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

3 Menghua Lyu^{1,2*}, Shiyu Wang^{1,2*}, Kai Gao^{1,2}, Longlong Wang^{1,2}, Bin Li^{2,3}, Lei Tian^{1,2#}

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

5 ² BGI-Shenzhen, Shenzhen, China 518083

6 ³ Shanghai Institute of Immunology, Shanghai JiaoTong University School of Medicine, Shanghai
7 200025, China

8

9 *These authors contributed equally to this work.

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

11 Abstract

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

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

25 Introduction

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

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 repertoire¹ and had a deleterious effect on immune senescence and health outcomes in the elderly². 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 pools³⁻⁵, 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 only⁶. 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 IFN γ and TNF⁷, to promote the activation of CD8+ T cells, recruit innate immune cells including natural killer and monocytes to inflammatory sites, and directly inhibit virus replication⁸. 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

56 pathway¹¹. This pathway is based on the recognition of CTL to target cells in an MHC-II
 57 dependent manner¹², where the MHC-II is upregulated in epithelial cells upon CMV infection.
 58 Although there have been advances in understanding CD4-CTL function in CMV infection, their
 59 derivation is still unclear. Common views developed from other infection diseases indicate the
 60 origination of CD4-CTLs from effector cells. Recently, evidences^{13,14,15} from studies on
 61 transcriptome factors suggested these cells can also differentiate from naïve cells directly. This
 62 indicates that it may be possible to prepare CD4-CTL for immune therapy from naïve cells.

63 Although some studies tried to unveil Treg function in CMV infection, it is still controversial to
 64 conclude. In humans, Treg cells from CMV-seropositive individuals attenuated the proliferation of
 65 autologous CD8+ T cells and, to a lesser extent, CD4+ T cells in response to CMV virus ex vivo
 66 stimulation using PD-1 pathway¹⁶. However, in CMV reactivating patients received HSCT, CMV
 67 reactivation does not correlate with the numerical reconstitution of CD4+CD25^{high}CD127⁻ Tregs,
 68 and conventional T cells in these patients expressed high level of Ki67 indicating that their
 69 activation function is unimpaired¹⁷. Selectively removing Treg in animal models is a classical
 70 method to verify Treg function at infectious situations¹⁸, and have been used to identify the
 71 negative function of Treg in some anti-viral immunities. However, these experiments failed to
 72 conclude Treg function in CMV infection. In murine, eight months after infection, Treg cells
 73 deletion decreased MCMV reaction in the spleen but enhance the reactivation in the salivary gland
 74 ¹⁹. By analyzing the cell composition and cytokine secretion, the Treg deletion was found to
 75 expanded CD4+ T cells in these two tissues, excepting of an increased production of IL-10 in the
 76 salivary gland which is inferred to disturb the conventional T cells function in the salivary gland
 77 and be in charge of the enhancing viral replication.

78 T cell receptor (TCR)-specific signaling pathway is essential to generate an effective antiviral
 79 immunity. To identify T cells bearing these antigen-specific TCRs, labeling them with antibodies
 80 targeting IFN- γ and other markers, such as CD69, as well as Fluorescence-activated cell sorting
 81 (FACS) are used in previous studies. In the past decade, CD154 is found to be a TCR-singling
 82 specific marker, and labeling it only is able to identify cells which should be labeling with
 83 multiple markers²⁰⁻²². Therefore, methods based on CD154 have been developed and employed by
 84 studies focusing on the antigen specificity of TCRs^{23,24}

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

97 Results

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

99 Circulating antigen-specific T cells are rare in peripheral blood during the latent stage of CMV
100 infection, comprising on average 10% of both the CD4 and CD8 memory compartments in blood¹.
101 To isolate CMV-specific CD4⁺T cells, we cultured PBMCs with and without CMV-pp65 peptides
102 for 24 hours and sorted CD3⁺CD154⁺ cells by flow cytometry²⁴⁻²⁷. We sorted CMV-specific
103 CD4⁺T cells from three donors and then pooled them together for single-cell mRNA-seq and paired
104 VDJ-seq using the 10 × Chromium platform (Figure 1A). Control CD4⁺T cells were acquired by
105 lymphocyte sorting with FSC/SSC. PBMCs from the three donors were subject to bulk RNA-seq
106 for subsequent SNP (single-nucleotide polymorphism) calling, and the sample identity of each cell
107 was deconvoluted with these natural genetic variations²⁸.

108 After stringent quality control and filtering by multiple criteria, RNA-seq data were obtained from
109 2847 and 6493 single cells from the two libraries (CMV and control), detecting a mean of 3041
110 and 1947 genes per cell, respectively. Productive VDJ sequences were obtained for 1271 CMV
111 cells and 3557 control cells. The unsupervised clustering of all cells from CMV and control
112 samples suggested that there are multiple subsets of CMV CD4⁺T cells (Figure1. B, C, D). CMV

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

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

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

128 Subpopulations of CD4+ T cell were further identified with canonical markers (Supplementary
129 Table2) and projected onto the UMAP embeddings (Figure 2A,2B). Stimulated CD4+ T cells were
130 clustered into five groups: naïve (17.01% of total CMV CD4+ T cells), memory-like (11.20%),
131 Treg (56.68%), effector memory (Tem, 3.65%), CTL (11.45%). Unstimulated CD4+ T cells were
132 clustered into four subgroups: naïve (55.30%), memory-like (41.98%), Treg (1.78%), and CTL
133 (0.94%) (Figure 2C). The ratio of unstimulated naïve and memory CD4+ T cells is consistent with
134 previous FACS data³², implying the classification using the scRNA-seq data in this study has a
135 good consistence with that of FACS. Obviously, Treg, CTL, and Tem are significantly enriched in
136 CMV, and Treg is the largest subset. This distribution is further identified in cells from each donor,
137 although cells from donor 3 were limited (1.58% of total CD4+T cells) (Supplement Figure 1 A,
138 1B).

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

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 receptors³⁶ (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 them³⁷. 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 α (*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 Tem²³. 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

170 124 and 36 were upregulated and downregulated with a log2-fold change > 1 respectively. These
 171 124 up-regulated genes comprised a group for encoding the cytokines and chemokines (*LTA*, *MIF*,
 172 *IL32*, *CXCL10* and *CCL4L2*), a group involving in metabolic [e.g., *GAPDH*, *PKM*, *ENO1*, *TPH1*,
 173 *PGK1*] (Figure 3E)^{40,41}, and a group regulating protein synthesis (e.g., *WARS*, *SEC61G*, *EIF5A*).
 174 These phenomena support the cell activation⁴². We also found that naïve T cells increased the
 175 expression of calcium binding proteins encoded by S100 family genes(e.g., *S100A4*, *S100A11* and
 176 *S100A10*) and cytoskeleton related protein(e.g., *ACTG1*, *ACTB*, *TUBB*, *PFN1*, *MYO1G*), which
 177 were reported in response to TCR engagement by antigen^{43,44}. Besides, many regulatory markers
 178 (e.g., *GITR* [*TNFRSF18*], *CISH*, *SOC31*, *TIGIT*) and cell apoptosis regulation markers (e.g.,
 179 *LGALS1*, *FAM162A*, *CFLAR*, *FAS*, *CDKN1A*) were strongly upregulated to maintain immune
 180 balance^{45,46}, although their expression level differs in cells at different differentiation period^{47,48}.
 181 These 36 down-regulated genes including *CD127*, *CD27* and *CD69L*, and were consistent with
 182 previous studies upon T cells activation⁴⁸.

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

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

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

4A). The analyses on CDR3 from TCR alpha chain (TRA) and TRB further verified the clonality in CD4-CTL and Tem. Two expanded clones in CD4-CTL and one in Tem were observed in Donor1, and four in CD4-CTL and one in Tem was observed in Donor2 (Figure 4B). The abnormal features were also exhibited in CDR3 length in naïve, CD4-CTL and Tem, where CDR3 in TRA and TRB from these subsets are shorten in CMV comparing to those in control (Figure 4C). These imply that the repertoire of CD4-CTL and Tem may be largely skewed. To identify clones targeting the same antigens among cell subsets, we used GLIPH2²³ to cluster clones from CMV and control. We identified convergences between memory and Treg cells, and those between Tem and CTLs in CMV. Interestingly, the convergent TCR clone in CD4-CTL from CMV just composed a minor part of CTL from control in Donor 1, and no convergent clone was found between them in donor 2, suggesting that CMV pp65-reactivated CTLs is rare in total CTLs. Furthermore, the dominated clones in CTLs almost derived from Tem cells in Donor 2, while in donor 1, only a small part of CTLs clones was shared with Tem (Figure 4D). This phenomenon indicates that Tem and CD4-CTL cells shared antigen specificity, and Tem may have a tight relationship with CD4-CTL in differentiation. Meanwhile, we did not find either expanded gene combination nor expanded CDR3 clones in Treg, implying Treg may be unspecific for antigens and activated in TCR-independent manners.

Discussion

CD154 is an effective marker to combine with single-cell mRNA sequencing and accomplishment high-throughput analysis of virus antigen-specific T cells cytometry²⁴⁻²⁷. Traditional research methods based on secreted cytokines, such as IFNG or TNF, to testing cmv-specific T cells, proven to perform well⁴⁹⁻⁵¹. However, it was of limited application in combination with sc-mRNA sequencing for a reason of cell damage caused by intracellular staining. Another popular method uses Peptide-MHC multimer to isolate antigen-specific T cells according to specific binding of TCR with pMHC allows detailed TCR and phenotypic analysis of cells using single-cell technologies^{52,53,54}. However, the decrease of TCR expression in activated T cells leads to the preference of relative low antigen-specific T cells bound with tetramer, and the selection of multimer-binding CD4+ T cells may bias our understanding of the phenotype of antigen-specific CD4+ T cells⁵⁵. The fact that 83.8% CD4+ T cells in CMV high express *CD154* (*CD40LG*)

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

231 Bystander activation of CD4+ T cells is less well studied compared with CD8+T cells , but it was
232 demonstrated that unrelated memory CD4 T cells can be activated after a recall tetanus
233 vaccination via bystander activation⁵⁶ and multiple cytokines sharing the common receptor
234 gamma chain can induce CD154/CD40 ligand expression by human CD4+ T lymphocytes via a
235 cyclosporin A-resistant pathway⁵⁷. In our data, CD4+ memory-like T cells are under an
236 environment of IFNG and IL2, and they are prone to be activated by these cytokines. Besides, we
237 find neither clonal expansion nor CDR3 length change in CD4+ memory-like T cell. So, it is
238 difficult to conclude whether these CD4+ memory-like T cells are CMV pp65 antigen-specific.

239 Our data indicated that the CMV-reactivated Treg had heterogeneously inhibitory functions. *LAG3*
240 and *CTLA4* are classical markers of Treg, by which Treg completely contact MHC-II and
241 CD80/CD86 respectively to repress the activation of conventional T cells.
242 Perforin/granzyme-induced apoptosis is a main pathway used by cytolytic cells to kill target
243 cells^{58,59}, and they commonly expressed simultaneously. In our study, Treg highly expressed
244 *GZMB*, *SRGN* (encoding an element protein for maintain granzyme storage), but their *PRFI*
245 expression is limited. One explanation is that a few perforins is enough to facilitate the entrance of
246 granzyme into target cells. The other is that granzyme B can induce cell death in a
247 perforin-independent manner⁶⁰ where it mediate a cleavage of extracellular matrix to reduce the
248 adhesion of immune cells and result in their death. In addition to these classical inhibitory
249 manners, we observed the expression of *LGALS1* and *LGALS3*, encoding Gal-1 and Gal-3
250 respectively, which may also participate in Treg immunosuppressive activity⁶¹. In previous studies,
251 disruption of Gal-1 attenuated the immunoexpressing effect of Treg cells⁶² and Gal-1 from Treg
252 induced the dysfunction of effector T cells and modulated their transient calcium influx⁶³. This
253 regulatory mechanism is not limited to Gal-1 but also employed by Gal-3 in Treg⁶⁴.

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

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 cells⁷¹. 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 Treg⁷² 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 migration⁷³. 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 carriers⁷⁴. 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 Treg⁷⁵. However, more investigations should be taken to unveil CD70 function in Treg.

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

294 CD4⁺ T cells are central organizers in immune responses. CMV-stimulated CD4⁺ T cell secretes
 295 chemokines and effector molecules to recruit and activate or assist other immune cells to
 296 orchestral an antiviral response. For example, HCMV-specific CD4⁺ T cells overall express high
 297 levels of chemokines *CCL4* and the CMV-specific naïve CD4⁺ T cells express high levels of
 298 *CXCL10*. *CCL4* help recruit immune cells such as macrophages, NK cells, monocytes as well as
 299 activated T cells to the site of infection⁷⁶⁻⁷⁹. *CXCL10* help HCMV-specific CD4⁺ T to recruit
 300 HCMV-specific CD8⁺ T cells who high expressed CXCR3 (chemokine receptors of *CXCL10*)⁸⁰ to
 301 the Inflammatory site. In addition, *IFNG* high expressed by CMV-specific CD4⁺ T could help in
 302 direct anti-viral activity and plays an important role in activating immune cells (such as B cells, T
 303 cells, NK cells, and macrophage) and augmenting antigen presentation (such as DC, B cells, and
 304 macrophages)^{81,82}. The high expressed cytotoxic marker (*LTA*, *GZMB*) also implies possible direct
 305 antiviral effects of CD4⁺ T Cells. Since HCMV has acquired extensive mechanisms of immune
 306 evasion, which include the downregulation of MHC class I molecules on infected cells^{83,84}.
 307 whereas, effector CD4⁺ T cells require no MHC class I molecules to eliminate target cells, these
 308 effector lymphocytes may have evolved to support effector CD8⁺ T cells in suppression of HCMV
 309 infection¹. Notably, in our data, almost all (99.1%) CMV-specific CD4⁺ T cell highly expressed
 310 *IFNG* compare with few (1.6%) expressed in control CD4⁺ T with logFC = 4.576. These results
 311 go nicely in line with the fact that CD4 T cells are important in maintain CD8 T-cell activity
 312 during prolonged infection, and their role in helping the antiviral antibody response may also be

313 essential⁸².

314 Unlike other latent-virus specific T cells (such as EBV, HSV), CMV CD4+ T cell does not show
 315 an exhaustion phenotype. Both EBV-specific CD8+ and CD4+ T cells are highly susceptible to
 316 apoptosis, while CMV CD4+ T cells are not. Akin to the cmv CD4+ T cell response, the majority
 317 of activated EBV-specific CD4+ T cells exhibit Th1-type response and express both perforin,
 318 granzyme B, CD107a^{85,86} and lack the lymphoid homing markers CCR7 and CD62L⁸⁷. It was
 319 speculated that EBV-specific CD4+ T cells might highly effective against MHC-II positive EBV
 320 infected B cells. CD4+CD25+ Treg cells are also found in HSV-1 specific CD4+ T cells⁸⁸. HSV⁸⁹:
 321 HSV-1 can establish a latent infection in TG of host, while CD4+ and CD8+ T cells can control
 322 the reactivation of the HSV-1 by surrounding latently infected neurons. CMV CD4+ T cells
 323 exhibit downregulation expression of IL7R, which is also observed in other chronic infections
 324 such as LCMV clone 13 and HIV infection⁹⁰⁻⁹³. This phenomenon is consistent with the previous
 325 reports that T cells fail to re-express the IL7R⁹⁴, when they are continuously stimulated.

326 Methods and Materials

327 PBMC preparation

328 We obtained peripheral blood from three CMV IgG-positive, healthy donors through a research
 329 protocol approved by the Beijing genomics institution-Shenzhen (BGI-Shenzhen) Institutional
 330 Review Board (IRB). Peripheral blood mononuclear cell (PBMCs) were immediately isolated
 331 from blood collected with EDTA blood collection tube by density centrifuge method with
 332 Histopaque-1077 (Sigma, Cat. 10771) within two hours⁹⁵, resuspended in 4°C cryopreservation
 333 medium consisting 90% fetal bovine serum (FBS, HYCLONE, Cat. sh30084.03) and 10%
 334 Dimethyl sulfoxide (DMSO, Sigma, Cat. D4540) and then placed in Mr. Frosty (Thermo
 335 Scientific) in -80°C container. Samples were then moved to liquid nitrogen for a long-time
 336 storage.

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

340 PBMC stimulation

341 Frozen PBMC from liquid nitrogen were immediately thawed in 37°C water and resuspended in
 342 complete medium (RPMI 1640 medium, 10% NEAA and 2% autologous plasma; RPMI 1640 and
 343 NEAA were purchased from ThermoFisher with Cat. 72400120 and Cat. 11140050) to a final
 344 density of 1×10^7 per milliliter (ml). We moved 150 microliter (ul) of cell suspension with three
 345 repetitions to each well in the 96-well U-plate (Falcon) and incubated them at 37°C for two hours.
 346 Then 75ul culture supernatant in each well was replaced by 75ul stimulation medium, and gently
 347 mixed. Cells were cultured in an incubator with 5% CO₂ at 37°C for 24 hours.

348 The stimulation medium included RPMI 1640 medium (without serum), anti-CD28 (2ug/ml,
 349 Clone G28.5, Genetex, Cat. GTX14148), anti-CD40 (2ug/ml, Clone HB14, Miltenyi, Cat.
 350 130-094-133) with/without CMV peptide (1.2 nmol/ml per peptide). To preserve the surface
 351 expression of CD154 on activated T cells, we used anti-CD40 to inhibit the interaction of surface
 352 CD154 with its counterpart CD40 as described in the previous study²². CMV pp65 peptide was
 353 purchased from Miltenyi (Cat. 130-093-438) and diluted in sterile water.

354 Enrichment of CMV pp65-specific T cells

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

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

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

individuals were mixed separately, diluted with PBS to a final concentration ~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 manufacturer's protocol. Sequences within these libraries were ligated with BGISEQ adapters⁷⁰ (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 Archive⁹⁶ of CNGBdb⁹⁷ 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

388 The R package Seurat⁹⁸ 3.1.5 was used to integrate and analyze datasets from CMV and control.
 389 The merged expression matrix was firstly filtered following the Seurat recommendation^{99,100} and a
 390 total of 8671 cells with unique UMI was obtained. Unsupervised clustering was conducted with
 391 Seurat with the parameter res = 0.5, it revealed a total of 16 clusters. We used mRNA biomarkers
 392 (Supplemental Table 2) obtained from recently published articles¹⁰¹⁻¹⁰³ to identify these clusters,
 393 and GSEA analysis was conducted to identify cluster3(see method Gene set enrichment analysis).
 394 TCR repertoire datasets were also combined to identify putative mucosal associated invariant T
 395 (MAIT) cells and $\gamma\delta$ T. Based on the expression of a typical TCR^{104,105}(Supplemental Table 2), a
 396 17th cluster (putative mucosal associated invariant T, MAIT) was emerged from cluster12. Based
 397 on whether they express a TCR a chain, and meanwhile express $\gamma\delta$ biomarkers (CD3D, CD3E,
 398 TRDC, TRGC1, TRGC2), 2111 $\gamma\delta$ T cells were identified and excluded from further analysis.
 399 The above classification was consistent with hematopoietic differentiation and the previously
 400 published t-SNE plots of PBMC scRNA-seq¹⁰⁶⁻¹⁰⁸.

401

402 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
$\gamma\delta$ T	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)

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

403

404 Differential express gene (DEG) analysis

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

411 Quality control metrics and filtering

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

415 Identifying the sample identity of each droplet

416 Firstly, transcriptome of each donors' peripheral blood mononuclear cell was sequenced on the
 417 BGI-SEQ500 platform with the sequencing type SE200. Raw data with 10G per sample was
 418 obtained. Then, we followed the best practices workflows recommend by
 419 GATK([https://gatk.broadinstitute.org/hc/en-us/articles/360035531192-RNAseq-short-variant-](https://gatk.broadinstitute.org/hc/en-us/articles/360035531192-RNAseq-short-variant-discovery-SNPs-Indels-)
 420 [discovery-SNPs-Indels-](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
 421 containing the genotype (GT) to assign each barcode to a specific sample. The VCF file and BAM
 422 files produced by cellranger2 was passed to the demuxlet software to deconvolute sample
 423 identity²⁸. Finally, we assign the best guess of the samples' identity to their corresponding donor
 424 and each "possible" or "ambiguous" droplets as unclear.

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 package¹¹¹(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](ftp://ftp.broadinstitute.org/pub/gsea/gene_sets/c7.all.v7.1.symbols.gmt)).

434

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	CASSRRTGIPTDTQYF	memory
	CASSRRTGIPTDTQYF	Treg
7	CASSLDIQETQYF	memory
	CASSLDIQETQYF	Treg

444

445 Donor 2

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⁺

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

471 Figure 3: CMV pp65 stimulated CD4+ naïve T cells show obvious activation characteristics. (A)
 472 Volcano plot showing the log2FoldChange(x-axis) and $-\log_{10}(p_val_adj)$ (y-axis) for differential
 473 expression gene between CMV and control naïve CD4+ T cells. Genes with log2FoldChange > 1
 474 and adjusted p value < 0.05 are more highly expressed among CMV and highlighted in red and
 475 labeled with their names. Genes with log2FoldChange < - 1 and adjusted p value < 0.05 are
 476 downregulated in CMV and highlighted in blue and labeled with their names. (B) GO analysis of
 477 DEGs between CMV naïve CD4+ T and that of control. The Top 20 enriched GO terms are
 478 ordered on the y-axis. X-axis represents the gene percentage in enriched GO terms. Sizes of the
 479 dots represent the number of genes included in each GO term. The color gradient of dots
 480 represents the adjusted P-values of each enriched GO term. (C) and (D) Pseudotime analysis of the
 481 five CD4+ T cell subsets in CMV. These cells were colored by cell cluster (C) and cell state (D)
 482 respectively. (E) Violin plots of exemplary feature gene expressions of the naïve CD4+ T cells
 483 from CMV and control. These feature genes were classified and labeled with their group name on
 484 the left.

485 Figure 4: TCR repertoire analysis of the five CD4+ T cell subsets from donor 1 and donor 2.
 486 Frequencies of gene combination of TRAV_TRAJ_TRBV_TRDJ (A) and the combination of
 487 CDR3a_CDR3b (B) of five stimulated CD4+ T cell subsets from each donor. X-axis is the ranked
 488 top 10 gene combination of each cluster. Y-axis is the count of each clonotype. (C) Distribution of
 489 amino acid (A.A.) length of TCR alpha and TCR beta of five CD4+ T cell subsets in CMV and
 490 control from the two donors. (D) TCR convergence of each cluster from two donors were
 491 analyzed with GLIPH2²³ (left: donor 1, right: donor 2). Clonotypes are sorted by descending order
 492 with their sequence counts, and the wide on circus indicate the sequence count per clonotype.

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

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

- 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 Nect2 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).

- 543 21. Meier, S., Stark, R., Frentsch, M. & Thiel, A. The influence of different stimulation conditions
544 on the assessment of antigen-induced CD154 expression on CD4+ T cells. *Cytom. Part A* **73**,
545 1035–1042 (2008).
- 546 22. Frentsch, M. *et al.* Direct access to CD4+ T cells specific for defined antigens according to
547 CD154 expression. *Nat. Med.* **11**, 1118–1124 (2005).
- 548 23. Huang, H., Wang, C., Rubelt, F., Scriba, T. J. & Davis, M. M. Analyzing the Mycobacterium
549 tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide
550 antigen screening. *Nat. Biotechnol.* (2020) doi:10.1038/s41587-020-0505-4.
- 551 24. Klinik, M. *et al.* CD154 , a marker of antigen-specific stimulation of CD4 T cells , is associated
552 with response to treatment in patients with chronic HCV infection. *J. Viral Hepat.* (2011)
553 doi:10.1111/j.1365-2893.2010.01430.x.
- 554 25. Frentsch, M. *et al.* Direct access to CD4 + T cells specific for defined antigens according to
555 CD154 expression. *Nat. Methods* **11**, 1118–1124 (2005).
- 556 26. Chattopadhyay, P. K., Yu, J. & Roederer, M. A live-cell assay to detect antigen-specific CD4 +
557 T cells with diverse cytokine profiles. *Nat. Med.* **11**, 1113–1117 (2005).
- 558 27. Kirchhoff, D. *et al.* Identification and isolation of murine antigen-reactive T cells according to
559 CD154 expression. 2370–2377 (2007) doi:10.1002/eji.200737322.
- 560 28. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic
561 variation. *Nat. Biotechnol.* **36**, 89–94 (2018).
- 562 29. Kang, S. & Brown, H. M. Direct Antiviral Mechanisms of. **18**, 1–15 (2018).
- 563 30. Samuel, C. E. Antiviral Actions of Interferons. **14**, 778–809 (2001).
- 564 31. Yu, F. *et al.* The transcription factor Bhlhe40 is a switch of inflammatory versus
565 antiinflammatory Th1 cell fate determination. **215**, (2018).
- 566 32. Lanio, N., Sarmiento, E., Gallego, A. & Carbone, J. Immunophenotypic profile of T cells in

- 567 common variable immunodeficiency: is there an association with different clinical findings?
568 *Allergol. Immunopathol. (Madr)*. **37**, 14–20 (2009).
- 569 33. Davicino, R. C. *et al.* Galectin-1–Driven Tolerogenic Programs Aggravate Yersinia
570 enterocolitica Infection by Repressing Antibacterial Immunity . *J. Immunol.* **199**, 1382–1392
571 (2017).
- 572 34. Probst-Kepper, M., Kröger, A., Garritsen, H. S. P. & Buer, J. Perspectives on regulatory T cell
573 therapies. *Transfus. Med. Hemotherapy* **36**, 302–308 (2009).
- 574 35. Lasoudris, F. *et al.* IL4I1: An inhibitor of the CD8+ antitumor T-cell response in vivo.
575 *European Journal of Immunology* vol. 41 1629–1638 (2011).
- 576 36. Yi, G. *et al.* Single-cell RNA-seq unveils critical regulators of human FOXP3 + regulatory T
577 cell stability. *Sci. Bull.* **65**, 1114–1124 (2020).
- 578 37. Russell, J. H. & Ley, T. J. L YMPHOCYTE -M EDIATED C YTOTOXICITY. (2002)
579 doi:10.1146/annurev.immunol.20.100201.131730.
- 580 38. Oh, D. Y. *et al.* Intratumoral CD4 + T Cells Mediate Anti-tumor Cytotoxicity in Human
581 Bladder Cancer Article Intratumoral CD4 + T Cells Mediate Anti-tumor Cytotoxicity in
582 Human Bladder Cancer. *Cell* 1–14 (2020) doi:10.1016/j.cell.2020.05.017.
- 583 39. Chan, K. *et al.* The Roles of MHC Class II, CD40, and B7 Costimulation in CTL Induction by
584 Plasmid DNA. *J. Immunol.* **166**, 3061–3066 (2001).
- 585 40. MacIver, N. J., Michalek, R. D. & Rathmell, J. C. Metabolic Regulation of T Lymphocytes.
586 *Annu. Rev. Immunol.* **31**, 259–283 (2013).
- 587 41. Maciolek, J. A., Alex Pasternak, J. & Wilson, H. L. Metabolism of activated T lymphocytes.
588 *Current Opinion in Immunology* vol. 27 60–74 (2014).
- 589 42. Ahern, T. & Kay, J. E. Protein synthesis and ribosome activation during the early stages of
590 phytohemagglutinin lymphocyte stimulation. *Exp. Cell Res.* **92**, 513–515 (1975).

- 591 43. Liu, X. *et al.* T cell receptor-induced nuclear factor κ B (NF- κ B) 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).

- 615 54. Fan, H. C., Fu, G. K. & Fodor, S. P. A. Combinatorial labeling of single cells for gene
616 expression cytometry. *Science* (80-.). **347**, 1258367 (2015).
- 617 55. Fuchs, Y. F. *et al.* Gene Expression-Based Identification of Antigen-Responsive CD8+ T Cells
618 on a Single-Cell Level. *Front. Immunol.* **10**, 1–15 (2019).
- 619 56. Whiteside, S. K., Snook, J. P., Williams, M. A. & Weis, J. J. Bystander T Cells: A Balancing
620 Act of Friends and Foes. *Trends Immunol.* **xx**, 1–15 (2018).
- 621 57. Fayen, J. D. & Western, C. Multiple cytokines sharing the common receptor c chain can induce
622 CD154 / CD40 ligand expression by human CD4 + T lymphocytes via a cyclosporin
623 A-resistant pathway. *Immunology* (2001).
- 624 58. Voskoboinik, I., Whisstock, J. C. & Trapani, J. A. Perforin and granzymes: Function,
625 dysfunction and human pathology. *Nat. Rev. Immunol.* **15**, 388–400 (2015).
- 626 59. Raja, S. M. *et al.* Cytotoxic Cell Granule-mediated Apoptosis. *J. Biol. Chem.* **277**,
627 49523–49530 (2002).
- 628 60. Gondek, D. C., Lu, L.-F., Quezada, S. A., Sakaguchi, S. & Noelle, R. J. Cutting Edge:
629 Contact-Mediated Suppression by CD4 + CD25 + Regulatory Cells Involves a Granzyme
630 B-Dependent, Perforin-Independent Mechanism . *J. Immunol.* **174**, 1783–1786 (2005).
- 631 61. Blidner, A. G., Méndez-Huergo, S. P., Cagnoni, A. J. & Rabinovich, G. A. Re-wiring
632 regulatory cell networks in immunity by galectin-glycan interactions. *FEBS Lett.* **589**,
633 3407–3418 (2015).
- 634 62. Baatar, D. *et al.* Tregs utilize β -galactoside-binding protein to transiently inhibit PI3K/p21ras
635 activity of human CD8+ T cells to block their TCR-mediated ERK activity and proliferation.
636 *Brain. Behav. Immun.* **23**, 1028–1037 (2009).
- 637 63. Wang, J. *et al.* Cross-Linking of GM1 Ganglioside by Galectin-I Mediates Regulatory T Cell
638 Activity Involving TRPC5 Channel Activation: Possible Role in Suppressing Experimental
639 Autoimmune Encephalomyelitis. *J. Immunol.* **182**, 4036–4045 (2009).

- 640 64. Ocklenburg, F. *et al.* UBD, a downstream element of FOXP3, allows the identification of
641 LGALS3, a new marker of human regulatory T cells. *Lab. Invest.* **86**, 724–737 (2006).
- 642 65. Dominguez-Villar, M., Baecher-Allan, C. M. & Hafler, D. A. Identification of T helper type
643 1-"like, Foxp3 + regulatory T cells in human autoimmune disease. *Nat. Med.* **17**, 673–675
644 (2011).
- 645 66. McClymont, S. A. *et al.* Plasticity of Human Regulatory T Cells in Healthy Subjects and
646 Patients with Type 1 Diabetes. *J. Immunol.* **186**, 3918–3926 (2011).
- 647 67. Rubtsov, Y. P. *et al.* Stability of the regulatory T cell lineage in vivo. *Science (80-.).* **329**,
648 1667–1671 (2010).
- 649 68. Daniel, V., Trojan, K., Adamek, M. & Opelz, G. IFN β + Treg in-vivo and in-vitro represent
650 both activated nTreg and peripherally induced aTreg and remain phenotypically stable in-vitro
651 after removal of the stimulus. *BMC Immunol.* **16**, 1–13 (2015).
- 652 69. Lucca, L. E. *et al.* TIGIT signaling restores suppressor function of Th1 Tregs. *JCI Insight* **4**,
653 (2019).
- 654 70. Yi, G. *et al.* Single-cell RNA-seq unveils critical regulators of human FOXP3+ regulatory T
655 cell stability. *Sci. Bull.* (2020) doi:10.1016/j.scib.2020.01.002.
- 656 71. Yu, X. *et al.* The surface protein TIGIT suppresses T cell activation by promoting the
657 generation of mature immunoregulatory dendritic cells. *Nat. Immunol.* **10**, 48–57 (2009).
- 658 72. Koch, M. A. *et al.* The transcription factor T-bet controls regulatory T cell homeostasis and
659 function during type 1 inflammation. *Nat. Immunol.* **10**, 595–602 (2009).
- 660 73. Piao, W. *et al.* Regulatory T Cells Condition Lymphatic Endothelia for Enhanced
661 Transendothelial Migration. *Cell Rep.* **30**, 1052-1062.e5 (2020).
- 662 74. Broadley, I., Pera, A., Morrow, G., Davies, K. A. & Kern, F. Expansions of cytotoxic
663 CD4+CD28- T cells drive excess cardiovascular mortality in rheumatoid arthritis and other

664 chronic inflammatory conditions and are triggered by CMV infection. *Front. Immunol.* **8**, 1–10
665 (2017).

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

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

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

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

676 79. Cells, C. D. T. *et al.* Identification the Major HIV-Suppressive Factors Produced by. *Science*
677 (80-.). (1987).

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

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

682 82. Whitmire, J. K. Induction and function of virus-specific CD4+ T cell responses. *Virology* vol.
683 411 216–228 (2011).

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

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

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

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

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

694 87. Long, H. M. *et al.* Mhc ii tetramers visualize human cd4+t cell responses to epstein-barr virus
695 infection and demonstrate atypical kinetics of the nuclear antigen ebna1 response. *J. Exp. Med.*
696 **210**, 933–949 (2013).

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

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

701 90. Wherry, E. J., Barber, D. L., Kaeck, S. M., Blattman, J. N. & Ahmed, R. *Antigen-independent*
702 *memory CD8 T cells do not develop during chronic viral infection.* vol. 101
703 www.pnas.org/cgi/doi/10.1073/pnas.0407192101 (2004).

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

707 92. Koesters, S. A. *et al.* IL-7Ra expression on CD4+ T lymphocytes decreases with HIV disease
708 progression and inversely correlates with immune activation. *Eur. J. Immunol.* **36**, 336–344
709 (2006).

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

- 713 94. Lang, K. S. *et al.* Inverse correlation between IL-7 receptor expression and CD8 T cell
714 exhaustion during persistent antigen stimulation. *Eur. J. Immunol.* **35**, 738–745 (2005).
- 715 95. Wang, L. *et al.* A Comprehensive Analysis of the T and B Lymphocytes Repertoire Shaped by
716 HIV Vaccines. *Front. Immunol.* **9**, 2194 (2018).
- 717 96. Guo, X. *et al.* CNSA: a data repository for archiving omics data. *Database (Oxford)*. **2020**, 1–6
718 (2020).
- 719 97. Chen, F. *et al.* CNGBdb; China National GeneBank DataBase. **42**, 799–809 (2020).
- 720 98. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of
721 single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
- 722 99. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21
723 (2019).
- 724 100. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell
725 transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**,
726 411–420 (2018).
- 727 101. Pizzolato, G. *et al.* Single-cell RNA sequencing unveils the shared and the distinct cytotoxic
728 hallmarks of human TCRV δ 1 and TCRV δ 2 $\gamma\delta$ T lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.*
729 **116**, 11906–11915 (2019).
- 730 102. Szabo, P. A. *et al.* Single-cell transcriptomics of human T cells reveals tissue and activation
731 signatures in health and disease. *Nat. Commun.* **10**, (2019).
- 732 103. Zhao, Y. *et al.* Single-cell transcriptomic landscape of nucleated cells in umbilical cord blood.
733 *Gigascience* **8**, 1–15 (2019).
- 734 104. Carnero Contentti, E., Farez, M. F. & Correale, J. Mucosal-Associated Invariant T Cell
735 Features and TCR Repertoire Characteristics During the Course of Multiple Sclerosis. *Front.*
736 *Immunol.* **10**, 1–17 (2019).

- 737 105. Kurioka, A., Walker, L. J., Klenerman, P. & Willberg, C. B. MAIT cells: new guardians of the
738 liver. *Clin. Transl. Immunol.* **5**, e98 (2016).
- 739 106. Fan, X. *et al.* Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse
740 preimplantation embryos. *Genome Biol.* **16**, (2015).
- 741 107. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat.*
742 *Methods* **14**, 865–868 (2017).
- 743 108. Zheng, C. *et al.* Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell
744 Sequencing. *Cell* **169**, 1342–1356.e16 (2017).
- 745 109. Szabo, P. A. *et al.* Single-cell transcriptomics of human T cells reveals tissue and activation
746 signatures in health and disease. *Nat. Commun.* **10**, (2019).
- 747 110. Hashimoto, K. *et al.* Single-cell transcriptomics reveals expansion of cytotoxic CD4 T cells in
748 supercentenarians. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 24242–24251 (2019).
- 749 111. Among, T., Clusters, G. & Yu, G. clusterProfiler: an R Package for Comparing Biological.
750 **16**, 284–287 (2012).
- 751 112. Trapnell, C. *et al.* letters The dynamics and regulators of cell fate decisions are revealed by
752 pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, (2014).

753

Figure1

Figure2

Figure2 E

Figure3

Figure3 E