The repercussions of recognition: imprints of T cells on the tumor microenvironment

Slagter, M.

CHAPTER 5

Distinct spatiotemporal dynamics of CD8$^+$ T cell derived cytokines in the tumor microenvironment

Mirjam E. Hoekstra[#], Maarten Slagter[#], Jos Urbanus, Mireille Toebes, Nadine Slingerland, Iris de Rink, Roelof J. C. Kluin, Marja Nieuwland, Ron Kerkhoven, Lodewyk F. A. Wessels and Ton N. Schumacher

[#] These authors contributed equally.

Cancer Cell, 2024. DOI: 10.1016/j.ccell.2023.12.010

Abstract

Cells in the tumor microenvironment (TME) influence each other through the secretion and sensing of soluble mediators, such as cytokines and chemokines. While signaling of interferon γ (IFN-γ) and tumor necrosis factor α (TNF-α) is integral to anti-tumor immune responses, our understanding of the spatiotemporal behavior of these cytokines is limited. Here, we describe a single cell transcriptome-based approach to infer which single or combined signals an individual cell has received. We demonstrate that, contrary to expectations, CD8$^+$ T cell-derived IFN-γis the dominant modifier of the TME relative to TNF-α. Furthermore, we demonstrate that cell pools that show abundant IFN-γsensing are characterized by decreased expression of TGF-β-induced genes, consistent with IFN-γ-mediated TME remodeling. Collectively, these data provide evidence that CD8$^+$ T cell-secreted

cytokines should be categorized into local and global tissue modifiers, and describe a broadly applicable approach to dissect cytokine and chemokine modulation of the tumor microenvironment.

Introduction

Tumors are composed of a diversity of interacting cell types, including tumor cells, fibroblasts, endothelial cells, and a variety of immune cell types. A first type of interactions between the cell populations that jointly form the TME is formed by direct cell-cell contacts, and to describe the effects of such cellular interactions, technologies such as PIC-seq, which allows RNA-sequencing (RNA-seq) of physically interacting cell pairs, have been developed[1]. Importantly, next to such direct cell – cell interactions, cells also use soluble factors, such as cytokines, chemokines, and growth factors, to influence the state of surrounding tissue cells[2].

One of the major cytokine-producing cell compartments in tumor tissue is formed by the $CD8^+$ cytotoxic T cell pool, and $CD8^+$ T cells have been demonstrated to play a central role in both immune checkpoint blockade[3] and adoptive T cell therapies[4-6]. Upon encounter of antigen-expressing target cells, $CD8^+$ T cells release lytic granules containing cytotoxic molecules such as perforin and granzymes in the synapse that is formed between the interacting cells. In addition, T-cell receptor (TCR) signaling leads to the secretion of the cytokines interferon γ (IFN-γ), tumor necrosis factor α (TNF-α) and Interleukin 2 (IL-2), which can, on their own or jointly, induce large-scale alterations in the transcriptome of cells that sense these factors. For example, IFN-γ receptor (IFN-γR) signaling has been demonstrated to result in increased expression of components of the antigen presentation pathway, enhances expression of immune checkpoint molecules, and can promote recruitment of other immune cells through production of chemokines such as CXCL9, 10 and 11[7,8]. In addition, IFN-γ and TNF-α have been demonstrated to regulate the activation and maturation state of, amongst others, macrophages and dendritic cells[9-12]. Furthermore, both IFN-γR and TNF-α receptor (TNF-αR) signaling can, in a context-dependent fashion, contribute to tumor cell senescence[13], apoptosis[8,14] and ferroptosis[15]. Finally, besides their direct effects on tumor cells, both IFN-γ and TNF-α can also be critical for tumor control through their effects on stromal cells in the tumor vasculature[16-20].

In spite of the central role of T cell-produced cytokines in the modulation of cell behavior in the TME, our understanding of the spatiotemporal behavior of $CD8^+$ T cell derived cytokines is limited. Specifically, it has not been established whether these signaling molecules differ in their capacity to not only influence target cells in the immediate vicinity of sites of antigen recognition, but also modulate the behavior of cells in the tumor tissue in a more global manner. Earlier work has demonstrated that cytokines may either be secreted in a multidirectional fashion, or can selectively be released in the immune synapse, analogous to the focused release of lytic granules[21]. Specifically, following TCR triggering, membrane-bound TNF-α has been shown to be distributed equally over

the cell membrane, as demonstrated by live imaging of TNF-α on activated murine CD4$^+$ T cells[21]. In contrast, in clusters of T cells and antigen-presenting cells, IFN-γ, IL-2, IL-4, and IL-5 were all shown to be localized at the microtubule organizing center (MTOC), consistent with directional release[21-hyperlinkcs2323]. Based on this postulated difference in mode of secretion, a more profound effect of TNF-α relative to, for instance, IFN-γ or IL-2 on cells that are distant from the site of antigen recognition could be expected. However, as the size of the cell field in which productive cytokine sensing can occur is also influenced by other parameters, such as cytokine half-life, receptor-mediated clearance, and binding to extracellular components, it has been difficult to predict the extent of long range sensing of different cytokines in the TME[24].

Evidence that cytokines can reach (remote) bystander cells that cannot be recognized by T cells directly has been obtained in a number of studies in both viral infection and tumor models[25-30]. Specifically, T cell-secreted IFN-γ in skin and lymph nodes was shown to induce expression of IFN-γ-responsive genes in large regions outside the parasite or virus infected areas[26,27,31]. Likewise, secretion of IFN-γ and TNF-α by CD4$^+$ T cells in tumors has been demonstrated to induce senescence in tumor cells that cannot be directly recognized by T cells[13]. In case of CD8$^+$ T cells, long-range sensing of IFN-γ has been observed by, amongst other, intravital imaging of fluorescent IFN-γR-signaling reporters in mosaic tumors that contain both antigen-positive and antigen-negative tumor areas. Using such fluorescent reporter systems, it was demonstrated that a large fraction of bystander cells senses IFN-γ upon intratumoral T-cell activation, and that IFN-γ sensing can occur in tumor cells at distances over hundreds of micrometers from the site of T cell activation[29,30]. Collectively, these data on individual cytokines provide an incentive to develop technology to measure and deconvolute the joint effects of multiple cytokines on the TME. In the present study, we set out to generate a strategy that allows the analysis of the effects of a broad set of cytokines simultaneously, and also provides information on the timing of such cytokine exposure. The data obtained demonstrate that, contrary to what would be predicted based on their mode of secretion, IFN-γ is the dominant T cell-secreted modifier of the TME.

Results

In order to measure cytokine sensing in the TME in a manner that is independent of genetic reporter systems, we explored whether gene expression signatures can reliably inform on the type and duration of cytokine exposure. Towards this goal, we exposed human ovarian carcinoma (OVCAR5) cells to different cytokines or cytokine combinations for 2-24 hours and analyzed transcriptomes by bulk RNA-seq (RNA-seq). In line with expectations, this revealed large groups of genes that were selectively induced by IFN-γ (such as *HLA-DRA* and *IRF1*), or TNF-α (such as CCL20 and MMP9). In addition, a gene set was identified that was either preferentially or exclusively induced by the combination of these cytokines (e.g., *UBD* and *CXCL9*) (Figure 5.1a). Furthermore, the relative expression of individual genes in these gene sets provided rich information on cytokine exposure time, distin-

guishing genes with a 'burst-like' expression pattern (e.g., *CCL20* after TNF-α exposure), and genes for which expression showed an exponential increase over time (e.g., *MMP9* after TNF-α exposure) (Figure 5.1a). While gene expression data were rich in terms of the nature and duration of cytokine exposure (Figure 5.1a-b), no substantial differences in gene expression were observed as a function of cytokine concentration beyond a limited concentration range (~10-fold range). Importantly, cell culture medium derived from T cell-tumor cell co-cultures contained high levels of IFN-γ (>100 ng/ml) and TNF-α (± 1 ng/ml) and induced a gene expression profile that was highly similar to that observed upon dual IFN-γ *plus* TNF-α exposure (Figure S5.1).

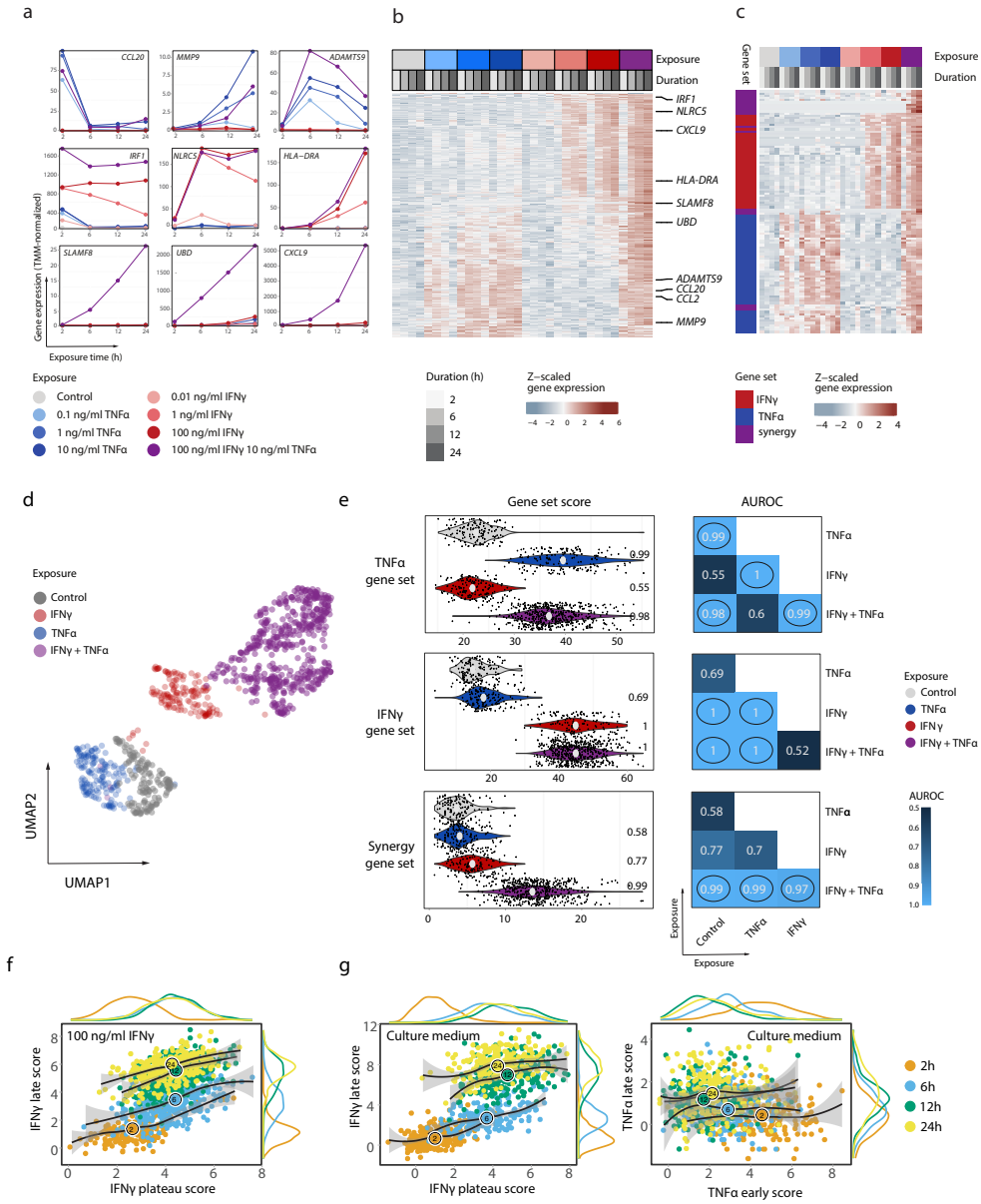

Figure 5.1: Gene expression reports on cytokine exposure

a mRNA expression profiles of selected genes in OVCAR5 cells exposed to indicated concentrations of IFN-γ, TNF-α or their combination, for the indicated duration. Top, middle, and bottom panels depict genes that are primarily responsive to TNF-α, IFN-γ, or TNF-α *plus* IFN-γ, respectively.

b Heatmap of bulk gene expression values inferred from OVCAR5 cells exposed to indicated concentrations of TNF-α, IFN-γ or TNF-α *plus* IFN-γ, for indicated durations. Unsupervised hierarchical clustering of data (shown are the 612 genes from the 'cytokine-responsive class', see Figure S5.2b), groups samples by exposure type and then by exposure duration. **c** Heatmap of bulk gene-expression values for mono-responsive genes and synergy genes inferred from *in vitro* stimulated OVCAR5 cells, as in a and b. Unsupervised hierarchical clustering of gene expression data shows a nearly full agreement with assigned gene classes (cluster purity of 0.86).

d UMAP of single cell RNA-seq data of OVCAR5 cells stimulated with indicated recombinant cytokines for 24h.
e Gene set scores for single cell RNA-seq data of *in vitro* cytokine stimulated cells as in a. Left panels: dots represent gene set scores of individual cells; violins represent densities of score distributions. Area under the receiver operator curve (AUROC) values, quantifying how well experimental conditions can be distinguished from the control condition, are depicted. Right panels: Heatmaps showing pairwise distinguishability of indicated experimental conditions (axes) using gene set scores, as quantified using AUROC values. Comparisons for which indicated gene sets are designed to show separation are encircled.
f IFN-γ plateau versus IFN-γ late gene set scores (see Methods) for OVCAR5 cells stimulated with IFN-γ (100ng/ml) for the indicated times. Black lines are LOESS-smoothed curves representing local averages, one per stimulus duration. The ratio of each of the two gene set scores informs on duration of cytokine exposure.
g IFN-γ plateau versus IFN-γ late gene set scores and TNF-α early versus TNF-α late gene set scores for OVCAR5 cells stimulated with culture medium obtained from T cell -tumor cell co-cultures (as in Figure S5.1) for the indicated times as in **F**.

To be able to assign duration and type of cytokine exposure to individual cells, we compiled a set of cytokine informative genes through a combination of model training and manual curation. In brief, to efficiently expand a seed set of ~80 informative genes identified by manual selection, all genes were annotated with descriptive features (see methods), developed to discern cytokine-responsive from unresponsive genes (Figure S5.2a). We next iteratively expanded from this initial seed set by training a machine learning model with the objective of predicting gene classes (e.g., 'cytokine-responsive' and 'cytokine-unresponsive') for all yet unclassified genes (Figure S5.2b-c). After 10 iterations of model training, gene class prediction, and correction of predicted classes, a set of 612 cytokine-responsive genes was obtained. (Figure S5.2d). 90 of these could be classified as mono-responsive to either IFN-γ (n=40) or TNF-α (n=50), i.e., with only a single cytokine eliciting a response and IFN-γ *plus* TNF-α eliciting a response that did not substantially deviate from the response to the main excitatory cytokine. This property renders these genes especially useful when simultaneously assessing the spreading behavior of both cytokines (Figure S5.2e and Figure 5.1c) and allows for the summing of expression values for component genes to infer cytokine stimulus (see below). Importantly, comparison of these gene sets to the TNF-α and IFN-γ Hallmark gene sets[32] that are frequently used to evaluate signaling revealed only modest overlap (Figure S5.3a). Whereas the newly developed gene sets consisted solely of genes that responded strongly and specifically to the cytokine they were assigned to in OVCAR5 cells, the Hallmark gene sets showed considerable overlap. Furthermore, a sizeable number of TNF-α hallmark genes was shown to respond to IFN-γ and vice versa, creating the potential for incorrect signal inference (Figure S5.3b-d). Next to the set of mono-responsive genes, a set of synergy genes, which are selectively expressed in the presence of both IFN-γ and TNF-α, was identified, providing an independent means to measure co-occurrence of IFN-γR and TNF-αR signaling (Figure 5.1c). Finally, analysis of gene expression dynamics demonstrated that this approach can inform on cytokine exposure duration (Figure S5.3e).

Having established a number of cytokine-and time-informative gene sets, we tested whether gene expression upon cytokine exposure is informative in single cell (sc) transcriptome data. In an unsupervised analysis, cells exposed to activating concentrations of either IFN-γ, TNF-α, or their com-

bination formed separated clusters, both from control cells and each other, (100% rejection rate on kBET-test[33] with stimuli as batches; median silhouette width: 0.17) (Figure 5.1d). To quantify signal strength for each exposure-specific gene set, we subsequently calculated cell expression scores for all genes that were included in either gene set. Using this strategy on single cell RNA-seq data from cells exposed to activating concentrations of IFN-γ or IFNα revealed a near perfect separation of IFN-γ and TNF-α exposed cells from control cells, as well as a clear separation between cells exposed to the two different stimuli (Figure 5.1e). In addition, exposure to the combination of IFN-γ *plus* TNF-α could be identified with high precision, both by analysis of the separate IFN-γand TNF-α gene set scores, and by use of the IFN-γ *plus* TNF-α synergy gene set (Figure 5.1e). Furthermore, ability to correctly assign cytokine stimuli was not affected by the experimental strategy (protease digestion, flow cytometric sorting) required to obtain single cell information from tumor material (Figure S5.3f). Finally, use of time-informative gene sets on *in vitro* cultured cells exposed to recombinant cytokines or culture medium from T cell-tumor cell co-cultures at different time points demonstrated the ability of this technique to also infer stimulus duration from single cell data (Figure 5.1f-g and Table S5.6).

Having established methodology for TNF-α and IFN-γ exposure inference in single cells, we subsequently set out to measure the degree of T cell-secreted cytokine sensing by tumor cells in the TME *in vivo*. To this purpose, OVCAR5 tumors that were composed of a large fraction of antigen negative ('bystander') tumor cells that could serve as cytokine sensing reporter cells, plus a small fraction of tumor cells that form targets for neoantigen-specific T cells[30] were established in NSG-β2m$^{-/-}$ mice (Figure 5.2a). Following treatment of mice bearing such mosaic tumors with TCR-transduced $CDK4_{R>L}$ neoantigen-specific $CD8^+$ T cells, infiltration of $CD8^+$ T cells into tumor tissue is observed and upon target cell recognition, cytokine production is initiated in tumor regions composed of antigen-positive tumor cells[30]. Note that in this setup, all subsequent analyses of cytokine-specific gene set scores by single cell RNA-seq were restricted to antigen negative bystander cells that cannot be recognized by T cells, and hence weren't influenced by direct cell-to-cell killing of antigen-positive tumor cells.

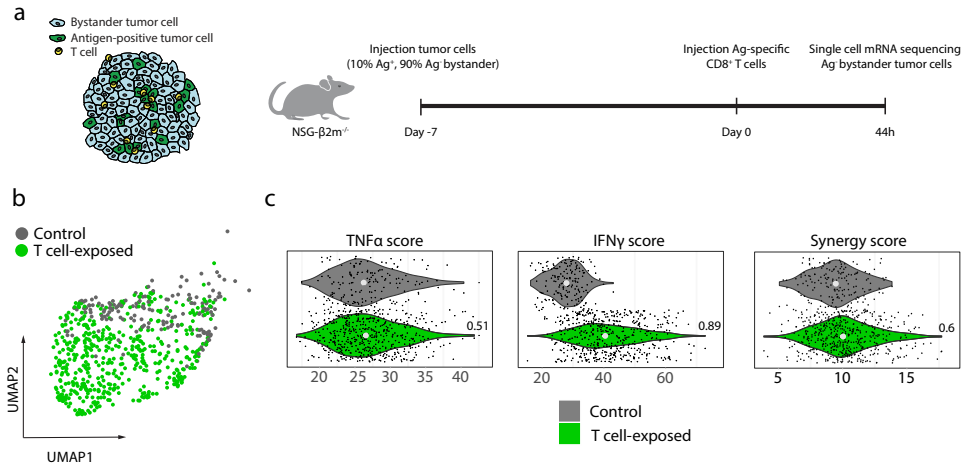

Figure 5.2: Frequent IFN-γ but not TNF-α sensing by bystander tumor cells.
a NSGβ2m$^{-/-}$ mice injected subcutaneously with a mixture of 10% CDK4$_{R>L}$ antigen expressing and 90% bystander OVCAR5 tumor cells were treated with either PBS (control) or CDK4$_{R>L}$-specific CD8$^+$ T cells after tumor establishment. Tumors were harvested 44h after treatment, and bystander tumor cells were analyzed by single cell RNA-seq.
b UMAP of single cells based on gene expression in the "cytokine-responsive" gene class, as described in Figure S5.2b.
c TNF-α, IFN-γ, and synergy gene set scores of single cells derived from OVCAR5 tumors. Dots represent gene set scores of individual cells, violins represent densities of score distributions. Numeric values reflect AUROC values that quantify separability between experimental conditions. Note that IFN-γ, but not TNF-α, gene set scores are increased in the T cell-exposed condition as compared to the control condition.

Having established that T cell infiltration and activity are first detected around 16 hours and increase up to 44 hours after T cell transfer (Figure S5.4a), we analyzed bystander tumor cells by single cell RNA-seq at 16-44 hours after T cell transfer. Unsupervised clustering of bystander cells derived from 44h T cell-exposed tumors and control tumors demonstrated a separation of a large fraction of bystander cells obtained from T cell-exposed tumors (Figure 5.2b). Thus, the presence of a tumor-reactive CD8$^+$ T cell compartment substantially modifies the transcriptome of a considerable part of bystander tumor cells in the TME. Importantly, assignment of cells to different cytokine exposure conditions revealed that a large fraction (70.4%) of bystander tumor cells in T cell-treated mice showed a pronounced IFN-γ gene set score, whereas such IFN-γ sensing was largely absent in tumor cells from control mice (5.3% of cells) (Figure 5.2c and Figure S5.4b). In contrast, presence of a tumor-reactive T cell compartment did not measurably increase the fraction of TNF-α sensing cells, with 5.3% of tumor cells classified as TNF-α sensing in both control and T cell-exposed tumors (Figure 5.2c and Figure S5.4b). In addition, tumor cells displaying high IFN-γ gene set scores did not show elevated TNF-α gene set scores (Fig S4b). As a second test of *in vivo* TNF-α exposure, we calculated synergy gene set scores, which independently inform on the sensing of the combination

of IFN-γ *plus* TNF-α (Figures 5.2c and S5.4b). Also by this metric, the presence of a tumor-reactive T-cell compartment did not result in TNF-α sensing by an appreciable fraction of bystander tumor cells. As a majority of TNF-α genes displays a burst-like, early, expression pattern, we next assessed cytokine sensing at the 16h timepoint, at which measurable T cell infiltration is just visible (Figure 5.3a). Already at this time point, a subset of bystander tumor cells derived from T cell-exposed tumors separated from bystander tumor cells in control tumors (Figure 5.3b). However, neither the use of the entire TNF-α mono reporter gene set (Figure 5.3c, left), nor the use of the TNF-α early time-informative gene set, showed an appreciable TNF-α sensing signal (Figure 5.3d, right). Note that T cells derived from such tumors did retain the capacity to produce TNF-α, indicating that the lack of an appreciable tumor cell population that showed TNF-α sensing was not explained by impaired cytokine production (Figure S5.4c). As a side note, application of the time reporting IFN-γ-responsive gene sets demonstrated that tumor cells isolated 44 hours after T-cell infusion that show a given expression of the IFN-γ "plateau gene set score" on average showed a slightly increased expression of the IFN-γ "late gene set score", as compared to tumor cells analyzed 16 hours after T-cell infusion (Figure 5.3d, left). To test whether bystander tumor cells did retain the capacity to respond to TNF-α *in vivo* when this cytokine is present, we intratumorally injected tumors with recombinant cytokines. Importantly, an evident TNF-α signal was observed upon injection of either TNF-α (with 53.4% of cells surpassing 95th percentile of control) or TNF-α *plus* IFN-γ (51.8%) (Figure 5.3e and Figure S5.4d), Furthermore, a pronounced synergy signal was selectively observed upon intratumoral injection of TNF-α *plus* IFN-γ (49.4% of cells), likewise indicating a high sensitivity to detect TNF-αR signaling *in vivo*.

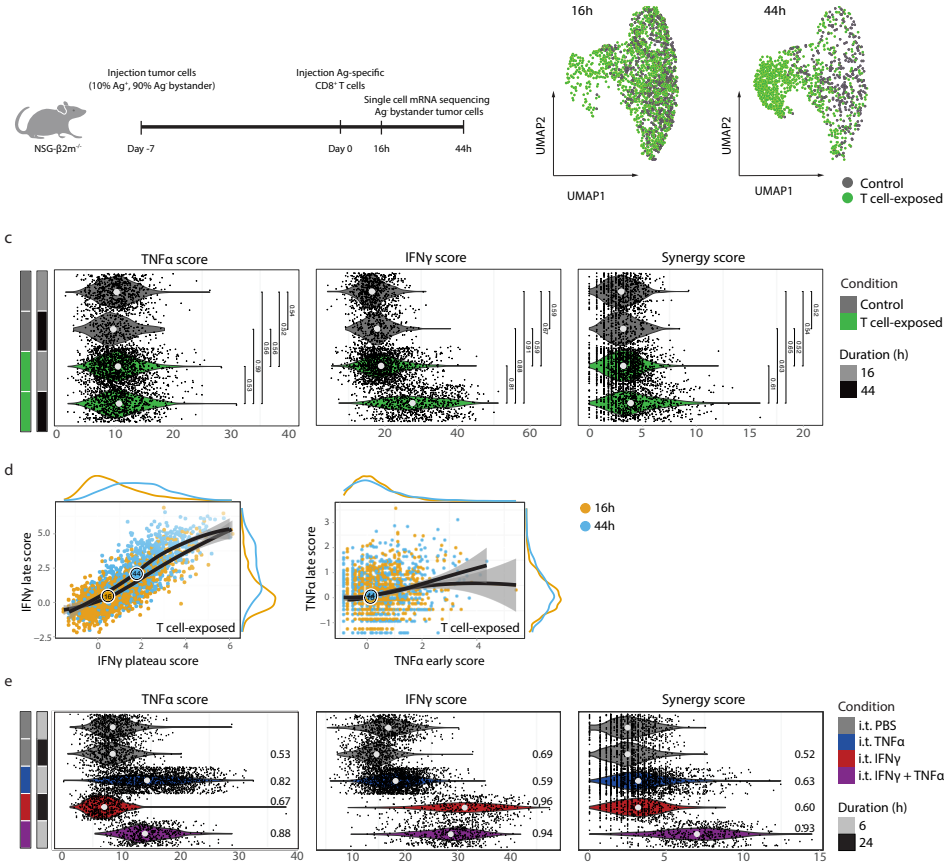

Figure 5.3: Appreciable IFN-γ but not TNF-α sensing by bystander tumor cells early after T-cell activation.

a NSG-β2m$^{-/-}$ mice injected subcutaneously with a mixture of 10% CDK4$_{R>L}$ antigen expressing and 90% bystander OVCAR5 tumor cells were treated with PBS (control) or with CDK4$_{R>L}$-specific CD8^{+} T cells after tumor establishment. Tumors were harvested 16h or 44h after treatment, and bystander tumor cells were subjected to single cell RNA-seq.

b UMAP of single cell sequencing data from bystander cells of control and T cell-exposed OVCAR5 tumors harvested 16 or 44h after treatment, based on genes in the "cytokine-responsive" class (Figure S5.2b).

c Violin plots of TNF-α, IFN-γ, and synergy gene set scores of cells derived from OVCAR5 tumors, as described in a. Dots represent gene set scores of individual cells, violins represent densities of score distributions. Numeric values reflect AUROC values that quantify separability between experimental conditions.

d Scatter plots of time-informative gene sets for *in vivo* single cell data described in a. To remove gene expression effects due to exposure duration independent from T-cell exposure, depicted gene set scores were normalized to control (PBS treated) counterparts in a duration-matched fashion (Methods).

e TNF-α, IFN-γ, and synergy gene set scores of OVCAR5 tumor cells derived from tumors injected with indicated recombinant cytokines. Cytokine exposure times were chosen based on maximal change in expression of cytokine specific responsive genes after in *vitro* cytokine exposure (Figure 5.1b). Numeric values reflect AUROC values that quantify separability between experimental conditions, as in Figure 5.1e.

To test whether the observed difference in long-range IFN-γ and TNF-α sensing also occurs in syngeneic tumor models, in which not only tumor cells but also infiltrating immune cells can respond to T cell-secreted cytokines[34,35], we compiled responsive gene sets from cytokine-stimulated mouse NRAS mutant melanoma (NMM) cells (Figure S5.5a). Application of these gene sets to bystander cells derived from mosaic NMM tumors demonstrated that a large fraction of bystander tumor cells (38.0%) responded to IFN-γ in T cell-treated mice, as compared to bystander tumor cells in mice that did not receive antigen-specific T cells (5.1%), or in mice in which antigen was lacking (5.7%). In contrast, presence of a tumor-reactive $CD8^+$ T-cell compartment did not induce sensing of TNF-α by bystander tumor cells (4.5% responding cells vs 5.1% and 5.7% in the two controls, respectively) (Figure 5.4a-b, Figure S5.5b). For NMM tumors, but not for OVCAR5 tumors, it is possible that cytokine-induced cell death results in a slight underestimate of the fraction of cells encountering the combination of IFN-γ and TNF-α signals. However, for both models, the observed bias towards IFN-γ sensing was not predominantly explained by TNF-α induced cell death (Figure S5.5c). Collectively, these data demonstrate in two different mouse models, and using gene sets that either report on the sensing of individual cytokines or on the combination of IFN-γ *plus* TNF-α, that widespread sensing is restricted to T cell-derived IFN-γ.

The ability to identify bystander tumor cells that have sensed IFN-γ *in vivo* makes it possible to test whether such sensing is associated with additional changes in cell state. To explore this, we used Milo[36] to identify transcriptionally similar cells (so-called neighborhoods) in the mouse NMM melanoma data. 64 out of 128 neighborhoods were enriched for bystander tumor cells derived from T cell-exposed tumors (hereafter referred to as 'T cell-exposed neighborhoods') relative to bystander cells from the PBS control condition (figure 5.4c) jointly comprising 74.3% of bystander tumor cells from T cell-exposed tumors. As a control, none of these neighborhoods were enriched or depleted for bystander tumor cells derived from T cell treated tumors in which antigen was lacking (Fig 4d, top). As expected, T cell-exposed neighborhoods showed a prominent IFN-γ-sensing profile but were also characterized by reduced expression of a second gene set that showed considerable overlap with genes induced by *in vitro* TGF-β stimulation of NMM melanoma cells (Figure 5.4d). Analysis of selected TGF-β responsive genes from bulk MNN RNAseq data (Figure S5.5a bottom panel) showed that a majority of these (47 out of 70) were negatively correlated with T-cell pressure (Figure 5.4e). However, a large fraction of TGF-β induced genes also appeared to show reduced expression upon IFN-γ exposure (Figure S5.5a), making it difficult to unambiguously ascribe this transcriptional response to lowered TGF-β sensing in T cell neighborhoods using solely gene signatures. To disentangle the transcriptional effects of co-occurring cytokines that regulate partly overlapping gene sets, we employed transcriptional deconvolution of cell neighborhoods, aiming to reconstruct their transcriptomes by algorithmically identifying optimal mixing weights of whole-transcriptome bulk RNAseq profiles (Methods), akin to the CIBERSORT approach[37]. A high similarity was observed between control neighborhoods and profiles of TGF-β-stimulated and unstimulated cells. In contrast, T cell-exposed neighborhoods more strongly resembled IFN-γ expression profiles (Figure 5.4f). Notably, omission of TGF-β-exposed reference profiles from this analysis increased recon-

struction error and predominantly did so for control neighborhoods (Figure 5.4g, left). In contrast, exclusion of IFN-γ-profiles specifically increased reconstruction error of T cell-exposed neighborhoods (Figure 5.4g, right). Collectively, these data demonstrate that T-cell pressure modulates bystander tumor cells towards transcriptional activity that is consistent with IFN-γ-sensing and with reduced TGF-β-sensing.

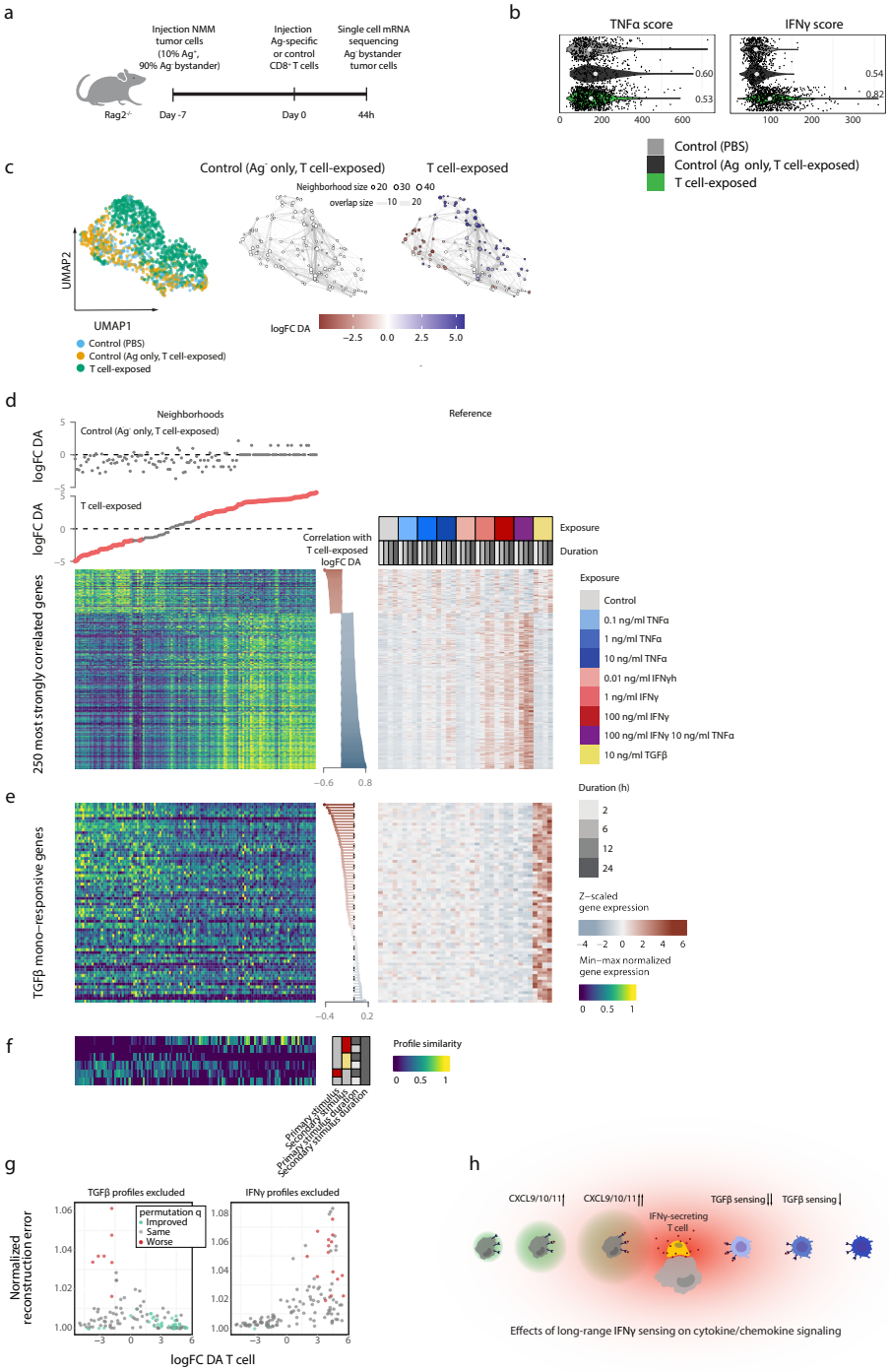

Figure 5.4: Figure 4. Frequent IFN-γ sensing in a syngeneic tumor model and relationship with reduced TGF-β sensing.

a Rag2$^{-/-}$ mice were injected subcutaneously with a mixture of 10% OVA antigen expressing and 90% Ag$^-$ bystander NMM tumor cells, or with Ag$^-$ NMM tumor cells only, and, following tumor establishment, were

treated with either PBS (control) or OT-1 $CD8^+$ T cells, as indicated. Ag^- bystander tumor cells were harvested for single cell RNA-seq analysis 44h after treatment.
b TNF-α and IFN-γ gene set scores, determined using the genes shown in a, for the T cell-exposed condition (green) and the two control conditions (T cell-exposed -Ag^- bystander NMM tumor cells only tumors, and PBS treated tumors, shades of gray). Formatting as in Figure 5.1e.
c Left panel: UMAP of NMM melanoma single cell data, as described in panels **a** & **b**. Middle and right panels: a Milo model[36] was fitted to the data to test for enrichment or depletion for any of the experimental conditions in neighborhoods of transcriptionally similar cells. Non-significantly imbalanced neighborhoods (Spatial FDR > 0.05), as well as homogeneous neighborhoods, are colored white.
d Left panel: heatmap of top 250 genes (rows) most strongly correlated (Spearman correlation) with enrichment for the T cell-exposed condition in cell state neighborhoods (columns) of transcriptionally similar cells. Depicted values are neighborhood averages. Neighborhoods are ordered according to compositional enrichment of cells from the T cell-exposed condition. Top panels show log fold change in differential abundance (logFC DA) for the indicated experimental condition relative to control condition. Right panel: heatmap showing bulk RNAseq gene expression profiles of NMM cells exposed to indicated cytokines for the same genes as in the heatmap in the left panel, ordered identically.
e As in **d**, but for TGF-β responsive genes selected on bulk RNAseq data.
f Deconvolution mixing weights of neighborhoods in an independent bulk RNAseq experiment (Figure S5.5a, right). Neighborhoods ordered as in **d**. Only the 6 out of 28 most highly selected reference profiles are shown, jointly comprising 94% of all assigned similarity.
g (Left) Increase in reconstruction error when the 17 reference profiles with TGF-β are omitted as compared to when all 28 profiles are included. Permutation testing was employed to test whether increase in reconstruction error could be explained by a lower number of reference profiles (Methods). (Right) As left, but omitting the 17 reference profiles with IFN-γ.
h Model visualizing secondary effects of long range IFN-γ sensing. In parallel to the mechanism in which long range IFN-γ sensing leads to generation of, for instance, CXCL9/10/11 chemokine fields and subsequent increased immune cell infiltration, long range IFN-γ sensing may result in secondary changes in the TME by decreasing TGF-β sensing.

Discussion

Next to signaling events induced by direct cell–cell contact, tumor cell behavior is modulated through the sensing of soluble mediators, such as chemokines and cytokines, offering possibilities for long range communication. Here, we describe and validate a single cell sequencing-based approach to identify such long-range communication, and also the secondary changes that are associated with it. Key components of this approach are the generation of bespoke gene sets that report on cell exposure to a given cytokine or cytokine combination, and also the employment of multivariate modeling in case signal-specific reporter genes are unavailable. We demonstrate in both humanized and syngeneic tumor models that $CD8^+$ T cells predominantly modulate the behavior of the tumor mass through IFN-γ release, while no substantial evidence for widespread TNF-α sensing is obtained. This lack of TNF-α sensing is observed in spite of ongoing IFN-γ sensing by a large fraction of the tumor mass, and hence continuous T-cell activity (note that a single intratumoral application of IFN-γ results in just a transient burst of IFN-γ sensing, Figure S5.4d and S5.5a). In addition, we note that bystander tumor cells do retain the capacity to respond to TNF-α *in vivo* when this signal is artifi-

cially provided (Figure 5.3e). In theory, the methodology that we describe is subject to inferential bias in case the studied signals influence cell survival. In our case, such cell survival effects do not form a significant confounder (Figure S5.5c), but it is important to be aware of this possibility when examining other cell models. The cumulative $CD8^+$ T cell-derived IFN-γ and TNF-α levels will vary across tumors, depending on e.g., the fraction of tumor cells presenting relevant antigen and T cell density. Importantly though, we consider it likely that, unless local TME signals would differentially influence production of either cytokine, the ratio between T cell produced TNF-α and IFN-γ will be constant due to their shared dependence on TCR-triggering. Hence, our observation of a differential reach of $CD8^+$ T cell-derived IFN-γ and TNF-α is expected to generalize to unseen settings.

Prior work has demonstrated that exposure to IFN-γ and TNF-α can influence tumor control by, for instance, inhibiting growth of antigen loss variants[13,18,29,30] and modifying behavior of tumor stromal cells[16,38]. Here we demonstrate that, whereas T cell-derived IFN-γ modulates the behavior of a large fraction of antigen-negative cells in the TME, such global effects are not observed for TNF-α. These observations lead us to propose a distinction between cytokines that act as local versus global modifiers of the TME. Importantly, modification of tumor growth through the global TME modifier IFN-γ may be expected to already occur in settings in which T cell activity is heterogeneous and restricted to smaller areas of the tumor cell mass. In contrast, the effect of local TME modifiers such as TNF-α may be most apparent in case of a stronger and more homogeneous intratumoral T cell response. Notably, should other cell types show a drastically different ratio of IFN-γ and TNF-α production, it would be of interest to evaluate differential sensing in these settings.

Finally, the ability to detect sensing of individual cytokines and chemokines makes it possible to determine whether such sensing is associated with additional alterations in cell state. In the current work, we demonstrate that cell states induced by T-cell activity are not only consistent with abundant IFN-γ sensing but also with decreased TGF-β-induced gene expression. Conceivably, intratumoral IFN-γ sensing could result in reduced availability of bioactive TGF-β, for instance through induction of a more pro-inflammatory macrophage state[39,40]. In addition, we provide evidence for mutual negative regulation between the two cytokines (Figure 5.4h). These data add to an emerging view on the role of tumor-reactive $CD8^+$ T cells, in which TCR signaling induced cytokine secretion, and in particular IFN-γ secretion, results in a global alteration of the tumor micromilieu.

Methods

Tumor cell culture and viral transductions

Human ovarian carcinoma OVCAR5 cells (F. Scheeren, The Netherlands Cancer Institute, The Netherlands) and murine $NRAS^{Q61R}$ mutant melanoma cells (NMM) (Norman Sharpless, University of North Carolina, USA,[41]) were cultured at 37°C/ 5% CO_2 in IMDM (Gibco) supplemented with 10%

FCS (Sigma), 100 U/ml penicillin (Roche), 100 μg/ml streptomycin (Roche), and GlutaMax (Gibco, 1x). Identity of OVCAR5 cells was validated by short tandem repeat analysis, STR data were not available for the NMM cell line. The following vectors were utilized: the $CDK4_{R>L}$-GFP-pMX, CMV-pMX , and IGS-PCDH vectors as described in [30], the OVA-mPlumb-pLenti vector as described in [42], and the $CDK4_{R>L}$ specific TCR (clone 17, NKI12)-pMP71 vector, as described in [43]. For retroviral transduction of human cells and mouse cells, FLYRD18 packaging cells (ECACC no. 95091902), and Phoenix-ECO packaging cells (ATCC, CRL-3214), were plated into 6 well plate dishes at 0.5×10^6 cells per well, respectively. After 24h, cells were transfected with 3 μg of one of the indicated retroviral vectors using X-tremeGENE (Roche), according to the manufacturer's protocol. After 48h, virus supernatant was harvested, filtered through a 0.45-μm filter and added to tumor cells in the presence of 8 μg/ml polybrene (Sigma) in a 1:1 dilution in medium. For lentiviral transductions, HEK293T cells (ATCC, CRL-3216) were plated at $3x10^6$ cells per 10 cm dish. After 24h, cells were transfected with 8 μg of one of the above indicated lentiviral plasmids, plus the lentiviral packaging and envelope plasmids psPAX (Addgene #12260) and pMD2.G (Addgene #12259) (3 μg each) using X-tremeGENE (Roche), according to the manufacturer's protocol. 2-3 days after transfection, supernatant of transfected cells was harvested, filtered through 0.45-μm filters, and added to OVCAR5 or NMM cells in a 1:1 dilution in medium. Antigen-positive GFP^+ OVCAR5 cells were generated by retroviral transduction with the pMX-$CDK4_{R>L}$-GFP vector. Antigen-positive $mPlumb^+$ NMM cells were generated by lentiviral transduction with the pLenti-OVA-mPlumb vector. Antigen-negative CFP^+ bystander OVCAR5 and NMM cells were generated by retroviral transduction with the pMX-CFP vector. After transduction, indicated cell populations were sorted on a FACSaria Fusion (BD biosciences) to >90% purity. Ag^-CFP^+IGS reporter cells were generated as described [30].

Generation and culture of TCR-modified T cells

Retroviral transduction and culture of human T cells was performed as described previously [30]. To obtain murine GFP^+ OT-1 $CD8^+$ T cells, spleens from C57BL/6;UBC-GFP;OT-I mice were passed through 70 μm strainers (Falcon) to obtain single cell suspensions. Splenocytes were then negatively enriched with the Mouse CD8 T Lymphocyte Enrichment Set (BD Biosciences) and activated at $1x10^6$ cells per 24 well for 48h with 2μg/ml Concanavalin A (Merck) in RPMI 1640 supplemented with 8% FCS, penicillin/streptomycin, 50μM β-mercapto-ethanol (Gibco), 10 ng/ml IL-2 (Immunotools), 0.5 ng/ml IL-7 (Immunotools) and 1 ng/ml IL-15 (Immunotools). After 48h, cells were spun down and taken up in fresh medium at a concentration $1x10^6$ cells/ml. Cells were kept at a concentration of $1x10^6$/ml, with refreshment of media every 24h, for 2-5 days before adoptive transfer.

In vitro cytokine stimulation

Ag^-CFP^+ tumor cells were plated at 200,000 cells per well in 6-well plates for 24h and were then treated with either human IFN-γ (Invitrogen), human TNF-α (Peprotech), murine IFN-γ (Thermo

Fisher), murine TNF-α (Peprotech) or murine TGF-β (ebioscience) the indicated combination, or were treated with culture medium from T cell -tumor cell co-cultures (see below), at the indicated concentrations or dilutions. At the indicated times, cells were harvested and used for bulk RNA-seq or single cell RNA-seq, as indicated. For bulk RNA-seq, cells were lysed in RLT lysis buffer (Qiagen) and stored at -80°C before sequencing. The bulk mRNA data sets obtained from these *in vitro* cytokine-stimulated cells at different time points are referred to as the "OVCAR5 bulk RNA-seq reference data set" and "NMM bulk RNA-seq reference data set" throughout the manuscript. For single cell RNA-seq, cells were stained with TotalSeq Hashtag antibodies (TotalSeq-B, Biolegend) and pooled, using an equal number of cells from each sample, to form one pool of cells for single cell RNA-seq analysis. Cell death and total cell counts were analyzed at 16h or 44h after treatment by IR-Dye staining and subsequent flow cytometry using AccuCountBlank 15.2-μm beads (Spherotech). To obtain culture medium from T cell -tumor cell co-cultures**,** Ag^+ GFP^+ OVCAR5 tumor cells were plated at $2x10^6$ per 10cm culture dish. After 1 day, $4x10^6$ $CDK4_{R>L}$-specific $CD8^+$ T cells were added, and culture medium was harvested after 24h, filtered through a 0.45 μm filter (GE) and stored at -80°C.

Mice

NOD-scid $Il2r\gamma^{null}$ $\beta2m^{null}$ (NSG-$\beta2m^{-/-}$), C57BL/6;RAG2 KO, C57BL/6;UBC-GFP, and C57BL/6;OT-I mice were obtained from Jackson Laboratories. UBC-GFP and OT-I mice were crossed to obtain GFP-OT-I donor mice for adoptive cell transfer experiments. All animal experiments were approved by the Animal Welfare Committee of The Netherlands Cancer Institute (NKI), in accordance with national guidelines. All animals were maintained in the animal department of NKI, housed in individually ventilated cage (IVC) systems under specific pathogen-free conditions and received food and water ad libitum. Mice were used at 8 to 26 weeks of age.

In vivo tumor models

$8x10^6$ OVCAR5 cells or $2x10^5$ NMM cells were injected subcutaneously into the flank of NSG-$\beta2m^{-/-}$ mice or C57BL/6;$RAG2^{-/-}$ mice, respectively, in 50 μ PBS (Gibco) and 50 μl matrigel (Corning), using the indicated mixtures of tumor cell variants. At day 7-8 after tumor inoculation, tumor-bearing mice received an intravenous injection of either 100μl PBS or 5 x 10^6 $CDK4_{R>L}$ TCR-transduced $CD8^+$ T cells or GFP^+ OT-1 $CD8^+$ T cells in 100 μl PBS, as indicated. On days 0, 1 and 2 after T cell transfer, NSG-$\beta2m^{-/-}$ mice received injections of 7.2 x 10^5 IU IL-2 dissolved in 200 μl PBS, twice daily, with an interval of 6–12 h between injections, to support T-cell engraftment. At the indicated times after T-cell transfer, mice were sacrificed and tumors were harvested. Harvested tumors were manually minced and enzymatically digested in RPMI medium (Gibco) supplemented with 5 Wünsch units/ml TH Liberase/ml (Roche) 25 μg/ml DNase I (Roche) for 20 min at 37 °C under continuous shaking. Subsequently, cell digests were filtered through a 70-μm strainer (Falcon)

and single-cell suspensions were stained with IR-Dye (Invitrogen) and TotalSeq Hastag antibodies (TotalSeq-A or B, Biolegend). Cells from each sample were combined at equal numbers, and CFP^+ (i.e., bystander) tumor cells were sorted from this cell pool on a FACSaria Fusion (BD biosciences) and analyzed by single cell RNA-seq (see below). To measure cytokine production of intratumoral T cells *ex-vivo*, tumor cells from digested tumors were subsequently cultured *in vitro* in the presence of Ag^- CFP^+ tumor cells (for digests from control tumors) or $CDK4_{R>L}$ Ag^+ tumor cells (for digests from Ag^+/Ag^- mixed tumors) for 3 hours in the presence of Golgi-plug (BD biosciences) to block cytokine secretion. Cells were subsequently stained for intracellular IFN-γ and TNF-α and analyzed by flow cytometry.

Intratumoral cytokine injections

Where indicated, tumors of $\geq$150 mm^3 size were intratumorally injected with 15ul PBS containing the indicated cytokines (100 ng IFN-γ, 10 ng TNF-α or 100 ng IFN-γ *plus* 10 ng TNF-α per ml of tumor mass), using a Veo™ insulin syringe with a BD Ultra-Fine™ 6mm x 31G needle (BD biosciences). At the indicated times after injection, tumors were harvested, digested, sorted, and bystander tumor cells were analyzed by single cell RNA-seq as described below.

Flow cytometry

For analysis of immune infiltrates in NSG-$\beta 2m^{-/-}$ mice, cells were stained with fluorochrome-labeled anti-human CD3 antibody (clone OKT3; BD biosciences) in FACS buffer (PBS supplemented with 0.5% w/v bovine serum albumin (Sigma) and EDTA (2 mM, Life Technologies)) for 20–30 min at 4 °C, while protected from light. For analysis of cytokine secretion of intratumoral T cells, cells were stained for anti-human CD3 (clone OKT3; BD biosciences), anti-mouse TCRβ constant domain (clone H57-597; BD Biosciences), anti-human IFN-γ (clone 4S.B3; BD Biosciences) and anti-human TNF-α (clone MAb11, eBioscience). After incubation, cells were washed twice with FACS buffer before resuspension in FACS buffer for analysis. IR-Dye (Invitrogen) was used to allow for live cell selection.

Bulk RNA-seq

Total RNA was isolated using the RNeasy Mini Kit (74106, Qiagen), including an on-column DNase digestion (79254, Qiagen), according to the manufacturer's instructions. RNA quality and quantity was assessed on the 2100 Bioanalyzer instrument, following the manufacturer's instructions "Agilent RNA 6000 Nano" (G2938-90034, Agilent Technologies). Total RNA samples having RIN values>8 were subjected to TruSeq stranded mRNA library preparation, according to the manufacturer's instructions (Document # 1000000040498, Illumina). Stranded mRNA libraries were ana-

lyzed on a 2100 Bioanalyzer instrument, following the manufacturer's protocol "Agilent DNA 7500 kit" (G2938-90024, Agilent Technologies), diluted to 10nM and pooled equimolar into multiplex sequencing pools for sequencing on HiSeq 2500 and NovaSeq 6000 instruments (Illumina). HiSeq 2500 single-end sequencing was performed using 65 cycles for Read 1, and 10 cycles for Read i7, using HiSeq SR Cluster Kit v4 cBot (GD-401-4001, Illumina) and HiSeq SBS Kit V4 50 cycle kit (FC-401-4002, Illumina). NovaSeq 6000 paired-end sequencing was performed using 54 cycles for Read 1, 19 cycles for Read i7, 10 cycles for Read i5, and 54 cycles for Read 2, using the NovaSeq 6000 SP Reagent Kit v1.5 (100 cycles) (20028401, Illumina).

Single cell gene expression library generation and sequencing

Single cell suspensions were diluted to a final concentration of 1,000 cells/μl in 1xPBS containing 0.04% weight/volume BSA. The Chromium Controller platform of 10X Genomics was used for single cell partitioning and barcoding. Per single cell suspension, each cell's transcriptome was barcoded during reverse transcription, pooled cDNA was amplified and Single Cell 3' Gene Expression libraries and Cell Hashing libraries via Feature barcode technology were prepared, according to the manufacturer's protocol (CG000183, CG000206 and CG000317, 10X Genomics). All libraries were quantified on a 2100 Bioanalyzer Instrument following the Agilent Technologies Protocol (Agilent DNA 7500 kit, G2938-90024). Sequence library pools were composed and quantified by qPCR, according to the KAPA Library Quantification Kit Illumina® Platforms protocol (KR0405, KAPA Biosystems). HiSeq 2500, NextSeq 550 or NovaSeq 6000 Illumina sequencing systems were used for paired-end sequencing of the Single Cell 3' Gene Expression libraries and Cell Hashing libraries, respectively, at a sequencing depth of between 20,000-60,000 reads /cell and approximately 3,500 reads/cell. HiSeq 2500 paired-end sequencing was performed using 100 cycles for Read 1, 8 cycles for Read i7, and 100 cycles for Read 2, using HiSeq PE Cluster Kit V4 (PE-401-4001, Illumina) and multiple HiSeq SBS Kit V4 50 cycle kits (FC-401-4002, Illumina). NextSeq 550 paired-end sequencing was performed using 28 cycles for Read 1, 10 cycles for Read i7, and 54 cycles for Read 2, using the NextSeq 500/550 High Output Kit v2.5 (75 Cycles) (20024906, Illumina) and NextSeq 500/550 Mid Output Kit v2.5 (150 Cycles) (20024904, Illumina). Novaseq 6000 paired-end sequencing was performed using 28 cycles for Read 1, 10 cycles for each Read i7 and i5, and 90 cycles for Read 2, using the NovaSeq 6000 S2 Reagent Kit v1.5 (100 cycles) (20028316, Illumina). Gene expression and antibody sequencing reads were mapped to the GRCh38 human reference genome (refdata-cellranger-GRCh38-3.0.0) and antibody reference sequences, respectively, using CellRanger Version 5.0.1 in multi mode (10x Genomics) with default parameters. The genomic sequence of the Katushka fluorescent protein (named as ENSG00000555555) was added to the human reference prior to mapping.

Bulk RNA-seq data preprocessing

Bulk raw read counts of human OVCAR5 samples were mapped to CellRanger's reference transcriptome refdata-cellranger-GRCh38-3.0.0 using version 0.9 of the Nextflow core kallisto pipeline (`https://github.com/cbcrg/kallisto-nf`) with kallisto 0.46.2. Read count distributions were a priori assessed using an awk script (as adapted from `https://www.biostars.org/p/243552/`). Transcript counts were collapsed to genes using the R package txImport (version 1.20.0) and Ensembl gene identifiers (IDs) were converted to HUGO gene names using a home-made lookup table generated from the CellRanger gtf file. Gene read counts were loaded into the R Seurat package (version 4.1.0), then TMM-library size normalized using edgeR (version 3.36.0) and finally additionally corrected using Seurat's regularized negative binomial model regression, in the same manner as the single cell data (SCT normalization[44]). The last step had the effect of setting lowly expressed genes (fewer than 1 unit of TMM-normalized gene expression) to zero and also reduced the contrast in expression between highly and lowly expressed genes. Bulk sequencing data of murine NMM samples were aligned paired-end, strand and transcriptome aware, with hisat2[45] against GRCm38. Counts per gene were made using itreecount (https://github.com/NKI-GCF/itreecount) and annotated using ensembl gtf version 87.

Selection of cytokine-responsive genes

To identify cytokine-responsive genes for the OVCAR5 cell line using machine learning, we first devised a set of gene-characterizing features. A limma voom model[46] of the form $y \sim X_c + X_d$, where y reflects per gene expression levels, Xs are design matrices, c reflect stimulus nature and concentration (one term for each combination of a tested stimulus and concentration), and *d* reflects stimulus exposure duration, was fitted to the 'OVCAR5 bulk RNA-seq reference data set' described in the section '*in vitro cytokine stimulation*' and Figure 5.1. Duration coefficients were included to absorb confounding duration gene expression dynamics that were independent of the nature of stimulus. Interaction terms between stimuli and durations were not included as the design matrices would not be full-rank, as we had exactly one replicate per experimental condition. We collected the following statistics:

1. Maximum t-statistics and effect sizes for the X_c terms were extracted from the fitted limma object.

2. A cytokine specificity score, capturing the relative response to either TNF-α or IFN-γ, which was computed as $|\beta_{\text{TNF-}\alpha}|/(|\beta_{\text{TNF-}\alpha}|+|\beta_{\text{IFN-}\gamma}|)$, where indicated coefficients reflect those of the highest tested concentrations of indicated cytokines (10 ng/ml for TNF-α and 100 ng/ml for IFN-γ). We have not noticed any genes responding more strongly to lower concentrations of a given stimulus, justifying this approach.

3. A set of genes responded more strongly to the combination of IFN-γ and TNF-α than expected based on their response to these stimuli in isolation. To describe this synergistic behavior, a com-

bined TNF-α*plus* IFN-γ synergy gene set score was computed as

$$\beta_{\text{IFN-}\gamma\text{+TNF-}\alpha}/(\beta_{\text{TNF-}\alpha} + \beta_{\text{IFN-}\gamma}) - 1$$

where indicated coefficients again reflect the highest tested concentrations.

4. As some genes responded very strongly to stimuli but only at specific timepoints and such effects would get diluted in the aforementioned limma model (which can be interpreted as an estimated average effect of concentration across the different tested exposure durations), single time point statistics were additionally extracted. For each stimulus and time point, the log2 fold difference with the duration-matched unstimulated control sample was computed, resulting in 3 additional statistics (IFN-γ, TNF-α and IFN-γ *plus* TNF-α), as well as the maximum of these three statistics.

5. Genes for which one stimulus was consistently higher across the four tested timepoints appeared more informative than genes for which this was more variable. To describe this, the maximum number of exposure durations for which any given cytokine and concentration yielded the highest or lowest response was evaluated, yielding another two integral statistics ranging between 1 and 4.

6. With the same goal in mind, the Pearson correlation between all three pairs of consecutive timepoints (2 & 6, 6 & 12, 12 & 24) across the different stimuli was recorded and summarized by the median across the three different sets. High-scoring genes on this metric will have high similarity in the ordering of stimuli in terms of effectuated gene expression across exposure durations. Line plots (as in Figure S5.2c) of such genes will appear ordered, i.e., with a low degree of line crossing.

7. The maximum (log2-transformed) gene expression for each gene in the TMM-normalized expression data across samples was extracted, as well as the difference between the maximum and minimum gene expression values across samples.

The above features were computed for all 33,514 detected genes. As simple thresholding using these statistics gave suboptimal results (data not shown), the set of gene classifications was augmented from an original manually-classified set of genes (n = 80) in an iterative process of i) random forest model fitting on already classified genes, with gene class as the response variable and the aforementioned features as explanatory variables ii) class prediction for previously unclassified genes and iii) manual curation of these model predictions (Figure S5.2a). For step i), classification random forests were trained using the ranger engine (ranger package version 0.13.1) in R the tidymodels library (version 0.1.4) with importance set to 'impurity', using the aforementioned gene descriptive statistics. The mtry parameter was optimized using 3-fold cross validation on a random, unique sample of 75% of the already classified genes, leaving 25% of the genes for validation purposes. A final model was trained on all the training data using optimal hyperparameters. The model was trained to discern between the following gene classes: 'cytokine-unresponsive' (a gene for which none of the evaluated exposures leads to clearly elevated gene expression as judged by inspection of line-plot as in Figure S5.2c) and 'cytokine-responsive' (responsive to at least one cytokine stimulus). The latter class was

then subdivided into the following classes: 'mono-responsive (a cytokine-responsive gene responding strongly to one of the two tested cytokines but not to the other, and for which the combination of TNF-α *plus* IFN-γ does not behave differently from the dominant cytokine), 'synergistically-responsive' (a cytokine-responsive gene responding *solely* to the combination of TNF-α *plus* IFN-γ and not to the individual cytokines), 'other synergy' (a cytokine-responsive gene that shows a moderate degree of stimulus synergy but for which individual cytokines also effectuate noticeable gene expression), 'anti-synergy' (a cytokine-responsive gene whose response to the combination of stimuli is weaker than to the sum of the individual stimuli), 'lowly expressed' (possibly too lowly expressed to be reliably detectable in single cell data) and 'cytokine-responsive, other' (responsive to cytokine exposure, but not fitting to any of the aforementioned class descriptions). Step ii) the final model was used to predict classes for all previously unclassified genes. Step iii) all 'cytokine-responsive' genes (and some 'cytokine-unresponsive' genes as well, to ensure the absence of false negative predictions) were inspected and predictions were adjusted where needed. After 10 iterations of model training and prediction curation, 612 cytokine-responsive genes were acquired (Figure S5.2d), at which point the yield of informative additional genes per additional cycle had slowed down to just a handful, suggesting nearly full extraction of all cytokine-responsive genes. For the purposes of this study, in which the 'other synergy', 'anti-synergy' and 'lowly expressed' classes are superfluous, genes belonging to these classes were next reclassified as 'cytokine-responsive, other' (Figure S5.2b). In addition, the 'mono-responsive' class was partitioned into 'IFN-γ mono-responsive' (genes with cytokine specificity score <= .5) and 'TNF-αmono-responsive' (genes with cytokine specificity score > .5). Additionally, the following 'time-informative' gene sets were compiled by manual sub-selection from all 612 cytokine-responsive genes: IFN-γ late and TNF-α late (cytokine-responsive genes with a most pronounced response at 12-24h of stimulation), IFN-γ plateau (IFN-γ-responsive genes rising in expression until 6 hours, after which they remain constant) and TNF-α early (genes responding most pronounced at 2 hours of stimulation). TNF-α plateau, as well as IFN-γ early, mono-responsive genes could not be identified.

For the selection of genes for the IFN-γ and TNF-α gene sets for the murine NMM cell line, genes were prioritized using a simplification of the analysis done on the human OVCAR5 data. An identical limma model was fitted to the MNN bulk RNA-seq reference data set, a candidate gene list was generated based on the resulting moderated *t*-statistics and candidate genes were then manually filtered for cytokine-unresponsive genes by inspecting their 'line' plots (as in Figure 5.1a), resulting in a total of 134 mono-responsive genes. TGF-β-responsive genes in the Milo analysis (Figure 5.4e) were identified by filtering based on limma moderated *t*-statistics, with a limma model as described above applied (Figure S5.5a), resulting in one *t*-statistic for each of the evaluated single stimuli per gene. Genes were included if they were found mono-responsive to TGF-β, i.e., a) a *t*-statistic surpassing 4.2 for TGF-β and b) below 1 for all other evaluated single stimuli. Genes were then filtered for biological, as opposed to purely technical, expression variation using the modelGeneVar function in the R package scran (package version 1.20.1), which was called with log2(cpm + 1) transformed data and with exposure duration meta information as the function's argument to 'block'. Genes were required

to have bio > 0 and FDR <= 10^-7, resulting in a total of 73 TGF-β mono-responsive genes.

Human OVCAR5 single cell sequencing data preprocessing

CellRanger UMI count data was loaded into R Seurat objects (version 4.1.0), and cells with less than 1,000 detected RNA features were filtered out. Across single cell sequencing experiments, either a single or two hashtag bar codes per sample (i.e., experimental condition) were employed, the latter to allow inclusion of a larger number of samples than the number of available hashtagging antibodies. Sample assignment for experiments employing single hashtagging was done using Seurat's (package version 4.1.0) HTODemux functionality on the CLR-normalized barcode hashtag data using default settings. Sample assignment for experiments employing double hashtagging was done using a custom functionality. First, HTOs were CLR normalized using Seurat's NormalizeData. Next, the product of normalized hashtag counts was computed for each theoretically possible combination of two different hashtag antibodies. Cells were assigned to the sample corresponding to the highest product of CLR-normalized hashtag counts. Low-confidence assignments were then filtered out based on the fold difference between the dominant and second to dominant hashtag combination. A threshold value of 2 for this statistic was picked by comparing cells for which the dominant combination of hashtags was expected (i.e., a combination included in the experimental design) and those for which it was not. Cells were additionally filtered for a maximum mitochondrial content of 30%. SCT total UMI count normalization was performed on the remaining set of cells[44]. We next identified the most variable features in the experiment (Seurat's FindVariableFeatures with default settings) and performed principal component analysis with 10 principal components (sufficient, as indicated by scree plots, obtained using Seurat's RunPCA, default settings) and a UMAP (Seurat's RunUMAP, default settings) over the principal component scores. Next, outlying clusters in the UMAP were automatically identified using density-based clustering on the UMAP cell coordinates with DBSCAN (fpc package, version 2.2.9) using parameters: eps = .6 and MinPts = 6. Small DBSCAN clusters (less than $1/(5c)$ cells, where c is the number of detected DBSCAN clusters) were then marked as outlying clusters and removed from downstream analyses. Using differential gene expression (DGE) analysis (Seurat's FindMarkers) and GSEA with (R 'fgsea' package, version 1.18.0) with REACTOME pathways[47] on the logFC-ranked list of differentially expressed genes (filtered first FDR-adjusted p-value <= .1) between the outlying cluster and the main body of cells (i.e., the composite of non-outlying clusters), we identified these outlying clusters (0% -2.7% of cells across experiments) to likely consist of keratinocytes and/or fibroblasts, characterized by high *KRT81, S100A9* and *LDHB* counts. SCT normalization was redone on the remaining cells. Using the intersection of cytokine-responsive genes and detectable genes for each experiment, PCA was recomputed with 10 PCs and UMAPs were recomputed based on the PC scores.

Murine NMM single cell sequencing data preprocessing

The murine NMM data were preprocessed identically to the human OVCAR5 single cell data. Inspecting the initial UMAP, we noticed two clusters of cells. Characterizing the smaller cluster (DBSCAN cluster 2, 21.5% of cells), we noticed a 7.2-fold difference in mean UMIs between clusters (47,753 versus 6,654 mean UMI), suggestive of the presence of dying cells and/or cell fragments in the second cluster. Cells in this cluster were then removed, and SCT normalization, PCA and UMAP computation was repeated on the remaining cells.

Heatmap visualization of bulk RNA-seq data

TMM and SCT normalized data were log2-transformed and subsequently Z-scaled across genes/features. Unless indicated otherwise, rows (genes) and columns (samples) were subsequently clustered using complete linkage hierarchical clustering (hclust function in R), using Spearman correlation distance (computed as $1 - c$, where c is the correlation between two samples) for genes (rows) and Euclidean distance for samples (columns).

Gene set scores for single cell data

For both OVCAR5 and NMM cell lines, gene set scores were computed by summing SCT-normalized expression values of a gene set's member genes. Where indicated, gene set scores were normalized to the distribution of gene set scores of a relevant, duration-matched control condition with the following transformation: $f(x) = x - m(x_c)/IQR(x_c)$, where m represents the median, x_c is a vector of scores for the same gene set for a reference condition and $IQR(x_c)$ is the interquartile range in that reference population.

Quantification of separability of experimental conditions using gene set scores

AUROC values in main text and figures represent the area under the receiver operator curve from a binary classifier aiming to separate the two indicated experimental conditions with the indicated gene set score as the sole explanatory variable. These scores were computed using the roc_auc_vec function in the R yardstick package (version 0.0.9). AUROC values range between 0.5 (signifying no separation between experimental conditions) and 1 (signifying complete separation between experimental conditions). To distinguish more than two classes at a time, as in Figure 5.2b, we trained SVM models using the Gaussian kernel on 75% of the data, leaving the remaining 25% as validation data, with the R tidymodels framework (package version 0.1.4). Hyperparameters cost and rbf_sigma were optimized in a 10-fold cross validation within the training data. Optimal parameters were then used to train a final model on the full set of training data, of which the performance was sub-

sequently evaluated on the validation data. Included confusion matrices show the correspondence between actual and predicted class labels in the validation data, using a model trained on independent data, ignoring the fact that single cell transcriptomes in the training and validation sets were jointly preprocessed.

Milo neighborhood analysis of single cell data

A PCA of the murine single cell data was computed using the 2,000 most variable genes (as identified using Seurat's FindVariableGenes with default settings) and 20 principal components. A KNN graph of the data was constructed on the principal component scores with Milo's (package version 1.0.0) buildGraph function using k = 10 and d = 20. Neighborhoods were defined using makeNhoods, with refined= TRUE, prop = 0.1 and identical settings for k and d as aforementioned. Neighborhoods were tested for differential abundance of the three experimental conditions using the testNhoods function, with a design matrix that was obtained from a metadata table with the following R model formula: ~1 + experimental condition, resulting in an intercept term and a regression coefficient for the second and third experimental conditions. In visualizations of the KNN graph, effect sizes (logFC DA) associated with a Spatial FDR > 0.05 were whitened.

Deconvolution analysis of single cell neighborhood expression

A gene expression matrix of g genes (rows) by n single cell neighborhoods (SCNs, columns) in linear space was extracted using the Milo nhoodExpression class method, TMM-transformed to normalize library sizes and finally log10-transformed, i.e., $M_N = log_{10}(TMM(X) + 1)$, wherein X is the output of nhoodExpression and M_N is the processed matrix of SCN expression. Similarly, a $g\ x\ 28$ reference profile matrix M_R consisting of bulk reference RNASeq libraries with the 28 conditions/columns unstimulated 2h -unstimulated 24h (i.e., no stimulation for 2 hours followed by no stimulation for another 24 hours), unstimulated 2h -10 ng/ml TGFb 24h, unstimulated 2h -100 ng/ml IFNy 24h, 10 ng/ml TGFb 2h -unstimulated 24h, 10 ng/ml TGFb 2h -100 ng/ml IFNy 24h, 100 ng/ml IFNy 2h -unstimulated 24h, 100 ng/ml IFNy 2h -10 ng/ml TGFb 24h, unstimulated 6h -unstimulated 24h, unstimulated 6h -10 ng/ml TGFb 24h, unstimulated 6h -100 ng/ml IFNy 24h, 10 ng/ml TGFb 6h -unstimulated 24h, 10 ng/ml TGFb 6h -100 ng/ml IFNy 24h, 100 ng/ml IFNy 6h -unstimulated 24h, 100 ng/ml IFNy 6h -10 ng/ml TGFb 24h, unstimulated 12h -unstimulated 24h, unstimulated 12h -10 ng/ml TGFb 24h, unstimulated 12h -100 ng/ml IFNy 24h, 10 ng/ml TGFb 12h -unstimulated 24h, 10 ng/ml TGFb 12h -100 ng/ml IFNy 24h, 100 ng/ml IFNy 12h -unstimulated 24h, 100 ng/ml IFNy 12h -10 ng/ml TGFb 24h, unstimulated 24h -unstimulated 24h, unstimulated 24h -10 ng/ml TGFb 24h, unstimulated 24h -100 ng/ml IFNy 24h, unstimulated 24h -100 ng/ml IFNy 10 ng/ml TGFb 24h, 10 ng/ml TGFb 24h -unstimulated 24h, 10 ng/ml TGFb 24h -100 ng/ml IFNy 24h, 100 ng/ml IFNy 24h -unstimulated 24h, 100 ng/ml IFNy 24h -10 ng/ml TGFb 24h in the columns and genes in the rows) were TMM-normalized and then similarly log10-transformed. The

matrices were identically row-ordered (genes), with as rows the intersection of detected genes in the two unprocessed source matrices ($g = 14,620$) such that they predominantly consisted of genes that are minimally or not responsive to the cytokines of interest.

To estimate a SCN's stimulus exposure, its transcriptome was modeled as a linear combination of the columns (samples) of M_R using lasso *l1*-penalized multivariate regression with the regression coefficients (β) constrained to be larger than or equal to 0, such that they can be interpreted as mixing weights. The cv.glmnet function from the R package glmnet was used for this regression[48], as well as to optimize the lambda penalty for regression coefficients with arguments: lower.limits=c(0), family='gaussian', alpha=.99. The mean squared error between the original transcriptome (y) and reconstructed transcriptome ($x\ \beta$), i.e,. $E[(y - x\ \beta)^2]$, was then extracted from the cv.glmnet object with lambda set to the value that minimized the cross validation error. To assess the importance of a subset of reference samples (columns) c for the reconstruction of any particular SCN, the reconstruction error was recomputed with the remaining columns of M_R after having removed samples/columns c. In this, a large increase in error can be interpreted as an indication that the samples c contain co-variation in gene expression that cannot be accommodated by any of the remaining reference samples. To quantify the possibility that reconstruction error was simply raised by providing the optimization algorithm a smaller set of basis vectors to work with (i.e., by limiting the span of M_R), permutation testing was employed. Specifically, the reconstruction error was first computed with 1,000 random selections of s reference profiles, where s is the number of reference profiles that were *not* associated with the reference profiles (columns) of which the importance is assessed. For example, to assess the importance of TGF-β stimulation, *s* is the number of profiles obtained from experiments where no TGF-β stimulation was applied (i.e. only unstimulated and IFN-γ stimulated samples, $s = 11$, Figure 5.4g, left). Similarly, to assess the importance of IFN-γ stimulation, *s* represents the number of profiles obtained from experiments where no IFN-γ stimulation was applied (i.e., only unstimulated and TGF-β-stimulated samples, $s = 11$, Figure 5.4g, right). The quantile of the observed error in the distribution of permutation errors was then acquired using the R function ecdf. SCN reconstruction was finally labeled as worse (or better) than expected if the observed error was in the 97.5st or higher (or 2.5th or lower) percentile of the permutation distribution.

Software and code availability

All preprocessing and analyses were done in R (version 4.1.0), frequently employing tidyverse packages (version 1.3.1). Row-or column-annotated heatmaps were made using the ComplexHeatmap package (version 2.9.3). Remaining plots were made using the ggplot2 package (version 3.3.5). All data processing and time-intensive computation was done inside of the R targets pipelining system[49]. Code and notebooks can be accessed at
www.github.com/slagtermaarten/cyto_inference.

Acknowledgements

We thank K. Bresser, A. M. van der Leun, L. Kok and S. Mourragui for input and valuable discussions. We thank staff of the NKI Genomics Core facility for technical support, along with the NKI Research High Performance Computing, Animal Intervention, and Flow Cytometry facilities. This work was supported by a Boehringer Ingelheim Fonds PhD Fellowship (to M. E. H.), ERC AdG SENSIT, grant agreement ID 742259 (to T. N. M. S.) and institutional funding of the Netherlands Cancer Institute by the Dutch Cancer Society.

Author contributions

M.E.H. conceived the study together with M.S. and T.N.S, performed wet lab experiments and wrote the manuscript with M.S., L.F.A.W. and T.N.S.. M.S. performed and interpreted bioinformatic analyses. J.U. & M.T. aided in wet lab experiments. N.S. performed exploratory data analysis. R.K and M.N, designed and performed sequencing experiments and I.R. & R.J.C.K performed data preprocessing. L.F.A.W. and T.N.S. supervised bioinformatic analyses, and T.N.S. supervised wet-lab experiments. All authors have read and approved the manuscript.

Declaration of interests

L.F.A.W. received project funding for unrelated work from Bristoll-Myers-Squibb. T.N.S. is advisor for Allogene Therapeutics, Asher Bio, Merus, Neogene Therapeutics, and Scenic Biotech; is a stockholder in Allogene Therapeutics, Asher Bio, Cell Control, Celsius, Merus, and Scenic Biotech; and is venture partner at Third Rock Ventures, all outside of the current work.

Bibliography

1. Giladi, A. *et al.* Dissecting cellular crosstalk by sequencing physically interacting cells. *Nat Biotechnol* **38**, 629-637, doi:10.1038/s41587-020-0442-2 (2020).

2. Altan-Bonnet, G. & Mukherjee, R. Cytokine-mediated communication: a quantitative appraisal of immune complexity. *Nat Rev Immunol* **19**, 205-217, doi:10.1038/s41577-019-0131-x (2019).

3. Raskov, H., Orhan, A., Christensen, J. P. & Gogenur, I. Cytotoxic CD8(+) T cells in cancer and cancer immunotherapy. *Br J Cancer* **124**, 359-367, doi:10.1038/s41416-020-01048-4 (2021).

4. Rosenberg, S. A. *et al.* Durable complete responses in heavily pretreated patients with metastatic melanoma using T-cell transfer immunotherapy. *Clin Cancer Res* **17**, 4550-4557, doi:10.1158/1078-0432.CCR-11-0116 (2011).

5. Dudley, M. E. *et al.* Randomized selection design trial evaluating CD8+-enriched versus unselected tumor-infiltrating lymphocytes for adoptive cell therapy for patients with melanoma. *J Clin Oncol* **31**, 2152-2159, doi:10.1200/JCO.2012.46.6441 (2013).

6. Morotti, M. *et al.* Promises and challenges of adoptive T-cell therapies for solid tumours. *Br J Cancer* **124**, 1759-1776, doi:10.1038/s41416-021-01353-6 (2021).

7. Mojic, M., Takeda, K. & Hayakawa, Y. The Dark Side of IFN-gamma: Its Role in Promoting Cancer Immunoevasion. *Int J Mol Sci* **19**, doi:10.3390/ijms19010089 (2017).

8. Castro, F., Cardoso, A. P., Goncalves, R. M., Serre, K. & Oliveira, M. J. Interferon-Gamma at the Crossroads of Tumor Immune Surveillance or Evasion. *Front Immunol* **9**, 847, doi:10.3389/fimmu.2018.00847 (2018).

9. Mosser, D. M. & Edwards, J. P. Exploring the full spectrum of macrophage activation. *Nat Rev Immunol* **8**, 958-969, doi:10.1038/nri2448 (2008).

10. Jorgovanovic, D., Song, M., Wang, L. & Zhang, Y. Roles of IFN-gamma in tumor progression and regression: a review. *Biomark Res* **8**, 49, doi:10.1186/s40364-020-00228-x (2020).

11. Parameswaran, N. & Patial, S. Tumor necrosis factor-alpha signaling in macrophages. *Crit Rev Eukaryot Gene Expr* **20**, 87-103, doi:10.1615/critreveukargeneexpr.v20.i2.10 (2010).

12. Trevejo, J. M. *et al.* TNF-alpha -dependent maturation of local dendritic cells is critical for activating the adaptive immune response to virus infection. *Proc Natl Acad Sci U S A* **98**, 12162-12167, doi:10.1073/pnas.211423598 (2001).

13. Braumuller, H. *et al.* T-helper-1-cell cytokines drive cancer into senescence. *Nature* **494**, 361-365, doi:10.1038/nature11824 (2013).

14. Montfort, A. *et al.* The TNF Paradox in Cancer Progression and Immunotherapy. *Front Immunol* **10**, 1818, doi:10.3389/fimmu.2019.01818 (2019).

15. Wang, W. *et al.* CD8(+) T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature* **569**, 270-274, doi:10.1038/s41586-019-1170-y (2019).

16. Kammertoens, T. *et al.* Tumour ischaemia by interferon-gamma resembles physiological blood vessel regression. *Nature* **545**, 98-102, doi:10.1038/nature22311 (2017).

17. Briesemeister, D. *et al.* Tumor rejection by local interferon gamma induction in established tumors is associated with blood vessel destruction and necrosis. *Int J Cancer* **128**, 371-378, doi:10.1002/ijc.25350 (2011).

18. Zhang, B., Karrison, T., Rowley, D. A. & Schreiber, H. IFN-gamma- and TNF-dependent bystander eradication of antigen-loss variants in established mouse cancers. *J Clin Invest* **118**, 1398-1404, doi:10.1172/JCI33522 (2008).

19. Spiotto, M. T., Rowley, D. A. & Schreiber, H. Bystander elimination of antigen loss variants in established tumors. *Nat Med* **10**, 294-298, doi:10.1038/nm999 (2004).

20. Spiotto, M. T. & Schreiber, H. Rapid destruction of the tumor microenvironment by CTLs recognizing cancer-specific antigens cross-presented by stromal cells. *Cancer Immun* **5**, 8 (2005).

21. Huse, M., Lillemeier, B. F., Kuhns, M. S., Chen, D. S. & Davis, M. M. T cells use two directionally distinct pathways for cytokine secretion. *Nat Immunol* **7**, 247-255, doi:10.1038/ni1304 (2006).

22. Kupfer, A., Mosmann, T. R. & Kupfer, H. Polarized expression of cytokines in cell conjugates of helper T cells and splenic B cells. *Proc Natl Acad Sci U S A* **88**, 775-779, doi:10.1073/pnas.88.3.775 (1991).

23. Sanderson, N. S. *et al.* Cytotoxic immunological synapses do not restrict the action of interferon-gamma to antigenic target cells. *Proc Natl Acad Sci U S A* **109**, 7835-7840, doi:10.1073/pnas.1116058109 (2012).

24. Hoekstra, M. E., Vijver, S. V. & Schumacher, T. N. Modulation of the tumor microenvironment by CD8(+) T cell-derived cytokines. *Curr Opin Immunol* **69**, 65-71, doi:10.1016/j.coi.2021.03.016 (2021).

25. Perona-Wright, G., Mohrs, K. & Mohrs, M. Sustained signaling by canonical helper T cell cytokines throughout the reactive lymph node. *Nat Immunol* **11**, 520-526, doi:10.1038/ni.1866 (2010).

26. Ariotti, S. *et al.* T cell memory. Skin-resident memory CD8(+) T cells trigger a state of tissue-wide pathogen alert. *Science* **346**, 101-105, doi:10.1126/science.1254803 (2014).

27. Schenkel, J. M. *et al.* T cell memory. Resident memory CD8 T cells trigger protective innate and adaptive immune responses. *Science* **346**, 98-101, doi:10.1126/science.1254536 (2014).

28. Muller, A. J. *et al.* CD4+ T cells rely on a cytokine gradient to control intracellular pathogens beyond sites of antigen presentation. *Immunity* **37**, 147-157, doi:10.1016/j.immuni.2012.05.015 (2012).

29. Thibaut, R. *et al.* Bystander IFN-gamma activity promotes widespread and sustained cytokine signaling altering the tumor microenvironment. *Nat Cancer* **1**, 302-314, doi:10.1038/s43018-020-0038-2 (2020).

30. Hoekstra, M. E. *et al.* Long-distance modulation of bystander tumor cells by CD8(+) T cell-secreted IFNgamma. *Nat Cancer* **1**, 291-301, doi:10.1038/s43018-020-0036-4 (2020).

31. Beck, R. J., Slagter, M. & Beltman, J. B. Contact-Dependent Killing by Cytotoxic T Lymphocytes Is Insufficient for EL4 Tumor Regression In Vivo. *Cancer Res* **79**, 3406-3416, doi:10.1158/0008-5472.CAN-18-3147 (2019).

32. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425, doi:10.1016/j.cels.2015.12.004 (2015).

33. Buttner, M., Miao, Z., Wolf, F. A., Teichmann, S. A. & Theis, F. J. A test metric for assessing single-cell RNA-seq batch correction. *Nat Methods* **16**, 43-49, doi:10.1038/s41592-018-0254-1 (2019).

34. Fukunaga, R., Sokawa, Y. & Nagata, S. Constitutive production of human interferons by mouse cells with bovine papillomavirus as a vector. *Proc Natl Acad Sci U S A* **81**, 5086-5090, doi:10.1073/pnas.81.16.5086 (1984).

35. Savan, R., Ravichandran, S., Collins, J. R., Sakai, M. & Young, H. A. Structural conservation of interferon gamma among vertebrates. *Cytokine Growth Factor Rev* **20**, 115-124, doi:10.1016/j.cytogfr.2009.02.006 (2009).

36. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat Biotechnol* **40**, 245-253, doi:10.1038/s41587-021-01033-z (2022).

37. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* **12**, 453-457, doi:10.1038/nmeth.3337 (2015).

38. Garnier, L. *et al.* IFN-gamma-dependent tumor-antigen cross-presentation by lymphatic endothelial cells promotes their killing by T cells and inhibits metastasis. *Sci Adv* **8**, eabl5162, doi:10.1126/sciadv.abl5162 (2022).

39. Boutilier, A. J. & Elsawa, S. F. Macrophage Polarization States in the Tumor Microenvironment. *Int J Mol Sci* **22**, doi:10.3390/ijms22136995 (2021).

40. Wen, F. Q. *et al.* Interleukin-4- and interleukin-13-enhanced transforming growth factor-beta2 production in cultured human bronchial epithelial cells is attenuated by interferon-gamma. *Am J Respir Cell Mol Biol* **26**, 484-490, doi:10.1165/ajrcmb.26.4.4784 (2002).

41. Burd, C. E. *et al.* Mutation-specific RAS oncogenicity explains NRAS codon 61 selection in melanoma. *Cancer Discov* **4**, 1418-1429, doi:10.1158/2159-8290.CD-14-0729 (2014).

42. Vredevoogd, D. W. *et al.* Augmenting Immunotherapy Impact by Lowering Tumor TNF Cytotoxicity Threshold. *Cell* **178**, 585-599 e515, doi:10.1016/j.cell.2019.06.014 (2019).

43. van Rooij, N. *et al.* Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma. *J Clin Oncol* **31**, e439-442, doi:10.1200/JCO.2012.47.7521 (2013).

44. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**, 296, doi:10.1186/s13059-019-1874-1 (2019).

45. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-915, doi:10.1038/s41587-019-0201-4 (2019).

46. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29, doi:10.1186/gb-2014-15-2-r29 (2014).

47. Gillespie, M. *et al.* The reactome pathway knowledgebase 2022. *Nucleic Acids Res* **50**, D687-D692, doi:10.1093/nar/gkab1028 (2022).

48. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* **33**, 1-22 (2010).

49. Landau, W. M. The targets R package: a dynamic Make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. . *Journal of Open Source Software* **6(57)**, doi: 10.21105/joss.02959 (2021).

Supplemental Figures

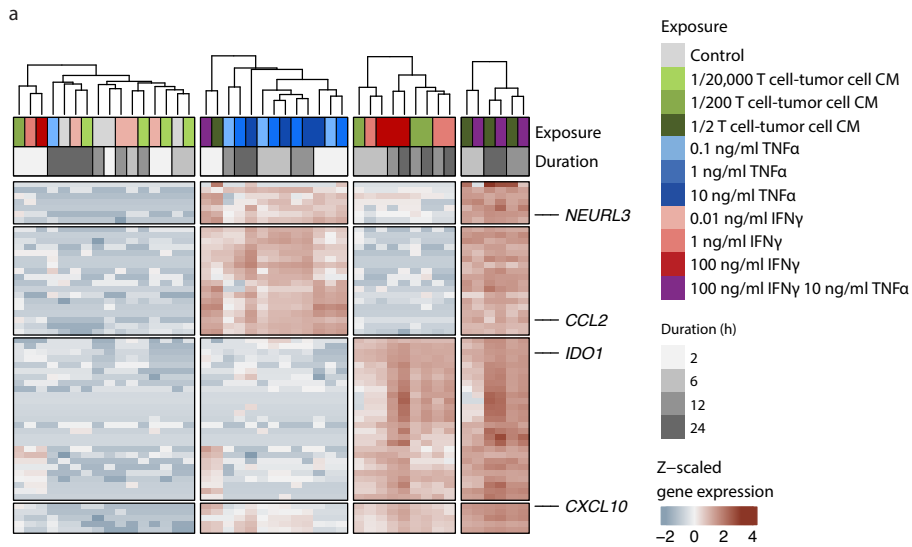

Figure S5.1: Gene expression profiles of tumor cells exposed to IFN-γ *plus* TNF-α or T cell - tumor cell co-culture medium.
OVCAR5 tumor cells were stimulated with recombinant IFN-γ, TNF-α, or IFN-γ*plus* TNF-α (as in Figure 5.1b), or with tissue culture medium (CM) obtained from co-cultures of $CDK4_{R>L}$ Ag^+ OVCAR5 tumor cells and $CDK4_{R>L}$-specific $CD8^+$ T cells, for the indicated times and concentrations. Unsupervised hierarchical clustering of bulk RNAseq data, with genes filtered for a variance of at least 0.4 across the included samples in the library size normalized but unscaled representation of the data. Note that T cell-tumor co-culture medium-exposed samples co-cluster with TNF-α*plus* IFN-γ-stimulated samples.

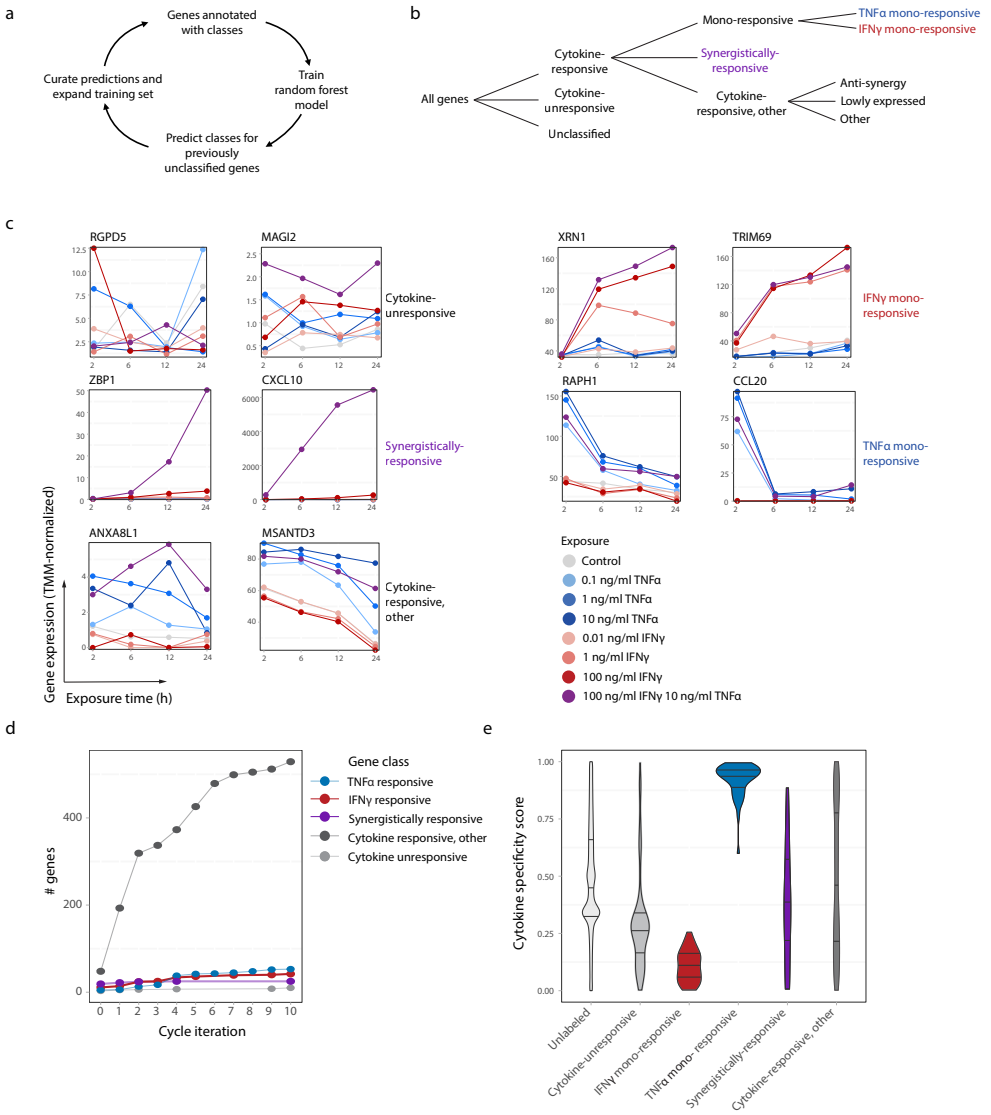

Figure S5.2: Model prediction and curation cycle to identify cytokine-responsive genes.
Starting with a small set of handpicked genes, a machine learning model trained on this initial seed set was used to efficiently identify additional cytokine-responsive genes of the indicated classes. Line plots (as shown in c) of each gene predicted to be cytokine-responsive were manually inspected and gene classes were corrected where necessary. **a** Hierarchy of gene class assignment. Members of the 'TNF-α mono-responsive' (blue) and 'IFN-γ mono-responsive' (red) gene classes were used to calculate cytokine gene set scores. Synergistically-responsive genes (purple), responding solely to the combination of TNF-α *plus* IFN-γ and not to the individual cytokines, form the 'synergy gene set' used to calculate synergy gene set scores.
b RNA expression profiles of two randomly selected members of the indicated gene classes, as described in b. Expression profiles depict bulk RNAseq data from OVCAR5 cells exposed *in vitro* to the indicated concentration of recombinant IFN-γ, TNF-αor IFN-γ *plus* TNF-α, for the indicated duration. **c** Gene counts per class after each prediction and curation cycle. Flattening of the curves with increasing numbers of cycles suggests that most cytokine-responsive genes have been identified. **d** The cytokine specificity score (See Methods section

'Selection of cytokine-responsive genes'), quantifying differential transcriptional responsiveness to either TNF-α (value of 1) or IFN-γ (value of 0), for members of the indicated gene classes.

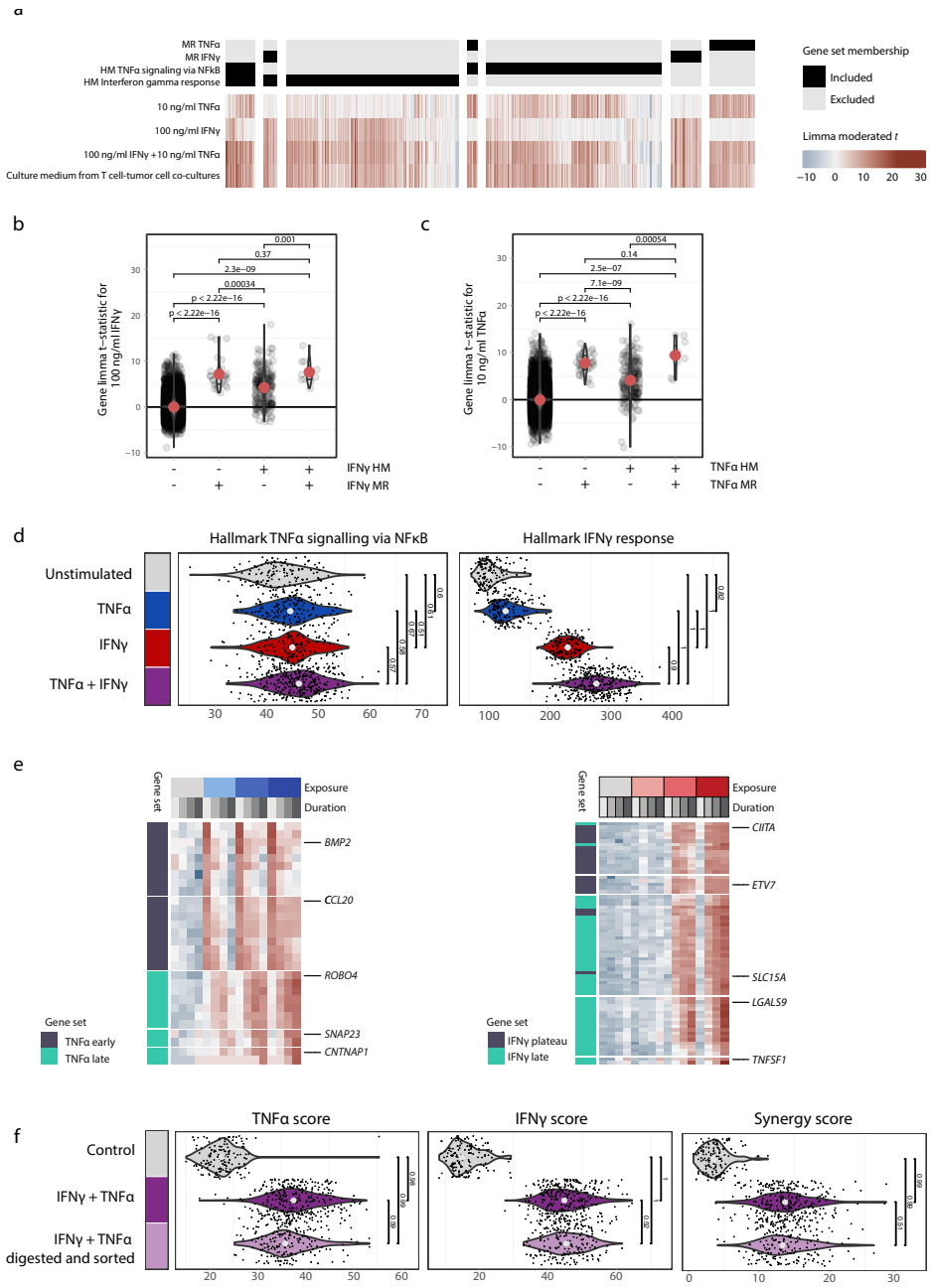

Figure S5.3: Validation and characterization of cytokine-responsive genes in bulk RNA-seq data. **a** Limma moderated *t*-statistics for activating concentrations of cytokines (100ng/ml IFN-γ, 10ng/ml TNF-α) in the OVCAR5 bulk RNAseq reference data set (See methods section *'In vitro cytokine stimulation'* and Figure

5.1). Depicted genes are present in one of the mono-responsive (MR) gene sets identified in this work and/or in one or both of the Hallmark 'Interferon gamma response' and 'TNF-αsignaling via NFkB' gene sets [32]. Note that whereas MR genes were selected based on their response to exactly one stimulus in OVCAR5 cells (among the cytokines tested), the majority of Hallmark genes did not qualify as MR, frequently responding to multiple cytokines, or not responding to the purported stimulus in OVCAR5 cells.
b Comparison of limma moderated *t*-statistics between genes that are in- or excluded from the Hallmark Interferon gamma response (HM) gene set and in- or excluded from the OVCAR5-specific IFN-γ MR set. Black dots represent individual genes; red dots denote medians. Displayed comparisons are Wilcoxon rank sum test p-values.
c As in **b**, but with the Hallmark 'TNF-α signaling via NFkB' and OVCAR5 TNF-α MR gene sets.
d As in Figure 5.1e, but with the Hallmark gene sets instead of the OVCAR5 customized ones. The 'TNF-α signaling via NFkB' set shows virtually no response to TNF-α stimulation, whereas the 'Interferon gamma response' responds more strongly to the combination of TNF-α and IFN-γ than it does to IFN-γ alone.
e Heatmaps of time-informative genes for TNF-α (left panel) and IFN-γ (right panel). 'Early' genes reach maximal expression after 2 hours of cytokine exposure, with reduced expression at later timepoints. 'Plateau' genes reach maximal expression at 6 hours of cytokine exposure, with relatively constant expression at later timepoints. 'Late' genes show a continuing increase in expression up to 24 hours of cytokine exposure.
f Violin plots depicting gene set scores of *in vitro* cultured and cytokine stimulated (100 ng/ml IFN-γ and 10 ng/ml TNF-α for 24h) OVCAR5 cells that were or were not exposed to single cell digestion and fluorescence activated cell sorting. Dots represent gene set scores of individual cells, violins represent densities of score distributions. Values represent AUROC values, denoting the separability of cells under the indicated conditions (with 0.5 indicating no separation and 1 indicating a complete separation).

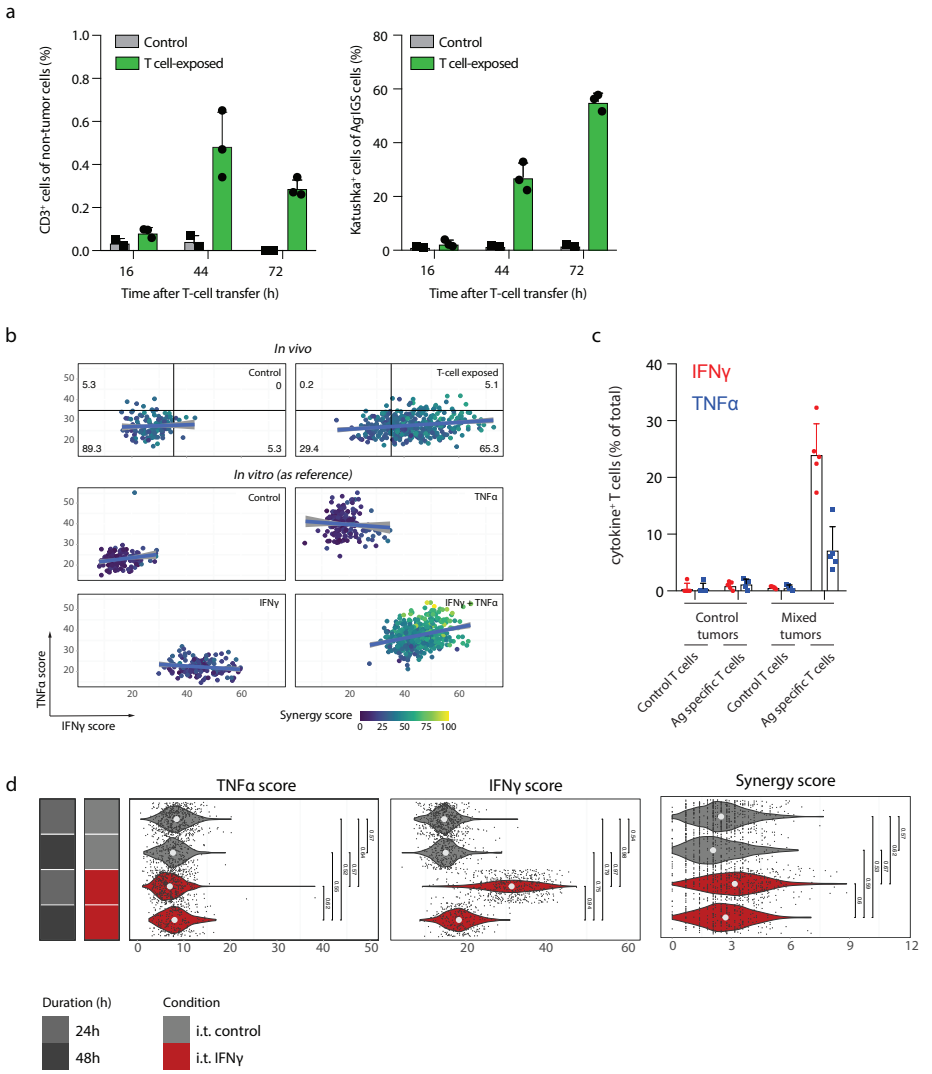

Figure S5.4: Quantification of T-cell infiltration and functionality *in vivo*.
a NSG-β2m$^{-/-}$ mice were injected subcutaneously with a mixture of 10% CDK4$_{R>L}$ antigen expressing (Ag^{+}) OVCAR5 cells and 90% Ag^{-} CFP^{+} bystander OVCAR5 tumor cells carrying a fluorescence based IFN-γ sensing (IGS) reporter. After tumor establishment, tumors were treated by intravenous injection of PBS (control) or CDK4$_{R>L}$ specific CD8^{+} T cells, and tumors were harvested at the indicated timepoint after treatment. Left panel: percentage of CD3^{+} cells of non-tumor cells, as measured by flow cytometry. Right panel: fraction of Ag^{-} bystander tumor cells with an activated IGS reporter (katushka^{+}) from the total bystander tumor cell population, as measured by flow cytometry.
b Upper two panels: scatter plots of TNF-α and IFN-γ gene set scores from OVCAR5 cells derived from control and T cell-exposed tumors. The horizontal and vertical lines demarcate the 95th percentiles of the *in vivo*

control condition for the TNF-α gene set score and IFN-γ gene set score, respectively. Bottom four panels: scatter plots of TNF-α and IFN-γ gene set scores from OVCAR5 cells that have been stimulated *in vitro* with recombinant cytokines, as described in Figure 5.1, serving as a reference for the two top panels. Note that the synergy gene set score (color scale) is selectively elevated in cells that show high expression of both the TNF-α and IFN-γ mono-responsive gene sets.

c Tumor-infiltrated T cells retain the capacity to produce IFN-γ and TNF-α. NSG-β2m$^{-/-}$ mice were injected subcutaneously with either a mixture of 10% antigen expressing (Ag^{+}) OVCAR5 cells and 90% Ag^{-} CFP^{+} bystander OVCAR5 tumor cells (mixed tumors) or with Ag^{-} CFP^{+} tumor cells (control tumors). After tumor establishment, mice were treated by intravenous injection of a mixture of control CD8^{+} T cells (40%) and specific CD8^{+} T cells (60%), and tumors were harvested 44h after treatment. Digested tumors were subsequently cultured in vitro in the presence of Ag^{-} CFP^{+} tumor cells (for digests from control tumors) or Ag^{+} tumor cells (for digests from Ag^{+}/Ag^{-} mixed tumors) for 3 hours in the presence of Golgi-plug to evaluate capacity for continued cytokine production. Subsequently, cells were stained for intracellular IFN-γ and TNF-α and analyzed by flow cytometry. Each dot represents T cells derived from one tumor, bar graphs show mean of the indicated groups + SD, *n*=5 mice per group. Data of one experiment is depicted. Note that specific CD8^{+} T cells derived from mixed tumors retain the capacity to produce TNF-α and IFN-γ.

d Gene set scores of OVCAR5 tumor cells obtained from subcutaneous OVCAR5 tumors isolated 24h or 48h after intratumoral injection of IFN-γ or PBS (control). Dots represent gene set scores of individual cells, violins represent densities of score distributions. Values represent AUROC values, denoting the separability of cells under the indicated conditions (with 0.5 indicating no separation and 1 indicating a complete separation). Note that a single injection of IFN-γ leads to a transient increase in IFN-γ gene set scores that is nearly reversed at 48h.

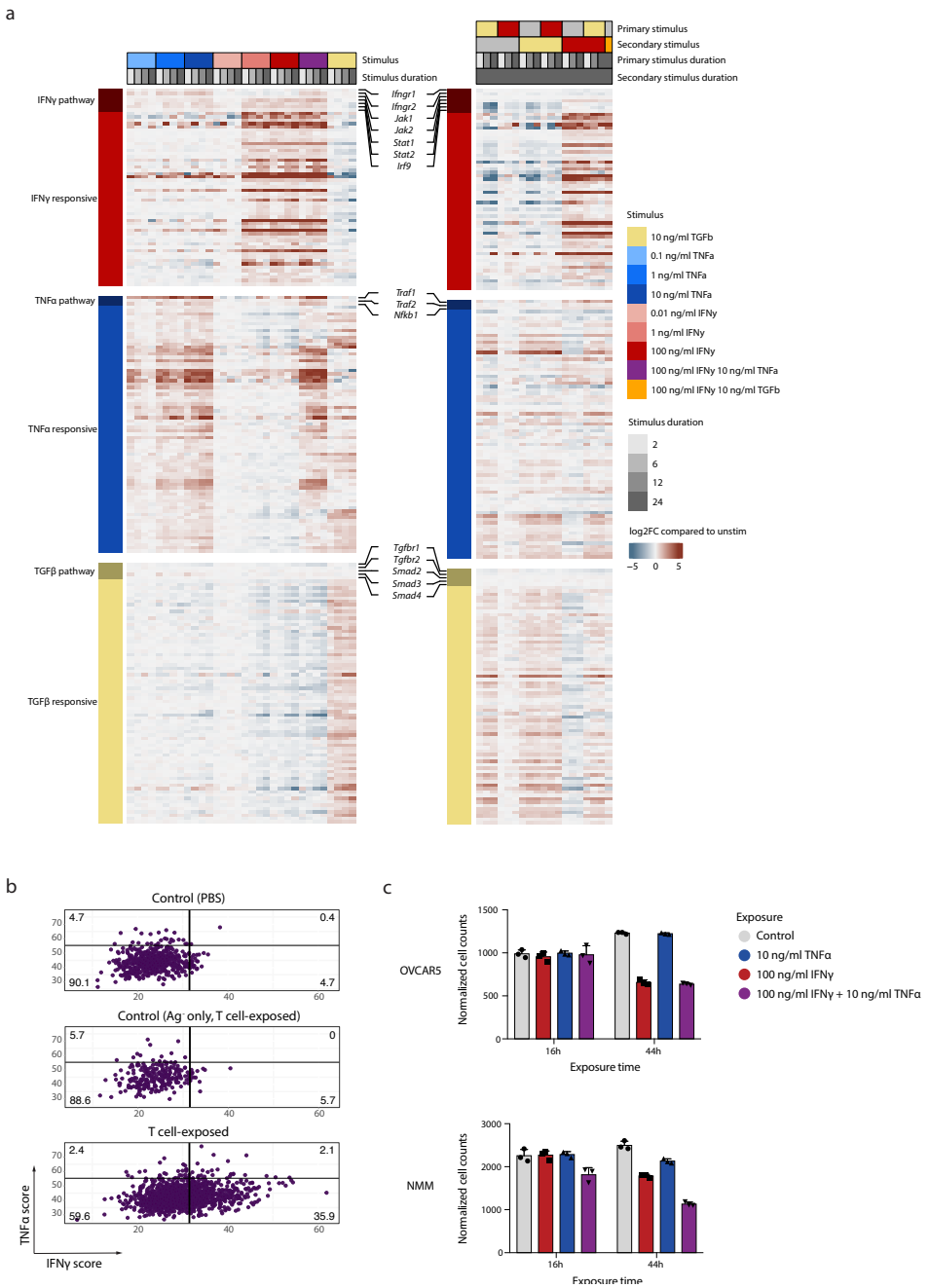

Figure S5.5: Characterization of cytokine-responsiveness in a syngeneic tumor model.
a Heatmap of bulk RNAseq data of murine NMM melanoma cells exposed *in vitro* to the indicated cytokines. Left and right show two separate RNAseq experiments, with different (combinations of) stimuli. Depicted 'Responsive' genes were manually selected for mono-responsiveness to either IFN-γ, TNF-α or TGF-β. Shown are the ratios between expression in cytokine stimulated and control condition for the same duration of time. Control conditions are hence not shown. **b** Scatter plot of IFN-γ vs. TNF-α gene set scores of data described in

Figure 5.4a-b, in which every point represents a cell. Plots are partitioned based on the 95th percentiles of the first control condition (top panel). **c** Quantification of cytokine mediated growth inhibition of OVCAR5 and NMM bystander tumor cells *in vitro*. Relative cell counts of CFP^+ OVCAR5 (left) and CFP^+ NMM (right) tumor cells after incubation in the absence or presence of recombinant IFN-γ TNF-α or IFN-γ *plus* TNF-α for the indicated time periods. Cells were analyzed by flow cytometry and cell counts were normalized to counting beads. Bar graph shows mean of 3 technical replicates. Representative data of two independent experiments are depicted. The cell inhibition that is observed upon IFN-γ exposure (in either the presence or absence of TNF-α) for OVCAR5 at 44 hours indicates that analysis of single cell RNA-seq data may somewhat (up to a factor of 2) underestimate IFN-γ sensing at this timepoint for this cell line (as 'sensing' cells are at a disadvantage). The cell inhibition that is observed upon IFN-γ exposure and that is enhanced by additional TNF-αexposure for NMM at 44 hours indicates that analysis of single cell RNA-seq data may somewhat underestimate IFN-γ (up to a factor of 1.3) and IFN-γ-TNF-α (up to a factor of 2) sensing at this timepoint for this cell line. Note that such cell inhibition does not influence estimates at the 16hr timepoint.

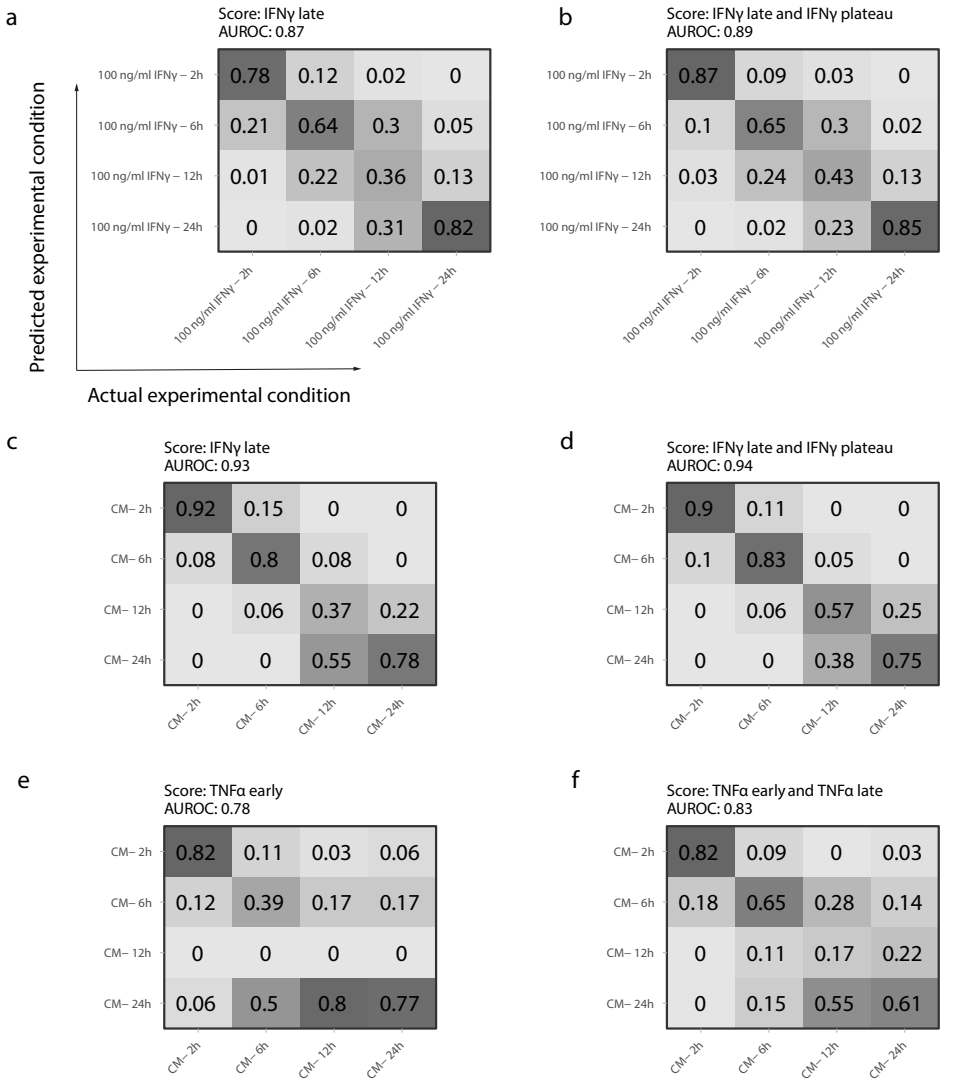

Figure S5.6: Quantification of exposure duration separability using time-informative gene sets
Quantification of distinguishability of experimental conditions using the indicated time-informative gene sets in Figure 5.1d. Included are so-called confusion matrices resulting from SVMs applied to held-out data (See Methods section '*Quantification of separability of experimental conditions using gene set scores*'), showing the fraction of assigned classes (vertical axes) for cells from each of the indicated experimental conditions (horizontal axes). For a perfect classifier, the diagonals of these matrices are 1 and all off-diagonal entries are 0.
a Cells were stimulated with IFN-γ (100 ng/ml) *in vitro* for the indicated durations. Data depict separability of exposure durations solely using the IFN-γ late gene set score.
b As in a, but now including the IFN-γ plateau gene set score as an additional explanatory variable, resulting in a modest increase in predictive performance (classifier AUROC of .89 from .87).
c As in a, but now using cells stimulated with culture medium (CM) from T cell - tumor cell co-cultures for the indicated durations.

d As in c, but including the IFN-γ plateau gene set, for a modest increase in predictive performance.
e As in c, but employing the TNF-α early gene set.
f As in e, but additionally including the TNF-α late gene set for a modest increase in predictive performance.