TOX is expressed by exhausted and polyfunctional human effector memory CD8$^+$ T cells

Takuya Sekine[1,*], André Perez-Potti[1,*], Son Nguyen[2,3,*], Jean-Baptiste Gorin[1], Vincent Wu[2,3], Emma Gostick[4], Sian Llewellyn-Lacey[4], Sara Falck-Jones[5], Sindhu Vangeti[5], Meng Yu[5], Anna Smed-Sörensen[5], Ahmed Gaballa[6], Michael Uhlin[7,8], Johan K. Sandberg[1], Christian Brander[9,10,11], Piotr Nowak[12], Paul A. Goepfert[13], David A. Price[4,14], Michael R. Betts[2,3], Marcus Buggert[1,#]

[1]Center for Infectious Medicine, Department of Medicine Huddinge, Karolinska Institutet, Stockholm, Sweden

[2]Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[3]Institute for Immunology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[4]Division of Infection and Immunity, Cardiff University School of Medicine, University Hospital of Wales, Cardiff, UK

[5]Division of Immunology and Allergy, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden

[6]Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden

[7]Department of Applied Physics, Science for Life Laboratory, Royal Institute of Technology, Stockholm, Sweden

[8]Department of Immunology and Transfusion Medicine, Karolinska University Hospital, Stockholm, Sweden

[9]IrsiCaixa AIDS Research Institute, Badalona, Spain

[10]Universitat de Vic – Universitat Central de Catalunya, Vic, Spain

[11]Institució Catalana de Recerca i Estudis Avançats, ICREA, Barcelona, Spain

[12]Division of Clinical Microbiology, Department of Laboratory Medicine, Karolinska Institutet, Stockholm, Sweden

[13]Division of Infectious Diseases, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA

[14]Systems Immunity Research Institute, Cardiff University School of Medicine, University Hospital of Wales, Cardiff, UK

#Correspondence: Marcus Buggert, PhD, Karolinska Institutet, Alfred Nobels Allé 8, 141 52 Stockholm, Sweden. Email: marcus.buggert@ki.se. Tel: +46-739245224.

*These authors contributed equally to this work.

ABSTRACT

$CD8^+$ T cell exhaustion is a hallmark of many cancers and chronic infections. In mice, T cell factor 1 (TCF-1) preserves the functionality of exhausted $CD8^+$ T cells, whereas thymocyte selection-associated HMG box (TOX) is required for the epigenetic remodeling and survival of exhausted $CD8^+$ T cells. However, it has remained unclear to what extent these transcription factors play analogous roles in humans. In this study, we mapped the expression of TOX and TCF-1 as a function of differentiation and specificity in the human $CD8^+$ T cell landscape. Effector memory $CD8^+$ T cells generally expressed TOX, whereas naive and early-differentiated memory $CD8^+$ T cells generally expressed TCF-1. Cytolytic gene and protein expression signatures were also defined by the expression of TOX. In the context of a relentless immune challenge, exhausted human immunodeficiency virus (HIV)-specific $CD8^+$ T cells commonly expressed TOX, often in clusters with various activation markers and inhibitory receptors, and rarely expressed TCF-1. However, memory $CD8^+$ T cells specific for cytomegalovirus (CMV) or Epstein-Barr virus (EBV) also expressed TOX, either with or without TCF-1. A similar phenotype was observed among HIV-specific $CD8^+$ T cells from individuals who maintained exceptional immune control of viral replication. Collectively, these data suggest that TOX identifies exhausted and functionally replete human effector memory $CD8^+$ T cells specific for chronic viruses, such as CMV, EBV, and HIV.

INTRODUCTION

In the setting of many cancers and chronic viral infections, ongoing antigen exposure can lead to $CD8^+$ T cell exhaustion, a phenomenon characterized by profound epigenetic and transcriptional changes associated with a progressive loss of effector functions and the upregulation of various inhibitory receptors (IRs). This process likely evolved to prevent excessive immune activation, but it remains a major stumbling block in the quest to develop more effective immunotherapies and vaccines. The functional consequences of exhaustion can be overcome to some extent via the administration of checkpoint inhibitors, which block signals transduced by IRs, such as CTLA-4 and PD-1. Although therapeutic interventions based on this concept have revolutionized the treatment of various cancers in recent years, clinical responses are unpredictable and vary considerably among individuals with the same diagnosis. A more detailed understanding of $CD8^+$ T cell exhaustion is therefore required to inform translational efforts in the fields of immunoprophylaxis and immunotherapy (*1*).

It is widely held that the process of exhaustion drives $CD8^+$ T cells into a transcriptionally distinct lineage (*1*) that encompasses 'precursor exhausted' and 'terminally exhausted' subsets within a unique differentiation spectrum (*2*). Precursor exhausted $CD8^+$ T cells are thought to self-renew and differentiate into terminally exhausted $CD8^+$ T cells, which display effector functionality. Direct comparison of these subsets has further shown that precursor exhausted $CD8^+$ T cells less commonly express IRs and more commonly express certain transcription factors, including T cell factor 1 (TCF-1) (*2*). TCF-1 is a high mobility group (HMG) box transcription factor that plays a critical role in the differentiation and survival of memory $CD8^+$ T cells under physiological conditions (*3, 4*). Moreover, TCF-1 maintains stemness and sustains $CD8^+$ T cell responses against chronic viral infections (*5-7*).

There is no definitive marker that identifies exhausted $CD8^+$ T cells. However, recent work has shown that the transcription factor thymocyte selection-associated HMG box (TOX) distinguishes exhausted from memory $CD8^+$ T cells in various mouse models (*8-12*). These studies have demonstrated that TOX is absolutely required for the epigenetic remodeling and survival of exhausted $CD8^+$ T cells (*8-12*). In the classic dichotomy, infection with lymphocytic choriomeningitis virus (LCMV) clone 13 leads to chronic antigen stimulation and generates exhausted TOX^+ virus-specific $CD8^+$ T cell populations, which are clearly distinct from the memory $CD8^+$ T cell

populations that emerge in response to infection with the non-persistent strain LCMV Armstrong (*8, 9*). Tumor-specific $CD8^+$ T cells similarly express TOX and exhibit functional impairment associated with the upregulation of IRs (*9, 10*). It has nonetheless remained unclear how these data apply to acute and chronic viral infections in humans. This is a highly pertinent issue in the context of global health, because immune control of viral replication has been associated with the functional attributes of $CD8^+$ T cells (*13-17*).

Human immunodeficiency virus (HIV) infection is characterized by sustained viral replication and immune activation, leading to a progressive loss of $CD4^+$ T cells and the eventual development of AIDS in the absence of antiretroviral therapy (ART). $CD8^+$ T cells are essential for immune control of HIV (*18*). In a majority of cases, however, immune control is partial at best, and HIV-specific $CD8^+$ T cells display the hallmarks of exhaustion, including dysfunctionality, impaired proliferation, and the expression of various IRs, such as PD-1, TIGIT, 2B4, CD39, and Tim-3 (*13, 15, 19-22*). In contrast, rare individuals maintain highly functional HIV-specific $CD8^+$ T cell populations, which associate with effective viral suppression *(13, 23, 24).* Specific transcription factors, such as Eomes and BATF, have previously been linked with $CD8^+$ T cell exhaustion in the context of HIV infection (*15, 25*). It has nonetheless remained unclear to date how this process relates to the expression of TOX and TCF-1.

We investigated the expression of TOX and TCF-1 as a function of differentiation and specificity in the human $CD8^+$ T cell landscape. TOX was expressed primarily in effector memory $CD8^+$ T cells, whereas TCF-1 was expressed primarily in naive and early-differentiated memory $CD8^+$ T cells. Progressive HIV disease was associated with increased expression of TOX, together with various activation markers and IRs, and decreased expression of TCF-1. However, functionally replete $CD8^+$ T cells specific for cytomegalovirus (CMV) or Epstein-Barr virus (EBV), which are controlled effectively by the immune system, also expressed intermediate to high levels of TOX. A similar pattern was observed for HIV-specific $CD8^+$ T cells in cases where viral replication was controlled in the absence of ART. Collectively, these findings suggest a nuanced model in humans, wherein antigen-driven $CD8^+$ T cell exhaustion is not necessarily defined solely by the expression of TOX.

RESULTS

TOX and TCF-1 are differentially expressed among circulating CD8$^+$ T cells

Recent studies have examined the roles of TOX and TCF-1 in mouse models of chronic antigen-driven CD8$^+$ T cell exhaustion (*8-12*). However, the expression patterns of these transcription factors, especially in combination, are less well defined in humans. To address this issue, we compared the transcriptomes of naive T (T_N), central memory T (T_{CM}), effector memory T (T_{EM}), and effector memory RA T (T_{EMRA}) cells flow-sorted directly *ex vivo* as phenotypically distinct CD8$^+$ subsets from peripheral blood samples donated by HIV$^-$ individuals (Figure S1A). Analysis of the RNA-sequencing (RNA-seq) data revealed a core signature of genes that were differentially expressed among these CD8$^+$ T cell subsets (Table S1). In particular, *Tox* and *Tcf7*, which encode TOX and TCF-1, respectively, were among the top identified genes that distinguished naive and memory CD8$^+$ T cells (Figure 1A). *Tox* transcript levels were significantly elevated in T_{CM}, T_{EM}, and T_{EMRA} cells relative to T_N cells, whereas *Tcf7* transcripts were significantly elevated in T_N cells relative to T_{EM} and T_{EMRA} cells (Figure 1B). We then employed the Assay of Transposase-Accessible Chromatin using sequencing (ATAC-seq) to characterize the epigenetic landscape of resting T_N, T_{EM}, and T_{EMRA} cells. The *Tox* and *Tcf7* loci both contained open chromatin clusters that distinguished T_N cells from T_{EM} and T_{EMRA} cells (Figure 1C). Importantly, TCF-1 binding motifs were more readily accessible in the open chromatin regions (OCRs) of T_N cells relative to the OCRs of T_{EM} and T_{EMRA} cells (Figure 1D). Equivalent sequence-based analyses were not possible for TOX, which is thought to bind DNA in a structure-dependent manner (*26*).

Differential expression of TOX and TCF-1 between naive and memory CD8$^+$ T cells was validated at the protein level using Western blots (Figure 1E, S1B, and S1C). Flow cytometric analyses further confirmed that TOX was predominantly expressed in T_{EM} and T_{EMRA} cells, whereas TCF-1 was predominantly expressed in T_N and T_{CM} cells (Figure 1F). A previous study found that tonic antigen-driven stimulation triggered the expression of TOX (*9*). We corroborated this result and traced the expression patterns of TOX and TCF-1 in flow-sorted T_N, T_{CM}, T_{EM}, and T_{EMRA} cells labeled with a stable fluorescent dye. Most of the proliferating cells in each subset expressed TOX but not TCF-1 in response to stimulation via CD3 and CD28 (Figure 1G and 1H).

Collectively, these results show that TOX and TCF-1 are differentially expressed and regulated among naive and memory subsets of human $CD8^+$ T cells under physiological conditions.

Expression of *Tox and Tcf7* separates $CD8^+$ T cells into distinct clusters

To extend these findings, we analyzed a publicly available single-cell RNA-seq (scRNA-seq) dataset acquired from $CD8^+$ T cells (10x Genomics Repository). Non-linear relationships among individual $CD8^+$ T cells were assessed using Uniform Manifold Approximation and Projection (UMAP). *Tox* and *Tcf7* transcripts were readily detectable and distinctly expressed at the single-cell level (Figure 2A and S2A). We then binned $CD8^+$ T cells into Tox^+ and $Tcf7^+$ events (Figure 2B). Tox^+ cells expressed certain *IR* genes, including *Pdcd1, Tigit, Cd244, and Lag3*, whereas $Tcf7^+$ cells expressed genes associated with naive and early-differentiated memory T cells, including *Ccr7, Il7r, Sell, Nell2, Lef1, Bach2, Myc,* and *Id3* (Figure 2B and S2B). Multiple effector memory genes were also differentially expressed between Tox^+ and $Tcf7^+$ cells, including *Gzmb, Gzma, Prf1, Cx3cr1, Tbx21, Eomes, Zeb2,* and *Prdm1* (Figure 2B and Table S2).

Further analysis of the same dataset using Seurat identified eight different clusters on the basis of differential gene expression (Figure 2C). Two naive-like clusters (0 and 1) with the highest expression levels of *Tcf7* also expressed high levels of *Ccr7, Lef1*, and *Sell* (Table S3). In contrast, three effector memory-like clusters (5, 6, and 7) with the highest expression levels of *Tox* were enriched for effector-related genes, including *Ccl5, Gzma, Nkg7, Gzmh, Klrd1, Prf1, Gnly,* and *Fgfbp2* (Figure 2D and Table S3). Although not part of the core signature, certain IR genes, including *Pdcd1*, *Tigit*, *Lag3*, and *Lilrb1,* were also prefentially expressed among these Tox^+ clusters (Figure S3).

Collectively, these findings highlight an association between *Tox* and the expression of genes that encode various immune cell effector proteins and IRs.

TOX and TCF-1 are expressed in $CD8^+$ T cell clusters

To validate the scRNA-seq data acquired from $CD8^+$ T cells, we designed a 28-color flow cytometry panel to assess the expression of differentially regulated transcripts at the protein level in relation to TOX and TCF-1. Data were again concatenated and analyzed using UMAP (Figure S3A and S3B). Manual gating identified T_N, T_{CM}, T_{EM}, and T_{EMRA} clusters in the UMAP space (Figure 3A). Distinct topographical regions

were then delineated by overlaying the expression of TOX and TCF-1 (Figure 3B). Phenograph analysis revealed 27 unique cell clusters in the UMAP space (Figure 3C and S3C). TCF-1 was coexpressed with CCR7, CD45RA, and CD28 (Figure 3D). In contrast, TOX was expressed in clusters with increased expression of cytolytic proteins, namely granzyme B (GzmB) and perforin, and certain IRs, including PD-1, TIGIT, 2B4, and CD39 (Figure 3E). Moreover, specific TOX clusters exhibited high expression levels of cytolytic proteins (cluster 5), IRs (cluster 27), or both cytolytic proteins and IRs (cluster 26) (Figure 3E). Manual gating analysis further confirmed that almost every single TOX^+ cell expressed GzmB, perforin, PD-1, TIGIT, or 2B4 (Figure S3C).

Although TOX and TCF-1 were largely expressed in different cell clusters, we also identified a progenitor-like $TOX^+TCF\text{-}1^+$ population and a highly differentiated $TOX^+TCF\text{-}1^-$ population (Figure 3F). $TOX^+TCF\text{-}1^+$ cells expressed higher levels of the early differentiation markers CCR7 and CD28, as well as some IRs (PD-1 and TIGIT), relative to $TOX^+TCF\text{-}1^-$ cells, whereas $TOX^+TCF\text{-}1^-$ cells expressed higher levels of perforin, GzmB, CD38, 2B4, and Tim-3 relative to $TOX^+TCF\text{-}1^+$ cells (Figure 3F). $GzmB^+perforin^+$ cells were also more prevalent than either $GzmB^+perforin^-$ or $GzmB^-perforin^-$ cells in the $TOX^+TCF\text{-}1^-$ compartment (Figure S3D).

Collectively, these observations demonstrate the existence of distinct TOX^+ subpopulations in the $CD8^+$ T cell landscape, some of which lack expression of IRs.

HIV infection is associated with increased expression of TOX

To understand how an ongoing systemic infection impacts the balance between TOX and TCF-1 expression in human memory $CD8^+$ T cells, we recruited HIV^+ viremic individuals and HIV^+ aviremic individuals on ART (Table S4). In line with previous studies (*15, 19, 22, 25*), we found that some IRs, including PD-1, TIGIT, and 2B4, were expressed at higher frequencies on memory $CD8^+$ T cells from HIV^+ viremic donors relative to HIV^- donors (Figure 4A). Circulating memory $CD8^+$ T cells also more commonly expressed TOX and less commonly expressed TCF-1 in the context of HIV (Figure 4B). Increased frequencies of $TOX^+TCF\text{-}1^-$ cells were detected in HIV^+ viremic donors relative to HIV^- and HIV^+ aviremic donors (Figure 4B). A similar pattern was observed for $TOX^-TCF\text{-}1^-$ cells (Figure 4B), which typically displayed high levels of activation (data not shown). Moreover, $TOX^+TCF\text{-}1^+$ cells persisted at elevated frequencies in HIV^+ aviremic donors relative to HIV^- donors, whereas HIV^-

donors harbored increased frequencies of $TOX^{-}TCF\text{-}1^{+}$ cells relative to HIV^{+} viremic and HIV^{+} aviremic donors (Figure 4B).

Memory $CD8^{+}$ T cells that expressed PD-1, TIGIT, or 2B4 mostly coexpressed TOX, both in the absence and presence of actively replicating HIV (Figure 4C). However, $PD\text{-}1^{+}TIGIT^{+}2B4^{+}$ cells were more heavily skewed toward a $TOX^{+}TCF\text{-}1^{-}$ phenotype in HIV^{+} viremic donors relative to HIV^{-} donors (Figure S4A). We also noted that many TOX^{+} cells expressed cytolytic molecules and that not all TOX^{+} cells expressed IRs (Figure 4D).

In further analyses, we concatenated the multiparametric data and compared memory $CD8^{+}$ T cells from HIV^{-} donors with memory $CD8^{+}$ T cells from HIV^{+} viremic donors via UMAP (Figure 4E). A balanced TOX *versus* TCF-1 expression pattern was observed in HIV^{-} donors, whereas a skewed TOX^{+} phenotype was found in HIV^{+} viremic donors (Figure 4E and 4F). Specific markers that distinguished TOX^{+} and $TCF\text{-}1^{+}$ clusters (Figure 3) were then analyzed using Phenograph (Figure 4F and S5B). Again, TCF-1 was coexpressed with CCR7 and CD28, whereas TOX was expressed in clusters with increased expression of cytolytic proteins and/or IRs (Figure 4G). Memory $CD8^{+}$ T cells from HIV^{-} donors were found commonly in clusters with high expression levels of TCF-1 (clusters 17 and 9) and occasionally in clusters with detectable expression levels of TOX (clusters 18 and 10) (Figure 4H and 4I). TOX^{+} clusters in HIV^{-} donors nonetheless tended to express lower levels of PD-1, TIGIT, and CD38 relative to TOX^{+} clusters in HIV^{+} viremic donors (clusters 20, 2, 4, 6, 13, 3, 11, 23, 1, and 21), whereas the dominant TOX^{+} clusters in HIV^{+} viremic donors expressed higher levels of cytolytic proteins and IRs relative to the dominant TOX^{+} clusters in HIV^{-} donors (Figure 4H and 4I).

Collectively, these data show that memory $CD8^{+}$ T cells are skewed into a continuum of TOX^{+} phenotypes in the setting of chronic infection with HIV.

Memory $CD8^{+}$ T cells specific for chronic viral antigens express TOX

Studies in mice have investigated the expression of TOX and TCF-1 in LCMV infection models (*8, 9, 12*). However, it has remained unclear whether viral infections induce the expression of TOX in human $CD8^{+}$ T cells, which are similarly prone to antigen-driven exhaustion via a process that is not necessarily associated with the expression of IRs (*17, 27*). We therefore investigated the expression of TOX and TCF-1 in human $CD8^{+}$ T cell populations specific for influenza virus (Flu), which

establishes acute infections, CMV or EBV, which establish persistent infections controlled by the immune system, or HIV, which establishes persistent infections and continues to replicate vigorously in the absence of ART (Figure 5A and S5C). In line with studies of mice infected with LCMV (*8, 9, 12*), we found that Flu-specific $CD8^+$ T cells rarely expressed TOX and commonly expressed TCF-1, whereas $CD8^+$ T cells specific for CMV, EBV, or HIV commonly expressed TOX with or without TCF-1 (Figure 5B and S5D).

Data concatenation and UMAP visualization revealed specificity-based clustering of $CD8^+$ T cells (Figure 5C). Flu-specific $CD8^+$ T cells were found mainly in cluster 9, which displayed high expression levels of TCF-1 and CCR7 and low to intermediate expression levels of TOX, cytolytic proteins, and IRs (Figure 4H, 5D, and 5E). Many EBV-specific $CD8^+$ T cells also localized to cluster 9, but additional clusters were identified with detectable expression of TOX and variable expression levels of cytolytic proteins and IRs (clusters 12, 10, 2, 4, 6, 7, and 8) (Figure 4H, 5D, and 5E). In contrast, CMV-specific $CD8^+$ T cells predominated in clusters with high expression levels of TOX and cytolytic proteins and intermediate to high expression levels of IRs (clusters 11, 18, 21, and 23), whereas HIV-specific $CD8^+$ T cells were most prevalent in clusters 1, 2, 3, 4, 5, and 6, which displayed high expression levels of TOX, low to intermediate expression levels of TCF-1, and variable expression levels of cytolytic proteins and IRs (Figure 4H, 5D, and 5E).

To determine the functional relevance of TOX, we stimulated PBMCs with pools of peptides corresponding to known optimal epitopes derived from Flu, CMV, EBV, and HIV and measured GzmB and perforin alongside surface mobilization of CD107a and the intracellular production of IFN-γ, TNF, and IL-2. Virus-specific $CD8^+$ T cells that produced IL-2, an early-differentiated function, rarely expressed TOX, whereas virus-specific $CD8^+$ T cells that were loaded with cytolytic proteins and produced the effector cytokines IFN-γ and TNF commonly expressed TOX (Figure 5F).

Collectively, these observations demonstrate that $CD8^+$ T cells with different functional profiles and viral specificities occupy distinct phenotypic niches with differential expression of TOX and TCF-1.

HIV-specific memory $CD8^+$ T cells from elite controllers exhibit increased expression of TCF-1

To correlate these findings with immune efficacy, we recruited HIV$^+$ elite controllers (ECs), defined as individuals who maintain undetectable levels of plasma viral RNA in the absence of ART (Table S4). As expected, we found that PD-1, CD38, and CD39 were expressed at higher frequencies among HIV-specific CD8$^+$ T cells from HIV$^+$ viremic donors relative to HIV-specific CD8$^+$ T cells from ECs (Figure 6A). Irrespective of disease status, a vast majority of HIV-specific CD8$^+$ T cells expressed TOX (Figure 6B). However, TCF-1$^+$ cells were present at higher frequencies among HIV-specific CD8$^+$ T cells from ECs relative to HIV-specific CD8$^+$ T cells from HIV$^+$ viremic donors (Figure 6B).

In response to cognate peptide stimulation, higher frequencies of HIV-specific CD8$^+$ T cells from ECs produced TNF and IL-2 relative to HIV-specific CD8$^+$ T cells from either HIV$^+$ viremic donors or HIV$^+$ aviremic donors on ART (Figure 6C and S6A). Moreover, TNF$^+$ HIV-specific CD8$^+$ T cells expressed lower levels of TOX on a per cell basis relative to TNF$^-$ HIV-specific CD8$^+$ T cells (Figure S6B), IL-2$^+$ HIV-specific CD8$^+$ T cells expressed lower levels of TOX and higher levels of TCF-1 on a per cell basis relative to IL-2$^-$ HIV-specific CD8$^+$ T cells, and CD107a$^+$ HIV-specific CD8$^+$ T cells expressed lower levels of TCF-1 on a per cell basis relative to CD107a$^-$ HIV-specific CD8$^+$ T cells (Figure 6C). In addition, most GzmB$^+$perforin$^+$IFN-γ^+ HIV-specific CD8$^+$ T cells expressed TOX without TCF-1, whereas most IL-2$^+$ HIV-specific CD8$^+$ T cells expressed TCF-1 (Figure 6D).

Collectively, these results identify a functional dichotomy associated with the differential expression of TOX and TCF-1 that correlates with immune control of HIV.

DISCUSSION

In this study, we investigated the extent to which two HMG box transcription factors, TOX and TCF-1, shape the processes of exhaustion and memory differentiation among subpopulations of human $CD8^+$ T cells. Under physiological conditions, effector memory $CD8^+$ T cells primarily expressed TOX, whereas naive and early-differentiated memory $CD8^+$ T cells primarily expressed TCF-1. Cytolytic gene and protein expression signatures among lymphocyte lineages in general were also defined by the expression of TOX. In the context of ongoing viral replication, dysfunctional HIV-specific $CD8^+$ T cells commonly expressed TOX, which clustered with various activation markers and IRs, and rarely expressed TCF-1. However, functionally competent memory $CD8^+$ T cells specific for CMV or EBV also expressed TOX. A similar phenotype was observed among HIV-specific $CD8^+$ T cells from ECs.

TOX and TCF-1 were originally described as transcriptional regulators of lymphocyte development and maturation (*28-31*). More recent studies have shown that TOX is expressed in most exhausted $CD8^+$ T cells specific for various cancers and chronic viruses (*8-12*), whereas TCF-1 is upregulated in precursor exhausted $CD8^+$ T cells (*5-7*). Importantly, TOX and TCF-1 are both required for the survival of murine $CD8^+$ T cells that recognize chronic viral antigens (*8, 12*). In humans, TCF-1 has been associated with early-differentiated $CD8^+$ T cells (*32, 33*), whereas TOX has been associated with exhuasted $CD8^+$ T cells (*8, 9, 34, 35*). However, these correlative studies did not assess both transcription factors simultaneously in relation to different aspects of $CD8^+$ T cell immunobiology. We identified three discrete memory $CD8^+$ T cell populations in humans based on the expression of TOX and TCF-1. Early-differentiated T_{CM} and T_{EM} cells displayed a $TOX^-TCF\text{-}1^+$ phenotype with negligible expression of IRs. A vast majority of Flu-specific $CD8^+$ T cells were encapsulated within this population, consistent with the acute memory phenotype described in mice infected with LCMV (*9*). Most healthy donors also harbored a progenitor-like $TOX^+TCF\text{-}1^+$ subset, which commonly expressed CCR7 and CD28, and an effector-like $TOX^+TCF\text{-}1^-$ subset, characterized by the expression of cytolytic proteins. Moreover, IRs were expressed primarily in conjunction with TOX, but not all TOX^+ cells expressed IRs. This observation suggests that exhaustion *per se* is not necessarily defined solely by TOX. Indeed, we found that TOX was frequently expressed in CMV-specific $CD8^+$ T cells, which are highly cytolytic and functionally replete (*15, 36, 37*). These effector properties have also been shown to confer protection against simian immunodeficiency virus in nonhuman primate models *(38,*

39). Accordingly, TOX appears to define biologically protective human effector memory $CD8^+$ T cells, driven to a functional optimum by CMV.

In contrast to laboratory mice, humans are commonly infected with many different chronic viruses, including herpesviruses, such as CMV and EBV, polyomaviruses, and endogenous retroviruses (*40*). Ongoing antigen exposure is therefore likely universal across various specificities in healthy donors, potentially explaining the widespread occurrence of TOX^+ memory $CD8^+$ T cells. These less virulent chronic viruses have evolved in synergistic equilibrium with the host, such that the human immune system as it exists today represents the unfinished product of countless adaptations over millions of years (*41*). Accordingly, TOX may serve as a positive regulator of survival that maintains highly functional and protective $CD8^+$ T cell responses directed against ineradicable pathogens (*17, 42, 43*).

Previous studies have shown that exhaustion is stably imprinted at the epigenetic level (*17, 43*). Our data support the notion that a similar phenomenon applies in the case of TOX. For example, we found that HIV-specific $CD8^+$ T cells expressed TOX more commonly than Flu-specific $CD8^+$ T cells, even in donors with low plasma burdens of viral RNA. In addition, TOX was expressed less frequently in $CD8^+$ T cells from HIV^- donors relative to $CD8^+$ T cells from HIV^+ aviremic donors on ART. Of particular note, we also found that TCF-1 was expressed at relatively high frequencies alongside TOX in HIV-specific $CD8^+$ T cells from ECs. These cells typically produce multiple cytokines and proliferate vigorously in response to antigen encounter (*13, 16*). Moreover, we noted that IL-2 production was associated with increased expression levels of TCF-1 and decreased expression levels of TOX, whereas most cytolytic HIV-specific $CD8^+$ T cells displayed a $TOX^+TCF\text{-}1^-$ phenotype. A functional spectrum of HIV-specific $CD8^+$ T cells can therefore be defined based on the transcription factors TOX and TCF-1.

Our collective dataset provides an expression atlas for TOX and TCF-1 in the human $CD8^+$ T cell landscape under physiological conditions and under challenge from a relentless pathogen. Importantly, we found that TOX demarcates cytolytic effector $CD8^+$ T cells, which reside primarily in the T_{EM} and T_{EMRA} compartments, as well as exhausted $CD8^+$ T cells. On the basis of these findings, we propose that TOX is a universal regulator of human memory $CD8^+$ T cells specific for chronic viruses, such as CMV, EBV, and HIV.

MATERIALS AND METHODS

Samples

Peripheral blood was collected from individuals classified as HIV^- (n = 30), HIV^+ viremic (VL > 1,000 RNA copies/mL; n = 23), HIV^+ aviremic on ART (VL < 50 RNA copies/mL for > 1 year; n = 17), and HIV^+ elite controllers (VL < 50 RNA copies/mL for at least three consecutive visits in the absence of ART; n = 10). Recruitment occurred at four sites under protocols approved by the relevant Institutional Review Boards: Karolinska University Hospital (Stockholm, Sweden), Asociación Civil Impacta Salud y Educación (Lima, Peru), University of Alabama at Birmingham (Birmingham, AL, USA), and Massachusetts General Hospital (Boston, MA, USA). All samples were collected in accordance with the principles of the Declaration of Helsinki. Donor groups and clinical parameters are summarized in Table S4.

Flow cytometry

PBMCs were isolated from whole blood via standard density gradient centrifugation and cryopreserved in liquid nitrogen. Cryopreserved PBMC samples were thawed, resuspended at 2 x 10^6 cells/mL, and rested overnight at 37°C in complete medium (RPMI 1640 medium supplemented with 10% fetal bovine serum, 1% L-glutamine, and 1% penicillin/streptomycin) in the presence of 10 U/mL DNAse I (Roche) as described previously (*15, 44, 45*). Cells were then washed in fluorescence-activated cell sorting (FACS) buffer (phosphate-buffered saline [PBS] supplemented with 2% fetal bovine serum and 2 μM EDTA) and stained with MHC class I tetramers and/or a directly conjugated antibody specific for CCR7 (clone G043H7, BioLegend) for 10 minutes at 37°C. Other surface markers were detected via the subsequent addition of an optimized panel of directly conjugated antibodies for 20 minutes at room temperature, and viable cells were identified by exclusion using a LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Thermo Fisher Scientific). Cells were then washed again in FACS buffer and fixed/permeabilized using a FoxP3 / Transcription Factor Staining Buffer Set (eBioscience). Intracellular markers were detected via the subsequent addition of an optimized panel of directly conjugated antibodies for 1 hour at 4°C. Stained cells were fixed in PBS containing 1% paraformaldehyde (Biotium) and stored at 4°C. All samples were acquired within 24 hours using a FACSymphony A5 (BD Biosciences). Data were analyzed with FlowJo software version 10.6.1 (FlowJo LLC).

Antibodies

Directly conjugated antibodies with the following specificities were used in flow cytometry experiments: CD3–BV605, CD3–BV650, or CD3–BUV805 (clone UCHT1), CD8–BUV395 (clone RPA-T8), CD14–APC-Cy7 (clone MϕP9), CD15–BUV395 (clone W6D3), CD16–A700 (clone 3G8), CD19–BUV737 (clone SJ25C1), CD28–BUV563 (clone CD28.2), CD38–BUV496 (clone HIT2), CD49a–BUV615 (clone SR84), CD69–BV750 (clone FN50), CD95–BB630 (clone DX2), CD107a–PE-CF594 (clone H4A3), CXCR5–APC-R700 (clone RF8B2), granzyme B–A700 or granzyme B–BB790 (clone GB11), IFN-γ–FITC or IFN-γ–PE (clone B27), IL-2–APC-R700 (clone MQ1-17H12), Ki-67–BB660 (clone B56), Lag-3–BUV661 (clone T47-530), perforin–BB700 (clone dG9), TCRγδ–PE-CF594 (clone B1), TIGIT–BUV737 (clone 741182), and TNF–BV650 (clone Mab11) from BD Biosciences; CCR7–APC-Cy7 (clone G043H7), CD8–BV570 or CD8–BV605 (clone RPA-T8), CD14–BV510 (clone M5E2), CD19–BV510 (clone HIB19), CD45RA–BV570 or CD45RA–BV650 (clone HI100), CD39–BV711 (clone A1), CD56–BV711 (clone HCD56), CD103–BV605 (clone Ber-ACT8), CD127–BV421 or CD127–BV785 (clone A019D5), CD161–BV605 (clone HP-3G10), PD-1–PE-Cy7 (clone EH12.2H7), perforin–BV421 (clone B-D48), Tim-3–BV650 (clone F38-2E2), TCR Vα7.2–PE (clone 3C10), and 2B4–PE/Dazzle 594 (clone C1.7) from BioLegend; TCF-1–AF488 or TCF-1–PE (clone C63D9) from Cell Signaling Technology; TOX–AF647 (clone REA473) from Miltenyi Biotec; and CD4–PE-Cy5.5 (clone S3.5) and CD25–PE-Cy5 (clone CD25-3G10) from Thermo Fisher Scientific.

Peptides

Pools of peptides corresponding to known optimal epitopes derived from CMV (n = 14), EBV (n = 26), Flu (n = 17), and HIV (n = 22) were purchased from Peptides & Elephants GmbH. All peptides were synthesized at a purity of > 95%. Lyophilized peptides were reconstituted at 100 mg/mL in DMSO and further diluted to 100 μg/mL in PBS.

Tetramers

MHC class I tetramers conjugated to BV421 or PE were used to detect $CD8^+$ T cells with the following specificities: CMV NLVPMVATV (NV9/HLA-A*0201), EBV GLCTLVAML (GL9/HLA-A*0201), Flu GILGFVFTL (GL9/HLA-A*0201), HIV FLGKIWPSHK (FK10/HLA-A*0201), HIV ILKEPVHGV (IV9/HLA-A*0201), HIV SLYNTVATL (SL9/HLA-A*0201), EBV RVRAYTYSK (RK9/HLA-A*0301), Flu

RVLSFIKGTK (RK10/HLA-A*0301), CMV QYDPVAALFL (QL10/HLA-A*2402), EBV TYGPVFMCL (TL9/HLA-A*2402), HIV KYKLKHIVW (KW9/HLA-A*2402), HIV RYPLTFGW (RW8/HLA-A*2402), CMV TPRVTGGGAM (TM10/HLA-B*0702), HIV GPGHKARVL (GL9/HLA-B*0702), EBV RAKFKQLL (RL8/HLA-B*0801), HIV EIYKRWII (EI8/HLA-B*0801), HIV KRWIILGLNK (KK10/HLA-B*2705), HIV ISPRTLNAW (IW9/HLA-B*5701), HIV KAFSPEVIPMF (KF11/HLA-B*5701), and HIV QASQEVKNW (QW9/HLA-B*5701). All tetramers were generated in house as described previously (*46*).

Functional assays

PBMCs were seeded in complete medium at 2 x 10^6 cells/mL in 96-well V-bottom plates (Corning) with unconjugated anti-CD28 (clone L293) and anti-CD49d (clone L25) (each at 3 μL/mL; BD Biosciences), brefeldin A (1 μL/mL; Sigma-Aldrich), monensin (0.7 μL/mL; BD Biosciences), anti-CD107a–PE-CF594 (clone H4A3; BD Biosciences), and the relevant viral peptides (each at a final concentration of 0.5 μg/mL). Unstimulated negative controls were included in each assay. Cells were analyzed by flow cytometry after incubation for 6 hours at 37°C.

Proliferation assay

$CD8^+$ T cell subsets were flow-sorted using a MA900 Multi-Application Cell Sorter (Sony Biotechnology), labeled with CellTrace Violet (0.5 μM; Thermo Fisher Scientific), and resuspended in complete medium in 96-well U-bottom plates (Corning) with IL-2 (100 IU/mL; PeproTech) and ImmunoCult Human CD3/CD28 T Cell Activator (5 μL/mL; STEMCELL Technologies). IL-2 was replenished on day 3. Cells were analyzed by flow cytometry after incubation for 5 days at 37°C.

Activation of the calcineurin pathway

Naive $CD8^+$ T cells were cultured in complete medium at various densities with ImmunoCult Human CD3/CD28 T Cell Activator (5 μL/mL; STEMCELL Technologies) or ionomycin (100 ng/mL; Sigma-Aldrich) and/or phorbol myristate acetate (10 ng/mL; Sigma-Aldrich). Unstimulated negative controls were included in each assay. Cells were analyzed by flow cytometry after incubation for 24 hours at 37°C.

UMAP and PhenoGraph

FCS3.0 data files were imported into FlowJo software version 10.6.0 (FlowJo LLC). All samples were compensated electronically. Dimensionality reduction was

performed using the FlowJo plugin UMAP version 2.2 (FlowJo LLC). The downsample version 3.0.0 plugin and concatenation tool was used to visualize multiparametric data from comparable numbers of total $CD8^+$ T cells per healthy donor (n = 4). Comparable numbers of memory $CD8^+$ T cells from HIV^- (n = 2) and HIV^+ individuals (n = 2) were similarly applied to UMAP. The following parameters were used in these analyses: metric = euclidean, nearest neighbors = 15, and minimum distance = 0.5. Clusters of phenotypically related cells were then detected using Phenograph version 0.2.1. The following markers were considered in each case: CCR7, CD25, CD28, CD38, CD39, CD45RA, CD49a, CD95, CD103, CXCR5, granzyme B, Ki-67, Lag-3, perforin, PD-1, TCF-1, TIGIT, Tim-3, TOX, and 2B4. Phenograph clustering was performed using the coordinates from UMAP. Plots were generated using Prism version 8.2.0 (GraphPad Software Inc.).

RNA-seq

Naive and memory $CD8^+$ T cells (250 cells/subset) were flow-sorted directly into lysis buffer using a FACSAria II (BD Biosciences). RNA-seq libraries were prepared from snap-frozen lysates using a SMART-Seq v4 Ultra Low Input RNA Kit (Takara). Briefly, 3' oligo-dt primers were hybridized to the poly(A) tails of mRNA molecules, and cDNA was generated using SMARTScribe Reverse Transcriptase (Takara) and preamplified using SeqAmp DNA Polymerase (Takara). Cleanup was performed using Agencourt AMPure XP Beads (Beckman Coulter). cDNA was quantified using Qubit 3.0 (Thermo Fisher Scientific), and fragment sizes were evaluated using a 2100 BioAnalyzer (Agilent). PCR products were then indexed using a Nextera XT DNA Library Prep Kit (Illumina). Briefly, transcripts were tagmented using Amplicon Tagment Mix (Illumina) and indexed using a Nextera Index Kit (Illumina). Cleanup was performed again using Agencourt AMPure XP Beads (Beckman Coulter). Libraries were pooled, quantified, and sequenced across 75 base pairs (bp) using a paired-end approach with a 150-cycle high-output flow cell on a NextSeq 550 (Illumina). Three biological replicates were sequenced per experiment. Fastq files from replicate sequencing runs were concatenated and aligned using STAR software version 2.5.2a and hg38. Depth ranged from 8 million to 13.6 million reads per sample. Raw counts were normalized using DESeq2 (Bioconductor). Normalized data were then transformed to Z-scores and analyzed for differential expression using the limma package in R.

ATAC-seq

Naive and memory $CD8^+$ T cells (30,000–50,000 cells/subset) were flow-sorted directly into complete medium using a FACSAria II (BD Biosciences). ATAC-seq was carried out as described previously (*45*). Briefly, cells were pelleted, washed in PBS, and treated with lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% IGEPAL CA-630). Nuclear pellets were resuspended in transposition buffer containing Tn5 transposase (Illumina) for 30 minutes at 37°C. Tagmented DNA was purified using a MinElute Reaction Cleanup Kit (Qiagen). Amplified libraries were purified using a QIAQuick PCR Purification Kit (Qiagen) and subjected to paired-end sequencing (38 bp + 37 bp) on a NextSeq 550 (Illumina). Two biological replicates were sequenced per experiment. ATAC-seq peaks were called out in each replicate using the (*-p 1e-7 --nolambda --nomodel*) function in MACS2. A read count table was generated and used to identify differentially enriched regions among subsets using DESeq2 with the parameters fold change > 2 and $P < 0.05$. *De novo Tcf7* and *Tox* motif analyses were carried out using HOMER. Peaks were visualized using Integrative Genomics Viewer software version 2.3.93.

Single-cell RNA-seq

Single-cell RNA-seq data were acquired from the 10x Genomics Repository (https://support.10xgenomics.com/single-cell-vdj/datasets/). Expression matrices were reconstructed from Feature-Barcode Matrices (https://github.com/Jx-b/10x_notebooks/blob/master/CellRanger%20output%20processing.ipynb). UMAP plots and differential gene expression profiles were generated using the Loupe Cell Browser (10x Genomics). Four healthy donors were analyzed for differential expression of genes between the Tox^+Tcf7^- and Tox^-Tcf7^+ $CD8^+$ T cell subsets (total n = 245,883 $CD8^+$ T cells). Heatmaps and violin plots were generated from Tox^+ (n = 3,187) and $Tcf7^+$ $CD8^+$ T cells (n = 12,952) using the Python Open Source Graphing Library (Plotly). Lineage analyses were conducted using a total of 7,865 PBMCs. Raw read counts were analyzed using Monocle 3 (*47-49*). Data were processed using the “preprocess_cds” function (num_dim = 100). The number of dimensions was reduced using “reduce_dimension”, and cells were clustered using “cluster_cells”. UMAP plots were generated using “plot_cells” (with “scale_colour_distiller” overwrite for plots containing overlays of specific gene expression). Pseudotime trajectories were generated using “learn_graph”.

Western blot

Naive and memory $CD8^+$ T cells were sorted either directly or after stimulation into complete medium using a MA900 Multi-Application Cell Sorter (Sony Biotechnology).

Cells were then pelleted, washed twice in ice-cold PBS, and lyzed in RIPA Lysis and Extraction Buffer (Thermo Fisher Scientific). Protein content was quantified using a Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Normalized amounts of protein in PBS were mixed with 4x loading buffer (62.5 mM Tris-HCl pH 6.8, 2% [w/v] SDS, 10% glycerol, 0.04 M DTT, and 0.01% [w/v] bromophenol blue) and heated for 10 minutes at 90°C. Equal volumes were loaded onto 10% Precast Polyacrylamide Gels (Bio-Rad). Gel electrophoresis was performed for 60 minutes at 130 V. Proteins were transferred to a polyvinylidene fluoride membrane over 1 hour at 100 V using a Mini-PROTEAN Tetra Cell and Mini Trans-Blot Module (Bio-Rad). Membranes were blocked with 5% skimmed milk in TBS-T (150 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.1% Tween-20) and incubated with the relevant antibodies overnight at 4°C. The following antibodies were used in these experiments: anti-GAPDH (clone 6C5) diluted at 1:2,000 (Abcam), anti-TCF1 (clone 655202) diluted at 1:1,000 (BioLegend), and anti-TOX (rabbit polyclonal ab155768) diluted at 1:700 (Abcam). Membranes were then washed in TBS-T, incubated with secondary antibodies conjugated to horseradish peroxidase (diluted at 1:4,000) for 1 hour at room temperature, and developed using West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific). Antibody dilutions were prepared using 2% milk in TBS-T. Blots were imaged using a G:BOX Imaging System (Syngene).

Statistics

Differences between unmatched groups were compared using an unpaired t-test or the Mann-Whitney *U* test, and differences between matched groups were compared using a paired t-test or the Wilcoxon signed-rank test. Correlations were assessed using the Pearson correlation or the Spearman rank correlation. Non-parametric tests were used if the data were not distributed normally according to the Shapiro-Wilk normality test. All analyses were performed using R studio or Prism 7.0 (GraphPad). Phenotypic relationships within multivariate data sets were visualized using FlowJo software version 10.6.1 (FlowJo LLC).

ACKNOWLEDGEMENTS

We express our gratitude to all donors, health care personnel, study coordinators, administrators, and laboratory managers involved in this work.

FUNDING

M.B. was supported by the Swedish Research Council, the Karolinska Institutet, the Swedish Society for Medical Research, the Jeansson Stiftelser, the Åke Wibergs Stiftelse, the Swedish Society of Medicine, the Cancerfonden, the Barncancerfonden, the Magnus Bergvalls Stiftelse, the Lars Hiertas Stiftelse, the Swedish Physician against AIDS foundation, the Jonas Söderquist Stiftelse, and the Clas Groschinskys Minnesfond. Additional support was provided by R01 and R56 grants from the National Institutes of Health (AI076066, AI118694, and AI106481 to M.R.B.) and the Penn Center for AIDS Research (AI045008). D.A.P. is a Welcome Trust Senior Investigator (100326/Z/12/Z).

AUTHOR CONTRIBUTIONS

M.B. conceived the project; T.S., A.P.P., S.N., J.P.G., V.W., S.F.J., S.V., M.Y., A.S.S., A.G., D.A.P., M.R.B., and M.B. designed and performed experiments; T.S., A.P.P., S.N., J.B.G., V.W., M.U., and M.B. analyzed data; J.K.S., C.B., P.N., E.G., S.L.L., P.A.G., D.A.P., M.R.B., and M.B. provided critical resources; D.A.P., M.R.B., and M.B. supervised experiments; T.S., A.P.P., S.N., M.R.B., and M.B. drafted the manuscript; C.B., P.A.G., D.A.P., M.R.B., and M.B. edited the manuscript.

COMPETING INTERESTS

The authors declare that they have no competing financial interests, patents, patent applications, or material transfer agreements associated with this study.

DATA AND MATERIALS AVAILABILITY

The sequence datasets reported in this paper have been deposited in the Gene Expression Omnibus under accession no. XXX.

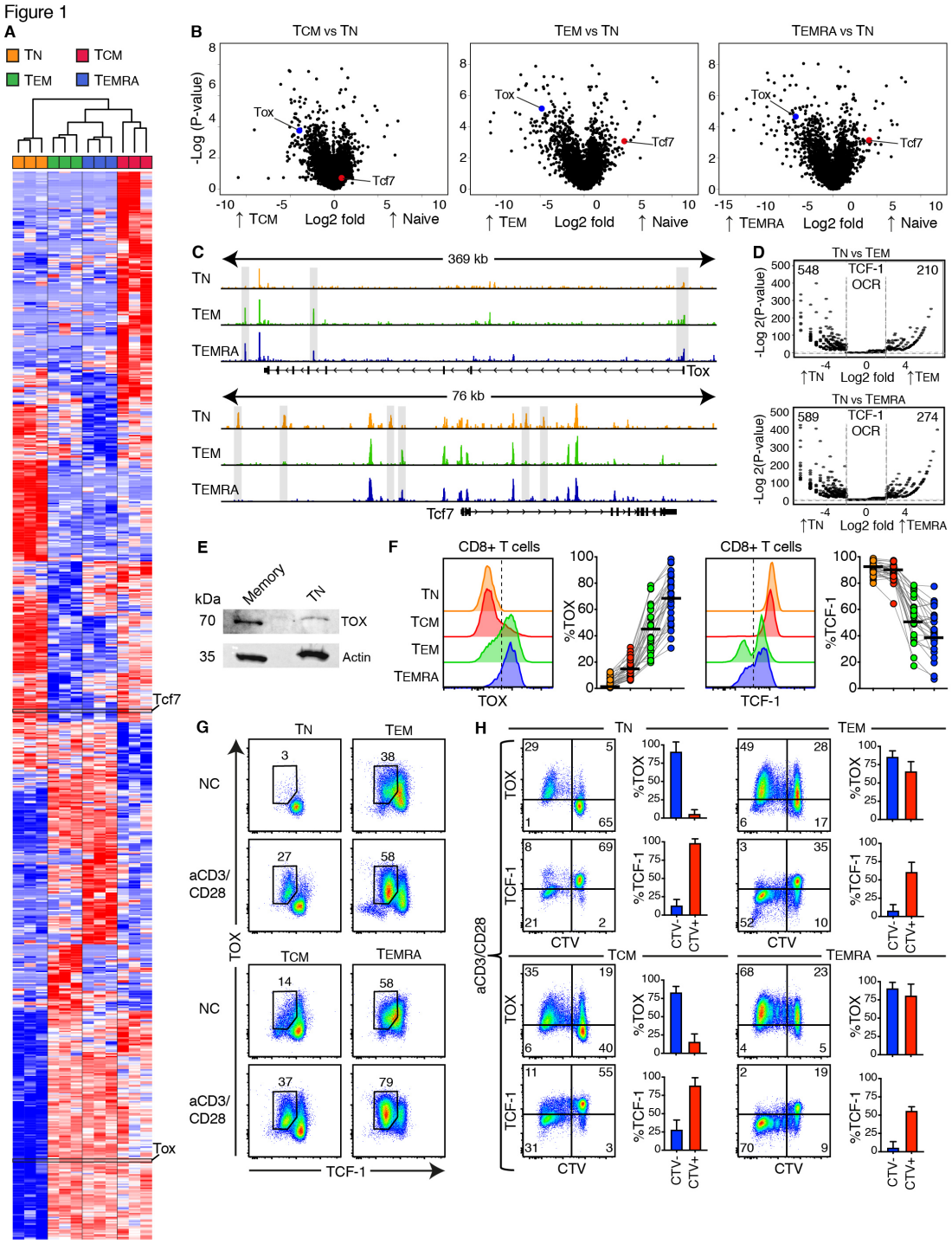

Figure 1. Expression of *Tox* and *Tcf7* in CD8$^+$ T cell subsets. (**A**) RNA-seq heatmap showing differentially expressed genes (fold change > 2; $P < 0.05$) among T_N, T_{CM}, T_{EM}, and T_{EMRA} cells (n = 3 healthy donors). (**B**) Volcano plots comparing gene expression between T_N cells and each subset of memory CD8$^+$ T cells (n = 1 healthy donor). (**C**) ATAC-seq tracks showing enrichment of OCRs adjacent to *Tox* (top) and *Tcf7* (bottom) for T_N, T_{EM}, and T_{EMRA} cells (n = 1 healthy donor). (**D**) Enrichment of the *Tcf7* motif in OCRs comparing T_N *versus* T_{EM} cells and T_N *versus* T_{EMRA} cells (n = 1 healthy donor). (**E**) Western blot analysis of TOX expression in T_N

versus memory (non-naive) $CD8^+$ T cells. Actin was included as a loading control. (**F**) Representative histograms and donor-matched graphs showing percent expression of TOX and TCF-1 in T_N, T_{CM}, T_{EM}, and T_{EMRA} cells. (**G**) Expression of TOX and TCF-1 in T_N, T_{CM}, T_{EM}, and T_{EMRA} cells after stimulation with ImmunoCult Human CD3/CD28 T Cell Activator for 5 days (aCD3/CD28). NC, negative control. (**H**) Details as in G. Proliferating cells were identified using CellTrace Violet (CTV).

Figure 2

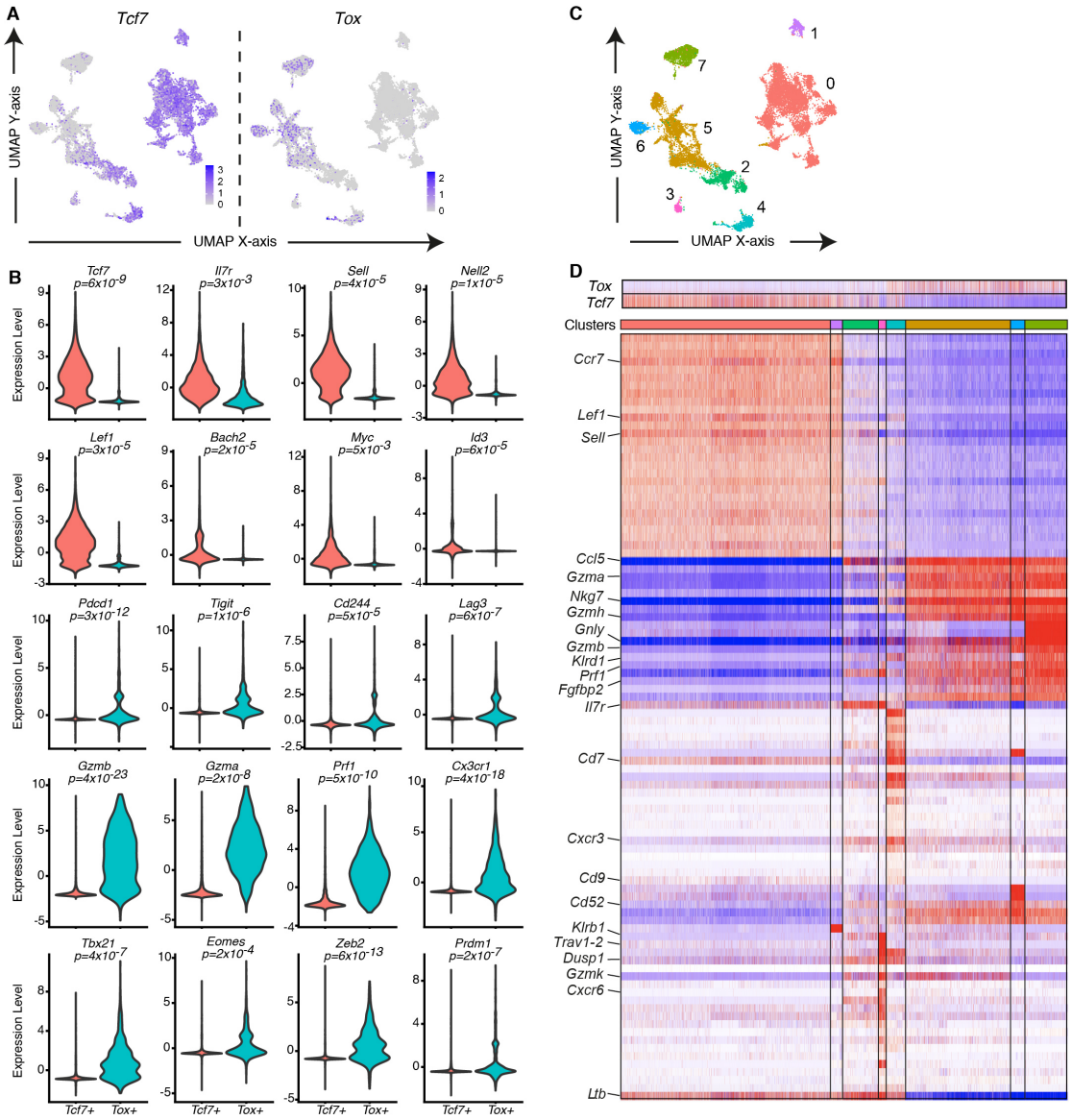

Figure 2. Expression of *Tox* and *Tcf7* in the CD8$^+$ T cell landscape. (**A**) scRNA-seq analysis of circulating CD8$^+$ T cells (n = 2 healthy donors). The UMAP plot illustrates the distribution of *Tox*$^+$ (>1 read) and *Tcf7*$^+$ cells (>1 read). (**B**) Violin plots showing immune-related genes that were differentially expressed between *Tox*$^+$ and *Tcf7*$^+$ cells (n = 2 healthy donors). Differentially expressed genes were identified using a threshold of $P < 0.05$. (**C**) Clustering analysis of the same dataset using Seurat. (**D**) Heatmap showing genes that were differentially expressed among the clusters identified in C. Highlighted genes encode proteins with known differential expression between naive and memory T cell subsets.

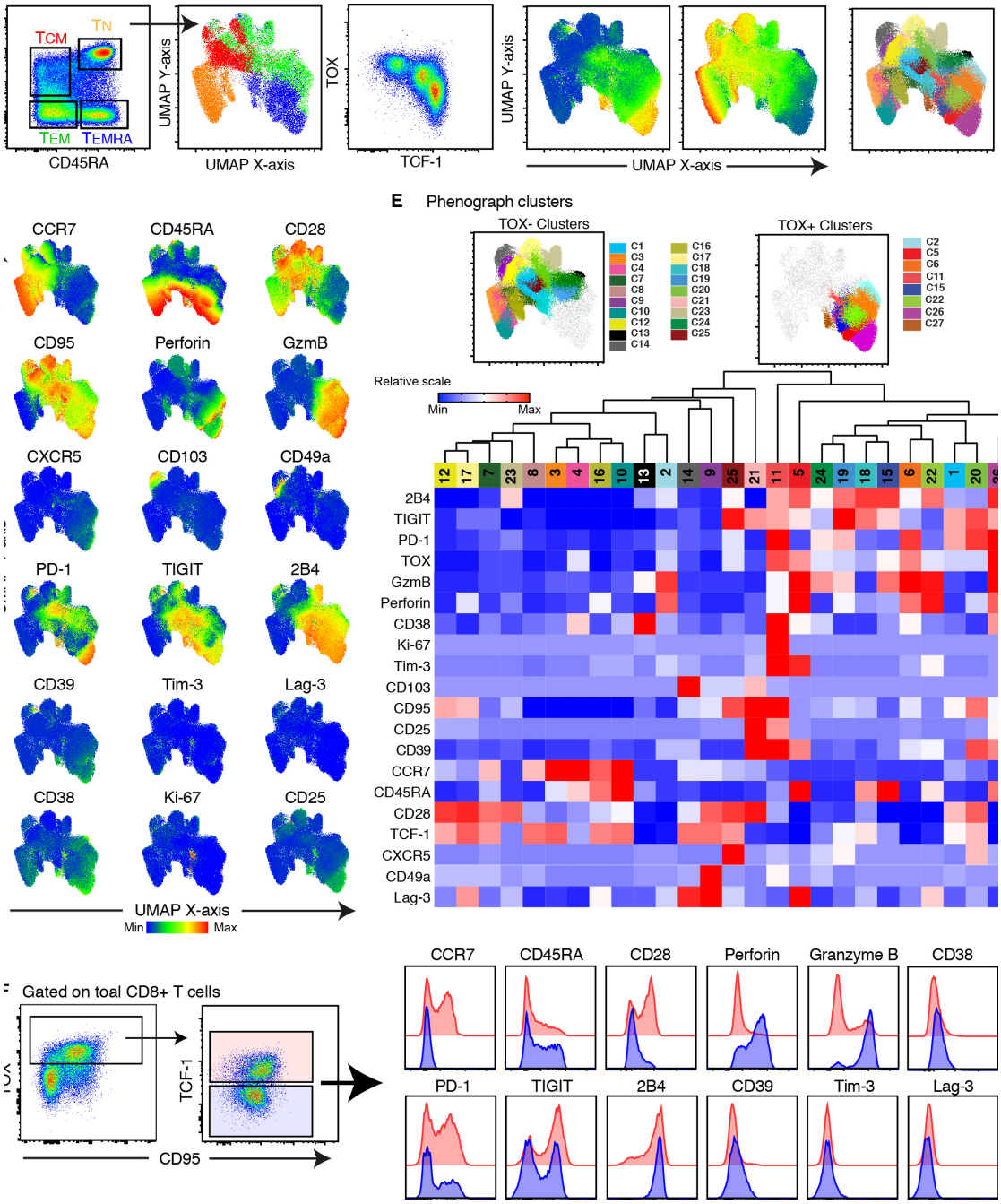

Figure 3. Expression of TOX and TCF-1 in CD8$^+$ T cell clusters.

(**A**) Left: flow cytometric gating strategy for the identification of T_N, T_{CM}, T_{EM}, and T_{EMRA} cells. Right: subset distribution on the UMAP plot generated from bulk CD8$^+$ T cells after data concatenation (n = 4 healthy donors). (**B**) Left: expression of TOX and TCF-1 in bulk CD8$^+$ T cells. Right: expression intensities projected on the corresponding UMAP plots. (**C**) The same UMAP display with subpopulations colored using Phenograph (n = 27 clusters). (**D**) UMAP plots showing expression patterns for the indicated markers. (**E**) Hierarchical clustering of expression intensity (z-score) for each of the indicated markers in each cluster derived using Phenograph. (**F**) Top left: flow cytometric gating strategy for the separation of TOX$^+$

cells into TCF-1^{-} and TCF-1^{+} populations. Top right: histograms showing expression of the indicated markers among TOX^{+}TCF-1^{+} (red) and TOX^{+}TCF-1^{-} cells (blue). Bottom: scatter plots showing expression frequencies for the indicated markers among TOX^{+}TCF-1^{+} and TOX^{+}TCF-1^{-} cells from all healthy donors. *P < 0.05, **P < 0.01, ***P < 0.001.

Figure 4

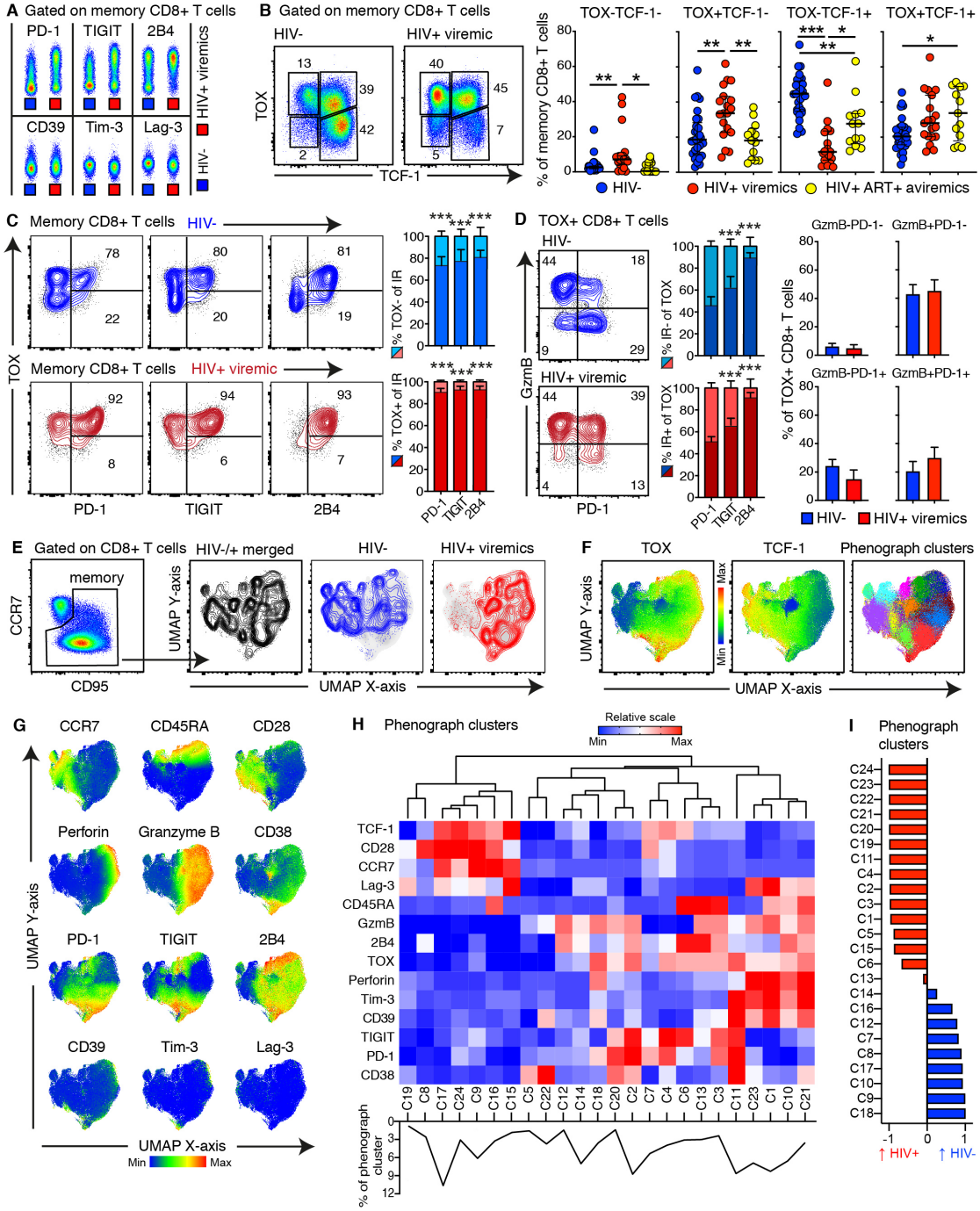

Figure 4. Expression of TOX and TCF-1 in donors infected with HIV. (**A**) Expression of the indicated IRs on memory CD8$^+$ T cells from a representative HIV$^-$ donor (blue boxes) and a representative HIV$^+$ viremic donor (red boxes). (**B**) Left: expression of TOX and TCF-1 in memory CD8$^+$ T cells. Right: scatter plots showing TOX$^{+/-}$TCF-1$^{+/-}$ population frequencies among memory CD8$^+$ T cells from HIV$^-$ (blue), HIV$^+$ viremic (red), and HIV$^+$ aviremic donors (yellow). (**C**) Left: coexpression of TOX *versus* PD-1, TIGIT, and 2B4 among memory CD8$^+$ T cells from a representative HIV$^-$ donor (top) and a representative HIV$^+$ viremic donor (bottom). Right: stack bars showing the frequencies of TOX$^-$ and TOX$^+$ cells among IR$^+$

memory $CD8^+$ T cells (n = 30 HIV^- donors in blue and n = 17 HIV^+ viremic donors in red). (**D**) Left: coexpression of PD-1 *versus* GzmB among $CD8^+$ T cells pre-gated on TOX. Middle: stack bars showing the frequencies of IR^- and IR^+ cells among TOX^+ memory $CD8^+$ T cells (n = 30 HIV^- donors in blue and n = 17 HIV^+ viremic donors in red). Right: bar graphs showing the distribution of TOX^+ cells among PD-$1^{+/-}$GzmB$^{+/-}$ populations of $CD8^+$ T cells (n = 30 HIV^- donors and n = 17 HIV^+ viremic donors). (**E**) Left: flow cytometric gating strategy for the identification of memory $CD8^+$ T cells, designed to include stem cell-like clusters. Right: UMAP plots generated from memory $CD8^+$ T cells after data concatenation. The colored plots show the distribution of memory $CD8^+$ T cells in the UMAP space for HIV^- (n = 2; blue) and HIV^+ viremic donors (n = 2; red). (**F**) Left: UMAP plots showing expression intensities for TOX and TCF-1. Right: the same UMAP display with subpopulations colored using Phenograph (n = 24 clusters). (**G**) UMAP plots showing expression patterns for the indicated markers. (**H**) Hierarchical clustering of expression intensity (z-score) for each of the indicated markers in each cluster derived using Phenograph. (**I**) Distribution of clusters between HIV^- (n = 2; blue) and HIV^+ viremic donors (n = 2; red). **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

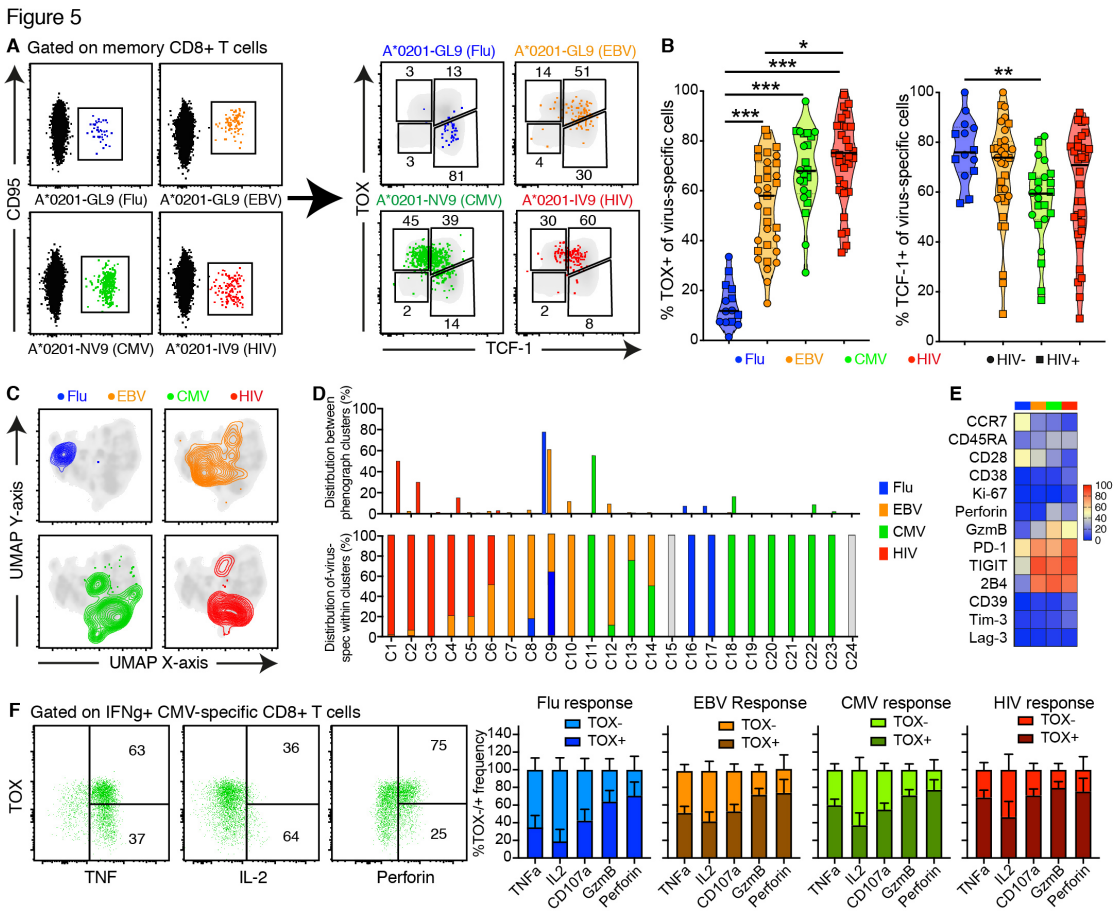

Figure 5. Expression of TOX and TCF-1 in virus-specific CD8$^+$ T cells. (**A**) Left: flow cytometric identification of memory CD8$^+$ T cells specific for Flu (blue), EBV (orange), CMV (green), or HIV (red) using MHC class I tetramers. Right: overlays showing the distribution of each virus-specific population in the TOX$^{+/-}$TCF-1$^{+/-}$ landscape. (**B**) Percent expression of TOX (left) and TCF-1 (right) in virus-specific memory CD8$^+$ T cells. (**C**) Distribution of virus-specific memory CD8$^+$ T cells on the UMAP plot derived in Figure 4 (total n = 4 healthy donors). (**D**) Top: distribution of virus-specific memory CD8$^+$ T cells among clusters derived using Phenograph. Bottom: summed distribution of virus-specific memory CD8$^+$ T cells in each cluster derived using Phenograph. Total n = 4 healthy donors. (**E**) Percent expression of the indicated markers among virus-specific memory CD8$^+$ T cells (total n = 45 healthy donors). (**F**) Left: coexpression of TOX *versus* TNF, IL-2, and perforin among CMV-specific CD8$^+$ T cells pre-gated on IFN-γ. Right: stack bars showing the frequencies of TOX$^-$ and TOX$^+$ cells among IFN-γ$^+$ virus-specific CD8$^+$ T cells (Flu: n = 25 healthy donors; EBV: n = 27 healthy donors; CMV: n = 30 healthy donors; HIV: n = 25 healthy donors). **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

Figure 6

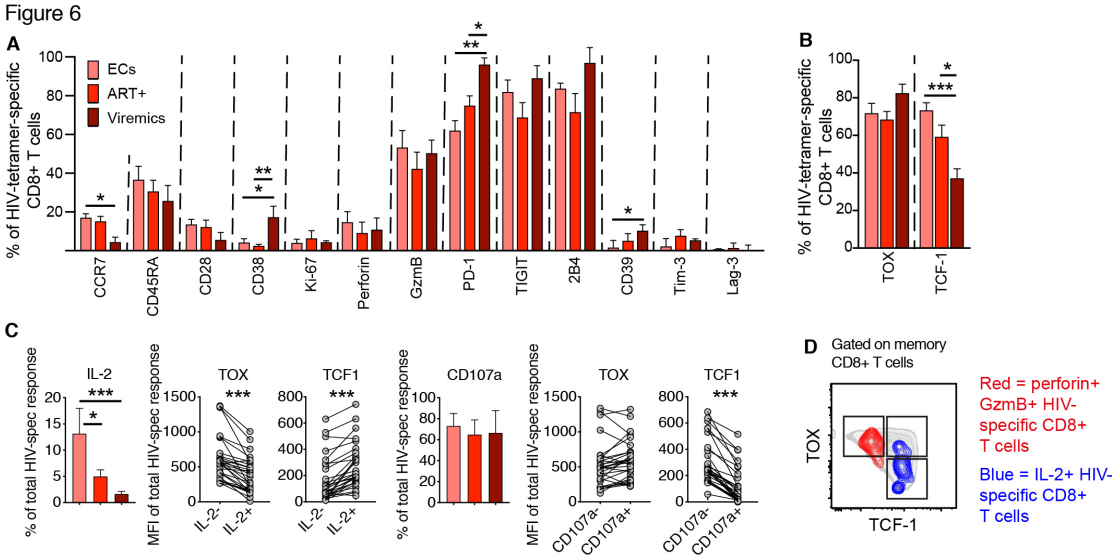

Fig. 6. Expression of TOX and TCF-1 in functional HIV-specific CD8$^+$ T cells. (**A**) Percent expression of the indicated markers among tetramer-defined HIV-specific CD8$^+$ T cells from ECs (n = 11; light red), HIV$^+$ aviremic donors on ART (n = 15; red), and HIV$^+$ viremic donors (n = 16; dark red). (**B**) Percent expression of TOX and TCF-1 in tetramer-defined HIV-specific CD8$^+$ T cells from the donor groups in A. (**C**) Colored graphs: percent frequencies of IL-2$^+$ (left) and CD107a$^+$ cells (right) among all responsive HIV-specific CD8$^+$ T cells from the donor groups in A after stimulation of PBMCs with pools of peptides corresponding to known optimal epitopes derived from HIV. Donor-matched graphs: expression intensities of TOX and TCF-1 in IL-2$^-$ *versus* IL-2$^+$ (left) and CD107a$^-$ *versus* CD107a$^+$ HIV-specific CD8$^+$ T cells (right). (**D**) Expression of TOX and TCF-1 in GzmB$^+$perforin$^+$IFN-γ^+ (red) and IL-2$^+$ HIV-specific CD8$^+$ T cells (blue). MFI, median fluorescence intensity. *P < 0.05, **P < 0.01, ***P < 0.001.

REFERENCES

1. L. M. McLane, M. S. Abdel-Hakeem, E. J. Wherry, CD8 T cell exhaustion during chronic viral infection and cancer. *Annu Rev Immunol* **37**, 457–495 (2019).
2. A. Kallies, D. Zehn, D. T. Utzschneider, Precursor exhausted T cells: key to successful immunotherapy? *Nat Rev Immunol* **20**, 128–136 (2020).
3. G. Jeannet *et al.*, Essential role of the Wnt pathway effector Tcf-1 for the establishment of functional CD8 T cell memory. *Proc Natl Acad Sci U S A* **107**, 9777–9782 (2010).
4. X. Zhou *et al.*, Differentiation and persistence of memory $CD8^+$ T cells depend on T cell factor 1. *Immunity* **33**, 229–240 (2010).
5. S. J. Im *et al.*, Defining $CD8^+$ T cells that provide the proliferative burst after PD-1 therapy. *Nature* **537**, 417–421 (2016).
6. D. T. Utzschneider *et al.*, T cell factor 1-expressing memory-like $CD8^+$ T cells sustain the immune response to chronic viral infections. *Immunity* **45**, 415–427 (2016).
7. T. Wu *et al.*, The TCF1-Bcl6 axis counteracts type I interferon to repress exhaustion and maintain T cell stemness. *Sci Immunol* **1**, eaai8593 (2016).
8. F. Alfei *et al.*, TOX reinforces the phenotype and longevity of exhausted T cells in chronic viral infection. *Nature* **571**, 265–269 (2019).
9. O. Khan *et al.*, TOX transcriptionally and epigenetically programs $CD8^+$ T cell exhaustion. *Nature* **571**, 211–218 (2019).
10. A. C. Scott *et al.*, TOX is a critical regulator of tumour-specific T cell differentiation. *Nature* **571**, 270–274 (2019).
11. H. Seo *et al.*, TOX and TOX2 transcription factors cooperate with NR4A transcription factors to impose $CD8^+$ T cell exhaustion. *Proc Natl Acad Sci U S A* **116**, 12410–12415 (2019).
12. C. Yao *et al.*, Single-cell RNA-seq reveals TOX as a key regulator of $CD8^+$ T cell persistence in chronic infection. *Nat Immunol* **20**, 890–901 (2019).
13. M. R. Betts *et al.*, HIV nonprogressors preferentially maintain highly functional HIV-specific $CD8^+$ T cells. *Blood* **107**, 4781–4789 (2006).
14. C. U. Blank *et al.*, Defining 'T cell exhaustion'. *Nat Rev Immunol* **19**, 665–674 (2019).
15. M. Buggert *et al.*, T-bet and Eomes are differentially linked to the exhausted phenotype of $CD8^+$ T cells in HIV infection. *PLoS Pathog* **10**, e1004251 (2014).
16. S. A. Migueles *et al.*, Lytic granule loading of $CD8^+$ T cells is required for HIV-infected cell elimination associated with immune control. *Immunity* **29**, 1009–1021 (2008).
17. D. T. Utzschneider *et al.*, T cells maintain an exhausted phenotype after antigen withdrawal and population reexpansion. *Nat Immunol* **14**, 603–610 (2013).
18. M. Buggert, A. S. Japp, M. R. Betts, Everything in its right place: resident memory $CD8^+$ T cell immunosurveillance of HIV infection. *Curr Opin HIV AIDS* **14**, 93–99 (2019).
19. C. L. Day *et al.*, PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. *Nature* **443**, 350–354 (2006).
20. P. K. Gupta *et al.*, CD39 expression identifies terminally exhausted $CD8^+$ T cells. *PLoS Pathog* **11**, e1005177 (2015).

21. R. B. Jones *et al.*, Tim-3 expression defines a novel population of dysfunctional T cells with highly elevated frequencies in progressive HIV-1 infection. *J Exp Med* **205**, 2763–2779 (2008).
22. J. Tauriainen *et al.*, Perturbed $CD8^+$ T cell TIGIT/CD226/PVR axis despite early initiation of antiretroviral treatment in HIV infected individuals. *Sci Rep* **7**, 40354 (2017).
23. Y. Fukazawa *et al.*, B cell follicle sanctuary permits persistent productive simian immunodeficiency virus infection in elite controllers. *Nat Med* **21**, 132–139 (2015).
24. A. Saez-Cirion *et al.*, HIV controllers exhibit potent CD8 T cell capacity to suppress HIV infection ex vivo and peculiar cytotoxic T lymphocyte activation phenotype. *Proc Natl Acad Sci U S A* **104**, 6776–6781 (2007).
25. M. Quigley *et al.*, Transcriptional analysis of HIV-specific $CD8^+$ T cells shows that PD-1 inhibits T cell function by upregulating BATF. *Nat Med* **16**, 1147–1151 (2010).
26. E. O'Flaherty, J. Kaye, TOX defines a conserved subfamily of HMG-box proteins. *BMC Genomics* **4**, 13 (2003).
27. J. Duraiswamy *et al.*, Phenotype, function, and gene expression profiles of programmed death-1^{hi} CD8 T cells in healthy human adults. *J Immunol* **186**, 4200–4212 (2011).
28. P. Aliahmad, B. de la Torre, J. Kaye, Shared dependence on the DNA-binding factor TOX for the development of lymphoid tissue-inducer cell and NK cell lineages. *Nat Immunol* **11**, 945–952 (2010).
29. P. Aliahmad, J. Kaye, Development of all CD4 T lineages requires nuclear factor TOX. *J Exp Med* **205**, 245–256 (2008).
30. J. L. Johnson *et al.*, Lineage-determining transcription factor TCF-1 initiates the epigenetic identity of T cells. *Immunity* **48**, 243–257 e210 (2018).
31. C. R. Seehus *et al.*, The development of innate lymphoid cells requires TOX-dependent generation of a common innate lymphoid cell progenitor. *Nat Immunol* **16**, 599–608 (2015).
32. R. Kratchmarov, A. M. Magun, S. L. Reiner, TCF1 expression marks self-renewing human $CD8^+$ T cells. *Blood Adv* **2**, 1685–1690 (2018).
33. M. Miron *et al.*, Human lymph nodes maintain TCF-1^{hi} memory T cells with high functional potential and clonal diversity throughout life. *J Immunol* **201**, 2132–2140 (2018).
34. B. Bengsch *et al.*, Epigenomic-guided mass cytometry profiling reveals disease-specific features of exhausted CD8 T cells. *Immunity* **48**, 1029–1045 e1025 (2018).
35. D. Wieland *et al.*, $TCF1^+$ hepatitis C virus-specific $CD8^+$ T cells are maintained after cessation of chronic antigen stimulation. *Nat Commun* **8**, 15050 (2017).
36. S. G. Hansen *et al.*, Effector memory T cell responses are associated with protection of rhesus monkeys from mucosal simian immunodeficiency virus challenge. *Nat Med* **15**, 293–299 (2009).
37. A. Harari, F. Bellutti Enders, C. Cellerai, P. A. Bart, G. Pantaleo, Distinct profiles of cytotoxic granules in memory CD8 T cells correlate with function, differentiation stage, and antigen exposure. *J Virol* **83**, 2862–2871 (2009).
38. S. G. Hansen *et al.*, Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature* **473**, 523–527 (2011).

39. S. G. Hansen *et al.*, Immune clearance of highly pathogenic SIV infection. *Nature* **502**, 100–104 (2013).
40. H. W. Virgin, E. J. Wherry, R. Ahmed, Redefining chronic viral infection. *Cell* **138**, 30–50 (2009).
41. D. J. McGeoch, F. J. Rixon, A. J. Davison, Topics in herpesvirus genomics and evolution. *Virus Res* **117**, 90–104 (2006).
42. D. E. Speiser *et al.*, T cell differentiation in chronic infection and cancer: functional adaptation or exhaustion? *Nat Rev Immunol* **14**, 768–774 (2014).
43. K. E. Pauken *et al.*, Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 blockade. *Science* **354**, 1160–1165 (2016).
44. M. Buggert *et al.*, Limited immune surveillance in lymphoid tissue by cytolytic $CD4^+$ T cells during health and HIV disease. *PLoS Pathog* **14**, e1006973 (2018).
45. M. Buggert *et al.*, Identification and characterization of HIV-specific resident memory $CD8^+$ T cells in human lymphoid tissue. *Sci Immunol* **3**, eaar4526 (2018).
46. D. A. Price *et al.*, Avidity for antigen shapes clonal dominance in $CD8^+$ T cell populations specific for persistent DNA viruses. *J Exp Med* **202**, 1349–1361 (2005).
47. J. Cao *et al.*, The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
48. J. H. Levine *et al.*, Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197 (2015).
49. X. Qiu *et al.*, Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* **14**, 979–982 (2017).