Distinct populations of antigen specific tissue resident CD8 T cells in human cervix mucosa

Tao Peng,1,2 Khamson Phasouk,1 Emily Bossard,1 Alexis Klock,3 Lei Jin,1 Kerry J. Laing,3 Christine Johnston,3 Noel A. Williams,1 Julie L. Czartoski,1 Dana Varon,3 Annalysa N. Long,4 Jason H. Bielas,4 Thomas Snyder,5 Harlan Robins,5 David M. Koelle,1,2,3,6,7 M Juliana McElrath,1,2,3,6 Anna Wald,1,2,3,8 Lawrence Corey,1,2,3* Jia Zhu 1,2* 

1Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA 
2Department of Laboratory Medicine and Pathology, University of Washington, Seattle, Washington, USA 
3Department of Medicine, University of Washington, Seattle, Washington, USA 
4Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA 
5Adaptive Biotechnologies, Seattle, Washington, USA 
6Department of Global Health, University of Washington, Seattle, Washington, USA 
7Benaroya Research Institute, Seattle, Washington, USA 
8Department of Epidemiology, University of Washington, Seattle, Washington, USA 

*: Equal contribution

Correspondence: 
Tao Peng, Ph.D. 
Box 358080, Department of Laboratory Medicine and Pathology 
University of Washington 
Seattle, WA  98195 
E-mail: tpeng@uw.edu; (o) 206-667-5861

Lawrence Corey, MD 
1100 Fairview Ave N 
Mail stop E3-300 
Fred Hutchison Cancer Research Center 
Seattle, WA  98109 
E-mail: lcorey@fredhutch.org; (o) 206-667-6770

Jia Zhu, Ph.D. 
Box 358080, Department of Laboratory Medicine and Pathology 
University of Washington 
Seattle, WA  98195 
E-mail: jiazhu@uw.edu; (o) 206-667-2015

Conflicts of interest: The authors declare no competing financial interests.
Abstract (192 words)

The ectocervix is part of the lower female reproductive tract (FRT), which is susceptible to sexually transmitted infections (STI). Comprehensive knowledge of the phenotypes and T cell receptor (TCR) repertoire of tissue resident memory T cells (TRM) in human FRT is lacking. We have taken single-cell RNA sequencing approaches to simultaneously define gene expression and TCR clonotypes of the human ectocervix. There are significantly more CD8 than CD4 T cells. Unsupervised clustering and trajectory analysis identify distinct populations of CD8 T cells with IFNG$^{hi}$GZMB$^{low}$CD69$^{hi}$CD103$^{low}$ or IFNG$^{low}$GZMB$^{hi}$CD69$^{med}$CD103$^{hi}$ phenotypes. Little overlap was seen between their TCR repertoires. Immunofluorescent staining shows that CD103$^+$ CD8 TRM cells preferentially localize in epithelium while CD69$^+$ CD8 TRM distribute evenly in epithelium and stroma. Ex vivo assays indicate up to 14% of cervical CD8 TRM clonotypes are HSV-2 reactive in HSV-2-seropositive persons, reflecting physiologically relevant localization. Our studies identify subgroups of CD8 TRM in the human ectocervix that exhibit distinct expression of antiviral defense and tissue residency markers, anatomic locations, and TCR repertoires that target anatomically relevant viral antigens. Optimization of the location, number, and function of FRT TRM is an important approach for improving host defenses to STI.
Introduction

The lower female reproductive tract (FRT) includes the vagina and ectocervix, which are the major entry sites in females for sexually transmitted infections (STI) including Herpes Simplex Virus type 2 (HSV-2), Human Immunodeficiency Virus type 1 (HIV-1) and certain Human papillomavirus (HPV) species. In the case of HSV-2, about 16% of US adult women are HSV-2 seropositive (1, 2) and nearly all have occasional FRT HSV-2 shedding (3, 4). It is generally assumed that the FRT mucosal immune system is critical to protect against STI. Successful prophylactic HPV vaccines have been developed (5) yet such vaccines for HSV and HIV are currently unavailable. Since several clinical trials of antibody-based vaccines for HSV-2 have failed to show significant protection (6), it is believed that some combination of CD4, CD8 T cells and B cells are necessary to protect against or control HSV-2 infection (7). Our studies using human genital skin biopsies from subjects with recurrent HSV-2 infection have shown that CD8 T cells at the dermal-epidermal junction (DEJ CD8) have the hallmarks of tissue resident memory (TRM) CD8 T cells and that these DEJ CD8 likely play critical roles in controlling recurrent HSV-2 infection in the genital skin area (8, 9). Using cervical cytobrush and biopsy specimens, several studies have shown that the human cervix possesses HSV-reactive CD4 and CD8 T cells (10-12). Studies using a mouse vaginal infection model with attenuated HSV-2 mutants have shown that both CD4 and CD8 T cells are necessary to control HSV-2 infection (13, 14). HIV-1 specific cytotoxic CD4 and CD8 T cells localize to the cervix in HIV-infected women (15, 16) and FRT CD4 T cells may mediate HIV infection and transmission (17, 18). It is believed that the FRT CD8 CTL responses offer some protection against HIV-1 infection. However, there is a major knowledge gap in understanding T cell responses at steady-state levels in human cervix mucosa.

The past decade has defined the importance of TRM as the most abundant memory T cell subset, and the first line of defense against pathogens (9, 13, 19-21). TRM reside in peripheral tissues with limited recirculation, a feature distinguishing them from recirculating effector memory T cells (Tem).
subsets are usually distinct and heterogeneous in different peripheral tissue types such as small intestine, lung, skin, liver and FRT (22-24). Transmembrane receptors including CD69, S1PR1 and CCR7 have been shown to regulate TRM residency (9, 25-27). Integrin molecules such as CD103 (αEβ7), CD49a (α1β1) and LFA-1 (αLβ2) are known to regulate tissue residency and local immune responses of TRM. CD103 is enriched in CD8 TRM in human and mouse skin epidermis (28, 29), CD49a regulates tissue residency of CD8 TRM in skin and lung tissues (30, 31) while LFA-1 is required for CD8 TRM persistency in liver and lung tissues (32-34). Transcription factors like TCF7 (TCF-1), KLF2, RUNX3, ID3, ZNF683 (Hobit), PRDM1 (Blimp1), EOMES and TBX21 (T-bet) have been shown to regulate tissue residency, memory and effector functions of TRM in mouse models (24, 26, 35-38). However, there is a need to discover reliable markers to define T cell subsets in specific human peripheral tissues.

By flow cytometry, CD69+ and/or CD103+ CD4 and CD8 TRM have been shown to exist in the human cervix and/or other parts of FRT (11, 39-41). To facilitate vaccine development against pathogens such as HIV and HSV-2, there is a major need to define phenotypes, TCR clonotypes and anatomic locations of different T cell subsets in human cervix mucosa. We obtained human ectocervical biopsies and sought to define phenotypes and TCR clonotypes of CD8 T cells in the cervix using single cell immune profiling based on combined gene expression and paired α and β chain TCR sequencing. Our single cell gene expression analysis revealed two distinct populations of CD8 TRM (cytolytic versus non-cytolytic) with differential expression of TRM markers (CD103, CD49a, LFA-1 and CD69). These two subsets of CD8 TRM have distinct anatomic locations and dominant TCR clonotypes in the human cervix. To probe antigenic targets of these CD8 TRM populations among HSV-2 seropositive women, we discovered that up to 14% of the clonotypes from both CD8 TRM subsets in the cervix are shared with HSV-2 reactive CD8 T cell clonotypes in matched blood. These results suggest that different populations of CD8 TRM are providing distinct defenses to viral infection in the human cervix.
Results

Distinct populations of CD8 TRM cells with differential expression of tissue residency markers and cytolytic and non-cytolytic antiviral genes in human cervix mucosa.

To define phenotypes and clonotypes of CD8 T cells in the human cervix, we obtained ten ectocervical biopsies from eight HSV-2 seropositive individuals, sequenced 23,426 cells for 5’ gene expression and obtained TCR-vdj data (V, D and J gene usage, CDR3 sequences and clonotype frequency (clone count)) from 10,102 cells out of the 23,426 cells (Top part, Table 1). We successfully generated single cell gene expression and TCR-vdj data using four different methods to process fresh cervical biopsies for sequence library construction: whole suspension, whole suspension with one-hour plating to remove a significant number of fibroblast cells, CD8 negative selection and CD8 negative selection with prior one-hour plating (details in Methods). We used CD8A+CD3D+, CD4+CD3D+ and SFRP2+ to mark CD8 and CD4 T cells, and fibroblast cells (42), respectively. NK cells were defined as CD8A^low^CD3D+. In both whole suspension and CD8 negative selection, one-hour plating significantly removed fibroblast cells and enriched CD8 and CD4 T cell populations in the single cell libraries (Figure 1). Other cell types such as B-cells, macrophages and vascular endothelial cells were found in whole suspension with or without one-hour plating.

Paired t-tests between the numbers of CD8 T cells (CD8A^+^CD3D^+^) and the numbers of CD4 T cells (CD4^+^CD3D^+^) from the eight cervical biopsies (C1, C2, C5, C6.1, C6.2, C7.1, C7.2, C8.1; specimens described in Table 1) suggest that there are significantly more CD8 T cells in the cervix than CD4 T cells (p-value = 0.003) (Figure 2A), which is consistent with previously published enumeration of CD4 and CD8 T cells by flow cytometry in human cervix tissues (43, 44).

Unsupervised clustering analysis of single cell gene expression data from these eight cervical samples identifies 17 distinct clusters of cells. Analysis of expression of cell type specific markers cross the 17 clusters identifies 12 major cell types [CD8 T cells (2 subtypes), CD4 T cells (2 subtypes), NK cells, B-cells,
macrophages (2 subtypes), dendritic cells, fibroblast cells (3 subtypes), epithelial cells, vascular endothelial cells, lymphatic endothelial cells, vascular smooth muscle cells, erythrocytes] (Left panel, Figure 2B & Figure 2C). Individual cell types from different cervix samples align well within clusters, implying that difference among the cell types is greater than the biopsy difference (Right panel, Figure 2B).

Statistical analysis identified significantly expressed genes in individual clusters. CD8 T cells are the major contributors to the expression of IFN-γ, granzyme B (GZMB) and perforin (PRF1), while CD4 T cells dominate the expression of FOXP3, a lineage transcription factor for CD4+ T regulatory cells (Treg), and CD40LG, a known molecule that regulates interactions between CD4+ T cells and many other cell types including B-cells and DC (45) (Figure 2C and data not shown). NK cells have low levels of CD8A expression, little CD3D and IFN-γ expression, and high levels of GZMB.

Statistical tests of differentially expressed genes (DEG) between the two clusters of CD8 T cells showed one with significantly higher levels of GZMB, GNLY (Granulysin), ITGAE (CD103 α), ITGA1 (CD49a α), IL7R (CD127) and ZNF683 (Hobit) while the other has much higher IFN-γ, TNF (TNF α), CCL3, CCL4, CD69, ITGB2 (LFA-1 β) and EOMES (Figure 2D). We labeled the two clusters of CD8 T cells as TRM GZMB/ITGAE Hi and TRM IFNG/CD69 Hi, respectively. Granzyme B (GZMB) and interferon γ (IFNG) encode important cytolytic and non-cytolytic T cell effector molecules, respectively. ITGAE, ITGA1 and ITGB2 encode the α subunits of CD103 and CD49a and the β subunit of LFA-1, respectively. These three integrin molecules are known to play roles in the establishment of T-cell tissue residency in peripheral tissues. ZNF683 is a transcription factor known to regulate tissue residency of T-cells, IL7R and EOMES play critical roles in generation of memory T-cells and GNLY encodes a cytotoxic molecule (46). Further analysis of expression of both α and β subunits of CD103, CD49a and LFA-1 in TRM IFNG/CD69 Hi and TRM GZMB/ITGAE Hi suggests that CD103 αE (ITGAE) and CD49a α1 (ITGA1) but not LFA-1 αL (ITGAL) are significantly higher in TRM GZMB/ITGAE Hi, while LFA-1 β2 (ITGB2) but not CD103 β7 (ITGB7) or CD49a
β1 (ITGB1) are significantly higher in TRM IFNG/CD69 Hi (Figure 2E). In summary, the data suggest that there are two distinct populations of CD8 TRM in the human cervix with differential expression of tissue residency markers and cytolytic and non-cytolytic antiviral genes.

To further define potentially different populations of CD8 T cells in the cervix, we purified CD8 T cells using CD8 negative immunomagnetic selection. We generated single cell gene expression data for 4,266 cells in which 3,491 cells had detectable CD8A expression in one cervix sample (C4). Unsupervised clustering analysis identified 16 clusters of cells, suggesting that CD8 T cells in this cervical specimen have different subtypes (Top panel, Figure 3A). To define the subtypes of CD8 T cells in the cervix, we chose well-characterized genes in categories of cytolytic (GZMA, GZMB, PRF1 and GNLY), non-cytolytic (IFNG, TNF and CCL5), tissue residency markers (CCR7, S1PR1, CD69, ITGAE, ITGB7, ITGA1, ITGB1, ITGAL, & ITGB2), T-cell survival (IL7R) and transcription factors known to regulate T-cell tissue residency (ZNF683, RUNX3 & EOMES). CD69 is an activation marker for T cells and a TRM marker in peripheral tissues and all the CD8 T cells had moderate to high levels of CD69 expression and cells in cluster 7 had the highest CD69 expression (Bottom panel, Figure 3A). Cells in cluster 7 also had the highest IFN-γ expression, high levels of GZMA yet low levels of GZMB and GNLY, and low levels of CCR7 and S1PR1, two receptors that are known to regulate T-cell exit from tissue (26, 47, 48). They also had high levels of ITGAL and ITGB2 (α & β subunits of LFA-1) yet low levels of ITGAE, ITGA1 and IL7R. The group of CD8 T cells in cluster 7 likely represent a type of TRM for immunosurveillance. We labelled them as IFN-γ high and CD69 high TRM (TRM IFNG/CD69 Hi as described in Figure 2D). In contrast, cells in cluster 1 had the highest GZMB, ITGAE and IL7R yet low levels of IFN-γ expression; they also expressed moderate levels of CCR7 and S1PR1. We labelled them as GZMB high and ITGAE high TRM (TRM GZMB/ITGAE Hi as described in Figure 2D). Consistent with the results described in Figure 2D, ZNF683 (Hobit) and RUNX3 are higher in TRM GZMB/ITGAE while EOMES is higher in TRM IFNG/CD69 Hi. In contrast, CD8 T cells in cluster 0 had little expression of cytolytic granzymes (GZMA, GZMB & PRF1) or IFN-γ yet they have high
levels of expression of \textit{CCR7}, \textit{S1PR1} \& \textit{IL7R} yet little expression of \textit{ITGAE}, \textit{ITGA1} or \textit{ITGAL} (Bottom panel, \textbf{Figure 3A}). CD8 T cells in cluster 0 appear to be a type of cells that have the potential to exit the cervix. Cells in the other clusters had mixed gene expression patterns of these groups of genes, suggesting that CD8 T cells in the human cervical tissues are heterogeneous. To define the relationship of cells in different clusters, we performed trajectory analysis of the entire group of cells. Clustering analysis partitioned cells into three groups and expression patterns of individual genes (\textit{IFNG}, \textit{TNF}, \textit{CD69}, \textit{GZMB}, \textit{PRF1}, \textit{ITAGE}, \textit{CCR7} and \textit{S1PR1}) in partition showed that CD8 TRM cells with high levels of \textit{IFNG}, \textit{TNF}, and \textit{CD69} were grouped in the left end of partition 2 while CD8 TRM with high levels of \textit{GZMB}, \textit{PRF1} and \textit{ITGAE} were grouped in the middle section of partition 1 and those with high levels of \textit{CCR7} and \textit{S1PR1} in the right end of partition 1 (Top panel, \textbf{Figure 3B}). By setting root nodes at the right ends of partition 1 and 2, two trajectory curves placed cells in partition 1 & 2 in pseudo time, a measure of how CD8 T cells in the human cervix relate to one another biologically (Bottom panel, \textbf{Figure 3B}). Detailed trajectory analysis of purified CD8+ T cells from the cervix suggests that TRM IFNG/CD69 Hi and TRM GZMB/ITAGE Hi are distinct populations of TRM in the human ectocervix.

\textbf{Distinct anatomic location of different populations of CD8 TRM in human cervix mucosa.}

To verify \textit{in vivo} distribution of two CD8 TRM subsets in the human cervix described in \textbf{Figures 2 \& 3}, and to confirm that these cell populations were also found in HSV-2 seronegative women, we obtained a separate set of eight ectocervix biopsies from seven subjects plus saved portions of four cervix biopsies which were used for single cell immune profiling (Bottom part, \textbf{Table 1}). Five cervix biopsies (C1, C5, C7.2, C8.1 \& C8.2) were from HSV-2 seropositive individuals while seven biopsies (C9 to C15) were from HSV-2 seronegative individuals. We performed double immunofluorescent staining using CD8A and CD103 antibodies, or CD8A and CD69 antibodies to analyze anatomic locations of CD8A$^+$CD103$^+$ T cells or CD8A$^+$CD69$^+$ T cells in these cervix samples. Most CD8A$^+$CD103$^+$ T cells locate in epithelium while CD8A$^+$CD69$^+$ T cells distribute evenly from epithelium to underlying stroma area (\textbf{Figure 4A}, and
Supplementary Figures 1 and 2). We did the double staining with eleven of the twelve cervix samples and counted the number of CD8A+ , CD8A*CD103+ and CD8A*CD69+ T-cells in epithelium and stroma. On average for the five HSV-2 seropositive samples, 62% and 12% of CD8A+ T cells are CD103+ in epithelium and underlying stroma, respectively and this difference is statistically significant (p-value = 0.0001, paired t-test, n = 5) (Top row, Figure 4B). For the six HSV-2 seronegative samples, the difference is also significant (93% versus 49%, p-value = 0.0029, paired t-test, n = 6) (Bottom row, Figure 4B). In contrast, similar percentages of CD8A+ T cells are CD69+ in epithelium and stroma (75% versus 77% for HSV-2 seropositive and 85% versus 87% for HSV-2 seronegative). Densities of CD8A+ T cells are similar in both areas (Figure 4B). Combination of RNA FISH (fluorescence in situ hybridization) for GZMB or IFNG and IHC (immunohistochemistry) for CD103 or CD69 demonstrated GZMB+CD103+ cells in epithelium and IFNG+CD69+ cells in underlying stroma area of the cervix biopsy tissues (n = 3) (Figure 4C). The distinct anatomic locations of CD8A*CD103*GZMB+ and CD8A*CD69*IFNG+ T cells in the human cervix suggest that the two CD8 TRM subsets (IFN-γhiCD69hiCD103low versus GZMBhiCD103hiCD69med) revealed by the single cell gene expression analysis may have different anatomic locations in the human cervix.

The two CD8 TRM subsets have distinct expanded clonotypes in human cervix mucosa.

We then analyzed the distribution of TCR clonotypes in the two clusters of CD8 TRM (IFNG/CD69 Hi and TRM GZMB/ITAGE Hi) from three cervical samples (C4, C2 and C5). Both types of TRM are oligoclonal, however, the TRM GZMB/ITAGE Hi subset had higher levels of clonal expansion (Left panel, Figure 5). For C4, 56 clonotypes out of the 365 unique clonotypes from TRM GZMB/ITAGE Hi were detected in more than one cells (expanded) while 19 of the 142 clonotypes in the TRM IFNG/CD69 Hi subset were expanded. There was no overlap between these expanded clonotypes (Top row, middle panel, Figure 5). Only 5 clonotypes were shared between TRM GZMB/ITAGE Hi and TRM IFNG/CD69 Hi (Top row, right panel, Figure 5). For C2, 26 clonotypes out of the 117 unique clonotypes from TRM GZMB/ITAGE Hi, and 17 clonotypes out of the 154 unique clonotypes from TRM IFNG/CD69 Hi were expanded and these two
groups of expanded clonotypes do not have any overlap either (Middle row, middle panel, Figure 5).

Only 7 singletons from the 154 unique clonotypes from TRM IFNG/CD69 Hi were shared with the 117 unique clonotypes from TRM GZMB/ITAGE Hi (Middle row, right panel, Figure 5). For C5, 4 out of the 32 expanded clonotypes plus 1 singleton from TRM GZMB/ITAGE Hi were shared with 5 singletons in TRM IFNG/CD69 Hi (Bottom row, Figure 5). The distinct expanded clonotypes of TRM IFNG/CD69 Hi and TRM GZMB/ITAGE Hi in the three cervix samples suggest that the two subsets of TRM described in Figures 2 & 3 have different roles in host defense and separate prior histories responsible for their placement.

**Expanded clonotypes are largely from CD8 T cells and clonally expanded T cells exhibit high levels of gene expression for granzymes and chemokines in human cervix mucosa.**

We selected CD8A+ and CD4+ cells and identified their clonotypes in the four cervical samples (C2, C5, C1 and C6.1. In all except C6.1, CD8 T cells had more diverse TCR repertoires than did CD4 T cells. In all the four cervical samples, CD8 T cells had higher percentages of expanded clonotypes (clonotype frequency >= 2) (Figure 6A). The results suggest that expanded clonotypes of T cells in the human cervix are mostly from CD8 T cells.

Since CD8 T cells in the cervix appear to be the major contributors to cytolytic and non-cytolytic gene expression and expanded TCR clonotypes in the human cervix, we evaluated gene expression differences between cells with expanded clonotypes (clonotype frequency ≥2) and those with singletons (clonotype frequency = 1). In specimen C4, we obtained gene expression data from > 4000 cells and TCR-vdj data from > 3000 cells. Comparison of gene expression between 1531 cells with clonotype frequency with ≥ 2 and 1780 cells with singletons showed that GZMH, GZMB, GZMA, NKG7 (A gene encoding a granule protein) and CCL4 were significantly higher in cells with expanded clonotypes. Comparison of gene expression between 462 cells with expanded clonotypes and 220 cells with singletons from C3 showed that the same set of genes except NKG7 (GZMH, GZMB, GZMA and CCL4) were significantly
higher in cells with expanded clonotypes (Left two panels, Figure 6B). We found similar results in two additional cervix samples (Right two panels, Figure 6B). These results suggest that clonally expanded T cells possess higher levels of gene expression for cytotoxicity and chemokines than do TCR singleton T cells.

Both CD8 TRM subsets in the human cervix target HSV-2 antigens.

The cervix is a site of HSV-2 infection and chronic shedding. Previous work has recovered HSV-2-specific T cells from this site (10, 11). To define if the T cell clonotypes associated with two distinct populations of CD8 TRM (TRM GZMB/ITAGE Hi and TRM IFNG/CD69 Hi as described in Figures 2-5) in the human cervix were reactive with a local infectious pathogen, we identified HSV-2 reactive TCR β (TCRB) sequences in the blood and sought TCR overlap with cervical T cells. HSV-2-reactive CD8 T cells were purified from blood using expression of CD137 as an activation induced marker (AIM) after cross-presentation of HSV-2 by autologous moDC (49). Genomic DNA prepared from sorted CD137\textsuperscript{hi} CD8 T cells was used for bulk TCRB CDR3 sequencing in a quantitative platform in which productive TCRB CDR3 reads correspond to individual cells (50). We compared such blood TCRB CDR3 sequences from three HSV-2 seropositive subjects (Pt4, Pt5 and Pt7) to the clonotypes associated with TRM GZMB/ITAGE Hi and TRM IFNG/CD69 Hi in matching cervical specimens (C4, C5 and C7.1). A matched TCRB clonotype is defined as matched TCRB CDR3 amino acid sequences and VDJ gene usage. We found that both subsets of CD8 TRM had high levels of HSV-2-specific TCRs. The levels were 2.75% and 10.20%, 6.35% and 3.64%, and 14.21% and 12.35% HSV-2 reactive clonotypes from C4, C5 and C7.1 cervix biopsy, respectively, for TRM GZMB/ITAGE Hi and TRM IFNG/CD69 Hi (Table 2). These abundances are far higher than the abundance of HSV specific CD8 T cells in peripheral blood (49, 51). The results suggest that both populations of CD8 TRM in the human cervix are enriched for T cells reactive with a chronic viral pathogen that provides local intermittent antigen exposure.
Discussion

The human ectocervix is a gateway to upper FRT and it is susceptible to several sexually transmitted viral infections including three well-studied viruses: HSV, HIV and HPV. Other viral infection such as CMV and EBV in the human cervix has been described (52, 53). Defining the complexities of the host immune response in this mucosal tissue is important in increasing our understanding how the host can contain exogenous pathogens. We sought to define CD8 T cell functional subsets and their TCR clonotypes in human cervix utilizing ectocervical biopsies and unbiased single cell immune profiling strategies. Our data uncovered two major subsets of CD8 TRM in the human cervix mucosa with distinct phenotypes: one with IFNGhiGZMBlowCD69hiCD103loLFA-1hi versus the other with IFNGlowGZMBhiCD69medCD103hiLFA-1low. Interestingly, TCR clonotypes of the two populations have little overlap. Immunofluorescent staining reveals that CD103+ CD8 TRM cells preferentially localize in epithelium while CD69+ CD8 TRM distribute evenly in epithelium and underlying stroma. We believe that this is the first report on CD8 TRM subsets with differential expression of cytolytic and non-cytolytic genes and tissue resident markers; as well as the association between these CD8 TRM subsets with distinct expanded clonotypes and anatomic locations. By comparing to HSV-2 reactive clonotypes from matching blood, we demonstrated that both CD8 TRM populations in the human cervix may contain HSV-2 specific T cells.

CD103 and CD49a have been shown to regulate tissue residency of CD8 TRM in human and mouse skin epidermis (30, 31), LFA-1 regulates TRM tissue residency in liver and lung tissues (32-34) and CD103+CD69+ T-cells have been shown to reside in human cervix tissues (11, 41, 44). We showed here that IFNGhiCD69hi CD8 TRM express high levels of LFA-1 but low levels of CD103 and CD49a. In contrast, GZMBhiCD103hi CD8 TRM express high levels of CD103 and CD49a but low levels of LFA-1, and they preferentially locate in epithelium. Although our studies suggest that both populations of CD8 TRM likely target HSV-2 antigens, it is unclear how the two distinct subsets of CD8 TRM perform immune surveillance against diverse pathogens in the human cervix. Our immunofluorescent staining analysis
showed that the two populations of CD8 TRM had distinct anatomic locations in the cervical specimens from both HSV-2 seropositive and seronegative female subjects. However, our single cell immune profiling analysis was performed in the cervical samples only from HSV-2 seropositive individuals as we do not have fresh tissues required for single cell immune profiling in HSV-2 seronegative persons.

TRM GZMB/ITGAE Hi express high levels of CXCR3, a chemokine receptor for CXCL9 and CXCL10. Both chemokines are highly inducible in epithelial cells by IFN-γ (54, 55). We hypothesize that IFN-γ high TRM cells produce IFN-γ, which induces expression of CXCL9/CXCL10 in epithelial cells, which in turn induce chemotaxis of GZMB/ITGAE high TRM cells to the epithelial surface of the cervix. The importance of IFN-γ in TRM has been nicely demonstrated in mouse models for antiviral regulation and chemotaxis of other immune cells in peripheral tissues (56, 57). Assuming some of IFN-γ high TRM cells in the human cervix among HSV-2 seropositive individuals are specific for HSV-2 antigens, it would be of interest to define HSV-2 antigens targeted by these IFN-γ high TRM cells, as they would be potential vaccine candidates for recurrent HSV-2 infection. In fact, a prime-pull strategy using CXCL9/CXCL10 has been nicely demonstrated to increase TRM in mouse vagina to control HSV infection (58).

Human TRM subsets have been demonstrated by dye efflux (22). Two recent publications have used single cell gene expression to elucidate the subsets of small intestine intraepithelial lymphocytes (siIEL) in mouse models (23, 24). ID3\textsuperscript{hi}Blimp1\textsuperscript{lo} siIEL are suggested to be stem-like TRM. Cells in cluster 0 shown in Figure 3A have high levels of TCF7, ID3 and ZNF683 (Hobit) and low levels of PRDM1 (Blimp1) (data not shown) so we hypothesize that they may be stem-like TRM cells. Their entire TCR repertoire (663 singletons) do not have overlap with those in cluster 7 (TRM IFNG/CD69 Hi) and only two of the singletons are found in cluster 1 (TRM GZMB/ITGAE Hi) (singletons as well) (data not shown). It is unclear what are the potential functions of stem-like TRM in the human cervix mucosa.
Single cell gene expression analysis of purified CD8 T cells clearly suggests that CD8 T cells in the human cervix have a highly heterogeneous population in terms of their gene expression and clonotypes. A recent report showed that TRM cells in mouse intestine could recirculate with predilection to re-establish tissue residency in original tissues upon local reinfection (59). This study supports our observation that CD8 TRM in the human cervix may be highly dynamic in terms of their development, differentiation, and antigenic specificities.

The expanded TCR clonotypes in the human cervix are mostly from CD8 T cells. Memory T cells proliferate not only in secondary lymphoid organs (SLO) but also on-site in peripheral tissues (60-62). How much SLO and local proliferation contribute to clonal expansion of CD8 T cells in human cervix is currently unknown. Our previous studies have shown the distinct anatomic distribution of CD8 and CD4 T cells in human genital skin during recurrent HSV-2 infection: CD8 T cells are closer to the dermal-epidermal junction where HSV-2 emerges from nerve endings to infect basal keratinocytes than CD4 T cells, which are found deeper in the dermal area (9, 63). This observation leads us to hypothesize that on-site proliferation likely contributes to the dominant clonotypes of CD8 T cells in the human cervix.

Could we leverage CD8 TRM subsets to make therapeutic vaccines against genital herpes infection? TRM cells have been shown to contribute to vaccine effects against malaria in humans and tuberculosis in Rhesus macaques (64, 65). Cervical biopsies may be a useful tool to define whether such goals can be achieved. This report provides strong evidence to suggest that CD8 TRM subsets are critical local immune components to perform on-site immune-surveillance against infection such as recurrent HSV-2 infection. It is therefore rational to leverage different CD8 TRM subsets as vaccine targets to combat genital herpes infection and other chronic infections in the FRT.
Methods

Human cervix biopsy acquisition. For 10x single cell immune profiling, female adults (P1 to P8 in Table 1) were recruited at the University of Washington Virology Research Clinic in Seattle, WA. The biopsy protocol was approved by University of Washington IRB committee (STUDY ID: STUDY00004709), and all participants provided written consent. All participants were HSV-2 seropositive and HIV seronegative. Biopsy procedures were conducted as described previously (11). Biopsy procedures were performed between menstrual cycles to prevent menstrual blood in cervical samples. Biopsies (ectocervix tissue) were obtained using a Baby Tischler Biopsy Forceps (Wallach Surgical, Trumbull, CT, USA) with a 4.2 x 2.3 mm bite size. We used lidocaine if more than one biopsy were obtained or if subjects requested it. It would be 0.5 ml of 1 % lidocaine injected into the cervical site prior to obtaining biopsy. The biopsy was placed in 5 ml of RPMI-CVX (RPMI 1640 medium supplemented with 5 mM L-glutamine, 50 U/ml penicillin, 50 µg/ml streptomycin, and 10% human serum, 10 µl/ml amphotericin B, and 0.5 µl/ml ciprofloxacin). All biopsies were transported to the laboratory on wet ice and processed within four hours of collection.

Five cervix biopsies from four HSV-2 seropositive subjects were tested by highly sensitive PCR methods to determine whether they had HSV-2 DNA (66, 67). Three cervix biopsies were negative and two had very low copies of HSV-2 DNA (Supplementary Table 1).

Cervix biopsies from P9 to P15 in Table 1 were similarly obtained in Seattle Vaccine Trials Unit at Fred Hutch. All participants were HSV-2 and HIV seronegative, and tests for chlamydia trachomatis, Neisseria gonorrhoeae and trichomonas vaginalis were also negative. Informed consent was obtained from all participants and was approved by Fred Hutchinson Cancer Research Center institutional review board (IR5640). Two ectocervical biopsies were obtained using a Baby Tischler Biopsy Forceps (Wallach Surgical, Trumbull, CT, USA) without lidocaine. All biopsies were placed in 5 ml of RPMI-CVX, transported to the laboratory on wet ice and processed within four hours of collection.
Generation of single cell suspension from human cervix biopsies. Cervical biopsies were sliced into 3 to 5 strips with a sterile scalpel (Size 10). The cervical strips were transferred into a sterile Falcon™ tube containing 3 ml of freshly prepared collagenase solution (1 mg/ml) prewarmed to 37°C (Collagenase from Clostridium histolyticum, Sigma-Aldrich, Cat # C6885-500mg), and 3 ul of DNase (1U/ul, Sigma-Aldrich, Cat # DN25-10mg). Afterwards, the 8 ml Falcon™ tube was placed in a plastic bag with a zip lock, and then placed in a 200 RPM shaker at 37°C for 30 min. After digestion, the biopsies were subjected to mechanical agitation by pipetting up and down 10-15 times using a 16-gauge blunt-end needle attached to a 3 ml syringe. Once agitation is complete, the solution was dispensed through a 70-micron cell strainer placed on a sterile 50 ml conical tube. R15 RPMI media was used to further wash the 8 ml Falcon™ tube to collect any residual cells and dispensed the wash through the 70-micron cell strainer. The R15 RPMI media wash was repeated 3-4 times until the final volume in the 50 ml conical tube reached 20 ml. The 50 ml conical tube was then centrifuged for 10 min at 250 g. After centrifuging, media was removed with a serological pipet, cells were gently resuspended with 500 ul of ACK Lysing Buffer (Gibco™ A10492-01) and incubated at RT for 3 min. After RBC lysis, two rounds of washing with 20 ml of 5% FBS RPMI media to neutralize ACK Lysing Buffer and centrifuging for 10 min at 250 g was performed. After the second wash, the cell pellet was resuspended in 3 ml of R15 RPMI media, dispensed in one well of a 6-well culture plate (Falcon® 353046), and incubated for 1 hour in a 37°C incubator with 5% CO2 to allow for fibroblasts to settle and adhere. After 1 hour, cells that remain suspended in R15 RPMI media were collected and subsequently centrifuged for 10 min at 250 g and washed in 1x PBS. After two washes in 1x PBS, cells were counted via a hemocytometer and diluted to a concentration of 1000 cells/ul for 10x Genomics processing. For CD8 negative selection, cell was lysed in ACK Lysing Buffer, neutralized and washed two times with 5% FBS RPMI media, and with or without one-hour plating as described above, cells were further purified following the protocol from EasySep™ Human CD8 T Cell Isolation Kit (STEMCELL, Cat # 17953). Once CD8 T Cells were isolated, they were
washed two times with 1x PBS, counted via a hemocytometer, and diluted to 1000 cells/μl for 10x Genomics.

10x 5’ gene expression and TCR-vdj library construction. Single cells from cervix biopsies were loaded onto the 10x Genomics Chromium Single Cell Chip. A target of 17,000 cells were loaded to account for ~60% cell capture rates. Cells were loaded according to 10x Genomics Manufacturer’s instructions for Chip A to generate Gel Bead-In-Emulsion (GEMS, 10x genomics) using barcoded gel beads. RNA containing GEMS underwent barcoded cDNA synthesis. Subsequent library construction followed 10x Genomics 5’ Single Cell Immune Profiling Manufacturer’s Instructions. Libraries were constructed using the Chromium Next GEM Single Cell 5’ Library & Gel Beads Kit v1.1 with additional Chromium Single Cell V(D)J Enrichment Kit, Human T Cells kit for TCR Libraries. Samples were quantified using the KAPA qPCR Quantification Kit and pooled equimolar for Illumina Sequencing. All libraries were pooled together with a 4:1 ratio between Gene Expression and TCR libraries.

Single-cell RNA-seq data processing and analysis. The sequencing of 10x libraries for 5’ GEX and TCR-vdj for the ten cervix biopsies from eight individuals and cell ranger analysis were performed at the shared resource at Fred Hutch. Cell ranger 3.0.2 with human genome GRCh38 for gene expression and GRCh38-alts-ensembl for TCR-vdj were used in the initial steps including quality control, alignment to human genomes and counting of aligned sequence reads for individual genes. We manually examined summary HTML files for the ten gene expression libraries and the ten TCR-vdj libraries generated from ten cervix samples. All the libraries have valid barcodes (> 85%), fraction reads in cells (> 80%) and reads mapped to genome (> 50%). The summary of number of cells sequenced and number of cells with TCR-vdj data for the ten cervix samples is described in Table 1. The initial analysis was performed in 10x cell Loupe browser and vdj Loupe browser. The MTX files from the output of cell ranger were imported into Monocle3 (68), Seurat (version 3.1.2) (69) or Scanpy (version 1.4.4) (70) for further analysis. In all the three methods, we filtered out cells with less than 200 genes detected and genes that are only detected
in less than 3 cells. We further filtered out cells with percentages of mitochondria genes more than 15%.

pyVDJ (https://github.com/veghp/pyVDJ) and Scirpy (71) is used to link TCR-vdj to gene expression in Scanpy. To perform unsupervised clustering analysis of single cell gene expression data, we use UMAP (Uniform Manifold Approximation and Projection) or tSNE (t-distributed Stochastic Neighbor Embedding) as dimension reduction methods and Leiden (an algorithm for methods of community detection in large networks) as the default clustering method. Trajectory analysis described in Figure 3 was performed in Monocle3. For the analysis described in Figure 2, batch correction among the eight cervical samples from the six subjects using BBKNN methods in Scanpy were performed before clustering analysis (72).

**Immunofluorescence staining and image analysis**

The immunofluorescent staining was performed as previously described (9, 73). Cervical biopsies were cut in half and embedded in Tissue-Tek O.C.T. (Sakura Finetek, Thermo Fisher), and immediately frozen in 2-methylbutane on dry ice and stored at -80°C. Frozen tissue blocks were cut into 8 µm sections and stored at -80°C. Slides were thawed and fixed in acetone for 15 minutes at -20°C and then left to dry for 30 minutes at room temperature. Slides were then quenched in 3% H2O2 for 60 min at RT then in blocking buffer (PBS containing 5% normal human and goat sera) for 1 hour followed by an incubation of primary antibodies (anti-CD69 or anti-CD103). Anti-CD69 (eBioscience, Cat# 14-0699-82) was done followed with the TSB goat anti-mouse protocol (ThermoFisher, Cat# B40912) and CD8-AF647 (BD Biosciences, 557708) was stained ON at 4°C in block solution. Anti-CD103 (Abcam, Cat # 224202) was done followed with the TSB goat anti-rabbit 555 protocol (ThermoFisher, Cat# B40923) and CD8-AF647 was stained ON at 4°C in block solution. Cell nuclei were counter-stained with DAPI (ThermoFisher, Cat# D3571). Slides were mounted with Prolong Gold (Invitrogen Cat# P36930) and cured at least 24 hours prior to imaging.
Images were analyzed in Fiji, an open source image processing package based on ImageJ. Area of epithelium and stroma are analyzed separately. The image is then split into DAPI + CD69 and DAPI + CD8. Each cell type is counted in the epithelium and stroma separately using Cell Counter. An all color combined image is counted for double positive cells in each area. The same analysis was done for CD103 + CD8 images.

**Combination of fluorescent in situ hybridization (FISH) and immunohistochemical staining (IHC)**

Fresh frozen cervix biopsies were cryo-sectioned into 10-mm slides and fixed with 4% paraformaldehyde and dehydrated in ethanol. Dehydrated slides were pretreated by hydrogen peroxide and 0.3% Triton before FISH probe hybridization using RNAscope® Multiplex Fluorescent Reagent Kit v2 (Advanced Cell Diagnostics, cat# 323100). FISH Probes include IFNG and GZMB (Advanced Cell Diagnostics, cat# 3105 and 31051_C2 for IFNG and cat# 445971 for GZMB). Probe hybridized slides were washed by washing buffer (Advanced Cell Diagnostics, cat# 323100) before IHC using AF488 tyramide SuperBoost Kit (Invitrogen, cat# B40912). Slides were incubated with block buffer for 60 minutes at room temperature (RT) and then with primary antibodies for 60 minutes at RT. Primary antibodies include CD69 (eBioscience, cat# 14-0669-82) and CD103 (Abcam, cat# ab224202). After washing with PBS with 0.1% Tween-20, the slides were incubated with poly-HRP-conjugated secondary antibodies for 60 minutes at RT. Slides were washed with 1x PBS with 0.1% Tween-20 and incubated with tyramide for 5 minutes at RT. Slides were counterstained with DAPI (ThemoFisher Scientific, cat# D3571) and mounted in Prolong Gold Antifade Mountant (Thermo Fisher Scientific, cat# P36930).

**Definition of blood CD8 TCR clonotypes reactive with HSV-2.**

HSV-2 specific CD8 T cells were identified and isolated using a method similar to that previously reported for HSV-1 (49). In brief, PBMC were cryopreserved from venous anticoagulated blood using standard methods. 2 x 10^5 monocyte-derived dendritic cells (moDC)—generated from adherent PBMC in
the presence of IL-4 and GM-CSF as described (74)—were incubated for 4 h at 37°C/5% CO₂ with 2 x 10⁵ UV-irradiated (Stratalinker XL1000, 180,000 microjoules) HeLa cells that had either been infected overnight with HSV-2 strain 186 (MOI 2.5) or left uninfected (mock). Autologous CD8 T cells were negatively selected (STEMCELL EasySep™ Human CD8 T Cell Isolation Kit) and incubated at 1-2 x 10⁶ T cells per well with the moDC/HeLa mix. PHA-P (1.6 µg/mL, Remel) was used as a positive control. T cells were harvested after 18-20 h incubation and stained with anti-CD3-ECD (UCHT1, Beckman Coulter), anti-CD8-FITC (3B5, ThermoFisher), anti-CD137-APC (4B4-1, BD), and 7-AAD (BD). Live (7-AAD negative), single CD3+CD8+ lymphocytes that expressed CD137 were sorted using a BD FACS Aria II (UW Cell Analysis Facility), pelleted, and stored at -80°C. DNA was extracted from cell pellets using the DNeasy Blood and Tissue Kit (Qiagen) following the manufacturer’s instructions, which was submitted to Adaptive Biotechnologies for TCRB sequencing.

**Data deposit.** The raw data for single cell gene expression and TCR clonotypes were deposited in NCBI GEO (GSE173231). The TCRB data for blood CD137⁺CD8⁺ cells described in Table 2 were available here:

https://clients.adaptivebiotech.com/pub/peng-2021-jcii

**Statistics.** Two sample T tests (paired or unpaired) were performed to derive the p-values described in Figures 2, 4 and 6.

**Study approval.** Written informed consent was received from participants prior to inclusion in the study. As described in Table 1, the cervix specimen (C1 to C8) were obtained at University of Washington Virology Research Clinic with IRB approval (STUDY ID: STUDY00004709). The cervix specimen (C9 to C15) were obtained at Seattle Vaccine Trials Unit at Fred Hutch with IRB approval (IR5640).
Acknowledgements: We thank M. Miner for editing. We also thank our study participants. This work was supported by grants from the National Institutes of Health (R01 AI134878 & R01 AI111780). This work was also supported by National Institutes of Health Contract 75N93019C00063 (DMK). The authors declare no competing financial interests.

Author Contributions: LC, JZ and TP conceived the study. TP did the analysis and wrote the paper. KP did the tissue digestion, and EB and ANL from JB’s laboratory did the 10x library construction. AK and LJ did immunofluorescent staining. CJ, AW, and DV performed the cervix biopsy procedures at University of Washington. NAW and JC from MJM’s group prepared the cervix biopsy at Fred Hutch. DK and KL isolated HSV-2-reactive CD8 T cells from PBMC. LC, JZ, DK and KL contributed to writing the paper.
References


47. Bromley SK, Thomas SY, and Luster AD. Chemokine receptor CCR7 guides T cell exit from peripheral tissues and entry into afferent lymphatics. *Nat Immunol.* 2005;6(9):895-901.


Figure 1. Plating whole single cell suspension for one hour before single cell RNA-seq library construction significantly removed fibroblast cells and enriched CD8 and CD4 T cells in the human cervix. CD8A, CD4, and SFRP2 are markers for CD8 and CD4 T cells, and fibroblast cells, respectively. CD3D is a marker for both CD8 and CD4 T cells. Comparison of CD8A, CD4, CD3D and SFRP2 gene expression in UMAP between C1 (whole suspension) and C2 (whole suspension with one-hour plating to deplete adherent cells), and between C3 (CD8 negative selection) and C4 (CD8 negative selection with one-hour plating) is marked by line segments with arrows at both ends.
Figure 2. Two subsets of CD8 TRM with differential expression of tissue residency markers and cytolytic and non-cytolytic genes in the human cervix. Eight cervix samples from the six subjects described in Table 1 were used: P1, P2, P5, P6 (2 samples), P7 (2 samples) & P8. A. Numbers of CD8 and CD4 T cells in the human cervix. CD8+ or CD4+ and CD3D+ cells were marked as CD8 and CD4 T cells, respectively. B. Unsupervised clustering analysis of single cell gene expression data from the 8 cervical samples identified 17 clusters of cells (Left panel). The names of individual clusters were based on the expression of cell type specific markers. Display of contribution of the eight cervical samples to individual clusters of cells (Right panel). C. UMAP to display individual gene expression of CD8A, CD4, CD3D, IFNG, GZMB, FOXP3, CD79A, HLA-DRA, SFRP2, KRT5, PECAM1, PROX1, TAGLN & HBB in individual clusters of cells from the eight cervical samples. CD8A, CD4, CD3D, IFNG, GZMB are used to mark T cells and their cytolytic and non-cytolytic gene expression; FOXP3, CD79A, HLA-DRA, SFRP2, KRT5, PECAM1, PROX1, TAGLN & HBB are used to mark regulatory T cells (Treg), B-cells, Macrophages/Monocytes/Dendritic cells, fibroblast cells, epithelial cells, vascular endothelial cells, lymphatic endothelial cells, fibroblast cells, epithelial cells, vascular smooth muscle cells and erythrocytes, respectively. D. UMAP to display gene expression of CD8A, GZMB, GNLY (granulysin), ITGA1 (CD103 a subunit), ITGA1 (CD49a subunit), IL7R, ZNF683 (Hobit), IFNG, TNF, CCL3, CCL4, CD69, ITGB2 (LFA-1 a subunit) & EOMES. Except CD8A, all the other genes are significantly differentially expressed between the two clusters of CD8 TRM (t-test p values < 10^-5). The IFNG/CD69 high cluster of cells were labeled as TRM IFNG/CD69 Hi and the GZMB/ITGA1 high cluster of cells were labeled as TRM GZMB/ITGA1 Hi. E. Expression of ITGA1 & ITGB7 (CD103), ITGA1 & ITGB1 (CD49a) and ITGA1 & ITGB2 (LFA-1) in TRM IFNG/CD69 Hi and TRM GZMB/ITGA1 Hi. * Three genes (ITGA1e, ITGA1a and ITGB2) are significantly differentially expressed between the two subsets of CD8 TRM.
Figure 3. Trajectory analysis of purified CD8 T cells in the human cervix. A. Clustering analysis of single cells. Top panel: expression of CD8A in different clusters; clusters 1 & 7 are marked and labeled as TRM GZMB/ITGAE Hi and TRM IFNG/CD69 Hi, respectively. Bottom panel: heatmap to show gene expression of functional categories (CD8, cytolytic, non-cytolytic, TRM marker, Survival and TRM transcription factor) in different clusters. B. Trajectory analysis of single cells. Top panel: gene expression of individual cells in partition. Genes for IFNG, TNF, CD69, GZMB, PRF1, ITGAE, CCR7, S1PR1 and CD8A are shown. Bottom panel: trajectory analysis in pseudo time. Two trajectory curves were generated with two root nodes marked in partition 1 and 2. The results shown in A & B were generated from CD8 T cells which were negatively selected from the cervical sample C4.
Figure 4. Distinct anatomic location of CD103+GZMB+CD8A+ and CD69+IFNG+CD8A+ T-cells in the human cervix. A. Double immunofluorescent staining of cervix biopsies (C5 and C9) with CD8A and CD103 antibodies, or CD8A and CD69 antibodies. B. Quantitation of CD103+CD8A+, CD69+CD8A+ and density of CD8A+ cells in epithelium and stroma from individual cervix biopsies (top row, n = 5, C1, C5, C7.2, C8.1 and C8.2 from HSV-2 seropositive subjects; bottom row: n = 6, C9 to C15 from HSV-2 seronegative subjects). C. CD103+GZMB+ cells in epithelium and CD69+IFNG+ cells in stroma of cervix biopsies (n = 3). Top panels: RNA FISH for GZMB and immunofluorescent staining for CD103; Bottom panels: RNA FISH for IFNG and immunofluorescent staining for CD69. Scale bar = 50 µm.
Figure 5. Single cell TCR clonotype analysis of the human cervix. Top, middle, and bottom rows are from C4, C2 & C5 cervical samples, respectively. Left: display of vdj clone count and expression of five genes (CD8A, IFNG, CD69, GZMB & ITGAE) in the two clusters of cells representing TRM IFNG/CD69 Hi and TRM GZMB/ITGAE Hi in UMAP. VDJ clone count refers to the number of cells that share the same TCR clonotype. Middle: graphs to show TCR clonotype size (number of cells sharing the same clonotype) and TCR clonotype overlaps between TRM GZMB/ITGAE Hi and TRM IFNG/CD69 Hi. Individual circles represent individual clonotypes that are shared between the two subsets of CD8 TRM. Half circles represent individual clonotypes that are unique to TRM GZMB/ITGAE Hi (X-axis) or TRM IFNG/CD69 Hi (Y-axis). Right: Venn diagrams show clonotype overlaps between TRM GZMB/ITGAE Hi and TRM IFNG/CD69 Hi. The two numbers in the middle of circles in the Venn diagrams refer to numbers of unique clonotypes associated with TRM GZMB/ITGAE Hi and TRM IFNG/CD69 Hi, respectively. The numbers in the middle of the Venn diagrams refer to numbers of shared TCR clonotypes between TRM GZMB/ITGAE Hi and TRM IFNG/CD69 Hi.
Figure 6. The expanded clonotypes in the human cervix are weighted towards CD8 T cells. **A.** Frequency of individual clonotypes in CD8A⁺ and CD4⁺ T cells from the four cervix samples (top row: C2 and C5; bottom row: C1 and C6.1). X-axis: individual clonotypes. Y-axis: cell barcode frequency for individual clonotypes. The tables in individual graphs show numbers of cells which are CD8A⁺ or CD4⁺, numbers of unique TCR clonotypes and percentages of clonotypes whose frequency are >= 2. **B.** Differential gene expression between T-cells with clonotype frequency >= 2 and those with singletons in the four cervical samples (C4, C3, C6.1 & C1). Top graphs: distribution of clonotypes with frequency >= 2 and those with singletons in tSNE plots. Bottom tables: lists of genes that are differentially expressed between T-cells with clonotype frequency >= 2 and those with singletons (P-value <= 0.05).
Table 1

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>Biopsy</th>
<th>Biopsy code</th>
<th>Tissue processing method</th>
<th>Methods</th>
<th># of cells sequenced (23426 total)</th>
<th># of cells with TCR (10102 total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>54.6</td>
<td>Cervix</td>
<td>C1</td>
<td>Whole lysate</td>
<td>single cell sequencing</td>
<td>1880</td>
<td>460</td>
</tr>
<tr>
<td>P2</td>
<td>55</td>
<td>Cervix</td>
<td>C2</td>
<td>Whole lysate/fibroblast cell reduction</td>
<td>single cell sequencing</td>
<td>1775</td>
<td>946</td>
</tr>
<tr>
<td>P3</td>
<td>35.7</td>
<td>Cervix</td>
<td>C3</td>
<td>CD8 negative selection</td>
<td>single cell sequencing</td>
<td>3211</td>
<td>742</td>
</tr>
<tr>
<td>P4</td>
<td>25.8</td>
<td>Cervix</td>
<td>C4</td>
<td>Fibroblast cell reduction/CD8 negative selection</td>
<td>single cell sequencing</td>
<td>4266</td>
<td>3343</td>
</tr>
<tr>
<td>P5</td>
<td>55.7</td>
<td>Cervix</td>
<td>C5</td>
<td>Whole lysate/fibroblast cell reduction</td>
<td>single cell sequencing</td>
<td>3169</td>
<td>1404</td>
</tr>
<tr>
<td>P6</td>
<td>57.3</td>
<td>Cervix</td>
<td>C6.1</td>
<td>Whole lysate/fibroblast cell reduction</td>
<td>single cell sequencing</td>
<td>1448</td>
<td>402</td>
</tr>
<tr>
<td>P7</td>
<td>43</td>
<td>Cervix</td>
<td>C6.2</td>
<td>Whole lysate/fibroblast cell reduction</td>
<td>single cell sequencing</td>
<td>1487</td>
<td>130</td>
</tr>
<tr>
<td>P7</td>
<td>43</td>
<td>Cervix</td>
<td>C7.1</td>
<td>Whole lysate/fibroblast cell reduction</td>
<td>single cell sequencing</td>
<td>2606</td>
<td>1237</td>
</tr>
<tr>
<td>P7</td>
<td>43</td>
<td>Cervix</td>
<td>C7.2</td>
<td>Whole lysate/fibroblast cell reduction</td>
<td>single cell sequencing</td>
<td>1710</td>
<td>811</td>
</tr>
<tr>
<td>P8</td>
<td>57.8</td>
<td>Cervix</td>
<td>C8.1</td>
<td>Whole lysate/fibroblast cell reduction</td>
<td>single cell sequencing</td>
<td>1874</td>
<td>567</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>Biopsy</th>
<th>Biopsy code</th>
<th>Tissue processing method</th>
<th>Methods</th>
<th>CD8A/CD103 staining</th>
<th>CD8A/CD69 staining</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>54.6</td>
<td>Cervix</td>
<td>C1</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>P5</td>
<td>55.7</td>
<td>Cervix</td>
<td>C5</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>P7</td>
<td>43</td>
<td>Cervix</td>
<td>C7.2</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>P8</td>
<td>57.8</td>
<td>Cervix</td>
<td>C8.1</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>P8</td>
<td>57.8</td>
<td>Cervix</td>
<td>C8.2</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>P9</td>
<td>25</td>
<td>Cervix</td>
<td>C9</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td>P10</td>
<td>30</td>
<td>Cervix</td>
<td>C10</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>P11</td>
<td>36</td>
<td>Cervix</td>
<td>C11</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>NA</td>
<td>Yes</td>
</tr>
<tr>
<td>P12</td>
<td>23</td>
<td>Cervix</td>
<td>C12</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>P13</td>
<td>29</td>
<td>Cervix</td>
<td>C13</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>P14</td>
<td>27</td>
<td>Cervix</td>
<td>C14</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>P15</td>
<td>30</td>
<td>Cervix</td>
<td>C15</td>
<td>Fresh frozen</td>
<td>Immunofluorescent staining</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note: for subjects P6, P7 & P8, two cervix biopsies were taken 8 weeks apart. C1, C5, C7.2 & C8.1 were divided for single cell sequencing and immunofluorescent staining.
Table 2

<table>
<thead>
<tr>
<th>Cervix</th>
<th>Number of unique clonotypes</th>
<th>Matching on TCRB CDR3 amino acid sequences + VDJ gene usage</th>
<th>PBMC</th>
<th>Number of unique clonotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4_GZMB^{hi}ITGAE^{hi}</td>
<td>255</td>
<td>7</td>
<td>Pt4_PBMC_CD8^{hi}CD137^{hi}</td>
<td>103</td>
</tr>
<tr>
<td>C4_IFNG^{hi}CD69^{hi}</td>
<td>49</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5_GZMB^{hi}ITGAE^{hi}</td>
<td>126</td>
<td>8</td>
<td>Pt5_PBMC_CD8^{hi}CD137^{hi}</td>
<td>89</td>
</tr>
<tr>
<td>C5_IFNG^{hi}CD69^{hi}</td>
<td>55</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7.1_GZMB^{hi}ITGAE^{hi}</td>
<td>190</td>
<td>27</td>
<td>Pt7_PBMC_CD8^{hi}CD137^{hi}</td>
<td>167</td>
</tr>
<tr>
<td>C7.1_IFNG^{hi}CD69^{hi}</td>
<td>81</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>