











Tumor editing suppresses innate and adaptive antitumor immunity and is reversed by inhibiting DNA methylation

Received: 6 July 2023

Accepted: 13 July 2024

Published online: 21 August 2024

 Check for updates

Ying Zhang ^{1,2,3,4,13} ✉, Pourya Naderi Yeganeh^{5,6,13}, Haiwei Zhang^{1,2}, Simon Yuan Wang ^{2,7}, Zhouyihan Li^{1,2,11}, Bowen Gu ^{1,2}, Dian-Jang Lee ^{1,2}, Zhibin Zhang ^{1,2,11}, Athanasios Ploumakis^{6,8}, Ming Shi ^{9,10}, Hao Wu ^{1,9}, Eric Lieberman Greer ^{2,7,12}, Winston Hide ^{5,6} ✉ & Judy Lieberman ^{1,2} ✉

Cancer cells edit gene expression to evade immunosurveillance. However, genome-wide studies of gene editing during early tumorigenesis are lacking. Here we used single-cell RNA sequencing in a breast cancer genetically engineered mouse model (GEMM) to identify edited genes without bias. Late tumors repressed antitumor immunity genes, reducing infiltrating immune cells and tumor–immune cell communications. Innate immune genes, especially interferon-stimulated genes, dominated the list of downregulated tumor genes, while genes that regulate cell-intrinsic malignancy were mostly unedited. Naive and activated CD8⁺ T cells in early tumors were replaced with exhausted or precursor-exhausted cells in late tumors. Repression of immune genes was reversed by inhibiting DNA methylation using low-dose decitabine, which suppressed tumor growth and restored immune control, increasing the number, functionality and memory of tumor-infiltrating lymphocytes and reducing the number of myeloid suppressor cells. Decitabine induced important interferon, pyroptosis and necroptosis genes, inflammatory cell death and immune control in GEMM and implanted breast and melanoma tumors.

The clinical successes of immunotherapy clearly show that immune cells can control tumor growth. However, most solid cancers do not respond to current immunotherapy. Although immunosurveillance constrains and eliminates some newly formed immunogenic tumors, the tumors that grow to cause disease quickly evolve to resist immunosurveillance and evade immune control^{1,2}. Landmark papers by Schreiber and colleagues^{1,2} postulated that tumors edit their gene expression to avoid immune recognition and elimination. They defined immunoediting by its functional consequences: reducing immune recognition and control. Editing could occur by gene mutation, amplification, deletion and/or epigenetic modification of gene expression. The immune evasive tumor clones could also mold the tumor microenvironment (TME) to dampen cytotoxic lymphocyte function, further promoting immune evasion¹.

Previous immunoediting studies mostly examined genes and/or pathways that activate antitumor T cells. Downregulation of antigen processing and presentation, upregulation of checkpoint receptor ligands, immunosuppressive cytokines and ‘don’t eat me’ signals, poor recruitment and exclusion of immune cells from the TME and an immunosuppressive TME are known strategies of tumor immune evasion^{2–5}. Tumor antigens are both mutated and epigenetically silenced in developing tumors from immunocompetent, but not immunodeficient, mice, suggesting immune selection leverages both genetic and epigenetic mechanisms to reduce tumor immunoreactivity^{6,7}.

Stimulation of T cell effector function and memory depends on danger signals that activate innate immunity to recruit immune cells to sites of danger and allow immune cells to distinguish self from nonself,

A full list of affiliations appears at the end of the paper. ✉ e-mail: ying.zhang@pku.edu.cn; whide@bidmc.harvard.edu; judy.lieberman@childrens.harvard.edu

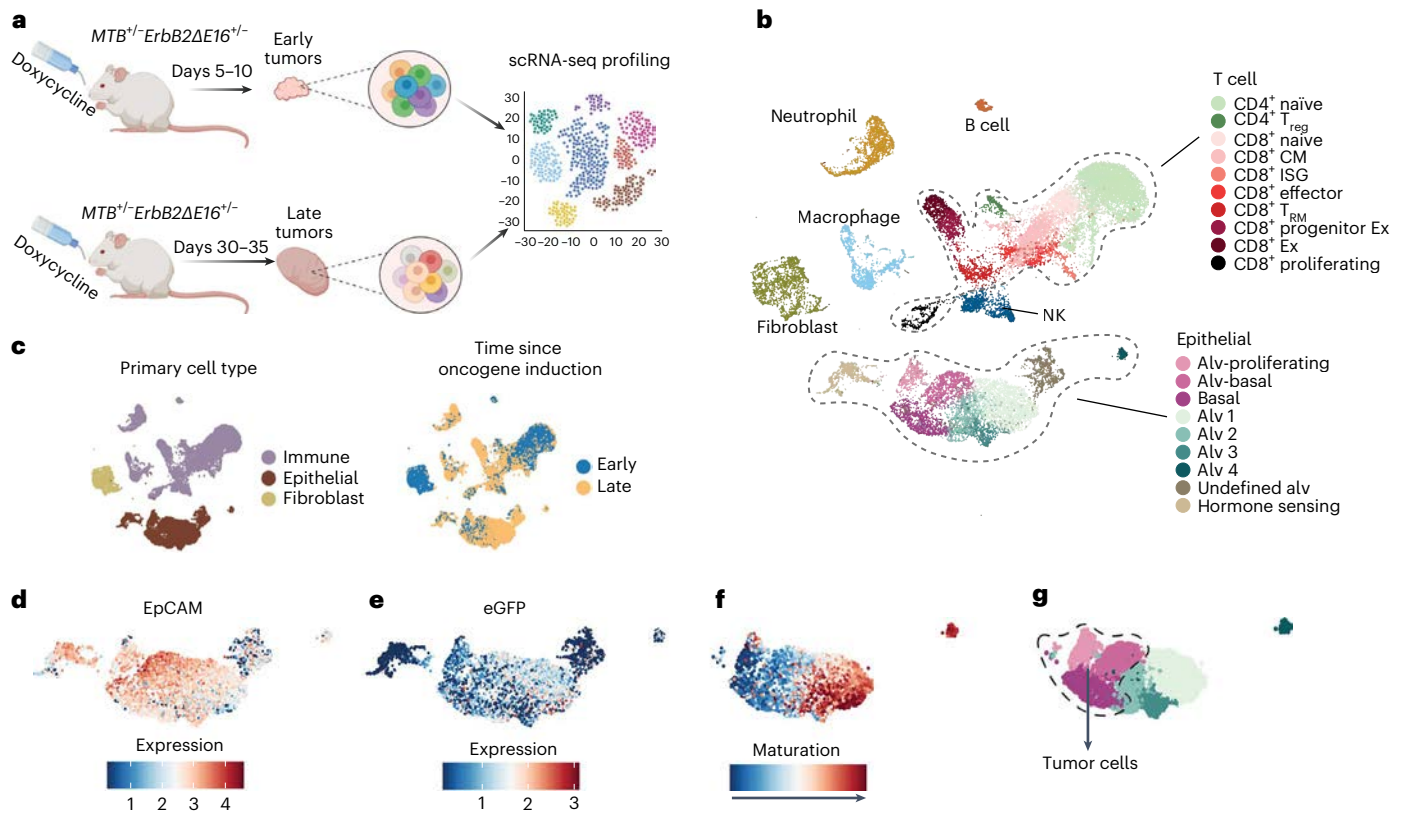


Fig. 1 | Cell mapping by scRNA-seq of early and late tumors in a breast cancer GEMM. **a**, Experimental set up: 8-week-old female *ErbB2ΔE16* transgenic mice were given doxycycline in the drinking water to induce breast tumors, which were harvested at early (days 5–10, $n = 4$) or late (days 30–35, $n = 4$) time points and dissociated into single cells, which were analyzed for scRNA-seq. **b**, Uniform Manifold Approximation and Projection (UMAP) plot of 24,798 cells and 20,499 genes in early and late GEMM libraries: 23 primary cell subtypes were identified using unsupervised clustering algorithms and manual inspection of known cell markers. **c**, UMAP plots showing primary cell types (left) or distribution of

cells in early and late tumors (right). Primary cell types were EpCAM⁺ epithelial cells (early tumor $n = 927$, late tumor $n = 6,582$), CD45⁺ immune cells (early tumor $n = 7,152$, late tumor $n = 8,539$) and FAP⁺/DCN⁺ fibroblasts (early tumor $n = 1,546$, late tumor $n = 52$). **d–g**, Epithelial clusters were classified as tumor and nontumor (**g**) based on expression of EpCAM (**d**) and eGFP (**e**) and maturation status of EpCAM⁺eGFP⁺ cells (**f**) as determined by CytoTRACE. Subtype markers are shown in Extended Data Figs. 1 and 2. **a**, Created using [BioRender.com](https://www.biorender.com). UMAP coordinates and cell metadata are provided in Source Data Fig. 1. CM, central memory; Ex, exhausted; Alv, Alveolar.

or for tumors, normal from transformed cells⁸. Without a danger signal, lymphocytes quickly become tolerant of the tumor. DNA promoter hypermethylation to suppress innate immune interferon (IFN), necroptosis and pyroptosis pathways that signal danger has been described in some cancers^{9,10}.

As emerging tumors are difficult to detect, genome-wide studies of gene expression changes during early tumorigenesis have not been performed. How much tumor gene editing involves suppression of immune genes versus genes that promote intrinsic cell malignancy is unclear.

Genetically engineered mouse model (GEMM) tumors are attractive models because they arise naturally in the tissue, more closely resemble human cancers than cancer cell lines, can be used to study tumor immunity and immunotherapy and have inducible oncogenes that synchronize tumorigenesis and luciferase or fluorescent reporters to identify newly transformed cells. GEMM are notoriously resistant to immunotherapy and thus provide stringent therapeutic models.

Here we took advantage of an aggressive breast cancer GEMM generated using an inducible MMTV-driven *ErbB2ΔE16* oncogene to identify without bias gene expression that is suppressed during early tumorigenesis. Single-cell RNA sequencing (scRNA-seq) comparing early GEMM tumors 1 week after oncogene induction with established tumors 1 month later revealed that most repressed genes were involved in immune signaling, especially in innate immunity, including IFNs and IFN-stimulated genes (ISGs). Many repressed immune genes in late

tumors could be derepressed by inhibiting DNA methyltransferases (DNMTs), reconstituting immunosurveillance.

Results

Cellular map of breast tumor editing

Doxycycline treatment of adult female *ErbB2ΔE16*^{+/−}*MTB*^{+/−} mice¹¹ induces mouse mammary gland tumor virus (MMTV)-driven mammary expression of *ErbB2* lacking exon 16, deleting a 16 amino acid juxtamembrane sequence, to produce constitutively active HER2 and aggressive breast tumors. Within 7–28 days, 80–90% of mice develop metastatic HER2⁺ breast tumors that are resistant to checkpoint blockade (CPB)¹². To identify early changes during tumorigenesis, scRNA-seq libraries were prepared from tumors harvested 5–10 days (early) or 30–35 days (late) ($n = 4$ samples per group) after oncogene induction (Fig. 1a; data available at https://pouryany.shinyapps.io/scgemm_app/). After raw data were assessed for quality using Seurat¹³, 24,798 cells that expressed 20,499 genes were analyzed by unsupervised clustering and manual inspection of canonical markers for annotation. Cells clustered into three major cell types—epithelial cell adhesion molecule (EpCAM⁺) epithelial cells, CD45⁺ immune cells and Fap⁺/Dcn⁺ fibroblasts—and 23 subclusters (Fig. 1b, c, Supplementary Fig. 1 and Supplementary Tables 1 and 2).

Epithelial cells separated into nine subclusters: four luminal-alveolar subgroups, Alv1–4 (*Elf5*⁺*Csn2*⁺); hormone sensing (*Prlr*⁺*Esr1*⁺); basal (*Krt14*⁺); alveolar-basal, with both alveolar and basal characteristics (*Aldh1a3*⁺) and proliferating and/or cycling alveolar (*Mki67*⁺)

cells (Fig. 1b,d–g, Extended Data Fig. 1, Supplementary Fig. 1 and Supplementary Table 3)^{14,15}. EpCAM, enhanced green fluorescent protein (eGFP) and HER2 (*ErbB2*) expression and the CytoTRACE algorithm that stratifies cells based on differentiation/maturation¹⁶ were used to identify normal, transforming and transformed epithelial subclusters (Fig. 1d–g). Fewer than 12% of the hormone sensing subcluster expressed eGFP or HER2 and it contained mainly normal epithelial cells (Extended Data Fig. 1c). Although Alv1 and 2 frequently expressed EpCAM, eGFP and HER2, all the Alv1–4 subtypes expressed more maturation markers than the proliferating, alv-basal and basal subclusters, which were less differentiated. Based on these data, our subsequent analyses focused on the poorly differentiated proliferating, alv-basal and basal subclusters, defining them as tumor cells. The other more differentiated subclusters, which also expressed *EpCAM*, *eGFP* and *ErbB2*, likely contained both normal mammary epithelia and transforming cells. The subcluster abundances in early versus late tumor cells were not significantly different (Extended Data Fig. 1d and Supplementary Table 4).

Downregulation of immune-related tumor genes during editing

Differentially expressed genes (DEG) between late and early tumor cells were identified in pooled alv-basal, basal and proliferating subclusters, using a linear model (MAST). In total, 229 genes were downregulated and 128 genes were upregulated in late versus early samples (false discovery rate (FDR) <0.05, log fold change (llogFC) >0.25) (Fig. 2a and Supplementary Table 5). Prominent downregulated genes were IFNs and ISGs ($n = 144$, ISGs queried from the Interferome database¹⁷), including nine guanylate binding proteins, involved in intracellular infection defense (Supplementary Table 6 and Supplementary Methods). Gene Ontology (GO) enrichment analysis of downregulated tumor genes revealed 179 significant, mostly immune-associated GO terms (q value < 0.05; Fig. 2b and Supplementary Table 7).

The most over-represented terms were involved in innate immunity, especially type I and II IFNs (IFN1 and IFN2) and infection defense. Previously described adaptive immune pathways responsible for antigen processing and presentation and lymphocyte killing, were also highly over-represented. Only one of the top 40 most over-represented GO terms, oxidative phosphorylation, was not directly associated with immunity. This GO term captured multiple genes in mitochondrial electron transport chain complexes I, III, IV and V. Hallmark gene set enrichment analysis in late versus early tumor DEGs also identified the most prominently downregulated pathways as type I and II IFNs and oxidative phosphorylation (Supplementary Table 8). Shared genes in all 179 GO terms significantly downregulated in late tumors were also analyzed using EnrichmentMap. Thirteen major clusters of inter-related terms broadly identified functional groups covering adaptive immunity, innate immunity and metabolism (Fig. 2c). The GO term cluster with the largest number of nodes was associated with cell killing, T cell-mediated immunity and adaptive immunity. The second largest cluster contained terms associated with mitochondrial oxidative phosphorylation and ATP synthesis (Fig. 2c and Extended Data Fig. 1e). Other large clusters focused on innate immunity, type I IFNs, cytokine and chemokine signaling, and adaptive immunity, antigen processing and presentation. Significantly upregulated genes in late tumors concentrated in Hallmark gene sets related to oncogenic signaling, such as KRAS and p53, and to terms related to cell proliferation and survival, metabolism, stress responses or immune inhibition (Supplementary Tables 8 and 9). Apoptosis-related genes became more prominent in late tumors. Metastasis was not among the top upregulated processes. The enrichment terms of downregulated genes had much lower P values than the enrichment terms of upregulated genes. Taken together, GO term enrichment analysis indicated that tumor editing is predominantly immunoeediting to suppress tumor immunity, rather than gene alterations that foster cell proliferation or invasiveness.

Low inflammatory gene expression in early tumors

Pyroptosis and necroptosis cause immunogenic cell death (ICD), which promotes tumor immunogenicity^{18,19}. However, genes involved in these pathways were not DEG in early versus late tumors. Key genes in these pathways (*Gsdmd*, *Gsdme*, *Nlrp3* in pyroptosis, *Ripk3* in necroptosis) were poorly expressed in both early and late tumors (Supplementary Fig. 2a). Their low expression led to low statistical power for analyzing differential expression. Only 1% of tumor cells had any sequencing reads for *Nlrp3*, 7% for *Gsdmd*, and 9% *Ripk3* in early or late tumors (Supplementary Fig. 2b). Although *Gsdme* had associated read counts in 44% of tumor cells, its expression was consistently low. To examine these genes with higher sensitivity, CD45⁺ EpCAM⁺ eGFP⁺ tumor cells were sorted from early and late tumors and key necroptosis, pyroptosis and ISG genes were analyzed by quantitative PCR with reverse transcription (RT–qPCR) (Extended Data Fig. 2a,b). These inflammatory genes were poorly expressed in both early and late GEMM tumors and expression of most genes was significantly lower in late tumors, except for *Gsdmd* and *Gsdme*, which were poorly expressed in both. These genes were also poorly expressed in normal mammary epithelia in five independent scRNA-seq studies (Supplementary Fig. 2c and Supplementary Methods)¹⁵. *Gsdme* was not detected above background in any normal epithelial cell subtype and *Gsdmd*, *Nlrp3* and *Ripk3* showed extremely low expression and only 0–9% of cells of each subtype had any sequencing reads. These data suggest that necroptosis and pyroptosis genes were not significantly edited in this nonmucosal GEMM because their expression was already suppressed in this tissue.

TILs change in late tumors

The scRNA-seq analysis classified tumor-infiltrating immune cells into five major clusters: B cells (*Cd79a*⁺), neutrophils (*S100a9*⁺), macrophages (*Fcgr3*⁺ *Cd74*⁺ *Lyz2*⁺), natural-killer (NK) (*Ncr1*⁺) and T cells (*Cd3d*⁺) (Fig. 1b and Supplementary Fig. 1). Ten T cell subclusters were resolved: two CD4⁺ subsets (regulatory T cells (T_{reg}) and naive CD4⁺ T cells) and eight CD8⁺ subtypes (naive (*Ccr7*⁺ *Sell*⁺ *Il7r*⁺ *Tcf7*⁺ *Bach2*⁺), effector (*Cd44*⁺ *Ctsw*⁺ *Nkg7*⁺), proliferating (*Cd44*⁺ *Mki67*⁺ *Top2a*⁺ *Stmn1*⁺), tissue-resident memory-like²⁰ (TRM, *Cd44*⁺ *Itgae*⁺ *Xcl1*⁺ *Chn2*⁺), central memory (*Cd44*⁺ *Ccr7*⁺ *Sell*⁺ *Il7r*⁺ *Tcf7*⁺), progenitor exhausted (*Cd44*⁺ *Tcf7*⁺ *Bach2*⁺ *Pdcd1*⁺), exhausted (*Cd44*⁺ *Tcf7*⁺ *Bach2*⁺ *Fcer1g*⁺ *Pdcd1*^{hi} *Lag3*^{hi} *Cd244a*^{hi}) and a subset with high ISG expression^{21,22} (CD8⁺ ISG⁺ T cells, *Isg15*⁺ *Ifit1*⁺ *Irf7*⁺) (Figs. 1b and 3a, Extended Data Figs. 3a–c and 4 and Supplementary Table 10). CD8⁺ TRM tumor-infiltrating lymphocytes (TILs) survive longer, proliferate more and kill tumor cells more efficiently than other subsets^{20,23}. Activated TRM cells produce IFN γ and tumor necrosis factor (TNF), which drive dendritic cell maturation to boost tumor immunity²⁴. Progenitor exhausted CD8⁺ TIL live longer, proliferate more and serve as a reservoir for exhausted CD8⁺ TIL that have higher coinhibitory receptor gene expression, short lifespan and limited proliferation²⁵. The CD8 ISG⁺ TIL cluster in humans was postulated to be an IFN-induced activation state preceding exhaustion²².

T and NK cell characteristics of early and late samples were next compared by Dirichlet multinomial regression. The proportion of CD4⁺ and CD8⁺ naive T cells significantly declined, while progenitor exhausted and exhausted CD8⁺ T cells increased in late tumors (Fig. 3b and Supplementary Table 4). Phenotypic and functional marker genes of CD4⁺ and CD8⁺ T cells and NK within the tumor showed consistent changes (Fig. 3c and Supplementary Table 11). All three subsets showed reduced *Klf2*, a transcription factor controlling naive lymphocyte homing to secondary lymphoid tissue. Late tumor CD4⁺ and CD8⁺ TIL significantly reduced expression of naive and memory cell marker transcripts, including *Ccr7*, *Cd69* and *Cd28*. CD8⁺ T cells as a whole also downregulated *Tcf7*, encoding TCF1, required for T cell stemness, memory formation and CD8 effector functions, and upregulated a large panel of coinhibitory receptor genes (*Pdcd1* (encoding PD-1), *Lag3*, *Cd244a*, *Tigit*, *Ctla4*, *Cd160*)^{26,27}. NK TIL showed changes characteristic of exhaustion, including significantly reduced *Eomes*

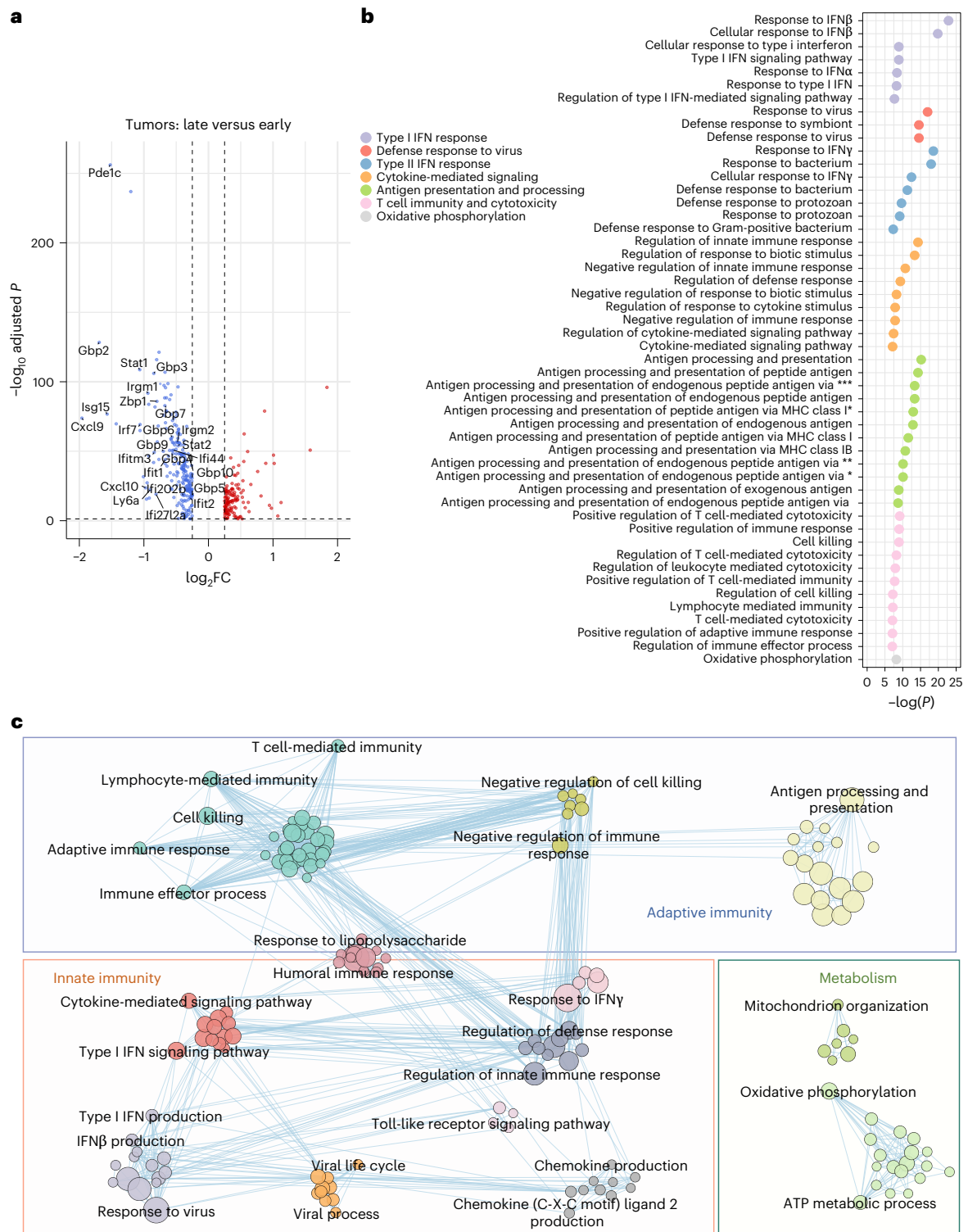


Fig. 2 | Tumor cells downregulate immune-related genes during early tumorigenesis. **a**, Volcano plot of DEGs between early and late tumors. Generalized linear models (MAST test) were used to determine two-sided *P* values and define DEGs ($\log_{2}FC > 0.25$, FDR adjusted *P* < 0.05). With this cut-off, 229 genes were downregulated and 128 were upregulated in late tumors compared to early tumors. **b**, Top 50 GO terms with most significant enrichment *P* values associated with downregulated genes in late tumors ($q < 0.05$). GO terms groups were manually annotated. **c**, Interaction network of significant GO terms ($q < 0.05$) associated with downregulated genes in late

versus early tumors. The network was generated using the EnrichmentMap method based on the similarity of GO terms and shared genes. Colors represent unsupervised clusters identified by graph clustering. Unsupervised clusters were manually classified into broader groups to represent adaptive immunity, innate immunity and metabolism. Node sizes are proportional to the significance of enrichment ($-\log(P)$). **b, c**, Enrichment statistic: one-sided Fisher's exact test. *P* values were adjusted for multiple comparisons using the *q* value method. Network *P* values, node and edge information are provided in Source Data Fig. 2.

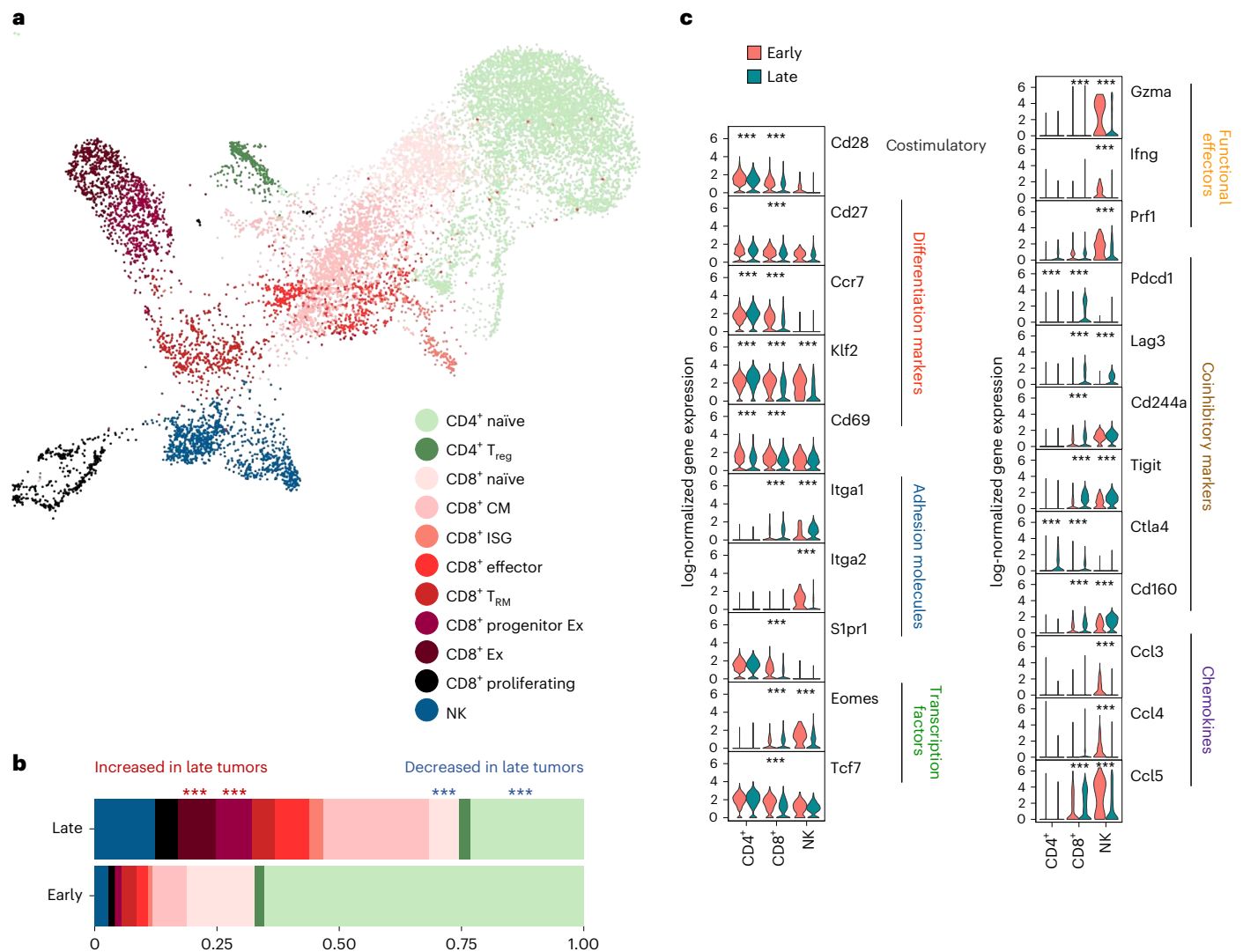


Fig. 3 | Tumor-infiltrating T and NK cells from late tumors show signs of exhaustion. **a**, UMAP of T and NK cells. **b**, Proportions of tumor-infiltrating T and NK cell clusters in early and late tumors. Asterisks denote significant differences (FDR adjusted $P < 0.05$) between the two groups, determined by two-sided Dirichlet multinomial test. Color code as in **a**. **c**, Comparison of expression of key markers by tumor-infiltrating T and NK cells in early and late-stage tumors. Markers correspond to costimulatory molecules (black), differentiation and

activation markers (red), adhesion molecules (blue), transcription factors (green), functional effectors (orange), exhaustion markers (brown) and chemokines (purple). Generalized linear models (MAST test) were used to determine two-sided P values of differential expression. The y axis represents log-normalized gene expression, based on scaled and log transformed scRNA-seq read counts in each cell. ***denotes FDR adjusted $P < 0.05$ and $|\log FC| > 0.25$. **b,c**, P values are provided in Source Data Fig. 3.

(NK functionality master regulator) and NK effector genes (*Gzma*, *Ifng*, *Prf1*), and increased coinhibitory receptor genes (*Lag3*, *Tigit*, *Cd160*) (Fig. 3c and Supplementary Table 11). The changes in gene expression within CD8⁺ T cell subclusters were less prominent, likely because the subcluster designation already identified distinct functionalities (Extended Data Fig. 3d and Supplementary Table 11). Thus, the scRNA-seq data captured the evolution of T cell exhaustion in early tumorigenesis.

Tumor editing disrupts tumor-immune cell communication

The downregulated transcripts suggested that tumor editing might disrupt communication between tumor cells and TIL. CellChat²⁸ was used to compare inferred cell-cell interactions in the early versus late scRNA-seq data, based on changes in ligand-receptor gene coexpression in all cells. More than 107 significant cellular communication pathways were active in early or late tumors (Fig. 4a,b and Extended Data Fig. 5). Here, 44 pathways were inferred to be significantly more active in early samples while 29 processes had higher activity in late samples

(Wilcoxon rank sum test $P < 0.05$) (Extended Data Fig. 5b). Thirteen pathways active in early samples were inactive in late tumors including important innate (type II IFNs, LIGHT (TNFSF14), IL2 signaling) and adaptive (CD80 costimulation, major histocompatibility complex (MHC)-II binding) immune pathways. NK expressed IFNII, which activates macrophages and suppress tumor growth, which was sensed by a broad range of tumor subtypes and most immune cells, only in early tumors. Similarly, CD8⁺ TRM cells expressed LIGHT, a proinflammatory cytokine that stimulates T cells and activates NF- κ B to increase innate immunity and/or trigger apoptosis, which was sensed only in early tumors by all tumor and some immune subclusters.

As cytotoxic NK and CD8⁺ TIL are key to inhibiting tumor growth, we further analyzed inferred cell communication pathways between NK or CD8⁺ TRM and tumor cells (Fig. 4c, Extended Data Fig. 6 and Supplementary Fig. 3). Multiple ligand-receptor interactions in early tumors decreased or became undetectable in late tumors, including macrophage migration inhibitory factor (MIF), ITGB2-ICAM1, IL-7, TNF and GzMA, which are involved in immune cell recruitment

and functions. When we combined cell–cell communication and DEG analysis (Fig. 4d), loss of IFNII signaling between NK and tumor cells could be attributed to loss of IFN γ transcripts in late sample NK. MIF interacts with its receptor CD74 to initiate a signaling cascade that activates NF- κ B and promotes lymphocyte survival²⁹. Loss of MIF signaling in late tumors could be attributed to reduced tumor cell transcripts for CD74, which heterodimerizes with CD44.

The inferred cellular communications between individual CD8⁺ TIL subtypes and tumor cells lost in late samples included MHC-I–CD8 interactions between tumors and multiple CD8⁺ TIL subtypes and MIF, IL2, CXCL and CX3C signaling pathways in late samples (Extended Data Fig. 6). Comparing interactions between pooled tumor cells and pooled CD8⁺ TIL, potential significant increases in coinhibitory signaling were found in late tumors (Supplementary Fig. 4). Inferred interactions between pooled CD8⁺ TIL with tumor cells and tumor-infiltrating myeloid cells suggested enhanced coinhibitory signaling in late tumors (PD-1–PD-L1, TIGIT–PVR/Nectin 2, CTLA-4–ICOSL/CD86), was mostly due to increased coinhibitory receptor expression on CD8⁺ TIL, rather than increased immune checkpoint ligands on tumor cells or tumor-infiltrating myeloid cells (Fig. 3c and Supplementary Figs. 4–7). Other checkpoint inhibitory receptor–ligand encoding genes (*Pdcd1lg* (PD-L2), *Fgl1* (LAG3 ligand)) or *Lgals9* (galectin 9 ligand of Tim3) were either not detected or barely detected (Supplementary Fig. 7). Collectively, our data suggest that changes in both tumor and infiltrating immune cells reduce immunosurveillance.

DAC increases GEMM tumor immunogenicity and immune control

DNMT promoter methylation represses gene expression. To assess whether DNA hypermethylation reduces tumor immunity, GEMM bearing late tumors (average tumor size, 300 mm³, ~1 month after doxycycline induction) were treated with decitabine (DAC, 5 mg kg⁻¹ ×3 in week 1, 2 mg kg⁻¹ ×3 in week 2) or PBS control. Two mice bearing the largest tumors (~1,000 mm³) were assigned to receive DAC to test its effectiveness against large tumors. The highest dose tested (5 mg kg⁻¹) was the allometrically scaled mouse dose corresponding to the low dose (15 mg m⁻² per day) used to treat elderly hematological malignancy, which achieves nanomolar human plasma concentrations^{30,31}. While the average tumor volume of PBS-treated mice increased fourfold within 18 days, the tumor volume of mice that received DAC decreased more than eightfold and two mice with the smallest tumors (of six DAC treated) became tumor-free (Fig. 5a). DAC created a highly immunogenic TME with increased CD8⁺, CD4⁺ and NK TIL (Fig. 5b). The proportion of CD8⁺CD103⁺ TRM cells increased roughly fourfold (Fig. 5c)²³. DAC-treated tumors had twice as many classical type 1 dendritic cells (cDC1) (Fig. 5d), which optimally prime TRM cells³². By contrast, immunosuppressive myeloid-derived suppressor cells (MDSCs) and tumor-associated macrophages were strongly reduced in DAC-treated tumors (Fig. 5e). Furthermore, in DAC-treated mice, many more NK TIL expressed granzyme B (Gzmb) and perforin (Pfn), and more CD8⁺

and NK TIL produced IFN γ and TNF after ex vivo stimulation (Fig. 5f,g). Thus, inhibiting DNA methylation strongly suppressed GEMM tumor growth and reconstituted antitumor immunity.

Key innate immune genes in GEMM were also compared by RT–qPCR analysis of sorted CD45⁺EpCAM⁺GFP⁺ late tumor cells collected 1 day following the last DAC or vehicle treatment (Extended Data Fig. 2c). DAC-upregulated *Ripk1* and *Ripk3* (necroptosis), *Gsdmd* and *Gsdme* (pyroptosis), *Irf7* (IFN pathway) and *Cxcl9* (chemokine). Some other innate immune genes in GEMM tumor cells showed a DAC-related increase that was not statistically significant.

DAC suppresses 4T1 tumors and enhances antitumor immunity

Mice bearing palpable orthotopic 4T1 triple-negative breast tumors were also treated with three doses of DAC (0.2, 1 or 5 mg kg⁻¹, three times a week for 2 weeks). All three doses significantly enhanced tumor immunity and delayed tumor growth, but the higher doses were more effective (Fig. 6a and Extended Data Fig. 7). The two higher doses increased CD8⁺ total and TRM TIL (Extended Data Fig. 7a–c). DAC also enhanced cDC1s in the TME (Extended Data Fig. 7d) and reduced MDSCs (Extended Data Fig. 7e). Both higher doses also strongly increased Gzmb-, IFN γ - and TNF-producing CD8⁺ TIL and CD8⁺ TIL polyfunctionality (Extended Data Fig. 7f–g). Thus, low-dose DAC can reinvigorate tumor immunity in two aggressive breast cancer models.

DAC derepresses silenced immune genes in multiple pathways

RT–qPCR also assessed the effect of in vitro low-dose DAC (0.1, 0.5, 1 μ M) on 18 key type I IFN and inflammation pathway messenger RNAs (mRNAs) in 4T1 (Fig. 6b). These low DAC concentrations maintained cell viability without altering cell cycle profile. DAC potently upregulated most of these genes in a dose-dependent manner (Extended Data Fig. 8a,b). Most of the same genes were significantly upregulated in 4T1 knocked down for *Dnmt1* or *Dnmt3b*, confirming that inhibiting DNMT derepresses innate immune genes (Extended Data Fig. 8c,d).

RNA-seq of 4T1 treated in vitro for 48 hours with 1 μ M DAC or PBS (Fig. 6c and Supplementary Tables 12 and 13) was also compared. As expected, DAC predominantly upregulated gene expression: 2,622 genes were significantly upregulated and 717 genes were downregulated ($|\log_{2}FC| > 1$ and FDR < 0.0001). Concordant with the RNA-seq data, assay for transposase-accessible chromatin with sequencing (ATAC-seq) and reduced representation bisulfite sequencing (RRBS) showed regions of DAC-induced increased chromatin accessibility (Extended Data Fig. 8e and Supplementary Table 14) and DNA hypomethylation (Supplementary Table 15 and Supplementary Methods). Upregulated innate immune genes that correlated well with both DNA hypomethylation and increased accessibility of promoter regions included *Ddx58*, *Irf7* and *Ripk3* (Extended Data Fig. 8e). Genes upregulated by DAC in 4T1 and downregulated in late GEMM tumors significantly overlapped: among genes detected in both datasets, 84 of 211 downregulated genes in the GEMM were upregulated in DAC-treated

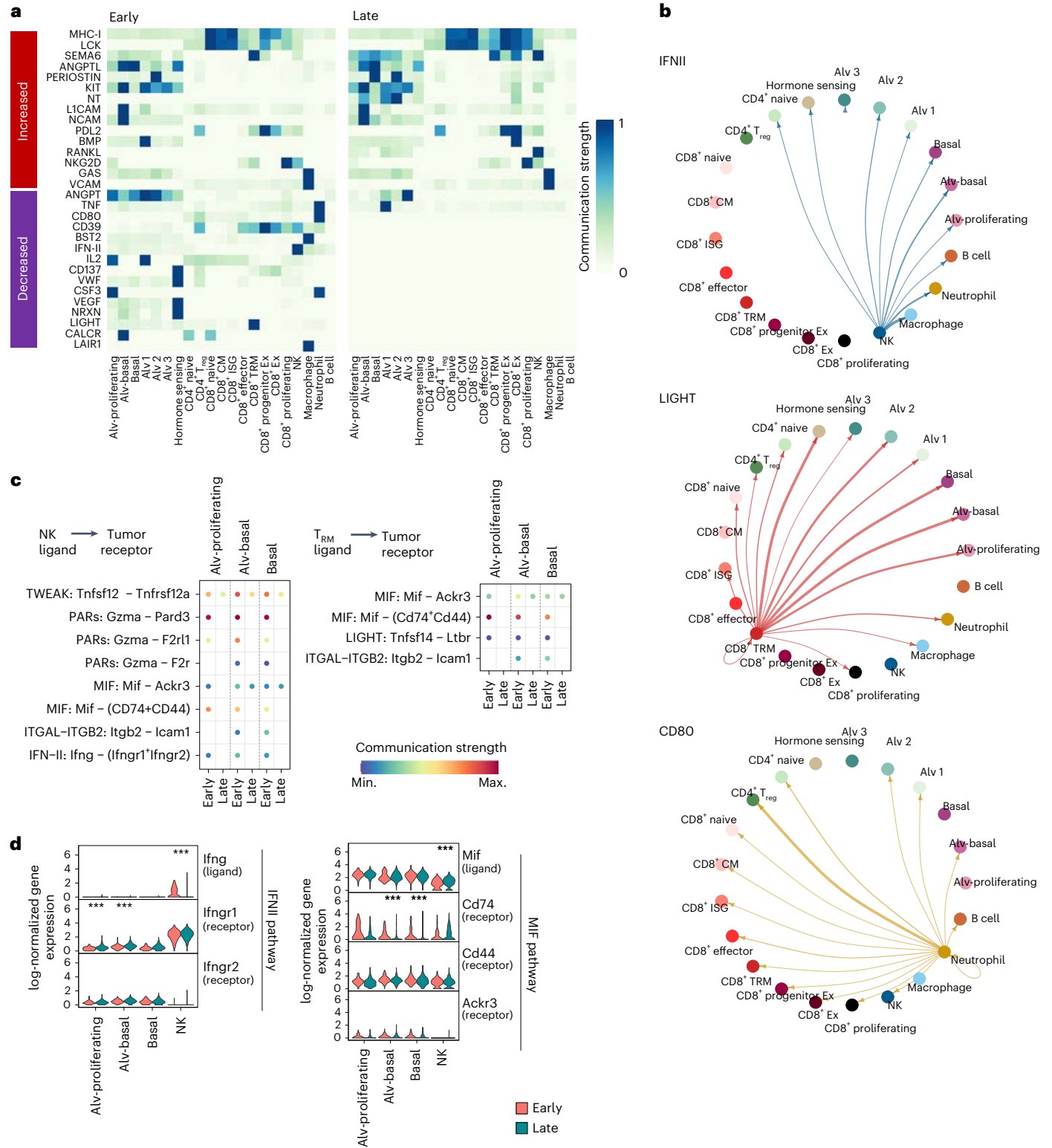
Fig. 4 | Disrupted tumor–immune cell communication during tumor progression. **a**, Heatmap of top 30 pathways with most significant changes in inferred communications between early and late tumors analyzed using CellChat software. Rows denote communication pathways. Columns denote cell types that participate in active interactions as either the sender (expressing ligands) or receiver (expressing receptors) of signaling pathways. Communication strength is assigned based on coexpression of ligand–receptor pairs. Non-zero strength values denote significant inferred communication events, identified using the CellChat one-sided permutation test ($P < 0.05$). **b**, Selected inferred cell–cell communication pathways with active signaling in early samples that were lost in late samples. Arrows represent significant interactions between ligands (source nodes) and receptors (target nodes). Line colors represent the source of signaling and line widths represent the strength of the communication. Significant interactions were computed using the CellChat one-sided permutation test

($P < 0.05$). **c**, Comparison of selected signals from NK cells (left) or CD8⁺ TRM cells (right) to tumor cell subclusters in early versus late tumors. Rows represent predicted ligand–receptor interactions. Solid points represent significant interactions ($P < 0.05$). Communication strength is related to coexpression of ligand–receptor pairs. P values represent the significance of the communication strength values, which were calculated using the CellChat one-sided permutation test. **d**, Violin plots show normalized expression of ligands and receptors of IFNII (left) and MIF (right) pathway inferred interactions between tumor cells and NK and CD8 TRM in early and late tumors. The y axis represents log-normalized gene expression, based on scaled and log transformed scRNA-seq read counts in each cell. Generalized linear models (MAST test) were used to determine two-sided P values of differential expression. ***denotes FDR adjusted $P < 0.05$ and $|\log_{2}FC| > 0.25$. P values are provided in Source Data Fig. 4. Max., maximum; min., minimum; PARs, protease activated receptors.

4T1 (40%, $P = 7 \times 10^{-27}$) (Fig. 6d). Most of these overlapping genes participate in IFN or other cytokine signaling, antigen presentation or cell killing (Extended Data Fig. 8f). A large fraction of DAC-upregulated genes play a key role in IFN α induction: 95% of overlapping genes were ISGs (Fig. 6d, Supplementary Table 16 and Supplementary Methods). DNMT inhibition induces endogenous retrovirus (ERV) demethylation, which activates RNA responses to induce IFN α signaling and promote tumor immunotherapy responsiveness⁹. However, DAC treatment of 4T1 did not consistently increase ERV expression (Supplementary Fig. 8

and Supplementary Methods, ERVs queried from the gEVE database³³), suggesting that low-dose DAC may activate IFN α signaling by acting on genes other than ERVs.

Among genes expressed in both 4T1 RNA-seq and GEMM scRNA-seq, 1,539 DAC-upregulated genes that were not significantly suppressed in late tumors were also highly enriched in immune-related GO terms (Fig. 6e, Extended Data Fig. 8f and Supplementary Table 16). Unsupervised EnrichmentMap analysis of the top 150 GO terms enriched in 4T1 DAC-upregulated genes showed that clusters



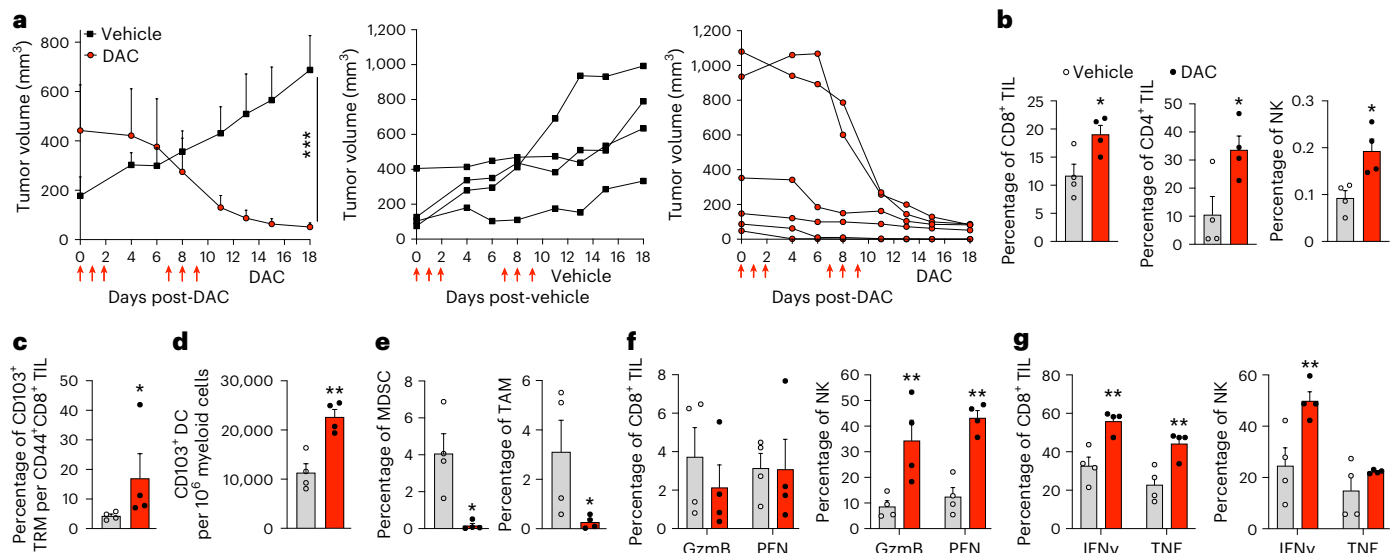


Fig. 5 | DAC inhibits breast tumor growth and promotes antitumor immunity.

a–g. Effect of DAC on tumor growth and immune response to breast tumors in *ErbB2ΔEx16* mice induced by doxycycline 1 month earlier. DAC (5 mg kg⁻¹ in week 1; 2 mg kg⁻¹ in week 2) or PBS were administered on days indicated by red arrows. **a**, Mean tumor growth in mice (left panel) and tumor growth curves of individual vehicle-treated control mice (middle) and DAC-treated mice (right). **b**, Percentage of tumor-infiltrating CD8⁺ T cells (left), CD4⁺ T cells (middle) or NK cells (right) of live cells in mice treated with vehicle or DAC. **c**, Percentage of CD103⁺ TRM within antigen-experienced CD44⁺ CD8⁺ TIL in mice treated with vehicle or DAC. **d**, Numbers of CD103⁺ dendritic cells (DC) per 10⁶ myeloid cells in tumors. **e**, Percentage of Gr-1⁺ CD11b⁺ MDSC (left) or CD11b⁺ F4/80⁺ tumor-associated macrophages (TAM) (right) of live cells in tumors from DAC or

vehicle-treated mice. **f**, Percentage of CD8⁺ TIL (left) or NK TIL (right) expressing GzmB and PFN in tumors from vehicle or DAC-treated mice. **g**, Percentage of CD8⁺ TIL (left) or NK TIL (right) producing IFN γ or TNF after ex vivo activation with phorbol myristate acetate (PMA) and ionomycin in tumors from vehicle or DAC-treated mice. Data shown are mean \pm s.e.m. and are representative of two independent experiments. **a**, Tumor volumes on day 18 were compared by Student's *t*-test (vehicle (PBS) *n* = 4, DAC *n* = 6); **b–e**, Student's *t*-test or Mann–Whitney test. **f, g**, Multiple *t*-test with type I error correction with Holm–Sidak method (vehicle (PBS) *n* = 4, DAC *n* = 4 since two DAC-treated mice were cured) **P* \leq 0.05, ***P* \leq 0.01, ****P* \leq 0.001, *****P* \leq 0.0001. Exact *P* values are provided in the Source Data.

of GO terms were also mostly immune-related (14 of 15 functional modules) (Fig. 6e). This over-representation of immune-related genes in the DAC-upregulated genes was unexpected given the prominent role of DNA methylation in suppressing alternate cell differentiation states. In fact, the one significant derepressed gene module that was not immune-related contained genes involved in meiosis and germ cell state, which did not appear to be related to tumorigenesis, but rather reflected the role of DNA hypermethylation in repressing the germline in somatic cells. Therefore, in a triple-negative breast cancer (TNBC) cell line orthotopic implant, DAC strongly preferentially derepressed immune-related genes that could promote tumor immunity.

DAC activity in 4T1 depends on cytotoxic lymphocytes

Next, we investigated what immune pathways contribute to improved DAC-mediated tumor control. Tumor cell elimination by killer lymphocytes, CD8⁺ and NK TIL, depends on PFN to deliver death-inducing granzymes into tumors. DAC protection against 4T1 orthotopic tumors was ablated in PFN-deficient mice, indicating the critical role of cytotoxic lymphocyte killing in DAC protection (Fig. 7a).

DAC activity against B16 depends on multiple innate pathways

Because the Type I IFN pathway was the most enriched innate immune pathway derepressed by DAC, we compared DAC treatment of wild-type (WT) and *Ifnar1*^{-/-} mice to determine whether Type I IFNs were required for DAC-mediated protection. Because *Ifnar1*^{-/-} mice were only available in the C57/BL6 background and a mouse TNBC cell line in this background was not available, we compared DAC's effect in WT and *Ifnar1*^{-/-} C57BL/6 mice bearing B16F10 melanoma (Fig. 7b). Tumors in IFNAR1-deficient mice were still strongly inhibited by DAC. However, B16 grew about twice as fast in *Ifnar1*^{-/-} mice as in WT mice, either with or without DAC (vehicle group: WT versus *Ifnar1*^{-/-}:

P = 0.044 by two-way analysis of variance (ANOVA) with Šidák's multiple-comparisons test). B16 tumor-infiltrating immune cell numbers were comparable in DAC-treated WT and *Ifnar1*^{-/-} mice, although expression of PFN, GzmB and IFN γ by CD8⁺ and NK TIL was reduced in *Ifnar1*^{-/-} mice. Thus, although DAC's efficacy only weakly depended on IFN1 signaling, DAC enhanced the antitumor functionality of killer lymphocytes.

To assess whether B16 protection is mediated by DAC induction of necroptosis or pyroptosis, mice bearing palpable B16 tumors were treated with two cycles of low-dose DAC (1 mg kg⁻¹) or PBS and tumor lysates were analyzed by immunoblotting for activation of proteins that trigger different forms of cell death: caspase 3 (apoptosis in cells not expressing GSDME, pyroptosis in cells expressing GSDME), GSDME (pyroptosis) and phospho-mixed lineage kinase domain-like protein (pMLKL) (necroptosis) (Fig. 7c). All three death mediators were activated in DAC-treated, compared to vehicle-treated, melanoma. Cells undergoing inflammatory cell death, but not apoptosis, permeabilize their cell membrane quickly and take up cell-impermeant dyes, such as propidium iodide and 4,6-diamidino-2-phenylindole (DAPI). To quantify DAC-induced inflammatory cell death in vivo, mice bearing B16-eGFP tumors were treated with two cycles of DAC or PBS, injected with propidium iodide intravenously 48 hours after the last injection and tumors were harvested 10 min later. DAPI was added to the staining solution to label cells with permeabilized membranes. DAC-treated tumors contained significantly more GFP⁺ tumor cells that took up both dyes than control mouse tumors (Fig. 7d). DAC-induced inflammatory cell death was also measured in mice bearing B16-eGFP tumors knocked out or not for *Gsdme*, *Gsdmd* or *Ripk3* (Extended Data Fig. 9a–e). Although only a few percent of untreated WT B16 tumors took up the dyes, ~20% of DAC-treated WT B16 tumor cells took up both dyes, indicating a high frequency were undergoing inflammatory cell death.

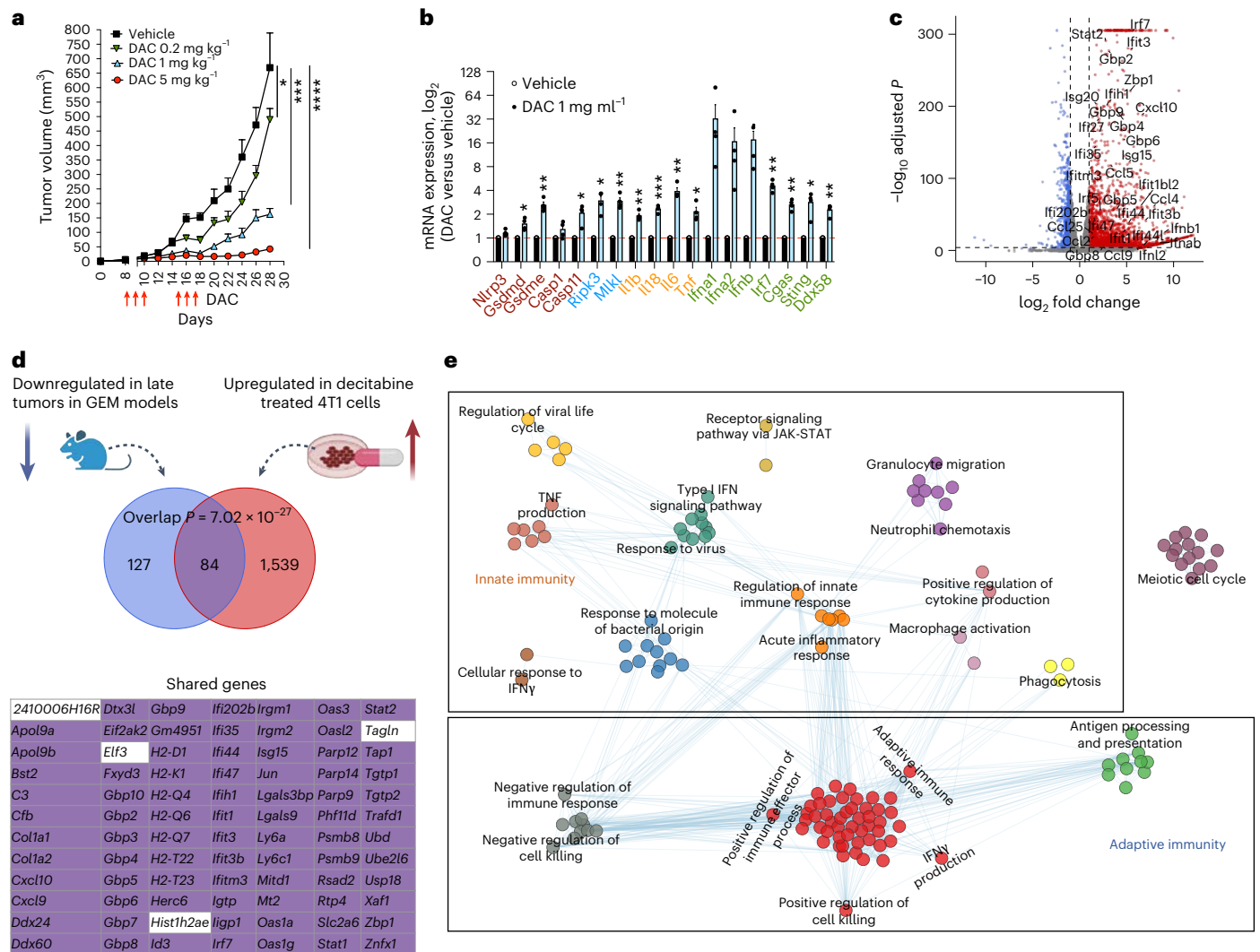


Fig. 6 | DAC-upregulated genes overlap with genes downregulated in late tumors. **a**, Growth of orthotopic 4T1 tumors in mice treated with vehicle (PBS) ($n = 7$) or DAC for two cycles, starting when tumors were palpable (0.2 mg kg⁻¹, $n = 8$; 1 mg kg⁻¹, $n = 5$; 5 mg kg⁻¹, $n = 7$). Red arrows indicate treatment. **b**, Relative mRNA expression of immune-related genes in orthotopic 4T1 tumors from mice that received two cycles of DAC (1 mg kg⁻¹) compared to untreated control mice (PBS) ($n = 4$ per group), analyzed by RT-qPCR. Genes whose products participate in pyroptosis (red), necroptosis (blue), inflammatory cytokines (orange) and the type I IFN pathway (green) were analyzed. mRNA of each target gene was normalized to *Gapdh* and then the ratio in the DAC-treated group was normalized to the ratio in the vehicle control group. **c**, Volcano plot of significant DEGs in 4T1 after in vitro exposure to DAC. Two-sided P values were determined by the DESeq2 algorithm (Wald test). P values were adjusted for multiple comparisons using the FDR method. **d**, Overlap between genes significantly downregulated during immunoeediting of GEMM breast tumors and genes significantly upregulated by DAC treatment of 4T1 breast tumor cells. The Venn diagram represents genes that are expressed in both 4T1 RNA-seq and GEMM scRNA-seq.

Statistics were calculated using Fisher's exact test. Numbers in red and blue part indicate the nonoverlapping genes. Lower table: 84 overlapping genes downregulated in late-stage transformed tumor cells and upregulated in vitro DAC-treated 4T1. ISGs are highlighted in purple. **e**, Interaction network of the top 150 GO terms significantly associated ($q < 0.05$) with DAC-upregulated genes in 4T1. The network was determined by EnrichmentMap based on the similarity of GO terms and shared genes. Colors represent unsupervised clusters identified by graph clustering. Unsupervised clusters were manually classified into broader functional groups. Enrichment statistic: one-sided Fisher's exact test. P values were adjusted for multiple comparison using the q value method. Data shown are mean + s.e.m. and are representative of two independent experiments. **a**, Tumor growth curves were compared by calculating the area-under-the-curve values for each sample followed by one-way ANOVA with Holm-Sidak's multiple comparisons. **b**, Multiple t -tests. FDR is calculated by two-stage step-up method of Benjamini, Krieger and Yekutieli. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. **d**, Created using [Biorender.com](https://biorender.com). Exact P values are provided in the Source Data.

About half as many DAC-treated *Gsdme*^{-/-} or *Ripk3*^{-/-} tumors took up dye, compared to DAC-treated WT or *Gsdmd*^{-/-} tumors (Fig. 7d). Thus, DAC induced both GSDME-mediated pyroptosis and RIPK3-dependent necroptosis, two types of ICD.

To determine which innate immune pathways contribute to DAC's antitumor activity in B16, mice bearing palpable WT, *Irf3*^{-/-} *Irf7*^{-/-}, *Gsdme*^{-/-}, *Ripk3*^{-/-} or *Gsdmd*^{-/-} B16 tumors were treated with two cycles of DAC or PBS. *Irf3*^{-/-} *Irf7*^{-/-} cells lack the key transcription factors for IFN signaling. DAC significantly delayed the growth of WT and

Gsdmd^{-/-} tumors (Fig. 6e, $P = 0.02$). DAC's antitumor effect ($P = 0.012$ in mice bearing WT tumors) was no longer significant against *Irf3*^{-/-} *Irf7*^{-/-}, *Gsdme*^{-/-} or *Ripk3*^{-/-} tumors (Fig. 7e, $P = 0.18$, 0.16 or 0.23, respectively). Furthermore, DAC-induced increases in CD8⁺ TIL and GzmB in both CD8⁺ and NK TIL attenuated in all knockout tumors (Extended Data Fig. 9f). Thus, multiple innate immune pathways—type I IFNs, GSDME-mediated pyroptosis and RIPK3-mediated necroptosis—contribute to DAC-mediated growth suppression and immunostimulation of B16.

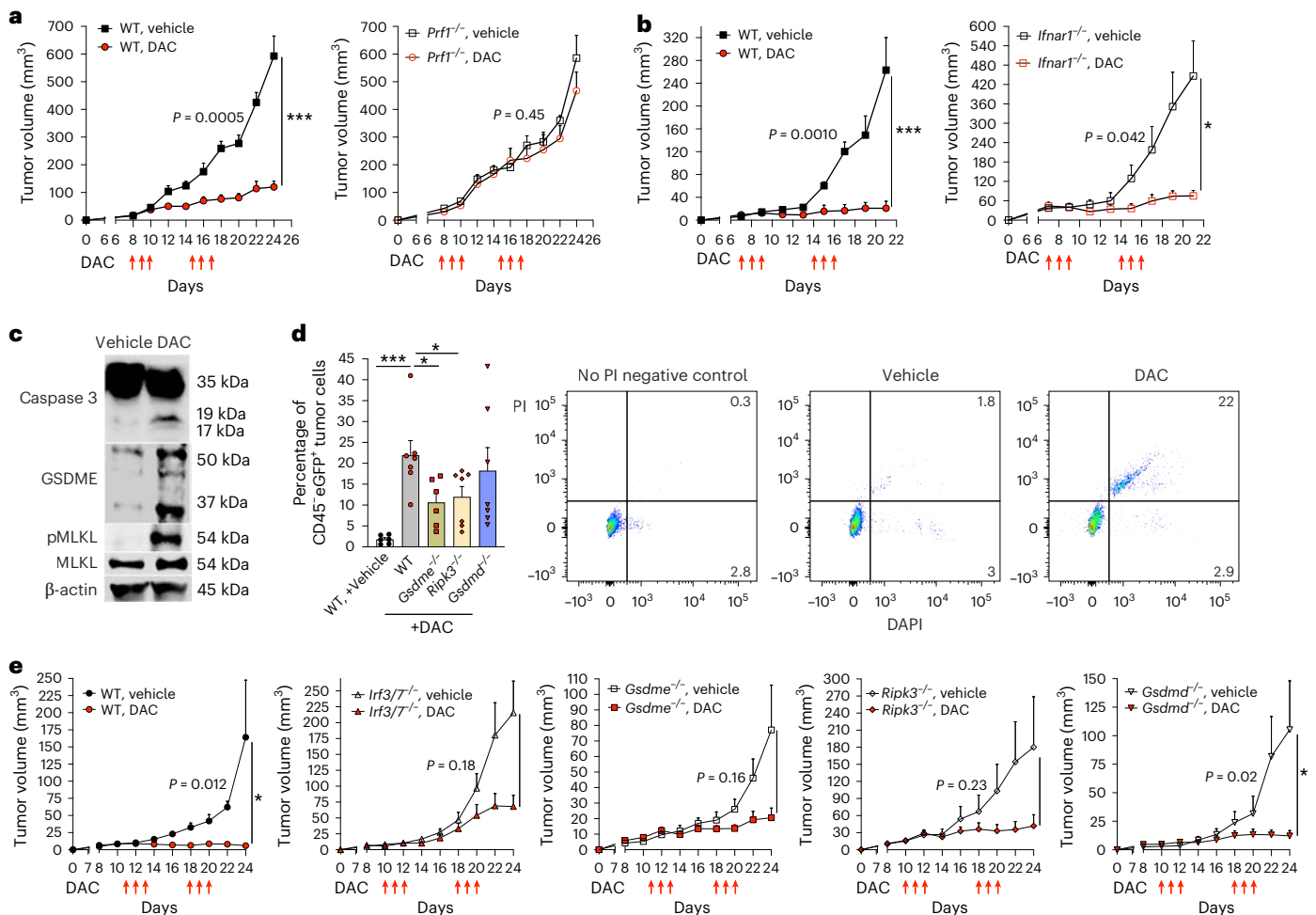


Fig. 7 | Inflammatory cell death and IFN signaling contribute to DAC tumor control. **a**, Growth of orthotopic 4T1 tumors in WT (left) or *Prf1*^{-/-} (right) mice treated with vehicle (PBS) or DAC (1 mg kg⁻¹) for two cycles when tumors became palpable (*n* = 5–7 per group). **b**, Growth of subcutaneous B16F10 tumors in WT (left) or *Ifnar1*^{-/-} (right) mice treated with vehicle or DAC (1 mg kg⁻¹) for two cycles when tumors became palpable (*n* = 5 per group). **c**, Mice challenged subcutaneously with B16F10 tumors were treated with vehicle (PBS) or DAC, 1 mg kg⁻¹ for two cycles, starting when tumors were palpable. Shown are immunoblots of representative tumor lysates probed for caspase 3, GSDME, pMLKL, MLKL and β-actin loading control. Experiment was repeated twice with similar results. **d**, Mice implanted with WT, *Gsdme*^{-/-}, *Ripk3*^{-/-} or *Gsdmd*^{-/-} B16F10-eGFP tumors were treated with PBS (vehicle) or DAC (1 mg kg⁻¹) for two cycles, starting when tumors became palpable. 48 h after the last DAC injection, mice were injected intravenously with propidium iodide (PI) and

ethanized 10 min later. Shown are the percentage of tumor cells that stained for DNA dyes, propidium iodide and DAPI (DAPI⁺PI⁺), to identify cells undergoing necrotic cell death (*n* = 6–7 per group). Right shows representative flow plots of propidium iodide and DAPI staining of CD45⁺ eGFP⁺ B16F10 tumor cells after vehicle or DAC treatment of WT tumor-bearing mice. **e**, Growth of WT, *Gsdme*^{-/-}, *Ripk3*^{-/-}, *Irf3*^{-/-} *Irf7*^{-/-} or *Gsdmd*^{-/-} B16F10 tumor cells in mice treated with vehicle or DAC (1 mg kg⁻¹) for two cycles (*n* = 5–7 per group). **a, e**, Red arrows indicate time of treatment. Data shown are mean + s.e.m. and are representative of two independent experiments. **a, b, e**, Tumor growth curves were compared by calculating the area-under-the-curve values for each sample followed by Student's *t*-test. **d**, One-way ANOVA with Fisher's least significant difference test. **P* ≤ 0.05, ***P* ≤ 0.01, ****P* ≤ 0.001, *****P* ≤ 0.0001. Exact *P* values are provided in the Source Data.

Discussion

This study used an aggressive breast cancer GEMM to investigate without bias changes in tumor and immune cell gene expression in newly formed tumors. Tumor cell gene expression changes were concentrated on suppressing genes that allow the tumor to be recognized by both innate and adaptive immune pathways, rather than on altering genes that promote tumor cell-intrinsic proliferation. Thus, early tumor gene editing is truly ‘immunoediting’ to reduce tumor immunogenicity, at least in this model. Many key immune pathway genes were epigenetically repressed early in tumor formation and derepressed in late tumors by inhibiting DNA methylation. Alterations of type I and II IFN pathway genes and ISGs were especially prominent. Our study suggests that immunoediting mainly depends on epigenetic modulation since genes suppressed during immunoediting and those derepressed by low-dose DAC overlap, and inhibiting DNMTs restores immune control.

Consistently, DAC increases DNA demethylation and chromatin accessibility of immune genes. This study revealed how TIL change their phenotype, gene expression, functionality and interactions with tumor cells early in tumorigenesis. T cell exhaustion occurred rapidly, within 1 month of oncogene induction.

Capturing gene expression in early tumorigenesis is technically challenging, as tumor evolution initially happens in a few cells that are difficult to detect. No previous study has used single-cell resolution to characterize immunoediting changes starting when tumors are just detectable. GEMM provide an especially valuable resource for studying tumor immunoediting because tumors arise in situ soon after oncogene induction and can be synchronized and captured very early. By contrast, transplantable tumor cell lines have been passaged in mice for many generations and have already been edited at the time of tumor implantation. Human patient-derived xenograft tumors

are also derived from already edited detectable tumors. Moreover, immunocompetent patient-derived xenograft mouse models are not available to study tumor changes under immune selection pressure. Studying immunoeediting in humans is challenging but might be possible in patients who are genetically cancer-prone who undergo frequent screening.

GEMM can be used to get snapshots of how evolving changes in tumor cells affect other cells in the TME. This study is just a beginning. Future GEMM studies with scRNA-seq at more time points could provide a more detailed picture of the evolution of immunoeediting and how it creates a TME that supports tumor persistence, vascularization and invasiveness. It could capture the activation and clonal expansion of tumor-specific T cells and reveal how immunoeediting drives T and NK cell dysfunction, myeloid cell suppression and tumor evasion. Studies in other GEMM could determine whether the prominence of innate and adaptive immune gene editing is a common feature of different cancers and describe how different tumors respond to immunosurveillance or other challenges specific to their tissue of origin. This study did not have adequate sequencing depth to capture changes in myeloid cells or stroma. More in-depth sequencing or sequencing of selected cell subsets will be needed to evaluate tumor interactions with these other TME cells. Spatial transcriptomics could provide a way to understand the molecular basis of immune cell exclusion and what distinguishes the tumors that coexist with infiltrating immune cells or organize tertiary lymphoid structures.

Depleting specific immune cell subsets before tumor induction or DAC treatment will provide information about the relative contributions of different immune cell types to immunosurveillance and immunotherapeutic protection. Previously, administration of tumor-targeted small-interfering RNAs (siRNAs) that knockdown genes involved in cancer immunity showed that these GEMM tumors can be made immunogenic and immune protection depended on cytotoxic lymphocytes¹². Comparing gene editing in immunodeficient mice could also uncover how much of tumor editing is driven by avoiding immunosurveillance versus other selective pressures.

Immune stimulatory gene hypermethylation during immunoeediting is remarkable in the context of the increased hypomethylation of genomic DNA in cancers^{34,35}. In a recent prostate cancer study, circulating tumor cells contained more large hypomethylated genomic regions than normal prostate cells³⁶. However, some genes within these hypomethylated regions were persistently repressed in most prostate cancers, but not derepressed by a DNMT inhibitor. These persistently repressed genes were enriched for genes that participate in lipid antigen processing and presentation and cellular IFN responses that could be derepressed by inhibiting EZH2, suggesting that other epigenetic mechanisms besides DNA methylation contribute to immune gene repression. Histone modifications could inhibit gene expression even when DNA methylation is decreased and other epigenetic drugs that derepress gene expression may increase tumor immunity on their own or increase the effect of DAC.

In late tumors, the only silenced gene module that was not directly immune-related involved mitochondrial metabolism and oxidative phosphorylation. Tumor metabolic reprogramming, which choreographs the switch from oxidative phosphorylation to glycolysis, also has an immunoeediting function that supports immune evasion since it depletes glucose in the TME that is needed for T cell function³⁷. A recent study that compared melanoma cell line tumor editing in immunocompetent and deficient mice uncovered an IFN γ driven pathway of tumor immune evasion that drives metabolic reprogramming³⁸. How much metabolic reprogramming of tumors is driven by the increased metabolic needs of the tumor versus the need to suppress lymphocyte killing of tumor cells could be delineated by comparing tumor editing in immunocompetent mice versus mice deficient in functional killer lymphocytes (that is, WT versus *Prfl*^{-/-} or NSG (NOD scid gamma) mice).

Danger signals are needed to induce cytotoxic and memory T cells, both of which are critical for immune tumor control³⁹. These danger signals can be provided by IFN or inflammatory innate immune pathways and their effects are amplified by downstream activation of ISGs, inflammatory cytokines and chemokines. Although the importance of innate immune production of IFNI in promoting antitumor immunity is well established and confirmed in this study, tumor immunology studies have paid little attention to type II IFNs and inflammatory pathways as danger signals to help generate effective, long-lived anti-tumor immunity. Although the effectiveness of IFNI versus other innate immune responses in tumors has not been directly compared, inducing acute inflammation, which is a stronger danger signal for recruiting and activating immune cells, might more potently induce antitumor protection. So far, clinical attempts to activate IFNI using STING agonists have been disappointing, perhaps partly because chronic IFNI signaling promotes PD-L1 expression and reduces antigen-specific T cell trafficking, expansion, metabolic reprogramming and function^{40,41}.

IFNI signaling was the most significantly enriched among immunoeedited genes in this breast cancer GEMM and in the overlap between immunoeedited and DAC-derepressed genes, providing additional evidence that tumor IFNI contributes to immunosurveillance. IFN induction in tumors is mostly triggered by mislocalized DNA in the cytosol, which could arise from genomic DNA (released from the nucleus in some tumors because of chromosomal instability or DNA damage), mitochondrial DNA or reverse transcription of endogenous retroelements⁴². This study did not analyze cytosolic DNA, except for ERVs, which are usually excluded from RNA-seq analysis because their repetitive sequences make them difficult to map to the genome. We did not find consistent ERV activation in DAC-treated 4T1 tumors, which suggested that ERVs may not be the dominant source of DNA stimulating IFNI in newly formed breast tumors. Cytoplasmic DNA sequencing, especially using long-read techniques, could help determine the source of IFN-stimulating DNA and whether and how it is altered during tumor editing.

Among other innate immune pathways, necroptosis- and pyroptosis-associated genes were not DEG in early versus late tumors. These genes were either not expressed or only weakly expressed in both early and late GEMM tumor cells, as well as in the underlying mammary tissue (Supplementary Fig. 2), so they did not need to be edited. Even so, sensitive RT-qPCR found some key genes in these pathways were repressed in late tumors. Because the breast is not a mucosal tissue, both the normal tissue and breast tumors may not be as prone to become inflammatory as tumors arising at tissue sites constantly exposed to infection. In fact, the B16 melanoma knockout experiments (Fig. 7b–e) strongly suggest that necroptosis and pyroptosis genes and pathways play an important role in immunosurveillance of this tumor. It will be worth looking at immunoeediting of innate immune genes in the future in GEMM tumors that arise from tissues, such as the skin or colon, in which innate immunity is more active.

Most known innate immune pathways are repressed either ab initio (necroptosis, pyroptosis) or by immunoeediting (type I and II IFN) in this breast cancer GEMM. This checkpoint blockade (CPB)-resistant GEMM¹² came under immune control after treatment with low-dose DAC. Given the profound antitumor immunity induced by DAC, it will be worthwhile testing whether CPB synergizes with low-dose DAC in CPB-resistant tumors. Immunoeediting disrupts many pathways besides T cell costimulation that immune cells use to recognize and eliminate tumors. Reversing the pathways suppressed during immunoeediting would be an attractive strategy for increasing the proportion of immunotherapy-responsive cancers. Low-dose DAC potentially provides a safe way to reverse many of the immunoeediting changes that disrupt immunosurveillance. Costimulation was suppressed during immunoeediting of this GEMM but was only one of many edited pathways. Disrupting IFNI signaling by B16 tumors, by knocking out *Irf3* and *Irf7* in the tumor or using *Ifnar1*^{-/-} mice, had a modest effect on DAC efficacy, suggesting that DAC-induced other immune activities that

compensated for IFN γ signaling loss. DAC-induced IFN γ signaling and release of danger-associated molecular patterns, such as HMGB1 and ATP, from pyroptotic and necroptotic cells, could also improve dendritic cell maturation and function to improve CD8 $^+$ T cell immunity. Low-dose DAC activated two potent inducers of ICD in B16: GSDME-mediated pyroptosis and RIPK3- and MLK1-mediated necroptosis. Knocking out either *Gsdme* or *Ripk3* had a modest but significant effect on immune control of the tumor. These results, taken together, indicate that innate immune pathways are redundant for reinducing effective immune surveillance. Any of these pathways can serve as danger signals to promote lymphocyte effector function, but some are likely more potent than others and combining them (as DAC treatment does) will likely be more potent than inducing any one pathway on its own.

Recent breast cancer studies identified CD8 $^+$ TIL subpopulations that coexpress cytotoxic proteins with innate and innate-like lymphocyte receptors, including $\gamma\delta$ T cell antigen receptors (TCRs), NK activating receptors and Fc ϵ R (refs. 21,43,44). These TIL could potentially use these activating receptors rather than conventional TCRs to bypass the requirement for MHC-I and tumor-specific or -associated antigens, pathways that tumor cells immunoedit. Our CD8 $^+$ TRM TIL contained cells that express innate and innate-like receptors (Extended Data Fig. 4, Supplementary Figs. 9 and 10 and Supplementary Table 17). Cells in our CD8 $^+$ TRM TIL and *Fcer1g* $^+$ innate-like T cells with high cytotoxic potential (ILTCKs) described by Chou et al.⁴¹ both express *Fcerg1*, *Gzmb* and *Klrb1c* (encoding NK1.1) (Supplementary Methods)^{21,43}. However, CD8 $^+$ TRM TIL markers in our study, such as *Chn2*, were not expressed in ILTCKs, suggesting these may be distinct cell populations. *Chn2* $^+$ CD8 $^+$ TRM TIL in a TNBC tumor cell line mouse model were linked to CPB responsiveness in both the mouse model and patients with breast cancer²⁰. Future studies with more in-depth sequencing of sorted CD45 $^+$ cells coupled with TCR profiling and functional assays could identify subclusters of effector cells that might be novel immunotherapy targets, especially for immune desert tumors that do not express antigens recognized by TCRs⁴⁴.

The profound improvement in immune control by DAC in CPB-resistant tumors suggests that further DAC studies are warranted, even though clinical studies of DNMT inhibitors in patients with solid tumors have been discouraging⁴⁵. However, this study and other pre-clinical studies suggest that derepressing tumor gene expression requires lower concentrations than the doses given to patients. Higher doses of cytidine analogs inhibit tumor cell proliferation and induce apoptosis and also directly inhibit cytotoxic lymphocyte proliferation and effector functions⁴⁶. How much of the clinical benefit of DAC in hematological malignancy is immune-related is unknown. DAC can potentiate NK in patients with leukemia by increasing NK activating receptor expression, IFN γ production and cytotoxicity at low doses⁴⁷. However higher concentrations kill NK. One possible way to improve DAC effectiveness is to test low doses that do not cause lymphocyte cell cycle arrest or death. Other reasons for lack of efficacy could be that most trials involved small numbers of heavily pretreated patients and did not stratify patients based on gene methylation or immune properties of the tumor⁴⁸. Therefore, it is too early to reject using DNMT inhibitors for improving immunotherapy for solid tumors. Recent clinical studies suggest that DAC has antitumor efficacy when combined with other anticancer therapies for ovarian cancer, recurrent metastatic cervical cancer or non-small cell lung cancer^{45,49,50}. Patients with DAC-treated hepatocellular carcinoma developed necrotic tumor lesions with associated inflammation and CD8 $^+$ T cell infiltration⁵¹, suggesting that DAC can mobilize antitumor immunity by inducing tumor inflammation.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions

and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41590-024-01932-8>.

References

- Gonzalez, H., Hagerling, C. & Werb, Z. Roles of the immune system in cancer: from tumor initiation to metastatic progression. *Genes Dev.* **32**, 1267–1284 (2018).
- Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat. Immunol.* **3**, 991–998 (2002).
- Dunn, G. P., Old, L. J. & Schreiber, R. D. The immunobiology of cancer immunosurveillance and immunoediting. *Immunity* **21**, 137–148 (2004).
- Schreiber, R. D., Old, L. J. & Smyth, M. J. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science* **331**, 1565–1570 (2011).
- O'Donnell, J. S., Teng, M. W. L. & Smyth, M. J. Cancer immunoediting and resistance to T cell-based immunotherapy. *Nat. Rev. Clin. Oncol.* **16**, 151–167 (2019).
- Matsushita, H. et al. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* **482**, 400–404 (2012).
- DuPage, M. et al. Expression of tumour-specific antigens underlies cancer immunoediting. *Nature* **482**, 405–409 (2012).
- Chaplin, D. D. Overview of the immune response. *J. Allergy Clin. Immunol.* **125**, S3–S23 (2010).
- Chiappinelli, K. B. et al. Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell* **162**, 974–986 (2015).
- Tan, Y. et al. Somatic epigenetic silencing of RIPK3 inactivates necroptosis and contributes to chemoresistance in malignant mesothelioma. *Clin. Cancer Res.* **27**, 1200–1213 (2021).
- Turpin, J. et al. The ErbB2DeltaEx16 splice variant is a major oncogenic driver in breast cancer that promotes a pro-metastatic tumor microenvironment. *Oncogene* **35**, 6053–6064 (2016).
- Zhang, Y. et al. Immunotherapy for breast cancer using EpCAM aptamer tumor-targeted gene knockdown. *Proc. Natl Acad. Sci. USA* **118**, e2022830118 (2021).
- Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
- Cristea, S. & Polyak, K. Dissecting the mammary gland one cell at a time. *Nat. Commun.* **9**, 2473 (2018).
- Saeki, K. et al. Mammary cell gene expression atlas links epithelial cell remodeling events to breast carcinogenesis. *Commun. Biol.* **4**, 660 (2021).
- Gulati, G. S. et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405–411 (2020).
- Samarajiwa, S. A., Forster, S., Auchettl, K. & Hertzog, P. J. INTERFEROME: the database of interferon regulated genes. *Nucleic Acids Res.* **37**, D852–D857 (2009).
- Zhang, Z. et al. Gasdermin E suppresses tumour growth by activating anti-tumour immunity. *Nature* **579**, 415–420 (2020).
- Snyder, A. G. et al. Intratumoral activation of the necroptotic pathway components RIPK1 and RIPK3 potentiates antitumor immunity. *Sci. Immunol.* **4**, eaaw2004 (2019).
- Virassamy, B. et al. Intratumoral CD8 $^+$ T cells with a tissue-resident memory phenotype mediate local immunity and immune checkpoint responses in breast cancer. *Cancer Cell* **41**, 585–601.e8 (2023).
- Chou, C. et al. Programme of self-reactive innate-like T cell-mediated cancer immunity. *Nature* **605**, 139–145 (2022).
- Zheng, L. et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* **374**, abe6474 (2021).
- Ganesan, A.-P. et al. Tissue-resident memory features are linked to the magnitude of cytotoxic T cell responses in human lung cancer. *Nat. Immunol.* **18**, 940–950 (2017).

24. Schenkel, J. M. et al. Resident memory CD8 T cells trigger protective innate and adaptive immune responses. *Science* **346**, 98–101 (2014).
25. Miller, B. C. et al. Subsets of exhausted CD8⁺ T cells differentially mediate tumor control and respond to checkpoint blockade. *Nat. Immunol.* **20**, 326–336 (2019).
26. Vodnala, S. K. et al. T cell stemness and dysfunction in tumors are triggered by a common mechanism. *Science* **363**, eaau0135 (2019).
27. Morrish, E. & Ruland, J. Cytotoxic FCER1G⁺ innate-like T cells: new potential for tumour immunotherapy. *Signal Transduct. Target. Ther.* **7**, 204 (2022).
28. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
29. Gore, Y. et al. Macrophage migration inhibitory factor induces B cell survival by activation of a CD74-CD44 receptor complex. *J. Biol. Chem.* **283**, 2784–2792 (2008).
30. Cashen, A. F., Shah, A. K., Todt, L., Fisher, N. & DiPersio, J. Pharmacokinetics of decitabine administered as a 3-h infusion to patients with acute myeloid leukemia (AML) or myelodysplastic syndrome (MDS). *Cancer Chemother. Pharmacol.* **61**, 759–766 (2008).
31. Nair, A. B. & Jacob, S. A simple practice guide for dose conversion between animals and human. *J. Basic Clin. Pharm.* **7**, 27–31 (2016).
32. Iborra, S. et al. Optimal generation of tissue-resident but not circulating memory T cells during viral infection requires crosspriming by DNGR-1⁺ dendritic cells. *Immunity* **45**, 847–860 (2016).
33. Nakagawa, S. & Takahashi, M. U. gEVE: a genome-based endogenous viral element database provides comprehensive viral protein-coding sequences in mammalian genomes. *Database* **2016**, baw087 (2016).
34. Hon, G. C. et al. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res.* **22**, 246–258 (2012).
35. Johnstone, S. E. et al. Large-scale topological changes restrain malignant progression in colorectal cancer. *Cell* **182**, 1474–1489. e23 (2020).
36. Guo, H. et al. DNA hypomethylation silences anti-tumor immune genes in early prostate cancer and CTCs. *Cell* <https://doi.org/10.1016/j.cell.2023.05.028> (2023).
37. Chang, C.-H. et al. Metabolic competition in the tumor microenvironment is a driver of cancer progression. *Cell* **162**, 1229–1241 (2015).
38. Tsai, C.-H. et al. Immunoediting instructs tumor metabolic reprogramming to support immune evasion. *Cell Metab.* **35**, 118–133. e7 (2023).
39. Waldman, A. D., Fritz, J. M. & Lenardo, M. J. A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat. Rev. Immunol.* **20**, 651–668 (2020).
40. Teijaro, J. R. et al. Persistent LCMV infection is controlled by blockade of type I interferon signaling. *Science* **340**, 207–211 (2013).
41. Kuhl, N. et al. STING agonism turns human T cells into interferon-producing cells but impedes their functionality. *EMBO Rep.* **24**, e55536 (2023).
42. Motwani, M., Pesiridis, S. & Fitzgerald, K. A. DNA sensing by the cGAS-STING pathway in health and disease. *Nat. Rev. Genet.* **20**, 657–674 (2019).
43. Dadi, S. et al. Cancer immunosurveillance by tissue-resident innate lymphoid cells and innate-like T cells. *Cell* **164**, 365–377 (2016).
44. Blanco-Heredia, J. et al. Converging and evolving immuno-genomic routes toward immune escape in breast cancer. *Nat. Commun.* **15**, 1302 (2024).
45. Juergens, R. A. et al. Combination epigenetic therapy has efficacy in patients with refractory advanced non-small cell lung cancer. *Cancer Discov.* **1**, 598–607 (2011).
46. Stübiger, T. et al. 5-azacytidine promotes an inhibitory T-cell phenotype and impairs immune mediated antileukemic activity. *Mediators Inflamm.* **2014**, 418292 (2014).
47. Li, X. et al. Concentration-dependent decitabine effects on primary NK cells viability, phenotype, and function in the absence of obvious NK cells proliferation—original article. *Front. Pharmacol.* **12**, 755662 (2021).
48. Linnekamp, J. F., Butter, R., Spijker, R., Medema, J. P. & van Laarhoven, H. W. M. Clinical and biological effects of demethylating agents on solid tumours—a systematic review. *Cancer Treat. Rev.* **54**, 10–23 (2017).
49. Fang, F. et al. A phase 1 and pharmacodynamic study of decitabine in combination with carboplatin in patients with recurrent, platinum-resistant, epithelial ovarian cancer. *Cancer* **116**, 4043–4053 (2010).
50. Pohlmann, P. et al. Phase II trial of cisplatin plus decitabine, a new DNA hypomethylating agent, in patients with advanced squamous cell carcinoma of the cervix. *Am. J. Clin. Oncol.* **25**, 496–501 (2002).
51. Mei, Q. et al. An open-label, single-arm, phase I/II study of lower-dose decitabine based therapy in patients with advanced hepatocellular carcinoma. *Oncotarget* **6**, 16698–16711 (2015).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2024

¹Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA. ²Department of Pediatrics, Harvard Medical School, Boston, MA, USA. ³Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education, School of Life Sciences, Peking University, Beijing, China. ⁴Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China. ⁵Department of Pathology, Beth Israel Deaconess Medical Center, Boston, MA, USA. ⁶Harvard Medical School, Boston, MA, USA. ⁷Division of Newborn Medicine, Boston Children's Hospital, Boston, MA, USA. ⁸Spatial Technologies Unit, Beth Israel Deaconess Medical Center, Boston, MA, USA. ⁹Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. ¹⁰School of Life Science and Technology, Harbin Institute of Technology, Harbin, China. ¹¹Present address: Department of Immunology, University of Texas MD Anderson Cancer Center, Houston, TX, USA. ¹²Present address: Departments of Pediatrics and Genetics, Washington University School of Medicine in St. Louis, St. Louis, MO, USA. ¹³These authors contributed equally: Ying Zhang, Pourya Naderi Yeganeh. ✉ e-mail: ying.zhang@pku.edu.cn; white@bidmc.harvard.edu; judy.lieberman@childrens.harvard.edu

Methods

Mice

Mouse strains were maintained with a 12 h/12 h light/dark cycle in the specific pathogen-free animal facility of Harvard Medical School, where the ambient temperature was maintained at 20 °C with 50% humidity. Mice were fed with PicoLab 5058 Irradiated Rodent Diet (Scott Pharma Inc., part number 3005752-220). Tumor experiments used female mice between 6 and 8 weeks old. Mice with similar tumor volumes were randomly allocated to PBS or DAC treatment groups, except where indicated. All animal experiments were conducted in compliance with institutional ethical regulations following protocols approved by the Harvard Medical School Institutional Animal Care and Use Committee. BALB/c and C57BL/6 mice were purchased from The Jackson Laboratory. The B6.129S2-*Irfnar1*^{tm1Agt}/Mmjax (*Irfnar1*^{-/-}) mice and *Prf1*^{-/-} mice were purchased from The Jackson Laboratory and bred in house. *ErbB2ΔEx16*^{+/-} mice and *MTB*^{+/-} mice (harboring the MMTV/reverse tetracycline transactivator transgene) were generously provided by W.J. Muller (McGill University). They were cross-bred to generate *ErbB2ΔEx16*^{+/-}*MTB*^{+/-} double transgenic mice. Genotyping was performed with *ErbB2ΔEx16* (forward primer: 5'-GTGACCTGTTTGGACCGGA-3', reverse primer: 5'-TCTCCGCATCGTGTACTTCC-3') and *MTB* (forward primer: 5'-ACCGTACTCGCAATTCCAAGGG-3', reverse primer: 5'-TGCCGCCATTATTACGACAAGC-3').

Cell lines and cell culture

Human embryonic kidney 293T (HEK293T) cells were purchased from the American Type Culture Collection (ATCC CRL-3216). B16F10 cells were a kind gift from G. Freeman (Dana-Farber Cancer Institute) and were originally from ATCC (CRL-6475). 4T1 cells were provided by F. Miller (Wayne State University). B16F10-eGFP cells were generated by transducing B16F10 cells with a lentiviral vector expressing GFP (Amsbio) and sorting for GFP⁺ cells. HEK293T, B16F10, B16F10-eGFP and 4T1 cells were cultured in Dulbecco's modified Eagle medium supplemented with 10% heat-inactivated fetal bovine serum (FBS) (VWR), 6 mM HEPES, 1.6 mM L-glutamine, 50 μM 2-mercaptoethanol, 100 U ml⁻¹ penicillin G and 100 μg ml⁻¹ streptomycin sulfate (Sigma-Aldrich). All cell lines were verified to be free of mycoplasma by PCR and were authenticated by morphology.

Animal studies

For tumor challenge experiments, B16F10, B16F10-eGFP, B16F10 control, *Gsdme*^{+/-}, *Gsdmd*^{+/-}, *Ripk3*^{+/-} or *Irf3*^{+/-}*Irf7*^{+/-} cells (10⁶ cells per mouse) were injected subcutaneously into the right flank of C57BL/6 mice. 4T1 cells (5 × 10⁴ cells per mouse) were injected into the fourth mammary fat pad of BALB/c mice. For DAC or vehicle (PBS) treatment, mice bearing palpable tumors were injected intraperitoneally (i.p.) with PBS or DAC for three consecutive days per week for 2 weeks. The highest dose given to mice, 5 mg kg⁻¹, was converted from the 15 mg m⁻² per day of treatment schedule for patients based on the FDA recommendation (<https://www.fda.gov/media/72309/download>). For tumor induction in *ErbB2ΔEx16*^{+/-}*MTB*^{+/-} mice, 8-week-old double transgenic females were given 2 mg ml⁻¹ doxycycline (Sigma-Aldrich) in the drinking water throughout the study. Mice treated with doxycycline for -14 days were enrolled into PBS or DAC treatment groups, with a similar average tumor volume for each group. Two mice with extra-large tumors (~1,000 mm³) were assigned to the DAC group to assess efficacy of DAC against established large tumors. Tumor growth was measured every other day. When the largest diameter of one or more control group tumors reached 10 mm (the maximal tumor size permitted by our ethics committee), all mice in the experiment were euthanized for analysis. For propidium iodide uptake assay, tumor-bearing mice treated with PBS or DAC were injected intravenously with 2.5 mg kg⁻¹ propidium iodide in 100 μl per mouse and mice were euthanized 10 min later for analysis.

Tumor digestion and preparation of single-cell suspension

For flow cytometry and immune cell functional analyses, tumors were cut into small pieces and dissociated in RPMI medium supplemented with 2 mg ml⁻¹ collagenase D, 100 μg ml⁻¹ DNase I and 2% FBS with agitation for 30 min at 37 °C. Samples were then homogenized, filtered through 40 μm strainers and washed twice with Leibovitz's L-15 medium.

Antibody staining and flow cytometry

Immune cells isolated from mice were stained with anti-CD45-PerCPCy5.5 or -Pacific Blue, CD8-FITC, -PerCPCy5.5 or Alexa700, CD4-PerCPCy5.5 or -PE-Cy7, CD44-PerCPCy5.5 or Pacific Blue, CD69-PE-Cy7, Dx5-FITC or -Pacific Blue, NKp46-APC, Gr-1-PE, CD11b-Alexa700, CD11c-Brilliant Violet 421 or Pacific Blue, CD103-Alexa700, F4/80-PE-Cy7, MHCII-Alexa700, EpCAM-PerCPCy5.5 or PE-Cy7. For GzmB or PFN staining, cells were first stained for cell-surface markers, fixed and permeabilized with fixation and permeabilization buffer (BD BioSciences) and stained with anti-GzmB-PacBlue and PFN-PE. For intracellular cytokine staining of ex vivo stimulated lymphocytes, 10⁶ cells per sample in RPMI medium containing 2% FBS were stimulated with phorbol myristate acetate (50 ng ml⁻¹, Sigma), ionomycin (2 μg ml⁻¹, Sigma) and Golgiplug (1.5 μg ml⁻¹, BD BioSciences) for 4 h. Cells were then fixed, permeabilized and stained with antibodies to IFNγ-Pacific Blue and TNF-PE-Cy7 after. All antibodies were used at a dilution of 1:100. Data were analyzed using BD FACSCanto II and FlowJo V.10. An example of the immune cell gating strategy is shown in Supplementary Fig. 11.

RNA isolation and RT-qPCR

Total RNA was extracted with TRIzol and a Direct-zol RNA miniprep kit (Zymo). RNA concentrations were quantified using a NanoDrop 2000 Spectrophotometer. Complementary DNA (cDNA) was synthesized using the High-capacity cDNA reverse transcription kit and cDNAs were amplified with primers corresponding to the target or housekeeping gene *Gapdh* (Supplementary Table 18), SsoFast EvaGreen Supermix and a Bio-Rad C1000 Thermal Cycler.

Generating gene knockout cell lines

B16F10 *Gsdme*^{+/-} or *Ripk3*^{+/-} cells were generated using the Edit-R pre-designed lentiviral single-guide RNA (sgRNA) system (Horizon). Cells stably expressing the sgRNA were selected with puromycin (InvivoGen) and were transiently transfected with an Alt-R S.p. Cas9 Expression Plasmid (IDT) using nucleofection (Lonza). Single clones were established by limiting dilution. B16F10 *Gsdmd*^{+/-} or *Irf3*^{+/-}*Irf7*^{+/-} cells were generated using the pGuide-it-ZsGreen1 system following the manufacturer's protocol (Takara). For each target gene, at least two knockout clones were tested for experiments with similar results. The sgRNA target sequences used for data shown in this paper are: *Gsdme*: TGTGAGTACATCTCCAGGG; *Ripk3*: TCCAGTGGGACTTCGTGTCC; *Gsdmd*: CAGCAGAGGCGATCTCATTCCGG; *Irf3*: GCATGGAAACCCCGAAACCG; *Irf7*: CTACGACCGAAATGCTTCAGGG.

Immunoblot

Cells were lysed in RIPA buffer supplemented with Halt protease and phosphatase inhibitor cocktail (100×, Thermo Fisher Scientific). Tumor tissues were flash frozen in liquid nitrogen before homogenization with RIPA. Protein lysates were analyzed by SDS-PAGE as previously described¹⁸. Primary antibodies used were anti-GSDME (cat. no. ab215191, Abcam, 1:1,000), anti-GSDMD (cat. no. ab209845, Abcam), anti-RIPK3 (cat. no. 2283, ProSci), anti-MLKL (cat. no. 37705s, Cell Signaling), anti-phospho-MLKL (cat. no. 37333s, Cell Signaling), anti-IRF3 (cat. no. 4302s, Cell Signaling), anti-caspase 3 (cat. no. 9662s, Cell Signaling) and anti-β-actin (cat. no. JLA20, Developmental Studies Hybridoma Bank). All primary antibodies were used at a dilution of 1:1,000.

Immunofluorescent imaging

Tumors were fixed with 10% formalin for 24 h and stored in 70% ethanol at 4 °C. Fixed tissues were embedded in paraffin and cut into 5 µm sections by the Dana-Farber Cancer Institute Rodent Histopathology Core. Sections were blocked with 5% normal donkey serum and stained with anti-CD8α (1:300, 4SM15, Thermo Fisher Scientific) and anti-EpCAM (1:300, 21050-1-AP, Proteintech) primary antibodies overnight. Secondary antibodies were donkey antirat Alexa Fluor 488 and donkey antirabbit Alexa Fluor 568 diluted 1:250 (Thermo Fisher Scientific). Sections were counterstained with Hoechst 33342 (Thermo Fisher Scientific) and imaged with a Zeiss LSM 880 Confocal laser scanning microscope. Images were analyzed by ZEN Black and ImageJ.

Isolation of CD45⁺EpCAM⁺ tumor cells from GEMM

Tumor samples were digested in RPMI medium containing 100 µg ml⁻¹ Liberase, 100 µg ml⁻¹ DNase I (Merck Millipore) and 2% v/v FBS with shaking at 200 rpm for 30 min at 37 °C. Red blood cells were removed by 1× ACK Buffer (155 mM NH₄Cl, 10 mM KHCO₃ and 0.1 mM EDTA). Dead cells were removed by MojoSort mouse dead cell removal kit (BioLegend). Live cells were enriched for CD45⁺ cells and then EpCAM⁺ cells with CD45 and EpCAM microbeads, respectively (Miltenyi Biotec). Enriched cells were stained for CD45-Alexa700, EpCAM-PE-Cy7, CD8-APC (BioLegend) and then sorted using a BD Aria III sorter. The first 5,000 EpCAM⁺CD45⁺CD8⁻ sorted cells were used to assess cell purity. Once purity (>90%) had been confirmed, subsequent sorted cells were directly sorted into 5 ml of RNA protect reagent (Qiagen) and ~1 × 10⁵ EpCAM⁺CD45⁺CD8⁻ cells were immediately subjected to RNA extraction using RNeasy Plus Micro Kit (Qiagen) for RT-qPCR assays.

scRNA-seq sample and library preparation

Tumors induced in 8-week-old *ErbB2ΔEx16^{+/+}MTB^{+/+}* female mice were collected in two batches. The first batch contained two early tumors and two late tumors, each from one mouse. The second batch contained two early and two late tumor samples, with each sample of early tumors pooled from three mice. Live cells from tumors were enriched twice with the dead cell removal kit. Single-cell suspensions with more than 80% viability were resuspended in Ca²⁺ and Mg²⁺-free PBS with 0.5% BSA and kept on ice. Cell concentration was adjusted to 1.4 × 10⁶ cells per ml and 10 ml was loaded onto a full skirted 96-well Eppendorf plate. The 10× Genomics Chromium Connect automated platform was used for automated single-cell library construction using the Chromium Next GEM Single Cell 5' Kit v.2 (10X Genomics, cat. no. PN-1000263), using 14 cDNA and 16 SI-PCR cycles. Approximately 14,000 live cells per sample were loaded into the system, with an expectation to encapsulate at least 8,000 cells per sample. Indexing was performed on the same instrument using the Dual Index Kit TT Set A, 96 rxns (10X Genomics, cat. no. PN-1000215). Gel beads-in-emulsion generation was assessed after 45 min and was found to be optimal for all samples. cDNA and gene expression libraries were quantified using a Qubit 3.0 fluorometer, and quality was assessed using HS DNA chips and an Agilent 2100 Bioanalyzer. cDNA quality control and final library quality control were performed using the Fragment Analyzer (Agilent Technologies) NGS Fragment Kit (1–6,000 basepairs) 500 (DNF-473-0500) to obtain library traces and concentrations. Libraries were sequenced at the Bauer core facility at Harvard University using the NovaSeq 6000 S4 platform (Illumina), one lane per batch of samples. Roughly 220,000 reads were obtained per cell, capturing the expression of ~2,300 transcripts per cell.

scRNA-seq preprocessing and initial quality control

scRNA-seq reads were mapped to a customized genome using the cellranger v.7.0.0 suite and an extended version of the mouse genome (mm10, GENCODE vM23/Ensembl 98) with an added eGFP sequence,

retrieved from the National Center for Biotechnology Information (NCBI portal) (<https://www.ncbi.nlm.nih.gov/nucore/JQ064510.1>). The Seurat package was used to process an initial set of 29,861 cells from the scRNA-seq data¹³. Genes with ≤30 reads across all cells were removed from analysis. Cells were removed from analysis if they had: <500 detected genes, <1,000 reads, mitochondrial content >25% or ribosomal content <1%.

scRNA-seq clustering

The scRNA-seq expression matrix was scaled to 10,000 reads per cell and log-normalized using Seurat¹³. The most variable genes ($n = 3,000$) were determined using variance stabilizing transformation. Principal component analysis was performed on all genes. Batch effects were removed from the principal components using Harmony⁵². Cells were clustered using the Louvain algorithm on a 20-nearest neighbor graph from the top 30 Harmony embeddings (resolution parameter of 2)⁵³. Clusters were projected onto two dimensions using uniform manifold approximation and projection (UMAP) using ten nearest neighbors, and a minimum distance of 0.3 (ref. 54). Outlier clusters were identified and discarded by manual inspection of UMAP projections.

scRNA-seq cell annotation

A final set of 40 clusters was used for cell annotation. Known cell markers were used to define cluster labels as detailed in Fig. 1. Markers of unsupervised clusters were determined separately for immune and epithelial cells using differential gene expression analysis. Genes in each cluster were compared against all other subclusters of the same type (for example, immune cells) for differential expression using a generalized linear model, MAST, implemented in Seurat^{13,55}. Positive and negative markers of each cluster were determined as genes with significantly upregulated or downregulated expression (FDR < 0.05 and logFC > 0.25). Genes with fewer than a total of 20 reads in broad cell classes (for example, epithelial) were excluded. Ribosomal- and mitochondrial-associated genes were excluded from marker identification. Immune cells were categorized as B cells (*Cd79⁺*), T cells (*Cd3d⁺*), NK cells (*Ncr1⁺*), macrophages (*Cd14⁺*, *Lyz2⁺*, *Cd74⁺*) and neutrophils (*S100a9⁺*, *Cd14⁺*). Subsets of T cells were defined as: Cd4⁺ (*Cd4⁺*); Cd4⁺ T_{reg} (*Cd4⁺Foxp3⁺*); Cd8⁺ (*Cd8a⁺*); Naive (*Ccr7⁺*, *Cd62l⁺* (*Sell*) and *Cd127⁺* (*Il7r*)), central memory (*Ccr7⁺*, *Cd62l⁺* and *Cd44⁺*), ISG (*Isg15⁺*, *Ifit1⁺* and *Irf7⁺*), Effector (*Ccr7⁺*, *Cd62l⁺*, *Cd127⁺* and *Cd44⁺*), TRM (*Cd103⁺* (*Iltgae*) and *Fcer1g⁺*), Progenitor Exhausted (*Tcf7⁺*, *Bach2⁺* and *Pdcd1⁺*), Exhausted (*Tcf7⁺* and *Pdcd1⁺*) (Supplementary Table 19) and proliferating (*Mki67⁺*). Mammary epithelial cells were defined as in Saeki et al. and Gray et al.^{15,56}: luminal-alveolar (Alv1-4; *Csn3⁺* and *Elf5⁺*), alveolar proliferating (*Mki67⁺*), basal (*Krt14⁺*) alveolar-basal-luminal cells (Alv-basal, *Aldh1a3⁺*) and hormone sensing (*Esr1⁺*, *Krt19⁺* and *Prlr⁺*). Annotated scRNA-seq expression profiles were deposited in a public R Shiny application using the ShinyCell package⁵⁷.

Cell differentiation status was determined using CytoTRACE, applied to all EpCAM⁺, HER2⁺ and eGFP⁺ epithelial cells, using raw counts and default parameters¹⁶. Clusters with lower maturation scores were defined as tumor cells.

Differential gene expression analysis of scRNA-seq data

DEGs in late versus early samples were determined for each cell subtype using generalized linear models, MAST⁵⁵. Additionally, DEG were determined between late and early samples in pooled tumor cells (Alv-basal, Alv-proliferating and basal) as one group. All comparisons were adjusted for the number of reads per cell (nCount_RNA)⁵⁵. DEG were determined using FDR < 0.05 and |logFC| > 0.25 cut-offs. Genes with <20 reads in broad cell classes (for example, immune) were excluded. Ribosomal and mitochondrial genes were excluded from analysis. Additional DEG analysis was performed comparing late and early samples by pooling cell subtypes into overall CD8⁺ and CD4⁺ T cell groups.

Functional characterization of scRNA-seq data

DEGs in pooled tumor cells were characterized separately for upregulated and downregulated genes. Enrichment analysis was performed using a Fisher's exact test on a background of all genes with at least 20 reads across HER2⁺ epithelial cells using clusterProfiler⁵⁸. Enrichment analysis was performed separately on GO Biological Process terms and as Hallmark gene sets from the MSigDB⁵⁹. *P* values were adjusted for multiple comparisons using the *q* value method⁶⁰. A cut-off of $q < 0.05$ was used to determine significantly enriched terms. To find broad classes of GO functions, we used EnrichmentMap, implemented as a plugin in the Cytoscape suite^{61–63}. EnrichmentMap determines similarity of GO terms based on shared genes and constructs a network where nodes represent GO terms and edges represent the Jaccard Index of shared genes between nodes. A cut-off of >0.375 to define edges and Markov cluster algorithm were used to find highly similar groups of GO terms. Isolated nodes, that is, without any edges, were excluded from analysis.

Cell type proportion analysis

Changes in cell type abundance between early and late samples were determined using Dirichlet multinomial regression using the DirichletReg R package as described in refs. 64,65. Samples with a small number of cells were removed from analysis. Dirichlet regression cannot directly work with zero frequency categories⁶⁶. To overcome this constraint, a customized normalization was used to assign a small constant frequency to cell types with zero observations in individual samples. Cell types with differential proportions in early versus late samples were determined using an adjusted *P* value cut-off (FDR < 0.05).

Cell–cell communication

Inferred cell–cell communications were analyzed using CellChat²⁸. Initial analysis was performed on cell subtypes in early and late samples separately. Fibroblasts, undefined cells and Alv4 were excluded due to insufficient numbers in early or late categories. The mouse database of CellChat (CellChatDB.mouse), including interactions from 'secreted signaling', 'ECM-receptor' and 'cell–cell contact', was used for analysis. The standard protocol of CellChat was applied to normalized scRNA-seq counts using the following functions 'identifyOverExpressedGenes', 'identifyOverExpressedInteractions', 'computeCommunProb', 'filterCommunication' and 'computeCommunProbPathway'. Cell–cell communication objects of late and early samples were merged for comparison. The CellChat analysis also compared inferred interactions between pooled tumor cells, CD8⁺ T cells and myeloid cells in early and late samples.

Bulk RNA sequencing sample preparation

Here, 4T1 cells were treated with either PBS or 1 μ M DAC for 48 h. Dead cells were removed from each sample using a dead cell removal kit (Miltenyi Biotec). The same batch of cells were also used for ATAC-seq and RRBS data analysis (Supplementary Methods). Total RNA was extracted from live cells using TRIzol and Direct-zol RNA miniprep kit (Zymo). Three biological replicates per condition were used for RNA sequencing library preparations. All RNA samples had RNA integrity number >9 as measured by an Agilent 2100 Bioanalyzer. mRNA was enriched using NEXTFLEX Poly(A) beads 2.0 and standard mRNA libraries were prepared with NEXTFLEX Rapid Directional RNA-Seq Kit v.2.0. quality control and mapping of RNA-seq in Supplementary Methods.

Differential gene expression analysis of bulk RNA-seq data

DEGs were determined using DESeq2 R package⁶⁷. Genes with fewer than ten reads across all samples were excluded. *P* values were adjusted for multiple comparisons using the FDR method⁶⁸. DEG in 4T1 cells were determined using FDR < 0.0001 and $|\logFC| > 1$ cut-offs.

Functional characterization of bulk RNA-seq data

Enrichment analysis was performed using the clusterProfiler package for upregulated genes in DAC-treated cells compared to controls⁵⁸. Enrichment analysis was performed on GO Biological Process terms. Significance of GO terms was evaluated as described above. To find broad classes of GO functions, we used EnrichmentMap, applied to the top 150 GO terms, as described above with a cut-off of >0.5 to define edges⁶².

Comparative analysis of bulk and scRNA-seq

Significantly upregulated genes in DAC-treated 4T1 cells from RNA-seq data were compared to genes that were significantly downregulated in pooled late tumors compared to early tumors from scRNA-seq data. Overlap between bulk and scRNA-seq datasets was assessed using Fisher's exact test. GO enrichment analysis of the overlapping genes was performed as described above. Enrichment analysis was also performed on genes that were upregulated in DAC-treated 4T1 cells but were not downregulated in late pooled tumors from scRNA-seq.

Statistics

No statistical methods were used to predetermine sample sizes but our sample sizes are similar to those reported in previous publications^{12,18}. Data collection and analysis were not blinded. No animal or data point was excluded from the analyses. Methods of statistical tests were chosen based on normality test on each set of data. Differences between two groups were calculated by Student's *t*-test or Mann–Whitney *U*-test unless noted in figure legends. Multiple comparisons between two groups were performed by multiple *t*-test with type I error correction or Welch's *t*-test. One-way ANOVA or Kruskal–Wallis test was used to calculate differences among multiple populations with one independent variable. Two-way ANOVA was used for experiments with two independent variables. Differences between tumor growth curves were compared by first calculating the area-under-the-curve values for each sample and then comparing different groups using the Student's *t*-test or one-way ANOVA. Type I errors were corrected by Holm–Sidak or Dunn's multiple-comparisons test or FDR calculated by the two-stage step-up method of Benjamini, Krieger and Yekutieli⁶⁹ unless noted otherwise in figure legends. A nonparametric test was used when data did not follow a normal distribution. All statistical tests used were two-sided unless indicated. For scRNA-seq data, significance cut-off was set at FDR adjusted $P \leq 0.05$, *** denotes FDR ≤ 0.05 . Statistical methods for analyzing single-cell and bulk RNA-seq, ATAC-seq and RRBS-seq are described above and in Supplementary Methods. For all other figures, Significance was set at $P \leq 0.05$, * $P \leq 0.05$, * $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. Statistical analyses were performed using GraphPad Prism v.10.

Reporting summary

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

Data availability

RNA-seq, ATAC-seq and RRBS datasets have been deposited in the Gene Expression Omnibus (GEO) under the accession code [GSE212029](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE212029). The scRNA-seq data (raw and processed) are available on the GEO portal of the NCBI GEO under the accession code [GSE234917](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE234917). Additional data accessed for this paper include an atlas of 50,000 normal murine mammary epithelial cells (accessed through online portal: <https://mouse-mammary-epithelium-integrated.cells.ucsc.edu>), a single-cell dataset from breast tumor tissues of MMTV-PyMT mice (using GEO Accession ID [GSE195937](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE195937)), a list of known ERV genes from the gVEV database (<http://geve.med.u-tokai.ac.jp/>, accessed December 2021) and a list of ISGs from the Interferome database (<https://interferome.org/>, accessed September 2022). A ShinyApp for interactive exploration of the scRNA-seq is available at https://pouryany.shinyapps.io/scgemm_app/. Source data are provided with this paper.

Code availability

Analyses presented in this study and codes to generate the figures are publicly available at: https://github.com/pouryany/scRNAseq_GEMM.

References

52. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
53. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).
54. McInnes, L. et al. (UMAP: Uniform Manifold Approximation and Projection). *J. Open Source Softw.* <https://doi.org/10.21105/joss.00861> (2018).
55. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015).
56. Gray, G. K. et al. A human breast atlas integrating single-cell proteomics and transcriptomics. *Dev. Cell* **57**, 1400–1420.e7 (2022).
57. Ouyang, J. F., Kamaraj, U. S., Cao, E. Y. & Rackham, O. J. L. ShinyCell: simple and sharable visualization of single-cell gene expression data. *Bioinformatics* **37**, 3374–3376 (2021).
58. Wu, T. et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).
59. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
60. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100**, 9440–9445 (2003).
61. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* **5**, e13984 (2010).
62. Reimand, J. et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.* **14**, 482–517 (2019).
63. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
64. Maier, M. DirichletReg: Dirichlet Regression for Compositional Data in R. in *Research Report Series / Department of Statistics and Mathematics*, no. 125 <https://doi.org/10.57938/ad3142d3-2fcd-4c37-aec6-8e0bd7d077e1> (Vienna University of Economics and Business, 2014).
65. Smillie, C. S. et al. Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. *Cell* **178**, 714–730.e22 (2019).
66. Douma, J. C. & Weedon, J. T. Analysing continuous proportions in ecology and evolution: a practical introduction to beta and Dirichlet regression. *Methods Ecol. Evol.* **10**, 1412–1430 (2019).
67. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
68. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).
69. Benjamini, Y., Krieger, A. M. & Yekutieli, D. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**, 491–507 (2006).

Acknowledgements

This work was supported by US NIH grant nos. K99 CA255841 (Y.Z.), NIH DP2 AG055947 (E.L.G.), Charles King fellowship (S.Y.W.), Cancer Research Institute fellowship (H.Z.) and NIH R01 CA240955 (J.L.). We thank W. J. Muller (Department of Biochemistry, McGill University) for *ErbB2ΔEx16^{-/-}* and *MTB^{-/-}* mice, I. Vlachos and the Spatial Technologies Unit at the Harvard Initiative for RNA Medicine for help with scRNA-seq studies, and members of the Lieberman laboratory for helpful discussions.

Author contributions

Y.Z. and J.L. conceived the study. Y.Z., P.N.Y., W.H. and J.L. designed experiments and wrote the manuscript. Y.Z., S.Y.W., H. Z., B.G., Z.L., D.-J.L., Z.Z., A.P. and M.S. performed all experiments. Y.Z., P.N.Y., S.Y.W., E.L.G., H.W. and W.H. performed data analysis. J.L. and W.H. supervised experiments and data analysis.

Competing interests

The authors declare no competing interests.

Additional information

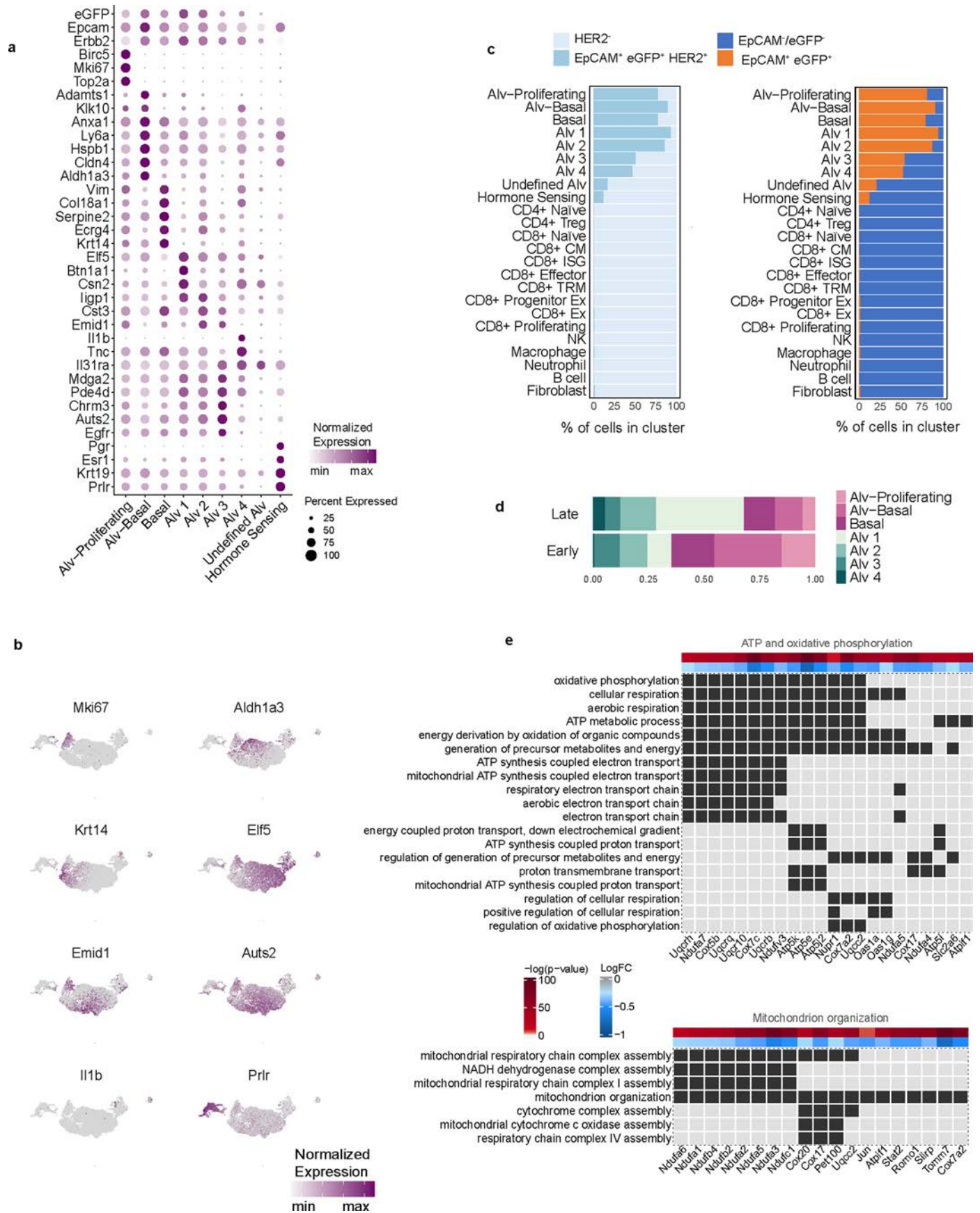
Extended data is available for this paper at <https://doi.org/10.1038/s41590-024-01932-8>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41590-024-01932-8>.

Correspondence and requests for materials should be addressed to Ying Zhang, Winston Hide or Judy Lieberman.

Peer review information *Nature Immunology* thanks Drew Pardoll and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Nick Bernard was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

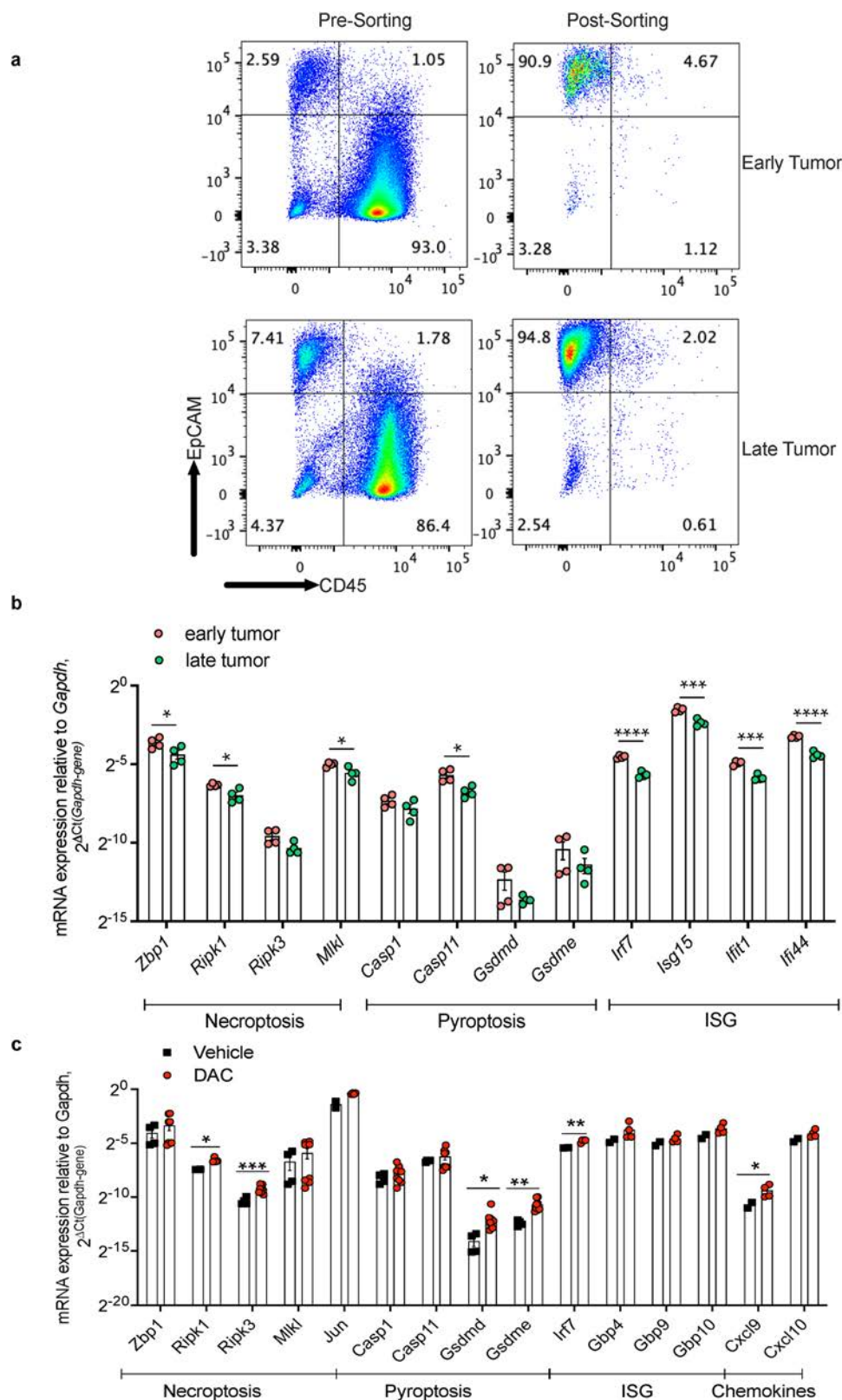
Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Tumor cell clustering and characterization. a, Dot plot showing expression of selected markers for the 9 EpCAM⁺ epithelial cell clusters. Key markers for each cluster are shown in dark purple. **b**, UMAP plots showing expression of key markers in epithelial cells. **c**, Percentage of *ErbB2*⁺ (left) and *EpCAM*⁺*eGFP*⁺ cells (right) in each cell cluster. **d**, Relative proportion of EpCAM⁺ epithelial cell subclusters in early and late tumors. **e**, Assignment of individual differentially expressed genes associated with ATP production/

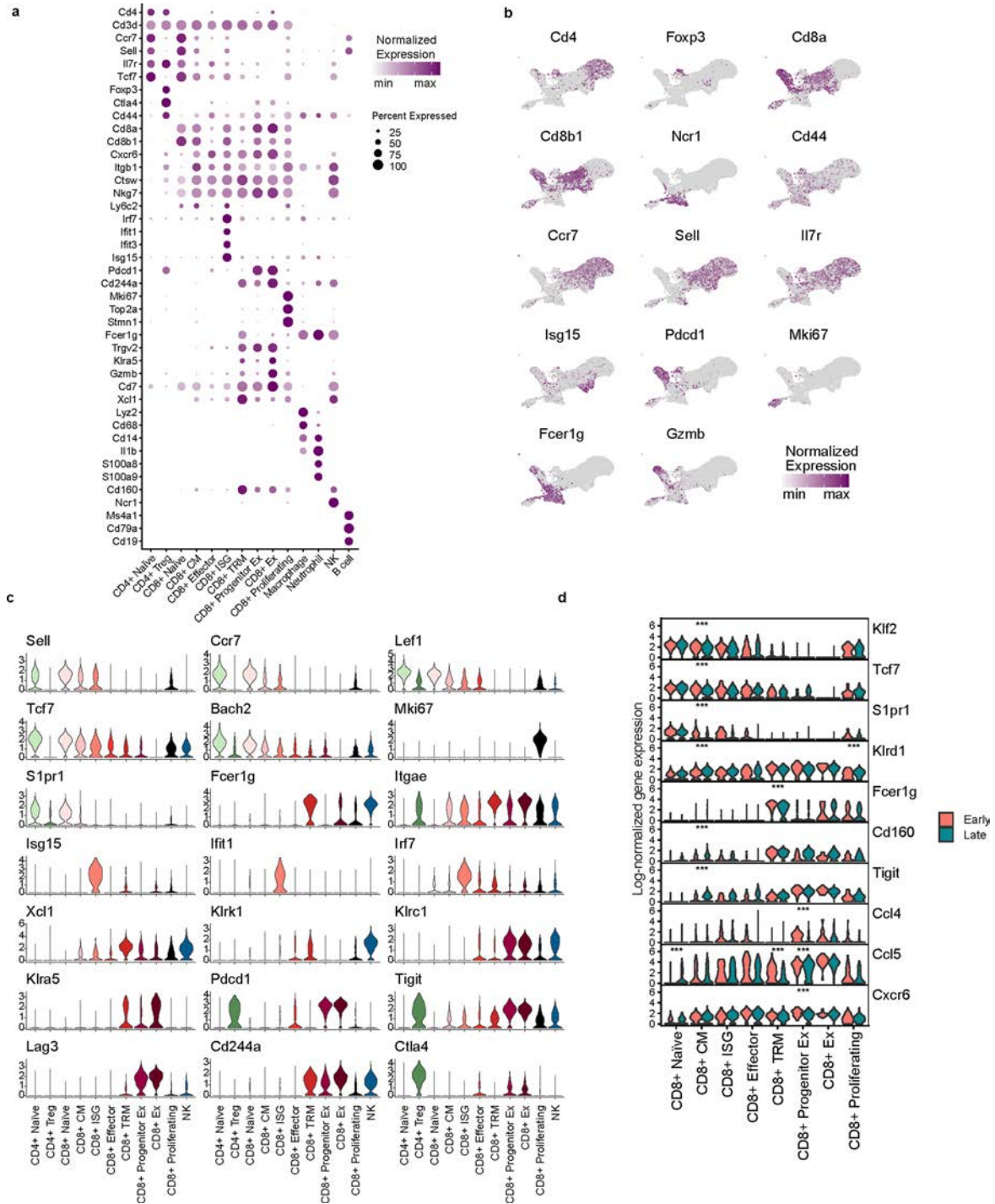
oxidative phosphorylation (top) and mitochondrial organization (bottom) to EnrichmentMap clusters (Fig. 2). Black squares represent membership in cluster. P-value and logFC correspond to comparison of late tumors versus early tumors, identified using two-sided generalized linear models (MAST test). Rows denote significant GO terms associated with differentially expressed genes in columns. P-values are provided in Source Data Extended Data Fig. 1.



Extended Data Fig. 2 | Key innate immune gene expression changes in GEMM tumor cells during tumorigenesis and following DAC treatment.

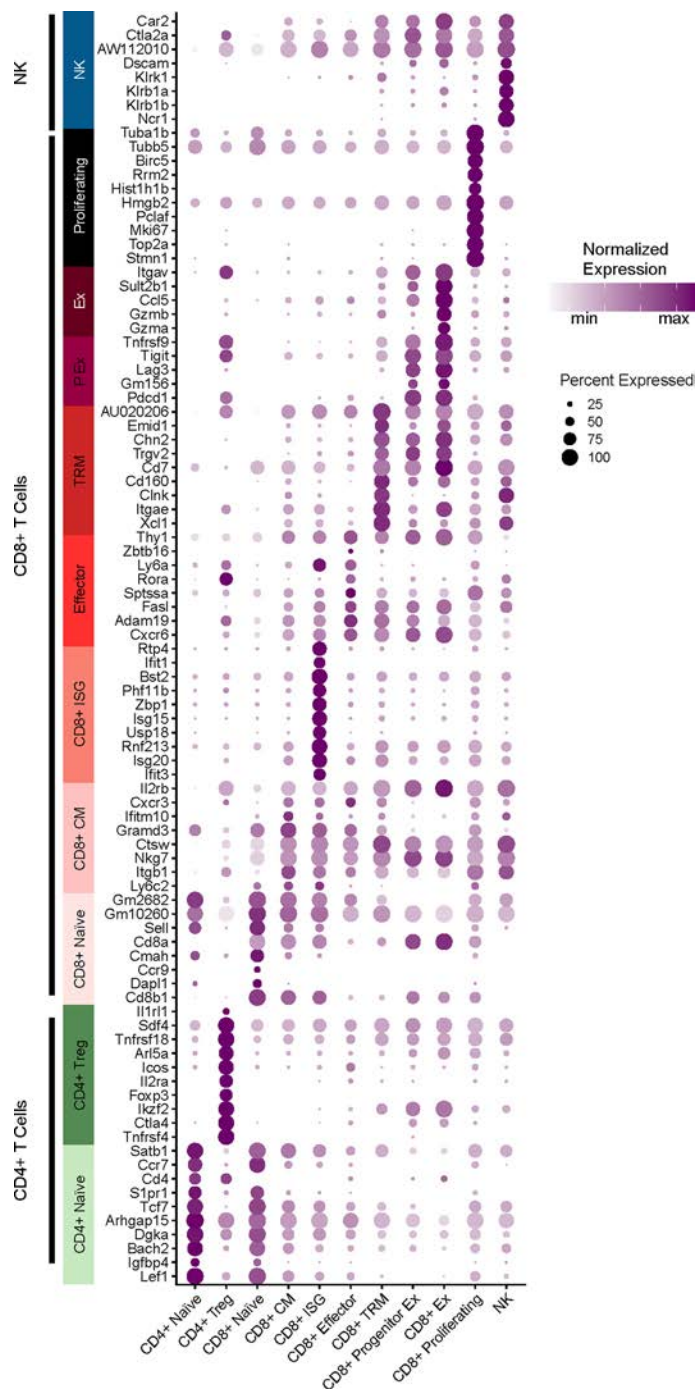
a, Representative flow plots of CD45⁺EpCAM⁺ tumor cells before and after microbead enrichment and sorting. **b**, qRT-PCR analysis of mRNA expression of genes involved in inflammatory cell death and interferon signaling, relative to *Gapdh*, in CD45⁺EpCAM⁺eGFP⁺ tumor cells sorted from early or late GEMM tumors. *n* = 4 samples/group, each early sample was pooled from three mice. **c**, qRT-PCR analysis of mRNA expression of genes involved in inflammatory cell

death and interferon signaling, relative to *Gapdh*, in CD45⁺EpCAM⁺eGFP⁺ tumor cells sorted from GEMM tumors treated with PBS (Vehicle) or DAC for 2 cycles. The effect of DAC treatment on tumor growth and tumor-infiltrating immune cells is shown in Fig. 5. Vehicle: *n* = 2–4; DAC: *n* = 4–8. Data shown are mean \pm s.e.m. values on a log scale and are representative of at least two experiments. (**b–c**) Welch's *t*-test. **p* \leq 0.05, ***p* \leq 0.01, ****p* \leq 0.001, *****p* \leq 0.0001. Exact *p*-values are provided in Source Data.



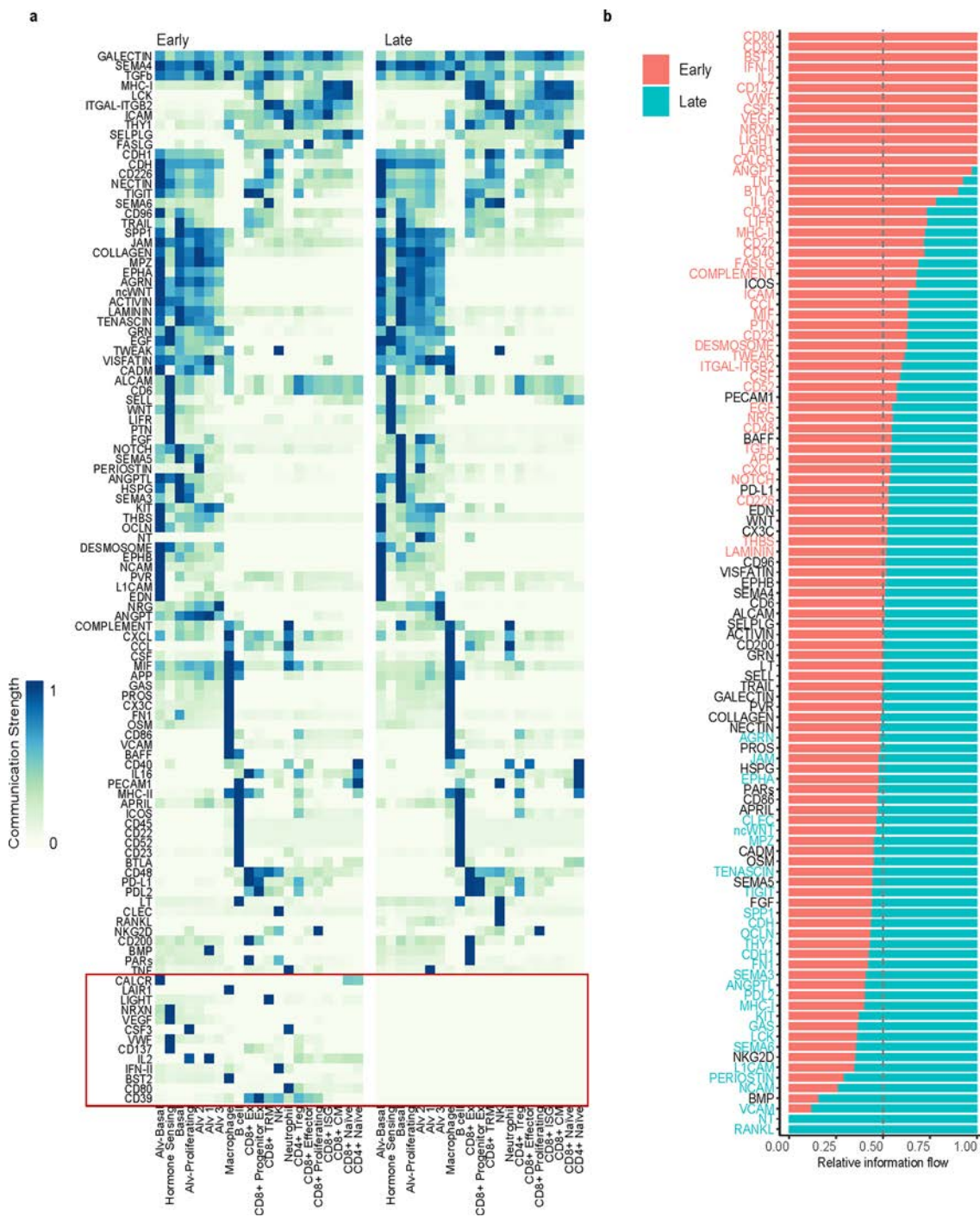
Extended Data Fig. 3 | Immune cell clustering and characterization. **a**, Bubble plot showing expression of selected markers in CD45⁺ immune cell clusters. Key markers for each cluster are shown in dark purple. **b**, UMAP plots show expression of key markers for T cell subsets and NK. **c**, Violin plots of normalized expression of key markers of T cell subsets and NK cells. The y-axis represents log-normalized gene expression, based on scaling and log transforming scRNA-seq read counts in each cell. **d**, Comparison of expression of selected markers in tumor-infiltrating CD8⁺ T cell subsets in early vs late tumors, including

markers of differentiation (*Klf2*), adhesion (*S1pr1*), transcription factors (*Tcf7*), functional effectors (*Klrd1*, *Fcer1g*), exhaustion (*CD160*, *Tigit*), chemokine and chemokine receptors (*Ccl4*, *Ccl5*, *Cxcr6*). The y-axis represents log-normalized gene expression, based on scaling and log transforming scRNA-seq read counts in each cell. (**b-d**) Generalized linear models (MAST test) were used to determine two-sided p-values of differential expression. *** denotes FDR adjusted $p < 0.05$ and $|\text{LogFC}| > 0.25$. p-values are provided in Source Data Extended Data Fig. 3.



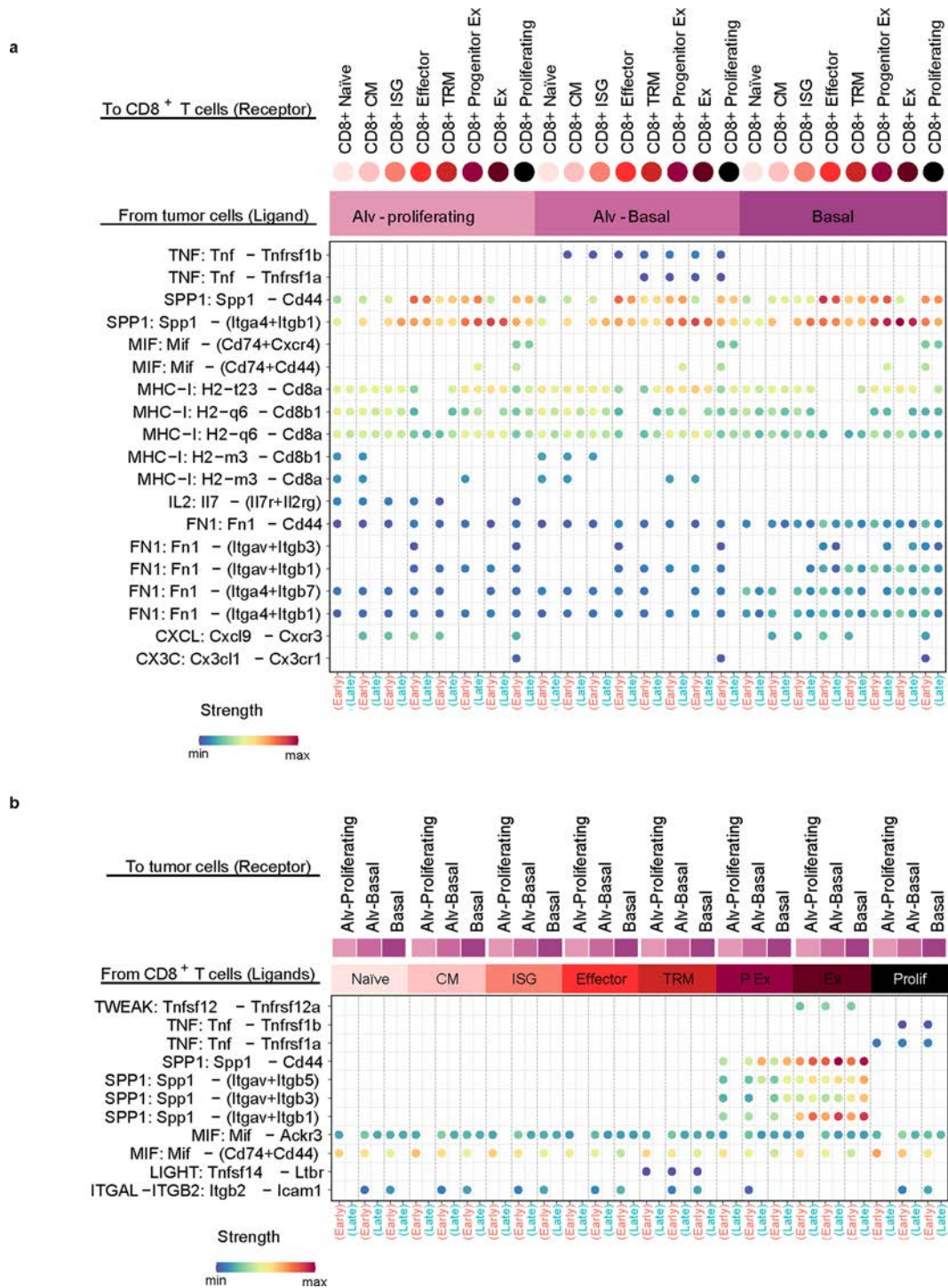
Extended Data Fig. 4 | Expression of top marker genes for each T and NK cell cluster. Bubble plot shows the top 10 marker genes for each T cell subset and NK. Marker expression was compared to that in all other lymphocyte populations (adjusted $P < 0.05$ and $|\text{LogFC}| > 0.25$). Top marker genes were selected based on adjusted p-value rankings. Genes that mark more than one population are displayed only once (for example, *Cd8b1* for CD8⁺ Naïve

and CM; *Tigit* for Progenitor Exhausted and Exhausted). P Ex: Progenitor Exhausted. Ex: Exhausted. A full list of genes with associated statistics is available in Supplementary Table 10. Generalized linear models (MAST test) were used to determine two-sided p-values of differential expression for marker identification. P-values were adjusted for multiple comparisons using the FDR method. p-values are provided in Source Data Extended Data Fig. 4.



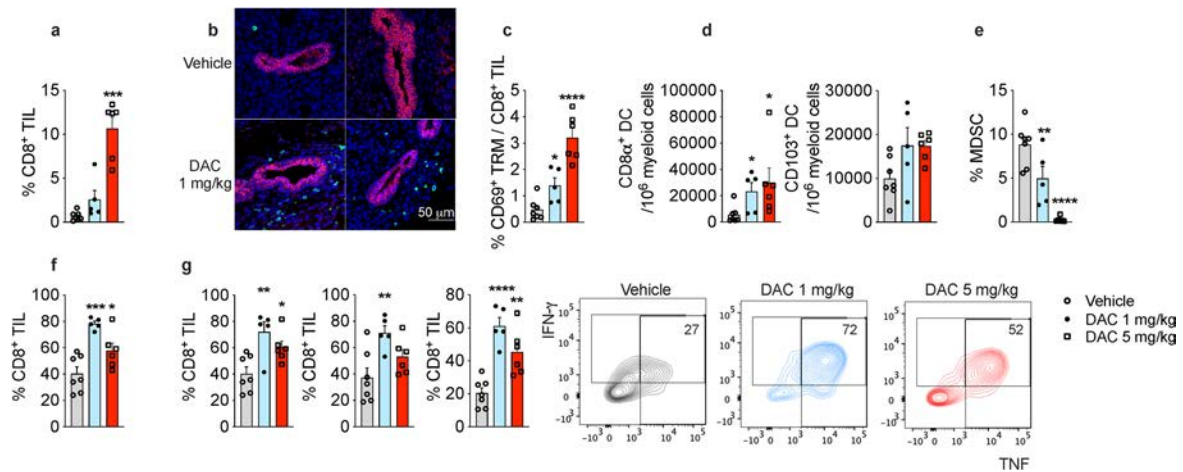
Extended Data Fig. 5 | Activity of cell-cell communication pathways in early and late tumors. Inferred cell-cell communication pathways were identified and evaluated using the CellChat algorithm. **a**, Heatmap of 107 pathways with significantly active communication in early or late samples. Red box denotes pathways silenced in late tumors. Rows denote communication pathways. Columns denote cell types communicating as either the sender (expressing ligands) or receiver (expressing receptors) of pathways. Communication strength is related to co-expression of ligand-receptor gene pairs.

Non-zero strength values denote significant communications, identified using the CellChat one-sided permutation test ($p < 0.05$). **b**, Relative changes in inferred communication between early and late tumors. Colored rows denote significant overall changes in information flow. 44 pathways had significantly higher activity in early samples (red), 29 had significantly higher activity in late samples (aqua) ($p < 0.05$). Statistical difference was evaluated by Wilcoxon test, comparing the pathway changes between late and early samples. p-values are provided in Source Data Extended Data Fig. 5.



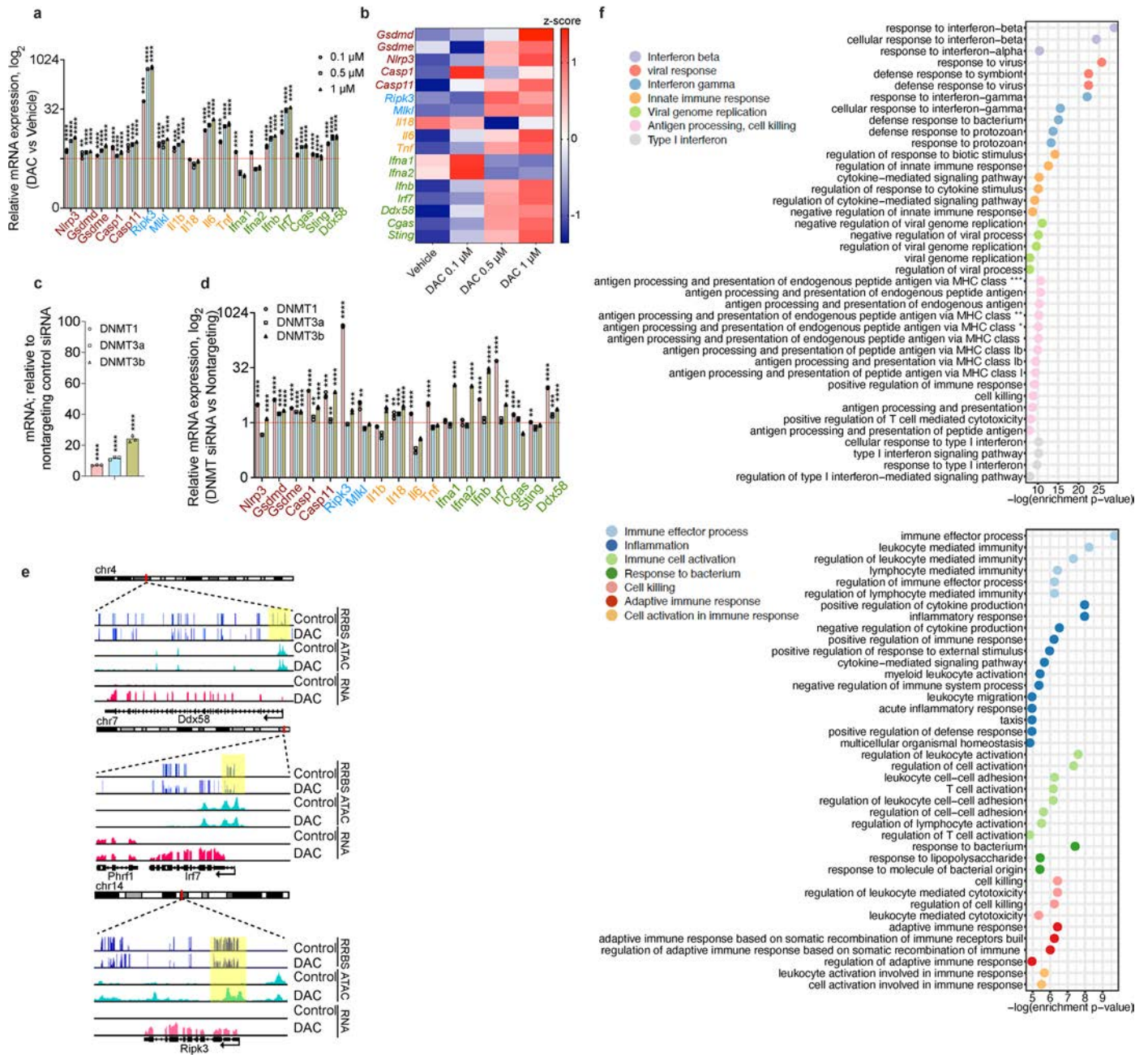
Extended Data Fig. 6 | Inferred cell-cell communication events lost in late tumors. **a**, Selected inferred cell-cell communications from subtypes of tumor cells (expressing ligands) to subtypes of CD8⁺ T cells (expressing receptors) that show loss of ligand-receptor interactions in late samples. **b**, Cell-cell communication from CD8⁺ T cells (expressing ligands) to tumor cells (expressing receptors). Row names represent pathways of specific ligand and receptor combinations.

Dots represent significant (p-value < 0.05) inferred communications corresponding to communication pathways (rows) and respective cells (columns). Communication strength is related to co-expression of ligand-receptor pairs. P-values represent the significance of the communication strength computed using the CellChat one-sided permutation test.



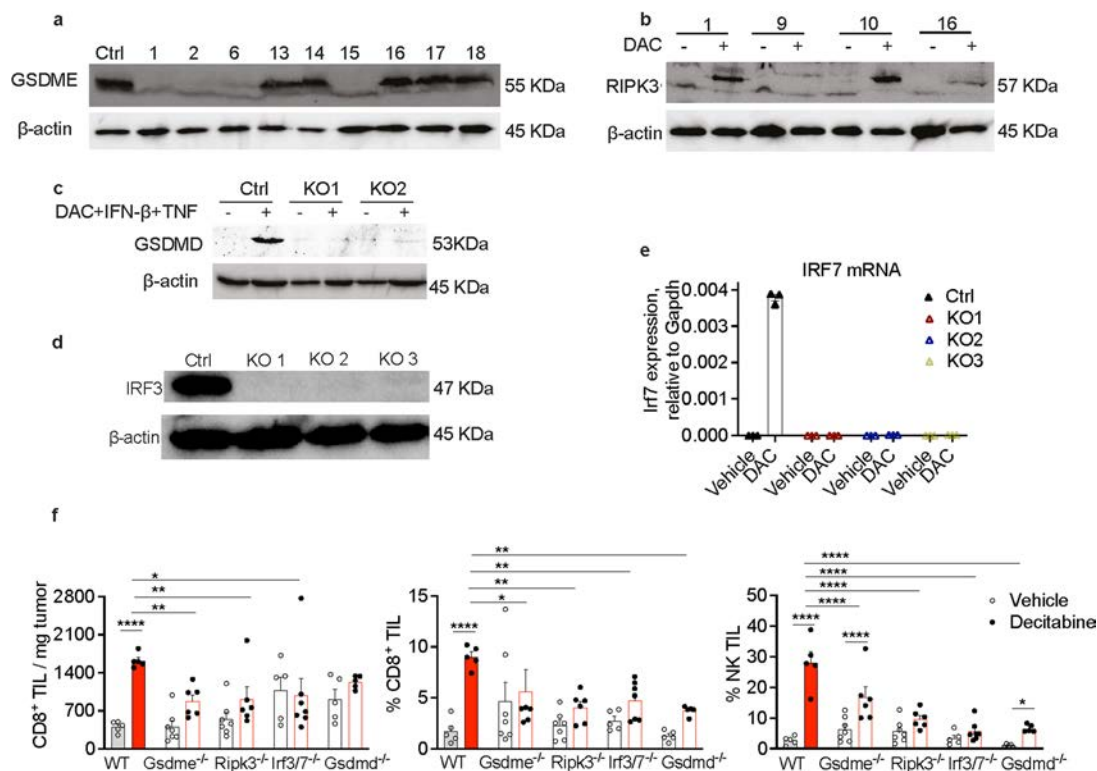
Extended Data Fig. 7 | Effect of DAC on immune response to orthotopic 4T1 tumors. **a**, Percentage of CD8⁺ TIL in tumors from mice treated with vehicle (PBS) or DAC. **b**, Immunofluorescent staining of CD8⁺ TIL in 4T1 orthotopic tumors in mice treated with two cycles of vehicle (upper panels) or DAC (lower panels). 20x image. Red: EpCAM; Green: CD8; Blue: Hoechst. **c**, Percentage of CD69⁺ tissue-resident memory (T_{RM}) TIL in CD8⁺ TIL in tumors from vehicle or DAC treated mice. **d**, Numbers of CD8α⁺ dendritic cells (DC) (left) or CD103⁺ DC (right) per 10⁶ myeloid cells in tumors from vehicle or DAC treated mice. **e**, Percentage of Gr-1⁺CD11b⁺ myeloid-derived suppressor cells (MDSC) in tumors from vehicle or DAC treated mice. **f**, Percentage of CD8⁺ TIL expressing Gzmb in

tumors from vehicle or DAC treated mice. **g**, Percentage of CD8⁺ TIL producing IFN-γ (left), TNF (middle), or both cytokines (right) after ex vivo activation with PMA and ionomycin in tumors from vehicle or DAC treated mice. Flow plots at extreme right show representative samples for each group. Data shown are mean + s.e.m. and are representative of two independent experiments. (**a**, **c**–**g**) Vehicle (PBS): n = 7, 1 mg/kg: n = 5; 5 mg/kg: n = 6. One mouse in 5 mg/kg group died prior to study endpoint. **a**, **d**: Kruskal-Wallis test with Dunn's multiple comparison test. **c**, **e**–**g**: One-way ANOVA followed by Holm-Sidak's multiple comparisons. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001. Exact p-values are provided in Source Data.



Extended Data Fig. 8 | DAC activates multiple immune pathways through gene demethylation and increased chromatin accessibility. **a**, Induction of selected immune genes involved in inflammatory cell death, inflammation and type I interferon signaling in 4T1 cells after in vitro exposure for 48 hr to increasing concentrations of DAC. Data show relative mRNA expression assessed by qRT-PCR (target gene relative to *Gapdh*) in samples treated with DAC or vehicle (PBS), *n* = 3/group. Genes are involved in pyroptosis (red), necroptosis (blue), inflammatory cytokines (orange), or the type I interferon pathway (green). Dotted red line indicates 2-fold increase in mRNA expression following DAC treatment. **b**, Heatmap of the z-score-normalized expression of immune genes in 4T1 cells treated with vehicle or increasing concentrations of DAC. **c, d**, Effect of *Dnmt* gene knockdown in 4T1 cell line on DNMT expression (**c**) and innate immune gene expression (**d**), *n* = 3/group. Samples were assessed by qRT-PCR 48 h after in vitro siRNA transfection and normalized to control cells transfected with nontargeting control siRNA. Dotted red line indicates 2-fold increase in mRNA expression following DNMT silencing. **e**, Location of hypomethylated regions (by RBSS, purple bars), more accessible regions (by ATAC-seq, cyan

bars), and upregulated genes (by scRNA-seq, red bars) in 4T1 cells. Right column: zoomed in snapshots show representative innate immune genes *Ddx58*, *Irf7*, and *Ripk3*. The yellow highlighted regions indicate upregulated genes that correlate with hypomethylated DNA and more accessible promoters. **f**, Top: top 40 Gene Ontology terms associated with the overlapping 84 genes that were both derepressed by DAC in 4T1 and edited in the GEMM, as shown in Fig. 6d. Bottom: top 40 Gene Ontology terms associated with the 1539 genes upregulated in DAC-treated 4T1 cells that were not significantly edited in the GEMM, as shown in Fig. 6d. GO terms were grouped using unsupervised hierarchical clustering and annotated based on most representative functions. Row names with asterisks denote truncated names. Enrichment statistic: one-sided Fisher's exact test. P-values were adjusted for multiple-comparisons using the q-value method. **a, c, d**: Multiple t-test. Type I errors were corrected by two-stage step-up method of Benjamini, Krieger and Yekutieli. **P* ≤ 0.05, ***P* ≤ 0.01, ****P* ≤ 0.001, *****P* ≤ 0.0001. For **a** and **d**, only significantly upregulated data was indicated. Exact p-values are provided in Source Data.



Extended Data Fig. 9 | *Gsdme*, *Ripk3*, and *Irf3/Irf7* knockout in B16 and its effect on TIL from DAC-treated B16. **a-b**, Immunoblots of B16 clones probed for GSDME (**a**) and RIPK3 (**b**) and actin loading control. *Gsdme*^{-/-} clones 1 and 15 and *Ripk3*^{-/-} clone 9 were selected for in vivo studies. **a**, GSDME and actin in B16F10 cells infected with nontargeting lentivector (Ctrl) or clones knocked out for *Gsdme*. **b**, RIPK3 in DAC-treated B16F10 Ctrl cells or clones knocked out for *Ripk3*. **c**, B16F10 control cells or clones knocked out for *Gsdmd* were treated with DAC plus IFN-β and TNF to induce GSDMD expression. Immunoblots probed for GSDMD and actin are shown. **d**, Immunoblots probed for IRF3 and actin in B16F10 Ctrl cells or clones knocked out for *Irf3*. **e**, *Irf7* mRNA in vehicle (PBS) or

DAC treated B16 F10 ctrl cells or clones knocked out for *Irf3* and *Irf7*, assessed 48 h following DAC treatment by qRT-PCR. n = 3/group. **f**, Left: Mean numbers of CD8⁺ TIL per milligram of WT, *Gsdme*^{-/-}, *Ripk3*^{-/-}, *Irf3*^{-/-}*Irf7*^{-/-} or *Gsdmd*^{-/-} B16F10 tumors in mice treated with vehicle (PBS) or DAC (n = 5–7/group). Middle and Right: Percentages of CD8⁺ (middle) or NK TIL (right) expressing GzmB in each group of tumors from mice treated with vehicle (black bars) or DAC (red bars). Data shown are mean ± s.e.m. and are representative of at least two experiments. (**f**) Two-way ANOVA followed by Fisher's Least Significant Difference (LSD) test. *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. Exact p-values are provided in Source Data.

Reporting Summary

Nature Portfolio 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 Portfolio 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 Flow cytometry analysis and cell sorting with BD FACSCanto II and BD FACSAria II, respectively. Data Acquisition with BD FACS Diva software (version 8.0.1)
Immunofluorescent imaging: Zeiss Zen Blue (version 3.1)
Western Blot: BIO-RAD Image Lab Touch Software (Version 1.2)

Data analysis Flowjo (Version 10.8.1) was used to analyze flow cytometry data. The Cellranger software suite (v. 7.0.0, 10X Genomics) was used to map scRNA-seq reads to a customized mm10 genome. Seurat (v. 4.1.0) library was used to import and preprocess scRNA-seq data. Harmony (v. 0.1.0) library was used for removal of batch-effect adjustment for downstream clustering and network construction. Differential expression analysis of scRNA-seq clusters was performed using MAST, implemented in the Seurat package. CytoTRACE (v. 0.3.3) was used to determine cell maturation. DirichletReg (v. 0.7.1) was used for cell proportion analysis. CellChat (v. 1.5.0) was used for analyzing cell-cell communication. The Interferome webportal (interferome.org) was used to determine interferon regulated genes (accessed 09-2022). ClusterProfiler (v. 4.3.1) was used for enrichment analysis. All other statistical analysis, visualization and data processing was done using standard R (v. 4.2.0) packages. Cytoscape (v. 3.9.1) was used to run EnrichmentMap and perform network clustering of Gene Ontology terms.

Additional analysis was performed on two publicly available external scRNA-seq datasets: 1- An atlas of 50K normal murine mammary epithelial cells (Saeki et al. 2021). 2- A single cell dataset from breast tumor tissues of MMTV-PyMT mice (Chou et al. 2022).

For bulk RNA-seq data, the fastq files were trimmed for the sequencing adaptors using CutAdapt (version 2.5) or Trimmomatic (0.36). The sequencing reads were aligned to the mouse genome (mm10) by Bowtie2 in RSEM (v. 1.3.3) pipeline with the following parameters bowtie2 -q --phred33 --sensitive --dpad 0 --gbar 99999999 --mp 1,1 --np 1 --score-min L,0,-0.1 -l 1 -X 1000 --no-mixed --no-discordant -p 4 -k 200.

Differentially expressed genes were determined using DESeq2 R package (v. 1.35.0). ClusterProfiler (v. 4.3.1) was used for enrichment analysis.

Sequencing reads from ATAC-seq were mapped to the mouse genome using Bowtie2 (v. 2.3.4.3) and peak calls were performed using HMMRATAC (<https://github.com/LiuLabUB/HMMRATAC>) software package.

For RRBS-seq data, sequencing reads were mapped to genome and methylation calls were performed using Bismark (v. 0.23.0). Differentially methylated regions (DMR) between DAC-treated and control cell lines were determined using the Metilene software package (v. 0.2-8).

Analyses presented in this study and codes to generate the figures are publicly available at: https://github.com/pouryany/scRNAseq_GEMM.

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 Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

RNA-seq, ATAC-seq, and RRBS datasets are deposited at Gene Expression Omnibus (GSE212029). The single-cell RNA-seq data (raw and processed) are available on the Gene Expression Omnibus portal of the National Center for Biotechnology Information (NCBI GEO) via the accession code GSE234917. Additional data used in this paper include an atlas of 50K normal murine mammary epithelial cells¹⁶ (accessed through online portal: <https://mouse-mammary-epithelium-integrated.cells.ucsc.edu>), a single cell dataset from breast tumor tissues of MMTV-PyMT mice²¹ (using GEO Accession ID GSE195937), a list of known endogenous retroviral (ERV) genes from the gEVE database⁶⁸ (<http://geve.med.u-tokai.ac.jp/>, accessed 12-2021), and a list of interferon stimulated genes from the Interferome database⁶⁹ (<https://interferome.org/>, accessed 09-2022). Source data are provided with this paper. A ShinyApp for interactive exploration of the scRNA-seq is available at https://pouryany.shinyapps.io/scgemm_app/. Analyses presented in this study and codes to generate the figures are publicly available at: https://github.com/pouryany/scRNAseq_GEMM.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Reporting on race, ethnicity, or other socially relevant groupings

Population characteristics

Recruitment

Ethics oversight

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

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

Data exclusions

Replication

Replication	one failure, it was performed another time to confirm the results to support conclusions stated in the manuscript. Data presented in the manuscript is from a single representative experiment.
Randomization	Animals were assigned randomly to experimental and control groups. Random allocation and quantitative measurement using instruments and kits in our experiments minimized biased assessments.
Blinding	Blinding was not possible for experiments, since for each assay performed, the same scientist was responsible to prepare, acquire and analyze data.

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
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

Antibody name and flurochrome Supplier name Catalog number Clone #

Anti-mouse CD45-PerCP-Cy5.5 BioLegend 103132 Clone# 30-F11

Anti-mouse CD45-Pacific Blue BioLegend 157212 Clone# S18009F

Anti-mouse CD8a FITC BioLegend 100705 Clone# 53-6.7

Anti-mouse CD8a PerCP-Cy5.5 BioLegend 100734 Clone# 53-6.7

Anti-mouse CD8b Alexa Fluor 700 BioLegend 126617 Clone# YTS156.7.7

Anti-mouseCD4 PerCP-Cy5.5 BioLegend 100434 Clone# GK1.5

Anti-mouseCD4 PE-Cy7 BioLegend 100422 Clone# GK1.5

Anti-mouse/human CD44 PerCP-Cy5.5 BioLegend 103032 Clone# IM7

Anti-mouse/human CD44 Pacific Blue BioLegend 103020 Clone# IM7

Anti-mouse CD69 PE-Cy7 BioLegend 104511 Clone# H1.2F3

Anti-mouse CD49b FITC BioLegend 108906 Clone# DX5

Anti-mouse CD49b Pacific Blue BioLegend 108918 Clone# DX5

Anti-mouse Nkp46 APC BioLegend 137608 Clone# 29A1.4

Anti-mouse Ly-6G/Ly-6C (Gr-1) Antibody PE BioLegend 108408 Clone# RB6-8C5

Anti-mouse/human CD11b Alexa Fluro700 BioLegend 101222 Clone# M1/70

Anti-mouse CD11c Brilliant Violet 421 BioLegend 117343 Clone# N418

Anti-mouse CD11c Pacific Blue BioLegend 117322 Clone# N418

Anti-mouse CD103 Alexa Fluor700 BioLegend 121442 Clone# 2E7

Anti-mouse F4/80 PE-Cy7 BioLegend 123114 Clone# BM8

Anti-mouse I-A/I-E Alexa Fluor 700 BioLegend 107621 Clone# M5/114.15.2

Anti-human/mouse Granzyme B Pacific Blue BioLegend 515408 Clone# GB11

Anti-mouse Perforin PE BioLegend 154306 Clone# S16009A

Anti-mouse IFN-g Pacific Blue BioLegend 505818 Clone# XMG1.2

Anti-mouse TNF PE-Cy7 BioLegend 506323 Clone# MP6-XT22

Goat anti-Rat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 488 Thermo Fisher Scientific A-11006

Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 568 Thermo Fisher Scientific A10042

CD8a Monoclonal Antibody, eBioScience Thermo Fisher Scientific 14-0808-80 Clone# 4SM15

EpCAM/CD326 Polyclonal antibody Proteintech 21050-1-AP Clone# N/A

Anti-mouse CD326 (Ep-CAM)-PerCP-Cy5.5 BioLegend 118220 Clone# G8.8

Anti-mouse CD326 (Ep-CAM) PE-Cy7 BioLegend 118216 Clone# G8.8

Recombinant anti-DFNA5/GSDME Abcam ab215191 Clone# EPR19859

Recombinant anti-GSDMD Abcam ab209845 Clone# EPR19828

RIP3 Antibody ProSci 2283 Clone# N/A

MLKL (D6W1K) Rabbit mAb Cell Signaling Technology 37705s Clone# D6W1K

Phospho-MLKL (Ser345) (D6E3G) Rabbit mAb Cell Signaling Technology 37333s Clone# D6E3G

IRF-3 (D83B9) Rabbit mAb Cell Signaling Technology 4302s Clone# D83B9

Caspase-3 Antibody Cell Signaling Technology 9662s Clone# N/A

Anti-beta-actin (JLA20) Developmental Studies Hybridoma Bank N/A Clone# JLA20

Validation

Validation for all antibodies is available on vendor websites. All Abs were acquired from commercial sources. In addition, we validated each antibody and dye as per the manufacturer's instructions. Antibodies were all validated by staining for

antigen-expressing cells but no staining for non-expressing cells. The dilution of antibodies and dyes were titrated to obtain the best concentration to use that stained only antigen-positive cells but not negative cells. Gates were set using unstained samples or samples stained with isotype control antibodies.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	HEK293T cells were purchased from the American Type Culture Collection (ATCC). B16F10 cells were kind gifts of Gordon Freeman (Dana-Farber Cancer Institute), which was originally purchased from ATCC. 4T1 cells were provided by Fred Miller (Wayne State University). B16F10-eGFP cells were generated by transducing B16F10 cells with lentiviral vector expressing GFP (Amsbio) and sorted for GFP+ cells. B16F10 Gsdme ^{-/-} cells, B16F10 Gsdmd ^{-/-} cells, B16F10 Irf3 ^{-/-} Irf7 ^{-/-} cells, and B16F10 Ripk3 ^{-/-} cells were generated by CRISPR/Cas9 gene knockout in the lab.
Authentication	The original cell lines were not authenticated beyond purchasing them from vendors or acquiring them through collaborators. B16F10-eGFP cells were positively selected for GFP+ cells by sorting and confirmed by imaging analysis. The B16F10 Gsdme ^{-/-} cells, B16F10 Gsdmd ^{-/-} cells, B16F10 Irf3 ^{-/-} Irf7 ^{-/-} cells, and B16F10 Ripk3 ^{-/-} cells were verified by western blot, qRT-PCR and ELISA test (for B16F10 Irf3 ^{-/-} Irf7 ^{-/-} cells), and sequencing.
Mycoplasma contamination	All cell lines were tested by PCR and were negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.

Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	Details about the mice used in this study have been included in the Methods section of this manuscript. Female BALB/c and C57BL/6 mice that are 6-8 weeks old were purchased from The Jackson Laboratory. The B6.129S2-Irfar1tm1Agt/Mmjax (Irfar1 ^{-/-}) mice and Prf1 ^{-/-} mice were purchased from The Jackson Laboratory and bred in house. ErbB2ΔEx16 ^{+/-} transgenic mice and MTB ^{+/-} transgenic mice (harboring the MMTV/reverse tetracycline transactivator transgene) were generously provided by William J. Muller (McGill University). All mice used for experiments were 6-8 weeks old. 8 week old ErbB2ΔEx16 ^{+/-} transgenic mice were used for tumor induction.
Wild animals	No wild animals were used in the study.
Reporting on sex	As this study mainly focused on breast cancer, only female mice were selected for experiments.
Field-collected samples	No field collected samples were used in the study.
Ethics oversight	All procedures were conducted using protocols approved by the Harvard Medical School IACUC.

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 isolate tumor-infiltrating immune cells for flow cytometry and functional analyses, we cut tumors into small pieces and digested the samples with RPMI medium supplemented with 2 mg/ml collagenase D, 100 µg/ml DNase I and 2% FBS with agitation for 30 min at 37 C. Samples were then homogenized, filtered through 40 µm strainers and washed with Leibovitz's L-15 medium twice.
Instrument	Flow cytometry was performed using a BD FACSCanto II analyzer.
Software	Data were acquired by BD FACSDiva software, then analyzed using Flowjo software V.10 (Becton Dickinson).

Cell population abundance

The purity of each cell population analyzed was evaluated by flow cytometry after staining with the appropriate antibodies (CD45+ for tumor-infiltrating immune cells).

Gating strategy

FSC/SSC gates were used to select mononuclear cells. FSC/FSH gates were then used to gate on single cells. Cells were then gated on live cells. Live cells were further gated on CD45+ leukocytes and then different populations of immune cells based on their expression of distinct markers. When needed, isotype control antibody staining was used to define positive/negative cell populations.

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