T CELL MEMORY

Early precursors and molecular determinants of tissue-resident memory CD8+ T lymphocytes revealed by single-cell RNA sequencing

Nadia S. Kurd1*,‡, Zhaoren He1,3*, Tiani L. Louis1, J. Justin Milner3, Kyla D. Omilusik3, Wenhai Jin2, Matthew S. Tsai1, Christella E. Widjaja2, Jad N. Kanbar1, Jocelyn G. Olvera1, Tiffani Tysl1, Lauren K. Quezada1, Brigid S. Boland1, Wendy J. Huang2, Cornelis Murre3, Ananda W. Goldrath3, Gene W. Yeo2,4,5, John T. Chang1,5,6§

During an immune response to microbial infection, CD8+ T cells give rise to distinct classes of cellular progeny that coordinate mediate clearance of the pathogen (effector cells) and provide long-lasting protection against reinfection (memory cells). Considerable heterogeneity has been previously described within the long-lived circulating memory (T_{circ}) cells that mediate potent protection within nonlymphoid tissues. Here, we used single-cell RNA sequencing to examine the gene expression patterns of individual CD8+ T cells responding to microbial challenge, differentiating into secondary effector cells upon reinfection, effector memory (T_{EM}) cells that mediate immediate pathogen control via rapid and potent "effector" function, and tissue-resident memory (T_{RM}) cells with distinct homing and migratory properties. Together, these findings may inform strategies to optimize CD8+ T cell responses to protect against microbial infection and cancer.

INTRODUCTION

CD8+ T cells responding to microbial challenge differentiate into distinct subsets of cellular progeny with unique migratory and functional properties that coordinate mediate clearance of the pathogen (effector cells) and provide long-lasting protection against reinfection (memory cells). Considerable heterogeneity has been previously described within the long-lived circulating memory T cell pool (1–3). Whereas central memory (T_{CM}) cells exhibit greater self-renewal and plasticity with the ability to rapidly proliferate and differentiate into secondary effector cells upon reinfection, effector memory (T_{EM}) cells provide immediate pathogen control via rapid and potent "effector" function. Moreover, recent studies have revealed additional heterogeneity within the classically defined T_{EM} cell population, including long-lived effector (LLE) cells and peripheral memory cells, which can be distinguished by distinct surface molecule expression and trafficking properties (1, 2, 4–6). In addition to these circulating memory T cell populations, a noncirculating subset, termed tissue-resident memory (T_{RM}) cells, has recently been described (7). T_{RM} cells are found in most tissues and positioned at key barrier surfaces, such as the skin and intestinal epithelium, where they play critical roles in limiting early pathogen spread and controlling infection, and also help to control the outgrowth of cancer cells (8–11). Whereas heterogeneity within the circulating CD8+ T cell memory population has been well characterized, it remains unclear whether the tissue-resident CD8+ T cell population might also be composed of distinct subsets that play unique roles in mediating protective immunity.

Recent studies have begun to illuminate the mechanisms regulating T_{RM} cell differentiation, function, and survival. Activation of naïve CD8+ T cells occurs in the spleen or draining lymph nodes, resulting in the up-regulation of key transcription factors including Blimp-1 (12). Recruitment of activated CD8+ T cells to nonlymphoid tissue sites is mediated by chemokine receptors that promote tissue entry, such as CCR9 and CXCR3 (12–14). Upon entry to tissue, CD8+ T cells undergo transcriptional changes that enforce tissue residency, in part by dampening expression of receptors that promote return to circulation such as CCR7 and S1PR1 (14), and begin to direct the T_{RM} cell differentiation program. These changes include up-regulation of transcription factors such as Hobi, which, together with Blimp-1, repress genes associated with recirculation, including Klf2, Sipri1, and Ccr7; and down-regulation of the T-box transcription factors T-bet and Eomes, enabling transforming growth factor–β (TGF-β) responsiveness (12–15). TGF-β signals within the tissue induce expression of CD103, a key factor for tissue retention, whereas low levels of T-bet expression are required for interleukin-15 (IL-15) responsiveness, which plays an important role for long-term survival of T_{RM} cells in some tissues (13, 15). However, additional unidentified regulators likely contribute to coordinating the T_{RM} cell differentiation program. Moreover, although it has been shown that T_{RM} cells preferentially arise from circulating precursors with low expression of killer cell lectin like receptor G1 (KLRG1) (13, 14), it remains unknown whether all cells that enter the tissue are destined to persist and continue their differentiation into T_{RM} cells or whether T_{RM} cells are derived from a specific subset of precursors within the tissue at early time points.

Several studies have described a core transcriptional signature that is shared among T_{RM} cells from distinct tissues, highlighting several key transcriptional regulators of T_{RM} cell differentiation, including Hobi, Blimp-1, and Runx3 (12, 16–19). However, most of these prior studies have focused on relatively late time points after the T_{RM} cell population has been well established, providing...
only a snapshot of the gene expression patterns used by T_{RM} cells and potentially missing early events important for their differentiation. Moreover, because these studies relied on RNA sequencing of bulk cell populations, potential heterogeneity representing distinct functional subsets or intermediate states of differentiation may have been missed.

Single-cell RNA sequencing (scRNA-seq) is a powerful approach that can reveal heterogeneity within cell populations and has been used extensively in recent studies to probe the dynamic gene expression patterns within a wide range of immune cell types in health and disease (20–29). This approach has allowed for the elucidation of new cell subsets and states, such as highly efficacious subpopulations of tumor-infiltrating lymphocytes and early states of differentiation for circulating effector and memory CD8^+ T cells (20, 22, 23).

Here, we used scRNA-seq to generate a single-cell transcriptomic dataset elucidating the dynamic gene expression patterns of individual CD8^+ T cells in the spleen and small intestine intraepithelial lymphocyte (siIEL) compartment throughout the course of their differentiation in response to viral infection. These analyses demonstrate that the circulating and tissue-resident subtypes of memory CD8^+ T cells use highly overlapping patterns of gene expression, revealing a core transcriptional program used by both subtypes throughout their differentiation. In addition, these analyses elucidated sets of genes with kinetics and expression patterns unique to circulating or T_{RM} CD8^+ T cells that may contribute to the specification of each distinct subtype. These data also revealed transcriptional heterogeneity within the siIEL CD8^+ T cell population throughout its differentiation. We show that within the established T_{RM} cell population, this molecular heterogeneity reflects functionally distinct subsets that were previously unknown. Moreover, we identify a distinct subset within the CD8^+ T cell pool within the tissue early in infection that likely represents the precursors of T_{RM} cells. These findings should inform future studies aimed at improving our understanding of T_{RM} cell differentiation and function, which may contribute to our ability to better manipulate CD8^+ T cell responses to protect against infection and cancer.

RESULTS
scRNA-seq analyses of circulating and siIEL CD8^+ T cells responding to viral infection

We used a scRNA-seq approach to investigate the transcriptional changes that occur in circulating and siIEL CD8^+ T cells responding to viral infection. We adoptively transferred P14 CD8^+ CD45.1^+ T lymphocytes, which have transgenic expression of a T cell receptor (TCR) that recognizes an immunodominant epitope of lymphocytic choriomeningitis virus (LCMV), into congenic CD45.2^+ wild-type (WT) recipients that were infected with the Armstrong strain of LCMV 1 day later. We FACS (fluorescence-activated cell sorting)-sorted naïve (CD62L^hi CD44^lo) P14 T cells from spleens of uninfected mice and activated donor P14 T cells (CD44^hi) from the spleens and siIEL compartments of recipient mice at 11 time points after infection and performed scRNA-seq using the 10x Genomics Chromium platform (Fig. 1A and fig. S1).

To relate changes in CD8^+ T cell populations over time and across tissues, we merged the data from all time points and anatomic sites and performed unsupervised t-distributed stochastic neighbor embedding (tsNE) and uniform manifold approximation and projection (UMAP) analyses (Fig. 1, B to D, and fig. S2). CD8^+ T cells from the siIEL compartment clustered distinctly from splenic CD8^+ T cells at all time points (Fig. 1, B to D, and fig. S2), consistent with the distinct transcriptional profile of T_{RM} cells that has been previously reported (12, 13, 16, 18, 19). For example, expression of genes previously associated with T_{RM} cells, such as Cd69 and Itgae, was strongly enriched among siIEL CD8^+ T cells compared with splenic CD8^+ T cells, whereas expression of genes that promote tissue egress and recirculation, such as Klj2 and S1pr1, was strongly enriched among splenic cells compared with siIEL cells (Fig. 1E). The divergence in gene expression profiles between splenic and siIEL CD8^+ T cells was evident as early as day 4 after infection (Fig. 1, B to D, and fig. S2), the earliest time point at which CD8^+ T cells can be detected within the intestinal epithelium (30), indicating that CD8^+ T cells begin to change their transcriptional profile rapidly upon entry into tissue. Differential expression analyses revealed that 928 genes were more highly expressed in day 4 splenic CD8^+ T cells and 1103 genes were more highly expressed in day 4 siIEL CD8^+ T cells, including Cd69 and Itgae (Fig. 1, E and F; fig. S3; and table S1). Genes more highly expressed by siIEL CD8^+ T cells included those associated with processes known to be important for establishment and maintenance of T_{RM} cells, including integrins and cell adhesion molecules (Itga1, Itgb2, Itgal, Itgb7, Itgax, and Jam1), regulators of cell trafficing (Ccr9 and Cxcr3) and TGF-β signaling (Tgfβ1, Tgfbr2, Smad7, Skil, and Smurf2), the tissue damage receptor P2x7 (31), and fatty acid–binding proteins (Fabp1, Fabp2, and Fabp6) (Fig. 1F, fig. S3, and table S1) (32). In addition, genes associated with other aspects of T cell function, including cytokines (Il2 and Il10), cytokine receptors (Il4ra, Il2rg, Il10rb, and Il2r1), chemokines (Ccl3, Ccl4, Ccl5, Cxcl10, and Cxcl11), effector molecules (Gzma, Gzmk, and Fasl), and both costimulatory (Icos and Cd28) and inhibitory (Cila4, Tigit, and Lag3) receptors, were more highly expressed by siIEL CD8^+ T cells. We also observed that genes encoding components and downstream mediators of the TCR signaling pathway (Zap70, Itk, Lats2, Rasgrf1, Fyb, Nr4a1, Nr4a2, Nfat1, Irf4, Ikkf, and Cds) regulators of intracellular calcium (Orai1, Orai2, Sri, and Rrad), and regulators of nuclear factor κB (NF-κB) signaling (Nfkbia, Nfkbiz, Rel, Ikkbb, Pin1, and Tnfaip3) were more highly expressed among siIEL CD8^+ T cells. Many of these differentially expressed (DE) genes were transcription factors with no previously reported role in T_{RM} cells (Ikzf2, Ikzf3, Gata2, Irf4, and Ikd2). Among the genes more highly expressed by splenic CD8^+ T cells were Batf and Zeb2 transcription factors that have been shown to regulate effector CD8^+ T cell differentiation (33, 34); genes encoding for transcription factors, such as Batf3, that have not been previously implicated in CD8^+ T cell differentiation; and genes encoding the high-affinity IL-2 receptor (IL-2Rα) and the OX40 costimulatory receptor (Fig. 1F, fig. S3, and table S1). Thus, these analyses reveal previously unappreciated signaling pathways and transcription factors that may represent early regulators of circulating versus siIEL CD8^+ T cell differentiation.

Gene ontology (GO) analyses revealed that genes associated with DNA replication and cell cycle regulation were enriched among splenic CD8^+ T cells, suggesting that siIEL CD8^+ T cells may be less proliferative than splenic CD8^+ T cells (table S1). Assessment of cell cycle status inferred from transcriptional analyses suggested that although both splenic and siIEL CD8^+ T cells were actively dividing at day 4 after infection, siIEL CD8^+ T cells had stopped proliferating by day 7 after infection, whereas splenic CD8^+ T cells continued proliferating until 7 to 10 days after infection (Fig. 1G).
These findings indicated that siIEL CD8\(^+\) T cells become quiescent more rapidly after activation compared with splenic CD8\(^+\) T cells. Together, these data reveal that CD8\(^+\) T cells that enter the siIEL compartment receive signals that alter their transcriptional profile rapidly after arrival, indicating that the T\(_{RM}\) cell fate may begin to be specified earlier than previously appreciated and may serve as a useful resource for identifying early molecular determinants of T\(_{RM}\) cell differentiation.

**Shared and tissue-specific components of gene expression programs in circulating and siIEL CD8\(^+\) T cells**

We next sought to elucidate changes in gene expression programs over time in CD8\(^+\) T cells responding to viral infection and to understand how gene expression programs in circulating versus siIEL CD8\(^+\) T cells relate to one another. We analyzed splenic and siIEL CD8\(^+\) T cells separately and performed weighted gene coexpression network analyses on each set of cells, defining tissue-specific modules of genes that exhibited similar patterns of expression over time in splenic (fig. S4, A and B, and table S2) or siIEL (fig. S4, C and D, and table S2) CD8\(^+\) T cells. We identified 10 and 8 distinct gene expression modules among splenic and siIEL CD8\(^+\) T cells, respectively (fig. S4, A to D, and table S2).

Although recent studies have primarily highlighted the differences in gene expression between circulating and tissue-resident CD8\(^+\) T cells (12, 13, 16, 18, 19), we found substantial overlap in the gene expression patterns used by differentiating splenic and siIEL CD8\(^+\) T cells. For example, 38% of the genes in Spleen Module 1 were shared with IEL Module 2; genes in these early differentiation modules were characterized by a decrease in expression, relative to that by naïve cells, after T cell activation (fig. S4, A, B, and E). Similarly, 40% of the genes in IEL Module 4 were shared with Spleen Module 4; genes in these intermediate differentiation modules were characterized by high expression during the peak of infection, followed by a subsequent decrease. Last, 33% of the genes in IEL Module 7 were shared with Spleen Module 10; genes in these late differentiation modules were characterized by a
reduction in expression after T cell activation, followed by increasing expression that was sustained at late time points (fig. S4, B, D, and E, and table S2). These observations suggested that splenic and siEL CD8+ T cells share a core transcriptional program throughout their differentiation, consistent with their common cytolytic lymphocyte functions (fig. S4E and table S2), and highlight the potential value of this dataset to reveal how specific genes or pathways, including those known to play key roles in regulating circulating memory CD8+ T cells, may act similarly or disparately in regulating T RM cell differentiation.

To elucidate distinct regulators of the circulating memory and T RM cell differentiation programs, we again performed weighted gene coexpression network analyses but, this time, analyzed splenic and siEL CD8+ T cells together. This approach enabled us to directly compare the level of representation of each module in cells from each anatomic site at each time point. Fifteen modules representing distinct temporal patterns of gene expression were defined and annotated as Combined Modules 1 to 15 (Fig. 2, A to C, and table S3). This analysis confirmed our observation that splenic and siEL CD8+ T cells share a core transcriptional program, as many Combined Modules were similarly represented in splenic and siEL CD8+ T cells throughout most of their differentiation (Fig. 2, B and C). Moreover, this analysis confirmed our finding that siEL CD8+ T cells were transcriptionally distinct from splenic CD8+ T cells as early as day 4 after infection (Fig. 1, B to F), because we noted that many modules (Combined Modules 2, 3, 5, 7, 11, and 12) were differentially represented within splenic and siEL CD8+ T cells at day 4 after infection but not at later time points (Fig. 2, B and C). For example, genes in one such module, Combined Module 12, were enriched in several pathways known to play roles in circulating CD8+ T cell memory cells, including autophagy, protein ubiquitination, and proteasome protein catabolism (table S3) (35, 36). These findings suggested that differentiating siEL CD8+ T cells may begin to acquire certain aspects of a memory-like transcriptional profile more rapidly than splenic CD8+ T cells.

These analyses also defined several modules that were significantly and consistently differentially enriched in splenic versus siEL CD8+ T cells over time. For example, Combined Modules 6 and 15 were enriched in siEL compared with splenic CD8+ T cells; genes in these modules were highly up-regulated in siEL CD8+ T cells throughout their differentiation (Fig. 2, B and C). Combined Modules 6 and 15 contained 157 and 371 genes, respectively, some of which have been previously implicated in promoting differentiation of circulating memory CD8+ T cells at the expense of T RM cells (Fig. 2, B and C, and table S3). For example, these spleen-enriched modules included Klf2, Sipr1, Ly6c1, and Ly6c2, which regulate trafficking and egress, and Emes, a transcription factor that negatively regulates T RM cell differentiation (fig. S5 and table S3) (12, 14, 15, 30). These spleen-enriched modules also included several cell surface markers, including Cx3cr1 and Klrk1, that have been associated with subsets of circulating effector CD8+ T cells. In addition, the presence of genes encoding specific cytokine receptors (Il18r1, Il18r1ap, and Il17rα) in these spleen-enriched modules suggests that signaling through these receptors might negatively regulate T RM cell differentiation and maintenance.

To demonstrate that genes identified by these analyses represent functionally important regulators of T RM cell differentiation, we selected for further studies three genes found in Combined Module 15, nuclear receptor subfamily 4 group A member 2 (Nr4a2), and two activating protein-1 (AP-1) transcription factor subunits, Junb and Fosl2. These genes exhibited increased expression in siEL relative to splenic CD8+ T cells at all time points (Fig. 2D). Nr4a2 is an orphan nuclear receptor that belongs to the Nr4a family of transcription factors, plays a central role in the development of regulatory T cells and pathogenic T helper 17 (TH17) cells (40–43), and has been implicated as an important driver of exhaustion in CD8+ T cells (44). Junb and Fosl2 play important roles in regulating TH17 cell identity and pathogenicity (45–47), and Junb has also been implicated as part of the effector CD8+ T cell transcriptional program (34). Nr4a2, Junb, and Fosl2 have not been previously reported to regulate the differentiation of T RM cells. To investigate whether these genes play a role in T RM cell differentiation, we transduced P14 CD8+CD45.1+ T cells with retroviruses encoding short hairpin RNA (shRNA) targeting Nr4a2, Junb, or Fosl2 and transduced P14 CD8+CD45.1+ T cells with retroviruses encoding non-targeting shRNA. Cells were mixed at a 1:1 ratio and adoptively transferred into CD45.2 recipient mice that were subsequently infected with LCMV (fig. S6A). Relative to nontargeting controls, knockdown (KD) of Nr4a2, Junb, or Fosl2 resulted in a greater...
Fig. 2. Identification of a siIEL CD8+ T cell–enriched gene expression profile. (A) Gene expression patterns of single splenic and siIEL CD8+ T cells responding to infection, where each row represents an individual cell, grouped by tissue within each time point, and each column represents an individual gene, grouped by module, represented as expression relative to the mean expression among all cells. Weighted gene coexpression network analyses of splenic and siIEL CD8+ T cells considered together were performed to derive gene modules. (B and C) Representation of each module among single CD8+ T cells in both the spleen and siIEL compartment over time, relative to the mean representation among all cells, depicted as a representation score superimposed on individual cells (B) or as violin plots showing the relative representation of each module in individual splenic (teal) versus siIEL CD8+ T cells (coral) (C). (D) Violin plots depicting gene expression patterns of known or putative regulators of TEM cell differentiation (represented as transcripts per million (TPM)) selected from siIEL CD8+ T cell–enriched modules, among single splenic (teal) or siIEL (coral) CD8+ T cells over time. (E) CD45.1+ P14 T cells were transduced with retrovirus encoding shRNA targeting Nr4a2, Junb, or Fosl2 (KD) and mixed with CD45.1+ P14 T cells transduced with shRNA encoding control (nontarget) shRNA at a 1:1 ratio of KD/control cells before adoptive transfer into CD45.2+ hosts that were subsequently infected with LCMV. Twenty-two to 26 days later, splenic and siIEL CD8+ T cells were analyzed by flow cytometry. Ratio of KD/control P14 T cells within transduced P14 CD8+ T cells in the spleen (Tcirc) or within CD69+CD103+ siIEL P14 CD8+ T cells (TRM), normalized to the ratio of KD/control cells at the time of transfer. Values are normalized to Tcirc. *p < 0.05 and ****p < 0.0001 (Student's two-tailed t test). Data were pooled from two independent experiments, with means and SEM of n = 8 mice per gene, where each circle represents an individual mouse.

A reduction in siIEL TRM cells than splenic memory CD8+ T cells (Fig. 2E and fig. S6B). Together, these results highlight key transcriptional differences between circulating and siIEL CD8+ T cells undergoing differentiation and demonstrate the potential value of this dataset in identifying previously unknown genes and pathways involved in specifying TEM cell versus circulating memory CD8+ T cell fates.

scRNA-seq analyses highlight circulating CD8+ T cell heterogeneity and reveal previously unappreciated heterogeneity within the siIEL CD8+ T cell pool

We next investigated potential heterogeneity of differentiating splenic and siIEL CD8+ T cells. We observed that individual splenic CD8+ T cells at certain time points expressed genes previously associated with CD8+ T cell subsets, such as terminal effector (TE) and memory precursor (MP) phenotype cells at the peak of infection, and TEM, TCM, and LLE cell subsets at later time points after infection (Fig. 3, A and B). For example, the majority of cells at days 6 and 7 after infection expressed high levels of Klrk1, likely representing TE cells, whereas a smaller number of cells exhibited lower expression of Klrk1 and higher expression of Tcf7 and Bcl2, likely representing MP cells (48–50). We also observed that certain cells at later time points expressed genes previously associated with TCM (Il7r, Tcf7, Sell, Bcl2, and Cx3cr1), LLE (Klrk1, Cx3cr1, and Zeb2), and TEM (intermediate levels of Cx3cr1, Zeb2, Il7r, Tcf7, and Bcl2) cells (Fig. 3, A and B) (1, 3). Mapping individual splenic CD8+ T cells to TE, MP, TCM, TEM, and LLE transcriptional signatures defined by bulk RNA-seq signatures of sorted cells from each subset revealed that heterogeneity observed among splenic CD8+ T cells might represent cells differentiating into each of these memory cell subsets (Fig. 3, C and D). Low, intermediate, or high expression of Cx3cr1 at day 6 after infection generally corresponded to cells with TCM, TEM, or LLE profiles, respectively, consistent with previously reported Cx3cr1low, Cx3cr1int, and Cx3cr1hi subsets with distinct functional capacities and terminal differentiation potentials that can be found at the peak of infection (Fig. 3, B and D) (51). These results demonstrate the concurrent presence of these memory subsets within the circulating CD8+ T cell memory pool, consistent with previous studies (4–6, 52), and support findings suggesting that these memory
CD8⁺ T cell subsets may begin to diverge in their differentiation pathways earlier than is widely appreciated (51). These analyses also uncovered a notable transcriptional divergence between cells at days 60 versus 90 after infection (Fig. 3A). Prior studies have shown that T_CM cells are the predominant circulating memory subset at late time points, but whether and how T_CM cells change transcriptionally at late time points has not been investigated in depth. Compared with day 60 T_CM cells, day 90 T_CM cells expressed higher levels of *Eomes* and lower levels of genes encoding cytolytic granules such as *Gzma* and *Gzmb*, suggesting a progressive loss of cytotoxic capability over time (table S4). Moreover, these analyses highlighted a number of genes without known functions in T_CM cell differentiation and/or maintenance, such as *Hspa1a* and *Hspa1b*, which encode members of a family of conserved heat shock proteins (HSP70) that plays a role in protecting against cellular stressors, including hypoxia, temperature aberrations, pH alterations, and oxidative stress (53).

Although heterogeneity within circulating CD8⁺ T cells has been relatively well described, whether heterogeneity exists within tissue-resident CD8⁺ T cell populations remains unclear. To investigate potential heterogeneity within the siIEL CD8⁺ T cell compartment, we defined cell clusters across the entire dataset, revealing 40 distinct clusters, 17 of which were found in the siIEL compartment (Fig. 4A). We also performed partition-based graph abstraction (PAGA) trajectory analysis (54) to infer relationships between the siIEL clusters we defined (fig. S7). We observed that multiple clusters of siIEL CD8⁺ T cells were present within several individual time points, suggesting that distinct subsets might exist within the siIEL CD8⁺ T cell population (Fig. 4B). For example, at day 60 after infection, two major cell clusters, Clusters 3 and 29, were observed. These clusters exhibited differential expression of 236 genes, including several transcription factors, suggesting that these clusters were regulated by distinct gene expression programs (table S5). For example, Cluster 3 cells had higher expression of *Id3*, *Jun*, *Fos*, *Klf2*, and *Mye*, whereas Cluster 29 cells had higher expression of transcription factors including *Bcl6*, *Zeb2*, and *Klf3*. Cluster 3 cells also exhibited higher expression of several genes associated with T_RM cell function, including cytokines and chemokines (*Ifng*, *Tnf*, *Ccl3*, and *Ccl4*), suggesting that Cluster 3 cells might be better poised to produce cytokines rapidly upon reinfecion (Fig. 4C and table S5). Cluster 3 cells also had higher transcript levels of the transcription factor *Id3*, and a recent study identified an *Id3*hi T_RM cell subset with distinct functional capabilities including an enhanced capacity to produce cytokines (55). These analyses suggest that distinct costimulatory and other signaling pathways might be responsible for regulating each of these cell clusters and raise the possibility that these cell clusters might have distinct functional capacities. Consistent with this possibility, PAGA analysis revealed that whereas Cluster 3 cells were linked to

---

**Fig. 3. scRNA-seq analyses highlight heterogeneity among circulating CD8⁺ T cells.** (A) tSNE analysis of splenic CD8⁺ T cells; each plot represents an individual time point with each color representing a specific time point, as in Fig. 1B. (B) Relative expression (compared with mean expression among all spleen cells) of known regulators of circulating CD8⁺ T cell differentiation superimposed on individual spleen cells. (C and D) Similarity of gene expression programs among single spleen cells to transcriptional signatures (derived from bulk RNA-seq profiles from FACS-sorted cells) of previously defined circulating CD8⁺ T cell subsets. (E) Similarity of gene expression programs among single spleen cells to TE (D7 KLRG1hiIL-7Rα; red) or MP (D7 KLRG1hiIL-7Rα; blue) transcriptional signatures. (D) Similarity of gene expression programs among single spleen cells to LLE (D35 CD62LloIL-7Rα; red), T_EM (D35 CD62LloIL-7Rα; blue), and T_CM (D35 CD62LloIL-7Rα; green) cell transcriptional signatures.
Fig. 4. scRNA-seq analyses reveal functionally distinct subsets within the siIEL CD8+ T cell pool. (A and B) Clustering analyses of siIEL CD8+ T cells from all time points (A) or cells at days 4, 60, and 90 after infection (B); numbers represent different cluster annotations. (C and D) Expression of selected genes within single siIEL CD8+ T cells at day 60 (C) or day 90 (D) after infection, relative to the mean expression among siIEL CD8+ T cells at the indicated time point, demonstrating differential gene expression between Clusters 3 and 29 (C) or between Clusters 17 and 19 (D). (E to I) P14 CD8+ T cells were transferred into congenically distinct recipients 1 day before infection with LCMV. Splenic and siIEL CD8+ T cells were harvested at day 30 or 90 after infection. (E) Representative flow cytometry plots displaying distribution of CD28 expression among total P14 T cells in the spleen (top) or CD69+CD103+ CD8+ T cells in the siIEL compartment (bottom). Numbers represent the percentage of total P14 T cells in each gate, and graphs to the right demonstrate the quantification of CD28lo (filled circles) and CD28hi (open circles) cells within the indicated population at each time point. Data are representative of (flow cytometry plots; left) or compiled from (graphs; right) three independent experiments where n = 2 to 5 mice per experiment (10 to 11 total mice). (F) Quantification of CD127 (IL-7Rα) expression by CD28lo (gray) or CD28hi (black) subsets within the spleen (top) or siIEL (bottom) at day 30 after infection, with representative flow cytometry plots shown on the left. Numbers represent mean fluorescence intensity (MFI) of CD127 within the indicated population. Data are representative of (left) or compiled from (graph on right side) n = 3 independent experiments with n = 7 total mice, where each circle represents an individual mouse. (G) Representative flow cytometry plots displaying (top) distribution of CD28 expression at day 30 after infection among total H-2D<sup>D</sup> GP<sub>33–41</sub> tetramer+ endogenous CD8+ T cells in the spleen (left) or among H-2D<sup>D</sup> GP<sub>33–41</sub> tetramer+ endogenous CD69<sup>+</sup>CD103<sup>+</sup> CD8+ T cells in the siIEL compartment (right) or displaying (bottom) expression of CD127 by CD28<sup>lo</sup> (gray) or CD28<sup>hi</sup> (black) subsets within the indicated compartment. Numbers represent the percentage of H-2D<sup>D</sup> GP<sub>33–41</sub> tetramer+ cells in each gate (top) or MFI of CD127 expression in the indicated population (bottom), quantified in the graph (right), where n = 4 individual mice. (H and I) At day 30 after infection, siIEL P14 T cells were sorted into CD127<sup>lo</sup> and CD127<sup>hi</sup> subsets and cultured in the presence of GP<sub>33–41</sub> Peptide in vitro. Quantification (right) of the proportion of CD127<sup>lo</sup> or CD127<sup>hi</sup> P14 siIEL CD8+ T cells producing IFN-γ and TNF-α (H) or IL-2 and TNF-α (I) as shown in representative flow cytometry plots (left). Data are representative of two independent experiments, with means and SEM of n = 4 wells of cultured cells per phenotype, derived from two separate pools of sorted cells plated in duplicate. *P < 0.05, **P < 0.01, and ***P < 0.0001 [paired t test (F and G) or Student’s two-tailed t test (H and I)].

Day 90 clusters, Cluster 29 cells were not, suggesting that Cluster 29 cells might represent a terminal state (fig. S7). These findings are consistent with increased expression of the transcription factor Zeb2, which is known to promote the terminal differentiation of circulating CD8+ T cells (33), by Cluster 29, and the reciprocal increased expression among Cluster 3 cells of the transcription factor...
factor Id3, which is known to regulate the development of long-lived circulating memory CD8^+ T cells (56, 57).

In addition, at day 90 after infection, two major cell clusters, Clusters 17 and 19, were observed; these clusters exhibited differential expression of 361 genes (table S6). Cluster 17 cells exhibited higher expression of 243 genes, including transcription factors (Bhlhe40, Ddx5, Foxo1, Irf4, Junb, Nr3c1, Nr4a1, Nr4a2, Nr4a3, Pnrc1, and Rora) and mediators of cytokine signaling (Ifngr1, Il1r1, Il21r, Il4ra, Stat3, Stat5a, and Stat4). Cluster 17 cells also exhibited higher expression of Zfp36l2, which encodes for a zinc finger protein involved in the regulation of mRNA decay (58), and the transcript for interferon-γ (IFN-γ), as well as genes associated with pathways that were highly represented within TRM cell–enriched modules, such as the regulation of NF-κB signaling (Nfkbia, Nfkbid, Tnfaf1p, Pim1, Rel, Nfkb1, Nfkbie, and Nfkbiz) and costimulatory and inhibitory molecules (Icos and Cita4) (Fig. 4D and table S6). Cluster 17 cells also exhibited higher expression of the costimulatory receptor Cd28, suggesting that costimulation through this receptor could specifically influence the differentiation or responsiveness of a specific subset of TRM cells. By contrast, Cluster 19 exhibited higher expression of 117 genes, including the tissue-homing molecule Cxcr3, interferon response genes (Trim14 and Ifit1b1l), and Pik3i1p1, a negative regulator of T cell activation (Fig. 4D and table S6). Together, these findings suggested that specific cytokine and/or costimulatory signals might be responsible for maintaining or directing the differentiation of Cluster 17 cells. Moreover, whereas Cluster 17 cells might be better poised to produce cytokines rapidly upon reinfection, Cluster 19 cells might be more likely to respond immediately to type I interferon signals within the tissue. These data suggest that multiple subsets with distinct transcriptional programs and functional capacities may exist within the TRM cell pool.

We next performed flow cytometry to determine whether the heterogeneity revealed by scRNA-seq analyses could be discerned at the protein level. CD45.1^+ P14 CD8^+ T cells were adoptively transferred into congenic CD45.2^+ hosts that were infected with LCMV 1 day later. At days 30 and 90 after infection, the spleens and siIEL compartments of recipient mice were harvested for analysis. At day 30 after infection, both splenic and siIEL CD8^+ T cells could be divided into CD28^hi and CD28^lo populations (Fig. 4E). In contrast, at day 90 after infection, splenic P14 CD8^+ T cells uniformly expressed high levels of CD28, but heterogeneity was retained within the siIEL P14 CD8^+ T cell population (Fig. 4E), in line with data revealing multiple clusters within the TRM cell population with differential expression of Cd28 at the mRNA level at day 90 after infection (Fig. 4D). We also noted that high expression of IL-7Ra (Cd127), which was observed in the functionally distinct Id3^hi TRM cell subset described in a recent study (55), also correlated with high expression of Cd28 (Fig. 4F). We also observed this heterogeneity among endogenous, siIEL CD8^+ H-2D^b GP33-41 tetramer^+ cells, with CD28^hi cells expressing higher levels of Cd127 (Fig. 4G). To investigate whether heterogeneity in expression of these markers might reflect functional heterogeneity, we sorted CD127^hi and CD127^lo populations from the siIEL CD8^+ T cell pool at day 30 after infection and measured their capacities to produce cytokines when rechallenged with cognate antigen in vitro. We found that CD127^hi cells produced higher levels of the cytokines IFN-γ, tumor necrosis factor–α (TNF-α), and IL-2 in response to restimulation than did CD127^lo cells (Fig. 4, H and I), consistent with enhanced cytokine production in Id3^hi TRM cells (55), and the higher level of Ifnγ transcript detected within CD28^hi cells at day 90 after infection (Fig. 4, C and D). Although it remains possible that differences in the ability to produce cytokine upon restimulation may reflect different responses to ex vivo conditions, the observation that these two populations respond differently suggests that they are functionally distinct. Together, these data demonstrate that the heterogeneity revealed by scRNA-seq analyses within the TRM cell pool at late time points after infection may be functionally important.

The observation that expression of a number of transcriptional regulators, including Ddx5, Junb, Nr4a2, and Pnrc1, as well as the RNA binding protein Zfp36l2, correlated with Cd28 expression (Fig. 4D), raised the possibility that these genes might represent regulators of TRM cell heterogeneity. To investigate this possibility, we transduced P14 CD8^+ CD45.1^+ T cells with retroviruses encoding shRNA targeting Junb, Nr4a2, Pnrc1, or Zfp36l2 before adoptive transfer into congenic CD45.2^+ hosts that were subsequently infected with LCMV. Knockdown of Nr4a2 and Junb resulted in a loss of total TRM cells, whereas knockdown of Pnrc1 or Zfp36l2 did not affect the size of the total TRM cell population (fig. S6C). However, knockdown of each of these genes resulted in a reduction in the proportion of siIEL TRM (CD69^+ CD103^+) cells expressing high levels of CD28 compared with congenically controlled, cotransfected control P14 CD8^+ T cells transduced with retroviruses encoding nontargeting shRNA (Fig. 5A and fig. S6C). In addition, we cotransfected P14 CD8^+ T cells from mice with a T cell–specific deletion of Ddx5 (Ddx5^−/−/CD4cre^+/−) and congenically distinct control P14 CD8^+ T cells (Ddx5^+/+CD4cre^−−) into recipients infected with LCMV 1 day later and found that although Ddx5 deletion resulted in a severe depletion of both the total circulating memory T cell and total TRM cell populations, the proportion of CD28^hi cells was further reduced among the remaining Ddx5^−/− TRM cell population (Fig. 5B and fig. S6C). Moreover, consistent with this decrease in the proportion of CD28^hi cells, Ddx5^−/− siIEL, but not splenic, CD8^+ T cells exhibited a reduced ability to produce IFN-γ and TNF-α upon restimulation with cognate antigen in vitro (Fig. 5C). Together, these findings suggest that, similar to Prdm1 and Id3 (55), Ddx5, Junb, Nr4a2, Pnrc1, and Zfp36l2 regulate the differentiation of transcriptionally and functionally distinct TRM cell subsets.

Although it is known that two distinct populations with different memory potentials, distinguished by high or low expression of the high-affinity IL-2 receptor, IL-2Ra (Cd25) (59, 60), exist within the circulating CD8^+ T cell pool early in infection, whether the CD8^+ T cell pool that seeds the tissue early in infection is heterogeneous remained unknown. To determine whether heterogeneity of siIEL CD8^+ T cells can be discerned early after infection, we next examined the earliest time point, day 4 after infection, at which CD8^+ T cells could be detected in the siIEL compartment. In addition to the heterogeneity observed at late time points after infection (Fig. 4B), two major clusters, Clusters 16 and 20, were evident at day 4 after infection and exhibited differential expression of 332 genes (Figs. 4B, 5B, and 6A and table S7). Only 17 of these genes were more highly expressed by Cluster 16 cells, such as interferon response genes Ifit1, Ifit1b1l, and Ifit3; Pik3i1p1, a negative regulator of TCR signaling; Mxd4, a Myc-antagonist that promotes survival in T cells (61); and Kdm5b, a histone demethylase (Fig. 6A and table S7). GO analyses revealed that the genes expressed more highly by Cluster 20 cells were overrepresented in biologic processes such as DNA replication, mitotic spindle and nucleosome assembly, cytokinesis, and cell division; specific genes included Cdc7, Cdc25b, Cenpk,

**Zfp36l2** and **Ddx5** of CD8+ T cells during the effector phase of the immune response

Ddx5 activates TNF-α and IFN-α (bottom), as shown in representative flow cytometry plots (top).

Before adoptive transfer into CD45.2+ hosts that were subsequently infected with LCMV, WT siIEL P14 CD8+ T cells, shown as quantification of the proportion of cells producing TNF-α and IFN-α (bottom), indicates a less terminally differentiated phenotype. Moreover, mice that had received IL-2Rαlo siIEL CD8+ T cells had a higher proportion of CD69hiCD103+ CD8+ T cells within the siIEL compartment at both days 7 and 30 (Fig. 6G), suggesting that IL-2Rαlo siIEL CD8+ T cells at day 4 after infection are more likely to give rise to bona fide TRM cells that persist long-term. Together, these data suggest that, analogous to IL-2Rαlo and IL-2Rαhi populations within the circulating CD8+ T cell pool early in infection (59, 60), two functionally distinct CD8+ T cell populations are present within the siIEL compartment at day 4 after infection: a IL-2Rαhi subset with a more terminally differentiated phenotype that likely represents a transient tissue effector population and a IL-2Rαlo subset with higher memory potential that likely contains the precursors of TRM cells. Cluster 20 cells, corresponding to the IL-2Rαlo population, are more proliferative, exhibit a transcriptional profile that includes factors that are associated with the effector CD8+ T cell fate (Ezh2 and Dnmt1), and are more likely to give rise to terminally differentiated effector-like CD8+ T cells. In contrast, Cluster 16 cells, corresponding to the IL-2Rαlo subset, may be more quiescent and better poised to respond to type I interferons, have greater survival capacity, and have greater potential to give rise to long-lived TRM CD8+ T cells. Together, these results demonstrate previously unappreciated transcriptional and functional heterogeneity within the siIEL CD8+ T cell pool at multiple states of differentiation.

DISCUSSION

Recent studies have begun to elucidate key regulators of TRM cell differentiation, function, and survival, such as the transcription factors Blimp1, Hobit, and Runx3. Our scRNA-seq analyses have identified a number of additional putative regulators of TRM cell differentiation. For example, **Nr4a2**, **Junb**, and **Fosl2** were among the 528 genes that were substantially enriched in siIEL CD8+ T cells relative to previously shown to promote effector CD8+ T cell differentiation by mediating the repression of memory-associated genes (20, 63). Consistent with these findings, PAGA trajectory analysis (Fig. S7) revealed that whereas Cluster 16 cells were linked to day 7 clusters, Cluster 20 cells were not, suggesting that Cluster 20 cells might be undergoing terminal differentiation, whereas Cluster 16 cells might represent precursors to siIEL CD8+ TRM cells.

To formally test this possibility, we adoptively transferred P14 CD8+CD45.1+ T cells into CD45.2+ hosts infected with LCMV 1 day later and harvested the siIEL compartment at day 4 after infection for flow cytometric analysis. We found that, in addition to IL-2Rαhi and IL-2Rαlo subsets within the early circulating CD8+ T cell population, there was also heterogeneity in IL-2Rα expression among day 4 CD8+ T cells within the small intestinal epithelial tissue; this heterogeneity was also evident among endogenous siIEL CD8+ H-2Dβ GP33-41 tetramer+ cells at day 4 after infection (Fig. 6C). We also found that protein expression of IL-2Rα and Ezh2 was correlated (Fig. 6D) and took advantage of this observation to test whether these populations within the tissue have distinct differentiation potentials. CD45.1+ CD8+IL-2Rαhi or CD8+IL-2Rαlo cells from the siIEL compartment were FACS-sorted at 4 days after infection and adoptively transferred intravenously into CD45.2+ recipient mice infected with LCMV 1 day before transfer; analysis of siIEL CD8+ T cells was performed 7 or 30 days later (Fig. 6E). P14 CD8+ T cells within the siIEL compartment of mice that had received IL-2Rαlo cells expressed lower levels of KLRG1 at day 7 after infection (Fig. 6F), indicating a less terminally differentiated phenotype. Moreover, mice that had received IL-2Rαlo P14 siIEL CD8+ T cells had a higher proportion of CD69hiCD103+ CD8+ T cells within the siIEL compartment at both days 7 and 30 (Fig. 6G), suggesting that IL-2Rαlo siIEL CD8+ T cells at day 4 after infection are more likely to give rise to bona fide TRM cells that persist long-term. Together, these data suggest that, analogous to IL-2Rαlo and IL-2Rαhi populations within the circulating CD8+ T cell pool early in infection (59, 60), two functionally distinct CD8+ T cell populations are present within the siIEL compartment at day 4 after infection: a IL-2Rαhi subset with a more terminally differentiated phenotype that likely represents a transient tissue effector population and a IL-2Rαlo subset with higher memory potential that likely contains the precursors of TRM cells. Cluster 20 cells, corresponding to the IL-2Rαlo population, are more proliferative, exhibit a transcriptional profile that includes factors that are associated with the effector CD8+ T cell fate (Ezh2 and Dnmt1), and are more likely to give rise to terminally differentiated effector-like CD8+ T cells. In contrast, Cluster 16 cells, corresponding to the IL-2Rαlo subset, may be more quiescent and better poised to respond to type I interferons, have greater survival capacity, and have greater potential to give rise to long-lived TRM CD8+ T cells. Together, these results demonstrate previously unappreciated transcriptional and functional heterogeneity within the siIEL CD8+ T cell pool at multiple states of differentiation.

Weights of CD69+CD103+ expressing high levels of CD28 (bottom) as shown in representative flow cytometry plots (top). Data are representative of one to two experiments, with means and SEM of n = 4 mice per group, where each circle represents an individual mouse. (B and C) P14 CD8+ T cells from mice with T cell–specific deletion of Ddx5 (Ddx5−/− CD4-Cre−) were cotransferred at a 1:1 ratio with congenically distinct control P14 CD8+ T cells (CD69hiCD103−) expressing high levels of CD28 (bottom) as shown in representative flow cytometry plots (top). Data are compiled from (B) or representative of (C) two independent experiments with means and SEM of n = 4 to 12 mice, where each circle represents an individual mouse. *P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001 (paired t test).

Cenpm, Cdc20, and Cdk1 (Fig. 6A and table S7). Moreover, additional analyses confirmed that a greater proportion of Cluster 20 cells were in the G2-M phases of the cell cycle compared with Cluster 16 cells, indicating that cells in Cluster 20 were more actively proliferating (Fig. 6B). Cluster 20 cells also expressed higher levels of Dmnt1, a DNA methyltransferase that is critical for the expansion of CD8+ T cells during the effector phase of the immune response (62), along with several components of the polycomb repressive complex 2, including Suz12 and Ezh2 (Fig. 6A), which negatively regulates gene expression via histone methylation and has been
Fig. 6. T_RM cell precursors identified within the siIEL CD8^+ T cell pool early in infection. (A) Expression of selected genes within single siIEL CD8^+ T cells at day 4 after infection, relative to the mean expression among all cells, demonstrating differential gene expression between cells in Clusters 16 and 20. (B) Cell cycle status of siIEL CD8^+ T cells at day 4 after infection, inferred from transcriptional profiles. (C and D) P14 CD8^+ T cells were adoptively transferred into congenic hosts 1 day before infection with LCMV. Splenic and siIEL CD8^+ T cells were analyzed by flow cytometry at day 4 after infection. Data are representative of two independent experiments, with means and SEM of n = 2 to 3 mice per experiment. (C) Representative flow cytometry plots displaying distribution of IL-2Rα expression among total siIEL P14 CD8^+ T cells at day 4 after infection (left) or among total H-2D^b^GP33-41 tetramer^+ endogenous siIEL CD8^+ T cells at day 4 after infection in mice that had not received adoptive transfer of P14 CD8^+ T cells (right). (D) Quantification of Ezh2 expression among IL-2Rα^lo^ (gray) and IL-2Rα^hi^ (blue) P14 siIEL CD8^+ T cells, as gated in (C), with representative flow cytometry plot (left) and quantification (right). Numbers represent MFI of Ezh2 within the indicated population. Data are representative of (plot to left) or compiled from (graph to right) two independent experiments, with n = 4 total mice, where each dot represents an individual mouse. (E) Schematic of experimental setup. CD45.1^+^P14 CD8^+^ T cells were adoptively transferred into CD45.2^+^ congenic hosts 1 day before infection with LCMV Armstrong. CD45.1^+^P14 CD8^+^ T cells were collected from the siIEL compartment at day 4 after infection and sorted on the basis of expression of IL-2Rα as shown in (C). IL-2Rα^hi^ and IL-2Rα^lo^ populations were adaptively transferred into secondary hosts that had been infected with LCMV 1 day before transfer. Seven or 30 days later, siIEL P14 CD8^+^ T cells were analyzed by flow cytometry. (F) Quantification of the proportion of KLRG1^hi^ P14 CD8^+^ T cells present within the siIEL compartment at day 7 after infection. Data were pooled from two independent experiments, with means and SEM of n = 7 (IL-2Rα^hi^) or 5 (IL-2Rα^lo^), where each circle represents an individual mouse. (G) Quantification of the percentage of siIEL CD45.1^+^P14 CD8^+^ CD69^+^CD103^+^ T cells derived from IL-2Rα^lo^ (gray) or IL-2Rα^hi^ (blue) day 4 siIEL cells at day 7 or day 30 after transfer, with representative flow cytometry plots shown to the left. Day 7 data were pooled from two independent experiments, with means and SEM of n = 5 (IL-2Rα^lo^) or 6 (IL-2Rα^hi^) hosts, where each circle represents an individual mouse. Day 30 data were pooled from three independent experiments, with means and SEM of n = 4 (IL-2Rα^lo^) or 12 (IL-2Rα^hi^) hosts, where each circle represents an individual mouse. *P < 0.05 and **P < 0.01 [paired t test (D) or Student’s two-tailed t test (F and G)].
spleen and bone marrow of naive mice, we found that knockdown of Nr4a2, Junb, or Fosl2 resulted in impaired TRM cell differentiation. Junb and Fosl2 encode for AP-1 dimerization partners that have been reported to repress T-bet expression in TH17 cells (46, 47). Given that down-regulation of T-bet expression is important for early establishment of the TRM cell transcriptional program (15), it is tempting to speculate that Junb and Fosl2 may regulate TRM cell differentiation by cooperatively regulating T-bet expression. Because Fosl2 has been previously reported to positively regulate Smad3 (47), a key component of the TGF-β signaling pathway, Fosl2 may also promote TRM cell differentiation through effects on TGF-β signaling.

In addition to identifying specific putative regulators of TRM cell differentiation, our analyses have implicated new pathways that may control aspects of TRM cell biology. For example, we observed that splenic CD8+ T cells exhibited increased levels of transcripts associated with TCR signaling, even after infection had been cleared and antigen was no longer present, suggesting that CD8+ T cells within tissues are transcriptionally poised for rapid responsiveness to TCR stimulation. Because TRM cells have been reported to exhibit a decreased ability to scan for antigens owing to reduced motility relative to circulating CD8+ T cells (64, 65), such enhanced TCR responsiveness might represent a potential mechanism enabling TRM cells to mount robust protective responses even in the face of limiting amounts of cognate antigen. Moreover, we observed that differentiating siIEL CD8+ T cells also exhibited sustained expression of genes encoding inhibitory receptors, including Ctla4, Tigit, and Lag3, consistent with prior reports (13, 18), which may provide a balance against higher basal TCR responsiveness and enable TRM cells to avoid excessive responses that might lead to autoimmune pathology (66).

Our analyses also provide new insights regarding TRM cell ontogeny, which has not been as well studied as that of circulating memory CD8+ T cells (20, 59, 67). Other than the observations that TRM cells are derived from circulating cells lacking high expression of KLRG1 (13, 16) and share a common clonal origin as TCM cells (68), very little is known about the precursors of TRM cells after their arrival in the tissue. We found that siIEL CD8+ T cells were transcriptionally distinct from splenic CD8+ T cells at 4 days after infection, the earliest time point at which these cells can be detected within the small intestinal tissue (30). One interpretation of this finding is that transcriptional changes induced by the local tissue microenvironment occur rapidly upon tissue entry. Alternatively, CD8+ T cells that seed tissues may represent precommitted TRM cell precursors that have already acquired some aspects of the TRM cell transcriptional program. Although there were distinct clusters of splenic CD8+ T cells observed at day 3 after infection, none of these clusters appeared to be transcriptionally similar to siIEL CD8+ T cells. This finding suggests that the TRM cell transcriptional program may not be initiated until after the cells have entered the tissue and does not support the hypothesis that precommitted TRM precursor cells are formed within the circulating CD8+ T cell pool. However, it has been previously shown that priming by DNGR-1+ dendritic cells in the draining lymph nodes is an important step for TRM cell differentiation in vaccinia virus infection in the skin and influenza A infection in the lung (69). These findings suggest that the mechanisms underlying early TRM cell fate specification might differ for localized infections and might even be distinct among different tissues or pathogens. A priming step that specifically induces a TRM cell precursor population that is poised to home to a specific tissue could be particularly advantageous for a localized infection actively occurring within that tissue. In contrast, a more generalized mechanism of TRM cell specification, in which less committed precursor cells enter tissues and differentiate into TRM cells only in response to local tissue signals, might be more permissive for seeding of multiple, diverse tissues, as occurs in a systemic infection. Alternatively, it is possible that both models of TRM cell fate specification could occur simultaneously during an infection, whereby each pathway induces functionally distinct subsets of TRM cells.

In addition to finding substantial transcriptional differences between splenic and siIEL CD8+ T cells at day 4 after infection, we also observed two transcriptionally distinct cell clusters within the siIEL CD8+ T cell pool at day 4 after infection. Cells from one cluster expressed higher levels of genes associated with proliferation and effector differentiation and were less likely to develop into long-lived CD69hiCD103hi TRM cells when transferred into new recipients. By contrast, cells from the other cluster, marked by lower expression of IL-2Ra, exhibited a greater potential to give rise to TRM cells upon adoptive transfer. Although it remains unknown whether this heterogeneity is a specific feature of the siIEL compartment or generalizable to TRM cell differentiation in other nonlymphoid tissues, our analyses establish that analogous to the IL-2Ra- population present within the circulating CD8+ T cell pool early in infection that contains the precursors of circulating memory cells (59, 60), bona fide TRM cell precursors can be found within the IL-2Ra+ subset present within the siIEL compartment at day 4 after infection. This finding represents an important step in understanding the ontogeny of TRM cells after their arrival in the tissue that will likely inform future studies to identify early determinants of TRM cell fate specification.

Whereas the heterogeneity within the circulating memory CD8+ T cell pool is well characterized, heterogeneity among TRM cells is much less clear. Some studies have reported tissue-specific differences in CD69 and CD103 expression by CD8+ T cells, but it is unknown whether this heterogeneity represents functionally distinct subsets (18, 70–72). Our analyses identify two transcriptionally distinct cell clusters at days 60 and 90 after infection that exhibit unique functional capacities. One population, distinguished by higher expression of CD28 and IL-7Ra, expressed high levels of transcripts encoding inflammatory cytokines and exhibited an enhanced ability to produce cytokines in response to restimulation. In addition, this population exhibited higher expression of Klf2, which promotes tissue egress and recirculation. By contrast, the other population, characterized by lower CD28 and IL-7Ra expression, exhibited higher expression of transcripts encoding molecules that promote tissue retention, such as Cxcr3 and Klf3 (13, 73). Recent studies have revealed that the “tissue residency” of TRM cells is not as permanent as previously thought; upon reactivation, TRM cells harbor the potential to leave the tissues and join the circulating memory pool or form TRM cell populations within secondary lymphoid organs (74, 75). Our findings suggest that this functionality may be restricted to a specific subset of TRM cells that is better poised to leave the tissue. A recent study found that a subset of TRM cells with high expression of the transcription factor Id3, which we also found to be more highly expressed within the CD28hi TRM population, exhibits an enhanced capacity to survive and give rise to circulating memory cells when adoptively transferred into new hosts (55). By contrast, the second TRM cell subset might be more prone to remain within the tissue and mediate protective responses directly within...
the tissue microenvironment. Moreover, in addition to revealing functionally distinct subsets within the TRM cell pool, our transcriptional analyses have also elucidated a number of factors regulating the differentiation of these subsets. Analogous to our findings, a recent study provided evidence for distinct subsets of human TRM cells, one with a greater potential for cytokine production and another with a higher proliferative potential in response to TCR stimulation (76).

Overall, our work has resulted in a single-cell transcriptomic dataset encompassing the gene expression patterns of circulating and siEL CD8+ T cells in response to viral infection. Our study reveals a core transcriptional program that is shared between circulating memory and TRM cells in addition to key differences in the kinetics and magnitude of gene expression between these two memory cell subtypes, which should serve as a useful resource for elucidating new genes and pathways regulating TRM cell differentiation. Our analyses demonstrate that CD8+ T cells within the siEL compartment become transcriptionally distinct from circulating cells rapidly upon entry into tissue and identify a subset of early siEL CD8+ T cells enriched for precursors of TRM cells. Moreover, our study reveals previously unperturbed molecular and functional heterogeneity within the TRM cell pool, underscoring the power and necessity of using a single-cell approach. This dataset should inform future studies aimed at improving our understanding of CD8+ T cell differentiation and function, which may lead to strategies to optimize CD8+ T cell responses to protect against microbial infection and cancer.

MATERIALS AND METHODS

Study design

The purpose of this study was to gain a broader understanding of the gene expression patterns that regulate CD8+ T cell differentiation and heterogeneity in response to viral infection. To this end, we performed scRNA-seq on CD8+ T cells in the spleen and small intestinal epithelium at various time points after viral infection to generate a dataset analyzing gene expression patterns over time in individual CD8+ T cells. Analysis of this dataset revealed a subset of genes whose expression was enriched in CD8+ T cells from the tissue, and the role of several of these factors in regulating the differentiation of TRM CD8+ T cells was validated by flow cytometric analysis of CD8+ T cells with shRNA-mediated knockdown of these genes. In addition, validation of putative subsets of CD8+ T cells within the tissue was performed by flow cytometric analysis, including sorting of individual putative subsets and assessment of function after ex vivo restimulation or adoptive transfer into secondary hosts. Information regarding the sample size and number of replicates for each experiment can be found in the relevant figure legend. All findings were successfully reproduced. No sample size calculations were performed; sample sizes were selected on the basis of previous studies performed in our lab. Mice were randomly allocated into groups before adoptive transfer of sorted IL-2Rαhi or IL-2Rαlo intraepithelial P14 T cells (Fig. 6). For scRNA-seq experiments, mice were randomly selected for P14 T cell harvesting at specific time points after infection. Randomization and blinding are not relevant to shRNA KD experiments or experiments involving cotransfer of Ddxs5−/− and WT CD8+ T cells, because P14 T cells transduced with retrovirus encoding shRNA against genes of interest and congenically distinct P14 T cells transduced with control retrovirus, or Ddxs5−/− and WT CD8+ T cells, were cotransferred into the same animals (Figs. 2 and 5). For assessment of phenotype of adoptively transferred sorted IL-2Rαhi or IL-2Rαlo intraepithelial P14 T cells (Fig. 6), the investigator was aware of the cell type transferred into each recipient. No data were excluded from analysis, except for recipient mice that had rejected adoptively transferred P14 CD8+ T cells.

Mice

All mice were housed under specific pathogen-free conditions in an American Association of Laboratory Animal Care–approved facility at the University of California, San Diego (UCSD), and all procedures were approved by the UCSD Institutional Animal Care and Use Committee. WT C57BL6/J (CD45.2+) and P14 TCR transgenic (CD45.1 or CD45.1.2), both maintained on a C57BL6/J background) mice were bred at UCSD or purchased from the Jackson Laboratories. Ddxs5−/− mice were obtained from Dr. Frances Fuller-Pace’s laboratory (University of Dundee) and have been previously described (77). To obtain congenically distinct P14 Ddxs5−/−/CD4-Cre+ and P14 Ddxs5−/−/CD4-Cre− mice, Ddxs5−/− mice were crossed to P14 CD4-Cre+ mice (either CD45.1+ or CD45.2+). All mice were used from 6 to 9 weeks of age, male mice were used as recipients, and male or female mice were used as donors in adoptive transfer experiments.

Antibodies, flow cytometry, and cell sorting

Cells were stained for 10 min on ice with the following antibodies: Vα2 (B20.1), CD8α (53–6.7), CD8β (YTS156.7.7), CD45.1 (A20), CD45.2 (104), CD44 (1 M7), CX3CR1 (SA011F11), CD127 (A7R34), CD27 (LG.3A10), CD69 (H1.2F3), CD103 (2E7), CD28 (37.51), CD25 (PC61), and KLRG1 (2F1/KLRG1) all purchased from BioLegend. In some experiments, cells were stained with H-2Dβ GP33–41 tetramer [obtained from the National Institutes of Health (NIH) Tetramer Core Facility] for 1 hour at room temperature before staining with cell surface antibodies. Samples were then stained in Fixable Viability Dye eFluor780 (Thermo Fisher Scientific) or Ghost Violet 510 (Tonbo Biosciences) at 1:1000 on ice for 30 min. For staining for Ezh2, cells were fixed and permeabilized using the FoxP3/Transcription Factor Staining Buffer Kit (Thermo Fisher Scientific) on ice for 30 min. For staining for TNF-α (MP6-XT22), and IL-2 (JES6-5H4) antibodies (all from BioLegend), 6 to 9 weeks of age, male mice were used as recipients, and male or female mice were used as donors in adoptive transfer experiments.

Naïve T cell transfer and infection

Splenocytes were collected from naïve CD45.1+ or CD45.1.2+ P14 mice and stained with antibodies against Vα2, CD8, and CD45.1.
CD8⁺ T cell isolation

For isolation of CD8⁺ T cells from the spleen, spleens were collected and dissociated to yield a cell suspension before treatment with Red Blood Cell Lysing Buffer Hybri-Max (Sigma-Aldrich). For isolation of CD8⁺ T cells from the small intestinal epithelium, Peyer’s patches were removed, and the tissue was cut longitudinally and washed of luminal contents. The tissue was then cut into 1-cm pieces that were incubated while shaking in DTE buffer [dithioerythritol (1 mM); Thermo Fisher Scientific) in 10% Hanks’ balanced salt solution and 10% Hepes bicarbonate] at 37°C for 30 min. Cells within the supernatant were collected and passed through a 44/67% Percoll density gradient to enrich for lymphocytes.

10x Genomics library preparation and sequencing

Activated P14 T cells (CD8⁺Vα2⁺CD45.1⁺CD44⁺) were sorted from the spleen or siIEL compartment and resuspended in phosphate-buffered saline (PBS) + 0.04% (w/v) bovine serum albumin. About 10,000 cells per sample were loaded into Single A Chips (10x Genomics) and partitioned into Gel Bead In-Emulsions in a Chromium Controller (10x Genomics). Single-cell RNA libraries were prepared according to the 10x Genomics Chromium Single Cell 3’ Reagent Kits v2 User Guide and sequenced on a HiSeq 4000 (Illumina).

Generation of shRNA-encoding retrovirus, transduction of CD8⁺ T cells, and adoptive transfer for analysis of gene KD

shERWOOD-designed UltramiR sequences targeting Junb, Nr4a2, Prc1, Fosl2, or Zfp36l2 (KD) or control (nontargeting) constructs in LMP-d Ametrine vector were purchased from transOMIC technologies. To generate retroviral particles, human embryonic kidney-293T cells were plated in 10-cm plates 1 day before transfection and individually transfected with 10 µg of each shRNA retroviral construct and 5 µg of pCL-Eco using TransIT-LT1 (Mirus). The retroviral supernatant was collected and pooled at 48 and 72 hours after transfection, in vitro–activated P14 T cells were transduced with retrovirus encoding KD shRNA, and congenically distinct in vitro–activated P14 T cells were transduced with control retrovirus. To activate P14 T cells in vitro, lymphocytes were collected from spleens and lymph nodes of naïve CD45.1⁺ and CD45.1.2⁺ P14 TCR transgenic mice and negatively enriched for CD8⁺ T cells using the CD8α⁺ T Cell Isolation Kit and LS MACS Columns (Miltenyi Biotec). CD8⁺ T cells (1 × 10⁶ per well) were plated in 48-well flat-bottom plates and precoated with goat anti-hamster IgG (100 µg/ml; H⁺L, Thermo Fisher Scientific), followed by 5 µg/ml each anti-CD3 (clone 3C11, BioXCell) and anti-CD28 (clone 37.51, BioXCell). Eighteen hours after activation, cells were transduced with retroviral supernatant supplemented with polybrene (8 µg/ml; Millipore) by spinfection for 90 min at 900 rcf (relative centrifugal force) at room temperature. After spinfection, the retroviral supernatant was removed and replaced with culture medium [Iscove’s modified Dulbecco’s medium + 10% fetal bovine serum (v/v) + 2 mM glutamine + penicillin (100 U/ml) + streptomycin (100 µg/ml) + 55 mM β-mercaptoethanol], and cells were rested for 2 hours at 37°C. Cells transduced with each individual construct were then pooled, washed three times with PBS, counted, and mixed to obtain a 1:1 ratio of transduced KD and transduced control cells, based on previously tested transduction efficiency of each retrovirus. Total P14 T cells (5 × 10⁶) were then adoptively transferred into congenically distinct hosts that were infected with 2 × 10⁶ PFU LCMV-Arm 1 hour later. Cells (1 × 10⁶) from this mixture were returned to culture with recombinant IL-2 (100 U/ml) and analyzed 18 hours later by flow cytometry to determine the input ratio of transduced KD/control P14 T cells. Twenty-two to 26 days after infection, splenocytes and siIEL were collected, as described above, and analyzed by flow cytometry to determine the ratio of KD/control cells within each subset of transduced P14 T cells.

Statistical analysis of flow cytometry data

Statistical analysis of flow cytometry data was performed using Prism software (GraphPad). P < 0.05 was considered significant. Statistical details for each experiment are provided within the relevant figure legends and the raw data file.

Quantitative and statistical analysis of scRNA-seq data

scRNA-seq mapping

 Reads from scRNA-seq were aligned to mm10 and collapsed into unique molecular identifier (UMI) counts using the 10x Genomics Cell Ranger software (version 2.1.0). All samples had sufficient numbers of genes detected (>1000), a high percentage of reads mapped to the genome (>70%), and a sufficient number of cells detected (>1000).

Cell and gene filtering

Raw cell-reads were loaded to R using the cellrangerRkit package. The scRNA-seq dataset was then further filtered on the basis of gene numbers and mitochondria gene counts to total counts ratio. To ensure that the samples with more cells would not dominate the downstream analysis, we randomly selected a portion of the cells that passed filtering for downstream analysis. We randomly selected ~2000 cells from each library for downstream analysis. After cell filtering and sampling, we filtered genes by removing genes that did not express >1 UMI in more than 1% of the total cells.

scRNA-seq dataset normalization and preprocessing

Five cell-gene matrices were generated as follows:

1) Raw UMI matrix.

2) UPM matrix. The raw UMI matrix was normalized to get UMIs per million reads (UPM) and was then log₂-transformed. All downstream differential analysis was based on the UPM matrix. The prediction models were also based on the UPM matrix, as other normalizations are very time consuming for large datasets.

3) MAGIC (Markov Affinity–Based Graph Imputation of Cells) matrix. The UPM matrix was further permuted by MAGIC (78). R package Rmagic 1.0.0 was used, and all options were kept as default. MAGIC aims to correct the dropout effect of scRNA-seq data; thus, we used the MAGIC-corrected matrix for visualizing the gene expression pattern rather than using the UPM matrix. All gene expression heatmaps and gene expression overlay on tSNE plots were based on the MAGIC matrix.

4) Supercell matrix. We merged 50 cells to create a “super” cell and used the supercell matrix as the input for WGCNA (weighted gene correlation network analysis) and cell-type annotation analysis.
This approach enabled us to bypass the issue of gene dropouts with scRNA-seq and is equivalent to performing WGCNA on thousands of pseudobulk samples. We first calculated the mutual nearest neighbor network with \( k \) set to 15, and then cells that were not mutual nearest neighbors with any other cells were removed as outliers. We randomly selected "n" cells in the UPM matrix as the seed for supercells. The expression of each supercell was equal to the average expression of its seed and the 50 nearest neighbor cells of its seed. We derived 7400 supercells from the dataset, so each single cell was covered \(~10\) times.

**scRNA-seq dataset dimension reduction**

Top variable genes, principal components analysis (PCA), and tSNE were calculated by Seurat version 2.3.4 functions: FindVariableGenes, RunPCA, and RunTSNE (79). Only the top 3000 genes were considered in the PCA calculation and only the top 25 PCs were used in tSNE. Louvain clustering was performed by Seurat’s FindClusters function based on the top 25 PCs, with resolution set to 2. UMAP was calculated by R packages umap_0.2 with default setting.

**Differential gene expression analysis**

DE genes were identified by performing pairwise comparison using two-sided Wilcoxon test and the FindAllMarkers function of Seurat. The threshold for DE genes was \( P < 0.05\) and an absolute fold change of \(>2\).

**WGCNA analysis**

We performed WGCNA (version 1.63) analysis on spleen cells alone, siIEL cells alone, or all cells considered together. The supercell matrices were used as the input to boost performance. Only the top 5000 variable genes were considered in this analysis. SoftPower was used as the input to boost performance. Only the top 25 PCs were used in tSNE. Louvain clustering was performed by Seurat’s FindClusters function based on the top 25 PCs, with resolution set to 2. UMAP was calculated by R packages umap_0.2 with default setting.

**Annotating single cells with bulk RNA-seq signatures**

The log2 transcripts per million data from bulk RNA-seq datasets were compared with the scRNA-seq supercell matrix. Bulk cell population RNA-seq samples were first grouped into different sets according to their mutual similarities. For each bulk RNA-seq sample set, the mean expression was first calculated. The first correlation was calculated between all the supercells and the mean expression from the bulk RNA-seq dataset. On the basis of the distribution of the first correlation, we were able to identify a group of supercells that were most similar to the mean expression of the bulk sample. To further identify the small differences between bulk RNA-seq expression within a given set, we removed the set mean from the bulk RNA-seq and the mean from the most similar group of supercells and then calculated the second correlation between the supercells and bulk RNA-seq. On the basis of the second correlation, we annotated the supercells with each bulk sample label.

**PAGA trajectory analysis**

The single-cell trajectories were constructed using the PAGA algorithm (54) implemented in the scanpy package (80). The UMI counts of the filtered cells and genes (filtered by the same approach mentioned above) were normalized by sequence depth before trajectory construction. In the construction, the knn-graphs were built with the scanpy.pp.neighbors function (\text{n\_neighbors} = 7 and \text{n\_pcs} = 20), and PAGA graphs were constructed using the scanpy.tl.paga function with default settings. In the PAGA graphs, each node represents a cell partition (a group of cells), and the edges between nodes represent the connection between these nodes measured by PAGA, whereas the strength of the edge describes the degree of connection (connectivity), which varies from 0 to 1. Only the edges with connectivity larger than 0.2 were shown in the final PAGA graphs, and graphs were based on clusters defined with Louvain clustering.

**SUPPLEMENTARY MATERIALS**

immunology.sciencemag.org/cgi/content/full/5/47/eaaz6894/DC1

Fig. S1. Sorting of P14 CD8\(^+\) T cells from spleen and siEL for scRNA-seq.

Fig. S2. Circulating and siEL CD8\(^+\) T cells are transcriptionally distinct at all time points after infection.

Fig. S3. Differential gene expression of circulating and siEL CD8\(^+\) T cells at day 4 after infection.

Fig. S4. Shared core transcriptional program among splenic and siEL CD8\(^+\) T cells.

Fig. S5. Components of the T\(_{RM}\) cell-enriched transcriptional signature.

Fig. S6. shRNA knockdown of putative regulators of T\(_{RM}\) cell differentiation.

Fig. S7. PAGA trajectory analysis of T\(_{RM}\) cell differentiation.

Table S1. DE genes by splenic and siEL CD8\(^+\) T cells at day 4 after infection.

Table S2. Genes in Spleen and IEL Modules.

Table S3. Genes in Combined Modules.

Table S4. DE genes by day 60 versus day 90 T\(_{RM}\)-like cells.

Table S5. DE genes by Cluster 3 versus Cluster 29 cells.

Table S6. DE genes by Cluster 17 versus Cluster 19 cells.

Table S7. DE genes by Cluster 16 versus Cluster 20 cells.

Table S8. Raw data file (Excel spreadsheet).

View/request a protocol for this paper from Bio-protocol.

**REFERENCES AND NOTES**


Early precursors and molecular determinants of tissue-resident memory CD8+ T lymphocytes revealed by single-cell RNA sequencing

Nadia S. Kurd, Zhaoren He, Tiani L. Louis, J. Justin Milner, Kýla D. Omilusik, Wenhao Jin, Matthew S. Tsai, Christella E. Widjaja, Jad N. Kanbar, Jocelyn G. Olivera, Tiffani Tysl, Lauren K. Quezada, Brigid S. Boland, Wendy J. Huang, Cornelis Murre, Ananda W. Goldrath, Gene W. Yeo and John T. Chang

DOI: 10.1126/sciimmunol.aaz6894

Stepping down resident memory lane

The antigen-specific CD8+ T response to microbial infection includes the differentiation of a subset of CD8+ T cells into tissue-resident memory (T_RM) cells that stop circulating and become confined within a nonlymphoid tissue. Kurd et al. used single-cell RNA sequencing of mouse CD8+ T cells at multiple time points during the first 90 days after viral infection to characterize how this differentiation process unfolds in the small intestine and tracks the emergence of heterogeneity among T_RM cells. They found evidence for T_RM cell precursors in the intestine by 4 days after infection and identified several putative regulators of T_RM cell differentiation. The results of this study provide a valuable transcriptomic atlas that will facilitate further investigation into immune functions provided by T_RM cells.