

ARTICLE

# High-dimensional single cell analysis identifies stem-like cytotoxic CD8<sup>+</sup> T cells infiltrating human tumors

Jolanda Brummelman<sup>1\*</sup>, Emilia M.C. Mazza<sup>1\*</sup>, Giorgia Alvisi<sup>1</sup>, Federico S. Colombo<sup>2</sup>, Andrea Grilli<sup>3,4</sup>, Joanna Mikulak<sup>5,6</sup>, Domenico Mavilio<sup>5,6</sup>, Marco Alloisio<sup>7</sup>, Francesco Ferrari<sup>8</sup>, Egesta Lopci<sup>9</sup>, Pierluigi Novellis<sup>7</sup>, Giulia Veronesi<sup>7</sup>, and Enrico Lugli<sup>1,2</sup>

**CD8<sup>+</sup> T cells infiltrating tumors are largely dysfunctional, but whether a subset maintains superior functionality remains ill defined. By high-dimensional single cell analysis of millions of CD8<sup>+</sup> T cells from 53 individuals with lung cancer, we defined those subsets that are enriched in tumors compared with cancer-free tissues and blood. Besides exhausted and activated cells, we identified CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells with a partial exhausted phenotype, while retaining gene networks responsible for stem-like plasticity and cytotoxicity, as revealed by single cell sequencing of the whole transcriptome. *Ex vivo*, CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells displayed enhanced self-renewal and multipotency compared with more differentiated subsets and were largely polyfunctional. Analysis of inhibitory and costimulatory receptors revealed PD-1, TIGIT, and CD27 as possible targets of immunotherapy. We thus demonstrate a hierarchy of differentiation in the context of T cell exhaustion in human cancer similar to that of chronically infected mice, which is further shown to disappear with disease progression.**

## Introduction

Despite the abundance of tumor-infiltrating lymphocytes (TILs) that has been associated with a better prognosis in several human cancers, the anti-tumor response mediated by immune cells is insufficient to induce tumor regression, mainly due to the potent immunosuppression found in the tumor microenvironment (TME; Croci et al., 2007; Gajewski et al., 2013). Among TILs, the number, localization, and quality of cytotoxic CD8<sup>+</sup> T cells have been shown to play a major role in this regard (Galon et al., 2013). Specifically, increased infiltration of CD8<sup>+</sup> T cells displaying an effector phenotype and expressing killer molecules such as granzyme B at the tumor invasive margin has been associated with favorable prognosis in colorectal (Pagès et al., 2005; Galon et al., 2006; Bindea et al., 2013) and other types of cancer (Schumacher et al., 2001; Hamanishi et al., 2007; Ganesan et al., 2017). Immunosuppression within the TME comprises several mechanisms that include, but are not limited to, the presence of suppressive populations such as tumor-associated macrophages/myeloid cells and CD4<sup>+</sup> regulatory T cells, chronic antigenic stimulation, and inhibitory metabolites, cytokines, and ligands. All together, they generate a state of dysfunction in T cells also known as T cell exhaustion, characterized by poor proliferative capacity, dimin-

ished cytokine production and killing function, and increased expression of several inhibitory receptors on the cell surface (Wherry, 2011). Current immunotherapeutic strategies targeting inhibitory receptors such as Programmed Death-1 (PD-1) and Cytotoxic T Lymphocyte Antigen-4 (CTLA-4) by blocking antibodies reinvigorate T cells and further boost T cell infiltration, expansion, and effector functions (Leach et al., 1996; Freeman et al., 2000; Sharma and Allison, 2015; Wu et al., 2016; Wei et al., 2017), thus resulting in long-term disease stabilization or, in some cases, tumor regression. However, responses are only seen in a subset of patients (Hodi et al., 2010; Topalian et al., 2012; Jacquilot et al., 2017; Krieg et al., 2018), hence highlighting the need for improved strategies.

The CD8<sup>+</sup> T cell compartment in peripheral blood and tissues is largely diverse, comprising several subsets with different degrees of specialization in phenotype, function, and gene expression (Mahnke et al., 2013; Farber et al., 2014). Recent application of high-content single cell technologies at the level of the whole transcriptome such as single cell RNA sequencing (scRNA-seq; Tirosh et al., 2016; Zheng et al., 2017) and cytometry by time-of-flight (CyTOF; Chevrier et al., 2017; Lavin et al., 2017) suggested

<sup>1</sup>Laboratory of Translational Immunology, Humanitas Clinical and Research Center, Rozzano, Milan, Italy; <sup>2</sup>Humanitas Flow Cytometry Core, Humanitas Clinical and Research Center, Rozzano, Milan, Italy; <sup>3</sup>Department of Biological Sciences, University of Modena and Reggio Emilia, Modena, Italy; <sup>4</sup>PhD Program of Molecular and Translational Medicine, Department of Medical Biotechnology and Translational Medicine, University of Milan, Segrate, Italy; <sup>5</sup>Unit of Clinical and Experimental Immunology, Humanitas Clinical and Research Center, Rozzano, Milan, Italy; <sup>6</sup>Department of Medical Biotechnologies and Translational Medicine, University of Milan, Milan, Italy; <sup>7</sup>Division of Thoracic Surgery, Humanitas Clinical and Research Center, Rozzano, Milan, Italy; <sup>8</sup>IFOM, the FIRC Institute of Molecular Oncology, Milan, Italy; <sup>9</sup>Nuclear Medicine Department, Humanitas Clinical and Research Hospital, Milan, Italy.

\*J. Brummelman and E.M.C. Mazza contributed equally to this paper; Correspondence to Enrico Lugli: [enrico.lugli@humanitasresearch.it](mailto:enrico.lugli@humanitasresearch.it).

© 2018 Brummelman 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/>).

that also CD8<sup>+</sup> TILs are functionally heterogeneous, displaying different levels of T cell activation and exhaustion (Tirosh et al., 2016; Chevrier et al., 2017). Differential analysis of these states, which may coexist in the same T cell or be mutually exclusive, led to the identification of new molecular regulators of exhaustion, such as the transcription factor GATA-3, which inhibits T cell effector function (Singer et al., 2017). Along the same lines, recent data obtained from chronically infected mice showed that the exhausted PD-1<sup>+</sup> T cell compartment is organized in a hierarchy of differentiation, comprising less-differentiated cells with intermediate levels of PD-1 (PD-1<sup>int</sup>) that retain self-renewal capacity and differentiation potential upon PD-1 blockade and PD-1 high (PD-1<sup>hi</sup>) cells that are terminally differentiated and functionally inefficient (He et al., 2016; Im et al., 2016; Leong et al., 2016; Utzschneider et al., 2016). This precursor-progeny relationship is reminiscent of that seen in the blood and lymphoid tissues in physiological situations, where CD8<sup>+</sup> T memory stem cells (T<sub>SCM</sub>) are at the apex of the differentiation program and are currently considered a major reservoir of long-term immunity (Gattinoni et al., 2011; Lugli et al., 2013a; Fuertes Marraco et al., 2015; Oliveira et al., 2015; Akondy et al., 2017). Whether a stepwise differentiation program is also present in the context of T cell exhaustion in human tumors is still not clear.

In this study, 27-parameter flow cytometry applied to primary non-small cell lung cancer (NSCLC) samples, the adjacent cancer-free tissue and the blood enabled us to investigate millions of single CD8<sup>+</sup> T cells and thus to identify rare immunophenotypes that are exclusively present within the tumor. In combination with scRNA-seq, we report the occurrence of stem-like CD8<sup>+</sup> T cells that lack the canonical T<sub>SCM</sub> phenotype of circulating T cells. Rather, they simultaneously display a partial exhausted phenotype along with potent effector potential and enhanced self-renewal capacity. We further show that these CXCR5<sup>+</sup> TILs are lost with disease progression.

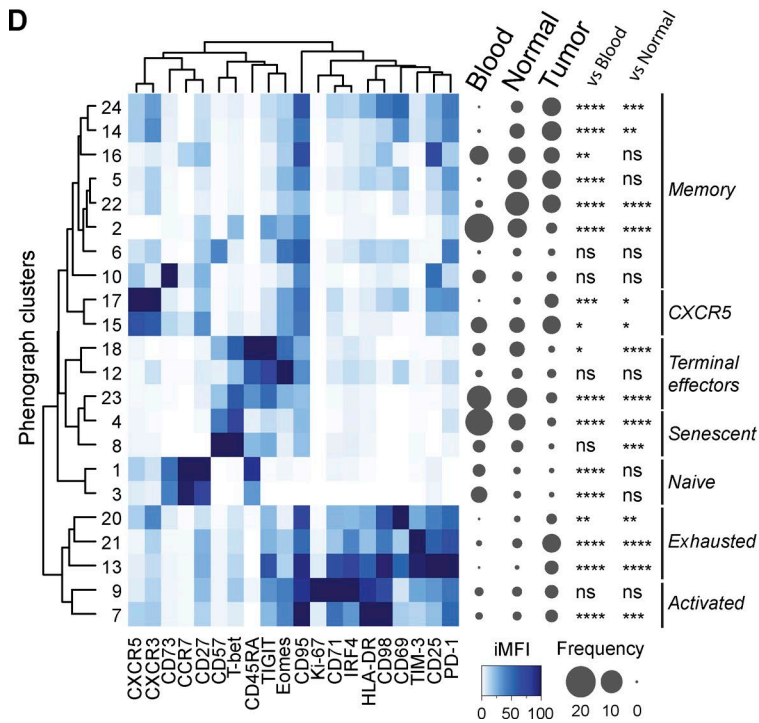
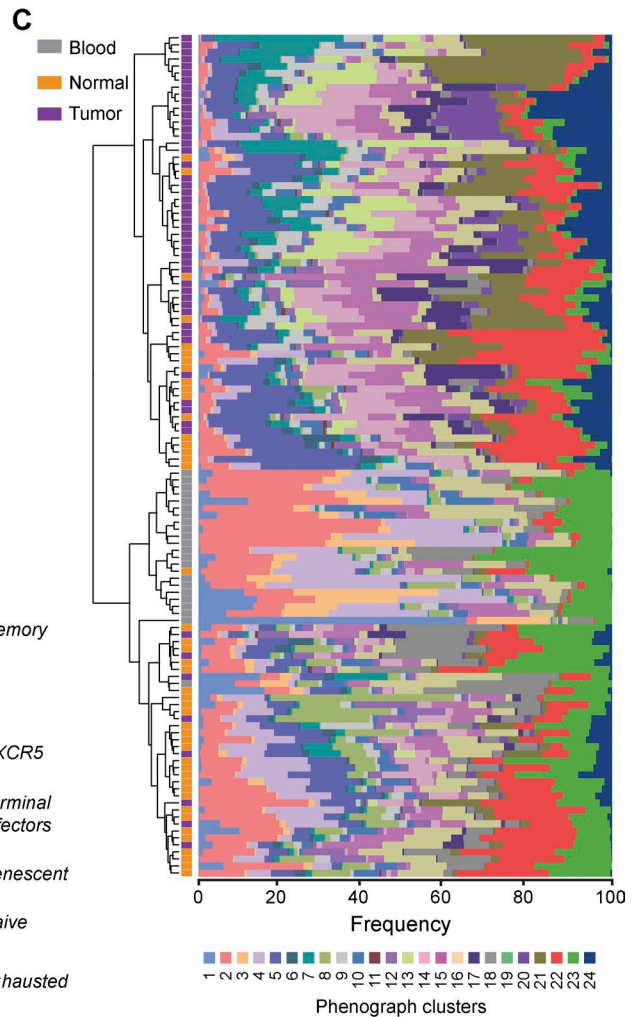
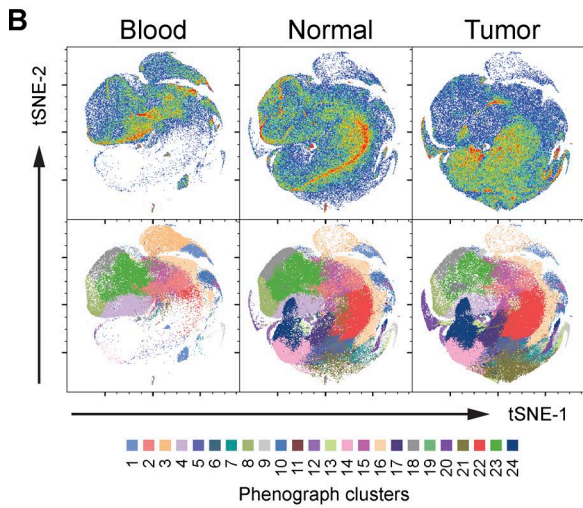
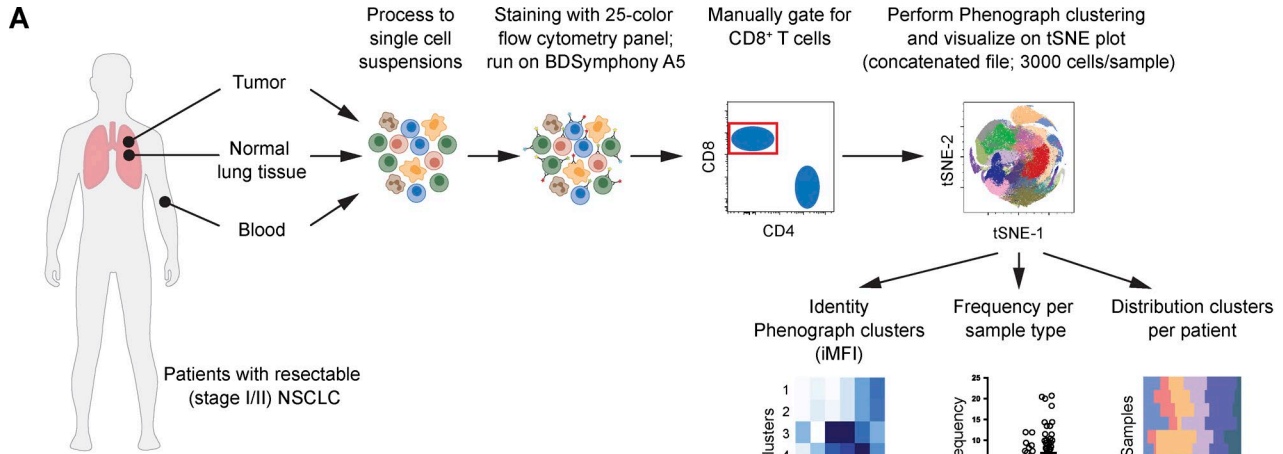
## Results

### High-dimensional single cell analysis of NSCLC-infiltrating CD8<sup>+</sup> T cells identifies tumor-enriched immunophenotypes

To better characterize the diversity occurring at the level of the CD8<sup>+</sup> T cell compartment specifically within the tumor, we developed a polychromatic flow cytometry panel capable of simultaneously investigating 27 parameters on millions of single cells from resected tumor ( $n = 53$ ), paired adjacent cancer-free lung tissue (hereafter referred to as “normal”;  $n = 45$ ), and peripheral blood ( $n = 22$ ) samples of treatment-naïve, NSCLC patients (Fig. 1A and Table 1). A list of the markers included in the panel, providing information on differentiation, activation, proliferation, and exhaustion status of T cells (see Mahnke et al., 2013), is reported in Table S1. Initial data representation with t-distributed stochastic neighboring embedding (tSNE), used to reduce the dimensionality of the 27-parameter dataset (Amir et al., 2013), revealed that single cells from the three districts are substantially diverse in terms of their CD8<sup>+</sup> T cell profile (Fig. 1B, top row). To gain more insight on the immunophenotypes that preferentially characterize CD8<sup>+</sup> T cells from the different specimens, we concatenated

CD8<sup>+</sup> T cell fluorescent data from all samples and analyzed them with Phenograph (Fig. 1A), an unsupervised clustering algorithm suitable for high-dimensional single cell data (Levine et al., 2015). Phenograph identified 24 different CD8<sup>+</sup> putative subpopulations (“clusters;” visualized on top of the tSNE plot; Fig. 1B, bottom row), whose frequencies are subsequently determined in each sample type. Hierarchical metaclustering of these data clearly separated blood and tumor specimens, indicating that they are substantially different at the high-dimensional single cell level (Fig. 1C). Instead, normal tissues were split into two major groups that were more similar to either the blood or the tumor samples.

To define the phenotypic identity of immunophenotypes resulting from Phenograph clustering, we calculated the integrated median fluorescence intensity (iMFI) values (Darrah et al., 2007) of each marker in each Phenograph cluster and visualized values in a heat map, along with the median frequency of the same clusters in the blood, normal tissues, and tumor samples (Fig. 1D). Hierarchical metaclustering of Phenograph clusters (rows) and markers (columns) grouped subpopulations with similar immunophenotypes. Two of these clusters (1 and 3) were mostly present within the blood and displayed a phenotypic identity coincident with naive T cells, i.e., characterized by high levels of CD45RA, CCR7, and CD27 and by the lack of multiple memory markers, such as CD95, PD-1, and T-bet. Other clusters, including 24, 14, 5, and 22, were virtually absent from blood while present in comparable abundance in both the normal tissue and tumor samples, indicating that they are lung-specific subsets. Despite displaying a common CD95<sup>+</sup> CD27<sup>lo</sup> T-bet<sup>-</sup> memory phenotype, these were mainly characterized by variable levels of CD69 and CXCR3 expression, among others, thereby indicating tissue-residency (Wong et al., 2016). The abundance of these phenotypic clusters was counterbalanced by the absence of CD27<sup>int</sup> T-bet<sup>int</sup> CD69<sup>-</sup> transitional memory cells (cluster 2) that instead are highly frequent in the blood. Importantly, the tumor lacked multiple subsets of CD45RA<sup>+/−</sup> CCR7<sup>-</sup> terminal effector cells (clusters 18, 23, 4, and 8), characterized by variable levels of CD57 and T-bet expression, that are known to be highly cytolytic, but replicative senescent (Takata and Takiguchi, 2006; Chattopadhyay et al., 2009; Mahnke et al., 2013), leading us to conclude that the tumor is largely devoid of anti-tumor effectors. On the contrary, multiple clusters were specifically enriched in the tumor samples. These included both exhausted CD8<sup>+</sup> T cell clusters 21, 13, and, to a lesser extent, 20, characterized by high expression of activation markers HLA-DR, CD25, and IRF4, and different combinations in expression of the inhibitory receptors PD-1, TIM-3, and TIGIT. Activated CD8<sup>+</sup> T cell cluster 7, expressing only high levels of the activation markers, but not the proliferation marker Ki-67, was also enriched in lung tumors. Remarkably, Phenograph identified two memory clusters expressing CXCR5, i.e., 15 and 17, the latter featuring the highest expression. In addition, only the frequency of cluster 17, characterized by a CD69<sup>+</sup> PD-1<sup>int</sup> CD45RA<sup>-</sup> CCR7<sup>-</sup> effector-like phenotype, but lacking additional activation and exhaustion markers, was most different between the tumor and the adjacent lung tissue or the blood, suggesting this is a tumor-specific subpopulation (Fig. 1D and Fig. S1).



### NSCLC tumors are enriched in CXCR5<sup>+</sup> CD8<sup>+</sup> T cells

It has been recently reported that chronically infected mice harbor, besides CXCR5<sup>+</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells that are exhausted and thus incapable of persisting in the long term, CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cell progenitors that are PD-1<sup>int</sup> and, upon PD-1 blockade, are responsible for controlling chronic viral infections (He et al., 2016; Im et al., 2016; Leong et al., 2016; Utzschneider et al., 2016; Wu et al., 2016). Therefore, these cells share phenotypic features with human CXCR5<sup>+</sup> CD8<sup>+</sup> TILs identified in Phenograph clusters 15 and 17. We initially confirmed that results obtained with Phenograph clustering could be obtained by standard approaches and indeed found that CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells isolated by manual gating are enriched within tumors (Fig. 2, A and B), indicating that unsupervised clustering of multidimensional single cell data are a reliable tool for the identification of phenotypically different subpopulations. In addition to the CXCR5<sup>+</sup> TIM-3<sup>-</sup> and CXCR5<sup>+</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells, our manual analysis identified a third subset of CD8<sup>+</sup> TILs lacking both CXCR5 and TIM-3 expression that is infrequent in chronically infected mice, indicating that human TILs are characterized by additional phenotypic complexity. As the major difference between clusters 15 and 17 is not only CXCR5 abundance, but also expression of the tissue residency marker CD69, we next sought to determine to which extent CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells in tumors express CD69. The vast majority of these cells were CD69<sup>+</sup> compared with those from normal tissue and blood samples (Figs. 2, C–E). Instead, CD69<sup>-</sup> CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells were relatively infrequent in both normal lung tissues and tumors, and their abundance decreased only slightly compared with blood (Fig. 2 F), thereby indicating that Phenograph overestimates the frequency of this phenotype. As a whole, lung cancer-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells were largely different compared with those from blood and the normal tissue at the phenotypic level, as they displayed a uniform CD45RA<sup>-</sup>CCR7<sup>-</sup>, effector memory-like phenotype, further characterized by increased expression of CD27 and PD-1, high levels of CXCR3, and lacking CD57 and T-bet (Fig. S1). In addition, CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells could be detected in primary colorectal carcinomas (1<sup>st</sup> colorectal carcinoma [CRC]; Fig. 2 G) and colorectal liver metastases (Fig. 2 H) at comparable frequencies, and, similarly to NSCLC-infiltrating TILs, they were for the vast majority CD69<sup>+</sup>.

We next determined how the CXCR5<sup>+</sup> CD8<sup>+</sup> TILs relate to other subpopulations within lung tumors. Pearson correlation analysis between the frequency of the CD69<sup>+</sup> CXCR5<sup>+</sup> cluster 17 and the additional clusters identified by Phenograph revealed a significant negative correlation with cluster 7 (HLA-DR<sup>+</sup> CD98<sup>hi</sup>

activated cells), cluster 9 (Ki-67<sup>+</sup> proliferating cells), and cluster 13 (exhausted cells; Fig. 2 I). Confirming what was previously revealed by bioinformatic analysis, Ki-67 was poorly expressed by CXCR5<sup>+</sup> CD8<sup>+</sup> T cells as determined by manual gating, while it increased progressively in CXCR5<sup>+</sup> TIM-3<sup>-</sup> and CXCR5<sup>+</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> TILs (Fig. 2, J and K). As a whole, tumor-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells are a quiescent population with hallmarks of tissue residency, but not of exhaustion and activation.

### Single cell transcriptomics defines the memory identity and direct cytotoxic potential of CXCR5<sup>+</sup> CD8<sup>+</sup> T cells

Due to their enhanced self-renewal capacity and multipotency in the context of chronic infection, CXCR5<sup>+</sup> CD8<sup>+</sup> T cells resemble stem-like memory T cells (He et al., 2016; Im et al., 2016; Leong et al., 2016; Utzschneider et al., 2016). The presence of a population with stem-like characteristics in human tumors accessible to PD-1 or, more in general, checkpoint blockade, would be extremely beneficial for the local anti-tumor immune response. Canonical human T<sub>SCM</sub> cells found in lymphoid tissues and circulation bear a CD45RA<sup>+</sup>CCR7<sup>+</sup>CD27<sup>+</sup>CD95<sup>+</sup> phenotype (Zhang et al., 2005; Gattinoni et al., 2011; Lugli et al., 2013a) that is largely different from that of the identified CXCR5<sup>+</sup> CD8<sup>+</sup> TILs (Fig. 1 D). These T<sub>SCM</sub> cells were virtually absent in NSCLC TILs and in lung tissues (Fig. S2), indicating that the lung does not support their development or infiltration and subsequent maintenance. As mouse CXCR5<sup>+</sup> CD8<sup>+</sup> T cells highly respond to PD-1 blockade, we sought to determine whether human tumor-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells have shared features. The gene expression profile of CXCR5<sup>+</sup> CD8<sup>+</sup> T cells from chronically infected mice is similar to that of CD4<sup>+</sup> T follicular helper (T<sub>fh</sub>) cells and of CD8<sup>+</sup> T cell memory precursors (He et al., 2016; Im et al., 2016; Leong et al., 2016; Utzschneider et al., 2016), suggesting they interact with B cells in lymphoid tissues and retain a long-lived memory potential, respectively. Despite different tumor types having largely different biology, infiltrating CD8<sup>+</sup> T cells share overlapping phenotypes, leading us to reason that a publicly available scRNA-seq dataset of CD45<sup>+</sup> leukocytes isolated from melanoma patients could better define the identity of CXCR5<sup>+</sup> CD8<sup>+</sup> T cells at the global transcriptomic level (Tirosh et al., 2016). In this dataset, a total of 912 cells were labeled as CD8<sup>+</sup>, of which 58 expressed CXCR5, thus comprising 6.4% of the melanoma-infiltrating CD8<sup>+</sup> T cells, a frequency similar to that of NSCLC-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells. Initial exploratory analysis by Pearson correlation between the expression of CXCR5 and of all the genes in the matrix revealed that transcripts typically present in long-lived memory T cells such as TCF7 (encoding TCF-1), WNT2, MYB, CD28, CCR7, BCL6,

Figure 1. **High-dimensional single cell analysis of tumor-infiltrating CD8<sup>+</sup> T cells in NSCLC identifies tumor-enriched immunophenotypes.** (A) Illustration depicting the experimental workflow. Analysis is performed on peripheral blood ( $n = 22$ ), normal lung tissue ( $n = 45$ ), and tumor ( $n = 53$ ) from NSCLC patients. Data were obtained in six independent experiments. (B) tSNE analysis of concatenated CD8<sup>+</sup> T cells (3,000 cells/sample) from peripheral blood ( $n = 22$ ), normal lung tissue ( $n = 45$ ), and tumor ( $n = 53$ ) samples from NSCLC patients (top panel) and color-coded tSNE map depicting clusters identified by Phenograph (bottom panel). (C) Hierarchical metaclustering (using Ward's method) of the frequencies of the 24 Phenograph clusters in single blood (gray;  $n = 22$ ), normal tissue (orange;  $n = 45$ ), and tumor (purple;  $n = 53$ ) samples. (D) Heat map showing the iMFI of specific markers in discrete Phenograph clusters identified in B. Phenograph clusters (rows) and markers (columns) are hierarchically metaclustered using Ward's method to group subpopulations with similar immunophenotypes. The median frequency of each Phenograph cluster within the blood ( $n = 22$ ), normal lung tissue ( $n = 45$ ), and tumor ( $n = 53$ ) samples are depicted by balloon plots. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$  tumor versus blood and normal tissue samples; two-way ANOVA with Bonferroni post-hoc test.

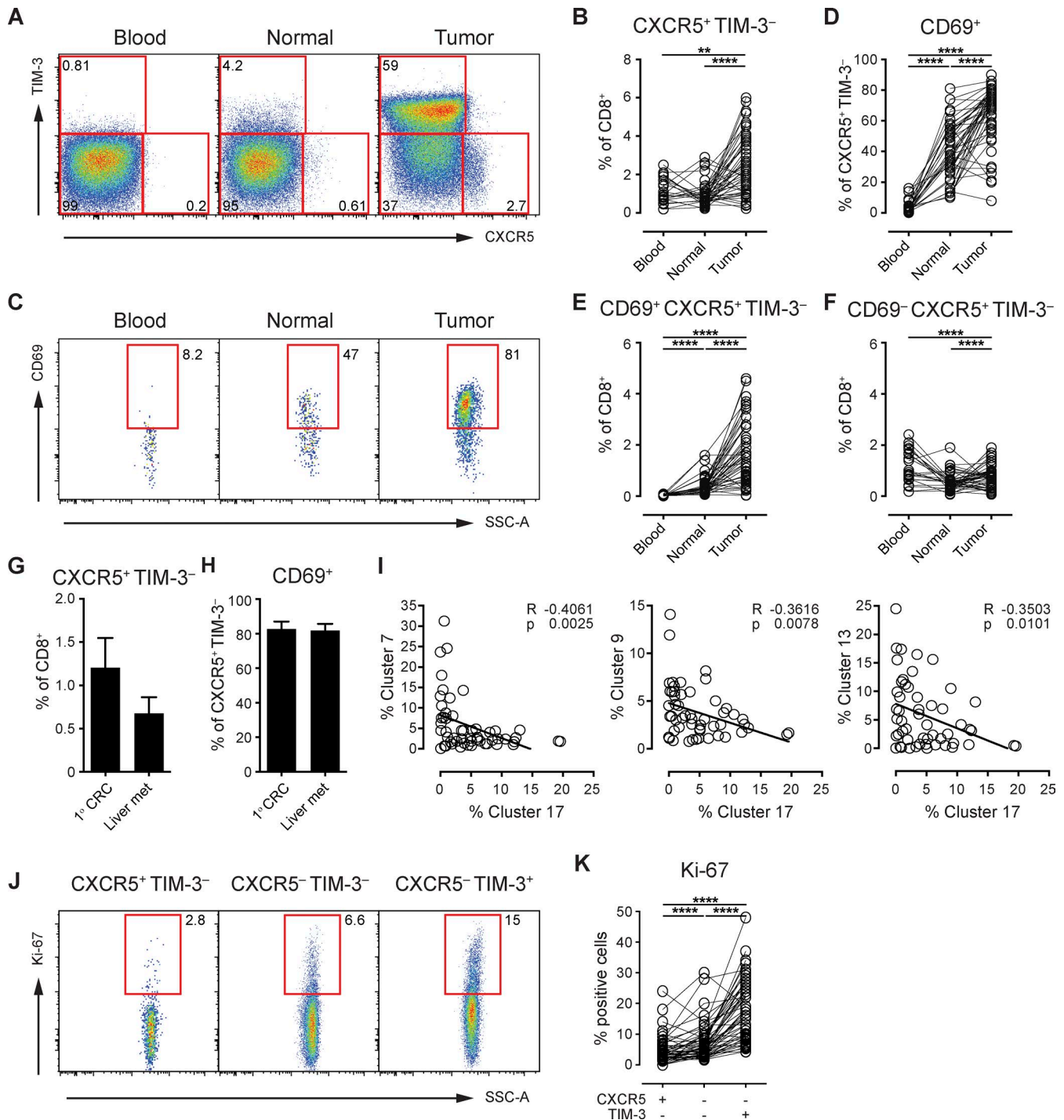
Table 1. NSCLC patients' characteristics

Parameter	Subdivision	Patients (n = 53)
Sex - no. (%)	Female	27 (50.9)
	Male	26 (49.1)
Age - mean (SD)		68 (8.2)
Smoking status - no. (%)	Smoker	29 (54.7)
	Former	17 (32.1)
	Non-smoker	6 (11.3)
	Unknown	1 (1.9)
Tumor type - no. (%)	ADK lepidic	15 (28.3)
	ADK solid & acinar	12 (22.6)
	ADK papillary	8 (15.1)
	ADK micropapillary	2 (3.8)
	ADK mucinous	2 (3.8)
	SSC	12 (22.6)
	Other	2 (3.8)
pStage - no. (%)	IA2	5 (9.4)
	IA3	6 (11.3)
	IB	7 (13.2)
	IIA	4 (7.5)
	IIB	15 (28.3)
	IIIA	7 (13.2)
	IIIB	6 (11.3)
	IVA	2 (3.8)
	Tumor location - no. (%)	R
RLL		8 (15.1)
RML		2 (3.8)
RUL		15 (28.3)
RML+RLL		2 (3.8)
L		7 (13.2)
LLL		7 (13.2)
LUL		10 (18.9)
Results PET-CT scan available - no. (%)	Yes	33 (62.3)
	No	20 (37.7)

and *SELL* (encoding CD62L) are correlated with *CXCR5* expression, while those present in exhausted or more differentiated effector CD8<sup>+</sup> T cells such as *HAVCR2* (encoding TIM-3), *GZMB*, *PRFI*, *IFNG*, and *PRDMI* (encoding Blimp-1) are anti-correlated (Fig. 3 A). To confirm these results, we further subdivided CD8<sup>+</sup> TILs from this dataset in three subsets based on *CXCR5* and *HAVCR2* expression levels. In particular, based on the combination of these two markers, we were able to identify cells that are CXCR5<sup>+</sup> and express low levels of *HAVCR2* (encoding TIM-3; see Materials and methods), CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup> T cells (reflecting those that were identified using flow cytometry). As the vast majority of CXCR5<sup>+</sup> CD8<sup>+</sup> T cells from melanoma expressed

*CD69* mRNA (Fig. S3 A), in line with protein data of lung tumors, we did not further subdivide between *CD69*<sup>-</sup> and *CD69*<sup>+</sup> for the analysis of gene expression (Fig. 3 A). In this way, we confirmed the preferential expression of *TCF7*, *EOMES*, *BCL6*, *FOXO1*, and *CD28* by CXCR5<sup>+</sup> T cells (Fig. 3, B and C) and revealing additional Tfh cell-associated genes such as *SH2D1A* (encoding SAP) and *Ly9* (Fig. S3 B). Notably, a large proportion of these cells also expressed killer molecules such as *GZMA*, *GZMK*, *GZMM*, and *PRFI* (the latter at slightly lower levels compared with CXCR5<sup>-</sup> TIM-3<sup>+</sup>). Importantly, the same cells expressed inhibitory receptors such as *PDCD1* (encoding PD-1) and *TIGIT*, albeit at lower levels compared with CXCR5<sup>-</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells, whereas they lacked *LAG3* and *CTLA4* (Fig. 3 C). Besides *CD28*, also co-stimulatory receptors *CD27* and *TNFRSF9* (encoding 4-1BB) are expressed by CXCR5<sup>+</sup> CD8<sup>+</sup> T cells, yet the latter at a lower level compared with CXCR5<sup>-</sup> TIM-3<sup>+</sup> cells. Overall, CXCR5<sup>-</sup> TIM-3<sup>-</sup> cells had a pattern of gene expression that is intermediate between these two subsets (Fig. 3, B and C).

Gene set enrichment analysis (GSEA) revealed significant overrepresentation of germinal-center CD4<sup>+</sup> Tfh and CD8<sup>+</sup> naive, but not CD8<sup>+</sup> effector T cell gene signatures within the CXCR5<sup>+</sup> compared with CXCR5<sup>-</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells (Fig. 3 D). Instead, naive and effector gene signatures were simultaneously more enriched when compared with CXCR5<sup>-</sup> TIM-3<sup>-</sup> (Fig. S3 C), further indicating that CXCR5<sup>+</sup> CD8<sup>+</sup> T cells have features of long-lived memory progenitors that are preferentially armed for effector functions. Most importantly, the genetic signature of murine CXCR5<sup>+</sup> TIM-3<sup>-</sup> cells isolated from chronically infected mice as from Im et al. (2016) was highly enriched in human CXCR5<sup>+</sup> TILs compared with CXCR5<sup>-</sup> TIM-3<sup>+</sup> (Fig. 3 D) and CXCR5<sup>-</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells (Fig. S3 C), further highlighting the similarity of human and murine subsets. In line with their follicular identity, the abundance of CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells positively correlated with that of CD4<sup>+</sup> CXCR5<sup>+</sup> (comprising Tfh cells) and, at a lesser extent, of CD19<sup>+</sup> B cells within the tumor (Fig. S3, D-G). As expected, the frequency of CD4<sup>+</sup> Tfh highly correlated with that of CD19<sup>+</sup> B cells, confirming their mutual relationship also in lung cancer (Fig. S3 H). Several of the transcripts found differentially expressed in melanoma-infiltrating CD8<sup>+</sup> T cell subsets by scRNA-seq were then confirmed at the protein level in the NSCLC samples by using flow cytometry (*TCF-1*, *Eomes*, *CD27*, *CD28*, and *PD-1*; Fig. 3, E and F), thus further indicating that CXCR5<sup>+</sup> CD8<sup>+</sup> TILs share similar identity among different tumors. This flow cytometric analysis revealed additional differentially expressed markers that could not be measured by scRNA-seq due to low levels of transcripts in all populations (unpublished data), including T-bet (lower in CXCR5<sup>+</sup>) and CXCR3 (higher in CXCR5<sup>+</sup>). Most notably, CXCR3 is a signature marker of T<sub>SCM</sub> cells (Gattinoni et al., 2011). Fig. S3 (I-K) additionally shows the phenotype of CD8<sup>+</sup> TILs compared with naive (T<sub>N</sub>) and effector memory T (T<sub>EM</sub>) cells from the blood of a healthy donor, confirming that CXCR5<sup>+</sup> TIM-3<sup>-</sup> cells coexpress markers of T cell precursors along with intermediate expression of PD-1 and TIGIT. We conclude that CXCR5<sup>+</sup> CD8<sup>+</sup> TILs are armed for effector functions while retaining features of memory cells and express molecules that are candidate for immunotherapeutic intervention, such as PD-1, TIGIT, and CD27, but not LAG-3, CTLA-4, or TIM-3.

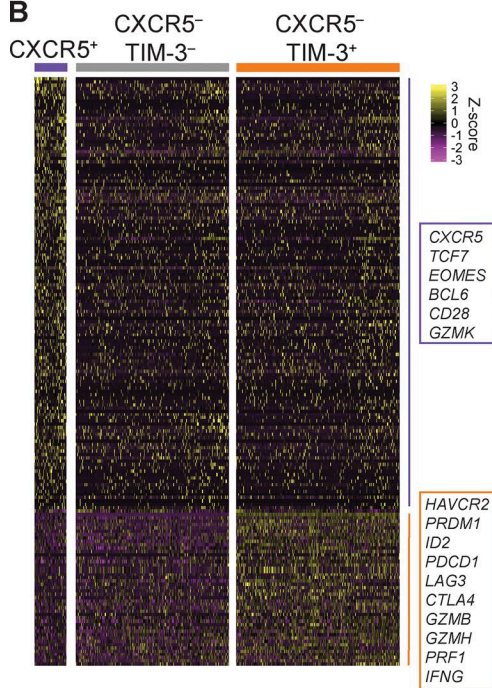


**Figure 2. NSCLC tumors are enriched in CXCR5<sup>+</sup> CD8<sup>+</sup> T cells.** (A) Representative dot plots showing manual gating analysis of CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells from a blood, normal lung tissue, and tumor sample. Numbers in the dot plots indicate the percentage of cells identified by the gates. (B) Summary of the frequency of CXCR5<sup>+</sup> TIM-3<sup>-</sup> of CD8<sup>+</sup> T cells from blood (*n* = 22), normal lung tissue (*n* = 45), and tumor (*n* = 53) samples as determined in A. Data were obtained in six independent experiments. (C) Representative dot plots showing CD69 expression by the CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells from a blood, normal lung tissue, and tumor sample. Numbers in the dot plots indicate the percentage of cells identified by the gates. (D–F) Summary of the frequency of CD69<sup>+</sup> within CXCR5<sup>+</sup> TIM-3<sup>-</sup> cells (D) and CD69<sup>+</sup> CXCR5<sup>+</sup> TIM-3<sup>-</sup> (E) and CD69<sup>-</sup> CXCR5<sup>+</sup> TIM-3<sup>-</sup> within CD8<sup>+</sup> T cells (F) from blood (*n* = 22), normal lung tissue (*n* = 45), and tumor (*n* = 53) samples, as determined in C. Data were obtained in six independent experiments. \*\*, *P* < 0.01; \*\*\*\*, *P* < 0.0001; paired Wilcoxon *t* test. (G and H) Frequency of CXCR5<sup>+</sup> TIM-3<sup>-</sup> (G) and CD69<sup>+</sup> expression in CXCR5<sup>+</sup> TIM-3<sup>-</sup> (H) CD8<sup>+</sup> T cells in primary CRC (1<sup>o</sup>CRC; *n* = 6) and colorectal liver metastasis (Liver met; *n* = 7). Data are obtained in one experiment and are shown as mean ± SEM. (I) Pearson correlation between the frequency of Phenograph cluster 17 (CD69<sup>+</sup> CXCR5<sup>+</sup>) with cluster 7 (HLA-DR<sup>+</sup> CD98<sup>hi</sup> activated cells), cluster 9 (Ki-67<sup>+</sup>, proliferating cells), and cluster 13 (exhausted cells). *P* values are shown in each panel. (J) Representative dot plots showing Ki-67 expression by the indicated CD8<sup>+</sup> T cell subsets from a tumor sample. Numbers in the dot plots indicate the percentage of cells identified by the gates. (K) Summary of the frequency of Ki-67<sup>+</sup> cells within the indicated CD8<sup>+</sup> T cell subsets from all tumor samples (*n* = 53), as determined in J. \*\*\*\*, *P* < 0.0001; paired Wilcoxon *t* test. Data were obtained in six independent experiments.

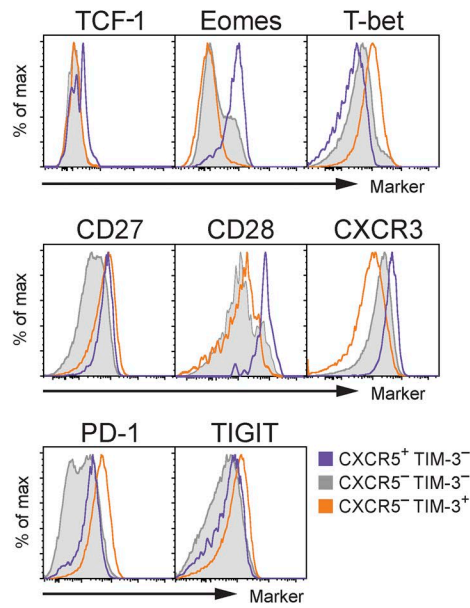
**A**

Rank	Gene	Score	P value
1	<i>CXCR5</i>	1	0
126	<i>TCF7</i>	0.1817	4.22E-06
140	<i>CD28</i>	0.1777	6.87E-06
145	<i>MYB</i>	0.1766	7.84E-06
232	<i>GZMK</i>	0.1578	6.65E-05
259	<i>WNT2</i>	0.1535	1.06E-04
448	<i>CCR7</i>	0.1335	7.62E-04
649	<i>BCL6</i>	0.1216	2.19E-03
660	<i>SELL</i>	0.1298	2.32E-03
849	<i>EOMES</i>	0.1128	4.48E-03
21130	<i>PRF1</i>	-0.0813	4.09E-02
21138	<i>IFNG</i>	-0.0848	3.28E-02
21173	<i>PRDM1</i>	-0.1255	1.56E-03
21176	<i>HAVCR2</i>	-0.1282	1.22E-03
21178	<i>GZMB</i>	-0.1350	6.61E-04

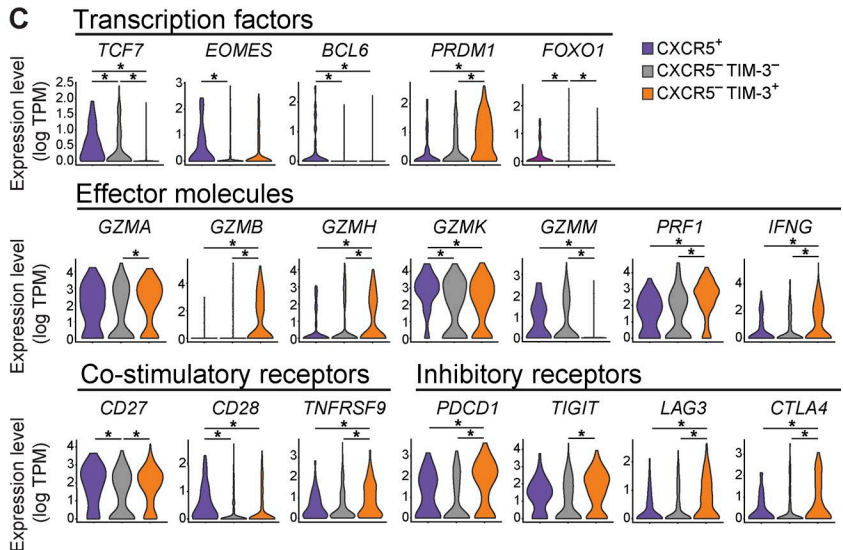
**B**



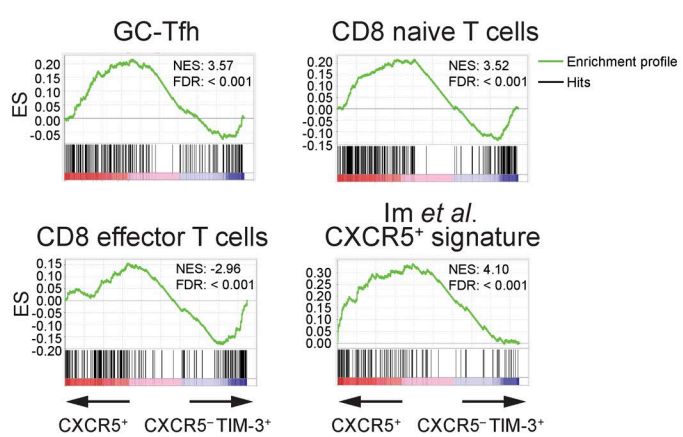
**E**



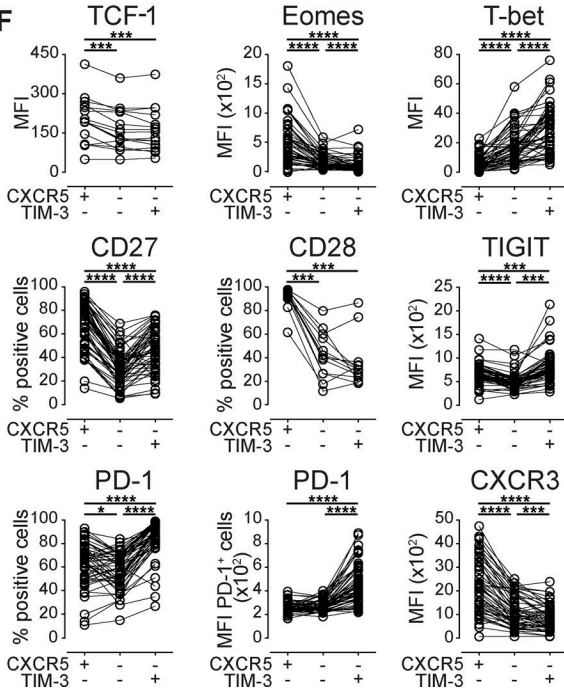
**C**



**D**



**F**



### Tumor-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells have rapid effector functions and are polyfunctional

We next sought to determine whether the lack of *IFNG*, *GZMH*, and *GZMB* would result in decreased effector functions by CXCR5<sup>+</sup> CD8<sup>+</sup> TILs compared with other subsets. All sort-purified T cell subsets rapidly produced IFN $\gamma$ , IL-2, TNF cytokines, and the degranulation marker CD107a upon PMA/ionomycin stimulation *ex vivo* (Fig. 4, A and B). In line with their exhausted phenotype, CXCR5<sup>-</sup> TIM-3<sup>-</sup> were the less functional in this regard, while CXCR5<sup>+</sup> TIM-3<sup>-</sup> expressed the highest level of CD107a (Fig. 4 B). Combinatorial analysis revealed the preferential coproduction of the four effector molecules by CXCR5<sup>+</sup> TIM-3<sup>-</sup> compared with both CXCR5<sup>-</sup> TIM-3<sup>-</sup> and CXCR5<sup>-</sup> TIM-3<sup>+</sup> cells (Fig. 4, C and D). These data indicate that CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> TILs retain rapid effector functions *ex vivo* and display enhanced polyfunctionality compared with other TIL subsets.

### Tumor-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells are endowed with enhanced stemness

We next tested whether NSCLC-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells retain enhanced stem cell-like properties as suggested by their molecular signature. A hallmark of human memory stem cells is their capability to generate more effector progeny while self-renewing. Therefore, we sort-purified CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells from tumor samples by flow cytometry (Fig. 5 A) and stimulated them with  $\alpha$ CD3/28 antibody-conjugated beads in combination with IL-2 and IL-15 or IL-15 alone to determine their multipotency and self-renewal capabilities, respectively (Gattinoni et al., 2011). Fig. 5 (B and D) shows that CXCR5<sup>+</sup> TIM-3<sup>-</sup> and CXCR5<sup>-</sup> TIM-3<sup>-</sup> cells proliferated more vigorously than CXCR5<sup>-</sup> TIM-3<sup>+</sup> cells in response to TCR stimulation. These proliferating subsets displayed high levels of HLA-DR and CD25 activation markers and low levels of CD27 and CD28 (Fig. 5 K and Fig. S4), suggesting further differentiation. In line with murine data of adoptively transferred T cells in hosts bearing a chronic infection, only CXCR5<sup>+</sup> TIM-3<sup>-</sup> cells were able to generate all subsets of T cells identified by combinations of CXCR5 and TIM-3 expression (Fig. 5, E and H), thereby suggesting multipotency. In addition, these cells displayed lower levels of TIM-3 on the cell surface compared with other subsets (Fig. 5 G). CXCR5<sup>+</sup> TIM-3<sup>-</sup> proliferated more than CXCR5<sup>-</sup> TIM-3<sup>+</sup> cells also in response to IL-15 (Fig. 5 D). A similar trend could be seen for CXCR5<sup>-</sup> TIM-3<sup>-</sup>, although not reaching statistical significance. Importantly, the three subsets generated progenies with dramatic differences at the phenotypic level. While CXCR5<sup>+</sup>

TIM-3<sup>-</sup> largely retained their original phenotype (Fig. 5, F and I), characterized by high levels of CD27 and CD28 costimulatory molecules and low levels of CD45RA (Fig. 5 L and Fig. S4), proliferating CXCR5<sup>-</sup> TIM-3<sup>-</sup> cells, the vast majority of which was CD45RA<sup>-</sup> upon *ex vivo* purification, displayed high levels of CD45RA and, at a lesser extent, TIM-3, indicating differentiation toward the terminal effector phenotype (Fig. 5, J and L; and Fig. S4). Most notably, CXCR5 could not be detected in these cells, as well as in CXCR5<sup>-</sup> TIM-3<sup>+</sup> exhausted cells in response to IL-15, hence confirming the unidirectional CXCR5<sup>+</sup> to CXCR5<sup>-</sup> conversion. Thus, the tumor-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells have features of multipotency and stemness, as previously described for circulating T<sub>SCM</sub> cells, in response to TCR-dependent and homeostatic stimulation, respectively.

### Loss of CXCR5<sup>+</sup> CD8<sup>+</sup> TILs with lung cancer disease progression

The enhanced stemness of the tumor-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells suggest that these cells might be beneficial for the anti-tumor CD8<sup>+</sup> T cell response, but their role in disease progression is still not defined. The maximum standardized uptake value (SUV<sub>max</sub>), as obtained from positron emission tomography (PET) before surgery, is an indicator of the glycolytic capacity of the tumor and is considered a negative prognostic parameter for NSCLC (Takeuchi et al., 2014; Liu et al., 2016; Lopci et al., 2016). Indeed, patients with a pathological stage (pStage) I tumor, defined according to the international TNM classification, displayed a lower SUV<sub>max</sub> compared with patients with pStage II or III in our cohort (Fig. 6 A). Therefore, we divided patients in two groups on the basis of the median value for this parameter across the entire cohort (high and low groups) for which both PET values and high-dimensional single cell profiling data were available ( $n = 33$ ). We next determined the frequency of cluster 17 (i.e., CD8<sup>+</sup> T cells with a preferential CD69<sup>+</sup> CXCR5<sup>+</sup> phenotype) in the tumor samples and revealed that this cluster is more prevalent in low glycolytic tumors (Fig. 6 B). Accordingly, we found that the same subpopulation was more prevalent within NSCLC samples of patients with early versus late disease stage (Fig. 6 C). In conclusion, the abundance of CXCR5<sup>+</sup> CD8<sup>+</sup> T cells in the tumor negatively correlates with disease progression.

### Discussion

We applied high-dimensional single cell approaches to analyze the tumor-infiltrating CD8<sup>+</sup> T cells and define those phenotypic subpopulations that are preferentially enriched within the tumor

Figure 3. **Single cell transcriptomics defines the early differentiated memory identity and cytotoxic potential of CXCR5<sup>+</sup> CD8<sup>+</sup> T cells.** (A) Pearson correlation analysis of *CXCR5* profile-correlated genes in the public available scRNA-seq dataset from metastatic melanoma-infiltrating CD8<sup>+</sup> T cells ( $n = 912$ ) from Tirosch et al. (2016). (B) Heat map depicting the Z score of expression values of genes that characterize melanoma-infiltrating CXCR5<sup>+</sup> ( $n = 58$ ), CXCR5<sup>-</sup> TIM-3<sup>-</sup> ( $n = 278$ ), and CXCR5<sup>-</sup> TIM-3<sup>+</sup> ( $n = 297$ ) CD8<sup>+</sup> T cells from the same dataset as in A. (C) Violin plots showing expression probability distributions of indicated genes within the infiltrating CXCR5<sup>+</sup> ( $n = 58$ ), CXCR5<sup>-</sup> TIM-3<sup>-</sup> ( $n = 278$ ), and CXCR5<sup>-</sup> TIM-3<sup>+</sup> ( $n = 297$ ) CD8<sup>+</sup> T cells, as defined in B. \*,  $P < 0.05$ ; Wilcoxon rank sum test. (D) GSEA of the infiltrating CXCR5<sup>+</sup> versus CXCR5<sup>-</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells for the indicated gene sets. ES, enrichment score; NES, normalized enrichment score; FDR, false discovery rate. (E) Representative histograms depicting expression of different markers on tumor-infiltrating CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells as determined by flow cytometry. (F) Summary of the results shown in C. All MFIs refer to marker expression in bulk populations identified by CXCR5 and TIM3 expression, except for PD-1, where MFI refers to the PD-1<sup>+</sup> fraction. All graphs show measurements from 53 donors obtained in six independent experiments, except for TCF-1 ( $n = 16$ ; four independent experiments) and CD28 ( $n = 13$ ; three independent experiments). \*,  $P < 0.05$ ; \*\*\*,  $P < 0.001$ ; \*\*\*\*,  $P < 0.0001$ ; paired Wilcoxon *t* test.



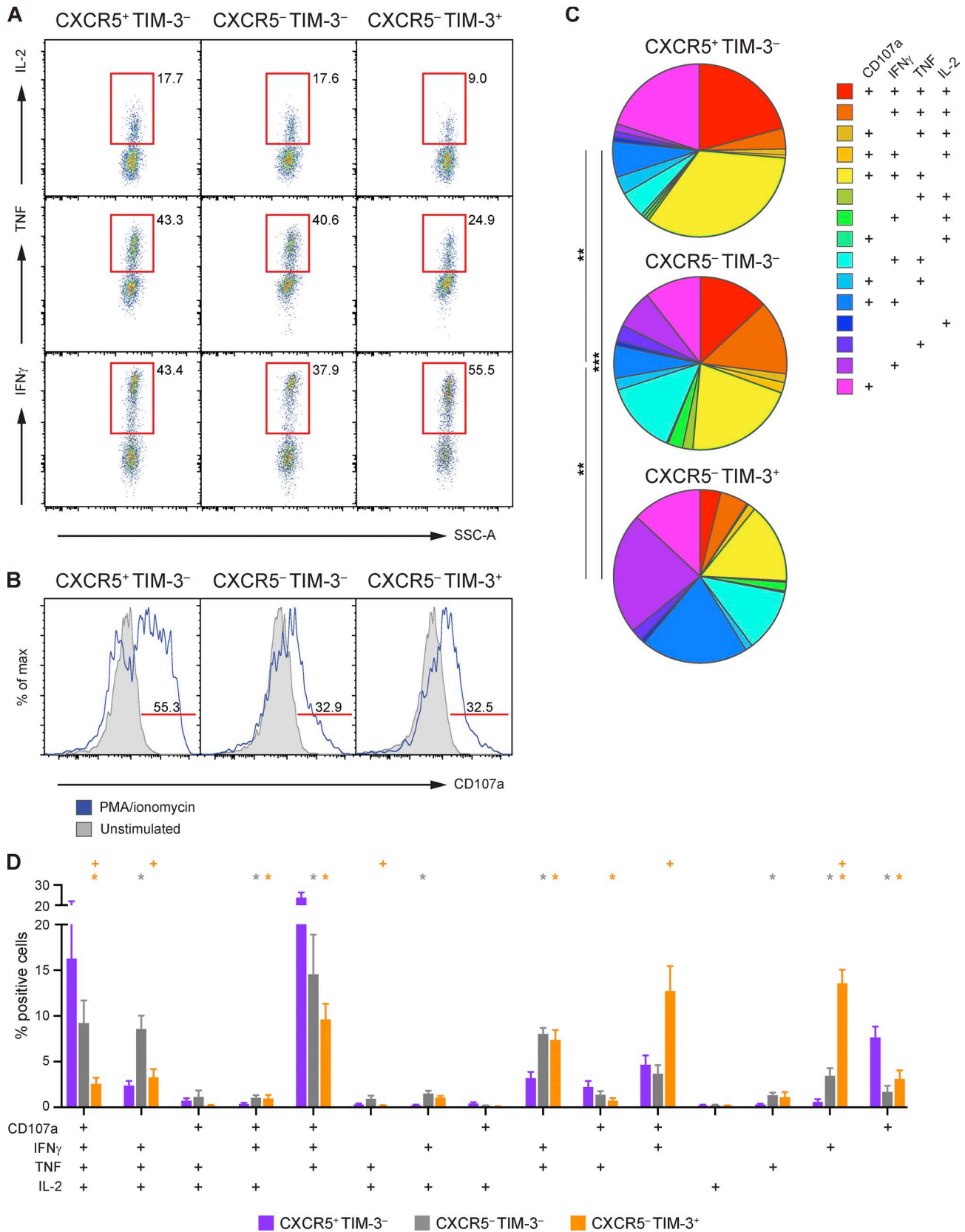


Figure 4. **CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells are efficient in rapidly producing effector molecules.** (A–D) Sorted CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> TILs were stimulated for 3 h with PMA/ionomycin, and CD107a expression and cytokine production was determined by flow cytometry ( $n = 7$ ; all performed in independent experiments). (A) Representative dot plots showing the analysis of IL-2, TNF, and IFN $\gamma$  production by NSCLC tumor-infiltrating CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells following short-term PMA/ionomycin stimulation. Numbers in the dot plots indicate the percentage of cells identified by the gates. (B) Representative histograms showing the expression of CD107a by the indicated subsets after short-term PMA/ionomycin

compared with the blood and the adjacent cancer-free tissue. We adopted a novel flow cytometry technology capable to investigate 27 parameters simultaneously in millions of single T cells, a number still difficult to reach by CyTOF or scRNA-seq due to throughput and cost, respectively, and thus were able to identify also those rare subsets that are present in the TME. Specifically, we focused on CD8<sup>+</sup> T cells that are amenable to immune checkpoint blockade (i.e., those that have traits of exhaustion) while retaining functional capacity, so to exploit their enhanced anti-tumor response. In this way, we revealed the presence of a subset of CD8<sup>+</sup> T cells featuring a CXCR5<sup>+</sup> TIM-3<sup>-</sup> phenotype, long-lived memory characteristics, CD4<sup>+</sup> Tfh-shared gene expression, and enhanced stem cell-like properties when compared with more differentiated subsets at the same site. The presence of these T cells exclusively in the TME is accompanied by increased frequencies of activated and exhausted phenotypes at the expense of CD27<sup>+</sup> memory T cells and highly cytolytic, CD45RA<sup>+</sup> T cells, which instead are normally abundant within the adjacent cancer-free lung tissue from the patient (Fig. 1 D). Our high-dimensional analysis identified a subset of nonneoplastic lung tissues that was more similar to blood samples rather than the tumors, possibly suggesting contamination from the circulating blood. Despite not being formally excluded, this is considered to be minor given the virtual absence of true naive cells (clusters 1 and 3) along with the presence of CD45RA<sup>int/+</sup> CCR7<sup>-</sup> CD57<sup>+</sup> phenotypes (clusters 18, 23, 4, and 8) that are normally abundant in lung tissues from healthy individuals (Thome et al., 2014) and of additional subsets that are rare in the blood (clusters 24, 14, 5, and 22). Overall, our data further confirm that the tumor is largely devoid of subpopulations capable to kill in an effective way, while supporting features of T cell exhaustion since the early disease stage. These results are in line with the recent characterization of a cohort of 73 renal cell carcinoma patients and a smaller cohort of 18 NSCLC patients by CyTOF, where patients at different stages of the disease were characterized by similar immune signatures at the level of T cells, including that of T cell exhaustion (Chevrier et al., 2017; Lavin et al., 2017).

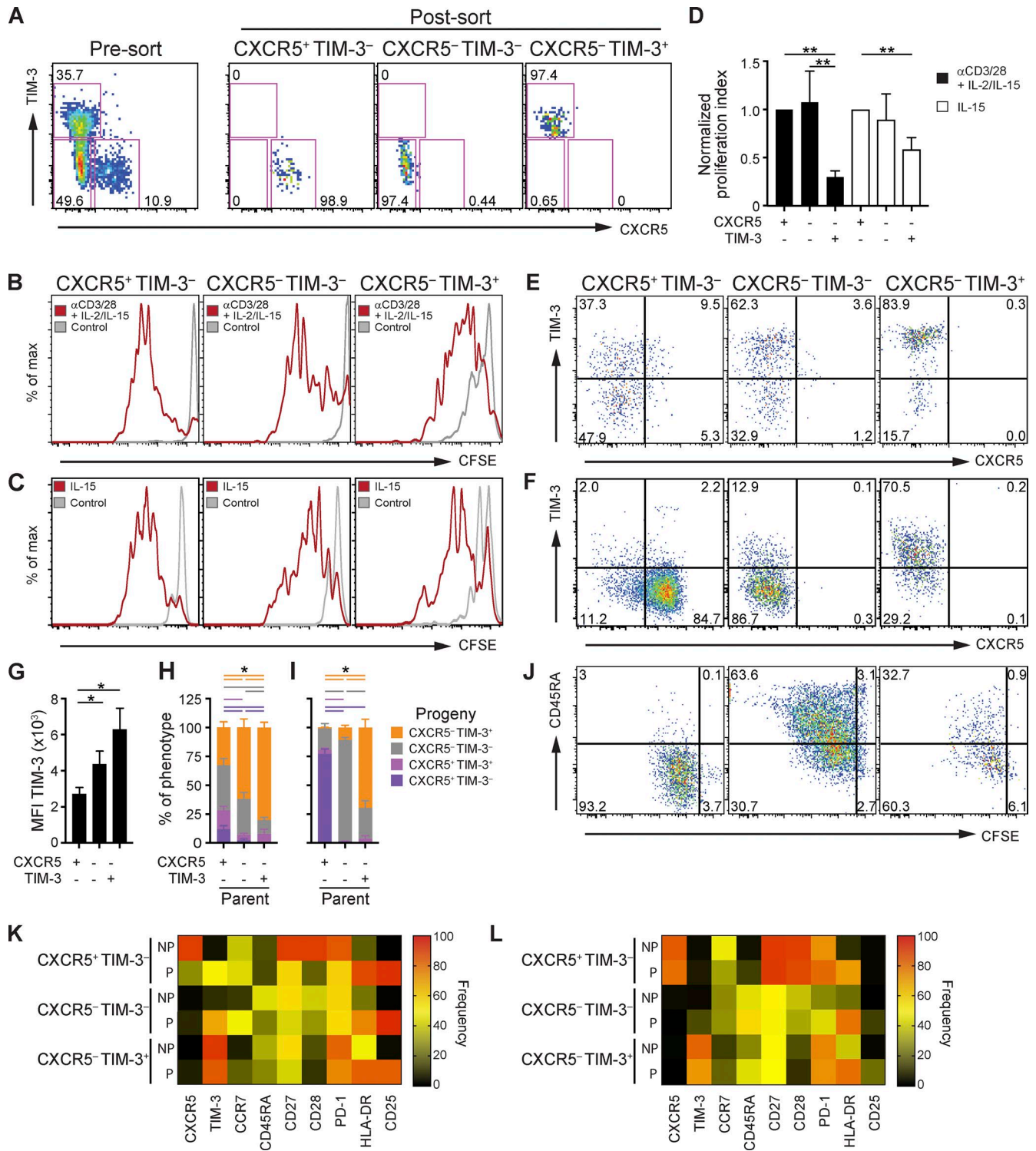
The CXCR5<sup>+</sup> population described here is the human equivalent of the partially exhausted, CXCR5<sup>+</sup> PD-1<sup>int</sup> CD8<sup>+</sup> T cells expressing the transcription factor TCF-1, that, in chronically infected mice, are preferentially responsible for viral control following PD-1 blockade (He et al., 2016; Im et al., 2016; Leong et al., 2016; Utzschneider et al., 2016; Wu et al., 2016). Although, we could not directly address this enhanced responsiveness to PD-1 blockade by the tumor-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells due to practical reasons (i.e., low number of cells obtained after sorting), the flow cytometry and scRNA-seq data support this concept. By scRNA-seq of melanoma TILs and subsequent validation by high-dimensional flow cytometry and functional assays in lung TILs, we confirmed the follicular identity, a dif-

ferentiation pattern similar to that of long-lived memory T cells, and partial exhaustion of single CXCR5<sup>+</sup> CD8<sup>+</sup> T cells and additionally revealed the simultaneous coexistence of stem cell-like behavior, polyfunctional effector molecule production, and cytotoxic potential. This is important because immune activation will result in immediate effector functions that can persist in the long term. Stem-like properties of T cells were shown to be beneficial in multiple instances, such as in the context of adoptive cell transfer where long-lived memory T cells are generally associated with superior anti-tumor responses (Gattinoni et al., 2005, 2011; Chapuis et al., 2013). Although CXCR5<sup>+</sup> CD8<sup>+</sup> T cells display a maturation phenotype that is largely different from that of canonical T<sub>SCM</sub> cells, i.e., CD45RA<sup>-</sup> CCR7<sup>-</sup>, effector memory-like with tissue-resident characteristics, they also share some T<sub>SCM</sub> properties, including the expression of CD27, CD28 and CXCR3 surface markers, and Eomes, but not T-bet transcription factors, self-renewal, and multipotency (Gattinoni et al., 2011; Lugli et al., 2013a; Fuertes Marraco et al., 2015; Oliveira et al., 2015; Pilipow et al., 2015; Akondy et al., 2017). Moreover, only CXCR5<sup>+</sup> CD8<sup>+</sup> TILs express TCF-1, a master regulator of persistence that is critical for the formation and maintenance of these cells in chronically infected mice. Despite acknowledging the limitations of in vitro assays to define a precursors-progeny relationship in humans, our data support the concept that a hierarchy of differentiation is also present at the tumor level in the context of exhaustion, albeit featuring different players, i.e., less differentiated CXCR5<sup>+</sup> TIM-3<sup>-</sup> stem-like CD8<sup>+</sup> T cells giving rise to CXCR5<sup>-</sup> and TIM-3<sup>+</sup> effector subsets upon triggering of the TCR.

Fine mapping of the expression of inhibitory and costimulatory receptors with single cell resolution by both flow cytometry and scRNA-seq indicated that PD-1 and TIGIT could be preferential targets of immunotherapy with checkpoint inhibitors, while CD27 with immunostimulating antibodies favor effector differentiation (Ahrends et al., 2016). Differently, blockade directed to LAG-3 and CTLA-4, which are present on CXCR5<sup>+</sup> CD8<sup>+</sup> T cells from chronically infected mice, but are poorly expressed on human CXCR5<sup>+</sup> CD8<sup>+</sup> TILs, would be relatively inefficient. Despite not being formally demonstrated due to the difficulty in identifying tumor antigens, the presence of traits of exhaustion and the accumulation exclusively within tumors suggest that CXCR5<sup>+</sup> CD8<sup>+</sup> TILs are tumor-specific. It is at the moment unclear why CXCR5<sup>+</sup> CD8<sup>+</sup> T cells mainly form/accumulate in the TME compared with the cancer-free lung. Their follicular cell identity and their positive correlation with the abundance of Tfh and B cells within the tumor suggest they localize at the border of B cell areas in tertiary lymphoid structures, which have been shown in multiple instances to be associated with enhanced effector responses and increased survival in lung cancer (Dieu-Nosjean et al., 2008).

In conclusion, CXCR5<sup>+</sup> CD8<sup>+</sup> TILs are the ideal cells to be relieved from suppression as they serve as precursors of more

stimulation. Numbers in the histograms indicate the percentage of cells identified by the gates. (C) Pie charts representing the proportion of CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> TILs expressing and producing different combinations of CD107a, IL-2, IFN $\gamma$ , and TNF after PMA/ionomycin stimulation ( $n = 7$ ). Frequencies were corrected by background subtraction as determined in unstimulated controls. \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; partial permutation tests, using SPICE software. (D) Frequency of CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> TILs expressing and producing different combinations of CD107a, IL-2, IFN $\gamma$ , and TNF after PMA/ionomycin stimulation ( $n = 7$ ). Data are given as mean  $\pm$  SEM. \*,  $P < 0.05$  versus CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> TILs; †,  $P < 0.05$  versus CXCR5<sup>-</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> TILs, paired Wilcoxon  $t$  test.



**Figure 5. Tumor-infiltrating CXCR5<sup>+</sup> CD8<sup>+</sup> T cells are endowed with enhanced stemness.** (A) Representative dot plots depicting the purity of sorted CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>+</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells from a NSCLC tumor sample. Numbers in the dot plots indicate the percentage of cells identified by the gates. (B–L) Sorted CFSE-stained CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>+</sup> TIM-3<sup>+</sup> CD8<sup>+</sup> T cells were stimulated with αCD3/CD28 beads in combination with IL-2 and IL-15 (10 ng/ml, each) for 6 d (*n* = 9; all performed in independent experiments) to determine effector differentiation or with IL-15 (50 ng/ml) for 11 d (*n* = 7; performed in six independent experiments) to determine their self-renewal capacity. As a non-proliferating negative control, cells were kept in 1 ng/ml IL-15. (B and C) Representative histograms showing CFSE dilution in sorted subsets after αCD3/CD28 + IL-2/IL-15 (B) or IL-15 (C) stimulation. (D) Summary of the proliferation index, calculated as MFI non-proliferating fraction / MFI proliferating fraction × % cells with diluted CFSE, after αCD3/CD28 + IL-2/IL-15 (*n* = 9) or IL-15 (*n* = 7) stimulation normalized for the value obtained for the CXCR5<sup>+</sup> TIM-3<sup>-</sup> subset. Data are given as mean ± SEM; \*\*, *P* < 0.01; paired Wilcoxon *t* test. (E and F) Representative dot plots depicting the expression of CXCR5 and TIM-3 by the proliferated cells of the different sorted subsets after αCD3/CD28 + IL-2/IL-15 (E) or IL-15 (F) stimulation. Numbers in the dot plots indicate the percentage of cells identified by the gates. (G) TIM-3 MFI of the proliferating cells (CFSE diluted) of the different sorted subsets after αCD3/CD28 + IL-2/IL-15 stimulation (*n* = 9). Data are given as mean ± SEM; paired

differentiated effectors with enhanced functionality and persistence. However, they tend to be lost with disease progression. More differentiated CXCR5<sup>-</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells retain some degree of functionality following stimulation despite their decreased self-renewal and degranulation potential and may serve as precursors of effector cells upon PD-1 blockade in advanced disease. However, this phenotype is poorly represented in the murine antigen-specific CD8<sup>+</sup> T cell pool generated following chronic infection (He et al., 2016; Im et al., 2016; Leong et al., 2016; Utzschneider et al., 2016; Wu et al., 2016), thus raising the question whether tumor-specific CD8<sup>+</sup> T cells are contained in the CXCR5<sup>-</sup> TIM-3<sup>-</sup> subset and can be primed intratumorally. Future studies should address whether tumor antigen-specific stem-like cells also exist in lymphoid tissues, and whether they are an additional reservoir of anti-tumor T cell responses.

## Materials and methods

### Study design

All human experiments were approved by the Humanitas Clinical and Research Center Internal Review Board (approval nos. 1501, 1021, and 168/18). All patients provided written informed consent. In this research, we included 53 patients with diagnosis of NSCLC with clinical stage I or II who underwent a lung major resection (anatomical segmentectomy, lobectomy, or pneumonectomy) plus radical lymph node dissection with radical intent at the Humanitas Cancer Center. Tumor ( $n = 53$ ) and adjacent cancer-free lung tissue samples from the same surgical specimen ( $n = 45$ ) were obtained in the operating room in a sterile field by the surgeon while blood sample collection (10 ml;  $n = 22$ ) were obtained by venipuncture at induction of anesthesia. Human intestinal specimens were obtained from patients with CRC undergoing resection surgery and liver tumor tissues from patients undergoing hepatectomy for hepatic metastatic disease from CRC. None of patients received preoperative chemotherapy or radiotherapy. Details on the patients' characteristics are indicated in Table 1. Information on the pathological stage (pStage), determined by an institutional pathologist (some patients were restaged as III following examination of the tumor), was available for all patients, while results of the preoperative fluorodeoxyglucose (FDG) PET-computed tomography (CT) scan were available for 33 patients. PET scans were acquired ~60 min after FDG administration in patients fasting for at least 6 h. Whole body images were obtained from the base of the skull to mid-thigh by means of integrated PET-CT tomographs. Reconstructed images were displayed on a GE ADW4.6 workstation (GE Healthcare). SUVmax was determined as the highest pixel value within the tumor mass that was identified with three-dimensional volumes of interest, as previously described (Lopci et al., 2016). SUVmax high and low groups (Fig. 6) were defined as those patients whose

tumors have values that are above or below the median of the distribution, respectively.

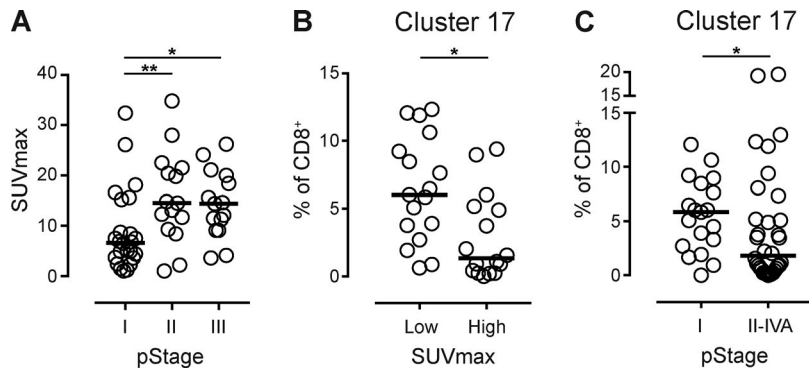
### Sample collection and processing

Blood samples were collected in Vacutainer EDTA Tubes (BD), while the tumor and normal tissue samples were collected in RPMI-1640 medium supplemented with 10% FBS (Sigma-Aldrich), 1% penicillin-streptomycin, and 1% Ultra-glutamine (both from Lonza), hereafter referred to as R10. Peripheral blood mononuclear cells were isolated from the blood samples or from healthy donors (buffy coats) by Lympholyte-H (Cedarlane) density separation and frozen in liquid nitrogen according to standard procedures, as described previously (Lugli et al., 2013b). The tumor and normal tissue samples were processed into single cell suspensions by mechanical disaggregation. The samples were cut into small fragments in a size of ~1 mm<sup>3</sup>, after which further disaggregation was performed using gentleMACS dissociator (Miltenyi Biotec). Subsequently, samples were passed through a 70- $\mu$ m cell strainer (Falcon) and washed with physiological saline solution (NaCl 0.9%; Baxter). After centrifugation (300 g, 15 min), erythrocytes lysis was performed by incubating the cell suspension for 5 min in ACK lysis buffer (made in house; 155 mM NH<sub>4</sub>Cl, 10 mM KHCO<sub>3</sub>, and 0.1 mM EDTA-2Na [all from Sigma-Aldrich] diluted in H<sub>2</sub>O), and washed again with saline. CRC samples were mechanically minced in small pieces, washed in HBSS without Ca<sup>2+</sup> and Mg<sup>2+</sup> (Lonza) supplemented with 1% penicillin/streptomycin/amphotericin (Invitrogen) and followed enzymatic digestion with 1.5  $\mu$ g/ml of collagenase IV (Life Technologies), 20  $\mu$ g/ml of hyaluronidase, and 50  $\mu$ g/ml of DNase (both from Sigma-Aldrich) for 2 h at 37°C/5% CO<sub>2</sub>. Lymphocytes were separated by 30%/70% Percoll gradient centrifugation. Liver metastasis were cut in small pieces followed by enzymatic digestion in gentleMACS dissociator with 2 mg/ml collagenase D (Roche Diagnostic) for 45 min at 37°C/5% CO<sub>2</sub>. Lymphocytes were isolated by 30%/70% Percoll density gradient. The cells were either frozen in liquid nitrogen according to standard procedures or, in case of sorting by flow cytometry, NSCLC leukocytes were further purified by centrifugation (790 g for 30 min without brake) on a 40%/70% Percoll (Sigma-Aldrich) gradient. After isolation of the interphase between the two layers, the cells were washed with HBSS without Ca<sup>2+</sup> or Mg<sup>2+</sup> (Lonza) and stained with fluorescently conjugated monoclonal antibodies (mAbs).

### Polychromatic flow cytometry for the detection of surface and intracellular antigens

For the ex vivo immunophenotyping experiments, frozen samples were used and were thawed in R10 supplemented with 20  $\mu$ g/ml DNase I from bovine pancreas (Sigma-Aldrich). After extensive washing with PBS (Sigma-Aldrich), the cells were stained immediately with the Zombie Aqua Fixable Viability kit (BioLegend) for 15 min at room temperature. Then, the cells were

Wilcoxon *t* test. **(H and I)** Percentage of CFSE diluted cells with a given phenotype following stimulation with  $\alpha$ CD3/CD28 + IL-2/IL-15 (H;  $n = 9$ ) or IL-15 (I;  $n = 7$ ). Data are given as mean  $\pm$  SEM. Horizontal bars indicate significant comparisons. Bar colors refer to progenies with specific phenotypes as indicated in the legend. Paired Wilcoxon *t* test. **(J)** Representative dot plots showing CD45RA expression and CFSE dilution by IL-15-stimulated cells. Numbers in the dot plots indicate the percentage of cells identified by the gates. **(K and L)** Heat maps showing the median frequency of the indicated markers by the different  $\alpha$ CD3/CD28 + IL-2/IL-15 (K;  $n = 9$ ) or IL-15 (L;  $n = 7$ ) stimulated CD8<sup>+</sup> T subsets. NP, non-proliferating cells from the negative control; P, proliferating cells (CFSE diluted) after stimulation. Comparisons that are statistically significant are indicated in Fig. S4. \*,  $P < 0.05$ .



**Figure 6. Loss of CXCR5<sup>+</sup> CD8<sup>+</sup> TILs with lung cancer disease progression.** (A) SUVmax values among NSCLC patients with pStage I ( $n = 24$ ), II ( $n = 15$ ), or III ( $n = 15$ ); \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; Mann-Whitney  $t$  test. (B) Based on the SUVmax values, samples were divided into SUVmax low and high with the median SUVmax value as cut-off. Frequency of Phenograph cluster 17 (CD69<sup>+</sup>CXCR5<sup>+</sup>) is depicted in the SUVmax low ( $n = 17$ ) and high ( $n = 16$ ) groups. \*,  $P < 0.05$ ; Mann-Whitney  $t$  test. (C) Frequency of Phenograph cluster 17 (CD69<sup>+</sup>CXCR5<sup>+</sup>) in tumor samples from patients with NSCLC pStage I ( $n = 19$ ) versus II-IVA ( $n = 34$ ). \*,  $P < 0.05$ ; Mann-Whitney  $t$  test. Data were obtained in six independent experiments. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; Mann-Whitney  $t$  test.

washed and stained with the combination of mAbs purchased from either BD Biosciences, BioLegend, or eBioscience, as listed in Table S1. mAbs were previously titrated to define the optimal concentration, as described (Lugli et al., 2013b). Chemokine receptors were stained for 20 min at 37°C, while all other surface markers (except CD3) were stained for 20 min at room temperature. Intracellular detection of CD3, Ki-67, and transcription factors was performed following fixation of cells with the FoxP3/transcription factor staining buffer set (eBioscience) according to manufacturer's instructions and by incubating with specific mAbs for 30 min at 4°C.

In parallel, all samples were stained with a panel of mAbs directed to reveal the major immune populations, such as T cells (CD3<sup>+</sup>), B cells (CD19<sup>+</sup>), myeloid (CD11b<sup>+</sup>), and NK subsets (CD16, CD56, and NKp46). Additional panels used for further characterizing the CXCR5<sup>+</sup> CD8<sup>+</sup> T cell subset are listed in Table S1. Stainings were performed as described above, except for intracellular staining of TCF-1, for which cells were fixed with Transcription Factor Buffer set (BD Biosciences) according to the manufacturer's instructions. For detection of cytokines, cells were fixed with BD Cytofix/Cytoperm Fixation/Permeabilization Solution kit (BD Biosciences) according to the manufacturer's instructions. CD107a-specific antibody was added in the culture medium during the stimulation.

All data were acquired on a FACS Symphony A5 flow cytometer (BD Biosciences) equipped with five lasers (UV, 350 nm; violet, 405 nm; blue, 488; yellow/green, 561 nm; red, 640 nm; all tuned at 100 mW, except UV tuned at 60 mW) and capable to detect 30 parameters. Flow cytometry data were compensated in FlowJo by using single stained controls (BD Compbeads incubated with fluorescently conjugated antibodies), as described (Lugli et al., 2017).

#### High-dimensional flow cytometry data analysis.

Flow Cytometry Standard (FCS) 3.0 files were imported into FlowJo software version 9, and analyzed by standard gating to remove aggregates and dead cells, and identify CD3<sup>+</sup> CD8<sup>+</sup> T cells. 3,000 CD8<sup>+</sup> T cells per sample were subsequently imported in FlowJo version 10, biexponentially transformed, and exported for further analysis in R (version 3.3.3) by a custom-made script that makes use of Bioconductor libraries and R statistical packages. Blood, normal lung tissue, and tumor samples were labeled with a unique computational barcode for further identification and concatenated in a single matrix by using the "cytof\_exprsMerge

function" (Chen et al., 2016). Data were analyzed using the Phenograph algorithm coded in the cytofkit package (version 1.6.5; Chen et al., 2016). Parameter K, indicating the number of nearest neighbors identified in the first iteration of the algorithm, was set at 60. Phenograph clusters were visualized using tSNE (cytofkit package, with default parameters). Clusters representing <0.5% were disregarded in subsequent analysis. The data were then reorganized and saved as new FCS files, one for each cluster, that were further analyzed in FlowJo to determine the frequency of positive cells for each marker and the corresponding MFI. These values were multiplied to derive the integrated MFI (iMFI, rescaled to values from 0 to 100; Fig. 1 D). The heat map, showing the iMFI of each marker per cluster, and the subsequent metaclustering was performed using the gplots R package. Hierarchical metaclustering of all samples, based on the frequency of Phenograph clusters (Fig. 1 C), was performed in R based on the Euclidean distance and Ward-linkage.

#### scRNA-seq data processing and in silico sorting of CD8<sup>+</sup> T cell subsets

Normalized scRNA-seq counts were downloaded from Gene Expression Omnibus dataset (GSE72056). Analysis was restricted to the cells labeled as "T cells," as previously defined by Tirosh et al. (2016). T cells were separated into CD4<sup>+</sup> and CD8<sup>+</sup> based on the normalized expression levels (E) of CD4 (E > 4), and CD8 (average of CD8A and CD8B, E > 4). The correlation between the expression profile of CXCR5 in the CD8<sup>+</sup> TILs ( $n = 912$ ), and that of all other genes present in the matrix was calculated using the "cor.test" function in R with Pearson as correlation methods. The CD8<sup>+</sup> TILs ( $n = 912$ ) were further divided in three groups based on CXCR5 and HAVCR2 (encoding TIM-3) expression levels. Specifically, we identified T cells that were CXCR5<sup>+</sup> (E > 0; including all cells irrespectively of HAVCR2 expression, which was negligible in this subset; unpublished data), CXCR5<sup>-</sup> TIM-3<sup>+</sup> (E = 0 and E ≥ 5.930667, respectively), and CXCR5<sup>-</sup> TIM-3<sup>-</sup> (E = 0 and E ≤ 2.1527, respectively). Thresholds of HAVCR2 (TIM-3) expression were determined by calculating the 33.3 and 66.7 percentile of the HAVCR2 mRNA distribution. Therefore, if HAVCR2 mRNA levels were <33.3 percentile, cells were labeled as CXCR5<sup>-</sup> TIM-3<sup>-</sup>; otherwise, if >66.7 percentile, as CXCR5<sup>-</sup> TIM-3<sup>+</sup>. Differentially expressed genes in each pairwise comparison of CD8<sup>+</sup> T cell subsets were determined by the "FindMarkers" function coded in the Seurat R package (version 2.1.0; Satija et al., 2015) with default parameters except for the

“logfc.threshold” that was set equal to 0. Subset-specific genes were defined as up-regulated in the pairwise comparisons between a cell subset and the other two. In this way we obtained 555, 46, and 1 specific gene for the CXCR5<sup>+</sup>, CXCR5<sup>-</sup> TIM-3<sup>+</sup>, and CXCR5<sup>-</sup> TIM-3<sup>-</sup> populations, respectively. The heat map shown in Fig. 3 B was generated by using the “DoHeatmap” function coded in the Seurat R package (for the CXCR5<sup>+</sup> subset, only the first 100 differentially expressed genes with the highest fold change resulting from the comparison with the CXCR5<sup>-</sup> TIM-3<sup>+</sup> or CXCR5<sup>-</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells are shown). Violin plots, depicting the gene expression probability distributions across subsets, were obtained with the Seurat R package.

### Microarray data analysis

Microarray probe fluorescence signals of the publicly available dataset GSE84105 (sample IDs used for this analysis: GSM2227309, GSM2227310, GSM2227311, GSM2227312, GSM2227313, and GSM2227314) were converted to expression values using robust multiarray average procedure RMA (Irizarry et al., 2003) of Bioconductor affy package. Specifically, fluorescence intensities were background adjusted, normalized using quantile normalization and log<sub>2</sub> expression values for a total of 18,075 custom probe sets calculated using median polish summarization and custom chip definition files for Affymetrix Mouse Genome 430 2.0 arrays based on Entrez genes (Mouse4302\_Mm\_ENTREZG version 21.0.0). To identify genes that are differentially expressed, we compared the expression profiles of CXCR5<sup>+</sup> TIM-3<sup>-</sup> versus CXCR5<sup>-</sup> TIM-3<sup>+</sup> samples using limma algorithm coded in the same R package (Ritchie et al., 2015). By setting the false discovery rate and fold change thresholds at 0.05 and 2, respectively, we obtained 246 differentially expressed genes. Mouse Entrez IDs were converted in the corresponding human homologous genes using the HUGO Gene Nomenclature Committee database, and they were subsequently used for GSEA.

### GSEA

Over-representation analysis was performed using the GSEA tool described in Subramanian et al. (2005). GSEA was applied on the entire list of genes that compose the single cell RNaseq expression matrix (GSE72056). Genes were ranked based on their fold change calculated between groups of each pairwise comparison and analyzed by GSEA in preranked mode. We adopted the “classic” enrichment statistic, the recommended approach for RNA-sequencing data. Gene sets of interest (“GSE10239 NAI VE VS MEMORY CD8 TCELL UP,” “GSE21380 NON TFH VS GER MINAL CENTER TFH CD4 TCELL DN,” and “GSE10239 NAI VE VS DAY4.5 EFF CD8 TCELL DN”) were retrieved from the immunological signatures collection (c7.all.v6.1) of the Molecular Signatures Database.

### Proliferation assay

CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup>, pregated as CD8<sup>+</sup> Aqua<sup>+</sup>, were isolated from freshly minced tumors with a FACSAria cell sorter (BD Biosciences) and stained with CellTrace CFSE kit (final concentration: 2 μM; Thermo Fisher) according to the manufacturer's protocol. Subsequently, cells were plated in R10 in U-bottom 96-well plates (10,000 cells/well) and stim-

ulated with either αCD3/CD28 Dynabeads (Life Technologies; 2 beads/cell) in combination with IL-2 and IL-15 (Peprotech; 10 ng/ml each) or IL-15 alone (50 ng/ml) for 6 and 11 d at 37°C, respectively. T cells kept in low dose of IL-15 (1 ng/ml) served as non-proliferating, negative control. Cells were then stained with mAbs (Table S1) as described above. Proliferation index depicted in Fig. 5 D was calculated as follows: Proliferation index = MFI non-proliferating fraction / MFI proliferating fraction × % cells with diluted CFSE.

### Analysis of effector molecules production

Freshly flow cytometry-sorted CXCR5<sup>+</sup> TIM-3<sup>-</sup>, CXCR5<sup>-</sup> TIM-3<sup>-</sup>, and CXCR5<sup>-</sup> TIM-3<sup>+</sup> were plated in R10 in U-bottom 96-well plates (10,000 cells/well) and stimulated with 10 ng/ml PMA and 0.5 μg/ml Ionomycin (both from Sigma-Aldrich) for 3 h at 37°C in the presence of Golgiplug (BD Biosciences) according to the manufacturer's instructions. The antibody for the detection of degranulation marker CD107a was added to the culture medium. Cells were then stained with mAbs (Table S1) as described above.

### Statistical analysis

Statistical analyses were performed using GraphPad Prism version 7, unless specified otherwise. Significance of differences for the frequency of single Phenograph clusters among blood, normal tissue, and tumor samples was determined using two-way ANOVA with Bonferroni post-hoc test. To compare distributions of manually gated subsets significance was determined by paired Wilcoxon *t* test, unless otherwise specified in the figure legends. Correlation of data were determined by calculating the Pearson correlation coefficient. P values <0.05 were considered statistically significant.

### Online supplemental material

Fig. S1 shows the phenotypic analysis of CXCR5<sup>+</sup> CD8<sup>+</sup> T cells. Fig. S2 shows that bona fide CD8<sup>+</sup> T<sub>SCM</sub> cells are absent in lung tumors and adjacent tissues. Fig. S3 shows that the gene signature of CXCR5<sup>+</sup> CD8<sup>+</sup> T cells resembles that of both CD4<sup>+</sup> T<sub>fh</sub> cells and memory CD8<sup>+</sup> T cells. Fig. S4 shows that proliferating CXCR5<sup>+</sup> TIM-3<sup>-</sup> CD8<sup>+</sup> T cells are self-renewing and multipotent. Table S1 lists all of the antibodies for flow cytometry used in this study.

### Acknowledgments

We wish to thank the laboratory of Massimiliano Pagani (INGM, Italy) for helping with single cell RNA-sequencing data analysis, Gianluca Rotta, and Jens Fleischer (BD Biosciences) for helping with FACSSymphony A5 instrument setup; Matteo Donadon and Guido Torzilli (Liver Unit, Humanitas) for colorectal liver metastasis samples; Antonino Spinelli (Colon and Rectal Unit, Humanitas) for colorectal carcinoma samples; Paola Spaggiari and Massimo Roncalli (Pathology Unit, Humanitas) for histological analysis; and Prof. Alberto Mantovani (Humanitas University) and the members of the Laboratory of Translational Immunology for critical discussion.

This work has been supported by grants from the Associazione Italiana per la Ricerca sul Cancro (AIRC; grants IG20607 to E. Lugli, IG14687 to D. Mavilio, and MFAG189323 to E. Lopci), Bris-

tol Myers Squibb (preclinical grant no. CA209-9YM to E. Lugli), the Italian Ministry of Health (grant PE-2016-02363915 to D. Mavilio), and the Humanitas Clinical and Research Center grant to E. Lugli and D. Mavilio. J. Brummelman is a recipient of the “Fondo di beneficenza Intesa San Paolo” fellowship from AIRC. E.M.C. Mazza and P. Novellis are recipients of the 2017 Fondazione Umberto Veronesi postdoctoral fellowship. E. Lugli is an International Society for the Advancement of Cytometry Marylou Ingram scholar. Purchase of the BD FACSSymphony A5 has been in part defrayed by a grant from Italian Ministry of Health (agreement no. 82/2015).

The Laboratory of Translational Immunology receives reagents in kind from BD Biosciences as part of a collaborative research agreement. The authors have no additional financial interests.

Author contributions: E. Lugli and G. Veronesi wrote the protocol for ethical evaluation. M. Alloisio, P. Novellis, and G. Veronesi performed the surgeries. E. Lopci performed and generated the FDG PET-CT scans. J. Brummelman and G. Alvisi processed the patients’ samples. J. Brummelman and E. Lugli designed the experiments. J. Brummelman and G. Alvisi carried out the experiments. F.S. Colombo performed the quality check of the FACSSymphony A5 and assisted with sort experiments. J. Brummelman and G. Alvisi analyzed the flow cytometry data. E.M.C. Mazza performed the tSNE and Phenograph analysis and processed the single cell RNA-sequencing data with assistance of F. Ferrari and A. Grilli. D. Mavilio and J. Mikulak performed analysis on colorectal and liver metastasis tumors. J. Brummelman and E. Lugli interpreted the data and wrote the manuscript with input of all authors.

Submitted: 11 April 2018

Revised: 6 July 2018

Accepted: 16 August 2018

## References

Ahrends, T., N. Bąbała, Y. Xiao, H. Yagita, H. van Eenennaam, and J. Borst. 2016. CD27 Agonism Plus PD-1 Blockade Recapitulates CD4+ T-cell Help in Therapeutic Anticancer Vaccination. *Cancer Res.* 76:2921–2931. <https://doi.org/10.1158/0008-5472.CAN-15-3130>

Akondy, R.S., M. Fitch, S. Edepuganti, S. Yang, H.T. Kissick, K.W. Li, B.A. Youngblood, H.A. Abdelsamed, D.J. McGuire, K.W. Cohen, et al. 2017. Origin and differentiation of human memory CD8 T cells after vaccination. *Nature.* 552:362–367. <https://doi.org/10.1038/nature24633>

Amir, E.-A.D., K.L. Davis, M.D. Tadmor, E.F. Simonds, J.H. Levine, S.C. Bendall, D.K. Shenfeld, S. Krishnaswamy, G.P. Nolan, and D. Pe’er. 2013. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nat. Biotechnol.* 31:545–552. <https://doi.org/10.1038/nbt.2594>

Bindea, G., B. Mlecnik, M. Tosolini, A. Kirilovsky, M. Waldner, A.C. Obenauf, H. Angell, T. Fredriksen, L. Lafontaine, A. Berger, et al. 2013. Spatiotemporal 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>

Chapuis, A.G., G.B. Ragnarsson, H.N. Nguyen, C.N. Chaney, J.S. Pufnock, T.M. Schmitt, N. Duerkopp, I.M. Roberts, G.L. Pogosov, W.Y. Ho, et al. 2013. Transferred WT1-reactive CD8+ T cells can mediate antileukemic activity and persist in post-transplant patients. *Sci. Transl. Med.* 5:174ra27. <https://doi.org/10.1126/scitranslmed.3004916>

Chattopadhyay, P.K., M.R. Betts, D.A. Price, E. Gostick, H. Horton, M. Roederer, and S.C. De Rosa. 2009. The cytolytic enzymes granzyme A, gran-

zyme B, and perforin: expression patterns, cell distribution, and their relationship to cell maturity and bright CD57 expression. *J. Leukoc. Biol.* 85:88–97. <https://doi.org/10.1189/jlb.0208107>

Chen, H., M.C. Lau, M.T. Wong, E.W. Newell, M. Poidinger, and J. Chen. 2016. Cytokit: A Bioconductor Package for an Integrated Mass Cytometry Data Analysis Pipeline. *PLOS Comput. Biol.* 12:e1005112. <https://doi.org/10.1371/journal.pcbi.1005112>

Chevrier, S., J.H. Levine, V.R.T. Zanotelli, K. Silina, D. Schulz, M. Bacac, C.H. Ries, L. Ailles, M.A.S. Jewett, H. Moch, et al. 2017. An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell.* 169:736–749.e18. <https://doi.org/10.1016/j.cell.2017.04.016>

Croci, D.O., M.F. Zacarias Fluck, M.J. Rico, P. Matar, G.A. Rabinovich, and O.G. Scharovsky. 2007. Dynamic cross-talk between tumor and immune cells in orchestrating the immunosuppressive network at the tumor micro-environment. *Cancer Immunol. Immunother.* 56:1687–1700. <https://doi.org/10.1007/s00262-007-0343-y>

Darrah, P.A., D.T. Patel, P.M. De Luca, R.W. Lindsay, D.F. Davey, B.J. Flynn, S.T. Hoff, P. Andersen, S.G. Reed, S.L. Morris, et al. 2007. Multifunctional TH1 cells define a correlate of vaccine-mediated protection against Leishmania major. *Nat. Med.* 13:843–850. <https://doi.org/10.1038/nm1592>

Dieu-Nosjean, M.C., M. Antoine, C. Danel, D. Heudes, M. Wislez, V. Poulot, N. Rabbe, L. Laurans, E. Tartour, L. de Chaisemartin, et al. 2008. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *J. Clin. Oncol.* 26:4410–4417. <https://doi.org/10.1200/JCO.2007.15.0284>

Farber, D.L., N.A. Yudanin, and N.P. Restifo. 2014. Human memory T cells: generation, compartmentalization and homeostasis. *Nat. Rev. Immunol.* 14:24–35. <https://doi.org/10.1038/nri3567>

Freeman, G.J., A.J. Long, Y. Iwai, K. Bourque, T. Chernova, H. Nishimura, L.J. Fitz, N. Malenkovich, T. Okazaki, M.C. Byrne, et al. 2000. Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *J. Exp. Med.* 192:1027–1034. <https://doi.org/10.1084/jem.192.7.1027>

Fuertes Marraco, S.A., C. Soneson, L. Cagnon, P.O. Gannon, M. Allard, S. Abed Maillard, N. Montandon, N. Rufer, S. Waldvogel, M. Delorenzi, and D.E. Speiser. 2015. Long-lasting stem cell-like memory CD8+ T cells with a naïve-like profile upon yellow fever vaccination. *Sci. Transl. Med.* 7:282ra48. <https://doi.org/10.1126/scitranslmed.aaa3700>

Gajewski, T.F., H. Schreiber, and Y.X. Fu. 2013. Innate and adaptive immune cells in the tumor microenvironment. *Nat. Immunol.* 14:1014–1022. <https://doi.org/10.1038/ni.2703>

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>

Galon, J., H.K. Angell, D. Bedognetti, and F.M. Marincola. 2013. The continuum of cancer immunosurveillance: prognostic, predictive, and mechanistic signatures. *Immunity.* 39:11–26. <https://doi.org/10.1016/j.immuni.2013.07.008>

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>

Gattinoni, L., C.A. Klebanoff, D.C. Palmer, C. Wrzesinski, K. Kerstann, Z. Yu, S.E. Finkelstein, M.R. Theoret, S.A. Rosenberg, and N.P. Restifo. 2005. Acquisition of full effector function in vitro paradoxically impairs the in vivo antitumor efficacy of adoptively transferred CD8+ T cells. *J. Clin. Invest.* 115:1616–1626. <https://doi.org/10.1172/JCI24480>

Gattinoni, L., E. Lugli, Y. Ji, Z. Pos, C.M. Paulos, M.F. Quigley, J.R. Almeida, E. Gostick, Z. Yu, C. Carpenito, et al. 2011. A human memory T cell subset with stem cell-like properties. *Nat. Med.* 17:1290–1297. <https://doi.org/10.1038/nm.2446>

Hamamishi, J., M. Mandai, M. Iwasaki, T. Okazaki, Y. Tanaka, K. Yamaguchi, T. Higuchi, H. Yagi, K. Takakura, N. Minato, et al. 2007. Programmed cell death 1 ligand 1 and tumor-infiltrating CD8+ T lymphocytes are prognostic factors of human ovarian cancer. *Proc. Natl. Acad. Sci. USA.* 104:3360–3365. <https://doi.org/10.1073/pnas.0611533104>

He, R., S. Hou, C. Liu, A. Zhang, Q. Bai, M. Han, Y. Yang, G. Wei, T. Shen, X. Yang, et al. 2016. Follicular CXCR5-expressing CD8(+) T cells curtail chronic viral infection. *Nature.* 537:412–416. <https://doi.org/10.1038/nature19317>

Hodi, F.S., S.J. O’Day, D.F. McDermott, R.W. Weber, J.A. Sosman, J.B. Haanen, R. Gonzalez, C. Robert, D. Schadendorf, J.C. Hassel, et al. 2010. Improved

- survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* 363:711–723. <https://doi.org/10.1056/NEJMoa1003466>
- 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. Nastij, 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>
- Irizarry, R.A., B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, and T.P. Speed. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 4:249–264. <https://doi.org/10.1093/biostatistics/4.2.249>
- Jacquelot, N., M.P. Roberti, D.P. Enot, S. Rusakiewicz, N. Ternès, S. Jegou, D.M. Woods, A.L. Sodrè, M. Hansen, Y. Meirou, et al. 2017. Predictors of responses to immune checkpoint blockade in advanced melanoma. *Nat. Commun.* 8:592. <https://doi.org/10.1038/s41467-017-00608-2>
- Krieg, C., M. Nowicka, S. Guglietta, S. Schindler, F.J. Hartmann, L.M. Weber, R. Dummer, M.D. Robinson, M.P. Levesque, and B. Becher. 2018. High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy. *Nat. Med.* 24:144–153. <https://doi.org/10.1038/nm.4466>
- Lavin, Y., S. Kobayashi, A. Leader, E.-A.D. Amir, N. Elefant, C. Bigenwald, R. Remark, R. Sweeney, C.D. Becker, J.H. Levine, et al. 2017. Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell*. 169:750–765.e17. <https://doi.org/10.1016/j.cell.2017.04.014>
- Leach, D.R., M.F. Krummel, and J.P. Allison. 1996. Enhancement of antitumor immunity by CTLA-4 blockade. *Science*. 271:1734–1736. <https://doi.org/10.1126/science.271.5256.1734>
- Leong, Y.A., Y. Chen, H.S. Ong, D. Wu, K. Man, C. Deleage, M. Minnich, B.J. Meckiff, Y. Wei, Z. Hou, et al. 2016. CXCR5(+) follicular cytotoxic T cells control viral infection in B cell follicles. *Nat. Immunol.* 17:1187–1196. <https://doi.org/10.1038/ni.3543>
- Levine, J.H., E.F. Simonds, S.C. Bendall, K.L. Davis, E.-A.D. Amir, M.D. Tadmor, O. Litvin, H.G. Fienberg, A. Jager, E.R. Zunder, et al. 2015. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell*. 162:184–197. <https://doi.org/10.1016/j.cell.2015.05.047>
- Liu, J., M. Dong, X. Sun, W. Li, L. Xing, and J. Yu. 2016. Prognostic Value of 18F-FDG PET/CT in Surgical Non-Small Cell Lung Cancer: A Meta-Analysis. *PLoS One*. 11:e0146195. <https://doi.org/10.1371/journal.pone.0146195>
- Lopci, E., L. Toschi, F. Grizzi, D. Rahal, L. Olivari, G.F. Castino, S. Marchetti, N. Cortese, D. Qehajaj, D. Pistillo, et al. 2016. Correlation of metabolic information on FDG-PET with tissue expression of immune markers in patients with non-small cell lung cancer (NSCLC) who are candidates for upfront surgery. *Eur. J. Nucl. Med. Mol. Imaging*. 43:1954–1961. <https://doi.org/10.1007/s00259-016-3425-2>
- Lugli, E., M.H. Dominguez, L. Gattinoni, P.K. Chattopadhyay, D.L. Bolton, K. Song, N.R. Klatt, J.M. Brenchley, M. Vaccari, E. Gostick, et al. 2013a. Superior T memory stem cell persistence supports long-lived T cell memory. *J. Clin. Invest.* 123:594–599.
- Lugli, E., L. Gattinoni, A. Roberto, D. Mavilio, D.A. Price, N.P. Restifo, and M. Roederer. 2013b. Identification, isolation and in vitro expansion of human and nonhuman primate T stem cell memory cells. *Nat. Protoc.* 8:33–42. <https://doi.org/10.1038/nprot.2012.143>
- Lugli, E., V. Zanon, D. Mavilio, and A. Roberto. 2017. FACS Analysis of Memory T Lymphocytes. *Methods Mol. Biol.* 1514:31–47. [https://doi.org/10.1007/978-1-4939-6548-9\\_3](https://doi.org/10.1007/978-1-4939-6548-9_3)
- Mahnke, Y.D., T.M. Brodie, F. Sallusto, M. Roederer, and E. Lugli. 2013. The who's who of T-cell differentiation: human memory T-cell subsets. *Eur. J. Immunol.* 43:2797–2809. <https://doi.org/10.1002/eji.201343751>
- Oliveira, G., E. Ruggiero, M.T. Stanghellini, N. Cieri, M. D'Agostino, R. Fronza, C. Lulay, F. Dionisio, S. Mastaglio, R. Greco, et al. 2015. Tracking genetically engineered lymphocytes long-term reveals the dynamics of T cell immunological memory. *Sci. Transl. Med.* 7:317ra198. <https://doi.org/10.1126/scitranslmed.aac8265>
- Pagès, F., A. Berger, M. Camus, F. Sanchez-Cabo, A. Costes, R. Molitor, B. Mlecnik, A. Kirilovsky, M. Nilsson, D. Damotte, et al. 2005. Effector memory T cells, early metastasis, and survival in colorectal cancer. *N. Engl. J. Med.* 353:2654–2666. <https://doi.org/10.1056/NEJMoa051424>
- Pilipow, K., A. Roberto, M. Roederer, T.A. Waldmann, D. Mavilio, and E. Lugli. 2015. IL15 and T-cell Stemness in T-cell-Based Cancer Immunotherapy. *Cancer Res.* 75:5187–5193. <https://doi.org/10.1158/0008-5472.CAN-15-1498>
- Ritchie, M.E., B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, and G.K. Smyth. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. <https://doi.org/10.1093/nar/gkv007>
- Satija, R., J.A. Farrell, D. Gennert, A.F. Schier, and A. Regev. 2015. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* 33:495–502. <https://doi.org/10.1038/nbt.3192>
- Schumacher, K., W. Haensch, C. Röefzaad, and P.M. Schlag. 2001. Prognostic significance of activated CD8(+) T cell infiltrations within esophageal carcinomas. *Cancer Res.* 61:3932–3936.
- Sharma, P., and J.P. Allison. 2015. The future of immune checkpoint therapy. *Science*. 348:56–61. <https://doi.org/10.1126/science.aaa8172>
- Singer, M., C. Wang, L. Cong, N.D. Marjanovic, M.S. Kowalczyk, H. Zhang, J. Nyman, K. Sakuishi, S. Kurtulus, D. Gennert, et al. 2017. A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells. *Cell*. 171:1221–1223. <https://doi.org/10.1016/j.cell.2017.11.006>
- Subramanian, A., P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA*. 102:15545–15550. <https://doi.org/10.1073/pnas.0506580102>
- Takata, H., and M. Takiguchi. 2006. Three memory subsets of human CD8+ T cells differently expressing three cytolytic effector molecules. *J. Immunol.* 177:4330–4340. <https://doi.org/10.4049/jimmunol.177.7.4330>
- Takeuchi, S., B. Khiewwan, P.S. Fox, S.G. Swisher, M.O. Rohren, R.L. Bassett Jr., and H.A. Macapinlac. 2014. Impact of initial PET/CT staging in terms of clinical stage, management plan, and prognosis in 592 patients with non-small-cell lung cancer. *Eur. J. Nucl. Med. Mol. Imaging*. 41:906–914. <https://doi.org/10.1007/s00259-013-2672-8>
- Thome, J.J., N. Yudanin, Y. Ohmura, M. Kubota, B. Grinsphun, T. Sathaliyawala, T. Kato, H. Lerner, Y. Shen, and D.L. Farber. 2014. Spatial map of human T cell compartmentalization and maintenance over decades of life. *Cell*. 157:814–828. <https://doi.org/10.1016/j.cell.2014.10.026>
- Tirosh, I., B. Izar, S.M. Prakadan, M.H. Wadsworth II, D. Treacy, J.J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, et al. 2016. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 352:189–196. <https://doi.org/10.1126/science.aad0501>
- Topalian, S.L., F.S. Hodi, J.R. Brahmer, S.N. Gettinger, D.C. Smith, D.F. McDermott, J.D. Powderly, R.D. Carvajal, J.A. Sosman, M.B. Atkins, et al. 2012. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* 366:2443–2454. <https://doi.org/10.1056/NEJMoa1200690>
- 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>
- Wei, S.C., J.H. Levine, A.P. Cogdill, Y. Zhao, N.A.S. Anang, M.C. Andrews, P. Sharma, J. Wang, J.A. Wargo, D. Pe'er, and J.P. Allison. 2017. Distinct Cellular Mechanisms Underlie Anti-CTLA-4 and Anti-PD-1 Checkpoint Blockade. *Cell*. 170:1120–1133.e17. <https://doi.org/10.1016/j.cell.2017.07.024>
- Wherry, E.J. 2011. T cell exhaustion. *Nat. Immunol.* 12:492–499. <https://doi.org/10.1038/ni.2035>
- Wong, M.T., D.E. Ong, F.S. Lim, K.W. Teng, N. McGovern, S. Narayanan, W.Q. Ho, D. Cerny, H.K. Tan, R. Anicete, et al. 2016. A High-Dimensional Atlas of Human T Cell Diversity Reveals Tissue-Specific Trafficking and Cytokine Signatures. *Immunity*. 45:442–456. <https://doi.org/10.1016/j.immuni.2016.07.007>
- Wu, T., Y. Ji, E.A. Moseman, H.C. Xu, M. Mangani, M. Kirby, S.M. Anderson, R. Handon, E. Kenyon, A. Elkhouloun, et al. 2016. The TCF1-Bcl6 axis counteracts type I interferon to repress exhaustion and maintain T cell stemness. *Sci. Immunol.* 1:eaa18593. <https://doi.org/10.1126/sciimmunol.aai8593>
- Zhang, Y., G. Joe, E. Hexner, J. Zhu, and S.G. Emerson. 2005. Host-reactive CD8+ memory stem cells in graft-versus-host disease. *Nat. Med.* 11:1299–1305. <https://doi.org/10.1038/nm1326>
- Zheng, C., L. Zheng, J.K. Yoo, H. Guo, Y. Zhang, X. Guo, B. Kang, R. Hu, J.Y. Huang, Q. Zhang, et al. 2017. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell*. 169:1342–1356.e16. <https://doi.org/10.1016/j.cell.2017.05.035>