frontiers in Immunology

ORIGINAL RESEARCH
published: 06 November 2019
doi: 10.3389/fimmu.2019.02568

Gene Expression-Based Identification of Antigen-Responsive CD8$^+$ T Cells on a Single-Cell Level

Yannick F. Fuchs[1†], Virag Sharma[1,2*†], Anne Eugster[1], Gloria Kraus[1], Robert Morgenstern[1], Andreas Dahl[3‡], Susanne Reinhardt[3], Andreas Petzold[3], Annett Lindner[1], Doreen Löbel[1] and Ezio Bonifacio[1,2,4*]*

[1] Faculty of Medicine, DFG Center for Regenerative Therapies Dresden, Technische Universität Dresden, Dresden, Germany, [2] German Center for Diabetes Research (DZD), Paul Langerhans Institute Dresden, Technische Universität Dresden, Dresden, Germany, [3] DRESDEN-Concept Genome Center c/o Center for Molecular and Cellular Bioengineering, Technische Universität Dresden, Dresden, Germany, [4] Institute of Diabetes and Obesity, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

OPEN ACCESS

Edited by:
Nick Gascoigne,
National University of Singapore, Singapore

Reviewed by:
Sid P. Kerkar,
Boehringer Ingelheim, United States
Guo Fu,
Xiamen University, China

***Correspondence:**
Yannick F. Fuchs
yannick.fuchs@tu-dresden.de
Virag Sharma
virag.sharma@mailbox.tu-dresden.de
Ezio Bonifacio
ezio.bonifacio@tu-dresden.de

[†] These authors have contributed equally to this work

[‡]ORCID:
Andreas Dahl
orcid.org/0000-0002-2668-8371

Specialty section:
This article was submitted to T Cell Biology, a section of the journal Frontiers in Immunology

Received: 21 June 2019
Accepted: 16 October 2019
Published: 06 November 2019

Citation:
Fuchs YF, Sharma V, Eugster A, Kraus G, Morgenstern R, Dahl A, Reinhardt S, Petzold A, Lindner A, Löbel D and Bonifacio E (2019) Gene Expression-Based Identification of Antigen-Responsive CD8$^+$ T Cells on a Single-Cell Level. Front. Immunol. 10:2568. doi: 10.3389/fimmu.2019.02568

CD8$^+$ T cells are important effectors of adaptive immunity against pathogens, tumors, and self antigens. Here, we asked how human cognate antigen-responsive CD8$^+$ T cells and their receptors could be identified in unselected single-cell gene expression data. Single-cell RNA sequencing and qPCR of dye-labeled antigen-specific cells identified large gene sets that were congruently up- or downregulated in virus-responsive CD8$^+$ T cells under different antigen presentation conditions. Combined expression of *TNFRSF9*, *XCL1*, *XCL2*, and *CRTAM* was the most distinct marker of virus-responsive cells on a single-cell level. Using transcriptomic data, we developed a machine learning-based classifier that provides sensitive and specific detection of virus-responsive CD8$^+$ T cells from unselected populations. Gene response profiles of CD8$^+$ T cells specific for the autoantigen islet-specific glucose-6-phosphatase catalytic subunit-related protein differed markedly from virus-specific cells. These findings provide single-cell gene expression parameters for comprehensive identification of rare antigen-responsive cells and T cell receptors.

Keywords: CD8$^+$ T cells, single-cell, antigen-responsive, gene-expression analysis, CTL (cytotoxic T lymphocyte), influenza matrix protein, CMV pp65, T cell receptor (TCR)

INTRODUCTION

CD8$^+$ T cells are integral to the clearance of virus-infected cells and the control of cell transformation. These attributes are exploited by therapies, such as vaccination against infection and immune therapies targeting cancer. CD8$^+$ T cells are also involved in the destruction of cells in some autoimmune diseases, such as type 1 diabetes, and in graft-vs.-host disease. In active immune responses, CD8$^+$ T cells undergo rapid clonal expansion and this expansion of activated CD8$^+$ T cells can be a marker of ongoing infection in immune-mediated disease. Therefore, it is conceivable that the detection and information provided by clonal expansions can be used to identify target antigens or disease-causing agents, and to develop therapies that exploit CD8$^+$ T cell specificity to control disease.

CD8+ T cell specificity is provided by T cell receptors (TCR), which recognize a cognate antigen peptide presented by major histocompatibility complex (MHC) class I molecules on the cell surface. This mechanism of antigen recognition has been used to develop fluorochrome-labeled, peptide-loaded MHC class I multimers and has led to sophisticated methods combining multiple fluorochromes and peptides to find and isolate antigen-specific CD8+ T cells (1–3). These reagents also allow detailed TCR and phenotypic analysis of cells using single-cell technologies (4–6). Multimers require high levels of TCR expression for the detection of antigen-specific cells (7). Although high TCR expression is a feature of CD8+ T cells in resting conditions, CD8+ T cells undergo numerous changes in gene and protein expression upon stimulation with their target peptide (8). In particular, downregulation of TCR is a consistent response to the engagement of the cognate peptide, conceivably for negative feedback control of T cell activity (9–11). Therefore, the selection of multimer-binding CD8+ T cells may bias our understanding of the phenotype of antigen-specific CD8+ T cells.

We have incorporated MHC class I multimers and peptide activation into an unbiased approach to analyze antigen-specific CD8+ T cells from peripheral blood. Here, we demonstrate marked downregulation of TCR upon stimulation with the cognate peptide and describe protein and gene sets that can be used to identify and isolate antigen-responsive CD8+ T cells. We show how these approaches can be combined with MHC multimers to obtain a comprehensive description of activated and non-activated antigen-specific CD8+ T cells, and how they can be used without MHC multimers for identification and in-depth gene expression profiling of cells responding to their cognate antigen. Additionally, we provide evidence that multimer binding (specificity) and peptide responsiveness are not identical.

RESULTS

Identification of Highly Pure Flu MP_{58-66}-Responsive CD8+ T Cells for Single-Cell Analysis

The activation of influenza matrix protein 1_{58-66} (Flu MP_{58-66})-specific CD8+ T cells by presentation of their cognate peptide compromised the ability to detect these cells using MHC class I multimers (**Figure 1A**), presumably due to downregulation of TCR (**Figure S1**). Therefore, we suspected that isolation of multimer-positive CD8+ T cells may miss recently activated cells, and we examined whether it was possible to complement MHC class I multimers with activation markers. To identify potential activation markers on the gene expression level, antigen-specific CD8+ T cells were isolated from carboxyfluorescein succinimidyl ester (CFSE)-labeled peripheral blood mononuclear cells (PBMCs) via Flu MP_{58-66} human leukocyte antigen (HLA) class I A*0201 multimers and directly flow sorted to unlabeled antigen-presenting cells (K562/A*0201 cells or autologous PBMCs; schematic in **Figure 1B**). Cells were stimulated overnight with the cognate peptide, mock peptide, or solvent, and CFSE-labeled cells were flow sorted as single cells (**Figure 1C**) and the gene expression of 75 selected genes analyzed. Consistent with the impaired detection of antigen-responsive cells via multimers when cells had been exposed to their cognate peptide, the multimer fluorescence intensity of dye-labeled cells incubated with their cognate antigen was lower than that of cells incubated with solvent or mock peptide (**Figure 1C**, top). Cells stimulated with the cognate peptide had distinct gene expression profiles compared with the control stimulated cells for each of three donors and either K562/A*0201 or PBMCs as the antigen-presenting cells (T-distributed stochastic neighbor embedding [t-SNE] plots, **Figure 1C**). As expected, the genes encoding known activation markers [CD137 (*TNFRSF9*), CD69, and CD25] and effector molecules (*IFNG*, *GZMB*, and *FASLG*) were upregulated in cells incubated with their cognate peptide Flu MP_{58-66} (**Figure 1C**; **Figure S2**) compared with cells incubated with mock peptide or solvent. Among the marker genes that best discriminated the cognate peptide-activated cells, upregulation of *TNFRSF9*, *REL*, *EGR2*, *SRP14*, *FASLG*, *GZMB*, *IFNG*, *CD69*, *HMGB1*, and *NAB2* and downregulation of *IL7R* were consistently detected in each of the donors (**Table S1**). Notably, the expression of some genes, including *IFNG*, was extremely heterogeneous between the responsive cells. Only minor differences were observed when K562/A*0201 cells were used as antigen-presenting cells compared with PBMCs (**Figure S2**), and the cognate peptide stimulated Flu MP_{58-66}-specific CD8+ T cells clustered together in the t-SNE analysis of the gene expression regardless of the type of antigen-presenting cell (**Figure 1C**). We also found that the profiles of a small number of cognate peptide-stimulated cells were similar to those of the mock peptide- and solvent-stimulated cells (e.g., donor #3), potentially reflecting heterogeneity among the pool of antigen-directed cells.

The findings using the target qPCR expression were extended using single-cell RNA sequencing (scRNAseq) of cells from donor #1 that were processed and isolated in the same manner. The TCR sequences of the Flu MP_{58-66} multimer-positive cells were dominated by *TRAJ42*01* (66.5%) and *TRBV19*01* (95.7%), as previously described for Flu MP_{58-66}-directed TCRs (4, 12–14) (**Table S2**). Again, the cognate peptide-stimulated cells were distinct from the mock peptide- and solvent-stimulated cells, and only a minority of cognate peptide-stimulated cells had gene expression profiles that were not distinguishable from the mock peptide- and solvent-stimulated cells (**Figure 1D**). The TCR sequences of these non-responsive cells included TCRs that were also detected in responsive cells, which suggests that these were unresponsive Flu-specific CD8+ T cells rather than an isolation artifact. In total, 2360 genes were differentially expressed (adjusted $p < 0.05$) upon stimulation with the cognate peptide relative to mock peptide or solvent stimulation (**Figure S3**; **Table S3**), of which 1940 had a >2 $\log_2$-fold change. Of these, 590 genes were differentially expressed (adjusted $p < 0.05$) in both K562/A*0201 cell- and PBMC-based peptide presentation. The top 50 differentially expressed genes are shown in **Figure 1E**. They include *TNFRSF9*, *CCL5*, *CCL4*, *EGR2*, *GZMB*, *IFNG*, *IL2RA*, and *IL7R*, which were also observed using the targeted gene panel. They also comprised

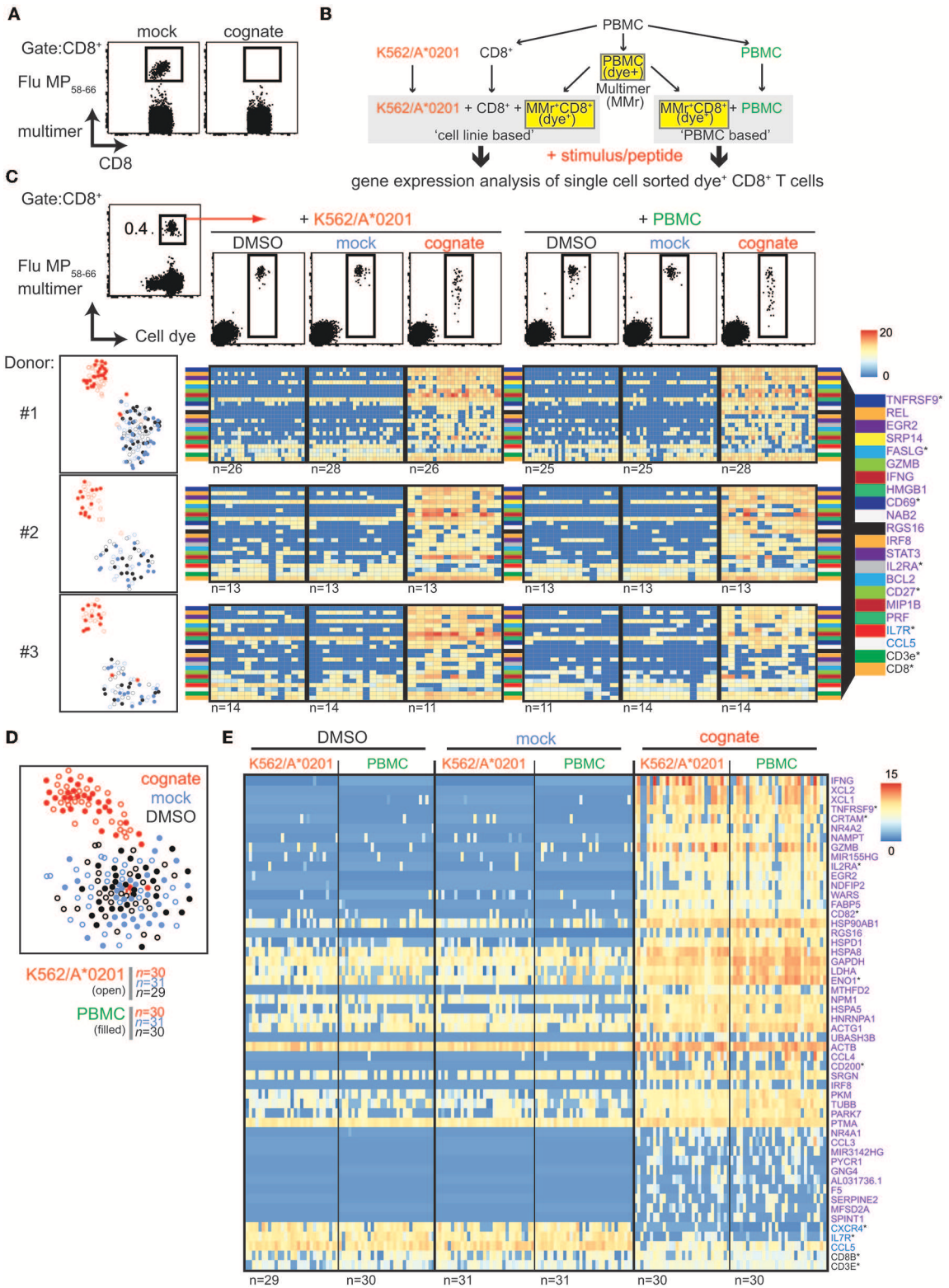

FIGURE 1 | Single-cell gene expression analysis of sorted Flu MP_{58-66}-responsive $CD8^+$ T cells. **(A)** Representative flow cytometry dot plots of PBMCs stained with HLA-A2 multimers loaded with Flu MP_{58-66} after incubating PBMCs for 20 h in the presence of mock ($IGRP_{265-273}$) or cognate (Flu MP_{58-66}) peptide. Plots show 5 × 10^4 cells in the CD8 gate. **(B)** Schematic work-flow of the dye-based activation assay. **(C)** *Top*: representative dot plots (donor #1) of multimer and cellular dye-stained

(Continued)

FIGURE 1 | PBMCs (left). Cells in the CD8 gate are shown. CD8 cells staining positive for the multimer and cell dye were sorted (red arrow) for use in assays using the K562/A*0201 cell line or autologous PBMCs for antigen presentation. After incubation with control stimuli (peptide solvent and mock peptide) or the cognate peptide for 20 h, $CD8^+$ T cells staining positive for the cell dye were sorted for single-cell targeted gene expression analysis. *Lower left*: t-SNE analysis for donors #1–3. Gene expression was analyzed in single cells following incubation with an antigen-presenting cell line (open circles) or PBMCs (filled circles) in the presence of solvent (black) or mock peptide (blue) as control stimuli or with the cognate peptide (red). *Lower right*: Heatmaps of the top 20 ranked differentially expressed genes in cognate peptide-stimulated cells relative to control-stimulated cells for donors #1–3. The numbers of analyzed cells are shown below the individual heatmaps. Genes marked with an asterisk encode proteins expressed on the cell surface. **(D)** t-SNE of scRNAseq gene expression data for Flu MP_{58-66}-directed $CD8^+$ T cells derived from donor #1 and incubated with K562/A*0201 (open) or autologous PBMCs (filled) in the presence of solvent (black) or mock peptide (blue) as control stimuli or with cognate peptide (red). **(E)** Heatmaps for the top 50 ranked differentially expressed genes in antigen-directed $CD8^+$ T cells incubated with cognate peptide relative to control stimuli (purple: upregulated genes; blue: downregulated genes; CD8B and CD3E are also shown). Data are shown for donor #1 following incubation with K562/A*0201 or autologous PBMCs. Genes were combined for the ranking. The numbers of analyzed cells are shown below the heatmaps.

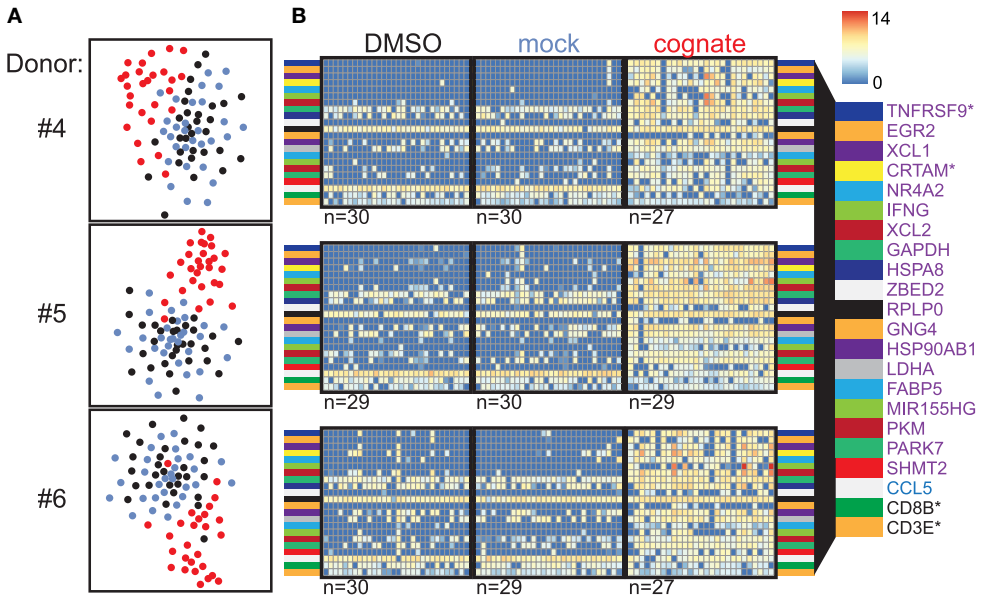

FIGURE 2 | Verification of marker genes in $CMVpp65_{495-503}$-responsive $CD8^+$ T cells. **(A)** t-SNE analysis of scRNAseq data for $CMVpp65_{495-503}$-directed $CD8^+$ T cells from donors #4–6 following incubation with autologous PBMCs in the presence of solvent (black) or mock peptide (blue) as control stimuli or with the cognate peptide (red). **(B)** Heatmaps of the top 20 ranked differentially expressed genes in CMV $pp65_{495-503}$-specific $CD8^+$ T cells incubated with cognate peptide relative to control stimuli for donors #4–6 (purple: upregulated genes; blue: downregulated genes; CD8B and CD3E are also shown). The numbers of analyzed cells are shown below the heatmaps. Genes marked with an asterisk encode proteins expressed on the cell surface.

genes encoding the cytokines X-C motif chemokine ligand (XCL)1 and XCL2 or cytotoxic and regulatory T cell molecule (CRTAM) that are involved in attraction (15) and adhesion to cross-presenting dendritic cells (16), genes important for metabolic function [e.g., *GAPDH*, *FABP5* (17)], and several other genes involved in protein synthesis supporting cell activation and expansion.

Verification of the Identified Marker Genes in $CMVpp65_{495-503}$-Responsive $CD8^+$ T Cells

To validate our findings obtained using Flu MP_{58-66}-specific $CD8^+$ T cells, we performed similar experiments using $CD8^+$ T cells specific to the dominant human cytomegalovirus (hCMV) structural protein pp65 (CMV $pp65_{495-503}$). CFSE-labeled multimer-isolated CMV-specific $CD8^+$ T cells were incubated with PBMCs loaded with CMV $pp65_{495-503}$ peptide or control antigen, and the CFSE-labeled cells were subsequently sorted for scRNAseq. The TCR repertoire of the isolated single cells resembled that expected for $CMVpp65_{495-503}$-specific $CD8^+$ T cells, and included the previously described enriched combinations of *TRAJ49*/*TRAV24* (donors #4–#6), *TRAJ50*/*TRAV35* (donors #4 and #5), *TRBV6-5*/*TRBJ1-2* (donors #4 and #5), and *TRBV12-4*/*TRBJ1-2* (donors #4 and #6) [(5); **Table S2**]. Again, cells stimulated with the cognate peptide were separated from the control-stimulated cells based on their gene expression profiles (**Figure 2A**), and the genes *TNFRSF9* (*CD137*), *XCL1*, *CRTAM*, *EGR2*, and *XCL2* were ranked highly as differentially expressed genes in all three donors (**Figure 2B**). The expression of 2067 genes was increased ($n = 1471$) or decreased ($n = 596$) in cells stimulated with the cognate peptide in all three donors (**Figure S4**; **Table S4**; **Figure 3A**).

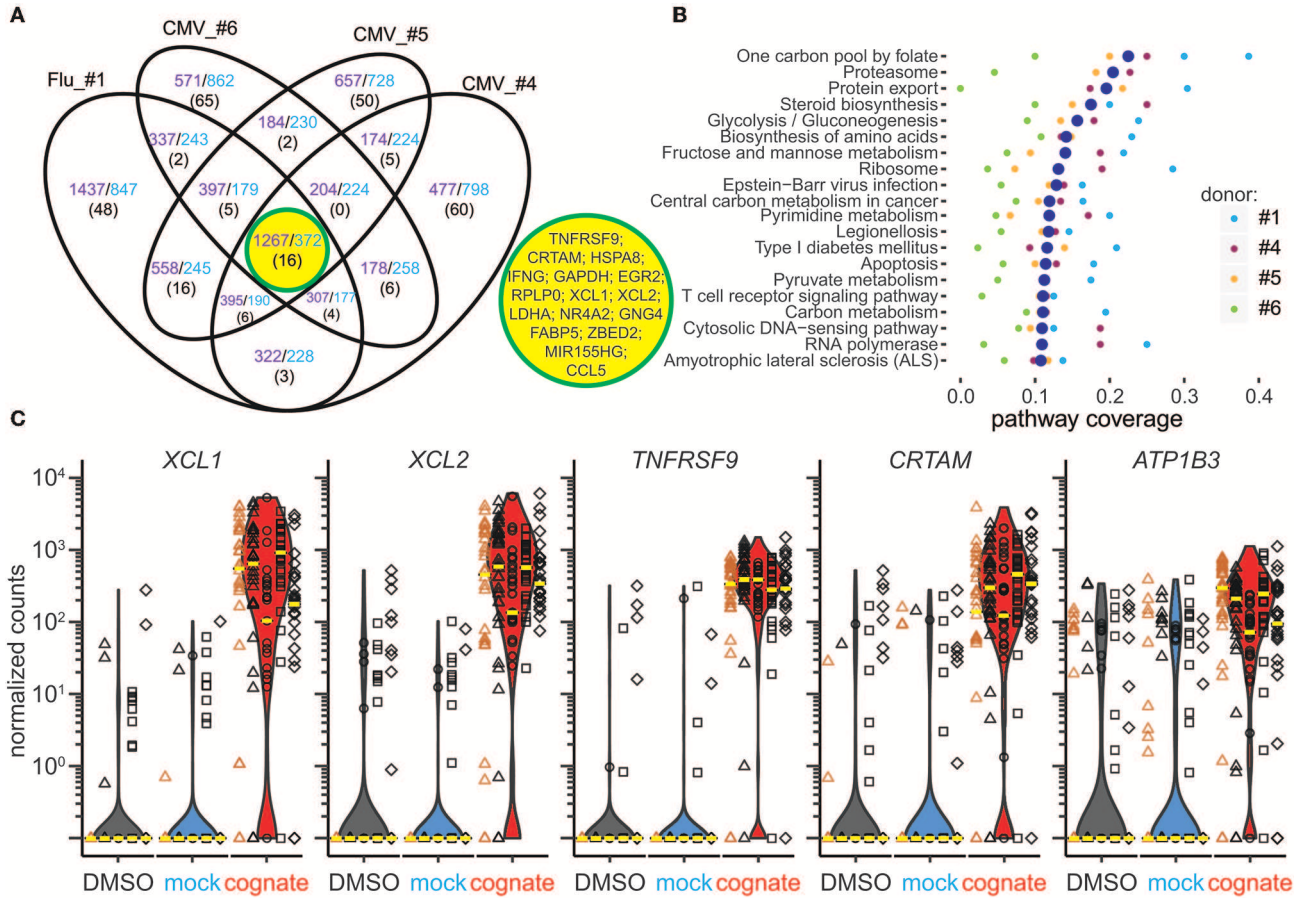

FIGURE 3 | Combined gene and pathway analysis of Flu MP_{58-66}- and $CMVpp65_{495-503}$-responsive $CD8^+$ T cells. **(A)** Venn diagram of overlapping genes upregulated (purple) or downregulated (blue) in Flu MP_{58-66}-directed (donor #1) or $CMVpp65_{495-503}$-directed $CD8^+$ T cells (donors #4–6) following stimulation with the cognate peptide relative to control stimuli. The number of genes belonging to the top 100 differentially expressed genes in each individual is shown in parenthesis. Sixteen genes were shared between all individuals, as shown in the central yellow circle. **(B)** Top 20 KEGG pathways identified based on the differentially expressed genes in individual donors following stimulation with cognate peptide relative to control stimuli. The pathways showing significant enrichment for these genes in at least one donor were included in the analysis and were ranked based on the median pathway coverage (large blue dots) for all donors. **(C)** Scatter/violin plots showing normalized counts of the five genes that best separated the cognate peptide-stimulated and control-stimulated cells in antigen-directed $CD8^+$ T cells incubated with cognate peptide (red violins) relative to solvent (black) or mock peptide (blue) in the presence of autologous PBMCs (black symbols) or K562/A*0201 (orange symbols). The *y*-axis represents the normalized read counts for individual Flu MP_{58-66}-specific cells from donor #1 (triangles) or $CMVpp65_{495-503}$-specific cells from donors #4 (circles), #5 (squares), and #6 (diamonds). The median values for individual donors and conditions are shown as yellow lines. Similar cell numbers were analyzed per condition for each donor, and data comprise counts for 148 cells incubated with DMSO, 152 cells incubated with mock peptide, and 143 cells stimulated with the cognate peptide.

Identification of Candidate Marker Genes Common to Flu MP_{58-66}-Responsive and $CMVpp65_{495-503}$-Responsive $CD8^+$ T Cells

A total of 1639 genes were up- or downregulated in both the Flu MP_{58-66} -responsive (donor #1) and $CMVpp65_{495-503}$-responsive (donors #4, #5, and #6) $CD8^+$ T cells (**Figure 3A**). We examined the top 100 differentially expressed genes in each of the donors and identified 16 that were shared between the CMV- and Flu-responsive cells. These included *TNFRSF9*, *CRTAM*, *XCL1*, *XCL2*, the transcription factors *EGR2* and *ZBED2,* the transcriptional regulator *NR4A2*, genes involved in metabolism (*GAPDH*, *LDHA*, and *FABP5*), the host gene for microRNA155 (*MIR155HG*), *RPLP0*, *HSPA8*, *GNG4*, *CCL5,* and *IFNG*.

Pathway analysis of the differentially expressed genes in antigen-responsive cells across donors and antigen specificities repeatedly demonstrated the enrichment of genes associated with cell metabolism, infection, amino acid biosynthesis, apoptosis, TCR, and cytokine signaling, protein processing, and antigen presentation (**Figure 3B**).

We reasoned that markers that may be suitable for identifying activated antigen-responsive $CD8^+$ T cells would be exclusively or most differentially expressed in either the cognate peptide-stimulated cells or the control mock peptide- and solvent-stimulated cells. Therefore, we ranked the genes

according to their ability to separate these two groups (**Table S5**). The top 10 ranked genes were *TNFRSF9*, *FABP5*, *CRTAM*, *EGR2*, *NR4A2*, *XCL1*, *XCL2*, *ATP1B3*, *RBPJ,* and *MIR155HG*. Among the top 50 separating genes, *XCL1*, *XCL2*, *TNFRSF9*, *CRTAM*, and *ATP1B3* showed the greatest differences in median expression levels between the cognate peptide- and control-stimulated cells (**Figure 3C**). The ability to separate the two groups and the large differences in expression support the use of these genes as potential transcriptional markers for antigen-responsive $CD8^+$ T cells.

We next assessed whether the top-ranking genes for separating the groups are also suitable for a wide range of peptide concentrations used for stimulation (**Figure S5**). Even with a 1,000-fold reduction in the peptide concentration, the expression of these genes differed markedly between mock-stimulated cells and cognate peptide-stimulated cells, suggesting that they mark antigen-responsive cells over a broad range of peptide concentrations.

CD137, CD82, and CD355 Surface Proteins Identify Antigen-Responsive Cells

Several of the top-ranked differentially expressed genes encode cell-surface proteins. One of these, CD137 (encoded by *TNFRSF9*), is a known cell-surface marker of activated $CD8^+$ T cells (18, 19). In accordance with our observations using Flu MP_{58-66}-specific cells, the detection of CMV $pp65_{495-503}$-directed cells via MHC class I multimers was impaired when the cells had been incubated with their cognate antigen. However, there was an increase in the frequency of CD137-stained $CD8^+$ T cells (**Figure 4A**; **Figure S6**). We compared the TCR repertoires between CMV multimer-sorted $CD8^+$ T cells from PBMCs stimulated with the control peptide (denoted antigen-specific $CD8^+$ T cells) and CD137-positive $CD8^+$ T cells from PBMCs stimulated with CMV $pp65_{495-503}$ (denoted antigen-responsive cells). There was extensive overlap between the two repertoires (**Figure 4B**; **Table S2**) indicating that both groups of cells display the same TCR specificity. As expected, there were marked differences in the gene expression profiles between antigen-specific and antigen-responsive cells, which included the expression of the key marker genes described above (e.g., *TNFRSF9*, *XCL1*, *XCL2*, and *CRTAM*; **Figure 4C**). The differentially expressed genes broadly overlapped with those identified in cells stimulated with cognate antigen in the dye-based assay (**Figure 4D**; **Table S6**), suggesting that multimer binding did not exert a major influence on the gene expression profiles in our dye-based $CD8^+$ T cell activation assays. Additionally, the gene expression profiles of multimer-sorted antigen-specific cells resembled those of multimer-sorted CFSE-labeled $CD8^+$ T cells incubated with control stimuli in previous assays (**Figure 4D**).

Other genes encoding cell-surface proteins were also found among the genes that were differentially expressed in cells responding to their cognate antigen as compared with control-stimulated cells. We used antibodies directed against nine of these potential surface markers (CD72, CD82, CD134, CD160, CD200, CD319, CD357, CD355/CRTAM, and SEMA4A) and compared their surface protein expression levels in multimer-positive antigen-specific (mock peptide-exposed) and $CD137^+$ antigen-responsive $CD8^+$ T cells (cognate peptide-stimulated). In addition to CD137 and CD25, the expression levels of both CD82 and CD355 were consistently increased in antigen-responsive cells (**Figure 4E**) for all six donors and both antigen specificities tested.

Marker-Based Identification of Antigen-Responsive $CD8^+$ T Cells Within Unselected Cells

To validate the identified gene expression markers of antigen-responsive cells further, we bulk-sorted $CD95^+CD8^+$ memory T cells from donor #4 after stimulating the cells with control (Flu MP_{58-66}) or CMV $pp65_{495-503}$ peptide, and we performed single-cell gene expression analysis and paired TCRα/β sequencing for over 4,000 cells in each condition. Using the TCR information retrieved from CMV-specific cells in previous scRNAseq experiments of donor #4 (**Table S2**), we were able to trace the CMV-directed cells within the bulk population. Among control peptide-stimulated cells, the CMV-directed cells were not easily distinguishable from other $CD8^+$ T cells based on their gene expression profiles (**Figure 5A**). However, when stimulated with their cognate peptide, the CMV TCR-bearing cells in the CMV stimulation conditions almost exclusively clustered together and were clearly separated from the thousands of other $CD8^+$ T cells based on their gene expression profiles. This cluster of cells was characterized by upregulation of the previously identified genes, including *XCL1*, *XCL2*, *TNFRSF9*, and *CRTAM* (**Figure 5B**).

Next, we used a machine learning algorithm (support-vector machine, SVM) to distinguish antigen-responsive from other cells using the top 10 marker genes identified in the cluster of antigen-responsive cells. The algorithm was trained using data obtained with the CMV peptide-stimulated $CD95^+CD8^+$ memory T cells to identify a hyperplane between antigen-responsive and non-responsive cells. We applied the algorithm to the scRNAseq datasets for donors #1–4. The algorithm correctly distinguished the cognate peptide-stimulated cells from the control-stimulated cells with high sensitivity (mean sensitivity 89.5%; range 86.7–96.6%) and specificity (mean 98.3%; range 94.9–100%) (**Figure 5D**). We also tested the algorithm on the Flu MP_{58-66} peptide-stimulated $CD95^+CD8^+$ memory T cells from donor #4. Using a total of 4,419 cells, the algorithm classified eight as antigen-responsive (**Figure 5C**). All eight cells were in close proximity within the t-SNE analysis and where the cluster of responsive cells was expected. Paired TCR information was available for five of the eight cells in this dataset and included the previously described public Flu MP_{58-66}-restricted TCR sequences (**Table S7**), suggesting that the algorithm can be used to identify the relevant antigen-specific and antigen-responsive $CD8^+$ T cells.

Using this independent approach, we confirmed that the identified genes are potent marker genes. Additionally, we showed that it was possible to identify antigen-responsive cells within a bulk of memory $CD8^+$ T cells, independently of MHC class I multimers.

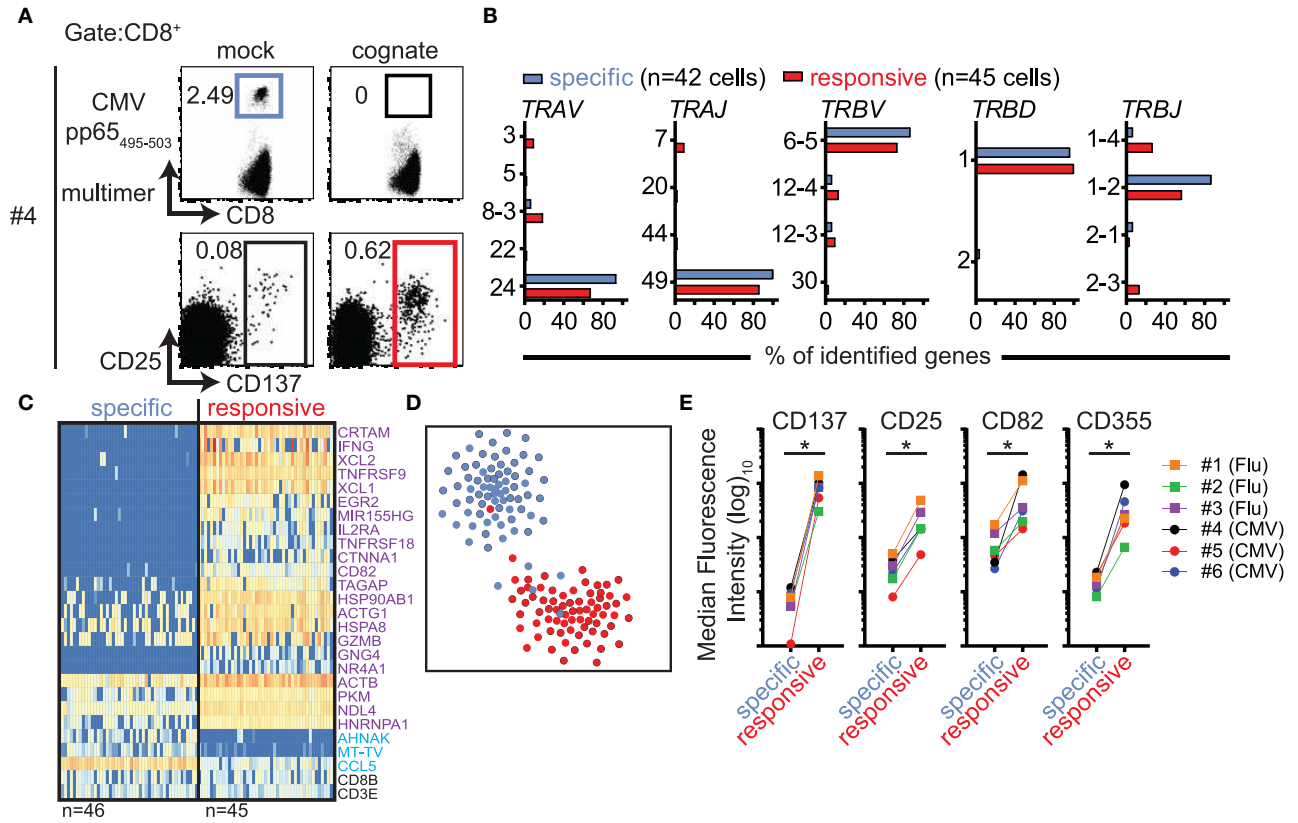

FIGURE 4 | CD137 expression marks antigen-responsive CD8+ T cells in the absence of multimer staining. **(A)** Representative flow cytometry dot plots of PBMCs from donor #4 stained with HLA-A2 multimers loaded with CMVpp65$_{495-503}$ peptide after incubating PBMCs for 20 h with mock (Flu MP$_{58-66}$) or cognate (CMVpp65$_{495-503}$) peptides (top) and corresponding plots showing CD25 and CD137 expression (bottom). All plots show 5 × 10^4 cells in the CD8 gate. **(B)** TCR repertoire analysis of CD8+ T cells subjected to single-cell sorting based on positive staining with CMVpp65$_{495-503}$ peptide-loaded HLA-A2 multimers from PBMCs stimulated with mock peptide (blue gate in **A**; antigen-specific) or based on CD137 expression after incubation with cognate peptide (representing cells in the red gate in **A**; antigen-responsive). The frequencies of genes used for the production of TCR α and β chains in each of the two populations of cells are shown. **(C)** Heatmaps of the top 25 ranked differentially expressed genes in CMV pp65$_{495-503}$-specific relative to CMV pp65$_{495-503}$-responsive CD8+ T cells from donor #4. The numbers of analyzed single cells are shown beneath the heatmaps. **(D)** t-SNE plot of gene expression data for single-cell-sorted antigen-specific (black outline filled blue) and antigen-responsive CD8+ T cells (black outline filled red). Data for CMV-directed cells incubated with mock (blue) or cognate (red) peptide in the dye-based activation assay are included as a reference. **(E)** Scatter plots of flow cytometric analyses comparing the median fluorescence intensities for the indicated surface markers between antigen-specific and antigen-responsive CD8+ T cells for donors #1–3 (donors with Flu-directed cells) and #4–6 (donors with CMV-directed cells).

Autoantigen-Specific CD8+ T Cells May Differ in Their Responsiveness

In addition to CD8+ T cells specific to viral antigens, we extended our study to autoreactive CD8+ T cells directed against the type 1 diabetes autoantigen islet-specific glucose-6-phosphatase catalytic subunit-related protein (IGRP). CD8+ T cells specific for the HLA-A*0201-restricted IGRP$_{265-273}$ peptide were previously discussed as potential effectors of the islet destructive process in patients with type 1 diabetes and in HLA transgenic mice (20–22). Furthermore, IGRP$_{265-273}$-directed CD8+ T cell clones were shown to kill peptide-loaded target cells in an antigen-specific manner (22, 23). We applied our dye-based CD8+ T cell activation assay to IGRP$_{265-273}$-directed CD8+ T cells from a previously described patient with high frequencies of both Flu MP$_{58-66}$- and IGRP$_{265-273}$-directed cells (**Figure 6A**) (23). In contrast to our findings with Flu- and CMV-directed cells, we did not observe a reduction in multimer fluorescence intensity or upregulation of CD137 protein after stimulation with cognate IGRP$_{265-273}$ peptide in the dye-positive or dye-negative cells (**Figure 6B**). As expected, CD137 was observed on the dye-negative CD8+ T cells stimulated with Flu MP$_{58-66}$ peptide (mock), suggesting that the response to the viral peptide was not impaired in this patient. The dye-positive cells displayed a restricted TCR repertoire and the dominant IGRP TCR clonotype previously described for this donor (**Table S2**) (23). IGRP cells stimulated with cognate peptide did not cluster separately from control-stimulated cells (**Figure 6C**), and only 321 genes were differentially expressed between cognate peptide- and control-stimulated cells, of which only six were shared between both types of antigen presentation. None of the top 50 differentially expressed genes were found with K562/A*0201- and PBMC-based antigen stimulation (**Table S8**; **Figure S7**). None of the top

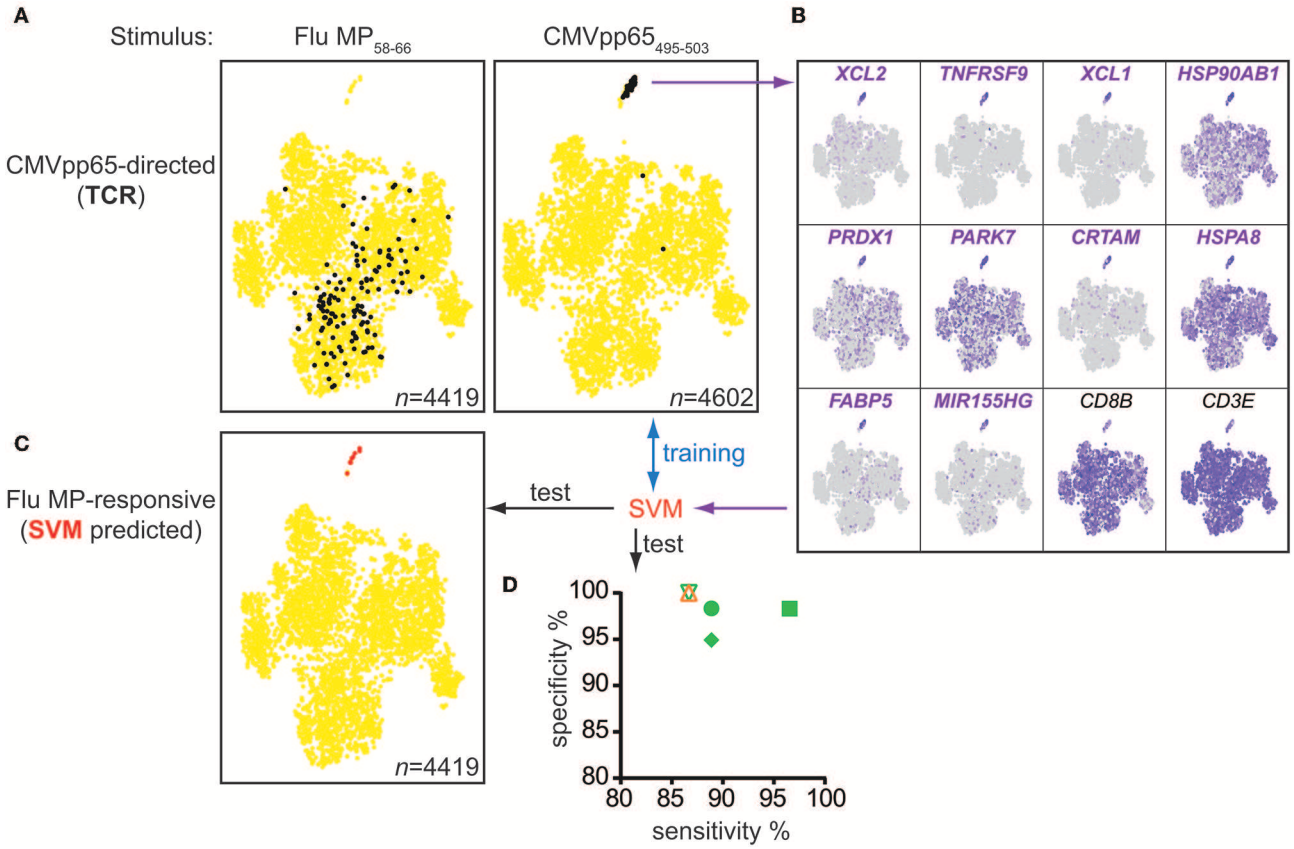

FIGURE 5 | Multimer-independent identification of antigen-responsive cells from bulk memory CD8+ T cells. **(A)** t-SNE plots of gene expression data from CD95+CD8+ memory cells sorted from the PBMCs of donor #4 incubated for 20 h with the indicated peptides. Cells expressing CMVpp65$_{495-503}$-directed TCRs, known from previous CMV multimer-based experiments, are shown in black. **(B)** t-SNE plots for CMV peptide-stimulated cells in **(A)** showing the expression levels of the top 10 genes marking the cluster of antigen-responsive cells and of the reference genes *CD3E* and *CD8B*. The 10 genes were used to train a machine learning algorithm (SVM), which was subsequently tested for its sensitivity and specificity to distinguish antigen-responsive cells from control stimulated-cells in the scRNAseq data sets of donors #4–6 (CMV-directed) and # 1 (Flu-directed) **(D)** and to predict Flu MP$_{58-66}$-responsive cells from bulk memory CD95+CD8+ T cells **(C)**.

16 genes previously identified for virus-responsive CD8+ T cells were differentially expressed between the IGRP cognate peptide- and control-stimulated cells (**Figure 6D**).

Thus, despite their strong binding to IGRP peptide-loaded multimers and distinct TCR repertoire, the antigen responsiveness of these antigen-specific cells is not analogous to that of virus-directed cells when tested in the same stimulatory setting. These findings suggest that certain autoantigen-specific CD8+ T cells show differing responses to their cognate peptide, are anergic, or may need divergent stimulatory conditions to mount a response similar to that of virus-specific CD8+ T cells.

DISCUSSION

The use of multimers to identify and subsequently characterize antigen-specific CD8+ T cells has advanced our knowledge of human TCRs and their phenotype. Here, we demonstrated that multimer-based identification of antigen-specific CD8+ T cells is adversely affected by recent activation, and we identified large gene sets that were congruently up- or downregulated in virus antigen-responsive CD8+ T cells across multiple viruses and donors and under different conditions of antigen presentation. We developed algorithms that accurately identify virus-responsive CD8+ T cells and can be applied broadly to single-cell RNA profiling data from antigen-stimulated cells.

Our approach to defining antigen-responsive profiles included proven models of well-established viral peptides presented by HLA-A*0201. We used two antigen presentation schemes and dominant peptides from two viruses. The antigen-specific cells that were used to establish the profiles were highly specific, as judged by their strong multimer binding and verified by TCR sequencing. Finally, we demonstrated a broad methodological applicability using three RNA profiling methods, and chose genes that identified virtually all of the multimer-selected antigen-responsive cells.

The limitations of this approach include a restricted set of antigenic peptides, which prevents us from generalizing our findings. In particular, the responsive profiles may be inappropriate for non-viral peptides, as we demonstrated for the islet β cell autoantigen IGRP. It is possible that the

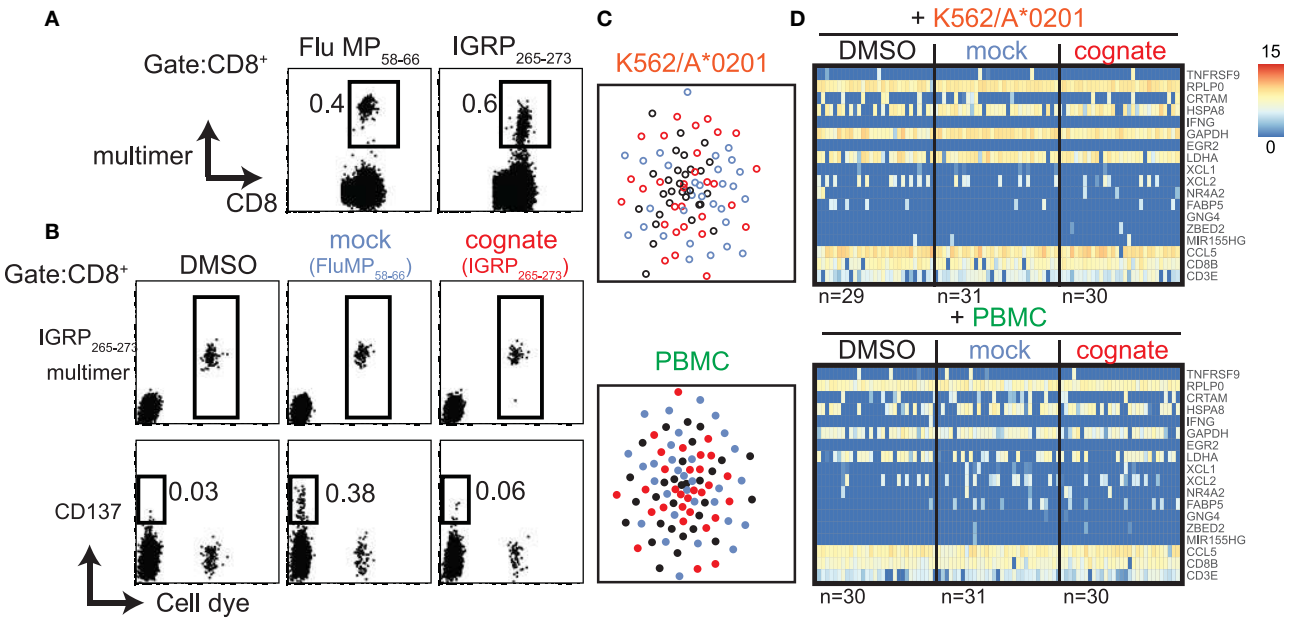

FIGURE 6 | Autoantigen-directed CD8+ T cells may differ in their responsiveness to the cognate peptide. **(A)** Representative flow cytometry dot plots of PBMCs from a donor with type 1 diabetes stained with HLA-A2 multimers loaded with Flu MP_{58-66} or $IGRP_{265-273}$ peptide. **(B)** Representative dot plots of bulk (cell dye-negative) and $IGRP_{265-273}$ directed (cell dye-positive) CD8+ T cells incubated for 20 h with the indicated stimuli using K562/A*0201 cells for antigen presentation in the dye-based CD8+ T cell activation assay. Cells in the CD8 gate are shown. **(C)** t-SNE plots of single-cell gene expression data from $IGRP_{265-273}$-directed CD8+ T cells after incubation for 20 h with the indicated antigen-presenting cells in the presence of solvent (black), mock peptide (blue), or cognate peptide (red). **(D)** Heatmaps of single-cell gene expression data for the previously identified 16 genes of antigen-responsive virus-directed CD8+ T cells derived from $IGRP_{265-273}$-directed CD8+ T cells stimulated with the control or cognate peptide (see **Figure 3A**). The reference genes *CD3E* and *CD8B* are also shown.

multimer-selected cells did not cover all phenotypes of antigen-specific CD8+ T cells, including cells with anergic or exhausted profiles. About 1% of the multimer-sorted virus-specific cells were not activated after cognate peptide presentation, and these cells could represent anergic or other T cell types. Finally, because our findings were based on *in vitro* stimulation, the profiles observed here may not represent those of CD8+ T cells activated *in vivo*.

Peptide stimulation of CD8+ T cells resulted in up- or downregulation of over 1,000 genes. Of particular relevance was the downregulation of TCR on the cell surface, which was also associated with a marked reduction in multimer binding. This is consistent with previous findings (9, 11), suggesting that the frequency of antigen-specific CD8+ T cells may be underestimated in conditions of chronic or recent *in vivo* activation when using multimers for quantification. Notably, we observed a marked reduction in multimer staining, even after stimulation with the cognate peptide at a concentration of 1 ng/mL, which is 10,000 times lower than the concentration typically used for *in vitro* stimulation. A bias against selection of recent or chronically activated CD8+ T cells by multimers may grossly underestimate the degree of CD8+ T cell activation in disease, and could confound recent findings in diseases such as type 1 diabetes (24).

Our objective was to identify a panel of genes and a classifier that would enable us to identify *in vitro*-stimulated, antigen-responsive CD8+ T cells from unselected populations. This requires high specificity because the antigen-responsive cells comprise a minority of cells. High specificity would be provided by genes that are consistently differentially expressed with high quantification in responsive cells. A set of 10 genes, identified by different methods, was used to develop a classifier using machine learning methods. The approach was used to identify antigen-responsive CD8+ T cells from single-cell data, and provides a tool for the future refinement of algorithms suitable for other virus antigen-responsive and other antigen-responsive cells.

At the protein level, we observed increased expression of identified markers CD82 and CRTAM (CD355), similar to the commonly used activation markers CD137 and CD25, in antigen-responsive cells. Our study found many other genes that are known to be transcriptionally increased or that encode proteins that are increased in antigen-responsive CD8+ T cells. These include XCL1 and XCL2 (8, 15, 25, 26), EGR2 (27, 28), NR4A2 (8), and MIR155HG (29, 30), among other proteins (31–33). Some of the identified genes encode proteins that are often used to isolate antigen-specific cells. IFNγ, which is used for cytokine capture of responsive CD8+ T cells (34), was strongly upregulated in the majority of the activated cells. However, it was noticeably absent in a proportion of the antigen-responsive cells, indicating that such isolation methods may miss some responsive cells. We also identified several genes that were differentially expressed in the antigen-responsive CD8+ T cells and, to our knowledge, have not been previously described in this context. These include genes encoding the transcription factor ZBED2,

which regulates transcription by RNA polymerase II, and was exclusively upregulated in many of the antigen-responsive cells; the guanine nucleotide binding plasma membrane associated protein GNG4; HSPA8; and the ribosomal protein RPLP0. These findings may help to understand the processes involved in the T cell-mediated antiviral response and provide new therapeutic targets.

A practical aspect of our study is that we accurately traced the virus-responsive cells identified in the transcriptomic profile back to their TCRs. This allowed us to perform broader antigen screening in combination with large-scale scRNAseq, where the responsive transcriptomic profile is used to map the antigen-specific TCR. An advantage of this approach is that the TCRs are representative of cells that can be robustly activated and permits selection of cells expressing desired effector molecules. This may be useful for improving the epitope selection process in peptide-based vaccination strategies or identifying virus-directed cells for adoptive therapy. Although we have not shown that similar profiles can identify tumor antigen-responsive $CD8^+$ T cells, it is likely that the identification of TCRs using a similar approach may facilitate the development of chimeric antigen receptor T cell-based therapies that are patient-specific. As an example, the profiles could be used in *in vitro* screens to identify TCRs of $CD8^+$ T cells that respond to neo-antigen peptides and could be useful for therapy. Of potential interest, we found no signature that could identify cells responsive to a known epitope of a β cell autoantigen, despite the high frequency of strong multimer-positive memory $CD8^+$ T cells in the donor. We previously reported that IGRP-directed clones generated from the same donor (and expressing the identical TCR) can kill peptide-loaded target cells *in vitro*, albeit with delayed kinetics (23). This suggests the possibility of an antigen-directed response of IGRP-directed cells to the target, at least in certain *in vitro* conditions. It is unclear whether autoantigen-specific $CD8^+$ T cells are generally less responsive than virus-specific cells or if our finding is limited to this antigen or this patient. A previous study showed that autoantigen-specific $CD8^+$ T cells obtained from peripheral blood were less responsive than those obtained from the pancreas (24). Autoantigen-directed cells may be benign in inactive stages of the disease or require additional factors for their activation, which could explain the differences observed between viral and autoantigen-directed responses using our assay setup. An intriguing possibility is that responses to neoantigens or hybrid peptides not normally expressed by self may be more similar to the ones observed for viral peptides and such features be used to distinguish pathogenic $CD8^+$ T cells and their target peptides. Additionally, low frequencies of β-cell directed $CD8^+$ T cells as observed using MHC class I multimers (35) make it difficult to obtain a broader picture of their genetic profiles and their antigen-specific response. The identification and characterization of autoantigen-specific $CD8^+$ T cells may, therefore, require a combination of multimer binding, peptide stimulation as we describe here, and functional confirmation of antigen specificity of the TCR in transduced reporter cell lines (36). Further studies are needed in patients with active autoimmune disease.

In conclusion, we provide findings that will allow the identification of antigen-responsive $CD8^+$ T cells and their TCRs in an unbiased manner. The findings can be applied to antigens where multimer reagents are unavailable, as an alternative to multimer-based cell selection or in combination with multimer selection. This should facilitate the development of $CD8^+$ T cell activation-based therapies.

MATERIALS AND METHODS

Subjects

Samples were obtained from six healthy adult blood donors or from an adult patient with type 1 diabetes; all had the HLA-A*0201 allele. All methods were performed in accordance with relevant guidelines and regulations. Samples were collected after obtaining informed consent and ethical committee approval (EK 240062016, TU Dresden and 5049/11, TU München).

Peptides

The Flu MP_{58-66} (GILGFVFTL), hCMV $pp65_{495-503}$ (NLVPMVATV) and $IGRP_{265-273}$ (VLFGLGFAI) peptides were purchased from Mimotopes or Panatecs at >95% purity, as confirmed by high-performance liquid chromatography and mass spectrometry. For the T cell assays, the peptides were dissolved in dimethyl sulfoxide (DMSO) to a concentration of 10 mg/mL and subsequently diluted in assay medium to final peptide concentrations of ≤10 μg/mL. The DMSO concentration in each assay was ≤0.1%.

Cell Staining and Flow Cytometry

Cells were stained using the following surface marker-directed monoclonal antibody–fluorochrome combinations: CD3-APC (clone HIT3a), CD3-AF488 (clone UCHT1), CD4 (clone SK3), CD8-APC-Cy7 (clone SK1), CD14-APC (clone M5E2), CD19-APC (clone HIB19), CD56-APC (clone B159), CD25-PE-Cy7 (clone M-A251), CD82-AF647 (clone 423524), and CD95-CF594 (clone DX-2) from BD Pharmingen; CD137-PerCP-eFluor710 (clone 4B4-1) and CD355-APC (clone Cr24.1) from Thermo Fisher Scientific; and CD16-APC (clone 3G8) from BioLegend. 7-Aminoactinomycin D (7AAD, BD Pharmingen) or Sytox Blue (Thermo Fisher Scientific) was used to exclude dead cells. Phycoerythrin-labeled HLA-A*0201 multimers loaded with Flu MP_{58-66} (GILGFVFTL), hCMV $pp65_{495-503}$ (NLVPMVATV), or $IGRP_{265-273}$ (VLFGLGFAI) were purchased from Immudex. Cells were acquired on a flow sorter (BD FACSAria Fusion, BD Biosciences) with FACSDiva 8 software (BD Biosciences) and analyzed using FlowJo 10 software (FlowJo LLC). Unless otherwise stated, cells were stained in phosphate-buffered saline (PBS) supplemented with 1% pooled human AB serum (PBS/1%AB) for 30 min on ice followed by at least two washing steps using the same buffer, and were then stained with a marker for dead cells before analysis. For CFSE labeling, PBMCs (4×10^7/mL) were resuspended in PBS at room temperature, mixed 1:1 with CFSE staining solution (1 μM CFSE in PBS), and the cell suspension was immediately vortexed for 10 s to achieve a homogenous CFSE distribution. Cells were incubated at 37°C for 10 min in the dark and washed with X-VIVO15 media

to remove excess CFSE. Multimer staining was carried out in PBS supplemented with 5% pooled AB serum for 10 min at room temperature followed by 25 min incubation on ice in the presence of the respective antibodies. Cells were subsequently washed with PBS/1%AB and incubated with 7AAD or Sytox Blue for 10 min immediately before flow cytometry. The gating strategies used for the analysis and sorting of multimer- or CFSE-stained cells, and for the isolation of antigen-responsive cells and $CD95^+$ $CD8^+$ memory T cells are illustrated in **Figure S8**.

$CD8^+$ T Cell Activation Assay

The activation assays were conducted using 200–500 CFSE-labeled antigen-specific $CD8^+$ T cells together with either 5×10^5 unlabeled PBMCs (PBMC-based assays) or a combination of 2.5×10^5 HLA-A*0201-expressing K562 cells [K562/A*0201 (37); kindly provided by Prof. Thomas Wölfel; Johannes Gutenberg Universität, Mainz, Germany] and 5×10^4 flow-sorted bulk $CD8^+$ T cells (cell line-based assays). To obtain CFSE-labeled multimer-specific $CD8^+$ T cells, the PBMCs were labeled with CFSE and subsequently incubated with the respective peptide-loaded HLA-A*0201 multimers to identify and sort highly pure antigen-specific $CD8^+$ T cells into the assay tubes. Unless stated otherwise, the cells were incubated with the cognate peptide (10 μg/mL), mock peptide (10 μg/mL), or solvent (DMSO) for 18–24 h. Subsequently, viable $CFSE^+$ $CD8^+$ T cells were single-cell-sorted for downstream gene expression analysis and TCR sequencing. For some experiments, antigen-responsive cells were sorted based on their expression of the activation marker CD137. In these cases, the assays were performed without prior CFSE labeling or multimer staining.

For droplet encapsulation sequencing experiments conducted on the 10x Genomics platform, PBMCs were incubated with Flu MP_{58-66} or hCMVpp65$_{495-503}$ peptide for 20 h and 6500 viable $CD95^+CD8^+$ memory $CD8^+$ T cells were sorted from each of these conditions and processed for gene expression and TCR analysis.

Targeted Single-Cell Gene Expression Analysis

Targeted gene expression analysis was done by single-cell qPCR as previously described (23) with some modifications. cDNA was synthesized using Quanta qScript TM cDNA Supermix directly on cells. Total cDNA was pre-amplified for 20 cycles ($1 \times 95°C$ for 8 min, 95°C for 45 s, 49°C with 0.3°C increment/cycle for 1 min, and 72°C for 1.5 min) and $1 \times 72°C$ for 7 min with TATAA GrandMaster Mix (TATAA Biocenter) in a final volume of 35 μL in the presence of the 75 primer pairs listed in **Table S9**. The final concentration of each primer was 25 nM. Pre-amplified cDNA (10 μL) was then treated with 1.2 U exonuclease I and expression was quantified by real-time PCR on a BioMarkTM HD System (Fluidigm Corporation) using the 96.96 Dynamic Array IFC, GE 96×96 Fast PCR+ Melt protocol, and SsoFast EvaGreen Supermix with Low ROX (Bio-Rad). The primer concentration was 5 μM in each assay.

Single-Cell RNA Sequencing

The scRNAseq workflow was based on the previously described Smart-seq2 protocol (38) with the following modifications. Single cells were flow-sorted into 96-well plates containing 2 μL of nuclease-free water with 0.2% Triton X-100 and 4 U murine RNase inhibitor (NEB), centrifuged, and frozen at −80°C. After thawing, 2 μL of the primer mix [5 mM dNTP (Invitrogen), 0.5 μM oligo-dT primer, 4 U murine RNase inhibitor] was added to each well. The reverse transcription reaction was performed as described (38), but with final concentrations of RNase inhibitor and Superscript II of 9 U and 90 U, respectively, at 42°C for 90 min, followed by an inactivation step at 70°C for 15 min. The number of pre-amplification PCR cycles was increased to 22 to ensure there was sufficient cDNA for downstream analysis. The amplified cDNA was purified using 1 × Sera-Mag SpeedBeads (GE Healthcare), and DNA was eluted in 12 μL nuclease-free water. The concentration of samples was measured using a plate reader (Infinite 200 PRO, Tecan) in 384 well black, flat-bottom, low-volume plates (Corning) using an AccuBlue Broad Range kit (Biotium). Then, 0.7 ng of pre-amplified cDNA was used for library preparation (Nextera DNA library preparation kit, Illumina) in a 5-μL reaction volume. Illumina indices were added during the PCR reaction [72°C for 3 min, 98°C for 30 s, 12 cycles of (98°C for 10 s, 63°C for 20 s, and 72°C for 1 min), and 72°C 5 min] with 1× KAPA Hifi HotStart Ready Mix and 0.7 μM of dual indexing primers. After PCR, the libraries were quantified with AccuBlue Broad Range kit, pooled in equimolar amounts, and purified twice with 1× Sera-Mag SpeedBeads. The libraries were sequenced on the NextSeq 500 Illumina platform to obtain 75 bp single-end reads aiming at an average sequencing depth of 0.5 million reads per cell.

Droplet Encapsulation Sequencing

The scRNAseq was performed on the 10× Genomics platform using Chromium Single Cell Immune profiling 5' v2 Reagent Kits according to the manufacturer's instructions. Gene expression libraries were sequenced on the NextSeq 500 Illumina platform using a high-output flow cell to obtain paired-end reads at the following read lengths to generate ∼160–180 million fragments: read 1, 26 cycles, i7 index, 8 cycles; read 2, 57 cycles. Similarly, TCR-enriched libraries were sequenced on the MiSeq Illumina platform to obtain 150-bp paired-end reads at a depth of 15 million total read pairs.

Data Processing and Analysis

Targeted Single-Cell Gene Expression Analysis

Raw data were analyzed using Fluidigm Real-Time PCR analysis software. Preprocessing and data analysis were conducted using KNIME 2.11.2 and RStudio Version 0.99.486. Preprocessing with a linear model to correct for confounding factors was conducted as previously described (39). To model the bimodal gene expression of single cells from T cell clones, the Hurdle model, a semi-continuous modeling framework, was applied to the preprocessed data (40). This allowed us to assess the differential expression profiles with respect to the frequency of expression and the positive expression mean via a likelihood ratio test. Genes were ranked by a score derived from the number

of comparisons with statistically significant differences in gene expression and the number of donors with statistically significant comparisons (0.5 points per significant comparison multiplied by the number of donors showing significant differences in expression of each gene). Genes with the same score were ranked based on the delta of their median expression differences between the cognate and control stimuli.

Single-Cell RNA-Sequencing

The scRNAseq adapter trimmed reads were mapped against the human hg38 reference genome using STAR (v2.5.4, default parameters) and Ensembl genes (v 91) as the gene model references for alignment to produce one BAM file per cell. The count per gene matrix was obtained using featureCounts (v1.6.2, *-Q 30, -p*). Differentially expressed genes were identified by performing pairwise comparison using two independent methods: SCDE (v2.8.0) (41) and DESeq2 (v1.20.0, using a zero-inflated negative binomial distribution) (42). Low-quality cells expressing only a few genes were filtered out from the counts matrix using the *clean.counts* function in SCDE (*min.lib.size* = 1,000, *min.reads* = 1, *min.detected* = 1). Only genes that were identified as differentially expressed by both tools were considered to be expressed differentially. Data from each experiment were visualized using the t-SNE algorithm implemented in the Rtsne R package. In all comparisons, the cells stimulated with cognate peptide were labeled as the test group, while cells incubated with solvent or mock peptide showed similar profiles, and thus were combined as a control group. Because the SCDE method does not output a *p*-value for each differentially expressed gene, a two-sided *p*-value for each gene was calculated from the Benjamini–Hochberg multiple testing corrected Z score (cZ) using the normal distribution as the null hypothesis. Genes were ranked by the product of their $\log_2$-fold change and the reciprocal of the adjusted *p*-value (values obtained from SCDE). For experiments using several donors, the genes were ranked individually for each donor and a composite list was generated by obtaining a mean rank for each gene across the multiple donors and then ranking the genes based on their mean rank. To identify genes that could distinguish cognate peptide-stimulated cells and control-stimulated cells in individual experiments, we calculated the separation index (SI) for each gene (g) using the following equations.

$$P\left(g, tg\right) = N(g)/N(tg), \tag{1}$$

where *tg* is the test group, and *N* is the number of cells. To further clarify, *N(g)* is the number of cells with a non-zero count for the particular gene, and *N(tg)* is a total number of cells in the test group. The same applies to the other groups analyzed.

$$P(g, sg) = N(g)/N(sg), \tag{2}$$

where *sg* is the solvent group.

$$P\left(g, mg\right) = N(g)/N(mg), \tag{3}$$

where *mg* is the mock group.

$$P\left(g, cg\right) = (P\left(g, sg\right) + P\left(g, mg\right))/2, \tag{4}$$

where *cg* is the control group.

$$SI\left(g\right) = \left|P\left(g, tg\right) - \overline{P}(g, cg)\right|. \tag{5}$$

The genes were ranked according to their SI in each experiment in descending order, and then ordered by the mean rank of all PBMC-based assays. For the top 50 genes, we determined the median expression difference between cognate peptide- and control-stimulated cells to identify genes that were potentially usable as marker genes.

The Enrichr package was used to determine whether the differentially expressed genes were preferentially enriched in a particular Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway. To identify potential pathways, we performed pathway enrichment analysis using the following workflow. First, all pathways that showed enrichment (adjusted $p < 0.05$) in at least one of the four donors (three CMV and one Flu) were listed together to generate a unique list of these pathways. Then, for each of these pathways, the coverage index was calculated for each donor as the number of differentially expressed genes belonging to a pathway divided by the total number of genes in this pathway. Finally, the pathways were ranked by the median coverage index for each pathway across the four donors.

Droplet Encapsulation Sequencing

The raw gene sequencing data were processed using the *count* command in Cell Ranger software (v2.1.0, *–expect-cells=3,000*) (10x Genomics). Gene annotation was filtered using the *mkgtf* command to include only protein-coding, lincRNA, and antisense gene features (*–attribute=gene_biotype:protein_coding, –attribute=gene_biotype:lincRNA, –attribute=gene_biotype:antisense*). The Seurat package (43) was used to analyze the gene expression data obtained in droplet encapsulation sequencing experiments. Briefly, all genes that were expressed in at least three cells and all cells that expressed ≥200 such genes were retained. The cells were then clustered based on the first 14 principal components (resolution parameter set to 0.6) and visualized using the resulting t-SNE (yielding 8 clusters) and combined with the TCR information retrieved from TCR-enriched libraries of the same experiment performed in parallel. CMVpp65$_{495-503}$-directed cells were identified within the CD95$^+$CD8$^+$ memory T cell single-cell data via their TCRα and β chain sequences (using TCR information of CMVpp65$_{495-503}$-directed cells retrieved in the preceding multimer-based scRNAseq experiments of the donor). To identify marker genes that separated the main cluster containing CMV-directed cells after stimulation with the cognate peptide from the remaining clusters, we used the *FindMarkers* function (test parameter set to *roc*). The top 10 marker genes for this main cluster were used to build a support-vector machine (SVM)-based classifier (https://github.com/bonifaciolab/classifier_AntigenResponsiveCD8TCells) to discriminate between responsive cells and non-responsive cells. For this purpose, we used the SVM implementation available with the package libSVM (v1.04). The log-normalized counts for the 10 marker genes were exported to a file, and responsive cells were labeled as true positives, whereas all other cells were labeled as true

negatives. This dataset was used to build the model, and the data obtained from the experiment with Flu MP_{58-66}-stimulated bulk-sorted $CD95^+CD8^+$ memory T cells was used as the test set. The model was also used to identify responsive cells obtained from other scRNAseq-based experiments. For this, the experiment read counts were normalized using the DESeq2 *counts* function (with normalized = TRUE parameter), and then used as test sets.

TCR Identification

TCR sequences were reconstructed from the transcriptome of each cell using the TraCeR algorithm (v0.6.0, default parameters) (44). Reads originating from the mRNA of the TCRs were extracted from the sequence data for each cell and were assembled into contigs to generate full-length TCRs. Only productive TCR chains, both α and β, were retained for each cell. To resolve instances where one cell had more than one productive α or β chain, the chain with greater expression (obtained by parsing the transcripts per million count) was retained as the productive chain for the respective cell. Sequences of productive chains were then uploaded to IMGT/HighV-QUEST (45) to identify the V, D, and J genes as well as the CDR3 region associated with each chain. The TCR sequences of the droplet encapsulation experiments were extracted using Cell Ranger software.

Additional Statistical Analyses

Wilcoxon's matched-pairs signed-rank test was used to compare median fluorescence intensities for the analyzed surface markers among antigen-specific and antigen-responsive $CD8^+$ T cells.

Software: FACS Diva 8, FlowJo 10, Graph Pad Prism 7, Tracer (v0.6.0), STAR (v2.5.4) libSVM (v1.04), featureCounts (v1.6.2), and R (v3.5.1).

R-packages: DESeq2 (v1.20.0), SCDE (v2.8.0), EnrichR (v1.0), Seurat (v2.3.4), Rtsne (v0.13), pheatmap (v1.0.10), and ggplot2 (v3.1.0).

DATA AVAILABILITY STATEMENT

Count matrices generated from raw single-cell data are available as excel files at: https://github.com/bonifaciolab/classifier_AntigenResponsiveCD8TCells.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethikkommission der TU Dresden; Ethikkommission der TU München. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YF, VS, and EB designed the studies. YF, GK, RM, SR, AL, and DL performed the experiments. YF, AE, AD, and EB supervised the work. VS, YF, AP, and AE analyzed the data. YF, VS, GK, and EB wrote and edited the manuscript.

FUNDING

This work was supported by JDRF fellowship 3-PDF-2016-188-A-N (to YF) and by grants from the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD e.V.). EB was supported by the DFG Research Center and Cluster of Excellence - Center for Regenerative Therapies Dresden (FZT 111) and BMBF/DLR with FKZ 01KU1502B (RESET AID).

ACKNOWLEDGMENTS

We thank the CMCB flow cytometry core facility, particularly A. Gompf and K. Bernhardt for technical support, and L. Ventola for administrative support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2019.02568/full#supplementary-material

Figure S1 | Impaired MHC class I multimer detection of CD8+ T cells responding to their cognate antigen.

Figure S2 | Gene expression analysis of Flu MP58-66-directed CD8+ T cells after incubation with peptide-loaded K562/A*0201 cells or autologous PBMCs.

Figure S3 | Comparison of gene expression between cognate peptide-stimulated and control-stimulated Flu MP58-66-directed CD8+ T cells.

Figure S4 | Comparison of gene expression between cognate peptide-stimulated and control-stimulated CMV pp65495-503-directed CD8+ T cells.

Figure S5 | Discriminative ability of antigen-responsive cells across a wide range of cognate peptide concentrations.

Figure S6 | Impaired MHC class I multimer detection and upregulation of CD137 in CD8+ T cells responding to their cognate antigen.

Figure S7 | IGRP265-273-directed CD8+ T cells markedly differ from virus-directed cells in their response to cognate antigen.

Figure S8 | Gating strategies for the flow sorting experiments.

Table S1 | Overview of statistical results of the targeted gene expression analysis.

Table S2 | T cell receptor information for the individual donors.

Table S3 | Differentially expressed genes in cognate antigen-responsive CD8+ T cells from donor 1.

Table S4 | Differentially expressed genes in cognate antigen-responsive CD8+ T cells from donors 4–6.

Table S5 | Separator index and ranking of separator genes.

Table S6 | Differentially expressed genes in CD137-expressing antigen-responsive cells.

Table S7 | SVM-predicted Flu MP58-66-responsive CD8+ T cells.

Table S8 | Differentially expressed genes in cognate antigen-stimulated CD8+ T cells from a donor with type 1 diabetes.

Table S9 | Primer pairs used for preamplification and qPCR.

REFERENCES

1. Andersen RS, Kvistborg P, Frosig TM, Pedersen NW, Lyngaa R, Bakker AH, et al. Parallel detection of antigen-specific T cell responses by combinatorial encoding of MHC multimers. *Nat Protoc.* (2012) 7:891–902. doi: 10.1038/nprot.2012.037
2. Unger WW, Velthuis J, Abreu JR, Laban S, Quinten E, Kester MG, Reker-Hadrup S, et al. Discovery of low-affinity preproinsulin epitopes and detection of autoreactive CD8 T-cells using combinatorial MHC multimers. *J Autoimmun.* (2011) 37:151–9. doi: 10.1016/j.jaut.2011.05.012
3. Hadrup SR, Bakker AH, Shu CJ, Andersen RS, van Veluw J, Hombrink P, et al. Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat Methods.* (2009) 6:520–6. doi: 10.1038/nmeth.1345
4. Glanville J, Huang H, Nau A, Hatton O, Wagar LE, Rubelt F, Ji X, et al. Identifying specificity groups in the T cell receptor repertoire. *Nature.* (2017) 547:94–8. doi: 10.1038/nature22976
5. Dash P, Fiore-Gartland AJ, Hertz T, Wang GC, Sharma S, Souquette A, et al. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature.* (2017) 547:89–93. doi: 10.1038/nature22383
6. Fan HC, Fu GK, Fodor SP. Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science.* (2015) 347:1258367 doi: 10.1126/science.1258367
7. Dolton G, Zervoudi E, Rius C, Wall A, Thomas HL, Fuller A, et al. Optimized Peptide-MHC multimer protocols for detection and isolation of autoimmune T-cells. *Front Immunol.* (2018) 9:1378. doi: 10.3389/fimmu.2018.01378
8. Wang M, Windgassen D, Papoutsakis ET. Comparative analysis of transcriptional profiling of CD3+, CD4+ and CD8+ T cells identifies novel immune response players in T-cell activation. *BMC Genomics.* (2008) 9:225. doi: 10.1186/1471-2164-9-225
9. Kao C, Daniels MA, Jameson SC. Loss of CD8 and TCR binding to Class I MHC ligands following T cell activation. *Int Immunol.* (2005) 17:1607–17. doi: 10.1093/intimm/dxh340
10. Alcover A, Alarcon B, Di Bartolo V. Cell biology of T cell receptor expression and regulation. *Ann Rev Immunol.* (2018) 36:103–25. doi: 10.1146/annurev-immunol-042617-053429
11. Viola A, Lanzavecchia A. T cell activation determined by T cell receptor number and tunable thresholds. *Science.* (1996) 273:104–6. doi: 10.1126/science.273.5271.104
12. Valkenburg SA, Josephs TM, Clemens EB, Grant EJ, Nguyen TH, Wang GC, et al. Molecular basis for universal HLA-A*0201-restricted CD8+ T-cell immunity against influenza viruses. *Proc Natl Acad Sci USA.* (2016) 113:4440–5. doi: 10.1073/pnas.1603106113
13. Naumov YN, Naumova EN, Yassai MB, Kota K, Welsh RM, Selin LK. Multiple glycines in TCR alpha-chains determine clonally diverse nature of human T cell memory to influenza A virus. *J Immunol.* (2008) 181:7407–19. doi: 10.4049/jimmunol.181.10.7407
14. Fuchs YF, Jainta GW, Kuhn D, Wilhelm C, Weigelt M, Karasinsky A, et al. Vagaries of the ELISpot assay: specific detection of antigen responsive cells requires purified CD8(+) T cells and MHC class I expressing antigen presenting cell lines. *Clin Immunol.* (2015) 157:216–25. doi: 10.1016/j.clim.2015.02.012
15. Brewitz A, Eickhoff S, Dahling S, Quast T, Bedoui S, Kroczek RA, et al. CD8(+) T cells orchestrate pDC-XCR1(+) dendritic cell spatial and functional cooperativity to optimize priming. *Immunity.* (2017) 46:205–19. doi: 10.1016/j.immuni.2017.01.003
16. Takeuchi A, Itoh Y, Takumi A, Ishihara C, Arase N, Yokosuka T, et al. CRTAM confers late-stage activation of CD8+ T cells to regulate retention within lymph node. *J Immunol.* (2009) 183:4220–8. doi: 10.4049/jimmunol.0901248
17. Pan Y, Tian T, Park CO, Lofftus SY, Mei S, Liu X, et al. Survival of tissue-resident memory T cells requires exogenous lipid uptake and metabolism. *Nature.* (2017) 543:252–6. doi: 10.1038/nature21379
18. Wolfl M, Kuball J, Ho WY, Nguyen H, Manley TJ, Bleakley M, Greenberg PD. Activation-induced expression of CD137 permits detection, isolation, and expansion of the full repertoire of CD8+ T cells responding to antigen without requiring knowledge of epitope specificities. *Blood.* (2007) 110:201–10. doi: 10.1182/blood-2006-11-056168
19. Wehler TC, Karg M, Distler E, Konur A, Nonn M, Meyer RG, et al. Rapid identification and sorting of viable virus-reactive CD4(+) and CD8(+) T cells based on antigen-triggered CD137 expression. *J Immunol Methods.* (2008) 339:23–37. doi: 10.1016/j.jim.2008.07.017
20. Coppieters KT, Dotta F, Amirian N, Campbell PD, Kay TW, Atkinson MA, et al. Demonstration of islet-autoreactive CD8 T cells in insulitic lesions from recent onset and long-term type 1 diabetes patients. *J Exp Med.* (2012) 209:51–60. doi: 10.1084/jem.20111187
21. Takaki T, Marron MP, Mathews CE, Guttmann ST, Bottino R, Trucco M, et al. HLA-A*0201-restricted T cells from humanized NOD mice recognize autoantigens of potential clinical relevance to type 1 diabetes. *J Immunol.* (2006) 176:3257–65. doi: 10.4049/jimmunol.176.5.3257
22. Unger WW, Pearson T, Abreu JR, Laban S, van der Slik AR, der Kracht SM, et al. Islet-specific CTL cloned from a type 1 diabetes patient cause beta-cell destruction after engraftment into HLA-A2 transgenic NOD/scid/IL2RG null mice. *PLoS ONE.* (2012) 7:e49213. doi: 10.1371/journal.pone.0049213
23. Fuchs YF, Eugster A, Dietz S, Sebelefsky C, Kuhn D, Wilhelm C, et al. CD8(+) T cells specific for the islet autoantigen IGRP are restricted in their T cell receptor chain usage. *Sci Rep.* (2017) 7:44661. doi: 10.1038/srep44661
24. Culina S, Lalanne AI, Afonso G, Cerosaletti K, Pinto S, Sebastiani G, et al. Islet-reactive CD8(+) T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. *Sci Immunol.* (2018) 3:eaao4013. doi: 10.1126/sciimmunol.aao4013
25. Matsuo K, Kitahata K, Kawabata F, Kamei M, Hara Y, Takamura S, et al. A highly active form of XCL1/Lymphotactin functions as an effective adjuvant to recruit cross-presenting dendritic cells for induction of effector and memory CD8(+) T cells. *Front Immunol.* (2018) 9:2775. doi: 10.3389/fimmu.2018.02775
26. Dorner BG, Scheffold A, Rolph MS, Huser MB, Kaufmann SH, Radbruch A, et al. MIP-1alpha, MIP-1beta, RANTES, and ATAC/lymphotactin function together with IFN-gamma as type 1 cytokines. *Proc Natl Acad Sci USA.* (2002) 99:6181–6. doi: 10.1073/pnas.092141999
27. Miao T, Symonds ALJ, Singh R, Symonds JD, Ogbe A, Omodho B, et al. Egr2 and 3 control adaptive immune responses by temporally uncoupling expansion from T cell differentiation. *J Exp Med.* (2017) 214:1787–808. doi: 10.1084/jem.20160553
28. Du N, Kwon H, Li P, West EE, Oh J, Liao W, et al. EGR2 is critical for peripheral naive T-cell differentiation and the T-cell response to influenza. *Proc Natl Acad Sci USA.* (2014) 111:16484–9. doi: 10.1073/pnas.1417215111
29. Dudda JC, Salaun B, Ji Y, Palmer DC, Monnot GC, Merck E, et al. MicroRNA-155 is required for effector CD8+ T cell responses to virus infection and cancer. *Immunity.* (2013) 38:742–53. doi: 10.1016/j.immuni.2012.12.006
30. Stelekati E, Chen Z, Manne S, Kurachi M, Ali MA, Lewy K, et al. Long-term persistence of exhausted CD8 T cells in chronic infection is regulated by MicroRNA-155. *Cell Rep.* (2018) 23:2142–56. doi: 10.1016/j.celrep.2018.04.038
31. Lippert E, Yowe DL, Gonzalo JA, Justice JP, Webster JM, Fedyk ER, et al. Role of regulator of G protein signaling 16 in inflammation-induced T lymphocyte migration and activation. *J Immunol.* (2003) 171:1542–55. doi: 10.4049/jimmunol.171.3.1542
32. Gubser PM, Bantug GR, Razik L, Fischer M, Dimeloe S, Hoenger G, et al. Rapid effector function of memory CD8+ T cells requires an immediate-early glycolytic switch. *Nat Immunol.* (2013) 14:1064–72. doi: 10.1038/ni.2687
33. Roge R, Thorsen J, Torring C, Ozbay A, Moller BK, Carstens J. Commonly used reference genes are actively regulated in *in vitro* stimulated lymphocytes. *Scand J Immuno.* (2007) 65:202–9. doi: 10.1111/j.1365-3083.2006.01879.x
34. Grifoni A, Costa-Ramos P, Pham J, Tian Y, Rosales SL, Seumois G, et al. Cutting edge: transcriptional profiling reveals multifunctional and cytotoxic antiviral responses of zika virus-specific CD8(+) T cells. *J Immunol.* (2018) 201:3487–91. doi: 10.4049/jimmunol.1801090
35. Gonzalez-Duque S, Azoury ME, Colli ML, Afonso G, Turatsinze JV, Nigi L, et al. Conventional and neo-antigenic peptides presented by β cells are targeted by circulating naive CD8+ T cells in type 1 diabetic and healthy donors. *Cell Metabo.* (2018) 28:946–60.e946. doi: 10.1016/j.cmet.2018.07.007
36. Williams T, Krovi HS, Landry LG, Crawford F, Jin N, Hohenstein A, et al. Development of T cell lines sensitive to antigen stimulation. *J Immunol Methods.* (2018) 462:65–73. doi: 10.1016/j.jim.2018.08.011

37. Britten CM, Meyer RG, Kreer T, Drexler I, Wolfel T, Herr W. The use of HLA-A*0201-transfected K562 as standard antigen-presenting cells for CD8(+) T lymphocytes in IFN-gamma ELISPOT assays. *J Immunol Methods.* (2002) 259:95–110. doi: 10.1016/S0022-1759(01)00499-9
38. Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc.* (2014) 9:171–81. doi: 10.1038/nprot.2014.006
39. Bonifacio E, Ziegler AG, Klingensmith G, Schober E, Bingley PJ, Rottenkolber M, et al. Effects of high-dose oral insulin on immune responses in children at high risk for type 1 diabetes: the Pre-POINT randomized clinical trial. *JAMA.* (2015) 313:1541–9. doi: 10.1001/jama.2015.2928
40. McDavid A, Dennis L, Danaher P, Finak G, Krouse M, Wang A, et al. Modeling bi-modality improves characterization of cell cycle on gene expression in single cells. *PLoS Computat Biol.* (2014) 10:e1003696. doi: 10.1371/journal.pcbi.1003696
41. Kharchenko PV, Silberstein L, Scadden DT. Bayesian approach to single-cell differential expression analysis. *Nat Methods.* (2014) 11:740–2. doi: 10.1038/nmeth.2967
42. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* (2014) 15:550. doi: 10.1186/s13059-014-0550-8
43. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* (2018) 36:411–20. doi: 10.1038/nbt.4096
44. Stubbington MJT, Lonnberg T, Proserpio V, Clare S, Speak AO, Dougan G, et al. T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods.* (2016) 13:329–32. doi: 10.1038/nmeth.3800
45. Alamyar E, Duroux P, Lefranc MP, Giudicelli V. IMGT((R)) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. *Methods Mol Biol.* (2012) 882:569–604. doi: 10.1007/978-1-61779-842-9_32

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.