Published in final edited form as:

Nat Immunol. 2019 August ; 20(8): 1059–1070. doi:10.1038/s41590-019-0418-x.

Altered differentiation is central to HIV-specific CD4+ T cell dysfunction in progressive disease

Antigoni Morou[1,2], **Elsa Brunet-Ratnasingham**[1,2], **Mathieu Dubé**[1,3], **Roxanne Charlebois**[1], **Eloi Mercier**[4], **Sam Darko**[5], **Nathalie Brassard**[1], **Krystelle Nganou-Makamdop**[5], **Sahaana Arumugam**[5], **Gabrielle Gendron-Lepage**[1,2], **Lifei Yang**[6], **Julia Niessl**[1,2,3], **Amy E. Baxter**[1,2,3,#], **James M. Billingsley**[7], **Premeela A. Rajakumar**[7], **François Lefebvre**[4], **R. Paul Johnson**[7], **Cécile Tremblay**[1,2], **Jean-Pierre Routy**[8], **Richard T. Wyatt**[3,6], **Andrés Finzi**[1,2], **Daniel C. Douek**[5], and **Daniel E. Kaufmann**[1,2,3,*]

[1]Research Centre of the Centre Hospitalier de l'Université de Montréal (CRCHUM), Montreal, Quebec, Canada.

[2]Université de Montréal, Montreal, Quebec, Canada.

[3]Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID), La Jolla, California, USA.

[4]Canadian Centre for Computational Genomics (C3G) - Montréal Node, Quebec, Canada.

[5]Human Immunology Section, Vaccine Research Center, NIAID, NIH, Bethesda, Maryland, USA.

[6]Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, California, USA.

[7]Yerkes National Primate Research Center and Emory University, Atlanta, Georgia, USA.

[8]Chronic Viral Illnesses Service and Division of Hematology, McGill University Health Centre, Montreal, Quebec, Canada.

Abstract

Dysfunction of virus-specific CD4+ T cells in chronic human infections is poorly understood. We performed genome-wide transcriptional analyses and functional assays of CD4+ T cells specific for human immunodeficiency virus (HIV) from HIV-infected people prior and after initiation of antiretroviral therapy (ART). A follicular helper T cell (T_{FH} cell)–like profile characterized HIV-

*Correspondence to: daniel.kaufmann@umontreal.ca.

AUTHOR CONTRIBUTIONS

A.M and D.E.K. designed the studies; A.M., E.B.R, R.C, N.B, S.A, G.G.L, L.Y and P.A.R performed experiments; M.D. provided input on manuscript content and data representation; A.M, E.M, S.D and F.L performed bioinformatics analyses; K.N.M, J.N, A.E.B, J.M.B, R.P.J, R.T.W, A.F and D.C.D provided technical expertise; C.T and J.P.R. obtained institutional review board approval and managed study participant recruitment; A.M and D.E.K. interpreted the data and wrote the paper with all co-authors' assistance.

#Current affiliation: University of Pennsylvania, Perelman School of Medicine, Philadelphia, USA

Life Sciences Reporting Summary.

Further information on research design is available in the Life Sciences Reporting Summary linked to this article.

Competing interests

The authors declare no competing interests

specific CD4+ T cells in viraemic infection. HIV-specific CD4+ T cells from people spontaneously controlling the virus (elite controllers) robustly expressed genes associated with the T_H1, T_H17 and T_H22 subsets of helper T cells. Viral suppression by ART resulted in a distinct transcriptional landscape, with a reduction in the expression of genes associated with T_{FH} cells but persistently low expression of genes associated with T_H1, T_H17 and T_H22 cells compared to the elite controller profile. Thus, altered differentiation is central to the impairment of HIV-specific CD4+ T cells and involves both gain of function and loss of function.

CD4+ T cell responses are critical for durable control of viral replication in chronic infections[1, 2]. Virus-specific CD4+ T cell immunity is of particular interest for HIV infection, which is characterized by functional impairment and destruction of CD4+ T cells. While classical models divided CD4+ T cells into distinct lineages, studies have demonstrated the importance of CD4+ T cell plasticity[3]. Persistent antigen and inflammatory signals cause impairment of antigen-specific responses, a state called immune exhaustion. Virus-specific CD8+ T cell exhaustion has been extensively investigated[4] and represents a *bona fide* cell differentiation program. These studies highlighted the relevance of genome-wide transcriptional studies to understand T cell impairment[5]. Compared to CD8+ T cells, less is known on CD4+ T cell dysfunction. Murine LCMV-specific CD4+ T cells in chronic infection, while exhibiting some characteristics shared with CD8+ T cells, also present distinct features[6, 7], including loss of a T_H1-signature[5] and skewing towards a T follicular helper (T_{FH}) phenotype[8, 9].

Whether findings in mice can be extrapolated to human HIV infection is unclear. Some features of virus-specific CD4+ T cells are shared between both infections: upregulation of co-inhibitory receptors are found in HIV progressors with ongoing viremia (chronic progressors, CP) and chronic LCMV infection[5, 10]. Rare subjects who spontaneously suppress HIV (elite controllers, EC) frequently exhibit robust virus-specific T_H1 responses[11] and strong proliferative capacity, similarly to mice infected with the acute strain of LCMV. However, HIV and LCMV are distinct viruses and there are notable differences between species in terms of T cell differentiation mechanisms, such as T_{FH} generation[12]. An issue of critical clinical relevance is the lack of restoration of effective anti-HIV immunity after suppressive antiretroviral therapy (ART): viral rebound is the rule after cessation of therapy. Whether persistent HIV-specific CD4+ T cell dysfunction on ART contribute to this failed response is an important, yet unresolved, question. The paucity of experimental tools capable of identifying highly heterogeneous antigen-specific CD4+ T cells has hampered the study of HIV-specific CD4+ T cell help. Intracellular cytokine assays (ICS) are of limited sensitivity for many non-T_H1 effector functions, and the use of HLA Class II tetramers in humans is constrained by availability, requirement for pre-defined epitopes and genetic diversity.

To determine key pathways and molecules that link HIV-specific T cell help to viral control, we here performed genome-wide transcriptional analyses and functional assays of HIV-specific CD4+ T cells from HIV-infected humans with diverse viral loads prior to ART initiation and followed a subgroup of them longitudinally after viral suppression on therapy.

RESULTS

Links between HIV-specific CD4 transcriptome profiles and viremia

To define molecular features that discriminate HIV-specific CD4$^+$ T cells in progressive vs. controlled infection, we performed a cross-sectional study of 38 chronically infected people who were untreated at the time of sampling. These included "elite controllers" (EC, HIV plasma viral load <50 vRNA copies/ml), "viremic controllers" (VC, viral load between 50 and 5,000 copies/ml) and "chronic progressors" (CP, viral load ≥5,000 copies/ml) (Participant characteristics: Supplementary Table 1).

We utilized an activation induced marker (AIM) assay to identify CD4$^+$ T cells specific for the Gag protein (hereafter termed HIV-specific CD4$^+$ T cells). *Ex vivo* stimulated HIV-specific CD4$^+$ T cells were identified by the co-upregulation of CD40L and CD69 on their surface after a 9-h stimulation with an HIV Gag peptide pool[13] (Fig. 1a, Supplementary Fig. 1a). Combining two markers enhanced detection of HIV-specific CD4$^+$ T cells by reducing background compared to CD40L alone (Supplementary Fig. 1b,c). This AIM assay overcomes limitations of cytokine-based techniques for detection of virus-specific cells, allows live-cell sorting and captures a broader Ag-specific CD4$^+$ T cell population (Supplementary Fig. 1d). There was no significant difference in the magnitude of HIV-specific CD4$^+$ T cell responses amongst cohorts or correlation with viremia or CD4 count (Fig. 1b, Supplementary Fig. 1e,f).

We reasoned that the AIM assay could reveal the transcriptome modulation in response to cognate antigen-mediated TCR signaling, reflecting the functional state of CD4$^+$ T cells. We thus performed microarray analysis of sorted ex vivo stimulated HIV-specific CD4$^+$ T cells. Pairwise comparisons between cohorts (Fig. 1c) revealed sets of differentially expressed genes (DEGs) between all three cohorts, with extensive dissimilarities between CPs and ECs. Fewer transcriptional differences were detected between the VC and EC groups. The majority of DEGs between VCs and ECs were also observed in the CPs versus ECs comparison (Fig. 1d). CPs and VCs showed the fewest DEGs, suggesting that viremia strongly impacts the transcriptional landscape of HIV-specific CD4$^+$ T cells.

We next used CAMERA (Correlation Adjusted Mean Rank gene set test)[14] enrichment analysis to investigate molecular mechanisms underlying HIV-specific T help alterations associated with disease status. We compared the identified DEGs to the Molecular Signatures Database (MSigDB) C5 collection of curated gene sets, which comprises the Gene Ontology (GO) terms associated with biological processes, molecular functions and cellular components (Fig. 1e). The "Positive regulation of T helper cell differentiation" signature was negatively associated with viremia. Consistently, the "Positive regulation of CD4$^+$ αβ T cell activation" and "Regulation of leukocyte migration" gene sets were significantly underrepresented in CP compared to both VCs and ECs, contrasting with upregulation of the MHC class II protein complex in CPs. Expression of genes associated with protein transport and/or processing was reduced in CPs. The pathways enriched with viremia included processes associated with chromatin modification, mitotic cell cycle, DNA packaging, chromosome segregation, response to stress, and apoptosis.

This approach identifies potential pathways associated with HIV-specific CD4+ T cell dysfunction, and points at a central role for altered T cell differentiation, regulation and activation in this impairment.

HIV-specific CD4 polarization is associated with disease status

Given the potential role for altered T cell differentiation in HIV-specific CD4+ T cell impairment, we determined features of CD4+ T cell differentiation associated with disease status. We compared by CAMERA our sets of DEGs with human T helper polarization signatures from the C7 collection MSigDB and publicly available microarray datasets. While the different conditions used in the generation of these signatures (e.g activated vs unactivated cells) may complicate the interpretation of such associations, we observed notable differences amongst cohorts. Compared to CPs, ECs and VCs displayed enriched T_H1 and T_H17 signatures[15, 16, 17] (Fig. 2a,b,e,f; Supplementary Fig. 2a; Supplementary Tables 2–4). Cardinal genes associated with T_H17 cell differentiation and function were upregulated in ECs (Fig. 2f). In viremic people (CPs, VCs), we noted a striking enrichment of a T_{FH} signature[18] (Fig. 2c,g, Supplementary Fig. 2a) compared to ECs. There was congruent upregulation of several key T_{FH}-associated genes compared to EC with the notable exception of Bcl6 that was instead upregulated in ECs (Fig. 2g, Supplementary Table 5). Although we did not observe a statistically significant enrichment of an LCMV-specific CD4+ T cell exhaustion signature[5] in CPs compared to ECs (Fig. 2d) (which might be due at least in part to the methodological differences), several of the DEGs enriched in CP were well-known exhaustion-associated genes (Fig. 2h, Supplementary Table 6). Several Interferon-stimulated genes (ISGs) upregulated in the LCMV model of exhaustion, such as *IFH1*, *IFIT3* and *IFI44*, were upregulated in ECs.

We next used qRT-PCR to further quantify levels of key DEGs identified by microarrays and selected additional CD4+ T cell genes. Principal component analysis discriminated CPs from ECs, whereas there was more overlap between CPs and VCs (Fig 2i). PC1 values significantly inversely correlated with viremia, showing that changes in the CD4+ T cell profile relate to antigen load (Fig 2j). The qRT-PCR data were consistent with the microarray results (Supplementary Table 7). They confirmed that genes associated with T_{FH} functions and B cell help (*CXCL13*, *SLAMF6*, *CD84*) and exhaustion (*TIGIT*, *CD200*, *LAG3*, *TOX*) were enriched in CPs. This contrasted with genes related to T_H17 and T_H22 functions (*IL17F*, *IL22*) and the interferon-induced chemokines (*CXCL9*, *CXCL10*, *CXCL11*), which were enriched in ECs (Fig 2i). Expression of selected transcription factors was consistent with these observations (Supplementary Fig 2b): ECs expressed significantly more *TBX21* and *AHR*, genes respectively coding for T_H1 and T_H22 master transcription factors, and there was in ECs a trend for enhanced expression of *RORC*, a master transcription factor of T_H17 differentiation. The behavior of the master T_{FH} transcription factor *BCL6* contrasted to most T_{FH} associated genes: *BCL6* was significantly upregulated in ECs compared to CPs in contrast to MAF, another transcription factor critical for T_{FH} differentiation, which was highly expressed in CPs. Helios (*IKZF2*), a transcription factor overexpressed in LCMV-specific CD4+ T cells in chronic infection, was also upregulated in CPs compared to ECs. Examining correlations between gene expression revealed positive and negative associations of genes, many not anticipated by classical dichotomic models of

T_H differentiation. These results are consistent with a complex spectrum of differentiation and functional profiles of HIV-specific $CD4^+$ T cells (Supplementary Fig 2c,d).

These data show higher expression of a T_{FH}-like signature in HIV-specific $CD4^+$ T cells in progressive infection, contrasting with enriched T_H1 and T_H17 signatures in subjects achieving spontaneous viral control.

Upregulation of classical T_{FH}-associated genes in $CXCR5_{neg}$ HIV-specific CD4 of CPs

We next investigated whether the enrichment of the T_{FH} signature in HIV-specific $CD4^+$ T cells of CPs compared to ECs was due to a higher fraction of HIV-specific circulating T_{FH} cells (cT_{FH}). We used the CD69/CD40L AIM assay to assess the percentages of HIV-specific $CXCR5^+CD45RA^-$ memory $CD4^+$ T cells ($CXCR5_{mem}$) in both cohorts (Fig. 3a and Supplementary Fig. 3a,b). As PD-1 expression on HIV-specific $CD4^+$ T cells correlates with viremia[19], PD-1 was omitted as a T_{FH} marker to avoid bias towards CPs. We observed only a trend for higher fractions of HIV-specific $CXCR5_{mem}$ in CPs than ECs, with considerable overlap between groups (Fig. 3b). In contrast, single subject GSEA (ssGSEA) scores showed a significant enrichment of the T_{FH} signature in CPs compared to ECs (Fig. 3c), consistent with the cohort-based CAMERA analyses of Fig. 2c,g. The T_{FH} signature enrichment by ssGSEA correlated directly with viremia (Fig. 3d). These data suggest that differences in percentage of $CXCR5_{mem}$ alone are not sufficient to explain the markedly enriched T_{FH} signature observed in HIV-specific $CD4^+$ T cells of CPs.

To define the cell subsets that contributed to the T_{FH} transcriptional signature observed in CPs, we applied microarray analysis to paired live-sorted subsets of $CXCR5_{mem}$ and $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells from CPs and ECs (Fig.3e; gating strategy Supplementary Fig. 3a). CXCR5 mRNA expression was consistent with protein expression (Fig. 3h). Differences in gene expression between CPs and ECs were more significant in the $CXCR5_{neg}$ than in the $CXCR5_{mem}$ subsets of HIV specific $CD4^+$ T cells (Fig. 3e). CPs displayed a marked enrichment of the T_{FH} signature not only in the $CXCR5_{mem}$ compartment, but also in the $CXCR5_{neg}$ HIV-specific $CD4^+$ T cell subset (Fig. 3f). We confirmed expression of several markers classically associated with T_{FH} cells in $CXCR5_{mem}$ and $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells at the mRNA level by qRT-PCR and at the protein level by flow cytometry. Consistent with the microarray results, upregulation of the TCR co-receptors PD1, TIGIT, CD200 and ICOS on $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells clearly differentiated CPs from ECs (Fig. 3g and Supplementary Fig. 3c). Surface markers *SLAMF6* and *CD84* (both associated with help to B cells) and *BTLA* (highly expressed in T_{FH} cells; Fig. 3h) were significantly higher in $CXCR5_{neg}$ $CD4^+$ T cells of CPs compared to ECs, while the mRNA expression of *BCL6* and *TCF7* did not differ in $CXCR5_{neg}$ $CD4^+$ T cells between the two groups. Transcriptional expression of *PRDM1* and *MAF* was significantly higher in $CXCR5_{neg}$ $CD4^+$ T cells of CPs compared to ECs (Fig. 3i).

In contrast to ECs, CXCR5 expression in CPs did not clearly delineate cells with a T_{FH} transcriptional profile: *ICOS*, *SLAMF6*, *BTLA*, *CD84*, *TIGIT*, *CD200* and the T_{FH}-associated transcription factor *cMAF* mRNA levels in $CXCR5_{neg}$ cells were equal or even superior to their $CXCR5_{mem}$ autologous counterparts (Fig. 3ghi, Supplementary Fig. 3c).

These data demonstrate the atypical upregulation of "T_{FH}-like" signature in the $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells of CPs compared to ECs.

Robust expression of T_{FH} cytokines in $CXCR5_{neg}$ HIV-specific CD4 of CP

We next sought to elucidate if the enriched T_{FH} signature in $CXCR5_{mem}$ and $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells of CPs compared to ECs is congruent with the enhancement of some T_{FH} functions and determine the phenotype of these effector cells. We assessed mRNA expression of the T_{FH} cytokines CXCL13 and IL-21 at single-cell resolution using a sensitive flow cytometric RNA Fluorescent *In Situ* Hybridization (RNA Flow FISH), as previously described[20, 21]. This assay is superior to standard ICS for detection of IL-21-producing cells[20]. There were increased frequencies of *CXCL13* $mRNA^+$ and *IL21* $mRNA^+$ HIV-specific $CD4^+$ T cells in CPs compared to ECs (Fig. 4a,b). This was a distinct feature of HIV-specific responses, as the expression after stimulation with the superantigen SEB was similar in both cohorts (Fig. 4c) and CMV pp65-specific $CD4^+$ T cells expressed less *CXCL13* and *IL21* than HIV-specific $CD4^+$ T cells in paired comparisons within the same humans (Supplementary Fig. 4a). As CD40L is not expressed on all Ag-specific $CD4^+$ T cells upon activation, we confirmed that the differential expression of *CXCL13* and *IL21* between CPs and ECs hold true without combining the RNA Flow FISH and AIM assays and without pre-gating on $CD40L^+$ cells (Supplementary Fig. S4b–i). Largely distinct $CD4^+$ T cell populations produced *CXCL13* and *IL21* mRNAs at the time point examined (Fig. 4d,e). We confirmed the increased CXCL13 production in CPs compared to ECs at the protein level, by measuring CXCL13 concentration by ELISA in the supernatant of CD8-depleted PMBCs after 48h stimulation with HIV Gag (Supplementary Fig. 5a). *CXCL13*$^+$ and *IL21*$^+$ cells expressed high levels of CXCR5 protein to compared to total HIV-specific $CD4^+$ T cells in ECs. In contrast, the levels of CXCR5 expression on *CXCL13*$^+$ and *IL21*$^+$ cell were comparable to the total population in CPs (Fig 4f). This gain of T_{FH}-like functions by $CXCR5_{neg}$ HIV-specific $CD4^+$ T cell in CPs was further apparent when we compared the ratios of *IL21*$^+$ and *CXCL13*$^+$ responses between the $CXCR5_{mem}$ and $CXCR5_{neg}$ subsets (Fig. 4g). As the inhibitory receptors PD-1 and TIGIT are markers of T_{FH} cells[12] and may also be associated with exhaustion, we examined their expression on *CXCL13* $mRNA^+$ and *IL21* $mRNA^+$ HIV-specific $CD4^+$ T cells. The majority of these cytokine-expressing cells were PD-1^+/$TIGIT^+$, particularly in CPs (Fig. 4h,i, Supplementary Fig. 5b,c.). Thus, the expression pattern of the inhibitory co-receptors PD-1 and TIGIT was consistent with T_{FH} skewing and may not imply exhaustion of these responses.

Finally, we explored possible links between HIV-specific $CD4^+$ T cell help and antibody (Ab) responses. We assessed plasma levels of p24-specific Abs by ELISA and Biolayer Interferometry (BLI) binding analysis. The results obtained by the two methods were highly correlated (Supplementary Fig. 5d). Although p24-specific Ab concentrations in the CP and EC groups were overall similar (Fig. 4j), we observed a significant correlation between the magnitude of HIV-specific $CD4^+$ T cell responses and p24-specific Abs, which was maintained when only the CP group was considered (Fig 4k). There was no significant association between $CXCR5_{mem}$ responses and p24-specific antibodies (Fig. 4l). However, there was a significant correlation between $CXCR5_{neg}$ HIV-specific $CD4^+$ T cell frequencies

and Ab responses in CPs, but not ECs (Fig. 4m), which might suggest that the $CXCR5_{neg}$ subset has T helper function to B cells *in vivo*.

These data demonstrate that dysregulated $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells in progressive infection gain functions and phenotypic features classically attributed to T_{FH} cells.

Mucosal immunity-related cytokines in HIV-specific CD4 of ECs

As we observed an enriched T_H17 signature in HIV-specific $CD4^+$ T cells in ECs compared to CPs, we characterized the expression of IL-17F and IL-22 at the single-cell level. RNA Flow FISH enabled superior detection of these cytokines in HIV-specific $CD4^+$ T cells compared to ICS (Fig. 5a,b). We observed more robust *IL17F* and *IL22* HIV-specific $CD4^+$ T cell responses in ECs compared to CPs (Fig 5c). In contrast, *IL17F* and *IL22* expression patterns in SEB- or CMV activated $CD4^+$ T cells did not differ between the groups (Fig. 5d, Supplementary Fig. S4a). We confirmed the strong IL-17F, IL-22 and IL-17A secretion by HIV-specific $CD4^+$ T cells in ECs at the protein level, in contrast to low or undetectable production in CPs (Fig. 5e, Supplementary Fig. 5e,f,g).

The *IL17F* $mRNA^+$ and *IL22* $mRNA^+$ subsets constituted two partially overlapping cell populations (Fig. 5f,g). In both CPs and ECs, these cells preferentially expressed CCR6, a chemokine receptor critical for recruitment of T_H17 and T_H22 cells to mucosal tissues (Fig. 5h). Expression of CXCR3, a marker of T_H1 cells also expressed on T_H1-T_H17 cells of mixed properties that are expanded in some inflammatory diseases[17], was low on *IL17F* and *IL22* $mRNA^+$ of both CPs and ECs (Fig 5i, Supplementary Fig. 5h). Performing the RNA Flow FISH without the AIM assay and $CD40L^+$ pre-gating showed again contrasting patterns of *IL17F* and *IL22* expression between CPs and ECs (Supplementary Figure S4b–i).

Systemic translocation of microbial products play a critical role in immune activation related to HIV pathogenesis[22]. Loss of gut mucosal integrity and T_H17 cell depletion have been associated with immune activation and disease progression[23]. We thus investigated possible links between HIV-specific T_H17 and T_H22 responses and systemic immune activation. We observed an inverse correlation between *IL22* mRNA expression by HIV-specific $CD4^+$ T cells and co-expression of the activation markers HLA-DR and CD38 on total $CD4^+$ and $CD8^+$ T cells (Fig. 5j, Supplementary Fig. 5i). We identified similar associations for *IL17F* (Fig. 5k, Supplementary Fig. 5j). We next investigated the relationship between preservation of $IL17^+$ and/or $IL22^+$ HIV-specific $CD4^+$ T cells and microbial translocation by deep sequencing of cell-free RNA fragments in plasma — a reflection of the part of the body microbiome that can access the circulation, the majority of which resides in the gut. The total levels of bacterial RNA reads in plasma did not significantly differ between the groups of CPs, ECs and uninfected control donors (UDs) (Fig. 5l). However, we observed a lower plasma bacterial RNA species diversity in CPs compared to the EC and UD groups (Fig. 5m). These observations may reflect low gut microbiome diversity, a state associated with systemic inflammation in several diseases[24] (Supplementary Fig. 5k,l). *IL22* and *IL17F* transcription in HIV-specific $CD4^+$ T cells was significantly correlated with bacterial RNA species diversity (Fig. 5n,o). CPs and ECs qualitatively differed in the bacterial RNA profile, forming largely distinct clusters (Supplementary Fig. S5m). *IL22* and *IL17F* mRNA levels in

HIV-specific CD4+ T cells inversely correlated with the prevalence of Proteobacteria, a phylum of Gram-negative bacteria that is increased in the gut microbiome in HIV-infected people[25] and has been proposed as a marker of dysbiosis[26] (Fig. 5p,q).

Thus, robust *IL22*+ and *IL17F*+ HIV-specific CD4+ T cells can be detected in the blood of ECs, while these responses are severely impaired in CPs. Associations of this deficiency with systemic immune activation, reduced plasma bacterial RNA species diversity with a skewing towards greater Proteobacteria abundance suggest that these findings in blood reflect perturbed gut homeostasis and loss of protective mucosal immunity.

ART results in a distinct transcriptome of HIV-specific CD4

Finally, we sought to determine whether the observed HIV-specific CD4+ T cell profiles were the consequences of ongoing viral antigen exposure, or the results of durable cell-fate decision programs that would persist after viral suppression on ART. We longitudinally followed a subgroup of 8 CPs and analyzed a second sample obtained while on ART (participant characteristics: Supplementary Table 8). Pairwise comparisons between the EC, CP pre-ART (CPpre) and CP post-ART (CPpost) cohorts (Fig 6a,b) showed sets of DEGs between all three groups, with fewer differences in the CPpre-CPpost and CPpost-EC than the CPpre-EC comparison, as may be expected. We next focused on selected GO terms and gene signatures identified as differentially expressed between CPs and ECs in Figs 1 and 2. While initiation of ART in CP modulated the activity of several apoptotic pathways regulating apoptosis and gene expression towards the levels observed in EC, differences in the "Positive regulation of T helper cell differentiation", "Positive regulation of CD4+ αβ T cell activation" and "Regulation of leukocyte migration" persisted (Fig 6c). We therefore examined how ART affected Thelper polarization signatures. Expression of the T_{FH} signature in HIV-specific CD4+ T cells was markedly reduced by ART, to levels similar to that found in EC (Fig 6d), while expression of the T_H1, T_H17 and LCMV exhaustion signatures was little affected by therapy: the differences clearly persisted between CPpost and EC (Fig 6e,f, Supplementary Fig. S6a). Expression of selected individual genes of interest was also consistent with these findings when expression levels in EC were considered as references (Fig 6g, Supplementary Fig. S6b and Supplementary Tables 9 and 10). While the results should be considered as trends given the limited size of the cohorts, expression of some genes appeared corrected by ART (e.g, *CXCL13*), whereas others are not restored at all (e.g. *IL22*).

These data demonstrate that suppressive ART results in a transcriptional landscape of HIV-specific CD4+ T cells distinct from the profiles of both untreated CPs and ECs. We identified gene modules susceptible to manipulation of antigen load, while others were unresponsive to antiviral therapy alone.

Discussion

The molecular features of virus-specific CD4+ T cells associated with pathogen containment are poorly understood in chronic human infections. We used genome-wide transcriptional analyses following *ex vivo* stimulation with HIV antigen to uncover differences in HIV-specific CD4+ T cell differentiation and function in chronically infected humans with

various viremia levels. An atypical enrichment of a T_{FH}-like signature in $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells characterized progressive infection. In contrast, elite control of viral replication was associated with robust HIV-specific T_H22 and T_H17 responses, which correlated negatively with immune activation and correlated positively with a greater diversity of bacterial RNA species in plasma and lower Proteobacteria abundance. None of these differences were detected in SEB or CMVpp65-reactive cells, revealing antigen-specific TCR signaling-driven processes. Suppressive ART in chronic progressors elicited a distinct HIV-specific $CD4^+$ T cell transcriptome, revealing in comparison to elite controllers, the presence of viremia sensitive (e.g, T_{FH} signature) and viremia unresponsive (e.g, T_H1 and T_H17 and T_H22 signatures) gene modules. These data suggest that HIV-specific $CD4^+$ T cell dysregulation does not represent a mere loss of Thelper functions, as classically defined for immune exhaustion, but also an adaptation to environmental cues combined with persistent changes, likely resulting from durable cell-fate decision programs.

At first sight, the marked T_{FH} signature[18] identified in CPs contrasts with the murine LCMV model, in which T_{FH} responses in chronic infection are critical for delayed viral control[8]. The extraordinary ability of HIV to generate escape mutations and elude autologous Ab neutralization contribute to these opposite associations between T_{FH} differentiation and disease outcome. A study showed that T_{FH} cells in chronic LCMV infection are capable of supporting generation of neutralizing Ab effective against the contemporaneous virus[9]. In contrast, the immune system lags behind the evolution of autologous HIV strains[27]. Kinetics of viral evolution in HIV infection may thus overcome the ability of adaptive T_{FH} skewing to contribute to delayed viral control. While technical limitations have to date hampered investigation of HIV-specific GC T_{FH} in lymph nodes, robust expansion of total GC T_{FH} populations was demonstrated in viremic HIV-infected people and SIV-infected macaques, albeit with qualitative defects[28, 29, 30, 31].

Genes classically associated with T_{FH} cells that are enriched in the $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells of CPs include *IL21* and *CXCL13*. HIV p24-specific Abs titers in plasma correlated with frequencies of $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells, suggesting that these responses may provide help to B cells *in vivo*. Studies in murine models have shown that CXCR5−/− $CD4^+$ T cells are able to take over some B helper functions[32, 33], resulting in the generation of low-affinity binding antibodies[9]. Acquisition of T_{FH}-like properties by $CXCR5_{neg}$ $CD4^+$ T cells has also been reported in rheumatoid arthritis[34] and cancer[35] and combines B help function with migration to inflamed tissues. These analogies suggest shared pathways of $CD4^+$ T cell differentiation between diverse diseases associated with chronic inflammation.

Damage to the gut mucosa with associated microbial translocation[22] and increased inflammatory properties of the enteric microbiota[36] are major drivers of HIV-associated chronic inflammation. We observed that HIV-specific T_H17 and T_H22 responses were abundant in ECs and rare in CPs. IL-17F and IL-17A are closely related cytokines recruiting neutrophils to sites of inflammation. IL-22 promotes intestinal barrier function[37], and supports complement-mediated elimination of pathobionts[38]. Lack of HIV-specific T_H17 and T_H22 responses was associated with $CD4^+$ and $CD8^+$ T cell activation and a reduced diversity of translocated bacterial RNA in plasma, with overrepresentation of Proteobacteria.

These results are consistent with studies in SIV infection showing preferential translocation of Proteobacteria compared to other phylums from the gut lumen into tissues, resulting in enhanced immune activation[39]. Our data do not establish causal links between features of HIV-specific CD4+ T cells and microbial translocation. It is more likely that lack of HIV-specific T_H-17/T_H22 responses and altered translocated microbiome both result from persistent insults to the gastrointestinal tract. The paucity of HIV-specific IL17+ and IL22+ CD4+ T cells observed in CPs may result from high susceptibility of T_H17 and T_H22 subsets to HIV infection[37, 40] and/or alteration of CD4+ T cell differentiation[41]. Together with the analysis of bacterial RNA fragments in plasma, these CD4+ T cell data suggest that studies of peripheral blood can provide useful insight into the systemic consequences of perturbed gut homeostasis. Probiotic and IL-21 therapy in ART-treated SIV-infected macaques was associated with enhanced polyfunctional T_H17 expansion and reduced markers of microbial translocation and dysbiosis compared to ART alone[42]. It will be important to determine if such interventions restore HIV-specific CD4+ T cell profiles.

This work raises questions that will need to be addressed in further studies. AIM assays require activation to enable cell identification. While we believe that analyzing the transcriptional profile modulated by TCR signaling is highly relevant for analysis of T cell responses, it will be important to investigate other populations of HIV-specific CD4+ T cells that do not co-upregulate CD69 and CD40L, for example by other AIM markers or by tetramers. Experiments in animal models will be required to define the ontogeny of $CXCR5_{neg}$ CD4+ T cells with T_{FH}-like features and whether they transitioned through a *bona fide* T_{FH} stage during their development. Longitudinal studies from the time of acute infection may identify events in the genesis or persistence of HIV-specific CD4+ T cell dysregulation. Investigations of human gut and lymph node biopsies, while challenging, would define the relationships between features of HIV-specific CD4+ T cells in blood and other anatomical compartments.

In summary, this work identifies extensive alterations in HIV-specific CD4+ T cell transcriptional profiles and demonstrates that skewed lineage differentiation is a key feature of altered CD4+ T cell help in the setting of high viral load, in which atypical gain of function is combined with classical exhaustion features. The data also reveal distinct features of HIV-specific CD4+ T cells after successful ART. The identified molecular features may therefore help direct selective interventions to improve immune function. Given the paucity of transcriptome datasets available on pathogen-specific CD4+ T cells in humans, our results will also provide useful benchmark comparisons for other infectious diseases and vaccine-induced responses.

ONLINE METHODS

Participants and Samples

Leukaphereses were obtained from study participants at the Montreal General Hospital, Montreal, Canada; at the Centre Hospitalier de l'Université de Montréal (CHUM) in Montreal, Canada; and at the Massachusetts General Hospital (MGH) in Boston, USA. The study was approved by the respective IRBs, written informed consent obtained from all participants prior to enrolment and PBMC samples obtained by leukapheresis. Samples were

collected between years 2013 and 2018. Subject characteristics are summarized in Supplementary Table 1. “Chronic progressors” (CP) were defined as people with a viral load of more than 5,000 vRNA copies per ml of plasma; “viremic controllers” (VC) were defined as people with more than 50 and less than 5,000 vRNA per ml plasma; and “elite controllers” (EC) were defined as subjects who spontaneously controlled viremia to below 50 RNA copies per ml plasma in the absence of therapy. CPs, VCs, and ECs.were either treatment naive or untreated for at least 3 months. Eight CPs were followed longitudinally and a second sample was obtained at least 6 months after initiation of ART therapy and suppression of viral load below 50 copies mL (Supplementary Table 8). PBMC were isolated by the Ficoll density gradient method and stored in liquid nitrogen.

Antibodies

For details of antibodies used, see Supplementary Tables 11–20 and descriptions of use below. In all cases, antibodies are monoclonal and raised in mice unless otherwise stated. All antibodies are commercially available, were validated by the suppliers and titrated using biological and/or isotype controls.

CD40L/CD69 AIM assay for live-cell sorting of HIV Gag-specific CD4+ T cells

Cryopreserved PBMCs were thawed and rested at 37°C for 3 hours. Cells were aliquoted into wells of a 24-well plate, at a total of 15×10^6 cells per well in RPMI medium supplemented with 10% Human AB (HAB) serum and PenStrep. For each assay, we used a no exogenous stimulation (No Ag) condition as control and an HIV-specific stimulation with an overlapping peptide pool corresponding to HIV Gag (JPT, PM-HIV-Gag ULTRA), or HCMVA pp65 (JPT, PM-PP65) at a final concentration of 0.5μg/ml/peptide. A CD40 blocking antibody (Miltenyi Biotec, clone HB14) was added to each well 15 minutes prior to stimulation at a final concentration of 0.5 μg/ml. Following a 9 hr stimulation, the cells were stained for 30 min at 4°C before live cell sorting on a FACS Aria cell sorter (BD BioSciences) equipped for handling of biohazardous material. Two panels of antibodies were used to isolate live cells subsets according to combinations of surface markers. In the first panel for isolation of total $CD69^+CD40L^+$ Gag-specific $CD4^+$ T cells, PBMCs were stained as indicated in Supplementary Table 11. The gating strategy for cell sorting is illustrated in Supplemental Figure 1. In the second panel used for isolation of CXCR5mem and CXCR5-subsets of $CD69^+CD40L^+$ Gag-specific $CD4^+$ T cells, PBMCs were stained as indicated in Supplementary Table 12. The gating strategy for cell sorting is illustrated in Supplementary Fig. 3. The FACS Aria was operated at 70 pounds per square inch with a 70-μm nozzle. For all populations, 5,000 cells were collected directly into RLT lysis buffer (Qiagen).

Phenotyping of Gag-specific $CD4^+$ T cells identified by the CD69/CD40L AIM assay.—PBMCs were thawed and plated in 24-well plates at 10×10^6/ml in RPMI $^+$ 10% HAB Serum (HAB) and rested for 3 hours at 37°C. Prior to stimulation, CD40 blocking antibody was added directly into culture at a final concentration of 0.5μg/ml, as well as antibodies against CXCR5, CCR6 and CXCR3. Cells were incubated for 15 min, then left unstimulated or stimulated either with an overlapping peptide pool corresponding to HIV Gag (JPT, PM-HIV-Gag ULTRA), at a final concentration of 0.5μg/ml/peptide, or with

Staphylococcal Enterotoxin B (SEB) as a positive control (1 μg/mL). Cells were stimulated for 9h, collected and stained for 30 min at 4°C with the antibodies in Supplementary Table 13. Cells were washed, fixed for 20 min in 2% paraformaldehyde, washed, and resuspended in 1% FCS/PBS for flow acquisition on a 5-laser LSRII (BD BioSciences).

Phenotyping of immune activation markers

PBMCs were thawed and rested overnight at 37°C. Next day, cells were washed and stained with viability dye (20 min at 4 C°) and surface markers (20 min at 4 C°) (Supplementary Table 14). Fluorescence minus one staining was used to define the cutoff for positivity. Cells were washed, fixed for 1h in 2% paraformaldehyde, washed, and resuspended in 1% FCS/PBS for flow acquisition on a 5-laser LSRII (BD BioSciences).

Standard intracellular cytokine staining (ICS) for cytokine measurements

PBMCs were incubated for 6 hours with an HIV Gag peptide pool at a concentration of 0.5μg/ml/peptide in the presence of 2.5 μg/ml brefeldin A (BD GolgiPlug) and (0.3μl/ml) monensin (BD GolgiStop). For delayed ICS assays, 2.5 μg/ml brefeldin A (BD GolgiPlug) and (0.3μl/ml) monensin (BD GolgiStop) were added 9h after stimulation with the Gag peptide pool and cells were cultured for 12h. Unstimulated cells were used as negative control and SEB (0.5) μg/ml as positive control. Cells were then stained with viability dye (20 min at 4 C°), surface markers (20 min at 4 C°), fixed with Fixation Solution (eBioscience), and stained for intracellular proteins in permeabilization 1X buffer (eBioscience) (30 min at 4 C°) as described in Supplementary Tables 15 and 16. Cells were acquired on an LSR II (BD Biosciences).

Gene Expression Analysis

Whole Transcriptome Analysis by microarrays—Total RNA was purified using the RNeasy Plus Micro Kit (Qiagen). RNA integrity was assessed using a 2100 Bioanalyzer (Agilent Technologies). Sense-strand cDNA was synthesized from total RNA using a fixed volume of 3μL. Fragmentation and labeling were performed to produce ss-DNA with the Applied Biosystems™ GeneChip® WT Pico Terminal Labeling Kit according to manufacturer's instructions (Applied Biosystems™). After fragmentation and labeling, 5 μg DNA target was hybridized on GeneChip® Clariom™ D, human (ThermoFisher Scientific) and incubated at 450C in the Genechip® Hybridization oven 640 (Affymetrix) for 17 hours at 60 rpm. GeneChips were then washed in a GeneChips® Fluidics Station 450 (Affymetrix) using Applied Biosystems™ Hybridization Wash and Stain kit according to the manufacturer's instructions (Applied Biosystems™). The microarrays were finally scanned on a GeneChip® scanner 3000 (Affymetrix).

Validation of transcriptional expression by Fluidigm—Total RNA was purified using the RNeasy Plus Micro Kit (Qiagen). cDNA was synthesized using all RNA available (or 1–5 ng) with the High-Capacity Reverse Transcription Kit with RNase Inhibitor (Life Technologies) (25° C for 10 min, 37° C for 120 min, 85° C for 5 min). cDNA equivalent to 1000 sorted cells was subjected to gene-specific preamplification using Taqman Preamp MasterMix (Applied Biosystems) and 96 pooled TaqMan Assays (Applied Biosystems) (Supplementary Tables 21–23) at final concentration 0.2X (95°C for 10 min, followed by 16

cycles of 95°C for 15 s and 60°C for 4 min). The preamplified cDNA was diluted 5-fold in DNA suspension buffer (Teknova) and was mixed with TaqMan Universal PCR Master mix (Life Technologies) and 20X GE sample loading reagent (Fluidigm). 20X Taqman assays were diluted 1:1 with 2X assay loading buffer (Fluidigm). Taqman assays mixtures were loaded onto a primed 96.96 Dynamic Array chip (Fluidigm). The chip was loaded into the IFC Controller, where each sample was mixed with each assay in every possible combination. The chip was transferred in a Biomark (Fluidigm) for real-time PCR amplification and fluorescence acquisition using single probe (FAM-MGB, reference: ROX) settings and the default hot-start protocol with 40 cycles. Cycle thresholds (Ct) were calculated using the Fluidigm BioMark software v1.4.2.

Combined cytokine/chemokine mRNA Flow FISH and protein staining assays

PBMCs were either left unstimulated or were stimulated with an HIV Gag peptide pool (JPT) or SEB for 15 hours after 15h min incubation with a CD40 blocking antibody (Miltenyi Biotec, clone HB14) at a final concentration of 0.5 μg/ml. For experiments for evaluation of mRNA expression without CD40 blocking antibody PBMCs were stimulated for 12h. After stimulation, cells were stained with a viability dye (20 min, 4°C, Fixable LiveDead, eBioscience) and then surface markers were labelled (30 min 4°C) with antibodies (See panels in Supplementary Tables 17–20). Samples were next subjected to the PrimeFlow RNA assay (Affymetrix/eBioscience) for specific mRNA detection as per manufacturer's instructions. All buffers and fixation reagents were provided with the kit, with the exception of flow cytometry staining buffer (2% FBS/PBS). After fixation and permeabilization, cytokine/chemokine mRNAs were labeled with sets of probe pairs, as listed in Supplementary Tables 17–20. The probes were diluted 1:5 in diluent and hybridized to the target mRNA for 2 hr at 40 °C. Samples were washed to remove excess probes and stored overnight in the presence of RNAsin. Signal amplification was achieved as previously described[20] by sequential 1.5 hr, 40 °C incubations with the pre-amplification and amplification mix. Amplified mRNA was labeled with fluorescently-tagged probes for 1 hr at 40 °C. A negative control probe and a positive control probe (against house-keeping gene RPL13A) were included in each experiment. Gates were set on the unstimulated control where appropriate or scrambled probes. Samples were acquired on an LSRII (BD Bioscience) or a BD LSRFortessa™ (experiments without CD40 blocking antibody). Analysis was performed using FlowJo (Treestar, V10) and Spice 5.35 software.

Measurements of cytokine secretion by beads arrays and ELISA

One million CD8 T cell-depleted PBMCs (Dynabeads CD8, StemCell) were incubated with an HIV Gag peptide pool (1μg/ml/peptide) or left unstimulated and incubated at 37°C in RPMI supplemented with 10% HAB and PenStrep. Cytokine secretion in supernatants was measured 48 hours after stimulation on a Luminex beads array platform, as previously described[19, 20]. Measurements were performed in duplicates using the Human High Sensitivity Cytokine Premixed Kit B (R&D Systems) for IL-2, IFN-γ, TNF-α, IL-17A, IL-22 and IL-31 or for IL-17F on a Bio-plex 200 array system (Bio-Rad Laboratories) per manufacturer' s instructions. CXCL13 was measured separately with the Human CXCL13 Quantikine ELISA Kit (R&D Systems), as per manufacturer's instructions.

Measurement of CD4i p24-specific antibodies by ELISA

The anti-p24 ELISA was performed as follows: 96 well plates (Thermo Scientific Nunc, FluoroNunc/LumiNunc, MaxiSorp Surface) were coated with 0.1μg/ml of recombinant p24 (NIH #12028) or BSA (Bioshop #ALB001.100) in PBS over night at 4°C. Plates were blocked for 90 min at room temperature (RT) with blocking buffer (TBS, Tween 0.1%, BSA 2%) and then washed 4 times with washing buffer (TBS, Tween 0.1%). Dilutions of human sera (1/1000, 1/3000, 1/6000) or rabbit anti-HIV p24 antiserum (NIH #4250) in washing buffer containing 0,1% of BSA were incubated for 2 h at RT. Plates were washed 4 times with washing buffer before incubation for 90 min at RT with HRP-conjugated secondary Abs: goat anti-human IgG HRP (Invitrogen #31413) or anti IgG rabbit HRP (Invitrogen #31462). Plates were then washed 4 times with washing buffer before revealing with standard ECL (Perkin Elmer #NEL105001EA) with a TriStar luminometer (LB 941, Berthold Technologies).

Plasma microbiome sequencing

Circulating total RNA and DNA were isolated from frozen plasma using RNAzolBD (MRCgene) according to the manufacturer's recommendations. After RNA fragmentation and reverse transcription using random hexamers, obtained transcripts as well as the isolated plasma DNA were individually incorporated into barcoded cDNA libraries based on the Illumina TruSeq platform. Libraries were validated by microelectrophoresis, quantified, pooled and clustered on Illumina patterned flow cells. Clustered flow cells were sequenced on an Illumina HiSeq 4000 in 75-base paired reads.

Gene expression data analysis

Microarrays: The preprocessing of the microarray data was done in R[43] v3.3.3 using the *oligo* package[44] v1.38.0. Specifically, the CEL files were loaded in R and the *rma* function was applied for background subtraction, normalization and summarization. Quality control was performed by inspecting various diagnostic plots of the intensity distribution and correlation structure of control and regular probes. After review of the QC plots, 3 outliers were removed. Combat (sva package 3.30.1) has been applied for batch normalization of samples using 4 replicates as reference for the normalization. The normalized data were used for subsequent subanalyses CPpre vs EC, CPpre vs CPpost, CPpost vs EC.

Differential expression: Differential analysis of gene expression was done by the R package *limma45*. Using the *lmFit* function, a linear model was fit to each gene separately. Moderated t-tests were performed between each group. P-values were adjusted using Benjamini-Hochberg method for multiple test correction.

CAMERA

T_H1 signature (GSE59295; T_H1 vs T_H2) and T_H17 signature (GSE49703; T_H17 vs T_H1) were available in C7 collection of MSigDB. The T_{FH} signature (GSE50391; CXCR5high CD45RO vs CXCR5-tonsil samples) and LCMV exhaustion signature (GSE41866; LCMV Clone 13 D30 vs Armstrong D30-Exhausted vs Memory) were created using the GEO2R online tool (https://www.ncbi.nlm.nih.gov/geo/geo2r/). For each signature, a gene set test

was performed using the *camera* function from the *limma* package, accounting for inter-gene correlation.

Venn diagrams were created using the online tool Eulerr.co, (R package version 4.1.0, https://cran.r-project.org/package=eulerr)46. Unsupervised hierarchical clustering was performed using the in heatmap2 function in R[43] on relative signal intensity values applying the Euclidean distance metric and complete linkage clustering method. Barcode plots were generated with the *barcodeplot* function from the *limma* package.

qRT-PCR results—Analysis of the qRT-PCR data obtained on the microfluidic platform was carried out using GenEx software (MultiD Analyses, URL: http://www.multid.se). Five endogenous control genes were included in the Fluidigm run and the stability of endogenous control genes across all experimental samples was evaluated applying the NormFinder algorithm[47] in GenEx. The mean expression of the most stable endogenous control genes was used for normalization and calculation of -ΔCt values. Principal component analysis, biplots and correlograms were created using the prcomp, fviz_pca_biplot and corrplot functions, respectively, in R programming language.

Statistical analyses of flow cytometry, beads array and qRT-PCR results

Statistical analyses were performed using Prism V6.0 (GraphPad) using non-parametric tests. Two-group comparisons were performed using the Mann-Whitney test and pairwise comparisons were performed using the Wilcoxon matched-pair test. For comparisons between three groups, Kruskal-Wallis or Friedman one-way ANOVA with Dunn's post-test was used. Permutation test (10000 permutations) was applied for pie-chart comparison using the SPICE software. For correlations, Spearman's R correlation coefficient was applied. Statistical tests were two-sided and $p<0.05$ were considered significant.

Bioinformatic analysis of deep sequencing results of cell-free RNA fragments in plasma

Raw, paired-end Illumina reads were first trimmed for low quality bases and adaptor contamination with Trimmomatic[48]. As a first pass to remove host-derived sequences, trimmed reads were first aligned against the HG38 genome with Bowtie 2[49] and then aligned to HG38 transcriptome with the splice aware aligner, STAR[50]. Unaligned paired-end reads were then assembled into contigs using Trinity[51]. Trinity contigs were then aligned against HG38 again to further remove host-derived sequences with BLAST+[52]. Remaining contigs were then aligned against the NCBI nt database[53] and quantified with Salmon[54]. Taxonomic information was assigned for each contig and TPM values were collapsed at the species level for each sample for downstream analysis. TPM values for each sample were collapsed on the Genus taxonomic level and Morisita-Horn Dissimilarity indexes were calculated on a pairwise basis using the R package, vegan[55]. The pairwise dissimilarity matrix was then clustered and the heatmap was produced with gplots[56].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank Josée Girouard, the clinical staff at the McGill University Health Centre in Montreal, the Ragon/MGH clinical and technical staff and all study participants for their invaluable role in this project; Dr Bruce Walker for providing clinical samples; Dr Dominique Gauchat and the CRCHUM Flow Cytometry Platform for technical assistance; Nathalie Hamel and McGill University And Génome Québec Innovation Centre for microarray analysis; Drs Jeremy Boss and Shane Crotty for their input on this manuscript. The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 IIIB p24 Recombinant Protein from ImmunoDX, LLC. This study was supported by the National Institutes of Health grants HL092565 (D.E.K), AI100663 CHAVI-ID (D.E.K, R.T.W; Consortium PI: D. Burton), OD011103 (R.P.J) and OD011132 (R.P.J); the Canadian Institutes for Health Research grants #137694 (D.E.K), #152977 (D.E.K), MOP-93770 (C.T) and foundation #352417 (A.F); the Canada Foundation for Innovation Program Leader grant #31756 (D.E.K); the FRQS AIDS and Infectious Diseases Network. This work was funded in part by the Intramural Program of the National Institutes of Health (D.C.D, S.D). D.E.K and C.T are supported by Senior Research Scholar Awards of the Quebec Health Research Fund (FRQS). J.P.R is the holder of Louis Lowenstein Chair in Hematology & Oncology, McGill University. A.F. is a Canada Research Chair on Retroviral Entry.

Data availability

Microarray data generated during the current study were deposited in the Gene Expression Omnibus (GEO) public depository with the accession number GSE128297 for the SuperSeries, and the following accession numbers for the Subseries; GSE129872 HIV-specific $CD4^+$ T cells samples from CPs, VCs and ECs), GSE128280 (CXCR5mem and CXCR5neg HIV-specific $CD4^+$ T cells from CPs and ECs), GSE128296 (HIV-specific $CD4^+$ T cells samples from CPs before/after ART and ECs). mRNA expression data by high-throughput qRT-PCR are available in Supplementary Material. All the datasets that support the findings of this study are available from the corresponding author on reasonable request.

REFERENCES

1. Laidlaw BJ, Craft JE & Kaech SM The multifaceted role of CD4(+) T cells in CD8(+) T cell memory. Nat Rev Immunol 16, 102–111 (2016). [PubMed: 26781939]
2. Swain SL, McKinstry KK & Strutt TM Expanding roles for CD4(+) T cells in immunity to viruses. Nat Rev Immunol 12, 136–148 (2012). [PubMed: 22266691]
3. O'Shea JJ & Paul WE Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells. Science 327, 1098–1102 (2010). [PubMed: 20185720]
4. Wherry EJ & Kurachi M Molecular and cellular insights into T cell exhaustion. Nat Rev Immunol 15, 486–499 (2015). [PubMed: 26205583]
5. Crawford A et al. Molecular and transcriptional basis of CD4(+) T cell dysfunction during chronic infection. Immunity 40, 289–302 (2014). [PubMed: 24530057]
6. Morou A, Palmer BE & Kaufmann DE Distinctive features of CD4+ T cell dysfunction in chronic viral infections. Curr Opin HIV AIDS 9, 446–451 (2014). [PubMed: 25023623]
7. Vella LA, Herati RS & Wherry EJ CD4(+) T Cell Differentiation in Chronic Viral Infections: The Tfh Perspective. Trends Mol Med 23, 1072–1087 (2017). [PubMed: 29137933]
8. Fahey LM et al. Viral persistence redirects CD4 T cell differentiation toward T follicular helper cells. J Exp Med 208, 987–999 (2011). [PubMed: 21536743]
9. Greczmiel U et al. Sustained T follicular helper cell response is essential for control of chronic viral infection. Sci Immunol 2 (2017).
10. Kaufmann DE et al. Upregulation of CTLA-4 by HIV-specific CD4+ T cells correlates with disease progression and defines a reversible immune dysfunction. Nat Immunol 8, 1246–1254 (2007). [PubMed: 17906628]
11. Ferre AL et al. HIV controllers with HLA-DRB1*13 and HLA-DQB1*06 alleles have strong, polyfunctional mucosal CD4+ T-cell responses. J Virol 84, 11020–11029 (2010). [PubMed: 20719952]
12. Crotty S T follicular helper cell differentiation, function, and roles in disease. Immunity 41, 529–542 (2014). [PubMed: 25367570]
13. Reiss S et al. Comparative analysis of activation induced marker (AIM) assays for sensitive identification of antigen-specific CD4 T cells. PLoS One 12, e0186998 (2017). [PubMed: 29065175]

14. Wu D & Smyth GK Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Res 40, e133 (2012). [PubMed: 22638577]
15. Ono C et al. Fluorescently activated cell sorting followed by microarray profiling of helper T cell subtypes from human peripheral blood. PLoS One 9, e111405 (2014). [PubMed: 25379667]
16. Zhang W et al. Effector CD4+ T cell expression signatures and immune-mediated disease associated genes. PLoS One 7, e38510 (2012). [PubMed: 22715389]
17. Ramesh R et al. Pro-inflammatory human Th17 cells selectively express P-glycoprotein and are refractory to glucocorticoids. J Exp Med 211, 89–104 (2014). [PubMed: 24395888]
18. Locci M et al. Human circulating PD-1+CXCR3-CXCR5+ memory Tfh cells are highly functional and correlate with broadly neutralizing HIV antibody responses. Immunity 39, 758–769 (2013). [PubMed: 24035365]
19. Porichis F et al. Responsiveness of HIV-specific CD4 T cells to PD-1 blockade. Blood 118, 965–974 (2011). [PubMed: 21652684]
20. Porichis F et al. High-throughput detection of miRNAs and gene-specific mRNA at the single-cell level by flow cytometry. Nat Commun 5, 5641 (2014). [PubMed: 25472703]
21. Baxter AE et al. Multiparametric characterization of rare HIV-infected cells using an RNA-flow FISH technique. Nat Protoc 12, 2029–2049 (2017). [PubMed: 28880280]
22. Brenchley JM et al. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. Nat Med 12, 1365–1371 (2006). [PubMed: 17115046]
23. Favre D et al. Critical loss of the balance between Th17 and T regulatory cell populations in pathogenic SIV infection. PLoS Pathog 5, e1000295 (2009). [PubMed: 19214220]
24. Levy M, Kolodziejczyk AA, Thaiss CA & Elinav E Dysbiosis and the immune system. Nat Rev Immunol 17, 219–232 (2017). [PubMed: 28260787]
25. Vujkovic-Cvijin I et al. Dysbiosis of the gut microbiota is associated with HIV disease progression and tryptophan catabolism. Sci Transl Med 5, 193ra191 (2013).
26. Shin NR, Whon TW & Bae JW Proteobacteria: microbial signature of dysbiosis in gut microbiota. Trends Biotechnol 33, 496–503 (2015). [PubMed: 26210164]
27. Richman DD, Wrin T, Little SJ & Petropoulos CJ Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proc Natl Acad Sci U S A 100, 4144–4149 (2003). [PubMed: 12644702]
28. Petrovas C et al. CD4 T follicular helper cell dynamics during SIV infection. J Clin Invest 122, 3281–3294 (2012). [PubMed: 22922258]
29. Cubas RA et al. Inadequate T follicular cell help impairs B cell immunity during HIV infection. Nat Med 19, 494–499 (2013). [PubMed: 23475201]
30. Lindqvist M et al. Expansion of HIV-specific T follicular helper cells in chronic HIV infection. J Clin Invest 122, 3271–3280 (2012). [PubMed: 22922259]
31. Perreau M et al. Follicular helper T cells serve as the major CD4 T cell compartment for HIV-1 infection, replication, and production. J Exp Med 210, 143–156 (2013). [PubMed: 23254284]
32. Miyauchi K et al. Protective neutralizing influenza antibody response in the absence of T follicular helper cells. Nat Immunol 17, 1447–1458 (2016). [PubMed: 27798619]
33. Kahan SM & Zajac AJ Late arising T follicular helper cells cultivate the B cell crop during chronic infections. Sci Immunol 2 (2017).
34. Rao DA et al. Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. Nature 542, 110–114 (2017). [PubMed: 28150777]
35. Gu-Trantien C et al. CD4(+) follicular helper T cell infiltration predicts breast cancer survival. J Clin Invest 123, 2873–2892 (2013). [PubMed: 23778140]
36. Neff CP et al. Fecal Microbiota Composition Drives Immune Activation in HIV-infected Individuals. EBioMedicine 30, 192–202 (2018). [PubMed: 29650491]
37. Kim CJ et al. A role for mucosal IL-22 production and Th22 cells in HIV-associated mucosal immunopathogenesis. Mucosal Immunol 5, 670–680 (2012). [PubMed: 22854709]
38. Hasegawa M et al. Interleukin-22 regulates the complement system to promote resistance against pathobionts after pathogen-induced intestinal damage. Immunity 41, 620–632 (2014). [PubMed: 25367575]

39. Klase Z et al. Dysbiotic bacteria translocate in progressive SIV infection. Mucosal Immunol 8, 1009–1020 (2015). [PubMed: 25586559]
40. El Hed A et al. Susceptibility of human Th17 cells to human immunodeficiency virus and their perturbation during infection. J Infect Dis 201, 843–854 (2010). [PubMed: 20144043]
41. Favre D et al. Tryptophan catabolism by indoleamine 2,3-dioxygenase 1 alters the balance of TH17 to regulatory T cells in HIV disease. Sci Transl Med 2, 32ra36 (2010).
42. Ortiz AM et al. IL-21 and probiotic therapy improve Th17 frequencies, microbial translocation, and microbiome in ARV-treated, SIV-infected macaques. Mucosal Immunol 9, 458–467 (2016). [PubMed: 26286233]
43. Team RCR: A language and environment for statistical computing. In: Computing, R.F.f.S., editor. Vienna, Austria; 2014.
44. Carvalho BS & Irizarry RA A framework for oligonucleotide microarray preprocessing. Bioinformatics 26, 2363–2367 (2010). [PubMed: 20688976]
45. Ritchie ME et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43, e47 (2015). [PubMed: 25605792]
46. Larsson J eulerr: Area-Proportional Euler and Venn Diagrams with Ellipses. R package version 4.1.0. 2018.
47. Andersen CL, Jensen JL & Orntoft TF Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 64, 5245–5250 (2004). [PubMed: 15289330]
48. Bolger AM, Lohse M & Usadel B Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014). [PubMed: 24695404]
49. Langmead B & Salzberg SL Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359 (2012). [PubMed: 22388286]
50. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]
51. Grabherr MG et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29, 644–652 (2011). [PubMed: 21572440]
52. Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ Basic local alignment search tool. J Mol Biol 215, 403–410 (1990). [PubMed: 2231712]
53. O'Leary NA et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44, D733–745 (2016). [PubMed: 26553804]
54. Patro R, Duggal G, Love MI, Irizarry RA & Kingsford C Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14, 417–419 (2017). [PubMed: 28263959]
55. Oksanen J et al. vegan: Community Ecology Package. R package version 2.4–3.; 2017.
56. Warnes G & al, e. gplots: Various R programming tools for plotting data. 2005.

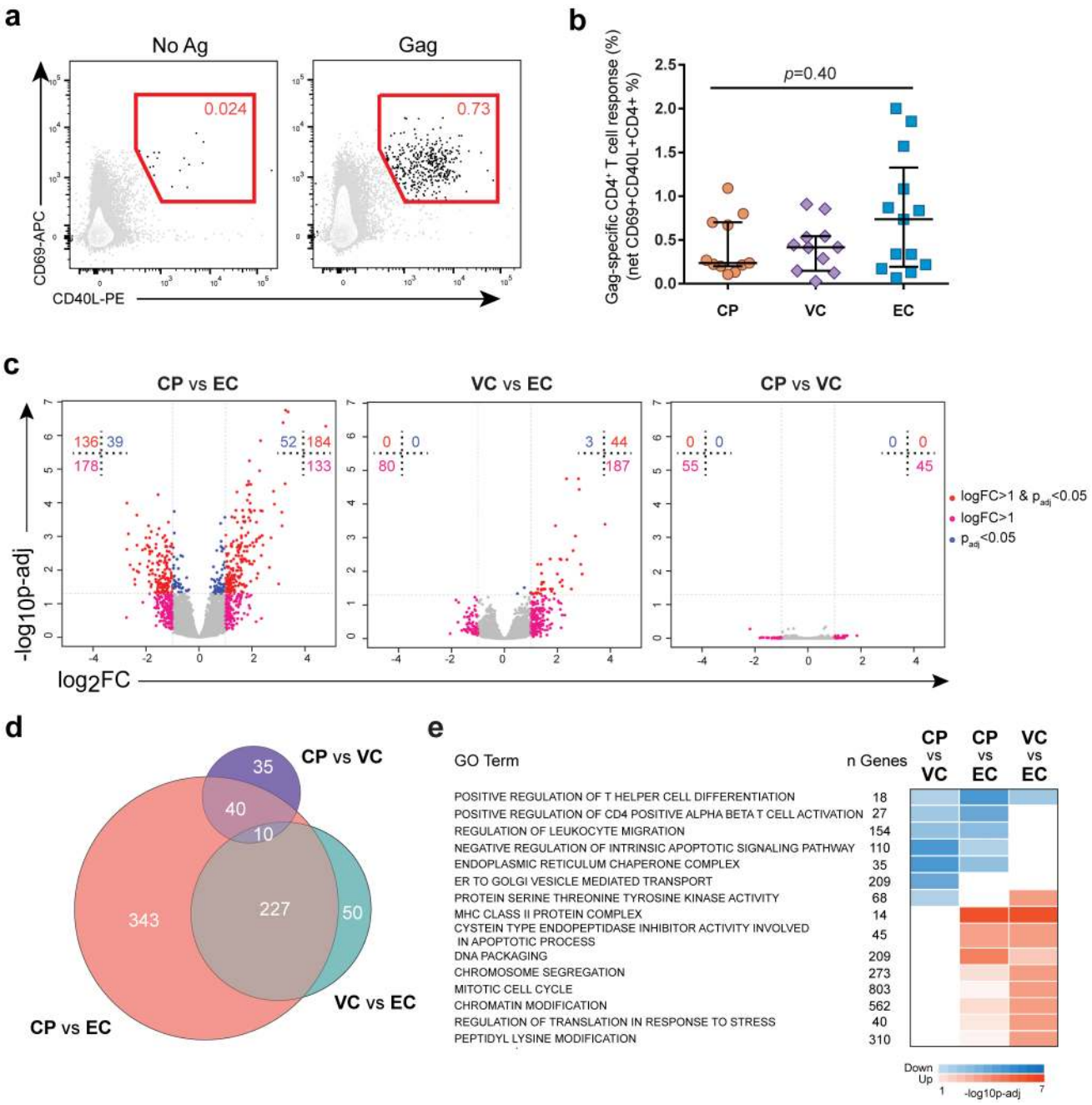

Figure 1. Deep coverage transcriptome analysis of HIV-specific CD4$^+$ T cells from untreated HIV-infected people with distinct disease status.

(a) Representative flow cytometry plots from an EC of detection of HIV-specific CD4$^+$ T cells by an activation-induced marker (AIM) assay based on co-upregulation of CD40L and CD69 9h after stimulation with an HIV Gag peptide pool. **(b)** Quantification of (a). Comparison of net frequency (background of No Antigen condition subtracted) of CD69$^+$CD40L$^+$ Gag-specific CD4$^+$ T cells amongst the CP, VC and EC groups by Kruskal-Wallis test. Horizontal lines and bars represent median and +/−interquartile range, respectively (n= 11CPs, 11 VCs, 13 ECs). **(c)** Paired volcano plot comparisons depicting

differentially expressed genes (DEG) as measured by microarray analysis of sorted HIV-specific $CD4^+$ T cells in the three groups. Red, pink and blue dots represent DEG with log2|FC| >1 and padj<0.05, $\log_2|FC|>1$ or padj<0.05, respectively. A two-sided moderated t-test followed by Benjamini-Hochberg method correction was applied. (n= 11 CPs, 9 VCs, 12 ECs). **(d)** Euler diagrams of DEG between each pairwise comparison of the three groups of participants ($\log_2FC> 1$, p<0.05); numbers represent DEGs. **(e)** Top significant enriched Gene Ontology terms from the curated C5 MSigDB via two-sided CAMERA analysis. Blue, Red and white color denote negative, positive or not significant (as defined by FDR<0.05 by Benjamini-Hochberg method), respectively.

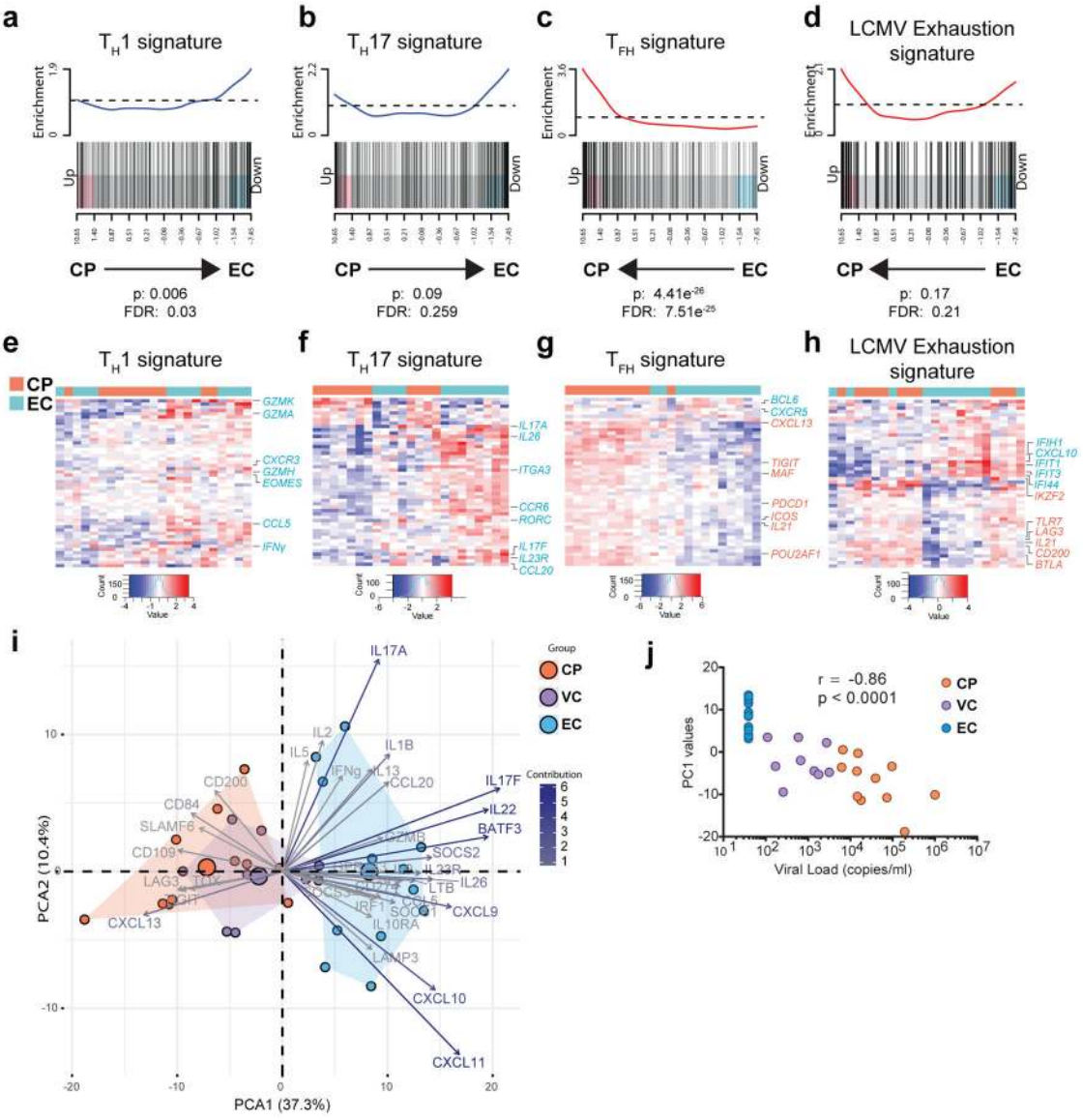

Figure 2. HIV-specific CD4$^+$ T cell polarization is associated with level of viral control.

(a-d) Barcode plots of enriched Thelper polarization and exhaustion gene sets by CAMERA in CP (n=11) vs EC (n=12) people. Red and blue lines denote positive and negative enrichment of CP genes vs. EC, respectively. (A) T_H1 signature: GSE59295 (T_H1 vs T_H2), (B) T_H17 signature: GSE49703 (T_H17 vs T_H1), (C) T_{FH} signature: GSE50391 (CXCR5high CD45RO vs CXCR5-tonsil samples), (d) LCMV exhaustion signature GSE41866 (LCMV Clone 13 D30 vs Armstrong D30-Exhausted vs Memory). Two sided p-value by CAMERA and FDR by Benjamini-Hochberg method. The arrows beneath the barcode plots depict the orientation of the enrichment between CP and EC. **(e-h)** Heatmaps of mRNA expression of the top 50 significant genes of each CD4$^+$ T cell signature in CPs and ECs. Red and blue font of gene names denote upregulated and downregulated mRNA expression in CPs vs ECs, respectively. **(i)** PCA representation of CPs (n=11), VCs (n=9) and ECs (n=12) based on transcriptional expression of 90 genes associated with CD4$^+$ T cell differentiation, effector

functions and exhaustion. Transcriptional expression (-ΔCt values) was evaluated by qRT-PCR (Fluidigm). The top 35 contributing genes are displayed and color-coded according to their contribution. Bigger dots represent the mean PC coordinates for each group. The numbers in parenthesis stand for the percentage of variance that the PC accounts for**. (j)** Correlation of PC1 values for each individual with viral load by two-tailed Spearman correlation (n= 11CPs, 9 VCs, 12 ECs).

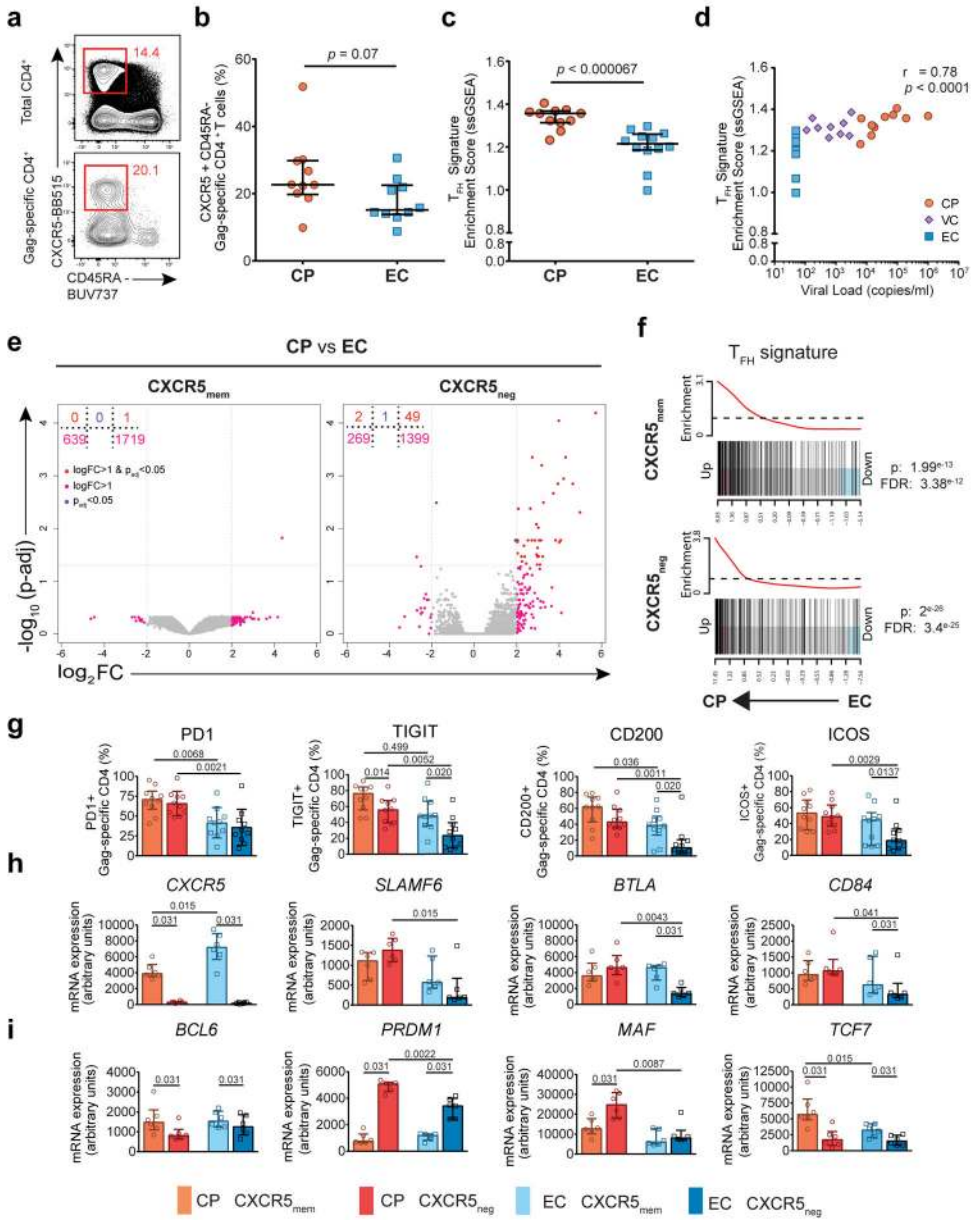

Figure 3. An atypical T_{FH}-like transcriptional signature in $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells discriminates chronic progressors from elite controllers.

(a) Representative flow cytometry plots from a CP illustrating the enrichment of cT_{FH} ($CXCR5^+$ CD45RA-) in $CD69^+CD40L^+$ Gag-specific $CD4^+$ T cells vs. total population in a CP subject, assessed 9h after stimulation with a Gag peptide pool. **(b,c)** Quantification from (a). Comparison between CPs and ECs of (**b**) the percentage of Gag-specific $CXCR5_{mem}$ assessed by flow cytometry (two-sided Mann-Whitney test) (n=10 CPs and 10 ECs); and **(c)** the enrichment score of a T_{FH} tonsil signature (GSE50391) by ssGSEA (two-sided Mann-Whitney test) (n_{CP}=11, n_{EC}=12). Bars represent the median +/−interquartile range. **(d)** Correlation of single-sample enrichment score for the T_{FH} tonsil signature with viral load by two-tailed Spearman test (n= 11 CPs, 9 VCs, 12 ECs). **(e)** Volcano plots depicting differential gene expression between CP and EC, amongst their live-sorted Gag-specific

$CXCR5_{mem}$ or $CXCR5_{neg}$ subsets, by microarray analysis. Red and pink dots represent DEG with $\log_2|FC|>1$ and padj<0.05 or $\log_2|FC| >1$, respectively, and the number of DEGs are displayed in the small tables. Two-sided moderated t-test followed by Benjamini-Hochberg correction method (padj) (n= 5 CPs and 3 ECs). **(f)** CAMERA enrichment analysis of the tonsil T_{FH} signature of $CXCR5_{mem}$ or $CXCR5_{neg}$ Gag-specific CD4 T cells of CPs compared to ECs (n= 5 CPs and 3 ECs). Two sided p-value by CAMERA and FDR by Benjamini-Hochberg method. **(g)** Comparison of expression of T_{FH} surface markers on $CXCR5_{mem}$ and $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells of CPs and ECs by flow cytometry (n= 10 CPs and 10 ECs). **(h)** mRNA expression of T_{FH}-related surface markers or **(i)** transcription factors by qRT-PCR (Fluidigm) (n= 6 CPs and 6 ECs) on sorted $CXCR5_{mem}$ and $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells of CPs and ECs. **(g,h,i)** Two-tailed Mann-Whitney test or for paired comparisons by two-tailed Wilcoxon matched-pairs signed-ranked test were used to verify significance. Only $p<0.05$ are displayed for clarity. Bars represent the median +/– interquartile range.

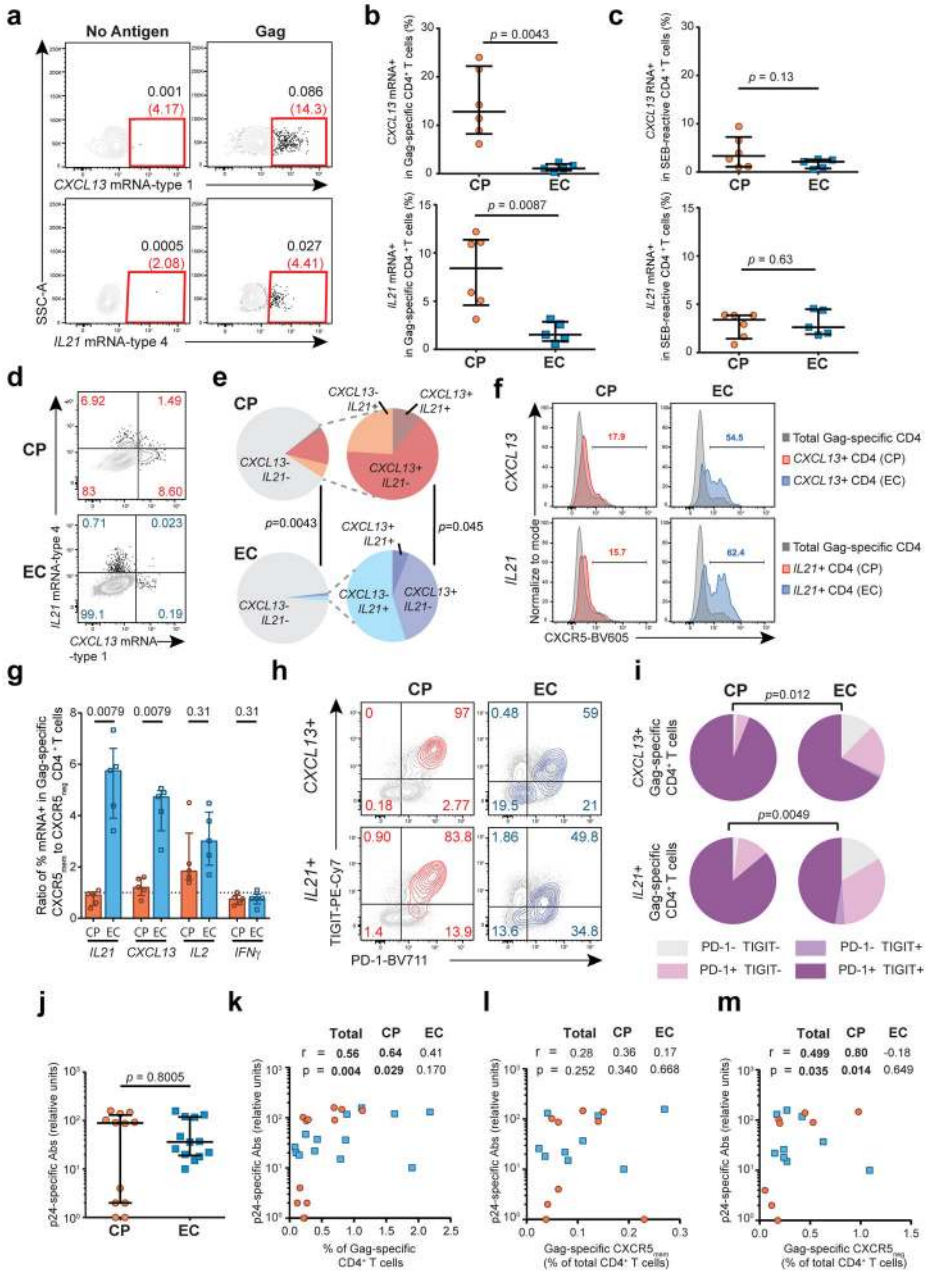

Figure 4. T_{FH} cytokine expression by $CXCR5_{neg}$ HIV-specific $CD4^+$ T cells of chronic progressors.

(a) Representative RNA flow FISH plots from an CP of *CXCL13* and *IL21* mRNA in $CD69^+CD40L^+$ $CD4^+$ T cells after a 15h stimulation with a Gag peptide pool. Numbers represent the frequency of $mRNA^+$ cells among (red) $CD69^+CD40L^+$ and (grey) total $CD4^+$ T cells (n= 6 CPs and 5 ECs) **(b,c)** Quantification of (a). Comparison of frequencies of *CXCL13* or *IL21* $mRNA^+$ $CD4^+$ T cells between CPs (n=6) and ECs (n=5) obtained as in **a** after Gag or SEB stimulation. Bars represent median frequencies +/−interquartile range (two-tailed Mann-Whitney test, **p<0.01). **(d,e)** RNA flow FISH analysis of co-expression of *CXCL13* and *IL21* mRNA by Gag-specific $CD4^+$ T cells with (**d**) representative plots

from CP and EC people and **(e)** related statistical analysis by permutation test (10000 permutations) in SPICE (n= 6 CPs and 5 ECs). Pie slices represent median frequency of CXCL13/IL21 mRNA$^+$ subpopulations. **(f)** Representative histograms from CP and EC people of CXCR5 expression on *CXCL13* or *IL21* mRNA$^+$ HIV-specific CD4$^+$ T cells overlaid on total HIV-specific CD4$^+$ T cells (grey). Numbers represent the percentages of CXCR5$^+$ cells within each cytokine$^+$ population (n= 6 CPs and 5 ECs). **(g)** Statistical analysis from (f) on 6 CPs and 5 ECs of the ratios between the $CXCR5_{mem}$ to the $CXCR5_{neg}$ subsets of frequencies of cytokine or chemokine mRNA$^+$ HIV-specific CD4$^+$ T cells (*CXCL13*, *IL21*, *IL2* and *IFNγ*) as assessed by RNA flow cytometry. Bars represent median frequencies +/−interquartile range (two-tailed Mann-Whitney test). **(h,i)** Co-expression patterns of PD-1 and TIGIT (surface protein Ab labeling) on *CXCL13* mRNA$^+$ or *IL21* mRNA$^+$ Gag-specific CD4$^+$ T cells identified by RNA flow FISH (n= 6 CPs and 5 ECs). (**h**) Representative examples of CP and EC people; cytokine/chemokine mRNA$^+$ HIV-specific CD4$^+$ T cells (red dots for CPs, blue dots for ECs) are overlaid on total HIV-specific CD4$^+$ T cell subpopulations (grey dots). Numbers represent the percentages of *CXCL13* RNA$^+$ or *IL21* mRNA$^+$ HIV-specific CD4$^+$ T cells located in each quadrant. **(i)** SPICE analysis of PD-1 and TIGIT expression from (h) on *CXCL13* mRNA$^+$ or *IL21* mRNA$^+$ HIV-specific CD4$^+$ T cells. (*p<0.05, **p<0.01 by permutation test, 10000 permutations). Slices represent median frequency of PD-1 and TIGIT subpopulations. **(j)** Comparison of p24-specific antibodies levels in plasma between CPs and ECs as assessed by ELISA. (two-tailed Mann-Whitney test (n= 12 CPs and 13). Bars represent medians +/− interquartile range. **(k,l,m)** Correlation between p24-specific antibodies levels measured by ELISA and frequencies of **k**) total HIV-specific CD4$^+$ T cells; **l)** $CXCR5_{mem}$ HIV-specific CD4$^+$ T cells and **m**) $CXCR5_{neg}$ HIV-specific CD4$^+$ T cells assessed by the CD69/CD40L AIM assay after a 9h stimulation with a Gag peptide pool (two-tailed Spearman test, n=12 CPs and 13 ECs).

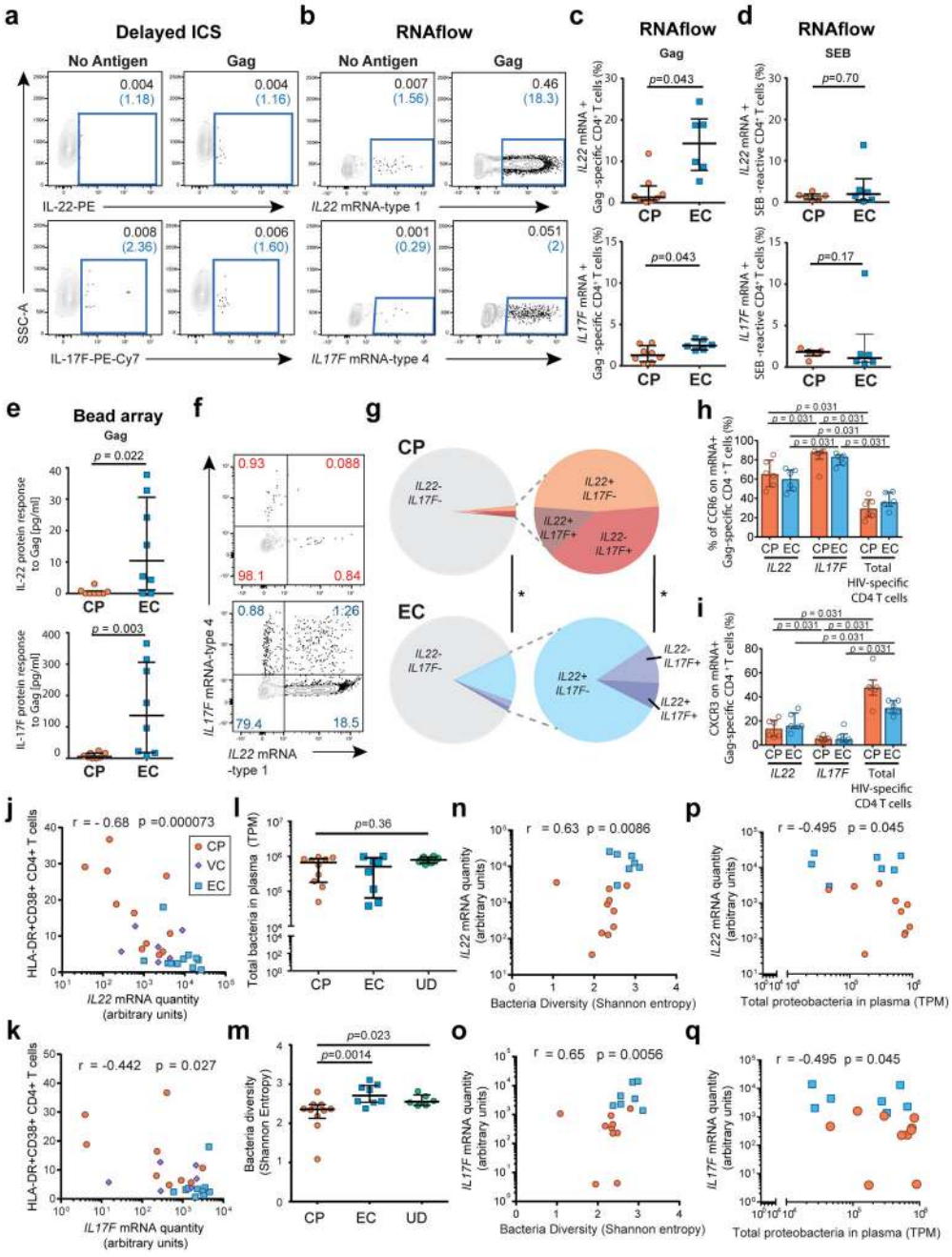

Figure 5. Increased expression of cytokines related to mucosal immunity in HIV-specific CD4$^+$ T cells of elite controllers compared to chronic progressors.

(a,b). Representative flow cytometry plots depicting detection of *IL22* and *IL17F* in CD69$^+$CD40L$^+$ CD4$^+$ T cells in one EC after stimulation with a Gag peptide pool by delayed ICS for cytokine protein and RNA flow FISH for cytokine mRNA, respectively. Numbers represent frequencies of mRNA$^+$ HIV-specific CD4$^+$ T cells among (blue) CD69$^+$CD40L$^+$ and (grey) total CD4$^+$ T cells (n= 8 CPs and 8 ECs). **(c,d)** Statistical comparison from (a) of frequencies of cytokine mRNA$^+$ antigen-specific CD4$^+$ T cells between CPs and EC after (c) a Gag peptide pool (n= 8 CPs and 8 ECs) or (d) SEB

stimulation (n=6 CPs and 8 ECs) (two-tailed Mann-Whitney test). **(e)** Statistical comparison of IL-22 and IL-17F protein levels detected by Luminex beads array in the supernatant of CD8-depleted PBMCs 48h after stimulation with a Gag peptide pool (n= 8 CPs and 8 ECs). **(c,d,e)** Bars represent median with +/−interquartile range. *p<0.05, **p<0.01 by two-tailed Mann-Whitney test. **(f)** Representative flow cytometry plots of co-expression patterns of *IL22* mRNA and *IL17F* mRNA in HIV-specific $CD4^+$ T cells in CP and EC people (n=6 CPs and ^ ECs). **(g)** Statistical analysis from (f) by SPICE (*p<0.05 by permutation test, 10000 permutations). Slices represent median frequency of *IL22*/*IL17F* $mRNA^+$ subpopulations of HIV-specific $CD4^+$ T cells (n= 6 CPs and 6 ECs). **(h,i)** Comparison of the frequencies of CCR6 and CXCR3 expression on *IL22* $mRNA^+$ and *IL17F* $mRNA^+$ HIV-specific $CD4^+$ T cells compared to total HIV-specific $CD4^+$ T cells in CPs (n=6) and ECs (n=6) by two-tailed Mann-Whitney test and by two-tailed Wilcoxon matched-pairs signed-ranked test. Only $p^{\wedge}0.05$ are displayed for clarity. Bars represent median with interquartile range. **(j,k)** Correlation of (**j**) *IL22* and (**k**) *IL-17F* mRNA levels assessed by qRT-PCR on sorted HIV Gag-specific $CD4^+$ T cells and $CD4^+$ T cell activation measured by HLA-DR and CD38 co-expression (two-tailed Spearman test; n= 10 CPs, 6 VCs, 12 ECs). (**l,m**) Statistical comparisons of microbial translocation in plasma in cohorts of CP, EC, and uninfected control donors (UD) (**l**) quantitation of bacterial RNA reads in plasma (transcripts per million, TPM); (**m**) bacterial RNA species diversity (Shannon entropy score) (Kruskal Wallis test; n= 10 CPs, 8 ECs and 6 UDs). Bars represent median +/−interquartile range. (**n,o**) Correlation between bacterial RNA species diversity in plasma and frequencies of (**n**) HIV-specific *IL22*$^+$ $CD4^+$ T cells and (**o**) HIV-specific *IL17*$^+$ $CD4^+$ T cells (two-tailed Spearman test; n=10 CPs and 8 ECs). (**p,q**) Correlation between abundance of Proteobacteria translocation and frequencies of (**p**) HIV-specific *IL22*$^+$ $CD4^+$ T cells and (**q**) HIV-specific *IL17*$^+$ $CD4^+$ T cells (two-tailed Spearman test; n=10 CPs and 8 ECs).

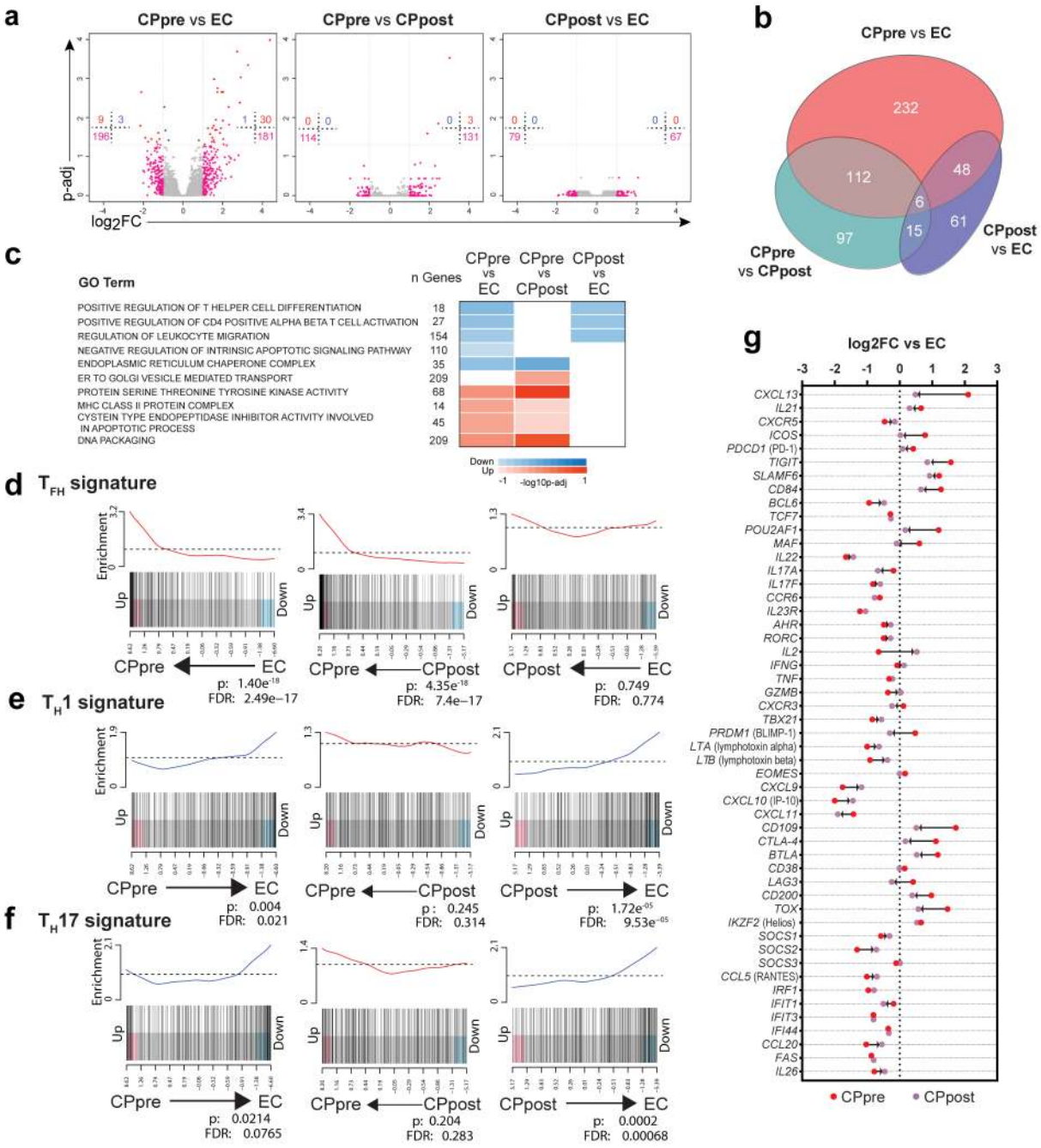

Figure 6. Distinct transcriptional imprint with reduction in T_{FH}-associated gene expression but poor correction of T_H1, T_H17 and T_H22 gene levels after suppression of viremia by ART

(a) Paired volcano plot comparisons depicting DEGs as measured by microarray analysis of sorted HIV-specific CD4+ T cells in 8 CPs before and after ART and 12 ECs. Red, pink and blue dots represent DEG with $\log_2|FC| > 1$ and padj<0.05, $\log_2 FC > 1$ or padj<0.05, respectively. Moderated two-sided t test followed by Benjamini-Hochberg method. **(b)** Euler diagrams of DEG between each pairwise comparison of the three groups of participants ($\log_2|FC| > 1$, $p < 0.05$); numbers represent DEGs. **(c)** Significant enriched Gene Ontology terms from the curated C5 MSigDB via CAMERA analysis. Blue, Red and white color

denote negative, positive or not significant (as defined by FDR<0.05, respectively. Two sided p-value by CAMERA and FDR by Benjamini-Hochberg method. **(d-f)** Barcode plots of enriched Thelper polarization gene sets by CAMERA in paired comparisons between the three groups (n= 8 CPpre 8 CPpost and 12 EC). Red and blue lines denote positive and negative enrichment, respectively. T_H1 signature: GSE59295 (T_H1 vs T_H2), T_H17 signature: GSE49703 (T_H17 vs T_H1), T_{FH} signature: GSE50391 (CXCR5high CD45RO vs CXCR5- tonsil samples). Two sided p-value by CAMERA and FDR by Benjamini-Hochberg method. **(g)** Differential expression of genes associated with CD4 T cell differentiation and function in HIV-specific $CD4^+$ T cells of CPs before and after treatment compared to ECs. Red and purple denote logarithmic fold change for comparisons CPpre. vs EC and CPpost vs EC, respectively, by microarray analysis in 8 CP before/after ART and 12 ECs.