The interaction of CD4 + helper T cells with dendritic cells shapes the tumor microenvironment and immune checkpoint blockade response

How does open access to this work benefit you?
Let us know @ library@weizmann.ac.il

Take down policy
The Weizmann Institute of Science has made every reasonable effort to ensure that Weizmann Institute of Science content complies with copyright restrictions. If you believe that the public display of this file breaches copyright please contact library@weizmann.ac.il providing details, and we will remove access to the work immediately and investigate your claim.

ARTICLES

Check for updates

The interaction of CD4+ helper T cells with dendritic cells shapes the tumor microenvironment and immune checkpoint blockade response

Merav Cohen [1,2,3,4,13] ✉, Amir Giladi [1,3,5,13], Oren Barboy [1,13], Pauline Hamon [3,4], Baoguo Li [1], Mor Zada [1,6], Anna Gurevich-Shapiro [1,6,7], Cristian Gabriel Beccaria [8], Eyal David [1], Barbara B. Maier [3,4], Mark Buckup [3,4], Iris Kamer [9], Aleksandra Deczkowska [1,10], Jessica Le Berichel [3,4], Jair Bar [6,9], Matteo Iannacone [8], Amos Tanay [11,13] ✉, Miriam Merad [3,4,12,13] ✉ and Ido Amit [1,13] ✉

Despite their key regulatory role and therapeutic potency, the molecular signatures of interactions between T cells and antigen-presenting myeloid cells within the tumor microenvironment remain poorly characterized. Here, we systematically characterize these interactions using RNA sequencing of physically interacting cells (PIC-seq) and find that CD4+PD-1+CXCL13+ T cells are a major interacting hub with antigen-presenting cells in the tumor microenvironment of human non-small cell lung carcinoma. We define this clonally expanded, tumor-specific and conserved T-cell subset as helper T tumor (Tht) cells. Reconstitution of Tht cells in vitro and in an ovalbumin-specific αβ TCR CD4+ T-cell mouse model, shows that the Tht program is primed in tumor-draining lymph nodes by dendritic cells presenting tumor antigens, and that their function is important for harnessing the antitumor response of anti-PD-1 treatment. Our molecular and functional findings support the modulation of Tht–dendritic cell interaction checkpoints as a major interventional strategy in immunotherapy.

The tumor microenvironment (TME) is a complex ecosystem, where cancer, immune and stromal cellular interactions influence tumor immunity and tumor growth[1]. Understanding cellular communications within the TME can be harnessed to enhance antitumor immune responses, as shown by the clinical success of immune checkpoint blockade (ICB) therapies that inhibit negative T-cell signals induced by cancer and antigen-presenting cells (APCs)[2–5]. However, the mechanisms of action of ICB therapies, and the factors contributing to their success in generating significant effector T-cell immunity are only partly understood. Better understanding of cellular communications within the TME is therefore essential for enhancing existing therapies and developing more-potent targeting schemes for immunotherapy.

T-cell–myeloid cell interactions determine the balance between inflammatory and tolerogenic immune function by providing instructive signals, including antigen presentation, cytokine production and co-stimulation, which shape T-cell response to threats[2,6,7]. Accordingly, physical interactions between different subsets of T cells and myeloid cells have been shown to control antitumor immunity[2,8,9]. For instance, while PD-1-expressing T cells can be

engaged by multiple cell types expressing PD-L1, such as dendritic cells (DCs), macrophages and cancer cells, it is the PD-1–PD-L1 axis provided in the context of T-cell–DC interactions that most significantly defines T-cell function in the TME[6]. These processes underline the need to unravel the cellular identities and molecular programs involved in physical interactions that control antitumor T-cell effector function at the tumor site[6,10].

Single-cell RNA-sequencing (scRNA-seq) technologies have led to better understanding of the heterogeneity of immune cell states within the TME[11–15]. Recent scRNA-seq studies across many cancer types, gave rise to a detailed atlas of recurrent CD8+ and CD4+ T-cell states in tumors, as well as to maps of tumor-associated macrophages (TAMs) and DCs[8,9,16–18]. Integration of T-cell receptor sequencing (TCR-seq) within such atlases suggested that some of the T-cell states, most notably the CD8+ dysfunctional/exhausted T cells, are clonally expanded in response to tumor neoantigens[11,19–21], but others may be bystander T cells. To further put these atlases in context, direct and unbiased quantification of the interaction between APCs and T cells within the TME is urgently needed. Currently, single-cell maps, even when combined with TCR-seq, can only

[1]Department of Immunology, Weizmann Institute of Science, Rehovot, Israel. [2]Department of Clinical Microbiology and Immunology, Sackler School of Medicine, Tel Aviv University, Tel-Aviv, Israel. [3]Department of Oncological Sciences, The Precision Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [4]The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [5]Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences and University Medical Center Utrecht, Utrecht, the Netherlands. [6]Sackler School of Medicine, Tel Aviv University, Tel-Aviv, Israel. [7]Division of Hematology, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel. [8]Division of Immunology, Transplantation and Infectious Diseases and Experimental Imaging Center, IRCCS San Raffaele Scientific Institute, and Vita-Salute San Raffaele University, Milan, Italy. [9]Institute of Oncology, Sheba Medical Center, Tel Hashomer, Israel. [10]Brain-Immune Communication Laboratory, Department of Immunology and Neuroscience, Institute Pasteur, Paris, France. [11]Department of Computer Science and Applied Mathematics, Department of Biological Regulation, Weizmann Institute of Science, Rehovot, Israel. [12]Human Immune Monitoring Center, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [13]These authors contributed equally: Merav Cohen, Amir Giladi, Oren Barboy, Amos Tanay, Miriam Merad, Ido Amit ✉e-mail: meravcohen@tauex.tau.ac.il; amos. tanay@weizmann.ac.il; miriam.merad@mssm.edu; ido.amit@weizmann.ac.il

provide hypotheses as to which potential interactions in the TME are driving the dynamics of effector or exhausted T cells, which are the prominent targets of recent successful immunotherapies. Here, we define a detailed cell–cell interaction network between T-cell and myeloid cell populations in the TME and healthy tissues. These data may serve as an important anchor to generate the next generation of T-cell/DC-based therapies.

Results

Characterization of myeloid and T-cell compartments in NSCLC.
To characterize and molecularly dissect physical interactions between myeloid cells and T cells in the TME, we applied PIC-seq technology[22] on clinical samples of early human non-small cell lung carcinoma (NSCLC) lesions. We profiled treatment-naive early NSCLC lesions that were surgically resected from ten patients and compared tumor-involved to adjacent tumor-free tissues (Supplementary Table 1). We optimized lung tissue digestion to preserve physiological cell conjugates and sorted single CD64+CD11c+ myeloid cells and TCRβ+ T cells, as well as CD64+CD11c+TCRβ+ conjugates of physically interacting cells (PICs; Fig. 1a,b and Methods). Overall, we sequenced 5,136 quality control (QC)-positive single TCRβ+ T cells and 5,626 QC-positive single CD11c+CD64+ myeloid cells collected from tumor and adjacent healthy tissues (Supplementary Table 2 and Extended Data Fig. 1a–d) and used the MetaCell package[23] to create a background model of single-cell states in the TME and healthy tissues (Fig. 1c–e). We controlled for cross-patient batch effects by confirming that each metacell includes cells from multiple patients (Extended Data Fig. 1e). We grouped T metacells, according to hallmark gene expression (Fig. 1c,d), into migratory T (*TCF7*, *CCR7* and *SELL*), naive T (*TCF7* and *IL7R*), proliferating T (*TOP2A*, *MKI67* and *STMN1*), CD4+ activated T (*CXCR4* and *CD69*), CD4+ memory T (*GZMB* and *CXCR4*), CD4+ regulatory T (T_reg; *FOXP3* and *IL2RA*), CD4+PD-1+CXCL13+ T (*CD4*, *CXCL13*, *PDCD1* and *IL21*), CD8+ T (*CD8A* and *CD8B*), CD8+ cytotoxic T (CTLs; *GNLY*, *GZMB* and *PRF1*) and CD8+ dysfunctional (exhausted) T cells (DysCD8; *GZMK*, *LAG3* and *HAVCR2*) (Fig. 1c,d and Extended Data Fig. 1f).

We similarly grouped myeloid cells into two subsets of monocytes, based on the high expression levels of the *VCAN* and *CD31* (*PECAM1*) genes (Fig. 1e), two 'MonMac' subsets (based on their high expression of CCR2 or CXCL10 and shared expression of monocyte and macrophage genes) and into two macrophage subsets: TAM subsets Mreg-Mac[16] (TREM2, GPNMB and APOE) and MMP9+ TAMs. Additional myeloid subsets included four groups of DCs: DC-expressing monocyte genes (MoDCs; expressing both monocytes (*VCAN*) and DCs (*CD1C*) genes), classical DC type I (cDC1; XCR1 and CLEC9A), classical DC type II (cDC2; CLEC10A, CD1C and BHLHE40) and mature DCs enriched in immunoregulatory molecules (mregDCs; FSCN1, CCL22 and CCL19)[8] (Fig. 1c,e). We found that the CD11c+CD64+ populations in some of the patients were enriched in natural killer (NK) cells (low in TRBC2 and CD8A, high in TRDC and KLRC1), which we grouped into two subsets of NCAM1+ NK and CX3CR1+ NK cells (Fig. 1e). For both the T-cell and myeloid cell compartments, we identified consistent differences in cell composition between tumor and adjacent tumor-free tissues (Fig. 1f). To support further such differences and to consolidate a common signature of the TME cellular composition, we computed TME versus healthy cell state ratios per-specimen and quantified correlations between subset abundances (Fig. 1g,h). Specifically, Mreg-Mac, TAMs, CXCL10+ MonMac, MoDCs, mreg-DCs and cDC2, together with T_reg, CD4+PD-1+CXCL13+ T and DysCD8 T cells were enriched in the TME, whereas the monocyte subsets and naive, proliferating and migratory T cells were under-represented (Fig. 1h). CD4+ memory and activated T-cell states, as well as CTLs were equally distributed across tumor and adjacent tissues. Notably, these results are in line with previous single-cell

analysis of NSCLC lesions[24–26] and therefore can serve as a baseline to investigate the molecular programs defining cellular interactions in the TME.

CD4+PD-1+CXCL13+ T cells are enriched in TME interactions.
We analyzed conjugates of CD64+CD11c+TCRβ+ PICs with heterotypic gene expression and removed putative T/NK cell PICs due to low specificity of their PIC-seq deconvolution (Extended Data Fig. 2a–d and Methods). Using an antibody-switching experiment we have previously applied[22], we quantified the frequency of real versus spurious doublets formed after human tumor and healthy lung tissue dissociation. For each tumor and healthy tissue, we digested two parallel tissue sections, staining the split samples for TCRβ and CD11c, each with a different combination of fluorophores (TCRβ–FITC/CD11C–PE and TCRβ–PE-Cy7/CD11c–APC-Cy7). We then pooled the two tumor and healthy samples together and analyzed them by FACS. We note that spurious doublets formed after dissociation, which combine fluorophores from parallel mixes, are rare, estimated at 2.5% and 4.4% of the CD64+CD11c+TCRβ+ PIC population in the TME and adjacent noninvolved tissue, respectively (Extended Data Fig. 2e). We overall modeled 839 QC-positive PICs by inference of their T-cell and myeloid cell identities as described (Extended Data Fig. 2f–j and Methods). The PIC-seq pipeline utilizes a detailed background model of the singlet populations contributing to the PIC conjugates to facilitate the estimation of interaction preferences and assign for each PIC the most probable pair of contributing singlet identities[22] (Fig. 2 and Methods). PICs isolated from tumor, but not healthy, lung tissue showed pronounced over-representation of CD4+PD-1+CXCL13+ T cells ($P = 1.03 \times 10^{-7}$) and depletion of CD8+ dysfunctional T cells ($P = 0.0067$) (Fig. 2a). The identification of interactive CD4+PD-1+CXCL13+ T cells within PICs was based on subset-specific genes, including *CXCL13*, *MAF*, *ZBED2*, *PRDM1*, *SNX9*, *SIRPG* and *IL21* (Fig. 2b). Of note, CD4+PD-1+CXCL13+ T cells were significantly over-represented in PICs when stratified across profiled patients (paired Mann–Whitney *U*-test; $P = 0.037$; Fig. 2c).

Myeloid cell contribution to T-cell/myeloid PICs was characterized by a reduction in the CD31+ monocyte subset and enrichment of the Mreg-Mac[16] subset in both healthy ($P = 0.0047$ and 0.002, respectively) and tumor tissues ($P = 0.0087$, 1.9×10^{-5}, respectively; Fig. 2d). Of note, we found significant enrichment in mregDC frequency ($P = 0.043$), alongside a depletion in cDC1 ($P = 0.0079$) in T-cell/myeloid PICs in tumor tissues compared to adjacent tissues (Fig. 2d). The mregDC annotation in PICs was based on expression of unique gene transcripts identified in singlets (such as *LAMP3*, *CCL22*, *CCL19*, *BIRC3*, *FSCN1* and *CCR7*; Fig. 2e), as we recently described[8]. The mregDC over-representation in PICs was observed also when stratified across patients (paired Mann–Whitney *U*-test; $P = 0.039$; Fig. 2f). Taken together, by investigating the cellular interaction repertoire within PICs, we identified CD4+PD-1+CXCL13+ T cells and mregDCs, two states that are enriched in the TME compared to healthy controls and engage specifically in T-cell–myeloid interactions in the TME, suggesting their potential TME interactive capacity and function.

Defining the Tht niche across tumor types. Given their enrichment in tumor PICs in NSCLC, we next asked whether the CXCL13+PD-1+ CD4+ T-cell program was also present in other tumor types, using scRNA-seq datasets obtained from 25 patients with melanoma[27] and 42 patients with breast cancer[28]. We annotated the T-cell states across tumor types and identified similar subsets to those we identified in NSCLC (Extended Data Fig. 3a and Methods). We subsequently derived a pan-tumor signature consisting of gene transcripts differentially expressed by CD4+PD-1+CXCL13+ cells compared to all other CD3+ T-cell populations across the different tumor types. We found that CD4+PD-1+CXCL13+ cells expressed
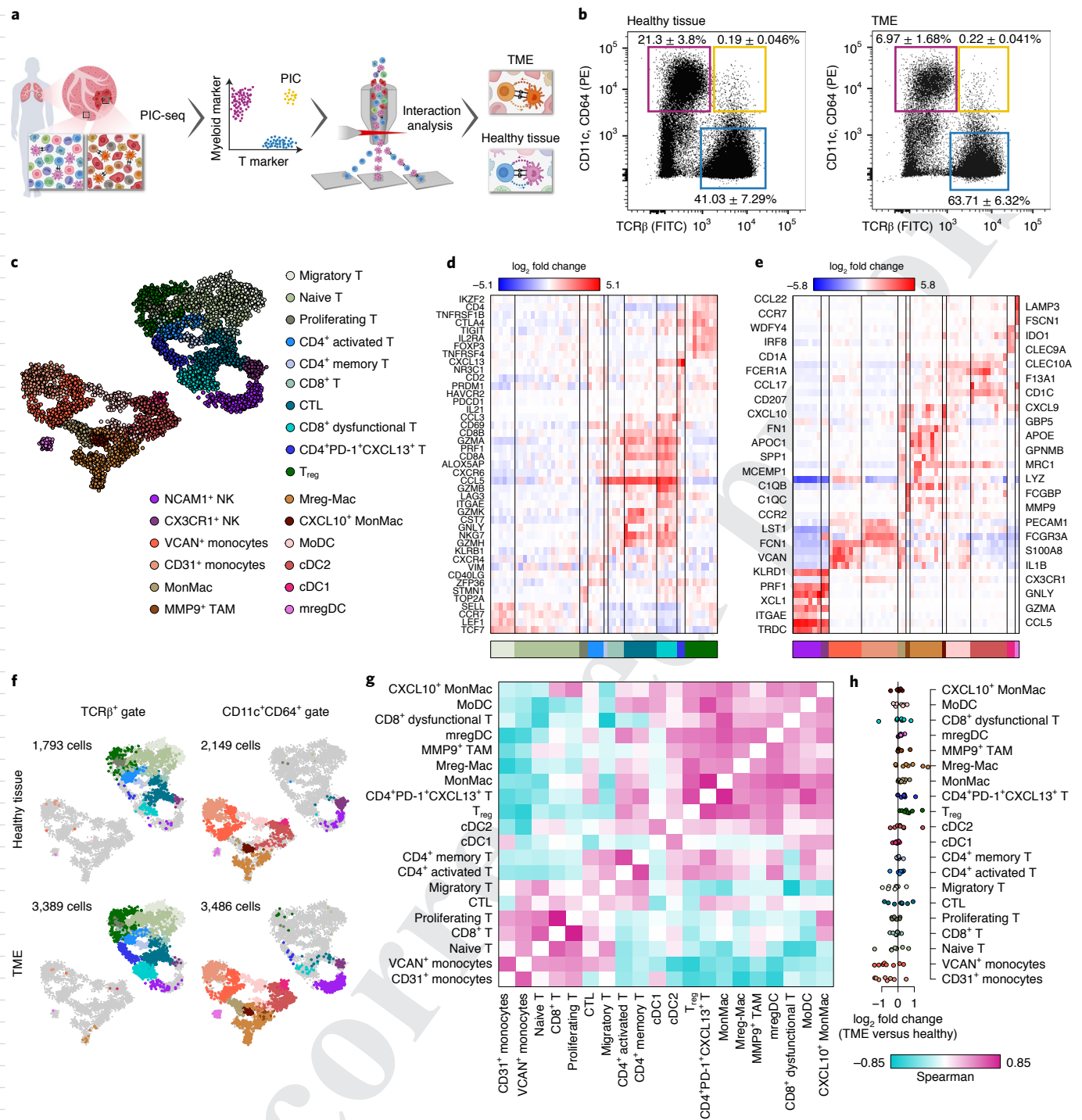
Fig. 1 | PIC-seq application on tumor and adjacent healthy tissues derived from stage-I biopsies of patients with NSCLC. a, Schematics of the experimental scheme using PIC-seq to characterize interactions between myeloid and T cells in human clinical specimens. **b,** Representative FACS plots of CD64⁺CD11c⁺ myeloid (purple) and TCRβ⁺ T cell (blue) singlets and CD64⁺CD11c⁺TCRβ⁺ PICs (orange) purified from healthy lung and TME of patients with NSCLC (*n* = 10 patients); population frequencies represent mean ± s.e.m. **c,** A two-dimensional (2D) representation of a MetaCell model of 5,136 TCRβ⁺ T cells and 5,626 CD11c⁺CD64⁺ myeloid cells from *n* = 10 patients with NSCLC, grouped into 22 T-cell and myeloid cell subsets. Dots represent single cells and dot colors are related to annotation of T-cell and myeloid cell subsets. **d,e,** Gene expression profiles of T cell (**d**) and myeloid (**e**) metacells. Values indicate enrichment (log₂ fold change) of a gene in a metacell over its median value across metacells. MetaCell annotation to 10 T-cell and 12 myeloid cell subtypes (bottom). **f,** Projection of TCRβ⁺ T or CD64⁺CD11c⁺ myeloid-sorted cells derived from healthy lung tissue or TME onto the 2D map shown in **c**. **g,** A pairwise gene Spearman correlation analysis of the representation of T-cell and myeloid cell subsets across samples derived from TME and adjacent healthy tissues of patients with NSCLC. **h,** Log₂ fold change between TME and healthy lung tissue percentage of each T-cell and myeloid cell subset across the profiled patients. Each circle represents one patient; *n* = 10 patients with NSCLC (Supplementary Table 1).

chemokines and cytokines (CXCL13, IL21, IFNG and CCL3), hallmark immune checkpoints (BTLA, PDCD1, TIGIT and CTLA4) and transcription factor genes (*MAF*, *HIF1A* and *ZBED2*) (Fig. 3a, Extended Data Fig. 3b,c, Supplementary Table 3 and Methods). This transcriptional program was specific to this subset and to PICs assigned to CD4+PD-1+CXCL13+ T cells in the NSCLC TME (Fig. 3b). We subsequently annotated this program as helper Tht cells.

To gain deeper insights into the molecular signature of Tht cells within PICs, we performed differential gene expression analysis between their observed and expected expression derived from the PIC-seq null model (Fig. 3c and Methods). We found major cytokines and chemokines (CXCL13, CCL4 and CCL5), immune checkpoint genes (*CTLA4* and *PDCD1*) and transcription factors (MAF) to be specifically upregulated in interacting Tht cells compared to singlet Tht cells (Fig. 3c,d), implicating that physically interacting myeloid cells are imprinting the unique Tht-cell signature.

In addition to our pan-cancer analysis revealing a conserved Tht-cell program across several human tumors, we also identified two Tht-cell states (Tht-I and Tht-II), evident in all analyzed cancer types (NSCLC, melanoma and breast cancers). Both groups have high expression of CD4 and CXCL13, whereas Tht-I cells exhibited higher expression of genes associated with cytotoxic activity and exhaustion (such as *GZMB*, *IFNG*, *LAG3* and *CXCR6*) and Tht-II cells expressed high levels of IGFL2, CPM and TCF7, suggesting a more naive state of Tht cell development (Extended Data Fig. 3d). Notably, in our NSCLC TME PIC data, Tht-I cells are significantly more interactive with myeloid cells compared to Tht-II cells ($P = 0.014$ and 0.55, respectively; two-tailed Mann–Whitney U-test; Extended Data Fig. 3e).

Tht cells define spatial niches in NSCLC tumors. We next explored the spatial distribution of Tht cells in lung sections of patients with NSCLC, including their relative proximity to myeloid cells. Chronic exposure to tumor antigens has been shown to lead to formation of tertiary lymphoid structures (TLSs) in tumor lesions[29,30]. As mregDCs are enriched in tumor-derived PICs, we investigated whether they are in physical proximity to Tht cells in tissue sections, by imaging the mregDC protein marker, DC-LAMP (Fig. 3e and LAMP3 in Fig. 2e). Notably, TLSs in early-stage NSCLC are composed of clusters of DC-LAMP+ mature DC and T cells, within T-cell areas adjacent to B-cell follicles[30]. Imaging of NSCLC tumor tissues identified CD4+ T cells expressing the Tht cell hallmark proteins CXCL13, PD-1, PRDM1 and BTLA (CD272; Extended Data Fig. 4a). These Tht cells were found to be located in close proximity to DC-LAMP+ DCs in NSCLC tumor tissues, suggesting that they form direct physical interactions. Specifically, in four of our profiled patients, we observed CD4+PD-1+ T cells and DC-LAMP+ DC interactions within CD4+ T-cell aggregates, raising the possibility that Tht cells might be involved in TLS response in the TME (Extended Data Fig. 4b). To study Tht cell spatial associa-

tion to other myeloid and lymphoid cell subsets in the TLS niche of the TME, we performed multiplexed immunohistochemical consecutive staining on a single slide (MICSSS) of formalin-fixed paraffin-embedded (FFPE) sections from patients with NSCLC (Methods). We first defined TLSs histologically, based on their high nuclear (4,6-diamidino-2-phenylindole (DAPI+)) density characteristic (Fig. 3e). While CD68+ macrophages and CD8+ T cells were enriched in non-TLS regions of the TME, CD3+CD8−PD-1+ (Tht) and CD3+CD8+PD-1+ (DysCD8) T cells, as well as CD20+ B cells and DC-LAMP+ DCs were spatially restricted to the TLS niche (Fig. 3e,f, Extended Data Fig. 4c and Methods). Of note, short-range community analysis confirmed compartmentalization within the TLS, with DC-LAMP+ DCs frequently observed in close proximity to Tht cells and DysCD8+ T cells, whereas CD20+ B cells formed segregated communities within the TLS (Extended Data Fig. 4d), suggesting that Tht cells and DCs interact in the TLS niche.

In agreement with our hypothesis that Tht cells are reactive to tumor-specific antigens, we found that these cells exhibited a high degree of clonality in breast and melanoma data, comparable to the tumor antigen-specific CD8 subset, dysfunctional (exhausted) CD8+ T cells (Fig. 3g and Extended Data Fig. 4e). Notably, Tht cell clones tended to be exclusive, exhibiting shared clonality between Tht-I and Tht-II, but with little contribution from other T-cell subsets, including CD8+ T cells or CD4+ T$_{reg}$ cells (Fig. 3h and Extended Data Fig. 4f). Together, our results, demonstrating the enrichment of clonally expanded CD4+PD-1+CXCL13+ Tht cells in physical interactions with mregDCs within the TLSs of tumor lesions, suggest that Tht cells are educated against tumor antigens.

Antigen specificity promotes a unique CD4+ T-cell state. To explore the mechanisms that induce tumor-specific CD4+ T-cell response to tumor antigen presentation, we made use of a model consisting of murine ovalbumin (OVA)-specific αβ TCR CD4+ T cells (OT-II) that recognize the OVA peptide. We isolated splenic CD11c+ DCs from C57BL/6 wild-type (WT) mice and cultured them for 2 h with mCherry-B16 melanoma cells expressing the OVA peptide (B16-OVA), before co-culturing them with splenic OT-II CD4+ T cells for 20 h. Partial co-cultures containing B16-OVA with splenic OT-II CD4+ T cells and monocultures of splenic OT-II CD4+ T cells served as controls. We performed MARS-seq on 1,860 QC-positive, single TCRβ+ T cells sorted from the cultures (Fig. 4a). Pairwise gene correlation analysis revealed two highly correlated gene modules enriched for genes related to immune activation and co-stimulation, reminiscent of the human Tht cell program (such as *Pdcd1*, *Tigit*, *Btla*, *Bhlhe40*, *Tnfrsf4*, *Cd82*, *Gng4*, *Tpi1* and *Pkm*) or for genes associated with a T-cell-naive state (such as *Sell*, *Ccr7* and *Tcf7*) (Fig. 4b). Expression of the activation signature was restricted to a small set of tumor-activated T cells, whereas the majority of the T cells were in the naive state (Fig. 4c). The frequency of tumor-activated T cells was highest in TCRβ+ T cells isolated from the cultures containing B16-OVA cancer cells, DCs and OT-II CD4+

**Fig. 2 | Interaction preferences of T-cell and myeloid cell subsets revealed by PIC-seq. a**, Distribution of T-cell subsets in TCRβ+ singlet T cells and CD64+CD11c+TCRβ+ PICs in healthy lung tissue and the TME of patients with NSCLC. Colors are related to annotation of T-cell subsets. Cells are downsampled so that T-cell and PIC numbers are equal per patient and then pooled from all profiled patients. False discovery rate (FDR)-adjusted two-tailed Fisher's exact test. **b**, Gene expression profiles of singlet T cells (left) and PICs (right) grouped by their MetaCell and PIC-seq assignment to T-cell subsets. Shown are genes supporting assignment of singlets and PICs to the CD4+PD-1+CXCL13+ T-cell subset. MetaCell and PIC-seq assignment of T-cell identities (bottom). **c**, Comparison of DysCD8+ and CD4+PD-1+CXCL13+ T-cell subset frequencies in singlets and PICs derived from healthy tissue and TME across all profiled patients. Two-tailed paired Mann–Whitney U-test; $n = 10$ patients with NSCLC. **d**, Distribution of myeloid subsets in CD64+CD11c+ singlet myeloid and CD64+CD11c+TCRβ+ PICs in healthy tissue and TME. Colors are related to annotation of myeloid cell subsets. Cells are downsampled so that T-cell and PIC numbers are equal per patient and then pooled from all profiled patients. FDR-adjusted two-tailed Fisher's exact test. **e**, Gene expression profiles of singlet myeloid cells (left) and PICs (right) grouped into their MetaCell and PIC-seq assignment to myeloid subsets. Shown are genes supporting assignment of singlets and PICs to the mregDC subset. MetaCell and PIC-seq assignment of myeloid identities (bottom). **f**, Comparison of different myeloid subset frequencies in singlets and PICs derived from healthy and tumor tissues across all profiled patients. Two-tailed paired Mann–Whitney U-test; $n = 10$ patients with NSCLC; *$P < 0.05$, **$P < 0.001$, ***$P < 10^{-5}$.

262   T cells and enriched in CD4+ T cells expressing high levels of PD-1
263   and CXCR5 cell-surface proteins, specifically derived from these
264   triple cultures (Fig. 4d,e).
265   To examine whether this activated T-cell gene signature is
266   restricted to the microenvironmental settings of tumor antigen
267   specificity or is general to immune conditions ensuing antigen pre-

sentation, we analyzed transcriptional states of OT-II T cells iso-
lated from the B16-OVA + DC cultures compared to co-cultures
with OVA-presenting DCs that were pre-exposed to lipopolysac-
charide (LPS) (Fig. 4f)[22]. We found that antigen presentation by the
LPS-activated OT-II T cells induced general immune activation genes
(such as *Mif*, *Il2ra*, *Srm*, *Irf4* and *Il2*). However, only tumor-specific
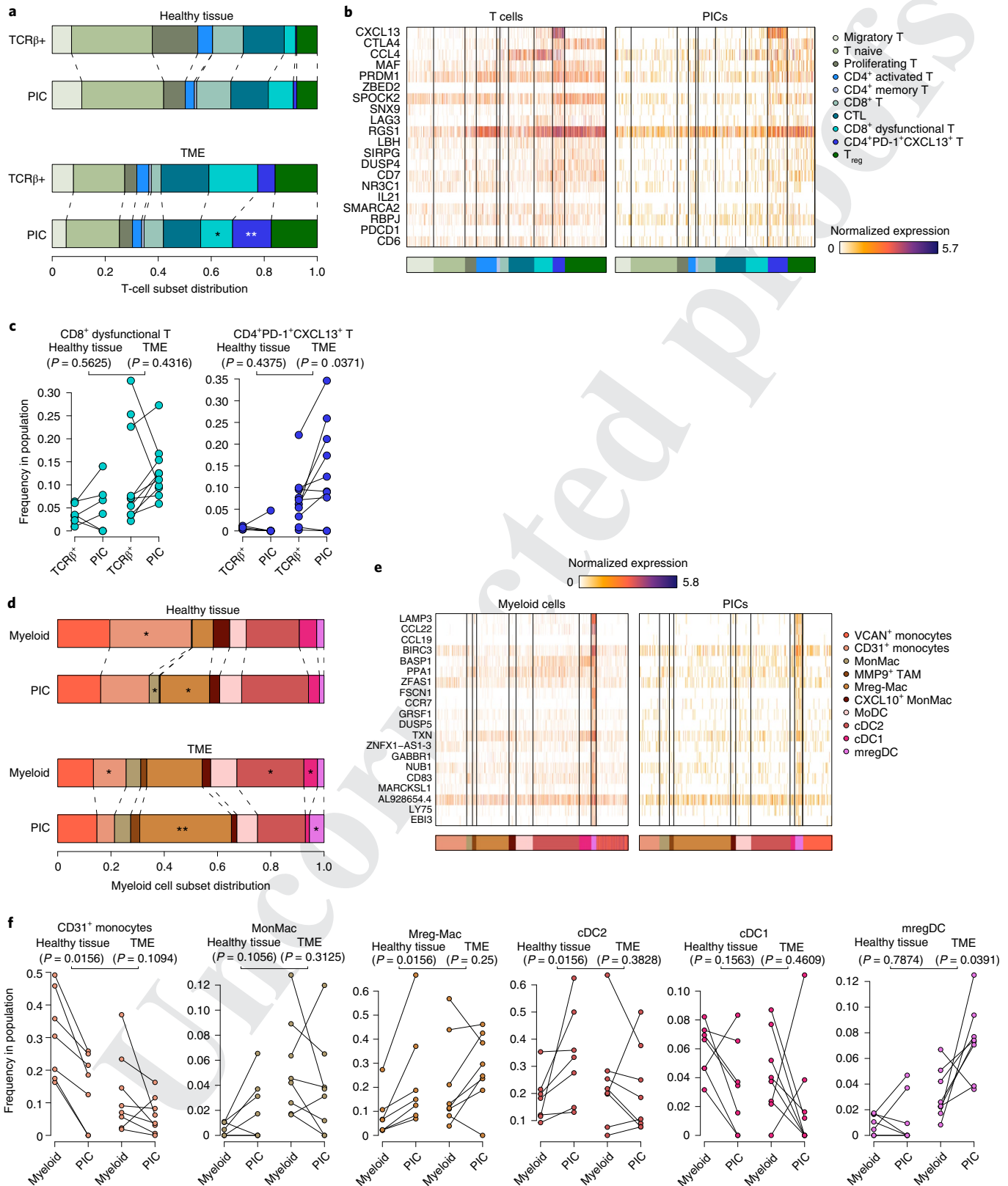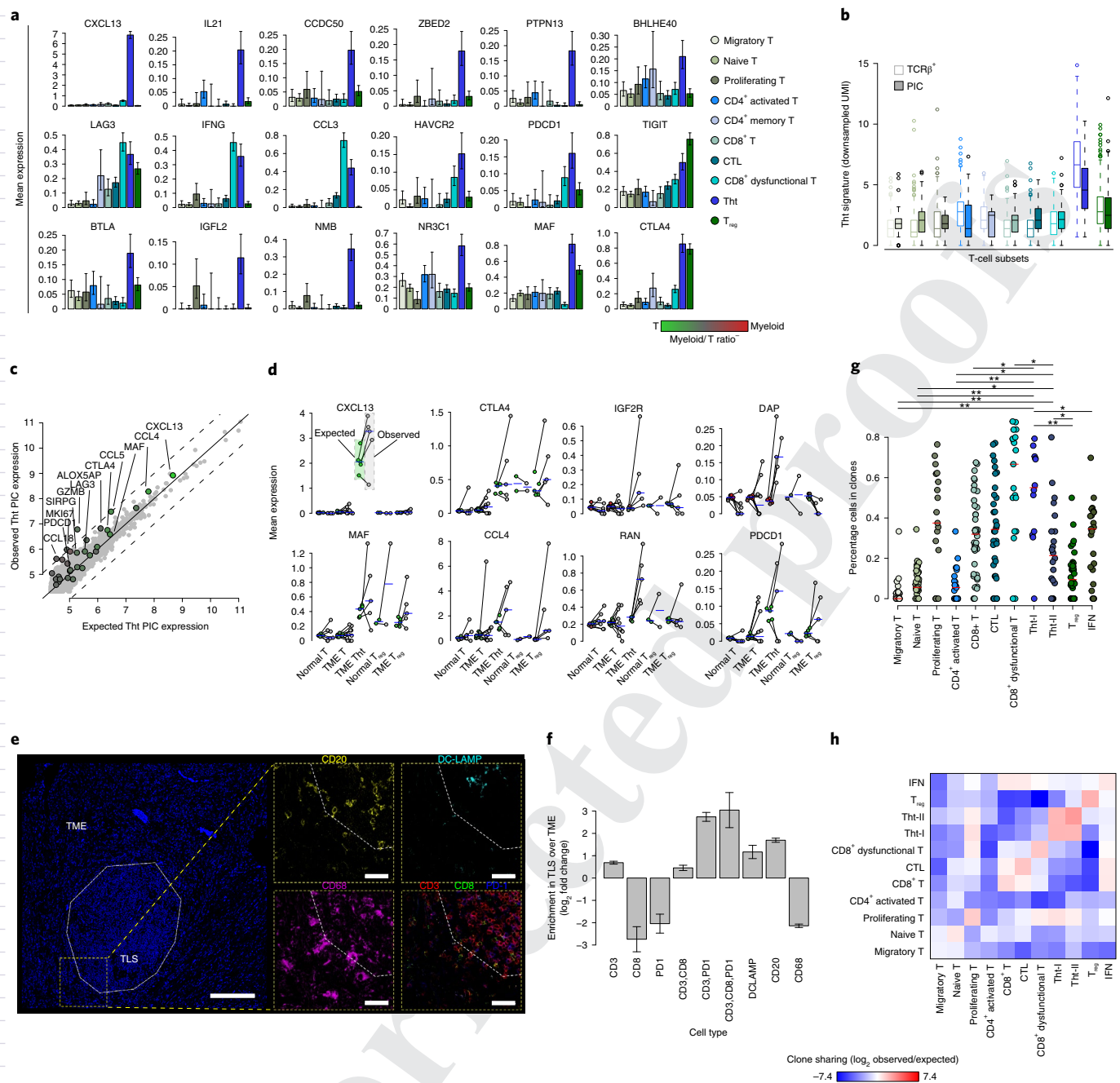
Fig. 3 | CD4+PD-1+CXCL13+ T cells present unique gene expression and interactive profile in TME. a, Mean normalized expression of CD4+PD-1+CXCL13+ T-cell genes across ten NSCLC T-cell subsets. Error bars indicate binomial 95% confidence intervals. $n = 3{,}371$ TME T cells over ten patients with NSCLC. **b**, Distribution of the pooled expression of the Tht (CD4+PD-1+CXCL13+) signature genes across all ten NSCLC T-cell populations in singlet (empty boxes) and PIC (full) states. The central mark in box plot is median, with 5th and 95th percentiles at the whiskers, 25th and 75th percentiles at the box and minima and maxima marked by dots; $n = 2{,}247$ T cells and 839 PICs over ten patients with NSCLC. **c**, Observed gene expression levels in PICs assigned to the CD4+PD-1+CXCL13+ identity, plotted against their expected levels. Highlighted genes are colored by their expected specificity to the T-cell (green) or myeloid (red) compartments ($\log_2$ fold change). **d**, Mean observed (gray, left) and expected (colored, right) gene expression levels in PICs grouped according to their T-cell identities in healthy and tumor tissues. Each connected pair of dots signifies a patient. Dot colors relate to their expected specificity in the T-cell (green) or myeloid cell (red) compartments as in **c**; cells from non-Tht or T_{reg} cell subsets are pooled together. Median value is marked with a blue line. **e**, MICSSS staining depicting a TLS inside a tumor section. White dashed lines indicate TLS boundaries. Scale bars, 100 μm (left), 20 μm (right). Image is representative of five TLSs from the same patient. **f**, Quantification of cell-type enrichment ($\log_2$ fold change) within the TLS compared to the TME. Cell types are determined by staining colocalization. $n = 5$ TLSs from the same patient; error bars represent 95% confidence intervals. **g**, Fraction of cells related to T-cell clones across different T-cell subsets (Extended Data Fig. 3a). FDR-adjusted two-tailed unpaired Mann–Whitney U-test. $n = 25$ breast cancer patients for whom TCR-seq data was available. **h**, The propensity of two cells from two breast cancer T-cell subsets to belong to the same clone. Data were calculated by sampling 10,000 pairs of cells and comparing clonal sharing characteristics to 10,000 pairs of cells sampled after shuffling clone identities. $*P < 0.05, **P < 0.001, ***P < 10^{-5}$.
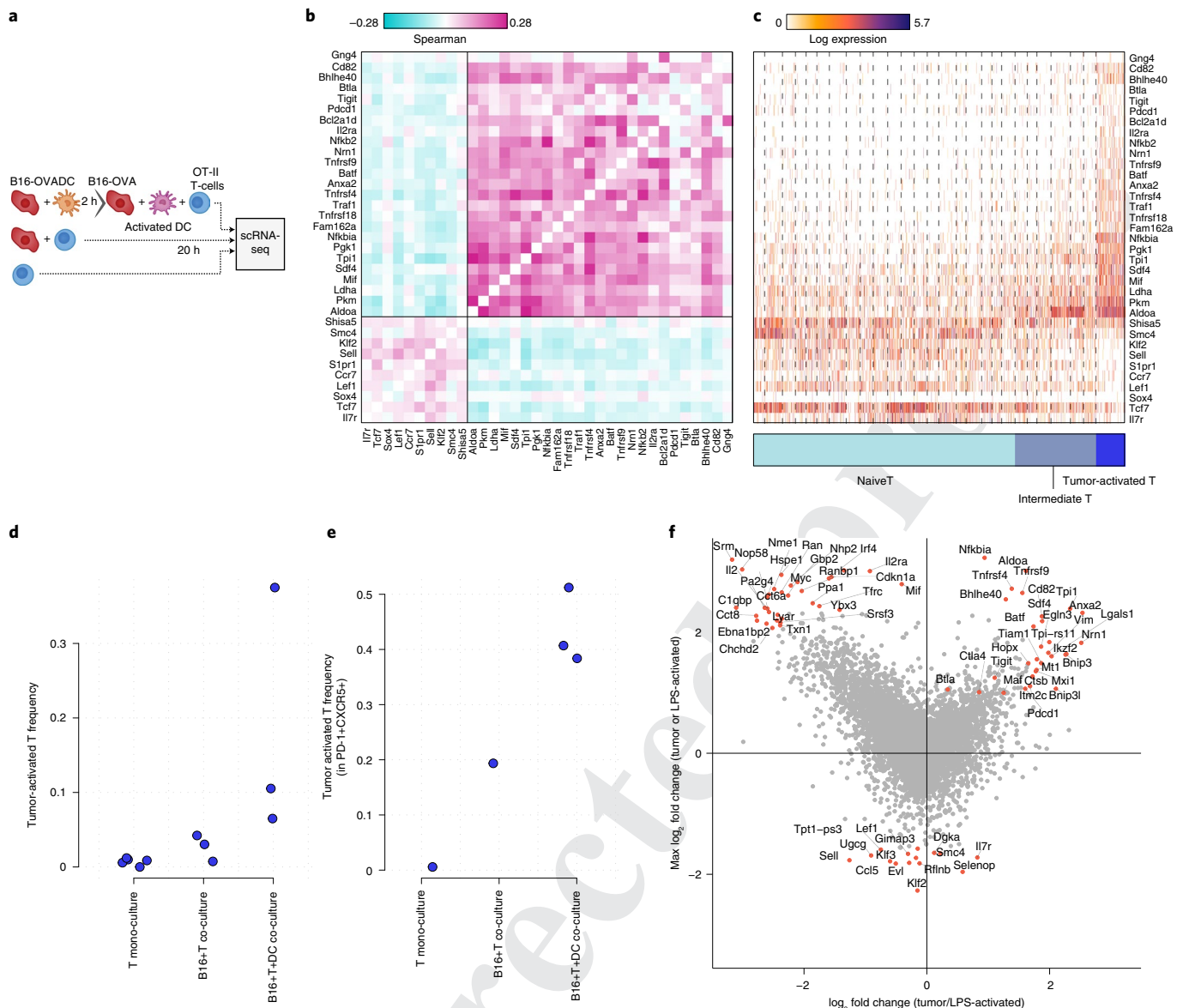
Fig. 4 | Differentiation to murine Tht cell state requires tumor antigen presentation by DCs. a, Schematics of the in vitro co-culture experiment. Splenic OT-II CD4+CD45.1+ T cells were sorted 20 h following culturing in three different conditions: (1) with B16-OVA and DCs, (2) with B16-OVA and (3) in a monoculture of OT-II CD4+CD45.1+ T cells. **b**, Pairwise gene Pearson correlations across T cells from all experimental conditions. **c**, Gene expression profiles of 1,860 single T cells collected from the three conditions, grouped into metacells. MetaCell annotation to three stages of T-cell differentiation (bottom). **d,e**, Frequency of the tumor-activated T-cell subset in the TCRβ+ (**d**) or TCRβ+CXCR5+PD-1+ (**e**) populations across experimental conditions. Each circle represents a biological replicate. **f**, Differential expression between tumor-activated OT-II T cells from **c** and OT-II T cells co-cultured with OVA-loaded DCs exposed to LPSs (log2 fold change, x axis). Gene expression of each activated subset (tumor or LPSs) was compared to naive T cells from the same experiment and the pairwise maximum log2 fold change, indicating upregulation in at least one condition, was calculated (y axis). Data summarize two independent experiments; n = 4 B16 cells + T cells + DCs; 4 B16 cells + T cells and 5 T-cell-alone independent cultures.

OT-II T cells expressed the co-stimulation and activation gene module (such as *Tpi1*, *Pkm*, *Bhlhe40*, *Btla*, *Ctla4*, *Tigit*, *Cd82*, *Gng4*, *Pdcd1*, *Tiam1* and *Maf*), which we call a mouse Tht-cell (mTht) signature (Fig. 4f and Supplementary Table 4). Notably, while the human and mouse Tht-cell programs are mainly conserved, *Cxcl13*, a hallmark gene in human Tht cells, was absent from the mTht cell signature. Taken together, our in vitro data demonstrate that the tumor-specific CD4+ T-cell gene module is driven in an antigen-specific manner upon interactions with tumor antigen-primed DCs.

mTht cells are restricted to tumor-draining lymph nodes and TME. To better understand the temporal and spatial differentiation

dynamics of the mTht cells in vivo, we utilized a mouse model in which we adoptively transferred splenic CD45.1+CD4+ OT-II T cells 6 d following subcutaneous injection of B16-OVA cells into WT C57/6J recipients (Fig. 5a). MARS-seq was performed on 12,154 QC-positive single TCRβ+ T cells isolated from tumor-draining lymph nodes (tdLNs) and peripheral cervical lymph nodes (cLNs), 10 and 17 d following tumor injection (Fig. 5a and Extended Data Fig. 5a). Ten days after tumor injection, CD45.1+TCRβ+ OT-II T cells were more frequent in the tdLNs than in the cLNs (P = 0.0078; two-tailed paired Mann–Whitney U-test; Extended Data Fig. 5b). We found that antigen-specific CD45.1+TCRβ+ OT-II T cells, but not polyclonal TCRβ+ T cells, upregulated the mTht cell signature

derived from the in vitro study (Fig. 5b and Supplementary Table 4; P (cLN) = 0.029 and P (dLN) = 0.00029; two-tailed Mann–Whitney U-test). This effect was observed in the tdLNs as early as day 10 after tumor injection. We found significant upregulation of the mTht cell program in CD45.1$^+$TCRβ$^+$ T cells from the tdLNs compared to the peripheral cLNs (tdLN versus cLN, P = 0.015). At a later time point, 17 d after injection, expression of the mTht cell program was diminished in the tdLNs (P = 0.023; Fig. 5b).

We found that tdLN CD45.1$^+$TCRβ$^+$ OT-II T cells, but not their cLN counterparts, upregulated a T-cell activation signature reminiscent of the mTht cell program that we observed in vitro, including *Tigit*, *Maf*, *Il21*, *Tiam1*, *Pkm* and *Tnfrsf4* (Fig. 5c and Supplementary Table 4). The in vivo mTht cell signature was associated with diminished expression of naive T-cell genes and correlated with expression of cell-cycle genes (Extended Data Fig. 5c–e). As Tht cells uniquely express genes considered as follicular helper T cell (T$_{FH}$) markers (such as *CXCL13* and *CXCR5*), we further compared the mTht cell signature to classical murine T$_{FH}$ cells. We applied an established mouse model of recombinant vesicular stomatitis virus (rVSV) infection (Methods), based on a previous study showing that CD4$^+$ T cells differentiated mostly to T$_{FH}$ cells upon infection with VSV[31]. We adoptively transferred naive SMARTA CD4$^+$CD45.1$^+$ T cells into C57BL/6 WT mice 24 h before subcutaneous intra-footpad infection with rVSV. On day 5 after infection, we isolated CD4$^+$CD45.1$^+$ polyclonal T cells and viral-specific CD4$^+$CD45.1$^+$ICOS$^+$PD-1$^+$CXCR5$^+$ T$_{FH}$ cells from footpad-draining popliteal lymph nodes and performed MARS-seq on 1,293 QC-positive single T cells. While we found high correlation between the mTht cell and rVSV-induced mouse T$_{FH}$ cell transcriptional programs, we also observed substantial differences related to mTht cell elevated expression of *Tnfrsf4*, *Il21*, *Eomes* and *Tiam1*, compared to both mouse T$_{FH}$ cells and in vitro LPS stimulated CD4$^+$ T cells (Fig. 5c–e). Notably, we could not detect *Cxcl13* expression in either cell type. We further verified that mTht and human Tht cells were homologous, in particular when compared to CD4$^+$ T$_{reg}$ cells and naive T cells (Fig. 5f). Compared to T$_{reg}$ cells, both human and mouse Tht cells lack expression of *Foxp3* and *Il2ra* and upregulate *Bhlhe40*, *Cd200*, *Hif1a*, *Tigit*, *Pdcd1*, *Tnfsf8*, *Il21* and *Nfatc1* (ref. [32]) (Fig. 5f). Notably, human Tht cells, but not mTht cells, also upregulated molecules related to T-cell cytotoxicity (such as GZMB, CCL5 and HOPX; Fig. 5f). Taken together, our results reveal the unique molecular signature of human and mouse Tht cells compared to other known CD4$^+$ T subsets, including, classical murine T$_{FH}$ and T$_{reg}$ cells.

We next explored the cell dynamics of mTht cells and their possible migration from tdLNs to the TME. For this purpose, we collected and profiled endogenous TCRβ$^+$ T cells and adoptively transferred OT-II CD45.1$^+$TCRβ$^+$ T cells from the TME on days 10 and 17, in addition to cells from the tdLNs and cLNs on day 10

following tumor injection. We used K-means clustering to identify 13 gene expression patterns that are either uniquely expressed or shared between the different conditions (cell types, tissues and time points; Fig. 5g). We found that tdLN-derived OT-II cells at day 10 maintained a naive state compared to tumor-infiltrating T cells (*Tcf7*, *Lef1*, *Sell* and *Ccr7*; clusters I and III; Fig. 5g), while also upregulating an mTht-related activation program (*Cd200*, *Btla* and *Nfatc1*; cluster II, *Tiam1*, *Hif1a* and *Il21*; cluster IV, *Maf*, *Batf* and cluster XII). Notably, the mTht cell gene signature first appeared in TME-derived OT-II infiltrating T cells only on day 17 (cluster IV and XII), with the upregulation of additional mTht-related genes (*Pkm*, *Mif* and *Tigit*; cluster XIII). Antigen-specific OT-II T cells from the TME were significantly different from their polyclonal, mostly CD8$^+$ T cell, counterparts (Fig. 5g).

Confocal imaging confirmed the presence of CD45.1$^+$PD-1$^+$ mTht cells in the tdLNs (Fig. 5h) and in the TME (Fig. 5i). Lower numbers of CD45.1$^+$ OT-II cells in cLNs compared to tdLNs, observed by FACS quantification and confocal imaging, further corroborated their accumulation specifically in the tumor niche (Extended Data Fig. 5b,f). Notably, mTht cells within the tdLN were localized in proximity and with physical contact to CD11c$^+$ myeloid cells (Fig. 5h).

We next investigated whether the mTht cell state was primed upon physical interactions and antigen presentation by myeloid cells in the tdLNs. Focusing on day 10 after injection when mTht cells accumulate in the tdLNs, we isolated and performed MARS-seq on 3,406 QC-positive CD11c$^+$ myeloid cells from the tdLNs and cLNs and analyzed them using the MetaCell package (Extended Data Fig. 6a). We additionally analyzed single TCRβ$^+$ and CD45.1$^+$TCRβ$^+$ T cells from the cLNs and tdLNs, grouping the metacells into naive T, proliferating T, bystander CD8$^+$ T, cytotoxic CD8$^+$ T, OT-II mTht$^-$ and OT-II mTht$^+$ cell subsets (according to whether they express high or low levels of the mTht cell gene signature; Supplementary Table 4 and Extended Data Fig. 6b). PIC-seq analysis of 1,939 TCRβ$^+$CD11c$^+$ PICs facilitated comparative composition and transcriptional analysis of myeloid and T cells interaction in tdLNs and cLNs (Extended Data Fig. 6c,d).

We observed significant depletion of naive and CD8$^+$ T cells in tdLN PICs (P = 0.012 and 0.0039, respectively, false discovery rate (FDR)-adjusted two-tailed Mann–Whitney U-test), alongside enrichment of OT-II mTht$^+$ (P = 0.0039) and cytotoxic CD8$^+$ T cells (P = 0.0039) in tdLN PICs (Extended Data Fig. 6e,f). Comparing gene expression observed in PICs relative to values expected by the singlet model in OT-II mTht$^+$ and OT-II mTht$^-$ (Extended Data Fig. 6g), revealed that OT-II mTht$^+$ PICs exclusively upregulated general T-cell activation genes in response to antigen stimulation (*Npm1*, *Mif* and *Tpi1*)[33,34]. Moreover, the OT-II mTht$^+$ PIC-specific gene signature was characterized by a pronounced induction of *Tigit*, *Eomes* and *Xcl1* (Extended Data Fig. 6g). These results suggest

Fig. 5 | Differentiation to mTht cell state is restricted to tdLNs and tumor site. a, Schematics of the experimental design. **b**, Distribution of the mTht cell signature in 11,292 total T (TCRβ$^+$) and OT-II T (CD45.1+TCRβ$^+$) single cells collected from the cLNs and tdLNs, at days 10 and 17 following tumor cell injection. The central mark in the box plot is the median, with 5th and 95th percentiles at the whiskers, 25th and 75th percentiles at the box and minima and maxima marked by dots. Circles indicate median signature in biological replicates. Data summarize two independent experiments; n = 6 day 9 tdLNs; 7 day 17 tdLNs; 8 day 10 cLNs and 3 day 17 cLNs. **c–e**, Differential gene expression (log$_2$ fold change) between CD45.1$^+$TCRβ$^+$ OT-II T cells from tdLNs and cLNs (**c**), LN-derived CD4$^+$CD45.1$^+$ICOS$^+$PD-1$^+$CXCR5$^+$ T$_{FH}$ cells after rVSV infection and CD45.1$^+$ T cells from non-infected mice (**d**) and between tumor-activated T cells and LPS-activated T cells from in vitro experiments (Fig. 4c) (**e**). **f**, Comparing human and mouse Tht cells. Differential gene expression (log$_2$ fold change) between the top 100 cells featuring highest tumor-activation signature in the tdLN and LN-derived T$_{reg}$ cells (left) or polyclonal tdLN TCRβ$^+$ cells (right; red), compared to human Tht cells versus T$_{reg}$ cells (left) or human Tht cells versus naive cells (right) from human breast TME (cyan, Extended Data Fig. 3a). **g**, K-means clustering of pooled tumor-activated OT-II T cells from day 10 tdLNs, day 10 TME and day 17 TME, as well as polyclonal T cells from day 10 and 17 TME. Values represent log$_2$ enrichment over the row mean; n = 4 day 10 TME; n = 2 day 17 TME. **h**, Representative confocal microscopy images of tdLNs (left) and tumor (right) sections extracted 10 and 17 d following tumor cell injection, stained for CD45.1$^+$ (OT-II) T cells, PD-1 and CD11c proteins. Scale bar, 30 µm; arrows indicate conjugates; asterisks indicate OT-II mThts. mCherry$^+$ B16-OVA tumor cells and PE-conjugated CD11c$^+$ DCs were identified by exclusive morphological properties. Inset highlights conjugates of mTht cells with CD11c$^+$ DCs. Images are representative of two independent experiments *P < 0.05, **P < 0.001, ***P < 10^{-5}.

a possible differentiation and activation trajectory into a mature mTht cell state, which is shown here to be dependent on the physical stimulation by antigen-presenting DCs.

Tht cells facilitate ICB antitumor response. PD-1-elevated expression in Tht cells (Fig. 3a), as well as Tht cell high affinity to tumor antigens, prompted us to investigate whether Tht cells play a promi-

nent role in effective anti-PD-1 therapy. Analysis of human breast cancer lesions before and 9 d after anti-PD-1 blockade[28] (Fig. 3) revealed a significant increase in both Tht-I and Tht-II numbers upon treatment across patients ($P = 0.019$ and 0.0078, respectively; two-tailed paired Mann–Whitney U-test; Fig. 6a). This result implicates an immediate and potentially direct response of both Tht-cell subsets to anti-PD-1 ICB. To directly test the function of mTht cells
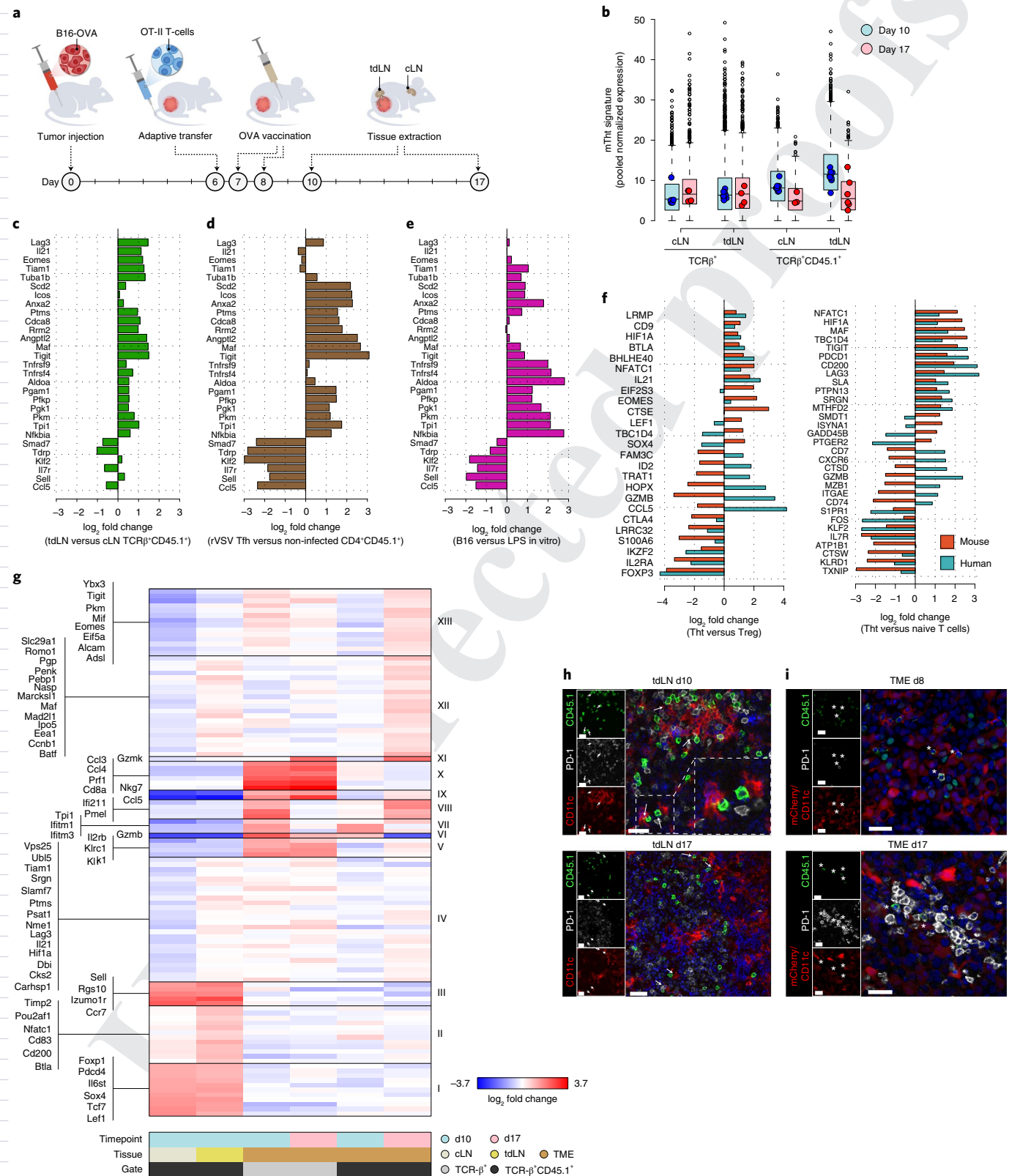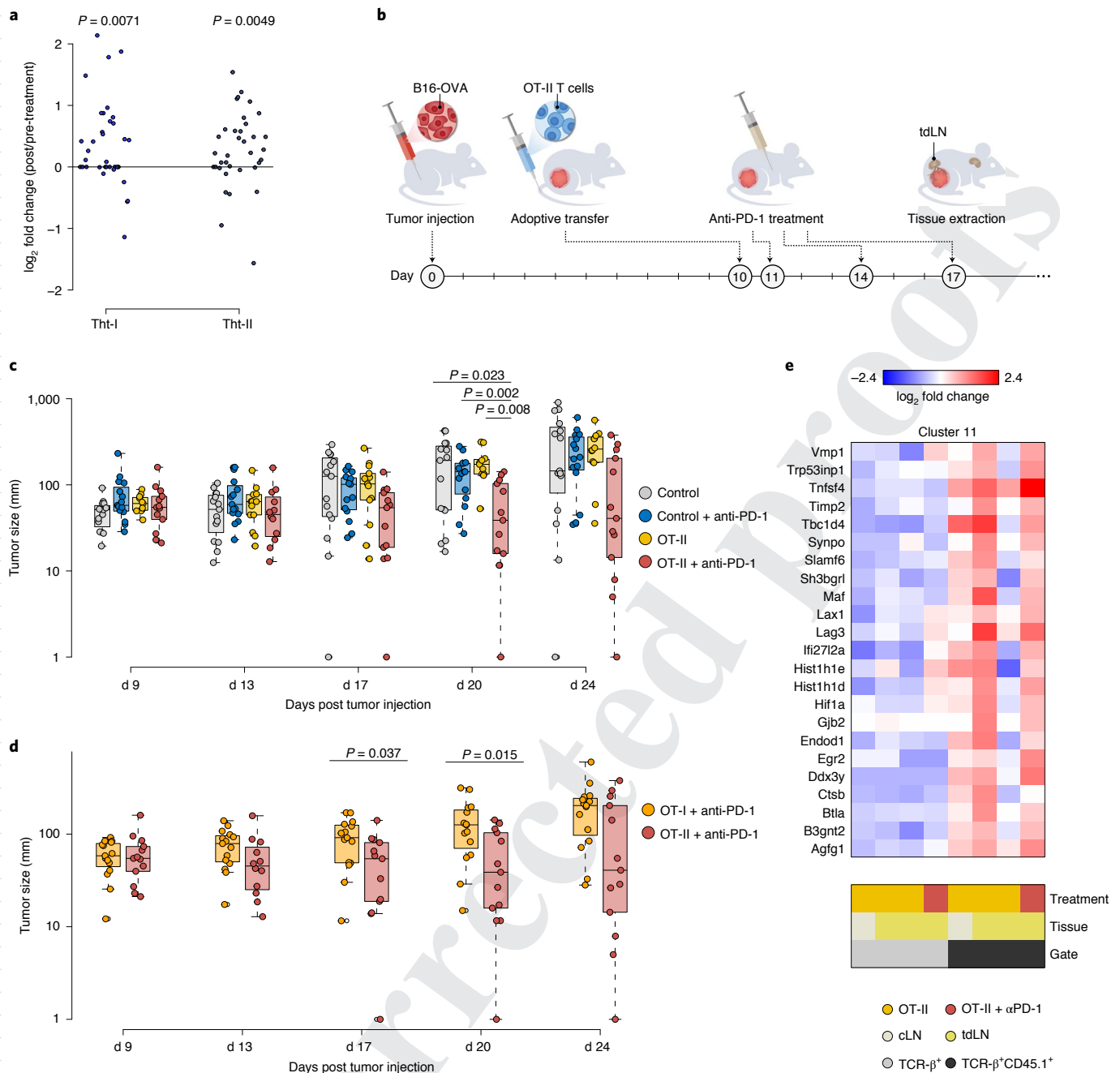
Fig. 6 | mTht cell role in response to anti-PD-1 immunotherapy. a, Changes in Tht-I and Tht-II numbers in breast cancer TME following anti-PD-1 treatment. Values represent log$_2$ fold change between post- and pre-treatment biopsies. Each dot represents one patient. Two-tailed paired Mann–Whitney U-test; $n = 36$ patients with breast cancer. **b**, Experimental design. Mice were injected subcutaneously with B16 melanoma cells presenting OVA. Adoptive transfer of OT-II CD4$^+$CD45.1$^+$T cells was performed 10 d after tumor cell injection and followed by three anti-PD-1 injections on days 11, 14 and 17. **c**, Tumor growth measurements of OT-II-transferred mice treated with anti-PD-1 (OT-II + anti-PD-1), compared to only anti-PD-1 treated mice (control + anti-PD-1), only OT-II-transferred mice (OT-II) and control WT mice (control); $n = 13$ OT-II + anti-PD-1; 13 OT-II; 15 control + anti-PD-1; and 15 control WT mice. **d**, Tumor growth measurements of OT-II- or OT-I-transferred B16-OVA tumor-bearing mice, treated with anti-PD-1. FDR-adjusted pairwise two-tailed Mann–Whitney U-test; $n = 13$ OT-II + anti-PD-1; and 15 OT-I + anti-PD-1 mice. The central mark in box plots is the median, with 5th and 95th percentiles at the whiskers, 25th and 75th percentiles at the box and minima and maxima marked by dots. FDR-adjusted pairwise two-tailed Mann–Whitney U-test. Data are representative of two independent experiments with similar results. **e**, K-means analysis of TCRβ$^+$ and CD45.1$^+$TCRβ$^+$ OT-II cells derived from tdLNs of mice treated with anti-PD-1 or control 17 d after tumor injection. Day 10 tdLN and cLN T cells from Fig. 5b were included for comparison. Shown are genes from cluster 11 (Extended Data Fig. 7b). Values represent log$_2$ fold change over the median.

in tumor response in the context of anti-PD-1 treatment, C57/6J WT mice were adoptively transferred with splenic CD45.1$^+$CD4$^+$ OT-II T cells 10 d following subcutaneous injection of B16-OVA

and then treated with anti-PD-1 at days 11, 14 and 17 following tumor injection (Fig. 6b). We found that OT-II-adoptively transferred mice treated with anti-PD-1 underwent significant tumor

545 reduction, starting from day 17 following tumor injection (Fig. 6c).
546 Notably, tumor response required the combined effect of mTht cells
547 and anti-PD-1 treatment, as tumors in mice that received only OT-II
548 or only PD-1 blockade grew similarly to control mice (Fig. 6c). Of
549 note, in this specific model, the effect of anti-PD-1 efficacy was sig-
550 nificantly greater when combined with CD4[+] OT-II compared to
551 CD8[+] OT-I T cells (Fig. 6d).
552 To better define the molecular reprograming of mTht cells upon
553 anti-PD-1 treatment, we performed MARS-seq on 2,867 polyclonal
554 TCRβ[+] and OT-II CD45.1[+]TCRβ[+] T cells from tdLNs of adoptively
555 transferred mice 17 d following tumor injection, with and without
556 administration of anti-PD-1. We used *K*-means clustering to iden-
557 tify coordinated transcriptional changes compared to day 10 cLN
558 and tdLN polyclonal and OT-II T cells (Extended Data Fig. 7a).
559 Anti-PD-1 treatment induced a strong type I interferon (IFN) gene
560 expression in both polyclonal and OT-II T cells (cluster 4; Extended
561 Data Fig. 7b). Of note, we observed that only OT-II T cells, in mice
562 treated with anti-PD-1, induced a gene expression program (such as
563 *Maf*, *Hif1a*, *Btla*, *Tnfsf4* and *Nfatc1*) reminiscent of day 10 mTht cells
564 from tdLNs (cluster 11; Fig. 6e and Extended Data Fig. 7c), suggest-
565 ing that anti-PD-1 blockade causes tumor-specific T cells to retain
566 their mTht cell signature for a longer duration in tdLNs. Together,
567 these results point to a direct involvement of Tht cells in the anti-
568 tumor efficacy of anti-PD-1 treatment and highlight the molecular
569 reprograming of mTht cells upon anti-PD-1 treatment.

570
Discussion
571
572 We reported the common physical interaction between CD4[+]PD-
573 1[+]CXCL13[+] T cells, which we named Tht cells, and DC-LAMP[+]
574 mregDCs in human NSCLC lesions. We show that Tht cells pref-
575 erentially interact with DCs than effector CD8[+] T cells within
576 TME. We further show, using an experimental mouse model, that
577 Tht cell differentiation, expansion and maintenance are correlated
578 with tumor antigen presentation in tdLNs and within the TME. The
579 unbiased nature of PIC-seq for unraveling cell–cell physical inter-
580 actions provides unexpected insights into the complex interaction
581 dynamics driving tumor immunity. Our data suggest the presence
582 of at least two distinct T-cell populations, CD8[+] effector/dysfunc-
583 tional and Tht cells, which respond specifically to tumor antigen
584 presentation by clonal expansion that reshapes the TME. However,
585 we still cannot describe how such interactions are combined with
586 additional active niches in the TME, such as T_reg, effector T, myeloid
587 and NK cells, to promote or repress tumor killing, in particular in
588 response to therapy.
589 Tht cells were previously classified as putative T_FH cells or
590 exhausted CD4[+] T cells in melanoma, colorectal and breast tum
591 ors[11,19,35,36], but their antigen specificity, molecular characterization,
592 function and specificity in the TME is poorly understood. Here we
593 showed that mouse Tht cells are characterized by a gene expression
594 signature shared by the classical mouse T_FH cells induced by rVSV
595 infection, with a few important distinctions. The Tht cells express a
596 unique gene signature consisting of B cells and DC-attraction and
597 activation cytokines (CXCL13, IL21 and XCL1) and immune check-
598 points, including high levels of PD-1, CTLA4 and TIGIT. The prev-
599 alence of the Tht cell signature in different human cancer types was
600 shown to correlate with better survival[37] or response to ICB ther-
601 apy[25] in colorectal cancer and with enhanced antitumor immunity
602 in breast cancer lesions[35,36]. The data we present here highlight Tht
603 cells as a potential hub of interaction with APCs in the TME and a
604 direct target of anti-PD-1 ICB, suggesting that the role of these cells
605 in promoting successful immunotherapy is likely to be involved
606 with their interactive capacity. Whether immunotherapy triggers
607 the depletion of the Tht cell niche and subsequent effector func-
608 tion de-repression or induction of de novo Tht cell differentiation
609 that promotes the helper activity and release of immunomodulating
610 cytokines in the TME, is still unclear. Some of these behaviors may

be linked to data from recent studies that unraveled the key contri-
butions of tissue-resident memory T cells to microbial and tumor
immunity[38,39], or to recent reports on induction of antigen-specific
CD4[+] T cells following neoantigen vaccination in glioblastoma[40].

Our data also unravel the centrality of mregDCs in TME orga-
nization. Two studies have showed that response to PD-1 blockade
requires CD28 engagement on T cells[41,42]. As CD28 is engaged by
B7, expressed mainly by APCs, these studies emphasize the need for
APCs for maximal response to PD-1 blockade. In the context of our
results, these findings may suggest that APCs may actively engage
or re-engage Tht cells upon PD-1 blockade, leading to their differ-
entiation or expansion rather than simply perturbing their static
functions in the TME. More in-depth T cell–DC interaction map-
ping using PIC-seq, as we introduce here, combined with the rap-
idly expanding set of tumor single-cell atlases and spatial single-cell
tumor maps[43–45] have the potential to transform our understanding
of the mechanisms that control durable response to immunotherapy
and guide the development of therapeutic strategies to harness this
knowledge.

Methods

Human samples. Tumor and adjacent healthy lung tissues were obtained from
surgical specimens of NSCLC (Supplementary Table 1) patients undergoing
resection at the Mount Sinai Medical Center after obtaining informed consent
in accordance with a protocol reviewed and approved by the IRB at the Icahn
School of Medicine at Mount Sinai and in collaboration with the Biorepository and
Department of Pathology and the Weizmann Institute of Science.

Mice. C57BL/6 WT female mice were obtained from Harlan. TCR-transgenic
OT-II male mice (harboring OVA-specific CD4[+] T cells) were a kind donation
from the laboratory of N. Friedman of the Weizmann Institute of Science. Mice
were housed under specific-pathogen-free conditions at the Animal Breeding
Center of the Weizmann Institute of Science and were used at age 8–10 weeks.
SMARTA[46] donor mice were obtained through the Swiss Immunological Mouse
Repository (SwImMR). All animals were handled according to the regulations
formulated by the Institutional Animal Care and Use Committee. Mice were
housed in conditions of 12/12 h dark/light cycle, 22 ± 1 °C ambient temperature
and 50 ± 10% humidity.

Cell lines. B16 melanoma cells (H2b), stably expressing chicken ovalbumin
(B16-OVA) and mCherry fluorescent labeling, were kindly provided by L.
Eisenbach's laboratory and maintained in DMEM (Invitrogen) supplemented
with 10% FCS, 1 mM L-glutamine, 100 U ml[−1] of penicillin and 100 mg ml[−1] of
streptomycin in a humidified 5% CO_2 atmosphere at 37 °C.

In vitro cultures. Cells were isolated from the spleens of 8-week-old C57BL/6
female mice. Splenocytes were washed and suspended in red blood lysis buffer
(Sigma-Aldrich) and DNase I (0.33 U ml[−1], Sigma-Aldrich), incubated for 5 min
at room temperature, washed twice with cold PBS, filtered through a 70-μm cell
strainer, centrifuged at 400*g* for 5 min at 4 °C and then resuspended in ice-cold
sorting buffer (PBS supplemented with 0.2 mM ethylenediaminetetraacetic acid,
pH 8, and 0.5% BSA). DCs derived from spleen tissues were enriched from the
single-cell suspension first by negative magnetic selection for CD19 and second
by positive magnetic selection for CD11c. Briefly, cells were incubated with CD19
MicroBeads (Miltenyi Biotech) for 15 min at 4 °C, washed and run through a
MACS column (Miltenyi Biotech). A negative fraction of cells was collected for
further incubation with CD11c MicroBeads Ultrapure (Miltenyi Biotech) for
10 min at 4 °C and a positive fraction of CD11c[+] cells was collected. After cell
counting 1×10^6 DCs were co-cultured with 1×10^6 B16-OVA cells in 96-well
U-bottomed tissue culture plates in 200 μl of standard medium RPMI-1640
supplemented with 10% fetal calf serum (FCS), 1 mM L-glutamine, 100 U ml[−1]
penicillin and 100 mg ml[−1] streptomycin (Biological Industries) for 2 h at 37 °C. In
parallel, we isolated CD4[+] T cells from spleens of CD45.1 OT-II male mice. Again,
cells were washed and suspended in red blood lysis buffer (Sigma-Aldrich) and
DNase I (0.33 U ml[−1], Sigma-Adrich), incubated for 5 min at room temperature,
washed twice with cold PBS, filtered through a 70-μm cell strainer and centrifuged

at 400g for 5 min at 4 °C. CD4$^+$ T cells were enriched by the CD4$^+$ T-cell Isolation kit, according to the manufacturer's instructions (Miltenyi Biotech). Briefly, splenocytes were incubated with biotin antibody cocktail for 5 min and afterward anti-biotin microbeads were added for an additional 10 min incubation. We collected the CD4$^+$ T cells, which were the unlabeled fraction. For culture experiments, we performed triple culture of CD4$^+$ T, DCs and B16-OVA cells, co-cultures of CD4$^+$ T and B16-OVA cells and monocultures of CD4$^+$ T cells. Triple-, co-cultured and monocultured cells were seeded at a concentration of 1×10^6 cells ml^{-1} (1:1 ratio in co-cultures) and following 20 h. All cultures were carried out in the standard medium RPMI-1640 supplemented with 10% FCS, 1 mM L-glutamine, 100 U ml^{-1} penicillin and 100 mg ml^{-1} streptomycin.

Mouse tumor model. For investigation of physical interactions in TME, tdLNs and cLNs, 2×10^6 B16-OVA tumor cells were suspended in 100 µl PBS and injected subcutaneously (s.c.) into 8-week-old female mice. CD4$^+$ T cells were isolated from splenocytes of OT-II male mice. Briefly, splenocytes were washed and suspended in red blood lysis buffer (Sigma-Aldrich) and DNase I (0.33 U ml^{-1}, Sigma-Aldrich), incubated for 5 min at room temperature, washed twice with cold PBS, filtered through a 70-µm cell strainer and centrifuged at 400g for 5 min at 4 °C. CD4$^+$ T cells were enriched by the CD4$^+$ T-cell Isolation kit, according to the manufacturer's instructions (Miltenyi Biotech). CD4$^+$ T cells were collected and intravenously injected (2×10^6 CD4$^+$ cells per mouse) 6 d following B16-OVA cell injection. Mice were then injected intraperitoneally (i.p.) with OVA 323-339 (10 µg per mouse; InvivoGen) diluted in PBS containing 30% alum adjuvant (Imject Alum adjuvant; Thermo Fisher Scientific), at days 7 and 8 following B16-OVA cell injection. Solid tumors, tdLNs and cLNs were collected 10 and 17 days following B16-OVA cell injection (Fig. 5a).

For in vivo functional experiments, 8–10-week-old female mice were injected s.c. with 2×10^6 B16-OVA tumor cells suspended in 100 µl PBS on their right flank. CD4$^+$ T cells were isolated from spleens of OT-II male mice and 3×10^6 cells were i.v. injected to mice 10 d following tumor cells injection. To examine the efficacy of PD-1 blockade in relation to OT-II T-cell function, mice were i.p. injected with 250 µg of anti-PD-1 (clone RMP1-14, BioXcell) or isotype control (rat IgG2a, BioXcell) on days 11, 14 and 17 following tumor cell injection. Tumor size was measured blindly to the conditions of the experiments by a caliper every 2 d and tumor volume was calculated by measuring two diameters and using the formula: $X^2 \times Y \times 0.52$ (X, smaller diameter; Y, larger diameter). Solid tumors, tdLNs and cLNs were also collected 17 d following tumor cell injection for scRNA-seq experiments (Fig. 6b). Mice were monitored so that the maximal tumor size of 1.5 cm in diameter was not exceeded. No statistical methods were used to predetermine sample sizes; our distribution of tumor sample sizes was similar to those reported in a previous publication[16].

Tissue dissociation. To derive a single-cell suspension in human NSCLC specimens, 0.1–0.4 g of tumor and adjacent healthy tissues were cut into small pieces and then mechanically dissociated by pipetting. Tissues were suspended with CO$_2$ Independent Medium (Thermo Fisher Scientific) supplemented with DNase (100 µg ml^{-1}, Sigma-Aldrich) and collagenase IV (0.5 mg ml^{-1}, Worthington) and incubated at 37 °C for 20 min, with frequent agitation.

To achieve single-cell suspensions in the murine model, tumors were cut into small pieces and suspended with RPMI-1640 supplemented with DNase (28 µg ml^{-1}, Sigma-Aldrich) and collagenase IV (1 mg ml^{-1}, Worthington). Tissues were homogenized by GentleMacs tissue homogenizer (Miltenyi Biotec) and incubated at 37 °C for 10 min, with frequent agitation. This tissue dissociation procedure was performed twice for each tumor.

To achieve single-cell suspensions from tdLNs and cLNs, tissues were digested in Iscove's modified Dulbecco's medium (Sigma-Aldrich). For mild dissociation, tissues were supplemented with Liberase-TL (100 µg ml^{-1}, Roche) and DNase I (100 µg ml^{-1}, Roche) and incubated with frequent agitation at 37 °C for 20 min.

After all dissociation procedures, cells were washed with cold PBS, filtered through a 70-µm/100-µm cell strainer and centrifuged at 380g for 5 min at 4 °C.

Flow cytometry and sorting. Cells were suspended in ice-cold sorting buffer (PBS supplemented with 0.2 mM ethylenediaminetetraacetic acid, pH 8, and 0.5% BSA) supplemented with anti-mouse CD16/32 (BD Bioscience) or anti-human TruStain FcX (BioLegend) to block Fc receptors before labeling with fluorescent antibodies against cell-surface epitopes. Murine samples were stained using the following anti-mouse antibodies: eFluor450-conjugated TER-119, eFluor450-conjugated NK1.1, PerCP Cy5.5-conjugated streptavidin and Pacific blue-conjugated CD19 (eBioscience); and PerCP Cy5.5-conjugated TCRβ, FITC-conjugated TCRβ, APC-Cy7-conjugated CD11c, APC-conjugated CD45.2, PE-Cy7-conjugated CD45.1, APC-conjugated CD279 (PD-1) and biotin-conjugated CXCR5 (BioLegend). Human samples were stained using the following anti-human antibodies: Pacific blue-conjugated CD235a, FITC-conjugated TCRβ, PE-conjugated CD11c and PE-conjugated CD64 (BioLegend); APC-conjugated CD45 (eBioscience) and Pacific blue-conjugated CD19 and PE-Cy7-conjugated CD56 (BD Biosciences). Before sorting, cells were stained with DAPI for evaluation of live/dead cells. Cell populations were sorted using either SORP-aria (BD Biosciences) or ARIA-III instrument (BD Biosciences) and analyzed using BD

FACSDIVA software (BD Biosciences) and FlowJo software. Isolated live cells were single-cell sorted into 384-well cell-capture plates containing 2 µl lysis solution and barcoded poly(T) reverse-transcription primers for single-cell RNA-seq[27,47]. Four empty wells were kept in each 384-well plate as a no-cell control during data analysis. Immediately after sorting, each plate was spun down to ensure cell immersion into the lysis solution and stored at −80 °C until processing.

rVSV infection model. Female mice at 8–10 weeks old were infected intra-footpad with 1×10^5 plaque-forming units (p.f.u.) of rVSV (a recombinant VSV expressing an LCMV glycoprotein recognized by SMARTA TCR-transgenic cells)[31]. Naive CD4$^+$ T cells from spleens of SMARTA female CD45.1 mice were negatively selected by magnetic isolation (Miltenyi Biotec), with purity >98%. SMARTA CD45.1$^+$ T cells (1×10^6) were injected i.v. into C57BL/6 CD45.2 recipients 1 d before intra-footpad rVSV infection.

Single-cell suspensions of footpad dLNs 5 d after infection were generated as described[31,48]. Single-cell populations from rVSV-infected mice (CD4$^+$CD45.1$^+$ or CD4$^+$CD45.1$^+$PD-1$^+$ICOS$^+$ cells) were sorted using the following flow cytometry antibodies: APC-CXCR5 (2G8; BD Biosciences), APC-Cy7-CD45.1 (A20; BioLegend), eFluor450-CD4 (RM4-5; eBioscience), BV605-ICOS (C398.4A; BioLegend), PE-PD-1 (J43; eBioscience), AxFl488-B220 (RA3-6B2; BioLegend), AxFl488-NK1.1 (PK136; BioLegend) and PE-Cy7-CD8a (53-6.7; BioLegend). Live/Dead Fixable Aqua Dead Cell Stain (Thermo Fisher Scientific) was used to exclude dead cells. Sorting was performed following exclusion of doublets, dead cells and B220$^+$ B cells, NK1.1$^+$ NK cells, CD8a$^+$ T cells and Ter119$^+$ erythrocytes.

Immunohistochemistry. For spatial examination in human specimens, FFPE sections were baked at 37 °C overnight, de-paraffinized in xylene and rehydrated in decreasing concentrations of ethanol. Tissue sections were incubated in citrate buffer (pH 6) for antigen retrieval at 95 °C for 30 min. After three PBS washes, we added blocking buffer (5% donkey serum in PBST and 0.1% Triton X-100) for 1 h at room temperature. After blocking, all primary antibodies were incubated at 4 °C overnight: rat DC-LAMP (1:200 dilution, Novusbio), goat CXCL13 (1:50 dilution, R&D Systems), rabbit CD4 (1:100 dilution, Novusbio), mouse CD272 (BTLA, 1:100 dilution, Abcam), rabbit PD-1 (1:100 dilution, Abcam) and rabbit PRDM1/Blimp1 (1:100 dilution, Abcam). After three PBST washes (0.01% Tween-20; Sigma-Aldrich) corresponding secondary antibodies were used simultaneously for 1 h at room temperature. After three PBST washes (0.01% Tween-20; Sigma-Aldrich), a TrueVIEW auto-fluorescence quenching kit (SP-8400) was applied before nuclei staining with DAPI for 6 min and coverslips were then mounted on slides with anti-fade mounting medium in the auto-fluorescence quenching kit. Mounted slides were kept in the dark before image acquisition.

For spatial examination in mouse specimens, frozen sections of tdLNs, cLNs and tumors were taken 10 and 17 d after B16-OVA cell injection. Lymph nodes and tumors were fixed in 4% PFA solution for 4 h and then transferred to 30% sucrose solution for 2 d. Tissues were embedded in Optimal Cutting Temperature compound (Sigma-Aldrich) and 10-µm sections were taken using a LEICA CM1950 machine. For visualization of T cells and DCs, following washes, the sections were first blocked with a blocking buffer solution (5% FBS, 1% BSA and 0.2% Triton) for 2 h at room temperature. Sections were incubated with primary antibodies overnight at 4 °C. The antibodies used were APC-conjugated CD279 (PD-1) (1:100 dilution; BioLegend), Alexa Fluor 488-conjugated CD45.1 (1:100 dilution; BioLegend) and PE-conjugated CD11c (1:100 dilution; BioLegend). Sections were washed three times with PBST (0.01% Tween-20; Sigma-Aldrich) and DAPI was added for 10 min to detect cell nuclei and before washing with PBST (0.01% Tween-20; Sigma-Aldrich). Sections were mounted with SlowFade (Invitrogen) and sealed with coverslips. Microscopic analysis was performed using a laser-scanning confocal microscope (Zeiss, LSM880). Images were acquired and processed using Imaris software (Bitplane).

Multiplex imaging analysis of human tissue sections. FFPE tissue sections (4 µm) were stained using the MICSSS protocol as previously described[49]. Briefly, slides were baked at 50 °C overnight, de-paraffinized in xylene and rehydrated in decreasing concentration of ethanol (100%, 90%, 70%, 50% and dH$_2$O). Sample slides were incubated in pH 6 or pH 9 of Target Retrieval Solution (Dako) at 95 °C for 30 min, then in 3% hydrogen peroxide for 15 min and in serum-free protein block solution (Dako) for 30 min. Primary antibody staining was performed using the optimized dilution during 1 h at room temperature or at 4 °C overnight, followed by signal amplification, using associated secondary antibody conjugated to horseradish peroxidase (HRP) during 30 min. Chromogenic revelation was performed using AEC (Vector). Tissue sections were counter-stained with hematoxylin, mounted with a glycerol-based mounting medium and finally scanned to obtain digital images (Aperio AT2, Leica). Then, the same slides were bleached and stained again, including specific blocking for previous similar species antibody staining (monovalent Fab fragment). Primary antibodies were used in the following order: anti-human PD-1 (polyclonal, Sigma-Aldrich), anti-human CD8 (clone C8/144B, Dako), anti-human CD3 (clone 2GV6, Ventana), anti-human CD68 (clone KP1, Dako), anti-human DC-LAMP (clone 1010E1.01, Novus Biologicals) and anti-human CD20 (clone L26, Dako).

MARS-seq library preparation. Single-cell libraries were prepared as previously described[27,47]. In brief, messenger RNA from cells sorted into cell-capture plates was barcoded and converted into complementary DNA and pooled using an automated pipeline. The pooled sample was then linearly amplified by T7 in vitro transcription and the resulting RNA fragmented and converted into a sequencing ready library by tagging the samples with pool barcodes and Illumina sequences during ligation, reverse transcription and PCR. Each pool of cells was tested for library quality and concentration assessed as described earlier. Primer barcodes and genes, for library preparations and for qPCR validations, were used according to the MARS-seq 2.0 protocol[27].

MARS-seq low level data processing. scRNA-seq libraries (pooled at equimolar concentration) were sequenced on an Illumina NextSeq 500 or NOVA-seq, at a median sequencing depth of 15,054 reads per cell. Sequences were mapped to the mouse genome (mm10) or human genome (hg38), demultiplexed and filtered as previously described[45,46] with the following adaptations. Mapping of reads was performed using HISAT (v.0.1.6); reads with multiple mapping positions were excluded. Reads were associated with genes if they were mapped to an exon, using the UCSC Genome Browser for reference. We estimated a median of 2% spurious unique molecular identifiers (UMIs) in the data using statistics on empty MARS-seq wells. Cells with <500 UMIs, >500,000 UMIs or >40% mitochondrial genes were excluded from analysis. We used the MetaCell package[23] to analyze all scRNA-seq data collected in this study and to derive background metacell covers of the single-cell populations for PIC-seq analysis, as described below. Default parameters were used unless otherwise stated.

PIC-seq summary. Assignment of PICs to their T-cell and myeloid cell identities was performed as previously described with necessary adjustments[22]. In short, PICs are modeled as a linear mixture of pairs of contributing cells. Each contributing cell (T cell or myeloid) belongs to a metacell from the respective T-cell or myeloid background models calculated over the singlet populations and its gene expression is sampled from the multinomial probability distribution of that metacell. The mixing factor, α, assigned for each PIC, denotes the fraction of UMIs contributed by the contributing T cell to that PIC.

The PIC-seq algorithm operates in two steps. First, it applies a linear regression model trained on synthetic PICs to infer α for each PIC. Second, it constructs all possible combinations of metacells from populations A and B mixed by α and calculates the expected gene expression distributions of these mixtures. A maximum likelihood estimator is applied on each PIC, to derive two metacells whose mixture is most likely to give rise to the PIC.

MetaCell model of human lymphocytes from NSCLC TME. We derived a MetaCell cover of TCRβ⁺ T and CD64⁺CD11c⁺ myeloid cells from ten matched biopsies of NSCLC tumors and adjacent healthy tissue. Mitochondrial genes and the highly variable immunoglobulin genes (*IGH*, *IGK* and IGL prefixes) were removed from the UMI tables. Gene features for MetaCell covers were selected using the parameter $T_{vm} = 0.2$, total UMI > 20 and more than three UMIs in at least three cells. We filtered the list of gene features used for MetaCell analysis from genes associated with cell cycle, immediate stress response and gene modules inducing strong patient-specific biases. To this end we first identified all genes with a correlation coefficient of at least 0.1 for one of the anchor genes *TOP2A*, *MKI67*, *PCNA*, *MCM4*, *UBE2C* (cell cycle), *HSPA1B*, *HSPA1A*, *HSP90AA1*, *DNAJB1*, *FOS*, *EGR1*, *IER3*, *FOSB* (stress response), *MTRNR2L8*, *RPS3AP5*, *MTRNR2L12*, *AC005912.1*, *RPS24P8*, *RPL10P9*, *MTCO2P12*, *MTND2P28*, *BIRC3*, *SYNE2*, *XIST*, *Y_RNA*, *SYNE1*, *FTH1*, *FTH1P8*, *RPS2P7* and *RPS27AP16* (patient specific). We then hierarchically clustered the correlation matrix between these genes (filtering genes with low coverage and computing correlation using a downsampled UMI matrix) and selected the gene clusters that contained the above anchor genes. We thus retained 373 genes as features. We used MetaCell to build a *k*-NN graph, perform boot-strapped co-clustering (500 iterations; resampling 70% of the cells in each iteration) and derive a cover of the co-clustering *k*-NN graph ($K = 50$). Outlier cells featuring gene expression higher than fourfold than the geometric mean in the metacells in at least one gene were discarded.

Detailed annotation of the different T and myeloid subsets was performed using hierarchical clustering of the MetaCell confusion matrix (Extended Data Fig. 1f) and supervised analysis of enriched genes according to literature (Fig. 1d,e). Two metacells, annotated as mast cells (by high expression of CPA3 and GATA2) and B cells (expressing CD79B) were removed from further analysis.

PIC-seq analysis of human lymphocytes. To derive genes that would serve as features for the linear regression, We selected the top 100 genes most correlated with total UMI count genes in the T cell and myeloid cell populations, as well the features in the MetaCell cover, retaining 498 features (Supplementary Table 5). Both synthetic and real PIC matrices were downsampled to 800 UMIs per cell (*numis* = 800). The R^2 value for estimating the mixing coefficient over the synthetic PIC was 65.8% (Extended Data Fig. 2a).

To derive genes for computing the maximum likelihood estimation (MLE) assignment, we chose the top 20 differential genes in each T cell (or myeloid cell)-associated metacell. We combined this set of genes with the genes used as

features in the metacell cover, but discarded genes that are highly differential both in the T cell and DC MetaCell models, as well as ribosomal and poorly annotated genes, retaining 337 genes (Supplementary Table 5). To validate the MLE assignment, we computed the error in assignments over 5,000 synthetic PIC (Extended Data Fig. 2b).

We further filtered putative PICs suspected as singlets as follows: We computed for each PIC its likelihood when modeled as a singlet originating from its assigned T (if $\alpha \geq 0.5$) or myeloid (if $\alpha < 0.5$) metacell. Putative PICs whose likelihoods when modeled as doublets were not greater than when modeled as singlets ($ll_{doublet} - ll_{singlet} \leq 0$) were determined as singlets and discarded from downstream analysis (Extended Data Fig. 2c). In addition, we noticed that PICs of TCRβ⁺ T cells and CD11c⁺CD64⁺ NK cells were poorly modeled by PIC-seq due to their high transcriptional resemblance ($R^2 = 40\%$ for α estimation; Extended Data Fig. 2a). Therefore, PICs assigned by PIC-seq to NK metacells were subsequently discarded from downstream analysis (Extended Data Fig. 2b).

Reanalysis of T cells from human melanoma and breast cancer. MetaCell analysis of 39,445 T cells from human melanoma and 67,538 was performed similarly to NSCLC. We found the same T-cell subsets in both cancer types, in addition to a subset featuring high expression of type I IFN response genes (*STAT1* and *IFIT1*; Extended Data Fig. 3a).

To derive a CD4⁺PD-1⁺CXCL13⁺ signature shared between NSCLC, breast cancer and melanoma, we pooled T cells from each tumor type by their T-cell subset annotation and extracted all genes differentially expressed between CD4⁺PD-1⁺CXCL13⁺ and any other T-cell subsets in any of the tumor types (fold change >2.8 or <−2.8, total of 763 genes; Supplementary Table 3). We then applied joint *K*-means clustering ($K = 38$) over normalized pooled gene expression profiles in both datasets (Extended Data Fig. 3b). Genes whose expression within Tht cells were twofold or higher than the pooled expression of all other T cells in all cancer types were used to define the human CD4⁺PD-1⁺CXCL13⁺ signature, resulting in a Tht cell signature of 32 genes (Fig. 3b and Supplementary Table 3).

To explore the clonal relationships of breast cancer (and melanoma) T cells, we used published clonal data based on TCR sequencing of 53,044 (5,596) T cells from 41 (21) patients[11]. We analyzed 25 (10) patient samples that contained TCR-seq data for at least seven annotated T-cell subsets and considered a cell to belong to a clone if it shared a TCR sequence with at least one other T cell from the same patient. To calculate clone sharing between T-cell subsets, we calculated the joint T-cell subset assignments of 10,000 sampled pairs of cells that share clonal data. We compared the results (log₂ fold change) to the joint T-cell subset assignments of 10,000 sampled pairs of cells that originate from the same patient but not necessarily from the same clone (Fig. 3h).

To explore Tht-I and Tht-II involvement in PICs, we extracted NSCLC PICs assigned to Tht cells and re-ran PIC-seq on a hybrid single-cell reference model, where the myeloid cells are derived from the NSCLC TME and the T cells are composed of breast Tht-I and Tht-II only.

MetaCell analysis of mouse scRNA-seq data. Two mouse scRNA-seq datasets were analyzed separately: (1) in vitro co-culture data and (2) an in vivo adoptive transfer experiment. For each dataset, we derived a MetaCell cover of all sequenced single cells. Mitochondrial genes and ERCC spike-ins were removed from the UMI tables. Gene features for MetaCell covers were selected using the parameter $T_{vm} = 0.1$, total UMI > 20 and more than three UMIs in at least three cells. We filtered the list of gene features used for MetaCell analysis from genes associated with cell cycle, tumor cells and erythrocytes and gene modules inducing strong batch-effect biases. To this end we first identified all genes with a correlation coefficient of at least 0.1 for one of the anchor genes: *Top2a*, *Mki67*, *Pcna*, *Mcm4*, *Ube2c*, *Hist1h1b* (cell cycle), *Gm10800*, *Gm10801*, *Gm22213*, *Gm22710*, *Trim30b*, *Trim30c*, *Gm10059*, *Tpt1-ps3*, *Gm7079*, *Ablim1*, *Mpp1*, *Gm24270*, *Gm10800*, *Gm23388*, *Bpifa2*, *Smr3a* (batch specific), *Mlana*, *Met* (tumor cells), *Hbb-a1* and *Hba-a2* (erythrocytes). We then hierarchically clustered the correlation matrix between these genes (filtering genes with low coverage and computing correlation using a downsampled UMI matrix) and selected gene clusters that contained the above anchor genes. We thus retained 287 (dataset I) and 460 (dataset II) as features (Supplementary Table 5). We used MetaCell to build a *k*-NN graph, perform boot-strapped co-clustering (500 iterations; resampling 70% of the cells in each iteration) and derived a cover of the co-clustering *k*-NN graph ($K = 30$ in dataset I and $K = 50$ in dataset II). Outlier cells featuring gene expression higher than fourfold than the geometric mean in the metacells in at least one gene were discarded. Metacells featuring high levels of *Cd79b* (B cells), *Mzb1* (plasma cells) or *Malat1* were annotated as outliers and removed from subsequent analysis.

To compare the tumor-activation gene signature to general T-cell activation, we used our previously published data of 1,531 single TCRβ⁺ T cells isolated from co-cultures of OT-II CD4⁺ T and DCs exposed to LPS and OVA peptide[22]. A total of 688 monoculated T cells from the same experiment were used as controls to derived the LPS-activation gene signature (Fig. 4f). To compare the tumor-activated T cells from the cLNs and dLNs to LN T_{reg} cells, we used our previously published data on T cells from lymph nodes[22], including specific enrichment of T_{reg} cells by sorting for ICOS⁺TCRβ⁺ and TIGIT⁺TCRβ⁺ T cells (Fig. 5e).

PIC-seq analysis of mouse model. To choose gene features for estimating the mixing factor, we first removed those with a strong cell-cycle signature and computed correlation with cell size as described above, retaining 464 genes (Supplementary Table 5). R^2 value was 61% (Extended Data Fig. 6c). Feature selection for the MLE assignment was performed similar to PIC-seq of human tumors, resulting in 337 genes used as features. To validate the MLE assignment, we computed the error in assignments over 5,000 synthetic PICs (Extended Data Fig. 6d).

Comparing observed and expected expression. In both human and mouse tumor models, we reconstructed the expected levels of a gene in each PIC as previously described[22]. In short, the expected expression of each gene in a certain PIC equals the α-weighted sum of the contribution from the T-cell part (which can be estimated from the characteristic multinomial distribution of the contributing T metacell) and the contribution from the myeloid cell part.

We used an FDR-adjusted chi-squared test to systematically scan for genes whose observed values diverge from expected in specific groups of PIC ($q < 10^{-5}$ in human PICs; Fig. 3c and $q < 10^{-4}$ in mouse PICs; Extended Data Fig. 6g).

Image analysis and quantification. Analysis of MICSSS images was performed with CellProfiler[50] and magick R package. MICSSS slides were scanned manually for areas containing TLSs, resulting in five different regions of interest from patient 5. For each region of interest, nuclei segmentation was performed on the DAPI channel with CellProfiler using the IdentifyPrimaryObjects method (typical diameter 10–30 pixels; 50-pixel window adaptive minimum cross-entropy thresholding method; 0.5 threshold-smoothing scale; distinguish clumped objects by intensity and draw boundaries by the propagate method) and the IdentifySecondaryObjects method with default parameters.

Multichannel images were loaded to R with the magick package, cropped and cleaned for background noise by manually picking ten unstained nuclei, subtracting the 99% quantile from the background distribution and rescaling. Cellular boundaries were determined by calculating the Voronoi diagram based on nuclei center with the deldir R package (Extended Data Fig. 4c). For each cell, channel intensity distribution was calculated, after subtracting the median values across channels. A cell was determined positive for a certain channel if its 20th highest pixel intensity was higher than 20. For cells positive for more than one channel, a cell was determined to express the channel with the highest total intensity, as well as coexpress any other channel displaying a >0.3 Pearson correlation to the leading channel (to exclude signals emanating from neighboring cells; Extended Data Fig. 4c).

Two cells were determined as neighbors if they shared a vertex in the nuclei Voronoi diagram. For the community analysis, all cells within a two-cell radius were considered a community.

Statistics. All experiments were performed in multiple distinct replicates (co-cultures or mice), as indicated in the text and figure legends. Detailed information about technical and biological replicates can be found in Extended Data Fig. 1 and Supplementary Table 2. All statistical tests were two-tailed. Statistical significance was computed in two principal ways: over pooled single-cell abundances (Fig. 2a,d and Extended Data Fig. 6e) using an FDR-adjusted Fisher's exact test. Cell subsets showing over-representation in PICs were further tested for reproducibility across patients or biological replicates using a Mann–Whitney U-test (Fig. 2b,e and Extended Data Fig. 6f). For tumor growth measurements, tumor size measurements were log-transformed and outliers were identified via the identify_outliers method of the rstatix package and removed. No other assumptions on data distribution were applied and all statistical tests were nonparametric. No statistical methods were used to predetermine sample sizes but our sample sizes are similar to those reported in previous publications[8,51]. All animal experiments were randomized before any experimental intervention.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

scRNA-seq data that support the findings of this study (including MARS-seq and PIC-seq of human NSCLC biopsies and mouse experiments), are deposited in the Gene Expression Omnibus under accession code GSE160903. Previously published scRNA-seq data and TCR-sequencing data that were reanalyzed here are available under accession codes GSE123139 (ref. [11]) and EGAD00001006608 (ref. [28]). Source data for all figures have been provided as Source Data files. All other materials and data supporting the findings of this study are available from the corresponding author on reasonable request. Source data are provided with this paper.

Code availability

The PIC-seq algorithm is available in the GitHub repository: https://github.com/aygoldberg/PIC-seq. All algorithms and auxiliary scripts used to analyze data and generate scripts are provided as supplementary software and will be deposited in the GitHub repository https://github.com/aygoldberg/NSCLS-PIC.

References

1. Chen, D. S. & Mellman, I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* **39**, 1–10 (2013).
2. Ferris, S. T. et al. cDC1 prime and are licensed by CD4(+) T cells to induce anti-tumour immunity. *Nature* **584**, 624–629 (2020).
3. Callahan, M. K., Postow, M. A. & Wolchok, J. D. Targeting T cell co-receptors for cancer therapy. *Immunity* **44**, 1069–1078 (2016).
4. Wherry, E. J. & Kurachi, M. Molecular and cellular insights into T cell exhaustion. *Nat. Rev. Immunol.* **15**, 486–499 (2015).
5. Topalian, S. L., Drake, C. G. & Pardoll, D. M. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell* **27**, 450–461 (2015).
6. Oh, S. A. et al. PD-L1 expression by dendritic cells is a key regulator of T-cell immunity in cancer. *Nature Cancer* **1**, 681–691 (2020).
7. Gajewski, T. F., Schreiber, H. & Fu, Y. X. Innate and adaptive immune cells in the tumor microenvironment. *Nat. Immunol.* **14**, 1014–1022 (2013).
8. Maier, B. et al. A conserved dendritic-cell regulatory program limits antitumour immunity. *Nature* **580**, 257–262 (2020).
9. Broz, M. L. et al. Dissecting the tumor myeloid compartment reveals rare activating antigen-presenting cells critical for T cell immunity. *Cancer Cell* **26**, 938 (2014).
10. Garris, C. S. et al. Successful anti-PD-1 cancer immunotherapy requires T cell-dendritic cell crosstalk involving the cytokines IFN-γ and IL-12. *Immunity* **49**, 1148–1161 (2018).
11. Li, H. et al. Dysfunctional CD8 T cells form a proliferative, dynamically regulated compartment within human melanoma. *Cell* **176**, 775–789 (2019).
12. Ledergor, G. et al. Single cell dissection of plasma cell heterogeneity in symptomatic and asymptomatic myeloma. *Nat. Med.* **24**, 1867–1876 (2018).
13. Jerby-Arnon, L. et al. A cancer cell program promotes T cell exclusion and resistance to checkpoint blockade. *Cell* **175**, 984–997 (2018).
14. Savas, P. et al. Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat. Med.* **24**, 986–993 (2018).
15. Filbin, M. G. et al. Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science* **360**, 331–335 (2018).
16. Katzenelenbogen, Y. et al. Coupled scRNA-seq and intracellular protein activity reveal an immunosuppressive role of TREM2 in cancer. *Cell* **182**, 872–885 (2020).
17. Zhang, L. et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell* **181**, 442–459 (2020).
18. Zilionis, R. et al. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* **50**, 1317–1334 (2019).
19. Zhang, L. et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature* **564**, 268–272 (2018).
20. Azizi, E. et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* **174**, 1293–1308 (2018).
21. Zheng, C. et al. Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. *Cell* **169**, 1342–1356 (2017).
22. Giladi, A. et al. Dissecting cellular crosstalk by sequencing physically interacting cells. *Nat. Biotechnol.* **38**, 629–637 (2020).
23. Baran, Y. et al. MetaCell: analysis of single-cell RNA-seq data using K-nn graph partitions. *Genome Biol.* **20**, 206 (2019).
24. Lavin, Y. et al. Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. *Cell* **169**, 750–765 (2017).
25. Guo, X. et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat. Med.* **24**, 978–985 (2018).
26. Leader, A. M. et al. CITEseq analysis of non-small-cell lung cancer lesions reveals an axis of immune cell activation associated with tumor antigen load and TP53 mutations. Preprint at *bioRxiv* https://doi.org/10.1101/2020.07.16.207605 (2020).
27. Keren-Shaul, H. et al. MARS-seq2.0: an experimental and analytical pipeline for indexed sorting combined with single-cell RNA sequencing. *Nat. Protoc.* **14**, 1841–1862 (2019).
28. Bassez, A. et al. A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. *Nat. Med.* **27**, 820–832 (2021).
29. Dieu-Nosjean, M. C. et al. Tertiary lymphoid structures, drivers of the anti-tumor responses in human cancers. *Immunol. Rev.* **271**, 260–275 (2016).
30. Dieu-Nosjean, M. C. et al. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *J. Clin. Oncol.* **26**, 4410–4417 (2008).
31. De Giovanni, M. et al. Spatiotemporal regulation of type I interferon expression determines the antiviral polarization of CD4(+) T cells. *Nat. Immunol.* **21**, 321–330 (2020).

32. Macian, F. NFAT proteins: key regulators of T-cell development and function. *Nat. Rev. Immunol.* **5**, 472–484 (2005).

33. Angiari, S. et al. Pharmacological activation of pyruvate kinase M2 inhibits CD4(+) T cell pathogenicity and suppresses autoimmunity. *Cell Metab.* **31**, 391–405 (2020).

34. Bacher, M. et al. An essential regulatory role for macrophage migration inhibitory factor in T-cell activation. *Proc. Natl Acad. Sci. USA* **93**, 7849–7854 (1996).

35. Gu-Trantien, C. et al. CXCL13-producing TFH cells link immune suppression and adaptive memory in human breast cancer. *JCI Insight* https://doi.org/10.1172/jci.insight.91487 (2017).

36. Gu-Trantien, C. et al. CD4(+) follicular helper T cell infiltration predicts breast cancer survival. *J. Clin. Invest.* **123**, 2873–2892 (2013).

37. Schurch, C. M. et al. Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *Cell* **182**, 1341–1359 (2020).

38. Amsen, D., van Gisbergen, K., Hombrink, P. & van Lier, R. A. W. Tissue-resident memory T cells at the center of immunity to solid tumors. *Nat. Immunol.* **19**, 538–546 (2018).

39. Schenkel, J. M. & Masopust, D. Tissue-resident memory T cells. *Immunity* **41**, 886–897 (2014).

40. Keskin, D. B. et al. Neoantigen vaccine generates intratumoral T cell responses in phase Ib glioblastoma trial. *Nature* **565**, 234–239 (2019).

41. Kamphorst, A. O. et al. Rescue of exhausted CD8 T cells by PD-1-targeted therapies is CD28-dependent. *Science* **355**, 1423–1427 (2017).

42. Hui, E. et al. T cell costimulatory receptor CD28 is a primary target for PD-1-mediated inhibition. *Science* **355**, 1428–1433 (2017).

43. Vickovic, S. et al. High-definition spatial transcriptomics for in situ tissue profiling. *Nat. Methods* **16**, 987–990 (2019).

44. Rodriques, S. G. et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).

45. Berglund, E. et al. Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nat. Commun.* **9**, 2419 (2018).

46. Oxenius, A., Bachmann, M. F., Zinkernagel, R. M. & Hengartner, H. Virus-specific MHC-class II-restricted TCR-transgenic mice: effects on humoral and cellular immune responses after viral infection. *Eur. J. Immunol.* **28**, 390–400 (1998).

47. Jaitin, D. A. et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014).

48. Sammicheli, S. et al. Inflammatory monocytes hinder antiviral B cell responses. *Sci. Immunol.* https://doi.org/10.1126/sciimmunol.aah6789 (2016).

49. Remark, R. et al. In-depth tissue profiling using multiplexed immunohistochemical consecutive staining on single slide. *Sci. Immunol.* **1**, aaf6925 (2016).

50. Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).

51. Zhou, T. et al. IL-18BP is a secreted immune checkpoint and barrier to IL-18 immunotherapy. *Nature* **583**, 609–614 (2020).

Acknowledgements

Author contributions

M.C. and I.A. conceived and designed the project, led all analyses and interpreted all experiments. A.G. conceived and designed the project, interpreted all experiments and performed computational analysis. M.C. and A.G. prepared the figures. M.C., A.G. and O.B. designed and performed experiments and analyzed the data. P.H., B.L., M.Z., A.G.-S., B.M., M.B., A.D. and J.L.B. performed experiments and analyzed the data. I.K. (under the supervision of J.B.) and C.G.B. (under the supervision of M.I.) performed experiments and analyzed data. M.C., A.G., A.T., M.M. and I.A. wrote the manuscript. A.T., M.M. and I.A. supervised the study.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s43018-022-00338-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43018-022-00338-5.

Correspondence and requests for materials should be addressed to Merav Cohen, Amos Tanay, Miriam Merad or Ido Amit.

Peer review information *Nature Cancer* thanks Sean Bendall and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
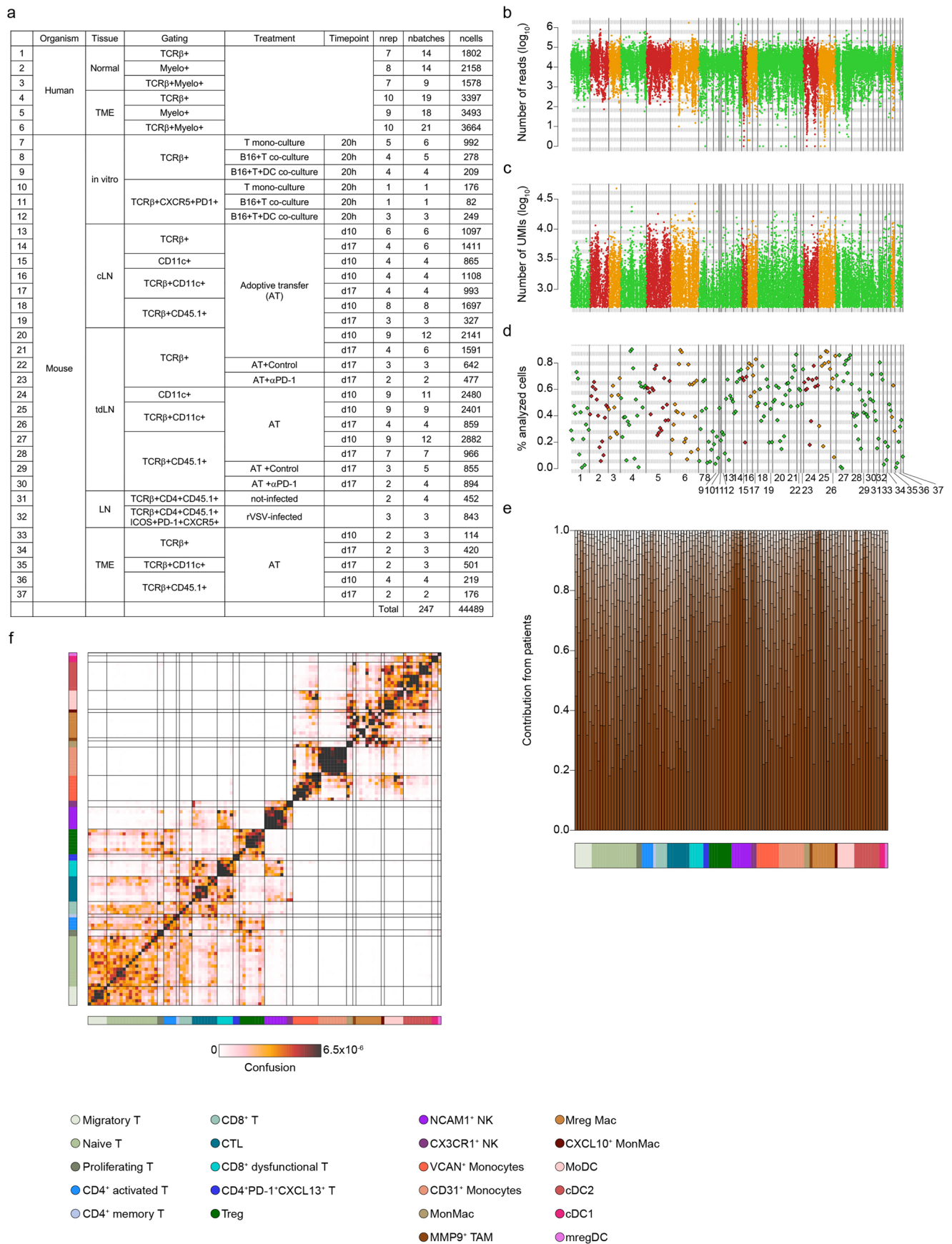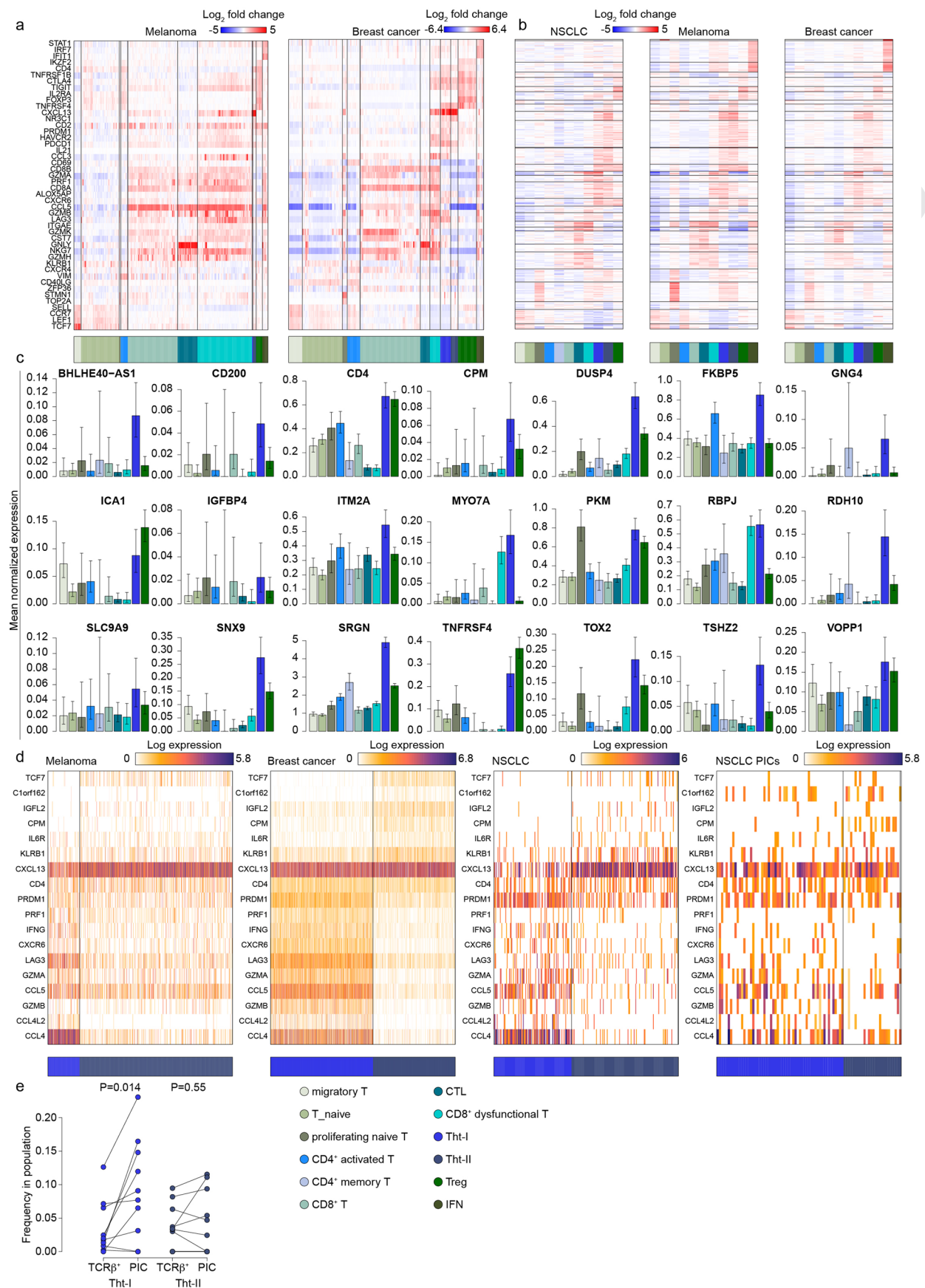
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Summary of samples and RNA-sequencing data. (a-d) Summary of all experimental samples, plates and cells, processed by MARS-seq and PIC-seq. "nrep" indicates number of biological replicates (patients, co-cultures or mice), "nbatches" indicates number of technical replicates, "ncells" indicates number of analyzed cells (Supplementary Table 2). Shown are (b) total number of Illumina reads, (c) total number of Unique Molecular Identifiers (UMIs) per cell, and (d) fraction of QC-positive cells retained for further analysis per technical replicate. (**e**) Patient contribution to each of the 112 metacells derived from TCRβ+ T cells and CD11c+CD64+ myeloid cells. Patients are ordered from top (brown) to least (white) contribution per metacell. (**f**) The confusion matrix of the MetaCell model shown in Fig. 1c. Entries denote for each pair of metacells the propensity of cells from both metacells to be clustered together in a bootstrap analysis. Bottom and left panels indicate metacell annotation to 22 T and myeloid subtypes.
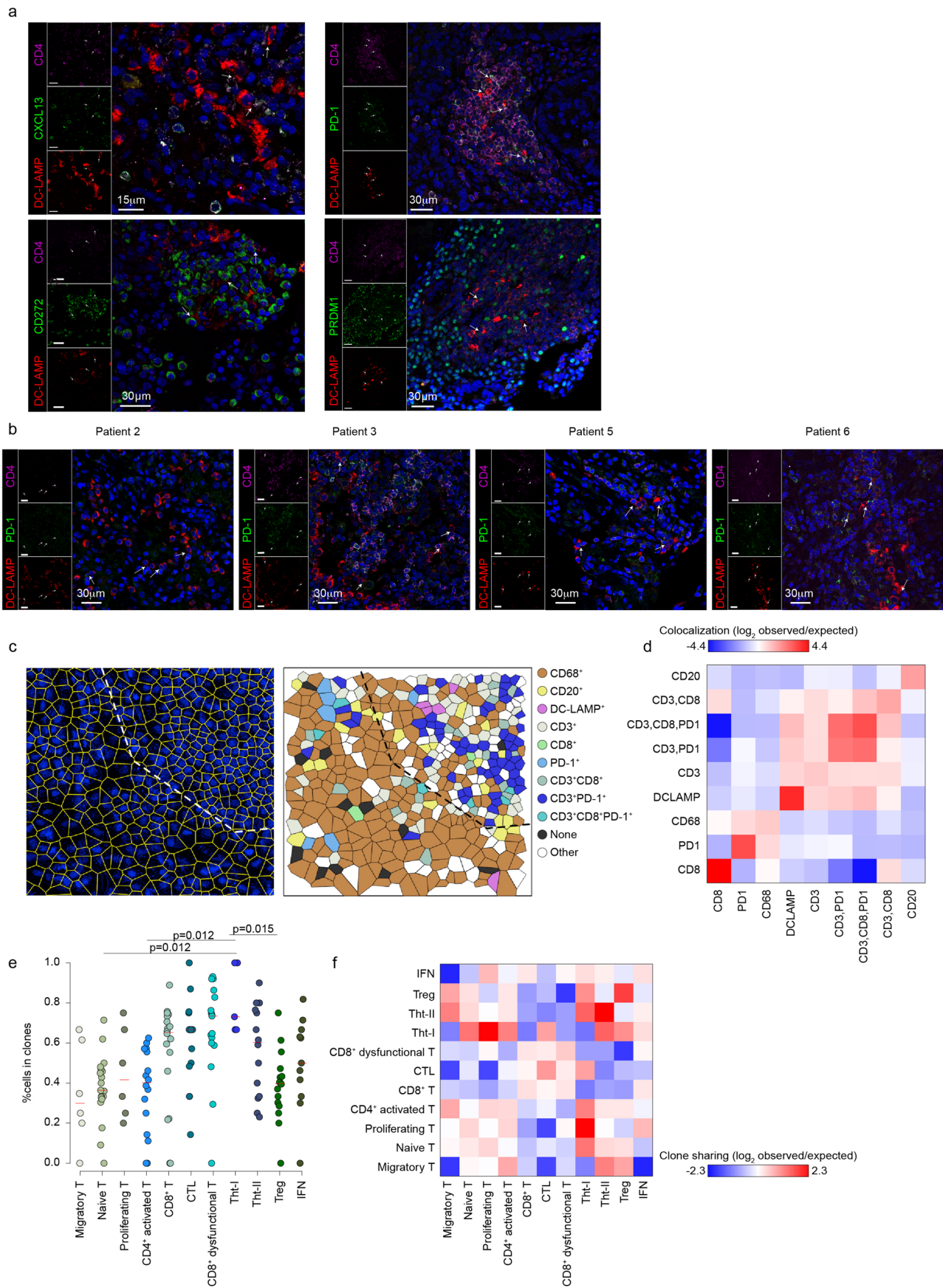
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | NSCLC PIC-seq quality controls. (**a**) Performance of the linear regression model estimating the mixing factor (α) of synthetic T-myeloid (left) and T-NK (right) PICs. (**b**) Performance of the T (left) and myeloid (right) metacell assignments of PIC-seq over 5,000 synthetic PICs. Each row summarizes all synthetic PICs originating from one metacell and their assignments to all metacells (columns). Data is row-normalized. (**c**) The cumulative distribution of the $ll_{doublets}$-$ll_{singlets}$ score, for PICs (orange) and for the T (green) and myeloid (red) singlet populations. The score indicates the gain in likelihood when each PIC is modeled as a doublet compared to its most likely singlet assignment (Methods). PICs whose scores were not positive were suspected as singlets and discarded from further analysis. (**d**) Fraction of retained PICs for each profiled patient (Supplementary Table 1). Numbers on top indicate absolute number of retained PICs per patient. (**e**) Flow cytometry dot-plots analysis (top) and quantification (bottom) of *in situ* and *in vitro* PICs formed before or after tissue dissociation. Values in brackets indicate the estimated spurious PIC frequency out of the PIC population, defined as PICs combining fluorophores from parallel samples. (**f**) Gene expression profiles of 10,762 single cells grouped into 22 transcriptional subsets. Top panel indicates whether a cell is derived from healthy tissue or TME. (**g**) Myeloid and T-cell subset identities of single cells in (f) (Fig. 1c). (**h**) Gene expression profiles of 839 QC-positive PICs, grouped by their contributing T-cell and myeloid identities, as determined by PIC-seq algorithm. Top panel as in (f). (**i**) Myeloid and T cell subset identity of PIC contributing cells in (g), as determined by PIC-seq algorithm. (**j**) Estimation of the relative UMI count from T cells (green), and myeloid (red) contributing to each PIC in (h), as inferred by PIC-seq algorithm (mixing factor, α).

Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Characterization of singlet and PIC-derived CD4⁺PD-1⁺CXCL13⁺ T cells in human TME. (a) Gene expression profiles of T cells from melanoma (left) and breast cancer (right). 36,341 cells from 21 melanoma patients grouped into 245 metacells, and 62,909 cells from 42 breast cancer patients grouped into 243 metacells are shown. Values indicate enrichment (log2 fold change) of a gene in a metacell over its median value across metacells. Annotation to T subsets is indicated below. (b) Joint K-mean clustering (K = 38) of 763 genes differentially expressed in CD4⁺PD-1⁺CXCL13⁺ cells across different T subsets derived from NSCLC, melanoma and breast tumors (Supplementary Table 3). (c) Mean normalized expression of genes upregulated by CD4⁺PD-1⁺CXCL13⁺ T derived from NSCLC TME across all T cell states. Error bars indicate binomial 95% confidence intervals of the estimated mean. n = 3371 TME T cells. (d) Gene expression profiles of single Tht cells divided into the Tht-I and Tht-II subsets. Shown are Tht cells from melanoma, breast cancer, NSCLC and NSCLC PICs assigned to the Tht identity. Tht PICs were further dissected to Tht-I and Tht-II using breast cancer Tht metacells as the T cell reference model (Methods). (e) Enrichment of Tht-I and Tht-II in the NSCLC PIC population compared to their frequency in the single cell TCRβ⁺ population. Two-tailed paired Mann–Whitney test; n = 10 NSCLC patients.
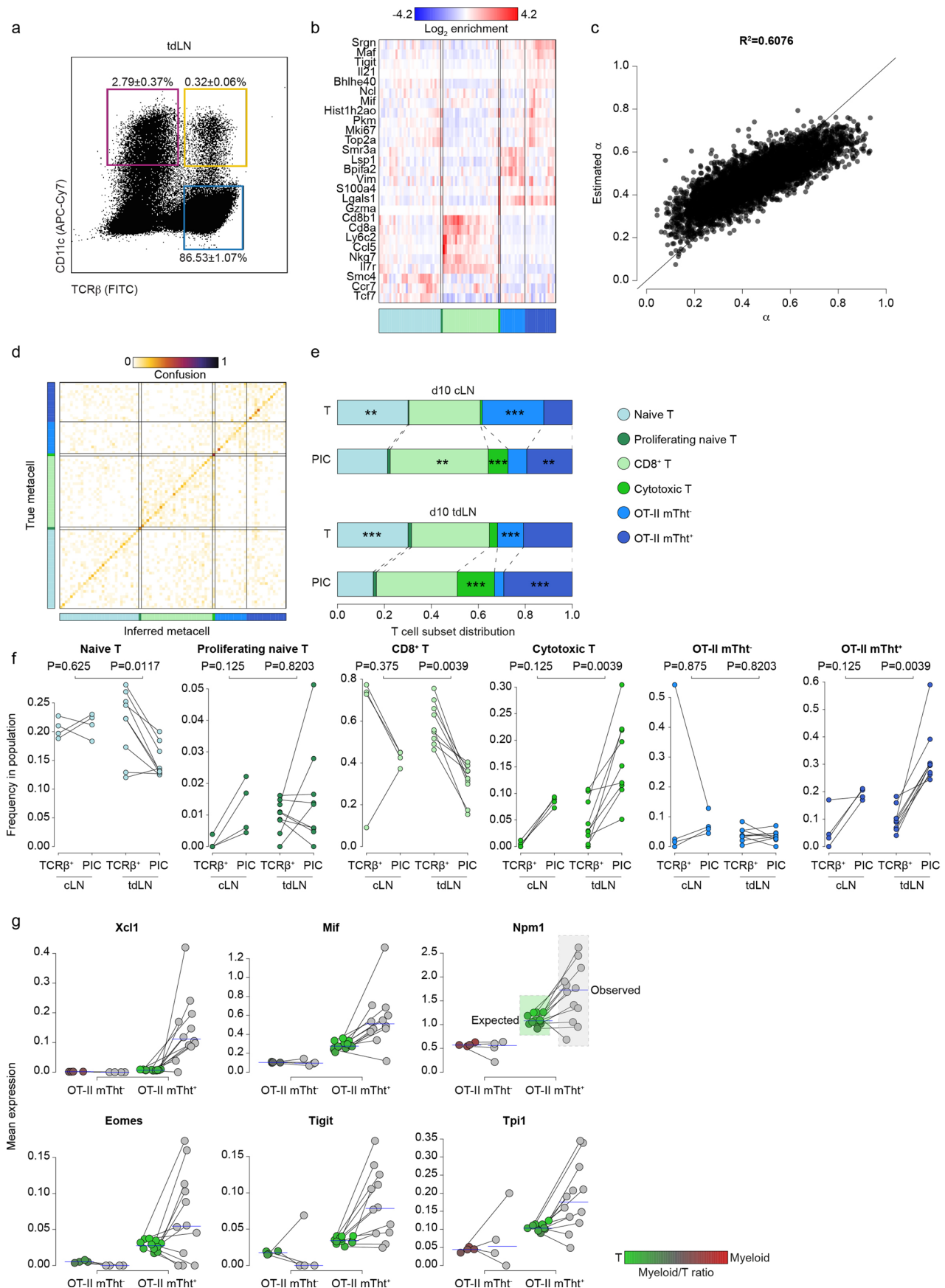
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Spatial and clonal properties of Tht cells. (a) Representative confocal microscopy images of tumor sections derived from one NSCLC patient stained for CD4, DC-LAMP and CD272 (BTLA) or PRDM1 proteins. **(b)** Representative confocal microscopy images of tumor sections derived from four additional NSCLC patients stained for CD4, DC-LAMP and PD-1 proteins. (a-b) Scale bar=30µm, and 15 µm in (a) top left panel; arrows indicate T-DC conjugates; images are representative of seven scanned patients. **(c)** Left – nuclear segmentation of the image depicted in Fig. 3e. Yellow markings outline the Voronoi diagram, enclosing all pixels sharing a nearest nucleus. Right – Cell type annotation of each segmented nuclei by expression and co-expression of marker intensity in the area of each nucleus' Voronoi structure (Methods). **(d)** Colocalization analysis. For each pair of cell types, we counted the number of occurrences the two cell types co-exist in a 2-cell radius community. Colors indicate $\log_2$ enrichment over the expected values. n=5 TLS from the same patient. **(e)** Fraction of cells related to T cell clones across different melanoma T cell subsets (Extended Data Fig. 3a) in melanoma patients for whom sufficient TCR-seq data was available. A cell is considered part of a clone if it shares a TCR sequence with at least one other T cell from the same patient. FDR-adjusted two-tailed unpaired Mann–Whitney test comparing Tht-I and II to other T subsets. n=10 melanoma patients. **(f)** A heat map depicting the propensity of two cells from two melanoma T subsets to belong to the same clone (clone sharing). Data was calculated by sampling 10,000 pairs of cells, and comparing clonal sharing characteristics to 10,000 pairs of cells sampled after shuffling clone identities, while preserving the number of clones and clone sizes per patient (Methods). *$P < 0.05$, **$P < 0.001$, ***$P < 10^{-5}$.
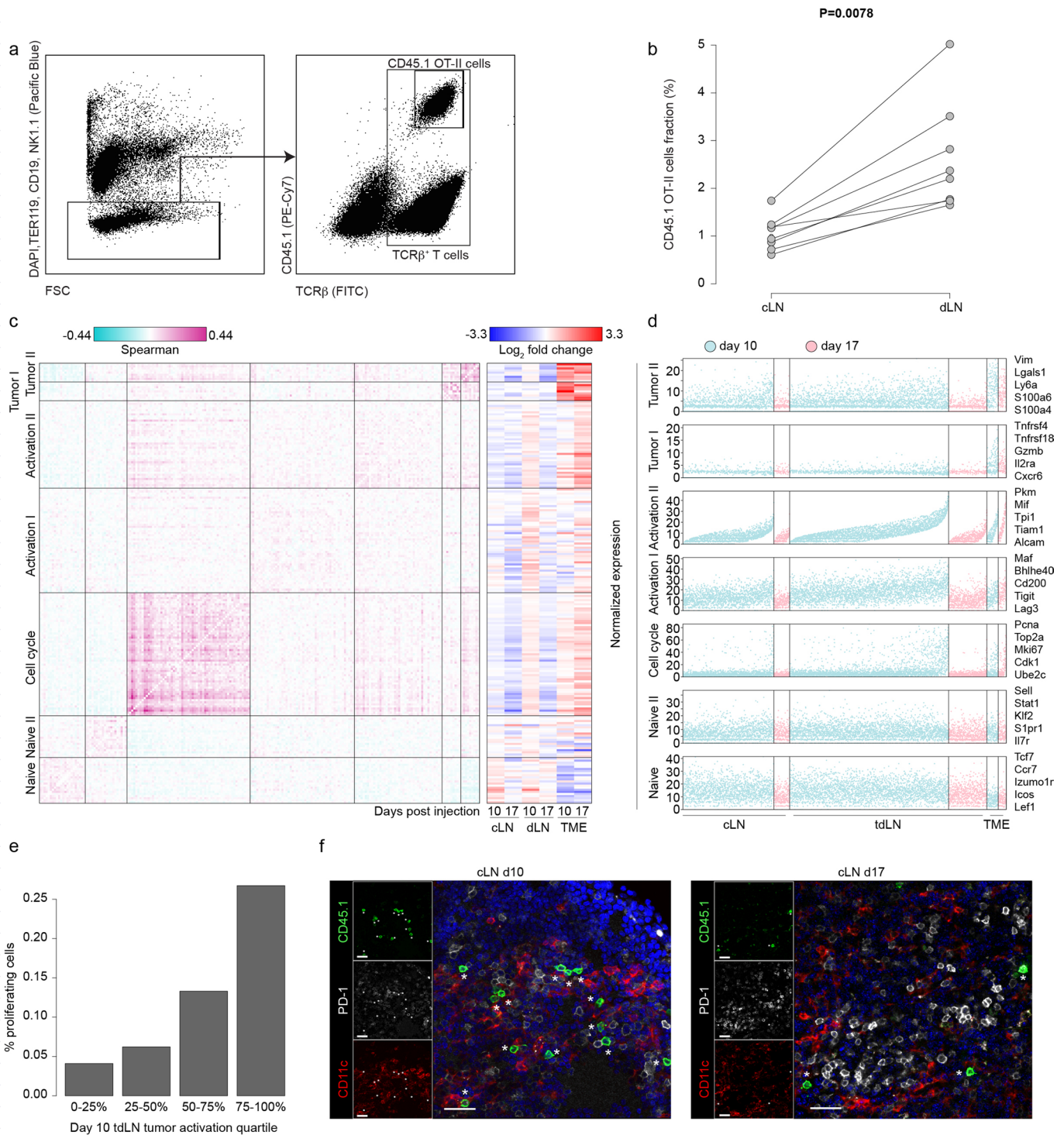
Extended Data Fig. 5 | Dynamic and spatial characterization of OT-II CD45.1+ T cells in murine TME niche and cLN. (**a**) Representative FACS plot showing gating strategy for isolation of adoptively transferred CD45.1+ OT-II cells and bystander polyclonal TCRβ+ T cells from tdLN. (**b**) Quantification of the percentage of CD45.1+ OT-II T cell out of the entire TCRβ+ population in matched cLN and tdLN tissues. Two-tailed paired Mann–Whitney test; n = 8 mice from two independent experiments (**c**) Gene module analysis. Shown are seven correlated gene modules derived from CD45.1+TCRβ+ OT-II T cells isolated from cLN, dLN and TME, 10- and 17 days post tumor cell injection. Left – pairwise Pearson correlation. Right – normalized pooled expression across different conditions. (**d**) Pooled expression of the seven gene modules from (c) across OT-II T cells derived from all conditions. Dots represent single cells and dot colors indicate time points. Representative genes from each module are depicted. (**e**) Fraction of proliferating cells in each quartile of the mTht activation signature as in Fig. 5b in dLN OT-II T cells. A cell was determined proliferating if it exhibited above-threshold expression of the cell-cycle module in (d). (**f**) Representative confocal microscopy images of cLN sections extracted 10- and 17 days following tumor cell injection, stained for CD45.1+ (OT-II) T cells, PD-1 and CD11c proteins. Scale bar=30μm; asterisks indicate OT-II mTht; images are representative of two independent experiments.

Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | PIC-seq in mouse model of tumor antigen specificity. (**a**) A representative FACS plot of CD11c$^+$ myeloid (purple), TCRβ$^+$ T (blue) singlets, and CD11c$^+$TCRβ$^+$ PICs (orange) purified from the tdLN (n = 6 independent experiments); Population frequencies represent mean ± s.e.m. (**b**) Gene expression of 7541 tdLN- and cLN-derived TCRβ$^+$ and CD45.1$^+$TCRβ$^+$ T cells isolated 10 days following tumor injection, grouped into 86 metacells. Bottom panel indicates cell annotations. (**c**) Performance of the linear regression model used to estimate the mixing factor (α) of synthetic T-myeloid PICs. (**d**) Performance of the T metacell assignments of PIC-seq over 5,000 synthetic PIC. Each row summarizes all synthetic PICs originating from one metacell and their assignments to metacells by the PIC-seq algorithm (columns). Data is row-normalized. (**e**) Distribution of T subsets in TCRβ$^+$ singlet T and CD11c$^+$TCRβ$^+$ PICs in cLN and tdLN 10 days following tumor cells injection. Cells are downsampled so that T and PIC numbers are equal per replicate and then pooled from all profiled patients. FDR-adjusted Two-tailed Fisher's exact test. (**f**) Comparison of different T subset frequencies in TCRβ$^+$ singlets and PICs isolated from cLN and tdLN 10 days following tumor cells injection, across all biological replicates. Two-tailed paired Mann–Whitney test. (**g**) Mean observed (gray) and expected (colored) gene expression levels in PICs of OT-II mTht$^-$ and mTht$^+$ subsets. Each connected pair of dots signifies a biological replicate; Dot colors relate to their specificity in the T (green) or myeloid (red) cell expected contributions. Groups with less than 10 cells were discarded. Median value is marked for each category. Data summarizes two independent experiments; (e-g) n = 4 day 10 cLN and 9 day 10 tdLN. *P < 0.05, **P < 0.001, ***P < 10^{-5}.
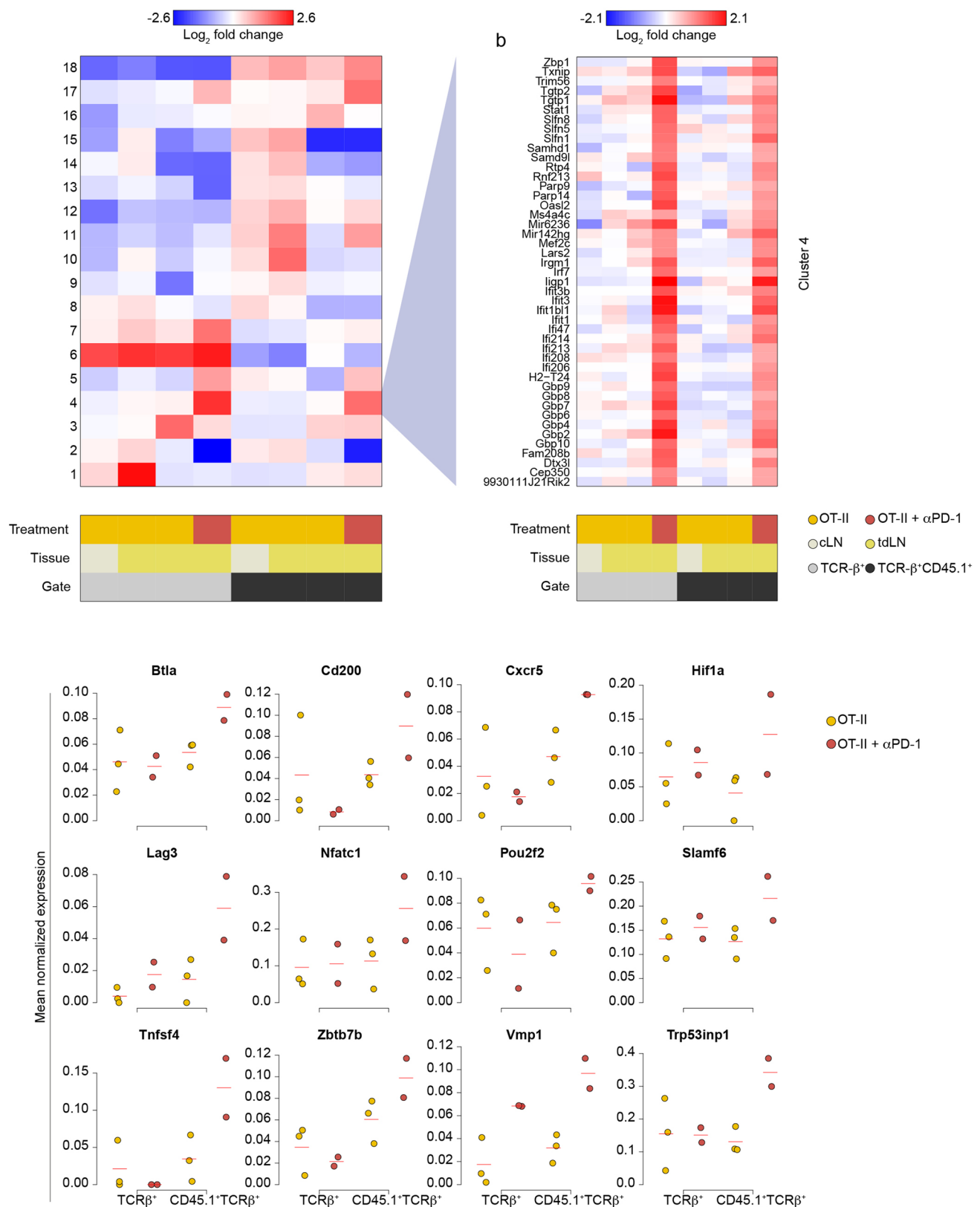
Extended Data Fig. 7 | Tht cells are involved in the anti-tumor effect of aPD-1 treatment. (a-b) K-means analysis of TCRβ[+] and CD45.1[+] TCRβ[+] OT-II cells derived from tdLN of mice with/without αPD-1 treatment, at day 17 following tumor injection. Day 10 tdLN and cLN T cells from Fig. 5b were included for comparison. Shown are cluster centers (a) and all genes from cluster 4 (enriched for type I Interferon response genes; (b)). Values represent $\log_2$ fold change over the median. **(c)** Mean normalized expression of key mTht-related genes from cluster 11 across the different samples. *$P < 0.05$, **$P < 0.001$, ***$P < 10^{-5}$.

QUERY FORM

AUTHOR:

The following queries have arisen during the editing of your manuscript. Please answer by making the requisite corrections directly in the e.proofing tool rather than marking them up on the PDF. This will ensure that your corrections are incorporated accurately and that your paper is published as quickly as possible.

Query No.	Nature of Query
Q1:	Please check your article carefully, coordinate with any co-authors and enter all final edits clearly in the eproof, remembering to save frequently. Once corrections are submitted, we cannot routinely make further changes to the article.
Q2:	Note that the eproof should be amended in only one browser window at any one time; otherwise changes will be overwritten.
Q3:	Author surnames have been highlighted. Please check these carefully and adjust if the first name or surname is marked up incorrectly. Note that changes here will affect indexing of your article in public repositories such as PubMed. Also, carefully check the spelling and numbering of all author names and affiliations, and the corresponding email address(es).
Q4:	You cannot alter accepted Supplementary Information files except for critical changes to scientific content. If you do resupply any files, please also provide a brief (but complete) list of changes. If these are not considered scientific changes, any altered Supplementary files will not be used, only the originally accepted version will be published.
Q5:	If applicable, please ensure that any accession codes and datasets whose DOIs or other identifiers are mentioned in the paper are scheduled for public release as soon as possible, we recommend within a few days of submitting your proof, and update the database record with publication details from this article once available.
Q6:	Your paper has been copy edited. Please review every sentence to ensure that it conveys your intended meaning; if changes are required, please provide further clarification rather than reverting to the original text. Please note that formatting (including hyphenation, Latin words, and any reference citations that might be mistaken for exponents) has been made consistent with our house style.
Q7:	Please ensure that genes are correctly distinguished from gene products: for genes, official gene symbols (e.g., NCBI Gene) for the relevant species should be used and italicized; gene products such as proteins and noncoding RNAs should not be italicized.
Q8:	Please note, we reserve 'significant' and its derivatives for statistical significance. Please reword where this is not the intended meaning (for example to important, notable, substantial); there are 3 instances throughout your text
Q9:	In the sentence beginning 'To better understand...' please check whether 'C57/6J' should be 'C57BL/6J'.
Q10:	In the sentence beginning 'To directly test the function of mTht cells..' please check whether 'C57/6J should be C57BL/6J.

QUERY FORM

	Nature Cancer
Manuscript ID	**[Art. Id: 338]**
Author	**Merav Cohen**

AUTHOR:

The following queries have arisen during the editing of your manuscript. Please answer by making the requisite corrections directly in the e.proofing tool rather than marking them up on the PDF. This will ensure that your corrections are incorporated accurately and that your paper is published as quickly as possible.

Query No.	*Nature of Query*
Q11:	In the sentence beginning 'To derive genes for computing the ...' please confirm that inserted definition of MLE is correct.
Q12:	Please check that all funders have been appropriately acknowledged and that all grant numbers are correct.
Q13:	If applicable, please ensure accession codes are scheduled for release on or before this article's scheduled publication date, and update the database record with publication details from this article once available.
Q14:	Please check that the Competing Interests declaration is correct as stated. If you declare competing interests, please check the full text of the declaration for accuracy and completeness.
Q15:	If ref. 26 (preprint) has now been published in final peer-reviewed form, please update the reference details if appropriate.

nature research

Corresponding author(s): Prof. Ido Amit

Last updated by author(s): Jan 16, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No open-source or custom code was used to collect data for this paper
Data analysis	For FACS analysis, we used the following sofrware: FACSDiva 7 FlowJo 10.4.2 All data analysis was performed in R (version 3.5.3). Image analysis was performed in CellProfiler (version 4.2.0). Data analysis was done with the published MetaCell package (version 0.3.41). PIC-seq algorithm can be found in the Github repository: https://github.com/aygoldberg/PIC-seq. List of R packages used in analysis and their version numbers: Package version glmnet 2.0-16 foreach 1.4.4 Matrix 1.2-18 compositions 1.40-2 bayesm 3.1-0.1 energy 1.7-5 robustbase 0.93-3 tensorA 0.36.1 gplots 3.0.1.1 plotrix 3.7-4 plyr 1.8.4

```
RANN    2.6.1
reshape2      1.4.3
KernSmooth    2.23-15
dendextend    1.9.0
Hmisc   4.2-0
Formula 1.2-3
survival      3.2-3
lattice 0.20-38
tglkmeans     0.2.0
ggrepel 0.8.1
ggplot2 3.3.2.9000
ape     5.2
scales  1.0.0
metacell      0.3.41
tgstat  2.3.5
misha   4.0.10
```

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Single cell RNA-seq data that support the findings of this study (including MARS-seq and PIC-seq of human NSCLC biopsies and mouse experiments), was deposited in the Gene Expression Omnibus (GEO) under accession code GSE160903. Previously published scRNA-seq data and TCR-sequencing data that were re-analysed here are available under accession code GSE123139 (melanoma REF) and EGAD00001006608 (breast REF).
Source data for all figures have been provided as Source Data files. All other materials and data supporting the findings of this study are available from the corresponding author on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications. All experiments were performed in multiple distinct replicates (co-cultures or mice), as indicated in the text and figure legends. For in vivo tumor growth experiments, sample size was chosen as in previous studies, and variation in tumor dimensions within the group was considered.
Data exclusions	To eliminate contamination of the PIC gate by single cells (singlets), putative PICs from human samples whose likelihoods when modeled as doublets were not greater than when modeled as singlets were determined as singlets and discarded from downstream analysis (See Methods, section "PIC-seq analysis of human lymphocytes"). For tumor growth measurements, tumor size measurements were log-transformed, and outliers were identified via the identify_outliers method of the rstatix package and removed.
Replication	In vivo tumor injection experiment was repeated twice, with the following biological replicates: n=6 day 9 tdLN, 7 day 17 tdLN, 8 day 10 cLN, and 3 day 17 cLN, 4 day 10 tumor and 2 day 17 tumor. Anti-PD1 treatment in vivo experiment was performed twice, assessed by tumor volume measurements, and yielded similar results. All replications were successful, showing similar results.
Randomization	All animal experiments were randomized before any experimental intervention.
Blinding	Tumor size was measured blindly to the conditions of the experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

eFluor450-conjugated TER-119, eFluor450-conjugated NK1.1, PerCP Cy5.5-conjugated streptavidin and Pacific blue-conjugated CD19 were purchased from eBioscience, PerCP Cy5.5-conjugated TCRβ, FITC-conjugated TCRβ, APC-Cy7-conjugated CD11c, APC-conjugated CD45.2, PE-Cy7-conjugated CD45.1, APC-conjugated CD279 (PD-1) and biotin-conjugated CXCR5 were purchased from Biolegend. Human samples were stained using the following anti-human antibodies: pacific blue-conjugated CD235a, FITC-conjugated TCRβ, PE-conjugated CD11c and PE-conjugated CD64 were purchased from Biolegend, APC-conjugated CD45, was purchased from eBioscience and pacific blue-conjugated CD19 and PE-Cy7-conjugated CD56 were purchased from BD Biosciences. rat DC-LAMP (1:200, Novusbio), goat CXCL13 (1:50, R&D Systems), rabbit CD4 (1:100, Novusbio), mouse CD272 (BTLA, 1:100, Abcam), rabbit PD-1 (1:100, Abcam), rabbit PRDM1/Blimp1 (1:100, Abcam). APC-conjugated CD279 (PD-1) (1:100; Biolegend), Alexa Fluor 488-conjugated CD45.1 (1:100; Biolegend) and PE-conjugated CD11c (1:100; Biolegend).

Validation

We used mouseFC-block / human FC-block inFACS experiments, and blocking solution in immuno-staining experiments. FACS antibodies were compared to unstained sample. Only FACS antibodies that yielded high signal and separation in the flow cytometer (fold change > 10) were used for sorting. Moreover, in the analysis we validated that antibodies used in the study were specific for unique cell types, and didn't exhibit any non-specific binding.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

B16 melanoma cells (H2b), stably expressing chicken ovalbumin (B16-OVA) and mCherry fluorescent labeling.

Authentication

By analyzing the single cell RNA-seq data, we authenticated the gene expression of the cancer cells.

Mycoplasma contamination

Cell line was negative for Mycoplasma.

Commonly misidentified lines
(See ICLAC register)

We used only identified cell line in the study.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

C57BL/6 WT 8-10 week old female mice and 8-10 week old male TCR-transgenic OT-II mice (harboring OVA-specific CD4+ T cells) were used in the study.

Wild animals

No wild animals were used in the study.

Field-collected samples

No field-collected samples were used in the study.

Ethics oversight

Mouse experiment protocol was approved by the Weizmann Institutional Animal Care and Use Committee (03150320-1 and 00580121-2), and by the Institutional Animal Committee of the San Raffaele Scientific Institute (670).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics

Samples of tumor and adjacent normal lung tissues were obtained from surgical specimens of non-small cell lung carcinoma (NSCLC) patients undergoing resection at the Mount Sinai Medical Center (New York, NY). All relevant metadata, including covariates, are reported in Supplementary Table 1.

Recruitment

Participants were recruited by obtaining informed consent in accordance with a protocol reviewed and approved by the Institutional Review Board at the Icahn School of Medicine at Mount Sinai (IRB Human Subjects Electronic Research Applications 10-00472 and 10-00135) and in collaboration with the Biorepository and Department of Pathology and the

Weizmann Institute of Science. Allocation of participants to the study was based on timing and biopsy avilability alone, and we are not aware of any selection bias regarding sample obtainment.

Ethics oversight	Icahn School of Medicine at Mount Sinai (IRB Human Subjects Electronic Research Applications 10-00472 and 10-00135) and in collaboration with the Biorepository and Department of Pathology and the Weizmann Institute of Science.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	To achieve single-cell suspension in the human NSCLC specimens, 0.1-0.4gr of tumor and adjacent normal tissues were cut into small pieces and then mechanically dissociated by pipetting. Tissues were suspended with CO2 Independent Medium (Thermo Fisher SCIENTIFIC) supplemented with DNase (100μg/ml, Sigma-Adrich) and collagenase IV (0.5mg/ml, Worthington), and incubated at 37oC for 20min, with frequent agitation. To achieve single-cell suspensions in the murine model, tumors were cut into small pieces, and suspended with RPMI- 1640 supplemented with DNase (28μg/ml, Sigma-Adrich) and collagenase IV (1mg/ml, Worthington). Tissues were homogenized by GentleMacs tissue homogenizer (Miltenyi Biotec), and incubated at 37oC for 10min, with frequent agitation. This tissue dissociation procedure was performed twice, for each tumor. To achieve single-cell suspensions from tdLN and cLN, tissues were digested in Isocoves modified Dulbeccos medium (IMDM; Sigma-Aldrich) medium. For mild dissociation, tissues were supplemented with Liberase-TL (100 μg/ml, Roche) and DNase I (100 μg/ml, Roche), and incubated with frequent agitation at 37oC for 20 min. After all dissociation procedures, cells were washed with cold PBS, filtered through a 100μm cell strainer, and centrifuged at 380g for 5 min at 4oC.
Instrument	Cell populations were sorted using either SORP-aria (BD Biosciences) or ARIA-III instrument (BD Biosciences), and analyzed using BD FACSDIVA software (BD Biosciences) and FlowJo software.
Software	Cell populations were sorted using either SORP-aria (BD Biosciences) or ARIA-III instrument (BD Biosciences), and analyzed using BD FACSDIVA software (BD Biosciences) and FlowJo software.
Cell population abundance	All sorted populations were processed and analyzed with scRNA-seq methods (MARS-seq and PIC-seq) and their cell abundances were calculated and reported in Figures 2a,d and 6c.
Gating strategy	human: FSC/SSC -> FSC/CD235a-CD19-CD56- -> TCRb/CD11cCD64 mouse: FSC/SSC -> FSC/TER119-CD19-NK1.- -> TCRb/CD11cCD64 mouse: FSC/SSC -> FSC/TER119-CD19-NK1.- -> TCRb/CD45.1

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.