

Supplementary Information

High-Throughput and High-Dimensional Single Cell Analysis of Antigen-Specific CD8⁺ T cells

Ke-Yue Ma^{1,10}, Alexandra A. Schonnesen^{2,10}, Chenfeng He², Amanda Y. Xia³, Eric Sun², Eunise Chen⁴, Katherine R Sebastian⁵, Yu-Wan Guo^{6,7}, Robert Balderas⁸, Mrinalini Kulkarni-Date⁵, Ning Jiang^{1,2,7,9,*}

¹Interdisciplinary Life Sciences Graduate Programs, the University of Texas at Austin, Austin, Texas, USA.

²Department of Biomedical Engineering, the University of Texas at Austin, Austin, Texas, USA.

³Department of Molecular Biosciences, the University of Texas at Austin, Austin, Texas, USA.

⁴Department of Nutritional Sciences, the University of Texas at Austin. Austin, Texas, USA.

⁵Department of Internal Medicine, Dell Medical School, the University of Texas at Austin, Austin, Texas, USA.

⁶McKetta Department of Chemical Engineering, the University of Texas at Austin. Austin, Texas, USA.

⁷Department of Oncology, Dell Medical School, the University of Texas at Austin, Austin, Texas, USA.

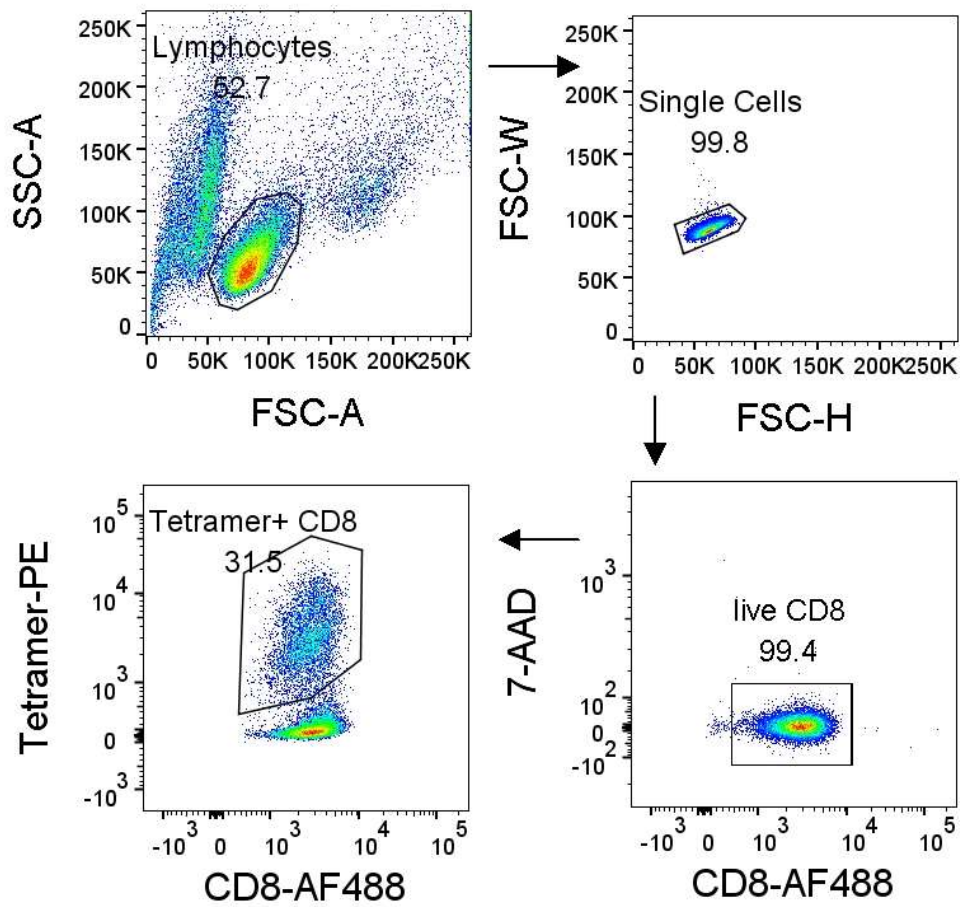
⁸Becton Dickinson, San Jose, California, USA.

⁹Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania, USA.

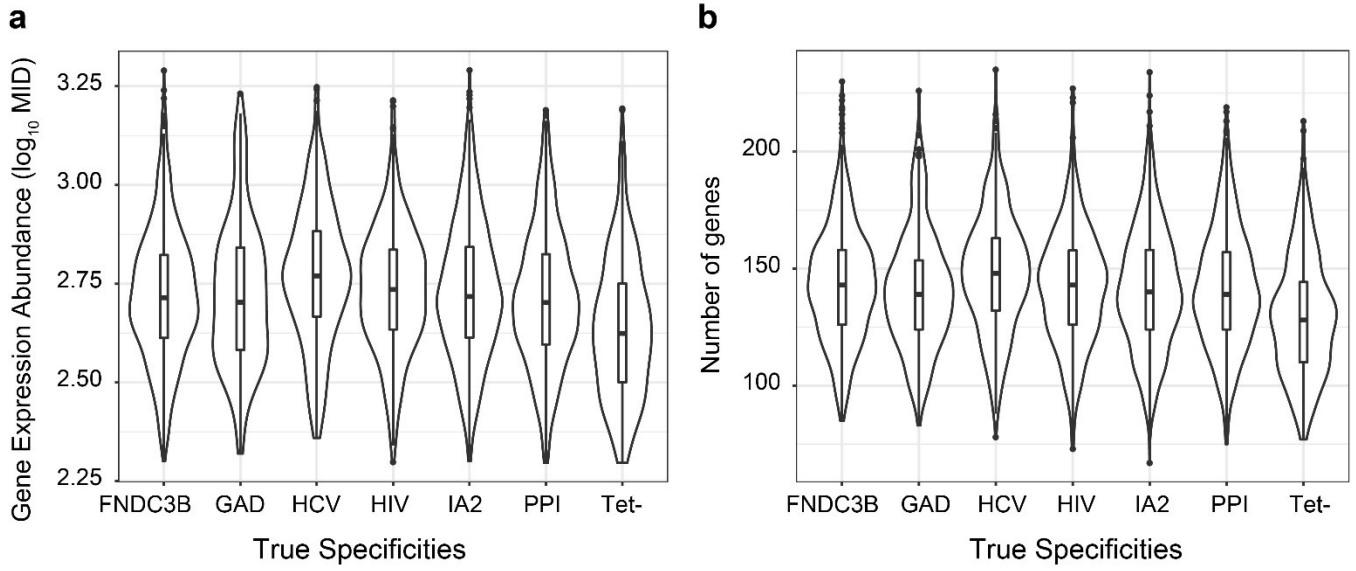
¹⁰These authors contributed equally

*Corresponding author: Ning Jiang, Ph.D.

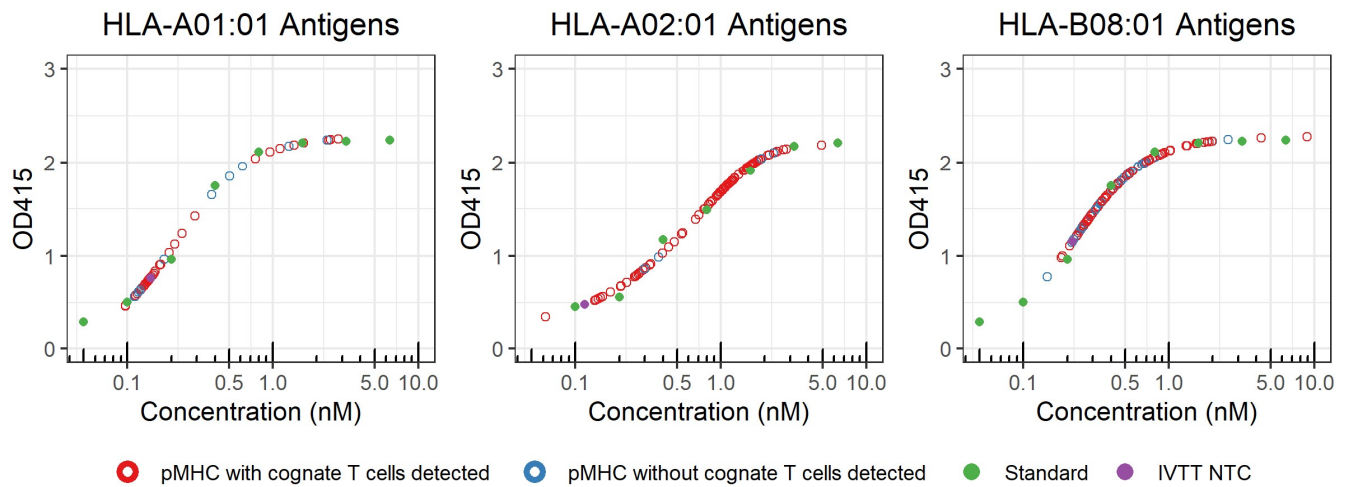
Email: jnjiang@seas.upenn.edu



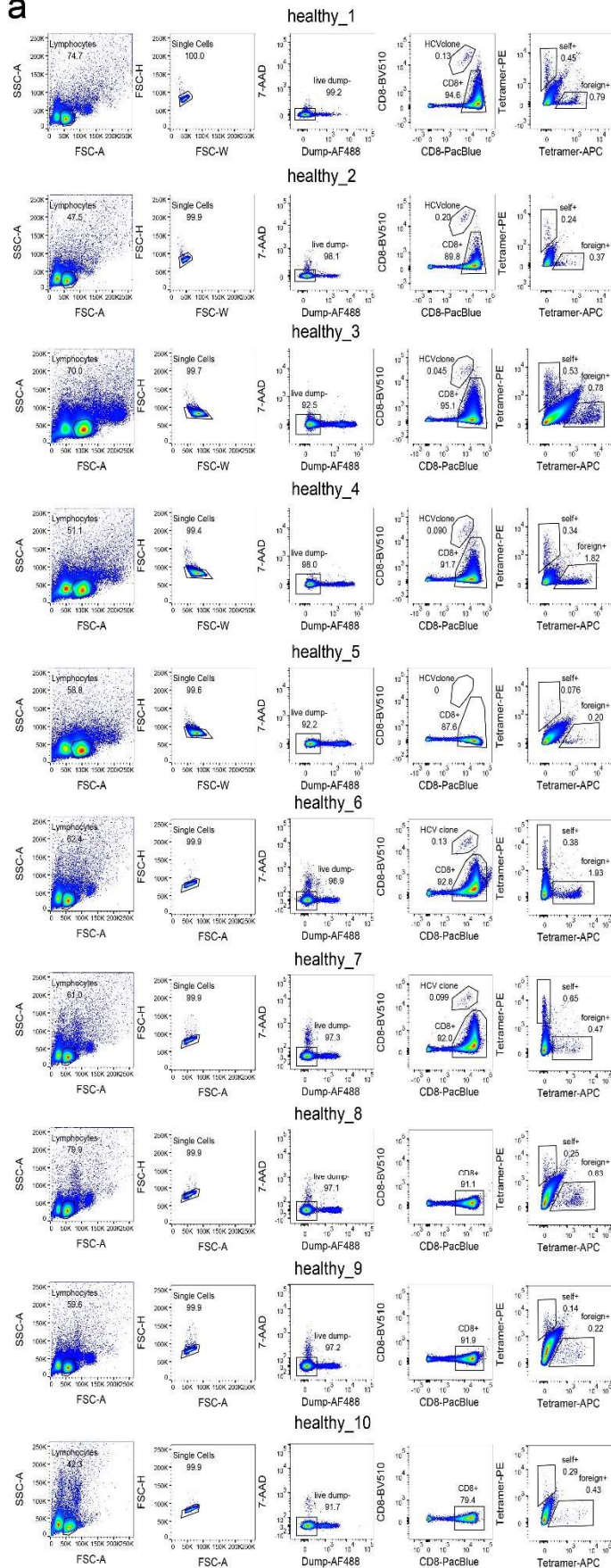
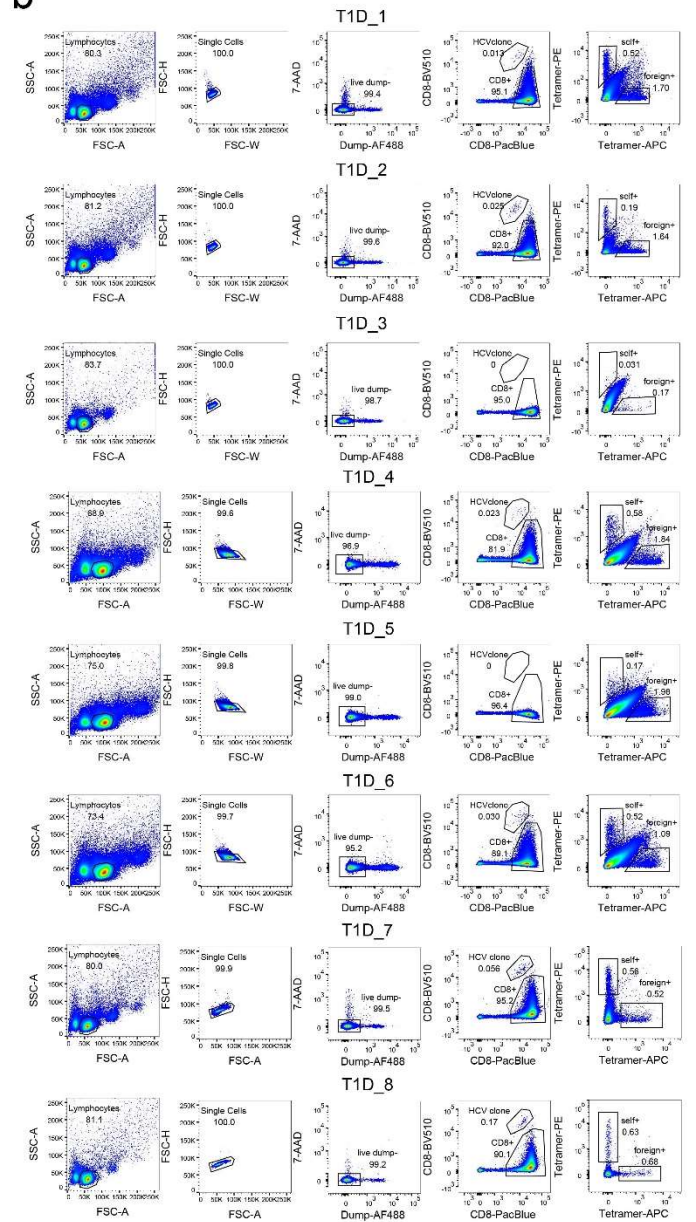
Supplementary Fig. 1. Gating strategy for sorting tetramer⁺ CD8⁺ T cells from a mixture of pMHC tetramer sorted polyclonal T cells cultured *in vitro*.



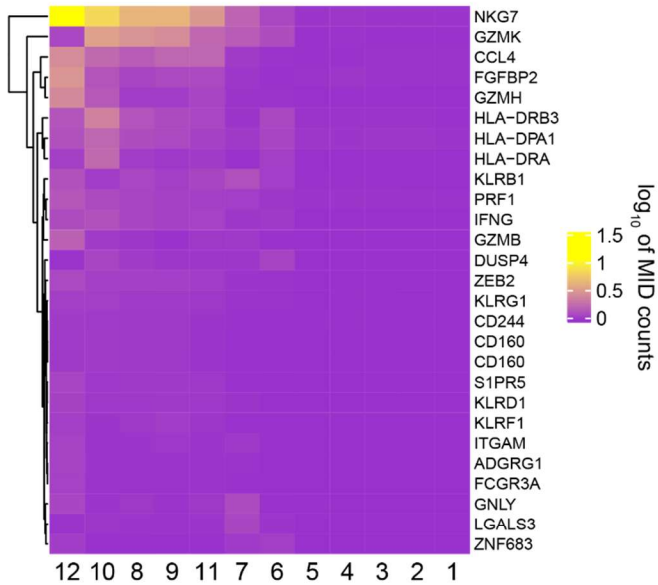
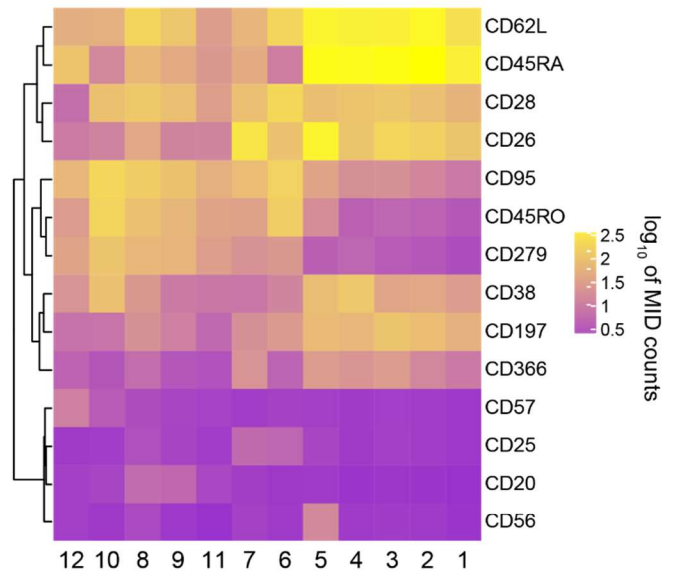
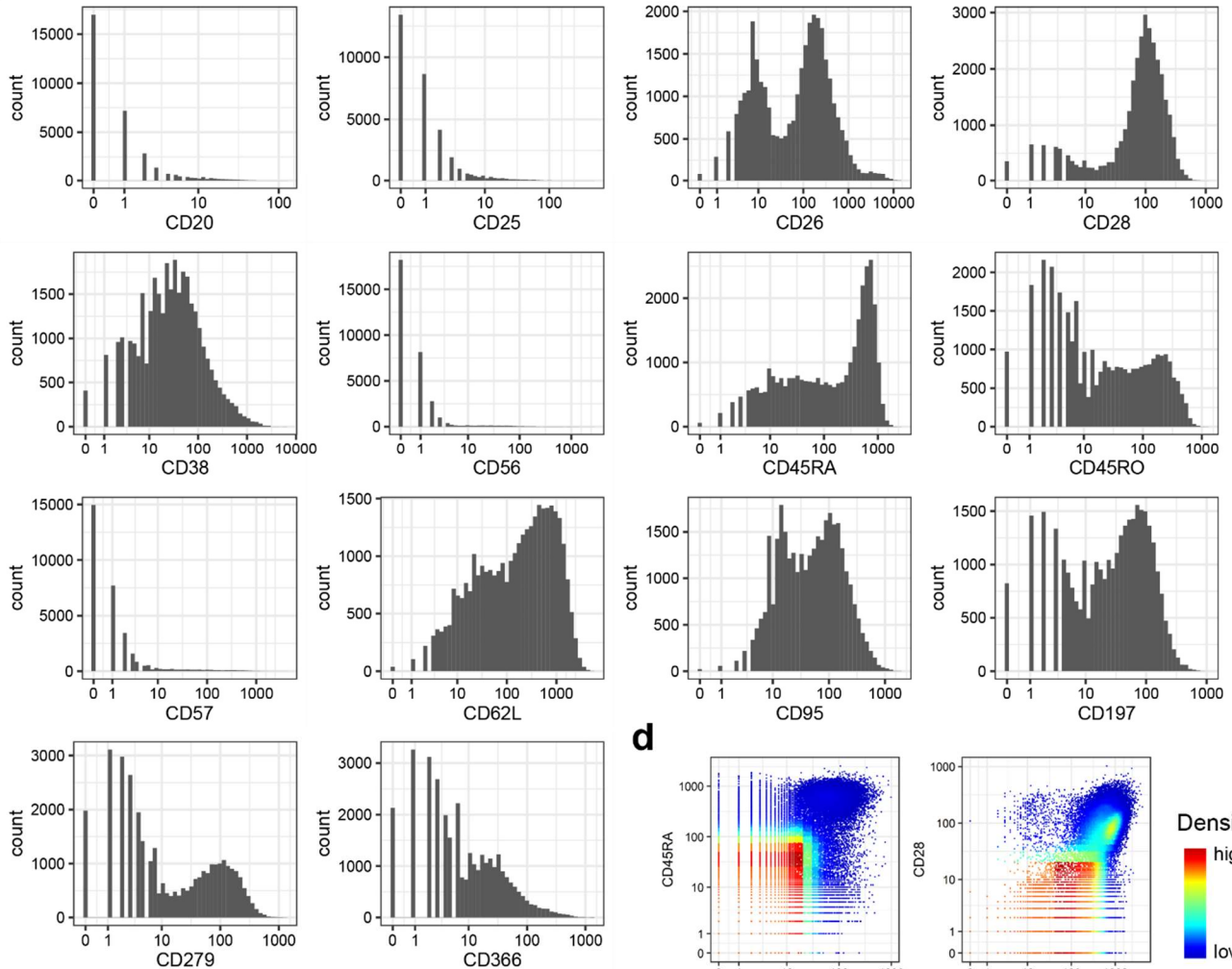
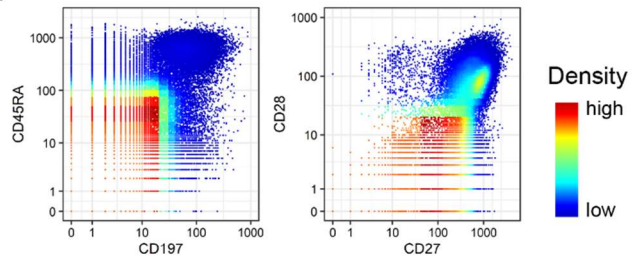
Supplementary Fig. 2. Distribution of (a) mRNA counts (\log_{10}) and (b) number of detected genes per cell among different antigen specific T cell populations.



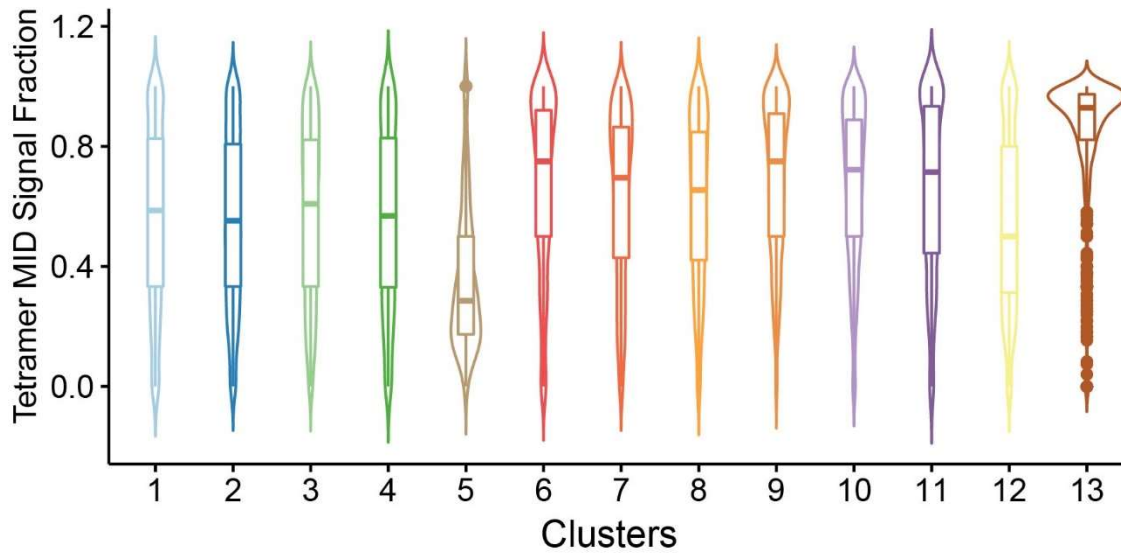
Supplementary Fig. 3. Concentration of pMHCs exchanged with IVTT generated peptides measured by ELISA. IVTT NTC: UV peptide exchange was done using non-templated control in IVTT reaction. Standard: HLA-A2 folded with KLVALGINAV peptide (HCVNS3-1406-1415), which was provided by NIH tetramer core.

a**b**

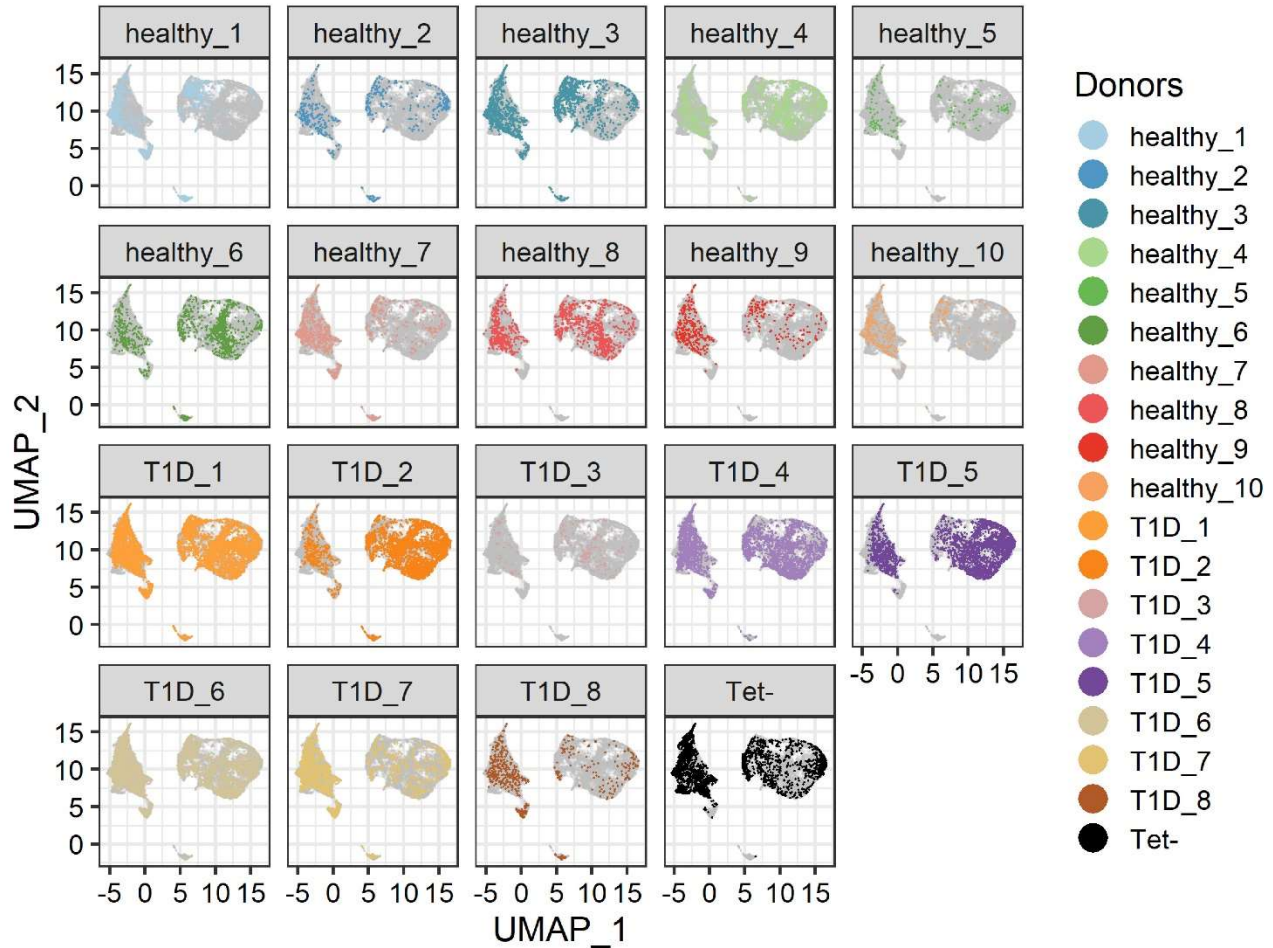
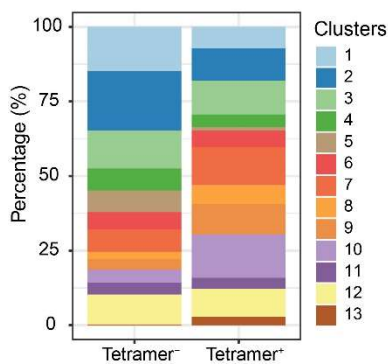
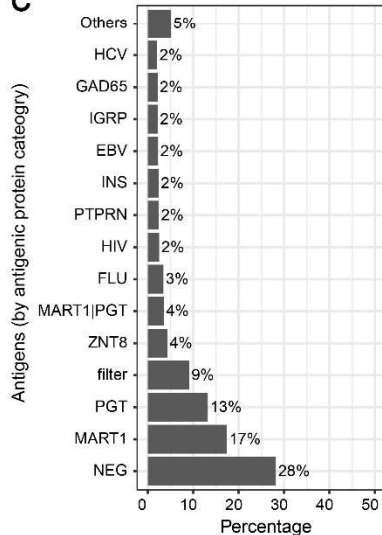
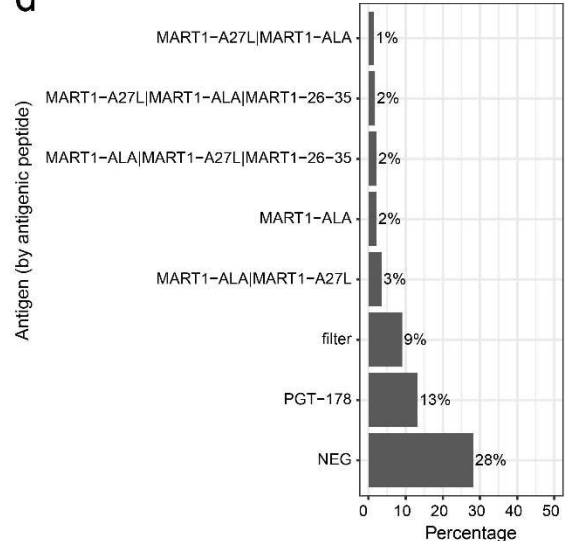
Supplementary Fig. 4. Gating strategy to sort tetramer⁺ T cells from primary CD8⁺ T cells enriched from PBMCs of (a) healthy and (b) T1D donors. Cultured HCV antigen-specific clone was spiked in primary CD8⁺ T cells whose HLA alleles include A02:01 and a different fluorescently labeled CD8⁺ antibody was used to pre-stain HCV antigen-specific clone.

a**b****c****d**

Supplementary Fig. 5. Absolute expression (\log_{10} of MID counts) of differentially expressed (a) genes and (b) surface-proteins among different clusters. Genes and surface-protein plotted here are the same set as in Fig. 3d. (c). Distributions of AbSeq MID counts for differentially expressed surface proteins shown in Fig.3d. (d). Density plot showing of all CD8 T cells by AbSeq MID counts of CD45RA vs. CD197, and CD28 vs. CD27, respectively. Colors represent the local density of cells on the two-dimension space.

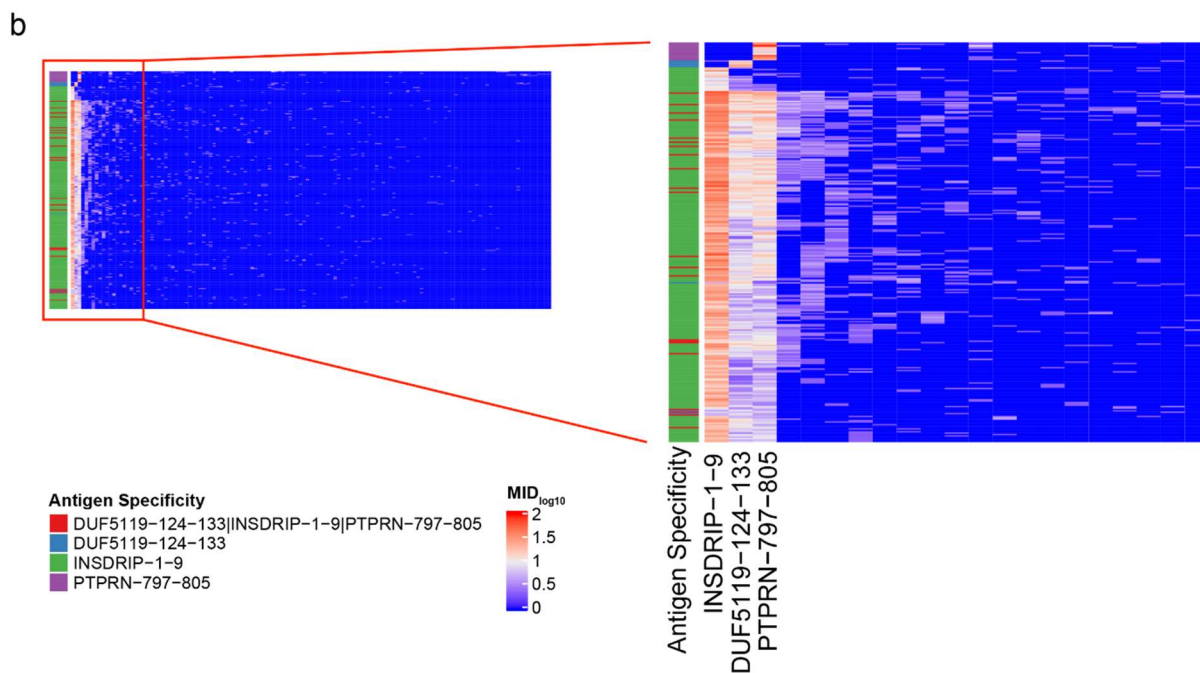
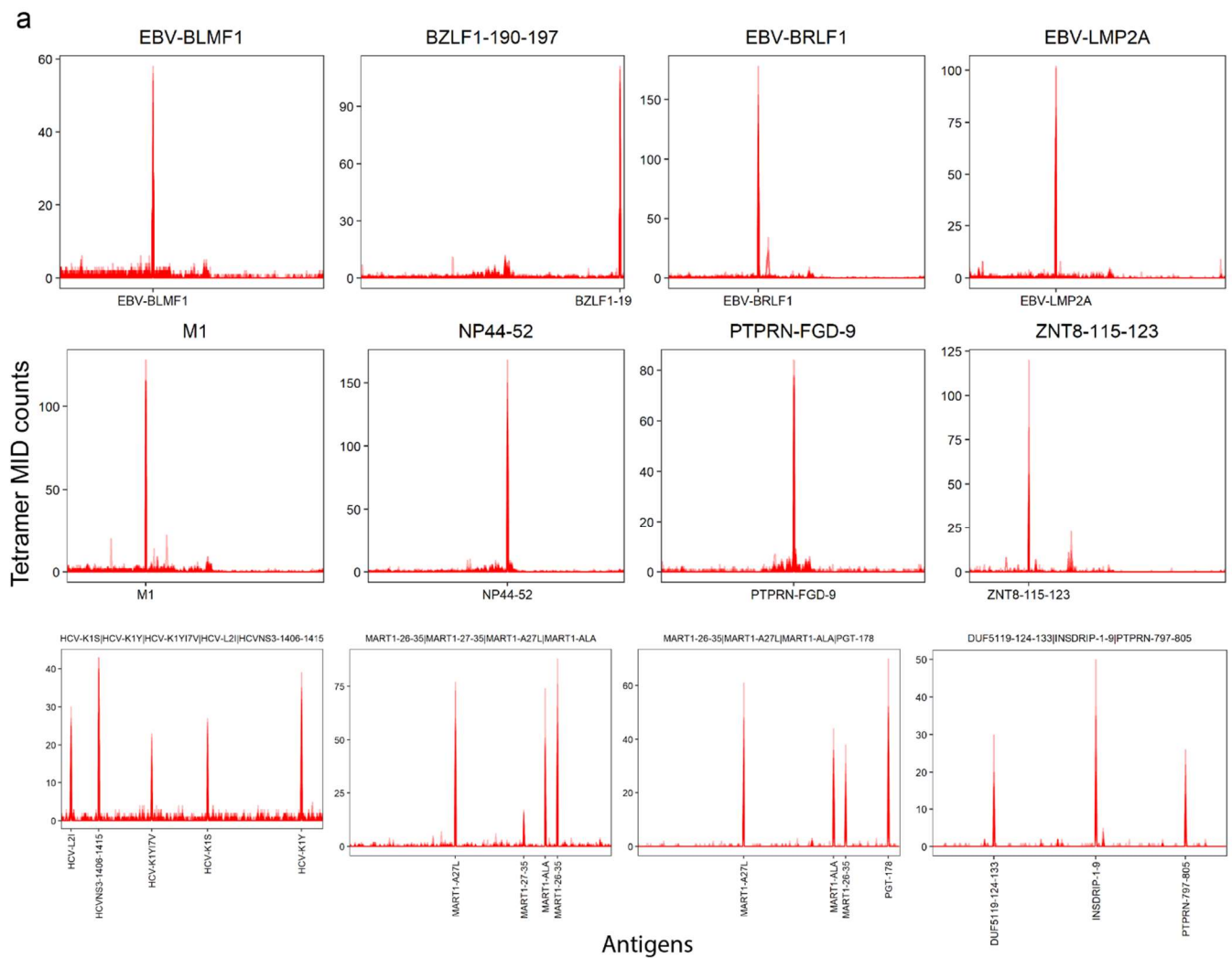


Supplementary Fig. 6. The distribution of tetramer DNA barcode MID signal fraction among different phenotype clusters combined from all donors. For each cell, tetramer MID signal fraction was defined as the fraction of cumulative MID count from putative binding antigens over cumulative MID count from all bound antigens (**See methods**).

a**b****c****d**

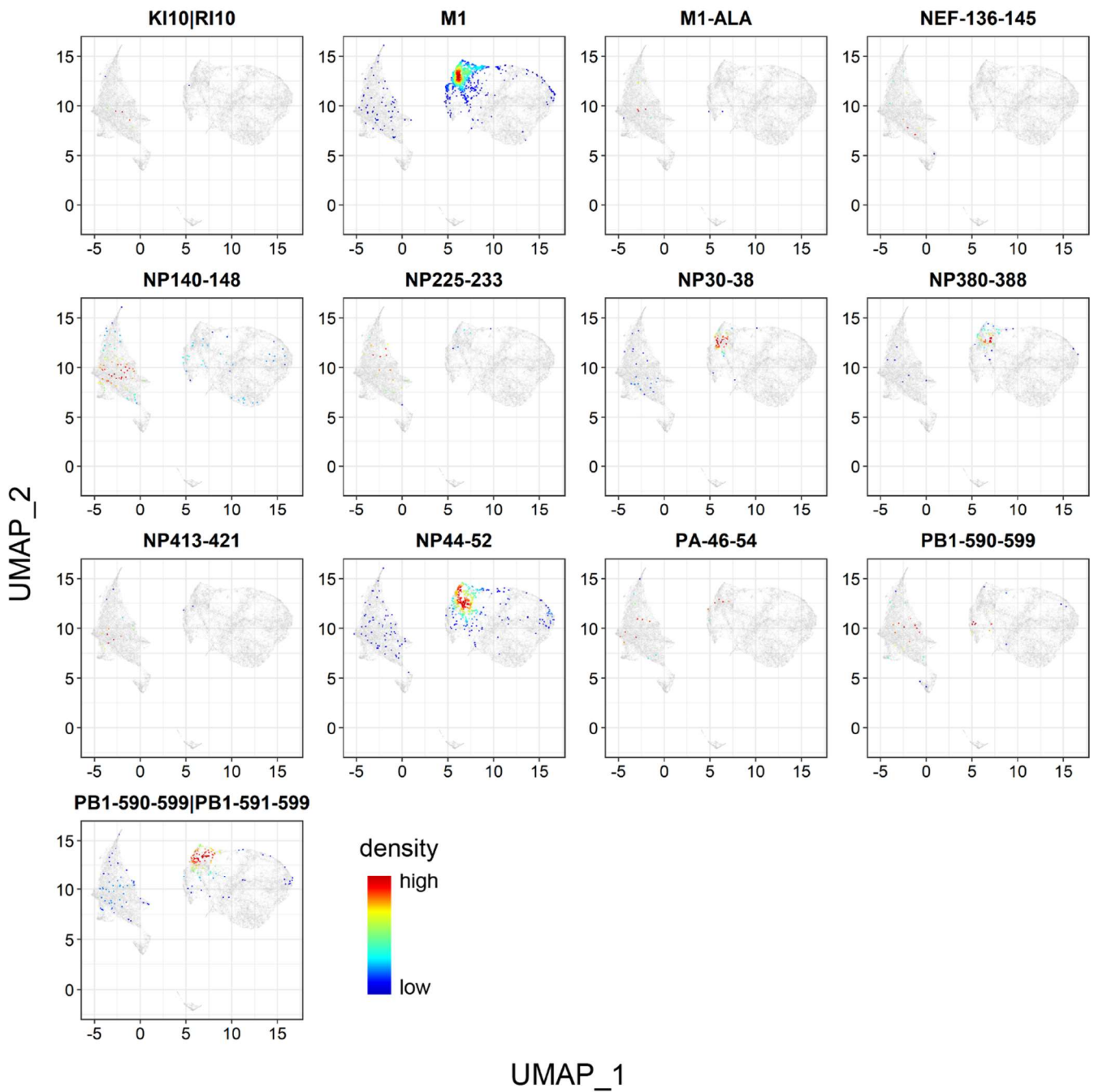
Supplementary Fig. 7. (a) UMAP projection of single cells among different donors. Grey dots represent all cells and colored dots are cells from different chips. **(b)** Comparison of distribution of phenotypes between Tetr⁺ and Tetr⁻ CD⁸⁺ T cells. **(c)** Percentage of naïve population in each antigen specific CD⁸⁺ T cell group. **(d)** Percentage breakdown of naïve CD⁸⁺ T cells among all major antigen specificities. Only antigen specificities with percentage greater than 1% within naïve population are shown. NEG cells were Tetr⁻ CD⁸⁺ T cells identified by Tetr⁺ MID counts. Cells were classified

into "filter" category based on following criteria: 1) more than one antigen bind to single cell, and these antigens are more than 3 amino acid distance away from each other; 2) correlation of tetramer MID between single cell and median of all cells with same TCR sequence is below 0.9, identified as described in Methods.

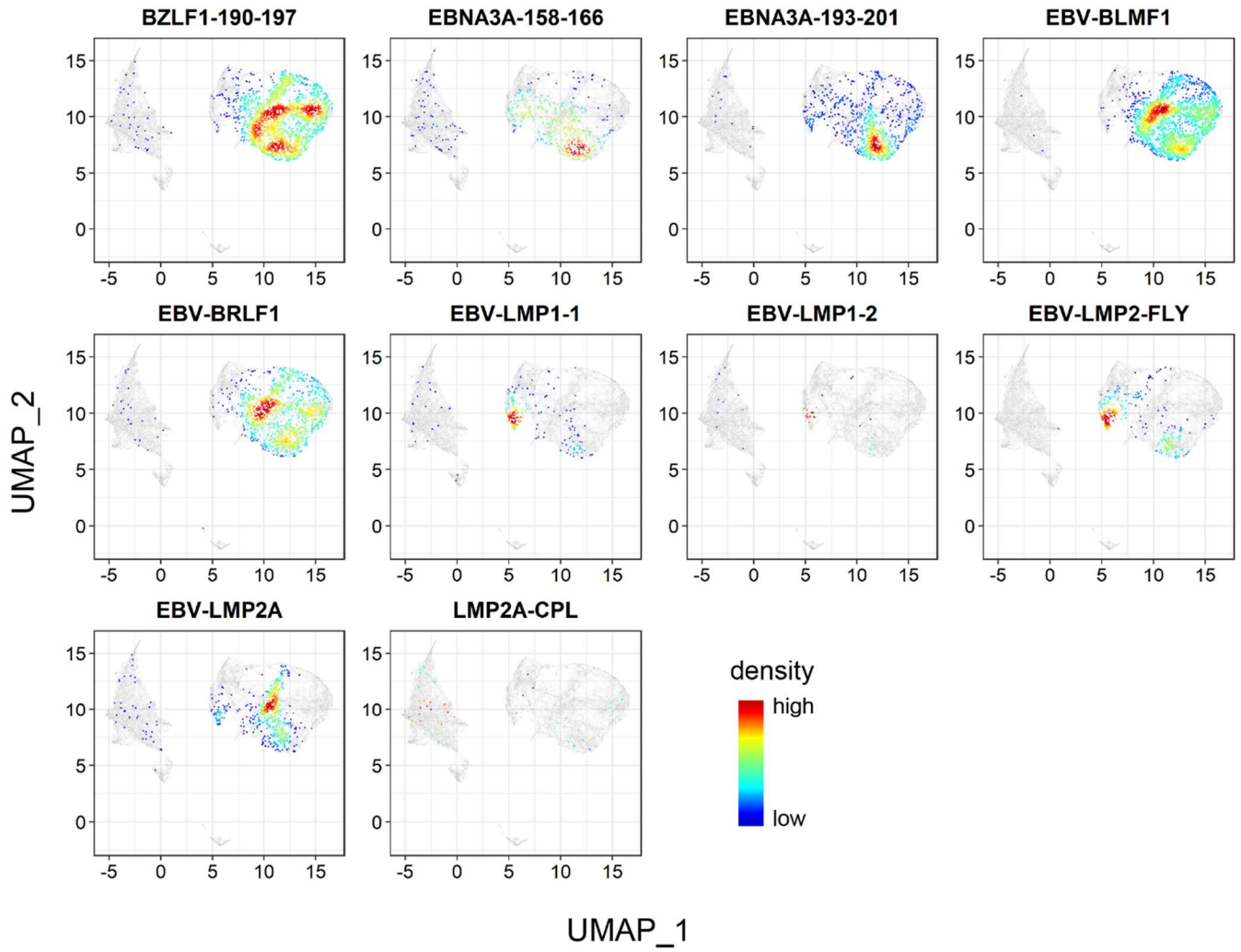


Supplementary Fig. 8. (a) Distribution of tetramer MID counts for eight antigen specificities, including EBV antigens (EBV-BLMF1, BZLF1-190-197, EBV-BRLF1, EBV-LMP2A), influenza viral antigens (M1,

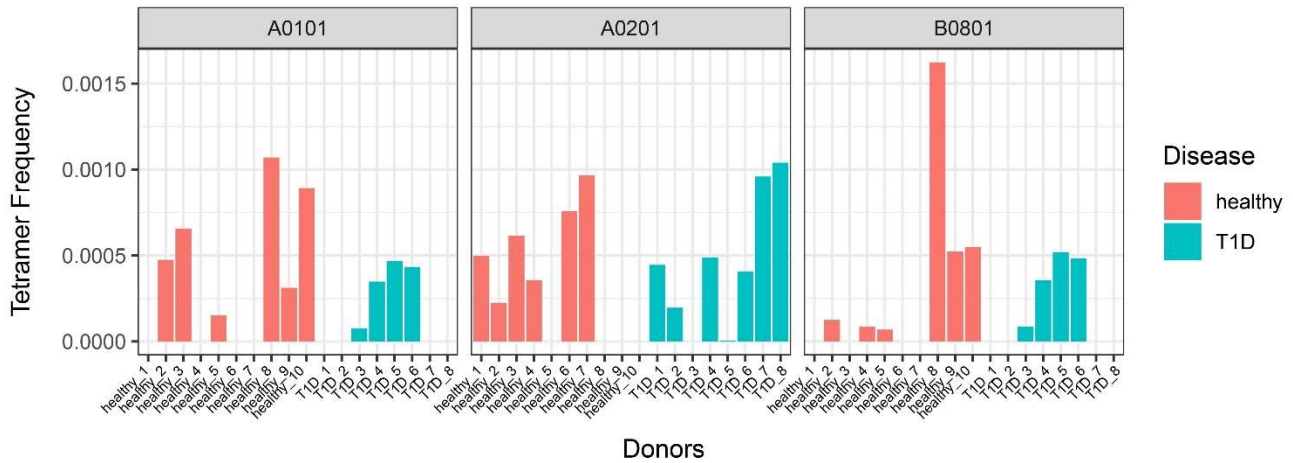
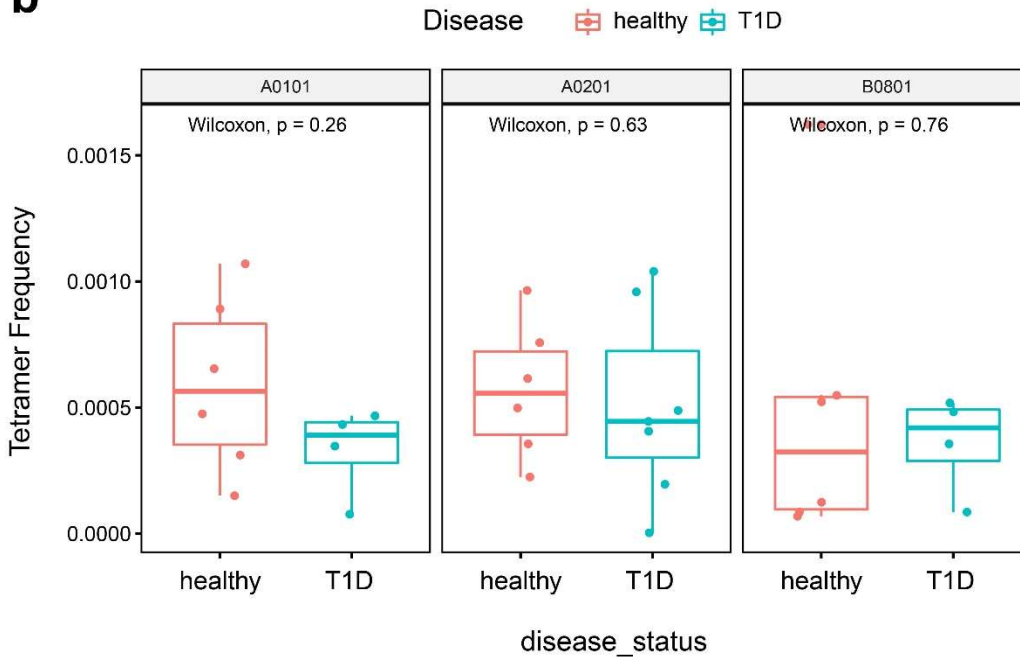
NP44-52), T1D associated antigens (PTPRN-FGD-9, ZNT8-115-123) or cross-reactive antigens (HCV, Mart1, and DUF5119-124-133/INSDRIP-1-9/PTPRN-797-805). For each cell in the group, the MID counts for each of the 280 antigens used in the experiment were tallied and then overlaid in the same order of the 280 antigens. Only the antigens emerge after the filter are labeled on the X axis and their position in the 280 antigen list were indicated by a tick on the x-axis. Each panel with a sharp single peak indicated single antigen-specificity, while panels with multiple sharp peaks indicated cross-reactive antigens. (b) Comparison of tetramer MID counts among DUF5119-124-133/INSDRIP-1-9/PTPRN-797-805 cross-reactive and single antigen specific cells.



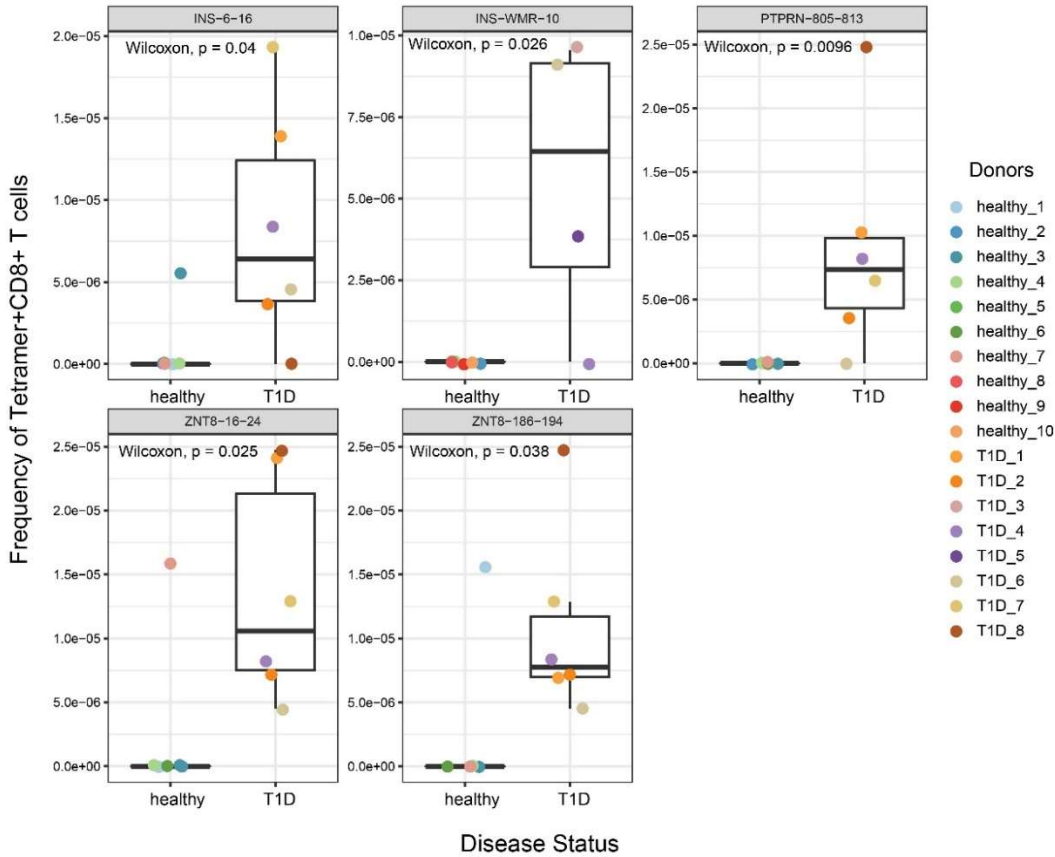
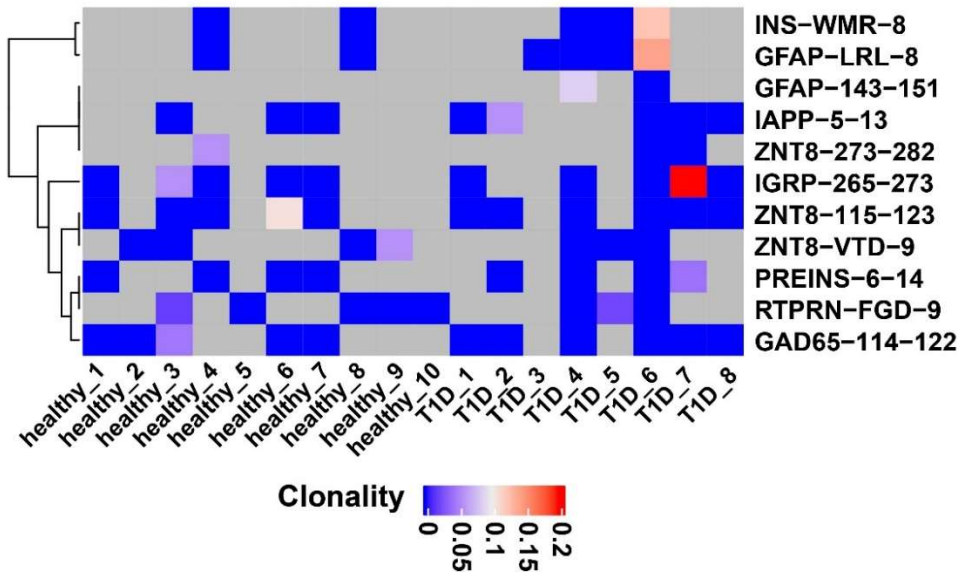
Supplementary Fig. 9. Density heatmap of the projection of influenza specific T cells on UMAP plot. Colors represent the local density of cells on the two-dimensional space.



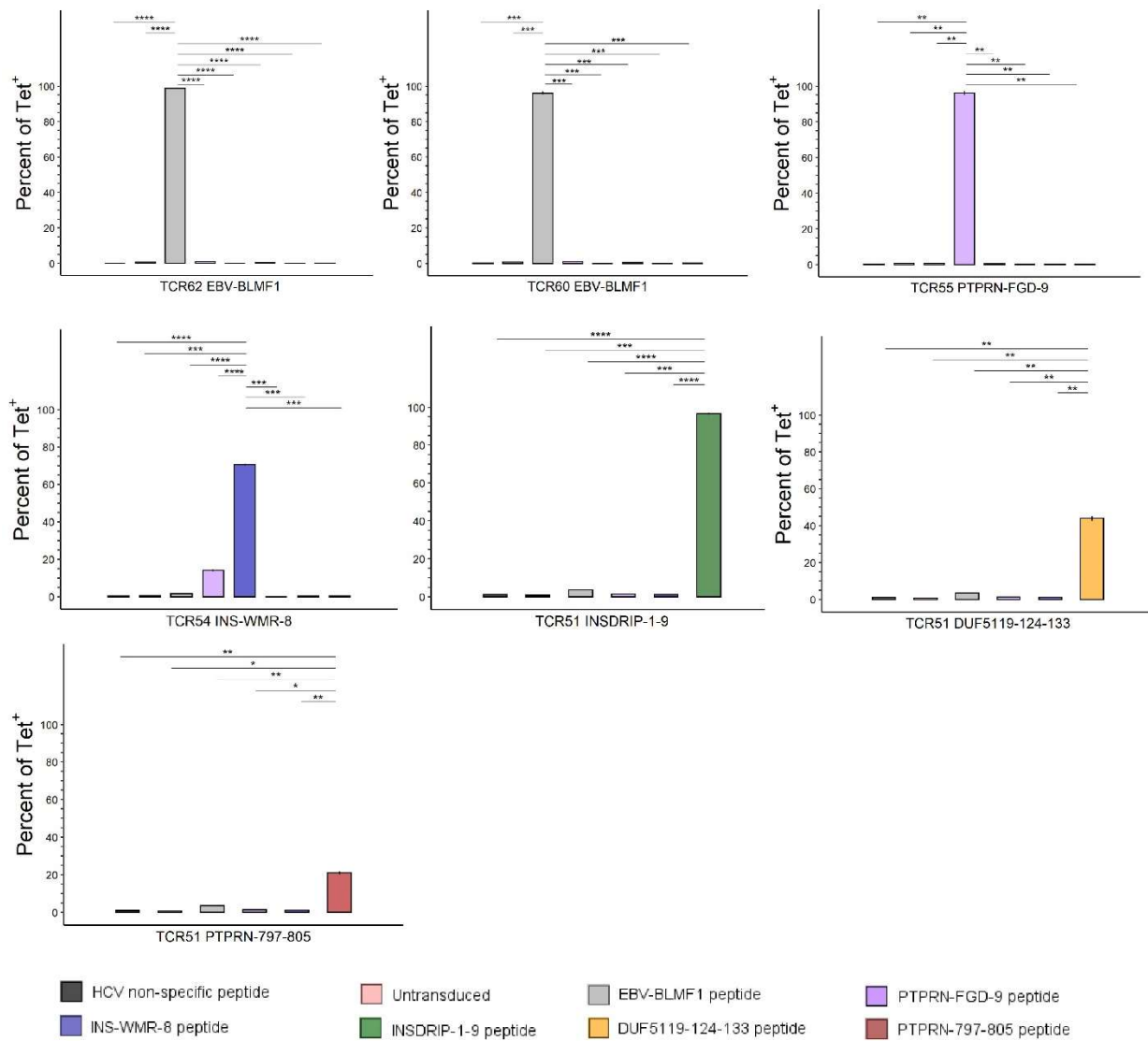
Supplementary Fig. 10. Density heatmap of the projection of EBV specific T cells on UMAP plot. Colors represent the local density of cells on the two-dimensional space.

a**b**

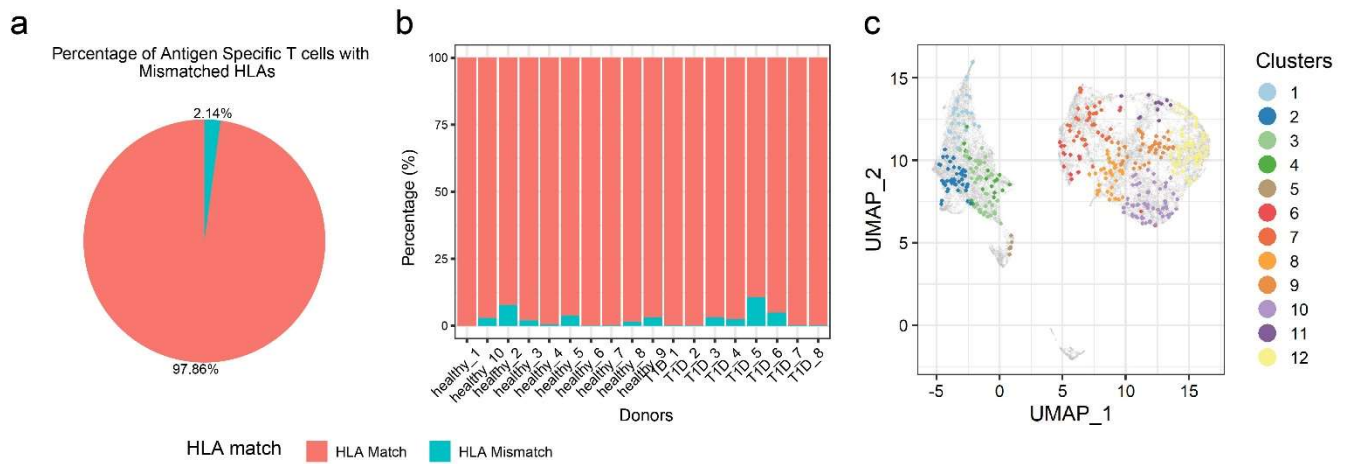
Supplementary Fig. 11. Frequency of total T1D autoantigen-specific CD8⁺ T cells in healthy subjects and T1D patients. **(a).** Frequency of T1D autoantigen tetramer⁺ CD8⁺ T cells in different donors for various HLA alleles. **(b).** Comparison of total T1D autoantigen tetramer⁺ CD8⁺ T cells between healthy and T1D donors for various HLA alleles. Wilcoxon non-parametric test was performed.

a**b**

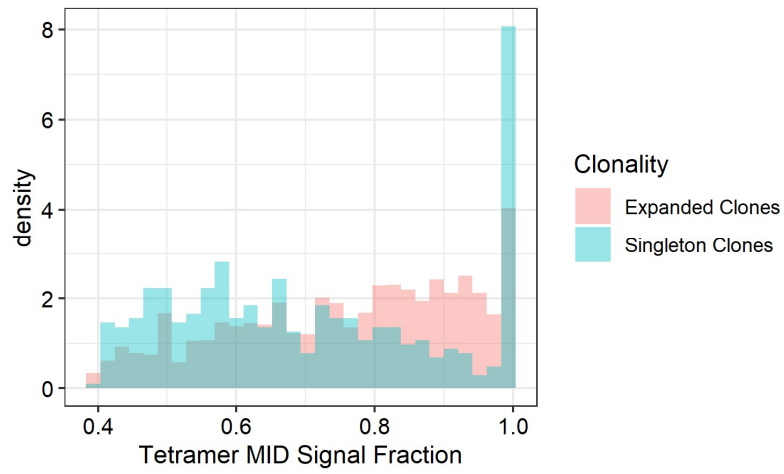
Supplementary Fig. 12. T1D autoantigens with different antigen-specific CD8⁺ T cell frequencies and clonality between healthy subjects and T1D patients. **(a).** Five T1D autoantigens were identified to have significant higher frequency of antigen specific T cells in peripheral blood when MID negative threshold was set to 15. Wilcoxon non-parametric test was performed. **(b).** TCR clonality heatmap of T1D antigenic specific T cells for each antigen/donor combination. Grey, no T cells were detected.



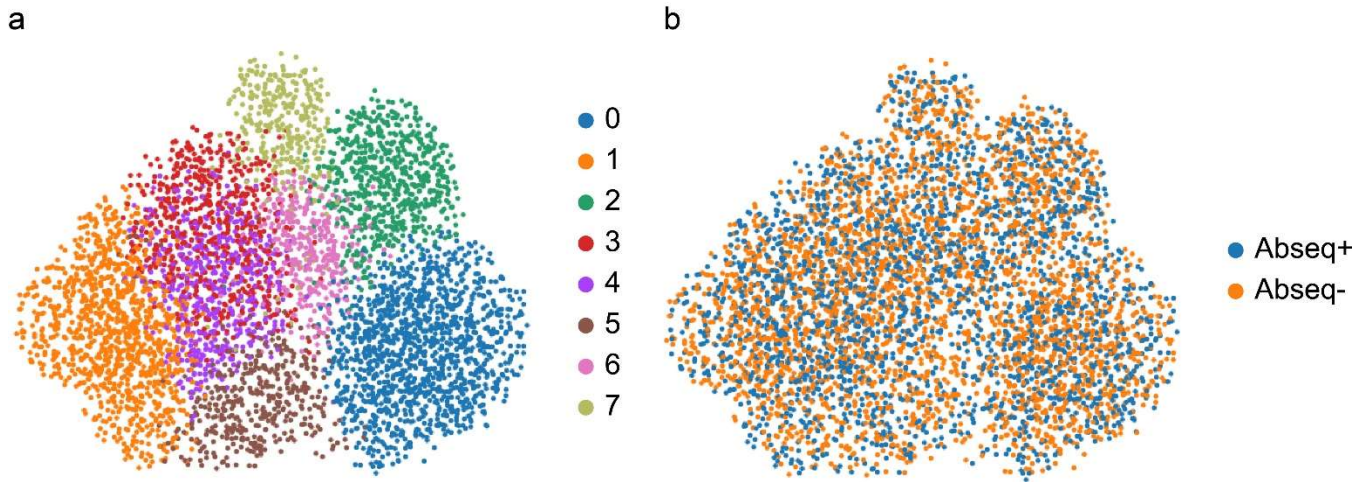
Supplementary Fig. 13. TCR specificity and cross-reactivity validation by tetramer staining. Bar plot showing the percentage of Tetramer⁺ cells gated on TCRβ^{hi} fraction of the cells, corresponding to **Fig 5b**. Tetramer staining experiments were performed in triplicates. Two-tailed Student's t-test was performed between cognate tetramer and each negative control for all TCRs. EBV-BLMF1: GLCTLVAML; INSDRIP-1-9: MLYQHLLPL; DUF5119-124-133: MVWGPDPPLYV; PTPRN-797-805: MVWESGCTV; PTPRN-FGD-9: FGDHPGHSY; INS-WMR-8: WMRLPLL. ns, not significant; *, p ≤ 0.05; **, p ≤ 0.01; ***, p ≤ 0.001; ****, p ≤ 0.0001.



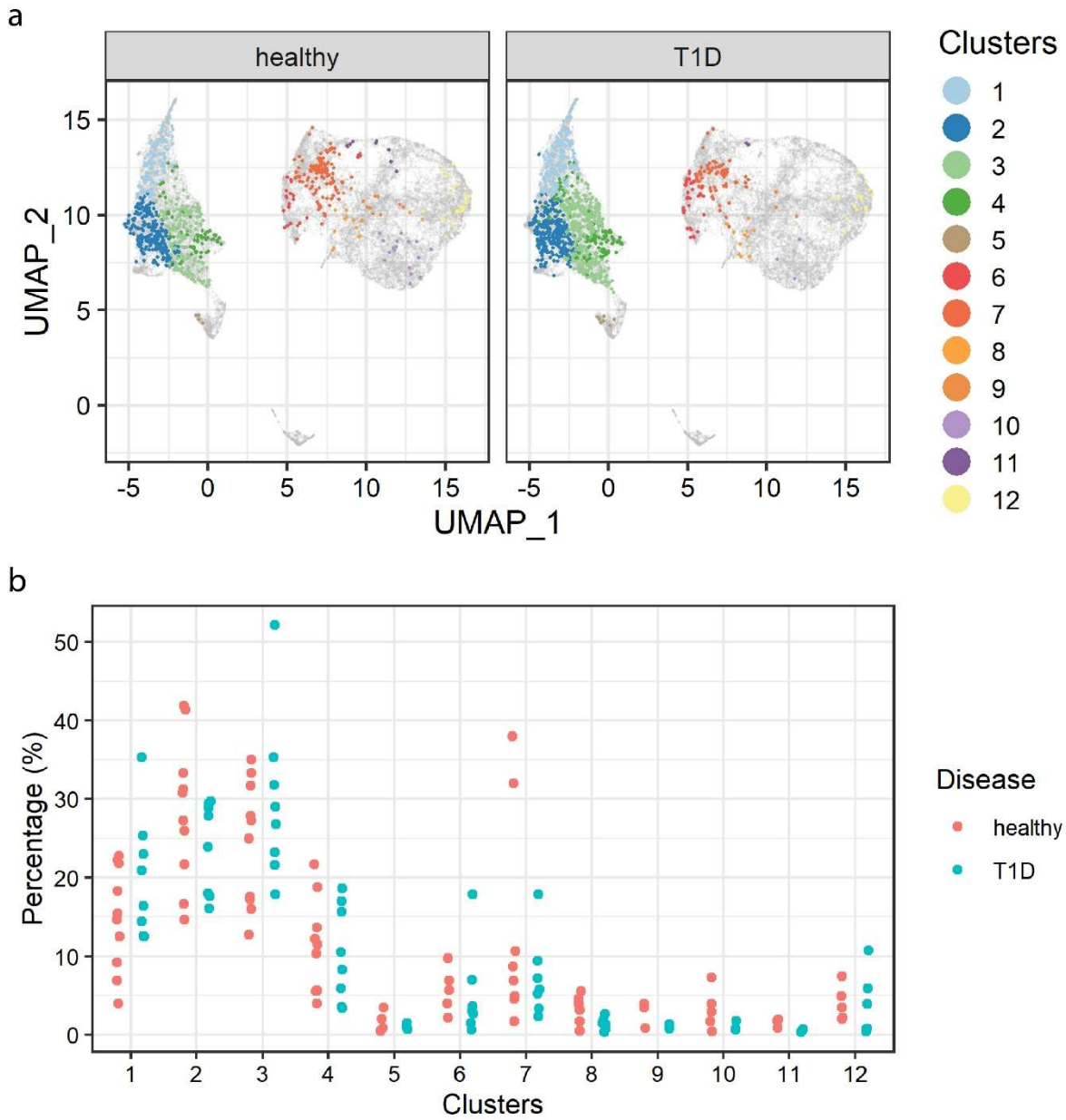
Supplementary Fig. 14. Analysis of T cells with bound antigen specificity being mismatched HLA alleles. **(a)** Summary of percentage of antigen specific T cells with mismatched HLA alleles in all donors. Combined percentage from two sources were presented (see Discussion). **(b)** Percentage of antigen specific T cells with mismatched HLA alleles in each donor. Combined percentage of two sources were presented (see Discussion). **(c)** Comparison of phenotypes of cells with mismatched HLA alleles with the overall population. Gray dots represent all CD8⁺ T cells.



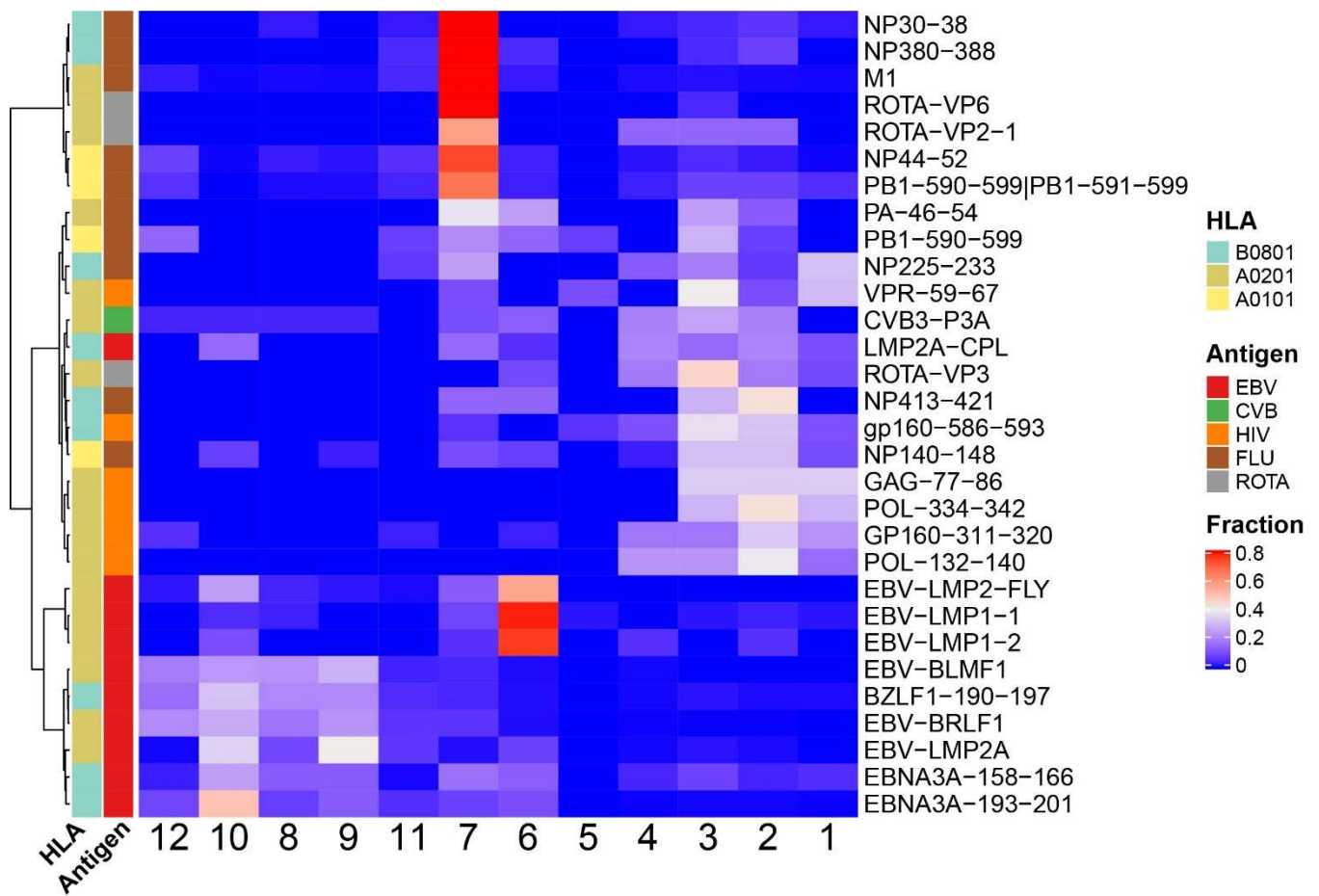
Supplementary Fig. 15. Distribution of tetramer MID signal fraction between expanded clones and singleton clones. Tetramer MID signal fraction was defined as the fraction of cumulative MID count from putative binding antigens over cumulative MID count from all bound antigens (see **Methods**).



Supplementary Fig.16: Comparison of single cell clustering using gene expression data between samples with or without AbSeq. (a) Single cell clustering colored by leiden cluster. (b) Single cell clustering colored by samples with or without AbSeq.

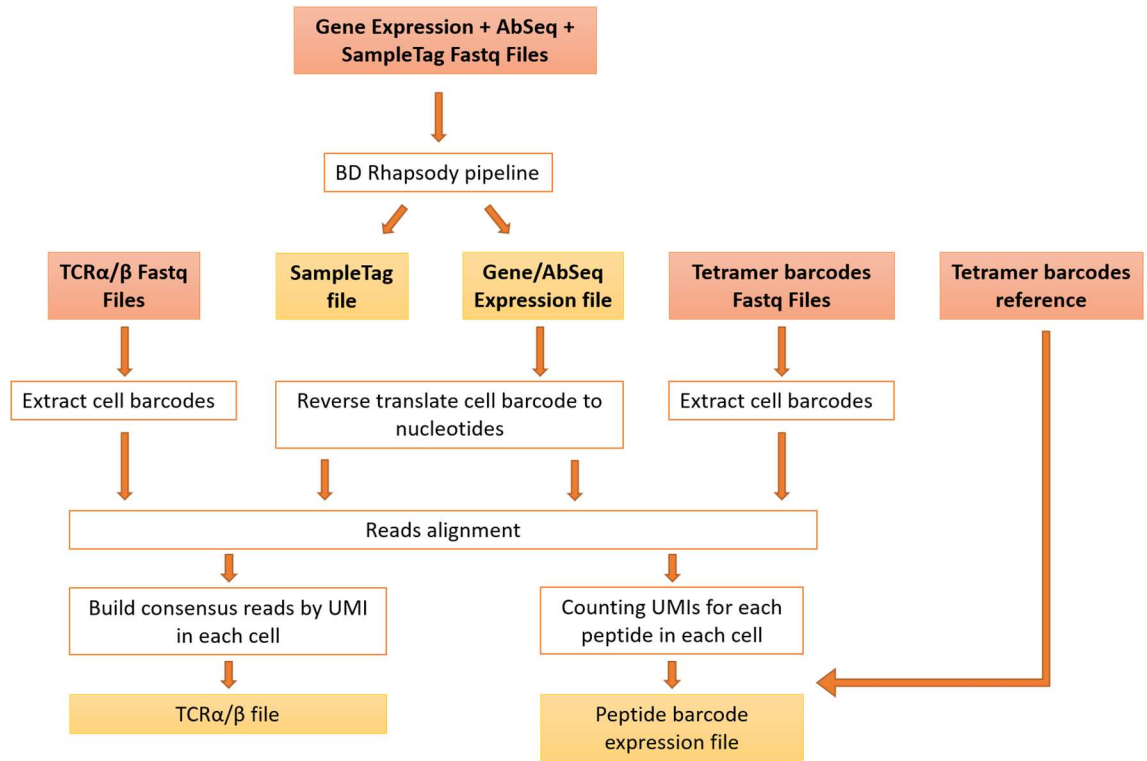


Supplementary Fig. 17. Comparison of T1D antigen specific CD8⁺ T cells between T1D patients and healthy subjects. (a) UMAP projection of T1D antigen specific CD8⁺ T cells in T1D patients and healthy subjects respectively. Colored dots are T1D antigen specific CD8⁺ T cells and gray dots are other cells. (b) Comparison of distribution of phenotypes among T1D antigen specific CD8⁺ T cells in each donor. Wilcox test was performed with no significance between T1D and healthy subjects in any cluster.



Supplementary Fig. 18. Distribution of viral antigen specific CD8⁺ T cells among 12 primary CD8⁺ T cells clusters in all 18 donors when tetramer MID negative threshold was set to 15 (**See methods**).

Legend



Supplementary Fig. 19. Diagrams illustrating the bioinformatic workflow to analyze target gene expression, AbSeq, SampleTag, tetramer DNA barcode, and TCR simultaneously for single cells.