bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Dissecting the landscape of activated CMV-stimulated CD4+ T cells in human by linking single-cell RNA-seq with T-cell receptor sequencing

Menghua Lyu[1,2*], Shiyu Wang[1,2*], Kai Gao[1,2], Longlong Wang[1,2], Bin Li[2,3], Lei Tian[1,2#]

[1] BGI Education Center, University of Chinese Academy of Sciences, Shenzhen 518083, China

[2] BGI-Shenzhen, Shenzhen, China 518083

[3] Shanghai Institute of Immunology, Shanghai JiaoTong University School of Medicine, Shanghai 200025, China

[*]These authors contributed equally to this work.

[#]Corresponding authors. Emails: LeiTian@genomics.cn, sci.tian@hotmail.com

## Abstract

CD4 T cell is crucial in CMV infection, but its role is still unclear during this process. Here, we present a single-cell RNA-seq together with T cell receptor (TCR) sequencing to screen the heterogenicity and potential function of CMV pp65 reactivated CD4+ T cell subsets from human peripheral blood, and unveil their potential interactions. Notably, Treg composed the major part of these reactivated cells. Treg gene expression data revealed multiple transcripts of both inflammatory and inhibitory functions. Additionally, we describe the detailed phenotypes of CMV-reactivated effector-memory (Tem), cytotoxic T (CTL), and naïve T cells at the single-cell resolution, and implied the direct derivation of CTL from naïve CD4+ T cells. By analyzing the TCR repertoire, we identified a clonality in stimulated Tem and CTLs, and a tight relationship of Tem and CTL showing a large share in TCR. This study provides clues for understanding the function of CD4+ T cells subsets and unveils their interaction in CMV infection, and may promote the development of CMV immunotherapy.

Key words: CMV pp65, Single-cell mRNA-seq, paired TCR-seq, CD4+ T cells

## Introduction

Cytomegaloviruses (CMV)/human herpesvirus 5 (HHV-5) infection is endemic in humans. Most

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

immunocompetent hosts show little to no clinical symptoms of primary infection and during persistent infection. Although its infection is regarded as asymptomatic, CMV hijacks the resources of the immune system throughout life by remaining latent and occasionally reactivate, eventually compromising on average approximately 10% of the entire T cell repertoire[1] and had a deleterious effect on immune senescence and health outcomes in the elderly[2]. In addition to this impact in healthy people, CMV infection can cause devastating consequences on immuno-compromised populations, such as the fetus and patients undergoing transplantation.

For these immunocompromised patients, reconstruction of CMV-specific T cells has emerged as an effective method to reduce CMV infection and reactivation. Data from patients who have received hematopoietic stem cell transplantations (HSCT) shows that the recovery from CMV disease correlates with the reconstitution of CD4+ and CD8+ T cell pools [3–5], and the recovery of CD4+ T cells is suggested as a prerequisite. The underlying mechanism may be that CMV-specific CD4+ T cells affiliate the expansion of CMV-specific CD8+ T cells, and lead a more effective clearance of serum virus compared to treatment with CD8+ T cells only[6]. Furthermore, infusion experiments with CD4+ T cells alone in immunocompromised mice is able to effectively repress CMV reactivation. These evidences suggest the key effector of CD4+ T cells in anti-CMV immunity, but CD4+ T is a heterogeneous group, and it is yet unclear of the function of these CD4+ T subsets in CMV infection and their interaction which hesitate the clinical application of adoptive immune therapy in CMV.

Therefore, studies on CD4+ T cell subsets have been performed respectively in past decades and reveal that CD4+ cytolytic cell (CD4-CTL), Treg and CD4+ memory T involve in the immune response to CMV infection in humans, nonhuman primates, and rodents. CD4-CTLs are firstly confirmed as a natural identity in chronic infections, such as LCMV, HBV, and CMV, and shows a strong antiviral effect in anti-CMV immunity by their helper functions and cytotoxicity. In terms of its helper function, CD4-CTL expressed cytokines, such as IFNγ and TNF[7], to promote the activation of CD8+ T cells, recruit innate immune cells including natural killer and monocytes to inflammatory sites, and directly inhibit virus replication[8]. As for their cytotoxicity, Fas/FasL pathway is utilized to mediate the death of infected B cells presenting viral epitopes with MHC-II[9,10]. Another cytotoxic mechanism used by CD4-CTL is via the perforin–granzyme

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pathway[11]. This pathway is based on the recognition of CTL to target cells in an MHC-II dependent manner[12], where the MHC-II is upregulated in epithelial cells upon CMV infection. Although there have been advances in understanding CD4-CTL function in CMV infection, their derivation is still unclear. Common views developed from other infection diseases indicate the origination of CD4-CTLs from effector cells. Recently, evidences[1314,15] from studies on transcriptome factors suggested these cells can also differentiate from naïve cells directly. This indicates that it may be possible to prepare CD4-CTL for immune therapy from naïve cells.

Although some studies tried to unveil Treg function in CMV infection, it is still controversial to conclude. In humans, Treg cells from CMV-seropositive individuals attenuated the proliferation of autologous CD8+ T cells and, to a lesser extent, CD4+ T cells in response to CMV virus ex vivo stimulation using PD-1 pathway[16]. However, in CMV reactivating patients received HSCT, CMV reactivation does not correlate with the numerical reconstitution of CD4+CD25highCD127- Tregs, and conventional T cells in these patients expressed high level of Ki67 indicating that their activation function is unimpaired[17]. Selectively removing Treg in animal models is a classical method to verify Treg function at infectious situations[18], and have been used to identify the negative function of Treg in some anti-viral immunities. However, these experiments failed to conclude Treg function in CMV infection. In murine, eight months after infection, Treg cells deletion decreased MCMV reaction in the spleen but enhance the reactivation in the salivary gland [19]. By analyzing the cell composition and cytokine secretion, the Treg deletion was found to expanded CD4+ T cells in these two tissues, excepting of an increased production of IL-10 in the salivary gland which is inferred to disturb the conventional T cells function in the salivary gland and be in charge of the enhancing viral replication.

T cell receptor (TCR)-specific signaling pathway is essential to generate an effective antiviral immunity. To identify T cells bearing these antigen-specific TCRs, labeling them with antibodies targeting IFN-γ and other markers, such as CD69, as well as Fluorescence-activated cell sorting (FACS) are used in previous studies. In the past decade, CD154 is found to be a TCR-singling specific marker, and labeling it only is able to identify cells which should be labeling with multiple markers[20–22]. Therefore, methods based on CD154 have been developed and employed by studies focusing on the antigen specificity of TCRs[23,24]

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To further unveil the potential function of CD4+ T cell subsets in CMV and understand their interactions, we performed a single-cell RNA and paired TCR sequencing on CMV pp65-specific CD4+ T cells from three healthy donors with a latent CMV infection. With a global view on CMV-specific CD4+ cells, we identified: 1) CMV-reactivated Treg cells accounted of a large proportion, and obtained a Th1 phenotype, enhanced migration ability and multiple inhibitory functions; 2) CD4-CTLs have a polyfunctional phenotypes; 3) CD4-CTL and effector memory T (Tem) experienced clonality and have a lager convergence in TCR repertoire; 4) a group of naïve cells expressed cytolytic factor. These findings exhibited the heterogeneity of CMV-reactivated CD4+ T cells, highlighted the balance between CMV-specific Treg and effector T cells, suggest that the composition of CD4+ T cells may be critical for adoptive T cells therapy for CMV. In summary, this study provides clues for understanding anti-CMV immunity and developing clinical therapy on CMV.

## Results

### CMV pp65-specific CD4+ T cells are characterized by typical antiviral profiles

Circulating antigen-specific T cells are rare in peripheral blood during the latent stage of CMV infection, comprising on average 10% of both the CD4 and CD8 memory compartments in blood[1]. To isolate CMV-specific CD4T cells, we cultured PBMCs with and without CMV-pp65 peptides for 24 hours and sorted CD3+CD154+ cells by flow cytometry[24–27]. We sorted CMV-specific CD4T cells from three donors and then pooled them together for single-cell mRNA-seq and paired VDJ-seq using the 10 × Chromium platform (Figure 1A). Control CD4T cells were acquired by lymphocyte sorting with FSC/SSC. PBMCs from the three donors were subject to bulk RNA-seq for subsequent SNP (single-nucleotide polymorphism) calling, and the sample identity of each cell was deconvoluted with these natural genetic variations[28].

After stringent quality control and filtering by multiple criteria, RNA-seq data were obtained from 2847 and 6493 single cells from the two libraries (CMV and control), detecting a mean of 3041 and 1947 genes per cell, respectively. Productive VDJ sequences were obtained for 1271 CMV cells and 3557 control cells. The unsupervised clustering of all cells from CMV and control samples suggested that there are multiple subsets of CMV CD4T cells (Figure1. B, C, D). CMV

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

stimulated CD4T cells and control CD4T cells that have both mRNA and VDJ data (CMV: 1200 cells, control: 1911 cells) were used for further analysis.

To reveal the potential function of overall CMV stimulated CD4+ T cells, we analyzed differentially expressed genes between CMV CD4T and control CD4T cells. CMV CD4+ T cells show a typical T cell activation profile including increased expression of *IL2RA*, *OX40 (TNFRSF4*), *MIR155HG*, *TNFRSF18*, *CD40LG* and *LGALS1*, and decreased expression of *IL7R* and *SELL*. These cells also express genes of inflammatory cytokines *IFNG* and *TNF*[29,30], T-bet independent IFN-γ production inducer *BHLHE40*[31], pro-inflammatory chemokine *CCL4*, and cytotoxic molecules *LTA* and *GZMB* (Figure 1E ). These results suggest that CMV CD4T cells are consist of several groups of activated multiple-cytokine-producing antiviral cells. This result is further confirmed by Gene Ontology (GO) analysis, where DEGs are significantly enriched in pathways such as T cell activation and response to tumor necrosis factors (Figure 1F). Consistent with previous reports using CD154 as a marker for antigen-specific CD4T cells, the cells we obtained here using the same strategy exhibit a typical activated anti-viral response.

Polyfunctionality profiles of CMV pp65-specific CD4 + T cell subsets

Subpopulations of CD4+ T cell were further identified with canonical markers (Supplementary Table2) and projected onto the UMAP embeddings (Figure 2A,2B). Stimulated CD4+ T cells were clustered into five groups: naïve (17.01% of total CMV CD4+ T cells), memory-like (11.20%), Treg (56.68%), effector memory (Tem, 3.65%), CTL (11.45%). Unstimulated CD4+ T cells were clustered into four subgroups: naïve (55.30%), memory-like (41.98%), Treg (1.78%), and CTL (0.94%) (Figure 2C). The ratio of unstimulated naïve and memory CD4+ T cells is consistent with previous FACS date[32], implying the classification using the scRNA-seq data in this study has a good consistence with that of FACS. Obviously, Treg, CTL, and Tem are significantly enriched in CMV, and Treg is the largest subset. This distribution is further identified in cells from each donor, although cells from donor 3 were limited (1.58% of total CD4+T cells) (Supplement Figure 1 A, 1B).

To investigate features of the five CD4+T cell subsets in CMV, we compared them with each other using the *FindAllMarkers* function. The Top10 DGEs were found to be different from each other,

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

141 indicating that these subsets may have distinct phenotypes (Figure 2D). Therefore, together with
142 the Top 10 DGEs and feature genes from literatures, we analyzed the phenotype of each subset.
143 We firstly analyzed the largest subset, Treg. These cells are *FOXP3+IFNG+TNF+*, and highly
144 expressed stable marker *SOCS1*, cytotoxic molecules (*LTA*, *LTB*), and a series of genes relate to
145 inhibition, such as *LGALS1*[33], *LGALS3*[34], *IL4I1*[35] and costimulator *CD70* (Figure 2E). Their
146 stability related cytokine receptors [36] (such as *IL2RA*, *TNFRSF18* [*GITR* ], *IFNGR2* ), and
147 inhibitory function related costimulatory molecules (such as *LAG3*, *CTLA4*, *TNFRSF4* [*OX40* ],
148 *TIGIT* ), are also abundantly expressed. The expression of chemokine receptor (*CCR4*, *CCR6*,
149 *CCR7* ) indicate their chemotaxis toward *CCL3*, *CCL5* (which is highly expressed by CTL and
150 Tem in our data) and homing ability to secondary lymphoid organs of Treg cells, the high
151 expression of *CCR6* and *CCL20* suggests that they can cluster in a self-sustaining positive
152 feedback loop.

153 Then, we analyzed CD4 Tem and CTL, since they exhibit similar expression profiles (Figure 2E).
154 Tem and CTL both expressed a high level of cytotoxic relative molecules (*GNLY*, *GZMB*, *GZMH*,
155 *CTSC*, *CST7*, *PRF1*, *NKG7*, *CTSB*, *FGFBP2*). Those similar expression of cytotoxic markers in
156 CTL and Tem indicate they may employ the same mechanism - the granule exocytosis pathway, to
157 initiate target cell apoptosis. This mechanism involves the regulated release of the contents of
158 cytotoxic granules ( such as *PRF1*, *GZMB*, *GZMH*, *GZMA*, *CTSC*, *GNLY*), into the immunological
159 synapse formed between the effector and target cell and kill them[37]. Besides, Tem expressed more
160 cytotoxic genes such as *LTB*, *GZMA*, *KLRB1*, and *KLRD1* than CTL, indicating the functional
161 spectrum of CD4 Tem is wider than that of CTL. CTL and Tem also abundantly expressed
162 chemokine genes (*CCL3*, *CCL4*, *CCL5*, *CCL3L3*, *CCL4L2* ), MHC  (*HLA-DPA1*, *HLA-DPB1*,
163 *HLA-DRB5*) and co-stimulators (*LAG3*, *CTLA4*, *OX40*, *PDCD1*), indicating they may attract
164 common targets to the inflammatory site and may kill them in another mechanism - MHC class
165 II-dependent fashion[38,39]. Previous studies reported that CTL may stem from Tem[23]. Subsequent
166 TCR repertoire analysis in our data also suggests that CTL may be originated from Tem.

167 CMV pp65 peptides pool exposed CD4+ naïve T cells in our data show obvious activation
168 characteristics. In total, 981 genes were differentially expressed (adjusted $p < 0.05$) upon
169 stimulation with the CMV peptides relative to control (Figure 3A; Supplement Table 3), of which

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

124 and 36 were upregulated and downregulated with a log2-fold change > 1 respectively. These 124 up-regulated genes comprised a group for encoding the cytokines and chemokines (*LTA*, *MIF*, *IL32*, *CXCL10* and *CCL4L2*), a group involving in metabolic [e.g., *GAPDH*, *PKM*, *ENO1*, *TPI1*, *PGK1*] (Figure 3E) [40,41], and a group regulating protein synthesis (e.g., *WARS*, *SEC61G*, *EIF5A*). These phenomena support the cell activation[42]. We also found that naïve T cells increased the expression of calcium binding proteins encoded by S100 family genes(e.g., *S100A4*, *S100A11* and *S100A10*) and cytoskeleton related protein(e.g., *ACTG1*, *ACTB*, *TUBB*, *PFN1*, *MYO1G*), which were reported in response to TCR engagement by antigen[43,44]. Besides, many regulatory markers (e.g., *GITR* [*TNFRSF18*], *CISH*, *SOCS1*, *TIGIT*) and cell apoptosis regulation markers (e.g., *LGALS1*, *FAM162A*, *CFLAR*, *FAS*, *CDKN1A*) were strongly upregulated to maintain immune balance[45,46], although their expression level differs in cells at different differentiation period [4746,47]. These 36 down-regulated genes including *CD127*, *CD27* and *CD69L*, and were consistent with previous studies upon T cells activation[48].

GO analysis of the differentially expressed genes between cells from CMV pp56-stimulated and control naïve cells demonstrated the significant enrichment of genes associated with T cell activation, protein processing, viral gene expression, cytokine signaling, and RNA processing ((Figure 3B). Pseudotime Analysis further indicated that CMV CD4+ naïve T cells may differentiate into cytotoxic, regulatory, or other effector helper T cells, which may be a reservoir of CD4+ effector T cells in response to CMV antigen stimulation (Figure 3C, 3D).

CMV pp65-specific CD4-CTLs shared antigen specificity with Tem

T cell receptor (TCR) repertoire reflects antigen-specificity for cells, and their antigen experience in effector and memory subsets. We therefore analyzed the features of TCR repertoire in each cell cluster from each donor to uncover their potential antigen specificity (donor 1, donor 2). The gene combinations used by each cluster was ranked, showing an obvious clonality of genes combination in CD4-CTL, and a slight genes expansion in Tem. The expanded gene combinations in CD4-CTL and Tem are different between two donors, where Donor1 preferred to use *TRAV5_TRAJ44_TRBV4-3_TRBJ2-1* in CD4-CTL and *TRAV1-2_TRAJ36_TRBV28_TRBJ1-2* in Tem, but Donor2 used *TRAV41_TRAJ49_TRBV6-5_TRBJ2-1* in both CD4-CTL and Tem(Figure

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

198 4A). The analyses on CDR3 from TCR alpha chain (TRA) and TRB further verified the clonality
199 in CD4-CTL and Tem. Two expanded clones in CD4-CTL and one in Tem were observed in
200 Donor1, and four in CD4-CTL and one in Tem was observed in Donor2 (Figure 4B). The
201 abnormal features were also exhibited in CDR3 length in naïve, CD4-CTL and Tem, where CDR3
202 in TRA and TRB from these subsets are shorten in CMV comparing to those in control (Figure
203 4C). These imply that the repertoire of CD4-CTL and Tem may be largely skewed. To identify
204 clones targeting the same antigens among cell subsets, we used GLIPH2[23] to cluster clones from
205 CMV and control. We identified convergences between memory and Treg cells, and those between
206 Tem and CTLs in CMV. Interestingly, the convergent TCR clone in CD4-CTL from CMV just
207 composed a minor part of CTL from control in Donor 1, and no convergent clone was found
208 between them in donor 2, suggesting that CMV pp65-reactivated CTLs is rare in total CTLs.
209 Furthermore, the dominated clones in CTLs almost derived from Tem cells in Donor 2, while in
210 donor 1, only a small part of CTLs clones was shared with Tem (Figure 4D). This phenomenon
211 indicates that Tem and CD4-CTL cells shared antigen specificity, and Tem may have a tight
212 relationship with CD4-CTL in differentiation. Meanwhile, we did not find either expanded gene
213 combination nor expanded CDR3 clones in Treg, implying Treg may be unspecific for antigens
214 and activated in TCR-independent manners.

215 Discussion

216 CD154 is an effective marker to combine with single-cell mRNA sequencing and accomplishment
217 high-throughput analysis of virus antigen-specific T cells cytometry[24–27]. Traditional research
218 methods based on secreted cytokines, such as IFNG or TNF, to testing cmv-specific T cells,
219 proven to perform well[49–51]. However, it was of limited application in combination with sc-mRNA
220 sequencing for a reason of cell damage caused by intracellular staining. Another popular method
221 uses Peptide-MHC multimer to isolate antigen-specific T cells according to specific binding of
222 TCR with pMHC allows detailed TCR and phenotypic analysis of cells using single-cell
223 technologies [52,53,54]. However, the decrease of TCR expression in activated T cells leads to the
224 preference of relative low antigen-specific T cells bound with tetramer, and the selection of
225 multimer-binding CD4+ T cells may bias our understanding of the phenotype of antigen-specific
226 CD4+ T cells[55]. The fact that 83.8% CD4+ T cells in CMV high express *CD154* (*CD40LG*)

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

compared with 17.4% low expressed in control indicate that CD154 is comparable with IFNG and TNF in discriminating antigen-specific CD4+ T cells (Figure 2c). Thus, CD154 is an effective marker to combine with single-cell mRNA sequencing and accomplishment a high-throughput technique platform for research on antigen activated CD4+ T cells.

Bystander activation of CD4+ T cells is less well studied compared with CD8+T cells , but it was demonstrated that unrelated memory CD4 T cells can be activated after a recall tetanus vaccination via bystander activation[56] and multiple cytokines sharing the common receptor gamma chain can induce CD154/CD40 ligand expression by human CD4+ T lymphocytes via a cyclosporin A-resistant pathway[57]. In our data, CD4+ memory-like T cells are under an environment of IFNG and IL2, and they are prone to be activated by these cytokines. Besides, we find neither clonal expansion nor CDR3 length change in CD4+ memory-like T cell. So, it is difficult to conclude whether these CD4+ memory-like T cells are CMV pp65 antigen-specific.

Our data indicated that the CMV-reactivated Treg had heterogeneously inhibitory functions. *LAG3* and *CTLA4* are classical markers of Treg, by which Treg completely contact MHC-II and CD80/CD86 respectively to repress the activation of conventional T cells. Perforin/granzyme-induced apoptosis is a main pathway used by cytolytic cells to kill target cells[58,59], and they commonly expressed simultaneously. In our study, Treg highly expressed *GZMB*, *SRGN* (encoding an element protein for maintain granzyme storage), but their *PRF1* expression is limited. One explanation is that a few perforins is enough to facilitate the entrance of granzyme into target cells. The other is that granzyme B can induce cell death in a perforin-independent manner[60] where it mediate a cleavage of extracellular matrix to reduce the adhesion of immune cells and result in their death. In addition to these classical inhibitory manners, we observed the expression of *LGALS1* and *LGALS3*, encoding Gal-1 and Gal-3 respectively, which may also participate in Treg immunosuppressive activity[61]. In previous studies, disruption of Gal-1 attenuated the immunoexpressing effect of Treg cells[62] and Gal-1 from Treg induced the dysfunction of effector T cells and modulated their transient calcium influx[63]. This regulatory mechanism is not limited to Gal-1 but also employed by Gal-3 in Treg[64].

The presentation of Th1-related cytokines in Treg has been observed in some studies, but whether

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

its appearance implied the stability of Treg inhibitory function is still controversial. In autoimmune diseases, the expression of effector cell cytokines often coupled with decreased inhibitory function in Treg[65,66]; while in other conditions, the expression of T-bet and inflammatory cytokines does not affect the inhibitory function[67,68]. TIGIT is a key regulator to maintain the stability of Treg inhibitory function[69,70], and is commonly expressed in activated naïve, memory and Treg cells[71]. In this study, Treg expressed TIGIT as similar as activated naïve T cells, implying Treg retained the inhibitory function under CMV infections, and expression of IFN-γ and TNF might enhance their inhibitory function on cells with Th1 phenotypes, including CD4-CTL, Tem and activated naïve T cells. Furthermore, *CCR6*, *CCR5*, *CCR7*, *CCL20*, *CCL3*, *CCL4* and *CCL5* were expressed by both Treg[72] and conventional cells, ensuring co-localization of these cells. Meanwhile, Treg unregulated *IL2RA*, implying an enhanced IL-2 signaling pathway. LTA is the downstream protein of IL-2RA, and its expression condition lymphatic endothelia for enhanced Treg transendothelial migration[73]. Notably, no convergent TCR clone was found between Treg and other CMV-reactivated CD4+ T cells, suggesting Treg may not function to inhibit other CMV-related CD4+ T cells in TCR-dependent manner. In together, Treg cells in CMV infection maintain their inhibitory function, and obtained reinforced inhibitory ability by enhanced migration function.

Another significance of Treg is its large proportion in CD4+ T cells. CD4-CTL in CMV is reported to increase cardiovascular mortality risk of CMV carriers[74]. Chronic infection and repeated activation of CMV may induced a continuous presentation of CD4-CTL and activation of other immune cells, and it is reasonable to induce an accumulation of regulatory cells to avoid their side effects. The large percentage of Treg in CMV-reactivated cells partly supported the hypothesis, but more studies are need to clarify whether the accumulation is necessary to protect individual or contributes to the persistent CMV infection. Additionally, it is novel to find the expression of *CD70* on Treg to our knowledge. *CD70* is commonly expressed on antigen-presenting cells as well as activated T cells to conform CD27–CD70 pathway to provide a costimulatory signal. In T cells, CD70 was showed to induce caspase-dependent apoptosis, and it may also perform a similar function in Treg[75]. However, more investigations should be taken to unveil CD70 function in Treg.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CMV-specific CD4+ Tem displays a distinct cytotoxic function as highly expressed a panel of canonical cytolytic molecules (*GZMB*, *GZMH*, *GZMA*, *PRF1*, *GNLY*, *NKG7*, *IFNG*, *CTSC*, *FGFBP2*, and *KLRB1*). GZMA, GZMB, and GZMH belong to granzyme, a subfamily of serine proteases that are function in mediated cell death. This simultaneous high expression of granzyme, PRF1, CTSC, and GNLY, may imply that CMV-specific CD4+ Tem is capable of kill target cells in a granzyme- and perforin- dependent manner. CMV-specific CD4+ Tem cells exert their function in peripheral target organs by the production of antimicrobial lymphokines, Inflammatory chemokine, and cytotoxin, thereby directly contributing to the containment of viral infection. These well prepared cytolytic particles enable CD4+ Tem to quickly start a war against the infection virus.

CD4+ T cells are central organizers in immune responses. CMV-stimulated CD4+ T cell secrets chemokines and effector molecules to recruit and activate or assist other immune cells to orchestral an antiviral response. For example, HCMV-specific CD4+ T cells overall express high levels of chemokines *CCL4* and the CMV-specific naïve CD4+ T cells express high levels of *CXCL10*. CCL4 help recruit immune cells such as macrophages, NK cells, monocytes as well as activated T cells to the site of infection[76–79]. CXCL10 help HCMV-specific CD4+ T to recruit HCMV-specific CD8+ T cells who high expressed CXCR3 (chemokine receptors of CXCL10)[80] to the Inflammatory site. In addition, IFNG high expressed by CMV-specific CD4+ T could help in direct anti-viral activity and plays an important role in activating immune cells (such as B cells, T cells, NK cells, and macrophage) and augmenting antigen presentation (such as DC, B cells, and macrophages)[81,82]. The high expressed cytotoxic marker (*LTA*, *GZMB*) also implies possible direct antiviral effects of CD4+ T Cells. Since HCMV has acquired extensive mechanisms of immune evasion, which include the downregulation of MHC class I molecules on infected cells[83,84]. whereas, effector CD4+ T cells require no MHC class I molecules to eliminate target cells, these effector lymphocytes may have evolved to support effector CD8+ T cells in suppression of HCMV infection[1]. Notably, in our data, almost all (99.1%) CMV-specific CD4+ T cell highly expressed IFNG compare with few (1.6%) expressed in control CD4+ T with logFC = 4.576. These results go nicely in line with the fact that CD4 T cells are important in maintain CD8 T-cell activity during prolonged infection, and their role in helping the antiviral antibody response may also be

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

essential[82].

Unlike other latent-virus specific T cells (such as EBV, HSV), CMV CD4+ T cell does not show an exhaustion phenotype. Both EBV-specific CD8+ and CD4+ T cells are highly susceptible to apoptosis, while CMV CD4+ T cells are not. Akin to the cmv CD4+ T cell response, the majority of activated EBV-specific CD4+ T cells exhibit Th1-type response and express both perforin, granzyme B, CD107a[85,86] and lack the lymphoid homing markers CCR7 and CD62L[87]. It was speculated that EBV-specific CD4+ T cells might highly effective against MHC-II positive EBV infected B cells. CD4+CD25+ Treg cells are also found in HSV-1 specific CD4+ T cells[88]. HSV[89]: HSV-1 can establish a latent infection in TG of host, while CD4+ and CD8+ T cells can control the reactivation of the HSV-1 by surrounding latently infected neurons. CMV CD4+ T cells exhibit downregulation expression of IL7R, which is also observed in other chronic infections such as LCMV clone 13 and HIV infection[90–93]. This phenomenon is consistent with the previous reports that T cells fail to re-express the IL7R[94], when they are continuously stimulated.

## Methods and Materials

### PBMC preparation

We obtained peripheral blood from three CMV IgG-positive, heathy donors though a research protocol proved by the Beijing genomics institution-Shenzhen (BGI-Shenzhen) Institutional Review Board (IRB). Peripheral blood mononuclear cell (PBMCs) were immediately isolated from blood collected with EDTA blood collection tube by density centrifuge method with Histopaque-1077 (Sigma, Cat. 10771) within two hours[95], resuspended in 4°C cryopreservation medium consisting 90% fetal bovine serum (FBS, HYCLONE , Cat. sh30084.03) and 10% Dimethyl sulfoxide (DMSO, Sigma, Cat. D4540) and then placed in Mr. Frosty (Thermo Scientific) in -80°C container. Samples were then moved to liquid nitrogen for a long-time storage.

Additionally, 2 ml peripheral blood from each donor was collected by blood collection tube without any additive, placed at room temperature for 30 minutes(min) and centrifuged for 10min at 2000g. Then plasma was collected and heat shocked for 30min at 55°C.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

PBMC stimulation

Frozen PBMC from liquid nitrogen were immediately thawed in 37°C water and resuspended in complete medium (RPMI 1640 medium, 10% NEAA and 2% autologous plasma; RPMI 1640 and NEAA were purchased from ThermoFisher with Cat. 72400120 and Cat. 11140050) to a final density of 1*10^7 per milliliter (ml). We moved 150 microliter (ul) of cell suspension with three repetitions to each well in the 96-well U-plate (Falcon) and incubated them at 37°C for two hours. Then 75ul culture supernatant in each well was replaced by 75ul stimulation medium, and gently mixed. Cells were cultured in an incubator with 5% $CO_2$ at 37°C for 24 hours.

The stimulation medium included RPMI 1640 medium (without serum), anti-CD28 (2ug/ml, Clone G28.5, Genetex, Cat. GTX14148), anti-CD40 (2ug/ml, Clone HB14, Miltenyi, Cat. 130-094-133) with/without CMV peptide (1.2 nmol/ml per peptide). To preserve the surface expression of CD154 on activated T cells, we used anti-CD40 to inhibit the interaction of surface CD154 with its counterpart CD40 as described in the previous study[22]. CMV pp65 peptide was purchased from Miltenyi (Cat. 130-093-438) and diluted in sterile water.

Enrichment of CMV pp65-specific T cells

Cells were collected and washed with FACS washing buffer (DPBS, 2% FBS and 1mM EDTA) for once and resuspended in staining buffer (FACS washing buffer with 10% human plasma and 1% BSA) containing antibodies against CD3, CD4, CD154 and CD69 (Supplementary Table 1). After been incubated on ice for 40 minutes, cells were washed with FACS washing buffer for twice, and resuspended in 100ul washing buffer. The stained cells were analyzed and sorted by a BD FACS Aria II cell sorter (BD Biosciences). For cells stimulated with CMV peptide, CD3+CD154+ cells were sorted as CMV-specific T cells; For unstimulating cells, monocytes and lymphocytes were sorted respectively and re-mixed as a control. The gating schedule for cells sorting was recorded by BD Aria II (Supplementary Fig.1), and FACS data was analyzed with Flowjo v10.0.7

Droplet generation, 10X RNA-seq and TCR-seq library preparation and sequencing

After been counted with C-Chip (inCYTO), CMV-specific cells and control cells from all three

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

individuals were mixed separately, diluted with PBS to a final concertation ~800/ul, and about 20,000 cells per reaction were loaded onto a Chromium Single Cell Chip (10x Genomics). The libraries for RNA-seq and TCR-seq were prepared using the Chromium Single Cell 5' Library & Gel Bead Kit v2, Chromium Single Cell V(D)J Human T Cell Enrichment Kit (10X Genomics) following the manufactory's protocol. Sequences within these libraries were ligated with BGIseq adapters [70](doi:10.1016/j.scib.2020.01.002), and then CMV and control libraries were loaded onto sequencing chip. The RNA-seq libraries were sequenced with an 8-base index read, a 26-base read 1 containing cell-identifying barcodes and unique molecular identifiers (UMIs), and a 100-base read 2 containing transcript sequences on BGIseq500; TCR-seq were sequenced with an 8-base index read, a 150-base read 1 containing cell-identifying barcodes, UMIs and insert started from V-gene region, and a 150-base read 2 containing insert from C-gene region. The raw data after sequencing was about 10+35 Gb per library for RNA-seq and 35+35 Gb for TCR-seq. The data that support the findings of this study have been deposited into CNGB Sequence Archive[96] of CNGBdb[97] with accession number CNP0001262.

Preprocessing single cell RNA-seq data

Raw data were split according to sample barcodes into CMV-stimulated (ST) and unstimulated library (CON), and then were filtered, blasted, aligned and qualified by Cellranger v2.2.0 with reference of refdata-cellranger-GRCh38-1.2.0 for RNA-seq data and Cellranger v3.0.0 with refdata-cellranger-vdj-GRCh38-alts-ensembl-2.0.0 for TCR-seq data. Other parameters were set as default in the software.

Supplementary Table 1. FACS antibodies

| Antigen | Clone | Fluorophore | Supplier | Dilution |
|---|---|---|---|---|
| CD3 | SK7 | FITC | BIOLEGEND | 1:100 |
| CD4 | RPAT4 | PerCP-Cy5.5 | EBIOSCIENCE | 1:200 |
| CD154 | TRAP-1 | PE | BD | 1:50 |
| CD69 | FN50 | BV421 | BIOLEGEND | 1:50 |

Data Integrating and cell clustering

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The R package Seurat[98] 3.1.5 was used to integrate and analyze datasets from CMV and control. The merged expression matrix was firstly filtered following the Seurat recommendation[99,100] and a total of 8671 cells with unique UMI was obtained. Unsupervised clustering was conducted with Seurat with the parameter res = 0.5, it revealed a total of 16 clusters. We used mRNA biomarkers (Supplemental Table 2) obtained from recently published articles[101–103] to identify these clusters, and GSEA analysis was conducted to identify cluster3(see method Gene set enrichment analysis). TCR repertoire datasets were also combined to identify putative mucosal associated invariant T (MAIT) cells and γδT. Based on the expression of a typical TCR [104,105](Supplemental Table 2), a 17th cluster (putative mucosal associated invariant T, MAIT) was emerged from cluster12. Based on whether they express a TCR a chain, and meanwhile express γδ biomarkers (CD3D, CD3E, TRDC, TRGC1, TRGC2), 2111 γδT cells were identified and excluded from further analysis. The above classification was consistent with hematopoietic differentiation and the previously published t-SNE plots of PBMC scRNA-seq[106–108].

Supplementary Table 2. cell type markers

| Cell Type | Markers |
|---|---|
| naïve CD4 T | CD3E, CD4, SELL, CD27, TCF7, CCR7 [102] |
| naïve CD8 T | CD3, CD8A, SELL, CD27[102] |
| γδT | CD3+CD4-CD8- \| CD3+CD4-CD8aa+, CD3D, CD3E, TRDC, TRGC1, TRGC2[101] |
| Memory Like T | CD3E, CD4, SELL, CD27 \| CCR7, SELL, TCF7 |
| Treg | CD3E+, TIGIT, CTLA4[108] \|FOXP3, IL2RA, CTLA4[109] |
| Tem | CD3E, CD4/CD8A, CD27, PRF1, GNLY |
| B | CD79A, CD19, CD79B, IGKC, IGHM, MS4A1, IGHD |
| NK | CD3E-,CD4+,FOXP3+,TNFRSF9+, NKG7, GNLY, NKG7, KLRD1, KLRC1[103] |
| active CD8 T | CD3E, CD8A, CD69, TNFRSF9(CD137) |

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

| | |
|---|---|
| CTL | GZMA, GZMB, PRF1, NKG7,CCR7, CD27, CD28, IL7R[110] |
| Monocyte | LYZ, S100A9, CD14,FGL2,MS4A7[103] |
| MAIT | TRAV1-2 / TRAJ33 , TRAV1-2 / TRAJ20 \| TRAV1-2 / TRAJ12[105] |

Differential express gene (DEG) analysis

DEG analysis was conducted by the function *FindMarkers* provided by *Seurat*. To character the features of CMV-specific CD4+ T cell response, we used a stricter standard to filter out DEGs between CMV and control CD4+ T cells according to the following standard: for upregulation genes in CMV, adjusted P-value < 0.05, log fold change >1, percentage of cells expressing the gene in CMV sample (pct.1) >0.8, percentage of cells expressing the gene in control (pct.2) < 0.2; for downregulation genes in CMV, adjusted P-value < 0.05, logFC >1, pct.1 <0.2 , pct.2 >0.8.

Quality control metrics and filtering

We used the CellRanger v2.2.0 software with the default settings to process the raw FASTQ files, align the sequencing reads to the GRCh38 transcriptome, and generate a filtered UMI expression profile for each droplet.

Identifying the sample identity of each droplet

Firstly, transcriptome of each donors' peripheral blood mononuclear cell was sequenced on the BGI-SEQ500 platform with the sequencing type SE200. Raw data with 10G per sample was obtained. Then, we followed the best practices workflows recommend by GATK(https://gatk.broadinstitute.org/hc/en-us/articles/360035531192-RNAseq-short-variant-discovery-SNPs-Indels-) to call SNP (Single-nucleotide polymorphism) and create VCF files containing the genotype (GT) to assign each barcode to a specific sample. The VCF file and BAM files produced by cellranger2 was passed to the demuxlet software to deconvolute sample identity[28]. Finally, we assign the best guess of the samples' identity to their corresponding donor and each "possible" or "ambiguous" droplets as unclear.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

425 Gene ontology analysis

426 To annotate the potential functions of the DEGs of each CD4+ T cell cluster , GO enrichment
427 analysis was conducted using the clusterProfiler R package[111](version 3.14.3) with the differential
428 expressed feature genes identified by Seurat. Top 20 Enriched pathways, ranked by normalized
429 enrichment score, with the FDR q-val ≤0. 05 were chosen and visualized.

430 Gene set enrichment analysis

431 Gene set enrichment analysis (GSEA, http://www.broad.mit.edu/gsea) was conducted with default
432 sets to deduce the cell type of cluster 3. The gene set collection used for GSEA was
433 c7.all.v7.1.symbols.gmt (ftp.broadinstitute.org://pub/gsea/gene_sets/c7.all.v7.1.symbols.gmt).

434

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

435 Pseudotime Analysis

436 A total of 1200 cells from cmv (include five clusters: naïve CD4 T, central memory T, effector
437 memory T, Treg, CTL) were used for pseudotime analysis. The analysis was performed with
438 Monocle2 (version 2.14.0) with UMI count expression data [112]. All the parameters were set as
439 default. Differential expression gene between the five clusters was identified with
440 differentialGeneTest function implemented in Monocle 2. The resulting genes with qval < 1e-5
441 was selected for ordering cells in pseudotime using reduceDimension with the DDRTree method
442 and orderCells functions.

443 Donor 1

| Cluster | TRB CDR3 | Subtype |
|---|---|---|
| 1 | CAISALAGHQTYNEQFF | memory |
| | CASSLAGHNGARELFF | Treg |
| | CASSPRLAGHTGELFF | Treg |
| 2 | CASSPRTKGASGRAVETQYF | memory |
| | CASSPRTKGASGRAVETQYF | Treg |
| 3 | CASSLWSTEDTQYF | memory |
| | CASSLWSTEDTQYF | Treg |
| 4 | CASSSLPSNYGYTF | CTL |
| | CASSSLPSNYGYTF | Tem |
| 5 | CASSEGNSGDNQPQHF | Treg |
| | CASSEGNSGDNQPQHF | Tem |
| 6 | CASSSRRTGIPTDTQYF | memory |
| | CASSSRRTGIPTDTQYF | Treg |
| 7 | CASSLDIQETQYF | memory |
| | CASSLDIQETQYF | Treg |

444

445 Donor 2

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

| Cluster | TRB CDR3 | Subtype |
|---|---|---|
| 1 | CASSQVGAELYGYTF | Treg |
| | CASSQVGTATEAFF | Treg |
| | CASSSQVGYGYTF | memory |
| 2 | CASSTTSGGYNEQFF | CTL |
| | CASSTTSGGYNEQFF | Tem |
| 3 | CSAREDRAWAPLHF | CTL |
| | CSAREDRAWAPLHF | Tem |

446

447 Figure 1: CMV pp65-specific CD4+ T cells are characterized by typical antiviral profiles. (A)
448 Experimental workflow for single-cell analysis of CD4+ T cells from PBMC of three donors
449 includes CMV pp65 in vitro stimulation and culture, CD154+ T cell sorting, and 5′ single-cell
450 RNA and paired T cell receptor sequencing. UMAP embeddings of merged scRNA-seq profiles
451 from control and stimulated (CMV) immune cells were plotted and colored by cell cluster (B) and
452 sample (C) respectively. (D) UMAP projections for the merged CD4+ T cells colored by
453 expression of *CD3E*, *CD4*, *SELL* (naive marker), *TCF7*, *CCR7*, *CD27*, *CD28*, *FOXP3* (Treg
454 marker), *TIGIT*, *IL2RA*, *CTLA4*, *PRF1* (cytotoxic marker), *GZMB*, *GNLY*, *NKG7*. Expression
455 values are normalized across CMV and control datasets. (E) Heat map of scaled mean gene
456 expression of the major canonical markers (columns) detected in different cell types in merged
457 cells of CMV and control (rows). (F) Dot plot of differential express genes (DEGs) shows both the
458 expression level and the percentage of CD4+ T cells in CMV and control. (G) Gene ontology (GO)
459 analysis of DEGs between CMV CD4+ T and that of control. The Top 20 enriched GO terms are
460 ordered on the y-axis. X-axis represents the gene percentage in enriched GO terms. Sizes of the
461 dots represent the number of genes included in each GO term. The color gradient of dots
462 represents the adjusted P-values of each enriched GO term.

463 Figure 2: Polyfunctionality profiles of CMV pp65-stimulated CD4 + T cell subsets. UMAP
464 embeddings of merged scRNA-seq profiles from control and stimulated (CMV) CD4+ T cells
465 were plotted and colored by cell cluster (A) and sample (B) respectively. Subpopulations of CD4+

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

T cell colored in (A) were identified with canonical markers described in Supplementary Table 2. (C) Distribution of the abundance of the five subsets in the CD4+ T cells of CMV and control datasets. (D) Heat map of these five subsets with the Top10 DEGs between each of them. (E) Violin plots of exemplary feature gene expressions of the five subsets. These feature genes were classified and labeled with their group name on the left.

Figure 3: CMV pp65 stimulated CD4+ naïve T cells show obvious activation characteristics. (A) Volcano plot showing the log2FoldChange(x-axis) and −log10(p_val_adj) (y-axis) for differential expression gene between CMV and control naïve CD4+ T cells. Genes with log2FoldChange > 1 and adjusted p value < 0.05 are more highly expressed among CMV and highlighted in red and labeled with their names. Genes with log2FoldChange <- 1 and adjusted p value < 0.05 are downregulated in CMV and highlighted in blue and labeled with their names. (B) GO analysis of DEGs between CMV naïve CD4+ T and that of control. The Top 20 enriched GO terms are ordered on the y-axis. X-axis represents the gene percentage in enriched GO terms. Sizes of the dots represent the number of genes included in each GO term. The color gradient of dots represents the adjusted P-values of each enriched GO term. (C) and (D) Pseudotime analysis of the five CD4+ T cell subsets in CMV. These cells were colored by cell cluster (C) and cell state (D) respectively. (E) Violin plots of exemplary feature gene expressions of the naïve CD4+ T cells from CMV and control. These feature genes were classified and labeled with their group name on the left.

Figure 4: TCR repertoire analysis of the five CD4+ T cell subsets from donor 1 and donor 2. Frequencies of gene combination of TRAV_TRAJ_TRBV_TRDJ (A) and the combination of CDR3a_CDR3b (B) of five stimulated CD4+ T cell subsets from each donor. X-axis is the ranked top 10 gene combination of each cluster. Y-axis is the count of each clonotype. (C) Distribution of amino acid (A.A.) length of TCR alpha and TCR beta of five CD4+ T cell subsets in CMV and control from the two donors. (D) TCR convergence of each cluster from two donors were analyzed with GLIPH2[23] (left: donor 1, right: donor 2). Clonotypes are sorted by descending order with their sequence counts, and the wide on circus indicate the sequence count per clonotype.

Supplement Figure 1: Distribution of stimulated CD4+T cells from each donor (n = 3). (A) UMAP

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

494 embeddings of CMV CD4+ T cells from each donor. *demuxlet*[28] was used to assign these cells to
495 each donor, the ambiguous droplet was assigned as "unclear". Proportions of cells from each
496 donor was showed on the left. The UMAP embeddings were colored (A) or split by donors(B)
497 respectively.

498 1. Sylwester, A. W. *et al*. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T
499 cells dominate the memory compartments of exposed subjects. *J. Exp. Med.* **202**, 673–685
500 (2005).

501 2. Pawelec, G. *et al*. Immunosenescence and Cytomegalovirus: where do we stand after a decade?
502 *Immun. Ageing* **7**, 13 (2010).

503 3. Lilleri, D. *et al*. Human cytomegalovirus-specific CD4+ and CD8+ T-cell reconstitution in
504 adult allogeneic hematopoietic stem cell transplant recipients and immune control of viral
505 infection. *Haematologica* **93**, 248–256 (2008).

506 4. Gabanti, E. *et al*. Human Cytomegalovirus (HCMV)-specific CD4+and CD8+ T cells are both
507 required for prevention of HCMV disease in seropositive solid-organ transplant recipients.
508 *PLoS One* **9**, (2014).

509 5. Gabanti, E. *et al*. Reconstitution of Human Cytomegalovirus-Specific CD4+ T Cells is Critical
510 for Control of Virus Reactivation in Hematopoietic Stem Cell Transplant Recipients but Does
511 Not Prevent Organ Infection. *Biol. Blood Marrow Transplant.* **21**, 2192–2202 (2015).

512 6. Einsele, H. *et al.* Infusion of cytomegalovirus (CMV)-specific T cells for the treatment of
513 CMV infection not responding to antiviral chemotherapy. *Blood* **99**, 3916–3922 (2002).

514 7. Sadeghi, M. *et al.* Dysregulated cytokine responses during cytomegalovirus infection in renal
515 transplant recipients. *Transplantation* **86**, 275–285 (2008).

516 8. Kang, S., Brown, H. M. & Hwang, S. Direct antiviral mechanisms of interferon-gamma.
517 *Immune Netw.* **18**, 1–15 (2018).

518 9. Malyshkina, A. *et al.* Fas Ligand-mediated cytotoxicity of CD4+ T cells during chronic

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

519 retrovirus infection. *Sci. Rep.* **7**, 1–10 (2017).

520 10. Zajac, A. J., Quinn, D. G., Cohen, P. L. & Frelinger, J. A. Fas-dependent CD4+ cytotoxic
521 T-cell-mediated pathogenesis during virus infection. *Proc. Natl. Acad. Sci. U. S. A.* **93**,
522 14730–14735 (1996).

523 11. Brown, D. M. Cytolytic CD4 cells: Direct mediators in infectious disease and malignancy. *Cell.*
524 *Immunol.* **262**, 89–95 (2010).

525 12. Lukacher, B. Y. A. E., Morrison, L. A., Braciale, V. L., Malissen, B. & Braciale, T. J.
526 CLONES. **162**, (1985).

527 13. Arase, N. *et al.* Heterotypic interaction of CRTAM with Necl2 induces cell adhesion on
528 activated NK cells and CD8+ T cells. *Int. Immunol.* **17**, 1227–1237 (2005).

529 14. Takeuchi, A. *et al.* CRTAM determines the CD4+ cytotoxic T lymphocyte lineage. *J. Exp. Med.*
530 **213**, 123–138 (2015).

531 15. Takeuchi, A. & Saito, T. CD4 CTL, a cytotoxic subset of CD4+ T cells, their differentiation
532 and function. *Frontiers in Immunology* vol. 8 (2017).

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

535 17. Velaga, S. *et al.* Reconstitution and phenotype of tregs in CMV reactivating patients following
536 allogeneic hematopoietic stem cell transplantation. *Immunol. Invest.* **42**, 18–35 (2013).

537 18. Rouse, B. T., Veiga-Parga, T. & Sehrawat, S. Tamara Veiga-Parga Sharvan Sehrawat Role of
538 regulatory T cells during virus infection. 182–196 (2013).

539 19. Almanan, M. *et al.* Tissue-specific control of latent CMV reactivation by regulatory T cells.
540 *PLoS Pathog.* **13**, (2017).

541 20. Chattopadhyay, P. K., Yu, J. & Roederer, M. A live-cell assay to detect antigen-specific CD4+
542 T cells with diverse cytokine profiles. *Nat. Med.* **11**, 1113–1117 (2005).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

21. Meier, S., Stark, R., Frentsch, M. & Thiel, A. The influence of different stimulation conditions on the assessment of antigen-induced CD154 expression on CD4+ T cells. *Cytom. Part A* **73**, 1035–1042 (2008).

22. Frentsch, M. *et al*. Direct access to CD4+ T cells specific for defined antigens according to CD154 expression. *Nat. Med.* **11**, 1118–1124 (2005).

23. Huang, H., Wang, C., Rubelt, F., Scriba, T. J. & Davis, M. M. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat. Biotechnol.* (2020) doi:10.1038/s41587-020-0505-4.

24. Klinik, M. *et al*. CD154 , a marker of antigen-specific stimulation of CD4 T cells , is associated with response to treatment in patients with chronic HCV infection. *J. Viral Hepat.* (2011) doi:10.1111/j.1365-2893.2010.01430.x.

25. Frentsch, M. *et al*. Direct access to CD4 + T cells specific for defined antigens according to CD154 expression. *Nat. Methods* **11**, 1118–1124 (2005).

26. Chattopadhyay, P. K., Yu, J. & Roederer, M. A live-cell assay to detect antigen-specific CD4 + T cells with diverse cytokine profiles. *Nat. Med.* **11**, 1113–1117 (2005).

27. Kirchhoff, D. *et al*. Identification and isolation of murine antigen-reactive T cells according to CD154 expression. 2370–2377 (2007) doi:10.1002/eji.200737322.

28. Kang, H. M. *et al*. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018).

29. Kang, S. & Brown, H. M. Direct Antiviral Mechanisms of. **18**, 1–15 (2018).

30. Samuel, C. E. Antiviral Actions of Interferons. **14**, 778–809 (2001).

31. Yu, F. *et al*. The transcription factor Bhlhe40 is a switch of inflammatory versus antiinflammatory Th1 cell fate determination. **215**, (2018).

32. Lanio, N., Sarmiento, E., Gallego, A. & Carbone, J. Immunophenotypic profile of T cells in

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

common variable immunodeficiency: is there an association with different clinical findings? *Allergol. Immunopathol. (Madr).* **37**, 14–20 (2009).

33. Davicino, R. C. *et al.* Galectin-1–Driven Tolerogenic Programs Aggravate Yersinia enterocolitica Infection by Repressing Antibacterial Immunity . *J. Immunol.* **199**, 1382–1392 (2017).

34. Probst-Kepper, M., Kröger, A., Garritsen, H. S. P. & Buer, J. Perspectives on regulatory T cell therapies. *Transfus. Med. Hemotherapy* **36**, 302–308 (2009).

35. Lasoudris, F. *et al.* IL4I1: An inhibitor of the CD8+ antitumor T-cell response in vivo. *European Journal of Immunology* vol. 41 1629–1638 (2011).

36. Yi, G. *et al.* Single-cell RNA-seq unveils critical regulators of human FOXP3 + regulatory T cell stability. *Sci. Bull.* **65**, 1114–1124 (2020).

37. Russell, J. H. & Ley, T. J. L YMPHOCYTE -M EDIATED C YTOTOXICITY. (2002) doi:10.1146/annurev.immunol.20.100201.131730.

38. Oh, D. Y. *et al.* Intratumoral CD4 + T Cells Mediate Anti-tumor Cytotoxicity in Human Bladder Cancer Article Intratumoral CD4 + T Cells Mediate Anti-tumor Cytotoxicity in Human Bladder Cancer. *Cell* 1–14 (2020) doi:10.1016/j.cell.2020.05.017.

39. Chan, K. *et al.* The Roles of MHC Class II, CD40, and B7 Costimulation in CTL Induction by Plasmid DNA. *J. Immunol.* **166**, 3061–3066 (2001).

40. MacIver, N. J., Michalek, R. D. & Rathmell, J. C. Metabolic Regulation of T Lymphocytes. *Annu. Rev. Immunol.* **31**, 259–283 (2013).

41. Maciolek, J. A., Alex Pasternak, J. & Wilson, H. L. Metabolism of activated T lymphocytes. *Current Opinion in Immunology* vol. 27 60–74 (2014).

42. Ahern, T. & Kay, J. E. Protein synthesis and ribosome activation during the early stages of phytohemagglutinin lymphocyte stimulation. *Exp. Cell Res.* **92**, 513–515 (1975).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

591 43. Liu, X. *et al.* T cell receptor-induced nuclear factor κb (NF-κB) signaling and transcriptional
592 activation are regulated by STIM1- and Orai1-mediated calcium entry. *J. Biol. Chem.* **291**,
593 8440–8452 (2016).

594 44. Kannan, Y. TEC and MAPK Kinase Signalling Pathways in T helper (TH) cell Development,
595 TH2 Differentiation and Allergic Asthma. *J. Clin. Cell. Immunol.* **01**, (2013).

596 45. Guram, K. *et al.* A threshold model for T-cell activation in the era of checkpoint blockade
597 immunotherapy. *Front. Immunol.* **10**, 1–20 (2019).

598 46. Rathmell, J. C. & Thompson, C. B. *Review Pathways of Apoptosis in Lymphocyte Development,*
599 *Homeostasis, and Disease*. *Cell* vol. 109 (2002).

600 47. Fuertes Marraco, S. A., Neubert, N. J., Verdeil, G. & Speiser, D. E. Inhibitory receptors beyond
601 T cell exhaustion. *Frontiers in Immunology* vol. 6 (2015).

602 48. Best, J. A. *et al.* Transcriptional insights into the CD8 + T cell response to infection and
603 memory T cell formation. *Nat. Immunol.* **14**, 404–412 (2013).

604 49. Pera, A. *et al.* CMV induces expansion of highly polyfunctional CD4 + T cell subset
605 coexpressing CD57 and CD154 . *J. Leukoc. Biol.* **101**, 555–566 (2017).

606 50. Gamadia, L. E. *et al.* Primary immune responses to human CMV: A critical role for
607 IFN-γ-producing CD4+ T cells in protection against CMV disease. *Blood* **101**, 2686–2692
608 (2003).

609 51. Gamadia, L. E., Rentenaar, R. J., Van Lier, R. A. W. & Ten Berge, I. J. M. Properties of CD4+
610 T cells in human cytomegalovirus infection. *Hum. Immunol.* **65**, 486–492 (2004).

611 52. Dash, P. *et al.* Quantifiable predictive features define epitope-specific T cell receptor
612 repertoires. *Nature* **547**, 89–93 (2017).

613 53. Glanville, J. *et al.* Identifying specificity groups in the T cell receptor repertoire. *Nature* **547**,
614 94–98 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

54. Fan, H. C., Fu, G. K. & Fodor, S. P. A. Combinatorial labeling of single cells for gene expression cytometry. *Science (80-. ).* **347**, 1258367 (2015).

55. Fuchs, Y. F. *et al.* Gene Expression-Based Identification of Antigen-Responsive CD8+ T Cells on a Single-Cell Level. *Front. Immunol.* **10**, 1–15 (2019).

56. Whiteside, S. K., Snook, J. P., Williams, M. A. & Weis, J. J. Bystander T Cells□: A Balancing Act of Friends and Foes. *Trends Immunol.* **xx**, 1–15 (2018).

57. Fayen, J. D. & Western, C. Multiple cytokines sharing the common receptor c chain can induce CD154 / CD40 ligand expression by human CD4 + T lymphocytes via a cyclosporin A-resistant pathway. *Immunology* (2001).

58. Voskoboinik, I., Whisstock, J. C. & Trapani, J. A. Perforin and granzymes: Function, dysfunction and human pathology. *Nat. Rev. Immunol.* **15**, 388–400 (2015).

59. Raja, S. M. *et al.* Cytotoxic Cell Granule-mediated Apoptosis. *J. Biol. Chem.* **277**, 49523–49530 (2002).

60. Gondek, D. C., Lu, L.-F., Quezada, S. A., Sakaguchi, S. & Noelle, R. J. Cutting Edge: Contact-Mediated Suppression by CD4 + CD25 + Regulatory Cells Involves a Granzyme B-Dependent, Perforin-Independent Mechanism . *J. Immunol.* **174**, 1783–1786 (2005).

61. Blidner, A. G., Méndez-Huergo, S. P., Cagnoni, A. J. & Rabinovich, G. A. Re-wiring regulatory cell networks in immunity by galectin-glycan interactions. *FEBS Lett.* **589**, 3407–3418 (2015).

62. Baatar, D. *et al.* Tregs utilize β-galactoside-binding protein to transiently inhibit PI3K/p21ras activity of human CD8+ T cells to block their TCR-mediated ERK activity and proliferation. *Brain. Behav. Immun.* **23**, 1028–1037 (2009).

63. Wang, J. *et al.* Cross-Linking of GM1 Ganglioside by Galectin-1 Mediates Regulatory T Cell Activity Involving TRPC5 Channel Activation: Possible Role in Suppressing Experimental Autoimmune Encephalomyelitis. *J. Immunol.* **182**, 4036–4045 (2009).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

64. Ocklenburg, F. *et al*. UBD, a downstream element of FOXP3, allows the identification of LGALS3, a new marker of human regulatory T cells. *Lab. Investig.* **86**, 724–737 (2006).

65. Dominguez-Villar, M., Baecher-Allan, C. M. & Hafler, D. A. Identification of T helper type 1-"like, Foxp3 + regulatory T cells in human autoimmune disease. *Nat. Med.* **17**, 673–675 (2011).

66. McClymont, S. A. *et al*. Plasticity of Human Regulatory T Cells in Healthy Subjects and Patients with Type 1 Diabetes. *J. Immunol.* **186**, 3918–3926 (2011).

67. Rubtsov, Y. P. *et al*. Stability of the regulatory T cell lineage in vivo. *Science (80-. ).* **329**, 1667–1671 (2010).

68. Daniel, V., Trojan, K., Adamek, M. & Opelz, G. IFNβ+ Treg in-vivo and in-vitro represent both activated nTreg and peripherally induced aTreg and remain phenotypically stable in-vitro after removal of the stimulus. *BMC Immunol.* **16**, 1–13 (2015).

69. Lucca, L. E. *et al*. TIGIT signaling restores suppressor function of Th1 Tregs. *JCI Insight* **4**, (2019).

70. Yi, G. *et al.* Single-cell RNA-seq unveils critical regulators of human FOXP3+ regulatory T cell stability. *Sci. Bull.* (2020) doi:10.1016/j.scib.2020.01.002.

71. Yu, X. *et al.* The surface protein TIGIT suppresses T cell activation by promoting the generation of mature immunoregulatory dendritic cells. *Nat. Immunol.* **10**, 48–57 (2009).

72. Koch, M. A. *et al*. The transcription factor T-bet controls regulatory T cell homeostasis and function during type 1 inflammation. *Nat. Immunol.* **10**, 595–602 (2009).

73. Piao, W. *et al*. Regulatory T Cells Condition Lymphatic Endothelia for Enhanced Transendothelial Migration. *Cell Rep*. **30**, 1052-1062.e5 (2020).

74. Broadley, I., Pera, A., Morrow, G., Davies, K. A. & Kern, F. Expansions of cytotoxic CD4+CD28- T cells drive excess cardiovascular mortality in rheumatoid arthritis and other

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

chronic inflammatory conditions and are triggered by CMV infection. *Front. Immunol.* **8**, 1–10 (2017).

75. O'Neill, R. E. *et al.* T Cell–Derived CD70 Delivers an Immune Checkpoint Function in Inflammatory T Cell Responses. *J. Immunol.* **199**, 3700–3710 (2017).

76. Lie Atemezem, A.□!, Mbemba, E., Vassy, R., Slimani, H. & Gattegno, L. *Human α 1-acid glycoprotein binds to CCR5 expressed on the plasma membrane of human primary macrophages. Biochem. J* vol. 356 (2001).

77. Slimani, H. *et al.* Interaction of RANTES with syndecan-1 and syndecan-4 expressed by human primary macrophages. *Biochim. Biophys. Acta - Biomembr.* **1617**, 80–88 (2003).

78. Raport, C. J., Gosling, J., Schweickart, V. L., Gray, P. W. & Charo, I. F. Molecular cloning and functional characterization of a novel human CC chemokine receptor (CCR5) for RANTES, MIP-1β, and MIP-1α. *J. Biol. Chem.* **271**, 17161–17166 (1996).

79. Cells, C. D. T. *et al.* Identification the Major HIV-Suppressive Factors Produced by. *Science (80-. ).* (1987).

80. Hertoghs, K. M. L. *et al.* Molecular profiling of cytomegalovirus-induced human CD8+ T cell differentiation. *J. Clin. Invest.* **120**, 4077–4090 (2010).

81. Billiau, A. & Matthys, P. Interferon-γ: A historical perspective. *Cytokine and Growth Factor Reviews* vol. 20 97–113 (2009).

82. Whitmire, J. K. Induction and function of virus-specific CD4+ T cell responses. *Virology* vol. 411 216–228 (2011).

83. Vieira Braga, F. A., Hertoghs, K. M. L., van Lier, R. A. W. & van Gisbergen, K. P. J. M. Molecular characterization of HCMV-specific immune responses: Parallels between CD8+ T cells, CD4+ T cells, and NK cells. *Eur. J. Immunol.* **45**, 2433–2445 (2015).

84. Kärre, K., Ljunggren, H. G., Piontek, G. & Kiessling, R. Selective rejection of H-2-deficient

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

lymphoma variants suggests alternative immune defence strategy. *Nature* **319**, 675–678 (1986).

85. Lam, J. K. P. *et al*. Emergence of CD4+ and CD8+ polyfunctional T cell responses against immunodominant lytic and latent EBV antigens in children with primary EBV infection. *Front. Microbiol*. **9**, 1–13 (2018).

86. Meckiff, B. J. *et al*. Primary EBV Infection Induces an Acute Wave of Activated Antigen-Specific Cytotoxic CD4 + T Cells . *J. Immunol*. **203**, 1276–1287 (2019).

87. Long, H. M. *et al*. Mhc ii tetramers visualize human cd4+t cell responses to epstein-barr virus infection and demonstrate atypical kinetics of the nuclear antigen ebna1 response. *J. Exp. Med*. **210**, 933–949 (2013).

88. Suvas, S., Kumaraguru, U., Pack, C. D., Lee, S. & Rouse, B. T. CD4+CD25+ T cells regulate virus-specific primary and memory CD8+ T cell responses. *J. Exp. Med*. **198**, 889–901 (2003).

89. Zhang, J., Liu, H. & Wei, B. Immune response of T cells during herpes simplex virus type 1 (HSV-1) infection. *Journal of Zhejiang University: Science B* vol. 18 277–288 (2017).

90. Wherry, E. J., Barber, D. L., Kaech, S. M., Blattman, J. N. & Ahmed, R. *Antigen-independent memory CD8 T cells do not develop during chronic viral infection*. vol. 101 www.pnas.orgcgidoi10.1073pnas.0407192101 (2004).

91. Fuller, M. J. *et al*. Cutting Edge: Emergence of CD127 high Functionally Competent Memory T Cells Is Compromised by High Viral Loads and Inadequate T Cell Help . *J. Immunol*. **174**, 5926–5930 (2005).

92. Koesters, S. A. *et al*. IL-7Rα expression on CD4+ T lymphocytes decreases with HIV disease progression and inversely correlates with immune activation. *Eur. J. Immunol*. **36**, 336–344 (2006).

93. Erin C. Dowd, M.D.a, Michael J. Frank, Ph.D.b, Anne Collins, Ph.D.c, James M. Goldd, and Deanna M. Barch, P. D. . HIV-specific CD8 T cells express low levels of IL-7Rα: Implications for HIV-specific T cell memory. *Physiol. Behav*. **176**, 139–148 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

94. Lang, K. S. *et al.* Inverse correlation between IL-7 receptor expression and CD8 T cell exhaustion during persistent antigen stimulation. *Eur. J. Immunol.* **35**, 738–745 (2005).

95. Wang, L. *et al.* A Comprehensive Analysis of the T and B Lymphocytes Repertoire Shaped by HIV Vaccines. *Front. Immunol.* **9**, 2194 (2018).

96. Guo, X. *et al.* CNSA: a data repository for archiving omics data. *Database (Oxford).* **2020**, 1–6 (2020).

97. Chen, F. *et al.* CNGBdb; China National GeneBank DataBase. **42**, 799–809 (2020).

98. 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 (2015).

99. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

100. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

101. Pizzolato, G. *et al.* Single-cell RNA sequencing unveils the shared and the distinct cytotoxic hallmarks of human TCRVδ1 and TCRVδ2 γδ T lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 11906–11915 (2019).

102. Szabo, P. A. *et al.* Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat. Commun.* **10**, (2019).

103. Zhao, Y. *et al.* Single-cell transcriptomic landscape of nucleated cells in umbilical cord blood. *Gigascience* **8**, 1–15 (2019).

104. Carnero Contentti, E., Farez, M. F. & Correale, J. Mucosal-Associated Invariant T Cell Features and TCR Repertoire Characteristics During the Course of Multiple Sclerosis. *Front. Immunol.* **10**, 1–17 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

105. Kurioka, A., Walker, L. J., Klenerman, P. & Willberg, C. B. MAIT cells: new guardians of the liver. *Clin. Transl. Immunol.* **5**, e98 (2016).

106. Fan, X. *et al.* Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol.* **16**, (2015).

107. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).

108. Zheng, C. *et al.* Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* **169**, 1342-1356.e16 (2017).

109. Szabo, P. A. *et al.* Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat. Commun.* **10**, (2019).

110. Hashimoto, K. *et al.* Single-cell transcriptomics reveals expansion of cytotoxic CD4 T cells in supercentenarians. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 24242–24251 (2019).

111. Among, T., Clusters, G. & Yu, G. clusterProfiler□: an R Package for Comparing Biological. **16**, 284–287 (2012).

112. Trapnell, C. *et al.* letters The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, (2014).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

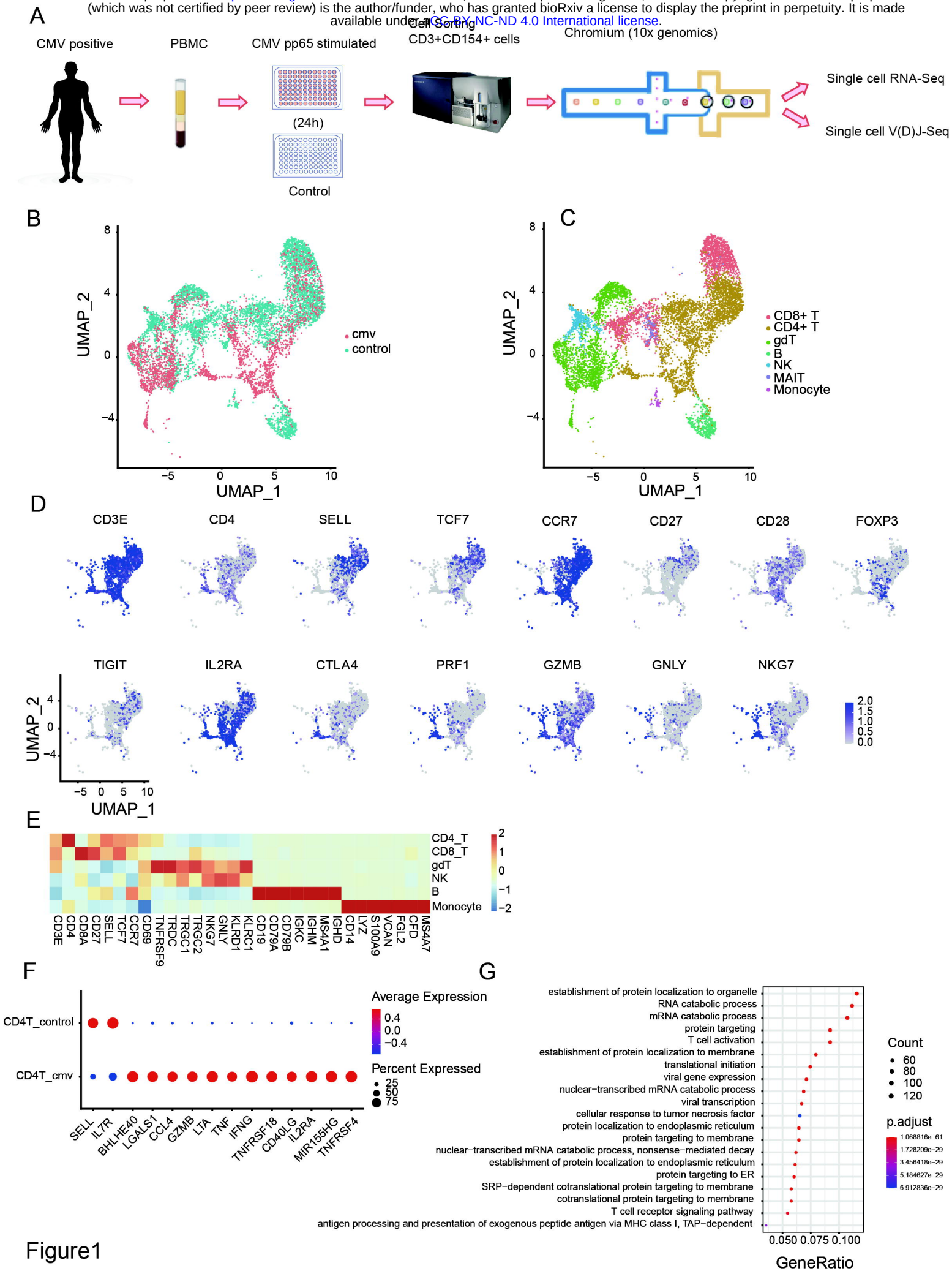
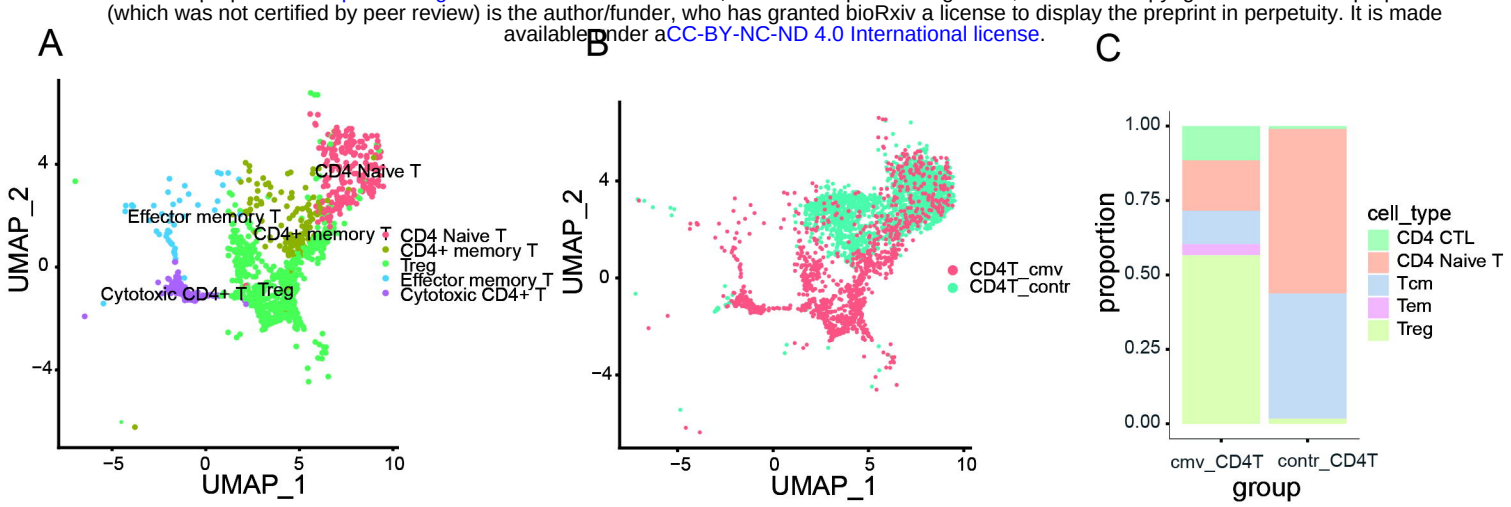



Figure1

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



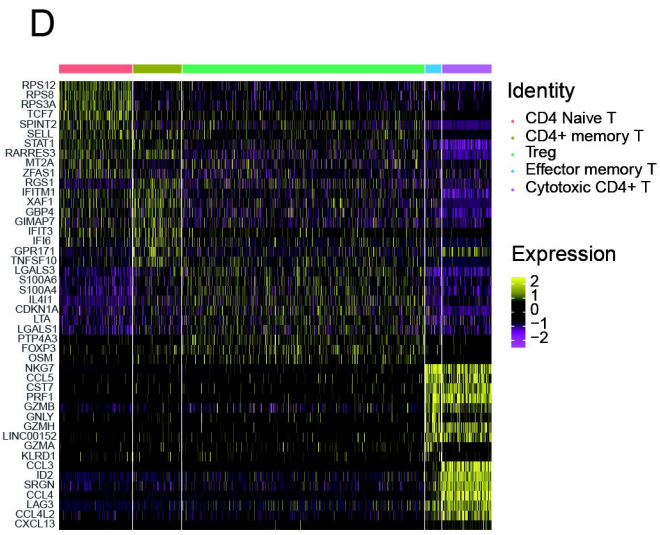




Figure2

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

E

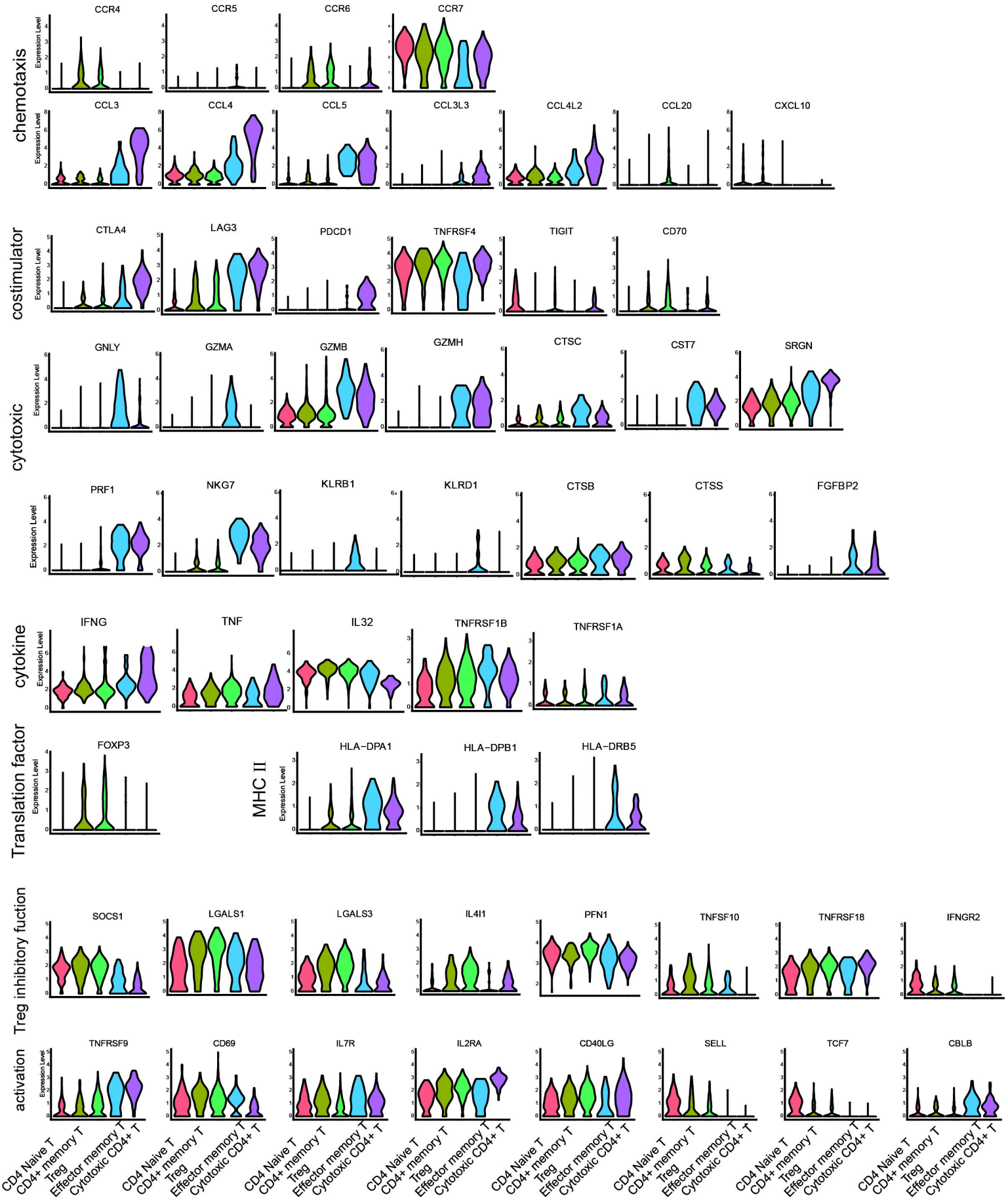
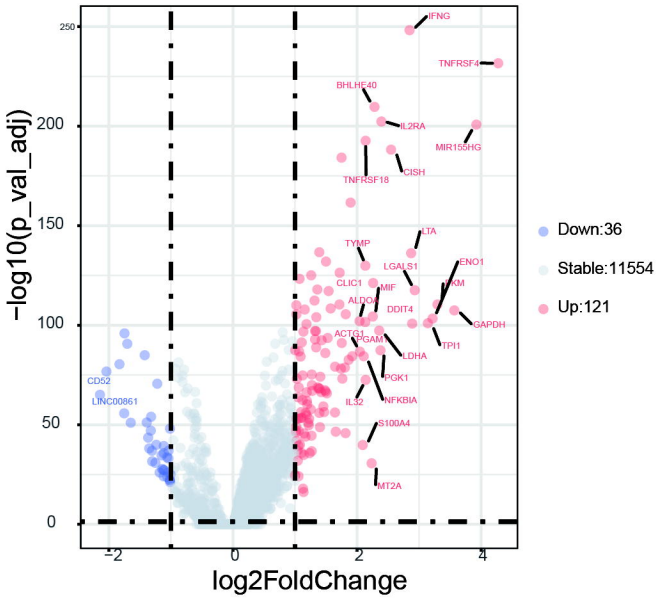


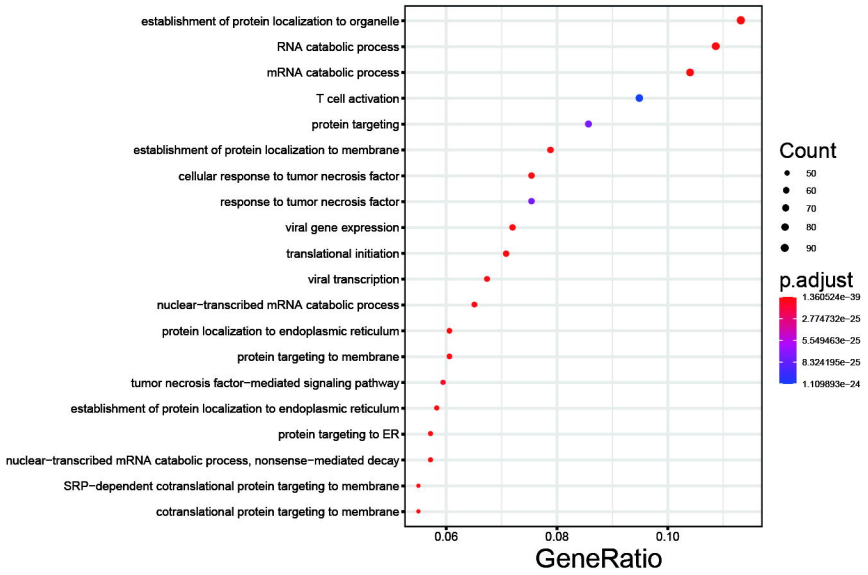

Figure2 E

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

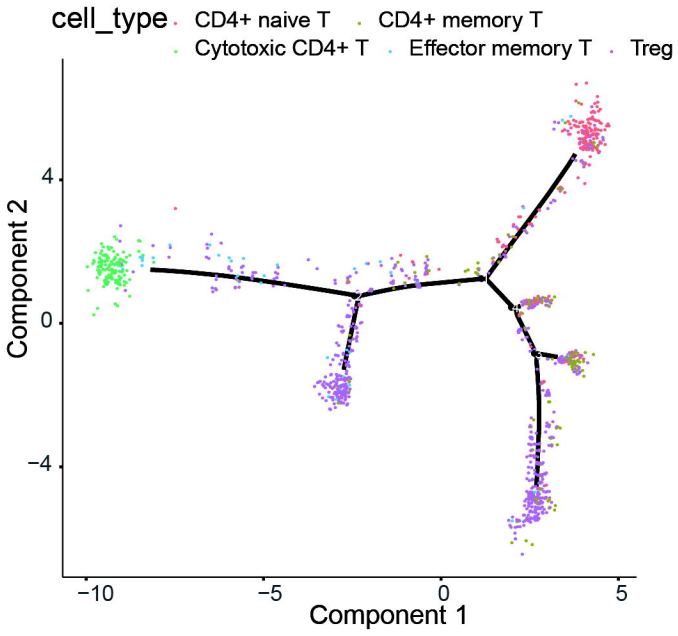
A




B




C




D

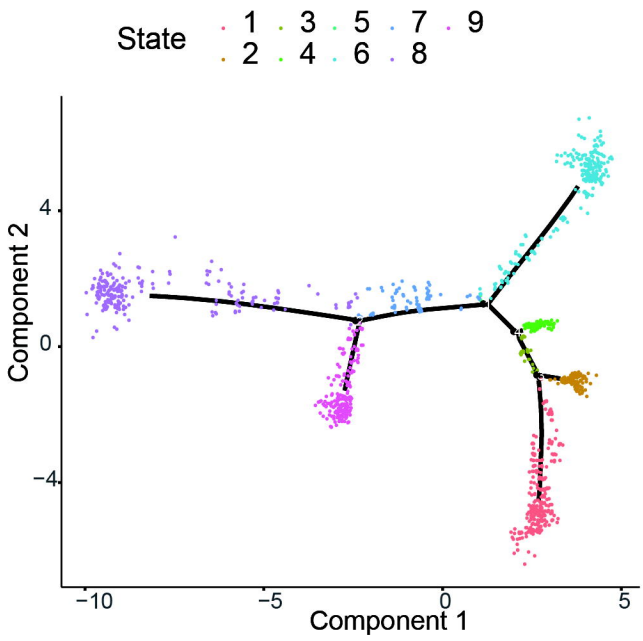



Figure3

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

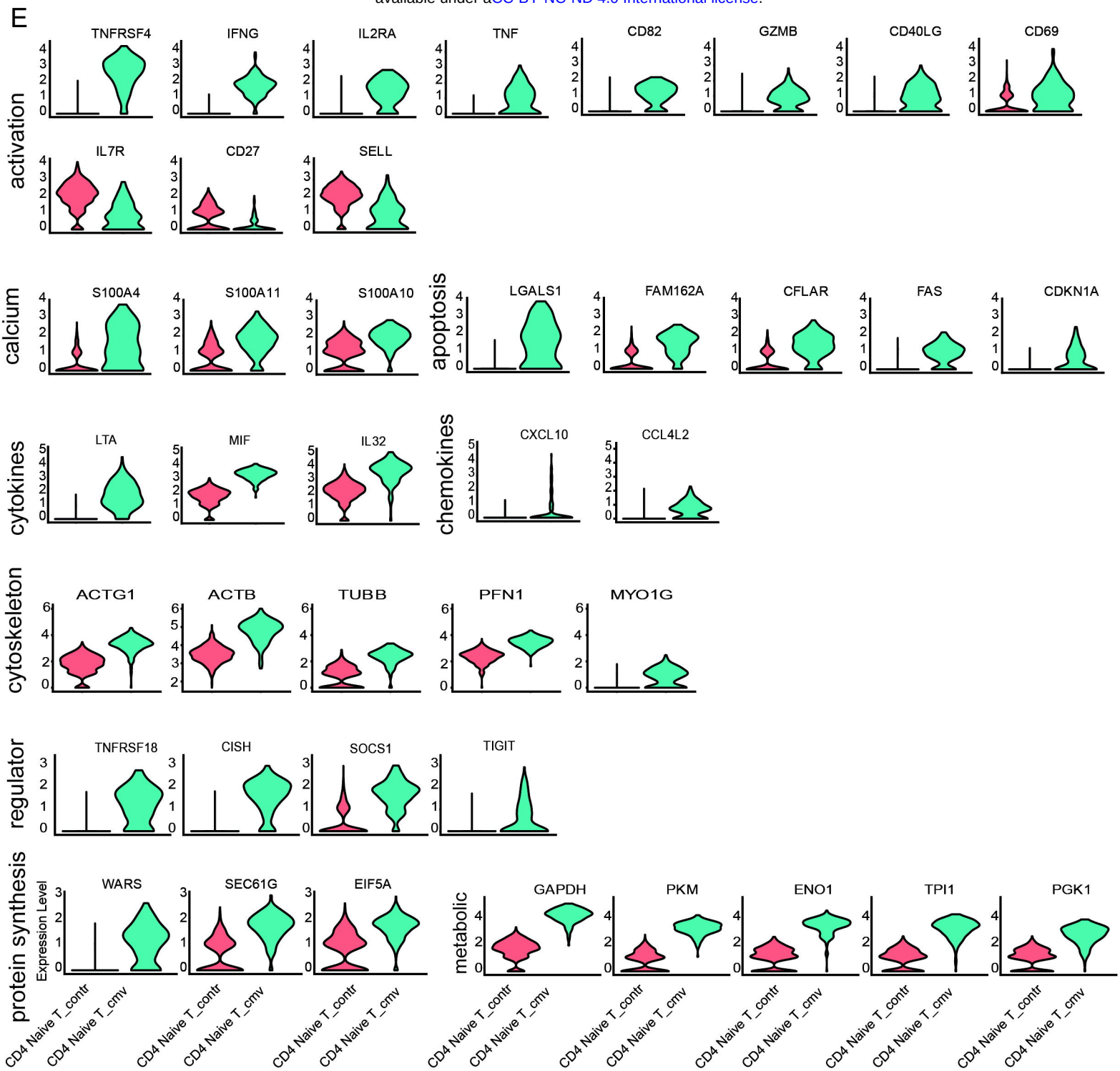
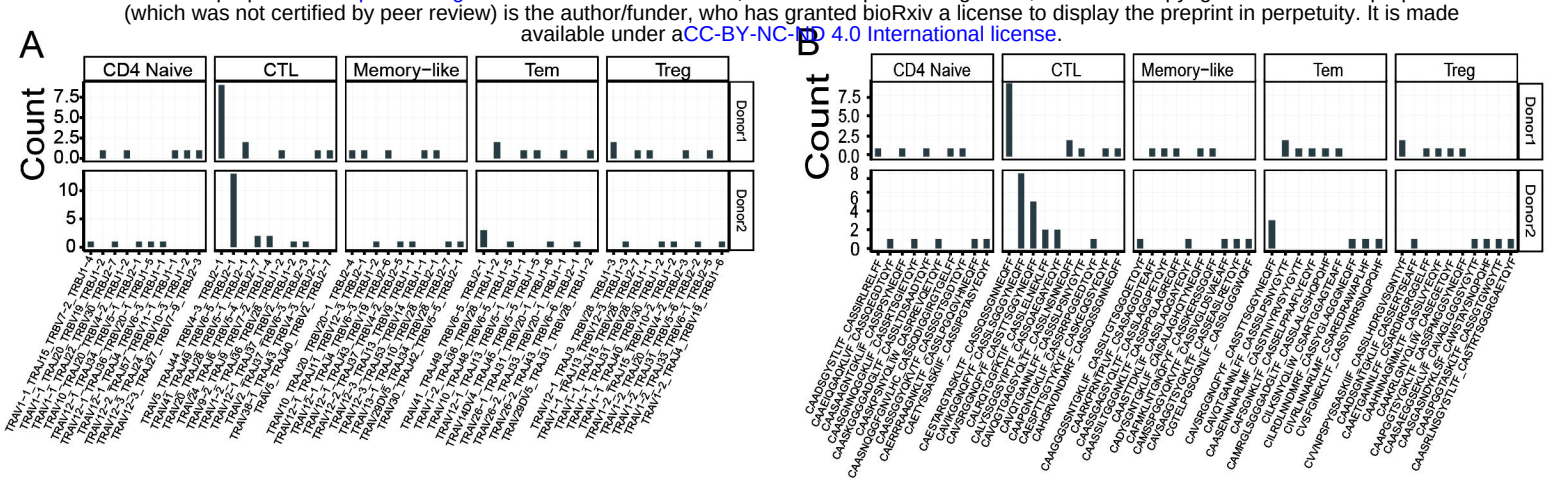
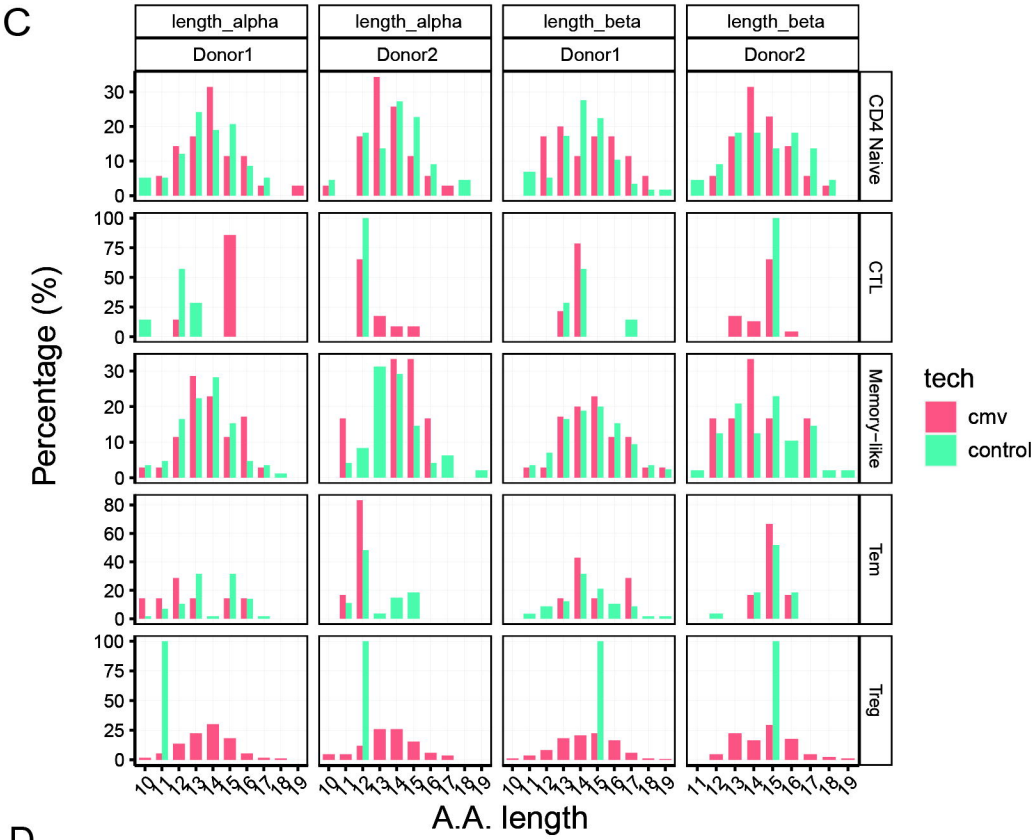



Figure3 E

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.30.268433; this version posted August 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



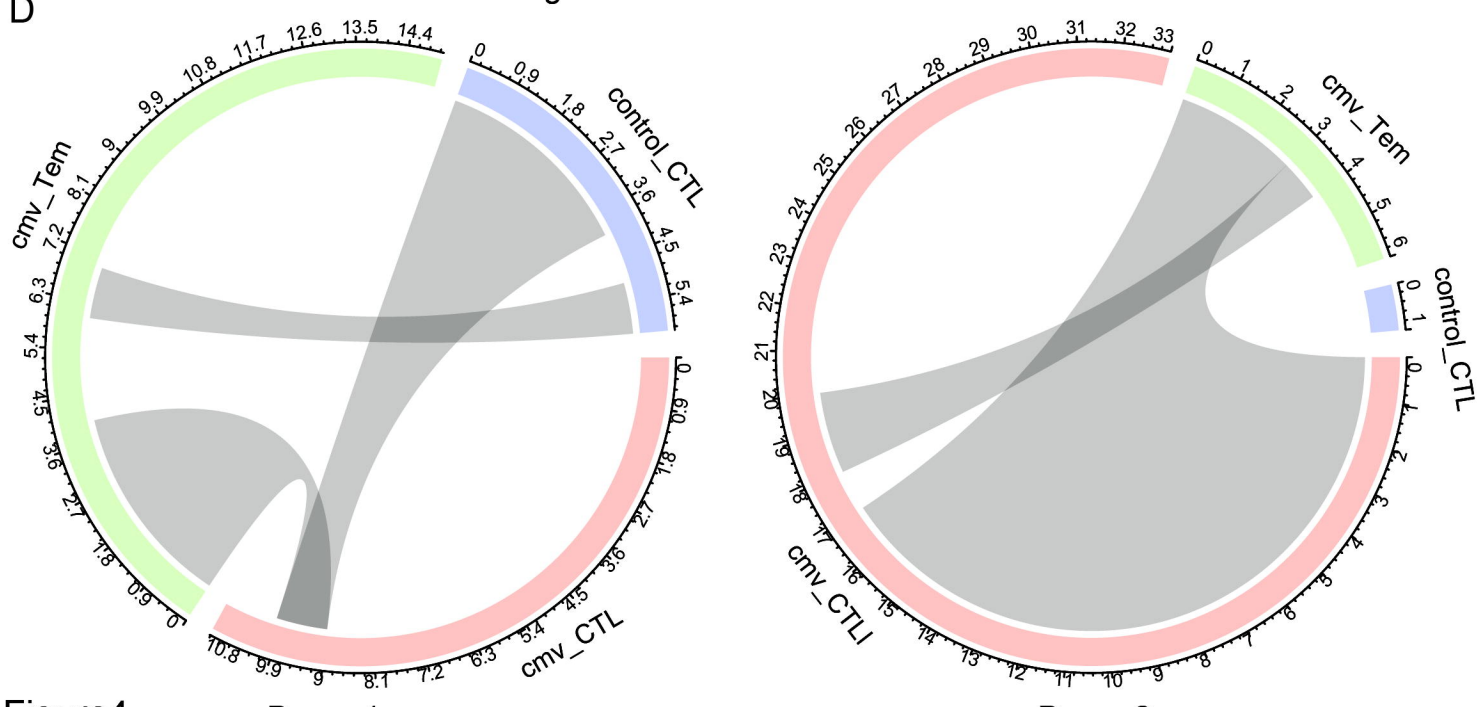




Figure4