs were selected using *ElbowPlot* and *JackStrawPlot* for downstream processing. This was used to build a shared nearest neighbour graph using *FindNeighbors*. The cells were clustered using *FindClusters*, with *algorithm=3* and *resolution=0.2*, determined using *clustree* function. For visualization and clustering, a 2D UMAP projection of the top 18 PCs was obtained using the *RunUMAP* function, with *metric="euclidean"* (Hao *et al.*, 2021).

The cell surface protein tag data was normalized using the *NormalizeData* function in Seurat with *normalization.method*="CLR". Briefly, for each cell, the hashtag counts were divided by the geometric mean of counts of all unique hashtags prior to log transformation.

Differential gene expression analysis

We identified differentially expressed genes using the MAST algorithm integrated within Seurat's *FindAllMarkers* function using default parameters on $\log_2$(CPM+1) counts (Hao *et al.*, 2021). Genes in a cluster were reported as differentially expressed compared to every other cluster if their $\log_2$ fold change of expression was > 0.25 and BH-adjusted P-value was < 0.05.

Gene module signature scores

Gene module scores were computed using the *AddModuleScore* function in Seurat, with default parameters (Hao *et al.*, 2021).

Trajectory inference

Single cell trajectories were computed using the slingshot algorithm within the dynverse/dyno (v0.1.2) suite of trajectory inference pipelines (Street *et al*, 2018). Raw and normalized UMI counts from cells of clusters 0, 2, 3, 4 and 7 were combined using the dynverse's wrap_expressions function. Cells were assigned groups by cluster and the number of expected start and end states were set to 1 using *add_prior_information*. Trajectory inference was run using *infer_trajectories* with default parameters, with method = ti_slingshot(). We did not specify any start or end cells for the trajectory.

Pathway enrichment analysis

DE genes between clusters 2 and 4 were fed to Ingenuity Pathway Analysis software from Qiagen. Enrichment pathways were sourced from Ingenuity core enrichment pathways.

TCR analysis

For each cell, all information about clonality, expressed TCR chains, their sequences and corresponding UMI counts were obtained as part of the standard 10X Genomics' cellranger multi pipeline (Table S5).

Assigning CD4 and CD8 T cells

For cells in cluster 2 and 4, we assigned cells as CD4 T cells, or CD8 T cells, first based on the *CD4* and *CD8A*/*CDB* transcript counts. Cells with > 0 counts for *CD4*, were assigned as $CD4^+$ T cell (333 cells). For CD8 T cells, we considered the cell as $CD8^+$ T cell (735 cells) if the sum of all three CD8 transcripts (*CD8A*, *CD8B*, *CD8B2*) was > 0 and *CD4* transcripts were undetectable. Cells (556) which could not be assigned as either CD4 or CD8 T cell based on this criterion, were further assigned as CD4- (274) and CD8- (12) T cell if they shared the clonotypes with already assigned CD4 or CD8 T cell. Those cells that could not be assigned to either CD4 or CD8 category based on either of these criteria, were called as NA (not assigned) cells (294). Based on these criteria, we had a total of 328 CD4 cells, 536 CD8 cells and 149 NA cells in cluster 2 (total 1013 cells) and 279 CD4 cells, 199 CD8 cells and 145 NA cells in cluster 5 (total 623 cells).

Statistical tests

All the statistical tests were done using GraphPad Prism v9.0. Mann Whitney u-test was performed on pair-wise comparison and *P*-value <0.05 was considered significant.

Data and Code availability

Scripts will be made available on GitHub (https://github.com/..). Sequencing data for this study will be deposited onto the Gene Expression Omnibus.

Figure Legends

Figure 1. Memory subset-specific expression of CD154 and CD137 upon TCR dependent T cell activation. (A) Schematic outline of the stimulation experiment. (B) Contour plots showing the gating strategy used to analyse activated T cells in the indicated category. (C) Bar graph shows the percentage of $CD137^+$ or $CD154^+$ cells in CD4 and CD8 T cell compartments at 6 hrs (top) and 24 hrs (bottom) after stimulation with αCD3/CD28 DynaBeads. (D) Bar graph shows the percentage of $CD137^+$ or $CD154^+$ cells in the indicated memory subset within CD4 and CD8 T cell compartments at 6 hrs (top) and 24 hrs (bottom) after stimulation with αCD3/CD28 DynaBeads. (E) Bar graph shows the percentage of $CD137^+$ or $CD154^+$ cells in the indicated memory subset within CD4 and CD8 T cell compartments after stimulation with αCD3/CD28 DynaBeads at both 6 hrs and 24 hrs (combined). (F) Bar graph shows the percentage of $CD154^+$ or $CD137^+$ activated cells in the indicated memory subset within CD4 and CD8 T cell compartments at both 6 hrs and 24 hrs (combined) after stimulation with hCMV peptide pools. (G) Bar graph shows the percentage of total activated cells ($CD154^+$ and/or $CD137^+$; combined) in the indicated memory subset within CD4 and CD8 T cell compartments at both 6 hrs and 24 hrs (combined) after stimulation with hCMV peptide pools. Each dot in the bar graphs (data points) represents a human donor. Data are mean ± SEM; significance was computed using Mann Whitney u-test; $*p<0.05$, $**p<0.005$, $***p<0.001$.

Figure 2. Single-cell transcriptome analysis of hCMV-reactive memory T cell subsets. (A) Schematic showing the study overview of isolating and single-cell RNA-seq of hCMV-specific T cells using 10X Genomics platform. Histograms (on the right) shows cells isolated using FACS for single-cell RNA-Seq from indicated T cell memory compartment from 10 donors. (B) 2D UMAP of single-cell RNA-Seq data from 6899 cells, using 2000 most variable transcripts, shows 9 distinct clusters of cells of hCMV-specific T cells. (C) 2D UMAP plots (top) and violin plots (bottom) showing expression of the indicated protein (oligo-tagged antibodies; adt). Colour scale shows expression levels of individual cells (UMAP) or mean expression in each cluster (violin plot). Only cells (6067) obtained from experiment 2 with 10 donors are shown. (D) Dot plot shows the mean expression (colour) and percent cells expressing (size) for selected marker gene transcripts in each cluster. (E) 2D UMAP plots of single-cell RNA-Seq data show the expression of the *CD4*, *CD8A* and *CD8B* transcripts. Scale represents expression level of the given transcript. Cells with zero counts for the indicated transcript are shown in grey colour. (F) 2D UMAP plots show the T_{CM}, T_{EM} and T_{EMRA} signature score for each cell. (G) 2D UMAP plots (left) of single-cell RNA-Seq data show the expression of the *CD154* (top) and *CD137* (bottom) transcripts. Scale represents expression level of the given transcript. Cells with zero counts for the indicated transcripts are shown in grey colour. Scatter plots (right) show the co-expression matrix for *CD154* (on X-axis), and *CD137* (on Y-axis) with *CD69* or *TNFRSF4* (colour scale: mean expression) transcripts. For violin plots in (C) and dot plots in (D) cluster with less than

1% of the total cells is not shown (cluster 8=56 cells). All expression counts for transcripts are $\log_2$ normalized (CPM+1).

Figure 3. hCMV-reactive T cell subtypes and their TCR clonal diversity. (A) Heatmap shows the expression of the most significantly enriched transcripts in each cluster. The top 200 transcripts are shown based on adjusted *P* value < 0.05, $\log_2$ fold change > 0.25 and > 10% difference in the percentage of cells expressing selected transcript between two groups of cells compared. Each column represents average expression of all cells for a given cluster. Colour scale represents z-score for the expression of each transcript across clusters. (B) Dot plot shows the mean expression (colour) and percentage of expressing cells (size) for selected marker gene transcript in each cluster. (C) 2D UMAP plots show the IFN response, T_{REG}, cytotoxicity and cell cycle signature score for each cell. (D) Violin plots show the normalized expression level ($\log_2$(CPM+1)) of the indicated differentially expressed transcript. Colour scale shows percentage of cells expressing the indicated transcript for each cluster. (E) Bar graph shows the percentage of cells with expanded clonotypes (frequency≥2) across 9 clusters. (F) Bar graph shows the frequency of 25 most expanded clonotypes across different clusters. For B and D, clusters with <1% of the total cells are not shown (cluster 8=56 cells).

Figure 4. Single-cell transcriptome analysis of hCMV-reactive T_{REGs}. (A) 2D UMAP plots of single-cell RNA-Seq data show the expression of the *FOXP3*, *IL2RA*, *IKZF2* (left) and joint expression of *FOXP3* with *IL2RA* and *FOXP3* with *IKZF2* (right 2 plots) transcripts. Scale represents $\log_2$ normalized (CPM+1) counts for expression values and scaled density for joint expression plots. Cells with zero counts for the indicated transcripts are shown in grey colour. (B) Bar graph shows the percentage of cells expressing *FOXP3* transcript ($\log_2$(CPM+1)>0) in clusters 1 and 5. (C) Dot plot shows the mean expression (colour) and percentage of expressing cells (size) for selected marker gene transcripts in indicated clusters. Clusters with <1% of the total cells are not shown (cluster 8=56 cells). (D) Scatter plots showing normalized co-expression level (log2(CPM+1)) between *FOXP3* and *GATA3* and *FOXP3* and *IL2RA* transcripts in the indicated clusters. (E) Bar graph shows the percentage of cells expressing *GATA3* or *IL1R2* transcripts ($\log_2$(CPM+1)>0) within *FOXP3* transcript expressing cells in clusters 1 and 5. (F) 2D UMAP plots show joint expression of *FOXP3 + GATA3*, *FOXP3 + CCR4* and *FOXP3 + IL1R2*, *FOXP3 + CCR8.*

Figure 5. Single-cell transcriptome analysis of hCMV-reactive CD4-CTLs and CD8-CTLs. (A) 2D UMAP plots of single-cell RNA-Seq data show the expression of the indicated transcripts. All expression counts for transcripts are $\log_2$ normalized (CPM+1). Cells with zero counts for the indicated transcripts are shown in grey colour. (B) Dot plot shows the mean expression (colour) and percentage of expressing cells (size) for selected marker gene transcripts in each cluster. (C) Violin plots showing the normalized expression level ($\log_2$(CPM+1)) of the indicated differentially expressed transcripts across 8 clusters. Colour scale shows percentage of cells expressing the indicated transcript for a given

cluster. (D) Trajectory analysis using slingshot shows the distribution of cells from indicated clusters (middle) on a pseudo-scale (left), and expression levels (scale) of the *GZMB*, *GNLY*, *CRTAM* and *KLRD1* transcripts. (E) 2D UMAP plot shows the cells with same clonotypes (colours) shared between cluster 3 and, clusters 2 and 4. (F) Pie chart shows the distribution of percentage of cells with top 20 expanded clonotypes in clusters 2 and 4. (G) Stacked bar graph shows the proportion of CD4, CD8 and NA (not assigned) T cells in each indicated cluster (refer to methods section for classification of cells). (H) Heatmap shows the normalized expression of 50 differentially expressed transcripts from each clusters 2 and 4 when compared to rest of the clusters. (I) Violin plots show the normalized expression level ($\log_2$(CPM+1)) of the indicated differentially expressed transcript across indicated T cell subsets in clusters 2 and 4. Colour scale shows percentage of cells expressing the indicated transcript. For B and C, clusters with <1% of the total cells are not shown (cluster 8=56 cells).

Supplementary figures.

Figure S1. hCMV reactivity across donors. (A) Graph shows the hCMV-specific IgG positivity (normalized absorbance) in 126 human donors. (B) Violin plot shows the normalized T cell reactivity to hCMV-specific peptide pools compared to unstimulated control. Donors above the horizontal dotted line for memory compartments were chosen for further single-cell genomics studies.

Figure S2. Single-cell RNA-Seq of hCMV-specific cells. (A) Stacked bar graph shows the distribution of cells from different donors across 9 clusters. (B) Stacked bar graph (left) and the 2D UMAPs (right) show the distribution of cells across 9 clusters from 2 different experiments/batch. Experiment 1 included 3 donors (D1, D2 and D7) and experiment 2 included all 10 donors. (C) Violin plots show the module enrichment score for T_{CM}, T_{EM}, T_{EMRA}, T_{CM}-T_{EM}, and T_{EM} -T_{EMRA} specific gene sets across 8 clusters. Colour scale indicates average module score. Clusters with <1% of the total cells are not shown (cluster 8=56 cells).

Figure S3. Functional characterization and TCR analysis of hCMV-specific T cells. (A) Violin plots show the module enrichment score for T_H1, T_H17, and $T_H1/17$ subset-specific gene sets across 8 clusters. Colour scale indicates average module score. Clusters with <1% of the total cells are not shown (cluster 8=56 cells). (B) UpSet plot shows the TCR clonotypes (frequency on y-axis) and sharing of the clonotypes (bottom, connected dots) across the 9 clusters. (C) Alluvial plot shows the clonally expanded cells (1234 cells) and their distribution and sharing across different clusters. Each connecting line represents a single cell.

Figure S4: hCMV-specific T_{REGs} show T_H2 characteristics. (A) Violin plot shows the module enrichment score for T_H2-specific gene set across 8 clusters. Colour scale indicates mean module score. Clusters with <1% of the total cells are not shown (cluster 8=56 cells)**.** (B) 2D UMAP plot shows joint

expression of T_H2-specific gene set and allergy induced T_{REG}-specific gene set. Scale represents density for joint expression plot.

Figure S5. Pathways enriched in hCMV-specific cytotoxic clusters. (A) Pathway enrichment analysis using Ingenuity Pathway Analysis (IPA) (Qiagen) shows enrichment of pathways in cluster 2 compared to cluster 4. (B) Pathway enrichment analysis using IPA shows enrichment of pathways in cluster 4 compared to cluster 2.

Supplementary tables.

Table S1. Details of cells recovered per cluster.

Table S2. List of reagents used in the study.

Table S3. Gene sets used for enrichment analysis.

Table S4. List of differentially expressed genes.

Table S5: TCR clonotypes of hCMV-specific T cells.

Acknowledgements

Funding: This work was supported by the DBT/Welcome Trust India Alliance Fellowship (grant number IA/I/18/2/504012) awarded to VSP and DBT NII core funding to VSP. RK is supported by DBT-JRF and NII for his PhD.

Author contribution: RK and VSP conceived, designed, and executed the study. RK and KS performed and analysed the experiments under the supervision of VSP, Genomics data was analysed by RK and SC under the supervision of GAA and VSP, AS supervised patient recruitment and sample collection, RK and VSP wrote the manuscript, which was further edited and approved by all authors.

Conflict: The authors declare no conflict of interest.

References

Almanan M, Raynor J, Sholl A, Wang M, Chougnet C, Cardin RD, Hildeman DA (2017) Tissue-specific control of latent CMV reactivation by regulatory T cells. *PLoS Pathog* 13: e1006507

Arlehamn CL, Seumois G, Gerasimova A, Huang C, Fu Z, Yue X, Sette A, Vijayanand P, Peters B (2014) Transcriptional profile of tuberculosis antigen-specific T cells reveals novel multifunctional features. *J Immunol* 193: 2931-2940

Awwad MHS, Mahmoud A, Bruns H, Echchannaoui H, Kriegsmann K, Lutz R, Raab MS, Bertsch U, Munder M, Jauch A *et al* (2021) Selective elimination of immunosuppressive T cells in patients with multiple myeloma. *Leukemia* 35: 2602-2615

Bahr F, Wehner R, Platzbecker U, Wermke M, Shayegi N, Middeke JM, Rollig C, Schetelig J, Ehninger G, Schmitz M *et al* (2013) Reconstitution of interleukin-17-producing T helper cells after allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant* 19: 357-365

Caccamo N, Milano S, Di Sano C, Cigna D, Ivanyi J, Krensky AM, Dieli F, Salerno A (2002) Identification of epitopes of Mycobacterium tuberculosis 16-kDa protein recognized by human leukocyte antigen-A*0201 CD8(+) T lymphocytes. *J Infect Dis* 186: 991-998

Cham CM, Gajewski TF (2005) Glucose availability regulates IFN-gamma production and p70S6 kinase activation in CD8+ effector T cells. *J Immunol* 174: 4670-4677

Chaudhuri L, Srivastava RK, Kos F, Shrikant PA (2016) Uncoupling protein 2 regulates metabolic reprogramming and fate of antigen-stimulated CD8+ T cells. *Cancer Immunol Immunother* 65: 869-874

Cheroutre H, Husain MM (2013) CD4 CTL: living up to the challenge. *Semin Immunol* 25: 273-281

Chu Y, Dai E, Li Y, Han G, Pei G, Ingram DR, Thakkar K, Qin JJ, Dang M, Le X *et al* (2023) Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. *Nat Med doi: 10.1038/s41591-023-02371-y*

Correa-Rocha R, Lopez-Abente J, Gutierrez C, Perez-Fernandez VA, Prieto-Sanchez A, Moreno-Guillen S, Munoz-Fernandez MA, Pion M (2018) CD72/CD100 and PD-1/PD-L1 markers are increased on T and B cells in HIV-1+ viremic individuals, and CD72/CD100 axis is correlated with T-cell exhaustion. *PLoS One* 13: e0203419

Crotty S (2019) T Follicular Helper Cell Biology: A Decade of Discovery and Diseases. *Immunity* 50: 1132-1148

Deaglio S, Dwyer KM, Gao W, Friedman D, Usheva A, Erat A, Chen JF, Enjyoji K, Linden J, Oukka M *et al* (2007) Adenosine generation catalyzed by CD39 and CD73 expressed on regulatory T cells mediates immune suppression. *J Exp Med* 204: 1257-1265

Eberlein J, Davenport B, Nguyen TT, Victorino F, Jhun K, van der Heide V, Kuleshov M, Ma'ayan A, Kedl R, Homann D (2020) Chemokine Signatures of Pathogen-Specific T Cells I: Effector T Cells. *J Immunol* 205: 2169-2187

Eschweiler S, Clarke J, Ramirez-Suastegui C, Panwar B, Madrigal A, Chee SJ, Karydis I, Woo E, Alzetani A, Elsheikh S *et al* (2021) Intratumoral follicular regulatory T cells curtail anti-PD-1 treatment efficacy. *Nat Immunol* 22: 1052-1063

Gandhi MK, Khanna R (2004) Human cytomegalovirus: clinical aspects, immune regulation, and emerging treatments. *Lancet Infect Dis* 4: 725-738

Gubser PM, Bantug GR, Razik L, Fischer M, Dimeloe S, Hoenger G, Durovic B, Jauch A, Hess C (2013) Rapid effector function of memory CD8+ T cells requires an immediate-early glycolytic switch. *Nat Immunol* 14: 1064-1072

Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M *et al* (2021) Integrated analysis of multimodal single-cell data. *Cell* 184: 3573-3587 e3529

Jayaratne HE, Wijeratne D, Fernando S, Kamaladasa A, Gomes L, Wijewickrama A, Ogg GS, Malavige GN (2018) Regulatory T-cells in acute dengue viral infection. *Immunology* 154: 89-97
Klenerman P, Oxenius A (2016) T cell responses to cytomegalovirus. *Nat Rev Immunol* 16: 367-377

Koch MA, Tucker-Heard G, Perdue NR, Killebrew JR, Urdahl KB, Campbell DJ (2009) The transcription factor T-bet controls regulatory T cell homeostasis and function during type 1 inflammation. *Nat Immunol* 10: 595-602

Kusnadi A, Ramirez-Suastegui C, Fajardo V, Chee SJ, Meckiff BJ, Simon H, Pelosi E, Seumois G, Ay F, Vijayanand P *et al* (2021) Severely ill COVID-19 patients display impaired exhaustion features in SARS-CoV-2-reactive CD8(+) T cells. *Sci Immunol* 6

Lim EY, Jackson SE, Wills MR (2020) The CD4+ T Cell Response to Human Cytomegalovirus in Healthy and Immunocompromised People. *Front Cell Infect Microbiol* 10: 202

Lyu M, Wang S, Gao K, Wang L, Zhu X, Liu Y, Wang M, Liu X, Li B, Tian L (2021) Dissecting the Landscape of Activated CMV-Stimulated CD4+ T Cells in Humans by Linking Single-Cell RNA-Seq With T-Cell Receptor Sequencing. *Front Immunol* 12: 779961

Mahnke YD, Brodie TM, Sallusto F, Roederer M, Lugli E (2013) The who's who of T-cell differentiation: human memory T-cell subsets. *Eur J Immunol* 43: 2797-2809

Meckiff BJ, Ramirez-Suastegui C, Fajardo V, Chee SJ, Kusnadi A, Simon H, Eschweiler S, Grifoni A, Pelosi E, Weiskopf D *et al* (2020) Imbalance of Regulatory and Cytotoxic SARS-CoV-2-Reactive CD4(+) T Cells in COVID-19. *Cell* 183: 1340-1353 e1316

Nowak A, Lock D, Bacher P, Hohnstein T, Vogt K, Gottfreund J, Giehr P, Polansky JK, Sawitzki B, Kaiser A *et al* (2018) CD137+CD154- Expression As a Regulatory T Cell (Treg)-Specific Activation Signature for Identification and Sorting of Stable Human Tregs from In Vitro Expansion Cultures. *Front Immunol* 9: 199

O'Shea JJ, Paul WE (2010) Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells. *Science* 327: 1098-1102

Omilusik KD, Best JA, Yu B, Goossens S, Weidemann A, Nguyen JV, Seuntjens E, Stryjewska A, Zweier C, Roychoudhuri R *et al* (2015) Transcriptional repressor ZEB2 promotes terminal differentiation of CD8+ effector and memory T cell populations during infection. *J Exp Med* 212: 2027-2039

Patil VS, Madrigal A, Schmiedel BJ, Clarke J, O'Rourke P, de Silva AD, Harris E, Peters B, Seumois G, Weiskopf D *et al* (2018) Precursors of human CD4(+) cytotoxic T lymphocytes identified by single-cell transcriptome analysis. *Sci Immunol* 3

Pena SV, Hanson DA, Carr BA, Goralski TJ, Krensky AM (1997) Processing, subcellular localization, and function of 519 (granulysin), a human late T cell activation molecule with homology to small, lytic, granule proteins. *J Immunol* 158: 2680-2688

Pena SV, Krensky AM (1997) Granulysin, a new human cytolytic granule-associated protein with possible involvement in cell-mediated cytotoxicity. *Semin Immunol* 9: 117-125

Poloni C, Schonhofer C, Ivison S, Levings MK, Steiner TS, Cook L (2023) T-cell activation-induced marker assays in health and disease. *Immunol Cell Biol*

Raffin C, Pignon P, Celse C, Debien E, Valmori D, Ayyoub M (2013) Human memory Helios-FOXP3+ regulatory T cells (Tregs) encompass induced Tregs that express Aiolos and respond to IL-1beta by downregulating their suppressor functions. *J Immunol* 191: 4619-4627

Rangel Rivera GO, Knochelmann HM, Dwyer CJ, Smith AS, Wyatt MM, Rivera-Reyes AM, Thaxton JE, Paulos CM (2021) Fundamentals of T Cell Metabolism and Strategies to Enhance Cancer Immunotherapy. *Front Immunol* 12: 645242

Sakaguchi S, Sakaguchi N, Asano M, Itoh M, Toda M (1995) Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J Immunol* 155: 1151-1164

Sallusto F (2016) Heterogeneity of Human CD4(+) T Cells Against Microbes. *Annu Rev Immunol* 34: 317-334

Scheinecker C, Goschl L, Bonelli M (2020) Treg cells in health and autoimmune diseases: New insights from single cell analysis. *J Autoimmun* 110: 102376

Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, Ha B, Altay G, Greenbaum JA, McVicker G *et al* (2018) Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell* 175: 1701-1715 e1716

Schober K, Voit F, Grassmann S, Muller TR, Eggert J, Jarosch S, Weissbrich B, Hoffmann P, Borkner L, Nio E *et al* (2020) Reverse TCR repertoire evolution toward dominant low-affinity clones during chronic CMV infection. *Nat Immunol* 21: 434-441

Seddiki N, Santner-Nanan B, Martinson J, Zaunders J, Sasson S, Landay A, Solomon M, Selby W, Alexander SI, Nanan R *et al* (2006) Expression of interleukin (IL)-2 and IL-7 receptors discriminates between human regulatory and activated T cells. *J Exp Med* 203: 1693-1700

Sefik E, Geva-Zatorsky N, Oh S, Konnikova L, Zemmour D, McGuire AM, Burzyn D, Ortiz-Lopez A, Lobera M, Yang J *et al* (2015) MUCOSAL IMMUNOLOGY. Individual intestinal symbionts induce a distinct population of RORgamma(+) regulatory T cells. *Science* 349: 993-997

Seumois G, Ramirez-Suastegui C, Schmiedel BJ, Liang S, Peters B, Sette A, Vijayanand P (2020) Single-cell transcriptomic analysis of allergen-specific T cells in allergy and asthma. *Sci Immunol* 5
Shevach EM (2009) Mechanisms of foxp3+ T regulatory cell-mediated suppression. *Immunity* 30: 636-645

Stegelmann F, Bastian M, Swoboda K, Bhat R, Kiessler V, Krensky AM, Roellinghoff M, Modlin RL, Stenger S (2005) Coordinate expression of CC chemokine ligand 5, granulysin, and perforin in CD8+ T cells provides a host defense mechanism against Mycobacterium tuberculosis. *J Immunol* 175: 7474-7483

Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung BZ, Mauck WM, 3rd, Smibert P, Satija R (2018) Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol* 19: 224

Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, Purdom E, Dudoit S (2018) Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* 19: 477
Sumida H, Cyster JG (2018) G-Protein Coupled Receptor 18 Contributes to Establishment of the CD8 Effector T Cell Compartment. *Front Immunol* 9: 660

Tovar-Salazar A, Weinberg A (2020) Understanding the mechanism of action of cytomegalovirus-induced regulatory T cells. *Virology* 547: 1-6

van Leeuwen EM, Remmerswaal EB, Vossen MT, Rowshani AT, Wertheim-van Dillen PM, van Lier RA, ten Berge IJ (2004) Emergence of a CD4+CD28- granzyme B+, cytomegalovirus-specific T cell subset after recovery of primary cytomegalovirus infection. *J Immunol* 173: 1834-1841

Velaga S, Ukena SN, Hopting M, Ivanyi P, Borchers S, Mischak-Weissinger EM, Hamwi I, Buchholz S, Ganser A, Franzke A (2013) Reconstitution and phenotype of Tregs in CMV reactivating patients following allogeneic hematopoietic stem cell transplantation. *Immunol Invest* 42: 18-35

Wagner L, Yang OO, Garcia-Zepeda EA, Ge Y, Kalams SA, Walker BD, Pasternack MS, Luster AD (1998) Beta-chemokines are released from HIV-1-specific cytolytic T-cell granules complexed to proteoglycans. *Nature* 391: 908-911

Wang Y, Su MA, Wan YY (2011) An essential role of the transcription factor GATA-3 for the function of regulatory T cells. *Immunity* 35: 337-348

Weiskopf D, Bangs DJ, Sidney J, Kolla RV, De Silva AD, de Silva AM, Crotty S, Peters B, Sette A (2015) 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

Xie MM, Fang S, Chen Q, Liu H, Wan J, Dent AL (2019) Follicular regulatory T cells inhibit the development of granzyme B-expressing follicular helper T cells. *JCI Insight* 4

Zhang L, Zhou R, Zhang W, Yao X, Li W, Xu L, Sun X, Zhao L (2019) Cysteine-rich intestinal protein 1 suppresses apoptosis and chemosensitivity to 5-fluorouracil in colorectal cancer through ubiquitin-mediated Fas degradation. *J Exp Clin Cancer Res* 38: 120

Zheng Y, Chaudhry A, Kas A, deRoos P, Kim JM, Chu TT, Corcoran L, Treuting P, Klein U, Rudensky AY (2009) Regulatory T-cell suppressor program co-opts transcription factor IRF4 to control T(H)2 responses. *Nature* 458: 351-356

Figure. 1

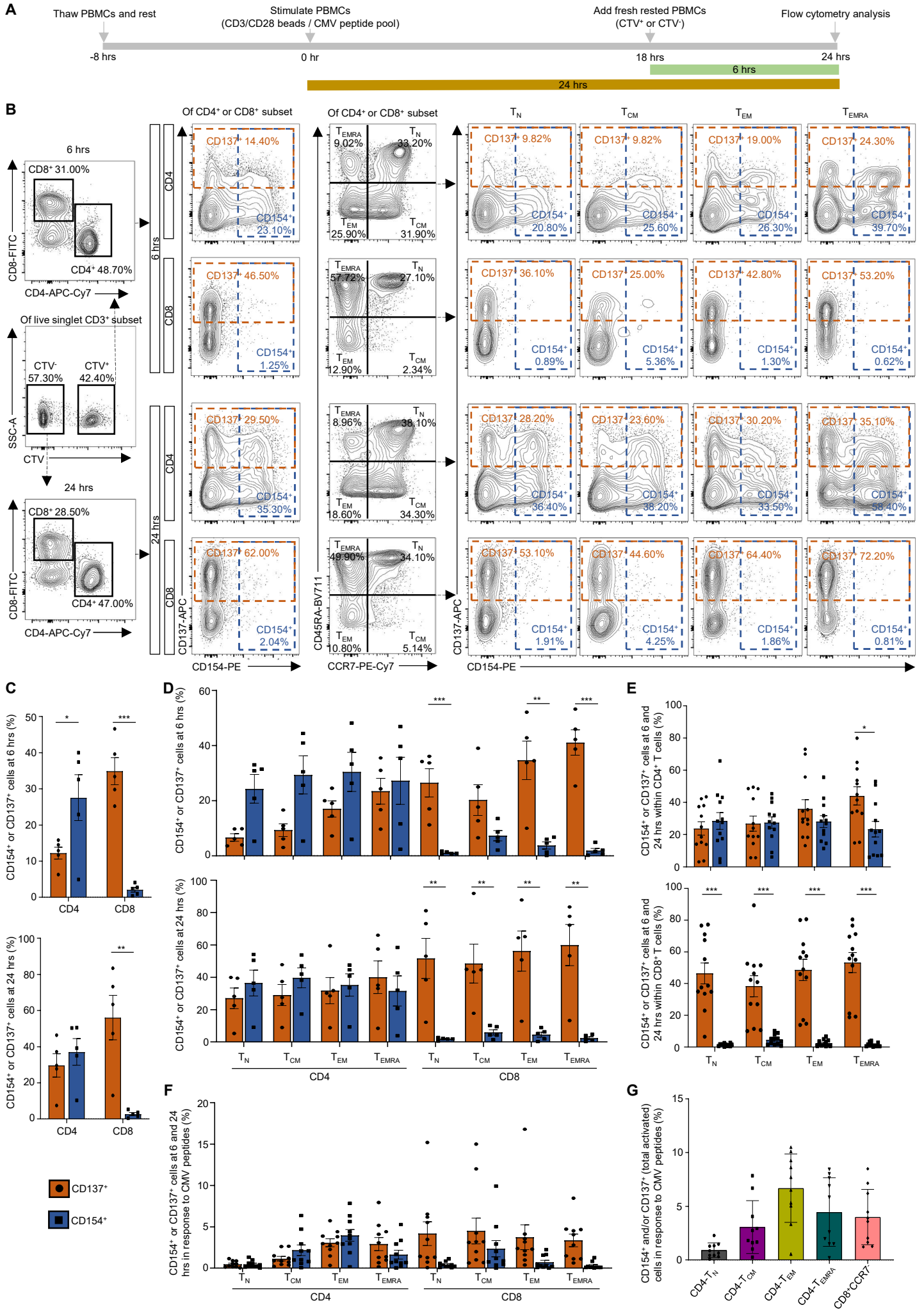

A
Thaw PBMCs and rest
Stimulate PBMCs
(CD3/CD28 beads / CMV peptide pool)
Add fresh rested PBMCs
(CTV+ or CTV-)
Flow cytometry analysis
-8 hrs
0 hr
18 hrs
24 hrs
6 hrs
24 hrs
B
Of CD4+ or CD8+ subset
Of CD4+ or CD8+ subset
TN
TCM
TEM
TEMRA
6 hrs
CD8+ 31.00%
CD4+ 48.70%
CD8-FITC
CD4-APC-Cy7
Of live singlet CD3+ subset
CTV- 57.30%
CTV+ 42.40%
SSC-A
CTV
24 hrs
CD8+ 28.50%
CD4+ 47.00%
CD4
CD8
CD137-APC
CD154-PE
CD45RA-BV711
CCR7-PE-Cy7
C
D
E
F
G
CD154+ or CD137+ cells at 6 hrs (%)
CD154+ or CD137+ cells at 24 hrs (%)
CD154+ or CD137+ cells at 6 and 24 hrs within CD4+ T cells (%)
CD154+ or CD137+ cells at 6 and 24 hrs within CD8+ T cells (%)
CD154+ or CD137+ cells at 6 and 24 hrs in response to CMV peptides (%)
CD154+ and/or CD137+ (total activated) cells in response to CMV peptides (%)
CD4-TN
CD4-TCM
CD4-TEM
CD4-TEMRA
CD8+CCR7-
CD137+
CD154+

Figure. 2

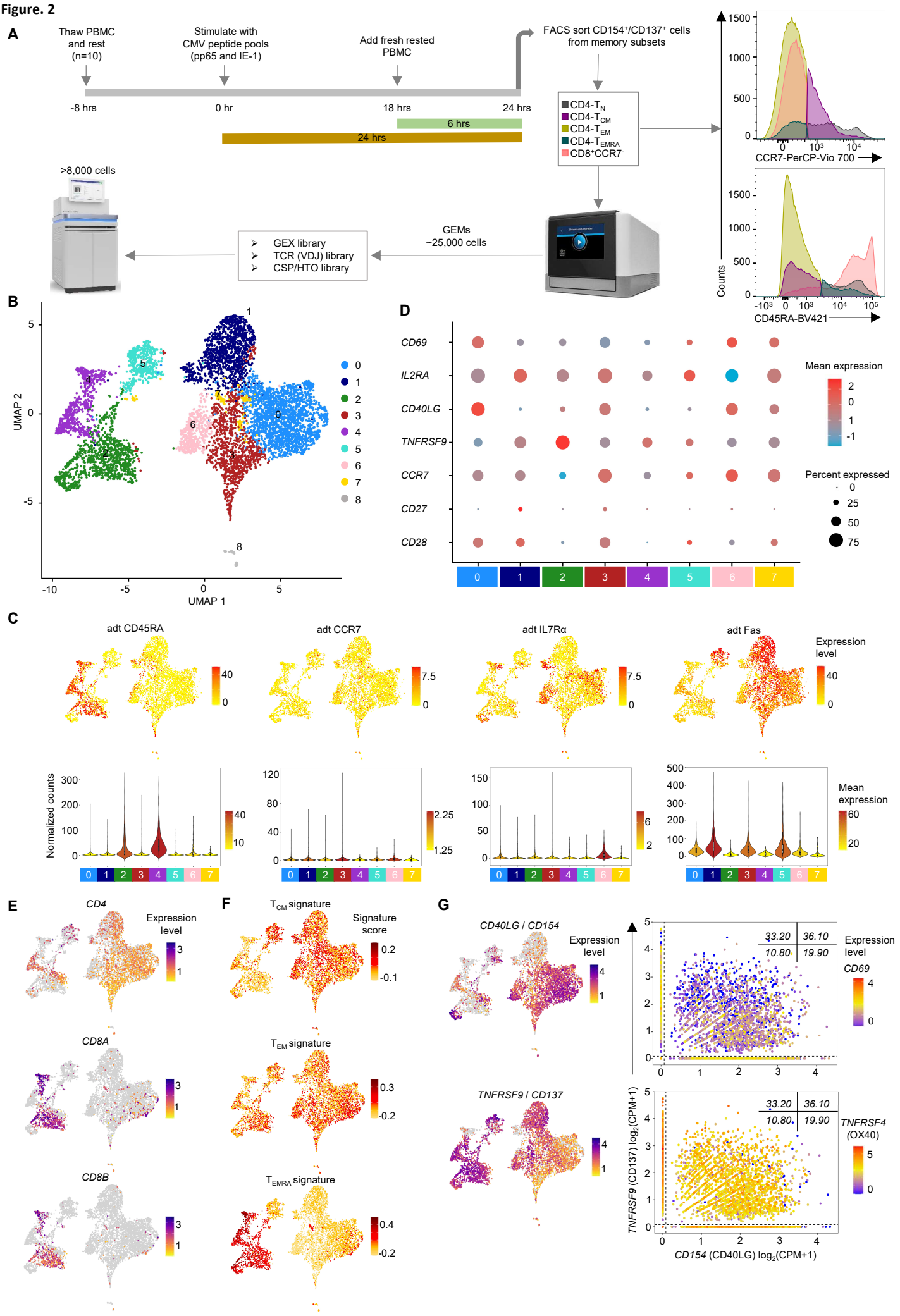

A
Thaw PBMC and rest (n=10)
Stimulate with CMV peptide pools (pp65 and IE-1)
Add fresh rested PBMC
FACS sort CD154+/CD137+ cells from memory subsets
-8 hrs
0 hr
18 hrs
24 hrs
6 hrs
24 hrs
CD4-TN
CD4-TCM
CD4-TEM
CD4-TEMRA
CD8+CCR7-
CCR7-PerCP-Vio 700
CD45RA-BV421
Counts
>8,000 cells
GEX library
TCR (VDJ) library
CSP/HTO library
GEMs ~25,000 cells
B
UMAP 1
UMAP 2
D
CD69
IL2RA
CD40LG
TNFRSF9
CCR7
CD27
CD28
Mean expression
Percent expressed
C
adt CD45RA
adt CCR7
adt IL7Rα
adt Fas
Expression level
Normalized counts
Mean expression
E
CD4
CD8A
CD8B
Expression level
F
TCM signature
TEM signature
TEMRA signature
Signature score
G
CD40LG / CD154
TNFRSF9 / CD137
Expression level
CD69
TNFRSF4 (OX40)
TNFRSF9 (CD137) log2(CPM+1)
CD154 (CD40LG) log2(CPM+1)
33.20
36.10
10.80
19.90

Figure. 3

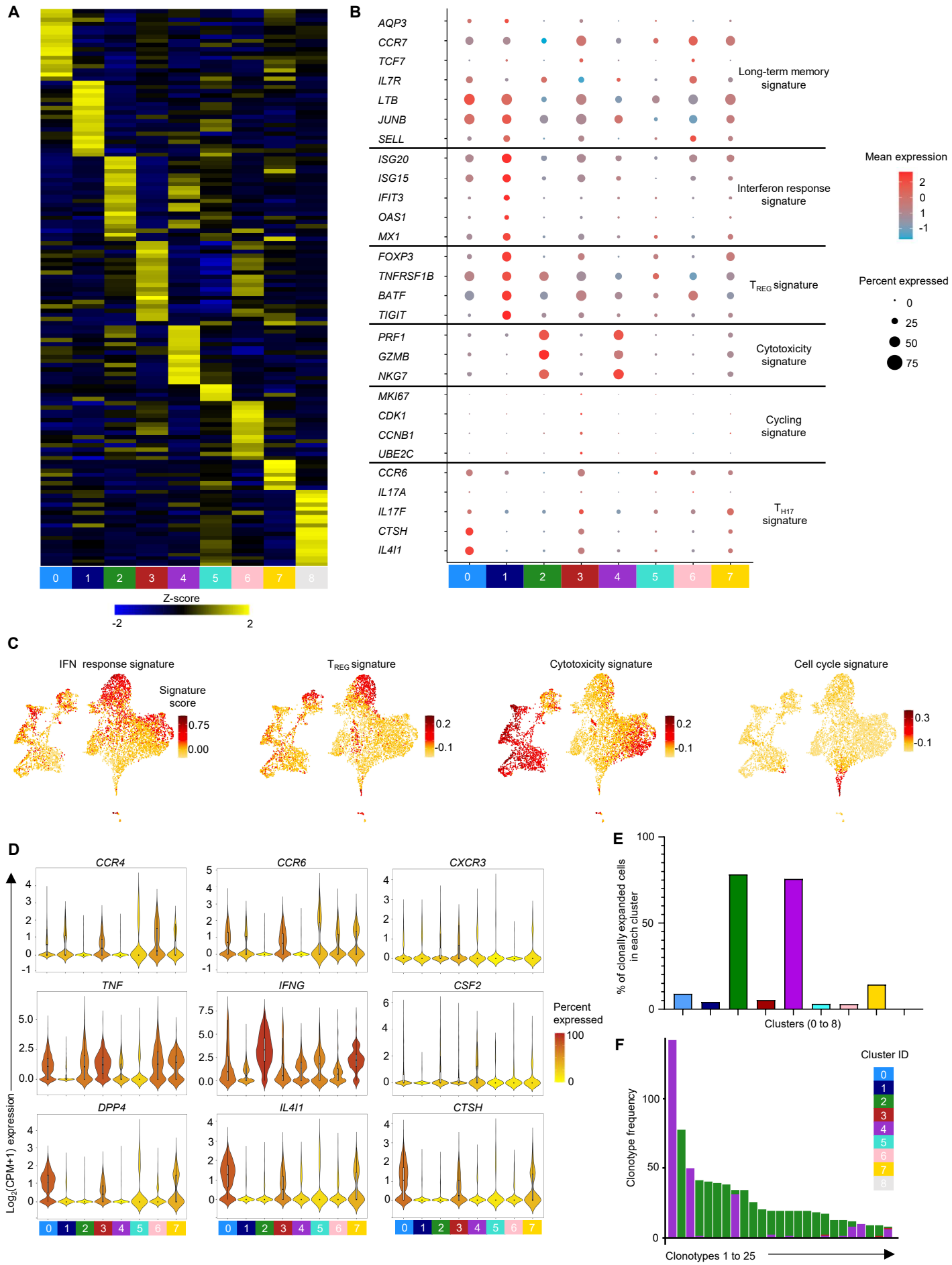

A
Z-score
-2
2
B
AQP3
CCR7
TCF7
IL7R
LTB
JUNB
SELL
Long-term memory signature
ISG20
ISG15
IFIT3
OAS1
MX1
Interferon response signature
FOXP3
TNFRSF1B
BATF
TIGIT
T_REG signature
PRF1
GZMB
NKG7
Cytotoxicity signature
MKI67
CDK1
CCNB1
UBE2C
Cycling signature
CCR6
IL17A
IL17F
CTSH
IL4I1
T_H17 signature
Mean expression
Percent expressed
C
IFN response signature
T_REG signature
Cytotoxicity signature
Cell cycle signature
Signature score
D
CCR4
CCR6
CXCR3
TNF
IFNG
CSF2
DPP4
IL4I1
CTSH
Log2(CPM+1) expression
Percent expressed
E
% of clonally expanded cells in each cluster
Clusters (0 to 8)
F
Clonotype frequency
Clonotypes 1 to 25
Cluster ID

Figure. 4

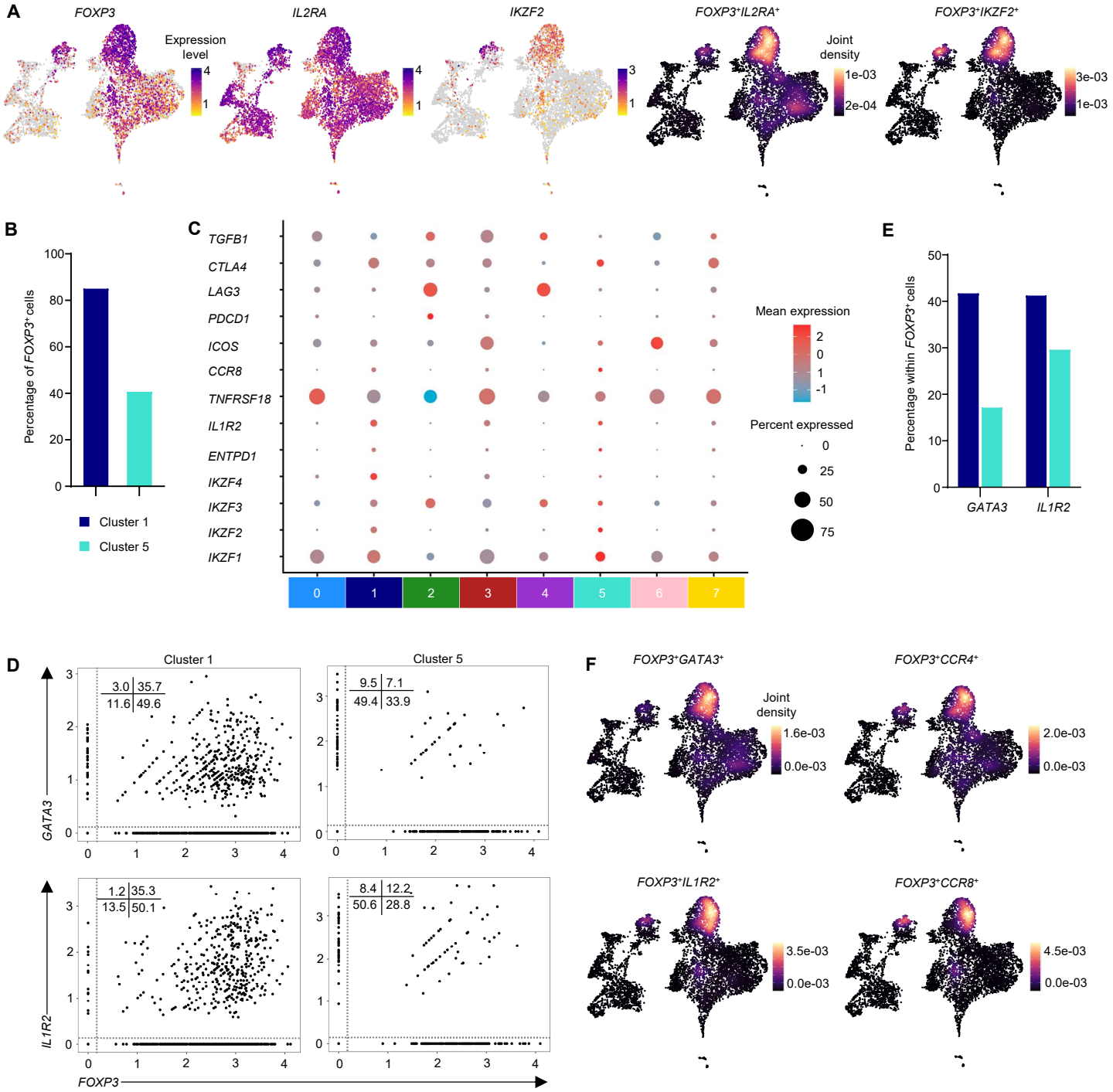
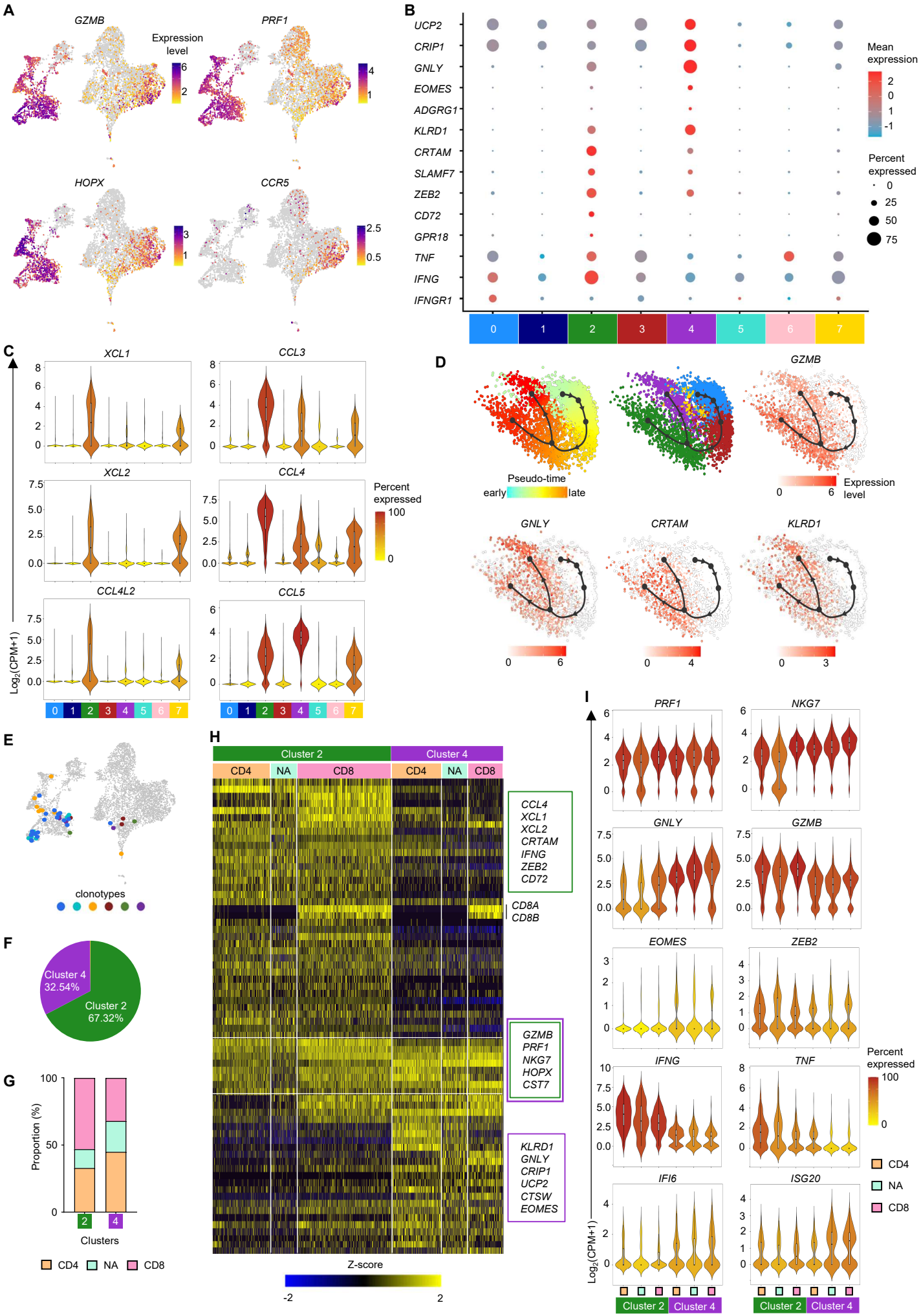

A
FOXP3
IL2RA
IKZF2
FOXP3+IL2RA+
FOXP3+IKZF2+
Expression level
Joint density
B
Percentage of FOXP3+ cells
Cluster 1
Cluster 5
C
TGFB1
CTLA4
LAG3
PDCD1
ICOS
CCR8
TNFRSF18
IL1R2
ENTPD1
IKZF4
IKZF3
IKZF2
IKZF1
Mean expression
Percent expressed
E
Percentage within FOXP3+ cells
GATA3
IL1R2
D
Cluster 1
Cluster 5
GATA3
IL1R2
FOXP3
F
FOXP3+GATA3+
FOXP3+CCR4+
FOXP3+IL1R2+
FOXP3+CCR8+
Joint density

Figure. 5


A
GZMB
PRF1
HOPX
CCR5
Expression level
B
UCP2
CRIP1
GNLY
EOMES
ADGRG1
KLRD1
CRTAM
SLAMF7
ZEB2
CD72
GPR18
TNF
IFNG
IFNGR1
Mean expression
Percent expressed
C
XCL1
CCL3
XCL2
CCL4
CCL4L2
CCL5
Percent expressed
Log2(CPM+1)
D
Pseudo-time
early
late
GZMB
Expression level
GNLY
CRTAM
KLRD1
E
clonotypes
F
Cluster 4 32.54%
Cluster 2 67.32%
G
Proportion (%)
Clusters
CD4
NA
CD8
H
Cluster 2
Cluster 4
CD4
NA
CD8
CCL4
XCL1
XCL2
CRTAM
IFNG
ZEB2
CD72
CD8A
CD8B
GZMB
PRF1
NKG7
HOPX
CST7
KLRD1
GNLY
CRIP1
UCP2
CTSW
EOMES
Z-score
I
PRF1
NKG7
GNLY
GZMB
EOMES
ZEB2
IFNG
TNF
IFI6
ISG20
Percent expressed
Log2(CPM+1)
Cluster 2
Cluster 4