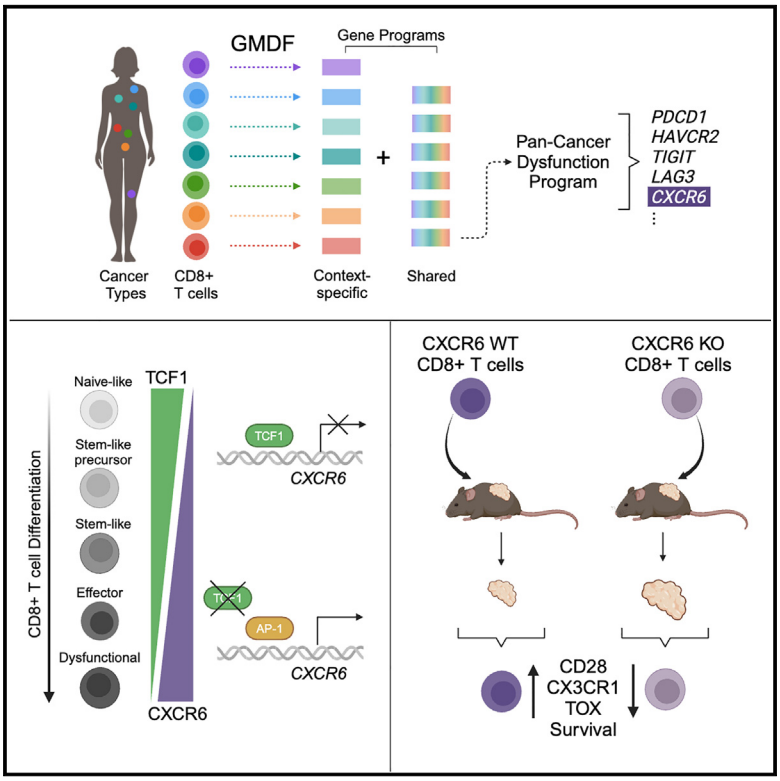
Cell Reports
Medicine

Article

Pan-cancer mapping of single CD8+ T cell profiles reveals a TCF1:CXCR6 axis regulating CD28 co-stimulation and anti-tumor immunity

Graphical abstract



Highlights

- Generalizable matrix decomposition framework identifies pan-cancer CD8+ T cell programs
- CXCR6 is a pan-cancer marker of chronically activated CD8+ T cells
- CXCR6 expression is driven by AP-1 and repressed by TCF1
- CXCR6 promotes CD8+ T cell survival and tumor control in part via promotion of CD28

Authors

Katherine Tooley, Livnat Jerby, Giulia Escobar, ..., Orit Rozenblatt-Rosen, Aviv Regev, Ana C. Anderson

Correspondence

ljerby@stanford.edu (L.J.), regev.aviv@gene.com (A.R.), acanderson@bwh.harvard.edu (A.C.A.)

In brief

Tooley et al. develop a computational method to identify pan-cancer CD8+ T cell gene programs and uncover a regulatory axis wherein TCF1 represses CXCR6 expression. They show CXCR6 expression is necessary for tumor control and promotes the survival of chronically activated anti-tumor T cells in part by promoting CD28 signaling.

Tooley et al., 2024, Cell Reports Medicine 5, 101640
July 16, 2024
https://doi.org/10.1016/j.xcrm.2024.101640

CellPress

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

Pan-cancer mapping of single CD8+ T cell profiles reveals a TCF1:CXCR6 axis regulating CD28 co-stimulation and anti-tumor immunity

Katherine Tooley,[1,2,8,10] Livnat Jerby,[3,4,5,10,*] Giulia Escobar,[1,8] S. Harsha Krovi,[1,8] Davide Mangani,[1,8] Gitanjali Dandekar,[1,8] Hanning Cheng,[1,8] Asaf Madi,[6] Ella Goldschmidt,[6] Conner Lambden,[1,8] Rajesh K. Krishnan,[1,8] Orit Rozenblatt-Rosen,[5,9] Aviv Regev,[5,7,9,11,*] and Ana C. Anderson[1,8,11,12,*]

[1]The Gene Lay Institute of Immunology and Inflammation of Brigham and Women's Hospital, Massachusetts General Hospital, and Harvard Medical School, Boston, MA, USA
[2]Division of Medical Sciences, Harvard Medical School, Boston, MA, USA
[3]Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA
[4]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA
[5]Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA
[6]Department of Pathology, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel
[7]Howard Hughes Medical Institute and Koch Institute of Integrative Cancer Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA
[8]Ann Romney Center for Neurologic Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
[9]Present address: Genentech, 1 DNA Way, South San Francisco, CA, USA
[10]These authors contributed equally
[11]These authors contributed equally
[12]Lead contact
*Correspondence: ljerby@stanford.edu (L.J.), regev.aviv@gene.com (A.R.), acanderson@bwh.harvard.edu (A.C.A.)
https://doi.org/10.1016/j.xcrm.2024.101640

SUMMARY

CD8+ T cells must persist and function in diverse tumor microenvironments to exert their effects. Thus, understanding common underlying expression programs could better inform the next generation of immunotherapies. We apply a generalizable matrix factorization algorithm that recovers both shared and context-specific expression programs from diverse datasets to a single-cell RNA sequencing (scRNA-seq) compendium of 33,161 CD8+ T cells from 132 patients with seven human cancers. Our meta-single-cell analyses uncover a pan-cancer T cell dysfunction program that predicts clinical non-response to checkpoint blockade in melanoma and highlights *CXCR6* as a pan-cancer marker of chronically activated T cells. *Cxcr6* is transactivated by AP-1 and repressed by TCF1. Using mouse models, we show that *Cxcr6* deletion in CD8+ T cells increases apoptosis of PD1+TIM3+ cells, dampens CD28 signaling, and compromises tumor growth control. Our study uncovers a TCF1:CXCR6 axis that counterbalances PD1-mediated suppression of CD8+ cell responses and is essential for effective anti-tumor immunity.

INTRODUCTION

CD8+ T cells in tumors span a spectrum of functional states reflecting their stemness, memory potential, effector function, and dysfunction or exhaustion status, all of which are shaped by both cell-intrinsic regulation and the tumor microenvironment (TME). Harnessing the effector potential of these tumor-infiltrating CD8+ T cells (TILs) is a goal of many cancer therapeutics. A foremost strategy is the blockade of immune checkpoint receptors, primarily CTLA4 and PD1, that are highly expressed on dysfunctional CD8+ TILs. Despite the success of immune checkpoint blockade (ICB) in cancers such as melanoma, lung, and renal,[1–4] only 12% of patients across all cancer indications are estimated to respond.[5]

The TME is a key determinant of ICB response,[6–8] but how it impacts T cell functionality remains incompletely understood. TMEs vary widely within and between cancer types, with differing degrees of vascularization, immune infiltration, and the presence of non-immune cell types like cancer-associated fibroblasts. These factors influence antigen presentation, immuno-regulatory signals, and the availability of oxygen and other nutrients, all known to impact T cell responses.[9,10] Thus, a greater understanding of CD8+ T cell states across multiple tumor types with varied TMEs is critical for informing the development of broadly efficacious therapies.

Single-cell genomics allows the study of T cell transcriptomes at an unprecedented resolution and has generated rich cell atlases[11–20] that uncover opportunities to identify novel regulators and circuits that determine T cell functionality. Here, we developed a dimensionality reduction method that identifies generalizable gene expression programs across cells, patients, and

datasets and applied it to single-cell RNA sequencing (scRNA-seq) data of CD8$^+$ T cells from 132 tumor samples spanning seven cancer types.[11–15,17–20] We uncovered a pan-cancer T cell dysfunction program whose expression level was predictive of failure to respond to ICB in a melanoma cohort. Among the top 10 program genes were the immune checkpoints *CTLA4*, *PDCD1*, and *HAVCR2*, as well as *CXCR6*, a chemokine receptor, whose ligand CXCL16 is expressed by myeloid cells in the TME. Investigating CXCR6 regulation, we found that it was induced by AP-1 and repressed by TCF1, a transcription factor critical for the maintenance of stem-like CD8$^+$ TILs.[21,22] *Cxcr6* deletion promoted apoptosis in PD1$^+$TIM3$^+$ CD8$^+$ TILs and decreased Tox, Bcl2, and CD28 expression, ultimately diminishing tumor growth control. As PD1 dampens T cell responses in part by suppressing CD28 signals,[23] our study uncovered a TCF1:CXCR6 regulatory axis that counterbalances PD1-mediated T cell suppression to support anti-tumor immunity.

RESULTS

GDMF: Generalizable matrix decomposition framework recovers shared programs across studies

The rapid increase in single-cell profiling studies opens opportunities to map cell states across dozens or hundreds of patients, as demonstrated by recent large-scale pan-cancer studies.[24,25] Such studies typically aggregate publicly available data that can vary in collection protocols and data pre-processing, leading to substantial technical variations. Supervised statistical models that identify differentially expressed genes can be designed to account for confounders, covariates, and hierarchical data structures. Conversely, unsupervised methods for extracting latent patterns representing gene programs often do not model this complexity, making their application across diverse datasets challenging, as patterns can reflect confounders rather than shared underlying biology.

To address this, we developed a generalizable matrix decomposition framework (GMDF) for unsupervised meta-analysis of large and diverse datasets. Given a set of scRNA-seq studies, GMDF identifies a low-dimensional representation of cell states across conditions by decomposing cell profiles to shared and context-specific metagenes or programs. The GMDF can incorporate covariates and contexts in the formulation and implementation of the dimensionality reduction task. ''Context'' is defined by the user based on the specific datasets and research question and can denote a specific cancer type, sex, treatment, cohort, etc. GMDF identifies expression programs (i.e., a set of co-expressed genes represented as a single meta-feature) based on variation within a dataset. If this within-dataset variation is repeatedly observed across multiple datasets, then it will be categorized as ''shared.'' If this within-dataset variation is observed only in datasets from a specific ''context,'' it will be identified as ''context-specific.'' By focusing on shared rather than context-specific programs, we can identify biologically meaningful variations that recur even across studies with different technical characteristics.

GMDF consists of two main steps. The first is formulated as a regularized, non-convex optimization problem that minimizes the reconstruction error (Figure 1A) and is solved using block coordinate descent. Second, GMDF applies a consensus approach, where multiple solutions obtained with different subsamples of the data and initializations are aggregated to a single robust consensus solution, thus mitigating the risk of local minima, non-unique solutions, and overfitting (STAR Methods).

In contrast to other methods, such as integrative non-negative matrix factorization (iNMF)[26] and linked inference of genomic experimental relationships (LIGER),[27] GMDF does not force any coupling of the shared and context-specific programs and thus outperforms previous methods in this task, shown using simulated data (Figures S1A–S1D, STAR Methods). Specifically, while LIGER and iNMF identify shared programs with context-specific extensions, GMDF provides sufficient flexibility in identifying both types of programs without enforcing any coupling between them. Comparing these methods, we find that enforcing a specific structure, as done by LIGER, results in suboptimal decompositions and misclassification of context-specific programs as shared.

GDMF identifies pan-cancer T cell expression programs across tumor types

To map CD8$^+$ T cell states across cancer types, we assembled a pan-cancer scRNA-seq compendium of nine datasets, spanning 33,161 CD8$^+$ TILs, collected from 132 cancer patients with melanoma, sarcoma, pancreatic, liver, colon, lung, and breast cancer.[11–15,17–20] We applied GMDF to this pan-cancer CD8$^+$ T cell compendium with cohort covariates (STAR Methods) to identify shared (''pan-cancer'') programs, i.e., a gene module that is repeatedly observed *within* each of the cohorts.

GMDF identified six pan-cancer programs that were shared across all nine cohorts of the compendium, which we annotated as naive/memory, activation/interferon response, AP1/stress response, cell cycle, mitochondrial metabolism, and chronic activation (Figures 1B; Tables S1 and S2). The chronic activation program contained genes associated with T cell dysfunction, including known immune checkpoints (*HAVCR2*, *TIGIT*, *PDCD1*, *CTLA4*, *LAG3*, and *ENTPD1*) and the transcription factor *TOX*.[28–30] In line with previous studies,[12,31] it also included effector genes (*GZMB*, *IFNG*, and *FASLG*), reflecting the coupling between T cell dysfunction and activation.[31–33]

To further examine the dysfunction-activation association, we computed an ''activation score'' for each cell based on its overall expression[12] (STAR Methods) of canonical effector function markers (*NKG7*, *CCL4*, *CST7*, *PRF1*, *GZMA*, *GZMB*, *IFNG*, and *CCL3*) minus its overall expression of canonical naive/memory T cell markers (*CCR7*, *TCF7*, *LEF1*, and *SELL*). Naive cells, with high expression of naive markers and low expression of effector markers, had low expression of the pan-cancer chronic activation program. However, as cells expressed higher activation scores, they were associated with increasingly higher values of the chronic activation program and showed more intercellular variation (Figure 1C).

To decouple activation and dysfunction, we first normalized the chronic activation GMDF score based on the expected value given the ''activation score'' (LOWESS regression line, Figure 1C, STAR Methods). Next, we annotated each T cell as naive/memory (low activation scores), effector (high activation and lower-than-expected chronic activation scores, i.e., below the

CellPress
OPEN ACCESS

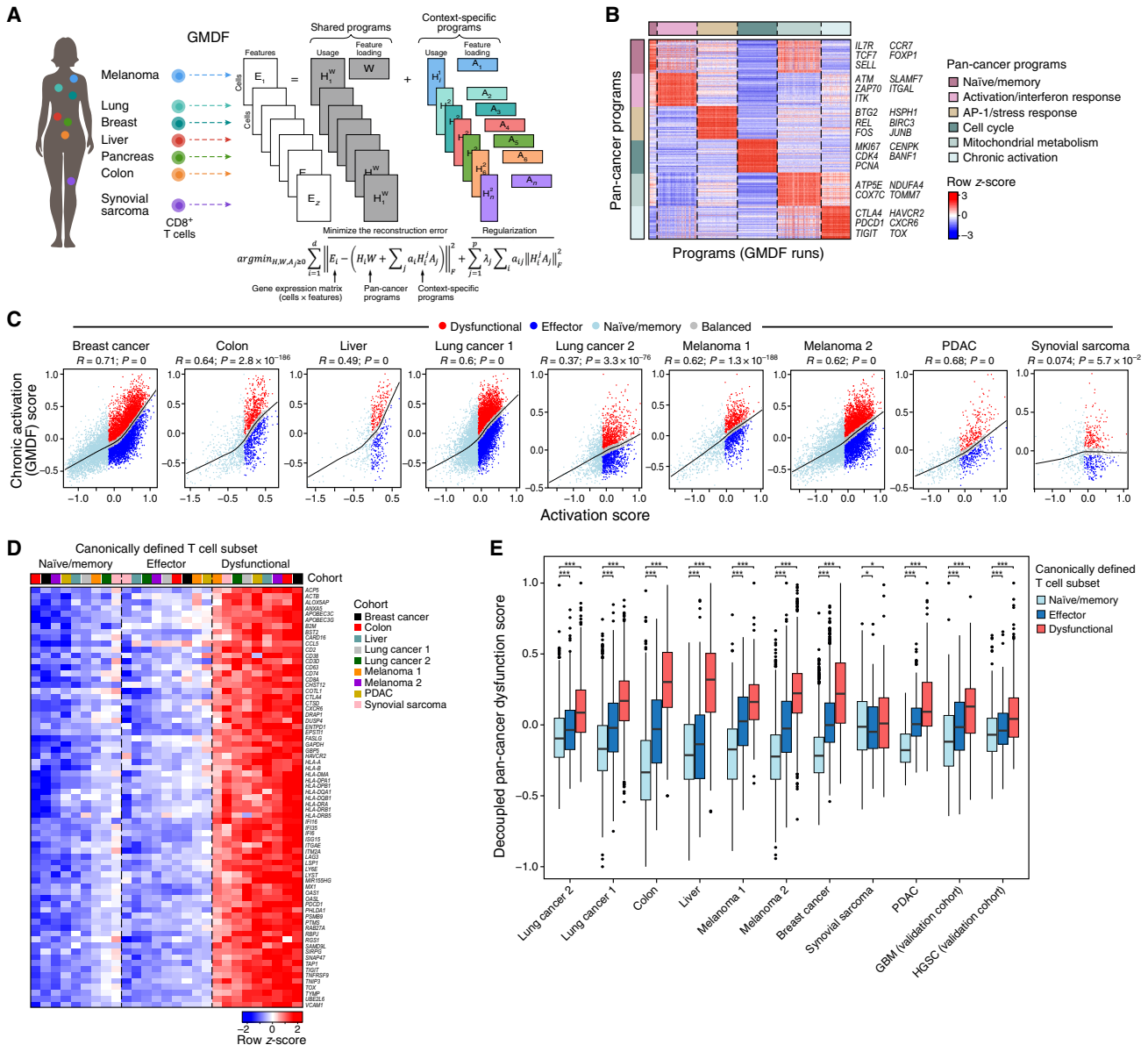


Figure 1. Pan-cancer approach reveals highly generalizable T cell expression programs

(A) Analysis approach. scRNA-seq data of CD8$^+$ TILs from seven human cancers (left) were analyzed with GMDF (center) revealing shared (pan-cancer) and context (tumor)-specific expression programs (right).

(B) GMDF identified 6 pan-cancer T cell expression programs. Top 100 genes (rows) of each program (color bar on the left) and their weights (W matrix) across programs from the different GMDF solutions (columns). Right: representative genes from each program.

(C) Chronic activation score (y axis) and activation score (x axis) for each CD8$^+$ T cell (dot) from each study (panels) with cells classified as dysfunctional, effector, or naive/memory by their signatures, or as "balanced," if their chronic activation score is at the expected level based on the LOWESS regression line (black). Pearson's R and association *p* value are shown.

(D) Average expression (row *Z* score, color bar) of genes (rows) from the decoupled pan-cancer dysfunction program in CD8$^+$ T cell subsets (columns) from 9 scRNA-seq studies stratified by expression of canonical markers (naive/memory: *CCR7*, *TCF7*, *LEF1*, and *SELL*; effector: *NKG7*, *CCL4*, *CST7*, *PRF1*, *GZMA*, *GZMB*, *IFNG*, and *CCL3*; dysfunctional: *PDCD1*, *TIGIT*, *HAVCR2*, *LAG3*, and *CTLA4*).

(E) Distribution of overall expression scores (y axis) of the decoupled pan-cancer dysfunction program (minus immune checkpoints: *PDCD1*, *TIGIT*, *HAVCR2*, *LAG3*, and *CTLA4*) in CD8$^+$ T cells stratified as in (D). Middle line: median; box edges: 25th and 75th percentiles, whiskers: most extreme points that do not exceed ± IQR*1.5; further outliers are marked individually. ***$p < 0.001$, mixed-effects test.

LOWESS regression line), or dysfunctional (high activation and higher-than-expected chronic activation scores, i.e., above the LOWESS regression line) (Figure 1C). Finally, using these three cell annotations and multilevel regression modeling (STAR Methods), we identified the unique features of each cell state across all cohorts, resulting in three ''decoupled pan-cancer programs'' of effector, dysfunction, and naive/memory cell states (Table S3). Indeed, when we stratified cells based on the expression of canonical naive/memory and effector markers (listed earlier), as well as dysfunction markers (*PDCD1*, *TIGIT*, *HAVCR2*, *LAG3*, and *CTLA4*), the decoupled pan-cancer programs reliably identified these canonically defined subsets (Figure S2A; Figure 1D). The decoupled pan-cancer dysfunction program consisted of 72 genes and notably did not include 35 genes from the original GMDF chronic activation program, including *GZMA*, *GZMB*, *IFNG*, *IL2RG*, *PRF1*, and *STAT1*, all of which are associated with T cell effector functions (Table S4), demonstrating the decoupling of the effector and dysfunction components.

To test the generalizability of the decoupled programs, we stratified the cells by expression of the canonical effector, naive/memory, and dysfunction markers (as aforementioned), and then examined whether using only the *de novo* genes in the three decoupled pan-cancer programs (i.e., after removing the canonical markers) could identify the pertaining cell population. Indeed, each of the decoupled programs marked the respective T cell populations across all cohorts, as well as in two additional studies of glioblastoma[16] and ovarian cancer[34,35] that were not included in the discovery compendium ($p < 6.82*10^{-14}$, $5.68*10^{-9}$, $4.26*10^{-8}$, mixed-effects test, for the dysfunction, effector, and naive/memory programs, respectively, in glioblastoma and similarly, $p < 1.38*10^{-35}$, $5.85*10^{-3}$, $3.53*10^{-14}$, in ovarian cancer, STAR Methods, Figures 1E; S2B and S2C).

The decoupled pan-cancer T cell dysfunction program predicts clinical non-response to immunotherapy

To test the clinical relevance of the pan-cancer programs, we assessed their overall expression in scRNA-seq data of $CD8^+$ TILs collected from patients with melanoma before and after ICB.[18] Each of the programs highlighted distinct populations of cells (Figure 2A). The decoupled pan-cancer dysfunction program was overexpressed in cells from non-responder patients compared to cells from responders (Figure 2A, $p = 6.37*10^{-50}$, mixed-effects test; pre- and post-treatment combined). This is consistent with the previously reported correlation of *ENTPD1* and *HAVCR2*, both members of the decoupled pan-cancer dysfunction program (Figures S3A and S3B), with failure to respond to ICB.[18] The decoupled pan-cancer dysfunction program also includes *SIRPG* (Figure S3C), the ligand of *CD47*, consistent with studies linking CD47 expression in cancer cells to T cell function[36–38] and lack of ICB response.[12] Notably, the decoupled naive/memory program, containing *TCF7* (Figure S3D), showed the opposite trend (Figure 2A, $p = 3.91*10^{-25}$, mixed-effects), consistent with previous reports.[18]

Importantly, expression of the decoupled pan-cancer dysfunction program in pre-treatment $CD8^+$ TILs was predictive of clinical non-response to ICB in patients with melanoma at both the single-cell (area under the ROC curve [AUC] = 0.72, $p = 9.7*10^{-83}$) and sample (AUC = 0.86, $p = 0.0075$, t test) levels (Figure 2B). In four patients with responding and non-responding lesions, the dysfunction program predicted non-response at the lesion level, showing significantly higher expression in cells from non-responding vs. responding lesions ($p < 1*10^{-30}$, Fisher combined test, Figure 2C). The decoupled pan-cancer dysfunction program shared features with previously reported signatures of T cell dysfunction in cancer and thus had similar predictive performance (Table S5). Notably, the GMDF chronic activation program (comprising both dysfunction and effector genes) was associated with lack of response. However, once decoupled into separate effector and dysfunction programs, they instead associated with response and non-response, respectively, supporting the value of decoupling (Table S5). We tested the predictive power of the pan-cancer dysfunction program in two additional scRNA-seq datasets: a breast cancer cohort[39] ($N = 11$ patients) and a basal ($N = 11$) or squamous cell carcinoma cohort ($n = 4$) treated with ICB.[40] The pan-cancer dysfunction program was not predictive of ICB response in these cohorts. This could reflect the lower number of patients compared to the melanoma cohort ($N = 32$), although these cohorts had more cells sampled per patient, or the more modest clinical responses observed in these cohorts (only one patient with complete response). Additional cohorts with larger numbers of ICB-treated patients across cancer types will be needed to compare the signature's predictive performance in a broader context.

CXCR6 is a high-ranking gene in the decoupled pan-cancer dysfunction program

Examining the decoupled pan-cancer dysfunction program genes for novel regulators, we noted *CXCR6* among the top 5 genes, with a high loading in the GMDF consensus solution. *CXCR6* was also the only gene that consistently marked dysfunctional T cells across all cancers examined (Figure 1D, Benjamini-Hochberg false discovery rate [FDR] <0.01, mixed-effects, in all cohorts). Previous studies have associated *CXCR6* with improved immune responses along with effector and memory T cell recruitment and trafficking.[41,42] More recent work has shown that CXCR6 is expressed in dysfunctional T cells in multiple cancer types[20,43] and in chronic viral infections.[44,45] Furthermore, CXCR6-mediated interactions between $CD8^+$ T cells and intra-tumoral dendritic cells (DCs) were shown to be important for sustained tumor control in a murine melanoma model.[46] Given these studies indicating the importance of CXCR6 for anti-tumor immunity, we pursued its study in more depth.

CXCR6 was predominantly expressed in T cells, and its ligand, CXCL16, was predominantly expressed by myeloid cells in the three tumor types in our compendium (sarcoma, lung, and melanoma) and an additional glioblastoma dataset where non-T cells were profiled[12–15,18,47] (Figure 3A). Based on these data and data from a breast cancer cohort,[11] we identified a program of genes co-expressed with *CXCL16* in macrophages (Figure S3E; Table S6). The program included complement genes (*C1QA*, *C1QB*, and *C1QC*), major histocompatibility complex (MHC) class II genes (*HLA-DMB* and *HLA-DOA*), and the co-stimulatory receptor *CD86*, indicating that CXCL16 marks activated cells engaged in antigen cross-presentation. Indeed,

CellPress
OPEN ACCESS

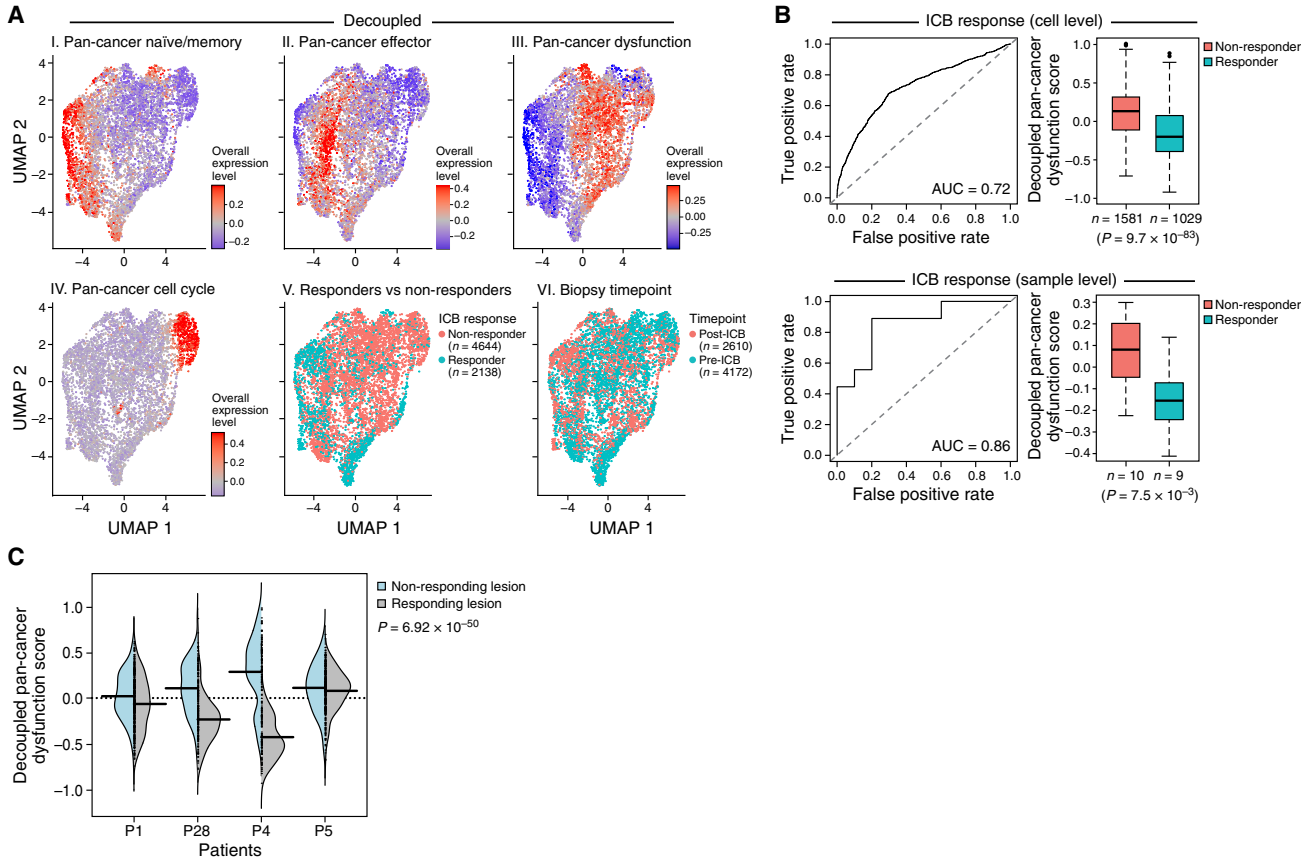


Figure 2. The pan-cancer T cell dysfunction program predicts ICB responses

(A) Uniform manifold approximation and projection (UMAP) embedding of CD8$^+$ T cell profiles (dots) from 48 melanoma tumors,[18] colored by the overall expression of the decoupled pan-cancer naive/memory (I), effector (II), and dysfunction (III), or pan-cancer cell cycle (IV) program, ICB response status (V), or pre- versus post-treatment biopsy status (VI).

(B) Left: true positive (y axis) and false positive (x axis) rates when predicting the clinical response of the tumor based on different levels of the pan-cancer dysfunction program score at either the CD8$^+$ T cell (top) or sample (bottom) level. Right: distribution of the overall expression (y axis) of the decoupled pan-cancer dysfunction program in responders or non-responders at the cell (top) and sample (bottom) level. Middle line: median; box edges: 25th and 75th percentiles, whiskers: most extreme points that do not exceed ± IQR*1.5; further outliers are marked individually.

(C) Distribution of overall expression of the decoupled pan-cancer dysfunction program in the responding and non-responding lesions of patients (x axis) with mixed responses. Statistical significance was determined by Student's t test (B, right) and Fisher combined test (C).

CXCL16 was co-expressed with antigen processing and presentation signatures in macrophages in all datasets (Figure 3B). These findings, along with data from mouse models showing that CD8$^+$ T cell CXCR6 expression optimized interactions with CXCL16-expressing dendritic cells,[46] indicated that CXCR6-CXCL16 interactions may mediate spatiotemporal control of chronically activated T cells in both mouse and human tumors.

CXCR6 expression tracks with T cell dysfunction and tumor progression in murine tumors

We next assessed CXCR6 expression in CD8$^+$ T cells in different mouse tumor models that vary in immunogenicity and in their expression of Ova as an ectopic tumor antigen. In CD8$^+$ TILs, the frequency of cells expressing CXCR6 increased from PD1$^-$ TIM3$^-$ (naive-like) to PD1$^+$TIM3$^-$ (stem-like PD1$^+$TCF1$^+$TIM3$^-$ and effector-like PD1$^+$TCF1$^-$TIM3$^-$) to PD1$^+$TIM3$^+$ (terminally dysfunctional) cells across all models (Figure 3C, top; Figure S3F). Tumor antigen-specific (Ova-Dextramer$^+$) cells showed the same pattern in Mc38Ova^{hi} but not in B16Ova tumors, where antigen-specific cells were less frequent (Figure 3C, bottom; Figure S3F, Figure S3G). Further, CXCR6 expression was higher in PD1$^+$TIM3$^+$ vs. PD1$^+$TIM3$^-$ cells among both total and tumor antigen-specific CD8$^+$ TILs (Figure S3H). CXCR6$^+$ TILs expressed more CX3CR1, associated with effector function,[48–51] Ki-67, and multiple co-inhibitory receptors (PD1, TIM3, LAG3, TIGIT, and CD39) and Tox,[28–30] all of which are associated with dysfunction (Figure 3D). Notably, CXCR6$^+$ T cells expressed lower TCF1 (Figure 3D), which represses effector CD8$^+$ T cell differentiation[52,53] and maintains stem-like cells that give rise to effector and, eventually, dysfunctional cells in both cancer and chronic virus infection models.[21,22,54,55] In line with the human pan-cancer compendium, CXCL16 was expressed by myeloid cells across the three tumor models, with the highest expression in MHCII^{hi} macrophages, followed by MHCII^{lo} macrophages, and DC subsets (Figure 3E).

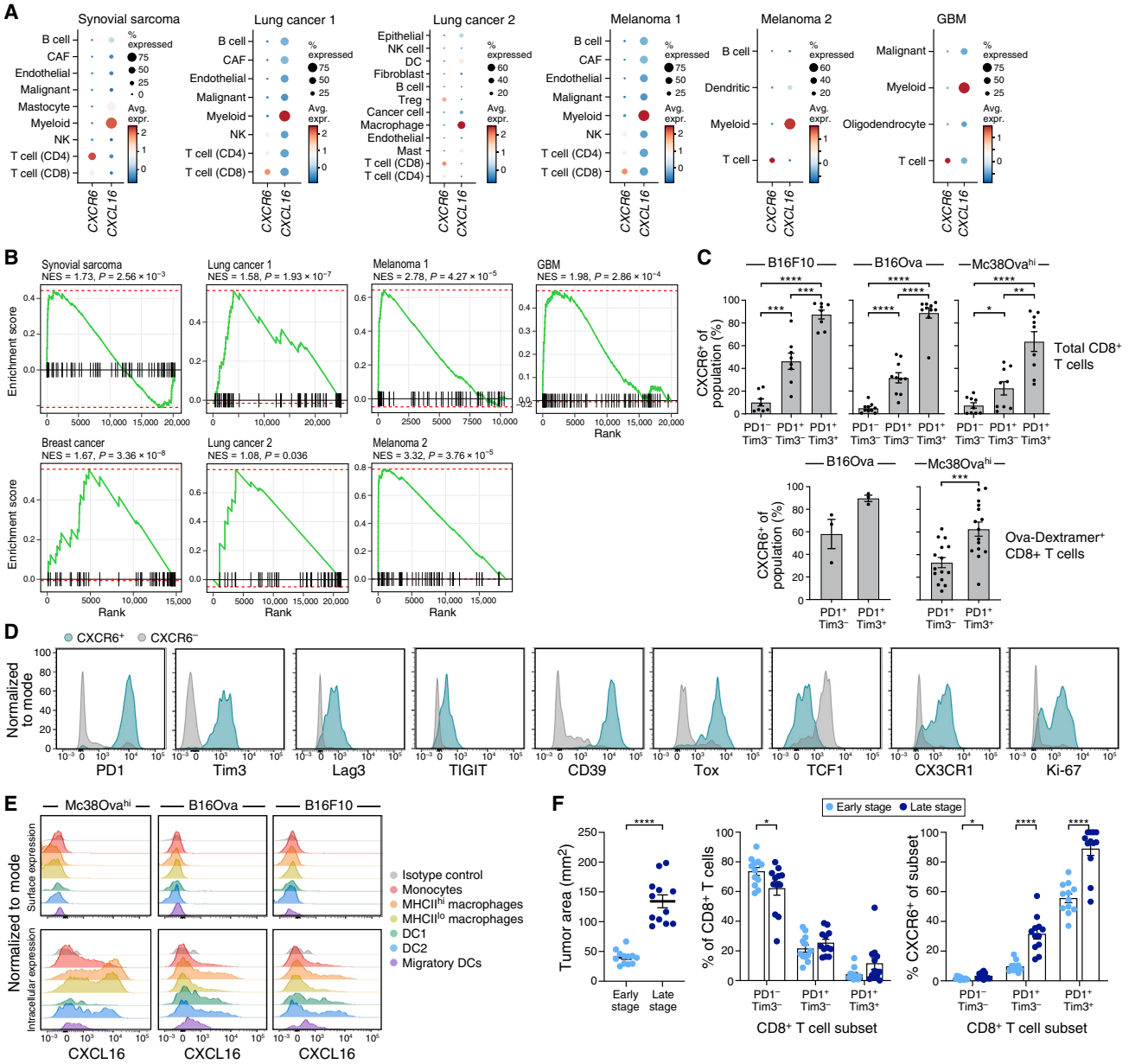


Figure 3. CXCR6 and CXCL16 expression in human and murine tumors tracks with dysfunctional T cells and myeloid cells, respectively

(A) Mean expression (color bar) and fraction of expressing cells (dot size) for *CXCR6* and *CXCL16* (columns) across cell types (rows) in different human tumor studies (panels).

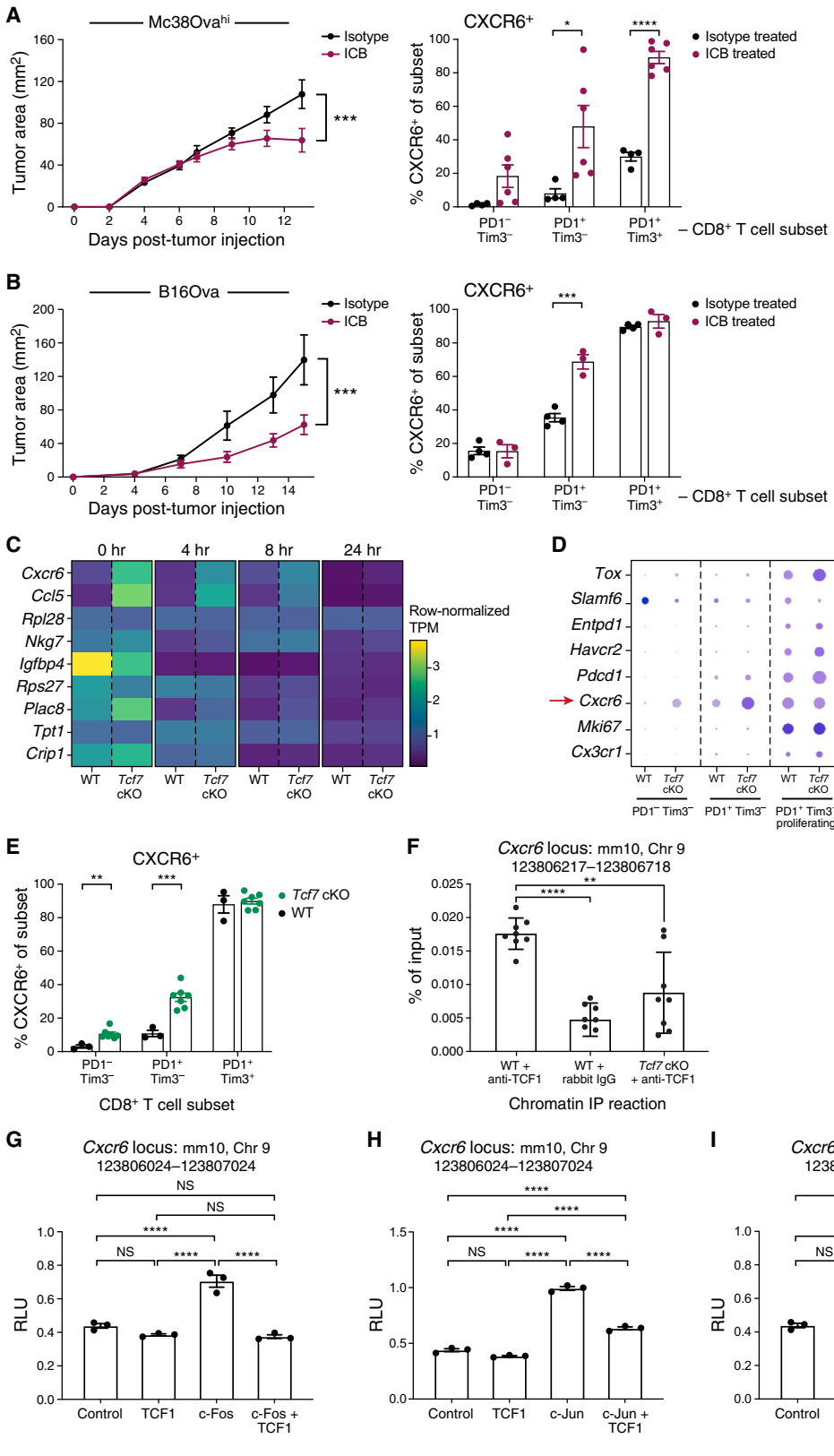
(B) GSEA plots obtained for the KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION gene set when compared against the ranking of genes based on their co-expression with *CXCL16* in macrophages in the indicated tumors.

(C) Frequency of CXCR6$^+$ cells (y axis, mean ± SEM) in the indicated subsets (x axis) of total (top) or Ova-Dextramer$^+$ (bottom) CD8$^+$ TILs harvested from B16F10, B16Ova, or Mc38Ova^{hi} tumors (top: B16F10, n = 8, 2 experiments combined. B16Ova, n = 10, 2 experiments combined. Mc38Ova^{hi}, n = 9, 2 experiments combined. Bottom: B16Ova, n = 3, 1 experiment. Mc38Ova^{hi}, n = 15, 2 experiments combined).

(D) Representative distributions of expression levels (x axis, fluorescence intensity) in CXCR6$^+$ and CXCR6$^-$ CD8$^+$ TILs from B16Ova tumors (n = 4, 1 experiment).

(E) Representative distributions of CXCL16 surface (top) or intracellular (bottom) expression (x axis, fluorescence intensity) in myeloid cells from Mc38Ova^{hi}, B16Ova, or B16F10 tumors (n = 4 per tumor, 1 experiment).

(F) Left: B16Ova tumor area (y axis, mean ± SEM) of early-stage or late-stage tumors. Middle: frequency (y axis, mean ± SEM) of PD1- and TIM3-expressing CD8$^+$ TILs (x axis) in early- or late-stage tumors. Right: frequency of CXCR6$^+$ cells (y axis, mean ± SEM) in the indicated subsets (x axis) of CD8$^+$ TILs from early- or late-stage B16Ova tumors (n = 12, 2 experiments combined). Statistical significance was determined by Student's unpaired t test (C, F). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.



(legend on next page)

These data point to CXCR6-CXCL16 mediating interactions between effector and dysfunctional PD1$^+$TIM3$^+$CD8$^+$ TILs with antigen-presenting cells in the TME.

We further characterized CXCR6 expression over the course of tumor progression, comparing CD8$^+$ TILs from B16Ova tumors harvested at early versus late stages of progression. While the proportions of PD1- and TIM3-expressing subsets remained mostly unchanged, CXCR6 expression within PD1$^+$TIM3$^-$ and PD1$^+$TIM3$^+$ cells increased with tumor progression (Figure 3F). Thus, CD8$^+$ T cells increasingly acquired CXCR6 expression as they transitioned to effector and finally terminally dysfunctional T cell states with tumor progression.

CXCR6 is up-regulated with ICB and is repressed by TCF1

The gradient of increasing CXCR6 expression from effector to dysfunctional CD8$^+$ TILs (Figures 3C; S3H) and a previous study showing its expression increases upon ICB in a mouse model[56] prompted us to examine whether CXCR6 increased in response to ICB in our human cohort. In CD8$^+$ TILs from patients with melanoma collected pre- and post-treatment,[18] CXCR6 expression increased following ICB treatment ($p = 4.63*10^{-4}$, mixed-effects) (Figure S4A). To test this further in a controlled setting, we examined two different tumor models (Mc38Ova^{hi} and B16Ova) and two different ICB therapies (anti-PD1 and anti-PD-L1 + anti-TIM3). In the immunogenic Mc38Ova^{hi} model, the fraction of CXCR6-expressing cells increased with ICB in both the PD1$^+$TIM3$^-$ and PD1$^+$TIM3$^+$ CD8$^+$ TIL populations (Figure 4A; Figure S4B). In the less immunogenic B16Ova model, the fraction of CXCR6-expressing cells increased upon ICB in the PD1$^+$TIM3$^-$ but not in the PD1$^+$TIM3$^+$ subset, where nearly all cells expressed CXCR6 prior to ICB (Figures 4B; S4C). We also observed an increase in the CXCR6 expression level within the subsets upon ICB in both models (Figures S4D and S4E). The PD1$^+$TIM3$^-$ population, containing both TCF1$^+$ stem-like and TCF1$^-$ effector-like cells, showed the most robust CXCR6 expression increase upon ICB. Interestingly, TCF1 and CXCR6 expression were mutually exclusive within this population (Figure S4F), and this was even more striking among antigen-specific cells (Figure S4G). Thus, the increased CXCR6 expression within this population likely reflected a shift toward TCF1$^-$ effector-like cells upon ICB.

Given the observed CXCR6 increase in CD8$^+$ TILs upon ICB (Figures 4A and 4B) and its anti-correlation with TCF1 (Figures S4F and S4G), we hypothesized that TCF1 regulated *Cxcr6* expression. To test this, we analyzed bulk RNA sequencing (RNA-seq) of CD8$^+$ TILs isolated from either wild-type (WT) mice or mice bearing conditional deletion of *Tcf7* (the gene encoding TCF1) in mature CD8$^+$ T cells (*Tcf7* cKO) over a time course of *in vitro* activation. CXCR6 was among the top 5 differentially expressed genes between WT and *Tcf7* cKO cells both prior to activation (time point 0, $p = 6.23*10^{-23}$, Wald test) and at subsequent time points, with the expression difference gradually decreasing over time (Figure 4C). Moreover, scRNA-seq of B16Ova tumors from WT and *Tcf7* cKO mice showed significantly higher CXCR6 expression in *Tcf7* cKO cells, specifically the PD1$^-$TIM3$^-$ ($p < 1*10^{-4}$, Wilcoxon test) and PD1$^+$TIM3$^-$ CD8$^+$ TILs ($p < 1*10^{-4}$, Wilcoxon test) (Figure 4D), which we confirmed at the protein level (Figure 4E).

These results suggested that *Tcf7*/TCF1 repressed CXCR6 expression in PD1$^-$TIM3$^-$ and PD1$^+$TIM3$^-$ CD8$^+$ T cells (where TCF1 is expressed; TCF1 levels become negligible once TIM3 is expressed), consistent with TCF1 acting as a transcriptional repressor.[57] Supporting direct regulation, chromatin immunoprecipitation (ChIP) sequencing data show TCF1 binding in the CXCR6 promoter region, which is also shown to be chromatin accessible in naive CD8$^+$ T cells[58,59] (Figure S4H). We confirmed TCF1 binding in this region (mm10, chr9: 123806217–123806718) by ChIP-PCR in CD8$^+$ T cells from WT and *Tcf7* cKO mice (Figure 4F). To test TCF1 repression of *Cxcr6* transcription, we identified candidate transcription factors with predicted binding sites in a 1-kb genomic region surrounding the predicted TCF1-binding sites (mm10, chr9: 123806024–123807024) (STAR Methods). Of these, 15 were expressed in CD8$^+$ TILs from B16F10 tumors,[33] including three members of the AP-1 transcription factor family, c-Fos, c-Jun, and JunD. We therefore tested if AP-1 transcription factors could induce *Cxcr6* expression and whether this induction could be repressed by TCF1. Using luciferase assays, we showed that c-Fos, c-Jun, and JunD induced *Cxcr6* transcription, but, upon addition of TCF1, transcription was repressed to control levels (Figures 4G–4I). Thus, TCF1 repressed AP-1-induced *Cxcr6* expression.

Figure 4. CXCR6 expression increases upon ICB and is repressed by TCF1

(A and B) Left: tumor area (y axis, mean ± SEM) over time (x axis) of Mc38Ova^{hi} (A) or B16Ova (B) implanted in wild-type (WT) mice and treated with anti-PD1 (A), anti-PD-L1 + anti-TIM3 (B), or isotype control. Right: frequency of CXCR6$^+$ cells (y axis, mean ± SEM) in the indicated subsets (x axis) of CD8$^+$ TILs harvested from tumors treated as above (A: $n = 4$–6 per group, 1 experiment. B: $n = 3$–4 per group, 1 experiment).

(C) Expression (row-normalized TPM, color bar) of top differentially expressed genes (rows) between WT OTI (E8i-Cre$^-$, *Tcf7*$^{FL/FL}$) and *Tcf7* cKO OTI (E8i-Cre$^+$, *Tcf7*$^{FL/FL}$) cells at different time points after T cell activation (columns) ($n = 3$, 1 experiment).

(D) Mean expression (color bar) and fraction of expressing cells (dot size) of key genes (rows) in different CD8$^+$ T cell clusters (columns) as determined by scRNA-seq of cells from B16Ova tumors implanted in WT or *Tcf7* cKO mice ($n = 3$ per group combined, 1 experiment).

(E) Frequency of CXCR6$^+$ cells (y axis, mean ± SEM) in the indicated subsets (x axis) of CD8$^+$ TILs from B16Ova tumors from WT (black) or *Tcf7* cKO (green) mice ($n = 3$–7 per group, representative of 2 experiments).

(F) Percent input (y axis, mean ± SEM) following chromatin immunoprecipitation (ChIP) PCR of the *Cxcr6* locus with anti-TCF1 or rabbit IgG control antibodies (x axis) in WT or *Tcf7* cKO CD8$^+$ T cells ($n = 8$, 5 experiments combined).

(G–I) Luciferase activity (RLU, relative light unit, y axis, mean ± SEM) in HEK293T cells transfected with *Cxcr6* locus-containing pGL4.10 luciferase reporters together with either empty vector (control) or vectors encoding the indicated transcription factors (x axis). Firefly luciferase activity is presented relative to constitutive Renilla luciferase activity ($n = 3$, representative of 2 experiments). Statistical significance was determined by linear mixed model (A left, B left), Student's unpaired t test (A right, B right, E, F comparing WT to *Tcf7* cKO anti-TCF1 samples), Student's paired t test (F comparing WT samples), or one-way ANOVA with Tukey's multiple comparisons test (G, H, I). NS = not significant, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$.

CellPress
OPEN ACCESS

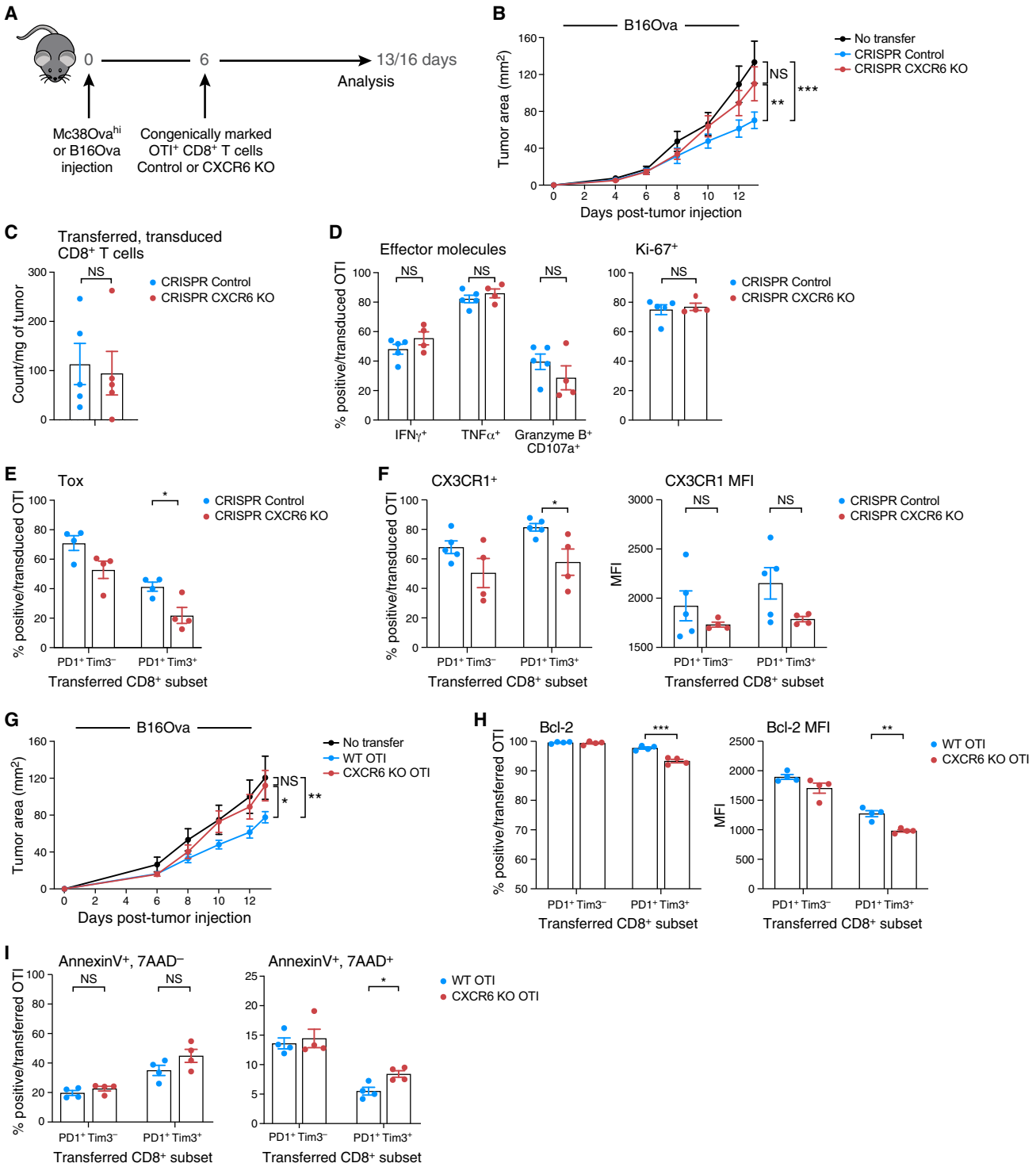


Figure 5. CXCR6 knockout reduces anti-tumor immunity

(A) Experimental design for adoptive transfer experiments.

(B–F) B16Ova tumor-bearing mice were adoptively transferred with CRISPR CXCR6 KO OTI T cells, CRISPR control OTI T cells, or no transfer control as indicated in (A). (B) Tumor area (y axis, mean ± SEM) over time (x axis) (n = 5 per group, representative of 2 experiments). (C) Number of transferred, transduced OTI cells per mg of tumor tissue (y axis, mean ± SEM) (n = 5 per group, representative of 2 experiments). (D) Left: frequency of IFNγ$^+$, TNF-α$^+$, or Granzyme B$^+$CD107a$^+$ cells (y axis, mean ± SEM) in transferred, transduced OTI cells after *ex vivo* activation with Ova 257–264. Right: frequency of Ki-67$^+$ cells (y axis, mean ± SEM) in transferred, transduced OTI cells (n = 4–5 per group, representative of 2 experiments). (E) Frequency of Tox$^+$ cells (y axis, mean ± SEM) in the indicated subsets

(legend continued on next page)

CXCR6 is required to maintain tumor control

To study the impact of CXCR6 expression on anti-tumor CD8$^+$ T cell function, we next utilized CRISPR-Cas9 to delete CXCR6 in mature OTI CD8$^+$ T cells and transferred them into mice bearing Ova-expressing tumors (Figure 5A). This approach deletes genes at the time of T cell activation, thus bypassing any compensatory mechanisms that might arise when genes are deleted during development. We confirmed that CXCR6 was efficiently deleted in CRISPR CXCR6 knockout (KO) OTI cells *in vivo* (Figure S5A) and found that CRISPR CXCR6 KO T cells failed to control tumor growth compared to CRISPR control T cells in both the B16Ova and Mc38Ova^{hi} models (Figures 5B; S5B), in line with previous reports.[46,56] We did not observe significant differences in the numbers of transduced cells within tumor tissue, although there was a trend toward fewer CRISPR CXCR6 KO OTI cells (Figures 5C; S5C). PD1 and TIM3 expression did not differ between CRISPR CXCR6 KO and CRISPR control OTI cells (Figure S5D). Despite the compromised tumor control by CRISPR CXCR6 KO OTI cells (Figure 5B; Figure S5B), functional analyses showed no major differences in their proliferation (Ki-67), pro-inflammatory cytokine production (IFNγ and TNF-α), or cytotoxic capacity (GranzymeB$^+$CD107a$^+$) (Figure 5D). However, Tox expression, known to impact the survival of CD8$^+$ TILs,[29] was reduced in PD1$^+$TIM3$^+$ CRISPR CXCR6 KO cells compared to CRISPR control cells (Figure 5E). There were also fewer CRISPR CXCR6 KO PD1$^+$TIM3$^+$ OTI cells expressing CX3CR1 and a trend toward lower CX3CR1 expression levels (Figure 5F).

We further confirmed that loss of CXCR6 compromised tumor growth control using CD8$^+$ T cells from OTI mice crossed with CXCR6 KO mice (Figure 5G). Again, there was a non-significant trend toward fewer transferred CXCR6 KO OTI cells in tumors (Figure S5E) and no difference in the proliferation, cytokine production, or cytotoxicity of transferred cells (Figure S5F). However, in line with our CRISPR data, the expression level of CX3CR1 was decreased in CXCR6 KO PD1$^+$TIM3$^+$ cells compared to WT OTI (Figure S5G, right), although the proportion of CX3CR1$^+$ cells was similar (Figure S5G, left).

As CX3CR1 expression has been associated with expression of the pro-survival protein Bcl2 in other immune cells,[60,61] we compared Bcl2 expression in our transferred cells. Although Bcl2 expression was high overall, the proportion of Bcl2-expressing cells and Bcl2 expression levels were significantly lower in PD1$^+$TIM3$^+$ CXCR6 KO OTI cells compared to WT OTI cells (Figure 5H). Consistent with a previous report that showed CXCR6-deficient CD8$^+$ T cells survived less in the TME,[46] there were more apoptotic (Annexin V$^+$, 7AAD$^+$) cells among PD1$^+$TIM3$^+$ CXCR6 KO OTI cells compared to WT OTI cells (Figure 5I). Collectively, these data support a model where CXCR6-CXCL16 interactions may regulate the survival of PD1$^+$TIM3$^+$ cells, possibly via alterations in CX3CR1 and Bcl2, such that CXCR6 KO cells are more apoptotic and therefore less effective at mediating tumor clearance.

Defective CD28 co-stimulation in CXCR6 KO cells

To further identify how CXCR6 signaling affects T cells, we performed RNA-seq on WT and CXCR6 KO OTI cells after co-culture with Ova-pulsed CXCL16$^+$ bone-marrow-derived dendritic cells (Figure S5H). The top three pathways enriched in WT vs. CXCR6 KO OTI cells were leukocyte activation, regulation of regulatory T cell differentiation, and co-stimulation by the CD28 family (Figures 6A and 6B; Table S7), whereas those enriched in CXCR6 KO OTI were metabolism of RNA, ribonucleoprotein complex biogenesis, and assembly (Figure 6C). Notably, the "costimulation by the CD28 pathway" (R-MMU-388841) was significantly higher in WT vs. CXCR6 KO OTI cells by gene set enrichment analysis (Figure 6D).

CD28 signaling is not only important for effective T cell co-stimulation and T cell survival,[62–64] but it is also critical for ICB response and was recently shown to be abrogated by PD1 signaling.[23,65] Given the increased apoptosis observed in CXCR6 KO T cells (Figure 5), we assessed CD28 expression in WT and CXCR6 KO OTI cells at early (day 3) and late (day 8) time points after co-transfer into the same tumor-bearing host (Figure 6E). Early post-transfer, WT and CXCR6 KO OTI cells were present in tumor tissue at similar frequencies, but later, all mice showed higher frequencies of WT cells (Figure 6F), in line with increased apoptosis in CXCR6 KO cells. Notably, these differences in survival were more readily observed in a co-transfer setting, likely reflecting the decreased capacity of CXCR6 KO cells to compete with WT cells for limiting growth and survival signals within the same TME. Importantly, although similar prior to transfer, the frequency of CD28$^+$ cells was lower among CXCR6 KO OTI cells at both early and late time points, as was the CD28 expression level per cell at the day 8 time point (Figure 6G), consistent with our RNA-seq data (Figure 6A). CD28 co-stimulation is directly linked to the pro-survival protein Bcl-xL.[62–64] Indeed, there was a decreased frequency of Bcl-xL-expressing cells and a lower Bcl-xL expression level in CXCR6 KO OTI cells (Figure 6H). Finally, there was lower CX3CR1 expression in CXCR6 KO OTI cells (Figure 6I), confirming our previous results (Figure 5F). Together, these data indicate that CD28 modulation may be one mechanism by which CXCR6 contributes to the preservation of anti-tumor T cell responses.

(x axis) of transferred, transduced OTI cells (n = 4 per group, 1 experiment). (F) Frequency of CX3CR1$^+$ cells (left, y axis, mean ± SEM) and CX3CR1 expression level in CX3CR1$^+$ cells (right, y axis, geometric mean fluorescence (MFI), mean ± SEM) in the indicated subsets (x axis) of transferred, transduced OTI cells (n = 4–5 per group, representative of 2 experiments).

(G–I) B16Ova-tumor bearing mice were adoptively transferred with CXCR6 knockout (KO) OTI cells, WT OTI cells, or no T cells. (G) Tumor area (y axis, mean ± SEM) over time (x axis) (n = 4–5 per group, 1 experiment). (H) Frequency of Bcl-2$^+$ cells (left, y axis, mean ± SEM) and level of Bcl-2 in Bcl-2$^+$ cells (right, y axis, geometric mean fluorescence (MFI), mean ± SEM) in the indicated subsets (x axis) of transferred OTI cells (n = 4 per group, 1 experiment). (I) Frequency of Annexin V$^+$, 7AAD$^-$ cells (left, y axis, mean ± SEM) and of Annexin V$^+$, 7AAD$^+$ cells (right, mean ± SEM) in the indicated subsets (x axis) of transferred OTI cells (n = 4 per group, 1 experiment). Statistical significance was determined by linear mixed model (B, G) and Student's unpaired t test (C–F, H, I). NS = not significant, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

DISCUSSION

Here, we devised and applied a new matrix factorization algorithm, GMDF, to map CD8$^+$ T cell gene expression programs across a spectrum of human cancers and identify a generalizable pan-cancer T cell dysfunction program. One of the top-ranking genes in the program was *CXCR6*, which we showed is negatively regulated by TCF1 and is predominantly expressed in PD1$^+$TIM3$^+$ terminally dysfunctional cells, where it may promote their survival through various mechanisms including promotion of CD28 co-stimulation and expression of Tox, Bcl2, and CX3CR1. Accordingly, CXCR6 deletion resulted in decreased tumor growth control, highlighting the contribution of dysfunctional T cells in mediating anti-tumor immunity and the importance of the TCF1:CXCR6 regulatory circuit.

GMDF incorporates mixed effects in unsupervised modeling to leverage the rapidly increasing number of single-cell studies through a unified pattern recognition framework. It identified shared programs as gene modules that are repeatedly observed in all datasets (i.e., a set of genes that co-vary within each one of the input datasets), as well as context-specific programs as gene modules that co-vary only in some of the datasets. These latent gene modules are identified by considering the co-variation of genes within each dataset.

While our study focuses on programs shared across CD8$^+$ TILs in various TMEs, context-specific programs could potentially identify ways to unleash immune responses in specific contexts. However, to uncover context-specific programs, data from multiple cohorts of the same tumor type are required, along with additional examination of GMDF's performance and robustness. Similarly, although expression of the decoupled pan-cancer dysfunction program was associated with ICB response in the melanoma cohort we studied, this connection, potentially in combination with other ICB resistance signatures in other cell types (e.g., in malignant cells[12]), should be further examined as larger and more statistically powered cohorts of ICB-treated patients across cancer types become available.

CXCR6 was one of the top members of the decoupled pan-cancer dysfunction program, whose expression predicted failure to respond to ICB in melanoma. Yet, CXCR6 expression increases in both patients with melanoma[18] and murine models upon ICB. Additionally, loss of CXCR6 dampened tumor control in mouse models. Our data therefore position CXCR6 as a positive regulator of dysfunctional T cells that sustains their limited functionality in the face of chronic activation in the TME. Thus, the pan-cancer dysfunction program contains both positive (e.g., *CXCR6*) and negative (e.g., *PDCD1*/PD1) regulators of T cell function and persistence in the TME, and their co-expression serves to tightly control chronically activated T cells.

CXCR6 KO CD8$^+$ T cells expressed lower CD28 and thus may receive less co-stimulation than their WT counterparts. A decrease in CD28 co-stimulation in CXCR6 KO CD8$^+$ T cells may not only compromise their survival through regulation of Bcl-xL[62–64] but also their ability to respond effectively to ICB, which is known to require CD28 signals.[65] Indeed, a separate study showed that CXCR6-deficient mice had a compromised response to ICB.[56] Our data are thus consistent with a model where the CXCR6 increase upon ICB reinforces CD28 signaling to counterbalance PD1-mediated inhibition.[23] This, together with a recent report showing that CXCR6 mediates interactions with dendritic cells to promote survival of effector cells,[46] portrays a combination of intrinsic and spatial signals that together elicit protective signaling via CXCR6 in chronically activated T cells.

CD8$^+$ T cells undergo an altered differentiation trajectory in the TME.[66,67] *Tcf7*/TCF1 is important early in this trajectory as it restrains effector differentiation and preserves a pool of stem-like CD8$^+$ T cells that seed the effector response upon ICB.[21,22,55] Our data show that TCF1 represses AP-1-driven CXCR6 expression. Thus, as stem-like CD8$^+$ T cells become activated and lose TCF1 expression, AP-1-mediated CXCR6 transcription can occur. CXCR6 expression further increases as cells progress along the effector differentiation trajectory and eventually become dysfunctional. CXCR6 likely functions to preserve and promote effector cells along this trajectory, having its greatest effects on dysfunctional cells, where CXCR6 levels are the highest. Further, TCF1$^+$ stem-like CD8$^+$ T cells have been shown to reside in specific tumoral niches.[68] Their positioning could be in part due to the repression of CXCR6, which has been shown to direct intra-tumoral localization of cells.[46] Thus, TCF1 suppression of CXCR6 in stem-like cells may sequester them in specialized niches that preserve their stemness and prevent effector differentiation. The TCF1:CXCR6 axis we describe may therefore orchestrate the differentiation trajectory of intra-tumoral CD8$^+$ T cells by directing their spatial localization.

Our study focused mainly on the T cell component of the CXCR6-CXCL16 interaction. However, CXCL16 is also expressed as a transmembrane protein.[69–71] CXCR6 binding to transmembrane CXCL16 can mediate cell adhesion[72] and reverse signal into CXCL16-expressing cells to modulate their phenotype as well.[73] Thus, future studies could investigate if CXCR6-CXCL16 interactions could modify CXCL16$^+$ APCs to promote their production of factors such as IL-15,[46] CX3CL1 (the ligand for CX3CR1), or B7 molecules to support the survival and maintenance of chronically activated CD8$^+$ TILs.

In conclusion, our study introduces a method for mixed-effects unsupervised modeling to leverage the large amount of available scRNA-seq data through a unified framework and reveals a regulatory axis essential for T cell-mediated tumor control.

Limitations of the study

Mice with an intact CD8$^+$ T cell compartment were used as recipients for adoptive transfer studies with antigen-specific T cells. Although we examined the endogenous CD8$^+$ T cells in these tumors and found no differences between the experimental and control groups, the contribution of endogenous CD8$^+$ T cells cannot be entirely ruled out. To interrogate the difference between WT and CXCR6 KO cells, we used an *in vitro* co-culture system to better control for differences in cell numbers, exposure to CXCL16 signals, and timing. However, investigating WT and CXCR6 KO cells and their interactions with CXCL16-expressing cells *in vivo* may provide a clearer picture of CXCR6

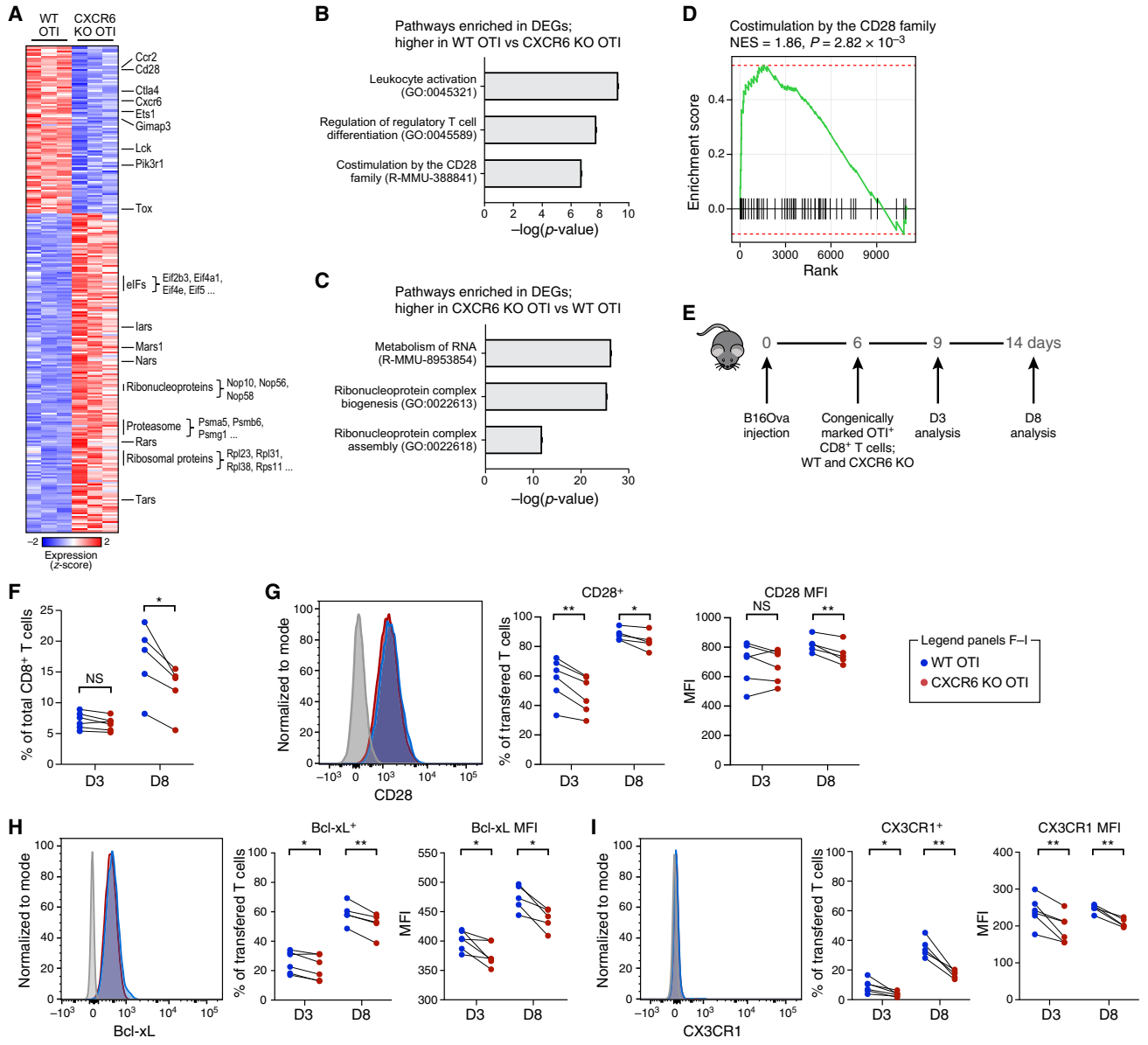


Figure 6. CXCR6 KO cells have reduced CD28 co-stimulation

(A) Heatmap showing differentially expressed genes (rows, *Z* score) between WT and CXCR6 KO OTI T cells (columns) sorted from T cell:DC co-cultures. Genes representative of the top 3 enriched pathways in WT and CXCR6 KO cells are labeled (n = 3, 1 experiment).

(B) The top 3 significantly enriched pathways based on genes that were significantly higher in WT compared to CXCR6 KO OTI T cells from T cell:DC co-cultures.

(C) The top 3 significantly enriched pathways based on genes that were significantly higher in CXCR6 KO compared to WT OTI T cells from T cell:DC co-cultures.

(D) GSEA plot for the "costimulation by the CD28 family" (R-MMU-388841) gene set when compared against the ranking of genes based on their differential expression between WT and CXCR6 KO T OTI cells from T cell:DC co-cultures.

(E) Experimental design for co-adoptive transfer experiments.

(F–I) WT and CXCR6 KO OTI T cells were co-adoptively transferred into B16Ova tumor-bearing mice as indicated in (E) (n = 5–6 per time point, 1 experiment). (F) Frequency (y axis, mean ± SEM) of transferred OTI cells on day 3 (D3) and day 8 (D8) post-transfer (x axis). (G) Expression level (left, x axis, fluorescence intensity) of CD28 in WT and CXCR6 KO OTI CD8$^+$ T cells versus fluorescence minus one (FMO, gray) prior to co-adoptive transfer. Frequency of CD28$^+$ cells (middle, y axis, mean ± SEM) and CD28 expression level in CD28$^+$ cells (right, y axis, geometric mean fluorescence [MFI]) in transferred OTI cells at day 3 (D3) and day 8 (D8) post-transfer. (H) Expression level (left, x axis, fluorescence intensity) of Bcl-xL in WT and CXCR6 KO OTI CD8$^+$ T cells versus FMO (gray) prior to co-adoptive transfer. Frequency of Bcl-xL$^+$ cells (middle, y axis, mean ± SEM) and Bcl-xL expression level in Bcl-xL$^+$ cells (right, y axis, geometric mean fluorescence [MFI]) in transferred OTI cells at day 3 (D3) and day 8 (D8) post-transfer. (I) Expression level (left, x axis, fluorescence intensity) of CX3CR1 in WT and CXCR6 KO OTI CD8$^+$ T cells versus FMO (gray) prior to co-adoptive transfer. Frequency of CX3CR1$^+$ cells (middle, y axis, mean ± SEM) and CX3CR1 expression level in CX3CR1$^+$ cells (right, y axis, geometric mean fluorescence [MFI]) in transferred OTI cells at day 3 (D3) and day 8 (D8) post-transfer. Statistical significance was determined by Student's paired t test (F–I). NS = not significant, *p < 0.05, **p < 0.01.

effects on CD8$^+$ T cells in the TME. Future studies should test the predictive power of the pan-cancer dysfunction program as single-cell or spatial data from larger cohorts of ICB-treated patients become available.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Mice
  - Mouse tumor cell lines
- METHOD DETAILS
  - Generalizable matrix decomposition framework (GMDF)
  - Obtaining a consensus GMDF solution
  - Batch correction
  - Comparing GMDF to LIGER and iNMF
  - Datasets and pre-processing
  - GMDF application to the pan-cancer CD8 T cell cohort
  - Decoupling convolved expression programs
  - CXCL16 pan-cancer macrophage program
  - Predictive performance of the pan-cancer dysfunction program
  - Tumor experiments with immune checkpoint blockade
  - TIL isolation
  - Flow cytometry
  - Adoptive cell transfers
  - Purification, lentiviral transduction, and *in vitro* culture of primary CD8$^+$ T cells for adoptive cell transfer experiments
  - sgRNA design, testing, and plasmid construction
  - Lentivirus production
  - *In vitro* time-course of T cell activation and bulk RNA-seq
  - ChIP-PCR
  - Mouse single cell RNA-seq
  - Mouse single cell RNA-seq analysis
  - Luciferase assays
  - T cell-DC co-culture, bulk RNA-seq, and pathway analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2024.101640.

ACKNOWLEDGMENTS

We thank Leslie Gaffney for help with figure preparation, Yufan Wu for computational input, and Geoffrey Fell, MS for statistical help. We thank the Human Tumor Atlas Network and Human Tumor Atlas Pilot Project consortium for their ongoing support. This work was supported by grants from the National Institutes of Health (R01CA187975; P01CA236749 to A.C.A.), the Klarman Cell Observatory at the Broad Institute, and the Howard Hughes Medical Institute (HHMI). A.C.A. is a recipient of the Brigham and Women's President's Scholar Award. A.R. was an investigator of HHMI. G.E. and L.J. were Cancer Research Institute Irvington Fellows supported by the Cancer Research Institute (G.E. CRI 2934). L.J. was a fellow of the Eric and Wendy Schmidt postdoctoral program and is a Chan Zuckerberg Biohub Investigator and a Paul Allen Distinguished Investigator. L.J. holds a Career Award at the Scientific Interface from Burroughs Wellcome Fund and a Career Award from the Ovarian Cancer Research Alliance.

AUTHOR CONTRIBUTIONS

Conceptualization, L.J., K.T., A.C.A., and A.R.; methodology, L.J., K.T., G.E., A.C.A., and A.R.; investigation, K.T., L.J., G.E., G.D., H.C., and D.M.; formal analysis, L.J., K.T., G.E., S.H.K., C.L., E.G., and A.M.; resources, L.J., K.T., G.E., G.D., H.C., and R.K.; writing, K.T., L.J., A.C.A., and A.R.; supervision, O.R.-R., L.J., A.C.A., and A.R.; funding acquisition, A.C.A. and A.R.

DECLARATION OF INTERESTS

A.C.A. is a member of the SAB for Tizona Therapeutics, Trishula Therapeutics, Compass Therapeutics, ExcepGen, and Zumutor Biologics, which have interests in cancer immunotherapy. A.C.A. is also a paid consultant for iTeos Therapeutics and Larkspur Biosciences. A.C.A.'s interests were reviewed and managed by the Brigham and Women's Hospital and Partners Healthcare in accordance with their conflict-of-interest policies. O.R.-R. is an employee of Genentech. A.R. is a co-founder and equity holder of Celsius Therapeutics, an equity holder in Immunitas, and was an SAB member of Thermo Fisher Scientific, Syros Pharmaceuticals, Neogene Therapeutics, and Asimov. From August 1, 2020, A.R. is an employee of Genentech and has equity in Roche. When at the Broad, A.R.'s interests were reviewed and managed by the Broad Institute, MIT and HHMI in accordance with their conflict-of-interest policies. A provisional patent application was filed including work in this manuscript.

Received: May 31, 2023
Revised: January 5, 2024
Accepted: June 11, 2024
Published: July 2, 2024

REFERENCES

1. Das, R., Verma, R., Sznol, M., Boddupalli, C.S., Gettinger, S.N., Kluger, H., Callahan, M., Wolchok, J.D., Halaban, R., Dhodapkar, M.V., and Dhodapkar, K.M. (2015). Combination therapy with anti-CTLA-4 and anti-PD-1 leads to distinct immunologic changes in vivo. J. Immunol. *194*, 950–959. https://doi.org/10.4049/jimmunol.1401686.
2. Huang, A.C., Postow, M.A., Orlowski, R.J., Mick, R., Bengsch, B., Manne, S., Xu, W., Harmon, S., Giles, J.R., Wenz, B., 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.
3. Kamphorst, A.O., Pillai, R.N., Yang, S., Nasti, T.H., Akondy, R.S., Wieland, A., Sica, G.L., Yu, K., Koenig, L., Patel, N.T., et al. (2017). Proliferation of PD-1+ CD8 T cells in peripheral blood after PD-1-targeted therapy in lung cancer patients. Proc. Natl. Acad. Sci. USA *114*, 4993–4998. https://doi.org/10.1073/pnas.1705327114.
4. Yuan, J., Gnjatic, S., Li, H., Powel, S., Gallardo, H.F., Ritter, E., Ku, G.Y., Jungbluth, A.A., Segal, N.H., Rasalan, T.S., et al. (2008). CTLA-4 Blockade Enhances Polyfunctional NY-ESO-1 Specific T Cell Responses in Metastatic Melanoma Patients with Clinical Benefit. Proc. Natl. Acad. Sci. USA *105*, 20410–20415.
5. Haslam, A., and Prasad, V. (2019). Estimation of the Percentage of US Patients With Cancer Who Are Eligible for and Respond to Checkpoint Inhibitor Immunotherapy Drugs. JAMA Netw. Open *2*, e192535. https://doi.org/10.1001/jamanetworkopen.2019.2535.
6. Hamid, O., Schmidt, H., Nissan, A., Ridolfi, L., Aamdal, S., Hansson, J., Guida, M., Hyams, D.M., Gómez, H., Bastholt, L., et al. (2011). A prospective phase II trial exploring the association between tumor microenvironment biomarkers and clinical activity of ipilimumab in advanced melanoma. J. Transl. Med. *9*, 204. https://doi.org/10.1186/1479-5876-9-204.
7. Hollern, D.P., Xu, N., Thennavan, A., Glodowski, C., Garcia-Recio, S., Mott, K.R., He, X., Garay, J.P., Carey-Ewend, K., Marron, D., et al. (2019). B Cells and T Follicular Helper Cells Mediate Response to Checkpoint Inhibitors in High Mutation Burden Mouse Models of Breast Cancer. Cell *179*, 1191–1206.e21. https://doi.org/10.1016/j.cell.2019.10.028.

8. Jiao, S., Subudhi, S.K., Aparicio, A., Ge, Z., Guan, B., Miura, Y., and Sharma, P. (2019). Differences in Tumor Microenvironment Dictate T Helper Lineage Polarization and Response to Immune Checkpoint Therapy. Cell *179*, 1177–1190.e13. https://doi.org/10.1016/j.cell.2019.10.029.

9. Anderson, K.G., Stromnes, I.M., and Greenberg, P.D. (2017). Obstacles Posed by the Tumor Microenvironment to T cell Activity: A Case for Synergistic Therapies. Cancer Cell *31*, 311–325. https://doi.org/10.1016/j.ccell.2017.02.008.

10. McLane, L.M., Abdel-Hakeem, M.S., and Wherry, E.J. (2019). CD8 T Cell Exhaustion During Chronic Viral Infection and Cancer. Annu. Rev. Immunol. *37*, 457–495. https://doi.org/10.1146/annurev-immunol-041015-055318.

11. Azizi, E., Carr, A.J., Plitas, G., Cornish, A.E., Konopacki, C., Prabhakaran, S., Nainys, J., Wu, K., Kiseliovas, V., Setty, M., et al. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. Cell *174*, 1293–1308.e36. https://doi.org/10.1016/j.cell.2018.05.060.

12. Jerby-Arnon, L., Shah, P., Cuoco, M.S., Rodman, C., Su, M.-J., Melms, J.C., Leeson, R., Kanodia, A., Mei, S., Lin, J.-R., et al. (2018). A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. Cell *175*, 984–997.e24. https://doi.org/10.1016/j.cell.2018.09.006.

13. Jerby-Arnon, L., Neftel, C., Shore, M.E., Weisman, H.R., Mathewson, N.D., McBride, M.J., Haas, B., Izar, B., Volorio, A., Boulay, G., et al. (2021). Opposing immune and genetic mechanisms shape oncogenic programs in synovial sarcoma. Nat. Med. *27*, 289–300. https://doi.org/10.1038/s41591-020-01212-6.

14. Kim, N., Kim, H.K., Lee, K., Hong, Y., Cho, J.H., Choi, J.W., Lee, J.-I., Suh, Y.-L., Ku, B.M., Eum, H.H., et al. (2020). Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. Nat. Commun. *11*, 2285. https://doi.org/10.1038/s41467-020-16164-1.

15. Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., Bassez, A., Decaluwé, H., Pircher, A., Van den Eynde, K., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. Nat. Med. *24*, 1277–1289. https://doi.org/10.1038/s41591-018-0096-5.

16. Mathewson, N.D., Ashenberg, O., Tirosh, I., Gritsch, S., Perez, E.M., Marx, S., Jerby-Arnon, L., Chanoch-Myers, R., Hara, T., Richman, A.R., et al. (2021). Inhibitory CD161 receptor identified in glioma-infiltrating T cells by single-cell analysis. Cell *184*, 1281–1298.e26. https://doi.org/10.1016/j.cell.2021.01.022.

17. Peng, J., Sun, B.-F., Chen, C.-Y., Zhou, J.-Y., Chen, Y.-S., Chen, H., Liu, L., Huang, D., Jiang, J., Cui, G.-S., et al. (2019). Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. Cell Res. *29*, 725–738. https://doi.org/10.1038/s41422-019-0195-y.

18. Sade-Feldman, M., Yizhak, K., Bjorgaard, S.L., Ray, J.P., de Boer, C.G., Jenkins, R.W., Lieb, D.J., Chen, J.H., Frederick, D.T., Barzily-Rokni, M., et al. (2018). Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. Cell *175*, 998–1013.e20. https://doi.org/10.1016/j.cell.2018.10.038.

19. Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q., Huang, J.Y., et al. (2018). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature *564*, 268–272. https://doi.org/10.1038/s41586-018-0694-x.

20. Zheng, C., Zheng, L., Yoo, J.-K., Guo, H., Zhang, Y., Guo, X., Kang, B., Hu, R., Huang, J.Y., Zhang, Q., 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.

21. Siddiqui, I., Schaeuble, K., Chennupati, V., Fuertes Marraco, S.A., Calderon-Copete, S., Pais Ferreira, D., Carmona, S.J., Scarpellino, L., Gfeller, D., Pradervand, S., 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.

22. Kurtulus, S., Madi, A., Escobar, G., Klapholz, M., Nyman, J., Christian, E., Pawlak, M., Dionne, D., Xia, J., Rozenblatt-Rosen, O., et al. (2019). Checkpoint Blockade Immunotherapy Induces Dynamic Changes in PD-1-CD8 + Tumor-Infiltrating T Cells. Immunity *50*, 181–194.e6. https://doi.org/10.1016/j.immuni.2018.11.014.

23. Hui, E., Cheung, J., Zhu, J., Su, X., Taylor, M.J., Wallweber, H.A., Sasmal, D.K., Huang, J., Kim, J.M., Mellman, I., and Vale, R.D. (2017). T cell costimulatory receptor CD28 is a primary target for PD-1–mediated inhibition. Science *355*, 1428–1433. https://doi.org/10.1126/science.aaf1292.

24. Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., et al. (2021). Pan-cancer single-cell landscape of tumor-infiltrating T cells. Science *374*, abe6474. https://doi.org/10.1126/science.abe6474.

25. Chu, Y., Dai, E., Li, Y., Han, G., Pei, G., Ingram, D.R., Thakkar, K., Qin, J.-J., Dang, M., Le, X., et al. (2023). Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. Nat. Med. *29*, 1550–1562. https://doi.org/10.1038/s41591-023-02371-y.

26. Yang, Z., and Michailidis, G. (2016). A non-negative matrix factorization method for detecting modules in heterogeneous omics multi-modal data. Bioinformatics *32*, 1–8. https://doi.org/10.1093/bioinformatics/btv544.

27. Welch, J.D., Kozareva, V., Ferreira, A., Vanderburg, C., Martin, C., and Macosko, E.Z. (2019). Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. Cell *177*, 1873–1887.e17. https://doi.org/10.1016/j.cell.2019.05.006.

28. Khan, O., Giles, J.R., McDonald, S., Manne, S., Ngiow, S.F., Patel, K.P., Werner, M.T., Huang, A.C., Alexander, K.A., Wu, J.E., et al. (2019). TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion. Nature *571*, 211–218. https://doi.org/10.1038/s41586-019-1325-x.

29. Scott, A.C., Dündar, F., Zumbo, P., Chandran, S.S., Klebanoff, C.A., Shakiba, M., Trivedi, P., Menocal, L., Appleby, H., Camara, S., et al. (2019). TOX is a critical regulator of tumour-specific T cell differentiation. Nature *571*, 270–274. https://doi.org/10.1038/s41586-019-1324-y.

30. Seo, H., Chen, J., González-Avalos, E., Samaniego-Castruita, D., Das, A., Wang, Y.H., López-Moyado, I.F., Georges, R.O., Zhang, W., Onodera, A., et al. (2019). TOX and TOX2 transcription factors cooperate with NR4A transcription factors to impose $CD8^+$ T cell exhaustion. Proc. Natl. Acad. Sci. USA *116*, 12410–12415. https://doi.org/10.1073/pnas.1905675116.

31. Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., 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.

32. Li, H., van der Leun, A.M., Yofe, I., Lubling, Y., Gelbard-Solodkin, D., van Akkooi, A.C.J., van den Braber, M., Rozeman, E.A., Haanen, J.B.A.G., Blank, C.U., 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.

33. Singer, M., Wang, C., Cong, L., Marjanovic, N.D., Kowalczyk, M.S., Zhang, H., Nyman, J., Sakuishi, K., Kurtulus, S., Gennert, D., et al. (2016). A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells. Cell *166*, 1500–1511.e9. https://doi.org/10.1016/j.cell.2016.08.052.

34. Qian, J., Olbrecht, S., Boeckx, B., Vos, H., Laoui, D., Etlioglu, E., Wauters, E., Pomella, V., Verbandt, S., Busschaert, P., et al. (2020). A pan-cancer blueprint of the heterogeneous tumor microenvironment revealed by single-cell profiling. Cell Res. *30*, 745–762. https://doi.org/10.1038/s41422-020-0355-0.

35. Olbrecht, S., Busschaert, P., Qian, J., Vanderstichele, A., Loverix, L., Van Gorp, T., Van Nieuwenhuysen, E., Han, S., Van den Broeck, A., Coosemans, A., et al. (2021). High-grade serous tubo-ovarian cancer refined with single-cell RNA sequencing: specific cell subtypes influence survival

and determine molecular subtype classification. Genome Med. *13*, 111. https://doi.org/10.1186/s13073-021-00922-x.

36. Bouguermouh, S., Van, V.Q., Martel, J., Gautier, P., Rubio, M., and Sarfati, M. (2008). CD47 Expression on T Cell Is a Self-Control Negative Regulator of Type 1 Immune Response. J. Immunol. *180*, 8073–8082. https://doi.org/10.4049/jimmunol.180.12.8073.
37. Frangieh, C.J., Melms, J.C., Thakore, P.I., Geiger-Schuller, K.R., Ho, P., Luoma, A.M., Cleary, B., Jerby-Arnon, L., Malu, S., Cuoco, M.S., et al. (2021). Multimodal pooled Perturb-CITE-seq screens in patient models define mechanisms of cancer immune evasion. Nat. Genet. *53*, 332–341. https://doi.org/10.1038/s41588-021-00779-1.
38. Liu, X., Pu, Y., Cron, K., Deng, L., Kline, J., Frazier, W.A., Xu, H., Peng, H., Fu, Y.-X., and Xu, M.M. (2015). CD47 blockade triggers T cell–mediated destruction of immunogenic tumors. Nat. Med. *21*, 1209–1215. https://doi.org/10.1038/nm.3931.
39. Zhang, Y., Chen, H., Mo, H., Hu, X., Gao, R., Zhao, Y., Liu, B., Niu, L., Sun, X., Yu, X., et al. (2021). Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. Cancer Cell *39*, 1578–1593.e8. https://doi.org/10.1016/j.ccell.2021.09.010.
40. Yost, K.E., Satpathy, A.T., Wells, D.K., Qi, Y., Wang, C., Kageyama, R., McNamara, K.L., Granja, J.M., Sarin, K.Y., Brown, R.A., et al. (2019). Clonal replacement of tumor-specific T cells following PD-1 blockade. Nat. Med. *25*, 1251–1259. https://doi.org/10.1038/s41591-019-0522-3.
41. Lee, L.N., Ronan, E.O., de Lara, C., Franken, K.L.M.C., Ottenhoff, T.H.M., Tchilian, E.Z., and Beverley, P.C.L. (2011). CXCR6 is a marker for protective antigen-specific cells in the lungs after intranasal immunization against Mycobacterium tuberculosis. Infect. Immun. *79*, 3328–3337. https://doi.org/10.1128/IAI.01133-10.
42. Sato, T., Thorlacius, H., Johnston, B., Staton, T.L., Xiang, W., Littman, D.R., and Butcher, E.C. (2005). Role for CXCR6 in recruitment of activated CD8+ lymphocytes to inflamed liver. J. Immunol. *174*, 277–283. https://doi.org/10.4049/jimmunol.174.1.277.
43. Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing, R., 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.
44. Daniel, B., Yost, K.E., Hsiung, S., Sandor, K., Xia, Y., Qi, Y., Hiam-Galvez, K.J., Black, M., J. Raposo, C., Shi, Q., et al. (2022). Divergent clonal differentiation trajectories of T cell exhaustion. Nat. Immunol. *23*, 1614–1627. https://doi.org/10.1038/s41590-022-01337-5.
45. Sandu, I., Cerletti, D., Oetiker, N., Borsa, M., Wagen, F., Spadafora, I., Welten, S.P.M., Stolz, U., Oxenius, A., and Claassen, M. (2020). Landscape of Exhausted Virus-Specific CD8 T Cells in Chronic LCMV Infection. Cell Rep. *32*, 108078. https://doi.org/10.1016/j.celrep.2020.108078.
46. Di Pilato, M., Kfuri-Rubens, R., Pruessmann, J.N., Ozga, A.J., Messemaker, M., Cadilha, B.L., Sivakumar, R., Cianciaruso, C., Warner, R.D., Marangoni, F., et al. (2021). CXCR6 positions cytotoxic T cells to receive critical survival signals in the tumor microenvironment. Cell *184*, 4512–4530.e22. https://doi.org/10.1016/j.cell.2021.07.015.
47. Neftel, C., Laffy, J., Filbin, M.G., Hara, T., Shore, M.E., Rahme, G.J., Richman, A.R., Silverbush, D., Shaw, M.L., Hebert, C.M., et al. (2019). An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. Cell *178*, 835–849.e21. https://doi.org/10.1016/j.cell.2019.06.024.
48. Gerlach, C., Moseman, E.A., Loughhead, S.M., Alvarez, D., Zwijnenburg, A.J., Waanders, L., Garg, R., de la Torre, J.C., and von Andrian, U.H. (2016). The Chemokine Receptor CX3CR1 Defines Three Antigen-Experienced CD8 T Cell Subsets with Distinct Roles in Immune Surveillance and Homeostasis. Immunity *45*, 1270–1284. https://doi.org/10.1016/j.immuni.2016.10.018.
49. Hudson, W.H., Gensheimer, J., Hashimoto, M., Wieland, A., Valanparambil, R.M., Li, P., Lin, J.-X., Konieczny, B.T., Im, S.J., Freeman, G.J., et al. (2019). Proliferating Transitory T Cells with an Effector-like Transcriptional Signature Emerge from PD-1+ Stem-like CD8+ T Cells during Chronic Infection. Immunity *51*, 1043–1058.e4. https://doi.org/10.1016/j.immuni.2019.11.002.
50. Zander, R., Schauder, D., Xin, G., Nguyen, C., Wu, X., Zajac, A., and Cui, W. (2019). CD4+ T Cell Help Is Required for the Formation of a Cytolytic CD8+ T Cell Subset that Protects against Chronic Infection and Cancer. Immunity *51*, 1028–1042.e4. https://doi.org/10.1016/j.immuni.2019.10.009.
51. van der Leun, A.M., Thommen, D.S., and Schumacher, T.N. (2020). CD8 + T cell states in human cancer: insights from single-cell analysis. Nat. Rev. Cancer *20*, 218–232. https://doi.org/10.1038/s41568-019-0235-4.
52. Danilo, M., Chennupati, V., Silva, J.G., Siegert, S., and Held, W. (2018). Suppression of Tcf1 by Inflammatory Cytokines Facilitates Effector CD8 T Cell Differentiation. Cell Rep. *22*, 2107–2117. https://doi.org/10.1016/j.celrep.2018.01.072.
53. Tiemessen, M.M., Baert, M.R.M., Kok, L., van Eggermond, M.C.J.A., van den Elsen, P.J., Arens, R., and Staal, F.J.T. (2014). T Cell Factor 1 Represses CD8+ Effector T Cell Formation and Function. J. Immunol. *193*, 5480–5487. https://doi.org/10.4049/jimmunol.1303417.
54. Im, S.J., Hashimoto, M., Gerner, M.Y., Lee, J., Kissick, H.T., Burger, M.C., Shan, Q., Hale, J.S., Lee, J., Nasti, T.H., 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.
55. Miller, B.C., Sen, D.R., Al Abosy, R., Bi, K., Virkud, Y.V., LaFleur, M.W., Yates, K.B., Lako, A., Felt, K., Naik, G.S., et al. (2019). Subsets of exhausted CD8+ T cells differentially mediate tumor control and respond to checkpoint blockade. Nat. Immunol. *20*, 326–336. https://doi.org/10.1038/s41590-019-0312-6.
56. Wang, B., Wang, Y., Sun, X., Deng, G., Huang, W., Wu, X., Gu, Y., Tian, Z., Fan, Z., Xu, Q., et al. (2021). CXCR6 is required for antitumor efficacy of intratumoral CD8+ T cell. J. Immunother. Cancer *9*, e003100. https://doi.org/10.1136/jitc-2021-003100.
57. Zhao, X., Shan, Q., and Xue, H.-H. (2022). TCF1 in T cell immunity: a broadened frontier. Nat. Rev. Immunol. *22*, 147–157. https://doi.org/10.1038/s41577-021-00563-6.
58. Steinke, F.C., Yu, S., Zhou, X., He, B., Yang, W., Zhou, B., Kawamoto, H., Zhu, J., Tan, K., and Xue, H.-H. (2014). TCF-1 and LEF-1 act upstream of Th-POK to promote the CD4(+) T cell fate and interact with Runx3 to silence Cd4 in CD8(+) T cells. Nat. Immunol. *15*, 646–656. https://doi.org/10.1038/ni.2897.
59. Philip, M., Fairchild, L., Sun, L., Horste, E.L., Camara, S., Shakiba, M., Scott, A.C., Viale, A., Lauer, P., Merghoub, T., et al. (2017). Chromatin states define tumour-specific T cell dysfunction and reprogramming. Nature *545*, 452–456. https://doi.org/10.1038/nature22367.
60. Landsman, L., Bar-On, L., Zernecke, A., Kim, K.-W., Krauthgamer, R., Shagdarsuren, E., Lira, S.A., Weissman, I.L., Weber, C., and Jung, S. (2009). CX3CR1 is required for monocyte homeostasis and atherogenesis by promoting cell survival. Blood *113*, 963–972. https://doi.org/10.1182/blood-2008-07-170787.
61. Zheng, J., Yang, M., Shao, J., Miao, Y., Han, J., and Du, J. (2013). Chemokine receptor CX3CR1 contributes to macrophage survival in tumor metastasis. Mol. Cancer *12*, 141. https://doi.org/10.1186/1476-4598-12-141.
62. Boise, L.H., Minn, A.J., Noel, P.J., June, C.H., Accavitti, M.A., Lindsten, T., and Thompson, C.B. (1995). CD28 costimulation can promote T cell survival by enhancing the expression of Bcl-XL. Immunity *3*, 87–98. https://doi.org/10.1016/1074-7613(95)90161-2.
63. Burr, J.S., Savage, N.D., Messah, G.E., Kimzey, S.L., Shaw, A.S., Arch, R.H., and Green, J.M. (2001). Cutting edge: distinct motifs within CD28 regulate T cell proliferation and induction of Bcl-XL. J. Immunol. *166*, 5331–5335. https://doi.org/10.4049/jimmunol.166.9.5331.
64. Okkenhaug, K., Wu, L., Garza, K.M., La Rose, J., Khoo, W., Odermatt, B., Mak, T.W., Ohashi, P.S., and Rottapel, R. (2001). A point mutation in CD28

distinguishes proliferative signals from survival signals. Nat. Immunol. *2*, 325–332. https://doi.org/10.1038/86327.

65. Kamphorst, A.O., Wieland, A., Nasti, T., Yang, S., Zhang, R., Barber, D.L., Konieczny, B.T., Daugherty, C.Z., Koenig, L., Yu, K., et al. (2017). Rescue of exhausted CD8 T cells by PD-1 targeted therapies is CD28-dependent. Science *355*, 1423–1427. https://doi.org/10.1126/science.aaf0683.

66. Philip, M., and Schietinger, A. (2022). CD8+ T cell differentiation and dysfunction in cancer. Nat. Rev. Immunol. *22*, 209–223. https://doi.org/10.1038/s41577-021-00574-3.

67. Tooley, K.A., Escobar, G., and Anderson, A.C. (2022). Spatial determinants of CD8+ T cell differentiation in cancer. Trends Cancer *8*, 642–654. https://doi.org/10.1016/j.trecan.2022.04.003.

68. Jansen, C.S., Prokhnevska, N., Master, V.A., Sanda, M.G., Carlisle, J.W., Bilen, M.A., Cardenas, M., Wilkinson, S., Lake, R., Sowalsky, A.G., et al. (2019). An intra-tumoral niche maintains and differentiates stem-like CD8 T cells. Nature *576*, 465–470. https://doi.org/10.1038/s41586-019-1836-5.

69. Abel, S., Hundhausen, C., Mentlein, R., Schulte, A., Berkhout, T.A., Broadway, N., Hartmann, D., Sedlacek, R., Dietrich, S., Muetze, B., et al. (2004). The transmembrane CXC-chemokine ligand 16 is induced by IFN-gamma and TNF-alpha and shed by the activity of the disintegrin-like metalloproteinase ADAM10. J. Immunol. *172*, 6362–6372. https://doi.org/10.4049/jimmunol.172.10.6362.

70. Gough, P.J., Garton, K.J., Wille, P.T., Rychlewski, M., Dempsey, P.J., and Raines, E.W. (2004). A disintegrin and metalloproteinase 10-mediated cleavage and shedding regulates the cell surface expression of CXC chemokine ligand 16. J. Immunol. *172*, 3678–3685. https://doi.org/10.4049/jimmunol.172.6.3678.

71. Schramme, A., Abdel-Bakky, M.S., Kämpfer-Kolb, N., Pfeilschifter, J., and Gutwein, P. (2008). The role of CXCL16 and its processing metalloproteinases ADAM10 and ADAM17 in the proliferation and migration of human mesangial cells. Biochem. Biophys. Res. Commun. *370*, 311–316. https://doi.org/10.1016/j.bbrc.2008.03.088.

72. Shimaoka, T., Nakayama, T., Fukumoto, N., Kume, N., Takahashi, S., Yamaguchi, J., Minami, M., Hayashida, K., Kita, T., Ohsumi, J., et al. (2004). Cell surface-anchored SR-PSOX/CXC chemokine ligand 16 mediates firm adhesion of CXC chemokine receptor 6-expressing cells. J. Leukoc. Biol. *75*, 267–274. https://doi.org/10.1189/jlb.1003465.

73. Adamski, V., Mentlein, R., Lucius, R., Synowitz, M., Held-Feindt, J., and Hattermann, K. (2017). The Chemokine Receptor CXCR6 Evokes Reverse Signaling via the Transmembrane Chemokine CXCL16. Int. J. Mol. Sci. *18*, 1468. https://doi.org/10.3390/ijms18071468.

74. Kim, J., He, Y., and Park, H. (2014). Algorithms for nonnegative matrix and tensor factorizations: a unified view based on block coordinate descent framework. J. Global Optim. *58*, 285–319. https://doi.org/10.1007/s10898-013-0035-4.

75. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Software *67*, 1–48.

76. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. J. Stat. Software *82* (*13*), 1–26. https://doi.org/10.18637/jss.v082.i13.

77. Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res. *42*, e168. https://doi.org/10.1093/nar/gku936.

78. Follenzi, A., Ailles, L.E., Bakovic, S., Geuna, M., and Naldini, L. (2000). Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. Nat. Genet. *25*, 217–222. https://doi.org/10.1038/76095.

79. Picelli, S., Björklund, Å.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods *10*, 1096–1098. https://doi.org/10.1038/nmeth.2639.

80. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf. *12*, 323. https://doi.org/10.1186/1471-2105-12-323.

81. Love, M.I., Huber, W., and Anders, S. (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.

82. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. J. Stat. Mech. *2008*, P10008–P10012.

83. Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, E.a.D., Tadmor, M.D., Litvin, O., Fienberg, H.G., Jager, A., Zunder, E.R., 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.

84. Shekhar, K., Lapan, S.W., Whitney, I.E., Tran, N.M., Macosko, E.Z., Kowalczyk, M., Adiconis, X., Levin, J.Z., Nemesh, J., Goldman, M., et al. (2016). Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. Cell *166*, 1308–1323.e30. https://doi.org/10.1016/j.cell.2016.07.054.

85. Farré, D., Roset, R., Huerta, M., Adsuara, J.E., Roselló, L., Albà, M.M., and Messeguer, X. (2003). Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. Nucleic Acids Res. *31*, 3651–3653. https://doi.org/10.1093/nar/gkg605.

86. Messeguer, X., Escudero, R., Farré, D., Núñez, O., Martínez, J., and Albà, M.M. (2002). PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics *18*, 333–334. https://doi.org/10.1093/bioinformatics/18.2.333.

87. Duong, E., Fessenden, T.B., Lutz, E., Dinter, T., Yim, L., Blatt, S., Bhutkar, A., Wittrup, K.D., and Spranger, S. (2022). Type I interferon activates MHC class I-dressed CD11b+ conventional dendritic cells to promote protective anti-tumor CD8+ T cell immunity. Immunity *55*, 308–323.e9. https://doi.org/10.1016/j.immuni.2021.10.020.

88. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523. https://doi.org/10.1038/s41467-019-09234-6.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Mouse CD8 BUV805 Clone 53-6.7	BD	Cat:# 612898 RRID: AB_2870186
Anti-Mouse CD11c Bv605 Clone N418	Biolegend	Cat:# 117334 RRID: AB_2562415
Anti-Mouse CD45.2 Percp5.5 Clone 104	Biolegend	Cat:# 109828 RRID: AB_893350
Anti-Mouse CD45.2 Bv785 Clone 104	Biolegend	Cat:# 109839 RRID: AB_2562604
Anti-Mouse CD24 BUV496 Clone M1/69	BD	Cat:# 612953; RRID: AB_2870229
Anti-Mouse CD45.1 APCFire-750 Clone A20	Biolegend	Cat:# 110752 RRID: AB_2629806
Anti-Mouse/Human Granzyme B FITC Clone GB11	Biolegend	Cat:# 515403 RRID: AB_2114575
Anti-Mouse/Human Granzyme B Pacific Blue Clone GB11	Biolegend	Cat:# 515408 RRID: AB_2562196
Anti-Mouse IFN-y BV421 Clone XMG1.2	Biolegend	Cat:# 505830 RRID: AB_2563105
Anti-Mouse IFN-y BV785 Clone XMG1.2	Biolegend	Cat:# 505838 RRID: AB_2629667
Anti-Mouse IL-2 Alexa Flour 700 Clone JES6-5H4	Biolegend	Cat:# 503818 RRID: AB_528931
Anti-Mouse CD45.1 Bv421 Clone A20	Biolegend	Cat:# 110732 RRID: AB_2562563
Anti-Mouse CD64 FITC Clone X54-5/7.1	Biolegend	Cat:# 139316 RRID: AB_2566556
Anti-Mouse PD1 Bv605 Clone 29F.1A12	Biolegend	Cat:# 135220 RRID: AB_2562616
Anti-Mouse CD8a BV605 Clone 53-6.7	Biolegend	Cat:# 100744 RRID: AB_2562609
Anti-Mouse Tim3 BUV395 Clone 5D12	BD	Cat:# 747620 RRID: AB_2744186
Anti-Mouse Tim3 APC Clone 5D12	BD	Cat:# 567164 RRID: AB_2916481
Anti-Mouse TNF-a PE-CY7 Clone MP6-XT22	Biolegend	Cat:# 506324 RRID: AB_2256076
Anti-Mouse CX3CR1 Bv785 Clone SA011F11	Biolegend	Cat:# 149029 RRID: AB_2565938
Anti-Mouse CX3CR1 Bv421 Clone SA011F11	Biolegend	Cat:# 149023 RRID: AB_2565706
Anti-Mouse CCR7 BV421 Clone 4B12	Biolegend	Cat:# 120119 RRID: AB_10897811
Anti-Mouse CD39 PE-Cy7 Clone Duha59	Biolegend	Cat:# 143806 RRID: AB_2563394
Anti-Mouse CD107a APC Clone ID4B	Biolegend	Cat:# 121614 RRID: AB_2234505
Anti-Mouse/Rat Bcl-2 PE-Cy7 Clone BCL/10C4	Biolegend	Cat:# 633512 RRID: AB_2565247

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Anti-Mouse/Rat/Human Bcl-XL PE Clone H-5	Santa Cruz Biotechnology	Cat:# sc-8392 PE; RRID: AB_626739
Anti-Mouse CXCR6 BV711 Clone SA051D1	Biolegend	Cat:# 151111 RRID: AB_2721558
Anti-Mouse CXCR6 PE Clone SA051D1	Biolegend	Cat:# 151104 RRID: AB_2566546
Anti-Mouse CD28 PE-CY7 Clone 37.51	Biolegend	Cat:# 102126 RRID: AB_2617011
Anti-Mouse Lag3 BV650 Clone C9B7W	Biolegend	Cat:# 125227 RRID: AB_2687209
Anti-Mouse I-A/I-E PerCP/Cy5.5 Clone M5/114.15.2	Biolegend	Cat:# 107626 RRID: AB_2191071
Anti-Mouse TCR B chain APC/Cy7 Clone H57-597	Biolegend	Cat:# 109220 RRID: AB_893624
Anti-Mouse CD90.1 (Thy1.1) APC/Fire 750 Clone OX-7	Biolegend	Cat:# 202544 RRID: AB_2650817
Anti-Mouse/Rat XCR1 BV785 Clone ZET	Biolegend	Cat:# 148225 RRID: AB_2783119
Anti-Mouse CD90.1 (Thy1.1) BUV496 Clone OX-7	BD	Cat:# 741110 RRID: AB_2870702
Anti-Mouse CXCL16 PE Clone 12-81	BD	Cat:# 566740 RRID: AB_2869842
Anti-Mouse Tim3 BV421 Clone 5D12	BD	Cat:# 747626 RRID: AB_2744192
Anti-Mouse Ly6C AF700 Clone HK1.4	Biolegend	Cat:# 128024 RRID: AB_10643270
Anti-Mouse CD45 BUV563 Clone 30-F11	BD	Cat:# 612924 RRID: AB_2722550
Anti-Mouse CD103 PE-CY7 Clone 2E7	Biolegend	Cat:# 121426 RRID: AB_2563691
Anti-Mouse TCR b-chain BUV737 Clone H57 597	BD	Cat:# 564799 RRID: AB_2738958
Anti-Mouse CD11b BUV737 Clone M1/70	eBioscience	Cat:# 367-0112-82 RRID: AB_2895933
Anti Ki-67 Alexa Fluor 700 Clone SolA15	eBioscience	Cat:# 56-5698-82 RRID: AB_2637480
Anti-Mouse TIGIT PerCP-eFluor 710 Clone GIGD7	eBioscience	Cat:# 46-9501-82 RRID: AB_11150967
OVA-dextramer PE	Immudex	Cat:# JD2163-PE
Anti-Mouse F4/80 APC Clone CI:A3-1	BioRad	Cat:# MCA497APC RRID: AB_2098197
Anti-Mouse/Human TCF1 PE Clone C63D9	Cell Signaling Technology	Cat:# 14456S RRID: AB_2798483
Anti-Mouse/Human TCF1 Clone C63D9	Cell Signaling Technology	Cat:# 2203S; RRID: AB_2199302
Normal Rabbit IgG #2729	Cell Signaling Technology	Cat:# 2729S; RRID: AB_1031062
Anti-Mouse/Human TOX APC Clone REA473	Miltenyi Biotec	Cat:# 130-118-335 RRID: AB_2751485
Anti-Mouse CD28 Clone PV-1	Bio X Cell	Cat:# BE0015-5 RRID: AB_1107628
Anti-Mouse CD3 Clone 145-2C11	Bio X Cell	Cat:# BP0001-1 RRID: AB_1107634
Anti-Mouse PDL1 Clone 10F.9G2	Bio X Cell	Cat:# BE0101 RRID: AB_10949073

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Anti-Mouse Tim3 Clone RMT3-23	Bio X Cell	Cat:# BE0115 RRID: AB_10949464
Anti-Mouse PD1 Clone RMP1-14	Bio X Cell	Cat:# BE0146; RRID: AB_10949053
Bacterial and virus strains		
One Shot™ Stbl3™ Chemically Competent E. coli	Invitrogen	Cat:#C737303
VSV-pseudotyped third-generation lentiviruses		N/A
Chemicals, peptides, and recombinant proteins		
DMEM	GIBCO	Cat#:11-965-118
RPMI	Thermo Scientific	Cat:# 11875119
NEAA	GIBCO	Cat:# 11140050
Sodium Pyruvate	GIBCO	Cat:# 11360070
Fetal Bovine Serum	SIGMA	Cat:#F2442
Collagenase D	Sigma-Aldrich	Cat:# 11088882001
HEPES Solution	Sigma-Aldrich	Cat:#H0887-100ML
L-Glutamine	GIBCO	Cat:# 25030081
Penicillin-Streptomycin	Sigma-Aldrich	Cat:#P4458-100ML
MEM Vitamin Solution	GIBCO	Cat:# 11120052
TCL Buffer	Qiagen	Cat:# 1031576
Protein A Agarose/Salmon Sperm DNA beads	Millipore Sigma	Cat:# 16-157
Polyjet	SignaGen Laboratories	Cat:# SL100688
DMXAA	Invivogen	Cat:# tlrl-dmx
Purified anti-mouse CD16/32	Biolegend	Cat:# 101302
Beta-mercaptoethanol	SIGMA	Cat:#M3148
PowerTrack SYBR Green Master Mix for qPCR	Applied Biosystems	Cat:# A46109
Recombinant mouse GM-CSF	Biolegend	Cat:# 576304
Recombinant mouse IL-2	Miltenyi Biotec	Cat:# 130-120-662
Recombinant human FLT3L-Ig	Bio X Cell	Cat:# BE0098
CountBright™ Absolute Counting Beads	Thermo Fisher	Cat:#C36950
Zombie UV fixable viability dye	Biolegend	Cat:# 423108
7-AAD Viability Staining Solution	Biolegend	Cat:# 420404
SIINFEKL peptide	Genscript	Cat:# RP10611
FastDigest Esp3I (BsmBI)	Thermo Scientific	Cat:# FERFD0454
Recombinant Mouse IL-12 protein	R&D Systems	Cat:# 419-ML-010/CF
Critical commercial assays		
FoxP3 Transcription Factor Staining Buffer Set	eBioscience	Cat:# 00-5523-00
CD8a microbeads	Miltenyi Biotec	Cat:# 130-117-044
QIAquick PCR Purification Kit	Qiagen	Cat:# 28104
NEBuilder HiFi DNA Assembly	New England Biolabs	Cat:#E5520S
Dual-Luciferase Reporter Assay System	Promega	Cat:#E1910
PE Annexin V Apoptosis Dectection Kit I	BD	Cat:# 559763
Deposited data		
Bulk RNA-seq data from *in vitro* T cell activation timecoursa	This manuscript	GEO: GSE234974
scRNA-seq data from mouse B16Ova	This manuscript	GEO: GSE234975
Bulk RNA-seq data from T cell:DC co-culture	This manuscript	GEO: GSE234976
Human Glioblastoma scRNA-seq data	Mathewson et al.[16]	GEO: GSE163108
Human Glioblastoma scRNA-seq data	Neftel et al.[47]	GEO: GSE131928

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human Melanoma scRNA-seq data	Jerby-Arnon et al.[12]	GEO: GSE115978
Human Melanoma scRNA-seq data	Sade-Feldman et al.[18]	GEO: GSE120575
Human Sarcoma scRNA-seq data	Jerby-Arnon et al.[13]	GEO: GSE131309
Human Lung Cancer scRNA-seq data	Kim et al.[14]	GEO: GSE131907
Human Lung Cancer scRNA-seq data	Lambrechts et al.[15]	Scope: https://gbiomed.kuleuven.be/scRNAseq-NSCLC
Human Breast Cancer scRNA-seq data	Azizi et al.[11]	GEO: GSE114724
Human Liver Cancer scRNA-seq data	Zheng et al.[20]	GEO: GSE98638
Human Colorectal Cancer scRNA-seq data	Zhang et al.[19]	GEO: GSE108989
Human PDAC scRNA-seq data	Peng et al.[17]	GSA: CRA001160
Human Ovarian Cancer scRNA-seq data	Qian et al.,[34] Olbrecht et al.[35]	http://blueprint.lambrechtslab.org
Experimental models: Cell lines		
B16F10 melanoma	ATCC	Cat: #CRL-6475
Mc38Ova^{hi} colon carcinoma	Laboratory of Dr. Nick Haining	N/A
B16Ova melanoma	Laboratory of Dr. Kai Wucherpfennig	N/A
293T/17 [HEK 293T/17]	ATCC	Cat: # CRL-11268
Experimental models: Organisms/strains		
C57BL/6 mice	Jackson laboratory	000664
C57BL/6 CD45.1 mice	Jackson laboratory	002014
OTI mice	Jackson laboratory	003831
CXCR6 KO	Jackson laboratory	005693
Cas9 mice	Jackson laboratory	028555
E8i-CRE+ x Tcf7fl/fl mice	Previously generated in our lab[5]	N/A
CXCR6 KO x OTI mice	This manuscript	N/A
Cas9 x OTI mice	This manuscript	N/A
Cas9 x CD45.1 x CD45.2 mice	This manuscript	N/A
Oligonucleotides		
Cxcr6 locus Chip-PCR forward primer	5′- GAGGCAGACCT TTAGTGAGCA	N/A
Cxcr6 locus Chip-PCR reverse primer	5′- TAGCTCGCACC GTATACACA	N/A
Recombinant DNA		
pCK005_U6-Sp-sgRNA_UP-TdTomato_WPRE	Addgene	Cat:# 85453
pGL4.10[luc2] Luciferase Reporter Vector	Promega	Cat:#E6651
Fos cDNA ORF clone in pcDNA3.1+/C-(K)-DYK vector (Clone ID OMu22131D)	Genscript	Cat: # NM_010234.3
JunD cDNA ORF clone (subcloned to include the first 144 base pairs) in pcDNA3.1+/C-(K)-DYK vector (Clone ID OMu14151D)	Genscript	N/A
Jun cDNA ORF clone in pcDNA3.1+/C-(K)-DYK vector (Clone ID OMu19159D)	Genscript	Cat: #NM_010591.2
TCF1 protein coding sequence (ENSMUST00000109071.3) in pcDNA3.1+/C-HA vector	Genscript	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
FlowJo software	Tree Star	https://flowjo.com
Metascape		https://metascape.org/gp/index.html#/main/step1
Morpheus	Broad Institute	https://software.broadinstitute.org/morpheus
Prism 9	GraphPad Software, Inc	https://www.graphpad.com/
CellRanger software	10X Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome
SEURAT V4.0.1		https://satijalab.org/seurat/
Promo3		https://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3
GPP sgRNA Designer	Broad Institute	https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design
Bioconductor package DESeq2 in R		https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Bioconductor package fgsea in R		https://bioconductor.org/packages/release/bioc/html/fgsea.html
TIDE: Tracking of Indels by Decomposition		http://shinyapps.datacurators.nl/tide/
Other		
CXCR6 sgRNA CRISPR sequence	5′- GCAGAGTACA GACAAACACC	N/A
GMDF Code and Instructions		https://github.com/livnatje/GMDF
Graphical Abstract web-design tool		Biorender.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ana C. Anderson (acanderson@bwh.harvard.edu).

Materials availability

Requests for plasmids generated in this study should be directed to the lead contact, Ana C. Anderson (acanderson@bwh.harvard.edu).

Data and code availability

- Single-cell RNA-seq and Bulk RNA-seq data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper analyzes existing, publicly available data. The accession numbers for these datasets are listed in the key resources table.
- All original code has been deposited on github. The link can be found in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice

6–8-week-old male or female C57BL/6 (Stock No. 000664), CXCR6 KO (Stock No. 005693), OT1 (Stock No. 003831), Cas9 (Stock No. 028555), E8i-Cre (Stock No. 008766), and CD45.1 (Stock No. 002014) transgenic mice were purchased from the Jackson Laboratory. *Tcf7* cKO (E8i-Cre^{+} *Tcf7*$^{FL/FL}$) mice were generated as previously described[22] and E8i-Cre^{-} x *Tcf7*$^{FL/FL}$ littermates were used as controls. CXCR6 KO, *Tcf7* cKO, and control mice were crossed to OT1. All mice were housed in a vivarium under SPF conditions, in cages of up to five mice, and fed a special rodent diet. All experiments involving laboratory animals were performed under protocols approved by the Harvard Medical Area Standing Committee on Animals (Boston, MA) and followed IACUC guidelines on the ethical care and use of animals.

Mouse tumor cell lines

B16F10 was purchased from ATCC. Mc38Ova^{hi} was generously provided by Dr. Nick Haining. B16Ova was kindly provided by Dr. Kai Wucherpfennig. All tumor cell lines were grown at 37°C with 10% CO_2. B16F10 and B16Ova were grown in RPMI supplemented with 10% FBS, glutamine (2mM), pen/strep, non-essential amino acids (1x), sodium pyruvate (1mM), and HEPES (5mM) and injected subcutaneously into the flanks of mice at 3-4x10^5 cells per mouse. Mc38Ova^{hi} cells were grown in DMEM supplemented with 10% FBS, glutamine (2mM), pen/strep, non-essential amino acids (1x), and sodium pyruvate (1mM) and injected subcutaneously into the flanks of mice at 1x10^6 cells per mouse. Post thaw, Mc38Ova^{hi} cells were cultured for one passage in medium containing puromycin (4 μg/mL) to select Ova-expressing cells followed by passage in normal medium.

METHOD DETAILS

Generalizable matrix decomposition framework (GMDF)

Given a set of d datasets collected under different conditions, GMDF decomposes each of the original datasets E_i (n_i x m) using latent shared (W) and context-specific (A_j) metagenes:

$$E_i \approx H_i W + \sum_j a_{ij} H_i^j A_j \quad \text{(Equation 1)}$$

where all the factor matrices are constrained to be non-negative. H_i, H_i^j are the usage matrices of the latent programs in dataset i. Note that W (k x m) is shared across all datasets (i = 1 … d), while A_j (k_j x m) is used only if the binary parameter a_{ij} is 1, denoting that condition j holds in sample i. A condition is defined by the user and depends on the specific structure and features of the data. It can denote the disease subtype, treatment status, data source, etc.

The optimization problem is then defined as

$$argmin_{H_i, H_i^j W, A_j \geq 0} \sum_{i=1}^{d} \left\| E_i - \left(H_i W + \sum_j a_{ij} H_i^j A_j \right) \right\|_F^2 + \sum_{j=1}^{p} \lambda_j \sum_i a_{ij} \left\| H_i^j A_j \right\|_F^2 \quad \text{(Equation 2)}$$

where λ_j is a tuning parameter which allows adjusting the size of context-specific effects.

To find a solution to the non-convex GMDF optimization problems, the variables are divided into blocks. Using block coordinate descent, the objective is iteratively minimized with respect to each block, while holding the others fixed.

$$(1)\ W = argmin_{W \geq 0} \left\| \begin{pmatrix} H_1 \\ \vdots \\ H_d \end{pmatrix} W - \begin{pmatrix} E_1 - \sum_j \sqrt{a_{1j}} H_1^j A_j \\ \vdots \\ E_d - \sum_j \sqrt{a_{dj}} H_d^j A_j \end{pmatrix} \right\|_2^F$$

$$(2)\ H_i = argmin_{H_i \geq 0} \left\| (W^T) H_i^T - \left(\left(E_i - \sum_j \sqrt{a_{ij}}\, H_i^j A_j \right)^T \right) \right\|_2^F$$

$$(3)\ A_g = argmin_{A_g \geq 0} \left\| \begin{pmatrix} \sqrt{a_{1g}} H_1^g \\ \vdots \\ \sqrt{a_{dg}} H_d^g \\ \sqrt{\lambda_g a_{1g}} H_1^g \\ \vdots \\ \sqrt{\lambda_g a_{dg}} H_d^g \end{pmatrix} A_g - \begin{pmatrix} E_1 - H_1 W - \sum_{j \neq g} \sqrt{a_{1j}} H_1^j A_j \\ \vdots \\ E_d - H_d W - \sum_{j \neq g} \sqrt{a_{dj}} H_d^j A_j \\ 0_{(\sum_i n_i) \times m} \end{pmatrix} \right\|_2^F$$, where A_g represents the programs specific to context g.

$$(4)\ H_i^g = argmin_{H_i \geq 0} \left\| \begin{pmatrix} \sqrt{a_{ig}} A_g^T \\ \sqrt{\lambda_g a_{ig}} A_g^T \end{pmatrix} (H_i^g)^T - \begin{pmatrix} \left(E_i - H_i W - \sum_{j \neq g} \sqrt{a_{ij}}\, H_i^j A_j \right)^T \\ 0_{m \times n_i} \end{pmatrix} \right\|_2^F$$, where H_i^g represents the usage of programs specific to context g in cohort i.

Each of the optimization subproblems above requires solving a nonnegative least-squares problem. Each of these subproblems is solved by using the fast block principal pivoting algorithm,[74] implemented using the Rcpp and LIGER packages in R. Convergence

criteria: After each iteration (1–4 described above), GMDF updates a convergence variable denoting the improvement in the score of the solution (computed as shown in Equation 2). The convergence variable drops semi-monotonically. Once it is below a pre-defined convergence threshold (with a default of $1*10^{-4}$), GMDF outputs the single run solution.

Obtaining a consensus GMDF solution

GMDF is non-convex, and therefore has multiple non-unique optimal solutions, in addition to sub-optimal solutions that can be obtained if reaching a local minimum. To alleviate these problems and generate a single robust solution, GMDF uses random subsampling of the data, and multiple initialization conditions to generate multiple solutions. The different solutions are then aggregated such that similar solutions that were repeatedly identified across various data sampling and initializations are combined. More specifically, all the shared programs (W matrixes), are clustered via hierarchical clustering, with Euclidean distance as the metric to evaluate similarity between programs. Clusters with less than 3 metagenes are removed and the others are converted to final metagenes by computing the mean loading of each gene in each cluster. The same procedure is performed for the A_j matrixes. This scheme mitigates outlier effects, reduces overfitting, and makes the final solution less dependent on the k and k_j parameters, identifying the number of the embeddings' dimensions directly from the data.

Batch correction

When applying GMDF to pan-cancer $CD8^+$ T cells mapping, E_i (Equation 1) includes only cells from the same cohort. Thus, GMDF does not make direct comparisons between cells of different datasets. The GMDF procedure further filters out batch-specific effects by factorizing the matrix to dataset-specific and -shared decompositions. The pan-cancer programs identified are unlikely to represent batch-effects across the studies, as they are repeatedly observed within each cohort. For future applications of GMDF, context specific programs will require several datasets from each context to distinguish between biologically meaningful context-specific program and second-order batch effects.

Comparing GMDF to LIGER and iNMF

To compare GMDF to LIGER[27] and iNMF,[26] data were simulated with shared and dataset-specific gene modules, where a shared gene module consists of genes that are correlated across the cells in all the datasets, and a dataset-specific module consists of genes that are correlated across the cells only in a specific dataset, while there is no correlation in the other datasets. Each dataset consisted of 200 cells, and module size was set to 100 genes. The results obtained with GMDF and LIGER/iNMF are shown in Figures S1A–S1D, with the ground truth denoted by the vertical color bar to the left. In contrast to LIGER/iNMF, GMDF identified all the programs and assigned them correctly as shared or specific to dataset 1 or 2 (Figures S1A–S1D). To further test LIGER, another analysis mode in the LIGER package was used, where, beyond decomposition, LIGER also identifies differentially expressed genes. Even when applied to identify differentially expressed genes across the different datasets, LIGER did not identify the context-specific modules.

Datasets and pre-processing

Processed data in the form of raw counts or transcripts per million (TPM) were obtained from the Gene Expression Omnibus (GEO), accession numbers: GSE163108 (glioblastoma[16]), GSE115978 (melanoma cohort 1[12]), GSE120575 (melanoma cohort 2[18]), GSE131309 (sarcoma[13]), GSE131907 (lung[14]), GSE114724 (breast cancer[11]), GSE98638 (hepatocellular carcinoma[20]), GSE108989 (colorectal cancer[19]), and GSE131928 (glioblastoma[47]). Pancreatic ductal adenocarcinoma (PDAC) TIL scRNA-seq data[17] was obtained from the Genome Sequence Archive (GSA), accession number CRA001160, project PRJCA001063. An additional lung cancer TIL scRNA-seq data[15] was downloaded from SCope (https://gbiomed.kuleuven.be/scRNAseq-NSCLC). Data from an Ovarian cancer cohort[34,35] was obtained from the website (http://blueprint.lambrechtslab.org). All data were converted to $\log_2(TPM/10 + 1)$, as previously described.[31] Throughout, Overall Expression was computed as previously described.[12] UMAPs (Figures 2 and S3) were obtained with the Seurat package, using the first 30 PCs obtained with top 2,000 most variable genes identified with the *FindVariableFeatures* function (default parameters).

GMDF application to the pan-cancer CD8 T cell cohort

GMDF was applied 100 times with different (random) initialization parameters and a random sub-sample of the data. To avoid genes with low detection rate, genes with detection rate below 0.1 in more than 4 datasets were omitted, resulting in 5,587 genes that were used as input for GMDF. For each run, 500 cells were subsampled from each cohort, and the topmost variable genes were selected, with variation threshold of 0.05 in the LIGER function selectGenes. GMDF parameters were 5 for k and k', convergence threshold $1*10^{-20}$, and $a_{ij} = 1$ if and only if $i = j$, to identify shared and cohort-specific programs. Following the 100 runs, a consensus solution was obtained as described above ("obtaining a consensus GMDF solution" section). Lastly, for each consensus shared program, the top 100 genes with the highest loading were used to define the pertaining program.

Decoupling convolved expression programs

T cell "activation" and "chronic activation" scores were computed separately for each dataset. Activation scores were defined as the Overall Expression[12] of the cytotoxic markers (*NKG7, CCL4, CST7, PRF1, GZMA, GZMB, IFNG, CCL3*) minus the Overall Expression

of memory/naive markers (*CCR7*, *TCF7*, *LEF1*, *SELL*). Chronic activation scores were defined as the Overall Expression of the pan-cancer chronic activation program identified by GMDF (Table S1). Chronic activation scores were then plotted as a function of the activation scores using a locally-weighted polynomial regression (LOWESS, black line in Figure 1C). The deviation from this regression line was used to classify CD8 T cells into four groups: Cells with a low activation score (below the 25th percentile) were classified as naive/memory-like cells, while the others were considered effector or dysfunctional, if their activation scores were (0.25 standard deviations) higher or lower than expected given their chronic activation scores, respectively. Cells with a high activation score (>25th percentile) that fit the regression line were annotated as "balanced" and were not used in the analysis below. In each dataset, genes that were significantly overexpressed in one subset of cells (e.g., dysfunctional) compared separately to each of the other two subsets (e.g., dysfunction vs. naive/memory and dysfunction vs. effector), were identified as markers of that cell subset in that dataset. These pairwise comparisons were performed with a multilevel regression model: $y_{ij} \sim N(\alpha_j + \beta x_{ij}, \sigma_1^2)$, $\alpha_j \sim N(\gamma_0, \sigma_2^2)$, where y_{ij} denotes the expression of the gene in cell i from sample j, x_{ij} denotes the total log-transformed number of reads detected in cell i in sample j, and α_j is the sample-specific intercept. The lme4[75] (https://CRAN.R-project.org/package=lme4) and lmerTest R packages[76] were used to fit the model, compute p-values, and identify the latent variables that maximize the posterior probability

$$argmax_{\alpha,\beta,\gamma,\sigma_1,\sigma_2} p(\alpha, \beta, \gamma, \sigma_1, \sigma_2 \mid y, x, u) \propto \prod_j^J \prod_i^{n_j} N(y_{ij} \mid \alpha_j + \beta x_{ij}, \sigma_1^2) \prod_j^J N(\alpha_j | \gamma_0, \sigma_2^2)$$

To aggregate the results across all datasets, pan-cancer p-values were obtained for each of the three pairwise comparisons using Fisher's combined probability test. Genes that were repeatedly identified as markers for one of the three states across all datasets and had a combined p-value $<10^{-10}$ were denoted as pan-cancer markers of the pertaining state.

To examine the model's ability to capture the continuous nature of T cell dysfunction, the HLM model was modified using a continuous (rather than binary) covariate to denote the dysfunction status based on the distance from the LOWESS regression line. When applied to identify dysfunction-specific genes in the melanoma cohort,[18] the procedure yielded very similar results to those of the original binary version (gene ranking similarity: $r_s = 0.906$, $p < 10^{-16}$, Spearman correlation).

CXCL16 pan-cancer macrophage program

scRNA-seq datasets of macrophages from melanoma,[12,18] sarcoma,[13] and glioblastoma,[47] breast cancer,[11] and lung cancer,[14,15] were analyzed to identify genes co-expressed with CXCL16. In each dataset genes co-expressed with CXCL16 were identified using partial Spearman correlation, accounting for the total number of genes detected in each cell. The CXCL16 program (Table S6) includes genes significantly co-expressed with CXCL16 (t-test, Benjamini-Hochberg FDR <0.05) in all seven datasets (Fisher's combined probability, Benjamini-Hochberg FDR $<1*10^{-3}$). In addition, in each dataset the gene ranks based on the partial Spearman correlation analyses were used as input for gene set enrichment analysis (GSEA) preformed using the fast GSEA package to obtain the normalized enrichment scores (NES) and plots shown in Figure 3B, using the KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION gene set.

Predictive performance of the pan-cancer dysfunction program

scRNA-seq data from melanoma samples[18] were used to examine the connection between the Overall Expression (OE) of a gene signature in CD8 T cells and ICB treatment status and response, using either (1) an HLM model, with patient-specific intercept, a covariate controlling for treatment status (i.e., pre/post ICB), and a binary covariate of response to treatment; or (2) an ROC curve using the average OE of the signature in the sample CD8 T cells to predict the response observed in a specific tumor, using only pre-treatment samples. The pan-cancer signatures identified here, and another set of T cells signatures previously identified in HCC[20] were used to examine the connection between T cell states and ICB response (Table S5).

Tumor experiments with immune checkpoint blockade

Tumor size was measured in two dimensions by caliper and is expressed as the product of two perpendicular diameters. In immune checkpoint blockade experiments, B16Ova tumors were treated with a combination of anti-PD-L1 (10F.9G2, 200μg per mouse) and anti-Tim-3 (RMT3-23, 200μg per mouse) antibodies or control immunoglobulin (Rat IgG2a and Rat IgG2b) i.p. on day 7 and day 10 post tumor implantation. Mc38Ova^{hi} tumors were treated with anti-PD1 (RMP1-14,100μg per mouse) antibody or control immunoglobulin (Rat IgG2a) i.p. on day 7 and day 10 post tumor implantation. Mice were then monitored every two days for tumor growth.

TIL isolation

TILs were isolated by dissecting the tumor mass and mincing the tumor tissue prior to digestion with Collagenase D (2.5 mg/mL) for 20 min at 37°C. Tumors were then dissociated to single cell suspensions by passing tissue through a 70μM filter prior to centrifugation and analysis by flow cytometry. For cytokine analysis, tumors were dissociated as above and cells underwent a density separation via percoll gradient before being plating for *ex vivo* stimulation.

Flow cytometry

Single cell suspensions were stained with fixable viability dye Zombie UV (Biolegend) and then incubated with anti-CD16/32 for 10 min prior to staining with antibodies against surface molecules. Antigen-specific $CD8^+$ T cells were identified by H-2Kb/SIINFEKL dextramer (Immudex) staining following the manufacturer's protocol. For intracellular staining, eBioscience Foxp3/transcription factor staining buffer set was used per manufacturer's protocol. For intra-cytoplasmic cytokine (ICC) staining of TILs, cells were stimulated *ex vivo* with 5 μg/mL OVA257-264 peptide (GenScript) for 4 h in the presence of Golgi stop (BD Biosciences) and Golgi Plug (BD Biosciences) prior to cell surface and ICC staining. Importantly, the antibody detecting CD107a was added to the cells during stimulation. Intracytoplasmic staining followed fixation and permeabilization. For intracellular staining of CXCL16, cells from tumors were cultured *ex vivo* for 4 h in the presence of Golgi Plug (BD Biosciences) prior to cell surface and intracellular staining. For quantitative flow cytometry we used absolute counting beads (Thermo Fisher Scientific, Cat. C36950) according to manufacturer's protocol. All data were collected on a BD Symphony (BD Biosciences) or Fortessa (BD Biosciences) and analyzed with FlowJo 10.7.1 software (TreeStar).

Adoptive cell transfers

Tumor cells were injected into the flanks of recipient mice as outlined above. On day 6 post tumor implantation, mice were randomized by tumor size and given T cells suspended in PBS via tail intravenous injections. For CRISPR KO experiments, $3x10^6$ (for B16Ova) or $2.25x10^6$ (for $Mc38Ova^{hi}$) CRISPR CXCR6 KO or CRISPR Control transduced cells were transferred into $Cas9^+CD45.1/.2$ recipients. For CXCR6 KO experiments, $7.5x10^5$ CXCR6 KO or Control cells were transferred. For adoptive co-transfer experiments, $2x10^5$ of both WT OTI CD45.1/.2 and CXCR6 KO OTI CD45.2 cells were transferred into the same CD45.1 tumor-bearing animal. Tumors were measured every two days until the day of tumor harvest when TILs were isolated and analyzed by flow cytometry.

Purification, lentiviral transduction, and *in vitro* culture of primary $CD8^+$ T cells for adoptive cell transfer experiments

For CRISPR KO transfers, $CD8^+$ T cells from spleens and lymph nodes of CD45.2 $Cas9^+$ OTI mice were isolated using CD8a microbeads (Miltenyi) and plated at a concentration of $5x10^5$ cells per well in the presence of IL-2 (6 ng/mL) in 24-well plates previously coated with 1 μg/mL anti-CD3/anti-CD28. The following day, cells were transduced at an MOI of 100 with lentiviral vectors. The vector contained an sgRNA targeting CXCR6 and a Thy1.1 gene to be used as a marker for transduction. A vector lacking the sgRNA but containing Thy1.1 was used as a control. Approximately 36 h after transduction, cells were moved to uncoated plates and cultured in the presence of IL-2 for an additional 3.5 days. On the day of transfer cells were stained with an antibody against Thy1.1 and FACs sorted to purify highly transduced cells. For CXCR6 KO adoptive transfer experiments, cells were isolated from wildtype OTI CD45.2 mice and CXCR6 KO OTI CD45.2 mice and activated as above. Cells were activated for 48 h before moving to uncoated plates and subsequently cultured in the presence of IL-2 for an additional 4 days prior to transfer. For adoptive co-transfer experiments, cells were isolated from wildtype OTI CD45.1/.2 and CXCR6 KO OTI CD45.2 mice and activated and cultured as above. T cells were cultured at 37°C with 10% CO_2 in DMEM containing 10% FBS, glutamine (2mM), pen/strep (100U/mL), non-essential amino acids (1x), sodium pyruvate (1.5mM), vitamins (1x), arginine, asparagine, folic acid (14μM), and 2-mercaptoethanol (57.2μM).

sgRNA design, testing, and plasmid construction

sgRNA guides were designed using the Broad Institute's GPP sgRNA designer portal (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design). Three candidate sgRNAs were chosen and each cloned into an sgRNA plasmid using BsmBI sites (Addgene #85453, modified with Thy1.1 in place of Tdtomato). Candidate sgRNA plasmids were transfected into doxycycline-inducible-Cas9 expressing murine 3T3 cells. After 48 h, transfection efficiency was determined and gDNA was collected. gDNA was used to perform a TIDE (Tracking of Indels by Decomposition) assay.[77] Briefly, PCR was used to amplify a roughly 700bp region around the expected sgRNA binding site. The PCR product was purified by agarose gel purification and Sanger sequenced. The resulting chromatogram was uploaded to the TIDE website for analysis and each sgRNA was evaluated compared to a control-transfected sample to determine the guide editing efficiency. The guide with the highest editing efficiency was chosen to move forward for lentivirus production (CXCR6 guide: 5′- GCAGAGTACAGACAAACACC).

Lentivirus production

VSV-pseudotyped third-generation lentiviruses were made as described previously.[78] Briefly, 293T cells were transiently co-transfected with five plasmids, then underwent a media change 12 h post-transfection. Cell supernatant was collected 26–30 h later and virus was concentrated by ultracentrifugation. Viral titers were determined on 293T cells by limiting dilution.

In vitro time-course of T cell activation and bulk RNA-seq

$CD8^+$ T cells from spleens and lymph nodes of wildtype OTI mice (E8i-Cre^- $Tcf7^{FL/FL}$ OTI^+) and *Tcf7* cKO OTI mice (E8i-Cre^+ $Tcf7^{FL/FL}$ OTI^+) ($n = 3$ per group) were isolated using CD8 microbeads (Miltenyi). After taking a timepoint 0 sample, cells were plated on plates previously coated with 1 μg/mL anti-CD3/anti-CD28 in the presence of murine IL2 (6 ng/mL). For each timepoint, 1,000 cells were sorted into 5μL of TCL buffer (Qiagen) supplemented with 1% 2-mercaptoethanol. Samples were processed using the SMART-Seq2 protocol,[79] and sequenced on an Illumina NextSeq.

RNA-seq reads were aligned using STAR to mouse genome version mm10, and expression levels were calculated using RSEM[80] using annotated transcripts (mm10), followed by further processing using the Bioconductor package DESeq2 in R.[81] Data was normalized using TPM normalization, and differentially expressed genes were defined using the differential expression pipeline on the raw counts with a single call to the function DESeq2 (reported *p*-values obtained via Wald test only for Benjamini-Hochberg FDR< 0.05).

ChIP-PCR

$CD8^+$ T cells from spleens and lymph nodes of wildtype (E8i-Cre^- $Tcf7^{FL/FL}$) TCR-transgenic and *Tcf7* cKO (E8i-Cre^+ $Tcf7^{FL/FL}$) TCR-transgenic mice were isolated using CD8 microbeads (Miltenyi). Cells were fixed for 10 min at room temperature with 1% formaldehyde and quenched with glycine. Cells were lysed with an SDS-lysis buffer prior to chromatin shearing via sonication. Sheared chromatin was "pre-cleared" using Protein A Agarose/Salmon Sperm DNA beads. Pre-cleared chromatin from the equivalent of $3.57x10^6$ cells was used in the chromatin IP reactions along with 3μg of either anti-TCF1 (C63D9) or Normal Rabbit IgG, both purchased from Cell Signaling. Protein-chromatin complexes were then bound with Protein A Agarose/Salmon Sperm DNA beads and underwent multiple rounds of washes. Finally, complexes were eluted off the beads, underwent reverse crosslinking with NaCl, and treated with proteinase K before isolating DNA using the QIAquick PCR Purification kit from Qiagen. DNA content from chromatin IPs was assessed using SYBR Green amplification with the following primers designed from the *Cxcr6* locus (mm10, chr9: 123806217–123806718): Fw-primer: 5′–GAGGCAGACCTTTAGTGAGCA–3'. Rev-primer: 5′–TAGCTCGCACCGTATACACA–3'. Results from Chromatin IP reactions were normalized to their "input" samples, or sample taken prior to Chromatin IP.

Mouse single cell RNA-seq

Wildtype (E8i-Cre^- $Tcf7^{FL/FL}$) and *Tcf7* cKO (E8i-Cre^+ $Tcf7^{FL/FL}$) mice were injected with $2.5x10^5$ B16Ova cells subcutaneously in their flanks. On day 13, tumors were harvested ($n = 3$ per group) and dissociated to single cell suspensions prior to FACS. $CD45^+$ $TCR\beta^+$ $CD8^+$ T cells (2,000 cells/mouse) were sorted along with additional immune cell subsets, and cells from mice from the same group were pooled and loaded for encapsulation for on the Chromium system (10x Genomics). Libraries were prepared using 5′ sequencing 10X Genomics kit v1 according to the Manufacturer's protocol. Libraries were sequenced on an Illumina HiSeq.

Mouse single cell RNA-seq analysis

Gene counts were obtained by aligning reads to the mm10 genome using CellRanger software (v1.3 10 3 Genomics). Cells were removed if they contained >5% mitochondrial transcripts or <200 genes, retaining 309 cells with 14,215 detected genes for further analysis. A log-transformed normalized count matrix was used as input for Principal Component Analysis (PCA) and the top 16 PCs were kept based on a drop in the proportion of variance explained by subsequent PCs. We confirmed that the resulting analyses were not particularly sensitive to this choice. Cells were clustered based on their top 16 PCs scores with the Louvain-Jaccard graph clustering algorithm,[82] as previously described[83,84] using the FindNeighbors ($k = 20$) followed by FindClusters (resolution = 0.5) functions with default parameters. scRNA-seq data was analyzed as describe above, with mixed-effect models used to account for confounders when examining differential expression of genes and gene programs. Analysis and plots were generated using Seurat package (version 4.0.1) for R with default parameters.

Luciferase assays

The murine *Cxcr6* genomic region (mm10, chr9:123806024-123807024) was chosen for luciferase assay testing, because it contained 500bp on either side of TCF1 predicted binding sites reported by Promo3.[85,86] To generate a list of potential regulators, this region was input into Promo3 with a maximum matrix dissimilarity rate of 10. The region was cloned into the luciferase reporter construct pGL4.10 (Promega E6651) using NEBuilder HiFi DNA Assembly (New England BioLabs). Using PolyJet (SignaGen Laboratories), HEK293T cells were transfected with the firefly luciferase reporter construct containing the CXCR6 region, the Renilla luciferase reporter as an internal control, and vectors expressing the indicated transcription factors. Transcription factor vectors were purchased from Genscript, all in the pcDNA3.1+/C-(K)-DYK vector backbone: Fos (Clone ID OMu22131D), JunD (Clone ID OMu14151D, subcloned to include the first 144 base pairs), and Jun (Clone ID OMu19159D). The Tcf1 vector was custom-made by Genscript based on the TCF1 protein coding sequence (ENSMUST00000109071.3) and put into the pcDNA3.1+/C-HA vector backbone. Cells were analyzed at 48 h post transfection with the dual luciferase assay kit (Promega E1910).

T cell-DC co-culture, bulk RNA-seq, and pathway analysis

$CD8^+$ T cells were isolated from the spleens and lymph nodes of wildtype CD45.2 OTI and CXCR6 KO CD45.2 OTI mice using CD8a microbeads (Miltenyi). $5x10^5$ cells per well were plated in the presence of murine IL-12 (10 ng/mL) in 24-well plates that were previously coated with 1 μg/mL anti-CD3/anti-CD28. After 48 h, media was changed to contain IL-2 (6 ng/mL) and cells were moved to uncoated plates. Cells were cultured in the presence of IL-2 for an additional 3 days before use in co-culture assay. Bone marrow-derived DCs from CD45.1 mice were generated with GM-CSF (5 ng/mL) and Flt-3L-Ig (human, 100 ng/mL) as previously described.[87] DCs were activated overnight with 2 μg/mL of the STING agonist DMXAA (Invitrogen, tlrl-dmx). The following morning, DCs were pulsed with 1nM of OVA257-264 peptide (GenScript) for 3 h at 37°C. $5x10^5$ Ova-pulsed DCs and 5×10^5 T cells were cultured together in 48 well plates. After 24 h, cells were stained with antibodies against CD45.2 and CD45.1 along with 7AAD for dead

cell removal. 5,000 CD45.2$^+$ cells were isolated by cell sorting into tubes containing 10μL TCL buffer (Qiagen) with 1% 2-mercaptoethanol. Samples were processed using the SMART-Seq2 protocol,[79] and sequenced on an Illumina NextSeq.

Bulk RNA-seq data were processed using nextflow-based RNA-seq pipeline with STAR and RSEM for alignment and transcript quantification ('–aligner star_rsem'), nf-core/tools v2.3.2 (https://github.com/nf-core/tools), bioconda, and nextflow v22.04.0 (https://github.com/nextflow-io/nextflow). Counts and transcripts per million (TPM) values were computed per sample. Estimated gene counts were provided as input into the Bioconductor R package DESeq2[81] for downstream differential gene expression analysis using default parameters. Differentially expressed genes (DEGs) were defined as those with an FDR <0.075 (Table S7). For pathway analysis, DEGs were input into Metascape[88] and run with default parameters. Significantly enriched pathways included: Leukocyte activation (GO:0045321), Regulation of regulatory T cell differentiation (GO:0045589), Costimulation by the CD28 family (R-MMU-388841), Metabolism of RNA (R-MMU-8953854), Ribonucleoprotein complex biogenesis (GO:0022613), and Ribonucleoprotein complex assembly (GO:0022618). Heatmaps were generated using Morpheus (https://software.broadinstitute.org/morpheus). Gene Set Enrichment Analysis was performed using the *fgsea* package in R with genes ranked using the Wald test statistic from the DESeq2 analysis. Gene sets from the mouse MSigDB category "REACTOME_COSTIMULATION_BY_THE_CD28_FAMILY" were interrogated to determine enrichment.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical tests are labeled for each figure in the figure legend. Significant differences between two groups were analyzed using GraphPad Prism using paired or unpaired two-tailed Student's t test. Tumor growth curves were analyzed using linear mixed effects models to test the trajectory of growth between various genotypes or treatments over time controlling for mouse. Values of $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ and $^{****}p < 0.0001$ were considered statistically significant.