

Single-cell transcriptomics of alloreactive CD4⁺ T cells over time reveals divergent fates during gut graft-versus-host disease

Jessica A. Engel,¹ Hyun Jae Lee,¹ Cameron G. Williams,¹ Rachel Kuns,¹ Stuart Olver,¹ Lianne I.M. Lansink,¹ Megan S.F. Soon,¹ Stacey B. Andersen,² Joseph E. Powell,^{3,4} Valentine Svensson,⁵ Sarah A. Teichmann,⁶ Geoffrey R. Hill,^{1,7,8} Antiopi Varelias,^{1,9} Motoko Koyama,⁷ and Ashraful Haque^{1,10}

¹QIMR Berghofer Medical Research Institute, Herston, Brisbane, Queensland, Australia. ²Institute for Molecular Bioscience, University of Queensland, St. Lucia, Queensland, Australia. ³Garvan-Weizmann Centre for Cellular Genomics, Sydney, New South Wales, Australia. ⁴UNSW Cellular Genomics Futures Institute, University of New South Wales, Sydney, New South Wales, Australia. ⁵Flagship Labs 60 Inc., Cambridge, Massachusetts, USA. ⁶Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, United Kingdom. ⁷Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. ⁸Division of Medical Oncology, University of Washington, Seattle, Washington, USA. ⁹Faculty of Medicine, University of Queensland, St. Lucia, Queensland, Australia. ¹⁰Department of Microbiology and Immunology, University of Melbourne, located at The Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia.

Acute gastrointestinal (GI) graft-versus-host disease (GVHD) is a primary determinant of mortality after allogeneic hematopoietic stem cell transplantation (alloSCT). The condition is mediated by alloreactive donor CD4⁺ T cells that differentiate into pathogenic subsets expressing IFN- γ , IL-17A, or GM-CSF and is regulated by subsets expressing IL-10 and/or Foxp3. Developmental relationships between Th cell states during priming in mesenteric lymph nodes (mLNs) and effector function in the GI tract remain undefined at genome scale. We applied scRNA-Seq and computational modeling to a mouse model of donor DC-mediated GVHD exacerbation, creating an atlas of putative CD4⁺ T cell differentiation pathways in vivo. Computational trajectory inference suggested emergence of pathogenic and regulatory states along a single developmental trajectory in mLNs. Importantly, we inferred an unexpected second trajectory, categorized by little proliferation or cytokine expression, reduced glycolysis, and high *tcf7* expression. TCF1^{hi} cells upregulated $\alpha 4\beta 7$ before gut migration and failed to express cytokines. These cells exhibited recall potential and plasticity following secondary transplantation, including cytokine or Foxp3 expression, but reduced T cell factor 1 (TCF1). Thus, scRNA-Seq suggested divergence of alloreactive CD4⁺ T cells into quiescent and effector states during gut GVHD exacerbation by donor DC, reflecting putative heterogeneous priming in vivo. These findings, which are potentially the first at a single-cell level during GVHD over time, may assist in examination of T cell differentiation in patients undergoing alloSCT.

Authorship note: JAE, HJL, and CGW are co-first authors.

Conflict of interest: The authors have declared that no conflict of interest exists.

Copyright: © 2020, American Society for Clinical Investigation.

Submitted: March 9, 2020

Accepted: May 21, 2020

Published: July 9, 2020.

Reference information: *JCI Insight*. 2020;5(13):e137990.
<https://doi.org/10.1172/jci.insight.137990>.

Introduction

Allogeneic hematopoietic stem cell transplantation (alloSCT) is a curative therapy for a range of acute leukemias, owing to the capacity of donor T cells within the transplant to kill tumor cells — also known as the graft-versus-leukemia (GVL) effect. Unfortunately, donor T cells can also damage noncancerous host tissue, particularly the gastrointestinal (GI) tract, liver, and skin, causing the serious condition, graft-versus-host disease (GVHD). Acute GVHD in the GI tract remains the primary determinant of GVHD severity and risk of death (1). Thus, a primary goal for alloSCT is the prevention of acute gut GVHD while preserving GVL. We previously showed in preclinical models that donor CD4⁺ T cells are initially activated by recipient nonhematopoietic antigen-presenting cells (APCs) within the gut, including epithelial cells that upregulate class II MHC molecules (2, 3). Donor-derived colonic DCs also prime donor CD4⁺ T cells in mesenteric lymph nodes (mLNs) and trigger Th cell differentiation, which serves to amplify and exacerbate GVHD (4). Using T cell receptor (TCR) transgenic T cells specific for a single allopeptide, TEa cells (5), we revealed that donor CD4⁺ T cells within the same alloSCT recipient differentiate into multiple cellular states that express proinflammatory and pathogenic Th1/Th17-associated cytokines, including IFN- γ and

IL-17A, or the master transcription factor for induced regulatory T (iTreg) cells, Foxp3 (4). Although cytokines such as IL-6 and IL-12 control differentiation of donor CD4⁺ T cells, fate-mapping studies based on the *IL-17a* promoter suggested complex dynamic relationships between apparent helper subsets (6). Thus, differentiation of alloreactive donor CD4⁺ T cells is characterized by complexity, both in terms of dynamics and multiple cellular states adopted, neither of which have been explored at genome scale.

scRNA-Seq enables unbiased genome-wide assessment of individual T cells without reliance on predetermined protein markers or genes. ScRNA-Seq was previously used to examine heterogeneity in CD4⁺ T cells isolated from *IL-17a* reporter mice undergoing experimental autoimmune encephalomyelitis (EAE) (7). Subsequently, scRNA-Seq was used to examine CD4⁺ T cell differentiation during house dust mite-induced allergy (8), protozoan parasite infection (9), as well as CD8⁺ T cells in viral infections and cancer (10–12). Many of these studies were cross-sectional, offering insight into heterogeneity among clonal TCR transgenic cells at a single time point. We previously examined CD4⁺ T cell transcriptomes over a range of time points during experimental malaria and employed computational approaches to reconstruct the dynamics of Th1 versus T follicular helper (Tfh) cell differentiation. Using an approach based on Bayesian Gaussian process latent variable modeling (bGPLVM), we identified a bifurcation point between 2 trajectories and revealed a role for T cell extrinsic factors in governing Th1/Tfh cell fate (9). More recently, we used scRNA-Seq to reveal heterogeneity and tissue adaptation of thymic Tregs and colonic CD4⁺ T cells in mice and humans during steady state (13, 14). Here, we examined donor DC-mediated differentiation of alloreactive donor CD4⁺ T cells during exacerbation of acute gut GVHD using droplet-based scRNA-Seq and computational modeling.

Results

Alloreactive donor CD4⁺ T cells acquire heterogeneous proinflammatory, regulatory, and uncharacterized states during the exacerbation phase of acute gut GVHD. We previously established a preclinical model of acute GVHD exacerbation (4), in which late donor CD4⁺ T cell responses were targeted to a single host-derived alloreactive peptide presented by donor class II MHC molecules. In this model, TEa TCR transgenic CD4⁺ T cells (B6 background) exhibit specificity for a BALB/c-derived Ea peptide from the class II MHC molecule, I-E^d, when presented by the donor (B6) class II MHC molecule, I-A^b (Figure 1A). Here, we exposed BALB/c mice to total body irradiation and provided an MHC-mismatched B6 bone marrow transplant (containing donor T cells). As in our previous study (4), we opted not to model the GVL effect by instilling leukemic cells, considering that tumor burdens would be relatively low during the exacerbation phase of GVHD. Twelve days later, once alloreactive donor APCs developed (4), TEa cells were transferred. By day 4, TEa cell priming occurred specifically within the mLNs and triggered differentiation into subsets with a strong capacity to express proinflammatory cytokines IFN- γ or IL-17A on restimulation, or in rarer cases (approximately 1%) express Foxp3 (Figure 1, A and B). We also examined the frequency of TEa cells expressing IFN- γ , IL-17A, or Foxp3 directly ex vivo, without further stimulation (Figure 1C). Although some TEa cells expressed IFN- γ , IL-17A, or Foxp3 directly ex vivo (Figure 1C), most did not at this early time point. Therefore, although Th cell differentiation occurred, use of 3 canonical Th markers was insufficient for characterizing the fate of most TEa cells in mLN.

Droplet-based single-cell RNA-Seq and computational modeling reveals divergent fates in donor CD4⁺ T cells. To examine donor CD4⁺ T cell differentiation without employing preselected markers, we opted for droplet-based scRNA-Seq. TEa cells were transferred into BALB/c mice 13 days after irradiation, provision of alloSCT, and subsequent GVHD initiation. TEa cells were then recovered from mLN at days 1, 2, 3, and 4 after transfer (Figure 2A and Supplemental Figure 1, A and B; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.137990DS1>). Nontransferred control TEa cells (day 0) were also examined. Flow cytometric assessment of sorted TEa cells, pooled from mice at each time point, revealed rapid, transient upregulation of CD69 and evidence of cell division by CFSE dilution by day 2 (Supplemental Figure 1B). This was followed on days 3 and 4 by complete loss of CFSE, indicative of dramatic clonal expansion, as well as substantial upregulation of the gut homing integrin, $\alpha 4\beta 7$, in many, but not all TEa cells. These data confirmed that TEa cell activation had occurred in mLN and suggested emerging heterogeneity. Therefore, TEa cells were processed for droplet-based scRNA-Seq. After excluding poor-quality single-cell transcriptomes (Supplemental Figure 2, A and B), we advanced 22,854 high-quality TEa samples for further analysis. We noted a substantial increase in the average number of genes detected per cell from day 0 through days 1 and 2, which dropped by day 4 (Supplemental Figure 2A). This was consistent with our previous study in which CD4⁺ T cells more than doubled the number of detected genes during clonal expansion (9).

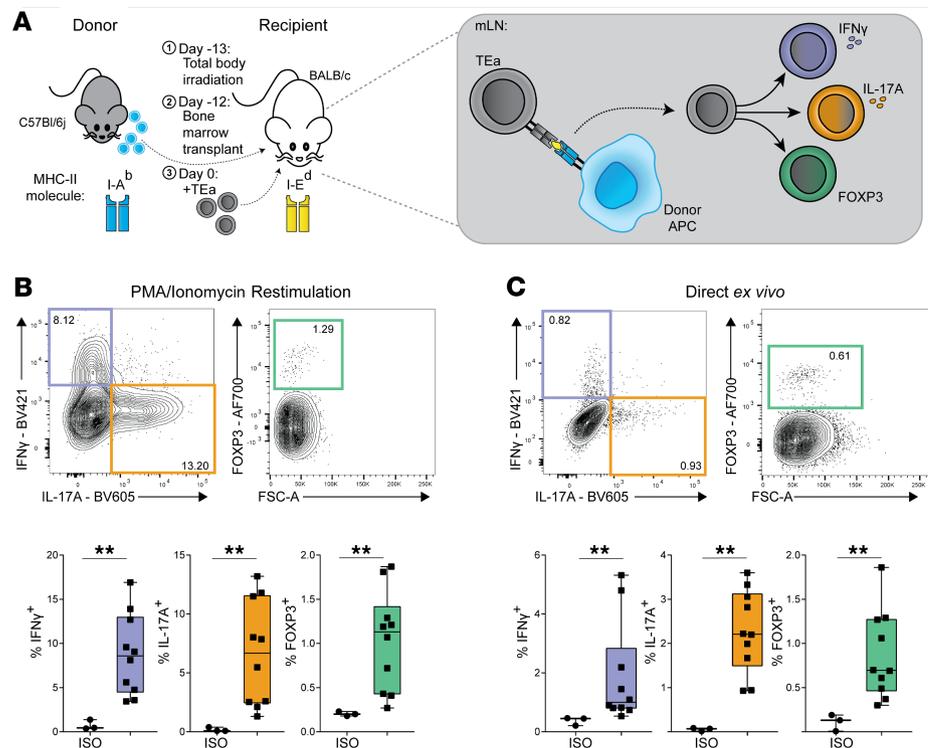


Figure 1. Alloreactive donor CD4⁺ TEa T cells acquire proinflammatory and regulatory states during acute GVHD exacerbation. (A) Schematic for model of acute gut-mediated GVHD exacerbation where donor CD4⁺ T cells respond to host alloantigenic peptide presented within donor MHC class II, which in turn drives CD4⁺ T cell expansion in the mesenteric lymph node (mLN). (B and C) Representative FACS plots for IFN- γ and IL-17A production and Foxp3 expression on TEa cells from the mLN at day 4 after transfer, with PMA/ionomycin restimulation (B) or directly ex vivo (C). Graphs represent IFN- γ ⁺, IL-17A⁺, or Foxp3⁺ TEa cell percentages. Data shown are combined from 3 replicate experiments ($n = 10$ mice) and are represented as box-and-whisker plots, with bounds from 25th to 75th percentile, median line, and whiskers ranging from minimum to maximum values. Statistical analysis was performed between the isotype control (ISO) and respective cytokine samples using a Mann-Whitney U test. ** $P \leq 0.01$. GVHD, graft-versus-host disease.

Uniform manifold approximation and projection (UMAP) (15) visualization after principal component analysis (PCA) suggested that TEa cells (days 0 and 1), although relatively homogeneous within their own time points, were distinct from each other, as well as cells assessed at days 2–4 (Supplemental Figure 2C). In contrast, from days 2 through 4, there was evidence for both transcriptomic overlap between cells as well as transcriptomic progression (Supplemental Figure 2C). Therefore, we sought to model potential developmental trajectories within the scRNA-Seq data. We used nonlinear probabilistic PCA, termed bGPLVM (9), which embedded the data in low-dimensional space and reordered transcriptomes independently of the time point of capture. Previously, we employed bGPLVM on a relatively small number of cells (<1000) (9). Here, it was evident that bGPLVM was scalable to approximately 25,000 cells. Running bGPLVM iteratively 10 times on the data set yielded similar learned embeddings, indicating the stability of the output (Supplemental Figure 3). Variability between cells was clearly observable along 3 latent variables, with interpretable variation also evident in the first 2 latent variables, thus permitting modeling in 2 dimensions (Figure 2B).

We next inferred differentiation trajectories based on transcriptomic similarity between cells. We initially used Slingshot (16), a top-ranked trajectory inference tool (17), on the bGPLVM embedding, with a start point specified in day 1 and with day 0 omitted owing to transcriptomic distance from the rest of the data (Supplemental Figure 4, A and B). Slingshot suggested 3 potential trajectories in the bGPLVM space, 2 of which (trajectories I and II) progressed through days 2, 3, and 4, whereas trajectory III terminated in day 2 (Figure 2B). Consistent with flow cytometric data (Supplemental Figure 1), *Cd69* expression gradually reduced along trajectories I and II, whereas *Itgb7*, encoding subunit $\beta 7$ of integrin $\alpha 4\beta 7$, increased (Figure 2, C and D). Notably, trajectory II ceased cell-cycling gene expression, including *Mki67* (encoding Ki-67) (Figure 2, C and D), and exhibited lower aerobic glycolysis gene expression compared with trajectory I.

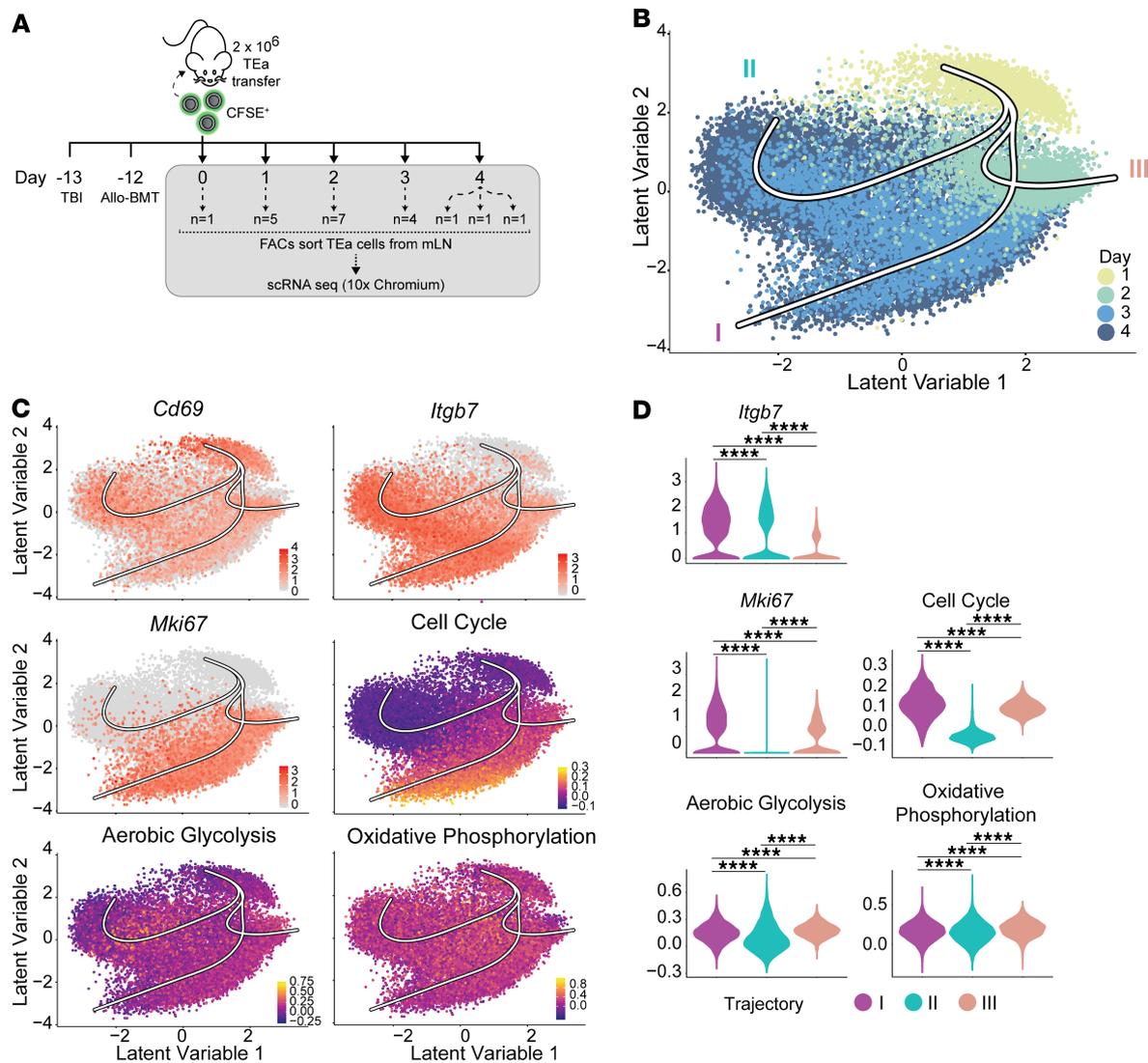


Figure 2. Droplet-based scRNA-Seq and computational modeling of TEa T cells over time suggests diverging fates in mLN. (A) Schematic of scRNA-Seq experiment used to study donor CD4⁺ T cell differentiation. TEa cells were recovered from the mLN at various time points and cells pooled from multiple mice (n = number of mice pooled; at day 4, samples from 3 mice ran separately) for droplet-based scRNA-Seq using the 10x Chromium. (B) bGPLVM visualization of TEa cells from days 1 to 4 overlaid with the developmental trajectories identified by Slingshot (trajectory I, II, and III). (C) Visualization of *Cd69*, *Mki67*, and *Itgb7* expression, or the cell cycle, aerobic glycolysis, and oxidative phosphorylation gene signature scores, in bGPLVM overlaid with trajectories. (D) Violin plots showing expression of genes or gene signature scores as described in (C) for each trajectory. Statistical analysis was performed using a Wilcoxon-rank sum test between trajectories. **** $P \leq 0.0001$. bGPLVM, Bayesian Gaussian process latent variable model; mLN, mesenteric lymph node.

We also used Slingshot on UMAP of PCs and found 4 trajectories, 2 of which largely superimposed on each other suggesting, as for bGPLVM, 3 main trajectories (Supplemental Figure 5A). Two of these progressed from days 2–4, with one largely devoid of cellular proliferation and the other exhibiting strong *Mki67* expression and other cycling genes (Supplemental Figure 5, B and C). To further test for trajectories, we employed other highly ranked—computational approaches (17), including Monocle 2, Velocyto (18), and PAGA (19). In all cases (Supplemental Figure 6), one main bifurcation event appeared within days 2–3, leading to 2 trajectories. Therefore, taken together, our analysis suggested 2 trajectories had emerged in TEa cells, which differed from each other in expression of genes, including those related to cell cycling and aerobic glycolysis.

Proinflammatory and regulatory effector fates emerge within one trajectory, revealing a second, quiescent Tcf7-expressing fate. We next assessed expression of the canonical markers *Ifng*, *Il17a*, and *Foxp3* within our bGPLVM/Slingshot model (Figure 3A). Surprisingly, the markers were expressed only in trajectory I. In addition, we noted significant *Il17a* and *Il17f* expression (Figure 3, *Il17f* not shown) in areas shared by

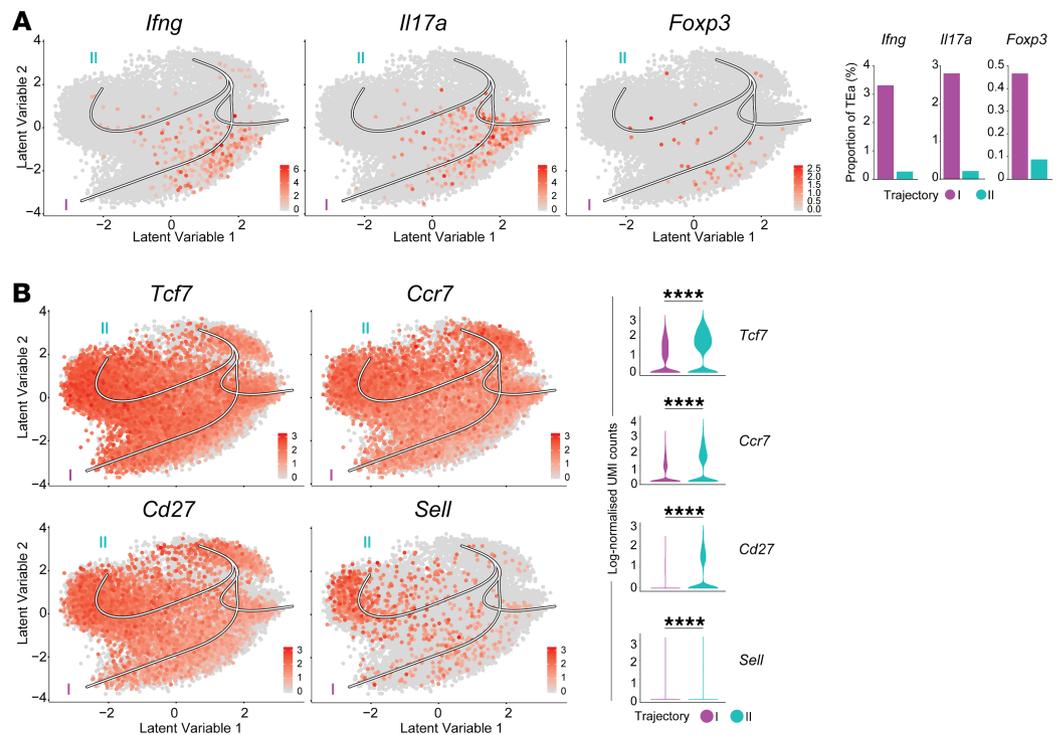


Figure 3. *Ifng*, *Il17a*, and *Foxp3* expressing fates emerge along one trajectory, identifying a second *Tcf7*^{hi} trajectory. (A) bGPLVM visualization of TEa cells expressing *Ifng*, *Il17a*, and *Foxp3* with Slingshot trajectories overlaid. Bar graphs show the proportion of cells within trajectory I or II that express *Ifng*, *Il17a*, or *Foxp3*. (B) bGPLVM visualization of TEa cells expressing *Tcf7*, *Ccr7*, *Cd27*, and *Sell* with Slingshot trajectories overlaid. Violin plots show the level of expression for each gene in trajectories I and II. Statistical analysis was performed using a Wilcoxon-rank sum test between trajectories. **** $P \leq 0.0001$. bGPLVM, Bayesian Gaussian process latent variable model

trajectories III and I (Figure 3A). These observations pointed to early, partly transient activation of the *Il17a* promoter at day 2, followed by *Ifng* and *Foxp3* upregulation. Most important, we saw little evidence of Th1, Th17, and iTreg cell types emerging from different developmental trajectories, either when examined using our previous Th gene modules or when relatively sparse *Ifng*, *Il17a*, and *Foxp3* expression were imputed using Adaptively-thresholded Low-Rank Approximation (ALRA) (20) (Supplemental Figure 7). Instead, our analysis suggested proinflammatory and regulatory fates had emerged within trajectory I.

Interestingly, trajectory II lacked *Ifng*, *Il17a*, and *Foxp3* expression (Figure 3A). Differential gene expression analysis between cells in trajectory II versus trajectory I revealed *Tcf7* as a top transcription factor associated with trajectory II, as well as elevated expression of central memory-associated *Ccr7* and *Sell* (encoding CD62L) (Figure 3B and Table 1 and Supplemental Figure 1). These observations were also seen in our UMAP/Slingshot model and in a second scRNA-Seq experimental repeat (Supplemental Figure 8 and Table 1). Thus, given an absence of cell cycle activity, lower gene expression of aerobic glycolysis, absence of proinflammatory or regulatory gene expression, and expression of memory or stem-like genes including *Tcf7*, we inferred that trajectory II contained quiescent TEa cells that had gone through a clonal burst in mLN, but had not acquired effector function.

We further examined genes expressed at a higher level in *Tcf7*^{hi} TEa cells compared with effector counterparts (Table 1 and Supplemental Table 1). We noted several genes previously reported to support T cell quiescence, including *Btg1*, which mediates deadenylation and degradation of mRNA (21), *Mxd4*, a Myc-antagonist (22), *Samhd1*, which mediates degradation of dNTPs required for DNA synthesis, *Laptm5*, which has been reported to mediate lysosomal targeting and degradation of components of the CD3 complex (23), and *Ms4a4b*, which may limit cell cycle time of T cells (24). Similarly, *Gimap3* encodes a GTP-binding protein linked with T cell longevity (25), whereas *Izumo1R*, encoding folate receptor 4, was recently reported on long-lived memory Tfh cells (26). In addition, we noted genes previously reported to control Th cell fate, including *Klf2* (27, 28), and trafficking, such as *Ccr7* and *Rgs10* (29, 30). Thus, our transcriptomic data were broadly consistent with *Tcf7*^{hi} TEa cells exhibiting a phenotype associated with central memory, quiescence, and longevity.

Table 1. Top 50 genes upregulated in cluster 0 in trajectory II compared with cluster 3 in trajectory I

Gene	Average log fold change	Adjusted P value
Malat1 ^A	2.07	0
Btg1 ^A	2.04	0
Ifi2712a ^A	1.79	0
Rgs10 ^A	1.62	2.45 × 10 ⁻²⁹²
Izumo1r ^A	1.59	4.81 × 10 ⁻¹²⁶
Igfbp4 ^A	1.54	8.47 × 10 ⁻⁶⁵
Tcf7 ^A	1.52	1.02 × 10 ⁻¹⁸⁷
Ypel3 ^A	1.42	1.97 × 10 ⁻¹³⁸
Npc2	1.29	6.60 × 10 ⁻²¹⁰
Shisa5	1.24	0
Cox7a2l ^A	1.19	2.52 × 10 ⁻²⁸²
Ms4a6b	1.17	2.83 × 10 ⁻¹⁷⁶
Ms4a4b	1.14	1.33 × 10 ⁻¹³⁶
Ccr7 ^A	1.11	3.58 × 10 ⁻⁶⁶
Gm26740	1.10	2.49 × 10 ⁻⁷⁴
Use1 ^A	1.04	1.20 × 10 ⁻⁹⁸
Limd2	1.03	1.75 × 10 ⁻²⁶⁴
Gm8369 ^A	1.01	3.42 × 10 ⁻⁵⁴
Laptm5	1.01	2.34 × 10 ⁻¹³⁹
Id3	1.01	5.28 × 10 ⁻²⁵
Gimap6	1.01	2.92 × 10 ⁻⁹⁸
Gbp2 ^A	1.00	3.15 × 10 ⁻¹³
Stat1 ^A	0.99	3.06 × 10 ⁻⁶⁷
Saraf	0.98	2.09 × 10 ⁻⁷⁹
Evl	0.97	7.43 × 10 ⁻⁵²
Sesn3 ^A	0.94	6.70 × 10 ⁻⁵²
Pold4	0.92	6.78 × 10 ⁻⁴⁰
Cd52	0.91	4.55 × 10 ⁻²⁰⁴
Gltscr2	0.90	3.75 × 10 ⁻¹³⁶
Gimap3	0.90	3.02 × 10 ⁻⁶³
Mxd4 ^A	0.89	4.91 × 10 ⁻⁴⁹
Fyb ^A	0.89	1.53 × 10 ⁻⁷⁹
Tspan32	0.89	1.58 × 10 ⁻³⁵
Slamf6 ^A	0.88	1.30 × 10 ⁻⁵⁵
Ltb	0.88	3.57 × 10 ⁻¹⁰⁸
Pfdn5	0.88	4.07 × 10 ⁻²⁰³
Samhd1	0.88	1.03 × 10 ⁻²⁹
Cd3d	0.86	1.82 × 10 ⁻²¹⁶
Cd27	0.85	2.21 × 10 ⁻³⁷
Hmha1	0.85	1.01 × 10 ⁻³²
Nsg2 ^A	0.84	4.70 × 10 ⁻⁴⁸
Eva1b	0.83	1.14 × 10 ⁻³⁰
Satb1 ^A	0.82	6.48 × 10 ⁻⁶⁴
Klf2	0.82	2.78 × 10 ⁻⁵⁵
Pnrc1 ^A	0.82	6.74 × 10 ⁻²⁷
Asap1 ^A	0.81	5.99 × 10 ⁻³⁰
Sepp1	0.81	5.30 × 10 ⁻²⁴
Ankrd12 ^A	0.80	1.24 × 10 ⁻¹⁹
Cd3g	0.80	1.84 × 10 ⁻¹⁶⁹
Ndr3	0.79	4.91 × 10 ⁻²⁵

^AGenes detected within the top 50 genes in an analogous comparison in a second independent scRNA-Seq experiment.

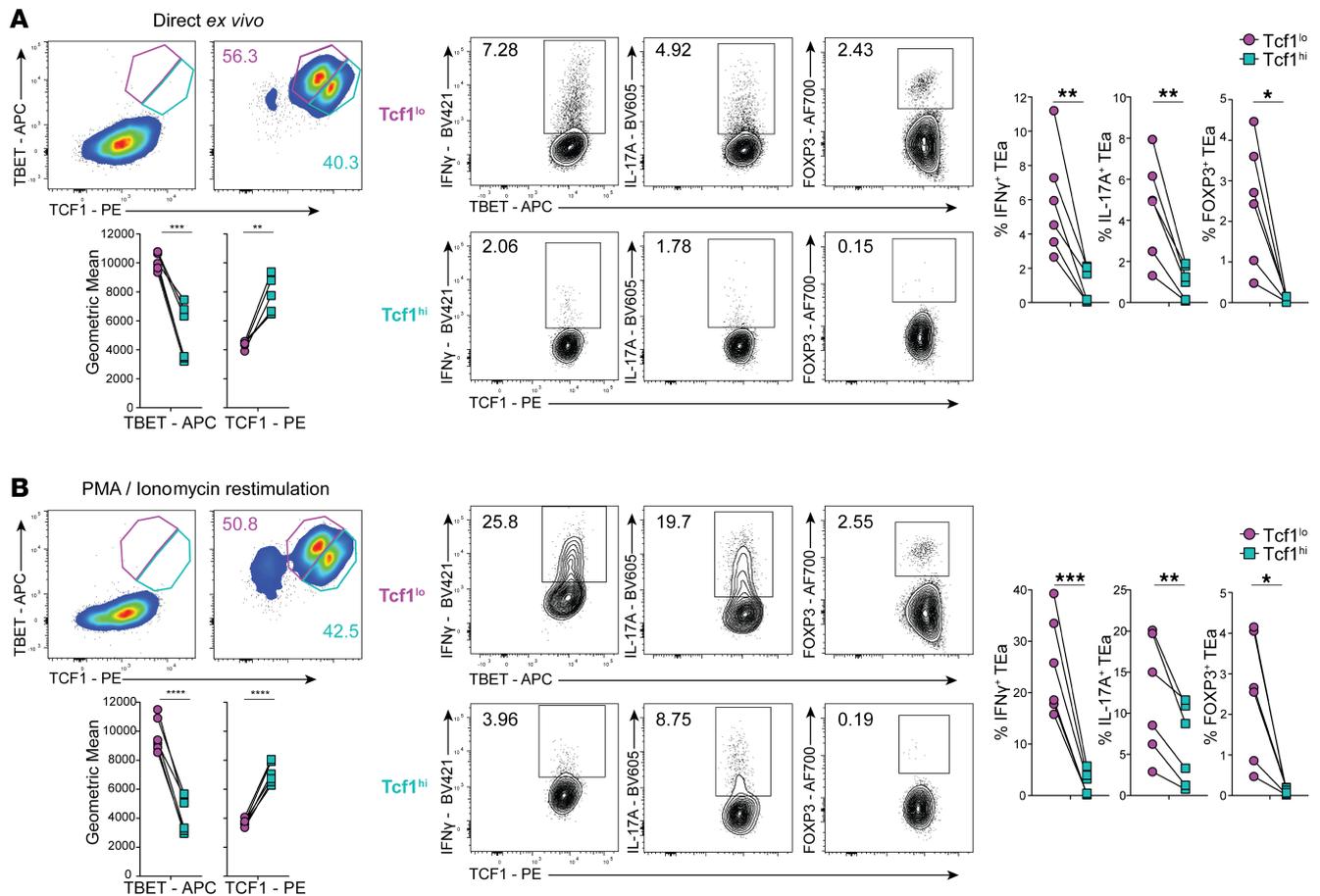


Figure 4. Clonally expanded TEa T cells acquire a TCF1^{hi} noneffector state in mLN. Representative FACS plots showing Tbet and Tcf1 expression in TEa cells at day 4 after transfer in the mLN directly ex vivo (A) or after restimulation with PMA/ionomycin (B). Graphs show the geometric mean fluorescence intensity of the Tcf1^{lo} (purple) or Tcf1^{hi} (turquoise) population. IFN- γ , IL-17A, and Foxp3 expression for the Tcf1^{lo} (purple) or Tcf1^{hi} (turquoise) population is also shown. Graphs show the percentage of IFN- γ ⁺, IL-17A⁺, or Foxp3⁺ TEa cells for the Tcf1^{lo} or Tcf1^{hi} population. Data shown are combined from 2 independent experiments ($n = 3$ mice/ experiment). Statistical analysis was performed using a paired t test. * $P < 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. mLN, mesenteric lymph node.

To test predictions from our transcriptomic models, we assessed mLN TEa cells at day 4 after transfer by flow cytometry. As expected (31), all TEa cells upregulated the Th1-associated lineage transcription factor, Tbet (Figure 4A). We also observed a clear bifurcation in the *Tcf7*-encoded protein expression, TCF1 (Figure 4A), with 1 population expressing higher levels of TCF1 and lower levels of Tbet than its counterpart (Figure 4A). We also noted that direct ex vivo expression of IFN- γ or IL-17A was substantially reduced in TCF1^{hi} cells relative to TCF1^{lo} counterparts. Notably, Foxp3 expression was absent in TCF1^{hi} cells relative to TCF1^{lo} cells (Figure 4A), and in vitro restimulation did not recover expression of these molecules (Figure 4B). Together, these data were consistent with scRNA-Seq prediction of the emergence of alloreactive CD4⁺ T cells that were quiescent and marked by high expression of TCF1.

Tcf7^{hi} alloreactive CD4⁺ T cells change minimally during migration from mLN to the gut. Although TCF1^{hi} TEa cells within the mLN failed to express Foxp3 or proinflammatory cytokines, IFN- γ and IL-17A, scRNA-Seq predicted their capacity to migrate to the gut owing to expression of the integrin gene *Itgb7* (Figure 2C). To test this finding and to explore the developmental relationships between mLN and gut-migrating TEa cells, we conducted a third scRNA-Seq experiment, examining TEa cells at day 5, both in mLN and in the gut intraepithelial lymphocyte (IEL) fraction (Figure 5A).

TEa cells were readily recovered from lamina propria (LP) and gut IEL fractions. However, given the longer protocol required to isolate cells from LP versus IEL, and the potential to interfere with transcriptome fidelity, we confined scRNA-Seq analysis to IEL TEa cells. To integrate day-5 data with our previous mLN data set, we repeated assessments of days 0 and 4. Finally, to control for possible technical variation in IEL induced by the isolation protocol, we treated day-5 mLN cells with and without the IEL isolation

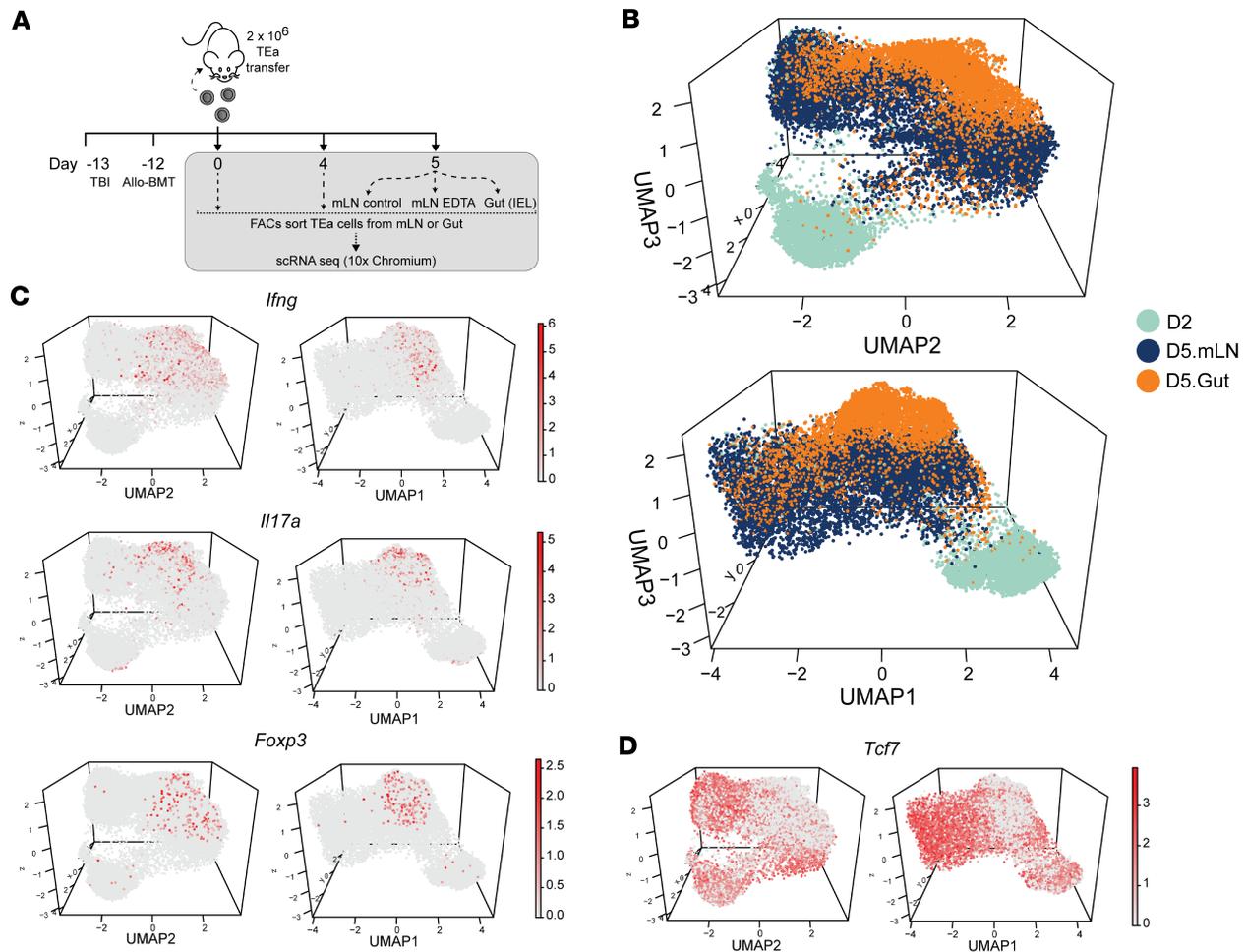


Figure 5. An integrated transcriptomic atlas of TEa T cells over time and across organs. (A) Schematic of scRNA-Seq experiment used to study the developmental relationships between TEa cells from the mLN and the gut (IEL). (B) Two perspectives of 3D UMAP visualization produced from 30 scVI latent variables representing all samples from days 2–5. Only day-2, day-5 gut, and day-5 mLN samples are shown for clarity. (C) *Ifng*, *Il17a*, and *Foxp3* expression in the same representation as (B). (D) *Tcf7* expression in the same representation as (B and C).

predigestion buffer. After scRNA-Seq, and quality control as before (Supplemental Figure 9A), we assessed 21,632 cells across all samples.

First, we noted no effect of the IEL isolation buffer on transcriptomes (Supplemental Figure 9B), indicating that direct comparison of cells from mLN and IEL was possible. Second, unsupervised clustering of day-5 IEL cells revealed 4 main clusters (0, 1, 2, and 3) and minor cluster 4, with *Ifng*, *Il17a*, and *Foxp3* upregulated across a broad area in clusters 0, 1, and 3 but not in cluster 2 (Supplemental Figure 9C). The frequency of TEa cells expressing proinflammatory cytokine genes was substantially elevated compared with mLN, with approximately 15% and approximately 40% of TEa cells in clusters 0, 1, and 3 expressing *Il17a* or *Ifng*, respectively (Supplemental Figure 9C). Pronounced expression also revealed that patterns for *Ifng* and *Il17a* expression were not identical in IEL TEa cells. Although *Ifng* was relatively uniform in expression across clusters 0, 1, and 3, *Il17a* was focused in specific areas. However, a clear distinction between Th1 and Th17 could not be drawn transcriptomically, consistent with previous data that alloreactive CD4⁺ T cells can coexpress IFN- γ and IL-17A at protein level. Interestingly, *Foxp3* was not tightly confined to a particular area in IEL cells, suggesting that iTregs were transcriptomically varied and could not be partitioned into a specific lineage separate from those expressing proinflammatory cytokines. Together, these data reveal that even after priming in mLN and migration to the gut, alloreactive effector CD4⁺ T cells remain transcriptomically similar to each other, regardless of their proinflammatory or regulatory phenotype.

Unsupervised clustering of IEL TEa cells also revealed cluster 2, which completely lacked *Ifng*, *Il17a*, and *Foxp3* expression, but was elevated for *Tcf7*, *Ccr7*, and *Cd27* expression relative to other clusters

(Supplemental Figure 9C). Cluster 4 contained a small population of cells that could not be reliably analyzed (Supplemental Figure 9D). Thus, our assessment of TEa cells in the gut revealed the presence of both effector and quiescent cell states, mirroring those seen in mLN.

To determine molecular relationships between TEa cells in mLN and the gut, we integrated our data sets across all time points and organs using single-cell variational inference (scVI) (32), which accounted for possible batch effects across independent experiments, and provided a temporal atlas of differentiation for alloreactive CD4⁺ T cells (Supplemental Figure 10, A and B). Output from scVI was first visualized via 3D UMAP. mLN TEa cells (days 0 and 4), regardless of experiment, occupied the same space as their time point counterparts, suggesting that any technical variation from different experiments, sequencing platforms, and protocols had been removed. Next, we noted that cells (days 0 and 1) existed in discrete transcriptional states, separate from each other and from day 2–5 cells (Supplemental Figure 10, A and B). This finding suggested TEa cells had undergone rapid and uniform change on initial exposure to alloantigen in mLN, with potential intermediate cellular states not captured by scRNA-Seq assessment at a single time point 24 hours after transfer. Similarly, we noted few transcriptomic intermediates between days 1 and 2, suggesting further uniform, rapid change during the second 24-hour period. Differential gene expression analysis between consecutive days revealed gene families associated first with ribosomal processes, then cellular division upregulated during the first 48-hour period of alloantigen exposure (Supplemental Table 2), consistent with initiation of clonal expansion.

Next, we noted from days 2–5 in mLN, a substantial, gradual increase in heterogeneity, with modest effector molecule expression, such as *Ifng*, *Ill17*, and *Foxp3* expression confined to one space, with quiescent *Tcf7^{hi}* cells occupying a separate space (Figure 5, B–D). Importantly, from our integrated model we inferred further transcriptomic change as effector cells migrated from mLN to the gut, including *Csf2* upregulation (encoding GM-CSF), Tr1-associated *Ill10* immune suppression, and increased *Ifng*, *Ill17a*, and *Foxp3* expression (Supplemental Figure 10C). This suggested differentiation into IFN- γ ⁺ or IL-17A⁺ proinflammatory effectors, Foxp3⁺ iTreg cells, or IL-10⁺ Tr1 cells, although initiated in secondary lymphoid tissue appeared to continue during and after migration to the gut. In contrast, we noted substantial transcriptomic overlap between *Tcf7^{hi}* cells in mLN versus IEL. Moreover, once *Tcf7^{hi}* cells had emerged in mLN by day 3, their phenotype altered very little either over the following 2 days, or during migration to the gut (Figure 5D). These data suggest that *Tcf7^{hi}* cells remained transcriptomically stable across different tissues, whereas effector cells, including those expressing *Ifng*, *Ill17a*, or *Foxp3*, underwent progressive transcriptomic change over this period.

Finally, using our scVI model, we investigated how quiescent alloreactive CD4⁺ T cells might develop in mLN during acute gut GVHD. Trajectory inference, for example using Slingshot on bGPLVM or UMAP embedding, assumes developmental changes in cells are gradual enough to capture intermediate states. However, although emergence of effector cells appeared gradual in the mLN and during migration to the gut, this was less clear for emergence of *Tcf7^{hi}* cells. Transcriptomic intermediate states were apparent between *Tcf7^{hi}* and *Tcf7^{lo}* cells by days 4–5 (Figure 5D), consistent with possible linear transitions from effector to memory. However, at day 2, before effector differentiation and clonal expansion, small numbers of cells appeared in the quiescent, low cell-cycling *Tcf7^{hi}* population (Figure 5D), suggesting an alternative mechanism of development unlinked to effector differentiation. Thus, scRNA-Seq analysis suggested that some quiescent *Tcf7^{hi}* alloreactive CD4⁺ T cells could have emerged rapidly within 48 hours of allopresentation in mLN, with developmental intermediates being difficult to capture in our experimental design. In summary, our integrated atlas of alloreactive CD4⁺ T cell differentiation (https://camerongw.github.io/Engel_Lee_Williams_Supplementary_File.html), revealed the emergence of proinflammatory, regulatory, and quiescent cell states within secondary lymphoid tissue, which emerged rapidly and evolved to differing degrees during migration to the gut.

Clonally expanded TCF1^{hi} CD4⁺ T cells can mount secondary effector responses in vivo. Finally, we sought to determine the functional potential of quiescent TCF1^{hi} TEa cells. We tested the hypothesis that TCF1^{hi} TEa cells could give rise to a secondary effector response and regenerate themselves within the gut. First, we noted these cells expressed less of the canonical coinhibitory markers *Pdcd1* (encoding PD-1), *Havr2* (encoding Tim-3), and *Tigit* but not *Lag3*, compared with effector TEa cells, consistent with the idea that TCF1^{hi} TEa cells might be responsive to reactivation in the gut (Supplemental Figure 11A). Second, regarding naive cells, *Tcf7^{hi}* cells expressed higher levels of both *Il6st* and *Il6ra* than effectors, which are required for classical IL-6 signaling that promotes CD4⁺ T cell responses in acute gut GVHD (Supplemental Figure 11B). These data suggested that *Tcf7^{hi}* cells might retain the capacity to respond to alloantigen.

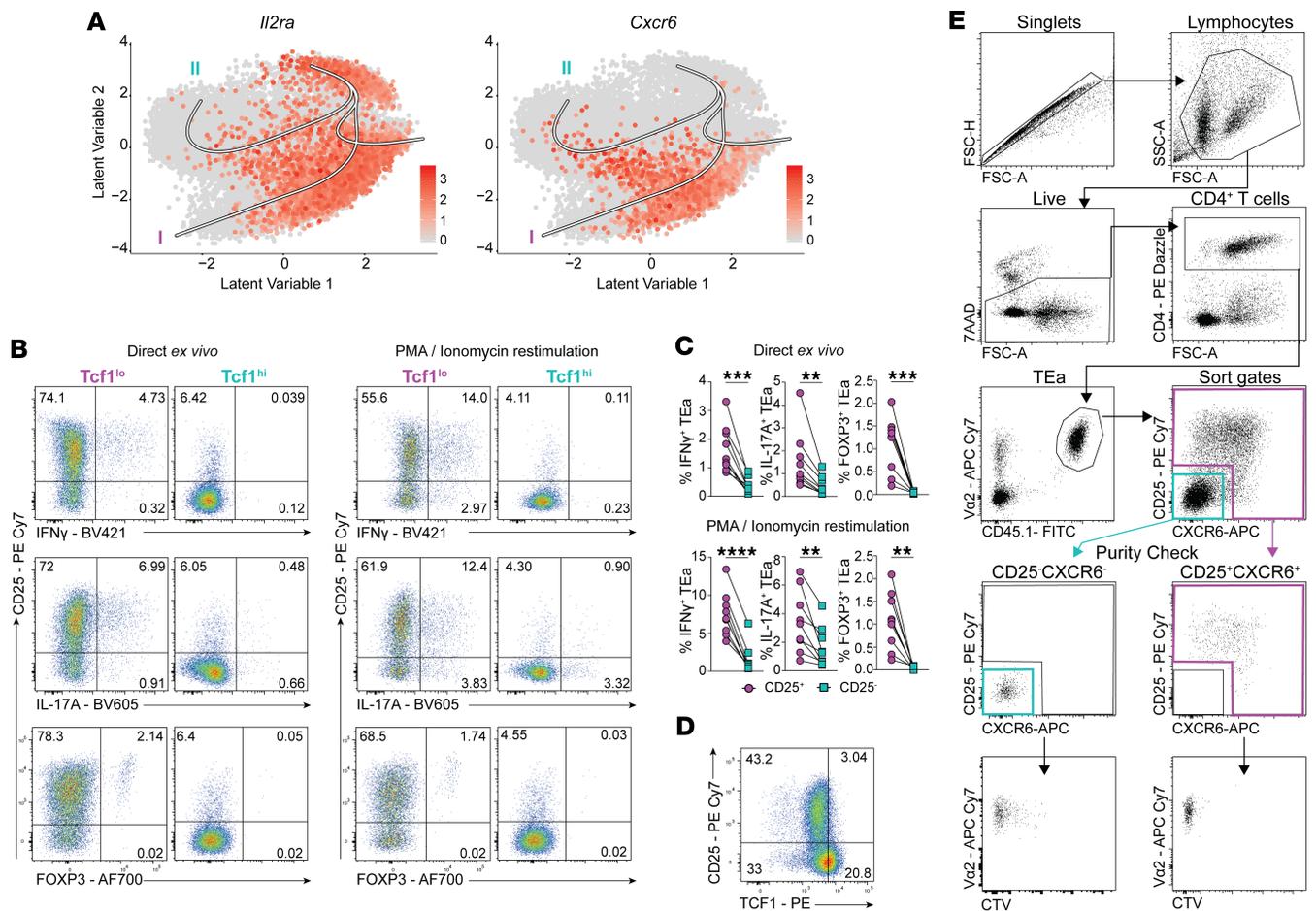


Figure 6. Defining a cell-sorting strategy for TCF1^{hi} TEa T cells. (A) bGPLVM visualization of TEa cells from the mLN from day 1–4 expressing *Il2ra* and *Cxcr6*, with Slingshot trajectories overlaid. (B) Representative flow cytometry plots showing the expression of CD25, IFN- γ , IL-17A, and Foxp3 on Tcf1^{lo} or Tcf1^{hi} TEa cells from the mLN at day 4 after transfer, directly ex vivo, or after restimulation with PMA/ionomycin. (C) Graphs show the percentage of CD25⁺ or CD25⁻ TEa cells that express IFN- γ , IL-17A and Foxp3. Data shown are combined from 2 independent experiments showing similar results ($n = 9$ mice). Statistical analysis was performed using a paired *t* test. * $P < 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. (D) Representative flow cytometry plot showing the expression of CD25 and TCF1 on TEa cells from the mLN at day 4 after transfer. (E) Flow cytometry gating strategy used to isolate CD25⁻CXCR6⁺ and CD25⁺CXCR6⁺ TEa cells from the mLN at day 4 after transfer. The corresponding purity checks and CTV expression for each population are shown. CTV, cell trace violet; mLN, mesenteric lymph node.

In the absence of a *Tcf7* reporter system, we designed a cell-sorting strategy to enrich for *Tcf7*^{hi} TEa cells and remove effector counterparts. We noted in our bGPLVM/Slingshot scRNA-Seq model that trajectory II cells expressed much lower levels of *IL2ra* (encoding CD25) and the Th1-associated chemokine receptor gene, *Cxcr6*, compared with trajectory I effector cells (Figure 6A). Consistent with this result, flow cytometric assessment of day-4 mLN revealed that IFN- γ /IL-17A production and Foxp3 expression were confined to CD25⁺ TCF1^{lo} TEa cells (Figure 6, B and C). Given that CD25⁻ TEa cells uniformly expressed higher levels of intracellular TCF1 compared with CD25⁺ counterparts (Figure 6D), we concluded that cell-sorting clonally expanded cells, based on the lack of CD25/CXCR6 expression and loss of CellTrace dye, provided a feasible alternative to employing a *Tcf7*-reporter (Figure 6E).

We sorted clonally expanded (CTV^{lo}) CD25⁻CXCR6⁺ TEa cells from mLN at day-4 after transfer, and transferred these or naive TEa comparator cells into irradiated BALB/c alloSCT recipients (Figure 7A and Figure 6E). Although our experiment was designed to directly compare the abilities of antigen-experienced TCF1^{hi} TEa and naive TEa cells to mount effector responses and regenerate themselves, we also transferred CD25⁺ and/or CXCR6⁺ counterparts into a third cohort of alloSCT recipients. Because of their lower prevalence, this was performed with 50% fewer cells than the other 2 groups. When assessed 4 days later, recipients of TCF1⁺ TEa cells harbored as many cells as those receiving naive TEa cells (Figure 7B). In contrast, recipients of TCF1⁻ TEa cells, albeit transferred with 50% fewer cells, harbored approximately

10% of the original input, compared with 40%–60% for naive and TCF1^{hi} cells. Whereas naive TEa cells diverged to give rise to both TCF1^{hi} and TCF1^{lo} cells, TCF1^{hi} TEa cells were less able to regenerate the TCF1^{hi} phenotype in a secondary transplant, and TCF1^{lo} TEa cells were almost incapable of doing so (Figure 7C). TCF1^{hi}-derived TEa cells were capable of expressing IFN- γ , IL-17A, or Foxp3 *ex vivo* and after restimulation (Figure 7, D and E), with IL-17A expression increased relative to primary responses by naive cells. Together, these data suggested that antigen-experienced TCF1⁺ CD4⁺ T cells, although quiescent after clonal expansion in a primary response, could mount secondary effector responses, but regenerated the TCF1^{hi} phenotype poorly. Thus, our data are consistent with the priming of quiescent, gut migratory, alloreactive CD4⁺ T cells, which retain effector potential.

Discussion

Although alloSCT is an established curative therapy for a range of hematological malignancies, a major limitation is acute GVHD, in which alloreactive naive donor T cells differentiate into proinflammatory effectors that damage the GI tract, liver, and skin (1). Cytokines such as IL-17A, IFN- γ , and GM-CSF produced by alloreactive Th1 and Th17 cells in the GI tract promote disease (33), whereas IL-10 produced by Tr1 and iTreg cells provide protection (31). An important goal in alloSCT is to preserve GVL effects while reducing GVHD. Key to this endeavor is consideration of the spatial and temporal differences between GVL and GVHD. Although GVL exerts beneficial effects in primary and secondary lymphoid organs, acute and lethal GVHD often occurs in the GI tract. By understanding CD4⁺ T cell differentiation in the gut after alloSCT, we may define new strategies to block pathogenic cellular states and encourage protective ones. Although CD4⁺ T cell differentiation has been explored at genome scale in infection, autoimmune, and allergy models (7–9), extrapolating to alloSCT remains challenging. For example, in the alloSCT setting, alloantigen is ubiquitous and constant, whereas pathogen-derived antigen may be more dynamic or transient. Second, alloSCT often features profound lymphopenia unlike other models. Hence, we specifically examined transcriptome dynamics of Th cell differentiation in mLN and the GI tract of mice after alloSCT.

By sampling transcriptomes from thousands of alloreactive CD4⁺ T cells of a single specificity across lymphoid and nonlymphoid gut-associated tissue, we detected cellular states expressing canonical Th1/Th17 cytokine genes, *Ifng* and *Il17a*, and the regulatory genes, *Foxp3* and *Il10*, at frequencies similar to that observed by flow cytometry. Notably, unbiased clustering and trajectory inference tools suggested substantial similarity between the transcriptomes of these effector subsets, particularly in lymphoid tissue but also in the gut. Subtle differences became more evident among gut-trafficked TEa cells that had stopped proliferating, with *Ifng* expressed more uniformly than either *Il17a* or *Foxp3*. Given that TEa cells can upregulate T-bet and IFN- γ but not IL-17A or Foxp3 in the complete absence of the class II MHC presentation (2, 4), our data are broadly consistent with a Th1-like state being a default program in the gut, which may be countered by alloantigen presentation via class II MHC molecule toward iTreg or Th17-like states. Nevertheless, our main inference from transcriptome dynamics was that proinflammatory states were not readily distinguished from immune-suppressive iTreg states. One question, however, is whether our inability to separate proinflammatory and regulatory states was attributed to reliance on droplet-based scRNA-Seq, without supplementing this with high-dimensional protein assessment via flow or mass cytometry, single-cell epigenomic assessment via scATAC-Seq, or scRNA-Seq at higher sequencing depths. Future experiments using high-dimensional flow cytometry will determine whether posttranscriptional regulation plays any role in distinguishing emerging effector states within secondary lymphoid tissue.

Our unbiased, single-cell genomic approach suggested an unexpected, apparent trajectory characterized by TCF1^{hi} expression, rapid shutdown of cellular proliferation, a lack of proinflammatory or immune-regulatory gene expression, an ability to migrate to the gut, and a capacity to mount a secondary recall response (Figure 8). In addition, many genes upregulated in these cells have been associated with T cell quiescence or longevity, including *Btg1*, *Samhd1*, *Mxd4*, *Laptm5*, *Gimap3*, and *Izumo1r* (21–23, 25, 26). Based on these observations, we infer TCF1^{hi} TEa cells to be generally quiescent memory or stem-like cells that emerged rapidly during alloSCT. Transcriptomic modeling suggested that *Tcf7*^{hi} cells could under certain circumstances arise from the cytokine-expressing effector lineage at day 3 or 4, which would be consistent with a linear model in which effector cells give rise directly to memory-like cells (34, 35). However, we also noted rare instances of *Tcf7*^{hi} cells emerging at day 2 after transfer, as clonal expansion was beginning and effector differentiation had yet to occur. We did not detect transcriptomic intermediates between this distal state and more naive cells, either because such states do not

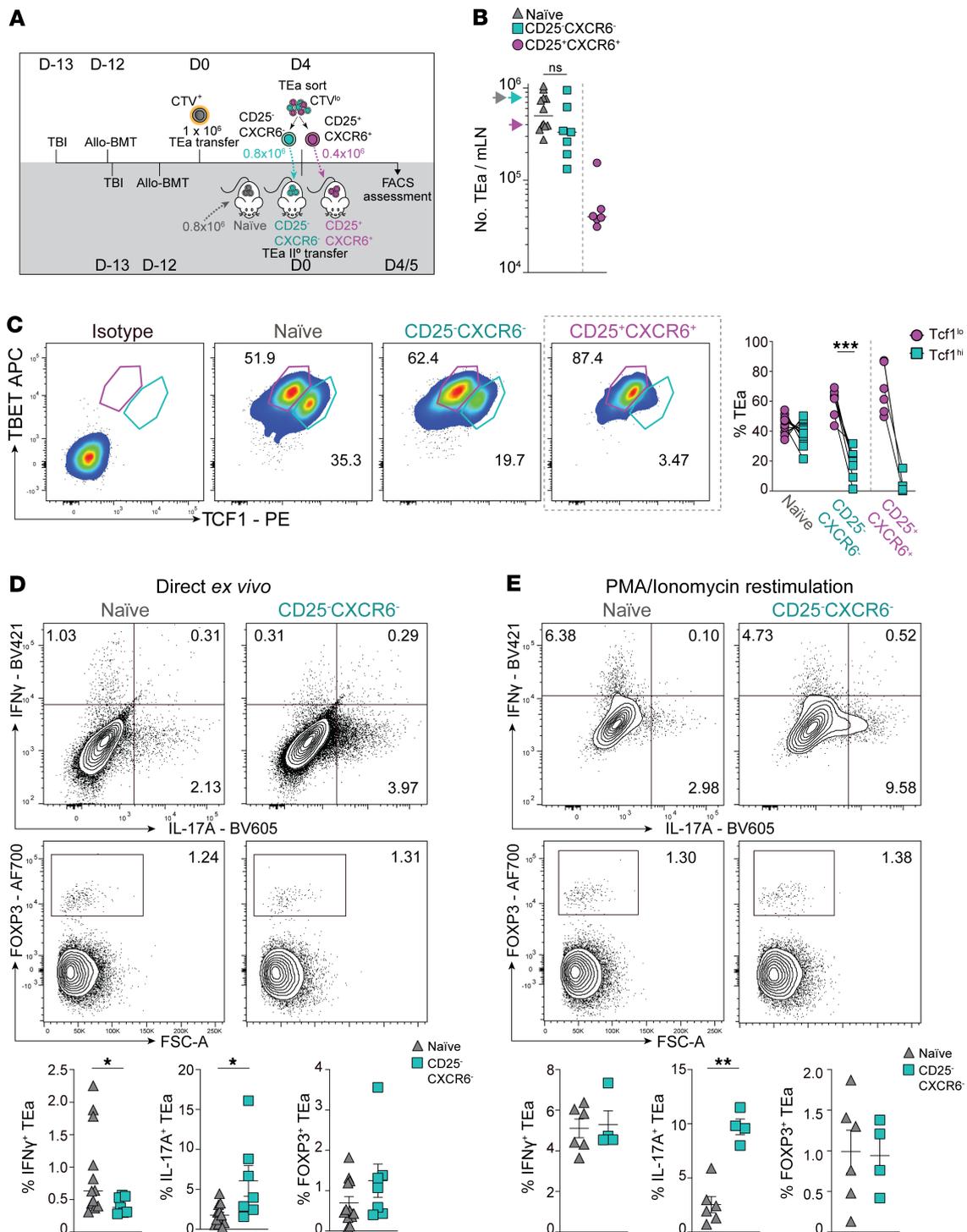


Figure 7. TCF1^{hi} TEa T cells can mount secondary effector responses in vivo. (A) Schematic of secondary TEa transfer experiment. On day 4 of the primary transplant, CD25⁺CXCR6⁻ or CD25⁺ and/or CXCR6⁺ (referred to as CD25⁺CXCR6⁺) TEa were FACS isolated from mLN. A total of 0.8 × 10⁶ CD25⁺CXCR6⁻ and 0.4 × 10⁶ CD25⁺CXCR6⁺ TEa (or control 0.8 × 10⁶ naive TEa) were transferred into recipient BALB/c mice that had received total body irradiation and a bone marrow transplant 13 and 12 days prior, respectively. FACS assessment of mLN TEa cells was performed on day 4 or 5 after secondary transfer. (B) Absolute numbers of TEa cells in the mLN at day 4 or 5 after secondary transfer. Arrows along the y-axis denote the number of TEa cells that were transferred per mouse for each group on day 0 (0.8 × 10⁶ naive TEa (gray) and CD25⁺CXCR6⁻ TEa (turquoise), 0.4 × 10⁶ CD25⁺CXCR6⁺ TEa (purple). (C) Representative flow cytometry plots showing the expression of Tbet and Tcf1 on TEa cells from the mLN at day 4 or 5 after secondary transfer. The graph shows the percentage of TEa cells that are Tcf1^{lo} or Tcf1^{hi} in each group. (D and E) Representative FACS plots showing IFN- γ , IL-17A, and Foxp3 expression on TEa cells after secondary transfer on day 4 or 5 from the mLN directly ex vivo (D) or after restimulation with PMA and ionomycin (E). Graphs show the percentage of IFN- γ ⁺, IL-17A⁺, and Foxp3⁺ TEa for each group. Data shown are combined from 2 independent experiments showing similar results (B–D: naive, n = 12; CD25⁺CXCR6⁻, n = 7, and CD25⁺CXCR6⁺, n = 6) or from one experiment only (E: naive, n = 6; CD25⁺CXCR6⁻, n = 4). Data are shown as the median (B) or mean \pm SEM (C–E). Statistical analysis was performed using a Mann-Whitney test (B, D, and E) or a paired t test (C). *P < 0.05, **P \leq 0.01, ***P \leq 0.001.

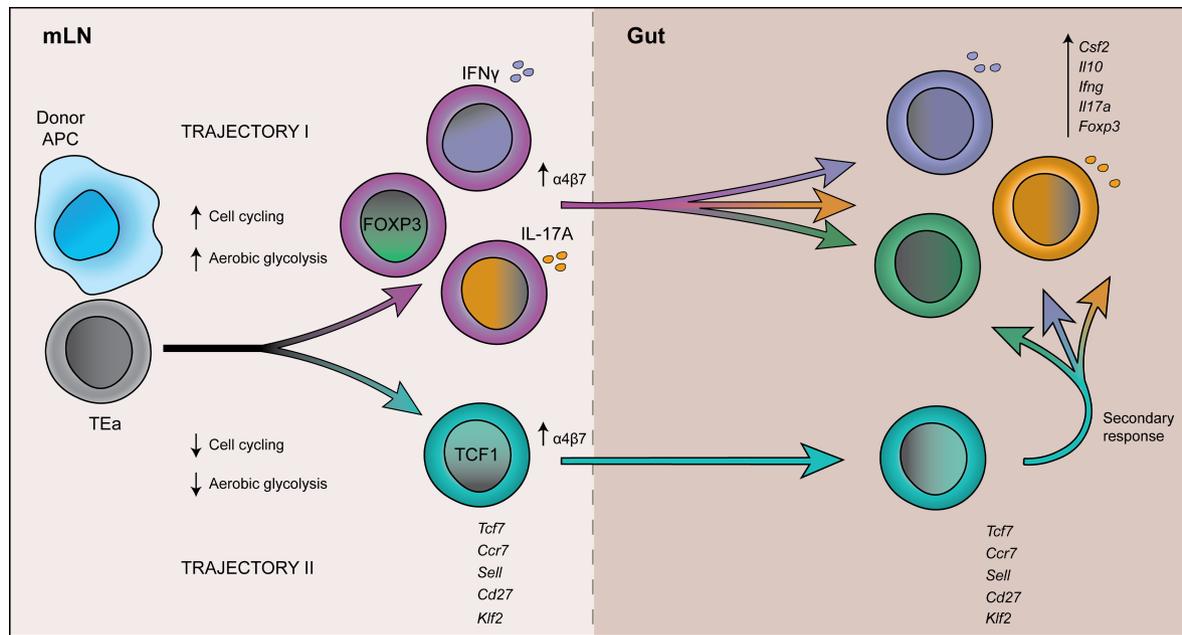


Figure 8. Schematic overview of alloreactive CD4⁺ T cell differentiation in the mLN and gut. Alloreactive CD4⁺ T cells (TEa) activated by donor antigen presenting cells differentiate along 2 developmental trajectories in the mLN. Trajectory I cells express high levels of cell-cycling and aerobic glycolysis genes and encompass pathogenic subsets that produce IFN- γ and IL-17A and regulatory subsets expressing FOXP3. Trajectory II cells are largely quiescent, expressing TCF1 and upregulating stemness-associated genes, including *Ccr7*, *Sell*, *Cd27*, and *Klf2*. Cells following both trajectories upregulate expression of $\alpha 4\beta 7$, allowing trafficking to the gut (IEL). Trajectory I cells continue to evolve transcriptomically during migration to the gut, by further upregulating expression of proinflammatory and regulatory genes. In contrast, trajectory II cells migrate to the gut, yet remain transcriptomically unchanged. Nevertheless, TCF1^{hi} cells are able to mount a secondary effector response.

exist, or because our study was not designed to detect such rare, transient events. Given that trajectory inference from scRNA-Seq data tends to rely on, and indeed assume, gradual transcriptomic change, it is likely that very rapid state changes cannot be mapped using this approach. Our data do not resolve the extent to which TCF1^{hi} CD4⁺ T cells emerge via gradual linear effector-memory transition versus a more rapid, branching process via asymmetric cell division, which was recently reported for similar cells during a respiratory virus infection model (36), and for CD8⁺ T cells during lymphocytic choriomeningitis virus infection (11). Given the presence of rare “pioneer”-like *Tcf7*^{hi} cells and apparent transcriptomic intermediates, we speculate that both of these developmental pathways may operate during alloSCT. New research tools are required to quantify the relative use of these mechanisms in vivo.

Antigen-experienced TCF1^{hi} cells have been reported frequently in recent CD8⁺ T cell studies (37–39), where the expression of TCF1 has been associated with long-term persistence of CD8⁺ T cells, either as memory or stem-like cells. Moreover, TCF1^{hi} CD8⁺ T cells are thought to represent the pool of cells that responds well to immune checkpoint blockade during cancer treatment (40–42). Thus, emerging dogma from CD8⁺ T cell studies is that antigen-experienced TCF1^{hi} T cells are functionally relevant. In addition, earlier reports suggested that some Th17 cells, defined by IL17A expression, might also express high levels of *Tcf7*, which was associated with their persistence (43). Our data reveal that alloreactive CD4⁺ T cells can also adopt a TCF1^{hi} state during GVHD exacerbation. Moreover, these cells may be functionally relevant at later times during GVHD. Thus, our data are broadly consistent with CD8⁺ T cell studies in suggesting that clonally expanded, antigen-experienced T cells can adopt a TCF1^{hi} state that may persist and function at later time points. Therefore, under certain circumstances, primary activation of small numbers of antigen-specific CD4⁺ T cells could give rise to a larger pool of highly plastic counterparts, with obvious implications for the magnitude and quality of secondary immune responses to the same antigen.

One question to consider from our transcriptomic modeling is how separate states, loosely referred to as “effector” or “quiescent”, emerge among clonal T cells during alloSCT. Heterogeneity among clonal TEa cells could have been induced via various nonmutually exclusive mechanisms, including asymmetric cell division (36), differential APC engagement, and differential access to early local cytokine signals. Given the stark difference in CD25 expression between trajectories, it appears feasible that

IL-2-signaling promotes effector function at the expense of quiescence. A recent study revealed that a reciprocal relationship in production and receipt of IL-2 controlled fate bifurcation in CD4⁺ T cells (44). Hence, we speculate that a similar mechanism might be acting during alloSCT. In our model, we previously reported that colonic CD103⁺ DCs played a crucial role in amplifying acute GVHD (4). It is possible that naive CD4⁺ T cells that failed to access these APC may have been programmed toward the quiescent TCF1^{hi} state. As part of this scenario, given that our model is characterized by profound lymphopenia, it is possible that homeostatic proliferation, perhaps via IL-7 signaling, might have partly contributed to the proliferation and stabilization of a quiescent cellular state. However, given that donor alloantigen presentation via class II MHC is important for supporting clonal expansion in this model, exposure to diverse donor APC may contribute to heterogeneity in CD4⁺ T cell differentiation.

A crucial element of our study design was its focus on CD4⁺ T cells of a single-specificity. Given that TCR sequence can influence effector fate (45), we intentionally reduced the complexity of our system to increase the likelihood of computationally modeling Th1/Th17/iTreg effector differentiation. Although numerous scRNA-Seq-based studies, including our own, have mapped cellular change in TCR transgenic T cells over time (7–12), an ongoing challenge has been to replicate this approach with highly diverse polyclonal T cell populations. We expect that increases in phenotypic diversity and heterogeneity in response kinetics will present additional computational challenges. Nevertheless, efforts to progress longitudinal genomic studies of T cell differentiation from TCR transgenic systems to polyclonal T cells are warranted, in particular to facilitate analogous studies in humans.

We envisage a model for development of Th1, Th17, Tr1, or iTreg-like states during alloSCT controlled by specific microanatomical T cell extrinsic factors, such as access to class II MHC presentation on different types of APC, or exposure to cytokines, including IL-6 and IL-12. Given the apparent continued maturation of proinflammatory TEa cells as they migrated from mLN to the gut, inferred from our scVI-based atlas, this raises the question of whether peripheral tissue signals in the gut, such as local cytokine-signaling, or unique cell-cell interactions contribute to this process. We speculate that spatial transcriptomic assessment of gut-located TEa cells will shed light on this matter. A further implication of our data is that conversion of emerging proinflammatory CD4⁺ T cells into protective nonpathogenic iTreg cells in the gut may be feasible, considering that developmental pathways are similar between the two. Conversely, our data also support those iTreg-based therapies with appropriate mitigation for the effects of reversion to proinflammatory states in vivo (46).

We and others have previously studied Th17 biology via fate-mapping using *IL-17a* promoter-driven Cre-mediated fluorescent tagging of cells (6, 47, 48). This binary approach is powerful, but does not differentiate between cells that might transiently express *Il17a* compared with those exhibiting prolonged expression. In our scVI transcriptomic model, we noted early transient expression of *Il17a* and *Il17f* at day 2, which disappeared only to reappear in some cells at day 5 in the gut. This suggests that CD4⁺ T cells can transiently express *IL17a* in mLN without becoming bona fide Th17 cells. Interestingly, a recent report employed an *Il17a*-Cre fate-mapping approach in the murine EAE model, and suggested 2 states existed among cells that had expressed *Il17a* at some point. One state was a quiescent, stem-like population that expressed TCF1 and CD27 (47). Hence, we propose that studies employing fate-mapping approaches should be interpreted with possible transient expression taken into account.

Our experiments revealed that clonally expanded TCF1^{hi} CD4⁺ T cells, which exhibited no effector function and shut down cellular proliferation, were capable of expressing proinflammatory cytokines or Foxp3 at a later date, as shown during secondary transplantation. This raises the hypothesis that after alloSCT, the gut is populated with quiescent alloreactive T cells that could influence disease outcome. Further experiments are required to examine the longer term persistence and functional relevance of TCF1^{hi} CD4⁺ T cells during GVHD. As a note of caution, however, the earlier hypothesis is founded on an exacerbation model, in which donor DC emerge after the initiation phase of GVHD and amplify symptoms. Therefore, whether the differentiation trajectories inferred in this study occur in alloreactive CD4⁺ T cells during GVHD initiation remains to be tested. Moreover, the possible influence of TCF1^{hi} CD4⁺ T cells on GVL warrants exploration, given the presence of these T cells in secondary lymphoid tissue and the pivotal role of alloSCT in mediating GVL. Future experiments should employ preclinical models and clinical samples to examine the possible relevance of our findings to the initiation phase of acute gut GVHD. Nevertheless, we view clonally expanded TCF1^{hi} CD4⁺ T cells as an opportunity to lodge immune-suppressive, possibly tissue-resident CD4⁺ T cells within the gut that could contribute to disease prevention after alloSCT. However, the potential for these

to produce pathogenic cytokines would clearly need to be addressed. In summary, our examination of transcriptome dynamics during GVHD exacerbation not only highlights possible developmental relationships between effector CD4⁺ T cells, but also suggests the existence of a quiescent, memory-like state that exhibits functional potential *in vivo*. Next, it will be important to interrogate our experimentally derived observations in preclinical models of acute GVHD initiation and in patient samples after alloSCT.

Methods

Mice

C57BL/6J and BALB/c mice were purchased from the Animal Resources Centre (Canning Vale, Australia), and transgenic TEa ($V\alpha 2^+$, $V\beta 6^+$, CD45.1⁺, CD90.1⁺) mice were bred in-house (4, 5). All mice were female between 6–12 weeks of age and were maintained under specific pathogen-free conditions within the animal facility at QIMR Berghofer Medical Research Institute.

Bone marrow transplantation

BALB/c mice were transplanted as previously described (4). Briefly, BALB/c mice received 900 cGy total body irradiation (TBI; ¹³⁷Cs source at 84 cGy/min) on day 13. On day 12, BALB/c mice were transplanted with 10×10^6 bone marrow cells and 2×10^5 FACS-purified T cells from C57BL/6J donor mice. On day 0, recipient BALB/c mice were injected intravenously with 1×10^6 to 2×10^6 FACS sorted TEa T cells ($V\beta 6^+V\alpha 2^+CD45.1^+$). Cell Trace CFSE Cell Proliferation (Life Technologies) and Violet Proliferation Dye 450 (VPD450; BD Biosciences) staining were performed, according to the manufacturer's protocol. For secondary transfer of TEa cells, CD25⁻CXCR6⁻ TEa cells or CD25⁺CXCR6⁺ TEa cells were FACS isolated on D4 from the mLN and then injected intravenously into BALB/c mice that had received total body irradiation 13 days prior and transplanted with bone marrow cells and purified T cells from C57BL/6J mice 12 days prior. For reasons of cell availability after sort, 0.8×10^6 CD25⁻CXCR6⁻ or 0.4×10^6 CD25⁺CXCR6⁺ TEa cells were transferred into respective groups of mice, and 0.8×10^6 naive TEa cells transferred as a reference control.

Cell isolation from small intestine and colon

Intraepithelial lymphocytes were isolated from the small intestine and colon of mice using the Lamina Propria Dissociation Kit, Mouse (Miltenyi Biotec), according to the manufacturer's protocol.

Flow cytometry

Cells were assessed for viability by staining with 7AAD (MilliporeSigma) or using a LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Life Technologies), according to the manufacturer's protocol. Before staining, cells were incubated with antibodies against CD16 and CD32 (2.4G2) to block Fc receptors. For surface staining, cells were incubated with various combinations of the following antibodies for 20 minutes at 4°C: CD4 (RM4-5); PerCPCy5.5, CD4 (RM4-5); PE Dazzle 594, $V\alpha 2$ (B20.1); APC Cy7, CD45.1 (A20); PE Cy7, CD69 (H1.2F3); PB, $\alpha 4\beta 7$, (KATK32); PE, CD25 (PC61); PECy7, CXCR6 (SA051D1); APC, Armenian hamster IgG isotype control; PB, rat IgG2a isotype control; PE, rat IgG1k isotype control; PECy7, rat IgG2bk isotype control; and APC (all BioLegend). For assessment of intracellular cytokine production and transcription factor expression, cells were incubated with or without ionomycin (500 ng/mL) and PMA (50 ng/mL) for 4 hours at 37°C. Brefeldin-A (5 μ g/mL) was added to cells after 1 hour of incubation. Intracellular staining was then performed using the eBioscience Foxp3/Transcription Factor Staining Buffer Set with the following antibodies: IFN- γ (XMG1.2); BV421, IL-17A (TC11-18H10.1); BV605, rat IgG1k isotype control; BV421, rat IgG1k isotype control; BV605, rat IgG2a isotype; AF700, mouse IgG1k isotype control; APC (all from BioLegend) Tcf1 (C63D9); PE, rabbit IgG isotype control, and PE (all from Cell Signaling Technology); T-bet (4B10); APC, Foxp3 (FJK-16S); and AF700 (all from eBioscience). Samples were acquired on a LSRII Fortessa Analyzer (BD Biosciences), and data analyzed using FlowJo software (Tree Star).

Single-cell RNA capture and sequencing

Three independent experiments were performed for sequencing: mGVHD1, mGVHD2, and mGVHD3. TEa cells were isolated by flow cytometry into a 1% BSA/PBS buffer. Approximately 8000 cells were loaded per channel onto a Chromium controller (10 \times Genomics) for generation of gel-bead-in-emulsions. Sequencing libraries were prepared using Single Cell 3' Reagent Kits v2 (mGVHD1 and mGVHD2) or

v3.0 (mGVHD3) (10× Genomics) and either sequenced on an Illumina NextSeq550 (mGVHD1 and mGVHD2) or converted using the MGIEasy Universal Library Conversion Kit (BGI) before sequencing on a MGISEQ-2000 instrument (BGI; mGVHD3).

Data availability

The full raw single-cell RNA sequencing data from this study have been submitted to the ArrayExpress database (<https://www.ebi.ac.uk/arrayexpress/>) (accession E-MTAB-9125).

scRNA-Seq data processing

FASTQ files were processed using “cellranger count” pipeline from Cell Ranger version 2.1.0 and 3.0.2 (10× Genomics) with 10× mouse genome 1.2.0 release as a reference. For BGI FASTQ files (mGVHD3), it was made compatible with Cell Ranger by reformatting file names and FASTQ headers using code from https://github.com/IMB-Computational-Genomics-Lab/BGIVsIllumina_scRNASeq (branch: master; commit ID: 89092f6) (49).

scRNA-Seq data quality control

Cells outside the thresholds of 200–6000 expressed genes and up to 15% mitochondrial content were removed. Further filtering was performed after unsupervised clustering of cells, in which clusters of cells with low Cd3 and Cd4 expression were removed. Cells that expressed Cd8a were globally removed from downstream analysis. Only genes expressed in 3 or more cells were considered.

Data transformation

scRNA-Seq data were normalized using the Seurat v2.3.4 “NormalizeData” function (50), in which UMI counts for each gene from each cell were divided by the total UMI counts from that cell, multiplied by the scale factor of 10,000, and natural log-transformed. Total UMI content and mitochondrial content per cell were considered unwanted sources of variation and removed by individual linear regression. Final residuals were then scaled to have mean feature expression of 0 and variation of 1 across cells.

Feature selection and dimensionality reduction

Highly variable genes (HVGs) were identified using the Seurat “FindVariableGene” function with default parameters. For each data subset used for dimensionality reduction, HVGs were computed individually and used as an input, unless otherwise specified. The number of HVGs used in each analysis is noted in the figure legends.

Principal component analysis (PCA) dimensionality reduction was performed using the Seurat “RunPCA” function, and the computed PCs were used to generate uniform manifold approximation and projection (UMAP) of scRNA-Seq data using the Seurat “RunUMAP” function. The number of PCs used in each analysis is noted in the figure legends.

bGPLVM dimensionality reduction was performed using GPfates v1.0.0 (9), in which data sets containing all genes after initial gene filtering step were used as input. Up to 5 latent variables were considered.

Integrated dimensionality reduction of mGVHD2 and mGVHD3 data sets was performed using scVI v0.4.1 (32). Each experiment was identified as separate batch. All parameters were kept at default except up to 30 latent variables were considered, 2 hidden layers were used for encoder and decoder neural networks, and up to 100 epochs were used to train the model. The computed latent variables were used as an input to generate UMAP using the Seurat “RunUMAP” or the umap-learn v0.3.5 python package (15).

Unsupervised clustering

The Seurat “FindClusters” function was used to perform unsupervised cell clustering. The resolution parameter and the number of PCs or variables used in each analysis are noted in the figure legends.

Trajectory inference

Slingshot. Trajectories were inferred through the mGVHD2 data set using Slingshot v0.99.12 (16). Slingshot requires clusters of cells and embeddings for those cells as input. bGPLVM latent variables 1 and 2 were used, and unbiased clustering based on these variables was also performed. A semisupervised approach was taken, in which a cluster with high proportion of day-1 cells was specified as a starting point (Supplemental Figure 4A).

Monocle. Trajectories were inferred through the mGVHD2 data set using Monocle v2.8.0 (51). PCA dimensionality reduction was performed, and the first 10 PCs were used as an input for unsupervised clustering using “plot_pc_variance_explained” and “clusterCells” functions, in which the number of clusters was specified ($n = 6$). Differential gene expression analysis was performed between clusters using “differentialGeneTest” function, and the list of significant genes ($q < 0.01$) was used as an input to order the cells using “orderCells” function.

PAGA. Trajectories were inferred through the mGVHD2 data set using SCANPY v1.4.4, which includes the PAGA trajectory inference algorithm (19). PCA dimensionality reduction was performed using the “scanpy.tl.pca” function with ARPACK SVD solver to aid computation. A neighborhood graph was computed via “scanpy.pp.neighbors” with the first 10 PCs, and the size of local neighborhood specified as 30. Unsupervised clustering of cells was performed using the “scanpy.tl.louvain” function with resolution 1.0. Finally, coarse-grained connectivity structures connecting the computed clusters of cells was mapped using “scanpy.tl.paga” with default parameters.

scVelo. RNA velocity analysis was performed using scVelo version 0.1.16 (52). Only cells (days 2–4) from the mGVHD2 data set were included in the analysis. All parameters were kept at default, except 3000 genes, 20 PCs, and 30 neighbors considered for RNA velocity estimation. Calculated velocity was projected onto precomputed bGPLVM embeddings.

Gene signature scoring

The Seurat “AddModuleScore” function was used to calculate gene signatures. The cell cycle score was calculated using 226 cell cycle genes derived from Cyclebase (53), the aerobic glycolysis score used 41 genes associated with the Gene Ontology (GO) ID GO:0006096, and the oxidative phosphorylation score used 30 genes associated with ID GO:0006119.

Gene expression imputation

Gene expression inference from missing data, or imputation, was performed using ALRA (initial release) (<https://github.com/KlugerLab/ALRA/>) (20) and MAGIC v1.5.0 (54).

Differential gene expression analysis and gene ontology term enrichment analysis

Differential gene expression analysis (DGEA) was performed using the Seurat “FindMarkers” function at default parameters. Comparisons were done as follows: (1) trajectory I and II cells (2) day 0 and 1 cells, and (3) day 1 and 2 cells. The first comparison was performed using mGVHD2 data set as input, and the second and third comparisons were performed using the integrated mGVHD2 and mGVHD3 data sets, respectively, as input. Genes from the second and third comparisons 2 and 3, with Bonferroni’s adjusted P value below 0.01 and average log fold change greater than 0.5, were considered as input for the gene ontology (GO) term enrichment analysis. GO terms were obtained from “org.Mm.eg.db” Bioconductor annotation package 99 (55). Fisher’s exact test was performed using “goana” function from edgeR (56) to identify overrepresented GO terms. Input genes for DGEA were used as the background gene set.

Other packages

scRNA-Seq data were primarily visualized with the ggplot2 v3.2.1 R package (57). Additional functions were provided by cowplot v0.9.2, in particular via Seurat’s inbuilt plotting features (50). 3D projections were created by the scatter plot3d v0.3-41 R package (58). Seurat v3.0.1 was used for conversion of data to loom format.

Statistics

Statistical analyses were performed using a Mann-Whitney test, Wilcoxon rank sum test, or a paired 2-tailed t test using Prism software (GraphPad Software Inc.). P values of less than 0.05 were considered significant.

Study approval

All animal procedures and protocols were approved (A0412-617M; P832) and monitored by the QIMR Berghofer Medical Research Institute Animal Ethics Committee.

Author contributions

JAE designed research studies, conducted wet-lab experiments, acquired the data, and analyzed the data. HJL and CGW designed research studies and performed bioinformatics analyses of the data. RK, SO, LIML, MSFS, SBA, AV, MK, and AH designed research studies and conducted experiments. VS provided input to the bGPLVM computational data modeling. JEP, SAT, GRH, AV, and MK provided valuable expertise to the project. SAT, MK, and AH conceived of the project and interpreted the results. AH, JAE, HJL, and CGW wrote and edited the manuscript and prepared figures. All authors discussed, commented on, and approved the final version of the manuscript.

Acknowledgments

This work was supported by research grants from the Australian National Health and Medical Research Council: AH, SAT, and MK were supported by Project and Ideas Grants (1126399 and 1180951). JEP is supported by an Investigator Grant (1175781). We would like to thank staff from the QIMR Berghofer flow cytometry facility for their expertise and assistance with experiments, and staff from QIMR Berghofer animal facility for animal care and husbandry. We acknowledge Paul Collins at the QIMR Berghofer sequencing facility and Scott Wood and his team at the QIMR Berghofer High Performance Computing facility for support with scRNA-Seq data management. We acknowledge helpful discussions with Kate Gartlan (QIMR Berghofer) on matters relating to T cell differentiation during alloSCT. We thank Kylie R. James (Wellcome Sanger Institute) for reading the manuscript and providing helpful comments on issues relating to CD4⁺ T cell differentiation. We acknowledge BGI Australia for their help in sequencing of 10[×] libraries on the MGISEQ-2000, especially Lynn Fink, Cheryll Ye, Shiu Wing In, Ivon Harliwong, Tian Wei, and Wei Han Min for their role in assisting with 10[×] library conversion, pooling design, sequencing and demultiplexing of data

Address correspondence to: Ashrafal Haque, Peter Doherty Institute for Infection and Immunity, 792 Elizabeth Street, The University of Melbourne, Victoria 3000, Australia. Phone: 61.3.9035.5206; Email: ashrafal.haque@unimelb.edu.au.

- Koyama M, Hill GR. The primacy of gastrointestinal tract antigen-presenting cells in lethal graft-versus-host disease. *Blood*. 2019;134(24):2139–2148.
- Koyama M, et al. Recipient nonhematopoietic antigen-presenting cells are sufficient to induce lethal acute graft-versus-host disease. *Nat Med*. 2011;18(1):135–142.
- Koyama M, et al. MHC class II antigen presentation by the intestinal epithelium initiates graft-versus-host disease and is influenced by the microbiota. *Immunity*. 2019;51(5):885–898.e7.
- Koyama M, et al. Donor colonic CD103⁺ dendritic cells determine the severity of acute graft-versus-host disease. *J Exp Med*. 2015;212(8):1303–1321.
- Grubin CE, Kovats S, deRoos P, Rudensky AY. Deficient positive selection of CD4 T cells in mice displaying altered repertoires of MHC class II-bound self-peptides. *Immunity*. 1997;7(2):197–208.
- Gartlan KH, et al. Th17 plasticity and transition toward a pathogenic cytokine signature are regulated by cyclosporine after allogeneic SCT. *Blood Adv*. 2017;1(6):341–351.
- Gaublomme JT, et al. Single-cell genomics unveils critical regulators of Th17 cell pathogenicity. *Cell*. 2015;163(6):1400–1412.
- Tibbitt CA, et al. Single-cell RNA sequencing of the T helper cell response to house dust mites defines a distinct gene expression signature in airway Th2 cells. *Immunity*. 2019;51(1):169–184.e5.
- Lonnberg T, et al. Single-cell RNA-seq and computational analysis using temporal mixture modelling resolves Th1/Tfh fate bifurcation in malaria. *Sci Immunol*. 2017;2(9):eaal2192.
- Chen Z, et al. TCF-1-centered transcriptional network drives an effector versus exhausted CD8 T cell-fate decision. *Immunity*. 2019;51(5):840–855.e5.
- Kakaradov B, et al. Early transcriptional and epigenetic regulation of CD8(+) T cell differentiation revealed by single-cell RNA sequencing. *Nat Immunol*. 2017;18(4):422–432.
- Li H, et al. Dysfunctional CD8 T cells form a proliferative, dynamically regulated compartment within human melanoma. *Cell*. 2019;176(4):775–789.e18.
- James KR, et al. Distinct microbial and immune niches of the human colon. *Nat Immunol*. 2020;21(3):343–353.
- Miragaia RJ, et al. Single-cell transcriptomics of regulatory T cells reveals trajectories of tissue adaptation. *Immunity*. 2019;50(2):493–504.e7.
- Becht E, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2019;37:38–44.
- Street K, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*. 2018;19(1):477.
- Saelens W, Cannoodt R, Todorov H, Saeys Y. A comparison of single-cell trajectory inference methods. *Nat Biotechnol*. 2019;37(5):547–554.
- La Manno G, et al. RNA velocity of single cells. *Nature*. 2018;560(7719):494–498.
- Wolf FA, et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol*. 2019;20(1):59.

20. Linderman LC, Zhao J, Kluger Y. Zero-preserving imputation of scRNA-seq data using low-rank approximation. bioRxiv. <https://www.biorxiv.org/content/10.1101/397588v1>. Published August 22, 2018. Accessed June 10, 2020.
21. Hwang SS, et al. mRNA destabilization by BTG1 and BTG2 maintains T cell quiescence. *Science*. 2020;367(6483):1255–1260.
22. Vasilevsky NA, Ruby CE, Hurlin PJ, Weinberg AD. OX40 engagement stabilizes Mxd4 and Mnt protein levels in antigen-stimulated T cells leading to an increase in cell survival. *Eur J Immunol*. 2011;41(4):1024–1034.
23. Kawai Y, Ouchida R, Yamasaki S, Dragone L, Tsubata T, Wang JY. LAPT5 promotes lysosomal degradation of intracellular CD3 ζ but not of cell surface CD3 ζ . *Immunol Cell Biol*. 2014;92(6):527–534.
24. Xu H, et al. MS4a4B, a CD20 homologue in T cells, inhibits T cell propagation by modulation of cell cycle. *PLoS One*. 2010;5(11):e13780.
25. Yano K, et al. Gimap3 and Gimap5 cooperate to maintain T-cell numbers in the mouse. *Eur J Immunol*. 2014;44(2):561–572.
26. Künzli M, et al. Long-lived T follicular helper cells retain plasticity and help sustain humoral immunity. *Sci Immunol*. 2020;5(45):eaay5552.
27. Lee JY, et al. The transcription factor KLF2 restrains CD4⁺T follicular helper cell differentiation. *Immunity*. 2015;42(2):252–264.
28. Weber JP, et al. ICOS maintains the T follicular helper cell phenotype by down-regulating Krüppel-like factor 2. *J Exp Med*. 2015;212(2):217–233.
29. García-Bernal D, Dios-Esponera A, Sotillo-Mallo E, García-Verdugo R, Arellano-Sánchez N, Teixidó J. RGS10 restricts upregulation by chemokines of T cell adhesion mediated by α 4 β 1 and α L β 2 integrins. *J Immunol*. 2011;187(3):1264–1272.
30. Lee JK, Kannarkat GT, Chung J, Joon Lee H, Graham KL, Tansey MG. RGS10 deficiency ameliorates the severity of disease in experimental autoimmune encephalomyelitis. *J Neuroinflammation*. 2016;13:24.
31. Zhang P, et al. Eomesodermin promotes the development of type 1 regulatory T (T_R1) cells. *Sci Immunol*. 2017;2(10):eaah7152.
32. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. *Nat Methods*. 2018;15(12):1053–1058.
33. Gartlan KH, et al. Donor T-cell-derived GM-CSF drives alloantigen presentation by dendritic cells in the gastrointestinal tract. *Blood Adv*. 2019;3(19):2859–2865.
34. Pepper M, Pagán AJ, Igyártó BZ, Taylor JJ, Jenkins MK. Opposing signals from the Bcl6 transcription factor and the interleukin-2 receptor generate T helper 1 central and effector memory cells. *Immunity*. 2011;35(4):583–595.
35. Tubo NJ, Fife BT, Pagan AJ, Kotov DI, Goldberg MF, Jenkins MK. Most microbe-specific naïve CD4⁺ T cells produce memory cells during infection. *Science*. 2016;351(6272):511–514.
36. Nish SA, et al. CD4⁺ T cell effector commitment coupled to self-renewal by asymmetric cell divisions. *J Exp Med*. 2017;214(1):39–47.
37. Utzschneider DT, et al. T cell factor 1-expressing memory-like CD8(+) T cells sustain the immune response to chronic viral infections. *Immunity*. 2016;45(2):415–427.
38. Wu T, et al. The TCF1-Bcl6 axis counteracts type I interferon to repress exhaustion and maintain T cell stemness. *Sci Immunol*. 2016;1(6):eaai8593.
39. Yao C, et al. Single-cell RNA-seq reveals TOX as a key regulator of CD8⁺ T cell persistence in chronic infection. *Nat Immunol*. 2019;20(7):890–901.
40. Im SJ, et al. Defining CD8⁺ T cells that provide the proliferative burst after PD-1 therapy. *Nature*. 2016;537(7620):417–421.
41. Jansen CS, et al. An intra-tumoral niche maintains and differentiates stem-like CD8 T cells. *Nature*. 2019;576(7787):465–470.
42. Siddiqui I, et al. Intratumoral Tcf1⁺PD-1⁺CD8⁺ T cells with stem-like properties promote tumor control in response to vaccination and checkpoint blockade immunotherapy. *Immunity*. 2019;50(1):195–211.e10.
43. Muranski P, et al. Th17 cells are long lived and retain a stem cell-like molecular signature. *Immunity*. 2011;35(6):972–985.
44. DiToro D, et al. Differential IL-2 expression defines developmental fates of follicular versus nonfollicular helper T cells. *Science*. 2018;361(6407):eaao2933.
45. Tubo NJ, et al. Single naïve CD4⁺ T cells from a diverse repertoire produce different effector cell types during infection. *Cell*. 2013;153(4):785–796.
46. Di Stasi A, et al. Inducible apoptosis as a safety switch for adoptive cell therapy. *N Engl J Med*. 2011;365(18):1673–1683.
47. Karmaus PWF, et al. Metabolic heterogeneity underlies reciprocal fates of T_H17 cell stemness and plasticity. *Nature*. 2019;565(7737):101–105.
48. Varelias A, et al. Lung parenchyma-derived IL-6 promotes IL-17A-dependent acute lung injury after allogeneic stem cell transplantation. *Blood*. 2015;125(15):2435–2444.
49. Senabouth A, et al. Comparative performance of the BGI Illumina sequencing technology for single-cell RNA-sequencing. *NAR Genom Bioinform*. 2020;2(2):lqaa034.
50. 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(5):411–420.
51. Trapnell C, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014;32(4):381–386.
52. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. BioRxiv. <https://www.biorxiv.org/content/10.1101/820936v1>. Published October 29, 2019. Accessed June 10, 2020.
53. Santos A, Wernersson R, Jensen LJ. Cyclebase 3.0: a multi-organism database on cell-cycle regulation and phenotypes. *Nucleic Acids Res*. 2015;43(Database issue):D1140–D1144.
54. van Dijk D, et al. Recovering gene interactions from single-cell data using data diffusion. *Cell*. 2018;174(3):716–729.e27.
55. Carlson M. org.Mm.eg.db: Genome wide annotation for Mouse. Bioconductor. <http://bioconductor.org/packages/release/data/annotation/html/org.Mm.eg.db.html>. Accessed June 10, 2020.
56. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–140.
57. Wickham H. *gg2plot: Elegant Graphics for Data Analysis*. New York, New York, USA: Springer-Verlag New York(4); 2009.
58. Ligges U, Mächler M. scatter plot3d- AnRPackage for visualizing multivariate data. *J Stat Soft*. 2003;8(11):1–20.