







ARTICLE

Single-cell transcriptomic analysis of tissue-resident memory T cells in human lung cancer

James Clarke^{1,2}, Bharat Panwar¹, Ariel Madrigal¹, Divya Singh¹, Ravindra Gujar¹, Oliver Wood², Serena J. Chee^{2,3}, Simon Eschweiler¹, Emma V. King^{2,4}, Amiera S. Awad^{3,5}, Christopher J. Hanley² , Katy J. McCann², Sourya Bhattacharyya¹, Edwin Woo³, Aiman Alzetani³, Grégory Seumois¹ , Gareth J. Thomas², Anusha-Preethi Ganesan¹, Peter S. Friedmann⁵, Tilman Sanchez-Elsner⁵ , Ferhat Ay¹ , Christian H. Ottensmeier^{2*} , and Pandurangan Vijayanand^{1,5,6*} 

High numbers of tissue-resident memory T (T_{RM}) cells are associated with better clinical outcomes in cancer patients. However, the molecular characteristics that drive their efficient immune response to tumors are poorly understood. Here, single-cell and bulk transcriptomic analysis of T_{RM} and non- T_{RM} cells present in tumor and normal lung tissue from patients with lung cancer revealed that PD-1-expressing T_{RM} cells in tumors were clonally expanded and enriched for transcripts linked to cell proliferation and cytotoxicity when compared with PD-1-expressing non- T_{RM} cells. This feature was more prominent in the T_{RM} cell subset coexpressing PD-1 and TIM-3, and it was validated by functional assays *ex vivo* and also reflected in their chromatin accessibility profile. This PD-1⁺TIM-3⁺ T_{RM} cell subset was enriched in responders to PD-1 inhibitors and in tumors with a greater magnitude of CTL responses. These data highlight that not all CTLs expressing PD-1 are dysfunctional; on the contrary, T_{RM} cells with PD-1 expression were enriched for features suggestive of superior functionality.

Introduction

In lung cancer and many other solid tumors, patient survival is positively correlated with an effective adaptive antitumor immune response (Galon et al., 2006). This response is mediated primarily by CD8⁺ CTLs. Because CTLs in tumors are chronically activated, they can become “exhausted,” a hyporesponsive state that, in the setting of infection, prevents inflammatory damage to healthy tissue (Wherry, 2011). Exhaustion involves up-regulation of surface molecules such as PD-1 and TIM-3, alongside a gradual diminution of functional and proliferative potential (Pardoll, 2012). Anti-PD-1 therapies have revolutionized cancer treatment by inducing durable responses in some patients (Robert et al., 2015). Given the association of PD-1 with exhaustion and the description of CTLs expressing PD-1 in human cancers, exhausted CTLs are generally assumed to be the cells reactivated by anti-PD-1 therapy, though definitive evidence for this is lacking in humans (Simon and Labarriere, 2017).

Though anti-PD-1 therapies can eradicate tumors in some patients, they also lead to serious “off-target” immune-mediated

adverse reactions (June et al., 2017), calling for research to identify features unique to tumor-reactive CTLs. One subset of CTLs that may harbor such distinctive properties are tissue-resident memory T (T_{RM}) cells which mediate the response to antitumor vaccines (Nizard et al., 2017) and facilitate rejection of tumors in animal models (Malik et al., 2017). T_{RM} cell responses have also recently been shown by our group (Ganesan et al., 2017) and others (Djenidi et al., 2015) to be associated with better survival in patients with solid tumors. The molecular features of T_{RM} cell responses have been characterized in infection models and involve rapid clonal expansion and up-regulation of molecules aiding recruitment and activation of additional immune cells alongside the conventional effector functions of CTLs (Schenkel and Masopust, 2014). To date, the properties of T_{RM} cells found in the background lung, compared with those in the tumor, are not fully elucidated. Furthermore, the properties of these cell subsets in the context of immunotherapy are still poorly understood. To address this question, we

¹La Jolla Institute for Immunology, La Jolla, CA; ²National Institute for Health Research and Cancer Research UK Southampton Experimental Cancer Medicine Center, National Institute for Health Research Southampton Biomedical Research Center, Cancer Sciences Unit, Faculty of Medicine, University of Southampton, Southampton, UK; ³Southampton University Hospitals National Health Service Foundation Trust, Southampton, UK; ⁴Department of Otolaryngology, Poole Hospital National Health Service Foundation Trust, Poole, Dorset, UK; ⁵Clinical and Experimental Sciences, National Institute for Health Research Southampton, Respiratory Biomedical Research Unit, University of Southampton, Faculty of Medicine, Southampton, UK; ⁶Department of Medicine, University of California San Diego, La Jolla, CA.

*C.H. Ottensmeier and P. Vijayanand contributed equally to this paper; Correspondence to Pandurangan Vijayanand: vijay@lji.org; Christian H. Ottensmeier: C.H.Ottensmeier@soton.ac.uk.

© 2019 Clarke et al. This article is distributed under the terms of an Attribution–Noncommercial–Share Alike–No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms/>). After six months it is available under a Creative Commons License (Attribution–Noncommercial–Share Alike 4.0 International license, as described at <https://creativecommons.org/licenses/by-nc-sa/4.0/>).

compared the transcriptome of T_{RM} and non- T_{RM} CTLs present in tumor and normal lung tissue samples from treatment-naïve patients with lung cancer. Furthermore, we investigated the same tissue-resident populations in head and neck squamous cell carcinoma (HNSCC) and during immunotherapy regimens.

Results

T_{RM} cells in human lungs are transcriptionally distinct from previously characterized T_{RM} cells

We analyzed the transcriptome of CTLs isolated from lung tumor and adjacent uninvolved lung tissue samples obtained from patients ($n = 30$) with treatment-naïve lung cancer (Table S1) sorted according to CD103 expression to separate T_{RM} cells from non- T_{RM} cells, as previously described (Ganesan et al., 2017). Lung CD103⁺ and CD103⁻ CTLs clustered separately and showed differential expression of nearly 700 transcripts, including several previously linked to T_{RM} cell phenotypes; we validated CD49A and KLRG1 at the protein level, as described previously (Fig. 1 A; Fig. S1, A–C; and Table S2; Hombrink et al., 2016; Ganesan et al., 2017). Gene set enrichment analysis (GSEA) showed that the pattern of expression of these transcripts correlated with a murine core tissue-residency signature in CTLs isolated from both lung and tumor samples (Fig. S2 D and Table S3; Milner et al., 2017). Together, these data confirm that CD103⁺ CTLs in human lungs and tumors are highly enriched for T_{RM} cells; for simplicity, hereafter, we refer to CD103⁺ CTLs as T_{RM} cells and CD103⁻ CTLs as non- T_{RM} cells.

We next compared differentially expressed transcripts between lung T_{RM} and non- T_{RM} cells with those reported for other T_{RM} cells. The comparison with gene signatures of human skin T_{RM} cells (Cheuk et al., 2017) and that of murine T_{RM} cells isolated from multiple organs (Milner et al., 2017) revealed limited, yet statistically significant overlap ($\leq 5\%$; Fig. 1, B and C), which suggested that core tissue-residency features were well preserved. However, those differentially expressed transcripts that were not preserved across organs or species were not significantly enriched (Fig. S1 E). Thus, the transcriptional program, outside of a core tissue-residency program of human lung T_{RM} cells, is distinct from that of human skin T_{RM} cells and murine T_{RM} cells present in several organs. Importantly, many of the features observed in human lung T_{RM} cells have not been previously reported (Table S2; Hombrink et al., 2016).

PD-1 expression is a feature of lung and tumor T_{RM} cells

We next asked if T_{RM} cells in lung tumors share tissue-residency features (Materials and methods) with T_{RM} cells in adjacent normal lung tissue. Nearly one third (89/306) of the T_{RM} properties (i.e., transcripts differentially expressed between CD103⁺ and CD103⁻ CTLs) in tumors were shared with those of normal lung T_{RM} cells (Fig. 1, D and E; and Table S4). Coexpression analysis (Fig. 1 F) and weighted gene coexpression network analysis (WGCNA; Fig. 1 G; Materials and methods) of the 89 “shared tissue residency” transcripts revealed a number of novel genes whose expression was highly correlated with known tissue-residency genes (*SIPRI*, *SIPR5*, *ITGA1* [CD49A], *ZNF693* [HOBIT], and *RBPJ*; Mackay et al., 2013, 2016; Hombrink

et al., 2016), suggesting that their products may also play important roles in the development, trafficking, or function of T_{RM} cells (Fig. 1, D–G; and Table S4). Notable examples encoding products likely to be involved in T_{RM} cell functionality, migration, or retention include *GPR25* (Kim et al., 2013), *SRGAP3* (Bacon et al., 2013), *AMICA1* (Witherden et al., 2010), *CAPG* (Parikh et al., 2003), *ADAMI9* (Huang et al., 2005), and *NUAK2* (Namiki et al., 2011; Fig. 1, F and G; and Fig. 1 H, upper panel).

Another important shared tissue-residency transcript was *PDCDI*, encoding PD-1 (Fig. 1 H, lower panel). We confirmed at the protein level that PD-1 is expressed at higher levels in both tumor and lung T_{RM} cells compared with non- T_{RM} cells (Fig. 1 I). Although PD-1 expression is considered typical of exhausted T cells, as well as activated cells (Pardoll, 2012), recent reports have suggested that high PD-1 expression is a tissue-residency feature of murine brain T_{RM} cells independent of antigen stimulation (Prasad et al., 2017; Shwetank et al., 2017) and of murine T_{RM} cells from multiple organ systems (Milner et al., 2017). In support of the conclusion that high expression of PD-1 reflects tissue residency rather than exhaustion, we found that when T_{RM} and non- T_{RM} cells isolated from both lung and tumor tissue were stimulated ex vivo (Materials and methods), they showed robust up-regulation of TCR-activation-induced genes (*NR4A1*, *CD69*, *TNFRSF9* [4-1BB], and *EGR2*) and cytokines (*TNF* and *IFNG*; Fig. 1, J and K; and Tables S1 and S5). In addition to *PDCDI*, shared tissue-residency transcripts included several genes (*SPRY1* [Collins et al., 2012], *TMIGD2* [Janakiram et al., 2015], *CLNK* [Utting et al., 2004], and *KLRCl* [Rapaport et al., 2015]) that encode products reported to play a regulatory role in other immune cell types (Fig. 1 H, lower panel). We speculate that the expression of these inhibitory molecules may restrain the functional activity of tumor T_{RM} cells and may represent targets for future immunotherapies. Overall, our transcriptomic analysis of T_{RM} cells has identified molecules that are potentially important for the function of T_{RM} cells and thus serves as an important resource for investigating the biology of human T_{RM} cells.

Tumor T_{RM} cells display greater clonal expansion

To identify features unique to tumor T_{RM} cells, we compared the transcriptome of T_{RM} cells and non- T_{RM} cells from both normal lung and tumors and detected 92 differentially expressed transcripts (Fig. 2 A and Table S4) specifically in this subset, hence termed “tumor T_{RM} -enriched transcripts.” Reactome pathway analysis of these tumor T_{RM} -enriched transcripts showed significant enrichment for transcripts encoding components of the canonical cell cycle, mitosis, and DNA replication machinery (Fig. 2 B and Table S4). The tumor T_{RM} cell subset thus appears to be highly enriched for proliferating CTLs, presumably responding to tumor-associated antigens (TAAs), despite PD-1 expression. Unique molecular identifier (UMI)-based TCR sequencing assays (Materials and methods) revealed that T_{RM} cells in tumors expressed a significantly more restricted TCR repertoire than non- T_{RM} cells in tumors, as shown by significantly lower Shannon–Wiener and inverse Simpson diversity indices (Fig. 2 C and Table S6). Furthermore, the tumor T_{RM} population contained a higher mean percentage of expanded

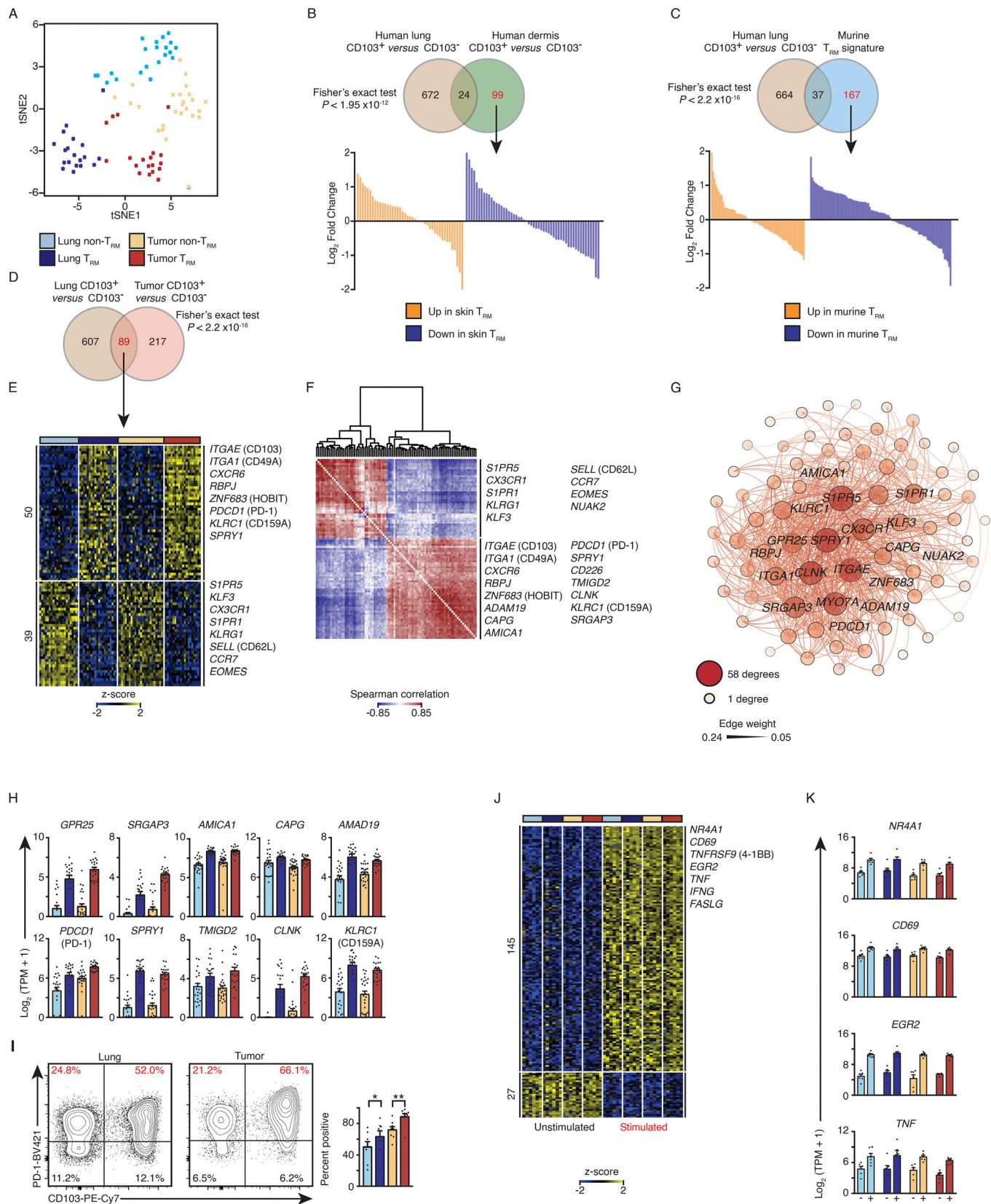


Figure 1. PD-1 expression is a feature of lung and tumor T_{RM} cells. (A) tSNE plot of tumor and lung CTL transcriptomes segregated by CD103 expression (lung non- T_{RM} = 21, lung T_{RM} = 20, tumor non- T_{RM} = 25, and tumor T_{RM} = 19). (B and C) Top: Venn diagrams showing overlap of transcripts differentially expressed in lung T_{RM} versus other previously characterized T_{RM} cells. Bottom: Waterfall plots represent the DESeq2 normalized fold change from the human lung comparison of genes not significantly (change twofold or less, with an adjusted P value of >0.05) differentially expressed between lung T_{RM} (CD103 $^{+}$) and

non- T_{RM} (CD103⁻) CTLs (marked in red font in the Venn diagram). **(D and E)** Venn diagram (D) and heat map (E) of RNA-seq analysis of 89 common transcripts (one per row) expressed differentially by lung T_{RM} versus lung non- T_{RM} and tumor T_{RM} versus tumor non- T_{RM} (pairwise comparison; change in expression of twofold with an adjusted P value of ≤ 0.05 [DESeq2 analysis; Benjamini-Hochberg test]), presented as row-wise z-scores of TPM counts; each column represents an individual sample; known T_{RM} or non- T_{RM} transcripts are indicated. The color scheme and number of samples are identical to A. **(F)** Spearman coexpression analysis of the 89 differentially expressed genes as in D and E; values are clustered with complete linkage. **(G)** WGCNA visualized in Gephi, the nodes are colored and sized according to the number of edges (connections), and the edge thickness is proportional to the edge weight (strength of the correlation). **(H)** Expression values according to RNA-seq data of the indicated differentially expressed genes shared by lung and tumor T_{RM} cells. Each symbol represents an individual sample, the bar represents the mean (colored as in A), and t-lines represents SEM. **(I)** Flow cytometry analysis of the expression of PD-1 versus that of CD103 on live and singlet-gated CD14⁻CD19⁻CD20⁻CD56⁻CD4⁻CD45⁺CD3⁺CD8⁺ cells obtained from lung cancer CTLs and matched paired lung CTLs. Right: Frequency of cell that express PD-1 in the indicated populations (*, $P \leq 0.05$; **, $P \leq 0.01$; $n = 8$; Wilcoxon rank-sum test); each symbol represents a sample, bars represent the mean, and t-lines represent SEM, colored as per A. **(J)** RNA-seq analysis of genes (row) commonly up- or down-regulated in the four cell types following 4 h of ex vivo stimulation; heat map as in E. **(K)** Bar graphs showing expression of transcripts in the indicated populations ($n = 6$ for all comparisons; as per J); left bar is ex vivo (-), and right bar is the expression profile following stimulation (+).

clonotypes (73% versus 52%, in tumor T_{RM} versus non- T_{RM} populations; Fig. 2 D). The most expanded clonotype in each patient comprised, on average, 19% of all T_{RM} cells with sequenced TCRs (three examples; Fig. 2 D, right, and Table S6), suggesting marked expansion of a single TAA-specific T cell clone in the tumor T_{RM} population, in keeping with recent publications (Guo et al., 2018; Savas et al., 2018).

Tumor T_{RM} -enriched transcripts that were highly correlated with cell cycle genes may encode products with important functions, as they are likely to reflect the molecular features of T_{RM} cells that are actively expanding in response to TAA (Table S4). *HAVCR2*, encoding the coinhibitory checkpoint molecule TIM-3, was most correlated and connected with cell cycle genes (Fig. 2, E–G). Thus, *TIM-3* expression may be a feature of lung tumor T_{RM} cells that is not linked to exhaustion but rather reflects a state of functionality, as the other transcripts that correlated with expression of *TIM-3* and cell cycle genes encode molecules that likely confer additional functionality, such as CD39 (encoded by *ENTPDI*; Pallett et al., 2017), CXCL13 (Bindea et al., 2013), CCL3 (Castellino et al., 2006), TNFSF4 (OX-40 ligand; Croft et al., 2009), and a marker of antigen-specific engagement (4-1BB; Bacher et al., 2016; Fig. 2, E–G). Robust expression of this set of molecules was not observed in either human lung T_{RM} cells or in the mouse T_{RM} signatures (Hombrink et al., 2016; Mackay et al., 2016; Milner et al., 2017), indicating that the human tumor T_{RM} population contains novel cell subsets.

Single-cell transcriptomic analysis reveals previously uncharacterized T_{RM} subsets

To determine whether tumor T_{RM} -enriched transcripts are expressed in all or only a subset of the tumor T_{RM} populations, we performed single-cell RNA sequencing (RNA-seq) assays in CD103⁺ and CD103⁻ CTLs isolated from tumor and matched adjacent normal lung tissue from an additional 12 patients with early-stage lung cancer (Table S1). Analysis of the ~12,000 single-cell transcriptomes revealed five clusters (Materials and methods) of T_{RM} cells and four clusters of non- T_{RM} cells with varying frequency per donor, highlighting the importance of studying multiple patient samples (Fig. 3, A and B; and Fig. S2, A–C). Among the five T_{RM} clusters, clusters 1–3 (light purple, purple, and blue, respectively) contained a greater proportion of the tumor T_{RM} population, while clusters 4 and 5 (green and red) contained more lung T_{RM} cells (Fig. 3 C). Most

strikingly, clusters 1–3 contained very few lung T_{RM} cells (Fig. 3 C). We infer that the tumor T_{RM} -enriched transcripts detected in our analysis of bulk populations (Fig. 2 A) were likely to be contributed by cells in these subsets.

In agreement with that conclusion, cells in cluster 1 (light purple) expressed high levels of the 25 cell cycle-related tumor T_{RM} -enriched transcripts (Fig. 3 D; Macosko et al., 2015), indicating that the enrichment of cell cycle transcripts in the bulk tumor T_{RM} population was contributed by this relatively small subset. Because these cells are actively proliferating, they likely represent TAA-specific cells. The majority of cells in this cycling cluster were from the tumor T_{RM} population (Fig. 3 E). These cells, as well as those in the larger cluster 2 (purple), were highly enriched for other prominent tumor T_{RM} -enriched transcripts like *HAVCR2* (*TIM-3*), including those encoding products that could confer additional functionality, such as CD39 (Ganesan et al., 2017), CXCL13 (Bindea et al., 2013), and CCL3 (Castellino et al., 2006; Fig. 3 F), consistent with recent reports (Guo et al., 2018; Savas et al., 2018), but with a noteworthy caveat that transcript expression does not necessarily reflect functionality. This shared expression pattern suggests that the cycling cluster (cluster 1, light purple) may represent cells in cluster 2 that are entering the cell cycle. Confirming this idea, cell-state hierarchy maps of all T_{RM} cells, constructed using Monocle2 (Trapnell et al., 2014; Materials and methods), revealed that cells in cluster 2 were most similar to the cycling T_{RM} cells (cluster 1, Fig. 3 G). Additionally, when we performed hierarchical clustering (Materials and methods) of these cells, we noted that the proliferating cluster 1 clustered more with cells assigned to cluster 2, than the other T_{RM} clusters (Fig. 3 H). This finding was corroborated when we calculated the average distance in principal component (PC) space (Materials and methods) between each cell in cluster 1 to the other T_{RM} clusters (Fig. S2 D).

T cells expressing *TCF7*, encoding the transcription factor TCF-1, are linked to “stemness” and have been shown to sustain T cell expansion and responses following anti-PD-1 therapy during chronic infections and in tumor models in mice (Im et al., 2016; Utzschneider et al., 2016). Applying unbiased differential expression analysis (Model-based Analysis of Single-cell Transcriptomics; MAST), we found that among the tumor-infiltrating CTLs *TCF7* expression was enriched in the *CCR7* (CD197)- and *SELL* (CD62L)-expressing non- T_{RM} subset, likely to reflect central memory cells (Figs. 3 B and S2 E; light orange), with no significant enrichment observed in any of the T_{RM} clusters

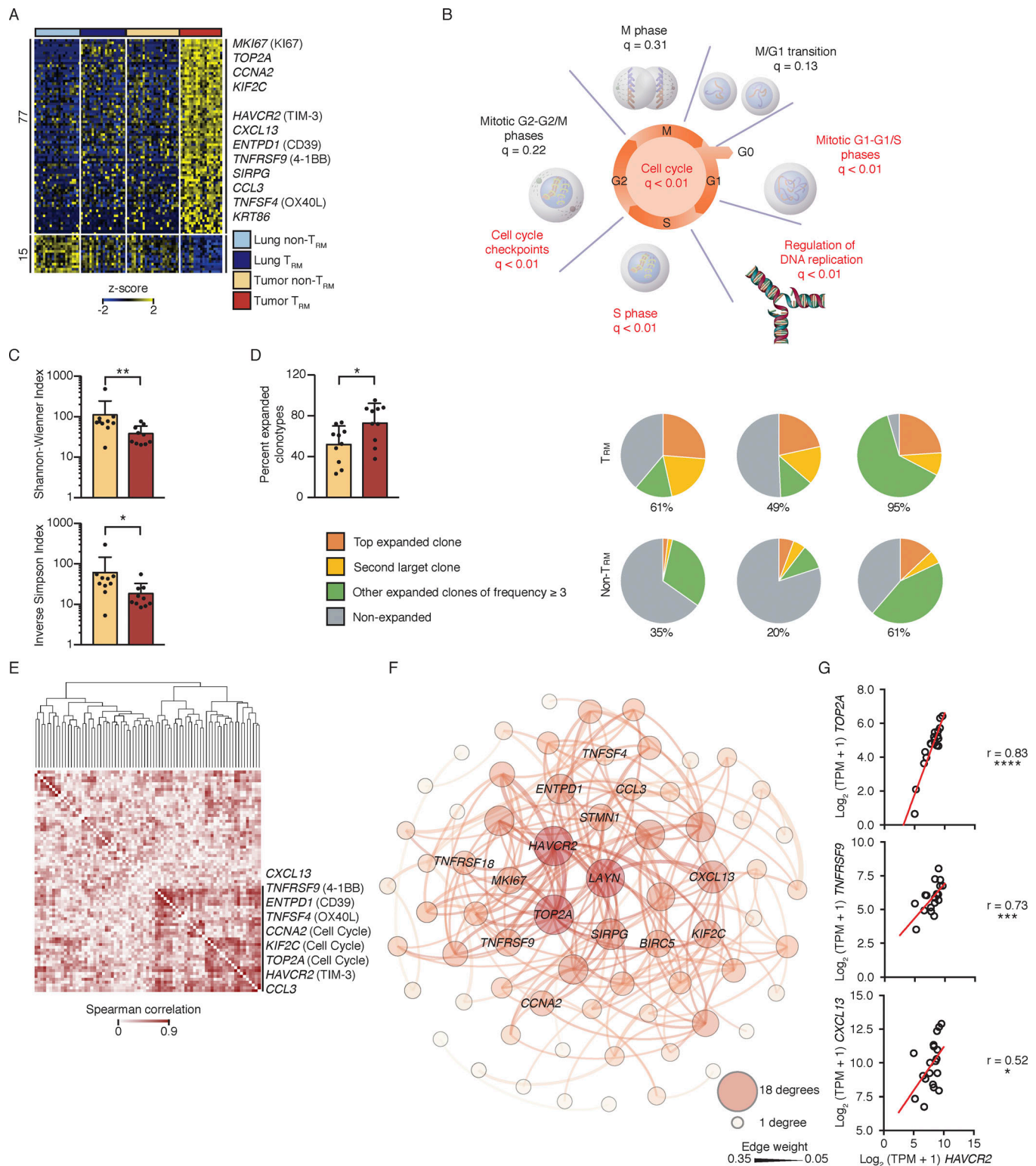


Figure 2. Tumor T_{RM} cells were clonally expanded. (A) RNA-seq analysis of transcripts (one per row) differentially expressed by tumor T_{RM} relative to lung T_{RM} , lung non- T_{RM} , and tumor non- T_{RM} (pairwise comparison; change in expression of twofold with an adjusted P value of ≤ 0.05 [DESeq2 analysis; Benjamini–Hochberg test]), presented as row-wise z-scores of TPM counts; each column represents an individual sample (lung non- T_{RM} = 21, lung T_{RM} = 20, tumor non- T_{RM} = 25, and tumor T_{RM} = 19). (B) Summary of overrepresentation analysis (using Reactome) of genes involved in the cell cycle that are differentially expressed by lung tumor T_{RM} cells relative to the other lung CTLs; q values represent FDR. (C) Shannon–Wiener diversity and inverse Simpson indices obtained using V(D) tools following TCR-seq analysis of β chains in tumor T_{RM} and tumor non- T_{RM} populations. Bars represent the mean, t-lines represent SEM, and symbols represent individual data points (*, $P \leq 0.05$; **, $P \leq 0.01$; $n = 10$ patients; Wilcoxon rank-sum test). (D) Left: Bar graphs show the percentage of total TCR β chains that were expanded (≥ 3 clonotypes). Bars represent the mean, t-lines represent SEM, and dots represent individual data points.

(*, $P \leq 0.05$; $n = 10$ patients). Right: Pie charts show the distribution of TCR β clonotypes based on clonal frequency. **(E)** Spearman coexpression analysis of the 77 genes up-regulated (A) in tumor T_{RM} cells; values are clustered with complete linkage. **(F)** WGCNA visualized in Gephi; the nodes are colored and sized according to the number of edges (connections), and the edge thickness is proportional to the edge weight (strength of correlation). **(G)** Correlation of the expression of *HAVCR2* (TIM-3) transcripts and the indicated transcripts in the tumor T_{RM} population; r indicates Spearman correlation value (*, $P \leq 0.05$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$; $n = 19$ patients).

(1 and 2) linked to cell proliferation (Fig. S2 E and Table S7) compared with other T_{RM} clusters. This finding is consistent with recent reports which showed that in mouse tumor models Tcf7⁺ tumor-infiltrating T cells were enriched for central memory, but not T_{RM} , cell features (Siddiqui et al., 2019), and in patients with melanoma, TCF7⁺ tumor-infiltrating T cells were mainly bystander cytotoxic T cells, whereas tumor-reactive cells were TCF7^{low/negative} (Li et al., 2019). Overall, our single-cell transcriptome uncovered additional phenotypically distinct subsets of lung and tumor T_{RM} cells that have not previously been described and are likely to play an important role in antitumor immune responses.

TIM-3⁺IL7R⁻ T_{RM} cell subset was enriched for transcripts linked to cytotoxicity

To dissect the molecular properties unique to tumor-infiltrating T_{RM} cells in each of the four larger T_{RM} clusters (clusters 2–5), we performed multiple pairwise single-cell differential gene expression analyses (Materials and methods). Over 250 differentially expressed genes showed higher expression in any one of the four clusters (Fig. 4 A and Table S7), indicating that cells in different clusters had divergent gene expression programs. For example, cells in cluster 3 were highly enriched for transcripts encoding heat shock proteins (e.g., *HSPA1A*, *HSPA1B*, and *HSP90AA1*), whereas cells in cluster 5, comprising T_{RM} cells from both normal lung and tumor tissue, expressed high levels of *IL7R*, which encodes the IL-7 receptor, a marker of memory precursor cells (Patil et al., 2018), and transcripts such as *GPR183* (Emgård et al., 2018), *MYADM* (Aranda et al., 2011), *VIM* (Niemenen et al., 2006), and *ANKRD28* (Tachibana et al., 2009), which encode proteins involved in cell migration and tissue homing (Fig. 4, A and B).

Because of their close relationship with cycling T_{RM} cells (Fig. 3, D, G, and H), we focused our analysis on the T_{RM} cells in cluster 2. The 91 transcripts enriched in cluster 2 compared with the other T_{RM} cell clusters (Fig. 4 A) included several that encoded products linked to cytotoxic activity such as *PRF1*, *GZMB*, *GZMA*, and *CTSW* (Patil et al., 2018), *RAB27A* (Stinchcombe et al., 2001), *ITGAE* (Franciszewicz et al., 2013), and *CRTAM* (Patil et al., 2018; Figs. 4 C and S3 A), as well as transcripts encoding effector cytokines and chemokines such as IFN- γ , *CCL3*, *CXCL13*, *IL-17A*, and *IL-26*. Cluster 2 also expressed high levels of transcripts encoding transcription factors known to promote the survival of memory or effector CTLs (ID2 [Yang et al., 2011], *STAT3* [Cui et al., 2011], *ZEB2* [Dominguez et al., 2015], and *ETS-1* [Muthusamy et al., 1995] or those that are involved in establishing and maintaining tissue residency (RBPJ, a key player in Notch signaling [Hombrink et al., 2016], and *BLIMP1* [Mackay et al., 2016], encoded by *PRDM1*; Figs. 4 A and S3 A). T_{RM} cells in cluster 2 also strongly expressed *ENTPDI* (Fig. 4, A and B), which

encodes CD39, an ectonucleotidase that cleaves ATP, which may protect this T_{RM} cell subset from ATP-induced cell death in the ATP-rich tumor microenvironment (Pallett et al., 2017) and has recently been shown to be enriched for tumor neoantigen-specific CTLs (Duhen et al., 2018; Simoni et al., 2018). This expression pattern likely confers highly effective and sustained antitumor immune function; in combination with earlier results, we conclude that this TIM-3⁺IL7R⁻ T_{RM} subset likely represents TAA-specific cells that were enriched for transcripts linked to cytotoxicity.

Intriguingly, T_{RM} cells in cluster 2 (TIM-3⁺IL7R⁻ subset) expressed the highest levels of *PDCDI* transcripts (Fig. 4 A) and were enriched for transcripts encoding other molecules linked to inhibitory functions such as TIM-3, TIGIT (Chan et al., 2014), and CTLA4 (Pardoll, 2012), and inhibitors of TCR-induced signaling and activation, like CBLB, SLAP, DUSP4, PTPN22, and NR3C1 (glucocorticoid receptor; Fig. 4, A–C; and Fig. S3 A; Huang et al., 2012; Xiao et al., 2015; Engler et al., 2017). Nonetheless, these T_{RM} cells expressed high transcript levels for cytotoxicity molecules (perforin, granzyme A, and granzyme B) and several costimulatory molecules such as 4-1BB, ICOS, and GITR (*TNFRSF18*; Figs. 4 C and S3 A; Pardoll, 2012). This coexpression program appeared to be specific to the tumor T_{RM} compartment, given it was also reflected in a SAVER (single-cell analysis via expression recovery)-imputed (Huang et al., 2018) coexpression profile being identified specifically in the T_{RM} subsets, but not the non- T_{RM} subsets (Fig. S3 B). In addition, direct comparison (Materials and methods) of the transcriptome of all *PDCDI*-expressing T_{RM} cells versus the non- T_{RM} cells present in tumors confirmed that *PDCDI*⁺ T_{RM} cells displayed significantly higher expression of transcripts linked to effector function (*IFNG*, *CXCL13*, *GZMB*, and *CCL3*) when compared with *PDCDI*⁺ non- T_{RM} cells and CTLs not expressing *PDCDI* (Fig. 4 D and Table S7). More specifically *PDCDI*⁺ T_{RM} cells that also coexpressed TIM-3 (cells in cluster 2) showed the highest expression levels of effector molecules compared with other subsets (Fig. 4 D). Overall, these findings agree with the bulk RNA-seq analysis, indicating that in T_{RM} cells, expression of particular inhibitory molecules, such as PD-1 and TIM-3, does not reflect exhaustion.

PD-1- and TIM-3-expressing tumor-infiltrating T_{RM} cells are not exhausted

To further address whether *PDCDI*-expressing T_{RM} cells in cluster 2 (TIM-3⁺IL7R⁻ T_{RM} cells) were exhausted or functionally active, we performed single-cell RNA-seq in tumor-infiltrating T_{RM} and non- T_{RM} cells using the more sensitive Smart-seq2 assay (Table S1). This also enabled paired transcriptomic and TCR clonotype analysis (Picelli et al., 2014; Patil et al., 2018). We reconstructed the TCR β chains (Materials and methods) in 80.5% of single cells, the TCR α chain in 76.6% of cells, and both

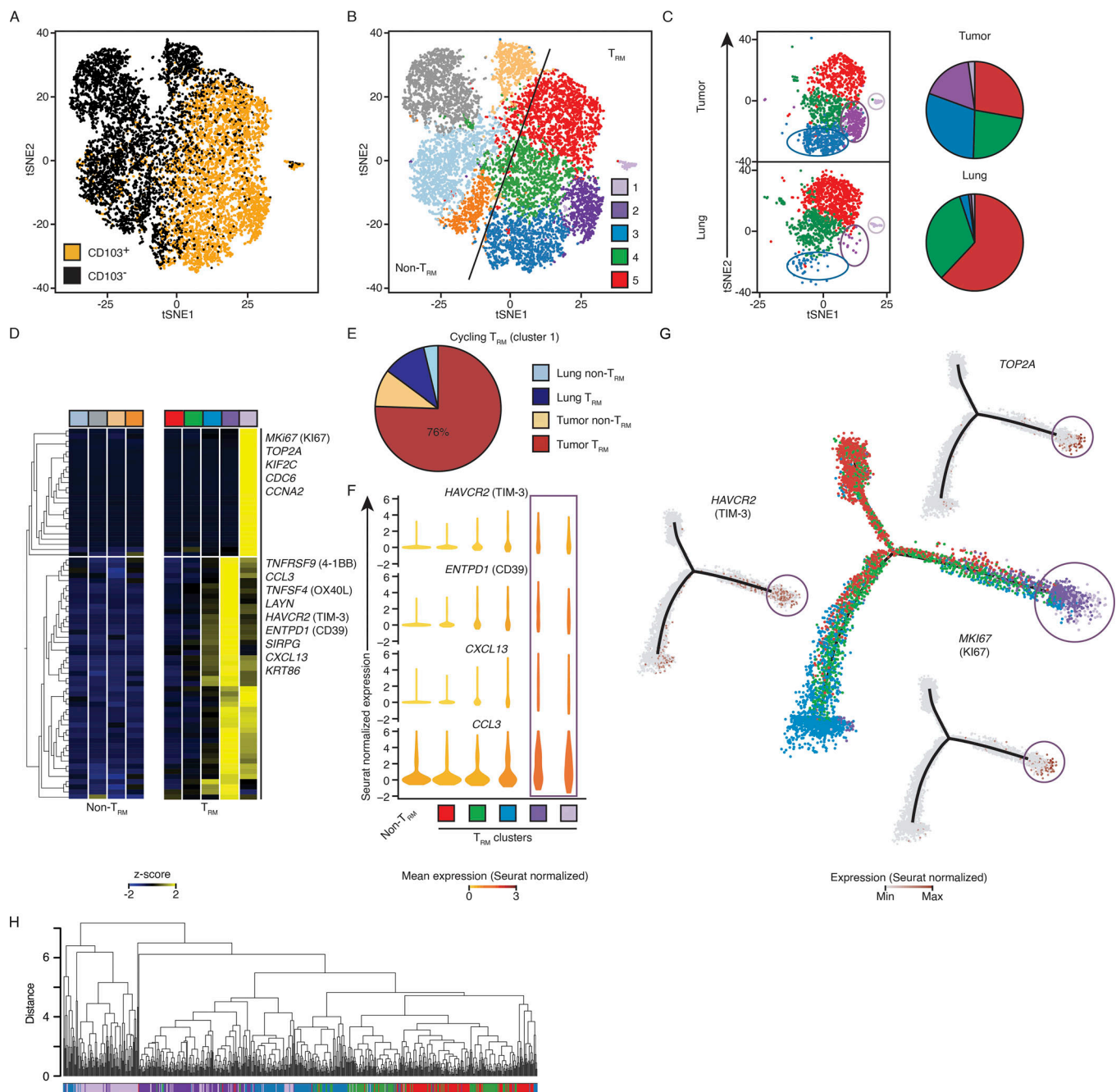


Figure 3. Single-cell transcriptomic analysis reveals previously uncharacterized T_{RM} subsets. (A) tSNE visualization of ~12,000 live and singlet-gated CD14⁺CD19⁺CD20⁺CD4⁺CD56⁺CD45⁺CD3⁺CD8⁺ single-cell transcriptomes obtained from 12 tumors and 6 matched normal lung samples. Each symbol represents a cell; color indicates protein expression of CD103 detected by flow cytometry. (B) Seurat clustering of cells in A, identifying nine clusters. (C) Cells from tumor and lung were randomly down-sampled to equivalent numbers of cells. Left: Distribution of T_{RM}-enriched clusters in tumor and lung. Right: Pie chart representing the relative proportions of cells in each T_{RM} cluster. (D) Expression of transcripts previously identified as up-regulated in the bulk tumor T_{RM} population (Fig. 2A) by each cluster; each column represents the average expression in a particular cluster. (E) Breakdown of cell type and tissue localization of cells defined as being in cluster 1 (shown in light purple in B and D). (F) Violin plots of expression of example tumor T_{RM} genes in each T_{RM}-enriched cluster (square below indicates the cluster type) and non-T_{RM} cells; shape represents the distribution of expression among cells and color represents the Seurat-normalized average expression counts. (G) Cell-state hierarchy maps generated by Monocle2 bioinformatic modeling of the T_{RM} clusters. In the center plot, each dot represents a cell colored according to Seurat-assigned cluster. Surrounding panels show relative Seurat-normalized expression of the indicated genes. (H) Cluster analysis of the location in PCA space for cells. Each cluster was randomly down sampled to the equivalent size of the smallest cluster ($n = 135$ cells per cluster, 675 total). The correlation method was Spearman, and the dataset was clustered with average linkage.

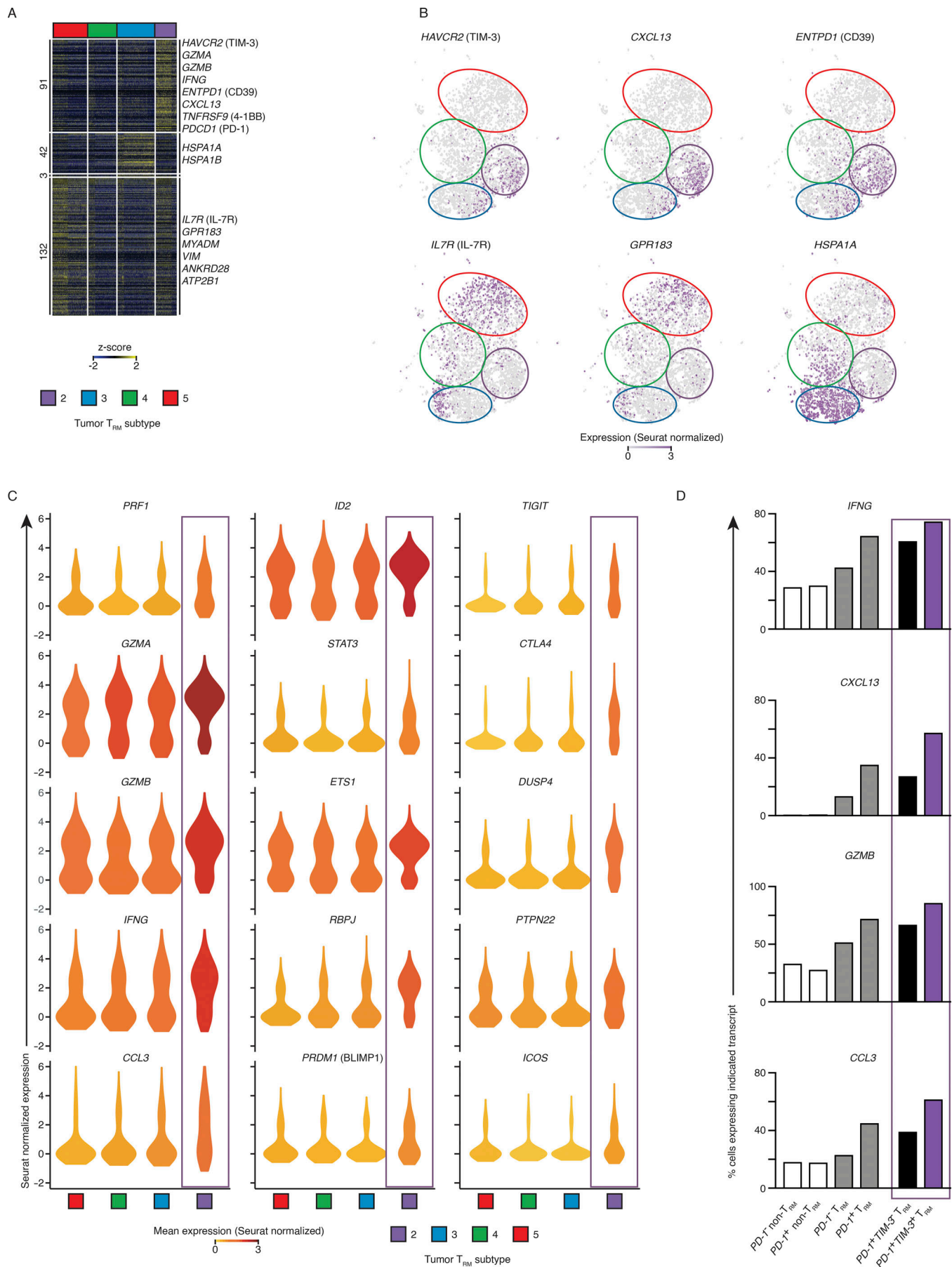


Figure 4. The $TIM-3^+IL7R^-$ T_{RM} subset was enriched for transcripts linked to cytotoxicity. (A) Single-cell RNA-seq analysis of transcripts (one per row) uniquely differentially expressed by each tumor T_{RM} subset in pairwise analysis compared with other clusters (adjusted P value of ≤ 0.01 ; MAST analysis), presented as row-wise z-scores of Seurat-normalized count; each column represents an individual cell. Horizontal breaks separate genes enriched in each of the four tumor T_{RM} subsets. (B) Seurat-normalized expression of indicated transcripts identified as differentially enriched in each cluster (as per A), overlaid across the tSNE plot, with expression levels represented by the color scale. (C) Violin plot of expression of functionally important genes identified as significantly enriched in a tumor T_{RM} subset; shape represents the distribution of expression among cells and color represents the Seurat-normalized average expression (cluster identification as per A). (D) Percentage of cells expressing the indicated transcripts in each population, where positive expression was defined as >1 Seurat-normalized count; " T_{RM} " corresponds to tumor-infiltrating CTLs isolated from clusters 2, 3, 4, and 5, and "non- T_{RM} " corresponds to the remaining cells not assigned into the proliferating T_{RM} cluster 1.

chains in 66.6% of cells (Table S8). As expected, clonally expanded tumor-infiltrating T_{RM} cells, which are likely to be reactive to TAA, were significantly enriched for genes specific to $TIM-3^+IL7R^-$ T_{RM} cells (Fig. 5 A and Table S7). Among tumor-infiltrating CTLs, a greater proportion of $TIM-3$ -expressing (Materials and methods) T_{RM} cells were clonally expanded compared with other T_{RM} and non- T_{RM} cells (Fig. 5 B). Furthermore, $TIM-3$ -expressing T_{RM} cells were significantly enriched for key effector cytokines and cytotoxicity transcripts (Table S9), despite expressing significantly higher levels of *PDCDI* (Fig. 5 C).

The higher sensitivity of the SMART-seq2 assay compared with the high-throughput 10X Genomics platform also allowed coexpression analysis due to lower dropout rates (Patil et al., 2018). Coexpression analysis showed that expression of *PDCDI* and *HAVCR2* ($TIM-3$) correlated with that of activation markers (*TNFRSF9* and *CD74*), *IFNG*, and cytotoxicity-related transcripts more strongly in T_{RM} cells than in non- T_{RM} cells (Fig. 5 D). Specifically, *IFNG* and *PDCDI* expression levels were better correlated in $TIM-3$ -expressing T_{RM} cells compared with all T_{RM} cells and non- T_{RM} cells (Fig. 5 D and Table S10), and the proportion of cells strongly coexpressing these transcripts was notably higher (30.3% versus 9.2% versus 0.6%; Fig. 5 E). Furthermore, in concordance with our high-throughput single-cell RNA-seq assays, this higher-resolution analysis verified that *IFNG*, alongside additional effector molecule-associated transcripts (*CXCL13*, *CCL3*, *GZMB*, and *PRFI*), were particularly enriched in $TIM-3^+$ T_{RM} cells versus $TIM-3^-$ T_{RM} cells and both *PDCDI*⁺ and *PDCDI*⁻ non- T_{RM} cells (Fig. 4 D; and Fig. 5, F and G). Overall, these results strongly support that *PD-1*- and $TIM-3$ -expressing tumor-infiltrating T_{RM} cells were not exhausted but instead were enriched for transcripts (*IFNG*, *PRFI*, and *GZMB*) encoding for molecules linked to effector functions.

In keeping with our transcriptomic assays, when stimulated ex vivo, the percentage of tumor-infiltrating T_{RM} cells that coexpressed *PD-1* (stained before stimulation) and effector cytokines was significantly higher when compared with non- T_{RM} CTLs (Fig. 5 H; gating previously reported; Ganesan et al., 2017). Analysis directly ex vivo demonstrated there was also greater coexpression of the *PD-1* and cytotoxic-associated proteins granzyme A and granzyme B in T_{RM} cells, when compared with non- T_{RM} CTLs in the tumor (Fig. 5 I). These data verify that *PD-1* expression in the T_{RM} subset of tumor-infiltrating CTLs does not necessarily reflect dysfunctional properties.

Surface $TIM-3^+IL-7R^-$ status uniquely characterizes a set of tumor T_{RM} cells

We next evaluated the protein expression of selected molecules to better discern the tumor-infiltrating T_{RM} subsets. Multiparameter protein analysis of CTLs (Materials and methods) present in tumors and adjacent normal lung revealed a subset of T_{RM} (*CD103*⁺) cells localized distinctly when the data were visualized in two-dimensional space (Fig. 6 A). This subset consisted of tumor T_{RM} cells only from tumor tissue (purple circle, Fig. 6 A), and uniquely expressed high levels of $TIM-3$ and lacked *IL-7R*, indicating that this cluster is the same as the $TIM-3$ -expressing tumor T_{RM} cluster (cluster 2) identified by single-cell RNA analysis (Fig. 6 B). Consistent with the single-cell transcriptome analysis, the $TIM-3$ -expressing T_{RM} cluster was unique to the T_{RM} cells isolated from the tumor and expressed higher levels of *CD39*, *PD-1*, and *4-1BB* (Fig. 6, A and C). *PD-1* and $TIM-3$ expression levels were also positively correlated with expression of *4-1BB*, which is expressed following TCR engagement by antigen (Bacher et al., 2016; Fig. 6, D and E), indicating that these cells are highly enriched for TAA-specific cells. $TIM-3$ -expressing CTLs were also detected among tumor-infiltrating T_{RM} cells isolated from both treatment-naïve lung cancer and HNSCC samples (Fig. 6 F), but not among non- T_{RM} cells in these treatment naïve tumors or T_{RM} cells in lung. Multicolor immunohistochemistry (IHC) was used to confirm the presence of $TIM-3$ -expressing T_{RM} cells in lung tumor samples, which also showed enrichment of this subset in $TIL^{hi}T_{RM}^{hi}$ "immune hot" tumors (Fig. 6, G and H; and Table S11). These findings confirmed, at the protein level, the restriction of the $TIM-3^+IL-7R^-$ T_{RM} subset to tumors from two cancer types.

Single-cell transcriptome analysis of CTLs from anti-*PD-1* responders

We next analyzed tumor-infiltrating T cells from 19 biopsies (Table S1) with known divergent responses to anti-*PD-1* therapy. Flow cytometry analysis of tumor T_{RM} cells isolated from responding patients before, during, and after treatment, showed increased proportion of $TIM-3^+IL-7R^-$ T_{RM} cells compared with the tumor T_{RM} cells from our cohort of treatment-naïve lung cancer patients and those not responding to anti-*PD-1* (median, ~70% versus ~24% and ~19%, respectively; Figs. 7 A and S4 A). This population also expressed high levels of *PD-1* in samples before anti-*PD-1* therapy that decreased after treatment, which is likely reflective of the clinical antibody blocking flow cytometric analysis (Huang et al., 2017; Fig. 7 B). Given this population had high expression of *PD-1* (Figs. 6, A–E; and Fig. 7 B), we

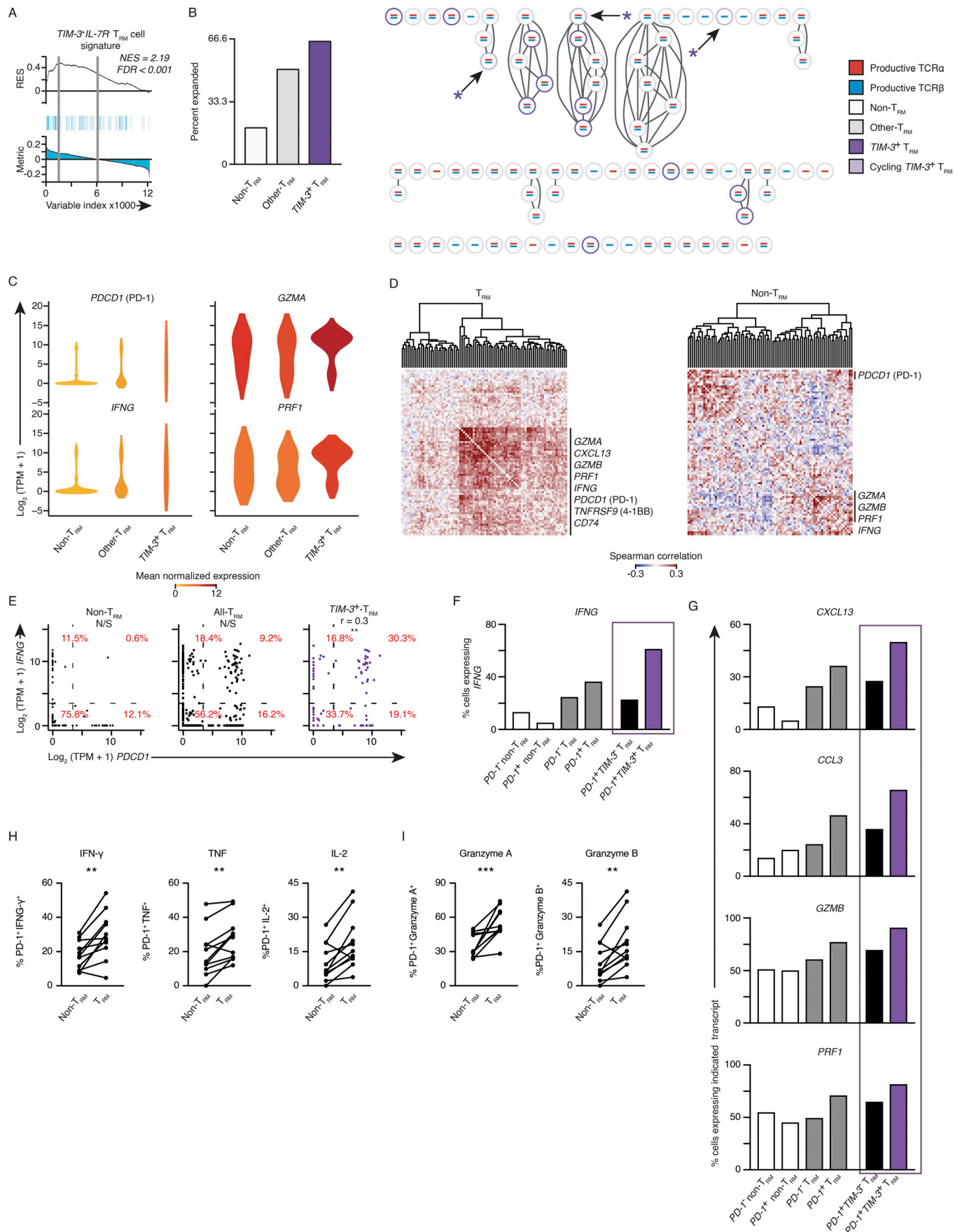


Figure 5. PD-1- and TIM-3-expressing tumor-infiltrating T_{RM} cells are not exhausted. (A) GSEA of the TIM-3⁺IL-7R⁻ T_{RM} subset signature in the transcriptome of clonally expanded tumor T_{RM} versus that of nonexpanded T_{RM} cells. Top: Running enrichment score (RES) for the gene set, from most enriched at the left to most underrepresented at the right. Middle: Positions of gene set members (blue vertical lines) in the ranked list of genes. Bottom: Value of the ranking metric. Values above the plot represent the normalized enrichment score (NES) and FDR-corrected significance. **(B)** Left: Percentage of cells that were clonally expanded in TIM-3⁺ (HAVCR2 ≥10 TPM counts) T_{RM} cells, remaining T_{RM} cells, and non-T_{RM} cells; clonal expansion was determined for cells from four and two patients for T_{RM} and non-T_{RM}, respectively. Right: Section of a clonotype network graph of cells from a representative patient. TIM-3⁺ (HAVCR2 ≥10 TPM counts) T_{RM} cells are marked with a purple circle; cells with >10 TPM counts expression of either MKI67 or TOP2A were considered cycling and denoted with an asterisk. **(C)** Violin plot of expression of indicated transcripts; shape represents the distribution of expression among cells and color represents average expression, calculated from the TPM counts (color coded as per B). **(D)** Spearman coexpression analysis of genes whose expression is enriched in the TIM-3⁺IL-7R⁻ T_{RM} cluster (Fig. 4 A) in tumor T_{RM} and non-T_{RM} populations, respectively; matrix is clustered according to complete linkage. **(E)** Correlation of PDCD1 and IFNG expression in non-T_{RM} cells, all T_{RM} cells, and then in TIM-3⁺ T_{RM}; each dot represents a cell. Percentages indicate the percentage of cells inside each of the graph sections (r indicates Spearman correlation value; **, P ≤ 0.01; N/S, no significance). **(F)** Percentage of cells expressing IFNG in each indicated population segregated on PD-1⁺ (PDCD1 ≥ 10 TPM counts). The final two bars are the T_{RM} population, as segregated by having expression of HAVCR2 (TIM-3) ≥ 10 TPM counts. **(G)** Percentage of cells expressing the indicated transcript as identified above the plot in each population (as per F). **(H)** Flow cytometry analysis of the percentage of PD-1⁺ T_{RM} and PD-1⁺ non-T_{RM} cells that express effector cytokines (as indicated above the graph) following 4 h of ex vivo stimulation. Gated on live and singlet-gated CD14⁻CD20⁻CD4⁻CD45⁺CD3⁺CD8⁺ cells obtained from lung cancer TILs, discriminated on CD103 expression (**, P ≤ 0.01; n = 11; Wilcoxon rank-sum test); each symbol represents a sample. Surface molecules (e.g., PD-1) were stained before stimulation. **(I)** Analysis of granzyme A and granzyme B directly ex vivo, gated and analyzed as per H (***, P ≤ 0.001; Wilcoxon rank-sum test).

concluded that these TIM-3⁺IL-7R⁻ T_{RM} cells are likely to be one of the key immune cell types that respond to anti-PD-1 therapy.

To comprehensively evaluate the molecular features and clonality of the CTLs (Fig. S4, A–C; and Materials and methods) responding to anti-PD-1 therapy, we performed paired single-cell transcriptomic and TCR analysis of CTLs isolated from biopsies both before and after therapy from two donors. This enabled us to maximize the usage of the material in these small, clinically difficult to obtain biopsies. Differential expression analysis of all CD8⁺ tumor-infiltrating CTLs revealed a significant enrichment of markers linked to cytotoxic function (*PRFI*, *GZMB*, and *GZMH*) and activation (*CD38*) in posttreatment samples when compared with pretreatment samples (Fig. 7, C and D; and Tables S1 and S12). Notably, we found increased expression of *ITGAE*, a marker of T_{RM} cells, in CTLs from posttreatment samples (Fig. 7, C and D). GSEA also showed that tumor-infiltrating T cells from posttreatment samples were enriched for T_{RM} features as well as those linked to the TIM-3⁺IL-7R⁻ T_{RM} subset (Fig. 7 E and Tables S4 and S7). Unbiased coexpression analysis of transcripts from posttreatment CTLs demonstrated that transcripts linked to cytotoxicity (*GZMH*) and activation (*CD38*) clustered together with the T_{RM} marker gene (*ITGAE*; Fig. 7 F and Table S12). Furthermore, we found several expanded TCR clones that were present before and after therapy (Fig. S4, D and E), which indicated that TIM-3-expressing T_{RM} cells (> 87% of the T_{RM} cells in pretreatment samples) with these clonotypes persisted in vivo for several weeks during treatment and largely maintained TIM-3 expression in posttreatment samples (>79% of the T_{RM} cells; Fig. S4 C). By restricting our analysis to these expanded clones, we found that the expression of *GZMB*, *PRFI*, *GZMH*, and *CD38* (Fig. S4 F) was increased in CTLs from posttreatment samples, which suggested that tumor-infiltrating CTLs with the same specificity displayed enhanced cytotoxic properties following anti-PD-1 treatment and that T_{RM} cells likely contributed to this feature.

To provide a further line of evidence for the functional potential of TIM-3⁺IL-7R⁻ T_{RM} cells and further characterize their epigenetic profile, we performed an assay for transposase-accessible chromatin using sequencing (OMNI-ATAC-seq;

Corces et al., 2017) on purified populations of tumor-infiltrating TIM3⁺IL7R⁻ T_{RM} and non-T_{RM} subsets pooled from lung cancer patients (n = 9; Table S1 and Fig. 6 B). These subsets clustered separately, highlighting the distinct chromatin accessibility profiles of these populations (Fig. 7 G). In keeping with our transcriptomic analyses (Fig. 1 E), we identified greater chromatin accessibility within 5 kb of the transcriptional start site of the *CD103* (*ITGAE*) and *KLF3* loci in the T_{RM} and non-T_{RM} compartment, respectively. Furthermore, consistent with single-cell transcriptional data, the TIM3⁺IL7R⁻ T_{RM} cells when compared with non-T_{RM} cells showed increased chromatin accessibility of genes encoding effector molecules such as granzyme B and IFN-γ, despite showing increased accessibility at the *PDCD1* (*PD-1*) and *TIM-3* (*HAVCR2*) loci (Fig. 7 H). Taken together, these epigenetic and transcriptomic data, combined with protein validation, highlight the potential functionality of TIM-3⁺IL-7R⁻ T_{RM} cells, which positively correlate with expression of PD-1 specifically in this subset.

Discussion

Our bulk and single-cell transcriptomic analysis showed that the molecular program of tumor-infiltrating T_{RM} cells is substantially distinct from that observed in the human background lung tissue or murine models. The most striking discovery was the identification of a TIM-3⁺IL-7R⁻ T_{RM} subset present exclusively in tumors. This subset expressed high levels of PD-1 and other molecules previously thought to reflect exhaustion. Surprisingly, however, they proliferated in the tumor milieu, were capable of robust up-regulation of TCR activation-induced genes, and exhibited a transcriptional program indicative of superior effector, survival, and tissue-residency properties. Functionality may not be truly reflected by transcript expression levels; hence, to support the conclusion that PD-1 expression does not reflect exhaustion in T_{RM} cells, we showed that the expression of the key effector cytokines, IL-2, TNF, and IFN-γ and the cytotoxicity molecules granzyme A and granzyme B was increased in PD-1-expressing T_{RM} cells when compared with non-T_{RM} cells. When compared with recent reports on

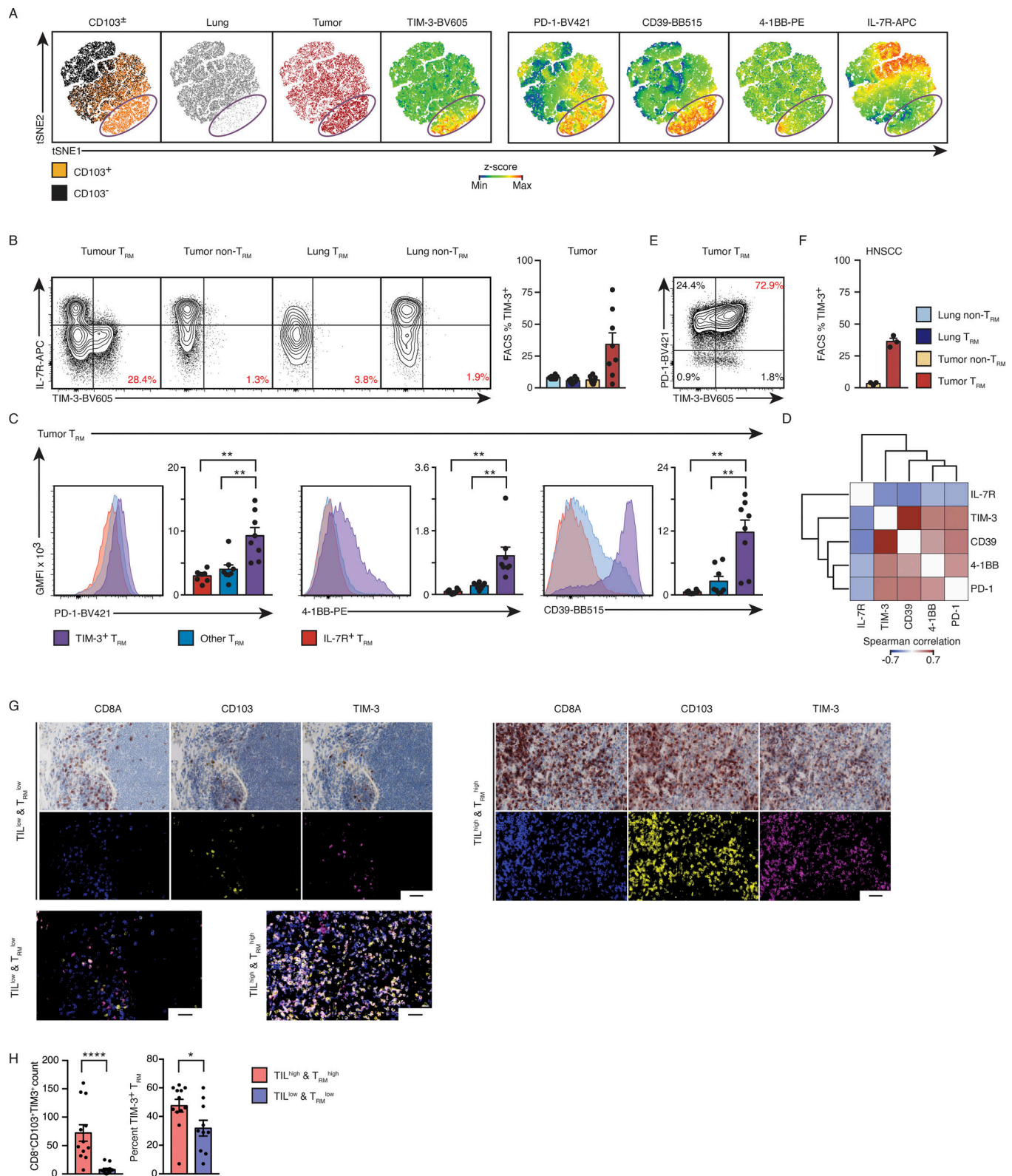


Figure 6. Surface TIM-3⁺IL-7R⁻ status uniquely characterizes a set of tumor T_{RM} cells. (A) tSNE visualization of flow cytometry data from 3,000 randomly selected live and singlet-gated CD14⁻CD19⁻CD20⁻CD56⁻CD4⁺CD45⁺CD3⁺CD8⁺ cells isolated from eight paired tumor and lung samples; each cell is represented by a dot colored as T_{RM} or non-T_{RM} (left), tumor or lung (second and third from left), and according to z-score expression value of the protein indicated above the plot (remaining panels). **(B)** Left: Contour plots show the expression of TIM-3 versus IL-7R in the cell type and tissue indicated above the plot; percentage of tumor CTLs (gated as above) in the indicated populations that express TIM-3 is shown. Right: Quantification of TIM-3⁺ in cells isolated from each tissue location; each symbol represents an individual sample. The small line indicates SEM, and bars represent the mean and are colored as indicated (*, $P \leq 0.05$; $n = 8, 3$). **(C)** Geometric mean fluorescent intensity (GMFI) of CD39, PD-1, and 4-1BB for each lung tumor-T_{RM} subset; bars represent the mean,

t-lines represent SEM, and each symbol represents data from individual samples (**, $P \leq 0.01$; $n = 8$; Wilcoxon rank-sum test); representative histograms are shown at left. **(D)** Coexpression analysis of flow cytometry data (C), as per Spearman correlation. **(E)** Contour plot highlighting the expression of PD-1 versus TIM-3 in CD14⁺CD19⁺CD20⁺CD56⁺CD4⁺CD45⁺CD3⁺CD8⁺CD103⁺ cells in a representative donor. **(F)** Analysis of three HNSCC samples, as per B. **(G)** Left: Representative multiplexed immunohistochemistry of CD8A, CD103, and TIM-3 in tumor specimens from a patient with TIL^{low}/T_{RM}^{low} tumor status; the representative false color image (middle) and the overlaid image (bottom) are shown. Right: As the left plots for a patient with TIL^{high}/T_{RM}^{high} tumor status (scale bar, 50 μ m). **(H)** Left: Quantification of the number of CD8A⁺CD103⁺TIM-3⁺ cells per region in biopsies defined as having a TIL^{high}/T_{RM}^{high} status versus TIL^{low}/T_{RM}^{low} status. Right: Percentage of CD8A⁺CD103⁺ CTLs expressing TIM-3 in each clinical subtype. Bars represent the mean, t-lines represent SEM, and symbols represent individual data points (*, $P \leq 0.05$; ****, $P \leq 0.0001$; $n = 21$; Mann-Whitney U test).

transcriptomic analysis of tumor-infiltrating lymphocytes (Guo et al., 2018; Savas et al., 2018), our in-depth study of T_{RM} cells in tumor and lungs clarifies the impact of PD-1 expression in different CTL subsets in the tumor and challenges the current dogma of PD-1 expression representing dysfunctional T cells in human tumors. While this protein-validated transcriptomic assessment was also corroborated by the results of a chromatin accessibility profile, an important caveat is that functional validation was performed using PMA and ionomycin stimulation, which does not fully reflect physiological TCR activation.

We defined a core set of genes commonly expressed in both lung and tumor T_{RM} cells, including a number of novel genes whose expression was highly correlated with known T_{RM} genes. Any of these genes may also be critically important for the development, trafficking, or function of lung or lung tumor-infiltrating T_{RM} cells. Some notable examples known or likely to have such functions are GPR25, whose closest homologue, GPR15 (Kim et al., 2013), enables homing of T cell subsets to and retention in the colon; AMICAI (Witherden et al., 2010), encoding JAML (junctional adhesion molecule-like), which contributes to the proliferation and cytokine release of skin-resident $\gamma\delta$ T cells; and SRGAP3, whose product functions in neuronal migration (Bacon et al., 2013). Thus, our study provides a valuable resource for defining molecules that are likely to be important for the development and function of human lung and tumor T_{RM} cells.

PDCDI was a prominent hit in the “shared lung tissue residency” gene list, and its expression was confirmed at the protein level in both lung and tumor T_{RM} cells. The fact that PD-1 was expressed by the majority of the T_{RM} cells isolated from uninflamed lung tissue of subjects with no active infection suggests that PD-1 might be constitutively expressed by these cells, as has been recently described for brain T_{RM} cells (Shwetank et al., 2017). Given that TIM-3⁺IL7R⁺ expressing tumor-infiltrating T_{RM} cells also express substantial levels of PD-1, we speculate that they may be the major cellular targets of anti-PD-1 therapy. We speculate that differences in the magnitude of this population of T_{RM}s could thus be an explanation for the variation in the clinical response to PD-1 inhibitors. We speculate further that the constitutive expression of PD-1 by the majority of T_{RM} cells in the lung tissue and presumably other organs (skin and pituitary gland) raises the possibility that anti-PD-1 therapy may nonspecifically activate potentially non-TAA-reactive T_{RM} cells to cause adverse immune reactions such as pneumonitis, dermatitis, and hypophysitis (June et al., 2017). Comprehensive analysis of the T_{RM} phenotype and TCR repertoire of CTLs present in tumor and organs affected by adverse reactions may substantiate these hypotheses. A recent study that compared PD-1^{high} versus other tumor-localized CTL populations demonstrated

that the presence of PD-1^{high} cells was predictive of response to anti-PD-1 therapy (Thommen et al., 2018). However, the authors did not segregate PD-1^{high} cells based on expression of T_{RM}-associated molecules; hence, this population of cells will have a mixture of PD-1⁺ non-T_{RM} cells and PD-1⁺ T_{RM} cells. Our findings have highlighted the profound differences in the properties of PD-1-expressing T_{RM} cells and non-T_{RM} cells. Hence, by studying a mixed population without delineating the contribution from PD-1-expressing T_{RM} cells, their study lacked the resolution to highlight the contribution of a specific T_{RM} subsets with transcriptomic features associated with increased functional properties and potential responsiveness to anti-PD-1 therapy.

Our findings also raise the question of which molecular players are essential for the generation and maintenance of this TIM-3⁺IL-R7⁺ subset of T_{RM} cells. Our analysis identified a number of potential transcription factors (e.g., STAT3, ID2, ZEB2, and ETS-1) and other molecules (e.g., PTPN22, DUSP4, LAYN, KRT86, and CD39) that are uniquely expressed in this subset and could thus be key players in their development or function, further underscoring the utility of the resource generated here.

The results herein also provide a rationale for assessing tumor T_{RM} subsets in both early- and late-phase studies of novel immunotherapies and cancer vaccines to provide early proof of efficacy as well as potential response biomarkers. The TIM-3⁺IL-R7⁺ T_{RM} subset can be readily isolated from tumor samples using the surface markers we identified and potentially expanded in vitro to screen and test for T_{RM}-targeted T cell therapies. In summary, our studies have provided rich molecular insights into the potential roles of human tumor-infiltrating T_{RM} subsets and thus should pave the way for developing novel approaches to improve T_{RM} immune responses in cancer.

Materials and methods

Ethics and sample processing

The Southampton and South West Hampshire Research Ethics Committee approved the study, and written informed consent was obtained from all subjects (Ganesan et al., 2017). Newly diagnosed, untreated patients with respiratory malignancies or HNSCC were prospectively recruited once referred (Chee et al., 2017; Wood et al., 2017). Freshly resected tumor tissue and, where available, matched adjacent nontumor tissue was obtained from lung cancer patients following surgical resection. HNSCC tumors were macroscopically dissected and slowly frozen in 90% FBS (Thermo Fisher Scientific) and 10% DMSO (Sigma) for storage, until samples could be prepared. Pre- and post-anti-PD-1 therapy samples were frozen as per HNSCC samples.

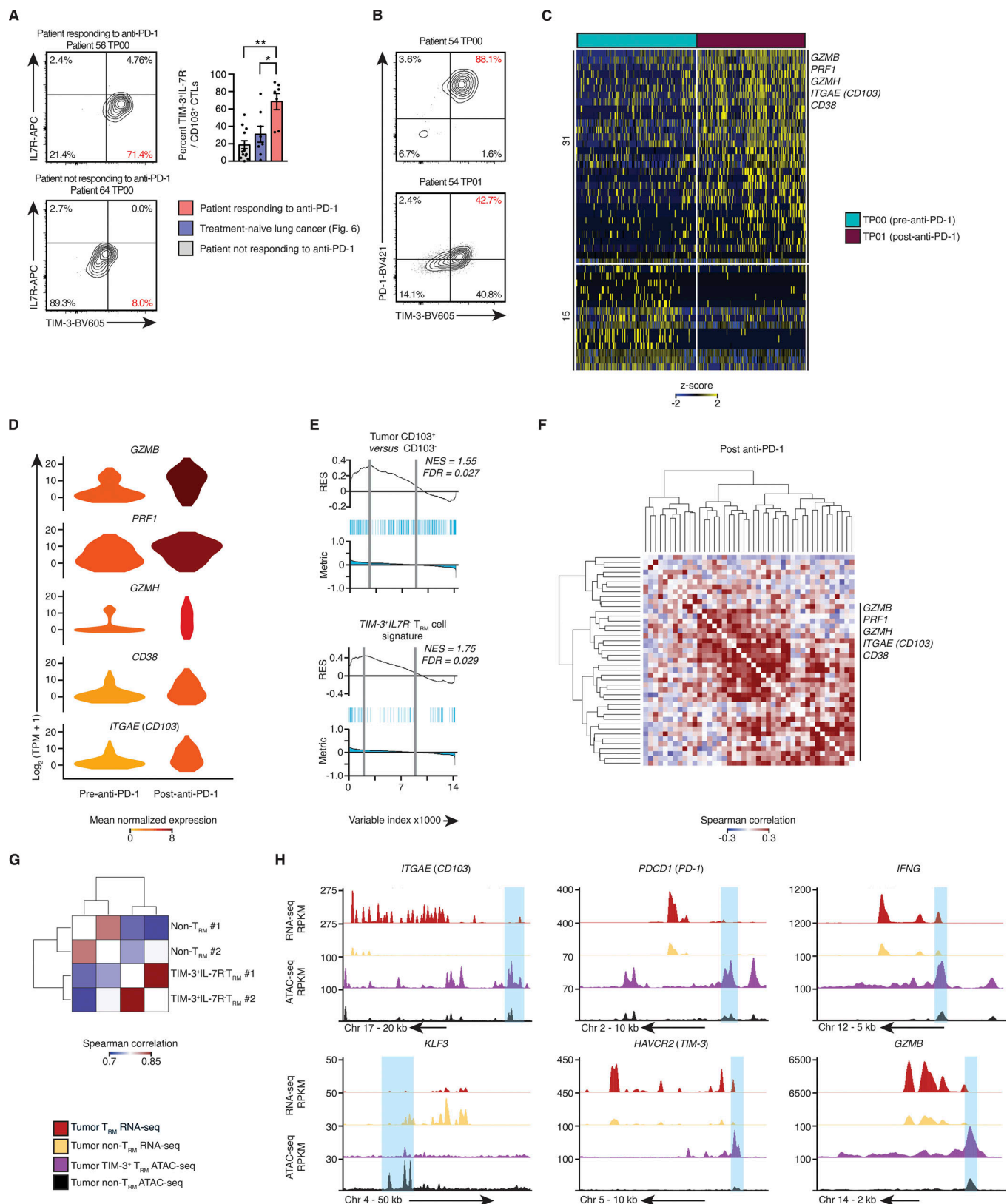


Figure 7. Single-cell transcriptome analysis of CTLs from anti-PD-1 responders. (A) Left: Contour plots show the expression of TIM-3 and IL-7R in CD14⁺CD19⁺CD20⁺CD4⁺CD45⁺CD3⁺CD8⁺CD103⁺ cells isolated from patients receiving anti-PD-1 treatment at the time point (TP) indicated above the plot; number in red indicates the percentage of tumor T_{RM} cells (CD8⁺CD103⁺) with TIM-3⁺IL-7R⁻ surface phenotype. Right: Quantification of the percentage of tumor-infiltrating TIM-3⁺IL-7R⁻ T_{RM} cells, isolated from the anti-PD-1 responding, nonresponding, and treatment-naïve patients. Bars represent the mean, t-lines represents SEM, and symbols represent individual data points (*, $P \leq 0.05$; **, $P \leq 0.01$; $n = 7, 8$, and 12 biopsies for responders, treatment-naïve

patients, and nonresponders, respectively; Mann–Whitney *U* test). **(B)** Contour plots demonstrate the expression of TIM-3 and PD-1 in the T_{RM} cells isolated from preimmunotherapy biopsies (gated as per Fig. 7 A). **(C)** Single-cell RNA-seq analysis of transcripts (one per row) differentially expressed between CTLs pre- and post-anti-PD-1 (MAST analysis), with an adjusted *P* value of ≤ 0.05 , presented as row-wise z-scores of TPM counts; each column represents a single cell ($n = 127$ and 151 cells, respectively). **(D)** Violin plot of expression of indicated transcripts differentially expressed between tumor-infiltrating CTLs isolated from pre- and post-anti-PD-1 treatment samples (as per C); shape represents the distribution of expression among cells, and color represents average expression, calculated from the TPM counts. **(E)** GSEA of the bulk tumor CD103⁺ versus CD103[−] transcriptional signature (Fig. 2 A) and TIM-3⁺IL7R[−] T_{RM} cell signature (Fig. 4 A) in tumor-infiltrating CTLs isolated from pre- and post-anti-PD-1 treatment samples. Top: RES for the gene set, from most enriched at the left to most underrepresented at the right. Middle: Positions of gene set members (blue vertical lines) in the ranked list of genes. Bottom: Value of the ranking metric. Values above the plot represent the NES and FDR-corrected significance. **(F)** Spearman coexpression analysis of transcripts enriched in tumor-infiltrating CTLs from post-anti-PD-1 treatment samples (C); matrix is clustered according to complete linkage. **(G)** Correlation analysis of all peaks identified in the OMNI-ATAC-seq libraries, pooled from nine donors across two experiments; cells were sorted on CD14[−]CD19[−]CD20[−]CD4[−]CD45⁺CD3⁺CD8⁺CD103⁺TIM-3⁺IL-7R[−] and CD14[−]CD19[−]CD20[−]CD4[−]CD45⁺CD3⁺CD8⁺CD103[−]. The matrix is clustered according to complete linkage. **(H)** University of California Santa Cruz genome browser tracks for key T_{RM} -associated gene loci as indicated above the tracks. RNA-seq tracks are merged from all purified bulk RNA-seq data, presented as reads per kilobase million (RPKM; as per Fig. 1 A; tumor non- $T_{RM} = 25$, tumor $T_{RM} = 19$; OMNI-ATAC-seq as per Fig. 7 G).

Flow cytometry of fresh material

Samples were processed as described previously (Wood et al., 2016; Ganesan et al., 2017). For sorting of fresh CTLs for population transcriptomic analysis, cells were first incubated at 4°C with Fc receptor (FcR) block (Miltenyi Biotec) for 10 min and then stained with a mixture of the following antibodies: anti-CD45-FITC (HI30; BioLegend), anti-CD4-PE (OKT4; BD Biosciences), anti-CD3-APC-Cy7 (SK7; BioLegend), anti-CD8A-PerCP-Cy5.5 (SK1; BioLegend), and anti-CD103-APC (Ber-ACT8; BioLegend) for 30 min at 4°C. Live/dead discrimination was by DAPI staining (Wood et al., 2016; Ganesan et al., 2017). CTLs were sorted based on CD103 expression using a BD FACS Aria-II (BD Biosciences) into ice-cold TRIzol LS reagent (Thermo Fisher Scientific).

Flow cytometry of cryopreserved material

For single-cell transcriptomic, stimulation assays, and phenotypic characterization, tumor and lung samples were first dispersed (as above) and cryopreserved in freezing media (50% complete RPMI (Fisher Scientific), 40% human decompartmented AB serum, 10% DMSO (both Sigma). Cryopreserved samples were thawed, washed twice with prewarmed (37°C) and room temperature MACS buffer, and prepared for staining as above. The second wash included an underlayer of FBS to help collect debris. The material was stained with a combination of anti-CD45-Alexa Fluor 700 (HI30; BioLegend), anti-CD3-APC-Cy7 (SK7; BioLegend), anti-CD8A-PerCP-Cy5.5 (SK1; BioLegend), and anti-CD103-PE-Cy7 (Ber-ACT8; BioLegend). Cells were counterstained where described with anti-CD19/20 (HIB19/2H7; BioLegend), anti-CD14 (HCD14; BioLegend), anti-CD56 (HCD56; BioLegend), and anti-CD4 (OKT4; BioLegend) for flow cytometric analysis and sorting. Live and dead cells were discriminated using propidium iodide (PI). Cells were stimulated for transcriptomic analysis when noted, with PMA and ionomycin, as previously described (Ganesan et al., 2017). For 10X single-cell transcriptomic analysis (10X Genomics), 1,500 cells each of CD103⁺ and CD103[−] CTLs from tumor and lung samples were sorted and mixed into 50% ice-cold PBS, 50% FBS (Sigma) on a BD Aria-II or Fusion cell sorter. From the 12 patients used for 10X genomics, matched lung tissue was used from 6 of these patients. For assessments of the bulk transcriptome following stimulation, CTLs were collected by sorting 200 cells into 8 μ l lysis buffer (Triton X-100 [0.1%] containing RNase inhibitor

[1 U/ μ l and deoxyribonucleotide triphosphate mix (2.5 mM) on an Aria-II (BD)]; for Smart-seq2-based single-cell analysis, CTLs were sorted as above, using single-cell purity, into 4 μ l lysis buffer on a BD Aria-II, as described previously (Picelli et al., 2014; Engel et al., 2016; Patil et al., 2018).

For tumor T_{RM} phenotyping in treatment-naïve patients, samples were analyzed on a FACS-Fusion (BD) following staining with anti-CD45-Alexa Fluor 700 (HI30; BioLegend), anti-CD3-APC-Cy7 (SK7; BioLegend), anti-CD8A-PerCP-Cy5.5 (SK1; BioLegend), anti-CD103-PE-Cy7 (Ber-ACT8; BioLegend), anti-CD127-APC (eBioRDR5; Thermo Fisher Scientific), anti-CD39-BB515 (TU66; BD), anti-4-1BB-PE (4B4-1; BioLegend), anti-PD-1-BV421 (EH12.1; BD), and anti-TIM-3-BV605 (F38-2E2; BioLegend). Cells were counterstained where described with anti-CD19/20-PE-Dazzle (HIB19/2H7; BioLegend), anti-CD14-PE-Dazzle (HCD14; BioLegend), anti-CD56-BV570 (HCD56; BioLegend), and anti-CD4-BV510 (OKT4; BioLegend). Dead cells were discriminated using PI. Phenotypic characterization of lung T_{RM} cells was completed using the antibodies above with anti-CD49A-PE (SR84; BD) and anti-KLRG1-APC (2F1/KLRG1; BioLegend) on an LSRII (BD) using a gating strategy previously described (Ganesan et al., 2017).

For paired analyses of patients before and after anti-PD-1 treatment, we collected tissue from patients with metastatic melanoma (patients 53–54; Tables S1 and S12) before the first dose and after 6 wk of immunotherapy. Patient 53 received ipilimumab at 3 mg/kg and nivolumab 1 mg/kg at three weekly intervals. Patient 54 was treated with pembrolizumab at 2 mg/kg, given once every 3 wk. Both patients achieved a complete remission. Biopsy specimens were cryopreserved in 90% FBS, 10% DMSO until data acquisition. Phenotypic characterization of T_{RM} cell samples before and after immunotherapy was completed by thawing the material and dispersing the biopsies using mechanical and enzymatic digestion, as described previously (Wood et al., 2016; Ganesan et al., 2017). Cells were stained as above and sorted into 2 μ l lysis buffer on a BD Aria-II as described above (Picelli et al., 2014; Engel et al., 2016; Patil et al., 2018). Live, singlet, CD14[−]CD19[−]CD20[−]CD45⁺CD3⁺ or CD14[−]CD19[−]CD20[−]CD45⁺CD3⁺CD8⁺ cells (depending upon the amount of the material available) were then index sorted on a BD Aria-II using the reagents described above. The data were then concatenated in FlowJo (v10.4.1) and, if required, CD8⁺CD4[−]CD103⁺ CTLs were then further isolated in silico based

on protein and/or gene expression data and the analysis completed as described in the RNA-seq section below. Flow cytometry analysis of the remaining donors were completed as above using anti-CD45-Alexa-R700 (HI30; BD), anti-CD3-PE-Dazzle (SK7; BioLegend), anti-CD20-APC-Cy7 (2H7; BioLegend), anti-CD14-APC-Cy7 (HCD14; BioLegend), anti-CD8A-PerCP-Cy5.5 (SK1; BioLegend), anti-CD103-Pe-Cy7 (Ber-ACT8; BioLegend), anti-CD127-APC (eBioRDR5; Thermo Fisher Scientific), anti-PD-1-BV421 (EH12.1; BD), and anti-TIM-3-BV605 (F38-2E2; BioLegend), and live/dead discrimination was completed with fixable/live dead (Fixable Viability Dye eFluor 780; Thermo Fisher Scientific). The cytometric analysis was completed on a Fusion cell sorter (BD). Samples that had <100 CTLs were removed from further analysis.

Flow cytometry-based intracellular protein validation was completed by thawing and washing samples as described above. The samples were incubated for 10 min at 4°C with FcR block as above and then stained using a combination of anti-CD45-Alexa-R700 (HI30; BD), anti-CD3-PE-Dazzle (SK7; BioLegend), anti-CD20-APC-Cy7 (2H7; BioLegend), anti-CD14-APC-Cy7 (HCD14; BioLegend), anti-CD4-BV510 (OKT4; BioLegend), anti-CD8A-PerCP-Cy5.5 (SK1; BioLegend), anti-CD103-Pe-Cy7 (Ber-ACT8; BioLegend), anti-CD127-APC (eBioRDR5; Thermo Fisher Scientific), anti-PD-1-BV421 (EH12.1; BD), and live/dead discrimination was completed with fixable/live dead (Fixable Viability Dye eFluor 780; Thermo Fisher Scientific). TIM-3 staining following ex vivo stimulation and fixation was affected, thus limiting our ability to study the intracellular cytokine profile of TIM-3⁺ cells directly. The sample was then washed and material for ex vivo quantification was immediately fixed (Fixation and Permeabilization Solution; BD) for 20 min at room temperature. The sample was then washed in permeabilization wash (Intracellular Staining Permeabilization Wash Buffer; BioLegend). The sample then received additional FcR blocking reagent and was stained with anti-granzyme B-PE (REA226; Miltenyi Biotec) and anti-Granzyme A-Alexa Fluor 647 (CB9; BioLegend) for 30 min at 4°C. Following this, the material was washed in Permeabilization Wash Buffer. For samples analyzed for ex vivo cytokine production, fixable/live dead was added after 3 h of the stimulation to account for any changes in cell viability during stimulation. For PMA/ionomycin analysis, the cell suspension was stimulated in for 4 h at 37°C in an incubator, at 5% CO₂, in 200 µl complete RPMI with Cell Activation Cocktail (with Brefeldin A; BioLegend) as per the manufacturer's recommendation. Following the addition of further FcR blocking reagent, cytokine staining was completed with anti-IL-2-PE (MQ1-17H12; BioLegend), anti-IFN-γ-BV785 (4S.B3; BioLegend), and anti-TNF-APC (MAb11; BioLegend). Data acquisition was completed on a Fortessa (BD), using a gating strategy previously described (Ganesan et al., 2017) and data were analyzed as above. One sample with <100 total CTLs quantified was removed.

All FACS data were analyzed in FlowJo (v10.4.1), and geometric-mean fluorescence intensity and population percentage data were exported and visualized in GraphPad Prism (v7.0a; Treestar). For t-distributed stochastic neighbor embedding (tSNE) and coexpression analysis of flow cytometry data, each sample was down-sampled to exactly 3,000 randomly

selected live and singlet-gated, CD14⁺CD19⁺CD20⁺CD4⁺CD56⁺CD45⁺CD3⁺CD8⁺ CTLs using the gating strategy described above, and 24,000 cells each from the lung and tumor samples were merged to yield 48,000 total cells. A tSNE plot was constructed using 1,000 permutations and default settings in FlowJo (v10.4.1), z-score expression was mean centered. Flow cytometry data were exported from FlowJo (using the channel values) and these data were imported into R for coexpression analysis (described below).

Bulk RNA-seq and TCR sequencing (TCR-seq)

Total RNA was purified using a miRNAeasy kit (Qiagen) from CD103⁺ and CD103[−] CTLs and quantified as described previously (Engel et al., 2016; Ganesan et al., 2017). For assessment of the stimulated transcriptome, RNA from ~100 sorted cells was used. Total RNA was amplified according to the Smart-seq2 protocol (Picelli et al., 2014). cDNA was purified using AMPure XP beads (0.9:1 ratio; Beckman Coulter). From this step, 1 ng cDNA was used to prepare a standard Nextera XT sequencing library (Nextera XT DNA sample preparation and index kits; Illumina; Ganesan et al., 2017). Samples were sequenced using an Illumina HiSeq2500 to obtain 50-bp single-end reads. For quality control, steps were included to determine total RNA quality and quantity, the optimal number of PCR preamplification cycles, and cDNA fragment size. Samples that failed quality control or had a low number of starting cells were eliminated from further sequencing and analysis. TCR-seq was performed as previously described (Patil et al., 2018) using Tru-seq single indexes (Illumina). Sequencing data were mapped and analyzed using MIGEC (molecular identifier groups-based error correction; Shugay et al., 2014) software (v1.2.7) with default settings, followed by V(D)J tools (v1.1.7) with default settings. Mapping quality control (QC) metrics are included in Tables S1 and S6.

10X single-cell RNA-seq

Samples were processed using 10X v2 chemistry as per the manufacturer's recommendations; 11 and 12 cycles were used for cDNA amplification and library preparation, respectively (Patil et al., 2018). To minimize experimental batch effects, cells were sorted from groups of six donors each day, and cells were pooled for 10X sequencing library preparation. Barcoded RNA was collected and processed following the manufacturer recommendations, as described previously. Libraries were sequenced on a HiSeq2500 and HiSeq4000 (Illumina) to obtain 100- and 32-bp paired-end reads using the following read length: read 1, 26 cycles; read 2, 98 cycles; and i7 index, 8 cycles (Table S1). Samples were pooled together, and DNA samples from whole blood were extracted using a high-salt method (Miller et al., 1988) and quantified using the Qubit 2.0 (Thermo Fisher Scientific). Genotyping was completed through the Infinium Multi-Ethnic Global-8 Kit (Illumina) following the manufacturer's instructions. Raw data from the genotyping analysis were exported using the Genotyping module and Plug-in PLINK Input Report Plug-in (v2.1.4) from GenomeStudio v2.0.4 (Illumina). Data quality was assessed using the snpQC package (Gondro et al., 2014) with R, and low-quality single nucleotide polymorphisms (SNPs) were detected; SNPs failing in >5% of the

samples and SNPs with Illumina's gene call scores <0.2 in $>10\%$ of the samples were flagged. Each subject's sex was matched with the genotype data, and flagged SNPs were removed for downstream analysis using PLINK (v1.90b3w; Purcell et al., 2007). Genetic multiplexing of barcoded single-cell RNA-seq was completed using Demuxlet (Kang et al., 2018) and matched with the Seurat output. Cells with ambiguous or doublet identification were removed from analysis of cluster and/or donor proportions.

Bulk RNA-seq analysis

Bulk RNA-seq data were mapped against the hg19 reference using TopHat (Trapnell et al., 2009; v2.0.9 (--library-type fr-unstranded --no-coverage-search) with FastQC (v0.11.2), Bowtie (v2.1.0.0; Langmead et al., 2009), and Samtools v0.1.19.0; Li and Durbin, 2009), and we employed htseq-count -m union -s no -t exon -i gene_name (part of the HTSeq framework, version v0.7.1; Anders et al., 2015). Trimmomatic (v0.36) was used to remove adapters (Bolger et al., 2014). Values throughout are displayed as \log_2 TPM (transcripts per million) counts; a value of 1 was added before log transformation. To identify genes expressed differentially by various cell types, we performed negative binomial tests for paired comparisons by using the Bioconductor package DESeq2 (Love et al., 2014; v1.14.1), disabling the default options for independent filtering and Cooks cutoff. We considered genes to be expressed differentially by any comparison when the DESeq2 analysis resulted in a Benjamini-Hochberg-adjusted P value of ≤ 0.05 and a fold change ≥ 2 . Union gene signatures were calculated using the online tool jVenn (Bardou et al., 2014), of which genes must have common directionality. GSEA, correlations, and heat maps were generated as previously described (Engel et al., 2016; Ganesan et al., 2017; Patil et al., 2018). Genes used in the GSEA are shown (Table S3). For the preservation of complementary signatures, read count data from Cheuk et al. (2017) were downloaded from Gene Expression Omnibus accession no. GSE83637, and differential expression analysis was completed as above. For the murine composite signature (Milner et al., 2017), orthologues between human and murine signatures were compared using Ensembl to facilitate converting from the murine to human signature; genes that had opposing expression changes were not considered conserved (Table S3). Reactome pathways were generated using the online tool (<https://reactome.org/>) for tumor T_{RM} -specific genes; a pathway was considered significantly different if the false discovery rate (FDR; q) value was ≤ 0.05 ; Table S4). Visualizations were generated in ggplot2 using custom scripts, while expression values were calculated using GraphPad Prism7 (v7.0a). For tSNE analysis, the data frame was filtered to genes with mean ≥ 1 TPM counts expression in at least one condition and visualizations created using the top 2,000 most variable genes, as calculated in DESeq2 (Love et al., 2014; v1.16.1); this allowed for unbiased visualization of the \log_2 (TPM counts + 1) data using package Rtsne (v0.13). Coexpression networks were generated in gplots (v3.0.1) using the heatmap2 function, while weighted correlation analysis was completed using WGCNA (Langfelder and Horvath, 2008; v1.61) from the \log_2 (TPM counts + 1) data matrix and the function

TOMsimilarityfromExpr (Beta = 5) and exportNetworkToCytoscape, weighted = true, threshold = 0.05. Highlighted genes were ordered as per the order in the correlation plot. Networks were generated in Gephi (v0.92; Mellone et al., 2016; Ottensmeier et al., 2016) using Fruchterman Reingold and Noverlap functions. The size and color were scaled according to the Average Degree as calculated in Gephi, while the edge width was scaled according to the WGCNA edge weight value. The statistical analysis of the overlap between gene sets was calculated in R (v3.5.0) using the fisher.test function (Stats v3.5.0) using the number of total quantified genes used for DESeq2, as the total value (20,231), with alternative="greater".

Single-cell RNA-seq analysis

Raw 10X data (Table S1) were processed as previously described (Patil et al., 2018), merging multiple sequencing runs using cellranger count function in cell ranger and then merging multiple cell types with cell ranger aggr (v2.0.2). The merged data were transferred to the R statistical environment for analysis using the package Seurat (v2.2.1; Macosko et al., 2015; Patil et al., 2018). Only cells expressing >200 genes and genes expressed in ≥ 3 cells were included in the analysis. The data were then log-normalized and scaled per cell and variable genes were detected. Transcriptomic data from each cell was then further normalized by the number of UMI-detected and mitochondrial genes. A PC analysis (PCA) was then run on variable genes, and the first eight PCs were selected for further analyses based on the standard deviation of PCs, as determined by an "elbow plot" in Seurat. Cells were clustered using the FindClusters function in Seurat with default settings, resolution = 0.6 and eight PCs. Differential expression between clusters was determined by converting the data to counts per million and analyzing cluster-specific differences using MAST (q ≤ 0.01 , v1.2.1; Finak et al., 2015; Patil et al., 2018; Soneson and Robinson, 2018). A gene was considered significantly different only if the gene was commonly positively enriched in every comparison for a singular cluster (Engel et al., 2016; Patil et al., 2018). Further visualizations of exported normalized data were generated using the Seurat package and custom R scripts. Cell-state hierarchy maps were generated using Monocle (v2.6.1; Trapnell et al., 2014) and default settings with expressionFamily = negbinomial.size(), lowerDetectionLimit = 1 and mean_expression ≥ 0.1 , including the most variable genes identified in Seurat for consistency. Average expression across a cell cluster was calculated using the AverageExpression function, and down-sampling was achieved using the SubsetData function (both in Seurat). Distance between clusters was calculated by calculating a particular cells location in PCA space (PC 1:3) using the function GetCellEmbeddings (in Seurat), and the values for each cell were then scaled per column (Scale function, core R) where described; finally, a distance matrix was calculated (dist function, core R, method = Euclidean). This matrix was filtered to the cells assigned to cluster 1, and the mean distance of each cell in cluster 1 to all cells in each of the remaining T_{RM} clusters (2, 3, 4, and 5) was calculated. The clustering analysis was completed using the hclust function in R (Stats, R v3.5.0) with average linkage and generated from the Spearman correlation analysis of each cell's

location in PCA space (as above). SAVER coexpression analysis (Huang et al., 2018) was completed on the raw-UMI counts of the T_{RM} cells (clusters 1–5) and the non- T_{RM} cells (remaining cells) using the function `saver` (v1.1.1) with `pred.genes.only = TRUE`, `estimates.only = FALSE` on transcripts assigned as uniquely enriched in cluster 2, removing genes not expressed in any cells in the non- T_{RM} compartment. Correlation values were isolated using the `cor.genes` function in SAVER and coexpression plots generated as described above.

Smart-seq2 single-cell analysis (Table S1) was completed as previously described using TraCeR (v0.5.1; Stubbington et al., 2016; Patil et al., 2018) and custom scripts to identify $\alpha\beta$ chains, showing only cells where both TCR chains were detected and to remove cells with low QC values as previously described (Patil et al., 2018). Here, we removed cells with <200,000 reads, and <30% of sequenced bases were assigned to untranslated regions and coding regions of mRNA. Samples were mapped as described previously (Patil et al., 2018), and the data were log transformed and displayed as normalized TPM counts; a value of 1 was added before log transformation. Visualizations were completed in ggplot2, Prism (v7.0a) and custom scripts in TraCeR. A cell was considered expanded when both the most highly expressed α and β TCR chain sequences matched other cells with the same stringent criteria. Cells were considered not expanded when neither α nor β TCR productive chain sequences matched those of any other cells. A cell was considered $PD-1^+$ or $TIM-3^+$ when the expression of *PDCDI* or *HAVCR2* was >10 TPM counts, while a cell was considered cycling if expression of cell cycle genes *TOP2A* and/or *MKI67* was >10 TPM counts. Differential expression profiling was completed using MAST (Finak et al., 2015; $q \leq 0.05$), as previously described (Patil et al., 2018).

Matched flow cytometry data were analyzed using FlowJo (v10.4.1), values and gates were exported into ggplot and “in-silico gates” were applied using custom scripts in R. Given ~85% of the $CD103^+$ cells were $TIM-3^+$ from our flow cytometry data, cells were broadly classified into T_{RM} or non- T_{RM} based on an individual cell’s protein expression (FACS gating) for patient 53. Where there was no available cell-specific associated protein data (patient 54), $CD3^+$ T cells were classified based on the lack of expression of *CD4* and *FOXP3* to enable removal of $CD4^+$ cells. Next, we stratified the single-cell transcriptomes into T_{RM} or non- T_{RM} cells when expression of T_{RM} -associated genes, *ITGAE* (*CD103*), *RBPJ*, and/or *ZNF683* (*HOBIT*) was >10 TPM counts; the classification of each single-cell library is summarized in Table S12. Differential gene expression analysis was completed as above.

Multiplex immunohistochemistry

Patients included in this cohort had a known diagnosis of lung cancer. 23 patients were selected in total, categorizing the donors using criteria previously reported (Ganesan et al., 2017). A multiplexed IHC method was used for repeated staining of a single paraffin-embedded tissue slide. Deparaffinization, rehydration, antigen retrieval, and IHC staining were performed using a Dako PT Link Autostainer. Antigen retrieval was performed using the EnVision FLEX Target Retrieval Solution, High pH (Agilent Dako) for all antibodies. The slide was first stained

with a standard primary antibody followed by an appropriate biotin-linked secondary antibody and HRP-conjugated streptavidin to amplify the signal. Peroxidase-labeled compounds were revealed using 3-amino-9-ethylcarbazole, an aqueous substrate that results in red staining, or 3,3'-diaminobenzidine (DAB) that results in brown staining and counterstained using hematoxylin (blue).

The slides were stained initially with Cytokeratin (prediluted, Clone AE1/AE3; Agilent Dako) and then sequentially with anti- $CD8\alpha$ (prediluted Kit IR62361-2; clone C8/144B; Agilent Dako), anti- $CD103$ (1:500; EPR4166(2); Abcam), and anti- $TIM-3$ (1:50; D5D5R; Cell Signaling Technology). The slides were scanned at high resolution using a Zeiss Axio Scan.Z1 with a 20 \times air immersion objective. Between each staining iteration, antigen retrieval was performed along with removal of the labile 3-amino-9-ethylcarbazole staining and denaturation of the preceding antibodies using a set of organic solvent-based destaining buffers as follows: 50% ethanol for 2 min, 100% ethanol for 2 min, 100% xylene for 2 min, 100% ethanol for 2 min, and 50% ethanol for 2 min. This process did not affect DAB staining. The process was repeated for each of the antibodies.

Bright-field images were separated into color channels in imaging processing software ImageJ FIJI (ImageJ Windows 64-bit final version; Schindelin et al., 2012). For the $TIL^{high}T_{RM}^{high}$ and $TIL^{low}T_{RM}^{low}$ tumors, the number of $CD8^+CD103^+TIM3^+$ cells were quantified manually. Two samples with ≤ 3 $CD8^+CD103^+$ CTLs quantified were removed to prevent calculating percentages of single events, resulting in a final number of 21 samples. These images were processed and combined to create pseudo-color multiplexed images. The raw counts for each protein, individually and together, are presented in Table S11 as the number of cells per 0.15 mm².

OMNI-ATAC-seq

CTLs were FACS sorted from cryopreserved lung cancer samples as described above using the following antibody cocktail: anti- $CD45$ -Alexa Fluor 700 (HI30; BioLegend), anti- $CD3$ -APC-Cy7 (SK7; BioLegend), anti- $CD8A$ -PerCP-Cy5.5 (SK1; BioLegend), anti- $CD103$ -PE-Cy7 (Ber-ACT8; BioLegend), anti- $CD127$ -APC (eBioRDR5; Thermo Fisher Scientific), and anti- $TIM-3$ -BV605 (F38-2E2; BioLegend). Cells were counter stained with anti- $CD19/20$ -PE-Dazzle (HIB19/2H7; BioLegend), anti- $CD14$ -PE-Dazzle (HCD14; BioLegend), and anti- $CD4$ -BV510 (OKT4; BioLegend). Dead cells were discriminated using PI. Samples were sorted into low-retention 1.5-ml Eppendorfs containing 250 μ l FBS and 250 μ l PBS. Three to six donors were pooled together to guarantee sufficient cell numbers. For each pool of cells, two or three technical replicates of 15,000–25,000 CTLs were generated for each library.

OMNI-ATAC-seq was performed as described in Corces et al. (2017), with minor modifications. Isolated nuclei were incubated with tagmentation mix (2 \times TD buffer, 2.5 μ l transposase enzyme from Nextera kit; Illumina) at 37°C for 30 min in a thermomixer, shaking at 1,000 RPM (Corces et al., 2017). Following tagmentation, the product was eluted in 0.1X Tris-EDTA buffer using DNA Clean and Concentrator-5 kit (Zymo). The purified product was preamplified for five cycles using Kappa 2X enzyme along

with Nextera indexes (Illumina), and based on quantitative PCR amplification, an additional seven cycles of amplification were performed for 20,000 cells. The PCR-amplified product was purified using DNA Clean and Concentrator-5 kit (Zymo), and size selection was done using AMPure XP beads (Beckman Coulter). Finally, concentration and quality of libraries were determined by picogreen and bioAnalyzer assays. Equimolar libraries were sequenced as above or on a NovaSeq 6000 for sequencing.

Next, technical replicates were randomly down sampled to between 25,000,000 and 40,000,000 total reads (Table S1) and merged using Bash scripts, resulting in two T_{RM} pools and two non- T_{RM} pools. These reads were mapped to hg19 with bowtie2 (v2.3.3.1). Chromosomes 1–22 and X were retained, chrY, chrM, and other arbitrary chromosome information based reads were removed. Samtools (v1.9) was used to get the uniquely mappable reads, and only reads with mapping quality score ≥ 30 were considered. Duplicate reads are removed using the “MarkDuplicates” utility of Picard tool (v 2.18.14). Before peak calling, tag align files were created by shifting forward strands by four bases and reverse strands by five bases (TN5 shift). Peaks were identified with MACS2 (v 2.1.1.20160309) using the function `-f BED -g 'hs' -q 0.01 --nomodel --nolambda --keep-dup all --shift -100 --extsize 200`. BamCoverage (v2.4.2) was used for converting bam files into bigwig and further UCSC track generation (same normalization across all ATAC-seq and RNA-seq samples), as per the following example: `bamCoverage -b TIL_103pos.bam -o TIL_103pos_NormCov.bw -of bigwig -bs 10 --normalizeTo1x 2864785220 --normalizeUsingRPKM -e 200`. The R package DiffBind (v2.2.12) was used to highlight differentially accessible peaks (based on DEseq2). R packages of org.Hs.eg.db (v3.4.0) and TxDb.Hsapiens.UCSC.hg19.knownGene (v3.2.2) were used to annotate peaks. Following differential expression, peaks were filtered to those within 5 kb of a transcription start site to focus directly on promoter accessibility. The correlation plot (Spearman) was completed as described above using all identified peaks. The plot was clustered according to complete linkage.

Accession codes

Sequencing data have been uploaded onto the Gene Expression Omnibus with the accession no. GSE111898.

Other statistical analysis

The significance of differences among matched samples was determined by a Wilcoxon rank-sum test for paired data or a Mann–Whitney *U* test for nonpaired data, unless otherwise stated. Statistical analyses were performed using GraphPad Prism7 (v7.0a). The Spearman correlation coefficient (*r* value) was used to access the significance of correlations between the levels of any two components of interest.

Online supplemental material

Fig. S1 includes additional analysis on lung-resident T_{RM} cells. Fig. S2 provides additional information regarding the breakdown per donor in each cluster and expression of *TCF7*. Fig. S3 demonstrates additional transcripts identified as coexpressed in

a particular cluster of tumor T_{RM} cells. Fig. S4 presents additional data regarding T_{RM} cells in the context of anti-PD-1 therapy. Table S1 provides matched clinical data for the patient samples used in this study. Table S2 includes the output and additional analysis of the differentially expressed genes in the lung T_{RM} cell population. Table S3 presents the genes used for GSEA. Table S4 provides further information regarding shared tissue-residency and tumor T_{RM} -enriched transcripts. Table S5 includes additional information regarding the transcriptome of T_{RM} cells following ex vivo stimulation. Table S6 comprises output of TCR-seq analysis. Table S7 covers the single-cell transcriptomic analysis of clusters of tumor T_{RM} cells. Table S8 contains the output of single cell TCR-seq analysis. Table S9 incorporates single-cell transcriptomic analysis of $TIM-3^+$ T_{RM} cells. Table S10 presents transcriptomic and protein coexpression analysis of tumor T_{RM} and non- T_{RM} cells. Table S11 provides raw data associated with the immunohistochemistry analysis. Table S12 contains single-cell transcriptomic and TCR-seq analysis of T_{RM} and non- T_{RM} cells before and after anti-PD-1 therapy.

Acknowledgments

We thank M. Chamberlain, K. Amer, D. Jeffrey, M. Lane, C. Fixmer, M. Lopez, N. Graham, M. Machado, T. Mellows, and B. Johnson for assistance with recruitment of study subjects and processing of samples. We thank D. Chudakov for sharing the detailed protocol for TCR-seq. We recognize C. Kim, L. Bogge-man, D. Hinz, C. Dillingham, and R. Simmons for their assistance with cell sorting using the FACSaria II Cell Sorter S10 RR027366; S. Liang, S. Lisette Rosales, and J. Greenbaum for assistance with library preparation, next-generation sequencing, and bioinformatics at the La Jolla Institute using the Illumina HiSeq 2500, and J. Moore, B. Schmiedel, and other members of the laboratory for their assistance with editing the figures and manuscript.

This study was supported by the Wessex Clinical Research Network and the National Institute for Health Research, UK (sample collection), the William K. Bowes Jr Foundation (P. Vijayanand), the Cancer Research UK Centres Network Accelerator Award grant A21998 (P. Vijayanand, T. Sanchez-Elsner, and C.H. Ottensmeier), the Faculty of Medicine of the University of Southampton (P. Vijayanand, T. Sanchez-Elsner, and C.H. Ottensmeier) and Cancer Research UK (J. Clarke, E.V. King, and C.H. Ottensmeier). Sequencing instrumentation was supported by the National Institutes of Health grant S10OD016262 and the NovaSeq6000 grant S10OD025052-01.

J. Clarke, S. Eschweiler, F. Ay, C.H. Ottensmeier, T. Sanchez-Elsner, and P. Vijayanand have filed a patent associated with the results presented here but have no additional competing financial interests. P. Vijayanand receives research funding from Pfizer for following up some of the targets identified in this study. The remaining authors declare no competing financial interests.

Author contributions: J. Clarke, A.-P. Ganesan, P.S. Friedmann, T. Sanchez-Elsner, C.H. Ottensmeier, and P. Vijayanand conceived of the work; J. Clarke, O. Wood, S. Eschweiler,

S.J. Chee, A. Alzetani, C.J. Hanley, and K.J. McCann performed cell isolations, phenotyping, and immunohistochemistry data analysis under the supervision of G.J. Thomas, P. Vijayanand, and C.H. Ottensmeier; S.J. Chee, A. Alzetani, K.J. McCann, E. Woo, and E.V. King assisted in patient recruitment, obtaining consent and sample collection; J. Clarke, R. Gujar, S. Bhattacharyya, and B. Panwar performed bulk RNA-seq and OMNI-ATAC-seq assays and analysis under the supervision of F. Ay and P. Vijayanand; J. Clarke, D. Singh, and A. Madrigal performed the single-cell experiments and analysis under the supervision of F. Ay, G. Seumois, and P. Vijayanand; J. Clarke wrote the first draft of the manuscript that was revised and edited by P.S. Friedmann, A.-P. Ganesan, F. Ay, C.H. Ottensmeier, and P. Vijayanand.

Submitted: 7 February 2019

Revised: 4 April 2019

Accepted: 23 May 2019

References

- Anders, S., P.T. Pyl, and W. Huber. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 31:166–169. <https://doi.org/10.1093/bioinformatics/btu638>
- Aranda, J.F., N. Reglero-Real, L. Kremer, B. Marcos-Ramiro, A. Ruiz-Sáenz, M. Calvo, C. Enrich, I. Correias, J. Millán, and M.A. Alonso. 2011. MYADM regulates Rac1 targeting to ordered membranes required for cell spreading and migration. *Mol. Biol. Cell*. 22:1252–1262. <https://doi.org/10.1091/mbc.e10-11-0910>
- Bacher, P., F. Heinrich, U. Stervbo, M. Nienen, M. Vahldieck, C. Iwert, K. Vogt, J. Kollet, N. Babel, B. Sawitzki, et al. 2016. Regulatory T Cell Specificity Directs Tolerance versus Allergy against Aeroantigens in Humans. *Cell*. 167:1067–1078.e16. <https://doi.org/10.1016/j.cell.2016.09.050>
- Bacon, C., V. Endris, and G.A. Rappold. 2013. The cellular function of srGAP3 and its role in neuronal morphogenesis. *Mech. Dev.* 130:391–395. <https://doi.org/10.1016/j.mod.2012.10.005>
- Bardou, P., J. Mariette, F. Escudié, C. Djemiel, and C. Klopp. 2014. jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics*. 15:293. <https://doi.org/10.1186/1471-2105-15-293>
- Bindea, G., B. Mlecnik, M. Tosolini, A. Kirilovsky, M. Waldner, A.C. Obenauf, H. Angell, T. Fredriksen, L. Lafontaine, A. Berger, et al. 2013. Spatio-temporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*. 39:782–795. <https://doi.org/10.1016/j.immuni.2013.10.003>
- Bolger, A.M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Castellino, F., A.Y. Huang, G. Altan-Bonnet, S. Stoll, C. Scheinecker, and R.N. Germain. 2006. Chemokines enhance immunity by guiding naive CD8+ T cells to sites of CD4+ T cell-dendritic cell interaction. *Nature*. 440: 890–895. <https://doi.org/10.1038/nature04651>
- Chan, C.J., L. Martinet, S. Gilfillan, F. Souza-Fonseca-Guimaraes, M.T. Chow, L. Town, D.S. Ritchie, M. Colonna, D.M. Andrews, and M.J. Smyth. 2014. The receptors CD96 and CD226 oppose each other in the regulation of natural killer cell functions. *Nat. Immunol.* 15:431–438. <https://doi.org/10.1038/ni.2850>
- Chee, S.J., M. Lopez, T. Mellows, S. Gankande, K.A. Moutasim, S. Harris, J. Clarke, P. Vijayanand, G.J. Thomas, and C.H. Ottensmeier. 2017. Evaluating the effect of immune cells on the outcome of patients with mesothelioma. *Br. J. Cancer*. 117:1341–1348. <https://doi.org/10.1038/bjc.2017.269>
- Cheuk, S., H. Schlums, I. Gallais Sérézal, E. Martini, S.C. Chiang, N. Marquardt, A. Gibbs, E. Detlofsson, A. Introini, M. Forkel, et al. 2017. CD49a Expression Defines Tissue-Resident CD8+ T Cells Poised for Cytotoxic Function in Human Skin. *Immunity*. 46:287–300. <https://doi.org/10.1016/j.immuni.2017.01.009>
- Collins, S., A. Waickman, A. Basson, A. Kupfer, J.D. Licht, M.R. Horton, and J.D. Powell. 2012. Regulation of CD4+ and CD8+ effector responses by Sprouty-1. *PLoS One*. 7:e49801. <https://doi.org/10.1371/journal.pone.0049801>
- Corces, M.R., A.E. Trevino, E.G. Hamilton, P.G. Greenside, N.A. Sinnott-Armstrong, S. Vesuna, A.T. Satpathy, A.J. Rubin, K.S. Montine, B. Wu, et al. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods*. 14:959–962. <https://doi.org/10.1038/nmeth.4396>
- Croft, M., T. So, W. Duan, and P. Soroosh. 2009. The significance of OX40 and OX40L to T-cell biology and immune disease. *Immunol. Rev.* 229:173–191. <https://doi.org/10.1111/j.1600-065X.2009.00766.x>
- Cui, W., Y. Liu, J.S. Weinstein, J. Craft, and S.M. Kaech. 2011. An interleukin-21-interleukin-10-STAT3 pathway is critical for functional maturation of memory CD8+ T cells. *Immunity*. 35:792–805. <https://doi.org/10.1016/j.immuni.2011.09.017>
- Djenidi, F., J. Adam, A. Goubar, A. Durgeau, G. Meurice, V. de Montpréville, P. Validire, B. Besse, and F. Mami-Chouaib. 2015. CD8+CD103+ tumor-infiltrating lymphocytes are tumor-specific tissue-resident memory T cells and a prognostic factor for survival in lung cancer patients. *J. Immunol.* 194:3475–3486. <https://doi.org/10.4049/jimmunol.1402711>
- Dominguez, C.X., R.A. Amezcua, T. Guan, H.D. Marshall, N.S. Joshi, S.H. Kleinstein, and S.M. Kaech. 2015. The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. *J. Exp. Med.* 212:2041–2056. <https://doi.org/10.1084/jem.20150186>
- Duhen, T., R. Duhen, R. Montler, J. Moses, T. Moudgil, N.F. de Miranda, C.P. Goodall, T.C. Blair, B.A. Fox, J.E. McDermott, et al. 2018. Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat. Commun.* 9:2724. <https://doi.org/10.1038/s41467-018-05072-0>
- Emgård, J., H. Kammoun, B. García-Cassani, J. Chesné, S.M. Parigi, J.-M. Jacob, H.-W. Cheng, E. Evren, S. Das, P. Czarnewski, et al. 2018. Oxysterol Sensing through the Receptor GPR183 Promotes the Lymphoid-Tissue-Inducing Function of Innate Lymphoid Cells and Colonic Inflammation. *Immunity*. 48:120–132.e8. <https://doi.org/10.1016/j.immuni.2017.11.020>
- Engel, I., G. Seumois, L. Chavez, D. Samaniego-Castruita, B. White, A. Chawla, D. Mock, P. Vijayanand, and M. Kronenberg. 2016. Innate-like functions of natural killer T cell subsets result from highly divergent gene programs. *Nat. Immunol.* 17:728–739. <https://doi.org/10.1038/ni.3437>
- Engler, J.B., N. Kursawe, M.E. Solano, K. Patas, S. Wehrmann, N. Heckmann, F. Lühder, H.M. Reichardt, P.C. Arck, S.M. Gold, and M.A. Friese. 2017. Glucocorticoid receptor in T cells mediates protection from autoimmunity in pregnancy. *Proc. Natl. Acad. Sci. USA*. 114:E181–E190. <https://doi.org/10.1073/pnas.1617151114>
- Finak, G., A. McDavid, M. Yajima, J. Deng, V. Gersuk, A.K. Shalek, C.K. Slichter, H.W. Miller, M.J. McElrath, M. Prlic, et al. 2015. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* 16:278. <https://doi.org/10.1186/s13059-015-0844-5>
- Franciszewicz, K., A. Le Floch, M. Boutet, I. Vergnon, A. Schmitt, and F. Mami-Chouaib. 2013. CD103 or LFA-1 engagement at the immune synapse between cytotoxic T cells and tumor cells promotes maturation and regulates T-cell effector functions. *Cancer Res.* 73:617–628. <https://doi.org/10.1158/0008-5472.CAN-12-2569>
- Galon, J., A. Costes, F. Sanchez-Cabo, A. Kirilovsky, B. Mlecnik, C. Lagorce-Pagès, M. Tosolini, M. Camus, A. Berger, P. Wind, et al. 2006. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*. 313:1960–1964. <https://doi.org/10.1126/science.1129139>
- Ganesan, A.-P., J. Clarke, O. Wood, E.M. Garrido-Martin, S.J. Chee, T. Mellows, D. Samaniego-Castruita, D. Singh, G. Seumois, A. Alzetani, et al. 2017. Tissue-resident memory features are linked to the magnitude of cytotoxic T cell responses in human lung cancer. *Nat. Immunol.* 18: 940–950. <https://doi.org/10.1038/ni.3775>
- Gondro, C., L.R. Porto-Neto, and S.H. Lee. 2014. SNPQC—an R pipeline for quality control of Illumina SNP genotyping array data. *Anim. Genet.* 45: 758–761. <https://doi.org/10.1111/age.12198>
- Guo, X., Y. Zhang, L. Zheng, C. Zheng, J. Song, Q. Zhang, B. Kang, Z. Liu, L. Jin, R. Xing, et al. 2018. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat. Med.* 24:978–985. <https://doi.org/10.1038/s41591-018-0045-3>
- Hombrink, P., C. Helbig, R.A. Backer, B. Piet, A.E. Oja, R. Stark, G. Brasser, A. Jongejan, R.E. Jonkers, B. Nota, et al. 2016. Programs for the persistence, vigilance and control of human CD8+ lung-resident memory T cells. *Nat. Immunol.* 17:1467–1478. <https://doi.org/10.1038/ni.3589>

- Huang, A.C., M.A. Postow, R.J. Orlowski, R. Mick, B. Bengsch, S. Manne, W. Xu, S. Harmon, J.R. Giles, B. Wenz, et al. 2017. T-cell invigoration to tumour burden ratio associated with anti-PD-1 response. *Nature*. 545: 60–65. <https://doi.org/10.1038/nature22079>
- Huang, C.-Y., Y.-C. Lin, W.-Y. Hsiao, F.-H. Liao, P.-Y. Huang, and T.-H. Tan. 2012. DUSP4 deficiency enhances CD25 expression and CD4⁺ T-cell proliferation without impeding T-cell development. *Eur. J. Immunol.* 42:476–488. <https://doi.org/10.1002/eji.201041295>
- Huang, J., L.C. Bridges, and J.M. White. 2005. Selective modulation of integrin-mediated cell migration by distinct ADAM family members. *Mol. Biol. Cell*. 16:4982–4991. <https://doi.org/10.1091/mbc.e05-03-0258>
- Huang, M., J. Wang, E. Torre, H. Dueck, S. Shaffer, R. Bonasio, J.I. Murray, A. Raj, M. Li, and N.R. Zhang. 2018. SAVER: gene expression recovery for single-cell RNA sequencing. *Nat. Methods*. 15:539–542. <https://doi.org/10.1038/s41592-018-0033-z>
- Im, S.J., M. Hashimoto, M.Y. Gerner, J. Lee, H.T. Kissick, M.C. Burger, Q. Shan, J.S. Hale, J. Lee, T.H. Nasti, et al. 2016. Defining CD8⁺ T cells that provide the proliferative burst after PD-1 therapy. *Nature*. 537:417–421. <https://doi.org/10.1038/nature19330>
- Janakiram, M., J.M. Chinai, A. Zhao, J.A. Sparano, and X. Zang. 2015. HHLA2 and TMIGD2: new immunotherapeutic targets of the B7 and CD28 families. *Oncotarget*. 4:e1026534. <https://doi.org/10.1080/2162402X.2015.1026534>
- June, C.H., J.T. Warshawer, and J.A. Bluestone. 2017. Is autoimmunity the Achilles' heel of cancer immunotherapy? *Nat. Med.* 23:540–547. <https://doi.org/10.1038/nm.4321>
- Kang, H.M., M. Subramaniam, S. Targ, M. Nguyen, L. Maliskova, E. McCarthy, E. Wan, S. Wong, L. Byrnes, C.M. Lanata, et al. 2018. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* 36:89–94. <https://doi.org/10.1038/nbt.4042>
- Kim, S.V., W.V. Xiang, C. Kwak, Y. Yang, X.W. Lin, M. Ota, U. Sarpel, D.B. Rifkin, R. Xu, and D.R. Littman. 2013. GPR15-mediated homing controls immune homeostasis in the large intestine mucosa. *Science*. 340: 1456–1459. <https://doi.org/10.1126/science.1237013>
- Langfelder, P., and S. Horvath. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 9:559. <https://doi.org/10.1186/1471-2105-9-559>
- Langmead, B., C. Trapnell, M. Pop, and S.L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. <https://doi.org/10.1186/gb-2009-10-3-r25>
- Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, H., A.M. van der Leun, I. Yofe, Y. Lubling, D. Gelbard-Solodkin, A.C.J. van Akkooi, M. van den Braber, E.A. Rozeman, J.B.A.G. Haanen, C.U. Blank, et al. 2019. Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell*. 176: 775–789.e18. <https://doi.org/10.1016/j.cell.2018.11.043>
- Love, M.I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15: 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Mackay, L.K., A. Rahimpour, J.Z. Ma, N. Collins, A.T. Stock, M.-L. Hafon, J. Vega-Ramos, P. Lauzurica, S.N. Mueller, T. Stefanovic, et al. 2013. The developmental pathway for CD103(+)CD8⁺ tissue-resident memory T cells of skin. *Nat. Immunol.* 14:1294–1301. <https://doi.org/10.1038/ni.2744>
- Mackay, L.K., M. Minnich, N.A.M. Kragten, Y. Liao, B. Nota, C. Seillet, A. Zaid, K. Man, S. Preston, D. Freestone, et al. 2016. Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science*. 352:459–463. <https://doi.org/10.1126/science.1252035>
- Macosko, E.Z., A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A.R. Bialas, N. Kamitaki, E.M. Martersteck, et al. 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*. 161:1202–1214. <https://doi.org/10.1016/j.cell.2015.05.002>
- Malik, B.T., K.T. Byrne, J.L. Vella, P. Zhang, T.B. Shabaneh, S.M. Steinberg, A.K. Molodtsov, J.S. Bowers, C.V. Angeles, C.M. Paulos, et al. 2017. Resident memory T cells in the skin mediate durable immunity to melanoma. *Sci. Immunol.* 2:6346. <https://doi.org/10.1126/sciimmunol.aam6346>
- Mellone, M., C.J. Hanley, S. Thirdborough, T. Mellows, E. Garcia, J. Woo, J. Tod, S. Frampton, V. Jenei, K.A. Moutasim, et al. 2016. Induction of fibroblast senescence generates a non-fibrogenic myofibroblast phenotype that differentially impacts on cancer prognosis. *Aging (Albany N.Y.)*. 9:114–132. <https://doi.org/10.18632/aging.101127>
- Miller, S.A., D.D. Dykes, and H.F. Polesky. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16:1215. <https://doi.org/10.1093/nar/16.3.1215>
- Milner, J.J., C. Toma, B. Yu, K. Zhang, K. Omilusik, A.T. Phan, D. Wang, A.J. Getzler, T. Nguyen, S. Crotty, et al. 2017. Runx3 programs CD8⁺ T cell residency in non-lymphoid tissues and tumours. *Nature*. 552:253–257. <https://doi.org/10.1038/nature24993>
- Muthusamy, N., K. Barton, and J.M. Leiden. 1995. Defective activation and survival of T cells lacking the Ets-1 transcription factor. *Nature*. 377: 639–642. <https://doi.org/10.1038/377639a0>
- Namiki, T., A. Tanemura, J.C. Valencia, S.G. Coelho, T. Passeron, M. Kawaguchi, W.D. Vieira, M. Ishikawa, W. Nishijima, T. Izumo, et al. 2011. AMP kinase-related kinase NAK2 affects tumor growth, migration, and clinical outcome of human melanoma. *Proc. Natl. Acad. Sci. USA*. 108:6597–6602. <https://doi.org/10.1073/pnas.1007694108>
- Nieminen, M., T. Henttinen, M. Merinen, F. Marttila-Ichihara, J.E. Eriksson, and S. Jalkanen. 2006. Vimentin function in lymphocyte adhesion and transcellular migration. *Nat. Cell Biol.* 8:156–162. <https://doi.org/10.1038/ncb1355>
- Nizard, M., H. Roussel, M.O. Diniz, S. Karaki, T. Tran, T. Voron, E. Dransart, F. Sandoval, M. Riquet, B. Rance, et al. 2017. Induction of resident memory T cells enhances the efficacy of cancer vaccine. *Nat. Commun.* 8:15221. <https://doi.org/10.1038/ncomms15221>
- Ottensmeier, C.H., K.L. Perry, E.L. Harden, J. Stasakova, V. Jenei, J. Fleming, O. Wood, J. Woo, C.H. Woelk, G.J. Thomas, and S.M. Thirdborough. 2016. Upregulated Glucose Metabolism Correlates Inversely with CD8⁺ T-cell Infiltration and Survival in Squamous Cell Carcinoma. *Cancer Res.* 76:4136–4148. <https://doi.org/10.1158/0008-5472.CAN-15-3121>
- Pallett, L.J., J. Davies, E.J. Colbeck, F. Robertson, N. Hansi, N.J.W. Easom, A.R. Burton, K.A. Stegmann, A. Schurich, L. Swadling, et al. 2017. IL-2^{high} tissue-resident T cells in the human liver: Sentinels for hepatotropic infection. *J. Exp. Med.* 214:1567–1580. <https://doi.org/10.1084/jem.20162115>
- Pardoll, D.M. 2012. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer*. 12:252–264. <https://doi.org/10.1038/nrc3239>
- Parikh, S.S., S.A. Litherland, M.J. Clare-Salzler, W. Li, P.A. Gulig, and F.S. Southwick. 2003. CapG(-/-) mice have specific host defense defects that render them more susceptible than CapG(+/+) mice to *Listeria monocytogenes* infection but not to *Salmonella enterica* serovar Typhimurium infection. *Infect. Immun.* 71:6582–6590. <https://doi.org/10.1128/IAI.71.11.6582-6590.2003>
- Patil, V.S., A. Madrigal, B.J. Schmiedel, J. Clarke, P. O'Rourke, A.D. de Silva, E. Harris, B. Peters, G. Seumois, D. Weiskopf, et al. 2018. Precursors of human CD4⁺ cytotoxic T lymphocytes identified by single-cell transcriptome analysis. *Sci. Immunol.* 3:8664. <https://doi.org/10.1126/sciimmunol.aan8664>
- Picelli, S., O.R. Faridani, Å.K. Björklund, G. Winberg, S. Sagasser, and R. Sandberg. 2014. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* 9:171–181. <https://doi.org/10.1038/nprot.2014.006>
- Prasad, S., S. Hu, W.S. Sheng, P. Chauhan, A. Singh, and J.R. Lokensgard. 2017. The PD-1: PD-L1 pathway promotes development of brain-resident memory T cells following acute viral encephalitis. *J. Neuroinflammation*. 14:82. <https://doi.org/10.1186/s12974-017-0860-3>
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J. Maller, P. Sklar, P.I.W. de Bakker, M.J. Daly, and P.C. Sham. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81:559–575. <https://doi.org/10.1086/519795>
- Rapaport, A.S., J. Schriewer, S. Gilfillan, E. Hembrador, R. Crump, B.F. Plougastel, Y. Wang, G. Le Fric, J. Gao, M. Cella, et al. 2015. The Inhibitory Receptor NKG2A Sustains Virus-Specific CD8⁺ T Cells in Response to a Lethal Poxvirus Infection. *Immunity*. 43:1112–1124. <https://doi.org/10.1016/j.immuni.2015.11.005>
- Robert, C., J. Schachter, G.V. Long, A. Arance, J.J. Grob, L. Mortier, A. Daud, M.S. Carlino, C. McNeil, M. Lotem, et al. KEYNOTE-006 investigators. 2015. Pembrolizumab versus Ipilimumab in Advanced Melanoma. *N. Engl. J. Med.* 372:2521–2532. <https://doi.org/10.1056/NEJMoa1503093>
- Savas, P., B. Virassamy, C. Ye, A. Salim, C.P. Mintoff, F. Caramia, R. Salgado, D.J. Byrne, Z.L. Teo, S. Dushyanthen, et al. Kathleen Cuninghame Foundation Consortium for Research into Familial Breast Cancer (kConFab). 2018. Single-cell profiling of breast cancer T cells reveals a

- tissue-resident memory subset associated with improved prognosis. *Nat. Med.* 24:986–993. <https://doi.org/10.1038/s41591-018-0078-7>
- Schenkel, J.M., and D. Masopust. 2014. Tissue-resident memory T cells. *Immunity*. 41:886–897. <https://doi.org/10.1016/j.immuni.2014.12.007>
- Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat. Methods*. 9: 676–682. <https://doi.org/10.1038/nmeth.2019>
- Shugay, M., O.V. Britanova, E.M. Merzlyak, M.A. Turchaninova, I.Z. Mamedov, T.R. Tuganbaev, D.A. Bolotin, D.B. Staroverov, E.V. Putintseva, K. Plevova, et al. 2014. Towards error-free profiling of immune repertoires. *Nat. Methods*. 11:653–655. <https://doi.org/10.1038/nmeth.2960>
- Shwetank, H.A., H.A. Abdelsamed, E.L. Frost, H.M. Schmitz, T.E. Mockus, B.A. Youngblood, and A.E. Lukacher. 2017. Maintenance of PD-1 on brain-resident memory CD8 T cells is antigen independent. *Immunol. Cell Biol.* 95:953–959. <https://doi.org/10.1038/icb.2017.62>
- Siddiqui, I., K. Schaeuble, V. Chennupati, S.A. Fuertes Marraco, S. Calderon-Copete, D. Pais Ferreira, S.J. Carmona, L. Scarpellino, D. Gfeller, S. Pradervand, et al. 2019. Intratumoral Tcf1⁺PD-1⁺CD8⁺ T Cells with Stem-like Properties Promote Tumor Control in Response to Vaccination and Checkpoint Blockade Immunotherapy. *Immunity*. 50: 195–211.e10. <https://doi.org/10.1016/j.immuni.2018.12.021>
- Simon, S., and N. Labarriere. 2017. PD-1 expression on tumor-specific T cells: Friend or foe for immunotherapy? *OncolImmunology*. 7:e1364828. <https://doi.org/10.1080/2162402X.2017.1364828>
- Simoni, Y., E. Becht, M. Fehlings, C.Y. Loh, S.-L. Koo, K.W.W. Teng, J.P.S. Yeong, R. Nahar, T. Zhang, H. Kared, et al. 2018. Bystander CD8⁺ T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature*. 557:575–579. <https://doi.org/10.1038/s41586-018-0130-2>
- Soneson, C., and M.D. Robinson. 2018. Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods*. 15:255–261. <https://doi.org/10.1038/nmeth.4612>
- Stinchcombe, J.C., D.C. Barral, E.H. Mules, S. Booth, A.N. Hume, L.M. Machesky, M.C. Seabra, and G.M. Griffiths. 2001. Rab27a is required for regulated secretion in cytotoxic T lymphocytes. *J. Cell Biol.* 152:825–834. <https://doi.org/10.1083/jcb.152.4.825>
- Stubbington, M.J.T., T. Lönnberg, V. Proserpio, S. Clare, A.O. Speak, G. Dougan, and S.A. Teichmann. 2016. T cell fate and clonality inference from single-cell transcriptomes. *Nat. Methods*. 13:329–332. <https://doi.org/10.1038/nmeth.3800>
- Tachibana, M., E. Kiyokawa, S. Hara, S. Iemura, T. Natsume, T. Manabe, and M. Matsuda. 2009. Ankyrin repeat domain 28 (ANKRD28), a novel binding partner of DOCK180, promotes cell migration by regulating focal adhesion formation. *Exp. Cell Res.* 315:863–876. <https://doi.org/10.1016/j.yexcr.2008.12.005>
- Thommen, D.S., V.H. Koelzer, P. Herzig, A. Roller, M. Trefny, S. Dimeloe, A. Kiialainen, J. Hanhart, C. Schill, C. Hess, et al. 2018. A transcriptionally and functionally distinct PD-1⁺ CD8⁺ T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. *Nat. Med.* 24:994–1004. <https://doi.org/10.1038/s41591-018-0057-z>
- Trapnell, C., L. Pachter, and S.L. Salzberg. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 25:1105–1111. <https://doi.org/10.1093/bioinformatics/btp120>
- Trapnell, C., D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N.J. Lennon, K.J. Livak, T.S. Mikkelsen, and J.L. Rinn. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32:381–386. <https://doi.org/10.1038/nbt.2859>
- Utting, O., B.J. Sedgmen, T.H. Watts, X. Shi, R. Rottapel, A. Iulianella, D. Lohnes, and A. Veillette. 2004. Immune functions in mice lacking Clnk, an SLP-76-related adaptor expressed in a subset of immune cells. *Mol. Cell Biol.* 24:6067–6075. <https://doi.org/10.1128/MCB.24.13.6067-6075.2004>
- Utzschneider, D.T., M. Charmoy, V. Chennupati, L. Pousse, D.P. Ferreira, S. Calderon-Copete, M. Danilo, F. Alfei, M. Hofmann, D. Wieland, et al. 2016. T Cell Factor 1-Expressing Memory-like CD8(+) T Cells Sustain the Immune Response to Chronic Viral Infections. *Immunity*. 45: 415–427. <https://doi.org/10.1016/j.immuni.2016.07.021>
- Wherry, E.J. 2011. T cell exhaustion. *Nat. Immunol.* 12:492–499. <https://doi.org/10.1038/ni.2035>
- Witherden, D.A., P. Verdino, S.E. Rieder, O. Garijo, R.E. Mills, L. Teyton, W.H. Fischer, I.A. Wilson, and W.L. Havran. 2010. The junctional adhesion molecule JAML is a costimulatory receptor for epithelial gammadelta T cell activation. *Science*. 329:1205–1210. <https://doi.org/10.1126/science.1192698>
- Wood, O., J. Woo, G. Seumo, N. Savelyeva, K.J. McCann, D. Singh, T. Jones, L. Peel, M.S. Breen, M. Ward, et al. SPARC Consortium. 2016. Gene expression analysis of TIL rich HPV-driven head and neck tumors reveals a distinct B-cell signature when compared to HPV independent tumors. *Oncotarget*. 7:56781–56797. <https://doi.org/10.18632/oncotarget.10788>
- Wood, O., J. Clarke, J. Woo, A.H. Mirza, C.H. Woelk, G.J. Thomas, P. Vijayanand, E. King, and C.H. Ottensmeier. 2017. Head and neck squamous cell carcinomas are characterized by a stable immune signature within the primary tumor over time and space. *Clin. Cancer Res.* 23:7641–7649. <https://doi.org/10.1158/1078-0432.CCR-17-0373>
- Xiao, Y., G. Qiao, J. Tang, R. Tang, H. Guo, S. Warwar, W.Y. Langdon, L. Tao, and J. Zhang. 2015. Protein Tyrosine Phosphatase SHP-1 Modulates T Cell Responses by Controlling Cbl-b Degradation. *J. Immunol.* 195: 4218–4227. <https://doi.org/10.4049/jimmunol.1501200>
- Yang, C.Y., J.A. Best, J. Knell, E. Yang, A.D. Sheridan, A.K. Jesionek, H.S. Li, R.R. Rivera, K.C. Lind, L.M. D'Cruz, et al. 2011. The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8⁺ T cell subsets. *Nat. Immunol.* 12:1221–1229. <https://doi.org/10.1038/ni.2158>